



JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF BUSINESS AND ECONOMICS

UNIVERSITY EXAMINATIONS: 2013/2014

COURSE CODE SCS 324

COURSE TITLE: STATISTICAL ANALYSIS USING SPSS

MAIN CAMPUS

DATE: AUGUST 2013

TIME: 2 HOURS

INSTRUCTIONS:

- 1. This paper contains FIVE questions**
- 2. Answer Question one(Compulsory) and any other two questions**
- 3. Write all answers in the book let provided**

For Question one (COMPULSORY- 30 Marks) use the sample questionnaire to answer questions

SAMPLE QUESTIONNAIRE

1. Sex _____male _____female
2. Age in years _____
3. Education level (please indicate the highest level of schooling that you completed)
_____year 10 _____year 12
_____University or College_____Post Grad_____

4. Are you currently on a diet to loss weight _____yes ____no

Please indicate how much you either agree or disagree with each of the following statements

5. _____i have little control over things that happen to me
6. _____i can do just about anything i set my mind to do
7. _____there's really no way i can solve some of the problems i have
8. _____there's little i can do to change many of the important things in my life
9. _____what happens to me in the future mostly depends on me
10. _____I often feel helpless in dealing with problems of life
11. _____sometimes i feel that am being pushed around in life

Question 1

- a) Outline clearly the steps one would undertake to prepare a data entry screen
{5 marks}
- b) Prepare a codebook for the sample questionnaire provided, detailing each of the variable names and codes to be used to prepare the data for entry in SPSS
{5 marks}
- c) Using the codebook you developed, describe how you would create value labels question 3 (education level) . describe the steps you would use in SPSS {3 marks}
- d) Describe how you would developed a new variable for age group from the raw age figures captured into SPSS
{5 marks}
- e) Clearly differentiate the four levels of categorical data i.e nominal, ordinal, interval and ratio
{4 marks}
- f) Indicate at least four types of statistical inference that one can make for a frequency distribution table with a histogram that has the normal curve
{4 marks}
- g) Indicate in real life scenarios what SPSS can be used to accomplish with specific examples. E.g Getting mean score in a national exam to rank schools performance
{4 marks}

Question 2 20marks

a) Give a brief description of the following:

- i. Multivariate {2 marks}
- ii. Bivaraite {2 marks}
- iii. Crosstabulation {2 marks}
- iv. Correlation {2 marks}
- v. Regression analysis {2 marks}

b) Discuss how you would conduct the following in SPSS

- i. Establish the strength of a relationship between two variables {4 marks}
- ii. Compare means of a given different samples of data {4 marks}
- iii. Give an indication of the shape of the distribution {2 marks}

Question 3 20marks

Use the diagram below to answer the questions that follow

Favor or Oppose Death Penalty for Murder

Respondent's Sex			Frequency	Percent	Valid Percent	Cumulative Percent
Male	Valid	Favor	502	78.3	82.7	82.7
		Oppose	105	16.4	17.3	100.0
		Total	607	94.7	100.0	
	Missing	DK	34	5.3		
	Total		641	100.0		
Female	Valid	Favor	572	66.6	73.2	73.2
		Oppose	209	24.3	26.8	100.0
		Total	781	90.9	100.0	
	Missing	DK	72	8.4		
		NA	6	.7		
		Total	78	9.1		
Total		859	100.0			

a) Give an inference of the information presented above {6 marks}

b) Explain when or under what conditions should the following t-tests be used in SPSS

- i. One sample t-test {3 marks}
- ii. Paired sample t-test {3 marks}
- iii. Independent sample t-test {3 marks}

c) What do the following stand for in SPSS

- i. cf
- ii. P-value

{5 marks}

Question 4 20marks

- a) Interpret the table below

{10 marks}

Respondent's Sex * Rap Music * RS Highest Degree Crosstabulation

RS Highest Deg			Rap Music					Total
			Like Ver Much	Like It	Mixed Feelings	Dislike It	Dislike Very Mu	
Less than HS	Respondent' Sex	Male	Count 5	11	14	30	55	115
		% within Respondent's S	4.3%	9.6%	12.2%	26.1%	47.8%	100.0%
	Female	Count 10	18	19	35	59	141	
		% within Respondent's S	7.1%	12.8%	13.5%	24.8%	41.8%	100.0%
	Total	Count 15	29	33	65	114	256	
		% within Respondent's S	5.9%	11.3%	12.9%	25.4%	44.5%	100.0%
High school	Respondent' Sex	Male	Count 9	36	50	87	110	292
		% within Respondent's S	3.1%	12.3%	17.1%	29.8%	37.7%	100.0%
	Female	Count 11	45	95	134	175	460	
		% within Respondent's S	2.4%	9.8%	20.7%	29.1%	38.0%	100.0%
	Total	Count 20	81	145	221	285	752	
		% within Respondent's S	2.7%	10.8%	19.3%	29.4%	37.9%	100.0%
Junior college	Respondent' Sex	Male	Count 1	4	4	13	14	36
		% within Respondent's S	2.8%	11.1%	11.1%	36.1%	38.9%	100.0%
	Female	Count 1	3	13	15	18	50	
		% within Respondent's S	2.0%	6.0%	26.0%	30.0%	36.0%	100.0%
	Total	Count 2	7	17	28	32	86	
		% within Respondent's S	2.3%	8.1%	19.8%	32.6%	37.2%	100.0%
Bachelor	Respondent' Sex	Male	Count 2	8	22	32	41	105
		% within Respondent's S	1.9%	7.6%	21.0%	30.5%	39.0%	100.0%
	Female	Count 2	11	30	27	52	122	
		% within Respondent's S	1.6%	9.0%	24.6%	22.1%	42.6%	100.0%
	Total	Count 4	19	52	59	93	227	
		% within Respondent's S	1.8%	8.4%	22.9%	26.0%	41.0%	100.0%
Graduate	Respondent' Sex	Male	Count 3	7	19	38	67	109
		% within Respondent's S		4.5%	10.4%	28.4%	56.7%	100.0%
	Female	Count 5	12	9	16	42	79	
		% within Respondent's S		11.9%	28.6%	21.4%	38.1%	100.0%
	Total	Count 8	19	28	54	109	188	
		% within Respondent's S		7.3%	17.4%	25.7%	49.5%	100.0%

b) Distinguish between the terms dependent and independent variable {4 marks}

c) If a quantitative variable measured at the numerical scale has been recorded into three categories, which measures are appropriate to describe the distribution

{6 marks}

Question 5 20marks

- a) Identify two statistical measures of deviation from normality of a given distribution. In the identified measures, illustrate using diagrams the shape of the distributions indicating instances of positive or negative deviation and give values (p-value) that are statistically accepted in SPSS for such deviations from normality. {8 marks}
- b) The following are the Linear Regression statistics from a study to find out the factors impacting on KCSE results in Kisumu District. The independent variables were **Number of Teachers(x_1)**, **Number of desks (x_2)**, **Number of Toilets (x_3)** and **Number of Books (x_4)**.

Table 1: Linear Regression Statistics

	Unstandardized Coefficients		
	B	Std Error	Sig.
(Constant)	77.247	37.989	0.020
Number of teachers	22.027	0.059	0.005
Number of Desks	1.643	7.929	0.243
Number of Toilets	0.152	0.286	0.068
Number of Books	26.456	8.627	0.045

a Dependent Variable: KCSE results P = 0.05

- i. Write the Regression equation line for this model {4 marks}
- ii. Rank the four Independent Variables (Number of Teachers, Number of desks , Number of Toilets and Number of Books in terms of the variable that best predicts KCSE results in Kisumu District. {4 marks}
- iii. Which of the independent variables are significantly related to KCSE results {4 marks}