

Reg. No. :

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

Question Paper Code: 99242

B.E./B.Tech. DEGREE EXAMINATION, APRIL 2024

Elective

Computer Science and Engineering

19UCS942 -BIG DATA ANALYTICS

(Régulations 2019)

Duration: Three hours

Maximum: 100 Marks

Answer All Questions

PART A - (10 x 2 = 20 Marks)

1. What is Big Data? Why Big Data? CO1-U
2. DFS can handle a large volume of data then why do we need Hadoop framework? CO1-U
3. How is the DataNode failure handled by NameNode? CO1-U
4. What is the fundamental difference between a MapReduce Split and a HDFS block? CO1-U
5. How to Create a Database in MongoDB? CO1-U
6. Can we change the block size? What is a block in HDFS and what is its default size in Hadoop 1 and Hadoop 2? CO2- App
7. Why MongoDB is a Document Oriented Storage? And why use MongoDB? CO1-U
8. Compare MongoDB with Cassandra. CO1-U
9. While creating a schema in MongoDB, what are the points need to be taken into consideration? CO2- App
10. Which command is used for inserting a document in MongoDB? CO2- U

PART – B (5 x 16= 80Marks)

11. (a) Why is big data important and describe some of the big data use cases CO2-App (16)
- Or
- (b) Briefly explain the Components of hadoop and application development in Hadoop CO2-App (16)

12. (a) Imagine that you are uploading a file of 500MB into HDFS. 100MB of data is successfully uploaded into HDFS and another client wants to read the uploaded data while the upload is still in progress. What will happen in such a scenario, will the 100 MB of data that is uploaded will it be displayed? CO2-App (16)
- Or
- (b) If 8TB is the available disk space per node (10 disks with 1 TB, 2 disk for operating system etc. were excluded.) and assuming initial data size is 600 TB. How will you estimate the number of data nodes (n)? CO2-App (16)
13. (a) Outline the steps involved to install and configure the MongoDB in Ubuntu. CO1-U (16)
- Or
- (b) Explain in details about the CRUD Operations with examples. CO1-U (16)
14. (a) Write a mapreduce program to count the occurrence of similar words in a file. use partitioner to partition key based on alphabets
 Input Data : Welcome to hadoop session
 Introduction to hadoop
 Introducing HIVE
 HIVE Session
 Pig Session CO1-U (16)
- (b) Write a mapreduce program to sort data by student name(Value). CO1-U (16)
- input data
 1001, john, 45
 1002, jack, 39
 1003, alex, 44
 1004, smith, 38Or
15. (a) Illustrate the Map Reduce architecture and Components of Map Reduce architecture with example. CO1-U (16)
- Or
- (b) Explain in detail about various components of Hadoop ecosystem CO1-U (16)