B. Sc. DEGREE END SEMESTER EXAMINATION - APRIL 2021

SEMESTER – 6: COMPUTER APPLICATION (STATISTICS)

COURSE: **15U6CRCST7: COMPUTER AIDED DATA ANALYSIS USING EXCEL & R**

(Common for Regular 2018 admission & Improvement 2017/Supplementary 2017/2016 /2015 admissions)

Time: Three Hours   Max Marks: 75

Part A

(Each question carries *five* marks, maximum marks from this part is *25*)

1. Write short notes on (i) Assignment operator in R, (ii) comparison operators in R, (iii) c function in R. Give examples.

2. How can we input data into R using excel? What are the two ways of access the variables included in the data set? Give examples.

3. Write the commands for creating a matrix in R. How we can create the matrix's row and column names. How we can access the elements of the matrix.

4. Write the codes in R for finding mean, median and mode from an ungrouped data set stored in an object x?

5. Distinguish between low-level function and high-level function used for drawing graphs in R. Give examples. Explain some of the essential arguments of the plot () function.

6. Discuss some of the useful built in function available in R. Give examples.

7. Write the codes in R to plot the scatter diagram and computing the correlation between two objects, "x" and "y." Write codes for finding variances of "x" and "y."

8. Write the codes in excel for conducting two independent sample t-test, how can we interpret the output. What are the essential assumptions for conducting this test?

Part B

(Each question carries *ten* marks, maximum marks from this part is *50*)

9. For the following frequency distribution, using R

Mid-point:	25	35	45	55	65	75	85	95
Frequency:	55	93	113	90	85	73	29	5

   a) Enter the data as dataframe.
   b) Add a column of cumulative frequency.
   c) Add a column of relative frequency (frequency/total frequency)
   d) Add a column of relativecumulative frequency (cumulativefrequency/total frequency)
   e) Plot cumulative frequency vs mid points.

10. Input the following data using read.table function.

Pre	5260	5470	5640	6180	6390	6515	6805	7515	7515	8230	8779
Post	3910	4220	3885	5160	5645	4680	5265	5971	6790	6900	7335

    a. Access 3rd element of the variable "pre"
    b. Access the variable "pre"

c. Access all the elements of the variable "pre" excepting 1st, 2nd, 3rd and 4th.

d. Use attach function and extract all elements of the variable post that are larger than 7000.

11. Compute mean and standard deviation for the following frequency distribution.

x: 94.5   104.5 114.5 124.5 134.5 144.5 154.5 164.5 174.5

f: 5   8   22   27   17   9   5   5   2   Using R.

12. The following table shows the heights of a sample of 10 fathers and their sons:

Height of father:170   167   162   163   167   166   169   171   164   165

Height of Son: 168   167   166   166   168   165   168   170   165   168

Draw the scatter diagram, find the coefficient of correlation between height of father and son and interpret the result (Using MS Excel).

13. A company producing light bulbs finds that the mean life span of the population of bulbs is 1200 hours with a standard deviation of 125 hours. A sample of 100 bulbs produced in a lot is found to have a mean life span of 1150 hours. Test whether the difference between the population and sample mean is statistically significant. (Using MS Excel)

14. To Fertilizers A and B were tried respectively on 10 and 8 randomly chosen experimental plots. The yields in the plots were as given below. Test whether there is a difference in the effects of the fertilizers as reflected in the mean yields. (Using R software)

Fertilizer   Yield

A   8   7.6   8.2   7.8   8.3   8.4   8.2   7.8   7.7   8

B   7.4   8.1   7.6   8.1   7.5   7.6   7.3   7.2

15. Consider the following data on the number of hours that 10 persons studied for statistics test and their scores on the test   (Using MS Excel)

Hours Studied (X):   4   9   10   14   4   7   12   22   1   17

Test Score (Y):   31   58   65   73   37   44   60   91   21   84

(a) Find the equation of the least squares line that approximates the regression the test scores on the number of hours studied.

(b) Predict the average test score of a person who studied 14 hours for the test.

****