

M.Sc.(IT & CA) Semester - 2 (CBCS) Examination
May/June-2021 (OLD COURSE)
INTRODUCTION TO BIG DATA AND HADOOP(CORE)

Time: 1:30 Hours**Marks: 42****Instructions:**

1. Figure to the right indicate marks.
 2. There are five questions in the question paper.
 3. Answer any three of the following questions.
-

- Q.1(A) Answer the following questions. (04)
1. What is Virtual Box?
 2. What is Big Data?
 3. Hadoop was developed by ____.
 4. BASE Stands for.
- Q.1(B) Answer the following question. (any one) (02)
1. Explain CRC with example.
 2. Difference between Structured and Unstructured Data.
- Q.1(C) Answer the following question. (any one) (03)
1. Explain 3V's of Big Data.
 2. Explain Rack awareness and Block Placement Policy.
- Q.1(D) Answer the following question. (any one) (05)
1. Explain HDFS Administration Commands.
 2. Describe HDFS Architecture.
- Q.2(A) Answer the following questions. (04)
1. What is MapReduce?
 2. What is Apache Pig?
 3. DAG Stands for.
 4. UDF Stands for.
- Q.2(B) Answer the following question. (any one) (02)
1. Difference between Slots and Containers.
 2. Explain Distributed Cache in Hadoop.
- Q.2(C) Answer the following question. (any one) (03)
1. Explain Hadoop Schedulers.
 2. Explain Architecture of Pig.
- Q.2(D) Answer the following question. (any one) (05)
1. Explain MapReduce Architecture.
 2. Explain Pig Diagnostic Operators in details.
- Q.3(A) Answer the following questions. (04)
1. What is Avro file?
 2. What is Hive?
 3. IDF Stands for.
 4. DML Stands for.
- Q.3(B) Answer the following question. (any one) (02)
1. Explain External and Managed table in Hive.
 2. Explain Hive Partitioning.

- Q.3(C) Answer the following question. (any one) (03)
1. Explain Hive View with example.
 2. Explain Map-Reduce Algorithms.
- Q.3(D) Answer the following question. (any one) (05)
1. Describe Architecture of Hive.
 2. Explain Hadoop Optimization Techniques in details.
- Q.4(A) Answer the following questions. (04)
1. What is HBase?
 2. What is Zookeeper in HBase?
 3. HBase is a Column-Oriented database. (True/False)
 4. Scan command is used to view the data in HTable. (True/False)
- Q.4(B) Answer the following question. (any one) (02)
1. Explain Data Model Operations in HBase.
 2. Explain Catching and Batching in HBase.
- Q.4(C) Answer the following question. (any one) (03)
1. Explain HBase Logical and Comparators Operators.
 2. Describe HBase Coprocessors.
- Q.4(D) Answer the following question. (any one) (05)
1. Explain Catalog Tables in HBase.
 2. Explain HBase Architecture.
- Q.5(A) Answer the following questions. (04)
1. What is Spark?
 2. Neo4j is a Graph Database. (True/False)
 3. RDD Stands for.
 4. SSH Stands for.
- Q.5(B) Answer the following question. (any one) (02)
1. What is Oozie?
 2. What is Flume?
- Q.5(C) Answer the following question. (any one) (03)
1. Explain Amazon EMR.
 2. What is Graph? Explain with example.
- Q.5(D) Answer the following question. (any one) (05)
1. Explain Spark Architecture.
 2. Explain spark RDD in details.
