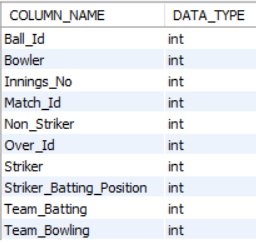
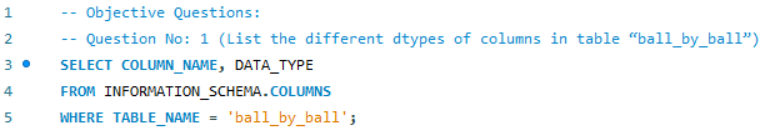
**Project: IPL Strategy For RCB**

**Objective Questions**

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

**Ans:** There are 10 columns with the data type integer.





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:**

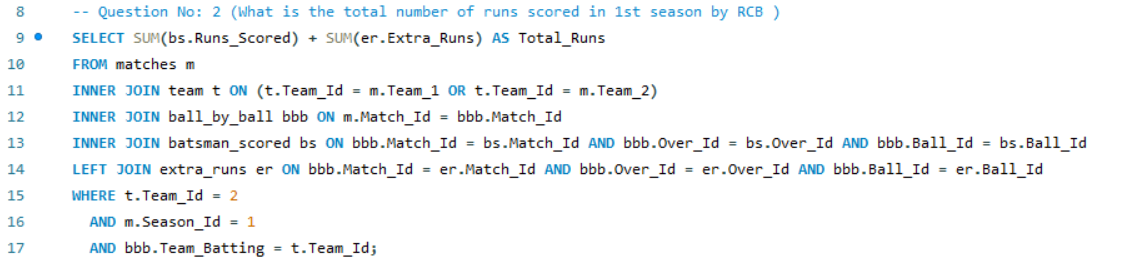
The following tables are required in order to determine Royal Challengers Bangalore's (RCB) total runs scored during the first IPL season, including additional runs:

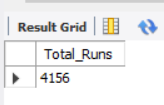
**Required Tables:**

1. **matches**: Contains match details, including the teams and season.
2. **team**: Provides team names and IDs.
3. **ball\_by\_ball**: Tracks each ball bowled, including the team batting.
4. **batsman\_scored**: Records runs scored by batsmen on each ball.
5. **extra\_runs**: Tracks additional runs like wides and no-balls.

**Query Approach:**

1. **Match Selection:** Filters matches where RCB played (Team\_Id = 2) in the first season (Season\_Id = 1).
2. **Joining Tables:** Joins relevant tables to link matches, balls, runs scored, and extra runs.
3. **Sum Runs:** Adds up runs scored by RCB's batsmen and any extra runs.
4. **Result:** The total runs scored by RCB in the first season.



****

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

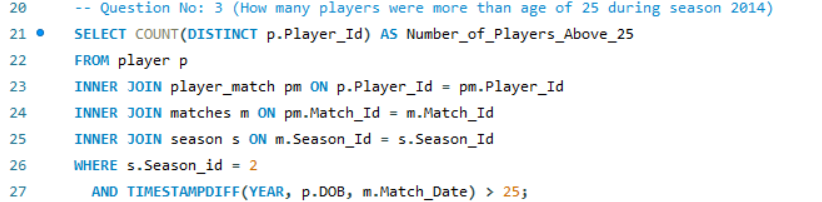
**Ans:**

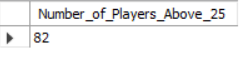
To find the number of players who were more than 25 years old during season 2014, we need the following tables:

1. **player** - To get the player's date of birth.
2. **player\_match** - To find which players participated in matches during season 2014.
3. **matches** - To find matches in season 2014.
4. **season** - To identify season 2014 and its matches.

**Query Approach:**

1. Identify all matches in season 2014.
2. Find all players who played in these matches.
3. Calculate the age of each player during season 2014.
4. Count the players who were older than 25.



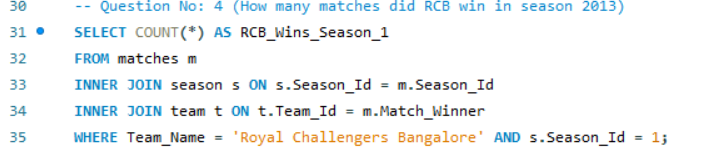


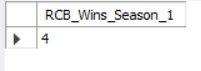
1. How many matches did RCB win in 2013?

**Ans:**

The following tables are required in order to calculate the number of games that Royal Challengers Bangalore (RCB) won in season one:   
1. **Matches:** To determine the winning team and other match details.   
2. **season:** To narrow down matches from the first season.   
3. **team:** To match RCB with the victorious team.   
  
**Query Method:**

1. Find matches from the first season using the season table.   
2. Sort the matches according to which RCB is the Match\_Winner.   
3. Determine the number of games in which RCB prevailed.   
The total number of games won by RCB in season 1 is provided by this approach.

****

****

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

**Ans:**

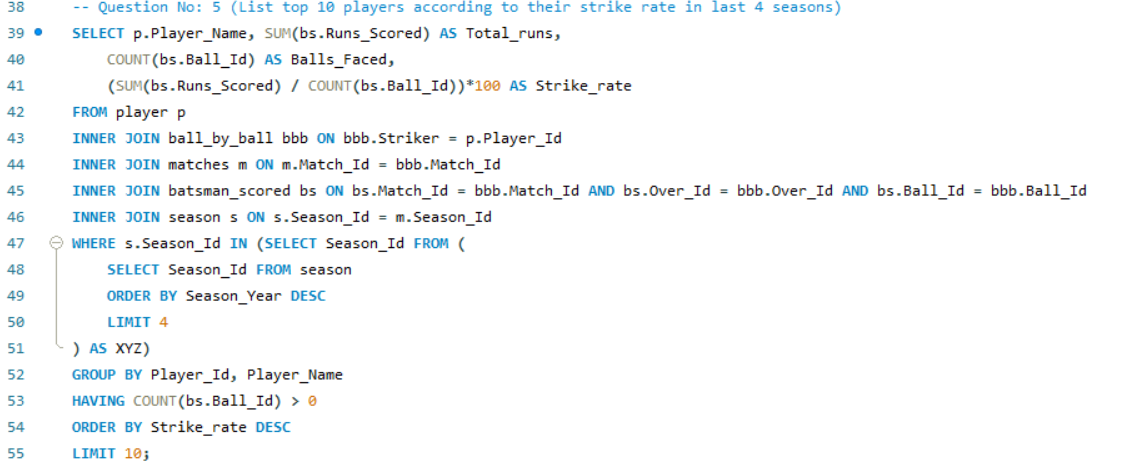
The following tables are needed in order to list the top 10 players by strike rate over the last four seasons:

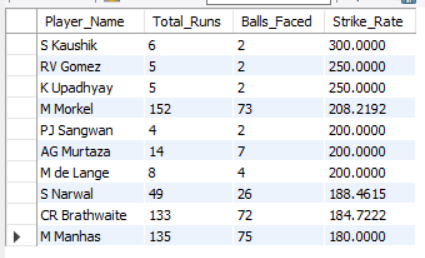
o **player:** Provides the player's name (Player\_Name) and Player\_Id to uniquely identify each player  
o **matches:** Filters matches in the last four seasons;   
o **season:** Assists in identifying the last four seasons using Season\_Id  
o **ball\_by\_ball:** Provides the player's details for each ball faced enabling us to determine the total number of balls faced  
o **batsman\_scored:** Provides the runs scored by the player on each ball, enabling us to determine the total runs.

**• Query Approach:**   
o We determine how many runs and balls each player has faced overall.  
Next, figure out the strike rate.  
o Based on this rating, we finally present the top 10.

**• Takeaways:**   
o Players with outstanding strike rates—more than 200 in the last four seasons—such as S Kaushik, RV Gomez, and K Upadhyay have demonstrated aggressive and effective scoring in few opportunities.   
o Despite not being specialized batsmen, bowlers like M Morkel and PJ Sangwan are included in the list because of their ability to swiftly score crucial runs.   
o Quick scoring by aggressive lower-order hitters like M de Lange and CR Brathwaite may change games, which is essential in shorter forms like T20.   
o The players on the list have a strike rate of more than 180, which indicates that they can speed up the innings and change the course of games under duress.

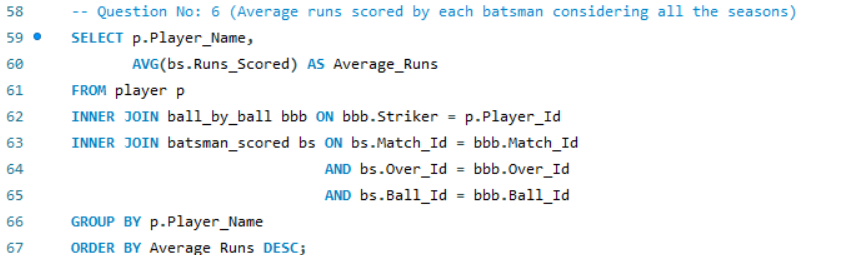
**• Suggestions:**   
o To increase the team's overall score, teams could use these high-strike-rate players in crucial overs, particularly in the last overs in Twenty20 matches.   
o In order to take advantage of these players' capacity to score runs rapidly under duress, strategic planning should concentrate on rotating them in middle-to-lower batting orders.

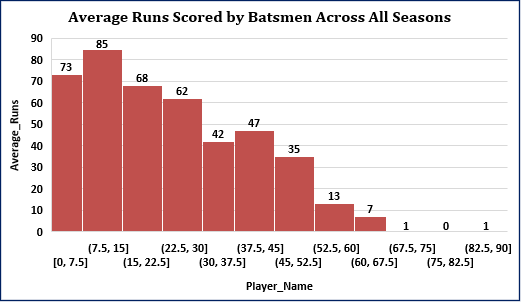




1. What are the average runs scored by each batsman considering all the seasons?

The following tables are required in order to determine the average number of runs scored by each hitter over all seasons:  
1. The player gives player IDs and names.  
2**. batsman\_scored:** Provides details on how many runs each batsman scored on each ball.   
3. **ball\_by\_ball:** Assigns a player to each run scored.   
  
**Query Method:**   
• **Joins:** To get each player's runs, the tables are connected.   
• **Calculations:** Each player's average of these runs is determined.   
• **Sorting:** Players with the highest averages are displayed first in the results.

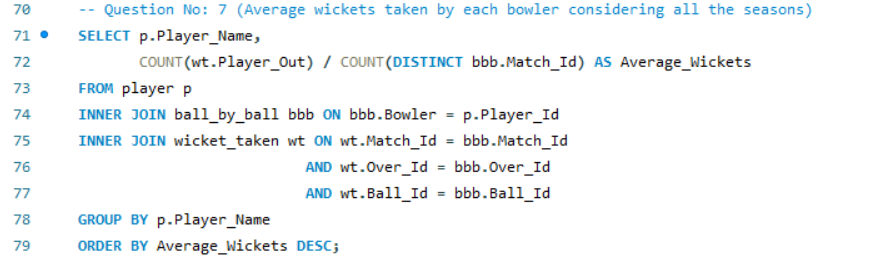


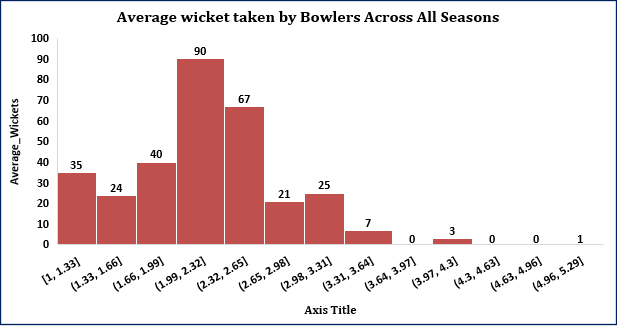


1. What are the average wickets taken by each bowler considering all the seasons?

**Ans:**

The following tables are required in order to determine the average number of wickets taken by each bowler over the course of all seasons:   
1. **Player:** Offers player IDs and names.   
2. **wicket\_taken:** Provides details on the number of wickets each bowler has taken on each ball.   
3. **Ball by ball:** identifies the bowler who delivered each ball.   
  
**Query Approach:**   
• **Joins:** To determine the number of wickets each bowler has taken, the tables are combined directly.   
• **Calculations:** The average number of wickets taken by each player is determined by dividing the total number of wickets by the number of matches.   
• **Sorting:** Players with the highest averages are displayed first in the results.



****

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

The following tables are required in order to compile a list of all the players whose average runs scored and wickets taken above the general average:

**• Tables Needed:**

o **Player:** Offers player IDs and names. o **batsman\_scored:** Provides details on the number of runs each batsman has scored. o **ball\_by\_ball:** Assigns each ball to the corresponding bowler and batsman. o **wicket\_taken:** Provides details on the number of wickets each bowler has taken.

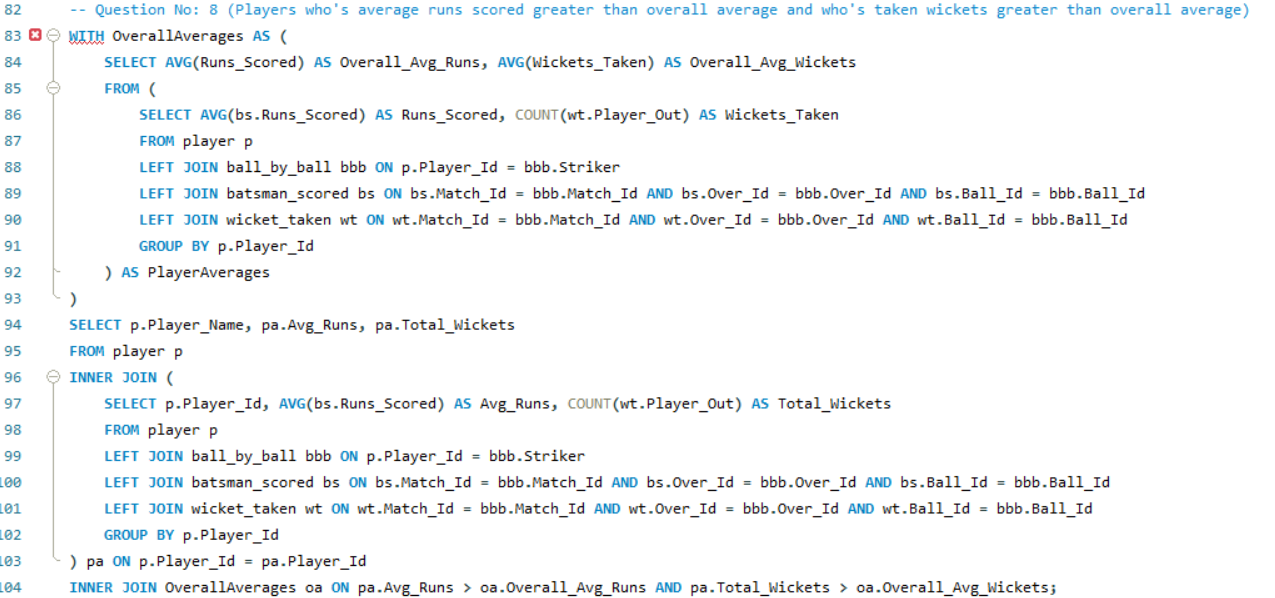
**• Query Method:**

o Compute Overall Averages: Find the average number of runs scored and wickets taken by each player overall.

o Examine and contrast player averages: Enumerate players who surpass the average number of runs and wickets.

**• Takeaways:**   
o Players who routinely surpass overall averages in runs and wickets, like as AB de Villiers, AD Russell, and CH Gayle, stand out for having a balanced influence with the bat and ball.  
o All-rounders with great versatility, like JA Morkel, JP Faulkner, and GJ Maxwell, make a substantial contribution to both run scoring and wicket-taking.  
o Batsmen like YK Pathan and Harbhajan Singh, who also make a substantial batsman contribution, give their teams a lot of versatility and balance.   
o The availability of several all-round players, such as SA Yadav and V Sehwag, demonstrates their ability to play both defensive (bowling) and offensive (batting) roles.

**• Suggestions:**   
o Because they can provide strategic depth and adaptability in a variety of match scenarios, teams should give preference to players who routinely perform better than the average in both batting and bowling.  
o To optimize team performance in a variety of match situations, concentrate on developing and keeping all-round players who can balance bowling and batting.



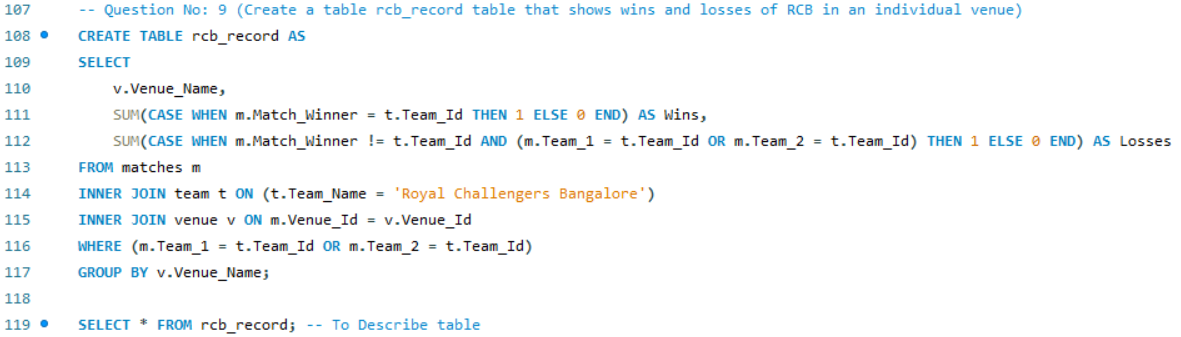
****

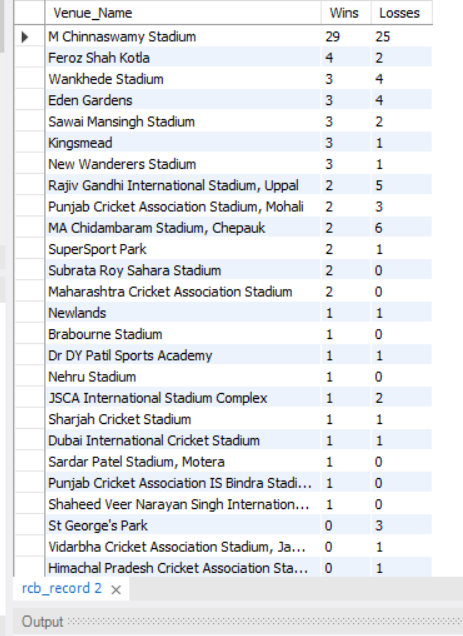
1. Create a table rcb\_record table that shows the wins and losses of RCB in an individual venue.

The following tables can be used to generate the rcb\_record table, which displays the RCB wins and losses at each venue:  
• **Tables are needed:**  
o **Matches:** To obtain match information, such as the location and outcome.   
To designate RCB and its games, as a team.   
o **venue:** To obtain information about the venue.

• **Query approach:**  
o **matches table:** Used to locate every RCB match, determine the venue, and identify the winner.   
o **Team table:** Used to make sure we're only paying attention to RCB games.   
o **Venue table:** Lists the names of the venues for every match.   
o **Wins:** Determined by adding up the instances in which RCB wins a game at each location.   
o **Losses:** Determined by keeping track of how many times RCB played at a location and lost.

**• Findings:**   
o **Home Dominance:** RCB plays 41% of their games at M Chinnaswamy Stadium, making it their most productive location.  
o **Diverse Locations:** The squad has played in a variety of locations, with no stadium accounting for more than 6% of their overall games.  
o **Spread-out Performance:** Important games are also played at venues like Feroz Shah Kotla and Wankhede Stadium, which come in second and third, respectively, with 6% and 4%.   
o **Low Impact at Many Venues:** Only 1% to 3% of RCB's games are played at a large number of stadiums, suggesting little success or exposure there.  
• **Suggestions:**   
o **Leverage Home Advantage:** Keep taking advantage of their great record at M Chinnaswamy Stadium, which stands as their highest success rate.   
o **Target Improvement at Secondary Venues:** Concentrate on methods to raise performance at locations that they regularly play at but might improve outcomes, such as Feroz Shah Kotla and Wankhede.





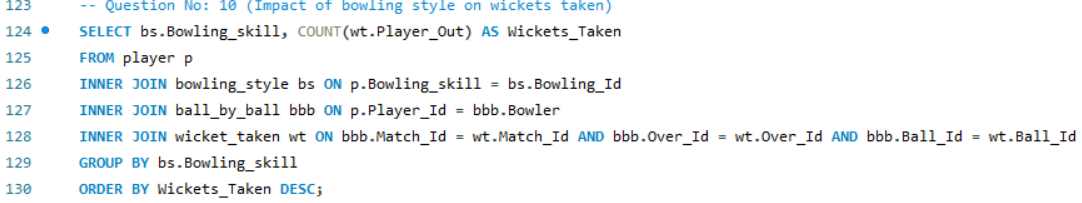
1. What is the impact of bowling style on wickets taken?

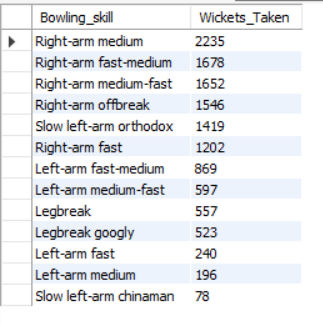
The following tables are necessary in order to analyze how bowling style affects the number of wickets taken:   
• **Player:** To learn more about the players, including how they bowl.   
• **bowling\_style:** To identify the kind of bowling style that each player is linked to.   
• **wicket\_taken:** To obtain information about each bowler's number of wickets taken.   
**• ball\_by\_ball:** To associate the bowler's wickets taken with each ball.

**Query Approach:**   
• Determine the style of each bowler.  
• Determine how many wickets each style of bowler has taken overall.  
• To evaluate the effects of various bowling techniques, compare the wicket counts.

**Insights:**  
1.With 2235 wickets taken, right-arm medium bowlers are the most successful, suggesting that this bowling style works well in a variety of match situations. By improving their medium pace alternatives for the right arm, RCB could take advantage of this.  
2. Due to their great effectiveness (1678 and 1546 wickets, respectively), right-arm fast-medium and off-break bowlers also pose a threat to hitters. Sometimes these methods are useful in crucial overs.   
3. With 1419 wickets taken by slow left-arm orthodox bowlers, spin options are still crucial, especially on slower pitches. To handle subcontinental circumstances, RCB can benefit from having flexible spin bowlers.

4. Leg break and googly bowlers are less effective at taking wickets, but they are useful in certain situations where they can take advantage of batting deficiencies, especially in the middle overs.   
**Suggestions:**   
1. To be versatile in a variety of match situations, RBC should think about a well-rounded bowling assault that prioritizes right-arm medium and fast-medium bowlers.   
2. In home circumstances, where spin is frequently more successful, and in critical overs, cultivating and incorporating quality spin bowlers—particularly left-arm orthodox spinners—can help obtain an advantage.



****

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

Based on the number of runs the team scored during the season and the number of wickets they took, the following tables are needed to determine whether the team's performance was better than it was the year before:   
• **Tables are needed:**  
o **team:** includes details on teams, such as Team\_Name and Team\_Id.  
o **player\_match:** Assigns players to teams and matches, enabling us to pair players with particular teams in particular games.   
o **matches:** Contains match-related information such as Match\_Id, Team\_1, Team\_2, and Season\_Id.   
o **batsman\_scored:** Keeps track of each match's runs scored by batsmen, identifiable by Match\_Id.   
o **wicket\_taken:** Lists, by Match\_Id, the number of wickets taken in each match.   
o **season:** Provides details about each season, such as Season\_Year and Season\_Id.

**• Query Method:**   
o Determine Team Performance by Season: o The query first determines each team's total score (TotalRuns) and total wickets (TotalWickets) for each season.  
o To do this, the team's matches are obtained by connecting the player\_match, team, and matches tables. Then, the runs from batsman\_scored are added up, and the wickets from wicket\_taken are counted.   
o Examine this year in comparison to last year:   
o To compare a team's performance over two consecutive seasons (current year vs. prior year), the query then self-joins the TeamPerformance CTE (Common Table Expression).

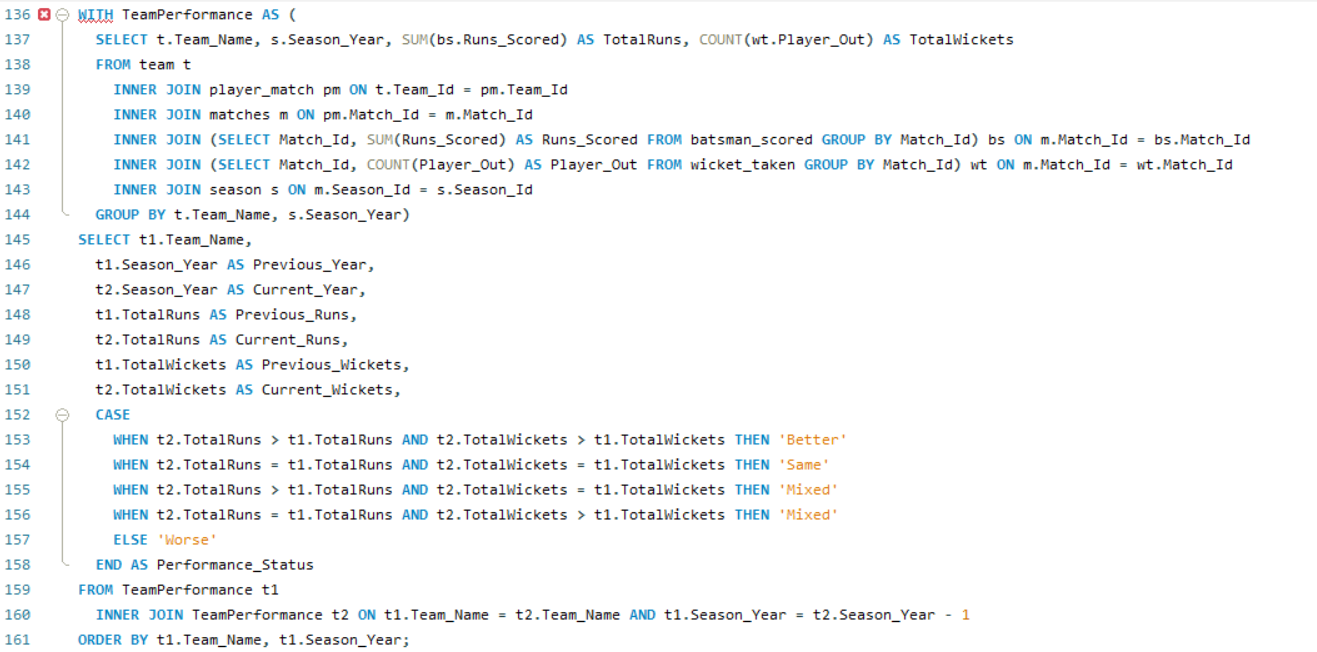
o The comparison of runs and wickets from the current and prior seasons is used to compute the Performance\_Status:   
o Better: If this year's scores and wickets taken surpass those of the previous year.  
o Same: If this year's runs scored and wickets taken are equivalent to those of the previous year.  
o Mixed: When one of the measures (wickets or runs) is higher than the previous year but the other is unchanged.   
o Worse: If this year's scores and wickets taken are lower than those of the previous year.

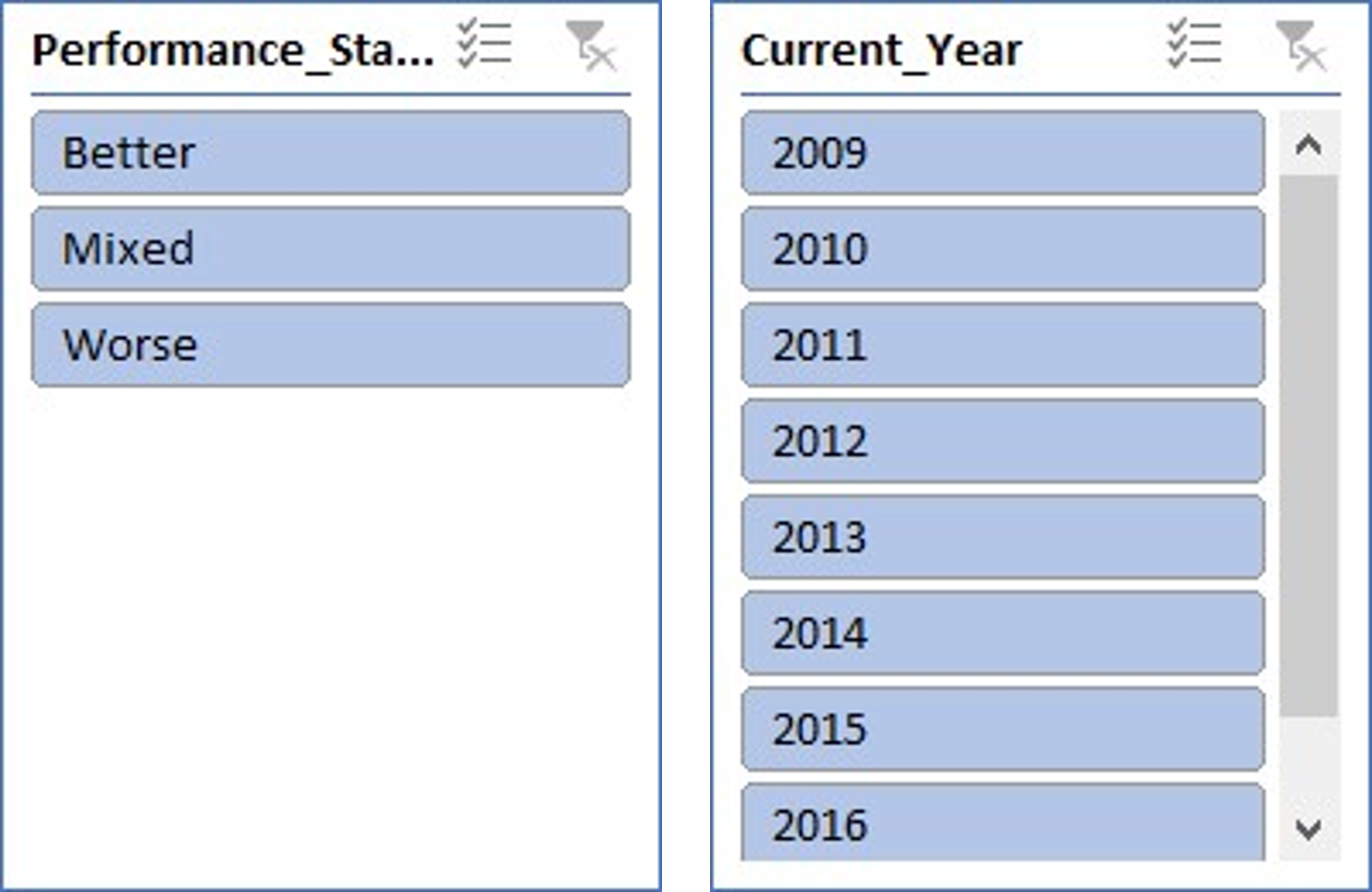
**Final Product:**   
o The query chooses the team name, the year of the current season, the year of the previous season, and the associated wickets and runs. Additionally, it gives the performance status according to the previously mentioned conditions.  
**• Perspectives:**  
o **Royal Challengers Bangalore (RCB) Displays Strong Growth:** RCB saw a notable gain in both runs scored and wickets taken, rising by 16,456 runs and 22 wickets, respectively, indicating a good upward trend in the team's overall performance.   
o **Consistent Top Performers:** Teams such as the Chennai Super Kings and Mumbai Indians exhibit steady, dependable play across the seasons with no rise in runs or wickets.

o **Delhi Daredevils Stagnation:** Although they are among the best teams in terms of runs and wickets, the Delhi Daredevils' performance may have plateaued since they have not increased their wicket total or run total.   
o **Up-and-Coming Teams:** Sunrisers Hyderabad and Pune Warriors exhibit notable percentage gains in runs and wickets, which may suggest their development and future ability to compete with more established teams.

**• Suggestions:**   
o **Pay Attention to Teams That Show Stagnation or decrease:** To buck these trends, teams like the Delhi Daredevils and Rajasthan Royals, which exhibit little to no improvement or even decrease, should reevaluate their player development, strategy, or tactics.

o **Leverage Growth for RCB and Other Improving Teams:** Teams who are improving in both categories, such as RCB, can capitalize on this momentum to improve their performance even more. They may also decide to make investments in bowling or all-rounders to keep up their upward trend.





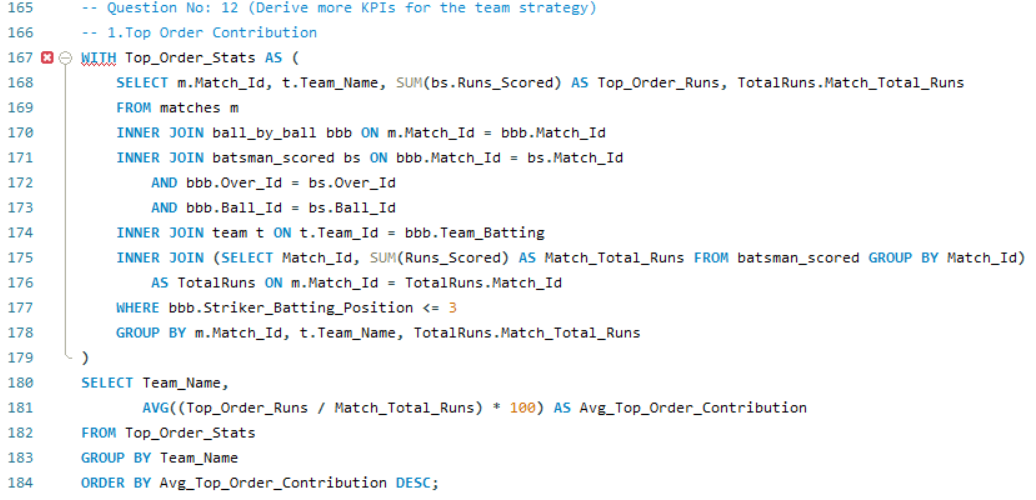
1. Can you derive more KPIs for the team strategy?

**Possible Team Strategy KPIs:**   
1. Contribution at the Top Order   
2. Boundary Frequency   
3. Performance of Powerplay   
4. Efficiency Over Death   
5. Venue-specific Win/Loss Ratio   
  
**To gain a better understanding, let's examine each one in detail:**

1. **Contribution at the Top Order   
   • Goal:**

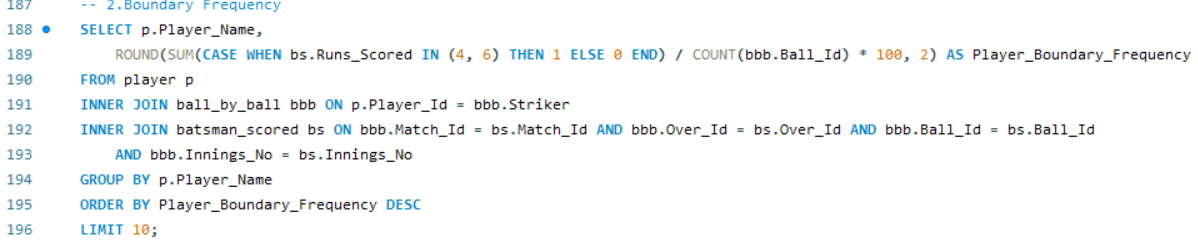
o To determine the average proportion of a team's total runs scored by its top three batsmen over all games.  
o It provides information about the relative contributions of the top three batsmen to the team's overall performance. **• Tables are needed:**o **Matches:** To recognize every match and connect it to the batting display.   
Ball by ball: To record the players' batting order and correlate it with the number of runs scored.   
o **batsman\_scored:** To obtain the precise number of runs that each batsman has scored on each ball.   
o **team:** To link the performance to a particular group.

**• Method:**   
 o **Find Matches & Teams:** To obtain match and team information, use the matches and team tables.  
 o **Sort Batsmen in Order:** To concentrate on the players in the top three batting positions, use ball\_by\_ball.  
 o **Determine Runs:** Using the batsman\_scored table, add up the runs that these top-order batsmen have scored.   
 o **Aggregate Data:** Determine the proportion of each match's total team runs that came from the top three batsmen.   
 o **Calculate Averages:** Lastly, take the average of these percentages for each team over all of their games.



**2. Boundary Frequency**  
• **Goal:** To evaluate the team's attacking play style, find out how frequently the team strikes boundaries (fours and sixes) per 100 balls.  
• **Tables are needed:**  
o **player:** Holds player data, such as Player\_Name and Player\_Id.   
o **ball\_by\_ball:** Keeps track of every ball bowled, including Team\_Batting and Striker (the player confronting the ball).   
o **Batsman\_scored:** records the number of runs scored on each ball, along with whether a boundary (4 or 6 runs) was reached.

* **Approach:**
* **Count Boundaries:** Calculate the total number of boundaries (fours and sixes) scored by each player.
* **Total Balls Faced:** Count the total number of balls each player faced.
* **Boundary Frequency Calculation:** Determine the boundary frequency by dividing the number of boundaries by the total balls faced, then multiply by 100.
* **Ranking:** Rank the players based on their boundary frequency and select the top 10.

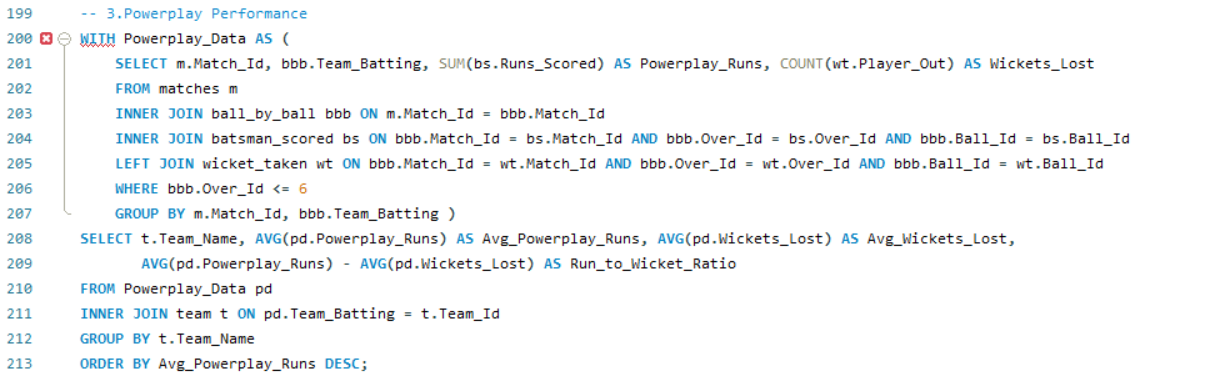


**3.Powerplay Performance**

* **Objective:**
* Teams with a high Run-to-Wicket ratio are not only scoring well but also managing to preserve their wickets, which is a sign of a well-rounded batting strategy.
* Teams with a low Run-to-Wicket ratio might need to reconsider their approach in the Powerplay, either by shuffling the batting order or by adopting a more conservative strategy initially.
* **Tables Required:**

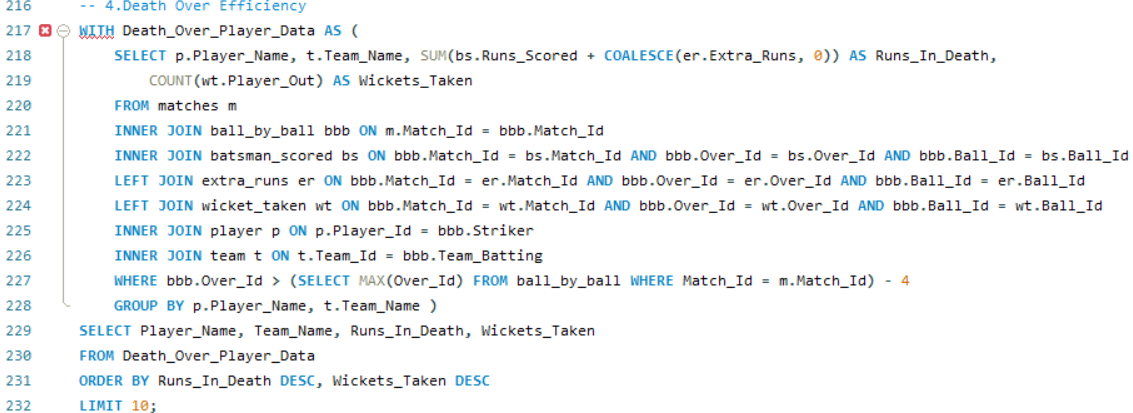
1. **matches**: Links the match details.
2. **ball\_by\_ball**: Retrieves balls bowled during the Powerplay overs (1-6).
3. **batsman\_scored**: Retrieves runs scored during the Powerplay.
4. **wicket\_taken**: Identifies wickets lost during the Powerplay.
5. **team**: Links to identify the teams by their IDs.

* **Approach:**
* **Calculate Powerplay Performance:**
  + First, calculate the total runs scored and wickets lost during the Powerplay (first 6 overs) for each match and team.
* **Aggregate Performance:**
  + Then, calculate the average Powerplay runs and wickets lost per team across all matches.
* **Compare Teams:**
  + Finally, compare the teams based on their average performance in the Powerplay, focusing on the balance between runs scored and wickets lost.



**4**. **Efficiency Over Death**

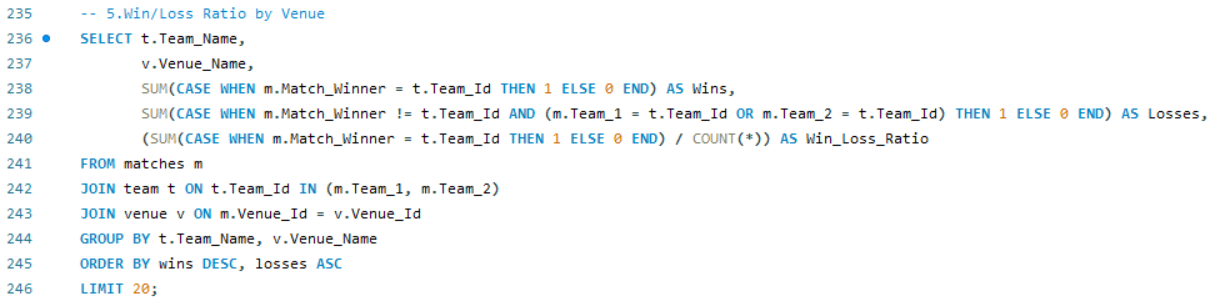
* • **Goal:** The top ten players according to the overall number of runs scored and wickets claimed in the final four overs of a match.   
  • **Derived Knowledge:**  
  o Key Finishers: The team's most dependable finishers under duress are highlighted by identifying the top 10 players who score the most runs in the final four overs.  
  o **Clutch Bowlers:** These players are also highlighted since they frequently take wickets in the final overs, showing the team who they can depend on to win crucial games.   
  o **Team Contribution:** You may determine which teams have the best death-over performers by tying player performance to their teams. This gives you information about the strategies and strengths of the entire team in the final overs.   
  • **Tables are needed:**   
  **matches**: Links to the match details.
* **ball\_by\_ball**: Used to identify balls bowled in the death overs.
* **batsman\_scored**: Retrieves runs scored in the death overs.
* **extra\_runs**: Includes extra runs given during the death overs.
* **wicket\_taken**: Identifies wickets taken in the death overs.
* **player**: Links player names to their performances.
* **team**: Identifies the teams to which the players belong.
* **Approach:**
  + **Aggregate Runs and Wickets**: Sum the runs scored and count the wickets taken by each player during the last 4 overs.
  + **Rank and Filter**: Rank players based on their death-over performance and filter to get the top



**5. Venue-specific Win/Loss Ratio**

• **Derived Insights:**   
o Teams that have won the most games and lost the fewest at particular locations, indicating the locations where teams excel.  
o Some teams routinely outperform others at particular locations, suggesting a possible home-field or venue-specific edge that may be important for strategic planning.   
• **Tables are needed:**   
o **matches:** To connect the match information, such as which teams participated and the outcome.   
o **team:** To obtain the names of the teams and the IDs that go with them.   
o **Venue:** To obtain the names and specifics of the locations where games were held.

**• Method:**   
o Determine how many victories and defeats each squad has at each location.  
0 Using these counts, calculate the win/loss ratio.  
o Sort teams according to how well they performed at various locations, giving the most victories priority.



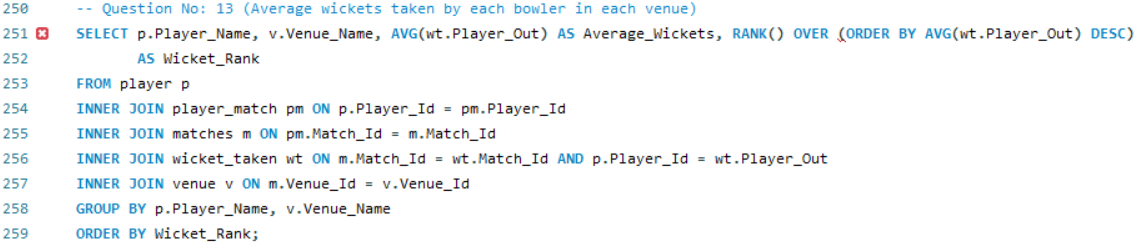
1. Using SQL, write a query to find out the average wickets taken by each bowler in each venue. Also, rank the gender according to the average value.

The following tables are needed to determine the average number of wickets taken by each bowler in each venue:   
• **Tables are needed:**  
o **player:** Use Player\_Name to obtain the bowler's name.  
o **player\_match:** To link every player to the games they participated in.   
o **Matches:** To connect with the venue and obtain match-related information.   
o **wicket\_taken:** To determine how many wickets each bowler has taken in a given game.   
o **Venue:** To find out the name of the location where the games were held.

**• Method:**

o **Determine Bowler and Venue:** To determine which bowler (player) took wickets at which location, the player, player\_match, matches, and venue tables are utilized.  
o **Determine Wickets:** The number of wickets taken by each bowler at various locations is determined using the wicket\_taken table.  
o **Determine Average:** The average number of wickets taken by each bowler at each site is determined using the AVG() method.   
o **Ranking:** The bowlers are ranked according to their average number of wickets taken using the RANK() function.

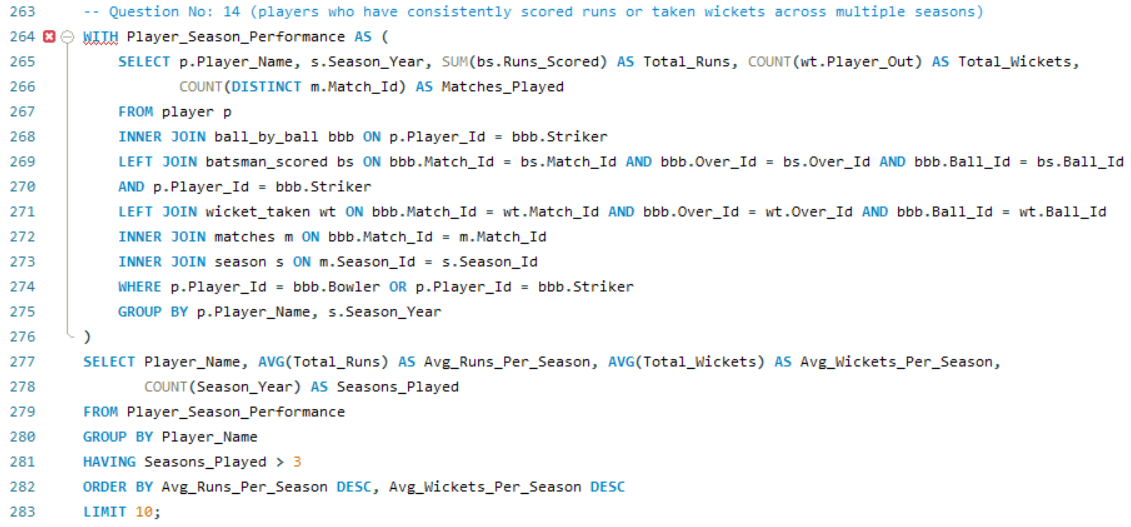
• **Acquired Valuable Knowledge:**   
o **Top Performers:** By capturing the highest average number of wickets, this query assists in finding bowlers who have continuously performed well across several venues.   
o **Venue Impact:** By examining this data, one can determine which locations are more favorable to particular bowlers because certain bowlers may perform better on average at a given location because of pitch conditions or other variables.   
o **Strategic Team Selection:** Using this data, teams can choose bowlers for upcoming games who have a track record of success at particular locations.



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

To get the players who have consistently performed well in past seasons, following tables are required:

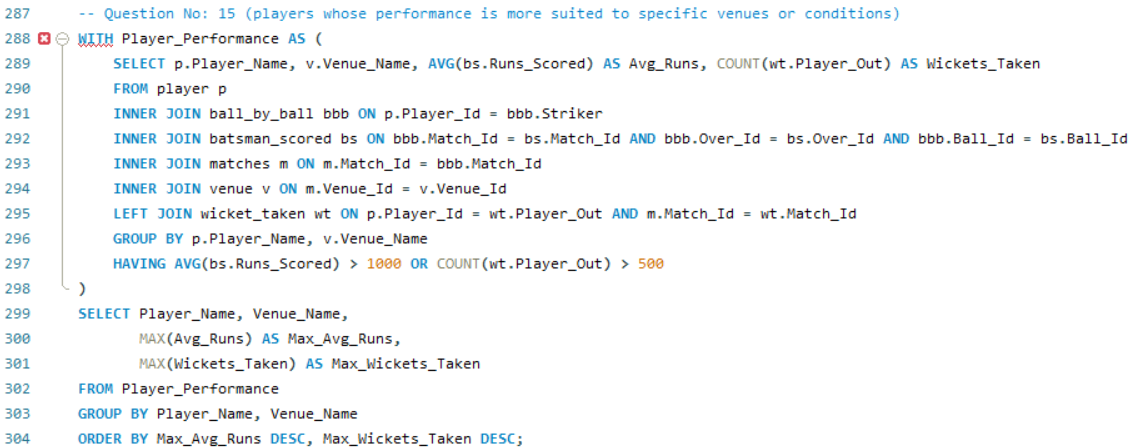
* **Tables Required:**
* **player**: To get player details.
* **player\_match**: To link players to specific matches.
* **matches**: To get match and season details.
* **season**: To identify the season year for each match.
* **ball\_by\_ball**: To track each ball, including which players were batting or bowling.
* **batsman\_scored**: To calculate the runs scored by each batsman.
* **wicket\_taken**: To determine the wickets taken by each bowler.
* **Approach:**
* **Aggregate Performance**: Calculate the total runs scored and wickets taken by each player per season.
* **Consistency Filter**: Focus on players with more than 3 seasons of participation.
* **Ranking**: Rank players based on their average runs and wickets per season to identify consistent performers.
* **Derived Insights:**
* **Consistency Over Time**: Identifies players who consistently perform well, making them valuable assets for their teams.
* **Balanced Performers**: Highlights players who contribute both with the bat and the ball across seasons.
* **Strategic Importance**: Provides insights into players who can be key in planning long-term team strategies.



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

To get players whose performance is more suited to specific venues or conditions, following tables are required:

* **Tables Required:**
* **player**: To identify player details.
* **ball\_by\_ball**: To link players with match events.
* **batsman\_scored**: To compute runs scored at venues.
* **wicket\_taken**: To count wickets taken by players at venues.
* **matches**: To connect the match details and venue information.
* **venue**: To determine venue names.
* **Query Approach:**
* Aggregate player performance at each venue, calculating the average runs scored and total wickets taken.
* Filter players who consistently score or take wickets above a certain threshold.
* Further aggregate the data to highlight top players by venue, limiting the results to the top performers.
* **Derived Insights:**
* **Top Performers by Venue:** Identify the top players at each venue who consistently score high or take wickets, making them key players for specific conditions.
* **Venue-Based Strategy:** Use the aggregated data to develop strategies for selecting players based on their performance at specific venues.
* **Player Specialization:** Understand which players excel under certain conditions, allowing for more precise team selection and game planning.

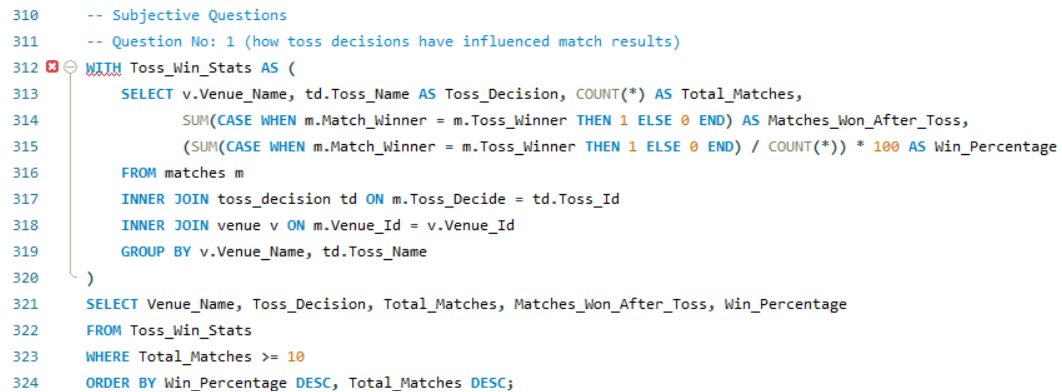


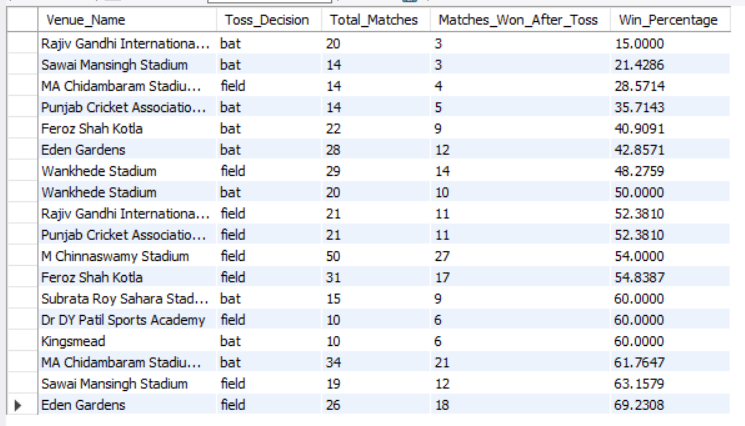
**Subjective Questions**

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?

We must examine data from three important tables in order to comprehend how toss choices have affected game outcomes.  
• **Tables Needed**  
o **matches:** To determine the winner of the toss, the winner of the match, and the connection between the location and the toss choices.   
o **toss\_decision:** To ascertain if the team that won the toss decided to bowl or bat.   
o **Venue:** To determine the match's location for venue-specific research.   
• **Method**   
o To concentrate on more pertinent insights, aggregate data to display only venues with a sizable number of matches (e.g., at least 10).   
o To find trends, figure out the win % for every throw choice made at these locations.   
o To identify the top ten most pertinent venue-toss choice combinations, sort and narrow down the results.

**• Derived Knowledge**   
o **Venue-Specific Impact:** There is a greater correlation between toss choices and match results at some venues, suggesting that these choices are important for match strategy.  
o **High Win Percentages:** At certain locations, winning the toss and choosing a particular course of action (bowling or batting) greatly raises the odds of winning.  
o **Strategic Importance:** To enhance their toss choices and match results, teams may find it useful to examine venue-specific data.   
**• Suggestions:**   
o **Data-Driven Decisions:** Teams should use venue-specific toss data to inform their choices, especially in locations where toss choices have a significant influence on game results.   
o **Venue Strategy:** To obtain a tactical edge, concentrate on locations where toss choices have impacted game outcomes in the past.

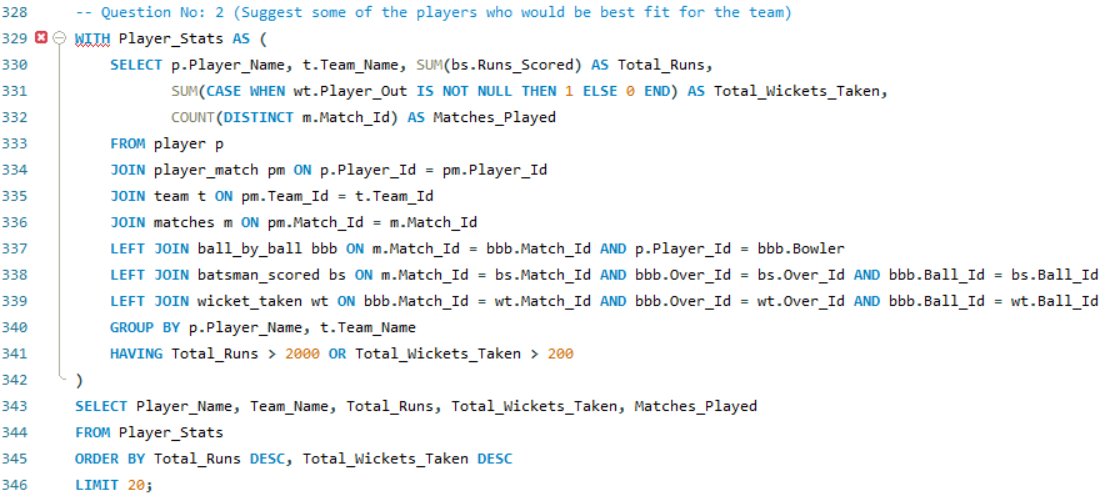




1. Suggest some of the players who would be best fit for the team.

To determine which individuals, based on their past batting and bowling performances, would be ideal additions to the team. We can identify players who often put up excellent performances by looking at player data like runs scored and wickets taken.   
• **Tables Needed**  
o **player:** To obtain player information.  
o **player\_match:** To associate players with teams and matches.   
o **team:** To obtain team details.   
o **matches:** To compile information about matches.   
o **Ball by ball:** To obtain bowler data and connect it to wickets.   
o **batsman\_scored:** To determine how many runs each player has scored.   
o **wicket\_taken:** To determine how many wickets a bowler has taken.

**• Method**   
o Determine each player's total runs scored and total wickets taken during all games.  
o Only players who have taken more than 50 wickets or scored more than 2000 runs should be included.  
o Return the top 20 players after sorting the results by total runs and total wickets taken.   
**• Derived Knowledge**   
o Players that score a lot of runs or take a lot of wickets are probably reliable players.   
o All-round players who can contribute in a variety of ways are identified with the use of batting and bowling statistics.   
o A targeted view of the best players is provided by filtering by run and wicket thresholds.   
**• Scenario-based recommendations**   
o To improve team balance, take into account hiring players that have good bowling and batting statistics.   
o Pay attention to players who have consistently performed across multiple matches, indicating reliability.

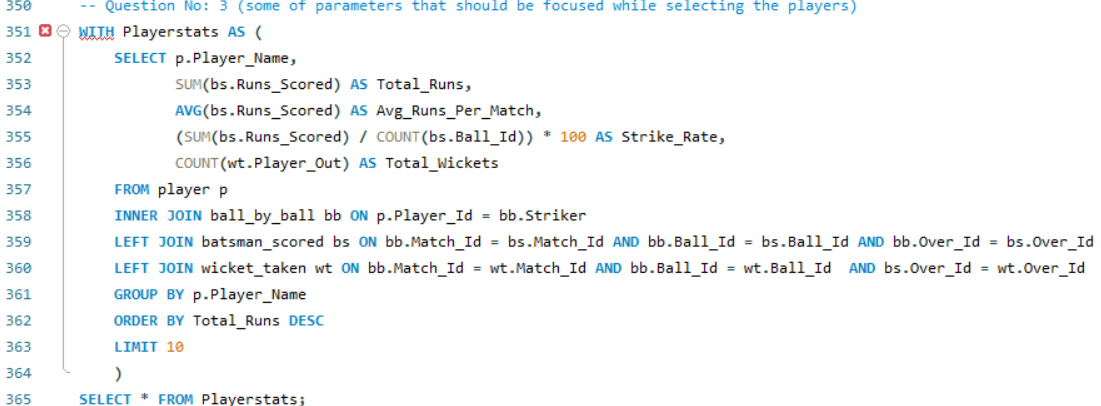


1. What are some of the parameters that should be focused on while selecting the players?

This query seeks to discover top-performing players by examining their bowling and batting statistics in order to concentrate on the important factors while choosing the best players. This comprises strike rate, total wickets taken, average runs per game, and total runs scored. The objective is to choose players who are excellent bowlers and batsmen who will make a major contribution to the team's performance.   
**• Tables Needed:**  
o **player:** To obtain the name and special identification of the player.  
**o ball\_by\_ball:** To monitor the results of each ball and determine which players are engaged.   
o **batsman\_scored:** To determine how many runs each player has scored.   
o **wicket\_taken:** To keep tabs on each player's number of wickets.

**• Method:**   
o **Player Details:** The player names are found using the player table.  
o **Ball-by-Ball Data:** Connect each ball event to a player by joining the ball\_by\_ball table.  
o **Batsman Runs:** Retrieve each player's runs using the batsman\_scored table.   
o **Wickets:** Determine how many wickets each player (bowler) has taken using the wicket\_taken table.   
o **Calculation:** Using total runs, determine the top 10 players' strike rate and average runs per game.   
o **Final Output:** Get the top 10 players by total runs and limit them.

**•** **Acquired Valuable Knowledge:**   
o **Top Run-Scorers:** Indicates which players have scored the most runs overall in the IPL, giving an indication of how well they batted.  
o **All-Rounders:** Important all-rounders who can take a lot of wickets and score a lot of runs are crucial for maintaining team balance.  
o **Impact of Strike Rate:** In limited-overs matches, especially in high-pressure overs, players with higher strike rates have a bigger impact.   
• **Suggestions:**   
o **Pay Attention to High Performers:** Players who regularly score runs and have high strike rates should be given priority because they are essential to winning games.   
o **All-Rounder Strategy:** Choose all-round players who make a substantial contribution with the bat and the ball to provide the team more compositional options.

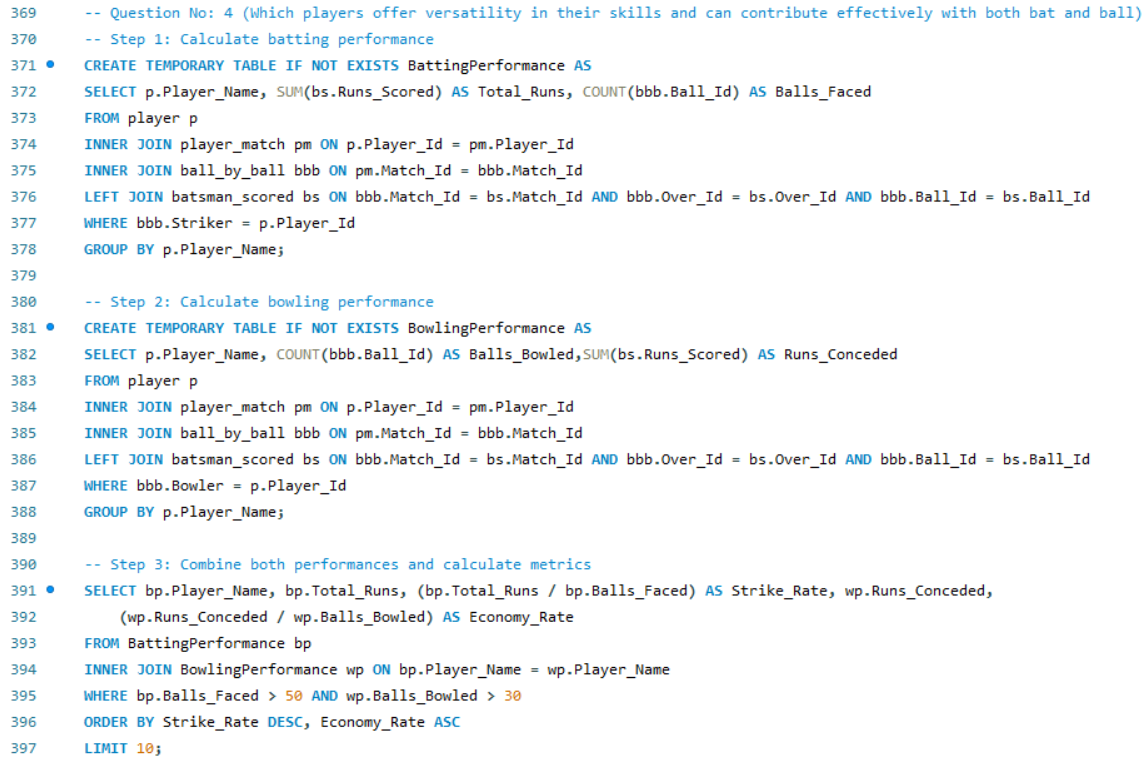




1. Which players offer versatility in their skills and can contribute effectively with both bat and ball? (can you visualize the data for the same)

The goal of this analysis is to find players who are excellent all-arounders who can make valuable contributions with the bat and the ball. We can identify versatile players who provide value in both batting and bowling by computing critical metrics for these performances; the following tables are needed for this:   
• **Tables Required:**  
**player:** includes player data, such as player\_name and player\_id.   
**player\_match:** Connects players to particular games.   
Details on balls bowled, batters faced, and wickets taken are provided by **ball\_by\_ball.**   
**batsman\_scored**: Keeps track of the batsman's runs.   
**• Method:**   
o **Step 1:** For each player who has been a striker, we add up the runs scored and count the balls faced to determine the batting performance.   
o **Step 2:** We count the number of balls bowled to determine the bowling performance and summing the runs conceded for each bowler.  
o **Step 3:** We combine the batting and bowling statistics for players who have faced more than 50 balls and bowled more than 30 balls, and calculate their strike rate and economy rate.

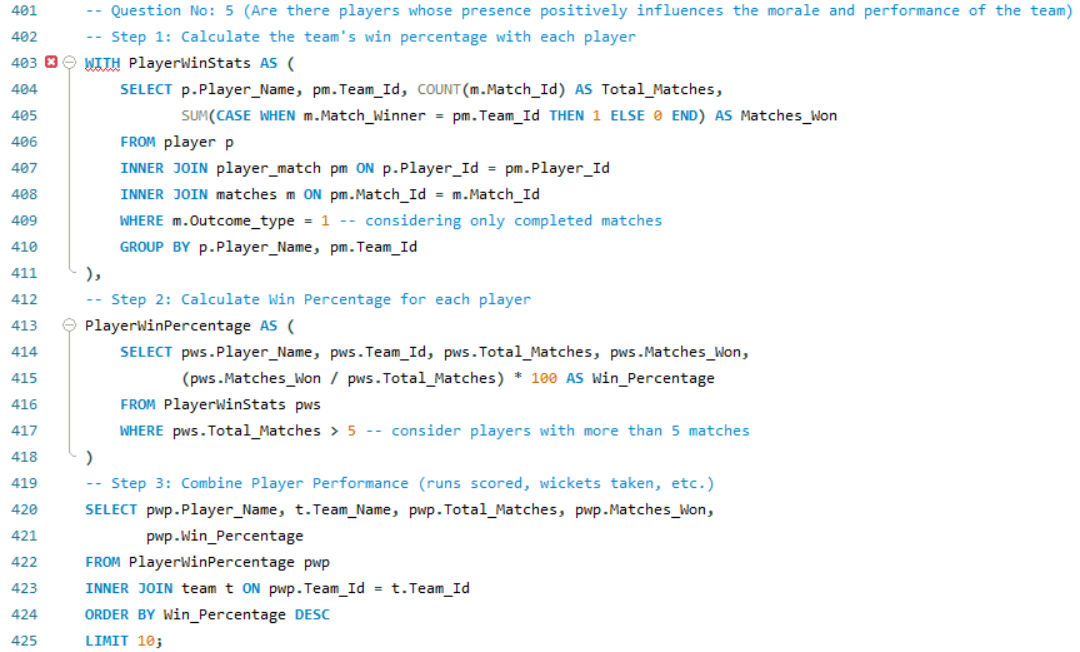
**• Derived Insights:**   
o This method finds players who are valuable all-rounders because they have contributed with the ball (low economy rates) and the bat (high strike rates).  
o We guarantee that the analysis is founded on consistent performances by concentrating on those with enough data points (balls faced/bowled).  
o The top 10 players that emerged can be readily seen in a graphic that compares their bowling and batting effectiveness.   
**• Suggestions:**   
o When choosing a balanced squad, these players ought to be given priority in order to guarantee both potent bowling and deep batting depth.   
o Make smart use of this information to place these individuals in the lineup for optimum impact as dependable bowlers and middle-order batsmen.

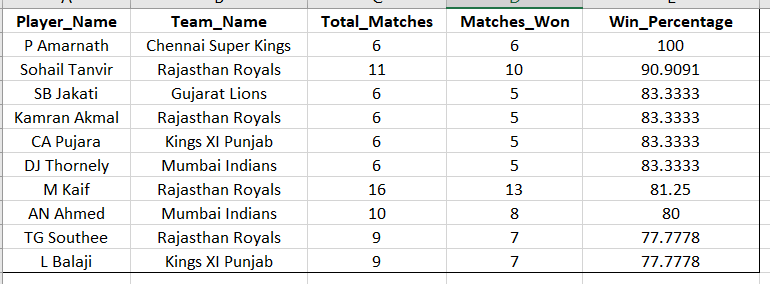


1. Are there players whose presence positively influences the morale and performance of the team? (justify your answer using visualization)

Through consistent 'Man of the Match' honors, this analysis seeks to identify players that have a good impact on team morale and performance. We identify players who make a substantial contribution to their team's performance by looking at their impact %.   
**• Tables Needed:**  
o **Player:** to obtain the ID and name of the gamer.  
o **matches:** To determine how many "Man of the Match" honors have been given out.   
o **player\_match:** To determine how many matches each player has played overall.   
**• Method:**   
o **Step 1:** Find the players who were named "Man of the Match" by joining them with matches.   
o **Step 2:** To determine how many matches each player has played, join with player\_match.   
o **Step 3:** Determine the impact % by dividing the total number of matches by the number of MoM awards.   
o **Step 4:** Arrange the outcomes by impact percentage to highlight players who have a positive influence.

**• Derived Insights:**   
o Players that consistently win "Man of the Match" honors, like as MS Dhoni and AB de Villiers, demonstrate their significant impact on team success.  
o When playing successfully, players with a greater Impact Percentage indicate they greatly increase team morale and success.  
o Restricting consideration to players who have received more than five MoM awards guarantees that only consistently significant performers are taken into account.   
**• Suggestions:**   
o Players having a high Impact Percentage should be taken into account as a means of enhancing team spirit during crucial games.   
o For optimum impact, use this data when choosing captains or key players in high-stakes matches.





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

To improve team balance before the IPL mega auction, RCB should concentrate on the following areas:   
o **Strengthen the Middle Order:** To provide stability and depth, get a dependable middle-order batsman.  
o **All-rounders:** To improve team balance, look for adaptable players who can contribute with the bat and the ball.  
o **Death-over Experts:** In the final overs, give preference to bowlers who are adept at yorkers and variations.   
o **Spin Reinforcement:** To manage the middle overs on pitches that are conducive to spin, acquire top-notch spinners.   
o **Invest in Up-and-Coming Talent:** To guarantee long-term team stability, concentrate on young, uncapped players with promise.

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

There are a number of contributing variables to high-scoring matches in the IPL and other T20 formats. These elements have a big influence on team tactics and spectatorship in addition to the on-field action.

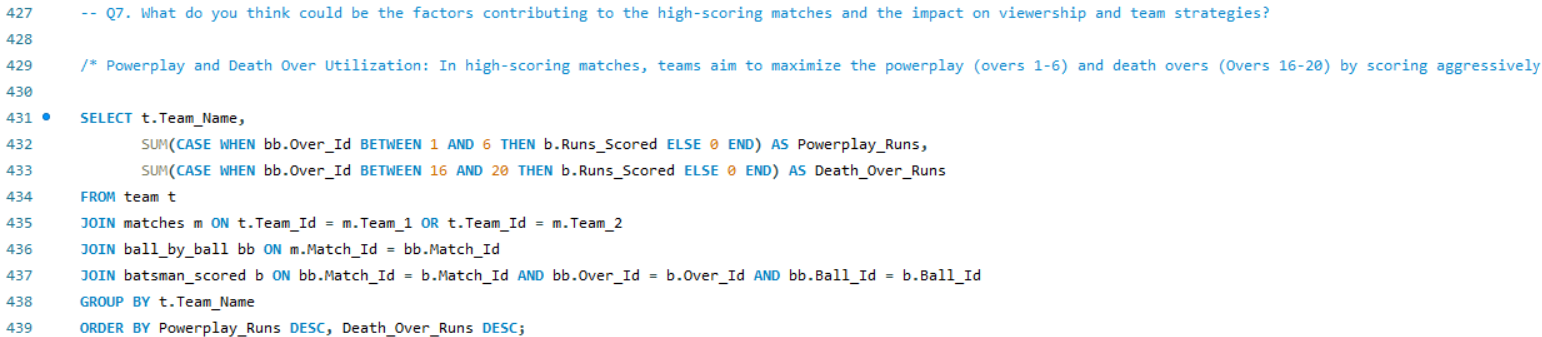
**High-scoring matches are caused by the following factors:**

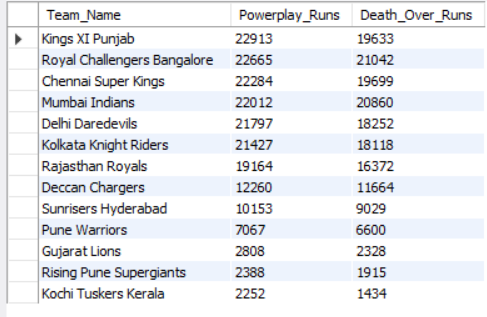
1. **Batting-Friendly Pitches:** A lot of T20 stadiums, particularly in India, have hard level fields that are ideal for batting. These pitches let batsmen play across the line and score runs rapidly while providing little support to bowlers, particularly pacers. As a result, high-scoring games are common.  
2. **Shorter Boundaries:** Batsmen may hit sixes and fours more easily at some venues because of their comparatively shorter boundaries. Even ill-timed shots frequently clear the ropes in smaller playing fields, increasing the number of runs scored.

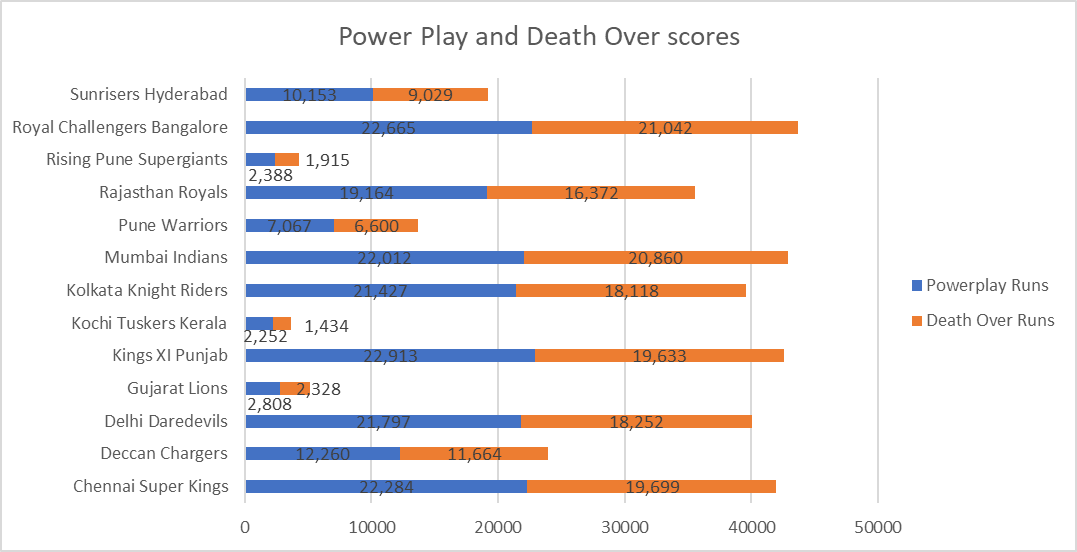
3. **Powerplay Overs and the Free Hit Rule:** Batsmen can attempt bold shots with little risk because of the fielding limits in the first six overs. The run rate is also increased by the free hit rule for no-balls, which allows batsmen to swing more freely without worrying about being dismissed.   
4. **Development of Batting Skills and Technology:** Contemporary cricket players are increasingly skilled in power hitting methods, and many have made investments to create 360-degree shots that defy conventional bowling, such as reverse sweeps and scoops. Furthermore, lighter, more powerful bats are now available thanks to developments in bat technology, which facilitates rapid scoring.

**Impact on Viewership:**   
1. **Higher Entertainment Value:** Because of the fast-paced action and frequent boundary-hitting, high-scoring contests typically draw larger audiences. Fans remain riveted to their displays until the very end because they enjoy watching thrilling chases and large scores. Higher television ratings and increased social media interaction are the results of this excitement spike.  
2. **Global Appeal:** T20 cricket has attracted a global following, especially in the IPL. Not just conventional cricket fans are drawn to the excitement of witnessing high totals and thrilling finishes. Thus, high-scoring games aid cricket in expanding its global fan base and market reach.

**Effect on Group Techniques:**   
1. **Aggressive Batting Strategy:** Rather than playing conservatively, teams are increasingly taking an aggressive batting strategy, striving for larger scores. As a result, power hitters who can increase the scoring rate even under duress are now the main focus.  
2. **Death-Bowling Specialization:** Teams must concentrate on death-over specialists—bowlers who can control the run rate in the last overs—because high-scoring games have made this crucial. These days, the secret to surviving the power-hitting assault is to use variations like yorkers, slower balls, and wide-angle deliveries.



****

****

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

We first identify teams playing at particular venues, tally their games and victories, then compute their win % at those locations in order to examine the effect of home field advantage on team performance. This informs performance-enhancing tactics and aids in evaluating how well teams use home circumstances. The tables needed for this study are listed below:   
**• Tables Needed**    
o **matches:** keeps track of match details, such as teams, location, match and toss winners.  
o **team:** Names the teams (Team\_1, Team\_2, and winners) that are playing in each match.   
o **venue:** Helps locate home grounds by storing match venue information.

**• Method**   
o **Home Team and Venue Identification**: As a stand-in for home ground presence, matches are sorted according to which teams are playing at particular locations and if the team won the toss.  
o **Match and Win Count:** Determine how many games each team has played at each location and how many of those games they have won.  
o **Win Percentage Calculation:** Determine each team's win percentage at each venue.   
o **Findings:** Sort the results by victory % in descending order and display those where the win percentage is not zero.

• **Derived Knowledge**   
o **Win Percentage as a Performance Indicator:** Teams who have a high home win percentage probably make good use of their home field.  
o **Significance of Toss:** For many teams, winning the toss and playing at home increases their chances of winning.  
o **RCB's Home Ground Performance:** RCB can evaluate its own home ground performance and pinpoint areas for development by contrasting it with those of other teams.

• **Suggestions**   
o **Improve Home Ground Tactics:** RCB should plan to perform better at home, with an emphasis on better adjusting to local conditions.  
o **Pay Attention to Toss Strategy:** Especially at home venues, better toss choices and strategies can result in more victories.

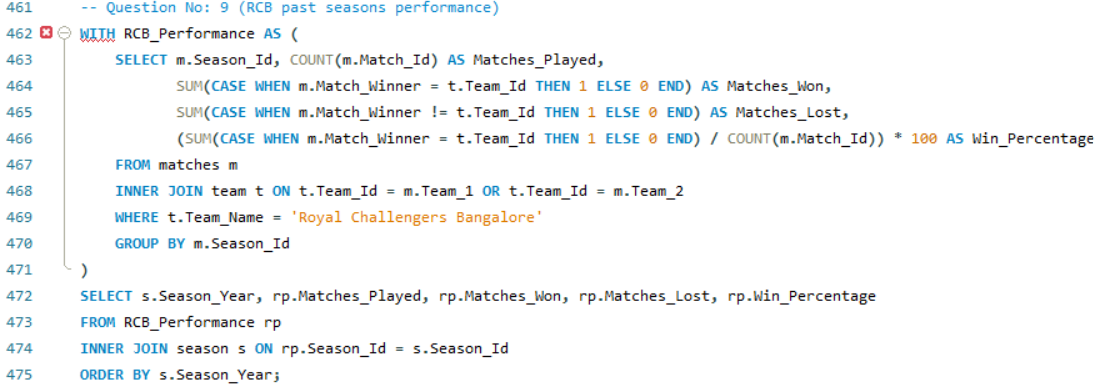


1. Come up with a visual and analytical analysis of the RCB's past season's performance and potential reasons for them not winning a trophy.

Royal Challengers Bangalore's past performance is examined in the following analysis, which focuses on the team's match results, win %, and important elements that might have affected their chances of taking home a title. The tables needed for the analysis are shown below.  
• **Necessary Tables**  
o **matches:** For information about teams, seasons, and winners.   
o **team:** To recognize RCB throughout games.   
o **season:** For performance data by season (year).   
• **Method**   
o **RCB filter Matches:** Extract all of the RCB matches.   
o **Determine Wins and Losses:** Compute the number of games played, victories, and defeats for every season.   
o **Win Percentage Calculation:** Determine each season's win percentage.   
o **Mapping the seasons:** To see results by season, join the season table with the performance data.

• **Derived Knowledge**   
o **Inconsistent Performance:** RCB has erratic play, with successful seasons interspersed with poorer ones.  
o **Close Matches:** Their momentum in crucial phases has been harmed by losing a number of close matches.  
o **High victory Percentage Not Converted:** They do poorly in knockout matches even though they have a strong victory rate in some seasons.   
**• Possible explanations for not taking home a trophy**   
o **Unbalanced Team:** excessive dependence on top-order batting at the expense of inconsistent middle-order and bowlers.   
o **Bowling Deficiencies:** RCB has trouble defending scores, especially in the final overs.   
o **Knockout Pressure:** In elimination and championship matches, they have frequently stumbled under duress.   
o **Injury and Squad Management:** The balance of the team has been impacted by the injuries to key players and dubious squad rotations.

**• Suggestions**   
o **Boost the bowling unit:** by concentrating on developing death-over experts.  
o **Increase Squad Depth:** The impact of injuries can be lessened by having a well-rounded squad with dependable bench strength.  
o **Enhance Close-Game Strategy:** Gaining momentum in the league stages can be achieved by concentrating on winning close games.



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

In the absence of the objective and subjective questions, I would analyze the IPL dataset using the following methodology, especially for a team like RCB:   
1. **Recognize the dataset:**  
• Start by carefully going over the IPL dataset to comprehend its organization, the tables that are accessible, and important columns like players, matches, teams, venues, and performances.  
• Determine key performance indicators for the team, including wins, losses, batting, bowling, and fielding statistics for each player, and venue-specific data.

2. **Specify the Analytical Objectives:**  
In the absence of predetermined queries, I would concentrate on producing insights in multiple areas:  
• **Team Performance Trends:** Examine the win/loss ratios, home/away performance, and tournament advancement of the team as a whole over the course of several seasons.  
• **Player Contributions:** Examine each player's contributions separately to find standout players, top run scorers, wicket-takers, and players with all-around skills.   
• **Player Performance Under Pressure:** Examine how players perform in high-pressure scenarios, death overs, and knockout matches.   
• **Team Weaknesses and Strengths:** Determine the team's enduring advantages (such batting power) and disadvantages (like middle-order stability, bowling, and fielding).   
• **Results and Patterns of Matches:** Keep an eye out for reoccurring trends in match outcomes, such as winning the toss, batting first, or playing at particular locations.

3. **Conduct Statistical Analysis:**

• Match Analysis: Examine statistics at the match level to determine the main elements influencing victories and defeats (e.g., toss decisions, chasing records, target scores).  
**• Player-Level Analysis:** Examine all-rounder efficacy, bowling economy, strike rates, and batting averages. Players who regularly contribute in both the batting and bowling departments should be highlighted.  
Examine home ground advantage and RCB's performance at home in comparison to away games.   
4. **Visualize Insights:**

• Produce charts and graphs using data visualization software (such as Excel, Power BI, or Tableau):   
o Seasonal tendencies in wins and losses.   
o Player performances (best bowlers, top scoring, etc.).   
o Comparisons between performance at home and away.   
o The effect of throwing on game results.

5. **Create Recommendations:**   
Provide practical suggestions based on the study:  
• The makeup of the squad and possible areas for development.  
• Techniques for triumphing in bouts with significant stakes.   
• Acquiring and keeping players for a well-rounded squad.   
  
Without depending on pre-formulated questions, this approach would produce a thorough study that would provide insights into RCB's overall performance and possible enhancements.

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

The matches database does not contain a column called Opponent\_Team. The query to substitute "Delhi\_Capitals" with "Delhi\_Daredevils" might resemble this if such a column existed:

