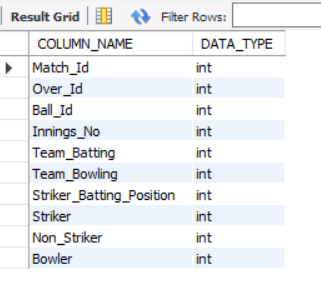
**Objective Questions**

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

Ans:

Data types: Data types are classifications of data that tell the compiler or interpreter how the programmer intends to use the data.

OUTPUT:



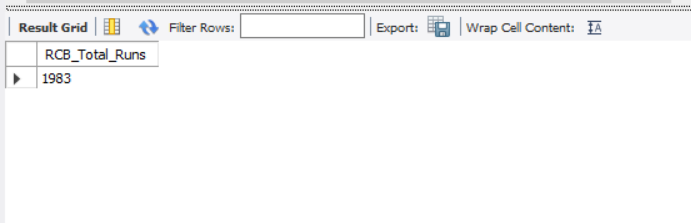
This will list all the column names and their corresponding data types for the table named 'ball\_by\_ball'. This can be useful for understanding the structure of the table, especially when preparing to write queries or perform data analysis on the table.

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

Ans: In this we calculate the total runs scored by RCB by summing up the scores (runs + extras) for each match.

* This script systematically aggregates and processes data to provide insights into RCB's performance, focusing on their scoring metrics in season 1.

OUTPUT::

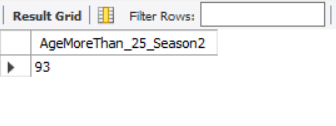


1. How many players were more than age of 25 during season 2 ?

Ans:

There are total 93 players are their whose age is more than 25 and present during season-2.

OUTPUT::



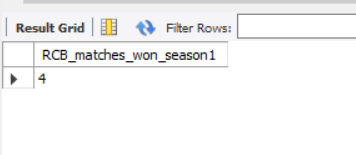
* Joins several tables (matches, player\_match, player, team, and season) to gather relevant data.
* Calculates the age of each player in Season 2 by subtracting the year of birth (DOB) from the season year.
* Filters the data to only include matches from Season 2.
* Selects unique players from the CTE who are older than 25 years.
* Counts these unique players and labels the result as AgeMoreThan\_25\_Season2.

1. How many matches did RCB win in season 1 ?

Ans :

In season one i.e 2008 RCB won only 4 matches in the tournament.

OUTPUT::

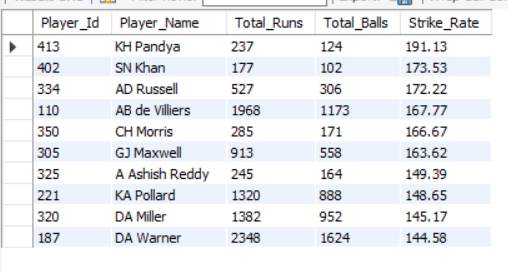


* Selects all matches from the matches table where the Season\_ID is 1.
* Filters these matches to only include those where the winning team (Match\_Winner) is the Royal Challengers Bangalore. This is done by comparing Match\_Winner with the Team\_Id of the Royal Challengers Bangalore, fetched from the team table.
* Counts the number of matches that meet the above criteria and labels the result as RCB\_season1\_won.

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

Ans:

OUTPUT::



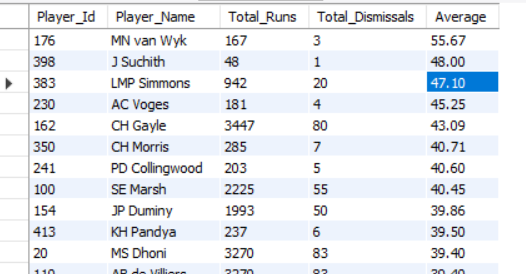
* Joins the results with the player table to get the player names.
* Filters to include only batsmen who have faced at least 100 balls.
* Calculates the strike rate as (Total\_Runs / Total\_Balls) \* 100 and orders the results by Strike\_Rate in descending order.
* Limits the results to the top 10 batsmen.

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

Ans :

For this we have to calculate (Total\_Runs/Total\_Dismiss) for all The season Then we find average runs scored by each batsman.

OUTPUT::



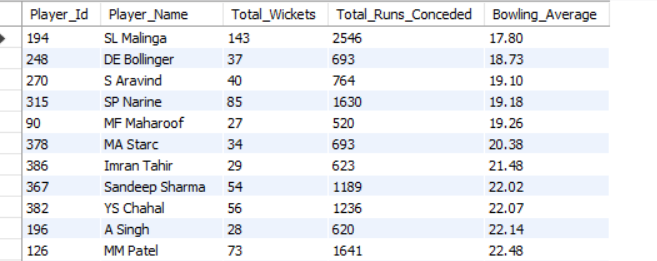
* Uses the Total\_Players\_Avg view to aggregate total runs and total dismissals for each player across all seasons.
* Calculates the batting average for each player as Total\_Runs / Total\_Dismissals.
* Joins with the player table to get player names.
* Selects and orders the players by batting average in descending order.

This creates a view that calculates the total runs, total dismissals, and batting average for each player across all IPL seasons and then queries this view to display the top players by batting average.

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

Ans :

OUTPUT::



This SQL script creates two views, highest\_wickets\_taken and runs\_conceded, and then uses these views to calculate the bowling average for bowlers with at least 25 wickets.

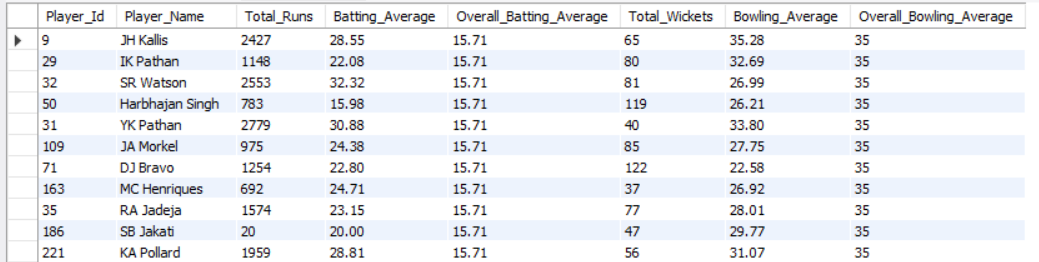
* Joins the highest\_wickets\_taken and runs\_conceded views to calculate the bowling average as Total\_Runs\_Conceded / Total\_Wickets.
* Filters the results to include only bowlers with at least 25 wickets.
* Orders the results by bowling average in ascending order to show the best bowling averages first.

In essence, the script identifies bowlers with at least 25 wickets and calculates their bowling averages, listing the bowlers with the best (lowest) averages.

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

Ans:

OUTPUT::



Summarizes total runs and total dismissals for each player from the Total\_Players\_Avg view.

* Selects players whose batting average is greater than the overall batting average.
* Joins the avg\_runs\_greater\_than\_batting\_avg and wickets\_greater\_than\_bowling\_avg views on Player\_Id.
* Selects and displays player details, including player ID, player name, total runs, batting average, overall batting average, total wickets, bowling average, and overall bowling average, for players who meet both criteria.

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

Ans:

OUTPUT::



This view provides a summary table showing how many matches RCB has won and lost at each venue they have played in.

This SQL script creates a view named rcb\_record\_table that summarizes the win-loss record of the Royal Challengers Bangalore (RCB) at each venue.

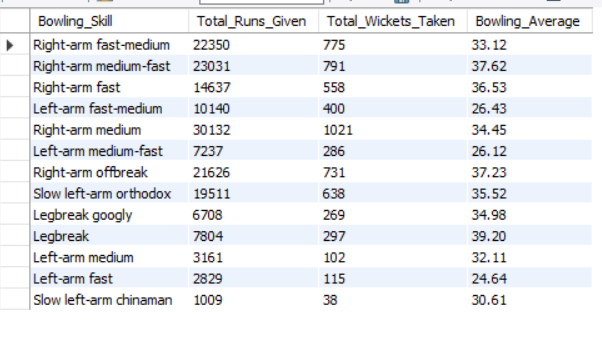
🡪Counts the number of matches won by RCB at each venue.

* Subquery a selects matches where RCB (Team\_Id = 2) played and won
* Selects all columns from the rcb\_record\_table view to display the win-loss record of RCB at each venue.

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

Ans:

OUTPUT::



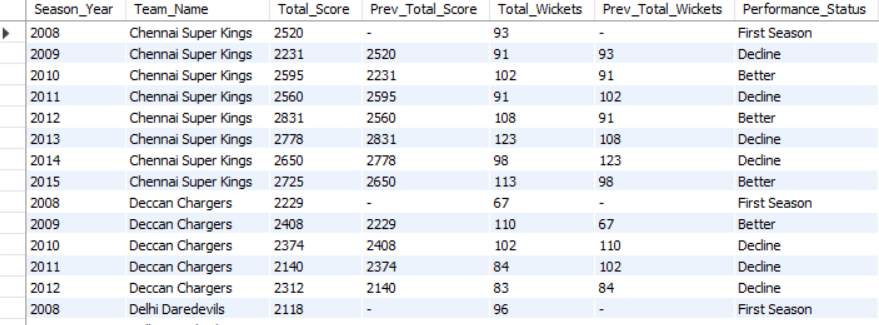
This query summarizes the bowling performance based on different bowling skills, showing the total runs conceded, total wickets taken, and the average bowling average for each bowling skill category.

* Selects columns from the CTE (cte) to display Bowling\_Skill, Total\_Runs\_Given (sum of Total\_Runs\_Conceded), Total\_Wickets\_Taken (sum of Total\_Wickets), and Bowling\_Average (average of Average).
* Groups the results by Bowling\_Skill.
* Rounds the Bowling\_Average to two decimal places using round(avg(Average), 2).

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

Ans:

OUTPUT::



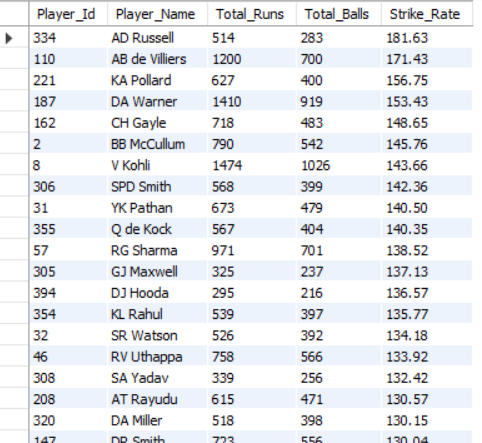
* Combines the first and second innings scores to calculate the total score for each team per season.
* Outputs the season year, team name, total score, previous total score, total wickets, previous total wickets, and performance status.
* Performance status is categorized as:
  + "First Season" for the team's first season.
  + "Better" if both total score and total wickets are higher than the previous season.
  + "Decline" otherwise.

This script tracks the performance of teams across seasons, evaluating whether they performed better or worse compared to the previous season based on their total scores and wickets.

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

Ans:

OUTPUT::



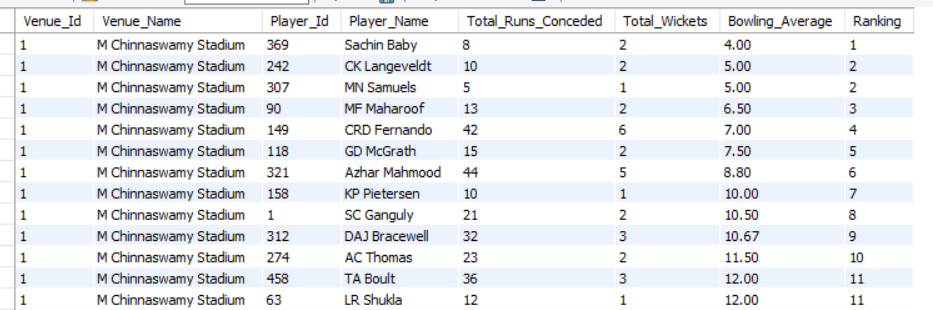
The top 20 players with the highest strike rates from the last two seasons, provided they faced at least 200 balls.

* Collects detailed ball-by-ball data, including match and season information.
* Joins ball\_by\_ball, batsman\_scored, matches, and season tables to gather necessary details such as runs scored per ball, striker, match ID, season ID, and year.
* Aggregates data to calculate the total runs scored (Total\_Runs) and the number of balls faced (Total\_Balls) for each striker across the filtered seasons.
* **Total\_Runs**: The total runs scored by the player in the last two seasons.
* **Total\_Balls**: The total number of balls faced by the player in the last two seasons.
* **Strike\_Rate**: The player's strike rate, calculated as (Total\_Runs / Total\_Balls) \* 100.

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

Ans:

OUTPUT::



calculates the bowling averages of bowlers by venue and ranks them based on their performance. It first creates two views, wickets\_taken\_by\_venue and runs\_conceded\_by\_venue, to aggregate data on wickets taken and runs conceded by bowlers at different venues. Finally, it combines these views to compute the bowling average for each bowler and assigns a rank to them at each venue.

**Bowling Average Calculation**:

* Joins wickets\_taken\_by\_venue and runs\_conceded\_by\_venue to calculate the bowling average (total runs conceded divided by total wickets taken) for each bowler at each venue.

**Ranking**:

* Uses the DENSE\_RANK function to rank bowlers at each venue based on their bowling average.

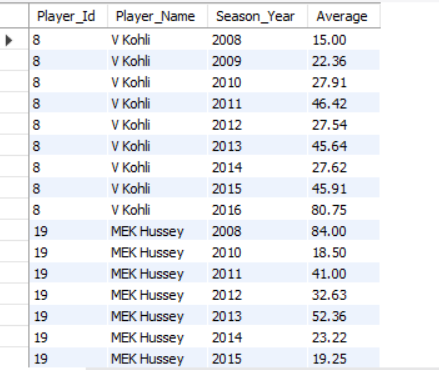
**Final Selection**:

* Orders the result by venue and rank.

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

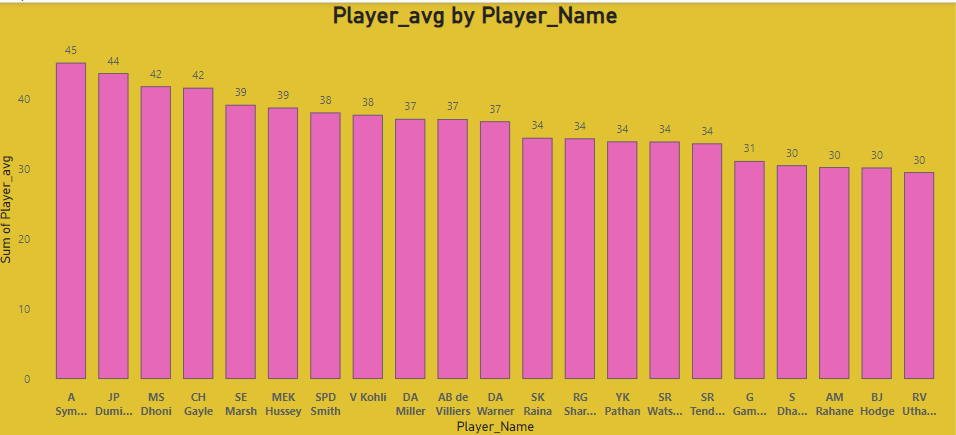
Ans:

OUTPUT::



This script calculates the batting averages of players and filters those who have a batting average above 30 in at least 4 seasons. It then retrieves the season-wise batting averages for these selected players, including their names. The final result is a list of player IDs, player names, season years, and batting averages, ordered by player ID and season year.

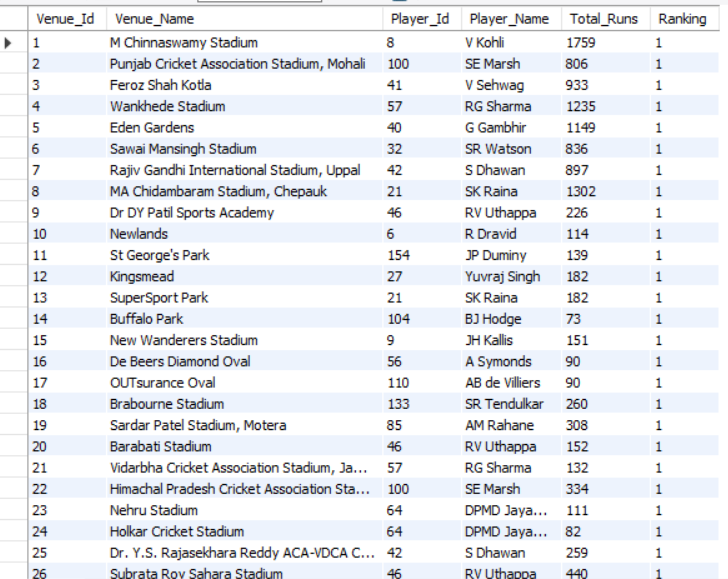
Visualisation is:



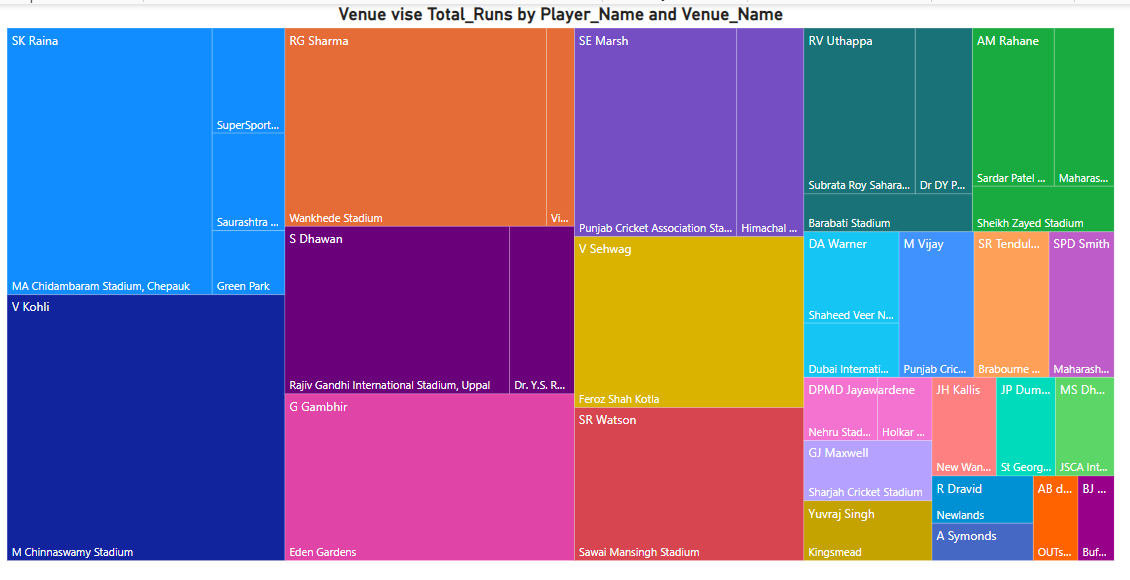
1. Are there players whose performance is more suited to specific venues or conditions? (how would you present this using charts?)

Ans:

OUTPUT::



This script identifies players who have achieved a batting average above 30 in at least 4 different seasons. It calculates the season-wise batting averages for these players and retrieves their names, displaying the data in an organized manner by player ID and season year.

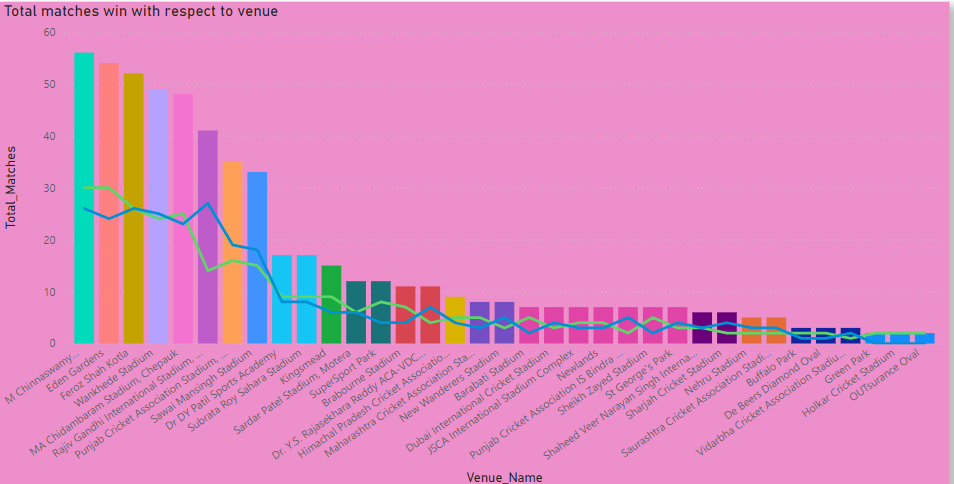


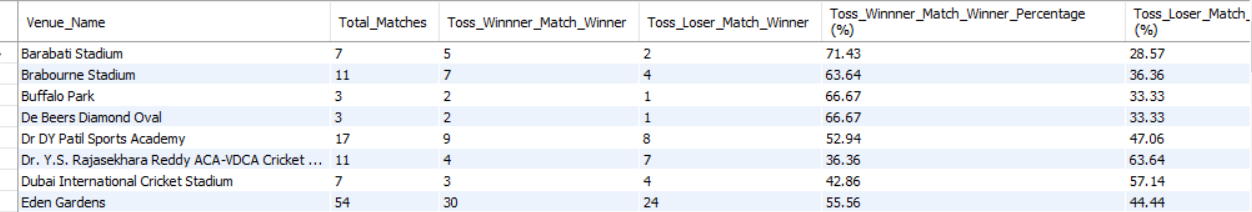
Subjective Questions

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

Ans:

OUTPUT:



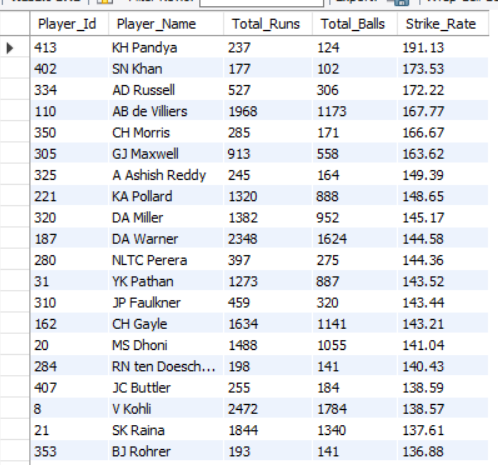


This will analyzes the impact of winning the toss on match outcomes in the IPL, both overall and by venue.

* Joins matches with toss\_decision to get match and toss details.
* Adds columns to indicate if the toss winner also won the match (Toss\_Win\_Results) or lost the match (Toss\_Loss\_Results).
* Counts total matches, matches where the toss winner won, and matches where the toss winner lost.
* Calculates the percentages of matches won by the toss winner and the toss loser.
* Joins with the venue table to get venue names.
* Displays the total matches, matches won by toss winners, matches won by toss losers, and their respective percentages for each venue, ordered by venue name.

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

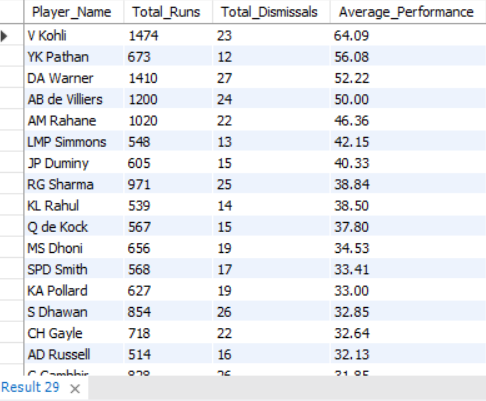
Ans: With respect to strike rate I opt this players …



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

Ans:

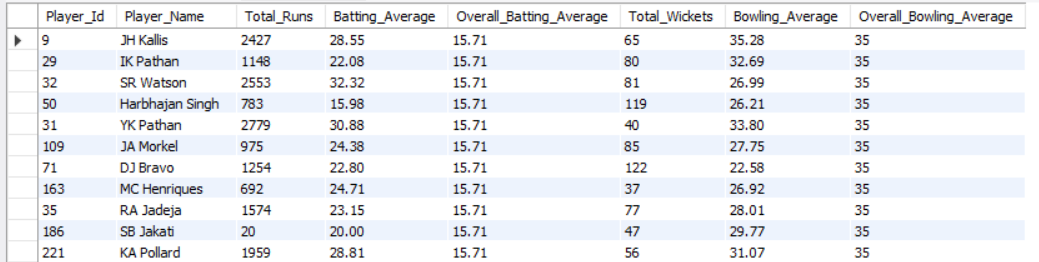
Here I take Average Performance of player in Past matches.

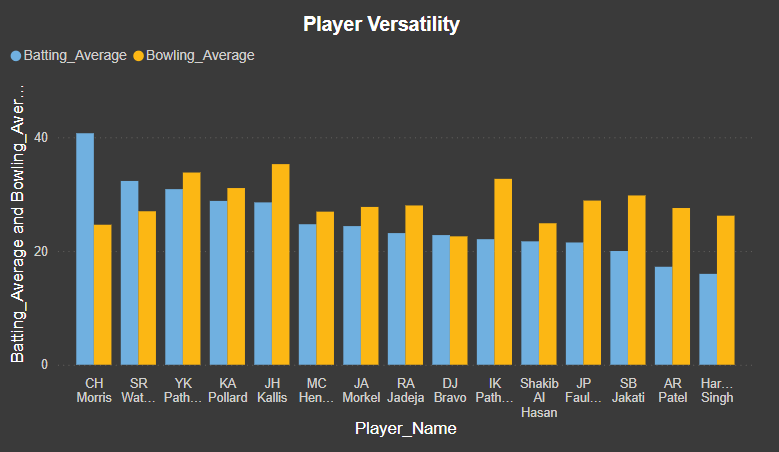


1. Which players offer versatility in their skills and can contribute effectively with both bat and ball? (can you visualize the data for the same)

Ans: The players who have average runs scored greater than overall

average and who have taken wickets greater than overall average.



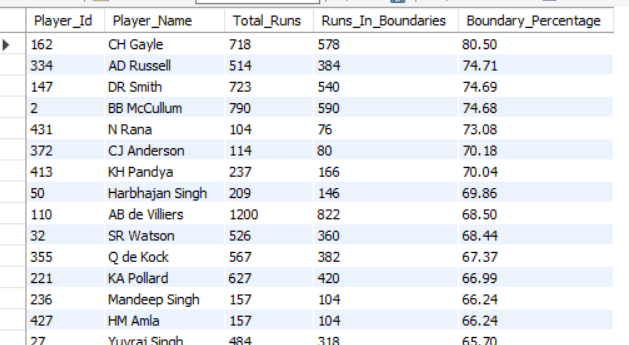


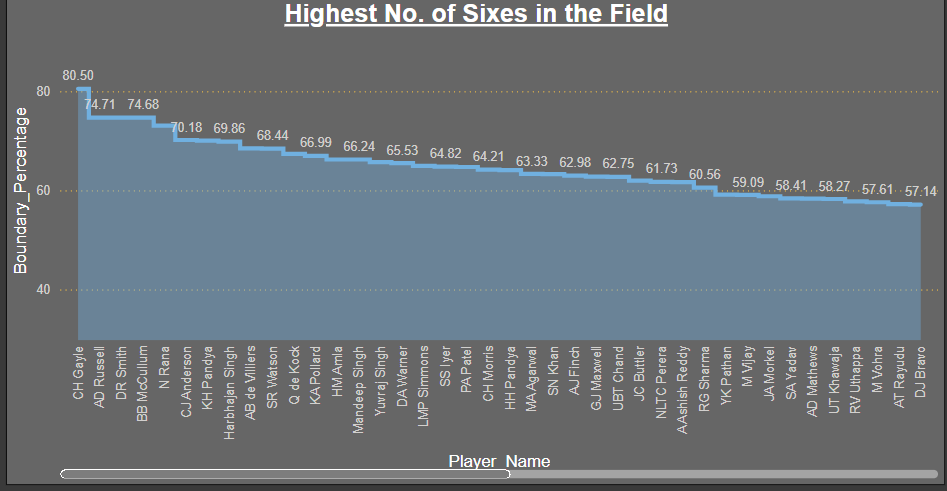
1. Are there players whose presence positively influences the morale and performance of the team? (justify your answer using visualisation)

Ans:

This SQL script calculates the percentage of runs scored from boundaries (fours and sixes) for players in the 2015 and 2016 IPL seasons, and filters players with a total run count of 100 or more. Here's a summary of each part:

It analyzes the performance of batsmen in the 2015 and 2016 IPL seasons, focusing on the percentage of their runs scored from boundaries (fours and sixes). It calculates the total runs and boundary runs for each player, filters for those who have scored 100 or more total runs, and orders the results by the percentage of runs from boundaries, in descending order. The output includes the player ID, player name, total runs, runs from boundaries, and the boundary percentage.

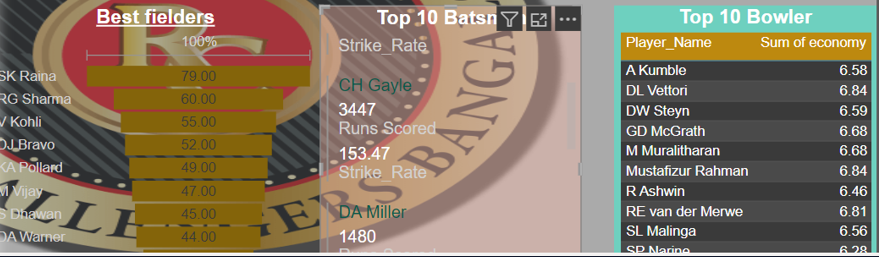




1. What would you suggest to RCB before going to mega auction ?

Ans: Before the mega auction, RCB (Royal Challengers Bangalore) should:

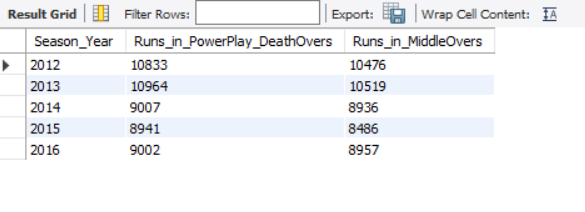
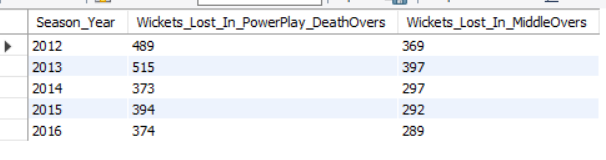
1. **Identify Core Players:** Retain key performers who form the team's backbone.
2. **Strengthen Bowling:** Focus on acquiring top-quality bowlers, especially death bowlers.
3. **Balance the Squad:** Ensure a good mix of experienced players and young talent.
4. **Fill Key Roles:** Address specific gaps, such as an all-rounder or a solid middle-order batsman.
5. **Budget Management:** Strategize spending to get the best value for money, avoiding overspending on marquee players.
6. **Analyze Opponents:** Study other teams' potential strategies to stay ahead in the auction.
7. **Emphasize Fitness:** Prioritize players with good fitness records to minimize injury risks.



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

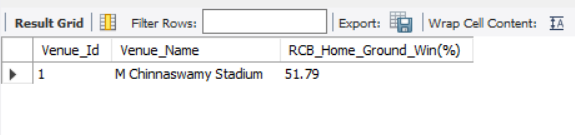
Ans: There can be two factors while dealing with high-scoring matches, the impact on viewership and team strategies: (We consider only last 5 seasons here)

-- i) Comparison between Runs Scored during Power Play (1 to 6 overs) & Death Overs (17 to 20 overs) and Runs Scored during Middle Overs (7 to 16 Overs):



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

Ans: As we see the RCB Performance in their own home ground is better in such cases And also their Batting Average is high comperative to their bowling.

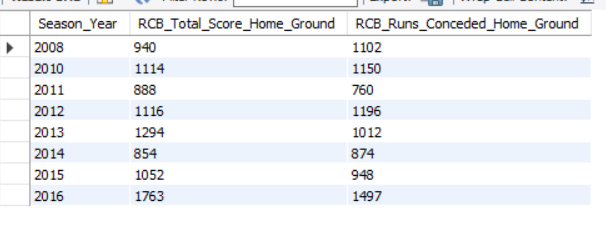


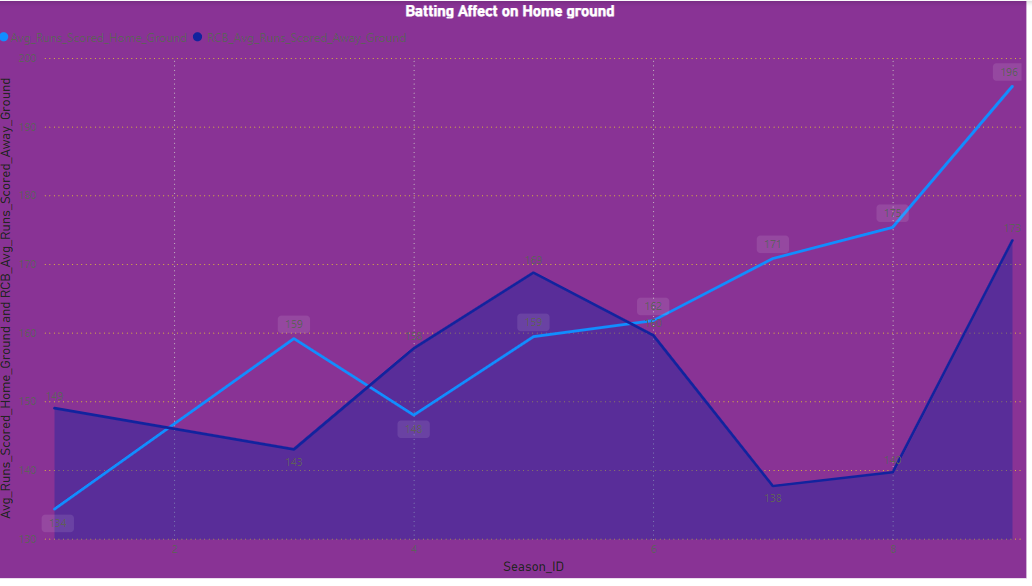
1. Each team will play 50% of the matches played in theirhome ground.
   1. On the behalf of 8 seasons, 4 times RCB batsmen have scored runs greater than RCB bowlers gave the runs.
   2. And 4 times RCB bowlers have conceded runs greater than RCB batsmen runs scored.
   3. Also RCB win percentage at their home ground is 51.79%.
   4. It is because of lack of quality bowlers in their bowling line up.
   5. So in the upcoming auction, RCB should opt mainly for wicket taking bowlers and also who bowls more dot balls in power plays and in death overs.
2. Come up with a visual and analytical analysis with the RCB past seasons performance and potential reasons for them not winning a trophy.

Ans:

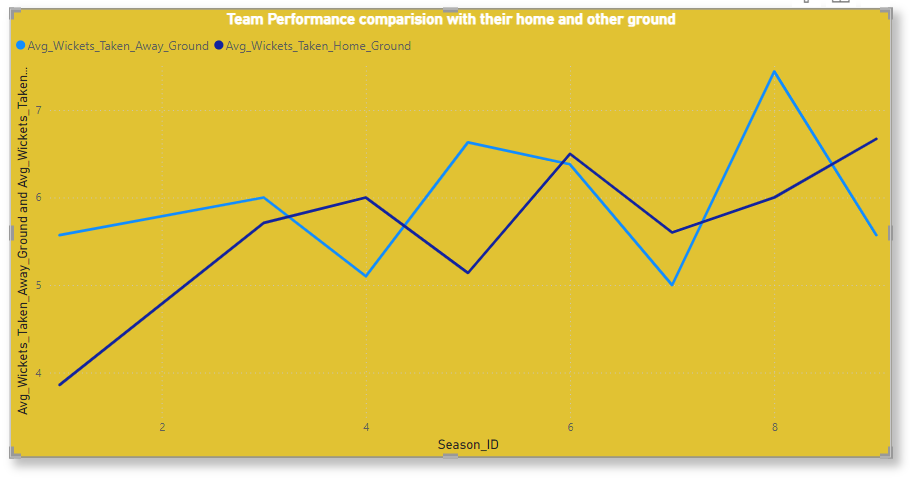
RCB Batters have scored well in all seasons in both Home and Away Grounds.

Considering the recent forms (last 3 seasons), the batters have done extremely well and justified their role.





2



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

Ans:**Team Balance**: Aim to achieve a balanced team with a mix of batsmen, bowlers, all-rounders, and wicketkeepers.

* **Role-Specific Needs**: Identify specific roles that need reinforcement, such as power hitters, finishers, death bowlers, spinners, etc.
* **Budget Management**: Plan to optimize the use of the available budget to maximize team strength without exceeding the salary .

**Objective Criteria**:

* **Batting**: Average, strike rate, consistency, ability to play under pressure.
* **Bowling**: Economy rate, strike rate, wicket-taking ability, performance in different phases (powerplay, middle overs, death overs).
* **Fielding**: Catches, run-outs, overall agility.
* **All-Round Ability**: Contribution with both bat and ball.

**Subjective Criteria**:

* **Current Form**: Recent performances in domestic and international cricket.
* **Injury History**: Fitness and injury records.
* **Team Fit**: Ability to gel with the existing squad and team culture.
* **Leadership and Experience**: Experience in high-pressure situations and potential for leadership roles.
* **This have been explained in Project Slide**.

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".

Ans:

1.In the given IPL Dataset there is no such column name i.e "Opponent\_Team" is Present.

2.So their will be no scope to replace the name in match column.

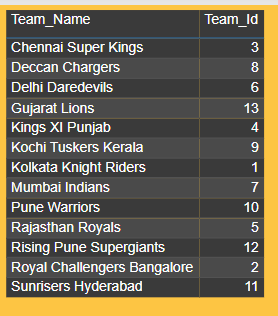
3. If we check Team Table

By :: Select \* from Team;

we get team name here we can change the team name from

"Delhi\_Capitals" to "Delhi\_Daredevils".

By:: **select Team\_Id, replace(Team\_name,"Delhi Daredevils","Delhi Capitals") as Team\_name from team;**

****