(ii) Create a suitable plot to study the distribution of avgscore for the whole class with appropriate chart label and axes labels. (3)

7. (a) Differentiate between sort() and order() functions with the help of an example. (3)

(b) What will be the output of the following R commands for a given vector x?

x <- c(2,7,–1,0,–4,70).

(i) x[which(x%%2==0)]

(ii) which.max(x) (2)

[This question paper contains 6 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 4582 **E**

Unique Paper Code : 32343408

Name of the Paper : Introduction to R Programming (SEC)

Name of the Course : **B.Sc. (Hons.) Computer Science**

Semester : IV

Duration : 2 Hours Maximum Marks : 25

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. All parts of Question **1 (Part A)** are compulsory.
3. Attempt any **three** questions from **Part B**.
4. **All** questions in **Part B** carry equal marks.

Section – A

1. (a) What is the purpose of attach() function in R? Give an example. (2)

(b) Write the name and syntax of the R function that is used to identify rows without any NA values. (2)

(c) Consider two R matrices as shown below : (2)

$$M1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ and } M2 = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$$

Write the output of the following R commands:

(i) M1*M2

(ii) M1 %*% M2

(d) What is the purpose of factors in R? Give an example. (2)

(e) Write the output of the following commands:

(i) rep(1:4, times=2)

(ii) rep(1:4, each=2) (2)

Section – B

2. (a) Write the output of the given R script. (2)

```
x <- c (1,0,NA,0,4)
p <- 0/x
is.nan(p)
is.na(p)
```


PCODE	PNAME	PQTY	PRICE

Write the R commands to perform the following :

(i) Load the R package to connect with MySQL database.

(ii) Connect to the database *'db1'*. Assume suitable login details.

(iii) Display the rows that have price greater than 20000 from the PRODUCT table. (2.5)

6. (a) Explain the purpose of skip and nrow arguments of read.table function. (2)

(b) Consider the'student'data frame as given below :

Roll No.	Score1	Score2
20/CS/02	78	87
20/CS/03	90	56
20/CS/33	89	43
20/CS/28	74	69

Write R commands to do the following :

(i) Add a new column 'avgscore' containing the average of score 1 and score 2 columns.

(b) Write the output for the following R script. (3)

```
f1 <- function(ob)
{
        length(ob)
}
L <- list(a=c(1,NA,2), b=-5:5))
print(L)
sapply(L,f1)
```

3. Consider the data file *"pollutant.csv"* as shown below : (1+2+2).

	City	Date	PM2.5
1	Delhi	01-04-2022	195
2	Delhi	02-04-2022	200
3	Mumbai	01-04-2022	110
4	Chennai	02-04-2022	90
5	Mumbai	02-04-2022	NA

Write the R commands for the following :

(i) Read the file *pollutant.csv* into a data frame 'pm25'.

(ii) Display the average PM2.5 level for the city 'Mumbai'.

(iii) Display the number of days when PM2.5 level of 'Mumbai' city was greater than 100.

4. (a) Define a function 'mysearch' in R that searches for an element in a given vector. The function should return the position of first occurrence of the element. If the element is not found, then the function should return –1. The prototype of the function is : (3)

mysearch(data,element)

Example:

mysearch(c(1,0,3,0,1),0) should return 2 i.e. index of first occurrence of 0

(b) Differentiate between paste and paste 0 functions in R with the help of an example. (2)

5. (a) List the compulsory and optional files required in a package directory structure. (2.5)

(b) Consider the structure of the PRODUCT table in the database 'db1' as given below.