

S.I.E.S. COLLEGE OF ARTS, SCIENCE AND COMMERCE (AUTONOMOUS)
SION WEST, MUMBAI – 400 022
SEMESTER END EXAMINATION 2023-2024

Class: MSc. Part-II
Subject: Data Science

Course name: Big Data Analytics

Day:

Semester: III

Date:

Max marks: 60

Paper: I

Course Code: SIPSDS31

Time:

N.B.

1. All questions are compulsory.
 2. Draw neat diagrams wherever necessary.
 3. Figures to the right indicate full marks.
-

Q1. Attempt the following (any TWO): (12)

- (A) Define term Big Data. What are the sources of Big Data? Why Big Data is important?
- (B) Discuss the Six V's of Big Data with a diagram.
- (C) Differentiate between Structured, Unstructured, and Semi-Structured data.
- (D) What are the challenges of handling of Big Data?

Q2. Attempt the following (any TWO): (12)

- (A) Write a note on the Association Rule with Apriori algorithm.
- (B) Define Machine Learning. Explain Unsupervised Learning.
- (C) Write a note on Supervised Learning.
- (D) Explain K-Nearest Neighbour in detail.

Q3. Attempt the following (any TWO): (12)

- (A) Explain the Decision Tree Technique in data mining.
- (B) How to Create an ARIMA Model for Time Series Analysis?
- (C) What is text mining? Explain the Text mining technique.
- (D) Write a note on Naïve Bayes.

Q4. Attempt the following (any TWO): (12)

- (A) What is Hadoop? Briefly describe the history of Hadoop.
- (B) Explain core components of Hadoop.

- (C) Briefly discuss Hadoop Application, Advantages and Disadvantages.
- (D) Write a note on Hadoop Ecosystem.

Q5. Attempt the following (any TWO): (12)

- (A) Differentiate between RDBMS and NoSQL.
- (B) Explain the term NoSQL. Write features of NoSQL.
- (C) Discuss the architecture of the Document database with MongoDB.
- (D) Explain the different ways of handling of Big data problems.