Roll No.											

PRESIDENCY UNIVERSITY

BENGALURU

End Term Examinations – December 2025

Date: 08-12- 2025 **Time:** 01:00pm – 04:00pm

School: SOE	**Program**: BCA	
Course Code: CSA3004	**Course Name**: Big Data Analytics	
Semester: V	**Max Marks**: 100	**Weightage**: 50%

CO - Levels	**CO1**	**CO2**	**CO3**
Marks	**30**	**40**	**30**

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **10Q x 2M=20M**

1.	Describe the commands for scanning, counting, and truncating tables in HBase.	**2 Marks**	**L1**	**CO2**
2.	Compare and contrast traditional data management approaches with Big Data approaches.	**2 Marks**	**L2**	**CO1**
3.	Identify the advantages and limitations of No-SQL databases in handling Big Data.	**2 Marks**	**L1**	**CO1**
4.	Explain the architecture of HDFS and the roles of Name Node and Data Node.	**2 Marks**	**L2**	**CO1**
5.	Explain do Combiners and Partitioners improve the efficiency of MapReduce jobs?	**2 Marks**	**L2**	**CO1**
6.	Describe the architecture of YARN and the roles of its various components.	**2 Marks**	**L2**	**CO2**
7.	Identify the different data types supported by Hive, and how are they used?	**2 Marks**	**L1**	**CO2**
8.	List and describe the common Hive DML (Data Manipulation Language) commands.	**2 Marks**	**L1**	**CO2**

9.	List the commands used for creating and listing tables in HBase.	2 Marks	L1	CO2
10.	Explain the difference between structured, unstructured, semi structured, and quasi-structured data.	2 Marks	L2	CO1

Part B

Answer the Questions. **Total Marks 80M**

11.	**a.**	Identify and explain three major challenges associated with Big Data.	**10 Marks**	**L2**	**CO1**
		Or			
12.	**a.**	Classify the process of file writing in HDFS.	**10 Marks**	**L2**	**CO1**

13.	**a.**	Explain the use of Hive for real-time data processing versus batch processing along with the limitations of Hive in handling real-time data, and how can they be addressed?	**10 Marks**	**L2**	**CO2**
		Or			
14.	**a.**	Describe the architecture of HBase and its key components.	**10 Marks**	**L2**	**CO2**

15.	**a.**	Explain the commands for disabling and enabling tables in HBase.	**10 Marks**	**L2**	**CO3**
		Or			
16.	**a.**	Describe the commands for scanning, counting, and truncating tables in HBase.	**10 Marks**	**L2**	**CO3**

17.	**a.**	Identify the "Four Vs" of Big Data and how do they contribute to its importance?	**10 Marks**	**L1**	**CO1**
		Or			
18.	**a.**	Outline the main types of No-SQL databases and how do they differ from relational databases?	**10 Marks**	**L1**	**CO1**

19.	**a.**	Describe the architecture of Hive and its key components.	**20 Marks**	**L2**	**CO2**
		Or			
20.	**a.**	Differentiate between big data and Traditional data.	**20 Marks**	**L2**	**CO2**

21.	**a.**	Outline the process of reading in HDFS.	**20 Marks**	**L1**	**CO3**
		Or			
22.	**a.**	Describe HDFS Federation and list how the architecture limitations of HDFS architecture are resolved by the same.	**20 Marks**	**L1**	**CO3**