

**GUJARAT TECHNOLOGICAL UNIVERSITY**  
**ME - SEMESTER-1 (NEW) EXAMINATION – WINTER 2018**

**Subject Code: 3712313****Date: 04/01/2019****Subject Name: Big Data Analytics****Time: 02:30 PM To 05:00 PM****Total Marks: 70****Instructions:**

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full mark

- Q.1** (a) What is Big Data? Mention its benefits and challenges. How Big Data can be used for Crowd sourcing analytics? (Discuss application) **07**
- (b) Explain the concept of Blocks and Heartbeat Message in HDFS Architecture. Explain HDFS Commands with example. **07**
- Q.2** (a) What is replica? Explain master-slave (MS) and Peer-to-Peer (P2P) replication. Explain sharding in MS and P2P. **07**
- (b) Write characteristics of Big Data. Apply Big data techniques to Credit Risk Management. **07**
- OR**
- (b) Discuss failure handling in Hadoop, Map Reduce and YARN in detail. **07**
- Q.3** (a) Draw Pig Architecture and explain components in detail. Discuss following Pig diagnostic operators in detail **07**
- i) Illustrate
- ii) Explain
- iii) Dump
- (b) Describe Read, Write (write in Memstore, write in HFile) and compaction mechanism in Hbase. **07**
- OR**
- Q.3** (a) Differentiate Map-Reduce and two stage Map-Reduce? Explain following phases of Map-Reduce with one common coding example. **07**
- (i) Map Phase (ii) Combiner Phase (iii) Shuffle and Sort Phase (iv) Reducer Phase. Write commands to compile and run the code.
- (b) Explain Hadoop Streaming and Hadoop Pipes in detail. **07**
- Q.4** (a) What is Apache Pig? List out Features and applications of Pig. Compare Pig with SQL, Hive and MapReduce. **07**
- (b) Explain CRUD operations in MongoDB with example query. **07**
- OR**
- Q.4** (a) Explain data Integrity and data Compression in HDFS. **07**
- (b) Explain CAP theorem and discuss BASE with respect to that. **07**
- Q.5** (a) Explain architecture and data model of Cassandra in detail. **07**
- (b) Explain Avro serialization and De-serialization with code generation. Write steps for Compiling and Running the Example Code. **07**
- OR**
- Q.5** (a) Write a short note on Unit Testing Hadoop MapReduce Jobs With MRUnit. **07**
- (b) What is Hive and HiveQL? Mention data types and file formats of Hive. Explain HiveQL data manipulation queries. **07**