Objective Questions

1. List the different dtypes of columns in table “ball\_by\_ball”

(using information schema)

Solution:- Here is the query ---

**select distinct data\_type from information\_schema.columns**

**where table\_name = 'Ball\_by\_Ball' and table\_schema = 'ipl';**

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

Solution:- The out of the query is – **2558** and the query for it is.

**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;**

**where table\_name = 'Ball\_by\_Ball' and table\_schema = 'ipl';**

****

1. How many players were more than the age of 25 during season 2014?

Solution:- There are 89 players above the age 25 in season of 2014

and here is the query---

**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;**



1. How many matches did RCB win in 2013?

Solution:- RCB won 9 matches in 2013 and here is the query---

**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;**

****

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

Solution:- Here is the query for the list of top 10 players---

**with season\_id\_lastfouryears as (**

**select distinct Season\_Id from Season where Season\_Year>=(select max(Season\_Year)-3 from Season)**

**),**

**cte as (**

**select distinct Match\_Id from Matches where Season\_Id in ( select Season\_Id from season\_id\_lastfouryears)**

**),**

**player\_stats as**

**(**

**select Striker,sum(Runs\_Scored) as total\_runs\_scored,count(Ball\_Id) as Total\_balls\_faced**

**from Ball\_by\_Ball**

**where Match\_Id in (select Match\_Id from cte)**

**group by Striker**

**),**

**Strike\_Rate as**

**(select Striker,total\_runs\_scored,Total\_balls\_faced, round((total\_runs\_scored\*100/Total\_balls\_faced),2) as Strike\_Rate**

**from player\_stats)**

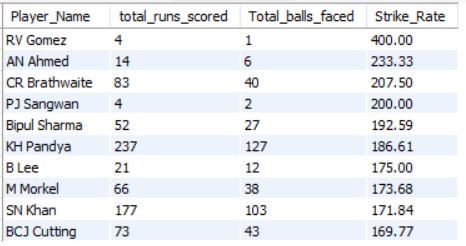
**select p.Player\_Name,s.total\_runs\_scored,s.Total\_balls\_faced,s.Strike\_Rate**

**from Strike\_Rate s join Player p**

**on s.Striker=p.Player\_Id**

**order by s.Strike\_Rate desc**

**limit 10;**

****

1. What are the average runs scored by each batsman considering all the seasons?

Solution:- Here is the query for average runs scored by each batsman ---

**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,**

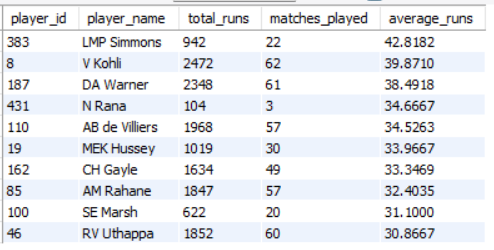
**coalesce(sum(runs\_scored) / nullif(count(distinct match\_id), 0), 0) as average\_runs**

**from cte**

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

**order by average\_runs desc**

**limit 10;**

****

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

Solution:- Here is the query ---

**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;**

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

Solution:- Here is the query ---

**-- -- -- For Batting -- -- --**

**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**

**),**

**cte1 as(**

**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**

**)**

**select \* from cte1 where**

**average\_runs>(select avg(average\_runs) from cte1);**

****

**-- -- -- For Bowling -- -- --**

**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**

**),**

**cte3 as (**

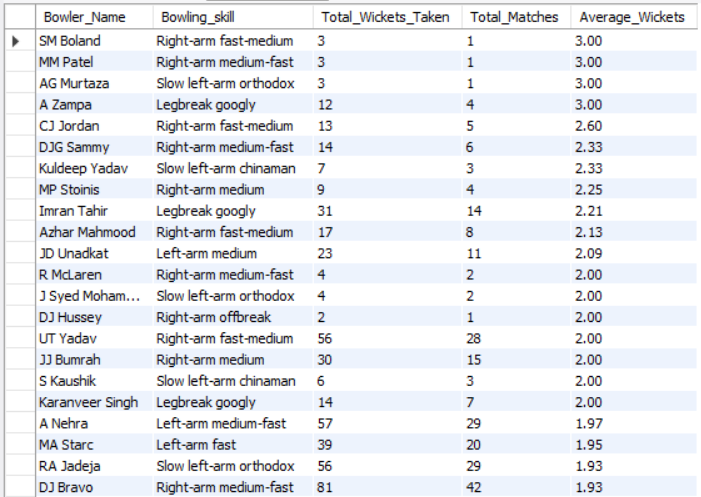
**select Bowler\_Name,Bowling\_skill,Total\_Wickets\_Taken,Total\_Matches,round((Total\_Wickets\_Taken)/(Total\_Matches),2) as Average\_Wickets**

**from cte2**

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

**)**

**select \* from cte3 where Average\_Wickets>(select avg(Average\_Wickets) from cte3);**



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

Solution:- Here is the query ---

**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'**

**)**

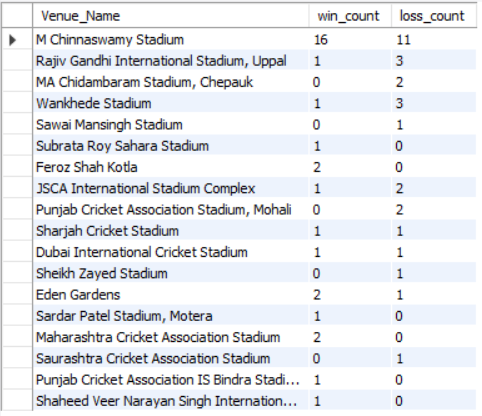
**select Venue\_Name, count(case when winner='Royal Challengers Bangalore' then 1 end) as win\_count,**

**count(case when winner != 'Royal Challengers Bangalore' then 1 end) as loss\_count**

**from cte1**

**group by Venue\_Name;**

**Table:-**



1. What is the impact of bowling style on wickets taken?

Solution:- Right-arm medium bowling, with its variations in pace and movement, is the most successful wicket-taking style. Off-spin and slow left-arm orthodox are also effective. While less common, left-arm pace and specialist spin (leg-break, chinaman) contribute, highlighting the value of bowling diversity.and here is the query ---

**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**

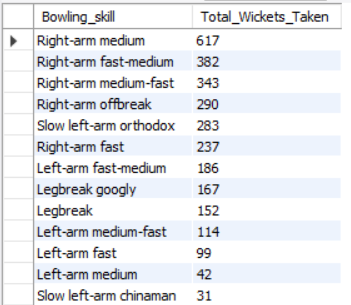
**)**

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

**from cte**

**group by Bowling\_skill**

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

****

1. Write the SQL query to provide a status of whether the performance of the team is better than the previous year's performance on the basis of the number of runs scored by the team in the season and the number of wickets taken

Solution:- Here is the query ---

**-- -- -- Number of Runs Scored -- -- --**

**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;**

**-- -- -- Number of Wickets Taken Year wise 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",**

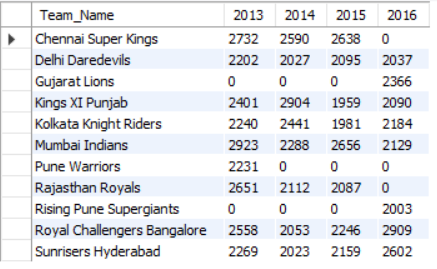
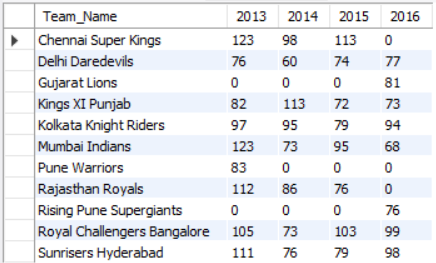
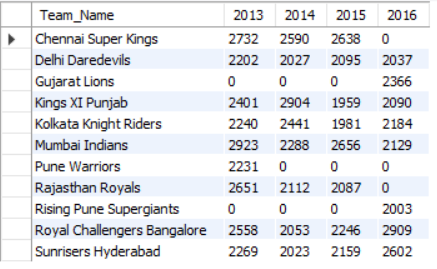
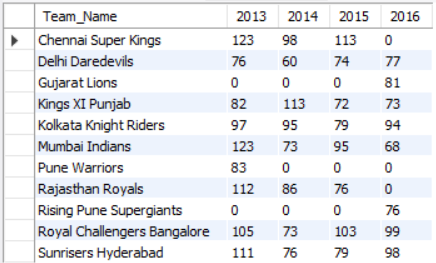
**sum(case when Year=2015 then Total\_Wickets\_Taken else 0 end) as "2015",**

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

**from cte2**

**group by Team\_Name**

**order by Team\_Name;**

**-- No. of Runs Scored** **per year-- -- No. of Wickets Taken** **per year--**

1. Can you derive more KPIs for the team strategy?

Solution:- here are the KPI’s that we can consider-

**KPI 1 :- Number of matches played, matches won, and win percentage for each team.**

**SELECT**

**t.Team\_Name,**

**COUNT(m.Match\_Id) as Total\_Matches,**

**sum(case when m.Match\_Winner = t.Team\_Id then 1 else 0 end) as Total\_Wins,**

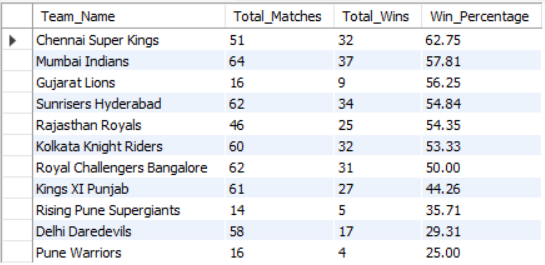
**round((sum(case when m.Match\_Winner = t.Team\_Id then 1 else 0 end) \* 100.0 / count(m.Match\_Id)),2) as Win\_Percentage**

**from Matches m**

**join Team t on m.Team\_1 = t.Team\_Id or m.Team\_2 = t.Team\_Id**

**group by t.Team\_Name**

**order by Win\_Percentage desc;**

****

**KPI 2 :- Players with the highest "Man of the Match" awards.**

**select**

**p.Player\_Name,**

**count(m.Man\_of\_the\_Match) as MOM\_Awards**

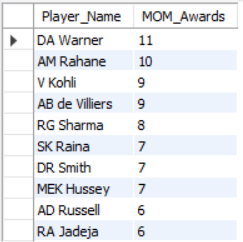
**from Matches m**

**join Player p on m.Man\_of\_the\_Match = p.Player\_Id**

**group by p.Player\_Name**

**order by MOM\_Awards desc**

**limit 10;**

****

**KPI 3 :- Top 10 player with highest runs scored.**

**select**

**p.Player\_Name,**

**sum(b.Runs\_Scored) as Total\_Runs**

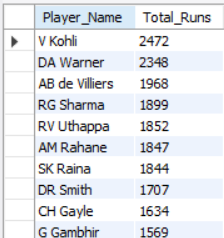
**from Ball\_by\_Ball b**

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

**group by p.Player\_Name**

**order by Total\_Runs desc**

**limit 10;**



**KPI 4 :- Top 10 player with highest wickets taken by each bowler.**

**select**

**p.Player\_Name,**

**count(w.Player\_Out) as Total\_Wickets**

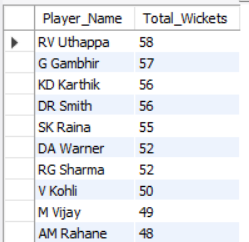
**from Wicket\_Taken w**

**join Player p on w.Player\_Out = p.Player\_Id**

**group by p.Player\_Name**

**order by Total\_Wickets desc**

**limit 10;**



**KPI 5 :- Toss wins and win percentage after winning the toss for each team.**

**select**

**t.Team\_Name,**

**count(m.Match\_Id) as Total\_Tosses\_Won,**

**sum(case when m.Toss\_Winner = m.Match\_Winner then 1 else 0 end) as Wins\_After\_Toss,**

**(sum(case when m.Toss\_Winner = m.Match\_Winner then 1 else 0 end) \* 100.0 / count(m.Match\_Id)) as Toss\_Win\_Percentage**

**from Matches m**

**join Team t on m.Toss\_Winner = t.Team\_Id**

**group by t.Team\_Name**

**order by Toss\_Win\_Percentage desc;**



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

Solution:- Here is the query ---

**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**

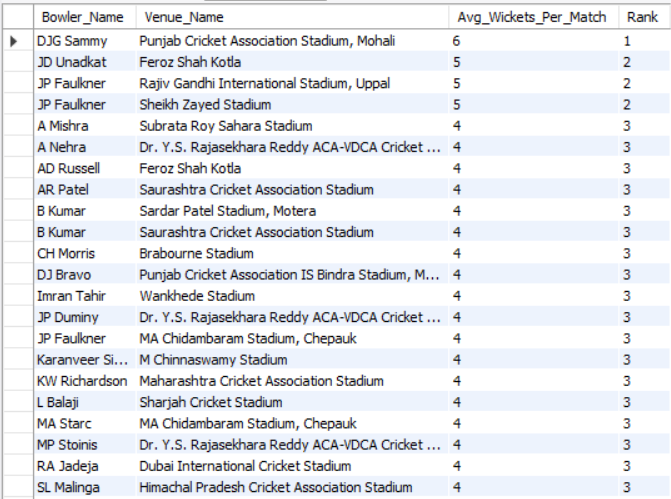
**Avg\_Wickets\_Per\_Match desc**

**)**

**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;**

****

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

Solution:- Here is the query ---

**-- -- -- With batting point of view -- -- --**

**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,**

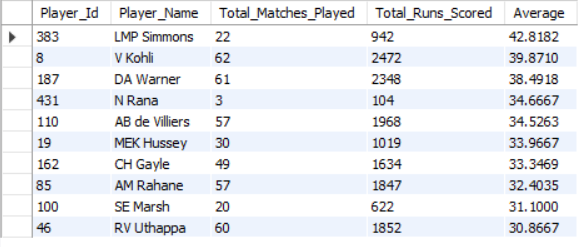
**sum(Runs\_Scored)/count(distinct Match\_Id) as Average**

**from cte**

**group by Striker,Player\_Name**

**order by Average desc**

**limit 10;**

****

****

**-- -- -- For 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,**

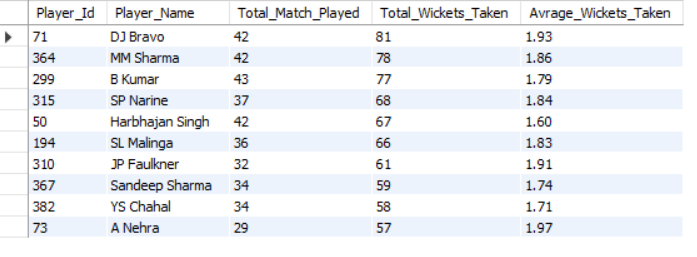
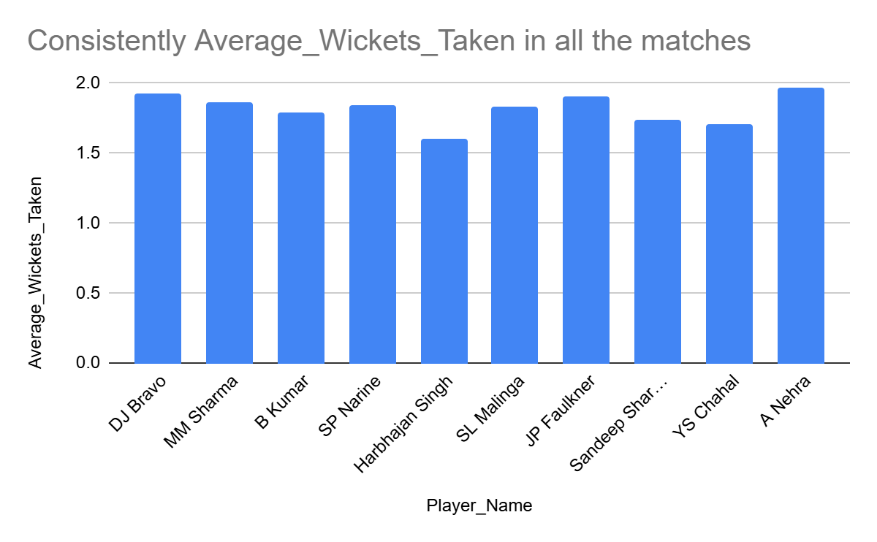
**round(count(\*) / count(distinct Match\_Id),2) as Avrage\_Wickets\_Taken**

**from cte**

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

**order by Total\_Wickets\_Taken desc**

**limit 10;**

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

Solution:- Here is the query ---

**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,**

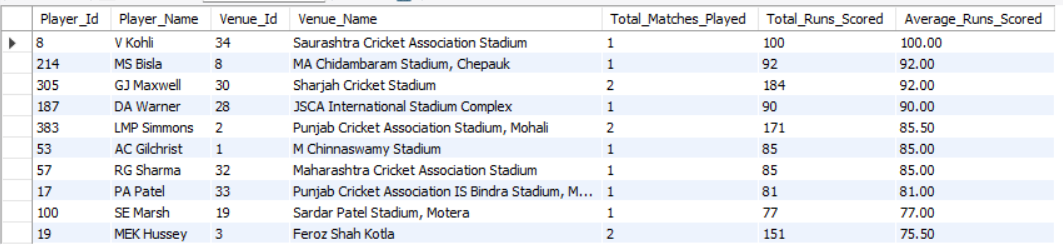
**round(sum(Runs\_Scored)/count(distinct Match\_Id),2) as Average\_Runs\_Scored**

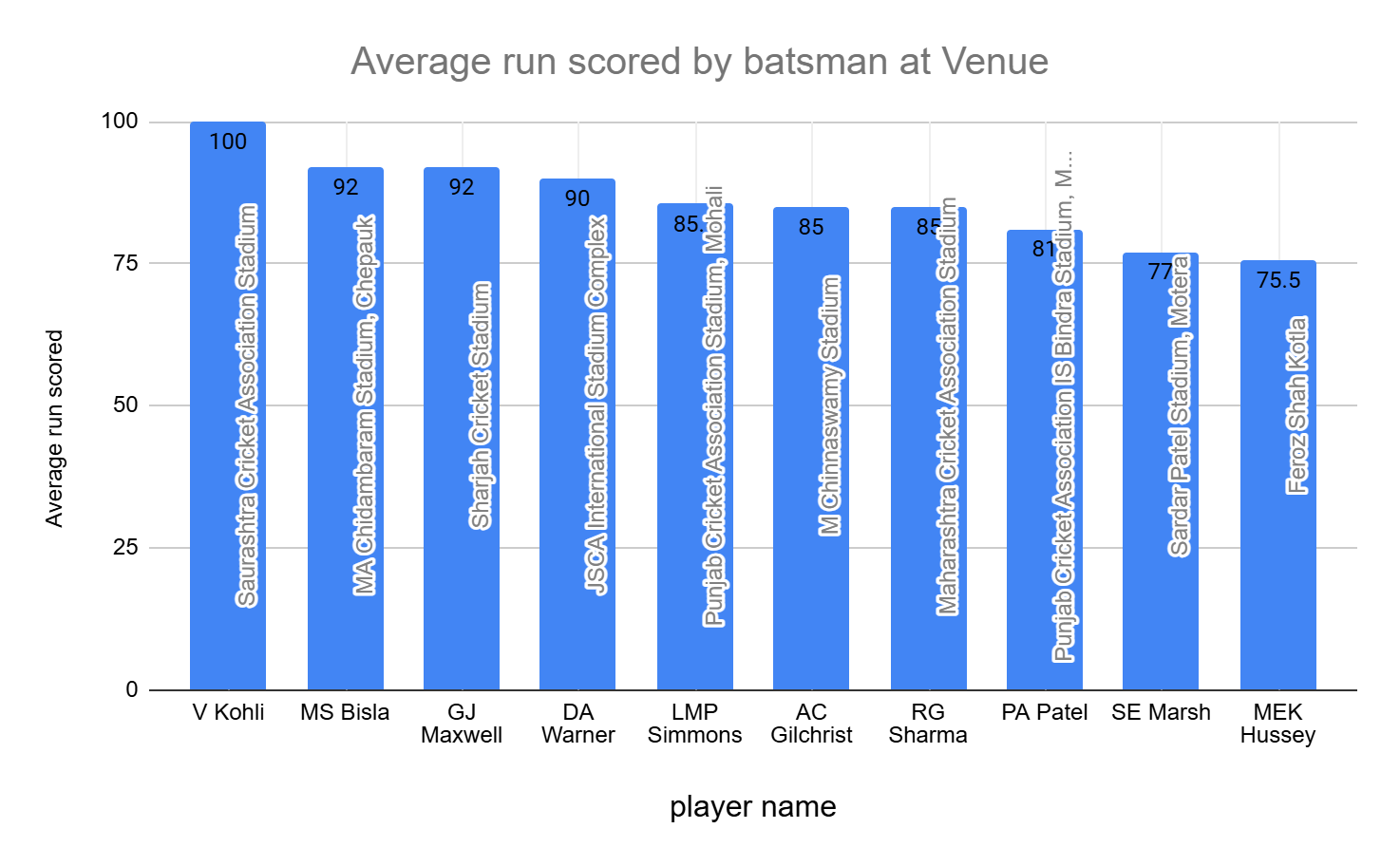
**from cte1**

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

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

**limit 10;**

****

****

**-- -- -- For 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,**

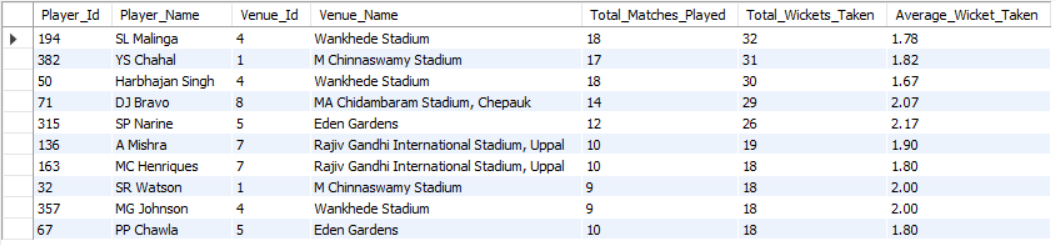
**count(\*) as Total\_Wickets\_Taken,**

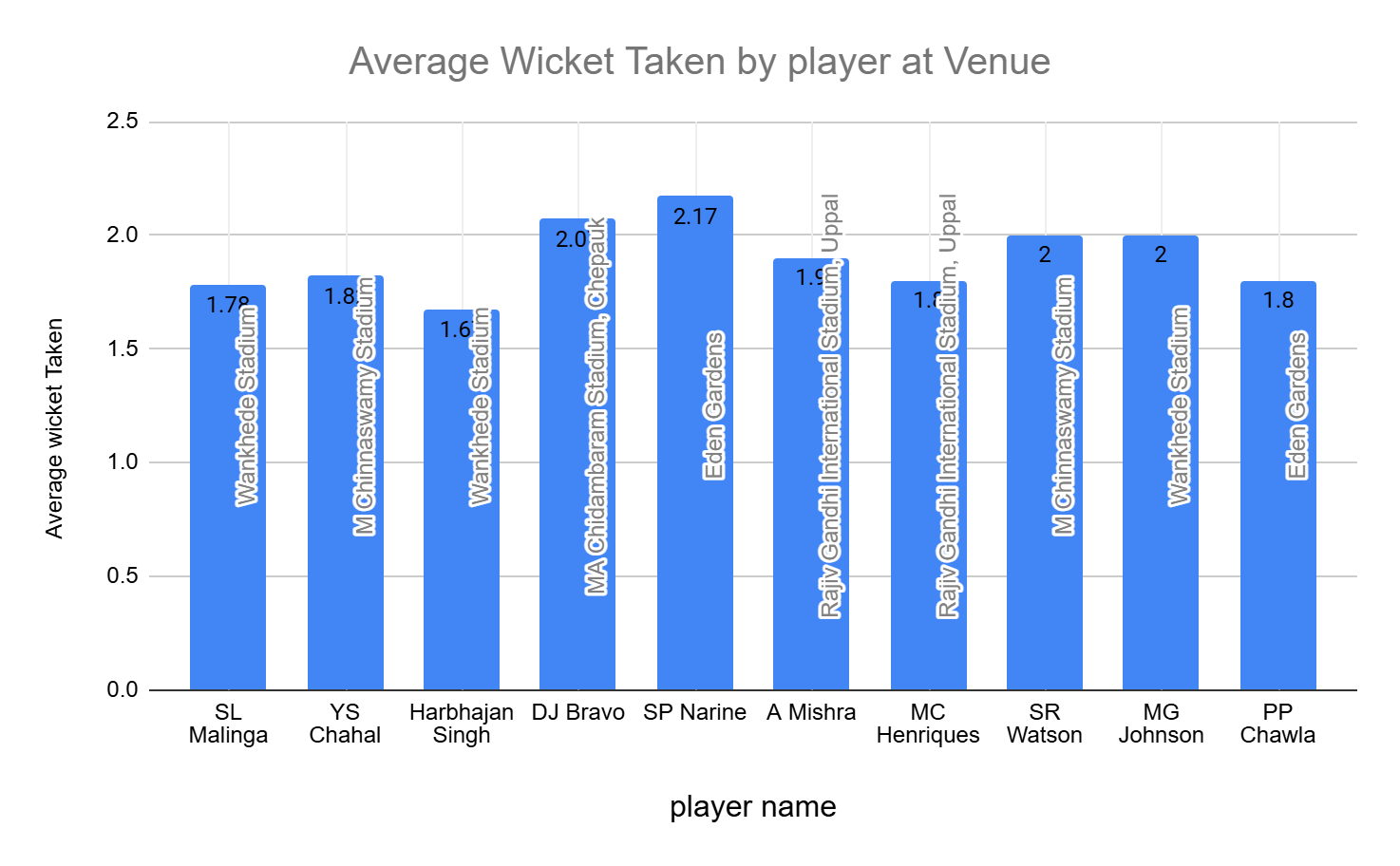
**round(count(\*) / count(distinct Match\_Id),2) as Average\_Wicket\_Taken**

**from cte1**

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

**order by Total\_Wickets\_Taken desc**

**limit 10;** ****

****

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?

Solution:- Here is the query---

**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;**

**Approach:**

* Used **Common Table Expressions (CTEs)** to categorize match results based on whether the toss-winning team won or lost the match.
* Aggregated results by venue to compare **Field First Wins vs. Bat First Wins**, highlighting venue-specific trends.

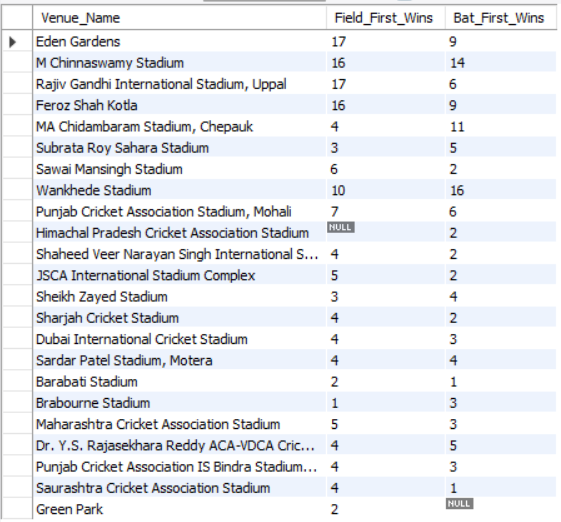
**Insights:**

* **Eden Gardens, Rajiv Gandhi International Stadium, and Feroz Shah Kotla** show a higher success rate for teams choosing to **field first**, indicating possible advantages like pitch conditions or dew factor.
* **MA Chidambaram Stadium and Wankhede Stadium** favor teams **batting first**, suggesting early pitch support or scoreboard pressure plays a role.

**Recommendations:**

* Teams should **analyze historical trends** before making a toss decision, as some venues favor specific strategies.
* Consider external factors such as **weather conditions, pitch type, and match format** to optimize decision-making.

**Visualizations:**

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

Solution:- Here is the query

**-- -- -- For Bowlers -- -- --**

**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;**

**-- -- -- For batting -- -- --**

**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;**

**Approach:**

* **Data-Driven Selection:** Identifies top performers using wickets taken and batting average.
* **Separate Analysis:** Differentiates bowlers and batsmen for role-specific evaluation.

**Insights:**

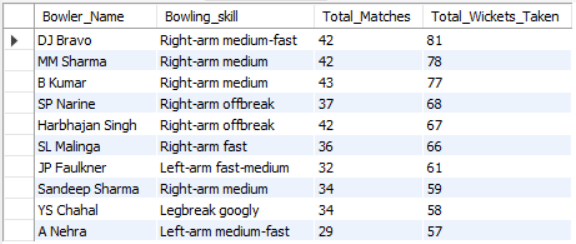
* **Top Wicket-Takers:** DJ Bravo, MM Sharma, and B Kumar consistently take wickets.
* **High Batting Average:** LMP Simmons and V Kohli excel in scoring runs consistently.

**Recommendations:**

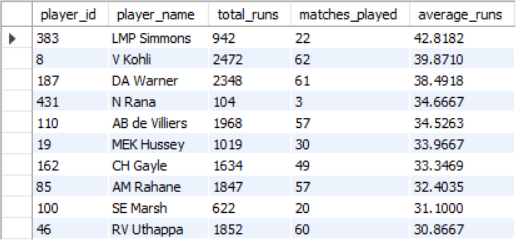
* **Balanced Team:** Select a mix of aggressive batsmen, consistent scorers, and effective bowlers.
* **Recent Form Matters:** Consider current form and match conditions, not just career stats.

**Visualizations:**

**-- -- -- For Bowlers -- -- --**

****

**-- -- -- For batting -- -- --**



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

Solution:- Here is the query

-- 1 -- Top 10 All-Rounders with highest Runs and highest Wickets

**WITH Batting AS (**

**SELECT p.Player\_Id, p.Player\_Name, SUM(bb.Runs\_Scored) AS Total\_Runs**

**FROM Ball\_by\_Ball bb**

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

**GROUP BY p.Player\_Id, p.Player\_Name**

**),**

**Bowling AS (**

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

**FROM Wicket\_Taken w**

**JOIN Player p ON w.Player\_Out = p.Player\_Id**

**GROUP BY p.Player\_Id**

**)**

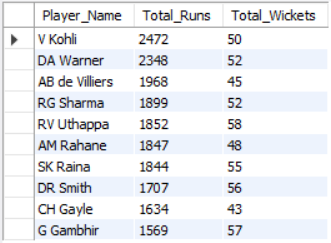
**SELECT b.Player\_Name, b.Total\_Runs, bw.Total\_Wickets**

**FROM Batting b**

**JOIN Bowling bw ON b.Player\_Id = bw.Player\_Id**

**WHERE b.Total\_Runs > 200 AND bw.Total\_Wickets > 10**

**ORDER BY b.Total\_Runs DESC , bw.Total\_Wickets DESC LIMIT 10;**



-- 2 -- Players with Best Strike Rate (Min 200 Runs)

**SELECT p.Player\_Name,**

**ROUND(SUM(bb.Runs\_Scored) \* 100.0 / SUM(bb.Ball\_Id), 2) AS Strike\_Rate**

**FROM Ball\_by\_Ball bb**

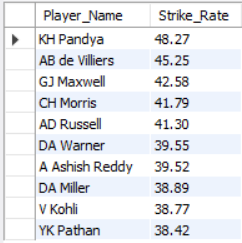
**JOIN Player p ON bb.Striker = p.Player\_Id**

**GROUP BY p.Player\_Name**

**HAVING SUM(bb.Runs\_Scored) > 200**

**ORDER BY Strike\_Rate DESC**

**LIMIT 10;**



-- 3 -- Most Economical Bowlers (Min 30 Overs Bowled)

**SELECT p.Player\_Name,**

**ROUND(SUM(bb.Runs\_Scored) / (COUNT(bb.Ball\_Id) / 6.0), 2) AS Economy\_Rate**

**FROM Ball\_by\_Ball bb**

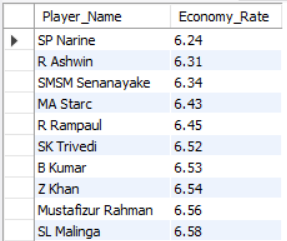
**JOIN Player p ON bb.Bowler = p.Player\_Id**

**GROUP BY p.Player\_Name**

**HAVING COUNT(bb.Ball\_Id) > 180 -- (180 balls = 30 overs)**

**ORDER BY Economy\_Rate ASC**

**LIMIT 10;**



**Approach:**

* **Data Aggregation:** Used SQL queries with GROUP BY and aggregate functions like SUM(), COUNT(), and ROUND() to calculate total runs, total wickets, strike rate, and economy rate.
* **Filtering & Ranking:** Applied HAVING clauses to ensure meaningful comparisons (e.g., minimum 200 runs for batsmen, 30 overs for bowlers) and used ORDER BY with LIMIT 10 to rank the top performers.

**Insights**:

* Balanced All-Rounders: **Players like V Kohli, DA Warner, and AB de Villiers excel in both batting and bowling, showing a rare balance of scoring high runs while taking crucial wickets.**
* Bowling Efficiency: **SP Narine and R Ashwin stand out as the most economical bowlers, maintaining low economy rates despite bowling a significant number of overs.**

#### **Recommendations:**

* **Strategic Player Utilization:** Teams should leverage high-strike-rate players like KH Pandya and AB de Villiers in powerplay and death overs to maximize scoring efficiency.
* **Bowling Attack Optimization:** Bowlers with low economy rates should be prioritized in middle overs to restrict the opposition's run rate and create pressure for wickets.

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)

Solution:- Here is the query—

**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, bs.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**

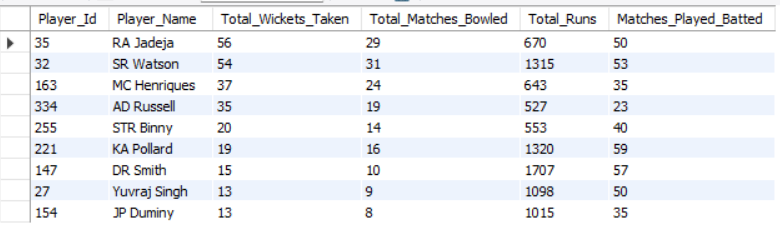
**JOIN Batting\_Data bt**

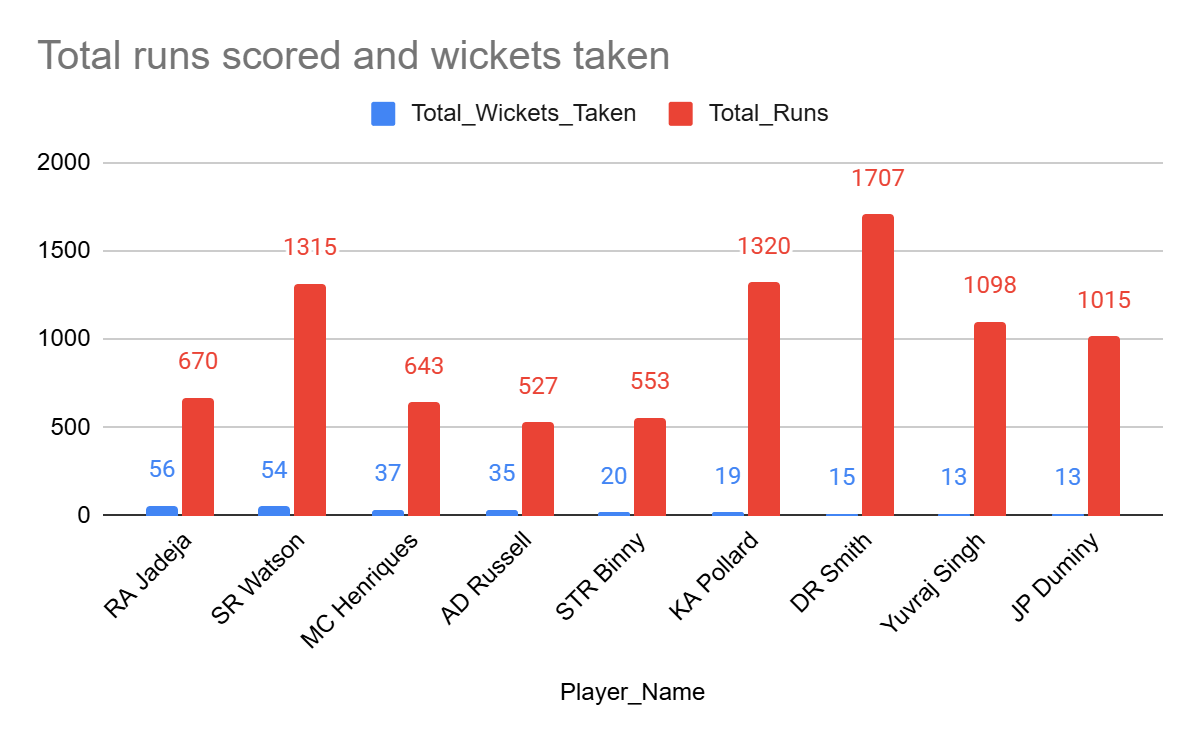
**ON b.Player\_Id = bt.player\_id**

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

**ORDER BY b.Total\_Wickets\_Taken DESC,**

**bt.Total\_Runs DESC;**





**Approach:**

* **Data Segmentation:** Separated batting and bowling statistics using CTEs (WITH clause) to calculate key metrics like total runs, wickets, and matches played.
* **Skill Filtering & Ranking:** Applied conditions (WHERE clause) to filter all-rounders with **more than 10 wickets and 500+ runs**, ranking them by total wickets and total runs.

#### **Insights:**

* Top-Performing All-Rounders: **Players like** RA Jadeja, SR Watson, and MC Henriques **consistently contribute in both batting and bowling, making them valuable assets.**
* Varied Contributions: **Some players like** KA Pollard and DR Smith **have high runs but relatively fewer wickets, indicating stronger batting dominance.**

#### **Recommendations:**

* Strategic Role Assignment: **Teams should** leverage true all-rounders **like SR Watson and AD Russell in high-pressure situations, as they can impact both innings.**
* Targeted Skill Development: **Players with** strong batting but lower wickets **(e.g., Pollard, Duminy) should focus on** enhancing bowling skills **to maximize their all-round effectiveness.**

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

Solution:- Here is the query

**-- -- -- Top 3 Batsman -- -- --**

**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,**

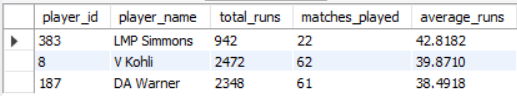
**coalesce(sum(runs\_scored) / nullif(count(distinct match\_id), 0), 0) as average\_runs**

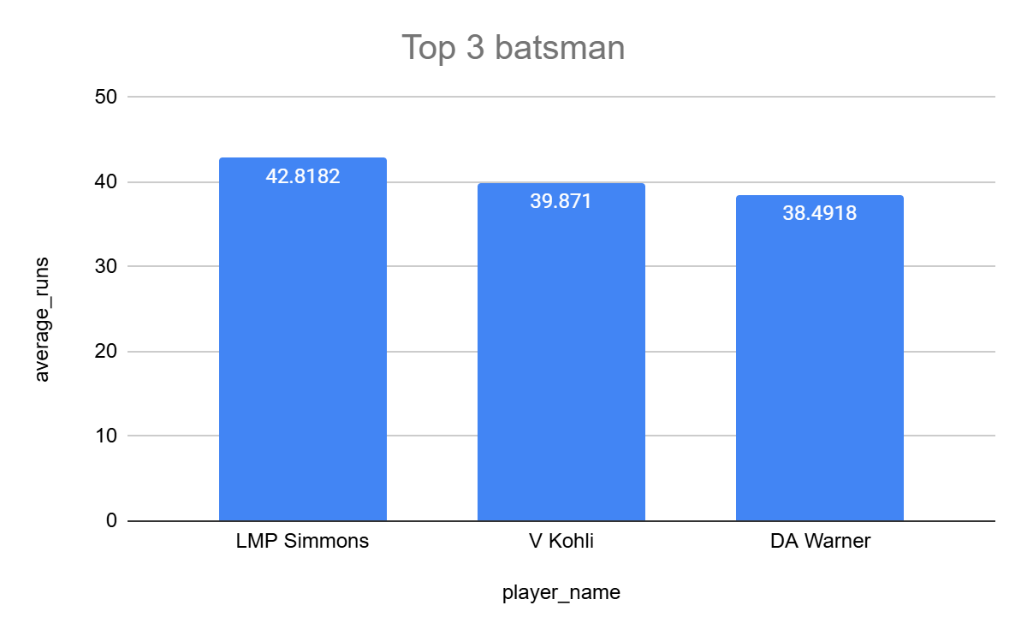
**from cte**

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

**order by average\_runs desc**

**limit 3;**

****

****

**-- -- -- Top 3 Bowlers -- -- --**

**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,**

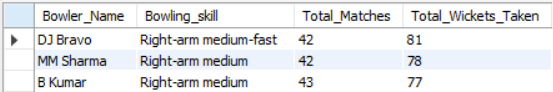
**COUNT(\*) AS Total\_Wickets\_Taken**

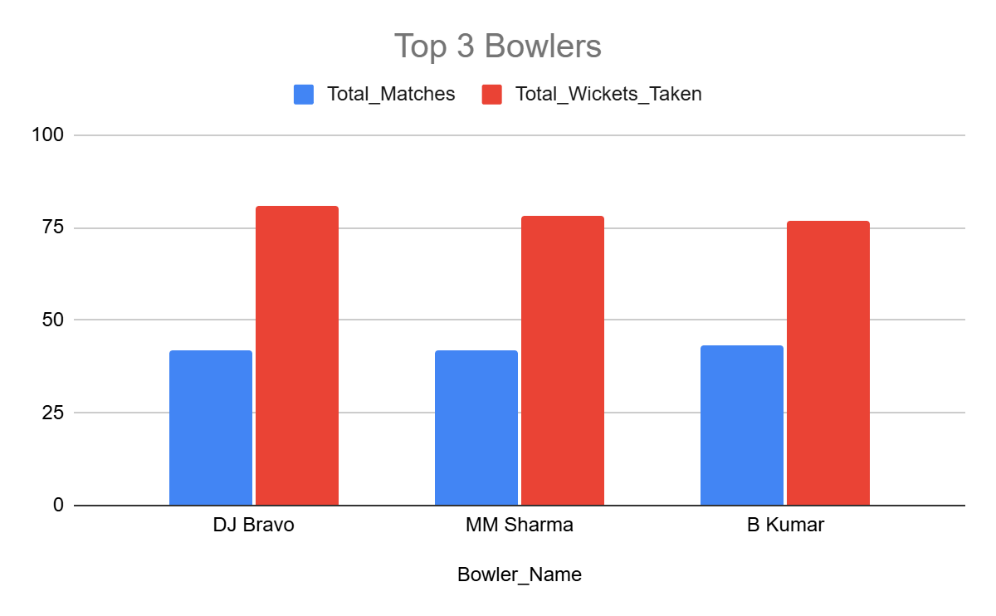
**FROM cte**

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

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

**limit 3;**

****

****

**-- -- -- Top 3 All Rounders -- -- --**

**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, bs.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**

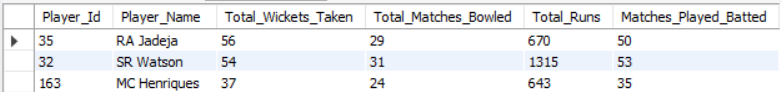
**JOIN Batting\_Data bt**

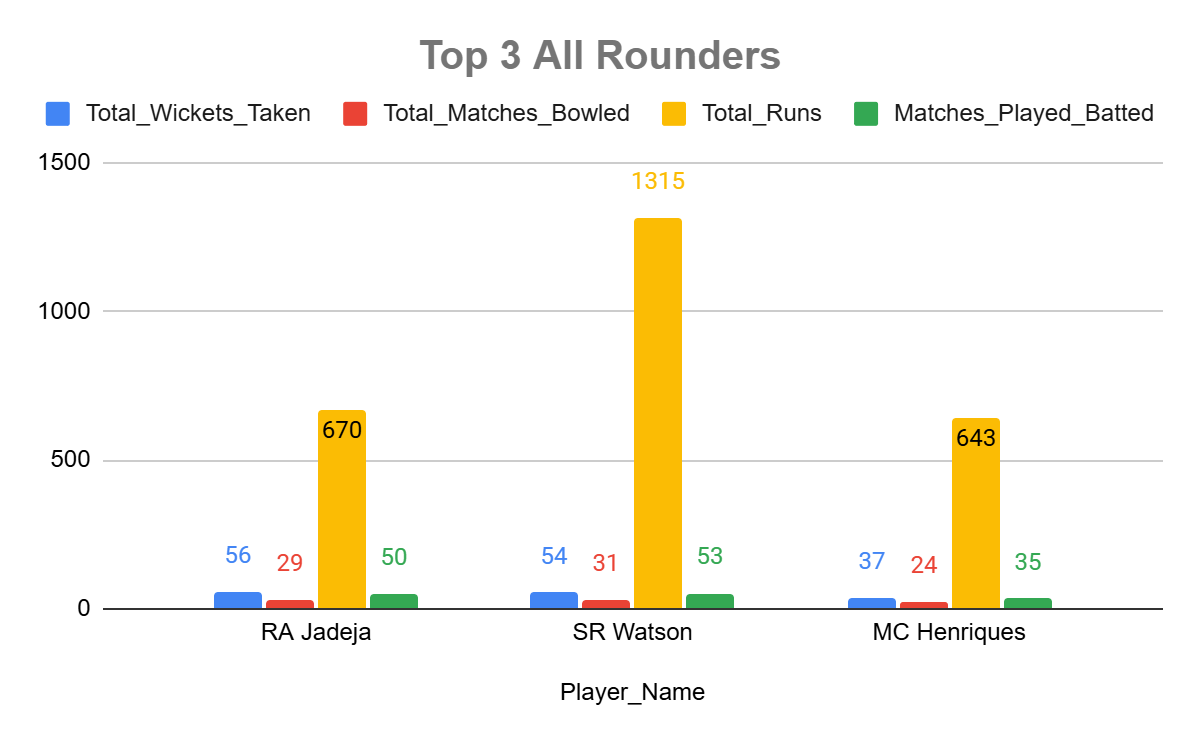
**ON b.Player\_Id = bt.player\_id**

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

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

**limit 3;**

****

****

#### **Approach:**

* Segmented Players into Three Categories: **Identified top** batsmen, bowlers, and all-rounders **based on performance metrics.**
* Performance Metrics: **Used** total runs, batting average, wickets taken, and matches played **to evaluate player impact.**
* Ranking & Selection: **Sorted results in descending order to determine the top** three **players in each category.**

#### **Insights:**

* **Key Players Uplifting Team Morale:**
  + **V Kohli & DA Warner** consistently perform with the bat, maintaining an **average above 38** and scoring **2000+ runs**, making them match-winners.
  + **DJ Bravo** (81 wickets) and **B Kumar** (77 wickets) play a crucial role in controlling the game with the ball.
  + **All-rounders like RA Jadeja & SR Watson** contribute significantly in both disciplines, making them invaluable assets.
* **Balanced Team Success:**
  + A team with impactful **batting, bowling, and all-round options** like Kohli, Bravo, and Watson creates a strong foundation for sustained performance.

#### **Recommendations:**

* Leverage Experience in Key Moments:
  + **Utilize** Kohli and Warner **in pressure chases to stabilize innings.**
  + **Deploy** Bravo and Kumar **in death overs for maximum bowling impact.**
* Optimize All-Rounders for Game Strategy:
  + **Players like** RA Jadeja & SR Watson **should be strategically placed in playing XIs to ensure depth in** both batting & bowling**, boosting team morale and adaptability.**

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

Solution:- Here is the query

**-- -- -- Top 10 All Rounders -- -- --**

**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, bs.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**

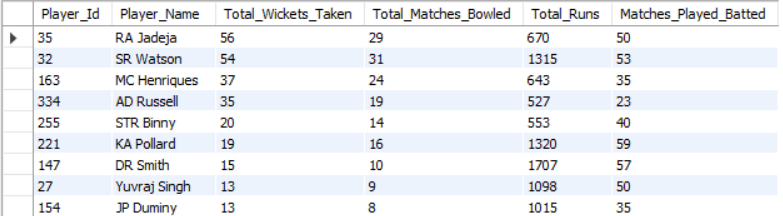
**FROM Bowling\_Data b**

**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;**

****

**-- -- -- death bowlers -- -- --**

**with cte as(**

**select b.Bowler,b.Over\_Id,b.Ball\_Id,b.Innings\_No,b.Match\_Id,(b.Runs\_Scored+coalesce(e.Extra\_Runs,0)) as Total\_Runs\_Scored**

**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**

**),**

**last\_five\_overs as (select \* from cte where Over\_Id>=15),**

**bowler\_stats AS (**

**SELECT Bowler, SUM(Total\_Runs\_Scored) AS Total\_Runs,**

**COUNT(\*) AS Total\_Balls**

**FROM last\_five\_overs**

**WHERE Total\_Runs\_Scored IS NOT NULL**

**GROUP BY Bowler**

**),**

**economy as (**

**SELECT Bowler,Total\_Runs,**

**round(Total\_Balls / 6.0,2) AS Overs\_Bowled,**

**round(Total\_Runs / (Total\_Balls / 6.0),2) AS Economy\_Rate**

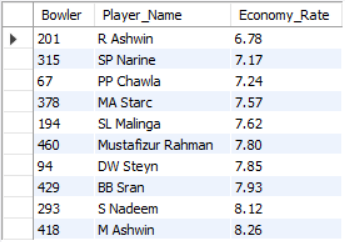
**FROM bowler\_stats**

**),**

**cte1 as(select \* from economy where Overs\_Bowled>10 order by Economy\_Rate limit 10)**

**select c.Bowler,p.Player\_Name,c.Economy\_Rate**

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

****

**-- -- -- Power Hitters -- -- --**

**with season\_id\_lastfouryears as (**

**select distinct Season\_Id from Season where Season\_Year>=(select max(Season\_Year)-3 from Season)**

**),**

**cte as (select distinct Match\_Id from Matches where Season\_Id in**

**( select Season\_Id from season\_id\_lastfouryears)),**

**player\_stats as**

**(select Striker,sum(Runs\_Scored) as total\_runs\_scored,count(Ball\_Id) as Total\_balls\_faced**

**from Ball\_by\_Ball**

**where Match\_Id in (select Match\_Id from cte)**

**group by Striker),**

**Strike\_Rate as**

**(select Striker,total\_runs\_scored,Total\_balls\_faced, round((total\_runs\_scored\*100/Total\_balls\_faced),2) as Strike\_Rate**

**from player\_stats)**

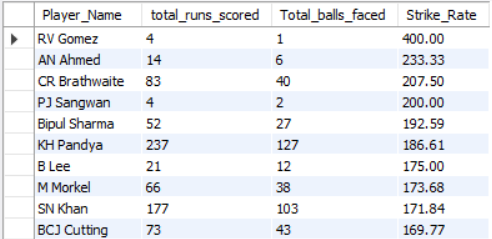
**select p.Player\_Name,s.total\_runs\_scored,s.Total\_balls\_faced,s.Strike\_Rate**

**from Strike\_Rate s join Player p**

**on s.Striker=p.Player\_Id**

**order by s.Strike\_Rate desc**

**limit 10;**



**Approach:**

* Analyzed data to identify top-performing **all-rounders, death bowlers, and power hitters** using SQL queries.
* Focused on key metrics such as **wickets taken, economy rate, strike rate, and total runs scored** to determine the best options for RCB.

**Insights:**

* **RA Jadeja, AD Russell, and SR Watson** stand out as top all-rounders, providing both **batting and bowling depth**.
* **R. Ashwin, Starc, and Mustafizur Rahman** are among the most economical death bowlers, crucial for RCB’s **late-over bowling strategy**.

**Recommendations:**

* **Target a balanced all-rounder** like **RA Jadeja or AD Russell**, who can contribute significantly with both bat and ball.
* **Strengthen the middle order** by acquiring power hitters like **KH Pandya or CR Brathwaite** to accelerate scoring in the final overs.

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

Solution:- Here is the query

**with cte as(**

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

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

**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 Match\_Id,Innings\_No,sum(Total\_Runs\_Scored) as Total\_Runs\_Scored**

**from cte**

**group by Match\_Id,Innings\_No**

**),**

**cte2 as (**

**select c1.Match\_Id,c1.Innings\_No,c1.Total\_Runs\_Scored,**

**m.Venue\_Id**

**from cte1 c1 join Matches m**

**on c1.Match\_Id=m.Match\_Id**

**),**

**cte3 as(**

**select Venue\_Id,Innings\_No,round(avg(Total\_Runs\_Scored),2) as Avg\_Runs\_Scored**

**from cte2**

**group by Venue\_Id,Innings\_No**

**),**

**cte4 as(**

**select Venue\_Id,**

**avg(case when Innings\_No=1 then Avg\_Runs\_Scored end) as First\_Inning\_Avg\_Score,**

**avg(case when Innings\_No=2 then Avg\_Runs\_Scored end) as Second\_Inning\_Avg\_Score**

**from cte3**

**group by Venue\_Id**

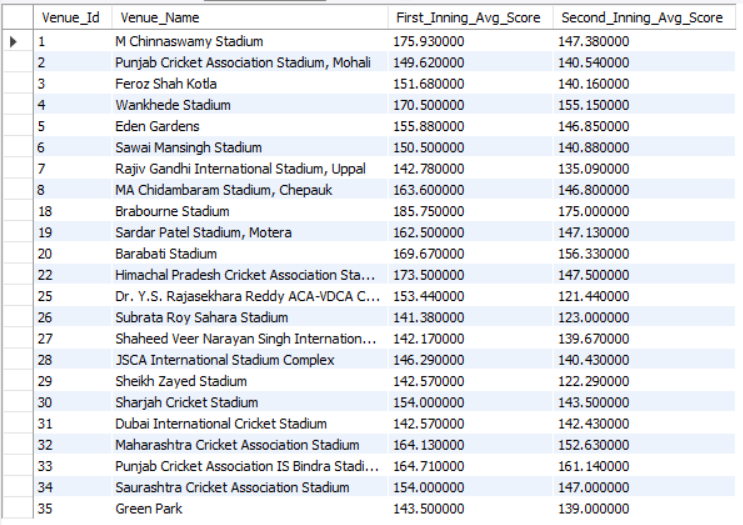
**)**

**select c4.Venue\_Id,v.Venue\_Name,c4.First\_Inning\_Avg\_Score,**

**c4.Second\_Inning\_Avg\_Score**

**from cte4 c4 join Venue v**

**on c4.Venue\_Id=v.Venue\_Id;**

****

**-- -- -- Players with Highest Runs -- -- --**

**with cte as (**

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

**b.Runs\_Scored**

**from Ball\_by\_Ball b**

**),**

**cte1 as (**

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

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

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

**from cte**

**group by Striker**

**order by Avg\_Runs desc**

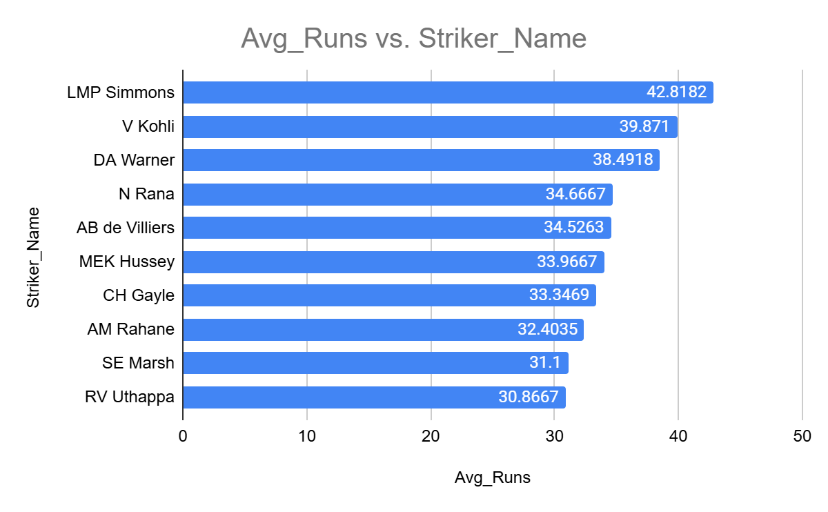
**limit 10**

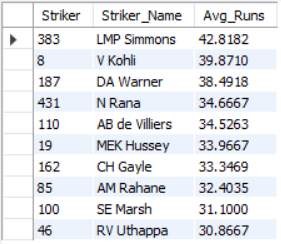
**)**

**select c1.Striker,p.Player\_Name as Striker\_Name,c1.Avg\_Runs**

**from cte1 c1 join Player p**

**on c1.Striker=p.Player\_Id;**

****

****

**-- -- -- Players with highest Strike Rates -- -- --**

**with cte as (**

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

**b.Runs\_Scored**

**from Ball\_by\_Ball b**

**),**

**cte1 as (**

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

**count(Ball\_Id) as Total\_Balls\_Played,**

**sum(Runs\_Scored)\*100/count(Ball\_Id) as Strike\_Rate**

**from cte**

**group by Striker**

**order by Strike\_Rate desc**

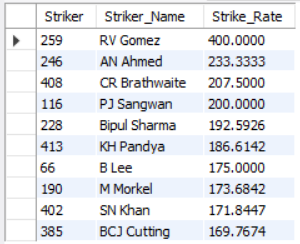
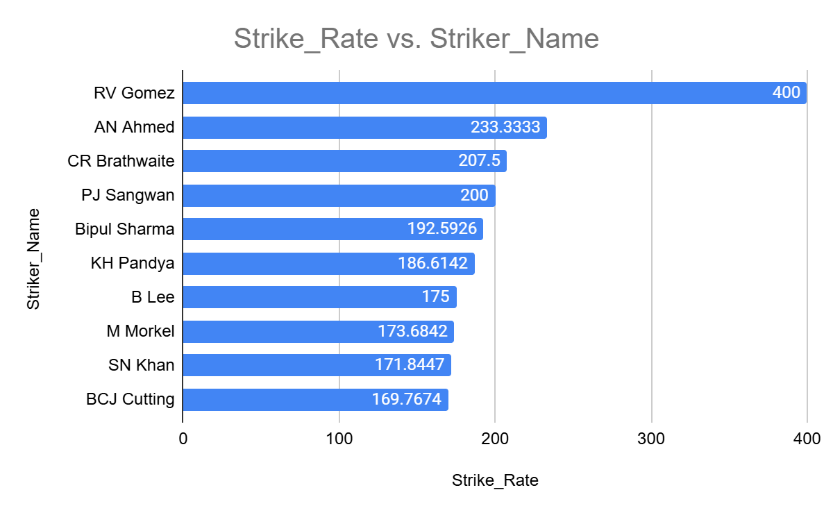
**limit 10**

**)**

**select c1.Striker,p.Player\_Name as Striker\_Name,c1.Strike\_Rate**

**from cte1 c1 join Player p**

**on c1.Striker=p.Player\_Id;**

****

### **Approach:**

* **Venue-Based Scoring Analysis:** Used SQL queries to calculate the average first and second innings scores at different venues to identify high-scoring grounds.
* **Player Performance Analysis:** Extracted top players based on average runs per match and strike rates to understand batting trends contributing to high scores.

### **Insights:**

* **Impact of Venue on Scoring:** Certain venues like M Chinnaswamy Stadium and Brabourne Stadium consistently produce high first and second innings scores, indicating batter-friendly conditions.
* **Aggressive Batting Trends:** Players with high strike rates (e.g., RV Gomez with 400.00) highlight the shift towards aggressive batting strategies in T20 matches.

### **Recommendations:**

* **Venue-Specific Strategies:** Teams should adapt their batting and bowling strategies based on historical scoring patterns at different venues.
* **Team Selection & Match Strategy:** Selecting power-hitters and aggressive batsmen for high-scoring venues can maximize match-winning potential.

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

Solution:- Here is the query

**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**

**),**

**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;**



**Approach:**

* **Home-Ground Performance Analysis:** Queried match data to identify the number of matches RCB played at M Chinnaswamy Stadium and their win count.
* **Win Percentage Calculation:** Compared RCB’s total matches and wins at their home ground to evaluate the impact of home-ground advantage.

**Insights:**

* **Strong Home Performance:** RCB has won **16 out of 27** matches at M Chinnaswamy Stadium, showing a **59.26% win rate**, indicating a significant home-ground advantage.
* **Favorable Conditions:** The venue is known for its **batting-friendly pitch**, which aligns well with RCB’s historically strong batting lineups.

**Recommendations:**

* **Leverage Pitch Conditions:** RCB should focus on **power-hitters and aggressive stroke players** to capitalize on the high-scoring nature of the ground.
* **Strategic Bowling Adjustments:** Given the batting-friendly conditions, investing in **death-over specialists and spinners** can help counter opposition strategies and maximize the home advantage.

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.

Solution:- Here is the query

**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,**

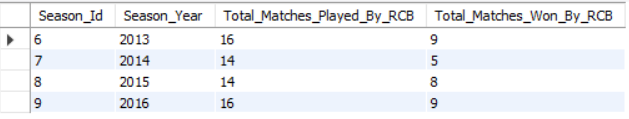
**sum(case when Match\_Winner\_Name='Royal Challengers Bangalore' then 1 end) as Total\_Matches\_Won\_By\_RCB**

**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;**

****

****

**-- -- -- Number of Runs Scored Yearwise by RCB -- -- --**

**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**

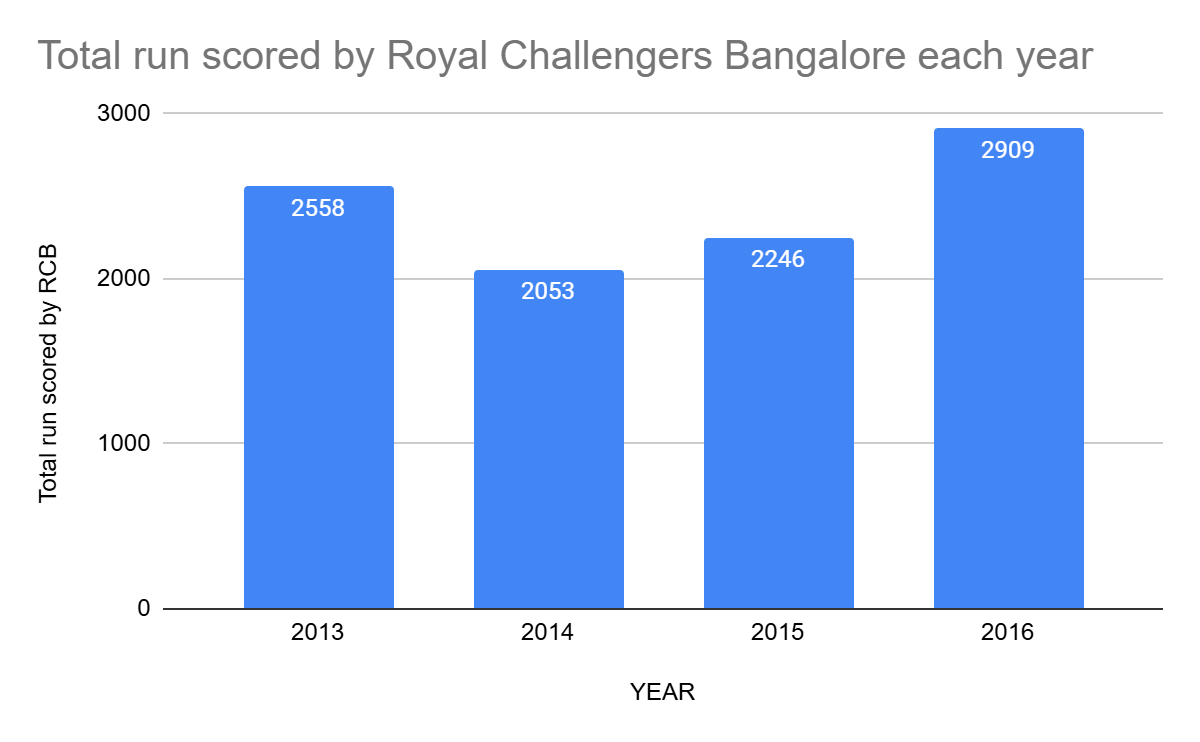
**group by team\_name**

**order by team\_name**

**)**

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

****

****

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

**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**

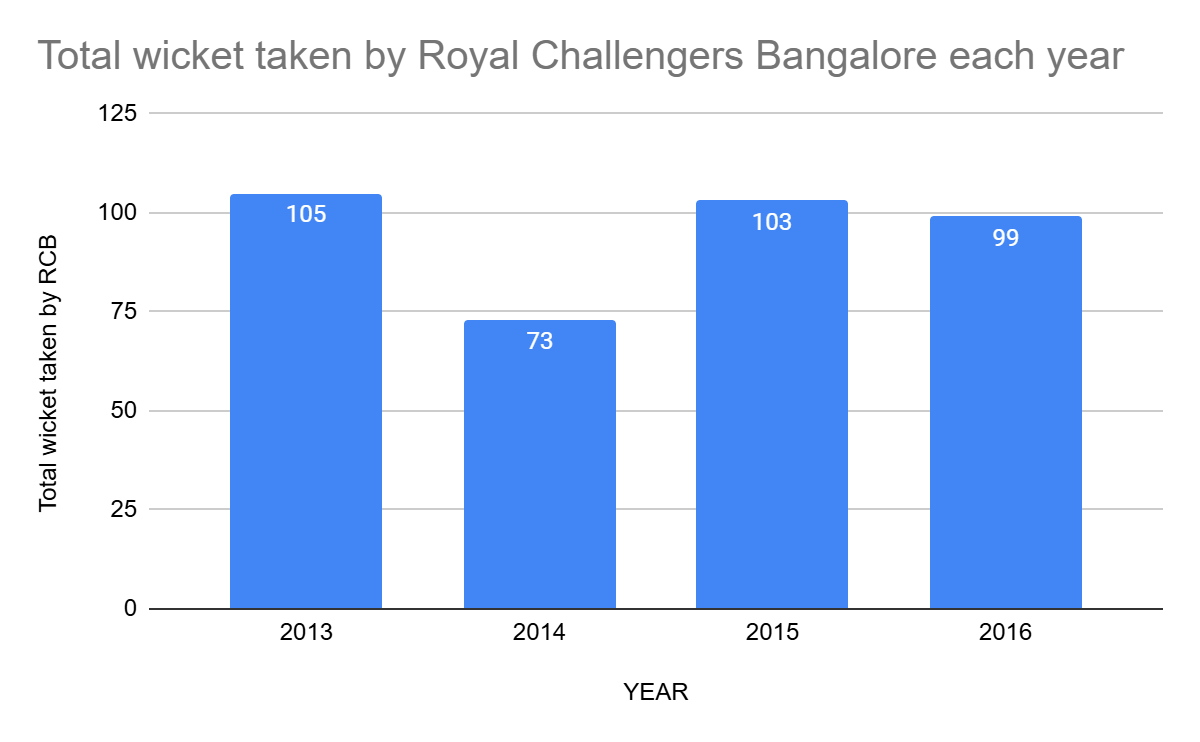
**group by Team\_Name**

**order by Team\_Name**

**)**

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

****

****

**-- -- -- venuewise performance for RCB -- -- --**

**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**

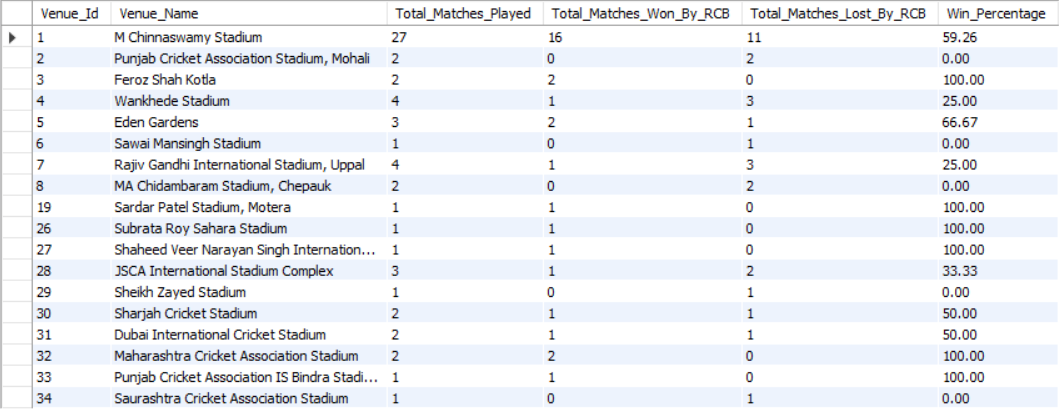
**)**

**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;**

****

**-- -- -- home/away performance -- -- --**

**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,**

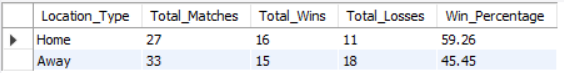
**ROUND(SUM(Total\_Wins) \* 100.0 / SUM(Total\_Matches\_Played), 2) AS Win\_Percentage**

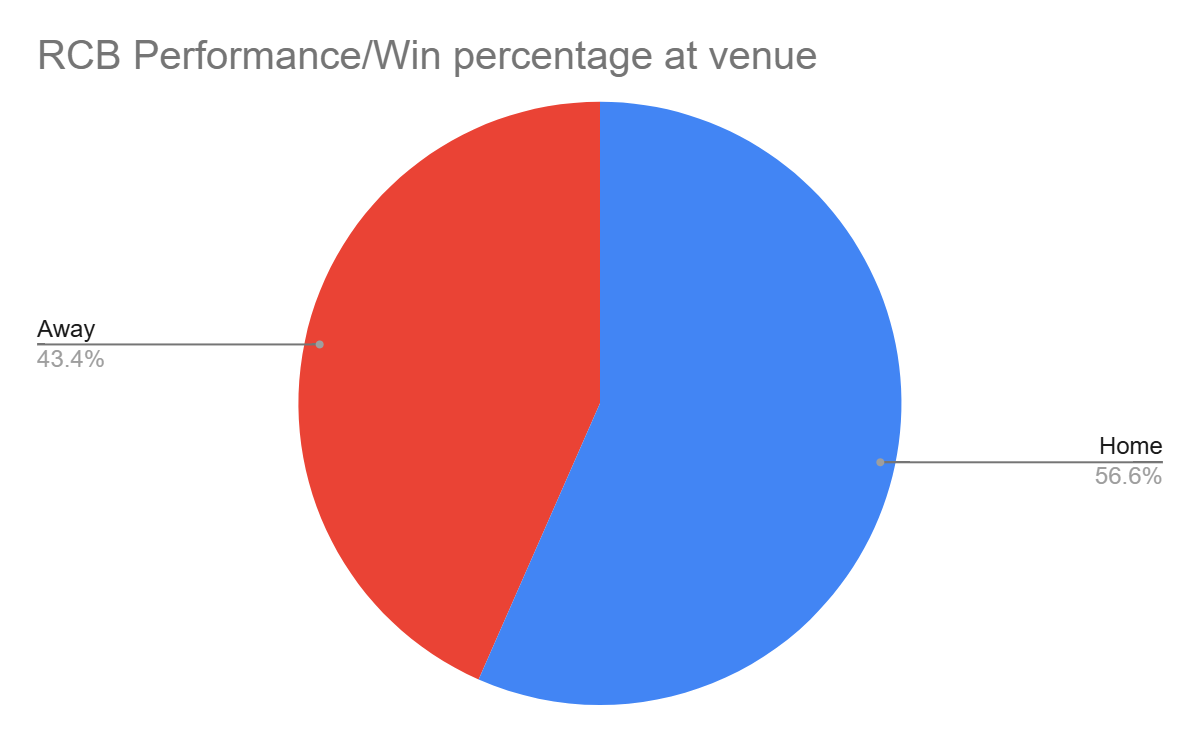
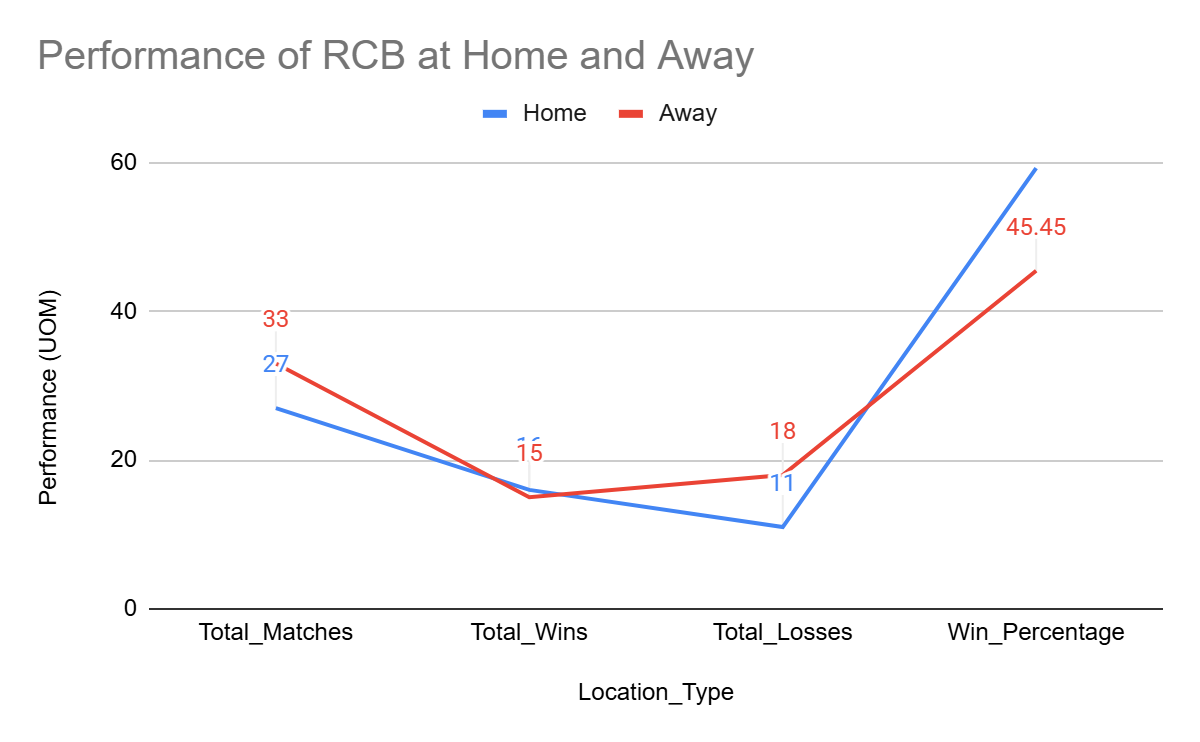
**FROM cte\_summary**

**GROUP BY Location\_Type**

**)**

**SELECT \* FROM final\_summary;**

****

****

**-- -- -- chasing/defending -- -- --**

**with cte as (**

**select m.Match\_Id,m.Team\_1,t.Team\_Name as Team1\_Name,m.Team\_2,t1.Team\_Name as Team2\_Name,m.Match\_Winner,t2.Team\_Name as Match\_Winner\_Name,**

**w.Win\_Type**

**from Matches m join Win\_By w**

**on m.Win\_Type=w.Win\_Id**

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

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

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

**),**

**cte1 as(**

**select Match\_Id,Team1\_Name,Team2\_Name,Match\_Winner\_Name,Win\_Type**

**from cte where Team1\_Name='Royal Challengers Bangalore' or Team2\_Name='Royal Challengers Bangalore'**

**)**

**select "RCB" as Team\_Name,count(case when Match\_Winner\_Name='Royal Challengers Bangalore' then Match\_Id end) as Total\_Wins,**

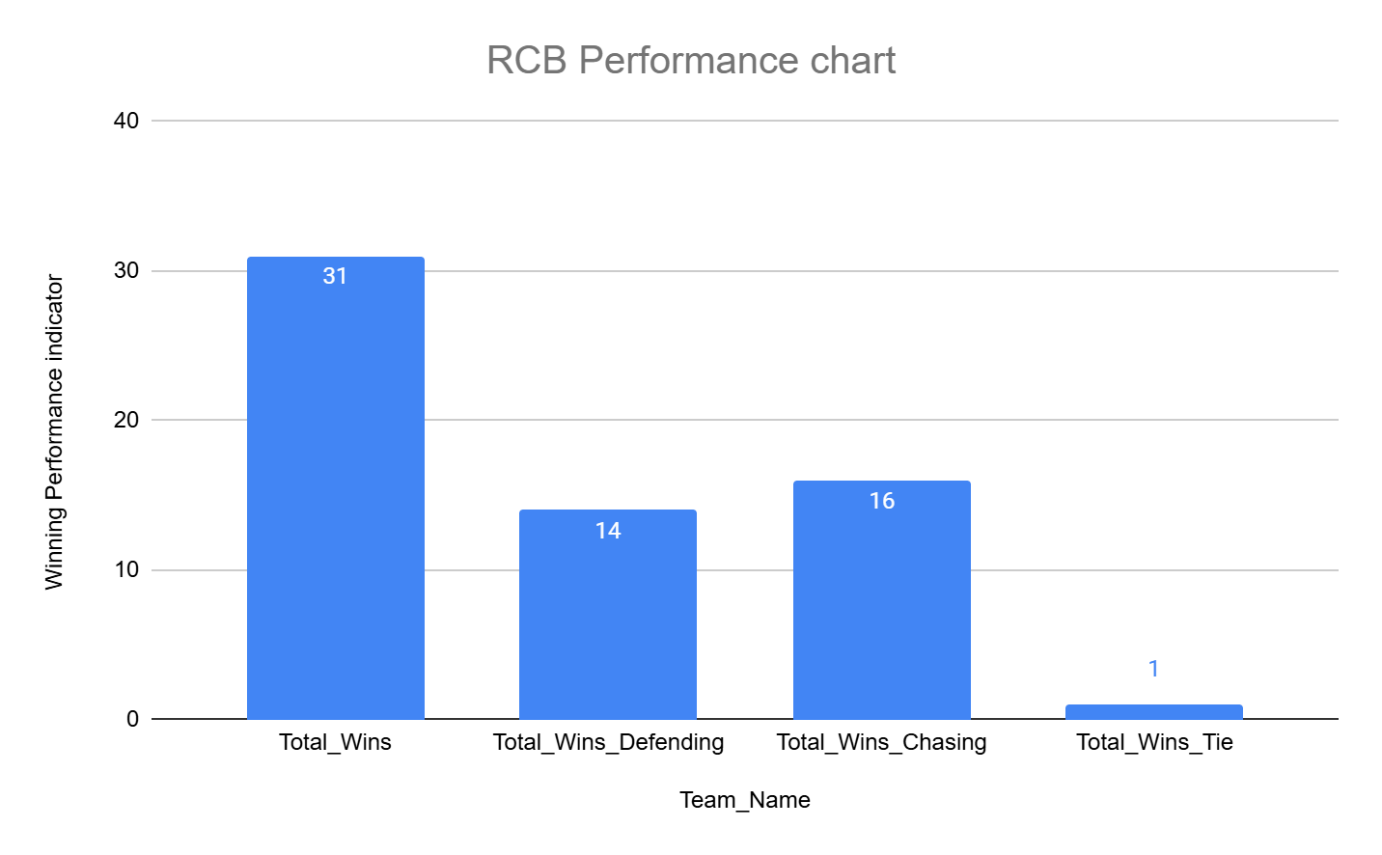
**count(case when Match\_Winner\_Name='Royal Challengers Bangalore' and Win\_Type='runs' then Match\_Id end) as Total\_Wins\_Defending,**

**count(case when Match\_Winner\_Name='Royal Challengers Bangalore' and Win\_Type='wickets' then Match\_Id end) as Total\_Wins\_Chasing,**

**count(case when Match\_Winner\_Name='Royal Challengers Bangalore' and Win\_Type='Tie' then Match\_Id end) as Total\_Wins\_Tie**

**from cte1;**

****

****

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

Solution:-

**Approach:**

If we didn’t have predefined questions, the best way to tackle this would be to step back and think—**what really impacts a team’s performance?** Instead of just crunching numbers, we’d dig into patterns, trends, and factors like venue conditions, home-ground advantage, and team consistency.

The first step would be to **organize the data properly**—ensuring match details, teams, venues, and results are all linked correctly. Then, we’d start exploring—how do teams perform in different stadiums? Do some venues favor batting more than bowling? Does playing at home really make a difference?

**Insights:**

Once we put everything together, some clear trends stand out:

* **Some stadiums are run-fests**—grounds like M Chinnaswamy and Brabourne consistently produce high scores, meaning teams need strong batting depth to dominate there.
* **RCB thrives at home—but there’s room for improvement.** At M Chinnaswamy, RCB has played **27 matches and won 16**, which is a decent **59.26% win rate**. Clearly, playing in front of their home crowd gives them an edge, but they aren’t untouchable.

**Recommendations:**

So, what can teams do with these insights?

1. **Play to the venue’s strengths.** If a stadium favors big scores, teams should focus on aggressive batting lineups and bowlers who can handle pressure in high-scoring games.
2. **RCB can push their home advantage even further.** Their batting is already strong, but their **death bowling has been a weak link**. Strengthening that area could turn more close games in their favor and boost their win percentage at home.

This is how I would processed, if I didn’t had any questions

1. 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".

Solution:- Here is the query

**UPDATE Matches SET Opponent\_Team = 'Delhi\_Daredevils' WHERE Opponent\_Team = 'Delhi\_Capitals';**