

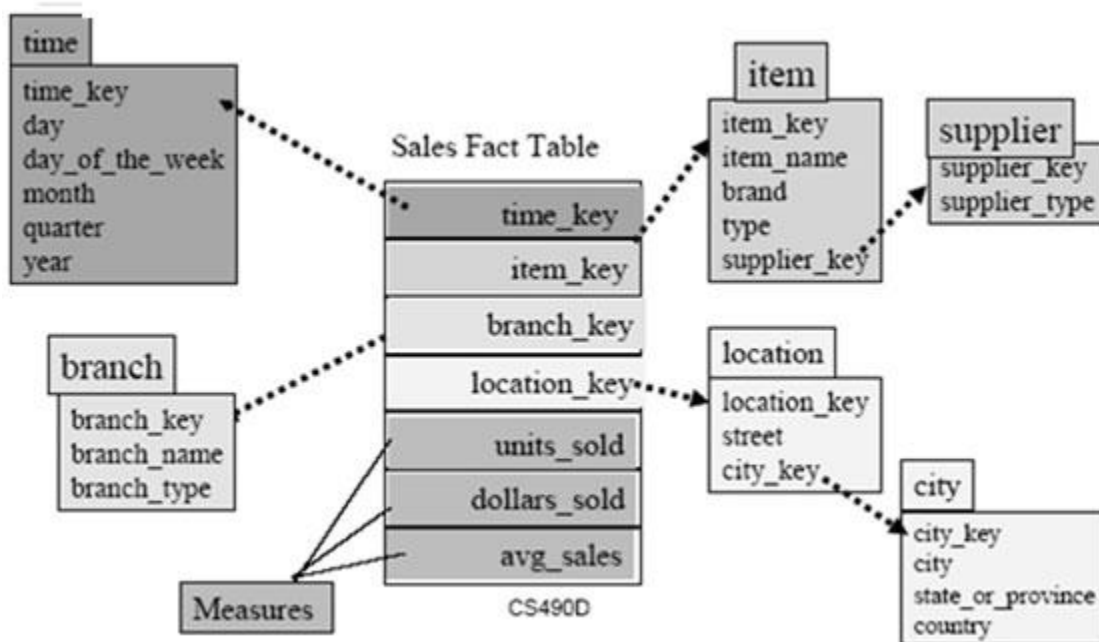
DEPARTMENT OF COMPUTER APPLICATIONS

MCA 4th Semester (2019 Batch)

Subject : **CA3405 Data Warehousing and Data Mining**

1. What is data warehousing? What are the characteristics of data warehouse?
2. Mention the various types of data available in data mining.
3. What is data mining? What are the various techniques of data mining? Explain in detail how data mining helps the business organizations.
4. What are the factors to be considered when comparing classification methods?
5. Explain data ware house architecture and operational data stores with neat diagram.
6. Discuss the importance of establishing standardized data mining query language .What are the potential benefits and challenges involved in such a task?
7. What is clustering? How does it differ from classification? Describe the following approaches to clustering methods: partitioning methods and hierarchical methods. Give an example for each.
8. Write down the applications of data warehousing.
9. What is a data mart? When is it appropriate? How data mart is different from data warehouse?
10. Discuss the typical OLAP operations in the multidimensional data models with the help of an example.
11. Why data cleaning is an important step in the discovery process? Explain various methods of data cleaning in detail.
12. What is spatial data mining? Give two applications of spatial data mining.
13. State various evaluation criteria that are essential for classification and prediction methods.
14. Discuss the functionality of Chameleon's clustering method with an example.
15. What do you mean by the concept hierarchy? Explain its importance in data mining.
16. What is meant by Multi level association rule? Discuss any two approaches for mining multilevel association rules with examples.
17. Define a spatial data cube. Discuss different types of dimensions in a spatial data cube.
18. What are the types of knowledge discovered during data mining? Explain with suitable example.

19. What is knowledge discovery? Explain in detail the steps involved in KDD process with an example?
20. List and explain the differences between OLAP and OLTP.
21. What is the difference between supervised and unsupervised learning? Give one example of each technique.
22. A data warehouse can be modeled by either a star schema or a snowflake schema. Briefly describe the similarities and the differences of the two models, and then analyze their advantages and disadvantages with regard to one another. Give your opinion of which might be more empirically useful and state the reasons for your opinion.
23. What is an artificial neural network? What is the difference between a feed-forward and back propagation network?
24. Following is a snowflake schema. Define the snowflake schema in DMQL.



25. On-Line Analytical Mining (OLAM) integrates on-line analytical processing with data mining and mining knowledge in multidimensional databases. Describe the importance of OLAM. Also explain an integrated OLAM and OLAP architecture with diagram.
26. What are Bayesian classifiers? Explain briefly Baye's theorem. Also explain how Naïve Bayesian classifier works?
27. Explain in detail the various backup and recovery procedures used in data warehouse.
28. What is genetic algorithm? Explain the process of genetic algorithm with the help of suitable flow diagram. Discuss the use of genetic algorithm in data warehouse and data mining.