



A Data-Driven Analysis of Two IPL Stars' Strengths and Weaknesses (1)

As a data scientist with a deep passion for cricket, I've always been fascinated by the stories that data can tell about the sport. The Indian Premier League (IPL), with its thrilling matches and dynamic performances, provides a rich dataset to explore and analyze. The excitement of T20 cricket, combined with the wealth of statistics it generates, presents a unique opportunity to delve into the nuances of player performance. My motivation for this analysis stems from my dual interests: leveraging data to uncover insights and celebrating the strategic prowess of top IPL players.

In this blog, I aim to dissect the performance of notable Indian batters, using Python libraries pandas and matplotlib to highlight their contributions and impact within the IPL. This not only satisfies my curiosity as a cricket enthusiast but also sharpens my skills as a data scientist, allowing me to contribute to the broader discourse on sports analytics. This blog marks my first foray into sports analytics and serves as an experiment in applying my data science skills to cricket. While I strive for accuracy, this is a learning experience, and the analysis might not capture every nuance perfectly.

We'll start our analysis by focusing on Suryakumar Yadav, who is recognized as the world's top T20 batter. While his ranking speaks volumes about his strengths, we'll delve deeper into his weaknesses to provide a comprehensive view of his performance. Using his detailed statistical data, we aim to explore various scenarios that highlight these areas. This approach will not only showcase the breadth of his cricketing capabilities but also allow us to examine the aspects of his game that might benefit from further improvement.

Dataset source : <https://www.kaggle.com/datasets/bhuvaneshprasad/ipl-dataset-2008-2023-only-dataset-you-need>

About the dataset :

This dataset contains historical match results, batting cards, bowling cards, fall of wickets cards, partnership cards, players info and ball by ball data for all IPL matches since 2008 till 2023. The dataset is split into multiple files, each containing a different type of data.

Files:

1. **ipl_historical.csv**: Contains match results data, such as the date, venue, teams involved, toss winner, toss decision, result, margin of victory, player of the match, etc.
2. **ipl_batting_card.csv**: Contains batting card data, such as the player name, team, batting position, runs scored, balls faced, fours, sixes, strike rate, dismissal method, etc.
3. **ipl_bowling_card.csv**: Contains bowling card data, such as the player name, team, bowling overs, maidens, runs conceded, wickets taken, economy rate, etc.
4. **ipl_fow_card.csv**: Contains fall of wicket data, such as the over number, wicket number, batsman name, dismissal method, etc.
5. **ipl_partnership_card.csv**: Contains partnership data, such as the partnership number, batsmen names, runs scored, balls faced, etc.
6. **ipl_players_info.csv**: Contains player information data, such as the player name, team, nationality, batting style, bowling style, etc.
7. **ipl_ball_by-ball_data.csv**: Contains ball-by-ball data for all IPL matches, such as the batsman name, bowler name, runs scored, extras, wicket ball, etc.

We begin by retrieving the player ID from [ipl_players_info.csv](#), specifically for Suryakumar Yadav, whose ID is 61990. This ID will be pivotal throughout our analysis. We then integrate annual data from [ipl_historical.csv](#) into [ipl_batting_card.csv](#). Primarily, we will utilize this comprehensive dataset for our entire analysis; it contains ball-by-ball details of runs scored by each batter from 2008 to 2023, including the bowler ID. This enables us to extract specific bowler details from [ipl_players_info.csv](#). We will specifically focus on the last three years of the IPL—2021, 2022, and 2023—to conduct our analysis.

We focus on two key attributes of a bowler: the "bowler-hand," categorized as left or right, and "bowler style," which is divided into several types.

```
['legbreak', 'right-arm offbreak', 'right-arm medium', nan,
'legbreak googly', 'right-arm fast-medium',
'slow left-arm orthodox', 'right-arm fast', 'left-arm fast-medium',
'left-arm medium', 'right-arm medium-fast', 'left-arm medium-fast',
'left-arm fast', 'left-arm wrist-spin']
```

A sample of dataset:

	match_id	year	bowler_id	bowler_style	bowler_arm	batsman_runs	iswicket	issix	isfour
3816	1359505	2023	59832.0	right-arm offbreak	right-arm	1.0	False	False	False
3820	1359505	2023	95219.0	right-arm leg break	right-arm	0.0	False	False	False
3821	1359505	2023	95219.0	right-arm leg break	right-arm	4.0	False	False	True
3822	1359505	2023	95219.0	right-arm leg break	right-arm	1.0	False	False	False
3824	1359505	2023	95219.0	right-arm leg break	right-arm	0.0	False	False	False
3827	1359505	2023	59832.0	right-arm offbreak	right-arm	4.0	False	False	True
3828	1359505	2023	59832.0	right-arm offbreak	right-arm	4.0	False	False	True
3829	1359505	2023	59832.0	right-arm offbreak	right-arm	4.0	False	False	True
3830	1359505	2023	59832.0	right-arm offbreak	right-arm	0.0	False	False	False
3831	1359505	2023	59832.0	right-arm offbreak	right-arm	1.0	False	False	False

Note : We clean the data with data only for this batter and remove wickets on freehit balls since it might affect the stats while accounting for strike rate ,average and dismissal ratio.

Stats summary for Suryakumar Yadav in last three years:

Basic stats

- Number of matches batted by this batter: 124
- total runs scored by the batter 3249
- Balls faced by the batter 2267
- Strike rate = 143.31
- Average - 35.60

Detailed Stats

- Fig 1 shows that percentage of runs Suryakumar Yadav has scored against each bowler style.

Percentage Runs scored by Suryakumar Yadav against each bowler style

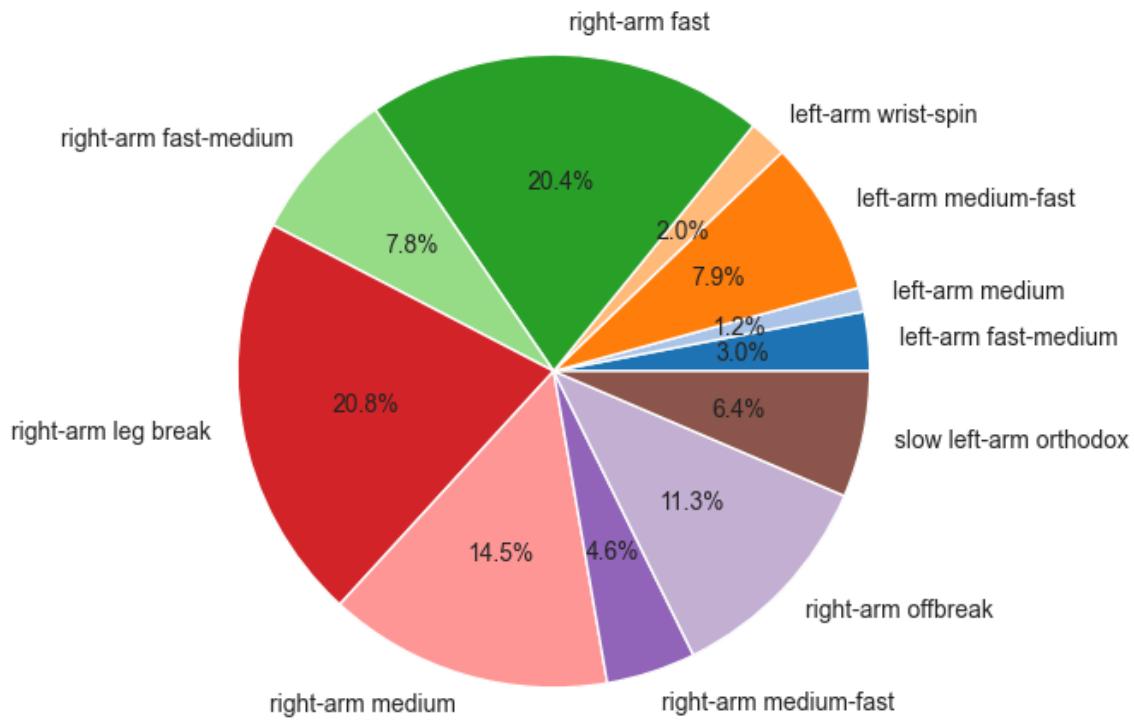


Fig 1 : Percentage of runs scoed by Suryakumar Yadav

- **Batting Strike Rate:** In T20 cricket, the batting strike rate (s/r) is a crucial metric, defined as the average number of runs a batter scores per 100 balls faced. A higher strike rate indicates that the batter is more effective at scoring quickly, making it a key statistic for evaluating performance in this fast-paced format.
 - Figure 2 indicates that Suryakumar Yadav's batting strike rate is notably lower against both slow left-arm orthodox and left-arm wrist spin bowlers. ~ 103.94

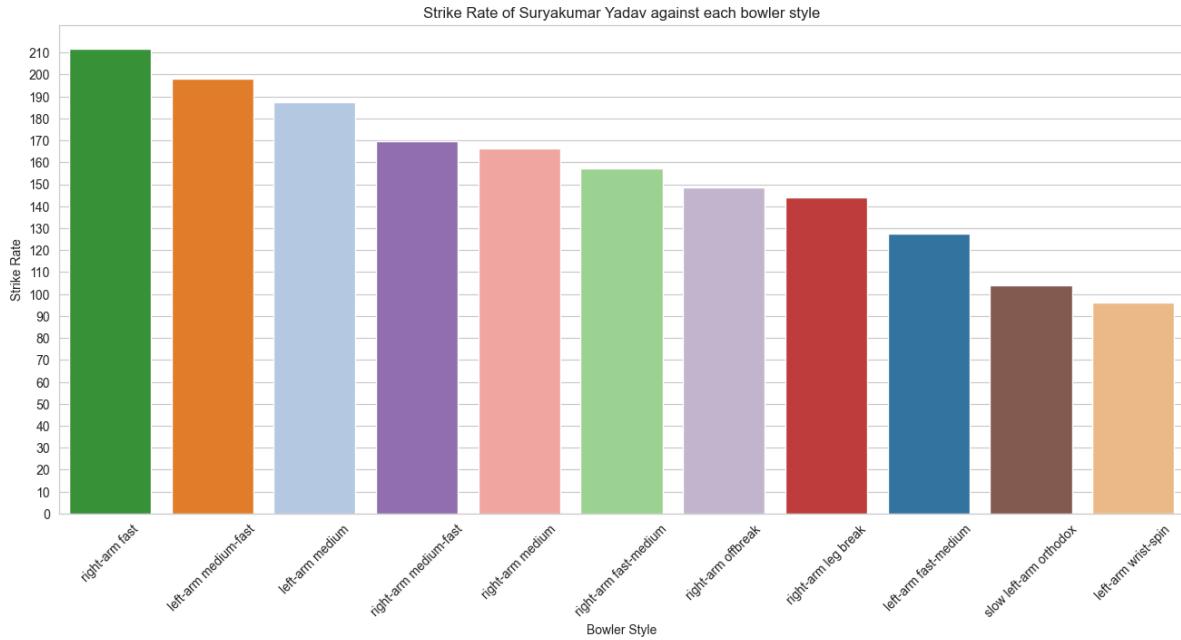


Figure 2: Strike rate against each bowler style

- From the data and Figure 3, it's evident that Suryakumar Yadav has faced relatively few deliveries against left-arm wrist spinners (~10 balls), but he has encountered a significant number of deliveries from slow left-arm orthodox bowlers (75 balls).
 - Therefore, we have decided to shift our focus and analysis toward his performance against slow left-arm orthodox bowling.

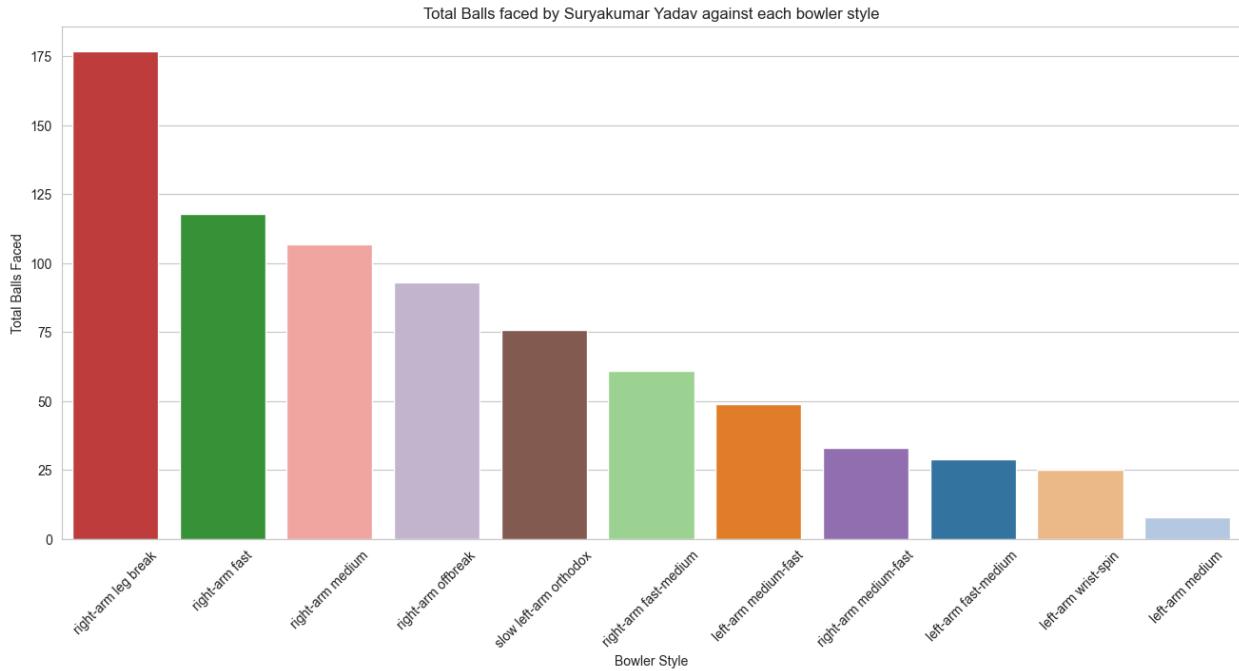


Figure 3 : Totals deliveries faced against each bowler.

- From Figure 4, we observe that the percentage of dismissals for Suryakumar Yadav is the second highest against slow left-arm orthodox bowlers, at approximately 16%. This suggests a significant vulnerability to slow left-arm orthodox bowling.



Percentage Dismissals of Suryakumar Yadav against each bowler style in last 3 years

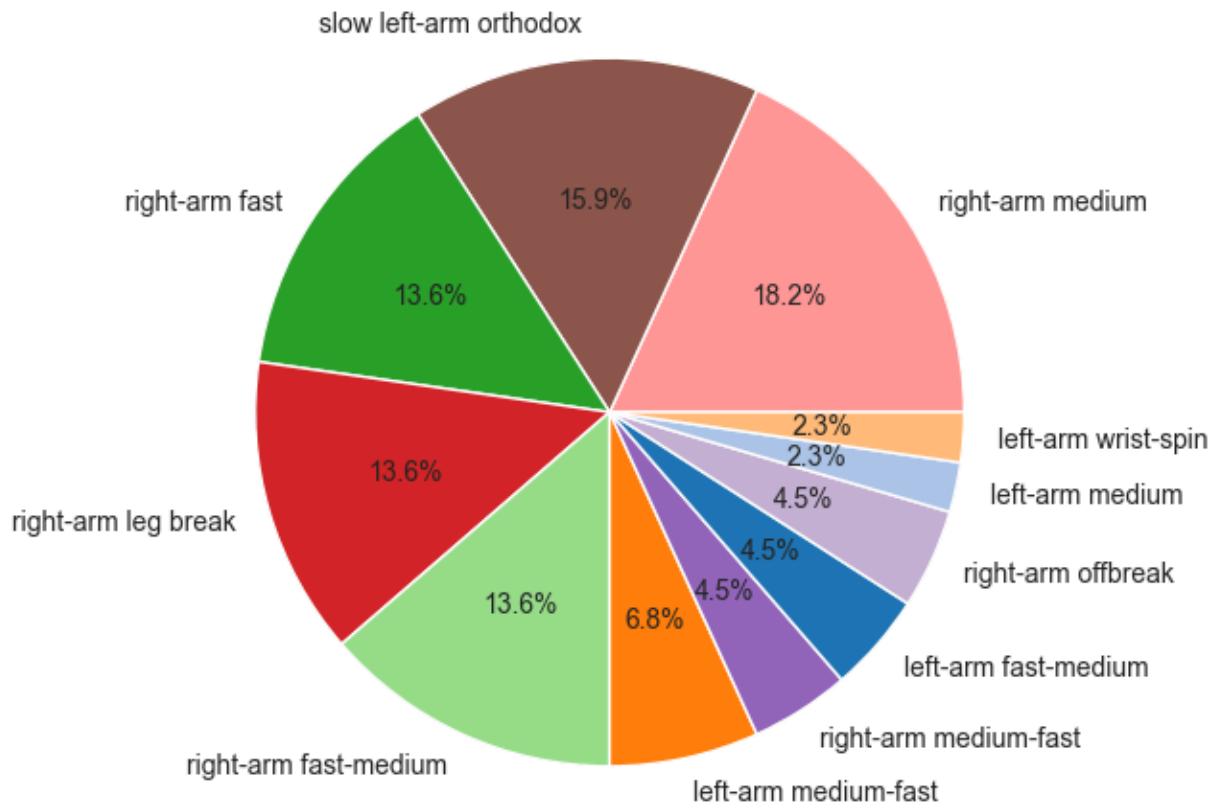


Figure 4 : Dismissal percentage

- Batting average :The batting average is a statistic used to evaluate a player's performance in cricket. It is calculated by dividing the total number of runs a player has scored by the number of times they have been out. Typically, this value is presented to two decimal places. A higher batting average indicates a player's greater consistency and effectiveness at scoring runs before getting out.
 - In our analysis of the overall performance across three years, Suryakumar Yadav has an average of 35.0. However, when we break down the averages by bowler style using data from Figure 5, it's evident that his lowest average, 13.7, occurs against slow left-arm orthodox bowling. This suggests a particular vulnerability to this style of bowling, which might warrant closer examination and targeted improvement in his gameplay.

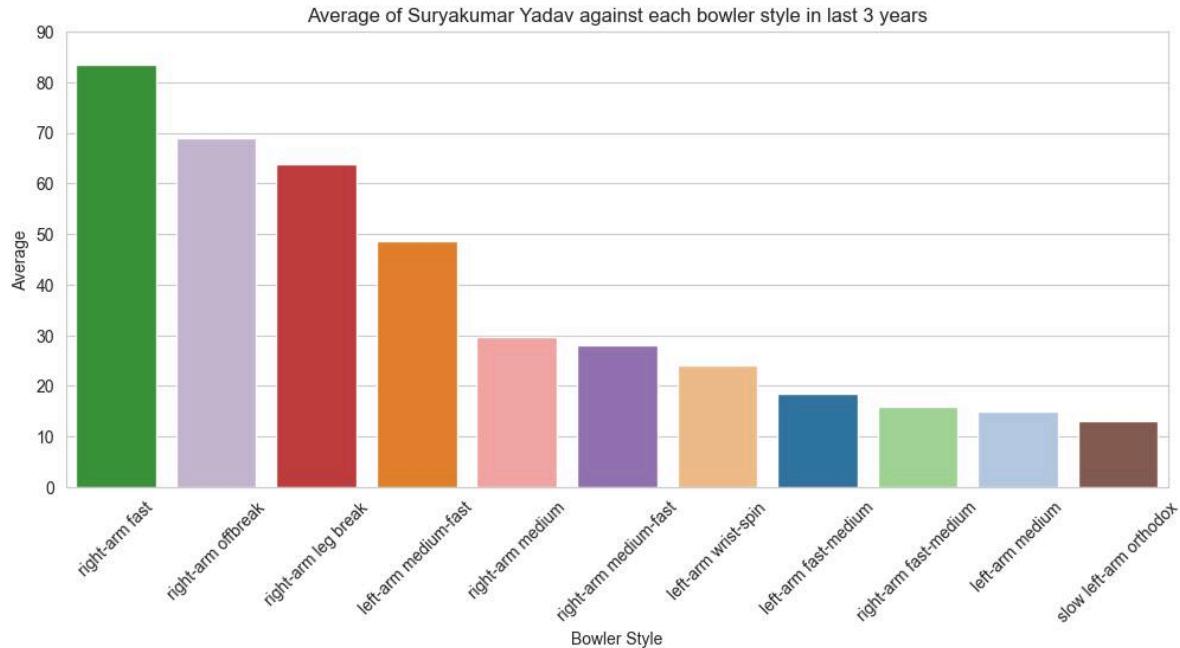


figure 5 : Average against each bowler style



Conclusion :

Suryakumar Yadav's overall performance against various bowlers has been exceptional, underscoring his status as the top T20 batter. However, his notably poorer performance against slow left-arm orthodox bowlers stands out. Despite this, it's important to consider that he achieved his number one ranking last year, which could suggest that statistics might not always tell the full story. To gain a more accurate understanding of whether he has addressed this weakness, we should not only consider his recent IPL performances but also analyze his performances across all formats in the last year. This broader view will help us determine if there has been a meaningful improvement against this specific type of bowling.

Repo : https://github.com/visakhpadmanabhan7/crick_analytics/blob/main/sky_analysis.ipynb