Module 1

1) How big data problems are handled by Hadoop system. 5M

2) What characteristics social networks makes it Big data? 5M

3) What is Big data? What is Hadoop? How they are linked? 5M

4) How Big data can be useful in developing Digital India? 5M

5) Mention the 4 characteristics of bigdata. Elaborate these characteristics w.r.to social media websites. 5M

6) Give difference between Traditional data management and analytics approach Versus Big data Approach. 5M

7) List down at least 4 different sources of bigdata from different domains and justify how they can be considered as bigdata applications.

8) Describe any five characteristics of Big Data 5M

9) Compare big data analytics with traditional data mining and warehousing system. 5M

10) What are the 3 V's of big data? Give two examples of big data case studies. Indicate which V's are satisfied by these case studies. 5M

11) What do you mean by the Hadoop Ecosystem? Describe any three components of typical Hadoop Ecosystem. 10M

12) Explain Hadoop Ecosystem with core components and explain its physical architecture. State limitations of Hadoop. 10M

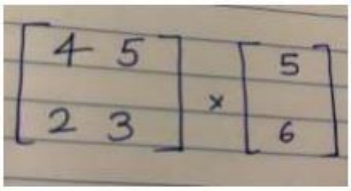
Module 2

1) Explain how Hadoop goals are covered in Hadoop distributed file system. 10M

2) Write pseudo code for Matrix vector Multiplication by MapReduce. Illustrate with an example showing all the steps. 10M

3) Apply Map Reduce Vector Multiplication algorithm to perform the following matrix

vector multiplication.



4) Show MapReduce implementation for joining two relations with example using pseudocode. 5M

5) What is MapReduce ? Explain how MapReduce works? What is shuffling in MapReduce? 10M

6) Describe the operations of shuffle and sort in the MapReduce framework with the help of one example. 5M

7) Explain the Map Reduce working and apply the working on the following document. "I like an apple and a banana. He likes an apple and a melon. I also like a melon." 5M

8) What is Combiner? When one should use combiner in MapReduce job? 5M

9) Explain how Hadoop's mapper and reducer work, with an example of performing any relational algebra operation using Map Reduce.

Module 3

1) Describe characteristics of NoSQL database. 10 M
2) What is NoSQL? Discuss any two architecture patterns of NoSQL.
3) List all variation of NoSQL database with two features of each and two examples of each. 5M
4) Differentiate between RDBMS and NoSQL Database. 10M (refer https://youtu.be/0buKQHokLK8)
5) When it comes to big data how NoSQL scores over RDBMS. 5M
6) Explain in detail any two Big data Applications based on NoSQL. 5M
7) Compare Key Value No-SQL datastore with Document based NoSQL datastore. 10M
8) What are the different architectural patterns in NoSQL? Explain Graph data store and Column FamilyStore patterns with relevant examples.
9) What do you understand by BASE properties in NoSQL database? Explain in detail any one NoSQL architecture pattern. Identify two applications that can use this pattern. 10M

Ans: Key value store - Amazon DynamoDB

Column family - Google's BigTable

Document Store MongoDB

Graph Store -Neo4j

10) Explain CAP theorem of NoSQL database. As No SQL database is not able to adopt ACID propertie scan we adopt NoSQL for traditional banking application? 5M
11) Explain in detail business drivers of NoSQL. 10M(Refer Moodle)
12) Explain four ways that NoSQL handles Big Data Problems. 10M(Refer Moodle)