NORMALIZATION

Normalization is the process of organizing data in a database. 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.

The main goal of database normalization is to restructure the logical model of a database to:

> Eliminate redundancy.
> Organize data efficiently.
> Reduce the potential for data anomalies (data anomalies are inconsistency in data stored as a result of database operation)

1st Normalization form:

> Each table has a primary key.
> The values in column of a table are atomic (no multi-value attributes).
> There are no repeating groups.

How we converted our database into 1st NF:

Step 1: Recognizing and eliminating the repeating groups in tables.

Step 2: Identifying the primary key for each table.

Step 3: Identifying all the dependencies in each table.

Below is songs table from our music database.

song_id	song_name	song_album	song_artist	song_genre	song_length	song_language	song_popularity
1	Beat It	Thriller 25 Super Deluxe Edition	Michael Jackson	Pop	3	English	81
2	Smooth Criminal	Bad 25th Anniversary	Michael Jackson	Pop	3	English	79
3	Gimme More	Blackout	Britney Spears	Pop	3	English	78
4	Eye of the Tiger	Rocky IV	Survivor	Rock	3	English	76
5	Everybody Wants To Rule The World	Songs From The Big Chair (Super Deluxe Edition)	Tears For Fears	R&B	3	English	86
6	Everybody Talks	Picture Show	Neon Trees	R&B	2	English	79

7	Earned It (Fifty Shades Of Grey)	Beauty Behind The Madness	The Weeknd	Pop	3	English	74
8	In Da Club	Get Rich Or Die Tryin'	50 Cent	Hip-Hop	2	English	82
9	Black and Yellow	Rolling Papers	Wiz Khalifa	Rap	3	English	75
10	Superman	The Eminem Show	Eminem, Dina Rae	Hip-Hop	4	English	0

All the records in the table are atomic and there are no repeating groups.

2nd Normalization Form:

> All requirements from 1st NF must be met.
> Redundant data across multiple rows of table must be moved to a separate table.
> The resulting table must be related to each other by use of foreign key.

Table is in 2nd Normal form when:

- It is in 1 Normal Form, and
- It includes no partial dependencies, i.e., no attribute is dependent on only a portion of primary key. In other words, every non-primary-key attribute is fully functionally dependent on the primary key.

How we converted our database into 2 NF:

After 1 NF,

Step 1: Writing each key component on a separate line.

Step 2: Assigning corresponding dependent attributes.

Below is the artist table from our music database

id	**name**	**age**	**country**	**recordLabel**
1	Michael Jackson	50	USA	Sony
2	Britney Spears	30	USA	Warner
4	Tears For Fears	34	Ireland	CBS
5	Neon Trees	23	Germany	Universal
6	The Weeknd	35	USA	BMG
7	50 Cent	36	Mexico	ARMIND
8	Wiz Khalifa	32	USA	Sony
9	Eminem	33	USA	Warner

For example, recordLable is dependent is fully functionally dependent on the primary key which is artist_id. And all the records in artist table are atomic and no repeating groups.

3rd Normalization Form:

> All requirements of 2 NF must be met.
> Eliminate the keys that do not depend on primary key. i.e., any key that is dependent not only on primary key but also on another field must be moved to another table.

Table is in 3rd Normal Form when following are true:

- It is in 2nd Normal Form.
- It contains no transitive dependencies.
- There are no fields or attributes that does not depend on the key, i.e., non-key attributes must be dependent on key(s) but and only on the key(s).

How we converted our database into 3rd Normal Form:

Step 1: Identified each new determinant (determinant is any attribute whose value determines other values within a row.

Step 2: Identifying attributes dependent on each determinant identified in step 1 and identify dependency.

Step 3: Removing dependent attributes from transitive dependency.

Below is the view from sings table.

song_name	**artist_name**
Beat It	Michael Jackson
Smooth Criminal - 2012 Remaster	Michael Jackson
Gimme More	Britney Spears
Everybody Wants To Rule The World	Tears For Fears
Everybody Talks	Neon Trees
Earned It (Fifty Shades Of Grey)	The Weeknd
In Da Club	50 Cent
Black and Yellow	Wiz Khalifa
In Da Club	50 Cent

'sings' table contains only songs_id and artist_id as attributes which are foreign keys from artist and songs table respectively, using those id's we have fetched the names of songs and artists accordingly.

This shows that there is no partial as well as transitive dependencies present in either songs or artist table.