

**MINING STUDENTS' PERFORMANCE IN UPSR USING
STATISTICS AND NEURAL NETWORKS**

NOR FAZIDA ABD. RAHMAN

**UNIVERSITI UTARA MALAYSIA
2008**

GA
2008
2008

MINING STUDENT'S PERFORMANCE IN UPSR USING STATISTICS AND
NEURAL NETWORKS

A thesis submitted to the Division of Applied Sciences, College of Arts and Sciences in
partial fulfillment of the requirements for the degree Master of Science (Information and
Communication Technology), Universiti Utara Malaysia.

by
Nor Fazida Abd Rahman

© Nor Fazida Abd Rahman, 2008. All rights reserved.



**KOLEJ SASTERA DAN SAINS
(College of Arts and Sciences)
Universiti Utara Malaysia**

**PERAKUAN KERJA KERTAS PROJEK
(Certificate of Project Paper)**

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certify that)

NOR FAZIDA ABD. RAHMAN

calon untuk Ijazah
(candidate for the degree of) **MSc. (ICT)**

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project paper of the following title)

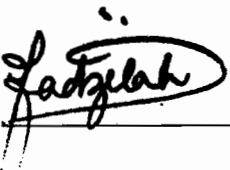
**MINING STUDENTS PERFORMANCE IN UPSR USING
STATISTICS AND NEURAL NETWORKS**

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project paper)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan
dan meliputi bidang ilmu dengan memuaskan.
(that the project paper acceptable in form and content, and that a satisfactory
knowledge of the field is covered by the project paper).

Nama Penyelia Utama
(Name of Main Supervisor): **ASSOC. PROF. FADZILAH SIRAJ**

Tandatangan
(Signature)

: 

Tarikh
(Date)

: 28/5/2008

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence by the Dean of College of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to

Division of Applied Sciences
College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman.

ABSTRACT (BAHASA MELAYU)

Pencapaian pelajar di dalam peperiksaan menjadi penanda aras yang penting dalam menentukan kualiti pendidikan di Malaysia. Data-data peperiksaan telah dikumpul mulai ujian-ujian bulanan yang telah dijalankan sehingga ke percubaan UPSR untuk diuji dengan peperiksaan UPSR yang sebenar. Ini juga melibatkan data-data lain yang berkaitan seperti latar belakang keluarga dan maklumat berkenaan persekolahan pelajar. Data mentah diproses serta dianalisa menggunakan kaedah Statistik. Kaedah Statistik memberikan analisis yang bernilai kepada model pencapaian. Kemudian, kombinasi unit input, unit tersembunyi dan unit output diuji untuk meramal pencapaian sebenar pelajar. 5 model diuji berdasarkan 5 matapelajaran teras untuk mengaitkannya dengan faktor-faktor lain menggunakan analisis diskriptif. Justeru, hubungan itu dikaji dengan teliti untuk mengukuhkan model jangkaan. Model akhir kes ini menggunakan kadar pembelajaran 0.1, kadar momentum 0.1, fungsi aktivasi Sigmoid, kriteria penghentian pembelajaran 100 pusingan dengan senibinanya 13 unit input, 2 unit tersembunyi dan 5 unit output. Keputusan yang telah diperolehi menunjukkan Rangkaian Neural mempunyai potensi yang tinggi untuk meramal pencapaian pelajar di masa hadapan.

ABSTRACT (ENGLISH)

Academic performance has become an important evidence of determining the quality in Malaysia's education system. The examination data is collected on the previous students' examinations yet to be tested for their coming UPSR. The other related data such as family background and schooling information are also involved. The raw data is preprocessed and analyzed using statistical method. The results from the statistical analysis indicate the significant contribution of these attributes to the achievement model. The combinations of input variables, hidden layer and output nodes are explored to predict the students' performance. Five models are constructed based on five subjects to relate them with other factors for the purpose of descriptive analysis. The relationship between examination results and other factors are investigated thoroughly to enhance the prediction model. The performance model obtained in this study uses parameters such as; learning rate 0.1, momentum rate 0.1, Sigmoid activation function, 100 epoch learning stopping criteria with its architecture, 13 inputs unit, 2 hidden units and 5 output units. The result indicates that Neural Networks has high potential to be used in predicting students' performance.

ACKNOWLEDGMENT

A million of thanks are due first to my supervisor, Associate Professor Fadzilah Siraj for patiently navigating and generously sharing her rich source of knowledge with me. A good teacher always guides her students in no matter situation they are.

Equally thankful to my family for all the love and support by giving me a beautiful place in their heart. Fortunately born in teachers' family, education is the top priority to me. In a process towards seeking for excellent and meaningful education, knowledge brings its own values. Heartfull of thanks also flow to my sister, Ihwana As'ad for her loving and caring towards me.

A warm thanks to my colleagues; SK Tan Sri Ghazali Jawi, SK Basia Lama, SK Sri Adika Raja and SK Mahkota Sari, Gerik, Perak for allowing me to use their schooling dataset as experimental tool. These research findings are specially dedicated for the sake of my sacred 'children'.

I would like to express my eternal gratitude to my dearest partner, Mohd Shukri Mohd Adam for his kindness and for everlasting bond of fellowship. Not least, my deep appreciation goes to all who lend a hand in realizing this project.

Nor Fazida Abd Rahman

Division of Applied Sciences

College of Arts and Sciences

2008

TABLE OF CONTENTS

Permission to Use	i
Abstract (Bahasa Melayu)	ii
Abstract (English)	iii
Acknowledgment	iv
Table of Contents	v
List of Figures	ix
List of Tables	x
List of Abbreviations	xi

CHAPTER 1: INTRODUCTION

1.1 Overview	1
1.2 Problem Statement	5
1.3 Objectives of The Study	6
1.4 Research Question.....	7
1.5 Scope of The Study	7
1.6 Significance of The Study	7
1.7 Thesis Organization	8

CHAPTER 2: LITERATURE REVIEW

2.1 Data Mining	9
2.1.1 Descriptive Analysis.....	10
2.1.2 PredictiveAnalysis.....	11

2.2 Students' Performance.....	17
2.3 Discussion	19

CHAPTER 3: METHODOLOGY

3.1 Methodology	20
3.1.1 Data Collection	21
3.1.2 Data Description	22
3.1.3 Data Propressing	23
3.1.4 Separate into Training and Test Sets	26
3.1.5 Define Networks Structure	26
3.1.6 Select Learning Algorithm	27
3.1.7 Set Parameter Values	29
3.1.8 Transform Data to Networks Inputs	30
3.1.9 Start Training and Revise Weights Stop Point	30
3.1.10 Stop and Test	32
3.1.11 Implementation	32

CHAPTER 4: RESULTS AND DISCUSSION

4.1 First Dimension Analysis	33
4.2 2 Dimension Analysis	89
4.3 Initial Study	101
4.4 Relationship Between Each Factors.....	147

4.5 Comparison Methods.....149

CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion150

5.2 Recommendations 151

REFERENCES

References152

APPENDICES

Appendix A: Raw Schooling Data

Appendix B: Numerical Data (xls)

LIST OF FIGURES

	PAGE
Figure 1.1 Interface of SMM	2
Figure 1.2 Interface of SAPR 16	2
Figure 2.1 Single layer perceptron network	11
Figure 3.1 Flow diagram of the development process of ANN	21
Figure 3.2 The background form	22
Figure 3.3 Data allocation	26
Figure 3.4 The architecture of multilayer perceptrons	27
Figure 3.5 MLP algorithm	28
Figure 3.6 Learning rate, momentum rate and epoch	28
Figure 3.7 Training and test process	31
Figure 3.8 Flow diagram process for multilayer perceptrons in Neural Connection	31
Figure 3.9 The main SPSS window	32
Figure 4.1 Gender distribution	34
Figure 4.2 Relationship between gender and BM1	35
Figure 4.3 Relationship between gender and BM2	36
Figure 4.4 Relationship between gender and BI	37
Figure 4.5 Relationship between gender and MT	38
Figure 4.6 Relationship between gender and SC	39
Figure 4.7 Attendance distribution	41

Figure 4.8	Relationship between attendance and BM1	42
Figure 4.9	Relationship between attendance and BM2	43
Figure 4.10	Relationship between attendance and BI	44
Figure 4.11	Relationship between attendance and MT	45
Figure 4.12	Relationship between attendance and SC	46
Figure 4.13	Co-curricular activities distribution	47
Figure 4.14	Relationship between co-curricular and BM1	48
Figure 4.15	Relationship between co-curricular and BM2	49
Figure 4.16	Relationship between co-curricular and BI	50
Figure 4.17	Relationship between co-curricular and MT	51
Figure 4.18	Relationship between co-curricular and SC	52
Figure 4.19	Position in family distribution	53
Figure 4.20	Relationship between position in family and BM1	55
Figure 4.21	Relationship between position in family and BM2	57
Figure 4.22	Relationship between position in family and BI	58
Figure 4.23	Relationship between position in family and MT	60
Figure 4.24	Relationship between position in family and SC	61
Figure 4.25	Father's job distribution	63
Figure 4.26	Relationship between father's job and BM1	64
Figure 4.27	Relationship between father's job and BM2	65
Figure 4.28	Relationship between father's job and BI	66
Figure 4.29	Relationship between father's job and MT	67
Figure 4.30	Relationship between father's job and SC	69

Figure 4.31	Mother's job distribution	70
Figure 4.32	Relationship between mother's job and BM1	71
Figure 4.33	Relationship between mother's job and BM2	72
Figure 4.34	Relationship between mother's job and BI	73
Figure 4.35	Relationship between mother's job and MT	75
Figure 4.36	Relationship between mother's job and SC	76
Figure 4.37	Family income percapita distribution	78
Figure 4.38	Relationship between family income percapita and BM1	79
Figure 4.39	Relationship between family income percapita and BM2	80
Figure 4.40	Relationship between family income percapita and BI	82
Figure 4.41	Relationship between family income percapita and MT	83
Figure 4.42	Relationship between family income percapita and SC	85
Figure 4.43	Relationship between gender and attendance	87
Figure 4.44	Relationship between gender and co-curricular	89
Figure 4.45	Relationship between same parents' job and attendance	97
Figure 4.46	Relationship between father's job with mother's job as housewife and attendance	99
Figure 4.47	Single parent and attendance	100

LIST OF TABLES

		PAGE
Table 3.1	The raw dataset	23
Table 3.2	The dataset variables	24
Table 3.3	The numeric coding	25
Table 3.4	The numerical data	25
Table 3.5	The attributes descriptions	30
Table 4.1	Frequency data	33
Table 4.2	Gender frequency data	35
Table 4.3	Gender and BM1 crosstabulation	35
Table 4.4	Gender and BM2 crosstabulation	36
Table 4.5	Gender and BI crosstabulation	37
Table 4.6	Gender and MT crosstabulation	38
Table 4.7	Gender and SC crosstabulation	39
Table 4.8	Attendance frequency table	40
Table 4.9	Attendance and BM1 crosstabulation	41
Table 4.10	Attendance and BM2 crosstabulation	43
Table 4.11	Attendance and BI crosstabulation	44
Table 4.12	Attendance and MT crosstabulation	45
Table 4.13	Attendance and SC crosstabulation	46
Table 4.14	Co-curricular frequency table	47
Table 4.15	Co-curricular and BM1 crosstabulation	48
Table 4.16	Co-curricular and BM2 crosstabulation	49

Table 4.17	Co-curricular and BI crosstabulation	50
Table 4.18	Co-curricular and MT crosstabulation	51
Table 4.19	Co-curricular and SC crosstabulation	52
Table 4.20	Position in family frequency table	53
Table 4.21	Position in family and BM1 crosstabulation	55
Table 4.22	Position in family and BM2 crosstabulation	56
Table 4.23	Position in family and BI crosstabulation	58
Table 4.24	Position in family and MT crosstabulation	59
Table 4.25	Position in family and SC crosstabulation	61
Table 4.26	Father's job frequency table	62
Table 4.27	Father's job and BM1 crosstabulation	63
Table 4.28	Father's job and BM2 crosstabulation	65
Table 4.29	Father's job and BI crosstabulation	66
Table 4.30	Father's job and MT crosstabulation	67
Table 4.31	Father's job and SC crosstabulation	68
Table 4.32	Mother's job frequency table	70
Table 4.33	Mother's job and BM1 crosstabulation	71
Table 4.34	Mother's job and BM2 crosstabulation	72
Table 4.35	Mother's job and BI crosstabulation	73
Table 4.36	Mother's job and MT crosstabulation	74
Table 4.37	Mother's job and SC crosstabulation	76
Table 4.38	Family income percapita frequency table	77
Table 4.39	Family income percapita and BM1 crosstabulation	78

Table 4.40	Family income percapita and BM2 crosstabulation	80
Table 4.41	Family income percapita and BI crosstabulation	81
Table 4.42	Family income percapita and MT crosstabulation	83
Table 4.43	Family income percapita and SC crosstabulation	85
Table 4.44	Gender and attendance frequency table	87
Table 4.45	Gender and attendance crosstabulation	87
Table 4.46	Gender and co-curricular activities frequency table	88
Table 4.47	Gender and co-curricular crosstabulation	88
Table 4.48	Father's job, mother's job and attendance frequency table	90
Table 4.49	Father's job, mother's job and attendance crosstabulation	91
Table 4.50	Father's job, mother's job and co-curricular frequency table	93
Table 4.51	Father's job, mother's job and co-curricular crosstabulation	94
Table 4.52	Same parents' job and attendance frequency table	96
Table 4.53	Same parents' job and attendance crosstabulation	96
Table 4.54	Father's job with mother's job as housewife and attendance frequency table	98
Table 4.55	Father's job with mother's job as housewife and attendance crosstabulation	98
Table 4.56	Single parent frequency table	100
Table 4.57	Single parent and attendance crosstabulation	100
Table 4.58	Academic factors frequency table	101
Table 4.59	BM1 model	102
Table 4.60	BM1 with gender model	103

Table 4.61	BM1 with attendance model	104
Table 4.62	BM1 with participation in co-curricular activities model	105
Table 4.63	BM1 with position in family model	106
Table 4.64	BM1 with father's job model	107
Table 4.65	BM1 with mother's job model	108
Table 4.66	BM1 with family income model	109
Table 4.67	BM1 with all attributes model	110
Table 4.68	BM2 model	111
Table 4.69	BM2 with gender model	112
Table 4.70	BM2 with attendance model	113
Table 4.71	BM2 with co-curricular model	114
Table 4.72	BM2 with position in family model	115
Table 4.73	BM2 with father's job model	116
Table 4.74	BM2 with mother's job model	117
Table 4.75	BM2 with family income model	118
Table 4.76	BM2 with all attributes model	119
Table 4.77	BI model	120
Table 4.78	BI with gender model	121
Table 4.79	BI with attendance model	122
Table 4.80	BI with co-curricular model	123
Table 4.81	BI with position in family model	124
Table 4.82	BI with father's job model	125
Table 4.83	BI with mother's job model	126

Table 4.84	BI with family income model	127
Table 4.85	BI with all attributes model	128
Table 4.86	MT model	129
Table 4.87	MT with gender model	130
Table 4.88	MT with attendance model	131
Table 4.89	MT with co-curricular model	132
Table 4.90	MT with position in family model	133
Table 4.91	MT with father's job model	134
Table 4.92	MT with mother's job model	135
Table 4.93	MT with family income model	136
Table 4.94	MT with all attributes model	137
Table 4.95	SC model	138
Table 4.96	SC with gender model	139
Table 4.97	SC with attendance model	140
Table 4.98	SC with co-curricular model	141
Table 4.99	SC with position in family model	142
Table 4.100	SC with father's job model	143
Table 4.101	SC with mother's job model	144
Table 4.102	SC with family income model	145
Table 4.103	SC with all attributes model	146
Table 4.104 (a)	Relationship between subjects and attributes	147
Table 4.104 (b)	Relationship between subjects and attributes	147
Table 4.105	Comparison between NN and regression	149

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Networks
BI	Bahasa Inggeris
BM1	Bahasa Melayu 1
BM2	Bahasa Melayu 2
BP	Backpropagation Algorithm
FF	Feed Forward Algorithm
MLP	Multilayer Perceptrons
MT	Mathematics
NC	Neural Connection
NN	Neural Networks
SAPR 16	Sistem Analisis Peperiksaan
SC	Science
SK	Sekolah Kebangsaan
SMM	Sistem Maklumat Murid
PKSR	Penilaian Kemajuan Berasaskan Sekolah
UPSR	Ujian Penilaian Sekolah Rendah

CHAPTER 1

INTRODUCTION

This section discusses the background of the study that consists of general overview on the education system in Malaysia, some brief description on the problem statements, objectives, scope and significance of this study. Finally, this section presents the thesis organization describing the structure of this report.

1.1 Overview

Realizing the importance of having a national database that comprises of graduate data, the Ministry of Higher Institutions of Learning has taken an initiative to form such a data warehouse. At the school level, 'Sistem Maklumat Murid' (SMM) is used for several years to collect information on students' background. SMM consists mainly information about demographic and co-curricular but not including their academic results. To support SMM, 'Analisis Peperiksaan 16' (SAPR 16) is used for the purpose of analyzing the students' results and ranks their performance accordingly. Even then, the analysis involves the use of descriptive approach of data mining technique. Little attempt has been made to include the use of forecasting model in managing the students'

Students' performance plays major role in determining the quality of our education system. Ujian Penilaian Sekolah Rendah (UPSR) is a public examination compulsory to be taken by Year 6 students in Malaysia. The performance gap is not only a school and classroom issue but also a national issue that must be addressed properly (Terry *et al.*, 2005). As young generation, their performance in school is important to determine which schools that these children are streamed to further their studies either to daily school, boarding school, semi boarding school or religious school. In effect, this will influence their career path in the future. Factors such as gender (Cripps, 1996; Bae *et al.*, 1997), attendance (Hayek *et al.*, 1999; Terry *et al.*, 2005), co-curricular activities (Henchey & Norman, 2001) and family background (Gibson & Margaret, 2003; Henchey & Norman, 2001) may influence their performance.

A study conducted by Erbe and Mach (2000) shows that the other school variables (school climate, students' background and others) make a significant contribution to the variance in Mathematics achievement in Chicago Public Secondary Schools. This indicates that examination result is not the only factor affecting students' performance.

Participation in co-curricular activities may also contribute into healthy communication and relationship among others. Actively involved students in co-curricular activities bring more benefits. It helps in expanding the idea and developing supportive relationship between teachers and their friends. A research by Gibson and Margaret (2003) at Hillside High School reveals that a sense of caring community in school is necessary to motivate the students. They are also encouraged to involve in school and

community activities as part of their academic motivation and support activities. Based on the research finding from Henchey and Norman (2001) at Canada High School, extra curricular activities in schools also take on special significance as integral part of learning experiences that school provides.

Family background like parents' job and socio economic status influence the life style. Moreover, it may affect student's achievement in the way of their capability to provide extra classes to their children. After school tutoring is also found as one of the factor that contributes to the students' achievement at Hillside High School (Gibson & Margaret, 2003). The students from high socio economic status tend to get outside good quality extra classes compared to low income family. Fortunately, low income background students tend to be concentrated in certain schools, which can depress achievement for all children at Indiana High School (Terry *et al.*, 2005). According to Henchey and Norman (2001), schools with students from high income homes tend to do better than schools with students from low income homes at Canada High School.

Students' gender becomes one of the main factors that influence the performance at Tennessee State University (Cripps, 1996). Female students tend to be hard work students and give full attention in every inch of their works. Male students usually actively involved in sport and pay less attention to their study. A study on "Women in Mathematics and Science" by Bae *et al.* (1997) found that boys and girls have similar Mathematics and Science proficiency scores at age 9 but the gap increased from time to time. In general, both gender had similar average Mathematics achievement but boys

had higher average Science achievement than girls. Additional research by Hayek *et al.*, (1999) examined that college activities contributed to continuous learning which would help enhancing students' performance.

A research at University of Illinois by Hostetler (1996) found that students' previous results had a strong correlation with their future results on predicting students' success in an introductory programming course. A research by Terry *et al.* (2005) at Indiana High School found that college attendance rates between African American and Hispanic students contributed to their performance.

To predict students' performance, forecasting techniques were employed either from statistical perspectives or computer science perspectives. Data mining techniques such as Neural Networks (NN) are commonly used to build forecasting model (Tsoukalas & Uhrig, 1997), (Huisken & Coffa, 2000) and (Khoa, Sakakibara & Nishikawa, 2006). On the other hand, regression analysis is generally used on the predicting students' success in programming course and predicting malignancy of ovarian tumors (Hostetler 1996), (Lu *et al.* 2001) and (Butcher & Muth 1985).

1.2 Problem Statement

Although SMM and SAPR 16 are currently being used in schools in Perak, these systems are used independently. As such, the information gathered from two systems only used for obtaining total raw marks and percentages. Hence, there is more room for data exploration and information seeking from the two databases. Therefore, this study

attempts to mine the students' background information with regard to their performance. In particular, the forecasting model will be obtained (if possible) and some explanation is sought using Statistical approach.

1.3 Objectives of The Study

The general objective of this study is to mine student's performance in UPSR using Statistics and NN method. No studies have yet been done on the application of Statistics and NN for UPSR performance in Malaysia. The specific objectives are as follows:

- a. To collect and preprocess previous students' data.

The process of collecting and preprocessing previous data are done for the purpose of data cleaning, data integration and data transformation.

- b. To mine students' data in order to build NN performance model.

The best five models for five subjects are selected to represent the best algorithm, activation function, momentum rate and learning rate that produced the highest accuracy under training phase.

- c. To evaluate the performance of NN model.

Due to the strength of each technique, both techniques (Statistics and NN) will be used in the analysis of the students' performance. The final NN models are used to show the percentage of accuracy for each factor.

1.4 Research Question

Based on the research objectives, the research question is:

- How do the academic and demographic factors influence students' performance in UPSR?

1.5 Scope of The Study

The students' examination and background dataset presented in this study consists of 420 Year 6 students of SK Tan Sri Ghazali Jawi, SK Basia Lama, SK Sri Adika Raja and SK Mahkota Sari , Gerik, Perak for the year 2006 until 2007 released by District Education Office. It is focused on Year 6 students only because they have to sit for UPSR. The examination dataset is taken from 'SAPR 16'. Grade is encoded using standard parameter set by Ministry of Education (A = 80-100, B = 60-79, C = 40-59, D = 20-39, E = 0-19). The attendance and family background dataset are taken from 'Sistem Maklumat Murid' (SMM). The other data such as co-curricular activities are obtained through the survey form distributed to the students.

1.6 Significance of The Study

The study contributes specially to the education field nowadays including students, parents and teachers. The related parties especially teachers can identify the weaknesses and plan for remedial strategies such as set up more extra classes, organize more useful co-curricular activities and make collaboration with their parents to guide their school attendance. The most important finding from this research is to provide the better

technique in mining students' performance for the purpose of improving any factors that contribute to their future performance.

As the schools in Perak use two systems provided by State Education Department and the Ministry Education of Malaysia, the findings from the study could highlights some useful information mined from the two systems. If the information found prove to be useful to the school in particular and the Ministry of Education at large, a new system needs to be developed so that the information could be processed automatically. The findings from this study could also be used to design specific program with respect to students' performance. Hence this may lead to a better management towards students' performance in school.

1.7 Thesis Organization

This thesis consists of five chapters. The first chapter is an introduction of this study, which includes the overview of this study, problem statement, objective of the study, scope of the study and significance of the study. Chapter 2 presents the literature review about Statistics, Neural Networks, and students' performance. Methodologies used in this study are described in Chapter Three while Chapter Four presents the findings and discussion. Finally, Chapter Five gives the conclusion of this study and some recommendations for the future study.

CHAPTER 2

LITERATURE REVIEW

This section discusses the literature on the selected area in this study including data mining, descriptive analysis, predictive analysis and students' performance. In addition, the discussion on regression technique is also presented.

2.1 Data Mining

Data mining is the analysis of observational datasets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner (Hand, Mannila & Smyth, 2001). It is a process of analyzing data from different perspectives and summarizing it into useful information.

Data mining tools allow users to analyze data from many different dimensions or angles, categorize and summarize the relationships of the identified variables. It extracts interesting knowledge from large databases, represents an important direction in the development of data and knowledge based system (Dai & Huang, 2007).

Generally, data mining task and model can be grouped into descriptive and predictive. The descriptive task normally involves data which do not have class or target. On the other hand, predictive task employs classification, prediction, time series and regression analysis. For supervised type of data, predictive approach is more appropriate to be used. In this study, each case pattern is represented with target.

Therefore, classification and regression techniques will be explored with the real data collected from schools in Hulu Perak district. Hence, the analysis using Statistics and NN are used to describe the data used in the study and the performance model built based on the conducted experiments.

2.1.1 Descriptive Analysis

Statistics is a mathematical science pertaining to the collection, analysis, interpretation or explanation and presentation of data (Wikipedia, 2008). The main purpose of statistical analysis is to summarize or describe a collection of data (descriptive statistics). It is usable to many disciplines such as social sciences, humanities, business, economy and others.

There are many techniques under statistical analysis such as Student's t-test, chi-square test, Analysis of variance (ANOVA), Mann-Whitney U, Factor Analysis, Correlation, Pearson product-moment correlation coefficient and Spearman's rank correlation coefficient. According to a study conducted by Butcher and Muth (1985) at West Virginia University about predicting students' performance in an introductory computer

science course, they used Statistical Analysis System to perform all statistical analysis such as variance analysis and correlation coefficients.

2.1.2 Predictive Analysis

Neural Networks is a set of connected input / output units where each connection has a weight associated with it. During the learning phase, the network learns by adjusting the weights so as to be able to predict the correct class label of the input samples. The most commonly used technique in data mining is artificial neural networks (ANN) which is non-linear predictive model that use forecasting to produce the outputs.

Huisken and Coffa (2000) stated that ANN is based on biological, like the human brain, by mimicking their architectural structure and information processing in an easy manner. NN can estimate any nonlinear function and is powerful for pattern recognition, classification and forecasting based on the study in forecasting the stock price (Khoa, Sakakibara and Nishikawa, 2006). NN has single layer perceptron and multi layer perceptrons network. Single layer perceptron consists of one or more artificial neurons in parallel. The neurons may be illustrated below:

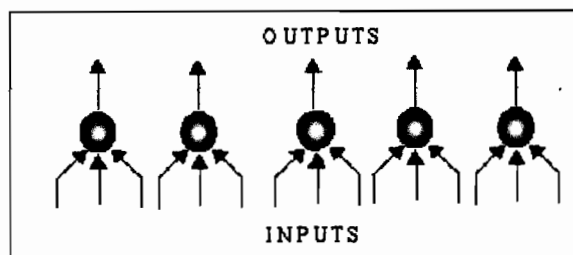


Figure 2.1: Single layer perceptron network

Each neuron in the layer provides one network output. Usually, it is connected to all of the external (or environmental) inputs. According to a study conducted by Perez, Gonzalez and Salinas (2000) on 2D pattern recognition problem, single layer perceptron improved the classification rate.

Multilayer perceptrons (MLP) have been applied successfully to solve some difficult problems by training them in supervised manner with error back-propagation (BP) algorithm (Haykin, 2007). Perceptrons with nonlinear threshold units connected in MLP are more powerful than a single perceptron. BP is an iterative procedure for minimization of error, with adjustments to the weights (Khoa, Sakakibara and Nishikawa, 2006). The network can be trained with available data to model the system.

Feed forward (FF) technique under NN is collected of processing units in a series of two or more layers. Wang and Mitrovic (2002) studied about the possibility of using FF to predict student's behaviour in intelligent educational systems. The prediction accuracy is high and can be further improved by editing the parameters used. Unfortunately, the information is not readily available in on-line situation.

Forecasting is a process that produces a set of outputs by a given set of variables. Forecasting using NN is to find an estimation of mapping between the input and output data through training. The trained NN are then used to predict the values for the future (Khoa, Sakakibara & Nishikawa, 2006).

NN technique has been used in many field such as predicting miner prospectivity, stock price, students' performance and others. A study by Fung *et al.*, (2005) explored the using of BP in predicting of miner prospectivity. It is simple to use and has shown to be robust and gives good results in most cases. BP is also used to forecast the time series stock price using NN (Khoa, Sakakibara & Nishikawa, 2006). The integration of profit and time factors to the training algorithm showed improvements comparing with the traditional training. A research by Huisken and Coffa (2000) stated that the time series analysis method used temporal data as input while the NN method used spatial data as input in predicting congestion on a short time. They found NN method gave the better performance in both the congestion prediction and the processor time.

Based on the research by Bose (2007), NN have created a new and advancing frontier in power electronics. NN applications are used in the intelligent control and estimation for power electronics and motor drives area. A study of NN in dielectric structure parameter determination by multifrequency methods by Drobakhin & Doronin (2006) found that NN produced satisfactory accuracy of estimation not only in point of training but in intermediate points. It used feed forward algorithm and sigmoid activation function as the architecture.

A survey of NN applications in automatic control by Chowdhury *et al.* (2003) found that NN offer some of the most effective control techniques. Neural modeling can also be very effective in the field of detection of faults in dynamic systems. Based on the study by Milovanovic *et al.* (2000) in neuromodeling for microwave design, NN technique

appears as a good alternative for classical CAD techniques. It was also found that ANN models can be successfully applied in more different areas of microwave engineering.

Another research of NN was in multimedia processing by Bojkovic *et al.* (2000). NN can offer a promising horizon for pattern detection/recognition, content based indexing and retrieval. Its technologies also present a solution to a broad spectrum of multimedia applications. NN is used for compression, denoising, contrast enhancement and segmentation.

An artificial NN is constructed to perform a critical task in auditing. A study by Chiu (1994) investigated a potential NN application in auditing. He found that the average difference between desired and predicted outputs was less than 10%. This indicates that ANN can perform the task with results consistent to those done by auditors.

NN application for robotic motion control by Fukuda *et al.* (1987) found that NN can get a desirable gain of the control system by learning and self organize the robust structure. They were also proposed Fuzzy Turbo to avoid the stagnation so that the NN can learn the system quickly. NN method is more applicable and adaptable for the hybrid control of robotic manipulators.

Wong (1993) in his study about AI and NN applications in power system found that they can become a useful technology for power engineering. The combination of Artificial Intelligence (AI) systems and tools for the needs of power engineering can be

achieved. The usefulness of these two techniques will be further expanded when they are able to be combined.

A study by Cripps (1996) found NN can be successfully applied in predicting degree program completion, earned hours and GPA. The NN model with feed forward architecture is used with back propagation learning. It also shows correlations of data and a high level of noise in student admission information.

Another research of NN implementation in education is conducted by Wang and Mitrovic (2002). They used NN to predict student's behaviour in intelligent educational systems. The simple feed forward network is used to make the prediction. The prediction accuracy is high and can be further improved. According to a study conducted by Fausett and Elwasif (1994), the back propagation NN is used to predict performance from test scores. The results gave reasonable predictions of students' performance in Calculus 1.

Regression analysis is used in many areas such as in predicting pollution, students' performance, sales, malignancy of ovarian tumors and many more. For example, the regression method is used to predict the pollution degree of outdoor insulator with the meteorological conditions. The results improved the reliability of the pollution prediction technique (Young, 2003).

A study conducted by Yusof and Syed Hassan (2000) used the linear regression analysis to find a simple equation relating between resonant frequency and number of turn of normal mode radiation pattern used in mobile communications. Results from the analysis show a good approximation to the measured frequency.

The regression analysis and NN are used to predict malignancy of ovarian tumors at University Hospital Leuven, Belgium has been carried out by Lu *et al.* (2002). They found that statistical analysis could be a great help in predicting input variables. In manufacturing, the regression is used to optimize or minimize defects in belt line molding process (Bon, Ogier & Razali, 2007).

Regression analysis has also been applied to education. Multiple regression modeling provides an elegant method of describing such relationships (Hand, Mannila & Smyth, 2001). It provides improved precision for prediction compared to simple linear regression. A study conducted by Hostetler (1996) used multiple regressions in his research on predicting student success in an introductory programming course at University of Illinois. He found that regression techniques may perform good result. Based on this study, prediction of students' performance from the presented model is a useful technique in counseling students.

Burtner (2004) used the regression analysis to show the differences in high school grade point averages and SAT scores account. There is less than 25% of the variability in the data describing student persistence toward an undergraduate degree. Another study

conducted by Streveler *et al* (2003) used step-wise regression to investigate the study strategies that have been practiced by academically successful students at Colorado School of Mines. The study discovered that 4 factors are correlated to the students' performance namely active learning, text anxiety, procrastination and lack of focus and specific study method.

2.2 Students' Performance

Students' performance consists of their achievement in examination. The factors that may influence their performance are examination results, co-curricular activities, socio economic status, gender, attendance and many more. The performance gap is not only a school and classroom issue but also a national issue that must be addressed properly (Terry *et al.*, 2005). Educators, parents and community must support students' learning towards their excellence.

Another study conducted by Erbe and Mach (2000) at Chicago Public Secondary Schools shows that the other school variables such as school climate, students' background and others give great contribution to the variance in Mathematics achievement. Socio economic variable is the major contributor to the variance.

Gibson and Margaret (2003) at Hillside High School stated in their research that a sense of caring community is necessary to motivate the students. The students are encouraged to become involved in school and community activities to support them in academic achievement. Based on the Henchey and Norman (2001) at Canada High School

research, extra curricular activities in schools take on special significance as integral part of learning experiences that school provides.

Socio economic status provides the good or bad facilities to the students. For example, after school tutoring is contributed to the students' achievement (Gibson & Margaret, 2003) at Hillside High School. Low income background students tend to be concentrated in certain schools, which can depress achievement for all children (Terry *et al.*, 2005). According to Henchey and Norman (2001) at Canada High School, schools with students from high income homes tend to do better than schools with students from low income homes.

Students' gender becomes one of the factors in contributing their level of performance (Cripps, 1996) at Tennessee State University. In attempts to predict degree completion, students' data such as gender, age, race and others were considered. A study on "Women in Mathematics and Science" by Bae *et al.*, (1997) found that boys and girls have similar Mathematics and Science proficiency scores at age 9 but the gap increased rapidly. In general, the gender difference in Mathematics scores is less than in Science scores.

Hostetler (1996) conducted research at University of Illinois finds that student's previous results had a strong correlation with their future results. The past academic success is a good predictor of current academic success. Attendance also contributes to the student's performance. It is accurately as a research by Terry *et al.* (2005) at Indiana

High School found that college attendance rates between African American and Hispanic students contributed to their performance.

The healthy co-curricular activities play main role in determining their performance. Additional research by Hayek *et al.* (1999) examined that college activities contributed to continuous learning which would help enhancing students' performance.

2.3 Discussion

In conclusion, two tools have been used to predict students' performance in UPSR such as Statistics and NN. Statistics is used to analyze and interpret the results. The multilayer perceptrons with back propagation which is also referred as back-propagation network (BPN) has shown its potential in predicting student's performance. As a result, this study will use multilayer perceptrons with back propagation technique in predicting students' performance. The use of Statistics techniques and NN in descriptive and modeling has been presented in the previous section. The NN has the potential to be used as a prediction tool. In many cases, NN accuracy outperforms statistical results. However, the role of Statistics in supplementing explanation to NN cannot be underestimated. It is used properly and the reliability of the results produces by NN can be further explained by Statistics.

CHAPTER 3

METHODOLOGY

This section presents the methodology used in this study. It also discusses the methodology adapted from Flow Diagram of The Development Process of ANN (Turban, 1992). The method presents data collection, preprocessing and analysis phases.

3.1 Methodology

The methodology is adapted from Flow Diagram of The Development Process of ANN (Turban, 1992) including data collection, preprocessing and analysis phases. The raw data was grouped, scaled and transformed into suitable inputs and outputs. The BPN used to train, validate and test for its accuracy. The methodology steps are shown below:

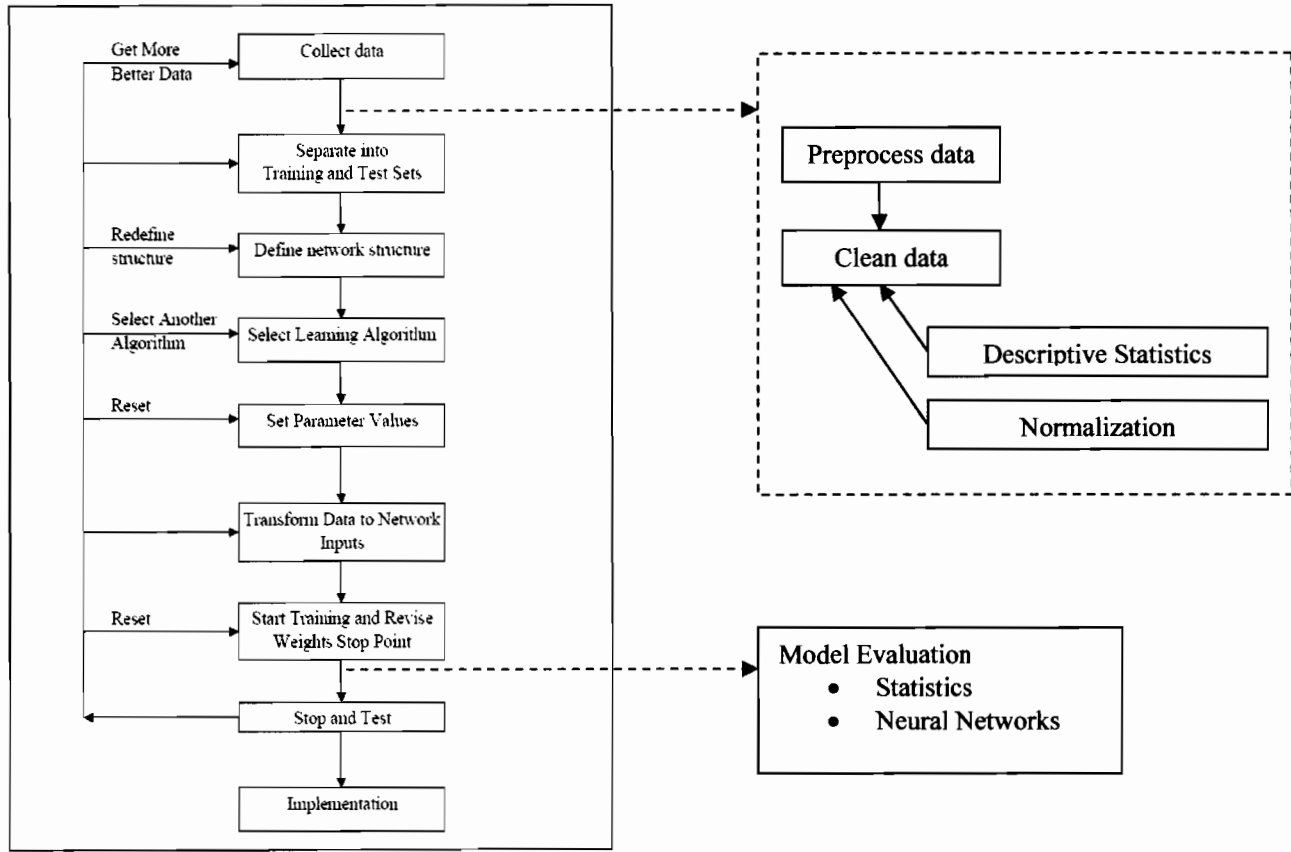


Figure 3.1: Flow diagram of the development process of ANN

3.1.1 Data Collection

About 420 students from various schools in Hulu Perak district are involved in the analysis. Besides SMM, the other instrument used to survey the students' background is as follows:

KAR2015		KEHASKIN 25622007	
ABDUL HAZIF ARIF BIN ABD RASHID		4	
DR TAHUN ENAM		2	
LJYA		2	
L L LELAKI		0	
01 MELAYU		0	
01 ISLAM		0	
TW WARGANEGARA		T TIDAK	
		J1 IBU BAPA	
MAKLUMAT PENJAGA		MAKLUMAT KOKURKULUM	
[REDACTED]			
A01 PENGAKAP	B02 PERSATUAN BAHASA MELAYU	F05 BOLA SEPAK	
SU BETAUSAHA	BN BENDAHARI	NAB Pengerusi	

Name :
 Class :
 Birth Certificate :
 Gender :
 Race :
 Co-curricular :
 Position :

Figure 3.2: The background form

The form is used to survey the other data that cannot be provided accurately by SMM. It is an alternative to seek for students' background information. Data preprocessing is done to keep up the good quality of data. Several points to consider about data preparation are data cleaning, data integration and data transformation. Data cleaning involves filling in missing values, smoothing noisy data, identify or remove outliers and resolve inconsistencies. Data integration is integration of multiple databases, data cubes or files. Data transformation is normalization and aggregation. The data is scaled to fall within a small and specific range using z-score normalization.

3.1.2 Data Description

Data description involves the collection of useful dataset. It differs for supervised or unsupervised learning. For supervised learning, it has a goal based on the choosing of dependent variable or target and specifying the data fields. For unsupervised learning, the goal is to group similar types of data or to identify the exceptions in a dataset. Defining studies must be done to specify the realm of data such as for training,

validation and test sets. After that, the sample size must be specified. It is not always necessary to mine an entire dataset. The raw dataset is shown below:

Table 3.1: The raw dataset

YEAR 6 STUDENTS EXAMINATION MARKS ENDED 2007																													
2007																													
GERAK BEMPU 2					PKSR 1					TRIAL UPSR					AVERAGE					TARGET									
AVERAGE	BM1	BM2	BN	MT	SC	AVERAGE	BM1	BM2	BN	MT	SC	AVERAGE	BM1	BM2	BN	MT	SC	AVERAGE	BM1	BM2	BN	MT	SC	BM1	BM2	BN	MT	SC	
94	68	54	97	98	76	77	52	78	89	89	70	76	89	98	86	80	75	85	77	71	83	97	74	B	B	A	A	B	
48	86	95	95	67	67	66	95	95	95	95	86	61	67	67	95	95	86	66	78	52	98	99	99	89	B	C	C	C	C
40	85	17	17	34	34	31	17	17	95	95	95	40	34	34	95	67	67	61	37	33	32	61	91	D	D	D	C	C	
21	34	95	95	17	34	39	86	86	95	17	17	32	86	86	95	34	17	95	91	80	90	22	28	C	B	C	D	D	
67	67	95	67	67	86	68	95	95	95	95	95	95	67	67	95	95	95	68	64	61	61	61	68	B	B	B	B	B	
38	34	34	95	95	17	39	95	95	95	95	95	95	95	95	67	67	67	62	41	42	98	92	37	C	C	C	C	D	
40	95	95	95	95	95	95	17	95	95	95	95	47	34	67	67	67	67	60	47	88	48	99	89	C	C	C	C	C	
46	67	95	34	95	95	53	95	95	95	95	95	58	67	67	95	95	62	94	61	91	90	90	90	C	B	C	C	C	
67	86	86	67	67	86	78	95	95	95	95	95	95	67	67	95	95	95	68	70	78	61	61	88	B	B	B	B	B	
82	86	86	67	67	67	75	95	95	95	95	95	95	95	95	67	67	95	60	74	74	63	70	98	B	B	B	B	B	
75	86	86	67	67	67	75	86	86	95	95	86	67	86	86	67	67	86	72	78	74	63	70	72	B	B	B	B	B	
20	95	34	34	95	34	42	95	95	95	95	95	95	95	95	67	67	67	62	67	42	41	52	91	C	C	C	C	C	
84	86	84	90	67	88	82	86	95	95	95	89	75	86	90	67	83	89	84	85	75	83	73	85	A	B	A	B	A	
62	86	86	67	67	86	78	95	95	17	17	95	40	95	95	34	17	95	43	68	86	91	47	88	B	B	C	B	B	
62	86	67	67	95	95	66	95	95	95	95	95	95	67	67	95	95	67	62	60	63	68	57	61	B	B	B	C	B	
68	95	34	95	95	34	47	95	95	95	86	86	67	95	67	67	86	86	72	99	91	88	70	61	C	B	B	B	B	
38	67	95	17	17	95	42	95	95	95	95	95	95	95	67	67	95	95	60	32	93	45	35	17	C	C	C	D	C	
78	86	67	67	67	67	71	17	17	86	17	95	38	34	34	86	34	95	49	68	88	72	98	61	B	C	B	C	B	
35	95	67	67	95	95	68	67	67	67	67	67	67	95	67	95	95	67	88	88	93	94	92	94	C	C	C	C	C	
62	67	67	95	95	95	68	67	86	86	67	67	75	95	86	86	67	95	70	61	72	88	63	61	B	B	B	B	B	
75	86	34	17	17	34	31	86	67	67	67	67	71	86	95	95	67	67	66	78	83	47	49	92	B	C	C	C	B	
31	95	95	95	95	95	95	67	67	67	67	86	71	95	95	67	67	86	66	82	98	94	91	97	C	C	C	C	C	
75	86	86	86	67	67	78	86	67	67	17	67	61	86	67	95	34	95	99	78	72	76	98	83	B	B	B	C	B	

3.1.3 Data Preprocessing

It involves data cleaning, data integration, data transformation, data reduction and data discretization. All the process must be done because data in the real world is dirty. Usually, the raw data is incomplete in terms of lacking attribute values, lacking certain attributes of interest or containing only aggregate data. The data is noisy because it contains errors or outliers. It is also inconsistent which is containing discrepancies in codes or name. The preprocessing process must be done because there is no quality in data showing no quality mining results. For cleaning data process, the missing values must be fill in, smooth noisy data, identify or remove outliers and resolve

inconsistencies. For data integration, the data from multiple databases, data cubes or files must be integrated. Normalization and aggregation are the example process under data transformation. Data reduction obtains reduced representation in volume but produces the same or similar analytical results. Data discretization involves part of data reduction but with particular importance especially for numerical data. The variables used in this dataset are shown below:

Table 3.2: The dataset variables

	A	B	C	D	E	F	G
1	YEAR 6 STUDENTS						
2	EXAMINATION MARKS ENDED 2007						
3							
4							
5							
6							
7	GENDER	ATTENDANCE	CO-CURRICULAR ACTIVITIES	POSITION IN FAMILY	FATHER'S JOB	MOTHER'S JOB	FAMILY PERCAPITA INCOME (RMB)
8	FEMALE	GOOD	ACTIVE	1	TEACHER	TEACHER	250
9	MALE	GOOD	ACTIVE	3	CLERK	NURSE	330
10	FEMALE	GOOD	ACTIVE	5	SOLDIER	HOUSEWIFE	67
11	MALE	GOOD	ACTIVE	3	ODD JOB	ODD JOB	35
12	MALE	GOOD	ACTIVE	2	RUBBER TAPPER	RUBBER TAPPER	67
13	MALE	GOOD	ACTIVE	5	DENTIST	NURSE	1200
14	MALE	BAD	PASSIVE	10	TECHNICIAN	HOUSEWIFE	240
15	MALE	BAD	PASSIVE	5	TEACHER	HOUSEWIFE	590
16	MALE	GOOD	ACTIVE	3	ODD JOB	FRUITS SELLER	140
17	MALE	GOOD	ACTIVE	2	POLICE	HOUSEWIFE	120
18	MALE	GOOD	ACTIVE	1	FISHMONGER	RUBBER TAPPER	230
19	MALE	GOOD	ACTIVE	1	ODD JOB	ODD JOB	120
20	MALE	GOOD	ACTIVE	5	DOCTOR	CLERK	1900
21	MALE	GOOD	ACTIVE	5	RUBBER TAPPER	RUBBER TAPPER	98

All figures must be changed into numerical according to their coding. The numeric coding is shown below:

Table 3.3: The numeric coding

Attributes	Types	Represent as
Gender	Male	1
	Female	2
Attendance	Good	1
	Bad	2
Co-curricular Activities	Active	1
	Passive	2
Position in Family	0	0
	1	1
	2	2
	3	3
	4	4
	5	5
	6	6
	7	7
	8	8
	9	9
Father's Job	Education	1
	Government Agencies	2
	Private Sector	3
	Odd Job	4
	Housewife	5
	Orphan	6
Mother's Job	Education	1
	Government Agencies	2
	Private Sector	3
	Odd Job	4
	Housewife	5
	Orphan	6
Family Income Percapita	RM 0 - 200	1
	RM 201 - 400	2
	RM 401 - 600	3
	RM 601 - 800	4
	RM 801 - 1000	5
	RM 1001 - 1200	6
	RM 1201 - 1400	7
	RM 1401 - 1600	8
RM 1601 - 1800	9	
RM 1800 - 2000	10	

data \ data_final \ GRADE \ Sheet3 \ CHILD NUMBER \ NUMBER \ INC

Then, the dataset will be changing into numerical forms for experiment sake. The numerical data is shown as below:

Table 3.4: The numerical data

	A	B	C	D	E	F	G
1	gender	attendance	co_curricular	position_family	father_job	mother_job	family_income
2							
3	2	1	1	1	1	1	3
4	1	1	1	3	2	2	2
5	2	1	1	5	2	5	1
6	1	1	1	3	4	4	1
7	1	1	1	2	4	4	1
8	1	1	1	6	2	2	3
9	1	2	2	10	2	5	2
10	1	2	2	5	1	5	3
11	1	1	1	3	4	4	1
12	1	1	1	2	2	5	1
13	1	1	1	1	4	4	2
14	1	1	1	1	4	4	1
15	1	1	1	5	2	2	10
16	1	1	1	6	4	4	1
17	1	1	1	3	4	4	1
18	1	1	1	7	6	4	1
19	1	2	2	7	2	2	10
20	1	1	1	3	4	4	1

3.1.4 Separate into Training and Test Sets

After data preprocessing is done, the data is divided into training, validation and test phase with the division of 80:10:10. The input set is randomized and divided into three distinct sets; 80% of the set is used for training, 10% for validation and the remaining 10% for testing. Training is used to train the networks, validation for monitoring the training process while test phase is to confirm the networks performance. The data allocation is shown below:

The screenshot shows a 'Data Allocation' dialog box with the following settings:

- File Order:** Random, Seed: 3
- Data Blocking:** None
- Mark remaining records as not used
- Include test records in range calculation
- Recalculate range information
- Data Sets (desired):**

	%	#
Training	80.	336
Validation	10.	42
Test	10.	42
Not used	?	?
Total	100	420
- Assignment:** Random, Seed: 3
- Test records at end

Buttons: OK, Cancel

Figure 3.3: Data allocation

3.1.5 Define Networks Structure

This study uses MLP with BPN networks. The input variables are fed into input an input layer which is then connected to a single hidden layer then finally to the output layer. The analysis for each subject is conducted separately. For example, for BM1, a total of 13 attributes are considered in this study represent grades based on investigation. The architecture of MLP networks is as follows:

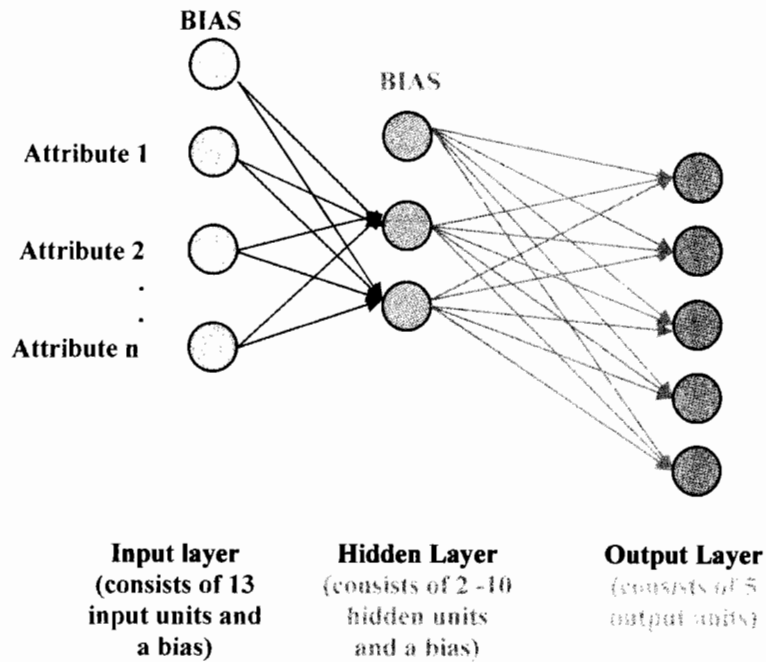


Figure 3.4: The architecture of multilayer perceptrons

A hidden layer is an intermediate layer that connects the input layer with output layer in MLP network. It increases the network complexity and produce better results. The output layer has five nodes to represent the expected grade for each subject. MLP is general purpose, flexible and used for non linear model that gives enough hidden neurons and enough data. It can approximate virtually any function to any desired degree of accuracy.

3.1.6 Select Learning Algorithm

Learning algorithm is used to show how well a neural network to solve a problem. It uses the BPN algorithm. It involves the FF of the input training pattern, the calculation of associated error and the adjustment of the weight. It is the powerful and flexible tool. Two important elements under BPN are learning rate and momentum. Learning rate

determines the magnitude of weight changes. Momentum rate allows the larger learning rate in faster convergence and minimizing the error.

Activation function is an element in neuron in the MLP. Its function is to smooth nonlinear function used to modify the summed inputs to the neurons. Three activation functions (Tanh, Sigmoid and Linear) are tested in developing the model. Those activation functions are tested to find out the best prediction. Then, Sigmoid function is

found as the best using $f(x) = \frac{1}{1 + \exp(-x)}$ formula. The learning algorithm used in this

study is shown below:

Figure 3.5: MLP algorithm

	1	2	3	4
Learning Coeff.	0.1	0.1	0.1	0.1
Momentum Coeff.	0.1	0.1	0.1	0.1
Max. Records	336	336	336	336
Max. Updates	336	336	336	100

Figure 3.6: Learning rate, momentum rate and epoch

3.1.7 Set Parameter Values

The analysis has to be done for every subject. Each subject has six input nodes taken from their two years results starting from Year 5 2006 until Year 6 2007. The inputs are as follows:

- i) pksr1_06 - Students' grade in PKSR 1 2006 (*A, B, C, D & E*).
- ii) pksr2_06 - Students' grade in PKSR 2 2006 (*A, B, C, D & E*).
- iii) gempur1_07 - Students' grade in Gerak Gempur 1 2007 (*A, B, C, D & E*).
- iv) gempur2_07 - Students' grade in Gerak Gempur 2 2007 (*A, B, C, D & E*).
- v) pksr1_07 - Students' grade in PKSR 1 2007 (*A, B, C, D & E*).
- vi) trial_07 - Students' grade in trial UPSR 2007 (*A, B, C, D & E*).
- vii) gender - Students' gender (Male/Female)
- viii) attendance - Students' attendance (Good/Bad)
- ix) co_curricular - Students' participation in co-curricular activities (Active/Passive)
- x) position_family – Students' position in family (1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and above)
- xi) father_job - Father's occupation (Education, Government Agencies, Private Sectors, Odd Job and Orphan)
- xii) mother_job - Mother's occupation (Education, Government Agencies, Odd Job, Housewife and Orphan)
- xiii) family_income - Average parents' income (RM0-200, RM201-400, RM401-600, RM601-800, RM801-1000, RM1001-1200,

RM1201-1400, RM1401-1600, RM1601-1800 and
 RM1801-2000)\

3.1.8 Transform Data to Network Inputs

The nodes details are shown below:

Table 3.5: The attributes descriptions

Num.	Physical Attribute	Features
1.	pk _{sr1} _06	Used as Input
2.	pk _{sr2} _06	Used as Input
3.	gempur ₁ _07	Used as Input
4.	gempur ₂ _07	Used as Input
5.	pk _{sr1} _07	Used as Input
6.	trial_07	Used as Input
7.	gender	Used as Input
8.	attendance	Used as Input
9.	co_curricular	Used as Input
10.	position_family	Used as Input
11.	father_job	Used as Input
12.	mother_job	Used as Input
13.	family_income	Used as Input
7.	A	Used as Output
8.	B	Used as Output
9.	C	Used as Output
10.	D	Used as Output
11.	E	Used as Output

3.1.9 Start Training and Revise Weights Stop Point

Regression analysis and NN are the tools for the experiments. Various learning rate, α , momentum, μ , and the number of units in the hidden layer, are used to test the performance of the network. Before training began, the preprocessed data is converted to CSV (comma delimited) format. The network has been trained to make the predictions. The process is shown below:

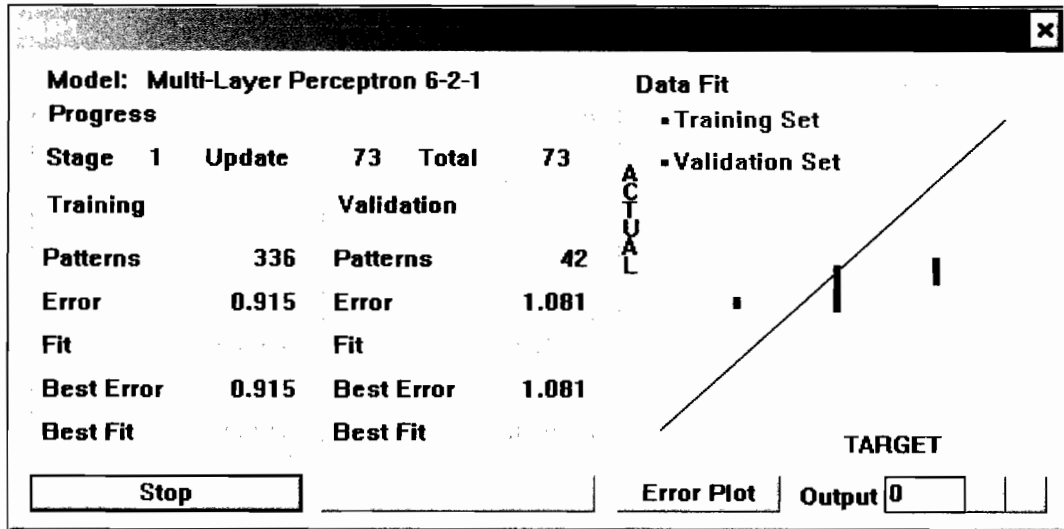


Figure 3.7: Training and test process

Training phase is done to choose the best model. Then, testing phase is done to test the predictive performance of the best model obtained. Both phases are done to determine the most suitable hidden unit, learning rate, momentum rate and activation function to produce the highest accuracy. The NN tool used for this project is Neural Connection (NC) and SPSS 10.0 software as shown below:

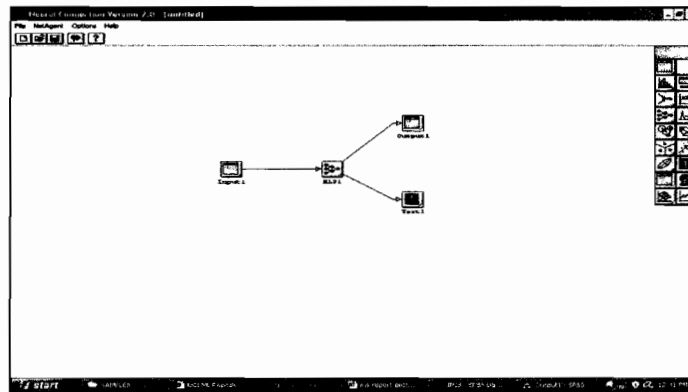


Figure 3.8: Flow diagram process for multilayer perceptrons in Neural Connection

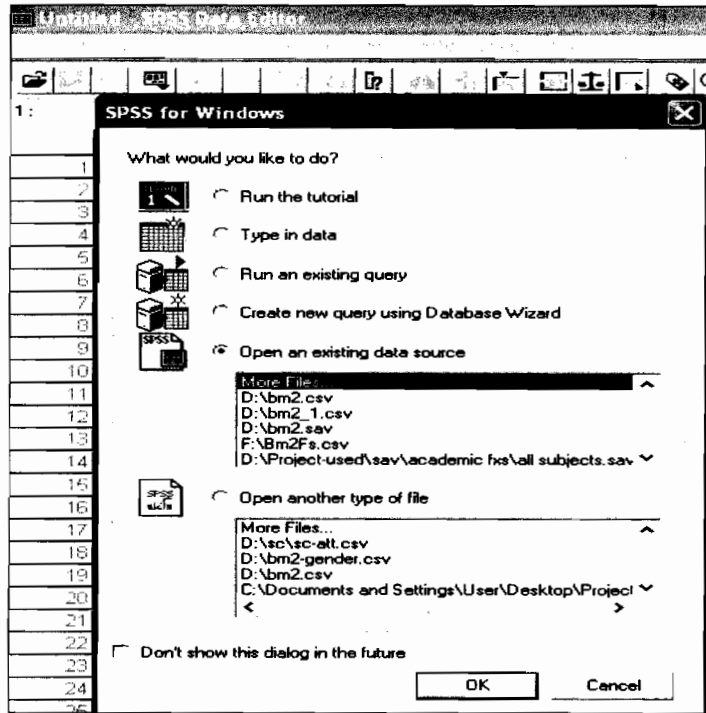


Figure 3.9: The main SPSS window

3.1.10 Stop and Test

In the heart of the process, it can be stop to continue testing for better result. The training process can be reset, reset the parameter values, select another algorithm, redefine structure and get better data to produce the better performance.

3.1.11 Implementation

After the whole process is completed, only then the best learning rate, momentum rate and activation function are found to produce the highest accuracy of the prediction. Then, this technique can be implemented in purpose of predicting students' performance in UPSR.

CHAPTER 4

RESULTS AND FINDINGS

This section presents the results and analysis obtained from the data collected. Various experiments have been conducted to explore the effects of each prediction variable against the targeted output.

4.1 First Dimension Analysis

This analysis is done to make sure the data does not have missing value. The analysis is shown below:

Table 4.1: Frequency data

Statistics								
		Gender	Attendance	Co-curricular	Position in Family	Father's Job	Mother's Job	Family Income Percapita
N	Valid	420	420	420	420	420	420	420
	Missing	0	0	0	0	0	0	0

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The experiment can be easily carried out because the data is clean. About 420 students from four schools in Hulu Perak district were tested. The gender allocation is shown as below:

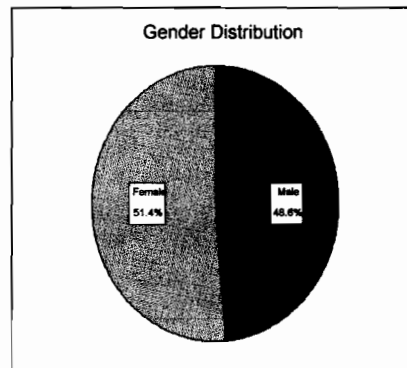


Figure 4.1: Gender distribution

About 51.4% or 214 female students involved in this study. Only 48.6% or 206 male students were being tested. There were 8 more female students than male students tested in this study.

One of the factors contributes to the students' performance is gender. The survey had been carried out to determine the relationship between gender and performance. The analysis of missing value is shown below:

Table 4.2: Gender frequency data

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * BM1	420	100.0%	0	.0%	420	100.0%
Gender * BM2	420	100.0%	0	.0%	420	100.0%
Gender * BI	420	100.0%	0	.0%	420	100.0%
Gender * MT	420	100.0%	0	.0%	420	100.0%
Gender * SC	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The relationship between gender and each subject is shown below:

Table 4.3: Gender and BM1 crosstabulation

GENDER * BM1 Crosstabulation										
BM1										
	A	% within BM1	B	% within BM1	C	% within BM1	D	% within BM1	Gender	Total
Male	21	10%	81	40%	94	46%	8	4%	204	49%
Female	22	10%	92	43%	91	42%	11	5%	216	51%
Total	43	10%	173	41%	185	44%	19	5%	420	100%

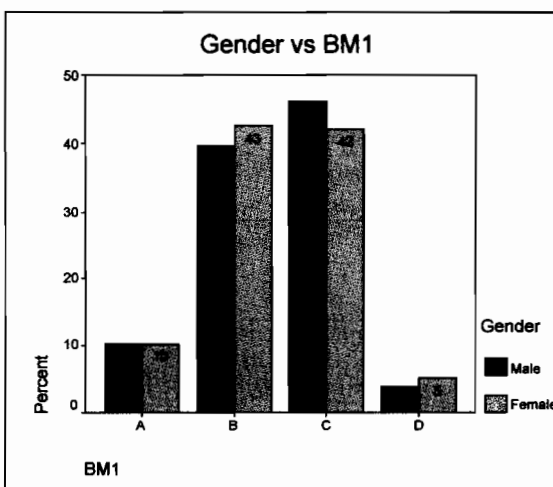


Figure 4.2: Relationship between gender and BM1

There were about 10% or 21 male students got A in Bahasa Melayu 1 (Pemahaman), while the same percent of female students or 22 got A in that subject. 40% or 81 male students got B in BM1 while 43% or 92 female students got B. About 46% or 94 male students and 42% or 91 female students got C in that subject. The lowest percent was given by grade D which was only 4% or 8 male students and 5% or 11 female students. The same percent of genders got A in BM1 while the female students gave the higher percent in grade B and D.

Table 4.4: Gender and BM2 Crosstabulation

GENDER * BM2 Crosstabulation													
BM2												Gender	Total
	A	% within BM2	B	% within BM2	C	% within BM2	D	% within BM2	E	% within BM2			
Male	9	4%	73	36%	98	48%	23	11%	1	0.5%	204	49%	
Female	6	3%	89	41%	101	47%	18	8%	2	1%	216	51%	
Total	15	4%	162	39%	199	47%	41	10%	3	1%	420	100%	

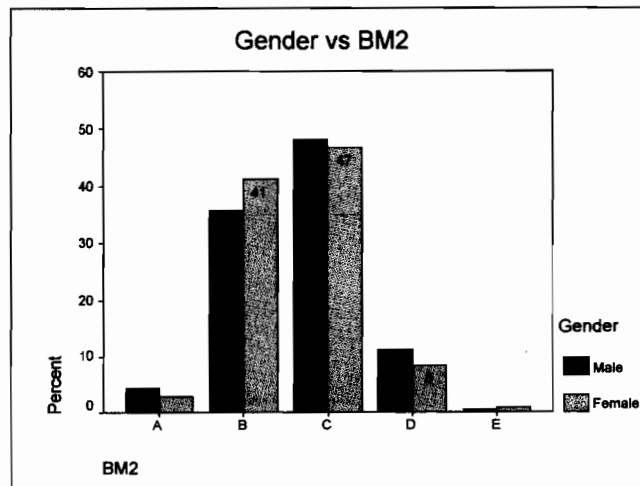


Figure 4.3: Relationship between gender and BM2

There were about 4% or 9 male students got A in Bahasa Melayu 2 (Penulisan), while 3% or 6 of female students got A in the that subject. 36% or 73 male students got B in BM2 while 41% or 89 female students got B. The highest percent was given by grade C where 48% or 98 male students and 47% or 101 female students. There were 11% or 23 of male students and 8% or 18 of female students got D. The lowest percent was given by grade E which were only 0.5% or 1 male student and 1% or 2 female students.

Table 4.5: Gender and BI crosstabulation

GENDER * BI Crosstabulation													
BI												Gender	Total
A	% within BI	B	% within BI	C	% within BI	D	% within BI	E	% within BI				
Male	4	2%	85	42%	88	43%	25	12%	2	1%	204	49%	
Female	12	6%	76	35%	101	47%	23	11%	4	2%	216	51%	
Total	16	4%	161	38%	189	45%	48	11%	6	1%	420	100%	

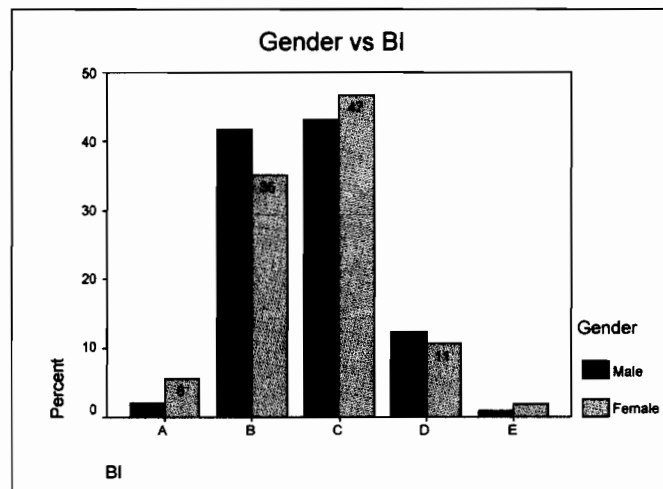


Figure 4.4: Relationship between gender and BI

There were about 2% or 4 male students got A in English, while 6% of female students or 12 got A. 42% or 85 male students got B in BI while 35% or 76 female students got

B. About 43% or 88 male students and 47% or 101 female students got C in that subject. There were 12% or 25 of male students and 11% or 23 of female students got D. The lowest percent was given by grade E which was only 1% or 2 male students and 2% or 4 female students. The female students gave the higher percent compared to male students for grade A, C and E. Male students gave higher percent for grade B and D.

Table 4.6: Gender and MT crosstabulation

GENDER * MT Crosstabulation													
MT												Gender	Total
A	% within MT	B	% within MT	C	% within MT	D	% within MT	E	% within MT				
Male	1	0.5%	74	36%	97	48%	33	16%	1	0.5%	204	49%	
Female	10	5%	59	27%	107	50%	34	16%	6	3%	216	51%	
Total	11	3%	133	32%	202	48%	67	16%	7	2%	420	100%	

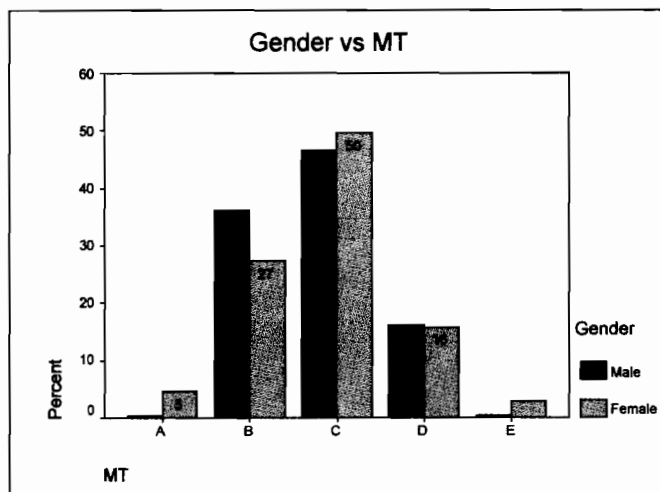


Figure 4.5: Relationship between gender and MT

There were about 0.5% or 1 male student got A in Mathematics, while 5% of female students or 10 got A. 36% or 74 male students got B in MT while 27% or 59 female students got B. About 48% or 97 male students and 50% or 107 female students got C

in that subject. There were the same percent of male and female students got grade D which was 16% or 33 and 34. The lowest percent was given by grade E which was only 0.5% or 1 male student and 3% or 6 female students. The female students gave the higher percent compared to male students for grade A, C and E. The male students gave the higher percent for grade B and D.

Table 4.7: Gender and SC crosstabulation

GENDER * SC Crosstabulation													
SC												Gender	Total
A	% within SC	B	% within SC	C	% within SC	D	% within SC	E	% within SC				
Male	5	2%	78	38%	102	50%	18	9%	1	0.5%	204	49%	
Female	13	6%	92	43%	98	45%	11	5%	2	1%	216	51%	
Total	18	4%	170	40%	200	48%	29	7%	3	1%	420	100%	

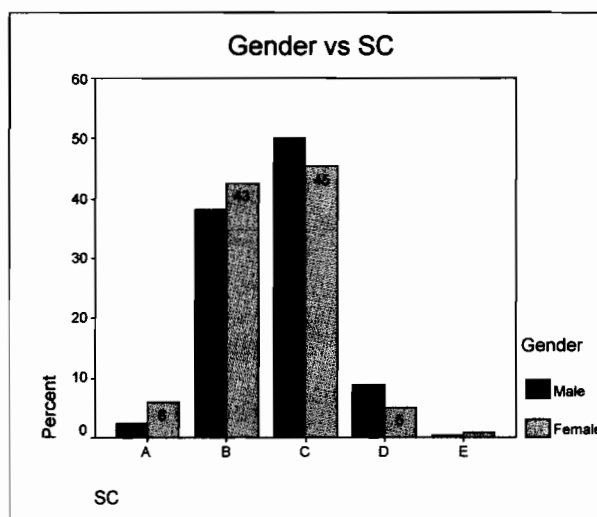


Figure 4.6: Relationship between gender and SC

There were about 2% or 5 male students got A in Science, while 6% of female students or 13 got A. 38% or 78 male students got B in SC while 43% or 92 female students got B. About 50% or 102 male students and 45% or 98 female students got C in that subject.

There were about 9% or 18 male students and 5% or 11 of female students got D. The lowest percent was given by grade E which was only 0.5% or 1 male student and 1% or 2 female students. The female students gave the higher percent compared to male students for grade A, B and E. The male students gave the higher percent for grade C and D.

The other factor contributes to the students' performance is attendance. The survey had been carried out to determine the relationship between attendance and performance. The analysis of missing value is shown below:

Table 4.8: Attendance frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Attendance * BM1	420	100.0%	0	.0%	420	100.0%
Attendance * BM2	420	100.0%	0	.0%	420	100.0%
Attendance * BI	420	100.0%	0	.0%	420	100.0%
Attendance * MT	420	100.0%	0	.0%	420	100.0%
Attendance * SC	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The attendance allocation either bad or good as shown below:

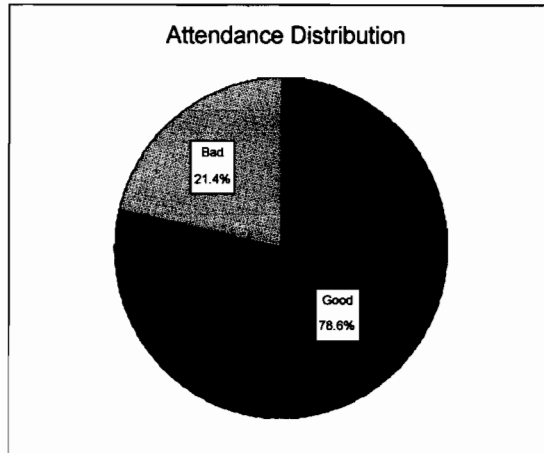


Figure 4.7: Attendance distribution

There were 78.6% or 330 students having good attendance. Bad attendance was only 21.4% or 90 students. More than half students had good attendance. The relationship between each subject and attendance is shown below:

Table 4.9: Attendance and BM1 crosstabulation

ATTENDANCE * BM1 Crosstabulation										
BM1										
	A	% within BM1	B	% within BM1	C	% within BM1	D	% within BM1	Attendance	Total
Good	39	12%	138	42%	142	43%	11	3%	330	79%
Bad	4	4%	35	39%	43	48%	8	9%	90	21%
Total	43	10%	173	41%	185	44%	19	5%	420	100%

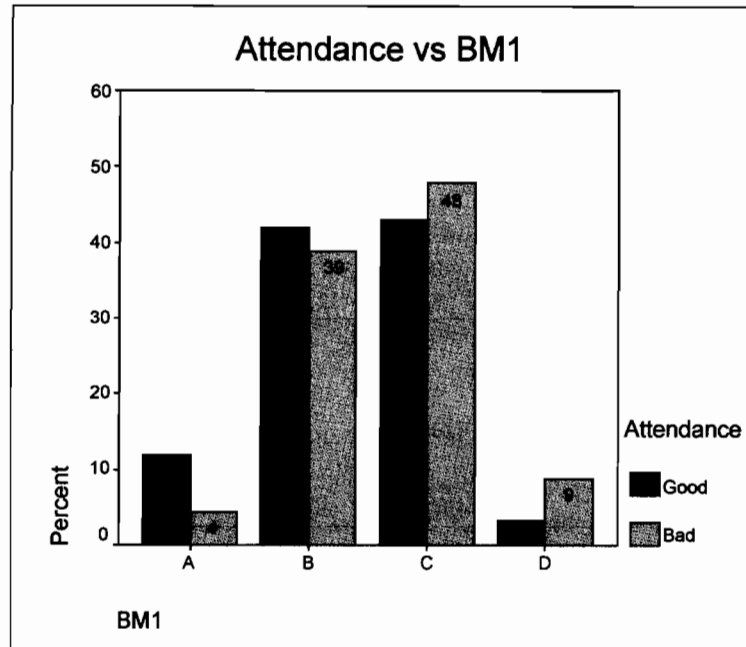


Figure 4.8: Relationship between attendance and BM1

There were about 12% or 39 of good attendance students got A in Bahasa Melayu 1 (Pemahaman), while 4% of bad attendance students or 4 got A. 42% or 138 of good attendance students got B in BM1 while 39% or 35 of bad attendance students got B. About 43% or 142 of good attendance students and 48% or 43 of bad attendance students got C in that subject. The lowest percent was given by grade D which was only 3% or 11 of good attendance students and 9% or 8 of bad attendance students. The good attendance students gave the higher percent for grade A and B. The bad attendance students gave the higher percent for grade C and D.

Table 4.10: Attendance and BM2 crosstabulation

ATTENDANCE * BM2 Crosstabulation												
BM2												
	A	% within BM2	B	% within BM2	C	% within BM2	D	% within BM2	E	% within BM2	Attendance	Total
Good	15	5%	136	41%	147	45%	31	9%	1	0.3%	330	79%
Bad	0	0%	26	29%	52	58%	10	11%	2	2%	90	21%
Total	15	4%	162	39%	199	47%	41	10%	3	0.7%	420	100%

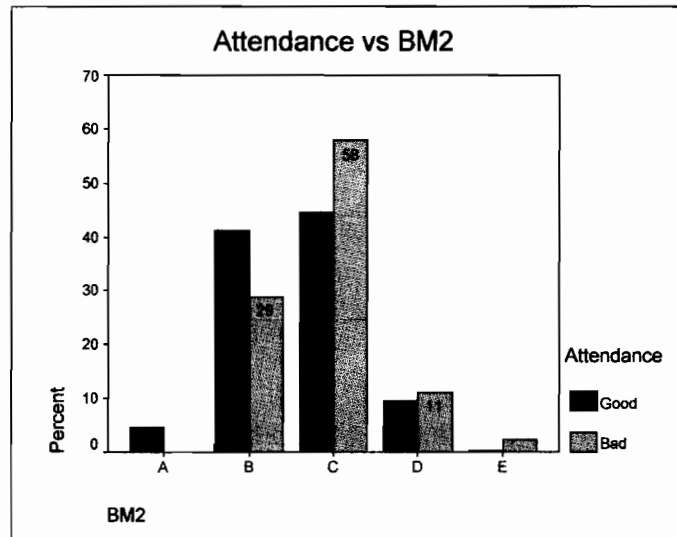


Figure 4.9: Relationship between attendance and BM2

There were about 5% or 15 of good attendance students got A in Bahasa Melayu 2 (Penulisan), while there was no grade A for bad attendance students. 41% or 136 of good attendance students got B in BM2 while 29% or 26 of bad attendance students got B. About 45% or 147 of good attendance students and 58% or 52 of bad attendance students got C in that subject. There were 9% or 31 of good attendance students and 11% or 10 bad attendance students got D. The lowest percent was given by grade E which was only 0.3% or 1 of good attendance students and 2% or 2 of bad attendance student. The good attendance students gave the higher percent for grade A and B. The bad attendance students gave the higher percent for grade C, D and E.

Table 4.11: Attendance and BI crosstabulation

ATTENDANCE * BI Crosstabulation												
BI												
	A	% within BI	B	% within BI	C	% within BI	D	% within BI	E	% within BI	Attendance	Total
Good	16	5%	141	43%	134	41%	36	11%	3	1%	330	79%
Bad	0	0%	20	22%	55	61%	12	13%	3	3%	90	21%
Total	16	4%	161	38%	189	45%	48	11%	6	1%	420	100%

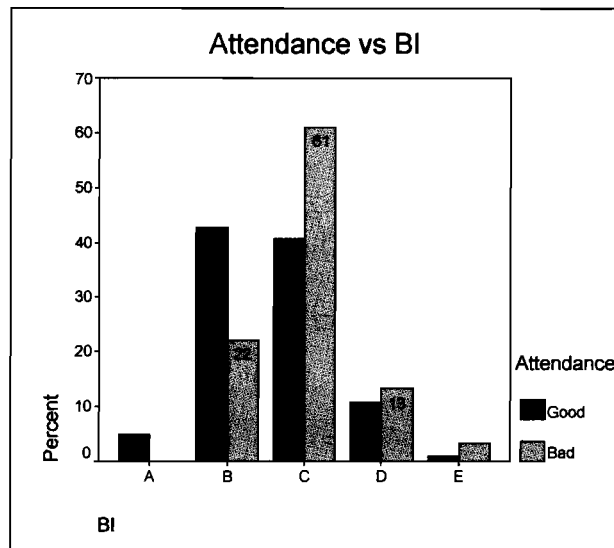


Figure 4.10: Relationship between attendance and BI

There were about 5% or 16 of good attendance students got A in English, while there was no grade A for bad attendance students. 43% or 141 of good attendance students got B in BI while 22% or 20 of bad attendance students got B. About 41% or 134 of good attendance students and 61% or 55 of bad attendance students got C in that subject. There were 11% or 36 of good attendance students and 13% or 12 of bad attendance students got D. The lowest percent was given by grade E which was only 1% or 3 of good attendance students and 3% or 3 of bad attendance students. The good attendance students gave the higher percent for grade A and B. The bad attendance students gave the higher percent for grade C, D and E.

Table 4.12: Attendance and MT crosstabulation

ATTENDANCE * MT Crosstabulation												
MT												
	A	% within MT	B	% within MT	C	% within MT	D	% within MT	E	% within MT	Attendance	Total
Good	11	3%	117	35%	158	48%	39	12%	5	2%	330	79%
Bad	0	0%	16	18%	44	49%	28	31%	2	2%	90	21%
Total	11	3%	133	32%	202	48%	67	16%	7	2%	420	100%

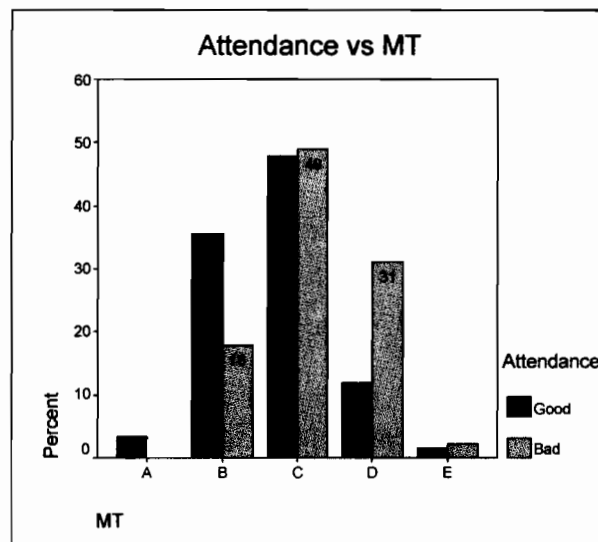


Figure 4.11: Relationship between attendance and MT

There were about 3% or 11 of good attendance students got A in Mathematics, while there was no grade A for bad attendance students. 35% or 117 of good attendance students got B in MT while 18% or 16 of bad attendance students got B. About 48% or 158 of good attendance students and 49% or 44 of bad attendance students got C in that subject. There were 12% or 39 of good attendance students and 31% or 28 of bad attendance students got D. The lowest percent was given by grade E which were only 2% or 5 of good attendance students and 2% or 2 of bad attendance students. The good attendance students gave the higher percent for grade A and B. The bad attendance students gave the higher percent for grade C, D and E.

Table 4.13: Attendance and SC crosstabulation

ATTENDANCE * SC Crosstabulation												
SC												
	A	% within SC	B	% within SC	C	% within SC	D	% within SC	E	% within SC	Attendance	Total
Good	17	5%	144	44%	151	46%	18	5%	0	0%	330	79%
Bad	1	1%	26	29%	49	54%	11	12%	3	3%	90	21%
Total	18	4%	170	40%	200	48%	29	7%	3	0.7%	420	100%

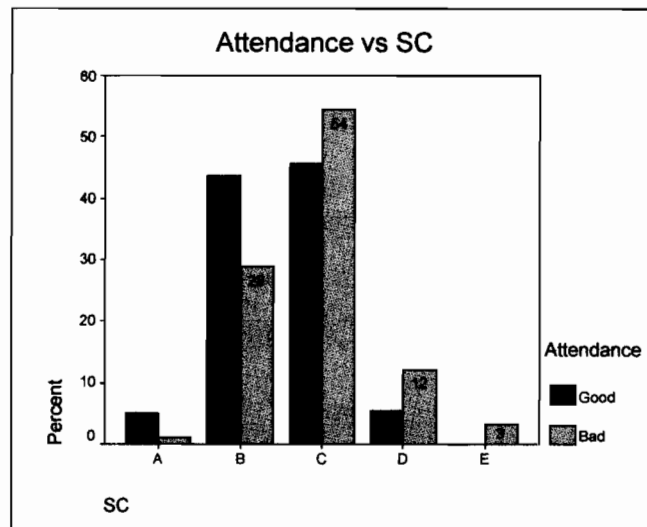


Figure 4.11: Relationship between attendance and Science

There were about 5% or 17 of good attendance students got A in Science, while 1% or 1 of bad attendance student got A in the same subject. 44% or 144 of good attendance students got B in SC while 29% or 26 of bad attendance students got B. About 46% or 151 of good attendance students and 54% or 49 of bad attendance students got C in that subject. There were 5% or 18 of good attendance students and 12% or 11 of bad attendance students got D. The lowest percent was given by grade E which was only 3% or 3 of bad attendance students. There was no good attendance students got E. The good attendance students gave the higher percent for grade A and B. Meanwhile, the bad attendance students gave the higher percent for grade C, D and E.

Co-curricular activities are one of the factors contributes to the students' performance. The survey had been carried out to determine the relationship between co-curricular activities and performance. The analysis of missing value is shown below:

Table 4.14: Co-curricular frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Co-curricular * BM1	420	100.0%	0	.0%	420	100.0%
Co-curricular * BM2	420	100.0%	0	.0%	420	100.0%
Co-curricular * BI	420	100.0%	0	.0%	420	100.0%
Co-curricular * MT	420	100.0%	0	.0%	420	100.0%
Co-curricular * SC	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The co-curricular allocation either active or passive as shown below:

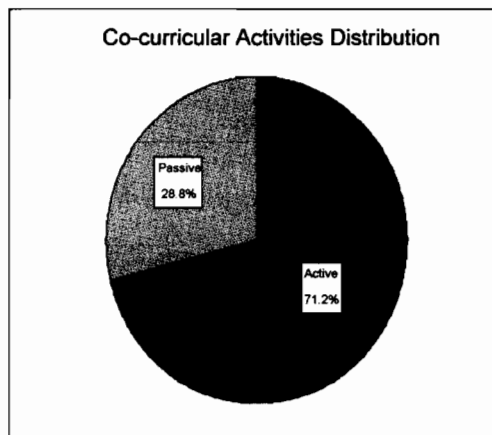


Figure 4.12: Co-curricular activities distribution

There are about 71.2% or 299 students actively involved in co-curricular activities. Meanwhile, 28.8% or 121 students are passively involved. The relationship between co-curricular activities and students' performance is shown below:

Table 4.15: Co-curricular and BM1 crosstabulation

CO-CURRICULAR * BM1 Crosstabulation										
BM1										
	A	% within BM1	B	% within BM1	C	% within BM1	D	% within BM1	Co-curricular	Total
Active	38	13%	122	41%	130	43%	9	3%	299	71%
Passive	5	4%	51	42%	55	45%	10	8%	121	29%
Total	43	10%	173	41%	185	44%	19	5%	420	100%

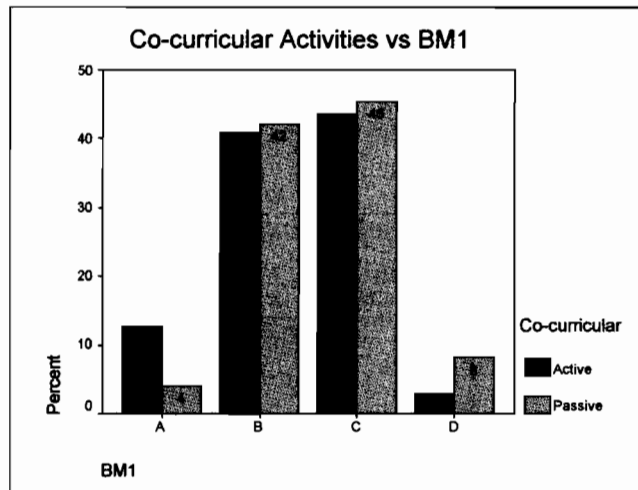


Figure 4.9: Relationship between co-curricular activities and BM1

There were about 13% or 38 of actively involved students got A in Bahasa Melayu 1 (Pemahaman), while there were 4% or 5 of passively involved students got A. Grade B in BM1 was presented by 41% or 122 of actively involved students and 42% or 51 of passively involved students. About 43% or 130 of actively involved students and 45% or 55 of passively involved students got C in that subject. There were 3% or 9 of

actively involved students and 8% or 10 of passively involved students got D. The highest percent was given by grade C while the lowest percent was given by grade D. The actively involved students gave the higher percent for grade A while passively involved students gave the higher percent for grade B, C and D.

Table 4.16: Co-curricular and BM2 crosstabulation

CO-CURRICULAR * BM2 Crosstabulation												
												BM2
	A	% within BM2	B	% within BM2	C	% within BM2	D	% within BM2	E	% within BM2	Co-curricular	Total
Active	15	5%	124	41%	130	43%	29	10%	1	0.3%	299	71%
Passive	0	0%	38	31%	69	57%	12	10%	2	2%	121	29%
Total	15	4%	162	39%	199	47%	41	10%	3	0.7%	420	100%

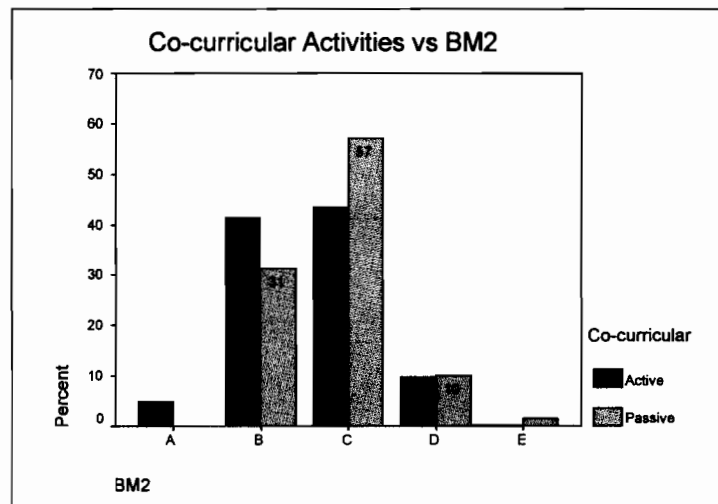


Figure 4.15: Relationship between co-curricular activities and BM2

There were about 5% or 15 of actively involved students got A in Bahasa Melayu 2 (Penulisan), while there was no passively involved students got A. Grade B in BM2 was presented by 41% or 124 of actively involved students and 31% or 38 of passively involved students. About 43% or 130 of actively involved students and 57% or 69 of

passively involved students got C in that subject. There were the same 10% or 29 of actively involved students and 12 of passively involved students got D. 0.3% or 1 of actively involved student and 2% or 2 of passively involved students got grade E. The highest percent was given by grade C while the lowest percent was given by grade E. The actively involved students gave the higher percent for grade A and B while passively involved students gave the higher percent for grade C and E.

Table 4.17: Co-curricular and BI crosstabulation

CO-CURRICULAR * BI Crosstabulation												
BI												
	A	% within BI	B	% within BI	C	% within BI	D	% within BI	E	% within BI	Co-curricular	Total
Active	16	5%	126	42%	122	41%	32	11%	3	1%	299	71%
Passive	0	0%	35	29%	67	55%	16	13%	3	2%	121	29%
Total	16	4%	161	38%	189	45%	48	11%	6	1%	420	100%

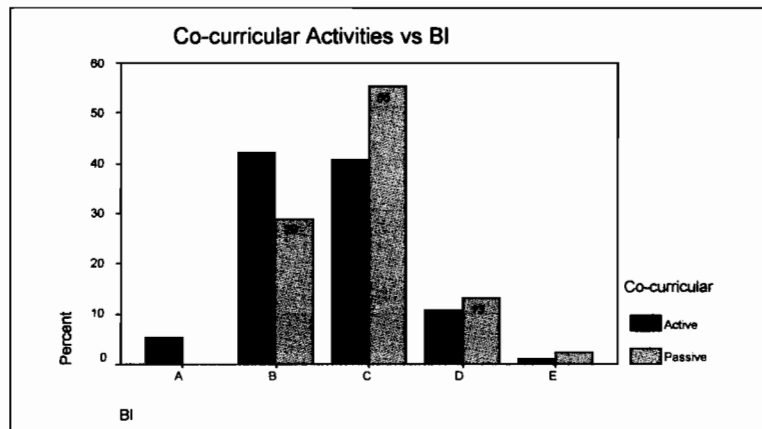


Figure 4.16: Relationship between co-curricular activities and BI

There were about 5% or 16 of actively involved students got A in English, while there was no passively involved students got A. Grade B in BI was presented by 42% or 126 of actively involved students and 29% or 35 of passively involved students. About 41%

or 122 of actively involved students and 55% or 67 of passively involved students got C in that subject. There were 11% or 32 of actively involved students and 13% or 16 of passively involved students got D. 1% or 3 of actively involved student and 2% or 3 of passively involved students got grade E. The highest percent was given by grade C while the lowest percent was given by grade E. The actively involved students gave the higher percent for grade A and B while passively involved students gave the higher percent for grade C, D and E.

Table 4.18: Co-curricular and MT crosstabulation

CO-CURRICULAR * MT Crosstabulation													
MT												Co-curricular	Total
	A	% within MT	B	% within MT	C	% within MT	D	% within MT	E	% within MT			
Active	9	3%	107	36%	144	48%	35	12%	4	1%	299	71%	
Passive	2	2%	26	21%	58	48%	32	26%	3	2%	121	29%	
Total	11	3%	133	32%	202	48%	67	16%	7	2%	420	100%	

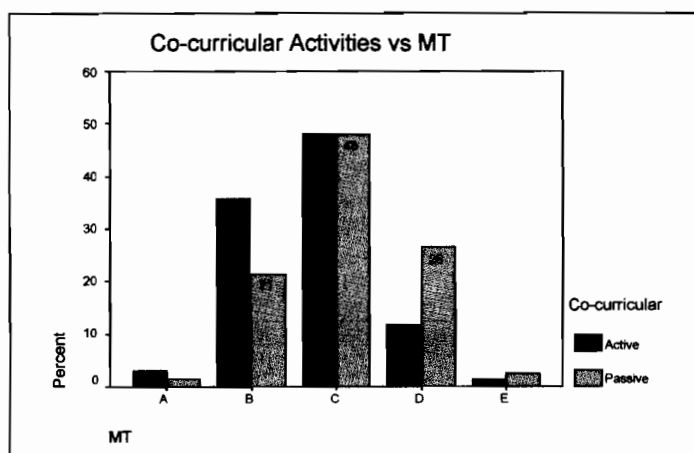


Figure 4.17: Relationship between co-curricular activities and MT

There were about 3% or 9 of actively involved students got A in Mathematics, while there were 2% or 2 of passively involved students got A. Grade B in MT was presented

by 36% or 107 of actively involved students and 21% or 26 of passively involved students. About 48% or 144 of actively involved students and also 48% or 58 of passively involved students got C in that subject. There were 12% or 35 of actively involved students and 26% or 32 of passively involved students got D. 1% or 4 of actively involved student and 2% or 3 of passively involved students got grade E. The highest percent was given by grade C while the lowest percent was given by grade E. The actively involved students gave the higher percent for grade A and B while the passively involved students gave the higher percent for grade D and E.

Table 4.19: Co-curricular and SC crosstabulation

CO-CURRICULAR * SC Crosstabulation												
SC												
	A	% within SC	B	% within SC	C	% within SC	D	% within SC	E	% within SC	Co-curricular	Total
Active	14	5%	135	45%	133	44%	17	6%	0	0%	299	71%
Passive	4	3%	35	29%	67	55%	12	10%	3	2%	121	29%
Total	18	4%	170	40%	200	48%	29	7%	3	1%	420	100%

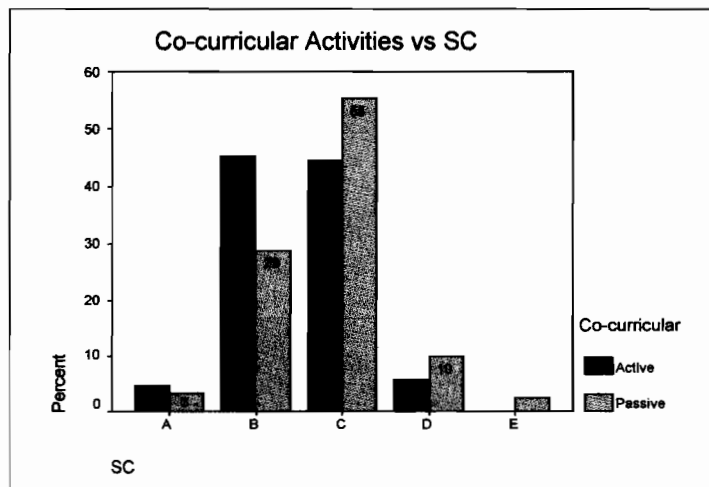


Figure 4.18: Relationship between co-curricular activities and SC

There were about 5% or 14 of actively involved students got A in Science, while there were 3% or 4 of passively involved students got A. Grade B in SC was presented by 45% or 135 of actively involved students and 29% or 35 of passively involved students. About 44% or 133 of actively involved students and 55% or 67 of passively involved students got C in that subject. There were 6% or 17 of actively involved students and 10% or 12 of passively involved students got D. There was no actively involved students got E while 2% or 3 of passively involved students got grade E. The highest percent was given by grade C while the lowest percent was given by grade E. The actively involved students gave the higher percent for grade A and B while the passively involved students gave the higher percent for grade C, D and E.

Position in family is one of the factors contributes to the students' performance. The survey had been carried out to determine the relationship between position in family and performance. The analysis of missing value is shown below:

Table 4.20: Position in family frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Position in Family * BM1	420	100.0%	0	.0%	420	100.0%
Position in Family * BM2	420	100.0%	0	.0%	420	100.0%
Position in Family * BI	420	100.0%	0	.0%	420	100.0%
Position in Family * MT	420	100.0%	0	.0%	420	100.0%
Position in Family * SC	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The position in family allocation is shown below:

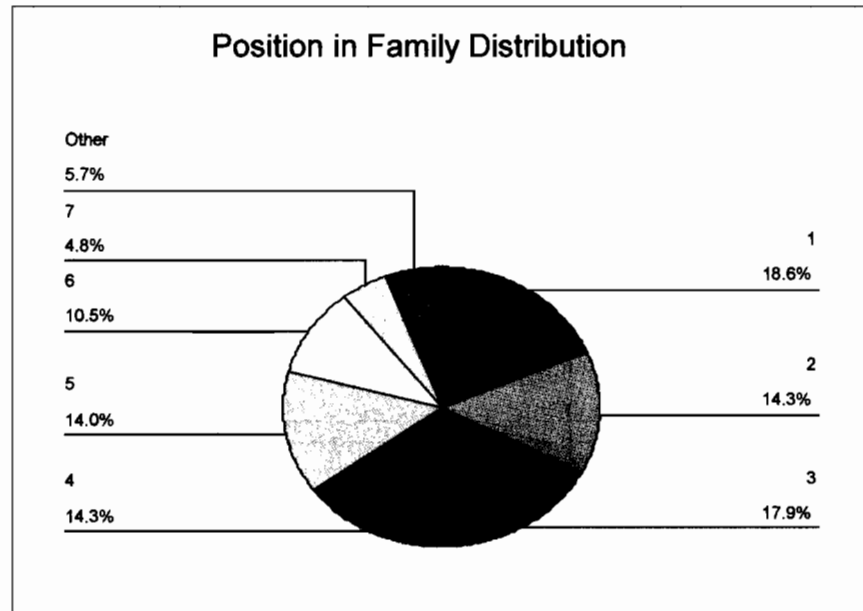


Figure 4.19: Position in family distribution

There were 18.6% or 78 of the students were the first children in the family while 14.3% or 60 students were the second children. About 17.9% or 75 students were the third children and 14.3% or 60 students were the forth children. The fifth children were presented by 14% or 59 students while the sixth children were only 10.5% or 44 students. The seventh children were 4.8% or 20 students and other than those positions were presented by 5.7% or 24 students. The highest percent was given by the first children and the lowest percent by the seventh children. The relationship between position in family and performance are shown below:

Table 4.21: Position in family and BM1 crosstabulation

POSITION IN FAMILY * BM1 Crosstabulation										
BM1										
	A	% within BM1	B	% within BM1	C	% within BM1	D	% within BM1	Position in Family	Total
1	8	10%	34	44%	31	40%	5	6%	78	19%
2	7	12%	25	42%	27	45%	1	2%	60	14%
3	6	8%	29	39%	37	49%	3	4%	75	18%
4	5	8%	22	37%	31	52%	2	3%	60	14%
5	7	12%	23	39%	22	37%	7	12%	59	14%
6	8	18%	22	50%	14	32%	0	0%	44	10%
7	1	5%	9	45%	10	50%	0	0%	20	5%
8	1	14%	2	29%	3	43%	1	14%	7	2%
9	0	0%	3	43%	4	57%	0	0%	7	2%
>=10	0	0%	4	40%	6	60%	0	0%	10	2%
Total	43	10%	173	41%	185	44%	19	5%	420	100%

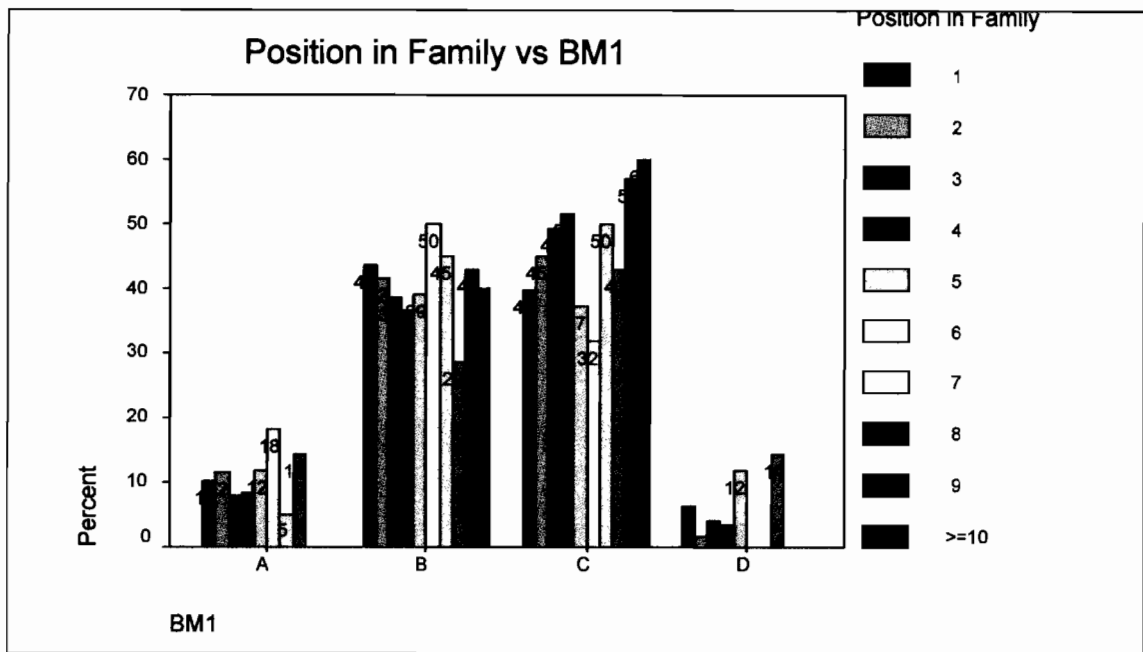


Figure 4.20: Relationship between position in family and BM1

There were about 10% or 8 students of the first children, 12% or 7 of the second children, both 8% or 6 and 5 of the third and forth children, 12% or 7 of the fifth children, 18% or 8 of the sixth children, 5% or 1 of the seventh children and 14% or 1 of the eighth children got A in Bahasa Melayu 1 (Pemahaman). Grade B was presented by

44% or 34 of the first children, 42% or 25 of the second children, 39% or 29 of the third children, 37% or 22 of the fourth children, 39% or 23 of the fifth children, 50% or 22 of the sixth children, 45% or 9 of the seventh children, 29% or 2 of the eighth children, 43% or 3 of the ninth children and 40% or 4 of the tenth or more children. 40% or 31 of the first children, 45% or 27 of the second children, 49 or 37 of the third children, 52% or 31 of the fourth children, 37% or 22 of the fifth children, 32% or 14 of the sixth children, 50% of the seventh children, 43% or 3 of the eighth children, 57% or 4 of the ninth children and 60% or 6 of the tenth or more children got C. Grade D was presented by 6% or 5 of the first children, 2% or 1 of the second child, 4% or 3 of the third children, 3% or 2 of the fourth children, 12% or 7 of the fifth children and 14% or 1 of the eighth child. The highest percent for grade A was given by the sixth children while the highest percent for grade D was eighth children.

Table 4.22: Position in family and BM2 crosstabulation

POSITION IN FAMILY * BM2 Crosstabulation												
BM2												
	A	% within BM2	B	% within BM2	C	% within BM2	D	% within BM2	E	% within BM2	Position in Family	Total
1	4	5%	31	40%	35	45%	8	10%	0	0%	78	19%
2	2	3%	25	42%	26	43%	7	12%	0	0%	60	14%
3	2	3%	27	36%	39	52%	6	8%	1	1%	75	18%
4	1	2%	21	35%	33	55%	4	7%	1	2%	60	14%
5	3	5%	23	39%	25	42%	7	12%	1	2%	59	14%
6	3	7%	21	48%	18	41%	2	5%	0	0%	44	10%
7	0	0%	8	40%	9	45%	3	15%	0	0%	20	5%
8	0	0%	0	0%	6	86%	1	14%	0	0%	7	2%
9	0	0%	2	29%	3	43%	2	29%	0	0%	7	2%
>=10	0	0%	4	40%	5	50%	1	10%	0	0%	10	2%
Total	15	4%	162	39%	199	47%	41	10%	3	1%	420	100%

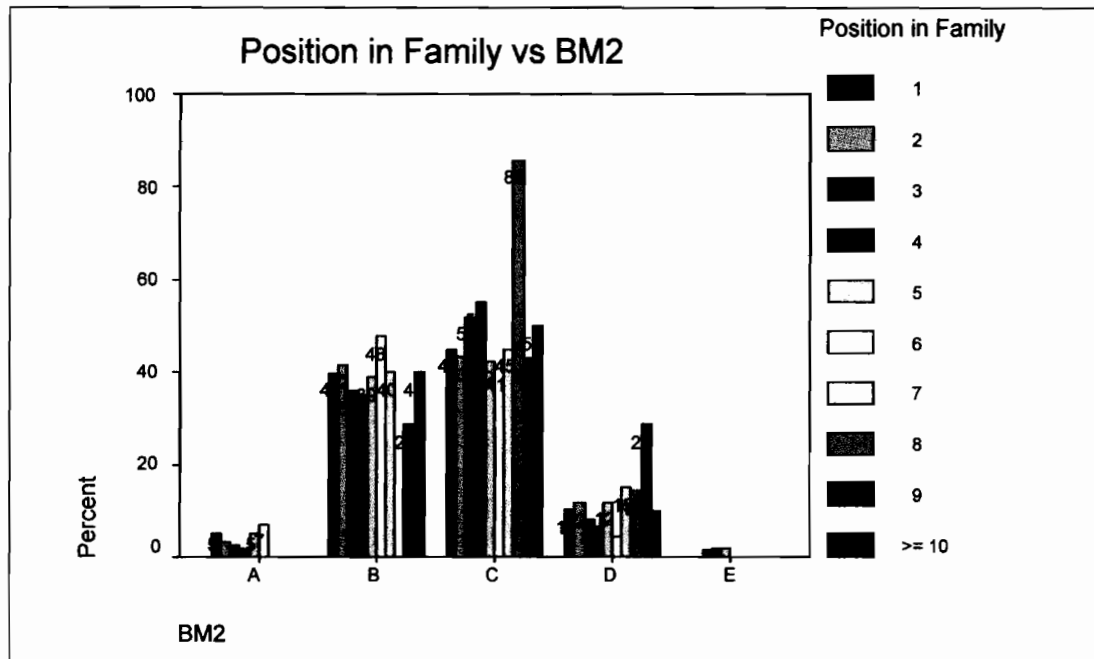


Figure 4.21: Relationship between position in family and BM2

There were about 5% or 4 students of the first children, 3% or 2 of the second children and third children. 2% or 1 of the forth child, 5% or 3 of the fifth children and 7% or 3 of the sixth children got A in Bahasa Melayu2 (Penulisan). Grade B was presented by 40% or 31 of the first children, 42% or 25 of the second children, 36% or 27 of the third children, 35% or 21 of the forth children, 39% or 23 of the fifth children, 48% or 21 of the sixth children, 40% or 8 of the seventh children, 29% or 2 of the ninth children and 40% or 4 of the tenth or more children. 45% or 35 of the first children, 43% or 26 of the second children, 52 or 39 of the third children, 55% or 33 of the forth children, 42% or 25 of the fifth children, 41% or 18 of the sixth children, 45% or 9 of the seventh children, 86% or 6 of the eighth children, 43% or 3 of the ninth children and 50% or 5 of the tenth or more children got C. Grade D was presented by 10% or 8 of the first children, 12% or 7 of the second children, 8% or 6 of the third children, 7% or 4 of the forth children, 12% or 7 of the fifth children, 5% or 2 of the sixth children, 15% or 3 of the seventh children, 14% or 1 of the eighth child, 29% or 2 of the ninth children and 10% or 1 of the tenth child or more. 1% or 1 of the third child and 2% or 1 of the forth and fifth children got grade E. The highest percent for grade A was given by the sixth children while the highest percent for grade E was forth and fifth children.

Table 4.23: Position in family and BI crosstabulation

POSITION IN FAMILY * BI Crosstabulation												
BI												
	A	% within BI	B	% within BI	C	% within BI	D	% within BI	E	% within BI	Position in Family	Total
1	5	6%	36	46%	34	44%	3	4%	0	0%	78	19%
2	4	7%	15	25%	33	55%	8	13%	0	0%	60	14%
3	3	4%	28	37%	36	48%	7	9%	1	1%	75	18%
4	1	2%	22	37%	29	48%	5	8%	3	5%	60	14%
5	3	5%	23	39%	19	32%	12	20%	2	3%	59	14%
6	0	0%	18	41%	22	50%	4	9%	0	0%	44	10%
7	0	0%	12	60%	6	30%	2	10%	0	0%	20	5%
8	0	0%	2	29%	3	43%	2	29%	0	0%	7	2%
9	0	0%	2	29%	1	14%	4	57%	0	0%	7	2%
>=10	0	0%	3	30%	6	60%	1	10%	0	0%	10	2%
Total	16	4%	161	38%	189	45%	48	11%	6	1%	420	100%

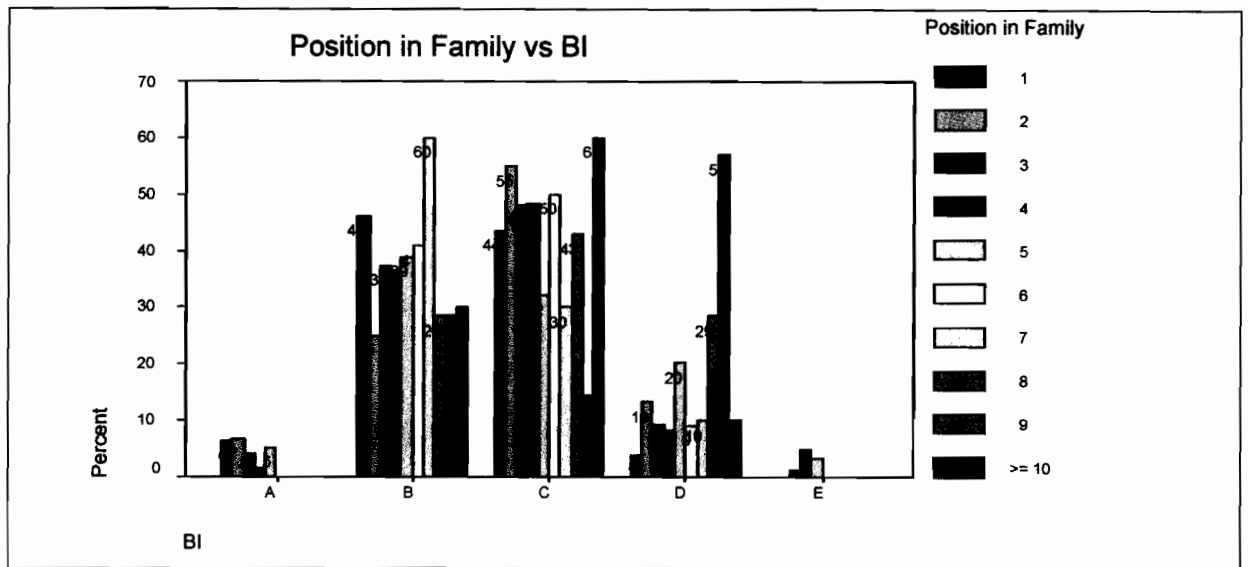


Figure 4.22: Relationship between position in family and BI

There were about 6% or 5 students of the first children, 7% or 4 of the second children, 4% or 3 of the third children, 2% or 1 of the forth child and 5% or 3 of the fifth children got A in English. Grade B was presented by 46% or 36 of the first children, 25% or 15 of the second children, 37% or 28 of the third children, 37% or 22 of the forth children, 39% or 23 of the fifth children, 41% or 18 of the sixth children, 60% or 12 of the seventh children, 29% or 2 of the eighth and ninth children and 30% or 3 of the tenth or

more children. 44% or 34 of the first children, 55% or 33 of the second children, 48% or 36 of the third children, 48% or 29 of the forth children, 32% or 19 of the fifth children, 50% or 22 of the sixth children, 30% or 6 of the seventh children, 43% or 3 of the eighth children, 14% or 1 of the ninth child and 60% or 6 of the tenth or more children got C. Grade D was presented by 4% or 3 of the first children, 13% or 8 of the second children, 9% or 7 of the third children, 8% or 5 of the forth children, 20% or 12 of the fifth children, 9% or 4 of the sixth children, 10% or 2 of the seventh children, 29% or 2 of the eighth children, 57% or 4 of the ninth children and 10% or 1 of the tenth child or more. 1% or 1 of the third child, 5% or 3 of the forth children and 3% or 2 of the fifth children got grade E. The highest percent for grade A was given by the second children while the highest percent for grade E was forth children.

Table 4.24: Position in family and MT crosstabulation

POSITION IN FAMILY * MT Crosstabulation												
MT												
	A	% within MT	B	% within MT	C	% within MT	D	% within MT	E	% within MT	Position in Family	Total
1	4	5%	27	35%	31	40%	15	19%	1	1%	78	19%
2	2	3%	15	25%	29	48%	14	23%	0	0%	60	14%
3	2	3%	19	25%	40	53%	11	15%	3	4%	75	18%
4	2	3%	20	33%	29	48%	8	13%	1	2%	60	14%
5	1	2%	19	32%	30	51%	9	15%	0	0%	59	14%
6	0	0%	15	34%	24	55%	3	7%	2	5%	44	10%
7	0	0%	10	50%	6	30%	4	20%	0	0%	20	5%
8	0	0%	4	57%	3	43%	0	0%	0	0%	7	2%
9	0	0%	3	43%	3	43%	1	14%	0	0%	7	2%
>=10	0	0%	1	10%	7	70%	2	20%	0	0%	10	2%
Total	11	3%	133	32%	202	48%	67	16%	7	2%	420	100%

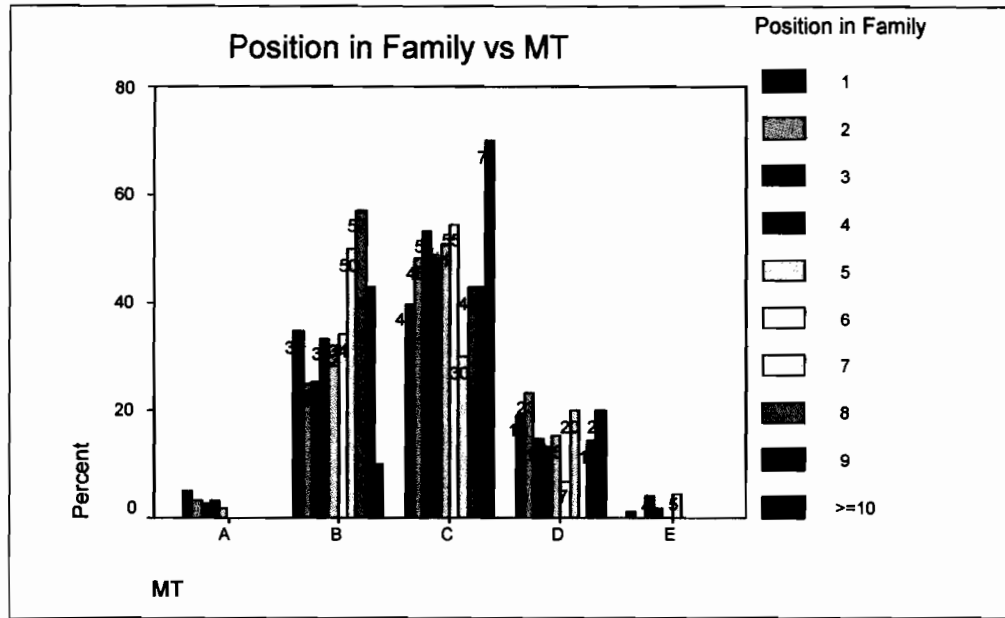


Figure 4.23: Relationship between position in family and MT

There were about 5% or 4 students of the first children, 3% or 2 of the second, third children and fourth children and 2% or 1 of the fifth child got A in Mathematics. Grade B was presented by 35% or 27 of the first children, 25% or 15 of the second children, 25% or 19 of the third children, 33% or 20 of the fourth children, 32% or 19 of the fifth children, 34% or 15 of the sixth children, 50% or 10 of the seventh children, 57% or 4 of the eighth children, 43% or 3 of the ninth children and 10% or 1 of the tenth or more child. 40% or 31 of the first children, 48% or 29 of the second children, 53 or 40 of the third children, 48% or 29 of the fourth children, 51% or 30 of the fifth children, 55% or 24 of the sixth children, 30% or 6 of the seventh children, 43% or 3 of the eighth and ninth children and 70% or 7 of the tenth or more children got C. Grade D was presented by 19% or 15 of the first children, 23% or 14 of the second children, 15% or 11 of the third children, 13% or 8 of the fourth children, 15% or 9 of the fifth children, 7% or 3 of the sixth children, 20% or 4 of the seventh children, 14% or 1 of the ninth child and 20% or 2 of the tenth children or more. 1% or 1 of the first child, 4% or 3 of the third children, 2% or 1 of the fourth child and 5% or 2 of the sixth children got grade E. The highest percent for grade A was given by the first children while the highest percent for grade E was sixth children.

Table 4.25: Position in family and SC crosstabulation

POSITION IN FAMILY * SC Crosstabulation												
SC												
	A	% within SC	B	% within SC	C	% within SC	D	% within SC	E	% within SC	Position in Family	Total
1	4	5%	31	40%	38	49%	4	5%	1	1%	78	19%
2	3	5%	29	48%	26	43%	2	3%	0	0%	60	14%
3	3	4%	29	39%	40	53%	3	4%	0	0%	75	18%
4	4	7%	22	37%	28	47%	5	8%	1	2%	60	14%
5	2	3%	23	39%	27	46%	6	10%	1	2%	59	14%
6	0	0%	22	50%	18	41%	4	9%	0	0%	44	10%
7	1	5%	6	30%	10	50%	3	15%	0	0%	20	5%
8	0	0%	1	14%	6	86%	0	0%	0	0%	7	2%
9	0	0%	4	57%	2	29%	1	14%	0	0%	7	2%
>=10	1	10%	3	30%	5	50%	1	10%	0	0%	10	2%
Total	18	4%	170	40%	200	48%	29	7%	3	1%	420	100%

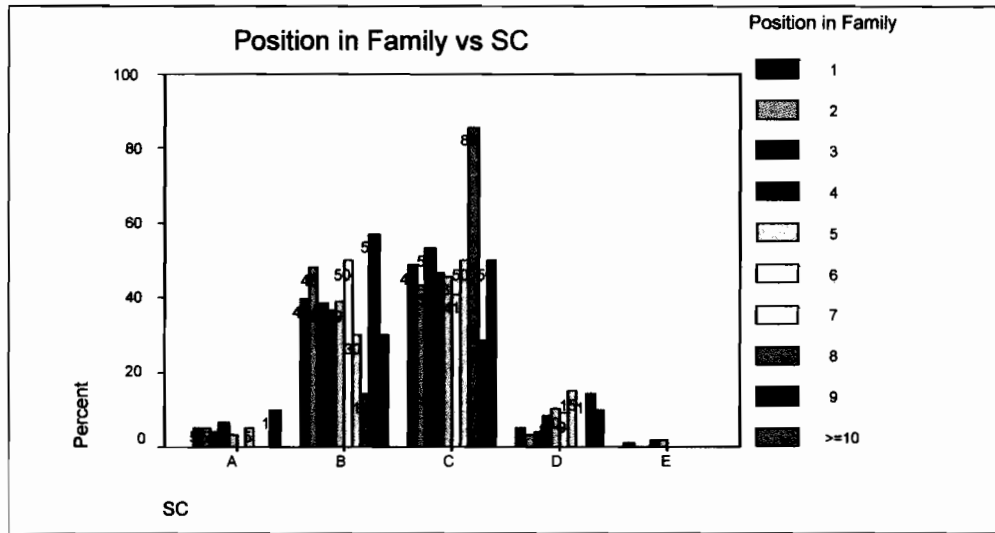


Figure 4.24: Relationship between position in family and SC

There were about 5% or 4 students of the first children, 5% or 3 of the second children, 4% or 3 of the third children, 7% or 4 of the fourth children, 3% or 2 of the fifth children, 5% or 1 of the seventh child and 10% or 1 of the tenth child or more got A in Science. Grade B was presented by 40% or 31 of the first children, 48% or 29 of the second children, 39% or 29 of the third children, 37% or 22 of the fourth children, 39% or 23 of the fifth children, 50% or 22 of the sixth children, 30% or 6 of the seventh children, 14% or 1 of the ninth child and 30% or 3 of the tenth or more children. 49% or 38 of the first

children, 43% or 26 of the second children, 53 or 40 of the third children, 47% or 28 of the fourth children, 46% or 27 of the fifth children, 41% or 18 of the sixth children, 50% or 10 of the seventh children, 86% or 6 of the eighth children, 29% or 2 of the ninth children and 50% or 5 of the tenth or more children got C. Grade D was presented by 5% or 4 of the first children, 3% or 2 of the second children, 4% or 3 of the third children, 8% or 5 of the fourth children, 10% or 6 of the fifth children, 9% or 4 of the sixth children, 15% or 3 of the seventh children, 14% or 1 of the ninth child and 10% or 1 of the tenth child or more. 1% or 1 of the first child, 2% or 1 of the fourth and fifth child got grade E. The highest percent for grade A was given by the tenth children or more while the highest percent for grade E was fourth and fifth children.

Father's job is one of the factors contributes to the students' performance. The survey had been carried out to determine the relationship between father's job and performance.

The analysis of missing value is shown below:

Table 4.26: Father's job frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Father's Job * BM1	420	100.0%	0	.0%	420	100.0%
Father's Job * BM2	420	100.0%	0	.0%	420	100.0%
Father's Job * BI	420	100.0%	0	.0%	420	100.0%
Father's Job * MT	420	100.0%	0	.0%	420	100.0%
Father's Job * SC	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The father's job allocation is shown below:

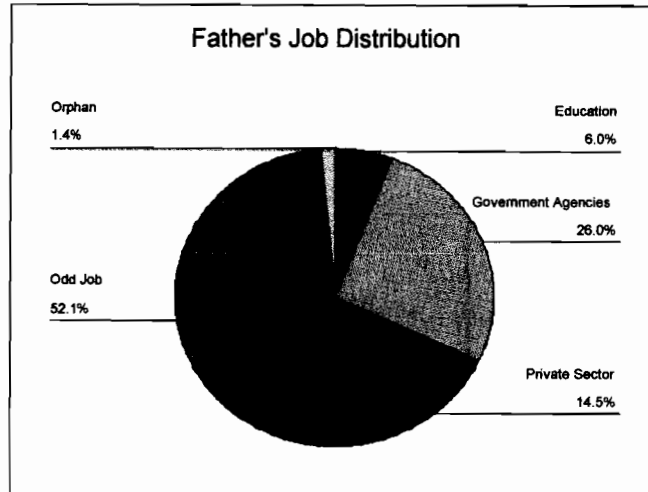


Figure 4.25: Father's job distribution

Odd job like rubber taper, farmer, fisherman and self jobs presented the highest percent (52.1%). The government agencies such as officer, banker, administrative and professional contributed 26%. About 14.5% of fathers' job was in private sectors like business, industrial workers and companies. There were 6.0% of the fathers' job was in education field such as teacher and lecturer. The lowest percent was given by orphan students (1.4%) only.

Table 4.27: Father's job and BM1 crosstabulation

FATHER'S JOB * BM1 Crosstabulation										
BM1										
	A	% within BM1	B	% within BM1	C	% within BM1	D	% within BM1	Father's Job	Total
Education	4	16%	10	40%	11	44%	0	0%	25	6%
Government Agencies	12	11%	51	47%	41	38%	0	0%	109	26%
Private Sector	10	16%	24	39%	23	38%	5	8%	61	15%
Odd Job	17	8%	86	39%	106	48%	4	2%	219	52%
Orphan	0	0%	2	33%	4	67%	0	0%	6	1%
Total	43	10%	173	41%	185	44%	19	5%	420	100%

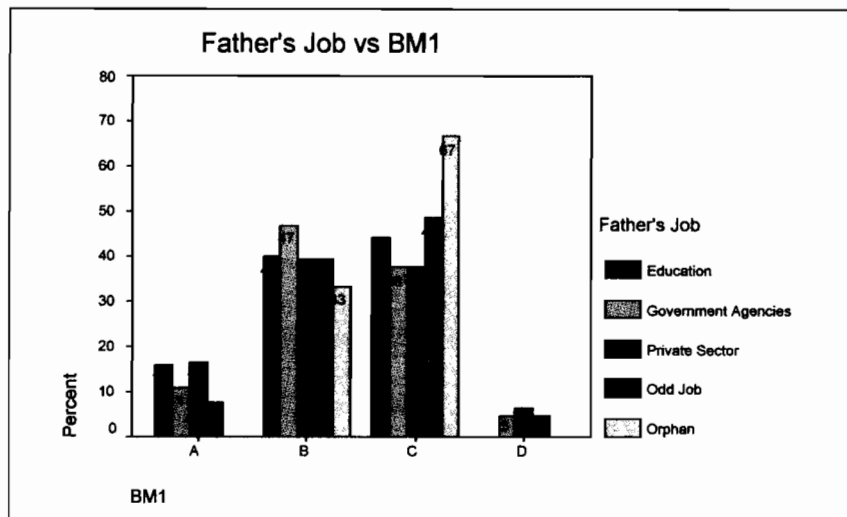


Figure 4.26: Relationship between father's job and BM1

There were 16% or 4 of fathers' job in education, 11% or 12 in government agencies, 16% or 10 in private sector and 8% or 17 as odd jobs got A in Bahasa Melayu 1 (Pemahaman). Grade B was presented by 40% or 10 fathers' job in education, 47% or 51 in government agencies, 39% or 24 in private sector, 39% or 86 as odd jobs and 33% or 2 were orphans students. About 44% or 11 fathers' job in education, 38% or 41 in government agencies, 38% or 23 in private sectors, 48% or 106 as odd jobs and 67% or 4 were orphans students got C. 8% or 5 fathers' job in private sectors and 2% or 4 as odd jobs got grade D. The highest percent for grade A was given by fathers' job in education and private sectors. The highest percent for grade D was fathers' job in private sectors.

Table 4.28: Father's job dan BM2 crosstabulation

FATHER'S JOB * BM2 Crosstabulation												
BM2												
	A	% within BM2	B	% within BM2	C	% within BM2	D	% within BM2	E	% within BM2	Father's Job	Total
Education	2	8%	14	56%	8	32%	1	4%	0	0%	25	6%
Government Agencies	6	6%	45	41%	50	46%	6	6%	2	2%	109	26%
Private Sector	4	7%	20	33%	33	54%	4	7%	0	0%	61	15%
Odd Job	3	1%	81	37%	104	47%	30	14%	1	0.5%	219	52%
Orphan	0	0%	2	33%	4	67%	0	0%	0	0%	6	1%
Total	15	4%	162	39%	199	47%	41	10%	3	1%	420	100%

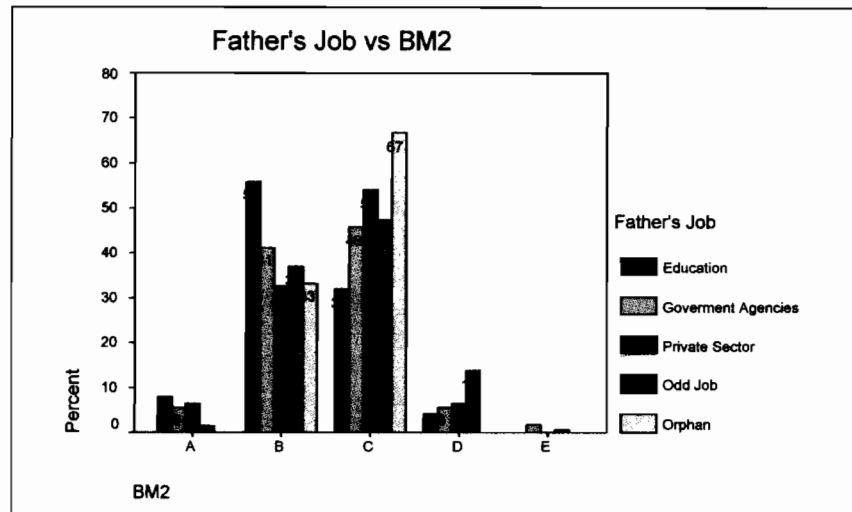


Figure 4.27: Relationship between father's job and BM2

There were 8% or 2 of fathers' job in education, 6% or 6 in government agencies, 7% or 4 in private sector and 1% or 3 as odd jobs got A in Bahasa Melayu 2 (Penulisan). Grade B was presented by 56% or 14 of fathers' job in education, 41% or 45 in government agencies, 33% or 20 in private sectors, 37% or 81 as odd jobs and 33% or 2 were orphans students. About 32% or 8 fathers' job in education, 46% or 50 in government agencies, 54% or 33 in private sectors, 47% or 104 as odd jobs and 67% or 4 were orphans students got C. 4% or 1 fathers' job in education, 6% or 6 in

government agencies, 7% or 4 in private sectors and 14% or 30 as odd jobs got grade D. 2% or 2 fathers' job in government agencies and 0.5% or 1 as odd job got grade E. The highest percent for grade A was given by fathers' job in education. The highest percent for grade E was fathers' job in government agencies.

Table 4.29: Father's job and BI crosstabulation

FATHER'S JOB * BI Crosstabulation												
BI												
	A	% within BI	B	% within BI	C	% within BI	D	% within BI	E	% within BI	Father's Job	Total
Education	3	12%	13	52%	7	28%	2	8%	0	0%	25	6%
Government Agencies	5	5%	49	45%	43	39%	10	9%	2	2%	109	26%
Private Sector	6	10%	21	34%	30	49%	4	7%	0	0%	61	15%
Odd Job	2	1%	75	34%	107	49%	31	14%	4	2%	219	52%
Orphan	0	0%	3	50%	2	33%	1	17%	0	0%	6	1%
Total	16	4%	161	38%	189	45%	48	11%	6	1%	420	100%

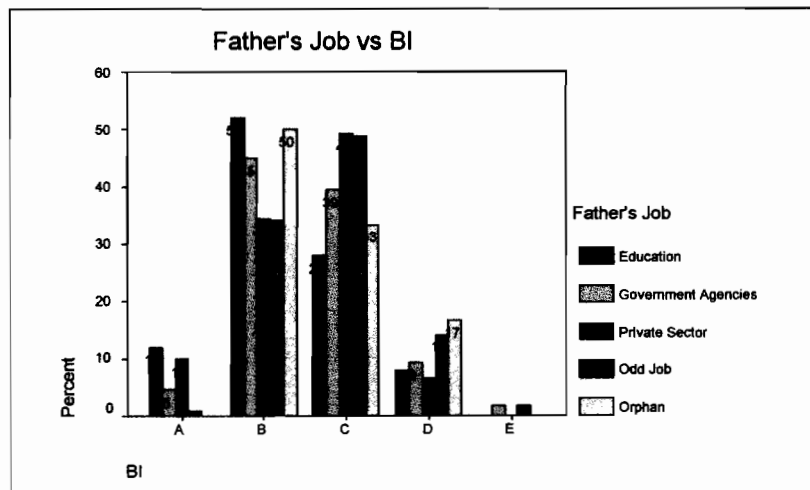


Figure 4.28: Relationship between father's job and BI

There were 12% or 3 of fathers' job in education, 5% or 5 in government agencies, 10% or 6 in private sector and 1% or 2 as odd jobs got A in English. Grade B was presented by 52% or 13 of fathers' job in education, 45% or 49 in government agencies, 34% or

21 in private sectors, 34% or 75 as odd jobs and 50% or 3 were orphans students. About 28% or 7 of fathers' job in education, 39% or 43 in government agencies, 49% or 30 in private sectors, 49% or 107 as odd jobs and 33% or 2 were orphans students got C. 8% or 2 of fathers' job in education, 9% or 10 in government agencies, 7% or 4 in private sectors, 14% or 31 as odd jobs and 17% or 1 was orphan student got grade D. 2% or 2 fathers' job in government agencies and 2% or 4 as odd jobs got grade E. The highest percent for grade A was given by fathers' job in education. The highest percent for grade E was fathers' job in government agencies and odd jobs.

Table 4.30: Father's job and MT crosstabulation

FATHER'S JOB * MT Crosstabulation													
MT												Father's Job	Total
A	% within MT	B	% within MT	C	% within MT	D	% within MT	E	% within MT				
Education	4	16%	7	28%	12	48%	1	4%	1	4%	25	6%	
Government Agencies	3	3%	39	36%	50	46%	14	13%	3	3%	109	26%	
Private Sector	3	5%	20	33%	29	48%	9	15%	0	0%	61	15%	
Odd Job	1	0.5%	65	30%	108	49%	42	19%	3	1%	219	52%	
Orphan	0	0%	2	33%	3	50%	1	17%	0	0%	6	1%	
Total	11	3%	133	32%	202	48%	67	16%	7	2%	420	100%	

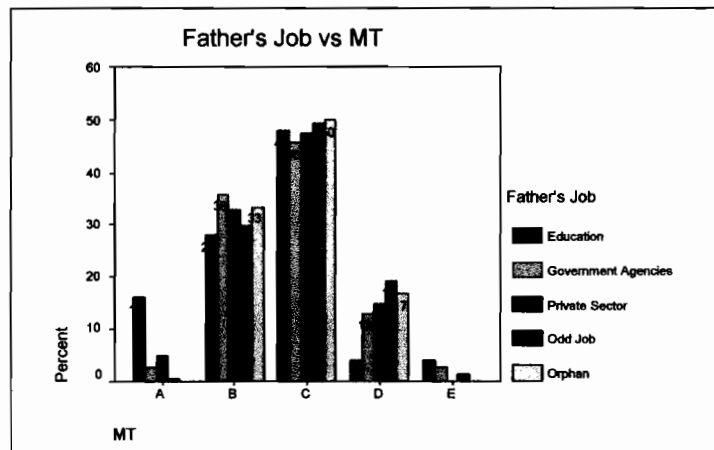


Figure 4.29: Relationship between father's job and MT

There were 16% or 4 of fathers' job in education, 3% or 3 in government agencies, 5% or 3 in private sector and 0.5% or 1 as odd job got A in Mathematics. Grade B was presented by 28% or 7 of fathers' job in education, 36% or 39 in government agencies, 33% or 20 in private sectors, 30% or 65 as odd jobs and 33% or 2 were orphans students. About 48% or 12 of fathers' job in education, 46% or 50 in government agencies, 48% or 29 in private sectors, 49% or 108 as odd jobs and 50% or 3 were orphans students got C. 4% or 1 of fathers' job in education, 13% or 14 in government agencies, 15% or 9 in private sectors, 19% or 42 as odd jobs and 17% or 1 was orphan student got grade D. 4% or 1 of fathers' job in education, 3% or 3 in government agencies and 1% or 3 as odd jobs got grade E. The highest percent for grade A and E were given by fathers' job in education.

Table 4.31: Father's job and SC crosstabulation

FATHER'S JOB * SC Crosstabulation												
SC												
	A	% within SC	B	% within SC	C	% within SC	D	% within SC	E	% within SC	Father's Job	Total
Education	3	12%	8	32%	12	48%	2	8%	0	0%	25	6%
Government Agencies	6	6%	37	34%	54	50%	11	10%	1	1%	109	26%
Private Sector	3	5%	26	43%	27	44%	5	8%	0	0%	61	15%
Odd Job	6	3%	97	44%	103	47%	11	5%	2	1%	219	52%
Orphan	0	0%	2	33%	4	67%	0	0%	0	0%	6	1%
Total	18	4%	170	40%	200	48%	29	7%	3	1%	420	100%

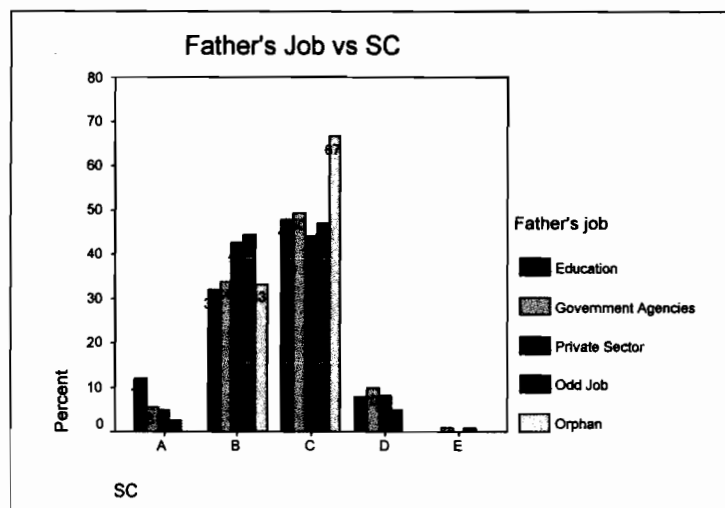


Figure 4.30: Relationship between father's job and SC

There were 12% or 3 of fathers' job in education, 6% or 6 in government agencies, 5% or 3 in private sectors and 3% or 6 as odd jobs got A in Science. Grade B was presented by 32% or 8 of fathers' job in education, 34% or 37 in government agencies, 43% or 26 in private sectors, 44% or 97 as odd jobs and 33% or 2 were orphans students. About 48% or 12 of fathers' job in education, 50% or 54 in government agencies, 44% or 27 in private sectors, 47% or 103 as odd jobs and 67% or 4 were orphans students got C. 8% or 2 of fathers' job in education, 10% or 11 in government agencies, 8% or 5 in private sectors, and 5% or 11 as odd jobs got grade D. 1% or 1 fathers' job in government agencies and 1% or 2 as odd jobs got grade E. The highest percent for grade A was given by fathers' job in education. The highest percent for grade E was fathers' job in government agencies and odd jobs.

Mother's job is one of the factors that contributes to the students' performance. The survey had been carried out to determine the relationship between mother's job and performance. The analysis of missing value is shown below:

Table 4.32: Mother's job frequency table

Case Processing Summary						
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
	Mother's Job * BM1	420	100.0%	0	.0%	420
Mother's Job * BM2	420	100.0%	0	.0%	420	100.0%
Mother's Job * BI	420	100.0%	0	.0%	420	100.0%
Mother's Job * MT	420	100.0%	0	.0%	420	100.0%
Mother's Job * SC	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The mother's job allocation is shown below:

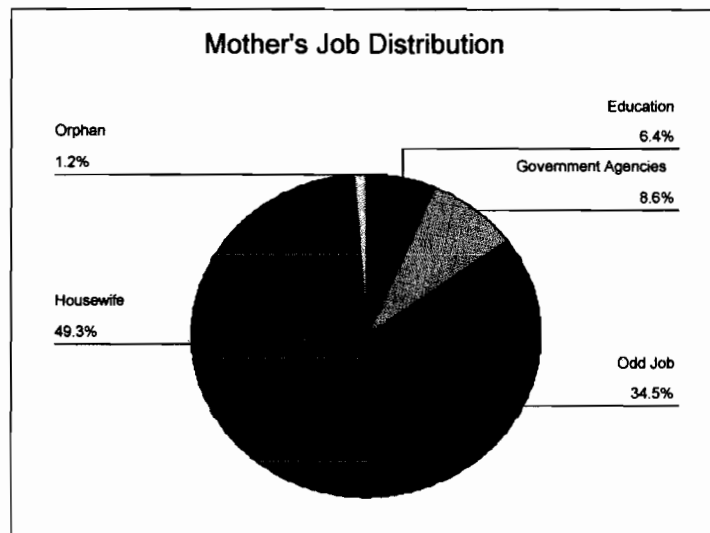


Figure 4.31: Mother's job distribution

The biggest percent for mothers' job was as housewife (49.3%). Odd job like rubber taper, farmer, fisherman and all self jobs presented 34.5%. The government agencies such as officer, banker, administrative and professional contributed 8.6%. There were 6.4% of the mothers' ' job was in education field such as teacher and lecturer. The lowest percent was given by orphan students (1.2%) only.

Table 4.33: Mother's job and BM1 crosstabulation

MOTHER'S JOB * BM1 Crosstabulation										
BM1										
	A	% within BM1	B	% within BM1	C	% within BM1	D	% within BM1	Mother's Job	Total
Education	5	19%	10	37%	11	41%	1	4%	27	6%
Government Agencies	5	14%	19	53%	12	33%	0	0%	36	9%
Odd Job	12	8%	53	37%	74	51%	6	4%	145	35%
Housewife	21	10%	89	43%	86	42%	11	5%	207	49%
Orphan	0	0%	2	40%	2	40%	1	20%	5	1%
Total	43	10%	173	41%	185	44%	19	5%	420	100%

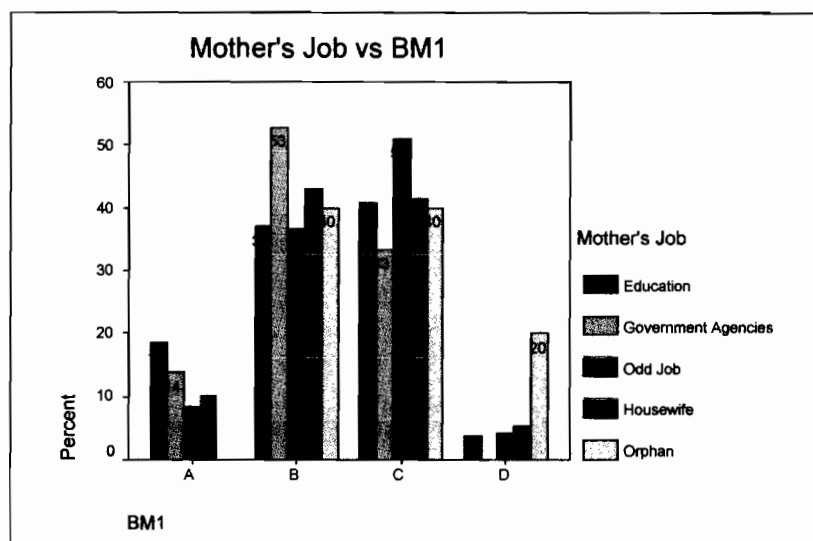


Figure 4.32: Relationship between mother's job and BM1

There were 19% or 5 of mothers' job in education, 14% or 5 in government agencies, 8% or 12 as odd jobs and 10% or 21 as housewives got A in Bahasa Melayu 1 (Pemahaman). Grade B was presented by 37% or 10 of mothers' job in education, 53% or 19 in government agencies, 37% or 53 as odd jobs, 43% or 89 as housewives and 40% or 2 were orphans students. About 41% or 11 of mothers' job in education, 33% or 12 in government agencies, 51% or 74 as odd jobs, 42% or 86 as housewives and 40% or 2 were orphans students got C. 4% or 1 of mothers' job in education, 4% or 6 as odd jobs,

5% or 11 as housewives and 20% or 1 was orphan student. The highest percent for grade A was given by mothers' job in education. The highest percent for grade D was orphan student.

Table 4.34: Mother's job and BM2 crosstabulation

MOTHER'S JOB * BM2 Crosstabulation												
BM2												
	A	% within BM2	B	% within BM2	C	% within BM2	D	% within BM2	E	% within BM2	Mother's Job	Total
Education	2	7%	15	56%	10	37%	0	0%	0	0%	27	6%
Government Agencies	1	3%	18	50%	14	39%	3	8%	0	0%	36	9%
Odd Job	3	2%	56	39%	66	46%	19	13%	1	1%	145	35%
Housewife	9	4%	70	34%	108	52%	18	9%	2	1%	207	49%
Orphan	0	0%	3	60%	1	20%	1	20%	0	0%	5	1%
Total	15	4%	162	39%	199	47%	41	10%	3	1%	420	100%

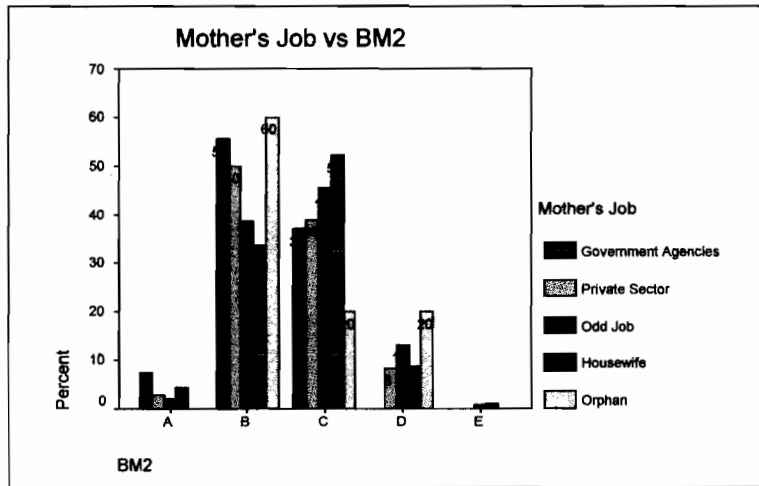


Figure 4.33: Relationship between mother's job and BM2

There were 7% or 2 of mothers' job in education, 3% or 1 in government agency, 2% or 3 as odd jobs and 4% or 9 as housewives got A in Bahasa Melayu 2 (Penulisan). Grade B was presented by 56% or 15 of mothers' job in education, 50% or 18 in government agencies, 39% or 56 as odd jobs, 34% or 70 as housewives and 60% or 3 were orphans

students. About 37% or 10 of mothers' job in education, 39% or 14 in government agencies, 46% or 66 as odd jobs, 52% or 108 as housewives and 20% or 1 was orphan student got C. 8% or 3 of mothers' job in government agencies, 13% or 19 as odd jobs, 9% or 18 as housewives and 20% or 1 was orphan student achieved grade D. About 1% or 1 of mothers' job as odd jobs and 1% or 2 as housewives got E. The highest percent for grade A was given by mothers' job in education. The highest percent for grade E was odd jobs and housewives.

Table 4.35: Mother's job and BI crosstabulation

MOTHER'S JOB * BI Crosstabulation												
BI												
	A	% within BI	B	% within BI	C	% within BI	D	% within BI	E	% within BI	Mother's Job	Total
Education	4	15%	14	52%	6	22%	2	7%	1	4%	27	6%
Government Agencies	1	3%	16	44%	17	47%	2	6%	0	0%	36	9%
Odd Job	1	1%	54	37%	68	47%	18	12%	4	3%	145	35%
Housewife	10	5%	75	36%	95	46%	26	13%	1	0.5%	207	49%
Orphan	0	0%	2	40%	3	60%	0	0%	0	0%	5	1%
Total	16	4%	161	38%	189	45%	48	11%	6	1%	420	100%

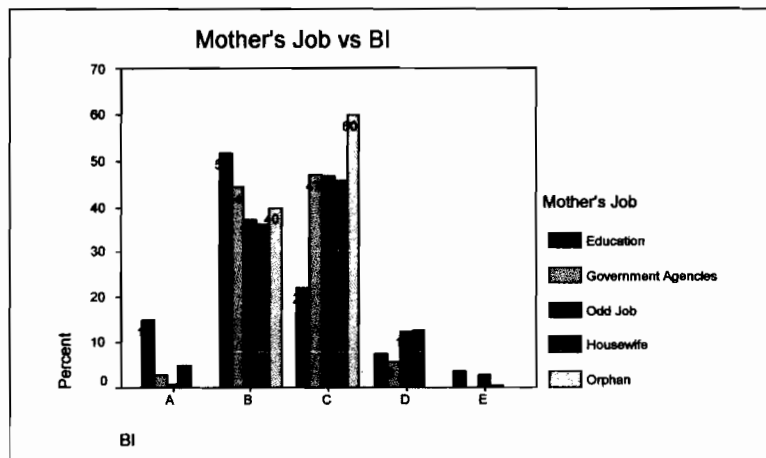


Figure 4.34: Relationship between mother's job and BI

There were 15% or 4 of mothers' job in education, 3% or 1 in government agency, 1% or 1 as odd job and 5% or 10 as housewives got A in English. Grade B was presented by 52% or 14 of mothers' job in education, 44% or 16 in government agencies, 37% or 54 as odd jobs, 36% or 75 as housewives and 40% or 2 were orphans students. About 22% or 6 of mothers' job in education, 47% or 17 in government agencies, 47% or 68 as odd jobs, 46% or 95 as housewives and 60% or 3 were orphan students got C. 7% or 2 of mothers' job in education, 6% or 2 in government agencies, 12% or 18 as odd jobs and 13% or 26 as housewives achieved grade D. About 4% or 1 of mothers' job as odd job, 3% or 4 as odd jobs and 0.5% or 1 as housewife got E. The highest percent for grade A and E was mothers' job in education.

Table 4.36: Mother's job and MT crosstabulation

MOTHER'S JOB * MT Crosstabulation												
MT												
	A	% within MT	B	% within MT	C	% within MT	D	% within MT	E	% within MT	Mother's Job	Total
Education	5	19%	9	33%	10	37%	1	4%	2	7%	27	6%
Government Agencies	1	3%	8	22%	23	64%	4	11%	0	0%	36	9%
Odd Job	0	0%	42	29%	73	50%	28	19%	2	1%	145	35%
Housewife	5	2%	73	35%	94	45%	32	15%	3	1%	207	49%
Orphan	0	0%	1	20%	2	40%	2	40%	0	0%	5	1%
Total	11	3%	133	32%	202	48%	67	16%	7	2%	420	100%

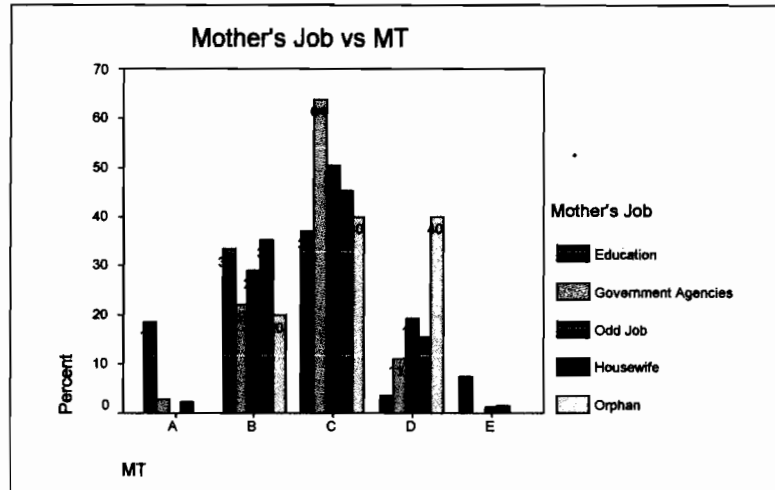


Figure 4.35: Relationship between mother's job and MT

There were 19% or 5 of mothers' job in education, 3% or 1 in government agency and 2% or 5 as housewives got A in Mathematics. Grade B was presented by 33% or 9 of mothers' job in education, 22% or 8 in government agencies, 29% or 42 as odd jobs, 35% or 73 as housewives and 20% or 1 was orphan student. About 37% or 10 of mothers' job in education, 64% or 23 in government agencies, 50% or 73 as odd jobs, 45% or 94 as housewives and 40% or 2 were orphan students got C. 4% or 1 of mothers' job in education, 11% or 4 in government agencies, 19% or 28 as odd jobs, 15% or 32 as housewives and 40% or 2 were orphan students achieved grade D. About 7% or 2 of mothers' job in education, 1% or 2 as odd jobs and 1% or 3 as housewives got E. The highest percent for grade A and E was mothers' job in education.

Table 4.37: Mother's job and SC crosstabulation

MOTHER'S JOB * SC Crosstabulation												
SC												
	A	% within SC	B	% within SC	C	% within SC	D	% within SC	E	% within SC	Mother's Job	Total
Education	4	15%	9	33%	12	44%	2	7%	0	0%	27	6%
Government Agencies	2	6%	14	39%	18	50%	2	6%	0	0%	36	9%
Odd Job	3	2%	63	43%	71	49%	6	4%	2	1%	145	35%
Housewife	9	4%	83	40%	96	46%	18	9%	1	0.5%	207	49%
Orphan	0	0%	1	20%	3	60%	1	20%	0	0%	5	1%
Total	18	4%	170	40%	200	48%	29	7%	3	1%	420	100%

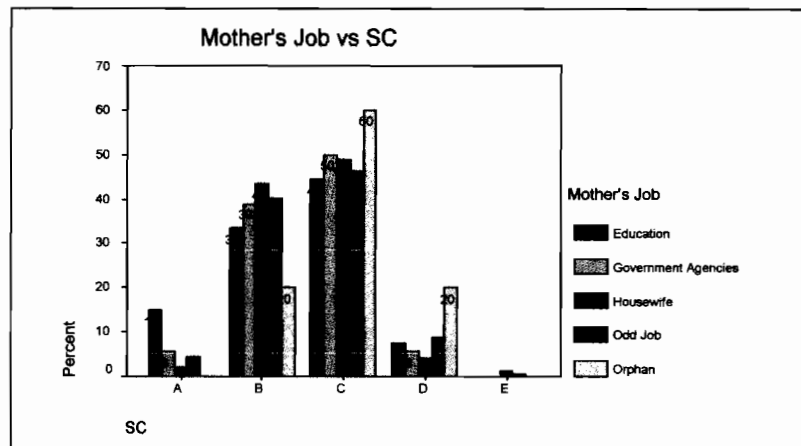


Figure 4.36: Relationship between mother's job and SC

There were 15% or 4 of mothers' job in education, 6% or 2 in government agencies, 2% or 3 as odd jobs and 4% or 9 as housewives got A in Science. Grade B was presented by 33% or 9 of mothers' job in education, 39% or 14 in government agencies, 43% or 63 as odd jobs, 40% or 83 as housewives and 20% or 1 was orphan student. About 44% or 12 of mothers' job in education, 50% or 18 in government agencies, 49% or 71 as odd jobs, 46% or 96 as housewives and 60% or 3 were orphan students got C. 7% or 2 of mothers' job in education, 6% or 2 in government agencies, 4% or 6 as odd jobs, 9% or 18 as housewives and 20% or 1 was orphan student achieved grade D. About 1% or

2 of mothers' job as odd jobs and 0.5% or 1 as housewife got E. The highest percent for grade A was mothers' job in education. The highest percent for grade E was odd jobs.

Family income percapita is one of the factors that contributes to the students' performance. The survey had been carried out to determine the relationship between family income percapita and performance. The analysis of missing value is shown below:

Table 4.38: Family income percapita frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Family Income Percapita * BM1	420	100.0%	0	.0%	420	100.0%
Family Income Percapita * BM2	420	100.0%	0	.0%	420	100.0%
Family Income Percapita * BI	420	100.0%	0	.0%	420	100.0%
Family Income Percapita * MT	420	100.0%	0	.0%	420	100.0%
Family Income Percapita * SC	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The family income percapita allocation is shown below:

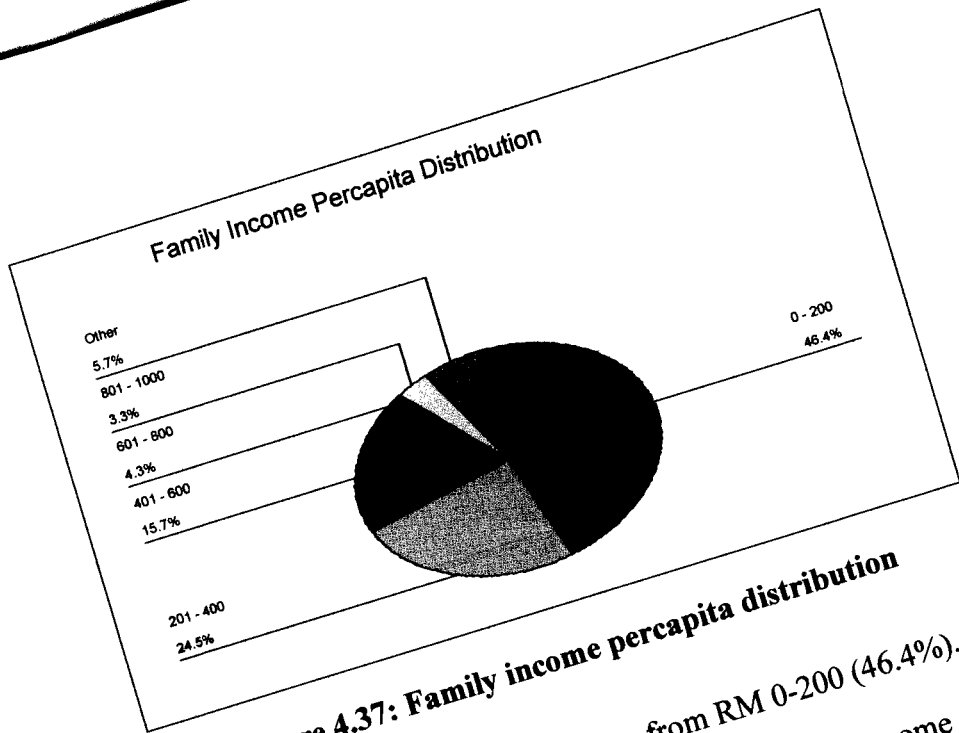


Figure 4.37: Family income percapita distribution

The biggest percent was given by income range from RM 0-200 (46.4%). Income range from RM 201-400 contributed 24.5%. 15.7% was given by income range from RM 401-600. Then, income range from RM 601-800 gave 4.3%. 3.3% was presented by income range from RM 801-1000. More than income range RM 1000 was represented as 5.7% from all the students.

Table 4.39: Family income percapita and BM1 crosstabulation

	FAMILY INCOME PERCAPITA * BM1 Crosstabulation						
	A	% within BM1	B	% within BM1	C	% within BM1	D
0-200	22	11%	69	35%	92	47%	
201-400	9	9%	43	42%	47	46%	
401-600	4	6%	31	47%	28	42%	
601-800	3	17%	8	44%	7		
801-1000	2	14%	8	57%	4		
1001-1200	0	0%	3	60%	2		
1201-1400	0	0%	1	100%			
1401-1600	0	0%	4	100%			
1601-1800	2	40%	0	0%			
1801-2000	1	11%	6	67%			
Total	43	10%	173				

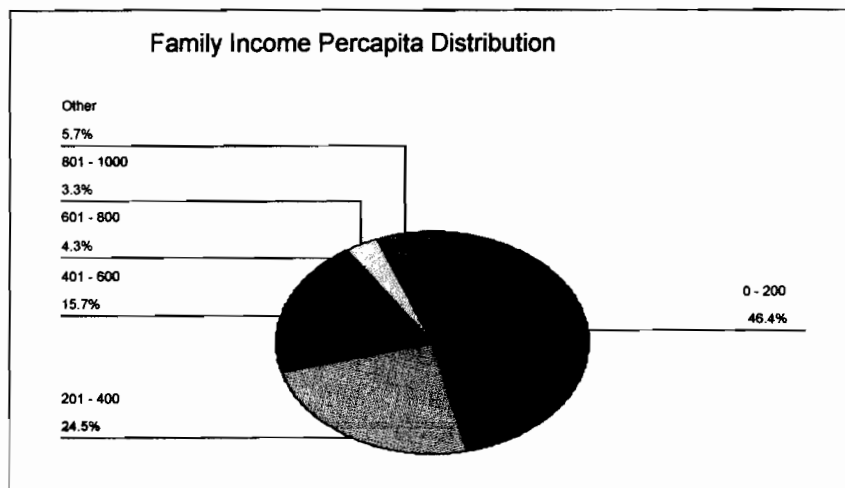


Figure 4.37: Family income percapita distribution

The biggest percent was given by income range from RM 0-200 (46.4%). Income range from RM 201-400 contributed 24.5%. 15.7% was given by income range from RM 401-600. Then, income range from RM 601-800 gave 4.3%. 3.3% was presented by income range from RM 801-1000. More than income range RM 1000 was represented as 5.7% from all the students.

Table 4.39: Family income percapita and BM1 crosstabulation

FAMILY INCOME PERCAPITA * BM1 Crosstabulation										
BM1										
	A	% within BM1	B	% within BM1	C	% within BM1	D	% within BM1	Family Income Percapita	Total
0-200	22	11%	69	35%	92	47%	12	6%	195	46%
201-400	9	9%	43	42%	47	46%	4	4%	103	25%
401-600	4	6%	31	47%	28	42%	3	5%	66	16%
601-800	3	17%	8	44%	7	39%	0	0%	18	4%
801-1000	2	14%	8	57%	4	29%	0	0%	14	3%
1001-1200	0	0%	3	60%	2	40%	0	0%	5	1%
1201-1400	0	0%	1	100%	0	0%	0	0%	1	0%
1401-1600	0	0%	4	100%	0	0%	0	0%	4	1%
1601-1800	2	40%	0	0%	3	60%	0	0%	5	1%
1801-2000	1	11%	6	67%	2	22%	0	0%	9	2%
Total	43	10%	173	41%	185	44%	19	5%	420	100%

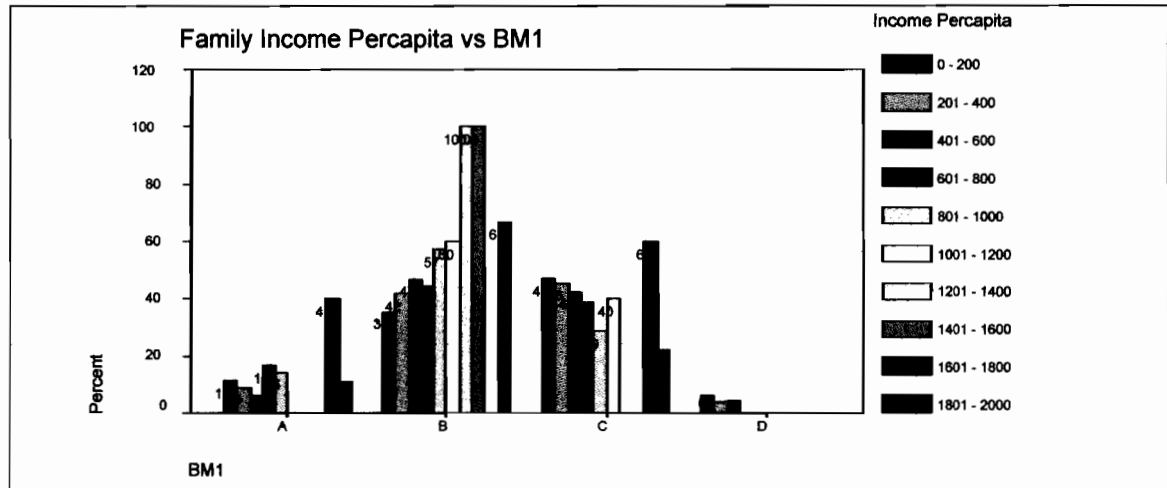


Figure 4.38: Relationship between family income percapita and BM1

There were 11% or 22 of income range RM0-200, 9% or 9 from RM201-400, 6% or 4 from RM401-600, 17% or 3 from RM601-800, 14% or 2 from RM801-1000, 40% or 2 from RM1601-1800 and 11% or 1 from RM1801-2000 got A for Bahasa Melayu 1 (Pemahaman). Grade B was presented by 35% or 69 from income range RM0-200, 42% or 43 from RM201-400, 47% or 31 from RM 401-600, 44% or 8 from RM 601-800, 57% or 8 from RM801-1000, 60% or 3 from RM1001-1200, 100% or 1 from RM1201-1400, 100% or 4 from RM1401-1600 and 67% or 6 from RM1801-2000. About 47% or 92 from income range Rm0-200, 46% or 47 from RM201-400, 42% or 28 from RM 401-600, 39% or 7 from RM601-800, 29% or 4 from RM801-1000, 40% or 2 from RM1001-1200, 60% or 3 from RM1601-1800 and 22% or 2 from RM1801-2000 achieved grade C. 6% or 12 from income range RM0-200, 4% or 4 from RM201-400 and 5% or 3 from RM401-600 got D. The highest percent for grade A was from income range RM1601-1800 while the highest percent for grade D was the students from family income range RM0-200.

Table 4.40: Family income percapita and BM2 crosstabulation

FAMILY INCOME PERCAPITA * BM2 Crosstabulation												
BM2												
	A	% within BM2	B	% within BM2	C	% within BM2	D	% within BM2	E	% within BM2	Family Income Percapita	Total
0-200	6	3%	69	35%	93	48%	26	13%	1	1%	195	46%
201-400	4	4%	36	35%	54	52%	8	8%	1	1%	103	25%
401-600	1	2%	30	45%	28	42%	7	11%	0	0%	66	16%
601-800	1	6%	7	39%	10	56%	0	0%	0	0%	18	4%
801-1000	1	7%	6	43%	7	50%	0	0%	0	0%	14	3%
1001-1200	0	0%	4	80%	1	20%	0	0%	0	0%	5	1%
1201-1400	0	0%	0	0%	1	100%	0	0%	0	0%	1	0%
1401-1600	0	0%	3	75%	1	25%	0	0%	0	0%	4	1%
1601-1800	2	40%	0	0%	2	40%	0	0%	1	20%	5	1%
1801-2000	0	0%	7	78%	2	22%	0	0%	0	0%	9	2%
Total	15	4%	162	39%	199	47%	41	10%	3	1%	420	100%

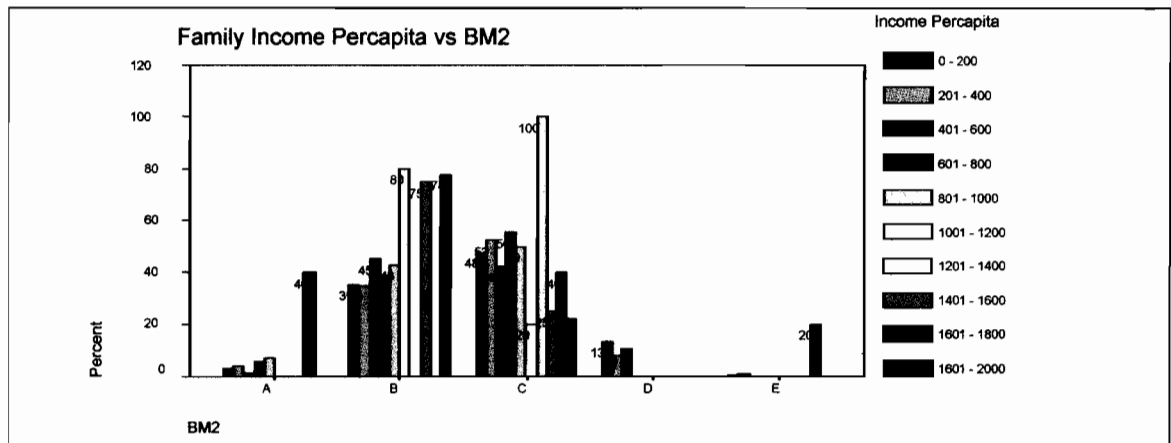


Figure 4.39: Relationship between family income percapita and BM2

There were 3% or 6 of family income range RM0-200, 4% or 4 from RM201-400, 2% or 1 from RM401-600, 6% or 1 from RM601-800, 7% or 1 from RM801-1000 and 40%

or 2 from RM1601-1800 got A for Bahasa Melayu 2 (Penulisan). Grade B was presented by 35% or 69 from income range RM0-200, 35% or 36 from RM201-400, 45% or 30 from RM 401-600, 39% or 7 from RM 601-800, 43% or 6 from RM801-1000, 80% or 4 from RM1001-1200, 75% or 3 from RM1401-1600 and 78% or 7 from RM1801-2000. About 48% or 93 from income range RM0-200, 52% or 54 from RM201-400, 42% or 28 from RM 401-600, 56% or 10 from RM601-800, 50% or 7 from RM801-1000, 20% or 1 from RM1001-1200, 100% or 1 from RM1201-1400, 25% or 1 from RM 1401-1600, 40% from RM1601-1800 and 22% or 2 from RM1801-2000 achieved grade C. 13% or 26 from income range RM0-200, 8% or 8 from RM201-400 and 11% or 7 from RM401-600 got D. About 1% or 1 student from family income range RM0-200, 1% or 1 from RM201-400 and 20% or 1 from RM1601-1800 presented E. The highest percent for grade A and E were the students from family income range RM1601-1800.

Table 4.41: Family income percapita and BI crosstabulation

FAMILY INCOME PERCAPITA * BI Crosstabulation												
BI												
	A	% within BI	B	% within BI	C	% within BI	D	% within BI	E	% within BI	Family Income Percapita	Total
0-200	5	3%	65	33%	92	47%	29	15%	4	2%	195	46%
201-400	6	6%	42	41%	44	43%	10	10%	1	1%	103	25%
401-600	2	3%	20	30%	38	58%	6	9%	0	0%	66	16%
601-800	0	0%	15	83%	2	11%	0	0%	1	6%	18	4%
801-1000	2	14%	8	57%	2	14%	2	14%	0	0%	14	3%
1001-1200	0	0%	1	20%	4	80%	0	0%	0	0%	5	1%
1201-1400	0	0%	0	0%	1	100%	0	0%	0	0%	1	0%
1401-1600	0	0%	3	75%	1	25%	0	0%	0	0%	4	1%
1601-1800	0	0%	1	20%	3	60%	1	20%	0	0%	5	1%
1801-2000	1	11%	6	67%	2	22%	0	0%	0	0%	9	2%
Total	16	4%	161	38%	189	45%	48	11%	6	1%	420	100%

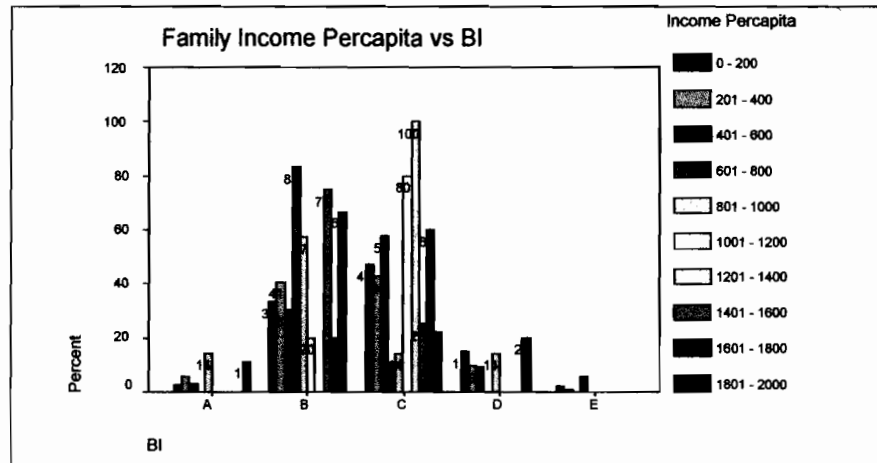


Figure 4.40: Relationship between family income percapita and BI

There were 3% or 5 of family income range RM0-200, 6% or 6 from RM201-400, 3% or 2 from RM401-600, 14% or 2 from RM801-1000 and 11% or 1 from RM1801-2000 got A for English. Grade B was presented by 33% or 65 from income range RM0-200, 41% or 42 from RM201-400, 30% or 20 from RM401-600, 83% or 15 from RM601-800, 57% or 8 from RM801-1000, 20% or 1 from RM1001-1200, 75% or 3 from RM1401-1600, 20% or 1 from RM1601-1800 and 67% or 6 from RM1801-2000. About 47% or 92 from income range RM0-200, 43% or 44 from RM201-400, 58% or 38 from RM 401-600, 11% or 2 from RM601-800, 14% or 2 from RM801-1000, 80% or 4 from RM1001-1200, 100% or 1 from RM1201-1400, 25% or 1 from RM1401-1600, 60% or 3 from RM1601-1800 and 22% or 2 from RM1801-2000 achieved grade C. 15% or 29 from income range RM0-200, 10% or 10 from RM201-400, 9% or 6 from RM401-600, 14% or 2 from RM801-1000 and 20% or 1 from RM1601-1800 got D. About 2% or 4 students from family income range RM0-200, 1% or 1 from RM201-400 and 6% or 1 from RM601-800 presented E. The highest percent for grade A was the students from

family income range RM801-1000 while for grade E was from income range RM601-800.

Table 4.42: Family income percapita and MT crosstabulation

FAMILY INCOME PERCAPITA * MT Crosstabulation												
MT												
	A	% within MT	B	% within MT	C	% within MT	D	% within MT	E	% within MT	Family Income Percapita	Total
0-200	1	1%	60	31%	90	46%	41	21%	3	2%	195	46%
201-400	5	5%	31	30%	50	49%	16	16%	1	1%	103	25%
401-600	2	3%	27	41%	32	48%	4	6%	1	2%	66	16%
601-800	0	0%	7	39%	10	56%	0	0%	1	6%	18	4%
801-1000	3	21%	3	21%	7	50%	0	0%	1	7%	14	3%
1001-1200	0	0%	0	0%	4	80%	1	20%	0	0%	5	1%
1201-1400	0	0%	0	0%	1	100%	0	0%	0	0%	1	0%
1401-1600	0	0%	0	0%	2	50%	2	50%	0	0%	4	1%
1601-1800	0	0%	0	0%	3	60%	2	40%	0	0%	5	1%
1801-2000	0	0%	5	56%	3	33%	1	11%	0	0%	9	2%
Total	11	3%	133	32%	202	48%	67	16%	7	2%	420	100%

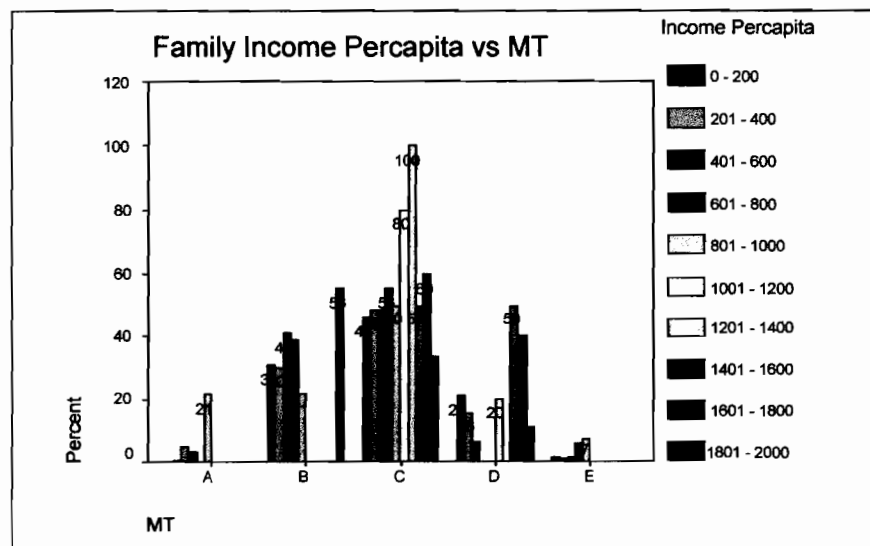


Figure 4.41: Relationship between family income percapita and MT

There were 1% or 1 of family income range RM0-200, 5% or 5 from RM201-400, 3% or 2 from RM401-600 and 21% or 3 from RM801-1000 got A for Mathematics. Grade B was presented by 31% or 60 from income range RM0-200, 30% or 31 from RM201-400, 41% or 27 from RM 401-600, 39% or 7 from RM 601-800, 21% or 3 from RM801-1000 and 56% or 5 from RM1801-2000. About 46% or 90 from income range RM0-200, 49% or 50 from RM201-400, 48% or 32 from RM 401-600, 56% or 10 from RM601-800, 50% or 7 from RM801-1000, 80% or 4 from RM1001-1200, 100% or 1 from RM1201-1400, 50% or 2 from RM 1401-1600, 60% or 3 from RM1601-1800 and 33% or 3 from RM1801-2000 achieved grade C. 21% or 41 from income range RM0-200, 16% or 16 from RM201-400, 6% or 4 from RM401-600, 20% or 1 from RM1001-1200, 50% or 2 from RM1401-1600, 40% or 2 from RM1601-1800 and 11% or 1 from RM1801-2000 got D. About 2% or 3 students from family income range RM0-200, 1% or 1 from RM201-400, 2% or 1 from RM401-600, 6% or 1 from RM601-800 and 7% or 1 from RM801-1000 presented E. The highest percent for grade A and E were the students from family income range RM801-1000.

Table 4.43: Family income percapita and SC crosstabulation

FAMILY INCOME PERCAPITA * SC Crosstabulation												
SC												
	A	% within SC	B	% within SC	C	% within SC	D	% within SC	E	% within SC	Family Income Percapita	Total
0-200	5	3%	83	43%	92	47%	13	7%	2	1%	195	46%
201-400	6	6%	43	42%	51	50%	3	3%	0	0%	103	25%
401-600	2	3%	31	47%	27	41%	6	9%	0	0%	66	16%
601-800	1	6%	6	33%	9	50%	2	11%	0	0%	18	4%
801-1000	3	21%	2	14%	9	64%	0	0%	0	0%	14	3%
1001-1200	0	0%	2	40%	1	20%	2	40%	0	0%	5	1%
1201-1400	0	0%	0	0%	1	100%	0	0%	0	0%	1	0%
1401-1600	0	0%	0	0%	3	75%	1	25%	0	0%	4	1%
1601-1800	0	0%	0	0%	3	60%	2	40%	0	0%	5	1%
1801-2000	1	11%	3	33%	4	44%	0	0%	1	11%	9	2%
Total	18	4%	170	40%	200	48%	29	7%	3	1%	420	100%

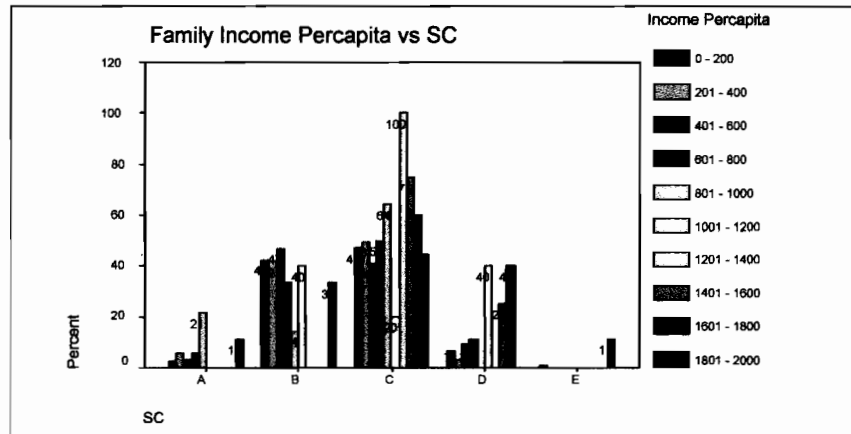


Figure 4.42: Relationship between family income percapita and SC

There were 3% or 5 of family income range RM0-200, 6% or 6 from RM201-400, 3% or 2 from RM401-600, 6% or 1 from RM601-800, 21% or 3 from RM801-1000 and 11% or 1 from RM1801-2000 got A for Science. Grade B was presented by 43% or 83

from income range RM0-200, 42% or 43 from RM201-400, 47% or 31 from RM 401-600, 33% or 6 from RM 601-800, 14% or 2 from RM801-1000, 40% or 2 from RM1001-1200 and 33% or 3 from RM1801-2000. About 47% or 92 from income range RM0-200, 50% or 51 from RM201-400, 41% or 27 from RM 401-600, 50% or 9 from RM601-800, 64% or 9 from RM801-1000, 80% or 4 from RM1001-1200, 100% or 1 from RM1201-1400, 25% or 1 from RM 1401-1600, 20% or 1 from RM1001-1200, 100% or 1 from RM1201-1400, 75% or 3 from Rm1401-1600, 60% or 3 from RM1601-1800 and 44% or 4 from RM1801-2000 achieved grade C. 7% or 13 from income range RM0-200, 3% or 3 from RM201-400, 9% or 6 from RM401-600, 11% or 2 from RM601-800 and 40% or 2 from RM1601-1800 got D. About 1% or 2 students from family income range RM0-200 and 11% or 1 from RM1801-2000 presented E. The highest percent for grade A was the students from family income range RM801-1000 while for grade E was from income range RM1801-2000.

Attendance and gender play the main roles to affect students' performance. The survey had been carried out to determine the relationship between attendance and gender. The analysis of missing value is shown below:

Table 4.44: Gender and attendance frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * Attendance	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The attendance versus gender allocation is shown below:

Table 4.45: Gender and attendance crosstabulation

GENDER vs ATTENDANCE Crosstabulation						
	Attendance				Gender	Total
	Good	% within Gender	Bad	% within Gender		
Male	164	80%	40	20%	204	49%
Female	166	77%	50	23%	216	51%
Total	330	79%	90	21%	420	100%

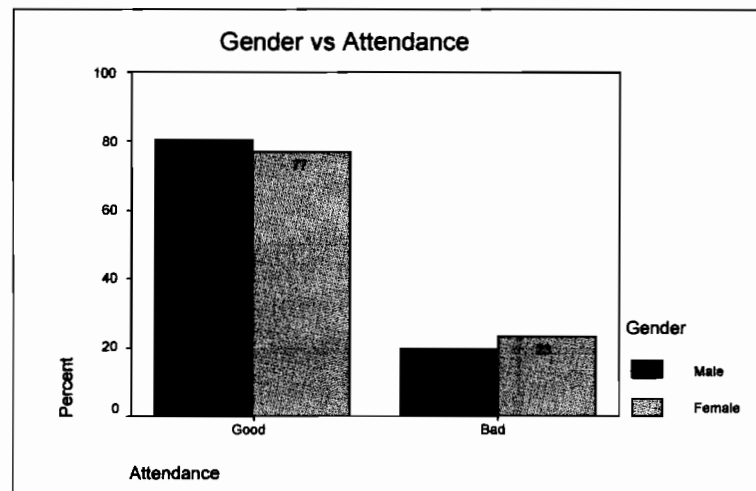


Figure 4.43: Relationship between gender and attendance

The good attendance given by male students was 80% or 164 while the bad attendance was represented as 20% or 40. There were 77% or 166 of female students with good attendance while 23% or 50 of bad attendance students. The good attendance was given higher by male students while the bad attendance was given higher by female students.

Participation in co-curricular activities and gender also affect students' performance. The survey had been carried out to determine the relationship between participation in co-curricular activities and gender. The analysis of missing value is shown below:

Table 4.46: Gender and co-curricular activities frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * Co-curricular	<i>420</i>	<i>100.0%</i>	<i>0</i>	<i>.0%</i>	<i>420</i>	<i>100.0%</i>

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The co-curricular activities and gender allocation is shown below:

Table 4.47: Gender and co-curricular crossstabulation

GENDER vs CO-CURRICULAR ACTIVITIES Crosstabulation						
	Co-curricular				Gender	Total
	Active	% within Gender	Passive	% within Gender		
Male	150	74%	54	26%	204	49%
Female	149	69%	67	31%	216	51%
Total	299	71%	121	29%	420	100%

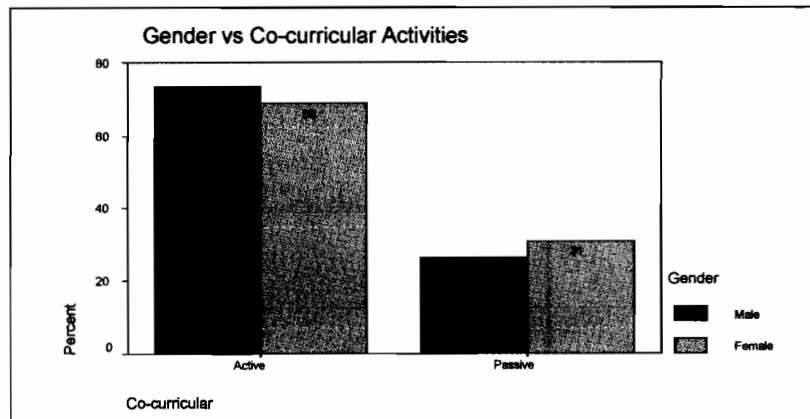


Figure 4.44: Relationship between gender and co-curricular

The male students those active in co-curricular activities gave 74% or 150. There were about 26% or 54 of male students which passive in co-curricular activities. 69% or 149 was represented as female students those active in co-curricular activities while 31% or 67 of passively involved in co-curricular activities. The highest percent was given by male students for actively involved in co-curricular activities and the highest percent for passively involved was given by female students.

4.2 2 Dimension Analysis

Attendance, father's job and mother's job showed the effects towards the students' performance. The survey had been carried out to determine the relationship between attendance, father's job, mother's job and performance. The analysis of missing value is shown below:

Table 4.48: Father's job, mother's job and attendance frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Father's Job *						
Mother's Job *						
Attendance	420	100.0%	0	.0%	420	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The attendance, father's job, mother's job allocation is shown below:

Table 4.49: Father's job, mother's job and attendance crosstabulation

		FATHER'S JOB * MOTHER'S JOB * ATTENDANCE Crosstabulation					
		Attendance			Attendance		
Father's Job	Mother's Job	Good	% within Attendance	Bad	% within Attendance	Parents' Job	Total
Education	Education	15	68%	1	33%	16	64%
	Government Agencies	1	5%	0	0%	1	4%
	Odd Job	0	0%	0	0%	0	0%
	Housewife	6	27%	2	67%	8	32%
	Orphan	0	0%	0	0%	0	0%
	Total	22	100%	3	100%	25	100%
Government Agencies	Education	7	8%	0	2100%	7	6%
	Government Agencies	19	22%	4	19%	23	21%
	Odd Job	8	9%	1	5%	9	8%
	Housewife	54	61%	16	76%	70	64%
	Orphan	0	0%	0	0%	0	0%
	Total	88	100%	21	100%	109	100%
Private Sectors	Education	1	2%	0	0%	1	2%
	Government Agencies	3	6%	0	0%	3	5%
	Odd Job	4	8%	0	0%	4	7%
	Housewife	40	82%	12	100%	52	85%
	Orphan	1	2%	0	0%	1	2%
	Total	49	100%	12	100%	61	100%
Odd Job	Education	1	1%	1	2%	2	1%
	Government Agencies	7	4%	1	2%	8	4%
	Odd Job	98	59%	30	56%	128	58%
	Housewife	57	35%	20	37%	77	35%
	Orphan	2	1%	2	4%	4	2%
	Total	165	100%	54	100%	219	100%
Orphan	Education	1	17%	0	0%	1	17%
	Government Agencies	1	17%	0	0%	1	17%
	Odd Job	4	67%	0	0%	4	67%
	Housewife	0	0%	0	0%	0	0%
	Orphan	0	0%	0	0%	0	0%
	Total	6	100%	0	0%	6	100%

For good attendance students with their mothers in education field, there were 94% or 15 students from fathers' job in education field, 100% or 7 from government agencies, 100% or 1 from private sector, 50% or 1 from odd job and 100% or 1 is orphan. 6% or 1 student from fathers' job in education field and 50% or 1 from odd job father was the bad attendance students. As their mothers in government agencies, for good attendance students were presented by 100% or of father's job in education, 83% or 19 in government agencies, 100% or 3 in private sectors, 88% or 7 as odd jobs and 100% or 1 was orphan. For bad attendance, there were 17% or 4 from government agencies and 13% or 1 as odd job. As odd jobs mothers, the good attendance students with fathers in government agencies was 89% or 8, 100% or 4 in private sectors, 77% or 98 as odd jobs and 100% or 4 were orphans. The bad attendance students were represented by 11% or 1 from father's job in government agency, 23% or 30 as odd jobs and 100% or 4 were orphans. As housewives mothers, the good attendance students were given by 75% or 6 from fathers' job in education, 77% or 54 from government agencies, 75% or 40 from private sectors and 74% or 57 as odd jobs. The bad attendance students were 25% or 2 from fathers' job in education, 23% or 16 in government agencies, 23% or 12 in private sectors and 26% or 20 as odd jobs. For orphan students with their fathers were in private sector was given as 100% or 1 and the odd jobs were represented as 50% or 2 (good attendance). 50% or 2 of bad attendance students were orphans as their fathers did odd jobs. Good attendance students gave the higher percent compared to bad attendance especially for the students with their mothers in education field and government agencies.

Participation in co-curricular activities, father's job and mother's job gave the big impact towards the students' performance. The survey had been carried out to determine the relationship between co-curricular activities, father's job, mother's job and performance. The analysis of missing value is shown below:

Table 4.50: Father's job, mother's job and co-curricular activities frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Father's Job *						
Mother's Job *						
Co-curricular	<i>420</i>	<i>100.0%</i>	<i>0</i>	<i>.0%</i>	<i>420</i>	<i>100.0%</i>

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The co-curricular activities, father's job and mother's job allocation is shown below:

Table 4.51: Father's job, mother's job and co-curricular crosstabulation

		FATHER'S JOB * MOTHER'S JOB * CO-CURRICULAR ACTIVITIES Crosstabulation						
		Co-curricular						
Father's Job	Mother's Job	Active	% within Co-curricular	Passive	% within Co-curricular	Parents' Job	Total	
Education	Education	12	67%	4	57%	16	64%	
	Government Agencies	1	6%	0	0%	1	4%	
	Odd Job	0	0%	0	0%	0	0%	
	Housewife	5	28%	3	43%	8	32%	
	Orphan	0	0%	0	0%	0	0%	
	Total	18	100%	7	100%	25	100%	
Government Agencies	Education	7	8%	0	0%	7	6%	
	Government Agencies	18	22%	5	19%	23	21%	
	Odd Job	8	10%	1	4%	9	8%	
	Housewife	50	60%	20	77%	70	64%	
	Orphan	0	0%	0	0%	0	0%	
	Total	83	100%	26	100%	109	100%	
Private Sectors	Education	1	2%	0	0%	1	2%	
	Government Agencies	2	4%	1	7%	3	5%	
	Odd Job	4	9%	0	0%	4	7%	
	Housewife	39	85%	13	87%	52	85%	
	Orphan	0	0%	1	7%	1	2%	
	Total	46	100%	15	100%	61	100%	
Odd Job	Education	1	1%	1	1%	2	1%	
	Government Agencies	7	5%	1	1%	8	4%	
	Odd Job	87	59%	41	58%	128	58%	
	Housewife	51	34%	26	37%	77	35%	
	Orphan	2	1%	2	3%	4	2%	
	Total	148	100%	71	100%	219	100%	
Orphan	Education	1	25%	0	0%	1	17%	
	Government Agencies	1	25%	0	0%	1	17%	
	Odd Job	2	50%	2	100%	4	67%	
	Housewife	0	0%	0	0%	0	0%	
	Orphan	0	0%	0	0%	0	0%	
	Total	4	100%	2	100%	6	100%	

For actively involved students in co-curricular activities with their mothers in education field, there were 75% or 12 students from fathers' job in education field, 100% or 7 from government agencies, 100% or 1 from private sector, 50% or 1 from odd job and 100% or 1 is orphan. For passively involved students, 25% or 4 students from fathers' job in education field and 50% or 1 from odd job fathers. As their mothers in government agencies, for actively involved students were presented by 100% or of fathers' job in education, 78% or 18 in government agencies, 67% or 2 in private sectors, 88% or 7 as odd jobs and 100% or 1 was orphan. For passively involved students, there were 22% or 5 from government agencies, 33% or 1 from private sector and 13% or 1 as odd job. As odd jobs mothers, the actively involved students with fathers in government agencies was 89% or 8, 100% or 4 in private sectors, 68% or 87 as odd jobs and 50% or 2 were orphans. The passively involved students were represented by 11% or 1 from father's job in government agency, 32% or 41 as odd jobs and 50% or 2 were orphans. As housewives mothers, the actively involved students were given by 63% or 5 from fathers' job in education, 71% or 50 from government agencies, 75% or 39 from private sectors and 66% or 51 as odd jobs. The passively involved students were 38% or 3 from fathers' job in education, 29% or 20 in government agencies, 25% or 13 in private sectors and 34% or 26 as odd jobs.

For orphan students with their fathers did the odd jobs were represented as 50% or 2 (actively involved). For the orphan students with their fathers in private sectors and odd jobs were 100% or 1 and 50% or 2 (passively involved). Actively involved students gave the higher percent compared to passively involved especially for the students with their mothers in education field and government agencies. Passively involved students gave the higher percent for orphan

students. School attendance, father's job and mother's job gave the big impact towards the students' performance. The survey had been carried out to determine the relationship between school attendance and parents' job towards the students' performance. The analysis of missing value is shown below:

Table 4.52: Same parents' job and attendance frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Same Parents' Job * Attendance	164	100.0%	0	.0%	164	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The school attendance and parents' job allocation is shown below:

Table 4.53: Same parents' job and attendance crosstabulation

SAME PARENTS' JOB vs ATTENDANCE Crosstabulation							
		Attendance				Same Parents' Job	Total
		Good	% within Same Parents' Job	Bad	% within Same Parents' Job		
Same Parents' Job	Education	15	12%	1	3%	16	10%
	Government Agencies	17	13%	3	9%	16	10%
	Odd Job	98	75%	30	88%	128	80%
	Total	126	100%	34	100%	160	100%

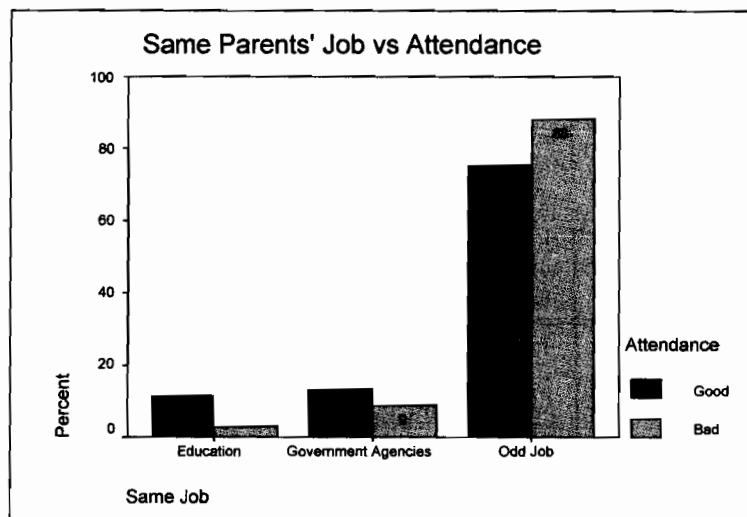


Figure 4.45: Relationship between same parents' job and attendance

There were 12% or 15 good attendance students from parents work in education field while 3% or 1 bad attendance students. 13% or 17 of good attendance students and 9% or 3 of bad attendance students came from parents in government agencies. The odd job gave 75% or 98 of good attendance and 88% or 30 of bad attendance students. Both parents form odd jobs gave the highest percent for good and bad attendance students.

School attendance, father's job and mother's job were among the main factors influenced the students' performance. The survey had been carried out to determine the relationship between school attendance, father's job and mother's job as housewife towards the performance. The analysis of missing value is shown below:

Table 4.54: Father's job with mother's job as housewife and attendance frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Father's Job with Mother's Job as Housewife * Attendance	208	100.0%	0	.0%	208	100.0%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The school attendance, father's job and mother's job as housewife allocation is shown below:

Table 4.55: Father's job with mother's job as housewife and attendance crosstabulation

FATHER'S JOB WITH MOTHER'S JOB AS HOUSEWIFE vs ATTENDANCE Crosstabulation							
		Attendance					
		Good	% within Father's Job	Bad	% within Father's Job	Father's Job	Total
Father's Job	Education	6	6%	2	2%	8	4%
	Government Agencies	40	39%	30	29%	70	34%
	Private Sector	21	20%	31	30%	52	25%
	Odd Job	36	35%	42	40%	78	38%
	Total	103	100%	105	100%	208	100%

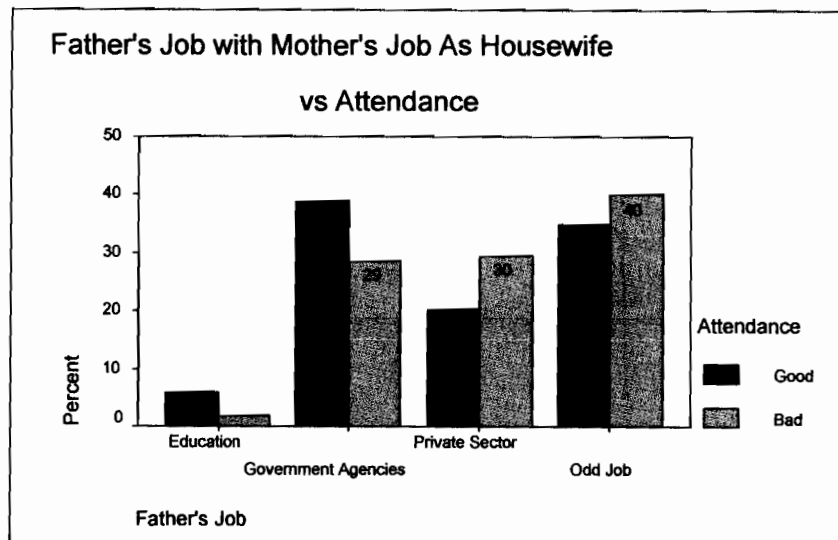


Figure 4.46: Relationship between father's job with mother's job as housewife and attendance

As housewives mothers and fathers in education field, the good attendance students were 6% or 6 while 2% or 2 of bad attendance students. 39% or 40 for good attendance and 29% or 30 for bad attendance students with their fathers work in government agencies. About 20% or 21 for good attendance and 30% or 31 for bad attendance students with fathers work in private sectors. Odd jobs fathers contributed 35% or 36 for good attendance and 40% or 42 for bad attendance students. The highest percent for good attendance was given by fathers in government agencies while for bad attendance students were odd jobs fathers.

School attendance and parents' status gave the big impact towards the students' performance. The survey had been carried out to determine the relationship between school attendance and parents' status either single mother or single father towards the students' performance. The analysis of missing value is shown below:

Table 4.56: Single parent frequency table

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Single Parent * Attendance	<i>10</i>	<i>100.0%</i>	<i>0</i>	<i>.0%</i>	<i>10</i>	<i>100.0%</i>

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The school attendance and parents' status allocation is shown below:

Table 4.57: Single parent and attendance crosstabulation

SINGLE PARENT vs ATTENDANCE Crosstabulation							
		Attendance					
		Good	% within Single Parent	Bad	% within Single Parent	Single Parent	Total
Single Parent	Single Mother	5	63%	0	0%	5	50%
	Single Father	3	38%	2	100%	5	50%
	Total	8	100%	2	100%	10	100%

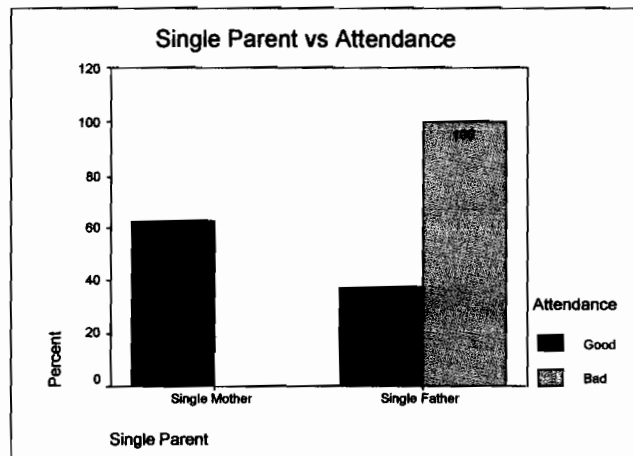


Figure 4.47: Single parent and attendance

There were 63% or 5 of good attendance students live with their only mother. For the children brought up by their only father, there were about 38% or 3 of good attendance students and 100% Or 2 of bad attendance students. The highest percent for good attendance students was given by single mother family while single father family for bad attendance students.

In order to develop a suitable model from the examination data, this chapter describes the results of the training and testing of NN using NC. Some experiments were carried out to determine the suitable parameters in order to provide the best result.

4.3 Initial Study

The analysis is done to make sure the data does not have missing value. It is shown below:

Table 4.58: Academic factors frequency table

Case Processing Summary						
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
BM1 * Actual BM1	420	100.00%	0	0.00%	420	100.00%
BM2 * Actual BM2	420	100.00%	0	0.00%	420	100.00%
BI * Actual BI	420	100.00%	0	0.00%	420	100.00%
MT * Actual MT	420	100.00%	0	0.00%	420	100.00%
SC * Actual SC	420	100.00%	0	0.00%	420	100.00%

Based on the above frequency table, clearly there is no missing value for all attributes in the study. The experiment can be easily carried out because the data is clean. Initial study was carried out to determine the number of hidden unit suitable for this study. The important parameter to control the training process is the number of epoch. If the number of epoch is too high, it will make the network memorizes the datasets and cause the over-fitting. As a result, a

good NN performance is not guaranteed. However, if the number of epoch is too small, the network will not be sufficiently trained. Therefore, several experiments need to be carried out to determine the suitable number of epoch to avoid the above situations. The results of training and test phase are shown below:

a) Bahasa Melayu 1

The final model for Bahasa Melayu 1 is:

Table 4.59: BM1 model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	3
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.2
Activation Function	Sigmoid
Stopping Criteria	100

The result using the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1      0    0    0    0
! 1.8+ 0      4    0    0    0
! 2.6+ 0      0    2    0    0
! 3.4+ 0      0    0   13   7
! 4.2+ 0      0    0    0   15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True           Predicted
```

```

! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 6    0    0    0    0
! 1.8+ 0    40   0    0    0
! 2.6+ 0    1    9    0    0
! 3.4+ 0    0    0   145  63
! 4.2+ 0    0    0    0   72
!
! Total number of targets : 336
! Total correct : 272
! Percentage correct : 80.95%

```

Bahasa Melayu 1 with Gender

The final model for Bahasa Melayu 1 and gender is:

Table 4.60: BM1 with gender model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.3
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	300

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    0    0    0
! 1.8+ 0    7    0    0    0
! 2.6+ 0    0    0    0    0
! 3.4+ 0    0    0   18   6
! 4.2+ 0    0    0    0   11
! Total number of targets : 42
! Total correct : 36
! Percentage correct : 85.71%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted

```

```

! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 8    0    0    0    0
! 1.8+ 0    39   0    0    0
! 2.6+ 0    0    11   0    0
! 3.4+ 0    0    1    140  62
! 4.2+ 0    0    0    0    75
!
! Total number of targets : 336
! Total correct : 273
! Percentage correct : 81.25%

```

Bahasa Melayu 1 with Attendance

The final model for Bahasa Melayu 1 and attendance is:

Table 4.62: BM1 with attendance model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    0    0    0    0
! 1.8+ 0    4    0    0    0
! 2.6+ 0    0    2    0    0
! 3.4+ 0    0    0    13   7
! 4.2+ 0    0    0    0    15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted

```

```

! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 6    0    0    0    0
! 1.8+ 0    40   0    0    0
! 2.6+ 0    1    9    0    0
! 3.4+ 0    0    0   145  63
! 4.2+ 0    0    0    0   72
! Total number of targets : 336
! Total correct : 272
! Percentage correct : 80.95%

```

Bahasa Melayu 1 with Co-curricular Activities

The final model for Bahasa Melayu 1 and co-curricular is:

Table 4.62: BM1 with participation in co-curricular activities model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	3
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    0    0    0    0
! 1.8+ 0    4    0    0    0
! 2.6+ 0    0    2    0    0
! 3.4+ 0    0    0   13   7
! 4.2+ 0    0    0    0   15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted

```

```

! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 6    0    0    0    0
! 1.8+ 0    40   0    0    0
! 2.6+ 0    1    9    0    0
! 3.4+ 0    0    0   145  63
! 4.2+ 0    0    0    0   72
! Total number of targets : 336
! Total correct : 272
! Percentage correct : 80.95%

```

Bahasa Melayu 1 with Position in Family

The final model for Bahasa Melayu 1 and position in family is:

Table 4.63: BM1 with position in family model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    0    0    0    0
! 1.8+ 0    4    0    0    0
! 2.6+ 0    0    2    0    0
! 3.4+ 0    0    0   13   7
! 4.2+ 0    0    0    0   15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted

```

```

! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 6    0    0    0    0
! 1.8+ 0    40   0    0    0
! 2.6+ 0    1    9    0    0
! 3.4+ 0    0    0   145  63
! 4.2+ 0    0    0    0   72
! Total number of targets : 336
! Total correct : 272
! Percentage correct : 80.95%

```

Bahasa Melayu 1 with Father's Job

The final model for Bahasa Melayu 1 and father's job is:

Table 4.64: BM1 with father's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    0    0    0    0
! 1.8+ 0    4    0    0    0
! 2.6+ 0    0    2    0    0
! 3.4+ 0    0    0   13   7
! 4.2+ 0    0    0    0   15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted

```

```

! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 6    0    0    0    0
! 1.8+ 0    40   0    0    0
! 2.6+ 0    1    9    0    0
! 3.4+ 0    0    0   145  63
! 4.2+ 0    0    0    0   72
! Total number of targets : 336
! Total correct : 272
! Percentage correct : 80.95%

```

Bahasa Melayu 1 with Mother's Job

The final model for Bahasa Melayu 1 and mother's job is:

Table 4.65: BM1 with mother's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    0    0    0    0
! 1.8+ 0    4    0    0    0
! 2.6+ 0    0    2    0    0
! 3.4+ 0    0    0   13   7
! 4.2+ 0    0    0    0   15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted

```

```

! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 6    0    0    0    0
! 1.8+ 0    40   0    0    0
! 2.6+ 0    1    9    0    0
! 3.4+ 0    0    0   145  63
! 4.2+ 0    0    0    0   72
! Total number of targets : 336
! Total correct : 272
! Percentage correct : 80.95%

```

Bahasa Melayu 1 with Family Income

The final model for Bahasa Melayu 1 and family income is:

Table 4.66: BM1 with family income model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	3
Output Unit	5
Weight Seed	2
Learning Rate	0.2
Momentum Rate	0.2
Activation Function	Sigmoid
Stopping Criteria	200

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! -----
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    0    0    0    0
! 1.8+ 0    4    0    0    0
! 2.6+ 0    0    2    0    0
! 3.4+ 0    0    0   13   7
! 4.2+ 0    0    0    0   15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted

```

```

! -----
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 6   0   0   0   0
! 1.8+ 0   40  0   0   0
! 2.6+ 0   0   10  0   0
! 3.4+ 0   0   0   145 63
! 4.2+ 0   0   0   0   72
! Total number of targets : 336
! Total correct : 273
! Percentage correct : 81.25%

```

Bahasa Melayu 1 with All Attributes

The final model for Bahasa Melayu 1 and all attributes is:

Table 4.67: BM1 with all attributes model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	12
Hidden Unit	1
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.2
Activation Function	Sigmoid
Stopping Criteria	300

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! -----
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1   0   0   0   0
! 1.8+ 0   4   0   0   0
! 2.6+ 0   0   2   0   0
! 3.4+ 0   0   0   13  7
! 4.2+ 0   0   0   0   15
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! -----
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+

```

```

! 1.0+ 6    0    0    0    0
! 1.8+ 0   40    0    0    0
! 2.6+ 0    1    9    0    0
! 3.4+ 0    0    0   145  63
! 4.2+ 0    0    0    0   72
! Total number of targets : 336
! Total correct : 272
! Percentage correct : 80.95%

```

b) Bahasa Melayu 2

The final model for Bahasa Melayu 2 is:

Table 4.68: BM2 model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	3
Output Unit	5
Weight Seed	2
Learning Rate	0.1
Momentum Rate	0.2
Activation Function	Sigmoid
Stopping Criteria	200

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      3.0+ 3.4+ 3.8+ 4.2+ 4.6+
! 3.0+ 10    0    0    0    0
! 3.4+ 0     0    0    0    0
! 3.8+ 0     2   12    2    0
! 4.2+ 0     0    0    0    0
! 4.6+ 0     0    0    0   16
! Total number of targets : 42
! Total correct : 38
! Percentage correct : 90.48%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      3.0+ 3.4+ 3.8+ 4.2+ 4.6+
! 3.0+ 83    1    0    0    0

```

```

! 3.4+ 0 0 0 0
! 3.8+ 0 15 112 14 11
! 4.2+ 0 0 0 0 0
! 4.6+ 0 0 3 0 97
! Total number of targets : 336
! Total correct : 292
! Percentage correct : 86.90%

```

Bahasa Melayu 2 with Gender

The final model for Bahasa Melayu 2 and gender is:

Table 4.69: BM2 with gender model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0 0 0 0 0
! 1.8+ 0 1 0 0 0
! 2.6+ 0 0 9 0 0
! 3.4+ 0 0 0 14 2
! 4.2+ 0 0 0 0 16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+

```

```

! 1.0+ 1 1 0 0 0
! 1.8+ 0 5 0 0 0
! 2.6+ 0 3 73 1 0
! 3.4+ 0 0 0 125 27
! 4.2+ 0 0 0 3 97
! Total number of targets : 336
! Total correct : 301
! Percentage correct : 89.58%

```

Bahasa Melayu 2 with Attendance

The final model for Bahasa Melayu 2 and attendance is:

Table 4.70: BM2 with attendance model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0 0 0 0 0
! 1.8+ 0 1 0 0 0
! 2.6+ 0 0 9 0 0
! 3.4+ 0 0 0 14 2
! 4.2+ 0 0 0 0 16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1 1 0 0 0

```

```

! 1.8+ 0    5    0    0    0
! 2.6+ 0    3   73    1    0
! 3.4+ 0    0    0   125  27
! 4.2+ 0    0    0    3   97
! Total number of targets : 336
! Total correct : 301
! Percentage correct : 89.58%

```

Bahasa Melayu 2 with Co-curricular Activities

The final model for Bahasa Melayu 2 and co-curricular activities is:

Table 4.71: BM2 with co-curricular model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    0    0    0
! 1.8+ 0    1    0    0    0
! 2.6+ 0    0    9    0    0
! 3.4+ 0    0    0   14   2
! 4.2+ 0    0    0    0   16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+

```

```

! 1.0+ 1    1    0    0    0
! 1.8+ 0    5    0    0    0
! 2.6+ 0    3    73   1    0
! 3.4+ 0    0    0   125  27
! 4.2+ 0    0    0    3    97
! Total number of targets : 336
! Total correct : 301
! Percentage correct : 89.58%

```

Bahasa Melayu 2 with Position in Family

The final model for Bahasa Melayu 2 and position in family is:

Table 4.72: BM2 with position in family model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    0    0    0
! 1.8+ 0    1    0    0    0
! 2.6+ 0    0    9    0    0
! 3.4+ 0    0    0   14   2
! 4.2+ 0    0    0    0   16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    1    0    0    0
! 1.8+ 0    5    0    0    0

```

```

! 2.6+ 0    3    73    1    0
! 3.4+ 0    0    0    125  27
! 4.2+ 0    0    0    3    97
! Total number of targets : 336
! Total correct : 301
! Percentage correct : 89.58%

```

Bahasa Melayu 2 with Father's Job

The final model for Bahasa Melayu 2 and father's job is:

Table 4.73: BM2 with father's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	3
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    0    0    0
! 1.8+ 0    1    0    0    0
! 2.6+ 0    0    9    0    0
! 3.4+ 0    0    0    14   2
! 4.2+ 0    0    0    0    16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    2    0    0    0
! 1.8+ 0    5    0    0    0

```

```

! 2.6+ 0    3    73    1    0
! 3.4+ 0    0    0    125  27
! 4.2+ 0    0    0    3    97
! Total number of targets : 336
! Total correct : 300
! Percentage correct : 89.29%

```

Bahasa Melayu 2 with Mother's Job

The final model for Bahasa Melayu 2 and mother's job is:

Table 4.74: BM2 with mother's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    0    0    0
! 1.8+ 0    1    0    0    0
! 2.6+ 0    0    9    0    0
! 3.4+ 0    0    0    14   2
! 4.2+ 0    0    0    0    16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    1    0    0    0
! 1.8+ 0    5    0    0    0
! 2.6+ 0    3    73   1    0
! 3.4+ 0    0    0    125  27

```

```

! 4.2+ 0    0    0    3    97
! Total number of targets : 336
! Total correct : 301
! Percentage correct : 89.58%

```

Bahasa Melayu 2 with Family Income

The final model for Bahasa Melayu 2 and family income is:

Table 4.75: BM2 with family income model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    0    0    0
! 1.8+ 0    1    0    0    0
! 2.6+ 0    0    9    0    0
! 3.4+ 0    0    0   14   2
! 4.2+ 0    0    0    0   16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%

```

Train

```

! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    2    0    0    0
! 1.8+ 0    5    0    0    0
! 2.6+ 0    0   76   1    0
! 3.4+ 0    0    0  125  27
! 4.2+ 0    0    0    3   97

```

```
! Total number of targets : 336
! Total correct : 303
! Percentage correct : 90.18%
```

Bahasa Melayu 2 with All Attributes

The final model for Bahasa Melayu 2 and all attributes is:

Table 4.76: BM2 with all attributes model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
!
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    0    0    0
! 1.8+ 0    1    0    0    0
! 2.6+ 0    0    9    0    0
! 3.4+ 0    0    0   14   2
! 4.2+ 0    0    0    0   16
! Total number of targets : 42
! Total correct : 40
! Percentage correct : 95.24%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    2    0    0    0
! 1.8+ 0    5    0    0    0
! 2.6+ 0    0   76   1    0
! 3.4+ 0    0    0  127  25
! 4.2+ 0    0    0    3   97
```

```
! Total number of targets : 336
! Total correct : 305
! Percentage correct : 90.77%
```

c) Bahasa Ingggris

The final model for Bahasa Ingggris is:

Table 4.77: BI model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	3
Output Unit	5
Weight Seed	3
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!          1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      1    0    0    0
! 1.8+ 0      5    1    0    0
! 2.6+ 0      1   12    0    0
! 3.4+ 0      0    0   10    0
! 4.2+ 0      2    0    1    9
! Total number of targets : 42
! Total correct : 36
! Percentage correct : 85.71%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!          1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0     10    0    0    0
! 1.8+ 0     32    6    0    0
! 2.6+ 0      8   106   0    0
! 3.4+ 0      8    4   101   4
! 4.2+ 0      0   12    5   40
```

! Total number of targets : 336
 ! Total correct : 279
 ! Percentage correct : **83.04%**

Bahasa Ingeris with Gender

The final model for Bahasa Ingeris and gender is:

Table 4.78: BI with gender model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      1      0      0      0
! 1.8+ 0      5      1      0      0
! 2.6+ 0      1     12      0      0
! 3.4+ 0      0      0      9      1
! 4.2+ 0      2      0      1      9
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0     10      0      0      0
! 1.8+ 0     32      6      0      0
! 2.6+ 0      8     106     0      0
! 3.4+ 0      8      4     99      6
! 4.2+ 0      0     12      3     42
! Total number of targets : 336
! Total correct : 279
```

! Percentage correct : **83.04%**

Bahasa Inggeris with Attendance

The final model for Bahasa Inggeris and attendance is:

Table 4.79: BI with attendance model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      1      0      0      0
! 1.8+ 0      5      1      0      0
! 2.6+ 0      1     12     0      0
! 3.4+ 0      0      0      9      1
! 4.2+ 0      2      0      1      9
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0     10     0     0     0
! 1.8+ 0     32     6     0     0
! 2.6+ 0     8     106    0     0
! 3.4+ 0     8      4     99     6
! 4.2+ 0     0     12     3     42
! Total number of targets : 336
! Total correct : 279
! Percentage correct : 83.04%
```

Bahasa Inggeris with Co-curricular Activities

The final model for Bahasa Inggeris and co-curricular activities is:

Table 4.80: BI with co-curricular model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	2
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	2100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    5    1    0    0
! 2.6+ 0    1   12   0    0
! 3.4+ 0    0    0    9    1
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    10   0    0    0
! 1.8+ 0    32   6    0    0
! 2.6+ 0    8   106  0    0
! 3.4+ 0    8    4   99   6
! 4.2+ 0    0   12   3   42
! Total number of targets : 336
! Total correct : 279
! Percentage correct : 83.04%
```

Bahasa Inggeris with Position in Family

The final model for Bahasa Inggeris and position in family is:

Table 4.81: BI with position in family model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	2
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    5    1    0    0
! 2.6+ 0    1   12    0    0
! 3.4+ 0    0    0    9    1
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    10   0    0    0
! 1.8+ 0    32   6    0    0
! 2.6+ 0    8   106  0    0
! 3.4+ 0    8    4   99   6
! 4.2+ 0    0   12   3   42
! Total number of targets : 336
! Total correct : 279
! Percentage correct : 83.04%
```

Bahasa Inggeris with Father's Job

The final model for Bahasa Inggeris and father's job is:

Table 4.82: BI with father's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    6    0    0    0
! 2.6+ 0    1   12   0    0
! 3.4+ 0    0    0    9    1
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 36
! Percentage correct : 85.71%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    9    1    0    0
! 1.8+ 0   34   4    0    0
! 2.6+ 0    8   106  0    0
! 3.4+ 0    7    5   92   13
! 4.2+ 0    0   12   1   44
! Total number of targets : 336
! Total correct : 276
! Percentage correct : 82.14%
```

Bahasa Inggeris with Mother's Job

The final model for Bahasa Inggeris and mother's job is:

Table 4.83: BI with mother's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 1    0    0    0    0
! 1.8+ 0    6    0    0    0
! 2.6+ 0    0   13   0    0
! 3.4+ 0    0    0   10   0
! 4.2+ 0    0    0    0   12
! Total number of targets : 42
! Total correct : 42
! Percentage correct : 100.00%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 10   0    0    0    0
! 1.8+ 0   38   0    0    0
! 2.6+ 0    0  114   0    0
! 3.4+ 0    0    0  117   0
! 4.2+ 0    0    0    0   57
! Total number of targets : 336
! Total correct : 336
! Percentage correct : 100.00%
```

Bahasa Inggeris with Family Income

The final model for Bahasa Inggeris and family income is:

Table 4.84: BI with family income model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      1    0    0    0
! 1.8+ 0      6    0    0    0
! 2.6+ 0      2    9    2    0
! 3.4+ 0      0    0    8    2
! 4.2+ 0      2    0    1    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      10   0    0    0
! 1.8+ 0      37   1    0    0
! 2.6+ 0      12  96    6    0
! 3.4+ 0      11   1   85   20
! 4.2+ 0       2   6    6   43
! Total number of targets : 336
! Total correct : 261
! Percentage correct : 77.68%
```

Bahasa Ingeris with All Attributes

The final model for Bahasa Ingeris and all attributes is:

Table 4.85: BI with all attributes model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      1    0    0    0
! 1.8+ 0      6    0    0    0
! 2.6+ 0      2    9    2    0
! 3.4+ 0      0    0    8    2
! 4.2+ 0      2    0    1    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True           Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      10   0    0    0
! 1.8+ 0      37   1    0    0
! 2.6+ 0      12  97   5    0
! 3.4+ 0      11   1   78  27
! 4.2+ 0       2   6    7   42
! Total number of targets : 336
! Total correct : 254
! Percentage correct : 75.60%
```

d) Mathematics

The final model for Mathematics is:

Table 4.86: MT model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! -----
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    1    0    0
! 1.8+ 0    0    4    0    0
! 2.6+ 0    0   13    1    0
! 3.4+ 0    0    1   10    0
! 4.2+ 0    1    2    0    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! -----
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    5    0    0
! 1.8+ 0    0   35    0    0
! 2.6+ 0    1  105    6    0
! 3.4+ 0    0   16  117    0
! 4.2+ 0    0   14    9   28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with Gender

The final model for Mathematics and gender is:

Table 4.87: MT with gender model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    1    0    0
! 1.8+ 0    0    4    0    0
! 2.6+ 0    0   13    1    0
! 3.4+ 0    0    1   10    0
! 4.2+ 0    1    2    0    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    5    0    0
! 1.8+ 0    0   35    0    0
! 2.6+ 0    1  105    6    0
! 3.4+ 0    0   16  117    0
! 4.2+ 0    0   14    9   28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with Attendance

The final model for Mathematics and attendance is:

Table 4.88: MT with attendance model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      0      1      0      0
! 1.8+ 0      0      4      0      0
! 2.6+ 0      0     13      1      0
! 3.4+ 0      0      1     10      0
! 4.2+ 0      1      2      0      9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      0      5      0      0
! 1.8+ 0      0     35      0      0
! 2.6+ 0      1    105      6      0
! 3.4+ 0      0     16     117      0
! 4.2+ 0      0     14      9     28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with Co-curricular Activities

The final model for Mathematics and co-curricular activities is:

Table 4.89: MT with co-curricular model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    1    0    0
! 1.8+ 0    0    4    0    0
! 2.6+ 0    0   13    1    0
! 3.4+ 0    0    1   10    0
! 4.2+ 0    1    2    0    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    5    0    0
! 1.8+ 0    0   35    0    0
! 2.6+ 0    1  105    6    0
! 3.4+ 0    0   16   117    0
! 4.2+ 0    0   14    9   28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with Position in Family

The final model for Mathematics and position in family is:

Table 4.90: MT with position in family model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    1    0    0
! 1.8+ 0    0    4    0    0
! 2.6+ 0    0   13    1    0
! 3.4+ 0    0    1   10    0
! 4.2+ 0    1    2    0    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    5    0    0
! 1.8+ 0    0   35    0    0
! 2.6+ 0    1  105    6    0
! 3.4+ 0    0   16  117    0
! 4.2+ 0    0   14    9   28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with Father's Job

The final model for Mathematics and father's job is:

Table 4.91: MT with father's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      0      1      0      0
! 1.8+ 0      0      4      0      0
! 2.6+ 0      0     13      1      0
! 3.4+ 0      0      1     10      0
! 4.2+ 0      1      2      0      9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!           1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      0      5      0      0
! 1.8+ 0      0     35      0      0
! 2.6+ 0      1    105      6      0
! 3.4+ 0      0     16     117      0
! 4.2+ 0      0     14      9     28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with Mother's Job

The final model for Mathematics and mother's job is:

Table 4.92: MT with mother's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    1    0    0
! 1.8+ 0    0    4    0    0
! 2.6+ 0    0   13    1    0
! 3.4+ 0    0    1   10    0
! 4.2+ 0    1    2    0    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    5    0    0
! 1.8+ 0    0   35    0    0
! 2.6+ 0    1  105    6    0
! 3.4+ 0    0   16  117    0
! 4.2+ 0    0   14    9   28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with Family Income

The final model for Mathematics and family income is:

Table 4.93: MT with family income model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    1    0    0
! 1.8+ 0    0    4    0    0
! 2.6+ 0    0   13    1    0
! 3.4+ 0    0    1   10    0
! 4.2+ 0    1    2    0    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    0    5    0    0
! 1.8+ 0    0   35    0    0
! 2.6+ 0    1  105    6    0
! 3.4+ 0    0   16  117    0
! 4.2+ 0    0   14    9   28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

Mathematics with All Attributes

The final model for Mathematics and all attributes is:

Table 4.94: MT with all attributes model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	3
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      0      1      0      0
! 1.8+ 0      0      4      0      0
! 2.6+ 0      0     13      1      0
! 3.4+ 0      0      1     10      0
! 4.2+ 0      1      2      0      9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0      0      5      0      0
! 1.8+ 0      0     35      0      0
! 2.6+ 0      1    105      6      0
! 3.4+ 0      0     16     117      0
! 4.2+ 0      0     14      9     28
! Total number of targets : 336
! Total correct : 250
! Percentage correct : 74.40%
```

e) Science

The final model for Science is:

Table 4.95: SC model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	2
Learning Rate	0.2
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result from the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!  
! True          Predicted  
! ----          -  
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+  
! 1.0+ 0    0    0    0    0  
! 1.8+ 0    1    1    0    0  
! 2.6+ 0    0   16    0    0  
! 3.4+ 0    1    0   10    0  
! 4.2+ 0    1    1    1   10  
! Total number of targets : 42  
! Total correct : 37  
! Percentage correct : 88.10%
```

Train

```
! ** Confusion Matrix For Output 1 **
!  
! True          Predicted  
! ----          -  
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+  
! 1.0+ 0    3    0    0    0  
! 1.8+ 0   10    5    0    0  
! 2.6+ 0    2   109    0    0  
! 3.4+ 0    4    4   133    1  
! 4.2+ 0    0   14    8   43  
! Total number of targets : 336  
! Total correct : 295  
! Percentage correct : 87.80%
```

Science with Gender

The final model for Science and gender is:

Table 4.96: SC with gender model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    5    1    0    0
! 2.6+ 0    1   12   0    0
! 3.4+ 0    0    0    9    1
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0   10    0    0    0
! 1.8+ 0   32    6    0    0
! 2.6+ 0    8   106   0    0
! 3.4+ 0    8    4   99    6
! 4.2+ 0    0   12    3   42
! Total number of targets : 336
! Total correct : 279
! Percentage correct : 83.04%
```

Science with Attendance

The final model for Science and attendance is:

Table 4.97: SC with attendance model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    5    1    0    0
! 2.6+ 0    1   12   0    0
! 3.4+ 0    0    0    9    1
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0   10    0    0    0
! 1.8+ 0   32    6    0    0
! 2.6+ 0    8   106   0    0
! 3.4+ 0    8    4   99    6
! 4.2+ 0    0   12    3   42
! Total number of targets : 336
! Total correct : 279
! Percentage correct : 83.04%
```

Science with Co-curricular Activities

The final model for Science and co-curricular activities is:

Table 4.98: SC with co-curricular model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	2
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	2100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    5    1    0    0
! 2.6+ 0    1   12    0    0
! 3.4+ 0    0    0    9    1
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 35
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0   10    0    0    0
! 1.8+ 0   32    6    0    0
! 2.6+ 0    8   106   0    0
! 3.4+ 0    8    4   99    6
! 4.2+ 0    0   12    3   42
! Total number of targets : 336
! Total correct : 279
! Percentage correct : 83.04%
```

Science with Position in Family

The final model for Science and position in family is:

Table 4.99: SC with position in family model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	2
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!  
! True          Predicted  
! ----          -  
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+  
! 1.0+ 0    1    0    0    0  
! 1.8+ 0    5    1    0    0  
! 2.6+ 0    1   12    0    0  
! 3.4+ 0    0    0    9    1  
! 4.2+ 0    2    0    1    9  
! Total number of targets : 42  
! Total correct : 35  
! Percentage correct : 83.33%
```

Train

```
! ** Confusion Matrix For Output 1 **
!  
! True          Predicted  
! ----          -  
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+  
! 1.0+ 0   10    0    0    0  
! 1.8+ 0   32    6    0    0  
! 2.6+ 0    8   106   0    0  
! 3.4+ 0    8    4   99    6  
! 4.2+ 0    0   12    3   42  
! Total number of targets : 336  
! Total correct : 279  
! Percentage correct : 83.04%
```

Science with Father's Job

The final model for Science and father's job is:

Table 4.100: SC with father's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    6    0    0    0
! 2.6+ 0    1   12   0    0
! 3.4+ 0    0    0    9    1
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 36
! Percentage correct : 85.71%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    9    1    0    0
! 1.8+ 0   34   4    0    0
! 2.6+ 0    8   106  0    0
! 3.4+ 0    7    5   92   13
! 4.2+ 0    0   12   1   44
! Total number of targets : 336
! Total correct : 276
! Percentage correct : 82.14%
```

Science with Mother's Job

The final model for Science and mother's job is:

Table 4.101: SC with mother's job model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!  
! True          Predicted  
! ----          -  
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+  
! 1.0+ 1    0    0    0    0  
! 1.8+ 0    6    0    0    0  
! 2.6+ 0    0   13   0    0  
! 3.4+ 0    0    0   10   0  
! 4.2+ 0    0    0    0   12  
! Total number of targets : 42  
! Total correct : 42  
! Percentage correct : 100.00%
```

Train

```
! ** Confusion Matrix For Output 1 **
!  
! True          Predicted  
! ----          -  
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+  
! 1.0+ 10   0    0    0    0  
! 1.8+ 0   38   0    0    0  
! 2.6+ 0    0  114   0    0  
! 3.4+ 0    0    0  117   0  
! 4.2+ 0    0    0    0   57  
! Total number of targets : 336  
! Total correct : 336  
! Percentage correct : 100.00%
```

Science with Family Income

The final model for Science and family income is:

Table 4.102: SC with family income model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	2
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    1    0    0    0
! 1.8+ 0    6    0    0    0
! 2.6+ 0    2    9    2    0
! 3.4+ 0    0    0    8    2
! 4.2+ 0    2    0    1    9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0    10   0    0    0
! 1.8+ 0    37   1    0    0
! 2.6+ 0    12  96   6    0
! 3.4+ 0    11   1   85   20
! 4.2+ 0     2   6    6   43
! Total number of targets : 336
! Total correct : 261
! Percentage correct : 77.68%
```

Science with All Attributes

The final model for Science and all attributes is:

Table 4.103: SC with all attributes model

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	6
Hidden Unit	1
Output Unit	5
Weight Seed	1
Learning Rate	0.1
Momentum Rate	0.1
Activation Function	Sigmoid
Stopping Criteria	100

The result for the above model is:

Test

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0   1   0   0   0
! 1.8+ 0   6   0   0   0
! 2.6+ 0   2   9   2   0
! 3.4+ 0   0   0   8   2
! 4.2+ 0   2   0   1   9
! Total number of targets : 42
! Total correct : 32
! Percentage correct : 76.19%
```

Train

```
! ** Confusion Matrix For Output 1 **
!
! True          Predicted
! ----          -
!      1.0+ 1.8+ 2.6+ 3.4+ 4.2+
! 1.0+ 0   10  0   0   0
! 1.8+ 0   37  1   0   0
! 2.6+ 0   12 97   5   0
! 3.4+ 0   11  1  78  27
! 4.2+ 0   2   6   7  42
! Total number of targets : 336
! Total correct : 254
! Percentage correct : 75.60%
```

4.4 Relationship Between Each Factors

The relationship between each subjects and attributes are shown below:

Table 4.104 (a): Relationship between subjects and attributes

Subjects	Academic		Gender		Attendance		Co-curricular		Position in Family	
	Train (%)	Test (%)	Train (%)	Test (%)	Train (%)	Test (%)	Train (%)	Test (%)	Train (%)	Test (%)
BM1	80.95	83.33	81.25	85.71	80.95	83.33	80.95	83.33	80.95	83.33
BM2	86.90	90.48	90.77	95.24	89.58	95.24	89.58	95.24	89.58	95.24
BI	83.04	85.71	83.04	83.33	83.04	83.33	83.04	83.33	83.04	83.33
MT	74.40	76.19	74.40	76.19	74.40	76.19	74.40	76.19	74.40	76.19
SC	87.80	88.10	83.04	83.33	83.04	83.33	83.04	83.33	83.04	83.33
Average	82.62	84.76	82.50	84.76	82.20	84.28	82.20	84.28	82.20	84.28

Table 4.104 (b): Relationship between subjects and attributes

Subjects	Father's Job		Mother's Job		Family Income		All	
	Train (%)	Test (%)	Train (%)	Test (%)	Train (%)	Test (%)	Train (%)	Test (%)
BM1	80.95	83.33	80.95	83.33	81.25	83.33	80.95	83.33
BM2	89.58	95.24	89.58	95.24	90.18	95.24	90.77	95.24
BI	82.14	85.71	100.00	100.00	77.68	76.19	75.60	76.19
MT	74.40	76.19	74.40	76.19	74.40	76.19	74.40	76.19
SC	82.14	85.71	100.00	100.00	77.68	76.19	75.60	76.19
Average	81.84	85.24	88.99	90.95	80.24	81.43	79.46	81.43

The results showed the high accuracy between each subject and attributes. All subjects have strong relationship with their factors based on the high percentage of accuracy. The academic factor showed that the past examination results bring the significant influence towards the future students' performance. The gender result showed the contribution to the students' performance.

It achieved the same accuracy like academic factor which is average 84.76% of accuracy. The male students gave the better result in BM1 and BM2 while the female students better than male in BI, MT and Sc. The attendance factor also brought the important effect towards the performance. It achieved average 84.28% of accuracy. The good attendance students gave the better results than the bad attendance students. The participation in co-curricular activities also affects the students' performance in future. The active students gave the better results compared to the passive students. The position in family also brought the meaningful contribution to the performance. The sixth position in family gave the highest result in BM1 and BM2 while the second position in family children gave the best result in BI.

The first position in family children are good in MT and the youngest children in family tend to excel in SC subject. Both co-curricular and position in family factors achieved the same percentage of accuracy which is average of 84.28%. The students with their father and mother in education field gave the better performance than the other parents' job. The father's job achieved average 85.24% while the mother's job got average 90.95% of accuracy. The students from medium family status tend to excel in their study the the other family economic status. The family income factor achieved average 81.43% of accuracy.

BM2, BI and SC gave the better result with more than 90% accuracy among the other subjects. However, the percentage of accuracy against BI and SC for mother's job factor is 100%. It means the mother's job factor influenced strongly on the students' performance. The percentage of accuracy for BM1 is 83.33% for all attributes except gender (85.71%). BM2 produced more

than 90% of accuracy while BI achieved around 80% of accuracy for all variables except around 70% for family income and 100% for mother's job. MT achieved the lowest percentage of accuracy which is around 70% for all factors. SC subject also achieved around 80% for all attributes except around 70% for family income and 100% for mother's job. The academic factors (examination) and demographics factors (gender, attendance, co-curricular activities, position in family, father's job, mother's job and family income) strongly influenced the students' performance in UPSR with 81.43% of accuracy.

4.5 Comparison Methods

The comparison between NN and regression is shown below:

Table 4.105: Comparison between NN and regression

Method	NN	Regression
Academic factors	84.76%	81.62%
Demographic factors	85.03%	50.08%

The NN analysis using academic factors achieved 84.76% of accuracy while the regression analysis is only 81.62%. NN with demographic factors obtained 34.95% higher than regression analysis. In conclusion, NN gave the better percentage of accuracy compared with regression analysis.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

This chapter wrapping up the contribution of the study, as well as presenting the problems and limitations encountered in conducting the study. In addition, the discussion on the recommendation for the future work is also presented.

5.1 Conclusion

This study proves that examination factors such as previous results and the other factors for example gender, attendance, co-curricular activities and family background help to brighten the students' performance in the future. There is a strong relationship between those factors. The NN is used to present a fairly good prediction of the students' performance.

Data of previous results, gender, attendances, co-curricular and family background are used in training the network to search the prediction model. The best result model then is tested for the next year prediction. This study shows that MLP can be used in modelling the students'

performance in the next UPSR. However, other data mining technique such as decision trees can be used in the future to extract rules from the data. Association rules could also be used to measure the association between attributes. The findings could be used to further enhance the strength of each attribute with prediction variables and among attributes or independent variables.

5.2 Recommendations

Further works can also be done on predicting performance in other public examinations like Penilaian Menengah Rendah (PMR), Sijil Pelajaran Malaysia (SPM), Sijil Tinggi Pelajaran Malaysia (STPM) and Sijil Tinggi Agama Malaysia (STAM). As a conclusion, the findings opened new avenues for future research in education field itself.

REFERENCES

- Larose D., (2006). *Data Mining Methods and Models*. New Jersey: John Wiley & Sons Inc, 93-149.
- Hand D., Mannila H., Smyth P., (2001). *Principles of Data Mining*. Cambridge: MIT Press.
- Haykin S., (2007). *Neural Networks: A Comprehensive Foundation*. New York: Macmillan Publishing, 156-175.
- Khoa N., Sakakibara K., Nishikawa I., (2006). Stock Price Forecasting Using Back Propagation Neural Networks with Time and Profit Based Adjusted Weight Factors. *Electronics & Communication Engineering Journal*, 5484 – 5488.
- Fung C., Iyer V., Brown W., Wong K., (2005). Comparing The Performance of Different Neural Networks Architectures for The Prediction of Mineral Prospectivity. *Electronics & Communication Engineering Journal*, vol. 1, 394 – 398.
- Wang T., Mitrovic A., (2002). Using Neural Networks to Predict Student's Performance. *Electronics & Communication Engineering Journal*, 969 – 973.
- Huisken G., Coffa A., (2000). Short-term Congestion Prediction: Comparing Time Series with Neural Networks. *Electronics & Communication Engineering Journal*, 66 – 69.
- Cripps A., (1996). Using Artificial Neural Nets to Predict Academic Performance. *American Psychological Association Journal*, 33 – 37.
- Hostetler T., (1996). Predicting Student Success in An Introductory Programming Course. *American Psychological Association Journal*, pp. 40 – 49.
- Butcher D., Muth W., (1985). Predicting Performance in An Introductory Computer Science Course. *American Psychological Association Journal* , 263 – 268.
- Bae, Yupin, Smith, Thomas M., (1997). Women in Mathematics and Science. *Education Resources Information Center Journal*.
- Gibson, Margaret A., (2003). Improving Graduation Outcomes for Migrant Students. *Education Resources Information Center Journal*.
- Terry E., Spradlin, Kirk R., Walcott C., Kloosterman P., Zaman K., McNabb S., Zapf J. & associates, (2005). Is The Achievement Gap in Indiana Narrowing. *Education Resources Information Center Journal*.

- Hayek, John C., Kuh, George D., (1999). College Activities and Environmental Factors Associated with The Development of Life Long Learning Competencies of College Seniors. *Education Resources Information Center Journal*.
- Erbe, Mach B., (2000). Correlates of School Achievement in Chicago Elementary Schools. *Education Resources Information Center Journal*.
- Henchey, Norman, (2002). Schools That Make A Difference : Final Report. Twelve Canadian Secondary Schools in Low Income Settings. *Education Resources Information Center Journal*.
- Dai, Huang, (2006). Data Mining Used in Rule Design for Active Database Systems. *Electronics & Communication Engineering Journal*.
- Young, Ho, Il., Ok, Suk, (2003). The Reliability of Pollution with Regression Analysis and The Possibility of Dispersion and Receptor Models. *Electronics & Communication Engineering Journal*.
- Yusof & Syed Hassan, (2000). Regression Analysis of Consonant Frequency over Number of Turn of Normal Mode Helical Antenna. *Electronics & Communication Engineering Journal*.
- Bon, Ogier, Razali, (2007). Modelling in Manufacturing Industry : Parameters Selection Using Regression Analysis. *Electronics & Communication Engineering Journal*.
- Perez, Gonzalez, Salinas, (2000). Neural versus Difference Equation Modeling for 2D Pattern Recognition Problems. *Electronics & Communication Engineering Journal*.
- Fausett, Elwasif, (1994). Predicting Performance From Test Scores Using Backpropagation and Counterpropagation . *Electronics & Communication Engineering Journal*.
- Burtner, (2004). Critical-To-Quality Factors Associated with Engineering Student Persistence: The Influence of Freshman Attitudes. *Electronics & Communication Engineering Journal*.
- Han, Cheng, Meng, (2003). Application of Four-layer Neural Network on Information Extraction. *Electronics & Communication Engineering Journal*.
- Walczak, (1994). Categorizing University Student Applicants with Neural Networks. *Electronics & Communication Engineering Journal*.
- Mahmoud E., (2004). Accuracy in Forecasting: A Survey. *Journal of Forecasting*.
- Friedman D., Montgomery D., (1985). Evaluation of The Predictive Performance of Biased Regression Estimators. *Journal of Forecasting*.

- Tsukamoto Y., Namatame A., (1996) Evolving Neural Network Models. *Electronics & Communication Engineering Journal*.
- Phua C., Alahakoon D., Lee V. (2000). Minority Report in Fraud Detection: Classification of Skewed Data. *Sigkdd Explorations*.
- Schultz R., (1984). The Implementation of Forecasting Models. *Journal of Forecasting*.
- Klein L., (1984). The Importance of The Forecast. *Journal of Forecasting*.
- Huss W., (1985). The Teachers/Practitioners Corner Comparative Analysis of Load Forecasting Techniques at a Southern Utility. *Journal of Forecasting*.
- Gruca T., Klemz B., Petersen E., (1999). Mining Sales Data Using A Neural Network Model of Market Response. *SIGKDD Explorations*.
- Lu C., Brabanter J., Huffel S., Vergote I., Timmeman D. (2001). Using Artificial Neural Networks To Predict Malignancy of Ovarian Tumors. *Proceedings of the 23rd Annual EMBS International Conference*.
- Atiya A., El-Shoura S., Shaheen S., El-Sherif M. (1999). "A Comparison Between Neural-Network Forecasting Techniques – Case Study: River Flow Forecasting" *Electronics & Communication Engineering Journal*.
- Ringwood J., Galvin G. (2002). Computer-Aided Learning in Artificial Neural Networks. *Electronics & Communication Engineering Journal*.
- Burke H., Rosen D., Goodman P., (1994). Comparing Artificial Neural Networks To The Other Statistical Methods For Medical Outcome Prediction. *Electronics & Communication Engineering Journal*.
- Sitte R., Sitte J. (2000). Analysis Of The Predictive Ability Of Time Delay Neural Networks Applied To The S&P 500 Time Series. *Electronics & Communication Engineering Journal*.
- Jiuzhen L., Jiaqing Z., (2002). Procedure Neural Networks With Supervised Learning. *Proceeding of the 9th International Conference on Neural Information Processing*.
- Ritschel, Pfeifer W., Grob (1994). Rating of Pattern Classifications in Multi-layer Perceptrons: Theoretical Background And Practical Results. *ACM*.
- Mahamud K., Bakar A., Norwawi N. (1998). Neural Network Modelling To Predict House Prices Performance. *Research Report*.

- Roy A., (2000). Artificial Neural Networks – A Science in Trouble. *SIGKDD Explorations*, 33.
- Chamillard A., (2006). Using Student Performance Predictions in A Computer Science Curriculum. *ACM*.
- Spradlin T., Kirk R., Walcott C., Kloosterman P., Zaman K., McNabb S., Zapf J., (2005). Is The Achievement Gap in Indiana Narrowing? *Center for Evaluation & Education Policy*.
- Sweet J., Rasher S., Abromitis B., Johnson E., (2004). Case Studies of High-Performing, High-Technology Schools. *North Central Regional Educational Laboratory*.
- Williams, Bugg D., (1998). Parent Involvement Gender Effects On Preadolescent Student Performance. *American Educational Research Association*.
- Verna, Ann M., Spina, Maria, (2002). Parental Processes and Self-Concepts Effect The Academic Achievement of Italian Students. *American Educational Research Association*.
- Sengupta S., (1995). A Comparative Study of Neural Network and Regression Analysis As Modelling Tools. *Electronics & Communication Engineering Journal*.
- Li Y., (1994). A General Linear-Regression Analysis Applied To The 3-Parameter Weibull Distribution. *Electronics & Communication Engineering Journal*.
- Itsuki R., Yajima H., Mizuno H., Kinukawa H. (1996). Application and Verification of Using Statistical Analysis Tool and Expert System Together in Multiple Regression Analysis. *Electronics & Communication Engineering Journal*.
- Fukuda T., Shibata T., Tokita M., Mitsuoka T. Neural Network Application for Robotic Motion Control. *Research Report*
- Wong K., (1993). Artificial Intelligence and Neural Network Applications In Power Systems. *Electronics & Communication Engineering Journal*.
- Means R., (1994). High Speed Parallel Hardware Performance Issues for Neural Network Applications. *Electronics & Communication Engineering Journal*.
- El-Sharkawi M., Huang S., (1994). Ancillary Techniques for Neural Network Applications. *Electronics & Communication Engineering Journal*.
- Vonk E., Jain L., Veelenturf L., (1995). Neural Network Applications. *Electronics & Communication Engineering Journal*.

- Sugiyama S, A Neural Network Application To Semantic and Logistic Recognition. *Electronics & Communication Engineering Journal*, 1995.
- Bershad N., Ibnkahla M., Castanie F., (1997). Statistical Analysis of A Two-Layer Backpropagation Algorithm Used For Modelling Nonlinear Memoryless Channels: The Single Neuron Case. *Electronics & Communication Engineering Journal*.
- Hansen J., Nelson R., (1997). Neural Networks and Traditional Time Series Methods: A Synergistic Combination in State Economic Forecasts. *Electronics & Communication Engineering Journal*.
- Ye Q., Liang B., Li Y., (2005). Amnestic Neural Network for Classification: Application on Stock Trend Prediction. *Electronics & Communication Engineering Journal*.
- Lisboa P., Taylor M., (1993). Workshop on Neural Network Applications and Tools. *Electronics & Communication Engineering Journal*.
- Tal B., (2003). Background Information On Our Neural Network – Based System of Leading Indicators. *CIBC World Markets*.
- Bojkovic Z., Milovanovic D., Mastorakis N., (2000). Neural Networks Applications For Multimedia Processing. *Electronics & Communication Engineering Journal*.
- Chowdhury F., Wahi P., Raina R., Kaminedi S., (2001). A Survey of Neural Networks Applications in Automatic Control. *Electronics & Communication Engineering Journal*.
- Rajagopalan R., Rajagopalan P., (1996). Applications of Neural Network in Manufacturing. *Electronics & Communication Engineering Journal*.
- Bose B., (2007). Neural Network Applications in Power Electronics and Motor Drives – An Introduction and Perspective. *Electronics & Communication Engineering Journal*.
- Wikipedia, (2008). Statistics. <http://en.wikipedia.org/wiki/Statistics>.

Appendix B: Numerical Data (xls)

YEAR	CLASS	NAME	plksr1_06				plksr2_06				gempar1_07				gempar2_07				plksr1_07													
			BMI	BM2	BI	MT	SC	AVG	BMI	BM2	BI	MT	SC	AVG	BMI	BM2	BI	MT	SC	AVG	BMI	BM2	BI	MT	SC							
6	1	1	81	73	79	81	78	78	71	52	80	85	63	70	98	70	80	90	82	84	68	54	87	98	76	77	52	79	89	89	70	
6	1	2	60	37	60	60	36	51	83	63	63	63	63	67	67	34	67	55	17	48	86	55	55	67	67	66	55	55	55	55	86	
6	1	3	36	60	36	60	60	50	63	14	14	34	34	32	34	32	34	17	34	40	55	17	17	34	34	31	17	17	55	55	55	
6	1	4	33	37	37	36	37	36	63	63	63	14	34	42	34	34	34	17	34	31	34	55	55	17	34	39	86	86	55	17	17	
6	1	5	65	60	60	60	60	61	63	63	63	63	83	67	67	67	67	67	67	67	67	55	67	67	86	68	55	55	55	55	55	
6	1	6	32	37	60	37	37	41	34	34	63	63	14	42	34	34	34	34	34	38	34	34	55	55	17	39	55	55	55	55	55	
6	1	7	60	60	36	60	36	50	63	63	63	63	63	63	55	55	17	55	17	40	55	55	55	55	55	55	17	55	55	55	55	
6	1	8	37	60	60	37	37	46	63	63	34	63	63	63	63	67	67	67	67	67	67	86	86	67	86	78	55	55	55	55	55	
6	1	9	60	60	60	60	60	60	83	83	83	83	83	75	67	67	67	67	67	67	86	86	67	86	78	55	55	55	55	55	55	
6	1	10	81	81	60	81	81	77	83	83	63	63	63	71	86	86	67	86	86	86	86	86	67	67	75	55	55	55	55	55	55	
6	1	11	60	81	60	81	60	68	83	83	63	63	63	71	67	86	67	86	67	75	86	86	67	67	75	86	55	55	55	55	86	
6	1	12	60	37	36	37	60	46	63	34	34	63	34	46	55	34	17	34	55	39	55	34	34	55	34	42	55	55	55	55	55	
6	1	13	81	79	83	81	81	81	83	83	80	63	85	79	86	80	83	86	87	84	86	83	90	67	85	82	86	55	95	55	85	85
6	1	14	60	60	60	60	60	60	83	83	63	63	83	75	67	55	67	55	67	62	86	86	67	67	86	78	55	55	17	17	55	
6	1	15	37	60	81	60	60	60	83	63	63	63	63	67	34	67	86	55	67	62	86	67	67	55	55	66	55	55	55	55	55	
6	1	16	60	60	81	60	60	64	63	34	63	63	34	51	67	55	86	67	68	55	34	55	55	55	34	47	55	55	55	86	86	
6	1	17	37	60	60	37	36	46	63	63	14	14	63	43	34	55	55	34	17	39	67	55	17	17	55	42	55	55	55	55	55	
6	1	18	81	81	60	81	60	73	83	63	63	63	63	67	86	67	86	67	78	67	86	67	67	67	71	17	17	17	86	17	55	
6	1	19	60	36	37	37	37	41	63	63	63	63	63	63	55	17	34	34	34	35	55	67	67	55	60	67	67	67	67	67	67	
6	1	20	60	60	60	60	60	60	63	63	63	63	63	63	55	67	55	67	62	67	67	67	55	55	60	67	86	86	67	67	67	
6	1	21	81	60	60	60	81	68	63	34	14	14	34	32	86	67	67	86	75	55	34	17	17	34	31	86	67	67	67	67	67	
6	1	22	37	60	37	36	36	41	63	63	63	63	63	63	63	55	34	17	17	31	55	55	55	55	55	55	67	67	67	67	67	
6	1	23	60	60	81	81	60	68	83	83	83	83	63	75	67	67	86	86	67	75	86	86	86	67	78	86	67	67	67	67	67	
6	1	24	60	37	60	36	60	51	63	63	34	14	34	42	55	34	55	17	55	43	55	34	17	34	39	67	67	67	67	67	67	
6	1	25	60	60	60	60	60	60	63	63	63	63	63	63	55	67	55	67	62	67	67	67	67	67	67	67	67	67	67	67	67	
6	1	26	81	60	60	81	81	73	83	63	63	63	63	67	86	67	86	67	86	76	86	67	67	55	66	67	17	17	17	86	67	
6	1	27	37	36	60	37	60	46	63	63	34	34	63	51	34	17	55	34	67	41	55	55	34	34	67	49	86	67	67	67	67	
6	1	28	60	60	37	60	60	55	14	34	34	63	63	42	55	67	34	67	58	17	17	34	34	55	55	39	67	67	67	67	67	
6	1	29	60	60	60	60	81	60	63	63	63	63	83	71	67	67	34	55	86	62	55	55	67	86	86	70	67	67	67	67	67	
6	1	30	60	81	60	60	37	60	63	63	63	63	83	71	67	86	55	55	34	59	67	67	55	86	86	72	67	67	67	67	67	
6	1	31	41	41	65	41	32	44	44	44	44	44	44	42	44	44	65	44	32	46	44	44	32	44	44	42	10	10	65	6	65	
6	2	32	41	41	65	81	81	62	60	60	44	44	44	50	44	44	65	90	90	67	65	65	44	44	44	42	65	65	10	6	65	
6	2	33	32	41	18	65	65	44	14	32	32	44	44	33	32	44	18	65	65	45	17	32	32	44	44	34	65	65	10	6	65	
6	2	34	65	65	41	32	81	57	44	44	60	44	44	47	65	65	44	32	90	59	44	44	65	44	44	48	65	10	65	65	65	
6	2	35	81	41	65	65	65	63	80	80	80	80	80	80	90	44	65	65	65	66	90	90	90	90	90	90	65	65	65	65	90	
6	2	36	65	41	41	41	32	44	44	60	60	44	44	50	65	44	44	44	32	46	44	65	65	44	44	52	65	65	65	65	65	
6	2	37	65	32	18	9	41	33	32	44	44	44	32	60	42	65	32	18	9	44	34	32	44	44	32	65	10	65	65	65	65	
6	2	38	65	65	81	81	32	65	44	60	44	44	44	45	65	65	90	90	32	68	44	65	44	44	44	46	90	90	65	65	65	
6	2	39	81	65	41	32	32	59	32	32	44	44	12	33	90	65	44	44	32	53	32	32	44	44	12	33	65	65	65	65	65	
6	2	40	65	41	41	65	65	55	44	44	60	60	80	58	65	44	44	65	65	57	44	44	65	65	90	62	65	65	10	6	65	
6	2	41	41	32	41	9	32	31	60	60	80	80	80	68	44	32	44	9	32	32	65	65	90	65	90	75	65	10	65	6	6	
6	2	42	65	65	65	81	81	71	44	44	60	60	80	58	65	65	65	65	90	75	44	44	65	65	90	62	90	90	65	65	65	
6	2	43	65	32	41	32	41	42	60	60	44	44	32	46	65	32	44	32	44	43	65	65	44	32	32	48	90	65	65	6	65	
6	2	44	41	41	41	81	81	62	80	60	60	80	68	44	44	44	65	90	90	67	90	65	65	65	10	65	65	65	65	65	65	
6	2	45	18	41	32	41	41	35	44	60	44	44	32	45	18	44	44	44	36	44	44	65	44	44	32	46	65	65	65	65	90	
6	2	46	65	65	81	81	81	71	44	44	60	44	44	50	65	65	65	90	90	75	44	44	65	65	44	52	65	65	65	65	65	
6	2	47	32	41	18	41	41	35	60	44	44	44	32	44	44	44	44	44	36	44	44	65	44	44	44	46	90	90	90	90	90	
6	2	48	65	41	65	65	65	32	54	44	44	80	80	62	65	44	65	65	32	54	44	44	90	90	65	67	65	65	6	65	6	
6	2	49	65	65	65	65	41	41	55	60	80	44	44	62	65	65	65	44	44	57	65	90	44	90	44	67	65	10	65	65	65	
6	2	50	41	41	41	32	32	37	80	80	44	44	60	44	62	44	44	32	39	62	44	90	44	90	44	67	65	10	6	65	65	
6	2	51	81	81	65	65	65	81	75	44	12	32	32	33	90	90	65	65	90	80	44	44	12	32	32	33	65	65	10	6	65	
6	2	52	65	41	81	81	41	62	80	80	44	44	80	60	62	65	44	90	44	67	90	44	44	90	65	67	65	65	65	65	65	

