

[This question paper contains 2 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 4602

L

Unique Paper Code : 2346000013

Name of the Paper : Big Data Analytics - I

Name of the Course : SEC (NEP-UGCF-2022)

Semester : VI

Duration : 1 Hours

Maximum Marks : 30

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The paper has two sections. All questions in "Section A" are compulsory.
3. Attempt any four questions from "Section B".
4. Parts of a question must be answered together.

Section A

1. (i) A _____ node acts as the Slave and is responsible for executing a Task assigned to it by the JobTracker. (1)
- (ii) HDFS provides a command-line interface called _____ to interact with HDFS. (1)
- (iii) _____ is the slave/worker node and holds the user data in the form of Data Blocks. (1)
- (iv) Output of the mapper is first written to the local disk for sorting and _____ process. (1)

P.T.O.

- (v) The data that can be processed, stored, and retrieved in a fixed format called _____ (1)
- (vi) The output of Map tasks is written in _____. (1)
- (vii) The daemons associated with the MapReduce phase are _____ and _____ (2)
- (viii) State whether True/False: (2)
- (i) Big data deals with high-volume, high-velocity and high-variety information assets.
 - (ii) The Big Data Analytics work on the unstructured data, where no specific pattern of the data is defined.

Section B

(Attempt any four)

- 2 Discuss Hadoop framework architecture. (5)
- 3 Explain the challenges of Big Data. (5)
- 4 Discuss the three V's of big data (5)
- 5 Explain Name node and Data node (5)
- 6 How data will be processed by MapReduce pipeline. Explain with the help of a diagram and an example. (5)
- 7 Discuss HDFS Architecture with the help of a diagram. (5)