

GUJARAT TECHNOLOGICAL UNIVERSITY**BE - SEMESTER-VII (NEW) EXAMINATION – SUMMER 2021****Subject Code:2171607****Date:03/08/2021****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.

		MARKS
Q.1	(a) Define Big Data. Explain with suitable example.	03
	(b) Explain 4 V's properties of Big Data.	04
	(c) Enlist the advantages of Hadoop. Draw and explain Hadoop Architecture with components.	07
Q.2	(a) Explain unstructured, semi-structured and structured data with one example of each.	03
	(b) What are the benefits of Big Data? Discuss challenges under Big Data.	04
	(c) Define HDFS. Discuss the HDFS Architecture and HDFS commands in brief.	07
OR		
	(c) Write Map Reduce code for counting occurrences of words in the input text file.	07
Q.3	(a) Explain in detail: Input and Output of MapReduce.	03
	(b) Explain CRUD operations in MongoDB with syntax.	04
	(c) Define RDD. Discuss transformations and actions in RDDs. State and explain RDD operations in brief.	07
OR		
Q.3	(a) Explain the following phases of MapReduce. i) Map Phase ii) Shuffle and Sort Phase iii) Reduce Phase	03
	(b) Implement Join operation using Hive.	04
	(c) Explain Pig Data Model in details. Discuss its effectiveness for data flow.	07
Q.4	(a) Define Zookeeper. List and explain its benefits.	03
	(b) Explain the terms: i) Metastore in Hive ii) setrep iii) Fair Scheduler iv) YARN	04
	(c) Discuss the concepts of regions in HBase and storing Big Data with HBase.	07
OR		
Q.4	(a) "Moving Computation is Cheaper than Moving Data" - Justify the sentence.	03
	(b) Differentiate: Apache pig Vs Map Reduce.	04
	(c) Draw and discuss the Architecture of Hive.	07
Q.5	(a) List out key elements of Ingress and Egress and explain them in detail.	03
	(b) Write a short note on NoSQL databases.	04
	(c) Discuss Big Data in Healthcare, Agriculture and Manufacturing.	07
OR		
Q.5	(a) Explain the terms: i) Replication Factor ii) Data Serialization iii) Heartbeat Message	03
	(b) List the differences between NoSQL and relational databases.	04
	(c) Discuss Big Data in Transportation, Education, Smart Cities.	07
