Introduction to Database Systems

Role of Data in Computer

Data

A collection of facts (numbers, words, measurements, observations, etc) that has been translated into a form that computers can process

Types of Data

Structured data

- Information stored DB
- Strict format

E.g. SQL Databases. Spreadsheets

Semi-structured data

- Data may have certain structure but not all information collected has identical structure
- Some attributes may exist in some of the entities of a particular type but not in others
- E.g. csv, JSON files

Unstructured data

- Very limited indication of data type
  - E.g., a simple text document

Why Study Databases?

Databases are useful

- Many **computing applications** deal with **large amounts of information**
- Database systems give a **set of tools** for storing, searching and managing this information
- Organizations use it to store company related data
- Social media platforms use it to store user related data.

What is a Database?

- "A set of information held in a com

information that is infrequently accessed and not likely to be modified.

- "One or more large structured sets of persistent data, usually associated with software to update and query the data"

  Free On-Line Dictionary of Computing

- "A collection of data arranged for ease and speed of search and retrieval"

  Dictionary.com

Databases

- Library catalogues
- Medical records
- Bank accounts
- Stock control
- Product catalogues
- Telephone directories
- Train timetables
- Airline bookings
- Credit card details
- Student records
- Customer histories

and so on…

Database Systems

- A database system consists of
  - Data (the database)
  - Software
  - Hardware
  - Users
- We focus mainly on the software

- Database systems allow users to
  - Store
  - Update
  - Retrieve
  - Organise
  - Protect

  their data.

Database Management Systems

- A database is a collection of information
- A database management system (DBMS) is the software than controls that information

- Examples:
  - Oracle
  - DB2 (IBM)
  - MS SQL Server
  - MS Access
  - Ingres
  - PostgreSQL
  - MySQL

What the DBMS does

- Provides users with
  - Data definition language (DDL)
  - Data manipulation language (DML)
  - Data control language (DCL)
- Often these are all the same language

- DBMS provides
  - Persistence
  - Concurrency (same everywhere)
  - Integrity (truthfulness)
  - Security
  - Data independence
- Data Dictionary
  - Describes the database itself

Types of SQL Statements

DDL stands for **Data Definition Languages (DDL).** The SQL statements that are used to define the database structure. Any **CREATE**, **DROP** and **ALTER** commands are examples of DDL SQL.

DML stands for **Data Manipulation Language**. The SQL statements that are in the DML class are **INSERT**, **UPDATE** and **DELETE**.

DCL stands for **Data Control Language,** it includes commands such as **GRANT** and **REVOKE** which mainly deals with the rights, permissions and other controls of the database system.

Data Dictionary - Metadata

- The dictionary or catalog stores information about the database itself
- This is data about data or ‘metadata’
- Almost every aspect of the DBMS uses the dictionary
- The dictionary holds
  - Descriptions of database objects (tables, users, rules, views, indexes,…)
  - Information about who is using which data (locks)

File Based Systems

- File based systems
  - Data is stored in files
  - Each file has a specific format
  - Programs that use these files depend on knowledge about that format
- Problems:
  - No standards
  - Data duplication
  - Data dependence
  - No provision for security, recovery.

Relational Systems

- Information is stored as *tuples* or *records* in *relations* or *tables*
- There is a sound mathematical theory of relations
- Most modern DBMS are based on the relational model

Relational Model: Definition and Properties

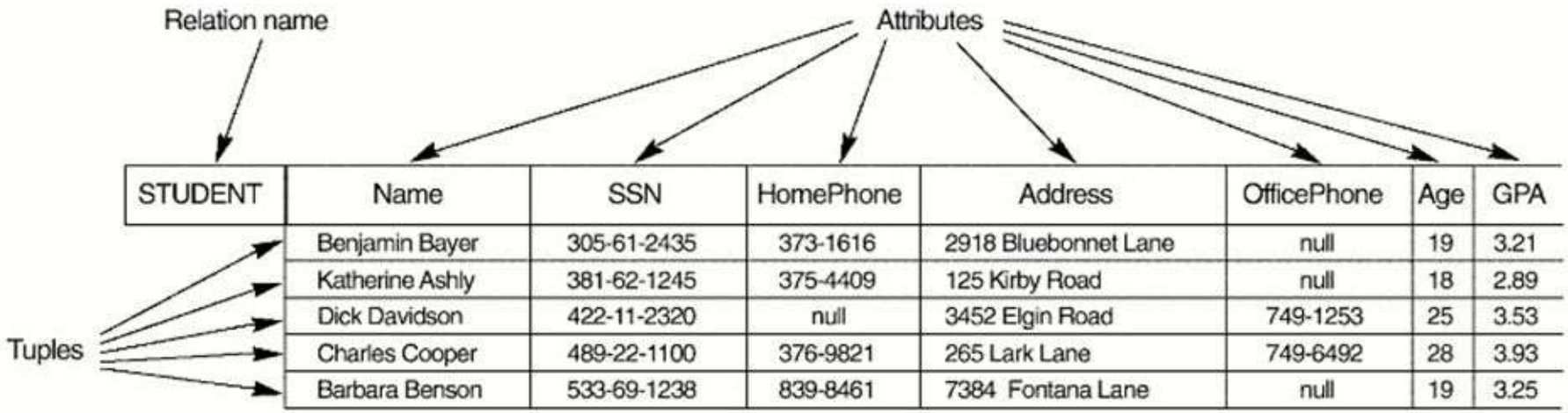
In original definition of Relational Model:

Tables are called **relations**;

Rows – **tuples**;

Column-names – **attributes**;

Data-type – **domain**.



STUDENT	Name	SSN	HomePhone	Address	OfficePhone	Age	GPA
	Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	null	19	3.21
	Katherine Ashly	381-62-1245	375-4409	125 Kirby Road	null	18	2.89
	Dick Davidson	422-11-2320	null	3452 Elgin Road	749-1253	25	3.53
	Charles Cooper	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
	Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	null	19	3.25