SMC-QR-8130

6|8

[This question paper contains 4 printed pages]

Your Roll No.....

Sr. No. of Question Paper :

Unique Paper Code : 42353605

Name of the Paper : SEC-4: Statistical Software-R

Name of the Course : **B.Sc. Mathematical Sc. / B.Sc. (Prog.)– CBCS (LOCF)**

Semester : **VI**

Duration : 2 Hours Maximum Marks : 38

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt all **5** questions, selecting any **2** parts from each.
3. All commands should be written using language R.

1. Total marks : 2x2 = 4

   Attempt any **two** of the following :

   (a) (i) Write a command to list all the variables defined ending with 'm'.
   (ii) Write "Jan", "Feb", "Mar", "Apr" as a factor.

   (b) Write command to compute (i) $\sin^{-1}(1)$ in degree (ii) $(2\pi/3) - 4$

   (c) Write the difference between lapply and sapply commands in R.

   (d) Read the following stem-and-leaf plot.

   The decimal point is 1 digit(s) to the right of the |

   6 | 017
   7 | 123357
   8 | 2699
   9 | 02

   Now give the data as vector 'V' with R command. Also write the R command to find the five basic quartiles of the vector 'V'.

2. Total marks : 3x2 = 6

Attempt any **two** of the following :

(a) Form a data frame of the following data and then convert it into a list

Response	20	14	30	70	19	91	40	25
Predictor	Closed	Closed	Closed	Open	Open	Open	Open	Open

(b) Make a list grass with data

mw: 10 13 15 11 13
umw: 6 8 7 8

Also create a data frame grass_1 by using stack() command and name the columns as 'rich' and 'graze'.

(c) Create a vector in R as

U: 4.3 6 7.5 6 4.1 2.8 5 7 9

Give R commands for converting U into integers, and the Tukey summary values.

(d) For the following data vectors:

Length = {7, 8, 9, 11.5}, and Height = {4, 9.5, 3.9, 2.5}.

Write R command to construct the data frame 'dimension'.

3. Total marks: 4x2 = 8

Attempt any **two** of the following :

(a) Put the following values into variable B: 2, 4, 7, 2, NA, 5, 6, 6, 9, 4, 5, 7, 3, 4, NA. Give R commands to find first five terms, the position of maximum item of 'B', and items less than equal to 7 & strictly more than 4.

(b) Give R command to form a 3x3 matrix: $\begin{pmatrix} 35 & 60 & 86 \\ 55 & 46 & 74 \\ 45 & 58 & 92 \end{pmatrix}$ with row names:

Zoology, Botany, Chemistry and column names: A1, A2, A3, respectively.

Also Find the sum of the rows of the matrix and the mean of the column 'A2'.

(c) Giving R command for the following matrix A:

	Alpha	Beta	Gamma
R1	6	5	3
R2	15	11	4
R3	9	12	2

Convert the matrix A into a data frame and display the data for Beta only.

(d) For the following two-dimensional data:

data_1	data_2	data_3
23	25	34
23	45	12
21	32	21
21	47	43

Write R command to display first and third row. Also, for above data draw bar chart with appropriate labels.

4. Total marks: 5x2 = 10

Giving R command to make a data frame 'DF' of two-column with a response variable (tree) and a predictor variable (site):

tree	8	5	2	9	1	4	5	6	7	7	4	3	3	5	5	4	6	4
site	x	x	y	x	y	z	y	y	y	z	y	y	y	z	y	z	y	x

Write R commands for any **two** of the following:

(a) Produce a histogram and also overlay a density plot of gray color having line width 1.5 units with gaussian kernel for the site 'y' only.

(b) Produce a histogram for 30 random normal variates data with mean and standard deviation from data in part (a), and also overlay two density plots of data one from part (a) and second from 30 normal variates with gaussian kernel, use different colors and line types.

(c) Shapiro-Wilk normality test and normal quantile-quantile plot with a straight line for the site 'y' only. Also using apply family command to produce Shapiro-Wilk normality test for the site 'x', 'y' and 'z' respectively.

(d) Create a Portable Network Graphics image of size 740 x 560 pixels to draw a box-whisker plot of the data frame 'DF' using blue color.

5. Total marks: 5x2 = 10

Giving R command to make a data frame 'df' with data:

x	3	7	12	11	13	8	14	6	15	5	17	19
y	4	16	18	13	19	17	20	14	20	11	12	15

Write R commands for any **two** of the following:

(a) Make a single vector 'rain' from *x* values of the data frame 'df' to draw a bar chart with month as names for the bars. Also label axes as 'month' and 'rainfall cm' with y limits 0 to 20.

(b) Make a single vector 'rain' from *x* values of the data frame 'df' to draw a pie chart clockwise with month as labels stating at 9 o'clock.

(c) Make a single vector 'rain' from *x* values of the data frame 'df' to draw a Cleveland dot plot with month as labels.

(d) To draw a scatter plot of data points (*x*, y) using gray color symbol '*' of 1.5 units size and axis labels with limits each 0 to 20.

Also add a line of best fit for the data.