

**GUJARAT TECHNOLOGICAL UNIVERSITY****BE - SEMESTER-VIII(NEW) EXAMINATION – SUMMER 2019****Subject Code:2180710****Date:17/05/2019****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.

- Q.1**
- (a) What is Big Data? List Advantages of it. **03**
  - (b) Explain characteristics of Big Data. **04**
  - (c) What is the role of a “combiner” in the Map Reduce framework? Explain with the help of an example. **07**
- Q.2**
- (a) What is Name node & Data node in Hadoop Architecture. **03**
  - (b) Explain traditional stream processing Architecture. **04**
  - (c) What is data serialization? With proper examples discuss and differentiate structured, unstructured and semi-structured data. **07**
- OR**
- Q.3**
- (c) Difference between HDFS and HBase. **07**
  - (a) Draw Map Reduce Execution pipeline architecture. **03**
  - (b) Explain working of Hive with proper steps and diagram. **04**
  - (c) What do you mean by HiveQL Data Definition Language? Explain any three HiveQL DDL command with its syntax and example. **07**
- OR**
- Q.3**
- (a) What is Zookeeper? List the Benefits and Limitations of it. **03**
  - (b) Explain Storage mechanism in HBase. **04**
  - (c) Explain the HiveQL-Select-Order By with suitable example. **07**
- Q.4**
- (a) Explain the following spark components i) MLib ii) GraphX **03**
  - (b) What is RDD in Apache Spark? Make a note on why RDD is better than Map Reduce data storage? **04**
  - (c) Explain Feature Extraction and Transformation in detail. **07**
- OR**
- Q.4**
- (a) What is Apache Spark? What are the advantages of using Apache Spark over Hadoop? **03**
  - (b) Explain in brief four major libraries of Apache Spark. **04**
  - (c) Compare: Map Reduce v/s Apache PIG. Explain in brief four major libraries of Apache Spark. **07**
- Q.5**
- (a) Compare: SQL v/s NoSQL. **03**
  - (b) Explain MongoDB shell and how to run the shell. **04**
  - (c) What is Mongo DB? Explain in brief key features of Mongo DB. Show basic CRUD operations in Mongo DB with proper example. **07**
- OR**
- Q.5**
- (a) Why NoSQL? Explain Advantages & Disadvantages. **03**
  - (b) Write a note on Use of NoSQL in Industry. **04**
  - (c) Explain Mongo DB>Create Database & Collection and Drop-Database & Collection. **07**

\*\*\*\*\*