

Reg. No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code	12602
---------------------	-------

B.E. / B.Tech. - DEGREE EXAMINATIONS, APRIL / MAY 2024

Seventh Semester

Computer Science and Engineering

20CSPC701 - BIG DATA ANALYTICS

Regulations - 2020

Duration: 3 Hours

Max. Marks: 100

PART - A (10 × 2 = 20 Marks)

Answer ALL Questions

	Marks	K-Level	CO
1. Define Brewer's Theorem / CAP Theorem.	2	K1	CO1
2. State few challenges with big data.	2	K1	CO1
3. List Hadoop configuration files.	2	K1	CO2
4. What do you mean by MapReduce?	2	K1	CO2
5. Compare Partitioning and Bucketing.	2	K2	CO3
6. Recall DDL (Data Definition Language) Statements.	2	K1	CO3
7. Show how NULL values are treated in MongoDB.	2	K2	CO4
8. Enumerate the features of MongoDB.	2	K1	CO4
9. Define HBase.	2	K1	CO5
10. List out the key Components of Oozie.	2	K1	CO5

PART - B (5 × 13 = 65 Marks)

Answer ALL Questions

11. a) Explain the terminologies used in big data environment.	13	K2	CO1
OR			
b) i) Discuss the responsibilities of Data Scientist.	4	K2	CO1
ii) What are the different phases of Data Analytics Lifecycle? Explain each in detail.	9	K2	CO1
12. a) Discuss a MapReduce program to find the no of words in a text file.	13	K2	CO2
OR			
b) Summarize the functions of Hadoop distributed file system with neat diagram.	13	K2	CO2
13. a) Discuss the following in pig.	13	K2	CO3
1. Parameter substitution			
2. Diagnostic Operator			
3. Complex data types			

OR

K1 – Remember; K2 – Understand; K3 – Apply; K4 – Analyze; K5 – Evaluate; K6 – Create

12602

- b) Show the output of following queries in hive. 13 K2 CO3
1. Create database
 2. To alter database properties
 3. To display all databases
 4. To describe databases
 5. To make a database as current working database

14. a) Explain CRUD (Create, Read Update, and Delete) operations in MongoDB. 13 K2 CO4

OR

- b) Describe the process of connecting MongoDB with NoSQL database using Jasperoft. 13 K2 CO4

15. a) Summarize the techniques used for Visual Data Analysis. 13 K2 CO5

OR

- b) Discuss in detail about HBase Architecture. 13 K2 CO5

PART - C (1 × 15 = 15 Marks)

16. a) Describe the process involved in Apache Streaming and Apache Processing. 15 K2 CO6

OR

- b) Discuss the functions of sqoop components and also explain how data import and export is performed. 15 K2 CO6