Roll	No
------	----



PRESIDENCY UNIVERSITY BENGALURU

SCHOOL OF INFORMATION SCIENCE MID TERM EXAMINATION - APR 2023

Semester : Semester IV - 2021 Course Code : CSA2021 Course Name : Sem IV - CSA2021 - Data Warehousing and Data Mining Program : BSD Date : 13-APR-2023 Time : 2:00PM - 3:30PM Max Marks : 50 Weightage : 25%

Instructions:

- (i) Read all questions carefully and answer accordingly.
- (ii) Question paper consists of 3 parts.
- (iii) Scientific and non-programmable calculator are permitted.
- (iv) Do not write any information on the question paper other than Roll Number.

PART A

ANSWER ALL THE QUESTIONS (5 X 2 = 10M) 1. Define Data warehouse? 2. Define Data mining . 3. Define Data Mart and mention it's advantages. 4. What are three factors considered for partial materialization of cuboids or subcubes? 4. What are three factors considered for partial materialization of cuboids or subcubes? 5. List any four commonly used OLAP operations for analyzing multidimensional data fromvarious perspectives in a data cube? (CO2) [Knowledge]

PART B

ANSWER ALL THE QUESTIONS

6. Imagine that a company has multiple departments with specific data needs, such as sales, marketing, and finance. Can you explain how a Data Mart could be used in this scenario to meet the unique data requirements of each department and mention the benefits it would provide for the company?

(CO1) [Comprehension]

(4 X 5 = 20M)

7. ABC Corporation has been collecting vast amounts of customer data over the years but is struggling to make sense of it all. The company wants to understand more about customer behavior and preferences, but analyzing the data manually is proving to be a monumental task. Explain why data mining is necessary in this scenario.

(CO1) [Comprehension]

8. Suppose you are working as a data analyst for a large manufacturing company. The company has a vast amount of transactional data stored in their database, including information about production, inventory, sales, and suppliers. The management team wants to analyze this data to gain insights into production trends across different time periods, items, locations, and suppliers. Draw the lattice of cuboids for the dimensions of time, item, location, and supplier to help company to gain insights.

(CO2) [Comprehension]

9. Assume you are a data analyst for AllElectronics, a large electronics retail company. The company has a vast amount of transactional data stored in their database, including information about sales, customers, products, and stores. The management team wants to improve the performance of the queries used to analyze the sales data. How would you use bitmap indexing to optimize the performance of the queries for the AllElectronics data set, and what benefits would this technique provide for the analysis of the data?

(CO2) [Comprehension]

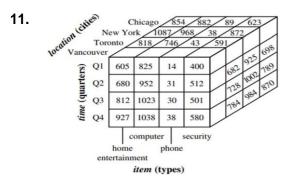
PART C

ANSWER ALL THE QUESTIONS

10. Imagine that a company has multiple departments with specific data needs, such as sales, marketing, and finance. Can you explain how different types of Data Marts, such as dependent and independent, could be used in this scenario to meet the unique data requirements of eachdepartment?

(CO1) [Application]

(2 X 10 = 20M)



Consider the above data cube and apply different OLAP operations such a Roll-

- up
- Drill-down
- · Slice and dice
- Pivot

(CO2) [Application]