Name: Krushnakumar Patle

Email: krishnapatle128@gmail.com

Batch: Data Engineering Batch-1

1. RDBMS
- Relational Database Management System (RDBMS) is a type of database management system that organizes data into tables, which consist of rows and columns.
- It is a type of database management system that organizes and manages data based on the principles of relational databases. In a relational database, data is stored in tables, and the relationships between the tables are defined by common fields, often referred to as keys.
- Common examples of RDBMS include MySQL, PostgreSQL, Oracle Database, Microsoft SQL Server, and SQLite. These systems are widely used in various applications, ranging from small-scale websites to large enterprise solutions, due to their ability to efficiently manage and organize structured data.

2. Data Warehousing
- Data warehousing is a concept and practice in the field of information technology that involves the collection, storage, and management of data from various sources within an organization.
- The goal of a data warehouse is to provide a centralized and unified repository of data for analysis and reporting.
- This helps organizations make informed decisions based on a comprehensive view of their data.

3. OLTP
- Online transaction processing (OLTP) is **a type of database system used for real-time execution of large volumes of database transactions by large numbers of people**.
- It is used for transaction-oriented applications, such as many operational systems, and is expected to respond to user requests and process them in real-time.
- OLTP systems are used for everyday transactions like online banking, online shopping, text messages, and account changes.
- These transactions are recorded and secured so that an enterprise can access the information anytime for accounting or reporting purposes.
- OLTP systems are capable of supporting transaction-oriented applications and are used for real-time execution of large volumes of database transactions by large numbers of people.

4. OLAP
- OLAP (Online Analytical Processing) is a computing method that enables users to extract and query data to analyze it from different points of view.
- It is used to perform multidimensional analysis at high speeds on large volumes of data from a data warehouse, data mart, or some other unified, centralized data store.
- OLAP databases are divided into one or more cubes, and each cube is organized and designed by a cube administrator to fit the way that you retrieve and analyze data.
- OLAP is part of the broader category of business intelligence, which also encompasses relational databases, report writing, and data mining.
- OLAP is designed to analyze multiple data dimensions at once, helping teams better understand the complex relationships in their data.

5. SQL and Its Features.
- SQL, or Structured Query Language, is a standardized programming language designed for managing and manipulating relational databases.
- It provides a set of commands for querying and modifying data stored in a relational database management system (RDBMS).
- SQL is used by developers, database administrators, and data analysts to interact with databases.
  - ❖ Features:
1. **Data Definition Language (DDL):**
   - DDL stands for Data Definition Language. It provides commands for defining relation schemas, deleting relations, and modifying relational schemas. All the structural operations performed on a database are controlled by this language.
   - SQL includes DDL commands (e.g., CREATE, ALTER, DROP) for defining and managing the structure of database objects such as tables, indexes, and views.
2. **Data Manipulation Language (DML):**
   - DML which stands for Data Manipulation Language is a computer programming language used for managing or manipulating data in a table.
   - SQL includes DML commands (e.g., INSERT, UPDATE, DELETE) for manipulating data within the tables.
3. **Declarative Language:**
   - SQL is a declarative language, meaning that users specify what results they want, and the database management system (DBMS) determines how to retrieve or manipulate the data. This is in contrast to procedural languages where users specify how to achieve a result.
4. **High-performance**
   - A large amount of data is retrieved quickly and efficiently. In addition, simple operations like inserting, deleting, and manipulating data can also be accomplished in a short amount of time. The reason SQL is so fast is that for a database product to be successful, it must return sets of data quickly when queried.
5. **Scalability**

- SQL database is vertically scalable, which means that you can increase the load on a single server by adding more RAM, SSDs, or CPUs.
- Due to the way data is stored (related tables vs unrelated collections), SQL databases generally support vertical scaling only – horizontal scaling is only possible for NoSQL databases.

6. **Security and authentication**
   - SQL Server includes several security-enabling features, including encrypted communication over SSL/TLS, Windows Data Protection API (DPAPI) to encrypt data at rest, authentication, and authorization.
   - Authentication means a process of identifying a user or a person based on their username and password. SQL Server authenticates its users by their credentials.
   - SQL Server supports two authentication modes, Windows authentication mode, and mixed mode.
   - Windows authentication is the default authentication method and is also known as integrated security because it is tightly integrated with Windows. Specific Windows user and group accounts are trusted to log in to SQL Server. Users who have already been authenticated do not need to provide additional credentials.
   - Mixed mode supports authentication both by Windows and by SQL Server. Usernames and passwords are maintained in SQL Server.

7. **Transaction Control Language:**
   - The propagation of a change in the database is called a transaction.
   - In DBMS, transactions are an important element, and to handle them, TCL(Transaction Control Language) is used, which includes commands such as commit, rollback, and savepoint.