Data Normalization Lesson: 1NF, 2NF, 3NF

Introduction:

Normalization is the process of organizing data to reduce redundancy and improve consistency. The main normal forms are 1NF, 2NF, and 3NF.

1. First Normal Form (1NF)

Definition:

- All values must be atomic (indivisible).

- No repeating groups or arrays.

- Each record must be unique.

Example Not in 1NF:

StudentID | Name | Courses

1 | John Doe | Math, English, Science

Corrected 1NF:

StudentID | Name | Course

1 | John Doe | Math

1 | John Doe | English

1 | John Doe | Science

2. Second Normal Form (2NF)

Definition:

- Must be in 1NF.

- No partial dependency (non-key fields must depend on the whole primary key).

Example Not in 2NF:

OrderID | ProductID | ProductName | Quantity

(ProductID -> ProductName creates partial dependency)

Corrected 2NF:

OrderDetails:

OrderID | ProductID | Quantity

Products:

ProductID | ProductName

3. Third Normal Form (3NF)

Definition:

- Must be in 2NF.

- No transitive dependency (non-key fields must not depend on other non-key fields).

Example Not in 3NF:

EmployeeID | EmployeeName | Department | DepartmentLocation

Corrected 3NF:

Employees:

EmployeeID | EmployeeName | Department

Departments:

Department | DepartmentLocation

Summary:

1NF: Atomic data, no repeating groups.

2NF: No partial dependency.

3NF: No transitive dependency.

Normalization helps ensure data integrity, reduce redundancy, and maintain consistent structure.