Seat No.: _____   Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY
MCA INTEGRATED– SEMESTER - VIII EXAMINATION- SUMMER-2023

Subject Code: 4480601   **Date: 27/06/2023**
Subject Name: Big Data
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. Use of simple calculators and non-programmable scientific calculators are permitted.

Q.1	**(a)**	Define. 1. OLTP 2. Filtering 3. Namenode 4. Big Table 5. HDFS 6. Hadoop 7. Data Mining	07
	(b)	What is data? Explain types of data in detail with example.	07
Q.2	**(a)**	State the phases of Big Data analytics lifecycle.	07
	(b)	Differentiate between traditional BI and Big Data BI	07
		OR	
	(b)	What is MongoDB? Explain how MongoDB achieves horizontal scaling	07
Q.3	**(a)**	What is MapReduce? State the key considerations while developing a MapReduce algorithm.	07
	(b)	Explain CAP and BASE theorem in detail.	07
		OR	
Q.3	**(a)**	Differentiate between 1. RDBMS and HADOOP 2. Task Parallelism and Data Parallelism	04 03
	(b)	Discuss A/B Testing, Correlation and Regression using suitable example.	07
Q.4	**(a)**	Explain HBase architecture in detail.	07
	(b)	Distinguish column oriented, document oriented, and key/value-oriented databases based on their features and storage architecture.	07
		OR	
Q.4	**(a)**	What is Business Architecture? Explain how Big Data is related to it at various levels of the organization.	07
	(b)	List and explain statistical analysis in detail.	07
Q.5	**(a)**	What is visual analysis? Explain its types in brief	07
	(b)	Explain any 7 commands of Hbase with example.	07
		OR	
Q.5	**(a)**	Explain Outlier Analysis and Semantic Analysis.	07
	(b)	Illustrate Classification and Clustering techniques.	07

****