GUJARAT TECHNOLOGICAL UNIVERSITY
BE - SEMESTER– VII (New) EXAMINATION – WINTER 2019

Subject Code: 2171607   **Date: 23/11/2019**

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. Enlist its importance over traditional database system.	**03**
	(b)	List out 4 V's and Explain in detail.	**04**
	(c)	List out the applications of Big data and explain them in details.	**07**
Q.2	**(a)**	Implement Join operation using Hive.	**03**
	(b)	List out key elements of Ingress and Egress and explain them in detail.	**04**
	(c)	Define HDFS. Discuss the HDFS architecture and HDFS commands in brief.	**07**

OR

	(c)	Write a short note on NoSQL databases. List the differences between NoSQL and relational databases.	**07**
Q.3	**(a)**	What is Zookeeper? List the benefits of it.	**03**
	(b)	Explain Aggregation and Indexing for MongoDB.	**04**
	(c)	What are the components of Spark? Also state the features of Spark.	**07**

OR

Q.3	**(a)**	Explain in detail: Input and Output of MapReduce.	**03**
	(b)	Write a short note on Pig.	**04**
	(c)	Explain CRUD operations in MongoDB.	**07**
Q.4	**(a)**	Explain any three HiveQL DDL command with its syntax and example.	**03**
	(b)	What is HBase? List out and explain the basic concepts of HBase in detail.	**04**
	(c)	Explain Replication and scaling feature of MongoDB.	**07**

OR

Q.4	**(a)**	Explain the following commands of HDFS: i) copyFromLocal ii) setrep iii) checksum	**03**
	(b)	As per your point of view, How Big data analytics can be useful in development of smart cities?	**04**
	(c)	Explain working of Hive with proper steps and diagram.	**07**
Q.5	**(a)**	Explain Metastore in Hive.	**03**
	(b)	Differentiate: Apache pig Vs Map Reduce.	**04**
	(c)	What are the problems related to Map Reduce data storage? How Apache Spark solves it using Resilient Distributed Dataset? Explain RDDs in detail.	**07**

OR

Q.5	**(a)**	Write a MapReduce code for Word Count.	**03**
	(b)	"Moving Computation is Cheaper than Moving Data", Justify the sentence.	**04**
	(c)	Explain Map-Reduce framework in detail. Draw the architectural diagram for Physical Organization of Computer Nodes.	**07**

****