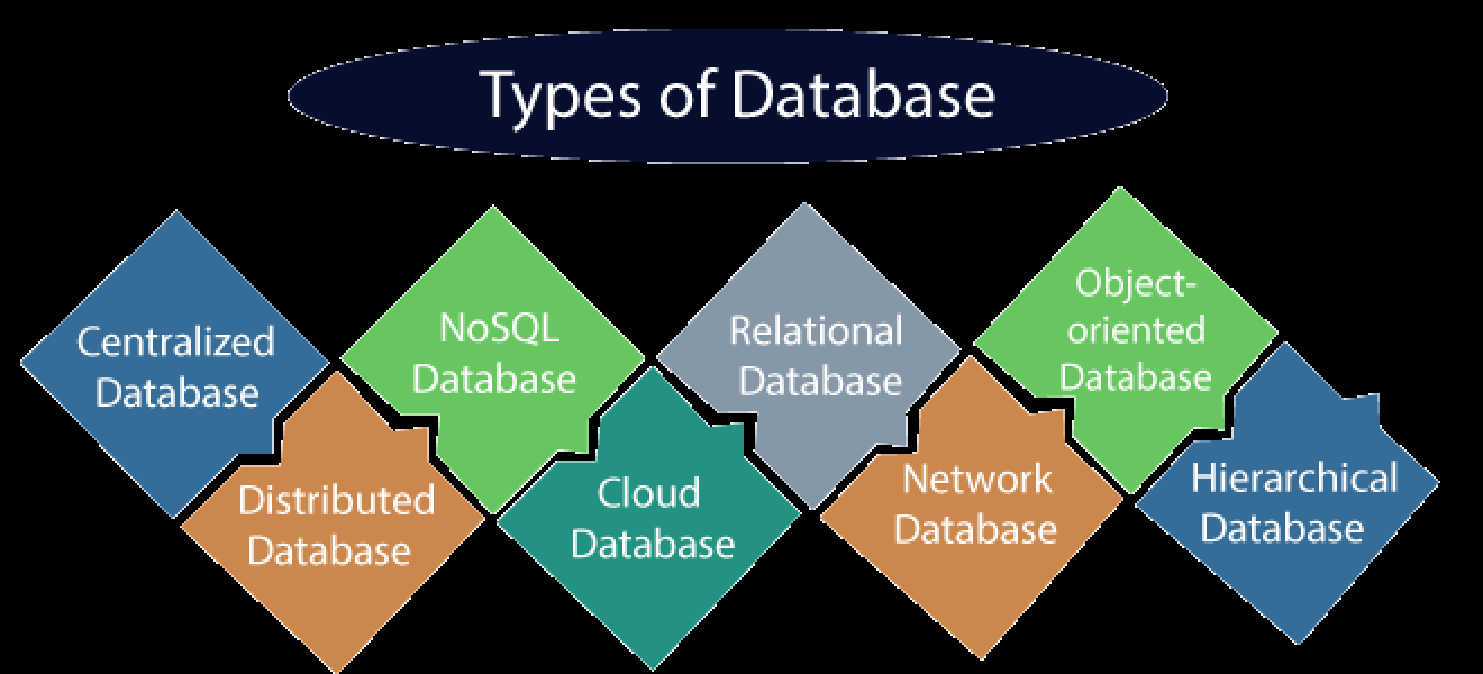
**Definitions:**

* The database is a collection of inter-related data which is used to retrieve, insert and delete the data efficiently. It is also used to organize the data in the form of a table, schema, views, and reports, etc.
* Database management system is a software which is used to manage the database.
* In dbms column/attribute and rows/records/tuple are the names.
* Data integrity is the overall accuracy, completeness, and consistency of data.
* Normalization is the process of organizing the data in the database by dividing the larger table into smaller and links them using relationships to reduce redundancy.

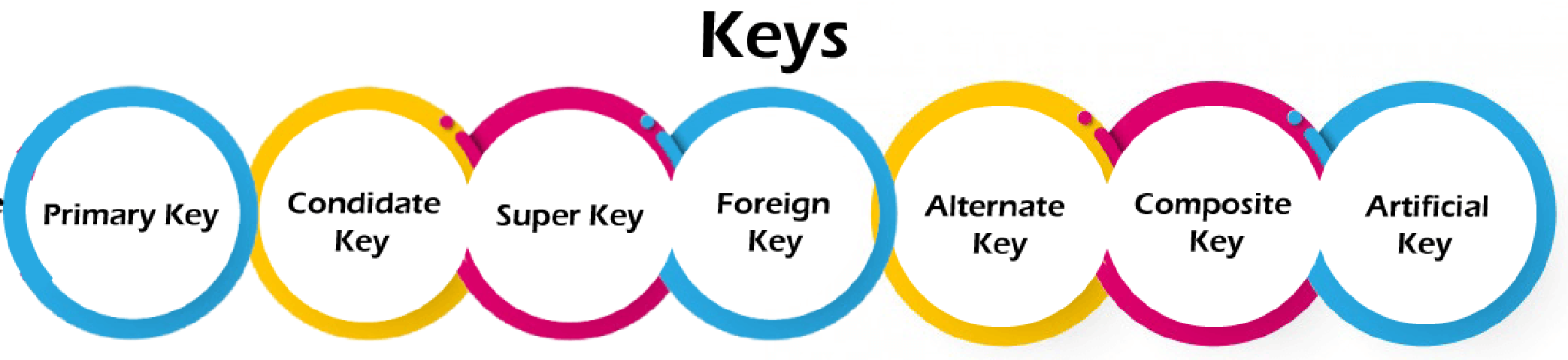


In 1 tier arch user can communicate with db server, in 2 tier arch user can communicate to db server via calls i.e jdbc ,in 3 tier user cannot communicate with db directly as there are serveral layers in between.

**ACID properties:**

* Atomicity: The term atomicity defines that the data remains atomic.
* Consistency: The word consistency means that the value should remain preserved always. which means if a change in the database is made, it should remain preserved always.
* Isolation: The term 'isolation' means separation. In DBMS, Isolation is the property of a database where no data should affect the other one and may occur concurrently.
* Durability: Durability ensures the permanency of something.

**Database Keys:**

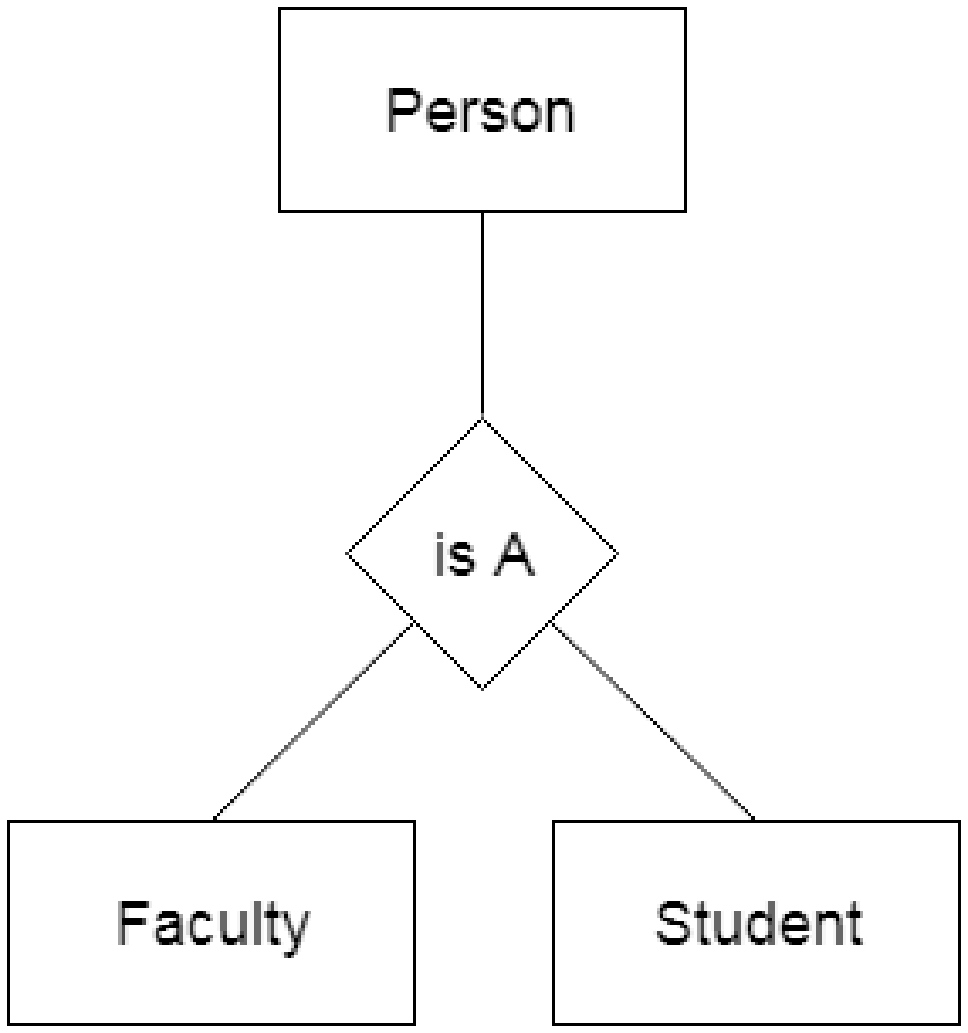


1. It is the first key used to identify one and only one instance of an entity uniquely. or Identify each and every record uniquely.
2. The attribute or set of attributes which can become primary key but not the primary key is called candidate keys, they all are the candidates for primary key
3. Super key is an attribute set that can uniquely identify a tuple. A super key is a superset of a candidate key.
4. Foreign keys are the column of the table used to point to the primary key of another table.
5. the total number of the alternate keys is the total number of candidate keys minus the primary key.
6. Whenever a primary key consists of more than one attribute, it is known as a composite key. This key is also known as Concatenated Key.
7. The key created using arbitrarily assigned data are known as artificial keys.

**Data Modeling:**

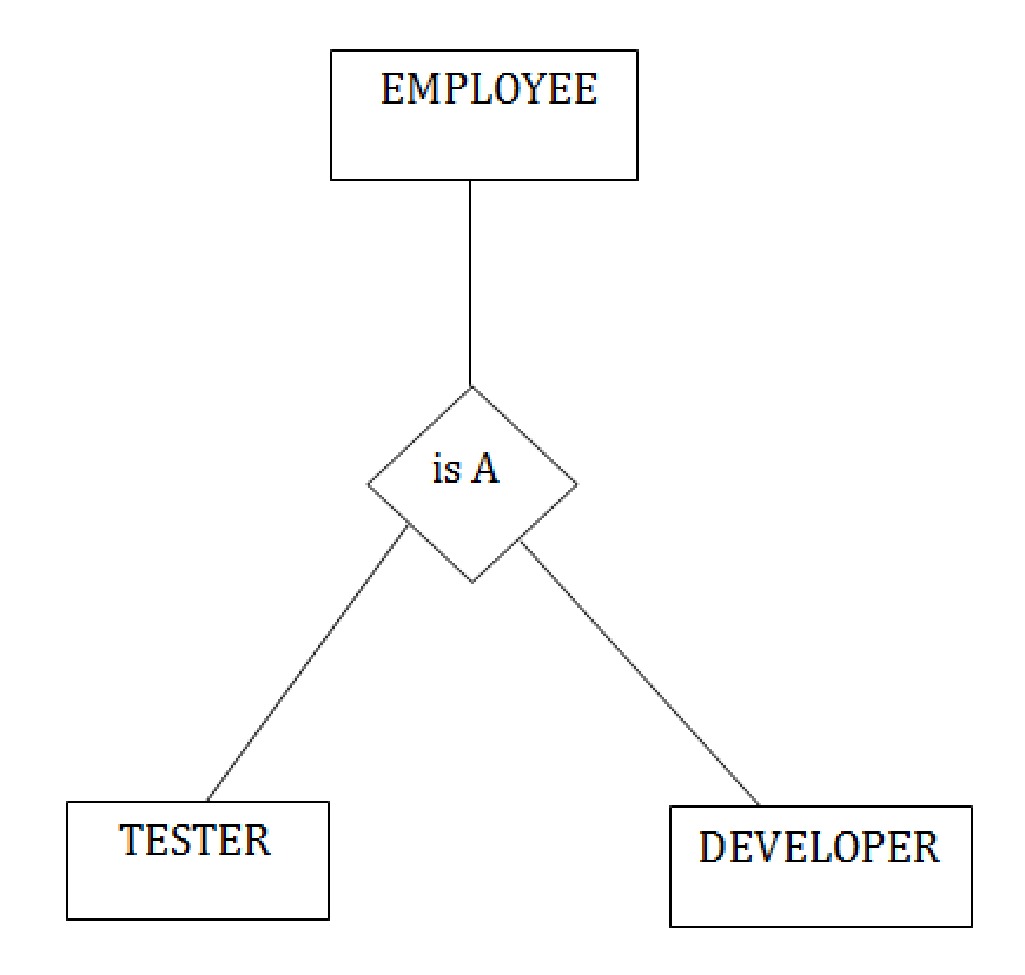
Generalization

bottom up

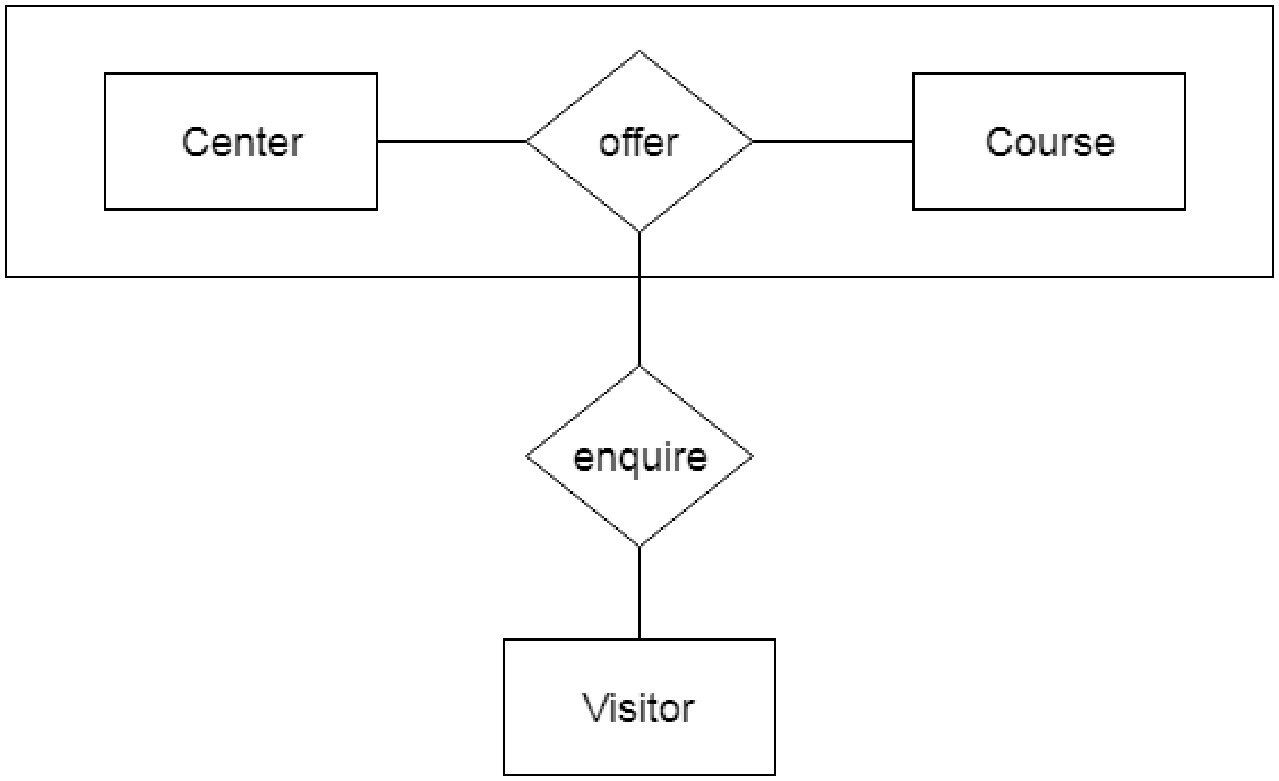


Specialization

top down

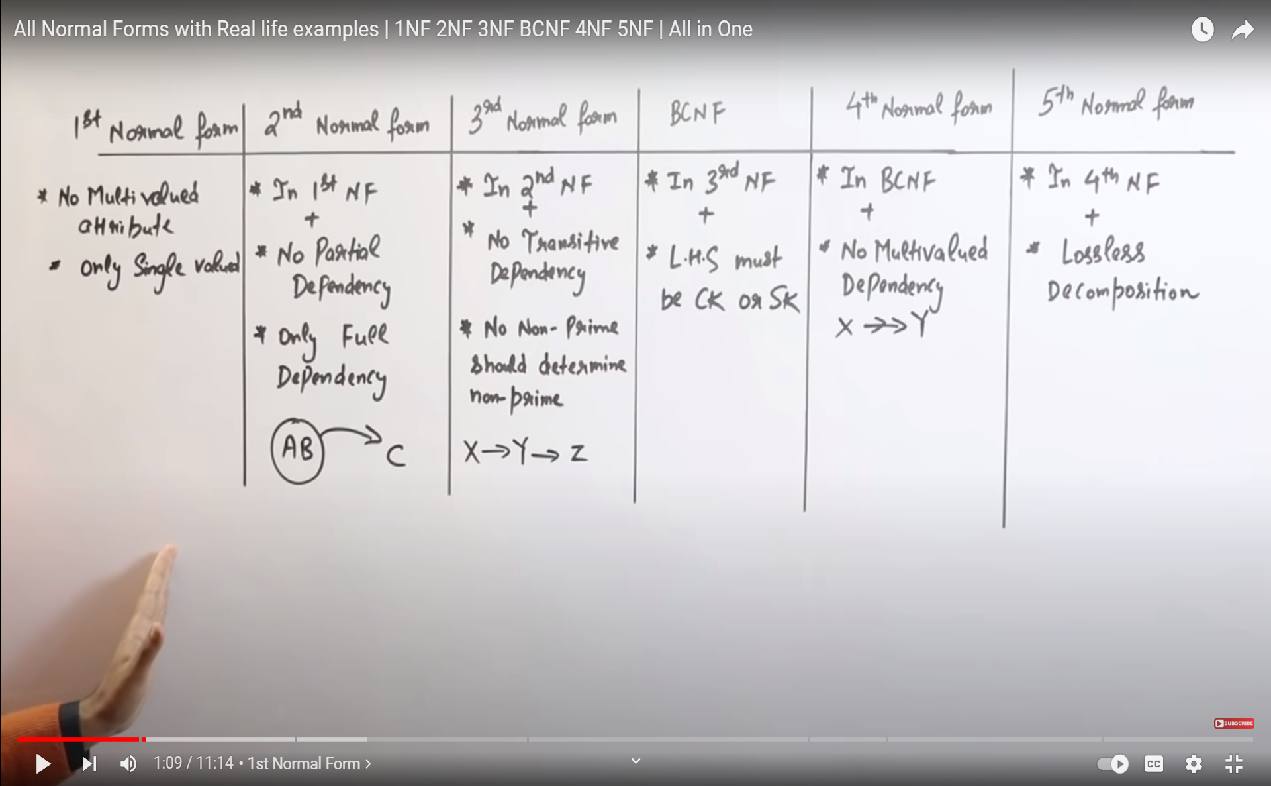


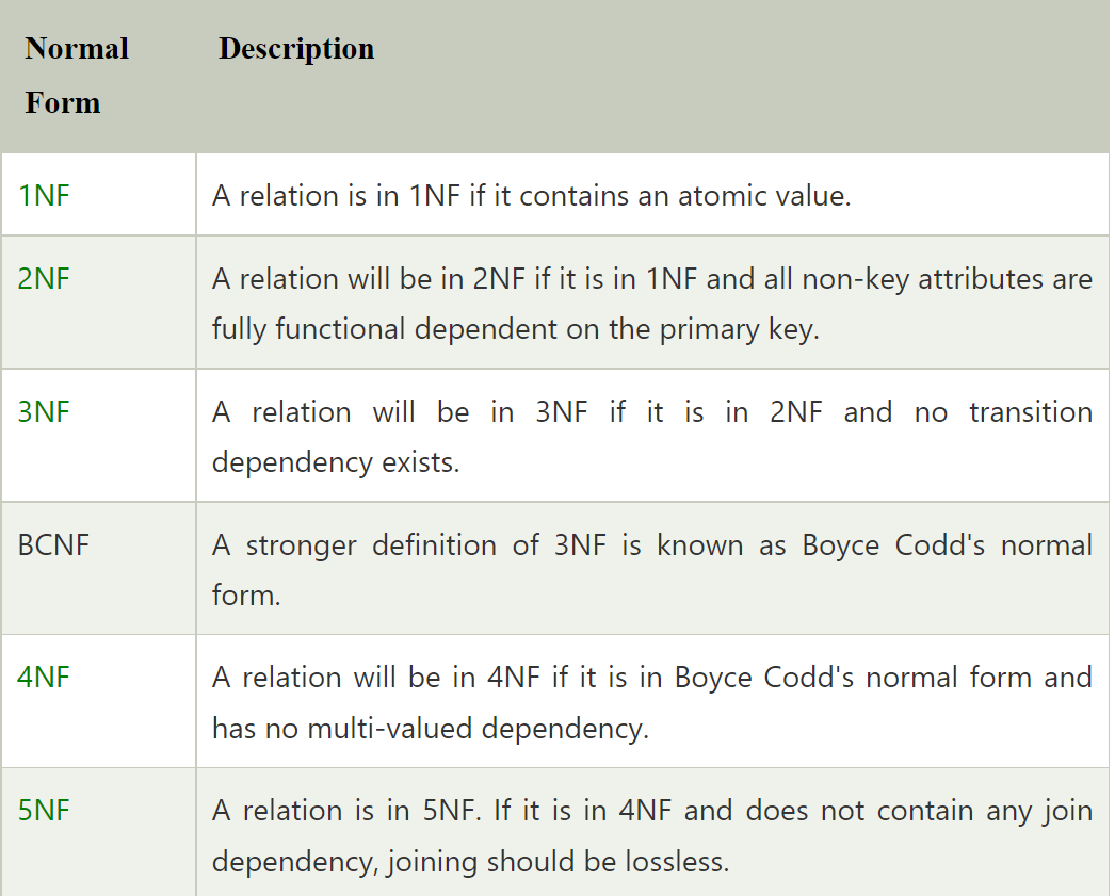
Aggregation



**Normalization:**

An attribute that is not part of any candidate key is known as non-prime attribute. An attribute that is a part of one of the candidate keys is known as prime attribute.

****



**where clause vs having clause:**

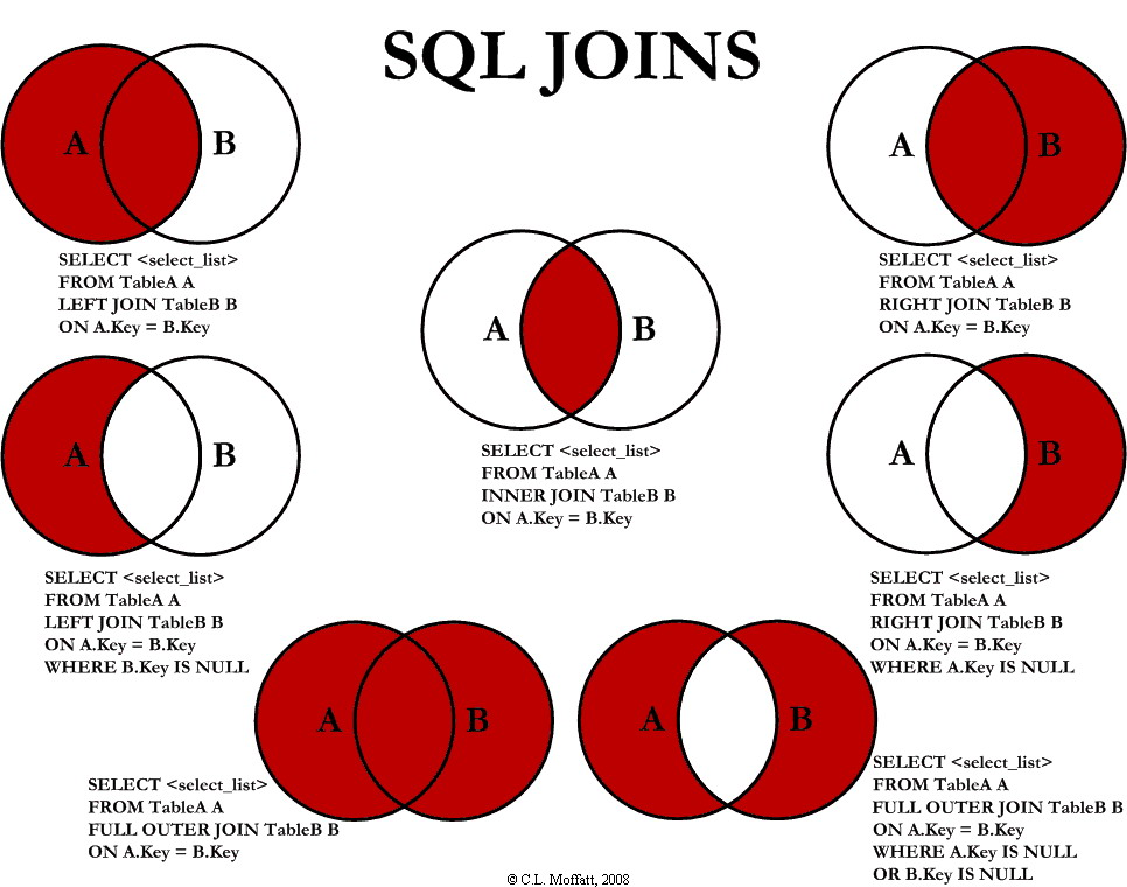
The main difference between them is that the WHERE clause is used to specify a condition for filtering records before any groupings are made, while the HAVING clause is used to specify a condition for filtering values from a group.

**group by vs order by:**

he GROUP BY clause is applicable when we want to use aggregate functions to more than one set of rows. The ORDER BY clause is applicable when we want to get the data obtained by a query in the sorting order

**Joins:**

In outer joins, all the related data from both the tables are combined correctly, plus all the remaining rows from one table. In full outer joins, all data are combined wherever possible.



**SQL commands:**

These SQL commands are mainly categorized into four categories as:

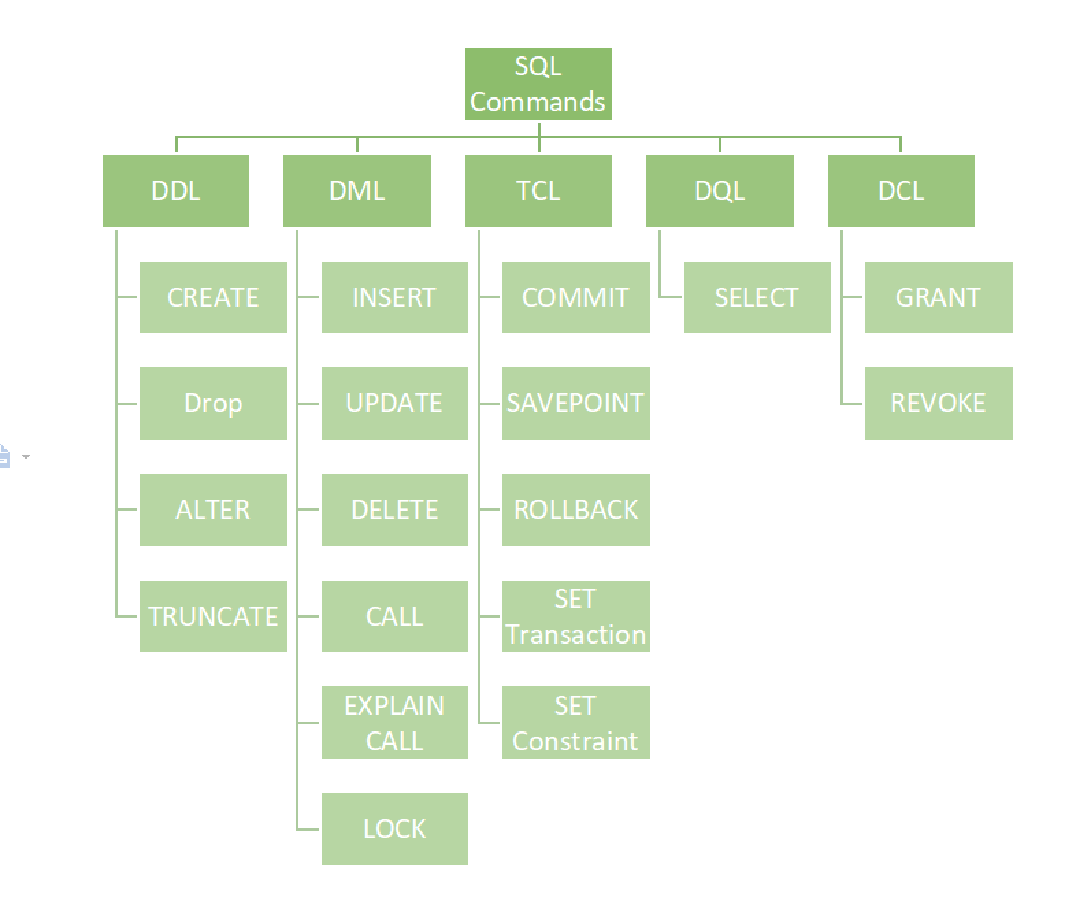
DDL – Data Definition Language

DQl – Data Query Language

DML – Data Manipulation Language

DCL – Data Control Language

\*\*TCL – Transaction Control Language



**Cursors:**

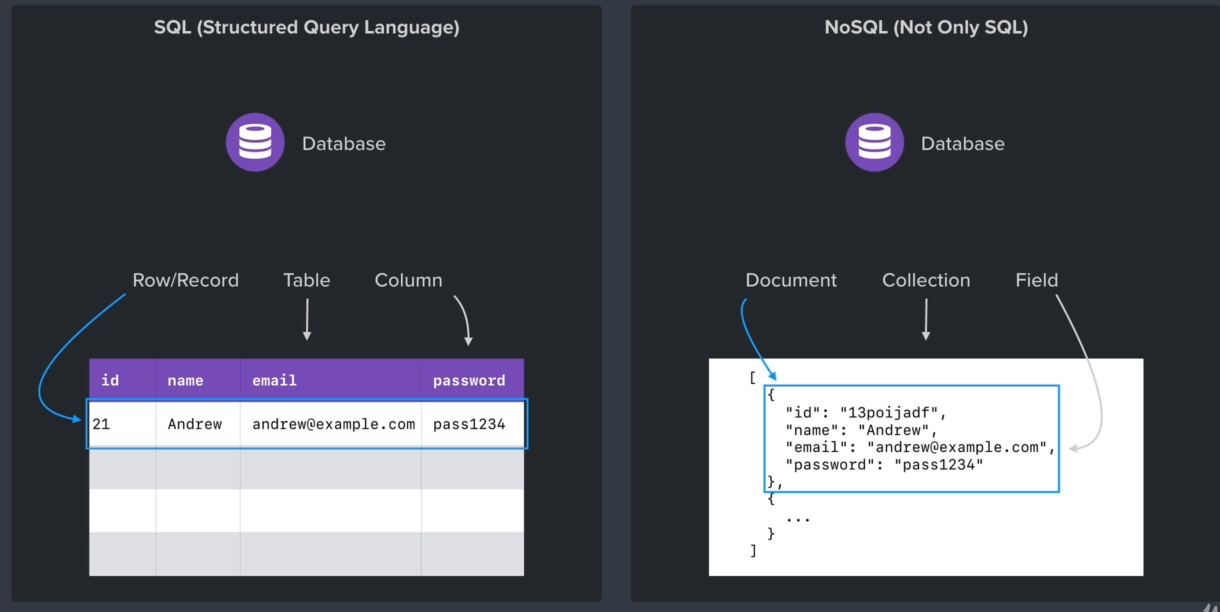
A CURSOR is a temporary area which holds the result set of a SELECT command returning multiple record(s). The record in the CURSOR can then be processed one by one.

**Purpose of stored Procedures:**

* If an action or series of actions is to be executed Purpose of a stored procedure
* repeatedly then we can create it as a procedure and compile it for future execution. It will be stored in the database in compiled form, thus providing improved performance
* Data manipulation can be implemented in the business layer of the application, but integration with other systems becomes difficult with this approach. Instead, by using stored procedures, integration becomes much easier

**SQL Queries:** [**https://www.javatpoint.com/sql-tutorial**](https://www.javatpoint.com/sql-tutorial)

**Joins:** [**https://afteracademy.com/blog/what-is-join-in-dbms-and-what-are-its-types/**](https://afteracademy.com/blog/what-is-join-in-dbms-and-what-are-its-types/)

****

Instead of using traditional id, mongodb uses global id (gid) that gives mongodb to scale well in the distributed systems aim is to avoid ID collision as there may be more then one database running at the same time

We can use mongodb library to generate ids of its own rather than server generated ids..