Reg. No.															

Question Paper Code	13847

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV / DEC 2025

Seventh Semester

Computer Science and Engineering

(Common to Computer Science and Engineering (IoT))

20CSPC701 - BIG DATA ANALYTICS

Regulations - 2020

Duration: 3 Hours | Max. Marks: 100

PART - A (MCQ) (10 × 1 = 10 Marks)

Answer ALL Questions

		Marks	K – Level	CO
1.	Which of the following is not a characteristic of Big Data? (a) Volume (b) Velocity (c) Variety (d) Visibility	1	K1	CO1
2.	In the BASE model, "E" stands for: (a) Extensible (b) Eventual consistency (c) Execution (d) Evolution	1	K1	CO1
3.	Hadoop is best suited for: (a) Small datasets (b) Distributed and large datasets (c) Only structured data (d) Only relational databases	1	K1	CO2
4.	Which Hadoop ecosystem tool is used for workflow scheduling? (a) Hive (b) Pig (c) Oozie (d) HBase	1	K1	CO2
5.	Which of the following is a Hive file format? (a) RCFile (b) ORC (c) TextFile (d) All of the above	1	K1	CO3
6.	In Pig, the Relational Operators include: (a) FILTER, FOREACH, GROUP, JOIN (b) SELECT, INSERT, DELETE (c) MAP, REDUCE (d) STORE, LOAD	1	K1	CO3
7.	In MongoDB, a document is analogous to: (a) Column (b) Row (c) Table (d) Schema	1	K1	CO4
8.	In MongoDB, the equivalent of a table in RDBMS is called: (a) Collection (b) Document (c) Row (d) Field	1	K1	CO4
9.	The primary role of Zookeeper in HBase is: (a) Data storage (b) Coordination and synchronization (c) Query processing (d) Visualization	1	K1	CO5
10.	Which of the following is a NoSQL database? (a) MySQL (b) MongoDB (c) PostgreSQL (d) Oracle	1	K1	CO6

PART - B (12 × 2 = 24 Marks)

Answer ALL Questions

		Marks	K – Level	CO
11.	Interpret any two challenges in handling Big Data.	2	K2	CO1
12.	What is the difference between Big Data and Traditional BI?	2	K1	CO1
13.	What is the role of YARN in Hadoop?	2	K1	CO2
14.	How does Hadoop achieve fault tolerance?	2	K1	CO2
15.	Define Hive RCFile.	2	K1	CO3
16.	What is Piggy Bank?	2	K1	CO3
17.	Define a MongoDB document.	2	K1	CO4
18.	Interpret the command sort() in MongoDB.	2	K2	CO4
19.	Compare Sqoop import and export.	2	K2	CO5
20.	Define Lambda Architecture.	2	K1	CO5
21.	Illustrate the advantages of NoSQL with Example.	2	K2	CO6
22.	List common NoSQL vendors.	2	K1	CO6

PART - C (6 × 11 = 66 Marks)

Answer ALL Questions

23.	a)	Discuss the evolution and definition of Big Data and how it differs from traditional data.	*11*	*K2*	*CO1*
		OR			
	b)	Explain the architecture of a Data Warehouse and its relationship with Hadoop Environment.	*11*	*K2*	*CO1*
24.	a)	Explain the Hadoop architecture with a neat diagram.	*11*	*K2*	*CO2*
		OR			
	b)	Explain MapReduce framework with an example workflow.	*11*	*K2*	*CO2*
25.	a)	Explain Hive query language statements including joins, sub-queries, aggregation, GROUP BY, and HAVING.	*11*	*K2*	*CO3*
		OR			
	b)	Explain Pig Latin overview, primitive and complex data types.	*11*	*K2*	*CO3*
26.	a)	Construct the basic data types in MongoDB with examples.	*11*	*K3*	*CO4*
		OR			
	b)	Utilize Jaspersoft studio to prepare a Jasper Report using the database created using MongoDB.	*11*	*K3*	*CO4*
27.	a)	Explain in detail about different types of data visualization techniques.	*11*	*K2*	*CO5*
		OR			
	b)	Discuss in detail about Spark architecture and its components with suitable diagram.	*11*	*K2*	*CO5*
28.	a)	Identify the difference between HBASE and RDBMS.	*11*	*K3*	*CO6*
		OR			
	b)	Demonstrate the sqoop architecture and build a routine to perform data import and export.	*11*	*K3*	*CO6*