

Code No.:(R22DS515PE)

R22

H.T.No.

8 R

**CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS**

**III-B.TECH-I-Semester End Examinations (Regular) - December- 2024
DATA SCIENCE APPLICATIONS
(CSD)**

[Time: 3 Hours]

[Max. Marks: 60]

Note: This question paper contains two parts A and B.

Part A is compulsory which carries 10 marks. Answer all questions in Part A.

Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks.

PART-A

(10 Marks)

1. a) What are the roles of data analyst? [1M]
- b) List the challenges faced by data scientist. [1M]
- c) What is data reduction? [1M]
- d) How can data reduction techniques improve computational efficiency? [1M]
- e) Define descriptive statistics. [1M]
- f) Define ANOVA. [1M]
- g) What is the significance of distribution plots? [1M]
- h) Mention the role of plots in assessing regression model assumptions. [1M]
- i) Explain about generalization error. [1M]
- j) How can cross-validation aid in selecting the best model? [1M]

PART-B

(50 Marks)

2. Compare and contrast data science with traditional data analysis techniques. [10M]
- OR**
3. How do data preparation and feature engineering impact the success of a data science project? [10M]
 4. Explain the importance of data collection strategies in the success of a data science project. [10M]
- OR**
5. Describe the role of data pre-processing in the data science lifecycle. [10M]
 6. How can outliers influence the mean and standard deviation? Illustrate with examples. [10M]
- OR**
7. How do skewness and kurtosis help identify potential anomalies in datasets? [10M]
 8. Discuss the limitations of in-sample evaluation and the importance of Decision making in regression analysis. [10M]
- OR**
9. How do pipelines simplify the workflow of polynomial regression and other machine learning tasks? [10M]
 10. How can a balance between overfitting and underfitting be achieved during model training? [10M]
- OR**
11. Why are out-of-sample evaluation metrics crucial for assessing model performance? [10M]
