NORMALIZATION:(26-11-2024)
==========

What is Normalization ?
==========

- it is a technique which is used to decompose(i.e divide) a table data into multiple tables.

Where we use Normalization?
==========

- DB designing level.

Why Normalization ?
==========

EX: Branch_Student_Details

STID	SNAME	BRANCH	HOD	OFFICE_NUMBER
1021	smith	cse	Mr.x	040-22334455
1022	allen	cse	Mr.y	040-22334455
1023	ward	cse	Mr.x	040-22334455
1024	miller	cse	Mr.y	040-22334455
1025	jones	cse	Mr.x	040-22334455

Disadvantages:
==========

- Dataredundancy problem(duplicate data).
- It occupied more memory.
- Data inconsistency problem(i.e irregular data).
- Insertion problem.
- Updation problem.
- Deletion problem.

- To overcome the above problems we need to use a technique is known as "Normalization".

Solution:
==========

(pk) Branch_Details Student_Details (fk)
==========

```
======================
```

Bocde	Bname	HOD	Office_number	Stid	Sname	Bcode
1	cse	Mr.x	040-22334455	1021	smith	1
				1022	allen	1
				1023	ward	1
				1024	miller	1
				1025	jones	1
				1026	scott	1

```
======================
```

Advantages:

```
=============
```

- To avoid dataredundancy problem(i.e no duplicate data).
- It occupied less memory.
- To avoid data inconsistency problem(i.e regular data).
- To avoid Insertion problem.
- To avoid Updation problem.
- To avoid Deletion problem.

Types of Normaliztion forms:

```
========================
```

- First normal form(1NF)
- Second normal form(2NF)
- Third normal form(3NF)
- Boyce-codd normal form(BCNF)
- Fourth normal form(4NF)
- Fifth normal form(5NF)

First normal form(1NF):

```
==================
```

- For a table to be in the First Normal Form, it should follow the following 4 rules:
  1. Each column should contain atomic value (atomic = single value).
  2. A column should contain values that are same datatype.
  3. All the columns in a table should have unique names.
  4. The order in which data is stored, does not matter.

EX: Student_details

```
=======================
```

Stid	Sname	Bcode
1021	smith	2
1022	allen	1
1023	ward	3

27-11-2024:
==========
Second normal form(2NF):
======================

- For a table to be in the Second Normal Form, it must satisfy two conditions:
    1. The table should be in the First Normal Form.
    2. There should be no Partial Dependency.

WHAT IS DEPENDENCY:
==================

- IN A TABLE IF NON-KEY COLUMNS (NON PRIMARY KEY) ARE DEPENDS ON KEY COLUMN (PRIMARY KEY) THEN IT IS CALLED AS FULLY DEPENDENCY / FUNCTIONAL
DEPENDENCY.

(PK)
EX: STID SNAME BRANCH ADDRESS

- Here, “STID “IS A KEY COLUMN and “SNAME”,”BRANCH”,”ADDRESS” ARE NON-KEY COLUMNS.
- These non-key columns are linked with key column is STID.so that in this table there is no partial dependency columns.

WHAT IS PARTIAL DEPENDENCY:
==========================

- IN A TABLE IF NON-KEY COLUMN DEPENDS ON PART OF THE KEY COLUMN, THEN IT IS CALLED AS PARTIAL DEPENDENCY.

<PRIMARY KEY (stu_id,sub_id) / COMPOSITE PRIMARY KEY>
EX: STU_ID SUB_ID STU_MARKS TEACHER

- Here, “STU_ID and SUB_ID “ IS A KEY COLUMNS - ” STU_MARKS”,” TEACHER” ARE NON-KEY
COLUMNS. THEN “TEACHER” DEPENDS ON “SUB_ID” BUT NOT “STU_ID” COLUMN.
- Here we found a partial dependency column is "TEACHER" so that we need to do decompose a table like below,

Subject_Table
==========
(pk)
SUB_ID SUB_NAME TEACHER

Student_table
==========
(pk) (fk)
STU_ID STU_MARKS SUB_ID

Third normal form(3NF):
==========

- For a table to be in the third normal form there is two conditions.
    1. It should be in the Second Normal form.
    2. And it should not have Transitive Dependency.

TRANSITIVE DEPENDENCY:
==========

- IN TABLE IF NON-KEY COLUMN DEPENDS ON ANOTHER NON-KEY COLUMN, THEN IT IS CALLED AS TRANSITIVE DEPENDENCY.

EX:

|--------CPK----------|

STUDENT_ID	SUBJECT_ID	STU_MARKS	EXAM_NAME	TOTAL_MARKS

- Here, “STU_ID and SUB_ID “ ARE KEY COLUMNS . ” EXAM_NAME”,” TOTAL_MARKS”
ARE NON-KEY COLUMNS. THEN “TOTAL_MARKS” DEPENDS ON “EXAM_NAME” BUT NOT “STU_ID and SUB_ID” COLUMNS.
- Here we found transitive dependency columns are "EXAM_NAME" and "TOTAL_MARKS"
so that we need to do decompose the above table into multiple tabbles.

(pk) Exam_Table (cpk) Score_Table (fk)

EXAM_ID	EXAM_NAME	TOTAL_MARKS	STUDENT_ID	SUBJECT_ID	STU_MARKS	EXAM_ID

28-11-2024
==========
Boyce-codd normal form(BCNF):
==========

- For a table to satisfy the Boyce - Codd Normal Form, it should satisfy the following two conditions:
    1. It should be in the Third Normal Form.
    2. And, for any dependency A → B, A should be a super key.

SUPER KEY:
==========

- A COLUMN (OR) COMBNATION OF COLUMNS WHICH ARE UNIQUELY IDENTIFYING
A ROW IN A TABLE IS CALLED AS SUPER KEY.

CANDIDATE KEY:
==============

- A MINIMAL SUPER KEY WHICH IS UNIQUELY IDENTIFYING A ROW IN A TABLE IS CALLED AS CANDIDATE KEY.

(OR)

- A SUPER KEY WHICH IS SUBSET OF ANOTHER SUPER KEY, BUT THE COMBINATION
OF SUPER KEYS ARE NOT A CANDIDATE KEY.

EX:

STUDENT TABLE

STUDENT_ID	NAME	BRANCH	MAILID	REG_NUMBER

Super key columns:
==========

student_id		student_id + mailid		
mailid		mailid + reg_number		student_id + mailid + reg_number
reg_number		reg_number + student_id		

Candidate key columns:
==========

student_id
mailid
reg_number

Ex:

Professor Table

|----------cpk ----------|

PROFESSOR_ID	SUBJECT(B)	PROFESSOR(A)
1	java	p.java
2	java	p.java

- Here, PROFESSOR column depends on SUBJECT so that PROFESSOR should be

super key but not a super key.

- Now to make a PROFESSOR column is a super key and SUBJECT is non-super key column in the table like below,

Professor Table

|----------cpk----------|

professor_id	professor	Subject
1	p.java	java
2	p.java	java

5. Fourth normal form(4NF):

==========

- For a table to satisfy the Fourth Normal Form, it should satisfy the following two conditions:
  1. It should be in the Boyce-Codd Normal Form.
  2. A table does not contain more than one independent multi valued attribute / Multi Valued Dependency.

Multi valued Dependency:

==========

- In a table one column same value mapping with multiple values of another column is called as multi valued dependency.

EX:

COLLEGE ENROLLMENT TABLE (5NF)

STUDENT_ID	COURSE	HOBBY
1	ORACLE	Cricket
1	JAVA	Reading
1	C#	Hockey

Mapping with multiple values of columns: (Decomposing table)

==========

Course_details (4NF)

==========

STUDENT_ID	COURSE
1	oracle
1	java

Hobbies_details(4NF)

==========

STUDENT_ID	HOBBY
1	cricket
1	reading

1	c#	1	hockey

Fifth Normal Form (5NF):

==========

- If a table is having multi valued attributes and also that table cannot decomposed into multiple tables are called as fifth normal form.

EX:

COLLEGE ENROLLMENT TABLE (5NF)

STUDENT_ID	COURSE	HOBBY
1	ORACLE	Cricket
1	JAVA	Reading
1	C#	Hockey

==========