Preview

- Data is the lifeblood of business, and it comes in a huge variety of formats — everything from strictly formed relational databases to your last post on Facebook.
- Databases are involved in almost all facets and activities of our daily lives: from school to work, medical care, government, nonprofit organizations.
- Databases evolved from the need to manage large amounts of data in an organized and efficient manner

Data vs Information

- Data is a **collection of raw, unorganised 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.

Data vs Information

- Information is **processed, organised and structured data**. It provides context for data and enables decision making. For example, a single customer's sale at a restaurant is data – this becomes information when the business is able to identify the most popular or least popular dish.

Types of Data

- **Structured -** This is the data that can be decomposed into a structured format, such as rows and columns. Eg: Relational Data
- **Semi Structured -** Here, data has some form of organization, but not as much as Structured.
  Eg: XML Files
- **Unstructured -** The data has no format or organization. It is difficult to access and manipulate such data.
  Eg: Word files

What is DBMS?

- A database management system (or DBMS) is essentially nothing more than a computerized data-keeping system.
- Users of the system are given facilities to perform several kinds of operations on such a system for either manipulation of the data in the database or the management of the database structure itself.
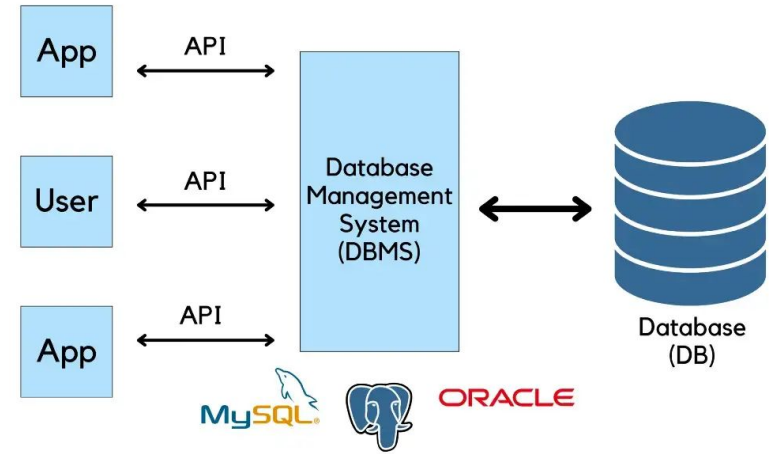- Database Management Systems (DBMSs) are categorized according to their data structures or types.
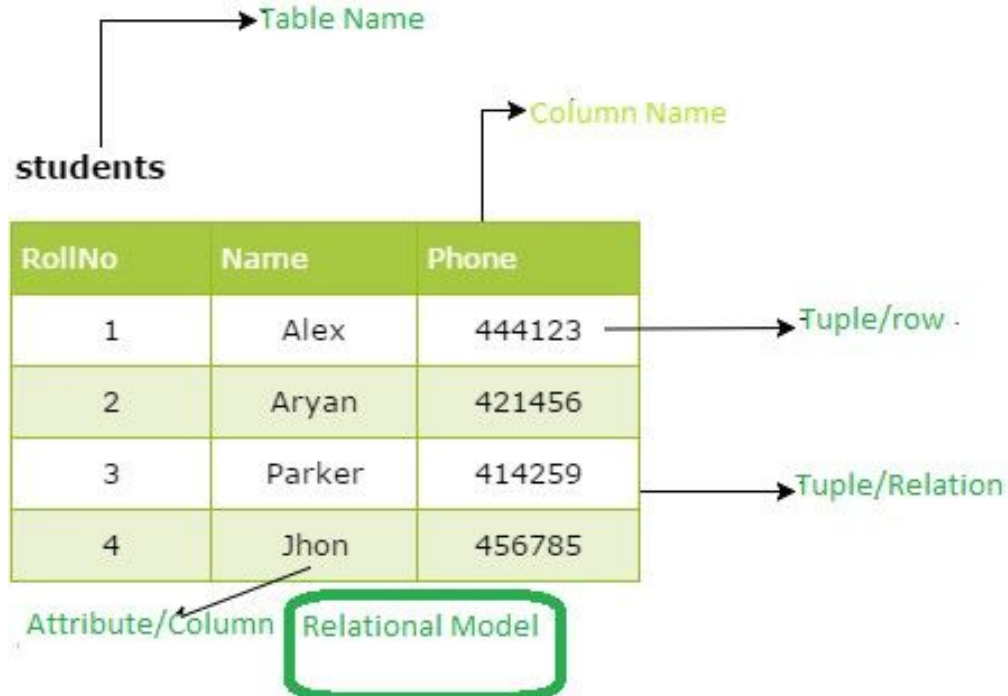
DBMS vs Flat files

DBMS vs. Flat File

DBMS	Flat File Management System
Multi-user access	It does not support multi-user access
Design to fulfill the need for small and large businesses	It is only limited to smaller DBMS system.
Remove redundancy and Integrity	Redundancy and Integrity issues
Expensive. But in the long term Total Cost of Ownership is cheap	It's cheaper
Easy to implement complicated transactions	No support for complicated transactions

Types of Databases

- Relational database
- Object-oriented database
- Hierarchical database
- Network database

RDBMS

Table Name

students

Column Name

RollNo	Name	Phone
1	Alex	444123
2	Aryan	421456
3	Parker	414259
4	Jhon	456785

Tuple/row

Tuple/Relation

Attribute/Column

Relational Model

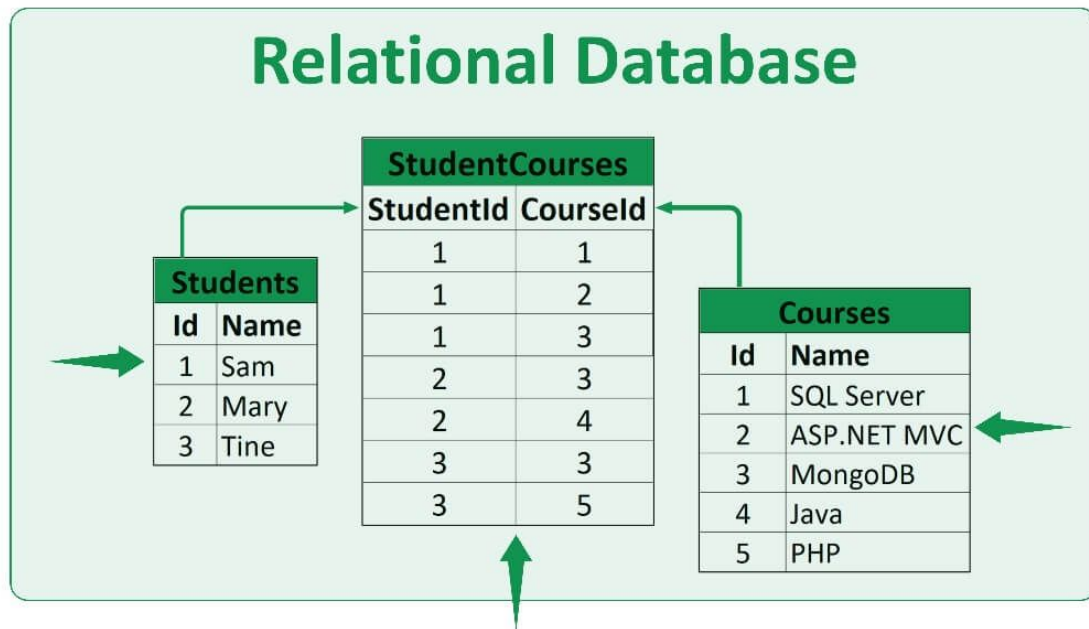https://www.learntek.org/blog/types-of-databases/

What is SQL ?

Structured Query Language

- manages relational databases and performs various operations on the data in them.
- maps different tables to get data from various places.
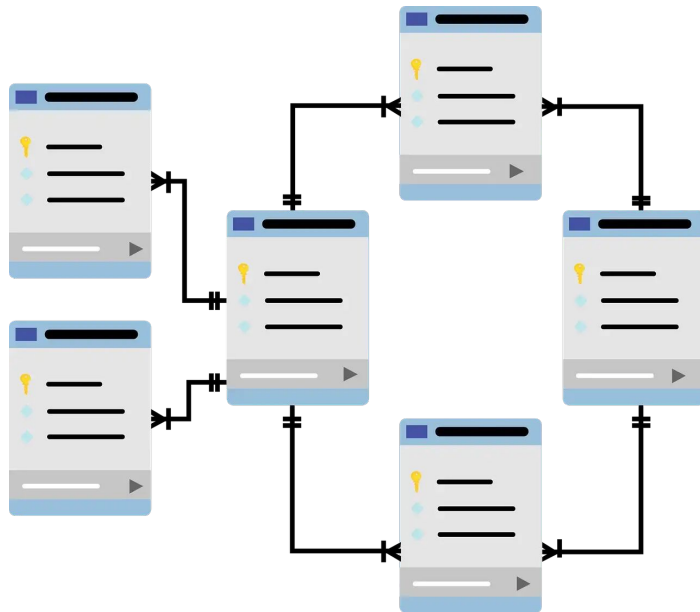
What is SQL ?

Relational and non relational databases

WOMEN WHO
CODE®
/data-science

Why use SQL for Data Science?

- Most commonly required skill

- Extract data needed for analysis with queries from databases

- Manage RDBMS

- Perform analytics and prepare data for preprocessing models.
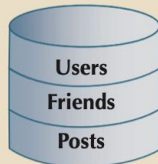
A Day In Susan's Life

See how many databases she interacts with each day

Before leaving for work, Susan checks her Facebook and Twitter accounts

Where is the data about the friends and groups stored?

Where are the "likes" stored and what would they be used for?

Users
Friends
Posts

On her lunch break, she picks up her prescription at the pharmacy

Where is the pharmacy inventory data stored?

What data about each product will be in the inventory data?

What data is kept about each customer and where is it stored?

Products
Sales
Customers

After work, Susan goes to the grocery store

Where is the product data stored?

Is the product quantity in stock updated at checkout?

Does she pay with a credit card?

Products
Sales
Customers

At night, she plans for a trip and buys airline tickets and hotel reservations online

Where does the online travel website get the airline and hotel data from?

What customer data would be kept by the website?

Where would the customer data be stored?

Flights
Hotels
Customers

Then she makes a few online purchases

Where are the product and stock data stored?

Where does the system get the data to generate product "recommendations" to the customer?

Where would credit card information be stored?

Products
Sales
Customers

Source: *Database Systems Design, Implementation, & Management | 13th Edition*
Carlos Coronel, Steven Morris

WOMEN WHO **CODE**®
/data-science

Series Overview

1. **Basics - Part 1** ✅

2. **Basics - Part 2 [January 20]**

   ○ high level CRUD and how SQL is used for these operations

   ○ basic structure of a SQL query - select, from, where, limit

   ○ conditional logic - >, <, equal, not equal, in, like, between

   ○ is null, and, or, not

Series Overview

3. **Intermediate - Part 1 [January 27]**

    - count, count distinct

    - sum, avg, min, max

    - group by, having, order by

    - case statements

4. **Intermediate - Part 2 [February 3]**

    - Joins - inner, left, right, outer, self join

Series Overview

5. **Advanced [February 10]**

   ○ Multiple joins

   ○ 2. Subqueries

   ○ 3. Window functions

   ○ 4. Set Operators

   ○ 5. CTE