

## Agenda

- 1) Select & query functions. (continued) ✓
- 2) Joins
- 3) Aggregates.

## Announcement

20th October → Holiday  
19th Oct. → Wed.  
18th Oct → Tues.

} Class.

## Students table

Name, LName, batchId, psp

Get all students of batch-id = 2 where psp  $\geq 60$  and psp  $\leq 90$

Send your answers.

Between → When you have statements like  $a \geq x$  AND  $a \leq y$

Select \* from students where batch-id = 2 AND

(psp Between 60 and 90)

Note:- Between is  $\geq$  and  $\leq$ . inclusive.

## # LIKE Operator

Whenever we compare strings we have diverse use cases.

e.g. Scales Batches table

Let's say you have only this table.

→ Get all the batches that started in August.

### Solution:-

- 1) % (percent)
- 2) \_ (underscore)

Select \* from batches

where name LIKE 'hello';

→ Gets me all the rows with name = 'hello'  
It does exact match.

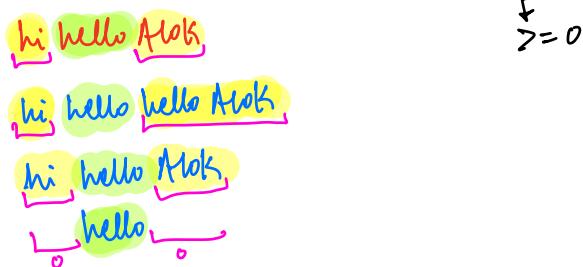
Batches.

id	Name
1	August Intermediate 2022
2	Aug Adv 22
3	May - -
4	June - -
5	July - -

e.g. 1) select \* from batches

where name LIKE '%hello%';

% means any # of characters.



$$\downarrow \\ \geq 0$$

2) \_ means any single char.

name like '\_ hello %'

'hello I am' → No.

'hi hello AloK' → No.

'I hello AloK' → Yes  
↑  
one char.

\* The most common usage of LIKE

→ find all containing ('x') in all the data.

→ select \* from table where column like '%x%'!

## # IS NULL.

IS NULL → If you have to check the value of a column to be null.

Don't say  $x = \text{NULL}$ . Rather use IS NULL

## # ORDER BY

By default, data sorted in ascending according to PK.

Query: Get all students where .batch-id=2 and sort by psp in descending order.

Note:- You can orderBy single/multiple columns.

Select \*  
from students  
orderBy (psp, timeOfSolving, fName)

	psp	time of solving (min)
Karthick	90	30
Srikanth	90	30

- \* By default, orderBy sorts in ascending order.
- \* If a tie happens after all columns mentioned in order by clause, the tie breaker happens with PK.
- \* To sort in descending order.

Select \*  
from students  
orderBy (psp) DESC

## # LIMIT

Select \* from students → first 5 students based on PK.  
Limit 5;

Select \* from students  
orderBy psp  
Limit 5 → bottom 5 students by psp

Select \* from students  
orderBy psp desc  
Limit 5 → top 5 students by psp

## # JOINS

### Scalier DB.

Students

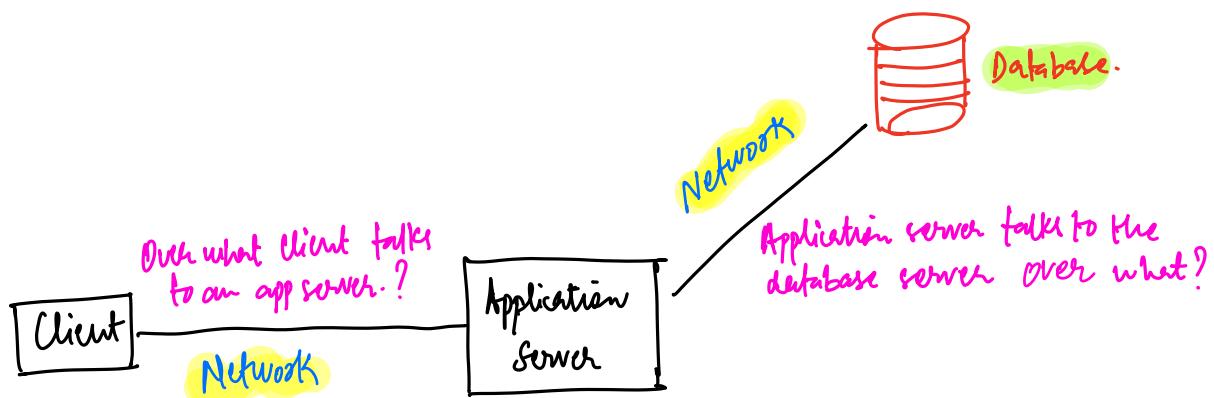
id	Name	email	password	batch-id

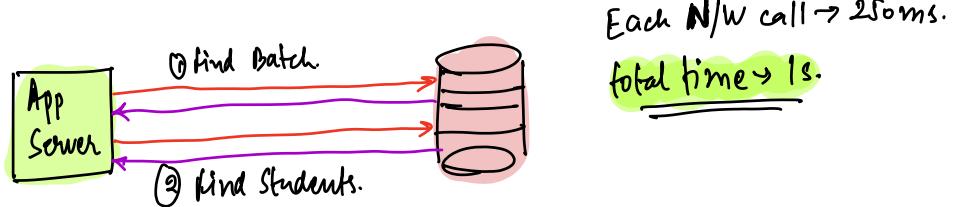
Batches

id	Name

Ques:- find all the students in 'Aprat Intermediate' Batch.

Solution:- a) Go to batches table and find the batch-id with given name.  
b) Find all the students with the batch-id. in students table.





Stu id.	st. Name	st. Email	st.Batch Id	st. Batch Name.

**Bad Schema Table (Denormalized)**

Issue:- Redundancy.

Benefit:- Querying will be faster.

Break:- 8:40.

We have these 2 tables.

<u>Students</u>				
<u>id</u>	<u>Name</u>	<u>email</u>	<u>password</u>	<u>batch-id</u>
1.	Alok	alok@gmail.com	123	1
2.	Sumit	sumit	456	1
3.	Manju	manju	8910	1
4.	Neha	neha	Abc	2
5.	Rahul	rahul	naamjunaHoga	1

<u>Batches</u>	
<u>id</u>	<u>Name</u>
1	Dec 21
2	Feb 22

Students }  
JOIN  
BATCHES.

<u>Students</u>		
<u>id</u>	<u>Name</u>	<u>batch-id</u>
1.	Alok	1
2.	Sumit	1
3.	Manju	1
4.	Neha	2
5.	Rahul	1

<u>Batches</u>	
<u>id</u>	<u>Name</u>
1	Dec 21
2	Feb 22

JOIN

<u>id</u>	<u>Name</u>	<u>st.-batch-id</u>	<u>batch-id</u>	<u>batch-name</u>
1.	Alok	1	1	Dec 21
2.	Sumit	1	1	Dec 21
3.	Manju	1	1	Dec 21
4.	Neha	2	2	Feb 22
5.	Rahul	1	1	Dec 21

# col = # nof cols in students + # cols in batches.

Select \* from 'this joined table'  
where batch-name = 'Dec 21'

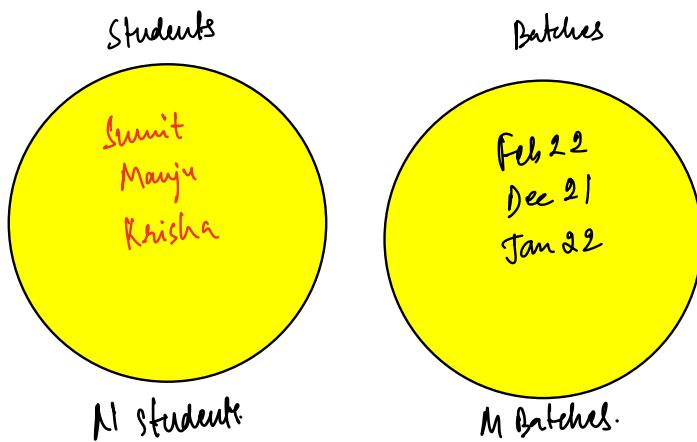
## # Types of Joins

i) Inner Join → Combines rows of both tables based on given condition.

Students		
id	Name	batch-id

Batches	
id	Name

Select \*  
from students  
inner join batches  
on students.batch-id = batches.id;  
where batches.name = "Dec 21"



for every row of students:  
for every row of batches:  
if the join condition is satisfied:  
pick and put this merged  
row in o/p

$$T.C = O(N \times M)$$

Let's take an example.

Students table

id	name

Phone numbers table

st-id	phone-num

Q) Print all the students with all their phone nos.

select name, phone-num

from students

inner join phone\_numbers

on students.id = phone\_numbers.std-id.

join

## 2) Self Join

Students.		
id	name	peer-reviewer-id
1	Manju	4
2	Krishna	4
3	Rahul	1
4	Aksha	2

PeerReviewers Table.

id	Name



Q) Print the name of every student with the name of their peer.

Q1 format →

id	name	peer-name

this is how I want the Q1.

```
Select *
from Students
join PeerReviewer p
on s.peerId = p.id
```

```
Select id, s.name, p.name
from Students s
join Students p
on s.peerId = p.id
```

Self join is nothing but inner join with same table.

Students.		
id	name	peer-reviewer-id
1	Manju	4
2	Krishna	4
3	Rahul	1
4	Aksha	2

Students.		
id	name	peer-reviewer-id
1	Manju	4
2	Krishna	4
3	Rahul	1
4	Aksha	2



Do a inner join b/w these 2 tables  
(Same table) → self join.

### 3) Compound Joins

Joining multiple tables.

Scalar DB

Students

id	Name	batch-id

batches

id	ins-id

instructors.

id	Name

Ques Print st.Name, ins.Name for every student

Ques

Students

<u>id</u>	<u>Name</u>	<u>batch-id</u>
1.	Alok	1
2.	Sunmit	1
3.	Manju	NULL
4.	Neha	2
5.	Rahul	1

Recently  
joined -  
No batch.

Batches

<u>id</u>	<u>Name</u>
1	Dec 21
2	Feb 22
3	Jan 22

Print the name of every student with the name of the batch.

If a student is not present in any batch, print their name with NULL as the batch name.

## # Outer Joins

- left Joins
- Right Joins
- full Joins.

1) **left Join** → If a row of left side doesn't match with any row of right side, it will be added in final answer with NULL.

Students

id	Name	batch-id
1.	Alok	1
2.	Sumit	1
3.	Manju	NULL
4.	Neha	2
5.	Rahul	1

Batches

id	Name
1	Dec 21
2	Feb 22
3	Jan 22

Recently joined -  
No batch.

↓ **left Join** these 2 tables.

id	Name	St.-batch-id	batch-id	batch-name
1.	Alok	1	1	Dec 21
2.	Sumit	1	1	Dec 21
3.	Manju	NULL	NULL	NULL
4.	Neha	2	2	Feb 22
5.	Rahul	1	1	Dec 21

2) Right Join → All rows of right side are returned even if they do not match with any row of left side

The diagram shows two tables: Students and Batches. The Students table has columns **id**, **Name**, and **batch-id**. It contains 5 rows with data: (1, Atok, 1), (2, Summit, 1), (3, Manju, NULL), (4, Neha, 2), and (5, Rahul, 1). A yellow circle highlights the **batch-id** column. A pink arrow points from the text "Recently joined - No batch." to the cell containing **NULL**. The Batches table has columns **id** and **Name**. It contains 3 rows with data: (1, Dec 21), (2, Feb 22), and (3, Jan 22). Below the tables is a blue box containing the text "Right Join b/w these 2 tables". An arrow points from this box down to the resulting table.

<u>Students</u>	<u>Batches</u>			
<u>id</u>	<u>Name</u>	<u>batch-id</u>	<u>id</u>	<u>Name</u>
1.	Atok	1	1	Dec 21
2.	Summit	1	2	Feb 22
3.	Manju	NULL	3	Jan 22
4.	Neha	2		
5.	Rahul	1		

<u>id</u>	<u>Name</u>	<u>st-batch-id</u>	<u>batch-id</u>	<u>batch-name</u>
1.	Atok	1	1	Dec 21
2.	Summit	1	1	Dec 21
4.	Neha	2	2	Feb 22
5.	Rahul	1	1	Dec 21
NULL	NULL	NULL	3	Jan 22

3) FULL JOIN

If any row of any side is not present put that with a **NULL**.

Thursday

- Complete Joins
- Start Aggregates and Built-in functions.