Printed pages: 02 | Sub Code:RCAE45

Paper Id: 214512 | Roll No:

MCA

(SEM V) THEORY EXAMINATION 2018-19

BIG DATA

Time: 3 Hours | *Total Marks: 70*

Note: 1. Attempt all Sections. If require any missing data; then choose suitably.
2. Any special paper specific instruction.

SECTION A

1. **Attempt *all* questions in brief** **2 x 7 = 14**

a. What is Big data? Discuss it in terms of volume and velocity.
b. What are the advantages of Hadoop?
c. Explain Metastore in HIVE.
d. Discuss big data in healthcare and medicine.
e. Define Serialization in Hadoop.
f. What is Zookeeper?
g. What is the necessity of PIG Latin?

SECTION B

2. **Attempt any *three* of the following:** **7 x 3 = 21**

a. What are the benefits of Big Data? Discuss challenges under Big Data. How Big Data Analytics can be useful in the development of smart cities.
b. Write a short note on NoSQL databases. List the differences between NoSQL and relational databases?
c. Draw and explain HDFS Architecture. Explain the function of NameNode and DataNode. What is a Secondary Namenode? Is it a substitute to the Namenode?
d. Explain "Shuffle & Sort" phase and "Reducer Phase" in MapReduce.
e. What are views in HIVE? What is the difference between internal and external tables in HIVE?

SECTION C

3. **Attempt any *one* part of the following:** **7 x 1 = 7**

(a) What are structures, unstructured and semi-structured data? Explain with examples.
(b) What are the different modes in which Hadoop can be installed and what is the use of each mode from application and developer point of view?

4. **Attempt any *one* part of the following:** **7 x 1 = 7**

(a) Define the role of combiner and partitioner in a map reduces application.
(b) Explain Master slave and peer-peer replication in detail.

5. **Attempt any *one* part of the following:** **7 x 1 = 7**

(a) How does HDFS ensure data Integrity in a Hadoop Cluster?
(b) Explain Avro file-based data structures in detail.

6. **Attempt any *one* part of the following:** **7 x 1 = 7**

(a) Discuss Hadoop YARN in detail with failures in classic MapReduce.
(b) Explain briefly about Input format and Output format in MapReduce.

7. **Attempt any *one* part of the following:** **7 x 1 = 7**

(a) What are the components of Pig Execution Environment?
(b) Explain Storage mechanism in HBase. Write a query to create a table in HBase.

****