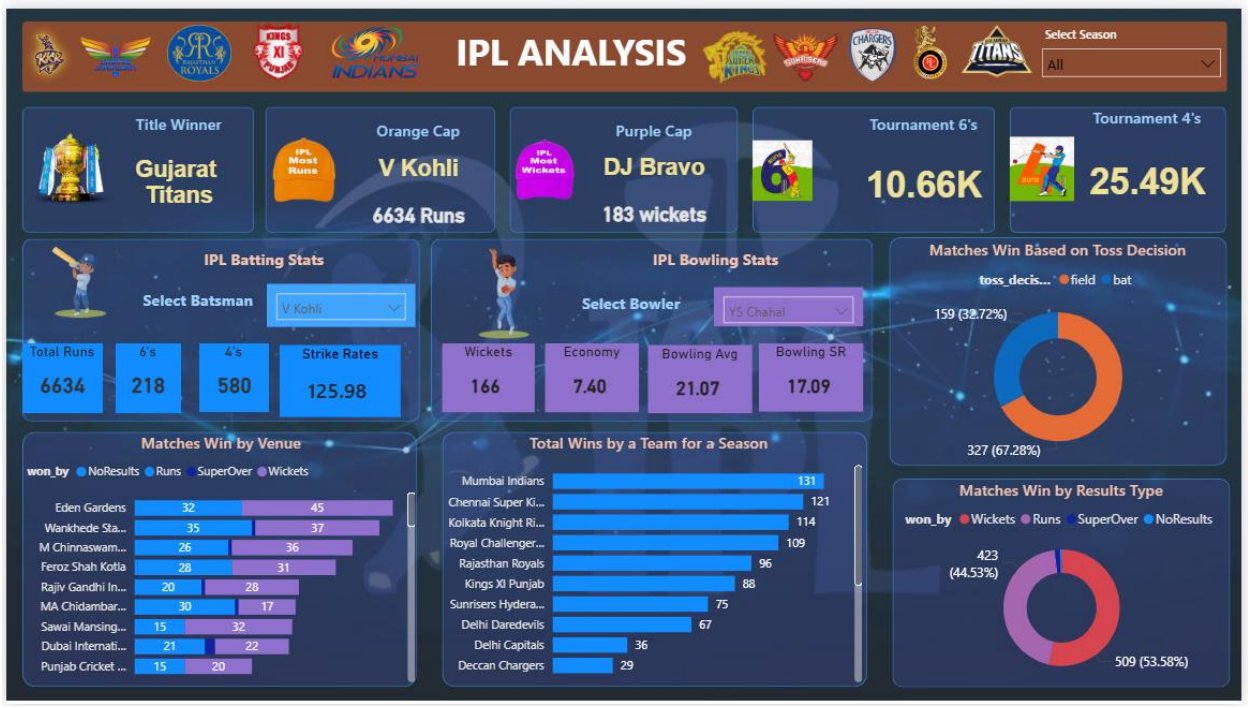
IPL Cricket Analysis (2008–2022) – Full Project Report

1. Project Overview

This project presents a comprehensive analysis of the Indian Premier League (IPL) from 2008 to 2022 using SQL, Python, and Power BI.

It covers ball-by-ball and match-level datasets to extract meaningful trends about batting, bowling, teams, venues, and tournament patterns.

An interactive dashboard was developed to give users deep insights into 15 years of IPL history.



2. Objectives

The key goals of the project were:

- To clean and structure IPL data using SQL & Python
- To perform exploratory data analysis (EDA) on match and ball-by-ball records
- To build DAX measures for advanced cricket statistics
- To visualize trends through an interactive Power BI dashboard
- To identify top performers, winning patterns, and venue-based behaviors

3. Datasets Used

1) **IPL Matches Dataset (2008–2022)**

   Contains match-level information such as:

   - Teams, venues, toss details

- Match results (runs / wickets)
- Man of the Match
- Umpires, seasons

2) **IPL Ball-by-Ball Dataset (2008–2022)**
   Contains:
   - Every ball bowled
   - Batter, bowler, runs scored
   - Wickets, extras, dismissals
   - Over-by-over match progress

4. Tools & Technologies

- **SQL (PostgreSQL):** Data loading, cleaning, joins, aggregations
- **Python:** Data preprocessing, feature engineering
- **Power BI:** Visualization, DAX measures, slicers, interactions

5. Data Cleaning & Transformation

SQL Steps

- Removed duplicates and null values
- Standardized team names
- Joined matches & ball-by-ball tables
- Calculated match aggregates:
  - Total runs per player
  - Wickets per bowler
  - Venue-based win patterns
  - Toss decision impact

Python Steps

- Outlier detection
- Data formatting for Power BI
- Exported cleaned datasets as CSV for import

6. DAX Measures in Power BI

Key measures created:

Batting

- **Total Runs**
- **Strike Rate**
- **Total 4s and 6s**
- **Average Runs per Innings**

Bowling

- **Total Wickets**
- **Economy Rate**
- **Bowling Average**
- **Strike Rate**

Team & Tournament

- **Wins by venue**
- **Wins by season**
- **Match results by type (runs/wickets/super over)**
- **Toss impact %**

7. Dashboard Features (From Screenshot)

I. Title, Orange Cap, Purple Cap

- **Gujarat Titans** as champions
- **V Kohli** with **6634 runs (Orange Cap)**
- **DJ Bravo** with **183 wickets (Purple Cap)**

II. Batting Analysis Panel

- Select any batsman from dropdown
- Shows:
  - Total Runs
  - 4s, 6s
  - Strike Rate

III. Bowling Analysis Panel

- Select any bowler
- Displays:
  - Wickets
  - Bowling Average
  - Economy
  - Bowling Strike Rate

IV. . Tournament Totals

- **10.66K total sixes**
- **25.49K total fours**

V. Toss Decision vs Match Result

- Pie chart showing % of wins after batting/fielding first
- Insight: Fielding-first gives a slight advantage in IPL

VI. Venue-Based Analysis

Shows wins based on:

- Runs

- Wickets
- Super Over
- No Result

For each major stadium (Eden Gardens, Wankhede, Chinnaswamy, etc.)

VII. Team Performance across Seasons

- Mumbai Indians dominate with **131 wins**
- Chennai Super Kings follow with **121 wins**
- Other teams ranked accordingly

VIII. Win by Result Type

Pie chart showing:

- Wins by wickets
- Wins by runs
- Super over wins
- No result matches

8. Key Insights

a. Batting Insights

- Virat Kohli is the most consistent scorer from 2008 to 2022
- Strike rates and boundary counts highlight aggressive players

b. Bowling Insights

- Bravo leads wickets but Chahal and Malinga show better economy and SR
- Wrist spinners dominate the wicket tally

c. Team Insights

- Mumbai Indians are the most successful franchise ever
- Toss decisions significantly influence match outcomes

d. Venue Insights

- Wankhede and Eden Gardens deliver the highest number of close finishes
- Certain grounds favor pace vs spin bowling

9. Conclusion

This project offers a **complete analytical view** of IPL history using modern data engineering and BI tools.

The final dashboard helps stakeholders understand:

- Player form
- Team consistency
- Venue effects
- Tournament trends

It acts as a powerful tool for analysts, cricket fans, and fantasy league strategists.

10. Future Enhancements

- Predictive modeling (Match winner prediction using ML)
- Player comparison tool
- Live API integration for real-time dashboards