



# PRESIDENCY UNIVERSITY

BENGALURU

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

## End - Term Examinations - December 2025

Date: 09- 12-2025

Time: 01:00pm - 04:00pm

<b>School:</b> SOCSE	<b>Program:</b> B.Tech		
<b>Course Code :</b> CSE3509	<b>Course Name:</b> Emerging technologies in Big Data		
<b>Semester:</b> VII	<b>Max Marks:</b> 100	<b>Weightage:</b> 50%	

CO - Levels	C01	C02	C03	C04	C05
<b>Marks</b>	<b>26</b>	<b>46</b>	<b>48</b>		

### 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.	Define Big Data and list its key characteristics.	2 Marks	L1	C01
2.	What are the key challenges Big Data addresses?	2 Marks	L1	C01
3.	Explain structured, unstructured and semi-structured data.	2 Marks	L2	C01
4.	Name two core components of HDFS and their roles.	2 Marks	L1	C02
5.	What is the purpose of MapReduce in Hadoop?	2 Marks	L1	C02
6.	Describe the YARN architecture components.	2 Marks	L2	C02
7.	What role does Sqoop play in Hadoop?	2 Marks	L1	C03
8.	Define Hive and its primary use.	2 Marks	L1	C03
9.	What are basic Hive table partitioning and bucketing?	2 Marks	L2	C03
10.	Outline key features of Apache Spark.	2 Marks	L2	C03

## Part B

Answer the Questions.

Total Marks 80M

11.	a.	Compare traditional data processing with Big Data approaches and explain Hadoop's role in addressing challenges.	10 Marks	L2	C01
	b.	Discuss the significance of the Four Vs of Big Data.	10 Marks	L2	C01
<b>Or</b>					
12.	a.	Explain real-world use cases where Big Data analytics benefits organizations.	10 Marks	L2	C01
	b.	Describe the challenges faced by traditional data systems and how Hadoop resolves them.	10 Marks	L2	C01
<b>Or</b>					
13.	a.	Describe the HDFS architecture with emphasis on NameNode and DataNode roles.	10 Marks	L3	C02
	b.	Explain YARN architecture and scheduler policies like FIFO and Capacity schedulers.	10 Marks	L2	C02
<b>Or</b>					
14.	a.	Discuss key features introduced in Hadoop 2.0 and their benefits.	10 Marks	L2	C02
	b.	Explain the MapReduce execution pipeline highlighting the roles of combiner and partitioner.	10 Marks	L2	C02
<b>Or</b>					
15.	a.	Explain Hive data types, partitioning and bucketing with examples.	10 Marks	L2	C03
	b.	Describe the functions of Sqoop and Flume, illustrating with use cases.	10 Marks	L3	C03
<b>Or</b>					
16.	a.	Discuss Hive DDL and DML commands including sorting and joining tables with examples.	10 Marks	L3	C03
	b.	Explain how Spark SQL integrates with RDDs and the use of user-defined functions (UDFs).	10 Marks	L3	C03
<b>Or</b>					
17.	a.	Explain the high availability feature of NameNode in Hadoop 2.0 and how it addresses single point of failure.	10 Marks	L3	C02
	b.	Describe the roles of ResourceManager and NodeManager in YARN.	10 Marks	L3	C02
<b>Or</b>					
18.	a.	Illustrate how Hive handles complex queries using joins with examples.	10 Marks	L3	C03
	b.	Discuss the use of user-defined functions (UDFs) in Spark SQL with an example.	10 Marks	L3	C03