Printed Pages:02 **Sub Code: KCA-022**

Paper Id: | 2 | 3 | 2 | 3 | 4 | 2 | **Roll No.** | | | | | | | | | | | | |

MCA
(SEM III) THEORY EXAMINATION 2022-23
BIG DATA

Time: 3 Hours ***Total Marks: 100***

Note: Attempt all Sections. If you require any missing data, then choose suitably.

SECTION A

1. Attempt *all* questions in brief. 2x10 = 20

a) Define types of digital data.
b) What is big data privacy and ethics?
c) Write history of hadoop.
d) What is real world Map Reduce?
e) What is HDFS concept?
f) What is Avro and file-based data structures?
g) Describe Hadoop ecosystem components.
h) State the need of NoSQL databases?
i) Describe the comparison of Pig with databases.
j) What is a zookeeper?

SECTION B

2. Attempt any *three* of the following: 10x3 = 30

a) Explain the concept big data with its architecture, and technological components.
b) Define anatomy of map reduce job run.
c) How does hdfs store, read, and write files.
d) Differentiate in following in reference with proper examples-
   (i) SQL and NOSQL
   (ii) MapReduce and YARN
   (iii)HBase and HDFS
   (iv)Spark and SCALA
e) Describe applications on big data using Pig, Hive, and HBase.

SECTION C

3. Attempt any *one* part of the following: 10x1 = 10

a) Analyze nature of data, and challenges of convolutional systems over intelligent data analysis? Also differentiate between analysis and reporting.
b) Describe 5 v's of data and big data features in detail.

4. Attempt any *one* part of the following: 10x1 = 10

a) Describe Apache Hadoop, Hadoop streaming, and Hadoop pipes.
b) Explain how does map reduce works. Also explain unit tests with MR unit, test data and local tests.

QP23DP1_029 | 20-02-2023 13:26:33 | 125.21.249.98

QP23DP1_029 | 20-02-2023 13:26:33 | 125.21.249.98

5. Attempt any *one* part of the following: 10x1 = 10

a) Explain the procedure of setting up a Hadoop cluster, cluster specification, cluster setup and installation.

b) Define data ingest with flume and scoop. Also, explain the concept of Hadoop in the cloud.

6. Attempt any *one* part of the following: 10x1 = 10

a) What are spark applications and how to install spark.

b) Describe inheritance, functions and closures concept of scale.

7. Attempt any *one* part of the following: 10x1 = 10

a) What do you mean by apache hive architecture? Explain hive installation.

b) What do you mean by infosphere, BigInsights and big sheets?

QP23DP1_029 | 20-02-2023 13:26:33 | 125.21.249.98

QP23DP1_029 | 20-02-2023 13:26:33 | 125.21.249.98