

Subqueries, views and query optimisation

- ① Sub-queries
 - correlated
 - ② Views
 - implementation
 - ③ Query optimisation
 - life cycle of a query
 - EXPLAIN
 - execution plan
 - Scans
-

① Window functions

② Stored procedures

③ Triggers

④ Advanced — shadow paging
— WAL

⑤ privileges

HLD — no-sql
— sharding

OS → Locking

Sub queries

→ I want all the students
whose IQ is > Average IQ

S ID	Name	IQ
1	A	120
2	B	150
3	C	150

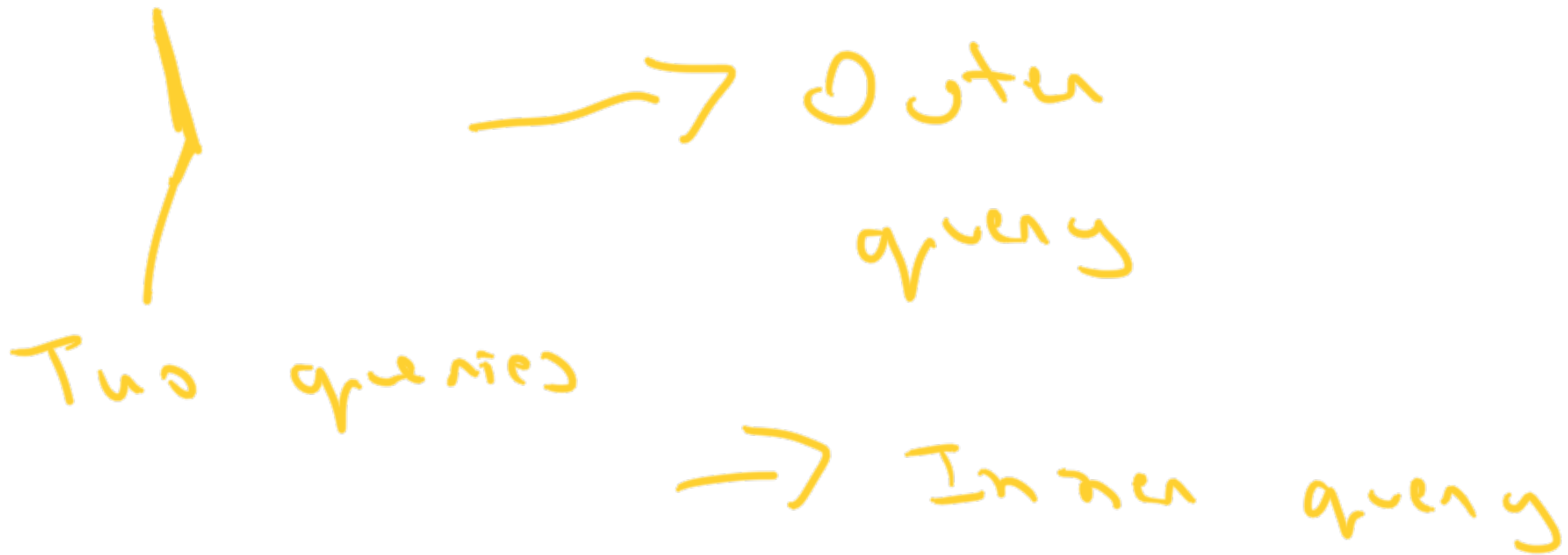
140

Select * From students where IQ > (select avg(IQ) from students)

SELECT ... FROM ...
WHERE IQ > AVG(IQ) * X

- ① Get average IQ
- ② use average IQ in where clause

Sub queries



② Select * from students
where IQ >

(Select AVG(IQ)
FROM students;
);

① Allows you to run multiple SQL
statements at once

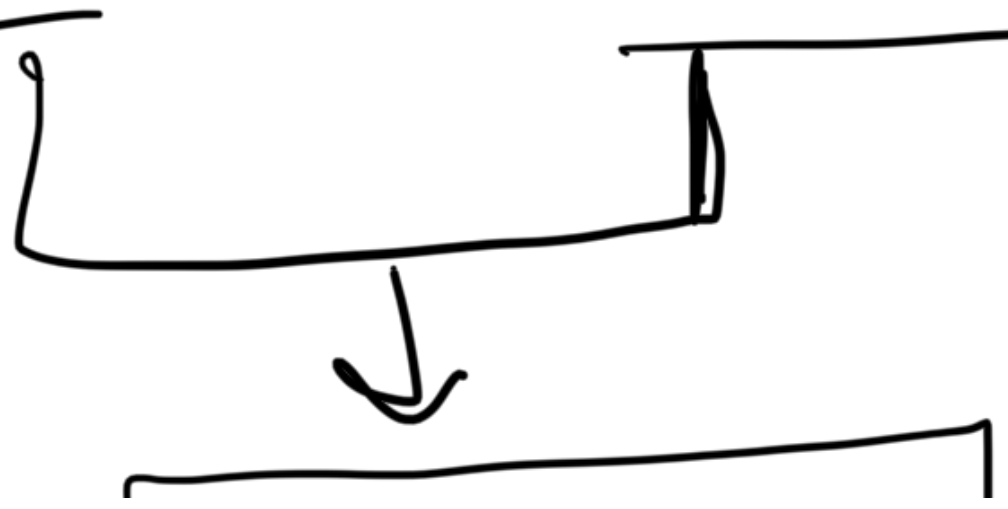
questions a...

② Intuitive + easy to read

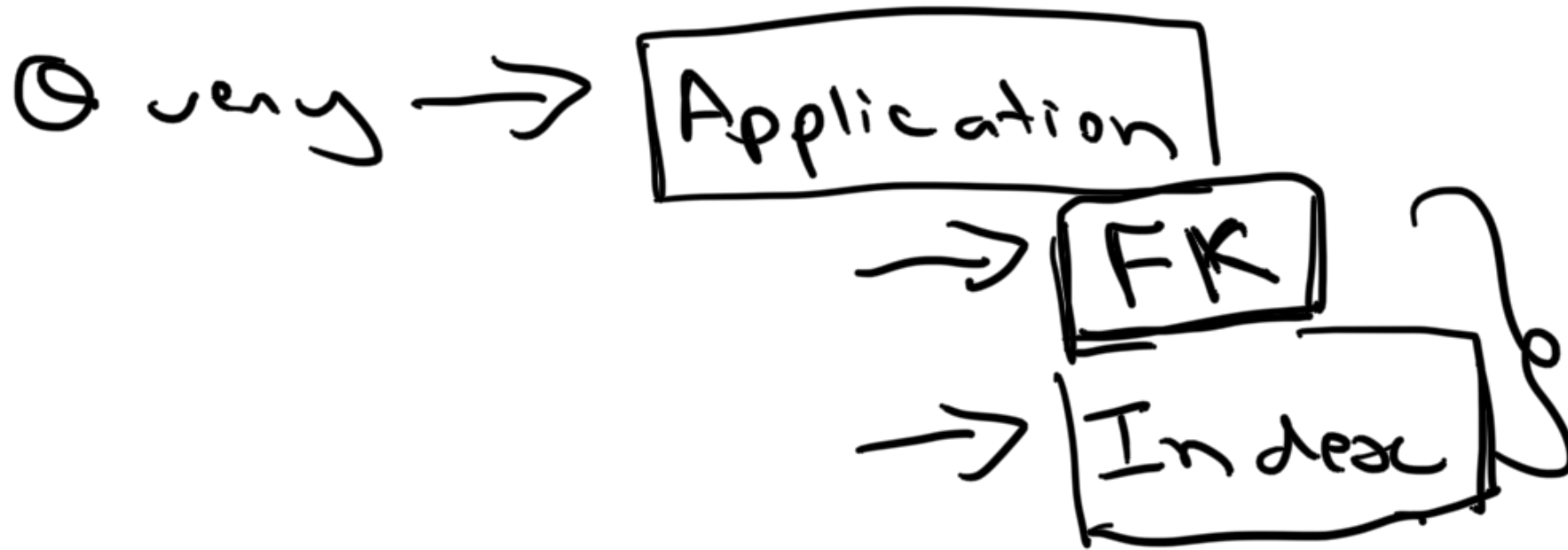
5:57 - 6:02
= 10:32

Students

Batches



| batch-id | → (FK)



→ Adhoc → no need if not frequent

Student S → id
→ email
→ phone } → in dears

APN → New relc
→ sentry
→ SLOW queries }

Join → 10 times a day

when → 10 00

Student →

name, phone

Where



name = ?

AND

phone = ?

↑

phone = ?

AND

name = ?

↓

(Student up here)



Student

phone X

Student

phone = null

email, phone, name

email, phone

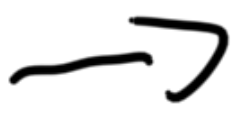
email

Views

D at a base



D ata



analysts

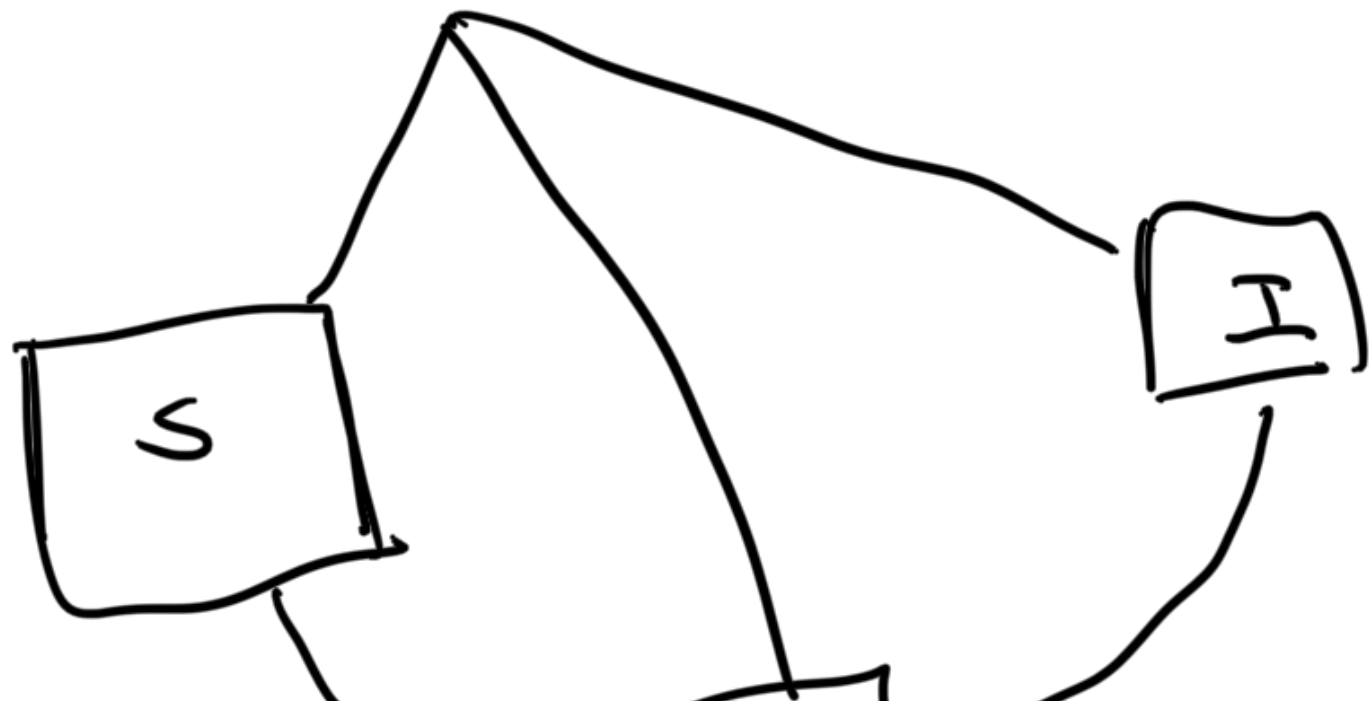


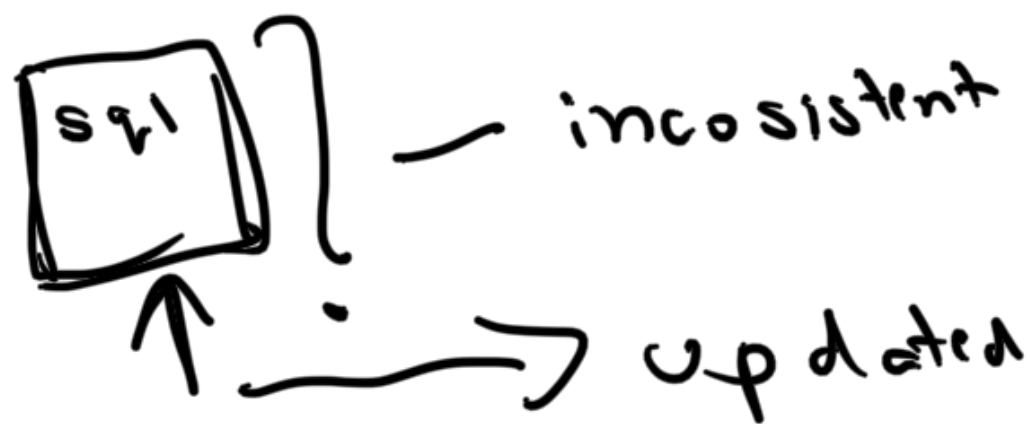
data scientists



normalized

Excel





② Security / abstraction

→ IF I want to get data
I need to know about the

structure / schema.

③

Denormalised

→ easy to work with

Student - batch

↳ anomalies ✓

↳ inconsistency ✓

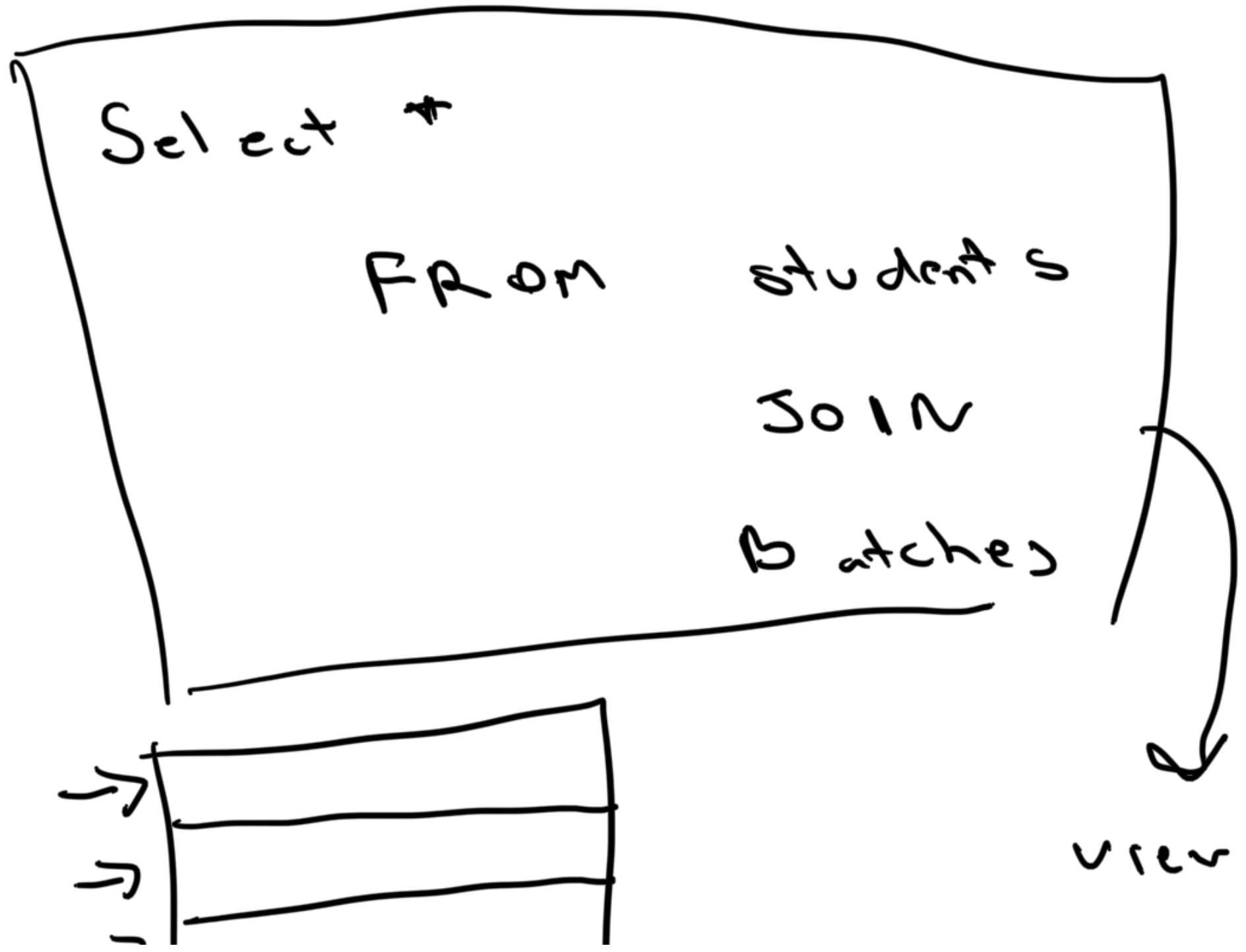
↳ data duplication

→ waste of space

View

View

↳ a temp. place to store the results of a query





View → abstraction
over your query

Student_batch

query

Select * from student_batch;



JOIN ...
batches

ALIAS

FOR A QUERY

No need to

① Rewrite the query

② Hide the complexity

③ Security → View → Third party

④ Maintainability is improved

→ s_name, batch_name → batch

→ UPDATE VIEW
↓
RECREATE

Developer



Creates the table

Data Analyst



Creates the query



Store query in a file

Data analyst

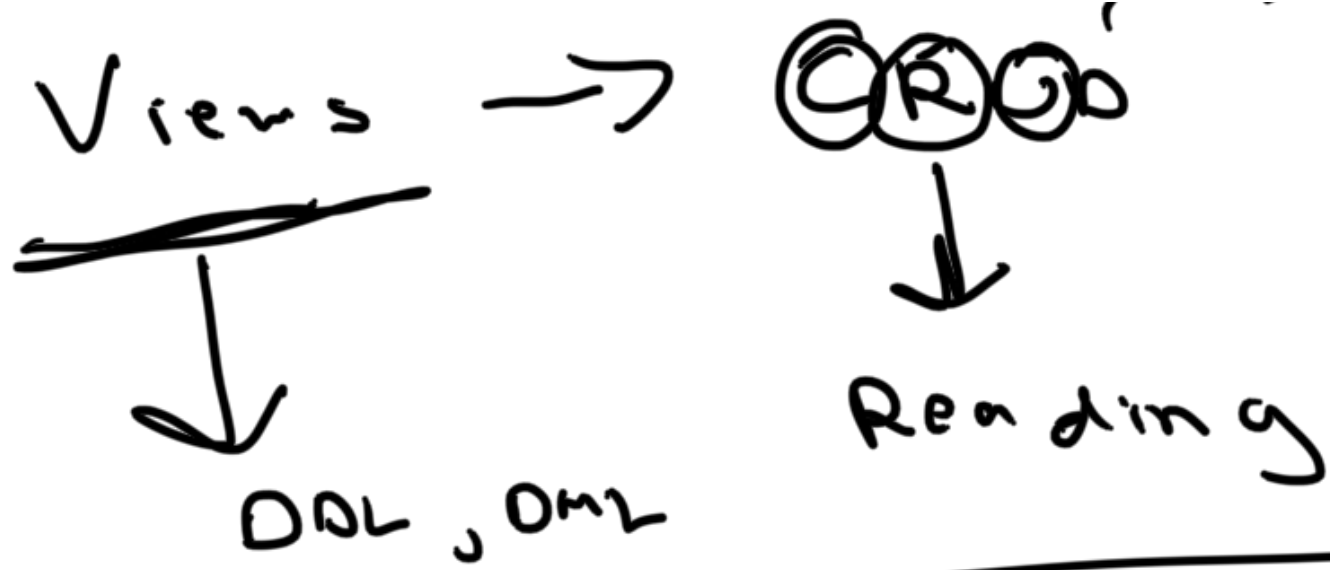
↳ Create the query

Developer

↳ Maintain the database

→ abstraction for a query
→ result set from the query

↳ limitations



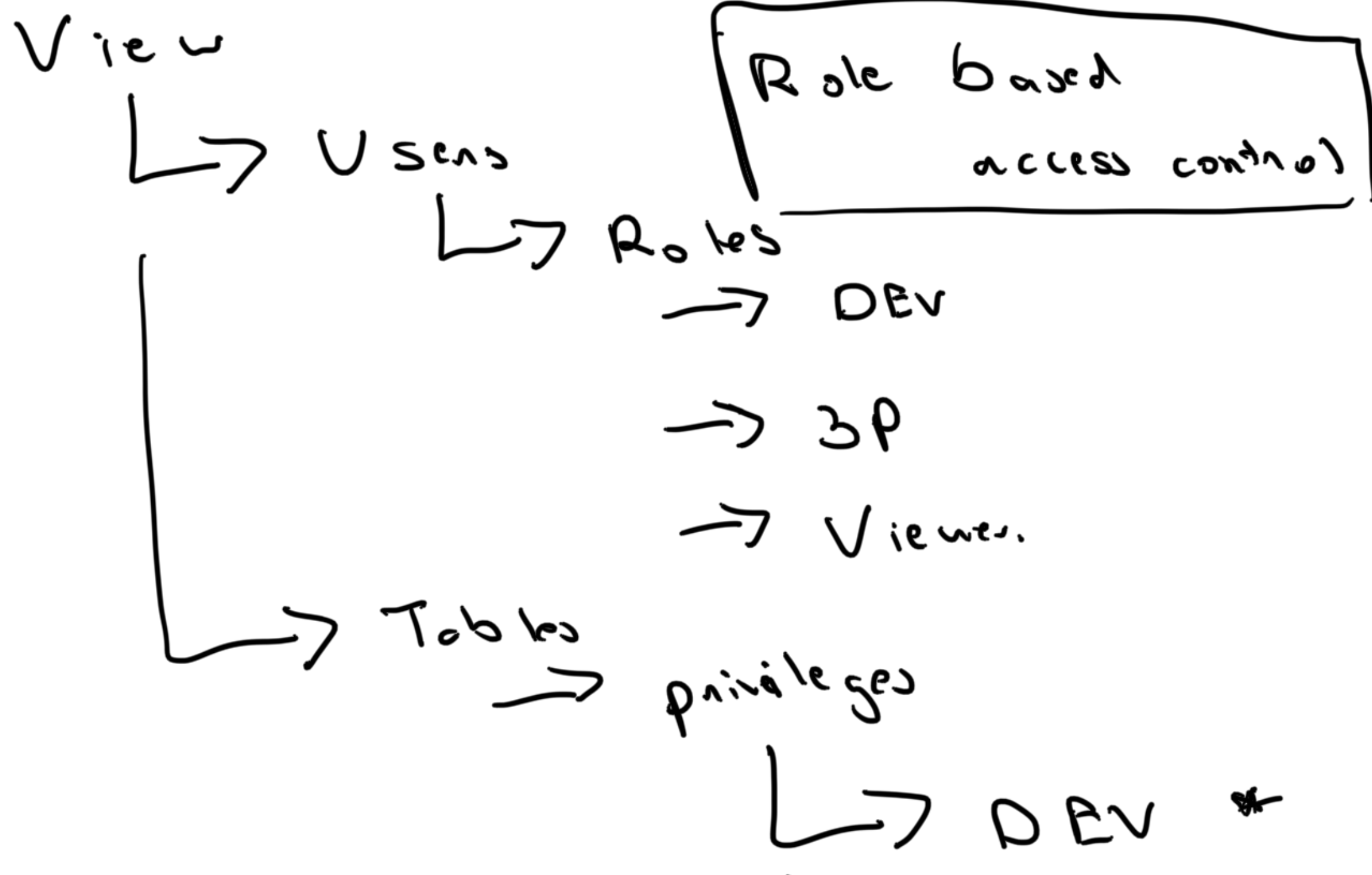
↻
CREATE VIEW NAME AS

Select _____
FROM _____
WHERE _____

;



Select * from view_name



↳ 3P VIEWS

* Normalised

* NOT easy to work with

* JOIN

Student	batch

* Run queries every time X

* How do we maintain? X



Out of sync

→ tied to



enum

a version of
schema

* Denorm + table

→ waste of space

→ maintain the table

→ inconsistency

Views

Application

DB



①

Filtered

rows

100

②

pagination

①

WHERE

CROSS JOIN

WHERE

②

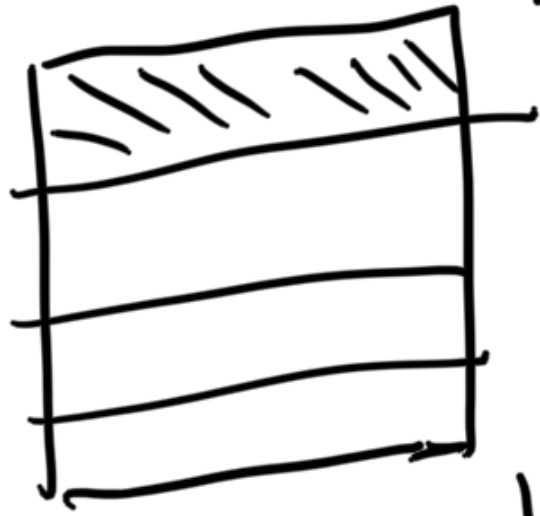
Sub queries

FROM



WHERE

Join

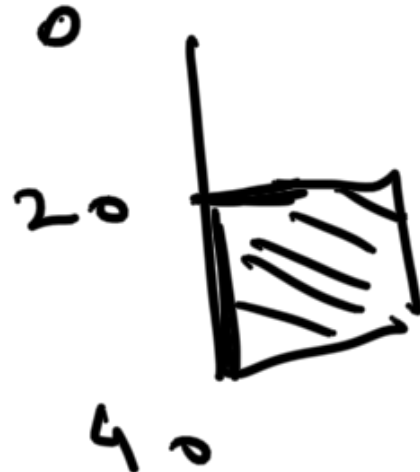


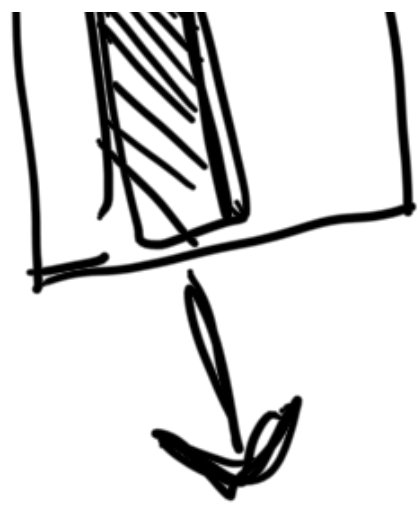
10000

5th Node



LIMIT 20
OFFSET 20





- Group by
- Subset