MAASAI MARA UNIVERSITY

REGULAR UNIVERSITY EXAMINATIONS
2022/2023 ACADEMIC YEAR
FOURTHYEAR FIRSTSEMESTER

SCHOOL OF BUSINESS AND ECONOMICS
BACHELOR OF SCIENCES IN FINANCIAL ECONOMICS),
BACHELOR OF SCIENCES IN ECONOMICS AND
STATISTICS, BACHELOR OF SCIENCES IN ECONOMICS,
BACHELOR OF AGRICULTURAL ECONOMICS, AND
BACHELOR OF ARTS IN AGRICULTURAL ECONOMICS

COURSE CODE: ECO 4106

COURSE TITLE: DATA ANALYSIS AND COMPUTER APPLICATIONS IN ECONOMICS

DATE: 8TH DECEMBER, 2022   TIME: 1100-1300

QUESTION 1

a) Food consumption score (FCS) is a household food security indicator that considers how diverse the household diet is and how frequently different food groups are taken by households. Households with FCS of 0-21 are considered to have a **Poor FCS**. Those with 21.5-35 are considered to have **Borderline FCS**, while those with FCS of 35.5 and above are considered to have **Acceptable FCS**. The World Bank commissioned a study in the Republic of South Sudan to assess the FCS for male-headed and female-headed households and summarized their measurements of FCS in bar charts as presented in Annex 1. Use the presentation to answer question 1 (a)(i)-iii)

i) Al Renk, Malut, Aryat, Bor, Tonga, and KajoKaii are classified as low-conflict districts. How many low-conflict districts have female-headed households with Acceptable FCS? (2 marks)

ii) List the districts in the country with Borderline FCS for the male-headed households (2 marks)

iii) For the districts that are not considered as low-conflict, what is the difference in FCS between the male- and female-headed households with the lowest scores? (1 mark)

b) The data summary below is the GNP per capita for 63 randomly picked countriesin Africa and Northern Asia. Use the output to answer the subsequent questions.

```
     Variable |        Obs       Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       gnppc |         63    8674.857    10634.68        370      39980

. sugnppc, detail

                         GNP per capita
-------------------------------------------------------------
      Percentiles      Smallest
 1%          370            370
 5%          410            380
10%          740            380       Obs                   63
25%         1360            410       Sum of Wgt.           63

50%         3360                      Mean            8674.857
                         Largest      Std. Dev.       10634.68
75%        14100          29240
90%        25580          33040       Variance        1.130008
95%        29240          34310       Skewness         1.30502
99%        39980          39980       Kurtosis        3.382168
```

i) Given the kurtosis and skewness coefficients, is the data from where the summary was made normally distributed? (Give detailed reasons for your response) (4 marks)

ii) Construct a 95 percent confidence interval for the mean GNP per capita. (3 marks)

iii) What is the GNP per capita range for the 63 countries? (1 mark)

c) With examples, explain the following terms:

i) A categorical variable (2 marks)

ii) Univariate regression (2 marks)

iii) Secondary data (2 marks)

d) Give any (3) benefits that a researcher gets if he opts to collect primary data as opposed to secondary? (6 marks)

QUESTION 2

a) The World Bank is about to finance the construction of 124 water reservoirs (dams and boreholes) in the arid and semi-arid lands (ASAL) counties in Kenya. The reservoirs will be of different capacities ranging from 12,029 m³ to 18,324 m³. After discussions with the beneficiary communities, the water needs assessment has been done (in m³) and summarized in a stem-and-leaf display below. Use the information to answer the questions below.

```
12*** | 029,036,057,085,105,139,165,183,187
12*** | 253,266,278,290,313,326,345,359,381
12*** | 429,445,482,549,572,580,595
12*** | 604,604,629,638,668,668,677,760,766
12*** | 807,808,880,898,903,940,949,949,953,965,974
13*** | 017,068,090,124,142,163,168,173,174,185
13*** | 273,302,314,334,362,374,387
13*** | 403,422,435,436,452,463,491,496,501,525,540,556,560,597
13*** | 609,644,649,659,672,680,701,757
13*** | 841,845,861,940,974
14*** | 053,078,112,113
14*** | 237,243,256,308
14*** | 434,450,497,522,545
14*** | 670
14*** | 849,868,879,905
15*** | 115,175,191
15*** | 436,468
16*** | 091,138
16*** | 315,356
16*** | 504,506
16*** | 940
17*** | 046
17*** | 240
17*** | 466
17*** | 653
18*** | 324
```

i) It takes (on average) $10 to generate one cubic meter of water. What will be the expenditure on construction of sites with capacities above 16,400 m³. (2 marks)

ii) According to the beneficiary needs assessment, water reservoirs with capacities above 14,600 will support irrigation activities. How many of the reservoirs will (upon completion) support irrigation activities? (2 marks)

iii) You are given that there are a total of 12,800 beneficiaries in the water reservoirs with capacities between 12,800 m³ and 12,900 m³. What is the amount of water availed per beneficiary (in m³/beneficiary).? (2 marks)

iv) State one demerit of using stem-and-leaf displays in presenting data (1 mark)

b) The statement of the null and its alternative hypothesis requires there is no possibility of creating a competing hypothesis other than the two. Look at the following sets of hypotheses and state if they are correctly stated by writing "Correctly stated" or "Incorrectly stated". Give a reason for your response.

i) $H_0: \mu = 10, H_a: \mu = 20$ (2 marks)

ii) $H_0: \mu = 0.8, H_a: \mu > 0.2$ (2 marks)

iii) $H_0: \mu = 12.3, H_a: \mu < 13.2$ (2 marks)

iv) $H_0: \mu = 0.2, H_a: \mu = 18$ (2 marks)

QUESTION 3

a) BOMA, an international NGO operating in Kalobeyei and Kakuma refugee settlements in Turkana County, is driving a project called Mama Pika Initiative (MPI). MPI promotes the substitution of carbon-dense energy sources with green solutions. It conducted a survey to understand the cooking duration per household (in hours) per week between households that are living in the camp (1) and those living off the camp (0). Data was collected from 236 households (81 off-the-camp and 155 in-the-camp). A two-sample test was done to understand if there is a difference in the cooking hours per week between the two populations. The results are given as follows:

```
Two-sample t test with equal variances
-----------------------------------------------------------------------------
  Group |     Obs        Mean    Std. Err.    Std. Dev.    [95% Conf. Interval]
--------+--------------------------------------------------------------------
      0 |      81    3.703704    .334809      3.013281
      1 |     155    5.406452                 3.645308    4.828032    5.984871
--------+--------------------------------------------------------------------
combined|     236    4.822034    .229733      3.529225    4.369435    5.274633
--------+--------------------------------------------------------------------
   diff |           -1.702748    .4719518                -2.632565   -.7729304
-----------------------------------------------------------------------------
   diff = mean(0) - mean(1)                                  t =   -3.6079
Ho: diff = 0                                   degrees of freedom =       234

   Ha: diff < 0                  Ha: diff != 0                  Ha: diff > 0
Pr(T < t) = 0.0002       Pr(|T| > |t|) = 0.0004         Pr(T > t) = 0.9998
```

i)   From the output above, state the null hypothesis   (1 mark)

ii) Interpret the two means (for the in-the-camp and off-the-camp living households).   (2 marks)

iii) The confidence interval for the mean weekly cooking hours for the off-the-camp dwellers is deliberately not given. Compute it.   (2 marks)

iv) Compute the standard error of the mean weekly cooking hours for the in-the-camp dwellers.   (1 mark)

v)   What conclusions can you draw with regard to the difference in the mean weekly cooking hours between the two populations.   (1 marks)

b) Assume that you were leading the research in the BOMA case in (a) above. Explain any 2 factors that you would probe in an attempt to explain the difference in the cooking hours.   (4 marks)

c) Highlight any 2 challenges that you may experience in conducting research in a refugee camp.   (4 marks)

QUESTION 4

a) JP Anderson is motor company that sells excavators. The pricing of an excavator is based on three variables: tons of sand displaced per liter of gasoline (dpl), chassis length (chl) in inches, and bucket capacity (bcp) in cubic centimeters. A market analyst run ordinary regression of the excavator price (in dollars) on these three main determinants of price and found the following:

```
. reg price dplchlbcp

      Source |       SS           df       MS        Number of obs   =        74
-------------+----------------------------------      F(3, 70)        =     12.14
       Model |   217367689         3   72455896.3     Prob > F        =    0.0000
```

4

```
     Residual |  417697707         70   5967110.1   R-squared      =      0.3423
-------------+--------------------------------       Adj R-squared  =      0.3141
        Total |  635065396         73   8699525.97   Root MSE       =      2442.8


--------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+------------------------------------------------------------------
dpl|  -139.0814    82.20966    -1.69   0.095    -303.0434     24.88062
chl |   59.61193   23.90525     2.49   0.015     11.93442    107.2894
bcp |   2644.771   761.8912     3.47   0.001    1125.227    4164.315
       _cons |  -2861.984    6026.6    -0.47   0.636    -14881.66     9157.69
--------------------------------------------------------------------------------
```

i)   From the regression table above, what percent of variations in the excavator   (1 mark)
     price are explained by *dpl*, *chl*, and *bcp*?

ii) Interpret the three coefficients for *dpl*, *chl*, and *bcp*.   (3 marks)

iii) Sketch the normal curve for the *dpl* variable. Label the coefficient and the   (3 marks)
     lower and upper confidence intervals.

iv) Would a parallelly done research that gives the coefficients of *dpl*, *chl*, and   (3 marks)
     *bcp*as 100, 90, and 1280 respectively be contradicting the outcome of the
     above analysis?

b)   Sketch and label the following types of kurtoses and skewness:

   i)   Leptokurtic, mesokurtic, and platykurtic distribution (in one diagram)   (3 marks)

   ii)   Right-skewed and left-skewed distribution (in one diagram)   (2 marks)

QUESTION 5

a)   You are planning to conduct a study in Kericho to establish the preferences   (10 marks)
     of poultry farmers of a certain poultry breed. Explain the five (5) steps you
     would take to conduct this research.

b)   In this study, as you are stepping out to conduct the research, list any 5   (5 marks)
     variables that you would closely probe as determinants of the farmers'
     preferences of a given poultry breed.

///END///

ANNEXES

Annex 1: Food consumption scores for 22 districts in the Republic of South Sudan (Male-headed households and Female-headed households)

Male-headed households	Female-headed households

District	Male-headed score	District	Female-headed score
Al Renk	41.0	Al Renk	37.4
Magwi	39.5	Magwi	36.0
Twi/Twic East &..	34.4	Twi/Twic..	35.7
Tonj	33.5	Tonj	34.8
Mundri	32.1	Mundri	34.1
Sobat	31.5	Sobat	30.1
Shokodom	29.7	Shokodom	28.3
Malut	24.7	Malut	27.0
Aryat	24.3	Aryat	26.3
Al Mayom	23.4	Al Mayom	16.5
Kapoeta	22.1	Kapoeta	15.0
Raja	17.8	Raja	13.8
Bor	17.6	Bor	7.2
Nahr Jur	17.2	Nahr Jur	6.5
Terkaka	14.7	Terkaka	6.3
Tonga	13.4	Tonga	5.2
Kajo Kaii	8.7	Kajo Kaii	4.7
Juba	8.3	Juba	4.4
Tombura	7.4	Tombura	4.2
Rabkona	6.9	Rabkona	3.6
Aweil	6.9	Aweil	3.4
Gogrial	6.7	Gogrial	3.2