22SSP645B

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc DEGREE EXAMINATION MAY 2025
(Sixth Semester)

Branch – **SOFTWARE SYSTEMS (Five Year Integrated)**

MAJOR ELECTIVE COURSE – II: BIG DATA ANALYTICS

Time: Three Hours Maximum: 50 Marks

SECTION-A (5 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** marks (5 x 1 = 5)

1. What is Big Data?
   (i) Data smaller than 1GB (ii) Large-scale, complex data
   (iii) Traditional relational data (iv) Only structured data

2 Which of the following is an advantage of Hadoop?
   (i) Centralized processing (ii) Scalability
   (iii) High latency (iv) Only supports SQL queries

3 Which NoSQL database is commonly used for real-time applications?
   (i) MySQL (ii) PostgreSQL
   (iii) Oracle (iv) Cassandra

4 Where is MapReduce primarily applied?
   (i) Cloud storage (ii) Distributed data processing
   (iii) Data visualization (iv) Small dataset processing

5 Which of the following is a Big Data application domain?
   (i) Healthcare (ii) Social media
   (iii) Banking (iv) All of the above

SECTION - B (15 Marks)
Answer **ALL** Questions
ALL Questions Carry **EQUAL** Marks (5 x 3 = 15)

6 a Explain the role of Big Data in modern enterprises.
OR
b Classify the different types of data in Big Data analytics.

7 a Illustrate the role of Hadoop Distributed File System (HDFS).
OR
b Discuss the advantages of using the Hadoop ecosystem.

8 a Show how NoSQL differs from traditional databases.
OR
b Analyze the need for MapReduce in Big Data processing.

Cont...

9 a Apply Pig in solving complex data transformations.

OR

b Determine the performance benefits of Hive over SQL.

10 a Evaluate the impact of Big Data in financial analytics.

OR

b Recommend a Big Data strategy for business intelligence.

SECTION - C (30 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks (5 x 6 = 30)

11 a Compare structured, semi-structured, and unstructured data in Big Data.

OR

b Create a case study on the use of Big Data in e-commerce.

12 a Assess the performance of HDFS in handling large datasets.

OR

b Criticize the limitations of Hadoop in real-time analytics.

13 a Design a schema for storing Big Data in NoSQL databases.

OR

b Interpret the working principles of MapReduce.

14 a Construct a practical Pig script for analyzing web log data.

OR

b Plan a data warehouse solution using Hive.

15 a Justify the application of Big Data in recommendation systems.

OR

b Survey emerging trends in Big Data technologies.

Z-Z-Z END