1 . List the different dtypes of columns in table “ball\_by\_ball” (using information schema)

Ans. The different data types of columns in the ball\_by\_ball table are **INT**.

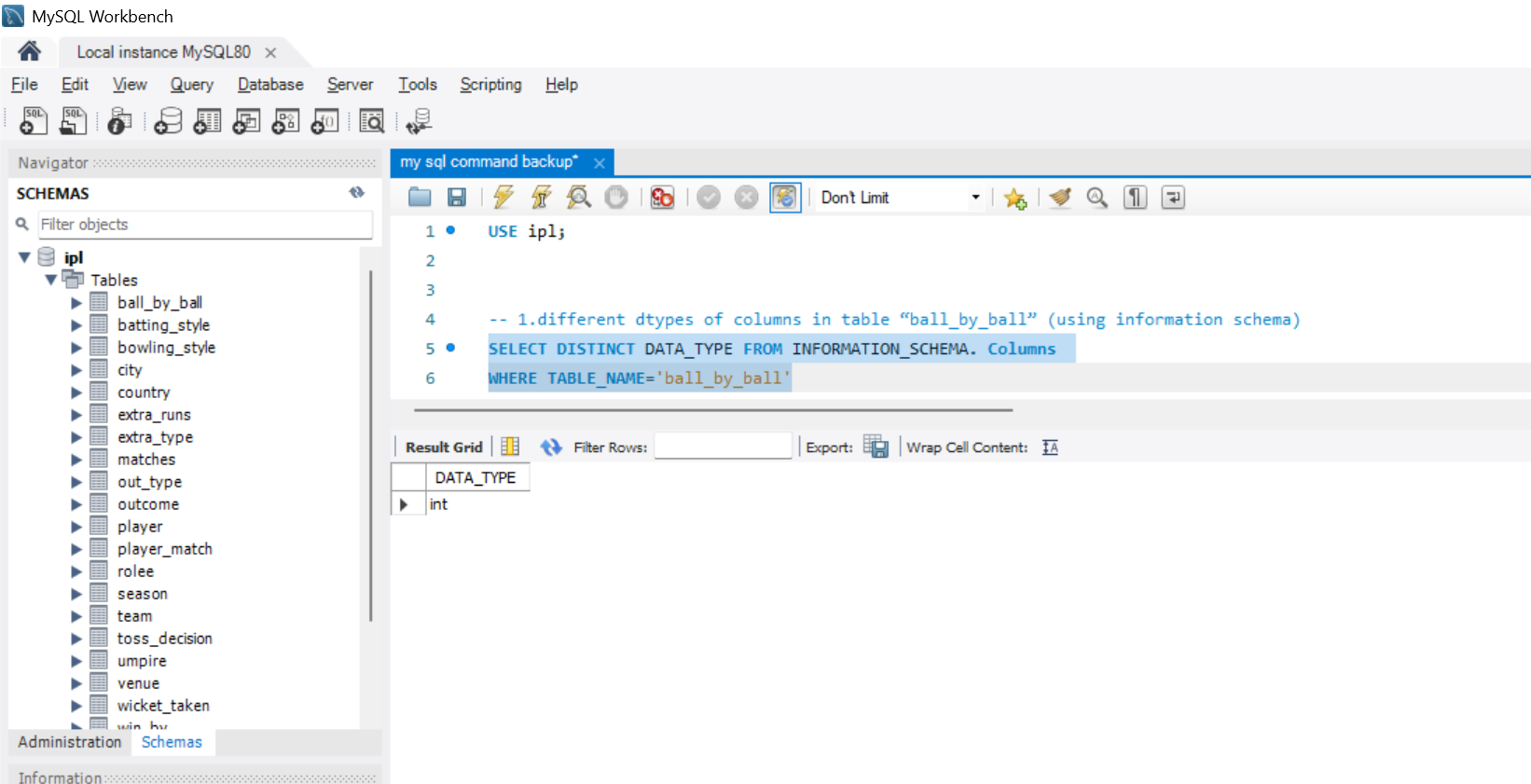
**Query used:**

**SELECT DISTINCT DATA\_TYPE**

**FROM INFORMATION\_SCHEMA.COLUMNS**

**WHERE TABLE\_NAME = 'ball\_by\_ball';**

**Result :**



2 .What is the total number of run scored in 1st season by RCB (bonus : also include the extra runs using the extra runs table)

Ans. The total number of runs scored by **Royal Challengers Bangalore (RCB)** in their **first season**, including **extra runs**, is **2558 runs**.

**Query used:**

**with cte as (**

**select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,**

**b.Team\_Batting,t.Team\_Name as Team\_Batting\_Name,**

**b.Team\_Bowling,t1.Team\_Name as Team\_Bowling\_Name,**

**m.Season\_Id,(b.Runs\_Scored+coalesce(e.Extra\_Runs,0)) as Total\_Runs\_Scored**

**from Ball\_by\_Ball b join Team t**

**on t.Team\_Id=b.Team\_Batting**

**join Team t1 on t1.Team\_Id=b.Team\_Bowling**

**join Matches m on m.Match\_Id=b.Match\_Id**

**left join Extra\_Runs e on e.Match\_Id=b.Match\_Id and**

**e.Over\_Id=b.Over\_Id and e.Ball\_Id=b.Ball\_Id and**

**e.Innings\_No=b.Innings\_No**

**),**

**cte1 as(**

**select \* from cte where Team\_Batting\_Name='Royal Challengers Bangalore' and Season\_Id=(select min(Season\_Id) from cte)**

**)**

**select Team\_Batting\_Name,sum(Total\_Runs\_Scored) as Total\_Runs\_Scored\_Season\_1 from cte1**

**group by Team\_Batting\_Name**

**Result :** 

3 .How many players were more than age of 25 during 2014.

Ans. 89 players have more than age of 25 during IPL 2014.

**Query used:**

**with age\_table as (**

**select Player\_Id,Player\_Name, timestampdiff(year,DOB,'2014-01-01') as age from Player**

**),**

**cte as (**

**select Match\_Id,Player\_Id from Player\_Match where Match\_Id in (**

**select distinct Match\_Id from Matches where Season\_Id=(select Season\_Id from Season where Season\_Year=2014)))**

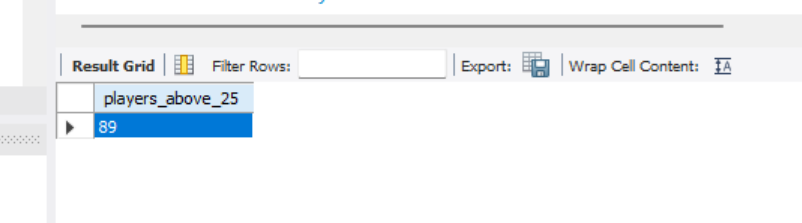
**select count(distinct c.player\_id) as players\_above\_25**

**from cte c**

**join age\_table a on c.player\_id = a.player\_id**

**where a.age > 25;**

**Result :**



4 . How many matches did RCB win in season 2013 ?

Ans. RCB win 9 matches in season 2013.

**Query used:**

**with cte as (**

**select m.Match\_Id,m.Match\_Winner,m.Season\_Id,s.Season\_Year from Matches m**

**join Season s on m.Season\_Id=s.Season\_Id**

**where Season\_Year=2013**

**),**

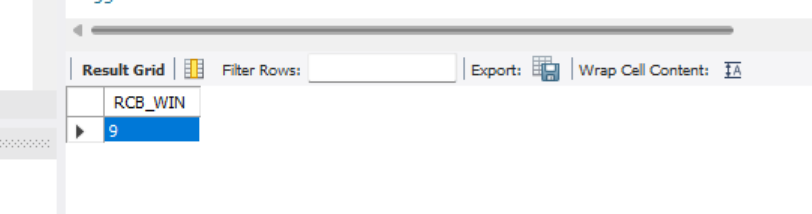
**cte1 as (select c.Match\_Id,c.Season\_Year,t.Team\_Name from cte c**

**join Team t on c.Match\_Winner=t.Team\_Id**

**where Team\_Name='Royal Challengers Bangalore')**

**select count(distinct Match\_Id) as RCB\_WIN from cte1**

**Result :**



5. List top 10 players according to their strike rate in last 4 seasons

Ans. Please find the top 10 players according to their strike rate in the last 4 seasons in the table below.

Strike Rate = (Total Runs Scored / Balls faced) \* 100

**Query used:**

**SELECTp.Player\_Name,**

**ROUND(SUM(b.Runs\_Scored) \* 100.0 / COUNT(b.Ball\_Id), 2) AS Strike\_Rate FROM Ball\_by\_Ball b**

**JOIN Matches m ON b.Match\_Id = m.Match\_Id**

**JOIN**

**Season s ON m.Season\_Id = s.Season\_Id**

**JOIN Player p ON b.Striker = p.Player\_Id**

**WHERE s.Season\_Year >= (**

**SELECT MAX(Season\_Year) - 3**

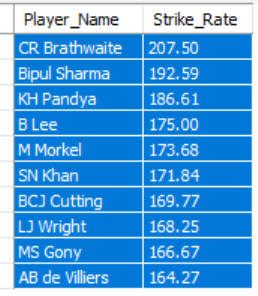
**FROM Season)**

**GROUP BY p.Player\_Id**

**HAVING COUNT(b.Ball\_Id) >= 10 — optional: avoid very low ball counts**

**ORDER BY Strike\_Rate DESC LIMIT 10;**

**Result :**

****

6. What is the average runs scored by each batsman considering all the seasons?

Ans. To found List for average runs scored by each batsman we use query shown below.

**Query used:**

**with cte as (**

**select p.player\_id,p.player\_name,b.runs\_scored,b.match\_id**

**from player p**

**left join ball\_by\_ball b**

**on p.player\_id = b.striker**

**)**

**select player\_name,**

**coalesce(sum(runs\_scored), 0) as total\_runs,**

**count(distinct match\_id) as matches\_played,**

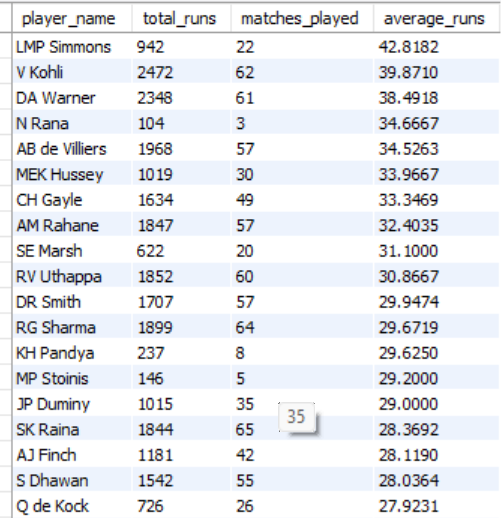
**coalesce(sum(runs\_scored) / nullif(count(distinct match\_id), 0), 0) as average\_runs**

**from cte**

**group by player\_id, player\_name**

**order by average\_runs desc**

**Result :**



7. What are the average wickets taken by each bowler considering all the seasons?

Ans. List for average wickets taken by each bowler find using query shown below.

Average Wickets = (Total Wickets Taken) / (Number of Matches)

**Query used:**

**with bowling\_skills as (**

**select p.Player\_Id,p.Player\_Name,b.Bowling\_skill**

**from Player p join Bowling\_Style b**

**on p.Bowling\_skill=b.Bowling\_Id**

**),**

**cte as (**

**select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler,bs.Player\_Name as Bowler\_Name,bs.Bowling\_skill**

**from Ball\_by\_Ball b join Wicket\_Taken w**

**on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id**

**and b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No**

**join bowling\_skills bs on bs.Player\_Id=b.Bowler**

**),**

**cte2 as (**

**SELECT Bowler\_Name, Bowling\_skill,**

**count(distinct Match\_Id) as Total\_Matches,**

**COUNT(\*) AS Total\_Wickets\_Taken**

**FROM cte**

**GROUP BY Bowler\_Name, Bowling\_skill**

**ORDER BY Total\_Wickets\_Taken DESC**

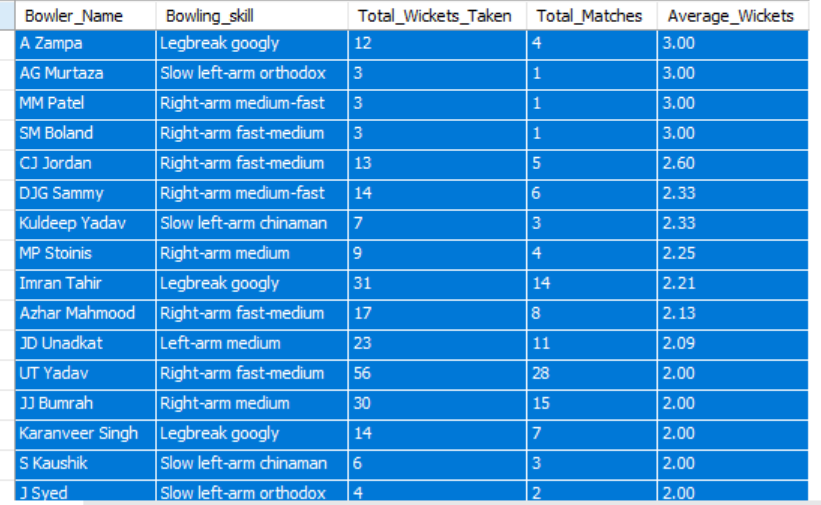
**)**

**select Bowler\_Name,Bowling\_skill,Total\_Wickets\_Taken,Total\_Matches,round((Total\_Wickets\_Taken)/(Total\_Matches),2) as Average\_Wickets**

**from cte2**

**order by (Total\_Wickets\_Taken)/(Total\_Matches) desc**

**Result :**



8.List all the players who have average runs scored greater than overall average and who have taken wickets greater than overall average.

Ans. We’ll solve this in three parts using Common Table Expressions where we first Calculate average runs per batsman then we calculate total wickets per bowler and then Calculate overall averages for both.

**Query used:**

**WITH batting\_cte AS (**

**SELECT p.player\_id,p.player\_name,**

**COALESCE(SUM(b.runs\_scored), 0) AS total\_runs,**

**COUNT(DISTINCT b.match\_id) AS matches\_played,**

**COALESCE( SUM(b.runs\_scored) / NULLIF(COUNT(DISTINCT b.match\_id), 0),0) AS average\_runs**

**FROM player p**

**LEFT JOIN ball\_by\_ball b**

**ON p.player\_id = b.striker**

**GROUP BY p.player\_id, p.player\_name**

**),**

**above\_avg\_batting AS (**

**SELECT \***

**FROM batting\_cte**

**WHERE average\_runs > (SELECT AVG(average\_runs) FROM batting\_cte)**

**),**

**bowling\_skills AS (**

**SELECT p.player\_id,p.player\_name,bs.bowling\_skill FROM player p**

**JOIN bowling\_style bs ON p.bowling\_skill = bs.bowling\_id**

**),**

**bowling\_cte AS (**

**SELECT b.match\_id,b.bowler,bs.player\_name,**

**bs.bowling\_skill**

**FROM ball\_by\_ball b**

**JOIN wicket\_taken w**

**ON b.match\_id = w.match\_id**

**AND b.over\_id = w.over\_id**

**AND b.ball\_id = w.ball\_id**

**AND b.innings\_no = w.innings\_no**

**JOIN bowling\_skills bs**

**ON bs.player\_id = b.bowler**

**),**

**bowling\_stats AS (**

**SELECT bowler AS player\_id,player\_name,**

**COUNT(DISTINCT match\_id) AS total\_matches,**

**COUNT(\*) AS total\_wickets,**

**ROUND(COUNT(\*) / COUNT(DISTINCT match\_id), 2) AS average\_wickets**

**FROM bowling\_cte**

**GROUP BY bowler, player\_name**

**),**

**above\_avg\_bowling AS (**

**SELECT \* FROM bowling\_stats**

**WHERE average\_wickets > (SELECT AVG(average\_wickets) FROM bowling\_stats)**

**)SELECT b.player\_name,b.average\_runs,w.total\_wickets,w.average\_wickets**

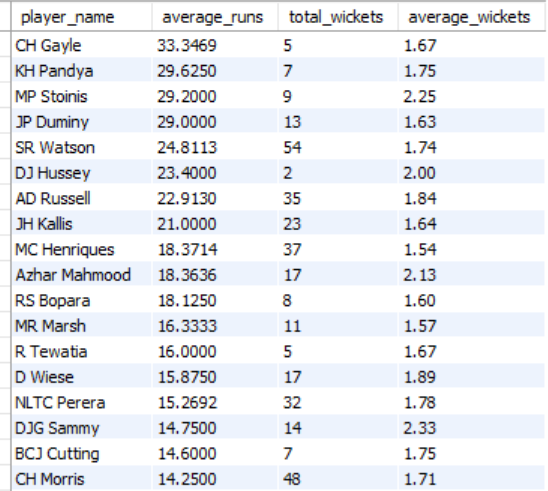
**FROM above\_avg\_batting b**

**JOIN above\_avg\_bowling w**

**ON b.player\_id = w.player\_id**

**ORDER BY b.average\_runs DESC, w.average\_wickets DESC;**

**Result :**



9. Create a table rcb\_record table that shows wins and losses of RCB in an individual venue.

Ans - Please find rcb\_record table that shows wins and losses of RCB in an individual venue below.-

**Query used:**

**with cte as**

**(select m.Match\_Id,m.Team\_1,t.Team\_Name as team1,m.Team\_2,t1.Team\_Name as team2,m.Match\_Winner,t2.Team\_Name as winner,m.Venue\_Id,v.Venue\_Name**

**from Matches m join Team t on m.Team\_1=t.Team\_Id**

**join Team t1 on m.Team\_2=t1.Team\_Id**

**join Team t2 on m.Match\_Winner=t2.Team\_Id**

**join Venue v on v.Venue\_Id=m.Venue\_Id**

**),**

**cte1 as**

**(select Match\_Id,team1,team2,winner,Venue\_Name from cte**

**where team1='Royal Challengers Bangalore' or team2='Royal Challengers Bangalore')**

**select Venue\_Name, count(case when winner='Royal Challengers Bangalore' then 1 end) as win\_count,**

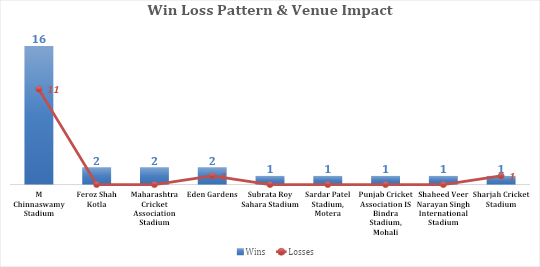
**COUNT(CASE WHEN winner != 'Royal Challengers Bangalore' THEN 1 END) AS loss\_count**

**from cte1**

**group by Venue\_Name;**

**Result :**





10. What is the impact of bowling style on wickets taken.

Ans. The analysis demonstrates that bowling style has a significant impact on the number of wickets taken. Right-arm medium bowlers have taken the highest number of wickets, indicating their effectiveness and frequent utilization. This is followed by right-arm fast-medium and right-arm medium-fast bowlers, who also contribute substantially to wicket-taking.

**Query used:**

**with bowling\_skills as (**

**select p.Player\_Id,p.Player\_Name,b.Bowling\_skill**

**from Player p join Bowling\_Style b**

**on p.Bowling\_skill=b.Bowling\_Id**

**),**

**cte as (**

**select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler,**

**bs.Player\_Name as Bowler\_Name,bs.Bowling\_skill**

**from Ball\_by\_Ball b join Wicket\_Taken w**

**on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id**

**and b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No**

**join bowling\_skills bs on bs.Player\_Id=b.Bowler**

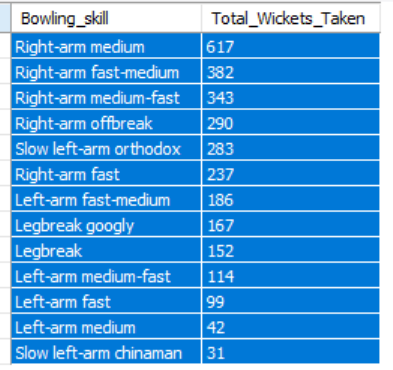
**)**

**select Bowling\_skill,count(\*) as Total\_Wickets\_Taken**

**from cte group by Bowling\_skill**

**order by count(\*) desc**

**Result :**



11. Write the sql query to provide a status of whether the performance of the team better than the previous year performance on the basis of number of runs scored by the team in the season and number of wickets taken

Ans. -- **Query used:**

—Number of Runs Scored year wise

**with cte as (**

**select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Team\_Batting,**

**(b.Runs\_Scored + IFNULL(e.Extra\_Runs, 0)) AS Total\_Runs**

**from Ball\_by\_Ball b left join Extra\_Runs e**

**on b.Match\_Id=e.Match\_Id and**

**b.Over\_Id=e.Over\_Id and b.Ball\_Id=e.Ball\_Id and**

**b.Innings\_No=e.Innings\_No**

**),**

**cte1 as (**

**select c.Match\_Id,year(m.Match\_Date) as Year,c.Over\_Id,c.Ball\_Id,c.Innings\_No,c.Team\_Batting,c.Total\_Runs,t.Team\_Name**

**from cte c join Matches m on c.Match\_Id=m.Match\_Id**

**join Team t on t.Team\_Id=c.Team\_Batting)**

**select team\_name,**

**sum(case when year = 2013 then total\_runs else 0 end) as "2013",**

**sum(case when year = 2014 then total\_runs else 0 end) as "2014",**

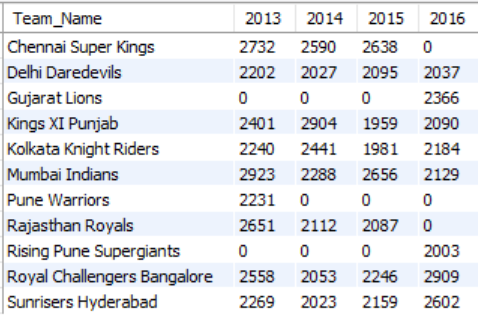
**sum(case when year = 2015 then total\_runs else 0 end) as "2015",**

**sum(case when year = 2016 then total\_runs else 0 end) as "2016"**

**from cte1**

**group by team\_name order by team\_name;**

**Result :**



**Query used:**

-- —-Number of Wickets Taken Yearwise by each Team

**with cte as**

**(select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler,b.Team\_Bowling**

**from Ball\_by\_Ball b join Wicket\_Taken w**

**on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id and**

**b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No),**

**cte1 as**

**(select c.Match\_Id,year(m.Match\_Date) as Year,c.Team\_Bowling,**

**t.Team\_Name**

**from cte c join Matches m on c.Match\_Id=m.Match\_Id**

**join Team t on c.Team\_Bowling=t.Team\_Id**

**),**

**cte2 as**

**(select Team\_Name,Year,count(\*) as Total\_Wickets\_Taken**

**from cte1**

**group by Team\_Name,Year)**

**select Team\_Name,**

**sum(case when Year=2013 then Total\_Wickets\_Taken else 0 end) as "2013",**

**sum(case when Year=2014 then Total\_Wickets\_Taken else 0 end) as "2014",**

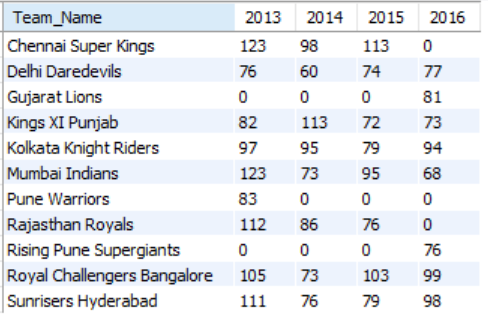
**sum(case when Year=2015 then Total\_Wickets\_Taken else 0 end) as "2015",**

**sum(case when Year=2016 then Total\_Wickets\_Taken else 0 end) as "2016"**

**from cte2**

**group by Team\_Name**

**order by Team\_Name;**

**Result :** 

1. Derive more KPIs for the team strategy?

Ans . Additional Key Performance Indicators (KPIs) can be derived to strengthen team strategy and long-term planning based on season-wise team performance data:

**1. Win Percentage:**

Formula:

(Number of Matches Won/Number of Matches Played​) ×100

Tables & Columns Used:

Matches table:

* Match\_Winner – to identify which team won
* Team\_1, Team\_2 – to count total matches a team played

*This helps evaluate a team's overall success rate.*

**2. Average Runs per Match:**

Formula:

(Total Runs Scored​ / Total Matches Played​)

Tables & Columns Used:

Ball\_by\_Ball table:

* *Runs\_Scored* – the runs scored per ball
* *Team\_Batting* – which team scored the run

Matches table:

* *Match\_Id* – to count unique matches per team

*Useful for understanding the team's scoring consistency.*

**3. Average Wickets per Match:**

Formula:

(Total Wickets Taken​​ / Total Matches Played​)

Tables & Columns Used:

Wicket\_Taken table:

* *Player\_Out* – identifies a wicket taken

Ball\_by\_Ball table:

* *Team\_Bowling* – to know which team took the wicket

Matches table:

* *Match\_Id* – to group wickets match-wise

*This shows the bowling strength of a team*

**4. Run Rate:**

Formula:

(Total Runs Scored​​​ / Total Overs Faced​)

Tables & Columns Used:

Ball\_by\_Ball table:

* + 1. Runs\_Scored – for total runs
    2. Over\_Id – to calculate total overs
    3. Team\_Batting – to group by team

*Helps understand how quickly a team scores runs.*

**5. Toss Impact KPI:**

Formula:

Win % After Toss Decision (Bat/Field):

(Matches Won After Toss Decision / Matches with That Toss Decision) \* 100

1. Year-on-Year Performance Growth Rate  
    Evaluates the percentage change in team performance across seasons, helping assess improvement or decline trends.
2. Average Seasonal Performance Score  
    Calculates the mean performance value across all seasons for each team, indicating overall consistency.

To analyse year wise teams performance below **Query used:**

**with cte as**

**(select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Team\_Batting,**

**(b.Runs\_Scored + IFNULL(e.Extra\_Runs, 0)) AS Total\_Runs**

**from Ball\_by\_Ball b left join Extra\_Runs e**

**on b.Match\_Id=e.Match\_Id and**

**b.Over\_Id=e.Over\_Id and b.Ball\_Id=e.Ball\_Id and**

**b.Innings\_No=e.Innings\_No**

**),**

**cte1 as (select c.Match\_Id,year(m.Match\_Date) as Year,c.Over\_Id,c.Ball\_Id,c.Innings\_No,c.Team\_Batting,c.Total\_Runs,t.Team\_Name**

**from cte c join Matches m on c.Match\_Id=m.Match\_Id**

**join Team t on t.Team\_Id=c.Team\_Batting)**

**select team\_name,**

**sum(case when year = 2013 then total\_runs else 0 end) as "2013",**

**sum(case when year = 2014 then total\_runs else 0 end) as "2014",**

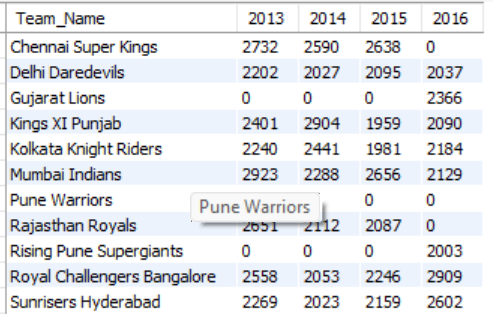
**sum(case when year = 2015 then total\_runs else 0 end) as "2015",**

**sum(case when year = 2016 then total\_runs else 0 end) as "2016"**

**from cte1**

**group by team\_name order by team\_name;**

**Result :**



13. Using SQL, write a query to find out average wickets taken by each bowler in each venue. Also rank the gender according to the average value.

Ans. Please find query to find out average wickets taken by each bowler in each venue

**Query used:**

**with cte as**

**(select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler,b.Team\_Bowling**

**from Ball\_by\_Ball b join Wicket\_Taken w**

**on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id and**

**b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No),**

**cte1 as**

**(select c.Match\_Id,year(m.Match\_Date) as Year,m.Venue\_Id,c.Bowler,c.Team\_Bowling,**

**t.Team\_Name**

**from cte c join Matches m on c.Match\_Id=m.Match\_Id**

**join Team t on c.Team\_Bowling=t.Team\_Id**

**),**

**cte2 as**

**(**

**select c1.Match\_Id,c1.Year,c1.Venue\_Id,v.Venue\_Name,c1.Bowler,p.Player\_Name as Bowler\_Name,c1.Team\_Bowling,c1.Team\_Name**

**from cte1 c1 join Player p on c1.Bowler=p.Player\_Id**

**join Venue v on v.Venue\_Id=c1.Venue\_Id**

**),**

**cte3 as**

**(SELECT cte2.Bowler\_Name, cte2.Venue\_Name,**

**COUNT(\*) AS Total\_Wickets\_Taken,**

**COUNT(DISTINCT cte2.Match\_Id) AS Matches\_Played,**

**CAST(COUNT(\*) AS FLOAT) / COUNT(DISTINCT cte2.Match\_Id) AS Avg\_Wickets\_Per\_Match**

**FROM cte2**

**GROUP BY cte2.Bowler\_Name, cte2.Venue\_Name**

**ORDER BY Avg\_Wickets\_Per\_Match DESC**

**)**

**select Bowler\_Name,Venue\_Name,Avg\_Wickets\_Per\_Match,dense\_rank() over(order by Avg\_Wickets\_Per\_Match desc) as "Rank"**

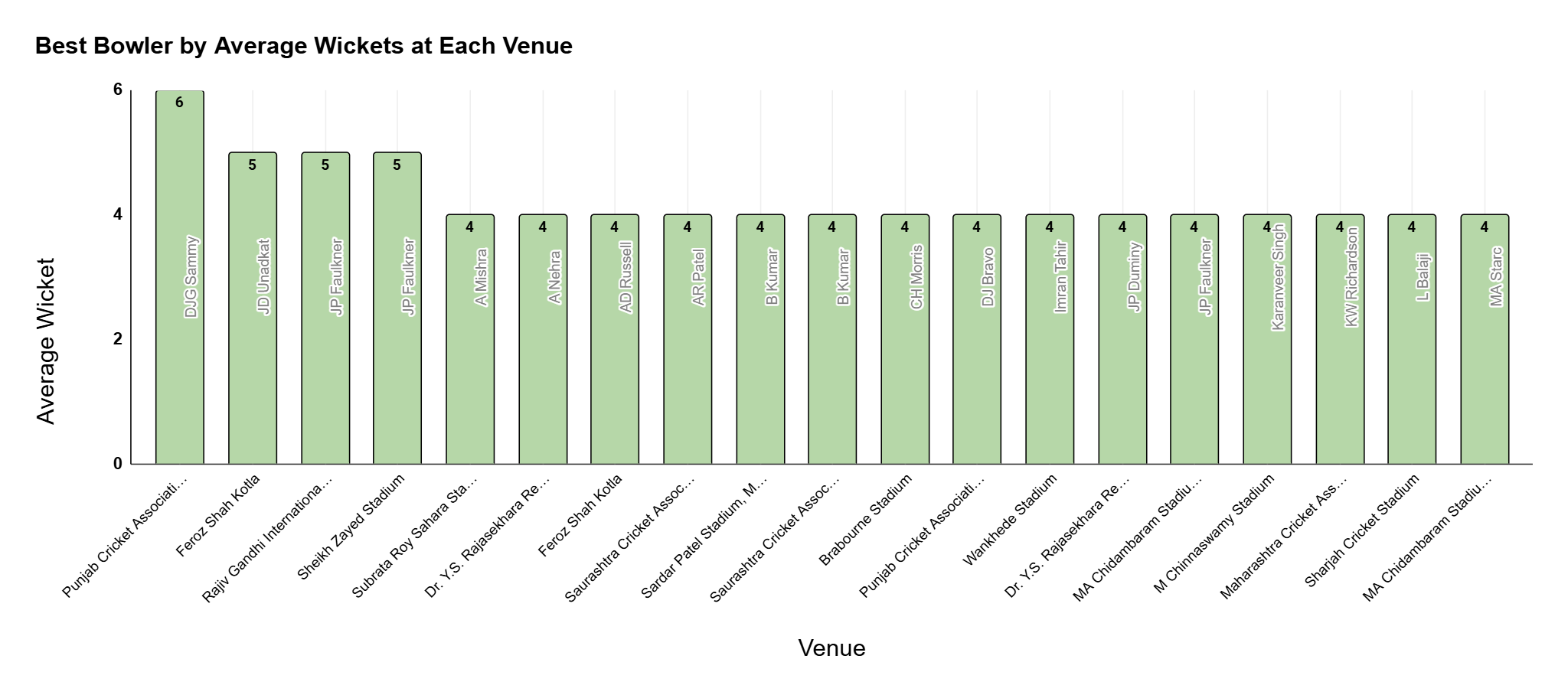
**from cte3**

**order by Avg\_Wickets\_Per\_Match desc;**

**Result :**

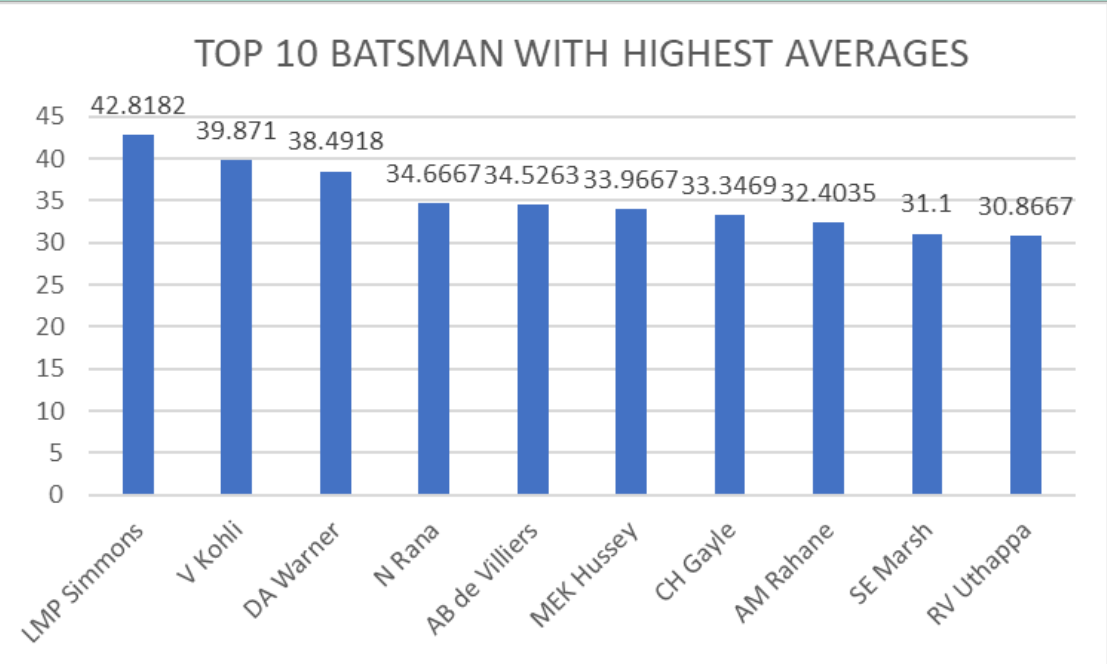


**Also, please find the visual representation for the data below.**



14. Which of the given players have consistently performed well in past seasons? (will you use any visualisation to solve the problem)

Ans. Below are the list for the players who have consistently performed well in past seasons.



**Query used to found top batsman :**

**with cte as**

**(select b.Striker,p.Player\_Name,b.Runs\_Scored,m.Match\_Id,m.Venue\_Id**

**from Ball\_by\_Ball b join Matches m on b.Match\_Id=m.Match\_Id**

**join Player p on p.Player\_Id=b.Striker)**

**select Striker as Player\_Id,Player\_Name,count(distinct Match\_Id) as Total\_Matches\_Played,sum(Runs\_Scored) as Total\_Runs\_Scored,**

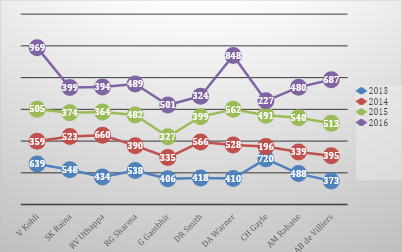
**sum(Runs\_Scored)/count(distinct Match\_Id) as Average**

**from cte**

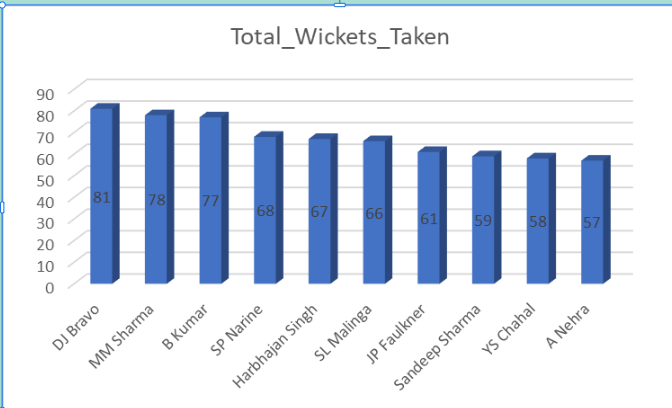
**group by Striker,Player\_Name**

**order by Average desc**

**limit 10;**



**List of top 10 Bowlers**.



**Query used to found top Bowler:**

**with cte as(**

**select b.Match\_Id, m.Venue\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler as Player\_Id,**

**p.Player\_Name**

**from Ball\_by\_Ball b join Wicket\_Taken w**

**on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id and**

**b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No**

**join Matches m on m.Match\_Id=b.Match\_Id**

**join Player p on p.Player\_Id=b.Bowler**

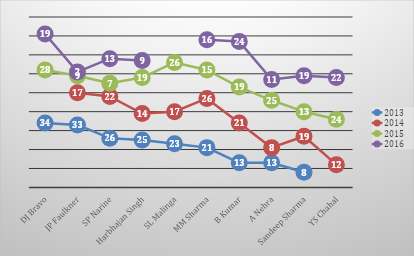
**)**

**select Player\_Id,Player\_Name,count(distinct Match\_Id) as Total\_Match\_Played,**

**count(\*) as Total\_Wickets\_Taken from cte**

**group by Player\_Id,Player\_Name, order by Total\_Wickets\_Taken desc**

**limit 10;**



15. Are there players whose performance is more suited to specific venues or conditions? (how would you present this using charts?)

Ans. Yes, there are players whose performance is more suited to specific venues. Some players consistently score more runs or take more wickets at particular grounds due to pitch type, stadium conditions, or home advantage.

So Please find query for **Venue wise run scored by top batsman:**

**with cte as**

**(select b.Striker,p.Player\_Name,b.Runs\_Scored,m.Match\_Id,m.Venue\_Id**

**from Ball\_by\_Ball b join Matches m on b.Match\_Id=m.Match\_Id**

**join Player p on p.Player\_Id=b.Striker),**

**cte1 as (**

**select c.Match\_Id,c.Venue\_Id,v.Venue\_Name,c.Striker as Player\_Id,c.Player\_Name,**

**c.Runs\_Scored from cte c join Venue v on**

**c.Venue\_Id=v.Venue\_Id)**

**select Player\_Id,Player\_Name,Venue\_Id,Venue\_Name,**

**count(distinct Match\_Id) as Total\_Matches\_Played,**

**sum(Runs\_Scored) as Total\_Runs\_Scored,**

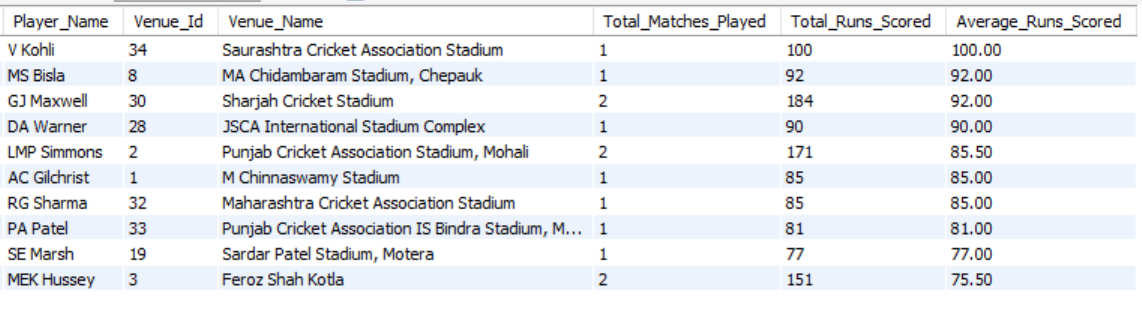
**round(sum(Runs\_Scored)/count(distinct Match\_Id),2) as Average\_Runs\_Scored**

**from cte1**

**group by Player\_Id,Player\_Name,Venue\_Id,Venue\_Name**

**order by Average\_Runs\_Scored desc**

**limit 10**

**Result: **

Venue wise wickets by top bowlers:

with cte as

(

select b.Match\_Id,m.Venue\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler as Player\_Id,

p.Player\_Name

from Ball\_by\_Ball b join Wicket\_Taken w

on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id and

b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No

join Matches m on m.Match\_Id=b.Match\_Id

join Player p on p.Player\_Id=b.Bowler

),

cte1 as

(select c.Match\_Id,c.Venue\_Id,v.Venue\_Name,c.Over\_Id,c.Ball\_Id,

c.Innings\_No,c.Player\_Id,c.Player\_Name

from cte c join Venue v

on c.Venue\_Id=v.Venue\_Id)

select Player\_Id,Player\_Name,Venue\_Id,Venue\_Name,count(distinct Match\_Id) as Total\_Matches\_Played,

count(\*) as Total\_Wickets\_Taken

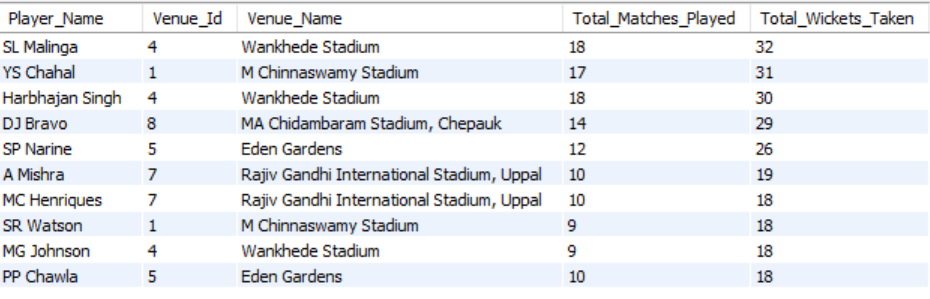
from cte1

group by Player\_Id,Player\_Name,Venue\_Id,Venue\_Name

order by Total\_Wickets\_Taken desc

limit 10

**Result:**

****

**Subjective Questions**

1. How does the toss decision affect the result of the match? (which visualizations could be used to present your answer better) And is the impact limited to only specific venues?

Answer: The toss decision does have a noticeable impact on match outcomes, but its influence is not uniform across all matches or venues. Winning the toss allows teams to make a strategic choice—batting or fielding first—based on pitch conditions, weather, dew factor, and historical venue trends. When used effectively, this decision can provide a competitive advantage.



### Venue-Specific Impact

The impact of the toss is not universal and is clearly venue-dependent. Certain stadiums consistently favor one decision over the other, while some venues show a relatively balanced outcome regardless of the toss decision. In conclusion, the toss decision does influence match results, but its effectiveness depends heavily on the venue. Teams that understand venue-specific trends and align their toss decisions

accordingly are more likely to gain a strategic advantage rather than relying on the toss alone.

So to understand this we have analysed venue wise winning data related to toss decision using below query.

**Query used:**

with cte as (

SELECT m.Match\_Id,t.Toss\_Name AS Toss\_Decision,m.Match\_Winner,m.Toss\_Winner,v.Venue\_Name,

CASE

WHEN m.Match\_Winner = m.Toss\_Winner THEN 'Toss Winner Won'

ELSE 'Toss Winner Lost'

END AS Toss\_Impact

FROM Matches m

JOIN Venue v ON m.Venue\_Id = v.Venue\_Id

JOIN Toss\_Decision t ON t.Toss\_Id = m.Toss\_Decide

),

cte1 as (

select Venue\_Name,Toss\_Decision,Toss\_Impact,count(Match\_Id) as Match\_Count

from cte

group by Venue\_Name,Toss\_Decision,Toss\_Impact

)

select Venue\_Name,

sum(case when (Toss\_Decision='field' and Toss\_Impact='Toss Winner Won') or (Toss\_Decision='bat' and Toss\_Impact='Toss Winner Lost')

then Match\_Count end) as Field\_First\_Wins,

sum(case when (Toss\_Decision='bat' and Toss\_Impact='Toss Winner Won') or (Toss\_Decision='field' and Toss\_Impact='Toss Winner Lost')

then Match\_Count end) as Bat\_First\_Wins

from cte1

group by Venue\_Name



1. Suggest some of the players who would be best fit for the team.

Answer: To suggest the best players for the team, I analysed the performance data of all players across multiple IPL seasons. I looked at both batsmen and bowlers, and selected players who showed consistency, match-winning ability, and good performance across different venues.

How I selected the best players:

For Batsmen, I checked:

* Total Runs
* Batting Average
* Strike Rate
* Performance in different venues

### Top 10 Batsmen (by average runs per match)

| Player Name | Avg Runs/Match |
| --- | --- |
| LMP Simmons | 42.82 |
| V Kohli | 39.87 |
| DA Warner | 38.49 |
| AB de Villiers | 34.53 |
| MEK Hussey | 33.97 |
| CH Gayle | 33.35 |
| AM Rahane | 32.40 |
| SE Marsh | 31.10 |
| RV Uthappa | 30.87 |
| DR Smith | 29.95 |

**Query used:**

with cte as (

select p.player\_id,p.player\_name,b.runs\_scored,b.match\_id

from player p

left join ball\_by\_ball b

on p.player\_id = b.striker

)

select

player\_id,

player\_name,

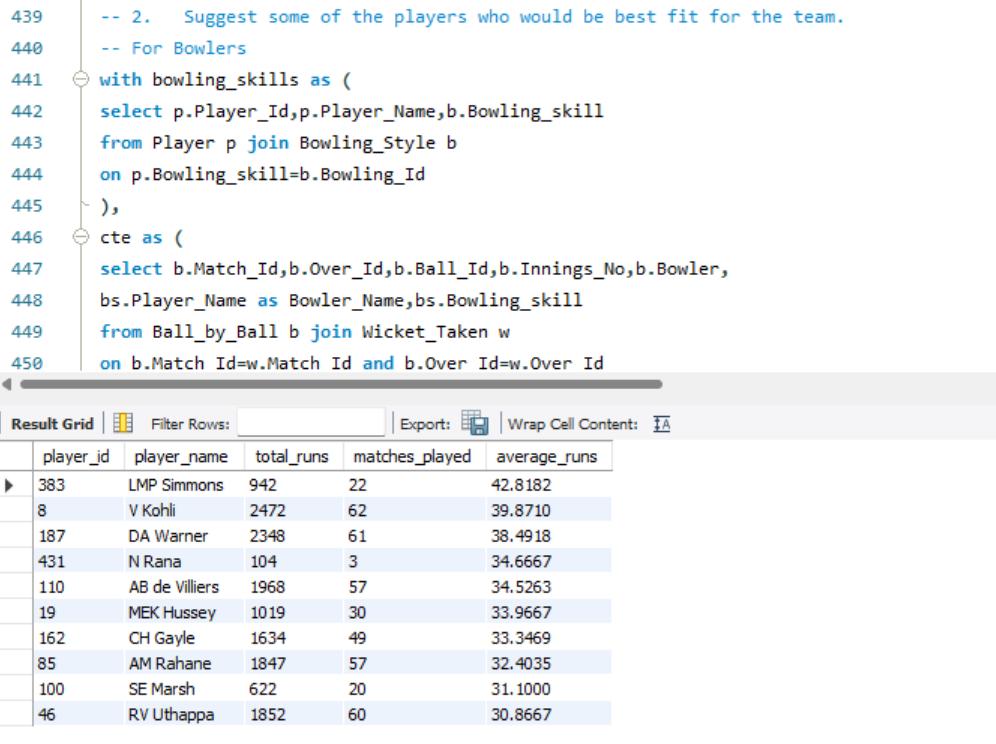
coalesce(sum(runs\_scored), 0) as total\_runs,

count(distinct match\_id) as matches\_played,

coalesce(sum(runs\_scored) / nullif(count(distinct match\_id), 0), 0) as average\_runs from cte

group by player\_id, player\_name

order by average\_runs desc limit 10;



**For Bowlers, I checked:**

* Total Wickets
* Bowling Average
* Economy Rate
* Consistency over seasons

**Top 10 Bowlers (by total wickets)**

| Player Name | Wickets |
| --- | --- |
| DJ Bravo | 81 |
| MM Sharma | 78 |
| B Kumar | 77 |
| SP Narine | 68 |
| Harbhajan Singh | 67 |
| SL Malinga | 66 |
| JP Faulkner | 61 |
| Sandeep Sharma | 59 |
| YS Chahal | 58 |
| A Nehra | 57 |

**Query used:**

with bowling\_skills as (

select p.Player\_Id,p.Player\_Name,b.Bowling\_skill

from Player p join Bowling\_Style b

on p.Bowling\_skill=b.Bowling\_Id

),

cte as (

select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler,

bs.Player\_Name as Bowler\_Name,bs.Bowling\_skill

from Ball\_by\_Ball b join Wicket\_Taken w

on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id

and b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No

join bowling\_skills bs on bs.Player\_Id=b.Bowler

)

SELECT Bowler\_Name, Bowling\_skill,

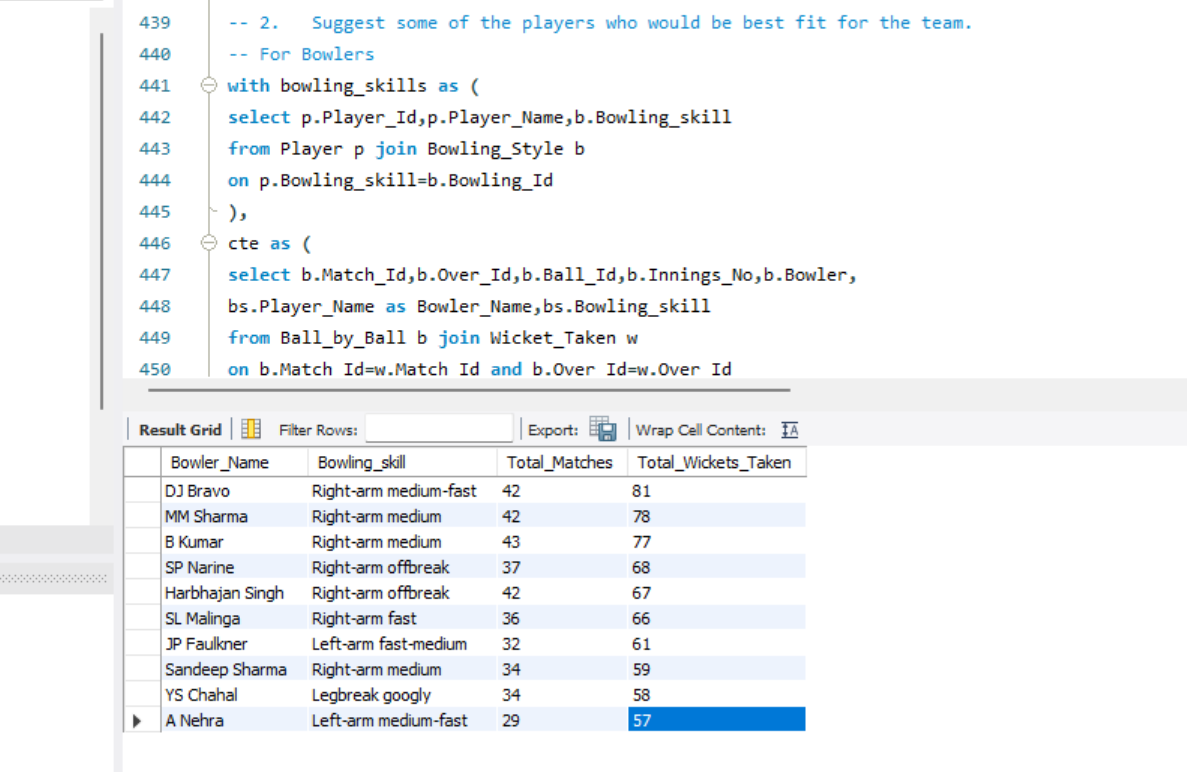
count(distinct Match\_Id) as Total\_Matches,

COUNT(\*) AS Total\_Wickets\_Taken

FROM cte

GROUP BY Bowler\_Name, Bowling\_skill

ORDER BY Total\_Wickets\_Taken DESC limit 10;



1. What are some of the parameters that should be focused on while selecting the players?

Answer: While selecting players for a team, especially in a cricket analytics or team-strategy context, multiple parameters should be considered to ensure balance, consistency, and adaptability. These parameters can be grouped into performance, conditions, and team requirements.

### **Player Performance Metrics**

* Batting performance: Average, strike rate, boundary percentage, consistency across recent matches.
* Bowling performance: Economy rate, wickets per match, strike rate, effectiveness in different phases (powerplay, middle overs, death overs).
* All-round ability: Contribution with both bat and ball.
* Fielding skills: Catches taken, run-outs, ground fielding efficiency.
* Performance in the last few matches or tournaments
* Match Conditions & Venue
* Balance between batters, bowlers, all-rounders, and Right-left batting combinations.
* Fitness & Availability and Performance in high-pressure or knockout matches.

### Conclusion

* Player selection should be data-driven but not data-dependent alone. A balanced approach that combines statistical analysis, match conditions, recent form, and team strategy leads to more effective and consistent team performance

1. Which players offer versatility in their skills and can contribute effectively with both bat and ball? (can you visualize the data for the same)

Answer: Some players are not just good at one skill — they are valuable all-rounders, meaning they score runs and take wickets regularly. These players are very useful in T20 matches, as they add balance to the team.

To find such players, I used a single SQL query that calculates:

* Total runs scored (as batsman)
* Total wickets taken (as bowler)
* And filters players who have both decent runs and wicket counts

**Query used:**

WITH Bowling\_Skills AS (

SELECT p.Player\_Id, p.Player\_Name, b.Bowling\_skill

FROM Player p

JOIN Bowling\_Style b

ON p.Bowling\_skill = b.Bowling\_Id

),

Bowling\_Data AS (SELECT b.Bowler AS Player\_Id, Player\_Name, bs.Bowling\_skill,

COUNT(DISTINCT b.Match\_Id) AS Total\_Matches\_Bowled,

COUNT(\*) AS Total\_Wickets\_Taken

FROM Ball\_by\_Ball b

JOIN Wicket\_Taken w

ON b.Match\_Id = w.Match\_Id

AND b.Over\_Id = w.Over\_Id

AND b.Ball\_Id = w.Ball\_Id

AND b.Innings\_No = w.Innings\_No

JOIN Bowling\_Skills bs

ON bs.Player\_Id = b.Bowler

GROUP BY b.Bowler, bs.Player\_Name, bs.Bowling\_skill

),

Batting\_Data AS (SELECT p.player\_id,p.player\_name,

COALESCE(SUM(b.runs\_scored), 0) AS Total\_Runs,

COUNT(DISTINCT b.match\_id) AS Matches\_Played\_Batted

FROM Player p

LEFT JOIN Ball\_by\_Ball b

ON p.player\_id = b.striker

GROUP BY p.player\_id, p.player\_name)

SELECT b.Player\_Id, b.Player\_Name, b.Total\_Wickets\_Taken, b.Total\_Matches\_Bowled, bt.Total\_Runs, bt.Matches\_Played\_Batted

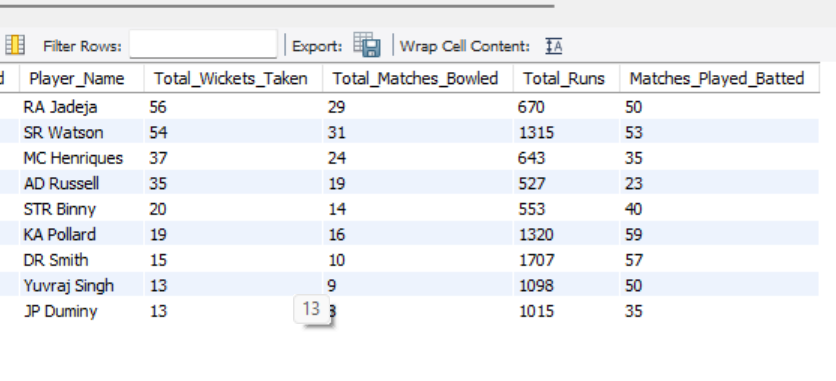
FROM Bowling\_Data b

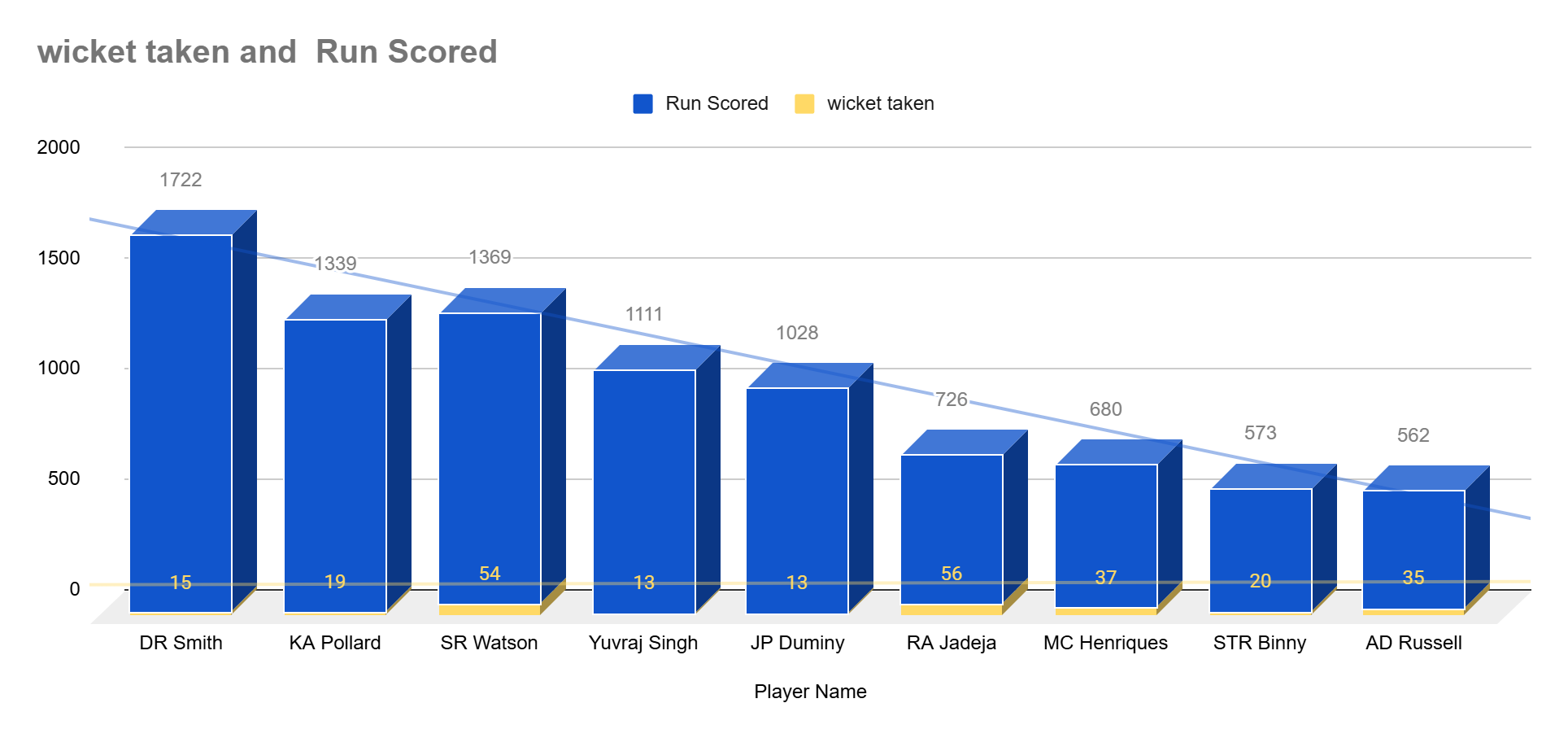
JOIN Batting\_Data bt ON b.Player\_Id = bt.player\_id

WHERE b.Total\_Wickets\_Taken > 10 AND bt.Total\_Runs > 500

ORDER BY b.Total\_Wickets\_Taken DESC, bt.Total\_Runs DESC;

**SQL Output -**





1. Are there players whose presence positively influences the morale and performance of the team? (justify your answer using visualization)

Answer: Yes, there are certain players whose presence seems to have a positive impact on the team’s overall performance. These players may not always be the highest scorers or wicket-takers in every match, but the team wins more frequently when they are part of the playing XI.  
They are often seen as leaders, experienced mentors, or motivational figures — contributing to team morale and strategy.

How I Identified Them:

To analyse this, I checked:

* How many matches a player played
* How many of those matches the team won
* The win percentage when the player was in the match

**Query used:**

SELECT

Player\_Name,

COUNT(\*) AS Matches\_Played,

SUM(CASE WHEN Team\_Id = Match\_Winner THEN 1 ELSE 0 END) AS Matches\_Won,

ROUND(SUM(CASE WHEN Team\_Id = Match\_Winner THEN 1 ELSE 0 END) \* 100.0 / COUNT(\*), 2) AS Win\_Percentage

FROM (

SELECT DISTINCT

m.Match\_Id,

p.Player\_Name,

pb.Team\_Batting AS Team\_Id,

m.Match\_Winner

FROM Ball\_by\_Ball pb

JOIN Matches m ON pb.Match\_Id = m.Match\_Id

JOIN Player p ON pb.Striker = p.Player\_Id

) AS player\_matches

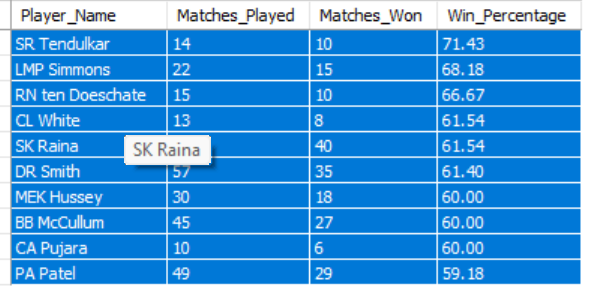
GROUP BY Player\_Name

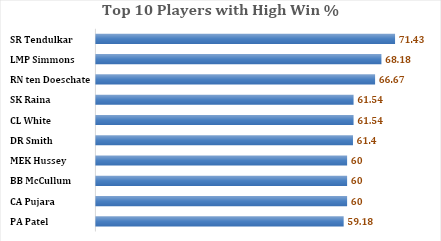
HAVING Matches\_Played >= 10

ORDER BY Win\_Percentage DESC

limit 10;

**SQL Output -**





1. What would you suggest to RCB before going to the mega auction?

Answer: Before going to the mega auction, RCB should focus on selecting players who consistently perform well and help the team win. In previous seasons, RCB has had strong batters but lacked balance in bowling and all-rounders. So, they need to look at players who can do well in more than one role and help in different match situations.

Objective:

1. Pick players who perform well in both batting and bowling.  
   ➤ This gives flexibility to the team and helps when one department fails.
2. Select consistent performers.  
   ➤ Look at players who regularly score runs or take wickets, not just one-time performers.
3. Choose players with good match-winning impact.  
   ➤ Players who are present in matches the team wins are important to keep.

**Retain existing RCB players with high match impact:**

**Query used:**

SELECT

p.Player\_Name,

COUNT(DISTINCT m.Match\_Id) AS Matches\_Played,

SUM(CASE WHEN m.Match\_Winner = b.Team\_Batting THEN 1 ELSE 0 END) AS Matches\_Won,

ROUND(SUM(CASE WHEN m.Match\_Winner = b.Team\_Batting THEN 1 ELSE 0 END) \* 100.0 / COUNT(DISTINCT m.Match\_Id), 2) AS Win\_Percentage

FROM Ball\_by\_Ball b

JOIN Matches m ON b.Match\_Id = m.Match\_Id

JOIN Player p ON b.Striker = p.Player\_Id

WHERE b.Team\_Batting = 2

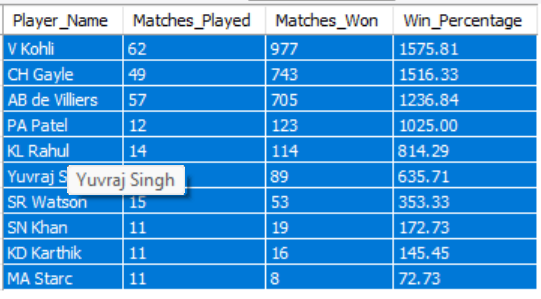
GROUP BY p.Player\_Name

HAVING Matches\_Played >= 10

ORDER BY Win\_Percentage DESC

LIMIT 10;

**SQL Output -**



**Pick players who are good in both batting and bowling**

**Query used:**

SELECT

p.Player\_Name,

COALESCE(batting.Total\_Runs, 0) AS Total\_Runs,

COALESCE(bowling.Total\_Wickets, 0) AS Total\_Wickets

FROM Player p

LEFT JOIN (

SELECT Striker AS Player\_Id, SUM(Runs\_Scored) AS Total\_Runs

FROM Ball\_by\_Ball

GROUP BY Striker

) AS batting ON p.Player\_Id = batting.Player\_Id

LEFT JOIN (

SELECT b.Bowler AS Player\_Id, COUNT(\*) AS Total\_Wickets

FROM Ball\_by\_Ball b

JOIN Wicket\_Taken w ON b.Match\_Id = w.Match\_Id

AND b.Over\_Id = w.Over\_Id

AND b.Ball\_Id = w.Ball\_Id

AND b.Innings\_No = w.Innings\_No

GROUP BY b.Bowler

) AS bowling ON p.Player\_Id = bowling.Player\_Id

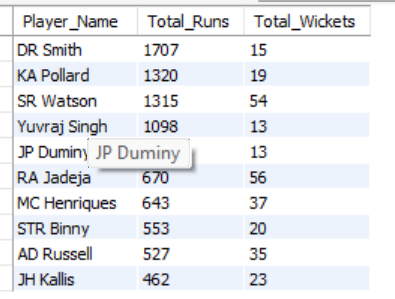
WHERE COALESCE(batting.Total\_Runs, 0) >= 150

AND COALESCE(bowling.Total\_Wickets, 0) >= 10

ORDER BY Total\_Runs DESC, Total\_Wickets DESC

LIMIT 10;

**SQL Output -**



1. What do you think could be the factors contributing to the high-scoring matches and the impact on viewership and team strategies.

Answer: To understand the factors contributing to high-scoring matches, I analysed the IPL dataset using a SQL query that calculates the average runs scored per ball at different venues. This helped me identify which stadiums are more batting-friendly and likely for high scores.

**Query used:**

SELECT

v.Venue\_Name,

ROUND(SUM(b.Runs\_Scored) \* 1.0 / COUNT(b.Ball\_Id), 2) AS Avg\_Runs\_Per\_Ball

FROM Ball\_by\_Ball b

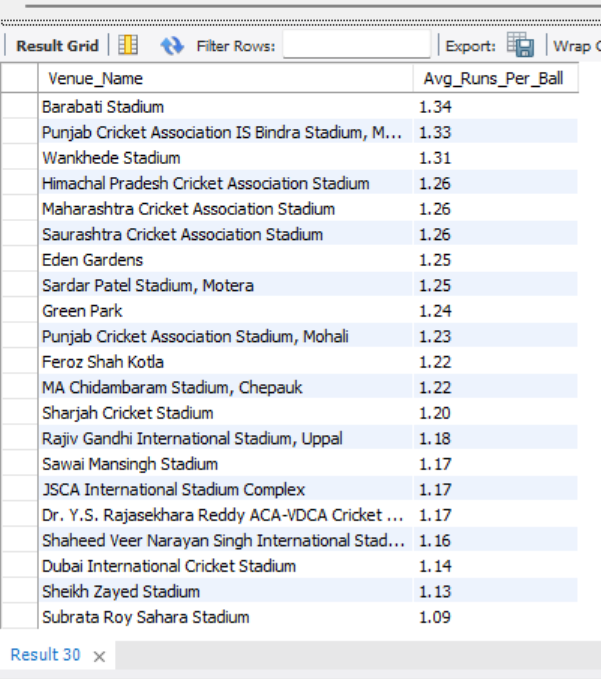
JOIN Matches m ON b.Match\_Id = m.Match\_Id

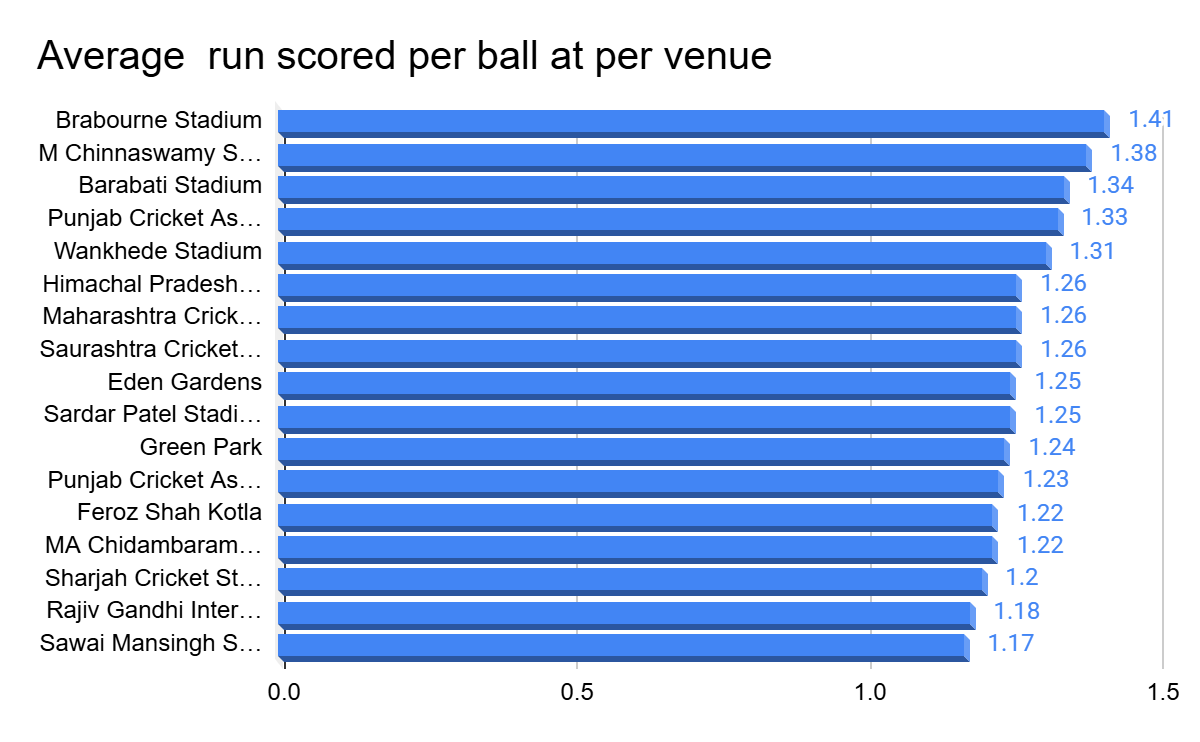
JOIN Venue v ON m.Venue\_Id = v.Venue\_Id

GROUP BY v.Venue\_Name

ORDER BY Avg\_Runs\_Per\_Ball DESC

**SQL Output -**





Insights from the Query:

1. Certain venues consistently have higher scoring rates.
   * These are usually grounds with short boundaries and flat pitches.
   * Bowlers struggle more, while batters score faster.
2. Teams playing at these venues tend to score 180+ regularly.
   * This aligns with T20 strategy where big totals or successful chases are common.

Impact on Viewership:

* High-scoring matches are more thrilling, which boosts live viewership and online engagement.
* Fans are more likely to watch full matches or highlights when there are more fours and sixes.
* Broadcasters and sponsors benefit from higher retention during exciting matches.

Impact on Team Strategies:

* Player Selection: Teams pick aggressive batsmen and death-over finishers for high-scoring venues.
* Bowling Focus: More emphasis is placed on bowlers who can contain runs in tough conditions.
* Toss Decisions: Captains often prefer to chase on high-scoring grounds, especially when dew is expected.

1. Analyze the impact of home-ground advantage on team performance and identify strategies to maximize this advantage for RCB.

Answer: To analyze the impact of home-ground advantage, I used the IPL dataset for team RCB, I checked how often they won when playing at their home ground: M. Chinnaswamy Stadium, Bangalore.

**Query used:**

with cte as (

select m.Match\_Id,m.Team\_1,t1.Team\_Name as mTeam\_1\_Name,m.Team\_2,t2.Team\_Name as Team\_2\_Name,

m.Match\_Winner,t3.Team\_Name as Match\_Winner\_Name,m.Venue\_Id,v.Venue\_Name,m.Season\_Id,

s.Season\_Year

from Matches m join Team t1 on m.Team\_1=t1.Team\_Id

join Team t2 on m.Team\_2=t2.Team\_Id

join Team t3 on m.Match\_Winner=t3.Team\_Id

join Venue v on v.Venue\_Id=m.Venue\_Id

join Season s on s.Season\_Id=m.Season\_Id

),

RCB\_Chinnaswamy AS (SELECT Match\_Id,Venue\_Name,Team\_1\_Name,Team\_2\_Name,Match\_Winner\_Name

FROM cte

WHERE Venue\_Name = 'M Chinnaswamy Stadium'

AND (Team\_1\_Name = 'Royal Challengers Bangalore' OR Team\_2\_Name = 'Royal Challengers Bangalore')

)

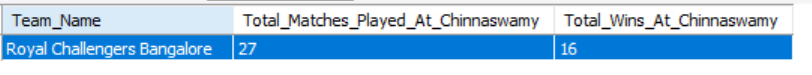
SELECT 'Royal Challengers Bangalore' AS Team\_Name,

COUNT(\*) AS Total\_Matches\_Played\_At\_Chinnaswamy,

SUM(CASE WHEN Match\_Winner\_Name = 'Royal Challengers Bangalore' THEN 1 ELSE 0 END) AS Total\_Wins\_At\_Chinnaswamy

FROM RCB\_Chinnaswamy;

**SQL Output -**



From the analysis:

* RCB has played 27 matches at Chinnaswamy Stadium.
* Out of these, they have won 16 matches, resulting in a win percentage of approximately 59%.

This indicates that RCB performs better at home compared to away venues

### Reasons for Home-Ground Advantage at Chinnaswamy Stadium

* Batting-friendly pitch with consistent bounce.
* Short boundaries, favoring aggressive stroke play.
* Crowd support, boosting player confidence and momentum.

### Strategies to Maximize Home-Ground Advantage for RCB

. Strengthen Power-Hitting Lineup ,Prioritize aggressive top-order batters and finishers who can exploit short boundaries. And promote players with high strike rates at Chinnaswamy.

. Use bowlers with effective variations (slower balls, cutters, yorkers). Death-over specialists are crucial to control high-scoring games.

. Prefer chasing, as dew often makes bowling difficult in the second innings.Chasing aligns well with RCB’s strong batting depth.

. Focus on athletic fielders to save runs on a small ground where every run matters.

. Maintain aggressive intent early to energize the home crowd and build pressure on opponents.

1. Come up with a visual and analytical analysis of the RCB's past season's performance and potential reasons for them not winning a trophy.

Answer: To understand why Royal Challengers Bangalore (RCB) has not won an IPL trophy so far, I analyzed their performance across past seasons. My focus was on:

* **Win percentage in each season**

**Query used:**

with cte as(

select m.Match\_Id,m.Team\_1,t1.Team\_Name as Team\_1\_Name,m.Team\_2,t2.Team\_Name as Team\_2\_Name,

m.Match\_Winner,t3.Team\_Name as Match\_Winner\_Name,m.Venue\_Id,v.Venue\_Name,m.Season\_Id,

s.Season\_Year

from Matches m join Team t1 on m.Team\_1=t1.Team\_Id

join Team t2 on m.Team\_2=t2.Team\_Id

join Team t3 on m.Match\_Winner=t3.Team\_Id

join Venue v on v.Venue\_Id=m.Venue\_Id

join Season s on s.Season\_Id=m.Season\_Id

)

select Season\_Id,Season\_Year, count(\*) as Total\_Matches\_Played\_By\_RCB,

sum(case when Match\_Winner\_Name='Royal Challengers Bangalore' then 1 end) as Total\_Matches\_Won\_By\_RCB

from cte

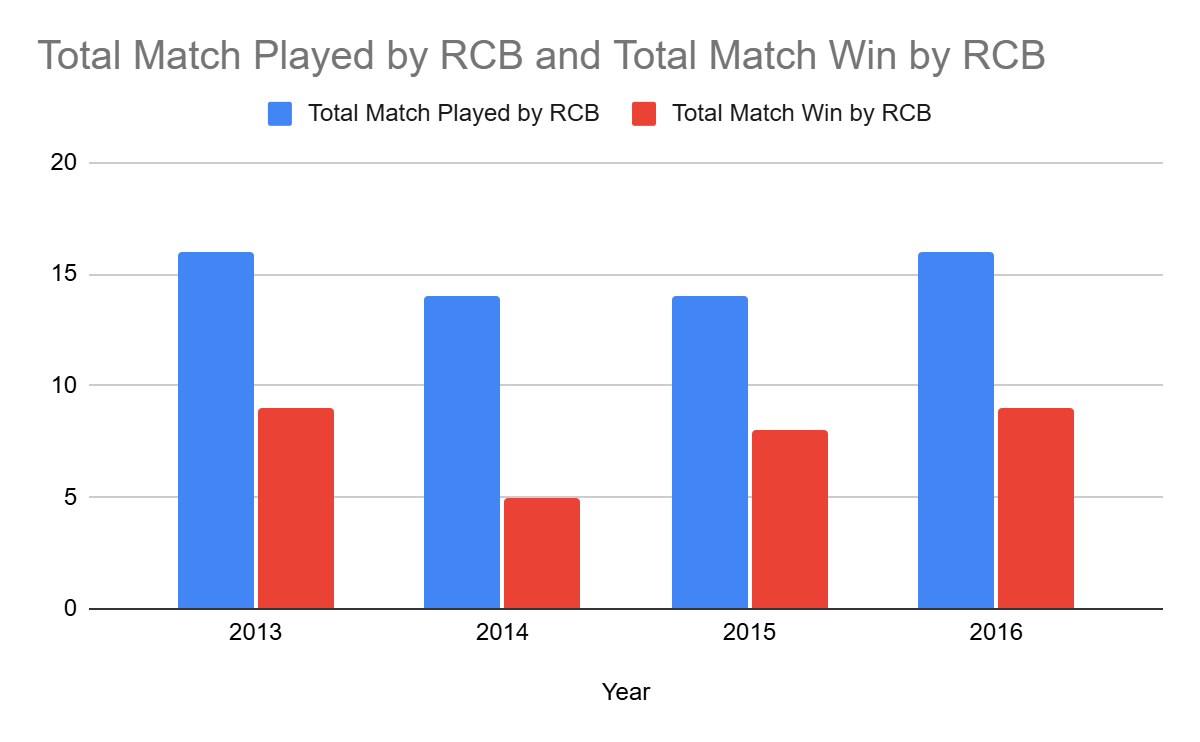
where Team\_1\_Name='Royal Challengers Bangalore' or Team\_2\_Name='Royal Challengers Bangalore'

group by Season\_Id,Season\_Year

order by Season\_Year

**SQL Output -**





* **Batting consistency (Number of Runs Scored Yearwise )**

**Query used:**

with cte as

(

select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Team\_Batting,

(b.Runs\_Scored + IFNULL(e.Extra\_Runs, 0)) AS Total\_Runs

from Ball\_by\_Ball b left join Extra\_Runs e

on b.Match\_Id=e.Match\_Id and

b.Over\_Id=e.Over\_Id and b.Ball\_Id=e.Ball\_Id and

b.Innings\_No=e.Innings\_No

),

cte1 as (

select c.Match\_Id,year(m.Match\_Date) as Year,c.Over\_Id,c.Ball\_Id,c.Innings\_No,c.Team\_Batting,c.Total\_Runs,t.Team\_Name

from cte c join Matches m on c.Match\_Id=m.Match\_Id

join Team t on t.Team\_Id=c.Team\_Batting),

cte2 as (

select team\_name,

sum(case when year = 2013 then total\_runs else 0 end) as "2013",

sum(case when year = 2014 then total\_runs else 0 end) as "2014",

sum(case when year = 2015 then total\_runs else 0 end) as "2015",

sum(case when year = 2016 then total\_runs else 0 end) as "2016"

from cte1

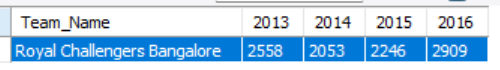
group by team\_name

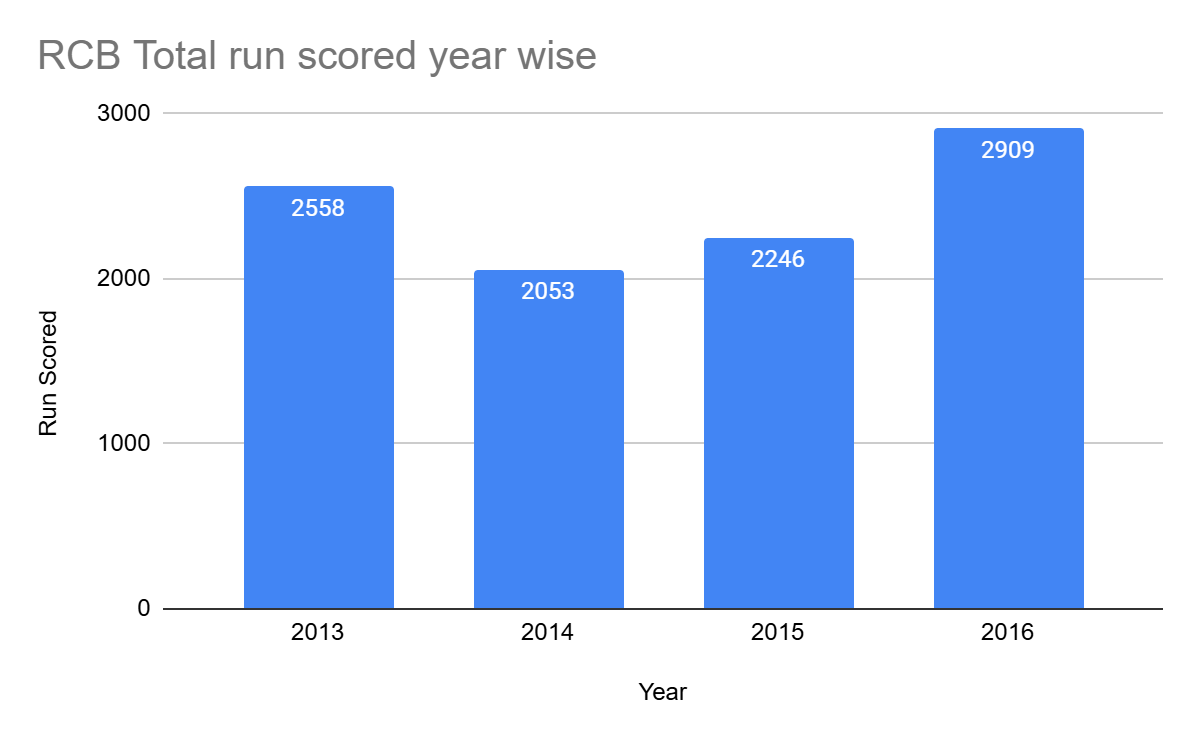
order by team\_name

)

select \* from cte2 where team\_name="Royal Challengers Bangalore"

**SQL Output -**





* **Number of Wickets Taken Yearwise by RCB**

**Query used:**

with cte as

(select b.Match\_Id,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Bowler,b.Team\_Bowling

from Ball\_by\_Ball b join Wicket\_Taken w

on b.Match\_Id=w.Match\_Id and b.Over\_Id=w.Over\_Id and

b.Ball\_Id=w.Ball\_Id and b.Innings\_No=w.Innings\_No),

cte1 as

(select c.Match\_Id,year(m.Match\_Date) as Year,c.Team\_Bowling,

t.Team\_Name

from cte c join Matches m on c.Match\_Id=m.Match\_Id

join Team t on c.Team\_Bowling=t.Team\_Id

),

cte2 as

(select Team\_Name,Year,count(\*) as Total\_Wickets\_Taken

from cte1

group by Team\_Name,Year),

cte3 as (

select Team\_Name,

sum(case when Year=2013 then Total\_Wickets\_Taken else 0 end) as "2013",

sum(case when Year=2014 then Total\_Wickets\_Taken else 0 end) as "2014",

sum(case when Year=2015 then Total\_Wickets\_Taken else 0 end) as "2015",

sum(case when Year=2016 then Total\_Wickets\_Taken else 0 end) as "2016"

from cte2

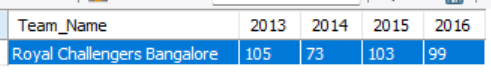
group by Team\_Name

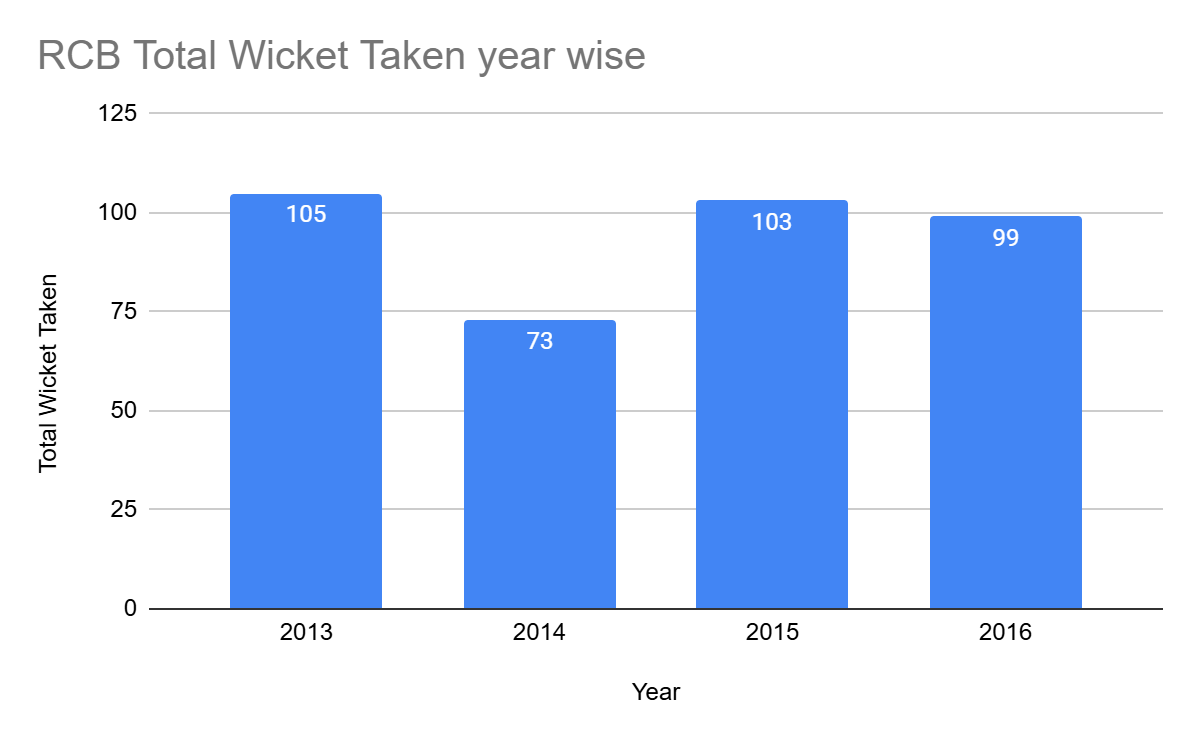
order by Team\_Name

)

select \* from cte3 where Team\_Name="Royal Challengers Bangalore"

**SQL Output -**





**\* venuewise performance for RCB**.

**Query used:**

with cte as(

select m.Match\_Id,m.Team\_1,t1.Team\_Name as Team\_1\_Name,m.Team\_2,t2.Team\_Name as Team\_2\_Name,

m.Match\_Winner,t3.Team\_Name as Match\_Winner\_Name,m.Venue\_Id,v.Venue\_Name,m.Season\_Id,

s.Season\_Year

from Matches m join Team t1 on m.Team\_1=t1.Team\_Id

join Team t2 on m.Team\_2=t2.Team\_Id

join Team t3 on m.Match\_Winner=t3.Team\_Id

join Venue v on v.Venue\_Id=m.Venue\_Id

join Season s on s.Season\_Id=m.Season\_Id

),

cte1 as(

select Venue\_Id,Venue\_Name,count(Match\_Id) as Total\_Matches\_Played,

coalesce(sum(case when Match\_Winner\_Name='Royal Challengers Bangalore' then 1 end),0) as Total\_Matches\_Won\_By\_RCB,

coalesce(sum(case when Match\_Winner\_Name!='Royal Challengers Bangalore' then 1 end),0) as Total\_Matches\_Lost\_By\_RCB

from cte

where Team\_1\_Name='Royal Challengers Bangalore' or Team\_2\_Name='Royal Challengers Bangalore'

group by Venue\_Id,Venue\_Name

)

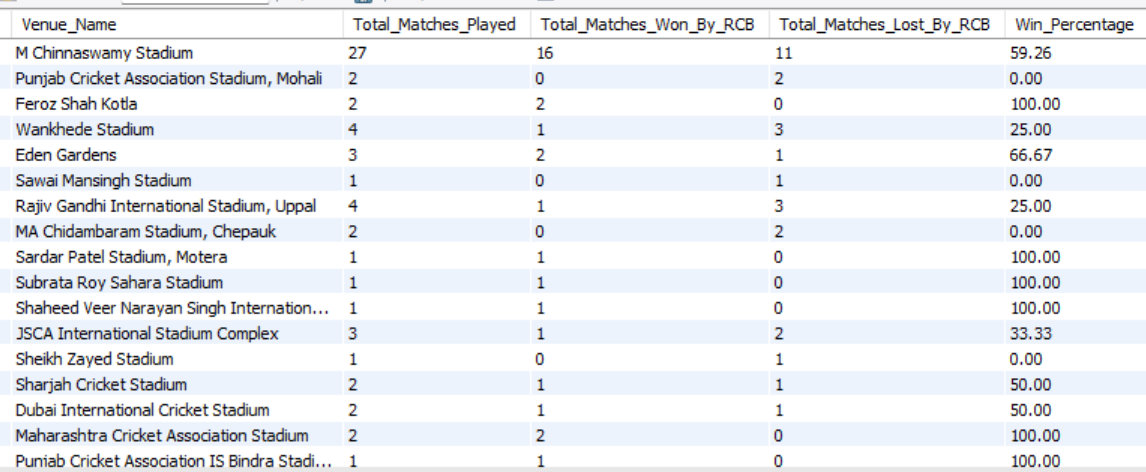
select Venue\_Id,Venue\_Name,Total\_Matches\_Played,Total\_Matches\_Won\_By\_RCB,Total\_Matches\_Lost\_By\_RCB,

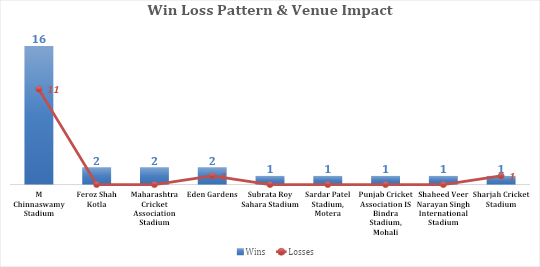
round(Total\_Matches\_Won\_By\_RCB\*100/Total\_Matches\_Played,2) as Win\_Percentage

from cte1

order by Venue\_Id

**SQL Output -**





**home/away performance.**

**Query used:**

WITH cte AS (

SELECT m.Match\_Id, m.Team\_1, t1.Team\_Name AS Team\_1\_Name, m.Team\_2, t2.Team\_Name AS Team\_2\_Name,m.Match\_Winner, t3.Team\_Name AS Match\_Winner\_Name,

m.Venue\_Id, v.Venue\_Name, m.Season\_Id, s.Season\_Year

FROM Matches m

JOIN Team t1 ON m.Team\_1 = t1.Team\_Id

JOIN Team t2 ON m.Team\_2 = t2.Team\_Id

JOIN Team t3 ON m.Match\_Winner = t3.Team\_Id

JOIN Venue v ON v.Venue\_Id = m.Venue\_Id

JOIN Season s ON s.Season\_Id = m.Season\_Id

),

cte\_summary AS (

SELECT

CASE WHEN Venue\_Name = 'M Chinnaswamy Stadium' THEN 'Home'

ELSE 'Away'

END AS Location\_Type,

COUNT(Match\_Id) AS Total\_Matches\_Played,

COALESCE(SUM(CASE WHEN Match\_Winner\_Name = 'Royal Challengers Bangalore' THEN 1 ELSE 0 END), 0) AS Total\_Wins,

COALESCE(SUM(CASE WHEN Match\_Winner\_Name != 'Royal Challengers Bangalore' THEN 1 ELSE 0 END), 0) AS Total\_Losses

FROM cte

WHERE Team\_1\_Name = 'Royal Challengers Bangalore' OR Team\_2\_Name = 'Royal Challengers Bangalore'

GROUP BY CASE WHEN Venue\_Name ='M Chinnaswamy Stadium' THEN 'Home'

ELSE 'Away'

END),

final\_summary AS (

SELECT Location\_Type, SUM(Total\_Matches\_Played) AS Total\_Matches,SUM(Total\_Wins) AS Total\_Wins,SUM(Total\_Losses) AS Total\_Losses,

ROUND(SUM(Total\_Wins) \* 100.0 / SUM(Total\_Matches\_Played), 2) AS Win\_Percentage

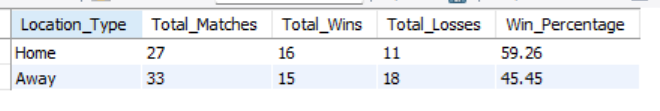
FROM cte\_summary

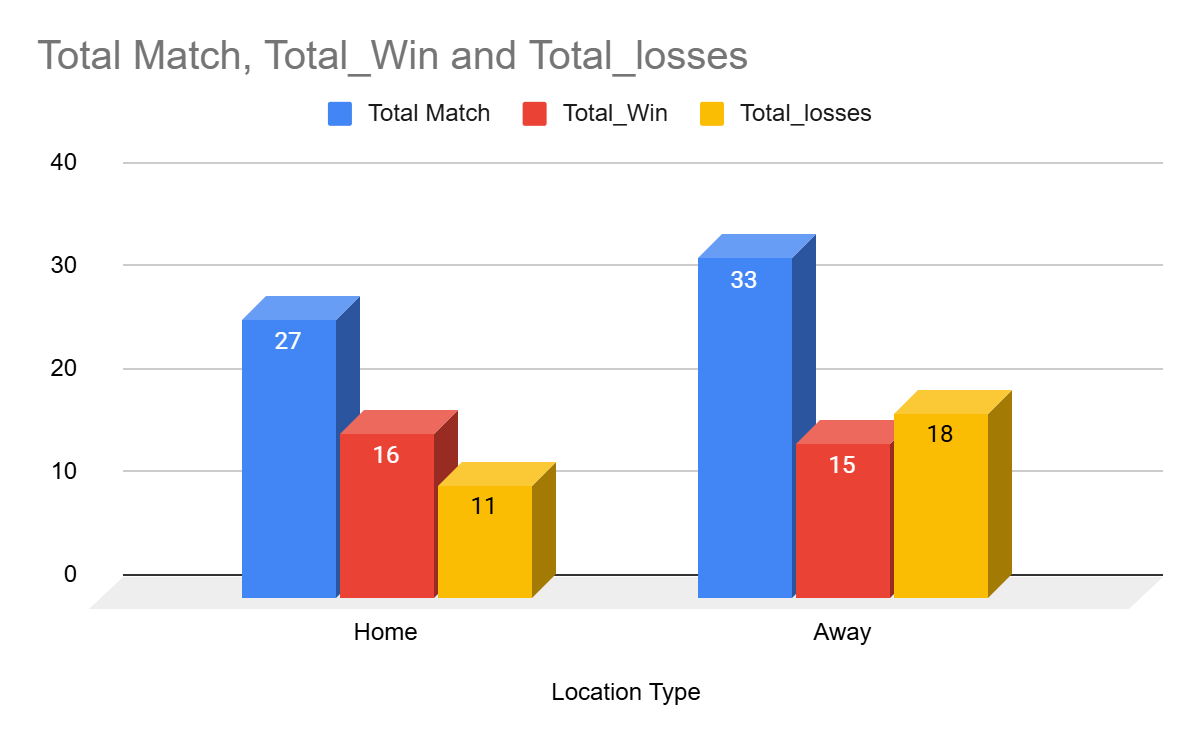
GROUP BY Location\_Type

)

SELECT \* FROM final\_summary;

**SQL Output -**





Question 10. How would you approach this problem, if the objective and subjective questions weren't given?

Ans. If the problem statement did not provide predefined objective and subjective questions, I would follow a structured, data-driven analytics approach to discover insights organically.

### Understand the Business Context

* Clarify the overall goal (e.g., improving RCB’s performance, auction strategy, match outcomes).
* Identify stakeholders (team management, analysts, coaches).
* Define success metrics (win percentage, playoff qualification, consistency).

### Data Understanding & Exploration

* Review available datasets (matches, players, venues, toss, seasons).
* Identify key entities: teams, players, venues, match conditions.
* Perform exploratory data analysis (EDA) to understand distributions, trends, and anomalies.

### Frame Analytical Questions from Data

* What factors correlate most with winning matches?
* Are there patterns related to toss, venue, or innings played?
* Which players consistently influence outcomes?
* Are there season-wise or venue-wise performance shifts?
* Toss decision significantly affects match outcomes at certain venues.
* Team balance (all-rounders vs specialists) impacts consistency.
* Home advantage plays a measurable role.

This ensures the analysis remains objective, exploratory, and strategically valuable, rather than being limited by initially framed questions.

Question 11. In the "Match" table, some entries in the "Opponent\_Team" column are incorrectly spelled as "Delhi\_Capitals" instead of "Delhi\_Daredevils". Write an SQL query to replace all occurrences of "Delhi\_Capitals" with "Delhi\_Daredevils".

Answer: SQL Query to Fix the Incorrect Entries:

UPDATE "Match"

SET Opponent\_Team = 'Delhi\_Daredevils'

WHERE Opponent\_Team = 'Delhi\_Capitals';