Seat No.: _____   Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY
MCA - SEMESTER– V EXAMINATION – WINTER 2019
Subject Code: 3650010   **Date: 19/11/2019**
Subject Name: Big Data Tools
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.**

Q.1	**(a)**	Define :	**07**
		1. Unstructured data	
		2. Sharding	
		3. Replication	
		4. Cluster	
		5. CAP theorem	
		6. Master – slave architecture	
		7. MongoDB	
	(b)	Give Syntax:	**07**
		1. aggregate in MongoDB.	
		2. Creating index in MongoDB.	
		3. Arranging data in ascending order in Pig.	
		4. Limit and skip in MongoDB.	
		5. Grouping the data in Pig.	
		6. Creating view in Hive.	
		7. Copy a local file into HDFS.	
Q.2	**(a)**	What is big data? List and explain characteristics of big data in detail.	**07**
	(b)	Write Pig Script for word count.	**07**
		OR	
	(b)	Write Hive query for word count.	**07**
Q.3	**(a)**	Give differences :	
		1. SQL, NoSQL and NewSQL	**04**
		2. Pig and Hive	**03**
	(b)	Explain Map – Reduce process with proper example.	**07**
		OR	
Q.3	**(a)**	Explain daemons of HDFS and Map Reduce.	**07**
	(b)	Explain mongoimport and mongexport in detail with syntax and e.g.	**07**
Q.4	**(a)**	Explain Hive Architecture in detail.	**07**
	(b)	What is NoSQL? Explain types of NoSQL databases in detail.	**07**
		OR	
Q.4	**(a)**	Explain Hadoop eco system in detail.	**07**
	(b)	What is Pig? Explain Pig Anatomy and Pig Philosophy in detail.	**07**
Q.5	**(a)**	What is spark? Give difference between Hadoop and Spark. Also explain RDD in detail.	**07**
	(b)	Explain static and dynamic partition in Hive.	**07**
		OR	
Q.5	**(a)**	Write short note on MLlib, Tensor Flow and Theone.	**07**
	(b)	Explain PiggyBank in detail with e.g.	**07**

****