RESEARCH METHODS
UGBS 602
QUESTIONS FROM 2019 TO 2023

UNIVERSITY OF GHANA

UNIVERSITY OF GHANA BUSINESS SCHOOL
DEPARTMENT OF OPERATIONS AND MANAGEMENT INFORMATION SYSTEMS
WEEKEND/REGULAR MBA: 2ND SEMESTER EXAMINATION: 2022/2023
UGBS 602: RESEARCH METHODS (3 CREDITS)

INSTRUCTIONS:
Answer **TWO (2) Questions** in **SECTION A (Theory, Concepts and Research Process)** and **ONE (1) Question** from **SECTION B (Quantitative Data Analysis)**. Answer **briefly** but **clearly** each part of a Question. Statistical Tables and Formulae are provided. **Use a separate Answer Booklet for each SECTION and Write the section letter on the front page of the Booklet. Write the name of your Lecturer on the front page of your answer booklet.**

TIME ALLOWED: 3 HOURS

SECTION A (Theory, Concepts and the Research Process): Answer Question 1 and any other question from this Section

QUESTION 1 (Compulsory)

a) In the extract below, correct all the in-text citation errors using the APA referencing style. Each citation is numbered for easy identification purposes. Write the correct citation as it should appear in the abstract, against each number in your answer booklet. Do **NOT** write the entire extract.
(15 marks)

E-commerce, a global trend for businesses to promote their operations and enhance sales is a trendy business model. E-commerce has been examined in the literature within the context of its contribution to competitive advantage **[1] (Boateng, R., Heeks, J, Molla, & Hinson, 2011);** the benefits it offers businesses **[2] (Carmichael – 2017);** the influence of firms' strategic values on its adoption; and the factors affecting it's adoption **[3] (Isaac Addo; 2012).** Further, studies have examined e-commerce business model among SMEs' in Ghana. These studies have appear to centre on the influences of e-commerce adoption the association between a firm's strategic value and e-commerce adoption **[4] (DeBerry-Spence., Marko., J. et al., 2008);** and the benefits of e-commerce to business **[5] (Xuhua, Elikem, Akaba, & Worwui-Brown, 2019; Boateng, R. (2016); Iddris, 2012). [6] (Boateng; R., 2016)** however observes that, studies on the influence ~~that the strategic~~ behaviour of firms have on the adoption of e-commerce has received less ~~literature~~ attention. This is a gap in the literature which demands research effort to understand; it

is imperative to understand the strategic behaviour of SMEs and whether these behaviours do affect their decision to adopt ecommerce [7] (Miles & Snow (1978).

Strategic behaviour has been widely studied in the literature. Several studies have validated the Miles and Snow typology as an appropriate theory for studying strategy behaviour. It is also interesting to note that Miles and Snow typology has been employed in many studies including e-commerce [8] (Svobodová & Rajchlová, 2020/Kumar, Fuksa, & Kumar, P 2018 et al.) because it is premeditated to apprehend and integrate organizational culture into a strategic framework. A merit of this theory is that, it hypothesizes that owner/manager personality and organizational culture have a strong influence on organizational strategy. However, studies relating to the influence that strategic behaviour has on e-commerce adoption focused on leadership variables (administrative) rather than technological and entrepreneurship factors [9] (Edu No Date). Therefore this study fills the gap by focusing on the four strategic behaviours (Analyser, Defender, Prospector and Reactor) described by Miles & Snow (1978) relative to their e-commerce adoption (technology). Examining the relationship between Miles and Snow's typology and firms' ecommerce adoption will help improve the understanding of the relative patterns of strategic behaviour developed by SMEs and how such patterns align with firms' technological orientations [10] - Giao, H. N. K. (2020).

b) In the reference list below, identify the source (e.g. Journal article, News Paper article etc) of each reference. (7 marks)

Reference List

i. Edwards, A. A., Steacy, L. M., Siegelman, N., Rigobon, V. M., Kearns, D. M., Rueckl, J. G., & Compton, D. L. (2022). Unpacking the unique relationship between set for variability and word reading development: Examining word- and child-level predictors of performance. *Journal of Educational Psychology*, 114(6), 1242–1256. https://doi.org/10.1037/edu0000696

ii. Giao, H. N. K. (2020). Customer Satisfaction at Tiki. vn E-Commerce Platform. *The Journal of Asian Finance, Economics, and Business*, 7(4), 173–183

iii. Boateng, R. (2016). Resources, electronic-commerce capabilities and electronic-commerce benefits: Conceptualizing the links. *Information Technology for Development*, 22(2), 242–264

iv. Roberts, S. (2020, April 9). Early string ties us to Neanderthals. *The New York Times*. https://www.nytimes.com/2020/04/09/science/neanderthals-fiber-string-math.html

v. Reynolds, G. (2019, April 9). Different strokes for athletic hearts. *The New York Times*. Page D4.

vi. Kaufman, K. A., Glass, C. R., & Pineau, T. R. (2018). *Mindful sport performance enhancement: Mental training for athletes and coaches*. American Psychological Association. https://doi.org/10.1037/0000048-000

vii. Zeleke, W. A., Hughes, T. L., & Drozda, N. (2020). Home–school collaboration to promote mind– body health. In C. Maykel & M. A. Bray (Eds.), *Promoting mind–body health in schools: Interventions for mental health professionals* (pp. 11–26). American Psychological Association.

viii. Horvath-Plyman, M. (2018). *Social media and the college student journey: An examination of how social media use impacts social capital and affects college choice, access, and transition* (Publication No. 10937367). [Doctoral research, New York University]. ProQuest.

c) Break down the reference below into its constituent parts. **(8 marks)**

Edwards, A. A., Steacy, L. M., Siegelman, N., Rigobon, V. M., Kearns, D. M., Rueckl, J. G., & Compton, D. L. (2022). Unpacking the unique relationship between set for variability and word reading development: Examining word- and child-level predictors of performance. *Journal of Educational Psychology*, 114(6), 1242–1256. https://doi.org/10.1037/edu0000696

QUESTION 2

Using the relevant framework, systematically demonstrate the key phases in the research process to be followed in conducting your MBA research. **(20 marks)**

QUESTION 3

a) *Qualitative researchers "study things in their natural settings, attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them" (Denzin & Lincoln, 2005, p.3).* Briefly discuss the circumstances under which a qualitative researcher will adopt the types of research below:

i. Phenomenology **(4 marks)**
ii. Ethnography **(4 marks)**
iii. Case Study **(4 marks)**

b) With relevant example(s), discuss the usefulness of Issue Gap, Method Gap and Context Gap as potential contributions to the literature. **(8 marks)**

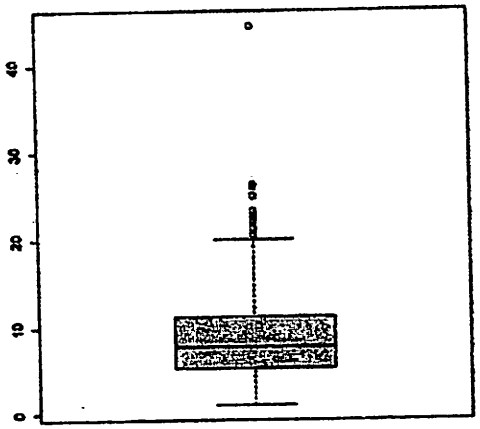
SECTION B:
ANSWER ONLY 1 OUT OF THE 2 QUESTIONS IN THIS SECTION

QUESTION 4

The Current Population Survey (CPS), sponsored jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics (BLS), is the primary source of labor force statistics for the U.S. population. CPS is used to supplement census information between census years. These data consist of a random sample of workers from the CPS85 (from the *mosaicData* package, obtained directly in *R* as **data(CPS85, package="mosaicData")**), with information on wages in US dollars per hour (*wage*) which is influenced by such predictors and other characteristics of the workers, including the number of years of education (*educ*); then *race*, a factor with levels NW (nonwhite) or W (white); *sex*, a factor with levels F and M; *hispanic*, a factor with levels Hisp or NH (Not Hispanic); *south*, a factor (to indicate if the worker is not from the south or from the south) with levels NS or S; *married*, to indicate whether the worker is Married or Single; *exper*, which is the number of years of work experience (inferred from age and educ); *union*, which shows whether the worker does not belong to

a union (Not) or does belong to it (Union); *age*, representing the age in years of a worker; and *sector*, which is a factor showing the type of sector the person works with levels: clerical, const (construction), manag (management), manuf (manufacturing), other, prof (professional), sale \ and service.

As part of your assessment, you are required to demonstrate your understanding of quantitative research and data analytics you have learnt. Results of the boxplot, correlation and regression analysis conducted in *R* software are presented below. Use them to answer the questions that follow:



```
> shapiro.test(CPS85$wage)

        Shapiro-Wilk normality test

data:  CPS85$wage
W = 0.8673, p-value < 2.2e-16
```

a) From the boxplot, provide two comments on shape and distribution of wage. Using the Shapiro Wilk's test of normality for wage, is wage normally distributed? Justify. (4 marks)

```
> table(CPS85$sector, CPS85$sex) #create contingency table

          F  M
clerical 76 21
const     0 20
manag    21 34
manuf    24 44
other     6 62
prof     52 53
sales    17 21
service  49 34
```

b) From the contingency table, how many clerical workers were males and how many females were in the sample? Justify. (3 marks)

```
        wage  educ exper   age   y
wage    1.00  0.38  0.09  0.18 NaN
educ    0.38  1.00 -0.35 -0.15 NaN
exper   0.09 -0.35  1.00  0.98 NaN
age     0.18 -0.15  0.98  1.00 NaN
y        NaN   NaN   NaN   NaN   1

n= 534


P
      wage   educ   exper  age    y
wage        0.0000 0.0443 0.0000
educ  0.0000       0.0000 0.0005
exper 0.0443 0.0000       0.0000
age   0.0000 0.0005 0.0000
y
```

c) Identify the pair of numerical variables that is mostly multicollinear. Is it significantly multicollinear? Justify (2 marks)

d) Which pair of numerical variables is least correlated with the outcome variable? (2 marks)

```
Call:
lm(formula = wage ~ educ + race + sex + hispanic + south + married +
    exper + union + age + sector, data = CPS85)

Residuals:
    Min      1Q  Median      3Q     Max
-11.504  -2.516  -0.614   1.945  34.898

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    -3.4138     6.7166  -0.508  0.61148
educ            0.7828     1.0901   0.718  0.47303
raceW           0.7869     0.5771   1.363  0.17333
sexM            1.8675     0.4232   4.412 1.24e-05 ***
hispanicNH      0.5959     0.8721   0.683  0.49471
southS         -0.6402     0.4200  -1.524  0.12804
marriedSingle  -0.3105     0.4127  -0.752  0.45212
exper           0.2095     1.0850   0.193  0.84693
unionUnion      1.5555     0.5139   3.027  0.00259 **
age            -0.1205     1.0841  -0.111  0.91157
sectorconst     1.0379     1.1304   0.918  0.35893
sectormanag     3.3040     0.7648   4.320 1.87e-05 ***
sectormanuf     0.4558     0.7308   0.624  0.53312
sectorother     0.3208     0.7621   0.421  0.67401
sectorprof      1.9849     0.6726   2.951  0.00331 **
sectorsales    -0.7497     0.8400  -0.893  0.37252
sectorservice  -0.7482     0.6640  -1.127  0.26031
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.295 on 517 degrees of freedom
Multiple R-squared:  0.3224,    Adjusted R-squared:  0.3014
F-statistic: 15.37 on 16 and 517 DF,  p-value: < 2.2e-16
```

e) Is the regression model well fit? Interpret the appropriate statistic. (2 marks)

f) Specify the regression model based on the *R* regression output. (5 marks)

g) Write the estimated equation for the regression estimates (5 marks)

h) Which regression estimate(s) is/are statistically significant? That is, check which variables are statistically significant or have predictive power to predict the target variable. (4 marks)

i) Determine the joint significance of the regression estimates and explain it (3 marks)

j) Interprete the coefficient of `marriedSingle` (3 marks)

k) Holding all other factors constant, would you advice a worker to work in the professional sector? Why? (3 marks)

l) What is the predicted wage differential between the wage of a 51-year old non-Hispanic male white worker with 15 years of education, single and living in the South, has 33 years of experience, does not belong to a union and working in the manufacturing sector, compared with the wage of a 40-year old Hispanic female non-white worker with 17 years of education, married and living in the South, has 29 years of experience, does not belong to a union and working in the service sector. What factors are driving the predicted difference? (14 marks)

QUESTION 5

a) We want to find out if caffeine affects heart rate. We measure the heart rate of some skewed after drinking decaffeinated coffee and that of another after drinking regular coffee. We get the following (somewhat exaggerated) results. Is there a statistically significant difference in the measured heart rates to justify if caffeine in coffee does affect heart rate? Assume a significance level of one percent.

Decaffeinated Coffee	93	67	73	42	70	69	68	69			
Regular Coffee	70	67	67	93	70	68	90				

i. Formulate the required hypotheses. (2 marks)
ii. What is the critical value of the test statistic? (2 marks)
iii. Compute the value of the test statistic. (9 marks)
iv. Provide a statistical and practical conclusion (3 marks)
v. What is the nonparametric analogue of the paired t test? (1 mark)
vi. When is a sample said to be small? (1 mark)
vii. Was the observed difference due to chance? Justify. (2 marks)

b) As part of his project work, Opeele would like to determine if there is a significant difference, at the five-significance level, among the salaries of five government workers in and around the University of Ghana – lecturers, nurses, administrators, police and lawyers. The sample information below shows the salaries (in thousands of Ghanaian cedis):

lecturers	nurses	administrators	police	lawyers
8	9	5	5	4
7	7	8	4	9
10	8	10		6
6		5		3
9				5

i. Formulate the research hypotheses? (2 marks)
ii. Determine if a statistically significant difference exist among the salaries of the five government workers (19 marks)
iii. What is the nonparametric analogue of the dependent ANOVA test? (2 marks)
iv. What is the parametric analogue of the Wilcoxon-signed rank test? (2 marks)
v. Which government workers needs the most salries? (2 marks)

c) You have been provided an extract of a questionnaire. For each question identify the type of data and the scale of measurement. (4 marks)

Background Information

1. According to the weatherman, what will the temperature be today? (1 mark)
2. What is your Gender? Male [ ] Female [ ] (1 mark)
3. Highest level of education attained: Diploma/HND [ ] Undergradpate [ ] Postgraduate [ ] Professional certificate [ ] Other (specify)..... (1 mark)
4. What time do you report to work?..... (1 mark)

UNIVERSITY OF GHANA

UNIVERSITY OF GHANA BUSINESS SCHOOL
DEPARTMENT OF OPERATIONS AND MANAGEMENT INFORMATION SYSTEMS
WEEKEND/REGULAR MBA: 2ND SEMESTER EXAMINATION: 2021/2022
UGBS 602: RESEARCH METHODS (3 CREDITS)

INSTRUCTIONS:

Answer **TWO (2) Questions** from **SECTION A (The Qualitative Section)** and **TWO (2) Questions** from **SECTION B (The Quantitative Section)**. Answer briefly but clearly each part of a Question. Statistical Tables and Formulae are provided. Use a separate answer booklet for each SECTION and write the Section Letter on the cover page of your answer booklets. Ensure that the name of your lecturer is also written on the cover page of your answer booklets.

TIME ALLOWED: 3 HOURS

SECTION A (QUALITATIVE PART): Answer Two (2) questions from this Section

QUESTION 1

You have taken the UGBS 602 (Research Methods) course this semester. Relative to the concepts and the theories learnt in the course, discuss how the knowledge gained from the course will be applied in the course of undertaking your final MBA research project. **(25 marks)**

QUESTION 2

a) On the basis of the purpose of research, discuss the three types of research that may fall under this classification. **(12 marks)**

b) Basic research does not differ from applied research. Do you agree to this claim? **(13 marks)**

QUESTION 3

a) Using the APA style referencing style (in-text), identify the from the extract below, the citations which have been wrongly styled and recommend the acceptable way to cite them. **(10 marks)**

"The Resource Based View (RBV) has been acknowledged as an organization's capability to firmly synchronize all resources, both physical and intellectual capital for value creation. The principal assumptions from RBV are that a firm can be profitable in a competitive market as long as it leverages on the bundle of strategic resources (Barney, 2001; Wang and Hajli, 2017). ..... The contention, however, is the level at which distinctive capabilities of organizations' resources are developed to support business needs. Hence, championing only the RBV alone will

not yield the needed business value from IT and other resources available to an organization (Barney et al., 2001). It is, therefore, essential to further concentrate on how the capabilities of IT resources and other standpoints support organizational processes. Various IT capabilities have been considered to include "managerial IT skills, technical IT skills, IT infrastructure, IT-enabled processes, relationship IT infrastructure, IT business experience and IT based-assets" (Garrison, G., Wakefield, R.L. and Kim, S. 2015; Fink, 2011; Bhatt and Grover, 2005). Significant findings from extant studies have suggested that deploying the applications of these capabilities alongside other resources support an organization's performance (Zhou, K., Liu, T. and Zhou, L. 2015; Bhatt and Grover, 2005)".

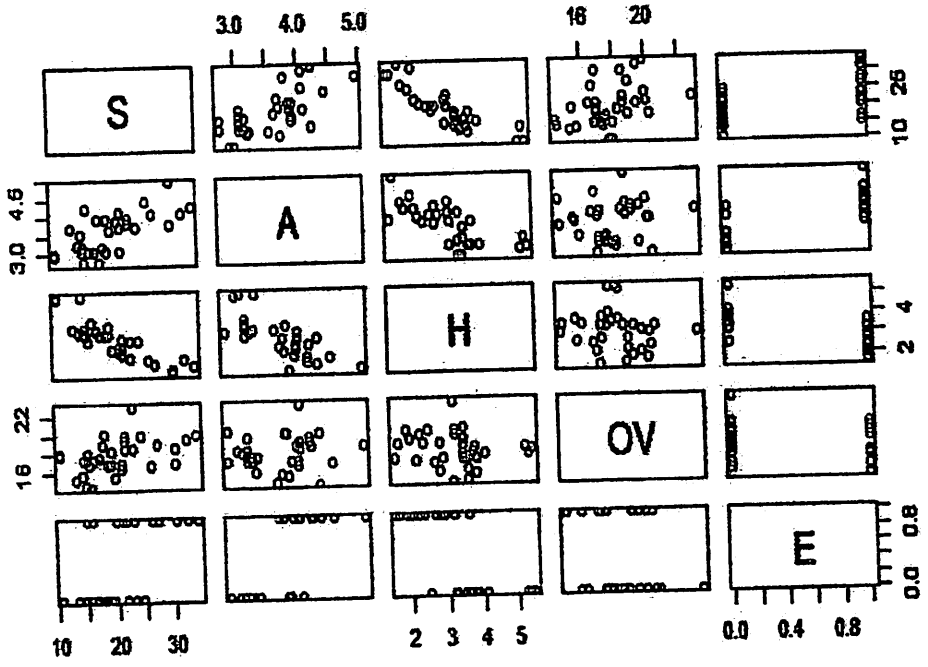
b) Briefly explain the concept of referencing in research. **(5 marks)**
c) Identify and briefly explain the stages in the literature review process. **(5 marks)**
d) Briefly explain why conducting a preliminary literature review is important. **(5 marks)**

SECTION B (QUANTITATIVE SECTION): Answer Two (2) questions from this Section. Question 4 is compulsory

QUESTION 4 (COMPULSORY) (30 MARKS)

The Ministry of Labour has received a lot of complaints from workers about the salaries they receive which the Ministry believes is not fair. The Ministry then wants to understand the factors that mostly influence the salary (S) received by a worker. Of the factors identified, the following are worthy to note: the educational level of workers (basic, secondary, first degree, master's degree and doctorate degree, where B=1 if a worker has basic education or 0 otherwise, SE=1 if a worker has secondary education or 0 otherwise, M=1 if a worker has masters or 0 otherwise, P=1 is a worker has a PHD or 0 otherwise), age of the workers (A), number of hours workers worked in day (H), number of overtime hours (OV), number of years of experience (E), nature of the job (where R=1 if the job is risky or 0 if not risky), type of job (where MA=1 if manufacturing or 0 if service), position held by the worker (where J=1 if the worker occupies a junior position or 0 if senior position), location of the job (urban, suburban and rural, where U= 1 if location is urban or 0 otherwise, SB=1 if location is suburban or 0 otherwise), gender of the worker (where F=1 if the worker is male or 0 if female), marital status (where SI=1 if single or 0 if married). Use the scatter plot matrix of selected variables, correlation matrix and dummy regression output from R below to answer the questions that follow).

a) From the scatter plot matrix, is the relationship between S and A, and OV and E neutral, negative or positive? **(2 marks)**
b) Which 3 pairs of variables are most strongly multicollinear? Using only the variance inflation factor (VIF) in the regression output, which one of the pairs of variables selected to be multicollinear could be deleted from the regression? **(3 marks)**



	S	B	SE	M	P	A	H	OV	E	R	MA	J	U	SB	F	SI
S	1.00															
B	0.80	1.00														
SE	-0.03	-0.38	1.00													
M	-0.85	-0.75	-0.21	1.00												
P	-0.78	-0.69	-0.19	0.79	1.00											
A	0.68	0.65	-0.01	-0.71	-0.45	1.00										
H	-0.87	-0.70	-0.06	0.89	0.66	-0.71	1.00									
OV	0.42	0.53	0.04	-0.43	-0.71	0.09	-0.17	1.00								
E	0.60	0.47	0.02	-0.59	-0.24	0.71	-0.69	-0.23	1.00							
R	-0.63	-0.55	-0.19	0.74	0.41	-0.83	0.66	-0.08	-0.78	1.00						
MA	0.58	0.53	0.21	-0.68	-0.66	0.66	-0.48	0.49	0.41	-0.73	1.00					
J	0.47	0.41	0.09	-0.42	-0.48	0.09	-0.40	0.50	0.18	-0.04	0.21	1.00				
U	0.26	0.36	-0.36	-0.12	-0.29	0.13	-0.25	0.13	-0.01	-0.09	0.03	-0.36	1.00			
SB	-0.21	-0.23	-0.17	0.12	0.16	-0.32	0.21	-0.03	-0.27	0.34	-0.25	-0.17	-0.22	1.00		
F	-0.49	-0.49	0.30	0.43	0.40	0.00	0.48	-0.34	-0.15	0.04	0.03	-0.36	-0.45	-0.22	1.00	
SI	-0.01	-0.13	0.34	-0.13	0.08	0.01	-0.08	-0.24	0.22	-0.17	-0.14	-0.10	-0.12	-0.06	-0.12	1.00

c) Specify the entire regression model. (4 marks)
d) Write the estimated equation for the regression (3 marks)
e) Determine the joint significance of the regression estimates (2 marks)
f) Is the regression model well fitted? (2 mark)
g) How many regression estimates are statistically insignificant? (3 marks)
h) Interpret the coefficient of secondary level of education of a worker (SE) (3 marks)
i) Holding all other factors constant, which type of educational level of a worker is expected to have the most potential effect on salary? (2 marks)
j) What is the predicted difference in the salary of a 30 year old single female who is a secondary educational worker, occupies a junior position and works for 6 hours a day with 4 hours overtime, has 8 years' experience in a manufacturing risky job located in a rural area. Compare this with the Salary of a 36 year old married male, who his masters degree senior worker, who works for 4 hours with no overtime, has 12 years' experience in the non-risky service industry located in an urban area. What categorical factors are driving the predicted difference? (9 marks)
k) Identify 2 other names used for the random noise in a regression model? (2 marks)

Regression output

Residuals:

Min	1Q	Median	3Q	Max
-3.6665	-1.1695	-0.1723	0.7962	4.6879

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	VIF
(Intercept)	30.13986	20.63780	1.460	0.1635	
B	2.16807	6.50671	0.333	0.7433	39.42
SE	-0.81958	4.94453	-0.166	0.8704	17.25
M	0.03583	0.03135	1.143	0.2698	60.36
P	-0.06037	0.03580	-1.686	0.1111	24.10
A	1.20419	2.44040	0.493	0.6284	6.81
H	4.65867	2.48779	-1.873	0.0005	23.69
OV	0.43395	0.91415	0.475	0.6414	10.67
E	0.33234	2.88450	0.115	0.9097	8.28
R	-0.89650	3.00672	-0.298	0.7694	13.44
MA	-2.96365	3.61633	-0.820	0.4245	8.75
J	-5.59275	5.58816	-1.001	0.3318	42.66
U	-5.43629	7.77674	-0.699	0.4946	50.00
SB	-2.76403	5.29226	-0.522	0.6086	14.32
F	-3.27243	6.38908	-0.512	0.6155	27.69
SI	-7.02194	7.50832	-0.935	0.3636	3.5

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.784 on 16 degrees of freedom
Multiple R-squared: 0.8899, Adjusted R-squared: 0.7866
F-statistic: 8.617 on 15 and 16 DF, p-value: 4.935e-05

QUESTION 5 (15 MARKS)

Below are the annual salaries (in thousands of GHS) of students who had degrees in economics, medicine, history, and nursing. A researcher wants to find out if there is a statistically significant difference in the mean salaries across the degree courses obtained.

Economics	Medicine	History	Nursing
58	67	49	55
69	81	52	41
60	70	48	49
65	59	53	57
71	61		58
74			

i. Formulate the research hypotheses for the above preamble? **(2 marks)**
ii. Compute the suitable test statistic to determine if a statistically significant difference exists among each of the four-degree categories. **(9 marks)**
iii. What is the nonparametric analogue of the repeated *t*-test statistic and the independent ANOVA? **(2 marks)**
iv. Rank from the lowest to the highest, the four-degree categories. **(2 marks)**

QUESTION 6 (15 MARKS)

We want to find out if caffeine affects heart rate. We take seven volunteers and measure their heart rate after drinking decaffeinated coffee. We take another six volunteers and measure their heart rate after drinking regular coffee. We get the following (somewhat exaggerated) results. A researcher wants to find out if there is a statistically significant difference in the measured heart rates to justify if caffeine in coffee does affect heart rate. Assume a significance level of one percent.

Decaffeinated Coffee	Regular Coffee
93	74
67	81
73	124
42	79
70	78
69	96
68	

i. Formulate the required hypotheses. **(2 marks)**
ii. What is the critical value of the test statistic? **(1 mark)**
iii. Compute the value of the test statistic. **(8 marks)**
iv. Provide a statistical and practical conclusion. **(3 marks)**
v. What is the nonparametric analogue of the paired *t*-test? **(1 mark)**

UNIVERSITY OF GHANA

REGULAR/WEEKEND MBA, SECOND SEMESTER EXAMINATIONS: 2020/2021
UGBS 602: RESEARCH METHODS (3 CREDITS)

INSTRUCTIONS:

Answer ALL Questions in **SECTION A (The Qualitative Part)** and ONE Question ONLY in **SECTION B (The Quantitative Part)**. Answer **Briefly** but **Clearly** Each Part of a Question. Statistical Tables and Formulae Are Provided. **Use A Separate Answer Booklet for Each SECTION and Write the Section Letter on the Front Page of the Booklet.**

TIME ALLOWED: TWO HOURS (2 HOURS)

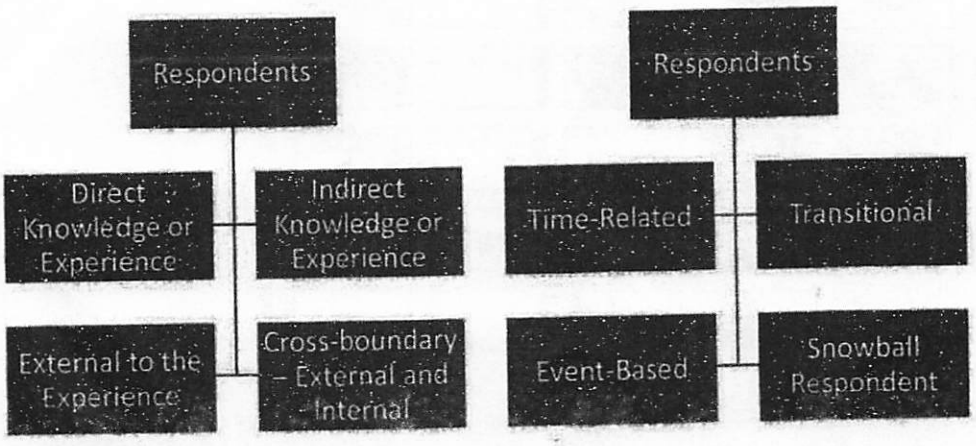
SECTION A:

ANSWER ALL QUESTIONS IN THIS SECTION

QUESTION 1

Blue Skies has been producing fresh fruit products from its factory near to Nsawam in Ghana since 1998. The company started by exporting premium quality freshly cut fruits to supermarkets in Europe before eventually diversifying to supply the local market with freshly squeezed 100% natural juice. You are a student carrying out research on the impact of COVID-19 on the operations of manufacturing companies, using Blue Skies Holdings as a case study. Using the respondent matrix below, categorise and justify your selection of respondents for the study.

Figure 1 Respondents Matrix for Data Collection



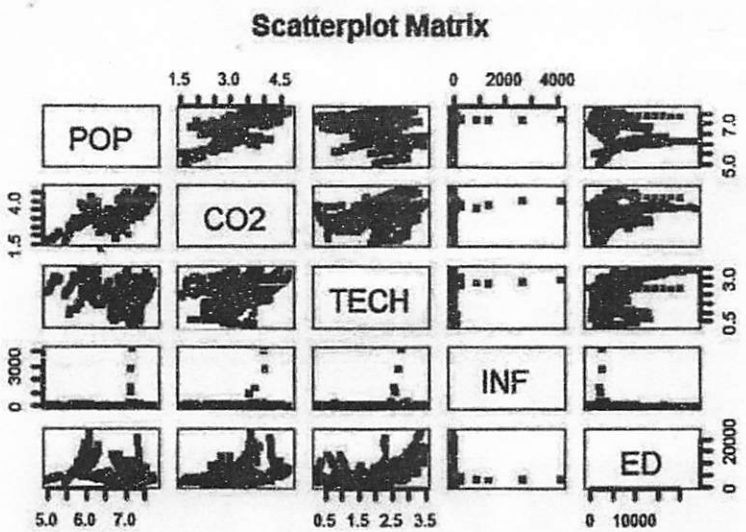
(50 marks)

SECTION B:
ANSWER ONLY 1 QUESTION OUT OF THE 2 IN THIS SECTION

QUESTION 2 (50 MARKS)

The United Nations Sustainable Development Goals (SDG) 7 seeks to ensure clean, affordable and sustainable energy for all by 2030 through cost effective production and usage of energy resources in order to boost economic growth of many states. To track the performance of states towards the achievement of this goal, the cost Malmquist productivity change index, **CMPI**, a performance measure, which measures the energy dynamic performance over time is computed. However, it is believed that some external factors beyond the control of these states may affect their **CMPI** over time. Therefore, the UN wants to determine how the following exogenous factors affect the CMPI of the states - a state's population (**POP**), the level of energy related carbon dioxide emissions (**CO2**), the level of technological innovation (**TECH**), inflation (**INF**), economic development of a state (**ED**), the state's affiliation to the Organization of the Petroleum Exporting Countries (where **OPEC**=1 if a state is an OPEC member or 0 otherwise), and finally, a state's affiliation to a regional bloc (AMU, EAC, ECCAS, ECOWAS, SADC; where **EAC**=1 if a state is a member of EAC or 0 otherwise, **ECCAS**=1 if a state is a member of ECCAS or 0 otherwise, **ECOWAS**=1 if a state is a member of ECOWAS or 0 otherwise, **SADC**=1 if a state is a member of SADC or 0 otherwise). Use the scatter plot matrix, the correlation with its p-value matrix and regression output to answer the related questions.

Scatterplot Matrix with selected variables:



a) From the scatter plot matrix, determine the relationship between POP and CO2, and between INF and ED. **(2 marks)**

Correlation Matrix:

Pearson correlations:

	CMPI	POP	CO2	TECH	INF	ED	OPEC	EAC	ECOWAS	ECCAS	SADC	AMU	VIF
CMPI	1.00												4.86
POP	0.03	1.00											6.40
CO2	-0.01	0.62	1.00										2.31
TECH	-0.09	-0.25	0.38	1.00									1.08
INF	0.03	0.05	0.06	0.04	1.00								1.92
ED	0.02	-0.31	0.27	0.49	-0.03	1.00							2.22
OPEC	-0.01	-0.03	0.25	0.33	0.24	0.34	1.00						3.26
EAC	0.06	0.15	0.00	-0.28	-0.04	-0.10	-0.17	1.00					3.26
ECOWAS	-0.02	0.00	-0.18	-0.14	-0.05	-0.19	-0.18	-0.41	1.00				3.26
ECCAS	-0.02	-0.17	-0.16	0.14	0.13	0.05	0.53	-0.32	-0.34	1.00			3.26
SADC	-0.01	-0.32	0.01	0.29	-0.02	0.25	-0.09	-0.21	-0.23	-0.17	1.00		3.26
AMU	-0.02	0.32	0.49	0.15	-0.01	0.12	-0.09	-0.21	-0.23	-0.17	-0.11	1.00	3.26

P-Value Matrix:

Pairwise two-sided p-values:

	CMPI	POP	CO2	TECH	INF	ED	OPEC	EAC	ECOWAS	ECCAS	SADC	AMU
CMPI												
POP	0.37											
CO2	0.76	<.0001										
TECH	0.00	<.0001	<.0001									
INF	0.33	0.10	0.04	0.18								
ED	0.54	<.0001	<.0001	<.0001	0.27							
OPEC	0.75	0.33	<.0001	<.0001	<.0001	<.0001						
EAC	0.05	<.0001	0.88	<.0001	0.16	0.00	<.0001					
ECOWAS	0.42	0.97	<.0001	<.0001	0.08	<.0001	<.0001	<.0001				
ECCAS	0.50	<.0001	<.0001	<.0001	<.0001	0.11	<.0001	<.0001	<.0001			
SADC	0.82	<.0001	0.63	<.0001	0.49	<.0001	0.00	<.0001	<.0001	<.0001		
AMU	0.57	<.0001	<.0001	<.0001	0.81	0.00	0.00	<.0001	<.0001	<.0001	0.00	

b) Which 2 pairs of variables are highly correlated with the outcome variable? Are they significant? (2 marks)
c) Which 3 pairs of variables are highly multicollinear? Are they significant? (3 marks)
d) Which pair of variables is the most highly correlated? Is it significant (2 marks)
e) Which variable is the best regressor? (1 mark)
f) If you were to specify the entire regression model in a journal article, how would you specify it? (4 marks)
g) Which two noise factors could be captured in the idiosyncratic error? (3 marks)
h) Write the estimated equation for the regression 3 marks)

Dummy Variable Regression Output:

Residuals:

Min	1Q	Median	3Q	Max
-0.7309	-0.0741	-0.0108	0.0491	2.0349

Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.9235	0.1121	8.2370	0.0000	***
POP	0.0366	0.0221	1.6530	0.0987	.
CO2	-0.0279	0.0227	-1.2260	0.2205	
TECH	-0.0352	0.0129	-2.7370	0.0063	**
INF	0.0000	0.0000	1.0690	0.2852	
ED	0.0000	0.0000	2.0310	0.0425	*
OPEC	0.0283	0.0357	0.7950	0.4270	
BlocEAC	-0.0331	0.0258	-1.2860	0.1988	
BlocECCAS	-0.0099	0.0295	-0.3370	0.7366	
BlocECOWAS	-0.0093	0.0239	-0.3900	0.6963	
BlocSADC	0.0543	0.0222	2.4470	0.0146	*

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1987 on 1062 degrees of freedom
Multiple R-squared: 0.03069, Adjusted R-squared: 0.02156

F-statistic: 3.363 on 10 and 1062 DF, p-value: 0.0002492

i) Are the coefficients jointly significance? Justify your answer (2 marks)
j) Using the VIF, which variable in the pairs of variables selected to be multicollinear may be deleted. (2 marks)
k) Is the regression model well fitted? Justify you answer. (2 marks)
l) What is the use of the coefficient of the regional bloc called BlocEAC? (3 marks)
m) Interpret the coefficient of CO2 (3 marks)
n) Distinguish between heteroskedasticity and multicollinearity. What could possibly the impact of the presence heteroskedasticity in the regression output? (3 marks)
o) What regression *command* would be used to test if the dependent variable was normally distributed? (Hint: assume the data set is called "energy") (3 marks)
p) What is the predicted difference in the performance or CMPI of a state with 4,000,000 population, 300 level of energy related carbon dioxide emissions, 40 level of technological innovation, 0.07 inflation rate, 36 economic development, is a non OPEC member state, and affiliated with AMU regional bloc, compared with the CMPI of another state with 4,000,000 population, 200 level of energy related carbon dioxide emissions, 100 level of technological innovation, 0.17 inflation rate, 60 economic development, is an OPEC member state, and affiliated with ECOWAS regional bloc. What factors are driving the difference? (10 marks)
q) What is the economic significance of the constant term in a regression model? (2 marks)

QUESTION 3 (50 MARKS)

a) A new travel and tour company wants to determine if there is a significant difference in their air ticket prices across Europe between cabin classes (economy, premium, business, royals, first class, presidential). A research student at the University of Ghana was presented with the sample information below to help them in their decisions

Economy	Premium	Business	Royals	First class	Presidential
4	6	5	10	10	12
7	7	6	9	7	13
3	6	8	8	10	11
3	7	4	8	9	
8	6	8	9		
5	8	8			
7					

i. Formulate the research hypotheses for the above preamble? (3 marks)
ii. Compute the suitable test statistic to determine if a statistically significant difference exist among the prices of the cabin classes. (13 marks)
iii. What is the nonparametric analogue of the repeated ANOVA test? (2 marks)
iv. Determine 4 assumptions underlying the test statistic computed in (ii). (4 marks)
v. Rank in a descending order from the most expensive to the cheapest, the prices of the companies. (3 marks)

b) A researcher, AGORO is conducting studies on Biomass of various trees. He wished to determine if the kilograms of biomass of male Juniper trees exceeded that of the female Juniper trees. So, he randomly selected some individuals of each gender from the field. He dried them to constant moisture, chipped them, and then weighed them to the nearest kg.

Male	Female
5	2
6	5
8	4
6	4
5	4
7	7
9	9
5	6
7	7
10	
9	

i. Formulate the required hypotheses. (2 marks)
ii. What is the critical value of the test statistic? (2 marks)
iii. Calculate the value of the test statistic. (10 marks)
iv. Provide a statistical and practical conclusion (4 marks)
v. Was the observed difference due to chance? Justify. (3 marks)

vi. What is the nonparametric analogue of the unpaired t test? (2 marks)
vii. Identify 2 assumptions undelying the test statistic used. (2 marks)

SOME USEFUL FORMULAE

SOME USEFUL FORMULAE

1. $t_{df} = t_{N-K-1,\alpha}$ 2. $t_{\hat{\beta}} = \frac{\hat{\beta}}{S.E.(\hat{\beta})}$ 3. $F = \frac{MSR}{MSE}$ 4. $F_{K,N-K-1,\alpha}$

5. $N-1 = TSS_{df}$ 6. $MSR = \frac{RSS}{K}$ 7. $R^2 = \frac{RSS}{TSS}$ 8. $\hat{\beta}_j \pm t_{n-K-1,\alpha} * S.E._{(\hat{\beta})}$

9. $\bar{R}^2 = 1-(1-R^2)\left(\frac{N-1}{N-K-1}\right)$ 10. $MSG = \frac{SSG}{K-1}$ 11. $MSW = \frac{SSW}{N-K}$

12. $F = \frac{MSG}{MSW}$ 13. $\bar{\bar{x}} = \frac{\sum_{i}^{K}(n_i\bar{x}_i)}{N}$ 14. $SSG = \sum_{i=1}^{K} n_i(\bar{x}_i - \bar{\bar{x}})^2 -$ 15. $t = \frac{\bar{d}}{S_d/\sqrt{n}}$

16. $SSW = \sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$ 17. $t = \frac{\frac{\sum d_i}{n}}{\frac{S_d}{\sqrt{n}}}$ 18. $t_{df} = t_{N-1,\alpha}$

19. $U_1 = R_1 - \frac{n_1(n_1+1)}{2}$ 20. $t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ 21. $U_2 = R_2 - \frac{n_2(n_2+1)}{2}$

22. $U_1 + U_2 = n_1 \times n_2$ 23. $\sum(R_1 + R_2) = \frac{n(n+1)}{2}$ 24. $U_{n1,n2,\alpha}$

25. $MSE = \frac{ESS}{N-K-1}$ 26. $F_{K-1,N-K,\alpha}$ 27. $\bar{d} = \frac{\sum d}{n}$ 28. $S_d = \sqrt{\frac{\sum(d-\bar{d})^2}{n-1}}$

UNIVERSITY OF GHANA

REGULAR/WEEKEND MBA, 2ND SEMESTER EXAMINATIONS: JUNE, 2019/2020
UGBS 602: RESEARCH METHODS (3 CREDITS)

SECTION A - QUALITATIVE PART

INSTRUCTIONS:
ANSWER ANY ONE OF THE FOLLOWING QUESTIONS in SECTION A.
SECTION A IS THE QUALITATIVE PART OF THE EXAMINATION. MAKE SURE YOU REFER TO SAKAI FOR SECTION B.

TIME ALLOWED: THREE DAYS

QUESTION A (50 MARKS)

The COVID-19 pandemic has been characterised by unprecedented change with very severe socio-economic consequences in the way of life, business and society. In response, governments, civil society, private sector firms, non-governmental institutions, entrepreneurs and citizens have employed a diversity of strategies to mitigate its effects. Write an essay of not less than 1,800 words (excluding references) on any of the following themes:

A. Examine impact of the COVID-19 pandemic on businesses, public sector institutions, non-governmental and civil society organisations, governments, or international development institutions in any African country of choice and discuss the strategies implemented by your selected entity/business/organisation/government to mitigate the effect brought by COVID-19.
B. Examine how the COVID-19 pandemic has changed or is changing entrepreneurship – are there any new start-ups birthed by the effects of COVID-19 and how are existing entrepreneurs or start-ups faring.
C. Examine how the COVID-19 pandemic is shaping political discourse and activities of political parties and the lessons/recommendations that can shape the future of politics in Africa.
D. Examine the experiences of healthcare personnel, frontline workers and government and non-governmental officials concerning how the COVID-19 pandemic and how these might necessitate changes in the administrative structures in healthcare delivery and the business of healthcare, in Africa.
E. Examine how the COVID-19 pandemic is redefining the "concept of the workplace" and the rise of telecommuting or working from home, and how all these changes affect human resource management activities, such as performance assessment, appraisals, recruitment, motivation, teamwork et cetera, in Africa.

F. Examine the effects of the COVID-19 pandemic on education institutions at any level, the response strategies by educational institutions and lessons/recommendations that can shape education in Africa.

G. Examine the perceptions of consumers and citizens towards misinformation, disinformation and fake news in the local media, international media, governments and unregulated news sources in relation to the COVID-19 pandemic.

H. Examine the effects of the COVID-19 lockdown and social distancing measures on online behaviour of citizens and consumers.

I. Examine the effects of the COVID-19 pandemic on the family unit and social relations between social groups or communities and the resulting effects on social behaviour in Africa.

J. In relation to the COVID-19 pandemic, examine the use of digital payments and other technologies to virtualise processes and transactions and the resulting effects on doing business in Africa.

Answers should be interesting and thought-provoking. You will be assessed by how you are able to use diverse resources – journal articles, conference papers, industry articles, periodicals, government reports, social media briefings, and others, to substantiate your arguments and provide evidence to your discussions. Be specific about the issue you want to address and do not digress.

- Kindly make sure that only RELEVANT diagrams, maps, figures and charts are used to substantiate arguments, where necessary. This is not compulsory.
- The list of references used should not be less than 10 in total.
- Answers should be typed, Font Style and Size – Times News Roman 12pts and Single Spacing. APA Style Referencing (6th or 7th edition) and British English Spelling.

- Turn overleaf for the structure of the essay and the maximum number of words per a section.

STRUCTURE OF THE ESSAY AND MAXIMUM NUMBER OF WORDS

TITLE OF PAPER

- not more than 12 words in title

ABSTRACT (100 words)

QUOTE (20 words):

- A Quote in reference to the pandemic which emphasises your paper. It can be from a leader, government official, professional or person involved or commenting on the issue from anywhere in the world. Make sure you provide reference for the quote.

INTRODUCTION (200 words)

- Introduce the COVID-19 pandemic
- Introduce the business/social phenomena
- Introduce the scope and context

PRIOR TO THE COVID-19 PANDEMIC (200 words)

- Describe the situation prior to the pandemic/ Business as Usual

SITUATION FACED THROUGH THE COVID-19 PANDEMIC (400 words)

- Describe the effects of the pandemic – a minimum of three effects should be discussed

ACTION TAKEN/STRATEGIC ACTIONS AND INFLUENCING FACTORS (300 words)

RESULTS ACHIEVED/NOT ACHIEVED (200 words)

LESSONS AND/OR RECOMMENDATIONS (300 words)

- Lessons and/or sound recommendations for practitioners and policy makers
- Recommendations for specific institutions, agencies and firms in the private, public and non-governmental sector
- What should Africa do to address future pandemics and uncertainties or the post-COVID-19 era?

REFERENCES

- The list of references used should not be less than 10 in total

APPENDIX A - METHODOLOGY (200 words)

- Briefly describe Methods used to collect primary from institutions or persons, and secondary data and resources from the Internet. If you conducted some brief interviews or reviewed websites or social media pages, kindly outline the data collection activities as detailed as possible.

UNIVERSITY OF GHANA

REGULAR/WEEKEND MBA, 2ND SEMESTER EXAMINATIONS: JUNE, 2019/2020
UGBS 602: RESEARCH METHODS (3 CREDITS)

SECTION B - QUANTITATIVE PART

INSTRUCTIONS:

Answer All Questions in This Section. Answer Briefly but Clearly Each Part of a Question. Statistical Tables and Formulae Are Provided. Answers Should Be Typed or Handwritten. SECTION B IS THE QUANTITATIVE PART OF THE EXAMINATION. MAKE SURE YOU REFER TO SAKAI FOR SECTION A.

TIME ALLOWED: THREE DAYS

QUESTION 1 (6 Marks)

A. Consider the following two independent random samples. Sample 1 (from Population 1): 12, 15, 18, 9. Sample 2 (from Population 2): 10, 11, 13, 11. In order to conduct a Wilcoxon Rank Sum test, what rank would you give to the two observations of "11" from Sample 2?

B. Perform a one-sample *t*-test using the following statistics at alpha=1% and aipha=5%:

$n = 5$
$\bar{x} = 3.871$
$s = 0.679$
The null hypothesis is $\mu = 5.0$

QUESTION 2 (5 Marks)

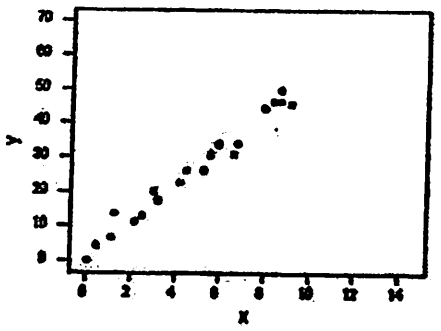
Distinguish between the following:

a. Parametric test and non-parametric test
b. Linearity and normality
c. Interpreting static variable and interpreting dynamic variable
d. P-value and α-value
e. Outliers and influential data points

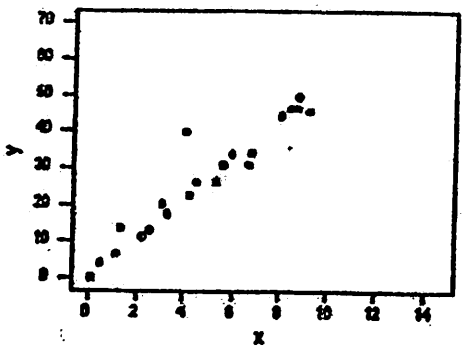
QUESTION 3 (4 Marks)

Based on question 2e above, do you think the following scatter plots contain any outliers or any influential data points? Justify your answers on each plot

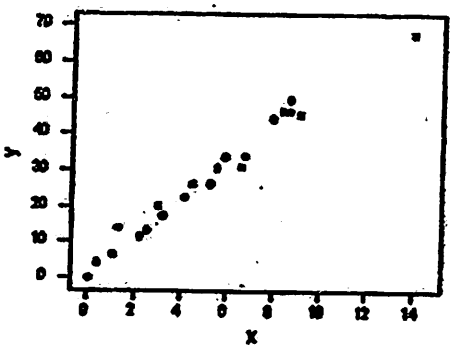
(i)



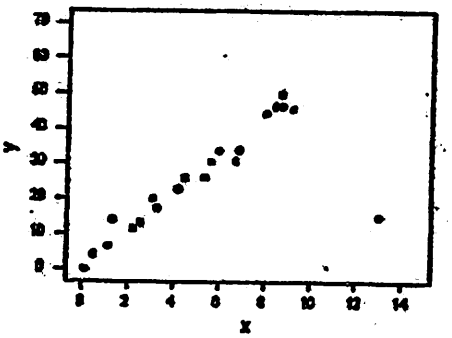
(ii)



(iii)



(iv)



QUESTION 4 (15 Mark)

A. SIC, an insurance company is reviewing its current policy rates. When originally setting the rates they wished that the average claim amount is below GHc1,800. But, they are concerned, now that the true mean is actually at least this amount, because the company could potentially lose a lot of money. They randomly selected 20 claims and found the mean to be GHc1,950 with a standard deviation of GHc500. Use $\alpha = 0.01$, to test if the insurance company should be concerned. Should the insurance company be concerned about their current policies? Justify using both the critical value and the p-value approaches and conclude practically. Is there any difference due to chance?

B. A random sample of starting annual salaries for Lecturers from 2 Ghanaian universities is shown below (in thousands of GHc):

UG	63	70	72	65	70	69		
KNUST	63	73	70	70	74	73	71	65

Formulate the hypotheses to determine if KNUST Lecturers are paid less than UG Lecturers? Justify using both the critical value and p-value approaches. Assume alpha=5%

QUESTION 5 (20 Mark)

A group of researchers want to predict life satisfaction score (LS) of some students seven years after University. The outcome variable is associated with variables that can be measured when the student was in the University. It is believed that LS will be contingent on age (A), gender (G) (which is 0=Male,1=Female), marital status (MS) (which is 0=No,1=Yes), income in University (Y) (in thousands), score on health inventory in University (H), number of children while in University (C), score on life satisfaction inventory in University (LSI), socio economic status of parents (SES), smoker (S) (0=No,1=Yes), score on spirituality inventory in University (SP), finish the program in University (F) (which is 0=No,1=Yes), and income seven years after University (YA) (in thousands). Use the correlation matrix to answer the questions that follow:

	LS	A	G	MS	Y	H	C	LSI	SES	S	SP	F	YA
LS	1.00												
A	0.04	1.00											
G	0.35	-0.02	1.00										
MS	0.45	-0.03	-0.10	1.00									
Y	0.31	0.60	-0.45	0.50	1.00								
H	0.39	0.10	-0.05	0.30	0.54	1.00							
C	-0.02	0.37	-0.02	0.33	0.55	0.43	1.00						
LSI	0.49	0.15	0.04	0.85	0.57	0.27	0.59	1.00					
SES	0.18	0.22	0.02	0.30	0.52	0.75	0.75	0.36	1.00				
S	-0.27	-0.16	0.16	-0.15	-0.51	-0.58	-0.49	-0.31	-0.55	1.00			
SP	0.31	-0.06	0.14	0.00	-0.13	-0.32	-0.34	-0.06	-0.16	0.12	1.00		
F	-0.23	0.15	0.17	-0.46	-0.24	-0.20	-0.31	-0.54	-0.15	-0.11	0.17	1.00	
YA	0.40	0.13	-0.10	0.04	0.28	0.07	-0.13	0.00	0.00	-0.30	0.34	0.50	1.00

a) Which 3 pairs of variables are weakly correlated with the regressand ?
b) Identify 3 pairs of variables that are mostly multicollinear
c) Which pair(s) of variable(s) show no correlation

Now, use the regression output to answer the questions that follow:

	Estimate	Std.Error	t-value	Pr(>|t|)	vif
(Intercept)	37.70	81.84	0.46	0.69	
A	-0.78	0.63	-1.24	0.34	18.09
GMale	-7.59	8.93	-0.85	0.49	28.57
MSYes	-17.05	6.06	-2.81	0.11	11.76
Y	0.60	0.82	0.73	0.54	97.34
H	-0.35	1.31	-0.27	0.81	114.68
C	-6.32	3.39	-1.86	0.20	12.82
LSI	0.74	0.28	2.66	0.12	14.20
SES	0.27	0.33	0.81	0.50	19.66
SYes	-4.57	12.95	-0.35	0.76	53.66
SP	0.21	0.21	0.99	0.43	9.11
FYes	-5.56	4.90	-1.13	0.37	6.77
YA	0.13	0.09	1.53	0.27	5.18

Residual standard error: 3.23 on 2 degrees of freedom (5 observations deleted due to missingness)

Multiple R-squared: 0.9781, Adjusted R-squared: 0.8469
F-statistic: 7.45 on 12 and 2 DF, p-value: 0.12

a. Specify the regression model and write the estimated regression equation
b. Which regressors are statistically different from 0? Justify why you selected those
c. How much higher is the life satisfaction score predicted to be if the score on spirituality inventory rises by 25?
d. Which one of the pairs of variables selected to be multicollinear may be deleted from the regression and why?
e. Interpret the coefficient of smoker (SYes) and income seven years after university
f. Determine and explain the joint significance and the reliability of the regression model
g. Suppose that two students, A and B, are both 29 years old but A is a female whereas B is a male. A is single while B is married. Both had an income of 4500 cedis when in the university and both have health inventory score of 21. Student A had two kids while in the university whereas student B had 4 kids. Both scored 16 on life satisfaction inventories as well as 23 on socioeconomic status of parents. Student A was a smoker but B was not. A scored 7 while B scored 12 on spirituality inventory. Both finished the university program but whiles a was earning 9000 Ghana cedis seven years after university, B was earning 6800. What is the predicted difference in the life satisfaction score of these two student seven years after university?

UNIVERSITY OF GHANA
(All rights reserved)
DEPARTMENT OF OPERATIONS AND MANAGEMENT INFORMATION SYSTEMS
REGULAR/WEEKEND MBA, SECOND SEMESTER EXAMINATIONS: MAY, 2019
UGBS 602: RESEARCH METHODS (3 CREDITS)

INSTRUCTIONS:

Answer ALL Questions in SECTION A (The Qualitative Part) and ONE Question ONLY in SECTION B (The Quantitative Part). Answer Briefly but Clearly Each Part of a Question. Statistical Tables and Formulae Are Provided. Use A Separate Answer Booklet For Each SECTION and Write the Section Letter on the Front Page of the Booklet.

TIME ALLOWED: TWO HOURS THIRTY-MINUTES (2.5 HOURS)

SECTION A:

ANSWER ALL QUESTIONS IN THIS SECTION

QUESTION 1

In reference to your term paper, select a topic of your choice and develop the following:

a) Research Topic
b) Research Title
c) Research Problem (brief including two references)
d) Research Purpose
e) Research Objectives
f) Research Methodology
   i. Research Approach
   ii. Study Location/Area
   iii. Study Population
   iv. Sampling Approach
   v. Data Collection Methods

(30 marks – Q1a to Q1f)

g) Supposing you have a *research grant* of GHS 6,000.00, develop a research budget covering the relevant activities to carry out this research.

(20 marks – Q1g)

SECTION B:
ANSWER ONLY 1 OUT OF THE 2 QUESTIONS IN THIS SECTION

QUESTION 2 (50 Marks)

a) The human resource manager of a certain institution wants to determine if the number of customer complaints differ based on the University completed by its customer service representatives so as to help in its recruitment decisions. Some employees were randomly selected and the number of customer complaints against them were recorded within a six-month period and grouped by the four universities attended by the reps as shown in table 1 below. Using a 5% significance level, does the complaints against the reps differ based on the schools they attended?

UG	KNUST	UCC	UDS	UBW
19	19	15	18	17
17	16	20	31	17
16	16	18	32	17
19	30	17	33	31
17		31		32
16		32		33
		33		

i. What statistical research hypotheses can be formulated? (2 marks)
ii. Compute the value of the test statistic (17 marks)
iii. Find the critical value of the test statistic. (3 marks)
iv. Conclude statistically and practically if a significant difference exists between the complaints against the reps in the various schools (4 marks)
v. What is the parametric analogue of the Kruskal-Wallis test? (2 marks)
vi. List 2 assumptions underlying the test employed here (2 marks)

b) A clinical psychologist wishes to compare two methods for reducing hostility levels in university students, and used a certain test (HLT) to measure the degree of hostility. A high score on the test indicates great hostility. The psychologist used 16 students who obtained high and nearly equal scores in the experiment. Eight were selected at random from among the 15 problem cases and treated with method 1 and the remaining 8 treated with method 2. All treatments were continued for a one-semester period. Each student was given the HLT test at the end of the semester, with the non-normally distributed results shown in the table below. Use $\alpha = 0.01$.

Method 1	Method 2
94	77
80	76
91	74
77	80
83	83
79	80

80	73
74	80

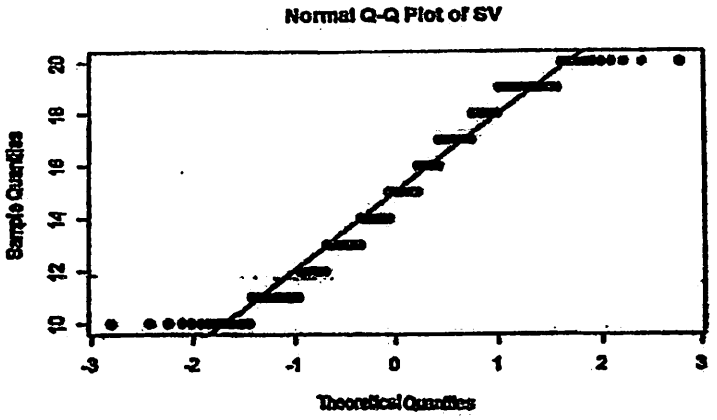
i. State the appropriate null and alternative hypotheses (2 marks)
ii. What is the critical value of the test statistic? (3 marks)
iii. Find the p-value (3 marks)
iv. Is there sufficient evidence to reject the null hypothesis that the two methods are equally effective? (10 marks)
v. What test statistic would have been used for this skewed distributed data if the psychologist was comparing four different methods? (2 marks)

QUESTION 3 (50 Marks)

In attempts to improve the productivity of her staff, the HR manager of ABC Insurance Brokers Ltd seeks to investigate the factors that affect the productivity (measured by ten thousands of sales volumes generated (SV)) of her employees. She has identified a number of factors that she believes explain the productivity of employees as follows; Years of working experience (E), level of commission (C), the Gender (where GM =1 if male and 0 otherwise), educational qualification (Diploma, Degree and Masters where D = 1 if diploma holder and 0 otherwise, M = 1 if qualification is Masters and 0 otherwise), Marital status (where S= 1 if single and 0 if married). Means of transport (Drives company car, Drives own car and Does not drive; where DC = 1 if employee drives company car and 0 otherwise, DO = 1 if employee drives own car and 0 otherwise), accommodation facility (where AC = 1 if employee lives in company provided accommodation and 0 if not). Use the correlation matrix and dummy regression output from *R software* to answer the questions that follow:

	SV	E	C	GM	D	M	S	DC	DO	AC
SV	1.00									
E	0.56	1.00								
C	0.77	0.85	1.00							
GM	-0.24	0.03	0.06	1.00						
D	-0.13	-0.04	-0.02	0.01	1.00					
M	0.30	0.06	0.04	-0.55	0.04	1.00				
S	0.47	0.08	0.08	-0.48	-0.08	0.38	1.00			
DC	0.56	0.06	0.07	0.87	0.02	-0.65	-0.58	1.00		
DO	0.40	-0.01	0.08	-0.07	0.01	0.08	0.07	-0.02	1.00	
AC	0.69	0.00	0.02	-0.65	-0.10	0.80	0.44	-0.75	0.07	1.00

a) From the matrix, identify 2 variables most strongly correlated with the regressand (2 marks)
b) Which 3 pairs of variables are mostly multicollinear? (3 marks)
c) Which 2 pairs of variables are most strongly correlated? (2 marks)
d) How is SV likely to respond if the commission level is increased? Justify (2 marks)
e) Using the normal Q-Q plot below, indicate if sales volume (SV) is normally distributed (2 marks)



Given the regression output below, answer the rest of the questions:

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.96
R Square	0.92
Adjusted R Square	0.92
Standard Error	0.84
Observations	200.00

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	12.00	1619.13	134.93	191.50	0.00
Residual	187.00	131.75	0.70		
Total	199.00	1750.88			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*VIF*
Intercept	12.54	0.44	28.70	0.00	
B	0.08	0.02	4.36	0.00	1.41
C	0.01	0.00	2.71	0.01	1.37
GM	-0.34	0.21	-1.60	0.11	1.03
D	-0.02	0.04	0.47	0.64	1.07
M	-0.13	0.06	2.26	0.03	3.06
S	0.00	0.01	-0.32	0.75	4.09
DC	1.90	0.21	9.04	0.00	3.12
DO	4.95	0.28	17.56	0.20	3.93
AC	2.28	0.20	11.54	0.00	3.58

f) Specify the entire regression model. (3 marks)
g) Estimate the regression equation for the model (3 marks)
h) Is the model reliable? Justify (3 marks)
i) Are the coefficients jointly significant? Justify (3 marks)

j) Using the variance inflation factors (VIF), which variable between the pairs of multicollinear variables may be dropped? Justify (3 marks)

k) Which coefficients are statistically not different from 0? Justify (3 marks)

l) Interpret the coefficient of working experience (3 marks)

m) Interpret the coefficient of Means of transport (Drives company car) (3 marks)

n) Based on the results, should the company provide the employees with accommodation? Justify (3 marks)

o) What is the predicted difference in the Sales volume of a married female employee who has been on the job for 10 years, holds a degree and drives her own car and that of a single male employee with 7 years working experience who has a masters degree and drives company car, if in both cases they are provided accommodation and receive 10% commission on the sales they generate? (7 marks)

p) You have been provided an extract of a questionnaire on entrepreneurial tendencies among workers. For each question identify the data type (categorical or numerical) and the scale of measurement. (5 marks)

i. What is your occupation?.....

ii. How satisfied are you with your current job? [1] very dissatisfied, ..., [5] very satisfied

iii. How much time do we get to spend with our families? .....

iv. What time do you normally close from work?.....

v. Do you feel very busy each day? .....

SOME USEFUL FORMULAE

SOME USEFUL FORMULAE

1. $t_{df} = t_{N-K-1,\alpha}$ 2. $t_{\beta} = \frac{\beta}{S.E.(\beta)}$ 3. $F = \frac{MSR}{MSE}$ 4. $F_{K,N-K-1,\alpha}$

5. $N-1 = TSS_{df}$ 6. $MSR = \frac{RSS}{K}$ 7. $R^2 = \frac{RSS}{TSS}$ 8. $\hat{\beta}_j \pm t_{n-K-1,\alpha} * S.E._{(\beta)}$

9. $\bar{R}^2 = 1-(1-R^2)\left(\frac{N-1}{N-K-1}\right)$ 10. $MSG = \frac{SSG}{K-1}$ 11. $MSW = \frac{SSW}{N-K}$

12. $F = \frac{MSG}{MSW}$ 13. $\bar{\bar{x}} = \frac{\sum_i^K (n_i \bar{x}_i)}{N}$ 14. $SSG = \sum_{i=1}^{K} n_i(\bar{x}_i - \bar{\bar{x}})^2 -$ 15. $t = \frac{\bar{d}}{S_d/\sqrt{n}}$

16. $SSW = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$ 17. $t = \frac{\frac{\sum d_i}{n}}{\frac{S_d}{\sqrt{n}}}$ 18. $t_{df} = t_{N-1,\alpha}$

19. $U_1 = R_1 - \frac{n_1(n_1+1)}{2}$ 20. $t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ 21. $U_2 = R_2 - \frac{n_2(n_2+1)}{2}$

22. $U_1 + U_2 = n_1 \times n_2$ 23. $\sum(R_1 + R_2) = \frac{n(n+1)}{2}$ 24. $U_{n1,n2,\alpha}$

25. $MSE = \frac{ESS}{N-K-1}$ 26. $F_{K-1,N-K,\alpha}$ 27. $\bar{d} = \frac{\sum d}{n}$ 28. $S_d = \sqrt{\frac{\sum(d-\bar{d})^2}{n-1}}$

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Critical Values of the Mann-Whitney U
(One-Tailed Testing)

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
	.01	–	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
	.01	–	–	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	.05	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
	.01	–	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	.05	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
	.01	–	1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	.05	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
	.01	0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	.05	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
	.01	0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9	.05	4	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
	.01	1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10	.05	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
	.01	1	3	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
11	.05	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
	.01	1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12	.05	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
	.01	2	5	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
13	.05	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
	.01	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	.05	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
	.01	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	.05	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
	.01	3	7	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
16	.05	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
	.01	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	.05	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
	.01	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	.05	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
	.01	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	.05	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
	.01	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	.05	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	.01	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Critical Values of the Mann-Whitney U
(Two-Tailed Testing)

n_2	α	n_1 3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	–	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	–	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	–	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	–	–	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	–	–	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	–	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	–	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	–	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	.05	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	.01	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	.05	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
	.01	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	.05	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
	.01	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	56	60
14	.05	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
	.01	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	.05	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
	.01	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	.05	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
	.01	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	.05	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
	.01	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	.05	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
	.01	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	.05	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
	.01	3	7	12	17	22	28	33	39	45	51	56	63	69	74	81	87	93	99
20	.05	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
	.01	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

Table of critical values for the F distribution (for use with ANOVA):

How to use this table:

There are two tables here. The first one gives critical values of F at the p = 0.05 level of significance.
The second table gives critical values of F at the p = 0.01 level of significance.
1. Obtain your F-ratio. This has (x,y) degrees of freedom associated with it.
2. Go along x columns, and down y rows. The point of intersection is your critical F-ratio.
3. If your obtained value of F is equal to or larger than this critical F-value, then your result is significant at that level of probability.
An example: I obtain an F ratio of 3.96 with (2, 24) degrees of freedom.
I go along 2 columns and down 24 rows. The critical value of F is 3.40. My obtained F-ratio is larger than this, and so I conclude that my obtained F-ratio is likely to occur by chance with a p<.05.

Critical values of F for the 0.05 significance level:

	1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.39	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.97	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.97	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.56	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.33	3.47	3.07	2.84	2.69	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.38	2.32	2.28
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.26
25	4.24	3.39	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.17
31	4.16	3.31	2.91	2.68	2.52	2.41	2.32	2.26	2.20	2.15
32	4.15	3.30	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14
33	4.14	3.29	2.89	2.66	2.50	2.39	2.30	2.24	2.18	2.13
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11

86	3.95	3.10	2.71	2.48	2.32	2.21	2.12	2.05	1.99	1.94
87	3.95	3.10	2.71	2.48	2.32	2.21	2.12	2.05	1.99	1.94
88	3.95	3.10	2.71	2.48	2.32	2.20	2.12	2.05	1.99	1.94
89	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
91	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.98	1.94
92	3.95	3.10	2.70	2.47	2.31	2.20	2.11	2.04	1.98	1.94
93	3.94	3.09	2.70	2.47	2.31	2.20	2.11	2.04	1.98	1.93
94	3.94	3.09	2.70	2.47	2.31	2.20	2.11	2.04	1.98	1.93
95	3.94	3.09	2.70	2.47	2.31	2.20	2.11	2.04	1.98	1.93
96	3.94	3.09	2.70	2.47	2.31	2.20	2.11	2.04	1.98	1.93
97	3.94	3.09	2.70	2.47	2.31	2.19	2.11	2.04	1.98	1.93
98	3.94	3.09	2.70	2.47	2.31	2.19	2.10	2.03	1.98	1.93
99	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.98	1.93
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.98	1.93

Critical values of F for the 0.01 significance level:

	1	2	3	4	5	6	7	8	9	10
1	4052.19	4999.52	5403.34	5624.62	5763.65	5858.97	5928.33	5981.10	6022.50	6055.85
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05
6	13.75	10.93	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
14	8.86	6.52	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.90	3.81
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59
18	8.29	6.01	5.09	4.58	4.25	4.02	3.84	3.71	3.60	3.51
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
23	7.88	5.66	4.77	4.26	3.94	3.71	3.54	3.41	3.30	3.21
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09
27	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.15	3.06
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03
29	7.60	5.42	4.54	4.05	3.73	3.50	3.33	3.20	3.09	3.01
30	7.56	5.39	4.51	4.02	3.70	3.47	3.31	3.17	3.07	2.98
31	7.53	5.36	4.48	3.99	3.68	3.45	3.28	3.15	3.04	2.96
32	7.50	5.34	4.46	3.97	3.65	3.43	3.26	3.13	3.02	2.93

33	7.47	5.31	4.44	3.95	3.63	3.41	3.24	3.11	3.00	2.91
34	7.44	5.29	4.42	3.93	3.61	3.39	3.22	3.09	2.98	2.89
35	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96	2.88
36	7.40	5.25	4.38	3.89	3.57	3.35	3.18	3.05	2.95	2.86
37	7.37	5.23	4.36	3.87	3.56	3.33	3.17	3.04	2.93	2.84
38	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.92	2.83
39	7.33	5.19	4.33	3.84	3.53	3.31	3.14	3.01	2.90	2.81
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
41	7.30	5.16	4.30	3.82	3.50	3.28	3.11	2.98	2.88	2.79
42	7.28	5.15	4.29	3.80	3.49	3.27	3.10	2.97	2.86	2.78
43	7.26	5.14	4.27	3.79	3.48	3.25	3.09	2.96	2.85	2.76
44	7.25	5.12	4.26	3.78	3.47	3.24	3.08	2.95	2.84	2.75
45	7.23	5.11	4.25	3.77	3.45	3.23	3.07	2.94	2.83	2.74
46	7.22	5.10	4.24	3.76	3.44	3.22	3.06	2.93	2.82	2.73
47	7.21	5.09	4.23	3.75	3.43	3.21	3.05	2.92	2.81	2.72
48	7.19	5.08	4.22	3.74	3.43	3.20	3.04	2.91	2.80	2.72
49	7.18	5.07	4.21	3.73	3.42	3.20	3.03	2.90	2.79	2.71
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.79	2.70
51	7.16	5.05	4.19	3.71	3.40	3.18	3.01	2.88	2.78	2.69
52	7.15	5.04	4.18	3.70	3.39	3.17	3.01	2.87	2.77	2.68
53	7.14	5.03	4.17	3.70	3.38	3.16	3.00	2.87	2.76	2.68
54	7.13	5.02	4.17	3.69	3.38	3.16	2.99	2.86	2.76	2.67
55	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75	2.66
56	7.11	5.01	4.15	3.67	3.36	3.14	2.98	2.85	2.74	2.66
57	7.10	5.00	4.15	3.67	3.36	3.14	2.97	2.84	2.74	2.65
58	7.09	4.99	4.14	3.66	3.35	3.13	2.97	2.84	2.73	2.64
59	7.09	4.98	4.13	3.66	3.35	3.12	2.96	2.83	2.72	2.64
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
61	7.07	4.97	4.12	3.64	3.33	3.11	2.95	2.82	2.71	2.63
62	7.06	4.97	4.11	3.64	3.33	3.11	2.94	2.81	2.71	2.62
63	7.06	4.96	4.11	3.63	3.32	3.10	2.94	2.81	2.70	2.62
64	7.05	4.95	4.10	3.63	3.32	3.10	2.93	2.80	2.70	2.61
65	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.80	2.69	2.61
66	7.04	4.94	4.09	3.62	3.31	3.09	2.92	2.79	2.69	2.60
67	7.03	4.94	4.09	3.61	3.30	3.08	2.92	2.79	2.68	2.60
68	7.02	4.93	4.08	3.61	3.30	3.08	2.91	2.79	2.68	2.59
69	7.02	4.93	4.08	3.60	3.30	3.08	2.91	2.78	2.68	2.59
70	7.01	4.92	4.07	3.60	3.29	3.07	2.91	2.78	2.67	2.59
71	7.01	4.92	4.07	3.60	3.29	3.07	2.90	2.77	2.67	2.58
72	7.00	4.91	4.07	3.59	3.28	3.06	2.90	2.77	2.66	2.58
73	7.00	4.91	4.06	3.59	3.28	3.06	2.90	2.77	2.66	2.57
74	6.99	4.90	4.06	3.58	3.28	3.06	2.89	2.76	2.66	2.57
75	6.99	4.90	4.05	3.58	3.27	3.05	2.89	2.76	2.65	2.57
76	6.98	4.90	4.05	3.58	3.27	3.05	2.88	2.76	2.65	2.56
77	6.98	4.89	4.05	3.57	3.27	3.05	2.88	2.75	2.65	2.56
78	6.97	4.89	4.04	3.57	3.26	3.04	2.88	2.75	2.64	2.56
79	6.97	4.88	4.04	3.57	3.26	3.04	2.87	2.75	2.64	2.55
80	6.96	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.64	2.55
81	6.96	4.88	4.03	3.56	3.25	3.03	2.87	2.74	2.63	2.55
82	6.95	4.87	4.03	3.56	3.25	3.03	2.87	2.74	2.63	2.55

83	6.95	4.87	4.03	3.55	3.25	3.03	2.86	2.73	2.63	2.54
84	6.95	4.87	4.02	3.55	3.24	3.03	2.86	2.73	2.63	2.54
85	6.94	4.86	4.02	3.55	3.24	3.02	2.86	2.73	2.62	2.54
86	6.94	4.86	4.02	3.55	3.24	3.02	2.85	2.73	2.62	2.53
87	6.94	4.86	4.02	3.54	3.24	3.02	2.85	2.72	2.62	2.53
88	6.93	4.86	4.01	3.54	3.23	3.01	2.85	2.72	2.62	2.53
89	6.93	4.85	4.01	3.54	3.23	3.01	2.85	2.72	2.61	2.53
90	6.93	4.85	4.01	3.54	3.23	3.01	2.85	2.72	2.61	2.52
91	6.92	4.85	4.00	3.53	3.23	3.01	2.84	2.71	2.61	2.52
92	6.92	4.84	4.00	3.53	3.22	3.00	2.84	2.71	2.61	2.52
93	6.92	4.84	4.00	3.53	3.22	3.00	2.84	2.71	2.60	2.52
94	6.91	4.84	4.00	3.53	3.22	3.00	2.84	2.71	2.60	2.52
95	6.91	4.84	4.00	3.52	3.22	3.00	2.83	2.70	2.60	2.51
96	6.91	4.83	3.99	3.52	3.21	3.00	2.83	2.70	2.60	2.51
97	6.90	4.83	3.99	3.52	3.21	2.99	2.83	2.70	2.60	2.51
98	6.90	4.83	3.99	3.52	3.21	2.99	2.83	2.70	2.59	2.51
99	6.90	4.83	3.99	3.52	3.21	2.99	2.83	2.70	2.59	2.51
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50

SOLUTION – RES METH

2022/2023

2022/2023 Question 4

a. From the box plot majority of the data are within the plot with few outliers. From the shaprio test, since the p-value is less than 0.05, we conclude that wage is normally distributed.

b. There were 76 females and 21 males

c. Age and exper (0.98), yes because the p-value (0.000) is less than alpha of 5%

d. Exper and wage (0.09)

e. No, the model is not reliable because the R-Squared (32%) is less than 50%. This means only 32% of the variation in wage is explained by all the independent variables.

f. $wage = \beta_0 + \beta_1 educ + \beta_2 raceW + \beta_3 sexM + \beta_4 hispanicNH + \beta_5 southS + \beta_6 marriedSingle + \beta_7 exper + \beta_8 unionUnion + \beta_9 age + \beta_{10} sectorconst + \beta_{11} sectormanag + \beta_{12} sectormanuf + \beta_{13} sectorother + \beta_{14} sectorprof + \beta_{15} sectorsales + \beta_{16} sectorservice + \varepsilon$

g. $wage = -3.41 + 0.78 educ + 0.79 raceW + 1.87 sexM + 0.60 hispanicNH - 0.64 southS - 0.31 marriedSingle + 0.21 exper + 1.56 unionUnion - 0.12 age + 1.04 sectorconst + 3.30 sectormanag + 0.46 sectormanuf + 0.32 sectorother + 1.98 sectorprof - 0.75 sectorsales - 0.75 sectorservice$

h.

i. SexM
ii. unionUnion
iii. sectormanag
iv. sectorprof

i. The regression estimates are jointly significant because the p-value of the F-Statistic (2.2e-16) is less than alpha (0.05)

j. On average, given race, sex, south, Hispanic, the number of years of work experience, union of a worker, age and sector of a worker, wage is predicted to be 0.31 units less if the worker is single than married

k. I will advise the worker to switch to management sector because the likely increase in wage (3.30) is higher than the likely increase in wage at the professional sector (1.98)

I.

	ESTIMATE	SCENARIO 1	SCENARIO 2
Intercept	-3.41		
educ	0.78	15	17
raceW	0.79	1	0
sexM	1.87	1	0
hispanicNH	0.60	1	0
southS	-0.64	1	1
marriedSingle	-0.31	1	0
exper	0.21	33	29
unionUnion	1.56	0	0
age	-0.12	51	40
sectorconst	1.04	0	0
sectormanag	3.30	0	0
sectormanuf	0.46	1	0
sectorother	0.32	0	0
sectorprof	1.98	0	0
sectorsales	-0.75	0	0
sectorservice	-0.75	0	1
wage		**11.87**	**9.75**

Scenario 1 is 2.12 units than than Scenario 2

Factors driving the

Number of years of work experience

Number of years of education

Age

Sex

Race

2022/2023 Question 5

a.

i. $H_0: m_1 = m_2$

$H_a: m_1 \neq m_2$

ii. $n_1 = 8 \quad n_2 = 7 \quad N = 15 \quad \propto = 0.01 \quad U_{n_1,n_2,\propto}$

from the Mann Whitney U table, $U_{8,7,0.01} = 6$

iii.

m_1	m_2	R_1	R_2
42		1	
67	67	3	3
	67		3
68	68	5.5	5.5
69		7.5	
69		7.5	
70	70	10	10
	70		10
73		12	
	90		13

93	93	14.5	14.5
Σ		61	59

$$Test\ 1: \sum(R_1 + R_2) = \frac{N(N+1)}{2}$$

$$(61 + 59) = \frac{15(15+1)}{2}$$

$$120 = 120$$

$$U_i = R_i - \frac{n_i(n_i+1)}{2}$$

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \equiv 61 - \frac{8(8+1)}{2} = 25$$

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} \equiv 59 - \frac{7(7+1)}{2} = 31$$

$$\therefore the\ test\ statistis\ is\ 25$$

$$Test\ 2: U_1 + U_2 = n_1 \times n_2$$

$$25 + 31 = 8 \times 7$$

$$56 = 56$$

iv. Since the calculated (25) is greater than the critical value (6) we fail to reject the null hypothesis and conclude that there is no statistically significant difference in the measure of heart rate and conclude that both regular and decaffeinated coffee affects heart rate

v. Wilcoxon signed-rank test

vi. When the test statistic is less than the critical value

vii. The observed difference was due to chance since we failed to reject the null hypothesis

b.

i. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

$H_a: \mu_i \neq \mu_j\ where\ ij = 1,2,3,4,5$

ii. $n_1 = 5 \quad n_2 = 3 \quad n_3 = 4 \quad n_4 = 2 \quad n_5 = 5 \quad N = 19 \quad k = 5$

$df_{BG} = k - 1 \equiv 5 - 1 = 4 \quad df_{WG} = N - k \equiv 19 - 5 = 14$

$F_{\propto,k-1,N-k} \equiv F_{0.05,4,14} = 3.11$

$\bar{x}_1 = \frac{8+7+10+6+9}{5} = 8 \quad \bar{x}_2 = \frac{9+7+8}{3} = 8 \quad \bar{x}_3 = \frac{5+8+10+5}{4} = 7$

$\bar{x}_4 = 4.5 \quad \bar{x}_5 = 5.4$

$\bar{\bar{x}} = \frac{8+7+10+6+9+9+7+8+5+8+10+5+5+4+4+9+6+3+5}{19} = 6.74$

$$SSG = \sum n_i(\bar{x}_i - \bar{\bar{x}})^2$$

$$SSG = 5(8-6.74)^2 + 3(8-6.74)^2 + 4(7-6.74)^2 + 2(4.5-6.74)^2 + 5(5.4-6.74)^2 = 31.98$$

$$SSW = \sum\sum (x_{ij} - \bar{x}_i)^2$$

$$SSW_1 = (8-8)^2 + (7-8)^2 + (10-8)^2 + (6-8)^2 + (9-8)^2 = 10$$
$$SSW_2 = (9-8)^2 + (7-8)^2 + (8-8)^2 = 2$$
$$SSW_3 = (5-7)^2 + (8-7)^2 + (10-7)^2 + (5-7)^2 = 18$$
$$SSW_4 = (5-4.5)^2 + (4-4.5)^2 = 0.5$$
$$SSW_5 = (4-5.4)^2 + (9-5.4)^2 + (6-5.4)^2 + (3-5.4)^2 + (5-5.4)^2 = 21.2$$
$$SSW = 10 + 2 + 18 + 0.5 + 21.2 = 51.70$$

$$MS_{BG} = \frac{SSG}{k-1} \equiv \frac{31.98}{4} = 7.995$$

$$MS_{WG} = \frac{SSW}{N-k} \equiv \frac{51.70}{14} = 3.693$$

$$F = \frac{msg}{msw} = \frac{7.995}{3.693} = 2.16$$

Since the F-calculated (2.16) is less than the critical value (3.34), we fail to reject the null hypothesis and conclude there is statistically no significant difference between the mean salaries of the five government workers at 5% significance level

iii. Kruskal-Wallis test
iv. Dependent paired t-test
v. Police since they have the least mean salary of GH₵4,500

c.

1. Numerical and interval
2. Categorical and nominal
3. Categorical and ordinal
4. Numerical and interval

2021/2022 Question 4

a. S and A is positive.
OV and E is negative

b. H and M (0.89), delete M because it has the highest VIF (60.36)
R and A (0.83), delete R because it has the highest VIF (13.44)
P and M (0.79), delete ED because it has the highest VIF (60.36)

c. $S = \beta_0 + \beta_1 B + \beta_2 SE + \beta_3 M + \beta_4 P + \beta_5 A + \beta_6 H + \beta_7 OV + \beta_8 E + \beta_9 R + \beta_{10} MA + \beta_{11} J + \beta_{12} U + \beta_{13} SB + \beta_{14} F + \beta_{15} SI + \varepsilon$

d. $S = 30.13986 + 2.16807B - 0.81958SE + 0.03583M - 0.06037P + 1.20419A + 4.65867H + 0.43395OV + 0.33234E - 0.89650R - 2.96365MA - 5.59275J - 5.43629U - 2.76403SB - 3.27243F - 7.02194SI$

e. The coefficients are jointly significant because the p-value (0.00004935) of the F-statistic is less than 5%

f. Yes, because Multiple R-Squared (88.99%) is more than 50%

g. 14 ($B, SE, M, P, A, V, E, R, MA, J, U, B, F$ and SI)

h. On average holding age of the workers, number of hours' workers worked in a day, number of overtime hours, years of experience, nature of job, type of job, position held, location of job, gender of worker, and the marital status constant, Salary is expected to be 0.81958 units less if level of education is secondary than degree.

i. Master (-0.85)

j.

	ESTIMATE	SCENARIO 1	SCENARIO 2
Intercept	30.13986		
B	2.16807	0	0
SE	-0.81958	1	0
M	0.03583	0	1
P	-0.06037	0	0
A	1.20419	30	36
H	4.65867	6	4
OV	0.43395	4	0
E	0.33234	8	12
R	-0.89650	1	0
MA	-2.96365	1	0
J	-5.59275	1	0
U	-5.43629	0	1
SB	-2.76403	0	0
F	-3.27243	0	1
SI	-7.02194	1	0
S		81.32	87.44

Scenario 1 is 6.12 less than Scenario 2 and the categorical factors driving the difference are; educational level, nature of job, type of job, position held, location of job, gender and marital status

k. Error term, unobserved heterogeneity, stochastic noise, idiosyncratic error

2021/2022 Question 5

i. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

$H_a: \mu_i \neq \mu_j \; where \; ij = 1,2,3,4$

ii. $n_1 = 6 \quad n_2 = 5 \quad n_3 = 4 \quad n_4 = 5 \quad N = 20 \quad k = 4$

$df_{BG} = k - 1 \equiv 4 - 1 = 3 \quad df_{WG} = N - k \equiv 20 - 4 = 16$

$F_{\alpha,k-1,N-k} = F_{0.05,3,16} = 3.24$

$\bar{x}_1 = \frac{58+69+60+65+71+74}{6} = 66.17 \quad \bar{x}_2 = \frac{67+81+70+59+61}{5} = 67.60$

$\bar{x}_3 = \frac{49+52+48+53}{4} = 50.50 \quad \bar{x}_4 = \frac{55+41+49+57+58}{5} = 52$

$$\bar{\bar{x}} = \frac{58+69+60+65+71+74+67+81+70+59+61+49+52+48+53+55+41+49+57+58}{20} = 59.85$$

$$SSG = \sum n_i(\bar{x}_i - \bar{\bar{x}})^2$$

$$SSG = 6(66.17 - 59.85)^2 + 5(67.60 - 59.85)^2 + 4(50.50 - 59.85)^2 + 5(52 - 59.85)^2 = 1{,}197.52$$

$$SSW = \sum\sum (x_{ij} - \bar{x}_i)^2$$

$$SSW_1 = (58 - 66.17)^2 + (69 - 66.17)^2 + (60 - 66.17)^2 + (65 - 66.17)^2 + (71 - 66.17)^2 + (74 - 66.17)^2 = 198.83$$

$$SSW_2 = (67 - 67.60)^2 + (81 - 67.60)^2 + (70 - 67.60)^2 + (59 - 67.60)^2 + (61 - 67.60)^2 = 303.20$$

$$SSW_3 = (49 - 50.50)^2 + (52 - 50.50)^2 + (48 - 50.50)^2 + (53 - 50.50)^2 = 17$$

$$SSW_4 = (55 - 52)^2 + (41 - 52)^2 + (49 - 52)^2 + (57 - 52)^2 + (58 - 52)^2 = 200$$

$$SSW = 198.83 + 303.20 + 17 + 200 = 719.03$$

$$MS_{BG} = \frac{SSG}{k-1} \equiv \frac{1{,}197.52}{3} = 399.17$$

$$MS_{WG} = \frac{SSW}{N-k} \equiv \frac{719.03}{16} = 44.94$$

$$F = \frac{msg}{msw} = \frac{399.17}{44.94} = 8.88$$

Since F-Test (8.88) > F-Critical (3.24), we reject the null hypothesis at 5% significance level and conclude that is there is significant statistical evidence to suggest the mean salary across the four degree courses are different

iii. Wilcoxon and Kruskal Wallis

iv. History
Nursing
Economics
Medicine

2021/2022 Question 6

i. $H_0: m_1 = m_2$
$H_a: m_1 \neq m_2$

ii. $n_1 = 8 \quad n_2 = 7 \quad N = 15 \quad \propto = 0.01 \quad U_{n_1,n_2,\propto}$
$from\ the\ Mann\ Whitney\ U\ table, U_{8,7,0.01} = 6$

iii.

m_1	m_2	R_1	R_2
42		1	
67	67	3	3

	67		3
68	68	5.5	5.5
69		7.5	
69		7.5	
70	70	10	10
	70		10
73		12	
	90		13
93	93	14.5	14.5
$\sum$		61	59

$$Test\ 1:\ \sum(R_1 + R_2) = \frac{N(N+1)}{2}$$

$$(61 + 59) = \frac{15(15+1)}{2}$$

$$120 = 120$$

$$U_i = R_i - \frac{n_i(n_i+1)}{2}$$

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \equiv 61 - \frac{8(8+1)}{2} = 25$$

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} \equiv 59 - \frac{7(7+1)}{2} = 31$$

$$\therefore the\ test\ statistis\ is\ 25$$

$$Test\ 2:\ U_1 + U_2 = n_1 \times n_2$$

$$25 + 31 = 8 \times 7$$

$$56 = 56$$

iv. Since the calculated (25) is greater than the critical value (6) we fail to reject the null hypothesis and conclude that there is no statistically significant difference in the measure of heart rate and conclude that both regular and decaffeinated coffee affects heart rate

v. Wilcoxon signed-rank test

2020/2021 Question 2 - SDG

a. POP and C02 is positive.
INF and ED is negative

b. TECH and CMPI -0.09 and EAC and CMPI 0.06
They are significant because their p-values are less than 5%

c. C02 and POP (0.62)
ECCAS and OPEC (0.53)
ED and TECH (0.49)
They are significant because their p-values are less than 5%

d. C02 and POP (0.62) Yes because the p-value is less than 5%

e. TECH (-0.09)

f. $CMPI = \beta_0 + \beta_1 POP + \beta_2 CO2 + \beta_3 TECH + \beta_4 INF + \beta_5 ED + \beta_6 OPEC + \beta_7 EAC + \beta_8 ECCAS + \beta_9 ECOWAS + \beta_{10} SDAC + \varepsilon$

g.
1. Level of literacy
2. Level of employment

h. $CMPI = 0.9235 + 0.0366POP - 0.0279CO2 - 0.0352TECH + 0.000INF + 0.000ED + 0.0283OPEC - 0.0331EAC - 0.0099ECCAS - 0.0093ECOWAS + 0.0543SDAC$

i. Yes, because the p-value of the F-statistics (0.0000) is less than 5%

j. Between C02 and POP (0.62), drop POP because it has the highest VIF (6.40)
Between ECCAS and OPEC (0.53), drop ECCAS because it has the highest VIF (3.26)
Between ED and TECH (0.49), drop ED because it has the highest VIF (1.92)

k. No, because the R-Square (0.03069) is less than 50%

l. Any time BLOC affiliation EAC changes, CMPI falls by 0.0331 units

m. On average holding POP, TECH, INF, ED, and state affiliation constant, a one-unit increase in C02 will lead to a 0.0279-unit decrease in CMPI

n. Heteroseedasticity talks about the normality of data while Multicollinearity talks about the relationship between independent variable. The impact is the outcome is skewed or biased

o. Attach(energy)
Shapiro.test(CMPI)

p. $CMPI = 0.9235 + 0.0366POP - 0.0279CO2 - 0.0352TECH + 0.000INF + 0.000ED + 0.0283OPEC - 0.0331EAC - 0.0099ECCAS - 0.0093ECOWAS + 0.0543SDAC$

	ESTIMATE	SCENARIO 1	SCENARIO 2
Intercept	0.9235		
POP	0.0366	4,000,000	4,000,000
C02	-0.0279	300	200
TECH	-0.0352	40	100
INF	0.0000	0.07	0.17
ED	0.0000	36	60
OPEC	0.0283	0	1
BLOCEAC	-0.0331	0	0
BLOCECCAS	-0.0099	0	0
BLOCECOWAS	-0.0093	0	1
BLOCSADC	0.0543	0	0
CMPI		**146,391.15**	**146,391.84**

Scenario 2 is 0.69 higher than Scenario 1

Factors driving the difference are, C02, TECH, INF, ED and State affiliation

q. It has no economic significance

2020/2021 Question 3

a.

i. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$

$H_a: \mu_i \neq \mu_j \; where \; ij = 1,2,3,4,5,6$

ii.

$n_1 = 7 \quad n_2 = 6 \quad n_3 = 6 \quad n_4 = 5 \quad n_5 = 4 \quad n_6 = 3 \quad N = 31 \quad k = 6$

$df_{BG} = k - 1 \equiv 6 - 1 = 5 \quad df_{WG} = N - k \equiv 31 - 6 = 25$

$F_{\alpha,k-1,N-k} \equiv F_{0.05,5,31} = 2.52$

$\bar{x}_1 = \frac{4+7+3+3+8+5+7}{7} = 5.29 \quad \bar{x}_2 = \frac{6+7+6+7+6+8}{6} = 6.67 \quad \bar{x}_3 = \frac{5+6+8+4+8+8}{6} = 6.5$

$\bar{x}_4 = 8.8 \quad \bar{x}_5 = 9 \quad \bar{x}_6 = 12$

$\bar{\bar{x}} = \frac{4+7+3+3+8+5+7+6+7+6+7+6+8+5+6+8+4+8+8+10+9+8+8+9+10+7+10+9+12+13+11}{31} = 7.48$

$$SSG = \sum n_i(\bar{x}_i - \bar{\bar{x}})^2$$

$SSG = 7(5.29 - 7.48)^2 + 6(6.67 - 7.48)^2 + 6(6.5 - 7.48)^2 + 5(8.8 - 7.48)^2 + 4(9 - 7.48)^2 + 3(12 - 7.48)^2 = 122.52$

$$SSW = \sum\sum (x_{ij} - \bar{x}_i)^2$$

$SSW_1 = (4 - 5.29)^2 + (7 - 5.29)^2 + (3 - 5.29)^2 + (3 - 5.29)^2 + (8 - 5.29)^2 + (5 - 5.29)^2 + (7 - 5.29)^2 = 20.18$

$SSW_2 = (6 - 6.67)^2 + (7 - 6.67)^2 + (6 - 6.67)^2 + (7 - 6.67)^2 + (6 - 6.67)^2 + (8 - 6.67)^2 = 3.33$

$SSW_3 = (5 - 6.5)^2 + (6 - 6.5)^2 + (8 - 6.5)^2 + (4 - 6.5)^2 + (8 - 6.5)^2 + (8 - 6.5)^2 = 15.50$

$SSW_4 = (10 - 8.8)^2 + (9 - 8.8)^2 + (8 - 8.8)^2 + (8 - 8.8)^2 + (9 - 8.8)^2 = 2.8$

$SSW_5 = (10 - 9)^2 + (7 - 9)^2 + (10 - 9)^2 + (9 - 9)^2 = 6$

$SSW_6 = (12 - 12)^2 + (13 - 12)^2 + (11 - 12)^2 = 2$

$$SSW = 20.18 + 3.33 + 15.5 + 2.8 + 6 + 2 = 49.81$$

$$MS_{BG} = \frac{SSG}{k-1} \equiv \frac{122.52}{5} = 24.50$$

$$MS_{WG} = \frac{SSW}{N-k} \equiv \frac{49.81}{25} = 1.99$$

$$F = \frac{msg}{msw} = \frac{24.50}{1.99} = 22.51$$

Since the F-calculated (22.51) is greater than the critical value (2.52), we reject the null hypothesis and conclude there is statistically significant difference between the mean ticket price of the cabin classes at 5% significance level

iii. Kruskal-Wallis

iv.

1. Categorical grouping
2. Normality
3. Independent random sampling
4. Homogeneity of variance for dependent variable

v. Presidential,
First class
Royals
Premium
Business
Economy

b.

i. $H_0: m_1 \leq m_2$

$H_a: m_1 > m_2$

ii. $n_1 = 11 \qquad n_2 = 9 \qquad N = 20 \qquad \propto = 0.05 \qquad U_{n_1,n_2,\propto}$

$from\ the\ Mann\ Whitney\ U\ table, U_{11,9,0.05} = 27$

iii.

m_1	m_2	R_1	R_2
	2		1
	4		3
	4		3
	4		3
5	5	6.5	6.5
5		6.5	
5		6.5	
6	6	10	10
6		10	
7	7	13.5	13.5
7	7	13.5	13.5
8		16	
9	9	18	18
9		18	
10		20	
Σ		138.5	71.5

$$Test\ 1: \sum (R_1 + R_2) = \frac{N(N+1)}{2}$$

$$(138.5 + 71.5) = \frac{20(20+1)}{2}$$

$$210 = 210$$

$$U_i = R_i - \frac{n_i(n_i+1)}{2}$$

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \equiv 138.5 - \frac{11(11+1)}{2} = 72.50$$

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} \equiv 71.5 - \frac{9(9+1)}{2} = 26.50$$

$\therefore$ *the test statistis is* 26.5

$$Test\ 2: U_1 + U_2 = n_1 \times n_2$$
$$72.5 + 26.50 = 11 \times 9$$
$$99 = 99$$

iv. Since the calculated (26.50) is less than the critical value (27) we reject the null hypothesis at 5% level of significance and conclude that there is statistically significant evidence that the kilogram of biomass of male Juniper trees exceeds that of the female.

v. The observed difference was not due to chance since we rejected the null hypothesis

vi. Binomial

vii. Random sampling and mutual exclusivity

May 2019

May 2019 Question 2

a.

i. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$
$H_a: \mu_i \neq \mu_j\ where\ ij = 1,2,3,4,5$

ii. $n_1 = 6 \quad n_2 = 4 \quad n_3 = 7 \quad n_4 = 4 \quad n_5 = 6 \quad N = 27 \quad k = 5$
$df_{BG} = k - 1 \equiv 5 - 1 = 4 \quad df_{WG} = N - k \equiv 27 - 5 = 22$

iii. $F_{\propto,k-1,N-k} \equiv F_{0.05,4,22} = 2.82$

iv. $\bar{x}_1 = \frac{19+17+16+19+17+16}{6} = 17.33 \qquad \bar{x}_2 = \frac{19+16+16+30}{4} = 20.25$

$\bar{x}_3 = \frac{15+20+18+11+31+32+33}{7} = 22.86 \qquad \bar{x}_4 = 28.5 \qquad \bar{x}_5 = 24.50$

$\bar{\bar{x}} = \frac{19+17+16+19+17+16+19+16+16+30+15+20+18+11+31+32+33+18+31+32+33+17+17+17+31+32+33}{27} =$

22.44

v. ANOVA

vi. Categorical Groupings
Independent variables across groups

b.

i. $H_0: m_1 = m_2$
$H_a: m_1 \neq m_2$

ii. $n_1 = 8 \qquad n_2 = 8 \qquad N = 16 \qquad \propto = 0.01 \qquad U_{n_1,n_2,\propto}$

from the Mann Whitney U table, $U_{8,8,0.01} = 7$

iii.

iv.

m_1	m_2	R_1	R_2
	73		1
74	74	2.5	2.5
	76		4
77	77	5.5	5.5
79		7	
80	80	10	10
80	80	10	10
	80		10
83	83	13.5	13.5
91		15	
94		16	
$\sum$		79.50	56.50

$$Test\ 1: \sum (R_1 + R_2) = \frac{N(N+1)}{2}$$

$$(79.50 + 56.50) = \frac{16(16+1)}{2}$$

$$136 = 136$$

$$U_i = R_i - \frac{n_i(n_i+1)}{2}$$

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \equiv 79.50 - \frac{8(8+1)}{2} = 43.5$$

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} \equiv 56.50 - \frac{8(8+1)}{2} = 20.5$$

$$\therefore the\ test\ statistis\ is\ 20.5$$

$$Test\ 2: U_1 + U_2 = n_1 \times n_2$$

$$43.50 + 20.5 = 8 \times 8$$

$$64 = 64$$

Since the test statistics (20.50) is greater than the critical value (7) at 1% significance level, we fail to reject the null hypothesis and conclude that there is statistically significant evidence to back the claims that both methods are equally effective

v. Kruskal-Wallis

May 2019 Question 3

a. C(0.77) & AC(0.69)

b. DC and GM (0.87)
C and E (0.85)
AC and DC (-0.75)

c. DC and GM (0.87)
C and E (0.85)

d. The sales volume will increase by 0.77 units per each unit increase in Commission

e. Sales volume is normally distributed since most of the data points are on the best fit line

f. $SV = \beta_0 + \beta_1 E + \beta_2 C + \beta_3 GM + \beta_4 D + \beta_5 M + \beta_6 S + \beta_7 DC + \beta_8 DO + \beta_9 AC + \varepsilon$

g. $SV = 12.54 + 0.08E + 0.01C - 0.34GM - 0.02D + 0.13M + 0.00S + 1.90DC + 4.95DO + 2.28AC$

h. Yes, because the R-Square (0.92) is greater than 50%

i. Yes, because the p-value of the F-statistics (0.00) is less than 5%

j. DC and GM (0.87), drop DC because it has the highest VIF (3.12)
C and E (0.85), drop E because it has the highest VIF (1.41)
AC and DC (-0.75), drop AC because it has the highest VIF (3.58)

k. GM, D, S & DO because their p-values are more than 5%

l. On average, given the level of commission, gender, educational qualification, marital status, means of transport and accommodation facility of an employee, a one-unit increase in the years of working experience will increase their sales volume by 0.08 units.

m. On average, given the years of working experience, level of commission, gender, educational qualification, marital status and accommodation facility of an employee, sales volume is predicted to be 1.90 units more if the employee drives company car than when the employee does not drive.

n. Yes, because workers in company provided accommodation less to 2.28 units in sales volume

o.

	ESTIMATE	SCENARIO 1	SCENARIO 2
Intercept	12.54		
E	0.08	10	7
C	0.01	0.1	0.1
GM	-0.34	0	1
D	-0.02	1	0
M	0.13	0	1
S	0.00	0	1
DC	1.90	0	1
DO	4.95	1	0
AC	2.28	1	1
SV		**20.55**	**17.07**

Scenario 2 is 3.48 less than Scenario 1

p.

i. Categorical and nominal
ii. Categorical and ordinal
iii. Numeric and interval
iv. Numeric and interval
v. Categorical and nominal

MARKETING MANAGEMENT
UGBS 614
QUESTIONS FROM 2019 TO 2023