

**Term 1 – Project**

**Subject:**

**Data Exploration & Visualization using Python I**

**GitHub Link: https://github.com/045005/045005\_IPL-23.git**

**Submitted by: Submitted to:**

**Agathiyan K Prof. Amarnath Mitra**

**045005**

**Index**

1. Project Objectives
2. Description of the Data
3. Analyzing the Batsmen List
   1. Analyzes based on runs
   2. Analysis based on 4s
   3. Analysis on based on 6s
   4. Finding the strongest batsman and team in terms of boundary
   5. Analysis based on Strike Rate and total average by each teams.
4. Analysis on the Bowlers Data
   1. Based on wickets taken
   2. Runs conceded.
   3. Total Economy Rate and Average by runs in each team.
   4. Analysis using the 3 wicket hauls, 5 wicket hauls, and maiden overs
5. Managerial Insights
6. **Project Objective**

The main objective of the project is to understand the latest stats of the teams that played in IPL’23. Using this analysis, we can further rank the team that played in the IPL based on several criteria.

1. **Introduction**

For all the codes used for this analysis, kindly refer to the ipnb file attached in my GitHub repository.

Here, I have scrapped data from 2 different URLs:-

<https://www.news18.com/cricketnext/ipl-2023/orange-cap-holder.html>

<https://www.news18.com/cricketnext/ipl-2023/purple-cap-holder.html>

The 1st URL is the data of all the batsmen who scored in the IPL 2023

Similarly, the 2nd URL is the data of all the bowlers who took wickets in the IPL 2023

First, I scrapped the batsmen data and saved them in a dataframe under the name df\_orange

Then I scrapped the bowler’s data and saved them in a dataframe under the name df\_purple

A screenshot of a sports schedule

Description automatically generated

A screenshot of a sports schedule

Description automatically generated

1. **Analyzing the batsmen list**

3.a. Analyzes based on runs

Here, I found the total runs scored in the entire IPL 2023.

Total runs by the top 10 batsmen

Total runs scored by each IPL team

And I compared the runs scored by each team and found that GT scored the most (2906).

Thus, GT is the most run-scoring team.

A screenshot of a computer

Description automatically generated

A graph with blue lines and white text

Description automatically generated

3.b. Analysis based on 4s

Here, I calculated the total number of 4s in IPL’23.

Of the 166 batsmen, 31 didn’t get a boundary (only 4).

And I segregated the range of the number of 4s and counted the batsmen in that range

A screenshot of a computer

Description automatically generated

A graph of a number of bars

Description automatically generated

Then calculated the number of 4s scored by each team.

A screenshot of a computer screen

Description automatically generated

A graph with blue lines

Description automatically generated

I also found out the top 4 scoring batsmen of each team.

A screen shot of a computer

Description automatically generated

3.c. Analysis based on 6s

Here, I calculated the total number of 6s in IPL’23.

Of the 166 batsmen, 54 didn’t get a boundary (only 6).

And I segregated the range of the number of 6s and counted the batsmen in that range.

A screenshot of a computer

Description automatically generated

A graph with blue squares

Description automatically generated

Then calculated the number of 6s scored by each team

A screenshot of a computer screen

Description automatically generated

A graph with blue lines

Description automatically generated

I also found out the top 6 scoring batsman of each team

A screenshot of a computer

Description automatically generated

3.d. Finding the strongest batsman and team in terms of boundary

Analyzing from the top 4 and 6 scoring batmen of each team, 4 players are topped in both the list.

A screenshot of a computer

Description automatically generated

Here, based on the highest average runs and highest runs, Shubham Gill is the Strongest Batsman in terms of Boundaries.

For teams, totals run scored only through boundaries by each team are

A screenshot of a computer program

Description automatically generated

A graph with blue lines

Description automatically generated

As per the findings, MI is the strongest team in scoring boundaries.

3.e. Analysis based on Strike Rate and total average by each team.

Here, I calculated the average strike rate and total average for each team. The formula used to calculate the average strike rate is the sum of all the strike rates for a given team divided by the total number of batsmen in that team. Similarly, for calculating the total average, the formula is the sum of the average of all the batsmen in that team divided by the total number of batsmen.

A screen shot of a computer

Description automatically generated

A graph of blue and orange bars

Description automatically generated

1. **Now the analysis of the bowlers' data**

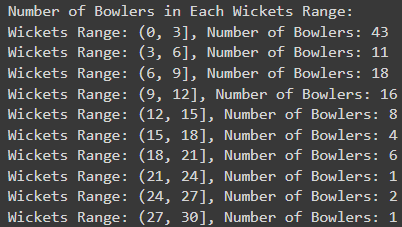
4.a. Based on wickets taken

Here, I calculated the total number of wickets taken in IPL’23 along with the top wicket-taker.

A screenshot of a computer screen

Description automatically generated

Then, I segregated the wickets taken into specific ranges and counted the number of bowlers in that range.



A graph of a number of crickets

Description automatically generated

Then, I found out the total number of wickets taken by each team in the 2023 season.

A screen shot of a computer

Description automatically generated

A graph with a line and a purple line

Description automatically generated

Here, for this data, we can say that the GT is most effective team in taking wickets.

4.b. Runs conceded.

I calculated the total runs that each team given. The formula I used here is by calculating the sum of the runs given by each bowler of that specific team.

A screenshot of a computer screen

Description automatically generated

A graph of a number of blue bars

Description automatically generated with medium confidence

As per the result, LSG team has the most efficient bowlers in terms of runs conceded.

4.c. Total Economy Rate and Average by Runs in each team.

Here, I calculated the average economy rate of each team. The formula used here is the sum of the economy rate of each bowler in a specific team. Similarly, I calculated the average by a bowler of each team.

A screen shot of a computer

Description automatically generated

A graph of a graph with different colored bars

Description automatically generated with medium confidence

Team with the best economy rate (lowest EC): CSK

Team with the best average (lowest AVG): LSG

4.d. Analysis using the 3 wicket hauls, 5 wicket hauls, and maiden overs.

Here, I calculated the total number of 3 wicket hauls, 5 wicket hauls and maiden overs by each team separately and compiled all the output in a single line graph.

A graph with numbers and lines

Description automatically generated

1. **Managerial Insights**

Batsmen Implications:

* Identify and retain batsmen who consistently perform well, especially those who score the most runs.
* Batsmen with a high number of boundaries (4s and 6s) are valuable. They can take advantage of field restrictions and powerplays to score quickly.
* Assess the distribution of top batsmen among different teams. This information can help in strategic bidding during auctions to ensure a balanced team.
* Ranking the teams based on several factors which include runs, power hitting (highest in boundaries), strike rate, and total average.

Bowlers Implications:

* Identify and retain bowlers who take a significant number of wickets (Purple Cap holders). Wicket-takers can put pressure on the opposition and change the course of the game.
* Balance wicket-taking with the economy rate (EC). Teams that can take wickets without leaking too many runs are crucial, especially in T20 cricket where runs are at a premium.
* Similarly Average (AVG) given by a bowler is also important for the team.
* In games like T20 it is important to score the most runs while batting. Just like scoring, conceding runs while bowling is also equally important. So, we can categorize the teams based on these factors too.
* Teams that can deliver maiden overs (MDNS) are essential to build pressure on the batsmen.