



SQL – ASSIGNMENT 1

FlipRobo Internship



APRIL 15, 2021

MAHABOOB BASHA SHAIK

Assignment 1 – Worksheet SQL

- 1) Option A & D
- 2) Option A & B
- 3) Option B
- 4) Option B
- 5) Option A
- 6) Option C
- 7) Option B
- 8) Option B
- 9) Option D
- 10) Option C

Subjective Answers:

- 11) The process of collecting and managing data from various sources to obtain meaningful insights for business. The data warehouse is used to connect and analysis data from various heterogeneous sources.

12)

OLAP : Online Analytical Processing	OLTP : Online Transaction Processing
1) Software tools that analyse data stored in a database	1) Transaction-oriented applications in a 3-tier architecture.
2) Single platform for all type of business analysis needs which includes planning, budgeting, forecasting, and analysis	2) administer day to day transactions of an organization.
3) Characterized by a large volume of data	3) Characterized by large numbers of short online transactions.
4) Created uniquely so that it can integrate different data sources for building a consolidated database	4) Uses traditional DBMS.

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

Subject Oriented

A data warehouse is subject oriented because it provides information around a subject rather than the organization's ongoing operations.

These subjects can be product, customers, suppliers, sales, revenue, etc. A data warehouse does not focus on the ongoing operations, rather it focuses on modelling and analysis of data for decision making.

Integrated

A data warehouse is constructed by integrating data from heterogeneous sources such as relational databases, flat files, etc.

This integration enhances the effective analysis of data.

Time Variant

The data collected in a data warehouse is identified with a particular time period.

The data in a data warehouse provides information from the historical point of view.

Non-volatile

Non-volatile means the previous data is not erased when new data is added to it.

The key characteristics of a data warehouse are as follows:

- Some data is denormalized for simplification and to improve performance.
- Queries often retrieve large amounts of data.
- Both planned and ad hoc queries are common.
- The data load is controlled.

14) 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 centre and the dimension tables at its peripheral representing the star's points.

Advantages of Star Schema –

Simpler Queries:

Join logic of star schema is quite cinch in compare to other join logic which are needed to fetch data from a transactional schema that is highly normalized.

Simplified Business Reporting Logic:

In compared to a transactional schema that is highly normalized, the star schema makes simpler common business reporting logic, such as as-of reporting and period-over-period.

Feeding Cubes:

Star schema is widely used by all OLAP systems to design OLAP cubes efficiently. In fact, major OLAP systems deliver a ROLAP mode of operation which can use a star schema as a source without designing a cube structure.

Disadvantages of Star Schema –

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

15) A SET is a string object that can have zero or more values, each of which must be chosen from a list of permitted values specified when the table is created. SET column values that consist of multiple set members are specified with members separated by commas (.). A consequence of this is that SET member values should not themselves contain commas.

- a. A SET column can have a maximum of 64 distinct members.
- b. SET values numerically, with the low-order bit of the stored value corresponding to the first set member.
- c. It can handle multiple values without much trouble.
- d. It allows us to compare multiple values without using complex JOIN operations.
- e. Its schema is simple that uses only one column instead of three different tables (for example, Person, Interest, Link the persons to particular interest) to store elements.
- f. It allows us to use binary functions for complex comparisons by comparing bit values on a single column.