**SQL (Clueless to Query Wizard)**

**(Beginner Level)**

1. Introduction

* What is SQL
* Why we learn SQL
* What is Database and Types
* SQL Commands
* SQL Components
* SQL Coding Style
* Setup your Environment

1. Query Data (Select)

* SELECT
* FROM
* WHERE
* ORDER BY
* GROUP BY
* HAVING
* DISTINCT
* TOP
* EXECUTION ORDER

1. Data Definition Language

* CREATE
* ALTER
* DROP
* Truncate
* Rename

1. Data Manipulation Language

* INSERT
* UPDATE
* DELETE

1. Filtering Data

* Comparison Operators
* Logical Operators
* BETWEEN
* IN
* LIKE

1. Combining Data

* Joining Data
* Basic Joins
* INNER Join
* LEFT Join
* RIGHT Join
* FULL Join
* Advanced Joins
* LEFT Anti Join
* RIGHT Anti Join
* FULL Anti Join
* Cross Join
* How to Choose Right Join
* How to Join Multiple tables
* SET Operators
* UNION
* UNION ALL
* EXCEPT
* INTERSECT

**(Intermediate Level)**

1. Row Level Functions

* String Function
* Number Functions
* Date & Time Functions
* Null Functions
* Case Statement

1. Aggregation & Analytical Functions

* Aggregate Functions
* Window Basics
* Window Aggregate Functions
* Window Ranking Functions
* Window Value Functions

**(Advanced Level)**

1. Advanced SQL Techniques

* Subqueries
* CTE
* Views
* CTAS & Temp tables
* Stored Procedures
* Triggers

1. Performance Optimization

* Indexes
* Partitions
* Performance Tips

1. AI & SQL
2. SQL Projects

* Data Warehousing Project
* Exploratory Data Analysis
* Advanced Data Analysis Project

**Person’s Data for Example:**

Everything generates data and it is everywhere your first name is data your mobile and everything inside the mobile is data, Car is as well generating a lot of data, Bank, Your Finance statements everything is data.

Now of course the question is where do we store our data personally?

**we store a lot of our data in like excels, spread sheets in a text file so you store a lot of your data in files.**

**Organization’s Data for Example:**

Now how about companies they have a lot of things that generate a lot of data that the products that they produce their customers as well generating lot of data and sales information and a lot of things so companies generate massive amount of data.

So now the big question is how they handle the data how they store it?

Of course, they cannot go and use like simple files. They need some bigger and stronger & Smarter and here where the database comes in so think about the database.

It’s like a container for storing data but instead of dumping files into folders the database organizes the data so it is easy to access to manage and to search so a database simply it is a container that stores data.

so now you might ask we are using database can’t we just use files like I do it personally well let me tell you.

why we use databases imaging that someone ask the following question go and find the Total Spending in your data? -

**I will go and open each and every file and combine the data then can give answer (but it is a messy process).**

**On Other hand we have Database if we want to ask a question it is very easy. All we have to do is talk to a database and ask a question, then the Database will give us answer with a result.**

**How can we talk to database or ask a database?**

We use **SQL:**

It is a language to talk with the Database. SQL Stands for Structured Query Language. With the help of SQL, we can ask DB to give the data and DB will return the required data. This way is very easy, simple & fast.

**companies prefer to DB instead of Excel sheets, Spread sheets or any other format.**

Reasons:

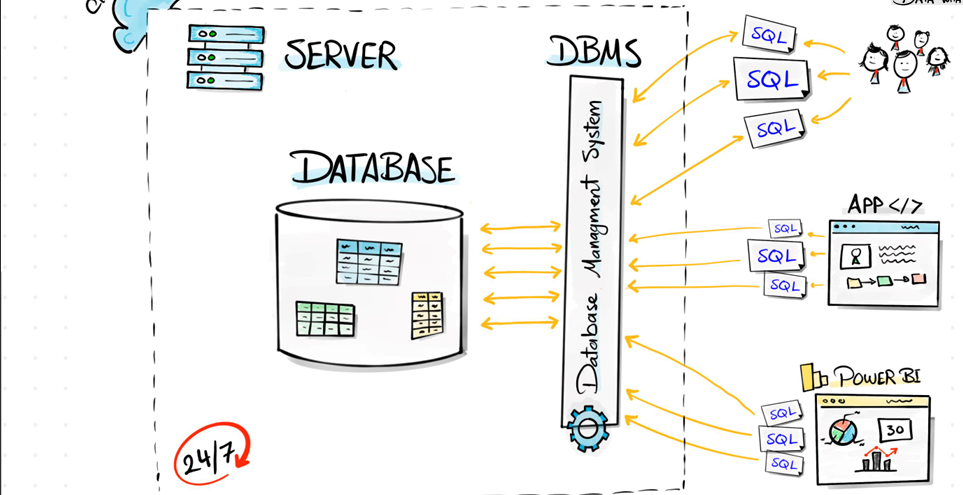
1. **DB can handle really huge amount of data**, sometimes we have like millions of data inside our DB but in the other side.
2. If you are storing your data **inside spreadsheets and you have like massive amount of data**. What can happen your **spread sheets going to just break**. They simply can’t handle the big data.
3. Another hand the reason why we use DB is that **it is just secure and it is safer to store important and critical data inside the DB** than storing it in Spread sheets and files.
4. **DB’s are secure and you can control who is accessing** what so it is just more **professional to store the data inside DB**.

**What we learned?**

* Most of the Companies stores their data inside a container called Database.
* For you in order to ask the questions and to talk your database you have to speak the language of SQL

|  |  |
| --- | --- |
| SQL Aspect | One-Line Explanation |
| What? | SQL is a standard language for managing and querying relational databases. |
| Why? | It enables efficient data storage, retrieval, and manipulation for applications and analytics. |
| When? | Used whenever structured data needs to be accessed, updated, or analyzed (e.g., web apps, reports). |
| Where? | Applied in databases like MySQL, PostgreSQL, and systems across industries (tech, finance, healthcare). |
| How? | By writing commands (e.g., SELECT, INSERT) to interact with database tables. |
| Who? | Used by developers, data analysts, DB admins, and business teams. |

What is DBMS & SQL Server?



Now, how it looks like usually in companies. So, we have our data inside **the DB** and then we have **multiple people with multiple roles that are just writing different SQLs** in order to talk to the data but now not only employees and people interact with the Database.

you could build a **website or an application that as well interacts with Database** by sending different SQLs and of course, depends on how many people are interacting with the application at the websites it might generate really massive number of SQLs that sends to the Database.

And not only that you might has as well **tools in order to do data visualizations** where you have like a dashboard or report may created by using **Power BI or Tableau and it is used by Stack holders and managers in order to make decisions** and as well those tools connected to the database and creating SQLs.

Now we can see we have lot of interactions with the database from **people, applications, tools a lot of things are generating SQLs and interacting with database**, but the database is just a container and storage rights.

we need **something software that manage all those requests** and that’s why we have something called **Database Management System (DBMS).** It is software that going to **manage all those requests to our Database and it going to make priority which SQL must be executes first** this software can as well manage the security whether the SQL is allowed to be executed in the first place.

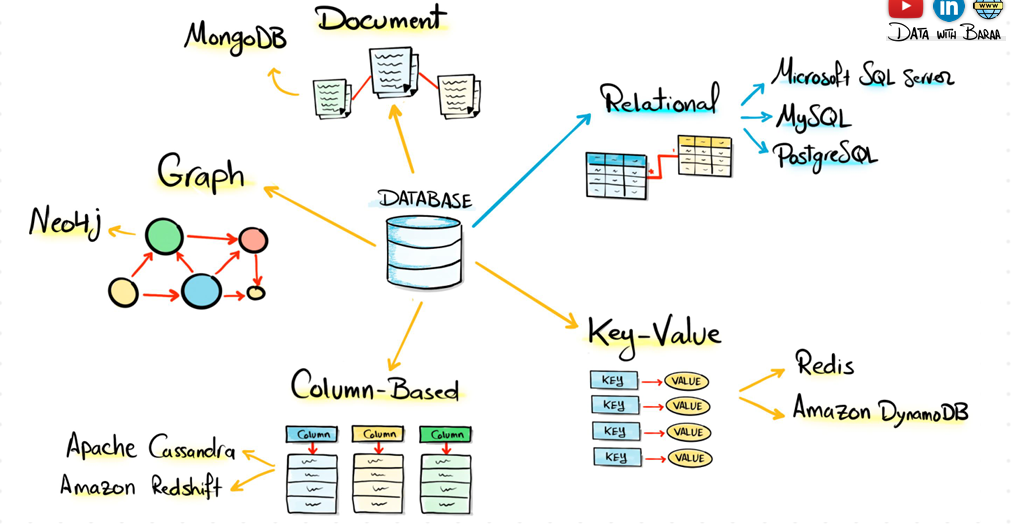
**Server it is like very powerful PC and as well it lives 24/7**. It is always available and here we can decide whether we’re going to have a server inside the company or we can use cloud services in order to run our database.

What we learn?

**Database: The Database is a container to store the data. SQL: The SQL it is a language in order to talk to the database. DBMS: It is the manager it manages the database. Server: it is the physical machine where the database lives.**

|  |  |
| --- | --- |
| Database Aspect | One-Line Explanation |
| What? | **A structured system to store, organize and manage data electronically.** |
| Why? | **To handle large amounts of data efficiently with security, consistency and fast access.** |
| When? | **Whenever data needs to be stored, retrieved, updated or analyzed (24/7 for most applications).** |
| Where? | **Used in all digital systems (apps, websites, enterprise software, IoT devices, etc.) across all industries.** |
| How? | **Through DBMS software (MySQL, MongoDB, etc.) using queries, APIs or direct connections.** |
| Who? | **Used by everyone indirectly; managed directly by developers, DBAs, data teams and organizations.** |

Types of Databases?



**Relational Databases:** Stores data in tables (like Excel sheets) with rows and columns. Tables can link to each other using keys. Best for: Structured data with clear relationships (e.g., customer orders, employee records). Example: MySQL, PostgreSQL, Microsoft SQL Server. Used in banking systems, e-commerce (Amazon), and school records.

**Key-Value Databases:** Stores data as simple key-value pairs (like a dictionary). Super-fast for lookups. Best for: Simple, fast data retrieval (e.g., caching, user sessions). Example: Redis, DynamoDB Used in shopping carts (Walmart), game leaderboards (Fortnite), and caching (Twitter feeds).

**Column-Base Databases:** Stores data in columns instead of rows, making it fast for analytics on large datasets. Best for: Big data, analytics (e.g., logs, financial reports). Example: Cassandra, Google Bigtable Used in stock market analysis (Bloomberg), IoT sensor data, and ad tracking (Facebook).

**Graph Base Databases:** Stores data as connected nodes (like a social network). Optimized for relationships. Best for: Complex relationships (e.g., social networks, fraud detection). Example: Neo4j, Amazon Neptune Used in LinkedIn connections, fraud detection (PayPal), and recommendation engines (Netflix).

**Document Databases:** Stores data in flexible JSON-like documents (not rigid tables). Best for: Unstructured or semi-structured data (e.g., user profiles, catalogs). Example: MongoDB, Firebase Used in content management (WordPress), mobile apps (Uber), and real-time data (weather apps).

**Quick Summary Table**

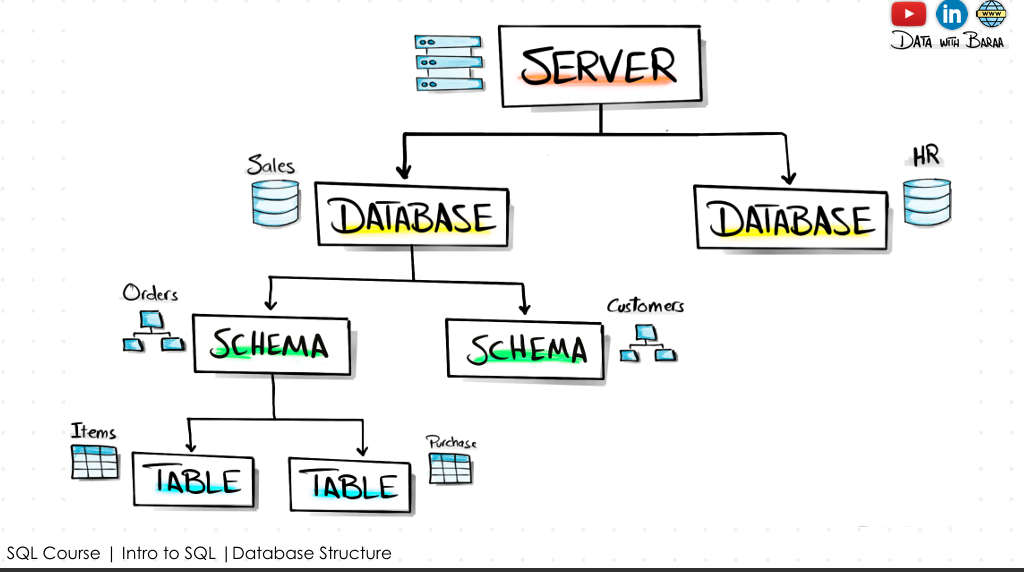
|  |  |  |  |
| --- | --- | --- | --- |
| Type | Stores Data As | Best For | Example Use Case |
| Relational | Tables (Rows/Cols) | Structured data (banking) | Amazon orders |
| Key-Value | Key → Value | Fast lookups (caching) | Redis session storage |
| Column-Based | Columns (not rows) | Big data analytics | Stock market data |
| Graph | Nodes & Connections | Social networks, fraud | LinkedIn connections |
| Document | JSON-like docs | Flexible data (apps, CMS) | MongoDB user profiles |

**Note: We are using Relational Databases. (Mostly used and powerful)**

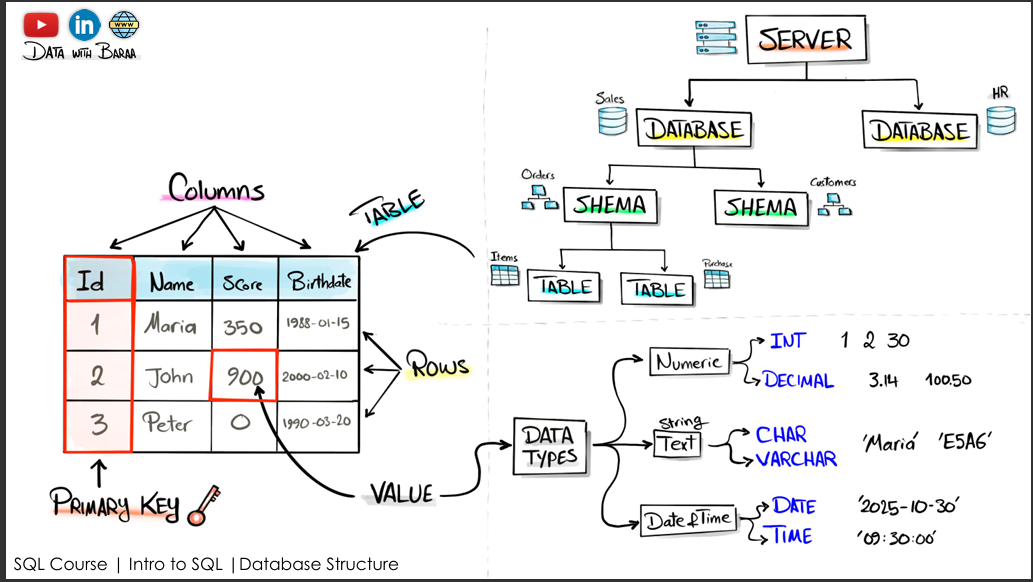
What is Database Structure?

**Server:** Databases are very structured and organized it has the following Hierarchy. The Starting point is the server as we all know it is a powerful PC and it is where the database lives and inside it we can have multiple databases (likes Sales, purchases, HR, Manufacturing) by that we understand that server can host multiple databases and as we learned a database is container of your data.

**Database:** Each database can have multiple Schemas.it is called like a category and we can call it a logical container that we can use it in order to group up related objects like let’s say we have hundreds of tables. so, you can split all the tables that has to do with the order in one schema and then another group of tables with the schema customers and so on it helps you to organize your tables and your objects in the database and now if you can go inside the schema you can have multiple tables like Tables.

**Table:** It is like a spreadsheet it organizes your data into the Columns.

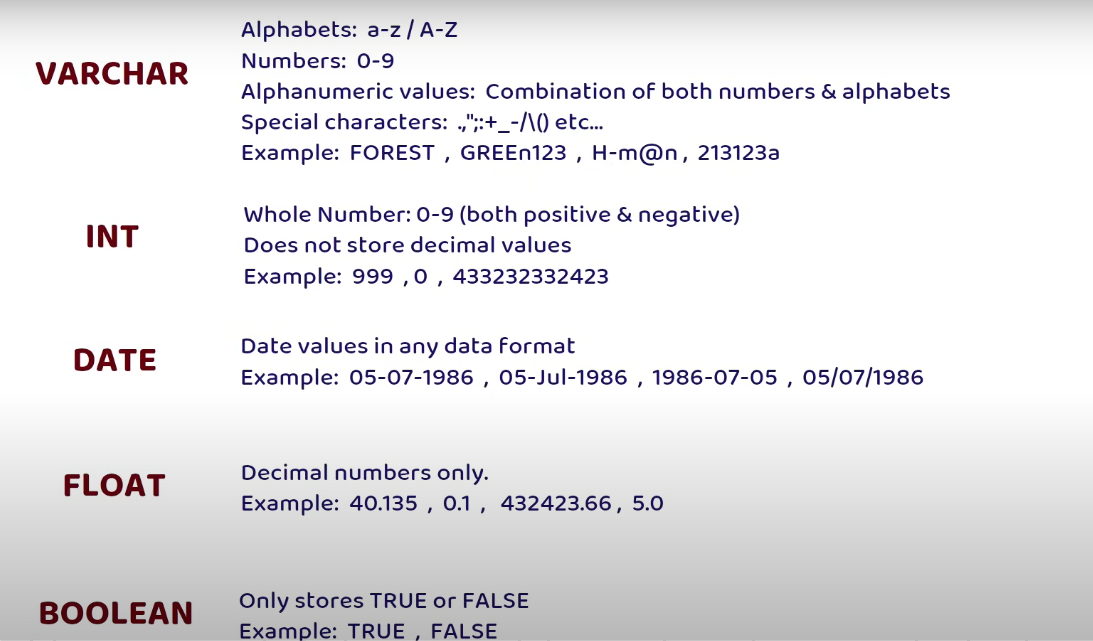
**Columns:** Columns define that you store it inside it. So, you can have one column about the CustomerID, another column about the Name, Score, Birthdate etc.., and each column is about one type of data and we call the columns as fields

**Rows:** We call them as records is where actually the data is stored. Like each record consists of CustomerID, Name, Score, Birthdate as above explained.

**Primary Key:** In each table there is like one very important column called Primary Key. It is always very important to have like one unique Identifier for each customer for each row and we use it for different purposes. In order to combine it with another table. In order to identify quickly one customer so it is Unique it’s like fingerprint and there is no two customers having the same ID.

**Data Types:** In a table each column stores the data with a specific data type. Data type means what kind of data we storing like (Numeric, String, Time, Float & Boolean).

* Numeric: **INT** for whole numbers (e.g., 1,3,5,10,42 etc.., **65000 times** larger than **SMALLINT**) Storing large counts (like website visitor counts) Primary keys that might grow beyond SMALLINT range Any value that could potentially exceed 32,767 When unsure about future growth needs. **DECIMAL** for exact decimals (e.g., 19.99,7.28,9.23) **SMALLINT** for smaller integers (e.g., 100. It stores much smaller than **INT** ) Storing ages (0-120) Quantity counts that won't exceed 32,767 Status codes or small numeric flags Any value known to stay below 32,767 to save space
* String: VARCHAR(n) for variable-length strings (e.g., 'Bill Gates') TEXT for long text (e.g., 'This is a long description...') CHAR(n) for fixed-length strings (e.g., 'US')
* Time (DATE / TIME) DATE for dates (e.g., '2023-12-31') TIME for times (e.g., '23:59:59') TIMESTAMP for date+time (e.g., '2023-12-31 23:59:59')
* FLOAT: FLOAT for approximate numbers (e.g., 3.14159) REAL for single-precision (e.g., 1.23456) DOUBLE PRECISION for double-precision (e.g., 1.23456789012345)
* BOOLEAN: BOOLEAN for true/false (e.g., TRUE)



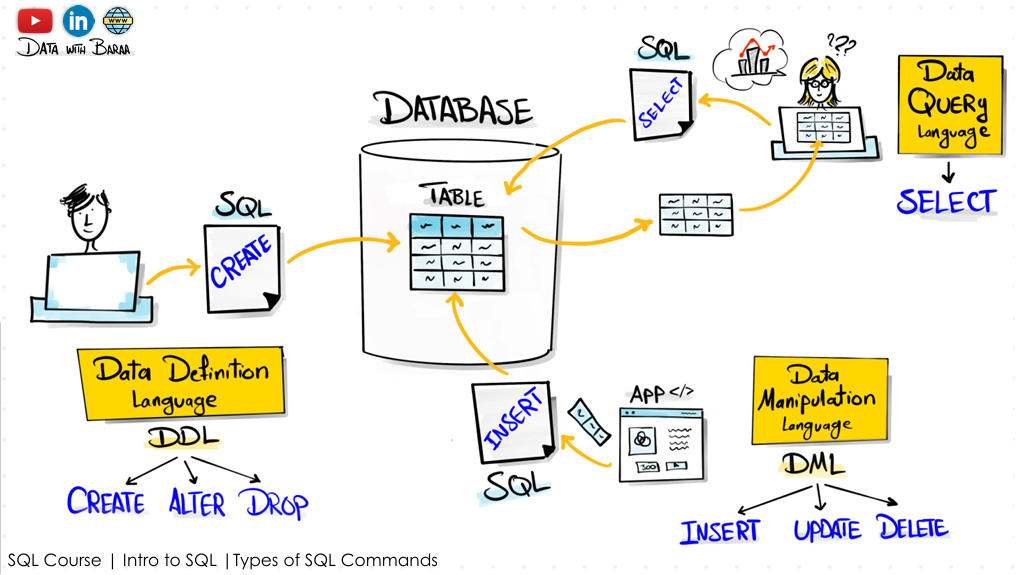
**COMMANDS: DDL (Data Definition Language):**

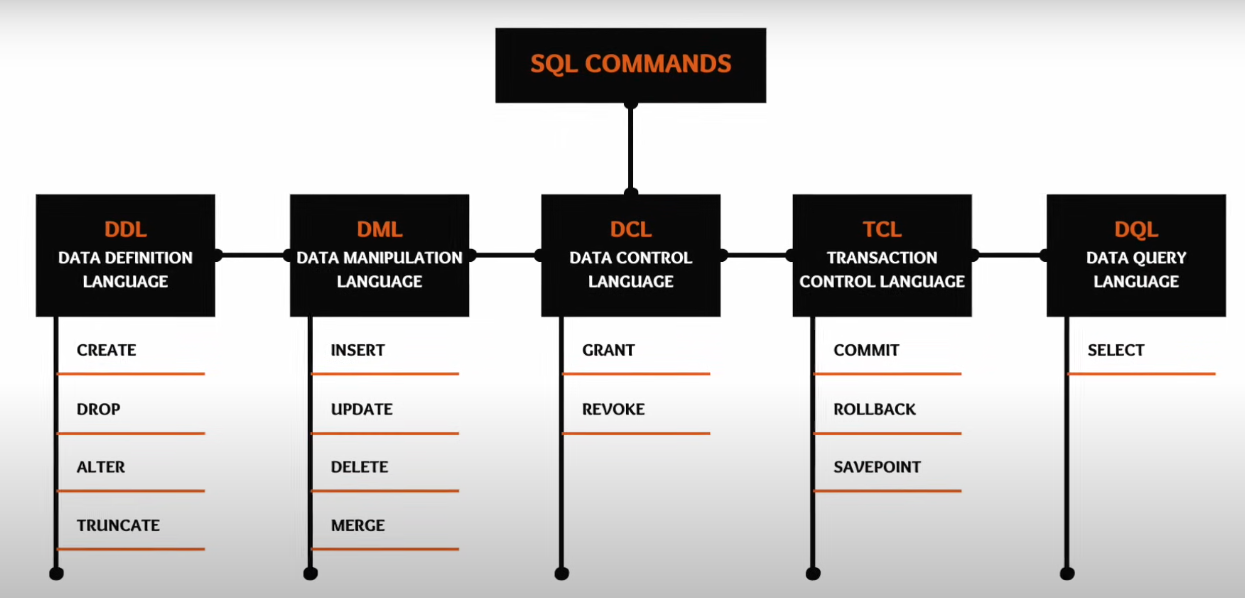
The Commands which we use to **create** something new. **Alter** to edit something that already exists and **Drop** order to delete something and if we use **Truncate** when want to delete the data but needs the Table structure.

**DML (Data Manipulation Language):**

The Commands which we use to **Insert** the new data into Database. **Update** in order to update an already existing data and **Delete** in order to go and delete the data from your table.

**DQL (Data Query Language):**

We use this command to Query the data by using **SELECT.**

Following are the Complete Commands which we have in SQL.

Why we need to learn SQL?

Top 3 Reasons to learn SQL:

1. **Talk to Data:** In order to talk to the data. We know most of the companies stores their data in databases. In order to work in Data related field, we have to know the SQL.
2. **Very High Demand:** If we go and check the Job Description of the Software Developer, Data Analyst, Data Engineer & Data Scientist we will find there they demand for SQL.
3. **Industry Standard:** If we check multiple modern data platforms and tools like Tableau, Power BI, Qlik View/Qlik Sense, KAFKA, Spark, Synapse will understand that there will be always a section where we have to enter SQL Codes. Because it is standard.

