

## CHAPTER 3

### PRELIMINARY DATA ANALYSIS

The preliminary data analyses are presented in two parts. The first part consists of descriptive summaries of all the variables measured in this study. The second part consists of a description of the associations between the outcome variable and each determinant of interest.

#### 1. Distributions of Variables

Histogram are used to describe data distributions. Figure 1 shows the size, mean, standard deviation, minimum and maximum of each variable together with frequency plots for the coded categorical variables (family status, family income, father's and mother's occupation, father's and mother's education, entrance score, residence, family status, family income, entrance score, residence, faculty and basic education). The means and standard deviations are not particularly meaningful for the nominal variables with three or more categories. The determinants of interest are the categorical variables, and the number of cases is 627. The outcome variable (university achievement) is continuous. The mean of university achievement is 2.725 and standard deviation is 0.436, and its minimum and maximum are 1.52 and 3.86, respectively. The coding for the data is as follow:

gender	1 = male, 2 = female
age group	0 = 16-17 year, 1 = 18 year, 2 = 19 year, 3 = 20+ (20 year or more than 20 year)
religion	1 = not muslim (Buddhism or Christian), 2 = muslim (muslim students are of interest in this study)
family status	1 = couple , 2 = single parent (father or mother deceased), 3 = separated (due to work commitments or other reasons)

family income	1 = <5,000 baht/month, 2 = 5,000-15,000 baht/month, 3 = >15,000 baht/month (more than 15,000 baht/month)
father educ	1 = no degree (Primary, secondary/high school, tertiary, college and certificate from university (1-2 years), 2 = degree (Bachelor degree, master degree or PhD.)
father occup	1 = govt (government official), 2 = priv emp (State Enterprise or company employee), 3 = priv owner (self employed or agriculture), 4 = others
mother educ	1 = no degree (primary, secondary/high school, tertiary, college and certificate from university (1-2 years), 2 = degree (Bachelor degree, master degree or PhD.)
mother occup	1 = govt (government official), 2 = priv emp (State Enterprise or company employee), 3 = priv owner (self employed or agriculture), 4 = others
entrance mode	1 = pooled, 2 = direct
entrance score	1 = <200 marks, 2 = 200-300 marks, 3 = >300 marks (more than 300 marks)
school	0 = formal, 1 = nonformal
residence	0 = local (in the South except the province in near local), 1 = near local (Pattani, Yala, Narathiwat, Songkhla, Satun, and Phatthalung), 2 = others (the North, Bangkok, near Bangkok, the East, and the West.)
school GPA	1 = <2, 2 = 2-3, 3 = >3 (more than 3)
faculty	15 = Faculty of Education, 30 = Faculty of Humanities and Social Sciences, 45 = Faculty of Science and Technology, 65 = College of Islamic Studies
basic educ	1 = science, 2 = language (Liberal Arts/Language), 3 = lang+maths (Liberal Arts/Mathematics), 4 = others

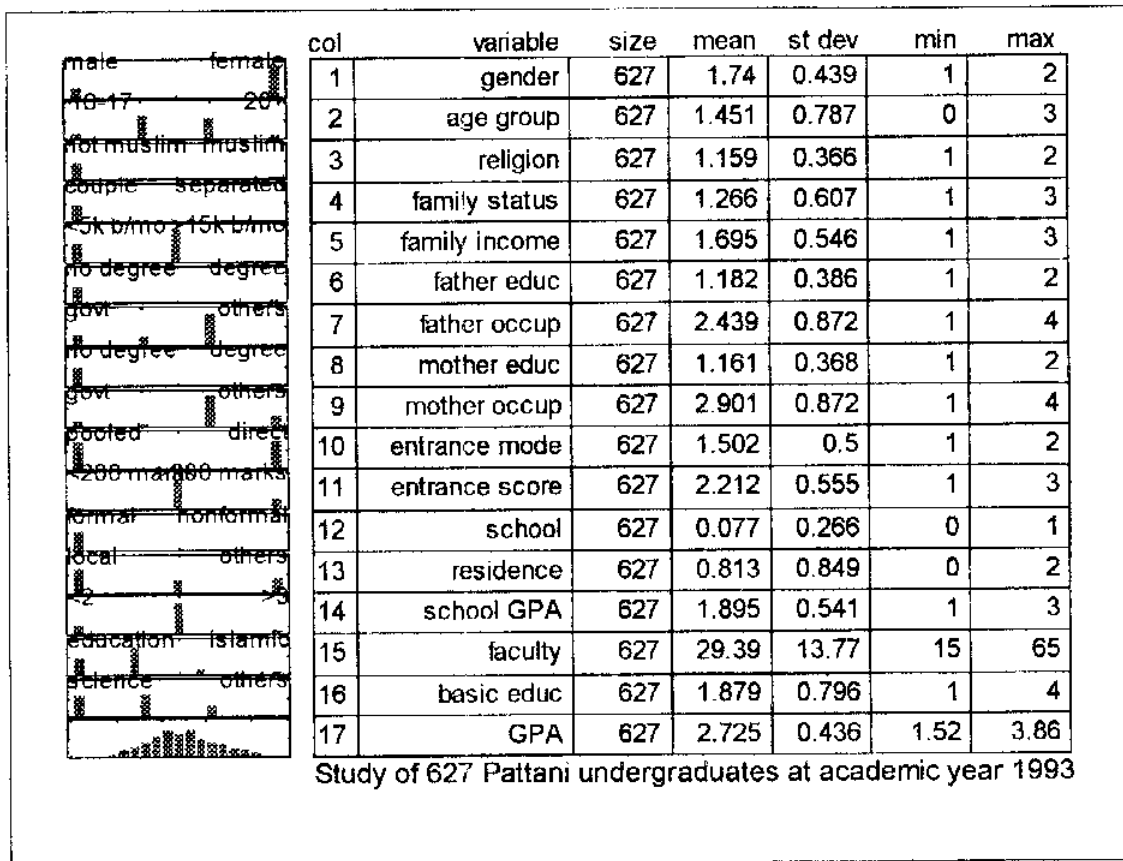


Figure 1: Histograms and numerical summaries means of data used

As shown in Figure 2, the histogram of university achievement is approximately symmetric, and these data do not require transformation. Note that, the distribution of the base 2 logarithms of university achievement is slightly skewed to the left.

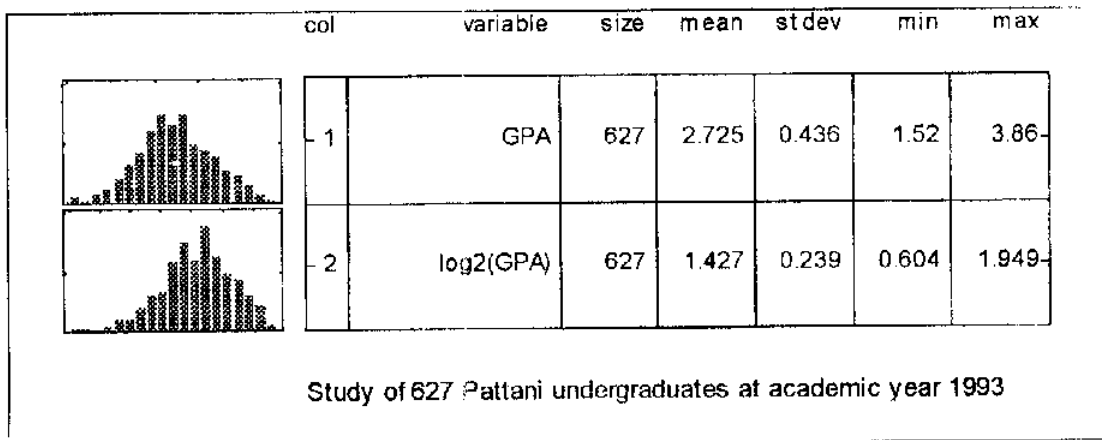


Figure 2: Histograms of outcome variable before and after log transformation

Table 4 shows the distribution of students by each determinant. A total of 464 students (74%) were female, and 163 students were male. Eighty-two percent of students were aged 18 or 19 years. The religious composition of the students comprised 84% Buddhists and Christians, and 16% Muslims. Only two students were Christian, so they are included with the larger group. Eighty-two percent of the students resided with both parents, 10% with a single parent (mother or father died), and 9% lived with a separated parent. Almost sixty-one percent had family incomes between 5,000-15,000 baht/month, and 35% were less than 5,000 baht/month. The proportion of fathers and mothers with no degree were 82% and 84% respectively. Fifty seven percent of fathers and 58% of mothers were private owners. The proportions pooled and direct entrance examination were 49.8% and 50.2%, respectively. For university entrance examination scores, 65% scored between 200 to 300 marks, 28% scored more than 300 marks and only 7% had less than 200 marks. Almost 92% finished formal school compared with 8% nonformal school. The number of students from the fourteen southern provinces was 449 of 627 students (71.6%). 80% had a school GPA greater than 2. The distribution of students by faculty was 50% Humanities and Social Sciences, 32% Education, 10% Science and Technology, and 8% College of Islamic Studies. The proportions studying science and language were 37% and 40%, respectively.

Table 4: Distribution of students by determinant variables

Determinants	Category	Count	Percentage	Mean	StDev
Gender	Male	163	26.0	2.61	0.467
	Female	464	74.0	2.77	0.417
Age group	16-17	65	10.4	2.72	0.431
	18	265	42.3	2.73	0.412
	19	246	39.2	2.75	0.456
	20*	51	8.1	2.56	0.444
Religion	Not Muslim	527	84.1	2.72	0.444
	Muslim	100	15.9	2.74	0.391
Family status	Couple	514	82.0	2.75	0.429
	Single parent	59	9.4	2.64	0.494
	Separated	54	8.6	2.58	0.402
Family income	<5,000 baht/month	218	34.8	2.70	0.453
	5,000-15,000 baht/month	382	60.9	2.74	0.429
	>15,000 baht/month	27	4.3	2.61	0.38
Father's education	No degree	513	81.8	2.73	0.433
	Degree	114	18.2	2.70	0.453
Father's occupation	Government	135	21.5	2.68	0.455
	Private employee	108	17.2	2.69	0.454
	Private owner	358	57.1	2.75	0.425
	Others	26	4.1	2.69	0.413
Mother's education	No degree	526	83.9	2.73	0.430
	Degree	101	16.1	2.68	0.465
Mother's occupation	Government	72	11.5	2.64	0.482
	Private employee	56	8.9	2.72	0.468
	Private owner	361	57.6	2.74	0.428
	Others	138	22.0	2.74	0.420
Entrance mode	Pooled	312	49.8	2.60	0.390
	Direct	315	50.2	2.84	0.447
Entrance score	<200 marks	44	7.0	2.53	0.378
	200-300 marks	406	64.9	2.70	0.441
	>300 marks	177	28.1	2.84	0.413
School	Formal	579	92.3	2.73	0.436
	Nonformal	48	7.7	2.63	0.432
Residence	Local	295	47.0	2.74	0.438
	Near local	154	24.6	2.80	0.475
	Others	178	28.4	2.64	0.385
School GPA	<2	128	20.4	2.54	0.436
	2-3	437	69.7	2.72	0.406
	>3	62	9.9	3.11	0.397
Faculty	Education	201	32.1	2.83	0.464
	Humanities and Social Science	316	50.4	2.67	0.408
	Science and Technology	61	9.7	2.67	0.471
	College of Islamic Studies	49	7.8	2.69	0.386
Basic education	Science	230	36.7	2.74	0.461
	Language	254	40.5	2.70	0.413
	Language and Mathematics	132	21.1	2.77	0.436
	Others	11	1.8	2.54	0.393

In this study the 627 subjects were classified into four age groups. The allocation to the four faculties and gender combinations is given in Table 5. The gender proportions were not the same within each age group. Age 18 years had the highest proportion.

*Table 5: Number of subjects classified by age group*

Age group	Education		Humanities& Soc.Science		Science & Technology		Islamic Studies		Total	
	male	female	male	female	male	female	male	female	male	female
16-17 years	4	14	2	37	2	4	1	1	9	56
18 years	18	52	31	125	7	19	6	7	62	203
19 years	20	79	21	83	15	10	8	10	64	182
20' years	6	8	7	10	3	1	12	4	28	23
Total	48	153	61	255	27	34	27	22	163	464

## 2. Associations between the Outcome and Each Determinant

A  $p$ -value is obtainable by using two-sample t-tests and one way analysis of variance (anova). Gender, religion, father's education, mother's education, entrance, and school are dichotomous determinants, so that two-sample t-test is the appropriate method for assessing the associations between these determinants and the outcome. The other determinants are multicategorical, so anova is used to assess their associations. The  $p$ -values are presented in Table 6.

11/1/2025  
11/1/2025  
11/1/2025

Table 6: P-values for the univariate associations

Determinants		p-value
<i>Stratification variables</i>		
Gender	t-statistic : 4.021	0.00006
Age group	F-statistic : 2.635	0.04897
Religion	t-statistic : 0.3313	0.74060
Residence	F-statistic : 5.43	0.00459
Family status	F-statistic : 4.871	0.00796
Family income	F-statistic : 1.56	0.21100
Father's education	t-statistic : 0.7044	0.48140
Father's occupation	F-statistic : 1.183	0.31530
Mother's education	t-statistic : 1.121	0.26270
Mother's occupation	F-statistic : 0.991	0.39650
<i>Predictor variables</i>		
Entrance mode	t-statistic : 7.196	0
Entrance score	F-statistic : 11.41	0.00001
School	t-statistic : 1.627	0.10420
School GPA	F-statistic : 39.95	0
Faculty	F-statistic : 5.591	0.00087
Basic education	F-statistic : 1.47	0.22160

The univariate data analysis indicates that gender, age group, family status, residence, and faculty are statistically significantly associated with achievement. School GPA, entrance mode, and university entrance examination score are very highly significant.

Box plots (Tukey, 1997) and 95% confidence intervals are used to show the relation between the significant predictor variables and mean achievement. Figure 3 shows box plots of achievement by *gender* (upper panel), and by *age group* (lower panel). The females had a mean achievement higher than the males. All age groups have similar spreads. The mean achievement is relatively constant with respect to age, except for those aged 20 years or more, where the mean achievement is markedly lower.

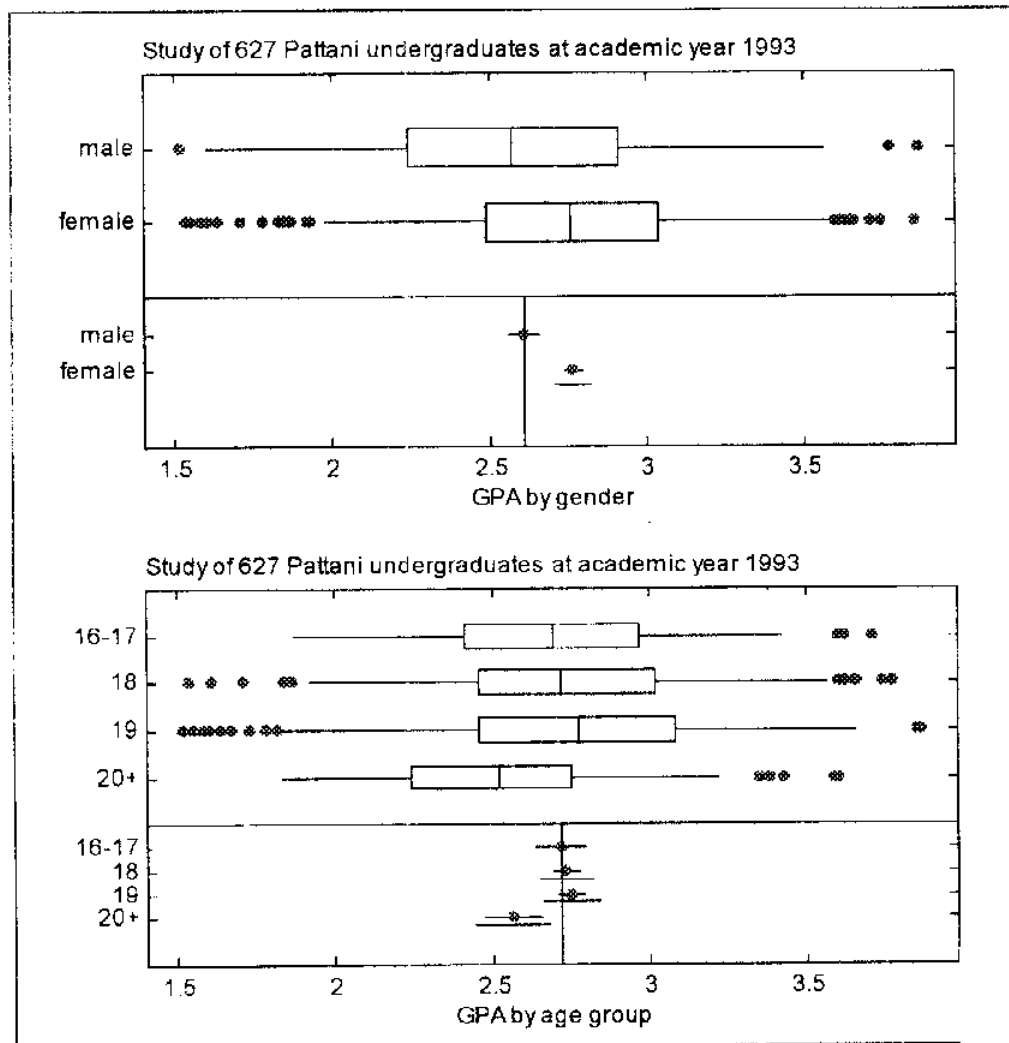


Figure 3: Box plots of achievement by gender, and age group



Figure 4 shows box plots and 95% confidence intervals of mean achievement by residence (upper panel), and by family status (lower panel). The students who came from the North, Bangkok, near Bangkok, the East, and the West had a mean achievement lower than those from the South. In the South, Pattani, Yala, Narathiwat, Songkhla, Satun, and Patthalung the mean achievement is slightly lower than in the other provinces. The multiple comparison of achievement stratified by family status may be summarised as follows:

couple > (single parent, separated)

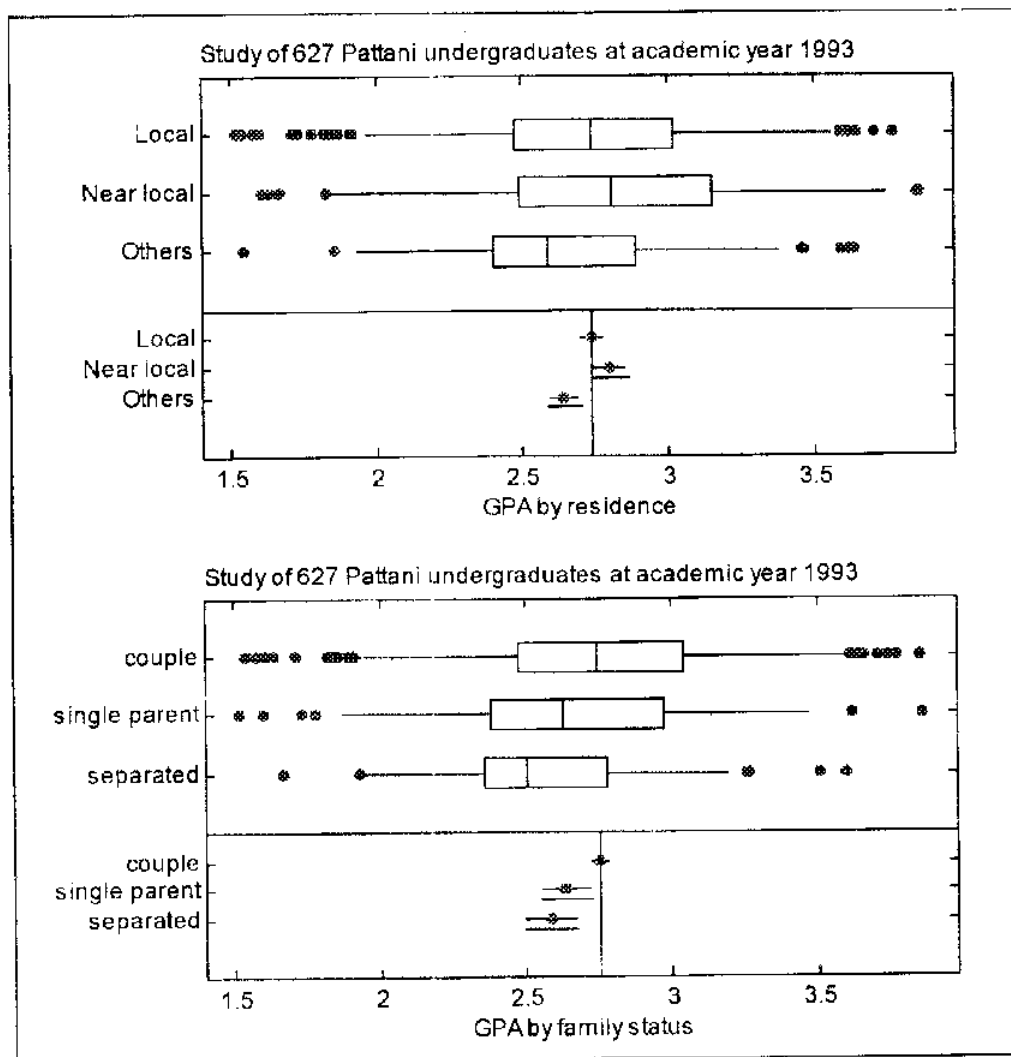


Figure 4: Box plots of achievement by residence, and family status

Figure 5 shows box plots and 95% confidence intervals of mean achievement by university enrollment. Entrance and university entrance examination scores are statistically significant. The students who passed the direct entrance examination had mean achievement scores substantially higher than those who did the pooled entrance examination. The achievement of the students clearly increases with university entrance examination score.

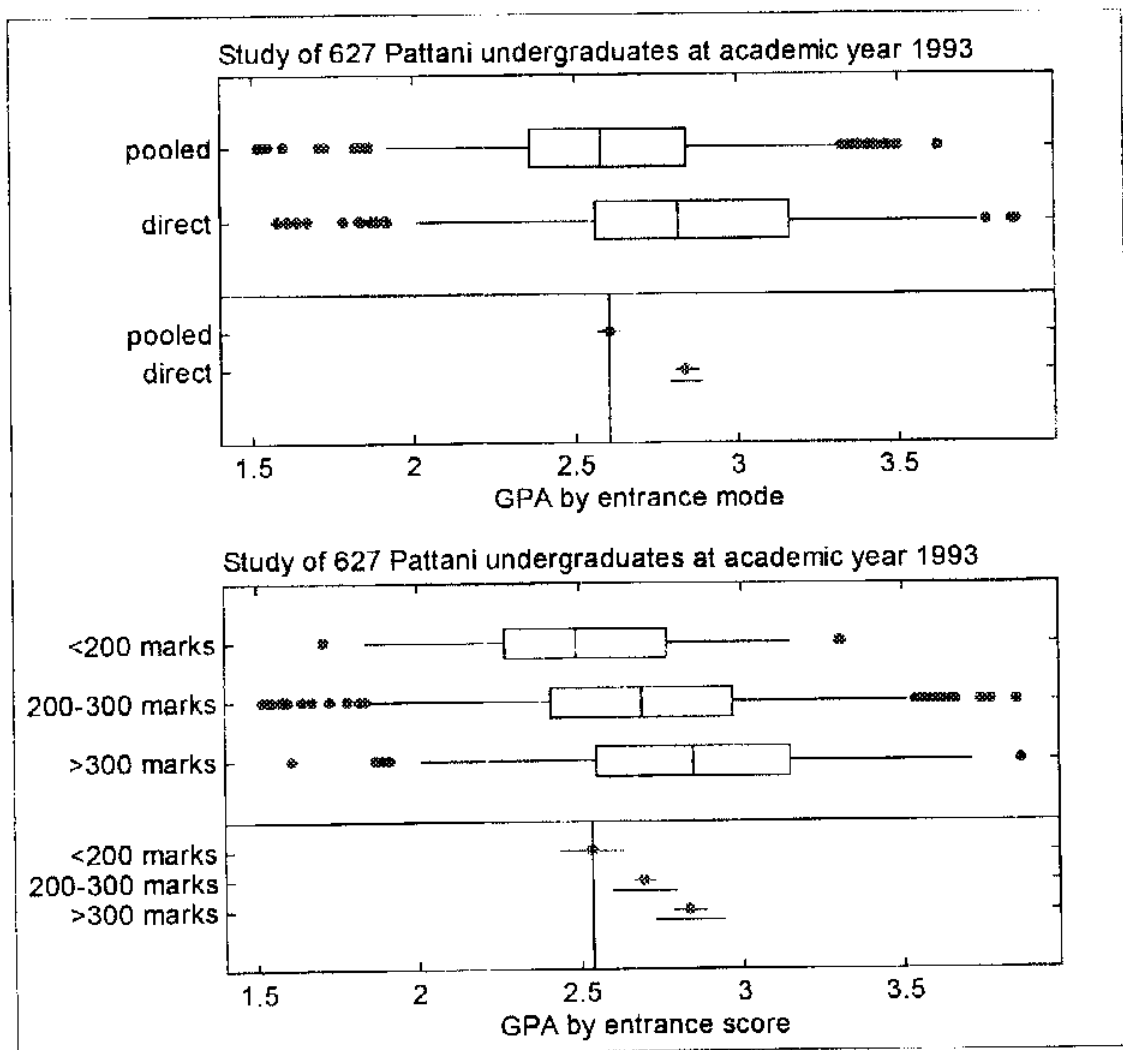


Figure 5: Box plots of achievement by university enrollment

Figure 6 shows the relation between mean achievement and the two other variables using box plots and 95% confidence intervals. The statistically significant variables are school grade point average, and faculty. The achievement of the students tends to increase with school GPA. Faculty of Education has a mean achievement higher than the other faculties. The multiple comparison is summarised as follows:

Education > (Islamic, Science and Technology, Humanities and Social Science)

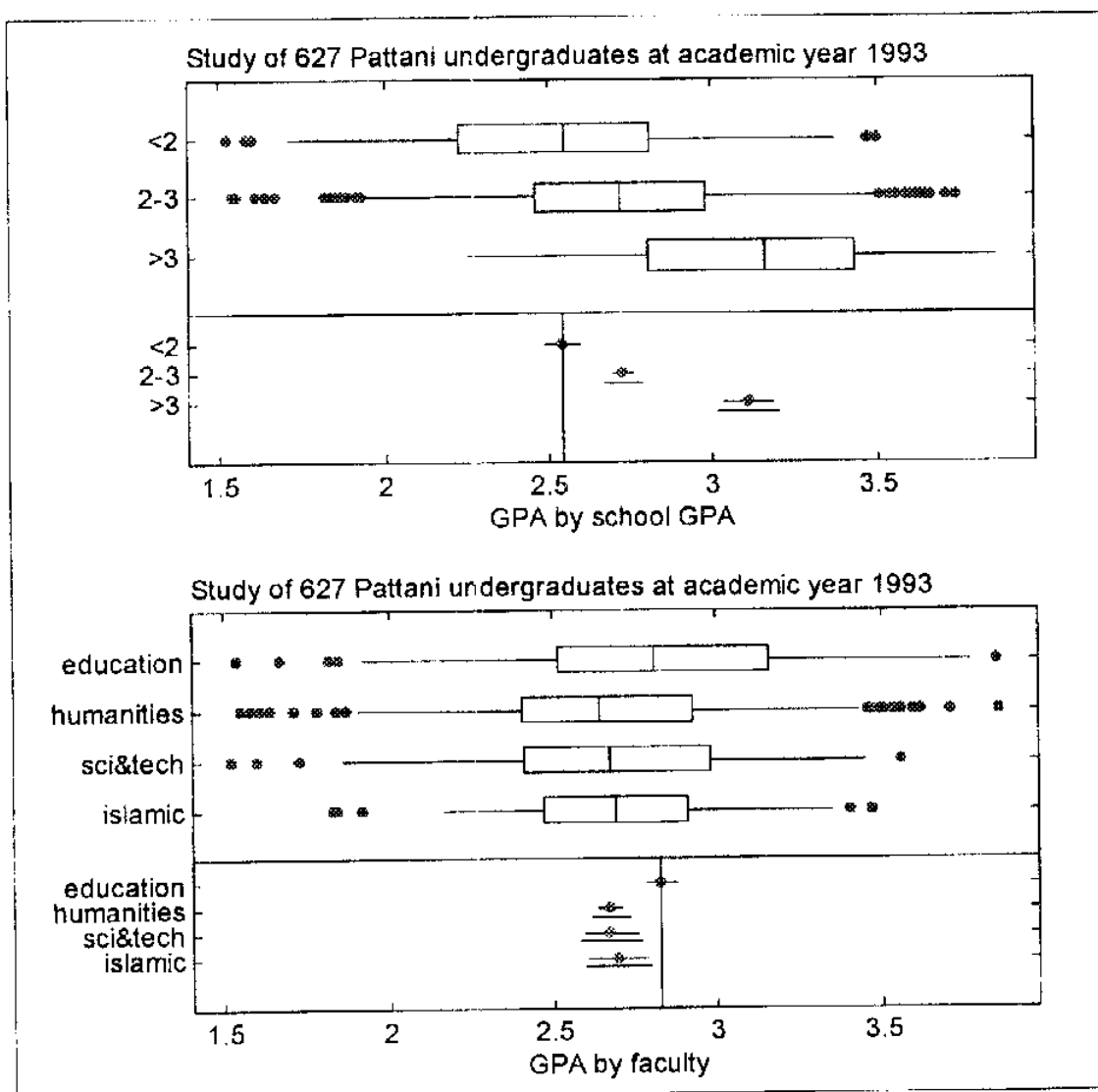


Figure 6: Box plots of achievement by educational background

Figure 7 shows the multivariate relation between mean achievement and school GPA, by age group and gender (x-label is school GPA, y-label is achievement, stratified in each age group). These plots show positive correlations between achievement and school GPA, with two exceptions as follows.

In age group 20+, there is no association for the males. In age group 16-17, there are no males with achievement of 3, so there is insufficient data.

In age group 18 and 19, the association for males is stronger than that for females. In other words, the effects of age and gender are not additive, and there is an interaction between gender and age.

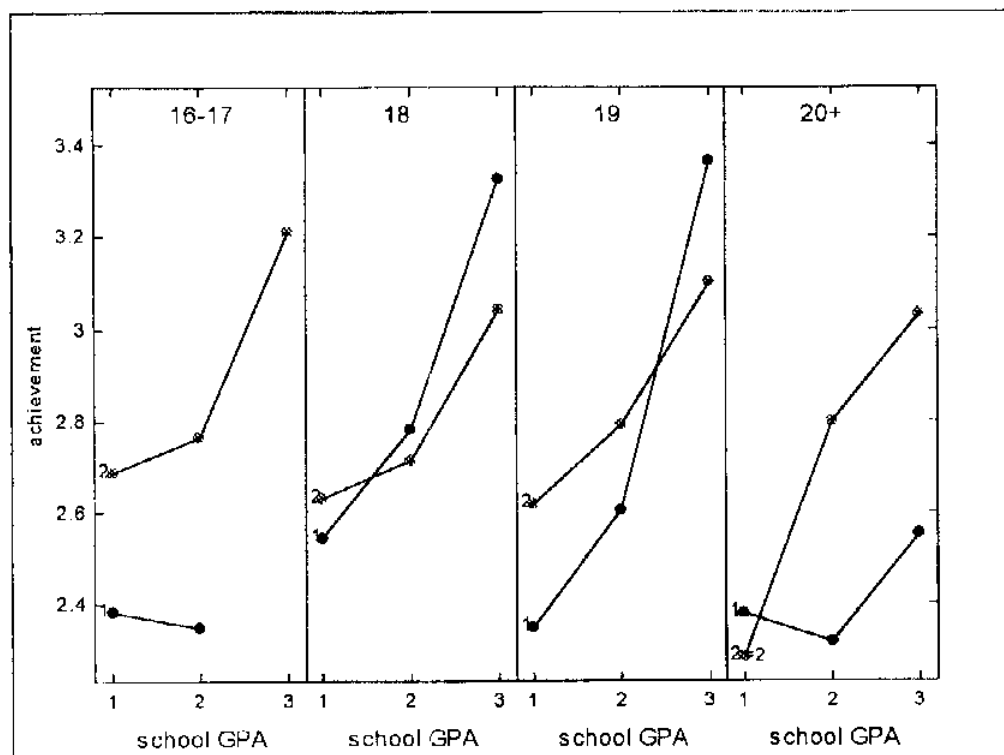


Figure 7: Achievement versus school GPA by age group & gender  
(male=1, female=2)

Figure 8 presents the multivariate relation between mean achievement and school GPA, by method of entrance and type of school (formal or nonformal school). X-label is school GPA, Y-label is achievement and stratified in each method of entrance. These plots show negative correlation between achievement and school GPA. There is insufficient data for nonformal students.

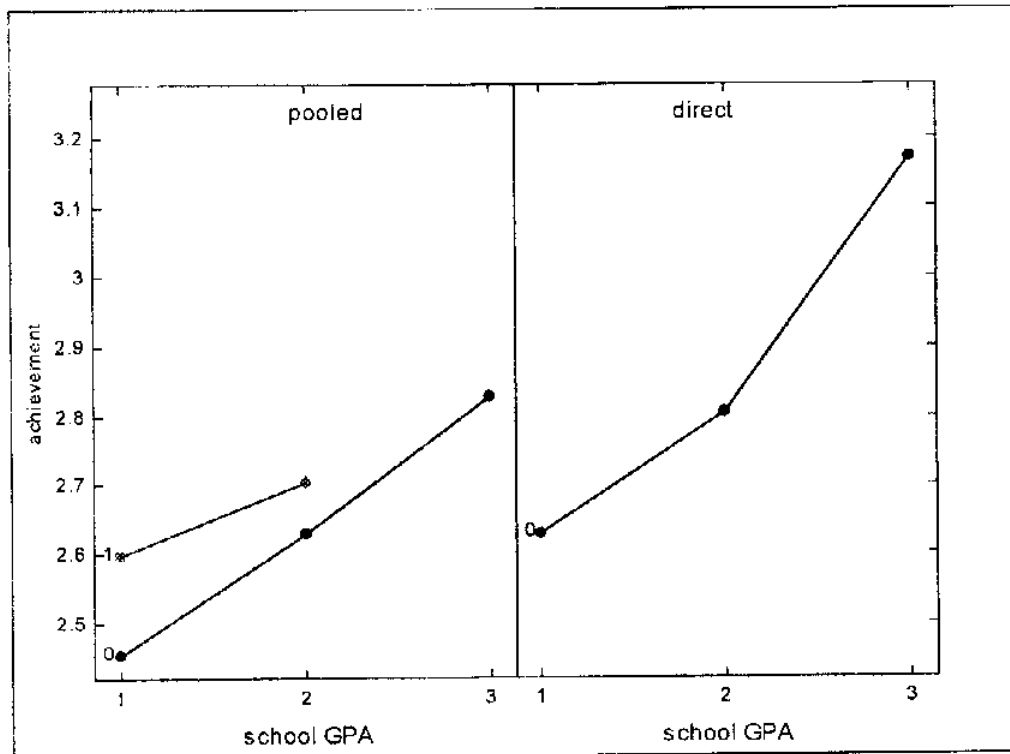


Figure 8: Achievement versus school GPA by method of entrance & type of school  
(formal=0, nonformal 1)