

# IPL Data Analysis & Dashboard Report

Om Jadhav | Data Analyst | AI-powered Analytics Specialist

## Executive Summary

This project delivers a full-stack analytical solution for the Indian Premier League (IPL) 2023 season, transforming granular ball-by-ball data into strategic insights. It includes robust data cleaning, relationship modelling, DAX-driven KPIs, and an interactive dashboard built for performance benchmarking, player evaluation, and decision support.

---

## Data Architecture & Cleaning

### Datasets Used

- matches.csv: Match-level metadata (teams, venue, toss, result)
- deliveries.csv: Ball-by-ball data (runs, wickets, player actions)

### Cleaning & Standardization

- **Team Name Normalization:**
    - Merged inconsistent entries like "Royal Challengers Bengaluru" and "Royal Challengers Bangalore" into a unified "Royal Challengers Bangalore"
    - Standardized "Pune Warriors" and "Rising Pune Supergiant" under "Pune Team" for consistency
  - **Null Handling:**
    - Removed rows with missing player names or match IDs
    - Imputed missing values where applicable using domain logic
  - **Data Types:**
    - Converted date columns to datetime format
    - Ensured numeric consistency for runs, wickets, and ball counts
- 

## Data Modeling

### Relationships

- **One-to-Many:**
  - matches (one) → deliveries (many) via match\_id
- **Cross Filtering:**
  - Enabled bi-directional filters for dynamic slicers (team, player, venue)
- **Calculated Columns:**

- Created derived fields for innings, over number, and dismissal types

---

## DAX Measures & KPIs

### Batsman Metrics

Batsman Strike Rate =

$(\text{SUM}(\text{deliveries}[\text{batsman\_runs}]) / \text{COUNT}(\text{deliveries}[\text{ball}])) * 100$

### Bowler Metrics

Bowler Economy =

$\text{VAR over\_} = \text{COUNT}(\text{deliveries}[\text{ball}]) / 6$

$\text{RETURN SUM}(\text{deliveries}[\text{total\_runs}]) / \text{over\_}$

Bowler Strike Rate =

$\text{COUNT}(\text{deliveries}[\text{ball}]) / \text{SUM}(\text{deliveries}[\text{is\_wicket}])$

### Cap Holders

Orange Cap Holder Runs =

$\text{CALCULATE}(\text{SUM}(\text{deliveries}[\text{batsman\_runs}])) \& \text{" Runs"}$

Purple Cap Holder Wickets =

$\text{CALCULATE}(\text{SUM}(\text{deliveries}[\text{is\_wicket}])) \& \text{" Wickets"}$

These measures powered dynamic cards and leaderboard visuals in the dashboard.

---

## Dashboard Highlights (For year 2023) :

### Tournament Summary

- **Winner:** Chennai Super Kings
- **Total Runs:** 25,688
- **Total Wickets:** 916
- **Total Sixes:** 2,175

### Player Insights

- **Top Batsmen:** Shubman Gill (890 runs), Virat Kohli, Faf du Plessis
- **Top Bowlers:** Mohit Sharma (31 wickets), Rashid Khan, Siraj
- **Player of the Match Leaders:** Gill & Kohli (4 awards each)

### Strategic Visuals

- **Toss Decision vs Win %:** Pie chart showing impact of batting/fielding first
  - **Team Wins:** Bar chart comparing win counts across franchises
  - **Performance Cards:** Dynamic KPIs for Orange/Purple Cap holders, strike rates, economies
- 

## Impact & Applications

- **Recruiter-Ready Portfolio:** Demonstrates advanced DAX, relationship modeling, and dashboarding
- **Sports Analytics Use Case:** Enables performance tracking and match strategy evaluation
- **Scalable Framework:** Can be extended to multi-season analysis or predictive modeling

## Achievements :

### Problem

The IPL 2023 dataset posed several analytical challenges typical of large-scale sports data: fragmented ball-by-ball and match-level tables, lack of pre-calculated performance metrics, and the need for dynamic, role-specific insights across players and teams. Additionally, the raw data lacked clarity in key areas such as strike rates, economy rates, and cap-holder identification, making it unsuitable for direct visualization or decision-making.

### Solution

To address these gaps, I designed a robust data model using Power BI, establishing one-to-many relationships between match and delivery datasets. I implemented custom DAX measures to calculate advanced KPIs including batsman strike rate, bowler economy, bowler strike rate, and dynamic cap-holder stats. These measures enabled real-time performance tracking and comparative analysis across roles. The dashboard was structured with modular visuals—KPI cards, leaderboards, pie charts, and filters—allowing users to explore trends interactively and intuitively.

### Result

The final dashboard delivers a comprehensive analytical view of IPL 2023, transforming raw data into actionable insights for fans, analysts, and decision-makers. It highlights top performers, strategic toss decisions, and match outcomes with precision, while showcasing technical depth in data modelling, DAX logic, and visual storytelling. This project not only solves the problem of fragmented sports data but also demonstrates my ability to build scalable, recruiter-ready analytics solutions with real-world impact.