

[6263]-393

B.E. (Artificial Intelligence and Data Science)
BIG DATA ANALYTICS
(2019 Pattern) (Semester - VIII) (417532B) (Elective - V)

*Time : 2½ Hours]**[Max. Marks : 70**Instructions to the candidates:*

- 1) Answer Q1 or Q2, Q3 or Q4, Q5 or Q6, Q7 or Q8.
- 2) Neat diagrams must be drawn wherever necessary.
- 3) Figures to the right side indicate full marks.
- 4) Use of non programmable electronic calculator is allowed.
- 5) Assume suitable/standard data if necessary.

- Q1)** a) What are the primary methods and functions for importing and exporting data in R and how can they be utilized effectively? [8]
 b) How can data analysts detect and address dirty data using visualizations and statistical techniques in R and what are the implications of anomalies in datasets on decision-making processes? [10]

OR

- Q2)** a) With the help of neat diagram explain the phases of Data Analytics life cycle. [8]
 b) What are the different attribute types in data analysis, and how are they categorized? How does R handle various data types, such as numeric, character, logical, and factors? [10]

- Q3)** a) Explain the following terms with respect to Exploratory data analysis. [10]
- i) Data Sourcing
 - ii) Data Cleaning
 - iii) Univariate analysis
 - iv) Bi-variate/Multivariate analysis
- b) What are the essential steps in data exploration and how do they contribute to uncovering insights and patterns within a dataset? [7]

OR

- Q4)** a) How do ensemble methods such as bagging, boosting, AdaBoost and Random Forest contribute to improving classification accuracy in machine learning models? [10]
 b) How does the utilization of a confusion matrix aid in the evaluation and selection of models. [7]

P.T.O

Q5) a) What are the challenges associated with visualizing big data and how do these challenges impact the effectiveness and efficiency of data analysis and decision-making processes? [10]

b) How does Tableau facilitate effective data visualization and what are some advanced techniques or features within Tableau that can be utilized to create insightful and interactive visualizations for complex datasets? [8]

OR

Q6) a) What are the key features and functionalities of the Google Chart API and how does it enable developers to create dynamic and interactive charts and visualizations for web applications? [8]

b) What are the various types of data visualization techniques available to data analysts? [10]

Q7) a) In what ways does financial data analytics leverage big data technologies to drive insights, mitigate risks and enhance decision-making processes within the financial industry? [10]

b) How does Apache HBase contribute to efficient data storage and retrieval in big data environments [7]

OR

Q8) a) How does the Semantria tool streamline the data collection process and what features or capabilities does it offer that make it a valuable asset for businesses seeking to gather and analyze large volumes of unstructured data from various sources? [10]

b) How does Mozenda serve as an effective data filtering and extraction tool? [7]

