UNIVERSITY EXAMINATIONS

2022/2023 ACADEMIC YEAR

THIRD YEAR SECOND SEMESTER

SPECIAL/ SUPPLEMENTARY EXAMINATION

FOR THE DEGREE OF BACHELOR OF SCIENCE

COURSE CODE: **STA 324/STA 348**

COURSE TITLE: STATISTICAL COMPUTING

DATE: 10/8/2023 **TIME:** 2:00 PM - 4:00 PM

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

This Paper Consists of 4 Printed Pages Please Turn Over

QUESTION ONE

(a) Enumerate 4 ways in which SPSS can transform data (4 marks)

(b) Income can be measured on several levels. Describe how income could be measured as an ordinal, interval and ratio measure. (6 marks)

(c) List with details ,three compulsory and five optional requirements in creating a data file (8 marks)

(d) A function Y is given by y=sin (x) and another function z is given by z= cos(x) Write m.scripts that can be executed in MATLAB to generate corresponding sine and cosine waves respectively (12 marks).

QUESTION TWO

A market researcher is interested in the coffee-drinking habits of males and females. He asks a sample of male and female office workers to record the number of cups of coffee they consume during a week.

a) Which parametric statistical technique could the researcher use to determine if males and females differ in terms of the number of cups of coffee consumed in a week? Justify your answer and describe how you would obtain this statistic using SPSS.

b) What are the key values you would look for in the output?

c) What assumptions should you check for when using the technique that you chose in question (a). above?

d) What non-parametric technique could be used to address this research question?

QUESTION THREE

a) Suppose you have the following Research Question

- To what extent does weight of a car in pounds predict miles per gallon in a U.S. dataset of 398 models of cars?

i) Are the two variables discrete or continuous?

ii) Are the two variables nominal, ordinal, interval or ratio scales?

iii) Which statistical procedure could we use to test the research question?

iv) What is the null hypothesis?

v) What is your expectation?

b) The regression SPSS output of the above question was as below. Use it to answer the following questions

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.807[a]	.651	.650	4.622

a. Predictors: (Constant), Vehicle Weight (lbs.)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15794.632	1	15794.632	739.503	.000[a]
	Residual	8457.943	396	21.358		
	Total	24252.575	397			

a. Predictors: (Constant), Vehicle Weight (lbs.)

b. Dependent Variable: Miles per Gallon

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	45.492	.841		54.110	.000
	Vehicle Weight (lbs.)	-.007	.000	-.807	-27.194	.000

a. Dependent Variable: Miles per Gallon

i) How much variance in Miles Per Gallon is explained by Car Weight?

ii) Is this variance explained significantly different to 0?

iii) What is the constant

iv) What is the slope?

v) Is the slope statistically significant?

vi) Write out the model regression equation

vii) What is the standardised regression coefficient for vehicle weight?

viii) If a car weighed 1000 pounds, what would be the predicted miles per gallon?

ix) What is the standard error of the estimate?

x) What would be the approximate 95% confidence interval of our prediction

QUESTION FOUR

A matrix B is given by B=[1,4,3,2;2,1,3,2;1,2,3,4;5,2,1,4]

(a). Describe how you can extract a 2×2 matrix from the third row and second column (4 marks).

(b). Explain how you can determine the Mean, Standard Deviation of the first two columns of matrix B (6 marks)

(c). The correlation coefficients of rows and columns in matrix B can be given by:

1.0000 -0.3149 -0.9685 0.4575
-0.3149 1.0000 0.1325 -0.2294
-0.9685 0.1325 1.0000 -0.5774
0.4575 -0.2294 -0.5774 1.0000

(i) . Describe how you can generate such correlation coefficients from matrix B using a MATLAB command. Explain statistical interpretations and importance of three sets of columns whose correlation coefficients are 1,0 and -1 respectively (10 marks)

QUESTION FIVE

ANOVA

Age of Respondent

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2784.189	2	1392.095	4.409	012
Within Groups	477048.215	1511	315.717		
Total	479832.404	1513			

Multiple Comparisons

Dependent Variable: Age of Respondent

LSD

(I) Race of Respondent	(J) Race of Respondent	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
White	Black	2.930*	1.344	.029	.29	5.57
	Other	5.609*	2.587	.030	.53	10.68
Black	White	-2.930*	1.344	.029	-5.57	-.29
	Other	2.679	2.828	.344	-2.87	8.23
Other	White	-5.609*	2.587	.030	-10.68	-.53
	Black	-2.679	2.828	.344	-8.23	2.87

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

Consider the SPSS output above

(a) State the Procedure for one-way between-groups ANOVA with post hoc tests

(b) Explain the interpretation of output from one-way between-groups ANOVA with post hoc tests