

Seat No.: _____

Enrolment No. _____

GUJARAT TECHNOLOGICAL UNIVERSITY

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

Subject Code: 2171607

Date: 29/04/2017

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

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 marks.

- Q.1** (a) What is Big data? Discuss it in terms of four dimensions, volume, velocity, variety and veracity. **07**
- (b) Define HDFS. Discuss the HDFS Architecture and HDFS Commands in brief. **07**
- Q.2** (a) What is RDD? Explain about transformations and actions in the context of RDDs. State and explain RDD operations in brief. **07**
- (b) Discuss big data in healthcare, transportation and medicine. **07**
- OR**
- (b) What is Hbase? Discuss in detail the data model and Implementation Aspect. **07**
- Q.3** (a) Write Short note on Hadoop Ecosystem also explain various elements of hadoop. **07**
- (b) Discuss Hadoop YARN in detail with failures in classic Map-reduce **07**
- OR**
- Q.3** (a) What is Spark? State the advantages of using Apache Spark over Hadoop MapReduce for big data processing with example. **07**
- (b) Elaborate on HivQL data manipulation queries in detail **07**
- Q.4** (a) Discuss the concept of regions in HBase and Storing Big data with HBase. **07**
- (b) How does HDFS ensure data Integrity in a Hadoop Cluster? **07**
- OR**
- Q.4** (a) Explain how HBase uses Zookeeper to Build Applications with Zookeeper. **07**
- (b) What are the Components of Spark? Also state the features of Spark. **07**
- Q.5** (a) Explain Pig data Model in detail and Discuss how it will help for effective data flow. **07**
- (b) Explain Map-reduce framework in detail. Draw the architectural diagram for Physical Organization of Compute Nodes. **07**
- OR**
- Q.5** (a) Draw and discuss the architecture of Hive in detail. **07**
- (b) Write a brief short note on: Spark Unified Stack **07**
