

Database normalization removes redundancy and ensures that data is logically stored. It removes insertion, deletion, and updation anomalies. Here I'll briefly explain the 5 normalization rules.

1. First Normal Form
2. Second Normal Form
3. Third Normal Form
4. BCNF: Boyce and Codd Normal Form
5. Fourth Normal Form

First Normal Form

A table is supposed to be in first normal form if,

- All the attributes are single-valued (atomic).
- All the columns have unique names.
- The order in which data is stored does not matter.

Example: In Shirt_Info table, Size attribute is not atomic. Hence it can be decomposed into Design_Info and Size_Info as shown in the image.

Shirt_Info			Design_Info		Size_Info	
Shirt_ID	Design	Size	Shirt_ID	Design	Shirt_ID	Size
100	Mickey Mouse	XXL,XL,M	100	Mickey Mouse	100	XXL
101	Flowers	M	101	Flowers	100	XL
102	Good Vibes	M,S	102	Good Vibes	100	M
103	Triangles	XL	103	Triangles	101	M
Primary Key: Shirt_ID			Primary Key: Shirt_ID		102	M
					102	S
					103	XL
					Primary Key: Shirt_ID, Size	

Second Normal Form

A table is supposed to be in second normal form if,

- It is in the 1st normal form.
- It does not have any partial dependency

Example: In Customer_Info table, Store_Name depends on Store_ID and not on Cust_ID. This is a partial dependency. Hence, Customer_Info is not in second normal form (though it satisfies 1NF). It can be decomposed into Customer_Data and Store_Data as shown below.

Customer_Info				Customer_Data		Store_Data	
Cust_ID	Store_ID	Customer_Name	Store_Name	Cust_ID	Customer Name	Store_ID	Store_Name
568	74896	Liam	JPM_1	568	Liam	74896	JPM_1
574	74896	Charles	JPM_1	574	Charles	25614	JPM_2
568	25614	Liam	JPM_2	145	Marcus	Primary Key: Store_ID	
235	25614	Cyrus	JPM_2	235	Cyrus		
Primary Key: Cust_ID, Store_ID				Primary Key: Cust_ID			

Second Normal Form Example

Third Normal Form

A table is supposed to be in third normal form if,

- It satisfies 2nd normal form.
- It does not have any transitive dependency.

Example: Movie_Info is in second normal form but it has a transitive dependency. Therefore, it is not in third normal form. It can be decomposed into Movie_Rating and Genre_Info as shown below.

Movie_Info				Movie_Rating			Genre_Info	
MID	GENRE_ID	GENRE	RATING	MID	RATING	GENRE_ID	GENRE_ID	GENRE
1	G1	Comedy	4.2	1	4.2	G1	G1	Comedy
2	G2	Romance	4.5	2	4.5	G2	G2	Romance
3	G3	Action	2.1	3	2.1	G3	G3	Action
4	G4	Horror	3.5	4	3.5	G4	G4	Horror
Primary Key: MID				Primary Key: MID			Primary Key: GENRE_ID	

In Movie_Info, MID -> GENRE_ID -> GENRE
Which implies MID -> GENRE. Hence, there is a transitive dependency ie. It is not in 3rd NF.

Third Normal Form Example

Boyce and Codd Normal Form (BCNF)

A table is supposed to be in BCNF if,

- It is in 3rd Normal Form.
- For every dependency $X \rightarrow Y$, X cannot be a non-prime attribute if Y is prime attribute (i.e. X should be a super key)

Example: Student_Info is not in BCNF because in the dependency Faculty -> Subject, Subject is prime attribute and Faculty is non-prime attribute. It can

Open in app ↗

Sign up

Sign in

Medium

Search

Write



Algorithm.

Student_Info			Student_Faculty_Info		Faculty_Info		
Student_ID	Subject	Faculty	SID	FID	FID	Faculty	Subject
101	DBMS	Dan	101	1	1	Dan	DBMS
101	Data Science	Blair	101	2	2	Blair	Data Science
102	DBMS	Cyrus	102	3	3	Cyrus	DBMS
103	Cloud Computing	Serena	103	4	4	Serena	Cloud Computing
104	DBMS	Dan	104	1	Primary Key: FID		
Primary Key: Student_ID, Subject			Primary Key: SID, FID				
(Assume that a faculty takes only one course.) In Student_Info, (Student_ID, Subject) -> Faculty But, Faculty -> Subject where Faculty is non-prime and Subject is prime attribute. Hence it is not in BCNF.							

BCNF Example

Fourth Normal Form

A table is supposed to be in fourth normal form if,

- It is in BCNF.
- It has no multi-valued dependency.

Multi-valued Dependency: For a dependency $A \twoheadrightarrow B$, if for a single value of A, multiple values of B exists, then the relation will be a multi-valued dependency.

Student_Info			Student_Course		Student_Hobby	
SID	Course	Hobby	SID	Course	SID	Hobby
1	Computer	Writing	1	Computer	1	Writing
1	Math	Coding	1	Math	1	Coding
2	Chemistry	Writing	2	Chemistry	2	Writing
3	Biology	Cricket	3	Biology	3	Cricket
4	Physics	Cooking	4	Physics	4	Cooking
			Primary Key: SID, Course		Primary Key: SID, Hobby	
In student_info, SID=1 gives more than one record. Hence, it is a multi-valued dependency.						

Fourth Normal Form Example