



Roll No.											
----------	--	--	--	--	--	--	--	--	--	--	--

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – October 2025

Date: 07-10-2025

Time: 11.45am to 01.15pm

School: SOIS	Program: BCA Data Science	
Course Code: CSA3004	Course Name: Big Data Analytics	
Semester: V	Max Marks: 50	Weightage: 25%

CO - Levels	CO1
Marks	50

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.

5Q x 2M=10M

1	Define Big Data and list some examples on the same.	2 Marks	L1	CO1
2	Name the characteristics of Big Data	2 Marks	L1	CO1
3	State the advantages that the MapReduce paradigm provides for processing large datasets?	2 Marks	L1	CO1
4	Classify Big Data Applications.	2 Marks	L2	CO1
5	Differentiate between SQL & No-SQL	2 Marks	L2	CO1

Part B

Answer the Questions.

Total Marks 40M

6.	a.	Distinguish the advantages that the MapReduce paradigm provides for processing large datasets?	10 Marks	L2	CO1
Or					
7.	a.	Explain the various daemons of Yarn.	10 Marks	L2	CO

							1
--	--	--	--	--	--	--	---

8.	a.	Classify the key motivations behind the development of Hadoop, and how has it evolved since its inception?	10 Marks	L2	CO 1
Or					
9.	a.	Examine how Secondary Name Node serves as the helping node to Name Node	10 Marks	L3	CO 1

10.	a.	Apply the scenario below and give an efficient approach to resolve it. Imagine you're working for a company that needs to understand the web traffic patterns on their site. To do this, you decide to use a sample dataset of web logs. By employing MapReduce, you can efficiently process and analyze this data to uncover insights about user behavior and site performance.	10 Marks	L3	CO 1
Or					
11.	a.	Apply the below scenario and implement a MapReduce job that processes these documents to count how many times each word appears across the entire dataset. Imagine you're working on a project where you need to analyze a large collection of text documents to determine the frequency of each word. To handle this efficiently, you decide to use a MapReduce approach. In this scenario, you'll use a sample dataset of text documents.	10 Marks	L3	CO 1

12.	a.	List the different features of Hadoop and give justifications for each.	10 Marks	L1	CO 1
Or					
13.	a.	Explain the architecture of HDFS in detail.	10 Marks	L2	CO 1