Objective Questions

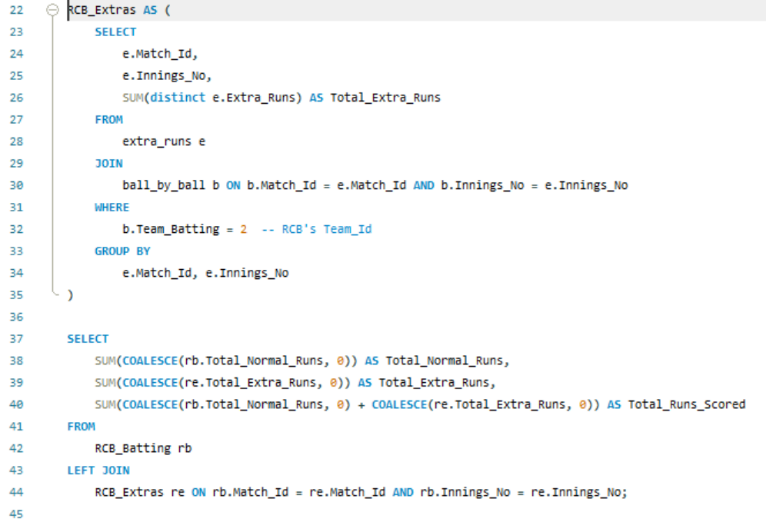
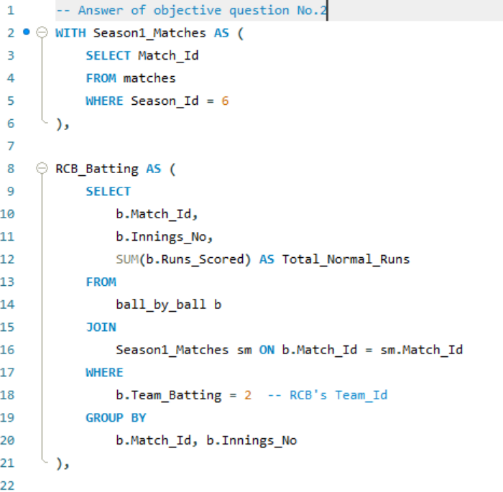
1. List the different dtypes of columns in table “ball\_by\_ball” (using information schema)

Ans: There is only 1 dtype, INTEGAR (INT)

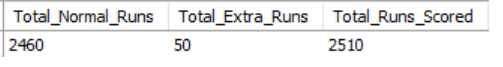
1. What is the total number of runs scored in 1st season by RCB (bonus: also include the extra runs using the extra runs table)  
   No season 1 data in matches table.

Ans: According to data 6th season is the 1st season. So, after assuming this I got total runs 2510 (extra-50+genuine runs 2460).

**Code:**



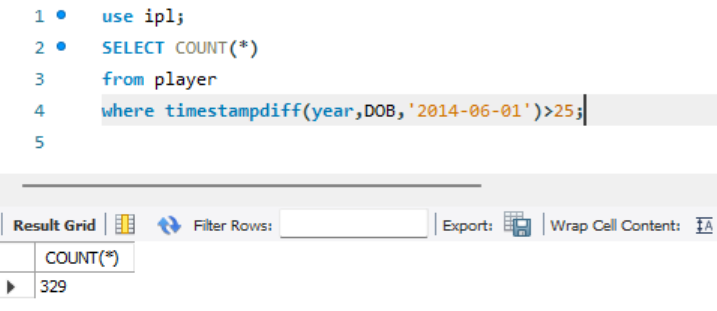
**Output:**



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

Ans: There were 329 players more than the age of 25 during season 2014.

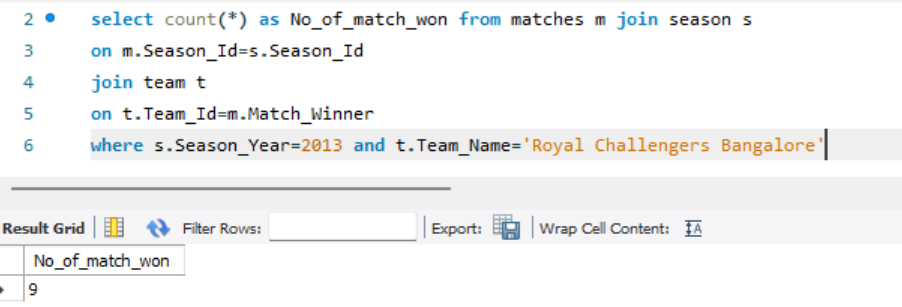
**Code & Output:**



1. How many matches did RCB win in 2013?

Ans: RCB won 9 matches in 2013.

**Code & Output:**



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

Ans: The top 10 players by strike rate are

1. RV Gomez

2. AN Ahmed

3. PJ Sangwan

4. DJ Hooda

5. M Morkel

6. B Lee

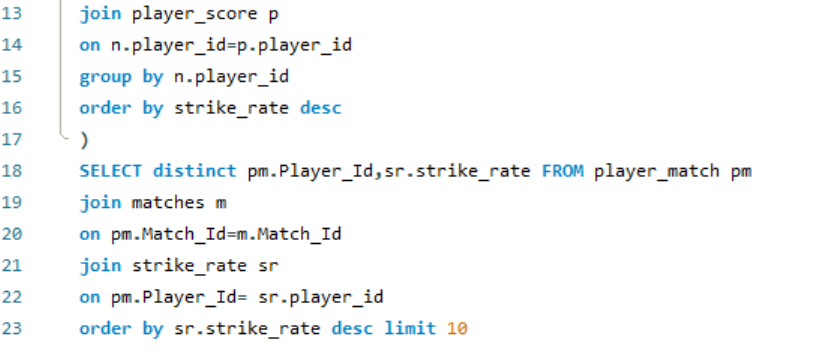
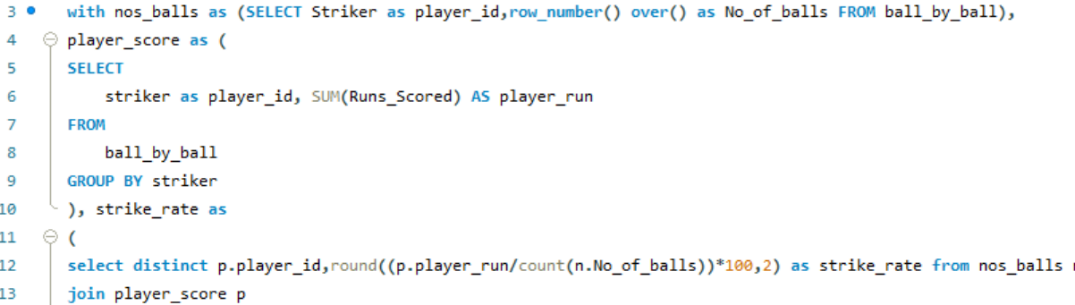
7. GJ Maxwell

8. HV Patel

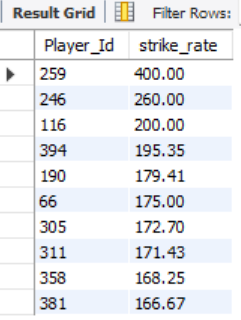
9. LJ Wright

10. Shivam Sharma

**Code:**



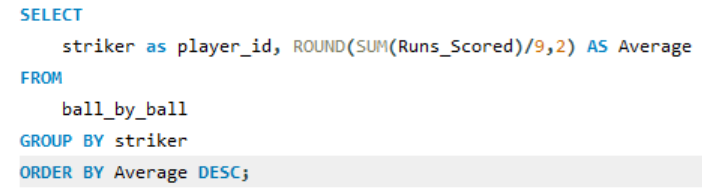
**Output:**



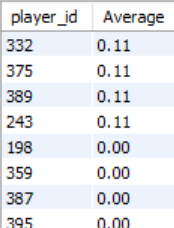
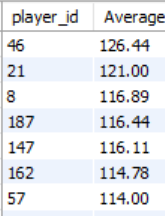
1. What are the average runs scored by each batsman considering all the seasons?

Ans: The average runs of each batsman considering all the seasons

**Code:**



**Output:**



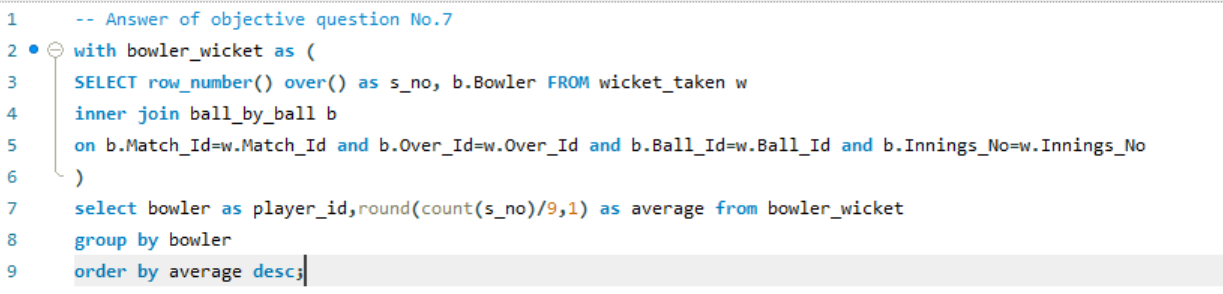
1. What are the average wickets taken by each bowler considering all the seasons?

Ans:

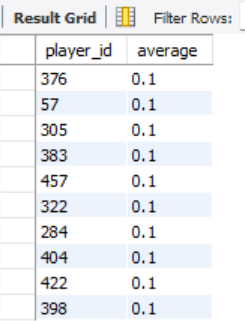
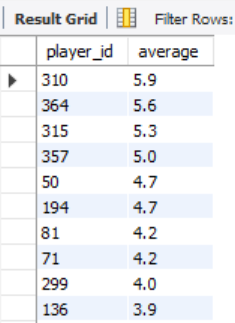
The average wicket taken by each bowler considering by all

Seasons are shown

**Code:**



**Output:**



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.

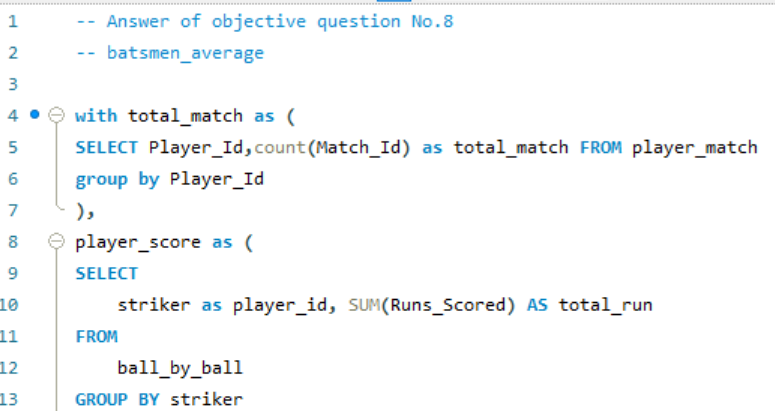
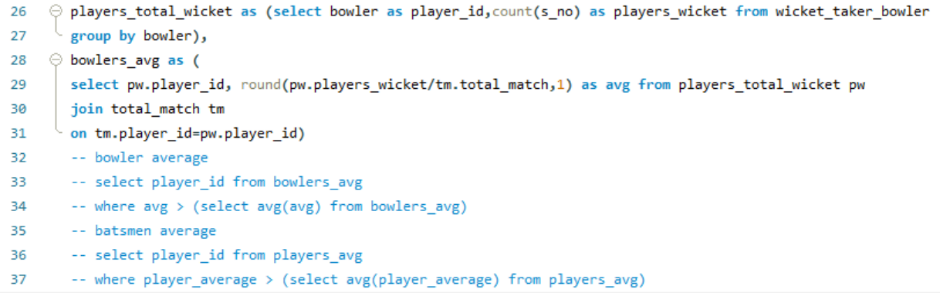
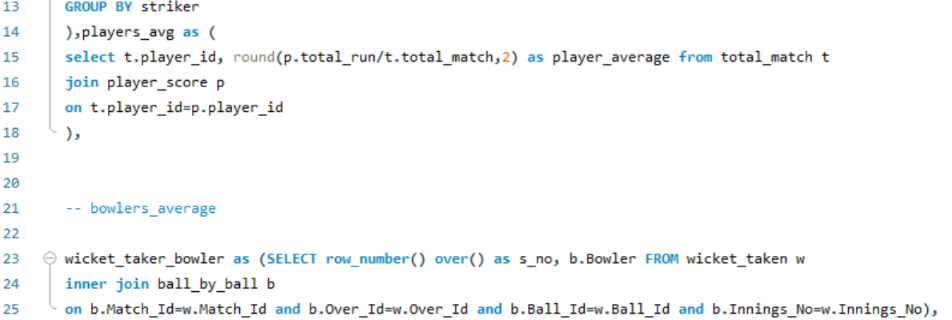
Ans:

The overall average run is 7.42 and overall average wicket is 0.59

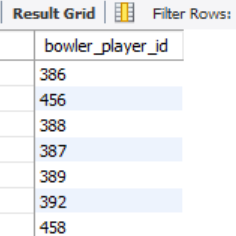
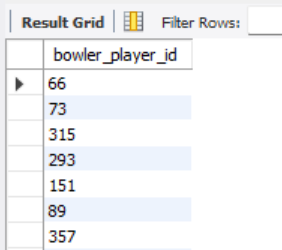
Players having average runs greater than overall average and

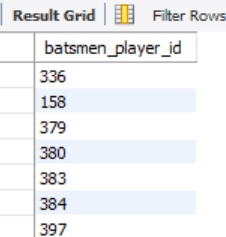
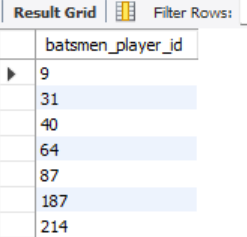
Players having average wicket greater than overall average.

**Code:**

**Output:**



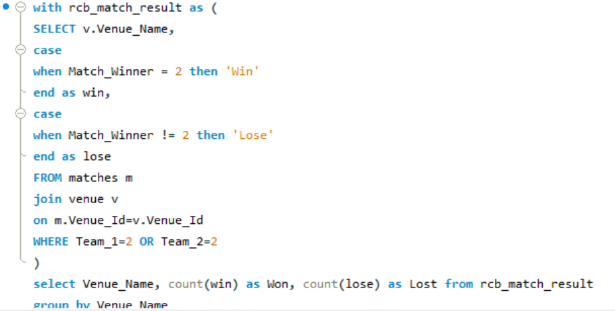


1. Create a table rcb\_record table that shows the wins and losses of RCB in an individual venue.

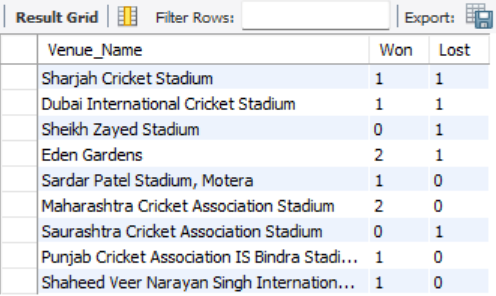
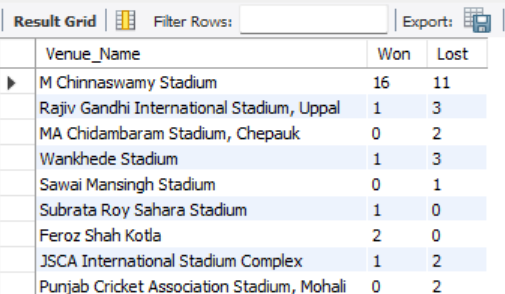
Ans:

RCB maximum no. of matches won at home ground chinnaswamy stadium and zero match won at sheikh Zayed stadium & saurastha cricket association stadium.

**Code:**



**Output:**

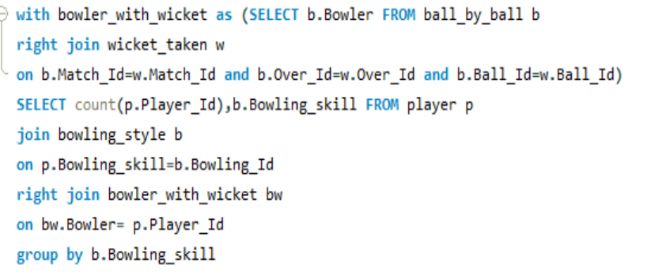


1. What is the impact of bowling style on wickets taken?

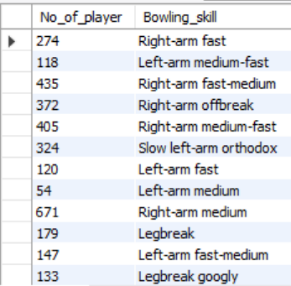
Ans: Right arm medium bowlers were taken highest wicket = 671

Where, Left arm medium bowlers were taken lowest wicket = 51

**Code:**



**Output:**



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:

I created a new column named as performance\_status.

It is showing status ‘improved’ if runs and wickets both more than

Previous season,

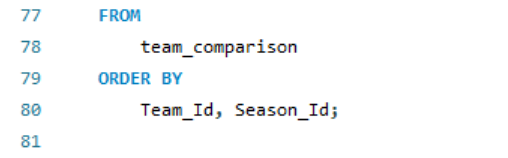
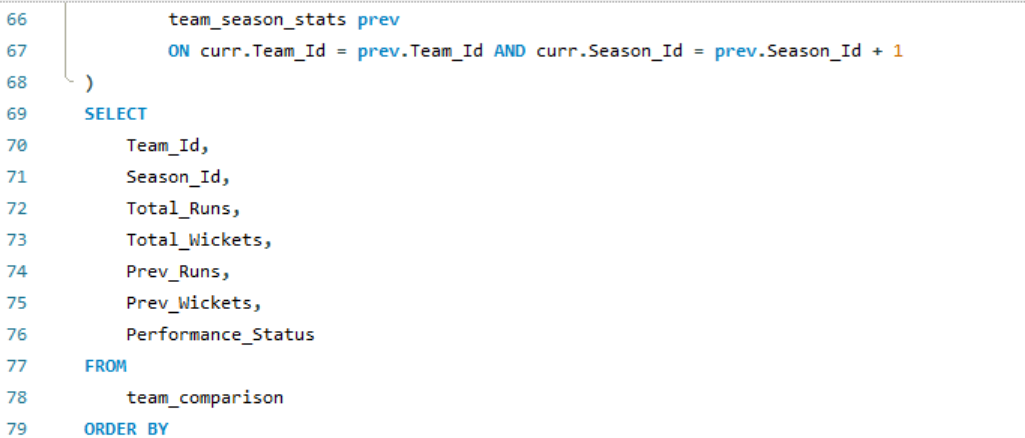
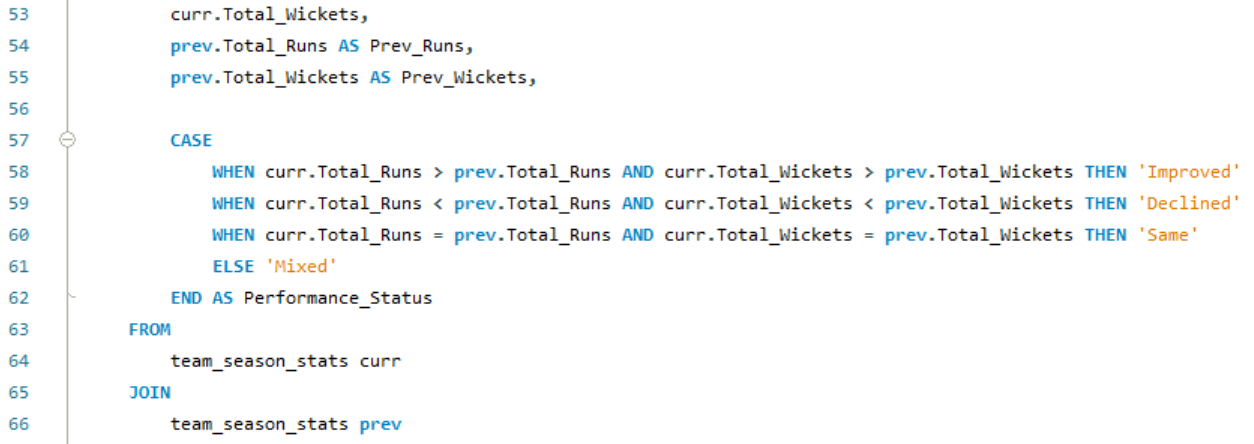
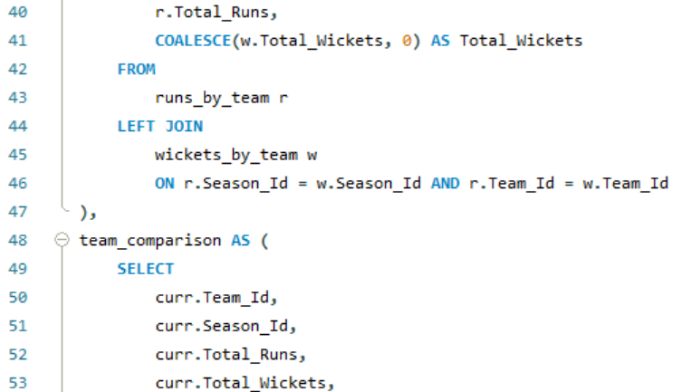
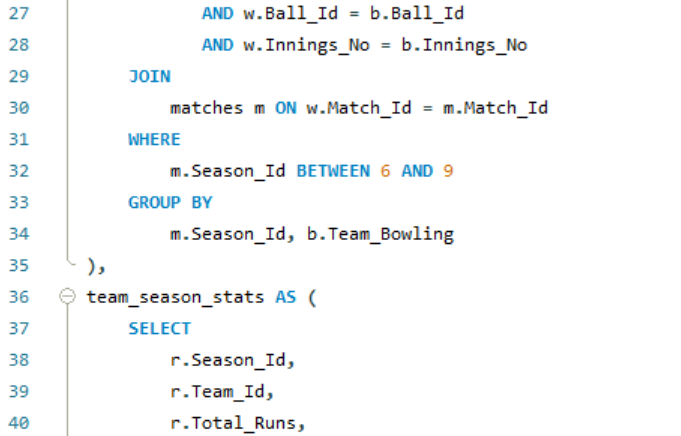
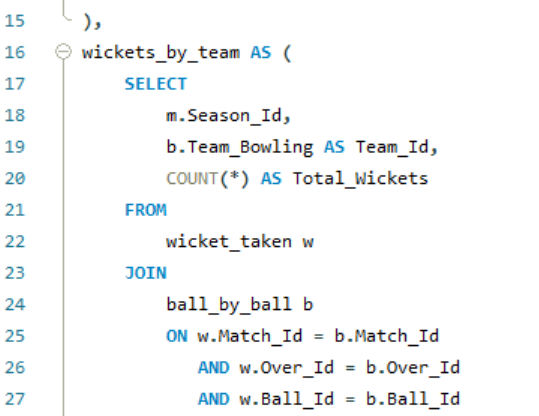
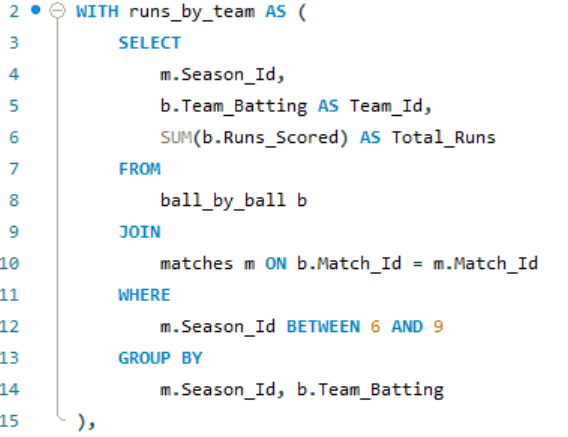
‘Declined’ if runs and wickets both less than

Previous season,

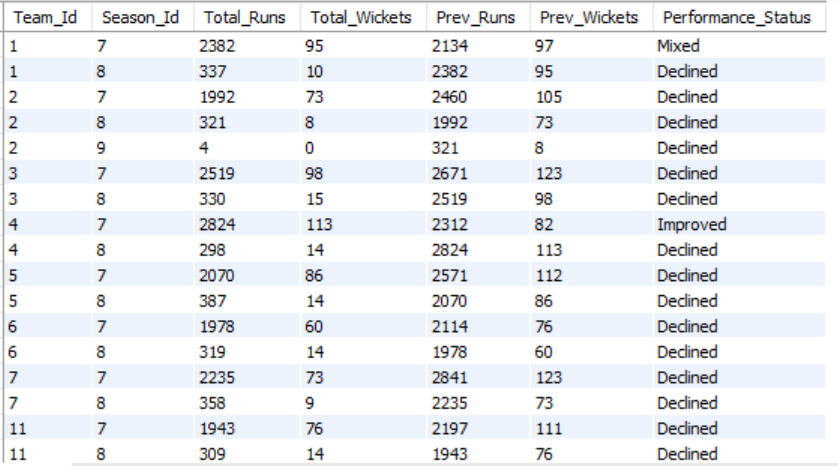
‘Same’ if both season runs and wickets were same,

Mixed if either one of runs & wickets more than previous season.

**Code:**



**Output:**

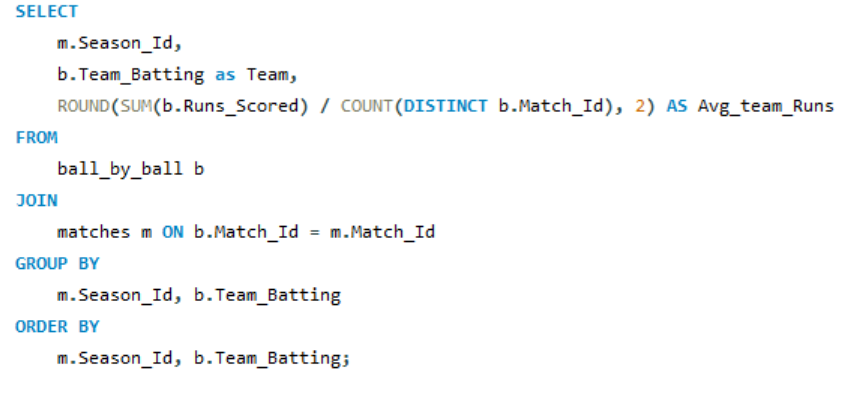


1. Can you derive more KPIs for the team strategy?

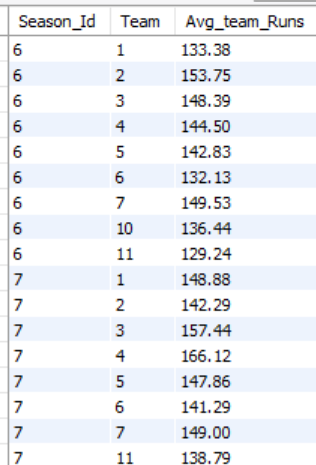
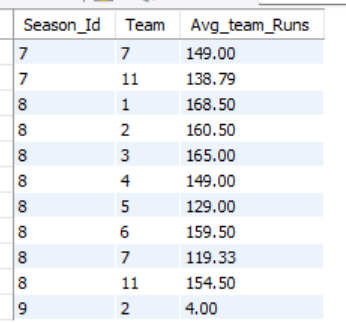
Ans: Yes, Average teams runs scored by each season

In season 6,7 and 8 RCB average runs in a inning are 153,142 and 160 respectively.

**Code:**

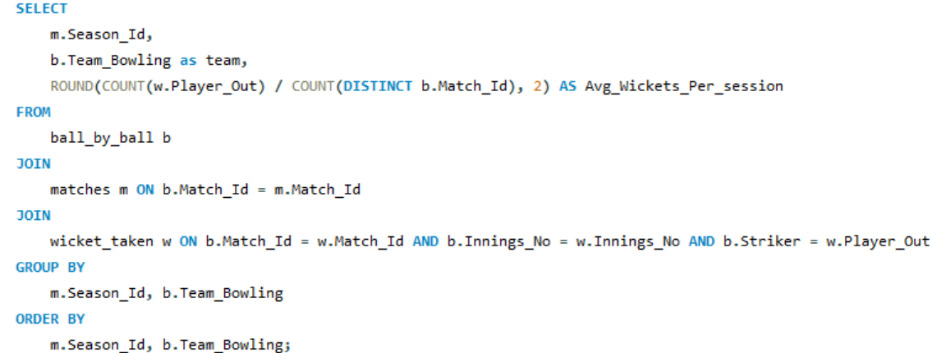


**Output:**

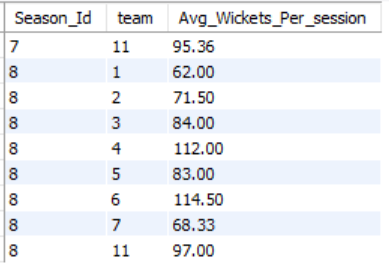
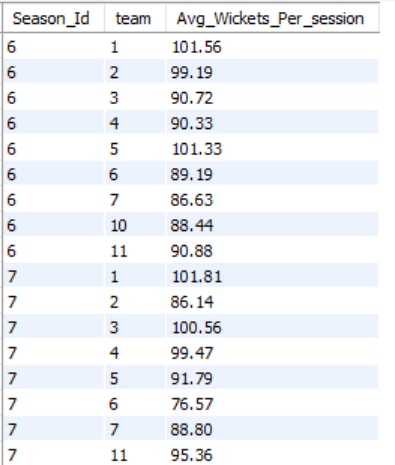
 

Average team wicket taken by each season

**Code:**



**Output:**

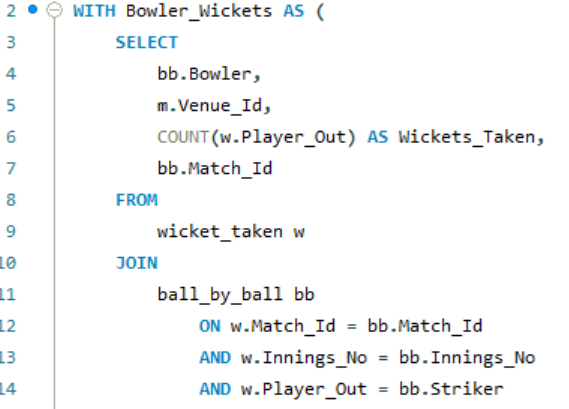
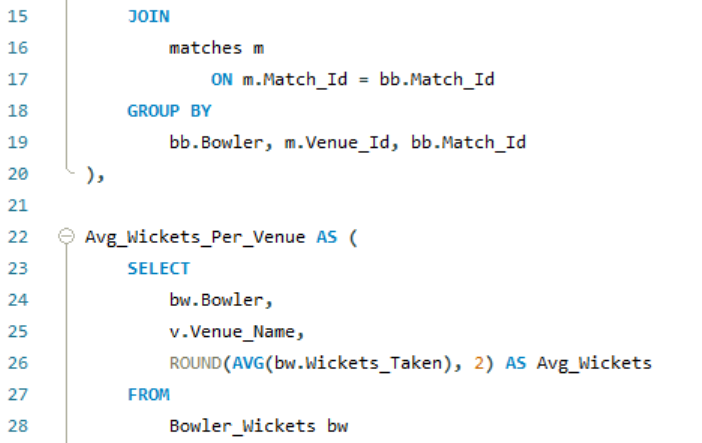


1. Using SQL, write a query to find out the average wickets taken by each bowler in each venue. Also, rank the gender according to the average value.

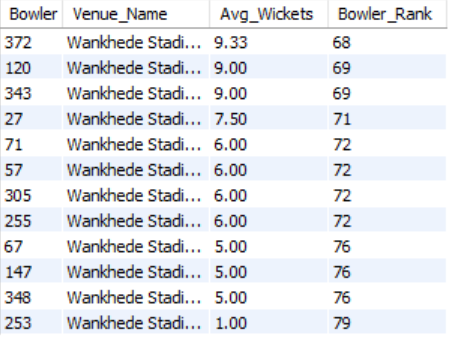
Ans:

The average wicket taken by each bowler in each venue are

**Code:**

**Output:**

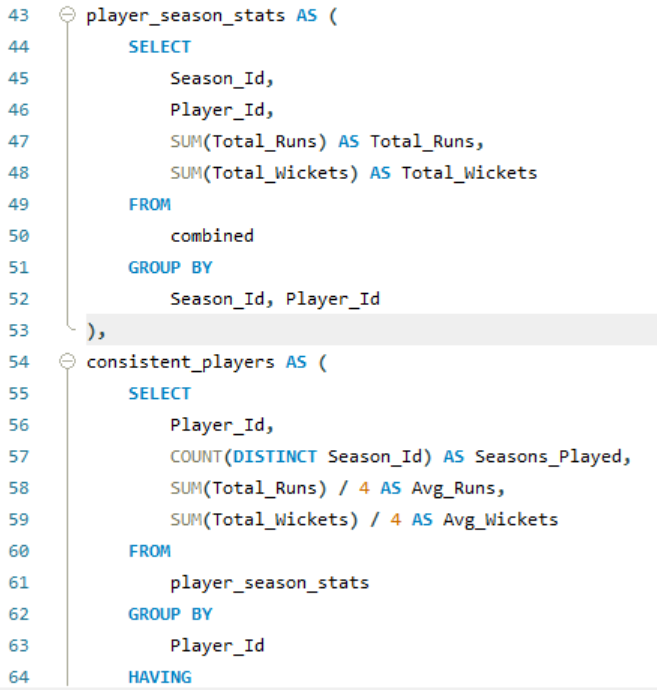
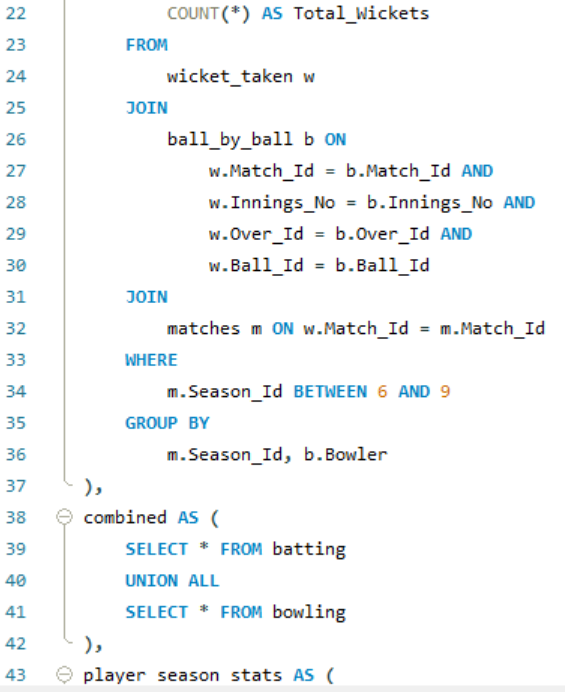
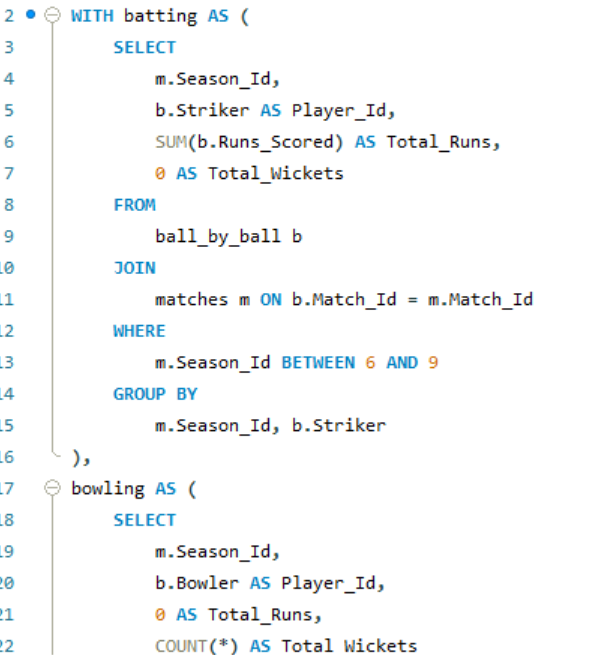


1. Which of the given players have consistently performed well in past seasons? (will you use any visualization to solve the problem)

Ans:

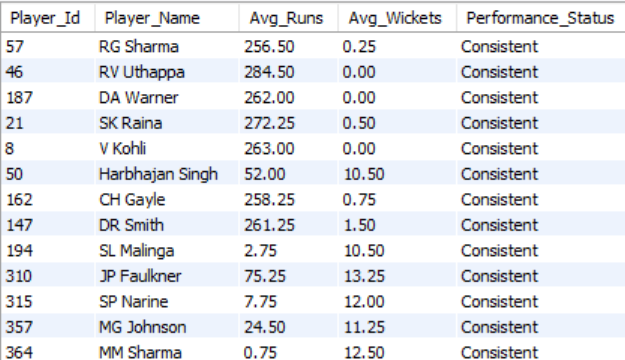
I have considered consistent player whose ‘average run per season’ is greater than ‘250’ or ‘average wicket per season’ is greater ‘10’.

**Code:**

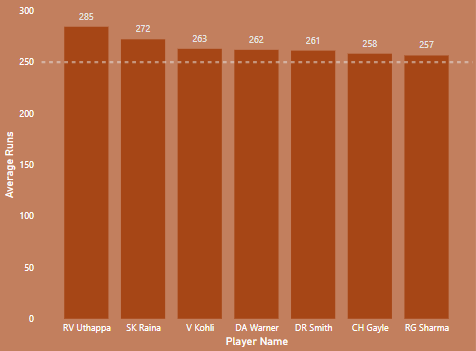
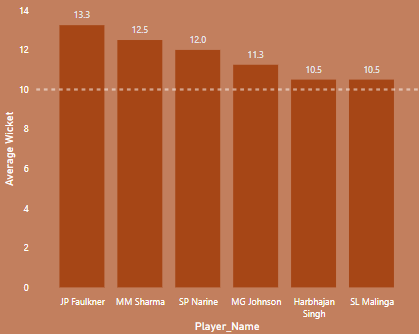




**Output:**



**Visualization:**

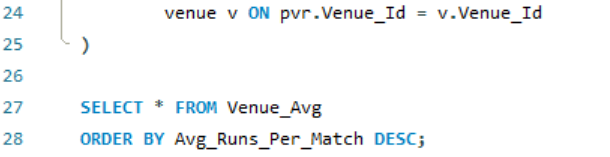
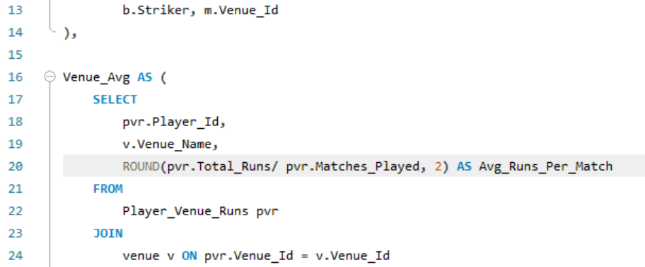
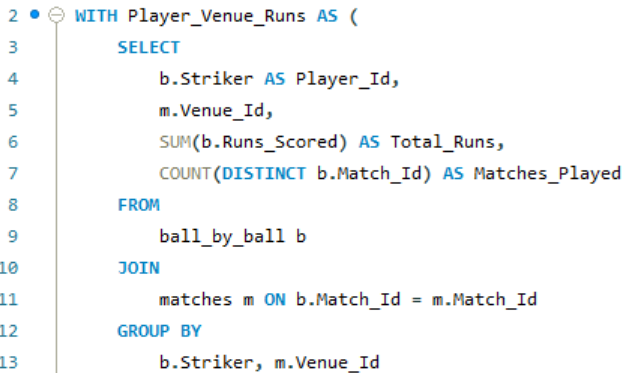
1. Are there players whose performance is more suited to specific venues or conditions? (how would you present this using charts?)

Ans:

Watson plays excellent with the average of 101 in MA

Chidambaram Stadium, Chepauk And simmons plays excellent with the average of 100 in Punjab Cricket Association Stadium, Mohali.

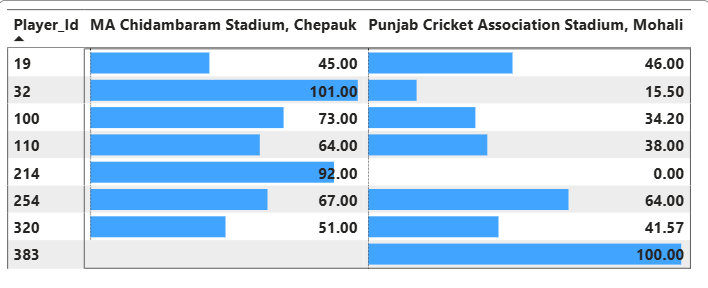
**Code:**



**Output:**



**Visualization:**

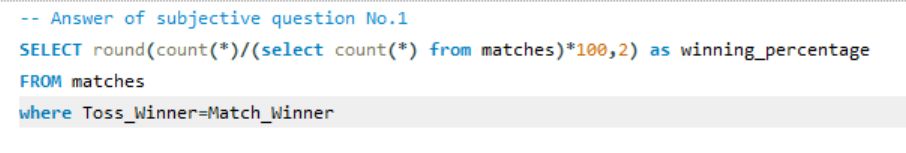


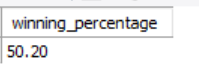
Subjective Questions

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?

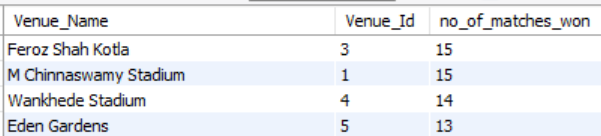
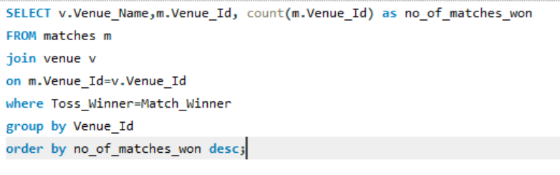
Ans:

**Code:**

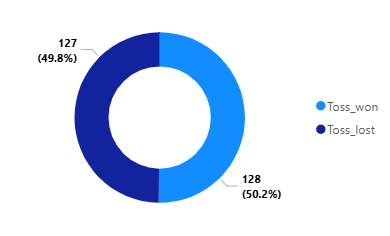




These 4 venue are highly impacted by toss decision



**Visualization:**



**Approach:**

Firstly, I counted the total no. of match where matches won when toss won and I divide by the total no. of matches. It gives the ratio. So, for percentage I multipied by 100.

**Insights:**

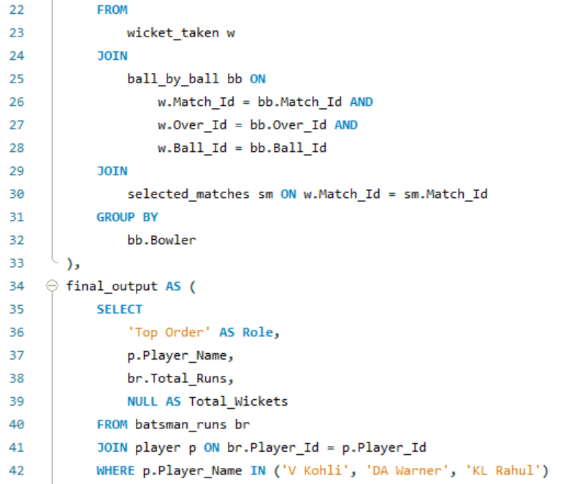
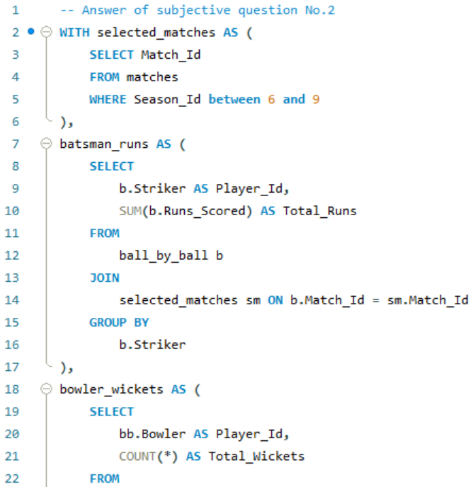
50.20% winning chances if you win the toss.

**Recommendation:**

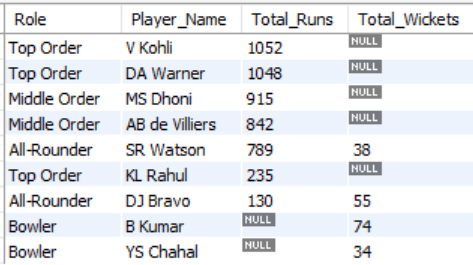
1. Pitch conditions: Early in a match, the pitch might be hard and flat — great for batting. Later, it might crack or assist spinners/seamers. So, captain might choose to bat first
2. Weather: Overcast skies can help swing bowlers, so a captain might choose to bowl first.
3. Dew factor: In day-night matches, dew can make the ball slippery and hard to grip later — so teams often prefer to chase.
4. Suggest some of the players who would be best fit for the team.

Ans:

**Code:**



**Output:**



**Approach & Insights:**

Players who best fit for the team are:

1. Virat Kohli

Reason - Insane form in 2016(season 9), leadership potential

Role – Top order batsman

1. David Warner

Reason - Consistent, aggressive, led SRH to title

Role – Opener

1. AB de Villiers

Reason - Finisher, explosive, consistent in 2016

Role - Middle-order Bat

1. Bhuvneshwar Kumar

Reason - Purple cap winner, excellent death bowling

Role - Swing Bowler

1. Andre Russell

Reason - Power hitter, wicket-taker

Role - All-rounder

1. Yuzvendra Chahal

Reason - Economical + reliable spinner, wicket-taker

Role - Leg Spinner

1. MS Dhoni

Reason - Calm finisher, great leader

Role - Wicket-keeper/Captain

1. Dwayne Bravo

Reason - Death bowling specialist, handy batsman

Role - All-rounder

1. Shane Watson

Reason – Experience (Australia captain), powerplay impact player

Role - All-rounder

1. KL Rahul

Reason - Young talent, played mature innings for RCB

Role - Flexible batsman (bat can any order)

1. Mustafizur Rahman

Reason - Surprise package, excellent slower balls

Role - Left-arm pacer

**Recommendation:**

Sample Team Composition:

Top Order: Kohli, Warner, Rahul

Middle Order: AB de Villiers, MS Dhoni (WK), Andre Russell

All-Rounders: Shane Watson, Dwayne bravo

Bowlers: Bhuvneshwar Kumar, Mustafizur Rahman,

Yuzvendra Chahal

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

Ans:

**Strike Rate**: We need batsmen who runs at >130 strike rate.

Economy: We need bowlers who concede <8 runs per over and take 2+ wickets regularly.

**Player of the match**: If his skill set overlaps with others, more numbers of potm will win.

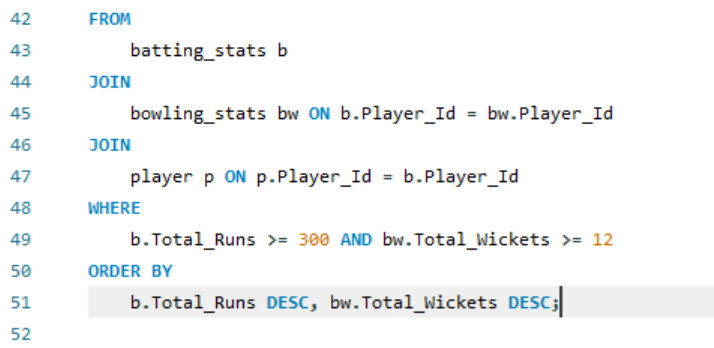
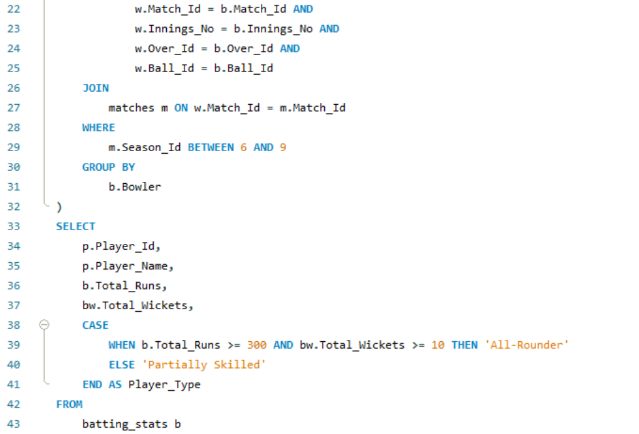
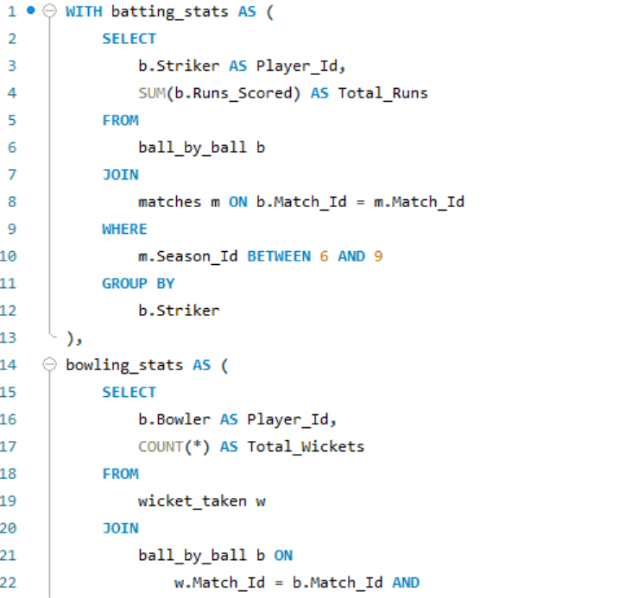
**Boom and Bust**: Avoid Boom and Bust pick means **perform explosively** (“boom”) or **completely fail** (“bust”) that leave gaps on off days; injuries strain the overseas quota.

**All-rounder**: Can bat anywhere 3-7 & bowl 2-4 overs? That’s Russell/Hardik category. **Fielding runs saved**, direct-hit run-outs, boundary stops; modern teams value 10–15 runs saved per match.

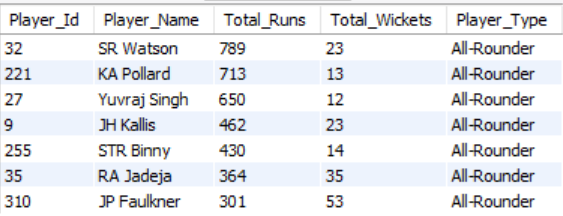
1. Which players offer versatility in their skills and can contribute effectively with both bat and ball? (can you visualize the data for the same)

Ans:

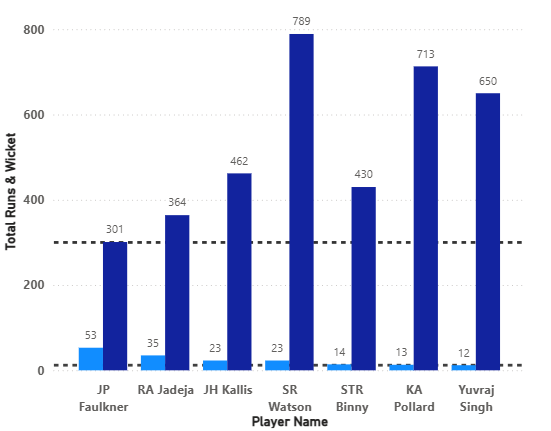
**Code:**



**Output:**



**Visualization:**



**Approach:**

I have taken versatile player who scored run at least 300 and

Wicket taken at least 12 & order by total runs and wickets in decreasing order.

Insights:

I got 7 players who played exceptionally well with both ball and bat.

Watson with 789 runs and 23 wickets & Faulkner with 301 runs and 53 wickets from Australia.  
Yuvraj Singh with 650 runs and 12 wickets & Jadeja with 364 runs and 34 wickets from India.

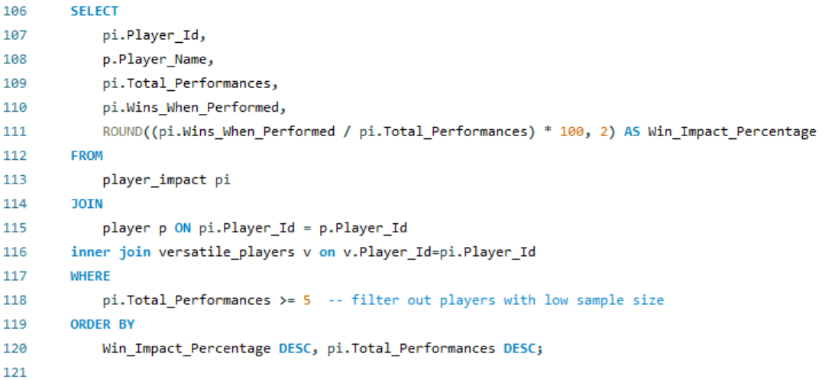
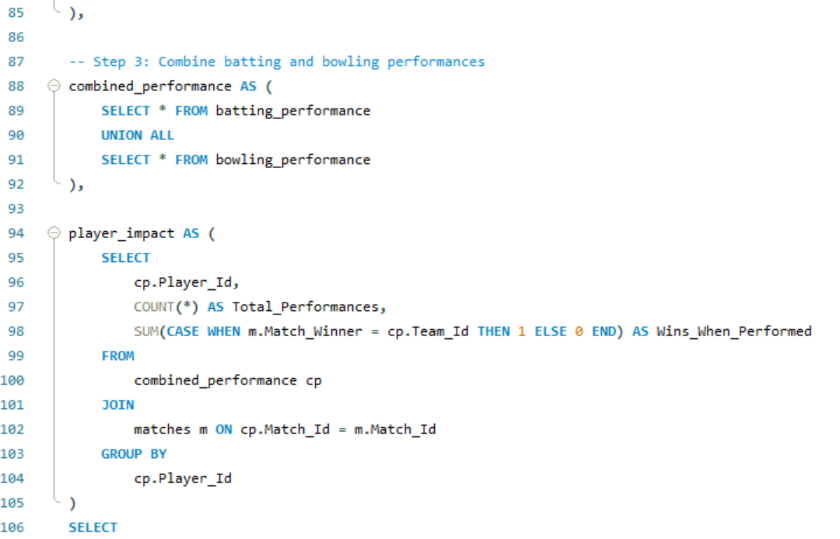
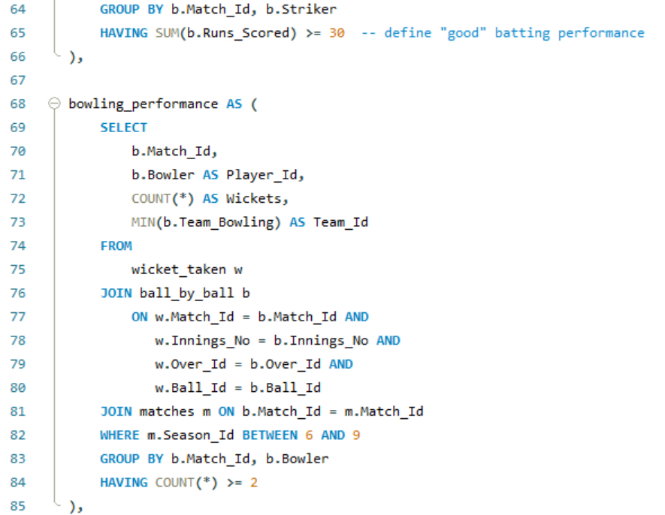
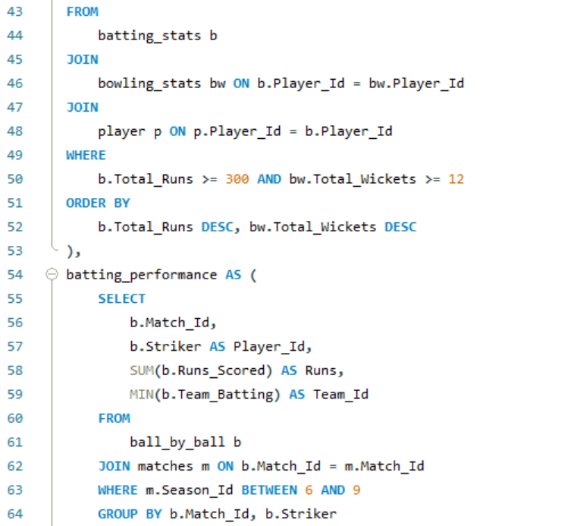
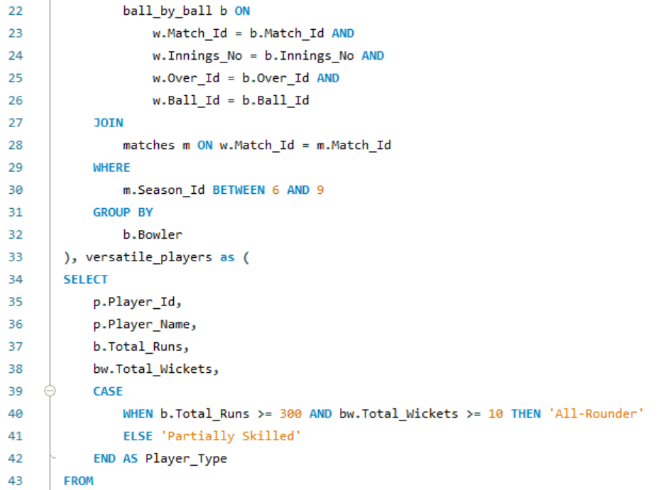
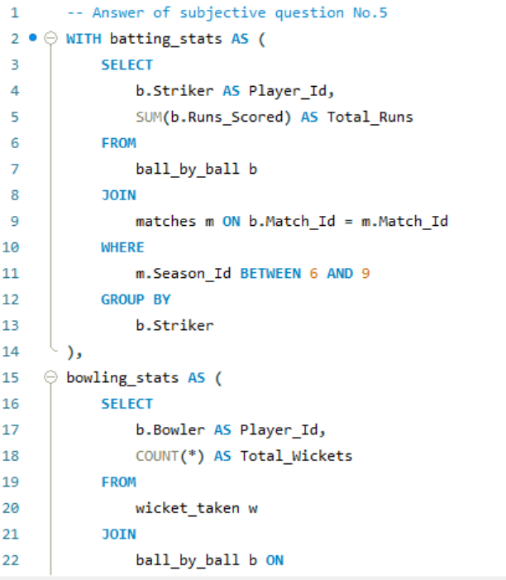
**Recommendation:**  
Try to bring more and more players from this list.

At least go for 2 overseas player and take maximum Indian players because they can contribute with both bat and ball.

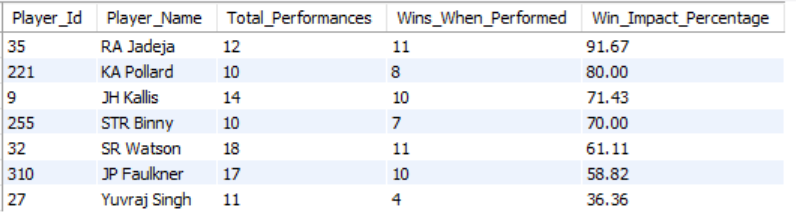
1. Are there players whose presence positively influences the morale and performance of the team? (justify your answer using visualization)

Ans:

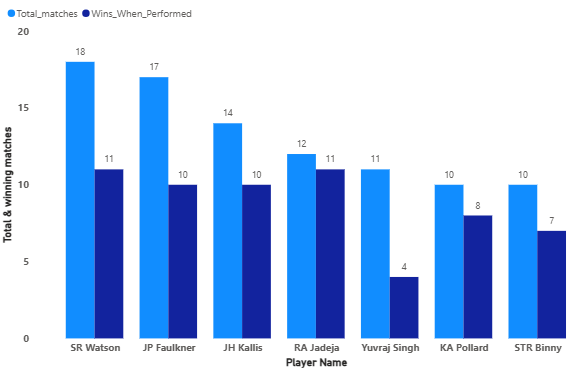
**Code:**



**Output:**



**Visualization:**



**Approach:**

I took as batsman they scored 30+ runs and as bowler they have taken wicket greater than and equal to 2 in a match and they have played at least 5 matches so I considered as good performer . So, I got to know all 7 players are influences positively

And performance of the team as well.

**Insights:**

High impact players- RA Jadeja with 91.67% win impact & KA Pollard with 80% win impact.

Low impact players- Yuvraj Singh with low 36% win impact

**Recommendation:**

Avoid overpaying for low-impact players: Players like Yuvraj Singh (36% win impact).

Prioritize high-impact performers: Players like RA Jadeja (91.67% win impact) and KA Pollard (80% win impact).

Target versatile all-rounders: High-impact all-rounders (Jadeja, Watson) contribute in multiple disciplines (bat & ball) and are crucial for team flexibility and balance.

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

Ans:

* Retention (Retain smart, not just stars)
  1. Virat Kohli - Icon player, consistent anchor
  2. Glenn Maxwell - Flexibility, spin-hitter, off-spin option
  3. Mohammed Siraj - Powerplay specialist, improved at death
  4. Faf du Plessis / Cameron Green – Leadership (Faf du Plessis) or Value for money (Green)
* Middle-order Indian batters (flexible, spin-capable, and free up overseas slots)

1. Nitish Rana
2. Washington Sunder
3. Shivam Dube

* Reliable death over bowlers (RCB’s biggest weakness year after year)

Overseas-

1. Mitchell Starc
2. Lockie Ferguson
3. Reece Topley

Indian-

1. Avesh Khan
2. Tushar Deshpande
3. Arshdeep Singh

* Venue-Specific Planning (Chinnaswamy = high-scoring)

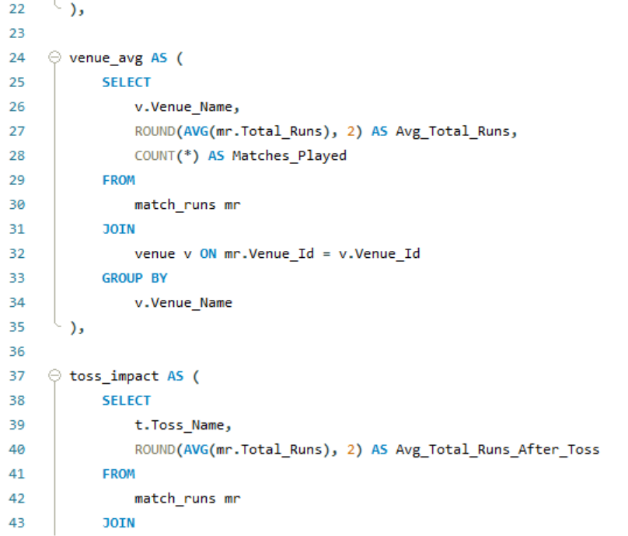
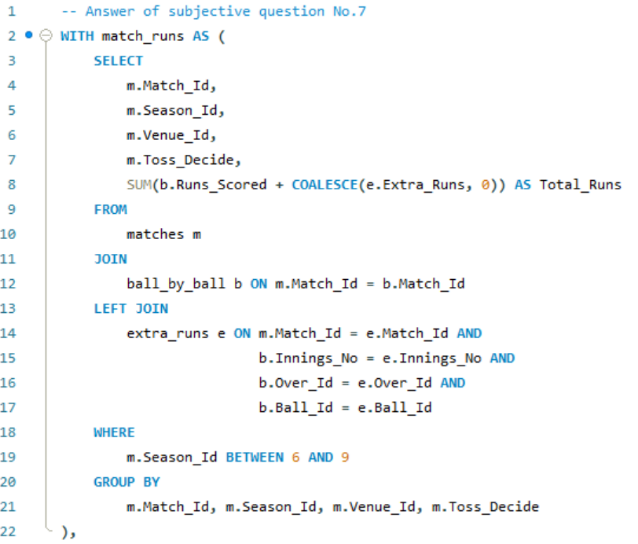
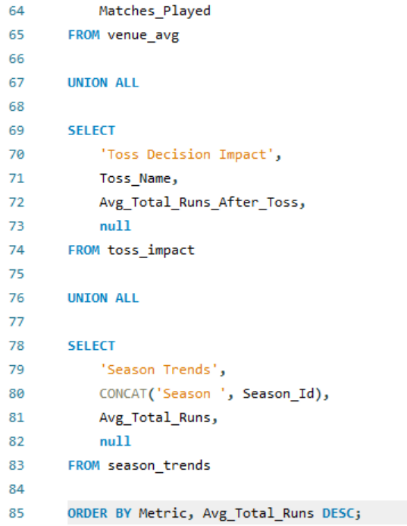
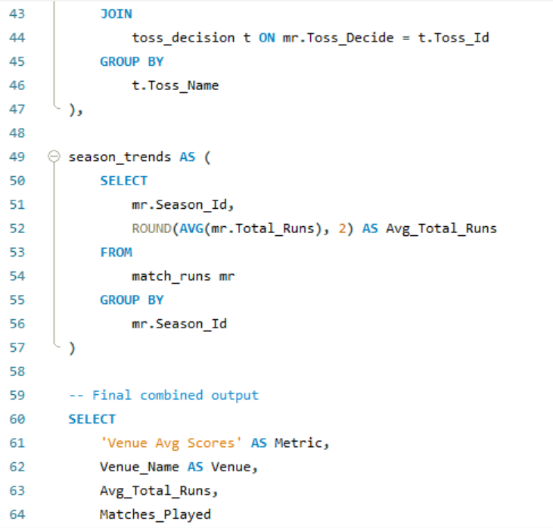
1. Build a **bat-deep team** with at least 8 batting options.
2. Stack **pace variations & cutters**, not express quicks alone.
3. Prefer **spinners who take wickets.**

|  |
| --- |
|  |

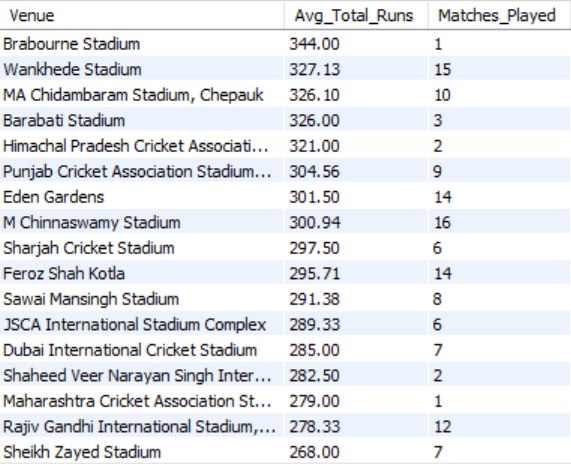
1. What do you think could be the factors contributing to the high-scoring matches and the impact on viewership and team strategies?

Ans:

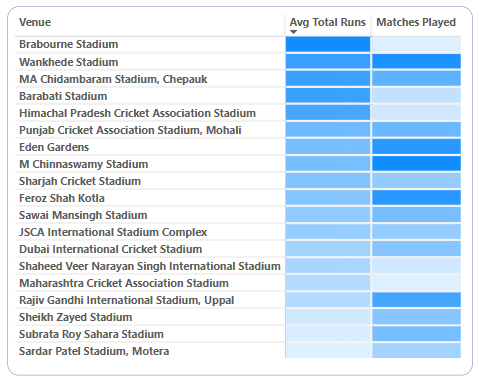
**Code:**

**Output:**



**Visualization:**



**Approach:**

I sum up all the runs for the both innings of a match  
and use average function, group by venue and also count no. of matches for each venue then I got average total runs venue wise and total matches for each venue.

**Insights:**

Highest average runs scored in a match was 344 runs at Brabourne stadium but only 1 match played where as 2nd highest average runs scored in a match was 327 runs at Wankhede stadium with 15 matches played.

Lowest average runs scored in a match was 263 runs at Sardar Patel Stadium, Motera with 5 matches played.

**Recommendation:**

Impact on viewership:

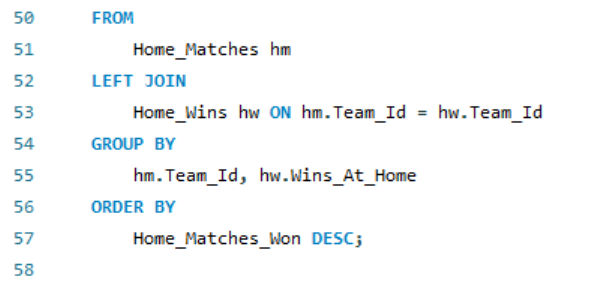
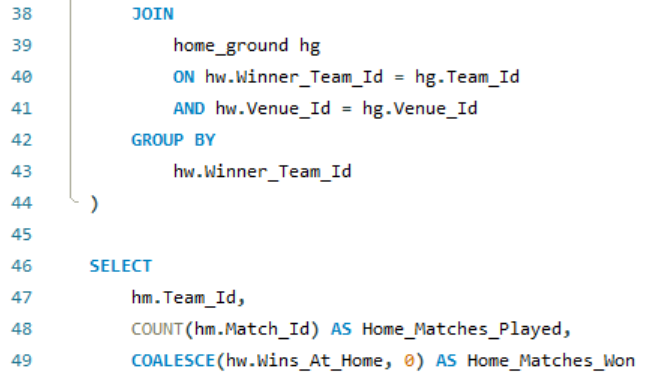
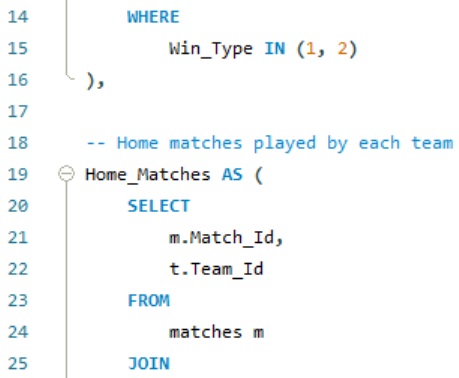
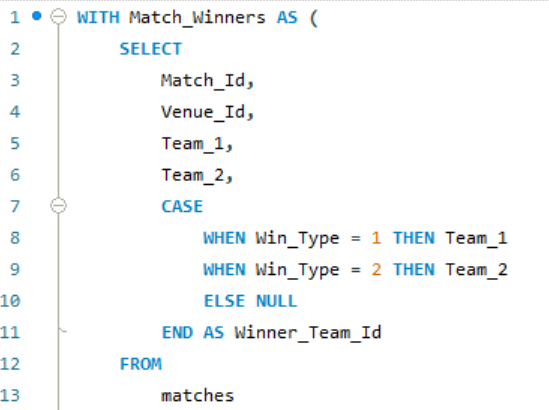
1. Big total means more boundaries and it engage more fans.
2. Sponsorship - Brands love ad slots in high-stakes games.
3. More TRP is gained when 200+ runs are chased.

Impact on team strategies:

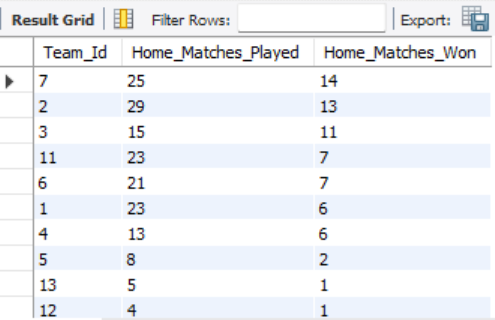
1. Team selection - More all-rounders and power-hitters over classical batters.
2. Toss Decisions – Prefer to chase, using dew to their advantage.
3. Analyze the impact of home-ground advantage on team performance and identify strategies to maximize this advantage for RCB.

Ans:

**Code:**



**Output:**



**Approach:**

I created a table name as “home\_ground”. In this table, I portrayed home ground for all the teams. Then with the help of this table found out for the all team, how many matches they have won in home ground and total matches played in home ground.

Insights:

Mumbai Indians and Chennai super king won more than 50%

matches in home ground whereas Gujarat Lions won only 20%

matches on their home ground.

**Recommendation:**

RCB’s home win rate (~45%) lags behind top teams like MI & CSK.

prioritize players with strong records at Chinnaswamy Stadium to capitalize on home conditions.

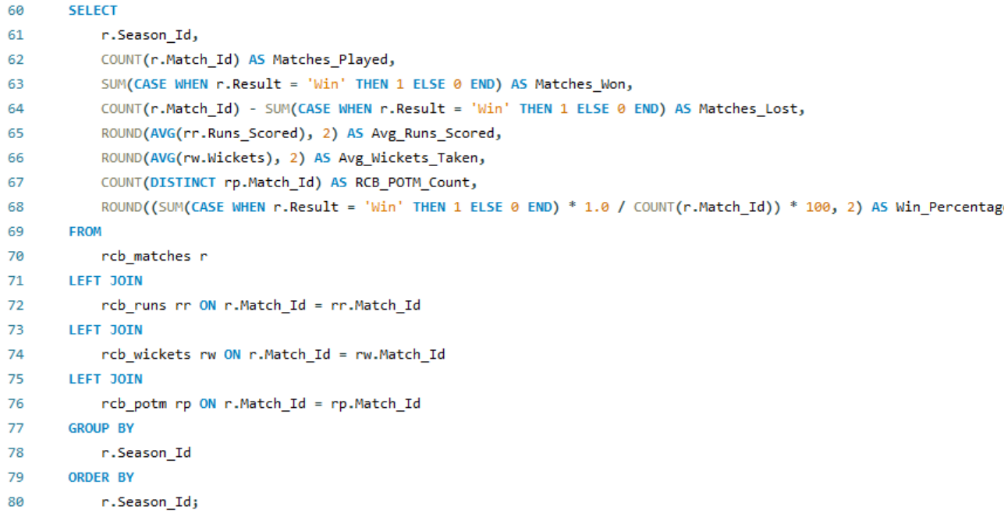
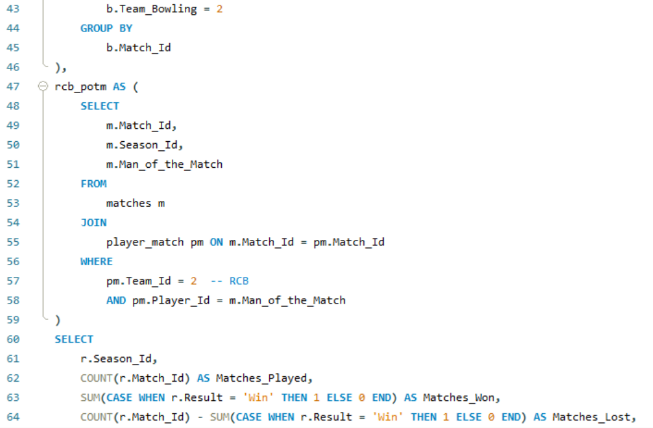
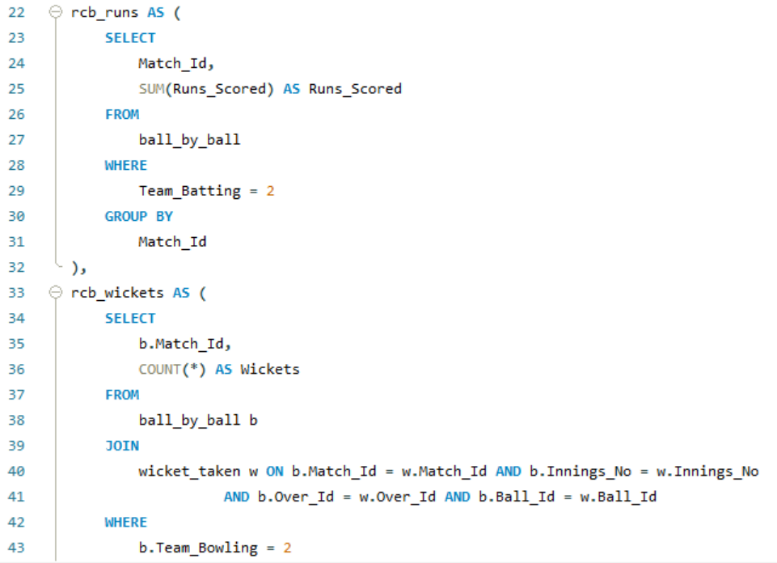
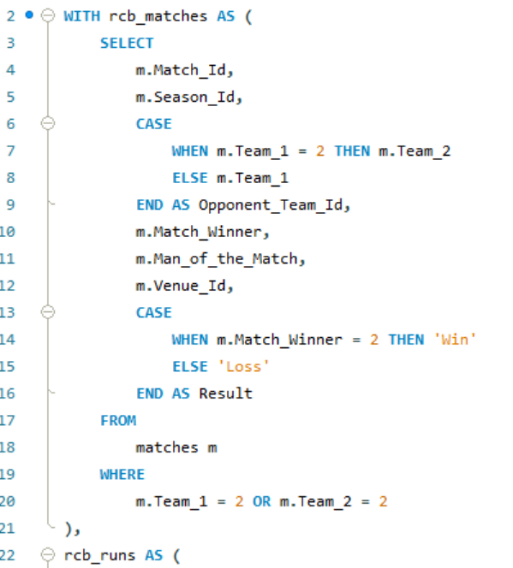
Chinnaswamy’s small boundaries demand aggressive batsmen who can exploit the venue’s high-scoring nature, increasing RCB’s ability to set/ chase big totals at home.

Since MI & CSK succeed at home consistently, study their player profiles and gameplans to adopt proven strategies suited for home dominance.

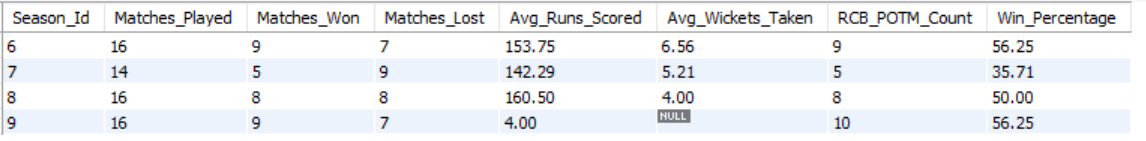
1. Come up with a visual and analytical analysis of the RCB's past season's performance and potential reasons for them not winning a trophy.

Ans:

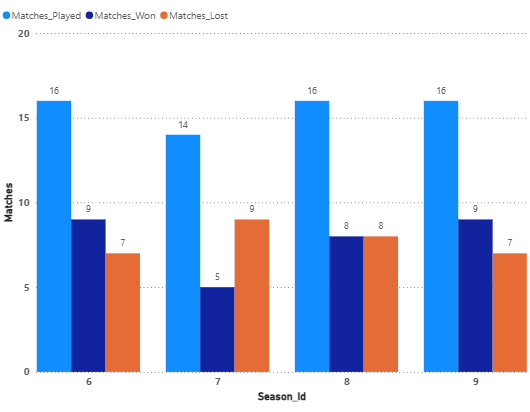
**Code:**



**Output:**



**Visualization:**



**Approach:**

I widely go with these following question:

The RCB past season’s performance include:

1.How many matches they played in every season?

2.How many matches they won every season?

3.How many matches they lost every season?

4.What is the average run scored in an inning by RCB in every

Season?

5.What is the average wicket taken in an inning by RCB in every

Season?

6.How many players received player of the man award in every

Season?

**Insight:**

1. 14-16 matches they played in every season.16 match means RCB qualified for play off.
2. Maximum 9 matches and minimum 5 matches RCB won in a season.
3. 142-160 average run scored in an inning by RCB in a season.
4. 4-6.5 average wicket taken in an inning by RCB in every season.
5. Maximum 10 man of the match award and minimum 5 man of the match award won by RCB in a season.

**Recommendation:**

1. Since RCB’s seasonal wins ranged widely (5–9 wins), invest in clutch players with proven ability to deliver under pressure, boosting chances of qualifying.
2. RCB’s average of just 4–6.5 wickets per innings, prioritize bowlers or bowling all-rounders who can reliably take breakthroughs, especially during middle overs.
3. Focus on players with a track record of earning Man of the Match awards (5–10 per season for RCB) — these are high-impact performers who can turn games single-handedly.
4. How would you approach this problem, if the objective and subjective questions weren't given?

Ans:

Step – 1

Define metrics – For batsmen

1. Total Runs scored by RCB from season 1.
2. Average runs scored by team in a match.
3. Consistency – No. of players who scored 30+ runs in how many innings.

For bowlers

* 1. Total wicket taken by RCB from season 1.
  2. Total wicket taken by each bowler.
  3. Average wicket taken by each bowler in a match.
  4. Economy rate of each bowler.

For all-rounder

1. Average wicket taken and Average run scored in a match.

Step – 2

Home ground and match analysis

1.Who performs well at Chinnaswamy (RCB's home)?

1. Who handles pressure (in playoff/close game performance)?
2. Who influence team morale (wins in matches where they did well)?
3. 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: There is no such column in ‘Match’ table in which incorrectly

Spelled as ‘Delhi Capital’

