

Chapter 3

Preliminary Data Analysis

In this chapter we described the preliminary data analysis based on the records obtained from English Vocabulary Skill Test and the university MIS (Management Information System) database of Prince of Songkla University, Pattani Campus. The subjects investigated in this study were 785 undergraduate students studying in the second semester of the 2005 academic year at Prince of Songkla University, Pattani Campus. In this chapter the frequency distributions of the variables were shown, and then the associations between these variables were presented.

3.1 Variable types

The roles of the variables may be classified as determinants and outcomes. The variables and their roles and data types are listed in Table 3.1.

As shown in Table 3.1, there were 10 variables. Five (gender, religion, degree duration, high school type, and home provinces) were binary, three (seniority level, faculty of study, and major field of study) were nominal, and the time taken to complete the test was taken to be continuous.

The outcome of interest in the study was English Vocabulary Skill Test score, which was of continuous data type.

Variables	Roles	Types
Subjects	identifier	
Gender	determinant	binary
Religion	determinant	binary
Degree duration	determinant	binary
Seniority level	determinant	nominal (4)
Faculty of study	determinant	nominal (6)
Major field of study	determinant	nominal (7)
High school type	determinant	binary
Home province	determinant	binary
Time taken to complete the test	determinant	continuous
English Vocabulary Skill Test score	outcome	continuous

Table 3.1: Variables and their roles and data types

3.2 Frequency distribution of the determinants

Table 3.2 shows the frequency distributions of the categorical determinants.

Of the 785 sample subjects, most of the undergraduate students were female (78.6%). The percentages in Islamic (50.3%) and other religions (49.7%) were approximately the same. Slightly more than one third (35.7%) of undergraduate students had finished university preparation at an Islamic religious high school. For home province, 59.6% came from the 5 Southern border provinces; Satun, Songkla, Pattani, Yala, and Narathiwat.

Most of undergraduate students (90.7%) studied in a 2 or 4 year degree program. The percentages in 1st, 2nd, 3rd, 4th or higher years were 25.9, 34.6, 21.4, and 18.1, respectively.

By faculty, 33% were from Science and Technology, 23.1% from Education, 20.5% from Humanities and Social Sciences, and 12.5% from the College of Islamic Studies.

In terms of the major field of study, the main majors comprised sciences (43.8%), languages (24.1%), and Islamic studies (13.2%).

Determinants	Category	Count	Percent
Gender	Female	617	78.6
	Male	168	21.4
Religion	Islamic	395	50.3
	Other	390	49.7
Degree duration	2 or 4 years	712	90.7
	5 years	73	9.3
Seniority level	1 st year	203	25.9
	2 nd year	272	34.6
	3 rd year	168	21.4
	4 th or higher years	142	18.1
Faculty of study	Education	181	23.1
	Humanities and Social Sciences	161	20.5
	Science and Technology	259	33.0
	College of Islamic Studies	98	12.5
	Communication Science	41	5.2
	Fine and Applied Arts	45	5.7
Major field of study	Science	344	43.8
	Communication Science	55	7.0
	Social Sciences	28	3.6
	Education	19	2.4
	Islamic Studies	104	13.2
	Arts	46	5.9
	Languages	189	24.1
High school type	Islamic religious school	280	35.7
	Other	505	64.3
Home province	5 Southern border provinces	468	59.6
	Other	317	40.4

Table 3.2: Frequency distributions of the categorical determinants

Figure 3.1 shows the distribution of the time taken by the students to complete the English Vocabulary Skill Test, measured in minutes. The time varied from 1 to 30 minutes, the average time that students taken to complete the test was nearly 6 minutes with standard deviation 3.3. However, most students spent 1 to 10 minutes to complete the test, 31.3% finished within 5 minutes, and a further 17.8% within 10 minutes. Only 4% took more than 10 minutes to finish.

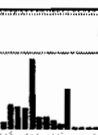
Variable name	Size	Graph	Minimum	Maximum	Mean	StDev	Skewness	Kurtosis
timeUsed	785		1	30.000	5.930	3.339	1.561	5.445

Figure 3.1: Distribution of the time taken to complete the test

3.3 Frequency distribution of the outcome

Figure 3.2 shows the distribution of the outcome variable (the English Vocabulary Skill Test scores).

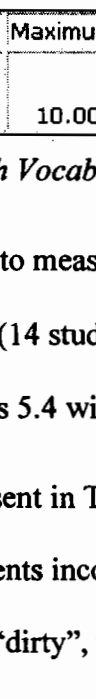
Variable name	Size	Graph	Minimum	Maximum	Mean	StDev	Skewness	Kurtosis
Total	785		0	10.000	5.439	2.492	-0.075	-0.739

Figure 3.2: Distribution of the English Vocabulary Skill Test scores

The English Vocabulary Skill Test scores used to measure the English Vocabulary Skill of undergraduate students, ranged from 0 (14 students) to 10 (39 students). The distribution was normal and the mean score was 5.4 with standard deviation 2.5.

The average scores on the individual items present in Table 3.3. It was ranged from 0.21 for the word “useful”, which most of students incorrectly matched to the opposite word, “useless”, to 0.85 for the word “dirty”, which students correctly

matched with the synonym word, “unclean”. The other eight items average scores were from 0.43 to 0.67.

Item	Mean	Standard Deviation
1. dirty	0.85	0.35
2. useful	0.21	0.41
3. disease	0.60	0.50
4. push	0.43	0.50
5. research	0.50	0.50
6. damage	0.44	0.50
7. propose	0.50	0.50
8. forecast	0.59	0.50
9. hazard	0.65	0.48
10. important	0.67	0.47

Table 3.3: Individual item mean scores of English Vocabulary Skill Test

3.4 Association between outcome and categorical determinants

In this section we investigated the association between outcome variable and categorical determinants.

Figure 3.3 shows the association between English Vocabulary Skill Test score and gender. The two-sample t-test did not give a significant result, and there was no evidence of a difference between male and female students in their scores on the English Vocabulary Skill Test.

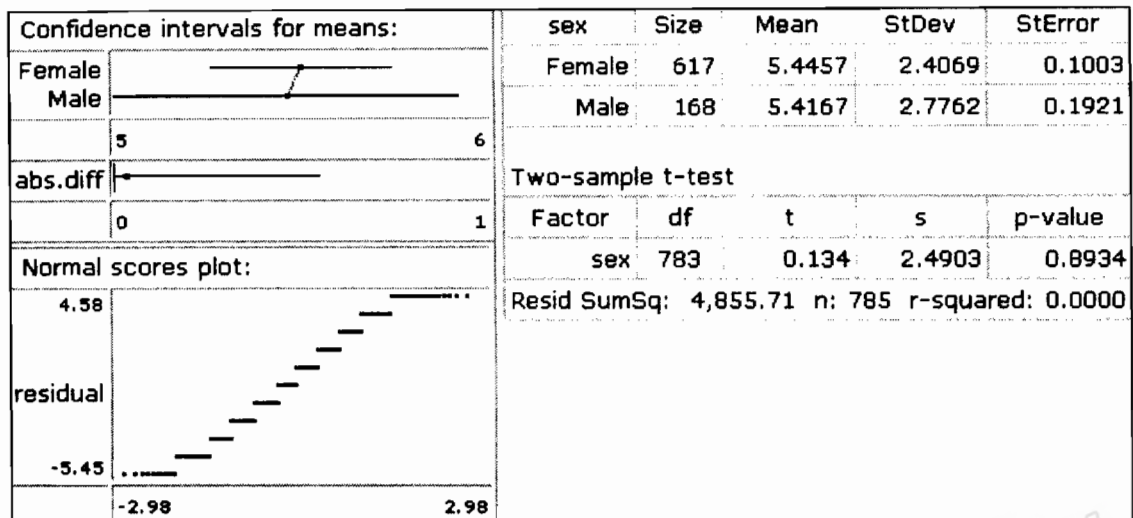


Figure 3.3: Association between English Vocabulary Skill Test score and gender

Figure 3.4 shows the association between English Vocabulary Skill Test score and religion. There was no evidence of a difference between Islamic and non-Islamic students in their scores on the English Vocabulary Skill test.

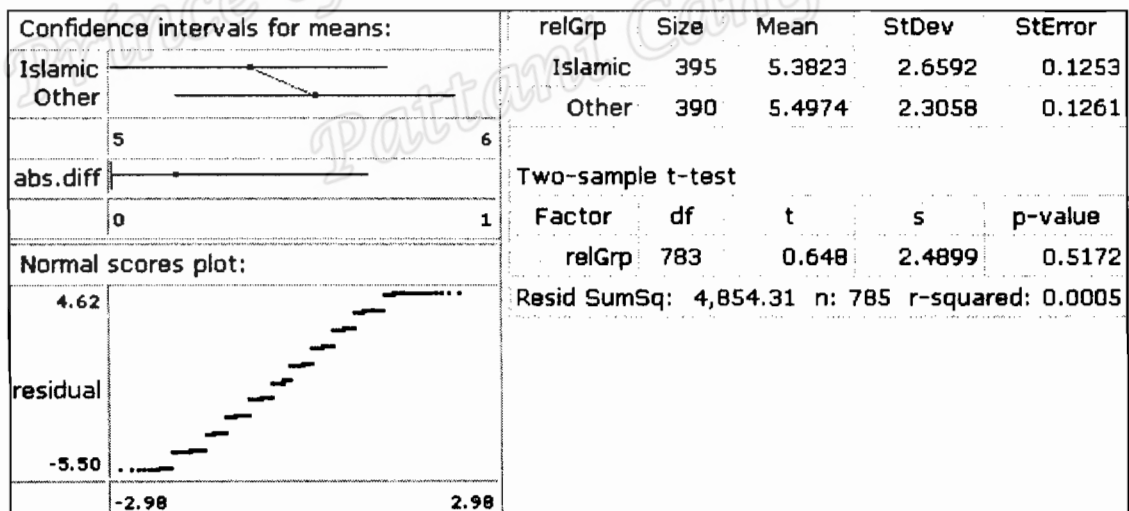


Figure 3.4: Association between English Vocabulary Skill Test score and religion

Figure 3.5 shows the association between English Vocabulary Skill Test score and the degree duration. The two-sample t-test gave a statistically significant association between English Vocabulary Skill Test score and the degree duration (p -value = 0.01).

It showed that the four-year-degree students did better than the five-year-degree students. They averaged 5.5 on the test, compared with an average score of 4.7 for the students doing five-year degrees.

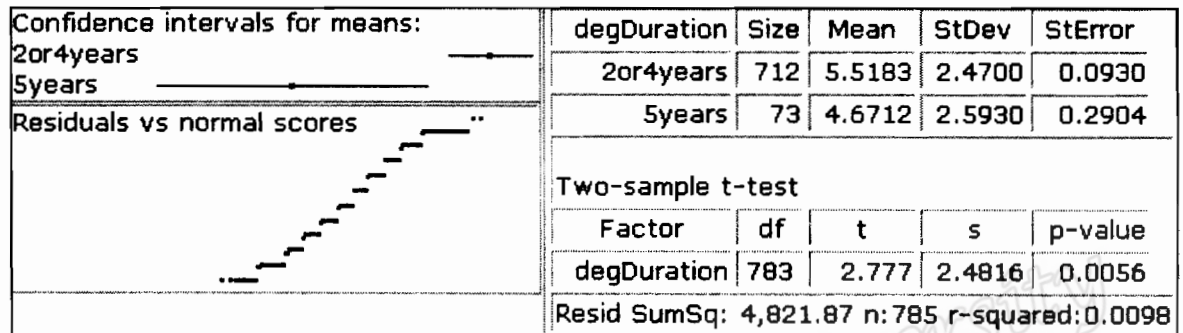


Figure 3.5: Association between English Vocabulary Skill Test score and degree duration

Figure 3.6 shows the association between English Vocabulary Skill Test score and the seniority level. It showed a highly statistically significant association between English Vocabulary Skill Test score and the seniority level (p -value < 0.0001). It presented that the second-year students scored less than the other students. They averaged 4.6 on the test, compared with average scores of 5.5 for the forth-year students, 5.9 for the first-year students, and 6.2 for the third-year students.

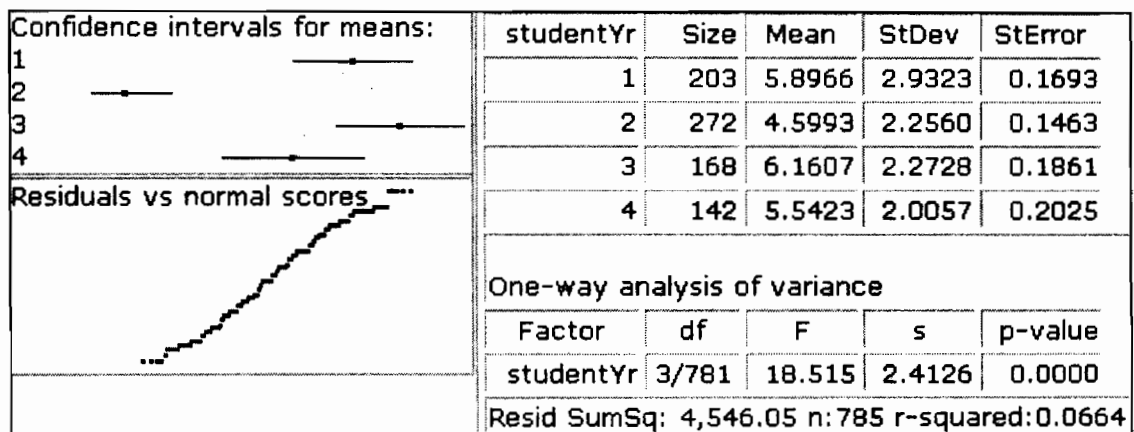


Figure 3.6: Association between English Vocabulary Skill Test score and seniority level

Figure 3.7 shows the association between English Vocabulary Skill Test score and the faculty of study. This association was statistically highly significant (p -value < 0.0001). The students enrolled in the College of Islamic Studies and in the Faculty of Communication Science did worse than other students (with average scores of 4.1), whereas those enrolled in the Faculty of Humanities and Social Sciences did better (with average scores 6.4).

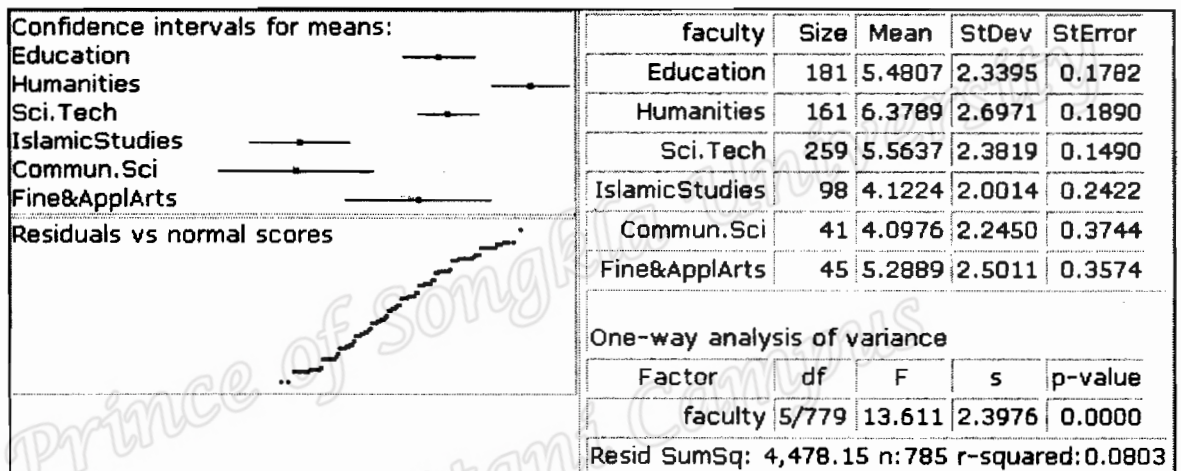


Figure 3.7: Association between English Vocabulary Skill Test score and faculty of study

Figure 3.8 shows the association between English Vocabulary Skill Test score and the major fields of study. This association was statistically highly significant (p -value < 0.0001). The students majoring in languages got higher scores than other students (with average scores 6.3), while those majoring in communication science got the lower scores (with average scores 3.8). The students majored in Islamic studies (with average scores 4.1) and social sciences (with average scores 4.6) got lower scores than the mean scores (5.4).

For the remainder got the same scores as the mean scores.

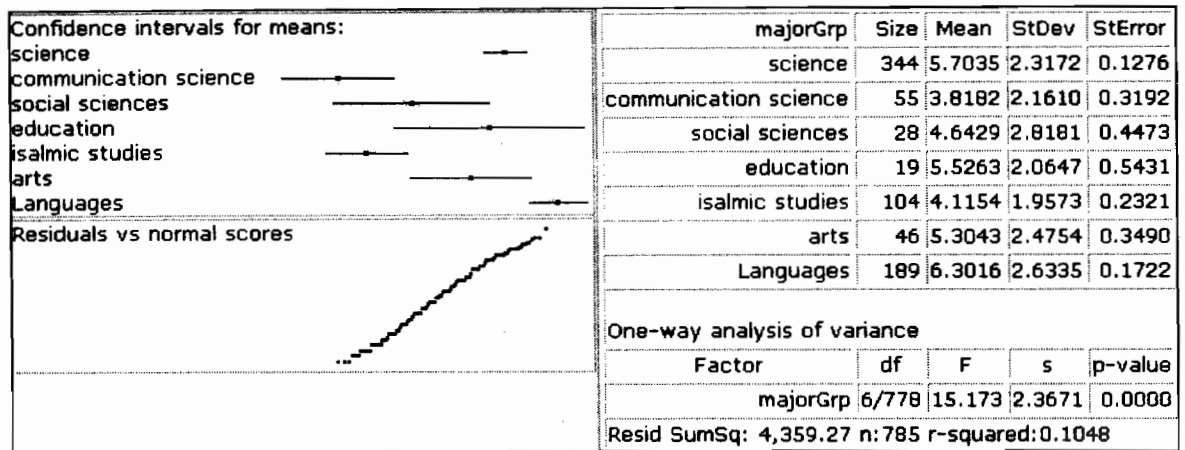


Figure 3.8: Association between English Vocabulary Skill Test score and major field of study

As the different faculties have similar majors. For example, students can get a science major both in the Faculty of Education and in the Faculty of Science and Technology.

So we created a new variable for “major field of study” containing nine categories.

Figure 3.9 shows the association between English Vocabulary Skill Test score and major field of study separated by faculties. This association was statistically highly significant (p -value < 0.0001). It presented that the 113 students majoring in any language in the Faculty of Humanities and Social Sciences averaged better than all other students on the English Vocabulary Skill Test. They averaged 7.5 on the test, compared with average scores ranging from 3.8 (for communication sciences majors) to 6.1 (for science majors in Education) for others (p -value < 0.0001).

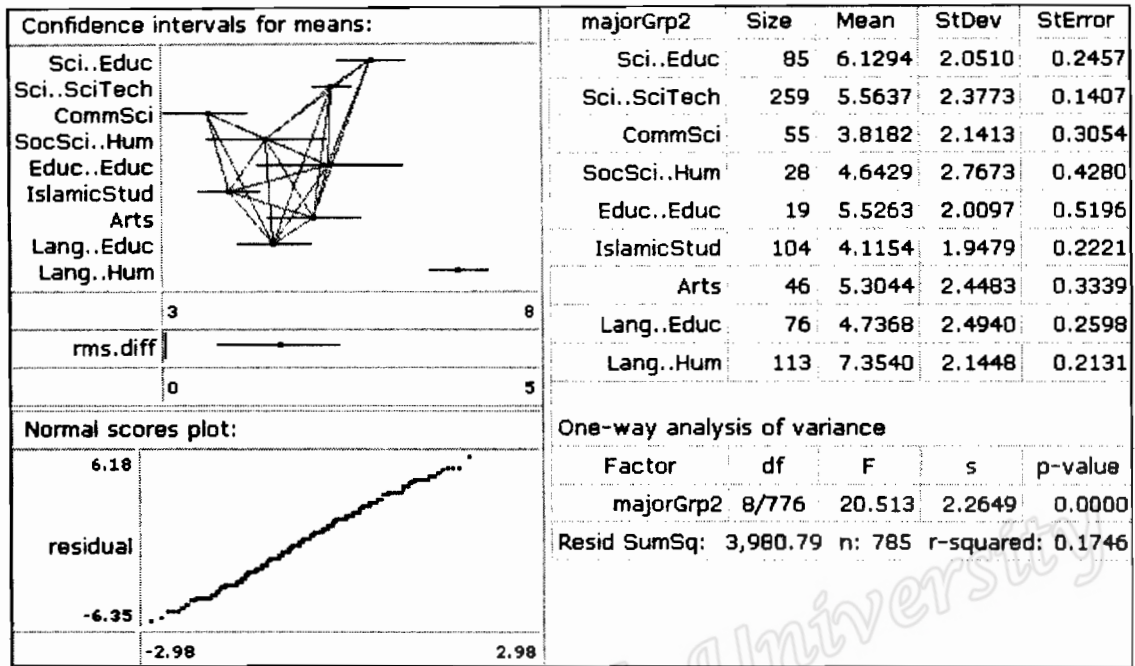


Figure 3.9: Association between English Vocabulary Skill Test score and major field of study separated by faculties

However, there were some students in the group of 113 who performed poorly (those majoring in Arabic and Chinese), and the 69 students majoring in English did best of all (averaging 8.2) as shown in the Figure 3.10.

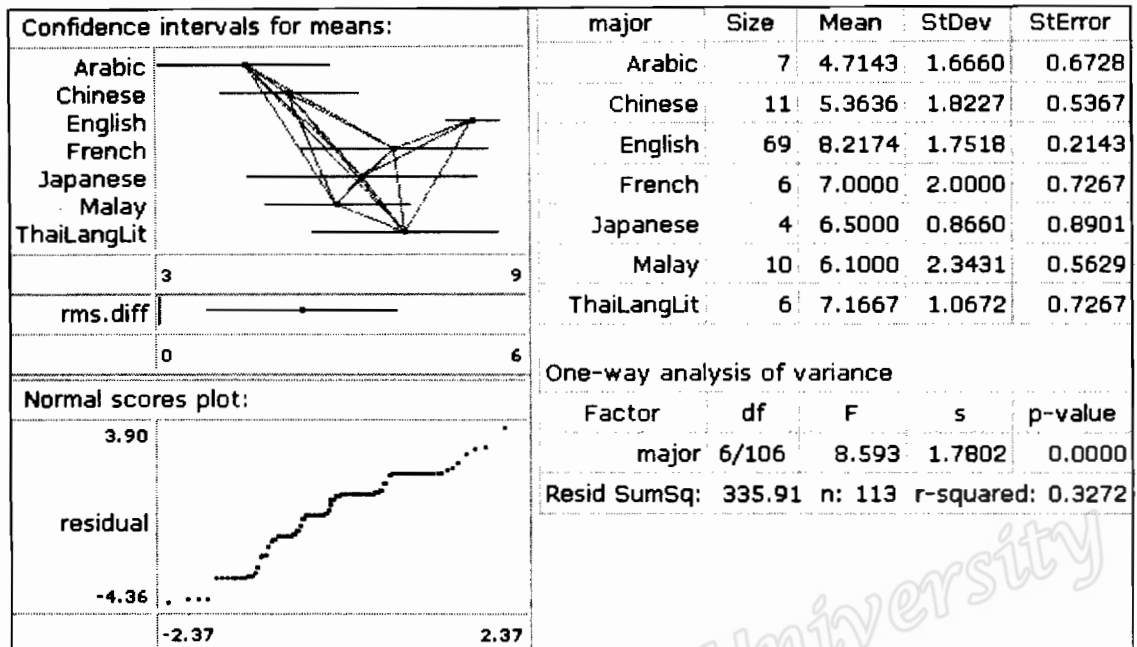


Figure 3.10: Association between English Vocabulary Skill Test score and languages major in the Faculty of Humanities and Social Sciences

For the students majoring in English, the 69 registered in the Faculty of Humanities and Social Sciences did better than the 35 in the Faculty of Education, who averaged only 5.6 (compared to 8.2) (p -value < 0.0001), as shown in the Figure 3.11.

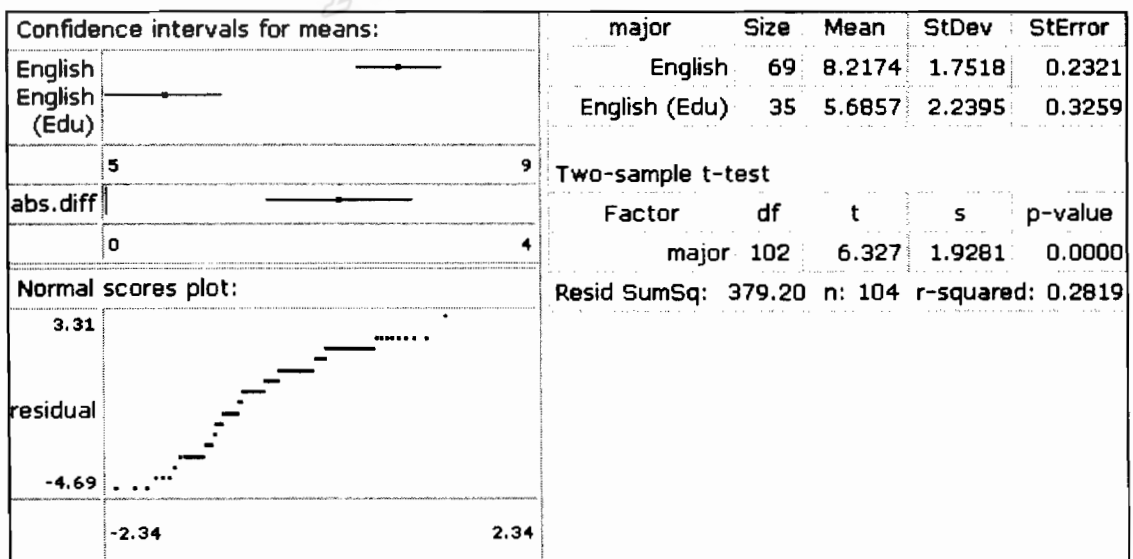


Figure 3.11: Association between English Vocabulary Skill Test score and English major in the Faculty of Humanities and Social Sciences and Faculty of Education

Figure 3.12 shows the association between English Vocabulary Skill Test score and the high school type. The two-sample t-test did not give a significant result, so there was no evidence of an association between the students who finished university preparation from the Islamic religious high school and other high schools.

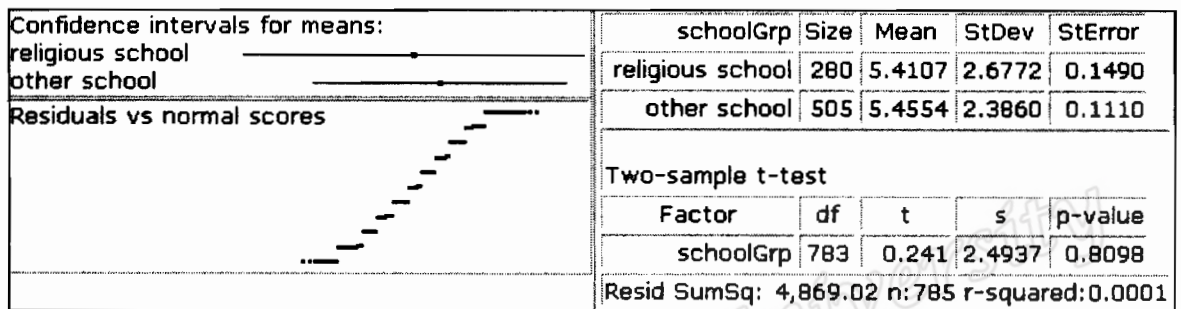


Figure 3.12: Association between English Vocabulary Skill Test score and high school type

Figure 3.13 shows the association between English Vocabulary Skill Test score and home province. The two-sample t-test did not give a significant result, so there was no evidence of an association between the students who lived in five Southern border provinces and other provinces.

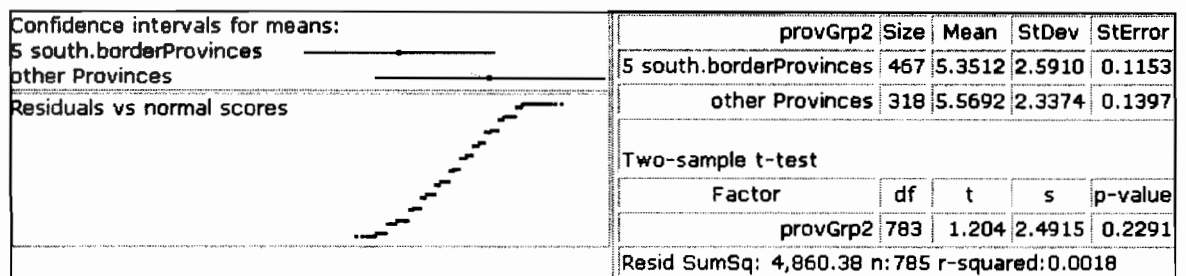


Figure 3.13: Association between English Vocabulary Skill Test score and home province

3.5 Association between outcome and continuous determinant

It remained to report the association between English Vocabulary Skill Test score and the time taken to complete the test.

As the outcome is continuous, the appropriate method to assessing this association is linear regression. Figure 3.13 shows the result. The score achieved in English Vocabulary Skill Test was found to be negatively correlated with the time taken to complete the test and the correlation coefficient was -0.072.

Linear Regression Analysis: Outcome = Total				
Determinant	Coefficient	StError	95% CI	p-value
Constant	5.7595	0.1810	5.4047, 6.1143	0.0000
timeUsed	-0.0540	0.0266	-0.1061, -0.0018	0.0428

r-sq: 0.0052 df: 783 RSS: 4,843.9134 s: 2.4872

Figure 3.14: Result from fitting the linear regression for the time taken to complete the test