Reg.No. \_\_\_\_\_\_\_\_\_\_\_\_

G:\logo and QP Template\logo 3 Feb 2018 final.tif

**End Semester Examination – Nov/Dec – 2018**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| **Code :** | **15CS3007** | **Duration :** | **3hrs** |
| **Sub. Name :** | **BIG DATA PLATFORMS** | **Max. marks :** | **100** |

**ANSWER ALL QUESTIONS (5 x 20 = 100 Marks)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Q. No.** | **Sub Div.** | **Questions** | **Course**  **Outcome** | **Marks** |
| 1. | a. | Elucidate the Characteristics of big data with appropriate examples and statistics. Share your experience as a customer on an e-commerce site. Comment on the big data that gets created on a typical e-commerce site. | CO1 | 10 |
| b. | In your university, you are getting streaming data from weather monitoring sensors. Also you have weather data for the past 3 years and spatial images of your region. You are requested to do an analysis by combining all these data and to make predictive analysis. In this scenario what type of database would you prefer? Discuss in detail with its types and advantages. | CO1 | 10 |
| (OR) | | | | |
| 2. | a. | You are at the library. You see a few students browsing through the library catalog on a kiosk. You see few students fill up the feedback form on the services offered by the library. Quite a few students are learning using e-learning content.   * Think for a while on the different types of data that are being generated in this scenario.Support your answer with logic * Analyze the challenges in dealing with unstructured data. | CO1 | 10 |
| b. | Is Data science and Business Intelligence same? State in your own words about Data science. What are the responsibilities of a Data Scientist? | CO1 | 5 |
|  | c. | Illustrate the Brewers’ theorem with suitable examples. | CO1 | 5 |
|  |  |  |  |  |
| 3. | a. | Compare and contrast “insert” and “save” in MongoDB. Give examples | CO3 | 5 |
| b. | State the relational operators available in MongoDB. Give syntax and illustrate each operator with an examples. | CO3 | 7 |
| c. | Why to use MongoDB? Enumerate the features of MongoDB. | CO3 | 8 |
| (OR) | | | | |
| 4. | a. | Explain Map Reduce Programming in MongoDB with a suitable example. | CO2 | 8 |
| b. | Consider the following collection called “student” and write the MongoDB commands for the following questions  Student (s-id,s-name, subject, DOB, grade, hobbies)   1. Create a collection called “student” and insert 5 documents 2. Display the documents from the student collection where hobbies is set to either “Singing” or “Dancing”. 3. Find the number of documents in the student collection wherein grade is “S”. 4. Display the fifth and sixth document from the student collection 5. Insert an array called phone\_nos into the students collection which contains 3 phone numbers. 6. Sort the documents based on grade in desending order and then on s-name in ascending order. | CO3 | 12 |
|  |  |  |  |  |
| 5. | a. | Paraphrase the main features of Cassandra with suitable diagrams and examples. Elaborate on Hinted handoffs and tunable consistency. | CO1 | 12 |
| b. | Tell what would happen if the following query is executed:  INSERT into userlogin (userid, psswd) values (1, Karunya) using TTL 30;  Describe what is meant by TTL. | CO1 | 4 |
| c. | Create a graph with 2 nodes with relationship as per the details given below using Neo4j.   * Customer node contains :id, name, dob properties * CreditCard node contains: id, number, cvv, expiredate properties * Customer to CreditCard Relationship: DO\_SHOPPING\_WITH * CreditCard to Customer Relationship: ASSOCIATED\_WITH | CO1 | 4 |
| (OR) | | | | |
| 6. | a. | Consider the following specification and write the Cassandra command for the given questions.  Employee(employee\_id,employee\_name,employee\_address,city,state, pstal\_code)   * Create a keyspace called “University” with replication factor-2 and simple strategy and create a “employee” table with the above mentioned fields, insert the necessary data * Insert a row into the employee table with new fields age and DOB. * Update the pstal\_code of employee\_id=1 to ‘641067' * Alter the employee table by adding a set called “skills” and update the values for “skills” * Alter the employee table by adding a map “todo” and update the values for “todo” * Display the details of customer living the city Coimbatore. | CO1 | 12 |
| b. | What is a graph? State the various features and advantages of Neo4j. | CO1 | 8 |
|  |  |  |  |  |
| 7. | a. | What were the limitations of Hadoop 1.0 architecture? Can you explain how it has been rectified in Hadoop 2 YARN with a neat diagram and illustration. | CO2 | 10 |
| b. | Explain the functions of master server and region servers in HBase with a neat diagram. | CO3 | 10 |
| (OR) | | | | |
| 8. | a. | What is the function of Sqoop? Discuss in detail with a neat diagram. | CO2 | 6 |
| b. | State the purpose of using PIG. Give a comparative study on PIG, Hive and Impala. | CO2 | 8 |
| c. | Explain the working principle of FLUME with a neat diagram. | CO2 | 6 |
|  | |  |  |  |
|  | | **Compulsory**: |  |  |
| 9. | a. | How shuffle and sort work in MapReduce system? Describe the process in the map side as well as in the reduce side with the suitable diagram. | CO2 | 7 |
| b. | Compare and contrast the fair scheduler and the capacity scheduler. | CO2 | 6 |
| c. | Illustrate how Hadoop runs a MapReduce job with a diagram and the necessary Java methods. | CO2 | 7 |