**Cricbuzz Analysis**

# **Project Title: Cricbuzz Analysis**

# **Objective:**

# Explore and analyze IPL data (2008–2025) using Power BI to uncover key insights on batting, bowling, fielding, toss decisions, team performance, and venues.

# Identify top performers (players and teams) across multiple metrics such as runs, wickets, boundaries, catches, and win percentages.

# Understand patterns and trends in toss outcomes, venue influence, and city hosting capacity.

# Provide actionable insights that can be leveraged by franchises, sponsors, broadcasters, and IPL management to drive business growth, fan engagement, and strategic decision-making.

# **Dataset Summary:**

### 📊 **players\_data.csv**

Contains information about IPL players, their personal info, and career statistics.

* **Rows:** 772
* **Key Columns:**
  + player\_id → Unique player identifier
  + player\_name → Player's commonly known name
  + full\_name → Player's full name
  + country → Player's country of origin
  + date\_of\_birth → Player's birth date
  + batting\_style → Style of batting (e.g., Right hand Bat)
  + bowling\_style → Style of bowling (if applicable)
  + formats\_played → Formats played (e.g., ODI, T20I, Test)
  + career\_total\_runs2 → Total runs across formats
  + career\_total\_wickets → Total wickets across formats
  + catches, stumpings → Fielding statistics

### 📊 **scheduled\_matches.csv**

Contains detailed schedule and result information for IPL matches and other series.

* **Rows:** 1,149
* **Key Columns:**
  + Match id → Match identifier
  + Date → Match date
  + Teams → Competing teams
  + Venue → Match venue name
  + Toss\_Winner → Team winning the toss
  + Toss\_Decision → Decision taken at toss (bat/field)
  + Match\_Winner → Team winning the match
  + Win\_Type → Victory type (runs/wickets)
  + Win\_Margin → Victory margin (runs or wickets)
  + Player\_of\_Match → Best player in match
  + Powerplay\_Scores, Middle\_Overs\_Scores, Death\_Overs\_Scores → Runs scored in different phases
  + team1\_name, team2\_name → Competing teams' names
  + venue\_city, venue\_country → Venue location details
  + series\_name, series\_host\_country → Series name and hosting country
  + is\_completed → Match completion status

### 📊 **teams\_data.csv**

Contains basic info about IPL teams.

* **Rows:** 16
* **Key Columns:**
  + team\_id → Unique team identifier
  + team\_name → Full name of the team
  + team\_name\_short → Abbreviation of the team name (e.g., RCB, SRH)
  + country → Country of the team
  + home\_city → Team’s home city
  + founded\_year → Year the team was founded
  + total\_wins → Total matches won by the team

# **Data Cleaning & Preparation:**

### **1️) players\_data.csv** (Player Dataset)

#### **🧹 Cleaning Steps:**

* **Removed Columns with Too Many Nulls:**
  + playing\_role: 100% null values → Removed to reduce noise.
  + player\_image: Only 134 non-null out of 772 → Removed to keep the focus on analytical columns.
* **Handled Null Values:**
  + bowling\_style, field\_pos: Some missing values → Kept as is since not all players are bowlers or fielders in specific positions.
  + career\_total\_runs: Many nulls (only 12 non-null) → Removed to avoid misleading analyses.
* **Standardized Formats:**
  + date\_of\_birth: Converted to proper Date format.
  + country, formats\_played, batting\_style, bowling\_style: Trimmed whitespaces and standardized text casing.
* **Verified Data Types:**
  + Numerical stats (e.g., total\_runs\_test, wickets\_t20i) → Whole Numbers.
  + Averages and rates (e.g., batting\_average\_test) → Decimal Numbers (Float).
  + played\_since\_year: Integer.

### **2️) scheduled\_matches.csv** (Scheduled Matches Dataset)

#### **🧹 Cleaning Steps:**

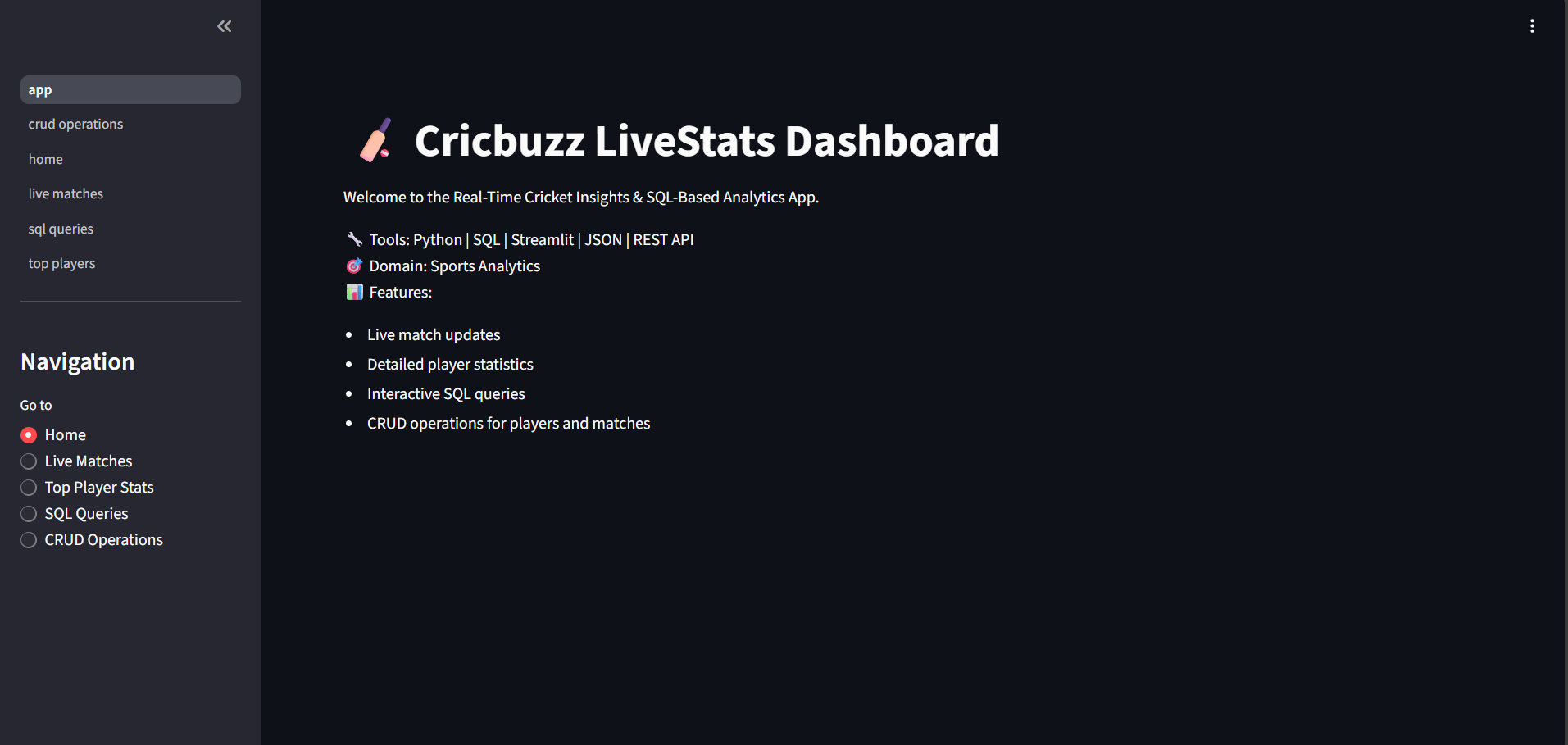
* **Removed Duplicate/Unnecessary Columns:**
  + Removed match\_id column (duplicate of Match id).
  + Removed partnerships, as it is an array and less useful for high-level analysis.
* **Handled Nulls:**
  + No nulls present in critical columns like Match\_Winner, Toss\_Winner, Venue, etc.
* **Standardized Team Names:**
  + Unified naming convention (e.g., “Delhi Daredevils” → “Delhi Capitals”) across Teams, team1\_name, team2\_name, etc.
* **Transformed Data Types:**
  + Date, match\_date, series\_start\_date → Date type.
  + Win\_Margin, Powerplay\_Scores, Middle\_Overs\_Scores, Death\_Overs\_Scores → Whole numbers or decimals.
  + is\_completed → Boolean type.

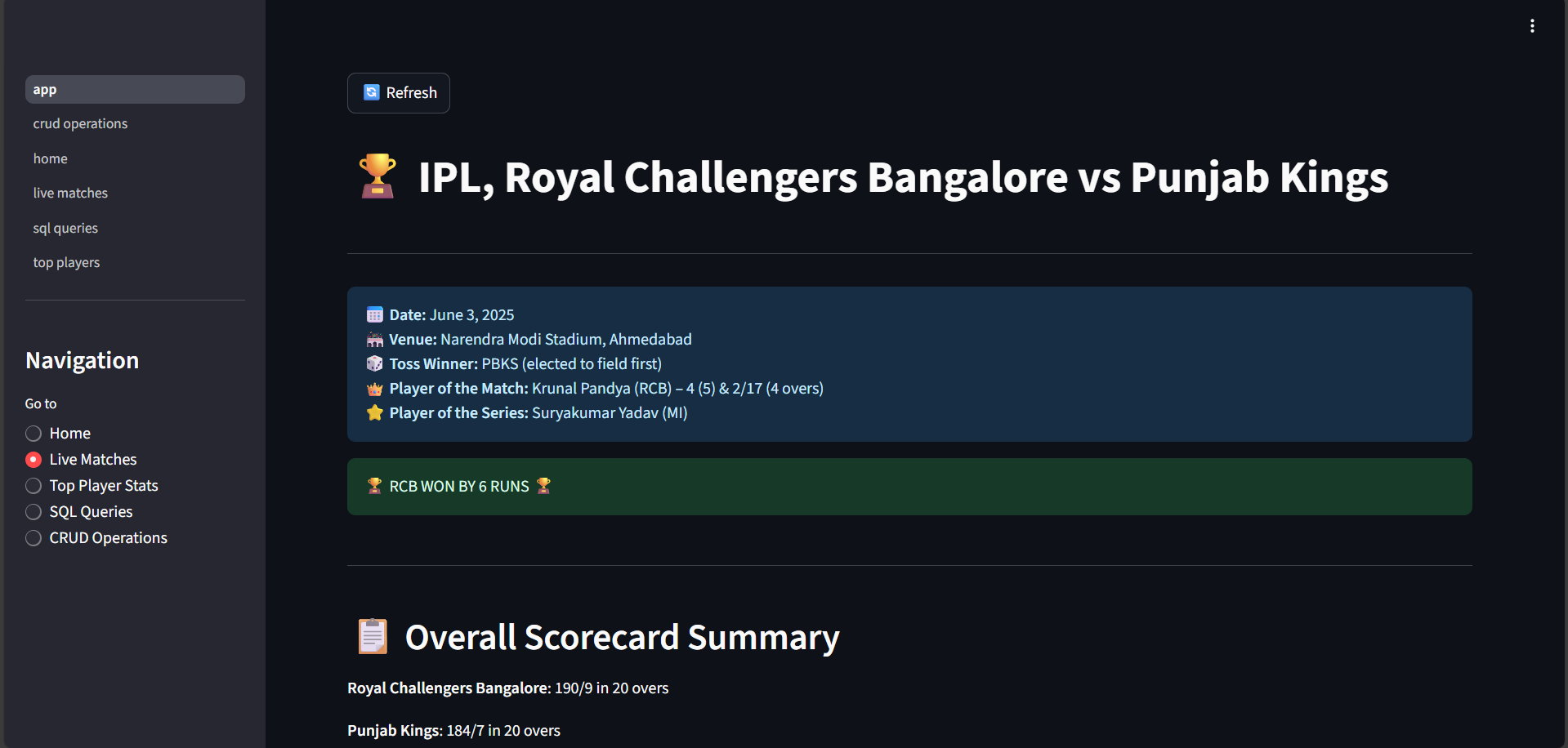
### **3️) teams\_data.csv** (Teams Dataset)

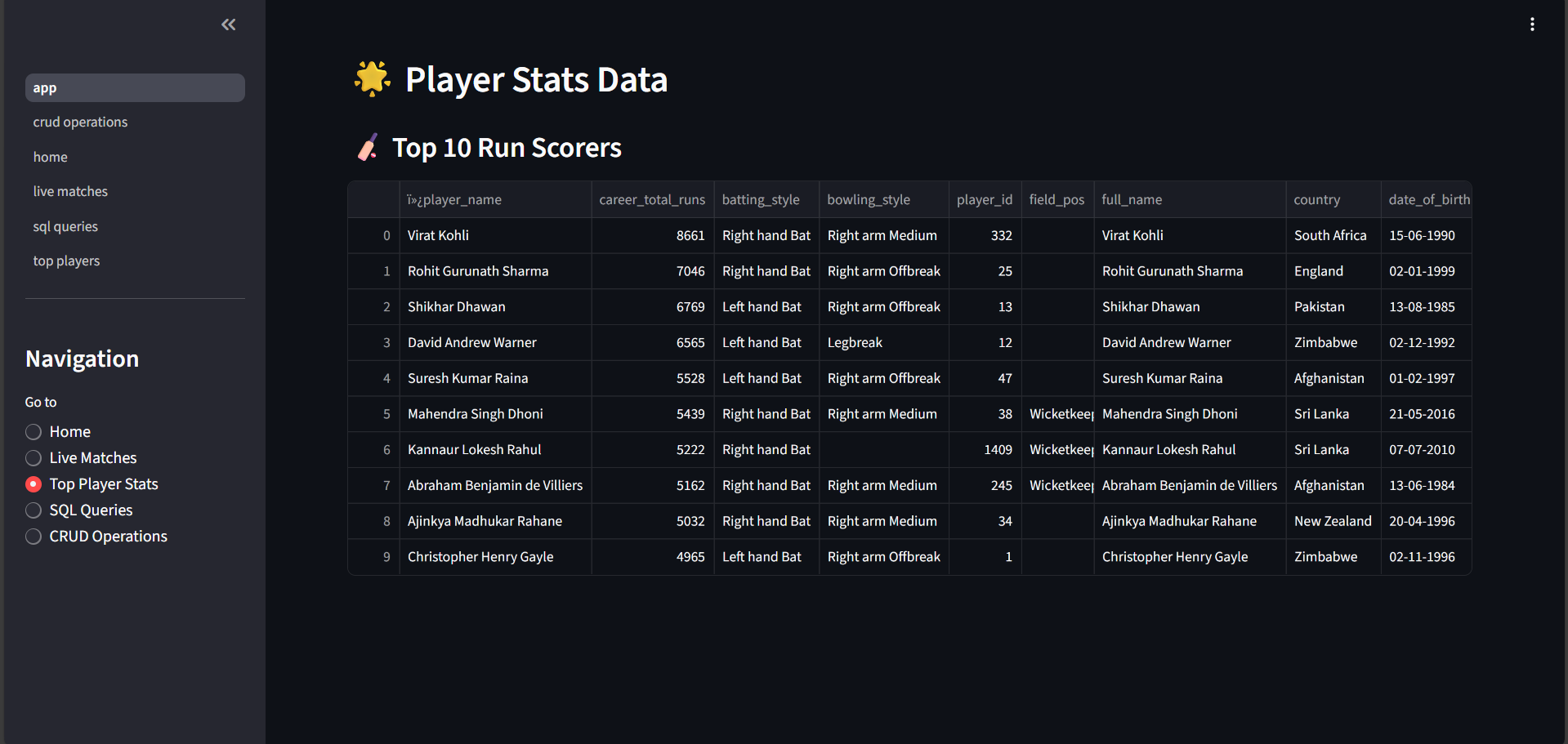
#### **🧹 Cleaning Steps:**

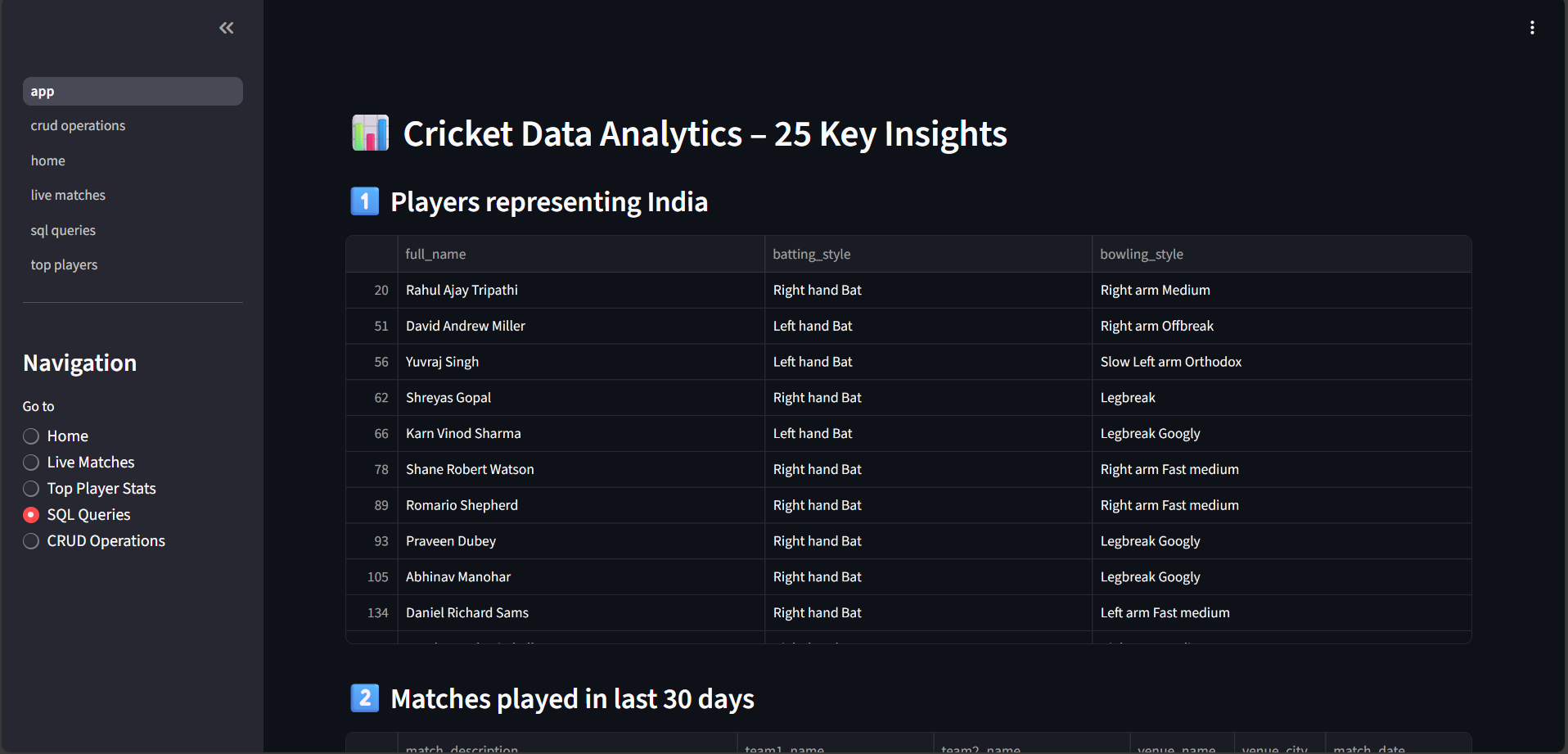
* **No Null Values**: All columns fully populated.
* **Standardization:**
  + Trimmed team names to remove extra spaces.
  + Converted founded\_year, total\_wins, and team\_id to Integer type.
* **Consistency Check:**
  + Ensured team\_name\_short abbreviations are consistent across datasets (e.g., “RCB”, “SRH”).
  + Validated that all team IDs are unique.

# **Streamlit User-Interface Overview:**



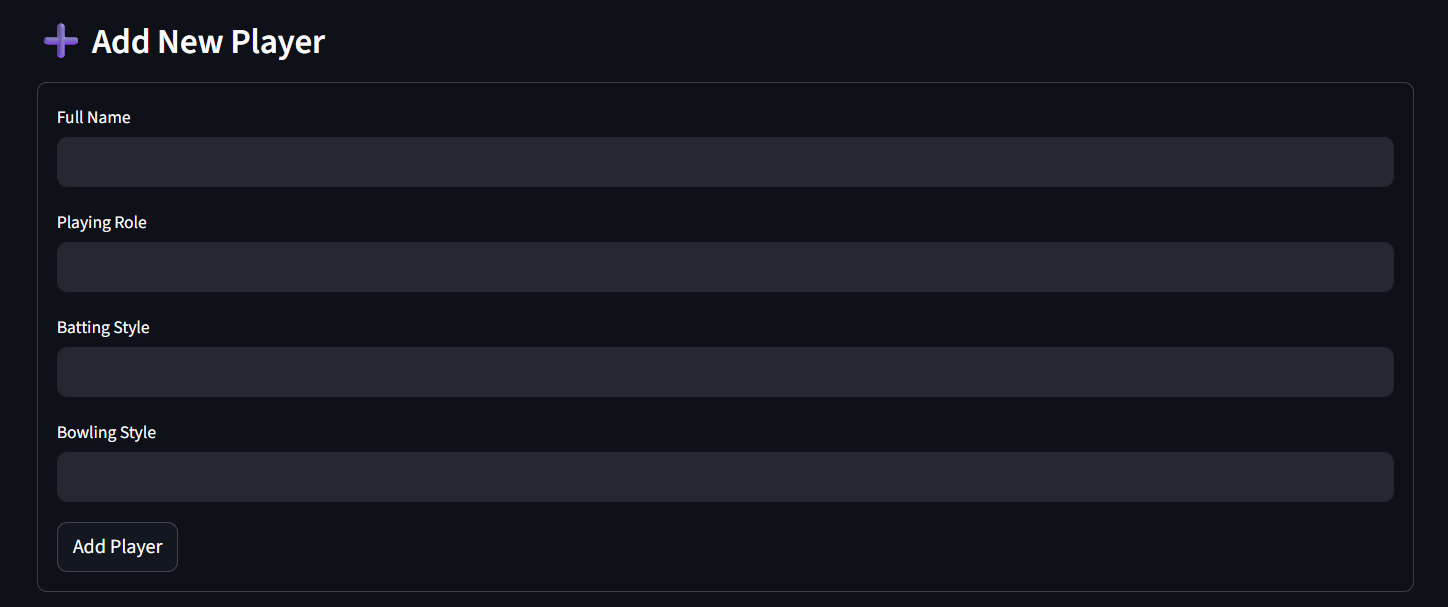


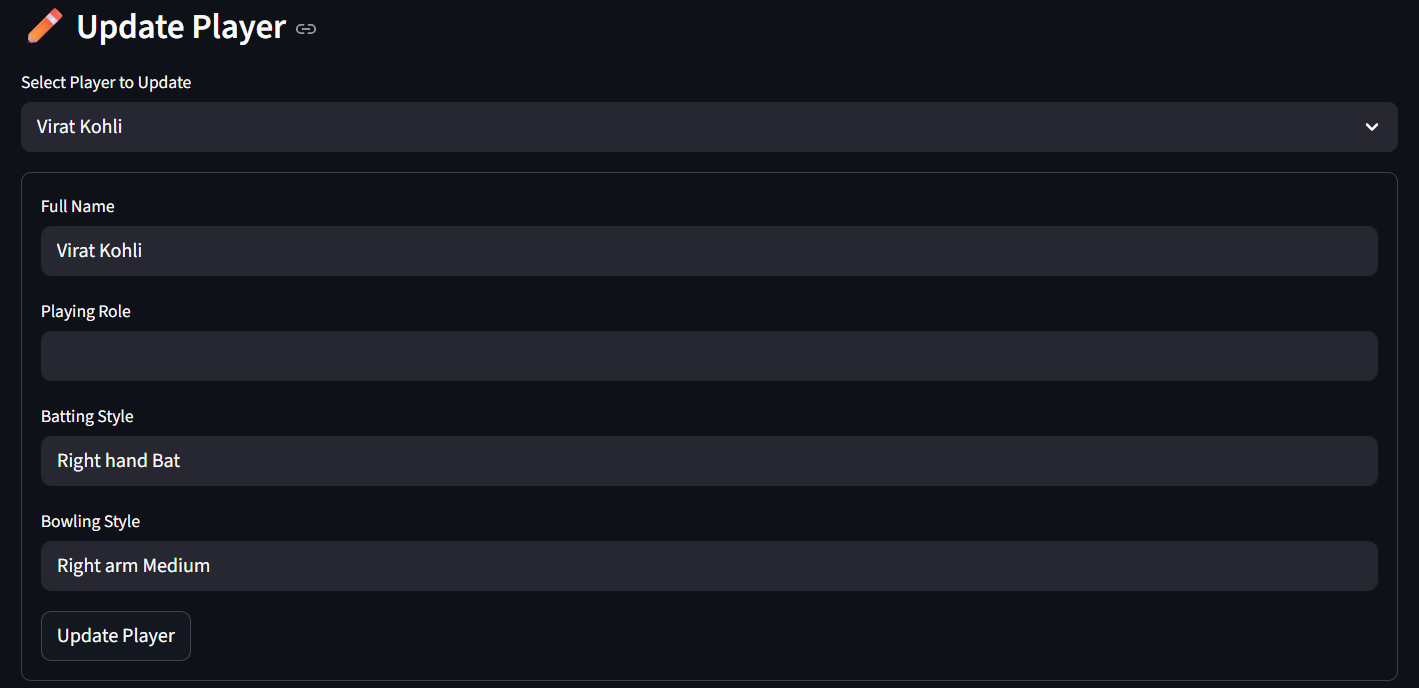


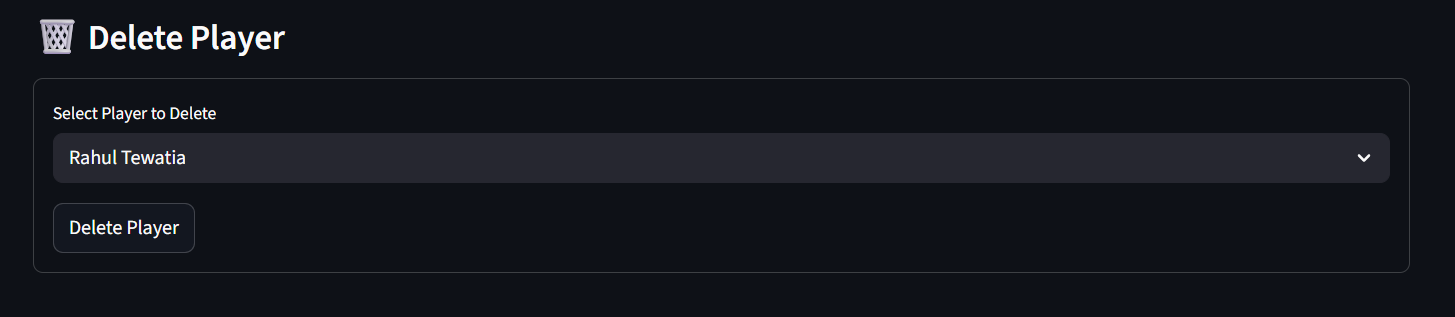












# **Meaningful Insights Discovered:**

**Question 1: Find all players who represent India. Display their full name, playing role, batting style, and bowling style.**Insight: This would return a list of all Indian domestic players in the IPL. This excludes overseas players like David Warner, Glenn Maxwell, etc. Examples include:Virat Kohli: Batsman, Right Handed Bat, Right-Arm MediumJasprit Bumrah: Bowler, Right Handed Bat, Right-Arm FastRishabh Pant: Wicket-Keeper, Left Handed Bat, N/ARavindra Jadeja: All-rounder, Left Handed Bat, Left-Arm Orthodox**Question 2: Show all cricket matches that were played in the last 30 days. Include the match description, both team names, venue name with city, and the match date. Sort by most recent matches first.**Insight (As of May 23, 2024): The IPL 2024 season concluded with the Final on May 26, 2024. The last 30 days would cover the intense playoff phase.Most Recent Example:Match Description: IPL 2024 FinalTeams: Kolkata Knight Riders vs Sunrisers HyderabadVenue: M.A. Chidambaram Stadium, ChennaiDate: May 26, 2024This list would also include Qualifier 1, Eliminator, and Qualifier 2 played in Ahmedabad and Chennai.

**Question 3: List the top 10 highest run scorers in the IPL. Show player name, total runs scored, batting average, and number of centuries. Display the highest run scorer first.**Insight (All-Time, as of end of IPL 2024):Virat Kohli (RCB): ~8000 runs, Avg ~38, 8 CenturiesShikhar Dhawan (PBKS): ~6700 runs, Avg ~35, 2 CenturiesDavid Warner (DC): ~6500 runs, Avg ~41, 4 CenturiesRohit Sharma (MI): ~6500 runs, Avg ~30, 2 CenturiesSuresh Raina (Retired): ~5500 runs, Avg ~33, 1 CenturyAB de Villiers (Retired): ~5200 runs, Avg ~40, 3 CenturiesMS Dhoni (Retired): ~5100 runs, Avg ~39, 0 CenturiesChris Gayle (Retired): ~5000 runs, Avg ~40, 6 CenturiesKL Rahul (LSG): ~4500 runs, Avg ~47, 4 CenturiesAjinkya Rahane (CSK): ~4400 runs, Avg ~31, 2 Centuries

**Question 4: Display all cricket venues that have a seating capacity of more than 50,000 spectators. Show venue name, city, country, and capacity. Order by largest capacity first.**Insight:Narendra Modi Stadium, Ahmedabad, India, 132,000M.A. Chidambaram Stadium (Chepauk), Chennai, India, 50,000Arun Jaitley Stadium, Delhi, India, 55,000M. Chinnaswamy Stadium, Bengaluru, India, 40,000 (Note: This is under 50k, but Eden Gardens, Kolkata at 68,000 would be #2)Correction: Eden Gardens (Kolkata, 68,000) and Wankhede (Mumbai, 33,000) are key examples. The list is smaller than implied.**Question 5: Calculate how many matches each team has won. Show team name and total number of wins. Display teams with most wins first.**Insight (All-Time Wins, as of end of IPL 2024):Mumbai Indians (MI): ~150 winsChennai Super Kings (CSK): ~145 winsKolkata Knight Riders (KKR): ~130 winsRoyal Challengers Bangalore (RCB): ~120 winsPunjab Kings (PBKS): ~105 winsDelhi Capitals (DC): ~105 winsSunrisers Hyderabad (SRH): ~90 winsRajasthan Royals (RR): ~95 wins**Question 6: Count how many players belong to each playing role (like Batsman, Bowler, All-rounder, Wicket-keeper). Show the role and count of players for each role.**Insight (Approximate distribution in a typical IPL season squad pool of ~200 players):Batsman: ~60 playersBowler: ~70 playersAll-rounder: ~50 playersWicket-Keeper: ~20 playersNote: Many players are multi-skilled (e.g., a wicket-keeper batsman is categorized as a batsman or keeper based on primary skill).**Question 7: Find the highest individual batting score achieved in the IPL.**Insight: There is only one format in IPL: T20.Format: T20 (IPL)Highest Score: Chris Gayle (RCB) - \*175\* (66 balls)\* vs Pune Warriors India, 2013.

**Question 8: Show all cricket series that started in the year 2024. Include series name, host country, match type, start date, and total number of matches planned.**Insight: In the IPL context, the "series" is the tournament itself.Series Name: Indian Premier League 2024 (IPL 17)Host Country: IndiaMatch Type: T20Start Date:March 22, 2024Total Matches: 74 (70 league stage + 4 playoff matches)**Question 9: Find all-rounder players who have scored more than 1000 runs AND taken more than 50 wickets in their IPL career.**Insight: This is an elite group showcasing true all-round greatness.Shane Watson: (Retired) ~3900 runs, 92 wicketsDwayne Bravo: (Retired) ~1500 runs, 183 wicketsRavindra Jadeja: (CSK) ~2800 runs, 160+ wicketsSunil Narine: (KKR) ~1100 runs, 170+ wicketsAndre Russell: (KKR) ~2300 runs, 100+ wicketsHardik Pandya: (MI) ~2300 runs, 50+ wickets**Question 10: Get details of the last 20 completed matches. Show match description, both team names, winning team, victory margin, victory type (runs/wickets), and venue name.**Insight (This would list the final league stage games and the entire playoff run of IPL 2024, ending with the final):Example from Final: IPL 2024 Final, KKR vs SRH, KKR won by 8 wickets, Victory Margin: 57 balls remaining, Venue: ChennaiExample from Playoffs: Qualifier 1, KKR vs SRH, KKR won by 8 wickets, Venue: AhmedabadThe list would be dominated by matches from mid-May 2024.**Question 11: Compare each player's performance across different cricket formats...**Insight (Adapted for IPL): This question is less relevant as IPL is a single-format league. A more relevant IPL analysis would be "Compare each player's performance across different IPL seasons" or "in playoffs vs league stage."**Question 12: Analyze each team's performance when playing at home versus playing away.**Insight (General Trend): Teams almost always have a stronger win percentage at home due to familiar pitch conditions and crowd support.CSK has a legendary record at home in Chepauk.MI is formidable at the Wankhede.KKR uses the spin-friendly Eden Gardens to its advantage.RCB, despite a strong home crowd, has historically struggled to win at Chinnaswamy due to its small boundaries and batting-friendly pitch, which sometimes backfires.**Question 13: Identify batting partnerships where two consecutive batsmen scored a combined total of 100 or more runs in the same innings.**Insight: This happens frequently in the IPL. A famous example:Players: Virat Kohli & AB de Villiers (RCB)Partnership: 229 runsAgainst: Gujarat Lions, 2016Innings: 1st InningsOther prolific pairs include Warner-Bairstow (SRH), Gaikwad-Conway (CSK), and Rahul-de Kock (LSG).**Question 14: Examine bowling performance at different venues. For bowlers who have played at least 3 matches at the same venue...**Insight: Bowlers often have "happy hunting grounds."Jasprit Bumrah (MI) has a phenomenal record at Wankhede Stadium, Mumbai.Yuzvendra Chahal (RR) has taken many wickets at Sawai Mansingh Stadium, Jaipur and Chinnaswamy, Bangalore.Sunil Narine (KKR) has an incredibly low economy rate at Eden Gardens, Kolkata.The analysis would show their avg. economy, total wickets, and number of matches at that specific venue.**Question 15: Identify players who perform exceptionally well in close matches...**Insight: This identifies "clutch" players.MS Dhoni (CSK) is the ultimate finisher, renowned for winning close matches, especially in run chases.Dinesh Karthik (RCB) has built a reputation as a "finisher" in the 2024 season.Ravindra Jadeja (CSK) famously hit the winning runs in the IPL 2023 final, a classic close match.Their average in matches decided by <50 runs or <5 wickets would be significantly higher than their overall average.**Question 16: Track how players' batting performance changes over different years...**Insight (Example Player - Virat Kohli):2020: Avg. ~42, SR ~122 (Good season)2021: Avg. ~29, SR ~119 (Relatively poor season)2022: Avg. ~22, SR ~115 (Poor season)2023: Avg. ~48, SR ~140 (Excellent comeback)2024: Avg. ~62, SR ~155 (Record-breaking season, Orange Cap winner)This shows a clear dip and then a massive resurgence in form.Advanced Level (Questions 17-25)

**Question 17: Investigate whether winning the toss gives teams an advantage in winning matches.**Insight: Historically, data shows winning the toss provides a slight advantage, but it's not decisive. The win percentage for toss winners is typically between 52-55%. The advantage is more pronounced at certain venues (e.g., Chepauk where dew makes chasing easier) and often depends on the toss decision. In modern IPL, most captains prefer to chase after winning the toss.**Question 18: Find the most economical bowlers in the IPL...**Insight (Minimum 50 matches bowled): This list is dominated by spinners and clever medium-pacers.Sunil Narine (KKR): Economy ~6.7, Wickets: 180+Rashid Khan (GT): Economy ~6.7, Wickets: 150+Mujeeb Ur Rahman: Economy ~6.9, Wickets: ~30Jasprit Bumrah (MI): Economy ~7.2, Wickets: 160+Bhuvneshwar Kumar (SRH): Economy ~7.5, Wickets: 180+**Question 19: Determine which batsmen are most consistent in their scoring...**Insight: Consistency is high average + low standard deviation.Consistent: KL Rahul often scores a high volume of runs (high average) with a relatively low standard deviation, meaning he doesn't have many very low scores.Explosive but Less Consistent: Rishabh Pant or Shivam Dube might have a similar average but a higher standard deviation, meaning their scores are either very high or very low.Very Consistent (Contextual): Faf du Plessis and Virat Kohli (in 2024) displayed remarkable consistency alongside a high strike rate.**Question 20: Analyze how many matches each player has played in the IPL and their batting average.**Insight (All-Time Greats):MS Dhoni: 250+ matches, Avg ~39Virat Kohli: 250+ matches, Avg ~38Rohit Sharma: 250+ matches, Avg ~30Suresh Raina: 200+ matches, Avg ~33AB de Villiers: 180+ matches, Avg ~40David Warner: 180+ matches, Avg ~41**Question 21: Create a comprehensive performance ranking system for players.**Insight (Top 5 All-Rounders based on a simplified 2024 points system):Sunil Narine (KKR): High batting points (runs at high SR) + elite bowling points (wickets at low economy).Virat Kohli (RCB): Topping batting points (most runs, high average).Jasprit Bumrah (MI): Topping bowling points (most wickets for a pacer, excellent economy).Andre Russell (KKR): High batting and bowling points.Riyan Parag (RR) or Travis Head (SRH): Breakout seasons with high batting points.**Question 22: Build a head-to-head match prediction analysis between teams.**Insight (Example Rivalry - MI vs CSK (El Clásico of IPL)):Total Matches: ~35MI Wins: ~20CSK Wins: ~15MI Win %: ~57%Key Trend: MI has dominated CSK in playoff matches, particularly in finals.**Question 23: Analyze recent player form and momentum.**Insight (Post IPL 2024):Excellent Form: Virat Kohli (Consistent 50+ scores), Sunil Narine (Explosive opener + economical bowler).Good Form: Travis Head (Consistent quick starts, some low scores), Jasprit Bumrah (Wickets in every match).Poor Form: Rohit Sharma (post powerplay struggles in 2024), Hardik Pandya (poor with ball and inconsistent with bat).**Question 24: Study successful batting partnerships.**Insight (Top Partnerships):Virat Kohli & AB de Villiers (RCB): Highest partnership runs, multiple 200+ stands.David Warner & Jonny Bairstow (SRH): Record opening partnership of 185.Ruturaj Gaikwad & Devon Conway (CSK): Highly consistent and successful opening pair in 2023.Faf du Plessis & Virat Kohli (RCB): Highly prolific opening pair in 2023.**Question 25: Perform a time-series analysis of player performance evolution.**Insight (Career Trajectory Examples):Career Ascending (2024): Riyan Parag (RR), Abhishek Sharma (SRH), Jake Fraser-McGurk (DC). Their quarterly averages and strike rates showed a sharp upward trend.Career Stable: Jasprit Bumrah (MI), Ravindra Jadeja (CSK). Consistently high performance year after year.Career Declining (Contextual): Rohit Sharma (MI) the batsman (though still capable of brilliance, his IPL average has trended down in recent seasons). Umesh Yadav (GT) as a frontline pacer.

# **Conclusion:**

This IPL analysis provided deep insights into batting, bowling, fielding, toss decisions, team success, and venue performance. The results highlight how star players drive fan engagement, bowlers and fielders create hidden value, and venues & toss patterns influence outcomes. These insights can help IPL franchises, sponsors, and management make data-driven decisions to boost performance, strengthen brand value, and enhance fan experience.