

## LEC-1: Introduction to DBMS

### 1. What is Data?

- a. Data is a collection of raw, unorganized facts and details like text, observations, figures, symbols, and descriptions of things etc.  
In other words, **data does not carry any specific purpose and has no significance by itself.**  
Moreover, data is measured in terms of bits and bytes – which are basic units of information in the context of computer storage and processing.
- b. Data can be recorded and doesn't have any meaning unless processed.

### 2. Types of Data

- a. **Quantitative**
  - i. Numerical form
  - ii. Weight, volume, cost of an item.
- b. **Qualitative**
  - i. Descriptive, but not numerical.
  - ii. Name, gender, hair color of a person.

### 3. What is Information?

- a. Info. Is **processed, organized, and structured data.**
- b. It provides **context of the data and enables decision making.**
- c. Processed data that make **sense** to us.
- d. Information is extracted from the data, by **analyzing and interpreting** pieces of data.
- e. E.g., you have data of all the people living in your locality, its Data, when you analyze and interpret the data and come to some conclusion that:
  - i. There are 100 senior citizens.
  - ii. The sex ratio is 1.1.
  - iii. Newborn babies are 100.These are information.

### 4. Data vs Information

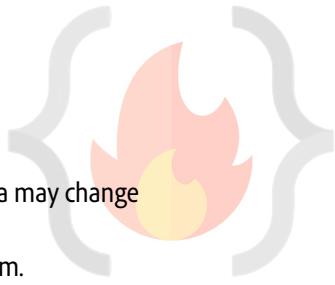
- a. Data is a collection of facts, while information puts those facts into context.
- b. While data is raw and unorganized, information is organized.
- c. Data points are individual and sometimes unrelated. Information maps out that data to provide a big-picture view of how it all fits together.
- d. Data, on its own, is meaningless. When it's analyzed and interpreted, it becomes meaningful information.
- e. Data does not depend on information; however, information depends on data.
- f. Data typically comes in the form of graphs, numbers, figures, or statistics. Information is typically presented through words, language, thoughts, and ideas.
- g. Data isn't sufficient for **decision-making**, but you can make decisions based on information.

### 5. What is Database?

- a. Database is an electronic place/system where data is stored in a way that it can be **easily accessed, managed, and updated.**
- b. To make real use Data, we need **Database management systems. (DBMS)**

### 6. What is DBMS?

- a. A database-management system (DBMS) 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.
- b. A DBMS is the database itself, along with all the software and functionality. It is used to perform different operations, like **addition, access, updating, and deletion** of the data.



- b. The overall design of the DB is called the DB **schema**.
  - c. Schema is **structural** description of data. Schema **doesn't change frequently**. Data may change frequently.
  - d. **DB schema** corresponds to the variable declarations (along with type) in a program.
  - e. We have 3 types of **Schemas: Physical, Logical**, several **view schemas** called subschemas.
  - f. Logical schema is most **important** in terms of its effect on application programs, as programmers construct apps by using logical schema.
  - g. **Physical data independence**, physical schema change should not affect logical schema/application programs.
3. **Data Models:**
- a. Provides a way to describe the **design** of a DB at **logical level**.
  - b. Underlying the structure of the DB is the Data Model; a collection of conceptual tools for describing **data, data relationships, data semantics & consistency constraints**.
  - c. E.g., ER model, Relational Model, **object-oriented** model, **object-relational** data model etc.
4. **Database Languages:**
- a. **Data definition language (DDL)** to specify the database schema.
  - b. **Data manipulation language (DML)** to express database queries and updates.
  - c. **Practically**, both language features are present in a single DB language, e.g., SQL language.
  - d. DDL
    - i. We specify consistency constraints, which must be checked, every time DB is updated.
  - e. DML
    - i. Data manipulation involves
      - 1. **Retrieval** of information stored in DB.
      - 2. **Insertion** of new information into DB.
      - 3. **Deletion** of information from the DB.
      - 4. **Updating** existing information stored in DB.
    - ii. **Query language**, a part of DML to specify statement requesting the retrieval of information.