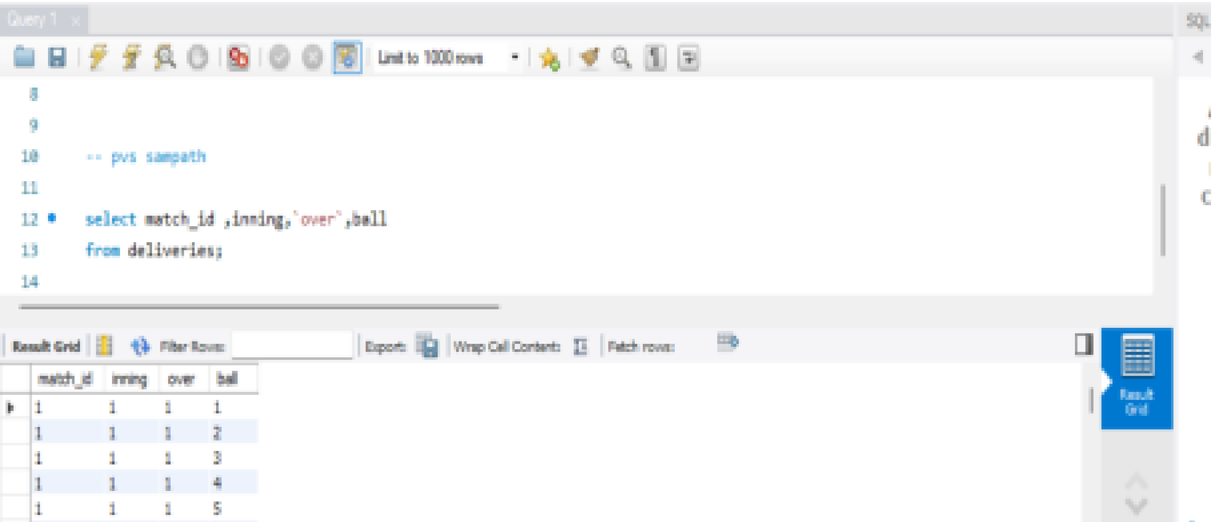
IPL Match Data Analysis using SQL

By Pvs sampath

1.Write a query to list all deliveries with their match_id, inning, over, and ball.

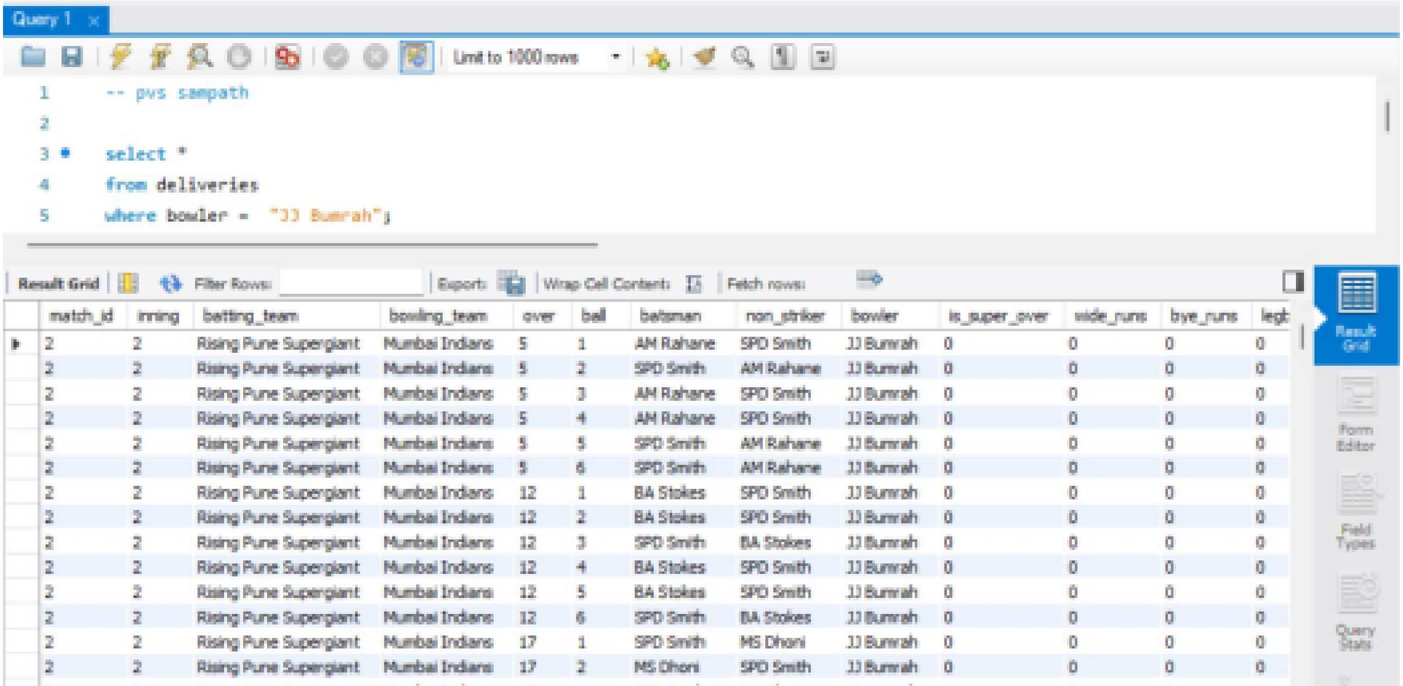


```
-- pvs sampath

select match_id ,inning,`over`,ball
from deliveries;
```

match_id	inning	over	ball
1	1	1	1
1	1	1	2
1	1	1	3
1	1	1	4
1	1	1	5

2.Write a query to display all columns for deliveries bowled by a specific bowler, JJ Bumrah.



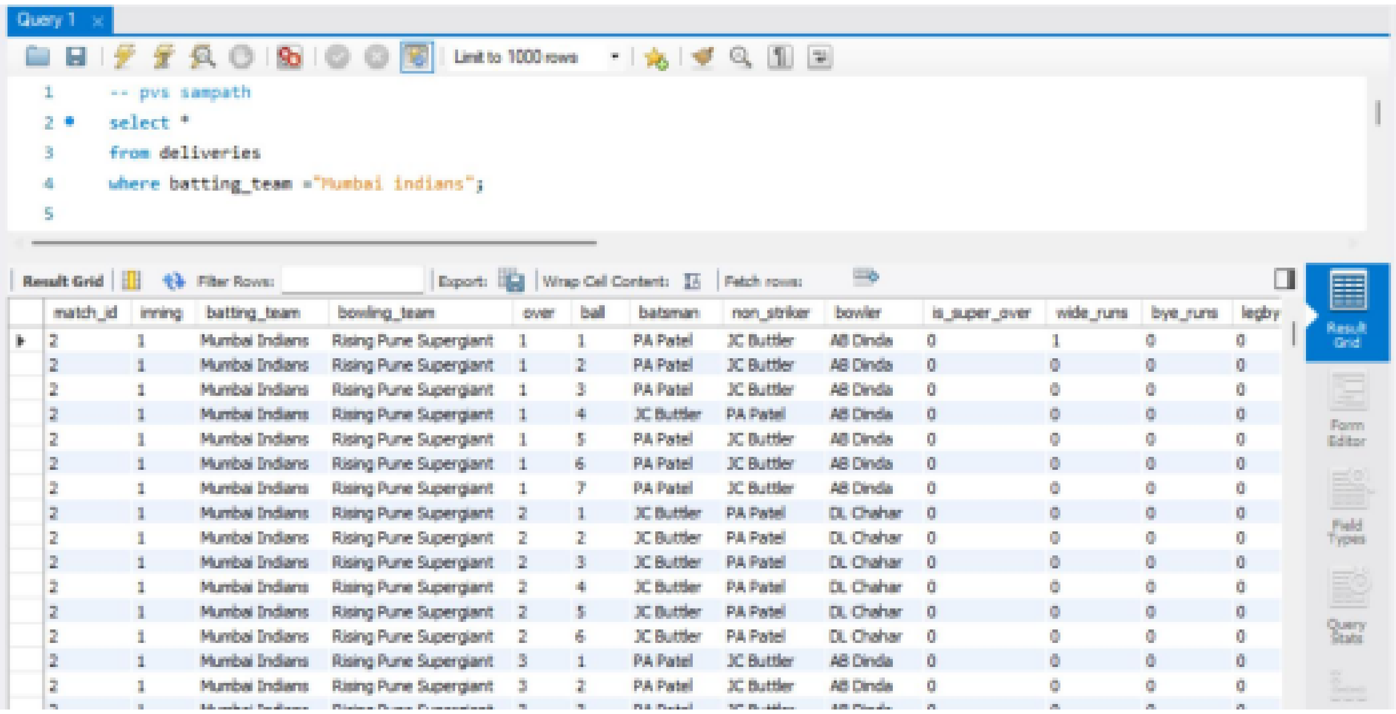
```
-- pvs sampath

select *
from deliveries
where bowler = "JJ Bumrah";
```

match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	wide_runs	bye_runs	legb
2	2	Rising Pune Supergiant	Mumbai Indians	5	1	AM Rahane	SPD Smith	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	5	2	SPD Smith	AM Rahane	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	5	3	AM Rahane	SPD Smith	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	5	4	AM Rahane	SPD Smith	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	5	5	SPD Smith	AM Rahane	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	5	6	SPD Smith	AM Rahane	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	12	1	BA Stokes	SPD Smith	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	12	2	BA Stokes	SPD Smith	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	12	3	SPD Smith	BA Stokes	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	12	4	BA Stokes	SPD Smith	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	12	5	BA Stokes	SPD Smith	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	12	6	SPD Smith	BA Stokes	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	17	1	SPD Smith	MS Dhoni	JJ Bumrah	0	0	0	0
2	2	Rising Pune Supergiant	Mumbai Indians	17	2	MS Dhoni	SPD Smith	JJ Bumrah	0	0	0	0

3.Write a query to list deliveries where the batting_team is

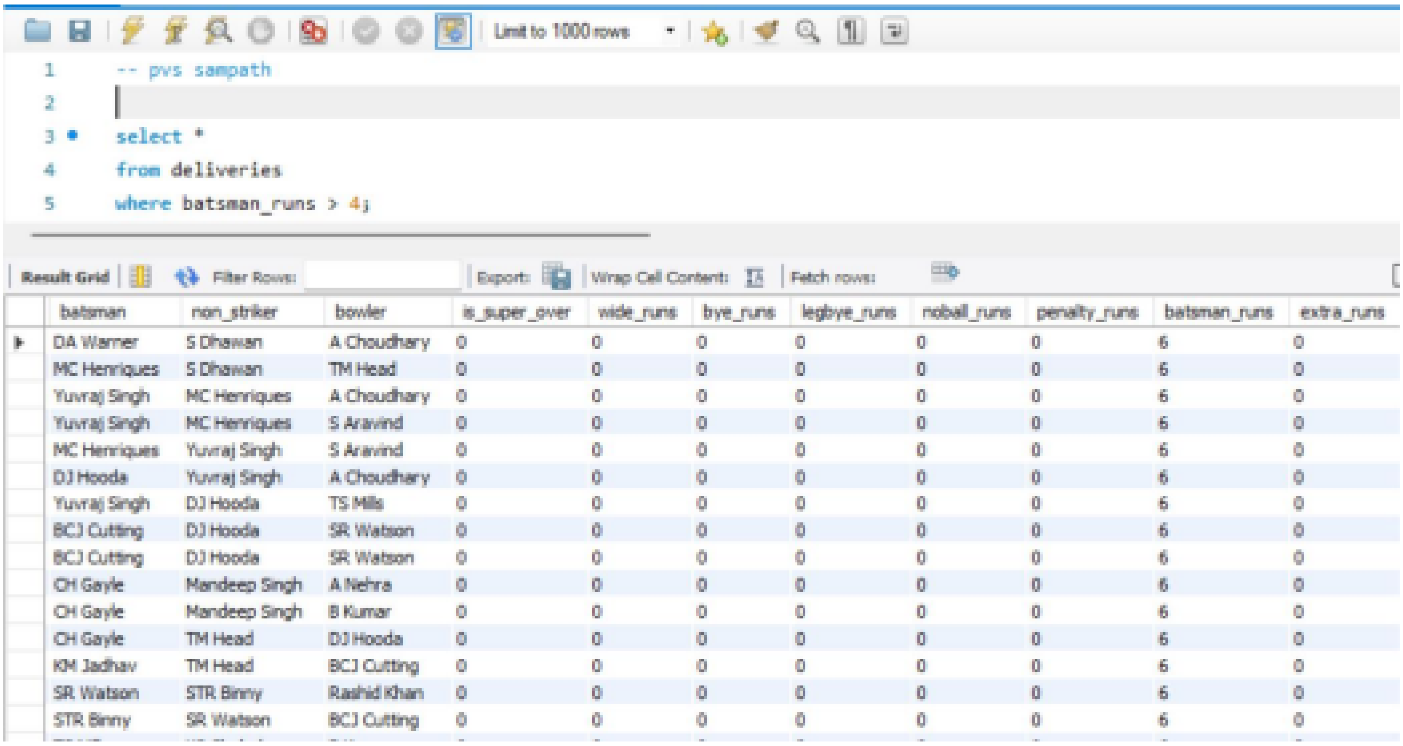
Mumbai Indians.



```
-- pvs sampath
select *
from deliveries
where batting_team ="Mumbai indians";
```

match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	wide_runs	bye_runs	legby
2	1	Mumbai Indians	Rising Pune Supergiant	1	1	PA Patel	JC Buttler	AB Dinda	0	1	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	1	2	PA Patel	JC Buttler	AB Dinda	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	1	3	PA Patel	JC Buttler	AB Dinda	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	1	4	JC Buttler	PA Patel	AB Dinda	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	1	5	PA Patel	JC Buttler	AB Dinda	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	1	6	PA Patel	JC Buttler	AB Dinda	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	1	7	PA Patel	JC Buttler	AB Dinda	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	2	1	JC Buttler	PA Patel	DL Chahar	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	2	2	JC Buttler	PA Patel	DL Chahar	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	2	3	JC Buttler	PA Patel	DL Chahar	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	2	4	JC Buttler	PA Patel	DL Chahar	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	2	5	JC Buttler	PA Patel	DL Chahar	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	2	6	JC Buttler	PA Patel	DL Chahar	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	3	1	PA Patel	JC Buttler	AB Dinda	0	0	0	0
2	1	Mumbai Indians	Rising Pune Supergiant	3	2	PA Patel	JC Buttler	AB Dinda	0	0	0	0

4.Write a query to find all deliveries where the batsman scored more than 4 runs off a single ball.

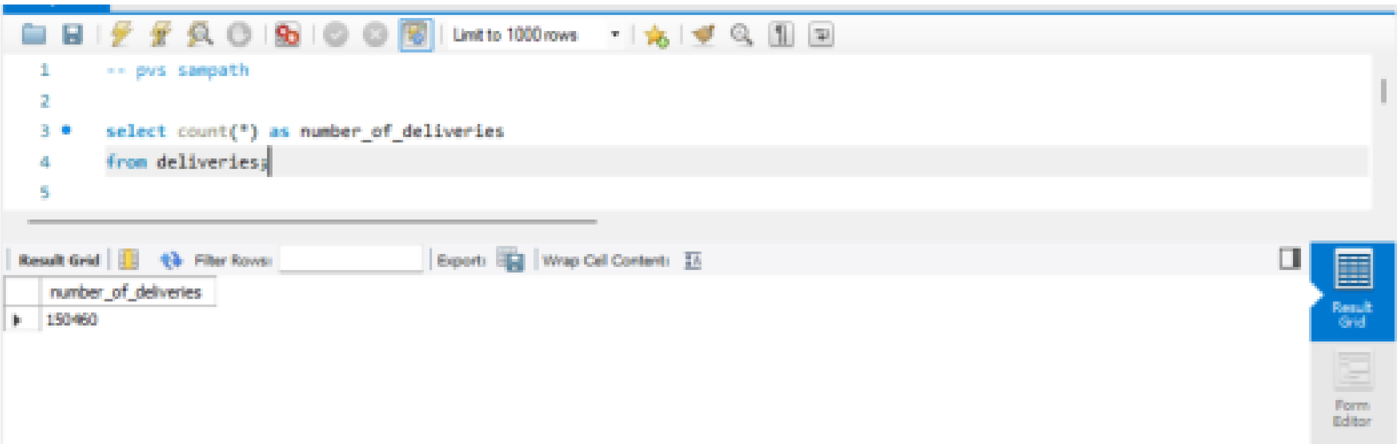


```
-- pvs sampath

select *
from deliveries
where batsman_runs > 4;
```

batsman	non_striker	bowler	is_super_over	wide_runs	bye_runs	legbye_runs	noball_runs	penalty_runs	batsman_runs	extra_runs
DA Warner	S Dhawan	A Choudhary	0	0	0	0	0	0	6	0
MC Henriques	S Dhawan	TM Head	0	0	0	0	0	0	6	0
Yuvraj Singh	MC Henriques	A Choudhary	0	0	0	0	0	0	6	0
Yuvraj Singh	MC Henriques	S Aravind	0	0	0	0	0	0	6	0
MC Henriques	Yuvraj Singh	S Aravind	0	0	0	0	0	0	6	0
DJ Hooda	Yuvraj Singh	A Choudhary	0	0	0	0	0	0	6	0
Yuvraj Singh	DJ Hooda	TS Mills	0	0	0	0	0	0	6	0
BCJ Cutting	DJ Hooda	SR Watson	0	0	0	0	0	0	6	0
BCJ Cutting	DJ Hooda	SR Watson	0	0	0	0	0	0	6	0
CH Gayle	Mandeep Singh	A Nehra	0	0	0	0	0	0	6	0
CH Gayle	Mandeep Singh	B Kumar	0	0	0	0	0	0	6	0
CH Gayle	TM Head	DJ Hooda	0	0	0	0	0	0	6	0
KM Jadhav	TM Head	BCJ Cutting	0	0	0	0	0	0	6	0
SR Watson	STR Binny	Rashid Khan	0	0	0	0	0	0	6	0
STR Binny	SR Watson	BCJ Cutting	0	0	0	0	0	0	6	0

5.Write a query to count the total number of deliveries in the dataset.

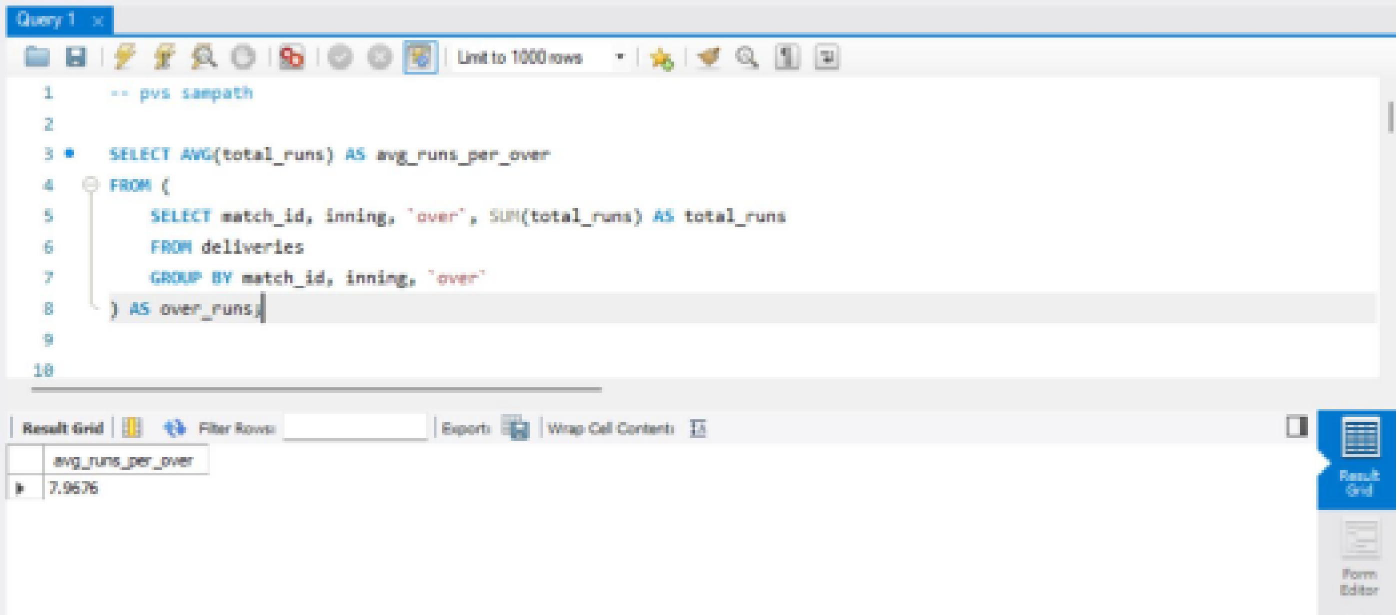


```
-- pvs sampath

select count(*) as number_of_deliveries
from deliveries;
```

```
number_of_deliveries
150460
```

6.Write a query to find the average number of runs scored per over

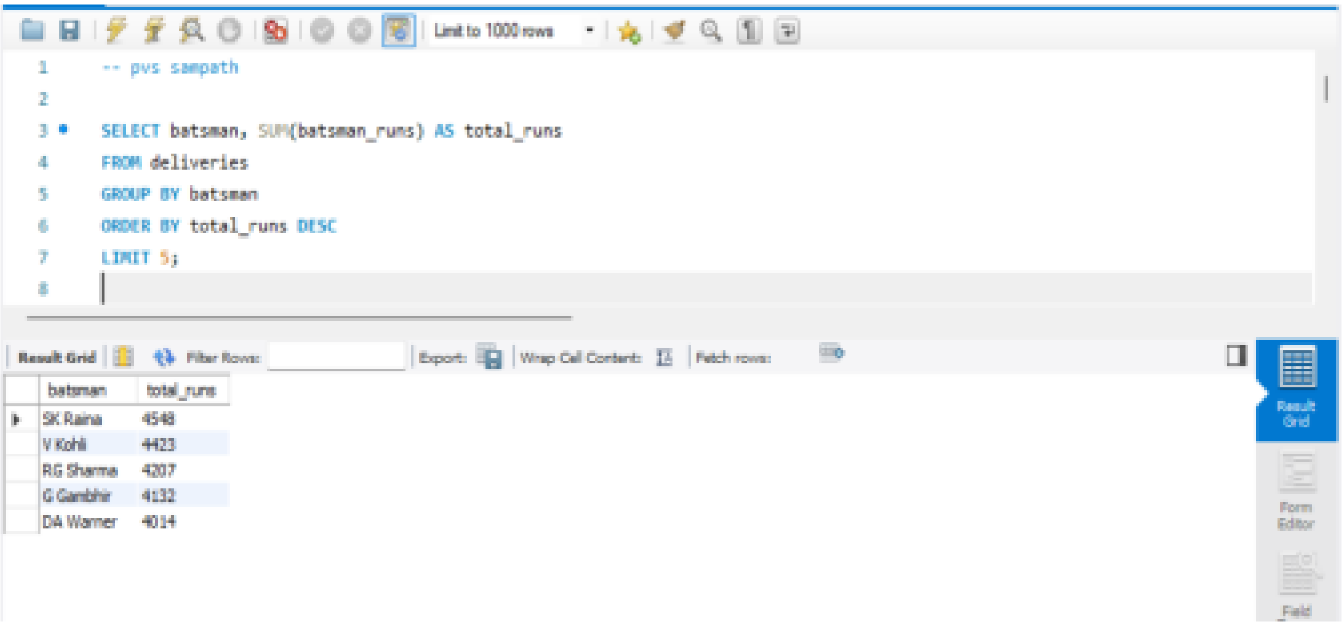


```
-- pvs sampath

SELECT AVG(total_runs) AS avg_runs_per_over
FROM (
    SELECT match_id, inning, `over`, SUM(total_runs) AS total_runs
    FROM deliveries
    GROUP BY match_id, inning, `over`
) AS over_runs;
```

```
avg_runs_per_over
7.9676
```

7.Write a query to list the top 5 batsmen with the highest total runs scored



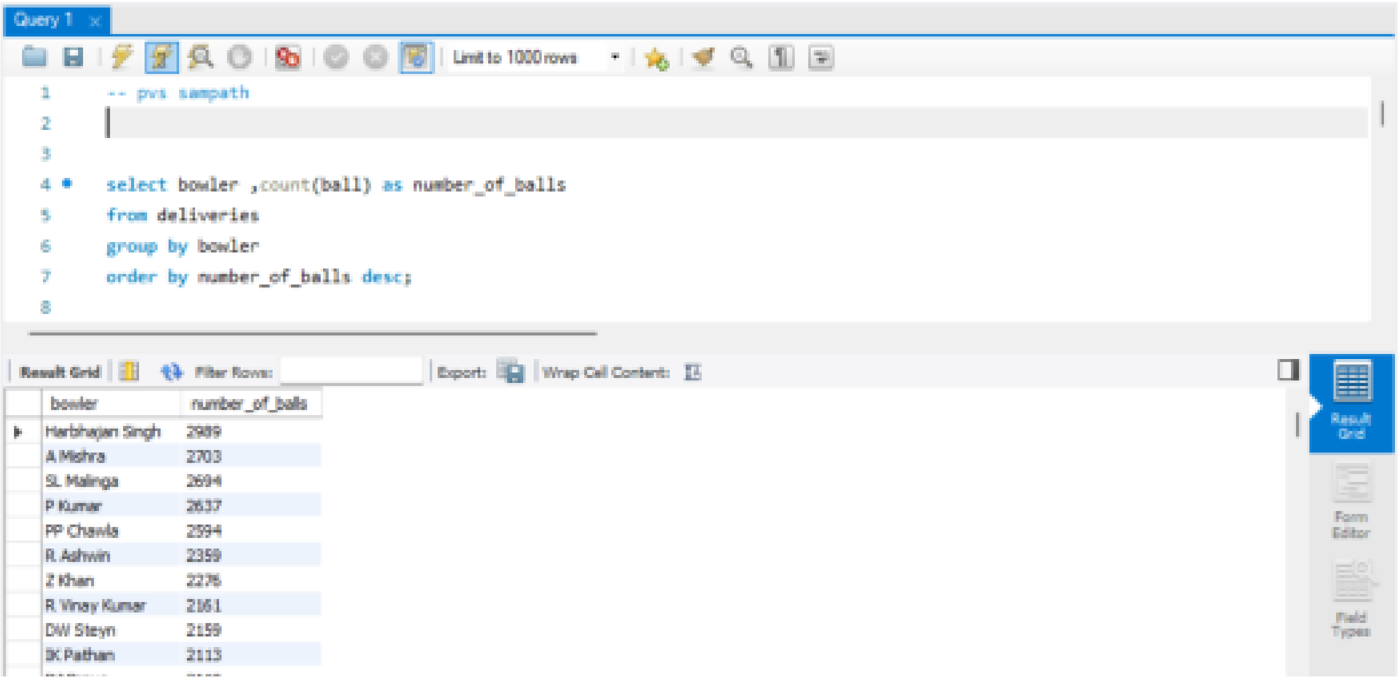
```sql
-- pvs sampath

SELECT batsman, SUM(batsman_runs) AS total_runs
FROM deliveries
GROUP BY batsman
ORDER BY total_runs DESC
LIMIT 5;
```

batsman	total_runs
SK Raina	4548
V Kohli	4423
RG Sharma	4207
G Gambhir	4132
DA Warner	4014

8.Write a query to list all pairs of bowlers who bowled in the same match. Include columns for match_id, bowler1, and bowler2.

9.Write a query to find the total number of deliveries bowled by each bowler. Display bowler and the count of deliveries.

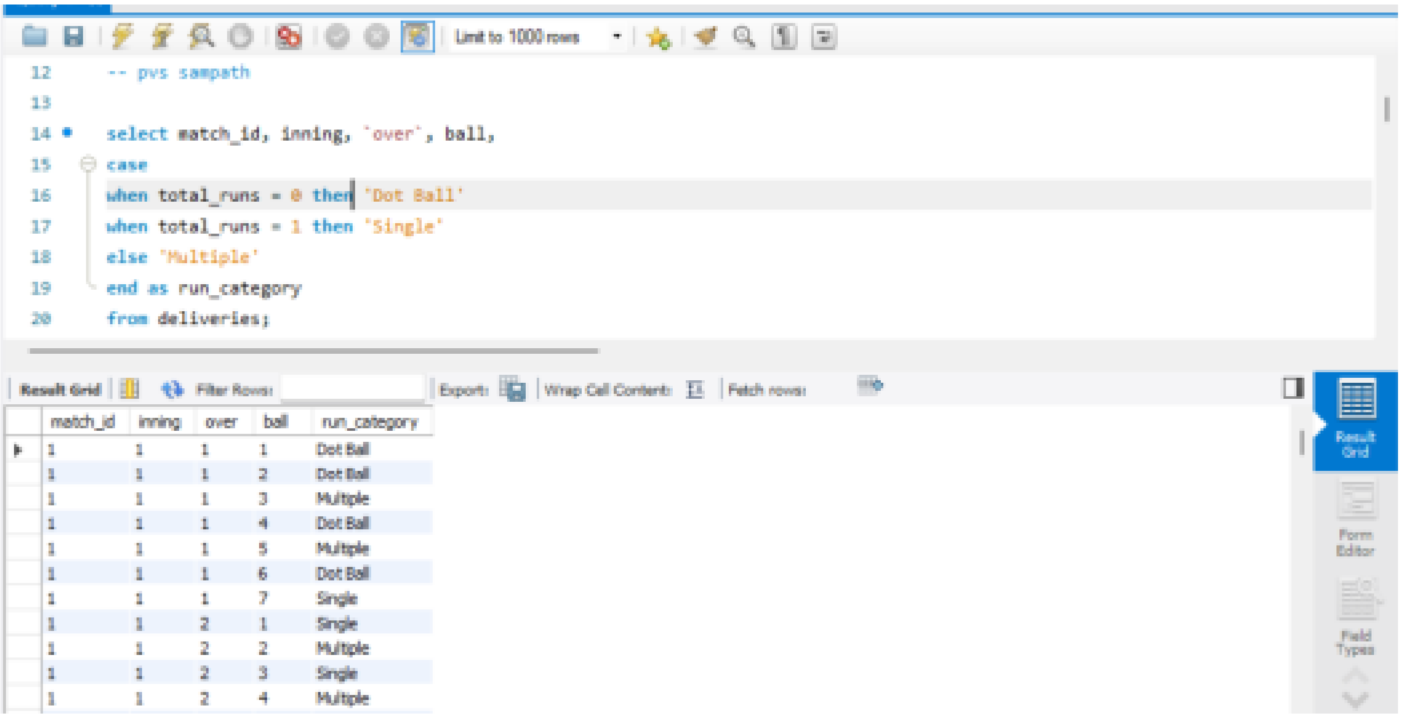


```sql
-- pvs sampath


select bowler ,count(ball) as number_of_balls
from deliveries
group by bowler
order by number_of_balls desc;
```

bowler	number_of_balls
Harbhajan Singh	2989
A Mishra	2703
SL Malinga	2694
P Kumar	2637
PP Chawla	2594
R Ashwin	2359
Z Khan	2276
R Vinay Kumar	2161
DW Steyn	2159
IK Pathan	2113

10. Write a query using a CASE statement to categorize deliveries into three categories based on runs scored: Dot Ball for 0 runs, Single for 1 run, and Multiple for more than 1 run.



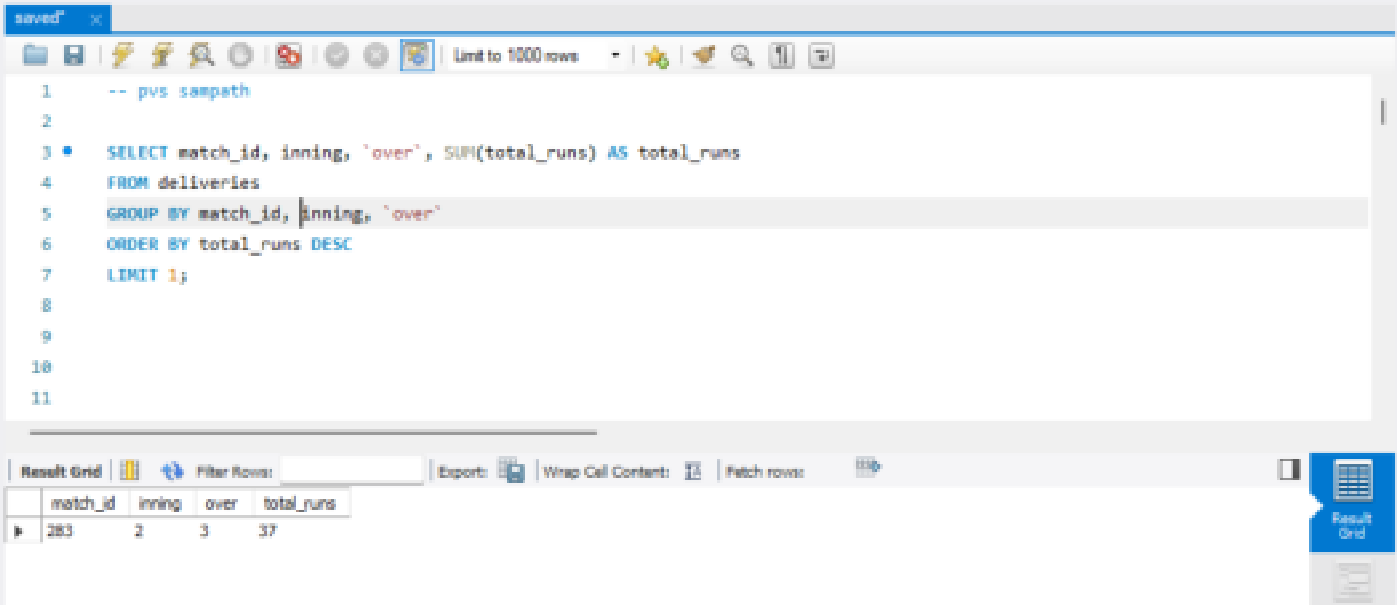
```sql
-- pvs sampath

select match_id, inning, `over`, ball,
case
when total_runs = 0 then 'Dot Ball'
when total_runs = 1 then 'Single'
else 'Multiple'
end as run_category
from deliveries;
```

match_id	inning	over	ball	run_category
1	1	1	1	Dot Ball
1	1	1	2	Dot Ball
1	1	1	3	Multiple
1	1	1	4	Dot Ball
1	1	1	5	Multiple
1	1	1	6	Dot Ball
1	1	1	7	Single
1	1	2	1	Single
1	1	2	2	Multiple
1	1	2	3	Single
1	1	2	4	Multiple

11. Write a query to add a new column is_boundary to the deliveries table that indicates if the delivery resulted in a boundary (4 or 6 runs).

12. Write a query using an advanced function to find the over with the highest total runs scored.

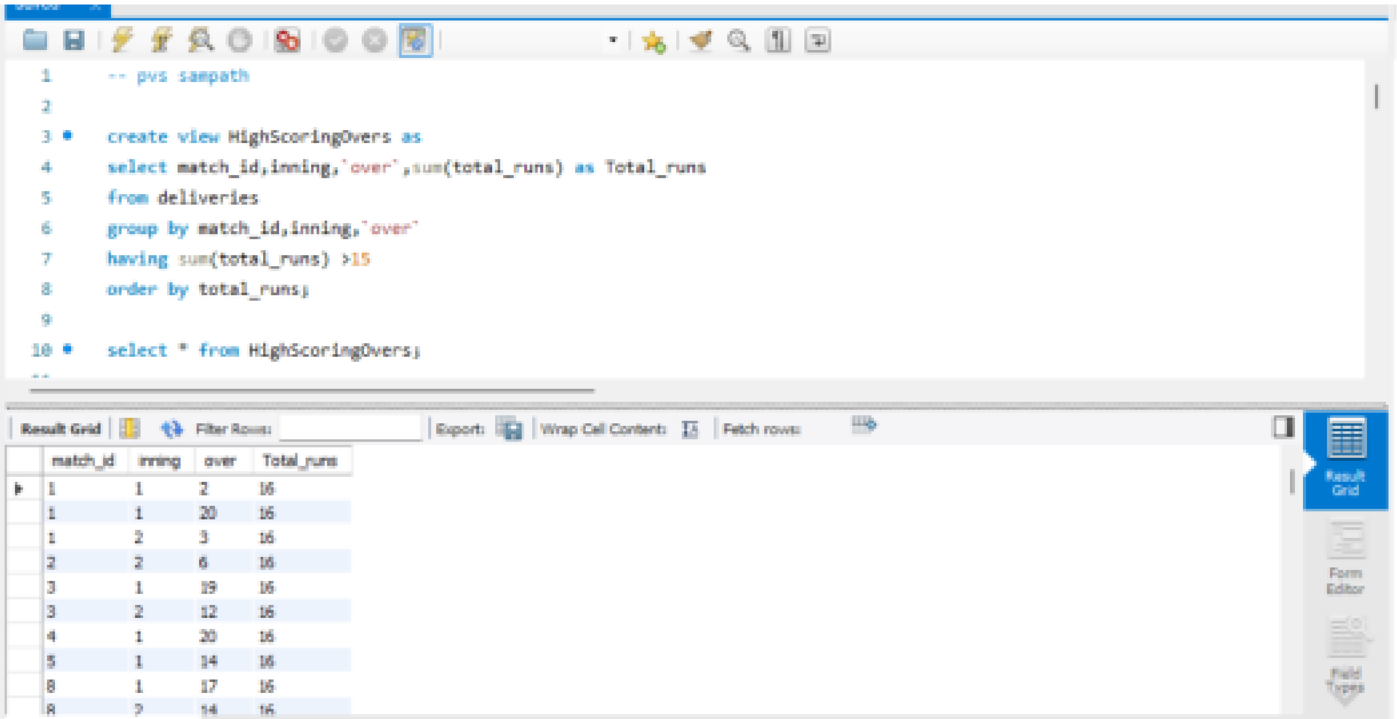


```sql
-- pvs sampath

SELECT match_id, inning, `over`, SUM(total_runs) AS total_runs
FROM deliveries
GROUP BY match_id, inning, `over`
ORDER BY total_runs DESC
LIMIT 1;
```

match_id	inning	over	total_runs
283	2	3	37

13. Create a view named HighScoringOvers that includes overs where the total runs scored are more than 15



```sql
-- pvs sampath

create view HighScoringOvers as
select match_id,inning,`over`,sum(total_runs) as Total_runs
from deliveries
group by match_id,inning,`over`
having sum(total_runs) >15
order by total_runs;

select * from HighScoringOvers;
```

match_id	inning	over	Total_runs
1	1	2	16
1	1	20	16
1	2	3	16
2	2	6	16
3	1	19	16
3	2	12	16
4	1	20	16
5	1	14	16
8	1	17	16
8	2	14	16

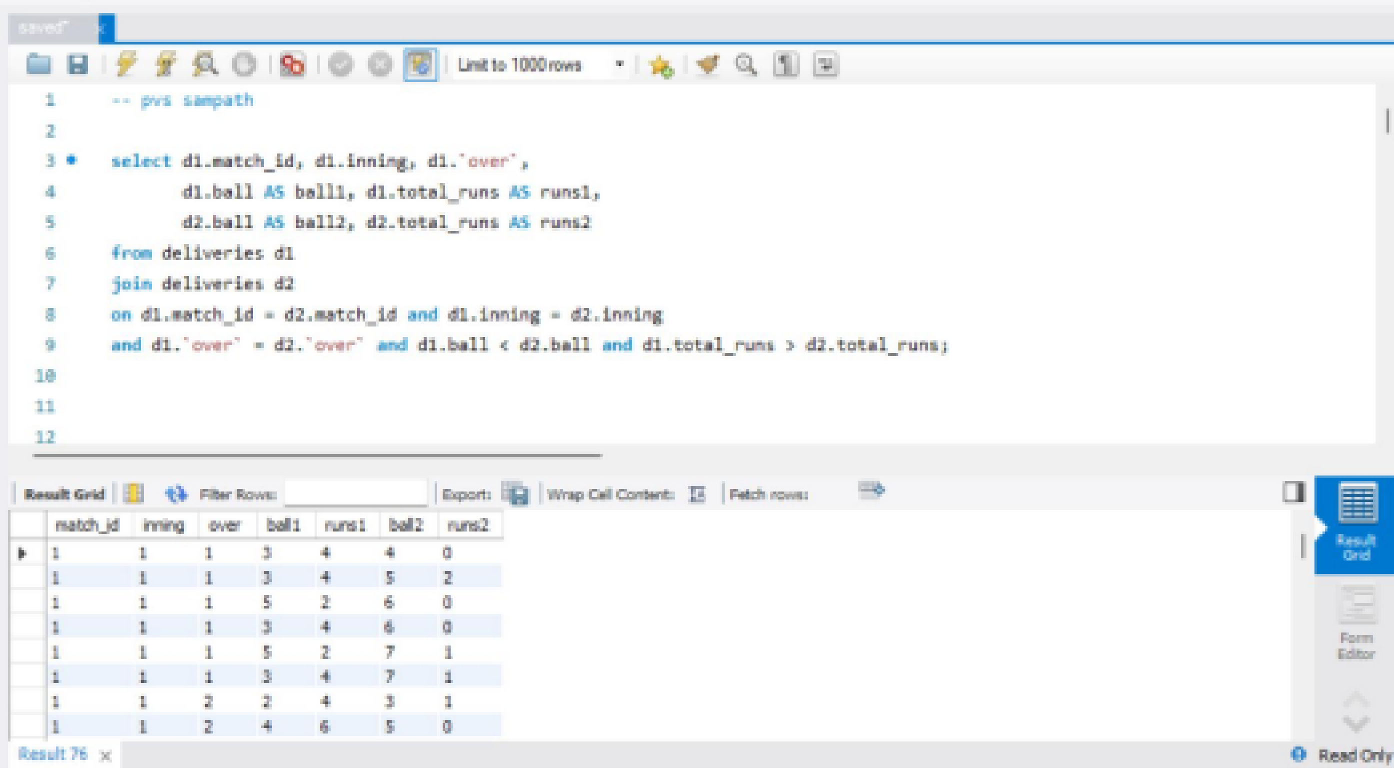
14. Write a query using a window function to rank deliveries based on their total_runs within each match_id.

15. Write a query to calculate the cumulative count of runs scored in each match sorted by over.

16. Write a stored procedure to update the runs scored on a delivery given its match_id, inning, over, ball, and new runs.

17. Write a query to find the bowler with the highest average runs conceded per over. Use subqueries and aggregate functions to achieve this.

18. Write a query to find pairs of deliveries in the same over where one delivery resulted in more runs than the other. Display columns for match_id, inning, over, ball1, runs1, ball2, and runs2.



```sql
-- pvs sampath

select d1.match_id, d1.inning, d1.`over`,
       d1.ball AS ball1, d1.total_runs AS runs1,
       d2.ball AS ball2, d2.total_runs AS runs2
from deliveries d1
join deliveries d2
on d1.match_id = d2.match_id and d1.inning = d2.inning
and d1.`over` = d2.`over` and d1.ball < d2.ball and d1.total_runs > d2.total_runs;
```

match_id	inning	over	ball1	runs1	ball2	runs2
1	1	1	3	4	4	0
1	1	1	3	4	5	2
1	1	1	5	2	6	0
1	1	1	3	4	6	0
1	1	1	5	2	7	1
1	1	1	3	4	7	1
1	1	2	2	4	3	1
1	1	2	4	6	5	0