Database Normalization Forms

By Dhandapani Yedappalli Krishnamurthi Sep 22, 2025

NF

Normalization in DBMS

👉 **Normalization** is the process of structuring a database to **reduce redundancy** and **improve data integrity**.
👉 Done using **Normal Forms (NF)**, introduced by **E.F. Codd**.

1. First Normal Form (1NF)

Rule:

- Data must be stored in **atomic (indivisible) values**.
- No **repeating groups** or arrays.

Example (Unnormalized Table):

Cust_ID | Cust_Name | Orders
1 | Anita | O101, O102
2 | Rahul | O103
❌ Problem: Orders column has multiple values.

In 1NF (Atomic Values):
Cust_ID | Cust_Name | Order_ID
1 | Anita | O101
1 | Anita | O102
2 | Rahul | O103

---

2. Second Normal Form (2NF)

Rule:

- Table must be in **1NF**.
- No **partial dependency** (non-key column depending on part of a composite key).

Example (Violates 2NF):

(Order_ID, Product_ID) → Primary Key

Order_ID | Product_ID | Product_Name | Cust_Name

O101 | P01 | Laptop | Anita

O101 | P02 | Mouse | Anita

❌ Problem: **Cust_Name** depends only on **Order_ID**, not the whole key.

Fix (2NF):

Orders: Order_ID | Cust_Name

OrderDetails: Order_ID | Product_ID | Product_Name

---

3. Third Normal Form (3NF)

Rule:

- Table must be in **2NF**.
- No **transitive dependency** (non-key attribute depending on another non-key).

Example (Violates 3NF):

Cust_ID | Cust_Name | City | Pincode

1 | Anita | Chennai | 600001

2 | Rahul | Bangalore | 560001

❌ Problem: **City → Pincode** (transitive dependency).

Fix (3NF):

Customers: Cust_ID | Cust_Name | City

CityInfo: City | Pincode

---

4. Boyce-Codd Normal Form (BCNF)

Rule:

- Table must be in **3NF**.
- For every **functional dependency (X → Y)**, X must be a **candidate key**.

Example (Violates BCNF):

Professor | Subject | Dept
Smith | DBMS | CS
Smith | SQL | CS
Jones | Networks| IT

❌ Problem: **Professor → Dept** but **Professor** is not a candidate key (since Subject also part of key).

Fix (BCNF):
ProfessorDept: Professor | Dept
ProfessorSubject: Professor | Subject

---

5. Fourth Normal Form (4NF)

Rule:

- Table must be in **BCNF**.
- No **multi-valued dependency**.

Example (Violates 4NF):

Student | Hobby | Language
Anita | Painting | English
Anita | Painting | Hindi
Anita | Dancing | English
Anita | Dancing | Hindi

❌ Problem: Hobby and Language are independent multi-valued facts.
Fix (4NF):
StudentHobbies: Student | Hobby
StudentLanguages: Student | Language

---

6. Fifth Normal Form (5NF / PJNF)

Rule:

- Table must be in **4NF**.
- No **join dependency** (facts should not be split unnecessarily).

👉 Rarely applied in practice. Used for very complex designs.

Summary Table

Normal Form	Rule	Example Fix
1NF	No repeating groups, atomic values	Split multiple orders into rows
2NF	No partial dependency	Split customer info from composite PK
3NF	No transitive dependency	Separate City → Pincode
BCNF	Every determinant must be a candidate key	Split Professor–Dept and Professor–Subject
4NF	No multi-valued dependency	Separate Hobbies and Languages
5NF	No join dependency	Rare in practice

In practice:

- Most databases are designed up to **3NF or BCNF**.
- Higher NFs (4NF, 5NF) are for specialized cases.