UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

Mining Survey Data for SWOT Analysis

by

Boonyarat Phadermrod

A thesis submitted for the degree of Doctor of Philosophy

November 4, 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
Electronics and Computer Science

Doctor of Philosophy

MINING SURVEY DATA FOR SWOT ANALYSIS

by Boonyarat Phadermrod

Strengths, Weaknesses, Opportunities and Threats (SWOT) analysis is one of the most important tools for strategic planning. The traditional method of conducting SWOT analysis does not prioritize and is likely to hold subjective views that may result in an improper strategic action. Accordingly, this research exploits Importance-Performance Analysis (IPA), a technique for measuring customers' satisfaction based on survey data, to systematically generate prioritized SWOT factors based on customers' perspectives which in turn produces more accurate information for strategic planning. This proposed approach is called IPA based SWOT analysis and its development issues discussed in this report are: (1) selecting a technique for measuring *importance* which is one of the two main aspects of IPA since currently there are no well-established approaches for measuring importance; and (2) identifying opportunities and threats since only strengths and weaknesses can be inferred from the IPA result.

The first issue is addressed by conducting an empirical comparison to analyse the performance of various techniques for measuring importance. Specifically, this thesis considers two data mining techniques namely Naïve Bayes and Bayesian Networks for measuring importance and compares their performance with other techniques namely Multiple Linear Regressions, Ordinal Logistic Regression and Back Propagation Neural Networks that have been used to derive the *importance* from the survey data. The comparison result measured against the evaluation metrics suggests that Multiple Linear Regressions is the most suitable technique for measuring *importance*.

Regarding the second issue, opportunities and threats were identified by comparing the IPA result of the target organisation with that of its competitor. Through the use of IPA based SWOT analysis, it is expected that an organisation can efficiently formulate strategic planning as the SWOT factors that should be maintained or improved can be clearly identified based on customers' viewpoints. The application of the IPA based SWOT analysis was illustrated and evaluated through a case study of Higher Education Institutions in Thailand. The evaluation results showed that SWOT analysis of the case

study has a high face validity and its quality is considered acceptable, thereby demonstrating the validity of this study. Although the application of IPA based SWOT analysis was illustrated in the specific field, it can be argued that IPA based SWOT analysis can be used widely in the other business areas where SWOT analysis has been seen to be applicable and the customer satisfaction surveys are generally conducted.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Boonyarat Phadermrod , declare that the thesis entitled *Mining Survey Data for SWOT Analysis* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as can be seen in section 1.4, chapter 1. Parts of this work are also currently under review as:
   - Phadermrod, B., Crowder, R. M., & Wills, G. B. (2016). Importance-Performance Analysis based SWOT analysis. *International Journal of Information Management.*
   - Phadermrod, B., Crowder, R. M., & Wills, G. B. (2016). Comparative analysis of importance measuring from customer satisfaction survey using data mining techniques. *Service Science.*

Signed:...

Date:..

Acknowledgements

Studying PhD abroad is one of the dreams that I would like to pursue. I really glad that I, finally, can make it become true. My PhD journey is filled with several moments of self-doubts and fear. I could not possible to achieve PhD without the support and guidance from many people.

Firstly, I would like to express my sincere gratitude to my advisors: Dr Richard Crowder and Dr Gary Wills, for supervision, time, and effort they put over the past few years and for encouraging me to grow as an independent researcher.

I would also like to thank my thesis committee: Dr David Millard and Dr Muthu Ramachandran, for their constructive comments which give a better version of my thesis and initiate the ideas of future research.

I gratefully acknowledge the funding from Royal Thai Government that gives me an opportunity for undertaking PhD and having a wonderful experience of living aboard. My sincere thank also goes to the staff of Thai Government Students' office for their support in solving problems I faced during the time in the UK.

I am grateful to the language experts for their help in checking the language consistency between an English-Thai version of questionnaires. I would like to thank all participants including staff and CPE students at Kasetsart University, and MBA students at Silpakorn University who so generously gave their time to response to questionnaires from part of my research.

I thank my friends in Thailand for their thoughts and well-wishes sending across miles. A million thanks to all Thai-Soton friends for being by my side and making me feel like we are not just friends, but we are family. I am also very thankful to all my colleagues at Mango for the good moments we had working together and for your friendship.

I would also like to express my special thanks to my parents, my brother and my sister for their unconditional love and for believing in me and supporting me to follow my dream. In particular, I would like to thank my uncle in law: Mr. David Steane for being my great supporter and a language tutor. Many thanks also go to the rest of my family for their thoughts, visits and continuous encouragement throughout this PhD.

And finally, my heartfelt thanks go to my better half: Mr Pramote Chopaka, for his constant motivation, patience, encouragement, and for his love during these years. I also thank him for letting me know that distance cannot keep as apart but create stronger bonds.

Abbreviations Used

AHP	Analytic Hierarchy Process
ANP	Analytic Network Process
ANOVA	Analysis of Variance
AUC	Area Under the ROC Curve
BNs	Bayesian Networks
BPNN	Back Propagation Neural Network
DEMATEL	DEcision MAking Trial and Evaluation Laboratory
DR	Direct-rating scales
HEIs	Higher Education Institutions
IPA	Importance-Performance Analysis
KDD	Knowledge Discovery in Databases
MAE	Mean Absolute Error
MI	Mutual Information
MLR	Multiple Linear Regression
NLP	Natural Language Processing
OLR	Ordinal Logistic Regression
R^2	Goodness of fit
RMSE	Root Mean Squared Error
ROC	Receiver Operating Characteristic
SEM	Structural Equation Modelling
SWOT	Strength, Weakness, Opportunities and Threats
SVM	Support Vector Machine
VIF	Variance Inflation Factor

Chapter 1

Introduction

In today's fierce business environment economic competition is now global and customer needs and technologies are changing rapidly requiring enterprises to adapt to these changes if they are to stay competitive. One of the critical factors for business success is strategic planning. Strategic planning enables enterprises to define their purpose and direction (Lawlor, 2005). In addition, Lawlor (2005) also states that '*Without strategic planning, businesses simply drift, and are always reacting to the pressure of the day*'. To create good strategic planning, enterprises need to have a clear understanding of their business which requires all business information, both internal and external, to be analysed.

Among the most important tools for strategic planning is SWOT analysis (Ying, 2010; Hill and Westbrook, 1997). SWOT analysis is a basic method for analysing and positioning an organisation's resources and environment in four regions: Strengths, Weaknesses, Opportunities and Threats (Samejima et al., 2006). By identifying these four fields, the organisation can recognize its core competencies for decision-making, planning and building strategies. SWOT analysis was devised by Albert S Humphrey as part of research at the Stanford Research Institute in the 1960s-1970s. This research was conducted by using data from Fortune 500 companies (Humphrey, 2005). Since then, SWOT analysis is widely used in both academic communities (Ghazinoory et al., 2011) and leading companies, for example Amazon.com, Inc. (Global Markets Direct, 2012) and eBay Inc. (Datamonitor, 2003).

Although SWOT analysis has been widely accepted as a tool for strategic planning (Oliver, 2000), in practice it cannot offer an efficient result and so that may cause an improper strategic action (Wilson and Gilligan, 2005; Coman and Ronen, 2009). This is because the traditional approach of SWOT analysis is based on qualitative analysis in which SWOT factors are likely to reflect the subjective views of managers or planner judgement and SWOT factors in each region are not ranked by the significance for company performance.

Additionally, in an increasingly competitive environment where customer orientation is essential in many businesses (Dejaeger et al., 2012), customer satisfaction has become one of the main indicators of business performance (Mihelis et al., 2001) that provides meaningful and objective feedback about customer preferences and expectations for organisations. Hence, the SWOT should be evaluated from the customer's perspective rather than the company point of view to ensure that the capabilities perceived by company are recognized and valued by the customers (Piercy and Giles, 1989; Wilson and Gilligan, 2005).

This deficiency in the traditional approach of SWOT analysis motivated our research to exploit the Importance-Performance Analysis (IPA), a technique for measuring customers' satisfaction from customer satisfaction surveys (Martilla and James, 1977; Matzler et al., 2003; Levenburg and Magal, 2005), to systematically generate prioritized SWOT factors based on customers' perspectives. This in turn produces more accurate information for strategic planning.

Specifically, strengths and weaknesses of the organisation are identified through an IPA matrix which is constructed on the basis of two main aspects of IPA which are *importance* and *performance*. Since, each quadrant in the IPA matrix can be interpreted as major/ minor strengths and weaknesses as shown in Figure 1.1 (Garver, 2003; Deng et al., 2008a; Silva and Fernandes, 2012; Hasoloan et al., 2012; Cugnata and Salini, 2013; Hosseini and Bideh, 2013) opportunities and threats are obtained by comparing the IPA matrix of the organisation with that of its competitor.

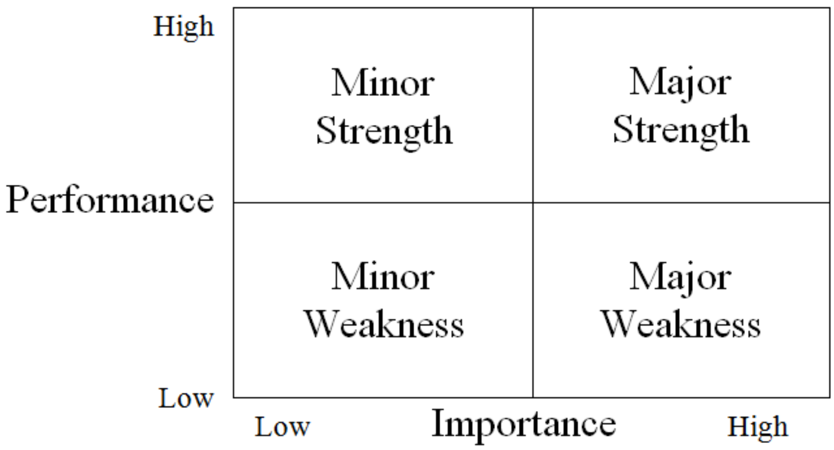


Figure 1.1: IPA matrix represented major/minor strengths and weaknesses adapted from Garver (2003)

1.1 Research challenges

As previously stated, the integration of IPA and SWOT will generate prioritized SWOT factors based on customers' perspectives. Therefore, this thesis adopts IPA for identifying strengths and weaknesses based on customers' perspectives collected through a customer satisfaction survey in order to improve the diagnostic power of SWOT analysis. To develop the framework for identifying SWOT analysis from IPA results, called Importance-Performance Analysis based SWOT analysis (IPA based SWOT analysis), three issues have to be considered:

1. To construct the IPA matrix, the *importance* and *performance* have to be measured as these are two main aspects of IPA. The technique for calculating *performance* is well-established by using a direct-rating scale whereas there are many approaches for indirectly measuring *importance* using statistical and data mining techniques. Since different *importance* measurement techniques will likely result in identifying dramatically different attributes for improvement, various techniques for measuring *importance* need to be compared against evaluation metrics in order to select an appropriate technique for measuring *importance*.

2. As only strength and weakness can be inferred from the IPA matrix, a possible solution to infer opportunity and threat from the IPA matrix need to be specified in order to provide a complete aspect of SWOT analysis.

3. Though SWOT analysis has been used in a wide range of subject areas and many studies have attempted to improve the limitations of traditional SWOT analysis, currently there are no direct methods and tools for validating the effectiveness and usability of SWOT analysis (Ayub et al., 2013).

1.2 Research questions

Based on the research challenges given in Section 1.1, research questions of this thesis are.

Research question 1. Which importance measure should be used in IPA?

Research sub-question 1.1. Which approaches for assessing *importance* work best in IPA: customer self-stated or implicitly derived importance measure?

Research sub-question 1.2. Which data mining technique is most appropriate for measuring *importance*?

Research question 2. How can IPA be applied to develop a SWOT analysis based on a customer satisfaction survey?

Research question 3. How good is the outcome of IPA based SWOT analysis?

Research sub-question 3.1. What is the staff level agreement on the outcome produced by IPA based SWOT analysis?

Research sub-question 3.2. What is a quality of outcome produced by IPA based SWOT analysis compare to the traditional SWOT analysis?

The first research question is aimed at conducting an empirical comparison of techniques for measuring *importance*. To address this research question several techniques that can be used for measuring importance were selected and the evaluation metrics were identified based on previous comparative studies. Further details of the empirical comparison as well as its results can be found in Chapter 5 and Chapter 6.

The second research question centred on designing a methodological framework for developing IPA based SWOT analysis. This framework serves as an outline of the main steps to be completed in order to obtain an organisation's SWOT from survey data. Further detail regarding the development of IPA based SWOT analysis can be found in Chapter 7.

The third research question is related to the evaluation of IPA based SWOT analysis through a case study using the survey research method. The details of evaluation methodology and results are described in Chapter 8 and Chapter 9 respectively.

In conclusion, it should be recognized that this work does not propose new algorithms but the empirical comparison was conducted to analyse the performance of various data mining techniques for measuring *importance*. The second topic being discussed in this work is about acquiring SWOT from the result of IPA and planning to evaluate the proposed framework by means of survey research methods.

1.3 Outline of the thesis

The overview of the following 10 chapters is as follows and Figure 1.2 shows how each chapter in this thesis relates to the others and the research questions.

Chapter 2 provides an in-depth overview of three principal topics relating to the research. The chapter begins with a brief introduction to SWOT and the research involved with the implementation of the SWOT analysis system. Then, this chapter provides general information about the process of knowledge discovery in databases and five well-known statistical and data mining techniques followed by the characteristic of satisfaction surveys and research involved mining for customer satisfaction.

Chapter 3 provides the background knowledge about Importance-Performance Analysis as it is a tool to analyse customer satisfaction in order to create SWOT. Three main topics regarding the IPA which are explained in this chapter are approaches that have

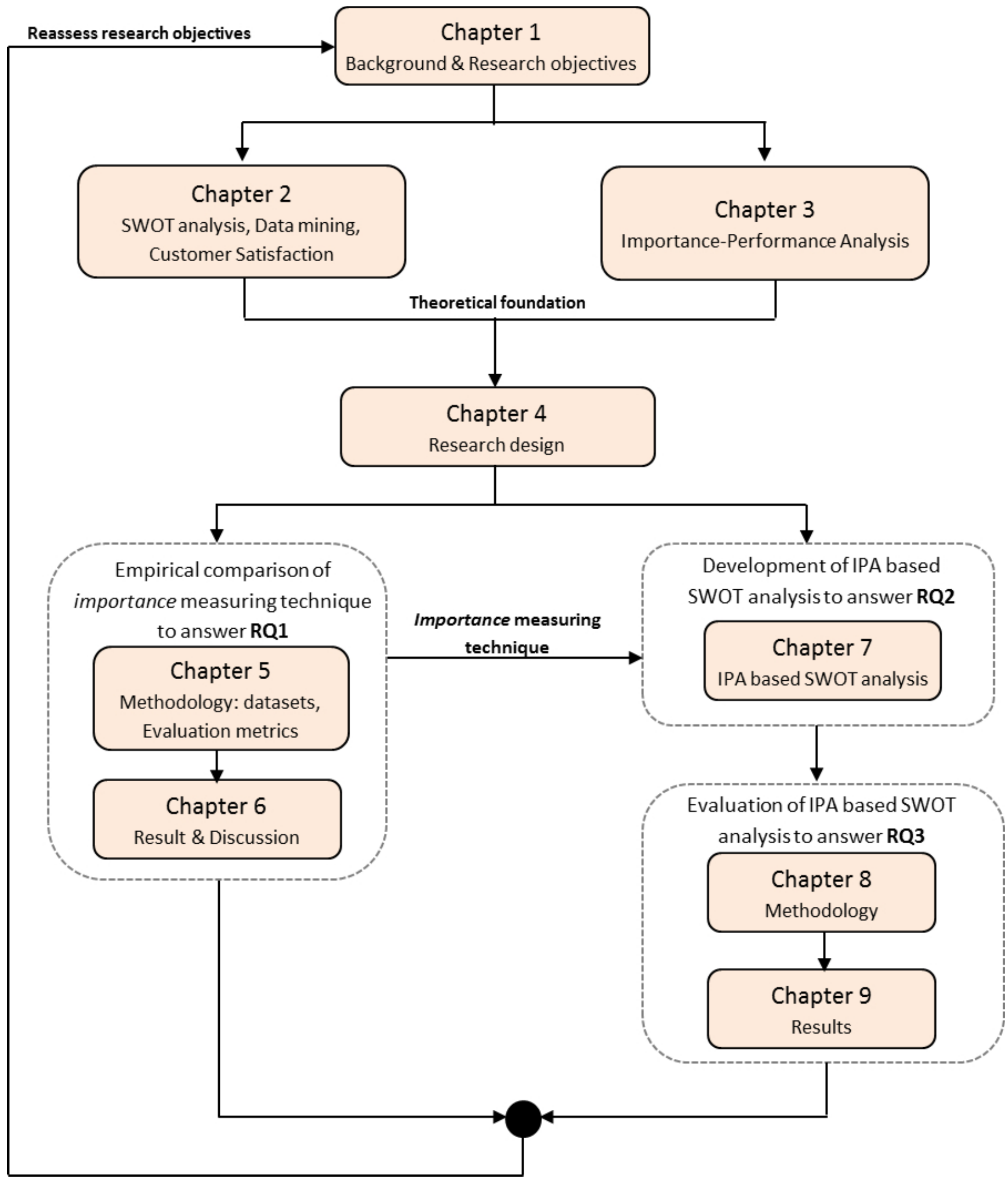


Figure 1.2: Diagram shows the relationship between the thesis chapters and the research questions

been used for measuring *importance*, past comparative studies of methods for measuring *importance* and methodology for deriving *importance* from results of two statistical techniques and three data mining techniques that are investigated in this thesis.

Having conducted the literature reviews in Chapter 2 and Chapter 3, Chapter 4 explains a procedural plan for how this research study is to be completed. Specifically, three main phases of research process namely "problem identification", "solution design", and "evaluation" are described with discussions of methodologies corresponding to each task of the research process and the rationale to select the methodology that is used to complete this research study.

Chapter 5 describes the empirical comparison of techniques for measuring *importance* in which each importance measurement technique is compared against evaluation metrics. Next, Chapter 6 reports the comparative results of the importance measurement techniques based on the three evaluation metrics and provides a discussion of comparative results. This leads to the justification to select the proper importance measurement technique.

Chapter 7 describes a framework of applying IPA for SWOT analysis. This framework serves as an outline of the main steps to be completed in order to obtain a company's SWOT from customer satisfaction survey. The key steps of the IPA based SWOT analysis are the IPA matrix construction in which a customer satisfaction survey is analysed to calculate the *importance* and *performance*, and SWOT factors identification based on the IPA matrix.

Chapter 8 explains methodologies related to the case study which was conducted to demonstrate and evaluate the IPA based SWOT analysis. Methodologies include the survey research method for conducting the three surveys and the implementation of IPA based SWOT analysis on a real case study. Subsequently, Chapter 9 reports the statistical analysis and evaluation results.

Chapter 10 discusses the results related to the exploration of the research questions as well as limitations of the study. And finally Chapter 11 provides the conclusion of the research as well as identifying directions for future work.

1.4 Publication list

The following peer reviewed conference papers and journal based on the work of this thesis have been accepted for publication:

1. Phadermrod, B., Crowder, R. M., & Wills, G. B. (2015). Attribute Importance Measure Based on Back-Propagation Neural Network: An Empirical Study. *International Journal of Computer and Electrical Engineering*, 7(2), pages 118-126.

2. Phadermrod, B., Crowder, R. M., & Wills, G. B. (2014). Developing SWOT Analysis from Customer Satisfaction Surveys. In *proceeding of 11th IEEE International Conference on e-Business Engineering (ICEBE)*, November, pages 97-104.

Chapter 2

Literature Review

The purpose of this chapter is to provide the background relating to the problem introduced in Chapter 1. The chapter begins with Section 2.1 that provides a brief introduction to SWOT and the research involved with the implementation of the SWOT analysis system. Section 2.2 gives general information about the process of knowledge discovery in databases and five well-known data mining techniques. Section 2.3 explains characteristics of satisfaction surveys and research involved mining for customer satisfaction. Finally, a summary of the chapter is provided in Section 2.4.

2.1 SWOT analysis

SWOT analysis is a practical analytical tool based on four fields: Strengths, Weaknesses, Opportunities and Threats (Hill and Westbrook, 1997). It is one of many tools that can be used in an organisation's strategic planning process. Other tools that are commonly used for strategy analysis are PEST analysis, Five Forces analysis, and 3C (Company-Customer-Competitor) analysis (Akiyoshi and Komoda, 2005).

Regarding the survey conducted by the Competitive Intelligent Foundation (Fehringer et al., 2006) which received responses from 520 competitive intelligent (CI) professionals, SWOT is the second-most frequently used analytic tool with 82.6% of respondents. It was ranked after competitor analysis with 83.2% of respondents. Additionally, the survey based on the answers supplied by the Chief Executive Officers of wide range organisations in the UK shows that SWOT analysis is the most widely applied strategic tool by organisations in the UK(Gunn and Williams, 2007). Recently, a survey about analytical methods used by enterprises in South African for environmental scanning also shows that SWOT analysis is the most frequently used analytic tool with 87% of respondents followed by competitor analysis with 85% of respondents (du Toit, 2016).

SWOT analysis offers a simple structured approach to identify an organisation's strengths and weaknesses and provides an external view of opportunities and threats (Dyson, 2004) which enables the organisations to construct strategies based on their strengths, minimize or eliminate their weaknesses, take advantage of opportunities and overcome threats to an organisation (Mojaveri and Fazlollahtabar, 2012).

2.1.1 SWOT analysis matrix

The SWOT analysis matrix is a 2 $\times$ 2 matrix as shown in Figure 2.1. The first row of the matrix represents Strengths and Weaknesses as internal factors that supporting and obstructing organisations to achieve their mission respectively. The second row represents Opportunities and Threats as the external factors that enable and disable organisations from accomplishing their mission respectively (Dyson, 2004). The internal factors are controllable factors within organisations, for example, finance, operations, and product whereas the external factors are uncontrollable factors arise from the external environment, for instance, political, economics, and social (Ghazinoory et al., 2011).

	Positive factors	Negative factors
Internal factors	Strengths	Weaknesses
External factors	Opportunities	Threats

Figure 2.1: The SWOT analysis matrix

2.1.2 Advantages and Disadvantages of SWOT analysis

Ghazinoory et al. (2011) and Nordmeyer (nd) describe the advantages and disadvantages of SWOT analysis. The main advantage of SWOT analysis is its simplicity thus allowing anyone with knowledge about the business to use it which can reduce costs from hiring external consultants. In addition, SWOT analysis can be applied in a wide range of subject areas - from health care, education, transportation, to the military. SWOT analysis also has some drawbacks. First, factors are described broadly, with unclear and ambiguous words and phrases that may make it difficult to apply SWOT result in strategic design. Second, factors in each quadrant are not ranked by their significance and, as a result, the organisation cannot determine the factors that are truly impacting on the company's goal. The third drawback is related to the subjective process in SWOT analysis in which the results represent opinions of the individuals who participate in the brainstorming session (instead of fact). In addition, there is no obligation to verify statements and opinions with data or analysis thus the results may lead to business decisions based on either unreliable or irrelevant data.

2.1.3 Quantitative SWOT analysis

According to the disadvantage in prioritization of SWOT factors, a number of researchers have proposed to enhance SWOT analysis with two multiple criteria decision-making (MCDM) techniques namely - Analytic Hierarchy Process (AHP) and Analytic Network Process (ANP).

The use of AHP in combination with SWOT was first proposed by Kurttila et al. (2000) and it has been referred to as A'WOT in Kangas et al. (2001). In A'WOT, first the SWOT analysis is carried out through a brainstorming session in which factors of the external and internal environment are identified. Then the AHP performs pair-wise comparisons among factors in order to determine the relative importance of SWOT factors thus these factors can be commensurable and prioritized. Specifically, the pair-wise comparisons are conducted in two levels: "Within each SWOT group" and "Between the four SWOT groups". The first provides the relative importance of the local factors (local priority) and the latter provides the relative importance of the SWOT groups (group priority). Subsequently, the total global priority of each SWOT factor is obtained by multiplying the local priorities of factors by the corresponding group priorities (Kurttila et al., 2000; Kangas et al., 2001). The SWOT-AHP method has been applied in various domains such as machine tool industry (Shinno et al., 2006), container ports (Chang and Huang, 2006) and tourism (Oreski, 2012).

Some researchers extended the SWOT-AHP by integrating it with fuzzy analytic hierarchy process called "fuzzy AHP-SWOT" (Lee and Lin, 2008; Zaerpour et al., 2008) or enhanced it with multiple criteria group decision-making (MCGDM) (Gao and Peng, 2011). In addition, some researchers utilized ANP instead of AHP in the SWOT framework which is called "ANP-SWOT" (Yüksel and Dagdeviren, 2007; Fouladgar et al., 2011).

The fuzzy AHP-SWOT allows decision-makers to provide fuzzy judgement in pair-wise comparisons instead of exact judgement in order to fully reflect a style of human thinking in real world decision problems (Zaerpour et al., 2008) which in turn produces the sensible quantitative values for the SWOT factors. The main steps for performing fuzzy AHP-SWOT is similar to the steps for performing AHP-SWOT except a method for calculating the relative importance from fuzzy values identified by decision makers.

The hybrid SWOT-MCGDM method quantifies SWOT factors based on the preference of multiple decision makers on SWOT factors and groups to provide more versatile information for evaluating the relative importance of SWOT factors (Gao and Peng, 2011).

To be specific, this method allows each decision maker to provide his/her personal preference of SWOT factors and groups in pair-wise comparison rather than asks multiple decision makers to discuss and provide exact group judgements.

Regarding the ANP-SWOT method, SWOT analysis is performed in connection with ANP, a generalization of the AHP, which enables measuring dependency among SWOT factors since generally the SWOT factors are dependent and related with one another but the AHP-SWOT assume that there is no dependence among the SWOT factors (Yüksel and Dagdeviren, 2007). The comparison result of AHP-SWOT and ANP-SWOT conducted in this paper shows that the relative importance of SWOT factors of these two methods is different due to the dependency among the SWOT factors. Additionally, Gorener (2012) also confirms that there are significant differences between AHP-SWOT and ANP-SWOT outcomes which can affect the strategy selection.

With the goal of making prioritized SWOT factors, all variations of quantitative SWOT analysis focus on determining and computing relative importance of SWOT factors. While this approach can make SWOT factors commensurable, they are produced solely on a company's perspective (without the customer's perspective). Although Piercy and Giles (1989) suggested the use of customer orientation in SWOT framework in order to make better use of the SWOT.

In consideration of the validation of these variations of quantitative SWOT analysis, a lack of validity testing is found. Most of the papers provided only a case study to demonstrate the application of their proposed approaches but they did not provide a concrete evaluation to test for the validity of the approaches. Yüksel and Dagdeviren (2007) discussed some difficulties with testing the validity and tested the validity of their proposed ANP-SWOT by comparing with the AHP-SWOT.

2.1.4 Research in SWOT analysis system

Although the growth in the number of papers about SWOT analysis shows that researchers are becoming increasingly interested in this topic, the major proportion of SWOT studies are focused on applying SWOT analysis to case studies in different areas and industries (Ghazinoory et al., 2011). There are a small number of published works aimed at developing a system that can assist users in performing a SWOT analysis.

One of the first articles in this field was published by Houben et al. (1999). This article exploited expert systems for implementing a prototype of a knowledge-based SWOT analysis system. However, the quality of rules for SWOT analysis is dependent on a number of experts as a source of knowledge. In addition, this research does not provide a complete view of SWOT analysis because it concentrates solely on the identification of strengths and weaknesses.

Other studies, Samejima et al. (2006) and Dai et al. (2011) exploit Text mining and Information Extraction to extract the keyword identified factors from unstructured data such as a product's press releases, financial news site, reports, and e-mail. This extracted information is then analysed in order to label as strength, weakness, and neutral by analysing their characteristics and locate in the SWOT matrix. However, Samejima et al. (2006) focus on analysing only press releases that described characteristics of a company's product. Thus, it can perform SWOT analysis in only one area of business management that is the product development. This is in contrast with Dai et al. (2011), whose work is able to extract factors related to the main criteria of SWOT analysis (such as technology, price, equipment, service, attitude, political change) and five parties regarding Five Forces Analysis (namely rivals, buyers, suppliers, substitutes, potential entrants) from diverse sources.

In 2013, a project proposal named e-SWOT[1] was submitted to EU's Seventh Framework Programme for Research (FP7) for funding. This project relied on various technologies such as social media analytics, text mining and sentiment analysis for constructing an automatic SWOT which will assist companies in assessing brands, products and their reputation globally.

Pai et al. (2013) developed an ontology-based SWOT analysis mechanism that analyses the structure of online Word-of-Mouth (WOM) appraisals and interprets them as the strengths, weaknesses, opportunities, and threats of a company. Specifically, this study extracts WOM appraisals from on-line resources then applies sentiment analysis in cooperation with Ontology to classify extracted appraisals into positive/negative appraisals. Then, both positive and negative appraisals are used to assess the strengths, weaknesses, opportunities, and threats of the target company. The system proposed in this work was evaluated regarding a user satisfaction questionnaire, which proved that the proposed method can be used to accommodate strategic planning.

Most of the studies described in this section are used customers' feedback as a source to create SWOT analysis. Thus, they are classified as customer-oriented SWOT analysis. This approach of generating SWOT makes better use of customers' feedback which helps companies to ensure that their core competency is recognized and valued by the customers. However, the drawback of no prioritized SWOT factors still exists as the SWOT factors produced by this approach are not measurable.

2.2 Data Mining

The requirement for turning large amount of data generated from various fields of human life application into valuable knowledge that can be utilized to support decision-making

[1] http://www.ideal-ist.eu/ps-es-82678

has formed the Knowledge Discovery in Databases (KDD) referred to as a set of steps for extracting useful patterns from large data repositories (Mannila, 1996). The core of KDD process is data mining which exploits data analysis and discovery algorithms for searching for interesting, useful patterns over the data (Fayyad et al., 1996b). Therefore, data mining is often used as a synonym for KDD (Mannila, 1996; Padhy et al., 2012)

The KDD process is an iterative activity which generally can be divided into three main tasks namely pre-processing, data mining and post-processing (Olaru and Wehenkel, 1999). The pre-processing consists of data selection, data transformation and data cleaning which identifies the dataset and chooses attributes, organises data in appropriate ways for the mining procedure, and removes irrelevant data as well as deals with missing values respectively. The post-processing consists of steps for assessing and interpreting mining results for ensuring the validity and reliability of discovered patterns (Fayyad et al., 1996b; Olaru and Wehenkel, 1999).

2.2.1 Data mining methods

Regarding the discovery goal, there are two types of data mining tasks: prediction and description. The former extracts patterns based on the current data to predict either unknown or future data and the latter extracts patterns for describing common characteristics of the data (Han and Kamber, 2000b).

These data mining tasks can be achieved by exploiting the following primary data mining methods (Fayyad et al., 1996a; Olaru and Wehenkel, 1999; Sahu et al., 2008).

1. ***Classification:*** this method infers relationship patterns between input data and predefined classes for classifying new unknown data into output classes. Well-known classification technique are Decision tree, Nearest neighbour, Naïve Bayesian and Neural networks.
2. ***Regression:*** this method creates mathematical formula that fit the input data for predicting behaviour of new datasets. This method is similar to classification in predicting unknown data but it can perform well on continuous numerical values rather than categorical data.
3. ***Clustering:*** this method measures the similarity between points of data based on their properties then forms clusters of data that share common properties. A well-known clustering technique is k-means clustering.
4. ***Association Rule Discovery:*** this method searches for co-occurrence patterns that can reveal the dependency between attributes of datasets. The classic application of Association Rule is identifying products of a supermarket that frequently are bought together to assist supermarket shelf management. Therefore, Association Rule is also known as ‘market basket analysis’.

5. ***Summarization:*** this method provides compact descriptions representing characteristics of an input dataset. These descriptions can be presented in numerical form such as means, and standard deviations or graphical form such as histograms, and scatter plots.

6. ***Deviation Detection:*** this method focuses on recognition of the significant changes in the data from previously measured or normative values. Thus, it can be used for discovering unusual behaviours or detecting new phenomenon contained in the data.

Fayyad et al. (1996a) and Sahu et al. (2008) provide more information including type of mining task and application of each method as described in Table 2.1.

Table 2.1: Data mining methods and their applications

Method	Mining Task	Application
Classification	Prediction	Fraud Detection, Recognition of Spam E-mail, Direct marketing
Regression	Prediction	Predicting sales amounts of product based on advertising expenditure, Estimating the survival probability of patient based on the diagnostic tests
Clustering	Description	Market Segmentation, Document Clustering
Association Rule Discovery	Description	Market basket analysis, Intrusion Detection
Summarization	Description	Automated report generation
Deviation Detection	Prediction	Intrusion Detection, Ecosystem Disturbances (predicting events like hurricanes and flood) (Tamberi, 2007)

These methods exploit different types of techniques such as statistics and machine learning. Each method has its own strengths and weaknesses meaning that each method typically suits some kinds of problems better than others (Fayyad et al., 1996a). Thus, choosing data mining methods that can offer useful and relevant knowledge is one of the challenges in developing data mining application (Singh et al., 2011).

One of the main tasks of this study is measuring *importance* as the relative strength of attributes that contributed to the overall customer satisfaction, which requires the methods that able to discover or extract relationship between input and predefined class. This requirement matches the goal of prediction and classification methods which is to create a model for representing a relationship between input and predefined class. While the other methods such as association rule focuses on discovering the dependency between attributes, and the clustering focuses on partition dataset into groups based on the attribute similarity. Hence, prediction and classification are the most suitable methods to be applied in measuring *importance*. Consequently, this thesis is mainly focused on prediction and classification techniques and their detail will be described in the next section.

2.2.2 Classification and Prediction Techniques

Classification and prediction are the two major methods of addressing prediction problems. The goal of these methods is to create a model which represents the relationship between attributes' values (independent variables) and target attribute value (dependent variable) and then use this model to predict the values of the target attribute of new data. The classification works well in predicting categorical (discrete) target attributes whereas the prediction works well in predicting continuous quantitative target attributes (Han and Kamber, 2000b). In the context of the classification, the target attribute is termed as class attribute and in the context of the prediction the target attribute is termed as predicted attribute.

Classification and prediction are a two-step process. The first step is the ***learning or training phase*** in which the training data is analysed in order to construct the classification and prediction models. Different classification and prediction techniques use different methods to build their models. The second step is the ***testing phase*** in which the predictive accuracy is evaluated by applying the model on a testing data to classify or predict the target attribute. The predictive accuracy is measured by considering the percentage of testing data that are correctly predicted by the model. If the model accuracy is considered acceptable, it can be deployed to predict the unseen data.

In this section, first the Multiple Linear Regressions which belongs to prediction methods are explained. Next, the four classification techniques namely - Ordinal Logistic Regressions, Back Propagation Neural Networks, Naïve Bayes and Bayesian Networks are explained.

2.2.2.1 Multiple Linear Regressions

Multiple Linear Regressions (MLR) is the statistical based, supervised learning technique in which the regression model is constructed based on input data in order to predict new instances of data. The regression model is represented as the linear combination of independent variables that are maximally correlated with the single dependent (target) variable. In general, suppose that there are k distinct independent variables, the equation of a regression model is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$$

where Y is the target variable, X_i is the i^{th} independent variable, β_0 is a regression intercept, β_i is the i^{th} regression coefficient (or regression weight) quantifies the association between each X_i and Y, and ε is a residual (error) term (Patel, 2003).

This regression equation expresses the contribution of each X_i to Y. In order to ensure that the regression model is a fit to the data, the primary task of MLR is to determine the best set of parameters β_i that minimize the difference between the actual outcome and the predicted outcome by using the method of least squares. The MLR can be thought as an extension of simple linear regression involved the single independent variable X and the dependent variable Y, which Y can be modelled as a linear function of X (Han and Kamber, 2000a).

Similar to other statistical techniques, the use of the MLR requires certain assumptions to be satisfied in order to ensure the regression model is valid. The assumptions of MLR that are identified as a primary concern in this research include linearity, normality, and collinearity. The linearity assumption assumes that the relationship between the independent variables and the dependent variable is linear, as appears in the name of MLR. If the relationship between independent variables and the dependent variable is not linear, the results of the regression analysis will be inaccurately estimated due to the regression model not fitting the data. A common method of checking linearity is the examination of residual (predicted minus actual values) plots (plots of the standardized residuals as a function of standardized predicted value) (Osborne and Waters, 2002). Normality assumes that the residuals are distributed normally. This assumption can be examined by using histograms as well as normal probability plots (Osborne and Waters, 2002). Collinearity refers to the assumption that the independent variables are uncorrelated. The presence of high correlations among the independent variables which is called multicollinearity can result in reducing power of the regression coefficients and misleading and unusual results (Keith, 2006). The existence of multicollinearity can be detected by observing a correlation matrix as well as calculating variance inflation factors (VIF).

2.2.2.2 Ordinal Logistic Regressions

Ordinal logistic regression or (ordinal regression) is an extension of a logistic regression that is especially used to predict an ordinal dependent variable with k responses category given one or more independent variables (O'Connell, 2006). Compared to the other types of regression such as MLR and binary logistic regression, OLR is the most appropriate and practical technique to analyse the effect of independent variables on a rank order dependent variable that cannot be assumed as a continuous measure or as normally distributed (Chen and Hughes, 2004).

The general equation of OLR is written in the following form (Piegorsch, 1992; Bender and Benner, 2000; Sentas et al., 2005):

$$l(c_j) = \alpha_j - \sum_{i=1}^{k} \beta_i x_i \tag{2.1}$$

where c_j is the cumulative probability for the jth category, α_j is the threshold for the jth category, $\beta_1 \ldots \beta_k$ are the regression coefficients, $x_1 \ldots x_k$ are the predictor variables, and k is the number of independent variables.

To estimate thresholds and coefficients, two major parameters in Equation 2.1 which are a number of independent variables k and the link function $l()$ need to be specified. The link function is a transformation of the cumulative probabilities that allows estimation of the model and is usually one of the following (Piegorsch, 1992): the logit, $log\{c/(1-c)\}$; the complementary log-log (cloglog), $log\{-log(1-c)\}$; the negative log-log, $-log\{-log(c)\}$; the probit, $\Phi^{-1}(c)$; and the cauchit, $tan(\pi(c-0.5))$.

There is no specific method to choose the link function that best fits a given dataset (Chen and Hughes, 2004). In general, different link functions are correspondent to the observed relative frequencies of the categories. The logit link is suitable when an ordinal dependent variable is evenly distributed among all categories; the cloglog link is appropriate when higher categories of ordinal dependent variable are more probable (Chen and Hughes, 2004) and the negative log-log link is appropriate when lower categories of ordinal dependent variable are more probable (Yay and Akıncı, 2009).

Among the five link functions, the first two: logit and cloglog are the most widely used (Bender and Benner, 2000; Chen and Hughes, 2004). The cumulative logit model for the k responses category of an ordinal dependent variable is written in the following form:

$$log\left(\frac{p(y \leq j|x)}{p(y > j|x)}\right) = \alpha_j - x\beta \quad (1 \leq j < k) \tag{2.2}$$

where j is a category being observed, y is the responses category of the dependent variable, x is a vector of independent variables, α_j is the cut-off point or threshold and β is a vector of logit coefficients (Fullerton, 2009). The left part of the equal sign in Equation 2.2 is the log of the cumulative odds which is the proportion of cumulative probability of being at or below a particular category (j) to the probability of being higher than that particular category (j) given the known independent variable.

While the logit was defined as the log of the odds, the clog-log is defined as the log of the negative log of the complementary probability (O'Connell, 2006) and the form of the cloglog link function is defined as follows:

$$log\left(-log(1-p(y \le j))\right) = \alpha_j - x\beta \tag{2.3}$$

The response variable in Equation 2.3 is transformed based on the conditional probabilities create Equation 2.4 in which its left part of the equal sign is the logs of the hazards which is the proportion of cumulative probability of being at a particular category (j) to the probability of being higher than that particular category (j) given the known independent variable.

$$log\left(-log\left(p(y = j|x)/p(y > j|x)\right)\right) = \alpha_j - x\beta \tag{2.4}$$

The quantity to the right of the equal sign of Equation 2.2 - 2.4 is the linear combination of the independent variables which is equivalent to the quantity to the right of the equal sign of Equation 2.1. Within these equations threshold (α) for each cumulative probability and the logit coefficients (β) are unknown to be estimated by means of the maximum likelihood method (Chen and Hughes, 2004).

With regard to the estimated logit coefficients, independent variables that influenced the dependent variable can be determined. Specifically, for continuous independent variables, the magnitude of the logit coefficients indicates that a one unit change on a specific independent variable affects the change of the odds (or relative risk) of the event occurrence by a factor of e^{β} , holding other independent variables as constant (Chen and Hughes, 2004).

In addition to the magnitude, the sign of logit coefficients should also be considered as it describes the direction of a relationship between independent and dependent variable. A positive coefficient indicated that the change of dependent variable is moved in the same direction of the change on a specific independent variable whereas a negative coefficient indicated that the change of dependent variable is moved in the opposite direction of the change on a specific independent variable.

As with other types of regression, the use of the OLR requires certain assumptions to be satisfied in order to ensure the regression model is valid. Instead of assumptions of normality, OLR model assumes that the corresponding regression coefficients in the link function are equal across all categories of ordinal responses (Bender and Benner, 2000). This is usually known as the assumption of proportional odds (or the assumption of parallel lines in SPSS). The proportionality assumption is the key assumption in OLR and the violation of this assumption results in biased coefficients, which may affect the implications for hypothesis testing (Fullerton, 2009).

2.2.2.3 Back Propagation Neural Networks

A BPNN model is one of the most widely used in neural computing for classification and prediction. Regarding Jost (1993) as cited in (Deng et al., 2008a), BPNN can be considered as an advanced multiple regression analysis that can accommodate complex and non-linear data relationships. BPNN is capable of producing an arbitrarily complex relationship between inputs and outputs and it is very robust technique that means BPNN still produces a good performance on noisy training data. However, BPNN also has some disadvantages which are the difficulty in interpreting the output due to its black-box nature, and the long training time (Svozil et al., 1997).

BPNN belongs to a ***multilayer feed-forward network*** and a supervised learning method that is trained with a ***back propagation*** learning algorithm (Turban et al., 2008). As the name implies, the structure of ***multilayer*** feed-forward network consists of multiple layers of neurons including one input layer, one or more hidden layers, and one output layer (Han and Kamber, 2000a). Although several hidden layers can be placed between the input and output layers, it is quite common to use only one hidden layer. Each layer contains neuron units. The number of neurons in the input and output layers are determined considering the input attributes and target attribute whereas the number of neurons in the hidden layer is determined through experimentation. Neurons of each layer are connected only to the neurons of the subsequent layer and the outputs from each layer are forwarded through the next subsequent layer only. The connection in the network cannot be in a backward direction and cannot skip a layer, hence the name ***feed-forward***. In addition, the network is ***fully connected*** which means all neurons at one layer are connected with each neuron in the next subsequent layer (Turban et al., 2008). A typical structure of multilayer feed-forward network is shown in Figure 2.2. This figure shows a 3-3-2 network structure; this means there are three neurons in the input layer and the hidden layer, and two neurons in the output layer.

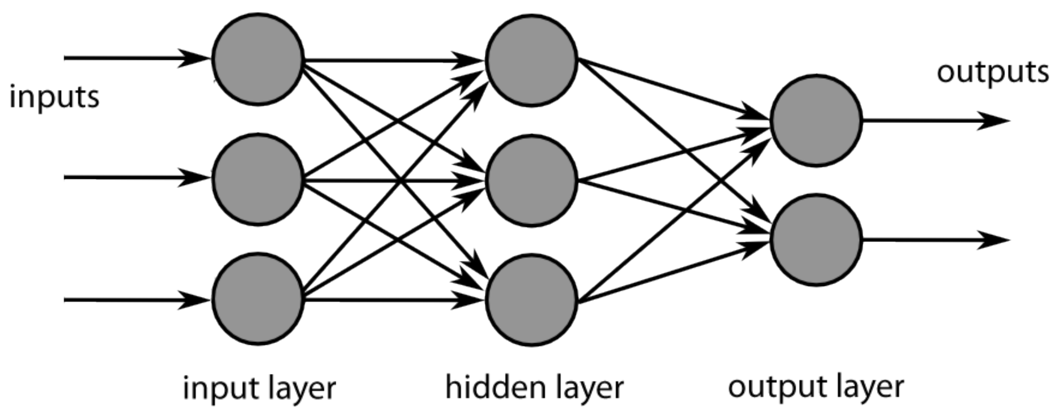


Figure 2.2: A 3-3-2 structure of a multilayer feed-forward network

Each connection between neurons has a weight associated with it. These weights express the relative strength of the input data that contributed to the output. During the learning phase, these weights are adjusted using a back propagation algorithm so that the output generated by the network matches the correct output. The process of training the multilayer feed-forward neural network is an iterative process involved adjusting weights to minimize the difference between the actual and the targets outputs (Chen et al., 2010).

The back propagation learning algorithm applies gradient descent to update the weights in order to minimize the error between the actual output of the network and its target output. The back propagation learning algorithm has four steps (Lippmann, 1987). First, all weights (denoted by W_{ij} in Figure 2.3) are set to some small random values near zero. Second, the input data (denoted by X_i in Figure 2.3) are fed into the input layer of network and the target outputs are specified. Third, the actual outputs (denoted by Y in Figure 2.3) are computed by calculating the weighted sum of each neuron and applying an activation function on the weighted sum. Some of the most commonly used activation functions are binary step function, sigmoid function and hyperbolic tangent function. Finally, an error is calculated as the difference between the target outputs (Z) and the actual outputs (Y), the error is then propagated backward through the network from the output layer to the input layer (see Figure 2.3). Then the weights are modified as the error is propagated and the processes are repeated at the second step until one of the terminated conditions is reached.

There are three conditions for learning termination (Han and Kamber, 2000a): (1) the error between the target output and actual output of the previous iteration drops below the pre-specified threshold; (2) classification error of the previous iteration is lower than the pre-specified threshold and (3) the number of training cycles reaches the pre-specified number of learning iterations.

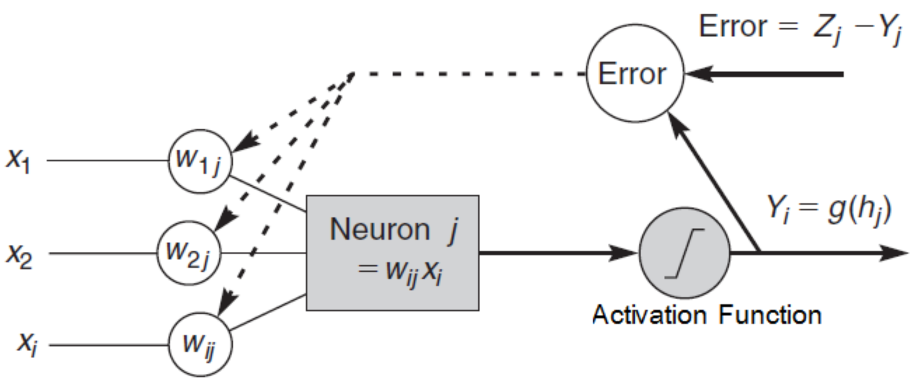


Figure 2.3: Back propagation of Errors for a Single Neuron (Turban et al., 2008)

The back propagation learning algorithm also has two parameters namely ***learning rate*** and ***momentum*** which can be adjusted to control the speed of reaching a solution. The learning rate is a constant typically having value between 0.0 and 1.0. It is an important parameter that controls how much the weights are changed at each iteration. A small value for the learning rate tends to slow down the learning process while a high value for the learning rate may cause network oscillation and inability to converge (Han and Kamber, 2000a). The momentum is a counterbalancing parameter which provides a balance to the learning rate by adding a proportion of the previous weight changes to the current weight changes (Turban et al., 2008).

2.2.2.4 Naïve Bayes

A Naïve Bayes classifier is a simple probabilistic classifier based on applying Bayes' theory with strong assumptions of independence among attributes. When represented as network, the structure of the Naïve Bayes network is shown as Figure 2.4. The root node of network represents class variable (C) and the leaf nodes represent independent variables or attributes ($A_1, A_2, \ldots, A_n$). In addition, there are no edges between attributes in the network as it captures the main assumption of Naïve Bayes classifier that every attribute is independent from others attributes (Friedman et al., 1997; Kotsiantis, 2007).

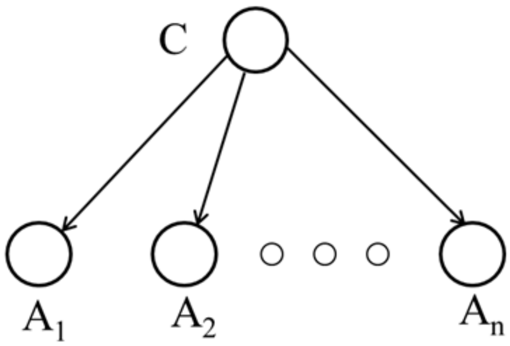


Figure 2.4: The structure of a Naïve Bayes network (Friedman et al., 1997)

Bayes' Theorem is a theorem of probability theory originally stated by the Reverend Thomas Bayes. According to Bayes' theorem, P(H|X) which is the posterior probability, or a *posteriori* probability of H conditioned on X can be computed in terms of probabilities P(X|H), P(H), and P(X) as Equation 2.5 (Han and Kamber, 2000a).

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \tag{2.5}$$

The three probabilities P(X|H), P(H), and P(X) may be able to estimated from the given data. P(X|H) is the a posterior probability of X conditioned on H which is referred to as Likelihood. P(H) is called the prior probability, or a *priori* probability since it is an initial probability value of H originally obtained before any condition is specified. P(X) is called Evidence, indicates the prior probability of X. Therefore, the Equation 2.5 can be written as $Posterior = \frac{Likelihood \times Prior}{Evidence}$.

In the context of classification problems, X is a data sample contained n attributes $X_1, X_2, \cdots, X_m$ with an unknown class label and H is a hypothesis that X belongs to class C. The goal is to determine P(H|X), the probability that sample X belongs to class C, given the attribute description of X. Suppose that there are m classes, $C_1, C_2, \cdots, C_m$. X will be labelled following to the class having the highest posterior probability, conditioned on X (Han and Kamber, 2000a).

Similar to other classification techniques, Naïve Bayes classification consists of training and testing phases. In the training phase, the distributions of P(H) and P(X_i|H) for each attribute are computed from the training data. P(H) indicates the probability of each class C_i and is computed as the number of records with class C_i divided by the number of all training records. For the discrete input attributes, P(X_i|H) is computed as the number of records of class C_i having attribute value u for X_i divided by the number of records of class C_i. These computed probabilities form the conditional probability tables, the output of the training phase (Han and Kamber, 2000a).

Using the conditional probability tables, a class of new data X' can be predicted in the training phase. The probabilities corresponding to the attribute value of X' in the conditional probability tables are used to compute the posterior probability of each class C_i regarding Bayes' theorem and then the class C_i having the highest posterior probability will be assigned as class of X' . Specifically, the posterior probability of each class C_i is calculated as the multiplication among the probability of input attributes P($X_i'|C_i$) and probability of class C_i since the independence assumption of Naïve Bayes is applied and the denominator of Bayes' theorem has the same value for all classes, referred to as constant so it can be omitted from the formula.

The main advantage of Naïve Bayes is its simplicity which results in low computational complexity. The training is very easy and fast using the empirical frequency and testing is straightforward - just looking up the conditional probability tables and performing probability multiplication. In addition, it is capable of handling large data sets very quickly and attains accurate results (Kotsiantis, 2007). Moreover, studies comparing classification algorithms (Domingos and Pazzani, 1997) have shown that Naïve Bayes produces higher accuracy than more sophisticated approaches including decision tree induction, instance-based learning, and rule induction in many domains. The main disadvantage of Naïve Bayes is its reliance on the assumption of independence among attributes which is often unsatisfied with real data and causes the accuracy of Naïve Bayes to be less than other more sophisticated techniques (such ANNs) (Kotsiantis, 2007).

2.2.2.5 Bayesian Networks

Bayesian Networks (BNs) known as belief networks, are directed acyclic graphs (DAG) with an associated set of probability tables (Heckerman, 1997). BNs consist of a qualitative part which represents variables as nodes and relationships among variables as directed edges in DAG, and a quantitative part which quantifies dependency intensity between variables in probabilistic terms (Ben-Gal, 2007). More specifically, for each variable (node) the probabilistic term is encoded in the conditional probability distribution that quantifies the influences of its parent nodes. For discrete variables, the conditional probability distribution of non-root nodes is captured in table called conditional probability tables (CPTs) which contain a local probability of node for each combination of the values of its parents whereas the conditional probability distribution of root nodes (the nodes without parents) depend solely on their probability distribution known as a prior probability distribution (Lee and Abbott, 2003).

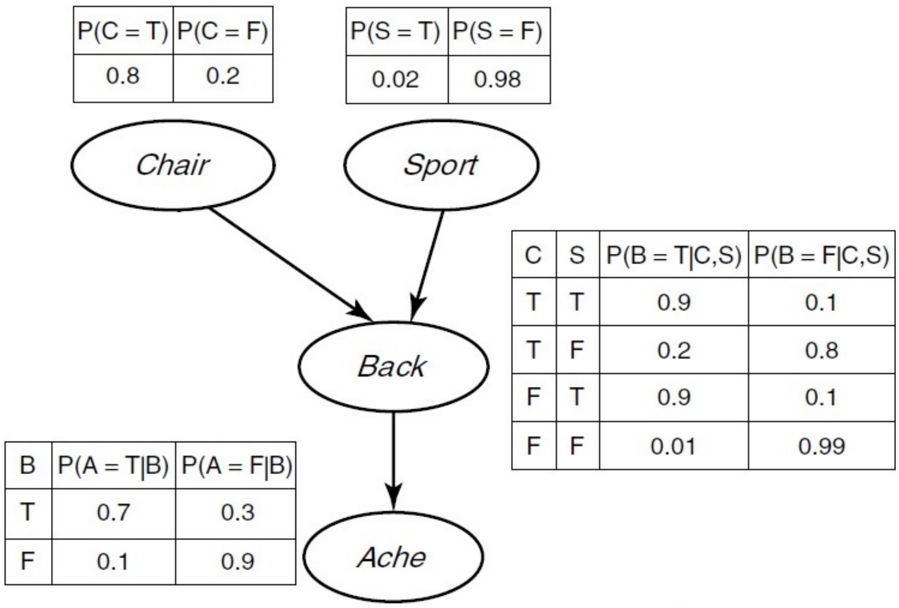


Figure 2.5: The backache BN example adapted from (Ben-Gal, 2007). Each node represents variable with its conditional probability table.

The example of BNs shown in Figure 2.5 illustrates that backache might be caused by wrong sport activity or an uncomfortable chair. In this example, the node *Chair* and *Sport* (denoted by C and S respectively) are the root nodes hence the probability tables associated with these nodes contain their prior probabilities. The nodes *Back* and *Ache* (denoted by B and A respectively) are the non-root nodes hence their CPTs contain conditional probabilities for all possible value of their parents. For example, the parent nodes (C,S) of *Back* have two values (True, False) therefore the CPT of *Back* contains four rows which each row represents probability of *Back* for each combination of values of node C and S.

BNs can be used to compute the conditional probability of one node, given the values of other nodes therefore BNs can be used as a classifier that predicts class value (dependent variable) regarding the posterior probability distribution of the class node given the values of input attributes (independent variables) (Cheng and Greiner, 2001). BNs improve the performance of Naïve Bayesian classifiers by avoiding the independence assumption (Friedman et al., 1997) because BNs have an ability to capture the relationships among input variables which provides tremendous value in exploring data.

Applying Bayesian network techniques to classification consists of a training (learning) and a testing phase. The BNs learning involves a model construction and BNs testing involves a model inference for predicting class (Cheng and Greiner, 2001). Generally, there are two major tasks in BNs learning from data: structure learning and parameter learning (Ben-Gal, 2007). The first aims to construct the DAG encoding the independencies among variables that are optimal for a given training data while the latter aims to construct the CPTs for the learned network. The BN structure learning is considered a harder problem than parameters learning since the number of possible structures is extremely large; finding the optimal structure of BN from data is considered as an NP-hard problem (Khanteymoori et al., 2009). Further information about structure learning can be found in Khanteymoori et al. (2009). The parameters learning task is estimating the conditional probability distribution of all nodes in the learned network from the training data. According to Lee and Abbott (2003), in the case of no missing values in the training data, the parameter estimation can be done simply by using the empirical conditional frequencies from the data given the network structure, which is similar to the probability computing involved in Naïve Bayes classification . In the case that missing data exists, the algorithm for parameter estimating is used and the most commonly used algorithm is the Expectation Maximization (EM) algorithm.

The result of the learning phase is a learned network representing an approximation to the probability distribution of all variables, and is used as a classifier to predict the class of the testing dataset in the testing phase. More precisely, the learned network is used to compute the posterior probability of values of the target (class) variable given the values of the input variables in the testing dataset. After that, the class value that attains the maximum posterior probability is assigned as the class label of the instance in the testing dataset (Friedman et al., 1997). Note that, the conditional independence statement encoded in the BN network is exploited in computing the posterior probability. This conditional independence states that each variable is independent of its non-descendants in the graph given the state of its parents. Thus, applying this statement provides an efficient way to compute the posterior probability by reducing the number of parameters needed to characterize a joint probability distribution (Ben-Gal, 2007).

The BNs offers several advantages over other types of classification technique. The major advantages of BNs are (Heckerman, 1997; Lee and Abbott, 2003): (1) the graphical diagram of BNs provides a better understanding about the inter-relationships among the variables which help a human expert to be able to modify a BNs model to obtain better predictive model; (2) BNs have an ability to capture the relationships among input variables which provides tremendous value in exploring data and (3) BNs allow a human expert to add their knowledge related to a domain in the form of network topology which contributes to the model efficiency; (4) BNs are flexible in handling the missing values and (5) BNs offer an efficient approach to avoid over fitting of data based on probability theory. One of the limitations of BNs is the ability to deal with continuous data thus this data has to be discretized before constructing BNs model (Nyberg et al., 2006).

2.3 Customer Satisfaction

Customer satisfaction has long been an area of interest in academic research and business practice, especially as it is postulated as one of the main indicators of business performance for the last decade. To provide a basic understanding about customer satisfaction, this section provides the review of customer satisfaction including the definitions and importance of customer satisfaction, measuring and analysis of customer satisfaction, and research on mining customer satisfaction data.

2.3.1 Overview of customer satisfaction

Customer satisfaction is an evaluative judgement of quality and intensity that a product or service itself, or a feature of it, does fulfil expectations (Bosnjak, 2014). Customer satisfaction can also be defined as "person's feeling of pleasure or disappointment which resulted from comparing a product's perceived performance or outcome against his/her expectations"(Kotler and Keller, 2011). Many studies on customer satisfaction have been defined with different definitions of customer satisfaction as organisations attempt to measure it. Accordingly, Millán and Esteban (2004) summarized a list of customer satisfaction definitions which can be referred to for further reading.

The underlying concepts of the various definitions are that (1) satisfaction is a post-consumption evaluative judgement which is a comparison of perceived outcomes of product or service with expectations (Pizam et al., 2016). Accordingly, satisfaction with a product or service is a construct that requires experience and use of a product or service. As a matter of fact, the customer satisfaction is actually about consumer satisfaction rather than about buyer satisfaction; (2) satisfaction is a feeling. It is a temporary attitude that can readily change given a group of circumstances (Hom, 2000). Hence, satisfaction is not a universal phenomenon and not every customer gets the same satisfaction out of the same product usage or service experience as he/she has different needs, objectives and personal experiences that influence his/her expectations (Pizam and Ellis, 1999).

In an increasingly competitive environment, customer orientation is essential in many businesses (Dejaeger et al., 2012). Accordingly, customer satisfaction has been getting importance in business practices especially for the last decade. It is postulated as one of the main indicators of business performance (Mihelis et al., 2001) as measuring customer satisfaction provides a reflection of an organisation's business activities. In addition, measuring customer satisfaction also provides meaningful and objective feedback about customer preferences and expectations which is useful for organisations in adjusting their operations and marketing to meet the customer needs and expectations.

The supporting evidence for the importance of customer satisfaction is that customer satisfaction measurements on a national scale such as the American Customer Satisfaction Index (ACSI)[2] which was established in 1994 to provide information regarding customer satisfaction with the quality of household products and services available to U.S. consumers (Fornell et al., 1996). Following the same line, several countries have developed the national measure of customer satisfaction for their own national economies such as the UK Customer Satisfaction Index (UKCSI)[3], South African Customer Satisfaction Index (SAcsi)[4] and Indian Customer Satisfaction Index (ICSI)[5], etc. These measurements of customer satisfaction on a national scale provide uniform and comparable information that allows for systematic benchmarking over time and across firms.

Furthermore, customer satisfaction is considered as one key to customer retention (Kotler and Keller, 2011) as a number of studies suggested that it positively influences on customer retention (Cengiz, 2010). A high level of satisfaction (with pleasurable experiences) is very likely to result in customer retention, product repurchase and the spread of positive word-of-mouth. In contrast, a very low level of satisfaction (with unpleasurable experiences) is very likely to result in abandoning the firm and even the spread of negative word (Kotler and Keller, 2011).

2.3.2 Measuring customer satisfaction

As previously explained, measuring customer satisfaction provides useful information on how an organisation delivers a product or service to the marketplace. In order to measure customer satisfaction, customer satisfaction surveys and questionnaires are used to determine customer attitudes and perceptions of the quality of the product or service.

Typically, in a social science context, satisfaction is referred to as a latent variable as it cannot be directly measured whereas variables that can be directly measured such as sex, age, weight are called manifest variables (Kenett and Salini, 2012, p.3). The latent variable is indirectly measured through a number of manifest variables. Each

[2] http://www.theacsi.org/
[3] https://www.instituteofcustomerservice.com
[4] http://www.sacsi.co.za/
[5] http://www.icsi.org.in/

manifest variable refers to one specific aspect of the latent variable. Taking the number of manifest variables together yields a meaning to the latent variable and builds the contents of the latent variable (Cassel, 2006).

Therefore, in order to design a customer satisfaction questionnaire, a set of manifest variables must be identified (Kenett and Salini, 2012, p.3) and these variables are usually formulated as a set of questions in the questionnaire (Figini and Giudici, 2007). Accordingly, for a questionnaire design, each component or attribute that is relevant for explaining the overall customer satisfaction should contain at least three questions (Cassel, 2006).

For example, in order to assess customer satisfaction in the private bank sector, it is necessary to identify attributes that characterize this type of service and then identify a list of questions for each attribute. According to Mihelis et al. (2001), some possible attributes of the bank sector are personnel of the bank, products, image of the bank, service and access. The possible questions attached to the attribute named personnel of the bank include all the characteristics concerning personnel such as skills and knowledge, responsiveness, communication and collaboration with customers, and friendliness.

Although satisfaction is best measured on a continuous scale, a Likert numeric scale is commonly used to measure satisfaction both in academic and business practice (Allen and Wilburn, 2002; Kenett and Salini, 2012, p.58). The value of scales is limited and it represents the degree of satisfaction perceived by the customer, which is typically ranges from the lowest point on the scale (representing the case when a customer is not satisfied at all) to the highest point (representing the case when a customer is completely satisfied).

In order to use the Likert scale, the choice of measurement scales has to be considered. The range of possible responses for a scale can vary and there has been much debate on an optimal number of responses. By now, five to seven-point scales are well established for closed-questions (Fink, 2003, p.57), (Kenett and Salini, 2012, p.137). The five-point scale is typically used although there are arguments in favour of the seven-point scale (Jamieson et al., 2004).

By using the Likert scale, responses are considered as ordinal data (Fink, 2003, p.52) which means that the response levels have a rank order and the interval between points on the Likert scale cannot be assumed equal (Jamieson et al., 2004). For example, no one can say that the distances between 'Completely satisfied' and 'Satisfied' is the same about the distances between 'Satisfied' and 'Neither satisfied or dissatisfied'.

Despite the fact that Likert-derived data is ordinal, it has become a common practice to treat ordinal scales as interval scales. Hence, it is possibly to employ parametric tests instead of non-parametric tests in analysis Likert-derived data. However, the researcher should consider the sample size and data distribution in determining whether parametric or non-parametric tests could be used to analyse such data (Jamieson et al., 2004).

2.3.3 Analysis of Customer satisfaction data

General data analysis techniques of customer satisfaction are statistical techniques such as descriptive statistics, regression and correlation analysis. Some advanced and modern techniques for analysis customer satisfaction data are Structural Equation Modelling (SEM), Bayesian Networks (BNs), etc.

2.3.3.1 Descriptive statistics

Descriptive statistics are the basic measures used to summarize and describe satisfaction survey data. The reporting of descriptive statistics generally consists of the percentage of individuals choosing a particular response to each question on a survey, measures of central tendency (such as the mean, median and mode), and measures of dispersion (such as the standard deviation, the variance, and the standard error) which describe the variability of response data (Goff et al., 2002).

Additionally, the reporting of descriptive statistics also includes a graphical display of data in the form of charts and graphs which facilitates an identification of patterns in the data. Commonly used charts and graphs are bar chart, pie chart and histogram. Note that the choice of descriptive statistics is dependent on the type of data being observed. Further detail about descriptive statistics can be found in (Field, 2009, p. 31-60; p.87-130).

While descriptive statistics provide some information that can be used during the preliminary analysis or used in conjunction with some advance statistics, they are not able to provide an in depth analysis of customer satisfaction (Grigoroudis and Siskos, 2009, p. 28) (such as key driver analysis). Therefore, an analysis of customer satisfaction should not stop at this analysis level.

2.3.3.2 Statistical and modern approaches

Multiple linear regression (MLR) is one of the most widely used statistical techniques for analysing customer satisfaction data. The technique is used to determine which attributes of product or service are the key drivers of customer satisfaction by discovering the relationship between the satisfaction/performance of the set of attributes (independent variables) and the overall customer satisfaction judgement (dependent variable) (Grigoroudis and Siskos, 2009, p.28). An overview of MLR has already been provided in Section 2.2.2.

One statistical technique that can be used to determine key driver attributes is correlational analysis. Key driver attributes are those that are most closely associated with overall satisfaction. Further detail about correlation can be found in (Field, 2009, p. 166-196).

Another statistical technique widely used in analysing customer satisfaction data is factor analysis. Its most common application is the reduction of a large number of attributes in the steps of designing a customer satisfaction survey.

Apart from these basic statistical techniques, some advance techniques for analysis customer satisfaction data are SEM which is a statistical technique for measuring relationships among latent variables (Grigoroudis and Siskos, 2009, p.34-40), conjoint analysis which is used to assess the effects of the trade-offs made by customers, when they evaluate their satisfaction for a particular product or service (Grigoroudis and Siskos, 2009, p.41). Additionally, there are some modern techniques in customer satisfaction data analysis such as BNs, CUB models and Rasch model. Details of these techniques can be found in (Kenett and Salini, 2012, p.193-213; p.231-279).

2.3.4 Research on mining customer satisfaction data

To gain better insight into customer satisfaction in various domains, some researchers have started applying data mining techniques in analysis of customer satisfaction data. Through the literature review, the summary of 10 publications is presented in Table 2.2. With regard to this table, the main objective of these previous works is to investigate which data mining technique offers the best predict customer satisfaction. Thus, the list of observed techniques and the evaluation metric are provided in these works.

The majority of the previous works conducted a comparison between the combination of two data mining techniques such as neural network, ordinal logistic regression (OLR), multiple linear regression (MLR), decision tree and Bayesian networks (BNs). An exception was the works by Azzalini et al. (2012), Dejaeger et al. (2012) and Nikolaos (2009) which conducted a comparison among three or more techniques.

Azzalini et al. (2012) conducted the comparison of eight data mining techniques in which some additional selected techniques are k-nearest neighbour, support vector machine, bagging trees and random forests. Dejaeger et al. (2012) conducted the comparison of four data mining techniques: OLR, neural network, decision tree and support vector machine (SVM). Nikolaos (2009) conducted the comparison of two data mining techniques namely neural networks and rule-induction with the other techniques in the field of customer satisfaction analysis namely Rough set and Multi-criteria decision analysis.

The common evaluation metrics used for comparing between or among the techniques are misclassification rate (the ratio of number of samples that are incorrectly classified to

Table 2.2: Publications that applied data mining techniques in analysis of customer satisfaction data

Citation(s)	Techniques	Dataset	Evaluation metrics	Result
Thomas and Galambos (2004)	MLR (stepwise) Decision tree	1698 responses to student satisfaction	n/a	The analysis of data using two techniques yield a comprehensive view for understanding student satisfaction. Since decision tree reveals a different perspective for understanding student satisfaction from that of MLR.
Nikolaos (2009)	Multi-criteria decision analysis Rough Sets Neural networks Rule-induction	524 responses to customer satisfaction of a big Greek shipping enterprise.	Correct classification percentage	Multi-criteria decision analysis offers better results on the prediction of customer satisfaction followed by rule-induction
Yay and Akıncı (2009)	OLR Neural networks	314 responses to student satisfaction	Correct classification percentage	Neural network has better satisfaction classification than OLR for train and test data.
Larasati et al. (2011)	OLR Neural networks	257 responses to customer satisfaction of restaurant	Misclassification rate	Neural networks is superior to OLR based on the different of misclassification rate resulted in training and testing.
Azzalini et al. (2012), p. 192-205	OLR MLR SVM Neural networks Decision tree k-nearest neighbour Bagging trees Random forests	4515 responses to customer satisfaction of IT company	Misclassification rate of dissatisfied customers Misclassification rate	MLR and OLR were recommended to marketing managers as they have specific optimality on the evaluation metrics.
Dejaeger et al. (2012)	OLR SVM Decision tree Neural networks	9300 of class evaluation response by students	Correct classification percentage Area under ROC curve Classifier comprehensibility	OLR performed best on two classifier performance evaluations

Table 2.2 (cont): Publications that applied data mining techniques in analysis of customer satisfaction data

Citation(s)	Techniques	Dataset	Evaluation metrics	Result
Huang (2012)	OLR Neural networks	425 responses to customer satisfaction of four hotels in Taiwan	Correct classification percentage	Neural network is more suitable than OLR for predicting customer satisfaction and repurchase behaviour.
Klicek et al. (2014)	BNs Neural networks	239 responses to customer satisfaction of a street festival	Correct classification percentage	A combination model of the two techniques is able to predict customer satisfaction on the high level of accuracy.
Perucca and Salini (2014)	OLR BNs	World Bank's Railways database.	Misclassification rate	BNs is superior to OLR based on the misclassification rate tested by using cross-validation.
Tama (2015)	Neural networks Decision tree	340 responses to customer satisfaction of fast-food restaurant	Correct classification percentage	Neural network has better satisfaction classification than decision tree.

the total number of samples) and correct classification percentage (the ratio of number of samples that are correctly classified to the total number of samples multiplied by 100). The best predicting customer satisfaction technique is the one with the lowest misclassification rate or the one with highest correct classification percentage.

For a comparison of results among data mining techniques, Yay and Akıncı (2009), Larasati et al. (2011) and Huang (2012) reported that neural network is superior to OLR in predicting customer satisfaction. This contradicts the results of Azzalini et al. (2012) and Dejaeger et al. (2012) which reported that OLR is superior to neural networks and the other techniques such as decision tree and SVM in predicting customer satisfaction.

Consider the comparison results of the other two works by Perucca and Salini (2014) and Tama (2015) which compared different sets of the techniques from the previously stated works. The first reported that BNs is superior to OLR in predicting customer satisfaction, while the latter reported that neural network is superior to decision tree in predicting customer satisfaction. In addition, the work by Nikolaos (2009) reported that Multi-criteria decision analysis offers better results on the prediction of customer satisfaction followed by the rule-induction.

The rest of the publications by Thomas and Galambos (2004) and Klicek et al. (2014) did not directly compared the two data mining techniques using evaluation metrics therefore the comparison results were not reported. In their work, Thomas and Galambos (2004) investigated the rules generated from linear regression and decision tree on student satisfaction data and concluded that the analysis of data using these two techniques yielded a comprehensive view for understanding student satisfaction. Klicek et al. (2014) proposed an approach to combine BNs with neural networks in predicting customer satisfaction. Specifically, BNs was used as a feature selection technique to reduce the number of input variables of the neural network. Based on the correct classification percentage, Klicek et al. (2014) reported that a hybrid technique was able to predict customer satisfaction at a high level of accuracy (over 90%).

These past studies of mining customer satisfaction focused on building a model to predict the customer satisfaction. They did not provide further steps for deriving *importance*, which is the main component of IPA, from the models of data mining techniques. The point of interest drawn from these studies is that data mining is a promising tool to discover some useful information from customer satisfaction data. These past studies also suggested some interesting data mining techniques which are neural network, BNs, decision tree and SVM that could be used to analyse and estimate *importance* from customer satisfaction data. In another word, the review of the past studies raised the idea of introducing new data mining techniques in the field of IPA studies.

2.4 Summary

In this chapter, three principal topics relating to the research are described namely – SWOT analysis, data mining and customer satisfaction. First, SWOT analysis is reviewed to examine the state of the art of the SWOT approach. The overview of SWOT analysis shows that this simple method is important as part of strategic planning. It supports companies to gain better insight into their internal and external business environments enabling them to generate good plans to achieve their business goals. The main advantage of SWOT analysis is its simplicity which has resulted in its continued use since the 1960s. Despite its advantage, there are typical shortcomings existing in the SWOT approach as it produces a superficial and imprecise list of factors, relies on subjective perception of a company's staff and lacks factor prioritization regarding the importance of each SWOT factor. Due to the disadvantage in prioritization of SWOT factors, a number of researchers proposed new variations of SWOT approaches that integrated SWOT with others quantitative methods such as AHP-SWOT, FAHP-SWOT and ANP-SWOT. These approaches make SWOT factors commensurable regarding their relative importance. However, both the quantitative SWOT and traditional SWOT approaches do not take customers' points of view into consideration even if it can ensure that the capabilities perceived by a company are recognized and valued by the customers. Therefore, this research aims to fill this gap in previous SWOT approaches. In term of the literature about the SWOT analysis system, techniques for developing the SWOT analysis system are mainly expert systems, text mining and sentiment analysis.

Data mining is the core step of KDD which aims to extract useful knowledge from a huge amount of data. Data mining relies on the disciplines of statistics and of machine learning in data analysis. Data mining methods can be divided in two groups: prediction and description. The predictive data mining method infers the extracted pattern of current data to predict the pattern of future data. Examples of predictive data mining methods are Classification and Regression. The descriptive data mining method exploits extracted patterns for describing common characteristics of data. Examples of descriptive data mining methods are Clustering and Association rule discovery. This report is mainly focused on methods of prediction and classification technique namely – MLR, OLR, BPNN, Naïve Bayes and BNs that will be applied to customer survey data to determine the company's attribute importance. For each technique, steps for training and testing classification or prediction model are described along with its advantages and disadvantages.

Customer satisfaction has long been an area of interest in academic research and business practice and has been getting importance in business practices especially for the last decade. Measuring customer satisfaction provides meaningful and objective feedback about customer preferences and expectations which is useful for organisations in adjusting their operations and marketing to meet the customer needs and expectations resulting in a competitive advantage. Typically, customer satisfaction is measured through the questionnaires using closed-ended or open-ended questions. This report is mainly focused on measuring customer satisfaction on a Likert scale through the closed-ended questions. Therefore, techniques for analysis customer satisfaction that were examined are basic statistical techniques such as descriptive statistics, regression analysis and some advanced statistical techniques such as SEM and conjoint analysis. In addition to the statistical techniques, many researchers have started applying data mining techniques to the analysis of customer satisfaction data with the aim of gaining better insight into customer satisfaction. A review of past research on mining customer satisfaction data showed that data mining techniques such as neural network, BNs, decision tree and SVM are promising techniques for discovering some useful information from customer satisfaction data.

Chapter 3

Importance-Performance Analysis

To provide additional background to the research, this chapter explains about Importance-Performance Analysis (IPA). The chapter begins with an overview of IPA (Section 3.1) where the main focus is the approaches that have been used for measuring *importance*. Section 3.2 provides the past comparative studies of method for measuring *importance* followed by the methodology for deriving *importance* by using well-known data mining techniques, Section 3.3.

3.1 Overview of Importance-Performance Analysis

Importance-Performance Analysis (IPA) is a technique proposed by Martilla and James (1977) for analysing customer satisfaction towards a company's product or service . For decades, IPA has been used as a tool for understanding customers' needs and desires so as to develop marketing strategies to respond to them. In a nutshell, IPA measures the satisfaction based on two components: '*performance*' and '*importance*'. The intersection of these two components creates a two-dimensional matrix that helps a company to identify improvement opportunities by discovering the attributes of a company's product or service that should be maintained or improved based on the customers' viewpoints (Garver, 2003).

IPA is widely used in many areas in which customer satisfaction is a key to a thriving business including higher education (Ford et al., 1999; Kitcharoen, 2004; Silva and Fernandes, 2012), tourism (Tarrant and Smith, 2002; Luo et al., 2010; Taplin, 2012a), government service (Seng Wong et al., 2011), health service (Gonçalves et al., 2014) , convenience store (Shieh and Wu, 2009) and banking service (Wu et al., 2012).

The advantage of IPA is based on its simplicity and ease of use. IPA provides an intuitive way to indicate strategic actions to be taken with respect to each attribute of a company's product or service (Tarrant and Smith, 2002). One of the major shortcomings of IPA (as it is a model in quadrants) is that a slight change in the position of an attribute will likely result in identifying dramatically different attributes for improvement (Bacon, 2003).

This shortcoming of IPA is related to two methodological issues of IPA which are I-P mapping partitions and direct vs. indirect measurement of importance (Lai and Hitchcock, 2015; Sever, 2015). The first issue is about approaches for identifying cut-off points of *importance* and *performance* that divide the IPA space. Further detail regarding this issue will be discussed in sub-section 3.1.2

The second issue is associated with the argument whether *importance* should be directly obtained from customer or indirectly derived on the basis of *performance* and overall customer satisfaction. Note that only the second issue is being focused and defined as the research question of this thesis. Further detail regarding the current approaches for measuring *importance* is explained in sub-section 3.1.4

3.1.1 IPA matrix

Since customer satisfaction is a function of customer perceptions, it involves the quality of the company's product or service and customer expectations. Therefore, IPA measures the satisfaction from customer satisfaction surveys based on two components of product or service attributes: *the importance of a product or service* to a customer and *the performance of a company* in providing that product or service (Martilla and James, 1977).

The intersection of these two components creates a two-dimensional matrix (Figure 3.1) where the *importance* is shown by the x-axis and the *performance* of the company by the y-axis.

Depending on cell location, the attributes related to a company's product or service are considered as major or minor strengths and weaknesses. The four quadrants in IPA are characterized by Silva and Fernandes (2010) as:

Quadrant 1. Keep up with the good work (high importance, high performance): this cell contains attributes that are perceived to be very important to customers, and the company seems to provide high levels of performance in these attributes. Thus these attributes are referred to as the major strengths and opportunities for achieving or maintaining competitive advantage;

Quadrant 2. Possible overkill (low importance, high performance): this cell contains attributes that are perceived as of low importance to customers, but the company seems to provide high levels of performance in these attributes. Thus these attributes are referred to as the minor strengths implying that resources

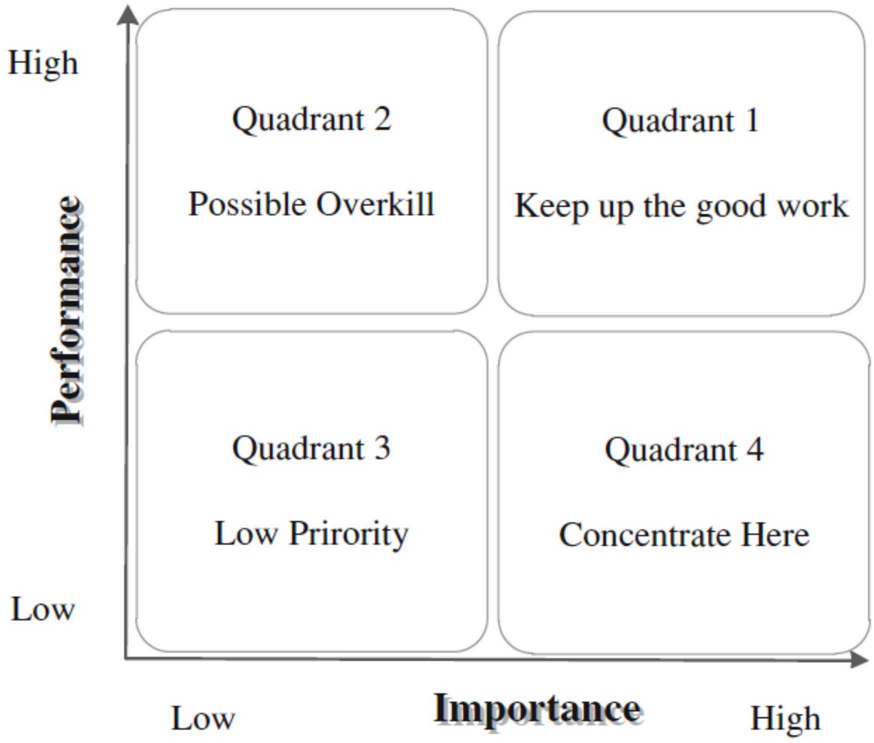


Figure 3.1: The IPA matrix (Hosseini and Bideh, 2013)

committed to factor in this quadrant would be better used in other quadrants in need of improved performance;

Quadrant 3. Low priority (low importance, low performance): this cell contains attributes that are rated as of low importance and low performance. Thus these attributes are referred to as the minor weaknesses that do not require a great deal of priority for improvement;

Quadrant 4. Concentrate here (high importance, low performance): this cell contains attributes that are perceived to be very important to customers but performance levels are fairly low. Thus these attributes are referred to as the major weaknesses that require immediate attention for improvement.

This IPA matrix facilitates the company to prioritize the product or service's attributes that need to be of concern in order to satisfy their customers. For example, the product or service's attributes that are indicated as major weaknesses (Quadrant 4) should receive top priority and be the focus of improvement efforts. In contrast, the product or service's attributes considered as major strengths (Quadrant 1) should be maintained and leveraged to ensure that the company continues to deliver good performance in these areas (Garver, 2003). Additionally, the IPA result can be interpreted as major/minor strengths and weaknesses (Deng et al., 2008a; Silva and Fernandes, 2012; Hasoloan et al., 2012; Hosseini and Bideh, 2013). Therefore, the IPA approach can be used to analyse customer surveys for SWOT analysis.

3.1.2 Developing the IPA matrix

Generally, data regarding customers' perceptions toward a company's product or service gathered via customer satisfaction surveys are analysed for constructing an IPA matrix. According to Duke and Mount (1996), the steps for developing an IPA matrix are described and shown in Figure 3.2.

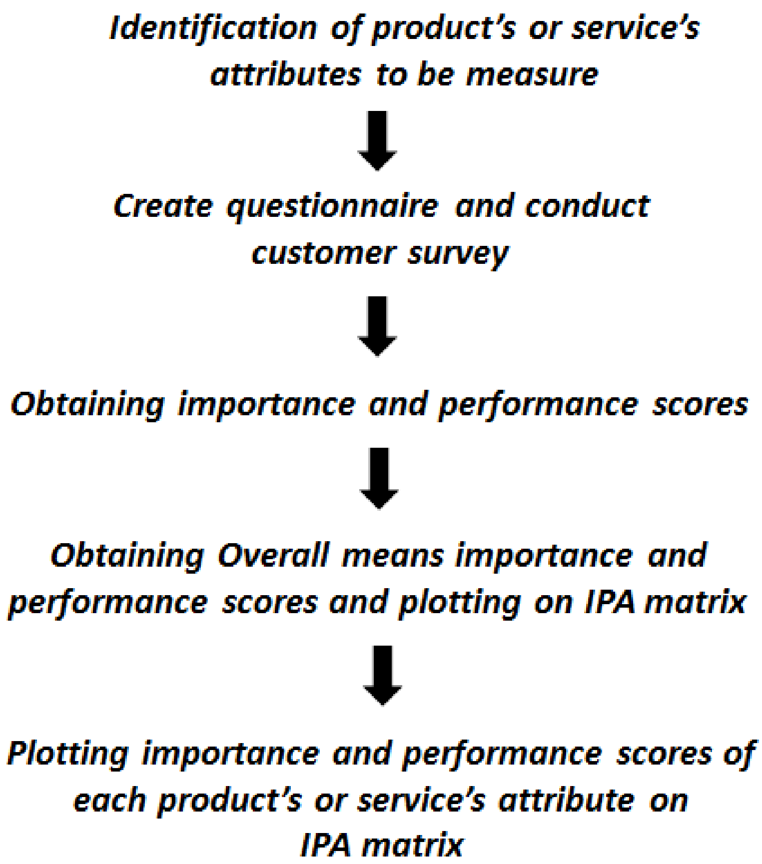


Figure 3.2: Steps for constructing IPA matrix

Step 1. Determine attributes to be measured: This is a critical step for conducting IPA since the usefulness of IPA will be severely limited, if evaluative factors important to the customer are overlooked. Typically, this attributes list is based on a thorough literature review in each application area or by interviews (Martilla and James, 1977).

Step 2. Create questionnaire and conduct customer survey: The product or service's attributes identified in the previous step are then used to develop questionnaires. Typically, the questionnaires contain questions that ask customers to rate the importance of the attribute ranging from "not important at all" to "highly important" and the performance of the attribute ranging from "very dissatisfied" to "very satisfied" in a 5-point or 7-point Likert scale. Next, the questionnaires are then administered to customers for obtaining their response on a product or service (Skok et al., 2001).

Step 3. Obtaining importance and performance: *performance* is typically calculated on a rating scale whereas *importance* can be calculated either on a rating scale (self-stated importance) or estimated on the basis of performance (implicitly derived importance) (Oliver,1997) as cited in (Eskildsen and Kristensen, 2006). In traditional IPA, both *importance* and *performance* are measured based on customer rating scales (Martilla and James, 1977). Other variations of IPA applied different methods to measure the *importance* and *performance* which will be described in the next sections.

Step 4. Plot importance and performance on matrix: Through this step, one of the approaches of I-P mapping partitions is identified to set the cut-off points that divide the IPA space into areas of different priorities. Then the identified attributes are plotted against each other, allowing the analyst to examine the attributes located in each area for generating an action plan.

Generally, there are three I-P mapping approaches that could be used to partition the IPA space: scale-centred quadrants approach, data-centred quadrants approach and diagonal line approach (Bacon, 2003). A scale-centred quadrants approach use the midpoint of the measurement scale as the cut-off points (e.g. the value of 3 on the 5-point Likert scale) (Garver, 2003) whereas a data-centred quadrants approach use either aggregate averages or median value of *importance* and *performance* as the cut-off points.

In a diagonal line approach, an upward 45° diagonal line is used to partition the IPA space (Bacon, 2003). All attributes located on the diagonal line have equal priorities for improvement since their *importance* is equal to *performance* ($I = P$). In case the x-axis represents *importance* and y-axis represents *performance*, attributes located above the diagonal line depict that their *importance* is less than *performance* ($I < P$); hence, they have the lower priorities for improvement than attributes located below the diagonal line ($I > P$).

In addition to these three approaches of I-P mapping partitions, Sever (2015) suggested the possible improvement of I-P mapping partitions by applying receiver operating characteristic (ROC) for evaluating and selecting optimal cut-off points of *importance* and *performance*.

As the majority of researchers use the data-centred quadrant for dividing the IPA space (Sever, 2015), this study will employ this approach for dividing the IPA space. Specifically, cut-off points on the x-axis and y-axis are defined as aggregate averages of *importance* and *performance* to generate a four-quadrant matrix. A further discussion on I-P mapping partition can be found in Azzopardi and Nash (2013); Lai and Hitchcock (2015); Sever (2015).

3.1.3 Approaches for measuring *performance*

Measuring *performance* in the product or service's attributes is the second necessary step in IPA (the first being the identification of product or service's attributes). Levels of performance are traditionally measured as actual performance and relative performance. Both types of performance are measured on customers'ratings in the customer satisfaction surveys. The difference between them is that the first does not take competitors' information into consideration whereas the latter is computed by comparing the company's performance to that of the best competitor (Garver, 2003). In today's competitive marketplace, the company needs to consider its performance relative to the competitors hence the performance should be measured on the basis of relative performance rather than actual performance.

The actual performance for each product or service's attribute can be straightforwardly computed as the mean value from all customers responding to the survey. The relative performance can be computed by subtracting the best competitor's mean performance from the company's mean performance which is called "gap analysis" (Lambert and Sharma, 1990) or dividing the company's mean performance by the best competitor's mean performance which is called "performance ratios" (Higgins, 1998). The result of gap analysis and performance ratio can be interpreted in three ways indicating that the performance of company is greater than, less than, or equal to the performance of its competitor. A positive performance gap and the performance ratio values that are greater than one represent that the company is performing better than its competitor whereas a negative performance gap and the performance ratio values that are less than one mean that the company has a market disadvantage. A performance gap valued zero and performance ratio values that fall in the region 1 ± 0.03 are indicative that the performance of the company is relatively equal to the performance of its competitor.

Another approach using relative performance is called "comparative scales" in which the comparative performance is measured directly on the survey containing questions that asked customers to compare the company's performance to that of the competitor. The possible answers might include "much better than competitor", "equal to competitor", and "much worse than competitor" (Garver, 2003).

3.1.4 Approaches for measuring *importance*

Measuring the *importance* is also a critical step for conducting IPA since different methods are likely to result in identifying dramatically different attributes for improvement.

Two different methods that are commonly used for estimating importance of attributes in IPA are customers' self-stated importance (explicit) and statistically inferred importance (implicit).

The customers' self-stated approach asks survey respondents to rate the *importance* of each product or service's attribute and calculates *importance* based on customer preference (e.g. rating scale, partial ranking). Although it is a commonly used approach, this method has some limitations. Primarily, adding questions for asking customers to rate *importance* increases the survey length which might affect the response rates (Garver, 2003). Secondly, self-stated importance tends to have low discrimination power as customers tend to consider that all attributes are very important (Gustafsson and Johnson, 2004).

Additionally, Matzler et al. (2004) stated that the traditional IPA has two assumptions that are erroneous in practice: (1) *Attribute performance and attribute importance are two independent variables.* (2) *The relationship between attribute performance and overall performance is linear and symmetric.* They pointed out that many researchers have proved that the relationship between attribute performance and attribute importance is causal which means changes in performance lead to changes in importance (Sampson and Showalter, 1999; Matzler et al., 2004) and the relationship between attribute performance and overall customer satisfaction (performance) is asymmetrical (Kano et al., 1984; Matzler and Sauerwein, 2002). Since these assumptions are erroneous in the real world, many researches argued that the customers' self-stated importance is not an adequate method to measure *importance* (Matzler et al., 2003; Deng et al., 2008b).

Consequently, the statistically inferred importance is introduced as a method to measure the *importance* by deriving the relationships between attributes' performances and overall performance measures such as overall satisfaction (Garver, 2003; Tontini and Silveira, 2007; Pezeshki et al., 2009). The different statistical methods for deriving *importance* measures are multiple regression analysis (MLR) (Matzler and Sauerwein, 2002; Pezeshki et al., 2009; Ho et al., 2012), ordinal logistic regression (OLR) (Larasati et al., 2011), structural equation models (SEMs) or partial correlation (Matzler et al., 2003; Deng et al., 2008b; Ban and Bogdan, 2013).

In regression analysis- MLR and OLR, an entire set of attributes' performance (independent variables) are regressed against the overall customer satisfaction (dependent variable). According to this method, the regression coefficients are considered as implicit *importance* which expresses the influence of product or services' attributes on the overall satisfaction. The higher the regression coefficients of attributes, the more important they are in customers' opinion (Pezeshki et al., 2009). SEMs are multivariate regression models in which each structural equation represents the cause-effect relationships among variables in the model. The structural parameters in the models can be estimated by a maximum likelihood (ML) model or partial least squares (PLS) (Gustafsson and

Johnson, 2004; Huang et al., 2007). These structural parameters indicate the effects of the attributes on the overall satisfaction. For partial correlation, the importance of the attributes is determined through the partial correlation coefficient by correlating each attribute's performance with overall satisfaction (Ban and Bogdan, 2013). Specifically, the correlation coefficient is a measure of the linear dependence between one attribute's performance (independent variable) and overall satisfaction (dependent variable) when other attribute's performances are eliminated.

The IPA is sensitive to the importance measurements employed as shown in an empirical study of the two importance measurements (stated and statistically inferred approach) conducted by Matzler et al. (2003). In this empirical study, the *importance* of an individual attribute, which was assessed by two forms of customers' self-stated importance (partial ranking and rating scale) are compared to the *importance* measured by statistically derived importance (correlation analysis). The results of this paper showed that the results from both methods are different and demonstrated that classical IPA (self-stated importance) could be misleading, since the self-stated method does not consider the relationship between attribute importance and overall satisfaction. Another empirical study on the importance measurements drawn by Musa et al. (2010) also reports that the difference of IPA results depended on how the *importance* is measured.

Although the statistically inferred importance is superior to the self-stated importance (Chu, 2002), the use of these statistical methods depends on some underlying assumptions that are always violated in the real world (Garver, 2003). First, data are relatively normal and second, the relationships between independent and dependent variables are linear. Additionally, both multiple regression analysis and structural equation models have another limitation that is the collinearity between the attributes (multicollinearity). The presence of multicollinearity can result in reducing the effects of the individual attributes on overall satisfaction (Hanson, 1992) which may lead to misleading and unusual results.

Recently, data mining techniques such as Artificial Neural Networks (ANNs) have become an alternative method for determining *importance*. ANNs implicitly detect complex non-linear relationships between independent (attributes' performance) and dependent variables (overall satisfaction). Moreover, ANNs are capable of dealing with collinearity between the independent variables. Since ANNs overcome limitations of traditional statistics, ANNs are a more suitable method to derive importance than statistical methods which can be biased and misleading (Deng et al., 2008a). For example, Krešić et al. (2013) conducted an empirical study that compares three methods for deriving importance namely correlation, multiple linear regression and back-propagation neural networks (BPNN). This study revealed that BPNN has superior predictive validity than multiple linear regression.

Several researchers have applied ANNs for obtaining *importance*, for instances Tsaur et al. (2002) measured *importance* by detecting the relationship between eight service aspects and attitudinal loyalty measures such as repurchasing intention using BPNN. These *importance* measure are then compared with those derived using a logistic regression model. Deng et al. (2008a) proposed a BPNN-based IPA that integrates the BPNN, three-factor theory and natural logarithmic transformation for customer satisfaction improvement. The BPNN model includes the 20 attributes as the input layer and overall customer satisfaction as the output layer. Hu et al. (2009) proposed a revised IPA approach that combines BPNN with decision making trial and evaluation laboratory (DEMATEL) studies in order to put forth a more reasonable methodology of analysis for importance and performance. In short, this paper calculated the integrated combinative importance by adding the *importance* derived by using BPNN with the values represented the cause-effect degree of the product or service's attributes regarding the concept of DEMATEL.

Chen et al. (2010) developed an Importance-Performance Strategy Matrix (IPSM) that measures *importance* by using BPNN and measured *performance* using the concept of the service gap (the difference between expected service and perceived service). Mikulić and Prebežac (2012) proposed an extended BPNN-based IPA that used both stated and implicit importance in decision-making. This paper defined *importance* as attributes' relevance and implicit importance measured by using BPNN as determinance. These two importance measurements are then used to construct a relevance-determinance matrix to provide an enriched informational basis for decision-making. In addition, a competitive performance matrix is constructed regarding a company's performance and a competitor's performance to ensure sufficient managerial action in a competitive environment. The result of this extended BPNN-IPA is compared to the result of conventional BPNN based IPA. The comparison revealed that the conventional BPNN based IPA could produce misleading results.

Hosseini and Bideh (2013) proposed a new framework called SOM-BPNN-IPA that applies self-organizing maps (SOM) for customers' segmentation then BPNN is used for measuring *importance* and developing individual IPA matrices for each market segment. The customers' segmentation that is incorporated into IPA considers the heterogeneity in customers' preferences and characteristics. Thus, this new framework could increase reliability and applicability of IPA results compared to a conventional BPNN-based IPA approach.

The summary of publications that implicitly derived importance using statistical and data mining techniques is shown in Table 3.1.

Table 3.1: Publications that implicitly derived *importance*

Citation(s)	***Method for measuring importance***	***Method for measuring performance***	***Remark***
Matzler et al. (2003)	partial correlation analysis	Actual performance (rating scale)	-
Deng et al. (2008b)	partial correlation analysis	Relative performance (performance ratio)	-
Ban and Bogdan (2013)	partial correlation analysis	Actual performance (rating scale)	Using the alternative partitions of the IPA grid
Pezeshki et al. (2009)	MLR	Actual performance (rating scale)	-
Ho et al. (2012)	the combination of MLR and DEMATEL	Actual performance (rating scale)	-
Deng et al. (2008a)	BPNN	Actual performance (rating scale)	-
Hu et al. (2009)	the combination of BPNN and DEMATEL	Actual performance (rating scale)	-
Chen et al. (2010)	BPNN	Actual performance (Service gap) gained from expected performance and perceived performance	-
Mikulić and Prebežac (2012)	stated and implicit importance using BPNN	Actual and relative performance (Gap analysis)	Provide 2 matrices: (1) Relevance-determinance matrix and (2) Competitive performance matrix
Hosseini and Bideh (2013)	BPNN	Actual performance (rating scale)	Provide individual IPAs for each customer segment clustered by SOM

3.2 Previous comparative studies of methods for measuring *importance*

Although the issue of implicitly derived versus self-stated importance has been debated for a period of time, as well as an increase in several methods used for measuring *importance*, there were a few empirical studies that compared the different methods against a set of evaluation criteria (Azzopardi and Nash, 2013). The summary of previous comparative studies is shown in Table 3.2 and their detail is explained in the following paragraphs.

Four previous empirical studies compared implicitly derived to self-stated importance measures. One of the first empirical studies was reported by Hauser (1991). Hauser compared three self-stated importance measures (including direct-rating scales, anchored scales and constant-sum scales) to one statistically derived importance measures using linear regression. These methods were compared on the basis of predictive validity (ability to predict actual overall customer satisfaction), relative predictive validity (ability to predict customer-reported priority), and face validity. The result of Hauser (1991) supported the use of stated importance measures since all three stated importance measures correlated highly with customer preference rank-order and had high face validity relative to that of the regression-based importance measure.

Consistent with the finding of Hauser (1991), the second study by Bacon (2003) reported that self-stated importance measures using direct-rating scales performed better than correlation-based and regression-based importance measures since, the relative predictive validity estimated as the mean of adjusted R-squared of direct-rating scales across 15 datasets was higher than that of correlation and regression. Although the self-stated importance measure was superior to the two statistically derived importance measures, its adjusted R-squared of half datasets was considered fairly low which suggested that it was a poor predictor of consumer priorities in those datasets.

The third empirical study by Gustafsson and Johnson (2004) reported the comparison of implicitly derived with self-stated importance measures on three datasets. Gustafsson and Johnson (2004) compared five statistically derived importance measures (including linear regression, normalized pairwise estimation, reflective/formative partial least squares, and principal components regression) to one self-stated importance measure using direct-rating scales. These importance measures were compared against three evaluation metrics which are predictive validity, diagnosticity (the ability to identify the consumers' most important attributes) and incidence of negative measures. Results showed that there is no clear winner among statistically derived importance measures. Each method of deriving *importance* has its own advantages and limitations. On the comparison of self-stated and implicitly derived importance measures, the results showed that the former reflected better attribute importance than the latter with respect to incidence of negative measures as the scale values are usually positive. The latter however, reflected better attribute importance than the former regarding both predictive validity and diagnosticity.

The fourth study by Taplin (2012b) provided additional evidence that self-stated relative importance was better than statistically derived importance measures using regression. Taplin (2012b) compared several models of absolute and relative self-stated importance with regression-based importance measures and the models using both methods with the predictive validity. The comparison of models using a single method showed that a regression-based importance measure yielded the highest predictive validity closely followed by the self-stated relative importance measure interpreted on the power scale (denoted as SumRPpow).

However, Taplin (2012b) suggested that regression gained an advantage of predicting on the same data from which it was calculated. Hence, a further test on Mean Squared Error (MSE) under a fair comparison using leave-one-out cross-validation was conducted and results showed that SumRPpow provided higher predictive accuracy than regression-based importance measure. This led to the conclusion to support the use of self-stated relative importance measures. Additionally, Taplin (2012b) reported that the importance measure based on both self-stated relative importance and statistically derived importance produced the highest predictive validity.

In contrast to the first four studies, some researchers concentrated solely on comparing either self-stated importance measures or implicitly derived importance measures. Cohen (2003) compared three self-stated importance measures including direct-rating scales, paired comparisons and Maximum Difference scaling (MaxDiff), and found that MaxDiff was superior to direct-rating scales and slightly superior to paired comparisons regarding to relative predictive validity and discriminating power (ability to provide discriminating measures).

In a similar vein, Chrzan and Golovashkina (2006) extended the previous comparative study by Cohen (2003) to add four more self-stated importance measures (including magnitude estimation, unbounded ratings, constant sum, and Q-sort) to compare with direct-rating scales and MaxDiff. These six self-stated importance measures were compared against three evaluation metrics which are task length (average time to complete task), discriminating power and predictive validity. The comparative results showed that MaxDiff has the highest predictive validity and greatest discriminating power in exchange for longest task length. Q-sort was the second highest ranked method as its predictive validity was very nearly as good as MaxDiff and its discriminating power was also ranked after MaxDiff. Q-sort also had the advantage of taking less time to complete than MaxDiff and other methods except direct-rating scales. The two lowest ranked self-stated importance methods were unbound rating and direct-rating scales since both methods had fairly low predictive validity and discriminating power.

A recent empirical study by Pokryshevskaya and Antipov (2014) focused on comparing nine statistical methods for deriving importance (including Pearson correlation, Kendall correlation, Spearman correlation, linear regression, partial least squares, multilayer perceptron, Shapley value decomposition, first-round effect, and Fields decomposition). These nine statistical methods for deriving importance were compared on the basis of diagnosticity (modified from (Gustafsson and Johnson, 2004)) and stability (a robustness of importance estimation to a small change of data). The comparison based on two real-world datasets showed that none of the methods appeared to be the best regarding both metrics simultaneously. Considering the average rank of methods according to diagnosticity and stability, the top three highest ranked methods were partial least squares, Shapley Value and Spearman correlation.

To recommend the best used method, Pokryshevskaya and Antipov (2014) carried out further analysis using Monte-Carlo iterations to vary weights of the two metrics from 0.4 to 0.6 and observed the percentage of being ranked in the top three for each method. Shapley value yielded the highest percentage among the methods thereby Pokryshevskaya and Antipov (2014) supported the use of Shapley value as a method for deriving *importance*.

Regarding the summary of previous comparative studies presented in Table 3.2, a number of times that each evaluation was used in the past comparative studies is shown in Table 3.3. Based on this table, evaluation metrics with the total number of studies greater than or equal to 2 were identified as the common evaluation metrics which were further observed in order to form the set of evaluation metrics of this study. Accordingly, 5 out of 8 evaluation metrics namely predictive validity, relative predictive validity, diagnosticity, an incidence of negative measures, and discriminative power were observed to decide whether or not they are suitable evaluation metrics in the context of the comparative study conducted in this thesis.

Consider the context of the comparative study of this thesis, two evaluation metrics namely relative predictive validity and incidence of negative measures were excluded from the set of evaluation metrics. The remaining was formed as the set of evaluation metrics of this study and their detail is provided in Section 5.2. The relative predictive validity was excluded from the set of evaluation metrics since the measurement of this metrics required the customer preference rank-order which is not contained in the datasets used in the comparative study of this thesis. The incidence of negative measures was excluded from the set of evaluation metrics since considering this metric provide an unfair comparison between regression-based techniques (such as MLR, OLR) that prone to produce negative measures with other techniques (such as direct-rating scales, BPNN, Naïve Bayes and BNs) that constantly produce positive measures.

Table 3.2: Summary of previous comparative studies

Publication	Type of method for measuring *importance*	Technique for measuring *importance*	Evaluation metrics
Hauser (1991)	Self-stated and Implicitly derived importance	Direct-rating scales Anchored scales Constant-sum scales Linear regression	Predictive validity Relative predictive validity Face validity
Bacon (2003)	Self-stated and Implicitly derived importance	Direct-rating scales Correlation Linear regression	Relative predictive validity
Gustafsson & Johnson (2004)	Self-stated and Implicitly derived importance	Direct-rating scales Linear regression Normalized pairwise estimation Reflective partial least squares Formative partial least squares Principal components regression	Predictive validity Diagnosticity Incidence of negative measures
Taplin (2012)	Self-stated and Implicitly derived importance	Direct-rating scales Relative direct-rating scales Linear regression	Predictive validity
Cohen (2003)	Self-stated importance	Direct-rating scales Paired comparisons MaxDiff	Relative predictive validity Discriminating power
Chrzan & Golovashkina (2006)	Self-stated importance	Direct-rating scales MaxDiff Magnitude estimation Unbounded ratings Constant sum Q-sort	Predictive validity Discriminating power Task length
Pokryshevskaya & Antipov (2014)	Implicitly derived importance	Pearson correlation Kendall correlation Spearman correlation Linear regression Partial least squares Multilayer perceptron Shapley value decomposition Fields decomposition First-round effect	Stability Diagnosticity Incidence of negative measures

Table 3.3: Evaluation metrics and their number of times used in the past comparative studies

Evaluation metrics	Description	Total number of studies
Predictive validity	an ability to predict actual overall customer satisfaction	4
Relative predictive validity	an ability to predict customer-reported priority	3
Diagnosticity	an ability to identify the consumers' most important attributes	2
Incidence of negative measures	a number of the negative coefficients for the regression-based method	2
Discriminating power	an ability to provide discriminating measures	2
Face validity	a subjective assessment of whether or not the technique can measure *importance*	1
Task length	an average time to complete task	1
Stability	a robustness of importance estimation to a small change of data	1

3.3 Methodology for deriving *importance*

In this section, methods for deriving implicit *importance* by using statistical and data mining techniques such as MLR, BPNN, Naïve Bayes and BNs are described. The first two techniques have been used to derive the *importance* from the survey data (Deng et al., 2008a; Pezeshki et al., 2009; Ho et al., 2012; Mikulić and Prebežac, 2012) while the last two techniques have never before been applied to derive the *importance*.

The main tasks for deriving *importance* are exploring the relationship between the performance rating of each attribute of company's product or service and the overall satisfaction, and estimating an impact of each attribute that has an effect on the overall customer satisfaction. A set of steps for deriving *importance* of each data mining technique are described in the following sub-sections.

3.3.1 Method for deriving *importance* based on MLR

Since MLR is an easy-to-use technique, the procedure for deriving *importance* based on MLR is simple and straightforward. First, the performance rating of each attribute and the overall satisfaction are defined as the independent and dependent variables respectively. Next, the regression model is trained and produces the regression coefficients.

MLR model is easy to interpret by considering the signs and magnitude of coefficients. The sign of coefficient represents the existence of negative or positive effects of the independent variables on the dependent variable. The magnitude of coefficient generally indicates how much a one unit increase in the independent variable results in a change of the dependent variable with all other variables held constant (Nathans et al., 2012). Therefore, the regression coefficients can be referred to as implicit *importance* which express the influence of attributes of company's product or service on the overall satisfaction.

3.3.2 Method for deriving *importance* based on OLR

Similar to MLR, the procedure for deriving *importance* based on OLR is simple and straightforward. First, the ordinal logistic model is build by specifying the performance rating of each attribute and the overall satisfaction are defined as the independent and dependent variables respectively. Additionally, the link function that would be the best fit to the data set is chosen. Two major link functions commonly used in the OLR model are logit and cloglog links(Chen and Hughes, 2004). Next, the regression coefficients are obtained from the trained model.

The coefficients obtained from the OLR model can also be referred to as implicit *importance* as with other types of regression. Specifically, the coefficient indicates that a one unit change on independent variable result in a change of the odds of dependent variable, holding other independent variables as constant(Chen and Hughes, 2004).

3.3.3 Method for deriving *importance* based on BPNN

In contrast to the result of MLR which can be referred to as implicit *importance*, the result of BPNN cannot be used directly as the *importance* hence there are two steps for deriving *importance* by using BPNN. The first step is the construction of a BPNN model in order to discover attributes of a company's product or service that has the major influence on overall satisfaction and selecting the model that produces the best performance. The second step is the computing of *importance* from the weights of the neural network. The contribution of each input to the outputs can be estimated through the connection weights since they are the links between the inputs and outputs. The magnitude of connection weights indicates the effect of each input on the predictive outputs where the input with a larger connection weight has more contributions to the output, and therefore is more important than the input with a smaller one (Olden and Jackson, 2002).

Step 1: Construct the BPNN prediction model for overall satisfaction

The BPNN model is modelled as one input layer, one hidden layer, and one output layer. The satisfaction ratings of each attribute of company's product or service are the neurons in the input layer and overall customer satisfaction is the only neuron in the output layer. For the hidden layer, the number of neurons is assigned by training the network with different configurations of hidden-layer neurons and selecting the best performing network.

After the BPNN architecture is specified, then the training parameters of the neural network such as activation function, learning rate and momentum are determined. Subsequently, several configurations of hidden-layer neurons are trained on the training dataset and the performance measured by three indicators: the mean absolute error (MAE), the root mean squared error (RMSE) and goodness of fit (R^2) in order to determine the number of neurons in the hidden layer. Note that, if the MAE and RMSE approach 0 this indicates that BPNN model has precise prediction ability while if R^2 is close to 1 this indicates that BPNN model has excellent goodness-of-fit (Deng et al., 2008a). Therefore the network configuration (input-hidden-output) that provides lowest MAE and RMSE, and highest R^2 is chosen as the final BPNN architecture.

Step 2: Acquiring the implicitly derived importance from BPNN's weights

After the BPNN model is trained as described in the previous step, the relative importance of each company's service attribute in terms of producing overall satisfaction are derived by applying Garson's algorithm proposed by Garson (1991). This algorithm is one of the two most widely used algorithms for calculating the input-hidden-output connection weight of BPNN (Mikulić et al., 2012) in order to determine the relative importance of each input variable in the neural network. Later, Glorfeld (1996) extended the procedure of Garson to deal with multi-neurons in the output layer.

Let I be the number of input-layer neurons, H be the number of hidden-layer neurons and, O be the set of output-layer neurons. Glorfeld (1996)'s steps for obtaining relative importance of each input variable in the neural network are described as follows:

- Computation of the input-hidden layer connection weight. For each hidden neuron, the proportion of each input-hidden connection weight (P_{ih}) is computed by dividing the absolute value of the input-hidden layer connection weight (W_{ih}) by the sum of the absolute value of the input-hidden layer connection weight of all input neurons, as shown in Equation 3.1.

$$P_{ih} = \frac{|W_{ih}|}{\sum_{i \in I} |W_{ih}|} \quad \forall i \in I, h \in H. \tag{3.1}$$

- Computation of the hidden-output layer connection weight. The total output weight for each hidden neuron (SO_h) is computed as the sum of the absolute value of the hidden-output layer connection weight (W_{oh}) of all output neurons, as shown in Equation 3.2.

$$SO_h = \sum_{o \in O} |W_{oh}| \quad \forall h \in H. \tag{3.2}$$

- Computation of input-hidden-output connection weight. The contribution of each neuron to the output via each hidden neuron (C_{ih}) is computed by multiplying the proportion of each input-hidden connection weight (P_{ih}) produced from Equation 3.1 by total output weight of the corresponding hidden neuron (SO_h) produced from Equation 3.2. The equation is shown below.

$$C_{ih} = P_{ih} \times SO_h \quad \forall i \in I, h \in H. \tag{3.3}$$

- Computation of the contribution for each input variable. For each input variable, the contribution (S_i) is computed as the sum of the contribution of each neuron (C_{ih}) produced from Equation 3.3 across all hidden neurons, as shown in Equation 3.4.

$$S_i = \sum_{h \in H} C_{ih} \quad \forall i \in I. \tag{3.4}$$

- Computation of relative importance for each input variable. After the sum of each input neurons contribution is computed, the relative importance (RI_i) is calculated following to the Equation 3.5. This equation formats the relative importance as a percentage (%) as the ratio of S_i to the overall contribution of all input neurons.

$$RI_i = \frac{S_i}{\sum_{i \in I} S_i} \times 100 \quad \forall i \in I. \tag{3.5}$$

In summary a method for deriving *importance* based on BPNN is the assessments of company' s service attributes are specified as input neurons and the overall satisfaction is specified as output neurons. BPNN with one hidden layer are trained and tested to establish the relationship between these service aspects and the overall satisfaction. In addition, a number of neurons in the hidden layer is assigned by experimental trial. After the BPNN model is trained, the *importance* can be calculated based on the extended Garson's algorithm proposed by Glorfeld (1996).

3.3.4 Method for deriving *importance* based on Naïve Bayes

There are two steps for conducting Naïve Bayes based IPA. The first step is the construction of a Naïve Bayes model in order to discover the attributes of company's product or service that have the major influence on overall customer satisfaction. The second step is the computing of *importance* from the conditional probability.

Step 1: Construct the Naïve Bayes prediction model for overall customer satisfaction

Similar to other techniques, the performance rating of each company's attribute and the overall satisfaction are defined as the independent and dependent variables for constructing the Naïve Bayes model. Note that, both independent and dependent variables must be discrete and mutually exclusive. As Naïve Bayes is a simple technique, its model can be constructed easily without a training parameter specified. Generally, the Naïve Bayes model contains information about probability of dependent (class) variable and the conditional probability distribution of each independent variable.

Step 2: Acquiring the importance of attributes from the conditional probability distribution

Given the conditional probability distribution table obtained from Naïve Bayes prediction model, the attribute importance can be computed based on the basis of Mutual Information (MI) which is measure of dependence between two random variables introduced by Shannon (1948). In particular, it quantifies the amount of information about one random variable given knowledge of another (Cover and Thomas, 2006).

In this study, the MI between performance of each company's attribute (X_i) and overall customer satisfaction (Y) is estimated as the reduction in uncertainty (or entropy) about X given knowledge of Y, see Equation 3.6 (Cover and Thomas, 2006).

$$MI(X_i, Y) = h(X_i) - h(X_i|Y) \tag{3.6}$$

where X_i represents satisfaction (performance) of each attribute of company's service, Y represents overall satisfaction, $h(X_i)$ is the entropy of X_i, and $h(X_i|Y)$ is the conditional entropy (the entropy of X_i conditioned on Y). The two entropies within Equation 3.6 can be calculated following to the Equation 3.7, 3.8 respectively.

The entropy of each attribute $h(X_i)$ is computed as Equation 3.7 (Shannon, 1948; Cover and Thomas, 2006).

$$h(X_i) = -\sum_{a=1}^{n} P(X_i = a)\, log_2 P(X_i = a) \tag{3.7}$$

where $P(X_i = a)$ is the probability of each attribute of company's service X_i with the satisfaction level corresponding to the n Likert scale ($a = 1, 2, \ldots, n$). Specifically, $P(X_i = a)$ is calculated from the conditional probability distribution table by dividing number of records of the level a for attribute X_i to the number of all records.

The entropy of X_i conditioned on Y, $h(X_i|Y)$ is computed as Equation 3.8 (Cover and Thomas, 2006).

$$h(X_i|Y) = \sum_{b=1}^{n} P(Y = b)\, h(X_i|Y = b) \tag{3.8}$$

where $P(Y = b)$ is the probability of the overall customer satisfaction Y with the satisfaction level corresponding to the n Likert scale ($b = 1, 2, \ldots, n$) calculated from the conditional probability distribution table by dividing number of records of the level b for overall customer satisfaction Y to the number of all records. And $h(X_i|Y = b)$ is the average conditional entropy of X_i computed from Equation 3.9.

$$h(X_i|Y = b) = -\sum_{a=1}^{n} P(X_i = a|Y = b)\, log_2 P(X_i = a|Y = b) \tag{3.9}$$

where $P(X_i = a|Y = b)$ is the distribution of $X_i = a$ condition on $Y = b$. $P(X_i = a|Y = b)$ is calculated from the conditional probability distribution table generated in previous step by dividing number of records of the level a for attribute X_i to the number of records of the level b for overall customer satisfaction Y.

In short, *importance* of each company's attribute is computed as MI by calculating the difference between attribute entropy $h(X_i)$ and conditional entropy $h(X_i|Y)$. MI is non-negative quantity in which high MI indicates strong correlation between two variables, whereas MI value of zero indicates uncorrelated variables (Shieh and Wu, 2011). Hence, it can be interpreted that the attribute with the highest MI is the most important.

3.3.5 Method for deriving *importance* based on BNs

As Naïve Bayes is a special case of BNs (Grossman and Domingos, 2004) in which each attribute has the class as its sole parent. Steps for conducting Bayesian Networks based IPA is similar to the steps for conducting Naïve Bayes based IPA which consist of predictive model construction, and measuring of *importance* based on conditional probability. These steps are described in the following sub-sections.

Step 1: Construct the BNs prediction model for overall customer satisfaction

Similar to Naïve Bayes, the performance ratings of each attribute of company's product or service are assigned as independent variables and the overall customer satisfaction is assigned as the class variable for constructing BNs model. The values of independent variables and class variable are discrete and mutually exclusive. In contrast with Naïve Bayes, several parameters have to be defined in order to construct BNs prediction model on the training data such as an estimator and search algorithm.

Step 2: Acquiring the importance of attributes from conditional probability distribution

Steps for acquiring the *importance* from the Bayesian model are similar to steps for acquiring the *importance* from the Naïve Bayes model described in Step 2 of section 3.3.4 which involved the calculation of MI based on the conditional probability distribution table obtained from BNs prediction model.

Note that the method for calculating the entropy of X_i, $h(X_i)$ and the conditional entropy $h(X_i|Y)$ from BNs is slightly more complicated than the method for calculating these two entropies from Naïve Bayes, as the network structure of BNs is more complicated than the network structure of Naïve Bayes. In detail, $P(X_i = a)$ and $P(X_i|Y = b)$ which are the main components for calculating entropy of X_i and the conditional entropy (see Equation 3.7 and 3.8), are calculated from conditional probability distribution through the marginalization process.

3.4 Summary

Three main topics regarding the IPA were explained in this chapter. Firstly, the overview of IPA revealed that IPA is a tool for understanding customer's needs and desires so that the company can manage resources to achieve the high level of customer's satisfaction. In a nutshell, IPA measures the satisfaction based on two components: *performance* and *importance*. The intersection of these two components creates a two-dimensional matrix that supports a company in discovering the attributes of a company's product or service that should be maintained or improved based on the customers' viewpoint.

Additionally, the overview of IPA revealed that a method for measuring *performance* is well-established by using direct rating from customers whereas there are many approaches for indirectly measuring *importance* using statistical methods and data mining since the customers' self-statements tend to produce company's attribute importance that are not sensible. The main idea of the indirect method for measuring importance is exploring customer survey data to discover the relationship between attribute's performance and the overall customer satisfaction. Regarding the literature survey, statistical methods that have been used for measuring *importance* are MLR, OLR, SEMs and partial correlation and BPNN, which is a predictive data mining technique, have recently become an alternative approach for determining a company's attribute importance.

Secondly, the past comparative studies of methods for measuring *importance* showed that there were a few empirical studies that compared the different methods against a set of evaluation criteria. And the results of the past comparative studies were not along the same lines, some studies supported the use of self-stated importance while some studies supported the use of implicitly derived importance. The review of the past comparative studies of methods for measuring *importance* also provided a background related to the evaluation metrics for setting-up the comparative experiment in Chapter 5.

Thirdly, methodologies for deriving *importance* were described to illustrate how the *importance* can be derived from results of two statistical techniques- MLR, OLR and three data mining techniques namely - BPNN, Naïve Bayes and BNs. MLR, OLR and BPNN have been used to derive the *importance* from the survey data while Naïve Bayes and BNs have never before been applied to derive the *importance*. For each technique, the main tasks for deriving *importance* are constructing the predictive model that discovers the relationship between the performance rating of each attribute and the overall customer satisfaction, and acquiring *importance* from an outcome of the predictive model. Specifically, methods for acquiring *importance* from an outcome of each technique are different since each technique produces a different form of outcome, for example; MLR and OLR produce regression coefficients, BPNN produces neural network weights, Naïve Bayes and BNs produce a conditional probability table.

Different importance measurement techniques will likely result in different IPA matrices. Therefore, the technique for measuring *importance* has to be carefully justified which leads to the empirical experiment set up in Chapter 5. Additionally, research process and research methodology that were conducted to complete this study are explained in the next chapter.

Chapter 4

Research Design

In the preceding chapters, the research goal and research questions as well as background knowledge of this thesis were described. Before proceeding to the chapter that presents the research findings, it is important to describe a procedural plan for how this research study is to be completed. Such a procedural plan is known as a research design and a well-defined research design is essential in providing valid research findings (Kumar, 2011).

Thus, the main function of this chapter is explaining a procedural plan which consists of the main tasks to be conducted throughout this study, and research methods to be used corresponding to each task. Section 4.1 provides an overview of the research design of this study. Section 4.2 - 4.5 provide detail for each phase of research design.

4.1 An overview of research design within this study

Among the three Information Technology (IT)-related disciplines which are Information Systems (IS), Computer Science and Computer Systems Engineering (Avison and Elliot, 2006), this research can be classified as IS research, an applied discipline, which emphasises the applications of technology to address the range of operational, managerial and strategic activities of organisations as this work applied knowledge in computer science to solve problems of organisations in conducting their business.

Consequently, the research work presented in this thesis was conducted based on a design science research process in IS. According to Kuechler and Vaishnavi (2005), design science research is a set of analytical techniques and perspectives for performing research in IS. It involves the design and evaluation of novel or innovative artefacts in order to solve observed problems of organisations (Hevner and Chatterjee, 2010).

A sequence of activities or steps in the design science research process can be structured in three main phases: "problem identification", "solution design", and "evaluation"(Offermann et al., 2009). For each phase, different researchers in IS defined a set of different steps, a comparison of design processes can be referred in Peffers et al. (2007); Offermann et al. (2009).

Through the review of a number of research articles in IS that contributed ideas for design processes, the research work presented in this thesis was conducted based on a design science research process suggested by Peffers et al. (2007) as it is one of the most cited studies in design science research and it provided well-defined activities to be applied in order to complete this research work.

According to Peffers et al. (2007), design science research process comprises of sequence of six activities: (1) problem identification and motivation, (2) definition of the objectives of a solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication. A detail of these activities is summarized in Table 4.1.

Using a design science research process in the context of this thesis is presented in Figure 4.1. In the first phase of the research process, the problem and motivation regarding traditional SWOT analysis was raised from an organisation's needs. Then the objective of a solution was defined as a new approach to conduct SWOT that is able to provide quantitative and customer-oriented SWOT factors.

In the solution design phase, an artefact in the form of methods, sets of steps used to perform tasks (Kuechler and Vaishnavi, 2005), was designed and developed to solve the problem defined in the first phase. The result of this phase is an approach named IPA based SWOT analysis. The main idea of this approach is to apply IPA for conducting SWOT analysis which produces SWOT factors based on results of customer satisfaction surveys and quantifies them based on two mains aspects of IPA which are *importance* and *performance*. The development of this approach consists of two main tasks namely IPA and IPA-SWOT. The former involves the selection of current statistical and data mining techniques used for implicitly derived importance. The latter involves the SWOT identification based on IPA in which the *importance* was derived by using the selected technique from the previous task.

In the evaluation phase, three surveys namely student satisfaction survey, staff evaluation of SWOT survey, and experienced users of SWOT survey were designed and conducted, as part of the case study of HEI in Thailand, following the survey research method. The results of these surveys were then analysed to demonstrate the application of IPA based SWOT and to evaluate the efficiency of IPA based SWOT .

Additional to the activities within the research processes presented in Figure 4.1, there are multiple research methods associated with the activities. The research methods that

Table 4.1: A summary of design science research activities presented in Peffers et al. (2007)

Phases	Design science research activities	Activity description	Output
Problem identification	Problem identification and motivation	Define and refine the research problem to assure its relevance and understanding	Research questions
	Definition of the objectives of a solution	Identify specific criteria that a solution for the problem should be met	
Solution design	Design and development	Develop a solution for the problem in the form of an artefact: constructs, models, methods, or instantiations (Hevner and Chatterjee, 2010)	Artefact
Evaluation	Demonstration	Use the developed artefact to solved problem in simulation, a case study, experimentation, or other appropriate activity to demonstrate its efficacy	
	Evaluation	Determining and measure how well the artefact performs by comparing the observed results with criteria or metrics	Evaluation metrics, Results
-	Communication	Report results, contributions, limitations, and new knowledge gained during the design and construction of artefact to research communities	Publications, PhD Thesis

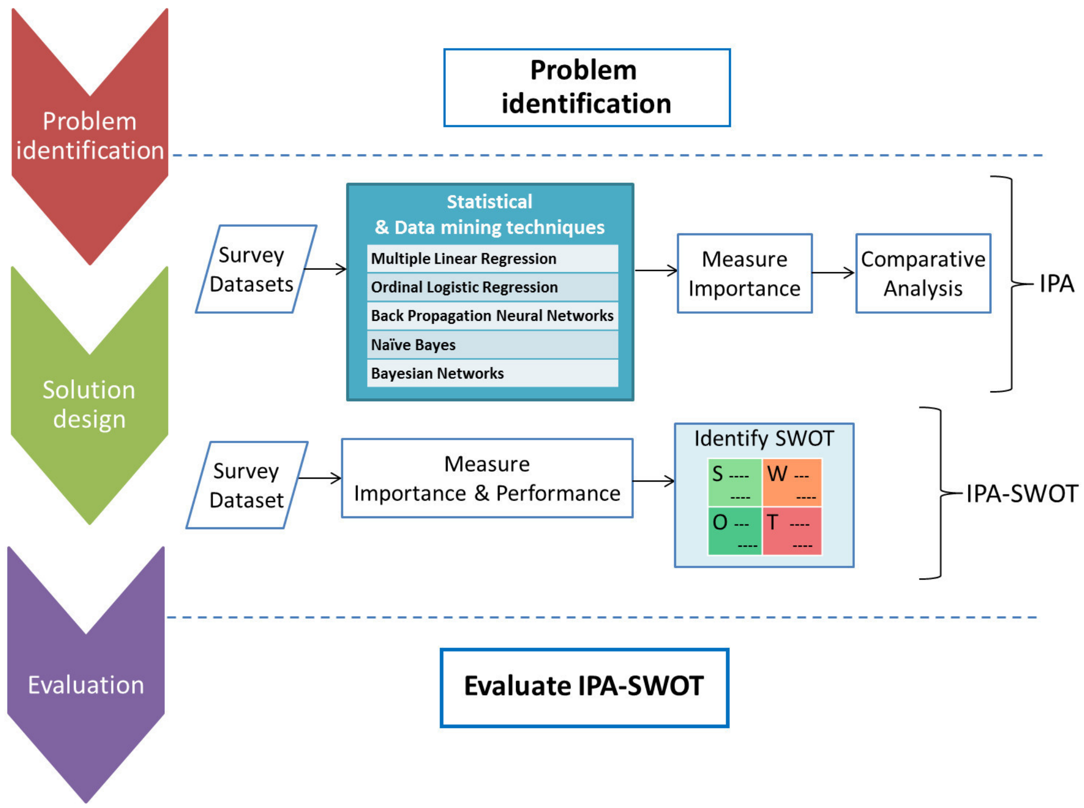


Figure 4.1: Research design scenario

were used to perform these activities were selected from a study on the methodologies applied in and accepted by seven leading MIS journals during 1993-2003 by Palvia et al. (2004). Specifically, the four selected research methods include library research, frameworks & conceptual models, survey and laboratory experiment. Their details are provided in Table 4.2.

Table 4.2: MIS Methodologies retained for this research based on Palvia et al. (2004)

Method	Definition
Library research	Research that is based mainly on review, summarize and synthesize existing literature.
Frameworks & conceptual models	Research that intends to develop a framework or a conceptual model
Survey	Research that uses predefined and structured questions to capture data from individuals
Laboratory experiment	Research in a simulated laboratory environment that manipulates and controls the various experimental variables and subjects

Library research is the first point of departure for a research journey as it establishes the need for the research. Library research also serves many important purposes related different activities in research process (Offermann et al., 2009). For example, the activities in problem identification phase employed library research to identify a problem as well as gain insight and understand the identified problem. The activity in the solution design phase employed library research to provide information relating to a possible solution to solve the problem. In this thesis, the library research was employed for accomplishing research activities in the first two phases.

According to Håkansson (2013), frameworks and conceptual models method is used for developing new concepts or interpreting existing concepts that are built upon the result of library research. In this thesis, conceptual research is represented in the form of the methodological framework for generating SWOT or artefact in the form of method in the context of design science research process.

Survey research involves the selection of a representative sample from a well-defined population and the use of data analysis techniques to generalize from that sample to the population (Pinsonneault and Kraemer, 1993). Normally, data for the survey research is collected from individuals by means of structured and predefined questions. The survey research was the most widely used research methodology by researchers in MIS research during 1993-2003 (Palvia et al., 2004) and is still in predominant use.

The survey research strengths lay in its ability of obtaining information from a large sample of the population and its ability to elicit information about attitude that is difficult to acquire using observational techniques (Glasow, 2005). The survey research weaknesses lay in its subjective nature in which bias in respondents or researcher may occur, the lack of response from respondents, and intentional misreporting by respondents (Glasow, 2005). In this thesis, survey research was employed for the purpose of evaluation of the method for generating SWOT.

Laboratory experiment involves examination of a phenomenon in a controlled setting. It is conducted to investigate how the variables in the study are related by manipulating one or more independent variables and measuring their impact on one or more outcome (dependent) variables (Easterbrook et al., 2008). In the case of this thesis, laboratory experiment was carried out to investigate the link between the type of statistical and data mining techniques and their performance on measuring *importance*. The degree of change in performance in different groups of techniques is then compared to identify the suitable technique for measuring *importance*.

One or more research methods may be employed to perform each activity within the research process. The connection between the research process, research methods, research questions, and thesis chapters is shown in Table 4.3 and details of each phase in the research process and related research methods are described in the following subsections.

Table 4.3: Connection between the design science research process, research methods, research questions, and thesis chapters

Phase of design science research process	Task	Research methods	Research questions	Thesis chapter
Problem identification	Problem Identification	Library research	-	2,3
Solution design	IPA	Library research Laboratory Experiment (Comparative experiment)	RQ 1	5,6
	IPA-SWOT	Library research Conceptual Framework	RQ 2	7
Evaluation	Demonstrate and Evaluate IPA-SWOT	Survey	RQ 3	8,9

4.2 Problem identification

Awareness of the problem began due to the author's recognition in an organisation context that the traditional way to conduct SWOT analysis through a brainstorming yields SWOT outcome from a personal attitude with generally no support evidence. Using this SWOT outcome to generate strategies would be improper. Then a library research was conducted with an aim of better understanding of SWOT analysis in which the main focus topics are: limitations of traditional SWOT analysis, and trends and application areas of SWOT analysis.

The starting points of the literature search were two literature review papers by Helms and Nixon (2010); Ghazinoory et al. (2011). The first review paper by Helms and Nixon (2010) presents the literature review of SWOT analysis from June, 1999 to June, 2009. The second review paper by Ghazinoory et al. (2011) presents a review of SWOT analysis related papers published in indexed journals from 1982 to the end of 2009. These two review papers pointed out the limitations of traditional SWOT analysis and as well as trends and application areas of SWOT analysis which provided evidence to support the need of a new approach to conduct SWOT analysis and evidence to show that SWOT analysis is still widely used as a tool for strategic planning and hence confirmed the significance of the topic.

A further literature search was conducted by following relevant references cited in the first two papers, to gain deeper insight of the problem of traditional SWOT analysis and existing approaches that try to improve the soundness of SWOT analysis. It was discovered that there are two main streams of SWOT approach: quantitative SWOT analysis (Kurttila et al., 2000; Kangas et al., 2001; Lee and Lin, 2008) and customer-oriented SWOT analysis (Dai et al., 2011; Pai et al., 2013). The former involves quantifying SWOT factors which can be prioritized. The latter involves extraction and analysis of customers' feedback as a source to create SWOT analysis which ensures that the capabilities perceived by a company are recognized and valued by the customers.

The literature search through these two approaches revealed that each approach focuses on improving a different aspect of SWOT analysis. While the first approach can make SWOT factors commensurable, these SWOT factors focus solely on a company's perspective and ignore the customer's perspective. On the other hand, the second approach generates SWOT factors from a customer's perspective but it has no means to prioritize SWOT factors. This can be identified as a gap between the two current approaches of SWOT analysis which will be filled by developing this research work. After the problem was identified as previously stated, the objective of a solution to the problem was then defined as a new approach to conduct SWOT that is able to provide quantitative and customer-oriented SWOT factors.

4.3 Solution design: IPA

Based on the results of the problem identification phase, a method was chosen as the design artefacts and customer satisfaction survey was chosen as a source of data for conducting SWOT analysis. Library research was also conducted in this phase to find a technique that can be employed to analyse customer satisfaction and develop SWOT analysis based upon its result. For this purpose, several publications related to customer satisfaction analysis were retrieved by querying scientific search engines such as the IEEE Explore and ACM Digital Library, and then following relevant references cited in the discovered papers. The review of these papers led to the decision of applying IPA to develop a new approach of SWOT analysis.

Consequently, a further library research was conducted to review the state-of-the-art of IPA starting with highly cited papers such as Martilla and James (1977); Matzler et al. (2003); Bacon (2003) then the literature search was gradually fanned out to other relevant papers. Through this process, one technical issue of IPA regarding an indirect method for measuring *importance* (Lai and Hitchcock, 2015) was chosen to be studied within this research work since, currently there are no well-established methods for measuring *importance* (Gustafsson and Johnson, 2004; Fontenot et al., 2007; Feng et al., 2014).

Focusing on the issue of techniques for measuring indirect importance, the experiment was set to compare three currently used indirect measurement techniques namely MLR, OLR, BPNN with the two new introduced techniques for measuring indirect importance namely Naïve Bayes, and BNs. BNs was introduced for deriving importance since it has been applied to analyse the customer satisfaction surveys (Salini and Kenett, 2009) which closely related to importance measure in IPA. Naïve Bayes was included in the comparative experiment as it is a special case of BNs (Grossman and Domingos, 2004).

In addition to the indirect importance measurement techniques to be compared in the experiment, the evaluation metrics used to decide the appropriate techniques for measuring *importance* were also defined through the review of previous comparative studies i.e. Hauser (1991); Bacon (2003); Gustafsson and Johnson (2004). Specifically, the three evaluation metrics were selected from eight evaluation metrics based on the number of times that it has been used in previous comparative studies.

Three datasets of customer satisfaction survey were used in this experiment. The first two datasets were published datasets which have been used by other researchers (Kenett and Salini, 2012, p. 19) and (Azzalini et al., 2012). The third dataset was a student satisfaction survey collected from a leading Thai University for demonstrating the use of the design artefact.

To ensure that the comparative experiment were properly conducted, two baseline experiments were conducted before starting the comparative experiment. Further details regarding the baseline experiments and comparative experiment set up, and the experimental results are presented in Chapter 5 and Chapter 6 respectively.

4.4 Solution design: IPA-SWOT

Based on the result of solution design phase for IPA, a technique for measuring *importance* was identified and applied in a methodological framework for developing SWOT analysis from IPA results, called IPA based SWOT analysis.

Before starting to design the methodological framework, the library research was conducted to review the connection between IPA and SWOT analysis and specify the possible way to acquire all SWOT aspects from the IPA results, known as IPA matrix. The review of literature indicated the correspondence between IPA and SWOT analysis since each quadrant of the IPA matrix can be interpreted as major/minor strength and weakness (Garver, 2003; Deng et al., 2008a; Silva and Fernandes, 2012; Hasoloan et al., 2012; Cugnata and Salini, 2013; Hosseini and Bideh, 2013).

Since, only strength and weakness can be inferred from the IPA matrix. Further library research was conducted to identify a possible solution to infer opportunity and threat from the IPA matrix which provided a complete aspect of SWOT analysis. This led to a decision to apply idea of Pai et al. (2013) to identify opportunity and threat based on the IPA matrix of a company's competitor as it is a sensible idea and can be practically applied in the context of this research. At the end of this phase, the methodological framework to develop SWOT analysis from IPA was designed, the detail is discussed in Chapter 7.

4.5 Evaluation through the survey research

After IPA based SWOT analysis was developed, it was executed in the context of a University in Thailand to demonstrate and evaluate its efficacy. Three surveys were developed for this purpose which are student satisfaction survey, staff evaluation of SWOT survey, and experienced users of SWOT survey. The first survey was developed to collect student satisfaction and then used as the data source to implement the IPA based SWOT analysis in a real-world context. The last two surveys were developed to evaluate the efficacy of IPA based SWOT analysis after implementing on data of the first survey, using quantitative measure since, currently there are no direct methods and tools for validating the effectiveness of SWOT analysis (Ayub et al., 2013).

Specifically, the staff evaluation of the IPA based SWOT analysis, was conducted using a survey to collect the staff's agreement on the analysis; to check whether the staff agree or disagree with the resulting SWOT analysis. The experienced users of SWOT survey involves the collection of MBA students' rating on the quality of outcome of IPA based SWOT analysis. In this section, the main aspects corresponding to the survey design are provided for which further details including pilot study, sample size and ethic approval related to each survey are provided in Chapter 8.

The main aspects of each survey including group of participant, survey media, and question type, contact method, and administration method are summarized in Table 4.4. And the detail of each aspect regarding possible methods that could have been used and the justification for the selection of specific method is provided in the following sub-sections.

Table 4.4: Detail of three surveys used for demonstration and evaluation of IPA based SWOT analysis

	Student satisfaction survey	**Staff evaluation of SWOT survey**	**Experienced users of SWOT survey**
Purpose:	Demonstration	Evaluation	Evaluation
Group of participant:	Undergraduate students	University staff	MBA students
Survey Media:	Written Survey	Written Survey	Written Survey
Survey Question type:	Closed-ended questions	Closed-ended questions	Closed-ended and open-ended questions
Contact method:	Personal contact (via class room)	E-mail	Personal contact (via class room)
Administration method:	Internet-mediated	Internet-mediated	Delivery and collection

4.5.1 Selecting survey media

Three types of survey media which are written survey, verbal survey and mixed mode survey (Glasow, 2005) were investigated to choose one that fitted-well with the context of this study. According to Glasow (2005), written survey is known as questionnaire and verbal survey is interview. Mixed mode survey is a combination of the written and verbal survey.

The written survey was chosen as the media to collect data for the three surveys for the reason that the data will be analysed by means of quantitative research using statistical and data mining techniques. While the use of a verbal survey allows a researcher to get more qualitative data, in particular, it is difficult to summarize and incorporate statistical analysis (Saunders et al., 2009).

Additionally, face-to-face interview is not applicable to use as a survey media of this study since the responses to the student satisfaction survey need to be anonymous to avoid intentional misreporting by respondents so as to the response to the staff evaluation of SWOT survey. While the telephone interview can keep the response anonymous, it takes time and cost to schedule and to establish a call for participants of this study.

4.5.2 Selecting survey questions

There are two types of survey questions: open-ended and closed-ended questions. The former allows respondents to formulate their own answers which enable the researcher to explore ideas of respondents while the latter forces respondents to choose an answer from a given set of responses (Glasow, 2005; Fink, 2012b).

Generally, the closed-ended questions require minimal thought and writing hence it is usually quicker and easier to answer than the open-ended questions which require greater thought to respond, and therefore, take a longer time to answer (Saunders et al., 2009). The results obtained from open-ended questions are also more difficult to compare and analyse (Glasow, 2005).

The closed-ended questions were used to collect quantitative data regarding student satisfaction as the specification of input data for IPA. Specifically, a survey that is suitable for applying IPA contains assessment of respondents satisfaction for an organisation's product or service measured by a Likert scale with either five or seven levels.

The closed-ended questions were also used to collect quantitative data of the second and third survey as the main purpose for data collection of these two surveys was to provide a quantitative measure to evaluate the IPA based SWOT analysis. In addition to the closed-ended questions, one open-ended question was used to ask the respondents of the third survey about their suggestion toward the IPA based SWOT analysis. This allowed respondents of the third survey who had experienced in generating SWOT give some useful ideas regarding IPA based SWOT analysis.

4.5.3 Selecting contact method

Possible ways to contact the respondents to take part in each survey of this study were investigated. The three possible contact methods are postal/e-mail, telephone, and personal contact. By considering contact methods that are the most convenient way to contact different group of respondents, e-mail and personal contact were chosen to contact the respondents to take part in this study.

Respondents of the first and the third survey who are undergraduate and graduate students (see Table 4.4) respectively were asked face to face to take part in the survey by the representative of the researcher at the end of their class. This guarantees that the majority of respondents were invited to take part in the survey which may increase the response rate, and allows the representative of the researcher to explain the objective of the survey.

To avoid the potential error that may occur during survey process since the researcher cannot physically be present to contact respondents and administer the survey by herself because of the limitation of place and time, the representative of the researcher had been already informed about the procedure to administer the survey and discussed the objective of the survey with the researcher prior to the survey administration.

Respondents of the second survey who are university staff including academic and non-academic staff were asked to take part in the survey by e-mail. This is a formal way to contact university staff and an easy way to follow-up their response. Besides, using e-mail is also convenient for respondents to answer when they are available.

Among the three possible contact methods, telephone was not chosen to be used although it is a quick method. This is because it involves a call cost and it is difficult to find phone number of all respondents especially with the large numbers of respondents. Besides, it is less formal to contact respondents who are university staff via telephone than e-mail.

4.5.4 Selecting administration method

The range of methods for collecting survey data can be classified into two major groups according to how a survey is administered and these are self-administered and interviewer-administered method. The written survey is collected through the self-administered method since the written survey is usually completed by respondents themselves whereas verbal survey is collected through the interviewer-administered method since it is usually carried out by interviewer who asks and records answer of respondents (Saunders et al., 2009).

Since only written survey is used in this study, only the self-administered method is considered as a major type for administering survey. The self-administered method is divided into three types: a delivery and collection, an internet-mediated, and postal/ e-mail. Main aspects of each method are compared in Table 4.5.

Through the investigation of characteristics for each mode of data collection, postal/e-mail was not a choice of administering survey within this study for the reason that this method takes the longest data collection period and it is practically difficult to obtain e-mail or address of all respondents. Therefore, two self-administered methods which are delivery and collection and internet-mediated were selected to be used in this study.

Specifically, data for the first and second survey were collected through online surveys developed by using iSurvey, a free to use survey generation tool for members of the University of Southampton. By using internet-mediated as the method of data collection, the anonymity of respondents can be assured and respondents can take their time to answer the questions on their own schedule. The anonymity and the time to consider questions were important factors if respondents were to provide honest and accurate data, and thereby improve the quality of responses (De Vaus, 2002a). Besides, respondents' answers are automatically stored in an electronic file ready for analysis.

Data for the third survey were collected using a delivery and collection method whereby each questionnaire was delivered and collected in the classroom since the questionnaire of the third survey is more complicated than that of the first two surveys. Using this delivery and collection method allows respondents to ask for clarification during completion from the survey administrator which reduces potential error such as uncompleted questions that can affect the quality of responses.

4.6 Summary

This chapter explains the research process and research methodology that have been used for completing this research study. By means of the design science research process, a sequence of the research process was divided into three main phases: "problem identification", "solution design", and "evaluation".

In the problem identification phase, the problem of this study was formulated as current SWOT analysis approaches have no means to generate and prioritize SWOT factors based on customer's perspective. This problem was first raised by the author's recognition in an organisation context and then it was established through the literature search.

In the solution design phase, a sets of step used to perform SWOT analysis that is able to provide quantitative and customer-oriented SWOT factors, was designed and developed which is called IPA based SWOT analysis. The main idea of this approach is to apply

Table 4.5: Comparison of the main aspects of three self-administered method

	Delivery and collection	**Internet-mediated**	**Postal/E-mail**
General description:	Paper & pencil survey, delivery in person to complete individually and collect later	Online survey, access through a web page to complete individually	Paper & pencil survey, delivery by postal or e-mail attachment to complete individually
Respondent's characteristics for which suitable:	Literate individuals	Computer-literate individuals	Literate and computer-literate individuals
Cost:	Medium - Printing & paper - On-site staff	Low - Hosting & software (Usually free)	Medium - Follow-up mailings - Printing, paper, envelopes - Stamps
Length of Data Collection Period:	Short (1 to 3 weeks)	Medium (2-6 weeks)	Long (4-8 weeks)
Response Rate:	Fair to Good	Poor to Good	Poor to Good
Suitable types of question:	Closed questions	Closed questions	Closed questions
Complexity of questionnaire:	May be complex	May be complex	Simple and clear

Source: Saunders et al. (2009); Fink (2012a); Blair et al. (2013)

IPA for conducting SWOT analysis which produces SWOT factors based on results of customer satisfaction survey and quantifies them based on two mains aspects of IPA which are *importance* and *performance*.

The development of the IPA based SWOT analysis consists of two main tasks namely IPA and IPA-SWOT. The former involves the comparative experiment method for selection the current statistical and data mining techniques used for implicitly derived importance. Further details regarding the experiment set up and the experimental results are presented in Chapter 5 and Chapter 6 respectively. The latter involves the construction of the IPA based SWOT analysis framework in other words it involves the identification SWOT factors based on IPA in which the *importance* was derived by using the selected technique from the previous task. Further detail regarding the development of SWOT analysis from IPA is designed and presented in Chapter 7.

In the evaluation phase, three surveys were conducted for the purpose of demonstrating the application of IPA based SWOT and to evaluate the efficiency of the approach. Through the review of methods related to the survey design, the written survey using closed-ended questions was selected as a survey media. Personal contact and e-mail were used as a method to contact the respondents to take part in the survey. The selection was made based on type of respondents which are students and university staff. After the respondents were contacted to take part in the survey, the first two surveys were administered to the respondents using internet-mediated method while the third survey was administered to the respondents using delivery and collection method. Further details of the survey design including pilot study, sample size and ethics approval related to each survey can be seen in Chapter 8.

Chapter 5

Experiment: Empirical comparison of importance measuring techniques

It was suggested in Chapter 3 that currently there are no well-established importance measurement techniques and different importance measurement techniques are likely to result in identifying dramatically different attributes for improvement. Such unsettled issue of IPA leads to the conducting of a comparative experiment described in this chapter.

The main function of this chapter is explaining how the comparative experiment is to be conducted in order to select a potential importance measurement technique to be used for SWOT analysis so as to answer the first research question "Which importance measure should be used in IPA?" The techniques for measuring *importance* being investigated are (1) Direct-rating scales (DR), (2) MLR, (3) OLR, (4) BPNN, (5) Naïve Bayes and (6) Bayesian Networks (BNs). The first four techniques have been used to measure *importance* while the last two have never before been applied to derive *importance* from the survey data. Attribute importance measured by these techniques will be compared against three evaluation metrics.

This chapter is organised as follows: section 5.1 reports the methodology and result of two preliminary experiments. Section 5.2 identifies criteria for evaluating techniques for determining *importance* namely predictive validity, discriminating power and diagnosticity power. Section 5.3 describes the datasets used in this experiment. Section 5.4 illustrates the methodology of empirical comparison.

5.1 Preliminary study

In order to ensure that the comparative study of importance measuring techniques described in the section 5.4 is conducted properly, two experiments were conducted as part of the preliminary study. Specifically, the first experiment involved the replication of Mutual Information (MI) based IPA and the second experiment involved the test of software configuration that used in the main experiment.

5.1.1 Replication of the IPA based MI

The objective of this experiment is to demonstrate that the author has gained good understanding regarding the IPA approach which can ensure that the method for calculating *importance* and *performance* is implemented correctly in this study. Through this experiment, Mutual Information (MI) based IPA proposed by Shieh and Wu (2011) was replicated on the dataset that used by Cugnata and Salini (2013) (henceforth, the referenced paper) then the IPA result from the replication was compared to that of the referenced paper.

According to Shieh and Wu (2011), IPA based MI matrix is constructed by calculating the *performance* as medians and calculating the *importance* as MI that measures the dependence between the performance of each attribute of a company's service and overall customer satisfaction. After the IPA based MI matrix is constructed, this matrix is then compared to the IPA matrix of the referenced paper by exploring the position of each attribute of a company's service on both IPA matrices.

The comparison result shows that all attributes of a company's service fall in the same quadrant of both IPA matrices and *performance* of two matrices are obviously the same. To confirm that both IPA matrices are similar, the t-test is conducted to test the mean different of *importance* computed in this experiment and *importance* of referenced paper. The t-test result shows that the means of the *importance* in this study and *importance* of referenced paper are not different at 95% confidence level.

Consequently, it can be reasonably concluded that the MI based IPA has been correctly replicated in this experiment. This also demonstrates that the author has gained good understanding regarding the IPA approach, therefore it can be ensured that the further experiment described in the next following sections are conducted properly. Further details regarding dataset, tools, methodologies as well as result of this experiment are described in the Appendix A.

5.1.2 Comparative Analysis of Classification Algorithms using WEKA

The objective of this experiment is to verify that WEKA 3.6.10, data mining tool used in this study, was correctly installed and configured. To do so, two classifiers used in this work for measuring *importance* - namely BPNN and BNs and another classifier named J48graft were implemented on the Pima Indians Diabetes (dataset from the UCI Machine Learning Repository[1]) using WEKA. Then, the accuracy of the three classifiers was compared to the accuracy of the same set of classifiers tested on the same dataset published by Rahman and Afroz (2013).

The paper published by Rahman and Afroz (2013) was selected as the referenced paper because this paper conducted the comparison of various classifiers using WEKA on the standard datasets in the field of data mining named UCI Machine Learning Repository. Additionally, two classifiers used in Rahman and Afroz (2013) namely BPNN and BNs are similar to the classifiers used in this research.

The results on both training and testing data showed that the accuracy of classifiers implemented through this experiment is similar to the accuracy of classifiers published by Rahman and Afroz (2013) for all classifiers. Thus, it could be concluded that WEKA 3.6.10 is installed and configured correctly so it can be used for further experiments. More details regarding dataset, tool, methodology and result of this experiment are described in Appendix B.

5.2 Evaluation Metrics

Building upon the previous comparative studies, three evaluation metrics were selected to use in this study which are predictive validity (ability to predict overall customer satisfaction) (Hauser, 1991; Gustafsson and Johnson, 2004; Chrzan and Golovashkina, 2006; Taplin, 2012b), diagnosticity (ability to identify the consumers' most important attributes)(Gustafsson and Johnson, 2004; Pokryshevskaya and Antipov, 2014), and discriminating power (ability to provide discriminating measures) (Cohen, 2003; Chrzan and Golovashkina, 2006).

5.2.1 Predictive validity

Predictive validity is a measure of how well a predicted overall customer satisfaction computed by each technique for measuring *importance* correlates with the actual overall customer satisfaction that has been provided. It was one of the most used metrics in the previous comparative studies. Since, the *importance* is measured by discovering the relationship between performance of an organisation's attributes and overall customer

[1] http://archive.ics.uci.edu/ml/datasets.html

satisfaction, the *importance* results of each technique should be compared based on the ability to predict overall customer satisfaction.

For each individual, the predicted overall customer satisfaction is computed as the sum of products between *performance* and *importance* regarding a multi-attribute attitude model, Equation 5.1.

$$\widehat{Y_j} = \sum_{i=1}^{n} P_{ij} I_i \tag{5.1}$$

$\widehat{Y_j}$ is the predicted overall customer satisfaction for individual j, n is number of attributes, P_{ij} is the rated *performance* for individual j on attribute i and I_i is the *importance* of attribute i obtained from each technique for measuring *importance*.

Generally, the value of correlation coefficients fall between -1 and $+1$ and the closer the correlation coefficient is to ± 1, the stronger the correlation (Field, 2009, p.170). It is expected that the overall customer satisfaction has positive correlations with the predicted overall satisfaction which means the value of actual overall customer satisfaction is moved in the same direction of the value of predicted overall customer satisfaction. Therefore, the model with the higher correlation coefficient is superior at predictive validity than the model with the lower correlation coefficient.

In addition, a baseline of predictive validity is generated by correlating the actual overall customer satisfaction and the predicted overall customer satisfaction computed as Equation 5.1 in which I_i is set to one indicated that each attribute is equally important. If the *importance* obtained from the selected technique is a valid indicator of the overall customer satisfaction then the correlation coefficient generated by correlating between the actual overall customer satisfaction and the predicted overall customer satisfaction computed from the *importance* measure of the technique should be greater than this baseline correlation coefficient.

5.2.2 Diagnosticity

Diagnosticity is an ability of the technique for measuring *importance* to identify just which attributes of an organisation's product or service are most important in affecting customer satisfaction (Gustafsson and Johnson, 2004) in other words it is the ability of the technique to distinguish the most important attribute from the less important attributes. Diagnosticity of each technique for measuring *importance* is assessed as the regression coefficients of the regression equation tested on the basis of relationship between *importance* and rank (Doyle et al., 1997).

The regression model to test the relationship between *importance* and rank is shown in Equation 5.2. This is an equation used in Pokryshevskaya and Antipov (2014) which suggested a modification of formula used in Doyle et al. (1997) from $Rank^2_{Residual}$ to $lnRank_{Residual}$.

$$Importance = \beta_1 Rank + \beta_2 lnRank_{Residual} + constant \quad (5.2)$$

Where $Importance$ is the *importance* measures of a particular technique, *Rank* is the rank order of attributes importance ranging from 1 (most important) to number of attribute (n) (least important), β_1 is a linear coefficient for *Rank*, $lnRank_{Residual}$ is a residual of natural-log rank, and β_2 is a non-linear coefficient for $lnRank_{Residual}$. Note that $lnRank_{Residual}$ is computed by subtracting the *lnRank* with the result from regressing *lnRank* on Rank. The use of $lnRank_{Residual}$ instead of *lnRank* reduces the collinearity between the *Rank* and *lnRank* which provides stable regression coefficients.

Only the non-linear coefficient (β_2) is used to evaluate the diagnostic power of each technique because it indicates how much the highest important value differs from the other important values. The case $\beta_2 > 0$, suggests that the technique for measuring *importance* is lacking in diagnosticity of the most important attribute in contrast to it having more diagnosticity of the least important attribute. Whereas $\beta_2 < 0$ suggests that the technique is able to distinguish the most important attribute and the technique with the larger magnitude of coefficient has a higher ability to diagnose the attribute with the most *importance* than the technique with a lower one. It can be simply stated that the lower the value of β_2, the better the diagnosticity.

5.2.3 Discriminating power

Discriminating power is a measure of how well a technique provides discriminating *importance* measures. This metric is employed since a good *importance* measurement technique should be able to distinguish a key driver attribute (high *performance* and high *importance*) from the other attributes. The multiplication of *performance* and *importance* for each attribute, as shown in Equation 5.3, is denoted as attributes' attitude, where A is set of attributes, P_a is *performance* of attribute a and I_a is *importance* of an attribute a.

$$Attitude_a = P_a \times I_a \quad \forall a \in A. \quad (5.3)$$

Discriminating power of each technique for measuring *importance* is assessed as the size of differences between attributes' attitude as a F-statistic obtained from a statistical test, "Repeated Measures Analysis of Variance". This produced F-statistic of each technique is comparable in which a larger F-statistic means more differentiation between attributes and a smaller F-statistic means less (Chrzan and Golovashkina, 2006).

Additionally, the baseline F-statistic is computed as the size of differences between attributes' attitude when *importance* is ignored ($I_a = 1$). It is expected that the attributes' importance acquired from each technique increases the size of difference between attributes' attitude hence the F-statistic of all techniques should be higher than the baseline F-statistic.

5.3 Datasets

Three datasets were used in this comparative study. The two datasets have been used by other researches (Kenett and Salini, 2012, p.19) and Azzalini et al. (2012), the last dataset was collected from a leading University in Thailand for the purpose of this study. The key features of the datasets are given in Table 5.1 and detail of each dataset is described in the following subsections.

Table 5.1: Basic characteristics of the datasets

	Dataset A	**Dataset B**	**Dataset C**
Source	(Kenett and Salini, 2012, p.19)	Azzalini et al. (2012)	Kasetsart University, Thailand
Business area	Media and telecommunication service	Software and consulting services	University
Scale type (performance)	5-point	10-point	5-point
Scale type (importance)	3-point	10-point	5-point
Number of records	266	4515	159
Number of attributes	7	14	7
Dependent variable	Overall satisfaction	Overall satisfaction	Overall satisfaction

5.3.1 Dataset A

This is 2010 customer satisfaction survey of a media and telecommunication service provider company, known as ABC, (Kenett and Salini, 2012, p.19). The dataset contains customer feedback from a questionnaire with 81 questions and the number of customers responding to this questionnaire is 266. Two parts of data in the dataset were selected for conducting IPA in this work including (1) assessment of overall satisfaction which is measured on a five-point Likert scale; (2) assessment of overall satisfaction level and importance level of six company's attributes - equipment and system, sales support, technical support, supplies and orders, purchasing support, and contracts and pricing.

The satisfaction level is measured based on five-point scale range from '1' (very dissatisfied) to '5' (very satisfied) and the importance level is measured based on three-point scale range from '1' (low importance) to '3' (high importance).

5.3.2 Dataset B

This dataset contains customers' response to a satisfaction survey of an IT Company producing and selling software and offering consulting services. Two parts of data in the dataset are selected for conducting IPA in this work including (1) assessment of overall satisfaction which is measured on a six-point Likert scale range from '1' (extremely satisfied) to '6' (extremely dissatisfied); (2) assessment of satisfaction level and importance level of 13 company's attributes which is measured based on 10-point scale range from '1' (extremely dissatisfied) to '10' (extremely satisfied) for satisfaction level and range from '1' (not at all important) to '10' (very important) for importance level .

5.3.3 Dataset C

This dataset were collected in a classroom via a questionnaire answered by 155 undergraduate students of Department of Computer Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University. Two parts of data in the dataset are selected for conducting IPA in this work including (1) assessment of overall satisfaction which is measured on a five-point Likert scale range from '1' (very satisfied) to '5' (very dissatisfied); (2) assessment of overall satisfaction level and importance level of six department's service attributes, for example, academic personel, teaching and learning, administration which is measured based on a five-point scale which a value range similar to the value range of overall customer satisfaction. Further information regarding this dataset and questionnaire can be found in Section 8.2 and Appendix G.1 respectively.

5.4 Methodology

Different techniques for deriving *importance* were selected to compare against each other and compare to one self-stated importance measure using direct-rating scales on criteria to evaluate the best importance measurement technique. The derived importance measurement techniques can be further categorized into two groups which are statistically inferred importance and data mining implicitly derived importance. MLR and OLR belong to statistically inferred importance. BPNN, Naïve Bayes and BNs are classified as data mining implicitly derived importance.

To compare the different *importance* measures, firstly the datasets have to be preprocessed. Secondly, *importance* is obtained from the datasets. For the direct-rating scales which is the self-stated importance method, the *importance* is calculated straightforwardly as the mean *importance* of each attribute. For the derived importance technique: MLR, OLR, BPNN, Naïve Bayes and BNs, the model of these techniques has to be implemented through the data analysis (model training) and evaluated using 10-fold cross validation[2]. Then the *importance* is derived from the outcomes of these models using the methods described in Section 3.3, Chapter 3.

The overall procedure of the empirical comparison is shown in Figure 5.1 and its detail will be described in sub-section 5.4.1 - 5.4.3 respectively.

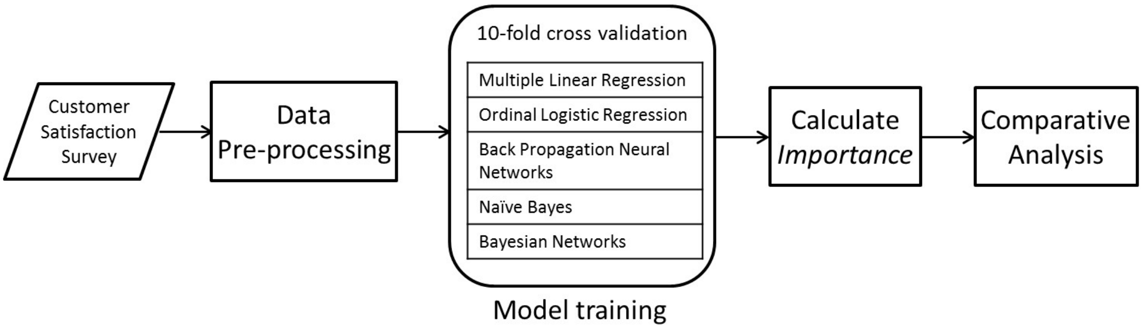


Figure 5.1: The overall procedure of the empirical comparison.

5.4.1 Data pre-processing

the main objective of data pre-processing is to prepare datasets as the input of statistical software and data mining tool used in this work. The main tasks of data pre-processing for this study consists of handling missing and inconsistent values, and formatting data.

To avoid missing values that may affect the quality of data mining model, a dataset reduction, the simplest solution for handling the missing values (Kaiser, 2014), is used in a data cleaning step. According to dataset reduction approach, a record that contains missing values in all selected attributes was removed. This elimination of the records with missing values is effective through all three datasets of this work since missing values occur only in a small percentage of records (less than 3%) as shown in Table 5.2.

Since *importance* is obtained on the basis of the relationship of *performance* and the overall customer satisfaction, the scoring scheme for the rating scale should be consistent among the value of overall customer satisfaction level, satisfaction level and *importance* level to ensure that the direction of movement between their values is similar. This work defines the higher rating scale means the higher level of satisfaction and *importance*.

[2]10-fold cross validation is k-fold cross validation, common technique for assessing accuracy of classifier or predictor, where k is set to 10 (Han and Kamber, 2000a). In 10-fold cross validation, the original dataset is divided into 10 mutually exclusive subsets (folds). The cross-validation process is then repeated 10 times. In each iteration, nine subsets (k-1) are used as training data and a single subset is retained as the test data; therefore, each subset used once for testing, nine times for training.

Table 5.2: Number of records across datasets

	Dataset A	**Dataset B**	**Dataset C**
Number of records (original)	266	4515	159
Number of deleted records	2	15	4
Number of records (after deleted records)	264	4500	155
Percentage of deleted records	0.75%	0.33%	2.50%

Therefore, the value of overall satisfaction level of Dataset B has to be reversed. To make it consistent to the value of satisfaction and *importance* level in the same dataset. The score is reversed by subtracting the original value from the highest value of response scale (in this case 6) plus 1. After reversing, the value of overall satisfaction level of Dataset B is ranged from '1' (extremely dissatisfied) to '6' (extremely satisfied). A similar process is also used to reverse the score of overall customer satisfaction level, satisfaction level and *importance* level in Dataset C.

Finally, datasets with no missing values and a consistent rating scale were converted into an ARFF (Attribute-Relation File Format) format as it is a transaction format of WEKA, the selected data mining tool of this research.

5.4.2 Model Training

In this study, two statistical techniques are implemented using SAS Enterprise Miner whereas three data mining techniques are implemented using WEKA. For each dataset, attributes' performance is selected as input attributes (independent variables) and the overall customer satisfaction is class attribute (dependent variable). Hence, there are 6, 13 and 6 input attributes for dataset A, B and C respectively (refer to Table 5.1).

According to Azzalini et al. (2012), these techniques can be divided into three groups according to a data type of the dependent variable (response variable) that is being predicted. MLR is a technique used to fit a model with data where the dependent variable is considered as quantitative discrete. Naïve Bayes and BNs are techniques used to fit a model with data that where dependent variable is considered as categorical (ignoring level order). OLR is a technique used to fit model with data where the dependent variable is considered as ordinal categorical.

Among the selected techniques, BPNN is the only technique that can fit both data types of dependent variable, considering the dependent variable as either categorical or quantitative variable. Hence, two BPNN models are implemented which are BPNN (regression) that treats the dependent variable as quantitative discrete and BPNN (classification) that treats the dependent variable as categorical. The former is BPNN model with one output neuron that is currently used to determine *importance* (Deng et al., 2008a; Chen et al., 2010; Mikulić and Prebežac, 2012). The latter is BPNN model with multiple output neurons. A summary of data type corresponds to each predictive model is shown in Table 5.3.

Table 5.3: List of model for measuring *importance* in this experiment

Model	Type of independent variables	Type of response (Dependent variable)
MLR	Quantitative discrete	Quantitative discrete
OLR	Quantitative discrete	Ordinal categorical
BPNN(regression)	Quantitative discrete	Quantitative discrete
BPNN(classification)	Quantitative discrete	Categorical
Naïve Bayes	Categorical	Categorical
BNs	Categorical	Categorical

Each model is trained using different parameter settings explained in sub-section 5.4.2.1. Subsequently, the quality of these model are measured and presented in the sub-section 5.4.2.2.

5.4.2.1 Parameter setting

Each technique has its own specific parameter to be specified in order to create the predictive model. These techniques with running parameters are given below:

Multiple Linear Regression (MLR): For MLR, both input and class attributes are numeric. MLR is a simple technique for which all parameters are set at their default value. One main parameter is method for regression which the forced entry (default option) is set. This means all input attributes are selected and entered into the model simultaneously. Before the MLR model was created from data, assumptions of MLR (such as normality, multicollinearity) were checked and found that majority of them were not violated (see Appendix C).

Ordinal Logistic Regression (OLR): For OLR, input attributes are numeric whereas class attribute is ordinal categorical for which the ordering of the categories is considered. To fit the ordinal logistic model, first the independent and dependent variables are specified then the *Complementary log-log* was selected as a link function since higher score (level of satisfaction) are more probable than lower score for all datasets. Other parameters were set at their default value. The proportional odds assumption was checked by the time of the model training and found that the assumption was not violated in two out of three datasets (see Appendix D).

Back-Propagation Neural Network (BPNN): The BPNN models were constructed for discovering the organisation's attribute that has the major influence on overall satisfaction. Recall that two types of BPNN model were implemented in this study corresponding to the type of dependent variable: BPNN(regression) or BPNN(classification). The former treat the dependent variable as numeric whereas the latter treat the dependent variable as categorical.

The BPNN contains three parts, including one input layer, one hidden layer, and one output layer. The performance of attributes is the neurons in the input layer; hence, there are 6, 13 and 6 input neurons for BPNN model of dataset A, B and C respectively. The output layer corresponds to the overall customer satisfaction in which BPNN(regression) has one output neuron whereas BPNN(classification) has n output neurons (n is possible values of overall customer satisfaction for each dataset); for example, possible class values of overall customer satisfaction of dataset A is 0-5 since it was measured in a five-point scale and missing value was treated as zero. Therefore, there are six output neurons for BPNN(classification) model of dataset A.

Both BPNN structure models are trained using the *MultilayerPerceptron classifier* in WEKA. Logistic (sigmoid) function and identity function are used as the activation functions for hidden and output neurons in BPNN(regression) respectively, whereas logistic functions are used as the activation functions for all neurons in BPNN(classification). Key parameters with their values for training these models across three datasets are shown in Table 5.4. The rest of the parameters such as *randomSeed*, *validationThreshold* are set regarding their default value in WEKA.

Table 5.4: Parameter setting of two BPNN structure models

Parameter	BPNN(regression)	BPNN(classification)
learningRate	0.3*	0.3
momentum	0.2	0.2
number of Epochs	5000*	500
decay	True	False

*For dataset A, learningRate = 0.7 and number of Epochs = 10000

The number of hidden-layer neurons of BPNN models for each dataset is assigned by training the model using the specified parameter setting and investigating the accuracy of the model with different configurations of hidden-layer neurons and then selecting the best performing network. Through the procedure for assigning the number of hidden-layer neurons of BPNN models, the number of hidden-layer neurons and the network structure for each BPNN model of three datasets is shown in Table 5.5. Further details on assigning number of hidden-layer neurons is explained in Appendix E.

Table 5.5: Number of hidden-layer neurons and network structure for each BPNN model of three datasets

Dataset	**Number of hidden-layer neurons**		**Neural Network structure model** (Input-Hidden-Output)	
	BPNN (regression)	BPNN (classification)	BPNN (regression)	BPNN (classification)
A	2	6	6-2-1	6-6-6
B	3	9	13-3-1	13-9-6
C	12	13	6-12-1	6-13-5

To give an example of both BPNN models, network architecture of BPNN(regression) and BPNN(classification) on dataset A is shown in Figure 5.2 and 5.3 respectively.

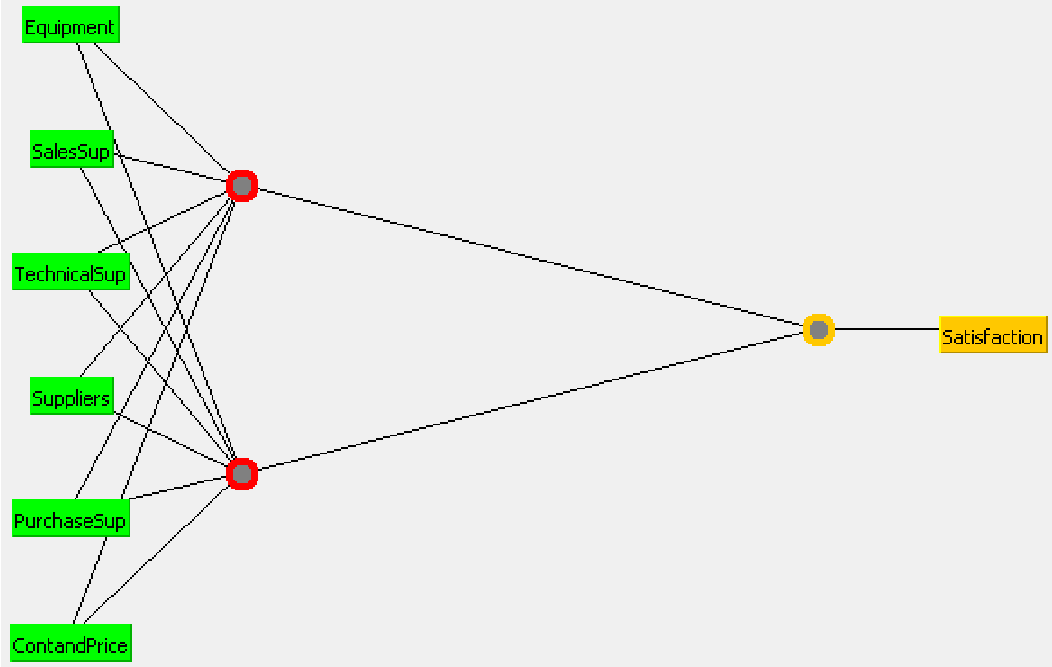


Figure 5.2: The 6-2-1 neural network as applied to dataset A.

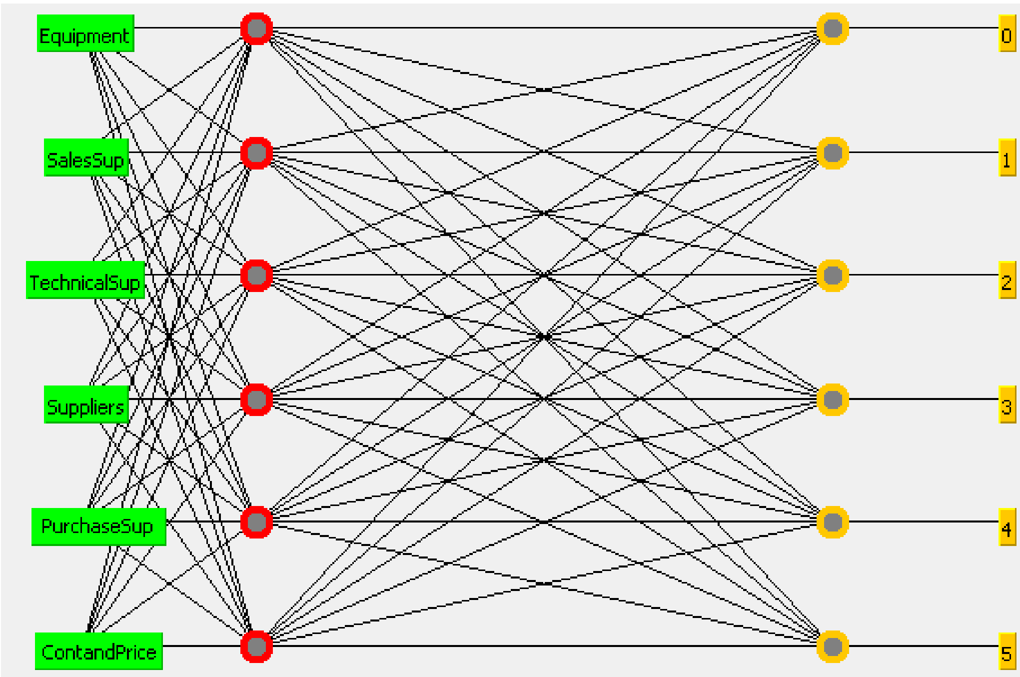


Figure 5.3: The 6-6-6 neural network as applied to dataset A.

Naïve Bayes: For Naïve Bayes, both input and class attributes of are discrete. As Naïve Bayes is a simple technique, its model can be constructed easily without a training parameter specified. An example of Naïve Bayes network is shown in Figure 5.4, which is the Naïve Bayes network of dataset A.

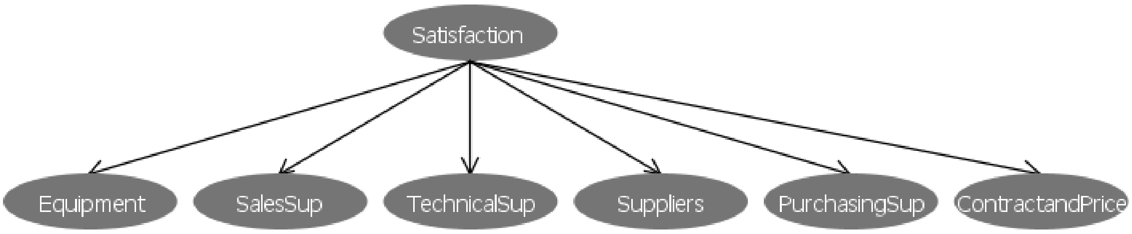


Figure 5.4: The structure of Naïve Bayes network as applied to dataset A.

Bayesian Networks (BNs): Similar to Naïve Bayes, the type of input data and class attributes for BNs are discrete. Several parameters have to be defined in order to construct the predictive model on the training data such as estimator and search algorithm. In this study, BNs structure is created by using *TAN* as search algorithm without using *ADTree* and the score type used to measure the quality of a network structure is *Akaike Information Criterion (AIC)*. Once the structure has been learned, the conditional probability tables of a BNs are estimated by using the *SimpleEstimator* with the *Alpha* value of 1, as the initial count of each value. An example of a Bayesian network is shown in Figure 5.5, which is the Bayesian network of dataset A.

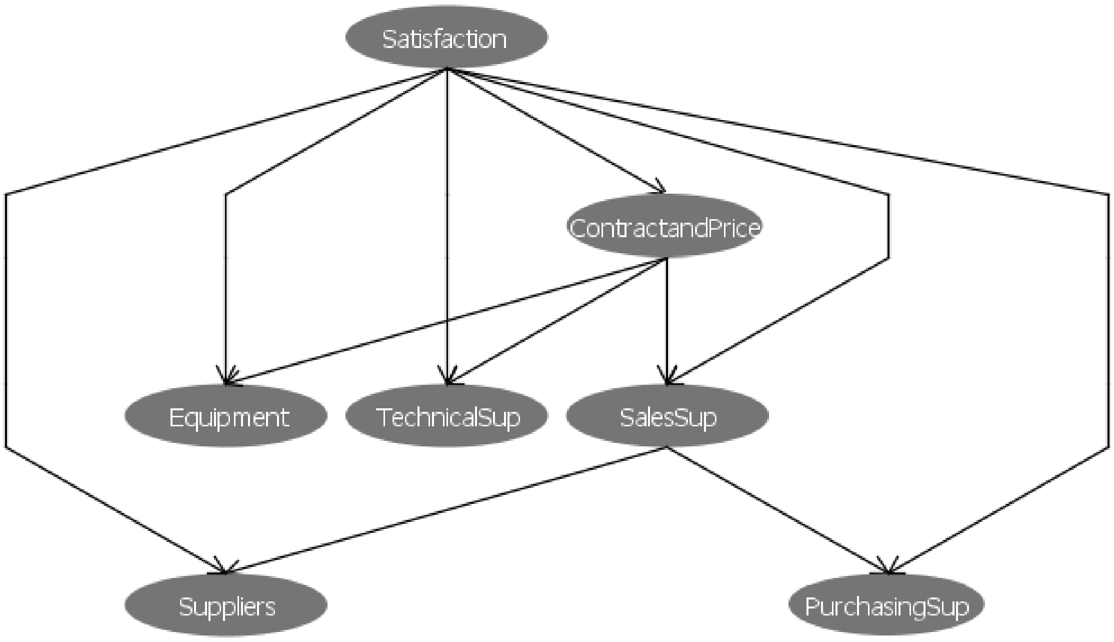


Figure 5.5: The structure of Bayesian network as applied to dataset A

5.4.2.2 Model quality

The quality of the model training may affect how well different techniques measure the *importance* since it will be derived from these models. Therefore, the quality of the model training in the previous sub-section will be evaluated in this section. In this experiment the quality of the model is measured in terms of accuracy (percentage of correct classification rate) and Root Mean Squared Error (RMSE) measured in training mode and testing mode that used 10-fold cross validation.

For each dataset, accuracy and RMSE of the model using different techniques will be compared against the baseline techniques namely ZeroR and OneR (Witten and Frank, 2005). As its name, ZeroR predicts class by considering only the class attribute (ignore input attributes). Class is assigned as the mean (for a numeric class) or the mode (for a nominal class). OneR predicts class by considering one attribute that yields the minimum-error for prediction.

Since accuracy is measured as percentage of correct classification rate, the higher accuracy is preferred to the lower one. Conversely, RMSE measures how much the predicted value differs from the actual value, therefore the lower RMSE is preferred to the higher one. In other word, the closer the RMSE is to zero, the more accurate the model is. It is expected that all selected techniques yield higher accuracy and lower RMSE than the two baseline techniques in both training and testing mode. Accuracy and RMSE of models of different techniques is shown in Table 5.6 - 5.8, for dataset A to dataset C respectively. Accuracy and RMSE comparison of models across three datasets is shown in Figure 5.6 and Figure 5.7.

According to the average accuracy of models shown in Table 5.6 - 5.8 and Figure 5.6, all models of the selected techniques yielded higher accuracy than the baseline accuracy produced by ZeroR. There is no dominating model that yields highest accuracy consistently across three dataset. BNs yielded the highest accuracy in dataset A while BPNN(classification) yielded the highest accuracy in the last two datasets. Among the selected techniques, OLR model yielded the lowest accuracy in dataset A and dataset C, while Naïve Bayes yielded the lowest accuracy in dataset B. The accuracy of these two models was lower than or equal to the accuracy of OneR, another baseline model, in the specified datasets.

Based on the average RMSE of models shown in Table 5.6 - 5.8 and Figure 5.7, it can clearly be seen that the model of MLR and BPNN(regression) yielded much higher error in terms of RMSE than the baseline RMSE (ZeroR and OneR) and models of the other techniques. MLR yielded the highest error rate across three datasets. BNs, OLR, and BPNN(classification) yielded the lowest error rate in terms of RMSE in dataset A, dataset B and dataset C respectively.

Table 5.6: Accuracy and RMSE of models using different techniques on dataset A

Model	**Train**		**Test** (10-fold cross validation)		**Average**	
	Accuracy	**RMSE**	**Accuracy**	**RMSE**	**Accuracy**	**RMSE**
ZeroR	44.70%	0.34	44.70%	0.34	44.70%	0.34
OneR	56.06%	0.38	54.92%	0.39	55.49%	0.39
MLR	-	0.79	-	0.82	-	0.81
OLR	54.17%	0.32	49.24%	0.40	51.71%	0.36
BPNN (regression)	-	0.77	-	0.80	-	0.79
BPNN (classification)	65.53%	0.28	46.97%	0.35	56.25%	0.32
Naïve Bayes	62.50%	0.29	53.03%	0.32	57.77%	0.31
BNs	72.73%	0.25	53.41%	0.32	63.07%	0.29

Table 5.7: Accuracy and RMSE of models using different techniques on dataset B

Model	**Train**		**Test** (10-fold cross validation)		**Average**	
	Accuracy	**RMSE**	**Accuracy**	**RMSE**	**Accuracy**	**RMSE**
ZeroR	68.62%	0.27	68.62%	0.29	68.62%	0.28
OneR	70.84%	0.31	69.89%	0.32	70.37%	0.32
MLR	-	0.51	-	0.51	-	0.51
OLR	70.47%	0.26	70.62%	0.26	70.55%	0.26
BPNN (regression)	-	0.50	-	0.50	-	0.50
BPNN (classification)	74.56%	0.28	70.22%	0.26	72.39%	0.27
Naïve Bayes	58.40%	0.36	57.80%	0.38	58.10%	0.37
BNs	73.20%	0.25	67.38%	0.28	70.29%	0.27

Table 5.8: Accuracy and RMSE of models using different techniques on dataset C

Model	**Train**		**Test** (10-fold cross validation)		**Average**	
	Accuracy	**RMSE**	**Accuracy**	**RMSE**	**Accuracy**	**RMSE**
ZeroR	63.87%	0.32	63.87%	0.32	63.87%	0.32
OneR	77.42%	0.30	76.77%	0.31	77.10%	0.31
MLR	-	0.38	-	0.73	-	0.55
OLR	82.58%	0.31	71.62%	0.40	77.10%	0.35
BPNN (regression)	-	0.37	-	0.40	-	0.39
BPNN (classification)	90.97%	0.18	80.00%	0.26	85.49%	0.22
Naïve Bayes	83.23%	0.24	81.29%	0.25	82.26%	0.25
BNs	86.45%	0.21	80.00%	0.25	83.23%	0.23

As a model which has a higher accuracy and lower error rate is preferable, two models: BNs and BPNN(classification) produced superior model quality than the other models. Both models consistently produced the highest accuracy and lowest RMSE across the three datasets. OLR and Naïve Bayes produced a moderate model quality. Although these two models produced model accuracy lower than or equal to that of OneR in some datasets, they yielded satisfactory RMSE. MLR and BPNN(regression) produced a low model quality, since both models yielded RMSE higher than the baseline RMSE. Notably, their RMSE measured on dataset A is considerably higher.

5.4.3 Obtaining *importance*

Recall that different techniques are categorized into three groups which are self-stated importance, statistically inferred importance and data mining implicitly derived importance. Each group has its own approach to obtain *importance*. For the direct-rating scales (DR) which is self-stated importance technique, the *importance* is calculated straightforwardly as the mean *importance* of each attribute. For the statistically inferred importance techniques, MLR and OLR, the coefficients of these two regression models can be referred to as implicit importance which expresses the influence of attributes on the overall customer satisfaction.

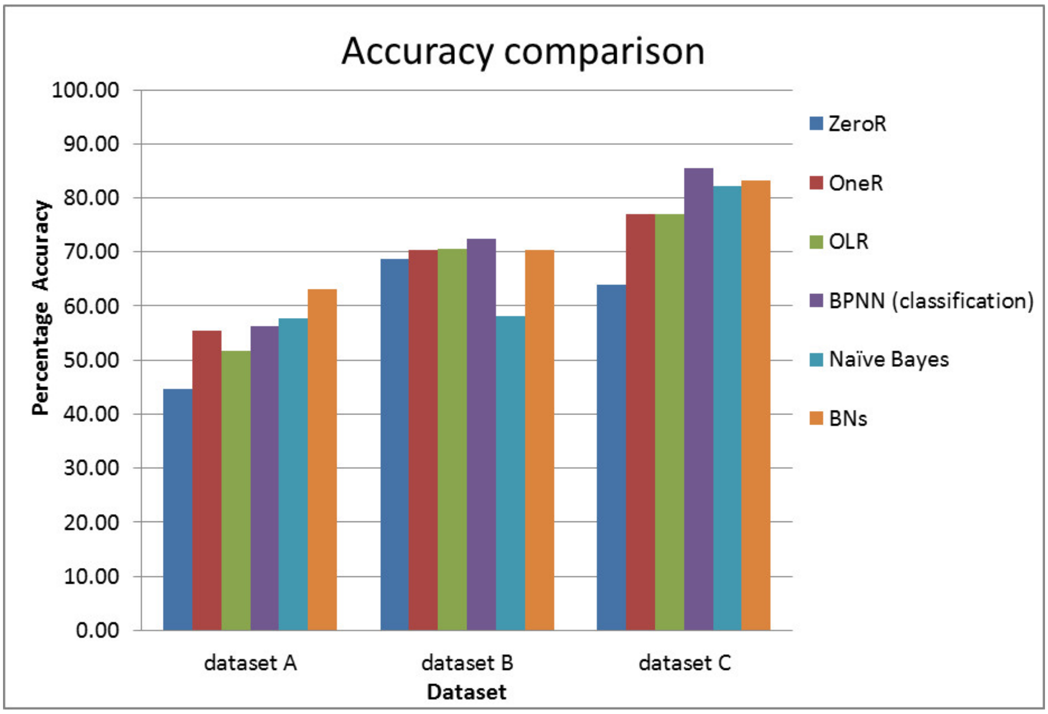


Figure 5.6: Percentage accuracy of different models across three datasets

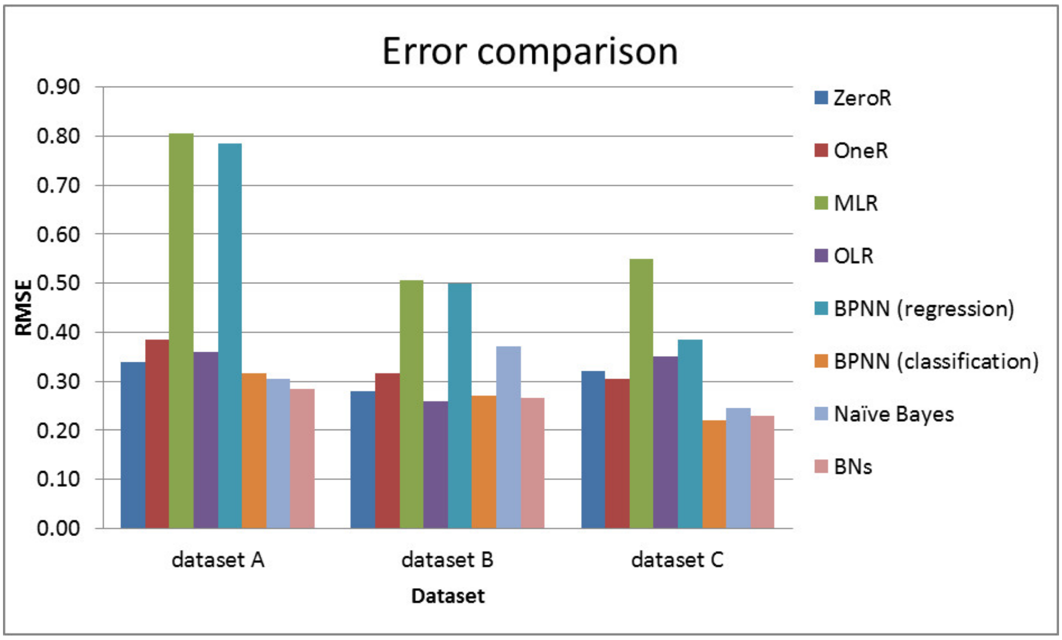


Figure 5.7: RMSE of different models across three datasets

For the data mining implicitly derived importance technique included BPNN, Naïve Bayes and BNs, the *importance* cannot be obtained directly from these data mining models due to the complexity of their outcome. Hence, a further step has to be implemented to calculate the *importance* from these model outcomes. Specifically, the *importance* can be obtained from the connection weights of BPNN model as explained in Step 2 of Section 3.3.3. The *importance* can be obtained from the conditional probability table of Naïve Bayes and BNs model as explained in Step 2 of Section 3.3.4 and Section 3.3.5. The computational example of *importance* obtained from Dataset A can be found in Appendix F.

For the purpose of comparison of the different *importance* measures, all the *importance* measured by each technique were expressed as a percentage contribution of attribute, with the negative *importance* estimated by MLR and OLR set to zero. This approach was used in previous comparative studies of *importance* measures (Gustafsson and Johnson, 2004; Pokryshevskaya and Antipov, 2014). In detail, the percentage contribution for each attribute x is calculated as the proportion of raw *importance* of attribute x to the summation of raw *importance* of all attributes. The *importance* obtained by each technique from the three datasets is shown in Table 5.9 - 5.11. The highest *importance* measure is represented as bold while the least *importance* measure is underlined.

Table 5.9: *Importance* measured by using different techniques on dataset A (percentage). Values in bold are the highest *importance* measure while values underlined are the lowest *importance* measure.

Attribute	**DR**	**MLR**	**OLR**	**BPNN (1)**	**BPNN (2)**	**Naïve Bayes**	**BNs**
Equipment and System	17.24	18.77	17.41	20.23	**28.90**	22.52	20.41
Sales Support	16.00	6.67	5.50	12.19	13.49	12.42	<u>3.58</u>
Technical Support	**17.68**	**40.82**	**47.10**	**25.59**	19.06	**22.80**	20.40
Supplies and Orders	16.70	14.97	12.89	14.33	8.79	14.36	13.66
Purchasing Support	<u>15.40</u>	<u>0.00</u>	<u>0.00</u>	<u>9.32</u>	<u>7.71</u>	<u>10.21</u>	7.43
Contracts and Pricing	16.97	18.77	17.11	18.34	22.04	17.69	**34.51**

(1) regression, (2) classification

MLR, OLR and BPNN(regression) constantly identify the same most important attribute (formatted as a bold font) across three datasets which are “Technical support”, “efficiency of products/services” and “Extra-curricular Activity”. Naïve Bayes identifies the same most important attribute as those three techniques on dataset A and dataset B whereas BPNN(classification) identifies the same most important attribute as those three techniques only on dataset C. BNs identifies differently the most important attribute from the other techniques across three datasets which are “Contracts and Pricing”, “problem solving” and “Computer Facility”.

Different techniques identify the least important attribute (formatted as underline) differently in three datasets. For dataset A, all techniques except BNs identify “Purchasing Support” as the least important attribute. For dataset B, MLR and OLR identify two attributes namely “expertise of personnel” and “helpfulness of personnel” as the least important attribute. BNs identifies “expertise of personnel” as the joint least important attribute with MLR and OLR, whereas BPNN(classification) and Naïve Bayes identify “helpfulness of personnel” as the least important attribute as do MLR and OLR.

Table 5.10: *Importance* measured by using different techniques on dataset B (percentage). Values in bold are the highest *importance* measure while values underlined are the lowest *importance* measure.

Attribute	**DR**	**MLR**	**OLR**	**BPNN (1)**	**BPNN (2)**	**Naïve Bayes**	**BNs**
V42-expertise of personnel	7.60	<u>0.00</u>	<u>0.00</u>	7.23	6.63	4.97	<u>2.82</u>
V43-consulting service capability	7.58	5.06	0.98	7.12	7.63	6.14	5.03
V44-problem solving	7.91	9.07	14.38	7.37	6.24	7.23	**27.66**
V45- reliability of products/services	**7.92**	10.48	2.59	9.11	9.47	8.35	5.31
V46-flexibility of products/services	<u>7.48</u>	5.18	5.82	4.30	8.42	8.91	6.21
V47-efficiency of products/services	7.85	**22.73**	**21.36**	**15.3**	8.54	**9.47**	9.46
V48-working speed of products	7.76	8.60	12.86	4.62	6.10	8.67	8.65
V49-helpfulness of personnel	7.62	<u>0.00</u>	<u>0.00</u>	12.28	<u>5.46</u>	<u>4.91</u>	2.84
V50-efficiency in serving customers	7.70	1.18	3.01	6.27	8.83	7.28	6.39
V51-predisposition towards customers' needs	7.54	3.18	6.22	5.75	6.93	7.99	4.75
V52-capacity to respond to customer's needs	7.74	15.67	17.81	7.38	6.03	8.98	8.98
V53-flexibility in making changes	7.59	4.71	4.57	<u>2.19</u>	8.19	8.43	5.17
V54-capacity for technological innovation	7.71	14.13	10.39	11.08	**11.53**	8.67	6.74

(1) regression, (2) classification

Table 5.11: *Importance* measured by using different techniques on dataset C (percentage). Values in bold are the highest *importance* measure while values underlined are the lowest *importance* measure.

Attribute	**DR**	**MLR**	**OLR**	**BPNN (1)**	**BPNN (2)**	**Naïve Bayes**	**BNs**
Teacher	**17.45**	15.84	27.43	20.50	<u>10.47</u>	14.57	11.65
Teaching	16.79	27.27	21.51	27.88	21.99	**22.28**	18.50
Administration	16.75	8.17	<u>1.73</u>	<u>5.28</u>	11.23	14.13	9.55
Computer Facility	17.23	<u>3.52</u>	13.20	9.49	15.18	<u>14.02</u>	**31.72**
Additional Service	16.24	9.60	5.71	7.26	13.96	14.26	<u>8.20</u>
Extra-curricular Activity	<u>15.55</u>	**35.60**	**30.43**	**29.59**	**27.17**	20.74	20.38

(1) regression, (2) classification

The other techniques: direct-rating scales and BPNN(regression) identify "flexibility of products/services" and "flexibility in making changes" as the least important attribute respectively. For dataset C, MLR and Naïve Bayes identify "Computer Facility" as the least important attribute whereas OLR and BPNN(regression) identify "Administration" as the least important attribute. Direct-rating scales, BPNN(classification) and BNs identify different attributes as the least important attribute which are "Extra-curricular Activity", "Teacher" and "Additional Service".

5.5 Summary

This chapter reports details of a comparative experiment which was conducted to select the technique to be used in the framework for conducting a SWOT analysis based on IPA as well as to help answer the first research question. First, the methodology of the two baseline experiments was explained and results of these baseline experiments suggested that the comparative experiment was properly conducted. Then, the three evaluation metrics for determining the best importance measure technique were described which included the method for calculating their value and their interpretation. Each evaluation metric evaluates each key aspect of importance measuring such as ability to discriminate the most importance attribute, and it was selected from commonly used evaluation metrics of previous comparative studies. Next, three datasets about customer satisfaction survey used in the comparative experiment were explained. Two datasets have been used by other research, while another dataset was collected from a leading University in Thailand for the purpose of this study. The main characteristic of data for applying IPA is containing an assessment of respondents' satisfaction for a company's product or service which is measured by Likert scales and containing an assessment of overall satisfaction on the Likert scales.

Lastly, procedures and techniques used in this comparative study were described. The techniques for measuring *importance* being investigated are direct-rating scales, MLR, OLR, BPNN, Naïve Bayes and BNs. To compare the different *importance* measures, firstly the datasets have to be pre-processed. Secondly, *importance* is obtained from the datasets. For the direct-rating scales which is the self-stated importance method, the *importance* is calculated straightforwardly as the mean *importance* of each attribute. For the derived importance technique: MLR, OLR, BPNN, Naïve Bayes and BNs, the model of these techniques had to be implemented through the data analysis (model training) and evaluated using 10-fold cross validation. Then the *importance* was derived from the outcomes of these models.

The importance measures obtained from different techniques identify differently the most important and the least important attributes. The next chapter will present the comparative results of the importance measurement techniques based on the three evaluation metrics. This leads to the justification to select the proper importance measurement technique.

Chapter 6

Results and Discussion: Empirical comparison of importance measuring

Given the three evaluation metrics and *importance* measures obtained from methodology explained in the previous chapter, this chapter presents the comparative results of the six importance measuring techniques and statistical analysis of the empirical comparison results. These results will help answer research question 1 and lead to the decision to select the technique to be used for measuring *importance* in the IPA based SWOT analysis framework.

This chapter is organised as follows: section 6.1 reports the test of convergent validity of all *importance* measures. Section 6.2 describes the comparative results followed by statistical analysis of these results (Section 6.3). Section 6.4 discusses the findings of this comparative experiment and section 6.5 provides the summary of the chapter.

6.1 The test of convergent validity of different importance measure

This section describes the result of the convergent validity of different importance measures tested on the three datasets and the findings based on these results.

6.1.1 Results of convergent validity for each dataset

To investigate whether different *importance* measures reflect the same construct, the test of convergent validity among techniques for measuring attribute importance is conducted by correlating the *importance* measured by each technique. The absolute correlation coefficient is used to examine a convergent validity in which a correlation of 0.35 is used as the cut-off level for concluding that there is a convergent validity between two techniques (Van Ittersum et al., 2007)

Following the approach to examine the convergent validity, the *importance* measured by each technique as shown in Tables 5.9 - 5.11 were correlated with each other by using Kendall correlation and the results of the convergent validity tested across three datasets are shown in Table 6.1 - 6.3.

Table 6.1: Convergence validity of importance measuring techniques (Dataset A)

Technique	**DR**[a]	**MLR**	**OLR**	**BPNN** (1)[b]	**BPNN** (2)[c]	**NB**[d]	**BNs**
DR	1.00						
MLR	0.97**	1.00					
OLR	1.00**	0.97**	1.00				
BPNN(1)	1.00**	0.97**	1.00**	1.00			
BPNN(2)	0.60	0.55	0.60	0.60	1.00		
NB	1.00**	0.97**	1.00**	1.00**	0.60	1.00	
BNs	0.47	0.55	0.47	0.47	0.60	0.47	1.00

** Indicates significant at $p < 0.01$

a - Direct-rating scales, b - BPNN(regression), c - BPNN(classification), d - Naïve Bayes

Regarding Table 6.1, correlations among the vector of six attributes' importance on different techniques are moderate and high. None of them are lower than 0.47 which is higher than 0.35. Thus, it can be concluded that there is convergent validity among techniques. Direct-rating scales, MLR, OLR, BPNN(regression), and Naïve Bayes are strongly and statistically significant correlated at the 0.01 level with each other. These techniques are moderately correlated with BPNN(classification) and BNs. BNs also has moderate correlation with BPNN(classification).

Table 6.2: Convergence validity of importance measuring techniques (Dataset B)

Technique	**DR**[a]	**MLR**	**OLR**	**BPNN** (1)[b]	**BPNN** (2)[c]	**NB**[d]	**BNs**
DR	1.00						
MLR	0.37	1.00					
OLR	0.30	0.66**	1.00				
BPNN(1)	0.41	0.30	0.07	1.00			
BPNN(2)	0.05	0.22	-0.01	-0.03	1.00		
NB	0.13	0.66**	0.63**	-0.05	0.26	1.00	
BNs	0.51*	0.66**	0.74**	0.13	0.18	0.51*	1.00

* Indicates significant at $p < 0.05$, ** Indicates significant at $p < 0.01$

a - Direct-rating scales, b - BPNN(regression), c - BPNN(classification), d - Naïve Bayes

Table 6.2 shows the lack of convergent validity among techniques since some correlation coefficients are lower than 0.35. The highest correlation is 0.74 indicates that OLR and BNs are strongly and statistically significant correlated at the 0.01 level with each other. OLR and MLR are also moderately and statistically significant correlated at the 0.01 level. Naïve Bayes and BNs are moderately and statistically significant correlated at the 0.05 level with each other. Naïve Bayes and BNs are also moderately and statistically significant correlated at the 0.01 level with MLR and OLR. All implicitly derived importance techniques are positively correlated with the customer self-stated importance technique (direct-rating scales), although the degree of correlations are considered low and moderate.

From table 6.2, the lowest correlation is -0.01 indicates that OLR is correlated with BPNN(classification) in the opposite direction. There are also two negative correlation coefficients indicating that BPNN(regression) is correlated with BPNN(classification) and Naïve Bayes in opposite direction. However the magnitude of correlation is close to zero, so it can also be inferred that BPNN(regression) is not correlated with BPNN(classification) and Naïve Bayes.

Table 6.3: Convergence validity of importance measuring techniques (Dataset C)

Technique	**DR**[a]	**MLR**	**OLR**	**BPNN** (1)[b]	**BPNN** (2)[c]	**NB**[d]	**BNs**
DR	1.00						
MLR	-0.33	1.00					
OLR	0.07	0.60	1.00				
BPNN(1)	-0.07	0.73*	0.87*	1.00			
BPNN(2)	-0.47	0.33	0.47	0.60	1.00		
NB	-0.20	0.87*	0.47	0.60	0.20	1.00	
BNs	0.20	0.20	0.33	0.47	0.33	0.07	1.00

* Indicates significant at $p < 0.05$

a - Direct-rating scales, b - BPNN(regression), c - BPNN(classification), d - Naïve Bayes

Table 6.3 shows the lack of convergent validity among techniques since some correlation coefficients are lower than 0.35. The highest correlation is 0.87 indicating that BPNN(regression) and OLR, MLR and Naïve Bayes are strongly and statistically significantly correlated at the 0.05 level with each other. BPNN(classification) is moderately correlated with OLR and BPNN(regression). BNs is weakly correlated to other techniques except BPNN(regression). They are moderately correlated with each other. MLR, BPNN(regression), BPNN(classification) and Naïve Bayes are negatively correlated with direct-rating scales. OLR and BNs are positively correlated with direct-rating scales although the degree of correlations are considered low.

6.1.2 Convergent Validity Findings

The test of convergent validity across the three datasets yielded mixed results. Only the convergent validity test on dataset A shows the convergent validity among the importance measured techniques, whereas other convergent validity tests showed the lack of convergent validity among the importance measured techniques. Considering the correlation between a pair of techniques, MLR and OLR are consistently strongly or moderately correlated with each other across the convergent validity test on three datasets. None of the correlation coefficients between MLR and OLR are lower than 0.35 which demonstrates the convergent validity between these two techniques. The convergent validity between MLR and OLR corresponds to the ground theory of them since, OLR is an extension of a logistic regression which is analogous to a linear regression. The result of convergent validity is also harmonized with the rank of attribute importance of these two techniques as they identified the same most important attribute in three datasets and identified the same least important attribute in two out of three datasets.

The convergent validity tested on dataset A and dataset C demonstrates the convergent validity between BPNN(regression) and BPNN(classification) since they are moderately correlated with each other. Both BPNN(regression) and BPNN(classification) are the model based on the Back Propagation Neural Network, however; BPNN(regression) has a higher correlation with MLR and OLR than with BPNN(classification), as shown in Tables 6.1-6.3. Since BPNN can be considered as an advance version of MLR and BPNN(regression) treating the dependent variable as quantitative discrete in the same way as MLR, whereas BPNN(classification) treats the dependent variable as categorical. Thereby, these two models of BPNN used different activation functions in the output layer: identity versus logistic function.

Although Naïve Bayes can be considered as a special case of BNs (Grossman and Domingos, 2004) and these two techniques are also treated the dependent variable as categorical, Naïve Bayes has a higher correlation with MLR and OLR than with BNs. As shown in Table 6.1-6.3, Naïve Bayes is consistently strongly or moderately correlated with MLR and OLR across the convergent validity tested on three datasets. However Naïve Bayes is moderately or weakly correlated with BNs with one of the tests of convergent validity which shows the lack of convergent validity between them.

Considering the convergent validity among groups of different techniques, customer self-stated and all implicitly derived importance measures which are statistically inferred importance and data mining implicitly derived importance, seem to have low convergent validity. Only the convergent validity tested on dataset A demonstrates the convergent validity among them. Other of them shows the lack of convergent validity. This supports the result of Tontini et al. (2014) that reported the lack of convergent validity between the customer self-stated importance and indirect importance measures (calculated using principal component regression). Among the implicitly derived importance measure techniques, BNs is the only technique that is consistently positively correlated with direct-rating scales across the convergent validity tested on three datasets.

These results lead to the conclusions that MLR and OLR produces similar *importance* measures. BPNN(regression) and Naïve Bayes produces *importance* measures closely agreeing to that of MLR and OLR. All implicitly derived importance measures tend to yield different *importance* measures from the customer self-stated measures. Another interesting finding is that a different data type of the dependent variable required by each technique does not directly effect the convergent validity among techniques. For example, MLR and OLR treat the dependent variable in different type but there is high convergent validity between them. In contrast, both Naïve Bayes and BNs treat the dependent variable as categorical but there is a lack of convergent validity between them.

6.2 Result of the empirical comparison

Importance measured by the six techniques, described in Sub-section 5.4.3 of the previous chapter, was tested against the three evaluation metrics, described in Section 5.2 of the previous chapter by using SPSS. The comparison results will be described in the following sub-sections.

6.2.1 Predictive validity

The analysis of predictive validity was evaluated in respondent-level data by correlating an actual overall customer satisfaction with the predicted overall customer satisfaction computed as a multi-attribute attitude model. The results of predictive validity tested across three datasets are shown in Table 6.4 .

Table 6.4: A measure of predictive validity as correlation coefficient between multi-attribute attitude model predictions with overall satisfaction. The higher correlation coefficient (close to 1) indicates the better predictive validity.

Technique	Dataset A	Dataset B	Dataset C	Average
Satisfaction rating*	0.585	0.633	0.760	0.659
Direct-rating scales	0.593	0.633	0.759	0.662
MLR	0.660	0.651	0.781	0.697
OLR	0.657	0.649	0.773	0.693
BPNN(regression)	0.634	0.631	0.779	0.681
BPNN(classification)	0.609	0.637	0.769	0.672
Naïve Bayes	0.626	0.640	0.768	0.678
BNs	0.617	0.633	0.746	0.665

*Baseline measurement of predictive validity

With regard to the correlation coefficients shown in Table 6.4, all techniques produced a positive correlation coefficient indicating that satisfaction levels of attributes are positively related with overall customer satisfaction. All correlation coefficients are statistically significant at the 0.01 level (1-tailed). These correlation coefficients ranged from 0.58 to 0.78 indicating that correlations of multi-attribute attitude model predictions with overall customer satisfaction are moderate and high. In other words, all techniques have a good predictive validity.

None of the techniques for measuring *importance* yield the average correlation coefficient across three datasets lower than the average correlation coefficient of the baseline predictive validity which is computed as multi-attribute attitude model without *importance*. Thus it clearly shows that the *importance* obtained from these techniques is a valid indicator of the overall customer satisfaction.

Based on the average correlation coefficient across three datasets, MLR yielded the highest correlation coefficient closely followed by OLR and the lowest correlation coefficient belongs to direct-rating scales which is slightly higher than the baseline correlation coefficient. Thus it can be concluded that MLR has a superior predictive validity than other techniques. However, the correlation of MLR has insignificant cardinal differences from that of other techniques especially OLR.

6.2.2 Diagnosticity

According to Pokryshevskaya and Antipov (2014)'s approach for evaluating diagnosticity, the importance measures obtained from each technique were regressed with rank resulting in the non-linear regression coefficients shown in Table 6.5.

Table 6.5: A measure of diagnosticity as non-linear regression coefficient (β_2) of the relationship between *importance* and rank. The lower the value of non-linear regression coefficient indicates the better diagnosticity.

Technique	**Dataset A**	**Dataset B**	**Dataset C**
Direct-rating scales	0.58	-0.06	0.62
MLR	-19.21	-7.36***	-15.79
OLR	-34.86	-6.34***	5.34
BPNN(regression)	-3.87	-3.51**	-1.38
BPNN(classification)	-6.02	-1.54**	-10.72*
Naïve Bayes	0.94	1.34*	-7.95
BNs	-10.34**	-14.71**	-14.99*

* Indicates significant at $p < 0.05$

** Indicates significant at $p < 0.01$

*** Indicates significant at $p < 0.001$

As the lower the value of non-linear regression coefficient indicates the better ability to differentiate the important attribute, OLR has the highest ability to diagnose the most important attribute among techniques evaluated on Dataset A. Whereas BNs and MLR has the highest ability to diagnose the attribute with the most importance among techniques evaluated on Dataset B and Dataset C respectively.

It is also clearly seen in Table 6.5 that direct-rating scales have a lack of diagnosticity of the most important attribute since this technique yielded a positive non-linear regression coefficient for two out of three datasets and has a negative non-linear regression coefficient with a small magnitude in another dataset. This outcome provided additional evidence to confirm the limitation of customers' self-stated importance measures using direct-rating scales that it tends to have low discrimination power as customers tend to consider that all attributes are very important (Gustafsson and Johnson, 2004).

On the comparisons with direct-rating scales which is a customer self-state importance measure, all of the implicitly derived importance techniques also have superior diagnosticity than the direct-rating scales since their non-linear regression coefficients are lower than that of direct-rating scales evaluated on most datasets.

Considering the implicitly derived importance techniques, only Naïve Bayes and OLR yield positive non-linear regression coefficients. The result of diagnosticity shows that Naïve Bayes has low ability to diagnose the attribute with the most importance in two out of three datasets whereas OLR has the lowest diagnosticity power on the dataset C. Other implicitly derived importance techniques consistently produced negative non-linear regression coefficients across three datasets indicating that most implicitly derived importance techniques have the ability to diagnose the attribute with the most important.

6.2.3 Discriminating power

The F-statistic for the repeated measured analysis of variance of each technique for measuring *importance* is shown in Table 6.6, and all show significant differences among attributes' attitude ($p < 0.001$).

Table 6.6: A measure of discriminating power as F-statistic between attribute discrimination. The larger F-statistic indicates the better discriminating power.

Technique	**Dataset A**	**Dataset B**	**Dataset C**
Satisfaction rating*	22	361	40
Direct-rating scales	48	477	75
MLR	1556	50267	2666
OLR	1863	48647	2918
BPNN(regression)	524	32074	2616
BPNN(classification)	715	9914	1112
Naïve Bayes	435	6641	544
BNs	836	40927	1181

*Baseline measurement of discriminating power

None of the techniques for measuring *importance* yield an F-statistic lower than baseline F-statistic which means attributes' importance acquired from each technique increase the size of difference between attributes' attitude. All implicitly derived importance techniques yield significant cardinal differences of F-statistic from the baseline F-statistic across all three datasets.

Recall that the larger F-statistic, indicates the more differentiation between attributes and a smaller F-statistic means less. Two techniques, MLR and OLR, clearly stand out as producing much larger values than the other measures. Specifically, OLR yielded the largest F-statistic in two out of three datasets which are dataset A and Dataset C while MLR yielded the largest F-statistic in dataset B. BNs yielded the third largest F-statistic in two out of three datasets. The direct-rating scales produced the lowest F-statistic which is close to the baseline F-statistic.

6.2.4 Summary of comparative results

Ranking of the techniques according to their predictive validity, diagnosticity and discriminating power presented in Table 6.7 (rank 1 corresponds to the best technique, rank 7– to the worst one), indicated that there is no clear winning technique since no one technique outperforms all the others across three metrics. MLR has the best predictive validity and diagnosticity whereas OLR has the best discrimination. However, considering the average rank of three evaluation metrics in Table 6.7 leads to the following findings which help answer the first research question:

(1) MLR is the best importance measurement technique based on the average rank of three evaluation metrics since it has the best predictive validity and diagnosticity power, and the second best discriminating power. OLR is the second best importance measurement technique as it has the best discriminating power and its predictive validity is nearly as good as MLR and its diagnosticity power is also ranked after MLR. BNs is the third best importance measurement technique as it has the second best diagnosticity power and the third best discriminating power.

(2) Direct-rating scales is the worst importance measurement technique since it is ranked last among the importance measurement techniques in all evaluation metrics. Based on this finding, it clearly shows that the implicitly derived importance techniques are better than the customer self-stated importance technique in term of these three evaluation metrics.

(3) Naïve Bayes is the second worst importance measurement technique and the worst implicitly derived importance technique based on the average rank of three evaluation metrics since, it has the lowest discriminative power and the lowest diagnosticity power among implicitly derived importance techniques and it has a moderate predictive validity.

(4) Considering two types of BPNN model, BPNN(regression) is ranked fourth while BPNN(classification) is ranked fifth among the importance measurement techniques based on the average rank of three evaluation metrics. This indicated that BPNN(regression) outperforms BPNN(classification) in measuring *importance*.

The first and second finding reveal the best and the worse importance measurement techniques based on the average rank of three evaluation metrics. These findings provide the answer to the first research question through its sub research questions. Specifically, the first finding provides the answer to sub research question 1.2 that MLR is the most appropriate technique for measuring *importance*. The second finding provides the answer to sub research question 1.1 that implicitly derived importance works better in measuring *importance* than customer self-stated importance. The other findings reveal the techniques that are ranked between the best and the worse importance measurement technique.

6.3 Statistical analysis of the results for the empirical comparison

To provide a statistical foundation for the findings in Section 6.2, the difference in the three evaluation metrics were statistically tested. Each metric results in different types of statistic which are correlation coefficients, regression coefficients and F-statistics respectively. Hence, the differences in the correlation coefficients, regression coefficients and F-statistics were tested corresponded to each metric.

6.3.1 Statistical test of Predictive validity

For each dataset, the significance tests for differences in predictive ability were tested following Williams's t-test for testing equality of dependent correlation coefficients as suggested in Chen and Popovich (2002) since all correlations were computed based on the same dataset and the sample size exceeds 20.

These differences in the correlations were tested in order to investigate whether the actual overall customer satisfaction (variable j) related to the predicted overall customer satisfaction of technique k were statistically significant different from that of technique h. The null hypothesis is stated as $r_{jk} = r_{jh}$ and the alternative hypothesis is $r_{jk} \neq r_{jh}$

The test results are shown in Table 6.8 - 6.10 as absolute t-values of the significance tests for differences in predictive validity for dataset A, B and C respectively. For each table, t-values represent the degree of difference between two correlation coefficients which is a measure of the predictive validity of the technique. The t-value of zero indicates no

Table 6.7: The ranking of techniques according to the tree evaluation metrics (The lower the rank, the better the*importance* measure)

Technique	**Rank by predictive validity**				**Rank by diagnosticity**				**Rank by discriminating power**				**Mean of all metrics**	**Rank**
	Dataset			**Mean**	**Dataset**			**Mean**	**Dataset**			**Mean**		
	A	**B**	**C**	**Rank**	**A**	**B**	**C**	**Rank**	**A**	**B**	**C**	**Rank**		
Direct-rating scales	7	5	6	6.0	6	6	6	6.0	7	7	7	7.0	6.33	7
MLR	1	1	1	1.0	2	2	1	1.7	2	1	2	1.7	1.44	1
OLR	2	2	3	2.3	1	3	7	3.7	1	2	1	1.3	2.44	2
BPNN(regression)	3	7	2	4.0	5	4	5	4.7	5	4	3	4.0	4.22	4
BPNN(classification)	6	4	4	4.7	4	5	3	4.0	4	5	5	4.7	4.44	5
Naïve Bayes	4	3	5	4.0	7	7	4	6.0	6	6	6	6.0	5.33	6
BNs	5	5	7	5.7	3	1	2	2.0	3	3	4	3.3	3.67	3

difference between predictive validity of the two techniques and the bigger the t-value, the more difference between predictive validity of the two techniques.

Additionally, to determine whether the difference between the predictive validity of the two techniques is statistically significant, the p-values of the tests were compared to the significance level of 0.05, 0.01 and 0.001. The symbol '*' indicates that p-value is less than the significance level of 0.05, the symbol '**' indicates that p-value is less than the significance level of 0.01 and the symbol '***' indicates that p-value is less than the significance level of 0.001.

Table 6.8: Significance tests for differences in predictive ability (Dataset A)

Technique	**SR**[a]	**DR**[b]	**MLR**	**OLR**	**BPNN** (1)[c]	**BPNN** (2)[d]	**NB**[e]	**BNs**
SR	-							
DR	3.67***	-						
MLR	3.54***	3.36***	-					
OLR	3.01**	2.81**	0.83	-				
BPNN(1)	5.13***	5.09***	2.11*	1.50	-			
BPNN(2)	1.90	1.38	2.85**	2.28*	2.94**	-		
NB	4.87***	4.52***	2.39*	1.79	3.82***	2.39*	-	
BNs	1.88	1.50	2.59*	2.04*	1.43	0.68	0.75	-

* Indicates significant at $p < 0.05$

** Indicates significant at $p < 0.01$

*** Indicates significant at $p < 0.001$

a - Satisfaction Rating (baseline measurement of predictive validity)

b - Direct-rating scales, c - BPNN(regression), d - BPNN(classification), e - Naïve Bayes

Table 6.4 and Table 6.8 show that correlations of most techniques except BPNN(classification) and BNs are statistically significantly different to the correlation of the baseline predictive validity. These tables also show that the highest correlation which belongs to MLR is significantly higher than the correlation of other techniques except OLR, the second best predictive validity technique.

The first column of Table 6.9 shows that correlations of most techniques are statistically significantly different to the baseline correlation except direct-rating scales and BNs since both techniques yield exactly the same correlation coefficient as the baseline correlation. Table 6.4 and Table 6.9 show that the highest correlation which belongs to MLR is statistically significantly higher than the correlation of other techniques at the 0.001 level of significance. These tables also show that the lowest correlation which belongs to BPNN(regression) is significantly lower than the correlation of other techniques except BNs. There are significantly differences between the correlations of techniques ranked in the top-5.

The first column of Table 6.10 shows that only the correlation produced by Naïve Bayes is statistically significantly different to the baseline correlation. Other techniques produce correlations that are not significantly different to baseline correlation. Considering the

Table 6.9: Significance tests for differences in predictive ability (Dataset B)

Technique	**SR**[a]	**DR**[b]	**MLR**	**OLR**	**BPNN (1)**[c]	**BPNN (2)**[d]	**NB**[e]	**BNs**
SR	-							
DR	0.00	-						
MLR	10.30***	10.77***	-					
OLR	10.01***	10.01***	2.28***	-				
BPNN(1)	2.24*	2.24*	11.47***	9.51***	-			
BPNN(2)	7.82***	7.82***	9.25***	7.90***	6.75***	-		
NB	13.90***	13.90***	8.23***	7.96***	7.88***	5.87***	-	
BNs	0.00	0.00	8.90***	10.01***	1.12	2.46*	4.83***	-

* Indicates significant at $p < 0.05$

** Indicates significant at $p < 0.01$

*** Indicates significant at $p < 0.001$

a - Satisfaction Rating (baseline measurement of predictive validity)

b - Direct-rating scales, c - BPNN(regression), d - BPNN(classification), e - Naïve Bayes

Table 6.10: Significance tests for differences in predictive ability (Dataset C)

Technique	**SR**[a]	**DR**[b]	**MLR**	**OLR**	**BPNN (1)**[c]	**BPNN (2)**[d]	**NB**[e]	**BNs**
SR	-							
DR	0.42	-						
MLR	1.93	1.94	-					
OLR	1.15	1.22	1.06	-				
BPNN(1)	1.92	1.97	0.44	1.53	-			
BPNN(2)	1.59	1.63	1.59	0.41	1.33	-		
NB	2.47*	2.79**	1.53	0.51	1.46	0.31	-	
BNs	1.65	1.53	2.47*	2.11*	2.70**	2.89**	2.64**	-

* Indicates significant at $p < 0.05$, ** Indicates significant at $p < 0.01$

a - Satisfaction Rating (baseline measurement of predictive validity), b - Direct-rating scales

c - BPNN(regression), d - BPNN(classification), e - Naïve Bayes

highest correlation which belong to MLR, this correlation is only significantly higher than the correlation of BNs at 0.05 level (see Table 6.4 and Table 6.10). Table 6.10 together with Table 6.4 also show that the lowest correlation which belongs to BNs is significantly lower than the correlation of other techniques except satisfaction rating and direct-rating scales.

Based on the correlation coefficients shown in Table 6.4 and their significance tests shown in Table 6.8 - 6.10, the following observations can be made.

(1) The implicitly derived importance techniques generally outperform the self-stated importance technique using direct-rating scales in term of predictive validity, although the differences are not entirely statistically significantly different for the three datasets.

(2) Among implicitly derived importance techniques, MLR is the best predictive validity technique since the correlations of MLR are significantly higher than most techniques

tested on two out of three datasets.

(3) Among implicitly derived importance techniques, BNs seem to have a low predictive validity since its correlations are not statistically significantly different from the baseline correlation tested across three datasets.

6.3.2 Statistical test of Diagnosticity

For each dataset, the significance tests for differences in diagnosticity were tested following an approach for the significance tests of two independent regression coefficients described in Kenny (1987) since the non-linear regression coefficients (Table 6.5) being tested were computed from importance measures of different techniques.

The null hypothesis to be tested is that a pair of regression coefficients is not significantly different from one another. This test of difference produces Student's t-statistic with N-4 degree of freedom where N is the sample size of two groups.

Table 6.11 - 6.13 show absolute t-values of the significance tests for differences in regression coefficients of dataset A, B and C respectively. T-values in each table represent the degree of difference between two regression coefficients which is a measure of diagnosticity of the technique. The t-value of zero indicates no differences between diagnosticity of the two techniques and the bigger the t-value, the more difference between diagnosticity of the two techniques.

Additionally, to determine whether the difference between diagnosticity of the two techniques is statistically significant, the p-values of the tests were compared to the significance level of 0.05, 0.01 and 0.001. The symbol '*' indicates that p-value is less than the significance level of 0.05, the symbol '**' indicates that p-value is less than the significance level of 0.01 and the symbol '***' indicates that p-value is less than the significance level of 0.001.

Table 6.11: Significance tests for differences in diagnosticity (Dataset A)

Technique	**DR**[a]	**MLR**	**OLR**	**BPNN** (1)[b]	**BPNN** (2)[c]	**NB**[d]	**BNs**
DR	-						
MLR	1.54	-					
OLR	2.42*	0.80	-				
BPNN(1)	2.47*	1.18	2.10	-			
BPNN(2)	1.97	0.99	1.92	0.57	-		
NB	0.11	1.52	2.39*	1.28	1.48	-	
BNs	1.45	0.60	1.49	0.84	0.53	1.37	-

* Indicates significant at $p < 0.05$

a - Direct-rating scales, b - BPNN(regression), c - BPNN(classification), d - Naïve Bayes

Table 6.12: Significance tests for differences in diagnosticity (Dataset B)

Technique	**DR**[a]	**MLR**	**OLR**	**BPNN** (1)[b]	**BPNN** (2)[c]	**NB**[d]	**BNs**
DR	-						
MLR	7.96***	-					
OLR	6.17***	0.74	-				
BPNN(1)	4.13***	3.11**	2.15*	-			
BPNN(2)	4.16***	5.92***	4.45***	2.16*	-		
NB	2.81*	8.35***	6.78***	4.99***	4.72***	-	
BNs	5.01***	2.40*	2.70*	3.68**	4.47***	5.41***	-

* Indicates significant at $p < 0.05$

** Indicates significant at $p < 0.01$

*** Indicates significant at $p < 0.001$

a - Direct-rating scales, b - BPNN(regression), c - BPNN(classification), d - Naïve Bayes

Table 6.13: Significance tests for differences in diagnosticity (Dataset C)

Technique	**DR**[a]	**MLR**	**OLR**	**BPNN** (1)[b]	**BPNN** (2)[c]	**NB**[d]	**BNs**
DR	-						
MLR	2.72*	-					
OLR	1.12	2.88*	-				
BPNN(1)	0.21	1.30	0.65	-			
BPNN(2)	3.42*	0.74	3.01*	0.94	-		
NB	1.85	1.03	2.13	0.63	0.49	-	
BNs	3.56*	0.11	3.36*	1.32	0.78	1.11	-

* Indicates significant at $p < 0.05$

a - Direct-rating scales, b - BPNN(regression), c - BPNN(classification), d - Naïve Bayes

Among three datasets, only dataset B with 13 observations (number of attributes used for building regression model) clearly shows significant differences between two regression coefficients of techniques as shown in Table 6.12 whereas Table 6.11 and Table 6.13 show that most regression coefficients of techniques tested on dataset A and dataset C were not significantly different. Since the power of the test difference between two regression coefficients is quite low (Kenny, 1987), even if the magnitude of coefficients tested on dataset A and C (Table 6.5) were quite different: with only six observations, significant differences between two regression coefficients were difficult to detect.

Based on the non-linear regression coefficient shown in Table 6.5 and their significance tests on three datasets shown in Table 6.11 - Table 6.13, the following observations can be made.

(1) Most implicitly derived importance techniques except Naïve Bayes generally have superior diagnosticity power than the self-stated importance technique using direct-rating scales, although the differences are not entirely significant different for the three datasets.

(2) Among implicitly derived importance techniques, MLR has the best diagnosticity power technique closely followed by BNs since MLR yields the top-2 best diagnosticity power while BNs yields the top-3 best diagnosticity power across the three datasets. The regression coefficients of MLR and BNs are significantly different from that of direct-rating scales in two out of the three datasets.

(3) Among implicitly derived importance techniques, Naïve Bayes has the lowest diagnosticity power since its magnitude of regression coefficients are lower than that of other implicitly derived importance techniques, although the differences are not entirely significant different for the three datasets.

6.3.3 Statistical test of Discriminating power

For each dataset, the significance tests for differences in discriminating power measured as F-statistics were tested by converting F-statistics to correlation coefficients (r) then conducting a test for equality of independent correlation coefficient. To convert F-statistics to r, first a partial eta squared (η_p^2) known as a measure of effect size was calculated from F-statistics using the formula described in Fritz et al. (2012). This η_p^2 can be also obtained directly from an output of SPSS.

Then r can be easily calculated as a square root of η_p^2 since it is a version of eta squared (η^2) which also known as the correlation ratio or R^2, and both η_p^2 and η^2 yield the same result in one-way ANOVAs (Levine and Hullett, 2002). Table 6.14 shows η_p^2 and r corresponding to F-statistics which is a measure of discriminating power of each technique across the three datasets. The closer the value of r to 1, the better discriminating power.

Table 6.14: partial eta squared and correlation coefficient corresponded to F-statistics of each technique

Technique	**Dataset A**		**Dataset B**		**Dataset C**	
	η_p^2	$\boldsymbol{r}$	η_p^2	$\boldsymbol{r}$	η_p^2	$\boldsymbol{r}$
Satisfaction rating*	0.078	0.279	0.074	0.272	0.205	0.453
Direct-rating scales	0.154	0.392	0.096	0.31	0.329	0.574
MLR	0.855	0.925	0.918	0.958	0.945	0.972
OLR	0.876	0.936	0.915	0.957	0.95	0.975
BPNN(regression)	0.666	0.816	0.877	0.936	0.944	0.972
BPNN(classification)	0.731	0.855	0.688	0.829	0.878	0.937
Naïve Bayes	0.623	0.789	0.596	0.772	0.78	0.883
BNs	0.761	0.872	0.901	0.949	0.885	0.941

* Baseline measurement of discriminating power

After r of each technique was computed for each dataset, the test for equality of several (more than two) independent correlation coefficients was conducted following an approach described in Kenny (1987). The null hypothesis is that the correlations computed in k groups are not significantly different from one another. This test of difference produces Chi-square statistic (χ^2) with $k - 1$ degree of freedom, where k is the number of correlation coefficients. If the Chi-square statistic exceeds the critical value in a Chi-squared distribution table, the null hypothesis is rejected otherwise the null hypothesis is retained.

In this study, eight groups of correlation coefficients were tested including the baseline correlation as shown in Table 6.14 therefore the degree of freedom was seven for all three datasets. The result of the test for equality of several independent correlation coefficients across three datasets is reported in Table 6.15.

Table 6.15: Chi-square statistic for the difference test of several independent correlations

	χ^2	**df**	**p-value**
Dataset A	485.41	7	< 0.001
Dataset B	14737.04	7	< 0.001
Dataset C	467.79	7	< 0.001

Table 6.15 shows that there is a significant difference among correlations produced by each technique at the 0.001 level of significance. Thus the tests for the equality of two independent correlation coefficients were conducted in order to investigate that (1) the baseline correlation as transform version of the F-statistic is statistically significantly different from correlation coefficients computed from F-statistic of each technique, and (2) the correlation coefficient computed from highest F-statistic is statistically significantly different from that of other techniques.

The tests for the equality of two independent correlation coefficients were conducted following an approach described in Kenny (1987). The null hypothesis to be tested is that the two correlation coefficients are not significantly different.

Table 6.16 - Table 6.18 show absolute z-values of the significance tests for differences in correlations (shown in Table 6.14) obtained from F-statistics in dataset A, B and C respectively. Z-values in each table represent the degree of difference between two correlation coefficients in units of the standard deviation. The z-value of zero indicates no difference between discriminating power of the two techniques and the bigger the z-value, the more difference between discriminating power of the two techniques

Additionally, to determine whether the difference between the discriminating power of the two techniques is statistically significant, the p-values of the tests were compared to the significance level of 0.05, 0.01 and 0.001. The symbol '*' indicates that p-value is less than the significance level of 0.05, the symbol '**' indicates that p-value is less than the significance level of 0.01 and the symbol '***' indicates that p-value is less than the significance level of 0.001.

Table 6.16: Significance tests for differences in correlations obtained from F-statistics (Dataset A)

Technique	SR[a]	DR[b]	MLR	OLR	BPNN (1)[c]	BPNN (2)[d]	NB[e]	BNs
SR	-							
DR	1.46	-						
MLR	15.26***	13.81***	-					
OLR	16.20***	14.74***	0.94	-				
BPNN(1)	9.80***	8.35***	5.46***	6.40***	-			
BPNN(2)	11.29***	9.83***	3.98***	4.92***	1.48	-		
NB	8.94***	7.48***	6.33***	7.27***	0.87	2.35*	-	
BNs	12.05***	10.59***	3.21**	4.15***	2.25*	0.76	3.11**	-

* Indicates significant at $p < 0.05$

** Indicates significant at $p < 0.01$

*** Indicates significant at $p < 0.001$

a - Satisfaction Rating (baseline measurement of discriminating power)

b - Direct-rating scales, c - BPNN(regression), d - BPNN(classification), e - Naïve Bayes

Table 6.17: Significance tests for differences in correlations obtained from F-statistics (Dataset B)

Technique	SR[a]	DR[b]	MLR	OLR	BPNN (1)[c]	BPNN (2)[d]	NB[e]	BNs
SR	-							
DR	1.97	-						
MLR	77.86***	75.89***	-					
OLR	77.29***	75.32***	0.57	-				
BPNN(1)	67.61***	65.64***	10.26***	9.69***	-			
BPNN(2)	42.96***	40.99***	34.90***	34.33***	24.65***	-		
NB	35.39***	33.42***	42.48***	41.91***	32.22***	7.57***	-	
BNs	73.15***	71.18***	4.71***	4.14***	5.54***	30.19***	37.76***	-

*** Indicates significant at $p < 0.001$

a - Satisfaction Rating (baseline measurement of discriminating power)

b - Direct-rating scales, c - BPNN(regression), d - BPNN(classification), e - Naïve Bayes

Based on the F-statistics shown in Table 6.6 and their significance tests shown in Table 6.16 - Table 6.18, the following observations can be made.

(1) All implicitly derived importance techniques yield F-statistics that are significantly higher than the baseline F-statistics at the 0.001 level of significance across three dataset. In other word, only the direct-rating scales yields F-statistic that is not significantly different from the baseline F-statistics.

(2) All implicitly derived importance techniques outperform the self-stated importance using direct-rating scales in term of discriminating power, since their F-statistics are significantly higher than the F-statistics of the direct-rating scales at the 0.001 level of significance for all three datasets.

Table 6.18: Significance tests for differences in correlations obtained from F-statistics (Dataset C)

Technique	**SR**[a]	**DR**[b]	**MLR**	**OLR**	**BPNN (1)**[c]	**BPNN (2)**[d]	**NB**[e]	**BNs**
SR	-							
DR	1.44	-						
MLR	14.29***	12.85***	-					
OLR	14.79***	13.35***	0.50	-				
BPNN(1)	14.29***	12.85***	0.00	0.50	-			
BPNN(2)	10.67***	9.24***	3.61***	4.11***	3.61***	-		
NB	7.85***	6.41***	6.43***	6.94***	6.43***	2.82**	-	
BNs	10.97***	9.53***	3.32**	3.82***	3.32**	0.29	3.12**	-

** Indicates significant at $p < 0.01$

*** Indicates significant at $p < 0.001$

a - Satisfaction Rating (baseline measurement of discriminating power)

b - Direct-rating scales, c - BPNN(regression), d - BPNN(classification), e - Naïve Bayes

(3) Among implicitly derived importance techniques, MLR and OLR have superior discriminating power than other techniques since their F-statistics are significantly higher than that of other technique and their F-statistics are not significantly different from one another.

(4) Among implicitly derived importance techniques, Naïve Bayes has the lowest discriminative power since its F-statistics are significantly lower than that of other implicitly derived importance techniques tested on two out of three dataset.

6.4 Discussion of the results for the empirical comparison

The findings of the comparative experiment described in Sub-section 6.2.4 and statistical analysis of the comparative results described in previous section lead to a discussion as follows:

- **Discussion about the best importance measurement technique**

  The first finding suggests that MLR is the best importance measurement technique followed by OLR and BNs. MLR is a winner in two out of three metrics and this result is strengthened by the statistical analysis of the evaluation metrics. It is also the only technique which is consistently ranked in the top-2 according to the three evaluation metrics across the three datasets. MLR gains an advantage in terms of predictive validity as its importance was calculated on the same data which was used for measuring correlation (Neslin, 1981) as with OLR, the second best performing technique.

OLR has high convergent validity with MLR. Hence, the evaluation result of their *importance* is slightly similar to each other. This is confirmed by the statistical test of differences of their results. OLR has the best discriminative power technique. However the difference is not statistically significantly different from MLR, the second best discriminative power of the technique, across the three datasets. On the other way round, the predictive validity of OLR is not statistically significantly lower than that of MLR in two out of the three datasets. The major difference of evaluation results between MLR and OLR is the diagnosticity power measured on Dataset C.

Despite the fact that both MLR and OLR are superior to the other techniques in terms of the three evaluation metrics, they produced negative coefficients in the first two datasets as shown in Table 5.9 and Table 5.10. Such negative coefficients result in uninterpretable importance measures (Gustafsson and Johnson, 2004). Since they indicate that the customer satisfaction is decrease as the attribute's performance is increase, which contradicts to an expectation of basic IPA studies that the customer satisfaction should increase as the attribute's performance is increase (Pokryshevskaya and Antipov, 2014). In another word, an attribute with a good performance should positively influent customer satisfaction, therefore, positive importance measures are strongly and theoretically expected. In consideration of the uninterpretable importance measures caused by the negative coefficients, all negative coefficients obtained from MLR and OLR in this study are set to zero according to the approach suggested by previous comparative studies of importance measures (Gustafsson and Johnson, 2004; Pokryshevskaya and Antipov, 2014).

The problem of negative regression coefficients may stem from the presence of multicollinearity among attributes. As Greene (2003) [p.57] suggested that multicollinearity in the regression method may cause the changes in the signs of coefficients (e.g. switches from positive to negative) that seem theoretically questionable. In this study, the presence of multicollinearity in the three datasets was checked by inspecting the variance inflation factor (VIF) statistic of each attribute in each dataset. The VIFs of all attributes in datasets (Appendix C) showed there was the presence of multicollinearity (VIF $>$ 1) however none of VIFs exceeds 10 which indicate the sign of serious multicollinearity requiring correction.

To avoid the issue of uninterpretable importance measures, a low proportion of negative coefficients among attributes are expected (Pokryshevskaya and Antipov, 2014). In another word, MLR and OLR would not recommend to use in measuring *importance* if they produced a high proportion of negative measures. Recall that there are 1/6 and 2/13 of negative coefficients obtained from MLR and OLR in dataset A and B (Table 5.9-5.10). These numbers indicate a low proportion of negative measures hence, the MLR and OLR are remained preferable choices to use in measuring *importance* of this study.

Among the top-3 importance measurement techniques, BNs is the only technique that is not prone to produce negative or uninterpretable importance measures. This is preferable in the context of most satisfaction studies (Pokryshevskaya and Antipov, 2014). Although, BNs has a low predictive validity which is about the same that of the baseline measurement. It is consistently ranked in the top-3 of diagnosticity power techniques and it is also the third best discriminate power technique.

The fact that the BNs is not a dominant importance measurement technique across three evaluation metrics is somewhat disappointing. Since all implicitly derived importance techniques measure the *importance* on the basis relationship of *performance* and the overall customer satisfaction. It is expected that the better the model to discover the relationship of performance and the overall customer satisfaction, the better the estimates of derived importance will be. Regarding the model accuracy shown in Table 5.6 - 5.8, BNs yields the top-2 best model accuracy measured by RMSE in the training and testing mode across the three datasets. Also, RMSE values of BNs are much lower than that of MLR.

The primary implication of this finding is that for the choice of technique to be used practically for creating IPA matrix in case there are no uninterpretable importance measures, for example- dataset C, MLR and OLR are the recommended techniques. Otherwise, BNs is a preferable technique.

- **Discussion about self-stated importance vs implicitly derived importance**

The second finding that direct-rating scales is the worst importance measurement technique, indicates that the implicitly derived importance techniques are superior at measuring *importance* to the self-stated importance using direct-rating scales as most implicitly derived importance techniques yielded better results than that of direct-rating scales in three evaluation metrics. Although the differences of the evaluation metrics are not significant different for all three datasets.

On the one hand, this finding supports the claim made by Neslin (1981) that the regression as the implicitly derived importance technique is superior to self-stated importance technique in terms of predictive ability. However on the other hand, this finding contradicts the research finding of Bacon (2003) and Gustafsson and Johnson (2004) which claimed that the self-stated importance technique is better than at least one implicitly derived importance technique - MLR.

Additionally, when comparing the predictive validity and discriminative power of self-stated importance and implicitly derived importance with the baseline measurement, it shows that *importance* measured by both types of techniques improve predictive validity and discriminative power of the baseline technique which considers each attribute is equally important when determining overall satisfaction. However the latter yields better improves of the two metrics than the former.

This minor finding confirms the superiority of implicitly derived importance over the self-stated importance and shows the merit of the *importance* measured by the techniques. In other words it shows that *importance* measured by the techniques is a valid predictor of overall customer satisfaction and a valid indicator for identifying key attributes that drive customer satisfaction as *importance* measures improve predictive validity and discriminative power when computed solely on performance.

It can be simply said that the presence of *importance* yielded a better predictive validity and discriminative power than when it was omitted (baseline measurement) and the *importance* measured by implicitly derived techniques even yielded better results than *importance* measured by self-stated technique. This result is confirmed by the statistical test of predictive validity and discriminative power of at least two out of the three datasets.

- **Discussion about statistically inferred importance vs data mining implicitly derived importance**

  The third finding suggests that Naïve Bayes is the second worst importance measurement technique and the worst implicitly derived importance technique based on the average rank of three evaluation metrics, and it ranked after BPNN(regression) and BPNN(classification). This primarily suggests that Naïve Bayes may not a suitable technique to be used for measuring *importance*.

  Although, using Naïve Bayes to measure *importance* does little to improve predictive ability from using customer self-stated importance measure, its major shortcoming is an ability to discriminate importance among attributes as evaluated by discrimination power and diagnosticity power. This is also a drawback of the customer self-stated technique. The reason for this outcome can be gleaned from the assumption of Naïve Bayes that all attributes contribute equally and independently to the class or the overall customer satisfaction in this context (Moran et al., 2009). Thereby the *importance* derived from the Naïve Bayes model in one attribute is not much different than the others and there is a small range of *importance* measures between the most important attribute and the least important attribute.

  Recall that the derived importance measurement techniques can be further categorized into two groups which are statistically inferred importance and data mining implicitly derived importance. This finding also suggests that the statistically inferred importance techniques which are MLR and OLR, are superior to all data mining implicitly derived importance techniques including BNs and BPNN(regression/classification), Naïve Bayes since, these data mining implicitly derived importance techniques are ranked from third to sixth among the importance measurement techniques respectively.

The finding that MLR is superior to both types of BPNN differ from Deng et al. (2008a) and Krešić et al. (2013) studies which reported that BPNN yielded a better improvement in estimating *importance* compared to MLR. However, these previous studies did not compare *importance* estimated by MLR and BPNN against evaluation metrics. These two previous studies compared only the model quality in learning pattern of data measured by Root Mean Squared Error (RMSE) and goodness-of-fit of the model between MLR and BPNN and concluded in favour of BPNN.

In conclusion, the third finding raised a case that the simple techniques - MLR and OLR work best. As MLR is an easy-to-use technique which makes it be a favourable choice to be used practically in the real world so as to OLR. While, the use of data mining implicitly derived importance techniques such as BPNN and BNs required a user who has good background knowledge about the techniques to specify several training parameters in order to get an accurate result. Additionally, *importance* can be obtained straightforwardly from the outcome of MLR and OLR whereas further steps are needed to estimate *importance* from the outcome of all data mining implicitly derived importance techniques.

- **Discussion about analytical techniques and type of dependent variable**

  Recall that the techniques for measuring importance fitted the model by considering the dependent variable which is overall satisfaction in three different data types: quantitative (interval scales), categorical, ordinal categorical.

  Theoretically, overall satisfaction measured in Likert scales is ordinal categorical variable (Jamieson et al., 2004). Therefore, it is expected that OLR, the technique that fits a model by treating the dependent variable as ordinal categorical data, should be the best importance measure technique. However, the first finding revealed that OLR is the second best importance measure technique and it is ranked after MLR that fits a model by treating the dependent variable as quantitative.

  Although MLR requires the dependent variable to be quantitative continuous which is not compatible with the ordinal nature of the overall satisfaction (Chen and Hughes, 2004), it is the commonly used technique for indirectly measuring importance (Lai and Hitchcock, 2015). Additionally, there has been a long controversy about treating Likert-derived data as quantitative (Jamieson et al., 2004). The first finding of this study supports the use of MLR and reveals that treating Likert-derived data as quantitative does not affect the *importance* estimation. It can be simply said that considering level order of dependent variable does not yield a better improvement of *importance* estimation.

  In addition, the fourth finding demonstrates that the *importance* estimation is not directly affected by different data types of the dependent variable required by each technique as it reports that BPNN(regression) outperforms BPNN(classification) at measuring *importance*. While it is expected that BPNN(classification) should be able to produce better importance measures than BPNN(regression) since, BPNN(classification) fits a model by treating the dependent variable as categorical which is more compatible with the nature of the overall satisfaction than the BPNN(regression) that fits a model by treating the dependent variable as quantitative.

In accordance with the implications drawn from the first and the fourth findings, comparative results of BNs and Naïve Bayes show that there is really a different ranked order between these two techniques which are 3^{rd} and 6^{th} respectively although both of them fit a model by treating the dependent variable as categorical.

Furthermore, different data types of the dependent variable required by each technique does not directly affect the convergent validity among techniques. For example, MLR and OLR treat the dependent variable as different data types but there is a high convergent validity between them. In contrast, both Naïve Bayes and BNs treat the dependent variable as categorical but there is a lack of convergent validity between them.

6.5 Summary

This chapter reports results of the comparative experiment which was conducted in the previous chapter to help answer the first research question. The empirical comparison results demonstrate that MLR is the best importance measure technique whereas direct-rating scales is the worst importance measure technique in terms of the three evaluation metrics. Hence, the author supports the use of implicitly derived importance measures instead of self-stated importance measures. The use of an implicitly derived importance technique not only provides reliable importance measures but also shortens the survey's length by removing questions about importance rating.

OLR is the second best importance measures technique which has high convergent validity with MLR. These two techniques are further classified as the statistically inferred importance measure whereas BNs, BPNN(regression/classification), and Naïve Bayes that ranked the third to sixth best importance measures are classified as data mining implicitly derived importance techniques. Thus, it can be concluded that the statistically inferred importance techniques are superior to the data mining implicitly derived importance techniques.

The statistical analysis of comparative results also confirmed the fact that MLR is the best importance measures technique since, the predictive validity of MLR is significantly higher than most techniques tested on two out of three datasets and discriminating power of MLR is significantly higher than that of other techniques except OLR tested across the three datasets. For the diagnosticity metric, diagnosticity of MLR is only significantly higher than that of other techniques except OLR in one dataset as significantly differences between two values of diagnosticity were difficult to detect in a small dataset. However, the diagnosticity of MLR is significantly different from that of direct-rating scales in two out of the three datasets.

Although MLR produced some negative regression coefficients, the proportion of them is considerably low. This drawback of MLR is not a case to be concerned and it is a trade-off for other advantages of MLR as stated in previous paragraph. Based on these findings, MLR is chosen to measure *importance* in the IPA based SWOT analysis framework. As MLR is an easy-to-use technique, the use of MLR facilitates the framework to be practically used in the real world. The details of the IPA based SWOT analysis framework are described in the next chapter.

Chapter 7

IPA based SWOT analysis

Based on the results of the previous chapter, a technique for measuring *importance* was identified and applied in a methodological framework for developing SWOT analysis from IPA results, called Importance-Performance Analysis based SWOT analysis (IPA based SWOT analysis). This framework serves as an outline of the main steps to be completed in order to obtain an organisation's SWOT from survey data.

This chapter introduces an IPA based SWOT analysis starting by describing its background in 7.1. Then, a framework of applying IPA for SWOT analysis is described in Section 7.2.

7.1 Background of IPA based SWOT analysis

Recall that there are three issues related to the development of the IPA based SWOT analysis: (1) selecting a technique for measuring *importance*, (2) identifying opportunities and threats from a customer satisfaction survey and (3) evaluation of IPA based SWOT analysis. This chapter focuses on addressing the second issue, while the first issue has already been addressed by conducting the empirical comparison of various techniques for measuring *importance* in which the methodology and results were explained in Chapter 5 and 6 respectively.

The second issue is associated with the second research question "How can IPA be applied to develop a SWOT analysis based on a customer satisfaction survey?" To address this research question, the library research was conducted to review the connection between IPA and SWOT analysis and specify the possible way to acquire all SWOT aspects from the IPA results, known as IPA matrix. The review of literature indicated the correspondence between IPA and SWOT analysis in two different ways. Several IPA papers directly interpret each quadrant of the IPA as each aspect of SWOT shown in Figure 7.1 (Duke and Mount, 1996; Kim and Oh, 2001; Luo et al., 2010).

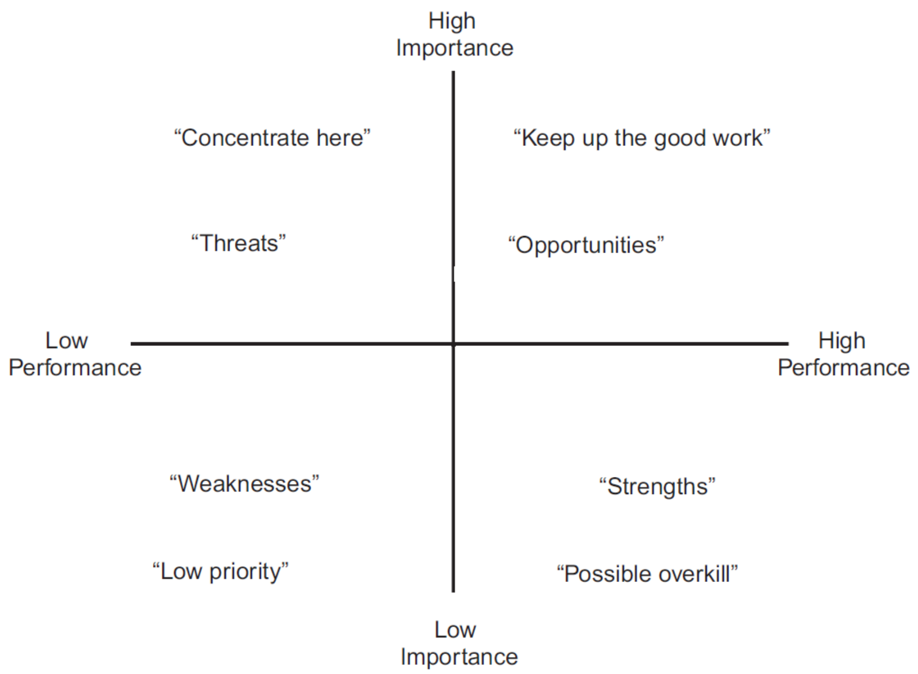


Figure 7.1: IPA matrix as SWOT (Kim and Oh, 2001)

However, this interpretation results in a mismatch between IPA and SWOT. For example, the “Possible overkill” quadrant of IPA matrix should not be interpreted as “Strength” in SWOT matrix (See Figure 7.1) since the organisation’s attributes located in that quadrant are not valued by customers thus the organisation should reallocate resources to other quadrants in need of improved performance. In contrast, the attributes defined as an organisation’s strength indicated that the organisation should reserve resources and efforts to maintain this level of performance. This example shows that the incorrect interpretation leads to the incorrect decision making regarding resources allocation. Therefore, the IPA result should not be interpreted as SWOT directly.

Besides, the review of literature indicated that most of IPA related papers only interpret each quadrant of IPA as major/minor strengths and weaknesses (Garver, 2003; Deng et al., 2008a; Silva and Fernandes, 2012; Hasoloan et al., 2012; Cugnata and Salini, 2013; Hosseini and Bideh, 2013) as previously described in Sub-section 3.1.1, Chapter 3.

Since only strengths and weaknesses can be inferred from the IPA matrix, this thesis proposes a framework to obtain all aspects of SWOT from the IPA matrix (See Figure 3.1). Briefly, the organisation’s strengths can be obtained from attributes of the organisation’s product or service located in “Quadrant1 - Keep up with the good work” and “Quadrant 2 - Possible overkill” of IPA matrix. Whereas the organisation’s weaknesses can be obtained from attributes of the organisation’s product or service located in “Quadrant3 - Low priority” and “Quadrant 4 - Concentrate here” of IPA matrix.

The same process is carried out to identify strengths and weaknesses of the competitor of target organisation. Then, opportunities and threats of the organisation can be identified by comparing strengths and weaknesses of organisation with that of its competitor based on the ideas of Pai et al. (2013) which stated that "*the strengths of competitor become the threats of the organisation and the weaknesses of competitors can become the opportunities of the organisation*".

7.2 IPA based SWOT analysis Framework

Main idea of this framework is to use IPA to analyse survey data of the organisation and its competitors, then the organisation's SWOT factor is derived from the IPA matrix as shown in Figure 7.2. The IPA based SWOT analysis comprises four steps:

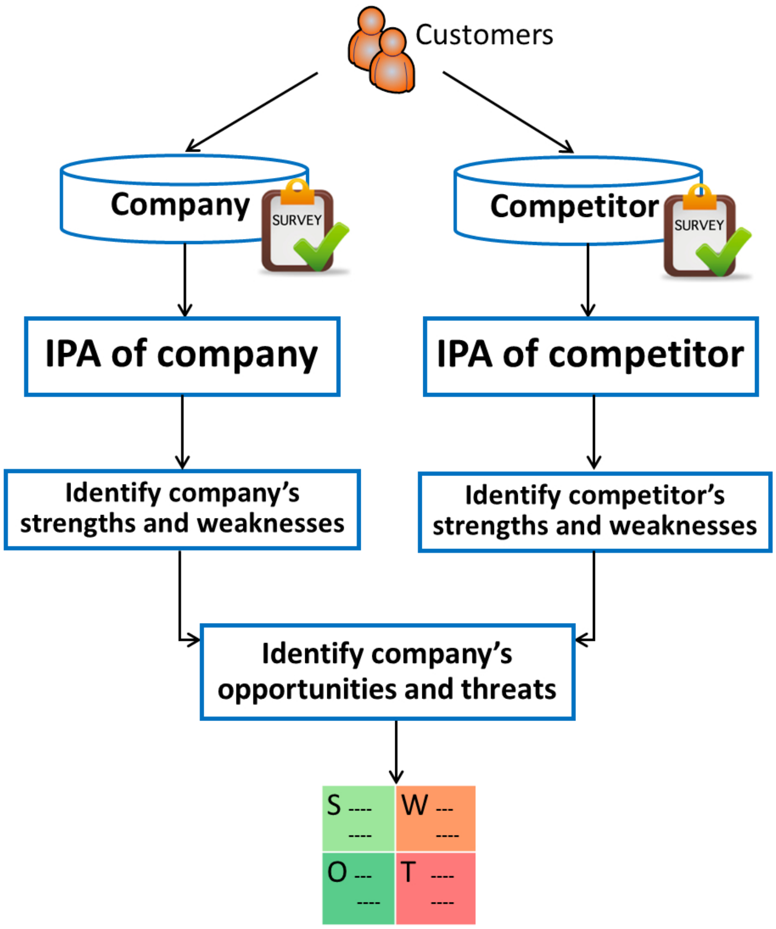


Figure 7.2: Proposed framework for applying IPA to SWOT analysis

Step 1: Undertake a customer satisfaction survey. First and foremost, attributes of an organisation's product or service are identified based on a thorough literature review in each application area or by interviews (Martilla and James, 1977; Skok et al., 2001; Levenburg and Magal, 2005). Then, a survey is developed regarding the identified attributes of an organisation's product or service.

Generally, a survey that is suitable for applying IPA consists of an assessment of respondents' satisfaction for an organisation's product or service which is measured by a Likert scale with either five or seven levels (Lai and Hitchcock, 2015). In addition, the survey should contain an assessment of overall satisfaction on the Likert scale. Two surveys using the same set of questions are required in this study. The first survey focuses on the service quality of the target organisation, while the second survey concentrates on the service quality of the organisation's competitor.

Step 2: Conduct an IPA on the customer survey. After the surveys are administered to customers of a target organisation and customers of an organisation's competitor, the customer survey data is processed to compute *importance* and *performance* for an individual attribute of an organisation's product or service. Procedures associated to this step are shown in Figure 7.3, and discussed below.

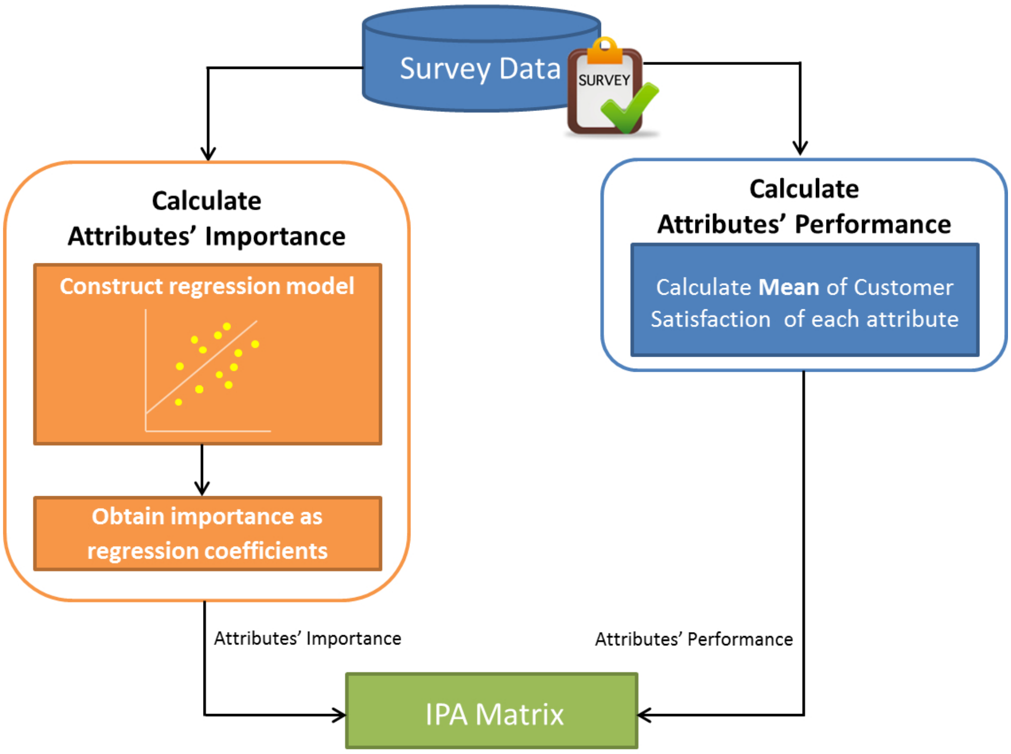


Figure 7.3: Steps for conducting IPA of customer survey data

- Calculate *importance*. Through this work, attributes' importance is derived from survey responses based on the relationships between attributes' performance and overall satisfaction instead of asking customers to rate the *importance*. Specifically, MLR is chosen to analyse the survey data and compute attribute' importance as MLR is the best implicitly derived importance method with regard to an empirical comparison results described in Chapter 6.

MLR is applied to the survey data to create a model for discovering the relationship between the attributes' performance of the product or services and overall satisfaction which reveals the attributes that influence the overall satisfaction. All attributes' performance are set as independent variables and overall satisfaction is set as dependent variable.

The regression coefficients obtained from the MLR model can be referred to as implicit importance since the regression coefficient generally indicates how much a one unit increase in the independent variable results in an increase or decrease in the dependent variable with all other variables held constant (Nathans et al., 2012).

- Calculate *performance*. The performance for each attribute of an organisation's products or services is computed by averaging performance ratings from all respondents to the questionnaire which is called "actual performance".

- Construct the IPA Matrix. The grand mean of attributes' importance and grand mean of attributes' performance are calculated, and then used to divide the IPA matrix into four quadrants. Finally, all attributes' importance and attributes' performance calculated in previous procedures are plotted on the x-axis and y-axis of IPA matrix respectively.

Step 3: Identify strengths and weaknesses through the IPA matrix. With regard to the IPA matrix produced in Step 2, an organisation's attributes located in Quadrant 1 and Quadrant 2 are identified as strengths as they have high performance, whereas an organisation's attributes located in Quadrant 3 and Quadrant 4 are identified as weaknesses as they are having low performance. Based on the same principle, strengths and weaknesses of the organisation's competitor are identified from the IPA matrix of competitor.

Step 4: Identify opportunities and threats through IPA matrix. By comparing the attributes of an organisation and its competitor that were previously labelled as strength and weakness, opportunities and threats of the organisation can be identified based on the ideas of Pai et al. Pai et al. (2013) which stated that "*the strengths of competitor become the threats of the organisation and the weaknesses of competitors can become the opportunities of the organisation*".

A summary of the identification for all aspects of SWOT and their managerial implication is presented in Table 7.1 and described as follows:

Strength (S). Attribute is labelled as an organisation's strength, since it is identified as a strength of an organisation and its competitor. This means both a target organisation and its competitor are performing well at providing this attribute. The organisation should maintain the performance of this attribute to ensure that the attribute is not turned into a threat when its performance is lower than that of the competitor.

Weakness (W). Attribute is labelled as an organisation's weakness, since it is identified as a weakness of an organisation and its competitor. This means both a target organisation and its competitor are not performing well at providing this attribute. The organisation should improve the performance of this attribute in order to obtain a competitive advantage in the target market over the competitor.

Opportunity (O). Attribute is labelled as an organisation's opportunity, since it is identified as strength of an organisation but it is identified as a weakness of a competitor. This means the competitor is not performing as well as the organisation at providing this attribute, implying that the organisation has a competitive advantage over the competitor. The organisation should maintain or leverage the performance of this attribute to stay competitive.

Threat (T). Attribute is labelled as an organisation's threat, since it is identified as a weakness of an organisation but it is identified as a strength of a competitor. This means the organisation is not performing as well as the competitor at providing this attribute, implying that the organisation has a competitive disadvantage to the competitor. Hence, an organisation should be aware of it and take an immediate action to improve the performance of this attribute in order to prevent a potential loss of profit.

Table 7.1: SWOT identification table

Strength - Weakness		SWOT aspect	Implication
Organisation	Competitor		
S	S	S	Head-to-head competition
	W	O	Competitive advantage
W	S	T	Competitive disadvantage
	W	W	Neglected opportunities

Additionally, each SWOT factor is weighted as the product of importance and performance. Specifically, a positive value of performance is assigned to the strength and opportunity factors since these factors have performance higher than or equal to the means of overall performance. On the other hand a negative value of performance is assigned to the weakness and threat factors since these factors have performance less than the means of overall performance. This weighting scheme enables factors in each SWOT aspect to be prioritized regarding the magnitude of the weight for which a factor with high magnitude that has higher priority in maintaining or improving than the lower one.

7.3 Summary

This chapter describes the IPA based SWOT analysis framework which offers a systematic way to analyse customer satisfaction surveys by using IPA in identifying SWOT, so as to answer the second research question. Through the steps explained in the proposed framework an organisation's SWOT can be obtained from survey data.

The key steps of the IPA based SWOT analysis are the IPA matrix construction in which a customer satisfaction survey is analysed to calculate the *importance* and *performance*, and SWOT factors identification based on the IPA matrix. Specifically, strengths and weaknesses are identified through an IPA matrix of the organisation, while opportunities and threats are obtained by comparing the IPA matrix of an organisation with that of its competitor.

By using IPA based SWOT analysis, the generated SWOT factors are not only measurable regarding the *importance* and *performance* but also meaningful as they are identified based on customers' points of view. A measurable SWOT factor enables an organisation to prioritize SWOT factors in creating an action plan while a customer oriented SWOT factor guarantees that the capabilities perceived by an organisation are recognized and valued by the customers. This facilitates an organisation to efficiently formulate strategic planning for maintaining or enhancing customer satisfaction, thereby gaining a competitive advantage.

To demonstrate the proficiency of the IPA based SWOT analysis in the real-world situation, the case study of Higher Education Institutions (HEIs) in Thailand are conducted and then the SWOT of the case study are evaluated. The details of case study and IPA based SWOT evaluation methodology are described in the next chapter.

Chapter 8

IPA based SWOT analysis: Evaluation Methodology

To demonstrate and evaluate the IPA based SWOT analysis which was introduced in the previous chapter, the case study of Higher Education Institutions (HEIs) in Thailand was conducted and is explained in this chapter. Three surveys were conducted as part of the case study following the survey research method. The first one is a student satisfaction survey which was conducted for the purpose of demonstrating how IPA based SWOT analysis can be used to generate SWOT. The last two surveys are staff evaluation of SWOT survey and experienced users of SWOT survey which were conducted for the purpose of evaluation of IPA based SWOT analysis so as to answer the third research question "How good is the outcome of IPA based SWOT analysis?"

The aim of this chapter is to describe the methodologies related to the case study including the survey research method for conducting the three surveys and the implementation of IPA based SWOT analysis on student satisfaction survey. This chapter is organised as follows: Section 8.1 provides background of the organisations involved in the case study. Section 8.2 describes the methodology for conducting student satisfaction survey. Section 8.3 illustrates the IPA based SWOT analysis framework implemented on the student satisfaction survey. Section 8.4 and Section 8.5 describes the methodology for conducting the staff evaluation of SWOT survey and experienced users of SWOT survey respectively.

8.1 Background of the case study

The department of Computer Engineering at Kamphaeng Saen campus (CPE-KU-KPS, henceforth KPS), Kasetsart University was selected as the case study of HEIs in Thailand, and its competitor is the department of Computer Engineering at Bangkhen Campus (CPE-KU-BKN, henceforth BKN).

KPS, the selected target organisation, was established in 2006 and located in the second campus of Kasetsart University. KPS's mission is to produce graduates in the field of computer engineering with quality, integrity and ethics, as well as support the country with know-how in the field of information technology. At present, the KPS still firmly commits to the mission and continuously improves teaching quality, research, and academic service to produce specialized graduates.

KPS was chosen as a target organisation because this department needed an improvement to be in the Thailand top 10 for computer engineering. However, KPS still has no concrete future plan since the department's future plan was formulated based on an imprecise SWOT list in which some ideas were raised from personal attitudes in the brainstorming with no supporting evidence or documents. The issues were observed based on the experience of the author who is one of the staff members in KPS.

BKN, the selected competitor, was established in 1989 and located on the main campus of Kasetsart University in the state capital. The BKN is in the Thailand top three for computer engineering. It has academic strengths and expertise in computer engineering which is guaranteed by both national and international awards, for instance, Skuba - a Small size robot soccer team- were the world champion team in RoboCup for the fourth consecutive year (2009-2012), an award-winning entertainment program of National Software Contest (NSC 2012).

BKN was chosen as a competitor of KPS because it has academic strengths and expertise in computer engineering. BKN is also ranked in the first place of the top computer engineering department in Thailand regarding the central university admissions test 2014[1], which guarantee that BKN has renowned teaching and research in computer engineering.

As with other universities, these two departments of Kasetsart University generally conduct student satisfaction surveys about the aspects of the department to assess the performance of the department. Data from these surveys can be further analysed to produce SWOT of KPS by applying the IPA based SWOT analysis framework which will provide useful information for formulating a concrete future plan for the department.

Through the case study, SWOT of KPS was generated regarding the IPA based SWOT analysis framework which was implemented on the student satisfaction surveys response

[1] www.cuas.or.th/document/57D_stat_rpass_web.pdf

by undergraduate students of KPS and BKN. Then, the SWOT of KPS was evaluated by using the staff evaluation of SWOT survey and experienced users of SWOT survey. Detail about the three surveys and implementation of IPA based SWOT analysis are described in the following sections.

8.2 Methodology for conducting student satisfaction survey

Regarding the first step of an IPA based SWOT framework (see section 7.2), a student satisfaction survey was undertaken in KPS and BKN. The results of the student satisfaction survey will be used as inputs of IPA. It is essential that the survey of the organisation and its competitor use the same questionnaire so that the IPA matrix of the competitor can be combined with the IPA matrix of the target organisation which fulfils all aspects of SWOT. Following the survey research method, tasks for undertaking student satisfaction survey are described in the following sub-sections.

8.2.1 Questionnaire design

The questionnaire used was divided into two sections. In the first section, students were asked to provide their perceived performance responses towards six departments'service attributes as well as the overall satisfaction. Subsequently, students were asked to provide their importance responses for each of the same attributes. This section contained closed-response questions, on a five-Likert scale. The second section refers to the demographic data of students such as academic year, gender and study program (regular/ special).

The six attributes that were selected based on a list of attributes defined in previous studies of student satisfaction (Siskos and Grigoroudis, 2002; Silva and Fernandes, 2011; Grebennikov and Shah, 2013) are shown below.

1. *Academic Personnel:* Teaching ability, Subject expertise, Friendliness of teaching staff;
2. *Teaching and Learning:* Lecture materials, e-learning resources, Fair assessment, Class size, and Teaching facilities and classroom condition;
3. *Administration:* Knowledge of rules and procedures of staff members, Friendliness of staff members, Ability of the staff members to provide services in a timely manner;
4. *Computer Facilities:* Quality and Availability of computer facilities, Availability of Internet access, and Availability of printing and photocopying facilities;

5. *Extra-Curricular Activities:* Exchange programs with foreign, Field trips, and Interpersonal Skills development activities etc.;

6. *Additional services:* Financial aid for students, Medical support to students, Library and Department's website.

Each attribute has a number of questions representing factors that drive student satisfaction. There are five questions related to the attribute named Academic Personnel and Administration. There are four questions related to the attribute named Computer Facilities and Additional services whereas, there are six and seven questions related to the attribute named Teaching and Learning, and Extra-Curricular Activities respectively. A complete questionnaire is presented in Appendix G.1.

8.2.2 Consistency of the translation between English and Thai version of questionnaire

Since the participants are Thai students, the questionnaire was translated from English into Thai. The consistency of the translation between the two languages was checked by four Thai native-speakers who earned postgraduate degrees in the UK in order to ensure that the Thai version of the questionnaire could be understood by participants while preserving the meaning of the original version.

In this step, four Thai native-speakers were asked individually to rate each statement or question on the degree of consistency between the two versions of languages by assigning a rating of $+1$ (consistency), 0 (undecided about whether Thai statement or question is consistent with the English version) or -1 (no consistency). They were also requested to give a comment to improve the statement or question in case that they assigned a rating of 0 or -1 to statement or question.

Subsequently, the total consistency score for each statement or question was calculated by adding up the rating of each Thai native-speaker. The results of the test of language consistency are shown in Appendix G.3. These results suggested that all Thai statements or questions could be used, because all had the total consistency score greater or equal to 3.

8.2.3 Sample size estimation

Data from student satisfaction was analysed to estimate *importance* and *performance* for each attribute of department which then can be used to produce SWOT. As previously described in Section 7.2, *importance* for the questions related to the attributes was obtained by using MLR. Hence, the minimum required sample size for data analysis using MLR is computed by using G*Power. The major input parameters and their value for estimating sample size are:

- Statistical test: multiple linear regression
- Alpha error probability: 0.05 - normal convention
- Power: 0.8 - normal convention
- Effect size f^2: 1 - to explore whether there is a relationship between attribute performance and overall satisfaction
- Number of predictor:4-7 (depending on number of questions associated to the attributes)

Given these input parameters, G*Power gives the minimum required sample size corresponded to the number of predictors (questions) for each attribute, as shown in Table 8.1. It can be drawn from this table that the minimum required sample size for the student satisfaction survey is 23 as it is the biggest minimum required sample size.

Table 8.1: Minimum required sample size for building multiple linear regression model of each attribute

Attribute	Number of questions /predictors	Minimum required sample size
Academic Personnel	5	20
Teaching and Learning	6	21
Administration	5	20
Computer Facilities	4	18
Extra-Curricular Activities	7	23
Additional services	4	18

8.2.4 Ethics

The student satisfaction survey was conducted under ethics number ERGO/FPSE/14268 which received approval of the Ethics Committee of the University of Southampton. This ethics number was presented in the participant information sheet (see detail in Appendix G.2) that was distributed to all students who took part in the survey.

8.2.5 Pilot study

Prior to the main study, the questionnaire was piloted among 32 undergraduate volunteers of KPS and BKN. They were asked to read the participant information sheet and then responded to the online questionnaires. The reliability of the questionnaire was assessed by Cronbach's alpha as one of the most frequently used methods for calculating internal consistency (Saunders et al., 2009). As a rule of thumb Cronbach's alpha value greater than 0.7 is considered reliable (De Vaus, 2002b). The reliability values for each attribute's performance is shown in Table 8.2. The Cronbach's alpha for each attribute of the department ranged from 0.86 to 0.97 as shown in Table 8.2. The Cronbach's alpha of all groups was greater than 0.70 thus it can be concluded that the questionnaire has good internal consistency.

Table 8.2: List of attributes used in the case study and Cronbach's alpha

Attribute	Question/Factor	Cronbach's alpha
Academic Personnel (Teacher1-5)	Teaching ability of teaching staff Subject expertise of teaching staff Friendliness of teaching staff Availability of teaching staff Advice and support in learning	0.86
Teaching and Learning (Teaching1-6)	Lecture materials e-learning resources Assessments (clarity and timely feedback) Class size Accurate and up-to-date unit content Teaching facilities and classroom condition	0.92
Administration (Admin1-5)	Knowledge of rules and procedures of staff Knowledge of the information about courses, exams, activities of staff Interest in solving student's problems by staff Friendliness of staff Ability of staff to provide services in a timely manner	0.97
Computer Facilities (CompFac1-4)	Quality of computer facilities Availability of computer facilities Availability of internet access Availability of printing and photocopying facilities	0.95
Extra-Curricular Activities (xActivity1-7)	Cultural exchange programs with foreign country Field trips Moral development activities Health development activities Interpersonal skills development activities Personal learning and thinking skills development activities Social volunteer activities	0.97
Additional Services (AddService1-4)	Financial aid for students Medical support to students Department website Library	0.96

8.2.6 Data collection

KPS and BKN students were asked face to face to take part in the survey by the representative of the researcher in their classroom during April and May of 2014, in the second semester of the 2014/2015 academic year. They were asked to read the participant information sheet and then responded to the online questionnaires available at: https://www.isurvey.soton.ac.uk/15462 (KPS version) and https://www.isurvey.soton.ac.uk/ 15415 (BKN version). The printed versions are shown in Appendix G.1.

A total of 155 and 43 valid questionnaires were collected for analysis from KPS and BKN respectively. Note that the sample size of both department were greater than 23 which is the minimum required sample size for data analysis using MLR computed by G*Power (see sub-section 8.2.3). Data characteristics of the samples from the two departments are shown in Table 8.3 and 8.4.

Table 8.3: General characteristics of sample CPE-KU-KPS students (N=155)

Characteristics	**Number**	**Percentage (%)**
Gender		
Male	91	58.7
Female	60	38.7
Missing	4	2.6
Study Level		
1	22	14.2
2	78	50.3
3	30	19.4
4	22	14.2
Other	1	0.6
Missing	2	1.3
Program		
Regular	95	61.3
Special	58	37.4
Missing	2	1.3

Table 8.3 lists general characteristics of sample KPS students who participated in the survey. The sample of 155 students comprise of 91 (58.7%) males and 60 (38.7%) females, with four missing entries accounting for the remaining 2.6%. The majority are the second year students which is about 50.3% of overall sample. 95 (61.3%) students are studied in a regular program and 58 (37.4%) students are studied in a special program.

Table 8.4 lists general characteristics of the sample of BKN students who participated in the survey. The sample of 43 students comprise of 36 (83.7%) males and 6 (14.0%) females, with one missing entry accounting for the remaining 2.3%. The majority are the second year students which is about 39.5% of the overall sample. 29 (67.4%) students studied in a regular program and 13 (30.2%) students studied in a special program.

Table 8.4: General characteristics of sample CPE-KU-BKN students (N=43)

Characteristics	Number	Percentage (%)
Gender		
Male	36	83.7
Female	6	14.0
Missing	1	2.3
Study Level		
1	1	2.3
2	17	39.5
3	9	20.9
4	14	32.6
Other	1	2.3
Missing	1	2.3
Program		
Regular	29	67.4
Special	13	30.2
Missing	1	2.3

8.2.7 Data analysis method

The student satisfaction survey data collected through the previous steps was then processed following step 2 - step 4 of IPA based SWOT framework as described in Section 7.2, chapter 7 in order to generate SWOT of KPS. Detail of the implementation of IPA based SWOT analysis on student satisfaction survey is described in the next section.

8.3 Implementation of IPA based SWOT analysis

This section describes how the student satisfaction survey was processed following step 2 - step 4 of IPA based SWOT framework as described in Section 7.2, chapter 7 in order to generate SWOT of KPS. Results produced from step 2 - step 4 of IPA based SWOT framework which are IPA matrices, strengths-weaknesses of the two departments, and SWOT of KPS are presented in Section 9.1, Chapter 9.

Regarding the step 2 of IPA based SWOT framework, data from the student satisfaction survey of KPS and BKN were analysed to create two IPA matrices. The main tasks within this step are calculating of *importance* and *performance* for individual factor related to the attribute of the department, and constructing the IPA matrix.

- Calculate *importance*. For each attribute, the MLR model was constructed using SPSS to obtain *importance* as a regression coefficient. Recall that there were six attributes of the department as shown in Table 8.2, hence six regression models were constructed corresponding to the number of attributes.

Each model has different set of independent variables which is the performance of factors (or questions) related to the attribute, but it has the same dependent variable which is the overall student satisfaction. For example, there were five independent variables involved in building the regression model of the attribute Academic Personnel namely “Teaching ability”, “Subject expertise”, “Friendliness”, “Availability”, and “Advice and support in learning” of teaching staff. All independent variables related to the six attributes are presented in Table 8.2.

All independent variables associated with attributes were entered into the model simultaneously as the forced entry is set for the method of regression. Once the regression models were trained, the *importance* was measured as the standardized regression coefficient of the model. Additionally, all *importance* was expressed as a percentage contribution of factor with the negative *importance* set to zero, which allows the *importance* measures of the KPS and BKN to be comparable. This approach has been used in some previous comparative studies of importance measures (Gustafsson and Johnson, 2004; Pokryshevskaya and Antipov, 2014).

Lastly, it is important to note that assumptions of MLR (such as normality, multicollinearity) tested on KPS and BKN surveys were checked and found not to be violated (see Appendix H and Appendix I). Thus, it can be assured that the MLR were properly conducted which results in reliable regression coefficients for *importance* obtained from the surveys of the two departments in the case study.

- Calculate *performance*. For each factor under the six attributes, the *performance* of KPS and BKN was calculated as a mean of satisfaction. The *performance* of the KPS and BKN were comparable since the satisfaction of the two departments was measured in the same scale of 5.

- Construct the IPA Matrix. The intersection between the x-axis and y-axis to generate a four-quadrant matrix was calculated as the grand mean of all *importance* and all *performance* of the KPS and BKN. The use of grand means provides a fair comparison between the IPA matrix of both departments. Then, the coordinates of *importance* and *performance* for each factor under the six attributes were plotted to generate complete IPA matrices of the two departments (see Table 9.1).

Regarding the step 3 of the IPA based SWOT framework, the IPA matrix of KPS and BKN were examined to identify strengths and weaknesses of the department. Factors located in Quadrant 1 and Quadrant 2 of IPA matrix are identified as strengths as their performance is higher than the grand mean of performance, whereas factors located in Quadrant 3 and Quadrant 4 of IPA matrix are identified as weaknesses as their performance is lower than the grand mean of performance.

Regarding the step 4 of the IPA based SWOT framework, the opportunities and threats of KPS were identified by comparing the factors of KPS and BKN that were previously labelled as strengths and weaknesses to generate a complete SWOT of KPS (see Table 9.2). This identification of opportunities and threats was conducted regarding the SWOT identification table presented in Table 7.1. In addition to the label of SWOT aspect, each factor also has the weight attached which was computed as the product of *importance* and *performance*, allowing the department to do factor prioritization within a SWOT group.

8.4 Methodology for conducting staff evaluation of SWOT survey

An IPA based SWOT analysis was completed based on student satisfaction toward services offered by KPS. The IPA-SWOT of KPS was then evaluated by staff at KPS, the staff all had worked at KPS for at least two years, hence they can be considered as domain experts. The aim of the evaluation was to answer sub research question 3.1, that is, do staff agree with the analysis. The staff's level of agreement on the IPA-SWOT of KPS provides face validity of the proposed approach. The evaluation also reveals how well the staff recognises organisation's competences given by their customers. Following the survey research method, tasks for undertaking staff evaluation of SWOT are described in the following sub-sections.

8.4.1 Questionnaire design

The staff evaluation of SWOT questionnaire was divided into three sections. The first and second section of the questionnaire comprised of closed questions asking a level of agreement of staff in the KPS towards the outcome of IPA based SWOT analysis which are strengths-weaknesses and opportunities-threats presented in Table 9.3. Each question was rated on a four-point Likert scale (1 = Completely disagree to 4 = Completely agree) without the midpoint that acts as a neutral option.

The even point scale is preferable to be used in this survey with regard to the reason that the author would prefer staff to make a definite choice whether they agree or disagree with the produced SWOT rather than choose the neutral option, in order to ensure that the evaluation result regarding SWOT output is valid. The use of an even point scale yields some advantages such as eliminating possible misinterpretation of the midpoint and revealing the inclination of the respondents in the middle.

Additionally, the use of an even point scale also reduces the social desirability bias as some respondents who actually lean toward a negative response but understated their standpoint by choosing the midpoint to avoid reporting what they perceive to be socially unacceptable (Garland, 1991). Specifically, the survey experiment by Garland (1991) showed that the absence of a midpoint has resulted in more negative ratings than were achieved when it was available. This result is consistent with the study of Johns (2005).

The last section of the questionnaire refers to the demographic data of KPS staff such as position, gender and number of working year. The printed questionnaire is shown in Appendix J.1.

8.4.2 Consistency of the translation between English and Thai version of questionnaire

Since the participants are Thai, the questionnaire was translated from English into Thai. The consistency of the translation between the two languages was checked by three Thai native-speakers who possess a good level of English in order to ensure that the Thai version of the questionnaire could be understood by participants while preserving the meaning of the original version.

In this step, three Thai native-speakers were asked individually to rate each statement or question on the degree of consistency between the two versions of languages by assigning a rating of +1 (consistency), 0 (undecided about whether Thai statement or question is consistent with the English version) or −1 (no consistency). They were also requested to give a comment to improve the statement or question in case that they assigned a rating of 0 or −1 to statement or question.

Subsequently, the total consistency score for each statement or question was calculated by adding up the rating of each Thai native-speaker. The results of the test of language consistency are shown in Appendix J.3. These results suggested that all Thai statement or question could be used, because all had the total consistency score greater or equal to 2.

8.4.3 Sample size estimation

To investigate whether staff agree or disagree with the outcome of IPA based SWOT analysis, the one-sample t-test was conducted to compare the level of agreement toward SWOT factor/ SWOT group with an acceptable threshold. The required minimum number of participants for the one-sample t-test computed by using G*Power was found to be 12. The major input parameters and their value for estimating sample size are:

- Statistical test: one-sample t-test
- Tail(s): Two

- Alpha error probability: 0.05 - normal convention
- Power: 0.8 - normal convention
- Effect size *d*: 0.9 - This represent a relatively large effect size for exploring whether the mean value of the level of agreement toward SWOT factor/ SWOT group is different from the threshold value.

In case that the assumption of one-sample t-test was violated, the Wilcoxon signed-rank test would be used instead of the one-sample t-test. The required minimum number of participants for the Wilcoxon signed-rank test computed by using G*Power was found to be 13. The major input parameters and their value for estimating sample size are:

- Statistical test: Wilcoxon signed-rank test (one sample case)
- Tail(s): Two
- Parent distribution: Normal
- Alpha error probability: 0.05 - normal convention
- Power: 0.8 - normal convention
- Effect size *d*: 0.9 - This represent a relatively large effect size for exploring whether the mean value of the level of agreement toward SWOT factor/ SWOT group is different from the threshold value.

8.4.4 Ethics

Staff evaluation of SWOT survey was conducted under ethics number ERGO/FPSE/ 18153 which was received approval of the Ethics Committee of the University of Southampton. This ethics number was presented in the participant information sheet (see detail in Appendix J.2) that distributed to all staff of KPS who took part in the survey.

8.4.5 Pilot study

Prior to the main study, the questionnaire was piloted among three KPS lecturers who are currently studying aboard. They were asked to read the participant information sheet and then responded to the online questionnaires available at: https://www.isurvey.soton.ac.uk/18170.

In addition to answering the questions in the three main sections of the questionnaire, participants were asked to complete an additional section of the questionnaire where they can give feedback about the questionnaire design and any other suggestions they had for improving the questionnaire. The participants' responses to this additional section of the questionnaire were shown in Table 8.5.

Table 8.5: Opinion regarding to the design of staff evaluation of SWOT questionnaire

Question	Yes (%)	No (%)	Suggestions/Comments
1. Are the instructions in Section I clear?	100.0	0.0	
2. Are the questions in Section I clear?	66.7	33.3	Please provide more description or example to make the questions more clear which will be increased an understanding of participant.
3. Are the instructions in Section II clear?	100.0	0.0	
4. Are the questions in Section II clear?	100.0	0.0	
5. Do you agree with the level of agreement used in the questionnaire?	33.3	66.7	It's quite difficult for the participants to respond if their real agreement is in the middle (neither agree or disagree). However, it is sensible to use this scale to collect data for a small group of sample.
6. Do you agree that definition of SWOT should be provided in the questionnaire?	66.7	33.3	
7. Are there any comments that you would like to make about the questionnaire?			Adding the definition of SWOT is the most important thing to do to improve the quality of questionnaire since most of staff in the department have different of understanding about SWOT.

Table 8.5 shows that there was no issue about the questions and instructions as they were clear, whereas issues about the scale of level of agreement were found. Although the participants preferred the odd point scale, an even point scale was still used in this study as the sample size was small and the author would prefer staff to make a definite choice whether they agree or disagree with the produced SWOT rather than choose the neutral option.

Apart from the issue about the scale, one suggestion for improving the questionnaire was made which was to provide the definition of SWOT to ensure that all participants of the main study have got the same understanding about SWOT. Based on this suggestion, the definition of SWOT was add to improve the questionnaire of the main study.

8.4.6 Data collection

KPS staff were asked to take part in the survey by e-mail during November 2015. They were asked to read the participant information sheet and then responded to the on-line questionnaires available at: https://www.isurvey.soton.ac.uk/18177. The printed versions is in Appendix J.1.

A total of 14 valid questionnaires were collected for analysis from total of 15 KPS staff which yield a 93.34% response rate. Note that the sample size was greater than 13 which is the minimum required sample size for data analysis using Wilcoxon signed-rank test computed by G*Power (see sub-section 8.4.3). Data characteristics of sample is shown in Table 8.6.

Table 8.6: General characteristics of sample CPE-KU-KPS staff (N=14)

Characteristics	**Number**	**Percentage (%)**
Gender		
Male	7	50.0
Female	7	50.0
Number of working years at CPE-KU-KPS		
< 1	0	0.0
1 - 3	3	21.4
4 - 6	6	42.9
7 - 10	5	35.7
> 10	0	0.0
Position		
Academic staff	9	64.3
Non-Academic staff	5	35.7

Table 8.6 lists general characteristics of sample KPS staff who participated in the survey. The sample of 14 staff comprise of 7 (50.0%) males and 7 (50.0%) females. The majority are staff with 4-6 working years which is about 42.9% of overall sample. Considering staff position, 9 (64.3%) are academic staff and 5 (35.7%) are non-academic staff.

8.4.7 Data analysis method

The tests whether staff agree or disagree with the outcome of IPA based SWOT analysis were conducted at individual and aggregate level. For the individual level, the level of agreement of KPS staff towards the produced SWOT factors which were presented in Table 9.3 (denoted as variable named SW_{1-15}, OT_{1-16}) were compared with an acceptable threshold. For the aggregate level, the average level of agreement of staff for each SWOT group (denoted as variable named S_{avg}, W_{avg}, O_{avg}, T_{avg}) was compared with an acceptable threshold.

The one-sample t-test was chosen to test the difference between the level of agreement of KPS staff towards outcome of IPA based SWOT analysis with an acceptable threshold. Subsequently, the assumptions of a one-sample t-test were tested in both individual and aggregate level of the outcome of IPA based SWOT analysis. The result of assumption testing of a one-sample t-test on each SWOT factor showed that most of variables SW_{1-15} and OT_{1-16} were violated the normality assumption (see Appendix K). On the contrary, the result of assumption testing of one-sample t-test on each SWOT group showed that all assumptions were met (see Appendix L).

Therefore, a non-parametric test named one-sample Wilcoxon signed rank was used instead of one-sample t-test to compare individual SWOT factor with an acceptable threshold of 3.0 out of 4.0. This threshold was set as staff who agreed with the produced SWOT would give a rating of 3 or above, the null and alternative hypotheses can be stated as follows:

$\mathbf{H_0}$: The median response is 3.0

$\mathbf{H_a}$: The median response is not 3.0

Since the assumptions were met, the one-sample t-test was used to compare the average level of agreement for each SWOT group with an acceptable threshold of 2.5 out of 4.0. This threshold was set according to the fact that the negative response was increased when using an even point scale as reported by Garland (1991), thus if staff mostly disagreed with the produced SWOT, the average of staff level agreement should be lower than 2.5. The null and alternative hypotheses can be stated as follows:

$\mathbf{H_0}$: The mean response is equal to 2.5

$\mathbf{H_a}$: The mean response is not equal to 2.5

In summary, the level of agreement of each SWOT factor was tested individually whether staff agreed or disagreed with it using a Wilcoxon signed rank since one assumption of one-sample t-test was violet. While, the aggregate level of SWOT factor as the average level of agreement for each SWOT group was tested whether staff agreed or disagreed with it using a one-sample t-test. Hypotheses of both tests were tested at the 5% significance level and their results are presented in Section 9.2, Chapter 9.

8.5 Methodology for conducting experienced users of SWOT survey

The aim for the development of this survey is to address sub research question 3.2 by assessing the outcome of SWOT produced by two approaches (1) traditional SWOT analysis through the brainstorming session (2) IPA based SWOT analysis. The traditional SWOT analysis was selected as the comparator of IPA based SWOT analysis as the author intended to compare the SWOT approach that is usually conducted by the target organisation of the case study with the proposed SWOT approach. Additionally, traditional SWOT analysis is practically in used SWOT approach as it is less complicated among the SWOT approaches. It is expected that the comparison between these two SWOT approaches will point out the differences between the currently used approach and the newly introduced approach of SWOT analysis.

8.5.1 Questionnaire design

The two questionnaires of experienced users of SWOT survey were divided into three sections. The first section provides information about the development of SWOT and consists of steps for generating SWOT and the produced SWOT factors. The second section provides a list of questions asking participants to assess the quality and perceived usefulness of SWOT presented in the first section of questionnaire. The last section of the questionnaire refers to the demographic data of MBA students such as gender, occupation and experience in using SWOT. The printed version of the two questionnaires are shown in Appendix M.1 (set A: IPA based SWOT analysis) and Appendix M.2 (set B: traditional SWOT analysis).

The differences between the two questionnaires are in the first and the second section of the questionnaire. In the first section, one describes information about the development of IPA based SWOT analysis (denoted as set A) whereas the other describes information about the development of traditional SWOT analysis. Specifically, the step for generating SWOT shown in questionnaire set A is described following the IPA based SWOT framework (see Section 7.2, Chapter 7) and the produced SWOT factors presented in questionnaire set A were obtained from the SWOT of KPS presented in Table 9.3. While, the step for generating SWOT shown in questionnaire set B is described following the traditional approach for conducting SWOT analysis and the produced SWOT factors presented in questionnaire set B were obtained from the brainstorming session of KPS staff.

In the second section, questionnaire set A comprises of 12 questions whereas questionnaire set B comprises of 10 questions. The first 10 questions are similar between questionnaire set A and set B. Each question was rated on a five-point Likert scale and its response was used to compare the difference of two SWOT approaches. The extra two questions of questionnaire set A were designed to ask participants about their opinion toward the IPA based SWOT analysis.

Some of the 10 questions used in both questionnaires were adapted from a few research studies that were intended to examine the effectiveness of SWOT analysis (Hill and Westbrook, 1997; Coman and Ronen, 2009; Pai et al., 2013) and some were created by the author to fulfil the criteria for assessing SWOT. The 10 questions are shown in Table 8.7. Question number 1-4 were designed for checking common flaws of the SWOT analysis outcome. Question number 5 and questions number 6-7 were designed for checking the comprehensibility and measurability of SWOT analysis outcome respectively. Question number 8 was designed for checking that the SWOT factors were defined based on a reasonable data source. And the last two questions were designed to check that the SWOT analysis outcome is actionable.

Table 8.7: List of questions in the experienced users of SWOT survey

Questions	Source
1. Strengths and weakness are explicit (clearly and unambiguously formulated).	Hill and Westbrook (1997)
2. The opportunities and the threats are explicit (clearly and unambiguously formulate).	Hill and Westbrook (1997)
3. All SWOT factors are not overgeneralisation.	Hill and Westbrook (1997)
4. The strengths and weakness, and the opportunities and threats are correctly classified as internal/external factors.	Hill and Westbrook (1997)
5. The SWOT factors comprehensively explain the department's situation.	Pai et al. (2013)
6. All SWOT factors are measurable.	
7. The order of SWOT factors makes decision-making easier.	
8. The data source for this SWOT analysis is reliable.	Pai et al. (2013)
9. The SWOT factors can be used as a starting point for strategic planning.	Coman and Ronen (2009)
10. The SWOT factors provide useful information that supports decision-making regarding strategic planning.	Pai et al. (2013)

8.5.2 Consistency of the translation between English and Thai version of questionnaire

Since the participants are Thai, the questionnaire was translated from English into Thai. The consistency of the translation between the two languages was checked by three Thai native-speakers who possess a good level of English in order to ensure that the Thai version of questionnaire could be understood by participants while preserving the meaning of the original version.

In this step, three Thai native-speakers were asked individually to rate each statement or question on the degree of consistency between the two versions of languages by assigning a rating of +1 (consistency), 0 (undecided about whether Thai statement or question is consistent with the English version) or −1 (no consistency). They were also requested to give a comment to improve the statement or question in case that they assigned a rating of 0 or -1 to statement or question.

Subsequently, the total consistency score for each statement or question was calculated by adding up the rating of each Thai native-speaker. The results of the test of language consistency are shown in Appendix M.4. These results suggested that most of Thai statements or question could be used as they had the total consistency score greater or equal to 2. However, a few of Thai statements needed some improvement as the total consistency score was less than 2 such as statement number 4.4-4.6 in Appendix M.4. The changes have been made for these Thai statements to create a questionnaire for the pilot study.

8.5.3 Sample size estimation

To investigate whether any significant difference exists between the rating score of traditional SWOT analysis and the IPA based SWOT analysis, the two independent sample t-test was conducted. The required minimum number of participants for the two independent sample t-test computed by using G*Power was found to be 42, which divided into two groups (21 participants in each group). The major input parameters and their value for estimating sample size are:

- Statistical test: two independent sample t-test
- Tail(s): Two
- Alpha error probability: 0.05 - normal convention
- Power: 0.8 - normal convention
- Effect size d: 0.9 - This represent a relatively large effect size for exploring whether the mean value of rating is different between two groups.

In case that the assumption of two independent sample t-test was violated, the Mann-Whitney U test would be used instead of the two independent sample t-test. The required minimum number of participants for the Mann-Whitney U test computed by using G*Power was found to be 44, which divided into two groups (22 participants in each group). The major input parameters and their value for estimating sample size are:

- Statistical test: Mann-Whitney U test (two groups)
- Tail(s): Two
- Parent distribution: Normal

- Alpha error probability: 0.05 - normal convention
- Power: 0.8 - normal convention
- Effect size *d*: 0.9 - This represent a relatively large effect size for exploring whether the mean value of rating is different between two groups.

8.5.4 Ethics

Experienced users of SWOT survey was conducted under ethics number ERGO/FPSE/ 18500 which received approval of the Ethics Committee of the University of Southampton. This ethics number was presented in the participant information sheet (see detail in Appendix M.3) that was distributed to all MBA students who took part in the survey.

8.5.5 Pilot study

After improving the consistency of the translation, the questions for evaluating quality of SWOT were examined for face validity by two experts who have experience in teaching and conducting SWOT analysis. It was found that the questions presented in this questionnaire appeared to be a valid measure for the SWOT quality. However, both of the experts raised their concern about the misclassification of SWOT factors in the case study.

Additional to the face validity, the questionnaire was piloted among four MBA students prior to the main study. They were asked to read the participant information sheet and then responded to the pilot questionnaires. Specifically, participants of the pilot study were divided into two groups of two participants to assess the two different SWOT approaches: IPA based SWOT analysis and traditional SWOT analysis.

Two groups of participants were asked to complete the three main sections of the questionnaire. Then, they were also asked to complete an additional section of the questionnaire where they can give feedback about the questionnaire design and any other suggestions they had for improving the questionnaire. The participants' responses to this additional section of the questionnaire are shown in Table 8.8.

Based on results of the pilot study shown in Table 8.8, the instruction of the introduction part of survey was rewritten and five of the SWOT factors presented in questionnaire set B were cut-down by KPS staff to improve the questionnaire of the main study.

8.5.6 Data collection

MBA students were asked face to face to take part in the survey by the representative of researcher in their classroom during December 2015. They were divided into two equal

Table 8.8: Opinion regarding to the design of questionnaire of experienced SWOT user

Question	Yes (%)	No (%)	Suggestions/Comments
1. Are the instructions in the introduction part of survey clear?	50.0	50.0	
2. Is the case study described in Section 1 clear?	50.0	50.0	
3. Are the case study, Figure and Table presented in Section 1 providing information for answering questions in Section 2?	75.0	25.0	
4. Is Figure 1 consistent with the case study and make it easier to understand the case study?	100.0	0.0	
5. Are the questions in Section 2 clear?	100.0	0.0	
6. Do you agree with the level of agreement used in the survey?	100.0	0.0	
7. Are the questions in Section 3 clear?	100.0	0.0	
8. Are there any comments that you would like to make about the questionnaire?	- It's take long time to complete the questionnaire due to the difficulty of questionnaire. This could be affect the accuracy of the questionnaire - There are too many SWOT factors in Table 1, suggested to reduce number of SWOT factors		

independent groups of 22 participants. The first group of the participants were asked to evaluate IPA based SWOT analysis through the questionnaire set A (Appendix M. 1). The second group of participants were asked to evaluate traditional SWOT analysis through the questionnaire set B (Appendix M.2).

Note that the sample size of 22 was equal to the minimum required sample size for data analysis using Mann-Whitney U test computed by G*Power (see sub-section 8.5.3). Data characteristics of the sample are shown in Table 8.9.

Table 8.9 lists general characteristics of MBA students who participated in the survey. For IPA based SWOT group, the sample of 22 MBA students comprise 3 (13.6%) males and 19 (86.4%) females. Half of the sample are full-time students and the other half are business owners (18.2%), others occupations included nurse, banker (13.6%) etc. 95.5% of the overall sample possess experience in using SWOT analysis in academia while 18.2% of the overall sample possess 1-4 years practical experience in using SWOT analysis. For the traditional SWOT group, the sample of 22 MBA students comprise of 6 (27.3%) males and 16 (72.7%) females. The majority are full-time students and bankers, and are about 36.4% and 31.8% of the overall sample respectively. All of the

Table 8.9: General characteristics of two groups of MBA students (N=22)

Charateristic	**IPA based SWOT**		**Traditional SWOT**	
	Number	**Percentage**	**Number**	**Percentage**
Gender				
Male	3	13.6	6	27.3
Female	19	86.4	16	72.7
Age				
21 - 25	6	27.3	7	31.8
26 - 30	14	63.6	12	54.5
31 - 35	2	9.1	3	13.6
> 35	0	0.0	0	0.0
Occupation				
Postgraduate student	11	50.0	8	36.4
Business owner	4	18.2	3	13.6
Government Officer	1	4.5	0	0.0
Staff/ Sale Manager	1	4.5	0	0.0
Staff/ Manager of Human Resources	1	4.5	1	4.5
Staff/ Manager of Marketing	1	4.5	2	9.1
Consultant	0	0.0	1	4.5
Other.	3	13.6	7	31.8
Experience in using SWOT analysis in academia				
No	0	0.0	0	0.0
Yes	21	95.5	22	100.0
Missing value	1	4.5	0	0.0
Practical experience in using SWOT analysis				
No	17	77.3	20	90.9
Yes	4	18.2	2	9.1
Missing value	1	4.5	0	0.0

sample possess experience in using SWOT analysis in academia while 9.1% of the overall sample possess 1-2 years practical experience in using SWOT analysis.

8.5.7 Data analysis method

To examine whether the IPA based SWOT analysis yielded a better quality of SWOT than the traditional SWOT analysis, the response of the MBA students on those two questionnaires were analysed by using the two independent sample t-test to figure out whether any significant difference exists between the mean ratings of two groups. Subsequently, the assumptions of the two independent sample t-test were tested and found that all variables represented the questions violated the normality assumption and some variables violated the assumption of homogeneity of variance (see Appendix N).

Therefore, a non-parametric test named Mann-Whitney U test was used instead of the two independent sample t-test to determine if there were statistically significant differences in the mean ranks of the rating score in terms of the two groups. Although the Mann-Whitney U test does not require any assumptions related to the distribution, there are two important assumptions that should be checked (Pallant, 2005). First, the two sample groups drawn from the population are random and second, the two sample group are independent of each other.

These two assumptions can be observed regarding the survey design. The participants were randomly selected from all MBA students at Faculty of Management Science, Silpakorn University, Thailand. They were also randomly assigned to each of the two groups. Each participant can not appear in more than one group and the behaviour of one participant does not influence the behaviour of another. Thereby, ratings from different participants are independent. These assumptions therefore were met.

In order to run a Mann-Whitney U test, rating for each question was set as the dependent variable and type of SWOT approach which is split into two groups: IPA based SWOT and Traditional SWOT was the independent variable. The null and alternative hypotheses of Mann-Whitney U test can be stated as follows:

$\mathbf{H_0}$: The distribution of ratings for the two groups are equal

$\mathbf{H_a}$: The distribution of ratings for the two groups are not equal or the mean ranks of the two groups are not equal.

Hypotheses of this test were tested at the 5% significance level and their results are presented in Section 9.3, Chapter 9.

8.6 Summary

This chapter describes details of the methodologies related to the case study which was conducted to demonstrate and evaluate the IPA based SWOT analysis. To demonstrate how the IPA based SWOT analysis can be used to generate SWOT, first the student satisfaction survey was undertaken at KPS and BKN following the survey research method. The data collected from the student satisfaction survey was then processed to generate SWOT of KPS regarding to the IPA based SWOT analysis framework. Through the framework, the student satisfaction survey data was analysed using multiple linear regression for calculating *importance* and computed the average as *performance* of each SWOT factor which generated the IPA matrix of KPS and BKN. These IPA matrices were then compared for classifying SWOT factors into four aspects of SWOT.

The SWOT of KPS was then evaluated by using the staff evaluation of SWOT survey and the experienced users of SWOT survey to address the third research question. Since, there are no direct methods and tools for validating the effectiveness of SWOT analysis, the first survey aims to evaluate the SWOT from the viewpoint of the domain users of SWOT in the case study which are the KPS staff. The participants of this survey were asked to rate their level of agreement towards the SWOT of KPS. To test whether staff agreed or disagreed with the SWOT of KPS, the data collected through this survey was analysed at individual and aggregate level. For the individual level, level of agreement for each SWOT factor of KPS was compared with an acceptable threshold using a one-sample Wilcoxon signed rank instead of one-sample t-test since the normality assumption of t-test was not met. For the aggregate level, SWOT factors of KPS were grouped into four aspects of SWOT and the average level of agreement for each group was computed. Subsequently, the one-sample t-test was used to compare the average level of agreement for each SWOT group with an acceptable threshold.

The latter survey aims to evaluate the quality of SWOT from the viewpoint of experienced users of SWOT. The MBA students who possess related knowledge or real-world experiences in using SWOT analysis were invited to take part in this survey. To evaluate how well the IPA based SWOT performs two versions of questionnaire were developed which are IPA based SWOT version and traditional SWOT version. The participants were divided equally into two groups. Participants in each group were asked to examine the SWOT presented in each version of the questionnaire and respond to the questions related to the quality of SWOT. The Mann-Whitney U test was used instead of the two independent sample t-test to compare rating of individual questions between these two groups since the normality assumption of t-test was not met.

The implementation of these methodologies produced SWOT of KPS, analytical results of the one-sample Wilcoxon signed rank test and the one-sample t-test, and analytical results of the Mann-Whitney U test which are presented together with the results interpretation in Chapter 9.

Chapter 9

IPA based SWOT analysis: Evaluation Results

By implementing the methodologies described in the previous chapter, SWOT of KPS and analytical results of SWOT evaluation were generated and are reported in this chapter. The aim of this chapter is to present SWOT of KPS which was generated following the IPA based SWOT analysis framework and describes the analytical results of the two surveys each of which is related to one of the sub research questions of the third research question.

The chapter begins with an illustration of SWOT of KPS in Section 9.1 followed by analytical results of staff evaluation on the SWOT of KPS presented in Section 9.2. Sections 9.3 reports analytical results of experienced users of SWOT analysis evaluation on the SWOT of KPS. The last section gives a summary of the chapter.

9.1 Results of IPA based SWOT analysis of the case study

The student satisfaction survey of KPS and BKN collected through the methodology described in Section 8.2 was analysed through the methodology described in Section 8.3 which produced the IPA matrix of the two departments and the SWOT of KPS. The detail of the two IPA matrices and SWOT of KPS is shown in the following sub-sections.

9.1.1 IPA matrix of the two departments

The IPA matrix of KPS and BKN was created individually by analysing data from the student satisfaction survey of the departments. Through the main tasks for constructing the IPA matrix comprised calculation of *importance* and *performance* for individual factor related to the attribute of the department, the two IPA matrices are presented in Table 9.1 and visually presented Figure 9.1-9.2.

The first column of this table presented a short name of factors and their description is provided in Table 8.2. These short names of factors are also used in the other tables of this chapter. The other columns presented *importance* (I), *performance* (P) and IPA quadrant (Q) of KPS and BKN respectively.

Table 9.1: Importance-Performance of the two departments. Each row represents *importance*, *performance* and IPA quadrant of the two departments. The second last row represents the mean of *importance* and *performance* for each department in which their overall mean (Grand Mean) is represented in the last row.

	CPE-KU-KPS			**CPE-KU-BKN**		
Factor	I	P	Q	I	P	Q
Teacher1	4.477	4.335	1	1.362	4.000	2
Teacher2	5.195	4.452	1	5.284	4.395	1
Teacher3	3.043	4.387	2	3.940	4.349	1
Teacher4	2.565	4.090	2	3.795	3.767	4
Teacher5	1.934	4.232	2	0.000	3.791	2
Teaching1	5.847	3.987	1	10.496	3.605	4
Teaching2	3.086	4.006	2	0.000	3.512	3
Teaching3	1.521	3.929	2	0.163	3.395	3
Teaching4	0.739	4.135	2	0.000	4.093	2
Teaching5	0.000	4.129	2	3.977	3.605	4
Teaching6	7.020	3.761	4	3.723	4.116	1
Admin1	0.000	3.877	2	0.708	4.000	2
Admin2	9.346	3.916	1	5.538	3.860	1
Admin3	0.000	3.761	3	0.000	4.070	2
Admin4	2.978	3.748	3	0.000	4.279	2
Admin5	2.652	3.781	3	8.989	4.116	1
CompFac1	4.043	3.697	4	0.000	3.674	3
CompFac2	5.868	3.684	4	6.955	3.651	4
CompFac3	0.000	3.568	3	3.033	3.884	2
CompFac4	5.151	3.426	4	3.850	2.698	4
xActivity1	1.326	3.716	3	4.739	3.535	4
xActivity2	1.739	3.819	2	7.009	3.233	4
xActivity3	2.760	3.852	2	0.000	3.302	3
xActivity4	4.608	3.774	4	0.000	3.209	3
xActivity5	0.630	3.916	2	8.407	3.674	4
xActivity6	6.390	3.961	1	0.399	3.767	3
xActivity7	0.000	3.916	2	2.451	3.209	3
AddService1	0.826	3.748	3	0.000	3.581	3
AddService2	5.803	3.729	4	5.266	3.605	4
AddService3	2.499	3.839	2	9.733	3.116	4
AddService4	7.955	3.729	4	0.182	2.930	3
Mean	3.226	3.900		3.226	3.678	
Grand Mean	**3.226**	**3.789**				

The red horizontal line and red vertical line in Figure 9.1-9.2 represent Grand Mean presented in Table 9.1, which is computed as overall mean of *importance* and *performance* of the two departments. The intersection of these lines created a four quadrants IPA matrix labelled as “Q1 - Keep up the good work”, “Q2 - Possible overkill”, “Q3 - Low priority”and “Q4 - Concentrate here”.

Based on IPA matrix of KPS in Figure 9.1, the majority of factors are located in Quadrant 2 and which are 13 out of 31 factors implying that resources committed to these factors would be better used in other quadrants in need of improved performance. There were seven factors located in Quadrant 4: Teaching6, ComFac1, ComFac2, ComFac4, xActivity4, AddService2, and AddService4 that require immediate attention for improvement to meet student satisfaction. There were five factors located in Quadrant 1: Teacher1, Teacher2, Teaching1, Admin2, and xActivity6 which are major factor for improving student satisfaction.

Based on IPA matrix of BKN in Figure 9.2, the majority factors are located in Quadrant 4 which are 10 out of 31 factors and they were competitive disadvantages for improving student satisfaction of BKN. Quadrant 1 located the 5 factors for improving student satisfaction of this department which were: Teacher2, Teacher3, Teaching6, Admin2, and Admin5. The two IPA matrices have a similar number of factors located in Quadrant 1 which were competitive advantages for improving student satisfaction but different numbers of factors located in Quadrant 4 which were competitive disadvantage for improving student satisfaction. KPS has a fewer number of factors that need to be concentrated on for improving student satisfaction than BKN. This indicates a sign that KPS may have a competitive advantage over BKN.

9.1.2 SWOT matrix of CPE-KU-KPS

The results of IPA shown in Table 9.1 is then interpreted as strengths and weaknesses. Factors located in Quadrant 1 and Quadrant 2 were identified as strengths as their performance was higher than the grand mean of performance, whereas factors located in Quadrant 3 and Quadrant 4 were identified as weaknesses as their performance was lower than the grand mean of performance. The identification of strength and weakness is shown in the second and third column of Table 9.2.

By comparing the factors of KPS and BKN that were previously labelled as strength (S) and weakness (W) in Table 9.2, opportunities (O) and threats (T) of KPS were then identified regarding the SWOT identification table presented in Table 7.1. Finally, the SWOT of KPS with factor weight is presented in the fourth and fifth column of Table 9.2.

Note that the factor weights were computed as the product of *importance* and *performance* of KPS shown in Table 9.1. Specifically, a positive value of *performance* was

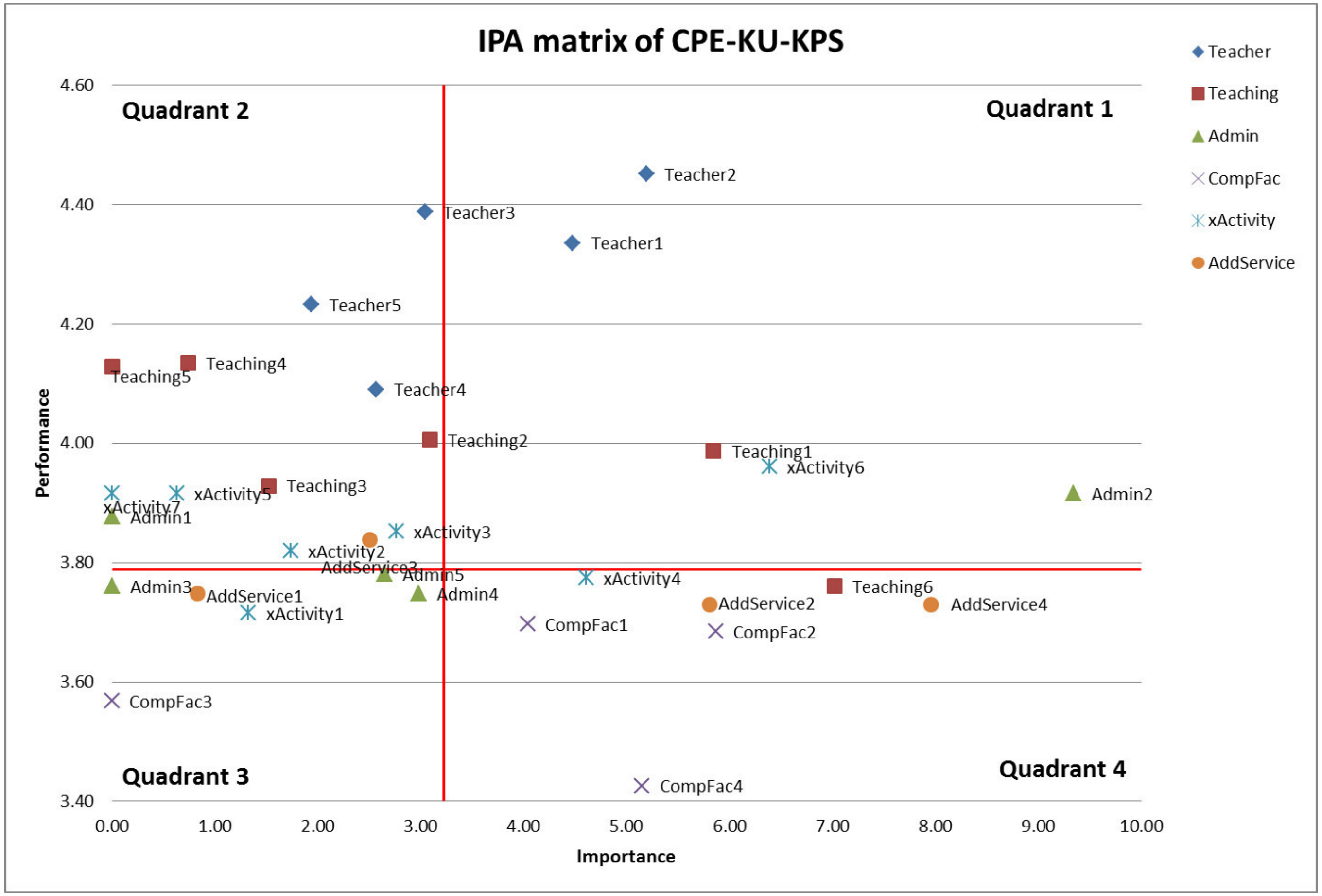


Figure 9.1: IPA matrix of CPE-KU-KPS. Each point represents a pair of *importance* and *performance* for each factor presented in the second and third column of Table 9.1. The IPA matrix is separated into four quadrants regarding to the intersection of the Grand Mean of *importance* and *performance* which represented as red horizontal line and red vertical line.

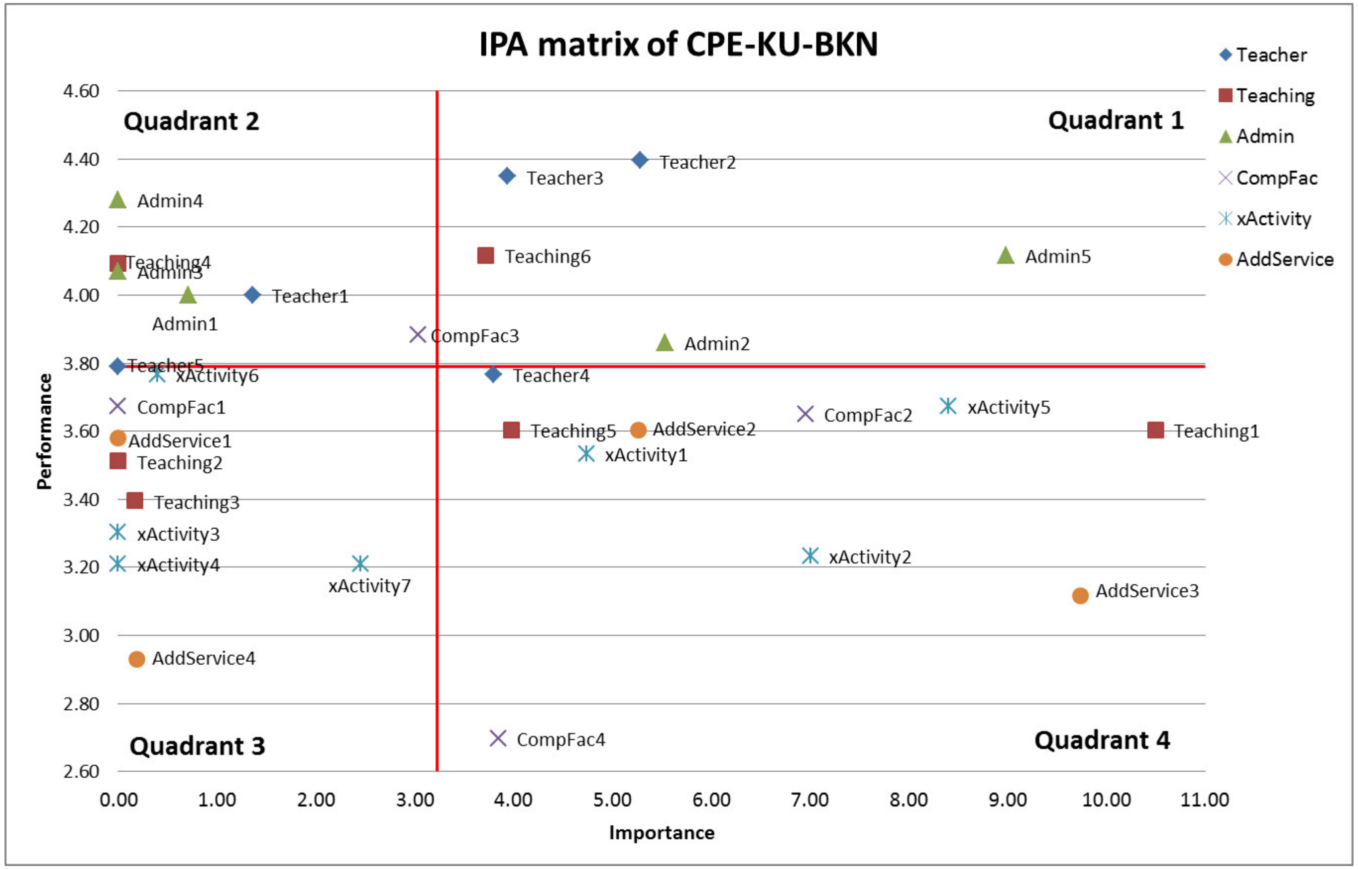


Figure 9.2: IPA matrix of CPE-KU-BKN. Each point represents a pair of *importance* and *performance* for each factor presented in the fifth and sixth column of Table 9.1. The IPA matrix is separated into four quadrants regarding to the intersection of the Grand Mean of *importance* and *performance* which represented as red horizontal line and red vertical line.

assigned to the strength and opportunity factors on the other hand a negative value of *performance* was assigned to the weakness and threat factors.

Table 9.2: Result of IPA based SWOT. Each row represents the factor which labelled as strength-weakness of the two departments, SWOT of CPE-KU-KPS, and its weight. For each factor, its weight is computed as the product of *importance* and *performance* of CPE-KU-KPS shown in Table 9.1

	Strength - Weakness		**SWOT**	**Weight**
Factor	CPE-KU-KPS	CPE-KU-BKN	**CPE-KU-KPS**	
Teacher1	S	S	S	19.411
Teacher2	S	S	S	23.124
Teacher3	S	S	S	13.349
Teacher4	S	W	O	10.490
Teacher5	S	S	S	8.187
Teaching1	S	W	O	23.310
Teaching2	S	W	O	12.364
Teaching3	S	W	O	5.978
Teaching4	S	S	S	3.056
Teaching5	S	W	O	0.000
Teaching6	W	S	T	-26.403
Admin1	S	S	S	0.000
Admin2	S	S	S	36.598
Admin3	W	S	T	0.000
Admin4	W	S	T	-11.160
Admin5	W	S	T	-10.026
CompFac1	W	W	W	-14.945
CompFac2	W	W	W	-21.619
CompFac3	W	S	T	0.000
CompFac4	W	W	W	-17.648
xActivity1	W	W	W	-4.927
xActivity2	S	W	O	6.640
xActivity3	S	W	O	10.633
xActivity4	W	W	W	-17.389
xActivity5	S	W	O	2.468
xActivity6	S	W	O	25.310
xActivity7	S	W	O	0.000
AddService1	W	W	W	-3.096
AddService2	W	W	W	-21.640
AddService3	S	W	O	9.595
AddService4	W	W	W	-29.663

The SWOT of KPS in Table 9.2 was also represented as a SWOT matrix and shown in Table 9.3. Within a SWOT group, factors are prioritized by their weight. For strength and opportunity, the factors with highest priority for maintaining or improving are "Admin2 - Knowledge of the information about courses, exams, activities of staff members" and "xActivity6 - Personal learning and thinking skills development activities" respectively. For weakness and threat, the factors with highest priority for improving are

"AddService4 - Library" and "Teaching6 - Teaching facilities and classroom condition" respectively.

In total, there are seven strengths, eight weaknesses, 11 opportunities and five threats of KPS that were identified based on the student satisfaction survey. The strengths are mainly in *Academic Personnel* as four out of five factors of *Academic Personnel* were identified as strengths. Most of the weaknesses are factors related to *Computer Facility* and *Additional Service.* Factors of *Teaching and Learning*, and *Extra-curricular activity* are mostly defined as opportunity where as factors of *Administration* are mostly defined as threat. It can be simply stated that KPS has *Academic personnel* strength and has great opportunity in *Teaching and Learning*, and *Extra-curricular activity.* The weakness of KPS is mainly related to *Computer Facility* and *Additional Service* and a threat of the department is related to *Administration.*

Table 9.3: SWOT matrix of CPE-KU-KPS

Strength		**Weakness**	
Factor	Weight	Factor	Weight
Admin2	36.598	AddService4	-29.663
Teacher2	23.124	AddService2	-21.640
Teacher1	19.411	CompFac2	-21.619
Teacher3	13.349	CompFac4	-17.648
Teacher5	8.187	xActivity4	-17.389
Teaching4	3.056	CompFac1	-14.945
Admin1	0.000	xActivity1	-4.927
		AddService1	-3.096
Opportunity		**Threat**	
Factor	Weight	Factor	Weight
xActivity6	25.310	Teaching6	-26.403
Teaching1	23.310	Admin4	-11.160
Teaching2	12.364	Admin5	-10.026
xActivity3	10.633	Admin3	0.000
Teacher4	10.490	CompFac3	0.000
AddService3	9.595		
xActivity2	6.640		
Teaching3	5.978		
xActivity5	2.468		
Teaching5	0.000		
xActivity7	0.000		

9.2 Results of staff evaluation on the SWOT of the case study

The staff evaluation of SWOT survey collected through the methodology described in Section 8.4 was analysed at both individual and aggregate level to test whether KPS staff agree or disagree with the SWOT of KPS presented in Table 9.3.

For the individual level, the level of agreement of KPS staff towards each SWOT factor was compared with a threshold of 3.0 using a one-sample Wilcoxon signed rank test. For the aggregate level, the average level of agreement of staff for each SWOT group was compared with a threshold of 2.5 using the one-sample t-test. Further details of the analytical method were provided in Sub-section 8.4.7 and the analytical results are shown in the following sub-sections.

9.2.1 Analytical results of staff evaluation on individual SWOT factor

Recall that there are 31 variables representing the level of agreement of KPS staff towards individual SWOT factor which were divided into 15 variables for strengths-weaknesses and 16 for opportunities-threats (denoted as variable named SW_{1-15}, OT_{1-16}). The hypothesis using the one-sample Wilcoxon signed rank was tested at the 5% significance level. If the p-value is less than 0.05 then the median is significantly different from 3.0. The results are shown in Table 9.4 and 9.6 for strength-weakness and opportunity-threat factors respectively.

Table 9.4: Result of one-sample Wilcoxon signed rank test for variable SW_{1-15}

Null Hypothesis	**Median**	**Sig. (2-tailed)**	**Decision regarding the null hypothesis**
The median of SW_1 equals 3.0	3.00	1.000	Retain
The median of SW_2 equals 3.0	3.00	0.180	Retain
The median of SW_3 equals 3.0	3.00	0.102	Retain
The median of SW_4 equals 3.0	3.00	0.083	Retain
The median of SW_5 equals 3.0	3.00	0.107	Retain
The median of SW_6 equals 3.0	2.00	0.003	**Reject**
The median of SW_7 equals 3.0	3.00	0.480	Retain
The median of SW_8 equals 3.0	2.50	0.096	Retain
The median of SW_9 equals 3.0	2.50	0.034	**Reject**
The median of SW_{10} equals 3.0	2.00	0.004	**Reject**
The median of SW_{11} equals 3.0	3.00	0.058	Retain
The median of SW_{12} equals 3.0	2.00	0.002	**Reject**
The median of SW_{13} equals 3.0	2.00	0.001	**Reject**
The median of SW_{14} equals 3.0	2.00	0.002	**Reject**
The median of SW_{15} equals 3.0	2.00	0.004	**Reject**

According to Table 9.4, there were 7 out of 15 variables represented strengths-weaknesses that have the median of staff agreement level statistically significantly different from 3.0. Among seven variables with their level of agreement of KPS staff lower than 3.0, six of them are weaknesses. Detail of these variables is shown in Table 9.5.

With regard to Table 9.4 and Table 9.5, it is clearly shown that staff agreed with more strength factors than weakness factors. As 6 out of 8 of weakness factors were not agreed by KPS staff whereas 1 out of 7 of strength factors were not agreed by KPS staff.

According to Table 9.6, 3 out of 16 variables represented opportunity factors that have the median of staff agreement level statistically significantly different from 3.0 namely

Table 9.5: List of strength-weakness variables that their level of staff agreement is lower than 3.0

Variable	Factor	Detail
SW_6	Admin1	Knowledge of rules and procedures of non-academic staff members
SW_9	CompFac2	Lack of availability of computer facilities for students
SW_{10}	CompFac4	Lack of availability of printing and photocopying facilities
SW_{12}	xActivity4	Poor arrangement of health development activities
SW_{13}	AddService1	Lack of financial aid provided for students
SW_{14}	AddService2	Lack of medical support provided for students
SW_{15}	AddService4	Lack of availability of library facilities

Table 9.6: Result of one-sample Wilcoxon signed rank test for variable OT_{1-16}

Null Hypothesis	Median	Sig. (2-tailed)	Decision regarding the null hypothesis
The median of OT_1 equals 3.0	3.00	0.655	Retain
The median of OT_2 equals 3.0	3.00	1.000	Retain
The median of OT_3 equals 3.0	3.00	0.014	**Reject**
The median of OT_4 equals 3.0	3.00	0.414	Retain
The median of OT_5 equals 3.0	3.50	0.008	**Reject**
The median of OT_6 equals 3.0	3.00	1.000	Retain
The median of OT_7 equals 3.0	3.00	0.317	Retain
The median of OT_8 equals 3.0	3.00	0.257	Retain
The median of OT_9 equals 3.0	3.00	0.102	Retain
The median of OT_{10} equals 3.0	3.00	0.414	Retain
The median of OT_{11} equals 3.0	3.00	0.180	Retain
The median of OT_{12} equals 3.0	3.00	0.025	**Reject**
The median of OT_{13} equals 3.0	3.00	0.564	Retain
The median of OT_{14} equals 3.0	3.00	0.083	Retain
The median of OT_{15} equals 3.0	3.00	0.083	Retain
The median of OT_{16} equals 3.0	3.00	1.000	Retain

OT_3, OT_5 and OT_{12}. For OT_3 and OT_{12}, the median of staff agreement level were lower than 3.0 whereas, for OT_5 the opposite was true: the median of staff agreement level were higher than 3.0. Detail of these variables is shown Table 9.7.

Table 9.7: List of opportunity variables that their level of staff agreement is different from 3.0

Variable	Factor	Detail
OT_3	Teaching2	Increasing E-learning resources to support student learning
OT_5	Teaching5	Increasing accurate and up-to-date course unit content
OT_{12}	xActivity3	Improving the arrangement of moral development activities

With regard to Table 9.6 and Table 9.7, KPS staff positively agreed with OT_5 and staff agreed with other variables except OT_3 and OT_{12}. KPS staff also agreed with all five variables of threat factors namely OT_6 - OT_{10}.

In summary, for each SWOT group only weakness factors have a percentage of agreement by KPS staff less than 50% while the other three groups have percentage of agreement by KPS staff up to 80% as shown in Table 9.8. It can be reasonably concluded that 3 out of 4 aspects of KPS SWOT have a high face validity.

Table 9.8: Summary of agreed SWOT factors

Group of SWOT	Variable	Number of factors	Number of agreed factors	Percentage of agreed factors
Strength	SW_1 - SW_7	7	6	85.71%
Weakness	SW_8 - SW_{15}	8	2	25.00%
Opportunity	OT_1 - OT_5, OT_{11} - OT_{16}	11	9	81.82%
Threat	OT_6 - OT_{10}	5	5	100.00%

9.2.2 Analytical results of staff evaluation on the group of SWOT factor

Recall that there are four variables named S_{avg}, W_{avg}, O_{avg}, T_{avg} representing the average level of agreement of staff for each SWOT group. The hypothesis using the one-sample t-test was tested at the 5% significance level. If the p-value is less than 0.05 then the mean response is significantly different from 2.5. The descriptive statistic of variables and the result of the one-sample t-test is shown in Table 9.9. Additionally, the means of response across the four variables is visually presented as simple error bars in Figure 9.3.

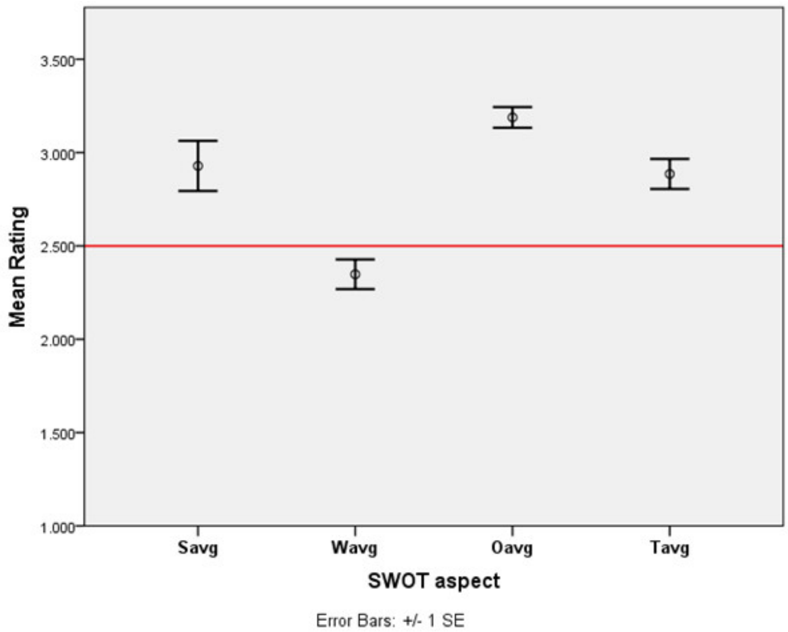


Figure 9.3: Simple error bar of the four variables of each SWOT group. The mean is represented by a dot, a line represents standard error of the mean, a red horizontal line is represented the threshold score of 2.5.

With regards to Table 9.9, average score of strength factors (M = 2.93, SD = 0.36) is higher than the threshold score of 2.5, a statistically significant mean difference of 0.43,

Table 9.9: Result of the One-sample t-test for each SWOT group (Test value is 2.5)

Variable	Mean	SD	Mean Difference	t	df	p (2-tailed)
S_{avg}	2.93	0.355	0.43	3.196	6	0.019
W_{avg}	2.35	0.224	-0.15	-1.916	7	0.097
O_{avg}	3.19	0.184	0.69	12.389	10	<0.001
T_{avg}	2.89	0.179	0.39	4.811	4	0.009

t(6) = 3.196, p = 0.019. Since $p < 0.05$, the null hypothesis is rejected and it can be concluded that the mean of average strength factors is statistically significantly different from the threshold score of 2.5.

Average score of weakness factors (M = 2.35, SD = 0.22) is lower than the threshold score of 2.5, a statistically significant mean difference of -0.15, t(7) = -1.916, p = 0.097. Since the p-value is greater than 0.05 therefore the null hypothesis is retained and it can be concluded that the mean of average weakness factors is not statistically significantly different from 2.5.

Average score of opportunity factors (M = 3.19, SD = 0.18) is higher than the threshold score of 2.5, a statistically significant mean difference of 0.69, t(10) = 12.389, p = 0.000. Since $p < 0.001$, the null hypothesis is rejected and it can be concluded that the mean of average opportunity factors is statistically significantly different from the threshold score of 2.5.

Average score of threat factors (M = 2.89, SD = 0.18) is higher than the threshold score of 2.5, a statistically significant mean difference of 0.39, t(4) = 4.811, p = 0.009. Since $p < 0.05$, the null hypothesis is rejected and it can be concluded that the mean of average threat factors is statistically significantly different from the threshold score of 2.5.

In summary, the average of agreement level toward weakness is slightly lower than, but not significantly different to the threshold score of 2.5 (p-value > 0.05). Thus it can be concluded that the mean of average weakness items is equal to 2.5 indicating that KPS staff seemed to agree with the weakness of KPS produced from IPA based SWOT analysis. The average of agreement level toward strength, opportunity and threat are statistically significantly different from the threshold score of 2.5 and their means are higher than 2.5 which also indicated that KPS staff seemed to agree with the strength, opportunity and threat of KPS produced from IPA based SWOT analysis.

Specifically, the average of agreement level towards strength and threat is close to 3.0 and the average of agreement level toward opportunity is higher than 3.0. To confirm that staff were mostly agreed with these three aspects of KPS SWOT, the one-sample t-test was further conducted with the threshold of 3.0 which means agree. The result is shown in Table 9.10 and the means of response across the three variables is visually presented as simple error bars in Figure 9.4.

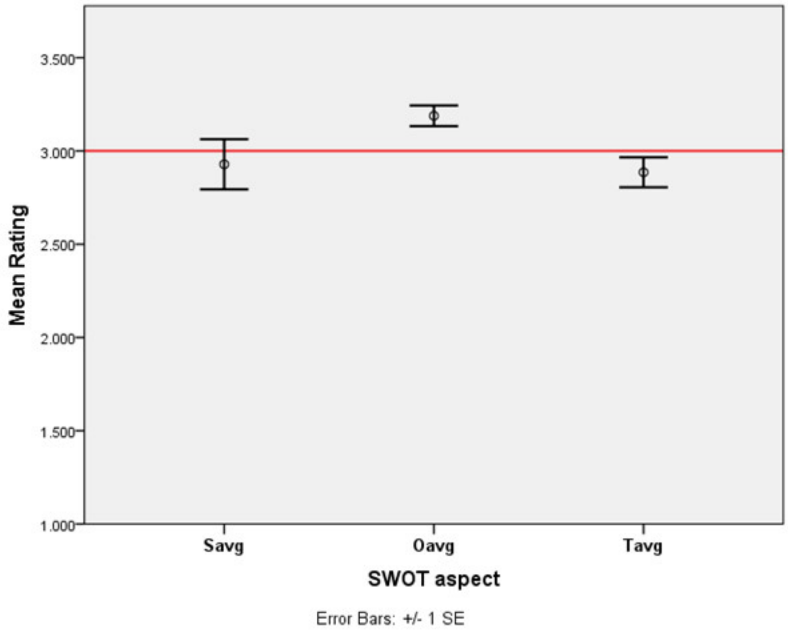


Figure 9.4: Simple error bar of the three variables of SWOT group. The mean is represented by a dot, a line represents standard error of the mean, a red horizontal line is represented the threshold score of 3.0

Table 9.10: Result of the One-sample t-test for three SWOT group (Test value is 3.0)

Variable	Mean	SD	Mean Difference	t	df	p (2-tailed)
S_{avg}	2.93	0.355	-0.07	-0.533	6	0.613
O_{avg}	3.19	0.184	0.19	3.390	10	0.007
T_{avg}	2.89	0.179	-0.11	-1.425	4	0.227

With regards to Table 9.10, the average of agreement level toward opportunity is statistically significantly different from 3.0 and its value is higher than 3.0, and the average of agreement level toward strength and threat is lower than, but not statistically significantly different to 3.0. It can be confirmed that KPS staff were mostly agreed with strength, opportunity and threat of KPS produced from IPA based SWOT analysis. Taken together, results from Table 9.9 and Table 9.10 suggest that KPS staff agreed with factors within each SWOT aspect.

The analytical results at both individual and aggregate level of SWOT factors provides answer to the sub research question 3.1. As they consistently indicated that KPS staff mostly agreed with strengths, opportunities and threats of their department and the analytical result of weaknesses in aggregate level showed that KPS staff were agreed on weaknesses. The analytical result of weaknesses in individual level showed that only 25.00% of weakness factors were agreed by staff. These results will be discussed in Chapter 10.

9.3 Results of SWOT experienced users evaluation on the SWOT of the case study

The experienced users of SWOT survey collected through the methodology described in Section 8.5 was analysed to determine whether there is a difference in rating score regarding the quality of SWOT of KPS produced through the two SWOT approaches: IPA based SWOT analysis and Traditional SWOT analysis.

Specifically, 44 participants of this survey were divided equally into two groups. The first group rated the 10 questions for IPA based SWOT analysis while the second group rated the 10 questions for traditional SWOT analysis. The mean and standard deviation of rating score of the two SWOT approaches are presented in Table 9.11.

Table 9.11: The mean and standard deviation of rating score on 10 questions regarding the two SWOT approaches

Question	**IPA based SWOT**		**Traditional SWOT**	
	Mean	Std. Deviation	Mean	Std. Deviation
1	3.364	0.790	3.227	0.752
2	3.318	0.568	3.500	0.673
3	3.364	0.790	3.227	0.752
4	3.000	0.926	3.136	0.834
5	3.364	0.902	3.909	0.426
6	3.864	0.560	3.409	0.734
7	4.045	0.576	3.273	0.827
8	3.409	1.008	4.182	0.501
9	3.500	0.740	4.091	0.526
10	3.409	0.854	4.091	0.294

The hypothesis was tested using Mann-Whitney U at the 5% significance level. If the asymptotic p-value is less than 0.05 then, there are statistically significant differences in the mean ranks of the rating score between the two groups. The results of Mann-Whitney U test are shown in Table 9.12.

The column *Mean Rank* in Table 9.12 indicates which group can be considered as having the higher rating score. For each question, the group with the higher mean rank has the higher rating score than another. The column *U statistic* and *Asymp. Sig* in Table 9.12 provides the test value and the asymptotic significance value of the test respectively. The following observations can be drawn from the interpretation of this table.

(1) The mean rank of rating score for the question numbers 1-4 between the two groups of SWOT approaches was not statistically significantly different. These questions are related to common flaws of the SWOT analysis outcome. Therefore, it can be concluded that the two SWOT approaches have about the same quality measured on the common flaws of the SWOT analysis.

Table 9.12: Results of the Mann-Whitney U test (N=22)

Questionnaire item	**Mean Rank**		**U statistic**	**Asymp. Sig. (2-tailed)**
	IPA based SWOT	Traditional SWOT		
1. Strengths and weakness are explicit (clearly and unambiguously formulated).	23.50	21.50	220.00	0.572
2. Opportunities and threats are explicit (clearly and unambiguously formulated).	20.89	24.11	206.50	0.348
3. SWOT factors are not overgeneralised	23.41	21.59	222.00	0.613
4. SWOT factors are not correctly classified as internal/external factors	21.48	23.52	219.50	0.577
5. SWOT factors are comprehensively explain the department's situation	18.48	26.52	153.50	0.015
6. All SWOT factors shown are measurable	25.89	19.11	167.50	0.043
7. The order of SWOT items makes decision-making easier.	27.84	17.16	124.50	0.002
8. The data source for this SWOT analysis is reliable	17.57	27.43	133.50	0.005
9. The SWOT items can be used as a starting point for strategic planning	18.09	26.91	145.00	0.006
10. The SWOT items provide useful information that supports decision-making regarding strategic planning	17.05	27.95	122.00	0.001

(2) The mean rank of rating score for the question numbers 5 and 8-10 in the Traditional SWOT analysis group was higher than and statistically significantly different from that of the IPA based SWOT analysis group. Question number 5 and number 8 are related to the comprehensibility of SWOT analysis outcome and reliability of data source of SWOT analysis respectively. Question number 9-10 are related to action-ability of SWOT analysis outcome. Hence, it can be concluded that Traditional SWOT analysis is better than IPA based SWOT analysis in terms of comprehensibility, reliability of data source, and action-ability.

(3) The rating score of the question numbers 6-7 in the IPA based SWOT analysis group was higher than and statistically significantly different from that of the Traditional SWOT analysis group. These questions are related to measurability of SWOT analysis outcome. Therefore, it can be concluded that IPA based SWOT analysis is better than Traditional SWOT analysis in terms of measurability.

Taken together, results from Table 9.11 and Table 9.12 suggest that quality for all evaluation aspects of the two SWOT approaches are considered acceptable as the average scores are centred around 3.00 to 4.00 and no SWOT approach is superior to another in all evaluation aspects. Hence, the answer to the sub research question 3.2. is that the quality of the outcome of IPA based SWOT analysis is considered acceptable. Discussions of these results will be presented in Chapter 10.

9.4 Summary

This chapter reports the results of the three surveys - student satisfaction survey, staff evaluation of SWOT survey and experienced users of SWOT survey which were conducted to demonstrate and evaluate the IPA based SWOT analysis so as to help answer the third research question. To demonstrate how the IPA based SWOT analysis can be used to generate SWOT, the student satisfaction surveys of KPS and BKN were analysed through the IPA based SWOT analysis framework which generated SWOT of KPS. The SWOT of KPS consists of factors classified as 7 strengths, 8 weaknesses, 11 opportunities and 5 threats. All SWOT factors are prioritized by their weight computed on the basis of importance-performance analysis. The factor with the larger magnitude of the weight has higher priority for maintenance or improvement than the factor with the smaller one.

For the purpose of evaluation, first the SWOT of KPS was assessed to determine whether KPS staff agreed or disagreed with it through the staff evaluation of SWOT survey. All data collected from the staff evaluation of SWOT survey was analysed at individual and aggregate level of SWOT factors using a one-sample Wilcoxon signed rank and the one-sample t-test respectively. The analytical results of these statistical tests showed that the staff of KPS agreed with SWOT of KPS. This suggested that the outcome of IPA based SWOT analysis accurately reflected the KPS's situation.

Second, the SWOT of KPS produced from IPA based SWOT analysis was compared with another version of SWOT of KPS produced from traditional SWOT approach, by experienced users of SWOT to determine whether there is a difference in rating score regarding the quality of SWOT of KPS produced through the two different SWOT approaches. All data collected from experienced users of SWOT survey was analysed using Mann-Whitney U test. The analytical results of Mann-Whitney U test suggest that no SWOT approach was superior to another in all evaluation aspects.

These results demonstrated that the IPA based SWOT analysis can be used to process and analyse customer satisfaction survey to generate SWOT that accurately reflects the organisation's situation. In addition, the quality of the outcome of IPA based SWOT analysis was considered acceptable compared with the outcome of traditional SWOT analysis. These results will be further discussed in Chapter 10.

Chapter 10

Discussion

Having conducted the empirical comparison of importance measure techniques and the case study in the previous chapters, this chapter reflects upon and discusses their results that help to answer the research questions of this study. Section 10.1 discusses results and methodology related to the empirical comparison of importance measure techniques. Section 10.2 discusses about pros and cons of IPA based SWOT analysis. Section 10.3 discusses evaluation results of IPA based SWOT analysis and finally Section 10.4 discusses the limitations of the research.

10.1 Discussion of Empirical comparison

There has been an ongoing argument over whether self-stated or implicitly derived methods should be used in measuring importance (Gustafsson and Johnson, 2004; Fontenot et al., 2007) or even a debate on the most appropriate method to obtain implicit importance (Feng et al., 2014; Lai and Hitchcock, 2015).

Particularly, over the years many different IPA variations have emerged with the main focus on a method to indirectly measure *importance*. Many researchers have claimed that their complicated mechanism works best without conducting an empirical comparison of evaluation metrics (Deng et al., 2008a; Hu et al., 2009; Krešić et al., 2013).

To date, there have been a few empirical studies that compared the different methods for measuring importance against a set of evaluation criteria (Azzopardi and Nash, 2013). Hence, the empirical comparison was set out with the aim of identifying a suitable technique for measuring importance in this thesis.

The results of the empirical comparison indicated MLR is the best importance measure technique whereas self-stated importance measure using direct-rating scales is the worst importance measure technique in terms of the three evaluation metrics. This finding supports the claim that implicitly derived importance measures should be used in measuring importance rather than the self-stated importance measure, and MLR is the most favourable technique for implicitly derived importance.

The finding in favour of MLR raises the case that a simple technique works best. However, the *importance* obtained from MLR needs to be used with caution, as this technique is likely to produce negative coefficients. A full discussion of the empirical results can be found in Section 6.4 while discussions related to the methodology of empirical comparison are provided in the following paragraphs.

- **Evaluation metrics**

  Evaluation metrics play a critical role in determining the best importance measurement techniques. Idealistically, the best comparison of alternative importance measurement techniques would be to develop different action plans according to importance measurement technique and observe business performance measures such as Return on Investment (ROI). However, considering time and cost to carry out such an experiment, it is practically hard or even impossible to carry out (Hauser, 1991).

  As a matter of fact, researchers of previous comparative studies specified a set of evaluation metrics based on their judgement on what is a good *importance* measure. Three evaluation metrics namely predictive validity, diagnosticity, and discriminating power which were most used in the previous comparative studies were selected to use in the present empirical comparison. The use of the multiple metrics not only allows the author to investigate different aspects of the importance measures but also maximizes the accuracy of the evaluation.

  Among the three selected evaluation metrics, predictive validity (ability to predict actual overall customer satisfaction) is the most used evaluation metrics in the previous comparative studies (Hauser, 1991; Gustafsson and Johnson, 2004; Chrzan and Golovashkina, 2006; Taplin, 2012b). However, it is considered to be problematic for a comparison as it allows MLR and OLR to gain an advantage over other techniques, because their *importance* was calculated on the same data which was used for measuring predictive validity as correlation coefficients (Neslin, 1981) therefore maximizing predictive validity.

  Concerning this issue, the relative predictive validity (ability to predict customer-reported priority) is a good alternative choice of the predictive validity since a prediction of customer preference rank-order provides a fair comparison among techniques. However, to measure this metric cause survey respondent to do an additional task which might affect survey response rate. Additionally, this present study used the two surveys from other research studies that did not provide the customer preference rank-order. Therefore, the relative predictive validity was not fit to use in the context of this empirical study.

Turning now to the weight of evaluation metrics, this is another key element that should be considered when the multiple metrics are used. Currently, either a set of metrics or the most critical criteria for identifying the importance measure technique has not been yet established. This study weighted all three evaluation metrics equally. This weighting scheme is similar to the weighting scheme used in the previous studies by Gustafsson and Johnson (2004); Chrzan and Golovashkina (2006).

On the other hand, this weighting scheme differs from the one used in the study by Pokryshevskaya and Antipov (2014). Their study used a more complicate weighing scheme in which the weights of evaluation metrics were varied using an iterative Monte-Carlo technique. This is because Pokryshevskaya and Antipov (2014) could not identify a clear winning technique if the two evaluation metrics were equally weighted, as there was not much difference between the ranked mean of the top six techniques. In contrast, there was a clear difference between the ranked mean of the top three techniques compared in this study (see Table 6.7). Hence, it is not necessary to conduct such a complicate weighting scheme in this study.

- **Use of the baseline evaluation metric**
  It is good practice to establish a baseline measurement for each metric. In this empirical study, the baseline measurements were computed under the criterion that all attributes are equally important in another word all have importance value equal to one.

  The baseline measurements then served as a point of reference to assess the effect of importance measures by different techniques. By comparing the value of evaluation metrics namely predictive validity and discrimination power of each technique with that of the baseline measurements, it shows that importance measures of all techniques do improve both predictive validity and discrimination power (see Table 6.4 and Table 6.6), thereby, demonstrating the merit of importance measure.

  The use of baseline measurements shows that the present empirical study was carefully conducted which reveals such an interesting finding as previously stated. In spite of its benefit, the majority of the previous studies did not set any baseline measurement except the work by Hauser (1991). Hence, this study recommends researchers in IPA field to establish baseline measurements when comparing importance measuring techniques.

- **Number of datasets used**
  With regard to number of datasets, this empirical study used three datasets in different business areas while the majority of the previous studies conducted the comparison on one or two dataset(s). The fact that the three datasets were used in this study allows the author to infer generality of the findings reported in Section 6.2.4 with confidence.

However, the size of the datasets used in this study is considered small (hundred records) and moderate (thousand records) when compared with the other context of a data mining application (Azzalini et al., 2012, p. 193). Further comparison of importance measuring techniques on a bigger scale of dataset would be beneficial for the IPA studies.

10.2 Discussion of IPA based SWOT analysis and its application

According to a demonstration of IPA based SWOT analysis application through a case study of HEIs, it shows that IPA based SWOT analysis can be used in supporting strategic planning in a real-world situation. Although the application of IPA based SWOT analysis in this study was illustrated in the specific field, the IPA based SWOT analysis can be used widely in the other business areas where SWOT analysis has been seen applicable and the customer satisfaction surveys are generally conducted.

As its name conveys, a distinctive characteristic of the IPA based SWOT analysis is that it utilizes IPA to systematically generate SWOT factors based on a customer satisfaction survey. To the best of the author's knowledge, this is the first study that applies IPA in generating SWOT factors.

The combined use of the IPA and SWOT analysis yields quantitative SWOT factors regarding the *importance* and *performance* obtained from customer satisfaction surveys. In addition, it yields customer-oriented SWOT factors which guarantee that the capabilities perceived by an organisation are recognized and valued by the customers and diminish bias of an organisation's staff. This facilitates an organisation to efficiently formulate strategic planning for maintaining or enhancing customer satisfaction, thereby gaining a competitive advantage.

It can be stated that IPA based SWOT analysis diminishes two major drawbacks of traditional SWOT analysis as well as bridges the gap of the current two main streams of SWOT approaches: quantitative SWOT analysis (Kurttila et al., 2000; Kangas et al., 2001; Lee and Lin, 2008) and customer-oriented SWOT analysis (Dai et al., 2011; Pai et al., 2013). The quantitative approach can make SWOT factors commensurable but they are identified solely on an organisation's perspective without considering the customer's perspective. On the other hand, the customer-oriented approach generates SWOT factors from a customer's perspective but it has no means to prioritize SWOT factors as they are extracted from unstructured data sources such as e-mail, online blog, and social media content.

Considering data sources for the different SWOT analysis approaches and their associated analytical techniques (Table 10.1), IPA based SWOT analysis is the only approach that uses structured data as an input data source while the others use unstructured data. The IPA based SWOT analysis performs a deeper analysis of data from customer satisfaction surveys that the organisation generally collects. Hence, it provides more credible SWOT factors than traditional and quantitative SWOT analysis that identify SWOT based on staff opinion usually with no supporting evidence.

The IPA based SWOT analysis also gains an advantage over the customer-oriented SWOT analysis in terms of simplicity since it requires less complicated analytical techniques for processing data than the customer-oriented SWOT analysis that requires computational linguistics to process customer opinion in the form of text, both at syntax and semantic level.

Table 10.1: Summary of data sources for the different SWOT analysis approaches and their associated analytical techniques

SWOT approaches	Data source	Data type	Analytical technique
Traditional SWOT analysis	Individual opinion of organisation's staff who attend a brain storming session	Unstructured data as personal opinion	n/a
Quantitative SWOT analysis	Same as above	Same as above	AHP, ANP
Customer-oriented SWOT analysis	e-mail, online blog, and social media content	Unstructured data, text	NLP, Text mining, Sentimental analysis
IPA based SWOT analysis	Customer satisfaction survey measured in Likert scales	Structured data, numeric	IPA

To take full advantages of IPA based SWOT analysis, all four steps for developing IPA based SWOT analysis (see Section 7.2) need to be carefully conducted. Among the four steps, the most critical one is the first step which involves the undertaking of a customer satisfaction survey. This is because a list of SWOT factors is dependent on the questionnaire items/questions that represent attributes of the organisation being surveyed. Therefore, the attributes need to be carefully selected to ensure that the evaluative attributes that are important to the customer are not overlooked as discussed by Martilla and James (1977) who suggested that attributes should be listed based on a thorough literature review in each application area or by interviews. The attributes associated with the student satisfaction survey conducted in this thesis were selected based on a list of attributes defined in previous studies of student satisfaction (Siskos and Grigoroudis, 2002; Silva and Fernandes, 2011; Grebennikov and Shah, 2013).

Another criteria in undertaking a customer satisfaction survey is that two identical questionnaires are required by IPA based SWOT analysis so that each factor of target organisation and its competitor can be compared and categorised into four aspects of SWOT. As the two organisations are willing to corporate in the case study, it is not difficult to fulfil this requirement in conducting the case study of this thesis. However, this requirement of the two identical questionnaires may not be practical which raises the issue of how can the organisation collect the satisfaction survey from its competitor's customers. One possible way to resolve this issue is to use customer satisfaction measurements on a national scale such as ACSI, UKCSI, and so forth, as a data source of IPA based SWOT analysis since, these measurements of customer satisfaction at a national level provide uniform and comparable information that allows for systematic benchmarking across firms.

10.3 Discussion of evaluation result of IPA based SWOT analysis

The two surveys described in Chapter 8 for the purpose of evaluating IPA based SWOT analysis. This section provides discussions of the results of these two surveys which were reported in Chapter 9.

10.3.1 Discussion of staff evaluation on the SWOT of the case study

A survey of KPS staff evaluation on SWOT of KPS, the results of IPA based SWOT analysis, set out to assess whether KPS staff agree or disagree with the SWOT of KPS. Data analysis was carried out at individual (each SWOT factor) and aggregate level (a group of SWOT factors).

The analytical results at both individual and aggregate level of SWOT factors consistently indicated that KPS staff mostly agreed with strengths, opportunities and threats of their department. However, analytical results toward the weaknesses are slightly different at the two analysis levels. The analytical result of weaknesses at aggregate level showed that KPS staff seem to agree on weaknesses whereas the opposite was true for the analytical result of weaknesses at the individual factor level.

The aggregate results suggest that KPS staff agreed with all aspects of the IPA based SWOT analysis for KPS. This demonstrates a validity of the IPA based SWOT method in a real-world case and confirms that it can thus be used in providing useful information for strategic planning. With regard to the results at the individual factor level, there is a low face validity on Weaknesses (W), as only 25.00% of weakness factors were agreed by staff. While other aspects of SWOT have high face validity, for example up to 80% of Strength (S), Opportunity (O) and Threat (T) factors were agreed by staff.

Although the result of the Weaknesses (W) factor at the individual factor level shows a presence of low face validity of IPA based SWOT analysis, it reveals another interesting finding that staff seem to blind on their organisation's weaknesses. The most likely explanation for the small percentage of staff agreement toward weaknesses is an unwillingness to admit individual weaknesses as highlighting weaknesses can draw attention to areas of the organisation that one has badly operated. Thus, one reacts in a state of denial[1] about them. According to Humphreys (2007), many senior managers deny to talk about weaknesses objectively as doing so imply criticism of the way that the organisation has been managed.

10.3.2 Discussion of SWOT experienced users evaluation on the SWOT of the case study

Experienced users of SWOT evaluation set out to assess whether there is a different in rating the quality of SWOT of KPS produced through the two SWOT approaches: IPA based SWOT analysis and traditional SWOT analysis. The descriptive statistics suggested that quality for all evaluation aspects of the two SWOT approaches was considered acceptable as the average scores ranged from moderate to good and the analytical results indicated that no SWOT approach was superior to another in all evaluation aspects.

Specifically, three findings can be drawn from the analytical results. The first finding suggests that common flaws of SWOT analysis were similar in both SWOT approaches. However, their rating scores regarding common flaws were considered acceptable.

Regarding the second finding, traditional SWOT analysis is better than IPA based SWOT analysis in term of comprehensibility, reliability of data source, and action-ability. This finding was unanticipated and suggests that comprehensibility, reliability of data source and action-ability of IPA based SWOT analysis are not clearly perceived by experienced SWOT users.

This result may be explained by the fact that information provided in the survey may not be enough for them to perceive these characteristics of IPA based SWOT analysis. Additionally, experienced SWOT users may be influenced by a primacy effect[2] as traditional SWOT analysis has been used for decades and experienced SWOT users are more familiar with it than IPA based SWOT analysis. In this circumstance, the experienced SWOT users might rate a higher score for the outcome of traditional SWOT relative to IPA based SWOT analysis. A recent study by Bansback et al. (2014), showed that primacy effect is also prominent in making decisions regarding the treatment options for patients.

[1]Denial is an ego-defence mechanism that unconsciously operates to reduce anxiety by refusing to perceive the unpleasant aspects of external reality (New World Encyclopedia, 2013)

[2]The primacy effect is a cognitive bias that results in the first introduced information to be remembered better or more easily, or to be more influential than information presented later on. (Kardes and Herr, 1990).

Finally, the third finding indicates that IPA based SWOT analysis is superior to traditional SWOT analysis in terms of measurability which enables SWOT factors to be prioritized. This finding provides evidence to confirm that IPA based SWOT analysis is able to diminish one of the great limitations of traditional SWOT analysis which is a lack of factor prioritization. Hence, it can be implied that using IPA based SWOT analysis provides more informative data for successful strategy formulation.

The combination of these finding provides some support for the soundness of IPA based SWOT analysis with regard to measurability and quality of its outcome although, it has less comprehensibility, reliability of data source and action-ability than the traditional SWOT analysis.

10.3.3 Discussion of sample size for the SWOT evaluation surveys

The numbers of KPS staff and MBA students who participated in the two evaluation surveys of IPA based SWOT analysis were 14 and 22 (per group) respectively. Although the numbers are relative small, they are approximately 93.34% and 91.67% of the population of the case study as the total number of staff member of KPS is 15 and the total number of MBA students is 48. These high proportions of responses indicate that the survey results are representative of the target population. This provides the author with the confidence in the findings based on evaluation survey results as the non-response bias is reduced (Fincham, 2008).

Considering the sample distribution, it was found that the sample distribution of the two surveys deviated from normal. For these reasons, the parametric statistics for comparing means such as t-tests were not used in analysing data of this study to avoid the incorrect findings. Although non-parametric statistics do not require any assumption regarding population distribution, they have two main shortcomings. The first is that non-parametric statistics tend to have less discriminative power than their analogous parametric statistics. For the tests of difference between groups, this may result in failure to detect the differences between groups when they actually exist (Pallant, 2005). The second drawback is that results of non-parametric statistics are often less easy to interpret than that of parametric statistics (Hoskin, 2012). With regards to these shortcomings of the non-parametric test, the evaluation results of this study therefore need to be interpreted with caution.

10.4 Limitations of the present study

It is important to note a number of limitations associated with the empirical study and the proposed framework presented in this study.

10.4.1 Limitations of the empirical study

The current empirical study has only compared several implicitly derived importance measures with a single customer self-stated importance measure named direct-rating scales. This is because direct-rating scales are less complicated customer self-stated importance measures than others such as constant sum, Q-sort, Maxdiff.

The simplicity of direct-rating scales make them commonly used and they have been chosen as a representative of a customer self-stated importance measure to compare with several implicitly derived importance measures in previous comparative studies (Gustafsson and Johnson, 2004; Bacon, 2003; Taplin, 2012b). Additionally, the main focus of this study is to investigate the statistical and data mining techniques for implicitly derived importance that could be ascertained in an absence of the customer self-stated importance measure.

The second limitation is that only two common statistical techniques namely MLR and OLR were included in the present empirical study as this study intended to investigate a new data mining technique that could be used in measuring importance and data mining is the area of research interested of the author rather than the advanced statistical techniques such as Principal Components Regression (PCR),Partial Least Squares (PLS), and Structural Equation Models (SEMs).

10.4.2 Limitations of the proposed framework

- **I-P mapping partition**
  Another technical issue of IPA referred to as I-P mapping partition (Lai and Hitchcock, 2015) was not investigated under the context of this study. Although different I-P mapping approaches may results in different IPA matrix and thereby a different SWOT outcome. Generally, there are three I-P mapping approaches that could be used to partition the IPA space: scale-centred quadrants approach, data-centred quadrants approach and diagonal line approach (Bacon, 2003).

  In this study, the IPA matrices are constructed subject to the use of 'data-centred quadrants approach' in which the mean of *importance* and *performance* are set as the cross-points. This approach offers a higher discriminative power than 'scale-centred quadrants approach' that use midpoint of the measurement scale as the cross-points (Garver, 2003). However, it has less predictive validity than diagonal line approach regarding to the empirical study conducted by Bacon (2003).

- **External Macro factors**
  This study focused on the customers' attitude hence, SWOT factors produced by IPA based SWOT analysis do not cover a broad range of factors especially the external macro factors such as economy, politics, technology, and trends that may

affect the organisation. This is also the limitation of another customer-oriented SWOT analysis by Pai et al. (2013) as well as the knowledge-based SWOT analysis system by Houben et al. (1999). In fact, the analysis of the external macro factors can be an issue of the traditional SWOT analysis if it is conducted by persons that lack the external situation's knowledge or conducted based on insufficient information.

Regarding this limitation, it should be noted that IPA based SWOT analysis is not intended to replace the traditional SWOT analysis but rather to provide a complete view of an organisation's situation from the customer side, while the traditional SWOT analysis provides information from the organisation side and information on external macro factors.

10.4.3 Limitation of evaluation of IPA based SWOT analysis

To evaluate a quality of outcome produced by IPA based SWOT analysis, a traditional SWOT analysis was used as a comparator. In this study, the traditional SWOT was produced through the brainstorming conducted by staff of the target company. The author let staff of the target company conducted the brainstorming in the way that they are get used to and it is highly likely that they did not take student satisfaction into consideration while they were generating SWOT.

Hence, comparing this traditional SWOT with the IPA based SWOT cannot clarify that the differences between them are relied on the methodology (brainstorming vs IPA based SWOT analysis) or data source (without vs with student satisfaction). It might be more useful if in the future the full scale of experiment would be conducted to compare (1) traditional SWOT analysis without student satisfaction, (2) traditional SWOT analysis with student satisfaction (in which staff would be given student satisfaction for consideration in their brainstorming session), and (3) IPA based SWOT analysis to help differentiate between SWOT approaches.

Chapter 11

Conclusion and Future work

SWOT analysis is one of the most important tools for strategic planning. It is the most widely applied strategic tool by both academic communities and business areas such as organisations in UK as well as the enterprises in South African (as described in Section 2.1). SWOT analysis offers a simple structured approach that helps organisations to gain better insight into their internal and external business environment. Despite its advantages in term of simplicity and widely used applications, the traditional approach of conducting SWOT analysis does not prioritize and is likely to hold subjective views that may result in an improper strategic action (see Sub-section 2.1.2).

Therefore, this study proposes an IPA based SWOT analysis that adopts the Importance-Performance Analysis (IPA), a technique for measuring customers' satisfaction from customer satisfaction surveys, to systematically generate SWOT factors corresponding to the customers' perspective. As mentioned in the literature review, this study bridges the gap of the current SWOT analysis approaches by providing both quantitative and customer-oriented SWOT factors which improves a deficiency of traditional SWOT analysis.

The key steps of the IPA based SWOT analysis are the IPA matrix construction in which a customer satisfaction survey is analysed to calculate the attributes' importance and the attributes' performance, and SWOT factors identification based on the IPA matrix. Specifically, strengths and weaknesses are identified through an IPA matrix of the organisation. Opportunities and threats are obtained by comparing the IPA matrix of an organisation with that of its competitor.

Through the use of IPA based SWOT analysis, it is expected that an organisation can efficiently formulate strategic planning as the SWOT factors that should be maintained or improved can be clearly identified based on customers' viewpoints. The IPA based SWOT analysis also makes the best use of data from customer satisfaction surveys that the organisation generally collects. The application of the IPA based SWOT analysis was illustrated and evaluated through a case study of HEIs in Thailand. The evaluation results showed that SWOT analysis of the case study have high face validity and was considered to have acceptable quality, thereby demonstrating the validity of IPA based SWOT analysis.

Three research questions were addressed through this research study. The findings for each research question are briefly presented as follows:

Research question 1. Which importance measure should be used in IPA?
The question was answered through the empirical comparison of different techniques for measuring importance: MLR, OLR, BPNN, Naïve Bayes, BNs, and direct-rating scales (Chapter 5).

Research sub-question 1.1. Which approaches for assessing *importance* work best in IPA: customer self-stated or implicitly derived importance measure?

The results of the empirical comparison indicated that an implicitly derived importance measure works best in IPA as the customer self-stated importance measure using direct-rating is ranked last among the importance measurement techniques in all evaluation metrics (Section 6.2).

Research sub-question 1.2. Which data mining technique is most appropriate for measuring *importance*?

The result of the empirical comparison indicated that MLR is the best importance measurement technique based on the average rank of three evaluation metrics, followed by OLR and BNs (Section 6.2).

Research question 2. How can IPA be applied to develop a SWOT analysis based on customer satisfaction surveys?
IPA based SWOT analysis was developed as a methodological framework to apply IPA in order to identify SWOT analysis based on customer satisfaction surveys (Chapter 7). The application of IPA based SWOT analysis was then demonstrated through the case study of one department of the leading university in Thailand, named KPS (Section 8.3, 9.1).

Research question 3. How good is the outcome of IPA based SWOT analysis?
The question was answered through the analytical results of the two surveys: staff evaluation of SWOT survey and experienced users of SWOT survey, each of which related to each sub research question.

Research sub-question 3.1. What is the staff level agreement on the outcome produced by IPA based SWOT analysis?

The staff of KPS were asked to rated their level of agreement toward the SWOT of their department that was produced by using IPA based SWOT analysis (Section 8.4). The staff evaluation was analysed using one-sample Wilcoxon signed rank and one-sample t-test, and the analytical results indicated that IPA based SWOT analysis produces the outcome that have high face validity as staff agreed with all aspects of SWOT outcome (Section 9.2).

Research sub-question 3.2. What is a quality of outcome produced by IPA based SWOT analysis compare to the traditional SWOT analysis?

The outcome of IPA based SWOT analysis and traditional SWOT analysis were assessed through several questions related to quality and perceived usefulness of SWOT by two groups of MBA students (Section 8.5). The responses of the MBA students were analysed by using the Mann-Whitney U test. The analytical results indicated that the quality of outcome of IPA based SWOT analysis is considered acceptable as it is about the same as that of traditional SWOT analysis (Section 9.3).

11.1 Contributions

This research proposes an alternative way to conduct SWOT analysis based on customer satisfaction surveys. The contributions of this research can be summarized as follows:

- **IPA based SWOT analysis**
  IPA based SWOT analysis was developed in order to serve as a set of steps for generating SWOT factors based on customer satisfaction surveys. A distinctive characteristic of the IPA based SWOT analysis is that it utilizes IPA to prioritize SWOT factors by means of *importance* and *performance*.

  The key steps of the IPA based SWOT analysis are the IPA matrix construction and SWOT factors identification based on the IPA matrix. The former involves the quantitative analysis of a customer satisfaction survey in order to calculate the *importance* and the *performance*. The latter involves the identification of strengths and weaknesses through an IPA matrix of the organisation and identification of opportunities and threats through a comparison between the IPA matrix of an organisation with that of its competitor.

  By using IPA based SWOT analysis, the generated SWOT factors are not only measurable but also meaningful as they are identified based on customers' points of view. A measurable SWOT factor enables an organisation to prioritize SWOT factors in creating an action plan while a customer oriented SWOT factor guarantees that the capabilities perceived by an organisation are recognized and valued by the customers. This facilitates an organisation to efficiently formulate strategic planning for maintaining or enhancing customer satisfaction, thereby gaining a competitive advantage.

- **Empirical comparison of importance measuring techniques**
  The empirical comparison of importance measuring techniques has two contributions worthy of attention. First, a review of past comparative studies revealed that, this is the first study that compared the two common statistical techniques for implicitly deriving importance which are MLR and OLR with the BPNN, Naïve Bayes and BNs as the new emerging techniques for implicitly deriving importance, using a set of evaluation metrics.

  Second, the comparison highlights possible significant differences between the techniques through different aspects of the importance measures and raises the case that the simple technique works best, as the results indicated that MLR outperformed other complicated techniques for implicitly derived importance such as BPNN and BNs.

  Based on these two intended contributions, it can be concluded that this comparative study can provide guidance for researchers and practitioners in applying statistical or data mining techniques for measuring *importance*. On top of that, it is expected that the present comparative study is able to provide new and useful information for business improvement and highlights a new research area of IPA.

11.2 Future work

Though contributions of this research are noticeable, some research directions could be explored further to make IPA based SWOT analysis become a best practice to conduct SWOT analysis. The five future research directions are ranked in the following priority for development. The first three are future work regarding IPA as it is a core component of the IPA based SWOT analysis, and the rest is future work regarding IPA based SWOT analysis as a whole.

11.2.1 Investigating another technical issue of IPA

As was pointed out in the limitations of the research (see Sections 10.4.2), I-P mapping is another technical issue of IPA that should be investigated in order to obtain a soundness IPA matrix thereby a more reliable SWOT outcome.

Although there are three I-P mapping approaches: scale-centred quadrants approach, data-centred quadrants approach and diagonal line approach, only the data-centred quadrants approach and diagonal line approach will be chosen for further investigation. This is because previous research has reported that the scale-centred quadrants approach has a major drawback regarding discriminative power (Garver, 2003).

11.2.2 Exploring the other evaluation metrics or justifying the weight of evaluation metrics

As discussed in the previous chapter (Section 10.1), evaluation metrics are the key criteria in determining the best importance measure techniques. Using different evaluation metrics may result in identifying difference techniques which may result in identifying dramatically different attributes for improvement (Tontini and Silveira, 2007). Hence, it is important to carefully select evaluation metrics based on the definition of what is a good importance measure.

However, currently the common agreement regarding a good importance measure has not yet clearly defined. A further literature review in marketing or even a survey of the opinions of experts in the field needs to be conducted to define the characteristic of a good importance measure. Consequently, possible evaluation metrics can be identified and weighted according to their contribution to the importance measure.

11.2.3 Considering the other data mining techniques or a combination of techniques in measuring *importance*

A literature review about research on mining customer satisfaction data (see Section 2.3.4) showed that apart from the five techniques used in this study, there are some interesting data mining techniques that could be used for measuring *importance* from customer satisfaction data such as a decision trees and a support vector machines (SVM).

A decision tree is one of the most widely used techniques in the context of Customer Relationship Management (CRM) including customer identification, customer attraction, customer retention, and customer development (Ngai et al., 2009). SVM is considered as the best classifier alongside neural network with respect to the accuracy (Duan and Da Xu, 2012).

Additionally, it would be interesting to observe a combination of the techniques in measuring *importance*. For example, Klicek et al. (2014) proposed a combination of Bayesian and neural networks to determine which attributes most influence the overall customer satisfaction and they reported that a combination of techniques outperformed the single statistical technique.

11.2.4 Developing a prototype system of IPA based SWOT analysis

Another possible area of future research direction would be to implement the prototype system of IPA based SWOT analysis for automatically generating SWOT based on a customer satisfaction survey. The proposed architecture consists of three main components namely User Interface, Analysis module and Data module, see Figure 11.1:

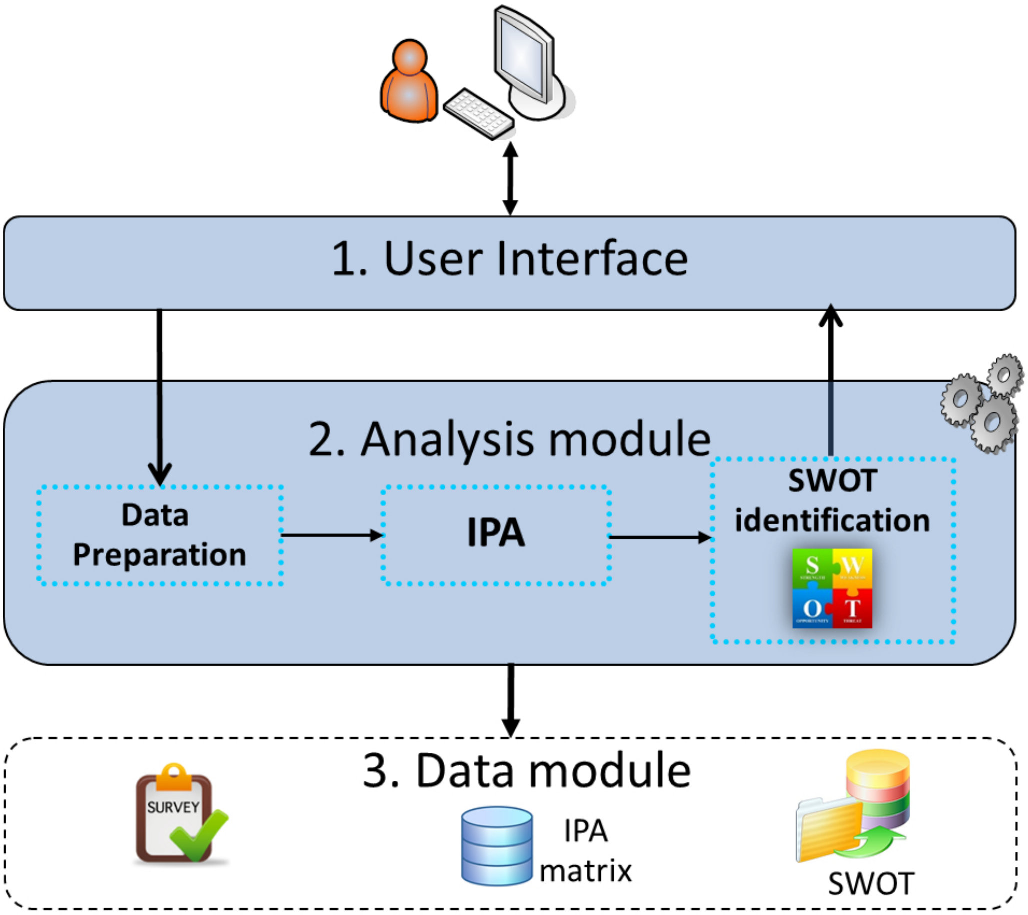


Figure 11.1: The proposed system architecture of IPA based SWOT analysis

1. **User Interface:** provides an interface for users to upload the customer satisfaction survey of their own organisation and that of their competitor. This component is also responsible for displaying the results of SWOT analysis.

2. **Analysis Module:** is responsible for analysis of survey data regarding the IPA based SWOT analysis. In this module, the survey data is first pre-processed and then *performance* is calculated as a mean score of satisfaction while *importance* is calculated by using the selected data mining technique in order to generate two IPA matrices, finally these two IPA matrices are compared to identify SWOT.

3. **Data Module:** is responsible for storing the results produced by the analysis module including the survey data, IPA matrices, and SWOT outcome for the purpose of re-examining these results.

Using the prototype system will reduce time to perform SWOT analysis and therefore facilitate the evaluation of IPA based SWOT analysis on a large scale. Specifically, the participants who take part in the evaluation will work on the prototype system and will be asked to give feedback regarding its operations as well as the SWOT outcome.

11.2.5 Investigating a method to identify macro external factor

As was stated in the limitations section (see Sections 10.4.2), SWOT produced by IPA based SWOT analysis was focused on customers' attitudes and it does not cover a broad range of factors especially the external macro factors such as economy, politics, technology, and trends. To resolve this limitation, the PESTEL analysis should be used in conjunction with SWOT analysis. PESTEL is an analysis framework of macro external factors (including Political, Economic, Social, Technological, Environmental, Legal) that might affect organisations.

Analogous to IPA based SWOT analysis, PESTEL analysis may be conducted based on the staff survey which contains a list of questions that represents all macro external factors asking staff to rate their *importance* and probability of occurrence. *Importance* indicates the impact level of the factor toward the organisation's performance which ranges from 'no impact' to 'very high impact'. Probability of occurrence indicates the possibility that the factor will have any impact on organisation which ranges from low to high probability (Free Management E-books, 2013).

Note that, in the early stage, the *importance* will be directly rated by staff. Later, it may be measured by using statistical or data mining techniques in analysis of a staff survey which is similar to the mechanism for implicitly deriving *importance* from customer satisfaction survey.

11.3 Summary of Thesis

The motivation of this thesis is the lack of a SWOT analysis framework that enables managerial staff in the organisation to systematically generate prioritized SWOT factors corresponding to the customers' perspective. With regard to this motivation, the main idea of the thesis is to integrate IPA into the SWOT analysis framework called IPA based SWOT analysis. The development of IPA based SWOT analysis that involved IPA matrix construction and SWOT factors identification based on the IPA matrix resulted in the following outcomes:

1) thorough analysis of the previous comparative studies of importance measuring techniques
2) establishment of two new techniques for measuring *importance* namely Naïve Bayes and BNs
3) methodology and result of the comparative study of four currently used techniques for measuring *importance* (MLR, OLR, BPNN, and direct-rating scales) with the two new techniques (Naïve Bayes and BNs)
4) development and evaluation of IPA based SWOT analysis

It is expected that the methodology of comparative study can be used to provide guidance as well as evaluation metrics for researchers and practitioners in applying statistical or data mining techniques for measuring *importance*. It is also expected that results of comparative study are able to provide a concrete answer to the argument regarding what is the most appropriate technique for measuring *importance*. Additionally, it is expected that two new techniques introduced for measuring *importance* in this study yield some interesting results and highlight a new research area of IPA.

Finally, the author has confidence in the IPA based SWOT analysis as it was carefully developed and evaluated by means of survey research method. Hence, the author expected that the IPA based SWOT analysis can be used practically for providing new and useful information for business improvement which facilitates an organisation to efficiently formulate strategic planning for maintaining or enhancing customer satisfaction. Another expectation is that the IPA based SWOT analysis can make the best use of data from customer satisfaction surveys that the organisation generally collects. On top of that, the author expected that IPA based SWOT analysis will become a best practice to conduct SWOT analysis after conducting further research work in the near future.

Appendix A

Replication of the IPA based Mutual Information

The main goal of this experiment is to compare an IPA based Mutual Information (MI) replicated in this study with the IPA result of the referenced paper in order to ensure that the method for calculating *importance* and *performance* is implemented correctly in this study. The MI will be applied to dataset A described in Section 5.3, Chapter 5, which is the same dataset used in Cugnata and Salini (2013) (henceforth, the referenced paper) so that this comparison is conducted based on the same method and dataset. Detail regarding tools, methodologies as well as result will be described in the next sections.

A.1 Tools

- ***Microsoft Excel 2010*** for analysing data in order to conduct IPA based MI and for constructing the IPA matrix.
- ***Plot Digitizer***[1] version 2.6.3 for estimating values of *importance* and *performance* shown in figure of the referenced paper.

A.2 Methodology

The methodology comprises three main steps. First, the IPA base MI is implemented following the approach proposed by Shieh and Wu (2011) on dataset A. Next, the *importance* and *performance* are extracted from the referenced paper. Finally, the *importance*

[1] http://plotdigitizer.sourceforge.net/

and *performance* computed in this experiment are compared to the *importance* and *performance* of the referenced paper.

- ***Conduct IPA based MI***
  Generally, steps for conducting IPA consists of measuring *importance* and *performance*, and constructing IPA matrix. According to Shieh and Wu (2011), the *performance* are calculated by using the median of respondents while the implicit *importance* are calculated by MI. A detail of steps for conducting IPA based MI is described as follows:

Step 1: Calculate *performance* of each attribute of company's service by computing the median value. For each attribute of a company's service, a blank or zero value is discarded and the rest of the rating values (1-5) are sorted by ascending order. Then the middle points of each company attribute are defined. In the case when the number of records is an even number the median value is the rating value at the middle point otherwise the median is computed by the average of the value of the middle point and the middle point + 1.

Step 2: Calculate the implicit *importance* by quantifying the dependency between the satisfaction of six attributes of company's service and the overall satisfaction of all attributes of company's service. A formula for calculating MI (Shieh and Wu, 2011) of each attribute of company's service is shown in Equation A.1.

$$MI(X_i, Y) = h(X_i) + h(Y) - h(X_i, Y) \tag{A.1}$$

where X_i represents satisfaction (performance) of each attribute of company's service, Y represents overall satisfaction, $h(X_i)$ is the entropy of X_i, $h(Y)$ is the entropy of Y, and $h(X_i, Y)$ is the joint entropy of X_i and Y. Each part of Equation A.1 which are $h(X_i)$, $h(Y)$ and $h(X_i, Y)$ shown in the Equation A.2-A.4 respectively.

Let $P(X_i = a)$ be the probability of each attribute of company's service X_i with the satisfaction level corresponding to the 5 Likert scale $a = 1, 2, \ldots, 5$, $h(X_i)$ can be computed as Equation A.2, where $P(X_i = a) > 0$. In case $P(X_i = a) = 0$, $log_2 P(X_i = a)$ is set to zero.

$$h(X_i) = -\sum_{a=1}^{5} P(X_i = a)\, log_2 P(X_i = a) \tag{A.2}$$

Let $P(Y = b)$ be the probability of the overall customer satisfaction Y with the satisfaction level corresponding to the 5 Likert scale $b = 1, 2, \ldots, 5$, $h(Y)$

can be computed as Equation A.3, where $P(Y = b) > 0$. In case $P(Y = b) = 0$, $log_2 P(Y = b)$ is set to zero.

$$h(Y) = -\sum_{b=1}^{5} P(Y = b)\, log_2 P(Y = b) \tag{A.3}$$

Let $P(X_i = a, Y = b)$ be the probability of each attribute of company's service X_i with the satisfaction level a and the overall customer satisfaction with the satisfaction level b. $h(X_i, Y)$ can be computed as Equation A.4, where $P(X_i = a, Y = b) > 0$. In case $P(X_i = a, Y = b) = 0$, $log_2 P(X_i = a, Y = b)$ is set to zero.

$$h(X_i, Y) = -\sum_{a=1}^{5}\sum_{b=1}^{5} P(X_i = a, Y = b)\, log_2 P(X_i = a, Y = b) \tag{A.4}$$

In short, *importance* of each attribute of company's service is computed by calculating attribute entropy $h(X_i)$, overall customer satisfaction entropy $h(Y)$ and joint entropy $h(X_i, Y)$. Since the level of satisfaction a, b represented as 5 Likert scale, five combination of $P(X_i = a)$ and $P(Y = b)$ have to be processed and accumulated in order to compute $h(X_i)$ and $h(Y)$ respectively. For calculating the joint entropy $h(X_i, Y)$, 25 combination of $P(X_i = a, Y = b)$ have to be computed and accumulated.

Step 3: Construct a 2×2 IPA matrix in which the x-axis represents *importance* and y-axis represents *performance*. In addition, means of all *importance* and means of all *performance* are used to divide the matrix into four quadrants. Finally, all *importance* and *performance* of attributes of company's service are plotted on the IPA matrix.

- ***Estimate performance and importance of the referenced paper***
  Since the referenced paper does not provide the actual value of *performance* and *importance*, these values were extracted from the IPA matrix of referenced paper by using a program named Plot Digitizer.

- ***Compare result of two previous steps using statistical test***
  The IPA results of the two previous steps are compared by exploring the position of each attribute of company's service on both IPA matrices. Besides, the T-test is conducted in order to guarantee that the IPA result of this study is similar to the result of referenced paper. Detail of procedure for conducting the T-test can be found in Section A.4.

A.3 IPA based MI results

According to steps for conducting the MI based IPA described in section A.2, the *importance* and *performance* for the six attributes of company's service are calculated and shown in Table A.1.

Table A.1: *performance* and *importance* of each attribute of company's service

Company Service Attribute	Performance (Median)	Importance (MI)
A1-Equipment and System	4	0.34
A2-Sales Support	3	0.19
A3-Technical Support	4	0.35
A4-Supplies and Orders	3	0.18
A5-Purchasing Support	4	0.18
A6-Contracts and Pricing	3	0.29
Average	**3.5**	**0.26**

According to Table A.1, the overall mean *performance* is 3.5 and the overall mean *importance* is 0.26 and these numbers were used to create the 2 × 2 IPA matrix. The *importance* and *performance* in Table A.1 were plotted on the IPA matrix as shown in Figure A.1.

The IPA matrix published in Cugnata and Salini (2013) is shown in Figure A.2. Similarly to Figure A.1, the horizontal line is the overall mean *performance* and the vertical line is the overall mean *importance*.

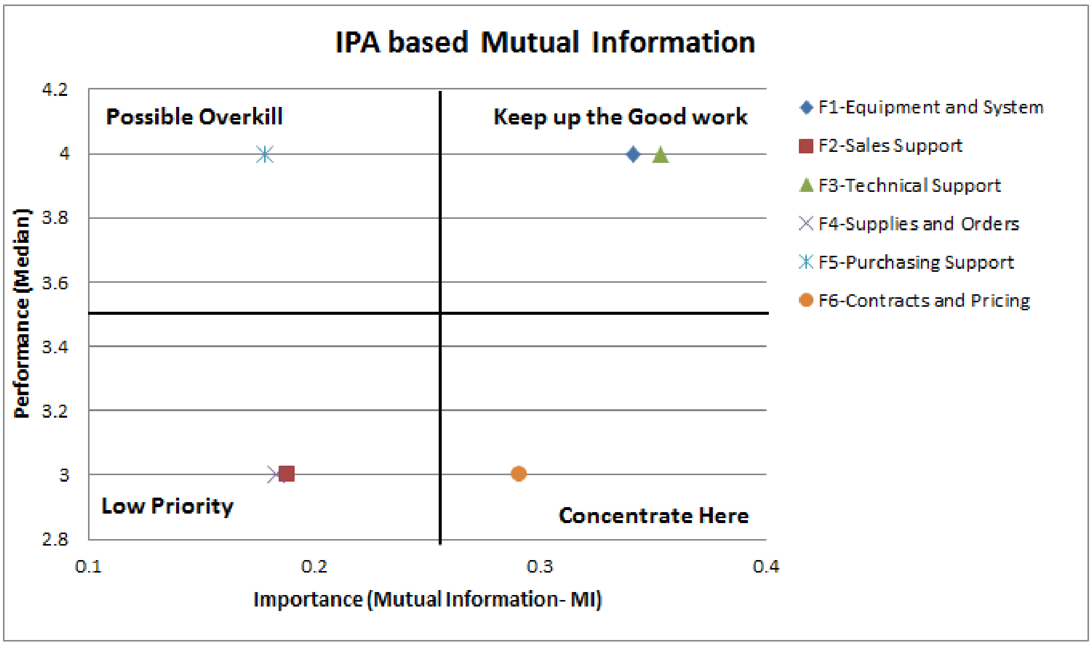


Figure A.1: IPA based MI matrix from experiment.

The comparison of the IPA results in Figure A.1 and Figure A.2 is shown in Table A. 2. This table compares the specified quadrant for each attribute of company's service of

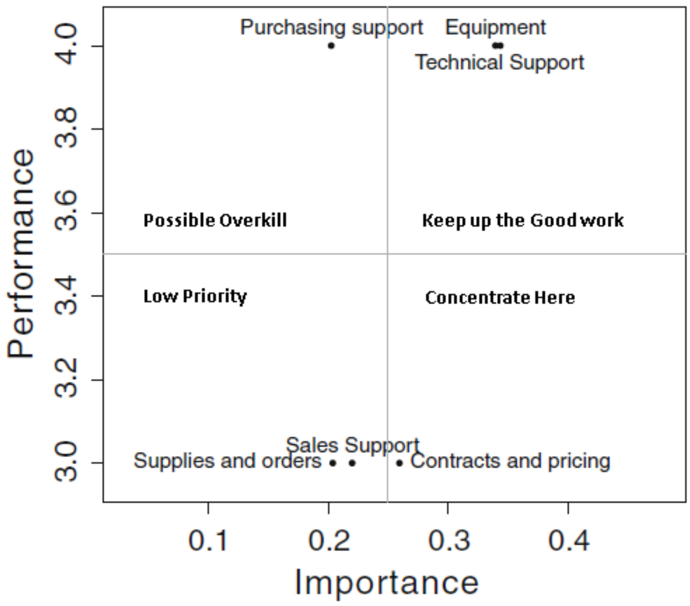


Figure A.2: IPA based MI matrix from referenced paper (Cugnata and Salini, 2013)

IPA result of this study with the IPA result of referenced paper in which '**K**' stands for "***keep up the good work***", '**P**' stands for "***possible overkill***", '**L**' stands for "***Low priority***"and '**C**' stands for "***Concentrate here***".

Table A.2: Comparison of IPA result for this study with that published in referenced paper

Company Service Attribute	**This study**	**Referenced paper** Cugnata and Salini (2013)
A1-Equipment and System	K	K
A2-Sales Support	L	L
A3-Technical Support	K	K
A4-Supplies and Orders	L	L
A5-Purchasing Support	P	P
A6-Contracts and Pricing	C	C

Table A.2 shows that the result of this study is similar to the result of referenced paper for all attributes of company's service. From this result, it can be reasonably concluded that the MI based IPA has been correctly implemented in this study.

In order to guarantee that the result of this study is similar to the result of the referenced paper the statistical test is conducted in section A.4. Subsequently, the *performance* and *importance* of referenced paper were extracted from Figure A.2 by using Plot Digitizer the result are shown in Table A.3.

Table A.3: Estimated *performance* and *importance* from Figure A.2

Company Service Attribute	Performance (Median)	Importance (MI)
F1-Equipment and System	4	0.34
F2-Sales Support	3	0.22
F3-Technical Support	4	0.34
F4-Supplies and Orders	3	0.20
F5-Purchasing Support	4	0.20
F6-Contracts and Pricing	3	0.26
Average	**3.5**	**0.26**

A.4 Statistical test of IPA based MI results

The T-test is conducted in this section in order to compare the difference between the means of the two datasets. Regarding Table A.1 and Table A.3, only the *importance* are selected for comparison because the *performance* of these datasets are obviously the same. Therefore, the input data for T-test are selected from Importance column of Table A.1 and Table A.3, and shown in Table A.4.

Table A.4: The *importance* from difference sources

***importance* (This study)**	***importance* (Referenced paper)**
0.34	0.34
0.19	0.22
0.35	0.34
0.18	0.20
0.18	0.20
0.29	0.26

Step 1: Establish the hypotheses. The aim of this test is to observe that the *importance* of this experiment either similar or different to the *importance* of the referenced paper. Therefore, the null hypothesis and the hypothesis are stated as follow:

H_0: $\mu_1 = \mu_2$ (means of the two groups are the same)

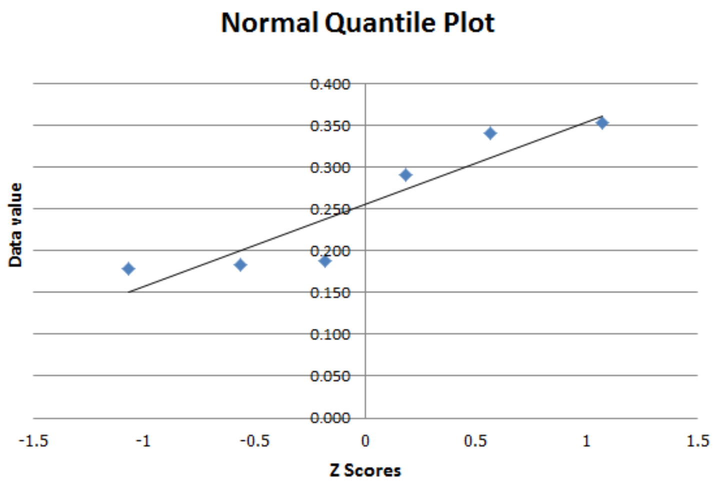
H_1: $\mu_1 \neq \mu_2$ (means of the two groups are difference)

Step 2: Preparing data for analysis. *importance* of each group are formatted in suitable format for Microsoft Excel 2010 as shown in Table A.4.

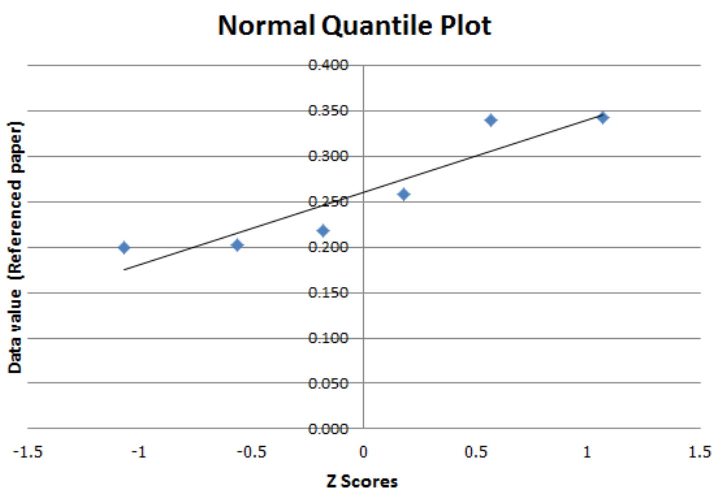
Step 3: Checking the assumptions of T-test. Before conducting the T-test, two assumptions - namely normality distribution and equality of variance have to be checked for the test to be accurate. Methods of testing the two assumptions are described in the following section:

In this study, the normality distribution is visually checked using the normal quantile plot. The normal quantile plots of value in each group are drawn in Figure A.

3(a) and A.3(b). All resulting plots are approximately a straight line, then it is plausible that the all datasets are normally distributed.



(a) The normal quantile plot of *importance*(This study)



(b) The normal quantile plot of *importance* (Referenced paper)

Figure A.3: The normal quantile plot of *importance* from different sources

Additionally, The straightness of the normal quantile plot can be measured by calculating the correlation coefficient (r) of the points in the plot. Formally, the hypothesis of normality is rejected if value of r less than the critical value for the quantile plot correlation coefficient for normality at level of significance (α). The result of correlation coefficient test for normality is shown in Table A.5.

The correlation coefficients from the normal quantile plot of two groups of *importance* are 0.927 and 0.928 respectively. These values are greater than 0.889 which is the critical value of correlation coefficient test for normality corresponding n=6 and $\alpha = 0.05$ (see Table A.5). Therefore, the hypothesis of normality cannot be rejected.

Table A.5: The result of correlation coefficient test for normality of the *importance*

Dataset	**Correlation Coefficient (r)**	**Critical Value Using $\alpha = 0.05$**	**Result**
importance (This Study)	0.927	0.889	Cannot reject the hypothesis of normality
importance (Referenced paper)	0.928	0.889	Cannot reject the hypothesis of normality

For the test of homogeneity of variances, the Levene's test is conducted to check this assumption before performing T-test. The Levene's test checks whether the variances of sample groups are statistically different. The test hypotheses are:

H_0: variances of the two groups are the same

H_1: variances of the two groups are difference

With regard to the Excel sheet for conducting Levene's test[2], the result of Levene's test is shown in Table A.6.

Table A.6: Levene's Statistic of two groups of *importance*

Source of Variation	**SS**	**df**	**MS**	**Levene's Statistic**	**Critical Value (α=0.05)**	**p-value**
Between Groups	0.001	1	0.001	0.952	4.965	0.352
Within Groups	0.012	10	0.001			

Referring to Table A.6, the Levene's statistic (0.952) is less than the critical value at 5% level of significance and p-value (0.352) is less than 0.05. Therefore, the null hypothesis cannot be rejected. The variances of the two groups are equal at a 95% confidence level.

Step 4: Computing the test statistic of T-test. After the assumptions have been ascertained then the independent T-test is conducted by using Data Analysis Tool of Microsoft Excel. In addition, the level of significance is set equal to 0.05. The result of the independent T-test is shown in Table A.7.

Step 5: Making a decision. Consider the critical values and the p-value, the null hypothesis cannot be rejected because t Stat value (-0.104) is less than T critical two-tail (2.228) and p-value (0.919) is greater than 0.05 as shown in Table A.7. That's mean the means of two datasets are not different at the 95% confidence level. With regard to this result, it can be reasonably concluded that the result of this study is similar to the result of the referenced paper which can be inferred that the MI based IPA has been correctly implemented in this study.

[2]www.stat.ufl.edu/ winner/computing/excel/levene.xls

Table A.7: T-test: Two-Sample Assuming Equal Variances

	***importance* (This study)**	***importance* (Referenced paper)**
Mean	0.256	0.260
Variance	0.007	0.004
Observations	6	6
Pooled Variance	0.006	
Hypothesized Mean Difference	0	
df	10	
t Stat	-0.104	
P(T$\leq$ t) two-tail	0.919	
t Critical two-tail	2.228	

Appendix B

Comparative Analysis of Classification Algorithms using WEKA

The main goal of this experiment is to verify that WEKA is installed and configured correctly so it can be used for further experiments. The paper published by Rahman and Afroz (2013) is selected as the referenced paper because this paper conducts the comparison of various classifiers using WEKA on the UCI Machine Learning Repository datasets. Additionally, two classifiers used in Rahman and Afroz (2013) are similar to the classifiers to be used in this research namely BPNN and BNs. More detail regarding dataset, tool, methodology and result will be described in next sections.

B.1 Dataset and Tool

Rahman and Afroz (2013) focus on the analysis of diabetes data and compare performance of different classification techniques in three data mining tools named WEKA, TANAGRA and MATLAB. Therefore, WEKA 3.6.10 and Pima Indians Diabetes, one of the UCI Machine Learning Repository datasets [1], are selected to conduct the experiment.

The Pima Indians Diabetes dataset contains personal data (e.g., age, number of times pregnant) and the results of medical examinations (e.g., blood pressure, body mass index) which are the input attributes and the diagnostic of diabetes is an output attribute which has two possible values (0 or 1) indicated that diabetes test is negative or positive. The details of this dataset are shown in Table B.1.

[1] http://archive.ics.uci.edu/ml/datasets.html

Table B.1: Detail of Pima Indians Diabetes Dataset

Number of Instances	Number of Attributes	Number of Classes	Attribute Type	Missing attributes values
768	9	2	Integer, Real	None

B.2 Methodology

In this section, the classifiers in the experiment are specified with their parameter setting. Additionally, the evaluation method and the performance metric are also specified in order to measure the performance of each classifiers.

- ***Classifiers***
  Three classifiers in WEKA namely J48graft, Multilayer perceptron (MLP), and BayesNet are selected from five classifiers used in Rahman and Afroz's paper. Default values of parameter are assigned to all parameters of each classifier as shown in Table B.2. Table B.2 also shows that all parameters of classifiers in this study have been assigned similarly to the parameters of classifiers described in Rahman and Afroz (2013).

Table B.2: Detail of Classifiers

Classifier Name	This study (Default value of WEKA)	Rahman and Afroz (2013)
J48graft	confidenceFactor = 0.25 minNumObj = 2 subtreeRaising = True unpruned = False	confidenceFactor = 0.25 minNumObj = 2 subtreeRaising = True unpruned = False
MLP	learningRate = 0.3 momentum = 0.2 randomSeed = 0 validationThreshold = 20 number of Epochs = 500 hiddenLayers = a	learningRate = 0.3/0.15 momentum = 0.2 randomSeed = 0 validationThreshold = 20 number of Epochs = 500
BayesNet	estimator =SimpleEstimator search algorithm = K2 ADTree = false	estimator =SimpleEstimator search algorithm = K2 ADTree = false

- ***Evaluation method***
  The evaluation is conducted both in training and testing phase. In the training phase, the whole dataset is used as training and testing data. In the testing phase, Rahman and Afroz (2013) apply the Percentage Split (also called Holdout method) mode. In more detail, the 66% data is used for training and the remaining data is for testing purposes therefore 507 instances are treated as training data and 261 instances are treated as testing data. Since the goal of this study is to compare

performance with the result published in Rahman and Afroz (2013), the Percentage Split is also used as test mode in this study with the option "preserve order for % split" to ensure that WEKA is randomly selected the same set of training and testing data used in (Rahman and Afroz, 2013).

- ***Performance metrics***
  Accuracy measured as percentage of correctly classified instances is used to evaluate performance of classifiers on training and testing data. Besides, others performance metrics derived from the confusion matrix like True-Positive Rate (TP rate), False-Positive rate (FP rate), and Precision, Recall, F-measure and ROC area are also used to measure performance of classifiers on training and testing data. Those set of performance metrics are also used in Rahman and Afroz (2013), thus they can use to compare the classifiers' result of this study and Rahman and Afroz (2013).

B.3 Result

After the three classifiers are applied on Diabetes dataset in training and testing phase, their performances are measured regarding metrics described in the previous section and reported in Table B.3, Table B.4 and Table B.5.

Table B.3: Accuracy of different classifiers

Classifier	Phase	Accuracy (%) of this study	Accuracy (%) of Rahman and Afroz (2013)
J48graft	Train	84.11	84.11
	Test	78.54	78.54
MLP	Train	80.60	80.60
	Test	77.78	77.78
BayesNet	Train	78.26	78.26
	Test	79.69	79.69

Table B.4: Different performance metrics in training phase

Metrics	Classifiers of this study			Classifiers of Rahman and Afroz (2013)		
	J48graft	MLP	BayesNet	J48graft	MLP	BayesNet
TP rate	0.841	0.806	0.783	0.841	0.806	0.783
FP rate	0.241	0.191	0.26	0.241	0.191	0.26
Precision	0.842	0.819	0.783	0.842	0.819	0.783
Recall	0.841	0.806	0.783	0.841	0.806	0.783
F-measure	0.836	0.809	0.783	0.836	0.809	0.783
ROC area	0.888	0.872	0.851	0.888	0.872	0.851

The results from Table B.3 show that the accuracy of this study is similar to accuracy published by Rahman and Afroz (2013) for all classifiers. Besides the accuracy, the

Table B.5: Different performance metrics in testing phase

Metrics	**Classifiers of this study**			**Classifiers of Rahman and Afroz (2013)**		
	J48graft	**MLP**	**BayesNet**	**J48graft**	**MLP**	**BayesNet**
TP rate	0.785	0.778	0.797	0.785	0.778	0.797
FP rate	0.189	0.306	0.253	0.189	0.306	0.253
Precision	0.816	0.774	0.799	0.816	0.774	0.799
Recall	0.785	0.778	0.797	0.785	0.778	0.797
F-measure	0.792	0.776	0.798	0.792	0.776	0.798
ROC area	0.803	0.813	0.848	0.803	0.813	0.848

results of this study measured by other performance metrics such as True-Positive Rate, False-Positive Rate, Precision, Recall, F-Measure, and ROC Area are similar to the results published by Rahman and Afroz (2013) as shown in Table B.4 and Table B.5. Based on these comparison results, it can be concluded that WEKA 3.6.10 is installed and configured correctly. Thus, this tool can be used for further data analysis for conducting IPA for SWOT analysis.

Appendix C

Assumption checking of the multiple linear regressions on 3 datasets

To ensure that the result of MLR is valid several assumptions of MLR such as normality, linearity were tested. Assumption testing result for each dataset were summarized in Table C.1 and explained in the following sections.

Table C.1: Summary of assumption checking of MLR on 3 datasets

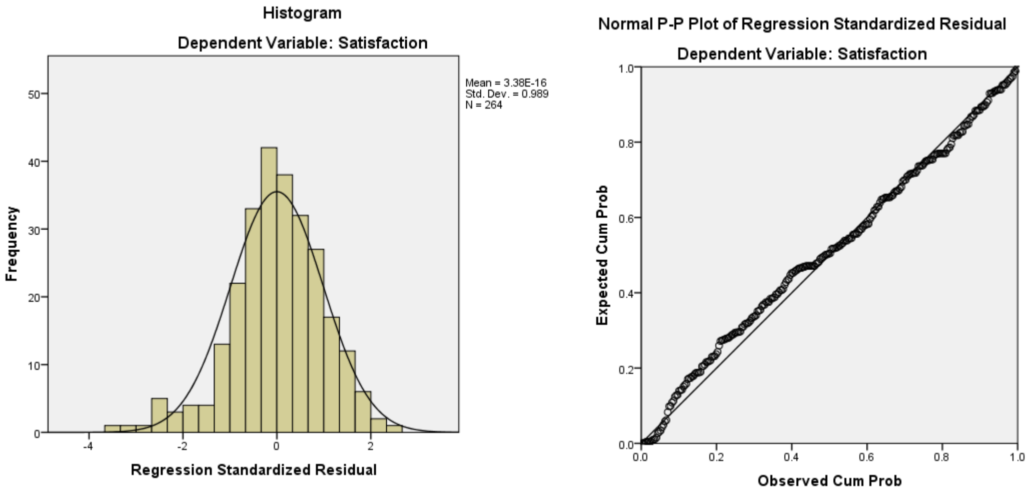
Assumptions	**Dataset**		
	A	**B**	**C**
Normality of residual	✓	(2)×	✓
Linearity	✓	✓	✓
Homoscedasticity of residual	(1)×	(1)×	✓
Independence of residual	✓	✓	✓
Multicollinearity	✓	✓	✓

With regard to Table C.1, a symbol ✓ indicates that the assumption was met and a symbol × indicates that the assumption was violated. According to the table it shows that three assumptions which are linearity, independence of residual, and no multicollinearity were met across three datasets. Whereas the other two assumptions: normality and homoscedasticity of residual were not met across three datasets. However, no further steps were performed to correct these assumptions and the rationales to support this decision are: (1) the regression is fairly robust to violation of homoscedasticity of residual. Beside violation of this assumption does not affect the regression coefficients which were the main focus of this study(Keith, 2014) and (2) the deviation from normality of residuals tested on dataset B was not the case to be concerned since the impact of departure from zero kurtosis is diminished in reasonably large sample (200+ cases: Tabachnick et al. (2001)).

C.1 Assumption checking of the multiple linear regressions on dataset A

Assumption 1: Normally distributed residual

The assumption of normality was visually tested by observing the histogram and normal P-P plot of the standardized residual. Figure C.1 showed that the histogram of residual appeared to look like a bell and the normal P-P plot was approximately a straight line, then it is plausible that the residual was normally distributed.



(a) Histogram plot of residual (b) Normality plot of residual

Figure C.1: Histogram and normality plots of residual (dataset A)

Consistent with the histogram and normality plot, the skew and kurtosis of unstandardized residual shown in Table C.2 were fell within a reasonable range to accept that data is reasonably close to normal which is -1.0 to $+1.0$ (George and Mallery, 2003).

Table C.2: Descriptive statistic of unstandardized residual (dataset A)

	N	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
Unstandardized Residual	264	0.000	0.7769	-0.533	0.15	1.032	0.299

Assumption 2: Linear relationship between independent and dependent variables

To check linearity assumption, a scatterplot of the standardized residual and standardized predicted value was drawn. The residual plot for predicted values of an outcome variable (Satisfaction) against the residuals (see Figure C.2) showed that the dots are not constantly spread over the horizontal line; however, there is no sign of any curve pattern. Therefore, it cannot conclude that the linearity assumption was violated.

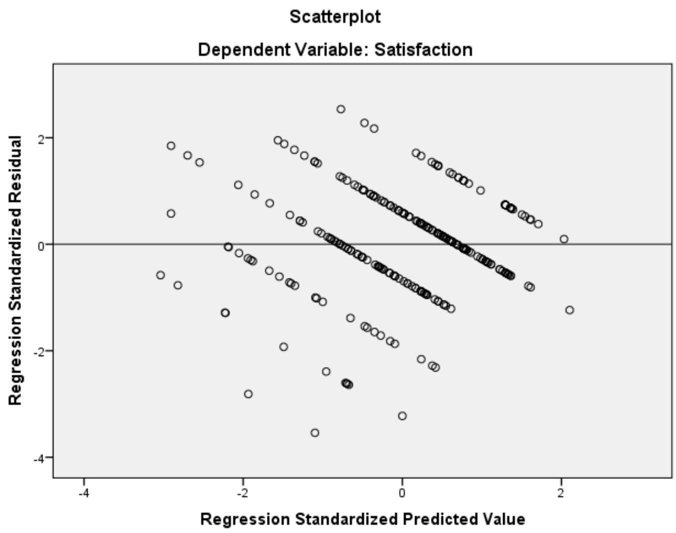


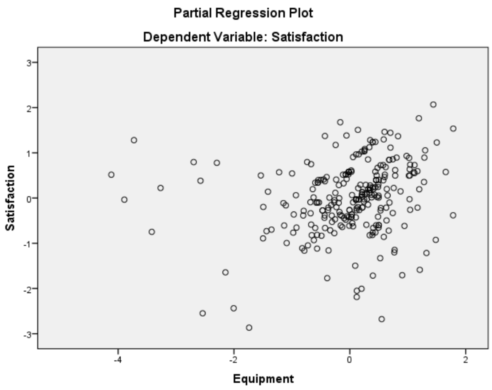
Figure C.2: Scatterplot of residual for the relationship between six independent variables and the Satisfaction (dataset A)

Non-linear relationships can also be detected by looking at the partial regression plots which are the scatterplots of the residuals of the outcome variable versus individual independent variables. All partial regression plots (Figure C.3) showed a fairly random pattern and a positive relationship to Satisfaction which indicated the presence of linearity.

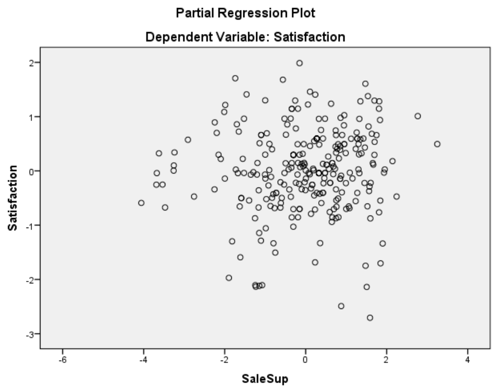
Apart from scatterplot and partial regression plot, it is useful to draw the scatterplot of the outcome variable against each independent variable for the examination of linearity. These scatterplots were drawn along with the curve fit of the linear model and non-linear models including: quadratic and cubic (see Figure C.4). The curve fit lines for each model of all scatterplots are appeared to be similar which indicated that the linear model is fit to the data. However, the curve fit lines for the non-linear models of the two independent variables named Equipment and TermCond are slightly different from that of the linear model.

To confirm the result of the graphical methods for evaluating linearity, for each independent variable the R-square of the four models are statistically compared through the test of equality of several independent correlation coefficients (r). Regarding a rule of thumb, a relationship is linear if the difference between the linear correlation coefficient (r) and the non-linear correlation coefficient is not significantly different.

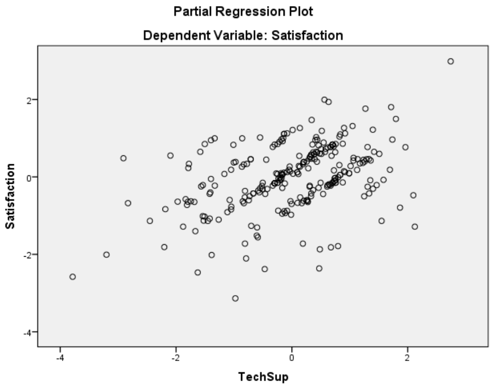
Table C.3 - C.8 show the model summary of linear/non-linear function of each independent variable against the Satisfaction associated to the scatterplots shown in Figure C.4. Given R-square from Table C.3 - C.8, the test for equality of several independent correlation coefficients was conducted following an approach described



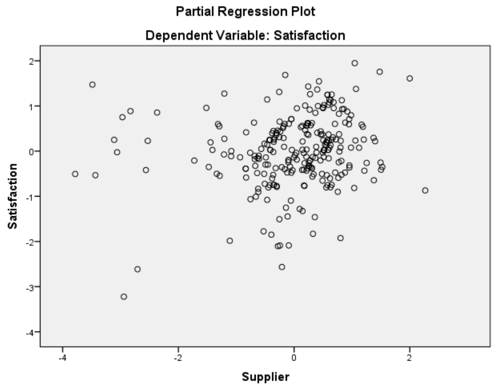
(a) Partial regression plot of Equipment against Satisfaction



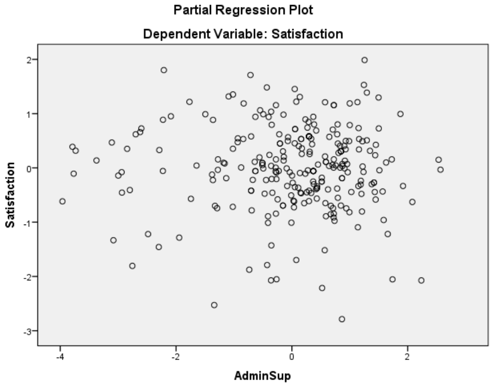
(b) Partial regression plot of SaleSup against Satisfaction



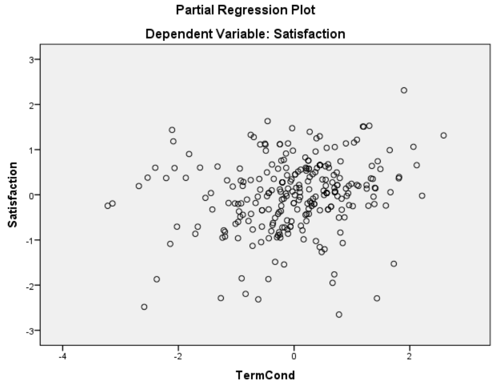
(c) Partial regression plot of TechSup against Satisfaction



(d) Partial regression plot of Supplier against Satisfaction



(e) Partial regression plot of AdminSup against Satisfaction



(f) Partial regression plot of TermCond against Satisfaction

Figure C.3: Partial regression plots of each independent variable against Satisfaction (dataset A)

in Kenny (1987). Then, the result of the test for equality of several independent correlation coefficients across six independent variables is reported in Table C.9.

Table C.9 shows that there is no significant difference among correlations of the models for each independent variable. Therefore, it's reasonable to conclude that

Table C.3: Model summary corresponded to the scatterplot of Equipment versus Satisfaction

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.155	48.068	1	262	0.000
Quadratic	0.204	33.353	2	261	0.000
Cubic	0.276	33.070	3	260	0.000

Table C.4: Model summary corresponded to the scatterplot of SaleSup versus Satisfaction

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.065	18.269	1	262	0.000
Quadratic	0.073	10.304	2	261	0.000
Cubic	0.121	11.976	3	260	0.000

Table C.5: Model summary corresponded to the scatterplot of TechSup versus Satisfaction

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.312	118.656	1	262	0.000
Quadratic	0.313	59.561	2	261	0.000
Cubic	0.317	40.184	3	260	0.000

Table C.6: Model summary corresponded to the scatterplot of Supplier versus Satisfaction

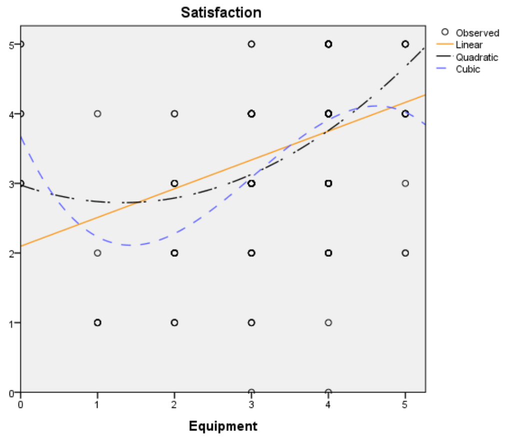
Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.136	41.137	1	262	0.000
Quadratic	0.165	25.776	2	261	0.000
Cubic	0.175	18.337	3	260	0.000

Table C.7: Model summary corresponded to the scatterplot of AdminSup versus Satisfaction

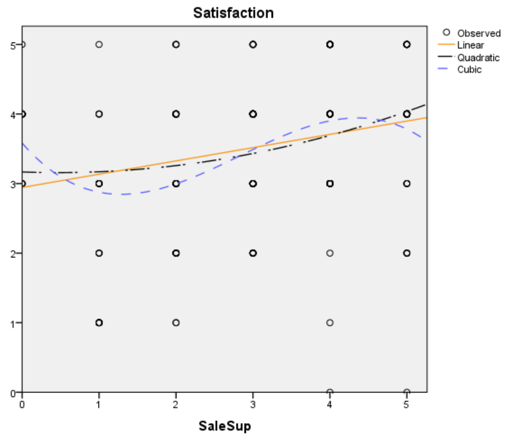
Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.042	11.403	1	262	0.001
Quadratic	0.052	7.111	2	261	0.001
Cubic	0.094	8.96	3	260	0.000

Table C.8: Model summary corresponded to the scatterplot of TermCond versus Satisfaction

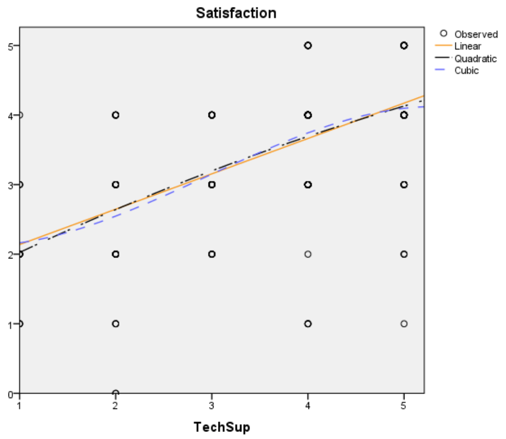
Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.163	51.025	1	262	0.000
Quadratic	0.183	29.17	2	261	0.000
Cubic	0.236	26.705	3	260	0.000



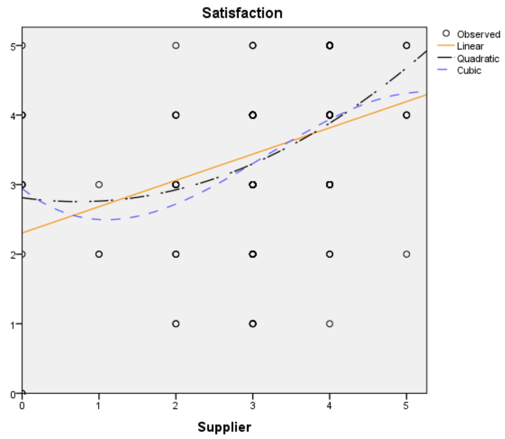
(a) Curve estimation of Equipment against Satisfaction



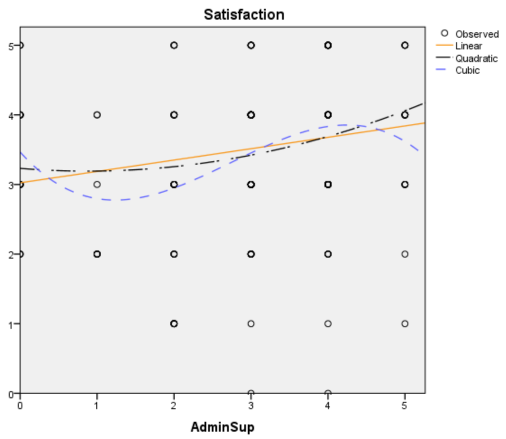
(b) Curve estimation of SaleSup against Satisfaction



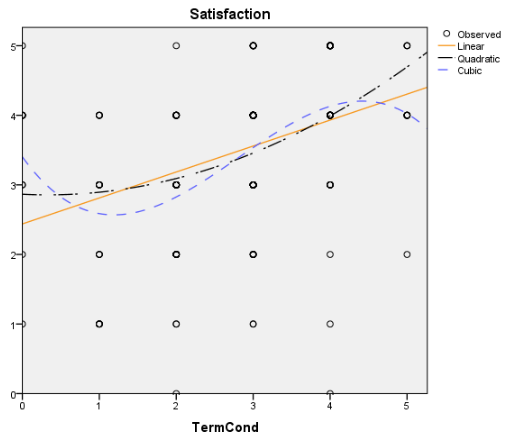
(c) Curve estimation of TechSup against Satisfaction



(d) Curve estimation of Supplier against Satisfaction



(e) Curve estimation of AdminSup against Satisfaction



(f) Curve estimation of TermCond against Satisfaction

Figure C.4: Scatterplot with model curve fit of each independent variable against Satisfaction (dataset A)

the relationship between each independent variable and the Satisfaction is linear.

Table C.9: Chi-square statistic for the difference test of several independent correlations (dataset A)

Independent variable	χ^2	**df**	**p-value**	**Equality**
Equipment	3.693	2	0.158	Yes
SaleSup	1.576	2	0.455	Yes
TechSup	0.006	2	0.997	Yes
Supplier	0.487	2	0.784	Yes
AdminSup	1.706	2	0.426	Yes
TermCond	1.455	2	0.483	Yes

Assumption 3: Homoscedasticity of residuals

The assumption is that the residuals at each level of the dependent variables had the same variance. There are graphical and non-graphical methods for checking homoscedasticity. For the graphical method, the scatterplot of the standardized residual and standardized predicted value as shown in Figure C.2 was reobserved. Although the dots were not constantly spread over the horizontal line, a funnel shaped pattern did not appear. It may be inferred that this assumption was met.

A non-graphical method using Breusch-Pagan test suggested that the variance of the residuals is not homogenous. The null hypothesis of Breusch-Pagan test is homoscedasticity. In this case, the null hypothesis is rejected (BP =23.95, df = 1, p-value =0.000).

Note that steps for conducting Breusch-Pagan test using SPSS is referred to Pryce (2002) and the validity of this test is depended on the normality of residual. The histogram and normality plot of residuals shown in Figure C.1 suggested that the residuals are fairly normal; therefore, the test result was reliable.

Even though this assumption was not met, further steps were not performed to correct this assumption because the regression is fairly robust to violation of this assumption. Beside violation of this assumption does not affect the regression coefficients which were the main focus of this study (Keith, 2014).

Assumption 4: Independence of residual (error)

The assumption is that the error is not correlated with any independent variables. The independence of residuals was checked through the Durbin-Watson statistic of the regression model. A conventionally acceptable range of this statistic is 1 - 3 in which the value close to 2 indicates no autocorrelation (Field, 2009). In this case, Durbin-Watson statistic was 1.99 indicated the lack of autocorrelation thereby the assumption was met.

Assumption 5: Multicollinearity

Under the assumption of no multicollinearity, the independent variables are not closely linearly related to each other. The degree of multicollinearity can be detected by inspecting the VIF statistic provided by SPSS. A VIF of 1 means that there is no correlation among the independent variables while VIF higher than 1 indicates a presence of multicollinearity among them. As a rule of thumb, VIF higher than 10 indicates the sign of serious multicollinearity (Field, 2009). The VIF for each independent variable was greater than 1 but lower than 10 (see Table C.10) hence it's can conclude that multicollinearity was present but there is no sign of serious multicollinearity which requires correction (such as removing redundant variables) within the dataset A.

Table C.10: Collinearity Statistics (VIF) of six independent variables (dataset A)

Independent variable	**Collinearity Statistics (VIF)**
Equipment	1.212
SaleSup	1.188
TechSup	1.251
Supplier	1.187
AdminSup	1.157
TermCond	1.348

C.2 Assumption checking of the multiple linear regressions on dataset B

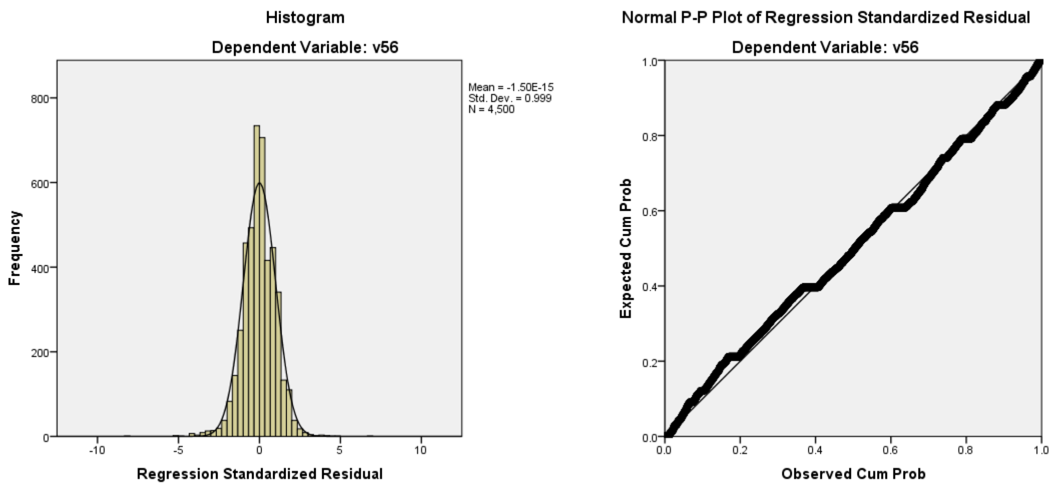
Assumption 1: Normally distributed residual

The assumption of normality was visually tested by observing the histogram and normal P-P plot of the standardized residual. Figure C.5 showed that the histogram of residual appeared to look like a bell but rather peak and the normal P-P plot was approximately a straight line, then it is plausible that the residual was normally distributed.

In addition to the histogram and normality plot, the skew and kurtosis of unstandardized residual were also observed to check the normality of residuals. Table C. 11 showed that only the skew was fell within reasonable range to accept that data is reasonably close to normal which is -1.0 to $+1.0$ (George and Mallery, 2003). Although the kurtosis was higher than 1.0, the impact of departure from zero kurtosis is diminished in reasonably large sample (200+ cases) (Tabachnick et al., 2001). Thus, the deviation from normality of residuals was not the case to be concerned.

Table C.11: Descriptive statistic of unstandardized residual (dataset B)

	N	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
Unstandardized Residual	4500	0.000	0.504	-0.273	0.037	2.88	0.073



(a) Histogram plot of residual (b) Normality plot of residual

Figure C.5: Histogram and normality plots of residual (dataset B)

Assumption 2: Linear relationship between independent and dependent variables

To check linearity assumption, a scatterplot of the standardized residual and standardized predicted value was drawn. The residual plot for predicted values of an outcome variable (V56) against the residuals (see Figure C.6) showed that the dots are not constantly spread over the horizontal line; however, there is no sign of any curve pattern. Therefore, it cannot conclude that the linearity assumption was violated.

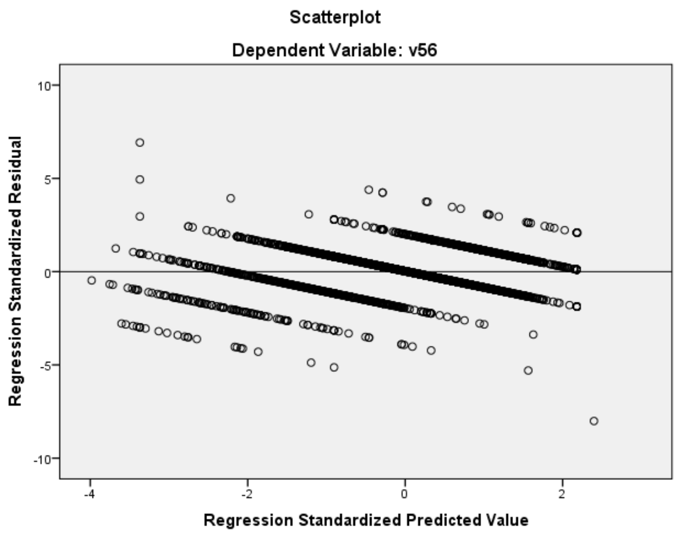
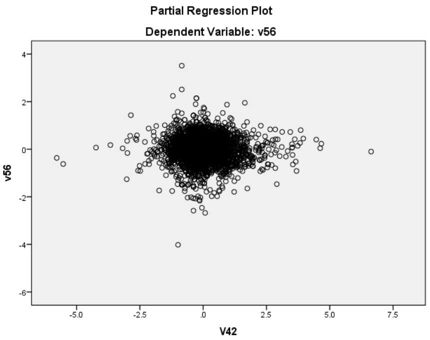
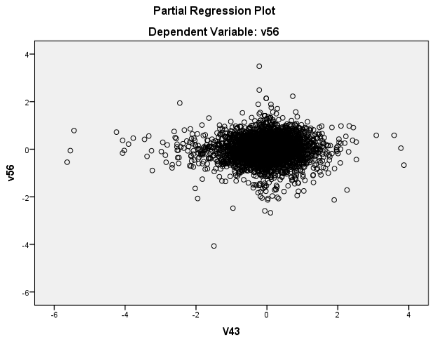


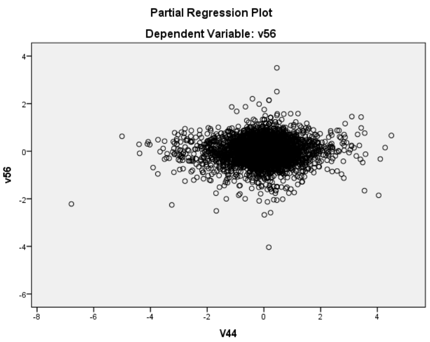
Figure C.6: Scatterplot of residual for the relationship between 13 independent variables and the V56 (dataset B)

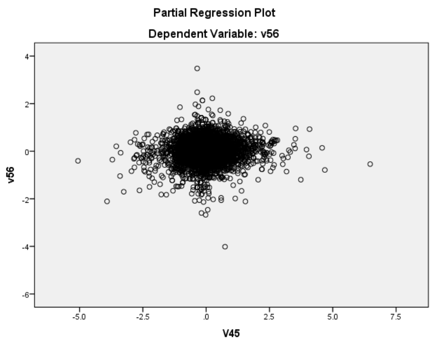
Non-linear relationships can also be detected by looking at the partial regression plots which are the scatterplots of the residuals of the outcome variable versus individual independent variables (V42 - V54). All partial regression plots (Figure C. 7 - C.8) showed a fairly random pattern to V56 which may indicate the presence of linearity.



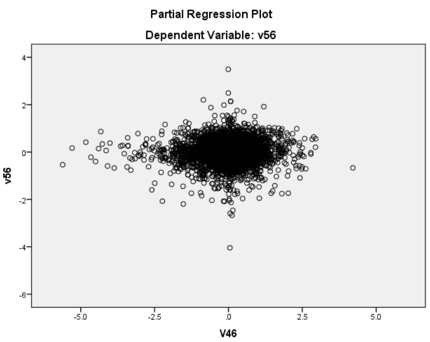


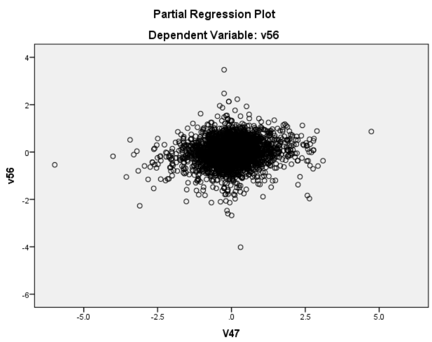
(a) Partial regression plot of V42 against V56 (b) Partial regression plot of V43 against V56





(c) Partial regression plot of V44 against V56 (d) Partial regression plot of V45 against V56





(e) Partial regression plot of V46 against V56 (f) Partial regression plot of V47 against V56

Figure C.7: Partial regression plots of independent variable V42-V47 against V56 (dataset B)

Apart from scatterplot and partial regression plot, it is useful to draw the scatterplot of the outcome variable against each independent variable for the examination of linearity. These scatterplots were drawn along with the curve fit of the linear model and non-linear models including: logarithmic, quadratic and cubic (see Figure C.9 - C.10). For each variable, the curve fit lines for linear model are appeared to be similar to that of quadratic and cubic model in the other word only the curve fit lines for logarithmic model are slightly different from the others.

Since the linearity cannot be confirmed using these graphs, the R-square of the four models for each independent variable shown in Table C.12 - C.24 are statistically compared through the test of equality for several independent correlation coefficients (r). The test for equality of several independent correlation coefficients was conducted following to an approach described in Kenny (1987). Then, the result of the test for equality of several independent correlation coefficients across 13 independent variables is reported in Table C.25.

Table C.12: Model summary corresponded to the scatterplot of V42 versus V56

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.197	1106.156	1	4498	0.000
Logarithmic	0.186	1030.152	1	4498	0.000
Quadratic	0.199	557.093	2	4497	0.000
Cubic	0.199	371.625	3	4496	0.000

Table C.13: Model summary corresponded to the scatterplot of V43 versus V56

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.246	1468.859	1	4498	0.000
Logarithmic	0.232	1356.8	1	4498	0.000
Quadratic	0.247	739.37	2	4497	0.000
Cubic	0.247	492.836	3	4496	0.000

Table C.14: Model summary corresponded to the scatterplot of V44 versus V56

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.293	1866.597	1	4498	0.000
Logarithmic	0.281	1758.288	1	4498	0.000
Quadratic	0.295	941.726	2	4497	0.000
Cubic	0.296	631.63	3	4496	0.000

Table C.15: Model summary corresponded to the scatterplot of V45 versus V56

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.336	2280.634	1	4498	0.000
Logarithmic	0.332	2234.325	1	4498	0.000
Quadratic	0.343	1174.4	2	4497	0.000
Cubic	0.344	785.072	3	4496	0.000

Table C.25 shows that there is no significant difference among correlations of the models for each independent variable. Therefore, it's reasonable to conclude that

Table C.16: Model summary corresponded to the scatterplot of V46 versus V56

Equation	Model Summary				
	R Square	F	df1	df2	p-value
Linear	0.349	2408.765	1	4498	0.000
Logarithmic	0.329	2209.976	1	4498	0.000
Quadratic	0.351	1216.013	2	4497	0.000
Cubic	0.351	811.069	3	4496	0.000

Table C.17: Model summary corresponded to the scatterplot of V47 versus V56

Equation	Model Summary				
	R Square	F	df1	df2	p-value
Linear	0.372	2666.125	1	4498	0.000
Logarithmic	0.367	2604.481	1	4498	0.000
Quadratic	0.378	1369.349	2	4497	0.000
Cubic	0.380	917.897	3	4496	0.000

Table C.18: Model summary corresponded to the scatterplot of V48 versus V56

Equation	Model Summary				
	R Square	F	df1	df2	p-value
Linear	0.339	2308.425	1	4498	0.000
Logarithmic	0.325	2164.941	1	4498	0.000
Quadratic	0.341	1163.588	2	4497	0.000
Cubic	0.342	780.348	3	4496	0.000

Table C.19: Model summary corresponded to the scatterplot of V49 versus V56

Equation	Model Summary				
	R Square	F	df1	df2	p-value
Linear	0.193	1072.92	1	4498	0.000
Logarithmic	0.179	980.881	1	4498	0.000
Quadratic	0.193	538.537	2	4497	0.000
Cubic	0.193	358.962	3	4496	0.000

Table C.20: Model summary corresponded to the scatterplot of V50 versus V56

Equation	Model Summary				
	R Square	F	df1	df2	p-value
Linear	0.286	1799.49	1	4498	0.000
Logarithmic	0.269	1656.898	1	4498	0.000
Quadratic	0.287	903.965	2	4497	0.000
Cubic	0.287	603.323	3	4496	0.000

Table C.21: Model summary corresponded to the scatterplot of V51 versus V56

Equation	Model Summary				
	R Square	F	df1	df2	p-value
Linear	0.311	2026.883	1	4498	0.000
Logarithmic	0.298	1908.444	1	4498	0.000
Quadratic	0.312	1021.957	2	4497	0.000
Cubic	0.313	682.567	3	4496	0.000

Table C.22: Model summary corresponded to the scatterplot of V52 versus V56

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.348	2399.94	1	4498	0.000
Logarithmic	0.343	2349.145	1	4498	0.000
Quadratic	0.354	1233.484	2	4497	0.000
Cubic	0.355	823.322	3	4496	0.000

Table C.23: Model summary corresponded to the scatterplot of V53 versus V56

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.326	2171.462	1	4498	0.000
Logarithmic	0.315	2065.275	1	4498	0.000
Quadratic	0.327	1094.153	2	4497	0.000
Cubic	0.330	737.211	3	4496	0.000

Table C.24: Model summary corresponded to the scatterplot of V54 versus V56

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.337	2285.828	1	4498	0.000
Logarithmic	0.334	2255.471	1	4498	0.000
Quadratic	0.342	1170.232	2	4497	0.000
Cubic	0.344	787.14	3	4496	0.000

Table C.25: Chi-square statistic for the difference test of several independent correlations (dataset B)

Independent variable	χ^2	**df**	**p-value**	**Equality**
V42	1.046	3	0.790	Yes
V43	1.315	3	0.725	Yes
V44	1.115	3	0.774	Yes
V45	0.750	3	0.861	Yes
V46	2.612	3	0.455	Yes
V47	0.803	3	0.849	Yes
V48	1.433	3	0.698	Yes
V49	1.341	3	0.719	Yes
V50	1.822	3	0.610	Yes
V51	1.137	3	0.768	Yes
V52	0.715	3	0.870	Yes
V53	0.980	3	0.806	Yes
V54	0.476	3	0.924	Yes

the relationship between each independent variable and the V56 is linear.

Assumption 3: Homoscedasticity of residuals

The assumption is that the residuals at each level of the dependent variables had the same variance. There are graphical and non-graphical methods for checking homoscedasticity. For the graphical method, the scatterplot of the standardized residual and standardized predicted value as shown in Figure C.6 was reobserved. Although the dots were spread out more at upper level of standardized predicted value than at lower level, a funnel shaped pattern did not appear. It may be inferred that heteroscedasticity was not present.

In contrast with the graphical method, a non-graphical method using Breusch-Pagan test suggested that the variance of the residuals is not homogenous. The null hypothesis of Breusch-Pagan test is homoscedasticity. In this case, the null hypothesis is rejected (BP =173.09, df = 1, p-value = 0.000).

Note that steps for conducting Breusch-Pagan test using SPSS is referred to Pryce (2002) and the validity of this test is depended on the normality of residual. The histogram and normality plot of residuals shown in Figure C.5 suggested that the residuals are fairly normal; therefore, the test result was reliable.

Even though this assumption was not met, further steps were not performed to correct this assumption because the regression is fairly robust to violation of this assumption. Beside violation of this assumption does not affect the regression coefficients which were the main focus of this study (Keith, 2014).

Assumption 4: Independence of residual (error)

The assumption is that the error is not correlated with any independent variables. The independence of residuals was checked through the Durbin-Watson statistic of the regression model. A conventionally acceptable range of this statistic is 1 - 3 in which the value close to 2 indicates no autocorrelation (Field, 2009). In this case, Durbin-Watson statistic was 1.96 indicated the lack of autocorrelation thereby the assumption was met.

Assumption 5: Multicollinearity

Under the assumption of no multicollinearity, the independent variables are not closely linearly related to each other. The degree of multicollinearity can be detected by inspecting the VIF statistic provided by SPSS. A VIF of 1 means that there is no correlation among the independent variables while VIF higher than 1 indicates a presence of multicollinearity among them. As a rule of thumb, VIF higher than 10 indicates the sign of serious multicollinearity (Field, 2009). The VIF for each independent variable was greater than 1 but lower than 10 (see Table C.26) hence it’s can conclude that multicollinearity was present but there is no sign of serious multicollinearity which requires correction (such as removing redundant variables) within the dataset B.

Table C.26: Collinearity Statistics (VIF) of 13 independent variables (dataset B)

Independent variable	**Collinearity Statistics (VIF)**
V42	4.294
V43	5.454
V44	4.244
V45	4.719
V46	5.114
V47	5.921
V48	4.398
V49	3.668
V50	5.230
V51	4.539
V52	5.381
V53	5.470
V54	4.191

C.3 Assumption checking of the multiple linear regressions on dataset C

Assumption 1: Normally distributed residual

The assumption of normality was visually tested by observing the histogram and normal P-P plot of the standardized residual. Figure C.11 showed that the histogram of residual appeared to look like a bell and the normal P-P plot was approximately a straight line, then it is plausible that the residual was normally distributed.

Consistent with the histogram and normality plot, the skew and kurtosis of unstandardized residual shown in Table C.27 were fell within a reasonable range to accept that data is reasonably close to normal which is -1.0 to $+1.0$ (George and Mallery, 2003).

Table C.27: Descriptive statistic of unstandardized residual (dataset C)

	N	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
Unstandardized Residual	155	0.000	0.3717	0.162	0.195	0.603	0.387

Assumption 2: Linear relationship between independent and dependent variables

To check linearity assumption, a scatterplot of the standardized residual and standardized predicted value was drawn. The residual plot for predicted values of an outcome variable (OCS student) against the residuals (see Figure C.12) showed that the dots are not constantly spread over the horizontal line; however, there is no sign of any curve pattern. Therefore, it cannot conclude that the linearity assumption was violated.

Non-linear relationships can also be detected by looking at the partial regression plots which are the scatterplots of the residuals of the outcome variable versus individual independent variables. All partial regression plots (Figure C.13) showed a fairly random pattern and a positive relationship to OCS student which indicated the presence of linearity.

Apart from scatterplot and partial regression plot, it is useful to draw the scatterplot of the outcome variable against each independent variable for the examination of linearity. These scatterplots were drawn along with the curve fit of the linear model and non-linear models including: quadratic and cubic (see Figure C.14). The curve fit lines for each model of all scatterplots are appeared to be similar which indicated that the linear model is fit to the data. However, the curve fit lines for the non-linear models of the two independent variables named Teaching and CompFac are slightly different from that of the linear model.

To confirm the result of the graphical methods for evaluating linearity, for each independent variable the R-square of the four models for each independent variable shown in Table C.28 - C.33 are statistically compared through the test of equality of several independent correlation coefficients (r). The test for equality of several independent correlation coefficients was conducted following an approach described in Kenny (1987). Then, the result of the test for equality of several independent correlation coefficients across six independent variables is reported in Table C.34.

Table C.28: Model summary corresponded to the scatterplot of Teacher versus OCS student

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.319	71.704	1	153	0.000
Logarithmic	0.311	68.988	1	153	0.000
Quadratic	0.319	35.674	2	152	0.000
Cubic	0.319	35.674	2	152	0.000

Table C.29: Model summary corresponded to the scatterplot of Teaching versus OCS student

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.413	107.439	1	153	0.000
Logarithmic	0.378	93.056	1	153	0.000
Quadratic	0.433	57.955	2	152	0.000
Cubic	0.429	57.076	2	152	0.000

Table C.30: Model summary corresponded to the scatterplot of Admin versus OCS student

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.298	64.811	1	153	0.000
Logarithmic	0.289	62.335	1	153	0.000
Quadratic	0.298	32.227	2	152	0.000
Cubic	0.298	21.359	3	151	0.000

Table C.31: Model summary corresponded to the scatterplot of CompFac versus OCS student

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.335	76.998	1	153	0.000
Logarithmic	0.301	66.031	1	153	0.000
Quadratic	0.339	38.958	2	152	0.000
Cubic	0.341	26.093	3	151	0.000

Table C.32: Model summary corresponded to the scatterplot of xActivity versus OCS student

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.467	134.029	1	153	0.000
Logarithmic	0.457	128.574	1	153	0.000
Quadratic	0.468	66.781	2	152	0.000
Cubic	0.468	66.849	2	152	0.000

Table C.33: Model summary corresponded to the scatterplot of AddService versus OCS student

Equation	**Model Summary**				
	R Square	F	df1	df2	p-value
Linear	0.341	79.144	1	153	0.000
Logarithmic	0.326	74.129	1	153	0.000
Quadratic	0.343	39.621	2	152	0.000
Cubic	0.343	39.663	2	152	0.000

Table C.34: Chi-square statistic for the difference test of several independent correlations (dataset C)

Independent variable	χ^2	**df**	**p-value**	**Equality**
Teacher	0.012	3	1.000	Yes
Teaching	0.499	3	0.919	Yes
Admin	0.016	3	0.999	Yes
CompFac	0.273	3	0.965	Yes
xActivity	0.024	3	0.999	Yes
AddService	0.052	3	0.997	Yes

Table C.34 shows that there is no significant difference among correlations of the models for each independent variable. Therefore, it's reasonable to conclude that the relationship between each independent variable and the OCS student is linear.

Assumption 3: Homoscedasticity of residuals

The assumption is that the residuals at each level of the dependent variables had the same variance. There are graphical and non-graphical methods for checking homoscedasticity. For the graphical method, the scatterplot of the standardized residual and standardized predicted value as shown in Figure C.12 was reobserved. Although the dots were not constantly spread over the horizontal line, a funnel shaped pattern did not appears. It may be inferred that this assumption was met.

Consistent with the graphical method, a non-graphical method using Breusch-Pagan test suggested that the variance of the residuals is homogenous. The null hypothesis of Breusch-Pagan test is homoscedasticity. In this case, the null hypothesis is retained (BP = 0.742, df = 1, p-value = 0.389).

Note that steps for conducting Breusch-Pagan test using SPSS is referred to Pryce (2002) and the validity of this test is depended on the normality of residual. The histogram and normality plot of residuals shown in Figure C.11 suggested that the residuals are fairly normal; therefore, the test result was reliable.

Assumption 4: Independence of residual (error)

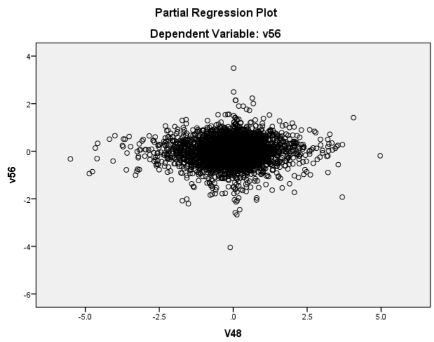
The assumption is that the error is not correlated with any independent variables. The independence of residuals was checked through the Durbin-Watson statistic of the regression model. A conventionally acceptable range of this statistic is 1 - 3 in which the value close to 2 indicates no autocorrelation (Field, 2009). In this case, Durbin-Watson statistic was 2.02 indicated the lack of autocorrelation thereby the assumption was met.

Assumption 5: Multicollinearity

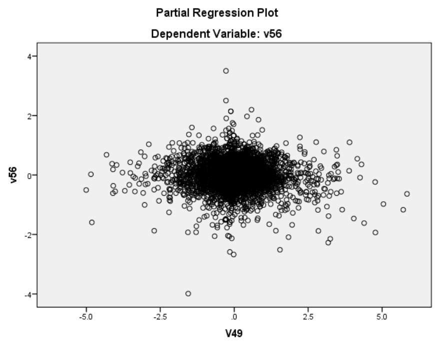
Under the assumption of no multicollinearity, the independent variables are not closely linearly related to each other. The degree of multicollinearity can be detected by inspecting the VIF statistic provided by SPSS. A VIF of 1 means that there is no correlation among the independent variables while VIF higher than 1 indicates a presence of multicollinearity among them. As a rule of thumb, VIF higher than 10 indicates the sign of serious multicollinearity (Field, 2009). The VIF for each independent variable was greater than 1 but lower than 10 (see Table C.35) hence it's can conclude that multicollinearity was present but there is no sign of serious multicollinearity which requires correction (such as removing redundant variables) within the dataset C.

Table C.35: Collinearity Statistics (VIF) of six independent variables (dataset C)

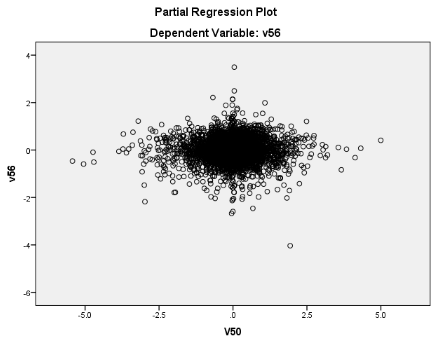
Independent variable	**Collinearity Statistics (VIF)**
Teacher	1.635
Teaching	2.058
Admin	1.807
CompFac	2.300
XActivity	2.353
AddService	2.193



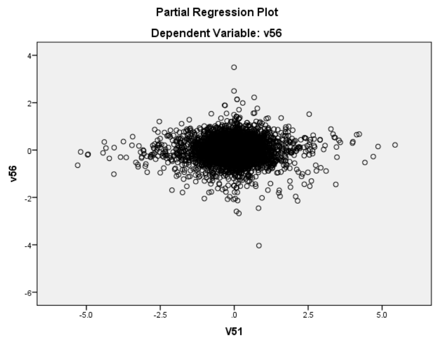
(a) Partial regression plot of V48 against V56



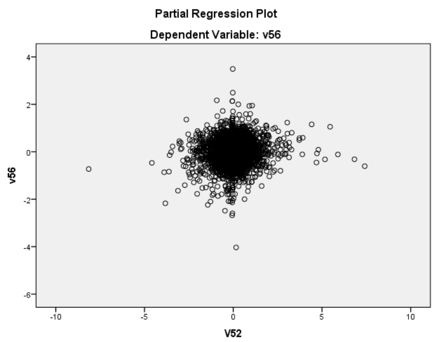
(b) Partial regression plot of V49 against V56



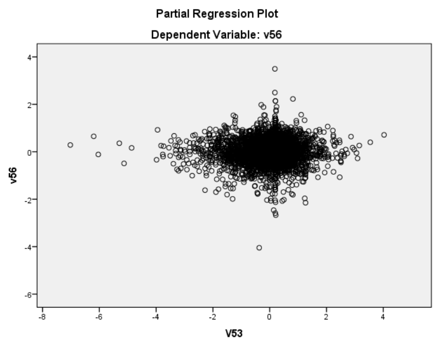
(c) Partial regression plot of V50 against V56



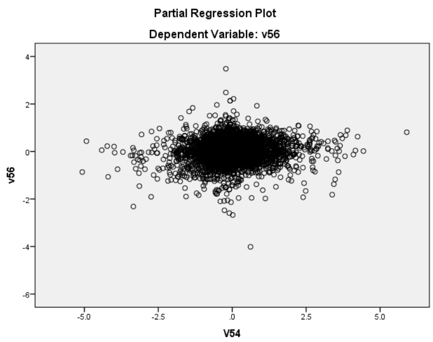
(d) Partial regression plot of V51 against V56



(e) Partial regression plot of V52 against V56

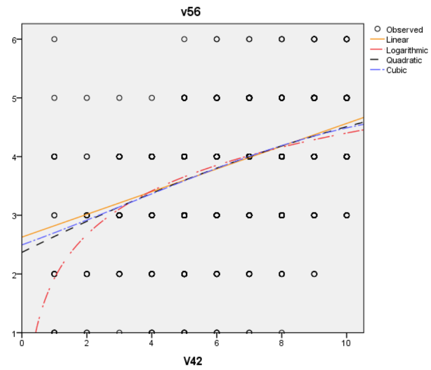


(f) Partial regression plot of V53 against V56

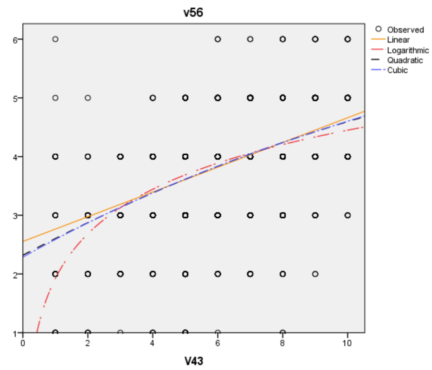


(g) Partial regression plot of V54 against V56

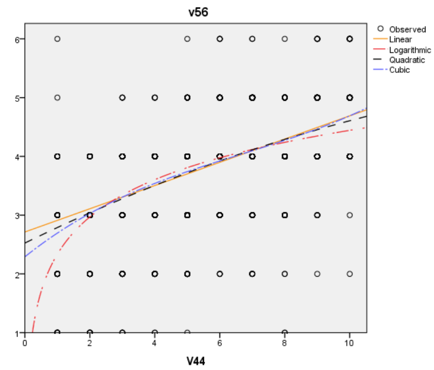
Figure C.8: Partial regression plots of independent variable V48-V54 against V56 (dataset B)



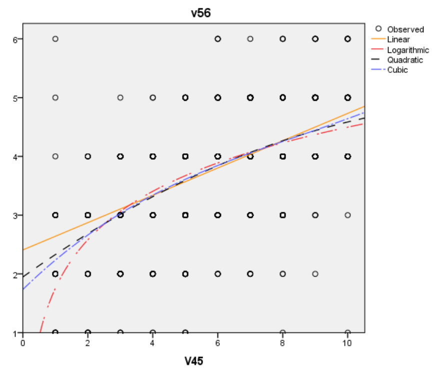
(a) Curve estimation of V42 against V56



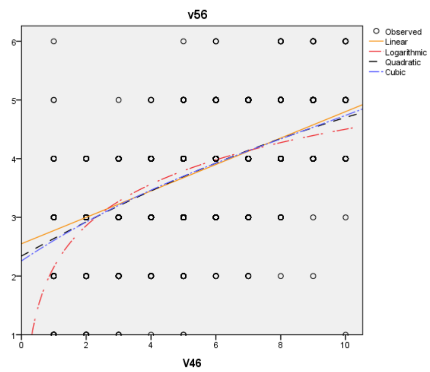
(b) Curve estimation of V43 against V56



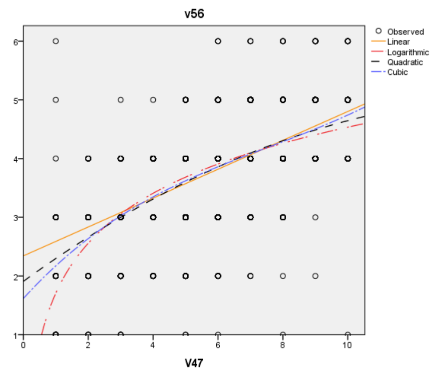
(c) Curve estimation of V44 against V56



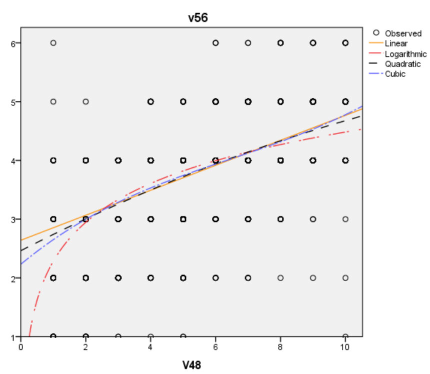
(d) Curve estimation of V45 against V56



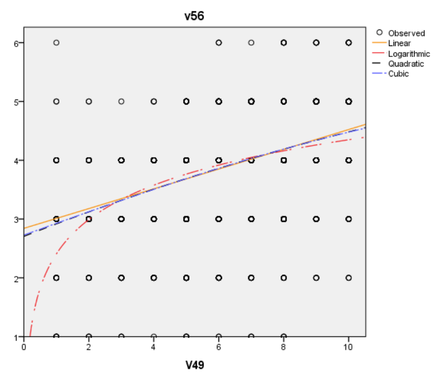
(e) Curve estimation of V46 against V56



(f) Curve estimation of V47 against V56

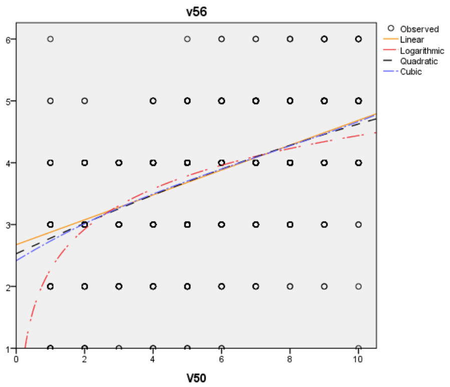


(g) Curve estimation of V48 against V56

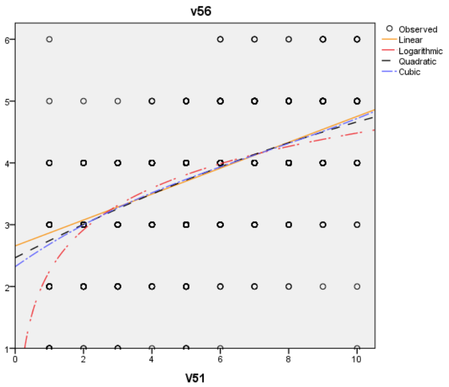


(h) Curve estimation of V49 against V56

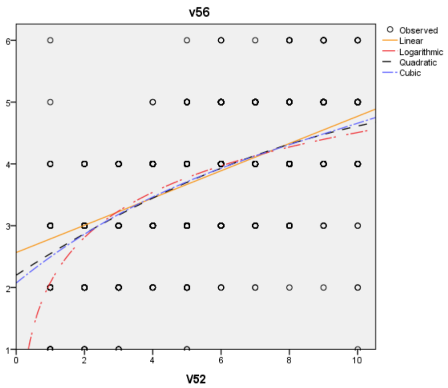
Figure C.9: Scatterplot with model curve fit of independent variable V42-V49 against V56 (dataset B)



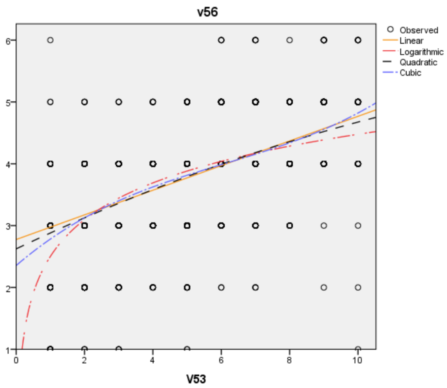
(a) Curve estimation of V50 against V56



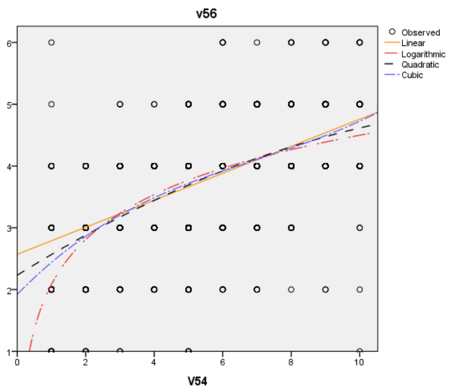
(b) Curve estimation of V51 against V56



(c) Curve estimation of V52 against V56

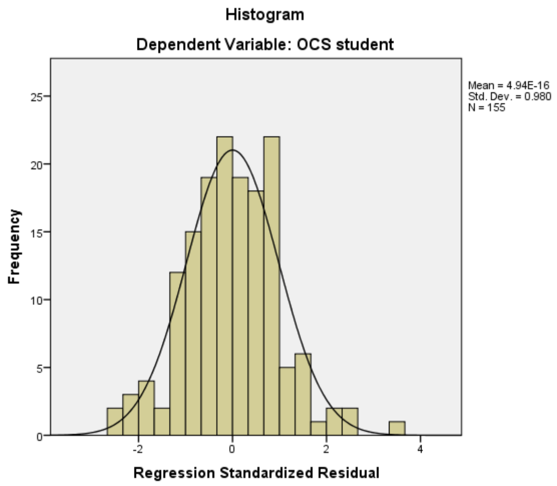


(d) Curve estimation of V53 against V56

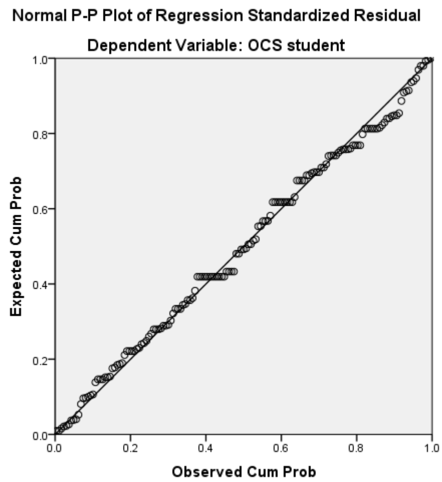


(e) Curve estimation of V544 against V56

Figure C.10: Scatterplot with model curve fit of independent variable V50-V54 against V56 (dataset B)



(a) Histogram plot of residual



(b) Normality plot of residual

Figure C.11: Histogram and normality plots of residual (dataset C)

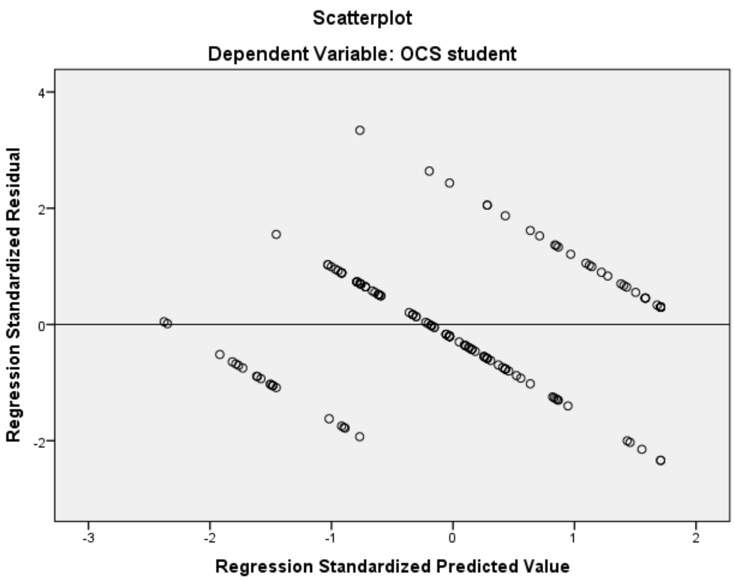
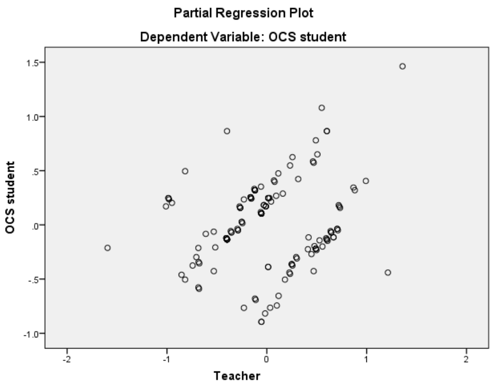
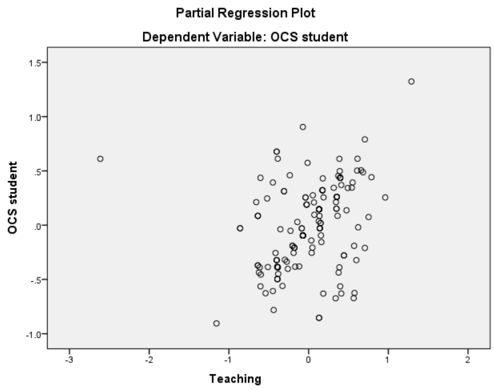


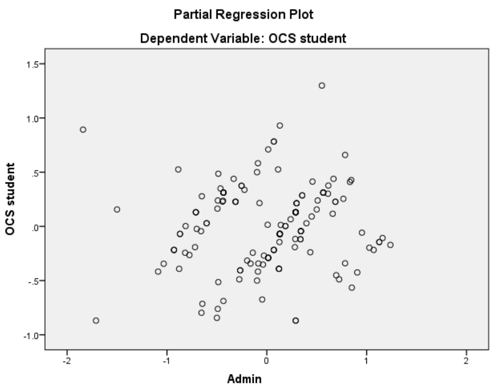
Figure C.12: Scatterplot of residual for the relationship between six independent variables and the OCS student (dataset C)



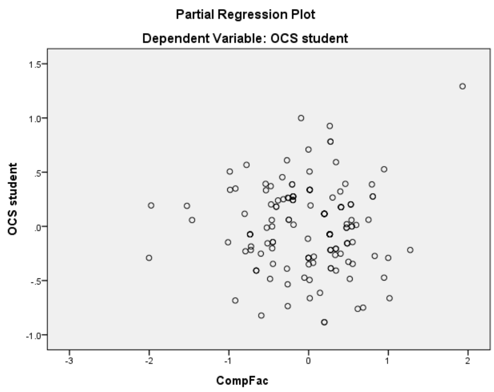
(a) Partial regression plot of Teacher against OCS student



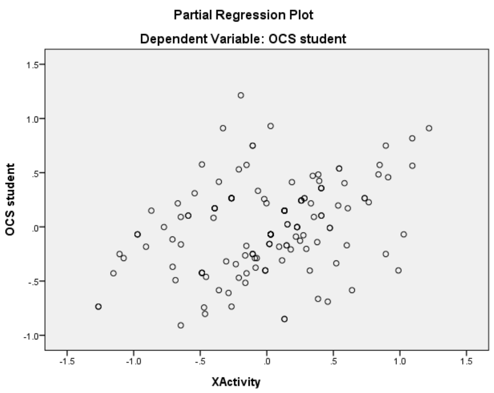
(b) Partial regression plot of Teaching against OCS student



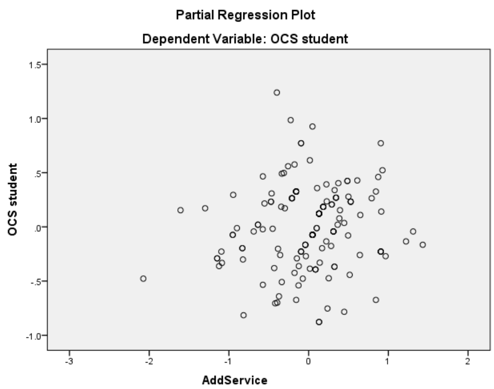
(c) Partial regression plot of Admin against OCS student



(d) Partial regression plot of CompFac against OCS student

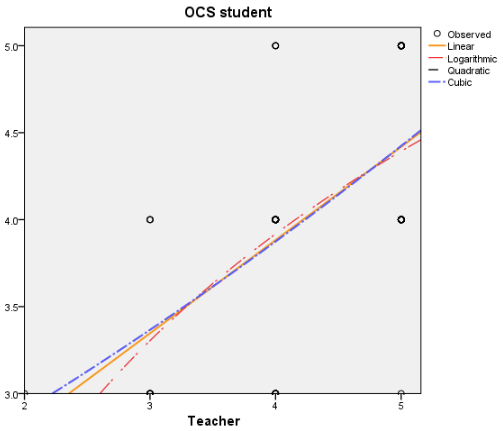


(e) Partial regression plot of xActivity against OCS student

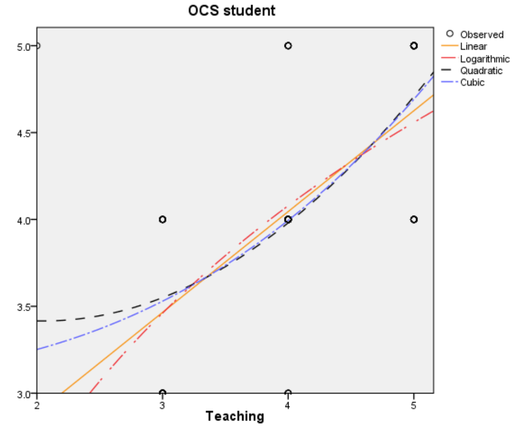


(f) Partial regression plot of AddService against OCS student

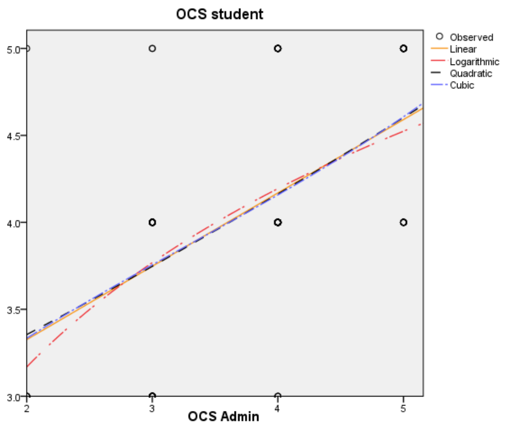
Figure C.13: Partial regression plots of each independent variable against OCS student (dataset C)



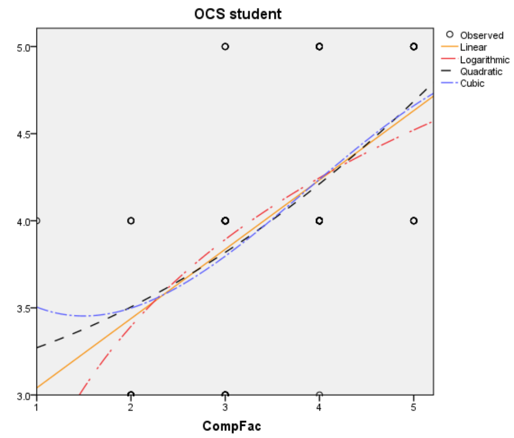
(a) Curve estimation of Teacher against OCS student



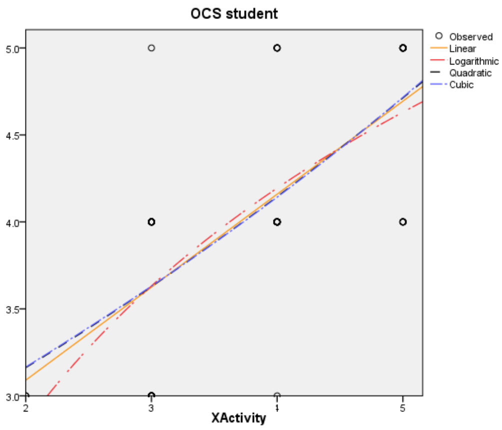
(b) Curve estimation of Teaching against OCS student



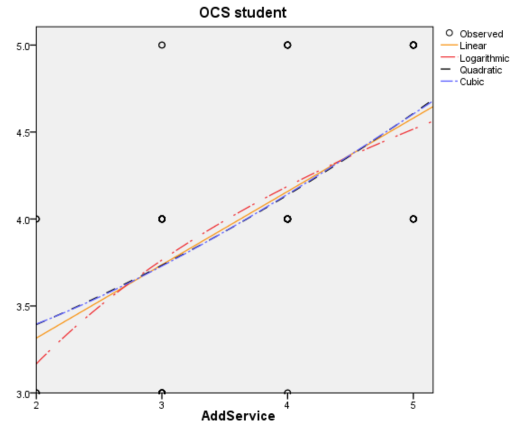
(c) Curve estimation of Admin against OCS student



(d) Curve estimation of CompFac against OCS student



(e) Curve estimation of xActivity against OCS student



(f) Curve estimation of AddService against OCS student

Figure C.14: Scatterplot with model curve fit of each independent variable against OCS student (dataset C)

Appendix D

Assumption checking of the ordinal logistic regressions on 3 datasets

To ensure that the result of OLR is valid one key assumption of OLR named proportional odds (or the assumption of parallel lines in SPSS) was tested. The null hypothesis for the test of proportional odds states that the corresponding regression coefficients in the link function are the same across all categories of ordinal responses.

The assumption is retained in case that the p-value (or Sig.) is greater than 0.05, otherwise the assumption is rejected. Assumption testing results for each dataset were summarized in Table D.1 and their detail was shown in Table D.2 - D.4.

Table D.1: Summary of assumption checking of OLR on 3 datasets

Assumptions	**Dataset**		
	A	**B**	**C**
Proportional odds	✓	✓	×

With regard to Table D.1, a symbol ✓ indicate that the assumption was met and a symbol × indicate that the assumption was violated. According to the table it shows that the assumption was met for the dataset A and dataset B, but the assumption was violate for the dataset C.

Even though the assumption of proportional odds was not met for the OLR model on dataset C, the other models of OLR such as Partial proportional Odds(PPO) and Proportional Odds With Partial Proportionality Constraints (POPPC) that relax the proportional odds assumption were not considered to use in this study. This is because the PPO and POPPC are more complicated than proportional odds that used in this study. As suggested by Fullerton (2009) that researchers should trade-off between accuracy and parsimony.

Table D.2: Test of proportional odds on dataset A

Model	**-2 Log Likelihood**	**Chi-Square**	**df**	**Sig.**
Null Hypothesis	464.753			
General	463.545	1.208	24	1.000

Table D.3: Test of proportional odds on dataset B

Model	**-2 Log Likelihood**	**Chi-Square**	**df**	**Sig.**
Null Hypothesis	0.000			
General	0.000	0.000	52	1.000

Table D.4: Test of proportional odds on dataset C

Model	**-2 Log Likelihood**	**Chi-Square**	**df**	**Sig.**
Null Hypothesis	48.315			
General	0.000	48.315	6	0.000

Appendix E

Assigning number of hidden-layer neurons

Following to Maren, Harston, and Pap (1990) cited in Deng et al. (2008a) the bound of neurons in the first hidden layer was between 2N + 1 and O(N + 1), where N is the number of input neurons and O is the number of output neurons, the bound of number hidden-layer neurons for each dataset is shown in Table E.1.

Table E.1: The bound of hidden-layer neurons

Dataset	**N**	**BPNN(regression)**			**BPNN(classification)**		
		O	**O(N+1)**	**2N+1**	**O**	**2N+1**	**O(N+1)**
A	6	1	7	13	6	13	42
B	13	1	14	27	6	27	84
C	6	1	7	13	5	13	35

Regarding Table E.1, the bounds of hidden-layer neurons for BPNN (regression) are {7, 13}, {14, 27} and {7, 13} for dataset A, B and C respectively. The bounds of hidden-layer neurons for BPNN (classification) are {13, 42}, {27, 84} and {13, 35} for dataset A, B and C respectively. In addition, several formulae for calculating number of hidden layers, shown in Table E.2, are also considered.

In order to determine the number of neurons in the hidden layer of the two BPNN structure models, several configurations of hidden-layer neurons were trained on the training dataset. Then, statistical comparisons of the performance of BPNN trained models were performed using *Experimenter* application in WEKA. Specifically, each given BPNN model was tested using 10-fold cross validation repeated 10 times for small dataset (dataset A and C), using percentage split for dataset B. Subsequently, the paired T-test was conducted (with 5% significance level) to check the null hypothesis that the mean difference of performance measures (e.g. accuracy) between the BPNN models is zero.

Table E.2: Formulas for fixing number of hidden-layer neurons

Dataset	BNNN structure Model	Number of Hidden neurons	
		? $\sqrt{N \times O}$	? $\frac{(4N^2+3)}{(N^2-8)}$
A	BPNN(regression)	2	5
	BPNN(classification)	6	5
B	BPNN(regression)	3	4
	BPNN(classification)	8	4
C	BPNN(regression)	2	5
	BPNN(classification)	5	5

E.1 Regression model of BPNN

For each dataset, the performances of BPNN(regression) network configurations within boundary in Table E.1 and fixing number in Table E.2 were measured by three indicators: the mean absolute error (MAE), the root mean squared error (RMSE) and goodness-of-fit (R^2) in the training phase. The average MAE and RMSE across different 10-folds accompanied by the variance (represented in bracket) were measured in the Testing phase.

Table E.3 shows the results of experimental BPNN(regression) models on dataset A. Regarding the statistical test on the result of 10-folds cross validation of this table, the difference of MAE and RMSE among these seven models was not statistically different at the 5% level of statistical significance. Therefore, the performance of training was considered to determine the number of neurons in the hidden layer of BPNN model. Finally, the neural network composed of two hidden neurons was selected since its RMSE value in training was the lowest one and its R^2 value was the highest one.

Table E.3: Results of experimental BPNN(regression) models on dataset A. The selected number of of neurons is represented in bold.

Number of hidden neurons	Training			Testing 10-cross validation (10 times)	
	MAE	RMSE	R^2	MAE	RMSE
2	0.588	**0.77**	**0.442**	0.617 (0.11)	0.802 (0.15)
5	0.584	0.771	0.438	0.612 (0.12)	0.805 (0.16)
7	0.582	0.773	0.426	0.614 (0.12)	0.809 (0.16)
8	0.582	0.77	0.429	0.611 (0.12)	0.807 (0.16)
10	0.583	0.778	0.419	0.613 (0.12)	0.810 (0.16)
12	0.584	0.782	0.416	0.618 (0.12)	0.817 (0.17)
13	0.585	0.781	0.415	0.610 (0.12)	0.810 (0.17)

According to the same procedure, experimental BPNN(regression) models to determine number of hidden-layer neurons on dataset B and dataset C was conducted and the results are shown in Table E.4 and Table E.5 respectively. Besides, the number of hidden-layer neurons within boundary and the number of hidden-layer neurons suggested by the formulae in Table E.2, these two experimental BPNN(regression) models were also included number of hidden-layer neurons according to the WEKA option which are N, O, N+O and (N+O)/2, where N is the number of input neurons and O is the number of output neurons.

Table E.4 shows the results of experimental BPNN(regression) models on dataset B. The performance of training among 10 models was about the same and there was no significantly difference of MAE and RMSE among the performance testing of percentage split of 10 models. Thus, the neural network composed of three hidden neurons was selected since this model produced the lowest MAE and RMSE although it was not statically different from other models.

Table E.4: Results of experimental BPNN(regression) models on dataset B. The selected number of of neurons is represented in bold.

Number of hidden neurons	**Training**			**Testing** Percentage split (10 times)	
	MAE	**RMSE**	R^2	**MAE**	**RMSE**
1	0.383	0.501	0.435	0.388(0.01)	0.507 (0.01)
3	0.38	0.499	0.441	**0.385 (0.01)**	**0.507 (0.01)**
4	0.38	0.499	0.441	0.385 (0.01)	0.508 (0.01)
7	0.38	0.499	0.441	0.386 (0.01)	0.507 (0.01)
13	0.38	0.499	0.44	0.387 (0.01)	0.507 (0.01)
14	0.38	0.499	0.441	0.387 (0.01)	0.507 (0.01)
15	0.38	0.500	0.440	0.387 (0.01)	0.507 (0.01)
18	0.38	0.499	0.442	0.387 (0.01)	0.507 (0.01)
21	0.38	0.499	0.441	0.387 (0.01)	0.507 (0.01)
25	0.38	0.499	0.441	0.387 (0.01)	0.507 (0.01)

Regarding Table E.5, the model with 12 hidden neurons yielded similar performance in training and testing as same as model with 13 hidden neurons, their MAE measured through 10-folds cross validation of this model was statistically different at the 5% level of statistical significance from other models with 1-3, 5-8 hidden-layer neuron(s). Both models also yielded lowest MAE and highest R^2 in training. Thus, either of them is qualified to be selected however considering the number of hidden neurons, the smaller number of hidden neurons was preferable. Thereby, the neural network composed of 12 hidden neurons was selected.

E.2 Classification model of BPNN

For the classification model of BPNN, networks with different number of hidden neurons within boundary in Table E.1 and fixing number in Table E.2, and number of hidden-layer neurons according to the WEKA option which are N, O, N+O and (N+O)/2

Table E.5: Results of experimental BPNN(regression) models on dataset C. The selected number of of neurons is represented in bold.

Number of hidden neurons	Training			Testing 10-cross validation (10 times)	
	MAE	RMSE	R^2	MAE	RMSE
1	0.313	0.371	0.61	0.33(0.05) *	0.40(0.08) *
2	0.311	0.37	0.613	0.33(0.05) *	0.39(0.08)
3	0.31	0.369	0.614	0.33(0.05) *	0.39(0.08)
5	0.31	0.369	0.614	0.33(0.05) *	0.39(0.08)
6	0.308	0.368	0.616	0.33(0.05) *	0.39(0.08)
7	0.309	0.369	0.615	0.33(0.05) *	0.39(0.08)
8	0.307	0.368	0.616	0.32(0.05) *	0.39(0.08)
10	0.308	0.368	0.616	0.32(0.05)	0.39(0.08)
11	0.307	0.368	0.617	0.32(0.05)	0.39(0.08)
12	0.306	0.368	0.617	0.32(0.05)	0.39(0.08)
13	0.306	0.368	0.617	0.32(0.05)	0.39(0.08)

* Significantly difference with the neural network composed of 12 hidden neurons

were trained. Then, their performance was measured by three indicators: percentage of accuracy, RMSE and area under the Receiver Operating Characteristic (ROC) curve (AUC). For the testing phase average percentage accuracy, RMSE and AUC across different evaluation 10 times accompanied by the variance are shown. For each dataset, the results of experimental BPNN(classification) models are shown in Table E.6 - Table E. 8.

The test of statistical significance on the result of 10-folds cross validation in Table E.6 shows that the RMSE and AUC of two models with five and six hidden neurons were statistically different from other network structures at the 5% level of statistical significance. Therefore, the training performance of five and six hidden neurons model were considered to determine the number of neurons in the hidden layer of the BPNN model. Regarding accuracy and RMSE in training, the neural network composed of six hidden neurons was selected since its RMSE value was lower than the RMSE value of the neural network composed of five hidden neurons and its accuracy was slightly greater than the accuracy of the five hidden neurons.

Table E.7 shows that RMSE on the result of percentage split of model with nine hidden neurons is statistical significance from other models with 19, 25 and 30 neurons at the 5% level of statistical significance. The performance of models with 4, 6, 8-9 and 13 was considered in training and testing, then the model with nine hidden neurons was selected since it yielded highest AUC in testing and higher accuracy in testing than the accuracy of model with 13 hidden neurons that yielded highest accuracy in the training.

Regarding the performance of model training and testing in Table E.8, the neural network composed of 13 hidden neurons was selected since the accuracy and RMSE in training among models was about the same but the model with 13 hidden neurons yielded

Table E.6: Results of experimental BPNN(classification) models on dataset A. The selected number of of neurons is represented in bold.

Number of hidden neurons	Training		Testing 10-cross validation (10 times)		
	Accuracy	**RMSE**	**Accuracy**	**RMSE**	**AUC**
5	64.394	0.285	47.081 (8.55)	0.342 (0.02)	0.700 (0.22)
6	**65.530**	**0.279**	48.299 (9.01)	0.342 (0.02)	0.685 (0.27)
12	77.652	0.249	46.425 (8.36)	0.362(0.03) *	0.402 (0.30) *
13	75.000	0.249	47.459 (8.29)	0.363(0.03) *	0.225 (0.24) *
15	78.030	0.233	46.936 (9.04)	0.363(0.03) *	0.429 (0.32) *
18	80.682	0.219	46.340 (8.96)	0.365(0.03) *	0.223 (0.25) *
20	81.061	0.224	45.486 (8.39)	0.368(0.03) *	0.154 (0.19) *
22	77.273	0.234	45.782 (9.42)	0.368(0.03) *	0.154 (0.14) *
24	80.682	0.221	46.115 (9.06)	0.366(0.03) *	0.225 (0.23) *
25	76.894	0.244	46.598 (8.94)	0.366(0.03) *	0.167 (0.23) *
28	78.030	0.237	45.084 (9.30)	0.367(0.03) *	0.158 (0.23) *
30	81.439	0.223	45.989 (9.44)	0.366(0.03) *	0.162 (0.18) *
33	80.303	0.223	46.142 (8.72)	0.366(0.03) *	0.156 (0.18) *
35	77.273	0.234	45.105 (9.64)	0.372(0.03) *	0.129 (0.22) *
38	78.030	0.237	46.060 (8.80)	0.368(0.03) *	0.137 (0.12) *
42	78.409	0.237	45.188 (8.47)	0.369(0.03) *	0.196 (0.26) *

* Significantly difference with the neural network composed of five hidden neurons

Table E.7: Results of experimental BPNN(classification) models on dataset B. The selected number of of neurons is represented in bold.

Number of hidden neurons	Training		Testing Percentage split (10 times)		
	Accuracy	**RMSE**	**Accuracy**	**RMSE**	**AUC**
4	72.911	0.252	69.901 (1.89)	0.261 (0.01)	0.927(0.06)
6	73.178	0.251	69.966 (1.66)	0.262 (0.00)*	0.932(0.04)
8	74.556	0.248	69.581 (1.63)	0.265 (0.00)	0.933(0.04)
9	74.556	0.247	69.731 (1.70)	0.265 (0.00)	**0.936(0.05)**
13	75.000	0.245	69.437 (1.78)	0.270 (0.01)	0.873(0.07)
19	76.267	0.241	69.326 (1.45)	0.273 (0.01) *	0.892 (0.07)
25	76.622	0.238	68.594 (1.73)	0.277 (0.01) *	0.882(0.07)
30	77.156	0.236	68.215 (1.60)	0.279 (0.01) *	0.690(0.33)

* Significantly difference with the neural network composed of nine hidden neurons

highest accuracy in testing using 10-cross validation. Although it is only statistical significance from one model with 35 neurons at the 5% level of statistical significance.

Table E.8: Results of experimental BPNN(classification) models on dataset C. The selected number of of neurons is represented in bold.

Number of hidden neurons	**Training**		**Testing** 10-cross validation (10 times)		
	Accuracy	**RMSE**	**Accuracy**	**RMSE**	**AUC**
5	89.032	0.187	79.71(8.69)	0.25(0.05)	0.86(0.09)
6	90.968	0.179	80.46(8.77)	0.25(0.06)	0.86(0.09)
11	90.968	0.178	79.50(9.72)	0.25(0.06)	0.86(0.09)
13	90.968	0.178	**80.47(9.45)**	0.25(0.06)	0.86(0.09)
15	90.968	0.179	78.86(10.11)	0.25(0.06)	0.85(0.09)
18	90.968	0.179	78.51(10.48)	0.26(0.06)	0.85(0.09)
20	90.968	0.176	77.37(10.56)	0.26(0.06)	0.85(0.09)
25	90.968	0.174	76.88(10.52)	0.26(0.06)	0.84(0.10)
30	90.968	0.173	75.57(11.34)	0.26(0.06)	0.84(0.10)
35	90.968	0.176	74.94(10.59) *	0.26(0.05)	0.84(0.10)

* Significantly difference with the neural network composed of 13 hidden neurons

Appendix F

Computational example of *importance* measures

F.1 Computational example for computing *importance* from Multiple Linear Regression

This section illustrates the computational example for obtaining *importance* shown in the third column of Table 5.9 from the regression coefficients shown in Figure F.1 and Table F.1.

The regression coefficients produced by SAS enterprise miner are unstandardized coefficients however the standardized coefficients are preferable. Hence, these unstandardized coefficients were converted into standardized coefficients by multiplying each unstandardized coefficient with the proportion of standard deviation of the attribute and standard deviation of Satisfaction (value = 1.035).

Importance of each attribute can be computed as percentage contribution of the standardized coefficients in which a negative one is set as zero. The standardized coefficients and *importance* are presented in the fourth and fifth column of Table F.1 respectively.

Table F.1: Unstandardized and standardized coefficients of MLR, and *importance* of attribute in dataset A

Attribute	**Unstandardized Coefficients**	**Std. Deviation**	**Standardized Coefficients**	**Importance**
Equipment and System	0.1915	0.987	0.183	18.769
Sales Support	0.0483	1.384	0.065	6.667
Technical Support	0.362	1.137	0.398	40.821
Supplies and Orders	0.1495	1.009	0.146	14.974
Purchasing Support	-0.00499	1.294	-0.006	0.000
Contracts and Pricing	0.1687	1.12	0.183	18.769

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.3968	0.2442	1.63	0.1054
Equipment	1	0.1915	0.0541	3.54	0.0005
SalesSup	1	0.0483	0.0382	1.27	0.2066
TechnicalSup	1	0.3620	0.0477	7.59	<.0001
Suppliers	1	0.1495	0.0523	2.86	0.0046
PurchaseSup	1	-0.00499	0.0403	-0.12	0.9016
CondandPrices	1	0.1687	0.0502	3.36	0.0009

Figure F.1: Result of multiple linear regression implemented on dataset A

F.2 Computational example for computing *importance* from Ordinal Linear Regression

This section illustrates the computational example for obtaining *importance* shown in the fourth column of Table 5.9 from the regression coefficients shown in Figure F.2.

Before the *importance* are calculated from regression coefficients, the sign of regression coefficients obtained from SAS are reversed and shown in the first column of Table F.2. This is because SAS internally changes the sign of coefficients in estimating coefficients depending on the option ascending order or descending order. *Importance* of each attribute can be computed as percentage contribution of the standardized coefficients in which a negative one is set as zero, and presented in the second column of Table F.2.

Table F.2: Estimated coefficients of OLR and *importance* of attribute in dataset A

Attribute	**Estimated Coefficients**	**Importance**
Equipment and System	0.231	17.408
Sales Support	0.073	5.501
Technical Support	0.625	47.099
Supplies and Orders	0.171	12.886
Purchasing Support	-0.046	0.000
Contracts and Pricing	0.227	17.106

F.3 Computational example for computing *importance* from neural network

This section illustrates the computational example for obtaining *importance* shown in the fifth column of Table 5.9 from the connection weights of neural network shown in Table F.3 based on the procedure to compute attribute's importance described in Section 3.3.3.

Given the connection weights of 6-2-1 neural network for dataset A as shown in Table F.3, the computation process for obtaining *importance* is as follows:

1. Input-Hidden layer computation. The proportion of each input-hidden connection weights shown in TableF.5 are calculated by dividing each absolute weight by the sum of absolute weights for all inputs corresponding to each hidden layer displayed in Table F.4.

e.g. $P_{11} = \frac{w_{11}}{S_{H1}} = \frac{0.459}{2.067} = 0.222$

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	0	1	-1.0453	0.7736	1.83	0.1766
Intercept	1	1	0.9073	0.4252	4.55	0.0329
Intercept	2	1	2.1604	0.3751	33.17	<.0001
Intercept	3	1	3.7043	0.4010	85.34	<.0001
Intercept	4	1	5.5438	0.4592	145.77	<.0001
Equipment		1	-0.2312	0.0701	10.59	0.0011
SalesSup		1	-0.0728	0.0587	1.54	0.2149
TechnicalSup		1	-0.6252	0.0811	59.88	<.0001
Suppliers		1	-0.1714	0.0705	5.72	0.0167
PurchaseSup		1	0.0459	0.0664	0.45	0.5022
CondandPrices		1	-0.2271	0.0750	9.17	0.0025

Figure F.2: Result of ordinal linear regression implemented on dataset A

Table F.3: neural weights of 6-2-1 network

Input/Hidden	**H1**	**H2**
Equipment and System	-0.45945	-0.51086
Sales Support	-0.61236	0.081693
Technical Support	-0.09916	-0.97187
Supplies and Orders	-0.15282	-0.47836
Purchasing Support	0.363219	-0.13334
Contracts and Pricing	-0.38013	-0.48767

Output/Hidden	**H1**	**H2**
Satisfaction	-0.69225	-1.32191

Table F.4: Absolute value of input-hidden connection weights

Input/Hidden	**H1**	**H2**
Equipment and System	$w_{11} = 0.459$	$w_{12} = 0.511$
Sales Support	$w_{21} = 0.612$	$w_{22} = 0.082$
Technical Support	$w_{31} = 0.099$	$w_{32} = 0.972$
Supplies and Orders	$w_{41} = 0.153$	$w_{42} = 0.478$
Purchasing Support	$w_{51} = 0.363$	$w_{52} = 0.133$
Contracts and Pricing	$w_{61} = 0.380$	$w_{62} = 0.488$
Sum	$S_{H1} = 2.067$	$S_{H2} = 2.664$

Table F.5: Proportion of input-hidden connection weights

Input/Hidden	**H1**	**H2**
Equipment and System	$P_{11} = 0.222$	$P_{12} = 0.192$
Sales Support	$P_{21} = 0.296$	$P_{22} = 0.031$
Technical Support	$P_{31} = 0.048$	$P_{32} = 0.365$
Supplies and Orders	$P_{41} = 0.074$	$P_{42} = 0.180$
Purchasing Support	$P_{51} = 0.176$	$P_{52} = 0.050$
Contracts and Pricing	$P_{61} = 0.184$	$P_{62} = 0.183$
Sum	1.000	1.000

2. Hidden-Output layer computation. The total hidden-output weights of each hidden neurons are computed as the sum of the absolute value of all outputs corresponding to each hidden layer. Since, there is only one output neuron in this example, the total sum is equal to the the absolute value of hidden-output weights as shown in Table F.6.

Table F.6: Absolute value of hidden-output connection weights

Output/Hidden	**H1**	**H2**
Satisfaction	$w_{11} = 0.692$	$w_{12} = 1.322$
Sum	$SO_1 = 0.692$	$SO_2 = 1.322$

3. Contribution of Input-Hidden weights and Hidden-Output weights. The contribution of each neuron to the output via each hidden neuron shown in Table F.7 is

computed by multiplying the proportion of each input-hidden connection weight (Table F.5) by total output weight of the corresponding hidden neuron (Table F.6).
e.g. C_{11}= $P_{11} \times SO_1 = 0.222 \times 0.692 = 0.154$

Table F.7: The contribution of each neuron to the output

Input/Hidden	**H1**	**H2**
Equipment and System	C_{11}= 0.154	C_{12} = 0.254
Sales Support	C_{21}= 0.205	C_{22} = 0.041
Technical Support	C_{31} = 0.033	C_{32} = 0.482
Supplies and Orders	C_{41} = 0.051	C_{42} = 0.237
Purchasing Support	C_{51} =0.122	C_{52} = 0.066
Contracts and Pricing	C_{61} =0.127	C_{62} = 0.242

4. Relative importance computation. The relative importance shown in Table F. 8 is calculated as the sum of the contribution of each neuron corresponding input layer (Table F.7) divided by overall contribution of all neurons in the network, and multiplied by 100.
e.g. $S_1 = C_{11} + C_{12} = 0.154 + 0.254 = 0.407$ and
$RI_1 = \frac{S_1}{(S_1+S_2+...+S_6)} \times 100 = \frac{0.407}{2.014} \times 100 = 20.226$

Table F.8: The relative importance of each attribute

Input	**Sum of contribution**	**Relative Importance (%)**
Equipment and System	S_1 = 0.407	RI_1 = 20.23
Sales Support	S_2 = 0.246	RI_2 = 12.19
Technical Support	S_3 = 0.515	RI_3 = 25.59
Supplies and Orders	S_4 = 0.289	RI_4 = 14.33
Purchasing Support	S_5 = 0.188	RI_5 = 9.32
Contracts and Pricing	S_6 = 0.369	RI_6 = 18.34
Sum	2.014	100.000

F.4 Computational example for computing *importance* from Naïve Bayes

This section illustrates the computational example for obtaining *importance* shown in the seventh column of Table 5.9 from the conditional probability distribution generated by Naïve Bayes. Regarding the model construction on dataset A, Naïve Bayes run by using WEKA generated six conditional probability tables corresponded to each attribute.

For each conditional probability table, the *importance* of each attribute is calculated based on the procedure to compute attribute's importance described in Step 2 of Section 3.3.4. The computational process is illustrated as follow:

1. Entropy of each attribute. The entropy of attribute $h(X_i)$ is computed as Equation 3.7, where $P(X_i = a)$ is the probability of each attribute with the satisfaction

Table F.9: The conditional probability table of attribute Equipment generated from WEKA

Equipment \Class	**0**	**1**	**2**	**3**	**4**	**5**	**Total**
0	1	1	1	4	3	3	13
1	1	5	2	1	2	1	12
2	1	4	9	6	3	1	24
3	2	4	7	41	21	4	79
4	2	2	9	22	87	27	149
5	1	1	3	2	8	8	23
Total class	8	17	31	76	124	44	300

level a. Since a has six possible level in a Likert scale ($a = 0, 1, \ldots, 5$), $P(X_i = a)$ for attribute Equipment (X_1) is calculated from the conditional probability table (Table F. 9) by dividing number of records of the level a (number in the last column of each row) to the number of all records.

This gives $P(X_1 = 0) = 13/300 = 0.043$, $P(X_1 = 1) = 12/300 = 0.040$, $P(X_1 = 2)$ $= 24/300 = 0.080$, $P(X_1 = 3) = 79/300 = 0.263$, $P(X_1 = 4) = 149/300 = 0.497$, and $P(X_1 = 5) = 23/300 = 0.077$. Using the same procedure, the $P(X_i = a)$ for the other five attributes are calculated and shown in Table F.10.

Table F.10: The probability density function for each attribute

X_i	**P(X_i=a)**	***a***					
		0	**1**	**2**	**3**	**4**	**5**
X1-Equipment and System	P(X_1=a)	0.043	0.040	0.080	0.263	0.497	0.077
X2-Sales Support	P(X_2=a)	0.073	0.093	0.153	0.250	0.280	0.150
X3-Technical Support	P(X_3=a)	0.020	0.063	0.127	0.147	0.377	0.267
X4-Supplies and Orders	P(X_4=a)	0.060	0.030	0.083	0.390	0.387	0.050
X5-Purchasing Support	P(X_5=a)	0.090	0.050	0.130	0.270	0.373	0.087
X6-Contracts and Pricing	P(X_6=a)	0.057	0.080	0.163	0.380	0.263	0.057

Given the probability density function for each attribute shown in Table F.10, the entropy of attribute $h(X_i)$ for $i = 1$ is computed as:

$$\begin{aligned} h(X_1) = &- 0.043 * \log_2(0.043) \; - 0.040 * \log_2(0.040) \; - 0.080 * \log_2(0.080) \; - 0.263 * \log_2(0.263) \\ &- 0.497 * \log_2(0.497) \; - 0.077 * \log_2(0.077) = 1.966 \end{aligned}$$

By the similar approach, the entropy of the other five attributes are calculated and shown in the second column of Table F.13

2. The conditional entropy. The entropy of X_i conditioned on Y, $h(X_i|Y)$ is computed as Equation 3.8. This equation can be broken into two main components which are $P(Y = b)$ (the probability of the overall customer satisfaction Y with the satisfaction level b) and $h(X_i|Y = b)$ (the average conditional entropy of X_i).

$P(Y = b)$ is a class distribution for the satisfaction level b ($b = 0, 1, \ldots, 5$), calculated from the conditional probability table (Table F.9) by dividing number of records of the level b (number in the last row of each column) to the number of all records. This gives $P(Y = 0) = 8/300 = 0.027$, $P(Y = 1) = 17/300 = 0.057$, $P(Y = 2) = 31/300 = 0.103$, $P(Y = 3) = 76/300 = 0.253$, $P(Y = 4) = 124/300 = 0.413$, and $P(Y = 5) = 44/300 = 0.147$.

$h(X_i|Y = b)$ can be computed from Equation 3.9, where $P(X_i = a|Y = b)$ is the distribution of $X_i = a$ condition on $Y = b$. $P(X_i = a|Y = b)$ is calculated from the conditional probability table by dividing number of records of the level a for attribute X_i to the number of records of the level b for overall customer satisfaction Y.

For instance, given the conditional probability table of attribute "Equipment" as shown in Table F.9, the computational of $P(X_1 = a|Y = b)$ would be $P(X_1 = 0|Y = 0) = 1/8 = 0.125$, $P(X_1 = 1|Y = 0) = 1/8 = 0.125$, ... , $P(X_1 = 0|Y = 1) = 1/17 = 0.059$, ..., $P(X_1 = 5|Y = 5) = 8/44 = 0.182$. In summary, there are 36 combinations in this computation process since both a and b have six possible levels. The specific results of $P(X_1 = a|Y = b)$ are provided in Table F.11.

Table F.11: Computational results of $P(X_1 = a|Y = b)$ (attribute Equipment)

	b					
a	**0**	**1**	**2**	**3**	**4**	**5**
0	0.125	0.059	0.032	0.053	0.024	0.068
1	0.125	0.294	0.065	0.013	0.016	0.023
2	0.125	0.235	0.290	0.079	0.024	0.023
3	0.250	0.235	0.226	0.539	0.169	0.091
4	0.250	0.118	0.290	0.289	0.702	0.614
5	0.125	0.059	0.097	0.026	0.065	0.182

Based on computational results of $P(X_1 = a|Y = b)$ shown in Table F.11, $h(X_i|Y = b)$ for $i = 1$ and $b = 0$ is computed as:

$$
\begin{aligned}
h(X_1|Y = 0) = & - 0.125 * \log_2(0.125) \ - 0.125 * \log_2(0.125) \ - 0.125 * \log_2(0.125) \\
& - 0.250 * \log_2(0.250) \ - 0.250 * \log_2(0.250) \ - 0.125 * \log_2(0.125) = 2.500
\end{aligned}
$$

By the similar approach, the entropy of $h(X_1|Y = b)$ for $b = 1, \ldots, 5$ are computed and shown in Table F.12.

Table F.12: The average conditional entropy of attribute Equipment , $h(X_1|Y = b)$

	b					
	0	**1**	**2**	**3**	**4**	**5**
$h(X_1|Y = b)$	2.500	2.346	2.262	1.731	1.404	1.706

Finally, the computational of $h(X_1|Y)$ can be calculated as a summation of product between $h(X_1|Y=b)$ and $P(Y=b)$ for $b=0,\ldots,5$ (see Equation 3.8):

$$\begin{aligned} h(X_1|Y) = {} & 2.500*0.027 \ + \ 2.346*0.057 \ + \ 2.262*0.103 \ + \ 1.731*0.253 \\ & + 1.404*0.413 \ + \ 1.706*0.147 = 1.702 \end{aligned}$$

Using the same procedure, the computational results of $h(X_i|Y)$ for all attribute are shown in the third column of Table F.13.

3. Mutual Information. Given the entropy and conditional entropy of each attribute (see the second and third column of Table F.13), mutual information of each attribute can be computed as Equation 3.6, and then *importance* of each attribute can be computed as percentage contribution of the mutual information. These two values are presented in the fourth and fifth column of the same table respectively.

Table F.13: Mutual Information and *importance* of attributes derived from Naïve Bayes

X_i	$h(X_i)$	$h(X_i\|Y)$	**MI(X_i,Y)**	**Importance**
X_1-Equipment and System	1.966	1.702	0.264	22.524
X_2-Sales Support	2.435	2.290	0.145	12.418
X_3-Technical Support	2.188	1.921	0.267	22.796
X_4-Supplies and Orders	1.970	1.802	0.168	14.359
X_5-Purchasing Support	2.258	2.138	0.120	10.211
X_6-Contracts and Pricing	2.225	2.018	0.207	17.691

F.5 Computational example for computing *importance* from Bayesian Networks

This section illustrates the computational example for obtaining *importance* shown in the last column of Table 5.9 from the conditional probability distribution generated by BNs. Similar to the process for deriving *importance* from Naïve Bayes, the *importance* of BNs are calculated from probability distribution table as Mutual Information (Equation 3.6). However, the method for calculating the entropy of X_i, $h(X_i)$ and the conditional entropy $h(X_i|Y)$ from BNs is slightly more complicated than the method for calculating these two entropies from Naïve Bayes, as the network structure of BNs is more complicated than the network structure of Naïve Bayes.

The conditional probability distribution tables represent a probability distribution of node depending on its parents. For example nodes represented attribute Contractand-Price and Equipment shown in Figure 5.5, the first has only one parent (Satisfaction) whereas the latter has two parents (Satisfaction and ContractandPrice) and their conditional probability distribution tables are shown as Figure F.3 and Figure F.4.

The table size of the conditional probability distribution of attribute ContractandPrice is 6 x 6 (see Figure F.3). Each cell indicates probability of ContractandPrice given Satisfaction where value of ContractandPrice is equal to a and value of Satisfaction is equal to b.

Probability Distribution Table For ContractandPrice

Satisfaction	0	1	2	3	4	5
0	0.125	0.125	0.25	0.125	0.25	0.125
1	0.118	0.353	0.176	0.176	0.118	0.059
2	0.065	0.129	0.419	0.258	0.065	0.065
3	0.039	0.092	0.237	0.526	0.092	0.013
4	0.056	0.04	0.089	0.403	0.347	0.065
5	0.045	0.023	0.045	0.273	0.523	0.091

Figure F.3: The conditional probability distribution of attribute ContractandPrice

The table size of the conditional probability distribution of Equipment is 36 x 6 (see Figure F.4). Number of rows in the table is the combination of all possible values of two parent nodes of Equipment. Both Satisfaction and ContractandPrice have 6 possible values 0-5 therefore their combination is 36. Each cell indicates probability of Equipment given Satisfaction and ContractandPrice where Equipment is equal to a, value of Satisfaction is equal to b and value of ContractandPrice is equal to c.

F.5.1 Pre-calculation procedure

There is one pre-calculation procedure to calculate the entropy of X_i, $h(X_i)$ and the conditional entropy $h(X_i|Y)$ from BNs, which is the calculation of $P(X_i = a)$ and $P(X_i|Y = b)$. Regarding the two types of conditional probability distribution tables, method for calculating $P(X_i = a)$ and $P(X_i|Y = b)$, can be divided in two cases: attribute with one parent (e.g. attribute ContractandPrice) and attribute with two parents (e.g. attribute Equipment).

1. Calculate $P(X_i = a)$ and $P(X_i|Y = b)$ on attribute X_i that has one parent

For the case of attribute with one parent, $P(X_i|Y = b)$ can be directly obtained as the conditional probability distribution table generated from WEKA. And then probability distribution of each attribute $P(X_i = a)$ can be computed based on the conditional probability distribution table by using the law of total probability[1](also known as marginalization).

[1] www.en.wikipedia.org/wiki/Law_of_total_probability

Probability Distribution Table For Equipment

Satisfaction	ContractandPrice	0	1	2	3	4	5
0	0	0.167	0.167	0.167	0.167	0.167	0.167
0	1	0.167	0.167	0.167	0.167	0.167	0.167
0	2	0.143	0.143	0.143	0.286	0.143	0.143
0	3	0.167	0.167	0.167	0.167	0.167	0.167
0	4	0.143	0.143	0.143	0.143	0.286	0.143
0	5	0.167	0.167	0.167	0.167	0.167	0.167
1	0	0.143	0.143	0.143	0.286	0.143	0.143
1	1	0.091	0.364	0.182	0.182	0.091	0.091
1	2	0.125	0.125	0.25	0.25	0.125	0.125
1	3	0.125	0.25	0.125	0.125	0.25	0.125
1	4	0.143	0.143	0.286	0.143	0.143	0.143
1	5	0.167	0.167	0.167	0.167	0.167	0.167
2	0	0.143	0.143	0.143	0.143	0.286	0.143
2	1	0.111	0.111	0.333	0.222	0.111	0.111
2	2	0.056	0.111	0.333	0.278	0.167	0.056
2	3	0.077	0.077	0.154	0.154	0.462	0.077
2	4	0.143	0.143	0.143	0.143	0.143	0.286
2	5	0.143	0.143	0.143	0.143	0.143	0.286
3	0	0.25	0.125	0.125	0.25	0.125	0.125
3	1	0.083	0.083	0.167	0.5	0.083	0.083
3	2	0.087	0.043	0.174	0.435	0.174	0.087
3	3	0.044	0.022	0.022	0.511	0.378	0.022
3	4	0.083	0.083	0.167	0.333	0.25	0.083
3	5	0.167	0.167	0.167	0.167	0.167	0.167
4	0	0.083	0.083	0.083	0.167	0.5	0.083
4	1	0.2	0.1	0.1	0.1	0.4	0.1
4	2	0.062	0.125	0.125	0.312	0.312	0.062
4	3	0.018	0.018	0.036	0.2	0.655	0.073
4	4	0.042	0.021	0.021	0.083	0.771	0.062
4	5	0.077	0.077	0.077	0.231	0.308	0.231
5	0	0.143	0.143	0.143	0.143	0.286	0.143
5	1	0.167	0.167	0.167	0.167	0.167	0.167
5	2	0.143	0.143	0.143	0.286	0.143	0.143
5	3	0.059	0.059	0.059	0.118	0.529	0.176
5	4	0.107	0.036	0.036	0.036	0.643	0.143
5	5	0.111	0.111	0.111	0.222	0.111	0.333

Figure F.4: The conditional probability distribution of attribute Equipment

For example, $P(X_i|Y = b)$ of attribute ContractandPrice is generated by WEKA and presented in Table F.14. Given Table F.14 and probability distribution of Satisfaction (Table F.15), probability distribution of attribute ContractandPrice (X_6) can be computed as follows:

$$P(X_6 = a) = \sum_{b \in Y} \; P(X_6 = a|Y = b) \; P(Y = b)$$

For instance, $P(X_6 = a)$ for a = 0 and a = 3 are computed as:

$$\begin{aligned} P(X_6 = 0) &= Cell_{00} * Col_0 \; + \; Cell_{01} * Col_1 + \; \ldots \; + Cell_{05} * Col_5 \\ &= 0.125 * 0.011 \; + \; 0.118 * 0.044 + \; \ldots \; + 0.045 * 0.144 = 0.055 \end{aligned}$$

Table F.14: The probability of attribute ContractandPrice conditioned on Satisfaction generated from WEKA

ContractandPrice (X_6)\Satisfaction (Y)	b					
a	**0**	**1**	**2**	**3**	**4**	**5**
0	0.125	0.118	0.065	0.039	0.056	0.045
1	0.125	0.353	0.129	0.092	0.040	0.023
2	0.250	0.176	0.419	0.237	0.089	0.045
3	0.125	0.176	0.258	0.526	0.403	0.273
4	0.250	0.118	0.065	0.092	0.347	0.523
5	0.125	0.059	0.065	0.013	0.065	0.091

Table F.15: The probability distribution of Satisfaction generated from WEKA

	0	**1**	**2**	**3**	**4**	**5**
Satisfaction (Y)	0.011	0.044	0.096	0.263	0.441	0.144

$$
\begin{aligned}
P(X_6 = 3) &= Cell_{30} * Col_0 \ + \ Cell_{31} * Col_1 + \ \dots \ + Cell_{35} * Col_5 \\
&= 0.125 * 0.011 \ + \ 0.176 * 0.044 + \ \dots \ + 0.273 * 0.144 = 0.390
\end{aligned}
$$

This computation is processed for other attribute's value which produced the table of probability distribution of attribute ContractandPrice shown in Table F.16.

Table F.16: The probability of attribute ContractandPrice

ContractandPrice (X_6)	**Probability**
0	**0.055**
1	0.075
2	0.159
3	**0.390**
4	0.267
5	0.055

2. Calculate $P(X_i = a)$ and $P(X_i|Y = b)$ on attribute X_i that has two parents

In case attribute X_i has two parents in which its conditional probability distribution contains value in this form $P(X_i|Y, X_j)$, $P(X_i|Y = b)$ can be computed by using the law of total probability as previously stated.

Considering attribute Equipment (X_1) which has two parents namely Satisfaction and ContractandPrice (X_6), $P(X_1 = a|Y = b)$ can be calculated from $P(X_1 = a|Y = b, X_6)$ as follows, where a and b are the possible value of attribute and Satisfaction which are 0-5.

$$P(X_1 = a|Y = b) = \sum_{c \in X_6} \ P(X_1 = a|Y = b, X_6) \ P(X_6)$$

Given the first six rows in which Satisfaction = 0 from 36 rows (Figure F.4) of the conditional probability distribution of attribute Equipment as shown in Table F.17 and probability distribution of attribute CondandPrice (X_6), shown in Table F.16, conditional probability of attribute $X_1 = a$ given $Y = b$ can be computed as the summation of product of value in $Cell_{bca}$ of Table F.17 and value in Row_c of Table F.16 where value range of c is 0-5.

Table F.17: Part of the probability distribution table of attribute Equipment conditioned on two parents generated from WEKA

Parent		**Probability distribution of Equipment**					
Satisfaction (b)	**ContractandPrice (c)**	**0**	**1**	**2** (a)	**3**	**4**	**5**
	0	0.167	0.167	0.167	0.167	0.167	0.167
	1	0.167	0.167	0.167	0.167	0.167	0.167
0	2	0.143	0.143	0.143	0.286	0.143	0.143
	3	0.167	0.167	0.167	0.167	0.167	0.167
	4	0.143	0.143	0.143	0.143	0.286	0.143
	5	0.167	0.167	0.167	0.167	0.167	0.167

For instance, $P(X_1 = a|Y = b)$ for $a = 0$ and $b = 0$ is computed as:

$$\begin{aligned} P(X_1 = 0|Y = 0) &= Cell_{000} * Row_0 \ + \ Cell_{010} * Row_1 + \ \ldots \ + Cell_{050} * Row_5 \\ &= 0.167 * 0.055 \ + \ 0.167 * 0.075 + \ \ldots \ + 0.167 * 0.055 \\ &= 0.009 \ + \ 0.012 \ + \ \ldots \ + \ 0.009 = 0.157 \end{aligned}$$

For instance, $P(X_1 = a|Y = b)$ for $a = 3$ and $b = 0$ is computed as:

$$\begin{aligned} P(X_1 = 3|Y = 0) &= Cell_{003} * Row_0 \ + \ Cell_{013} * Row_1 + \ Cell_{023} * Row_2 \ + \ \ldots \ + Cell_{053} * Row_5 \\ &= 0.167 * 0.055 \ + \ 0.167 * 0.075 + \ 0.286 * 0.159 \ + \ \ldots \ + 0.167 * 0.055 \\ &= 0.009 \ + \ 0.012 \ + \ 0.045 \ + \ \ldots \ + \ 0.009 = 0.179 \end{aligned}$$

This computation is processed for all combination of ContractandPrice and Satisfaction which produced the table of conditional probability distribution of attribute Equipment shown in Table F.18.

Table F.18: The conditional probability distribution table of attribute Equipment

Equipment (X_1) \ Satisfaction	**0**	**1**	**2**	**3**	**4**	**5**
0	**0.157**	0.13	0.101	0.082	0.052	0.101
1	0.157	0.2	0.11	0.06	0.049	0.082
2	0.157	0.195	0.192	0.109	0.056	0.082
3	**0.179**	0.165	0.175	0.417	0.179	0.133
4	0.195	0.179	0.276	0.264	0.585	0.435
5	0.157	0.13	0.147	0.067	0.08	0.168

The probability distribution of attribute Equipment $P(X_i = a)$ then can be computed from Table F.18 and Table F.15 regarding the procedure described in the Step 1 of the pre-calculation procedure.

For instance, $P(X_1 = a)$ for a = 0 and a = 3 are computed as:

$$\begin{aligned} P(X_1 = 0) &= Cell_{00} * Col_0 \ + \ Cell_{01} * Col_1 + \ \ldots \ + Cell_{05} * Col_5 \\ &= 0.157 * 0.011 \ + \ 0.130 * 0.044 + \ \ldots \ + 0.101 * 0.144 = 0.076 \end{aligned}$$

$$\begin{aligned} P(X_1 = 3) &= Cell_{30} * Col_0 \ + \ Cell_{31} * Col_1 + \ \ldots \ + Cell_{35} * Col_5 \\ &= 0.179 * 0.011 \ + \ 0.165 * 0.044 + \ \ldots \ + 0.133 * 0.144 = 0.234 \end{aligned}$$

This computation is processed for other attribute's value which produced the table of probability distribution of attribute Equipment shown in Table F.19.

Table F.19: The probability of attribute Equipment

Equipment (X_1)	**Probability**
0	**0.076**
1	0.070
2	0.094
3	**0.234**
4	0.427
5	0.099

F.5.2 Calculate *importance* as Mutual Information

1. Calculate Entropy of each attribute $h(X_i)$

Given the probability density function for attribute Equipment and CondandPrice shown in Table F.19 and Table F.16 respectively, the entropy of attribute $h(X_i)$ for $i = 1$ and $i = 6$ are computed following Equation 3.7 :

$$\begin{aligned} h(X_1) = &- 0.076 * \log_2(0.076) \ - 0.070 * \log_2(0.070) \ - 0.094 * \log_2(0.094) \\ &- 0.234 * \log_2(0.234) \ - 0.427 * \log_2(0.427) \ - 0.099 * \log_2(0.099) = 2.218 \end{aligned}$$

$$\begin{aligned} h(X_6) = &- 0.055 * \log_2(0.055) \ - 0.075 * \log_2(0.075) \ - 0.159 * \log_2(0.159) \\ &- 0.390 * \log_2(0.390) \ - 0.267 * \log_2(0.267) \ - 0.055 * \log_2(0.055) = 2.200 \end{aligned}$$

By the similar approach, the entropy of the other four attributes are calculated and shown in the second column of Table F.21

2. Calculate conditional entropy $h(X_i|Y)$

The entropy of X_i conditioned on Y, $h(X_i|Y)$ is computed as Equation 3.8. This equation can be broken into two main components which are $P(Y = b)$ (the probability of the overall customer satisfaction Y with the satisfaction level b) and $h(X_i|Y = b)$ (the average conditional entropy of X_i).

In the context of this computational example, $P(Y = b)$ is presented in Table F.15 and $h(X_i|Y = b)$ can be computed from Equation 3.9. Given Table F.15 and Table F. 18 represents conditional probability distribution table of attribute Equipment $P(X_1 = a|Y = b)$, conditional entropy of attribute Equipment $h(X_1|Y = b)$ for $b = 0$ can be computed as follows:

$$\begin{aligned} h(X_1|Y = 0) = &- 0.157 * \log_2(0.157) \ - 0.157 * \log_2(0.157) \ - 0.157 * \log_2(0.157) \\ &- 0.179 * \log_2(0.179) \ - 0.195 * \log_2(0.195) \ - 0.157 * \log_2(0.157) = 2.579 \end{aligned}$$

By the similar approach, the entropy of $h(X_1|Y = b)$ for $b = 1, \ldots, 5$ are computed and shown in Table F.20.

Table F.20: The average conditional entropy of attribute Equipment, $h(X_1|Y = b)$

	b					
	0	**1**	**2**	**3**	**4**	**5**
$h(X_1\|Y = b)$	**2.579**	2.564	2.499	2.184	1.855	2.266

Finally, the computational of $h(X_1|Y)$ can be calculated as a summation of product between $h(X_1|Y = b)$ and $P(Y = b)$ for $b = 0, \ldots, 5$ (see Equation 3.8):

$$\begin{aligned} h(X_1|Y) = &\ 2.579 * 0.011 \ + \ 2.564 * 0.044 \ + \ 2.499 * 0.096 \ + \ 2.184 * 0.263 \\ &+ 1.855 * 0.441 \ + \ 2.266 * 0.144 = 2.102 \end{aligned}$$

Using the same procedure, the computational results of $h(X_i|Y)$ for all attribute are shown in the third column of Table F.21.

3. Calculate Mutual Information. Given the entropy and conditional entropy of each attribute (see the second and third column of Table F.21), mutual information of each attribute can be computed as Equation 3.6, and then *importance* of each attribute can be computed as percentage contribution of the mutual information. These two values are presented in the fourth and fifth column of the same table respectively.

Table F.21: Mutual Information and *importance* of attributes derived from Bayesian Networks

X_i	$h(X_i)$	$h(X_i\|Y)$	**MI(X_i,Y)**	**Importance**
X_1-Equipment and System	**2.218**	**2.102**	0.115	20.409
X_2-Sales Support	2.466	2.445	0.020	3.582
X_3-Technical Support	2.374	2.259	0.115	20.401
X_4-Supplies and Orders	2.275	2.197	0.077	13.661
X_5-Purchasing Support	2.399	2.357	0.042	7.432
X_6-Contracts and Pricing	**2.200**	2.005	0.195	34.515

Appendix G

Student satisfaction survey material

G.1 Student Satisfaction Survey

Department of Computer Engineering, Faculty of Engineering at KamphaengSaen, Kasetsart University

The survey asks about your level of satisfaction with, and opinion on the importance of, aspects of the Department of Computer Engineering.

Section I. Level of Satisfaction and Importance with the aspects of Department of Computer Engineering

Direction: Considering your educational experience at the department of Computer Engineering at Kamphaeng Saen campus, please indicate your level of satisfaction with the aspects of this department by ticking the response that best corresponds to your opinion using the following scales:

Score	Interpretation on Level of Satisfaction
5	Very Satisfied
4	Satisfied
3	Neutral
2	Dissatisfied
1	Very Dissatisfied

Score	Interpretation on Level of Importance
5	Extremely important
4	Very important
3	Moderately important
2	Slightly important
1	Not important

1. Academic Personal	Level of Satisfaction					Level of Importance				
	5	4	3	2	1	5	4	3	2	1
Teaching ability of teaching staff										
Subject expertise of teaching staff										
Friendliness of the teaching staff										
Availability of teaching staff										
Advice and support in learning										
Overall satisfaction level with Academic Personal										

2. Teaching and Learning	Level of Satisfaction					Level of Importance				
	5	4	3	2	1	5	4	3	2	1
Lecture materials										
e-learning resources										
Assessments (clarity and timely feedback)										
Class size										
Accurate and up-to-date unit content										
Teaching facilities and classroom condition										
Overall satisfaction level with Teaching and Learning										

3. Administration	Level of Satisfaction					Level of Importance				
	5	4	3	2	1	5	4	3	2	1
Knowledge of rules and procedures of staff members										
Knowledge of the information about courses, exams, activities of staff members										
Interest in solving the problems of student by staff members										
Friendliness of staff members										
Ability of the staff members to provide services in a timely manner										
Overall satisfaction level with Administration										

4. Computer Facilities	Level of Satisfaction					Level of Importance				
	5	4	3	2	1	5	4	3	2	1
Quality of computer facilities (Hardware and Software)										
Availability of computer facilities										
Availability of internet access										
Availability of printing and photocopying facilities										
Overall satisfaction level with Computer Facilities										

5. Extra-Curricular Activities	Level of Satisfaction					Level of Importance				
	5	4	3	2	1	5	4	3	2	1
Cultural exchange programs with foreign countries										
Field trips										
Moral development activities										
Health development activities										
Interpersonal skills development activities										
Personal learning and thinking skills development activities										
Social volunteer activities										
Overall satisfaction level with Extra-Curricular Activities										

6. Additional services	**Level of Satisfaction**					**Level of Importance**				
	5	4	3	2	1	5	4	3	2	1
Financial aid for students										
Medical support to students										
Department website (Updated content and easy to find information)										
Library										
Overall satisfaction level with Additional services										

Section II. Personal Information

Direction: Please tick the circle that best corresponds to your answer for each question below

1. Gender ○ Male ○ Female
2. Study Level ○ 1 ○ 2 ○ 3 ○ 4 ○ other (specify)
3. Program ○ Regular ○ Special

Thank you for taking the time to complete this questionnaire

G.2 Participant Information sheet

Study Title: Mining Survey Data for SWOT analysis
Researcher: Boonyarat Phadermrod
Ethics number: ERGO/FPSE/14268

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

I am completing a PhD research about mining survey data for SWOT analysis which aimed to generate prioritized SWOT factors based on the customer's perception. To evaluate the proficiency of the proposed approach in the real-world situation, the case study of Department of Computer Engineering (DoCE), Kasetsart University will be conducted. Through this case study, your response on the level of satisfaction and importance with the aspects of the DoCE will be analysed to produce SWOT using data mining technique.

Why have I been chosen?

This research is trying to get a level of satisfaction with, and a level of the importance of, aspects of the DoCE. Thus, you are chosen as you are a student in the DoCE, Kasetsart University (Bangkhen Campus and Kamphaeng Saen campus).

What will happen to me if I take part?

The questionnaire will take no longer than 10 minutes to complete. There are three sections in the questionnaire. Section 1 and 2 asks about your level of satisfaction and importance with the aspects of the DoCE respectively. Section 3 asks about your general information.

Are there any benefits in my taking part?

By taking part, you have the opportunity to help us develop the new method to generate SWOT based on satisfaction survey which can be used in government, academic institution and private company. You're also help the DoCE to improve their quality of each aspect to meet your satisfaction.

Are there any risks involved?

There is no risk involved for participants completing the questionnaire.

Will my participation be confidential?

The name or student ID of the participants will not be taken and participation will be kept anonymous. All data will be safe in a protected computer. All will be destroyed once the research is completed.

What happens if I change my mind?

You have the right to withdraw from doing the questionnaire at any time by exiting the webpage.

What happens if something goes wrong?

If you have any concern or complaint with this research please contacts me (Boonyarat Phadermrod: bp6g12@soton.ac.uk). If you have a trouble with the online forms, please contact the ECS School Office on school@ecs.soton.ac.uk 02380 592909.

Where can I get more information?

If you would like more information on this research please feel free to contact me (Boonyarat Phadermrod: bp6g12@soton.ac.uk)

G.3 Result of language consistency translation test of student satisfaction survey

Remark: +1 the question/ instruction is consistent with the English version

0 undecided about whether Thai question/ instruction is consistent with the English version

−1 the question/ instruction is not consistent with the English version

Table G.1: Summary of language consistency translation test of student satisfaction survey

	Question/ Instruction (English)	**Question/ Instruction** (Thai)	**Language consistency score** Expert 1	Expert 2	Expert 3	Expert 4	**Total Score**	**Comment**
1	The survey asks about your opinion, including your level of satisfaction and importance toward the aspects of the Computer Engineering department.	แบบประเมินนี้สำรวจถึงความคิดเห็นของนิสิตเกี่ยวกับระดับความพึงพอใจ และระดับความสำคัญในด้านต่างๆ ของภาควิชาวิศวกรรมคอมพิวเตอร์	+1	+1	+1	0	3	Change “your opinion”to “student’s opinion”
2	Section I. Level of Satisfaction and Importance with the aspects of Computer Engineering department	ตอนที่ 1. ระดับความพึงพอใจ และระดับความสำคัญในด้านต่างๆ ของภาควิชาวิศวกรรมคอมพิวเตอร์	+1	+1	+1	+1	4	
3	Direction: Considering your educational experience at the Computer Engineering department, please indicate your level of satisfaction and importance with the aspects of this department by ticking the appropriate response using the following scales:	คำชี้แจง: กรุณาระบุระดับความพึงพอใจ และระดับความสำคัญในด้านต่างๆ ของภาควิชาวิศวกรรมคอมพิวเตอร์จากประสบการณ์การศึกษา ณ ภาควิชาวิศวกรรมคอมพิวเตอร์โดยทำเครื่องหมายถูก (✓) ในช่องของตัวเลขที่ตรงกับระดับความคิดเห็นของนิสิต (ความหมายของตัวเลขแสดงระดับความพึงพอใจ และ ความสำคัญอธิบายด้านล่าง)	+1	0	+1	+1	3	Change “appropriate response” to “the response that best corresponds to your opinion”
4	Score, Interpretation on Level of Satisfaction Very Satisfied Satisfied Neutral Dissatisfied Very Dissatisfied	ระดับความพึงพอใจ, คำอธิบายระดับความพึงพอใจ พึงพอใจมาก พึงพอใจ ปานกลาง ไม่พึงพอใจ ไม่พึงพอใจมาก	+1	+1	+1	+1	4	
5	Score, Interpretation on Level of Importance Extremely important Very important Moderately important Slightly important Not at all important	ระดับความสำคัญ, คำอธิบายระดับความสำคัญ สำคัญมากที่สุด สำคัญมาก ค่อนข้างสำคัญ สำคัญน้อย ไม่สำคัญ	+1	0	+1	+1	3	

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score				Total Score	Comment
			Expert 1	Expert 2	Expert 3	Expert 4		
6	1. Academic Personal - Teaching ability of teaching staff - Subject expertise of teaching staff - Friendliness of the teaching staff - Availability of teaching staff - Advice and support in learning - Overall satisfaction level with Academic Personal	1. อาจารย์ประจำภาควิชา - ความสามารถในการสอนของอาจารย์ - ความเชี่ยวชาญในเนื้อหาวิชาของอาจารย์ - ความเป็นกันเองของอาจารย์ - การให้โอกาสแก่นิสิตเข้าพบนอกเวลาเรียนของอาจารย์ - การให้คำปรึกษาและความช่วยเหลือในการเรียน - ระดับความพึงพอใจโดยรวมต่ออาจารย์ประจำภาควิชา	+1	+1	0	+1	3	Suggest to use the word “Lecturer ” since teaching staffs are included demonstrators.
7	2. Teaching and Learning - Lecture materials - e-learning resources - Assessments (clarity and timely feedback) - Class size - Accurate and Up-to-date unit content - Equipped teaching facilities and learning areas - Overall satisfaction level with Teaching and Learning	2. การเรียนการสอน - เอกสารประกอบการเรียนการสอน - สื่อการเรียนรู้อิเล็กทรอนิกส์ - การวัดประเมินผล (ชัดเจน และได้รับแจ้งภายในเวลาที่กำหนด) - จำนวนนิสิตในชั้นเรียน - ความถูกต้องและความทันสมัยของเนื้อหาวิชา - อุปกรณ์การเรียนการสอน และสภาพห้องเรียน - ระดับความพึงพอใจโดยรวมต่อการเรียนการสอนของภาควิชา	+1	+1	0	+1	3	Change to “Teaching facilities and classroom condition”
8	3. Administration - Knowledge of rules and procedures of staff members - Knowledge of the information about courses, exams, activities of staff members - Interest in solving the problems of student by staff members - Friendliness of staff members - Ability of the staff members to provide services in a timely manner - Overall satisfaction level with Administration	3. งานบริหารและธุรการ - ความรู้เรื่องข้อบังคับ ระเบียบมหาวิทยาลัยของเจ้าหน้าที่ - ความรู้เรื่องข้อมูลข่าวสารเกี่ยวกับ รายวิชา การสอบ และ กิจกรรมนิสิตของเจ้าหน้าที่ - ความใส่ใจในการแก้ปัญหาให้แก่นิสิตของเจ้าหน้าที่ - ความเป็นกันเองของเจ้าหน้าที่ - ความสามารถในการให้บริการนิสิตในระยะเวลาที่เหมาะสมของเจ้าหน้าที่ - ระดับความพึงพอใจโดยรวมต่องานบริหารและ ธุรการของภาควิชา	+1	+1	+1	+1	4	

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score Expert 1	Expert 2	Expert 3	Expert 4	Total Score	Comment
9	4. Computer Facilities - Quality of computer facilities (Hardware and Software) - Availability of computer facilities - Availability of Internet access - Availability of printing and photocopying facilities - Overall satisfaction level with Computer Facilities	4. บริการคอมพิวเตอร์ - คุณภาพของอุปกรณ์คอมพิวเตอร์ (ฮาร์ดแวร์และซอฟต์แวร์) - ความพร้อมให้บริการของอุปกรณ์คอมพิวเตอร์ - ความพร้อมให้บริการเชื่อมต่ออินเตอร์เน็ต - ความพร้อมให้บริการของเครื่องพิมพ์และเครื่องถ่ายเอกสาร - ระดับความพึงพอใจโดยรวมต่องานบริการคอมพิวเตอร์	+1	+1	+1	+1	4	
10	5. Extra-Curricular Activities - Cultural Exchange programs with foreign - Field trips - Moral development activities - Health development activities - Interpersonal skills development activities - Personal learning and thinking skills development activities - Social volunteer activities - Overall satisfaction level with Extra-Curricular Activities	5. กิจกรรมเสริมหลักสูตร - โครงการแลกเปลี่ยนวัฒนธรรมกับต่างประเทศ - ทัศนะศึกษาดูงาน - กิจกรรมพัฒนาคุณธรรม - กิจกรรมพัฒนาสุขภาพ - กิจกรรมพัฒนาความสัมพันธ์ระหว่างบุคคล - กิจกรรมพัฒนาทักษะการคิดและการเรียนรู้ - กิจกรรมเพื่อสังคม - ระดับความพึงพอใจโดยรวมต่อกิจกรรมเสริมหลักสูตรของภาควิชา	+1	+1	+1	+1	4	
11	6. Additional services - Financial aid for students - Medical support to students - Department website (Updated content and easy to find information) - Library - Overall satisfaction level with Additional services	6. บริการด้านอื่นๆ - ความช่วยเหลือด้านการเงินแก่นิสิต - การรักษาพยาบาลสำหรับนิสิต - เว็บไซต์ประจำภาควิชา (เนื้อหาทันสมัยและสามารถค้นหาข้อมูลได้ง่าย) - ห้องสมุด - ระดับความพึงพอใจโดยรวมต่อบริการด้านอื่นๆ ของภาควิชา	+1	+1	+1	+1	4	

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score				Total Score	Comment
			Expert 1	Expert 2	Expert 3	Expert 4		
12	Overall satisfaction level with your educational experience at Department of Computer Engineering	ระดับความพึงพอใจโดยรวมต่อประสบการณ์การศึกษา ณ ภาควิชาวิศวกรรมคอมพิวเตอร์	1	1	1	1	4	
13	Section II. Personal Information	ตอนที่ 2. ข้อมูลพื้นฐานเกี่ยวกับผู้ตอบแบบสอบถาม	1	1	1	1	4	
14	Direction: Please tick the circle that best corresponds to your answer for each question below	คำชี้แจง: กรุณาทำเครื่องหมายถูก (✓) ในวงกลมที่ตรงกับข้อมูลของนิสิตมากที่สุด	1	1	1	1	4	
15	Gender ○ Male ○ Female	เพศ ○ ชาย ○ หญิง	1	1	1	1	4	
16	Study Level	ชั้นปีที่ศึกษา	1	1	1	1	4	
17	Program ○ Regular ○ Special	โปรแกรมการศึกษา ○ ภาคปกติ ○ ภาคพิเศษ	1	1	1	1	4	

Appendix H

Assumption checking of the multiple linear regression on CPE-KU-KPS dataset

There were six attributes in CPE-KU-KPS dataset namely Academic Personal (Teacher), Teaching and Learning (Teaching), Administration (Admin), Computer Facilities (CompFac), Extra-Curricular Activities (xActivity) and Additional Services (AddService). To build regression model for each attribute, items related to the attribute were set as independent variables and the overall student satisfaction (OCSstudent) was set as the dependent variable.

The assumptions of multiple linear regression were checked for each attribute in the dataset to ensure that the result of regression was valid. Assumption testing results for each group of attribute are summarized in Table H.1 and explained in the following sections.

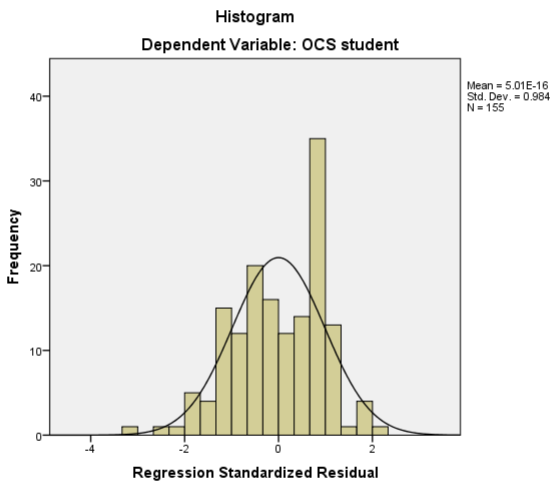
Table H.1: Summary of assumption checking of MLR on six attributes of CPE-KU-KPS dataset

Assumptions	**Attribute**					
	Teacher	Teaching	Admin	CompFac	xActivity	AddService
Normality of residual	✓	✓	✓	✓	✓	✓
Linearity	✓	✓	✓	✓	✓	✓
Homoscedasticity of residual	✓	✓	✓	✓	✓	✓
Independence of residual	✓	✓	✓	✓	✓	✓
Multicollinearity	✓	✓	✓	✓	✓	✓

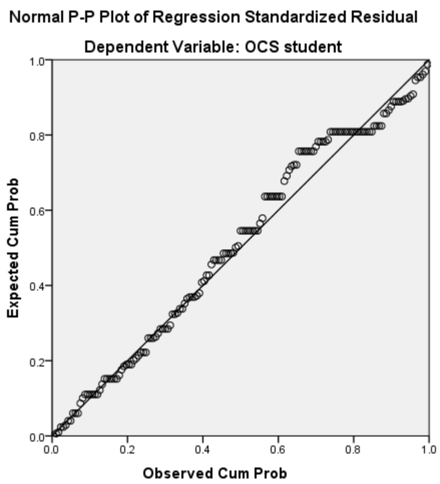
With regard to Table H.1, a symbol ✓ indicates that the assumption was met and a symbol × indicates that the assumption was violated. According to the table it shows that all assumptions were met across six attributes of the dataset.

H.1 Normally distributed residual

For each group of attribute, assumption of normality was visually tested by observing the histogram and normal P-P plot of the standardized residual. This assumption was also tested by observing the skewness and kurtosis of unstandardized residual in which skew and kurtosis values between −1.0 and +1.0 are reasonable range to accept that data is reasonably close to normal (George and Mallery, 2003). Histograms and normal P-P plots of the standardized residual for each group were shown in Figure H.1 - H. 6. And descriptive statistic of unstandardized residual for each group was shown in Table H.2.

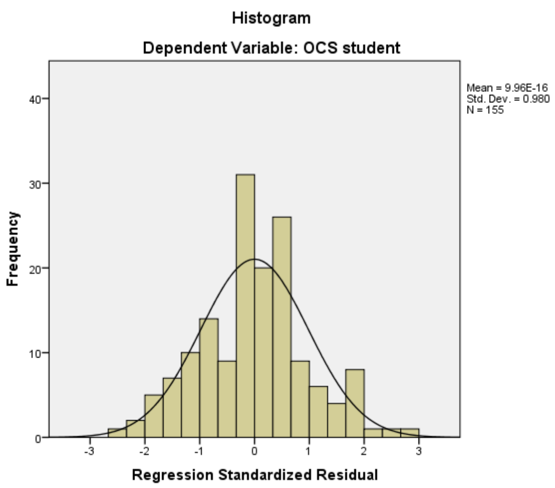


(a) Histogram plot of residual

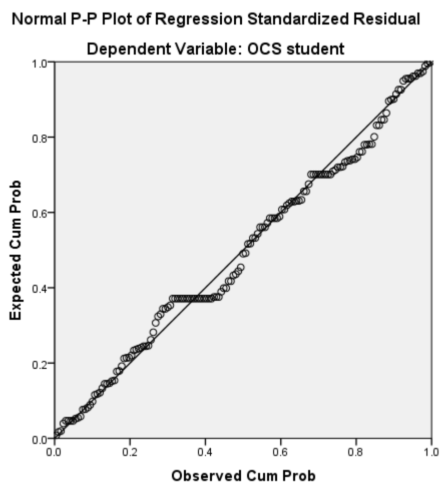


(b) Normality plot of residual

Figure H.1: Histogram and normality plots of residual (Teacher)

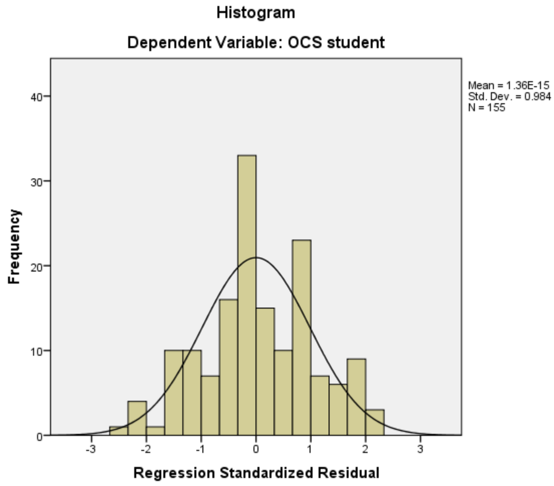


(a) Histogram plot of residual

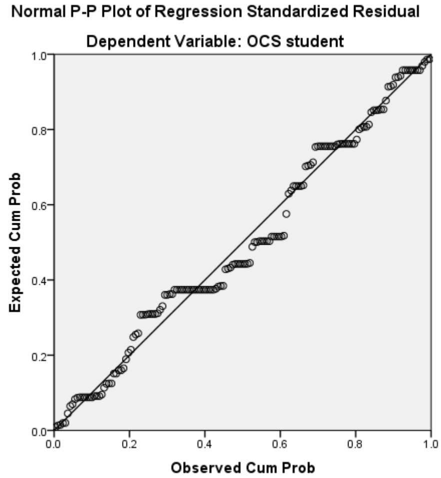


(b) Normality plot of residual

Figure H.2: Histogram and normality plots of residual (Teaching)

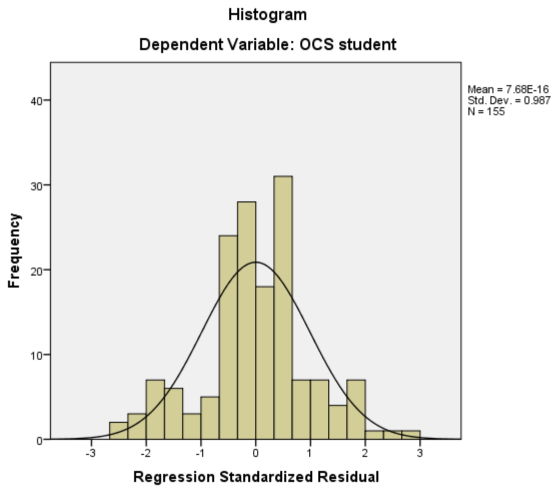


(a) Histogram plot of residual

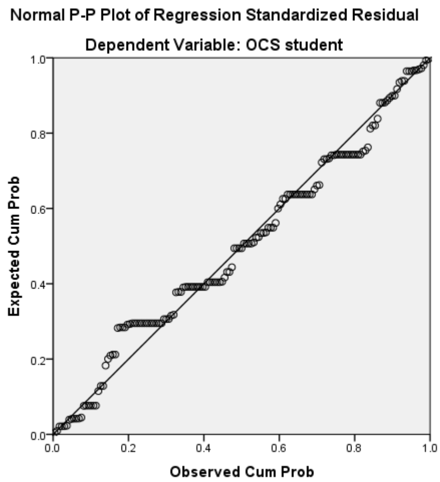


(b) Normality plot of residual

Figure H.3: Histogram and normality plots of residual (Admin)



(a) Histogram plot of residual

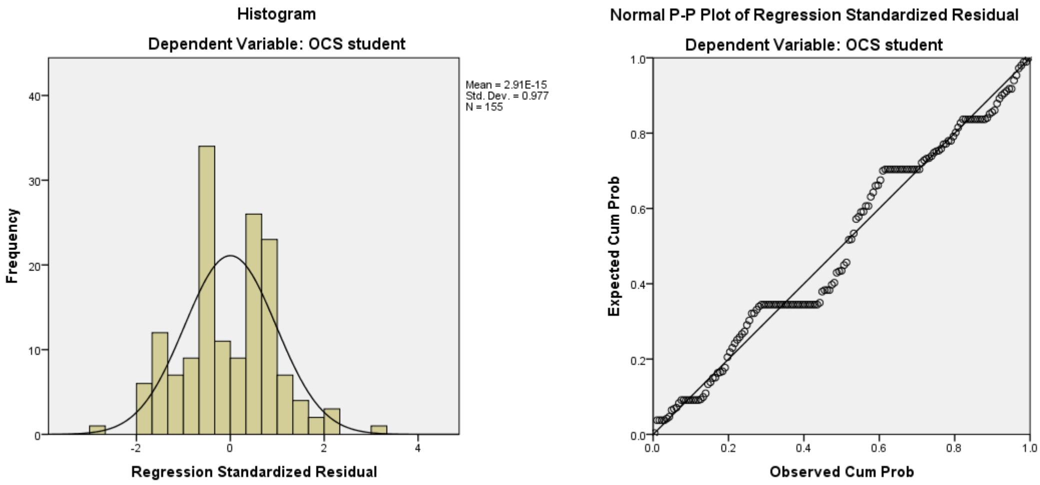


(b) Normality plot of residual

Figure H.4: Histogram and normality plots of residual (CompFac)

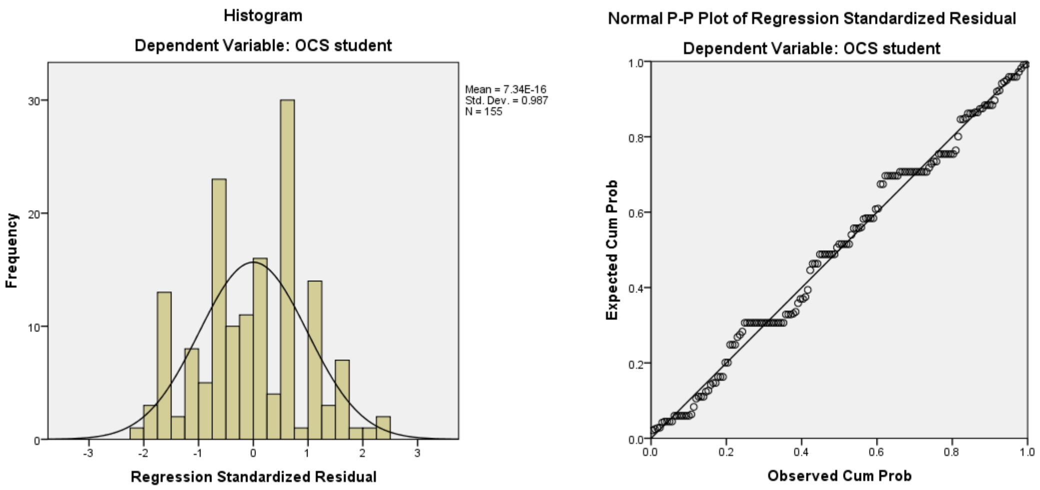
The histogram of four attributes namely Teacher (Figure H.1), Teaching (Figure H.2), CompFac (Figure H.4) and AddService (Figure H.6) were appeared to look like a bell and their normal P-P plots were approximately a straight line which suggested that the residuals of these groups were not deviated from normal. Corresponded to the histogram and P-P plots, skewness and kurtosis of these groups shown in Table H.2 were within the range to accept that their residuals were normally distributed.

For attribute Admin and xActivity, although their normal P-P plots were slightly departed from the diagonal line indicated a presence of non-normality of residuals (see Figure H.3 and Figure H.5), skewness and kurtosis of these two groups shown in Table H.2 were within the range to accept that their residuals were normally distributed.



(a) Histogram plot of residual (b) Normality plot of residual

Figure H.5: Histogram and normality plots of residual (xActivity)



(a) Histogram plot of residual (b) Normality plot of residual

Figure H.6: Histogram and normality plots of residual (AddService)

Thus, it was reasonable to conclude that the residuals of all groups in the dataset were normally distributed.

H.2 Linear relationship between independent and dependent variables

To check linearity assumption, a scatterplot of the standardized residual and standardized predicted value was drawn for each group of variables (see Figure H.7- H.12). The residual plot for predicted values of an outcome variable (OCS student) against the residuals of all attributes showed that the dots were not constantly spread over the

Table H.2: Descriptive statistic of unstandardized residual for each attribute of CPE-KU-KPS dataset

Unstandardized Residual	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
Teacher	155	0.00	0.4691	-0.435	0.195	-0.009	0.387
Teaching	155	0.00	0.4445	0.227	0.195	0.204	0.387
Admin	155	0.00	0.4809	0.048	0.195	-0.311	0.387
CompFac	155	0.00	0.474	-0.006	0.195	0.262	0.387
xActivity	155	0.00	0.4219	0.135	0.195	0.159	0.387
AddService	155	0.00	0.4397	0.004	0.195	-0.476	0.387

horizontal line; however, there was no sign of any curve pattern. Therefore, it cannot conclude that the linearity assumption was violated.

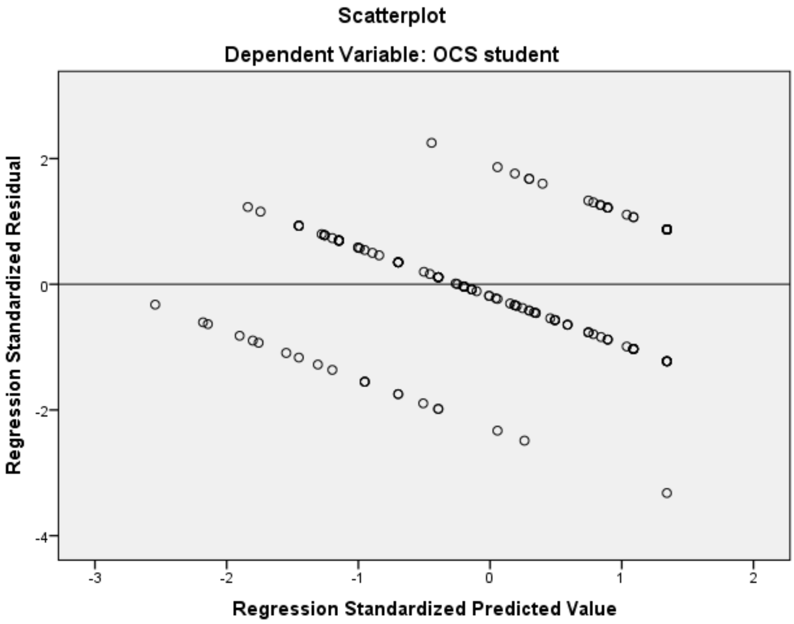


Figure H.7: Scatterplot of residual for the relationship between five independent variables of attribute Teacher and the OCS student

For each group of attribute, non-linear relationships can also be detected by adding curve components which are quadratic (variable^2) and cubic (variable^3) to the regression equation to see whether this increases the explained variance (R-square). Specifically, for each independent variable of each group, the R-square of the linear component and the two curve components was statistically compared through the test of equality for several independent correlation coefficients (r) following an approach described in Kenny (1987). If there is no departure from linearity in the relationship, it is expected to see that the difference between the linear correlation coefficient (r) and the non-linear (curve) correlation coefficient is not significantly different (p-value > 0.05).

Table H.3 shows the R-square of linear/non-linear components of each independent variable of attribute Teacher against the OCS student. Regarding this table, R-squares

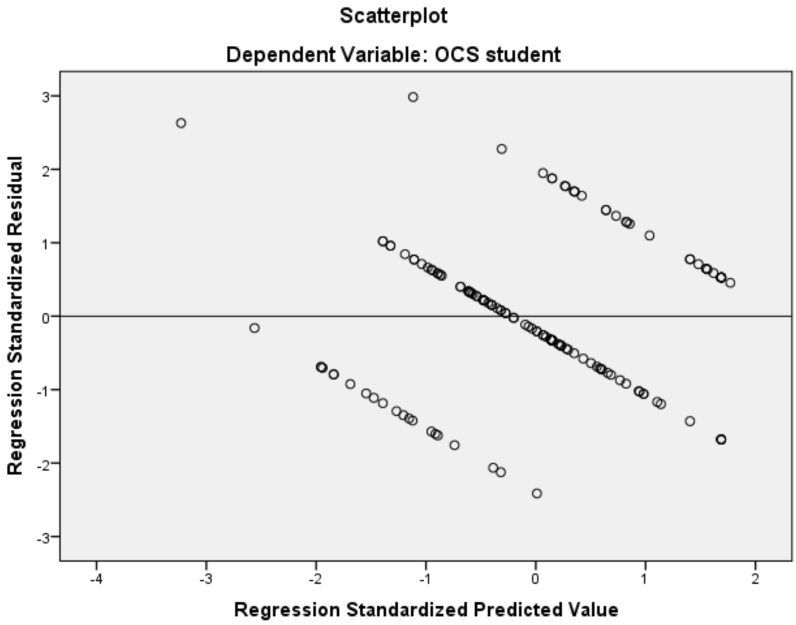


Figure H.8: Scatterplot of residual for the relationship between six independent variables of attribute Teaching and the OCS student

of the curve component of all variables of attribute Teacher were equal to or slightly higher than R-square of the linear component.

The test for equality of several independent correlation coefficients shown in Table H. 4 indicated that there was no significant difference among correlations of the linear/ non-linear components for each independent variable of attribute Teacher. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute Teacher and the OCS student is linear.

Table H.3: R-square of the components of regression equation (attribute Teacher)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
Teacher 1	0.260	0.261	0.261
Teacher 2	0.269	0.270	0.270
Teacher 3	0.189	0.189	0.189
Teacher 4	0.186	0.186	0.186
Teacher 5	0.207	0.209	0.209

Table H.5 shows the R-square of linear/non-linear components of each independent variable of attribute Teaching against the OCS student. Regarding this table, R-squares of the curve component of two variables named Teaching 1 and Teaching 6 were equal to or slightly higher than R-square of the linear component. R-squares of the curve component of the other four variables named Teaching 2-5 were greatly higher than R-squares of the linear component indicated a sign of non-linearity.

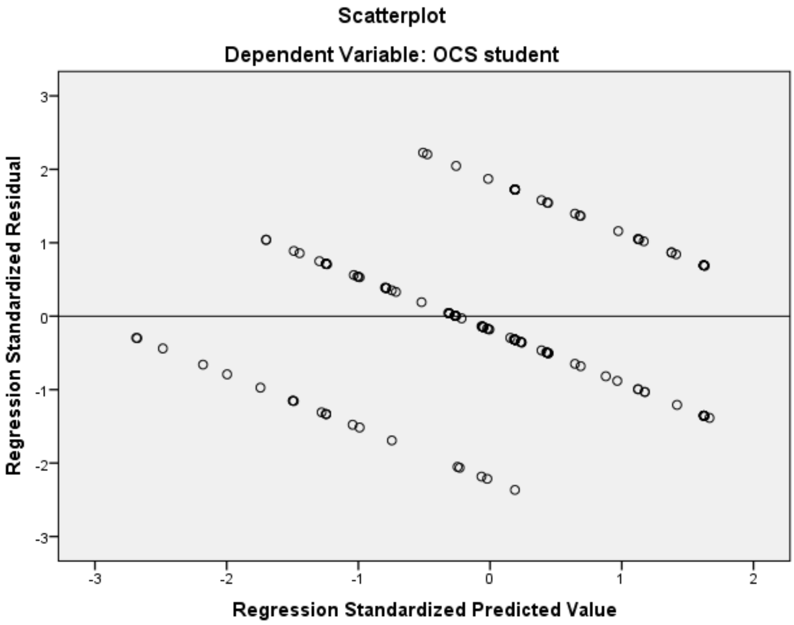


Figure H.9: Scatterplot of residual for the relationship between five independent variables of attribute Admin and the OCS student

Table H.4: Chi-square statistic for the difference test of equality of correlations (attribute Teacher)

Independent variable	χ^2	**df**	**p-value**	**Equality**
Teacher 1	0.000	2	1.000	Yes
Teacher 2	0.000	2	1.000	Yes
Teacher 3	0.000	2	1.000	Yes
Teacher 4	0.000	2	1.000	Yes
Teacher 5	0.001	2	1.000	Yes

However, considering the test for equality of several independent correlation coefficients reported in Table H.6, it showed that there was no significant difference among correlations of the linear/non-linear component for each independent variable of attribute Teaching. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute Teaching and the OCS student is linear.

Table H.5: R-square of the components of regression equation (attribute Teaching)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
Teaching 1	0.306	0.309	0.309
Teaching 2	0.256	0.283	0.288
Teaching 3	0.248	0.268	0.265
Teaching 4	0.188	0.197	0.198
Teaching 5	0.142	0.24	0.241
Teaching 6	0.341	0.345	0.346

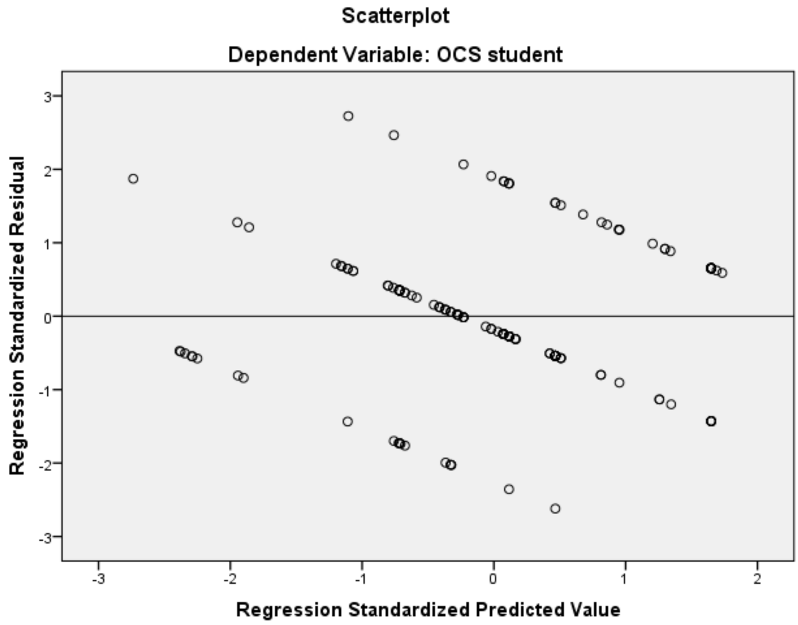


Figure H.10: Scatterplot of residual for the relationship between four independent variables of attribute CompFac and the OCS student

Table H.6: Chi-square statistic for the difference test of equality of correlations (attribute Teaching)

Independent variable	χ^2	df	p-value	Equality
Teaching 1	0.002	2	0.999	Yes
Teaching 2	0.156	2	0.925	Yes
Teaching 3	0.062	2	0.969	Yes
Teaching 4	0.018	2	0.991	Yes
Teaching 5	1.993	2	0.369	Yes
Teaching 6	0.004	2	0.998	Yes

Table H.7 shows the R-square of linear/non-linear components of each independent variable of attribute Admin against the OCS student. Regarding this table, R-squares of the curve component of most variables were equal to or slightly higher than R-square of the linear component. Only the variable named Admin 3 which its R-squares of the curve component were greatly higher than its R-squares of the linear component.

Given R-square from Table H.7, the test for equality of several independent correlation coefficients was conducted and the result was reported in Table H.8. The Table H.8 showed that there was no significant difference among correlations of the linear/non-linear component for each independent variable of attribute Admin. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute Admin and the OCS student is linear.

Table H.9 shows the R-square of linear/non-linear components of each independent variable of attribute CompFac against the OCS student. Regarding this table, R-squares

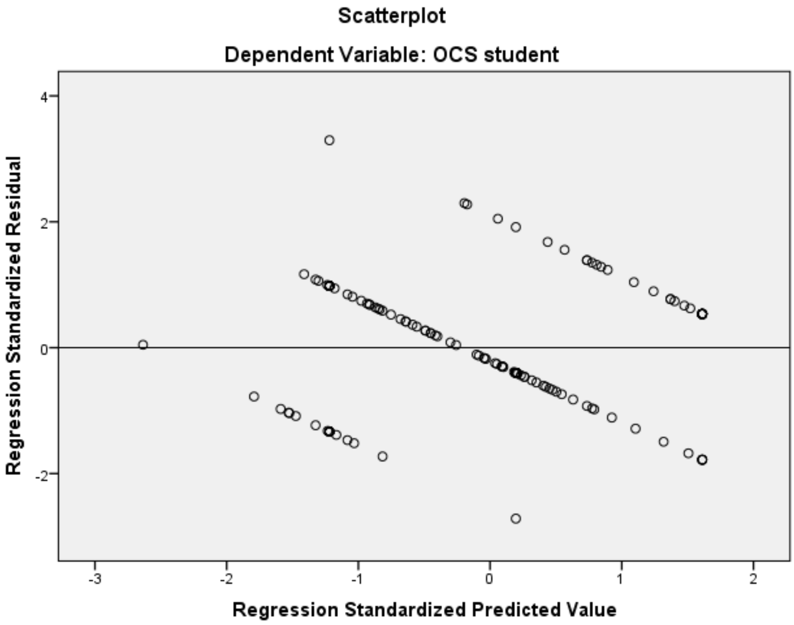


Figure H.11: Scatterplot of residual for the relationship between seven independent variables of attribute xActivity and the OCS student

Table H.7: R-square of the components of regression equation (attribute Admin)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
Admin 1	0.204	0.204	0.204
Admin 2	0.317	0.324	0.324
Admin 3	0.203	0.212	0.213
Admin 4	0.198	0.199	0.201
Admin 5	0.219	0.220	0.225

Table H.8: Chi-square statistic for the difference test of equality of correlations (attribute Admin)

Independent variable	χ^2	**df**	**p-value**	**Equality**
Admin 1	0.000	2	1.000	Yes
Admin 2	0.008	2	0.996	Yes
Admin 3	0.018	2	0.991	Yes
Admin 4	0.001	2	0.999	Yes
Admin 5	0.006	2	0.997	Yes

of the curve component of most variables were equal to or slightly higher than R-square of the linear component. Only the variable named CompFac 4 which its R-squares of the curve component were greatly higher than its R-squares of the linear component indicated a sign of non-linearity.

However, considering the test for equality of several independent correlation coefficients reported in Table H.10, it showed that there was no significant difference among correlations of the linear/non-linear component for each independent variable of attribute

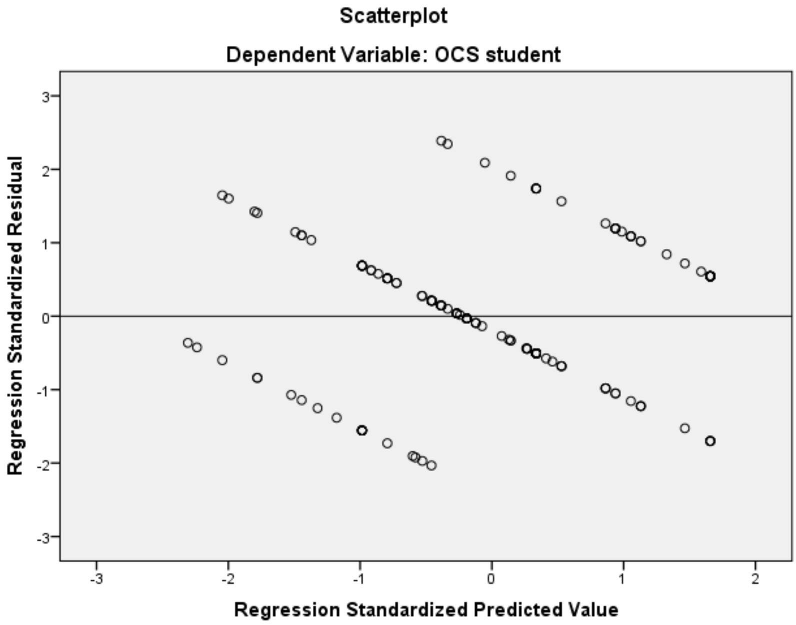


Figure H.12: Scatterplot of residual for the relationship between four independent variables of attribute AddService and the OCS student

CompFac. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute CompFac and the OCS student is linear.

Table H.9: R-square of the components of regression equation (attribute CompFac)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
CompFac 1	0.296	0.300	0.308
CompFac 2	0.317	0.322	0.324
CompFac 3	0.165	0.165	0.165
CompFac 4	0.278	0.281	0.294

Table H.10: Chi-square statistic for the difference test of equality of correlations (attribute CompFac)

Independent variable	χ^2	**df**	**p-value**	**Equality**
CompFac 1	0.019	2	0.990	Yes
CompFac 2	0.007	2	0.997	Yes
CompFac 3	0.000	2	1.000	Yes
CompFac 4	0.038	2	0.981	Yes

Table H.11 shows the R-square of linear/non-linear components of each independent variable of attribute xActivity against the OCS student. Regarding this table, R-squares of the curve component of most variables were equal to or slightly higher than R-square of the linear component. Only the variables named xActivity 4 and xActivity 7 which their R-squares of the curve component were critically higher than their R-squares of the linear component indicated a sign of non-linearity.

However, considering the test for equality of several independent correlation coefficients reported in Table H.12, it showed that there was no significant difference among correlations of the linear/non-linear components for each independent variable of attribute xActivity. Therefore, it was reasonably to conclude that the relationship between each independent variable of attribute xActivity and the OCS student is linear.

Table H.11: R-square of the components of regression equation (attribute xActivity)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
xActivity 1	0.311	0.313	0.314
xActivity 2	0.339	0.350	0.35
xActivity 3	0.415	0.419	0.419
xActivity 4	0.402	0.415	0.424
xActivity 5	0.371	0.371	0.372
xActivity 6	0.413	0.414	0.414
xActivity 7	0.286	0.297	0.296

Table H.12: Chi-square statistic for the difference test of equality of correlations (attribute xActivity)

Independent variable	χ^2	**df**	**p-value**	**Equality**
xActivity 1	0.001	2	0.999	Yes
xActivity 2	0.021	2	0.990	Yes
xActivity 3	0.003	2	0.999	Yes
xActivity 4	0.065	2	0.968	Yes
xActivity 5	0.000	2	1.000	Yes
xActivity 6	0.000	2	1.000	Yes
xActivity 7	0.019	2	0.990	Yes

Table H.13 shows the R-square of linear/non-linear components of each independent variable of attribute AddService against the OCS student. Regarding this table, R-squares of the curve component of all variables were greatly higher than R-square of the linear component indicated a sign of non-linearity.

However, considering the test for equality of several independent correlation coefficients reported in Table H.14, it showed that there was no significant difference among correlations of the linear/non-linear components for each independent variable of attribute AddService. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute AddService and the OCS student is linear.

In summary, the test of the addition of curve components of all variables in each group suggested that the linearity assumption was met.

Table H.13: R-square of the components of regression equation (attribute AddService)

Independent variable	R-Square		
	Linear	**Quadratic**	**Cubic**
AddService 1	0.236	0.250	0.260
AddService 2	0.334	0.345	0.347
AddService 3	0.257	0.264	0.282
AddService 4	0.378	0.394	0.410

Table H.14: Chi-square statistic for the difference test of equality of correlations (attribute AddService)

Independent variable	χ^2	df	p-value	Equality
AddService 1	0.079	2	0.961	Yes
AddService 2	0.025	2	0.988	Yes
AddService 3	0.088	2	0.957	Yes
AddService 4	0.135	2	0.935	Yes

H.3 Homoscedasticity of residuals

The assumption is that the residuals at each level of the dependent variables had the same variance. There are graphical and non-graphical methods for checking homoscedasticity. For the graphical method, the scatterplot of the standardized residual and standardized predicted value for each attribute (Figure H.7- H.12) were reobserved. Although the dots were not constantly spread over the horizontal line, a funnel shaped pattern did not appear. It may be inferred that this assumption was met.

For a non-graphical method, the Breusch-Pagan test for all group of attributes was conducted. The null hypothesis of Breusch-Pagan test is homoscedasticity. If p-value of the test less than 0.05, it is suggested that the null hypothesis is rejected.

Table H.15: Breusch-Pagan test for the homoscedasticity of all groups

Group of Variables	BP	df	p-value	Homoscedasticity
Teacher	2.328	1	0.127	Yes
Teaching	1.285	1	0.257	Yes
Admin	0.745	1	0.388	Yes
CompFac	0.080	1	0.777	Yes
xActivity	0.974	1	0.324	Yes
AddService	0.333	1	0.564	Yes

Consistent with the graphical method, a non-graphical method using Breusch-Pagan test for all group presented in Table H.15 is suggested that the variance of the residuals is homogenous. Thereby the assumption was met.

Note that steps for conducting Breusch-Pagan test using SPSS is referred to Pryce (2002) and the validity of this test is depended on the normality of residual. The z-score of skewness and kurtosis of residuals shown in Table H.2 suggested that the residuals of all groups were fairly normal; therefore, the test results were reliable.

H.4 Independence of residual

The assumption is that the error is not correlated with any independent variables. The independence of residuals was checked through the Durbin-Watson statistic of the regression model. A conventionally acceptable range of this statistic is 1 - 3 in which the value close to 2 indicates no autocorrelation (Field, 2009).

Table H.16: Durbin-Watson statistic of all attributes

Group of Variables	**Durbin-Watson statistic**	**Independence of residuals**
Teacher	1.977	Yes
Teaching	1.920	Yes
Admin	2.089	Yes
CompFac	1.962	Yes
xActivity	1.904	Yes
AddService	1.987	Yes

Table H.16 showed that Durbin-Watson statistic of all groups was within acceptable range indicated the lack of autocorrelation thereby the assumption was met.

H.5 Multicollinearity

Under the assumption of no multicollinearity, the independent variables are not closely linearly related to each other. The degree of multicollinearity can be detected by inspecting the VIF statistic provided by SPSS. A VIF of 1 means that there is no correlation among the independent variables while VIF higher than 1 indicates a presence of multicollinearity among them. As a rule of thumb, VIF higher than 10 indicates the sign of serious multicollinearity (Field, 2009). The VIF for each independent variable of all groups was greater than 1 but lower than 10 (see Table H.17- H.22) hence it's can conclude that multicollinearity was present but there is no sign of serious multicollinearity which requires correction (such as removing redundant variables) within the KPS dataset.

Table H.17: Collinearity Statistics (VIF) of five independent variables of attribute Teacher

Independent variable	Collinearity Statistics (VIF)
Teacher 1	1.890
Teacher 2	1.930
Teacher 3	1.578
Teacher 4	1.929
Teacher 5	2.073

Table H.18: Collinearity Statistics (VIF) of six independent variables of attribute Teaching

Independent variable	Collinearity Statistics (VIF)
Teaching 1	1.795
Teaching 2	1.826
Teaching 3	2.303
Teaching 4	2.011
Teaching 5	1.774
Teaching 6	2.305

Table H.19: Collinearity Statistics (VIF) of five independent variables of attribute Admin

Independent variable	Collinearity Statistics (VIF)
Admin 1	2.821
Admin 2	3.194
Admin 3	3.000
Admin 4	2.306
Admin 5	2.918

Table H.20: Collinearity Statistics (VIF) of four independent variables of attribute CompFac

Independent variable	Collinearity Statistics (VIF)
CompFac 1	3.074
CompFac 2	3.238
CompFac 3	1.935
CompFac 4	2.241

Table H.21: Collinearity Statistics (VIF) of seven independent variables of attribute xActivity

Independent variable	Collinearity Statistics (VIF)
xActivity 1	2.963
xActivity 2	3.237
xActivity 3	5.446
xActivity 4	4.157
xActivity 5	3.895
xActivity 6	3.068
xActivity 7	2.505

Table H.22: Collinearity Statistics (VIF) of four independent variables of attribute AddService

Independent variable	**Collinearity Statistics (VIF)**
AddService 1	2.124
AddService 2	2.304
AddService 3	1.831
AddService 4	1.864

Appendix I

Assumption checking of the multiple linear regression on CPE-KU-BKN dataset

There were six attributes in CPE-KU-BKN dataset namely Academic Personal (Teacher), Teaching and Learning (Teaching), Administration (Admin), Computer Facilities (CompFac), Extra-Curricular Activities (xActivity) and Additional Services (AddService). To build regression model for each attribute, items related to the attribute were set as independent variables and the overall student satisfaction (OCSstudent) was set as the dependent variable.

The assumptions of multiple linear regression were checked for each attribute in the dataset to ensure that the result of regression was valid. Assumption testing result for each group of attribute are summarized in Table I.1 and explained in the following sections.

Table I.1: Summary of assumption checking of MLR on six attributes of CPE-KU-BKN dataset

Assumptions	**Attribute**					
	Teacher	Teaching	Admin	CompFac	xActivity	AddService
Normality of residual	✓	✓	×	✓	✓	✓
Linearity	✓	✓	✓	✓	✓	✓
Homoscedasticity of residual	✓	✓	✓	✓	✓	✓
Independence of residual	✓	✓	✓	✓	✓	✓
Multicollinearity	✓	✓	✓	✓	✓	✓

With regard to Table I.1, a symbol ✓ indicates that the assumption was met and a symbol × indicates that the assumption was violated. According to the table it shows that most assumptions were met across six attributes of the dataset except assumption of normality for attribute Admin. Further steps were not performed to correct the assumption of normality of residual because the regression is fairly robust to violate of normality assumption. Beside violation of this assumption does not affect the regression coefficients which were the main focus of this study, the violation is only affected the significance tests for regression coefficients (Keith, 2014).

I.1 Normally distributed residual

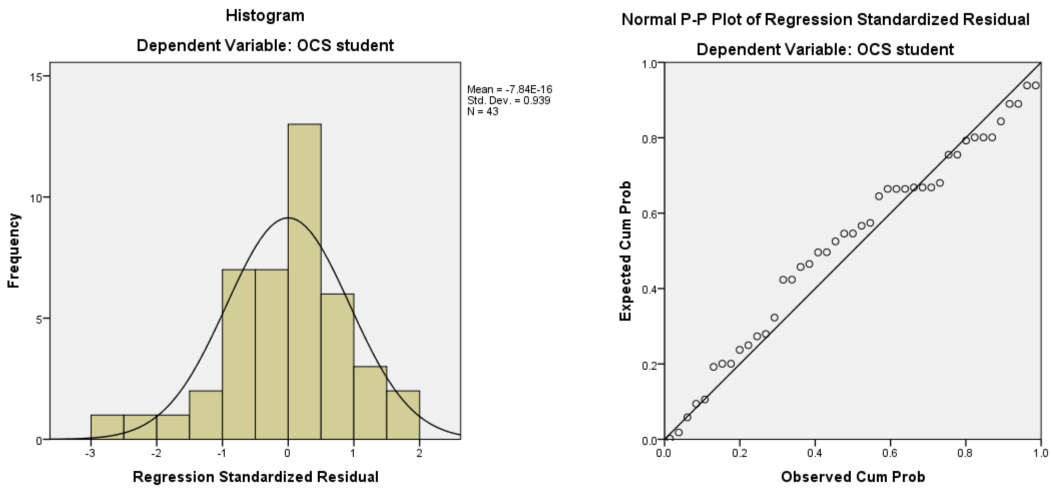
For each group of attribute, the assumption of normality was visually tested by observing the histogram and normal P-P plot of the standardized residual. This assumption was also tested by observing the skewness and kurtosis of unstandardized residual in which skew and kurtosis values between -1.0 and $+1.0$ are reasonable ranges to accept that data is reasonably close to normal (George and Mallery, 2003).

Since the dataset was rather small, the significance tests of skewness and kurtosis were also conducted by comparing the z-score of skewness and kurtosis against the value of z-score at the significant level. Regarding to Tabachnick et al. (2001), the conventional significant level of 0.001 was used to evaluate the significance of skewness and kurtosis with small to moderate samples and the z-score associated to this conventional significant level is 3.29. Thus, the absolute value of z-score that greater than 3.29 indicated that value of skewness and kurtosis is significantly different from zero which mean distribution is deviated from normal.

Histograms and normal P-P plots of the standardized residual for each group of attribute are shown in Figure I.1 - I.6. And descriptive statistic of unstandardized residual for each group is shown in Table I.2. The histogram of group Teacher (Figure I.1), Admin (Figure I.3), CompFac (Figure I.4) and xActivity (Figure I.5) appeared to be asymmetrical and their corresponding P-P plots were slightly deviated from the diagonal line which indicated that the distribution of residuals were not normal. In contrast, the histogram of the other two groups named Teaching (Figure I.2) and AddService (Figure I. 6) appeared to look like a bell and their normal P-P plots were approximately a straight line which suggested that the residuals of these variables were not deviated from normal.

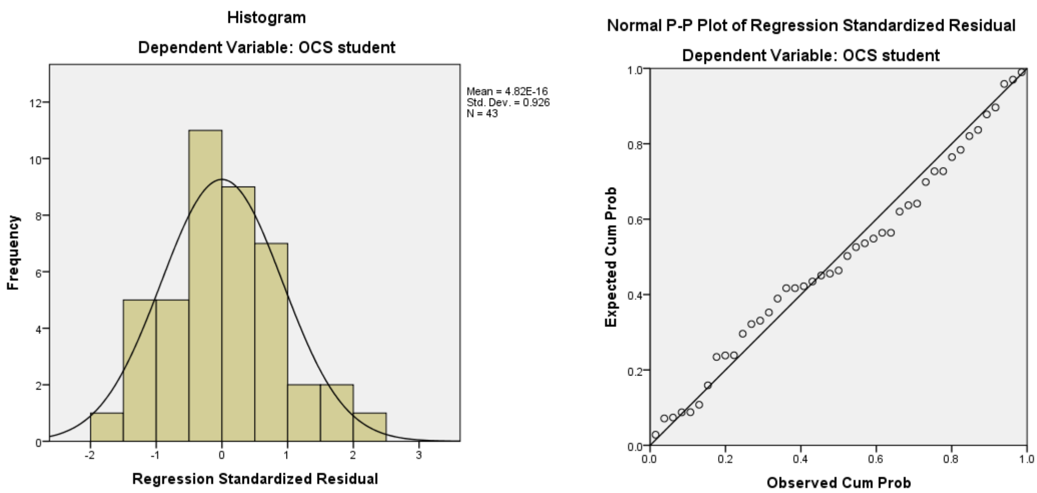
Table I.2: Descriptive statistic of unstandardized residual for each attribute of CPE-KU-BKN dataset

Unstandardized Residual	**N**	**Mean**	**Std. Deviation**	**Skewness**			**Kurtosis**		
				Statistic	Std. Error	z-score	Statistic	Std. Error	z-score
Teacher	43	0.000	0.557	-0.901	0.361	2.496	1.262	0.709	1.780
Teaching	43	0.000	0.456	0.274	0.361	0.759	0.181	0.709	0.255
Admin	43	0.000	0.619	-0.971	0.361	2.690	2.879	0.709	4.061
CompFac	43	0.000	0.629	-0.861	0.361	2.385	2.104	0.709	2.968
xActivity	43	0.000	0.396	0.483	0.361	1.338	-0.475	0.709	0.670
AddService	43	0.000	0.513	0.031	0.361	0.086	0.348	0.709	0.491



(a) Histogram plot of residual (b) Normality plot of residual

Figure I.1: Histogram and normality plots of residual (Teacher)

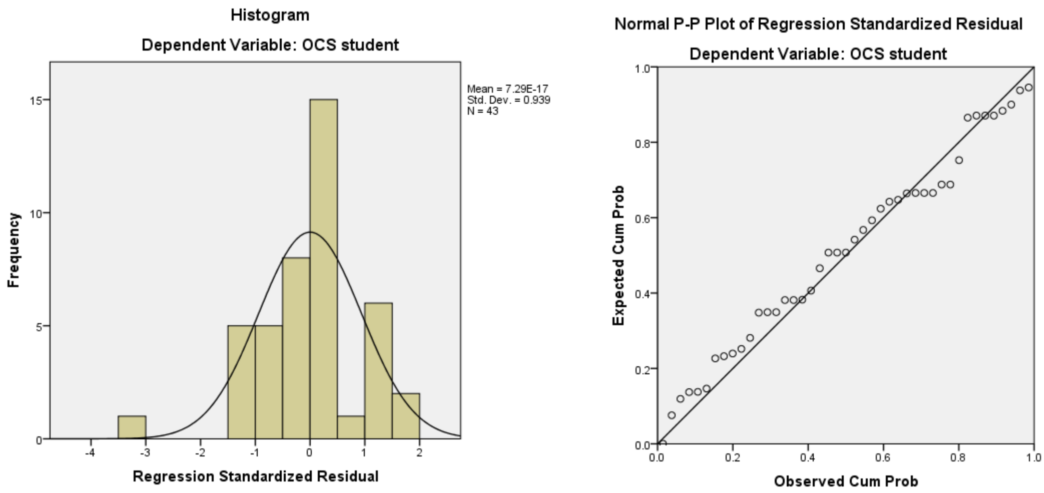


(a) Histogram plot of residual (b) Normality plot of residual

Figure I.2: Histogram and normality plots of residual (Teaching)

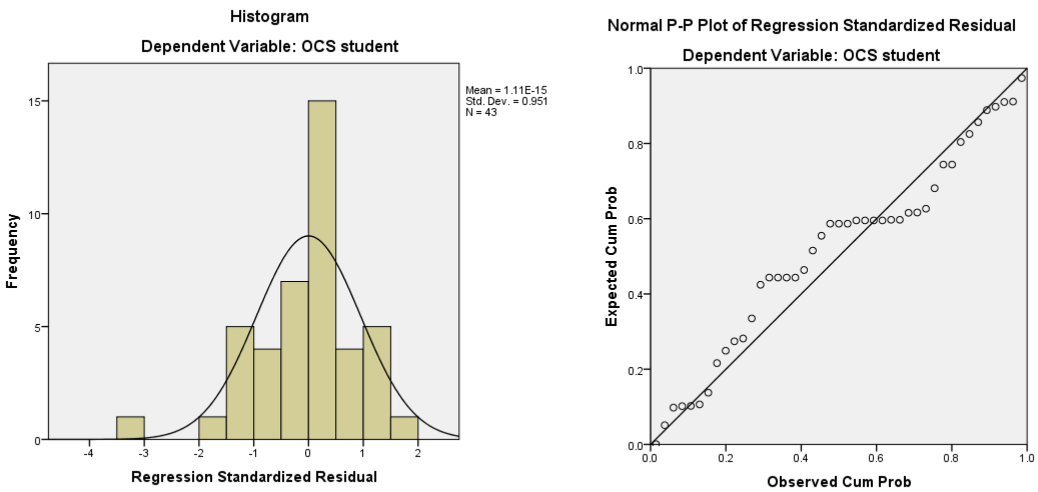
Regarding to Table I.2, skew and kurtosis of Teaching, xActivity and AddService were within the range to accept that their residuals were normally distributed. On the other hand, the kurtosis of Teacher, Admin and CompFac were greater than 1.0 indicated that the residuals were deviated from normal which corresponded to the histogram and P-P plots. Although the kurtosis of Teacher and CompFac were higher than 1.0, their z-scores were less than 3.29. Thus, it was reasonable to conclude that the residuals of these two groups were normally distributed.

Table I.2 also showed that kurtosis z-score of Admin was higher than 3.29 indicated that the residual of this group was not normally distributed. Even though the assumption of normality of residual was not met for Admin group, further steps were not performed



(a) Histogram plot of residual (b) Normality plot of residual

Figure I.3: Histogram and normality plots of residual (Admin)



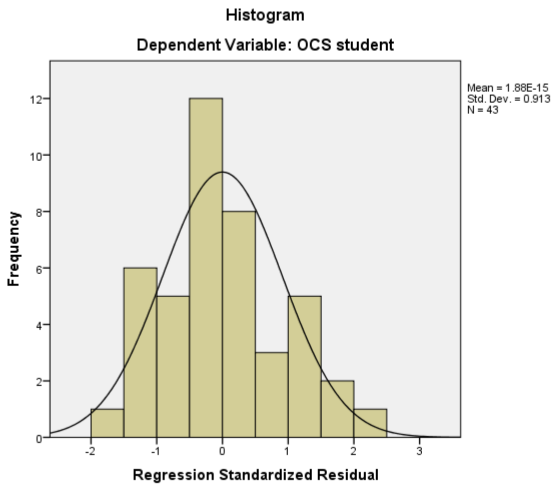
(a) Histogram plot of residual (b) Normality plot of residual

Figure I.4: Histogram and normality plots of residual (CompFac)

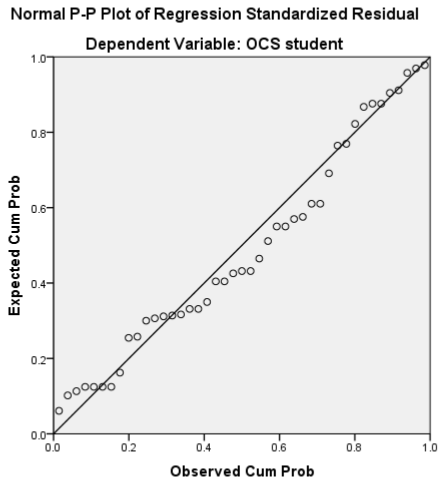
to correct this assumption because the regression is fairly robust to violate of normality assumption. Beside violation of this assumption doe not affect the regression coefficients which were the main focus of this study, the violation is only affected the significance tests for regression coefficients (Keith, 2014).

I.2 Linear relationship between independent and dependent variables

To check linearity assumption, a scatterplot of the standardized residual and standardized predicted value was drawn for each group of variable (see Figure I.7- I.12). The

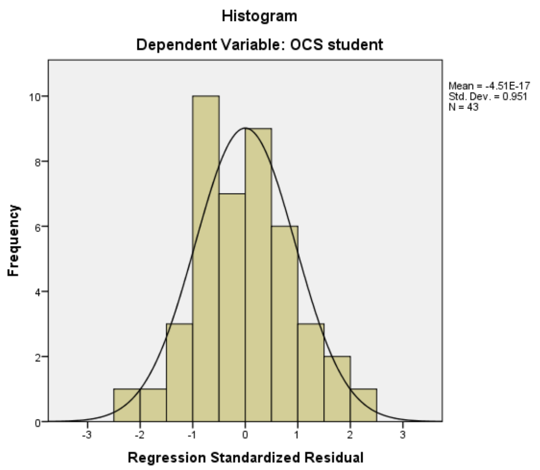


(a) Histogram plot of residual

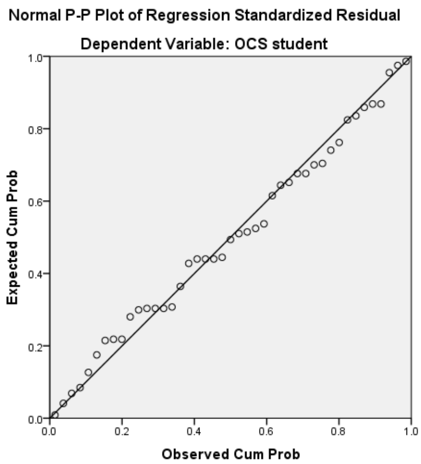


(b) Normality plot of residual

Figure I.5: Histogram and normality plots of residual (xActivity)



(a) Histogram plot of residual



(b) Normality plot of residual

Figure I.6: Histogram and normality plots of residual (AddService)

residual plot for predicted values of an outcome variable (OCS student) against the residuals of all attributes showed that the dots were not constantly spread over the horizontal line; however, there was no sign of any curve pattern. Therefore, it cannot conclude that the linearity assumption was violated.

For each group of attribute, non-linear relationships can also be detected by adding curve components which are quadratic (variable^2) and cubic (variable^3) to the regression equation to see whether this increases the explained variance (R-square). Specifically, for each independent variable of each group, the R-square of the linear component and the two curve components was statistically compared through the test of equality for several independent correlation coefficients (r) following an approach described in Kenny (1987). If there is no departure from linearity in the relationship, it is expected to see

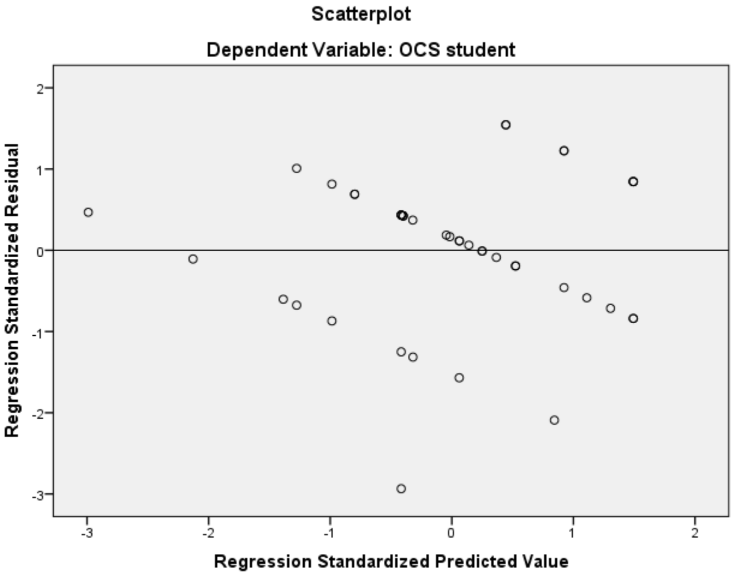


Figure I.7: Scatterplot of residual for the relationship between five independent variables of attribute Teacher and the OCS student

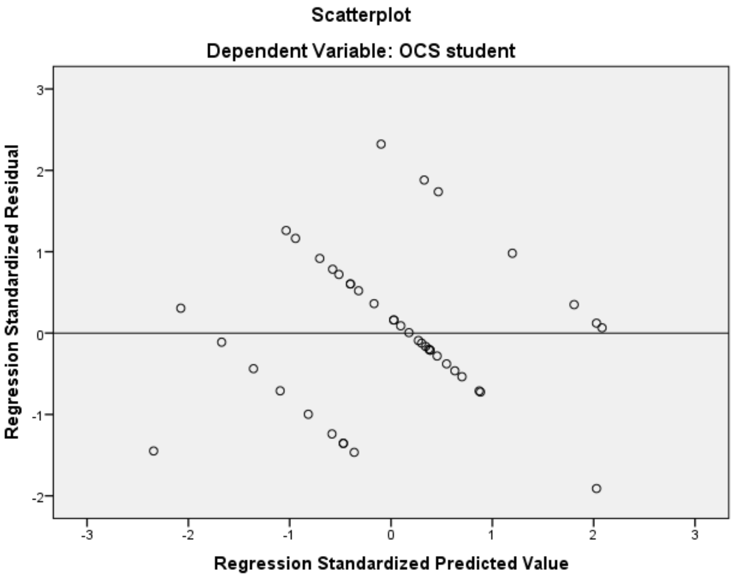


Figure I.8: Scatterplot of residual for the relationship between six independent variables of attribute Teaching and the OCS student

that the difference between the linear correlation coefficient (r) and the non-linear (curve) correlation coefficient is not significantly different (p-value > 0.05).

Table I.3 shows the R-square of linear/non-linear components of each independent variable of attribute Teacher against the OCS student. Regarding this table, R-squares of the curve component of all variables of attribute Teacher were equal to or slightly higher than R-square of the linear component. Only the variable named Teacher 4 which its

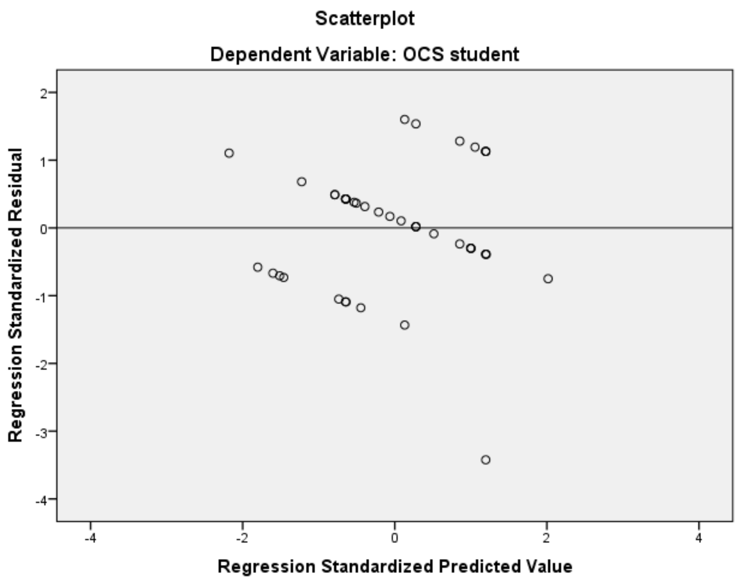


Figure I.9: Scatterplot of residual for the relationship between five independent variables of attribute Admin and the OCS student

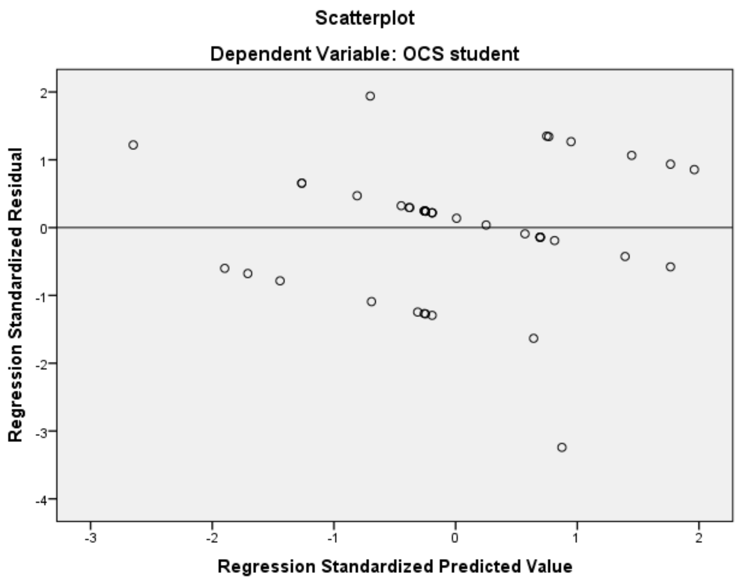


Figure I.10: Scatterplot of residual for the relationship between four independent variables of attribute CompFac and the OCS student

R-squares of the curve component were greatly higher than its R-squares of the linear component.

Given R-square from Table I.3, the test for equality of several independent correlation coefficients was conducted and the result was reported in Table I.4. The Table I.4 showed that there was no significant difference among correlations of the linear/non-linear components for each independent variable of attribute Teacher. Therefore, it

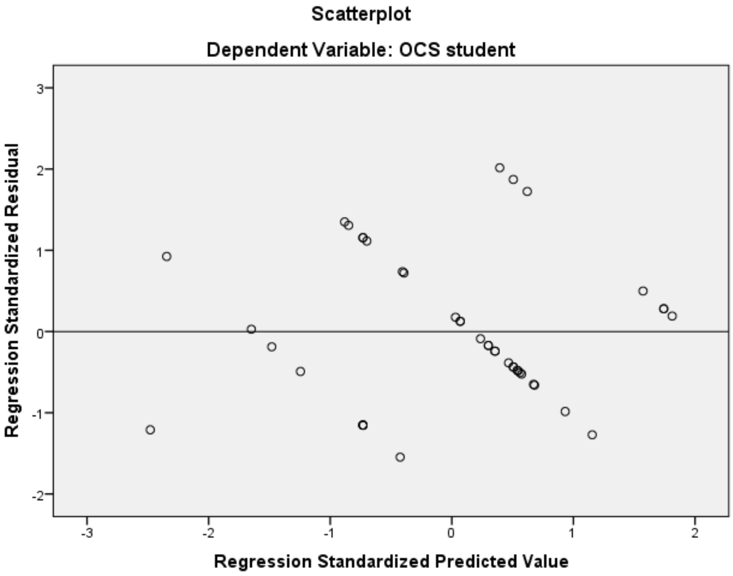


Figure I.11: Scatterplot of residual for the relationship between seven independent variables of attribute xActivity and the OCS student

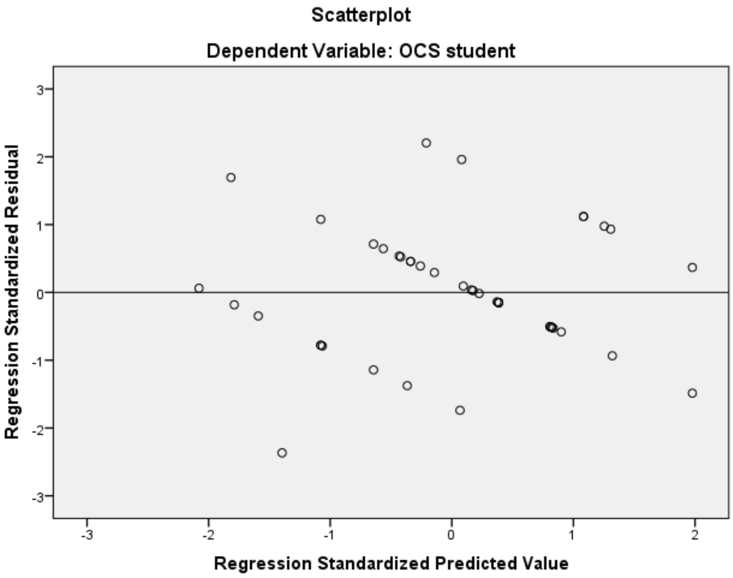


Figure I.12: Scatterplot of residual for the relationship between four independent variables of attribute AddService and the OCS student

was reasonable to conclude that the relationship between each independent variable of attribute Teacher and the OCS student is linear.

Table I.5 shows the R-square of linear/non-linear components of each independent variable of attribute Teaching against the OCS student. Regarding this table, R-squares of the curve component of two variables named Teaching 2 and Teaching 6 were equal to or slightly higher than R-square of the linear component. R-squares of the curve

Table I.3: R-square of the components of regression equation (attribute Teacher)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
Teacher 1	0.206	0.206	0.206
Teacher 2	0.205	0.209	0.209
Teacher 3	0.214	0.215	0.215
Teacher 4	0.163	0.174	0.179
Teacher 5	0.147	0.151	0.151

Table I.4: Chi-square statistic for the difference test of equality of correlations (attribute Teacher)

Independent variable	χ^2	**df**	**p-value**	**Equality**
Teacher 1	0.000	2	1.000	Yes
Teacher 2	0.001	2	1.000	Yes
Teacher 3	0.000	2	1.000	Yes
Teacher 4	0.011	2	0.994	Yes
Teacher 5	0.001	2	1.000	Yes

component of the other four variables named Teaching1, and Teaching 3-5 were greatly higher than R-squares of the linear component indicated a sign of non-linearity.

However, considering the test for equality of several independent correlation coefficients reported in Table I.6, it showed that there was no significant difference among correlations of the linear/non-linear components for each independent variable of attribute Teaching. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute Teaching and the OCS student is linear.

Table I.5: R-square of the components of regression equation (attribute Teaching)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
Teaching 1	0.461	0.480	0.483
Teaching 2	0.024	0.028	0.028
Teaching 3	0.162	0.190	0.196
Teaching 4	0.023	0.032	0.032
Teaching 5	0.253	0.264	0.286
Teaching 6	0.116	0.119	0.119

Table I.7 shows the R-square of linear/non-linear components of each independent variable of attribute Admin against the OCS student. Regarding this table, R-squares of the curve component of most variables were equal to or slightly higher than R-square of the linear component. Only the variable named Admin 5 which its R-squares of the curve component were greatly higher than its R-squares of the linear component.

Given R-square from Table I.7, the test for equality of several independent correlation coefficients was conducted and the result was reported in Table I.8. The Table I.8 showed

Table I.6: Chi-square statistic for the difference test of equality of correlations (attribute Teaching)

Independent variable	χ^2	**df**	**p-value**	**Equality**
Teaching 1	0.022	2	0.989	Yes
Teaching 2	0.004	2	0.998	Yes
Teaching 3	0.055	2	0.973	Yes
Teaching 4	0.021	2	0.990	Yes
Teaching 5	0.039	2	0.981	Yes
Teaching 6	0.001	2	1.000	Yes

that there was no significant difference among correlations of the linear/non-linear components for each independent variable of attribute Admin. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute Admin and the OCS student is linear.

Table I.7: R-square of the components of regression equation (attribute Admin)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
Admin 1	0.045	0.046	0.046
Admin 2	0.060	0.064	0.064
Admin 3	0.003	0.003	0.003
Admin 4	0.003	0.003	0.004
Admin 5	0.068	0.105	0.105

Table I.8: Chi-square statistic for the difference test of equality of correlations (attribute Admin)

Independent variable	χ^2	**df**	**p-value**	**Equality**
Admin 1	0.000	2	1.000	Yes
Admin 2	0.002	2	0.999	Yes
Admin 3	0.000	2	1.000	Yes
Admin 4	0.002	2	0.999	Yes
Admin 5	0.128	2	0.938	Yes

Table I.9 shows the R-square of linear/non-linear components of each independent variable of attribute CompFac against the OCS student. Regarding this table, R-squares of the curve component of most variables were equal to or slightly higher than R-square of the linear component. Only the variable named CompFac 1 which its R-squares of the curve component were critically higher than its R-squares of the linear component indicated a sign of non-linearity.

However, considering the test for equality of several independent correlation coefficients reported in Table I.10, it showed that there was no significant difference among correlations of the linear/non-linear components for each independent variable of attribute CompFac. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute CompFac and the OCS student is linear.

Table I.9: R-square of the components of regression equation (attribute CompFac)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
CompFac 1	0.005	0.163	0.188
CompFac 2	0.080	0.080	0.094
CompFac 3	0.057	0.079	0.079
CompFac 4	0.025	0.037	0.038

Table I.10: Chi-square statistic for the difference test of equality of correlations (attribute CompFac)

Independent variable	χ^2	**df**	**p-value**	**Equality**
CompFac 1	3.784	2	0.151	Yes
CompFac 2	0.018	2	0.991	Yes
CompFac 3	0.055	2	0.973	Yes
CompFac 4	0.036	2	0.982	Yes

Table I.11 shows the R-square of linear/non-linear components of each independent variable of attribute xActivity against the OCS student. Regarding this table, R-squares of the curve component of most variables were equal to or slightly higher than R-square of the linear component. Only the variables named xActivity 1 and xActivity 4 which their R-squares of the curve component were critically higher than their R-squares of the linear component indicated a sign of non-linearity.

However, considering the test for equality of several independent correlation coefficients reported in Table I.12, it showed that there was no significant difference among correlations of the linear/non-linear component for each independent variable of attribute xActivity. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute xActivity and the OCS student is linear.

Table I.11: R-square of the components of regression equation (attribute xActivity)

Independent variable	**R-Square**		
	Linear	**Quadratic**	**Cubic**
xActivity 1	0.105	0.318	0.477
xActivity 2	0.438	0.451	0.451
xActivity 3	0.155	0.170	0.179
xActivity 4	0.018	0.225	0.225
xActivity 5	0.376	0.381	0.381
xActivity 6	0.293	0.319	0.327
xActivity 7	0.310	0.329	0.350

Table I.13 shows the R-square of linear/non-linear component of each independent variable of attribute AddService against the OCS student. Regarding this table, R-squares

Table I.12: Chi-square statistic for the difference test of equality of correlations (attribute xActivity)

Independent variable	χ^2	df	p-value	Equality
xActivity 1	5.321	2	0.070	Yes
xActivity 2	0.008	2	0.996	Yes
xActivity 3	0.025	2	0.987	Yes
xActivity 4	3.864	2	0.145	Yes
xActivity 5	0.001	2	0.999	Yes
xActivity 6	0.043	2	0.979	Yes
xActivity 7	0.054	2	0.973	Yes

of the curve component of all variables were greatly higher than R-square of the linear component indicated a sign of non-linearity.

However, considering the test for equality of several independent correlation coefficients reported in Table I.14, it showed that there was no significant difference among correlations of the linear/non-linear components for each independent variable of attribute AddService. Therefore, it was reasonable to conclude that the relationship between each independent variable of attribute AddService and the OCS student is linear.

Table I.13: R-square of the components of regression equation (attribute AddService)

Independent variable	R-Square		
	Linear	Quadratic	Cubic
AddService 1	0.091	0.145	0.146
AddService 2	0.171	0.174	0.224
AddService 3	0.368	0.368	0.382
AddService 4	0.150	0.162	0.172

Table I.14: Chi-square statistic for the difference test of equality of correlations (attribute AddService)

Independent variable	χ^2	df	p-value	Equality
AddService 1	0.218	2	0.897	Yes
AddService 2	0.140	2	0.933	Yes
AddService 3	0.009	2	0.996	Yes
AddService 4	0.021	2	0.989	Yes

In summary, the test of the addition of curve components of all variables in each group suggested that the linearity assumption was met.

I.3 Homoscedasticity of residuals

The assumption is that the residuals at each level of the dependent variables had the same variance. There are graphical and non-graphical methods for checking homoscedasticity. For the graphical method, the scatterplot of the standardized residual and standardized predicted value for each attribute (Figure I.7- I.12) were reobserved. Although the dots were not constantly spread over the horizontal line, a funnel shaped pattern did not appear. It may be inferred that this assumption was met.

For a non-graphical method, the Breusch-Pagan test for all group of attributes was conducted. The null hypothesis of Breusch-Pagan test is homoscedasticity. If p-value of the test less than 0.05, it is suggested that the null hypothesis is rejected.

Table I.15: Breusch-Pagan test for the homoscedasticity of all groups

Group of Variables	BP	df	p-value	Homoscedasticity
Teacher	0.371	1	0.542	Yes
Teaching	0.100	1	0.751	Yes
Admin	3.454	1	0.063	Yes
CompFac	0.936	1	0.333	Yes
xActivity	0.883	1	0.347	Yes
AddService	0.812	1	0.368	Yes

Consistent with the graphical method, a non-graphical method using Breusch-Pagan test for all group which is presented in Table I.15 suggested that the variance of the residuals is homogenous. Thereby the assumption was met.

Note that steps for conducting Breusch-Pagan test using SPSS is referred to Pryce (2002) and the validity of this test is depended on the normality of residual. The z-score of skewness and kurtosis of residuals shown in Table I.2 suggested that the residuals of all groups were fairly normal; therefore, the test results were reliable.

I.4 Independence of residual

The assumption is that the error is not correlated with any independent variables. The independence of residuals was checked through the Durbin-Watson statistic of the regression model. A conventionally acceptable range of this statistic is 1 - 3 in which the value close to 2 indicates no autocorrelation (Field, 2009).

Table I.16 showed that Durbin-Watson statistic of all groups was within acceptable range indicated the lack of autocorrelation thereby the assumption was met.

Table I.16: Durbin-Watson statistic of all attributes

Group of Variables	Durbin-Watson statistic	Independence of residuals
Teacher	2.149	Yes
Teaching	1.313	Yes
Admin	2.083	Yes
CompFac	1.957	Yes
xActivity	2.196	Yes
AddService	2.051	Yes

I.5 Multicollinearity

Under the assumption of no multicollinearity, the independent variables are not closely linearly related to each other. The degree of multicollinearity can be detected by inspecting the VIF statistic provided by SPSS. A VIF of 1 means that there is no correlation among the independent variables while VIF higher than 1 indicates a presence of multicollinearity among them. As a rule of thumb, VIF higher than 10 indicates the sign of serious multicollinearity (Field, 2009). The VIF for each independent variable of all groups was greater than 1 but lower than 10 (see Table I.17- I.22) hence it's can conclude that multicollinearity was present but there is no sign of serious multicollinearity which requires correction (such as removing redundant variables) within the BKN dataset.

Table I.17: Collinearity Statistics (VIF) of five independent variables of attribute Teacher

Independent variable	Collinearity Statistics (VIF)
Teacher 1	2.276
Teacher 2	1.404
Teacher 3	1.998
Teacher 4	2.880
Teacher 5	3.109

Table I.18: Collinearity Statistics (VIF) of six independent variables of attribute Teaching

Independent variable	Collinearity Statistics (VIF)
Teaching 1	2.351
Teaching 2	1.257
Teaching 3	2.210
Teaching 4	1.334
Teaching 5	1.308
Teaching 6	1.306

Table I.19: Collinearity Statistics (VIF) of five independent variables of attribute Admin

Independent variable	Collinearity Statistics (VIF)
Admin 1	1.884
Admin 2	2.409
Admin 3	2.792
Admin 4	3.180
Admin 5	2.554

Table I.20: Collinearity Statistics (VIF) of four independent variables of attribute CompFac

Independent variable	Collinearity Statistics (VIF)
CompFac 1	2.526
CompFac 2	1.963
CompFac 3	1.252
CompFac 4	1.604

Table I.21: Collinearity Statistics (VIF) of seven independent variables of attribute xActivity

Independent variable	Collinearity Statistics (VIF)
xActivity 1	2.592
xActivity 2	2.239
xActivity 3	4.033
xActivity 4	3.551
xActivity 5	2.683
xActivity 6	3.070
xActivity 7	3.193

Table I.22: Collinearity Statistics (VIF) of four independent variables of attribute AddService

Independent variable	Collinearity Statistics (VIF)
AddService 1	1.866
AddService 2	2.238
AddService 3	1.194
AddService 4	1.912

Appendix J

Staff evaluation of SWOT survey material

J.1 Survey on the level of agreement toward SWOT of department

This survey forms part of a study into the development of SWOT based on customer satisfaction survey, in particular the case study consider Department of Computer Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University (CPE-KU-KPS). Definition of each aspect of SWOT is provided as follow:
Strengths refer to the internal operation that the department doing well
Weakness refer to the internal operation that the department poorly perform
Opportunities refer to external factors that the department can use its advantage or operation that department perform better than other departments in Kasetsart University.
Threats refer to external factors that may cause a problem to the department or operation that department poorly perform than other departments in Kasetsart University.

Section I. Strength-Weakness.
This section asks about your level of agreement towards the strengths and weaknesses of the CPE-KU-KPS that have been identified based on an earlier CPE-KU-KPS's student satisfaction survey.

Instruction: Please assess your level of agreement with the following sentences according to your view and experience by ticking the appropriate response using the following scales:

Score	Interpretation on Level of Agreement
4	Completely agree
3	Agree
2	Disagree
1	Completely disagree

Section II. Opportunity-Threat.

This section asks about your level of agreement towards the opportunities and threats in the CPE-KU-KPS that identified by comparing the CPE-KU-KPS's student satisfaction with the Department of Computer Engineering, Kasetsart University, Bangkhen campus (CPE-KU-BKN)'s student satisfaction.

Instruction: Please assess your level of agreement with the following sentences according to your view and experience by ticking the appropriate response.

Section III. Personal Information

Instruction: Please tick the circle that best corresponds to your answer for each question below

1. Gender ◯ Male ◯ Female
2. Number of working years at CPE-KU-KPS ◯ <1 ◯ 1-3 ◯ 4-6 ◯ 7-10 ◯ >10
3. Position ◯ Academic staff ◯ Non-Academic staff

Thank you very much for your help

Sentences about strengths and weaknesses	Level of Agreement			
	4	3	2	1
1. Teaching ability of teaching staff is a strength of the department				
2. Subject expertise of teaching staff is a strength of the department				
3. Friendliness of the teaching staff towards students is a strength of the department				
4. Ability of teaching staff to give advice and support to student learning is a strength of the department				
5. Appropriate number of students per class is a strength of the department				
6. Knowledge of rules and procedures of non-academic staff members is a strength of the department				
7. Knowledge of the information about courses, exams and activities of non-academic staff members is a strength of the department				
8. Poor quality of computer facilities for students (Hardware and Software) is a weakness of the department				
9. Lack of availability of computer facilities for students is a weakness of the department				
10. Lack of availability of printing and photocopying facilities is a weakness of the department				
11. Poor arrangement of cultural exchange programs with foreign countries is a weakness of the department				
12. Poor arrangement of health development activities is a weakness of the department				
13. Lack of financial aid provided for students is a weakness of the department				
14. Lack of medical support provided for students is a weakness of the department				
15. Lack of availability of library facilities is a weakness of the department				

Sentences about opportunities and threats	Level of Agreement			
	4	3	2	1
1. Increasing the availability of teaching staff to the students is an opportunity for the department				
2. Increasing the availability of lecture materials is an opportunity for the department				
3. Increasing E-learning resources to support student learning are an opportunity for the department				
4. Increasing the ability to give clear and timely feedback to students is an opportunity for the department				
5. Increasing accurate and up-to-date course unit content is an opportunity for the department				
6. Inadequate of teaching facilities and learning areas is a threat to the department				
7. Lack of interest in solving the students' problem by non-academic staff members is a threat to the department				
8. Lack of friendliness of non-academic staff members to students is a threat to the department				
9. Lack of ability of non-academic staff members to provide services in a timely manner is a threat to the department				
10. Lack of availability of internet access for students is a threat to the department				
11. Updated content and easy to find information on the department website is an opportunity for the department				
12. Improving the arrangement of field trips activities is an opportunity for the department				
13. Improving the arrangement of moral development activities is an opportunity for the department				
14. Improving the arrangement of interpersonal skills development activities is an opportunity for the department				
15. Improving the arrangement of personal learning and thinking skills development activities is an opportunity for the department				
16. Improving the arrangement of social volunteer activities is an opportunity for the department				

J.2 Participant Information sheet

Study Title: Mining Survey Data for SWOT analysis (validation phase I)
Researcher: Boonyarat Phadermrod
Ethics number: ERGO/FPSE/18153

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

I am completing a PhD research about mining survey data for SWOT analysis which aimed to generate prioritized SWOT factors based on the customer's perception. To evaluate the proficiency of the proposed approach in the real-world situation, the case study of Department of Computer Engineering (DoCE), Kasetsart University will be conducted. Through this case study, your response on the level of agreement with the SWOT aspects of the DoCE will be analysed to evaluate SWOT generated based on satisfaction survey.

Why have I been chosen?

At this stage, this research is trying to get a level of agreement with the SWOT aspects of the DoCE which generated based on student satisfaction survey of DoCE. Thus, you are chosen as you are staff members of DoCE, Kasetsart University.

What will happen to me if I take part?

The questionnaire will take no longer than 15 minutes to complete. There are three sections in the questionnaire. Section 1 and 2 asks about your level of agreement with Strength-Weakness and Opportunity-Threat of the DoCE respectively. Section 3 asks about your general information.

Are there any benefits in my taking part?

By taking part, you have the opportunity to help us develop the new method to generate SWOT based on satisfaction survey which can be used in government, academic institution and private company.

Are there any risks involved?

There is no risk involved for participants completing the questionnaire.

Will my participation be confidential?

The name of the participants will not be taken and participation will be kept anonymous. All data will be safe in a protected computer. All will be destroyed once the research is completed.

What happens if I change my mind?

You have the right to withdraw from doing the questionnaire at any time by exiting the webpage.

What happens if something goes wrong?

If you have any concern or complaint with this research please contacts me (Boonyarat Phadermrod: bp6g12@soton.ac.uk). If you have a trouble with the online forms, please contact the ECS School Office on school@ecs.soton.ac.uk 02380 592909.

Where can I get more information?

If you would like more information on this research please feel free to contact me (Boonyarat Phadermrod: bp6g12@soton.ac.uk)

J.3 Result of language consistency translation test of staff evaluation of SWOT survey

Remark: +1 the question/ instruction is consistent with the English version

0 undecided about whether Thai question/ instruction is consistent with the English version

−1 the question/ instruction is not consistent with the English version

Table J.1: Summary of language consistency translation test of staff evaluation of SWOT survey

	Question/ Instruction (English)	**Question/ Instruction** (Thai)	**Language consistency score**			**Total Score**	**Comment**
			Expert 1	Expert 2	Expert 3		
1	This questionnaire forms part of a study into the development of SWOT based on customer satisfaction survey, in particular the case study consider Department of Computer Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart university (CPE-KU-KPS).	แบบประเมินนี้เป็นส่วนหนึ่งของงานวิจัยเพื่อการสร้างจุดแข็ง จุดอ่อน โอกาส และ อุปสรรค (SWOT) จากผลสำรวจความพึงพอใจของลูกค้าโดยกรณีศึกษาในงานวิจัยนี้ คือภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ กำแพงแสน มหาวิทยาลัยเกษตรศาสตร์ (วศ.คอม กพส)	+1	+1	+1	3	Check the proper way to use the word SWOT in Thai context.
2	Section I. Strength-Weakness.	ตอนที่ 1. ระดับความคิดเห็นต่อจุดแข็งและจุดอ่อนของภาควิชาวิศวกรรมคอมพิวเตอร์	+1	+1	+1	3	
3	This section asks about your level of agreement towards the strengths and weaknesses of the CPE-KU-KPS that were identified based on an earlier CPE-KU-KPS's student satisfaction survey.	แบบสอบถามส่วนนี้สำรวจถึงระดับความคิดเห็นของท่านต่อจุดแข็ง (Strength) และจุดอ่อน (Weakness) ของภาควิชาวิศวกรรมคอมพิวเตอร์ซึ่งจุดแข็งและจุดอ่อนของภาควิชาเหล่านี้สร้างขึ้นจากผลการวิเคราะห์แบบประเมินความพึงพอใจของนิสิตต่อภาควิชา	+1	+1	+1	3	
4	Instruction: Please assess your level of agreement with the following sentences according to your view and experience by ticking the appropriate response:	คำชี้แจง: กรุณาระบุระดับความคิดเห็นของท่านต่อข้อความด้านล่างตามความคิดเห็นและประสบการณ์จากการทำงานของท่าน ณ ภาควิชาวิศวกรรมคอมพิวเตอร์ โดยทำเครื่องหมายถูก (✓) ในช่องที่ตรงกับระดับความคิดเห็นของท่าน	+1	+1	+1	3	
5	Completely agree Agree Disagree Completely disagree	เห็นด้วยอย่างยิ่ง เห็นด้วย ไม่เห็นด้วย ไม่เห็นด้วยอย่างยิ่ง	+1	+1	+1	3	
6	Teaching ability of teaching staff is a strength of the department	ความสามารถในการสอนของอาจารย์ผู้สอนถือเป็นจุดแข็งของภาควิชา	+1	+1	+1	3	
7	Subject expertise of teaching staff is a strength of the department	ความเชี่ยวชาญในเนื้อหาวิชาของอาจารย์ผู้สอนถือเป็นจุดแข็งของภาควิชา	+1	+1	+1	3	

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score Expert 1	Expert 2	Expert 3	Total Score	Comment
8	Friendliness of the teaching staff towards students is a strength of the department	ความเป็นกันเองของอาจารย์ผู้สอนต่อนิสิตถือเป็นจุดแข็งของภาควิชา	+1	+1	+1	3	
9	Ability of teaching staff to give advice and support student learning is a strength of the department	การให้คำปรึกษาและความช่วยเหลือในการเรียนของอาจารย์ผู้สอนต่อนิสิต ถือเป็นจุดแข็งของภาควิชา	+1	+1	+1	3	
10	Appropriate number of students per class is a strength of the department	ความเหมาะสมของจำนวนนิสิตในชั้นเรียนถือเป็น จุดแข็งของภาควิชา	+1	+1	+1	3	
11	Knowledge of rules and procedures of non-academic staff members is a strength of the department	ความรู้เรื่องข้อบังคับ ระเบียบมหาวิทยาลัยของเจ้าหน้าที่ ถือเป็นจุดแข็งของภาควิชา	+1	+1	+1	3	Add “ความชำนาญ” into statement
12	Knowledge of the information about courses, exams and activities of non-academic staff members is a strength of the department	ความรู้เรื่องข้อมูลข่าวสารเกี่ยวกับ รายวิชา การสอบ และ กิจกรรมนิสิตของเจ้าหน้าที่ ถือเป็นจุดแข็งของภาควิชา	+1	+1	+1	3	
13	Poor quality of computer facilities for students (Hardware and Software) is a weakness of the department	อุปกรณ์คอมพิวเตอร์ (ฮาร์ดแวร์และซอฟต์แวร์) ที่ด้อยคุณภาพ ถือเป็นจุดอ่อนของภาควิชา	+1	+1	+1	3	
14	Lack of availability of computer facilities for students is a weakness of the department	การขาดความพร้อมให้บริการของอุปกรณ์คอมพิวเตอร์ ถือเป็นจุดอ่อนของภาควิชา	+1	+1	0	2	Rephrase to make the statement more clear
15	Lack of availability of printing and photocopying facilities is a weakness of the department	การขาดความพร้อมให้บริการของเครื่องพิมพ์และเครื่อง ถ่ายเอกสาร ถือเป็นจุดอ่อนของภาควิชา	+1	+1	0	2	Rephrase to make the statement more clear
16	Poor arrangement of cultural exchange programs with foreign countries is a weakness of the department	การจัดโครงการแลกเปลี่ยนวัฒนธรรมกับต่างประเทศที่ด้อยประสิทธิภาพ ถือเป็นจุดอ่อนของภาควิชา	+1	+1	+1	3	
17	Poor arrangement of health development activities is a weakness of the department	การจัดกิจกรรมพัฒนาสุขภาพที่ด้อยประสิทธิภาพ ถือเป็นจุดอ่อนของภาควิชา	+1	+1	+1	3	

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score			Total Score	Comment
			Expert 1	Expert 2	Expert 3		
18	Lack of financial aid provided for students is a weakness of the department	ความช่วยเหลือด้านการเงินแก่นิสิตที่ไม่เพียงพอ ถือเป็น จุดอ่อนของภาควิชา	+1	+1	0	2	Rephrase to make the statement more clear
19	Lack of medical support provided for students is a weakness of the department	การให้ความช่วยเหลือด้านรักษาพยาบาลแก่นิสิตที่ไม่เพียงพอ ถือเป็นจุดอ่อนของภาควิชา	+1	+1	0	2	Rephrase to make the statement more clear
20	Lack of availability of library facilities is a weakness of the department	การขาดความพร้อมให้บริการของห้องสมุดถือเป็น จุดอ่อนของภาควิชา	+1	+1	+1	3	
21	Section II. Opportunity-Threat.	ตอนที่ 2. ระดับความคิดเห็นต่อโอกาสและอุปสรรคของภาควิชาวิศวกรรมคอมพิวเตอร์	+1	+1	+1	3	
22	This section asks about your level of agreement towards the opportunities and threats in the CPE-KU-KPS that identified by comparing the CPE-KU-KPS's student satisfaction with the Department of Computer Engineering, Kasetsart University, Bangkhen campus (CPE-KU-BKN)'s student satisfaction.	แบบสอบถามส่วนนี้ สำรวจถึงระดับความคิดเห็นของท่าน ต่อโอกาส (Opportunity) และอุปสรรค (Threat) ของภาควิชาวิศวกรรมคอมพิวเตอร์ซึ่งโอกาสและอุปสรรค ของภาควิชาเหล่านี้สร้างขึ้นจากการเปรียบเทียบผลการ วิเคราะห์แบบประเมินความพึงพอใจของนิสิต วศ.คอม กพส. ต่อภาควิชา กับผลการวิเคราะห์แบบประเมินความพึงพอใจของนิสิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ (วศ.คอม บางเขน) ต่อภาควิชา	+1	+1	+1	3	
23	Instruction: Please assess your level of agreement with the following sentences according to your view and experience by ticking the appropriate response.	คำชี้แจง: กรุณาระบุระดับความคิดเห็นของท่านต่อข้อความด้านล่างตามความคิดเห็นและประสบการณ์จากการทำงานของท่าน ณ ภาควิชาวิศวกรรมคอมพิวเตอร์ โดยทำเครื่องหมายถูก (✓) ในช่องที่ ตรงกับระดับความคิดเห็นของท่าน	+1	+1	+1	3	
24	Increasing the availability of teaching staff to the students is an opportunity for the department	การเพิ่มโอกาสในการให้นิสิตเข้าพบนอกเวลาเรียนของอาจารย์ผู้สอน ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	
25	Increasing the availability of lecture materials is an opportunity for the department	การเพิ่มความพร้อมของเอกสารประกอบการเรียนการสอน ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	
26	Increasing E-learning resources to support student learning are an opportunity for the department	การเพิ่มสื่อการเรียนรู้อิเล็กทรอนิกส์ (e-Learning) ที่สนับสนุนการเรียนของนิสิตถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score			Total Score	Comment
			Expert 1	Expert 2	Expert 3		
27	Increasing the ability to give clear and timely feedback to students is an opportunity for the department	การวัดประเมินผลที่ชัดเจนและได้รับแจ้งภายในเวลาที่กำหนด ถือเป็นโอกาสของภาควิชา	+1	+1	0	2	Change the word "ได้รับแจ้ง" to "แจ้งให้นิสิตทราบ"
28	Increasing accurate and up-to-date unit content is an opportunity for the department	การเพิ่มความถูกต้องและความทันสมัยของเนื้อหาวิชาที่สอน ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	
29	Inadequate of teaching facilities and learning areas is a threat to the department	อุปกรณ์การเรียนการสอน และ สภาพห้องเรียนที่ไม่เหมาะสมต่อการเรียนการสอน ถือเป็นอุปสรรค ของภาควิชา	+1	+1	+1	3	Cut the word "การเรียนการสอน" to make statement more clear
30	Lack of interest in solving the students' problem by non-academic staff members is a threat to the department	การขาดความใส่ใจในการแก้ปัญหาให้แก่นิสิตของเจ้าหน้าที่ ถือเป็นอุปสรรคของภาควิชา	+1	+1	+1	3	
31	Lack of friendliness of non-academic staff members to students is a threat to the department	การขาดความเป็นกันเองของเจ้าหน้าที่ ถือเป็นอุปสรรคของภาควิชา	+1	+1	+1	3	Change the word "ความเป็นกันเอง" to "มนุษย์สัมพันธ์ที่ดี"
32	Lack of ability of non-academic staff members to provide services in a timely manner is a threat to the department	การขาดความสามารถในการให้บริการนิสิตในระยะเวลา ที่เหมาะสมของเจ้าหน้าที่ ถือเป็นอุปสรรคของภาควิชา	+1	+1	+1	3	
33	Lack of availability of internet access for students is a threat to the department	การขาดความพร้อมให้บริการเชื่อมต่ออินเตอร์เน็ต ถือเป็นอุปสรรคของภาควิชา	+1	+1	+1	3	
34	Updated content and easy to find information on the department website is an opportunity for the department	เว็บไซต์ประจำภาควิชาที่มีเนื้อหาทันสมัย และสามารถค้นหาข้อมูลได้ง่าย ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	
35	Improving the arrangement of field trips activities is an opportunity for the department	การพัฒนาการจัดกิจกรรมทัศนะศึกษาดูงานให้ดีขึ้น ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	Change the word "พัฒนา" to "เพิ่มประสิทธิภาพ" and cut the word "ให้ดีขึ้น"
36	Improving the arrangement of moral development activities is an opportunity for the department	การพัฒนาการจัดกิจกรรมพัฒนาคุณธรรมให้ดีขึ้น ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	Same as 35

	Question/ Instruction (English)	**Question/ Instruction** (Thai)	**Language consistency score**			**Total Score**	**Comment**
			Expert 1	Expert 2	Expert 3		
37	Improving the arrangement of interpersonal skills development activities is an opportunity for the department	การพัฒนาการจัดกิจกรรมพัฒนาความสัมพันธ์ระหว่างบุคคลให้ดีขึ้น ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	Same as 35
38	Improving the arrangement of personal learning and thinking skills development activities is an opportunity for the department	การพัฒนาการจัดกิจกรรมพัฒนาทักษะการคิดและการเรียนรู้ให้ดีขึ้น ถือเป็นโอกาสของภาควิชา	+1	+1	+1	3	Same as 35
39	Improving the arrangement of social volunteer activities is an opportunity for the department	การพัฒนาการจัดกิจกรรมเพื่อสังคมให้ดีขึ้น ถือเป็นโอกาส ของภาควิชา	+1	+1	+1	3	Same as 35
40	Section III. Personal Information	ตอนที่ 3. ข้อมูลพื้นฐานเกี่ยวกับผู้ตอบแบบสอบถาม	+1	+1	+1	3	
41	Instruction: Please tick the circle that corresponds to your answer for each question below	คำชี้แจง: กรุณาทำเครื่องหมายถูก(✓) ในวงกลมที่ตรงกับข้อมูลของท่านมากที่สุด	+1	+1	+1	3	
42	Gender ○ Male ○ Female	เพศ ○ ชาย ○ หญิง	+1	+1	+1	3	
43	Number of working years at CPE-KU-KPS	จำนวนปีที่ทำงาน ณ ภาควิชาวิศวกรรมคอมพิวเตอร์	+1	+1	+1	3	
44	Position ○ Academic staff ○ Non-Academic staff	ตำแหน่ง ○ อาจารย์ ○ เจ้าหน้าที่	+1	+1	+1	3	

Appendix K

Assumption checking of one sample t-test for staff evaluation on SWOT factor

There are 31 variables represented the level of agreement of CPE-KU-KPS staff towards produced SWOT factor which divided into 15 variables for strengths-weaknesses and 16 for opportunities-threats (denoted as SW_{1-15}, OT_{1-16} respectively). Before the one sample t-test was conducted, the assumptions of one sample t-test were tested on each variable and the results are shown in the following sections.

K.1 Result of the assumptions test of one sample t-test on staff agreement level of SWOT factor

The one-sample t-test required three assumptions to be met as follows (Field, 2009):

Assumption 1: Variables should be interval/ ratio level.

All 31 variables represented the level of agreement of CPE-KU-KPS staff that were measured on a scale of 1 to 4. This assumption there was met.

Assumption 2: The data are independent.

Responses to the level of agreement of CPE-KU-KPS staff are independent because they come from a different person. This assumption there was met.

Assumption 3: Sampling distribution of the dataset is relatively normal.

The normality for each variable was tested by using Shapiro-Wilk in SPSS since the size of the dataset is smaller than 2000. A review of the Shapiro-Wilk test for normality for variable SW_{1-15}, OT_{1-16} are shown in Table K.1 - K.2 respectively.

Table K.1: Shapiro -Wilk test for normality of variable SW_{1-15}

Variable	**Shapiro-Wilk**			**Normality**
	Statistic	**df**	**p-value**	
SW_1	0.735	14	0.001	No
SW_2	0.750	14	0.001	No
SW_3	0.769	14	0.002	No
SW_4	0.516	14	0.000	No
SW_5	0.893	14	0.088	Yes
SW_6	0.616	14	0.000	No
SW_7	0.816	14	0.008	No
SW_8	0.773	14	0.002	No
SW_9	0.758	14	0.002	No
SW_{10}	0.769	14	0.002	No
SW_{11}	0.862	14	0.032	No
SW_{12}	0.750	14	0.001	No
SW_{13}	0.735	14	0.001	No
SW_{14}	0.750	14	0.001	No
SW_{15}	0.769	14	0.002	No

Table K.2: Shapiro -Wilk test for normality of variable OT_{1-16}

Variable	**Shapiro-Wilk**			**Normality**
	Statistic	**df**	**p-value**	
OT_1	0.779	14	0.003	No
OT_2	0.735	14	0.001	No
OT_3	0.639	14	0.000	No
OT_4	0.801	14	0.005	No
OT_5	0.646	14	0.000	No
OT_6	0.810	14	0.007	No
OT_7	0.850	14	0.022	No
OT_8	0.806	14	0.006	No
OT_9	0.769	14	0.002	No
OT_{10}	0.801	14	0.005	No
OT_{11}	0.750	14	0.001	No
OT_{12}	0.616	14	0.000	No
OT_{13}	0.652	14	0.000	No
OT_{14}	0.516	14	0.000	No
OT_{15}	0.516	14	0.000	No
OT_{16}	0.551	14	0.000	No

A review of the Shapiro-Wilk test for normality presented in Table K.1 and Table K.2 suggests that the data is non-normally distributed since p-values of most variables are less than 0.05. Only data of variable named SW_5 is normally distributed.

Addition to the normality test, skew and kurtosis of each variable are also observed and shown in Table K.3 - K.4 for variable SW_{1-15}, OT_{1-16} respectively. Regarding to George and Mallery (2003), skew and kurtosis values between -1.0 to $+1.0$ are reasonable ranges to accept that data is reasonably close to normal.

Most of the values of skewness and kurtosis reported in Table K.3 were within the range for reasonably concluded that data is relatively normal. Except skewness and kurtosis of SW_1, SW_4, SW_{6-7} and SW_{13} that values were outside the range. These values of skewness and kurtosis reported in Table K.3, and normality test reported in Table K.1, suggested that normality was not reasonable for some variables. Additionally, the dataset is relatively small, the central limit theorem cannot be referred to conclude that data is normally distributed. Thus, these variables need to be transformed to reduce the non-normal distribution.

Table K.3: Descriptive statistic of variable SW_{1-15}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
SW_1	14	3.000	0.555	0.000	0.597	1.330	1.154
SW_2	14	3.214	0.579	0.028	0.597	0.209	1.154
SW_3	14	3.286	0.611	-0.192	0.597	-0.258	1.154
SW_4	14	3.214	0.426	1.566	0.597	0.501	1.154
SW_5	14	2.571	0.938	-0.240	0.597	-0.491	1.154
SW_6	14	2.357	0.497	0.670	0.597	-1.838	1.154
SW_7	14	2.857	0.770	0.264	0.597	-1.123	1.154
SW_8	14	2.643	0.745	0.731	0.597	-0.637	1.154
SW_9	14	2.571	0.646	0.692	0.597	-0.252	1.154
SW_{10}	14	2.286	0.611	-0.192	0.597	-0.258	1.154
SW_{11}	14	2.571	0.756	-0.280	0.597	0.294	1.154
SW_{12}	14	2.214	0.579	0.028	0.597	0.209	1.154
SW_{13}	14	2.000	0.555	0.000	0.597	1.330	1.154
SW_{14}	14	2.214	0.579	0.028	0.597	0.209	1.154
SW_{15}	14	2.286	0.611	-0.192	0.597	-0.258	1.154

Table K.4 also shown that data of variables OT_{2-3}, OT_5, OT_{12-16} were not normally distributed since their skewness or kurtosis were outside the range. This consistency to the result of normality test reported in Table K.2, therefore these variables need to be transformed to reduce non-normal distribution in data.

In addition to the three assumptions, the significant outliers were investigated. For each variable, univariate outliers were first virtually detected through a box plot. In box plots of SPSS, outliers are indicated by small circle and extreme outliers are indicated

Table K.4: Descriptive statistic of variable OT_{1-16}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
OT_1	14	2.929	0.616	0.024	0.597	0.302	1.154
OT_2	14	3.000	0.555	0.000	0.597	1.330	1.154
OT_3	14	3.429	0.514	0.325	0.597	-2.241	1.154
OT_4	14	3.143	0.663	-0.151	0.597	-0.310	1.154
OT_5	14	3.500	0.519	0.000	0.597	-2.364	1.154
OT_6	14	3.000	0.679	0.000	0.597	-0.394	1.154
OT_7	14	2.786	0.802	-0.608	0.597	0.801	1.154
OT_8	14	2.786	0.699	0.321	0.597	-0.633	1.154
OT_9	14	2.714	0.611	0.192	0.597	-0.258	1.154
OT_{10}	14	3.143	0.663	-0.151	0.597	-0.310	1.154
OT_{11}	14	3.214	0.579	0.028	0.597	0.209	1.154
OT_{12}	14	3.357	0.497	0.670	0.597	-1.838	1.154
OT_{13}	14	3.071	0.475	0.308	0.597	2 .923	1.154
OT_{14}	14	3.214	0.426	1.566	0.597	0.501	1.154
OT_{15}	14	3.214	0.426	1.566	0.597	0.501	1.154
OT_{16}	14	3.000	0.392	0.000	0.597	6.500	1.154

by asterisks (*). Next to these outliers is the number of the case associated to the outlier. Please note that only extreme outliers are concerned in this study. Then a variable that its box plot presented an extreme outlier was re-observed to find outliers using z-scores.

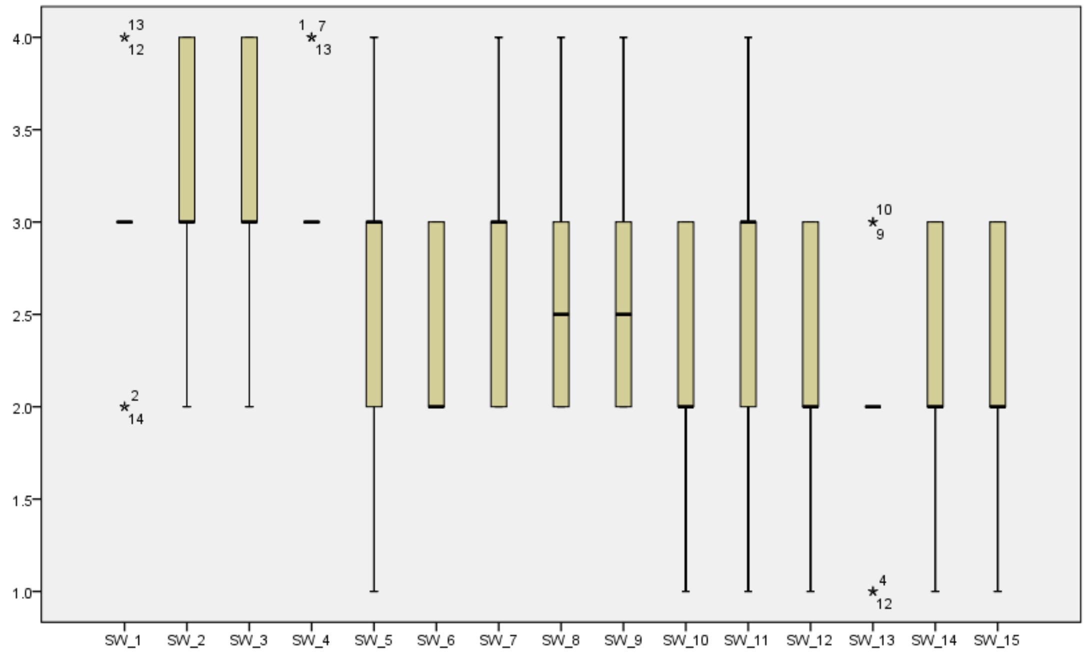


Figure K.1: Box plots of variables SW_{1-15}

There is a presence of extreme outliers in three variables in Figure K.1 named SW_1, SW_4 and SW_{13}. There were four extreme outliers for variable SW_1 and SW_{13} whereas there were three extreme outliers for variable SW_4. Figure K.2 also showed a presence

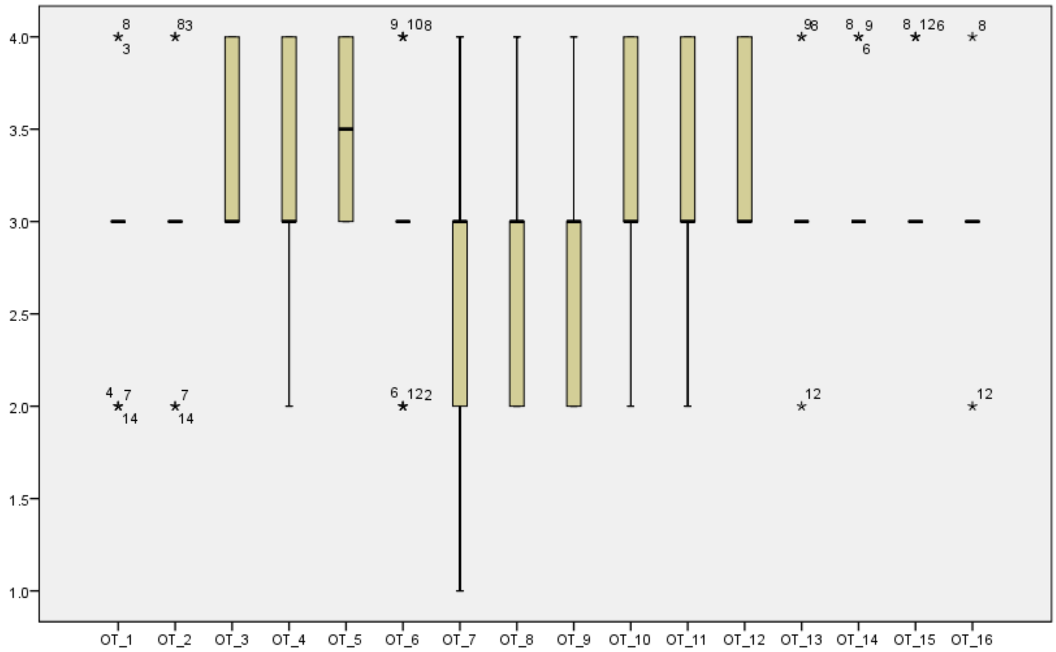


Figure K.2: Box plots of variables OT_{1-16}

of extreme outliers in seven variables named OT_{1-2}, OT_6 and OT_{13-16}. All of them have at least two extreme outliers.

Variable that its box plot presented an extreme outlier was re-observed to find outliers using z-scores. To look for outliers using z-scores, first the values of variable were converted to z-scores. Then z-scores were explored in which z-scores ±3.29 or beyond indicated a case of significant outlier (Tabachnick et al., 2001). This procedure was conducted using SPSS syntax provided by Field (2009) statistic book.

Regarding result produced by this syntax, it is expected to see majority of cases (95%) with absolute value less than 1.96, 5% (or less) with an absolute value greater than 1.96, and 1% (or less) with an absolute value greater than 2.58. Finally, it is expected that there are no cases above 3.29 which indicted significant outliers.

The results of identifying outliers using z-scores for variable named SW_1, SW_4 and SW_{13} are shown in Table K.5 to Table K.7 respectively. All of the cases of variable SW_1, SW_4 and SW_{13} have z-score less than 2.0 indicates no sign of outliers of these variables.

Table K.5: z-score frequency table of SW_1

	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

The results of identifying outliers using z-scores for variable named OT_{1-2}, OT_6 and OT_{13-16} are shown in Table K.8 to Table K.14 respectively. There were no sign of

Table K.6: z-score frequency table of SW_4

	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

Table K.7: z-score frequency table of SW_{13}

	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

outlier in five out of seven variables as these variables which are OT_{1-2}, OT_6, OT_{14-15} have z-score less than 2.0. Table K.11 and Table K.14 presented that there were some cases of variable OT_{13} and OT_{16} have z-score greater than 2.0 but still less than 3.29 indicated that there were no significant outliers. Hence, outliers were not a problem of these two variables.

Table K.8: z-score frequency table of OT_1

	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

Table K.9: z-score frequency table of OT_2

	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

Table K.10: z-score frequency table of OT_6

	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

Table K.11: z-score frequency table of OT_{13}

	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Absolute z-score less than 2.0	13	92.9	92.9	92.9
Absolute z-score greater than 1.96	1	7.1	7.1	100.0
Total	14	100.0		

Table K.12: z-score frequency table of OT_{14}

	Frequency	Percent	Valid Percent	Cumulative Percent
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

Table K.13: z-score frequency table of OT_{15}

	Frequency	Percent	Valid Percent	Cumulative Percent
Absolute z-score less than 2.0	14	100.0	100.0	100.0
Total	14	100.0		

Table K.14: z-score frequency table of OT_{16}

	Frequency	Percent	Valid Percent	Cumulative Percent
Absolute z-score less than 2.0	12	85.7	85.7	85.7
Absolute z-score greater than 1.96	2	14.3	14.3	100.0
Total	14	100.0		

K.2 Result of the normality test for one sample t-test on the transformed staff agreement level of SWOT factor

Although there was no sign of extreme outlier in data, some of variables in SW_{1-15} and OT_{1-16} were violated the normality assumption. Therefore, variables with skewed and peaked distribution need to be transformed, and others variables in SW_{1-15} and OT_{1-16} that have relatively normal distribution would also have to transform to be able to compare with others that violate assumption.

As most data had positive skewness, Log, square root, and reciprocal function were used to transform data. The results of normality test, skewness and kurtosis after transformed data using Log, square root, and reciprocal function are shown in Table K. 15-K.18, Table K.19-K.22, Table K.23-K.26 respectively.

After transform data using the three functions, the data still did not have normal distribution. As p-value of normality test was less than 0.05 for all variables and there were some variables that their skewness or kurtosis were outside the range. These variables still not met the assumptions of one-sample t-test therefore non-parametric test was recommended to be used instead of one-sample t-test.

Table K.15: Shapiro-Wilk test for normality of Log transformed for SW_{1-15}

Variable	Shapiro-Wilk			Normality
	Statistic	**df**	**p-value**	
$logSW_1$	0.726	14	0.001	No
$logSW_2$	0.748	14	0.001	No
$logSW_3$	0.765	14	0.002	No
$logSW_4$	0.516	14	0.000	No
$logSW_5$	0.862	14	0.032	No
$logSW_6$	0.616	14	0.000	No
$logSW_7$	0.813	14	0.007	No
$logSW_8$	0.773	14	0.002	No
$logSW_9$	0.756	14	0.001	No
$logSW_{10}$	0.760	14	0.002	No
$logSW_{11}$	0.835	14	0.014	No
$logSW_{12}$	0.743	14	0.001	No
$logSW_{13}$	0.720	14	0.001	No
$logSW_{14}$	0.743	14	0.001	No
$logSW_{15}$	0.760	14	0.002	No

Table K.16: Shapiro-Wilk test for normality of Log transformed for OT_{1-16}

Variable	Shapiro-Wilk			Normality
	Statistic	**df**	**p-value**	
$logOT_1$	0.771	14	0.002	No
$logOT_2$	0.726	14	0.001	No
$logOT_3$	0.639	14	0.000	No
$logOT_4$	0.794	14	0.004	No
$logOT_5$	0.646	14	0.000	No
$logOT_6$	0.803	14	0.005	No
$logOT_7$	0.808	14	0.006	No
$logOT_8$	0.802	14	0.005	No
$logOT_9$	0.761	14	0.002	No
$logOT_{10}$	0.794	14	0.004	No
$logOT_{11}$	0.748	14	0.001	No
$logOT_{12}$	0.616	14	0.000	No
$logOT_{13}$	0.649	14	0.000	No
$logOT_{14}$	0.516	14	0.000	No
$logOT_{15}$	0.516	14	0.000	No
$logOT_{16}$	0.544	14	0.000	No

Table K.17: Descriptive statistic of Log transformed for SW_{1-15}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
$logSW_1$	14	0.598	0.062	-0.561	0.597	1.461	1.154
$logSW_2$	14	0.621	0.061	-0.444	0.597	1.283	1.154
$logSW_3$	14	0.628	0.065	-0.597	0.597	0.863	1.154
$logSW_4$	14	0.623	0.041	1.566	0.597	0.501	1.154
$logSW_5$	14	0.537	0.125	-0.784	0.597	0.027	1.154
$logSW_6$	14	0.522	0.062	0.670	0.597	-1.838	1.154
$logSW_7$	14	0.578	0.087	0.027	0.597	-1.351	1.154
$logSW_8$	14	0.553	0.085	0.522	0.597	-1.174	1.154
$logSW_9$	14	0.547	0.076	0.453	0.597	-1.082	1.154
$logSW_{10}$	14	0.509	0.085	-0.761	0.597	1.367	1.154
$logSW_{11}$	14	0.543	0.100	-0.936	0.597	1.371	1.154
$logSW_{12}$	14	0.500	0.081	-0.633	0.597	1.777	1.154
$logSW_{13}$	14	0.470	0.084	-0.746	0.597	1.563	1.154
$logSW_{14}$	14	0.500	0.081	-0.633	0.597	1.777	1.154
$logSW_{15}$	14	0.509	0.085	-0.761	0.597	1.367	1.154

Table K.18: Descriptive statistic of Log transformed for OT_{1-16}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
$logOT_1$	14	0.589	0.070	-0.409	0.597	0.104	1.154
$logOT_2$	14	0.598	0.062	-0.561	0.597	1.461	1.154
$logOT_3$	14	0.644	0.050	0.325	0.597	-2.241	1.154
$logOT_4$	14	0.612	0.072	-0.544	0.597	0.207	1.154
$logOT_5$	14	0.651	0.050	0.000	0.597	-2.364	1.154
$logOT_6$	14	0.596	0.076	-0.367	0.597	-0.350	1.154
$logOT_7$	14	0.568	0.104	-1.315	0.597	2.374	1.154
$logOT_8$	14	0.571	0.080	0.034	0.597	-1.084	1.154
$logOT_9$	14	0.564	0.072	-0.108	0.597	-0.966	1.154
$logOT_{10}$	14	0.612	0.072	-0.544	0.597	0.207	1.154
$logOT_{11}$	14	0.621	0.061	-0.444	0.597	1.283	1.154
$logOT_{12}$	14	0.637	0.048	0.670	0.597	-1.838	1.154
$logOT_{13}$	14	0.607	0.051	-0.425	0.597	3.785	1.154
$logOT_{14}$	14	0.623	0.041	1.566	0.597	0.501	1.154
$logOT_{15}$	14	0.623	0.041	1.566	0.597	0.501	1.154
$logOT_{16}$	14	0.600	0.044	-0.950	0.597	6.932	1.154

Table K.19: Shapiro-Wilk test for normality of Square root transformed for SW_{1-15}

Variable	Shapiro-Wilk			Normality
	Statistic	df	p-value	
$sqrtSW_1$	0.731	14	0.001	No
$sqrtSW_2$	0.750	14	0.001	No
$sqrtSW_3$	0.768	14	0.002	No
$sqrtSW_4$	0.516	14	0.000	No
$sqrtSW_5$	0.874	14	0.047	No
$sqrtSW_6$	0.616	14	0.000	No
$sqrtSW_7$	0.815	14	0.008	No
$sqrtSW_8$	0.774	14	0.002	No
$sqrtSW_9$	0.757	14	0.002	No
$sqrtSW_{10}$	0.764	14	0.002	No
$sqrtSW_{11}$	0.846	14	0.019	No
$sqrtSW_{12}$	0.747	14	0.001	No
$sqrtSW_{13}$	0.726	14	0.001	No
$sqrtSW_{14}$	0.747	14	0.001	No
$sqrtSW_{15}$	0.764	14	0.002	No

Table K.20: Shapiro-Wilk test for normality of Square root transformed for OT_{1-16}

Variable	Shapiro-Wilk			Normality
	Statistic	df	p-value	
$sqrtOT_1$	0.775	14	0.002	No
$sqrtOT_2$	0.731	14	0.001	No
$sqrtOT_3$	0.639	14	0.000	No
$sqrtOT_4$	0.798	14	0.005	No
$sqrtOT_5$	0.646	14	0.000	No
$sqrtOT_6$	0.807	14	0.006	No
$sqrtOT_7$	0.823	14	0.010	No
$sqrtOT_8$	0.805	14	0.006	No
$sqrtOT_9$	0.764	14	0.002	No
$sqrtOT_{10}$	0.798	14	0.005	No
$sqrtOT_{11}$	0.75	14	0.001	No
$sqrtOT_{12}$	0.616	14	0.000	No
$sqrtOT_{13}$	0.652	14	0.000	No
$sqrtOT_{14}$	0.516	14	0.000	No
$sqrtOT_{15}$	0.516	14	0.000	No
$sqrtOT_{16}$	0.548	14	0.000	No

Table K.21: Descriptive statistic of Square root transformed for SW_{1-15}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
$sqrtSW_1$	14	1.725	0.163	-0.381	0.597	1.390	1.154
$sqrtSW_2$	14	1.786	0.164	-0.277	0.597	0.878	1.154
$sqrtSW_3$	14	1.805	0.172	-0.453	0.597	0.445	1.154
$sqrtSW_4$	14	1.789	0.114	1.566	0.597	0.501	1.154
$sqrtSW_5$	14	1.575	0.313	-0.643	0.597	-0.143	1.154
$sqrtSW_6$	14	1.528	0.158	0.670	0.597	-1.838	1.154
$sqrtSW_7$	14	1.676	0.228	0.102	0.597	-1.282	1.154
$sqrtSW_8$	14	1.611	0.223	0.586	0.597	-1.012	1.154
$sqrtSW_9$	14	1.592	0.197	0.523	0.597	-0.848	1.154
$sqrtSW_{10}$	14	1.498	0.211	-0.616	0.597	0.921	1.154
$sqrtSW_{11}$	14	1.585	0.250	-0.761	0.597	1.010	1.154
$sqrtSW_{12}$	14	1.475	0.201	-0.466	0.597	1.340	1.154
$sqrtSW_{13}$	14	1.400	0.204	-0.584	0.597	1.472	1.154
$sqrtSW_{14}$	14	1.475	0.201	-0.466	0.597	1.340	1.154
$sqrtSW_{15}$	14	1.498	0.211	-0.616	0.597	0.921	1.154

Table K.22: Descriptive statistic of Square root transformed for OT_{1-16}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
$sqrtOT_1$	14	1.702	0.183	-0.274	0.597	0.149	1.154
$sqrtOT_2$	14	1.725	0.163	-0.381	0.597	1.390	1.154
$sqrtOT_3$	14	1.847	0.138	0.325	0.597	-2.241	1.154
$sqrtOT_4$	14	1.763	0.191	-0.412	0.597	0.022	1.154
$sqrtOT_5$	14	1.866	0.139	0.000	0.597	-2.364	1.154
$sqrtOT_6$	14	1.721	0.199	-0.249	0.597	-0.374	1.154
$sqrtOT_7$	14	1.650	0.261	-1.124	0.597	1.877	1.154
$sqrtOT_8$	14	1.657	0.209	0.122	0.597	-0.953	1.154
$sqrtOT_9$	14	1.638	0.187	-0.021	0.597	-0.773	1.154
$sqrtOT_{10}$	14	1.763	0.191	-0.412	0.597	0.022	1.154
$sqrtOT_{11}$	14	1.786	0.164	-0.277	0.597	0.878	1.154
$sqrtOT_{12}$	14	1.828	0.133	0.670	0.597	-1.838	1.154
$sqrtOT_{13}$	14	1.748	0.136	-0.176	0.597	3.436	1.154
$sqrtOT_{14}$	14	1.790	0.114	1.566	0.597	0.501	1.154
$sqrtOT_{15}$	14	1.790	0.114	1.566	0.597	0.501	1.154
$sqrtOT_{16}$	14	1.729	0.115	-0.646	0.597	6.699	1.154

Table K.23: Shapiro-Wilk test for normality of Reciprocal transformed for SW_{1-15}

Variable	Shapiro-Wilk			Normality
	Statistic	df	p-value	
$newSW_1$	0.681	14	0.000	No
$newSW_2$	0.706	14	0.000	No
$newSW_3$	0.722	14	0.001	No
$newSW_4$	0.516	14	0.000	No
$newSW_5$	0.720	14	0.001	No
$newSW_6$	0.616	14	0.000	No
$newSW_7$	0.787	14	0.003	No
$newSW_8$	0.758	14	0.002	No
$newSW_9$	0.737	14	0.001	No
$newSW_{10}$	0.648	14	0.000	No
$newSW_{11}$	0.684	14	0.000	No
$newSW_{12}$	0.632	14	0.000	No
$newSW_{13}$	0.623	14	0.000	No
$newSW_{14}$	0.632	14	0.000	No
$newSW_{15}$	0.648	14	0.000	No

Table K.24: Shapiro-Wilk test for normality of Reciprocal transformed for OT_{1-16}

Variable	Shapiro-Wilk			Normality
	Statistic	df	p-value	
$newOT_1$	0.729	14	0.001	No
$newOT_2$	0.681	14	0.000	No
$newOT_3$	0.639	14	0.000	No
$newOT_4$	0.749	14	0.001	No
$newOT_5$	0.646	14	0.000	No
$newOT_6$	0.762	14	0.002	No
$newOT_7$	0.638	14	0.000	No
$newOT_8$	0.773	14	0.002	No
$newOT_9$	0.733	14	0.001	No
$newOT_{10}$	0.749	14	0.001	No
$newOT_{11}$	0.706	14	0.000	No
$newOT_{12}$	0.616	14	0.000	No
$newOT_{13}$	0.606	14	0.000	No
$newOT_{14}$	0.516	14	0.000	No
$newOT_{15}$	0.516	14	0.000	No
$newOT_{16}$	0.503	14	0.000	No

Table K.25: Descriptive statistic of Reciprocal transformed for SW_{1-15}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
$newSW_1$	14	0.345	0.072	1.361	0.597	2.122	1.154
$newSW_2$	14	0.321	0.064	1.442	0.597	4.238	1.154
$newSW_3$	14	0.316	0.067	1.482	0.597	3.847	1.154
$newSW_4$	14	0.316	0.036	-1.566	0.597	0.501	1.154
$newSW_5$	14	0.464	0.244	1.696	0.597	2.056	1.154
$newSW_6$	14	0.441	0.083	-0.670	0.597	-1.838	1.154
$newSW_7$	14	0.375	0.102	0.293	0.597	-1.618	1.154
$newSW_8$	14	0.405	0.103	-0.265	0.597	-1.796	1.154
$newSW_9$	14	0.411	0.095	-0.203	0.597	-1.855	1.154
$newSW_{10}$	14	0.476	0.171	2.312	0.597	7.185	1.154
$newSW_{11}$	14	0.435	0.185	2.375	0.597	6.952	1.154
$newSW_{12}$	14	0.488	0.166	2.332	0.597	7.539	1.154
$newSW_{13}$	14	0.548	0.201	1.824	0.597	2.795	1.154
$newSW_{14}$	14	0.488	0.166	2.332	0.597	7.539	1.154
$newSW_{15}$	14	0.476	0.171	2.312	0.597	7.185	1.154

Table K.26: Descriptive statistic of Reciprocal transformed for OT_{1-16}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
$newOT_1$	14	0.357	0.083	0.972	0.597	0.092	1.154
$newOT_2$	14	0.345	0.072	1.361	0.597	2.122	1.154
$newOT_3$	14	0.298	0.043	-0.325	0.597	-2.241	1.154
$newOT_4$	14	0.333	0.080	1.214	0.597	1.330	1.154
$newOT_5$	14	0.292	0.043	0.000	0.597	-2.364	1.154
$newOT_6$	14	0.351	0.088	0.899	0.597	-0.124	1.154
$newOT_7$	14	0.405	0.190	2.650	0.597	8.050	1.154
$newOT_8$	14	0.381	0.097	0.319	0.597	-1.542	1.154
$newOT_9$	14	0.387	0.090	0.421	0.597	-1.552	1.154
$newOT_{10}$	14	0.333	0.080	1.214	0.597	1.330	1.154
$newOT_{11}$	14	0.321	0.064	1.442	0.597	4.238	1.154
$newOT_{12}$	14	0.304	0.041	-0.670	0.597	-1.838	1.154
$newOT_{13}$	14	0.333	0.057	1.717	0.597	6.500	1.154
$newOT_{14}$	14	0.316	0.036	-1.566	0.597	0.501	1.154
$newOT_{15}$	14	0.316	0.036	-1.566	0.597	0.501	1.154
$newOT_{16}$	14	0.339	0.051	2.283	0.597	9.060	1.154

Appendix L

Assumption checking of one sample t-test for staff evaluation on the group of SWOT factor

There are 4 variables represented the average on level of agreement of CPE-KU-KPS staff towards each aspect of the produced SWOT (denoted as S_{avg}, W_{avg}, O_{avg}, T_{avg}). Before the one-sample t-test will be conducted to test that the average of staff agreement for each SWOT aspect is not significant from an acceptable threshold, the assumptions of one sample t-test were checked for each variable in the dataset to ensure that the result of t-test was valid.

The one-sample t-test required three assumptions to be met as follows (Field, 2009):

Assumption 1: Variables should be interval/ ratio level.

All 4 variables represented the average level of agreement of CPE-KU-KPS staff that were measured on a scale of 1 to 4. This assumption there was met.

Assumption 2: The data are independent.

Responses to the level of agreement of CPE-KU-KPS staff are independent because they come from a different person. This assumption there was met.

Assumption 3: Sampling distribution of the dataset is relatively normal. The normality for each variable was tested by using Shapiro-Wilk in SPSS since the size of the dataset is smaller than 2000. A review of the Shapiro-Wilk test for normality for variable S_{avg}, W_{avg}, O_{avg} and T_{avg} was shown in Table L.1.

A result of the Shapiro-Wilk test for normality presented in Table L.1 suggests that the data is normally distributed since p-values of all variables are greater than 0.05.

Table L.1: Shapiro -Wilk test for normality of variable S_{avg}, W_{avg}, O_{avg} and T_{avg}

Variable	**Shapiro-Wilk**			**Normality**
	Statistic	**df**	**p-value**	
S_{avg}	0.899	7	0.323	Yes
W_{avg}	0.907	8	0.332	Yes
O_{avg}	0.951	11	0.658	Yes
T_{avg}	0.881	5	0.314	Yes

Addition to the normality test, skew and kurtosis of each variable are also observed and shown in Table L.2. Regarding George and Mallery (2003), skew and kurtosis values between -1.0 to $+1.0$ are reasonable ranges to accept that data is reasonably close to normal.

Most of the values of skewness and kurtosis reported in Table L.2 were within the range for reasonably concluded that data is relatively normal which is consistent to the result of normality test.

Table L.2: Descriptive statistic of variable S_{avg}, W_{avg}, O_{avg} and T_{avg}

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
S_{avg}	7	2.929	0.355	-0.731	0.794	-0.905	1.587
W_{avg}	8	2.348	0.224	0.005	0.752	-1.107	1.481
O_{avg}	11	3.188	0.184	0.321	0.661	-0.852	1.279
T_{avg}	5	2.886	0.179	0.828	0.913	-1.217	2.000

In addition to the three assumptions, the significant outliers were investigated through a box plot. In box plots of SPSS, outliers are indicated by a small circle and extreme outliers are indicated by asterisks (*). Next to these outliers is the number of the case associated to the outlier. Figure L.1 shows no sign of extreme outliers for all variables.

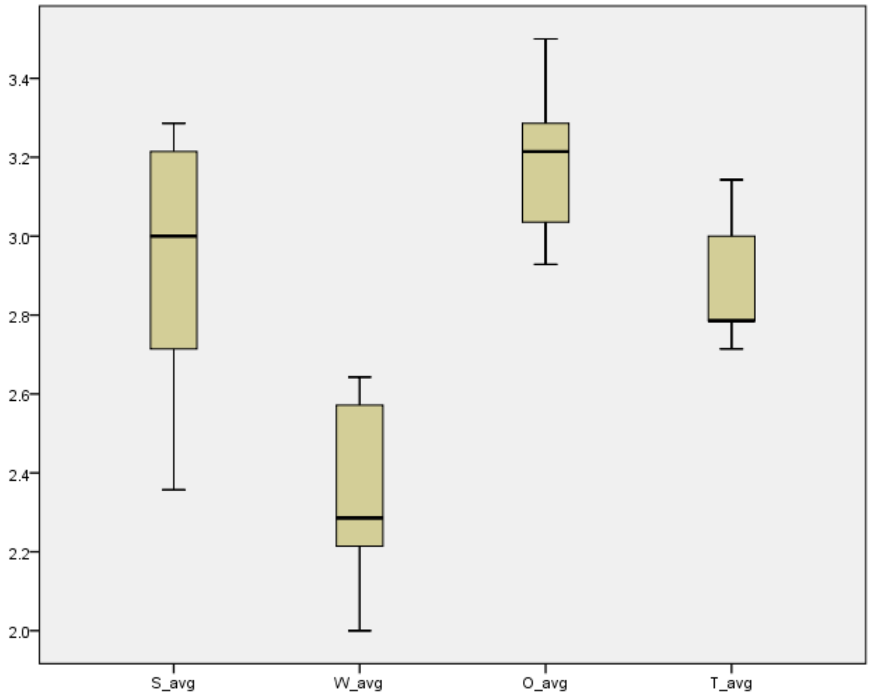


Figure L.1: Box plots of variables S_{avg}, W_{avg}, O_{avg} and T_{avg}

Appendix M

Experienced users of SWOT survey material

M.1 Questionnaire on the quality of an outcome of IPA based SWOT

This questionnaire forms part of a study into the development of SWOT based on customer satisfaction survey. The data collected by this questionnaire is used as a basis for evaluating the SWOT that has been created. Hence, this survey is very important for researchers and to the success of the whole research. Your information will be used for this research purpose only. Thank you very much for your time in completing this questionnaire.

There are three parts to this questionnaire:

Part I: A case study of creating SWOT

Part II: Questions regarding the quality of produced SWOT

Part III: Demographic questions asking the participant's background

Please read a case study described in Part I of questionnaire and then indicate the level of your agreement by ticking ✓ in the provided box in Part II of questionnaire. Finally, please provide some of your information in Part III of questionnaire.

I: A case study of creating SWOT

The department of IT innovation of XYZ University was established in 2006. The mission of the department is to produce graduates with quality, integrity and ethics, as well as support the country with know-how in the field of information technology. At present, the department of IT innovation still firmly commits to the mission and continuously improves teaching quality, research, and academic service to produce specialized graduates.

With the goal of being in the top 10 in Thailand, the department of IT innovation has to perform its role effectively. One way to assess how well the performance of the department is to conduct student satisfaction survey toward the aspects of the department; for example, Teacher, Computer Facilities etc. This assessment of student satisfaction can identify which of department's aspect meet or not meet the student satisfaction. Thus, the department can promote highly performing areas and improve those poorly performing areas.

To make the best use of data collected through questionnaire, a member of the department had come up with the idea of using the result of student satisfaction survey to identify SWOT of the department for formulating a concrete future plan. In order to create SWOT based on student satisfaction survey, one section asking students about their level of importance (need/ expectation) with the aspects of the department had been added to the survey. Based on the result of the survey, the strengths and weaknesses were identified and presented in Table M.1.

Opportunities and threats were identified under a supposition that "*the strengths of competitor become the threats of the company and the weaknesses of competitors can become the opportunities of the company*". Hence, the department of ICT of ABC University was identified as competitor. Then opportunities and threats were identified through the survey result of ICT department which collected by using the same questionnaire of IT innovation department. These opportunities and threats were presented in Table M. 1 and steps for generating SWOT were shown in Figure M.1.

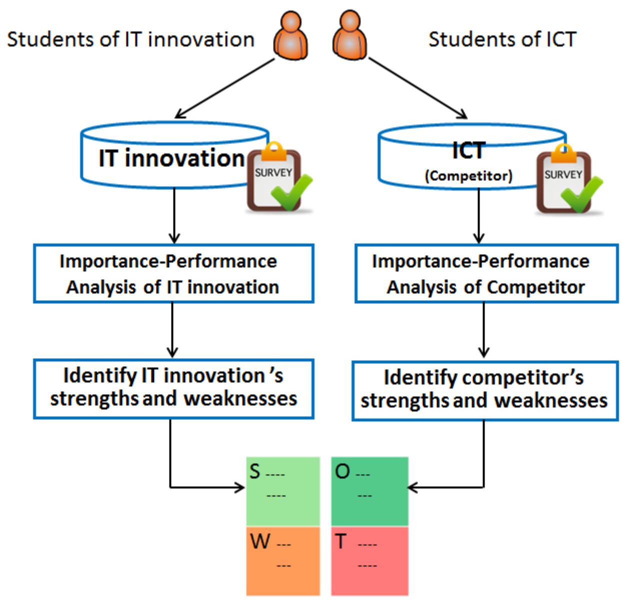


Figure M.1: Steps for generating SWOT based on customers satisfaction survey

Table M.1: SWOT of department of IT Innovation (order by weight)

SWOT groups	SWOT factors	Weight*
Strengths	Knowledge of the information about courses, exams, activities of non-academic staff	36.6
	Subject expertise of teaching staff	23.1
	Teaching ability of teaching staff	19.4
	Friendliness of teaching staff towards students	13.3
	Ability of teaching staff to give advice and support to student learning	8.2
	Appropriate number of students per class	3.1
	Knowledge of rules and procedures of non-academic staff	0.0
Weaknesses	Lack of availability of library facilities	-29.7
	Lack of medical support provided for students	-21.6
	Lack of availability of computer facilities for students	-21.6
	Lack of availability of printing and photocopying facilities for students	-17.6
	Poor arrangement of health development activities	-17.4
	Poor quality of computer facilities for students (Hardware and Software)	-14.9
	Poor arrangement of cultural exchange programs with foreign countries	-4.9
	Lack of financial aid provided for students	-3.1
Opportunities	Improving the arrangement of personal learning and thinking skills development activities	25.3
	Increasing the availability of lecture materials	23.3
	Increasing E-learning resources to support student learning	12.4
	Improving the arrangement of moral development activities	10.6
	Increasing the availability of teaching staff to the students	10.5
	Updated content and easy to find information on the department website	9.6
	Improving the arrangement of field trips activities	6.6
	Increasing the ability to give clear and timely feedback to students	6.0
	Improving the arrangement of interpersonal skills development activities	2.5
	Increasing accurate and up-to-date course unit content	0.0
	Improving the arrangement of social volunteer activities	0.0
Threats	Inadequate of teaching facilities and learning areas	-26.4
	Lack of friendliness of non-academic staff to students	-11.2
	Lack of ability of non-academic staff to provide services in a timely manner	-10.0
	Lack of interest in solving the students'problem by non-academic staff	0.0
	Lack of availability of internet access for students	0.0

* Compute by multiplying the importance by positive/negative performance (for strength, opportunity /weakness, threat)

II: Questions regarding the quality of produced SWOT

Please use SWOT of IT innovation department presented in Table 1 and the step for generating SWOT based on customer's satisfaction survey presented in Figure M.1 to answer the questions 1-12 by ticking ✓ in the box corresponded to your opinion.

1) Strengths and weakness shown in Table M.1 are explicit (clearly and unambiguously formulated).

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

2) Opportunities and threats shown in Table M.1 are explicit (clearly and unambiguously formulated).

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

3) How many SWOT items shown in Table M.1 that is overgeneralised (Too broad description)?

All of them (100%)	***Most of them (75%)***	***Half of them (50%)***	***Some of them (25%)***	***None of them (0%)***
☐	☐	☐	☐	☐

4) How many SWOT items shown in Table M.1 that is incorrectly classified as internal/ external factors?

All of them (100%)	***Most of them (75%)***	***Half of them (50%)***	***Some of them (25%)***	***None of them (0%)***
☐	☐	☐	☐	☐

5) The SWOT items shown in Table M.1 are comprehensively explain the department's situation

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

6) All SWOT items shown in Table M.1 are measurable.

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

7) The order of SWOT items shown in Table M.1 makes decision-making easier.

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

8) The data source for this SWOT analysis is reliable.

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

9) The SWOT items shown in Table M.1 can be used as a starting point for strategic planning.

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

10) The SWOT items shown in Table M.1 provide useful information that supports decision-making regarding strategic planning.

Strongly disagree	***Disagree***	***Neither agree nor disagree***	***Agree***	***Strongly agree***
☐	☐	☐	☐	☐

11) If you were taken part of developing SWOT for your organisation, how would you like to use this approach to create SWOT for your company/organisation?

Definitely not use	***Probably not use***	***Neither use nor not use***	***Probably use***	***Definitely use***
☐	☐	☐	☐	☐

12) What recommendation would you like to make regarding the development of SWOT based on customer satisfaction?

III: Demographic questions asking the participant's background

Please tick ✓ in the provided box that corresponds to your answer for each question below:

1. What is your gender?	☐ Male ☐ Prefer not to specify ☐ Female
2. What is your age range?	☐ 21 - 25 ☐ 31 - 35 ☐ 26 - 30 ☐ > 35
3. What is your occupation?	☐ Postgraduate student ☐ Business owner ☐ Government Officer ☐ Staff/ Sale Manager ☐ Staff/ Manager of Human Resources ☐ Staff/ Manager of Marketing ☐ Consultant ☐ Other. Please specify
4. Do you have any experience in using SWOT analysis in academia? ☐ No ☐ Yes	
5. Do you have any practical experience in using SWOT analysis as a basis for strategy development? ☐ No ☐ Yes If your answer to this question is "Yes", How many years of practical experience have you had in using SWOT analysis?	

Thank you very much for your participation in this study.

M.2 Questionnaire on the quality of an outcome of Traditional SWOT

This questionnaire forms part of a study into the development of SWOT based on customer satisfaction survey. The data collected by this questionnaire is used as a basis for evaluating the SWOT that has been created. Hence, this survey is very important for researchers and to the success of the whole research. Your information will be used for this research purpose only. Thank you very much for your time in completing this questionnaire.

There are three parts to this questionnaire:

Part I: A case study of creating SWOT

Part II: Questions regarding the quality of produced SWOT

Part III: Demographic questions asking the participant's background

Please read a case study described in Part I of questionnaire and then indicate the level of your agreement by ticking ✓ in the provided box in Part II of questionnaire. Finally, please provide some of your information in Part III of questionnaire.

I: A case study of creating SWOT

The department of IT innovation of XYZ University was established in 2006. The mission of the department is to produce graduates with quality, integrity and ethics, as well as support the country with know-how in the field of information technology. At present, the department of IT innovation still firmly commits to the mission and continuously improves teaching quality, research, and academic service to produce specialized graduates.

The department of IT innovation arranges an annual meeting in which staff were informed and discuss previous and future plans of the department. Staff were also brainstormed to create SWOT for each aspect of the department such as Teacher, Teaching, Computer Facilities etc. Steps for generating SWOT through brainstorming were shown in Figure M.2 and the produced SWOT factors were presented in Table M.2 - M.5.

The other parts of the questionnaire can be referred to the previous questionnaire shown in Section M.1.

Table M.2: SWOT of department of IT Innovation (Strengths)

Strengths
1. Academic staff possesses know-how in computer which can support the community.
2. Academic staff are in the same age so they can informally talk or discuss to each other
3. Academic staffs'age is not much different from students'age so students are not afraid to ask for some suggestions from academic staff. This establishes acquaintanceship between academic staff and student.
4. Department has an action plan correspondent to strategy of faculty and this plan is evaluated occasionally.
5. Non-academic staff are shared and learned each other tasks so they can work interchangeably.
6. Non-academic staff have capability to improve their potential in information technology which enable them to work efficiently.
7. Non-academic staff have loyalty to organisation and willing to do some public tasks of faculty
8. Service mind
9. Department has software and network laboratory that available for students of the department and others
10. There is a yearly budget for maintenance computer hardware which ensure the computer availability
11. Department has its own curriculum that combines subjects in computer and electronic which are still in need in the computer industrial
12. Department arranges the cooperative education program with the institution in Taiwan
13. Department arranges the field trip that allows students to visit the top computer company in every academic year
14. There are interpersonal skills development, academic, social volunteer activities
15. Department supports students to develop their knowledge by taking part of software development contest etc.
16. All useful information for students has been advertised regularly.
17. Department arranges activity to educate the alumni
18. Department provides scholarship for students regularly such as academic outstanding scholarship, working scholarship

Table M.3: SWOT of department of IT Innovation (Weaknesses)

Weaknesses
1. Department need to hire external lecturer because some internal academic staff are currently studying phd degree.
2. Academic staff has a lot teaching work load so they have less time to do research
3. Department has shortage of academic staff as some of them are currently studying PhD degree.
4. Information distribution and Information Technology related to the administration service has not developed adequately
5. Some of the staff are lack of proactive management. They often do work instantly.
6. Department faces limitation about software license which affects the computer service.
7. The network efficiency is inadequate due to the lack of IP address.
8. A limited number of books in the library are not meet the students'requirement. Poor management of library service such as not enough space, no new books are purchased, no mechanism to check the missing books.

Table M.4: SWOT of department of IT Innovation (Opportunities)

Opportunities
1. Most of academic staff live in campus which safe their time for travelling to office and enable them to work during night-time. They are also have better working environment than staffs working in other University. 2. Faculty has granted policy for academic staff with master degree to study PhD degree within 2 years since they have started working. 3. Students are learned by using the real and updated instrument which allows them to gain experience and develop their skill continuously. 4. Academic staff are subject expertise which is good for knowledge transfer and sharing. 5. Students live nearby campus which safe their time for travelling to University and increase an opportunity for learning. 6. Department is able to modify the subject in the curriculum in a specific direction as well as to make it fit with the knowledge and ability of new graduated academic staff. 7. Department supports staff to attend a meeting or training for improving their work efficiency. 8. By using the curriculum that combines subjects in computer and electronic, students have great variety of job opportunities. 9. Department has CCNA network laboratory which enables students to learn through the real network device instead of using the network simulation. This facilitates students to adapt their knowledge in real working situation. 10. Field trip in the oversea should be specified in the curriculum so students can visit the oversea company. 11. Department should establish the cooperation with local organisation or local industry. 12. Department has staff who responsible to advertise information to students via website and Facebook which allows students to receive information at a glance. 13. There is a clinic in campus in case that student has caught an accident. 14. Department supports student who has problem about tuition fee by allowing them to pay as monthly instalments

Table M.5: SWOT of department of IT Innovation (Threats)

Threats
1. Compare with other departments that have long been established, the department has no specific research group or direction as well as has no senior academic staff to be the mentor in doing research. 2. Environment and learning area is not encourage students to learn. 3. University regulation is less flexible to operate which affects the performance of department. 4. Managing team of the department have a lot teaching load so they have less time to fully manage the department. 5. Special program students are partially received student loan which is not cover the whole amount of tuition fee. 6. The budget for arranging activities of the department is restricted by finance regulation of university.

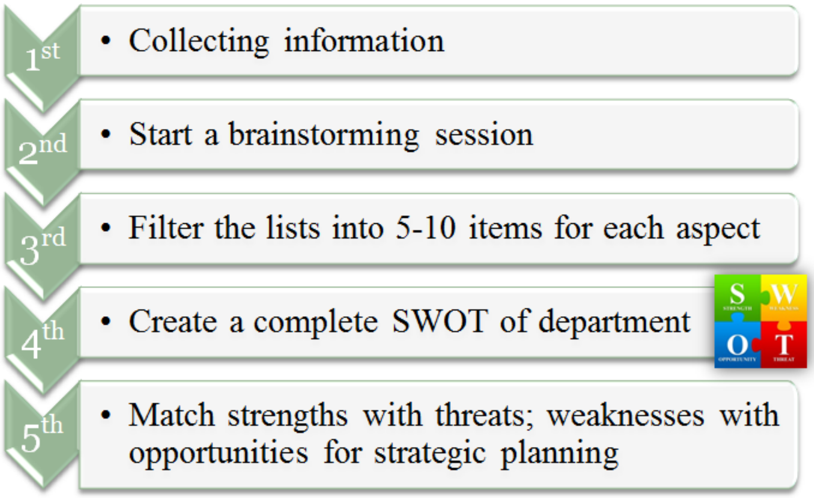


Figure M.2: Steps for generating SWOT through brainstorming

M.3 Participant Information sheet

Study Title: Mining Survey Data for SWOT analysis (validation phase II)
Researcher: Boonyarat Phadermrod
Ethics number: ERGO/FPSE/18500

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

I am completing a PhD research about mining survey data for SWOT analysis which aimed to generate prioritized SWOT factors based on the customer's perception. To evaluate the proficiency of the proposed approach in the real-world situation, the case study of one Higher Education Institution was conducted. Through this case study, your response on the level of agreement towards the quality of SWOT outcome will be analysed to compare the outcome of SWOT produced by two approaches (1) traditional SWOT analysis through the brainstorming session (2) IPA based SWOT analysis.

Why have I been chosen?

You have been approached to participate in this study because you are experienced SWOT user who has relevant knowledge in using SWOT analysis.

What will happen to me if I take part?

There are three sections in the questionnaire which will take no longer than 25 minutes to complete. You will be asked to read a case study and explore SWOT factors produced by two approaches. Then you will be asked to rate your level of agreement measured on the quality of SWOT. Finally, you will be asked to specify your background.

Are there any benefits in my taking part?

By taking part, you have the opportunity to assist the development of a new method to generate SWOT based on satisfaction survey which can be used in government, academic institution and private company.

Are there any risks involved?

There is no risk involved for participants completing the questionnaire.

Will my participation be confidential?

The contents of this questionnaire are absolutely confidential. The name of the participants will not be taken and participation will be kept anonymous. All data will be safe in a protected computer. All will be destroyed once the research is completed.

What happens if I change my mind?

You have the right to withdraw from doing the questionnaire at any time.

What happens if something goes wrong?

If you have any concern or complaint with this research please contacts me (Boonyarat Phadermrod: bp6g12@soton.ac.uk).

Where can I get more information?

If you would like more information on this research please feel free to contact me (Boonyarat Phadermrod: bp6g12@soton.ac.uk)

M.4 Result of language consistency translation test of experienced users of SWOT survey

Remark: +1 the question/ instruction is consistent with the English version

0 undecided about whether Thai question/ instruction is consistent with the English version

−1 the question/ instruction is not consistent with the English version

1) Introduction part of questionnaire

1.1	Question/Intruction (English)	Language consistency score			Total Score
		Expert 1	Expert 2	Expert 3	
	This questionnaire forms part of a study into the development of SWOT based on customer satisfaction survey. The data collected by this questionnaire is used as a basis for evaluating the SWOT that has been created. Hence, this survey is very important for researchers and to the success of the whole research. Your information will be used for this research purpose only. Thank you very much for your time in completing this questionnaire.	+1	+1	+1	3
1.1	**Question/Instruction (Thai)**	**Comment**			
	แบบสอบถามนี้เป็นส่วนหนึ่งของงานวิจัยเพื่อการสร้างจุดแข็ง จุดอ่อน โอกาส และ อุปสรรค (SWOT) จากผลการสำรวจ ความพึงพอใจของลูกค้า โดยที่ข้อมูลที่ได้จากการตอบแบบสอบถามนี้จะถูกนำไปใช้ในการประเมิน SWOT ที่สร้างขึ้น จากงานวิจัยนี้ ดังนั้นการตอบแบบสอบถามของท่าน จึงมีความสำคัญอย่างยิ่งต่อผู้วิจัย และความสมบูรณ์ของงานวิจัย ทั้งนี้ข้อมูลที่ ได้รับจากท่านจะถูกนำไป ใช้เพื่อการวิจัย เท่านั้น และท้ายที่สุด ผู้วิจัยขอขอบคุณท่านที่สละเวลา ตอบแบบสอบถามมา ณ ที่นี้	Reduce the word “วิจัย”			

1.2	Question/Intruction (English)	Language consistency score			Total Score
		Expert 1	Expert 2	Expert 3	
	There are three parts to this questionnaire: Part I: A case study of creating SWOT Part II: Questions regarding the produced SWOT Part III: Demographic questions asking the participant’s background	+1	0	+1	2
1.2	**Question/Instruction (Thai)**	**Comment**			
	แบบสอบถามนี้ ประกอบไปด้วย 3 ตอน ดังต่อไปนี้ ตอนที่ 1. กรณีศึกษาการสร้าง SWOT ขององค์กร ตอนที่ 2. คำถามเกี่ยวกับการประเมินคุณภาพ SWOT ตอนที่ 3. คำถามเกี่ยวกับข้อมูลพื้นฐานของผู้ตอบแบบสอบถาม				

1.3	Question/Intruction (English)	Language consistency score			Total Score
		Expert 1	Expert 2	Expert 3	
	Please read a case study described in Part I of questionnaire and then indicate the level of your agreement by ticking ✓ in the provided box in Part II of questionnaire. Finally, please provide some of your information in Part III of questionnaire.	+1	+1	+1	3
1.3	**Question/Instruction (Thai)**	**Comment**			
	คำชี้แจง: กรุณาอ่านกรณีศึกษาของการสร้าง SWOT ขององค์กรที่อธิบายในตอนที่ 1 ของแบบสอบถาม และตอบคำถาม เกี่ยวกับคุณภาพของ SWOT ในตอนที่ 2 จากนั้นตอบคำถามเกี่ยวกับข้อมูลพื้นฐานของท่านในตอนที่ 3	Add “โดยทำเครื่องหมาย ✓ ลงในช่อง”			

2) First part of questionnaire

<table>
<tr><th>2.1</th><th>Question/Intruction (English)</th><th colspan="3">Language consistency score</th><th rowspan="2">Total Score</th></tr>
<tr><th></th><th></th><th>Expert 1</th><th>Expert 2</th><th>Expert 3</th></tr>
<tr><td colspan="2">Part I: A case study of creating SWOT
The department of IT innovation of XYZ University was established in 2006. The mission of the department is to produce graduates with quality, integrity and ethics, as well as support the country with know-how in the field of information technology. At present, the department of IT innovation still firmly commits to the mission and continuously improves teaching quality, research, and academic service to produce specialized graduates.</td><td>+1</td><td>+1</td><td>+1</td><td>3</td></tr>
<tr><th>2.1</th><th>Question/Instruction (Thai)</th><th colspan="4">Comment</th></tr>
<tr><td colspan="2">ตอนที่ 1. กรณีศึกษาการสร้าง SWOT ขององค์กร
ภาควิชานวัตกรรมเทคโนโลยีสารสนเทศ (Information Technology Innovation) มหาวิทยาลัย XYZ ก่อตั้งขึ้นในปี พ.ศ. 2549 โดยพันธกิจของภาควิชา คือ ผลิตบัณฑิตที่มีความรู้ คุณธรรม และจริยธรรม สร้างสรรค์นวัตกรรม และถ่ายทอดเทคโนโลยีสู่สังคม ในปัจจุบันภาควิชานวัตกรรมเทคโนโลยีสารสนเทศได้ดำเนินงานตามพันธกิจดังกล่าว และยังคงพัฒนาคุณภาพ การเรียนการสอนและงานบริการวิชาการอย่างต่อเนื่อง เพื่อผลิตบัณฑิตที่มีความเชี่ยวชาญด้านเทคโนโลยีสารสนเทศ</td><td colspan="4">Add “อย่างมุ่งมั่น”</td></tr>
</table>

<table>
<tr><th>2.2</th><th>Question/Intruction (English)</th><th colspan="3">Language consistency score</th><th rowspan="2">Total Score</th></tr>
<tr><th></th><th></th><th>Expert 1</th><th>Expert 2</th><th>Expert 3</th></tr>
<tr><td colspan="2">With the goal of being in the top 10 in Thailand, the department of IT innovation has to perform its role effectively. One way to assess how well the performance of the department is to conduct student satisfaction survey toward the aspects of the department; for example, Teacher, Computer Facilities etc. This assessment of student satisfaction can identify which of department's aspect meet or not meet the student satisfaction. Thus, the department can promote highly performing areas and improve those poorly performing areas.</td><td>+1</td><td>+1</td><td>+1</td><td>3</td></tr>
<tr><th>2.2</th><th>Question/Instruction (Thai)</th><th colspan="4">Comment</th></tr>
<tr><td colspan="2">เพื่อให้บรรลุเป้าหมายที่จะเป็นภาควิชาเทคโนโลยีสารสนเทศชั้นนำ 1 ใน 10 ของประเทศไทย ภาควิชานวัตกรรมเทคโนโลยี สารสนเทศจะต้องดำเนินงานอย่างมีประสิทธิภาพ ซึ่งวิธีการหนึ่งในการประเมินประสิทธิภาพการดำเนินงานของภาควิชา (Performance) คือ การสอบถาม ความพึงพอใจของนักศึกษาต่อผลการดำเนินงานในด้านต่างๆ ของภาควิชา เช่น อาจารย์ประจำภาควิชา การเรียนการสอน บริการ คอมพิวเตอร์ เป็นต้น ผลประเมินความพึงพอใจของนักศึกษานี้สามารถชี้ให้เห็นว่าด้านใดของภาควิชาที่นักศึกษามีความพึงพอใจกับการดำเนินงาน และด้านใดของภาควิชาที่มีผลการดำเนินงาน ยังไม่ตรงกับความต้องการของนักศึกษา ซึ่งจะนำไปสู่การปรับปรุงประสิทธิภาพการดำเนินงานในส่วนที่ยังมีผลการดำเนินงาน ไม่ดีเท่าที่ควร ตลอดจนการพัฒนาประสิทธิภาพการดำเนินงานของด้านที่ดีอยู่แล้ว ให้ดียิ่งขึ้น</td><td colspan="4">Replace “สามารถ” with “จะบ่ง”</td></tr>
</table>

2) First part of questionnaire (cont)

2.3	Question/Intruction (English)	Language consistency score			Total Score
		Expert 1	Expert 2	Expert 3	
To make the best use of data collected through questionnaire, a member of the department had come up with the idea of using the result of student satisfaction survey to identify SWOT of the department for formulating a concrete future plan. In order to create SWOT base on student satisfaction survey, one section asking students about their level of importance (need/ expectation) with the aspects of the department had been added to the survey. Based on the result of the survey, the strengths and weaknesses were identified and presented in Table 1.		+1	+1	+1	3
2.3	Question/Instruction (Thai)	Comment			
นอกจากนี้ หนึ่งในสมาชิกของภาควิชาได้เสนอแนวคิดการวิเคราะห์ SWOT ของภาควิชาจากผลการประเมินความพึงพอใจของ นักศึกษาอีกด้วย โดย SWOT ที่ได้จะถูกใช้เป็นข้อมูลสำหรับการสร้างแผนพัฒนาภาควิชาในอนาคต ทั้งนี้ส่วนคำถามเกี่ยว กับความคิดเห็นของนักศึกษาต่อระดับความสำคัญ (Importance) ในด้านต่าง ๆ ของภาควิชาได้ถูกเพิ่มเติมลงในแบบประเมิน ความพึงพอใจของนักศึกษา เพื่อประเมินว่าด้านต่างๆ ของภาควิชาส่งผลต่อคุณภาพการศึกษาของนักศึกษามากหรือน้อยเพียงใด หลังจากที่ภาควิชาได้ทำการสำรวจความพึงพอใจของนักศึกษาโดยใช้แบบประเมินที่ได้ออกแบบมานั้น จุดแข็ง และ จุดอ่อนของภาควิชานวัตกรรมเทคโนโลยีสารสนเทศจึงถูกสร้างขึ้นจากผลการประเมินความพึงพอใจของนักศึกษา ดังแสดงใน ตารางที่ 1.					

2.4	Question/Intruction (English)	Language consistency score			Total Score
		Expert 1	Expert 2	Expert 3	
Opportunities and threats were identified under a supposition that “the strengths of competitor become the threats of the company and the weaknesses of competitors can become the opportunities of the company”. Hence, the department of ICT of ABC University was identified as competitor. Then opportunities and threats were identified through the survey result of ICT department which collected by using the same questionnaire of IT innovation department. These opportunities and threats were presented in Table 1 and steps for generating SWOT were shown in Figure 1.		+1	+1	+1	3
2.4	Question/Instruction (Thai)	Comment			
โอกาส และ อุปสรรค ถูกสร้างขึ้น ตามสมมุติฐานที่ว่า “จุดแข็งของคู่แข่งขององค์กร ถือเป็น อุปสรรคขององค์กร และ จุดอ่อน ของคู่แข่งขององค์กร ถือเป็น โอกาสขององค์กร” ในกรณีศึกษานี้ ภาควิชาเทคโนโลยีสารสนเทศและการสื่อสาร (Information and Communication Technology, ICT) มหาวิทยาลัย ABC ถูกเลือกให้เป็นคู่แข่งของภาควิชานวัตกรรมเทคโนโลยีสารสนเทศ ดังนั้นโอกาส และ อุปสรรค ของภาควิชานวัตกรรมเทคโนโลยีสารสนเทศ จึงถูกสร้างขึ้นจากผลการประเมินความพึงพอใจของ นักศึกษา ICT ที่เก็บข้อมูลโดยใช้ แบบสอบถามชุดเดียวกันกับที่ภาควิชานวัตกรรมเทคโนโลยีสารสนเทศใช้ โอกาส และ อุปสรรคของภาควิชานวัตกรรมเทคโนโลยีสารสนเทศที่ถูกสร้างขึ้นนั้นแสดงในตารางที่1 และ ขั้นตอนการสร้าง SWOT จาก แบบประเมินความพึงพอใจของลูกค้านั้นแสดงในรูปที่ 1.					

2) Second part of questionnaire

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score			Total Score	Comment
			Expert 1	Expert 2	Expert 3		
3. 1	II: Questions regarding the produced SWOT	ตอนที่ 2. คำถามเกี่ยวกับการประเมินคุณภาพ SWOT	+1	0	+1	2	
	Please use SWOT of IT innovation department presented in Table 1 and the step for generating SWOT based on customers'satisfaction survey presented in Figure 1 to answer the questions 1-12 by ticking ✓ in the box corresponded to your opinion.	คำชี้แจง: จาก SWOT ของกรณีศึกษาของภาควิชานวัตกรรมเทคโนโลยีสารสนเทศ แสดงในตารางที่ 1 และ ขั้นตอนการสร้าง SWOT แสดงในรูปที่ 1 กรุณาตอบ คำถามข้อ 1-12 โดยทำเครื่องหมายถูก (✓) ในช่อง ที่ตรงกับความคิดเห็นของท่านมากที่สุด					Add "จากแบบประเมินความพึงพอใจของลูกค้า"
3.2	Strengths and weakness shown in Table 1 are explicit (clearly and unambiguously formulated).	จุดแข็ง และ จุดอ่อน แสดงในตารางที่ 1 มีความชัดเจน เข้าใจง่าย ไม่กำกวม	+1	+1	+1	3	
3.3	Opportunities and threats shown in Table 1 are explicit (clearly and unambiguously formulated).	โอกาส และ อุปสรรค แสดงในตารางที่ 1 มีความชัดเจน เข้าใจง่าย ไม่กำกวม	+1	+1	+1	3	
3.4	How many SWOT items shown in Table 1 that is overgeneralised (Too broad description)?	จากตารางที่ 1 มีจำนวนรายการ SWOT กี่รายการที่มีความหมายกว้างเกินไป ไม่เฉพาะเจาะจง	+1	+1	+1	3	
3.5	How many SWOT items shown in Table 1 that is incorrectly classified as internal/external factors?	จากตารางที่ 1 มีจำนวนรายการ SWOT กี่รายการ ที่ถูกจัดกลุ่มเป็นปัจจัยภายใน/ภายนอก ที่ไม่ถูกต้อง	+1	+1	+1	3	
3.6	The SWOT items shown in Table 1 are comprehensively explain the department's situation	รายการ SWOT แสดงในตารางที่ 1 สามารถอธิบายสถานการณ์ขององค์กรได้ครอบคลุมรอบด้าน	+1	+1	+1	3	
3.7	All SWOT items shown in Table 1 are measurable.	รายการ SWOT แสดงในตารางที่ 1 สามารถวัด และเปรียบเทียบได้	+1	+1	+1	3	Add "ทั้งหมด"
3.8	The order of SWOT items shown in Table 1 makes decision-making easier.	การเรียงลำดับของรายการ SWOT แสดงในตารางที่ 1 ทำให้การตัดสินใจเกี่ยวกับแผนขององค์กรง่ายขึ้น	+1	+1	+1	3	Replace "ทำ" with "ช่วย"
3.9	The data source for this SWOT analysis is reliable.	แหล่งข้อมูลสำหรับวิเคราะห์ SWOT แสดงในตารางที่ 1 มีความน่าเชื่อถือ	+1	+1	+1	3	Replace "สำหรับ" with "ที่นำมาใช้เพื่อการ"

2) Second part of questionnaire (cont)

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score			Total Score	Comment
			Expert 1	Expert 2	Expert 3		
3.10	The SWOT items shown in Table 1 can be used as a starting point for strategic planning.	รายการ SWOT แสดงในตารางที่ 1 สามารถนำไปใช้สำหรับการวางแผนกลยุทธ์ขององค์กรได้	+1	+1	+1	3	
3.11	The SWOT items shown in Table 1 provide useful information that supports decision-making regarding strategic planning.	รายการ SWOT แสดงในตารางที่ 1 นำมาซึ่งข้อมูลที่ เป็นประโยชน์ต่อการตัดสินใจเกี่ยวกับแผนขององค์กร	+1	+1	+1	3	Replace “เกี่ยวกับแผน” with “วางแผนกลยุทธ์”
3.12	If you were taken part of developing SWOT for your organisation, how would you like to use this approach to create SWOT for your company / organisation?	หากท่านมีบทบาทในการสร้าง SWOT ในองค์กรของท่าน ท่านจะนำวิธีการสร้าง SWOT แสดงในรูปที่ 1 ไปใช้ในการสร้าง SWOT ของบริษัทหรือองค์กรของท่านหรือไม่	+1	-1	+1	1	Check the use of How and Yes/No question
3.13	What recommendation would you like to make regarding the development of SWOT based on customer satisfaction?	ข้อเสนอแนะของท่านต่อการสร้าง SWOT จากผลสำรวจ ความพึงพอใจของลูกค้า	+1	-1	+1	1	Change Thai sentence to be the question
3.14	Strongly disagree Disagree Neither agree nor disagree Agree Strongly agree	ไม่เห็นด้วยอย่างยิ่ง ไม่เห็นด้วย ไม่ใช่ทั้งเห็นด้วย หรือ ไม่เห็นด้วย เห็นด้วย เห็นด้วยอย่างยิ่ง	+1	+1	+1	3	
3.15	All of them (100%) Most of them (75%) Half of them (50%) Some of them (25%) None of them (0%)	ทั้งหมดของรายการ (100%) ส่วนใหญ่ของรายการ (75%) ครึ่งหนึ่งของรายการ (50%) บางส่วนของรายการ (25%) ไม่มี (0%)	+1	+1	+1	3	Replace “ทั้งหมด” with “ทุก”
3.16	Definitely not use Probably not use Neither use nor not use Probably use Definitely use	ไม่นำไปใช้ อย่างแน่นอน อาจจะ ไม่นำไปใช้ ไม่แน่ใจ อาจจะ นำไปใช้ นำไปใช้ อย่างแน่นอน	+1	+1	+1	3	

3) Third part of questionnaire

	Question/ Instruction (English)	Question/ Instruction (Thai)	Language consistency score			Total Score	Comment
			Expert 1	Expert 2	Expert 3		
4.1	III: Demographic questions asking the participant's background Please tick ✓ in the provided box that corresponds to your answer for each question below:	ตอนที่ 3. คำถามเกี่ยวกับข้อมูลพื้นฐานของ ผู้ตอบแบบสอบถาม คำชี้แจง: กรุณาทำเครื่องหมายถูก (✓) ในช่องที่ตรงกับข้อมูลของท่านมากที่สุด	+1	+1	+1	3	
4.2	What is your gender? ☐ Male ☐ Female ☐ Prefer not to specify	เพศ ☐ ชาย ☐ หญิง ☐ ไม่ต้องการระบุ	+1	+1	+1	3	
4.3	What is your age range?	อายุ	+1	+1	+1	3	Suggest to use "ช่วงอายุ"
4.4	What is your occupation? ☐ Students ☐ Business owners ☐ Government Officers ☐ Marketing/Sales Managers ☐ Personnel/Human Resources Managers ☐ Consultants ☐ Other. Please specify	อาชีพ ☐ นักศึกษา ระดับบัณฑิตศึกษา ☐ เจ้าของธุรกิจส่วนตัว ☐ พนักงานของรัฐ/ข้าราชการ ☐ พนักงานการตลาด/ผู้จัดการฝ่ายขาย ☐ พนักงานฝ่ายบุคคล/ผู้จัดการฝ่ายบุคคล ☐ ที่ปรึกษา ☐ อื่นๆ โปรดระบุ	+1	+1	-1	1	Check a proper word of occupation list
4.5	Do you have any experience in using SWOT analysis in academia? ☐ Yes ☐ No	ท่านเคยเรียนเกี่ยวกับการวิเคราะห์ SWOT ในสถาบันศึกษาหรือไม่? ☐ เคย ☐ ไม่เคย	0	0	-1	-1	Add "หรือมีประสบการณ์"
4.6	Do you have any practical experience in using SWOT analysis as a basis for strategy development? ☐ Yes ☐ No If your answer to this question is "Yes", How many years of practical experience have you had in using SWOT analysis?	ท่านมีประสบการณ์การวิเคราะห์ SWOT เพื่อสร้างแผนกลยุทธ์ในองค์กรหรือหน่วยงานของท่านหรือไม่ ☐ มี ☐ ไม่มี กรุณาระบุจำนวนปีที่ท่านมีประสบการณ์การวิเคราะห์ SWOT	0	+1	0	1	Check the consistency of the option of question 4.5 and 4.6

Appendix N

Assumption checking of two independent sample t-test on experienced users of SWOT evaluation

There are 10 variables represented the rating score of MBA students towards each question related to the quality of SWOT (denoted as $q_1 - q_{10}$). Before performing the t-test, assumptions of t-test must be met for the test to be accurate. The two independent sample t-test assumes that sample data are measured at least at the interval level, sample data is normally distributed, the variances of the two groups are equal (Homogeneity of variance) and the two groups of samples are independent from each other (Field, 2009). Detail regarding the assumptions checking is explained below:

Assumption 1: Variables should be interval/ ratio level.

All 10 variables represented the rating score of MBA students that were measured on a scale of 1 to 5. This assumption there was met.

Assumption 2: The data are independent.

Responses to the quality of SWOT of MBA students are independent because they come from different participants. There were also different participants in each group with no participant being in more than one group. This assumption there was met.

Assumption 3: Sampling distribution of the dataset is relatively normal. The normality for each variable within each group was tested by using Shapiro-Wilk in SPSS since the size of the dataset is smaller than 2000. A review of the Shapiro-Wilk test for normality for variable $q_1 - q_{10}$ is shown in Table N.1.

Table N.1: Shapiro -Wilk test for normality of variable $q_1 - q_{10}$

Variable		Shapiro-Wilk			Normality
		Statistic	**df**	**p-value**	
q1	IPA based SWOT	0.786	22	0.000	No
	Traditional SWOT	0.796	22	0.000	No
q2	IPA based SWOT	0.732	22	0.000	No
	Traditional SWOT	0.823	22	0.001	No
q3	IPA based SWOT	0.861	22	0.005	No
	Traditional SWOT	0.796	22	0.000	No
q4	IPA based SWOT	0.847	22	0.003	No
	Traditional SWOT	0.868	22	0.007	No
q5	IPA based SWOT	0.858	22	0.005	No
	Traditional SWOT	0.584	22	0.000	No
q6	IPA based SWOT	0.733	22	0.000	No
	Traditional SWOT	0.742	22	0.000	No
q7	IPA based SWOT	0.746	22	0.000	No
	Traditional SWOT	0.757	22	0.000	No
q8	IPA based SWOT	0.879	22	0.012	No
	Traditional SWOT	0.671	22	0.000	No
q9	IPA based SWOT	0.681	22	0.000	No
	Traditional SWOT	0.703	22	0.000	No
q10	IPA based SWOT	0.881	22	0.013	No
	Traditional SWOT	0.332	22	0.000	No

A result of the Shapiro-Wilk test for normality presented in Table N.1 suggests that the data is not normally distributed since p-values of all variables are less than 0.05.

Addition to the normality test, skew and kurtosis of each variable are also observed and shown in Table N.2 - N.3 for variable $q_1 - q_{10}$ of IPA based SWOT and Traditional SWOT group respectively. Regarding George and Mallery (2003), skew and kurtosis values between -1.0 to $+1.0$ are reasonable ranges to accept that data is reasonably close to normal.

Most of the values of skewness and kurtosis reported in Table N.2 and Table N. 3 were within the range for reasonably concluded that data is relatively normal. Except skewness and kurtosis of q_{10} and kurtosis of q_5 within Traditional SWOT group that values were outside the range.

Values of skewness and kurtosis reported in these tables, and normality test reported in Table N.1, suggested that normality was not reasonable for some variables. Additionally, the dataset is relatively small, the central limit theorem cannot be referred to conclude that data is normally distributed. Therefore, the assumption was not met.

Assumption 4: Homogeneity of variance. In addition to the normality assumptions, the equality of variances should be tested when examining the mean difference of

Table N.2: Descriptive statistic of variable $q_1 - q_{10}$ (IPA based SWOT group)

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
q1	22	3.364	0.790	-0.781	0.491	3.030	0.953
q2	22	3.318	0.568	-0.050	0.491	-0.506	0.953
q3	22	3.364	0.790	-0.142	0.491	-0.352	0.953
q4	22	3.000	0.926	0.396	0.491	-0.892	0.953
q5	22	3.364	0.902	-0.835	0.491	1.025	0.953
q6	22	3.864	0.560	-0.074	0.491	0.459	0.953
q7	22	4.045	0.576	0.014	0.491	0.510	0.953
q8	22	3.409	1.008	-0.034	0.491	-1.016	0.953
q9	22	3.500	0.740	-1.163	0.491	-0.019	0.953
q10	22	3.409	0.854	0.058	0.491	-0.399	0.953

Table N.3: Descriptive statistic of variable $q_1 - q_{10}$ (Traditional SWOT group)

Variable	**N**	**Mean**	**Std. Deviation**	**Skewness**		**Kurtosis**	
				Statistic	Std. Error	Statistic	Std. Error
q1	22	3.227	0.752	-0.413	0.491	-1.036	0.953
q2	22	3.500	0.673	0.000	0.491	0.043	0.953
q3	22	3.227	0.752	-0.413	0.491	-1.036	0.953
q4	22	3.136	0.834	0.269	0.491	-0.363	0.953
q5	22	3.909	0.426	-0.637	0.491	3.168	0.953
q6	22	3.409	0.734	-0.847	0.491	-0.538	0.953
q7	22	3.273	0.827	-0.574	0.491	-1.282	0.953
q8	22	4.182	0.501	0.413	0.491	0.752	0.953
q9	22	4.091	0.526	0.142	0.491	1.116	0.953
q10	22	4.091	0.294	3.059	0.491	8.085	0.953

two independent samples in order to guarantee that the difference between groups is not affected by variance. This assumption can be tested with Levene's test (Field, 2009).

Levene's test checks whether the variances of sample groups are statistically different. The Levene's test hypotheses are:

H_0: variances are the same

H_1: variances are different

For each question, Levene's test was conducted to test the equality of variance of two groups and the result of Levene's test is shown in Table N.4.

Referring to Table N.4, for the rating on variable $q_1 - q_4$, the variances were equal for IPA based SWOT and Traditional SWOT group since p-value of these variables were higher than 0.05. But for variable $q_5 - q_{10}$ the variances were significantly different in the two groups since p-value of these variables were less than 0.05. The assumption of homogeneity of variance there was not consistently met for all variables.

Table N.4: Test of Homogeneity of Variance between two groups (Based on Mean)

Variable	**Levene Statistic**	**df1**	**df2**	**p-value**	**Equal variance**
q1	0.025	1	42	0.875	Yes
q2	1.312	1	42	0.259	Yes
q3	0.101	1	42	0.752	Yes
q4	0.292	1	42	0.592	Yes
q5	13.402	1	42	0.001	No
q6	5.459	1	42	0.024	No
q7	9.496	1	42	0.004	No
q8	15.648	1	42	0.000	No
q9	7.194	1	42	0.010	No
q10	27.442	1	42	0.000	No

References

Akiyoshi, M. and Komoda, N. (2005). An analysis framework of enterprise documents for business strategy design. In *International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, volume 1, pages 65–69. IEEE.

Allen, D. R. and Wilburn, M. (2002). Why customer satisfaction. In *Linking customer and employee satisfaction to the bottom line*, pages 1–16. ASQ Quality Press.

Avison, D. and Elliot, S. (2006). Scoping the discipline of information systems. In *Information Systems: The State of the Field*, pages 3–18. John Wiley & Sons Chichester, UK.

Ayub, A., Razzaq, A., Aslam, M. S., and Iftekhar, H. (2013). A conceptual framework on evaluating swot analysis as the mediator in strategic marketing planning through marketing intelligence. *European Journal of Business and Social Sciences*, 2(1):91–98.

Azzalini, A., Scarpa, B., and Walton, G. (2012). *Data analysis and data mining: an introduction*. Oxford University Press.

Azzopardi, E. and Nash, R. (2013). A critical evaluation of importance-performance analysis. *Tourism Management*, 35:222–233.

Bacon, D. R. (2003). A comparison of approaches to importance-performance analysis. *International Journal of Market Research*, 45(1):55–72.

Ban, O.-I. and Bogdan, V. (2013). Using partial correlation coefficient to assess the importance of quality attributes: Study case for tourism romanian consumers. *Economics, Business and Management*, 1(1):21–24.

Bansback, N., Li, L. C., Lynd, L., and Bryan, S. (2014). Exploiting order effects to improve the quality of decisions. *Patient education and counseling*, 96(2):197–203.

Ben-Gal, I. (2007). Bayesian networks. In *Encyclopedia of statistics in quality and reliability*. Wiley Online Library.

Bender, R. and Benner, A. (2000). Calculating ordinal regression models in SAS and S-Plus. *Biometrical Journal*, 42(6):677–699.

Blair, J., Czaja, R. F., and Blair, E. A. (2013). Data collection I: Selecting a method. In *Designing Surveys: A Guide to Decisions and Procedures: A Guide to Decisions and Procedures*, pages 48–82. Sage Publications.

Bosnjak, M. (2014). Customer satisfaction. In *Encyclopedia of quality of life and well-being research*, pages 1409–1412. Springer Science+Business Media Dordrech.

Cassel, C. (2006). Measuring customer satisfaction, a methodological guidance. *Statistics Sweden*, 2(2).

Cengiz, E. (2010). Measuring customer satisfaction: Must or not? *Journal of Naval Science and Engineering*, 6(2):76–88.

Chang, H.-H. and Huang, W.-C. (2006). Application of a quantification SWOT analytical method. *Mathematical and Computer Modelling*, 43(1):158–169.

Chen, C.-K. and Hughes, J. (2004). Using ordinal regression model to analyze student satisfaction questionnaires. *IR applications*, 1(1):1–12.

Chen, L.-F., Lin, S.-P., and Lin, Y.-C. (2010). Application of IPA and Back-Propagation Neural Network to build a new service quality decision making model. *International Journal of Information and Management Sciences*, 21:31.

Chen, P. Y. and Popovich, P. M. (2002). The pearson product-moment correlation. In *Correlation: Parametric and nonparametric measures*, pages 9–24. Sage Publications.

Cheng, J. and Greiner, R. (2001). Learning bayesian belief network classifiers: Algorithms and system. In *Advances in Artificial Intelligence*, pages 141–151. Springer.

Chrzan, K. and Golovashkina, N. (2006). An empirical test of six stated importance measures. *International Journal of Market Research*, 48(6):717–740.

Chu, R. (2002). Stated-importance versus derived-importance customer satisfaction measurement. *Journal of Services Marketing*, 16(4):285–301.

Cohen, S. (2003). Maximum difference scaling: Improved measures of importance and preference for segmentation. In *Sawtooth Software Conference Proceedings, Sawtooth Software, Inc*, volume 530, pages 61–74.

Coman, A. and Ronen, B. (2009). Focused SWOT: diagnosing critical strengths and weaknesses. *International Journal of Production Research*, 47(20):5677–5689.

Cover, T. M. and Thomas, J. A. (2006). Entropy, relative entropy, and mutual information. In *Elements of information theory*, pages 13–54. Wiley-interscience.

Cugnata, F. and Salini, S. (2013). Model-based approach for importance–performance analysis. *Quality & Quantity*, pages 1–12.

Dai, Y., Kakkonen, T., and Sutinen, E. (2011). MinEDec: a decision-support model that combines text-mining technologies with two competitive intelligence analysis methods. *International Journal of Computer Information Systems and Industrial Management Applications*, 3:165–173.

Datamonitor (2003). eBay Inc. SWOT Analysis. `http://www.advisorgate.com/files/p10.pdf` Accessed (April 27, 2013).

De Vaus, D. (2002a). Administering questionnaires. In *Surveys in social research*, pages 122–144. Allen & Unwin.

De Vaus, D. (2002b). Building scales. In *Surveys in social research*, pages 180–199. Allen & Unwin.

Dejaeger, K., Goethals, F., Giangreco, A., Mola, L., and Baesens, B. (2012). Gaining insight into student satisfaction using comprehensible data mining techniques. *European Journal of Operational Research*, 218(2):548–562.

Deng, W.-J., Chen, W.-C., and Pei, W. (2008a). Back-propagation neural network based importance–performance analysis for determining critical service attributes. *Expert Systems with Applications*, 34(2):1115–1125.

Deng, W.-J., Kuo, Y.-F., and Chen, W.-C. (2008b). Revised importance–performance analysis: three-factor theory and benchmarking. *The Service Industries Journal*, 28(1):37–51.

Domingos, P. and Pazzani, M. (1997). On the optimality of the simple bayesian classifier under zero-one loss. *Machine learning*, 29(2-3):103–130.

Doyle, J. R., Green, R. H., and Bottomley, P. A. (1997). Judging relative importance: direct rating and point allocation are not equivalent. *Organizational behavior and human decision processes*, 70(1):65–72.

du Toit, A. (2016). Using environmental scanning to collect strategic information: A South African survey. *International Journal of Information Management*, 36(1):16–24.

Duan, L. and Da Xu, L. (2012). Business intelligence for enterprise systems: a survey. *IEEE Transactions on Industrial Informatics*, 8(3):679–687.

Duke, C. R. and Mount, A. S. (1996). Rediscovering performance-importance analysis of products. *Journal of Product & Brand Management*, 5(2):43–54.

Dyson, R. G. (2004). Strategic development and SWOT analysis at the University of Warwick. *European journal of operational research*, 152(3):631–640.

Easterbrook, S., Singer, J., Storey, M.-A., and Damian, D. (2008). Selecting empirical methods for software engineering research. In *Guide to advanced empirical software engineering*, pages 285–311. Springer.

Eskildsen, J. K. and Kristensen, K. (2006). Enhancing importance-performance analysis. *International Journal of Productivity and Performance Management*, 55(1):40–60.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996a). From data mining to knowledge discovery in databases. *AI magazine*, 17(3):37.

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., et al. (1996b). Knowledge discovery and data mining: Towards a unifying framework. *Knowledge Discovery and Data Mining*, pages 82–88.

Fehringer, D., Hohhof, B., and Johnson, T. (2006). State of the art: competitive intelligence. *Competitive Intelligence Foundation, Washington*.

Feng, M., Mangan, J., Wong, C., Xu, M., and Lalwani, C. (2014). Investigating the different approaches to importance–performance analysis. *The Service Industries Journal*, 34(12):1021–1041.

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Figini, S. and Giudici, P. (2007). Statistical models for customer satisfaction data. *Journal of Quality Technology and Quality Management*, 7(1):69–82.

Fincham, J. E. (2008). Response rates and responsiveness for surveys, standards, and the journal. *American journal of pharmaceutical education*, 72(2):43.

Fink, A. (2003). *How to ask survey questions*, volume 4. Sage Publications.

Fink, A. (2012a). Conducting surveys: Everyone is doing it. In *How to Conduct Surveys: A Step-by-Step Guide: A Step-by-Step Guide*, pages 1–28. Sage Publications.

Fink, A. (2012b). The survey form: questions, scales, and appearance. In *How to Conduct Surveys: A Step-by-Step Guide: A Step-by-Step Guide*, pages 29–56. Sage Publications.

Fontenot, G., Henke, L., Carson, K., and Carson, P. P. (2007). Techniques for determining importance: balancing scientific method and subjectivity. *Journal of Targeting, Measurement and Analysis for Marketing*, 15(3):170–180.

Ford, J. B., Joseph, M., and Joseph, B. (1999). Importance-performance analysis as a strategic tool for service marketers: the case of service quality perceptions of business students in new zealand and the usa. *Journal of Services Marketing*, 13(2):171–186.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., and Bryant, B. E. (1996). The american customer satisfaction index: nature, purpose, and findings. *the Journal of Marketing*, pages 7–18.

Fouladgar, M. M., Yakhchali, S. H., Yazdani-Chamzini, A., and Basiri, M. H. (2011). Evaluating the strategies of Iranian mining sector using a integrated model. In *2011 International Conference on Financial Management and Economics Proceedings*, pages 58–63.

Free Management E-books, . (2013). Pestle analysis: Strategy skills. `www.free-management-ebooks.com/dldebk/dlst-pestle.htm` Accessed (June 18, 2016).

Friedman, N., Geiger, D., and Goldszmidt, M. (1997). Bayesian network classifiers. *Machine learning*, 29(2-3):131–163.

Fritz, C. O., Morris, P. E., and Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1):2.

Fullerton, A. S. (2009). A conceptual framework for ordered logistic regression models. *Sociological methods & research*, 38(2):306–347.

Gao, C.-Y. and Peng, D.-H. (2011). Consolidating SWOT analysis with nonhomogeneous uncertain preference information. *Knowledge-Based Systems*, 24(6):796–808.

Garland, R. (1991). The mid-point on a rating scale: Is it desirable. *Marketing bulletin*, 2(1):66–70.

Garson, G. D. (1991). Interpreting neural-network connection weights. *AI Expert*, 6(4): 46–51.

Garver, M. S. (2003). Best practices in identifying customer-driven improvement opportunities. *Industrial Marketing Management*, 32(6):455–466.

George, D. and Mallery, M. (2003). *SPSS for windows step by step: A simple study guide and reference, 11.0 update.* Allyn and Bacon Boston.

Ghazinoory, S., Abdi, M., and Azadegan-Mehr, M. (2011). SWOT Methodology: A State-of-the-Art Review for the Past, A Framework for the Future. *Journal of Business Economics and Management*, 12(1):24–48.

Glasow, P. A. (2005). Fundamentals of survey research methodology. *MITRE Washington C3 Centre*, 18(April):1–12.

Global Markets Direct (2012). Amazon . com , Inc . - Financial and Strategic Analysis Review. `http://callisto.ggsrv.com` Accessed (April 27, 2013).

Glorfeld, L. W. (1996). A methodology for simplification and interpretation of backpropagation-based neural network models. *Expert Systems with Applications*, 10(1):37–54.

Goff, B. A., Schack, R. W., and Veth, J. (2002). Customer Satisfaction Handbook. Technical report, Charter Oak Group.

Gonçalves, J. R., Pinto, A., Batista, M. J., Pereira, A. C., and BoviAmbrosano, G. M. (2014). Importance-performance analysis: Revisiting a tool for the evaluation of clinical services. *Health (1949-4998)*, 6(5).

Gorener, A. (2012). Comparing ahp and anp: An application of strategic decisions making in a manufacturing company. *International Journal of Business and Social Science*, 3(11):194–208.

Grebennikov, L. and Shah, M. (2013). Monitoring trends in student satisfaction. *Tertiary Education and Management*, 19(4):301–322.

Greene, W. H. (2003). *Econometric analysis*. Prentice Hall.

Grigoroudis, E. and Siskos, Y. (2009). *Customer satisfaction evaluation: Methods for measuring and implementing service quality*, volume 139. Springer Science & Business Media.

Grossman, D. and Domingos, P. (2004). Learning bayesian network classifiers by maximizing conditional likelihood. In *Proceedings of the twenty-first international conference on Machine learning*, page 46. ACM.

Gunn, R. and Williams, W. (2007). Strategic tools: an empirical investigation into strategy in practice in the UK. *Strategic Change*, 16(5):201–216.

Gustafsson, A. and Johnson, M. D. (2004). Determining attribute importance in a service satisfaction model. *Journal of Service Research*, 7(2):124–141.

Håkansson, A. (2013). Portal of research methods and methodologies for research projects and degree projects. In *Proceedings of the International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS)*, pages 67–73. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Han, J. and Kamber, M. (2000a). classification and prediction. In *Data Mining:Concepts and Techniques*, chapter Chapter 6:, pages 380–394. the Morgan Kaufmann Series in data management systems.

Han, J. and Kamber, M. (2000b). Introduction. In *Data Mining:Concepts and Techniques*, pages 3–26. the Morgan Kaufmann Series in data management systems.

Hanson, R. (1992). Determining attribute importance. `http://www.quirks.com/articles/a1992/19921002.aspx?searchID=518894734&sort=5&pg=1`. Accessed (March 5, 2014).

Hasoloan, H. I. P., Yaakob, M. Y., and Yahaya, S. H. (2012). Revisited the importance and performance analysis (IPA) and Kano model for customer satisfaction measurement. *Global Engineers & Technologists Review*, 2(1):22–39.

Hauser, J. (1991). Comparison of importance measurement methodologies and their relationship to consumer satisfaction. Technical report, MIT Marketing Center Working Paper.

Heckerman, D. (1997). Bayesian networks for data mining. *Data mining and knowledge discovery*, 1(1):79–119.

Helms, M. M. and Nixon, J. (2010). Exploring SWOT analysis-where are we now? a review of academic research from the last decade. *Journal of strategy and management*, 3(3):215–251.

Hevner, A. and Chatterjee, S. (2010). Design research in information systems. In *Design research in information systems: Theory and Practice*, volume 22, pages 75–105. Springer Science & Business Media.

Higgins, K. T. (1998). The value of customer value analysis. *Marketing Research*, 10(4): 38–44.

Hill, T. and Westbrook, R. (1997). SWOT analysis: It's time for a product recall. *Long Range Planning*, 30(1):46–52.

Ho, L.-H., Feng, S.-Y., Lee, Y.-C., and Yen, T.-M. (2012). Using modified IPA to evaluate supplier's performance: Multiple regression analysis and DEMATEL approach. *Expert Systems with Applications*, 39(8):7102–7109.

Hom, W. (2000). An overview of customer satisfaction models. Technical report, California Community Colleges.

Hoskin, T. (2012). Parametric and nonparametric: Demystifying the terms. Technical report, Mayo Clinic, Department of Health Science.

Hosseini, S. Y. and Bideh, A. Z. (2013). A data mining approach for segmentation-based importance-performance analysis (SOM–BPNN–IPA): a new framework for developing customer retention strategies. *Service Business*, pages 1–18.

Houben, G., Lenie, K., and Vanhoof, K. (1999). A knowledge-based SWOT-analysis system as an instrument for strategic planning in small and medium sized enterprises. *Decision support systems*, 26(2):125–135.

Hu, H.-Y., Lee, Y.-C., Yen, T.-M., and Tsai, C.-H. (2009). Using BPNN and DEMATEL to modify importance–performance analysis model–a study of the computer industry. *Expert Systems with Applications*, 36(6):9969–9979.

Huang, H.-C. (2012). Research on the influential factors of customer satisfaction and post-purchase behavior for hotels: The multilayer perceptrons neural network approach and logistic regression analysis. *Advances in Information Sciences & Service Sciences*, 4(10):442–450.

Huang, H.-C., Chu, W., and Wang, W.-K. (2007). Strategic performance measurement and value drivers: evidence from international tourist hotels in an emerging economy. *The Service Industries Journal*, 27(8):1111–1128.

Humphrey, A. (2005). SWOT analysis for management consulting. *SRI Alumni Newsletter (SRI International).*

Humphreys, J. (2007). Weakness or opportunity. *MIT Sloan Management Review*, 48(3): 96.

Jamieson, S. et al. (2004). Likert scales: how to (ab) use them. *Medical education*, 38(12):1217–1218.

Johns, R. (2005). One size doesn't fit all: Selecting response scales for attitude items. *Journal of Elections, Public Opinion & Parties*, 15(2):237–264.

Kaiser, J. (2014). Dealing with missing values in data. *Journal of systems integration*, 5(1):42.

Kangas, J., Pesonen, M., Kurttila, M., and Kajanus, M. (2001). A'WOT: Integrating the AHP with SWOT analysis. In *Proceedings of the Sixth International Symposium on the Analytic Hierarchy Process (ISAHP)*, pages 2–4.

Kano, N., Seraku, N., Takahashi, F., and Tsuji, S. (1984). Attractive quality and must-be quality. *The Journal of the Japanese Society for Quality Control*, 14(2):39–48.

Kardes, F. R. and Herr, P. M. (1990). Order effects in consumer judgment, choice, and memory: The role of initial processing goals. *Advances in Consumer Research*, 17:541–546.

Keith, T. (2006). *Multiple regression and beyond*. Pearson Education Boston.

Keith, T. Z. (2014). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.

Kenett, R. and Salini, S. (2012). *Modern analysis of customer surveys: with applications using R*, volume 117. John Wiley & Sons.

Kenny, D. A. (1987). Testing measures of association. In *Statistics for the social and behavioral sciences*, pages 270–291. Little, Brown.

Khanteymoori, A., Homayounpour, M., and Menhaj, M. (2009). A bayesian network based approach for data classification using structural learning. In *Advances in Computer Science and Engineering*, pages 25–32. Springer.

Kim, B.-Y. and Oh, H. (2001). An extended application of importance-performance analysis. *Journal of Hospitality & Leisure Marketing*, 9(3-4):107–125.

Kitcharoen, K. (2004). The importance-performance analysis of service quality in administrative departments of private universities in thailand. *ABAC journal*, 24(3):20–46.

Klicek, B., Oreski, D., and Oreski, S. (2014). Bayesian and neural networks for customer satisfaction optimization. In *Proceedings of the 5th International Conference on Theoretical and Applied Mechanics*, pages 119–129.

Kotler, P. and Keller, K. (2011). Creating long-term loyalty relationships. In *Marketing management*, pages 122–149. Prentice Hall.

Kotsiantis, S. B. (2007). Supervised machine learning: a review of classification techniques. *Informatica (03505596)*, 31(3).

Krešić, D., Mikulić, J., and Kožić, I. (2013). Artificial neural network-based applications in travel and tourism research : A review and case study - working paper -. In *International Critical Tourism Studies Conference*.

Kuechler, B. and Vaishnavi, V. (2005). Design science research in information systems. `www.isworld.org/Researchdesign/drisISworld.htm` Accessed (February 27, 2016).

Kumar, R. (2011). The research design. In *Research methodology: a step-by-step guide for beginners*, pages 93–102. SAGE Publications.

Kurttila, M., Pesonen, M., Kangas, J., and Kajanus, M. (2000). Utilizing the analytic hierarchy process (AHP) in SWOT analysis – a hybrid method and its application to a forest-certification case. *Forest Policy and Economics*, 1(1):41–52.

Lai, I. K. W. and Hitchcock, M. (2015). Importance–performance analysis in tourism: A framework for researchers. *Tourism Management*, 48:242–267.

Lambert, D. M. and Sharma, A. (1990). A customer-based competitive analysis for logistics decisions. *International Journal of Physical Distribution & Logistics Management*, 20(1):17–24.

Larasati, A., DeYong, C., and Slevitch, L. (2011). Comparing neural network and ordinal logistic regression to analyze attitude responses. *Service Science*, 3(4):304–312.

Lawlor, J. E. (2005). The Importance of Strategic Planning. `http://www.practicaldecisions.com/strategic-planning.pdf` Accessed (April 27, 2013).

Lee, K.-l. and Lin, S.-c. (2008). A fuzzy quantified SWOT procedure for environmental evaluation of an international distribution center. *Information Sciences*, 178(2):531–549.

Lee, S.-M. and Abbott, P. A. (2003). Bayesian networks for knowledge discovery in large datasets: basics for nurse researchers. *Journal of Biomedical Informatics*, 36(4):389–399.

Levenburg, N. M. and Magal, S. R. (2005). Applying importance-performance analysis to evaluate e-business strategies among small firms. *E-service Journal*, 3(3):29–48.

Levine, T. R. and Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, 28(4):612–625.

Lihong, Z. and Yanping, Z. (2005). SWOT intellectual analytical system. *Journal of The China Society For Scientific and Technical Information*, 3:008.

Lippmann, R. P. (1987). An introduction to computing with neural nets. *ASSP Magazine, IEEE*, 4(2):4–22.

Luo, F., Zhong, Y., Zhao, S., and Zhang, Y. (2010). Visitors' satisfaction measurement in the national park: A case study of zhangjiajie national forest park. In *International Conference on Management and Service Science (MASS)*, pages 1–4. IEEE.

Mannila, H. (1996). Data mining: machine learning, statistics, and databases. In *Eighth International Conference on Scientific and Statistical Database Systems*, pages 2–9. IEEE.

Martilla, J. A. and James, J. C. (1977). Importance-performance analysis. *Journal of marketing*, 41(1).

Matzler, K., Bailom, F., Hinterhuber, H. H., Renzl, B., and Pichler, J. (2004). The asymmetric relationship between attribute-level performance and overall customer satisfaction: a reconsideration of the importance–performance analysis. *Industrial Marketing Management*, 33(4):271–277.

Matzler, K. and Sauerwein, E. (2002). The factor structure of customer satisfaction: an empirical test of the importance grid and the penalty-reward-contrast analysis. *International Journal of Service Industry Management*, 13(4):314–332.

Matzler, K., Sauerwein, E., and Heischmidt, K. (2003). Importance-performance analysis revisited: the role of the factor structure of customer satisfaction. *The Service Industries Journal*, 23(2):112–129.

Mihelis, G., Grigoroudis, E., Siskos, Y., Politis, Y., and Malandrakis, Y. (2001). Customer satisfaction measurement in the private bank sector. *European Journal of Operational Research*, 130(2):347–360.

Mikulić, J., Paunović, Z., and Prebežac, D. (2012). An extended neural network-based importance-performance analysis for enhancing wine fair experience. *Journal of Travel & Tourism Marketing*, 29(8):744–759.

Mikulić, J. and Prebežac, D. (2012). Accounting for dynamics in attribute-importance and for competitor performance to enhance reliability of BPNN-based importance–performance analysis. *Expert Systems with Applications*, 39(5):5144–5153.

Millán, Á. and Esteban, A. (2004). Development of a multiple-item scale for measuring customer satisfaction in travel agencies services. *Tourism Management*, 25(5):533–546.

Mojaveri, H. S. and Fazlollahtabar, H. (2012). Designing an integrated AHP based fuzzy expert system and SWOT analysis to prioritize development strategies of Iran agriculture. *Revista de Management Comparat International/ Review of International Comparative Management*, 13(1):117–129.

Moran, S., He, Y., and Liu, K. (2009). Choosing the best bayesian classifier: An empirical study. *International Journal of Computer Science*, 36(4):322–331.

Musa, R., Pallister, J., Robson, M., and Daud, N. M. (2010). Application of importance-performance analysis (IPA) to formulate customer satisfaction strategies in the direct sales industry in Malaysia. *Business Strategy Series*, 11(5):277–285.

Nathans, L. L., Oswald, F. L., and Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. *Practical Assessment, Research & Evaluation*, 17(9):2.

Neslin, S. A. (1981). Linking product features to perceptions: Self-stated versus statistically revealed importance weights. *Journal of Marketing Research*, pages 80–86.

New World Encyclopedia, . (2013). Defense mechanism – new world encyclopedia. `http://www.newworldencyclopedia.org/p/index.php?title=Defense_mechanism&oldid=972026` Accessed (June 18, 2016).

Ngai, E. W., Xiu, L., and Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert systems with applications*, 36(2):2592–2602.

Nikolaos, L. (2009). Customers satisfaction in shipping companies under artificial intelligence and multicriteria decision analysis. *Internal Journal of Computers*, 3(4):349–356.

Nordmeyer, B. (n.d.). Advantages & Disadvantages of SWOT Analysis. `smallbusiness.chron.com/advantages-amp-disadvantages-swot-analysis-41398.html` Accessed (March 5, 2013).

Nyberg, J. B., Marcot, B. G., and Sulyma, R. (2006). Using bayesian belief networks in adaptive management. *Canadian Journal of Forest Research*, 36(12):3104–3116.

O'Connell, A. A. (2006). *Logistic regression models for ordinal response variables*. 146. Sage publications.

Offermann, P., Levina, O., Schönherr, M., and Bub, U. (2009). Outline of a design science research process. In *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, pages 7–18. ACM.

Olaru, C. and Wehenkel, L. (1999). CAP tutorial data mining. In *IEEE Computer Applications in Power*, pages 19–25.

Olden, J. D. and Jackson, D. A. (2002). Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks. *Ecological modelling*, 154(1):135–150.

Oliver, R. W. (2000). Real time strategy: The real-time toolbox. *Journal of Business Strategy*, 21(2):7–10.

Oreski, D. (2012). Strategy development by using SWOT-AHP. *Technology Education Management Informatics*, 1(4):283–291.

Osborne, J. and Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8(2):1–9.

Padhy, N., Mishra, D., Panigrahi, R., et al. (2012). The survey of data mining applications and feature scope. *International Journal of Computer Science, Engineering and Information Technology*, 2(3):43–58.

Pai, M.-Y., Chu, H.-C., Wang, S.-C., and Chen, Y.-M. (2013). Ontology-based SWOT analysis method for electronic word-of-mouth. *Knowledge-Based Systems*, 50:134–150.

Pallant, J. (2005). Non-parametric statistics. In *SPSS survival manual: a step by step to data analysis using SPSS*, pages 286–299. Australia: Allen & Unwin.

Palvia, P., Leary, D., Mao, E., Midha, V., Pinjani, P., and Salam, A. (2004). Research methodologies in mis: an update. *Communications of the Association for Information Systems*, 14(1):526–542.

Patel, N. (2003). Multiple Linear Regression in Data mining. In *MIT OpenCourseware, Lecture Notes of 15.062 Data Mining*, pages 1–14. Massachusetts Institute of Technology.

Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3):45–77.

Perucca, G. and Salini, S. (2014). Travellers’ satisfaction with railway transport: A bayesian network approach. *QUALITY TECHNOLOGY AND QUANTITATIVE MANAGEMENT*, 11(1):71–84.

Pezeshki, V., Mousavi, A., and Grant, S. (2009). Importance-performance analysis of service attributes and its impact on decision making in the mobile telecommunication industry. *Measuring Business Excellence*, 13(1):82–92.

Piegorsch, W. W. (1992). Complementary log regression for generalized linear models. *The American Statistician*, 46(2):94–99.

Piercy, N. and Giles, W. (1989). Making SWOT analysis work. *Marketing Intelligence & Planning*, 7(5/6):5–7.

Pinsonneault, A. and Kraemer, K. (1993). Survey research methodology in management information systems: an assessment. *Journal of management information systems*, 10(2):75–105.

Pizam, A. and Ellis, T. (1999). Customer satisfaction and its measurement in hospitality enterprises. *International journal of contemporary hospitality management*, 11(7):326–339.

Pizam, A., Shapoval, V., Ellis, T., Okumus, F., and Okumus, F. (2016). Customer satisfaction and its measurement in hospitality enterprises: a revisit and update. *International Journal of Contemporary Hospitality Management*, 28(1).

Pokryshevskaya, E. and Antipov, E. (2014). A comparison of methods used to measure the importance of service attributes. *International Journal of Market Research*, 56(3): 283–296.

Pryce, G. (2002). Heteroscedasticity: Testing and correcting in spss. *University of Glasgow*.

Rahman, R. M. and Afroz, F. (2013). Comparison of various classification techniques using different data mining tools for diabetes diagnosis. *Journal of Software Engineering & Applications*, 6(3).

Sahu, H., Shrma, S., and Gondhalakar, S. (2008). A Brief Overview on Data Mining Survey. *International Journal of Computer Technology and Electronics Engineering (IJCTEE)*, 1(3):114–121.

Salini, S. and Kenett, R. S. (2009). Bayesian networks of customer satisfaction survey data. *Journal of Applied Statistics*, 36(11):1177–1189.

Samejima, M., Shimizu, Y., Akiyoshi, M., and Komoda, N. (2006). SWOT Analysis Support Tool for Verification of Business Strategy. In *2006 IEEE International Conference on Computational Cybernetics*, pages 1–4. IEEE.

Sampson, S. E. and Showalter, M. J. (1999). The performance-importance response function: Observations and implications. *Service Industries Journal*, 19(3):1–25.

Saunders, M., Lewis, P., and Thornhill, A. (2009). Collecting primary data using questionnaires. In *Research methods for business students*, pages 360–413. Pearson Education.

Seng Wong, M., Hideki, N., and George, P. (2011). The use of importance-performance analysis (IPA) in evaluating Japan's e-government services. *Journal of Theoretical and Applied Electronic Commerce Research*, 6(2):17–30.

Sentas, P., Angelis, L., Stamelos, I., and Bleris, G. (2005). Software productivity and effort prediction with ordinal regression. *Information and Software Technology*, 47(1): 17–29.

Sever, I. (2015). Importance-performance analysis: A valid management tool? *Tourism Management*, 48:43–53.

Shannon, C. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(July):379–423.

Shieh, J.-I. and Wu, H.-H. (2009). Applying importance-performance analysis to compare the changes of a convenient store. *Quality and Quantity*, 43(3):391–400.

Shieh, J.-I. and Wu, H.-H. (2011). Applying information-based methods in importance–performance analysis when the information of importance is unavailable. *Quality & Quantity*, 45(3):545–557.

Shinno, H., Yoshioka, H., Marpaung, S., and Hachiga, S. (2006). Quantitative SWOT analysis on global competitiveness of machine tool industry. *Journal of engineering design*, 17(03):251–258.

Silva, F. and Fernandes, P. O. (2010). Using importance-performance analysis in evaluating institutions of higher education: A case study. In *International Conference on Education and Management Technology*, pages 121–123. IEEE.

Silva, F. and Fernandes, P. O. (2011). Importance-performance analysis as a tool in evaluating higher education service quality: the empirical results of ESTiG (IPB). In *International Business Information Management Association Conference*, pages 306–315.

Silva, F. and Fernandes, P. O. (2012). Empirical study on the student satisfaction in higher education: Importance-satisfaction analysis. *Management*, 293(42):2.

Singh, S., Solanki, A., Trivedi, N., and Kumar, M. (2011). Data mining challenges and knowledge discovery in real life applications. In *3rd International Conference on Electronics Computer Technology (ICECT)*, volume 3, pages 279–283. IEEE.

Siskos, Y. and Grigoroudis, E. (2002). Measuring customer satisfaction for various services using multicriteria analysis. In *Aiding decisions with multiple criteria*, pages 457–482. Springer.

Skok, W., Kophamel, A., and Richardson, I. (2001). Diagnosing information systems success: importance–performance maps in the health club industry. *Information & Management*, 38(7):409–419.

Svozil, D., Kvasnicka, V., and Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and intelligent laboratory systems*, 39(1):43–62.

Tabachnick, B. G., Fidell, L. S., and Osterlind, S. J. (2001). Cleaning up your act screening data prior to analysis. In *Using multivariate statistics*, pages 56–110. Allyn and Bacon Boston.

Tama, B. A. (2015). Data mining for predicting customer satisfaction in fast-food restaurant. *Journal of Theoretical and Applied Information Technology*, 75(1):18–24.

Tamberi, F. (2007). Anomaly Detection. `http://www.cli.di.unipi.it/~tamberi/old/docs/tdm/anomaly-detection.pdf` Accessed (April 27, 2013).

Taplin, R. H. (2012a). Competitive importance-performance analysis of an australian wildlife park. *Tourism Management*, 33(1):29–37.

Taplin, R. H. (2012b). The value of self-stated attribute importance to overall satisfaction. *Tourism Management*, 33(2):295–304.

Tarrant, M. A. and Smith, E. K. (2002). The use of a modified importance-performance framework to examine visitor satisfaction with attributes of outdoor recreation settings. *Managing Leisure*, 7(2):69–82.

Thomas, E. H. and Galambos, N. (2004). What satisfies students? mining student-opinion data with regression and decision tree analysis. *Research in Higher Education*, 45(3):251–269.

Tontini, G., Picolo, J. D., and Silveira, A. (2014). Which incremental innovations should we offer? comparing importance–performance analysis with improvement-gaps analysis. *Total Quality Management & Business Excellence*, 25(7-8):705–719.

Tontini, G. and Silveira, A. (2007). Identification of satisfaction attributes using competitive analysis of the improvement gap. *International Journal of Operations & Production Management*, 27(5):482–500.

Tsaur, S.-H., Chiu, Y.-C., and Huang, C.-H. (2002). Determinants of guest loyalty to international tourist hotels—a neural network approach. *Tourism Management*, 23(4): 397–405.

Turban, E., Sharda, R., Jay E. Aronson, and David N. King. (2008). Neural Networks for Data Mining. In *Business intelligence: A managerial approach.*, chapter Online Chapter, pages W6–1. Upper Saddle River: Pearson Prentice Hall.

Van Ittersum, K., Pennings, J. M., Wansink, B., and Van Trijp, H. C. (2007). The validity of attribute-importance measurement: A review. *Journal of Business Research*, 60(11):1177–1190.

Wilson, R. M. and Gilligan, C. (2005). *Strategic marketing management: planning, implementation and control*. Routledge.

Witten, I. H. and Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Wu, C.-H., Lee, Y.-C., Cheng, Y.-C., and Tasi, S.-B. (2012). The use of importance-performance analysis (IPA) in evaluating bank services. In *9th International Conference on Service Systems and Service Management (ICSSSM)*, pages 654–657. IEEE.

Yay, M. and Akıncı, E. D. (2009). Application of ordinal logistic regression and artifical neural networks in a study of student satistaction. *Cypriot Journal of Educational Sciences*, 4(7):58–69.

Ying, Y. (2010). SWOT-TOPSIS Integration Method for Strategic Decision. In *2010 International Conference on E-Business and E-Government*, pages 1575–1578. IEEE.

Yüksel, İ. and Dagdeviren, M. (2007). Using the analytic network process (ANP) in a SWOT analysis–a case study for a textile firm. *Information Sciences*, 177(16):3364–3382.

Zaerpour, N., Rabbani, M., Gharehgozli, A. H., and Tavakkoli-Moghaddam, R. (2008). Make-to-order or make-to-stock decision by a novel hybrid approach. *Advanced Engineering Informatics*, 22(2):186–201.