



sRoll No.									
--------------	--	--	--	--	--	--	--	--	--

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – October 2025

Date: 08-10-2025

Time: 11.45am to 01.15pm

School: SOCSE	Program: CBD	
Course Code : CBD2508	Course Name: Big Data Technologies	
Semester: V	Max Marks: 50	Weightage: 25%

CO - Levels	C01	C02	C03	C04	C05
Marks	33	17			

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	List and explain types of bigdata.	2 Marks	L1	C01
2	Identify the responsibilities of Job Tracker	2 Marks	L1	C01
3	List the features of Hadoop 2.X version.	2 Marks	L1	C01
4	Identify the various daemons of Yarn	2 Marks	L1	C01
5	List two components of PIG	2 Marks	L1	C02

Part B

Answer the Questions.

Total Marks 40M

6.	a.	Describe the NoSQL Database. Explain the types of NoSQL.	5 Marks	L2	C01
	b.	with neat diagram Illustrate Hadoop Master/Slave Architecture and Daemon processes in Hadoop Core	5 Marks	L2	C01

Or

7.	a.	Distinguish the ways in which rack awareness improve fault tolerance in a distributed environment?	5 Marks	L2	C01
	b.	With neat diagram, Explain the various daemons of Yarn.	5 Marks	L2	C01

8.	a.	Student_ID	Batch	Semester	Course_Enrolled	15 Marks	L3	CO1
		PRN2019001	2019	5	ReactJS			
		PRN2018001	2018	7	Cloud Computing			
		PRN2018018	2018	7	ReactJS			
		PRN2019025	2019	5	Cloud Computing			
		PRN2019026	2019	5	Cloud Computing			
		PRN2018011	2018	7	DevOps			
		PRN2018021	2018	7	DevOps			
		PRN2018027	2018	5	ReactJS			
		PRN2019028	2019	7	DevOps			
		PRN2018033	2018	7	DevOps			
<p>The dataset below gives the details of student enrollments from 2018 and 2019 batches for the Value-Added Courses (VAC): ReactJS, Cloud Computing and DevOps.</p> <p>Write a Hadoop MapReduce application to process the data and to display the total student enrollments for each course.</p>								

Or

9.	a.	ICICI bank wants to analyse its customer data set to analyse transactions in different outlets. The data set contains information of CustomerId,City,Amount,Outlet. Write a map reduce program to find the maximum transaction amount for each outlet.	15 Marks	L3	CO1
		Dataset :			
		212,Bangalore,2000,Pantalo			
		216,Bangalore,8000,Reliance			
		112,Bangalore,9000,Pantalo			
		218,Bangalore,7000,Pantalo			
		221,Bangalore,10000,Reliance			
		316,Bangalore,9000,Reliance			
		331,Bangalore,9000,Max			
		116,Bangalore,2000,Reliance			
		296,Bangalore,6000,Max			

10.	a.	Explain the architecture of HBase with a neat diagram.	7 Marks	L2	CO2
		b. Explicate with an example in Hive i) Hive Joining tables ii) Hive bucketing.			
11.	a.	Differentiate between Hive & HBase	7 Marks	L2	CO2
		b. Describe the following Hadoop Ecosystem components: a) Apache Pig b) Apache Mahout			