

# SQL

## Complete Summary

### Part - 3

Instructor:  
Deepak Poonia  
MTech, IISc Bangalore  
GATE CSE AIR 53; AIR 67;  
AIR 107; AIR 206; AIR 256

DBMS Complete Course:

<https://www.goclasses.in/courses/Database-Management-Systems>

# NOTE :

**Complete Discrete Mathematics & Complete Engineering Mathematics Courses**, by GO Classes, are **FREE** for ALL learners.

Visit here to watch : <https://www.goclasses.in/s/store/>

SignUp/Login on Goclasses website for free and start learning.

**COMPLETE COURSE**  
**Engineering Mathematics**

**GATE CS IT**

**Freely Available Features**

- ✓ All Video Lectures
- ✓ Quizzes
- ✓ Homeworks, Practice Sets
- ✓ Annotated Notes
- ✓ Toppers' Hand Written Notes
- ✓ GATE PYQs Video Solutions



**SACHIN MITTAL SIR**



[www.goclasses.in](http://www.goclasses.in)



/ GoClasses

**COMPLETE COURSE**  
**Discrete Mathematics**

**GATE CS IT**

**Freely Available Features**

- ✓ All Video Lectures
- ✓ Quizzes, Homeworks
- ✓ Annotated Notes
- ✓ Toppers' Hand Written Notes
- ✓ Summary Lectures
- ✓ GATE PYQs Video Solutions



**DEEPAK POONIA SIR**



[www.goclasses.in](http://www.goclasses.in)



/ GoClasses

FREE

# COMPLETE COURSE Linear Algebra

**GATE DA**

## Freely Available Features

- ✓ All Video Lectures
- ✓ Quizzes
- ✓ Homeworks, Practice Sets
- ✓ Annotated Notes
- ✓ Toppers' Hand Written Notes
- ✓ GATE PYQs Video Solutions

**SACHIN MITTAL SIR**[www.goclasses.in](http://www.goclasses.in)**/ GoClasses**

We are on Telegram. Contact us for any help.

# Link in the Description!!

Join GO Classes **Doubt Discussion** Telegram Group :



**@GATECSE\_GOCLASSES**

We are on Telegram. Contact us for any help.

Join GO Classes [Telegram Channel](#), Username: **@GOCLASSES\_CSE**

Join GO Classes **Doubt Discussion** Telegram Group :

Username: **@GATECSE\_Goclasses**

(Any doubt related to Goclasses Courses can also be asked here.)

Join **GATEOverflow Doubt Discussion** Telegram Group :

Username: **@GateOverflow\_CSE**





# SQL

## Complete Summary

### Part - 3

Recap:

SQL

GROUP BY clause

# GROUP BY

- ❖ SELECT ... FROM ... WHERE ...  
GROUP BY *list\_of\_columns*;
- ❖ Example: find the average GPA for each age group
  - SELECT age, AVG(GPA)  
FROM Student  
GROUP BY age;

# SQL Query Template with GROUP BY:

```
SELECT b1, ..., bk, AGG1(e1) AS x1, ..., AGGm(em) AS xm  
FROM R  
GROUP BY b1, ..., bk
```



## SQL: Structured Query Language Grouping Queries

SQL SELECT statement has two more clauses to support grouping operations: GROUP BY and HAVING clauses.

### GROUP BY Clause

The syntax of a GROUP BY clause is:

**GROUP BY <AttributeName>, ... , <AttributeName>**

The GROUP BY clause is added to the SELECT statement after the WHERE clause (or, if there is no WHERE clause, after the FROM clause).

GROUP BY clause causes the DBMS to separate the Cartesian Product of all tables referenced in the FROM clause into groups of tuples.

Each group of tuples *must agree on all values* of attributes listed in the GROUP BY clause.

Describe the output of the following SQL Query.

Consider, for example the relational table

Student(firstName, lastName, gpa, class, grade, school)

The following query:

```
SELECT school , COUNT( * )  
FROM Student  
GROUP BY school ;
```

Grouping operation is used to **allow for computation and reporting of aggregate operations over groups** in the SELECT statement.

Consider, for example the relational table

Student(firstName, lastName, gpa, class, grade, school)

The following query:

```
SELECT school , COUNT(*)  
FROM Student  
GROUP BY school ;
```

Valid Query

Will output the number of students enrolled in each school.

Next Sub-Topic:

SQL

Having Clause

## Having Clause:

Used to filter groups based on the group properties.

---

AFTER Groups are formed, we can eliminate some  
groups using HAVING clause.

---

Having Clause applies on one Group at a time.

---

Output??

SELECT A, Min(B)  
FROM R  
GROUP BY A;  
HAVING Count(\*) > 1

R

A	B
1	'a'
2	'b'
3	'a'
1	'b'
2	'a'
1	'c'

Output??

A	Min(B)
1	'a'
2	'a'

Count(\*) = 3

↑  
Q<sub>1</sub>

SELECT A, Min(B)  
FROM R ①  
GROUP BY A; ②  
HAVING Count(\*) > 1 ③

R	
A	B
1	'a'
2	'b'
3	'a'
1	'b'
2	'a'
1	'c'

Count(\*) = 3

↑  
Q<sub>2</sub>

Count(\*) = 2

↑  
Q<sub>3</sub>

Group Q<sub>1</sub> passes ✓  
 Group Q<sub>2</sub> passes ✓  
 Group Q<sub>3</sub> fails

A	B	C	D
a1	b1	1	7
a1	b1	2	8
a2	b1	3	9
<del>a3</del>	<del>b2</del>	4	10
a2	b1	5	11
a1	b1	6	12

```
SELECT A, B, SUM(C), MAX(D)  
FROM R  
GROUP BY A, B;  
HAVING B = 'b1' AND Sum(C) > 8
```

R

A	B	C	D
a1	b1	1	7
a1	b1	2	8
a2	b1	3	9
a3	b2	4	10
a2	b1	5	11
a1	b1	6	12

$g_1 \rightarrow B = 'b1'; \text{sum}(C) = 9$

- ④ SELECT A, B, SUM(C), MAX(D)
- ① FROM R
- ② GROUP BY A, B;
- ③ HAVING B = 'b1' AND Sum(C) > 8

O/p:  $\begin{array}{l} \max \\ \text{A } B \text{ sum}(C) \end{array}$

a, b1, 9	12
----------	----

check for every group  
independently

$B = 'b1'; \text{sum}(C) = 8$

only  
 $g_1$  passes

Note: when we have Group by & Having clause ;

Number of Tuples in O/P = Number of Groups that pass

Note: 'WHERE' clause is applied before creating groups .

Note: when we have Group by & No Having clause ;

Number of Tuples in O/P = Number of Groups

Note: 'WHERE' clause is applied before creating groups .

## WHERE Vs HAVING:

for Tuples

① Filters Tuples

② Applies on  
Tuples individually  
& independently

③ Whole tuple selected  
or Rejected

for Groups

① Filter Groups

② Applies on Groups  
individually & indepen-  
dently

③ Whole group selected  
or Rejected

# HAVING

- Used to filter groups based on the group properties (e.g., aggregate values, GROUP BY column values)
- `SELECT ... FROM ... WHERE ... GROUP BY ... HAVING condition;`
  - Compute `FROM (x)`
  - Compute `WHERE ( $\sigma$ )`
  - Compute `GROUP BY`: group rows according to the values of GROUP BY columns
  - Compute `HAVING` (another  $\sigma$  over the groups)
  - Compute `SELECT ( $\pi$ )` for each group that passes HAVING

# HAVING examples

- List the average popularity for each age group with more than a hundred users
  - ```
SELECT age, AVG(pop)
  FROM User
 GROUP BY age
 HAVING COUNT(*) > 100;
```
- Find average popularity for each age group over 10
  - ```
SELECT age, AVG(pop)
  FROM User
 GROUP BY age
 HAVING age > 10;
```

## Rules for Having Clause:

- ① Can NOT use Having without Group by.
- ② Group By  $A_1, A_2, A_3$   
HAVING  $A_1, A_3$   $\underbrace{\text{agg}(A_4), \text{agg}(A_1)}$   

*Subset of Group By can be used unaggregated*

*Aggregation on any attribute*

Not Allowed:

Group by A, B, C

Having A, (D), sum(D), count(\*), avg(A)

Not Allowed

## Rules for Having Clause:

There are several rules we must remember about HAVING clauses:

- An aggregation in a HAVING clause applies only to the tuples of the group being tested.
- Any attribute of relations in the FROM clause may be aggregated in the HAVING clause, but only those attributes that are in the GROUP BY list may appear unaggregated in the HAVING clause (the same rule as for the SELECT clause).

## \*\*SQL-92 Rules for Having Clause:

Can NOT have Having Clause in the Absence of GROUP BY clause.

When you use HAVING clause without using GROUP BY clause in SQL

@GO\_classes

SQL-92  
~~Abba~~ nahi manenge

# Describe the output of the following SQL Query.

Consider, for example the relational table

Student(firstName, lastName, gpa, class, grade, school)

For example,

```
SELECT school , COUNT(*)
FROM Student
GROUP BY school
HAVING AVG(gpa) > 3.0;
```

Consider, for example the relational table

**Student(firstName, lastName, gpa, class, grade, school)**

For example,

```
SELECT school , COUNT(*)  
FROM Student  
GROUP BY school  
HAVING AVG(gpa) > 3.0;
```

returns the number of students enrolled in each school, which has average student GPA over 3.0.

## HAVING Clause

The GROUP BY clause transforms the Cartesian Product from a relation of tuples into a relation of groups of tuples.

The HAVING clause is to the groups is what the WHERE clause is to the individual tuples. It provides a condition, and filters out the groups that fail it.

The syntax of HAVING clause is:

```
SELECT b1 , . . . , bk , AGG1( e1 ) AS x1 , . . . , AGGm(em) AS xm  
FROM R  
GROUP BY b1 , . . . , bk  
HAVING C;
```

(Q31) Find the age of the youngest sailor for each rating level.

Sailors(*sid: integer*, *sname: string*, *rating: integer*, *age: real*)

Boats(*bid: integer*, *bname: string*, *color: string*)

Reserves(*sid: integer*, *bid: integer*, *day: date*)



(Q31) Find the age of the youngest sailor for each rating level.

→ Group by

Phrase for  
Grouping

- ✓ Sailors(sid: integer, sname: string, rating: integer, age: real)
- Boats(bid: integer, bname: string, color: string)
- Reserves(sid: integer, bid: integer, day: date)

SELECT rating, min(age)  
FROM Sailors  
Group by rating

Next Sub-Topic:

SQL

Order BY Clause

If we want the FINAL Output “to be displayed” in SORTED order, we can use ORDER BY clause.

By default, Ordering is in the Ascending Order.

Can use DESC for Descending Order.

```
SELECT <attribute and function list>
FROM <table list>
[ WHERE <condition> ]
[ GROUP BY <grouping attribute(s)> ]
[ HAVING <group condition> ]
[ ORDER BY <attribute list> ];
```

Conceptual Evaluation :

- ① FROM
- ② WHERE
- ③ Group by

$\equiv$    
 SELECT  $\overbrace{A, B, \dots}$  --> CLASSES  
 ORDER By  $\overbrace{A, B}$  DESC }  
 {  
 ④ Having  
 ⑤ SELECT  
 ⑥ DISTINCT  
 ⑦ ORDER by

R

	A	B
1		'a'
2		'b'
3		'a'
1		'b'
2		'a'
1		'c'

SELECT \*  
ORDER BY A;

≡ ORDER BY A ASC ;

1	c	} fine
1	a	
2	a	} fine
2	b	
3	a	

MULTIPLE  
RESULTS  
POSSIBLE.

R

	A	B
r <sub>1</sub>	1	'a'
r <sub>2</sub>	2	'b'
r <sub>3</sub>	3	'a'
r <sub>4</sub>	1	'b'
r <sub>5</sub>	2	'a'
r <sub>6</sub>	1	'c'

SELECT \*

order by B

= order by  
B Asc

3	a
1	a
2	a
2	b
1	b
	c

multiple  
results  
possible.

R

	A	B
$r_1$	1	'a'
$r_2$	2	'b'
$r_3$	3	'a'
$r_4$	1	'b'
$r_5$	2	'a'
$r_6$	1	'c'

other  
order possible

order by A, B DESC;

1	'c'	must
1	'b'	
2	'a'	
2	'b'	must
2	'a'	
3	'a'	

No



R

	A	B
r <sub>1</sub>	1	'a'
r <sub>2</sub>	2	'b'
r <sub>3</sub>	3	'a'
r <sub>4</sub>	1	'b'
r <sub>5</sub>	2	'a'
r <sub>6</sub>	1	'c'

Order by B DESC, A ASC;

1	c
1	b
2	b
1	a
2	a
3	a

No other order possible.

## 6.1.8 Ordering the Output

We may ask that the tuples produced by a query be presented in sorted order. The order may be based on the value of any attribute, with ties broken by the value of a second attribute, remaining ties broken by a third, and so on, as in the  $\tau$  operation of Section 5.2.6. To get output in sorted order, we may add to the select-from-where statement a clause:

**ORDER BY <list of attributes>**

The order is by default ascending, but we can get the output highest-first by appending the keyword **DESC** (for “descending”) to an attribute. Similarly, we can specify ascending order with the keyword **ASC**, but that word is unnecessary.

### 3.4.4 Ordering the Display of Tuples

SQL offers the user some control over the order in which tuples in a relation are displayed. The **order by** clause causes the tuples in the result of a query to appear in sorted order. To list in alphabetic order all instructors in the Physics department, we write:

```
select name
      from instructor
     where dept_name = 'Physics'
       order by name;
```

By default, the **order by** clause lists items in ascending order. To specify the sort order, we may specify **desc** for descending order or **asc** for ascending order. Furthermore, ordering can be performed on multiple attributes. Suppose that we wish to list the entire *instructor* relation in descending order of *salary*. If several



instructors have the same salary, we order them in ascending order by name. We express this query in SQL as follows:

```
select *
from instructor
order by salary desc, name asc;
```



The order-item-list is a list of order-items; an order-item is a column name, optionally followed by one of the keywords ASC or DESC. Every column mentioned in the ORDER BY clause attributes must also appear in the select-list of the query; otherwise it is not clear what columns we should sort on.

The keywords ASC or DESC that follow a column control whether the result should be sorted—with respect to that column—in ascending or descending order; the default is ASC. This clause is applied as the last step in evaluating the query.

## Conceptual Order of Evaluation of a Select Statement:

1. First the product of all tables in the **FROM clause** is formed.
2. The **WHERE clause** is then evaluated to eliminate rows that do not satisfy the search\_condition.
3. Next, the rows are grouped using the columns in the **GROUP BY clause**.
4. Then, Groups that do not satisfy the search\_condition in the **HAVING clause** are eliminated.
5. Next, the expressions in the **SELECT clause** target list are evaluated.
6. If the **DISTINCT** keyword is present in the select clause, duplicate rows are now eliminated.
7. Finally, the resulting rows are sorted according to the columns specified in the **ORDER BY clause**.

HW: Number of tuples in the Output ??

```
SELECT S.rating, MIN (S.age)
FROM Sailors S
WHERE S.age >= 18
GROUP BY S.rating
HAVING COUNT (*) > 1
```

sid	sname	rating	age
22	dustin	7	45.0
31	lubber	8	55.5
71	zorba	10	16.0
64	horatio	7	35.0
29	brutus	1	33.0
58	rusty	10	35.0

```
SELECT S.rating, MIN(S.age)
FROM Sailors S
WHERE S.age >= 18
GROUP BY S.rating
HAVING COUNT(*) > 1
```

rating	age
1	33.0
7	45.0
7	35.0
8	55.5
10	35.0

sid	sname	rating	age
22	dustin	7	45.0
31	lubber	8	55.5
71	zorba	10	16.0
64	horatio	7	35.0
29	brutus	1	33.0
58	rusty	10	35.0

2

rating	m-age	count
1	33.0	1
7	35.0	2
8	55.0	1
10	35.0	1

3

rating	
7	35.0

Next Topic:

SQL

Nested Queries

Nested Query:

# Query within Query.

## 6.3 Subqueries

In SQL, one query can be used in various ways to help in the evaluation of another. A query that is part of another is called a *subquery*. Subqueries can have subqueries, and so on, down as many levels as we wish.

main Query

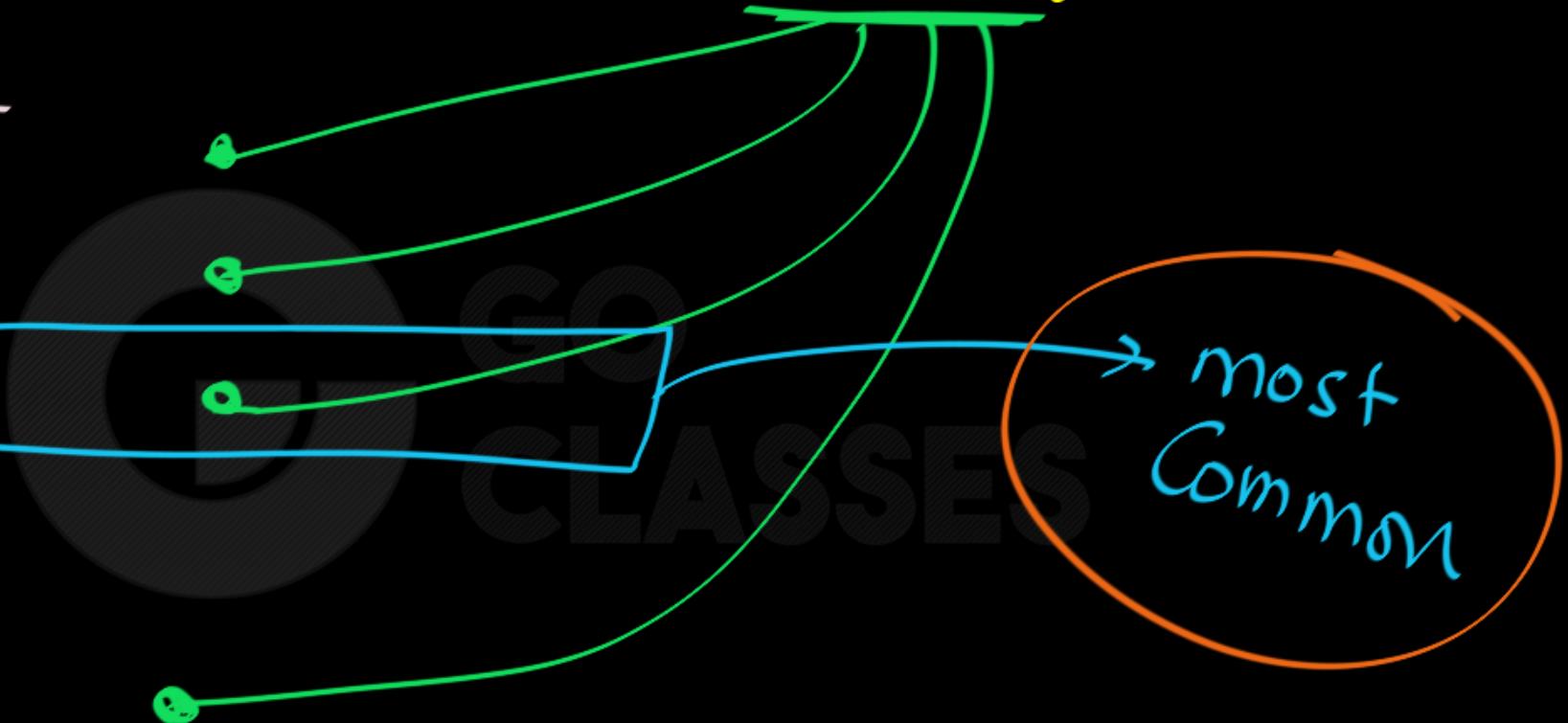
SELECT M.StarID, M.Name  
FROM MovieStar M  
WHERE M.StarID IN

Subquery

(SELECT S.StarID  
FROM StarsIn S  
WHERE MovieID=28)

Where can we insert a Subquery ??

SELECT  
FROM  
[ WHERE ]  
Group by  
Having



# Nested Queries

- A very powerful feature of SQL:

```
Select A1,A2,...,An
From R1,R2, ...,Rm
where condition
```

- A nested query is a query that has another query embedded with it.
  - A **SELECT, FROM, WHERE, or HAVING** clause can itself contain an SQL query!
  - Being part of the WHERE clause is the most common

## Two types of Sub-queries:

### 1. Simple Subquery:

Query without Co-relation with surrounding query..(NO dependence from outer query)

### 2. Co-related Subquery:

Query with Co-relation with surrounding query..(Dependence from outer query)

Movie(MovieID, Title, Year)

StarsIn(MovieID, StarID, role)

MovieStar(StarID, Name, Gender)

SELECT M.StarID, M.Name  
FROM MovieStar M  
WHERE M.StarID IN

main query

(SELECT S.StarID  
FROM StarsIn S  
WHERE MovieID=28)

Simple  
Subquery

Subquery

Movie(MovieID, Title, Year)  
StarsIn(MovieID, StarID, role)  
MovieStar(StarID, Name, Gender)

Scoping Rule :  
any attribute belongs  
to closest surrounding  
query.

SELECT FROM WHERE

movie

IN (SELECT S.StarID  
FROM StarsIn S  
WHERE MovieID=28)

main Query

```
SELECT cName, state  
FROM College C1  
WHERE exists
```

```
(SELECT *  
FROM College C2  
WHERE C2.state = C1.state AND  
C2.cName <> C1.cName);
```

Correlated Subquery

attribute of outer query

```
SELECT cName, state  
FROM College C1  
WHERE exists (SELECT *  
              FROM College C2  
              WHERE C2.state = C1.state AND  
                    C2.cName <> C1.cName);
```

“Evaluation of Subquery” in the WHERE clause:

## “Evaluation of Subquery” in the WHERE clause:

1. Simple Subquery in the WHERE clause:  
(Query without Co-relation)

The simplest subqueries can be evaluated once and for all, and the result used in a higher-level query.

“Evaluation of Subquery” in the WHERE clause:

## 2. Co-related Subquery:

Query with Co-relation with surrounding query..(Dependence from outer query)

Re-Computed for **EVERY** Row of Outer Table.

Take rows of Outer query one-by-one & evaluate the corelated subquery each time.

Nested Query concept is BEST Understood  
with the help of Examples.

But first, let us Study some SQL keywords.

Next Sub-Topic:

SQL

Understanding  
Some SQL Keywords

# Some SQL Keywords

**Set Comparator Keywords:**

**IN, ALL, EXISTS, ANY**

**NOT IN, NOT ALL, NOT EXISTS, NOT ANY**

Next Sub-Topic:

SQL

Some SQL Keywords

**IN, NOT IN**

IN  $\in \in$  membership keyword

NOT IN :  $\notin$

---

$s$  : a Tuple       $j$        $r$  : a Subquery

$s \text{ IN } r$  : True iff Tuple  $s$  is in  $r$ .

$s \text{ NOT IN } r$  : True iff Tuple  $s$  is Not in  $r$

---

$s$  : a Tuple ;  $R$  : a subquery

$s \text{ IN } R$   
 $s \text{ NOT IN } R$

}  $s, R$  must be  
compatible.

Tuple  $s (10, 'a', 20)$  then  $R$

int	char	int

Note :

Subquery R Result

10	a
10	b
20	b

10 IN

INVALID

NOT Compatible

# 1. IN , NOT IN keywords:

(set membership keywords)

The IN operator allows us to test whether a value or tuple is in a given set/bag;

$s \text{ IN } R$  is true if and only if  $s$  is equal to one of the tuples in  $R$ . Likewise,

$s \text{ NOT IN } R$  is true if and only if  $s$  is equal to no tuple in  $R$ .

$s \text{ IN } R$  is true if and only if  $s$  is equal to one of the values in  $R$ . Likewise,  $s \text{ NOT IN } R$  is true if and only if  $s$  is equal to no value in  $R$ .

- With  $\text{IN} (\in)$  and  $\text{NOT IN} (\notin)$  it is possible to check whether an attribute value appears in a set of values computed by another SQL **subquery**.

R	A	B
10	a	
20	a	
20	c	
30	a	

SELECT \*  
FROM R  
WHERE (A, B) IN R;

*Invalid*  
*must be*  
*a subquery*

R

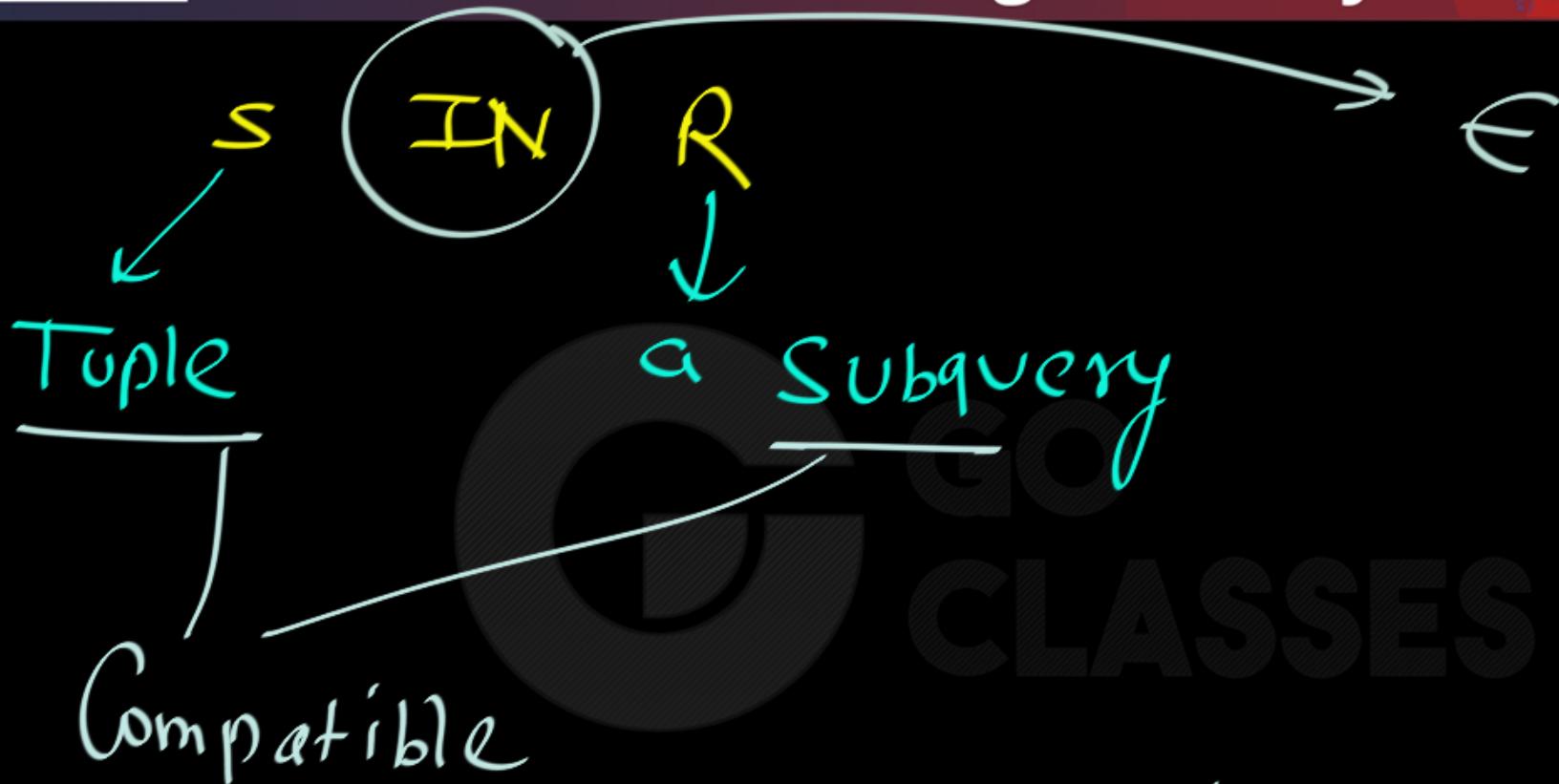
	A	B
10	a	
20	a	
20	c	
30	a	

① `SELECT * FROM R`  
② WHERE  $(A, B)$  IN  $\left( \begin{array}{l} \text{SELECT *} \\ \text{FROM R} \end{array} \right)$

Applies Tuple  
by Tuple

Final o/p = Same  
as R

Subquery  
 $\equiv R$



NOT  $IN \equiv \notin$

## STUDENTS

<u>SID</u>	FIRST	LAST	EMAIL
101	Ann	Smith	...
102	Michael	Jones	(null)
103	Richard	Turner	...
104	Maria	Brown	...

## RESULTS

<u>SID</u>	<u>CAT</u>	<u>ENO</u>	POINTS
101	H	1	10
101	H	2	8
101	M	1	12
102	H	1	9
102	H	2	9
102	M	1	10
103	H	1	5
103	M	1	7

## EXERCISES

<u>CAT</u>	<u>ENO</u>	TOPIC	MAXPT
H	1	Rel.Alg.	10
H	2	SQL	10
M	1	SQL	14



```
SELECT FIRST, LAST  
FROM STUDENTS  
WHERE SID NOT IN (SELECT SID  
                   FROM RESULTS  
                   WHERE CAT = 'H')
```

What is the output of this Query ??

```
SELECT FIRST, LAST  
FROM STUDENTS  
WHERE SID NOT IN (SELECT SID  
                   FROM RESULTS  
                   WHERE CAT = 'H')
```

Scoping Rule

## STUDENTS

<u>SID</u>	<u>FIRST</u>	<u>LAST</u>	<u>EMAIL</u>
101	Ann	Smith	...
102	Michael	Jones	(null)
103	Richard	Turner	...
104	Maria	Brown	...

→ pass

## EXERCISES

<u>CAT</u>	<u>ENO</u>	<u>TOPIC</u>	<u>MAXPT</u>
H	1	Rel.Alg.	10
H	2	SQL	10
M	1	SQL	14

## RESULTS

<u>SID</u>	<u>CAT</u>	<u>ENO</u>	<u>POINTS</u>
101	H	1	10
101	H	2	8
101	M	1	12
102	H	1	9
102	H	2	9
102	M	1	10
103	H	1	5
103	M	1	7



SELECT FIRST, LAST  
FROM STUDENTS  
WHERE SID NOT IN

↓  
Tuple by Tuple

{101, 101, 102, 102, 103}

(SELECT SID  
FROM RESULTS  
WHERE CAT = 'H')

O/P:

Maria	Brown
-------	-------

Subquery : Simple  
↓  
Execute once

## Students without any homework result.

```
SELECT FIRST, LAST  
FROM STUDENTS  
WHERE SID NOT IN (SELECT SID  
                    FROM RESULTS  
                    WHERE CAT = 'H' )
```

FIRST	LAST
Maria	Brown

Q: Consider the following Query on relation R (A,B).

```
SELECT *
FROM R
WHERE (A,B) IN ( R )
```

What is the output of it??

Q: Consider the following Query on relation R (A,B).

```
SELECT *  
FROM R  
WHERE (A,B) IN ( R )
```

Invalid

What is the output of it??

must be a  
Subquery

Consider the following Query: R (A,B)

SELECT \*  
FROM R  
WHERE (A,B) IN

SELECT \*  
FROM R )

Simple Subquery

$\exists R$

O/P: Same as R ✓

The IN operator allows us to test whether a value is in a given set of elements; an SQL query is used to generate the set to be tested.

There are a number of SQL operators that we can apply to a relation  $R$  and produce a boolean result. However, the relation  $R$  must be expressed as a subquery. As a trick, if we want to apply these operators to a stored table Foo, we can use the subquery (SELECT \* FROM Foo). The same trick works for union, intersection, and difference of relations.

# IN subqueries

- $x \text{ IN } (\text{subquery})$  checks if  $x$  is in the result of *subquery*
- Example: users at the same age as (some) Bart

- ```
SELECT *  
FROM User
```

What's Bart's age?

- ```
WHERE age IN (SELECT age
```

- ```
FROM User
```

- ```
WHERE name = 'Bart');
```

Next Sub-Topic:

SQL

Co-Related Sub-queries  
or Correlated Nested  
Queries

“Evaluation of Subquery” in the WHERE clause:

## 2. Co-related Subquery:

Query with Co-relation with surrounding query..(Dependence from outer query)

Re-Computed for **EVERY** Row of Outer Table.

---

Take rows of Outer query one-by-one & evaluate the corelated subquery each time.

Example:

```
SELECT *
FROM R
```

```
WHERE A > ( SELECT count(*)
              FROM S
              WHERE R.B = S.B)
```

R =  S = 

A	B	B	C	D
1	2	2	4	6
3	4	4	6	8
5	6	4	7	9

Example:  
SELECT \*  
FROM R  
WHERE A >

main query  
 $(\text{SELECT count(*)})$   
FROM S  
WHERE R.B = S.B)

Subquery : Correlated subquery

$R =$    
 $S =$

A	B
1	2
3	4
5	6

B	C	D
2	4	6
4	6	8
4	7	9

2. Suppose relations R(A,B) and S(B,C,D) are as follows:

Example:

```
SELECT *  
FROM R
```

WHERE A >

*Tuple by  
Tuple*

```
SELECT count(*)  
FROM S  
WHERE R.B = S.B)
```

R =

A	B
1	2
3	4
5	6

S =

B	C	D
2	4	6
4	6	8
4	7	9

for every Tuple of R, calculate inner query.

Example:

```
SELECT *
FROM R
```

WHERE A >

*checking  
for  
Tuple (1,2)*

<sup>1</sup>

`SELECT count(*)
FROM S
WHERE R.B = S.B`

*Count (\*)  
1*

fails

A	B
1	2
3	4
5	6

R =

B	C	D
2	4	6
4	6	8
4	7	9

S =

Example:

```
SELECT *
FROM R
```

WHERE A > (

3  

```
SELECT count(*)
FROM S
WHERE R.B = S.B)
```

Check for  
Tuple (3,4)

Passed

A	B
1	2
3	4
5	6

B	C	D
2	4	6
4	6	8
4	7	9



Count(\*)  
2

Example:

SELECT \*  
FROM R

WHERE A > (

Check for  
Tuple (5, 6)

Passed

R = 

A	B
1	2
3	4
5	6

 S = 

B	C	D
2	4	6
4	6	8
4	7	9

SELECT count(\*)  
FROM S  
WHERE R.B = S.B

Count(\*)  
0

Example:

SELECT \*  $\Rightarrow$  o/p:  
FROM R

A	B
3	4
5	6

R =   
S = 

A	B
1	2
3	4
5	6

B	C	D
2	4	6
4	6	8
4	7	9

WHERE A > ( SELECT count(\*)  
FROM S  
WHERE R.B = S.B )

Next Sub-Topic:

SQL

Some SQL Keywords

EXISTS , NOT EXISTS

EXIST : NOT a SQL keyword

EXISTS : keyword

---

Syntax :

EXISTS R

Subquery

Subquery

→ True iff R is Not Empty.

NOT EXISTS R

: True iff R is Empty.

# 1. EXISTS , NOT EXISTS keywords: (existence of a tuple)

EXISTS R is a condition that is true if and only if R is not empty.

NOT EXISTS R is true if and only if R is empty.

EXISTS R is a condition that is true if and only if R is not empty.

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *  
FROM R
```

```
WHERE EXISTS ( SELECT *  
                FROM S  
                WHERE B < C)
```

$$R = \begin{array}{|c|c|} \hline A & B \\ \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline 7 & 8 \\ \hline \end{array} \quad S = \begin{array}{|c|c|} \hline C & D \\ \hline 2 & 10 \\ \hline 4 & 12 \\ \hline 6 & 14 \\ \hline 8 & 16 \\ \hline \end{array}$$

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

$$R =$$

$$S =$$

Example: main query  
SELECT \* FROM R WHERE EXISTS (

( SELECT \* FROM S WHERE B < C )

Subquery : Correlated

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

SELECT \*  
FROM R

WHERE EXISTS ( SELECT \*  
FROM S  
WHERE B < C)

Applies  
Tuple by  
Tuple of R

R =	A B	S =	C D
	1 2		2 10
	3 4		4 12
	5 6		6 14
	7 8		8 16

Example:

```
SELECT *
FROM R
```

WHERE EXISTS

check  
Tuple  
(1,2)

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Passed

A	B
1	2
3	4
5	6
7	8

R =

C	D
2	10
4	12
6	14
8	16

S =

SELECT \*  
FROM S  
WHERE B < C)

C	D
4	12
6	14
8	16

Example:

```
SELECT *
FROM R
```

WHERE EXISTS

Check  
Tuple

(3,4)

True

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

( SELECT \*  
FROM S  
WHERE (B < C) )

C	D
6	14
8	16

Example:

```
SELECT *
FROM R
```

WHERE EXISTS

Check  
Tuple

(5,6)

True

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

```
( SELECT *
  FROM S
 WHERE B < C )
```

C	D
2	10
4	12

Example:

```
SELECT *
FROM R
```

WHERE EXISTS

check  
tuple  $(7, 8)$

$\nwarrow$   
 $\downarrow$   
 $\downarrow$   
*false*

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

( SELECT \*  
FROM S  
WHERE  $B < C$ )

C	D
2	10

$\nwarrow$   
*Empty  
Table*

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *  
FROM R
```

```
WHERE EXISTS ( SELECT *  
                FROM S  
                WHERE B < C)
```



O/P:

A	B
1	2
3	4
5	6

R =

A	B
1	2
3	4
5	6
7	8

S =

C	D
2	10
4	12
6	14
8	16

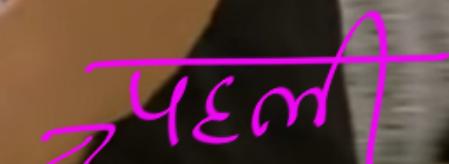
Every  
Row      Outer  
Query

**ja puch ke aa**

Correlated subquery

@GO\_Classes

# Simple Subquery (without correlation)

Main  aakhiri baar aaya hu

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *  
FROM R
```

$$R = \begin{array}{|c|c|}\hline A & B \\ \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline 7 & 8 \\ \hline\end{array} \quad S = \begin{array}{|c|c|}\hline C & D \\ \hline 2 & 10 \\ \hline 4 & 12 \\ \hline 6 & 14 \\ \hline 8 & 16 \\ \hline\end{array}$$

```
WHERE EXISTS ( SELECT count(*)  
                FROM S  
                WHERE B < C)
```

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example: *main query*  
SELECT \*  
FROM R  
WHERE

EXISTS ( SELECT count(\*)  
FROM S  
WHERE B < C)

R =	S =
A B	C D
1 2	2 10
3 4	4 12
5 6	6 14
7 8	8 16

*Correlated Subquery*

Example:

SELECT \*  
FROM R

WHERE EXISTS

True

Tuple by  
 Tuple  
(1,2) Tuple

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Passes

R =      S =

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

SELECT count(\*)  
 FROM S  
 WHERE B < C)

Count(\*)  
 3

Example:

```
SELECT *  
FROM R
```

```
WHERE EXISTS ( SELECT count(*)  
                FROM S  
                WHERE B < C)
```

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

The diagram shows two tables, R and S, on a light blue background. Table R (left) has columns A and B, with rows containing (1,2), (3,4), (5,6), and (7,8). Table S (right) has columns C and D, with rows containing (2,10), (4,12), (6,14), and (8,16). Handwritten annotations in red and purple are present: 'Pass' is written above the first three rows of R, and a purple circle highlights the last row of R (5,6). Red arrows point from the first three rows of R to the first three rows of S. In the center, 'R =' is circled in red and 'S =' is circled in purple.

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

Example:

```
SELECT *
FROM R
```

WHERE EXISTS

Check Tuple  
 $(7, 8)$

True

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Pass

$R =$

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

X

X

X

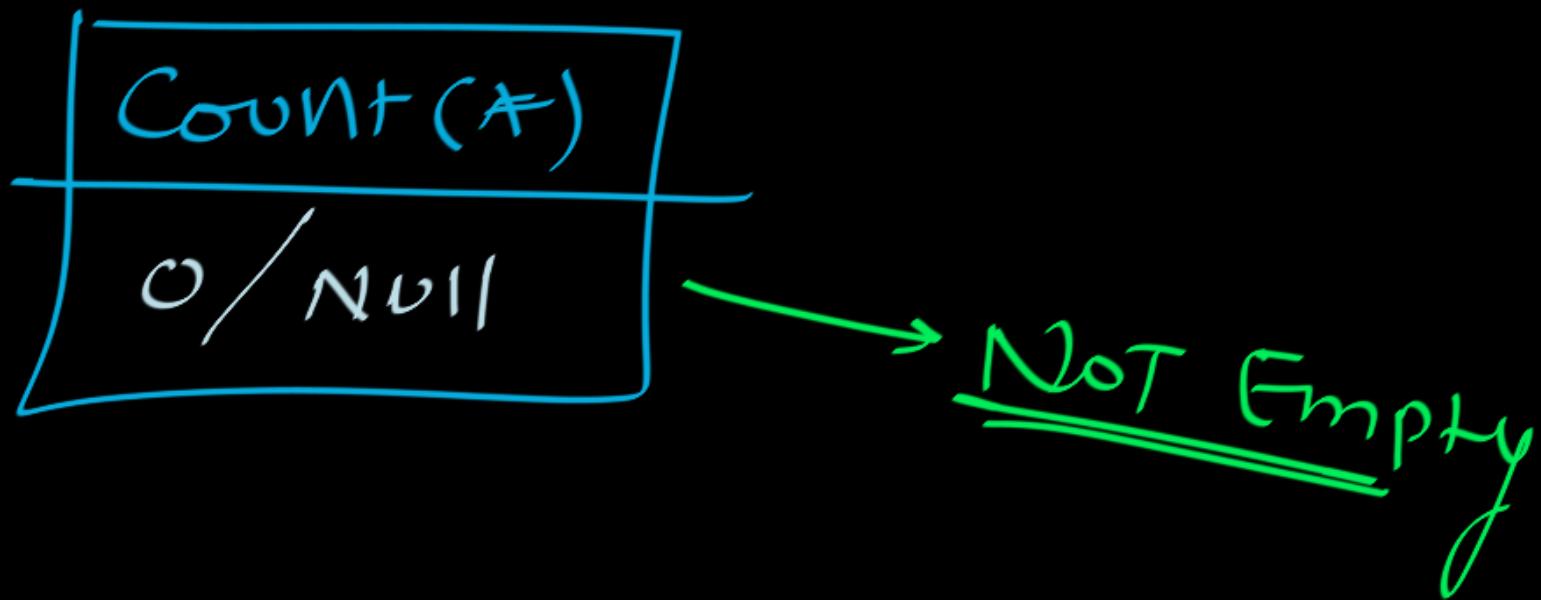
X

( SELECT count(\*)  
 FROM S  
 WHERE  $B < C$ )



NOT Empty

Empty Table : Containing No Tuple



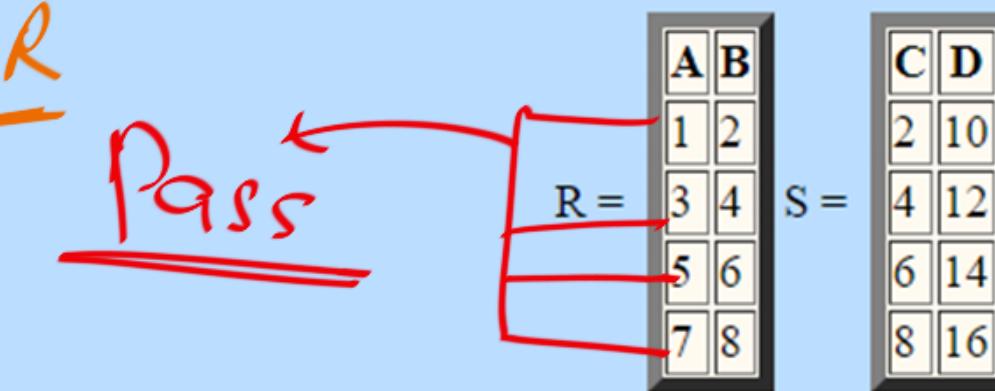
6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *
FROM R
```

```
WHERE EXISTS ( SELECT count(*)
                FROM S
                WHERE B < C)
```

*Always  
True*



Whatever A :

Count (\*)

Count (A)

Sum (A)

Avg (A)

min (A)

max (A)

Never Give  
Empty Result.

A

Empty

Count(\*) = 0

Count(A) = 0

Sum(A) = 0

min(A) = null

max(A) = null

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *  
FROM R
```

```
WHERE NOT EXISTS ( SELECT *  
FROM S  
WHERE B < C)
```

R =	A	B	S =	C	D
	1	2		2	10
	3	4		4	12
	5	6		6	14
	7	8		8	16

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

SELECT \*  
FROM R

*O/P*

A	B
7	8

*Pass*

R =      S =

A	B	C	D
1	2	2	10
3	4	4	12
5	6	6	14
7	8	8	16

WHERE NOT EXISTS ( SELECT \*  
FROM S  
WHERE B < C)

Tuple by  
Tuple

*True when  
Subquery Result is Empty*

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *  
FROM R
```

```
WHERE NOT EXISTS ( SELECT count(*)  
FROM S  
WHERE B < C)
```

R =	A	B	S =	C	D
	1	2		2	10
	3	4		4	12
	5	6		6	14
	7	8		8	16

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:  $\rightarrow \sigma_P =$

SELECT \*  
FROM R

WHERE NOT EXISTS

*Always  
False*

Empty Table

All fail

R =      S =

A	B	C	D
1	2	2	10
3	4	4	12
5	6	6	14
7	8	8	16

( SELECT count(\*)  
FROM S  
WHERE B < C)

*Never Empty  
Table*

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *  
FROM R
```

```
WHERE ( SELECT *  
        FROM S  
        WHERE B < C) >= 2 ;
```

$$R = \begin{array}{|c|c|}\hline A & B \\ \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline 7 & 8 \\ \hline\end{array} \quad S = \begin{array}{|c|c|}\hline C & D \\ \hline 2 & 10 \\ \hline 4 & 12 \\ \hline 6 & 14 \\ \hline 8 & 16 \\ \hline\end{array}$$

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *\nFROM R\nWHERE
```

```
( SELECT *\n  FROM S\n WHERE B < C ) >= 2 ;
```

a  
Table

Invalid Query

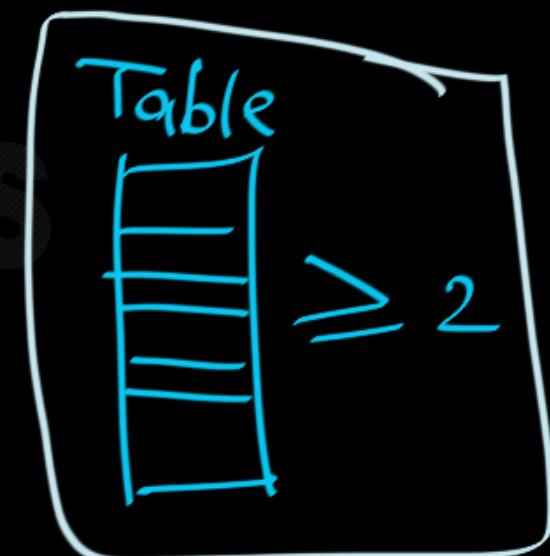
No aggregate

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

R =

S =



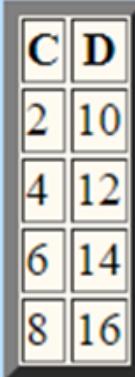
Invalid

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *
FROM R
```

```
WHERE ( SELECT count(*)
        FROM S
        WHERE B < C) >= 2 ;
```

R =  S = 

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *\nFROM R\nWHERE
```

```
(SELECT count(*)\nFROM S\nWHERE B < C)>= 2 ;
```

Simple  
Column, Simple Tuple

Simple aggregates

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

R =

S =

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

```
SELECT *
FROM R
```

```
WHERE ( SELECT count(*)
        FROM S
        WHERE B < C ) >= 2 ;
```

Subquery : Coselated ✓

A	B
1	2
3	4
5	6
7	8

C	D
2	10
4	12
6	14
8	16

R =

S =

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Example:

SELECT \*

FROM R

WHERE

↓  
Tuple  
by  
Tuple

( SELECT count(\*)  
FROM S  
WHERE  $B < C$ ) \geq 2 ;

Pass

$R =$

A	B
1	2
3	4
5	6
7	8

$S =$

C	D
2	10
4	12
6	14
8	16

✓  
✓  
✓

C	D
2	10
4	12
6	14
8	16

Count(\*)  
3  
 $\geq 2$

Value 3



DBMS

internally

Example:

SELECT \*  
FROM R

WHERE ( SELECT count(\*)  
FROM S  
WHERE B < C) >= 2 ;

<u>Q/P</u>	$\rho$						
	<table border="1"> <thead> <tr> <th>A</th><th>B</th></tr> </thead> <tbody> <tr> <td>1</td><td>2</td></tr> <tr> <td>3</td><td>4</td></tr> </tbody> </table>	A	B	1	2	3	4
A	B						
1	2						
3	4						

6. Here are two relations, R(A,B) and S(C,D). Their current values are:

Diagram illustrating the state of relations R and S. Relation R (A, B) has 8 rows with values (1, 2), (3, 4), (5, 6), and (7, 8). Relation S (C, D) has 8 rows with values (2, 10), (4, 12), (6, 14), and (8, 16). Handwritten annotations show 'pass' with arrows pointing to the first two rows of R, and 'fail' with an arrow pointing to the last row of R.

R =	A	B
pass	1	2
pass	3	4
fail	5	6
	7	8

S =

C	D
2	10
4	12
6	14
8	16

Next Sub-Topic:

SQL

Some SQL Keywords

UNIQUE, NOT UNIQUE

Closely related to EXISTS is the UNIQUE predicate.

When we apply UNIQUE to a subquery, it returns true if no row appears twice in the answer to the subquery, that is, there are no duplicates;

Syntax:

Unique  $R$

Not Unique  $R$

Subquery

---

Subquery

: True iff  $R$  has No  
Duplicate Tuples.

True iff  $R$  has  
Some Duplication.

Next Sub-Topic:

SQL

SOME SQL KEYWORDS

ALL , ANY/SOME

ALL

ANY / some  
key words

SUGGESTIONS

ANY some

Syntax:

$(<, =, \leq, >, \geq, \neq)$

$S < \underline{\text{ALL}} R$

Universal  
Quantifier ( $\forall$ )

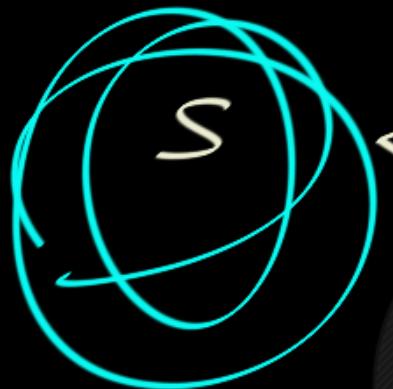
If R Empty  
then "ALL" is  
Always True.

Syntax:

$(<, =, \leq, >, \geq, \neq)$

$S < \cancel{\text{Any}} \ R$  :  $S$  must be  
less than  
 $\cancel{\text{Any}}$   $\overset{\text{Some}}{\curvearrowleft}$  value  
of  $R$

at least  
one

Syntax:

Existential  
Quantifier

at least one

If R Empty  
then "Some"  
is Always False

## Syntax:

NOT  $s < \text{ALL } R$  : True if  $s$  is  $\geq$  some value in  $R$ .

## Syntax:

NOT  $S \leq \text{Some } R$  ;  $S$  is  
Greater than  
All values  
of  $R$ .

### 5.4.3 Set-Comparison Operators

We have already seen the set-comparison operators EXISTS, IN, and UNIQUE, along with their negated versions. SQL also supports op ANY and op ALL, where op is one of the arithmetic comparison operators {<, <=, =, >, >=, >}. (SOME is also available, but it is just a synonym for ANY.)



3.  $s > \text{ALL } R$  is true if and only if  $s$  is greater than every value in unary relation  $R$ . Similarly, the  $>$  operator could be replaced by any of the other five comparison operators, with the analogous meaning:  $s$  stands in the stated relationship to every tuple in  $R$ . For instance,  $s <> \text{ALL } R$  is the same as  $s \text{ NOT IN } R$ .
4.  $s > \text{ANY } R$  is true if and only if  $s$  is greater than at least one value in unary relation  $R$ . Similarly, any of the other five comparisons could be used in place of  $>$ , with the meaning that  $s$  stands in the stated relationship to at least one tuple of  $R$ . For instance,  $s = \text{ANY } R$  is the same as  $s \text{ IN } R$ .

The **EXISTS**, **ALL**, and **ANY** operators can be negated by putting **NOT** in front of the entire expression, just like any other boolean-valued expression. Thus,  $\text{NOT EXISTS } R$  is true if and only if  $R$  is empty.  $\text{NOT } s \geq \text{ALL } R$  is true if and only if  $s$  is not the maximum value in  $R$ , and  $\text{NOT } s > \text{ANY } R$  is true if and only if  $s$  is the minimum value in  $R$ . We shall see several examples of the use of these operators shortly.

Next Sub-Topic:

SQL

Some more points...

Note that :

IN and NOT IN are equivalent to = ANY and <> ALL,  
respectively.

$$S [IN] R \equiv S [= ANY] R$$
$$S [NOT IN] R \equiv S [<> ALL] R$$

NOTE:

EXISTS and Aggregate in the inner query

: Take  
Care

EXISTS

Always  
True

SELECT Agg(A)

Never  
Empty  
Result

# Clicker nested question

co-related query - apply subquery  
on each tuple

Determine the result of:

```
SELECT Team, Day
FROM Scores S1
WHERE Runs <= ALL
    (SELECT Runs
     FROM Scores S2
     WHERE S1.Day = S2.Day )
```

Which of the following is in  
the result:

- A. (Carp, Sun)
- B. (Bay Stars, Sun)
- C. (Swallows, Mon)
- D. All of the above
- E. None of the above

## Scores:

Team	Day	Opponent	Runs
Dragons	Sun	Swallows	4
Tigers	Sun	Bay Stars	9
Carp	Sun	Giants	2
Swallows	Sun	Dragons	7
Bay Stars	Sun	Tigers	2
Giants	Sun	Carp	4
Dragons	Mon	Carp	6
Tigers	Mon	Bay Stars	5
Carp	Mon	Dragons	3
Swallows	Mon	Giants	0
Bay Stars	Mon	Tigers	7
Giants	Mon	Swallows	5

# Clicker nested question

Determine the result of:

SELECT Team, Day

FROM Scores S1

WHERE Runs <= ALL

(SELECT Runs  
FROM Scores S2  
WHERE S1.Day = S2.Day)

Scores:

Team	Day	Opponent	Runs
Dragons	Sun	Swallows	4
Tigers	Sun	Bay Stars	9
Carp	Sun	Giants	2
Swallows	Sun	Dragons	7
Bay Stars	Sun	Tigers	2
Giants	Sun	Carp	4
Dragons	Mon	Carp	6
Tigers	Mon	Bay Stars	5
Carp	Mon	Dragons	3
Swallows	Mon	Giants	0
Bay Stars	Mon	Tigers	7
Giants	Mon	Swallows	5

Correlated Subquery

# Clicker nested question *fails*

Determine the result of:

```
SELECT Team, Day
FROM Scores S1
WHERE Runs <= ALL
  (SELECT Runs
   FROM Scores S2
   WHERE S1.Day = S2.Day )
```

tuple  $t_1$ :  $4 \leq \text{ALL}$

4
9
2
7
2
4

all ✓  
Runs made  
on Sunday

Scores:

Team	Day	Opponent	Runs
Dragons	Sun	Swallows	4
Tigers	Sun	Bay Stars	9
Carp	Sun	Giants	2
Swallows	Sun	Dragons	7
Bay Stars	Sun	Tigers	2
Giants	Sun	Carp	4
Dragons	Mon	Carp	6
Tigers	Mon	Bay Stars	5
Carp	Mon	Dragons	3
Swallows	Mon	Giants	0
Bay Stars	Mon	Tigers	7
Giants	Mon	Swallows	5

# Clicker nested question

Determine the result of:

```
SELECT Team, Day
FROM Scores S1
WHERE Runs <= ALL
  (SELECT Runs
   FROM Scores S2
   WHERE S1.Day = S2.Day)
```

$t_1:$

3 Tuples

Scores:			
Team	Day	Opponent	Runs
Dragons	Sun	Swallows	4
Tigers	Sun	Bay Stars	9
Carp	Sun	Giants	2 ✓
Swallows	Sun	Dragons	7
Bay Stars	Sun	Tigers	2 ✓
Giants	Sun	Carp	4
Dragons	Mon	Carp	6
Tigers	Mon	Bay Stars	5
Carp	Mon	Dragons	3
Swallows	Mon	Giants	0 ✓
Bay Stars	Mon	Tigers	7
Giants	Mon	Swallows	5

We want those  
Tuples where Runs  
are Least on that  
Day.

# Clicker nested question

Determine the result of:

```
SELECT Team, Day  
FROM Scores S1  
WHERE Runs <= ALL  
(SELECT Runs  
FROM Scores S2  
WHERE S1.Day = S2.Day )
```

Which of the following is in the result:

- A. (Carp, Sun) ✓
- B. (Bay Stars, Sun) ✓
- C. (Swallows, Mon) ✓
- D. All of the above ✓
- E. None of the above

Scores:

Team	Day	Opponent	Runs
Dragons	Sun	Swallows	4
Tigers	Sun	Bay Stars	9
Carp	Sun	Giants	2 ✓
Swallows	Sun	Dragons	7
Bay Stars	Sun	Tigers	2 ✓
Giants	Sun	Carp	4
Dragons	Mon	Carp	6
Tigers	Mon	Bay Stars	5
Carp	Mon	Dragons	3
Swallows	Mon	Giants	0 ✓
Bay Stars	Mon	Tigers	7
Giants	Mon	Swallows	5

Determine the result of:

```
SELECT Team, Day  
FROM Scores S1  
WHERE Runs <= ALL  
(SELECT Runs  
FROM Scores S2  
WHERE S1.Day = S2.Day )
```

Which of the following is in the result:

- A. (Carp, Sun)
- B. (Bay Stars, Sun)
- C. (Swallows, Mon)
- D. All of the above
- E. None of the above

Correct

## Scores:

Team	Day	Opponent	Runs
Dragons	Sun	Swallows	4
Tigers	Sun	Bay Stars	9
Carp	Sun	Giants	2
Swallows	Sun	Dragons	7
Bay Stars	Sun	Tigers	2
Giants	Sun	Carp	4
Dragons	Mon	Carp	6
Tigers	Mon	Bay Stars	5
Carp	Mon	Dragons	3
Swallows	Mon	Giants	0
Bay Stars	Mon	Tigers	7
Giants	Mon	Swallows	5

Team/Day pairs such that the team scored the minimum number of runs for that day.

# Semantics of subqueries

- ```
SELECT *
  FROM Users AS u
 WHERE EXISTS (SELECT * FROM User
               WHERE name = 'Bart'
               AND age = u.age);
```
- For each row  $u$  in User
  - Evaluate the subquery with the value of  $u.age$
  - If the result of the subquery is not empty, output  $u.*$
- The DBMS query optimizer may choose to process the query in an equivalent, but more efficient way

# Scoping rule of subqueries

- To find out which table a column belongs to
  - Start with the immediately surrounding query
  - If not found, look in the one surrounding that; repeat if necessary
- Use *table\_name.column\_name* notation and AS (renaming) to avoid confusion

Next Sub-Topic:

SQL

Joins in SQL

## 5.1.6 Joins of Bags

Joining bags presents no surprises. We compare each tuple of one relation with each tuple of the other, decide whether or not this pair of tuples joins successfully, and if so we put the resulting tuple in the answer. When constructing the answer, we do not eliminate duplicate tuples.

Joins in Rel. Alg.  $\equiv$  Joins in SQL

Duplicates  
NOT Eliminated

| A | B |
|---|---|
| 1 | 2 |
| 1 | 2 |

(a) The relation  $R$ 

| B | C |
|---|---|
| 2 | 3 |
| 4 | 5 |
| 4 | 5 |

(b) The relation  $S$ 

| A | R.B | S.B | C |
|---|-----|-----|---|
| 1 | 2   | 2   | 3 |
| 1 | 2   | 2   | 3 |
| 1 | 2   | 4   | 5 |
| 1 | 2   | 4   | 5 |
| 1 | 2   | 4   | 5 |
| 1 | 2   | 4   | 5 |

(c) The product  $R \times S$ 

Figure 5.3: Computing the product of bags

Table  $R$  :  $i$  attributes,  $m$  tuples

$S$  :  $j$  attributes,  $n$  tuples

$R \times S$  :  $i + j$  attributes,  $mn$  tuples

| A | B |
|---|---|
| 1 | 2 |
| 1 | 2 |

(a) The relation  $R$

| B | C |
|---|---|
| 2 | 3 |
| 4 | 5 |
| 4 | 5 |

(b) The relation  $S$

$R \bowtie S$  :

Natural Join :

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 1 | 2 | 3 |

That is, tuple  $(1, 2)$  of  $R$  joins with  $(2, 3)$  of  $S$ . Since there are two copies of  $(1, 2)$  in  $R$  and one copy of  $(2, 3)$  in  $S$ , there are two pairs of tuples that join to give the tuple  $(1, 2, 3)$ . No other tuples from  $R$  and  $S$  join successfully.

| A | B |
|---|---|
| 1 | 2 |
| 1 | 2 |

# Database Management System

GO Classes

(a) The relation  $R$

| B | C |
|---|---|
| 2 | 3 |
| 4 | 5 |
| 4 | 5 |

(b) The relation  $S$

As another example on the same relations  $R$  and  $S$ , the theta-join

$$R \bowtie_{R.B < S.B} S$$

produces the bag

| A | R.B | S.B | C |
|---|-----|-----|---|
| 1 | 2   | 4   | 5 |
| 1 | 2   | 4   | 5 |
| 1 | 2   | 4   | 5 |
| 1 | 2   | 4   | 5 |

# Clicker outer join question

- Given:

Compute:

```
SELECT R.A, R.B, S.B, S.C, S.D  
FROM R FULL OUTER JOIN S  
    ON (R.A > S.B AND R.B = S.C)
```

- Which of the following tuples of R or S is dangling (and therefore needs to be padded in the outer join)?

- (1,2) of R
- (3,4) of R
- (2,4,6) of S
- All of the above
- None of the above

| R(A,B) |   | S(B,C,D) |   |   |
|--------|---|----------|---|---|
| A      | B | B        | C | D |
| 1      | 2 | 2        | 4 | 6 |
| 3      | 4 | 4        | 6 | 8 |
| 5      | 6 | 4        | 7 | 9 |

# Clicker outer join question

- Given:

Compute:

SELECT R.A, R.B, S.B, S.C, S.D  
FROM R FULL OUTER JOIN S  
ON (R.A > S.B AND R.B = S.C)

- Which of the following tuples of R or S is dangling (and therefore needs to be padded in the outer join)?

- (1,2) of R ✓
- (3,4) of R
- (2,4,6) of S
- All of the above

✓ E. (4,7,9) of S

Dangling Tuple

R(A,B) S(B,C,D)

| A | B | B | C | D |
|---|---|---|---|---|
| 1 | 2 | 2 | 4 | 6 |
| 3 | 4 | 4 | 6 | 8 |
| 5 | 6 | 4 | 7 | 9 |

Dangling

~~R~~ ~~S~~

$R.A > S.B \ \& \ R.B = S.C$

| A | R.B | S.B  | C    | D    |
|---|-----|------|------|------|
| 1 | 2   | NULL | NULL | NULL |
| 3 | 4   | 2    | 4    | 6    |
| 5 | 6   | 4    | 7    | 9    |
|   |     |      | NULL | NULL |

NULL NULL 4

## Dangling Tuple:

In the Join operations,

A tuple that fails to pair with any tuple of the other relation in a join is said to be a dangling tuple.

# Clicker outer join question

- Given:

Compute:

```
SELECT R.A, R.B, S.B, S.C, S.D  
FROM R FULL OUTER JOIN S  
    ON (R.A > S.B AND R.B = S.C)
```

- Which of the following tuples of R or S is dangling (and therefore needs to be padded in the outer join)?

- A. (1,2) of R
- B. (3,4) of R
- C. (2,4,6) of S
- D. All of the above
- E. None of the above

R(A,B) S(B,C,D)

| A | B | B | C | D |
|---|---|---|---|---|
| 1 | 2 | 2 | 4 | 6 |
| 3 | 4 | 4 | 6 | 8 |
| 5 | 6 | 4 | 7 | 9 |

| A    | B    | B    | C    | D    |
|------|------|------|------|------|
| 3    | 4    | 2    | 4    | 6    |
| 5    | 6    | 4    | 6    | 8    |
| 1    | 2    | NULL | NULL | NULL |
| NULL | NULL | 4    | 7    | 9    |

5.39.1 Sql: UGC NET CSE | August 2016 | Part 2 | Question: 20 top

Consider a database table  $R$  with attributes  $A$  and  $B$ . Which of the following  $SQL$  queries is illegal?

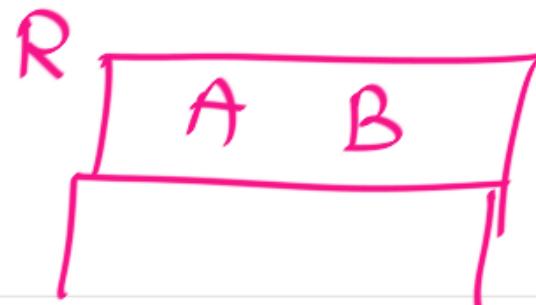
- A. SELECT A FROM R;
- C. SELECT A, COUNT(\*) FROM R GROUP BY A;
- B. **SELECT A, COUNT(\*) FROM R;**
- D. SELECT A, B, COUNT(\*) FROM R GROUP BY A, B;

ugcnetcse-aug2016-paper2 databases sql



## UGC NET CSE | August 2016 | Part 2 | ★

## Question: 20



1,862 views

asked Sep 24, 2016 • recategorized Oct 14, 2018 by Pooja Khatri



Consider a database table  $R$  with attributes  $A$  and  $B$ . Which of the following SQL queries is illegal?

- 2
- A. SELECT A FROM R;
  - B. SELECT A, COUNT(\*) FROM R;
  - C. SELECT A, COUNT(\*) FROM R GROUP BY A;
  - D. SELECT A, B, COUNT(\*) FROM R GROUP BY A, B;

{ SELECT A, Count(\*)  
From R  
})  
illegal

Invalid

Databases

#ugcnetcse-aug2016-paper2

#databases

#sql

Premium IN

go classes sql



## Database Management System



Mod 2 - Lecture 24:

# SQL GATE PYQs

- Deepak Poonia

(IISc Bangalore; GATE AIR 53 & 67)

www.goclasses.in

2:07:21

SQL GATE Questions - Complete Analysis | DBMS GATE CS PYQs | GO Classes | Deepak Poonia

4.4K views • 1 year ago

GO Classes for GATE CS

SQL GATE Questions - Complete Analysis | DBMS GATE CS PYQs | GO Classes | Deepak Poonia Crack #GateCSE



11 chapters SQL GATE Questions Session Begins | GATE CSE 2000 | Question: 2.26 - SQL...



www.goclasses.in