Data vs Information

- Data forms basis for information.
- Information is derived from data
-

Data Quality and Data Governance

Data quality can be assessed by

- Accuracy
- Relevance
- Timeliness
- Completeness
- Uniqueness(reduced redundancy)

Introduction to Database Management Systems (DBMS)

A DBMS is a software system designed to manage, store, retrieve, and manipulate data efficiently.
It serves as an interface between the database and end-users or application programs, ensuring that data is organized, accessible, and secure. DBMSs are essential for handling large volumes of data in a structured manner, enabling businesses and organizations to make data-driven decisions.

Key Concepts of DBMS

1. **Database**:
   - A database is a structured collection of data organized for easy access, management, and updating.
   - It can store various types of data, such as text, numbers, images, and more.

2. **Data:**
   - Raw facts or figures that can be processed to produce meaningful information.

3. **Schema**:
   - The blueprint or structure of the database, defining how data is organized and related.

- Includes tables, fields, relationships, and constraints.

4. **Tables**:
   - Data is stored in tables, which consist of rows (records) and columns (attributes).

5. **Query**:
   - A request to retrieve or manipulate data from the database, often written in a query language like SQL (Structured Query Language).

6. **Transaction**:
   - A sequence of operations performed on the database as a single logical unit, ensuring data integrity.

Functions of a DBMS

1. **Data Storage:**
   - Efficiently stores large amounts of data in a structured format.

2. **Data Retrieval:**
   - Allows users to retrieve data quickly using queries.

3. **Data Manipulation:**
   - Supports operations like insertion, deletion, and modification of data.

4. **Data Security:**
   - Protects data from unauthorized access and ensures privacy. Enforces user security and data privacy within a database

5. **Data Integrity:**
   - Ensures accuracy and consistency of data through constraints and rules.

6. **Concurrency Control: ()**
   - Manages simultaneous access to the database by multiple users using sophisticated algorithms.

7. **Backup and Recovery:**
   - Provides mechanisms to recover data in case of system failures.

8. **Data Dictionary management**
– Defines data elements and their relationships

Types of DBMS

1. **Relational DBMS (RDBMS)**:
   - Stores data in tables with rows and columns.
   - Uses SQL for querying and managing data.
   - Examples: **MySQL**, **PostgreSQL**, **Oracle**, **SQL Server**.

2. **NoSQL DBMS**:
   - Designed for unstructured or semi-structured data.
   - Types include document-based, key-value, column-family, and graph databases.
   - Examples: MongoDB, Cassandra, Redis, Neo4j.

Advantages of DBMS

1. **Data Sharing:**
   - Multiple users can access and share data simultaneously.

2. **Data Redundancy Control**:
   - Minimizes duplicate data, saving storage space.

3. **Data Consistency**:
   - Ensures data remains accurate and consistent across the database.

4. **Improved Security**:
   - Provides user authentication and access control mechanisms.

5. **Efficient Data Management**:
   - Simplifies data storage, retrieval, and manipulation.

6. **Scalability**:
   - Supports growing amounts of data and users.

Disadvantages of DBMS

1. **Complexity**:
   - Setting up and maintaining a DBMS can be complex and resource-intensive.

2. **Cost**:
   - Licensing, hardware, and maintenance costs can be high.

3. **Performance Overhead**:
   - Additional layers of abstraction can slow down performance.

4. **Vulnerability**:
   - Centralized databases are more susceptible to failures and attacks.

Types of Databases

- A single user DB - supports one user at a time
- Desktop - A single user database running on a personal computer
- Multi user - Supports multiple users at a time
- Workgroup - A multi user db that supports a small group of users(eg, a department)
- Enterprise DB - Multi user DB for a large group(eg, a campus database, facebook/instagram database)
- Centralised - supports data situated at a single site
- Distributed - Supports data situated across several sites.

The importance of Database design

- It defines the DBs expected use, Avoid redundancy
- Poorly designed database generates errors —> leads to bad decisions —> can lead to organisations failure(**REMEMBER!! Data vs Information**)

Data Redundancy(- keeping data in two or more places **)**

- Results in data inconsistency
- Data Anomalies(***inconsistencies***) develop when required changes in redundant data are not made successfully
  - Update anomalies
  - Deletion Anomalies
  - Insertion Anomalies

Data Modeling

- Simple representation of complex real world data structures

- In Data modelling we will have
  - Entities
  - Attributes
  - Relationship
  - Constraints

Business Rules

- Brief, precise, and unambiguous descriptions of a policies, procedures, or principles within a specific organisation
- Apply to any organisation that stores and uses data to generate information
- Description of operations that help to create and enforce actions within that organisation’s environment
- Must be rendered in writing
- Must be kept up to date
- Sometimes are external to the organisation
- Must be easy to understand and widely disseminated
- Describe characteristics of the data as viewed by the company

Translating business rules into data model components

Conclusion

A DBMS is a critical tool for managing data in today's data-driven world. It provides a systematic way to store, retrieve, and manipulate data while ensuring security, integrity, and efficiency. Whether it's a small business or a large enterprise, a DBMS is indispensable for handling data effectively and making informed decisions.