

1. What is the primary purpose of building the multidimensional model?

The primary purpose of building the multidimensional model is to allow customers to interrogate analytical questions associated with market or business trends, unlike relational databases which allow customers to access data in the form of queries. They allow users to rapidly receive answers to the requests which they made by creating and examining the data comparatively fast. ⁶

2. What is the need of data warehouses?

Data warehouses are used to store a huge amount of data, which is typically collected from multiple heterogeneous sources like files, DBMS, etc. The goal is to produce statistical results that may help in decision making. ¹

3. How can you tell the difference between fact and dimension tables?

Fact tables contain records that combine attributes from different dimension tables. These records allow users to analyze different aspects of their business, which can aid in decision-making and improving the business. Dimension tables, on the other hand, provide the context and background information for the measures recorded in the fact table. One of the main differences between fact tables and dimension tables is that dimension tables contain the attributes that the measures in the fact table are based on. [^10^]

4. What are major issues that will be faced in data warehouse implementation?

Some of the difficulties of implementing data warehouses are: the design, construction and implementation of the warehouse, manual data processing, administration of a data warehouse, quality and consistency of data, data loss during ETL testing, data incompatibility, and absence of business course information. ¹⁹ [^20^]

5. What are the challenges in ETL process?

Some of the common challenges in ETL process are network latency, unoptimized code, not enough resources, data quality issues, and long-term maintenance. ¹⁵

6. What is the difference between ROLAP and MOLAP?

ROLAP (Relational Online Analytical Processing) and MOLAP (Multidimensional Online Analytical Processing) are two types of OLAP (Online Analytical Processing) systems. ROLAP uses a relational database to store and manage data, while MOLAP uses a multidimensional database. ⁶

7. Why is feature selection important in data mining?

Feature selection is important in data mining because it helps to reduce the dimensionality of the data, which can improve the accuracy and efficiency of the data mining process. ⁶

8. What do you understand by Data Mining?

Data mining is the process of discovering patterns in large datasets. ⁶

9. What are the major issues in data mining?

Some of the major issues in data mining are data quality, data privacy, and data security. ⁶

10. List the major clustering methods.

Some of the major clustering methods are K-means clustering, hierarchical clustering, and density-based clustering. ⁶

11. List the major classification methods

Some of the major classification methods are decision trees, logistic regression, and support vector machines. ⁶

12. Illustrate support vector machine with example.

Support vector machine (SVM) is a supervised machine learning algorithm that can be used for classification or regression tasks. SVM works by finding the hyperplane that best separates the data into different classes. ⁶

13. What are the advantages of spatial database?

Some of the advantages of spatial databases are that they can handle large amounts of spatial data, they can perform spatial queries efficiently, and they can support spatial indexing. ⁶

14. What is temporal database? Explain temporal database with example.

A temporal database is a database that stores data relating to time instances. For example, a temporal database might store information about the history of changes made to a particular record. ⁶

15. Explain about text mining and discuss about the challenges in text mining

Text mining is the process of extracting useful information from unstructured text data. Some of the challenges in text mining are dealing with large volumes of data, dealing with noisy data, and dealing with data in different languages. ⁶