Code No.: AI742PE      R20   H.T.No. [ ][ ][8][R][ ][ ][ ][ ][ ]

# CMR ENGINEERING COLLEGE: : HYDERABAD
## UGC AUTONOMOUS
### IV–B.TECH–I–Semester End Examinations (Supply) – April - 2024
### DATA ANALYTICS AND VISUALIZATION
### (CSM)

[Time: 3 Hours]                                              [Max. Marks: 70]

**Note:** This question paper contains two parts A and B.
Part A is compulsory which carries 20 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 and may have a, b, c as sub questions.

## PART-A                                                    (20 Marks)

| | | |
|---|---|---|
| 1. a) | What is management data analysis? | [2M] |
| b) | What are outliers? Give an example. | [2M] |
| c) | What is the importance of Data analytics important? | [2M] |
| d) | List the Uses of Data modeling. | [2M] |
| e) | What is multi -collinearity. | [2M] |
| f) | What is F1-score? Write the formula to calculate F1-score. | [2M] |
| g) | Write two differences of supervised and unsupervised Learning. | [2M] |
| h) | What is entropy function in decision tree? Give an example. | [2M] |
| i) | List the different categorization of visualization methods. | [2M] |
| j) | Write about donut chart. | [2M] |

## PART-B                                                    (50 Marks)

2. With neat diagram illustrate the data and architecture design.            [10M]

**OR**

3. Explain about different stages in Data Processing.                        [10M]

4. Summarize the applications of Data Analytics with examples.              [10M]

**OR**

5. Define Imputations. Explain different types of Missing Data.             [10M]

6. Apply linear regression using the method of least squares to the following data and predict the crop yield for rain fall of 5 cm.     [10M]

| Rain fall(in cms) | 10.5 | 8.8 | 13.4 | 12.5 | 18.8 | 10.3 | 7.0 | 15.6 | 16 |
|---|---|---|---|---|---|---|---|---|---|
| Paddy yield(quintal per acre) | 30.3 | 46.2 | 58.8 | 59.0 | 82.4 | 49.2 | 31.9 | 76.0 | 78.8 |

**OR**

7. Explain different types of variables used in Regression modeling.         [10M]

8. Explain Decision tree classification algorithm.                           [10M]

**OR**

9. Explain with suitable example about ARIMA.                                [10M]

10. Illustrate with examples the different Geometric Projection Visualization Techniques.   [10M]

**OR**

11. Explain the challenges in visualizing complex data and relations and suggest suitable mechanisms to address them.   [10M]

\*\*\*\*\*\*\*\*\*\*\*\*