REGULAR UNIVERSITY EXAMINATIONS
2018/2019 ACADEMIC YEAR
SECOND **YEAR** *FIRST* **SEMESTER**

SCHOOL OF SCIENCE
BACHELOR OF SCIENCE IN APPLIED STATISTICS

COURSE CODE: STA 2220

COURSE TITLE: COMPUTATIONAL METHODS AND DATA ANALYSIS II

DATE: APRIL 2019   **TIME:**

INSTRUCTIONS TO CANDIDATES

1. Answer Question **ONE** and any other **TWO** questions
2. Show all your working and be neat
3. Do not write on the question paper

*This paper consists of **THREE** printed pages. Please turn over.*

QUESTION ONE (30 MARKS)

1. Define the following terms as used in R
    i.   A data frame   (1 marks)
    ii.   A matrix   (1marks)
    iii. A Character object   (1 marks)

2. The following is an example of a matrix with 2 rows and 3 columns $A = \begin{bmatrix} 2 & 4 & 3 \\ 1 & 5 & 7 \end{bmatrix}$

    Reproduce a memory representation of the matrix in R with the matrix function.

                                                                     (4 marks)
3. State and explain advantages of a spreadsheet over a hand calculator   (6marks)
4. Data may be entered into the cells of a worksheet in many types. State and Explain
                                                                     (5 marks)
5. State and explain the data editing and formatting features found in excel system
                                                                     (6marks)
6. In the data set painters, the pie chart of the School variable is a collection of pizza wedges showing the proportion of painters in each school.
    i.   Use R in applying the table function to produce the frequency distribution of School.   (4marks)
    ii.   With an extension of the program, write the program to colorize the pie chart

                                                                     (2 marks)

QUESTION TWO (20 MARKS)

1. If there are twelve cars crossing a bridge per minute on average
    i.   Write R program to find the probability of having seventeen or more cars crossing the bridge in a particular minute.   (5 marks)
    ii.   Find the percentage probability of having seventeen or more car crossing the bridge in a particular minute   (1 mark)
2. Define exponential distribution   (1 mark)
3. Suppose the mean checkout time of a supermarket cashier is three minutes. Use R to find the probability of a customer checkout being completed by the cashier in less than two minutes by applying the functionpexp of the exponential distribution with rate=1/3.

                                                                     (3marks)

$$8y + 16z = 0$$
4. Solve a system of equation using matrix $x - 3z = 1$ (10 marks)
$$-4x14y + 2z = 6$$

QUESTION THREE (20 MARKS)

1. State and explain three types of elementary operation which it can be used to transformed a system of linear equation to a simpler equivalent system (5marks)

2. State three steps of planning a function to extract every other element of a vector (3marks)

3. Find the correlation coefficient of the eruption duration and waiting time in the data set faithful. Observe if there is any linear relationship between the variables. (4marks)

4. Describe the following terms
   i. Probability distribution (2marks)
   ii. Normal distribution (2marks)
   iii. Binomial distribution (2marks)
   iv. Central Limit Theorem (2marks)

QUESTION FOUR (20 MARKS)

1. Define a relative frequency distribution (2 marks)

2. State and explain five principles of good coding (5marks)

3. Suppose there are twelve multiple choice questions in a Statistics class quiz. Each question has five possible answers, and only one of them is correct. Using R, write a program which find

   i. The probability of having exactly 4 correct answers if a student attempts to answer every question at random. (3 marks)

   ii. The probability of having four or less correct answers if a student attempts to answer every question at random. (5 marks)

   iii. Then give the probability percentage of four or less questions answered correctly by random in a twelve question multiple choice quiz (1 marks)

4. Assume that the test scores of a college entrance exam fits a normal distribution. Furthermore, the mean test score is 72, and the standard deviation is 15.2. Using R, write and a program which finds the percentage of students scoring 84 or more in the exam? (4 marks)