

SHRIMATHI DEVKUNVAR NANALAL BHATT VAISHNAV COLLEGE FOR WOMEN
(AUTONOMOUS)

(Affiliated to the University of Madras and Re-accredited with 'A+' Grade by NAAC)
Chromepet, Chennai - 600 044.

B.Sc.Computer Science - END SEMESTER EXAMINATIONS - NOV'2024

SEMESTER - V

24UCSET5B01 - Big Data Analytics

Total Duration : 1 Hrs.30 Mins.

Total Marks : 40

Section B

Answer any **TEN** questions ($10 \times 2 = 20$ Marks)

1. What is Cap theorem?
2. Define CQLSH.
3. What is a key feature of the replication pipeline process in HDFS?
4. What is the primary function of the MapReduce component in the Hadoop ecosystem?
5. Define Cassandra.
6. What is the main difference between partitioned and non-partitioned databases in Cloudant?
7. What are the primary benefits of Hadoop for big data processing and storage?
8. Write the primary benefit of RDD immutability in Spark.
9. Define NoSQL database.
10. What is the primary purpose of an index in MongoDB?
11. What is a directed acyclic graph (DAG) and its role in Spark's data analysis pipeline execution?
12. What is the primary function of the MapReduce component in the Hadoop ecosystem?

Section C

Answer any **FOUR** questions ($4 \times 5 = 20$ Marks)

13. Explain the process of sharding in NoSQL databases and its benefits.
14. How does MongoDB's flexible schema allow for easier adaptation and modification of the database structure?
15. Explain the main components of Hadoop Distributed File System (HDFS).

Contd...

16. Write an advantages of using Spark for big data processing.
17. Discuss the YARN as the cluster manager interface for Spark.
18. Describe the curl command line tool to make HTTP requests to the Cloudera API.
