

IPL Cricket Analysis & Prediction Project Report

Project Title: IPL Cricket Analysis

Submitted By: Hemashakthi M

1. Introduction

The IPL (Indian Premier League) is one of the most popular T20 cricket tournaments in the world.

This project aims to analyze IPL matches data and provide insights on team performance, player statistics, season trends, and predict match outcomes using simple data science techniques.

2. Dataset Details

We used two CSV files:

1. **matches.csv** – contains match-level information such as season, teams, winner, city, and umpires.
 - a. Rows: 636
 - b. Columns: 18
2. **deliveries.csv** – contains ball-by-ball information including batsman, bowler, runs, and dismissals.

Source: Kaggle IPL Dataset

3. Data Preprocessing

- Checked for missing values.
- Data was grouped and aggregated using Pandas groupby and value_counts.

4. Exploratory Data Analysis

4.1 Top Winning Teams

- Calculated total wins for each team.
- Displayed top 5 winning teams in a bar chart.

4.2 Top Batsmen

- Total runs per batsman were summed from `deliveries.csv`.
- Displayed top 5 batsmen with highest runs.

4.3 Top Bowlers

- Counted wickets for each bowler (excluding run-outs).
- Displayed top 5 bowlers by wickets.

4.4 Matches Per Season

- Counted number of matches played in each season.
- Displayed a bar chart showing trends over the years.

5. Prediction Methodology

- A simple rule-based prediction system was implemented:
 - The team with more historical wins between two teams is predicted as the winner.
 - If both teams have equal wins, the prediction shows: "Draw or Toss Factor."

6. Tools & Libraries Used

- Python 3.x
- Pandas Matplotlib

7. Insights Derived

1. Mumbai Indians, Chennai Super Kings, and Kolkata Knight Riders are historically top-performing teams.
2. Players like Virat Kohli, David Warner, and MS Dhoni feature among top run-scorers.
3. Bowlers like Lasith Malinga and Amit Mishra lead in wickets taken.
4. The number of matches per season has generally increased over time.
5. Teams with historically more wins have a slightly higher chance of winning future matches.

8. Conclusion

This project provides a beginner-friendly analysis of IPL matches and players.

It demonstrates how **Python and basic data science techniques** can generate actionable insights for fans, analysts, and fantasy league players.