**SQL Worksheet I**

1. A) and B)
2. A), B) and C)
3. B)
4. B)
5. A)
6. C)
7. B)
8. B)
9. B)
10. A)
11. A data warehouse is a system used for reporting and data analysis. They store current and historical data in one single place that are used for creating analytical reports for workers throughout the enterprise. 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.
12. OLAP: Online Analytical Processing, a category of software tools which provide analysis of data for business decisions. OLAP systems allow users to analyze database information from multiple database systems at one time.

**The primary objective is data analysis and not data processing**.

OLTP: Online transaction processing shortly known as OLTP 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**

1. Following are the characteristics of a Data Warehouse:

1. Subject oriented

A data warehouse is subject-oriented, as it provides information on a topic rather than the ongoing operations of organizations. Such issues may be inventory, promotion, storage, etc. Never does a data warehouse concentrate on the current processes. Instead, it emphasized modeling and analyzing decision-making data. It also provides a simple and succinct description of the particular subject by excluding details that would not be useful in helping the decision process.

2. Integrated

Integration in Data Warehouse means establishing a standard unit of measurement from the different databases for all the similar data. The data must also get stored in a simple and universally acceptable manner within the Data Warehouse. Through combining data from various sources such as a mainframe, relational databases, flat files, etc., a data warehouse is created. It must also keep the naming conventions, format, and coding consistent. Such an application assists in robust data analysis. Consistency must be maintained in naming conventions, measurements of characteristics, specification of encoding, etc.

3. Time-variant

Compared to operating systems, the time horizon for the data warehouse is quite extensive. The data collected in a data warehouse is acknowledged over a given period and provides historical information. It contains a temporal element, either explicitly or implicitly.

One such location in the record key system where Data Warehouse data shows time variation is. Each primary key contained with the DW should have an element of time either implicitly or explicitly. Just like the day, the month of the week, etc.

4. Non-volatile

Also, the data warehouse is non-volatile, meaning that prior data will not be erased when new data are entered into it. Data is read-only, only updated regularly. It also assists in analyzing historical data and in understanding what and when it happened. The transaction process, recovery, and competitiveness control mechanisms are not required. In the Data Warehouse environment, activities such as deleting, updating, and inserting that are performed in an operational application environment are omitted.

1. 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.
2. SETL databases provide an alternative to standard SQL databases. Each SETL database object is a potentially very large (up to hundreds of gigabytes) abstract object stored on disk. Each such database object resembles a map from a set of system generated 4- or 5-byte record identifiers to the records which the database stores, each of which is a fully general object of the SETL language, but which is itself not excessively large (e.g. no more than a few dozen megabytes.) The database records are systematically indexed for efficient access, using the abstract technique described below. Once a record of a SETL database has been accessed, all the operations of the SETL language can be applied to it. Conversely, any standard SETL object can be stored as a database record.