FirstRanker.com
Firstranker's choice
www.FirstRanker.com www.FirstRanker.com

Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–VII (NEW) EXAMINATION – WINTER 2017

Subject Code: 2171607 Date:02/11/2017

Subject Name: Big Data Analytics(Departmental Elective - II)

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)**	How Big Data Analytics can be useful in the development of smart cities.	**03**
	(b)	What is Big data? Discuss it in terms of volume and velocity.	**04**
	(c)	What are the advantages of Hadoop? Explain Hadoop Architecture and its Components with proper diagram.	**07**
Q.2	**(a)**	Explain Storage mechanism in HBase.	**03**
	(b)	In Map Reduce how Job Scheduling is done in case of the Fair Scheduler?	**04**
	(c)	Draw and explain HDFS Architecture. Explain the function of NameNode and DataNode. What is a Secondary Namenode? Is it a substitute to the Namenode?	**07**
		OR	
	(c)	Explain following commands of HDFS with syntax and at least one example of each. (i) get (ii) cp (iii) chown	**07**
Q.3	**(a)**	Explain Metastore in Hive.	**03**
	(b)	Explain Map-reduce framework in brief.	**04**
	(c)	Explain the concept of Blocks and Heartbeat Message in HDFS Architecture. What are the benefits of block transfer?	**07**
		OR	
Q.3	**(a)**	What is Zookeeper? List the benefits of it.	**03**
	(b)	Explain the 5 P's of Data science in brief.	**04**
	(c)	Explain Spark components in detail. Also list the features of spark.	**07**
Q.4	**(a)**	Differentiate NoSQL and relational databases?	**03**
	(b)	What do you mean by HiveQL Data Definition Language?	**04**
	(c)	Explain following for MongoDB. (i) Indexing (ii) Aggregation	**07**
		OR	
Q.4	**(a)**	Explain any three HiveQL DDL command with its syntax and example.	**03**
	(b)	Explain scaling feature of MongoDB.	**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**
Q.5	**(a)**	Compare Raw oriented and Column Oriented database structures.	**03**
	(b)	Discuss how Pig data model will help in effective data flow.	**04**
	(c)	Explain how HBase uses Zookeeper to Build Applications with Zookeeper.	**07**
		OR	

www.FirstRanker.com

FirstRanker.com
Firstranker's choice

Q.5	**(a)**	What is Spark?	**03**
	(b)	Differentiate: Apache pig Vs Map Reduce.	**04**
	(c)	Explain CRUD operations in MongoDB.	**07**

****

www.FirstRanker.com