**NORMALIZATION**

**What Is Normalization?**

* Generally, in a table, we will have a lot of redundant information which is not required, so it is better to divide this complex table into multiple smaller tables which contains only unique information.
* The process of table design to minimize the data redundancy is called normalization.
* In other words, Normalization in SQL is the process of organizing data to avoid duplication and redundancy.
* It includes the creation of tables, establishing relationships between them, and defining rules for those relationships.
* The main aim of Normalization is to add, delete or modify field that can be made in a single table.

**Types of Normalization:**

**First Normal Form (1NF)**

* A relation schema is in 1NF, if and only if:
* All attributes in the relation are atomic (indivisible value)
* And there are no repeating elements or group of elements.
* This is the most basic level of normalization. In 1NF, each table cell should contain only a single value, and each column should have a unique name. The first normal form helps to eliminate duplicate data and simplify queries.

**Second Normal Form (2NF)**

* A relation is said to be in 2NF, if and only if:
* It is in 1NF.
* And creates & places data subsets in an individual table and defines relationship between tables using primary key.
* 2NF eliminates redundant data by requiring that each non-key attribute be dependent on the primary key. This means that each column should be directly related to the primary key, and not to other columns.

**Third Normal Form (3NF)**

* A relation is said to be in 3NF if and only if:
* It is in 2NF.
* And removes those columns which are not related through the primary key.
* 3NF builds on 2NF by requiring that all non-key attributes are independent of each other. This means that each column should be directly related to the primary key, and not to any other columns in the same table.

**Boyce-Codd Normal Form (BCNF)**

* BCNF is a stricter form of 3NF that ensures that each determinant in a table is a candidate key. In other words, BCNF ensures that each non-key attribute is dependent only on the candidate key.

**Fourth Normal Form (4NF)**

* 4NF is a further refinement of BCNF that ensures that a table does not contain any multi-valued dependencies.

**Fifth Normal Form (5NF)**

* 5NF is the highest level of normalization and involves decomposing a table into smaller tables to remove data redundancy and improve data integrity.

**Advantages of Normalization**

* **Reduced data redundancy:** Normalization helps to eliminate duplicate data in tables, reducing the amount of storage space needed and improving database efficiency.
* **Improved data consistency:** Normalization ensures that data is stored in a consistent and organized manner, reducing the risk of data inconsistencies and errors.
* **Simplified database design:** Normalization provides guidelines for organizing tables and data relationships, making it easier to design and maintain a database.
* **Improved query performance:** Normalized tables are typically easier to search and retrieve data from, resulting in faster query performance.
* **Easier database maintenance:** Normalization reduces the complexity of a database by breaking it down into smaller, more manageable tables, making it easier to add, modify, and delete data.