

## An Analysis of Normalization Techniques in Database Management System

Ankita Jain, M.C.A. Scholar

[jainankita209@gmail.com](mailto:jainankita209@gmail.com)

Jayoti Vidyapeeth Women's University, Jaipur, Rajasthan

### Abstract

Database normalization or normalization is the process of organizing the columns (attributes) and tables (relations) of a relational database to minimize data redundancy. Normalization involves decomposing a table into less redundant (and smaller) tables without losing information; defining foreign keys in the old table referencing the primary keys of the new ones. The objective is to isolate data so that additions, deletions, and modifications of an attribute can be made in just one table and then propagated through the rest of the database using the defined foreign keys.

**Key Words:** Analysis, Database, Normalization, NF, BCNF, Redundant.

### 1. Introduction

If a database design is not perfect, it may contain anomalies, which are like a bad dream for any database administrator. Managing a database with anomalies is next to impossible.

- **Update anomalies** – If data items are scattered and are not linked to each other properly, then it could lead to strange situations. For example, when we try to update one data item having its copies scattered over several places, a few instances get updated properly while a few others are left with old values. Such instances leave the database in an inconsistent state.
- **Deletion anomalies** – we tried to delete a record, but parts of it was left undeleted because of unawareness, the data is also saved somewhere else.
- **Insert anomalies** – we tried to insert data in a record that does not exist at all.

Normalization is a method to remove all these anomalies and bring the database to a consistent state.

### 2. Normalization of Database

Database Normalization is a technique that organizing the data in database. Normalization is a process of decomposing table to eliminate data redundancy and useless characteristics i.e. Insertion, Update and Deletion Anomalies. It is a multi-step process that removing duplicated data from the relation tables. In other words the normalization is a process in which we renovate an un-normalized relation into relations.

Normalization is used for mainly two purposes,

Eliminate redundant (useless) data.

Ensuring data dependencies make sense i.e. data is logically stored [1].

#### Advantage of normalization

- Eliminate duplicate data that you will be able to reduce the size of the database.
- Good performance.
- Increase storage efficiency.
- It is conceptually clean and easy to maintain and change as your need [2].

#### Disadvantage of normalization

- With the increase the normal forms its performance gets slow.
- It is time consuming and complicated method in normalizing relations of advanced degree [3].

### 3. Types of Normalization

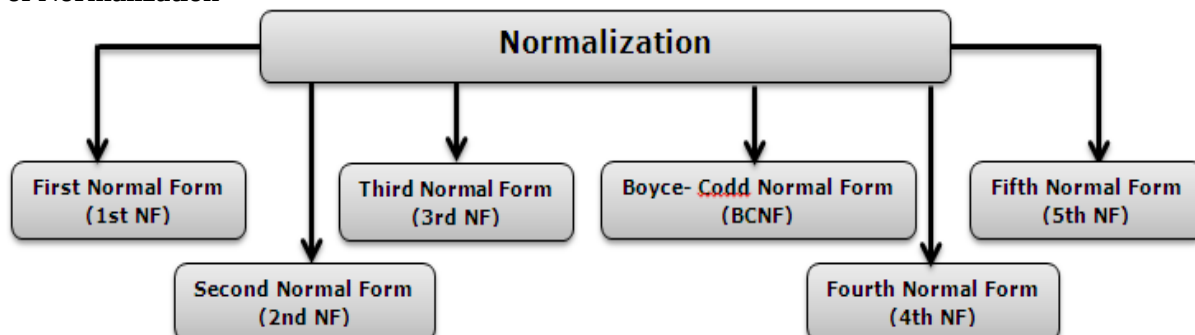


Figure 1: Normalization Flow Model

#### A. First Normal Form (1 NF)

- The table cells must be a single value.
- Eliminate repeated group in individual table.
- Create a separate table for each related data.
- Identify primary key in each table.

**Definition:** An entity is in the first normal form, it contains no repeating groups(data) as well as columns. An entity is in the first normal form if it contains no repeating groups.

| Name Of Student | Age | Subject        |
|-----------------|-----|----------------|
| Naman           | 20  | English, Hindi |
| Ajay            | 30  | Science        |
| Nishita         | 24  | Math           |

**Table 1: Student Record**

In First Normal Form, any row must not have a more than one value is saved. We must separate data into multiple rows [5].

| Age | Name of student | Subject |
|-----|-----------------|---------|
| 20  | Naman           | English |
| 20  | Naman           | Hindi   |
| 30  | Ajay            | Science |
| 24  | Nishita         | Math    |

**Table 2: Student record 1NF**

**B. Second Normal Form (2 NF):** In the Second Normal Form, remove partial dependency. A partial dependency means a type of functional dependency where an attribute is functionally dependent on its primary key. A relation is in 2NF if it is in 1NF and every non-key attribute is fully dependent on each candidate key.

| Name of Student | Age |
|-----------------|-----|
| Naman           | 20  |
| Ajay            | 30  |
| Nishita         | 24  |

**Table 2: Student Record**

| Name of Student | Subject |
|-----------------|---------|
| Naman           | English |
| Naman           | Hindi   |
| Ajay            | Science |
| Nishita         | Math    |

In Student record Table the candidate key will be **name of Student** column, because all other column i.e. **Age** is dependent on it.

### C. Third Normal form (3 NF)

- Remove transitive dependency.
- Transitive Dependency:- A type of functional dependency where an attribute is functionally dependent on an attribute other than the primary key.

A relation is in third normal form, if it is in 2NF and every non-key attribute of the relation is non-transitively dependent on each candidate key of the relation.

| Company  | City    | State | Zip   |
|----------|---------|-------|-------|
| ABC Ltd. | Mumbai  | MH    | 12344 |
| XYZ Ltd. | Noida   | UP    | 12346 |
| PQR Ltd. | Chennai | TN    | 67834 |

This table is not in the 3NF. In this example, the city and state are dependent upon the ZIP code. To place this table in 3NF, two separate tables are created -- one containing the company name and ZIP code and the other one containing city, state, ZIP code pairings.

| Company  | Zip   |
|----------|-------|
| ABC Ltd. | 12344 |
| XYZ Ltd. | 12346 |
| PQR Ltd. | 67834 |

| City    | State | Zip   |
|---------|-------|-------|
| Mumbai  | MH    | 12344 |
| Noida   | UP    | 12346 |
| Chennai | TN    | 67834 |

**D. Boyce-Coded Normal Form (BCNF):** if every determinant is a candidate key then it is a relation of Boyce-Coded Normal Form (BCNF). It is used in database normalization.

**Example:**

**Client Interview:**

| Client No | Interview Date | Interview Time | Staff No | Room No |
|-----------|----------------|----------------|----------|---------|
| Sr78      | 14-may-11      | 10:00          | SF6      | G101    |
| Sr76      | 17-may-11      | 12:50          | SF5      | G101    |
| Sr45      | 17-may-11      | 12:10          | SF45     | G102    |
| Sr56      | 06-july-11     | 10:20          | SF34     | G102    |

FD1 Client No, Interview Date -> Interview Time, Staff No, Room No (Primary Key)  
 FD2 Staff No, Interview Date, Interview Time -> Client No (Candidate key)  
 FD3 Room No, Interview Date, Interview Time -> Client No, Staff No (Candidate key)  
 FD4 Staff No, Interview Date -> Room No (not a candidate key).

To transform the Client Interview relation to BCNF, we must remove the violating functional dependency by creating two new relations called Interview and Staff Room as shown below.

Interview (Client No, Interview Date, Interview Time, Staff No), Staff Room (Staff No, Interview Date, Room No).

#### E. Fourth Normal Form (4th NF)

In Fourth Normal Form, Has no multiple sets of multi-valued dependency. In other terms, 4NF states no entity can have more than a single one-to-many relationship within an entity if the one-to-many attributes are independent of each other.

- Fourth Normal Form applies many-to-many relationships.

**Definition:** A table is in fourth normal form (4NF) if and only if it is in BCNF and contains no more than one multi-valued dependency.

#### F. Fifth Normal Form (5th NF)

In Fifth Normal Form

- A relation that has a join dependency cannot be decomposed by a projection into other relations without spurious results.
- A relation is in 5NF when its information content cannot be reconstructed from several smaller relations i.e. from relations having fewer attributes than the original relation

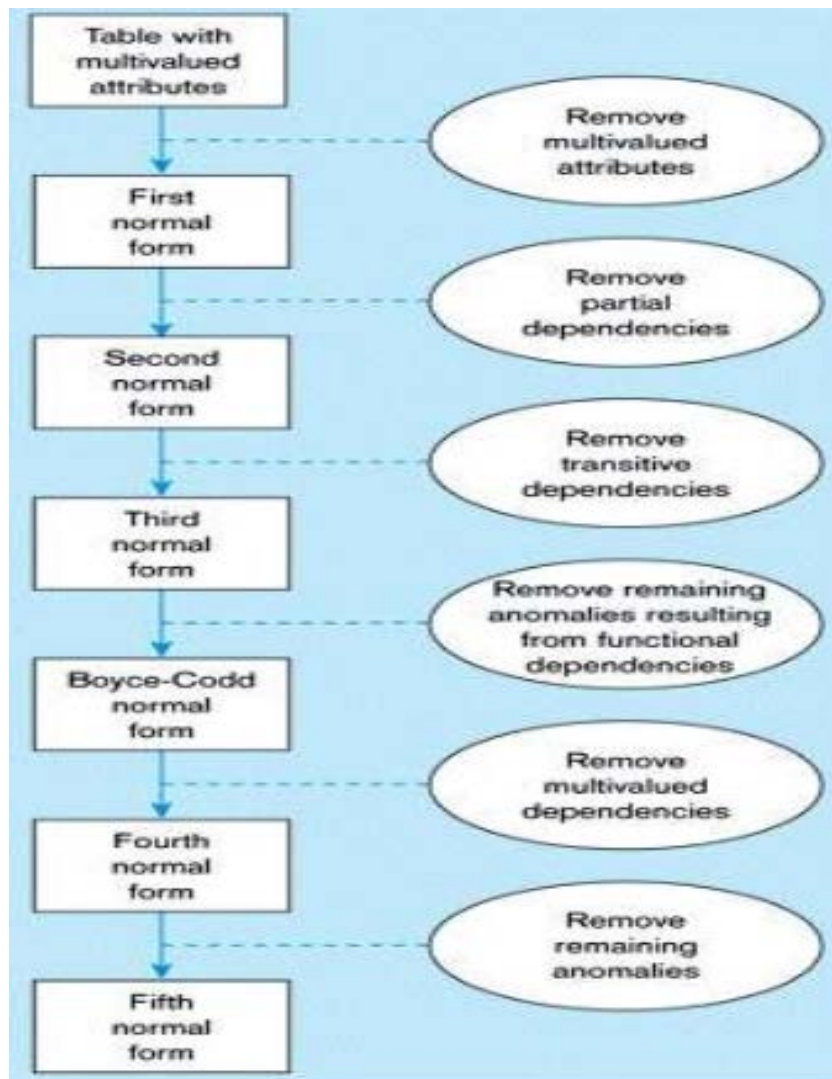


Figure 2: Steps in Normalization

#### 4. Reference

- [1]. <http://www.studytonight.com/dbms/database-normalization.php>.
- [2]. <http://technology.blurtit.com/595612/advantages-and-disadvantages-of-database-normalization>.
- [3]. <http://www.indiastudychannel.com/resources/162467-What-is-normalization-What-are-its-merits-demerits-and-types.aspx>.
- [4]. <http://www.csharpcorner.com/uploadfile/nipuntomar/normalization-and-its-types>.
- [5]. <http://www.studytonight.com/dbms/database-normalization.php>.
- [6]. <http://www.timlin.net/csm/cis363/Normalization.pdf>.