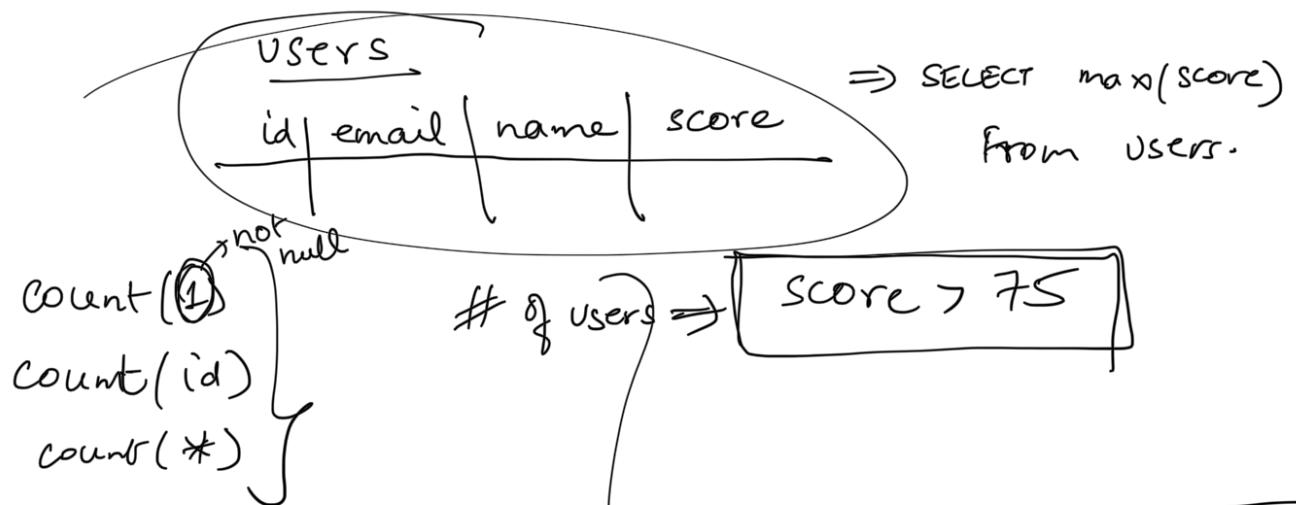


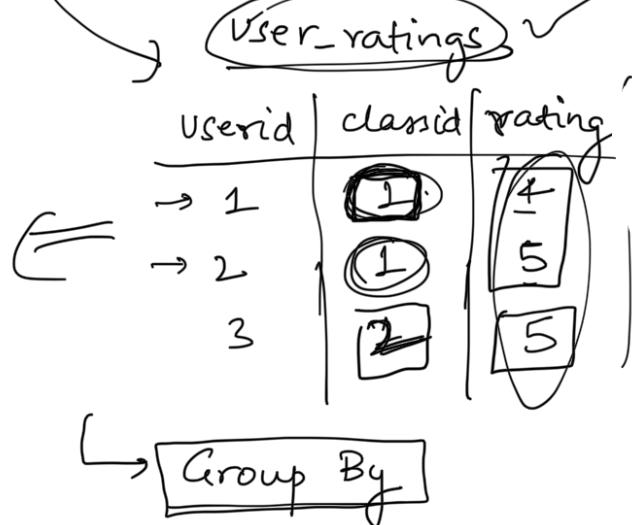
AGGREGATION & SUB-QUERIES

Group By

avg, sum, count, max, min

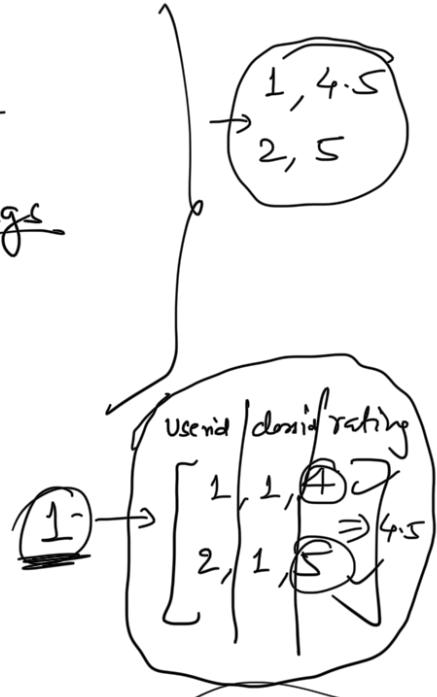


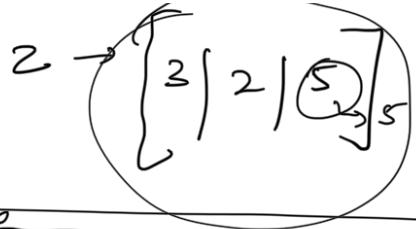
class_id	Avg rating
1	4.5
2	5



```
SELECT
    class_id, rating
FROM user_ratings
```

→ GROUP BY class_id





Users		
id	email	name
1		

Userid	attn %
1	0.5 \Rightarrow 50%
2	0 \Rightarrow 0%
3	1 \Rightarrow 100%

SELECT
(Userid, sum(attended)/count(attended))
 FROM user_classes
 GROUP BY User_id

Userid	classid	attended
1	1	0
1	2	1
1	3	1
1	4	0
2	1	0
2	2	0
2	2	0
2	4	0
3	1	1
3	2	1
3	2	1
3	2	1

film		
id	title	...

actor		
ia	name	...

film-actor		
film-id	actor-id	...

① Figure out tables which have info.

①.2 JOIN tables so that info

② Filter

③ Group by ?

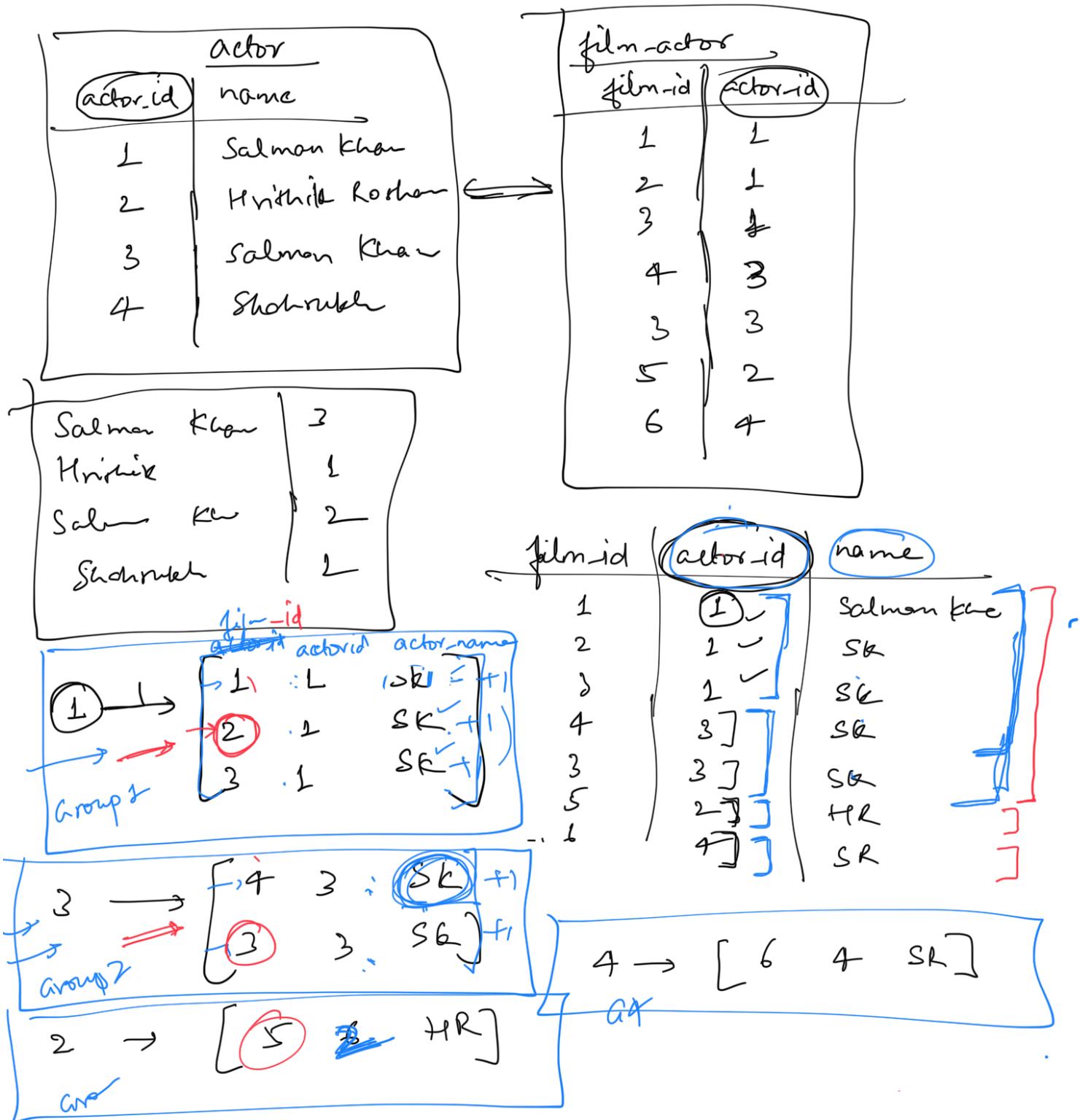
④ Print

actor-name	number-of-film

Group by

actor-name
actor-id

⑤ Limit, offset, order by -



SELECT
 (Actor-name), count(1)

SK, 3 col x

Group by colx

com	C
1	A
2	

Companies - Users

group by user-id

15

user-id	company-id	round-date	interview-round
1	1	2023-07-01	1
1	2	2023-07-03	1
1	1	2023-07-05	0
1	2	2023-07-02	0
2	2	2023-07-03	1
2	2	2023-07-08	1
2	2	2023-07-04	1
2	2	2023-07-08	1

User , company → # of rounds .

user-id	cid	# of rounds
1	1	3
1	2	2
2	2	4

Subqueries

employees

id	name	dept	salary
1	A	HR	1000
2	B	HR	100
3	C	HR	500
4	D	Tech	200
5	E	Tech	300
6	P	Tech	400

print id, name
for everyone
whose salary is
more than
avg salary in
the company

Step 1: calculate avg salary in the company

SELECT avg(salary) FROM employees.

— SELECT id, name
FROM employees
WHERE salary >

Subquery
↓
SELECT avg(salary) FROM employees

1. Choose tables

1.2 Do join if required

2. Apply filters (WHERE)

3. Group By.

3.1 Having ← group aggregate

④ Print ← table aggregate

⑤ Order by, limit, offset

Step 2: Calculate dept, avg(salary) in dept.

dept	avg-sal.
HR	533
Tech	300

SELECT

e.id, e.name

FROM employees e

JOIN

SELECT dept, Avg(salary) AS dept_avg
FROM employees GROUP BY dept) AS

tmp-table

ON (e.dept = tmp-table.dept
AND e.salary > tmp-table.
dept_avg)

SELECT *

FROM

(
= .
= .
) AS fancy-name
CROSS ...

Users				
id	email	name	batch_id	psp
1			1	80
2			1	39
3			1	100
4			1	75
5			2	60
6			2	70
7			2	90
8			2	55

batch_id = 1

batch_id = 2

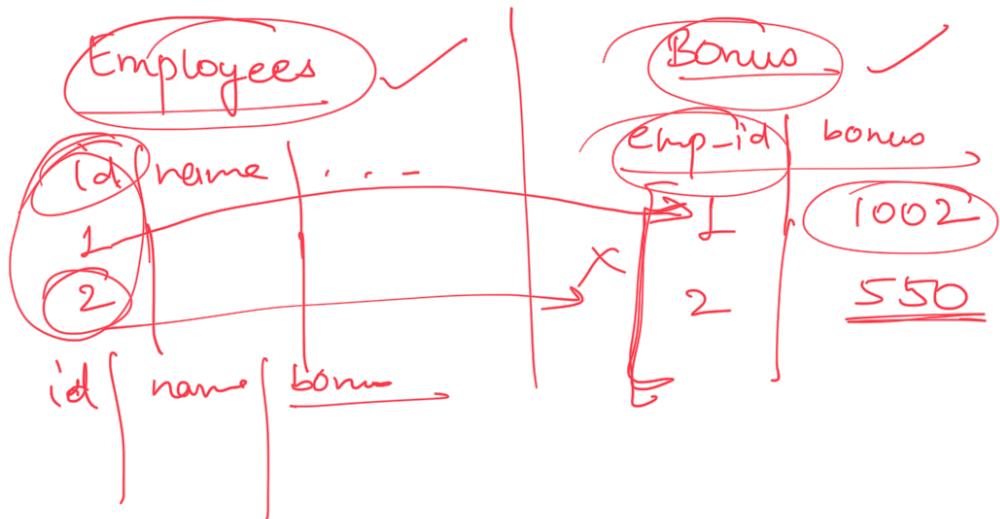
email of students in batch_id = 1 whose psp > max-psp in batch_id 2

SELECT
MAX(psp)
FROM users
WHERE batch_id = 2

⇒ max-psp in batch 2

SELECT email
FROM users
WHERE batch_id = 1

AND $p < p$ \Rightarrow (SELECT *
 WHERE max(psp) FROM users
 batch_id = 2)



Employees e

LEFT JOIN bonus b

ON e.id = b.emp_id
 AND b.bonus > 1000

1	abcde	1000
2	efghi	NULL

JOIN

