

**GUJARAT TECHNOLOGICAL UNIVERSITY****BE - SEMESTER-VII (NEW) EXAMINATION – SUMMER 2022****Subject Code:2171607****Date:01/06/2022****Subject Name:Big Data Analytics****Time:02:30 PM TO 05:00 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.

		<b>MARKS</b>	
<b>Q.1</b>	(a) Define big data. How Big Data and Hadoop are related?	<b>03</b>	
	(b) Discuss with an example that big data analytics is helpful in increasing business revenue.	<b>04</b>	
	(c) Explain HDFS Architecture with all its components in detail.	<b>07</b>	
<b>Q.2</b>	(a) Discuss various modes in which Hadoop runs.	<b>03</b>	
	(b) What are different configuration files in Hadoop? Explain their roles.	<b>04</b>	
	(c) Why data locality is required in Hadoop? Explain with an example.	<b>07</b>	
<b>OR</b>			
<b>Q.3</b>	(c) Explain following commands of HDFS with syntax and example: (i) get (ii) cp (iii) chown	<b>07</b>	
	(a) Discuss different types of tables available in Hive.	<b>03</b>	
	(b) Write a short note on Hive Data Types.	<b>04</b>	
<b>Q.3</b>	(c) What is HBase? List its features. Differentiate: (i) HBase and Hadoop (ii) HBase and RDBMS	<b>07</b>	
	<b>OR</b>		
	(a) Explain any three HiveQL DDL command with example.	<b>03</b>	
<b>Q.3</b>	(b) Write a short note on Zookeeper.	<b>04</b>	
	(c) What is Hive? List its features. Explain working of Hive with proper diagram.	<b>07</b>	
	(a) What is MongoDB? Explain its key features.	<b>03</b>	
<b>Q.4</b>	(b) Explain CRUD operations in MongoDB with example.	<b>04</b>	
	(c) What is NoSQL? List merits of NoSQL over SQL. Explain any two NoSQL database types with example.	<b>07</b>	
	<b>OR</b>		
<b>Q.4</b>	(a) Differentiate between NoSQL and RDBMS.	<b>03</b>	
	(b) Explain scaling feature of MongoDB.	<b>04</b>	
	(c) What is data serialization? Explain structured, unstructured and semi-structured data with example.	<b>07</b>	
<b>Q.5</b>	(a) Compare MongoDB and RDBMS.	<b>03</b>	
	(b) What is shuffling and sorting in Map Reduce. Give an example.	<b>04</b>	
	(c) Explain Spark Architecture with all its components in detail.	<b>07</b>	
<b>OR</b>			
<b>Q.5</b>	(a) Explain the role of JobTracker in Hadoop.	<b>03</b>	
	(b) Explain PIG architecture with proper diagram.	<b>04</b>	
	(c) What are the problems related to Map Reduce data storage? How Spark solves them? Explain.	<b>07</b>	

\*\*\*\*\*