

Code: CS7T3

**IV B.Tech - I Semester – Regular / Supplementary Examinations  
November 2016**

**BIG DATA CONCEPTS  
(COMPUTER SCIENCE AND ENGINEERING)**

Duration: 3 hours

Max. Marks: 70

Answer any FIVE questions. All questions carry equal marks

1. a) Explain in detail about data storage and analysis. 7 M  
b) Compare structured data and unstructured data. 7 M
2. With an example code write a program that mines weather data from the NCDC. 14 M
3. Define HDFS. Explain HDFS in detail. 14 M
4. Draw and Explain the internal structure of a sequence file format. 14 M
5. With an example show the dependencies needed for building and testing a map reduce application. 14 M
6. Explain map-reduce frame in detail. Draw the architecture diagram for physical Organization of Nodes. 14 M

7. Draw and discuss the output format class hierarchy. 14 M
8. Briefly explain the following:
- a) Map reduce library classes. 7 M
  - b) Map Side joins. 7 M