

M.Sc.(IT & CA) Semester - 2 (CBCS) Examination
August/September -2020 [NEW COURSE]
INTRODUCTION TO BIG DATA AND HADOOP (CORE)

Time: 2:00 Hours**Marks: 56****Instructions:**

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

Q.1(A) Attempt the following. (03)

1. HDFS stands for _____
2. Default block size in HDFS is (Hadoop 2.0) _____
3. Default replication factor in HDFS is _____

Q.1(B) Give the following answer (any two) (06)

1. Explain structured, semi structured and unstructured data?
2. Explain HDFS data read process.
3. Explain name node and secondary name node.

Q.1(C) Give the following answer (any one) (05)

1. Explain 3V characteristics of big data.
2. Explain HDFS write process.

Q.2(A) Attempt the following. (03)

1. Pig was developed by _____.
2. YARN stands for _____.
3. In pig _____ is used to display the output to the screen.

Q.2(B) Give the following answer (any two) (06)

1. Explain LOAD in pig.
2. Explain schedulers in MR.
3. Explain different mode of executing pig constructs.

Q.2(C) Give the following answer (any one) (05)

1. Explain map reduce architecture.
2. Explain input splits in pig using suitable example.

Q.3(A) Attempt the following. (03)

1. Hive is based suited for OLAP. (True/False)
2. HQL stands for _____.
3. DML stands for _____.

- Q.3(B) Give the following answer (any two) (06)
1. Explain managed table and external table.
 2. Explain partitions in hive.
 3. Explain standalone mode of MR.
- Q.3(C) Give the following answer (any one) (05)
1. Explain hive architecture.
 2. Explain optimization techniques in map reduce.
- Q.4(A) Give the following answers. (03)
1. HBase stores data in _____ form.
 2. HBase is NoSQL, distributed database. (True/False)
 3. What is HBase coprocessor?
- Q.4(B) Give the following answer (any two) (06)
1. Explain advanced HBase features.
 2. Explain zookeeper.
 3. Create a table in HBase using java API.
- Q.4(C) Give the following answer. (any one) (05)
1. Explain HBase architecture.
 2. Explain any five filters in HBase.
- Q.5(A) Give the following answers. (03)
1. RDD stands for _____.
 2. EMR stands for _____.
 3. EC2 stands for _____.
- Q.5(B) Give the following answer (any two) (06)
1. Explain spark architecture in detail.
 2. Explain RDD.
 3. Explain Oozie work flow.
- Q.5(C) Give the following answer (any one) (05)
1. Explain different mode of Hadoop cluster.
 2. Explain import export process RDBMS in HDFS using Sqoop.
