R22 Regulation *Subject code: 4O8GA*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

(Autonomous, Accredited by NAAC with 'A+' Grade)

B.Tech VIII Semester Regular Examinations, May 2026

FUNDAMENTALS OF DATA SCIENCE

(CSE (AI&ML))

Maximum Marks: 60 Date: 08.05.2026 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)	Write the concept of Data Science and its importance.	1	1	U
b)	Define the term "Datafication" with an example.	1	1	R
c)	Compare Binary Attributes and Ordinal Attributes?	1	2	U
d)	Define Interquartile Range.	1	2	R
e)	Explain the concept of factors and factor levels.	1	3	U
f)	Define Data Frame in R.	1	3	R
g)	Explain conditional statements in R.	1	4	U
h)	Give an example for conditional statement in R?	1	4	U
i)	Explain the use of scatter plots in data analysis.	1	5	U
j)	Define linear regression.	1	5	R

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	BTL
2	A. Explain Underfitting and Overfitting. B. Compare and analyze Big Data and Data Science with suitable examples.	5 5	1	U
	OR			
3	A. Explain the concept of statistical inference and its importance in Data Science. B. How do populations and samples play a role in this process?	5 5	1	U
4	A. Discuss measures of central tendency and apply mean, median, and mode to a dataset. B. Analyze different measures of dispersion such as range, variance, and standard deviation.	5 5	2	Ap
	OR			
5	A. Explain quartiles and interquartile range with simple example. B. Analyze the role of graphical representations in statistical data description.	5 5	2	Ap

6	A. Explain how to create a vector in R. Include the different ways to name vectors. B. Perform vector arithmetic (addition, subtraction, multiplication, and division). Provide examples to illustrate how vectors can be used.	5 5	3	U
	OR			
7	A. Explain matrices and arrays in R and demonstrate their creation and sub setting with examples. B. Discuss how to subset matrices using row and column indices, as well as how to perform matrix arithmetic.	5 5	3	U
8	A. Discuss the concept of nested functions in R, where one function is defined within another. B. Explain the benefits and challenges of using nested functions	5 5	4	Ap
	OR			
9	A. Discuss the process of loading and using R packages. B. Explain why packages are essential in R and how they extend the functionality of the base R environment.	5 5	4	Ap
10	A. Describe box plot and analyze how it helps in identifying outliers. B. Explain the role of data visualization in Data Science and its importance in decision-making.	5 5	5	Ap
	OR			
11	A. Compare and analyze simple linear regression and multiple linear regression. B. Explain the advantages and limitations of different graphical representations in data analysis.	5 5	5	Ap