SNS COLLEGE OF TECHNOLOGY

(An Autonomous Institution)

23MCT303 – DATA ANALYTICS IN AUTOMATION SYSTEM

QUESTION BANK

UNIT IV: MANAGING HADOOP FILE (CO4)

Short Answer Questions (2 Marks)

Q.No	Question	Bloom's Level	CO
1	What is Hadoop? State its two main components.	Remember	CO4
2	Define HDFS and list its key features.	Remember	CO4
3	Name the two phases of MapReduce.	Remember	CO4
4	List the three modes of Hadoop installation.	Remember	CO4
5	Differentiate between vertical and horizontal scalability.	Understand	CO4
6	What is the role of NameNode in HDFS?	Understand	CO4
7	State any two advantages of MapReduce paradigm.	Remember	CO4
8	What is a DataNode in HDFS?	Remember	CO4
9	Define fault tolerance in Hadoop.	Understand	CO4
10	Mention the purpose of secondary NameNode.	Remember	CO4

Long Answer Questions (13/14 Marks)

Q.No	Question	Bloom's Level	CO
1	Explain the anatomy of Hadoop including HDFS architecture and MapReduce framework in detail.	Apply/Analyze	CO4
2	Describe the working of MapReduce programming model with execution flow and key components.	Apply/Analyze	CO4
3	Discuss vertical and horizontal scalability in Hadoop. Explain how Hadoop achieves scalability and fault tolerance.	Analyze	CO4
4	Compare Stand-alone, Pseudo Distributed, and Fully Distributed modes of Hadoop installation with configuration details.	Apply/Analyze	CO4
5	Explain HDFS block storage, replication mechanism, and roles of NameNode and DataNode.	Apply/Analyze	CO4
6	Describe the complete process of executing a MapReduce job in Hadoop.	Apply	CO4
7	Discuss the advantages of using Hadoop for managing large-scale data in distributed environments.	Analyze	CO4
8	Explain how Hadoop ensures data reliability and high availability through its architecture.	Analyze	CO4
9	Describe the configuration steps for setting up Hadoop in Pseudo Distributed mode.	Apply/Create	CO4

Q.No	Question	Bloom's Level	CO
10	Discuss the key features of HDFS that make it suitable for big data storage and processing.	Analyze	CO4