

Roll No.

Total Pages : 2

BT-7/D-23

47420

DATA SCIENCE WITH R PROGRAMMING

Paper-PC-CS-AIML-401A

Time Allowed : 3 Hours]

[Maximum Marks : 75

Note : Attempt **five** questions in all, selecting at least **one** question from each Unit. All questions carry equal marks.

UNIT-I

1. (a) Discuss the various data science tools and the relationship of data science with other fields. 8
- (b) Explain Data analysis and its types in detail. 7
2. (a) What are the different data types and in R programming? Discuss them with syntax. 5
- (b) Explain the following with suitable examples :
 - (i) Data Cleaning
 - (ii) Data Reduction
 - (iii) Data Transformation
 - (iv) Data Discretization. 10

UNIT-II

3. (a) What are the different data visualization libraries in R programming? 7

- (b) Discuss the data visualization software and specialized tools. 8
4. (a) How is the position dependence and independence used to find the descriptive statistics? 7
- (b) Explain the Bayesian probability and probability distribution with examples. 8

UNIT-III

5. (a) Write a R script to find the mean, median and mode of a data. 9
- (b) Discuss the uses, advantages, and disadvantages of R. 6
6. (a) Name different data types in R. Write an R script to convert matrix to vector. 7
- (b) Explain data type conversion. Write an R script to convert vector to data frames. 8

UNIT-IV

7. (a) How to get and set a working directory for CSV files in R programming? 6
- (b) Explain Linear, polynomial and logistics method for regression in data science. 9
8. (a) Discuss the different clustering and classification methods. 9
- (b) Describe the Hidden Markov Model. 6