

Reg. No. :

--	--	--	--	--	--	--	--	--	--

Question Paper Code: 49809

B.E. / B.Tech. DEGREE EXAMINATION, APRIL 2019

Elective

Information Technology

14UIT909- BIG DATA AND ANALYTICS

(Regulation 2014)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 1 = 10 Marks)

- Unstructured data is
 - Table data
 - data heavy
 - Transactional data.
 - Combination of table and date
- Which one from the following is a not analytics method?
 - Regression
 - Monte Carlo
 - Spatial
 - Classification
- Most NoSQL databases support automatic _____, meaning that you get high availability and disaster recovery
 - Processing
 - Scalability
 - Replication
 - All of the mentioned
- Pipes use _____ as the channel for communication.
 - bus
 - sockets
 - data
 - control
- HDFS provides a command line interface called _____ used to interact with HDFS.
 - “HDFS Shell”
 - “FS Shell”
 - “DFS Shell”
 - “HDF Shell”
- The High availability of Namenode is achieved in HDFS2.x using _____
 - Polled Edit Logs
 - Synchronized Edit Logs
 - Shared Edit Logs
 - Edit Logs Replacement

7. When You are developing a combiner that takes as input Text keys, IntWritable values, and emits Text keys, IntWritable values. Which interface should your class implement?
 - (a) Combiner <Text, IntWritable,Text, IntWritable>
 - (b) Reducer <Text, IntWritable,Text, IntWritable>
 - (c) Combiner <Text,Text, IntWritable, IntWritable>
 - (d) Combiner <Text, Text, IntWritable, IntWritable>
8. Which of the following is web-based client software for MongoDB ?
 - (a) BI Studio
 - (b) Database Master
 - (c) Fang of Mongo
 - (d) Mongo3
9. Which of the following is a wide-column store?
 - (a) Cassandra
 - (b) Riak
 - (c) MongoDB
 - (d) Redis
10. _____ provides statistics on the per-collection level.
 - (a) Mongosniff
 - (b) Mongotop
 - (c) Mongooplog
 - (d) Mongofiles

PART - B (5 x 2 = 10 Marks)

11. Name the advanced big data analytics methods.
- 12 List out Pros and cons of relational databases.
13. Why Key Type need be both Writable and Comparable in Map-Reduce Programs.?
14. What is Hadoop streaming?
15. What is ZOO keeper?

PART - C (5 x 16 = 80 Marks)

16. (a) Explain the working of different fraud detecting approaches. (16)
- Or
- (b) Explain with illustrative examples Big data in healthcare.. (16)
17. (a) Illustrate with examples relations and aggregates (16)
- Or
- (b) Analyze with an example How Hadoop uses Scale out feature to improve the Performance (16)

18. (a) Detail the Read and Write operations in Hadoop Distributed File System (HDFS) and detail the Java API required for that. (16)

Or

(b) Analyze Hadoop streaming in detail with suitable examples (16)

19. (a) How does map reduce on YARN technology work? Details the failures in YARN over map reduce. (16)

Or

(b) Draw the architecture of Map Reduce process and explain the mapper and master in detail. (16)

20. (a) What are the CRUD operations in HBase? Among them explain Write and Read operations with supportive commands. (16)

Or

(b) What are the good rules and goals for creating data model in Cassandra? (16)

