

(c) Given the data

X: 3, 5, 7, 5, 3, 2, 6, 8, 8, 5, 6, 9, 4, 5, 7, 3, 4

(i) Write the command for drawing histogram in R.

(ii) Specify the breaks of the bar at 2, 5, 6, 8, and 9.

(iii) Colour the bars in Red and write the title HISTOGRAM.

(d) Explain Quantile-Quantile Plot with the help of example.

6. Attempt any two parts. Each part is of 3 marks.

The following data of count and speed are given below:

Count	speed
9	2
25	3
15	5
2	9
14	14
25	24
24	29
47	34

Suppose that the above data is expressed in the form of a data frame named FM.

(a) You have to express the information with the help of box plot. Write the syntax of the following in R software:

(i) Make box plot of count and speed.

(ii) Label box plot of count and speed as count and speed respectively.

(iii) Make the horizontal box plot of count and speed.

(b) Write the syntax of scatter plot for the following using the data frame FM:

(i) Scatter plot of FM.

(ii) Add line of best fit to scatter plot of FM.

(iii) Using ■ symbol, draw the scatter plot of FM

(c) Write a short note on Pair Plot.

(d) Month wise rain data of Delhi is given as:

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
rain	3	5	7	5	3	2	6	8	5	6	9	8

rain and month variables are defined below in R.

rain = c(3, 5, 7, 5, 3, 2, 6, 8, 5, 6, 9, 8)

month = c('Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec')

Draw the output of the following commands:

(i) plot(rain, type = 'b', axes = FALSE, xlab = 'Month', ylab = 'Rainfall cm')

(ii) axis(side = 1, at = 1:length(rain), labels = month)

(iii) axis(side = 2)

(iv) box ()

(500)

(13/5/2022 Eve)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2738

A

Unique Paper Code : 62353607

Name of the Paper : SEC – Statistical Software-R

Name of the Course : B.A. Programme

Semester : VI

Duration : 2 Hours

Maximum Marks : 38

Instructions for Candidates

- Write your Roll No. on the top immediately on receipt of this question paper.
- All the questions are compulsory.
- Attempt any two subparts of each part in question 1. Each subpart is of 1 mark and whole question 1 is of 8 marks.
- In questions 2 to 6, attempt any two parts out of four parts. Each part is of 3 marks and each question is of 6 marks.

1. Attempt any two subparts of each part. Each subpart is of 1 mark. This whole question carries 8 marks.

(a)

(i) If $a = 15$, $b = 3$ then $\text{print}(a\%b)$ will give _____.

(ii) Write the R syntax to find $\cos(45^\circ)$.

(iii) Write the R syntax to find $\log_3 10$.

(iv) Write the command to find the variance of data.

(b)

(i) hist() command is used for _____ (history, histogram).

(ii) _____ command can sort the data.(order(), rank()).

(iii) _____ command produces the sum values of columns. (columnSums() / colSums()).

(iv) _____ command is used to make scatter plot. (spplot() / plot())

P.T.O.

- (c) State whether TRUE or FALSE
- `rm()` command finds the defined variables.
 - `c(4 6 7 9)` gives a data frame.
 - `ls.str()` command finds the structure of all the defined variables.
 - If `x=c(2,34,56,43,23)`, then `sort(x)` will sort the data in descending order.
- (d) State whether TRUE or FALSE
- `read.csv(choose.file())` is used to read a file.
 - Quantile-Quantile plots are used for visualizing data in a straight line.
 - If `x` is a five item vector then `x[1:4]` command shows only first and fourth items.
 - `is.matrix(x)` command is used to verify if a given object '`x`' is a matrix data object.
2. Attempt any two parts. Each part is of 3 marks.
- (a) Differentiate between `concatenate(c)` and `scan` commands. Also use these commands to create simple data containing the text: Jan Feb Mar Apr.
- (b) Write a command to remove all the variables defined beginning with 'a' and ending with 's'. Differentiate between list and data frame.
- (c) Read data from the file "marks.csv". How to make a comment in R?
- (d) Create a vector
`x: 10, 2, 3.5, -12, NA, 4.2, NA`. How to check the type of this vector? Convert the type of this vector into integer.
3. Attempt any two parts. Each part is of 3 marks.
- (a) In each of the following, find the difference between the two commands
- `max(x)` and `which(x == max(x))`, where `x` is a vector of some numbers
 - `sort(x, na.last = NA)` and `sort(x, na.last = FALSE)`, where `x` is a vector of numbers along with NA entries
- (b) In each of the following, explain why there is error in the output.
- `Weekdays=c(Mon, Tue, Wed, Thu, Fri)`
 - `x=c(3, 2, 1, 4, 5); sort(x, decreasing = true)`
 - `mat=matrix(1:10, ncol = 3)`
- (c) Suppose that `Data=c(7,8,5,5,3,4,4,6,9,7)` is the vector of the 10 numbers. Describe the output of the following commands:
- `Data[1-length(Data)]`
 - `Data[3:7]`
 - `Data[Data>5]`
- (d) Illustrate using an example what the three commands `sort()`, `order()` and `rank()` do on vectors of numbers, and point out the differences between them.

4. Attempt any two parts. Each part is of 3 marks.
- (a) Write the commands to save the data in the following table as a matrix named 'Mat', and then add an extra column on the right of it whose entries are the sum of the entries in each row.

	C1	C2	C3	C4
R1	1	4	7	10
R2	2	5	8	11
R3	3	6	9	12

- (b) Write the commands to store the data in the following table as a data frame named 'Data1', and explain how would the output of the command `summary(Data1)` look like.

	Age	Gender
Bob	21	M
Alice	20	F
James	22	M
Henry	23	M

- (c) Suppose that the following data is stored in a data frame named 'Data2'. How would you sort the data in the descending order of the entries in the last column? How would you add information of Person-6 (Gender: M, Age: 33) into this data frame?

	Gender	Age
Person-1	M	27
Person-2	M	36
Person-3	F	41
Person-4	M	34
Person-5	F	23

- (d) What are `attach()` ... `detach()` and `with()` commands used for? Illustrate using examples.
5. Attempt any two parts. Each part is of 3 marks.
- (a) Write a short note on "the Shapiro-Wilk test". Explain it with the help of example.
- (b) Write the command of the following in R :
- Generate `n` random variables from the Poisson distribution with `lambda = 10`.
 - Generate `n` random variables using the binomial distribution having size = 20 and probability of each trial being a success is 0.6.
 - Generate 10 random variables from the Uniform distribution between 10 to 30.