WORKSHEET-1

**SQL**

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

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

11. What is data-warehouse?

Answer: Data warehouse is a system used for storing and reporting on data. The data typically originates in multiple systems, then it is moved into the data warehouse for long-term storage and analysis. This storage is structured such that users from many divisions or departments within an organization can access and analyse the data according to their needs. It support analytical reporting, structured and/or ad hoc queries, and decision making. Data warehousing involves data cleaning, data integration, and data consolidations. The information gathered in a warehouse can be used in any of the following domains –

1. Tuning Production Strategies
2. Customer Analysis
3. Operations Analysis

Data warehouses provides:

* A single point of access for all data, rather than requiring users to connect to dozens or even hundreds of systems individually
* An assurance of data quality
* A history of the data they store
* Separation between day-to-day operational systems and analytical systems, for security reasons

12. What is the difference between OLTP VS OLAP?

Answer: OLAP is Online Analytical Processing, a category of software tools which provide analysis of data for business decisions. OLAP systems allow users to analyse database information from multiple database systems at one time. The primary objective is data analysis and not data processing.

OLTP is Online transaction processing, it supports transaction-oriented applications in a 3-tier architecture. OLTP administers day to day transaction of an organization. The primary objective is data processing and not data analysis.

|  |  |
| --- | --- |
| OLAP | OLTP |
| Consists of historical data from various Databases. | Consists only operational current data. |
| It is subject oriented. Used for Data Mining, Analytics, Decision making,etc. | It is application oriented. Used for business tasks. |
| The data is used in planning, problem solving and decision making. | The data is used to perform day to day fundamental operations. |
| It reveals a snapshot of present business tasks. | It provides a multi-dimensional view of different business tasks. |
| Large amount of data is stored typically in TB, PB | The size of the data is relatively small as the historical data is archived. For ex MB, GB |
| Relatively slow as the amount of data involved is large. Queries may take hours. | Very Fast as the queries operate on 5% of the data. |
| It only need backup from time to time as compared to OLTP. | Backup and recovery process is maintained religiously |
| This data is generally managed by CEO, MD, GM. | This data is managed by clerks, managers. |
| Only read and rarely write operation. | Both read and write operations. |

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

Answer: Below are major characteristics of data warehouse:

Subject-oriented –  
A data warehouse is always a subject oriented as it delivers information about a theme instead of organization’s current operations. It can be achieved on specific theme.

Integrated –  
It is somewhere same as subject orientation which is made in a reliable format. Integration means founding a shared entity to scale the all similar data from the different databases. The data also required to be resided into various data warehouse in shared and generally granted manner.

Time-variant: Data is organized via time-periods (weekly, monthly, annually, etc.). In this data is maintained via different intervals of time such as weekly, monthly, or annually etc. It founds various time limit which are structured between the large datasets and are held in online transaction process (OLTP). The time limits for data warehouse is wide-ranged than that of operational systems

Non-Volatile –  
As the name defines the data resided in data warehouse is permanent. It also means that data is not erased or deleted when new data is inserted. It includes the mammoth quantity of data that is inserted into modification between the selected quantities on logical business. It evaluates the analysis within the technologies of warehouse.

14. What is Star-Schema??

Answer: A star schema is the simplest form of a dimensional model, in which data is organized into facts and dimensions.  A fact is an event that is counted or measured, such as a sale or login.  A dimension contains reference information about the fact, such as date, product, or customer. A star schema is diagrammed by surrounding each fact with its associated dimensions. The resulting diagram resembles a star.

Star schemas are optimized for querying large data sets and are used in data warehouses and data marts to support [OLAP](https://searchdatamanagement.techtarget.com/definition/OLAP) cubes, business intelligence and analytic applications, and [ad hoc](https://searchbusinessanalytics.techtarget.com/definition/ad-hoc-analysis) queries.

15. What do you mean by SETL?

Answer: SETL is Semantic extract-transform-load framework for semantic data warehouses

* SETL builds on Semantic Web (SW) standards and tools.
* SETL provides a number of powerful modules, classes, and methods for (dimensional and semantic) DW constructs and tasks.
* SETL supports semantic and traditional data sources, semantic integration, and creating or publishing a (MD) semantic DW.
* Using SETL, we perform a comprehensive experimental evaluation by producing a MD semantic DW that integrates a semantic and non-semantic data sources.
* The evaluation shows that SETL improves considerably over the competing solutions/tools in terms of productivity, and performance.