ONE PAGE SUMMARY

IT & Software › Database › Normalization

Introduction to Normalization

Mastering Normalization for Efficient Database Design

DEFINITION

> Normalization is the process of organizing data in a database to reduce redundancy and improve data integrity. It involves dividing large tables into smaller, related ones and defining relationships between them.

KEY CONCEPTS

First Normal Form (1NF)

Ensures that all columns contain atomic values and each entry is unique.

Second Normal Form (2NF)

Eliminates partial dependencies; all non-key attributes must depend on the entire primary key.

Third Normal Form (3NF)

Removes transitive dependencies; non-key attributes must depend only on the primary key.

Boyce-Codd Normal Form (BCNF)

A stronger version of 3NF that addresses certain types of anomalies not handled by 3NF.

Denormalization

The process of combining tables to improve read performance at the cost of redundancy.

EXAMPLES

→ Splitting a customer table into separate tables for customer details and orders.

→ Creating a product table that links to a suppliers table to avoid duplicate supplier data.

→ Using a junction table to manage many-to-many relationships between students and courses.

MEMORY TIPS

★ Remember 1NF as 'Atomic' for single values.

★ Think of 2NF as 'Whole' for complete dependencies.

★ 3NF can be recalled as 'Direct' since attributes depend directly on the primary key.

COMMON MISTAKES

✗ Confusing 1NF with having unique rows only, ignoring atomicity.

✗ Overlooking partial dependencies when moving to 2NF.

✗ Failing to recognize transitive dependencies in 3NF.

QUICK RECAP

Normalization is essential for structuring databases efficiently. It minimizes redundancy and enhances data integrity through various normal forms. Always check for dependencies to ensure your database is well-normalized.