JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL INFORMATICS AND INNOVATIVE SYSTEMS

UNIVERSITY EXAMINATION FOR THE DEGREE OF BACHER OF BUSINESS ADMNISTRATION

3RD YEAR 2ND SEMESTER 2013/2014 ACADEMIC YEAR

CENTRE: KISUMU

COURSE CODE: SCS 324

COURSE TITLE: STATISTICAL ANALYSIS WITH SPSS

EXAM VENUE: AH **STREAM: BBA**

DATE: 5/12/2013 **EXAM SESSION: 9.00 – 11.00 AM**

TIME: 1 ½ HOURS

Instructions:

1. **Answer question 1 (Compulsory) and ANY other 2 questions.**
2. **Candidates are advised not to write on the question paper.**
3. **Candidates must hand in their answer booklets to the invigilator while in the examination room.**

Question One

a. What's the importance of summary statistics using frequencies in SPSS. {5 marks}

. .

b. You manage a team that sells computer hardware to software development companies. At each company, your representatives have a primary contact. You have categorized these contacts by the department of the company in which they work (Development, Computer Services, Finance, Other, Don't Know). Use Frequencies to study the distribution of departments to see if it meshes with your goals.

Required

i. Explain the steps of coming up with a frequency distribution table in SPSS . {3 marks}

ii. How would you come up with a pie chart for the same information in a above using a similar procedure. {3 marks}

iii. On running the frequency distribution you got the following, interpret your findings. {4 marks}

Department

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Development	16	22.9	25.8	25.8
	Computer services	30	42.9	48.4	74.2
	Finance	13	18.6	21.0	95.2
	Other	3	4.3	4.8	100.0
	Total	62	88.6	100.0	
Missing	Don't know	8	11.4		
Total		70	100.0		

c. The shape of a distribution is very important in data analysis. Using illustrations explain what is meant by skewness and kurtosis giving their statistical significant values. {8 marks)

d. How would you go about converting a variable like age or income into a grouped variable like age group or income category in SPSS. {4 marks}

e. Write the multiple regression equation explaining what each letter stands for in the equation. {3 marks}

Question Two

a. In order to determine customer satisfaction rates, a retail company conducted surveys of 582 customers at 4 store locations. From the survey results, you found that the quality of customer service was the most important factor to a customer's overall satisfaction. Given this information, you want to test whether each of the store locations provides a similar and adequate level of customer service. You used the Crosstabs procedure to test the hypothesis that the levels of service satisfaction are constant across stores and obtained the following results. Interpret your findings. {5 marks}

Count

		Service satisfaction					
		Strongly Negative	Somewhat Negative	Neutral	Somewhat Positive	Strongly Positive	Total
Store	Store 1	25	20	38	30	33	146
	Store 2	26	30	34	27	19	136
	Store 3	15	20	41	33	29	138
	Store 4	27	35	44	22	34	162
Total		93	105	157	112	115	582

b. From the crosstabulation alone, you realized it was impossible to tell whether these differences were real or due to chance variation hence you used the chi-square test to ascertain your findings and obtained the following.

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	16.293[a]	12	.178
Likelihood Ratio	17.012	12	.149
Linear-by-Linear Association	.084	1	.772
N of Valid Cases	582		

Explain using the findings the relevance of the two-sided asymptotic significance of the
. chi-square statistic. {5 marks}

c. A public health researcher is studying smoking addiction in young people. He believes the data will show that heavier smokers began smoking at a younger age than lighter smokers

d. and is especially interested to know if the association is linear. Using the means procedure the following results were obtained.

i. Means of Age by Smoking Level

Age when first smoked a cigarette

# Cigarettes smoked per day past 30 days	Mean	N	Std. Deviation
1 to 5 cigarettes each day	15.81	1119	4.452
6 to 15 cigarettes (about 1/2 pack) each	15.89	1594	4.820
16 to 25 cigarettes (about 1 pack) each	15.63	1604	5.450
26 to 35 cigarettes (about 1 1/2 pk) eac	14.18	622	4.066
35 or more cigarettes (about 2 packs) ea	14.45	461	4.376
Total	15.48	5400	4.866

Interpret your findings above {5 marks}

ii. ANOVA and Tests of Linearity

			Sum of Squares	df	Mean Square	F	Sig.
Age when first smoked a cigarette * # Cigarettes smoked per day past 30 days	Between Groups	(Combined)	1974.095	4	493.524	21.158	.000
		Linearity	1321.500	1	1321.500	56.655	.000
		Deviation from Linearity	652.595	3	217.532	9.326	.000
	Within Groups		125841.1	5395	23.326		
	Total		127815.2	5399			

Interpret your findings above {5 marks}

Question Three

Use the sample questionnaire to answer questions BELOW

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*

Using the key 1 . Agree and 2. Disagree, Please indicate your opinion on 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*

a) Outline the steps one would undertake to prepare a data entry screen and prepare a codebook for the sample questionnaire provided above detailing each of the variable names and codes to be used to prepare the data for entry in SPSS. {8 marks}

b) Describe how you would developed a new variable for age group from the raw age figures captured into SPSS {3 marks}
c) Clearly differentiate the four levels of categorical data i.e nominal, ordinal, interval and ratio. {4 marks}
d) 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. {2 marks}
e) Why is a multiple response set important in SPSS for study questions with similar answers? Use above questionnaire to explain your answer. {3 marks}

Question Four

a) How would establish the following in SPSS for 2 variables. In each case clearly mention the statistical procedure and the relevant statistics

   i. Differences {4 marks}
   ii. Relationship {4 marks}
   iii. Association {4 marks}

b) With respect to normality discuss skewness and kurtosis illustrating your answer with sketches. {8 marks}

Question Five

Use the SPSS output for Linear Regression tables below to answer the following questions.

(a) Write down the linear regression equation indicatiing what each letter represents in the equation. {4 marks}

(b) What is the value of the standard error of the estimate? {3 marks}

(c) How many degrees of freedom are associated with the t-value for the line of regression? . {4 marks}

(d) What is the value of the correlation coefficient? {2 marks}

(e) Confidence and Prediction Interval {3 marks}

(f) What is the 95% confidence interval for the mean value of $\hat{Y}$ when x = ? {2 marks}

(g) What is the 95% prediction interval for $\hat{Y}$ when x = ? {2 marks}

Coefficients

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig
	B	Std Error	Beta		
Constant	2.129	.250	0.941	8.505	.000
Additive	.338	.050		6.821	.000

Model Summary

Model	R	R Square	Adjusted R Square	Std Error of the Estimate	Durbin-Watson
1	.941	.886	.867	.32121	2.321

ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig
1 Regression	4.801	1	4.801	46.532	.000
Residual	.619	6	.103		
Total	5.420	7			