



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

B.Tech III Semester Regular Examinations, February 2024
INTRODUCTION TO DATA SCIENCE
(CSE (DATA SCIENCE))

Maximum Marks: 60

Date: 23.02.2024 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		CO	Bloom Tx
All the following questions carry equal marks (10×1M=10 Marks)			
1.a)	Write the significance of Z-scores	1	L1
b)	What are data visualization types?	1	L1
c)	What are the different types of charts and visualizations?	2	L2
d)	What trade-offs do modelers face when choosing a modeling philosophy?	2	L2
e)	What methods are commonly used for measuring distances in data analysis?	3	L2
f)	What is the primary purpose of clustering in data analysis?	3	L2
g)	Why are startup and shutdown processes important in R?	4	L2
h)	What techniques are available for filtering data in R?	4	L2
i)	What is the primary purpose of lists in R?	5	L2
j)	Why the factors important in R?	5	L2
Part-B			
Answer All the following questions. (5×10M=50Marks)			Bloom Tx level
2	a) Define data science and explain its role in computer science and real-world applications. [5M] b) What is classification and regression? How do you assess the fit of a regression model to the data? [5M]	1	L2 L3
OR			
3	What advanced ranking techniques does the show cover, and how can they be applied in real-world scenarios? [10M]	1	L2
4	What design principles should be considered when developing a visualization aesthetic to ensure clarity and effectiveness? [10M]	2	L3
OR			
5	a) What is the concept of baseline models, and why are they essential in the modeling process? [5M] b) What are the key metrics used for evaluating the performance of mathematical models? [5M]	2	L2 L2

6	a) How does logistic regression differ from linear regression? [5M] b) How does regression involve parameter fitting, and what role do these parameters play? [5M]	3	L2 L3
	OR		
7	Explain the significance of the PageRank algorithm in web page ranking. [10M]	3	L2
8	a) What are some common operations that can be performed on vectors in R? [5M] b) Explain the concept of vectorized operations in R and their advantages in data analysis tasks? [5M]	4	L1 L2
	OR		
9	How are scalars, vectors, arrays, and matrices declared in R, and what distinguishes them from each other? [10M]	4	L2
10	What is a data frame? How do you create a data frame in R and what are the key functions for manipulating data frames? [10M]	5	L2
	OR		
11	a) How do you create factors, and what role do they play in statistical analysis? [5M] b) Explain the difference between ordered and unordered factors. [5M]	5	L4 L2