

Introduction to DBMS and SQL

What is data?

- ❖ Data is a collection of raw, unorganized facts and details like text, observations, figures, symbols, and descriptions that represent people, objects, events, or other entities.
- ❖ In other words, data does not carry any specific purpose and has no significance by itself.

What is information?

- ❖ Information is data that has been processed, organized, or structured.
- ❖ It provide meaning or context, making data useful for decision-making or understanding.
- ❖ Data becomes information through the process of analysis, interpretation, and organization.

Types of Data:

1. Quantitative or Numerical Data: Data represented by numbers and used for quantitative analysis. Ex: Weight, volume, cost of an item.
2. Qualitative or Categorical Data: Data that represents categories or groups and used for qualitative analysis. Ex: Name, gender, hair color of a person.
3. Binary Data: Data consists of a sequence of binary digits (bits, 0 or 1) .Ex : images, videos, audios, plots or graphs

What is a Database?

- ❖ A database is a structured collection of data that is organized and stored in a systematic way to facilitate data management and retrieval.
- ❖ To make real use Data, we need Database management systems. (DBMS)

How is Data Stored in a Database?

- ❖ Data in a database is typically stored in tables, which consist of rows and columns.
- ❖ Each row represents a record or observation (in terms of Machine Learning) , and each column represents a field or attribute (or Feature in terms of ML).

What is DBMS?

- ❖ DBMS stands for Database Management System.
- ❖ It is a collection of interrelated data and a set of programs to access those data.
- ❖ The collection of data, usually referred to as the database, contains information relevant to an enterprise. The primary goal of a DBMS is to provide a way to store and retrieve database information that is both convenient and efficient.
- ❖ It is software that enables users to define, create, maintain, and control access to databases.

DBMS vs. File System:

File-processing systems has major disadvantages.

- i. Data Redundancy and inconsistency
- ii. Difficulty in accessing data
- iii. Data isolation
- iv. Integrity problems
- v. Atomicity problems
- vi. Concurrent-access anomalies
- vii. Security problems

Above 7 are also the Advantages of DBMS.

Properties of an Ideal DBMS:

1. Integrity
2. Availability
3. Security
4. Independent of Application
5. Concurrency

Functions of DBMS:

1. Data Management - Store, retrieve and modify data .**(CRUD Operation)**
2. Integrity - Maintain accuracy of data .
3. Concurrency - Simultaneous data access for multiple users.
4. Transaction - Modification to database must either be successful or must not happen at all.
5. Security - Access to authorized users only.
6. Utilities - Data import/export, user management, backup, logging.

Advantages and Disadvantages of DBMS:

- ❖ Advantages: Data consistency, centralized data management, data security, data sharing, and data integrity, reduce access time, multiple user interface, backup.
- ❖ Disadvantages: Cost of implementation and maintenance, complexity, and potential performance issues.

Types of DBMS with Real-life Examples:

1. Relational Databases - Also known as SQL databases, these databases use a relational model to organize data into tables with rows and columns. (Example: MySQL, PostgreSQL , Oracle)
2. No-SQL Databases - These databases are designed to handle large amounts of unstructured or semi-structured data, such as documents, images, or videos. (Example: MongoDB, Cassandra)
3. Column Databases - These databases store data in columns rather than rows, making them well- suited for data warehousing and analytical applications. (Example: Amazon Redshift, Google BigQuery)
4. Graph Databases - These databases are used to store and query graph-structured data, such as social network connections or recommendation systems. (Neo4j, Amazon Neptune)

5. Key-value databases - These databases store data as a collection of keys and values, making them well-suited for caching and simple data storage needs (Redis and Amazon DynamoDB)

Application of DBMS:

1. Data Storage: A database is used to store large amounts of structured data, making it easily accessible, searchable, and retrievable.
2. Data Analysis: A database can be used to perform complex data analysis, generate reports, and provide insights into the data.
3. Record Keeping: A database is often used to keep track of important records, such as financial transactions, customer information, and inventory levels.
4. Web Applications: Databases are an essential component of many web applications, providing dynamic content and user management.

What is SQL:

- ❖ SQL (Structured Query Language) is a programming language used for managing and manipulating data in relational databases.
- ❖ It allows you to insert, update, retrieve, and delete data in a database.
- ❖ It is widely used for data management in many applications, websites, and businesses.
- ❖ In simple terms, SQL is used to communicate with and control databases.

References:

1. CampusX (<https://youtube.com/@campusx-official?si=qjDOxjLwl2pZSxaq>)
2. TechTFQ (<https://youtube.com/@techtfq?si=zDhSv8M-6l4ywhDb>)