MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

1. Create an cars.xlsx on the following data. Predict the mpg of a car with cyl =8 and hp=200 using Linear Regression. Display the proportion of cars with different numbers of gears.

Cars	**mpg**	**cyl**	**hp**
car1	21.0	6	110
car2	22.8	4	93
car3	21.4	6	110
car4	18.7	8	175
car5	18.1	6	105
car6	14.3	8	245

MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

2. Create an passenger.xlsx on the following data. Find all the female passengers whose Age is above 35.0 after removing the NA. Find the oldest female who survived. Draw the histogram for the distribution of age using ggplot2 and hist().

Survived	Sex	Age
1	female	38.0
1	female	26.0
1	female	35.0
0	male	35.0
0	male	*NA*
0	male	54.0
0	male	2.0
1	female	27.0
1	female	14.0
1	female	*NA*
1	female	58.0
0	male	20.0

MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

3. Create an student.xlsx on the following data. Build a decision tree model on result. Predict the possibility of Pass/Fail of a student whose Study_Hours = 6, Sleep_Hours = 5.

Study_Hours	Sleep_Hours	Result
2	8	Fail
5	6	Pass
1	7	Fail
7	5	Pass
10	4	Pass
3	7	Fail
6	6	Pass
8	5	Pass
4	6	Fail
9	4	Pass

MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

4. Create an fit.xlsx on the following data. Count the total number of data present into your dataset. Short the data groupwise on Age and fitness. Draw the histogram for the distribution of age using ggplot2 and hist().

Age	Fitness	Income	Miles
18	4	29562	112
19	3	31836	75
19	3	30699	66
19	3	32973	85
20	2	35247	47
20	3	32973	66
21	3	35247	75
21	3	32973	85
21	4	35247	141
21	3	37521	85

MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

5. a. Given the following parameters for the annual rainfall (in inches) in a certain region:

- Mean (μ) = 45 inches
- Standard Deviation (σ) = 6 inches

What is the probability that the annual rainfall will be between 40 and 50 inches?

b. Write a recursive user define function to print the Fibonacci Numbers up to a user given terms.

MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

6. Create an study.xlsx on the following data. Now Visualize the linear regression where Exam_Score is the dependent variable and Study_Hours is the independent variable. Predict the predict the exam score of a student who studies for 7.5 hours. Create a scatter plot to visualize the relationship of Exam_Score and Study_Hours using ggplot2 and simple plot().

Study_Hours	Exam_Score
1	50
2	55
3	65
4	70
5	75
6	78
7	85
8	88
9	90
10	95

MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

7. Create an product.xlsx on the following data. Now create a bar plot using ggplot2 and barplot() to visualize the sales data. Save the details of the product where the sales are more than 150 but less than equal to 250 into p1.csv file. Print A/B/C/D/E for the highest sales.

Product	Sales
Product A	150
Product B	200
Product C	100
Product D	250
Product E	175

MCKV Institute of Engineering

B.Tech. 5th Semester Practical Examination

Academic Year: 2024-25

Stream : CSE (DS)

Paper Name: Introduction to R Programming **Paper Code: PC-CS593**

Time allotted: 3 hours ***Full Marks: 60***

Marks Distribution
Source Code: 20
Output: 20
Viva Voce: 20

8. Create an cars.xlsx on the following data. Print the name of the car with highest mpg. Predict the mpg of a car with cyl =8 and hp=200 using Linear Regression. Calculate the Pearson correlation coefficient between cyl and hp and interpret the result.

Cars	mpg	cyl	hp
car1	21.0	6	110
car2	22.8	4	93
car3	21.4	6	110
car4	18.7	8	175
car5	18.1	6	105
car6	14.3	8	245