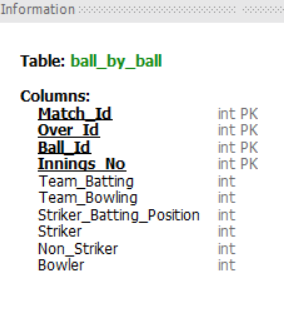


Objective Questions

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

Ans:- Different dtypes of column in ball\_by\_ball are:-

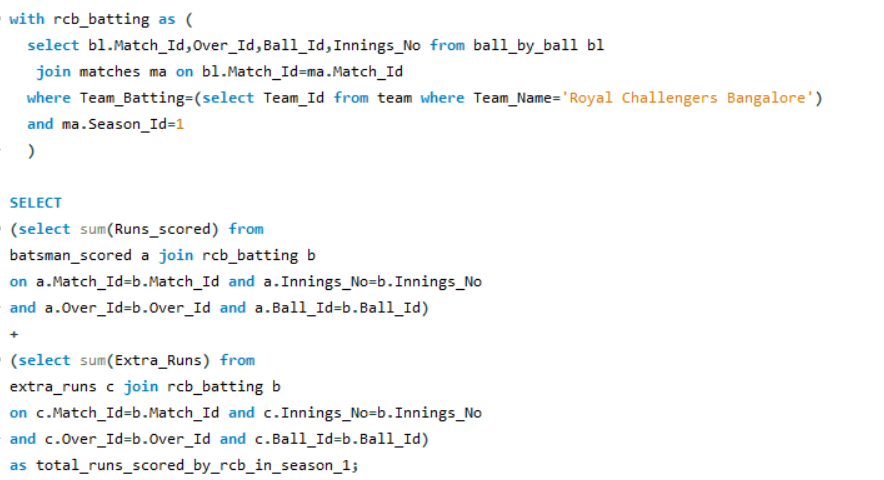


Q.2) What is the total number of run scored in 1st season by RCB (bonus : also include the extra runs using the extra runs table)

Ans:-

Approach:-In this question, I first created a CTE to filter only the matches in which RCB played and only the matches of season 1. Then joined the CTE with table “batsman\_scored” to get the runs\_scored and the did summation of runs\_scored column .Did the same with extra\_runs table.

**Query:-**

****

**Answer:-**

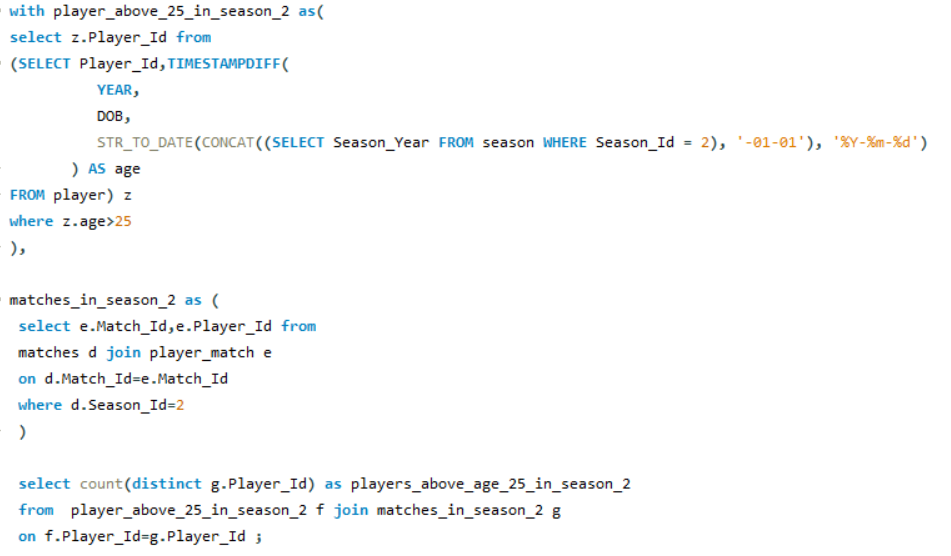


Q.3)How many players were more than age of 25 during season 2 ?

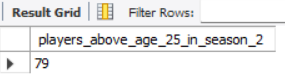
Ans:-

**Approach:-**  Firstly, I calculated the age of all players in 2009 (season 2) and then filtered those who were above 25 age. Then I joined matches and player\_match table to filter matches of season 2 and players palyed in that matches. And then I joined both of tables and selected the distinct count of player\_id.

**Query:-**

****

**Answer:-**

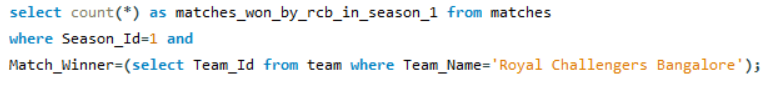


Q.4)How many matches did RCB win in season 1 ?

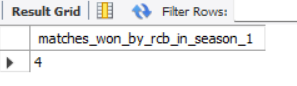
Ans:-

**Approach:-** I selected those matches from matches table where seaon id was 1 and match\_winner was RCB.

**Query:-**

****

**Answer:-**

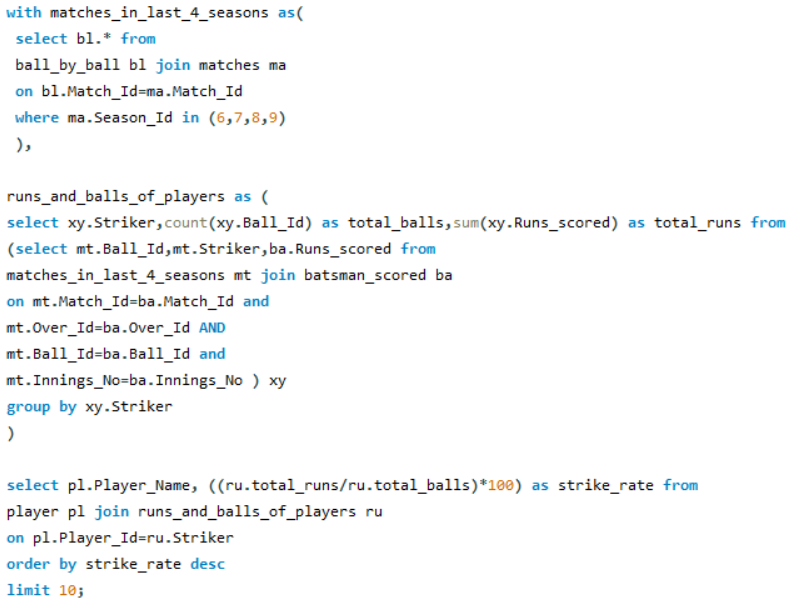


Q.5) List top 10 players according to their strike rate in last 4 seasons.

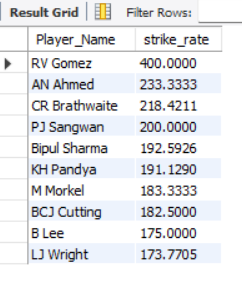
Ans:-

**Approach:-** I first filtered the match id from last 4 seasons and then joined them with ball\_by\_ball because this table contains the info about batsman. But the runs scored are in batsman\_scored so we join again with them and then aggeregate total balls played and runs scored by using group by player\_name and then dividing total run by total balls to get the strike rate, Also order by strike rate in descending order with a limit of 10.

**Query:-**

****

**Answer:-**

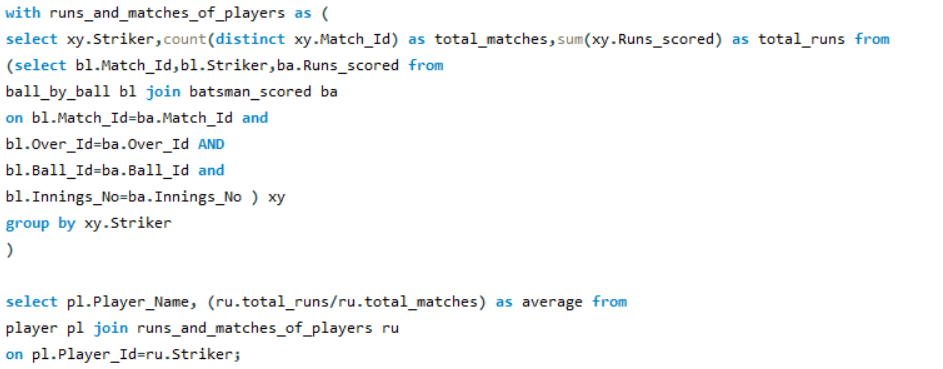


Q.6) What is the average runs scored by each batsman considering all the seasons?

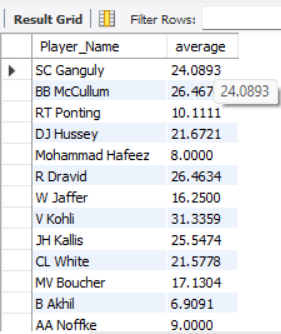
Ans:-

**Approach:-** In this I tried to aggeregate total matches played and total runs scored by every player and then divide them to get the average runs.

**Query:-**

****

**Answer:-** The table contains almost 434 rows but we are showing only few rows here.

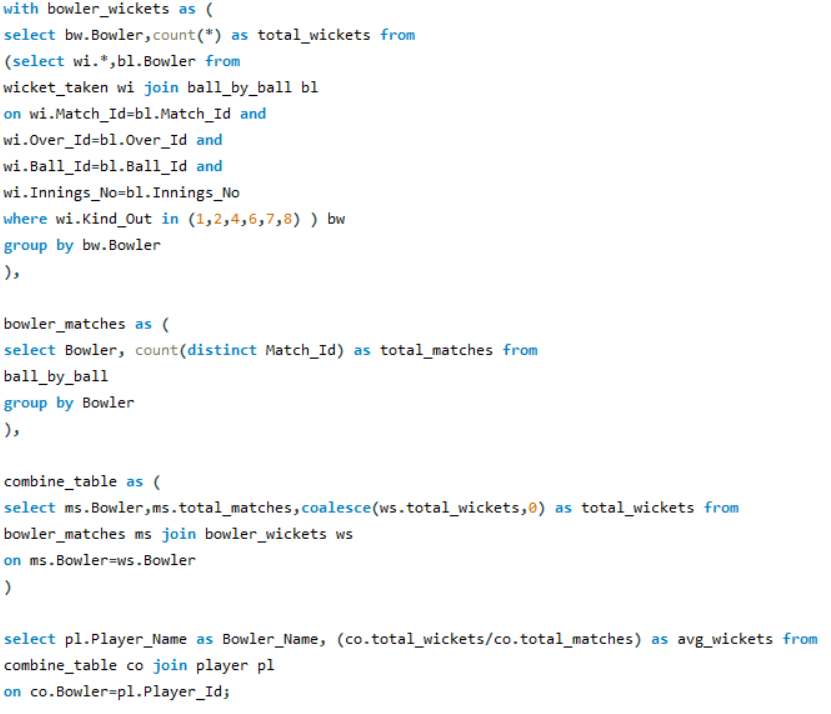


Q.7) What are the average wickets taken by each bowler considering all the seasons?

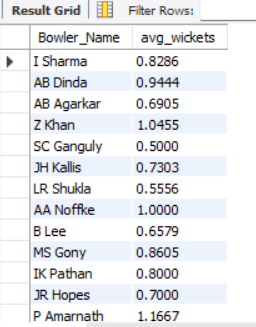
Ans:-

**Approach:-** In this I tried to aggeregate total matches played and total wickets taken by every player and then divide them to get the average wickets.

**Query:-**



**Answer:-** Here we are showing only few rows.

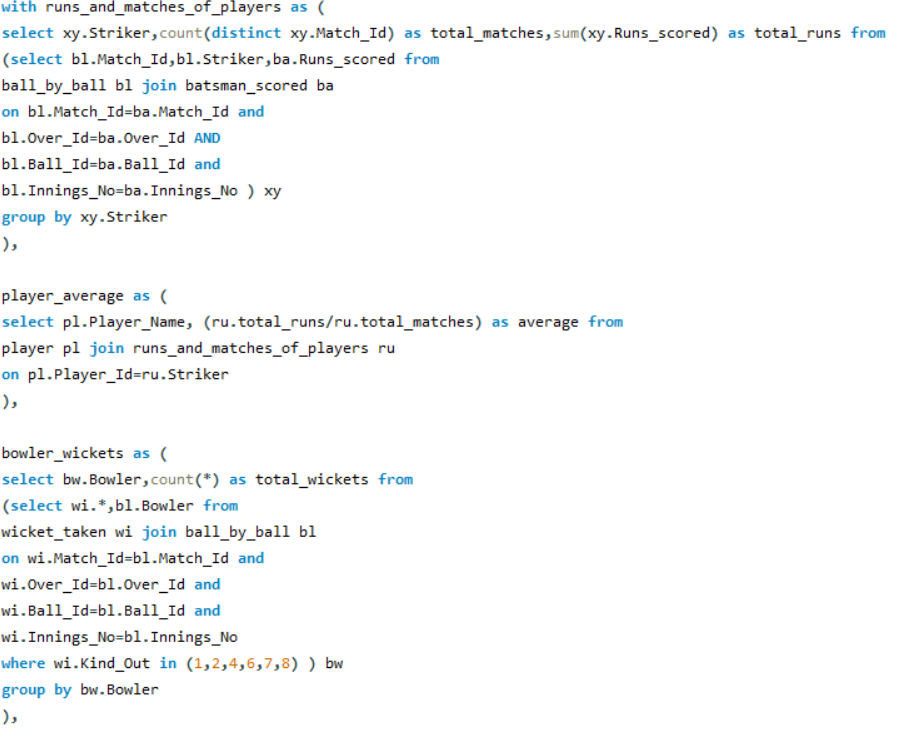


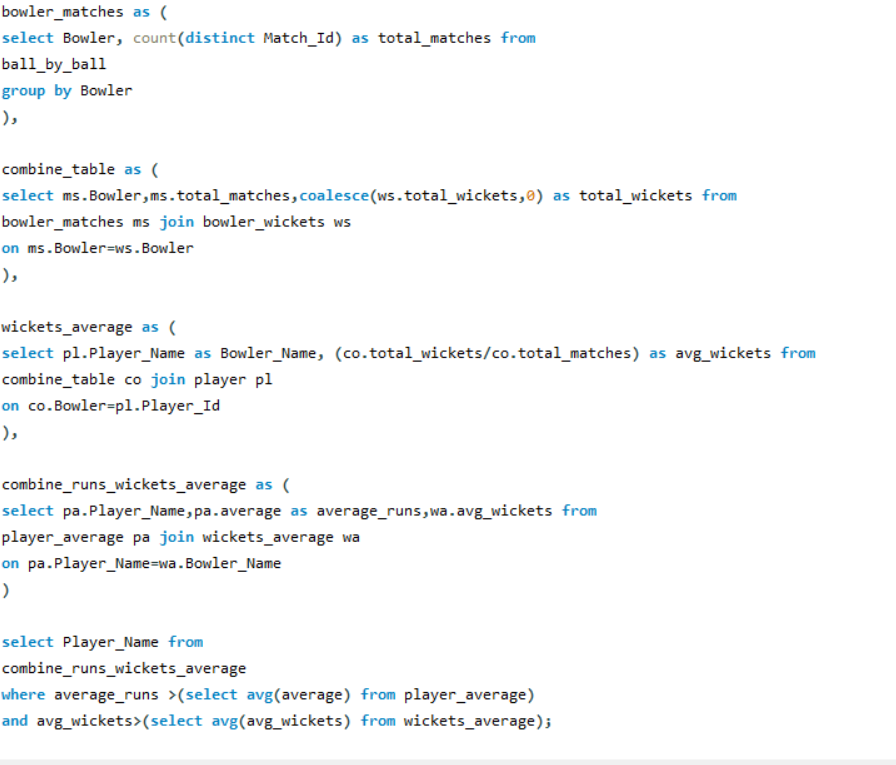
Q.8) List all the players who have average runs scored greater than overall average and who have taken wickets greater than overall average

Ans:-

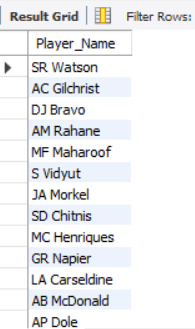
**Approach:-** In this I first tried to calculate average runs and average wickets for each player and then we will filter those who have average better than the overall average in both runs and wickets.

**Query:-** This query is veery long. So I will post 2 screenshots to show the query.





**Answer:-** There are 30 players in this list but I am showing only a few.

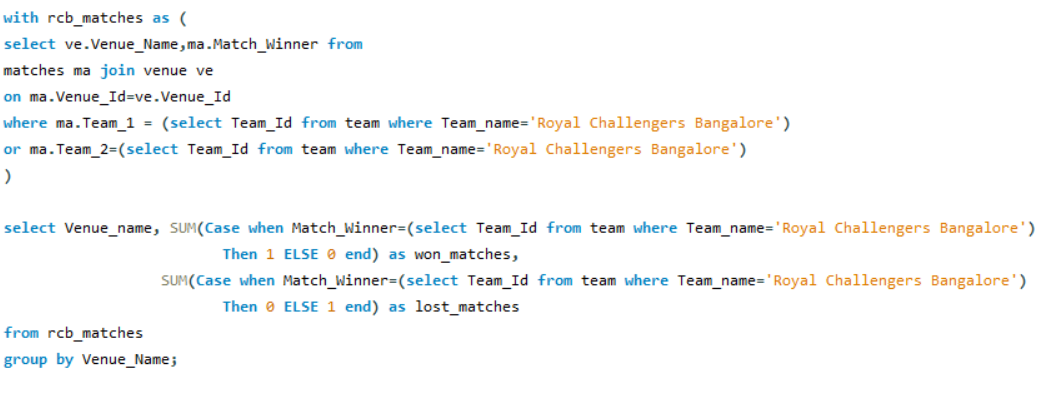


Q.9) Create a table rcb\_record table that shows wins and losses of RCB in an individual venue.

Ans:-

**Approach:-** In this , I aggregated data by venue name and used case function along with sum to calculate both won\_matches and lost\_matches.

**Query:-**

****

**Answer:-** I am showing only few venues here.

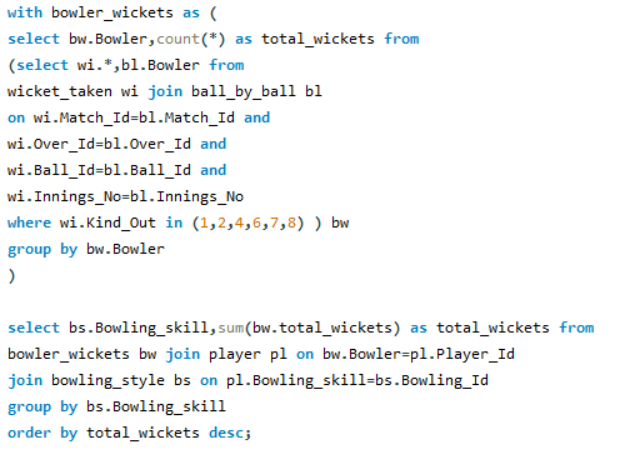


Q.10) What is the impact of bowling style on wickets taken?

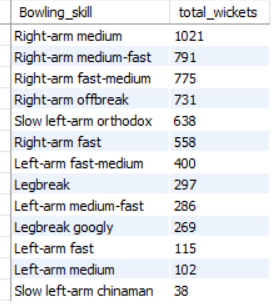
Ans:-

**Approach:-** In this I tried to aggregate wickets by using group by with bowling skill.

**Query:-**



**Answer:-**



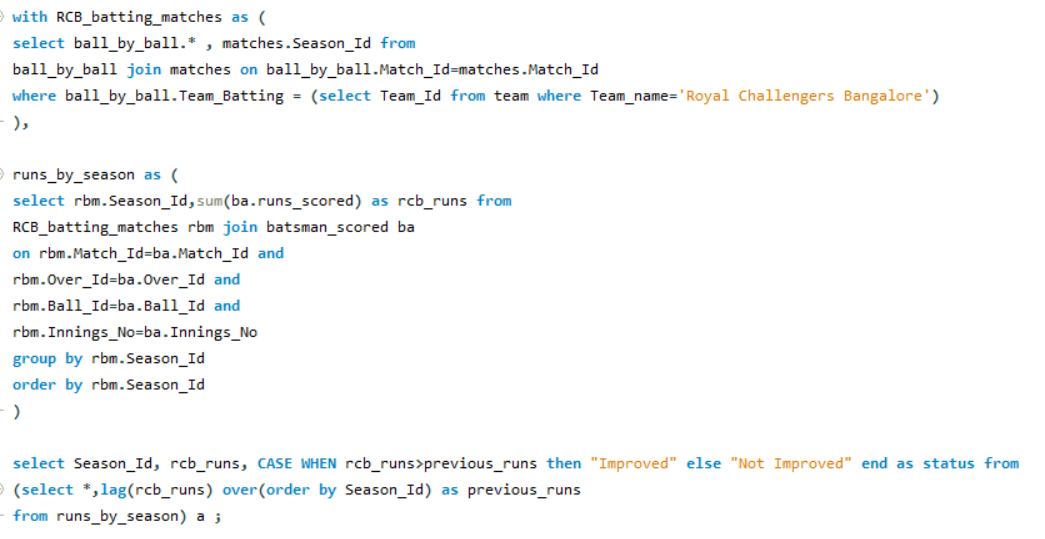
Q.11) Write the sql query to provide a status of whether the performance of the team better than the previous year performance on the basis of number of runs scored by the team in the season and number of wickets taken.

Ans:-

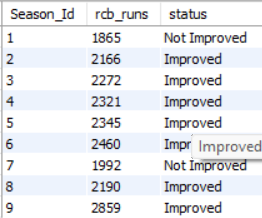
**Approach:-** I broke this question into 2 parts. In 1st part, we will check for runs and in 2nd we check for wickets.

1. **Status of Batting performance:-**

**Query:-**

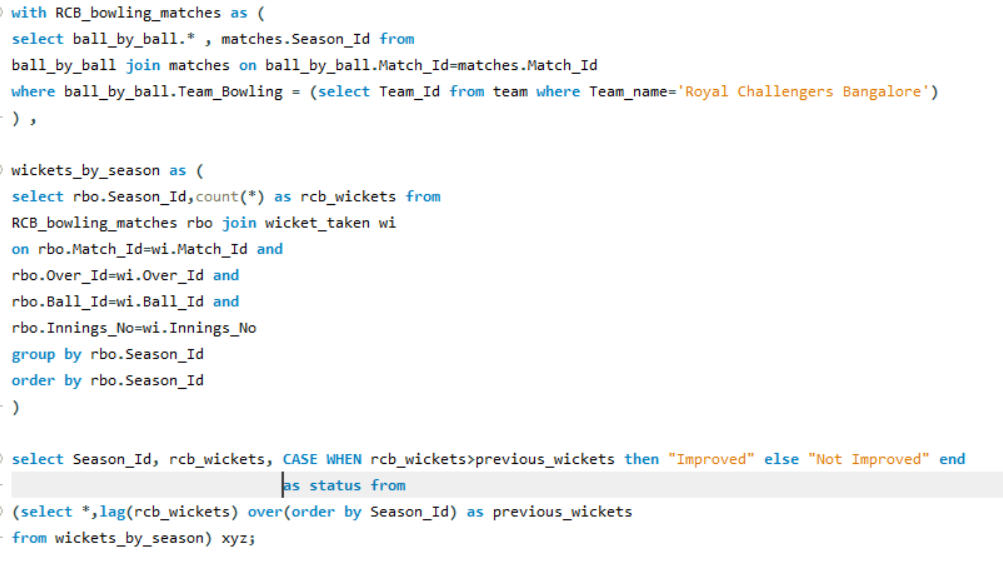
****

**Answer:-**

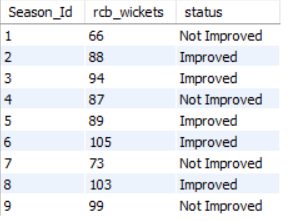
****

1. **Status of Bowling performance:-**

**Query:-**

****

**Answer:-**

****

Q.12) Can you derive more KPIs for the team strategy if possible?

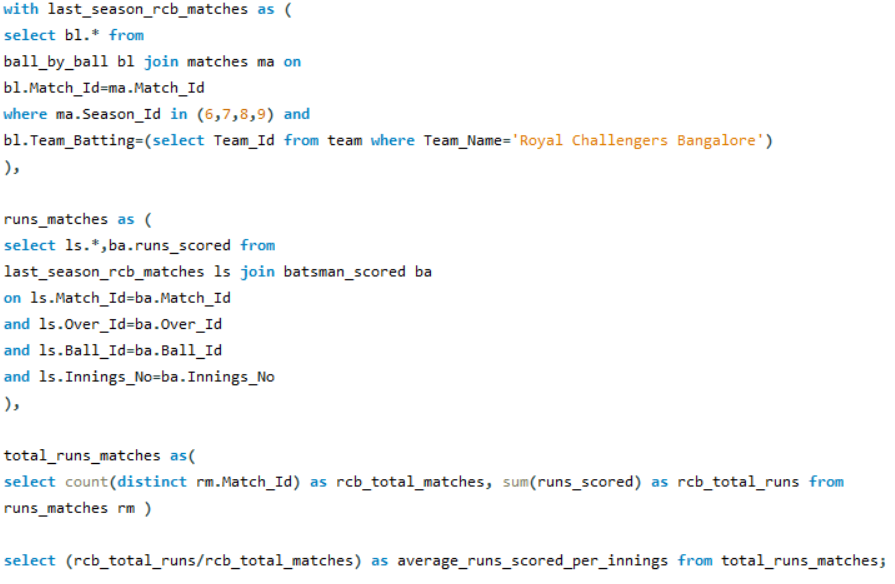
Ans:-

**Approach:-** In this question, I considered 2 KPIs, that is average runs

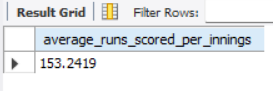
scored per innings and given per innings in the last 4 seasons to analyze both the batting and bowling performances.

**Average runs scored per innings:-** This KPI will help us understand how RCB’s batting unit is performing so that we can decide their strategy regarding batting before going to the auction.

**Query:-**

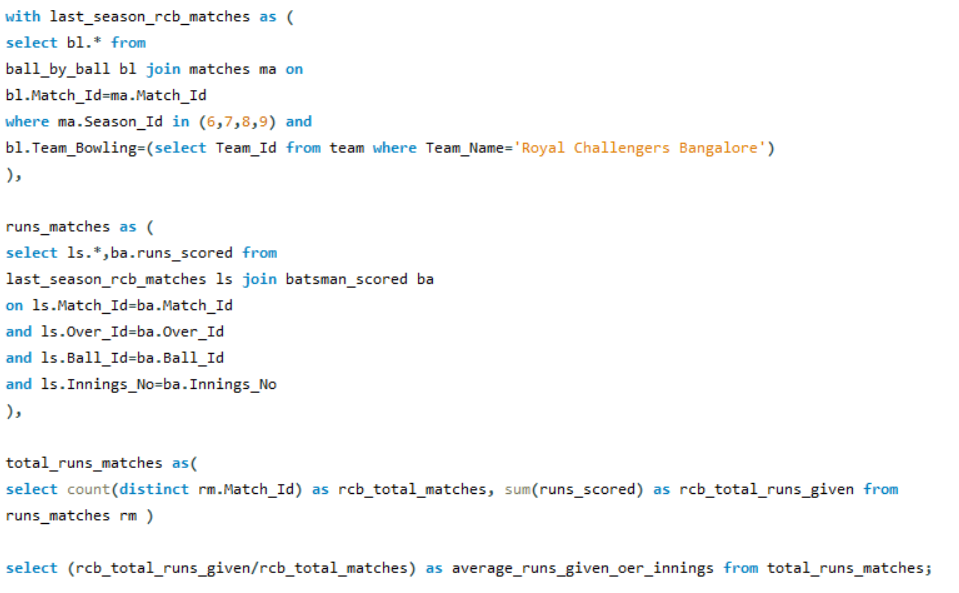
****

**KPI:-**



**Average runs given per innings:-** This KPI will help us understand how many runs are being conceded by RCB bowlers per inning to decide bowling strategy.

**Query:-**

****

**Answer:-**

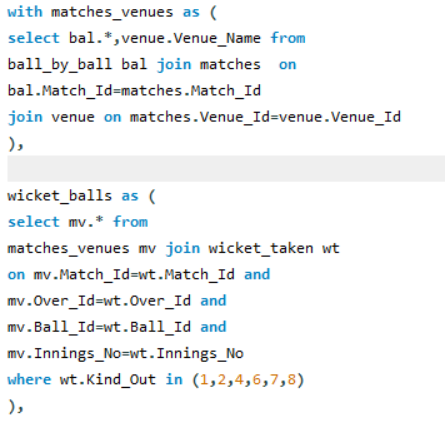


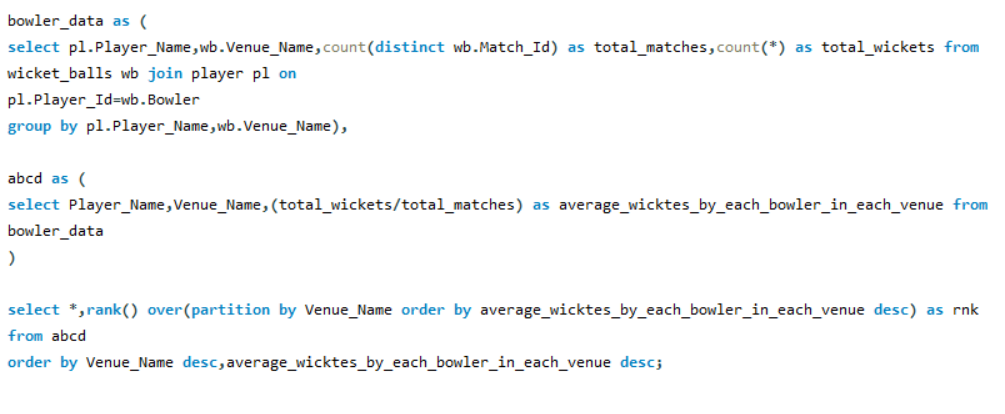
Q.13) Using SQL, write a query to find out average wickets taken by each bowler in each venue. Also rank the gender according to the average value.

Ans:-

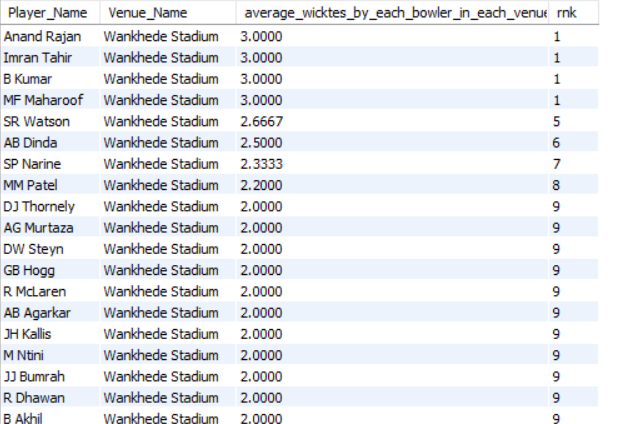
**Approach:-**In this question , I found out the total number of wickets taken by each bowler in each venue and also found total matches played by him in that venue. Divided them to get the average by each bowler per venue. Then ranked them according to the average and did partition by venue.

**Query:-**

****

****

Here is the starting rows of the table:-



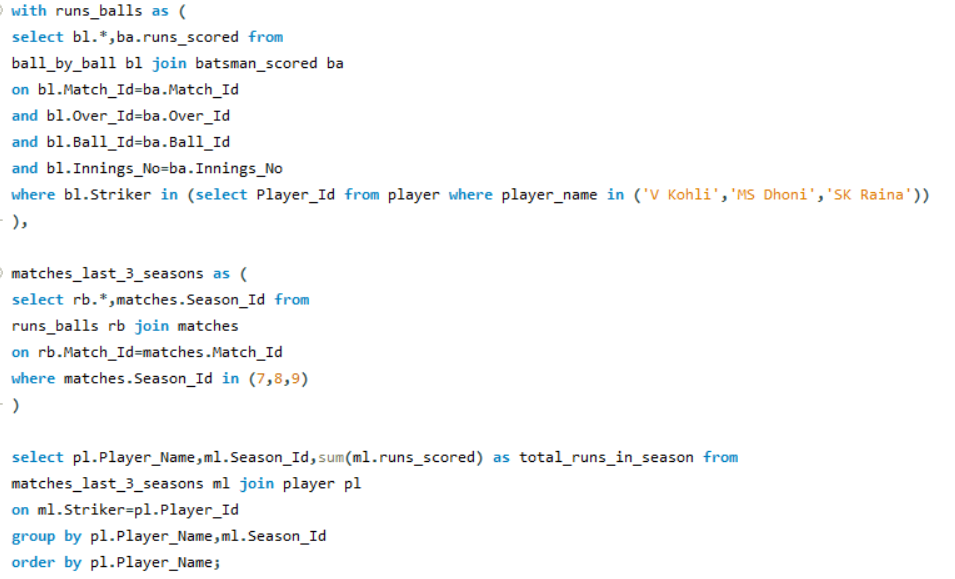
Q.14) Which of the given players have consistently performed well in past seasons? (will you use any visualisation to solve the problem)

Ans:-

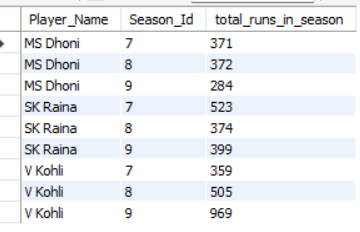
**Approach:-**To analyse this I will consider 3 batsman and 3 bowlers and analyse their performance in the last 3 seasons.

The 3 batsman which I am considering are:- Virat Kohli,MS Dhoni,Suresh Raina.

**Query:-**

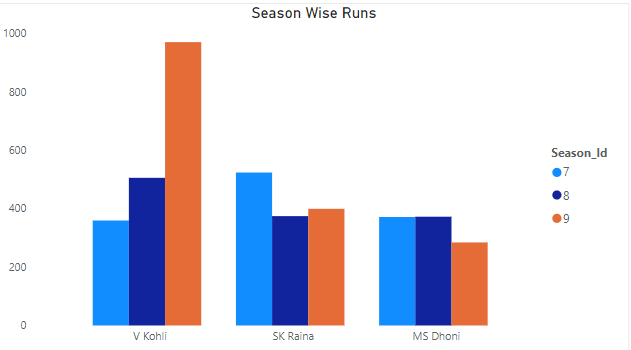
****

So their performances in the past 3 seasons are as :-



So here we can say that runs scored by both MS Dhoni and Suresh Raina has decreased in last 3 seasons but for Virat Kohli,it has increase at a very fast rate.

Lets try to visualise this data to analyse better:-



Q.15) Are there players whose performance is more suited to specific venues or conditions? (how would you present this using charts?)

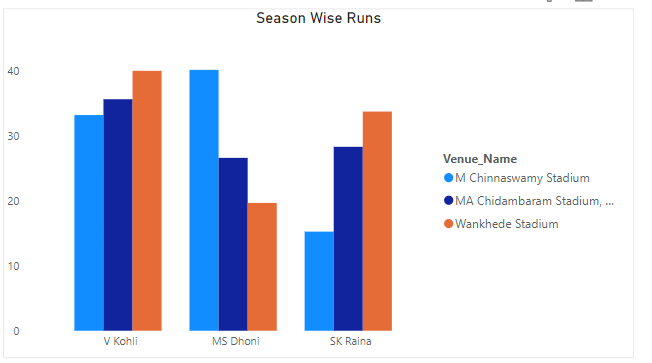
Ans:- In this question we will select 3 venues and 3 players and analyse their performance in these 3 venues. The 3 players I considered are Virat Kohli, Suresh Raina and Ms Dhoni, and the 3 venues are M Chinnaswamy Stadium, MA Chidambaram Stadium, Chepauk,

Wankhede Stadium.

Here is their performance in 3 venues:-



Lets use visualisation to analyse better:-



From this visual we can say that, the performance of both MS Dhoni and Suresh Raina depends on conditions and values while Virat Kohli’s performance is good in almost all conditions.

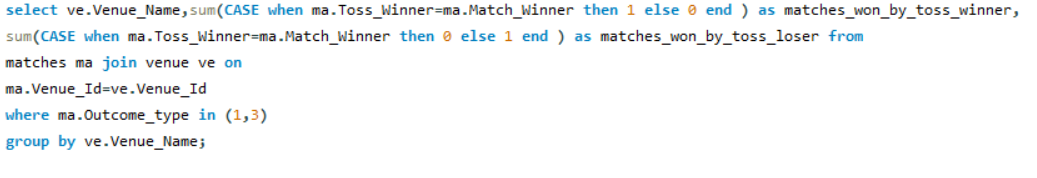
Subjective Questions

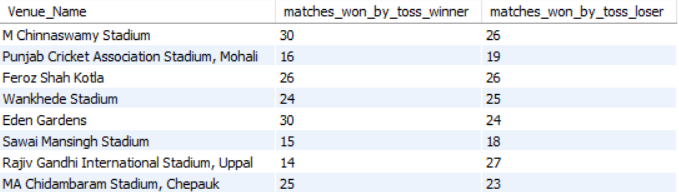
Q.1)How does toss decision have affected the result of the match ? (which visualisations could be used to better present your answer) And is the impact limited to only specific venues?

Ans:-

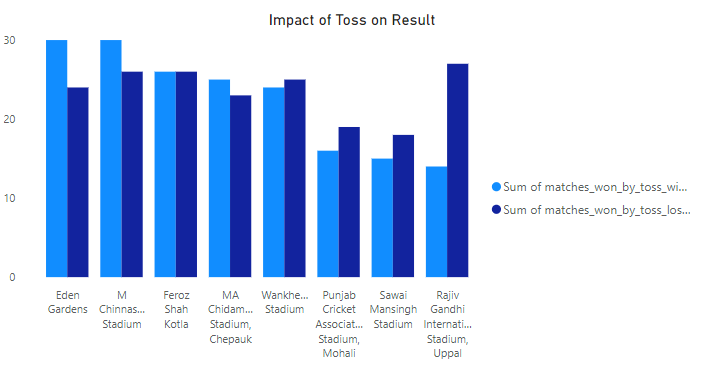
**Approach:-** I have considered only 8 venues because in these venues , sample data is large enough to analyse. In these venues, we will use SQL to find out how many matches are won by toss winner and how many by toss loser.

**Query:-**

****



Lets use a visualisation to analyse better:-



Based on the data and bar chart, here’s an analysis of how the toss decision influences match outcomes at various venues:

**Observations:**

1. **Venues where the toss winner has an advantage:**
   * **Eden Gardens**: The toss-winning teams have secured 30 wins, indicating a strong correlation between winning the toss and winning the match.
   * **M Chinnaswamy Stadium**: Out of 30 toss-winning matches, 26 also resulted in wins for the toss-winning teams. This suggests a significant impact of toss results on the outcome here.
   * **Wankhede Stadium**: Similarly, toss winners recorded 24 wins, and they won 25 matches overall, showing a balanced trend between toss outcomes and match results.
2. **Venues where the toss has less influence:**
   * **Feroz Shah Kotla**: Both toss winners and toss losers performed equally well, with 26 wins each. This implies that the toss may not significantly affect match results here.
   * **MA Chidambaram Stadium, Chepauk**: While toss winners achieved 25 wins, toss losers weren’t far behind with 23 wins. Thus, other factors might be more critical at this venue.
   * **Sawai Mansingh Stadium**: The distribution is fairly balanced, with toss losers having a slight edge (18 wins) compared to toss winners (15 wins). This suggests minimal toss impact.
3. **Notable trends:**
   * **Punjab Cricket Association Stadium, Mohali**: The data reflects that toss winners had fewer wins (16) compared to toss losers (19), indicating that losing the toss might not be a disadvantage here.
   * **Rajiv Gandhi International Stadium, Uppal**: Interestingly, toss losers have the upper hand with 27 wins, compared to only 14 wins by toss winners. This reverses the typical pattern where toss winners often have the advantage.

**Conclusion:**

The toss impact varies across venues. At some venues like Eden Gardens and Chinnaswamy Stadium, winning the toss appears to provide a significant advantage, potentially due to pitch conditions or weather that favors early decisions. In contrast, at venues like Feroz Shah Kotla and Rajiv Gandhi International Stadium, the toss has minimal or even a reversed influence, suggesting that other factors such as team strategy or player performance play a more prominent role.

This analysis highlights the importance of venue-specific conditions in cricket and how they interact with the toss to shape match outcomes.

4o

Q.2) Suggest some of the players who would be best fit for the team?

Ans:- I would suggest those players to RCB which are below or equal to the age of 27 and have taken either more wickets or scored more runs in last 2 seasons. Buying young talent in the auction is very necessary because successful teams like MI,CSK have built their team on these strategies.

**Approach for Selecting Players:**

**Step 1: Defining Key Filters**

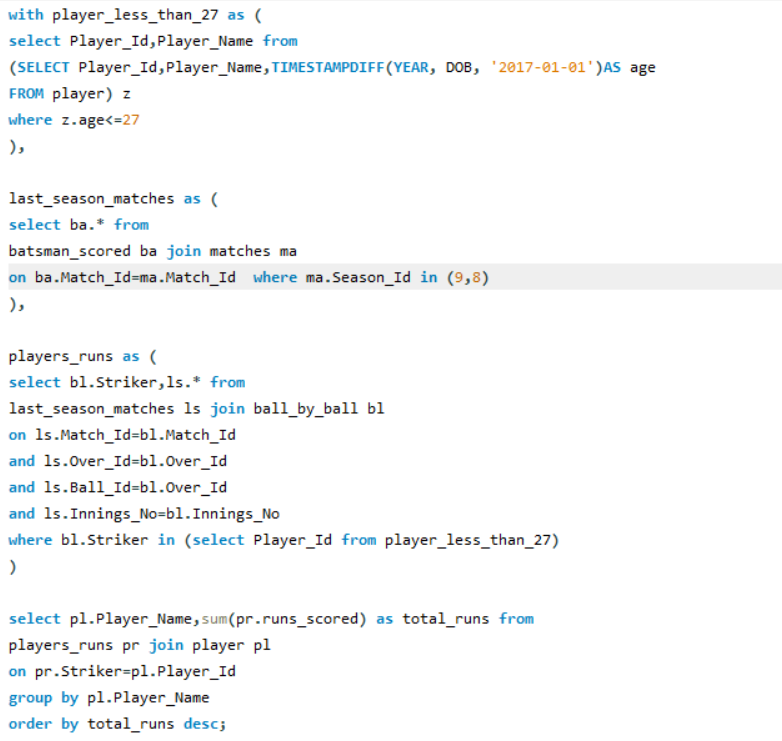
1. **Age Filter:**
   * **Include players aged 27 years or younger**.
   * This ensures RCB invests in young talent that can serve the team for several years, building consistency and core strength.
2. **Performance Filter (Last 2 Seasons):**
   * **Batters/All-rounders:** Players with **high run aggregates** in the past two seasons .
   * **Bowlers:** Players with **more wickets** (e.g., > 25 wickets) over the last two seasons.
   * This ensures that only consistent performers are targeted. The focus is on those excelling in their roles, adding immediate value to the squad.

**Step 2: Analyze the Dataset for Players**

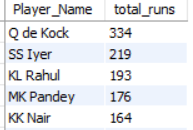
* **Collect Data:** Get player statistics for the past two seasons from using SQL queries for:
  + Total Runs (for batters and all-rounders)
  + Total Wickets (for bowlers and all-rounders)

**So the suggested batsman are:-**

**Query:-**

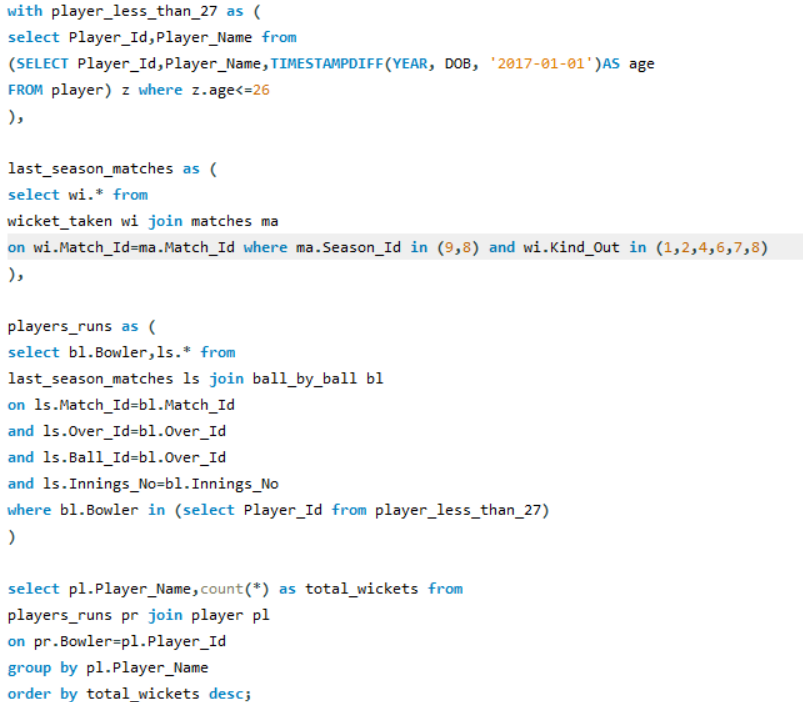
****

**Answer:-**

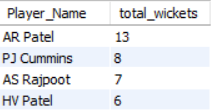


**The suggested bowlers are:-**

**Query:-**

****

**Answer:-**



Q.3) What are some of parameters that should be focused while selecting the players?

Ans:- The parameters that should be considered are:-

a) Strike rate of players while batting in the death overs(16-20).

b) Economy of the bowlers in death overs(16-20).

c) Focussing on players who can deliver both with the bat and bowl ,i.e., all-rounders.

d) Focussing on young talent and build the future team around them.

e) Dropping players who are above 35 and not performing.

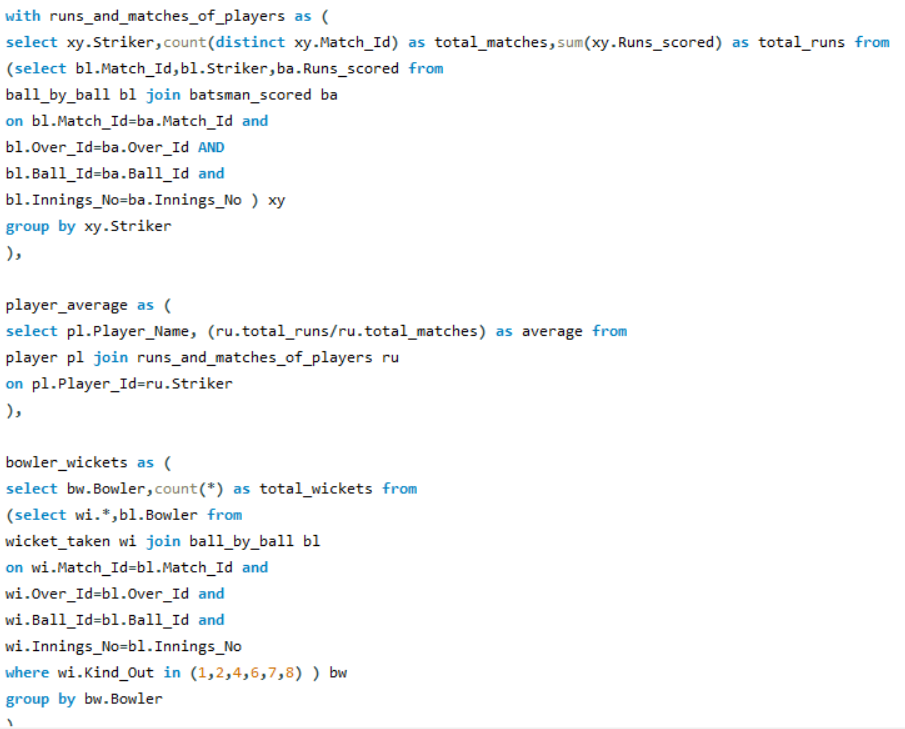
f) Selecting proper players for each position .

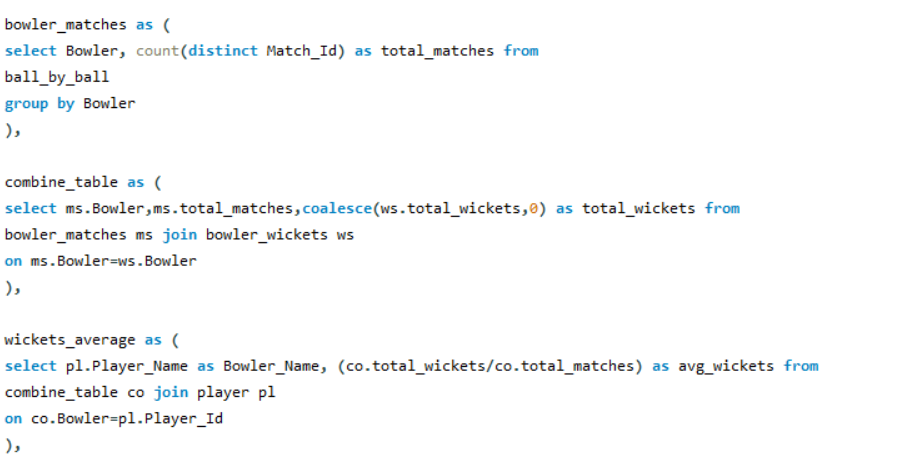
Q.4) Which players offer versatility in their skills and can contribute effectively with both bat and ball? (can you visualize the data for the same)

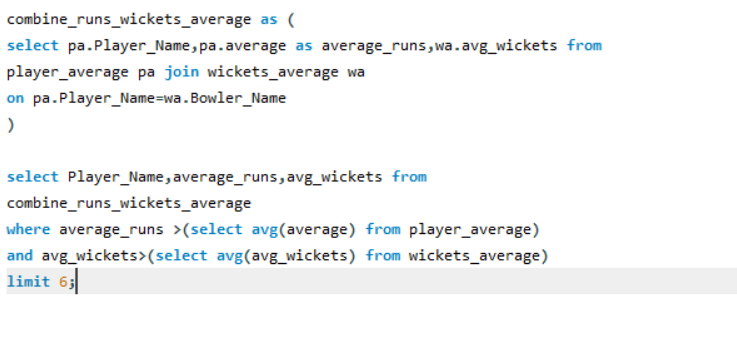
Ans:-

**Approach:-** I am considering players who have average runs scored greater than overall average and who have taken wickets greater than overall average.

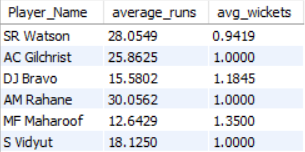
**Query:- The query is very big, so we will use 3 screenshots.**

****

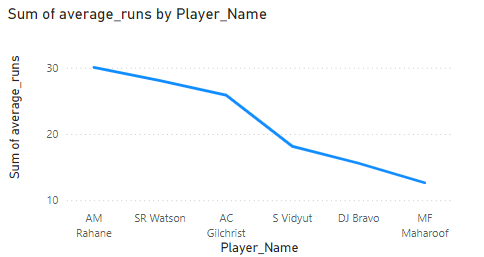
****

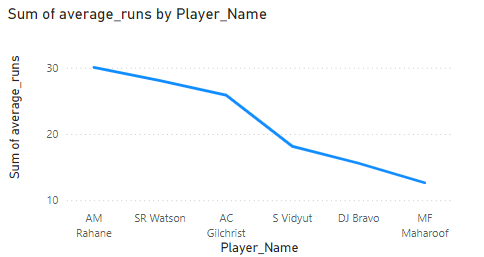
****

Here are some of the players who are versatile both with bat and bowl:-



Lets visualise this data:-





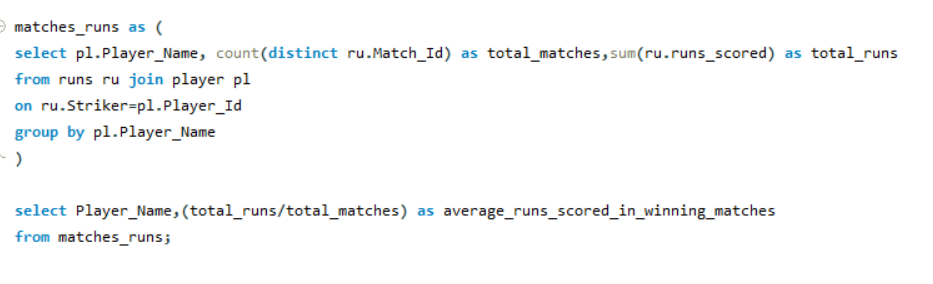
Q.5) Are there players whose presence positively influences the morale and performance of the team?

Ans:-

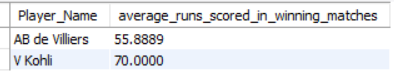
**Approach:-** In this question we will check the performance of 2 players, Virat Kohli and AB de Villiers to check whether their performance increase the morale or not. We will check their average in winning cause.

**Query:-**

****

****

**Answer:-**



From the provided data, we observe the **average runs scored in winning matches** by these two players:

* **Virat Kohli:** 70.00
* **AB de Villiers:** 55.89

These statistics suggest that both players have a significant positive impact on the team’s performance, especially during matches that result in victories. Let’s explore how their performances influence the **morale and overall performance of the team**.

**1. Setting an Example with Consistent Performance**

* **Kohli’s average of 70** in winning matches highlights his ability to rise to the occasion when the team needs him the most.
* **De Villiers’ average of 55.89** also showcases his consistency and reliability under pressure.
* Players who score high consistently in **winning causes** set a **benchmark for other players** to perform similarly, inspiring the entire squad to raise their game.

**2. Providing Stability and Confidence in Key Situations**

* Both Kohli and de Villiers are known for their **composure** in high-pressure situations. Their ability to stay at the crease for longer periods brings stability, which reduces the pressure on other batters.
* Other players, especially youngsters, are likely to feel more confident when batting alongside proven performers like these two, knowing they have a strong anchor at one end.

**3. Leading by Example and Boosting Morale**

* **Virat Kohli**, as a leader and former captain of RCB, not only performs with the bat but also sets high standards for discipline, fitness, and determination. His example motivates teammates to put in their best efforts.
* **De Villiers**, with his experience and ability to play unconventional shots under pressure, brings in a fearless brand of cricket that **raises the team’s fighting spirit**.

**4. Momentum Creation and Match-winning Impact**

* High individual averages in winning matches indicate that both players perform at their best when it matters the most. This habit of **contributing in crucial moments** creates momentum that carries through the team, helping them stay positive even in difficult situations.
* Knowing that players like Kohli or de Villiers are in good form gives the team confidence to push hard, knowing that they have match-winners to rely on.

**5. Inspiring and Mentoring Young Players**

* Both players are known to share their experience with younger teammates, guiding them on how to handle pressure. This has a direct impact on the **development and morale** of young players, ensuring a long-term positive influence on the team’s culture and performance.

**Conclusion:**

The high averages of Kohli (70.00) and de Villiers (55.89) in winning matches show that their individual performances are strongly aligned with the team’s success. Their ability to deliver consistently under pressure, provide stability, and lead by example significantly **boosts the morale and confidence of the entire team**. This makes them invaluable assets to RCB, not just as players but as motivators and mentors, contributing to both short-term wins and long-term team development.

Q.6) What would you suggest to RCB before going to mega auction ?

Ans:- I would suggest following to the RCB:-

a) Targeting good lower order batters (5,6,7) for better performance in death overs.

b) Targeting bowlers which can produce swing and seam with the all in powerplay and gives the early breakthrough.

c)Focussing on spinners which can control the pace of the game in middle overs and take wickets at regular intervals.

d) Focussing on bowlers which can produce different variations in death overs like yorker,knucle ball, cutter etc.

e) Targeting fitter players to improve fielding as we know “Catches Win Matches”.

f) Building a base around young talent with the existing senior base of players guiding them.

Q.7) What do you think could be the factors contributing to the high-scoring matches and the impact on viewership and team strategies?

Ans:- High-scoring matches in cricket, particularly in formats like T20, can be influenced by various factors that also impact viewership and team strategies. Here are some of the key contributing factors and their effects:

Factors Contributing to High-Scoring Matches:

**a) Flat Pitches and Smaller Boundaries:**

* Impact: Flat pitches with little assistance for bowlers result in more run-scoring opportunities, as the ball comes onto the bat nicely, making it easier for batsmen to play big shots. Smaller boundaries allow for more frequent sixes and fours.
* Viewership Impact: High-scoring games with more boundaries are typically more exciting for viewers, leading to increased viewership and fan engagement.
* Team Strategy: Teams often select more aggressive, power-hitting batsmen who can take advantage of such conditions. Bowlers focus more on variations like slower balls and yorkers to counter the conditions.

**b) Power Hitting and Batting Depth:**

* Impact: Modern teams focus heavily on power-hitting and having deeper batting line-ups, where even lower-order batsmen can clear the ropes. This increases the likelihood of consistently high scores.
* Viewership Impact: High entertainment value due to frequent big hits and run chases creates suspense and attracts casual viewers, boosting ratings and ticket sales.
* Team Strategy: Teams invest in all-rounders and big hitters, with a focus on maximizing the scoring rate at every stage of the innings, particularly in the last 5-10 overs. It also encourages aggressive strategies in batting orders.

**c) Fielding Restrictions (Powerplays):**

* Impact: Powerplays (the first 6 overs in T20 cricket, for example) limit the number of fielders outside the 30-yard circle, allowing batsmen to play more freely and score rapidly during this period.
* Viewership Impact: Early flurries of boundaries and aggressive batting during powerplays can draw in viewers right from the start of the match.
* Team Strategy: Teams often plan aggressive starts to maximize runs in powerplay overs, leading to higher scores. Top-order batsmen are selected for their ability to exploit this phase.

**d) Quality of Batsmen and Evolution of Techniques:**

* Impact: Modern batsmen have adapted to playing unconventional shots (like reverse sweeps, scoops, and switch hits), increasing their ability to score quickly. Improved batting techniques and fitness have also contributed to higher run totals.
* Viewership Impact: Spectacular and innovative shot-making keeps the audience engaged and excited throughout the match.
* Team Strategy: Teams may favor more versatile and aggressive batsmen who can play a variety of shots to keep the run rate high, even against quality bowling attacks.

**e) Increased Focus on T20 Analytics:**

* Impact: Data analytics are increasingly used to analyze bowlers’ weak spots and field placements, giving batsmen an advantage. Teams also use analysis to identify areas on the field where they can score more easily.
* Viewership Impact: Close matches, driven by optimized strategies, make the games more thrilling for the audience.
* Team Strategy: Teams use match-up data and strategy based on bowler weaknesses and ground dimensions to optimize their batting line-up and game plan.

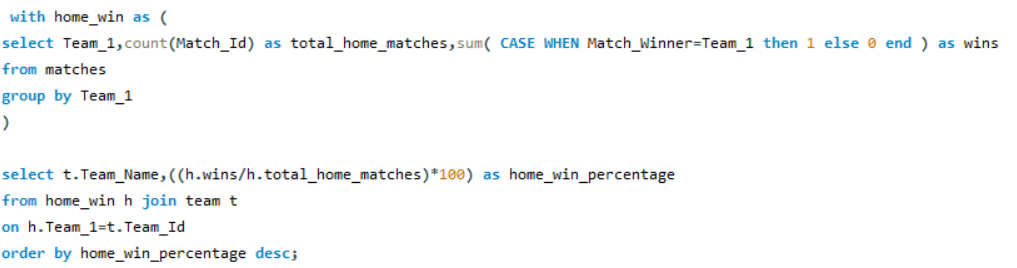
Q.8) Analyze the impact of home ground advantage on team performance and identify strategies to maximize this advantage for RCB.

Ans:-

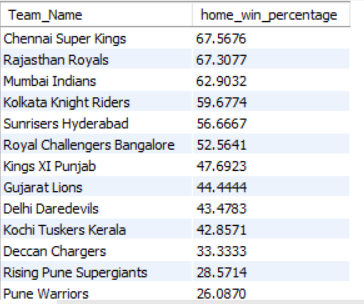
**Approach:-** Here we are using a information to make the query short.

The info is that the home team is always written as Team\_1 , you can even check that when Team\_1 is MI, then ground will be Wankhede stadium. So we are skipping this step of finding home ground.

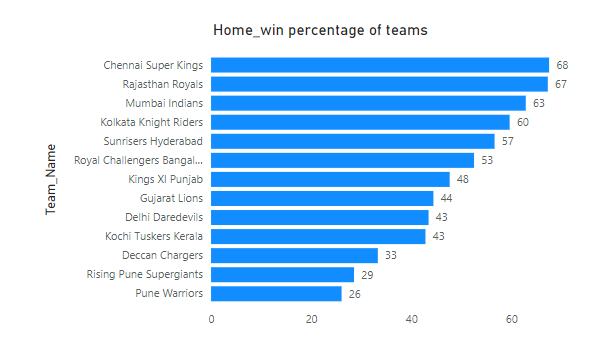
**Query:-**

****

Here is the data showing teams and their win percentage at home grounds:-



Lets try to visualise this:-



From this visual, we can conclude that CSK has the best performance in their home grounds while RCB is on 7th position in this list and hs failed to take any home advantage.

Now we will look at strategies to improve home win percentage of RCB.

To improve Royal Challengers Bangalore's (RCB) home-ground win percentage and maximize their advantage at the M. Chinnaswamy Stadium, here are several strategies that focus on leveraging their home conditions and addressing areas of weakness:

### **1. Leverage Ground Conditions (M. Chinnaswamy Stadium):**

* **Pitch & Boundary Characteristics**: The M. Chinnaswamy Stadium in Bangalore is known for its **flat pitch** and **short boundaries**, which typically favor aggressive batting and high-scoring matches. However, this can also be a challenge for bowlers.
* **Strategy**:
  + **Power Hitters**: Focus on building a team with explosive middle-order and lower-order power hitters who can take full advantage of the small boundaries and accelerate in the death overs.
  + **Bowling Variations**: Invest in bowlers who can **bowl economically in high-scoring conditions**, such as those skilled in **death bowling** (e.g., yorkers, slower balls, cutters). This can help limit opposition runs, especially in the last overs.

### **2. Improving Bowling Strategy:**

* **Tactical Bowling Selections**: Given that the stadium favors batsmen, RCB needs bowlers who excel in **pace variations**, **slower deliveries**, and **spinners who can turn the ball even on flat surfaces**. Historically, spinners have found some success on flat pitches by relying on flight, guile, and variations in pace.
* **Death Bowling**: Focus on improving death-over bowling by bringing in specialists who can consistently execute yorkers and slower balls. Teams like Mumbai Indians and Sunrisers Hyderabad have excelled at death bowling, which could be an area RCB can improve on.
* **Bowling in Powerplay**: Having bowlers who can take early wickets in the powerplay will reduce pressure on bowlers later in the innings. Early wickets can halt the opposition's momentum.

### **3. Aggressive Batting Power:**

* **Top-Order Explosiveness**: As RCB already has some explosive top-order players (like Virat Kohli, Faf du Plessis, etc.), they should continue to rely on players who can get off to a flying start. Openers should aim to score big in the first 6 overs (powerplay), making the most of fielding restrictions.
* **Lower-Order Finishers**: Having strong finishers who can clear boundaries in the last 5-6 overs is key to posting or chasing down large totals at the Chinnaswamy. Players like **Dinesh Karthik** or **Glenn Maxwell** are great examples of finishers who can dominate these conditions.

### **4. Strengthen Fielding and Catching:**

* **Home Crowd Energy**: The Bangalore crowd is passionate and loud. RCB needs to channel this energy into excellent fielding performances. Teams with strong fielding units often win close matches, so RCB should focus on **maximizing fielding efficiency**, especially in stopping boundaries and converting half-chances.
* **Boundary-Line Fielding**: Since the boundaries are shorter, RCB should employ athletic fielders along the boundary line to save critical runs and pull off catches.

### **5. Home Advantage - Playing to Strengths:**

* **Data-Driven Strategy**: Utilize **data analytics** to analyze past home-ground performances and identify weak spots. Detailed analysis of opposition teams' weaknesses (e.g., which bowlers or batsmen struggle in small-ground conditions) will help plan match strategies.
* **Opposition Analysis**: Focus on specific opponents. Analyze their key players and customize plans for each match. For example, some teams may struggle against spin, while others may have a weak death bowling unit.

Q.9) Come up with a visual and analytical analysis with the RCB past seasons performance and potential reasons for them not winning a trophy.

Ans:-

**Approach:-** We will comsider 3-4 factors here to analyse RCB's past performance. We know that top order batting is RCB's strenght , so our factors will be more related to lower order and bowling.

We will consider these factors:-

a)Total runs scored and wickets taken by RCB in previous seasons.

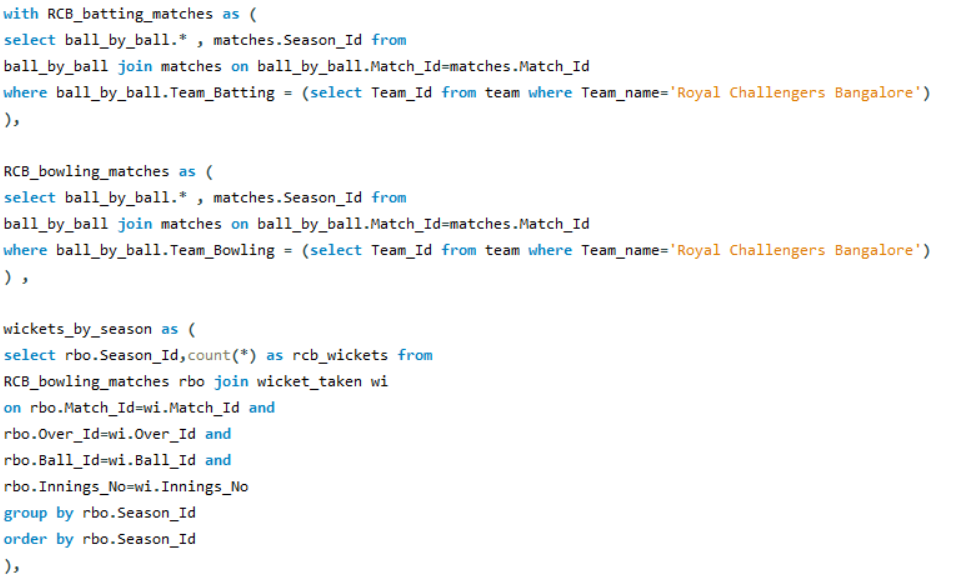
b) economy of RCB bowler in overs 16-20.

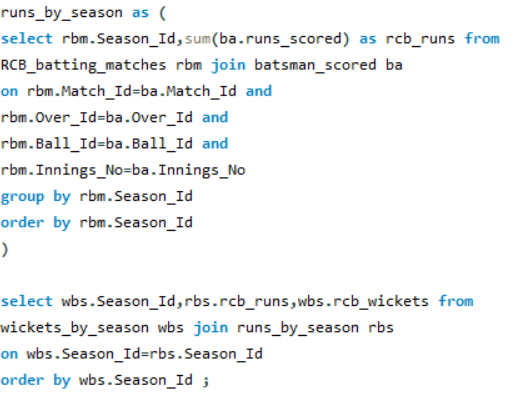
c) Wickets taken per inning by rcb.

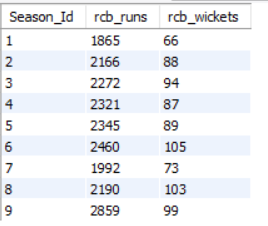
d) Averages of batting position 5,6,7.

**a) Total runs scored and wickets taken by RCB in previous seasons:-**

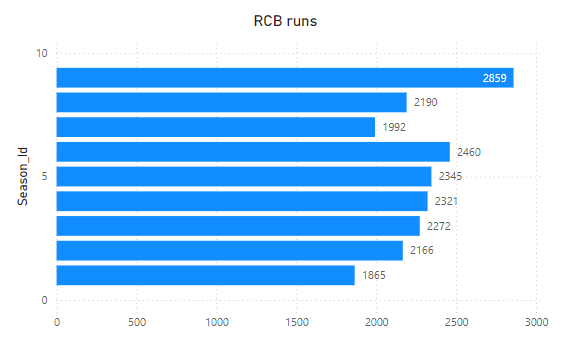
**Query:-**

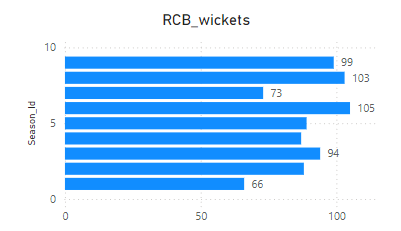
****

****

****

**Lets use a visual:-**

****

****

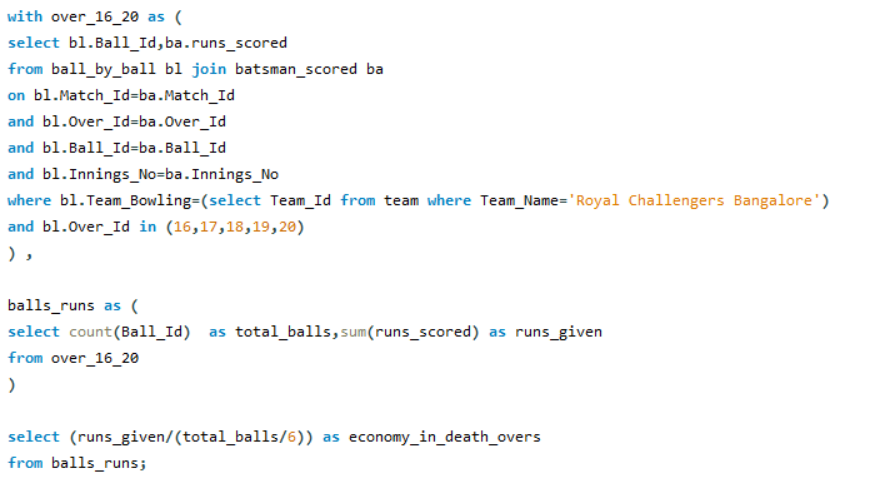
**Analysis:-** From both the table and visual, we can clearly see that RCB scored most runs in 2016 (season 9) ,which is also the season where RCB reached final.They also took close to 100 wickets in 2016.

Also the 2 seasons (1 and 7) where RCB scored less than 2000 runs and took less than 80 wickets, are the seasons in which they ended up in bottom 4 of the table.

So we can conclude that RCB did not win trophy because of their inconsistent performance, like sometimes their batters were good, sometimes their bowlers but as a unit they didn’t performed together and when they did, they almost reached the final in 2016.

**b) Economy of RCB bowler in overs 16-20 :-**

**Query:-**

****

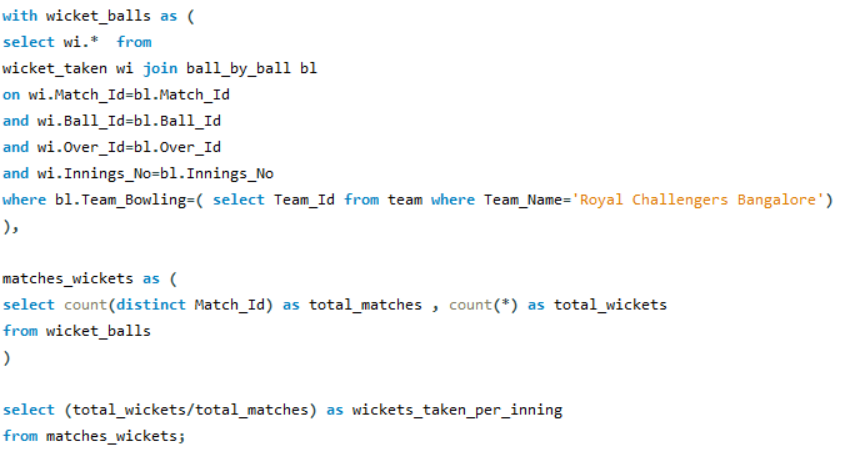
****

**Analysis:-** So the data shows that RCB bowlers are leaking ryan at the rate of almost 9.4 in death overs. This is a very high number since the data also considers initial seasons when the game was not that advanced and even 150 was a winning score.

This high economy rate is one of the main reasons of RCB not winning trophy as these are the overs where game changes.

**c) Wickets taken per inning by rcb:-**

**Query:-**

****

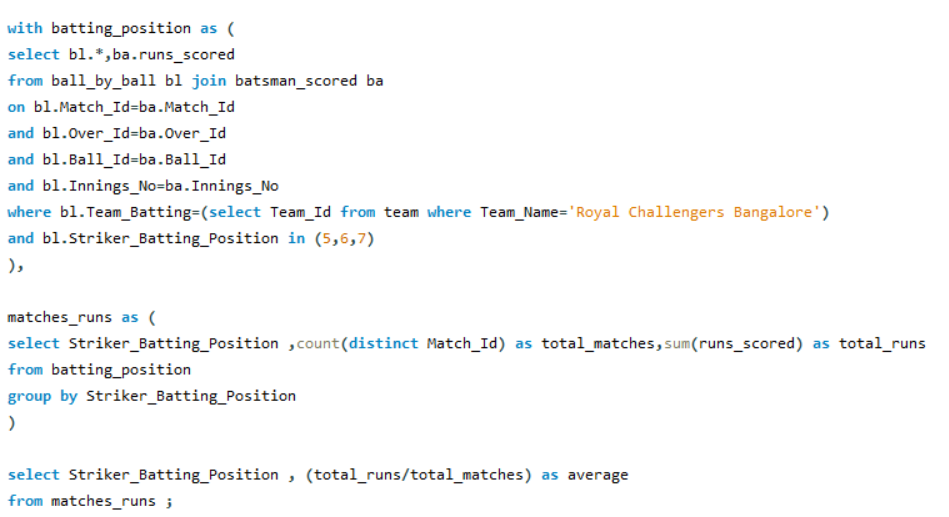
****

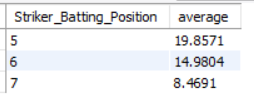
**Analysis:-** RCB bowlers are only taking around 6 wickets per inning.

This means that they are not able to restrict the team to a good total because these days batting extends very large due to presence of all rounders. This is another one of the many reasons of RCB not winning trophy.

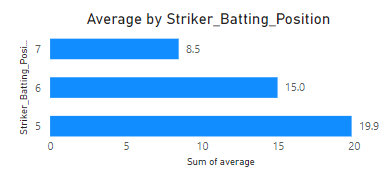
**d) Averages of batting position 5,6,7:-**

**Query:-**

****

****

**Lets use a visual now:-**

****

**Analysis:-** All the 3 positions of lower middle order have average less than 20. RCB’s top order contain players like Chris Gayle, Virat Kohli,AB de Villers. So when they perform, RCB usually wins matches but when they fail, there is no one in the lower middle order to bail them out.

Even if the top 3 performs, you need to finish the game which RCB lacks.

So, too much dependency on top order and unavailability of proper finishers and all rounders is one of the main reason of RCB not winning any trophy.

Q.10) How would you approach this problem, if the objective and subjective questions weren't given?

Ans:- Without specific objective and subjective questions, here's how I would approach the problem:

### **1. Define the Problem Statement:**

* **Objective**: The primary objective is to build a winning team for RCB that is competitive on the field and provides value for money during the 2017 mega player auction.
* **Sub-objectives**:
  + Identify top-performing players who will contribute to match victories.
  + Ensure that the players are reliable over the course of the tournament.
  + Maximize the team's budget by investing wisely in players who provide value for money (cost vs. performance).

### **2. Data Collection and Preparation:**

Collect relevant data to evaluate player performance, reliability, and auction costs. The key datasets would include:

* **Player Performance Data**:
  + Past performance data such as runs scored, wickets taken, strike rate, bowling economy, batting/bowling averages, and other metrics for both Indian and overseas players.
  + Fielding metrics like catches, run-outs, and overall impact on the field.
  + Consistency data such as performance trends over multiple seasons.
* **Player Availability Data**:
  + Injury history and fitness levels of players (to gauge reliability).
  + International commitments (to ensure the player will be available for the entire tournament).
* **Player Auction Data**:
  + Auction base price and final bid price from past auctions.
  + Historical trends on how certain player types (batsmen, all-rounders, spinners, fast bowlers) are valued in auctions.
* **Team Budget and Composition**:
  + Current team composition, including retained players and their salaries.
  + Remaining auction budget and gaps in the team that need to be filled (e.g., middle-order batsman, death bowler, etc.).

### **3. Exploratory Data Analysis (EDA):**

Perform a thorough analysis of the collected data to extract insights:

* **Player Performance Trends**: Analyze each player's past performance (e.g., runs, wickets, strike rate) to identify top performers and those who have shown consistency over multiple seasons.
* **Cost vs. Performance**: Evaluate players based on their auction price and performance. This will help identify players who are undervalued (high performance at a low cost) or overvalued (low performance at a high cost).
* **Reliability**: Assess player reliability by looking at injury history, number of matches played per season, and consistency in performance.

### **4. Deriving Key Performance Indicators (KPIs):**

From the data, derive key metrics that will guide decision-making during the auction:

* **Player Efficiency Index**: A combined metric that evaluates a player's contribution per dollar spent, balancing performance and auction cost.
* **Consistency Index**: Measures how consistently a player performs across seasons, giving more weight to players with steady performance.
* **Value-for-Money (VFM) Metric**: A ratio of a player’s past auction price to their performance (runs/wickets per match), helping to identify undervalued players.
* **Impact Index**: Measures the overall impact a player has on winning games, combining batting, bowling, and fielding contributions.
* **Injury and Availability Risk Score**: A score based on a player’s injury history, fitness levels, and potential international commitments to assess their reliability over the season.

### **5. Segmentation and Shortlisting of Players:**

Based on the KPIs, segment players into categories:

* **High-Value Players**: Players with top performance, consistency, and relatively lower costs.
* **Core Players**: Reliable, experienced players who form the backbone of the team.
* **Impact Players**: Players who may be expensive but have the potential to win games on their own (e.g., a power-hitting batsman or a death-overs bowler).
* **Value Picks**: Low-cost players with high potential, especially domestic talent, who can be developed over time.
* **Risky Players**: Players with high performance but also high injury or availability risks.

### **6. Formulate Auction Strategy:**

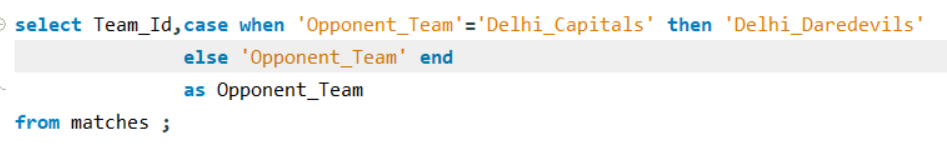
After shortlisting players, devise a strategy for the auction:

* **Prioritize Gaps**: Identify the team's weak spots (e.g., lack of a quality finisher or death bowler) and target players who can fill those gaps.
* **Budget Allocation**: Plan how to allocate the auction budget across player categories (e.g., prioritize top-performing bowlers over batsmen if bowling is the team's weak point).
* **Target Undervalued Players**: Focus on players who are likely to be undervalued in the auction but can deliver strong performances.
* **Minimize Injury Risk**: Avoid bidding aggressively on injury-prone players or those with major international commitments during the tournament.

**Summary:-** By following this structured approach, we can make informed decisions that balance player performance, reliability, and value, helping RCB build a competitive and cost-effective team for the tournament.

Q.11) 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:- I am attaching the screenshot of the query:-



### 