

4731

4

5. What is NoSQL database? How is it different from the relational database? (5)
6. Differentiate between structured, semi-structured, and unstructured data. Give an example of each. (5)

(100)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 4731

H

Unique Paper Code : 2346000013

Name of the Paper : Big Analytics I

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

Duration : 1 Hour

Maximum Marks : 30

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **All** parts in **Section A** are compulsory.
3. Attempt **ANY FOUR** questions from **Section B**.
4. Parts of a question should be answered together.

Section A

1. (a) Fill in the blanks : (8)

P.T.O.

- (i) _____ can best be described as a programming model used to develop Hadoop-based applications that can process massive amounts of data.
- (ii) MapReduce is implemented _____ in programming language.
- (iii) _____ node serves as the slave and is responsible for carrying out the tasks that have been assigned to it by the JobTracker.
- (iv) During start-up, the _____ loads the file system state from the fsimage and the EditLogs file.
- (v) _____ maps input key pairs to a set of intermediate value pairs.
- (vi) _____ is an example of relational database.
- (vii) _____ acts as a master node in a Hadoop cluster.

- (viii) _____ join is usually used when one data set is large and the other data set is small.

(b) State whether the following statements are true or false : (2)

- (i) Big data is often difficult to collect, but easy to analyze.
- (ii) Raw data should be processed only one time.

Section B

- 2. How NameNode failure is handled in Hadoop? (5)
- 3. What are the two main phases of a MapReduce operation? Explain with the help of an example. (5)
- 4. How do you copy data from the local file system to HDFS and vice-versa? (5)