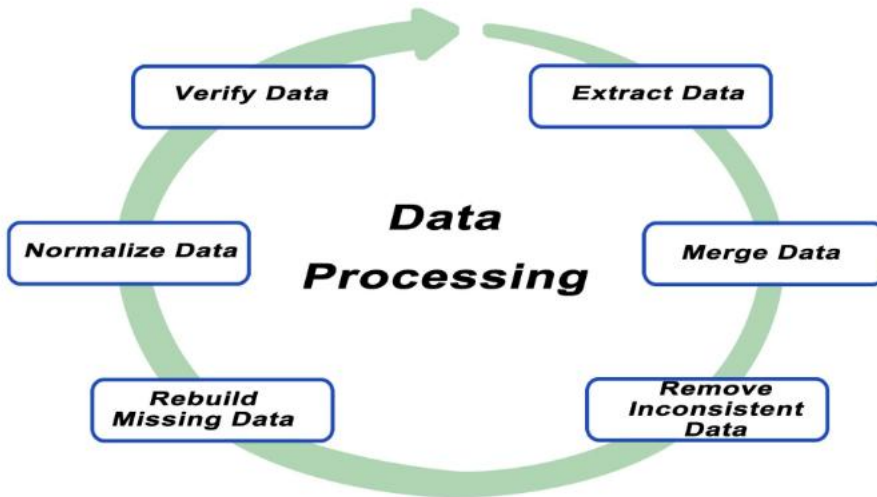


DATABASE NORMALIZATION

Database Normalization is the process of organising data in a database to avoid **data redundancy** (repeating the same info) and **data anomalies** (errors that happen when you update, delete, or insert data). At the same time, it is useful to improve database efficiency. It's important to create a scalable, reliable and performant databases.



AIMS OF NORMALIZATION

1. Eliminate redundancy: breaking data into logical tables and removing duplicate information.
2. Confirming data integrity: establishing clear relationships and dependencies between entities that maintains data accurate and reliable.
3. Prevent anomalies
4. Optimize Query Performance

Normalization reduces data redundancy and anomalies by dividing large, complex tables into smaller, related tables. To do so it uses a set of rules known as **normal forms**, each having distinct requirements that narrow down how the database would be designed, or in other words, each addresses specific types of redundancy and dependency issues. The normal forms have a hierarchy to respect. The normal forms get increasingly stricter.

First Normal Form (1NF)

It is the initial stage of the process.

- Eliminates duplicate columns and ensure every column contains only atomic, indivisible values of a single data type
- Makes sure that every row is unique
- Makes sure that there are no repeating groups in a row

Second Normal Form (2NF)

It meets all the 1NF requirements and ensure that all non-primary key attributes are fully functionally dependent on the primary key

Third Normal Form (3NF)

It meets all the 2NF requirements and ensures that all the non-key attributes do not depend on other non-key attributes

PRACTISE

1. Give the definition of the following terms

Data Redundancy:

Data Anomalies:

“Atomic” value:

Normal Form:

NF Requirements:

2. Complete the sentences with the appropriate words

1. Normalization aims to improve database e_____ while creating systems that are scalable, reliable, and p_____
2. The process involves dividing large, complex tables into s_____, r_____ t_____
3. Normal forms follow a specific h_____ and become s_____ as you progress.
4. To reach 1NF you must ensure that every row is u_____
5. 2NF requires that all non-primary keys attributes are full f_____ d_____ on the primary key.

3. Transforms the following sentences using the right technical expression

Informal: We need to make sure that data doesn't repeat itself

Technical:

Informal: We want to stop mistakes from happening when we change things

Technical:

Informal: Every box in the table should only have one piece of information

Technical:

Informal: We need to split the big, messy tables into smaller ones that are connected to each other

Technical:

Informal: You have to follow the rules in order because each step gets harder than the one before

Technical: