NORMALIZATION OF DATABASE

BY

SAKTHIVEL.A

NORMALIZATION

- In database, Normalization is used to remove the duplicate data and database anomalies from relational table.
- It also helps us to reduce data redundancy.

TYPES OF ANOMALIES:

- Insertion anomalies
- Deletion anomalies
- Updation anomalies

EXAMPLE FOR ANOMALIES

Std_id	Std_Name	Dept	HOD	Office_No
1	AAA	CSE	Mr.JD	1234
2	BBB	CSE	Mr.JD	1234
3	CCC	CSE	Mr.JD	1234

- When we insert a new row to this table, we need to insert repeated values for it which an **insertion anomaly.**
- When we delete the student details from this table the department details will also get deleted along with the student details which is a **deletion anomaly.**
- If we want to change the HOD name then we have to change it in all the rows which is an **updation anomaly.**

FIRST NORMAL FORM

A table is in First Normal form if it satisfies the following conditions:

- Each column should contain atomic values.
- A column should contain values in same data type.
- Each column should have a unique name.
- Order doesn’t matter.

EXAMPLE

Roll_No	Name	Course
101	AAA	C,C++
102	BBB	Java
103	CCC	Python,Java

TABLE IN 1NF

Roll_No	Name	Course
101	AAA	C
101	AAA	C++
102	BBB	Java
103	CCC	Python
103	CCC	Java

SECOND NORMAL FORM

- Table should be in 1NF
- No partial dependency should be there in the table.

PARTIAL DEPENDENCY:

Partial dependency is means, a non primary attribute depends on a part of a primary key (candidate key).

EXAMPLE

Std_id	Subject_id	Marks	Teacher
101	11	82	Mr.A
101	12	85	Mr.B
102	11	90	Mr.A

In the above table, (Std_id + Subject_id) are primary keys and teacher column partially depends only on Subject_id column.

TABLE IN 2NF

Std_id	Subject_id	Marks
101	11	82
101	12	85
102	11	90

Subject_id	Teacher
11	Mr.A
12	Mr.B

THIRD NORMAL FORM

- Table should be in 2NF.
- No Transitive dependency is allowed.

TRANSITIVE DEPENDENCY

A non primary attribute depends on another non primary attribute is called Transitive dependency.

EXAMPLE

Std_id	Subject_id	marks	Exam_Name	Total_Marks
101	10	52	Theory	60
102	11	17	Practicals	20
103	12	18	Assignments	20

In the above table, Total_Mart column which is a non-primary attribute depends on another non-primary attribute which is Exam_Name.

TABLE IN 3NF

Std_id	Subject_id	marks	Exam_Name
101	10	52	Theory
102	11	17	Practicals
103	12	18	Assignmants

Exam_Name	Total_Marks
Theory	60
Practicals	20
Assignmants	20

BOYCE CODD NORMAL FORM

- Table should be in 3NF.
- For any dependency A -> B, A should be a super key.

NOT ALLOWDED

No primary attribute depends on a non-primary attribute

EXAMPLE

Std_id	Subject	Professor
101	Java	P.Java1
101	C++	P.CPP
102	Java	P.Java2

In the above table, the primary attribute Subject depends on a non-primary attribute Professor.

TABLE IN BCNF

Std_id	P_id
101	1
101	2
102	3

P_id	Professor	Subject
1	P.Java1	Java
2	P.CPP	C++
3	P.Java2	Java

FOURTH NORMAL FORM

- Table should be in BCNF.
- It should not have multi-valued dependency.

CONDITIONS OF MULTI-VALUED DEPENDENCY

- A -> B, For a single value of A, more than one value of B exist.
- Table should have atleast 3 columns.
- For a table with A, B, C columns, B and C should be independent.

EXAMPLE

Std_id	Course	Hobby
101	Science	Cricket
101	Maths	Hockey
102	C++	Crickey
102	PHP	Hockey

In the above table, Std_id have more than one value in the other two columns and the other two columns (Course and Hobby) are having no relationship.

TABLE IN 4NF

Std_id	Course
101	Science
101	Maths
102	C++
102	PHP

Std_id	Hobby
101	Cricket
101	Hockey
102	Crickey
102	Hockey

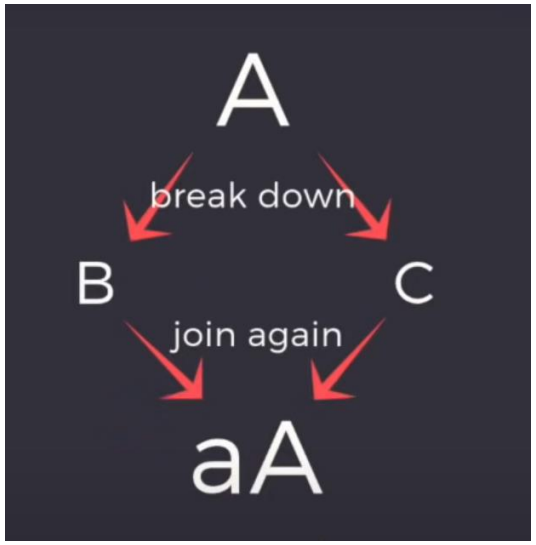
FIFTH NORMAL FORM

- Table should be in 4NF.
- It should not have join dependency.

JOIN DEPENDENCY

- If we break a table into two tables and again if we try to join that tables, if the joined table have a problems like
  - Additional informathon creation or
  - Information loss
- Then the table is having join dependency.

JOIN DEPENDENCY



EXAMPLE

Supplier	Product	Customer
S1	P1	C1
S1	P2	C2
S2	P3	C3
S2	P4	C4
S3	P1	C1

SPLITTING THE TABLE

Supplier	Product
S1	P1
S1	P2
S2	P3
S2	P4
S3	P1

Customer	Product
C1	P1
C2	P2
C3	P3
C4	P4

Supplier	Customer
S1	C1
S1	C2
S2	C3
S2	C4
S3	C1

PROBLEM

After splitting the table, if we try to get to know about an information “ S1 sells P1 to C1 “ it is losted.

THANK YOU