

WORKSHEET-1

SQL ASSIGNMENT

Q1. A) Create, D) ALTER

Q2. A) Update, B) Delete

Q3. B) Structured Query Language

Q4. B) Data Definition Language

Q5. A) Data Manipulation Language

Q6. C) Create Table A (B int, C float)

Q7. B) Alter Table A ADD COLUMN D float

Q8. B) Alter Table A Drop Column D

Q9. B) Alter Table A Alter Column D int

Q10. C) Alter Table A Add Primary key B

Q11. What is data-warehouse?

Ans. Data warehousing is a collection of tools and techniques using which more knowledge can be driven out from a large amount of data. This helps with the decision-making process and improving information resources.

Data warehouse is basically a database of unique data structures that allows relatively quick and easy performance of complex queries over a large amount of data. It is created from multiple heterogeneous sources.

Q12. What is the difference between OLTP VS OLAP?

Ans.

BASIS FOR COMPARISON	OLTP	OLAP
Basic	It is an online transactional system and manages database modification.	It is an online data retrieving and data analysis system.

BASIS FOR COMPARISON	OLTP	OLAP
Focus	Insert, Update, Delete information from the database.	Extract data for analyzing that helps in decision making.
Data	OLTP and its transactions are the original source of data.	Different OLTPs database becomes the source of data for OLAP.
Transaction	OLTP has short transactions.	OLAP has long transactions.
Time	The processing time of a transaction is comparatively less in OLTP.	The processing time of a transaction is comparatively more in OLAP.
Queries	Simpler queries.	Complex queries.
Normalization	Tables in OLTP database are normalized (3NF).	Tables in OLAP database are not normalized.
Integrity	OLTP database must maintain data integrity constraint.	OLAP database does not get frequently modified. Hence, data integrity is not affected.

Q13. What are the various characteristics of data-warehouse?

Ans. Characteristics of data warehouse are:

1. Subject-Oriented

A data warehouse is subject oriented as it offers information regarding a theme instead of companies' ongoing operations. These subjects can be sales, marketing, distributions, etc.

A data warehouse never focuses on the ongoing operations. Instead, it put emphasis on modelling and analysis of data for **decision making**. It also provides a simple and concise view around the specific subject by excluding data which not helpful to support the decision process.

2. Integrated

In Data Warehouse, integration means the establishment of a common unit of measure for all similar data from the dissimilar database. The data also needs to be stored in the Datawarehouse in common and universally acceptable manner.

A data warehouse is developed by integrating data from varied sources like a mainframe, relational databases, flat files, etc. Moreover, it must keep consistent naming conventions, format, and coding.

This integration helps in effective analysis of data. Consistency in naming conventions, attribute measures, encoding structure etc. have to be ensured.

3. Time-Variant

The time horizon for data warehouse is quite extensive compared with operational systems. The data collected in a data warehouse is recognized with a particular period and offers information from the historical point of view. It contains an element of time, explicitly or implicitly.

One such place where Datawarehouse data display time variance is in in the structure of the record key. Every primary key contained with the DW should have either implicitly or explicitly an element of time. Like the day, week month, etc.

Another aspect of time variance is that once data is inserted in the warehouse, it can't be updated or changed.

4. Non-volatile

Data warehouse is also non-volatile means the previous data is not erased when new data is entered in it.

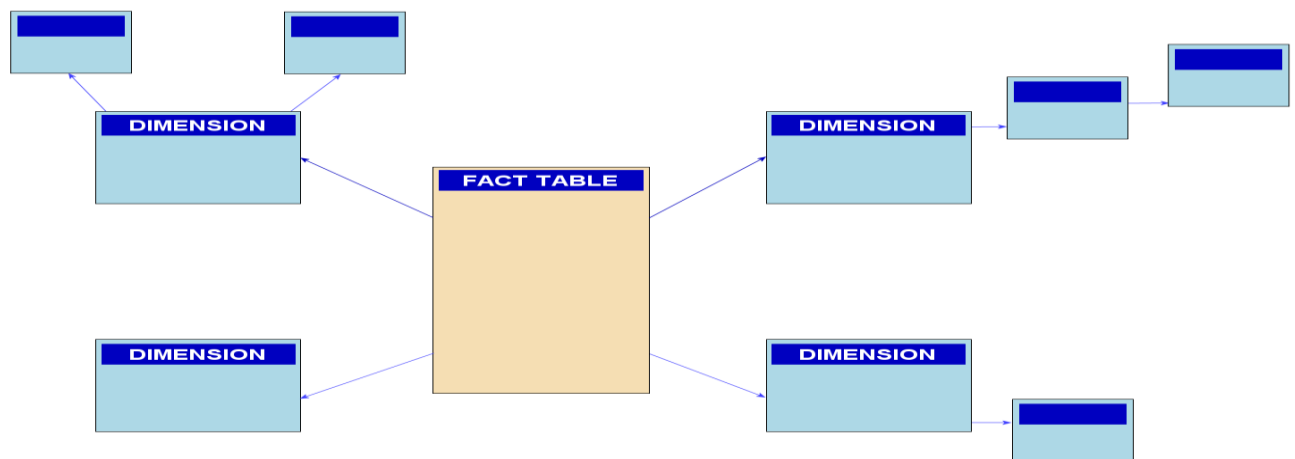
Data is read-only and periodically refreshed. This also helps to analyze historical data and understand what & when happened. It does not require transaction process, recovery and concurrency control mechanisms.

Activities like delete, update, and insert which are performed in an operational application environment are omitted in Data warehouse environment. Only two types of data operations performed in the Data Warehousing are

1. Data loading
2. Data access

Q14. What is Star-Schema?

Ans. The **star schema** is the simplest style of data mart schema and is the approach most widely used to develop data warehouses and dimensional data marts. The star schema consists of one or more fact tables referencing any number of dimension tables. The star schema is an important special case of the snowflake schema, and is more effective for handling simpler queries.



Q15. What do you mean by SETL?

Ans. SETL (SET Language) is a very high-level programming language based on the mathematical theory of sets. It was originally developed by (Jack) Jacob T. Schwartz at the New York University (NYU) Courant Institute of Mathematical Sciences in the late 1960s.

SETL provides two basic aggregate data types: unordered sets, and sequences (the latter also called tuples). The elements of sets and tuples can be of any arbitrary type, including sets and tuples themselves. Maps are provided as sets of pairs (i.e., tuples of length 2) and can have arbitrary domain and range types. Primitive operations in SETL include set membership, union, intersection, and power set construction, among others.

SETL provides quantified Boolean expressions constructed using the universal and existential quantifiers of first-order predicate logic.

SETL provides several iterators to produce a variety of loops over aggregate data structures.