**NORMALIZATION AND DENORMALIZATION**

**Normalization** Normalization is the process of organizing data in a database. It includes creating tables and establishing relationships between those tables according to rules designed both to protect the data and to make the database more flexible by eliminating redundancy and inconsistent dependency.

**Objectives**

* **Eliminate duplicate data**: Avoid problems that arise from updating redundant data
* **Ensure data dependencies make sense**: Establish connections between tables using primary and foreign keys
* **Facilitate easier querying**: Make it easier to query the database
* **Improve data integrity**: Ensure data is accurate, consistent, and reliable
* **Reduce redundancy**: Avoid storing redundant data
* **Make the database more flexible**: Eliminate inconsistent dependency

Three types of norms in normalization

* 1NF
* 2NF
* 3NF

1NF - Ensure that each column contains individual values .

Eliminate repeating groups in columns.

2NF - Removes partial dependencies.

3NF - Removes transitive dependencies.

Advantages

* Eliminates redundancy and saves storage.
* Improves data integrity and consistency.
* Simplifies maintenance and update.

Disadvantages

* Requires multiple joins to fetch related data , reducing , performance .
* More complex queries and shower retrieval.

**Denormalization**  Denormalization is a database technique where redundant data is intentionally added to a previously normalized database to improve read performance by reducing the number of joins needed to retrieve data, essentially combining data from multiple tables into a single table for faster querying, but at the cost of increased data redundancy and potential complexities in data maintenance and updates.

Example: Instead of maintaining separate tables , we combine them.

Order table: Before

|  |  |  |
| --- | --- | --- |
| Order id | Product id | Quantity |
| 1 | 101 | 2 |
| 2 | 102 | 1 |

Product table:

|  |  |
| --- | --- |
| Product id | Product name |
| 101 | Laptop |
| 102 | Phone |

After

|  |  |  |  |
| --- | --- | --- | --- |
| Order id | Product  id | Product  Name | Quantity |
| 1 | 101 | Laptop | 2 |
| 2 | 102 | Phone | 1 |

Advantages

* Faster read performance by reducing joins.
* Simpler queries and faster reporting.

Disadvantages

* Increases redundancy and storage requirements.
* Can lead to data inconsistency.
* More difficult to maintain updates and deletions.