

Code No.: AI732PE

R20

H.T.No.

8 R

**CMR ENGINEERING COLLEGE : HYDERABAD**  
**UGC AUTONOMOUS**  
**IV-B.TECH-I-Semester End Examinations (Supply) – April - 2024**  
**BIG DATA ANALYTICS**  
**(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 the role of Data Management? [2M]
- b) Define data quality. [2M]
- c) Write the history of Hadoop. [2M]
- d) How is Spark better than MapReduce? [2M]
- e) How do you run descriptive statistics in Excel? [2M]
- f) What technologies support Big Data analysis? [2M]
- g) List the applications of Machine Learning. [2M]
- h) What is a consistent hypothesis in terms of learning? [2M]
- i) Explain Tableau in data visualization. [2M]
- j) What is data analysis and visualization? [2M]

**PART-B**

**(50 Marks)**

2. List and explain the processes of Data Management. [10M]
- OR**
3. Explain the functions of Data Management. [10M]
4. Describe the knowledge management in detail. [10M]
- OR**
5. Discuss the decision models. [10M]
6. Explain the outlier detection process. [10M]
- OR**
7. What is Big Data? Explain the Data Storage and Analysis. [10M]
8. Illustrate the MAP hypotheses and consistent learners. [10M]
- OR**
9. List and explain the various types of the Machine Learning algorithms. [10M]
10. Explain the features of QlikView. [10M]
- OR**
11. List and explain the visualizations tolls. [10M]

\*\*\*\*\*