Normalization

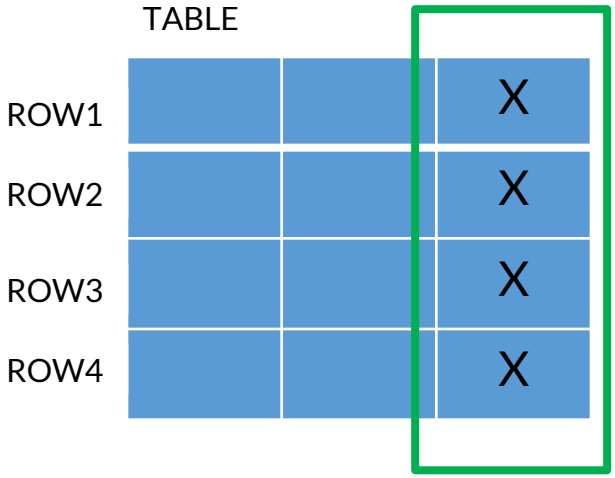
In this session, you will learn:

- Problems without Normalization
- How Normalization solves these problems

What is Normalization

- A technique of organizing the data into multiple related tables, to minimize data redundancy.

What is Data Redundancy?



- Repetition of data increases the size of database
- Other issues like
  - Insertion problems
  - Deletion problems
  - Updation problems

Why should we reduce it?

STUDENTS table

RollNo	Name	Branch	HOD	Dept_Tel
1	ABC	CSE	Mr. X	53337
2	XYZ	CSE	Mr. X	53337
3	PQR	CSE	Mr. X	53337
4	MNO	CSE	Mr. X	53337

These anomalies occur when it is not possible to insert data into a database because the required fields are missing. (field restricted to have no null constraint)

Insertion Anomaly

Reason for Data Repetition

- Two different but related data is stored in the same table

RollNo	Name	Branch	HOD	Dept_Tel
1	ABC	CSE	Mr. X	53337
2	XYZ	CSE	Mr. X	53337
3	PQR	CSE	Mr. X	53337
4	MNO	CSE	Mr. X	53337

STUDENTS table

RollNo	Name	Branch	HOD	Dept_Tel
1	ABC	CSE	Mr. X	53337
2	XYZ	CSE	Mr. X	53337
3	PQR	CSE	Mr. X	53337

STUDENTS table

RollNo	Name	Branch	HOD	Dept_Tel
1	ABC	CSE	Mr. X	53337

← Deleting this last student record

Branch information deleted along with Student data

Deletion Anomaly

STUDENTS table

RollNo	Name	Branch	HOD	Dept_Tel
1	ABC	CSE	~~Mr. X~~ Mr.Y	53337
2	XYZ	CSE	~~Mr. X~~	53337
3	PQR	CSE	~~Mr. X~~	53337
4	MNO	CSE	~~Mr. X~~	53337

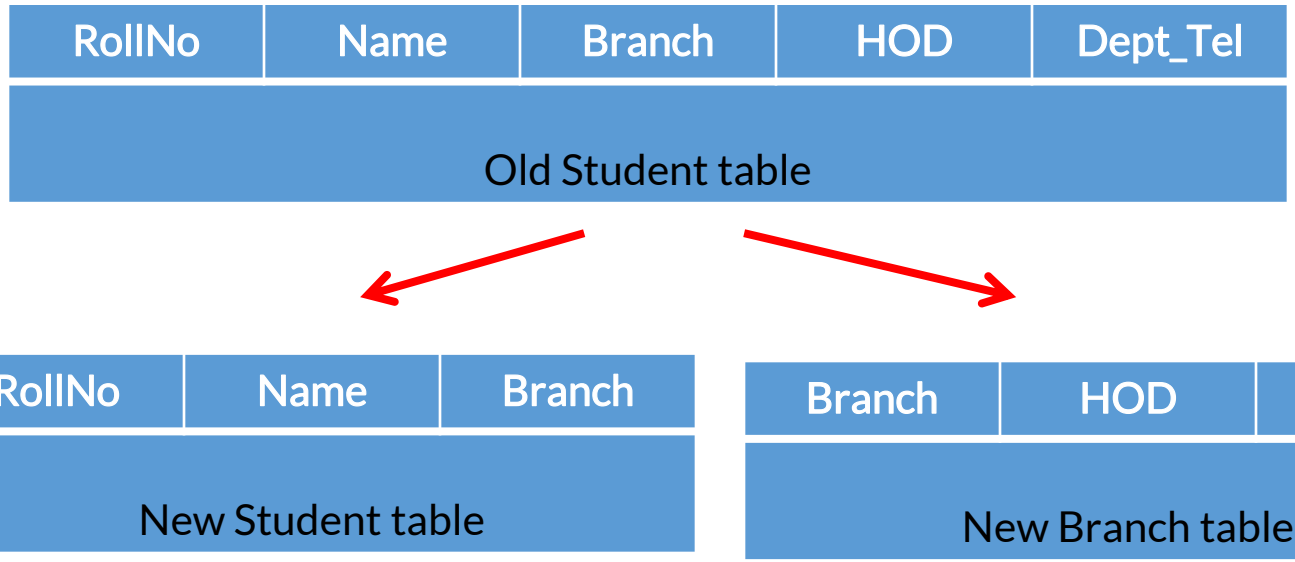
Updation Anomaly

It occur when the same data is repeated in multiple rows, and changes are made in some but not all instances

How Normalization will solve all these problems?

Student table

Student table + Branch table



Student table

RollNo	Name	Branch
1	ABC	CSE
2	XYZ	CSE

Branch table

Branch	HOD	Dept_Tel
CSE	Mr. Y	53337

Not Eliminating Redundancy

But Minimizing Redundancy

Student table

RollNo	Name	Branch
No data		
2	XYZ	CSE
3	PQR	CSE

Deleting this last student record

Branch table

Branch	HOD	Dept_Tel
CSE	~~Mr. Y~~ Mr.Y	~~53337~~43337

THANKS