Advanced SQL Queries

Author:- Azarudhin Usman Reference: Youtube- @AnalystAdithya

Introduction:

This document covers SQL topics like Aggregate functions, Joins, Windows Functions and joins.

-- Find numbers which occur consicutively at least 3 times.

Sample Output:

num integer
1
3

-- Query

```
select num from (with a as (select * from (
values
      (1,1),
      (2,1),
      (3,1),
      (4,2),
      (5,2),
      (6,3),(7,3), (8,3)
) as apple_store(id, num))
select num, lag(num)over(order by id) prevnum, lead(num)over(order by id)
nextnum from a) b
where num = prevnum and num = nextnum
```

Output:

```
select num from (with a as (select * from (
values
    (1,1),
    (2,1),
    (3,1),
    (4,2),
    (5,2),
    (6,3),(7,3), (8,3)
) as apple_store(id, num))
select num, lag(num)over(order by id) prevnum, lead(num)over(order by id) nextnum from a) b
where num = prevnum and num = nextnum
```

Data Output | Messages | Notifications

	num integer
1	1
2	3

-- To display the number which is present >= 4 times

-- Query

```
select distinct num from (with a as (select * from (
values
      (1,1),
      (2,1),
      (3,1),
      (4,1),
      (5,2),
      (6,3),(7,3), (8,3)
) as apple_store(id, num))
select id, num, count(*) over(partition by num) as num_count,
lag(num)over(order by id) prevnum, lead(num)over(order by id) nextnum from
a) b
where num_count >=4 and num = prevnum and num = nextnum
```

Output:

```
select distinct num from (with a as (select * from (
values
    (1,1),
    (2,1),
    (3,1),
    (4,1),
    (5,2),
    (6,3),(7,3), (8,3)
) as apple_store(id, num))
select id, num, count(*) over(partition by num) as num_count, lag(num)over(order by id) prevnum, lead(num)over
where num_count >=4 and num = prevnum and num = nextnum
```

Data Output | Messages | Notifications

	num integer
1	1

****

-- Display the deparment details where the employee has skilled in Excel and SQL

Sample Output:

department text	ename text	count bigint
Marketing	Andrew	2
SDE	Allen	2

-- Query

```
with a as (select * from (
values
      ('Marketing','Andrew','Excel'),
      ('Marketing','Andrew','SQL'),
      ('Marketing','Alex','Excel'),
      ('HR','John','Excel'),
      ('SDE','Allen','Power BI'),
      ('SDE','Allen','SQL'),
      ('SDE','Allen','Excel')
) as company (department, ename, skills))

select department, ename, count(distinct skills) from a
where skills in ('Excel','SQL')
group by department, ename
having count(distinct skills)=2
```

Output:

```sql
with a as (select * from (
values
    ('Marketing','Andrew','Excel'),
    ('Marketing','Andrew','SQL'),
    ('Marketing','Alex','Excel'),
    ('HR','John','Excel'),
    ('SDE','Allen','Power BI'),
    ('SDE','Allen','SQL'),
    ('SDE','Allen','Excel')
) as company (department, ename, skills))

select department, ename, count(distinct skills) from a
where skills in ('Excel','SQL')
group by department, ename
having count(distinct skills)=2
```

Data Output | Messages | Notifications

	department text	ename text	count bigint
1	Marketing	Andrew	2
2	SDE	Allen	2

-- Display the department, employee who has skilled in 2 technology other than Excel and SQL

-- Query

```sql
with a as (select * from (
values
      ('Marketing','Andrew','Excel'),
      ('Marketing','Andrew','SQL'),
      ('Marketing','Alex','Excel'),
      ('HR','John','Excel'),
      ('SDE','Allen','Power BI'),
      ('SDE','Allen','SQL'),
      ('SDE','Allen','Excel'),
      ('SDE','Allen','Python')
) as company (department, ename, skills))

select department, ename, count(distinct skills) from a
where skills not in  ('Excel','SQL')
group by department, ename
having count(distinct skills)=2
```

Output:

```
with a as (select * from (
values
    ('Marketing','Andrew','Excel'),
    ('Marketing','Andrew','SQL'),
    ('Marketing','Alex','Excel'),
    ('HR','John','Excel'),
    ('SDE','Allen','Power BI'),
    ('SDE','Allen','SQL'),
    ('SDE','Allen','Excel'),
    ('SDE','Allen','Python')
) as company (department, ename, skills))

select department, ename, count(distinct skills) from a
where skills not in  ('Excel','SQL')
group by department, ename
having count(distinct skills)=2
```

Data Output | Messages | Notifications

	department text	ename text	count bigint
1	SDE	Allen	2

-- Display the department which has more skilled employees

-- Query

```
with a as (select * from (
values
      ('Marketing','Andrew','Excel'),
      ('Marketing','Andrew','SQL'),
      ('Marketing','Alex','Excel'),
      ('HR','John','Excel'),
      ('SDE','Allen','Power BI'),
      ('SDE','Allen','SQL'),
      ('SDE','Allen','Excel'),
      ('SDE','Allen','Python')
) as company (department, ename, skills))

select department, ename, count(distinct skills) from a
group by department, ename
order by count(distinct skills) desc
```

Output:

```sql
with a as (select * from (
values
    ('Marketing','Andrew','Excel'),
    ('Marketing','Andrew','SQL'),
    ('Marketing','Alex','Excel'),
    ('HR','John','Excel'),
    ('SDE','Allen','Power BI'),
    ('SDE','Allen','SQL'),
    ('SDE','Allen','Excel'),
    ('SDE','Allen','Python')
) as company (department, ename, skills))

select department, ename, count(distinct skills) from a
group by department, ename
order by count(distinct skills) desc limit 1
```

Data Output | Messages | Notifications

	department text	ename text	count bigint
1	SDE	Allen	4

****

-- Find the paris of product combos in each order

Sample output:

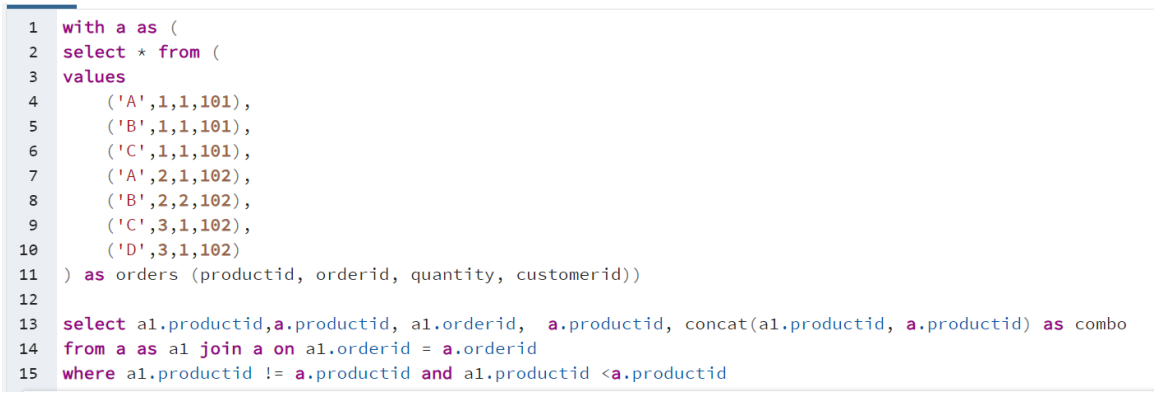
	productid text	productid text	orderid integer	productid text	combo text
1	A	C	1	C	AC
2	A	B	1	B	AB
3	B	C	1	C	BC
4	A	B	2	B	AB
5	C	D	3	D	CD

-- Query

```
with a as (
select * from (
values
      ('A',1,1,101),
      ('B',1,1,101),
      ('C',1,1,101),
      ('A',2,1,102),
      ('B',2,2,102),
      ('C',3,1,102),
      ('D',3,1,102)
) as orders (productid, orderid, quantity, customerid))

select a1.productid,a.productid, a1.orderid,  a.productid,
concat(a1.productid, a.productid) as combo
from a as a1 join a on a1.orderid = a.orderid
where a1.productid != a.productid and a1.productid <a.productid
```

Output:

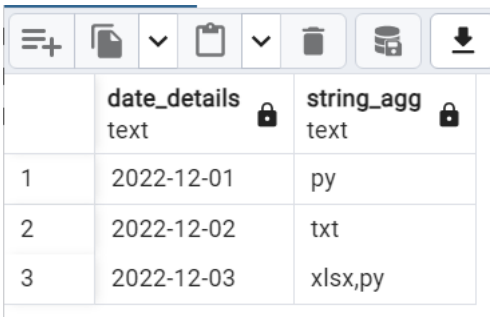


	productid text	productid text	orderid integer	productid text	combo text
1	A	C	1	C	AC
2	A	B	1	B	AB
3	B	C	1	C	BC
4	A	B	2	B	AB
5	C	D	3	D	CD

****

-- Find the file type with maximum changes on a given date

-- Sample output

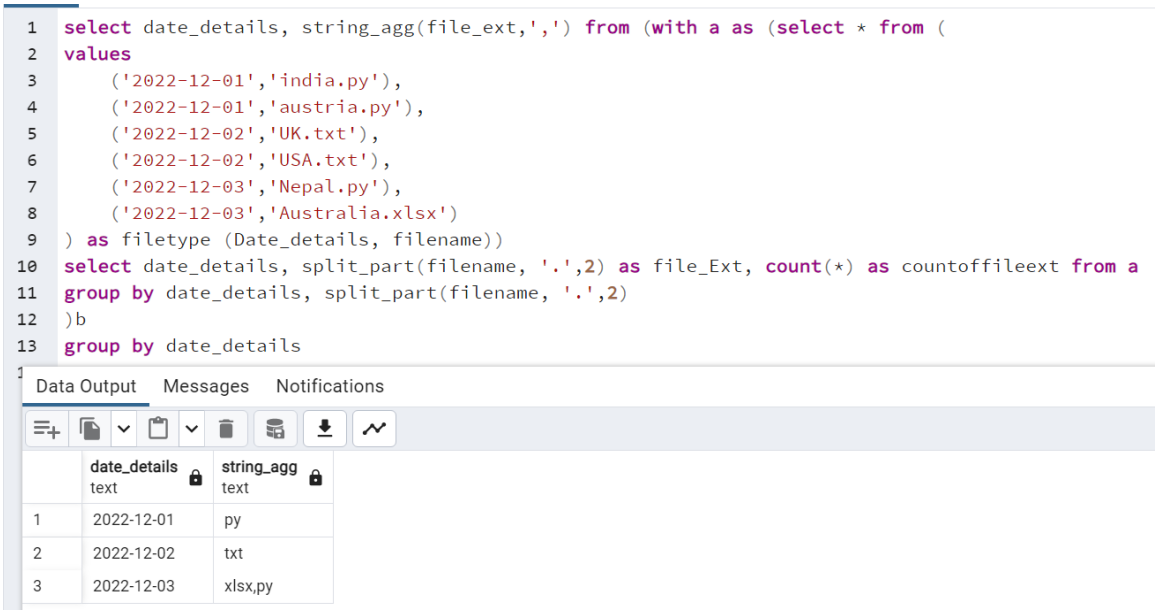


	date_details text	string_agg text
1	2022-12-01	py
2	2022-12-02	txt
3	2022-12-03	xlsx,py

-- Query

```
select date_details, string_agg(file_ext,',') from (with a as (select *
from (
values
      ('2022-12-01','india.py'),
      ('2022-12-01','austria.py'),
      ('2022-12-02','UK.txt'),
      ('2022-12-02','USA.txt'),
      ('2022-12-03','Nepal.py'),
      ('2022-12-03','Australia.xlsx')
) as filetype (Date_details, filename))
select date_details, split_part(filename, '.',2) as file_Ext, count(*) as
countoffileext from a
group by date_details, split_part(filename, '.',2)
)b
group by date_details
```

Output:



```
1  select date_details, string_agg(file_ext,',') from (with a as (select * from (
2  values
3      ('2022-12-01','india.py'),
4      ('2022-12-01','austria.py'),
5      ('2022-12-02','UK.txt'),
6      ('2022-12-02','USA.txt'),
7      ('2022-12-03','Nepal.py'),
8      ('2022-12-03','Australia.xlsx')
9  ) as filetype (Date_details, filename))
10 select date_details, split_part(filename, '.',2) as file_Ext, count(*) as countoffileext from a
11 group by date_details, split_part(filename, '.',2)
12 )b
13 group by date_details
```

Data Output | Messages | Notifications

	date_details text	string_agg text
1	2022-12-01	py
2	2022-12-02	txt
3	2022-12-03	xlsx,py

--Q: Display the filename which modified more number of times in a day.

-- Sample output

filename text	countoffileext text
USA.txt	2022-12-02 02:30:29
UK.txt	2022-12-02 12:30:30
india.py	2022-12-01 08:30:29
Nepal.py	2022-12-03 04:30:29

-- Query

```
select filename,countoffileext from (with a as (select * from (
values
      ('2022-12-01 06:30:29','india.py'),
      ('2022-12-01 08:30:29','india.py'),
      ('2022-12-02 12:30:29','UK.txt'),
      ('2022-12-02 02:30:29','USA.txt'),
      ('2022-12-03 04:30:29','Nepal.py'),
      ('2022-12-02 12:30:30','UK.txt')
) as filetype (Date_details, filename))
select filename, max(date_details) as countoffileext from a
group by filename
)b group by filename,countoffileext
```

Output:

```sql
select filename,countoffileext from (with a as (select * from (
values
    ('2022-12-01 06:30:29','india.py'),
    ('2022-12-01 08:30:29','india.py'),
    ('2022-12-02 12:30:29','UK.txt'),
    ('2022-12-02 02:30:29','USA.txt'),
    ('2022-12-03 04:30:29','Nepal.py'),
    ('2022-12-02 12:30:30','UK.txt')
) as filetype (Date_details, filename))
select filename, max(date_details) as countoffileext from a
group by filename
)b group by filename,countoffileext
```

Data Output | Messages | Notifications

	filename text	countoffileext text
1	USA.txt	2022-12-02 02:30:29
2	UK.txt	2022-12-02 12:30:30
3	india.py	2022-12-01 08:30:29
4	Nepal.py	2022-12-03 04:30:29

****

--Q Find the Students who scored more marks in Maths than Chemistry

-- Sample Output

std_name text	mathsmarks text	chemistrymarks text
Smith	95	90
Donna	98	90

-- Query

```sql
with a as (select * from (
values
      ('Allen','75 85 75'),
      ('Blake','80 85 88'),
      ('Smith','95 85 90'),
      ('Donna','98 85 90')
) as filetype (Std_name, Maths_Physics_Chemistry))

select std_name, split_part(Maths_Physics_Chemistry,' ',1) as mathsmarks,
```

```
split_part(Maths_Physics_Chemistry,' ',3) as Chemistrymarks
from a
where split_part(Maths_Physics_Chemistry,'
',1)>split_part(Maths_Physics_Chemistry,' ',3);
```

Output:

```
with a as (select * from (
values
    ('Allen','75 85 75'),
    ('Blake','80 85 88'),
    ('Smith','95 85 90'),
    ('Donna','98 85 90')
) as filetype (Std_name, Maths_Physics_Chemistry))

select std_name, split_part(Maths_Physics_Chemistry,' ',1) as mathsmarks,
split_part(Maths_Physics_Chemistry,' ',3) as Chemistrymarks
from a
where split_part(Maths_Physics_Chemistry,' ',1)>split_part(Maths_Physics_Chemistry,' ',3);
```

Data Output | Messages | Notifications

	std_name text	mathsmarks text	chemistrymarks text
1	Smith	95	90
2	Donna	98	90

-- Q Display the students who scored equal marks in both the subject.

-- Query

```
with a as (select * from (
values
      ('Allen','75 85 75'),
      ('Blake','80 85 88'),
      ('Smith','95 85 90'),
      ('Donna','98 85 90')
) as filetype (Std_name, Maths_Physics_Chemistry))

select std_name, split_part(Maths_Physics_Chemistry,' ',1) as mathsmarks,
split_part(Maths_Physics_Chemistry,' ',3) as Chemistrymarks
from a
where split_part(Maths_Physics_Chemistry,'
',1)=split_part(Maths_Physics_Chemistry,' ',3);
```

Note: Just by changing the logical operator in the where condition we can get the desired output.

Output:

```sql
with a as (select * from (
values
    ('Allen','75 85 75'),
    ('Blake','80 85 88'),
    ('Smith','95 85 90'),
    ('Donna','98 85 90')
) as filetype (Std_name, Maths_Physics_Chemistry))

select std_name, split_part(Maths_Physics_Chemistry,' ',1) as mathsmarks,
split_part(Maths_Physics_Chemistry,' ',3) as Chemistrymarks
from a
where split_part(Maths_Physics_Chemistry,' ',1)=split_part(Maths_Physics_Chemistry,' ',3);
```

Data Output | Messages | Notifications

	std_name text	mathsmarks text	chemistrymarks text
1	Allen	75	75

****

-- Q Students securing more marks than school average and less than class average.

-- Sample Output:

id integer	name text	sections text	marks integer	avg_school_mark numeric	avg_section_mark numeric
6	F	C	66	65.88	71.50

-- Query

```sql
select * from (with a as (select * from (
values
      (1,'A','A',80),(2,'B','A',80),
      (3,'C','B',35),(4,'D','B',50),
      (5,'E','B',60),(6,'F','C',66),
      (7,'G','C',77),(8,'H','D',79)
) as filetype (ID,NAME, SECTIONs, MARKS))
select *, round(avg(marks) over(order by marks rows
between unbounded preceding and unbounded following),2) as avg_school_mark,
round(avg(marks) over(partition by sections order by marks rows
between unbounded preceding and unbounded following),2) as avg_section_mark
from a) where marks > avg_school_mark and marks < avg_section_mark
```

Query:

```
select * from (with a as (select * from (
values
    (1,'A','A',80),(2,'B','A',80),
    (3,'C','B',35),(4,'D','B',50),
    (5,'E','B',60),(6,'F','C',66),
    (7,'G','C',77),(8,'H','D',79)
) as filetype (ID,NAME, SECTIONs, MARKS))
select *, round(avg(marks) over(order by marks rows
between unbounded preceding and unbounded following),2) as avg_school_mark,
round(avg(marks) over(partition by sections order by marks rows
between unbounded preceding and unbounded following),2) as avg_section_mark
from a) where marks > avg_school_mark and marks < avg_section_mark
```

Data Output | Messages | Notifications

	id integer	name text	sections text	marks integer	avg_school_mark numeric	avg_section_mark numeric
1	6	F	C	66	65.88	71.50

-- Display the student who scored less than school average and less than class average.

-- Sample output

id integer	name text	sections text	marks integer	avg_school_mark numeric	avg_section_mark numeric
3	C	B	35	65.88	48.33

-- Query

```
select * from (with a as (select * from (
values
      (1,'A','A',80),(2,'B','A',80),
      (3,'C','B',35),(4,'D','B',50),
      (5,'E','B',60),(6,'F','C',66),
      (7,'G','C',77),(8,'H','D',79)
) as filetype (ID,NAME, SECTIONs, MARKS))
select *, round(avg(marks) over(order by marks rows
between unbounded preceding and unbounded following),2) as avg_school_mark,
round(avg(marks) over(partition by sections order by marks rows
between unbounded preceding and unbounded following),2) as avg_section_mark
from a) where marks < avg_school_mark and marks < avg_section_mark
```

Output:

```sql
select * from (with a as (select * from (
values
    (1,'A','A',80),(2,'B','A',80),
    (3,'C','B',35),(4,'D','B',50),
    (5,'E','B',60),(6,'F','C',66),
    (7,'G','C',77),(8,'H','D',79)
) as filetype (ID,NAME, SECTIONs, MARKS))
select *, round(avg(marks) over(order by marks rows
between unbounded preceding and unbounded following),2) as avg_school_mark,
round(avg(marks) over(partition by sections order by marks rows
between unbounded preceding and unbounded following),2) as avg_section_mark
from a) where marks < avg_school_mark and marks < avg_section_mark
```

Data Output | Messages | Notifications

	id integer	name text	sections text	marks integer	avg_school_mark numeric	avg_section_mark numeric
1	3	C	B	35	65.88	48.33

****

--Q Actors and Directors combination in Indian movies.

-- Display the Actor and Director combination moives

-- Sample output

combomovies text
Madhavan - Mani Rathnam
Ram Charan - Raja Mouli

-- Query

```
select combomovies from( with a as (select * from (
values
      ('Madhavan','Mani
Rathnam','Alaipayudhe','Tamil'),('Madhavan','Mani
Rathnam','Guru','Hindi'),
      ('Ram Charan','Raja Mouli','RRR','Telugu'),('Ram Charan','Raja
Mouli','Mahadheera','Telugu'),
      ('Jr NTR','Raja Mouli','RRR','Telugu'),('Jr NTR','Director
X','Brindhavanam','Telugu'),
      ('SRK','Hirani','Dunky','Hindi'),('Amir
Khan','Hirani','PK','Hindi')
) as filetype (Actor,Director, Movie, Language))

select actor, director, count(*), concat(actor,' - ', director) as
```

```
combomovies from a
group by actor, director
having count(*)>1) b
```

Output:

```
select combomovies from( with a as (select * from (
values
    ('Madhavan','Mani Rathnam','Alaipayudhe','Tamil'),('Madhavan','Mani Rathnam','Guru','Hindi'),
    ('Ram Charan','Raja Mouli','RRR','Telugu'),('Ram Charan','Raja Mouli','Mahadheera','Telugu'),
    ('Jr NTR','Raja Mouli','RRR','Telugu'),('Jr NTR','Director X','Brindhavanam','Telugu'),
    ('SRK','Hirani','Dunky','Hindi'),('Amir Khan','Hirani','PK','Hindi')
) as filetype (Actor,Director, Movie, Language))

select actor, director, count(*), concat(actor,' - ', director) as combomovies from a
group by actor, director
having count(*)>1) b
```

Data Output | Messages | Notifications

	combomovies text
1	Madhavan - Mani Rathnam
2	Ram Charan - Raja Mouli

--Q Display the Actor, Director and Moviename who did more than one film.

-- Sample Output

	actor text	director text	movie text	language text
1	Madhavan	Mani Rathnam	Alaipayudhe	Tamil
2	Madhavan	Mani Rathnam	Guru	Hindi
3	Ram Charan	Raja Mouli	RRR	Telugu
4	Ram Charan	Raja Mouli	Mahadheera	Telugu

-- Query

```
select * into movies from (select * from (
values
      ('Madhavan','Mani
Rathnam','Alaipayudhe','Tamil'),('Madhavan','Mani
Rathnam','Guru','Hindi'),
      ('Ram Charan','Raja Mouli','RRR','Telugu'),('Ram Charan','Raja
Mouli','Mahadheera','Telugu'),
      ('Jr NTR','Raja Mouli','RRR','Telugu'),('Jr NTR','Director
X','Brindhavanam','Telugu'),
      ('SRK','Hirani','Dunky','Hindi'),('Amir
Khan','Hirani','PK','Hindi')
) as filetype (Actor,Director, Movie, Language)) a
```

```
select * from movies a where a.actor in (select b.actor from movies b where
a.actor=b.actor
                                                    group by b.actor,
b.director
                                                    having
count(*)>1)
```

Output:

```
select * into movies from (select * from (
values
    ('Madhavan','Mani Rathnam','Alaipayudhe','Tamil'),('Madhavan','Mani Rathnam','Guru','Hindi'),
    ('Ram Charan','Raja Mouli','RRR','Telugu'),('Ram Charan','Raja Mouli','Mahadheera','Telugu'),
    ('Jr NTR','Raja Mouli','RRR','Telugu'),('Jr NTR','Director X','Brindhavanam','Telugu'),
    ('SRK','Hirani','Dunky','Hindi'),('Amir Khan','Hirani','PK','Hindi')
) as filetype (Actor,Director, Movie, Language)) a

select * from movies a where a.actor in (select b.actor from movies b where a.actor=b.actor
                                         group by b.actor, b.director
                                         having count(*)>1)
```

Data Output | Messages | Notifications

	actor text	director text	movie text	language text
1	Madhavan	Mani Rathnam	Alaipayudhe	Tamil
2	Madhavan	Mani Rathnam	Guru	Hindi
3	Ram Charan	Raja Mouli	RRR	Telugu
4	Ram Charan	Raja Mouli	Mahadheera	Telugu

-- Display the Actor and director who did more than one movie in the same language.

-- Sample Output

	actor text	director text	movie text	language text
1	Ram Charan	Raja Mouli	RRR	Telugu
2	Ram Charan	Raja Mouli	Mahadheera	Telugu

-- Query

```
select * from movies a where a.actor in (select b.actor from movies b where
a.actor=b.actor and a.language = b.language
                                                    group by b.actor,
b.director, b.language
                                                    having
count(*)>1)
```

Output:

```
select * from movies a where a.actor in (select b.actor from movies b where a.actor=b.actor
                                         group by b.actor, b.director, b.language
                                         having count(*)>1)
```

Data Output | Messages | Notifications

	actor text	director text	movie text	language text
1	Ram Charan	Raja Mouli	RRR	Telugu
2	Ram Charan	Raja Mouli	Mahadheera	Telugu

-- Display the Director who directed the same hero in differenct languages.

-- Sample output

actor text	director text	movie text	language text
Ram Charan	Raja Mouli	RRR	Telugu
Ram Charan	Raja Mouli	Mahadheera	Telugu

-- Query

```
select * from movies a where a.actor in
(select b.actor from movies b where a.actor=b.actor and a.language
=b.language
group by b.director, b.actor)
```

Output:

```
select * from movies a where a.actor in
(select b.actor from movies b where a.actor=b.actor and a.language =b.language
group by b.director, b.actor)
```

Data Output | Messages | Notifications

	actor text	director text	movie text	language text
1	Ram Charan	Raja Mouli	RRR	Telugu
2	Ram Charan	Raja Mouli	Mahadheera	Telugu

-- Q Display the no of flims directed by Each director.

-- Query

```
select director, count(*) from movies
group by director
order by count(*) desc
```

Output:

```
10  select director, count(*) from movies
11  group by director
12  order by count(*) desc
```

Data Output | Messages | Notifications

	director text	count bigint
1	Raja Mouli	3
2	Hirani	3
3	Mani Rathnam	2
4	Director X	1

****