SQL WORKSHEET-1

11. What is Data Warehouse?

A Data warehouse is a Relational Database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data, but it can include data from other sources. A Data warehouse is typically used to connect and analyse business data from heterogeneous sources. It is electronic storage of a large amount of information by a business which is designed for query and analysis instead of transaction processing. It is a process of transforming data into information and making it available to users in a timely manner to make a difference.

12. What is the difference between OLTP VS OLAP?

Difference between OLTP and OLAP:

* Online Analytical Processing (OLAP) is a category of software tools that analyse data stored in a database whereas Online transaction processing (OLTP) supports transaction-oriented applications in a 3-tier architecture.
* OLAP creates a single platform for all type of business analysis needs which includes planning, budgeting, forecasting, and analysis while OLTP is useful to administer day to day transactions of an organization.
* OLAP is characterized by a large volume of data while OLTP is characterized by large numbers of short online transactions.
* In OLAP, data warehouse is created uniquely so that it can integrate different data sources for building a consolidated database whereas OLTP uses traditional DBMS.

13. What are the various characteristics of the Data Warehouse?

There are three prominent data warehouse characteristics:

* Integrated: The way data is extracted and transformed is uniform, regardless of the original source.
* Time-variant: Data is organized via time-periods (weekly, monthly, annually, etc.).
* Non-volatile: A data warehouse is not updated in real-time

14. What is Star Schema?

In [computing](https://en.wikipedia.org/wiki/Computing), the star schema is the simplest style of [data mart](https://en.wikipedia.org/wiki/Data_mart) [schema](https://en.wikipedia.org/wiki/Logical_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](https://en.wikipedia.org/wiki/Fact_table) referencing any number of [dimension tables](https://en.wikipedia.org/wiki/Dimension_(data_warehouse)). The star schema is an important special case of the [snowflake schema](https://en.wikipedia.org/wiki/Snowflake_schema), and is more effective for handling simpler queries. The star schema separates business process data into facts, which hold the measurable, quantitative data about a business, and dimensions which are descriptive attributes related to fact data. Examples of fact data include sales price, sale quantity, and time, distance, speed and weight measurements. Related dimension attribute examples include product models, product colors, product sizes, geographic locations, and salesperson names.

15. What do you mean by SETL?

**ETL** is a process that extracts the data from different source systems, then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system. Full form of ETL is Extract, Transform and Load. ETL is a recurring activity (daily, weekly, monthly) of a Data warehouse system and needs to be agile, automated, and well documented. ETL provides a method of moving the data from various sources into a data warehouse. As data sources change, the Data Warehouse will automatically update.