**WORKSHEET-1**

**SQL**

**Q1 and Q2 have one or more correct answer. Choose all the correct option to answer your question.**

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

**Answer:** A) Create D) ALTER

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

**Answer:** A) Update B) Delete C) Select

**Q3 to Q10 have only one correct answer. Choose the correct option to answer your question.**

3. Full form of SQL is:

**Answer:** B) Structured Query Language

4. Full form of DDL is:

**Answer:** B) Data Definition Language

5. DML is:

**Answer:** A) 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?

**Answer:** C) 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?

**Answer:** B) 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?

**Answer:** B) 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?

**Answer:** B) 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?

**Answer:** C) Alter Table A Add Primary key B

**Q11 to Q15 are subjective answer type questions, Answer them briefly.**

11. What is data-warehouse?

**Answer:** A data warehouse is a type of data management system that is designed to enable and support business intelligence (BI) activities, especially analytics. Data warehouses are solely intended to perform queries and analysis and often contain large amounts of historical data. The data within a data warehouse is usually derived from a wide range of sources such as application log files and transaction applications.

A data warehouse centralizes and consolidates large amounts of data from multiple sources. Its analytical capabilities allow organizations to derive valuable business insights from their data to improve decision-making. Over time, it builds a historical record that can be invaluable to data scientists and business analysts. Because of these capabilities, a data warehouse can be considered an organization’s “single source of truth.”

A typical data warehouse often includes the following elements:

* A relational database to store and manage data
* An extraction, loading, and transformation (ELT) solution for preparing the data for analysis
* Statistical analysis, reporting, and data mining capabilities
* Client analysis tools for visualizing and presenting data to business users
* Other, more sophisticated analytical applications that generate actionable
* information by applying machine learning and artificial intelligence (AI) algorithms

12. What is the difference between OLTP VS OLAP?

**Answer:** OLTP vs. OLAP: side-by-side comparison

OLTP is operational, while OLAP is informational. A glance at the key features of both kinds of processing illustrates their fundamental differences, and how they work together.

|  |  |  |
| --- | --- | --- |
|  | **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 |
| **Backup and recovery** | Regular backups required to ensure business continuity and meet legal and governance requirements | Lost data can be reloaded from OLTP database as needed in lieu of regular backups |
| **Productivity** | Increases productivity of end users | Increases productivity of business managers, data analysts, and executives |
| **Data view** | Lists day-to-day business transactions | Multi-dimensional view of enterprise data |
| **User examples** | Customer-facing personnel, clerks, online shoppers | Knowledge workers such as data analysts, business analysts, and executives |
| **Database design** | Normalized databases for efficiency | Denormalized databases for analysis |

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

**Answer:** 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. It is periodically updated via the uploading of data, protecting it from the influence of momentary change.

14. What is Star-Schema??

**Answer:** In the STAR Schema, the center of the star can have one fact table and a number of associated dimension tables. It is known as star schema as its structure resembles a star. The star schema is the simplest type of Data Warehouse schema. It is also known as Star Join Schema and is optimized for querying large data sets.

In the following example,the fact table is at the center which contains keys to every dimension table like Dealer\_ID, Model ID, Date\_ID, Product\_ID, Branch\_ID & other attributes like Units sold and revenue.

Characteristics of Star Schema:

* Every dimension in a star schema is represented with the only one-dimension table.
* The dimension table should contain the set of attributes.
* The dimension table is joined to the fact table using a foreign key
* The dimension table are not joined to each other
* Fact table would contain key and measure
* The Star schema is easy to understand and provides optimal disk usage.
* The dimension tables are not normalized. For instance, in the above figure, Country\_ID does not have Country lookup table as an OLTP design would have.
* The schema is widely supported by BI Tools

15. What do you mean by SETL?

**Answer:** Short for Set Theory as a Language (or Set Language), SETL is a high-level programming language that's based on the mathematical theory of sets. It was developed in the early 1970's by mathematician Professor J. Schwartz. SETL is an interpreted language with a syntax that is resembles C and, in many cases, similar to Perl. In SETL every statement is terminated by a semicolon. Variable names are case-insensitive and are automatically determined by their last assignment.