



**University of Sri Jayewardenepura**  
**Faculty of Humanities and Social Sciences**  
**Bachelor of Arts Fourth Year First Semester Examination**  
**August/September - 2018**  
**Economics**  
**ECON 4160 .03 – Applied Econometrics**

Time: Three hours (03)

**Answer any four (04) questions.** Calculators are allowed. Each question carries equal marks

1. I. What are the goals of data screening? [4 marks]  
 II. Why does a researcher need to look for bivariate screening? [4 marks]  
 III. Describe the uses of the following tools in data screening  
 a. Frequency table  
 b. Box plot  
 c. Histogram  
 d. Cross tabs

[4 marks]

IV. How do you resolve the following issues in a data set ?

- a. Errors in data entry  
 b. Outliers  
 c. Missing data

[4 marks]

2. I. What is the null hypothesis for an independent sample t test? How does it differ from one sample t test?

[4 marks]

II. In what situations should the paired samples t test be used rather than independent samples t test ?

[4 marks]

III. Briefly explain why there are two versions of t test results on SPSS output. How do you decide which one is more appropriate?

[6 marks]

IV. In what situations should the Mann-Whitney *U* Test ? Explain

[6 marks]

3. I. Interpret the SPSS outputs in 3A and 3B. Report the results for each set separately.

[12 marks]

**Output 3A,**

**Group Statistics**

	Religion group	N	Mean	Std. Deviation	Std. Error Mean
Female literacy	Muslim	25	48.72	25.203	5.041
	Catholic	32	77.09	20.039	3.542

Independent Samples Test

	Levene's Test for Equality of Variances	t-test for Equality of Means								
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Female literacy	Equal variances assumed	1.180	.282	-4.737	55	.000	-28.374	5.990	-40.377	-16.370
	Equal variances not assumed			-4.605	45.053	.000	-28.374	6.161	-40.782	-15.965

**Output 3B**

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre test	70.09	107	10.650	1.030
	Post test	78.34	107	22.883	2.212

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 Pre test & Post test	107	.865	.000

Paired Samples Test

	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Paired Differences				
				Lower	Upper			
Pair 1 Pre test - Post test	-8.243	14.671	1.418	-11.055	-5.431	-5.812	106	.000

- II. a. Describe the research situation related to the SPSS outputs in **3C**. [2 marks]
- b. Analyze the outputs in 3C and report the results. [6 marks]

Female literacy	Mean	Std. Deviation	N
Pre test	70.09	10.650	107
Post test	78.34	22.883	107

### Output 3C

#### Descriptives

Female literacy								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
					Catholic	32		
Muslim	25	48.72	25.203	5.041	38.32	59.12	14	100
Others	28	72.57	32.531	6.148	59.96	85.19	9	100
Total	85	67.26	28.607	3.103	61.09	73.43	9	100

#### ANOVA

Female literacy					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12477.690	2	6238.845	9.092	.000
Within Groups	56266.616	82	686.178		
Total	68744.306	84			

#### Multiple Comparisons

Dependent Variable: Female literacy

Tukey HSD

(I) Religion group	(J) Religion group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Catholic	Muslim	28.374*	6.992	.000	11.68	45.06
	Others	4.522	6.779	.783	-11.66	20.70
	Muslim	-28.374*	6.992	.000	-45.06	-11.68
Muslim	Others	-23.851*	7.208	.004	-41.06	-6.65
	Catholic	-4.522	6.779	.783	-20.70	11.66
Others	Muslim	23.851*	7.208	.004	6.65	41.06

\*. The mean difference is significant at the 0.05 level.

4. Answer the following questions by using the SPSS outputs in Output 4A

- I. Write the regression equation [2 marks]
- II. Assess the estimated model for goodness of fitness and overall significance [4 marks]
- III. Assess the individual predictors [4 marks]
- IV. Assess the validity of the model [4 marks]
- V. Report the results [6 marks]

**Output 4A**

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.799 <sup>a</sup>	.639	.626	23.4309	.639	47.785	3	81	.000	
2	.896 <sup>b</sup>	.802	.792	17.4500	.163	66.039	1	80	.000	2.062

a. Predictors: (Constant), People living in cities (%), Population increase (% per year)), Gross domestic product / capita

b. Predictors: (Constant), People living in cities (%), Population increase (% per year)), Gross domestic product / capita, Female literacy

c. Dependent Variable: Infant mortality (deaths per 1000 live births)

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	78703.008	3	26234.336	47.785	.000 <sup>b</sup>
	Residual	44469.505	81	549.006		
	Total	123172.513	84			
2	Regression	98812.227	4	24703.057	81.126	.000 <sup>c</sup>
	Residual	24360.287	80	304.504		
	Total	123172.513	84			

a. Dependent Variable: Infant mortality (deaths per 1000 live births)

b. Predictors: (Constant), People living in cities (%), Population increase (% per year)), Gross domestic product / capita

c. Predictors: (Constant), People living in cities (%), Population increase (% per year)), Gross domestic product / capita, Female literacy

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tol	VIF
1	(Constant)	82.64	8.728		9.46	.000		
	Gross domestic product / capita	-.001	.001	-.143	-1.66	.101	.601	1.663
	Population increase (% per year))	10.26	2.468	.299	4.15	.000	.864	1.157
	People living in cities (%)	-.89	.131	-.564	-6.80	.000	.649	1.540
2	(Constant)	138.85	9.491		14.62	.000		
	Gross domestic product / capita	-.002	.001	-.191	-2.97	.004	.596	1.677
	Population increase (% per year))	-1.86	2.367	-.054	-.78	.434	.521	1.919
	People living in cities (%)	-.34	.118	-.220	-2.94	.004	.442	2.263
	Female literacy	-.88	.109	-.661	-8.12	.000	.373	2.680

a. Dependent Variable: Infant mortality (deaths per 1000 live births)

5. I. Explain the different methods of multiple regression and their research situations. [6 marks]

II. A researcher is interested in examining the demand for health insurance in urban sector. Suggest him a suitable and possible econometric model with justifications. [6 marks]

III. **Output 5A** reveals the results related to internal migration decision. Interpret each output. [8 marks]

**Output 5A**

**Block 0: Beginning Block**

**Classification Table<sup>a,b</sup>**

	Observed	Predicted		
		Migrate		Percentage Correct
		.00	1.00	
Step 0	Migrate .00	0	42	.0
	1.00	0	42	100.0
	Overall Percentage			50.0

- a. Constant is included in the model.
- b. The cut value is .500

**Block 1: Method = Enter**

**Omnibus Tests of Model Coefficients**

	Chi-square	df	Sig.
Step	78.793	4	.000
Step 1 Block	78.793	4	.000
Model	78.793	4	.000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	37.656 <sup>a</sup>	.609	.811

- a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	1.932	8	.983

Classification Table<sup>a</sup>

	Observed	Predicted		
		Migrate		Percentage Correct
		.00	1.00	
Step 1	Migrate .00	37	5	88.1
	Migrate 1.00	3	39	92.9
	Overall Percentage			90.5

a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
income	-.463	.199	5.436	1	.020	.629
family_size	-.023	.016	1.946	1	.163	.978
Step 1 <sup>a</sup> education_level	2.231	1.108	4.052	1	.044	9.311
family_wealth	.118	.040	8.944	1	.003	1.125
Constant	-3.323	3.878	.735	1	.391	.036

\*\*\*\*\*

Step	Chi-square	df	Sig.
1	1.932	8	.983

Step	Chi-square	df	Sig.
1	1.932	8	.983