Database Design and Normalization - Easy Notes

1. Database Tables and Normalization

A database is a collection of tables that store data in an organized way. Each table has columns (fields) and rows (records).

However, if a database is not well designed, it can have data redundancy (repeating data) and inconsistencies. To avoid this, we use Normalization.

2. The Need for Normalization

Normalization helps to:

- Remove duplicate data.
- Reduce data redundancy.
- Improve data consistency.
- Make database updates easier.

3. The Normalization Process

Normalization is done in steps called Normal Forms (NF).

Functional Dependencies:

- A functional dependency means that one column (attribute) depends on another.

  Example: In a table of students, Student_ID -> Student_Name means that Student_ID uniquely determines Student_Name.

Minimal Set of Functional Dependencies:

- The smallest number of dependencies needed to describe the table structure.

4. Steps in Normalization

First Normal Form (1NF):

- Remove duplicate columns.
- Each column should have atomic (indivisible) values.
- Each row should have a unique identifier (Primary Key).

Second Normal Form (2NF):

- Meet 1NF requirements.
- Remove Partial Dependency (a non-key column should depend on the whole primary key, not just a part).

Third Normal Form (3NF):

- Meet 2NF requirements.
- Remove Transitive Dependency (non-key columns should depend only on the primary key).

5. Higher-Level Normal Forms

Boyce-Codd Normal Form (BCNF):

- Stronger than 3NF.
- Ensures there are no partial or transitive dependencies.

Fourth Normal Form (4NF):

- Removes multi-valued dependencies.

Fifth Normal Form (5NF):

- Deals with join dependencies.

6. Surrogate Key Considerations

A Surrogate Key is an artificial primary key (e.g., an auto-incremented ID) instead of using a natural key like a name or phone number.

7. Normalization and Database Design

- Helps in designing efficient, consistent, and error-free databases.
- Reduces storage waste and improves performance.
- Ensures data integrity.

Conclusion:

Normalization is a step-by-step process that helps organize database tables to remove redundancy and improve efficiency. It starts from 1NF (basic structure) to higher forms (BCNF, 4NF, and 5NF), ensuring a well-structured database.