**SQL WORKSHEET**

**Q1.** A, D

**Q2.** A, B, C

**Q3.** B

**Q4.** B

**Q5.** A

**Q6.** C

**Q7.** B

**Q8.** B

**Q9.** D

**Q10.** A

**Q11.** A **Data warehousing** 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.

**Q12.**

| **Parameters** | **OLTP** | **OLAP** |
| --- | --- | --- |
| **Process** | It is an online transactional system. It manages database modification. | OLAP is an online analysis and data retrieving process. |
| **Characteristic** | It is characterized by large numbers of short online transactions. | It is characterized by a large volume of data. |
| **Functionality** | OLTP is an online database modifying system. | OLAP is an online database query management system. |
| **Method** | OLTP uses traditional DBMS. | OLAP uses the data warehouse. |
| **Query** | Insert, Update, and Delete information from the database. | Mostly select operations |
| **Table** | Tables in OLTP database are normalized. | Tables in OLAP database are not normalized. |
| **Source** | OLTP and its transactions are the sources of data. | Different OLTP databases become the source of data for OLAP. |
| **Data Integrity** | OLTP database must maintain data integrity constraint. | OLAP database does not get frequently modified. Hence, data integrity is not an issue. |
| **Response time** | It's response time is in millisecond. | Response time in seconds to minutes. |
| **Data quality** | The data in the OLTP database is always detailed and organized. | The data in OLAP process might not be organized. |
| **Operation** | Allow read/write operations. | Only read and rarely write. |
| **Audience** | It is a market orientated process. | It is a customer orientated process. |

**Q13.** Data warehouse can be controlled when the user has a shared way of explaining the trends that are introduced as specific subject. Below are major characteristics of data warehouse:

1. **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. That means the data warehousing process is proposed to handle with a specific theme which is more defined. These themes can be sales, distributions, marketing etc.
2. **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. In addition, it must have reliable naming conventions, format and codes. Integration of data warehouse benefits in effective analysis of data. Reliability in naming conventions, column scaling, encoding structure etc. should be confirmed. Integration of data warehouse handles various subject related warehouse.
3. **Time-Variant –**  
    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. The time limits for data warehouse is wide-ranged than that of operational systems. The data resided in data warehouse is predictable with a specific interval of time and delivers information from the historical perspective. It comprises elements of time explicitly or implicitly. Another feature of time-variance is that once data is stored in the data warehouse then it cannot be modified, alter, or updated
4. **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 quantity on logical business. It evaluates the analysis within the technologies of warehouse.

**Q14. Star schema** is the fundamental schema among the data mart schema and it is simplest. This schema is widely used to develop or build a data warehouse and dimensional data marts. It includes one or more fact tables indexing any number of dimensional tables. The star schema is a necessary case of the snowflake schema. It is also efficient for handling basic queries.

It is said to be star as its physical model resembles to the star shape having a fact table at its center and the dimension tables at its peripheral representing the star’s points. Below is an example to demonstrate the Star Schema:

**Advantages of Star Schema –**

1. Simpler Queries
2. Simplified Business Reporting Logic
3. Feeding Cubes
4. Query performance gains

**Disadvantages of Star Schema –**

1. Data integrity is not enforced well since in a highly de-normalized schema state.
2. Not flexible in terms if analytical needs as a normalized data model.
3. Star schemas don’t reinforce many-to-many relationships within business entities – at least not frequently.

**Q15.** SETL (SET Language) is a very high-level programming language based on the mathematical theory of sets.