SQL – Worksheet-1

1. Which of the following is/are DDL commands in SQL?

A) Create B) Update

C) Delete D) ALTER

Answer : Update

2. Which of the following is/are DML commands in SQL?

A) Update B) Delete

C) Select D) Drop

Answer : Drop

3. Full form of SQL is:

A) Strut querying language B) Structured Query Language

C) Simple Query Language D) None of them

Answer : Structured Query Language

4. Full form of DDL is:

A) Descriptive Designed Language B) Data Definition Language

C) Data Descriptive Language D) None of the above.

Answer : Data Definition language

5. DML is:

A) Data Manipulation Language B) Data Management Language

C) Data Modeling Language D) None of these

Answer : Data Manipulation Language

6. Which of the following statements can be used to create a table with column B int type and C float type?

A) Table A (B int, C float) B) Create A (b int, C float)

C) Create Table A (B int,C float) D) All of them

Answer : Create Table A(B int , C float)

7. Which of the following statements can be used to add a column D (float type) to the table A created above?

A) Table A (D float) B) Alter Table A ADD COLUMN D float

C) Table ACB int, C float, D float) D) None of them

Answer : Alter Table A ADD COLUMN D float

8. Which of the following statements can be used to drop the column added in the above question?

A) Table A Drop D B) Alter Table A Drop Column D

C) Delete D from A D) None of them

Answer : Alter Table A Drop Column D

9. Which of the following statements can be used to change the data type (from float to int) of the column D of table A created in above questions?

A) Table A (D float int) B) Alter Table A Alter Column D int

C) Alter Table A D float int D) Alter table A Column D float to int

Answer : Alter Table A Alter Column D int

10. Suppose we want to make Column B of Table A as primary key of the table. By which of the following statements we can do it?

A) Alter Table A Add Constraint Primary Key B B) Alter table (B primary key)

C) Alter Table A Add Primary key B D) None of them

Answer : Alter Table A Add Primary key B

11.What is data-warehouse?

Answer : A Data Warehousing (DW) is process for collecting and managing data from varied sources to provide meaningful business insights. A Data warehouse is typically used to connect and analyze business data from heterogeneous sources. The data warehouse is the core of the BI system which is built for data analysis and reporting.
It is a blend of technologies and components which aids the strategic use of data. 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.

The decision support database (Data Warehouse) is maintained separately from the organization's operational database. However, the data warehouse is not a product but an environment. It is an architectural construct of an information system which provides users with current and historical decision support information which is difficult to access or present in the traditional operational data store.

You many know that a 3NF-designed database for an inventory system many have tables related to each other. For example, a report on current inventory information can include more than 12 joined conditions. This can quickly slow down the response time of the query and report. A data warehouse provides a new design which can help to reduce the response time and helps to enhance the performance of queries for reports and analytics.

Data warehouse system is also known by the following name:

- **Decision Support System (DSS)**
- **Executive Information System**
- **Management Information System**
- **Business Intelligence Solution**
- **Analytic Application**
- **Data Warehouse**

The Datawarehouse benefits users to understand and enhance their organization's performance. The need to warehouse data evolved as computer systems became more complex and needed to handle increasing amounts of Information. However, Data Warehousing is a not a new thing.

12.What is the difference between OLTP VS OLAP?

Answer : OLTP and OLAP: The two terms look similar but refer to different kinds of systems. Online transaction processing (OLTP) captures, stores, and processes data from

transactions in real time. Online analytical processing (OLAP) uses complex queries to analyze aggregated historical data from OLTP systems.

OLTP : An OLTP system captures and maintains transaction data in a database. Each transaction involves individual database records made up of multiple fields or columns. Examples include banking and credit card activity or retail checkout scanning.

In OLTP, the emphasis is on fast processing, because OLTP databases are read, written, and updated frequently. If a transaction fails, built-in system logic ensures data integrity.

OLAP: OLAP applies complex queries to large amounts of historical data, aggregated from OLTP databases and other sources, for data mining, analytics, and business intelligence projects. In OLAP, the emphasis is on response time to these complex queries. Each query involves one or more columns of data aggregated from many rows. Examples include year-over-year financial performance or marketing lead generation trends. OLAP databases and data warehouses give analysts and decision-makers the ability to use custom reporting tools to turn data into information. Query failure in OLAP does not interrupt or delay transaction processing for customers, but it can delay or impact the accuracy of business intelligence insights.

Let' have side by side comparison of OLTP & OLAP:

Characteristics	Handles a large number of small transactions	Handles large volumes of data with complex queries
Query types	Simple standardized queries	Complex queries
Operations	Based on INSERT, UPDATE, DELETE commands	Based on SELECT commands to aggregate data for reporting
Response time	Milliseconds	Seconds, minutes, or hours depending on the amount of data to process
Design	Industry-specific, such as retail, manufacturing, or banking	Subject-specific, such as sales, inventory, or marketing
Source	Transactions	Aggregated data from transactions
Purpose	Control and run essential business operations in real time	Plan, solve problems, support decisions, discover hidden insights

Data updates	Short, fast updates initiated by user	Data periodically refreshed with scheduled, long-running batch jobs
Space requirements	Generally small if historical data is archived	Generally large due to aggregating large datasets

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

Answer : Data warehouses are characterized by being:

Subject-oriented: A data warehouse typically provides information on a topic (such as a sales inventory or supply chain) rather than company operations.

Time-variant: Time variant keys (e.g., for the date, month, time) are typically present.

Integrated: A data warehouse combines data from various sources. These may include a cloud, relational databases, flat files, structured and semi-structured data, metadata, and master data. The sources are combined in a manner that's consistent, relatable, and ideally certifiable, providing a business with confidence in the data's quality.

Persistent and non-volatile: Prior data isn't deleted when new data is added. Historical data is preserved for comparisons, trends, and analytics.

Data warehouse components are engineered for speed. When results are accessible quickly, they can be analyzed on the fly.

Key attributes of most data warehouses are:

- They are often deployed as a central database for the enterprise.

- Provide ETL (extract, transform, load) data processing capability. The ETL process requires a staging area where data is transformed before it enters the data warehouse for analysis. NOTE: An alternate pattern could exist wherein data can be transformed upstream in data lakes as well and fed into a data warehouse
- Store metadata.
- Include access to reporting tools. BI tools such as PowerBI and Tableau may connect to the Warehouse through built-for-purpose drivers or they may leverage SQL for queries—albeit it is important to guard against unbounded queries that could impact performance. Often, there are a set of certified reports that are frequently and automatically refreshed.
- A cloud data warehouse can help you break down data silos and get faster time to value to support your organization's growth initiatives.

14. What is Star Schema?

Answer : A star schema is the elementary form of a dimensional model, in which data are organized into **facts** and **dimensions**. A fact is an event that is counted or measured, such as a sale or log in. A dimension includes reference data about the fact, such as date, item, or customer.

A star schema is a relational schema where a relational schema whose design represents a multidimensional data model. The star schema is the explicit data warehouse schema. It is known as **star schema** because the entity-relationship diagram of this schemas simulates a star, with points, diverge from a central table. The center of the schema consists of a large fact table, and the points of the star are the dimension tables.

Fact Tables

A table in a star schema which contains facts and connected to dimensions. A fact table has two types of columns: those that include fact and those that are foreign keys to the dimension table. The primary key of the fact tables is generally a composite key that is made up of all of its foreign keys.

A fact table might involve either detail level fact or fact that have been aggregated (fact tables that include aggregated fact are often instead called summary tables). A fact table generally contains facts with the same level of aggregation.

A star schema is a multi-dimensional data model used to organize data in a database so that it is easy to understand and analyze. Star schemas can be applied to data warehouses, databases, data marts, and other tools. The star schema design is optimized for querying large data sets.

A star schema is used to denormalize business data into dimensions (like time and product) and facts (like transactions in amounts and quantities).

A star schema has a single fact table in the center, containing business "facts" (like transaction amounts and quantities). The fact table connects to multiple other dimension tables along "dimensions" like time, or product. Star schemas enable users to slice and dice the data however they see fit, typically by joining two or more fact tables and dimension tables together.

Star schemas **denormalize** the data, which means adding redundant columns to some dimension tables to make querying and working with the data faster and easier. The purpose is to trade some redundancy (duplication of data) in the data model for increased query speed, by avoiding computationally expensive join operations.

In this model, the fact table is normalized but the dimensions tables are not. That is, data from the fact table exists only on the fact table, but dimensional tables may hold redundant data.

15.What do you mean by SETL?

Answer:

SET Theory Language: A programming language developed by Jack Schwartz in the early 1970s. It is based on set theory and used for mathematical and telecommunications applications.

SETL is a very-high level language with dynamic typing and dynamic data structures, based on the mathematical notion of set. It was designed in the very early 1970s by J. Schwartz – a renown mathematician, with the help of R. Dewar and others. The language introduced a fundamentally new paradigm in programming in which sets, ordered sets and maps are the principal data structures and the programs are expressed in terms of set constructors, set operations, and predicates on sets. The very name SETL is an abbreviation of 'SET Language'.

The set-oriented paradigm is based on the assumption that sets are as essential constructions in programming as they are in mathematics. SETL not only has extensive provision for programming with sets but also takes advantage of the syntactic tradition of abstract mathematics, the language's notation being very close to the one of set theory, thus making it possible to express many algorithms in a familiar, natural, and concise manner.

SETL programs are much more declarative than procedural. According to its author, the language should present 'an abstract but nevertheless executable notation for describing algorithms'. In his view, SETL should provide means for solving programming problems more or less at the semantic level of the problems themselves, postponing possible choices of more detailed encoding 'until logical structure is worked out'. This implies that SETL programs would typically be considered prototypes rather than real implementations. On the other hand, the implementations should be obtainable by staged refinement, also aided by the language.

The primitive datatypes of SETL include integer, floating-point, Boolean, atom, and string. Integers are of unlimited magnitude. Atoms are unique values produced by a dedicated operation and are used to tag other values as an aid in constructing 'data structuring maps'. Strings, although counting as a primitive datatype, are operationally similar to ordered sets.

Both sets and ordered sets (called tuples) are heterogeneous in SETL: values of all sorts can be members of the same structure, including other sets and tuples. Nesting is unrestricted. Beyond sets and tuples, maps are the only other data structure provided in the language. A map is actually a set (and is denoted no differently) – one whose elements are pairs, i.e. tuples of length two. The sets of the first and the second items of those pairs constitute the domain and the range of the map, correspondingly. Although maps are not a separate datatype, they are supported in SETL by specific operations, besides those for sets.