PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION MAY 2025
(Sixth Semester)

Branch- COMPUTER TECHNOLOGY

BIGDATA ANALYTICS

Time: Three Hours   Maximum: 50 Marks

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

1 Process of analysis of large and complicated data are called_____
   (i) Data mining   (ii) Data warehouse
   (iii) Big data analytics   (iv) Database

2. Which phase in mapreduce is responsible for data aggregation?
   (i) Reduce phase   (ii) Merge phase
   (iii) Map phase   (iv) Shuffle plhase

3. Find the following correct operations that will keep the rows that matches
   the values in the common column for both tables
   (i) Aggregation   (ii) Union
   (iii) Selection   (iv) Natural Join

4 Choose the correct technique to reduce the number of features in a dataset
   (i) Correlation Analysis   (ii) Dimensionality Reduction
   (iii) Feature Selection   (iv) Eigen Decomposition

5 Identify which type of data mainly handle large volumes
   of datain NoSQL databases
   (i) Structured   (ii) Semi-structured
   (iii) Unstructured   (iv) Text data

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

6 a. List the characteristics of Big data.Explain.
                    OR
   b. Describe the classifications in Data mining.

7 a. What are key-value pairs in Mapreduce?
                    OR
   b. Outline the overview of execution of a Mapreduce program.

8 a. How do you implement Mapreduce in Matrix vector multiplication.
                    OR
   b. Explain Union and Intersection operations using mapreduce.

9   a.   Find Root Mean Square Error in Dimensionality reduction to predict the error rate.

OR

       b.   Analyze the process of UV decomposition.

10   a.   Write a note on On-Disk storage of Big data.

OR

       b.   Discuss the key benefits and advantages of NoSQL Database.

SECTION -C (30 Marks)
Answer ALL questions
ALL questions carry EQUAL Marks   (5 x 6 = 30)

11   a.   Briefly explain the types of Big data.

OR

       b.   Discuss the Clustering techniques in Big data.

12   a.   What is Map and reduce phase? Explain its working process.

OR

       b.   Summarize the phases of reduce task and combiners.

13   a.   Classify various operations on relational algebra in Map reduce. Describe with example.

OR

       b.   Elucidate the concept of Matrix multiplication with one map reduce steps.

14   a.   How do you build a complete UV Decompositions algorithm? Explain.

OR

       b.   Inspect Optimizing an arbitrary element of dimensionality reduction.

15   a.   Enumerate Distributed File System of Big data technology.

OR

       b.   Narrate the database of NoSQL and NewSQL.

Z-Z-Z   END