**NORMALIZATIONS**

Database normalization is the process of structuring a relational database in accordance with a series of normal forms in order to reduce data redundancy and improve data integrity.

This includes creating tables and establishing relationships between those tables according to rules designed both to protect the data and to make the database more flexible by eliminating redundancy and inconsistent dependency.

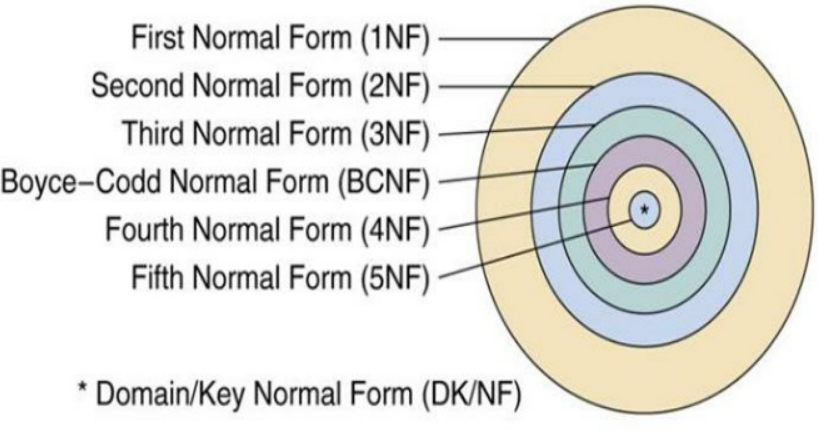
**Redundancy**

**A condition created within a database or data storage technology in which the same piece of data is held in two separate places**.

**Anomalies**

**A fault in a database that usually emerges as a result of shoddy planning and storing everything in a flat database**

* **Insertion Anomaly:** Insertion Anomaly refers to when one cannot insert a new tuple into a relationship due to lack of data.
* **Deletion Anomaly:** The delete anomaly refers to the situation where the deletion of data results in the unintended loss of some other important data.
* **Update Anomaly:** The update anomaly is when an update of a single data value requires multiple rows of data to be updated.



***1NF***

Relation is in first normal form if it does not contain any **composite** or **multi-valued attribute**. A relation is in first normal form if every attribute in that relation is singled valued attribute.

**If a relation contains a composite or multi-valued attribute, it violates the first normal form.**

example

ID Name Courses

------------------

1 A c1, c2

2 E c3

3 M C2, c3

(here column 1 and 3 has multi valued attribute)

In 1NF

ID Name Course

------------------

1 A c1

1 A c2

2 E c3

3 M c2

3 M c3

Here primary key can be (ID, Course) as composite key.

First Normal Form (1NF) does not eliminate redundancy, but rather, it’s that it eliminates repeating groups.

Instead of having multiple columns of the same kind of data in a record, (0NF or Unnormalized form) you remove the repeated information into a separate relation and represent them as rows. This is what constitutes 1NF.

**Partial dependency**

a nonprime attribute is functionally dependent on part of a candidate key. (A nonprime attribute is an attribute that's not part of any candidate key.)

**Suppose {A,B} → {C} but also {A} → {C}. Then {A,B} → {C} is partial; {C} is partially functionally dependent on {A,B};**

**2NF**

A relation must be **in first normal form** and relation **must not contain any partial dependency**.

STUD\_NO COURSE\_NO COURSE\_FEE

1 C1 1000

2 C2 1500

1 C4 2000

4 C3 1000

4 C1 1000

2 C5 2000

COURSE\_FEE cannot alone decide the value of COURSE\_NO or STUD\_NO;  
COURSE\_FEE together with STUD\_NO cannot decide the value of COURSE\_NO;  
COURSE\_FEE together with COURSE\_NO cannot decide the value of STUD\_NO;  
Hence,  
COURSE\_FEE would be a non-prime attribute, as it does not belong to the one only candidate key {STUD\_NO, COURSE\_NO} ;  
But, COURSE\_NO -> COURSE\_FEE, i.e., **COURSE\_FEE is dependent on COURSE\_NO, which is a proper subset of the candidate key (partial dependency).** Non-prime attribute COURSE\_FEE is dependent on a proper subset of the candidate key, which is a partial dependency and so this relation is not in 2NF.

**Table 1** **Table 2**

STUD\_NO COURSE\_NO COURSE\_NO COURSE\_FEE

1 C1 C1 1000

2 C2 C2 1500

1 C4 C3 1000

4 C3 C4 2000

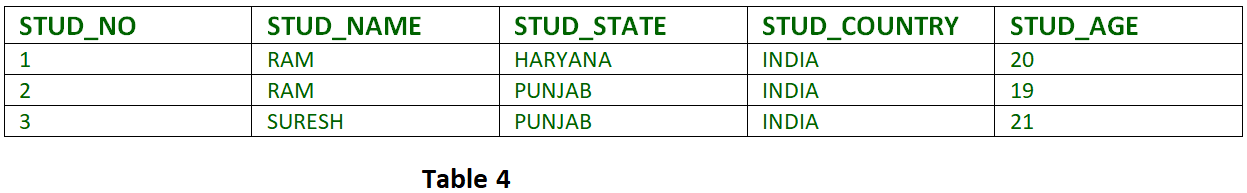
4 C1 C5 2000

2 C5

**3NF**

A relation that is in First and Second Normal Form and in which no non-primary-key attribute is transitively dependent on the primary key, then it is in Third Normal Form (3NF).

No non-key attribute should be dependent on another non key attribute.



FD set:  
{STUD\_NO -> STUD\_NAME, STUD\_NO -> STUD\_STATE, STUD\_STATE -> STUD\_COUNTRY, STUD\_NO -> STUD\_AGE}

Candidate Key:  
{STUD\_NO}

For this relation in table 4, STUD\_NO -> STUD\_STATE and STUD\_STATE -> STUD\_COUNTRY are true. So STUD\_COUNTRY is transitively dependent on STUD\_NO. It violates the third normal form. To convert it in third normal form, we will decompose the relation STUDENT (STUD\_NO, STUD\_NAME, STUD\_PHONE, STUD\_STATE, STUD\_COUNTRY\_STUD\_AGE) as:

STUDENT (STUD\_NO, STUD\_NAME, STUD\_PHONE, STUD\_STATE, STUD\_AGE)

STATE\_COUNTRY (STATE, COUNTRY)

**BCNF**

*(Boyce-Codd normal form)*

A relation is in BCNF if it’s already in 3NF and For all non-trivial functional dependency

X -> Y

Either x is a super key or Y is prime attribute