778 (M)

Your Roll No:...........................

Unique Paper Code	: 32343408
Name of the Paper	: Introduction to R Programming (SEC)
Name of the Course	: B.Sc.(H) Computer Science
Semester	: IV
Duration : 2 hours	Max. 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.

PART A

1. Answer the following questions:

   (a) What is a named vector? Give an example.   2

   (b) Differentiate between vector recycling and vector repetition.   2

   (c) Considering the list L1 as given below. State the difference between the   2
       statements 1 and 2.

       L1 ← list(1:5, 7, list(2.2,3.3,4.4))

       Statement 1: L1[1]

       Statement 2: L1[[1]]

   (d) Write R command to define a factor that stores days of the week. Assume   2
       the 'Monday' as first day of the week.

   (e) Write the R command to print average of the vector X as given below: 1
       X←c(3,-3,5,6,-2,NA,-4)

   (f) What is the purpose of the *fill* argument in the *read.table()* function?   1

PART B

2. Consider the file 'sampledata.csv' as shown below:

   Rollno;Name;Age;Course

   101;John;18;CS

   102;Tom;17;MT

   103;Joy;20;CS

   104;Mary;NA;MT

   Write the R commands to do the following:

   (a) Read the data from the file into a dataframe named *studf*.   2

   (b) Display the names of the students having Age greater than 18.   1

   (c) Display the average Age of the students.   1

   (d) Display the Rollno of students enrolled in the course CS.   1

3. Consider PRODUCT table stored in a MySQL database 'ecom'. The structure of the PRODUCT table is as given below:

   PRODUCT(PCODE, PNAME, QTY, PTYPE, PRICE)

   Write R command(s) to do the following: 5

   (a) Load relevant packages to connect to the database.

   (b) Connect to the 'ecom' database.

   (c) Display all products whose quantity is greater than 100.

   (d) Display the maximum price for each type of product.

   (e) Close the connection with the database.

4. (a) Differentiate between the functions `lapply` and `sapply`.   2

   (b) What will the output of the R script given below: 3

   ```
   x ← c(0,NA,NULL,2,0/0)
   ptint(length(x))
   print(is.na(x))
   print(is.nan(x))
   ```

5. Define a function `fact()` to calculate factorial of a given number and a vector named 'facts' containing factorial of first 5 numbers. Create a package 'myfact' containing the function `fact()` and vector `facts`.   5

6. Consider the dataframe 'stures' given below:

Rno	Name	Paper1	Paper2
1101	Ritu	78	70
1102	Shyam	87	91
1103	Neetu	65	55

Write the R commands to do the following:

(i) Display the name of the students in descending order of their score in Paper1. 1

(ii) Display the name of the students who have scored more than 75 in both papers. 1

(iii) Draw a suitable plot to show the frequency distribution of marks in Paper2. 1

(iv) Draw a barplot on Paper1 scores. Provide suitable chart title and axes labels. 2

7. (a) Assume there is a data file available on web via the URL http://www.abc.com/data/weather.csv. Write the command to download this file and save it as *myweatherdata.csv* on your local computer. 2

(b) Write R commands to do the following: 3

(i) Define a 4x5 integer matrix M1.

(ii) Display row-wise mean of the matrix M1.

(iii) Reshape the matrix M1 into 2 rows and 10 columns.