|  |  |
| --- | --- |
|  |  |
|  | **SQL worksheet**  1) a,d |
|  |  |
|  | 2) a.b |
|  |  |
|  | 3) b |
|  |  |
|  | 4) b |
|  |  |
|  | 5) a |
|  |  |
|  | 6) c |
|  |  |
|  | 7) b |
|  |  |
|  | 8) b |
|  |  |
|  | 9) b |
|  |  |
|  | 10) c |
|  |  |
|  | 11) A data warehouse is a large collection of business data used to help an organization make decisions. The large amount of data in data warehouses comes from different places such as internal applications such as marketing, sales, and finance; customer-facing apps; and external partner systems, among others. |
|  | On a technical level, a data warehouse periodically pulls data from those apps and systems; then, the data goes through formatting and import processes to match the data already in the warehouse. The data warehouse stores this processed data so it’s ready for decision makers to access. How frequently data pulls occur, or how data is formatted, etc., will vary depending on the needs of the organization. |
|  |  |
|  | 12) Following are the difference between OLAP and OLTP system: |
|  | a) Users: OLTP systems are designed for office worker while the OLAP systems are designed for decision-makers. Therefore while an OLTP method may be accessed by hundreds or even thousands of clients in a huge enterprise, an OLAP system is suitable to be accessed only by a select class of manager and may be used only by dozens of users. |
|  | b) Functions: OLTP systems are mission-critical. They provide day-to-day operations of an enterprise and are largely performance and availability driven. These operations carry out simple repetitive operations. OLAP systems are management-critical to support the decision of enterprise support tasks using detailed investigation. |
|  | c) Nature: Although SQL queries return a set of data, OLTP methods are designed to step one record at the time, for example, a data related to the user who may be on the phone or in the store. OLAP system is not designed to deal with individual customer records. Instead, they include queries that deal with many data at a time and provide summary or aggregate information to a manager. OLAP applications include data stored in a data warehouses that have been extracted from many tables and possibly from more than one enterprise database. |
|  | d) Design: OLTP database operations are designed to be application-oriented while OLAP operations are designed to be subject-oriented. OLTP systems view the enterprise record as a collection of tables (possibly based on an entity-relationship model). OLAP operations view enterprise information as multidimensional). |
|  | e) Data: OLTP systems usually deal only with the current status of data. For example, a record about an employee who left three years ago may not be feasible on the Human Resources System. The old data may have been achieved on some type of stable storage media and may not be accessible online. On the other hand, OLAP systems needed historical data over several years since trends are often essential in decision making. |
|  | f) Kind of use: OLTP methods are used for reading and writing operations while OLAP methods usually do not update the data. |
|  | g) Access Patterns: The access pattern of an OLTP system consist primarily of short, atomic transactions. Such a system needed concurrency control and recovery techniques. However, access to OLAP systems is mostly read-only operations because these data warehouses store historical information. |
|  | The biggest difference between an OLTP and OLAP system is the amount of data analyzed in a single transaction. Whereas an OLTP handles many concurrent customers and queries touching only a single data or limited collection of records at a time, an OLAP system must have the efficiency to operate on millions of data to answer a single query. |
|  |  |
|  | 13) A common way of introducing data warehousing is to refer to the characteristics of a data warehouse as set forth by William Inmon. Following are the characteristics of data-warehouse: |
|  | a) 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. |
|  | b) 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. |
|  | c) 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. |
|  | d) Non-volatile: |
|  | Non-volatile means the previous data is not erased when new data is added to it. |
|  |  |
|  | 14) In data warehousing and business intelligence (BI), 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 diagramed 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 cubes, business intelligence and analytic applications, and ad hoc queries. |
|  |  |
|  | 15) SETL (SET Language) is a very high-level programming language based on the mathematical theory of sets. It was originally developed by Jacob T. Schwartz at the New York University (NYU) Courant Institute of Mathematical Sciences in the late 1960s. |
|  | SETL provides two basic aggregate data types: unordered sets, and sequences. The elements of sets and tuples can be of any arbitrary type, including sets and tuples themselves. Maps are provided as sets of pairs and can have arbitrary domain and range types. Primitive operations in SETL include set membership, union, intersection, and power set construction, among others. |
|  | SETL provides quantified Boolean expressions constructed using the universal and existential quantifiers of first-order predicate logic. |
|  | SETL provides several iterators to produce a variety of loops over aggregate data structures. |