NORMALIZATION IN DBMS...

What is normalization?

- Normalization is the process of Organizing the data and the attributes of a database.
- It performed to reduce the data redundancy in a database & ensure that data is stored logically.
- Normalization is systematic approach of decomposing table to eliminate data redundancy and undesirable characteristics like insertion, update and delete.
- Normalization is multi-step process that puts data in tubular form & remove duplicate data from relation tables.

Table of office employees.

Id	NAME	ADDRESS	PROFESSION
101	Vishal	Pune	Developer
102	Vijay	Nagar	Accountant
103	Shree	Beed	Scientist
104	Pranev	Nasik	Manager
105	Saurabh	Mumbai	Clerk
106	Mayur	Thane	Operator.

In this table, we have data of office employees

1]. Insertion Anomaly :-

An insertion anomaly occurs in the relation database when some attributes or data items are inserted into database without existance of other attributes.

2]. Updation Anomaly :-

Updation Anomaly occurs when the same data item are repeated with the same values are not linked to each other.

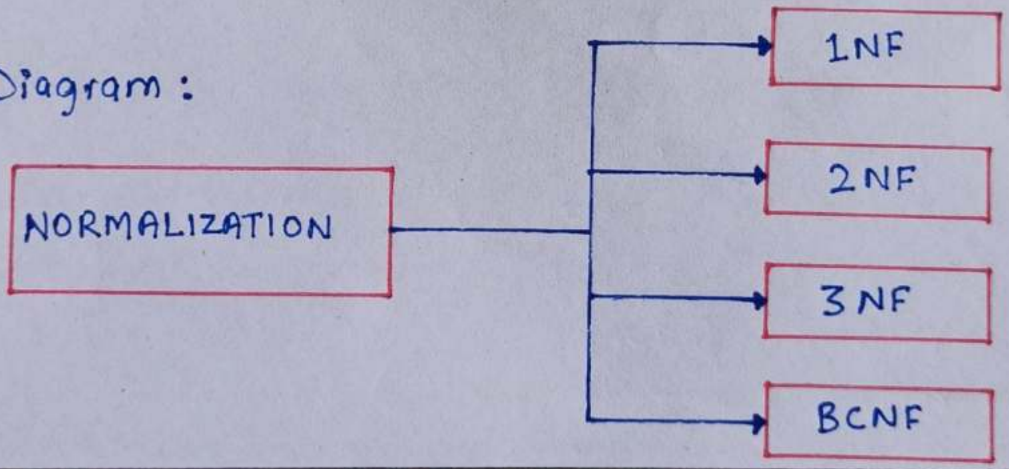
3]. Deletion Anomalies :-

Deletion Anomalies occurs when deleting one part of the data deletes the other necessary information from the database.

★ TYPES OF NORMALIZATION

i]. 1 NF
ii]. 2 NF
iii]. 3 NF
iv]. BCNF

- Diagram :



★ Function Dependancy :-

It is a relationship that exist between two sets of attributes of a relational table where one set of attributes can determine value of other set of attributes.

Denoted by x > y

∴ X is determinant and Y is dependant.

1]. 1NF (First Normal Form :-

In 1NF relation each table cell showd contain a single value. Each record looks like unique.

Player ID	Player Name	Game
01	Sham	Cricket, Hockey
02	Vaishnav	Football
03	Gourav	Basket ball.

Here in Game row we stored two games so it is multi-valued attribute. Its not 1NF relation.

Convert it into 1NF

Player ID	Player Name	Game
01	Sham	Cricket
01	Sham	Hockey
02	Vaishnav	Football
03	Gaurav	Basket ball.

Its simple method to store game saperately in 1NF. Now this is First Normal form.
1NF wants to store unique information in table without data repeatation.

2]. 2NF (Second Normal Form):-

In 2NF relation must be in 1NF. In the Second Normal form, all non-key attributes are fully functionally dependant on the Primary key.

Student ID	Specialization	Student Age
501	Physics	22
501	Math	22
502	Zoology	24
503	Sanskrit	27
503	Botany.	27

In this table, studentAge is depend on student ID which is subset of candidate key. Its is not 2NF relation.

Convert this table in two parts.

Student ID	StudentAge.
501	22
502	24
503	27

And

Student ID	Specialization.
501	Physics
501	Math
502	Zoology
503	Sanskrit
503	Botany.

The following two tables are satisfy the conditions of 2NF relation. It also in 1NF form and every non-prime attributes is dependant on primary key.

3]. 3NF (Third Normal Form):-

The relation in 3NF if it is in 2NF and no transition dependancy exist. Non-prime attribute is dependant on the primary key.

Student ID	Student Name	Student zipcode	StudentCity.
301	Aditya	422602	Pune
301	Aditya	411000	Kolhapur
302	Mandar	400001	Kalyan
303	Rushi	400099	Mumbai

Its not 3NF because student ID -> student City transitive dependancy.
Student zipcode is not super key & studentcity is not prime attribute.

Student ID	StudentName	Student Zipcode
301	Aditya	422602
301	Aditya	411000
302	Mandar	400001
303	Rushi	400099

< Student Location >

Student Zipcode	StudentCity.
422602	Pune
411000	Kolhapur
400001	Kalyan
400099	Mumbai

→ We converted the table into 3NF by converting it into two parts & they don't have transitive dependancy.

→ Some dependancies cause redundancy in data base.

→ Redundancy removed by BCNF.

4]. BCNF (Boyce - codd Normal form) :-

- Boyce - codd Normal form is next part of 3NF.
- Table must be in 3NF.
- Table in BCNF if every function dependancy X → Y X is superkey of table.

Employee code	Project ID	Project Leader
01	F03	Ajit
01	F01	Sanskar
02	F04	Rohan
03	F02	Pravin

- For non-trival functional dependancy, Project leader -> project , ProjectId is prime attribute but ProjectLeader is not prime attribute.
- For BCNF convert table into tree parts.

Employee code	Project Id.
01	F03
01	F01
02	F04
03	F02

ProjectLeader	Project Id
Ajit	F03
Sanskar	F01
Ronan	F04
Pravin	F02

Thus, we converted tables into BCNF. By factoring it.