**Module 2**

1. Explain with a neat diagram various system role in an HDFS development.
2. Write down the steps required to request to mapreduce and the types of process in mapreduce.
3. Explain the following with example:-
4. HDFS block replication
5. Rack awareness
6. HDFS safe node
7. Name node high availability
8. HDFS snapshot
9. Illustrate YARN based execution model and its functions with a neat diagram.
10. Write the command syntax for the following:
11. Copy files from HDFS
12. Copy files within HDFS
13. Delete a file within HDFS
14. Delete a directory in HDFS
15. Make a directory in HDFS
16. Copy files to HDFS
17. Explain with a neat diagram Apache Sqoop Import and Export Methods.
18. Explain why Apache Hive and Pig are used. Bring out the differences between Hive and Pig.
19. Depict the oozie DAG workflow jobs with suitable diagrams.

**Module 3**

1. List the characteristics of NoSQL data store. Show how does CAP and Base properties hold in NoSQL databases?
2. Classify and describe the different NoSQL data Architectural pattern.
3. Illustrate the different Shared-Nothing distribution Model.
4. Compare between NoSQL and SQL(RDBMS). Give the characteristics of NoSQL solutions for Big Data.
5. List and compare the features of BigTable, RC, ORC and Parquet data stores.
6. What are the characteristics of the object data store model?
7. With example explain key-value store.
8. Illustrate Schema less model and its characteristics.