University of Kelaniya – Faculty of Social Sciences

Academic Year 2016/2017

Bachelor of Arts Honors Degree First Year

Second Semester Examination – (January 2018)

Social Statistics

SOST 22444: Computer Applications for Social Statistics

Answer Four (04) Questions Selecting two (02) from each Parts

No. of Questions: 08 **Time: 03 Hours**

Part I

01. i) Briefly explain the uses of Ms. Excel (06 Marks)

ii) Identify the difference between LEN, LEFT, RIGHT, and MID functions (08 Marks)

iii. Following table shows the description about six (6) tourists' hotels

	A	B	C	D	E
1	Name of the Hotel	Board Type	Room Charge (per Day)	Available Rooms	Place
2	Blue Ocean	Half	8,000	8	Matara
3	Serendib	Full	16,000	10	Kandy
4	River View	Full	18,000	7	Kandy
5	Holiday Resort	Half	9,000	2	Matara
6	Coral Sands	Full	15,000	5	Matara
7	Jungle Resort	Half	6,000	6	Matara

By using the above table write the suitable function to calculate the,

a) Average room charge (per day) of a hotel in Matara which has at least 5 available rooms

(02 Marks)

b) Average room charge (per day) of a hotel in Matara which has full board type and available rooms are greater than 4 (02 Marks)

c) How many hotels have more than 6 available rooms (02 Marks)

02. i) Describe COUNT, COUNTA and COUNTBLANK in Ms. Excel (06 Marks)

ii) Briefly explain the two types of operators in MS. Excel (06 Marks)

iii) Define the meaning of "Absolute Reference" in Spreadsheet Program (03 Marks)

iv) What is meant by a function? Recognize the categories of functions in Ms. Excel (05 Marks)

03. i) Identify the four objects in Ms. Access (08 Marks)

iii) Define "Records and Fields" in MS. Access with suitable examples (04 Marks)

iv) Why Ms. Access is mostly used to create a database than Ms. Excel? Discuss (08 Marks)

04. i) What is meant by a Database? (04 Marks)

ii) Identify drawbacks of a Database (05 Marks)

iii) Following tables show the details of the Employee database

Employee_Personal_Data

Emp_ID	Empy_Name	Emp_Address	National-ID	Emp_Phone
2522	Jagath Romen	Colombo	8818528211V	071-5826546
9825	Nisha Kumari	Colombo	8522567212V	077-5689735
0023	Sandun Pathirana	Kaluthara	9066894523V	077-5897653
8482	Ajantha Wijewardhane	Gampaha	8956821359V	078-8659721

Employee-Salary_Details

Salary_ID	Emp_ID	Date	Amount	Bonus
25-22A	2522	24/12/2018	64,000	15,000
98-26A	9825	24/12/2018	70,000	10,000
00-23B	0023	15/12/2018	42,000	8,000
84-82C	8482	10/12/2018	25,000	5,000

By using the above tables,

a) Identify the Primary Key, Foreign Key, Super Keys and give reasons for selecting that fields as primary, super and foreign keys. (09 Marks)

b) Define the terms of Composite Key and Candidate Key (02 Marks)

Part II

05. i) Discuss the properties of the variable view in SPSS (10 Marks)

ii) Following tables shows the frequency distribution of Region of a study on international tourists.

Region

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	America	4	4.0	4.2	4.2
	E.Asia & Pacific	23	22.8	24.0	28.1
	Europe	51	50.5	53.1	81.3
	South Asia	18	17.8	18.8	100.0
	Total	96	95.0	100.0	
Missing	System	5	5.0		
Total		101	100.0		

a) Identify the difference between Percent & Valid Percent (04 Marks)

b) What is meant by "Missing Value" (02 Marks)

iii) Recognize the difference between "Recode into same variable" and "Recode into difference variable" in SPSS (04 Marks)

06. i) Describe the strengthens and weaknesses of SPSS (06 Marks)

ii) A clinic provides a dieting program to help their patients to lose weight and asks a research agency to investigate the effectiveness of the dieting program. The number of inward patients are 50.

a) The agency wants to take 15 inward patients randomly for the investigation. Write down the relevant steps of selecting the sample in SPSS (02 Marks)

b) Identify the suitable test to check the effectiveness of the dieting program (02 Marks)

c) Following table shows the results of the test carried out by the research agency to identify the effectiveness of the dieting program, Interpret the results (04 Marks)

	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t value	d.f.	p value
Before Weight - After_Weight	1.400	1.265	.400	.495	2.305	3.500	9	.007

(Variables: weight of the patients before the dieting program and weight of the patients after the dieting program)

iii) Following table shows the part of questionnaire related to a study of "Tourists Satisfaction about Travel Destination Attributes".

1. Gender			**2. Marital Status**				
Male	1		Single	1			
Female	2		Married	2		**3. Age**	

The forth question of the questionnaire is designed to identify the satisfaction level of "Galle" Destination.

4. What is your satisfaction of Galle destination?

1= Worst, 2=Bad, 3= Neutral , 4= Good 5= Excellent

FACTORS	**5**	**4**	**3**	**2**	**1**
Landscape (beautiful scenery and natural attractions)	5	4	3	2	1
Culture, religious & history (history, monument, heritage, arts, handcraft and ways of life of Local people,)	5	4	3	2	1
Entertainment (such as in pubs, casino, etc.)	5	4	3	2	1
Shopping	5	4	3	2	1
Accommodation	5	4	3	2	1
Food	5	4	3	2	1

a) If you are going to enter the answers for this questions into SPSS, describe the variables that you need in create the variable view (02 Marks)

b) Identify the types of Measures in each identified variables (04 Marks)

07. i) Following table shows the details of gender of householder according to the living area and income level.

Income Level	Area of Living	Gender of Householder	
		Male	Female
Lower	Rural	749	35
	Urban	233	133
Lower Middle	Rural	625	38
	Urban	330	303
Upper Middle	Rural	420	37
	Urban	374	467
Higher	Rural	153	26
	Urban	266	25
Total		3150	1064

a) Fill the below table by assuming that you have to enter this data into the SPSS.

Name	Type	Label	Values	Measures

(08 Marks)

ii) Identify the rules that you should consider when you give a variable name in SPSS (04 Marks)

iii) The following SPSS output shows the simple liner regression model to identify the relationship between exam score and the number of study hours.

Independent variable – Number of Study Hours

Dependent variable – Exams Score

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.830[a]	.697	.663	11.71589

a. Predictors: (Constant), Hours

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1625.833	1	1625.833	11.845	.003[b]
	Residual	2470.717	18	137.262		
	Total	4096.550	19			

a. Dependent Variable: Exam_Score
b. Predictors: (Constant), Hours

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	26.367	10.441		2.525	.021
	Hours	1.629	.473	.630	3.442	.003

a. Dependent Variable: Exam_Score

a) Determine the correlation between exam Score and Number of Hours (02 Marks)

b) Write the regression equation and define the parameters of the regression model (03 Marks)

c) Test the overall and the parameters significance at 5% significance level (03 Marks)

08. i) Identify the importance of STATA and SPSS as statistical data analysis software (10 Marks)

ii) Recognize the difference between "Within Column Relative Frequencies, Within Row Relative Frequencies and Relative Frequencies" in STATA (06 Marks)

iii) Explain the use of "All possible two-way tables" in STATA (02 Marks)

iv) Define the One-way table and Multiple one way table in STATA with suitable examples (02 Marks)