Roll No.												

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – October 2025

Date: 29-10-2025 **Time:** 02.30pm to 04.00pm

School: SOCSE	**Program:** B.Tech - ISE	
Course Code : CSE3509	**Course Name:** Emerging Technologies in Big Data	
Semester: VII	**Max Marks**: 50	**Weightage**: 25%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**	**CO5**
Marks	**26**	**24**			

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	What is the function of the NameNode in HDFS?	**2 Marks**	**L1**	**CO1**
2	Define structures, unstructured, and semi-structures data.	**2 Marks**	**L1**	**CO1**
3	Name two common schedulers used by YARN.	**2 Marks**	**L2**	**CO1**
4	Mention a basic Sqoop command for data import.	**2 Marks**	**L2**	**CO2**
5	What is ‘bucketing’ in Hive?	**2 Marks**	**L2**	**CO2**

Part B

Answer the Questions. **Total Marks 40M**

6.	**a.**	Apply your understanding of the Hadoop MapReduce execution pipeline to explain how combiners and partitioners optimize job performance.	**20 Marks**	**L3**	**CO1**
		Or			
7.	**a.**	Explain the architecture of HDFS, detailing the roles of NameNode, DataNode, block management, and replication. How does this architecture ensure data reliability and fault tolerance?	**14 Marks**	**L2**	**CO1**
	b.	Describe the file write and read process in HDFS.	**6 Marks**	**L2**	**CO1**

8.	**a.**	Explain Hive's data types, partitioning, and bucketing with examples.	**10 Marks**	**L2**	**CO2**
	b.	Demonstrate through examples how Hive joins work, including inner and outer joins.	**10 Marks**	**L2**	**CO2**
		Or			
9.	**a.**	Explain Sqoop’s import/export features, focusing on import of all tables and incremental data transfer. When would you prefer sqoop vs other ingestion tools.	**20 Marks**	**L2**	**CO2**