

GUJARAT TECHNOLOGICAL UNIVERSITY**BE - SEMESTER-VII (NEW) EXAMINATION – WINTER 2021****Subject Code:3170722****Date:23/12/2021****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) Explain types of Big Data	03
	(b) Describe Traditional vs. Big Data business approach.	04
	(c) What is Big Data? Explain Challenges of Conventional System.	07
Q.2	(a) Explain basic Components of Analyzing the Data with Hadoop.	03
	(b) What is Map Reduce and explain How Map Reduce Works?	04
	(c) Brief Anatomy of a Map Reduce Job run and Failures.	07
	OR	
	(c) Describe Map Reduce Types and Formats.	07
Q.3	(a) Explain NoSQL data architecture.	03
	(b) Elaborate Key-value stores, Graph stores, Column family stores & Document stores.	04
	(c) Describe analyzing big data with a shared-nothing architecture.	07
	OR	
Q.3	(a) Difference between master-slave versus peer-to-peer distribution models.	03
	(b) What is a big data NoSQL? Explain in details.	04
	(c) Which Four ways that NoSQL systems handle big data problems.	07
Q.4	(a) What is Stream Computing and Sampling Data in a Stream.	03
	(b) Write the application of RTAP.	04
	(c) Explain Stream Data Model and Architecture.	07
	OR	
Q.4	(a) How Graph Analytics used in Big Data.	03
	(b) Write a short note on Decaying Window.	04
	(c) Explain with example: How to perform Real Time Sentiment Analysis of any product.	07
Q.5	(a) What is the use of Pig and Hive in Big Data?	03
	(b) Describe data processing operators in Pig.	04
	(c) Describe HBase and ZooKeeper in details.	07
	OR	
Q.5	(a) Explain HIVE services.	03
	(b) Write application of writing Spark.	04
	(c) Describe any application that you know related to enhance particular business using big data and explain how it is important as a business prospective.	07

GUJARAT TECHNOLOGICAL UNIVERSITY**BE - SEMESTER–VII (NEW) EXAMINATION – SUMMER 2022****Subject Code:3170722****Date:18/06/2022****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.

- Q.1** (a) What is Big data Analytics? List applications of it. **03**
 (b) Explain 4 V's of Big data. **04**
 (c) What is Hadoop? Briefly explain the core components of it. **07**

- Q.2** (a) How is Big data and Hadoop related? **03**
 (b) How HDFS is different from traditional NFS? **04**
 (c) Draw and Explain HDFS architecture. How can you restart NameNode and all the daemons in Hadoop? **07**

OR

- (c) What is MapReduce? Explain working of various phases of MapReduce with appropriate example and diagram. **07**

- Q.3** (a) What do you mean by job scheduling in Hadoop? List different schedulers in Hadoop. **03**
 (b) What are WAL, MemStore, Hfile and Hlog in HBase? **04**
 (c) Explain the architecture of HBase. **07**

OR

- Q.3** (a) What is Zookeeper? **03**
 (b) Differentiate between HIVE and HBASE. **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**

- Q.4** (a) What is Pig? **03**
 (b) What are the features of MongoDB? **04**
 (c) Explain the concept of regions in HBase and storing Big data with HBase. **07**

OR

- Q.4** (a) How MongoDB is better than SQL database? **03**
 (b) Explain the Data Model of HBase. **04**
 (c) Explain Pig data Model in detail and Discuss how it will help for effective data flow. **07**

- Q.5** (a) Write a short note on Spark. **03**
 (b) Write difference between MongoDB and Hadoop. **04**
 (c) Explain CRUD operations in MongoDB. **07**

OR

- Q.5** (a) What are the features of MongoDB? **03**
 (b) Explain Replication and scaling feature of MongoDB. **04**
 (c) What is RDD? State and Explain RDD operations. **07**
