



Normalization

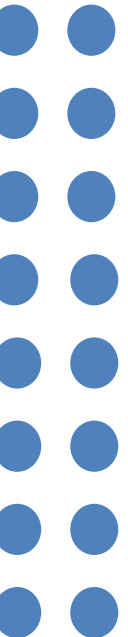


LESSON OBJECTIVES

- 
- Understand Normalization
 - Normalize tables efficiently in MySQL
- 

What is Normalization?

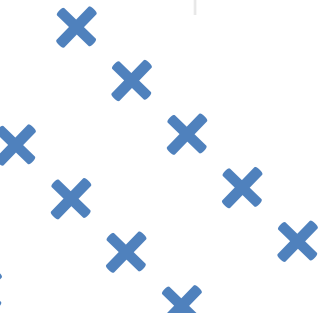
- Normalization is the process of organizing data within a database (relational database) to ensure data is stored logically.
- To organize data so that there are no anomalies.(Update, Delete, Insert)



What is a Key?

- A KEY is an attribute used to uniquely distinguish a record of a table(A unique identifier).
- A KEY could be an individual column or a combination of multiple columns.

Primary Key -->	id	roll_no	student_name	city	dept_id	<---Foreign Key
	1	22	Matthew	Cape Town	10	
	2	23	Matthew	Durban	10	
	3	24	Joel	Johannesburg	20	
	4	25	Candice	Cape Town	20	
	5	26	Ryan	Pretoria	25	



Types of Keys

- Primary Key
- Foreign Key
- Super Key
- Candidate Key



Primary Key

Primary Key



- is a set of one or more fields/columns of a table that uniquely identify a record in a database.
- It can not accept null or duplicate values.

	Student Table					
Primary Key -->	id	roll_no	student_name	city	dept_id	<---Foreign Key
	1	22	Matthew	Cape Town	10	
	2	23	Matthew	Durban	10	
	3	24	Joel	Johannesburg	20	
	4	25	Candice	Cape Town	20	
	5	26	Ryan	Pretoria	25	

Department Table	
10	Sound Engineering
20	Civil Engineering
25	Chemical Engineering
30	Electrical Engineering

Foreign Key

Foreign Key



- is a field in a database table that is a Primary Key of another table.

	Student Table					
Primary Key -->	id	roll_no	student_name	city	dept_id	<---Foreign Key
	1	22	Matthew	Cape Town	10	
	2	23	Tarran	Durban	10	
	3	24	Joel	Johannesburg	20	
	4	25	Candice	Cape Town	20	
	5	26	Ryan	Pretoria	25	

Department Table	
dept_id	dept_name
10	Sound Engineering
20	Civil Engineering
25	Chemical Engineering
30	Electrical Engineering

Super Key

Super Key



- is a set of one or more than one key that can be used to identify a record uniquely in a table.
- can have extra attributes that are redundant for distinct identification

	Student Table					
Primary Key -->	id	roll_no	student_name	city	dept_id	<---Foreign Key
	1	22	Matthew	Cape Town	10	
	2	23	Tarran	Durban	10	
	3	24	Joel	Johannesburg	20	
	4	25	Candice	Cape Town	20	
	5	26	Ryan	Pretoria	25	

{roll_no, student_name, city}
{id, city}
{id, student_name}
{roll_no, dept_id}
{roll_no, student_name}

Candidate Key

Candidate Key

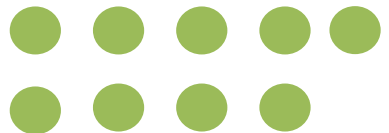
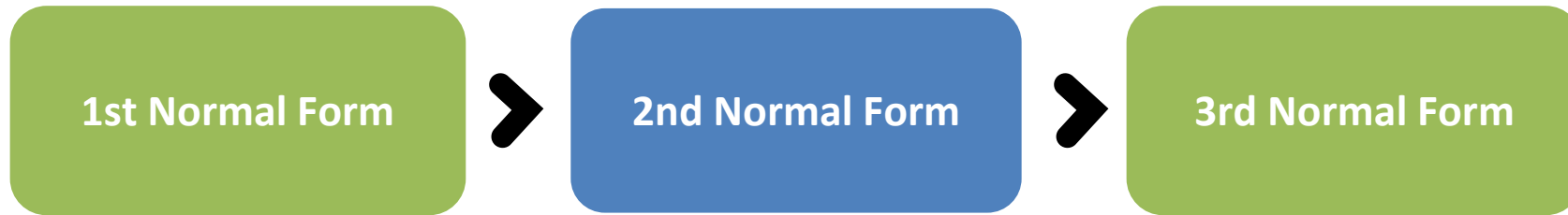


- is the minimal super key that can identify a record uniquely in a table.
- Each Candidate Key can work as a Primary Key

	Student Table					
Primary Key -->	id	roll_no	student_name	city	dept_id	<---Foreign Key
	1	22	Matthew	Cape Town	10	
	2	23	Tarran	Durban	10	
	3	24	Joel	Johannesburg	20	
	4	25	Candice	Cape Town	20	
	5	26	Ryan	Pretoria	25	

{roll_no, student_name, city}
{id, city}
{id, student_name}
{roll_no, dept_id}
{roll_no, student_name}

Types of Normal Forms



1st Normal Form

Step 1



- Each table cell should contain a single value.(Not a multivalued attribute)
- Each record should be unique.

id	student_name	course
22	Matthew	MENG1,CENG1
24	Joel	CENG1
25	Candice	SENG1
28	Ryan	EENG1



{PK}id	student_name	course
22	Matthew	MENG1
22	Matthew	CENG1
24	Joel	CENG1
25	Candice	SENG1
28	Ryan	EENG1

2nd Normal Form

Step 2



- Table should be in First Normal Form
- Columns should only depend on the primary key.
- It should not have a partial dependency.
 - If the field of any candidate key determines the non-prime attributes, it is called a partial dependency.

lecturer_id	course_name	course_price
I_1	Mathematics	R6000
I_5	Literature	R5000
I_3	Social Studies	R4000
I_3	Bible Studies	R3000
I_5	Theology	R3000
I_2	Astrophysics	R9000

3rd Normal Form

Step 3



- Table should be in Second Normal Form.
- It should not have any transitive dependencies.(If a field is dependent on any other column other than the primary key of the table)

lecturer_id	course_name	course_price	lecturer_city	lecturer_zipcode
I_1	Mathematics	R6000	Cape Town	7300
I_5	Literature	R5000	Johannesburg	2000
I_3	Social Studies	R4000	Pretoria	0100
I_3	Bible Studies	R3000	Pretoria	0100
I_5	Theology	R3000	Johannesburg	2000
I_2	Astrophysics	R9000	Durban	4000