A data collection methods psychology essay

3. 0 Introduction

In chapter three, all the data previously collected will be analysed to test against the hypotheses. The result generated will show the relationship between the 4 independent variables and stress. Research design, data collection methods, sample design, research instrument, construct measurement, data processing and data analysis of the study are also discussed in this chapter.

3. 1 Research Design

In this research, the researchers analyse the unit of undergraduates in Kuala Lumpur and Selangor. They are analyzed using cross-sectional studies based on the relationship between stress and undergraduates' stress factors. Survey is conducted to collect data because this research covers a large numbers of target respondents in different geographical areas (Sekaran, 2003). Besides that, research design can be altered during data collection stage of the survey (Groves, 1989). Moreover, survey is cost efficient and quick (Zikmund, 2003). Self-administered questionnaires will be distributed to the undergraduates in Kuala Lumpur and Selangor and collected immediately upon completion.

3. 2 Data Collection Methods

Primary and secondary data firm the basis of this research which provides helpful and appropriate information in this study. Primary data is gathered from the respondents through questionnaires. On the other hand, researchers collected information from online information databases as the sources of secondary data.

3. 2. 1 Primary Data

Malhotra (2007) defined primary data as the information collected for the first time specifically for a research project. These data collected originates from first-hand experience called as primary data (Zikmund, 2003).

Questionnaire is one of the standardized and familiar methods of most of the researchers where data are obtained without bias. As mentioned in Zikmund (2003), questionnaire is the main tool to gather opinions, perception, direct expression and descriptions from respondents and make the data more comparable and amenable for analysis. Hence, self-administered questionnaire is researchers' choice to collect primary data in order to increase the dependability and simplicity of data in this research.

According to Ministry of Higher Education (2010), most universities and colleges are located in Kuala Lumpur and Selangor (Appendix A). Hence, primary data was collected from respondents through 250 sets of questionnaires which were distributed randomly by hand in Kuala Lumpur and Selangor.

According to Ministry of Higher Education (2010), most of the undergraduates are located in selected target universities. The targeted universities included public universities such as Universiti Islam Antarabangsa Malaysia (UIAM), Universiti Malaya (UM), Universiti Putra Malaysia (UPM), Universiti Teknologi MARA (UiTM), and private universities such as Universiti Kuala Lumpur (UniKL), Unisel Universiti Selangor (UNISEL) and Universiti Tun Abdul Razak (UNITAR). Around 10 minutes are spent by each target respondent to complete the questionnaires which are immediately collected under the supervision of researchers.

3. 2. 2 Secondary Data

According to Sekaran (2003), secondary data is the information from previously published or compiled sources by researchers for certain purposes which are related to historical data to support our research study. Secondary data included scholarly books, articles in journals, textbooks, government and private sources and online information database. Furthermore, obtaining secondary data is easier, cost and time effective but such data may be obsolete and inaccurate to fulfil researchers' requirements in of present investigation compared to primary data (Zikmund, 2003).

From the beginning of research, secondary data is used to review the similar literature study from online information databases such as EBSCOhost, JSTOR and ProQuest and other databases provided by Universiti Tunku Abdul Rahman (UTAR)'s website. In addition, Internet search engine such as Google Scholar is also used by researchers to develop other related material in this research.

3. 3 Sampling Design

3. 3. 1 Target Population

This research focuses on the study of stress factors among undergraduates in Kuala Lumpur and Selangor. Therefore, the target population will be the undergraduates in Kuala Lumpur and Selangor.

3. 3. 2 Sampling Frame and Sampling Location

There is no sampling frame in this research and the questionnaires will be distributed at universities in Kuala Lumpur and Selangor. Since most of the universities (Ministry of Higher Education, 2010) are located in Kuala Lumpur and Selangor respectively, the researchers have chosen these two areas to conduct their study. The location of Malaysian universities will be showed in Appendix A.

3. 3. 3 Sampling Elements

The sampling elements are the undergraduates studying in universities in Kuala Lumpur and Selangor. They will be randomly selected for questionnaires survey as suicide cases for undergraduates in Malaysia are getting serious (Kok et al., 2011).

3. 3. 4 Sampling Techniques

Besides that, quota sampling will be targeted at public and private universities and at 64% (160 questionnaires) and 36% (90 questionnaires) respectively since the figures provided by the Ministry of Higher Education (2008) indicated that the numbers of undergraduates in public and private universities are 247, 881 and 140, 669 students respectively in 2007.

3. 3. 5 Sampling Size

To determine the regression of sample size, Green (1991) indicated that N> 50+8m were N is the sample size and m is the number of independent variables. Since, there are 4 independent variables in this study, therefore m is equal to 4. Hence, n= 50+8(4)= 82, which means that the sampling size required should be more than 82 target respondents. However, according to

Denscome (2007), usually the research sampling should involve between 30 to 250 cases. Therefore, researchers will distribute 250 questionnaires to the target respondents for more accuracy in results.

3. 4 Research Instrument

Questionnaires method is an efficient technique for collecting data in which the respondents need to answer a set of questions presented by the researcher in order to obtain information (Zikmund, Babin & Griffin, 2010). In addition, Sekaran (2003) defined that questionnaire is a pre-planned set of questions for respondents to answer which enables the researches to know exactly what is needed and how to evaluate the variables of interest.

Normally, the self-administered questionnaires can be personally distributed to the target respondents by hand. Besides, in order to ensure that the questionnaire's components and measurement scales are properly designed and are easily understood by respondents, a pilot test had to be conducted.

3. 4. 1 Questionnaires Design

Self-administered questionnaires are chosen to conduct the study in this research. Researchers distribute questionnaires to respondents and collect them after the response. This method is used because the direct contact with respondents. The whole process is done fast, efficiently, ensuring richness of data and involving a large numbers of respondents (Rattray & Jones, 2005). Questionnaires method is a better way to make comparison and judgment then interviewing respondents (Field, 2003).

Besides, closed-ended questions are used to construct standardised questionnaires in order to reduce bias and assist researchers in coding the information and simplify the data compilation process. English is international language and widely used in Malaysia, hence it is used in designing the questionnaires.

3. 4. 2 Pilot Test

According to Zikmund (2003), pilot test is a small assessment designed to investigate the logistics and obtain information prior to a larger study in order to improve the questionnaire's quality and efficiency. The pilot test can disclose the deficiencies of the questionnaires before proceeding to large scale studies.

Researchers have distributed 30 questionnaires to the target respondents from UTAR to test the reliability and internal consistency of the questionnaires in order to ensure whether it is comprehensive and able to implement the real survey.

3. 5 Constructs Measurement

3. 5. 1 Scale Measurement

In order to enrich the description in research, the most common way is to perform measurement which is a process of assigning numbers to the research objects (Brown, 2011). There are four types of measurement scale which includes nominal, ordinal, interval and ratio scale. Each of these scales has its own functions and helps to differentiate the variables. There are three types of measurement scales used in this research which are nominal, ordinal and interval.

Nominal scale is categorizing by assigning number labelling to objects based on its nature, for example measurement on genders. Subsequently, ordinal scale is categorizing data by ranking or order, where datum are expressed orderly such as firstly and followed by. Interval scale is to show order of things, however the equal intervals between objects represents equal differences. Likert scale is a form of interval scale where it is the most suitable to describe opinion measurement (Brown, 2011).

In this research, a five points Likert scale is used by dividing the interval scale into five points to distinguish the degree of agreement and disagreement in each statement. The pointer starts from number 1 to number 5 which represent strongly disagree, disagree, neutral, agree and strongly agree.

Moreover, the questionnaires are divided into Section A, Section B and Section C. Section A consists of demographic profile of the target respondents, the questions of which are designed to collect their demographic information such as gender, age, race, and so on. Section B investigates the factors of stress of the undergraduates using the 5 point Likert scale. The scale is used to examine the agreement of the four independent variables that influence students' perceptions about their selection of stress factors of stress that contributes to stress. In Section C, questions on current level of overall stress of undergraduates are being used to examine how significant the factors of stress will influence their stress by using the 5 point Likert scale.

3. 6 Data Processing

The raw data collected from survey questionnaire need to be analysed in order to provide useful information in resolving the research questions. All data collected are needed to undergo a series process before it can serve the purpose of being analyse. This data processing is a process that includes questionnaire checking, data editing, coding, transcribing and cleaning before the analysis made (Zikmund, 2003).

3. 6. 1 Questionnaire Checking

Questionnaire checking is a process that identifies, corrects, and resolves any errors or problems which may occur in the questionnaires. The questionnaires are being checked and amended numerously to ensure error free, followed by an approval from the supervisor to ensure the questions asked are of highest quality, appropriate and comprehensible by respondents.

3. 6. 2 Data Editing

Data editing is performed to ensure high accuracy of the data. Missing data is being discarded while extremely contrasting data is being removed to ensure consistency and to prevent any undesirable results outcome.

3. 6. 3 Data Coding

Statistical Package for Social Science (SPSS) program requires all data to be in code form, so the data are being coded before being keyed into. Therefore, prior to the questionnaire distribution, data coding are being done by assigning code to each option of answers in the questionnaire.

In the research questionnaire, respondents are required to fill up their demographic profiles. Coding was being made on each question with the first answer on each question coded as 1 while the second answer is coded as 2 and so on. For example, in Section A the first question is on genders and the answer is either male or female. The code for male is 1 while female is 2.

Subsequently, Section B requires respondent's opinion regarding five different stress factors in their daily life. The five different stress factors also known as independent variables include financial issues, academic issues, social support issues and time management issues. As for Section C, the questionnaire focuses questions regarding stress level for the past four weeks of respondents' life which are the dependent variables in the research. Both sections require respondents to respond to their level of agreement and disagreement by using the 5 Likert scale. The 5 levels of agreement are: strongly agree, agree, neutral, disagree and strongly disagree. The code for strongly agree is 5, agree is coded as 4, neutral is coded as 3, disagree coded as 2 and lastly strongly disagree is coded as 1.

3. 6. 4 Data Transcribing

After all questionnaires are retrieved, the answers to the questionnaires are converted into code data which are subsequently being keyed manually into Statistical Package for Social Science (SPSS) for further analysis.

3. 6. 5 Data Cleaning

With the help of Statistical Package for Social Science (SPSS), the checking process is enhanced ensuring consistent data. Any extreme data is easily identified and replaced with a more consistent data. Furthermore, missing data can to be detected thus ensuring no data is missed out during transcribing process.

3. 7 Data Analysis

The Computer program that is being used to analyze questionnaire data is Statistical Package for the Social Science program (SPSS) which includes descriptive analysis and inferential analysis. To ensure that target respondents have a clear understanding of the questionnaires, a pilot test will be conducted before distributing them to respondents.

3. 7. 1 Descriptive Analysis

Descriptive analysis consist of measurement of mode, mean, frequency, range, standard deviation and variance. The researchers use it to analyses the demographic data in the questionnaires (Han, 2010). The research data will be summarized by calculating the average, frequency, range and standard deviation. Consequently, mode, mean, frequency, range, standard deviation and variance are calculated based on the interval scale of 4 independent variables and dependent variable.

3. 7. 2 Reliability Test

A reliability test will be conducted to make sure the measures are free from error. Cronbach's Alpha reliability test model is selected to measure the reliability of the questionnaires and 0 to 1 is used to assess the reliability of data collected. According to Nunnally and Bernstein (1994)'s study (as cited in Krishnan & Ramasamy, 2011), an alpha of 0. 7 is an acceptable reliability coefficient level and acts as the minimum requirement for validity.

3. 7. 3 Normality Test

Normality test is used for the purpose of checking the normality of the data set (Archila, 2010). When the sampling size is more than 100, Kolmogorov-

Smirnov test is used to test the normality of the sample data. The p-value will indicate whether the data is significantly different. When the p-value is more than 0. 05 the data set could be assumed as normal (Islam, 2011).

3. 7. 4 Inferential Analysis

Inferential analysis using Pearson’s correlation and multiple is used to test whether the hypothesis is substantial.

3. 7. 4. 1 Pearson Correlation

Pearson’s correlation measure the degree of interrelation between the dependent and independent variables (Mudelsee, 2003). Doss and Kamery (2006) mentioned that Pearson’s correlation is used to test whether the relationship between variables exists. Hence, it is being used to determine whether the relationship between the financial issues, academic issues, social support issues, time management issues and stress faced by undergraduates exist. The relationship depends on Pearson’s correlation outcome. The range of outcome is between -1 to 1. Positive outcome indicates a positive relationship between variables and vice versa. 0 outcome indicates no relationship between both variables (Doss et al., 2006).

3. 7. 4. 2 Multiple Regressions

Multiple regression analysis will be used for identifying which independent variable influences dependent variable the most (Hair, Babin, Money & Samuel, 2003). Multiple regressions are able to examine the relationship between dependent and independent variables (Orme & Buehler, 2001). Hence, it is used to evaluate the relationship between financial issues, academic issues, social support issues, time management issues and stress faced by undergraduates.

Since the researcher used interval data measurement scale on dependent and independent variables where both variables are parametric, the general equation for the multiple regressions is (Yount, 2006).

Therefore, the multiple regressions equation in this research is developed as

3. 8 Conclusion

This chapter consists of research design, data collection methods, sampling design, research instrument, constructs measurement, data processing and data analysis which are used to collect data for analysis. In chapter four, the researchers will examine the pattern of and analysis of the result which are collected from relevant target respondents.