**Objective Questions**

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

ANS: Different data types are:

**Int**

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)**

ANS: **2558**

**Query:**

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

**Table:**



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

ANS: **89**

**Table:**

****

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?

**ANS: 9**

**Table:**

****

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

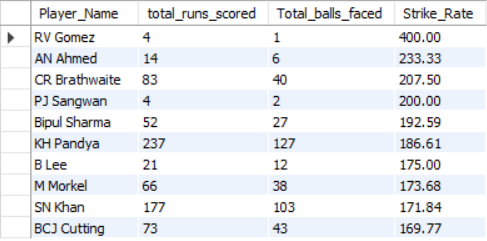
where Team\_Name='Royal Challengers Bangalore')

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**

ANS:

Table:



Query:

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

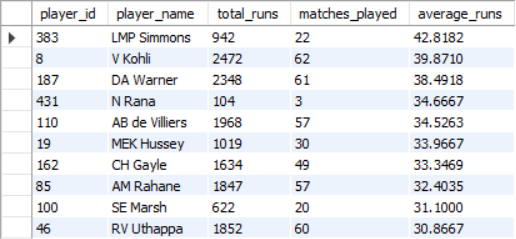
order by s.Strike\_Rate desc

limit 10

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

ANS:

**Table:**



**Query:**

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

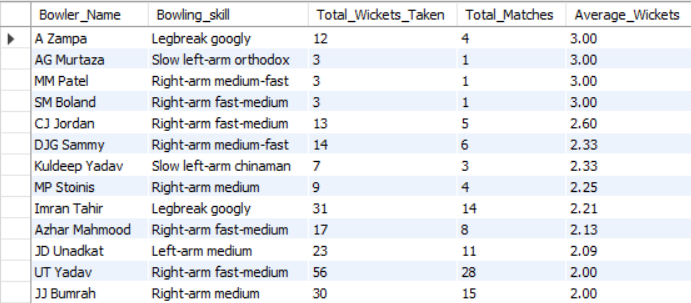
order by

average\_runs desc;

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

ANS:

Table:



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

)

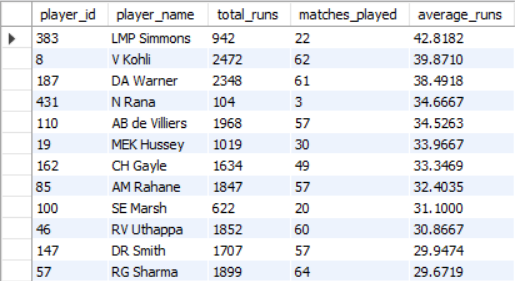
select Bowler\_Name,Bowling\_skill,Total\_Wickets\_Taken,Total\_Matches,round((Total\_Wickets\_Taken)/(Total\_Matches),2) as Average\_Wickets

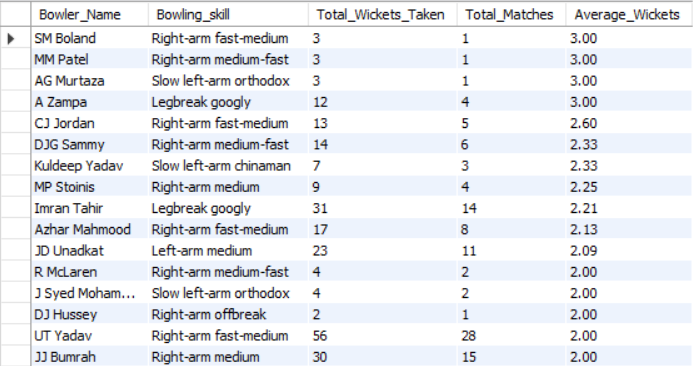
from cte2

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**

ANS:





Query:

**For 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

),

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 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

),

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

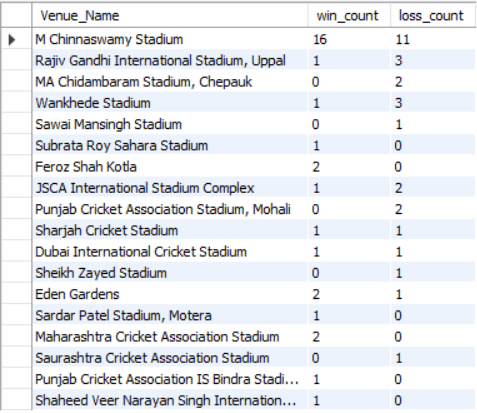
order by (Total\_Wickets\_Taken)/(Total\_Matches) desc

)

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.**

ANS:



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,

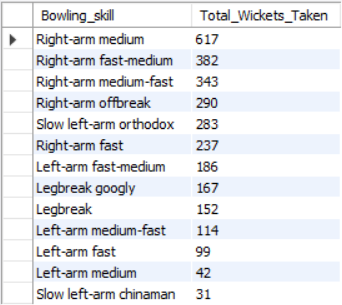
COUNT(CASE WHEN winner != 'Royal Challengers Bangalore' THEN 1 END) AS loss\_count

from cte1

group by Venue\_Name

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

ANS:



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

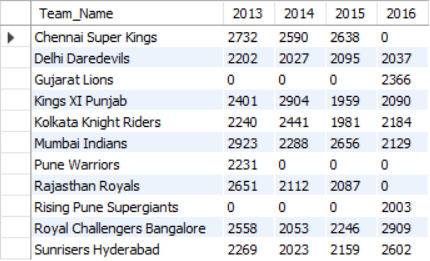
from cte

group by Bowling\_skill

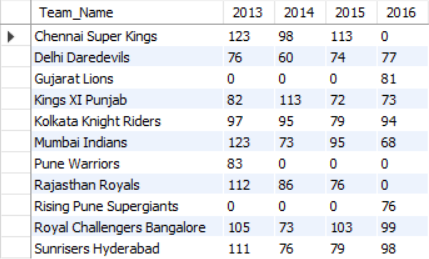
order by count(\*) desc

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

ANS: Total Runs Scored By Each Team Yearwise:



Total Wickets Taken By Each Team Yearwise:

+

Query for Yearwise total Runs Scored by each team :

**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;**

Query for 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"**

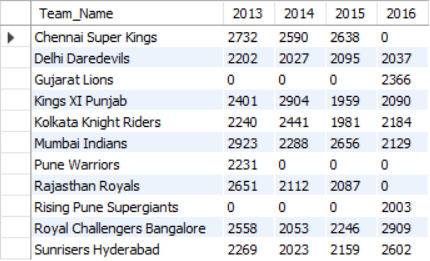
**from cte2**

**group by Team\_Name**

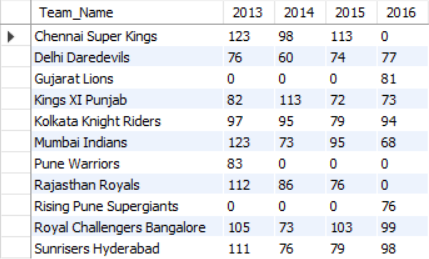
**order by Team\_Name**

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

ANS: Total Runs Scored:



Total Wickets Taken:



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.**

ANS:

Table:

These are the Top 10 Bowlers Name:



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"

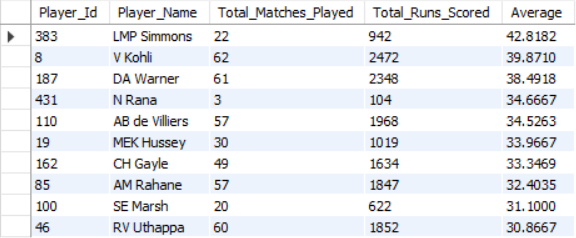
from cte3

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)**

ANS:

For Batsman:



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)**

**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**

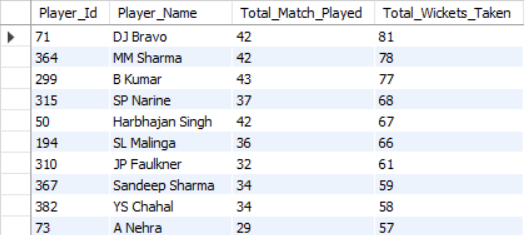
**from cte**

**group by Striker,Player\_Name**

**order by Average desc**

**limit 10;**

**For Bowler:**



Query:

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;

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

ANS: For Batsman: 

Visualization:

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:



Visualization:

Query:

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

**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?

ANS:

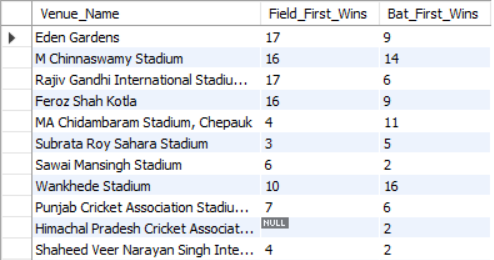
**Approach:**

To determine **how the toss decision affects the result of the match** and whether this impact is venue-specific, we will analyze the data focusing on:

1. **Field-First Wins**: Number of matches won when teams opted to field first.
2. **Bat-First Wins**: Number of matches won when teams opted to bat first.
3. Compare these statistics across all venues to identify trends and anomalies.

**Visualization:**

**These are some venues :**



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

**Insights:**

1. **Overall Trends**:
   * At certain venues, teams clearly prefer one decision over the other:
     + **Eden Gardens** and **Rajiv Gandhi International Stadium**: Field-first teams performed much better (17 wins vs. 9 and 17 vs. 6, respectively).
     + **MA Chidambaram Stadium** and **Wankhede Stadium**: Bat-first wins dominate significantly (11 vs. 4 at MA Chidambaram and 16 vs. 10 at Wankhede).
2. **Null or Missing Data**:
   * **Himachal Pradesh Cricket Association Stadium** and **Green Park** have incomplete data, indicating that these venues either have fewer matches or incomplete reporting. This should be factored into any conclusions.
3. **Balanced Venues**:
   * **M Chinnaswamy Stadium**: Near balance between field-first (16) and bat-first (14) wins.
   * **Sardar Patel Stadium**: Equal wins for both decisions (4 vs. 4).
4. **Venue-Specific Observations**:
   * Venues like **Sawai Mansingh Stadium** (6 field-first vs. 2 bat-first) show that field-first decisions yield better results.
   * Conversely, **Dr. Y.S. Rajasekhara Reddy Stadium** (5 bat-first vs. 4 field-first) shows a slight preference for bat-first.

**Suggestions:**

1. **Venue-Specific Strategies**:
   * At venues like **Eden Gardens** and **Rajiv Gandhi Stadium**, captains may favor fielding first based on historical results.
   * For venues like **MA Chidambaram Stadium**, teams may benefit more from batting first.
2. **Data Gaps**:
   * Investigate venues with NULL data (e.g., **Himachal Pradesh Stadium** and **Green Park**) to confirm why match outcomes are missing.
3. **General Patterns**:
   * Teams should assess pitch conditions and past trends at specific venues before making toss decisions.
   * For balanced venues, real-time factors like weather and team strengths may play a bigger role.
4. Suggest some of the players who would be best fit for the team.

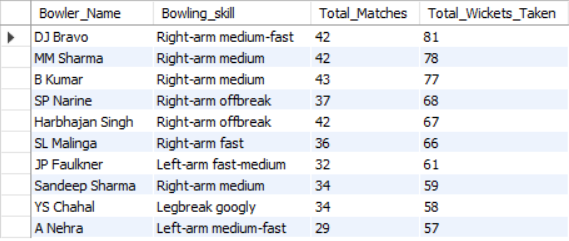
Ans:

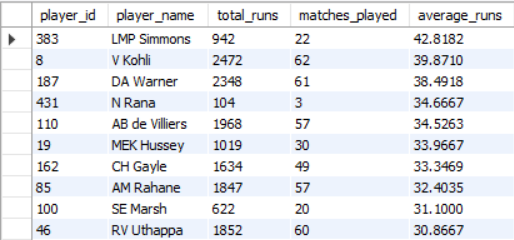
**Approach:**

To answer this question effectively, we need to determine the best-fit players based on the following key criteria:

1. Player Performance Metrics: Runs, wickets
2. Role-specific Selection:
   * Batsmen: Runs
   * Bowlers: Wickets

**Table:**





**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

**Insights:**

1. Batsmen Analysis

* High Performers by Average Runs:
  + LMP Simmons: Highest average runs (42.82), though he has played fewer matches (22). This makes him a reliable short-term performer.
  + V Kohli: Second-highest average (39.87) and highest total runs (2472) across 62 matches. His consistency makes him a valuable asset.
  + DA Warner: Averages 38.49 with 2348 runs in 61 matches, showing consistency and experience.
* Steady Contributors:
  + AB de Villiers: Averaging 34.52 with 1968 runs, combining experience and reliability.
  + CH Gayle: Known for his aggressive style with 1634 runs and 33.34 average.
* Insights:
  + Prioritize V Kohli and DA Warner for experienced and consistent run-making.
  + LMP Simmons could be a strong pick for explosive starts despite fewer matches.

2. Bowlers Analysis

* Wicket-Taking Leaders:
  + DJ Bravo: Highest wicket-taker (81 wickets in 42 matches). An all-rounder with strong wicket-taking ability.
  + MM Sharma (78 wickets) and B Kumar (77 wickets): Both excel in medium-pace bowling.
* Strike Bowlers:
  + SL Malinga: 66 wickets in just 36 matches, showcasing exceptional strike rates for limited games.
  + SP Narine and Harbhajan Singh: Both with right-arm offbreak skills and over 65 wickets.
* Insights:
  + Focus on DJ Bravo and SL Malinga for maximum wicket-taking abilities.
  + Include SP Narine for spin variety and middle-overs control.

**Suggestions:**

 Top Batsmen to Consider:

* V Kohli (high runs + consistency)
* DA Warner (proven track record)
* LMP Simmons (best average runs)

 Top Bowlers to Include:

* DJ Bravo (highest wickets)
* SL Malinga (strike bowler with fewer matches)
* SP Narine or Harbhajan Singh for spin options.

 Balanced Approach:

* Use top-order batsmen like V Kohli and explosive players like CH Gayle.
* Combine pace and spin options with bowlers like DJ Bravo and SP Narine for variety.

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

ANS:

**Approach**

To select the best players for a team, focus on **key performance parameters** derived from the provided data, which include:

1. **Batsmen**: Runs scored, matches played, and batting averages.
2. **Bowlers**: Wickets taken, matches played, and bowling skills.

By analyzing these metrics, we can evaluate player consistency, experience, and their ability to influence match outcomes.

**Insights**

**For Batsmen:**

* **Consistency**:
  + Measured using **batting average** (total runs/matches played).
  + Players like **LMP Simmons** (42.82) and **V Kohli** (39.87) stand out as consistent performers.
* **Total Runs**:
  + Total runs highlight players’ contributions over time.
  + **V Kohli** (2472 runs) and **DA Warner** (2348 runs) are reliable run-scorers over multiple matches.
* **Match Experience**:
  + More matches played indicate adaptability and reliability under pressure.
  + Players like **V Kohli** (62 matches) and **RV Uthappa** (60 matches) have significant experience.

**For Bowlers:**

* **Wicket-Taking Ability**:
  + Measured using **total wickets** taken in matches.
  + Players like **DJ Bravo** (81 wickets) and **MM Sharma** (78 wickets) excel in this area.
* **Strike Rates and Efficiency**:
  + Wickets in fewer matches show effective and match-winning bowling.
  + **SL Malinga** (66 wickets in 36 matches) is highly efficient.
* **Bowling Variety**:
  + Different **bowling skills** add versatility:
    - Fast bowlers: **SL Malinga**, **JP Faulkner**.
    - Spinners: **SP Narine**, **Harbhajan Singh**, **YS Chahal**.
    - Medium pacers: **B Kumar**, **MM Sharma**.

**Parameters to Focus On While Selecting Players**

1. **For Batsmen**:
   * **Batting Average**: Indicates consistency and reliability.
   * **Total Runs**: Helps assess overall contribution to the team.
   * **Match Experience**: Experienced players are dependable under pressure.
   * **Recent Performance**: Focus on the most recent stats to ensure form and fitness.
2. **For Bowlers**:
   * **Total Wickets**: Key indicator of their wicket-taking ability.
   * **Strike Rate/Efficiency**: Wickets taken in fewer matches highlight performance.
   * **Bowling Skills**: Include a mix of **pace** and **spin** for versatility.
   * **Match Conditions**: Consider bowlers suitable for venue conditions (e.g., spinners on turning pitches).
3. **Overall Team Balance**:
   * Maintain a **balance between batting and bowling strengths**.
   * Include versatile players like **all-rounders** (e.g., DJ Bravo) who contribute in multiple areas.
   * Ensure a mix of experienced players and young performers for long-term success.

**Suggestions**

1. **Top Batsmen**: Prioritize players with high averages and total runs like **V Kohli**, **DA Warner**, and **LMP Simmons**.
2. **Top Bowlers**: Select bowlers with strong wicket-taking ability like **DJ Bravo**, **SL Malinga**, and versatile spinners like **SP Narine**.
3. **All-Rounders**: Players like **JP Faulkner** and **DJ Bravo** add depth to both batting and bowling.
4. **Venue and Match Context**: Tailor the selection based on match conditions (e.g., spin-friendly pitches or pace-friendly tracks).
5. Which players offer versatility in their skills and can contribute effectively with both bat and ball? (can you visualize the data for the same)

ANS:

Approach:

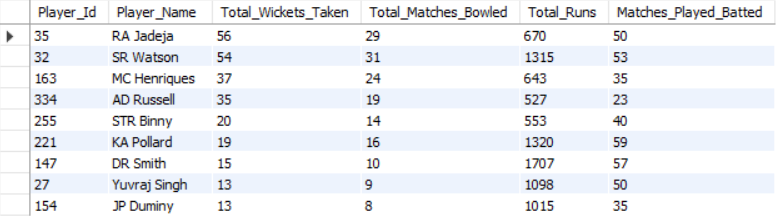
1. Set Thresholds for Significant Contribution:

* Define minimum thresholds for total runs and total wickets to classify a player as versatile. For example:
  + Total Runs > 500
  + Total Wickets > 20

2. Identify Versatile Players:

* Filter players who meet or exceed these thresholds in both metrics.

**Visualization:**



**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;

**Insights:**

1. **Players with High Versatility:**
   * Based on the metrics, identify standout players. For example:
     + **SR Watson**:
       - High run contribution (1315 runs in 53 matches = ~24.81 runs/match).
       - Strong bowling performance (54 wickets in 31 matches = ~1.74 wickets/match).
       - All-rounder score = 1.74+24.81=26.551.74 + 24.81 = 26.551.74+24.81=26.55.
     + **RA Jadeja**:
       - Consistent in both bowling (56 wickets in 29 matches = ~1.93 wickets/match) and batting (670 runs in 50 matches = ~13.4 runs/match).
       - All-rounder score = 1.93+13.4=15.331.93 + 13.4 = 15.331.93+13.4=15.33.
2. **Specialized Performers:**
   * **DR Smith:**
     + Outstanding in batting (1707 runs in 57 matches = ~29.93 runs/match) but less contribution in bowling (15 wickets in 10 matches = ~1.5 wickets/match).
     + Primary strength lies in batting.
   * **KA Pollard:**
     + Exceptional batting contribution (1320 runs in 59 matches = ~22.37 runs/match).
     + Limited bowling contribution (19 wickets in 16 matches = ~1.19 wickets/match).
3. **Underperformers in Versatility:**
   * Players like JP Duminy and Yuvraj Singh have decent batting contributions but are relatively weaker in bowling.

**Suggestions:**

1. **Role Identification:**
   * **SR Watson and RA Jadeja**: Ideal choices as all-rounders due to their balanced performance in both skills.
   * **DR Smith and KA Pollard**: Can be utilized as batting specialists, with occasional bowling when required.
2. **Player Utilization:**
   * For versatile players like **AD Russell**, ensure they get enough opportunities to contribute in both roles. His moderate contributions (527 runs, 35 wickets) indicate potential for further optimization.
3. **Training Focus:**
   * Players like **Yuvraj Singh** and **JP Duminy** can work on their bowling skills to improve their all-rounder impact.
4. **Team Strategy:**
   * Build a core team around versatile players like Watson and Jadeja, complemented by batting and bowling specialists to strengthen specific areas.
5. Are there players whose presence positively influences the morale and performance of the team? (justify your answer using visualization)

ANS:

**Approach:**

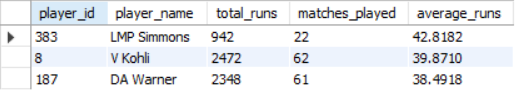
To identify players whose presence positively influences team morale and performance, we'll evaluate the top 3 batters, bowlers, and all-rounders using the following methodology:

Categorize Players into Roles:

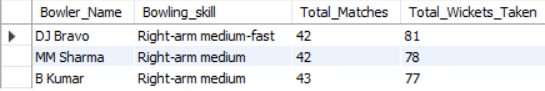
* Batters: Based on total runs scored and average runs.
* Bowlers: Based on total wickets taken.
* All-rounders: Based on their contributions in both batting (total runs) and bowling (total wickets).

**Visualization:**

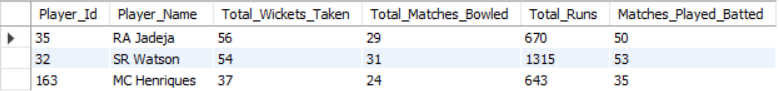
**Top 3 Batsman**



**Top 3 Bowlers**



**Top 3 All rounders**



**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

**Insights:**

1. Top 3 Batters:

* LMP Simmons: Highest average runs per match (42.81) despite playing fewer matches. A critical player for consistent batting performance.
* V Kohli: Exceptional overall contribution with the highest total runs (2472) in 62 matches, a reliable run-scorer.
* DA Warner: Second-highest total runs (2348) and consistency in 61 matches.

2. Top 3 Bowlers:

* DJ Bravo: Most wickets (81 in 42 matches) with right-arm medium-fast bowling. A consistent wicket-taker.
* MM Sharma: 78 wickets in 42 matches, another dependable bowler with right-arm medium pace.
* B Kumar: Slightly fewer wickets (77) in 43 matches, but still a valuable bowling asset.

3. Top 3 All-Rounders:

* SR Watson: Combined score = 1315+(10×54)=18551315 + (10 \times 54) = 18551315+(10×54)=1855. Exceptional with both bat and ball.
* RA Jadeja: Combined score = 670+(10×56)=1230670 + (10 \times 56) = 1230670+(10×56)=1230. Strong contributions in bowling and decent batting.
* MC Henriques: Combined score = 643+(10×37)=1013643 + (10 \times 37) = 1013643+(10×37)=1013. A balanced all-rounder but slightly below the top two.

**Suggestions:**

1. Boosting Morale through Key Players:
   * Batters:
     + Leverage V Kohli and DA Warner as core batters for chasing or building large totals.
     + Encourage LMP Simmons to play in high-pressure situations due to his consistency.
   * Bowlers:
     + Utilize DJ Bravo and MM Sharma as frontline bowlers in powerplay and death overs for maximum impact.
     + Employ B Kumar in middle overs to maintain pressure.
   * All-Rounders:
     + Assign leadership roles or pivotal moments to SR Watson and RA Jadeja for their ability to influence both batting and bowling outcomes.
2. Team Strategy:
   * Build the team's spine around versatile players like Watson and Jadeja, while ensuring strong support from specialists like Kohli, Simmons, Bravo, and Sharma.
3. What would you suggest to RCB before going to the mega auction?

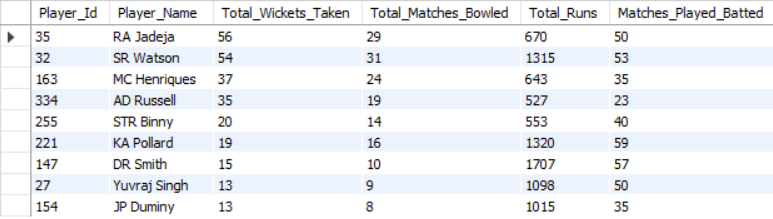
ANS:

**Approach:**

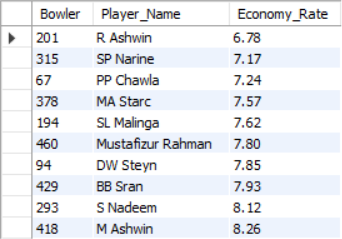
* **Strengths:** Analyze RCB's current squad to identify strong areas (e.g., Virat Kohli, Faf du Plessis for consistency, and proven finishers).
* **Weaknesses:** Historically, RCB has struggled with death bowling and finishing games. Address gaps like reliable death bowlers, consistent power hitters, and quality all-rounders.

**Visualization:**

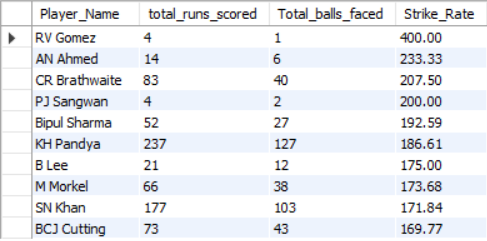
**Top 10 All Rounders:**



**Top 10 Death Bowlers:**



**Top 10 Power Hitters:**



**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

**Insights:**

1. Insights on All-Rounders:

* Versatile Players: RA Jadeja and SR Watson have the highest contributions with both bat and ball. Jadeja's ability to take 56 wickets in 29 matches and score 670 runs across 50 games makes him an ideal all-rounder.
* Explosive Finishers: AD Russell (527 runs in 19 matches) and KA Pollard (1320 runs in 59 matches) are power hitters capable of turning games around. Both are also handy bowlers with 35 and 19 wickets, respectively.
* Underrated Options: JP Duminy has a strike rate close to 125 (estimated from data) and can provide balance in the middle order with part-time spin.

Recommendation: Target all-rounders like Jadeja, Russell, or Watson who bring balance to both departments. For a budget-friendly option, STR Binny could be considered.

2. Insights on Power Hitters:

* High Strike Rates: Players like CR Brathwaite (207.5), AN Ahmed (233.33), and KH Pandya (186.61) have exceptional strike rates, making them valuable in the middle or death overs.
* Consistent Contributions: KH Pandya scored 237 runs from 127 balls with a strike rate of 186.61, making him a reliable finisher.
* Explosive but Situational: RV Gomez and PJ Sangwan have high strike rates but very few runs, indicating limited match exposure or opportunities.

Recommendation: Invest in players like KH Pandya or CR Brathwaite for finishing roles. SN Khan (strike rate: 171.84) is another excellent middle-order option.

3. Insights on Death Bowlers:

* Best Economy Rates: R Ashwin (6.78) and SP Narine (7.17) are the most economical bowlers, ideal for controlling runs during the powerplay and middle overs.
* Specialized Death Bowlers: MA Starc (7.57) and SL Malinga (7.62) have proven their ability to bowl at the death.
* Budget Options: S Nadeem (8.12) and M Ashwin (8.26) can be effective middle-overs spinners but are slightly expensive in terms of economy rate.

Recommendation: Focus on bowlers like Starc or Malinga for death overs. Narine can be a dual-purpose spinner and a pinch hitter if required.

**Suggestions:**

Key Targets:

* All-Rounders: AD Russell, KH Pandya, or Jadeja (if available) for match-winning capabilities.
* Power Hitters: CR Brathwaite, SN Khan, or KH Pandya to finish games with a high strike rate.
* Death Bowlers: MA Starc or SL Malinga for their expertise in closing out innings.

Balanced Squad:

* Invest in a mix of experienced and young players. Scout uncapped Indian talents for depth.
* Ensure a strong spin department with players like Ashwin or Narine for middle-over control.

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

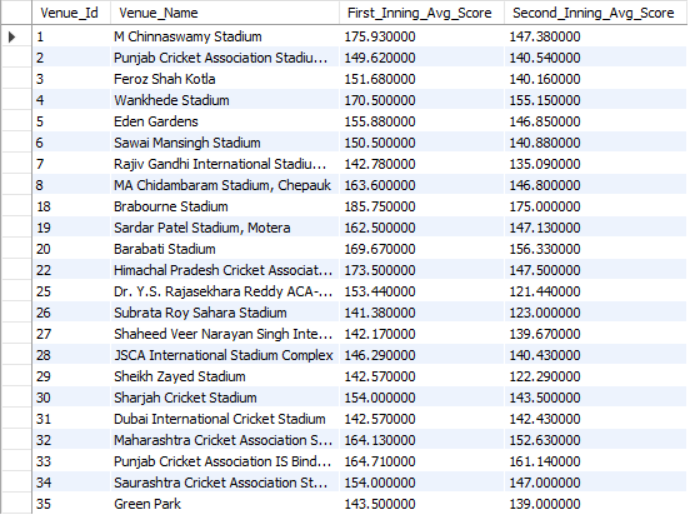
ANS:

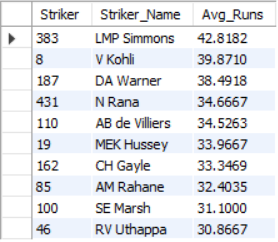
**Approach:**

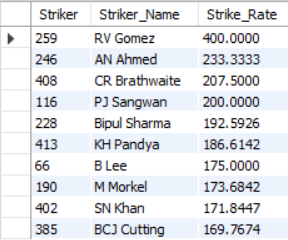
To address the factors contributing to high-scoring matches, analyze the following key areas:

1. **Venue Impact:**
   * Compare the average scores in the first and second innings across various venues.
   * Identify venues where scoring is consistently higher (e.g., M. Chinnaswamy Stadium, Brabourne Stadium).
2. **Player Performance:**
   * Evaluate strike rates and average runs of top-performing batsmen to identify their impact on high-scoring games.
   * Focus on power hitters with high strike rates (e.g., CR Brathwaite, KH Pandya).
3. **Bowling Economy:**
   * Assess the economy rates of bowlers at specific venues to understand whether bowlers are struggling.
4. **Team Strategy:**
   * Analyze whether teams are adapting strategies for high-scoring venues, such as batting first or stacking their lineups with power hitters.
5. **Impact on Viewership:**
   * Consider the correlation between high-scoring matches and increased audience interest, which can influence team strategies and player selections.

**Visualization:**







**Insights:**

1. High-Scoring Venues:
   * M. Chinnaswamy Stadium (175.93 first innings, 147.38 second innings): A batting paradise with short boundaries and flat pitches.
   * Brabourne Stadium (185.75 first innings, 175 second innings): Extremely high scores indicate limited bowler assistance.
   * Himachal Pradesh Cricket Association Stadium (173.5 first innings): Another high-scoring venue with favorable conditions for batsmen.
2. Key Players for High Scores:
   * Top Run Scorers:
     + Virat Kohli (2472 runs, 39.87 avg) and DA Warner (2348 runs, 38.49 avg) are consistent performers.
     + LMP Simmons (42.82 avg) stands out with his exceptional average.
   * Power Hitters:
     + CR Brathwaite (207.5 SR), KH Pandya (186.61 SR), and SN Khan (171.84 SR) can accelerate the scoring rate.
3. Challenging Venues for Bowlers:
   * Brabourne and M. Chinnaswamy are challenging due to flat pitches.
   * Second innings scoring is consistently high in Brabourne, indicating dew or favorable batting conditions.
4. Strategic Observations:
   * Teams batting first in high-scoring venues are setting massive targets.
   * Death bowlers with high economy rates at these venues need to be avoided or rotated.

**Suggestions for Team Strategies:**

1. **Player Selection:**
   * Focus on **power hitters** like KH Pandya and CR Brathwaite for venues like Brabourne and M. Chinnaswamy.
   * Include consistent run-scorers like Kohli and Warner to anchor the innings.
2. **Venue-Specific Strategy:**
   * In high-scoring venues:
     + Opt to **bat first** to leverage favorable conditions and post massive totals.
     + Prioritize spinners and all-rounders who can restrict runs in the middle overs.
   * In low-scoring venues:
     + Use disciplined seamers and restrict the opposition in the powerplay.
3. **Bowling Lineup:**
   * Invest in death bowlers like Starc or Malinga who can handle pressure.
   * Rotate bowlers strategically at flat-pitch venues.
4. **Impact on Viewership:**
   * High-scoring games create excitement and attract viewership. Teams should prioritize entertaining cricket by stacking their lineups with attacking batsmen.
5. Analyze the impact of home-ground advantage on team performance and identify strategies to maximize this advantage for RCB.

ANS:

**Approach:**

To analyze the impact of home-ground advantage for Royal Challengers Bangalore (RCB) and identify strategies to maximize it:

1. **Home Performance Analysis:**
   * Calculate RCB's win percentage at M. Chinnaswamy Stadium and compare it to their performance at other venues.
   * Evaluate player and team-specific strengths at Chinnaswamy.
2. **Pitch and Conditions:**
   * Understand the pitch behavior at Chinnaswamy, known for being a high-scoring venue with short boundaries.
   * Assess how weather (e.g., dew factor in second innings) influences the match outcomes.
3. **Player Impact:**
   * Identify key performers (batsmen, bowlers) who thrive at Chinnaswamy.
   * Assess death bowling efficiency and powerplay batting at the venue.
4. **Opponent Analysis:**
   * Study the performance of opponents at Chinnaswamy to develop counter-strategies.
5. **Fan Support and Viewership:**
   * Consider the psychological boost provided by the home crowd.
   * Explore strategies to leverage crowd support.

**Visualization:** 

**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;

**Insights:**

1. **Home-Ground Win Percentage:**
   * RCB has played 27 matches at Chinnaswamy, winning 16 of them.
   * **Win Percentage = (16/27) × 100 ≈ 59.26%.**
   * This indicates a significant home-ground advantage.
2. **High-Scoring Matches:**
   * Chinnaswamy's average first-inning score of 175.93 reflects its reputation as a batting-friendly venue.
   * Teams need to aim for aggressive batting strategies, especially in the first innings.
3. **Batsmen Impact:**
   * Power hitters like AB de Villiers, Chris Gayle, and Virat Kohli have historically dominated at this venue.
   * Identifying and retaining batsmen who can capitalize on short boundaries is critical.
4. **Bowling Weaknesses:**
   * The venue's conditions often challenge bowlers, especially in the death overs.
   * Death-over specialists like Mitchell Starc or yorker specialists are essential to restrict opposition scoring.
5. **Opposition Challenge:**
   * Opponents aware of the batting-friendly pitch tend to strengthen their batting lineups at Chinnaswamy.
   * RCB must focus on outscoring them and planning bowling rotations effectively.

**Suggestions to Maximize Home Advantage:**

1. **Strengthen Batting:**
   * Build a core lineup with power hitters capable of scoring 180+ consistently.
   * Leverage experienced players like Kohli and include explosive batsmen to dominate powerplays.
2. **Specialized Bowling Strategies:**
   * Focus on bowlers with tight control in death overs.
   * Include spinners who can slow down scoring in middle overs.
3. **Win Toss and Bat First:**
   * Teams batting first often succeed at Chinnaswamy due to scoreboard pressure in high-scoring games.
   * Use the dew factor to RCB's advantage by forcing opponents to bowl second.
4. **Retain Key Players:**
   * Retain players who have proven success at Chinnaswamy, like AB de Villiers, Gayle (if available), or similar replacements.
5. **Leverage Fan Support:**
   * Create a vibrant stadium atmosphere to boost team morale and intimidate opposition.
   * Engage fans pre-match to ensure maximum support.
6. **Pitch Preparation:**
   * Customize pitches to slightly favor spin or seam (depending on RCB's bowling strength) while maintaining batting-friendly conditions.
7. 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.

ANS:

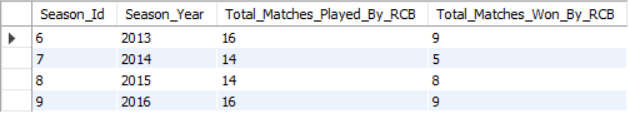
**Approach:**

To analyze RCB's performance across seasons and identify potential reasons for not winning a trophy:

1. **Performance Across Seasons:**
   * Analyze RCB's win percentage over the years to identify patterns.
   * Compare the number of wins against matches played for each season.
2. **Venue and Location Impact:**
   * Study RCB's home and away performances.
   * Focus on specific venues where RCB has struggled.
3. **Batting and Bowling Metrics:**
   * Compare total runs scored and wickets taken by RCB across seasons.
   * Analyze strike rates and bowling efficiency.
4. **Match Context:**
   * Evaluate RCB's ability to defend and chase totals.
   * Study performances in high-pressure games (finals, playoffs).

**Visualization:**

Year wise Total Matches Won By RCB



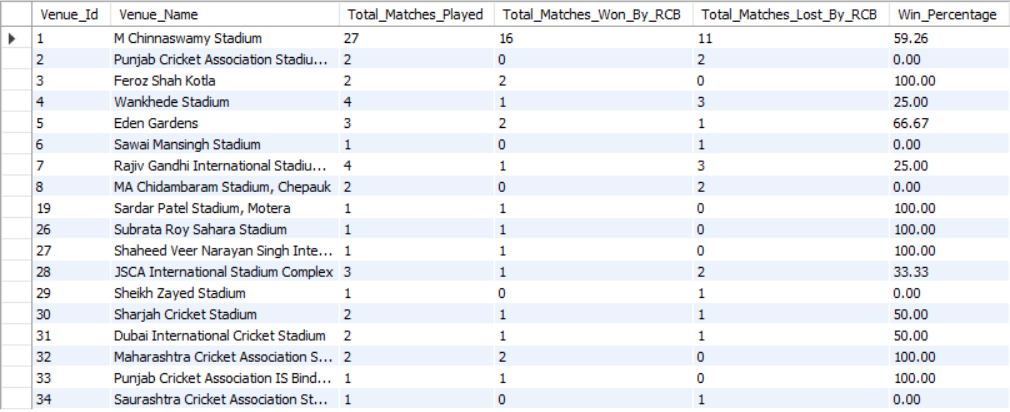
Year wise Total Runs Scored By RCB



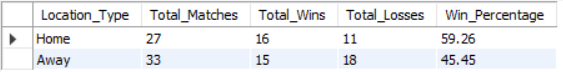
Year wise Total Wickets Taken By RCB



Venue wise Performance By RCB



home/away performance By RCB



chasing/defending performance By RCB



**Query:**

-- Past years performance by 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

)

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

**Insights:**

**1. Seasonal Performance Trends:**

* **2013:** 56.25% win rate (9/16 matches) – Decent performance but inconsistent batting/bowling.
* **2014:** 35.71% win rate (5/14 matches) – A poor season, likely due to bowling inefficiencies.
* **2015:** 57.14% win rate (8/14 matches) – Improved but fell short in playoffs.
* **2016:** 56.25% win rate (9/16 matches) – Strong season with a final appearance, driven by Kohli's outstanding form.

**2. Home and Away Performance:**

* Home win percentage: **59.26%.**
* Away win percentage: **45.45%.**
* RCB performs significantly better at home but struggles to replicate success in away matches.

**3. Venue-Specific Insights:**

* Strong performance at M. Chinnaswamy (59.26%) and venues like Feroz Shah Kotla, Motera, and MCA Stadium (100%).
* Struggles at Mohali, Chepauk, and Rajiv Gandhi Stadium (poor win rates).

**4. Batting Strengths and Bowling Weaknesses:**

* Total runs scored increased significantly in 2016 (2909), led by Virat Kohli's exceptional form.
* Strike rates are consistently high, indicating batting dominance.
* Bowling inefficiencies, particularly in death overs, often cost crucial matches.

**5. Match Context:**

* Wins while chasing: 16.
* Wins while defending: 14.
* Slightly better at chasing, possibly due to strong finishers like AB de Villiers.

**Suggestions to Improve RCB's Performance:**

1. **Strengthen Bowling Unit:**
   * Focus on acquiring experienced death-over specialists to handle high-pressure situations.
   * Retain spinners who can exploit slow tracks at away venues.
2. **Address Away Performance:**
   * Develop strategies tailored for challenging away venues like Chepauk and Mohali.
   * Conduct venue-specific training camps to prepare players for varied conditions.
3. **Optimize Team Composition:**
   * Leverage power hitters for home matches at Chinnaswamy to capitalize on batting-friendly conditions.
   * Include all-rounders for away matches to balance both departments.
4. **Defending Strategies:**
   * Improve the ability to defend totals by setting aggressive fields and rotating bowlers effectively.
   * Use data analytics to predict opposition's chasing patterns and counter them.
5. **Leadership and Motivation:**
   * Ensure clear and proactive captaincy during high-stakes games.
   * Engage in mental conditioning programs to enhance performance under pressure.
6. How would you approach this problem, if the objective and subjective questions weren't given?

ANS:

**Approach:**

1. **Understand the Context:**
   * The task involves analyzing RCB's performance in IPL across various parameters like venues, seasons, players, and match situations.
2. **Define Objectives:**
   * Without predefined questions, focus on:
     + Analyzing team performance trends.
     + Evaluating player contributions.
     + Identifying venue-based and situation-based factors affecting performance.
     + Proposing actionable strategies for improvement.

**Insights:**

1. **Team Performance Trends:**
   * RCB has shown variability across seasons, with strong performances (e.g., 2016) but lack of consistency.
   * Win percentage is significantly higher at home (59.26%) compared to away (45.45%), indicating a strong home-ground advantage.
2. **Player Contributions:**
   * Key players like Virat Kohli and AB de Villiers have been consistent performers, contributing heavily to the team’s success.
   * Lower contributions from bowlers in critical moments might explain the inability to defend totals effectively.
3. **Venue Impact:**
   * RCB performs exceptionally at Chinnaswamy Stadium (16 wins out of 27 matches) but struggles in venues like Punjab Cricket Association Stadium and MA Chidambaram Stadium.
4. **Match Situations:**
   * The team is slightly better at chasing (16 wins) compared to defending totals (14 wins), indicating the need for better strategies while setting targets.

**Suggestions:**

1. **Enhance Bowling Strength:**
   * Invest in death bowlers to improve performance while defending totals.
   * Focus on acquiring spinners for venues like Chepauk, where spin plays a crucial role.
2. **Leverage Home-Ground Advantage:**
   * Strategize to maximize runs at Chinnaswamy Stadium, given its high-scoring nature.
   * Train bowlers to counteract high-scoring trends at home.
3. **Improve Away Performance:**
   * Develop specific strategies for challenging venues like Mohali and Uppal by understanding their pitch behavior and playing conditions.
4. **Optimize Player Utilization:**
   * Rotate and rest key players to maintain their performance throughout the season.
   * Focus on middle-order consistency to avoid collapses in high-pressure matches.
5. **Analyze Toss Decisions:**
   * Strengthen chasing strategies, as the team has shown better success chasing targets.
   * For matches at spin-friendly venues, consider batting first to allow spinners to control the second innings.
6. 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".

ANS:

**Query:**

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