

**Code No: 185CX****JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD****B. Tech III Year I Semester Examinations, January - 2025****INTRODUCTION TO DATA SCIENCE****(Common to CSE(DS), AI&DS)****Time: 3 Hours****Max. Marks: 60****Note:** This question paper contains two parts A and B.

i) Part- A for 10 marks, ii) Part - B for 50 marks.

- Part-A is a compulsory question which consists of ten sub-questions from all units carrying equal marks.
- Part-B consists of ten questions (numbered from 2 to 11) carrying 10 marks each. From each unit, there are two questions and the student should answer one of them. Hence, the student should answer five questions from Part-B.

**PART- A****(10 Marks)**

- List the basic data types available in R. [1]
- What is the purpose of fitting a model in data science? [1]
- What does the term skewness refer to in a dataset? [1]
- Which statistical measure is most appropriate for ordinal data? [1]
- How do you extend a data frame in R? [1]
- What is a named list in R? [1]
- How do you load a package in R? [1]
- What is function scoping in R? [1]
- What is the role of a chart legend in data visualization? [1]
- What kind of data is best represented by a line graph? [1]

**PART - B****(50 Marks)**

2.a) Discuss the relationship between Big Data and Data Science with a suitable diagram.

b) Describe the key steps in the statistical modelling process. [5+5]

**OR**

3.a) How does fitting a model help in understanding a dataset? Give an example.

b) Describe the steps involved in setting up an R programming environment. [5+5]

4.a) Explain the different types of attributes with examples for each type.

b) How do you calculate the range and IQR for a given dataset? Give one example.

[5+5]

**OR**

5.a) Compare discrete and continuous attributes with relevant examples.

b) Compare and contrast variance and standard deviation as measures of data spread using examples. [5+5]

QA QA QA QA QA QA QA G

- 6.a) Explain how sorting and merging are handled in data frames with examples.  
b) State the process of converting a list into a vector in R with code. [5+5]

QA QA QA QA OR QA QA QA G

- 7.a) How do factors and data frames complement each other in R? Explain with an example.  
b) Describe the key operations that can be performed on vectors in R. [5+5]

- 8.a) Write an R program to find the factorial of a number using recursion.  
b) Explain the differences between mathematical functions and custom functions in R. [5+5]

QA QA QA QA OR QA QA QA G

- 9.a) Write an R program to find the highest common factor and greatest common divisor.  
b) Write an R program to show the multiplication table of a given number. [5+5]

- 10.a) Explain the steps involved in performing linear regression analysis.  
b) Discuss how outliers affect regression analysis and ways to handle them. [5+5]

QA QA QA QA OR QA QA QA G

- 11.a) How can a scatter plot in R be created to visualize relationships between variables? Write the R code for it.  
b) Discuss the advantages of using multiple linear regression over simple linear regression. [5+5]

QA QA QA QA ---ooOoo--- QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G