Normalization

- Normalization is the process of organizing the data in the database.
- Normalization is used to minimize the redundancy from a relation or set of relations. It is also used to eliminate the undesirable characteristics like Insertion, Update and Deletion Anomalies.
- Normalization divides the larger table into the smaller table and links them using relationship.
- The normal form is used to reduce redundancy from the database table.

EXAMPLE:

BEFORE NORMALIZATION:

FILM_ID	FILM_NAME	FILM_CATEGORY_ID	GENRE	ACTOR_ID	ACTOR_NAME
101	Titanic	1	Romance	1	Leonardo
102	Interstellar	2	Sci-Fi	2	Matthew
103	John Wick	3	Action	3	Reeves
104	John Wick 2	3	Action	3	Reeves
105	Iron Man	3	Action	4	Tony

AFTER NORMALIZATION:

FILM TABLE

FILM_ID	FILM_NAME	FILM_CATEGORY_ID	ACTOR_ID
101	Titanic	1	1
102	Interstellar	2	2
103	John Wick	3	3
104	John Wick 2	3	3
105	Iron Man	3	4

FILM CATEGORY TABLE

FILM_CATEGORY_ID	GENRE
1	Romance
2	Sci-Fi
3	Action
3	Action
3	Action

ACTOR TABLE

ACTOR_ID	ACTOR_NAME
1	Leonardo
2	Matthew
3	Reeves
3	Reeves
4	Tony

First Normal Form (1NF)

- A relation will be 1NF if it contains an atomic value.
- It states that an attribute of a table cannot hold multiple values. It must hold only single-valued attribute.
- First normal form disallows the multi-valued attribute, composite attribute, and their combinations.

EXAMPLE:

NOT In 1NF:

FILM_ID	FILM_NAME	ACTOR_ID
101	Titanic	1, 4
102	Interstellar	2, 3
103	John Wick	3, 4

IN 1NF:

FILM_ID	FILM_NAME	ACTION_ID
101	Titanic	1
101	Titanic	4
102	Interstellar	2
102	Interstellar	3
103	John Wick	3
103	John Wick	4

Second Normal Form (2NF)

- In the 2NF, relational must be in 1NF.
- In the second normal form, all non-key attributes are fully functional dependent on the primary key.

EXAMPLE:

NOT IN 2NF

FILM_CATEGORY_ID	ACTOR_ID	REMUNERATION	GENRE
1	1	1 CR	Romance
1	3	1.5 CR	Romance
2	2	1 CR	Sci-fi
3	3	2 CR	Action
3	2	1.5 CR	Action

IN 2NF

FILM_CATEGORY_ID	ACTOR_ID	REMUNERATION
1	1	1 CR
1	3	1.5 CR
2	2	1 CR
3	3	2 CR
3	2	1.5 CR

FILM_CATEGORY_ID	GENRE
1	Romance
2	Sci-Fi
3	Action

Third Normal Form (3NF)

- A relation will be in 3NF if it is in 2NF and not contain any transitive partial dependency.
- If there is no transitive dependency for non-prime attributes, then the relation must be in third normal form.

EXAMPLE:

NOT IN 3NF

FILM_ID	FILM_NAME	ACTOR_ID	RELEASE_YEAR	ACTOR_NAME
101	Titanic	1	2006	Leonardo
102	Interstellar	2	2012	Matthew
103	John Wick	3	2009	Reeves
104	John Wick 2	3	2011	Reeves
105	Iron Man	4	2019	Tony

IN 3NF

FILM_ID	FILM_NAME	ACTOR_ID	RELEASE_YEAR
101	Titanic	1	2006
102	Interstellar	2	2012
103	John Wick	3	2009
104	John Wick 2	3	2011
105	Iron Man	4	2019

ACTOR_ID	ACTOR_NAME
1	Leonardo
2	Matthew
3	Reeves
4	Tony

Boyce Codd normal form (BCNF)

- BCNF is the advance version of 3NF. It is stricter than 3NF.
- A table is in BCNF if every functional dependency X → Y, X is the super key of the table.
- For BCNF, the table should be in 3NF, and for every FD, LHS is super key.

EXAMPLE:

NOT IN BCNF:

FILM_CATEGORY_ID	ACTOR_ID	FILM_NAME
1	1	Titanic
2	2	Interstellar
3	3	John Wick
3	3	John Wick 2
1	4	Iron Man

IN BCNF:

FILM_ID	ACTOR_ID
101	1
102	2
103	3
104	3
105	4

FILM_ID	FILM_NAME	FILM_CATEGORY_ID
101	Titanic	1
102	Interstellar	2
103	John Wick	3
104	John Wick 2	3
105	Iron Man	1