

(3 Hours)

[Total Marks : 80]

N.B. :

- 1) Question No.1 is **compulsory**.
- 2) Attempt any **three** from the remaining **five** questions

- Q1 **Write short notes on following (any four):**
- [a] Explain the industry applications of Big Data. [05]
 - [b] Discuss the use of Pig Latin. [05]
 - [c] How is HBase helpful in Hadoop ecosystem? [05]
 - [d] Explain the benefits of Apache Kafka [05]
 - [e] Tableau as Big Data Visualization tool [05]
 - [f] Spark DAG [05]
- Q2 [a] Write short note on Mapper Task and Reducer Task [10]
[b] Explain HDFS Architecture in detail. [10]
- Q3 [a] Explain matrix multiplication mapper and reducer function for MapReduce with a help of a suitable example. [10]
[b] Explain the architecture of Apache Hive in Hadoop Ecosystem. List out various HiveQL builtin functions. [10]
- Q4 [a] Describe the main components and the working of Apache Pig. Discuss the Load() & Store() commands in Pig framework [10]
[b] Discuss HBASE data models used in Hadoop for Big Data Analytics. [10]
- Q5 [a] What is RDD? How is data partitioned in RDD? [10]
[b] Explore the process of visualizing data using Data Driven Documents. [10]
- Q6 [a] Discuss various challenges and approaches to big data visualization [10]
[b] What is Big Data and its types? List out the difference between Traditional Data Vs. Big Data with the help of example? [10]
