

JK Lakshmipat University Jaipur
INSTITUTE OF ENGINEERING AND TECHNOLOGY

End Term Examination, December 2023

B. Tech. (CSE), Semester VII, 2023-24

Roll No.....

CS1225: MULTI-DIMENSIONAL AND NOSQL
 DATABASES

Time: 2.5 hours

Max. Marks: 30

Instructions to students:

1. Do not write anything other than your roll number on question paper.
2. Mention all the assumptions for your answers clearly.
3. Attempt all questions

Q1	Assume there are 20 processors as part of a parallel distributed database system. Using the concept of range partitioning, 1. Perform SORTING on a column like price and a price range from 1000-5000. 2. Perform an equi JOIN on two relations R1(..,age,..) and R2(...,age,..). Assume that the two relations are distributed on another field called emp_id . Consider two cases: a. Size of R1 and R2 is almost the same. b. Size of R1 is much smaller than R2.	2.0 (LO4) 1.0+2.0 (LO 4)
Q2.1	Consider a system where a temperature sensor records temperature. Following data is collected about the sensor like id, name etc., time of generation of the event, event value, event id. Consider the following query: For the month of September 2022, display all data that has been generated by a sensor named 's11234' recording an event with id as 12. Design a columnFamily/Table for the query above and write the CQL table that you create.	2.0 (LO 5,6)
Q2.2	For each question below, answer the question with respect to Cassandra and CQL: a. The number of memtables in each Cassandra node is/are _____ b. Assume the following: Replication Factor is 5, READ consistency is QUORUM, WRITE consistency is ONE. Has the system achieved strong consistency? Justify c. <i>SELECT TTL (all Primary Key Columns) FROM Table1;</i> What will be the output? Justify.	1.0+1.0 +1.0 (LO 5,6)

Q3.1	Consider the following as one of the documents, stored in MongoDB, of a company called ABC Automobiles. <pre> { "InvoiceNumber": "GBPG8011", "TotalSalePrice": 89000, "SaleDate": "2015-04-30T00:00:00", "Customer": { "Name": "Wonderland Wheels", "CreditRisk": false, "Reseller": true }, "Address": { "Town": "London", "PostCode": "E7 4BR", "CountryName": "United Kingdom" }, "CustomerComments": ["Excellent", "Wonderful", "Superb"], "Salesdetails": [{ "LineItem": 1, "Make": "Porsche", "Model": "944", "SellingPrice": 8500, "LineItemDiscount": 50, "PurchaseCost": 6800, "RepairsCost": 250, "PartsCost": 225, "TransportInCost": 150 }, { "LineItem": 2, "Make": "Bentley", "Model": "Flying Spur", "SellingPrice": 80500, "LineItemDiscount": 500, "PurchaseCost": 64400, "RepairsCost": 500, "PartsCost": 750, "TransportInCost": 750 }] } </pre> ABC Automobiles buys second-hand cars, repairs them, and sells them either directly to customers or to other intermediate companies like “Wonderland Wheels”. Write MongoDB aggregation pipelines for the following: 1. For all the customers in the “town” of London, what is the “total sale price” that ABC automobiles has earned? 2. What is the average “selling price” per car “make” in the database?	1.5+1.5 (LO 6)
Q3.2	Explain the architecture of mongoDB with respect to a replica set. Describe in detail the components and the communication between the components in different scenarios.	2.0 (LO 5)
Q4.1	A husband and wife own two houses and 3 cars. What are the number of edges when it is modeled as a Property Graph versus as a Hyper Graph. Draw the two graphs	2.0 (LO 6)

Q4.2	Consider the following graph: <pre> CREATE (p:DRUG{name: "A", primaryUse: "Epilepsy"}) , (p1:DRUG{name: "B", primaryUse: "Analgesic"}) , (p2:DRUG{name: "C", primaryUse: "Epilepsy"}) , (p3:DRUG{name: "D", primaryUse: "Epilepsy"}) , (j1:COMPANY{name: "Pfizer"}) , (j2:COMPANY{name: "Novartis"}) , (s1:PATIENT{name: "Anna", DOB:Date("1999-05-04"), Gender:"F"}) , (s2:PATIENT{name: "Sandy", DOB:Date("1999-05-24"), Gender:"F"}) , (s3:PATIENT{name: "Rudy", DOB:Date("1999-1-26"), Gender:"M"}) , (p) -[r1:MADE_BY {SINCE:2015}]-> (j1) , (p1) -[r2:MADE_BY {SINCE:2017}]-> (j1) , (p2) -[r3:MADE_BY {SINCE:2017}]-> (j2) , (p3) -[r4:MADE_BY {SINCE:2017}]-> (j1) </pre> Write queries for Neo4j for the graph above: 1. Insert a new relationship between PATIENT Rudy and DRUG C that describes the fact that Rudy was prescribed drug C since 2022. 2. List the names of Drugs that have been made by Pfizer 3. Delete the node of Drug C.	1.0+1.0 +1.0 (LO 6)
Q5	Write the pseudocode to convert a star schema to a generic CASSANDRA data model OR Write the pseudocode to convert a star schema to fit a MONGODB database	5.0 (LO 7)
Q6	1. Both Data Warehouses and Data Lakes store data. How are they then different? 2. When would you use a Key-Value Data Store like Redis? Give two reasons. 3. Explain the terms multi-model database and polyglot persistence clearly.	2.0+1.0 +2.0 (LO 6,7)

