



TKRCET  
 Indian to Character • International in Excellence

*R22 Regulation* *Subject code:4E3HA*  
**TKR COLLEGE OF ENGINEERING AND TECHNOLOGY**  
 (Autonomous, Accredited by NAAC with 'A+' Grade)

**B.Tech III Semester Regular/Supplementary Examinations, December 2025**  
**INTRODUCTION TO DATA SCIENCE**  
*(CSE(DS))*

**Maximum Marks: 60**

Date:22.12.2025

Duration: 3 hours

- Note:**
1. This question paper contains two parts A and B.
  2. Part A is compulsory which carries 10 marks. Answer all questions in Part A.
  3. Part B consists of 5 Units. Answer any one full question from each unit.
  4. Each question carries 10 marks and may have a, b, c, d as sub questions.

**Part-A**

All the following questions carry equal marks (10X1M=10 Marks)		Marks	CO	BTL
1.a)	Explain the term properties of data with any one example.	1M	1	K2
b)	What is meant by classification in data science.	1M	1	K1
c)	What is a chart type in data visualization.	1M	2	K1
d)	Define Exploratory Data Analysis.	1M	2	K1
e)	What is linear regression.	1M	3	K1
f)	Define clustering.	1M	3	K1
g)	What is the purpose of the help function in R.	1M	4	K1
h)	What is a vector in R.	1M	4	K1
i)	What is a list in R.	1M	5	K1
j)	Explain factors in R programming.	1M	5	K2

**Part-B**

Answer All the following questions. (5X10M=50Marks)		Marks	CO	BTL
2	Analyze how <i>Computer Science</i> , <i>Data Science</i> , and <i>Real Science</i> differ in terms of goals, methodologies, and approaches to problem-solving. Using examples, explain how asking interesting questions from data helps bridge these domains and leads to meaningful insights.	10M	1	K4
OR				
3	Examine the role of scoring and ranking systems in data-driven decision making. Analyze how BMI, Z-scores, and normalization contribute to developing effective scoring systems, and evaluate the limitations of relying solely on these techniques for real-world ranking problems.	10M	1	K4
4	Explain the process of Exploratory Data Analysis (EDA). Describe how visualization aesthetics, chart types, and great visualizations support the understanding of data before model building.	10M	2	K2
OR				
5	Discuss the philosophies of mathematical modeling and explain the taxonomy of models. Describe the importance of baseline models and model evaluation environments in understanding model performance.	10M	2	K2

6	Explain linear and logistic regression and describe how better regression models are developed. Discuss the concept of regression as parameter fitting with suitable examples.	10M	3	K2
OR				
7	Describe how distances are measured in data analysis and explain the principles of nearest neighbor classification. Discuss the role of graphs, networks, and PageRank in analyzing relationships within data.	10M	3	K2
8	Apply your understanding of R programming basics to explain the use of functions, major data structures, and the help and example functions. Demonstrate how these components support performing regression analysis of exam grades.	10M	4	K3
OR				
9	Apply vector operations in R to real data examples by explaining scalars, vectors, arrays, and matrices. Illustrate how filtering, NA/NULL handling, vectorized operations, and vectorized if-then-else constructs are used in data processing.	10M	4	K3
10	Explain the creation and manipulation of matrices and arrays in R and describe how they are used to handle multidimensional scientific or statistical data.	10M	5	K2
OR				
11	Discuss lists, data frames, factors, and tables in R. Explain how each structure is used to store, organize, and analyze different types of data.	10M	5	K2