**Summary**

The analysis focused on evaluating IPL team performances over multiple seasons, considering match outcomes, win-loss ratios, total runs, and wickets. The dataset was cleaned, formatted, and visualized using Python libraries like Pandas, Matplotlib, and Seaborn.

**Objectives**

**Data Cleaning & Formatting** - Extract necessary columns, handle missing values, and preprocess data for meaningful insights.

**Match Outcome Analysis** - Visualize team wins/losses over seasons.

**Performance Trends** -Identify seasonal trends, run totals, wicket counts, and win-loss ratios.

**Visualization** – data is visualized using seaborn library.

**Activities and Tasks**

Match Outcome Analysis

1. **Task 1**: Visualize match outcomes (Win/Loss) across different years

Description:

The analyses of IPL match data to visualize the number of wins and losses for each team across different seasons.

2.Extract Relevant Columns: Only the necessary columns (‘season’, ‘team1’, ‘team2’, ‘winner’) are extracted for analysis.

3.Handle Missing Values: Rows with missing values in the ‘winner’ column are removed to ensure data completeness.

4.Calculate Wins per Season: Then groups the data by ‘season’ and ‘winner’ to count the number of wins for each team in each season. The result is stored in a Data Frame called ‘wins\_per\_season’.

5. Melt Data Frame: The **pd.melt** function is used to transform the Data Frame so that each row represents a team that played in a match, along with the season and the winner.

6.Mark Losses: A new column called ‘loss’ is created to indicate whether a team lost a match. This is done by checking if the team is not the winner if it is true then it will be marked as true

7.Calculate Losses per Season: Then groups the data by ‘season’ and ‘team’ to count the number of losses for each team in each season. The result is stored in a DataFrame called ‘losses\_per\_season’.

9.Visualize Wins and Losses: The use of ‘seaborn’ to create bar plots showing the number of wins and losses for each team per season.

1. **Task 2:** Analyse team performance based on historical data and seasonal trends

Description: To analyse the team performance and season trends we have to look at total run, wickets, win to loss ratio and seasonal performance

1. Extract Relevant Columns: Only the necessary columns (‘id’, ’season’, ‘team1’, ‘team2’, ‘winner’) are extracted for analysis.
2. Find the total runs scored by each team: group the deliveries data frame by ‘batting\_team’ and find the sum of ‘total\_runs’.
3. Find the total wicket by Each team: we have a ‘is\_wicket’ column so we can group that but before that we need to filter out the values i.e 0 wickets onces that is done we can group the column by bowling team and store it in a variable called **Wicket**.
4. Calculate win/loss ratio for each team: by grouping the ‘matches\_df’ by **winner**. Transform the ‘matches\_df’ with ‘pd.melt’ so that each row represent a team. then group the data by team to count the total number of matches played by each team then we merge the ‘wins and matches\_played’ to calculate the win/loss ratio. team\_performance = team\_performance[['batting\_team', 'total\_runs', 'wickets', 'win\_loss\_ratio']].
5. To understand the season trends, we have to group them by season and winner from ‘matches\_df’.

**Challenges and Solutions**

Came across multiple challenges while understanding the dataset as I am not very knowledgeable about cricket or the format of it working in a team helped me understand the about this I won’t say I am an expert but I now know that 1over is equal to 6 balls.

While performing my 1st task where I had to visualize outcome win/loss from different years where there was no such column as loss to solve It I first transformed the dataset using melt function for team1 and 2 then compared the team’s column with winner column e.g. if team is not a winner it will be true this can by using lambda which is a small anonymous function.

|  |  |  |
| --- | --- | --- |
| winner | team | loss |
| Kolkata Knight Riders | Royal Challengers Bangalore | True |
| Chennai Super Kings | Kings XI Punjab | True |
| Delhi Daredevils | Delhi Daredevils | False |

**Conclusion**

The analysis provided valuable insights into IPL team performance trends. By calculating and visualizing various performance metrics such as total runs, total wickets, win/loss percentage, and seasonal trends, the analysis provides valuable insights into team performance. The use of data visualization techniques helps in identifying patterns and trends, making it easier to understand the strengths and weaknesses of different teams over time.

**Winning Consistency:** Teams like Mumbai Indians and Chennai Super Kings showed stability over seasons, while others fluctuated.

**Batting vs. Bowling Impact:** Some teams excelled in scoring but lacked strong bowling performance, affecting their win-loss ratio.

**Recommendation**

* Conduct targeted training sessions to improve batting techniques and strategies.
* Focus on improving accuracy and variations to take crucial wickets.
* Encourage a culture of continuous learning and improvement within the team.
* Regularly practice fielding drills to enhance catching, throwing, and ground fielding skills.