

- N.B.:** 1) Question No.1 is compulsory.
2) Attempt any THREE from the remaining questions.
3) Figures to the right indicate full marks.

Write a short note on a following (Any Four)

- Q1.** (a) Need of Big Data Analytics [5]
(b) D3 and big data [5]
(c) CAP Theorem [5]
(d) Features of Apache Spark [5]
(e) Load and Store Functions of Apache Pig [5]
(f) Big Data Characteristics (any 5) [5]
- Q2.** (a) Explain Hadoop Distributed File System (HDFS) Architecture in detail. [10]
(b) Explain the Workflow using Resource Manager, Application Master & Node Master of Hadoop YARN with a suitable diagram. [10]
- Q3.** (a) What is Rack awareness in Hadoop HDFS? How to achieve fault tolerance with the help of Rack awareness? Explain with the help of a suitable example. [10]
(b) Explain matrix multiplication mapper and reducer function for MapReduce with a help of a suitable example. [10]
- Q4.** (a) List and explain types of NoSQL database with example of each? [10]
(b) Explain HBase Architecture and Components of HBase Architecture. [10]
- Q5.** (a) Explain the architecture of Apache Pig with help of diagram. [10]
(b) What is Apache Kafka? Explain the benefits and need of Apache Kafka. [10]
- Q6.** (a) What is big data and what are types of big data? Write differences between traditional data vs big data approach. [10]
(b) What are tools and benefits of Data Visualization also explain challenges of Big Data Visualization [10]
