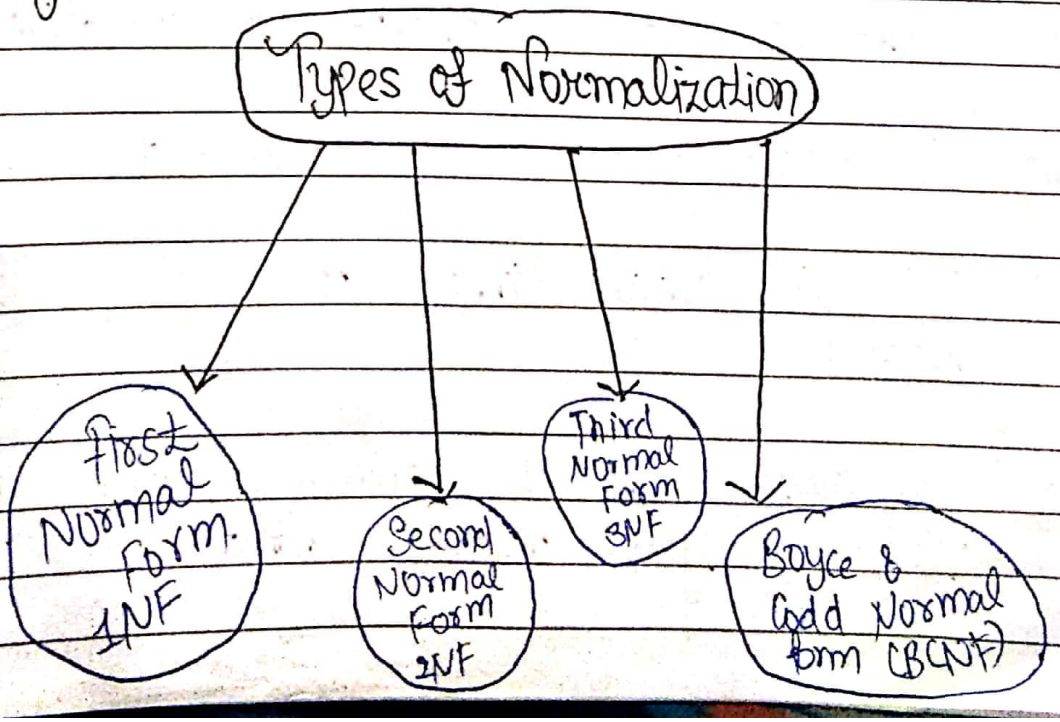
#DBMS Notes

__/__/____

Notes by :- JPWebdevelopers.in

* 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.
- Normalization is based on the concept of Normal forms.
- Normal forms are used to eliminate or reduce redundancy in database tables.
- It is used to avoid the data redundancy, insertion anomaly, update anomaly & delete anomaly



__/__/____

Five Normal forms These are 1NF, 2NF, 3NF, 4NF, 5NF where NF stands for Normal forms.

→ The 1NF, 2NF, 3NF were proposed by Dr. E.F. Codd. Later on 3NF known as BCNF was proposed by Boyce and Dr. codd.

→ The 4NF and 5NF Normal forms were proposed after BCNF.

* Advantages of Normalization :-

(i) Normalization helps to minimize the data redundancy
(ii) Better understanding of data.
(iii) More efficient data Structure.
(IV) More flexible data structure.
(V) Enforces the concept of Relational integrity

* Disadvantages of Normalization :-

(i) It is very time consuming and difficult process.

(ii) you cannot start building the database before you know what the user needs.

__/__/____

(i) First Normal Form (1NF):-

- All Relation is said to be in 1NF if and only if it follow the Rules:-
  - → A relation will be 1NF if it contains an atomic value.
  - → There are no repeating groups in the table.
  - → Table should not contain any multivalued Attribut
- It is a technique to Remove or Reduce Redundancy from a table.
- It must hold only single-valued attribute.

Example:- Relation EMPLOYEE is not in 1NF because of multi-valued attributes

EMP_ID	EMP_NAME	EMP_PHONE
14	John	7282626365
		9064 3828
20	Harry	7393702
12	Sam	8589 03022
		8574 7838

→ The decomposition of the EMPLOYEE table into 1NF :-

EMP_ID	EMP_NAME	EMP_PHONE
14	John	7282626365
14	John	9064 3828
20	Harry	7393702
12	Sam	8589 03022
12	Sam	8574 7838

__/__/____

(2) Second Normal Form :-

- 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: A school can store the data of teachers and the subjects they teach. In a school, a teacher can teach more than one subject.

Teacher table

TEACHER ID	SUBJECT	TEACHER_AGE
25	Java	30
25	Python	30
47	DBMS	35
83	Maths	38
83	Computer	38

- non-prime attribute is TEACHER_AGE is dependent on TEACHER_ID.

→ To convert the given table into 2NF, we decompose it into two tables.

__/__/____

TEACHER_ID	TEACHER AGE
25	30
47	35
83	38

TEACHER_ID	SUBJECT
25	Java
25	Python
47	DBMS
83	Maths
83	Computer

- To be in second normal form, a relation must be in 1NF and relation must not contain any partial dependency.

Partial Dependency :- A Partial dependency would occur whenever a non-Prime attribute depends functionally on Part of candidate key.

The 2NF (Second Normal form) eliminates the Partial dependency.

(3) Third Normal form (3NF) :-

- 3NF is an upgrade to second normal form.
- When a table is in the Second Normal form and has no transitive dependency, then it is the Third Normal Form.

Transitive Dependency :- When an indirect relationship causes functional dependency it is called Transitive Dependency.

Example: If $P \rightarrow Q$ and $Q \rightarrow R$ is true then $P \rightarrow R$ is a transitive dependency.

Rules for 3NF :-

(i) Table should be in 2NF.
(ii) There is no transitive dependency.

Example :- Consider a relation student (rollno, game, fee)

Rollno	Game	fee
1	Basketball	500
2	Basketball	500
3	Basketball	500
4	Cricket	600
5	Cricket	600
6	Tennis	400

__/__/____

F - (rollno, -> game, rollno -> fee, game -> fee).

Rollno$^+$ = (rollno, game, fee)
⇒ rollno is primary key.

→ fee has transitive dependency on rollno via game.

Anomalies
= = = =

(i) Insertion anomaly :- A new game can't be inserted into the table unless we get a student to play that game.

(ii) Deletion anomaly :- If rollno 6 is deleted from the table we also lost the complete information regarding tennis.

(iii) Updation anomaly :- To change the fee for basketball we need to make changes in more than one place.

Decomposition for 3NF :-

To overcome these anomalies, the student table should be divided into smaller tables.

So divide the student table into R1 (game, fee) and R2 (rollno, game).

R1

Rollno	Game
1	Basketball
2	Basketball
3	Basketball
4	Cricket
5	Cricket
6	tennis

Q2

Game	fee
Basketball	500
Cricket	600
Tennis	400

__/__/____

* Boyce Codd Normal Form (BCNF) :-

- BCNF is the advance version of 3NF. It is stricter than 3NF.
- A relation in BCNF, if it is already in 3NF and determinant is a candidate key.
- Determinant is a simple attribute or composite attributes which can uniquely determines the value of other attributes.

Example:- Table : ITEM

Item No.	Item_Name	Quantity	Price
1001	Pen	20	50
1001	Pencil	15	10
1004	Eraser	30	10
1007	Pen	25	50

→ In this, attribute Quantity is fully functionally dependent on (item_No, item_Name)

(item_No, item_Name) → Quantity

Composite determinant

→ The attribute Price is functionally dependent on item name

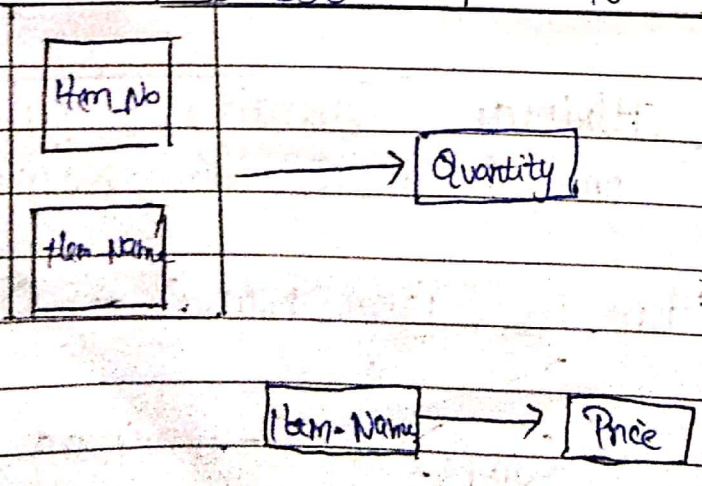
Item_Name ⇒ Price

↳ Simple attribute determinant.

→ The normalized relations / tables for above dependencies are shown below :-

Item_No	Item_Name	Quantity
1001	Pen	20
1001	Pencil	15
1004	Eraser	30
1007	Pen	25

Item_Name	Price
Pen	50
Pencil	10
Eraser	10



__/__/____

* Fourth Normal Form (4NF) :-

A Relation 'R' is in fourth Normal form (4NF) if it is in BCNF and has no multi value dependency.

Multi valued dependency :- A table is said to have multi-valued dependency, if the following conditions are true,

1. For a dependency A → B, if for a single value of A, multiple value of B exists, then the table may have multi-valued dependency.

2. A table should have atleast 3 Colomns for it to have a multi-valued dependency.

3. for a relation R(A,B,C) if there is a multi-valued dependency between A and B, then B and C should be independent of each other.

Example :- A →→ B and A →→ C

Name	department	Language
John	Sales	English
John	Sales	French
John	Marketing	English
Micheal	Sales	English
Micheal	Sales	French
Micheal	Sales	Japanese

__/__/____

In this Table :-
The attribute Department is multi valued dependent on attribute Name.

Name →→ Department

The attribute Language is multivalued dependent on attribute Name.

Name →→ Language.

→ So Convert the above relation (EMP) with multi valued dependencies into 4NF, move each multi-valued dependency pair to a new table.

Emp

Name	Department	Language
John	Sales	English
John	Sales	French
John	Marketing	English
Micheal	Sales	English
Micheal	Sales	French
Micheal	Sales	Japanese

↙ Table EMP_DEPT

Name	Department
John	Sales
John	Marketing
Micheal	Sales

→ Table EMP_LANG

Name	Language
John	English
John	French
Micheal	English
Micheal	French
Micheal	Japanese

__/__/____

(5) Fifth Normal form (5NF) :-

- A relation R in the fifth Normal form (5NF) if it is already in 4NF and has no join dependency and joining should be lossless.

- 5NF is satisfied when all the tables are broken into as many tables as possible in order to avoid redundancy.

- 5NF is also known as Project-Join Normal form (PJ/NF).

Example :-

SUBJECT	LECTURER	SEMESTER
Computer	Anshika	Sem 1
Computer	John	Sem 1
Math	John	Sem 1
Math	Palvi	Sem 2
DBMS	Ram	Sem 1

So to make the above table into 5NF, we can decompose it into three Relations P1, P2 & P3.

P1

Semester	Subject
Sem 1	Computer
Sem 1	Math
Sem 1	DBMS
Sem 2	Math

P2

Subject	Lecturer
Computer	Anshika
Computer	John
Math	John
Math	Palvi
DBMS	Ram

P3

Semester	Lecturer
Sem 1	Anshika
Sem 1	John
Sem 1	John
Sem 2	[redacted] Palvi
Sem 1	Ram