

GUJARAT TECHNOLOGICAL UNIVERSITY

BE- SEMESTER-VII EXAMINATION – WINTER 2025

Subject Code:2171607

Date:24-11-2025

Subject Name:Big Data Analytics

Time:10:30 AM TO 01:00 PM

Total Marks:70

Instructions:

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Simple and non-programmable scientific calculators are allowed.

		MARKS
Q.1	(a) As per your point of view, How Big data analytics can be useful in development of smart cities?	03
	(b) Discuss Four V's of Big Data.	04
	(c) Draw and explain HDFS Architecture. Explain the function of NameNode and DataNode. What is a Secondary Namenode? Is it a substitute to the Namenode?	07
Q.2	(a) Explain the following commands of HDFS: i) copyFromLocal ii) setrep iii) checksum	03
	(b) Discuss with an example that big data analytics is helpful in increasing business revenue.	04
	(c) Explain working of MapReduce with reference to 'WordCount' program for a file having input as below: Red apple Red wine Green apple Green peas Pink rose Blue whale Blue sky Green city Clean city	07
	OR	
	(c) Write a MapReduce code for Word Count.	07
Q.3	(a) "Moving Computation is Cheaper than Moving Data" - Justify the sentence.	03
	(b) What is Hadoop Ecosystem? Discuss various components of Hadoop Ecosystem.	04
	(c) Explain Spark Architecture with all its components in detail.	07
	OR	
Q.3	(a) What is HBase? Explain storage mechanism of HBase with an example.	03
	(b) Discuss Machine Learning with MLlib in SPARK.	04
	(c) What are the problems related to Map Reduce data storage? How Apache Spark solves it using Resilient Distributed Dataset? Explain RDDs in detail.	07
Q.4	(a) Explain any three HiveQL DDL command with its syntax and example.	03
	(b) Define Zookeeper. List and explain its benefits.	04
	(c) What is NoSQL database? List the differences between NoSQL and relational databases. Explain in brief various types of NoSQL databases in practice.	07

OR

- Q.4** (a) Mention advantages of using NoSQL databases. **03**
(b) Explain following terms with reference to MongoDB Query Language: Database, Collection, Document, Datatypes **04**
(c) Draw and explain Architecture of APACHE HIVE. Explain various data insertion techniques in HIVE with example. **07**

- Q.5** (a) Explain scaling feature of MongoDB. **03**
(b) Explain following in brief with respect to Mongo DB : **04**
1) Collections and documents
2) Indexing and retrieval
(c) Explain CRUD operations in MongoDB. **07**

OR

- Q.5** (a) Explain MongoDB sharding process. **03**
(b) Justify: “SPARK is faster than MapReduce”. **04**
(c) Explain following CRUD operations on MongoDB with necessary syntax or example: **07**
Create a database, Drop a database, Create and display collection, Insert a document in collection, Finding document from a collection, Updating a document, Deleting a document
