

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)
BSc DEGREE EXAMINATION MAY 2026
(Sixth Semester)

Branch – **COMPUTER TECHNOLOGY**

BIG DATA ANALYTICS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	Which of the following describes the “variety” characteristic of big data? (i) It denotes the different types of data (ii) It refers the speed of execution (iii) It used to store vast amount of data (iv) Quality of the data	K1	CO1
	2	Relate a Twitter messages is an example of _____ (i) Structured (ii) Unstructured (iii) Semi-structured (iv) Hybrid	K2	CO1
2	3	Find out which part of the MapReduce is responsible for processing one or more chunks of data and producing the output results. (i) Map task (ii) Job Task (iii) Mapper (iv) Reduce task	K1	CO2
	4	What is the role of the combiner in the MapReduce model? (i) Data Compression (ii) Data sorting (iii) Data Grouping (iv) Combining Key-value pairs on the mapper	K2	CO2
3	5	What are the two main tasks in the MapReduce algorithm? (i) Mapped, Reduce (ii) Mapping, Reduction (iii) Map, Reduction (iv) Map, Reduce	K1	CO3
	6	Choose the correct answer if given a relation R, partition its tuples according to their values in one set of attributes G. (i) Grouping (ii) Projections (iii) Sorting (iv) Aggregation	K2	CO3
4	7	Select the primary goal of dimensionality reduction. (i) Increase the number of features (ii) Reduce the number of features (iii) Remove the number of feature (iv) Increase model complexity	K1	CO4
	8	What does a lower RMSE value indicates in a model? (i) The model is less interpretable. (ii) The model underfits the data (iii) The model has higher accuracy (iv) The model has multicollinearity	K2	CO4
5	9	What is relation in RDBMS? (i) Key (ii) Value (iii) Table (iv) Datatype	K1	CO5
	10	Show which type of database they do not provide the same transaction and consistency support as exhibited by ACID compliant RDBMS. (i) Mango-DB (ii) NoSQL (iii) NewSQL (iv) PostgreSQL	K2	CO5

Cont...

SECTION - B (35 Marks)Answer **ALL** questions**ALL** questions carry **EQUAL** Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Classify the types of big data with example.	K2	CO1
		(OR)		
	11.b.	Outline the sources of big data.		
2	12.a.	Analyze the use of Map tasks in MapReduce Framework.	K4	CO2
		(OR)		
	12.b.	Simplify the details of MapReduce Execution.		
3	13.a.	Apply the various relational-algebra operations to perform queries.	K3	CO3
		(OR)		
	13.b.	Build matrix multiplication with one MapReduce step.		
4	14.a.	Construct the RMSE value for the Matrices U and V with all entries as 1.	K3	CO4
		(OR)		
	14.b.	Build a Complete UV-Decomposition Algorithm with four steps.		
5	15.a.	Explain briefly about a relational database is manually shared using application logic with neat sketch.	K2	CO5
		(OR)		
	15.b.	Summarize when In-memory storage devices is appropriate and inappropriate.		

SECTION - C (30 Marks)Answer **ANY THREE** questions**ALL** questions carry **EQUAL** Marks

(3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	Identify the Application of big data mining ?	K3	CO1
2	17	Illustrate the MapReduce computation for counting the number of occurrences for each word in a collection of documents.	K2	CO2
3	18	Inspect how to compute natural join by MapReduce?	K4	CO3
4	19	Apply UV Decomposition dimensionality reduction technique for 5-by-5 matrix.	K3	CO4
5	20	Analyze how to access data in a distributed file system with neat sketch.	K4	CO5

Z-Z-Z

END