NORMALIZATION

BASIC CONCEPT OF NORMALIZATION:

Normalization➔ It is a technique of organizing the data into multiple related tables, to minimize **DATA REDUNDANCY**..

DATA REDUNDANCY➔It is nothing but repetition of similar data at multiple places.

- Repetition of data increases the size of database.
- Repletion of data hence needs extra space.
- Othe issues like:
    - Insertion Problem➔ To insert redundant data for every new row is a data insertion problem.
    - Deletion Problem➔ Loss of a related dataset when some other dataset is deleted.
    - Updation Problem

- Normalization is the technique to Minimizing the data redundancy.
- It follows, divide and Rule.
- It divide the data into separate independent logical entities and relating them using a common unique name.
- Eg:
    - We have a student Table with column names “roll_no., std_nm, branch, hod, ph.no.”.
    - Normalization divide the table into two separate table i.e. Student Table(roll_no, std_nm, branch) & Branch table(branch, hod, ph.no).
    - If we update, delete, insert any row in table then automatically it update, delete, insert a row in another table.

Types of Normalization:

- 1st Normal Form
- 2nd normal Form
- 3rd Normal Form
- BCNF

First Normal Form:

S-1:

- Scalable Table design which can be easily extended.
- If your table is not even in 1st Normal Form, It is considered poor DB Design.

❖ Every table in your database should at least follow the 1st Normal Form, always.

How to achieve 1st Normal Form:

There are 4 basic rules that a table should follow to be in 1st Normal Form.

RULE-1:

- Each Column should contain atomic/single values.

RULE-2:

- A column should contain values that are of the same type.
- Do not inter-mix different types of values in any column.
- Eg:

  We cannot insert name in a DOB Column.

RULE-3:

- Each Column should have a Unique Name.
- Same names leads to confusion at the time of data retrieval.

RULE-4:

- Order in which data is saved doesn't matter.
- Using SQL query, you can easily fetch data in any order from a table.

Second Normal Form:

Conditions:

- The table should be in 1st Normal Form.
- The table should not have any Partial Dependency.

Third Normal Form:

Conditions:

- It should be in 2nd Normal form.
- It should not have Transitive Dependency.

TRANSITIVE DEPENDENCY➔ If a column depends on a Non-Primary Column.

Boyce-Codd Normal Form:

- It is called as BCNF or 3.5 Normal Form.

Conditions:

- It should be in the 3rd Normal Form.
- For any dependency A→B, A should be a **SUPER KEY**.

Fourth Normal Form:

Conditions:

- It should be in BCNF.
- It should not have Multi-Valued Dependency.

Three Conditions for Multi-Valued Dependency:

- A→B, for a single value of A, more than one value of B exist.
- Table should have at least 3 columns.
- For a this table with A,B,C columns, B & C should be independent.

Firth Normal Form:

- It also called as PJNF(Project Joint Normal Form)

Conditions:

- It should be in 4^{th} Normal Form.
- It should not have Joint Dependency.