IPL Data Analysis (2008–2025) Using Power BI

Student Name: MD. Shakawat Hossain Faravi

University: Green University of Bangladesh

Department: BSc in Computer Science and Engineering

Date: August 13, 2025

Abstract

This report presents a comprehensive analysis of IPL matches from 2008 to 2025 using Power BI. The project includes end-to-end data processing, data modeling, DAX measures, KPI calculation, and dashboard visualization for season, team, and player performance. Key metrics include Orange Cap, Purple Cap, top 4s/6s, matches won/lost, and overall team trends. The insights provide a detailed understanding of player and team performance across 17 IPL seasons.

Chapter 1

Project Overview

The Indian Premier League (IPL) is one of the most popular T20 cricket leagues globally. This project analyzes 17 seasons of IPL data to provide insights into team performance, player statistics, and match outcomes. The main objective is to create an interactive Power BI dashboard that provides dynamic insights to the user.

Objectives

- Analyze season-wise, team-wise, and player-wise performance.
- Build a professional interactive Power BI dashboard.
- Calculate key metrics: Orange Cap, Purple Cap, top 4s/6s, matches won/lost, and total points.
- Demonstrate best practices in data visualization and dashboard design.

Chapter 2

Data Sources

- **IPL Matches Data:** season, match_id, team1, team2, match_date, match_winner, venue, result.
- **Ball-by-Ball Data:** match_id, batter, bowler, batter_runs, extras, is_wicket, wicket_kind.
- **Teams Data:** team_name, image_url.
- **Players Data:** player_name, player_image.

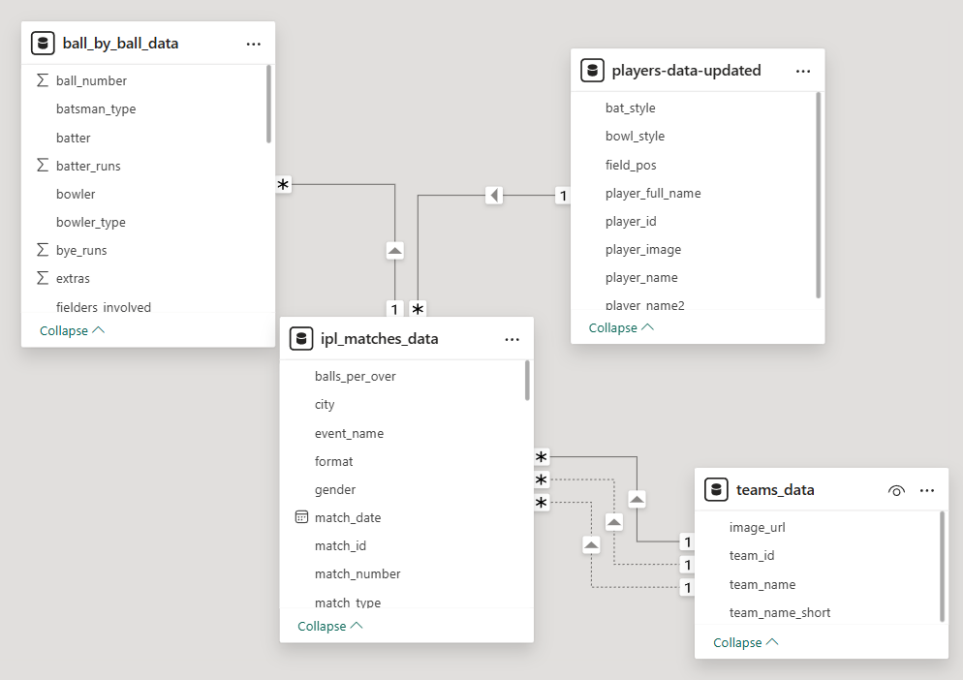

Figure 2.1: Power BI Data Model Relationships

Chapter 3

Data Preparation and Cleaning

- Imported datasets into Power BI using Power Query.
- Removed duplicates and blanks.
- Handled missing values and corrected data types.
- Created relationships between tables.

Problem: Some ball-by-ball rows lacked season information.
Solution: Used the `RELATED()` function in DAX to bring season info from the match table.

Chapter 4

Data Modeling

- Star schema design for efficient analysis.
- Fact table: `ball_by_ball_data`.
- Dimension tables: `ipl_matches_data`, `teams_data`, `players_data`.

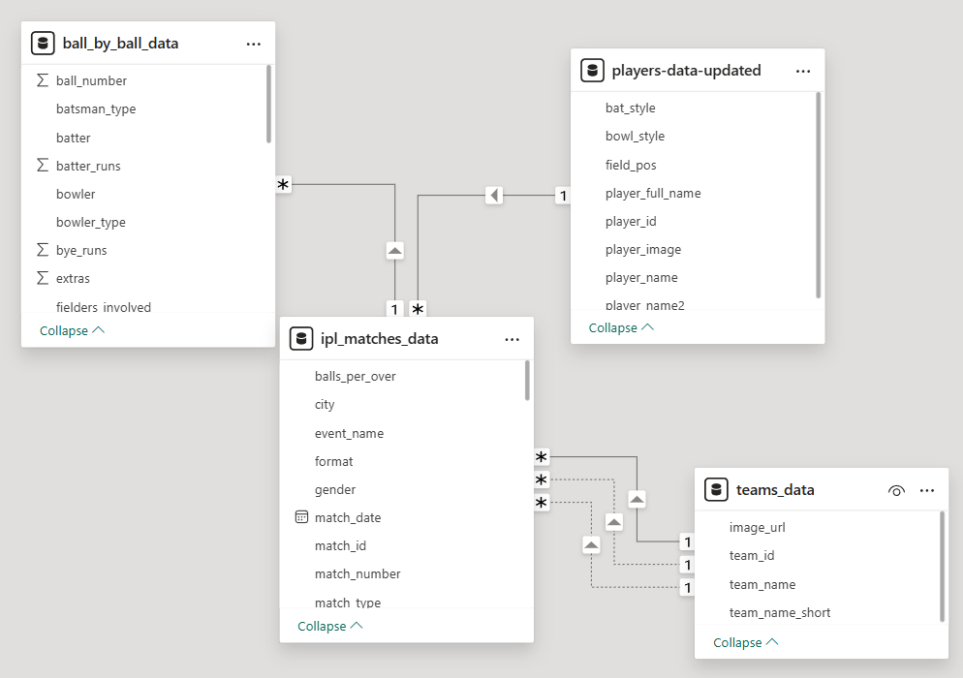

Figure 4.1: Star Schema for IPL Data

Chapter 5

DAX Measures and KPIs

5.1 Season Winner and Runner-Up

```
Season Winner =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
VAR FinalMatchDate = CALCULATE(MAX(ipl_matches_data[match_date]),
    ipl_matches_data[season] = SelectedSeason)
VAR FinalMatchWinner = CALCULATE(MAX(ipl_matches_data[match_winner]),
    ipl_matches_data[season] = SelectedSeason, ipl_matches_data[
    match_date] = FinalMatchDate)
RETURN FinalMatchWinner

Runner_Up =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
VAR FinalMatchDate = CALCULATE(MAX(ipl_matches_data[match_date]),
    ipl_matches_data[season] = SelectedSeason)
VAR FinalMatchWinner = CALCULATE(MAX(ipl_matches_data[match_winner]),
    ipl_matches_data[season] = SelectedSeason, ipl_matches_data[
    match_date] = FinalMatchDate)
VAR Team1 = CALCULATE(MAX(ipl_matches_data[team1]), ipl_matches_data[
    season] = SelectedSeason, ipl_matches_data[match_date] =
    FinalMatchDate)
VAR Team2 = CALCULATE(MAX(ipl_matches_data[team2]), ipl_matches_data[
    season] = SelectedSeason, ipl_matches_data[match_date] =
    FinalMatchDate)
RETURN IF(FinalMatchWinner = Team1, Team2, Team1)
```

5.2 Player Metrics

```
Centuries =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
VAR SeasonData = FILTER(ball_by_ball_data, RELATED(ipl_matches_data[
    season]) = SelectedSeason)
VAR BatterRuns = SUMMARIZE(SeasonData, ball_by_ball_data[match_id],
    ball_by_ball_data[batter], "TotalRuns", SUM(ball_by_ball_data[
    batter_runs]))
VAR CenturyCount = FILTER(BatterRuns, [TotalRuns] >= 100)
RETURN COUNTROWS(CenturyCount)
```

```
HalfCenturies =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
VAR SeasonData = FILTER(ball_by_ball_data, RELATED(ipl_matches_data[
    season]) = SelectedSeason)
VAR BatterRuns = SUMMARIZE(SeasonData, ball_by_ball_data[match_id],
    ball_by_ball_data[batter], "TotalRuns", SUM(ball_by_ball_data[
    batter_runs]))
VAR HalfCenturyCount = FILTER(BatterRuns, [TotalRuns] >= 50 && [
    TotalRuns] < 100)
RETURN COUNTROWS(HalfCenturyCount)
```

Chapter 6

Team Metrics

```
Matches Won =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
VAR CurrentTeam = SELECTEDVALUE(teams_data[team_name])
RETURN CALCULATE(COUNTROWS(ipl_matches_data), ipl_matches_data[season]
    = SelectedSeason, ipl_matches_data[match_winner] = CurrentTeam,
    ipl_matches_data[match_type] = "T20")
```

```
Matches Lost =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
-- Team1 lost
VAR Team1Lost = CALCULATE(COUNTROWS(ipl_matches_data), USERELATIONSHIP(
    ipl_matches_data[team1], teams_data[team_name]), ipl_matches_data[
    season]=SelectedSeason, ipl_matches_data[match_type]="T20",
    ipl_matches_data[match_winner]<>ipl_matches_data[team1])
-- Team2 lost
VAR Team2Lost = CALCULATE(COUNTROWS(ipl_matches_data), USERELATIONSHIP(
    ipl_matches_data[team2], teams_data[team_name]), ipl_matches_data[
    season]=SelectedSeason, ipl_matches_data[match_type]="T20",
    ipl_matches_data[match_winner]<>ipl_matches_data[team2])
RETURN Team1Lost + Team2Lost
```

Chapter 7

Dashboard and Visualization

This chapter presents the interactive Power BI dashboard developed for IPL data analysis. The dashboard contains dynamic slicers, KPI cards, tables, and charts to provide detailed insights into teams, players, and matches.

7.1 Dashboard Overview

The dashboard allows users to:

- Filter data dynamically by season, team, and player.
- View key performance indicators such as Orange Cap, Purple Cap, total 6s, total 4s, matches won/lost, and team points.
- Drillthrough for detailed player and team analysis.
- Navigate using bookmarks and interactive tooltips.

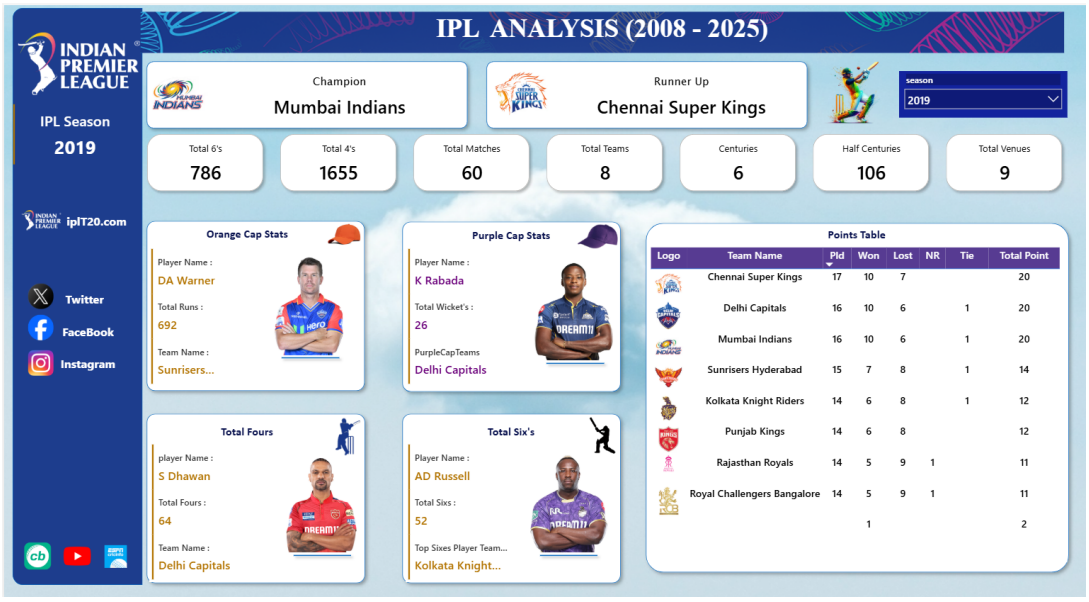

Figure 7.1: Interactive IPL Dashboard in Power BI

7.2 Orange Cap

Definition

The Orange Cap is awarded to the player who scores the most runs in a particular IPL season.

DAX Measure

```
Orange Cap =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
RETURN
TOPN(
    1,
    SUMMARIZE(
        ball_by_ball_data,
        ball_by_ball_data[batter],
        "TotalRuns", SUM(ball_by_ball_data[batter_runs])
    ),
    [TotalRuns], DESC
)
```

Visualization

The Orange Cap is displayed as a KPI card in the dashboard, often accompanied by the player image.

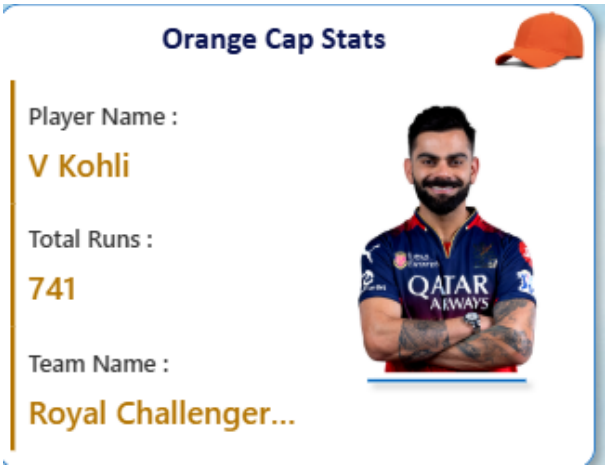

Figure 7.2: Orange Cap KPI Card

7.3 Purple Cap

Definition

The Purple Cap is awarded to the bowler who takes the most wickets in a season.

DAX Measure

```
Purple Cap =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
RETURN
TOPN(
    1,
    SUMMARIZE(
        ball_by_ball_data,
        ball_by_ball_data[bowler],
        "Wickets", SUMX(
            FILTER(ball_by_ball_data, ball_by_ball_data[is_wicket] = 1)
                ,
            1
        )
    ),
    [Wickets], DESC
)
```

Visualization

Displayed as a KPI card with bowler image for the selected season.

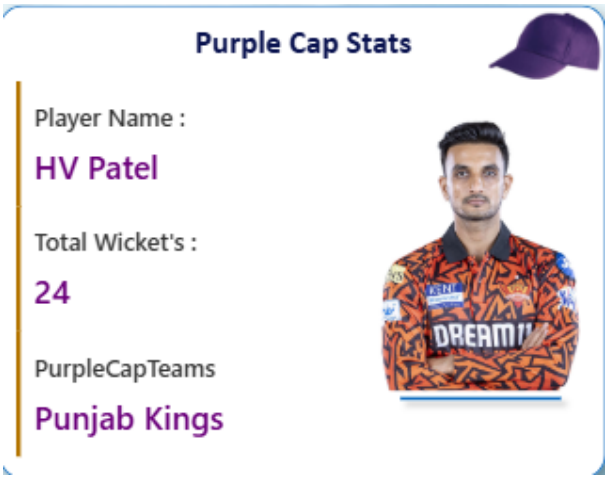

Figure 7.3: Purple Cap KPI Card

7.4 Total Sixes

Definition

This KPI shows the total number of sixes hit by a player or team in a season.

DAX Measure

```
Total Sixes =
CALCULATE(
    COUNTROWS(ball_by_ball_data),
    ball_by_ball_data[batter_runs] = 6)
```

Visualization

Displayed in a card or table showing top six-hitting players per season.

Figure 7.4: Total Sixes KPI Card

7.5 Total Fours

Definition

This KPI shows the total number of fours hit by a player or team in a season.

DAX Measure

```
Total Fours =
CALCULATE(
    COUNTROWS(ball_by_ball_data),
    ball_by_ball_data[batter_runs] = 4
)
```

Visualization

Displayed as a card or table highlighting top boundary scorers.

7.6 Points Table

Definition

The Points Table summarizes the performance of teams in a season, showing matches played, won, lost, points, and net run rate.

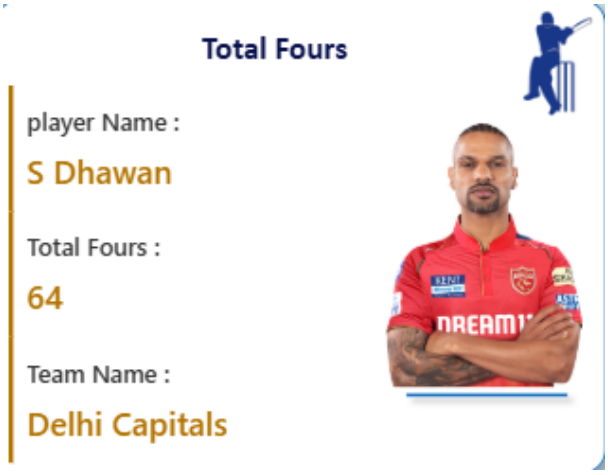

Figure 7.5: Total Fours KPI Card

DAX Measures

```
Matches Won =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
VAR CurrentTeam = SELECTEDVALUE(teams_data[team_name])
RETURN
CALCULATE(
    COUNTROWS(ipl_matches_data),
    ipl_matches_data[season] = SelectedSeason,
    ipl_matches_data[match_winner] = CurrentTeam,
    ipl_matches_data[match_type] = "T20"
)
```

```
Matches Lost =
VAR SelectedSeason = SELECTEDVALUE(ipl_matches_data[season])
VAR CurrentTeam = SELECTEDVALUE(teams_data[team_name])
RETURN
CALCULATE(
    COUNTROWS(ipl_matches_data),
    ipl_matches_data[season] = SelectedSeason,
    ipl_matches_data[match_winner] <> CurrentTeam,
    ipl_matches_data[match_type] = "T20"
)
```

Visualization

The Points Table is displayed as a matrix or table visualization with conditional formatting for easy comparison.

Points Table

Logo	Team Name	Pld	Won	Lost	NR	Tie	Total Point
	Sunrisers Hyderabad	17	10	7			20
	Chennai Super Kings	16	11	5			22
	Kolkata Knight Riders	16	9	7			18
	Rajasthan Royals	15	7	8			14
	Delhi Capitals	14	5	9			10
	Mumbai Indians	14	6	8			12
	Punjab Kings	14	6	8			12
	Royal Challengers Bangalore	14	6	8			12

Figure 7.6: IPL Points Table

Chapter 8

Insights and Findings

- Teams with consistent Orange Cap winners often dominated the season.
- Purple Cap analysis revealed top-performing bowlers per season.
- Correlation between top scorers and team success can be visualized via the dashboard.
- Season-wise match trends help in predicting potential team performance.

Chapter 9

Conclusion

This project demonstrates end-to-end data analysis using Power BI. From data cleaning to advanced DAX calculations, the interactive dashboard provides meaningful insights into IPL matches, teams, and players over 17 seasons. Future work can include predictive modeling and automated updates for ongoing seasons.

Chapter 10

References

- Kaggle IPL Datasets: `https://www.kaggle.com/`
- Power BI Documentation: `https://docs.microsoft.com/en-us/power-bi/`
- DAX Guide: `https://dax.guide/`