

Module-1

1. What is Big data and explain the characteristics of data.
2. With relevant examples, explain the characteristics of structured, semi structured and unstructured data.
3. What is Big Data Analytics? Explain its importance.
4. Explain the different ways of handling unstructured data
5. List the challenges of big data
6. Compare the following
 - i) Traditional business intelligence (BI) Vs Big data
 - ii) Analytic 1.0, analytics 2.0 and analytics 3.0
7. Differentiate descriptive, predictive, diagnostic and prescriptive analytics with examples
8. Compare SQL, NOSQL and NewSQL
9. Write a note on: In memory analytics, in-database processing, and shared-nothing architecture with its advantages
10. What is CAP theorem? Explain the trade-off between consistency and availability with respect to CAP theorem.
11. Describe the features of Hadoop.
12. List and explain the key advantages of Hadoop
13. Explain a typical Hadoop environment with its components.

Module-2

1. Explain the architecture of Hadoop Distributed File System (HDFS) with a neat diagram.
2. Compare RDBMS Vs Hadoop
3. With neat diagram, explain the anatomy of File Read and File Write in Hadoop
4. List and explain HDFS commands.
5. With the neat diagram, explain MapReduce programming work flow
6. Apply map reduce programming techniques and write the programs for the following
 - i) Word count
 - ii) Sorting student information based on student name
7. Write a note on YARN architecture
8. Explain the Master-Slave architecture of Hadoop with a neat diagram.

Module-3

1. Explain the different data types supported by MongoDB.
2. Explain the benefits of using MongoDB in Big Data applications.

3. With neat diagrams, explain replication and sharding in MongoDB
4. List and explain the benefits of using MongoDB in Big Data applications.