

Code: IT8T2D

**IV B.Tech - II Semester – Regular / Supplementary Examinations  
March 2019**

**BIG DATA ANALYTICS  
(INFORMATION TECHNOLOGY)**

Duration: 3 hours

Max. Marks: 70

**PART – A**

Answer *all* the questions. All questions carry equal marks

11x 2 = 22 M

1.

- a) Define the terms Petabyte and Exabyte in measuring Big data.
- b) What is Hadoop MapReduce?
- c) What is Namenode and its Responsibilities?
- d) List any 4 components of HDFS.
- e) What is Replication Factor?
- f) List any 4 AWK operations.
- g) What are the three primary phases of Reducer Class?
- h) List any three input formats in MapReduce.
- i) Define Partial sort in MapReduce.
- j) Differentiate between HiveQL and MySQL.
- k) How Hbase is related to HDFS?

## PART – B

Answer any **THREE** questions. All questions carry equal marks.

3 x 16 = 48 M

2. a) What is big data? Why big data is more important? Explain with real time examples. 8 M
- b) Explain the role of Apache Hadoop in analyzing Big Data. 8 M
3. a) Explain HDFS architecture with neat diagram. 8 M
- b) Explain how files are written to HDFS. 8 M
4. a) What is the role of Java MapReduce in Hadoop. 8 M
- b) Explain MapReduce WebUI. 8 M
5. How Hadoop runs a MapReduce job using the classic framework? Explain with neat diagram. 16 M
6. a) Compare the Hive with Traditional Databases. 8 M
- b) List and explain any 6 Data types of Hive with description. 8 M