INDIAN PREMIERE LEAGUE 2023 DATA ANALYSIS

Full Project Details with Source Code & Explanation

@sunilghanchi

About IPL 2023

- **Number of matches**: There were a total of 74 matches played in the IPL 2023 season, including the playoffs and final.
- **Number of players**: A total of 600 players played in the IPL 2023 season, including both Indian and overseas players.
- **Teams**: There were 10 teams competing in the IPL 2023 season:
  - Chennai Super Kings
  - Delhi Capitals
  - Gujarat Titans
  - Kolkata Knight Riders
  - Lucknow Super Giants
  - Mumbai Indians
  - Punjab Kings
  - Rajasthan Royals
  - Royal Challengers Bangalore
  - Sunrisers Hyderabad
- **Format**: The IPL 2023 season was played in a round-robin format, with each team playing each other team twice. The top four teams at the end of the league stage qualified for the playoffs.
- **Playoffs**: The IPL playoffs format consists of four stages: Qualifier 1, where the top two teams from the league stage compete for a direct spot in the final; Eliminator, between the third and fourth-placed teams, with the winner advancing to Qualifier 2; Qualifier 2, where the victor of the Eliminator faces the losing team from Qualifier 1 for the last final spot; and finally, the championship Final, pitting the winners of Qualifier 1 and Qualifier 2 against each other to determine the IPL champion.
- **Winner**: The Chennai Super Kings won the IPL 2023 season, defeating the Gujarat Titans in the Final.

Project Abstract

The IPL 2023 Data Analysis project aims to create an interactive and insightful dashboard using Power BI, focusing on player performance analysis across various positions in batting and bowling. By leveraging the power of data visualization and analytics, this project seeks to construct an unbeatable playing XI that excels in different aspects of the game.

The project involves collecting and cleaning player data for the IPL 2023 season, encompassing batting and bowling statistics. Through the creation of interactive dashboards in Power BI, the following key aspects will be addressed:

1. **Positional Analysis:** Players will be assessed based on their performances in different batting and bowling positions. The dashboard will provide a clear view of how players contribute to the team's success in various roles.
2. **Player Consistency and Form:** The dashboard will showcase players' consistency and recent form through dynamic visualizations, helping to identify consistent performers and those in peak form.
3. **Match Situation Performance:** By segmenting performance based on match situations such as powerplays, middle overs, and death overs, the dashboard will highlight players' strengths under specific conditions.
4. **Balancing the Team:** The dashboard will allow users to experiment with different combinations of players from various positions to create a balanced and potent playing XI.
5. **Captaincy Insights:** Leadership qualities and decision-making abilities of players will be visually presented, assisting in identifying potential captaincy candidates for the unbeatable XI.
6. **Emerging Talents:** The dashboard will provide insights into the performances of emerging talents, aiding in their identification and assessment.

There are many more things that we can analysis but in this project we are analysis the every position as Power Hitters/Openers, Middle Orders/Anchors, Lower Order/Finisher, All rounder, Fast Bowlers. These are the main things we are going to focus on and making dashboard.

The outcome of the IPL 2023 Data Analysis project will be an interactive Power BI dashboard that enables users to explore player statistics, make informed decisions, and construct their unbeatable playing XI. Users, such as team management, analysts, and cricket enthusiasts, will be able to engage with the dashboard, gaining valuable insights and constructing teams that leverage players' strengths effectively.
This project aims to enhance the understanding of player dynamics, support team selection decisions, and elevate the enjoyment of the IPL by offering an interactive platform to visualize and analyze player performances comprehensively.

Steps for Data Analysis

1. Data Collection and Import
2. Data Transformation or Data Preprocessing
3. Data Modeling
4. Calculations and Measures
5. Dashboard Creation
6. Adding Interactivity
7. Performance Optimization

1. Data Collection and Import

- Certainly, data collection and import is a crucial initial step in our project. For this analysis, we will gather IPL 2023 data from the ESPN Cricket website. By leveraging web scraping techniques, we will extract information such as team details, player statistics, match results, and other relevant metrics. This data will serve as the foundation for our Power BI analysis, enabling us to uncover insights into player performance, team dynamics, and overall trends throughout the tournament.
- We will utilize the Scrapy framework to scrape IPL 2023 data from the ESPN Cricket website. Scrapy is a powerful and efficient Python library specifically designed for web scraping. With Scrapy, we can define spider classes that navigate through web pages, extract structured data using CSS or XPath selectors, and store the collected data. By configuring Scrapy settings and pipelines, we can ensure efficient data extraction and transformation, allowing us to preprocess the scraped data using pandas and perform in-depth analysis using tools like Power BI.

Website

https://www.espncricinfo.com/records/tournament/team-match-results/indian-premier-league-2023-15129

Match results

Team 1	Team 2	Winner	Margin	Ground	Match Date	Scorecard
Titans	Super Kings	Titans	5 wickets	Ahmedabad	Mar 31, 2023	Twenty20
Punjab Kings	KKR	Punjab Kings	7 runs	Mohali	Apr 1, 2023	Twenty20
Super Giants	Capitals	Super Giants	50 runs	Lucknow	Apr 1, 2023	Twenty20
Sunrisers	Royals	Royals	72 runs	Hyderabad	Apr 2, 2023	Twenty20
RCB	Mumbai	RCB	8 wickets	Bengaluru	Apr 2, 2023	Twenty20
Super Kings	Super Giants	Super Kings	12 runs	Chennai	Apr 3, 2023	Twenty20
Capitals	Titans	Titans	6 wickets	Delhi	Apr 4, 2023	Twenty20
Royals	Punjab Kings	Punjab Kings	5 runs	Guwahati	Apr 5, 2023	Twenty20
KKR	RCB	KKR	81 runs	Eden Gardens	Apr 6, 2023	Twenty20
Super Giants	Sunrisers	Super Giants	5 wickets	Lucknow	Apr 7, 2023	Twenty20
Royals	Capitals	Royals	57 runs	Guwahati	Apr 8, 2023	Twenty20
Mumbai	Super Kings	Super Kings	7 wickets	Wankhede	Apr 8, 2023	Twenty20
Titans	KKR	KKR	3 wickets	Ahmedabad	Apr 9, 2023	Twenty20
Sunrisers	Punjab Kings	Sunrisers	8 wickets	Hyderabad	Apr 9, 2023	Twenty20
RCB	Super Giants	Super Giants	1 wicket	Bengaluru	Apr 10, 2023	Twenty20
Capitals	Mumbai	Mumbai	6 wickets	Delhi	Apr 11, 2023	Twenty20
Super Kings	Royals	Royals	3 runs	Chennai	Apr 12, 2023	Twenty20
Punjab Kings	Titans	Titans	6 wickets	Mohali	Apr 13, 2023	Twenty20

We will scrape this details from the website as Column names Team 1, Team 2, Winner, Margin, Ground, Match Date and Scoreboard link to go there and collect every match scoreboard with separate of batting and bowling

```python
import scrapy
import re

class EspncricinfoSpider(scrapy.Spider):
    name = "espncricinfo"
    allowed_domains = ["espncricinfo.com"]
    start_urls = ["https://www.espncricinfo.com/records/tournament/team-match-results/indian-premier-league-2023-15129"]
    header = {'User-Agent':'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110
    match_id = 1

    def __init__(self, *args, **kwargs):
        super(EspncricinfoSpider, self).__init__(*args, **kwargs)
        self.match_id_counter = 0

    def start_requests(self):
        for url in self.start_urls:
            yield scrapy.Request(url=url, callback=self.parse, headers=self.header)

    def parse(self, response):
        boards = response.xpath("//div[@class='ds-p-0']/div/div[@class='ds-overflow-x-auto ds-scrollbar-hide']")
        for rows in boards:
            row = rows.xpath(".//table/tbody/tr")
            for team in row:
                fteam = team.xpath(".//td[@class='ds-min-w-max']/span/a/span/text()").get()
                steam = team.xpath(".//td[@class='ds-min-w-max ds-text-right']/span/a[@class='ds-inline-flex ds-items-start ds-lea
                winner = team.xpath("(.//td[@class='ds-min-w-max ds-text-right'])[2]/span/text()").get()
                margin = team.xpath("(.//td[@class='ds-min-w-max ds-text-right'])[3]/span/text()").get()
                ground = team.xpath("(.//td[@class='ds-min-w-max ds-text-right'])[4]/span/a[@class='ds-inline-flex ds-items-start
                date = team.xpath(".//td[@class='ds-min-w-max ds-text-right ds-whitespace-nowrap']/span/text()").get()
```

You can see this we start our scrappy with scrapping and we are scrapping the data by X-Path and save the data in both CSV and JSON for better Visualize of data

```
Match_Id,Team 1,Team 2,Team,Winner,Margin,Ground,Date,Scoreboard Link
1,Titans,Super Kings,Titans Vs Super Kings,Titans,5 wickets,Ahmedabad,"Mar 31, 2023",https://www.espncricinfo.com/ser
2,Punjab Kings,KKR,Punjab Kings Vs KKR,Punjab Kings,7 runs,Mohali,"Apr 1, 2023",https://www.espncricinfo.com/series/i
3,Super Giants,Capitals,Super Giants Vs Capitals,Super Giants,50 runs,Lucknow,"Apr 1, 2023",https://www.espncricinfo.
4,Sunrisers,Royals,Sunrisers Vs Royals,Royals,72 runs,Hyderabad,"Apr 2, 2023",https://www.espncricinfo.com/series/ind
5,RCB,Mumbai,RCB Vs Mumbai,RCB,8 wickets,Bengaluru,"Apr 2, 2023",https://www.espncricinfo.com/series/indian-premier-l
6,Super Kings,Super Giants,Super Kings Vs Super Giants,Super Kings,12 runs,Chennai,"Apr 3, 2023",https://www.espncric
7,Capitals,Titans,Capitals Vs Titans,Titans,6 wickets,Delhi,"Apr 4, 2023",https://www.espncricinfo.com/series/indian-
8,Royals,Punjab Kings,Royals Vs Punjab Kings,Punjab Kings,5 runs,Guwahati,"Apr 5, 2023",https://www.espncricinfo.com/
9,KKR,RCB,KKR Vs RCB,KKR,81 runs,Eden Gardens,"Apr 6, 2023",https://www.espncricinfo.com/series/indian-premier-league
10,Super Giants,Sunrisers,Super Giants Vs Sunrisers,Super Giants,5 wickets,Lucknow,"Apr 7, 2023",https://www.espncric
11,Royals,Capitals,Royals Vs Capitals,Royals,57 runs,Guwahati,"Apr 8, 2023",https://www.espncricinfo.com/series/india
12,Mumbai,Super Kings,Mumbai Vs Super Kings,Super Kings,7 wickets,Wankhede,"Apr 8, 2023",https://www.espncricinfo.com
13,Titans,KKR,Titans Vs KKR,KKR,3 wickets,Ahmedabad,"Apr 9, 2023",https://www.espncricinfo.com/series/indian-premier-
14,Sunrisers,Punjab Kings,Sunrisers Vs Punjab Kings,Sunrisers,8 wickets,Hyderabad,"Apr 9, 2023",https://www.espncrici
15,RCB,Super Giants,RCB Vs Super Giants,Super Giants,1 wicket,Bengaluru,"Apr 10, 2023",https://www.espncricinfo.com/s
16,Capitals,Mumbai,Capitals Vs Mumbai,Mumbai,6 wickets,Delhi,"Apr 11, 2023",https://www.espncricinfo.com/series/india
17,Super Kings,Royals,Super Kings Vs Royals,Royals,3 runs,Chennai,"Apr 12, 2023",https://www.espncricinfo.com/series/
```


```
{
    "Match_Id": 1,
    "Team 1": "Titans",
    "Team 2": "Super Kings",
    "Team": "Titans Vs Super Kings",
    "Winner": "Titans",
    "Margin": "5 wickets",
    "Ground": "Ahmedabad",
    "Date": "Mar 31, 2023",
    "Scoreboard Link": "https://www.espncricinfo.com/series/indian-pr
},
```


The first one is CSV format and second one is JSON format. Both the Python and CSV files will be given to you.

Similarly we scrap the all batting and bowling states and also scrap the Player info with images

2. Data Transformation or Data Preprocessing

- Data transformation and preprocessing will be conducted using Python with the Pandas library. This step involves cleaning, structuring, and refining the IPL 2023 data collected from the ESPN Cricket website. By utilizing Pandas' powerful data manipulation functions, we will handle missing values, remove duplicates, reformat columns, and apply necessary calculations to derive meaningful insights. This preparation ensures that the data is in an optimal format for analysis and visualization in Power BI.
- After cleaning and structing the data we will ready to build the dashboard, let’s start to build the dashboard.

3. Data Modeling

In Power BI, we will further transform the scraped and preprocessed data to ensure it's in the required format for analysis. This includes converting date fields to the appropriate date format, handling missing values, and applying any necessary calculations or aggregations. Once the data is prepared, we will create a data model by establishing relationships between different tables if needed. This step is crucial for building meaningful visualizations and insights. With the data model in place, we can start creating interactive and informative visualizations using Power BI's drag-and-drop interface. These visualizations will allow us to analyze player performance, batting and bowling statistics, power hitters' impact, team dynamics, and other relevant insights from the IPL 2023 data.

4. Calculations and Measures

- We will calculate various key metrics and measures using Data Analysis Expressions (DAX) in Power BI. These calculations will include calculating total runs, total innings played, total overs bowled, total number of wides, total number of no-balls, as well as other batting and bowling-related measures. These DAX measures will provide us with valuable insights into player performance, team dynamics, and power hitters' impact on the matches. By aggregating and analyzing this data, we can identify trends, patterns, and formulate strategies for building an unbeatable player 11 for the IPL 2023 season.
- In this later we will also make custom position DAX measure and Player Selection, Color Callout. This 3 are important measures for interactivity.

❑ Dashboard Creation, Interactive Design and Performance

Creating an effective dashboard in Power BI involves several important considerations:

1. Clear Objective: Define the purpose of your dashboard. What insights or information do you want to convey? Keep this objective in mind throughout the design process.

2. Audience Analysis: Understand who will be using the dashboard. Tailor the design, layout, and content to meet the needs of your target audience.

3. Data Visualization: Choose the appropriate visualizations (tables, charts, graphs) that effectively communicate your data. Use different types of visuals for different types of data.

4. Layout and Composition: Organize the dashboard layout logically, placing related visuals together. Use grids, alignment, and whitespace to enhance readability.

5. Color Palette: Select a consistent color palette that aligns with your data and branding. Use color to highlight key information and trends, but avoid overwhelming the dashboard.

6. Interactivity: Use interactive features like slicers, filters, and drill-through actions to allow users to explore the data on their terms. This enhances engagement.

7. DAX Measures Integration: Incorporate the calculated DAX measures you've created into the visuals to provide dynamic and insightful data.

8. KPIs: Highlight key performance indicators (KPIs) prominently. These are the metrics that are most critical to your analysis.

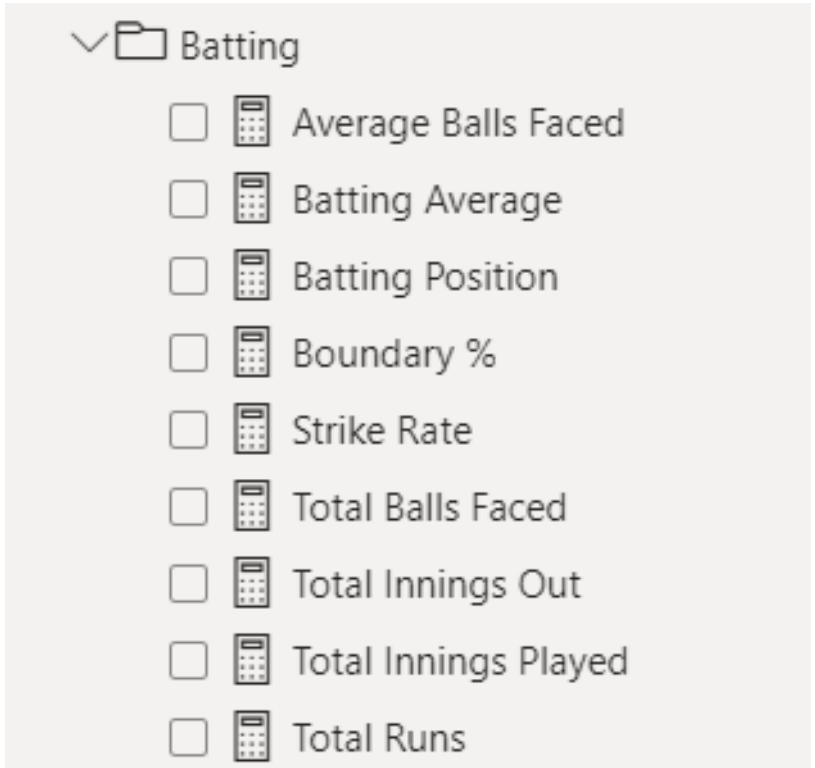
9. Storytelling: Organize the visuals in a way that tells a clear and cohesive story. Guide the users through the data, explaining trends and insights.

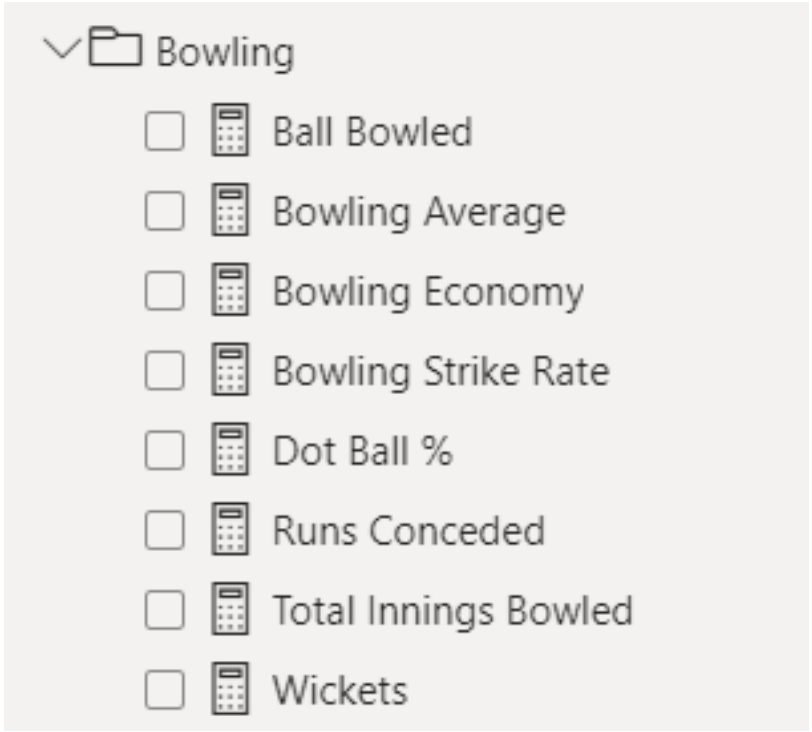
10. Responsive Design: Ensure your dashboard is responsive and looks good on different devices, including desktops, tablets, and smartphones.

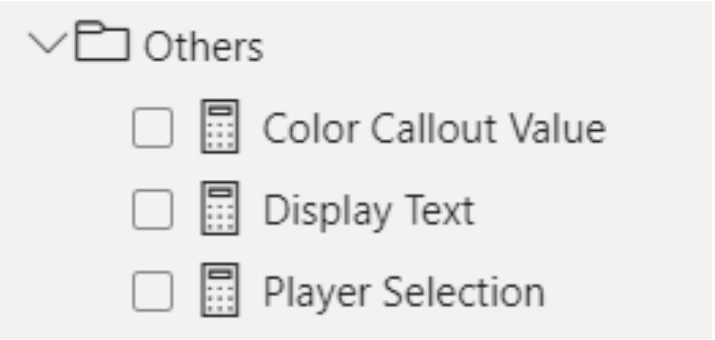
11. Testing: Test the dashboard thoroughly to ensure that all the interactions, filters, and calculations work as intended.

12. Feedback Loop: Gather feedback from users and stakeholders to make iterative improvements to the dashboard.

DAX Measures







Player Analysis | Final 12

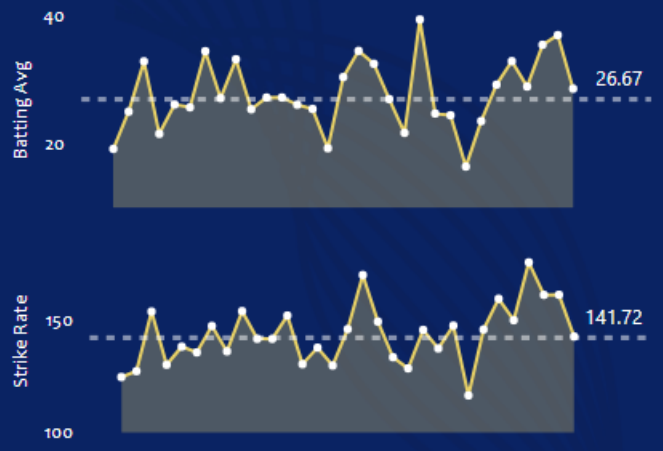
INDIAN PREMIER LEAGUE

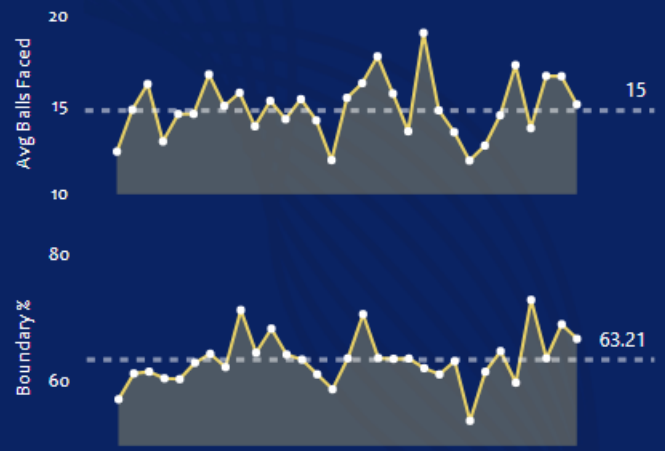
Power Hitters | Anchors | Finisher | All Rounder | Fast Bowlers

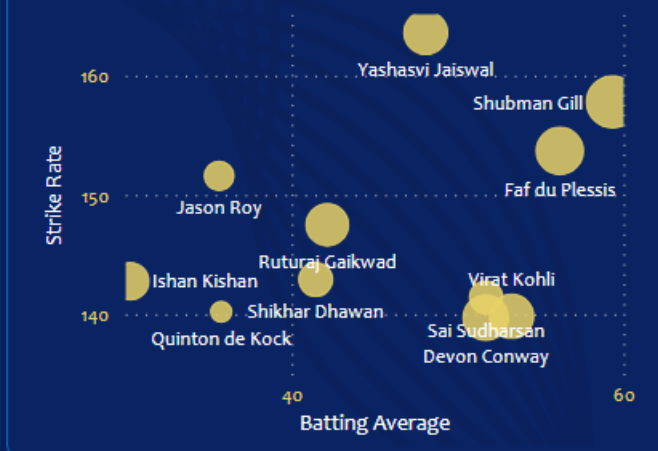
Power Hitters / Openers

Select Player(s) by clicking the player's name to see their individual or combined strength.

Name	Team	Batting Style	Innings Batted	Batting Position	Runs	Balls Faced	Batting S/R	Batting Average	Boundary %
Shubman Gill	Gujarat Titans	Right hand Bat	17	2	890	564	157.80	59.33	60.45
Faf du Plessis	Royal Challengers Bangalore	Right hand Bat	14	2	730	475	153.68	56.15	62.47
Devon Conway	Chennai Super Kings	Left hand Bat	15	2	672	481	139.71	51.69	61.90
Virat Kohli	Royal Challengers Bangalore	Right hand Bat	14	1	639	457	139.82	53.25	55.71
Yashasvi Jaiswal	Rajasthan Royals	Left hand Bat	14	1	625	382	163.61	48.08	77.44
Ruturaj Gaikwad	Chennai Super Kings	Right hand Bat	15	2	590	400	147.50	42.14	61.69







Player Analysis | Final 12

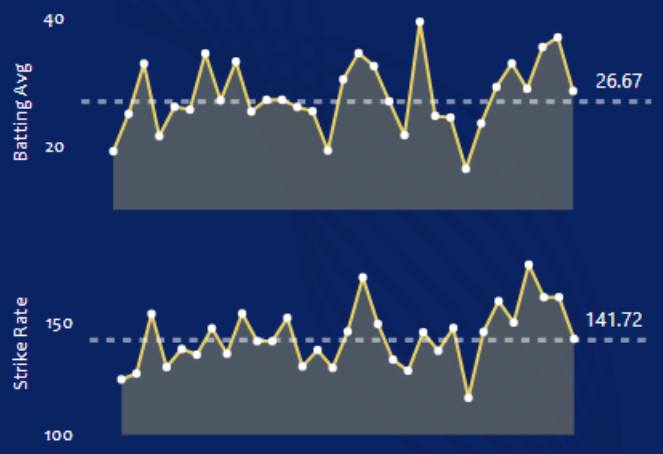
INDIAN PREMIER LEAGUE

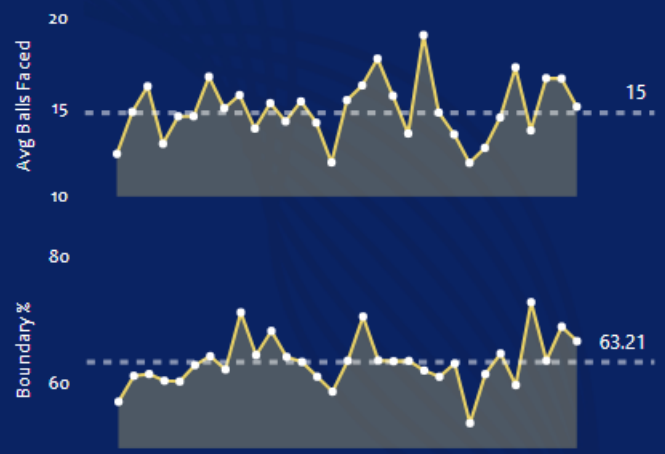
Power Hitters | Anchors | Finisher | All Rounder | Fast Bowlers

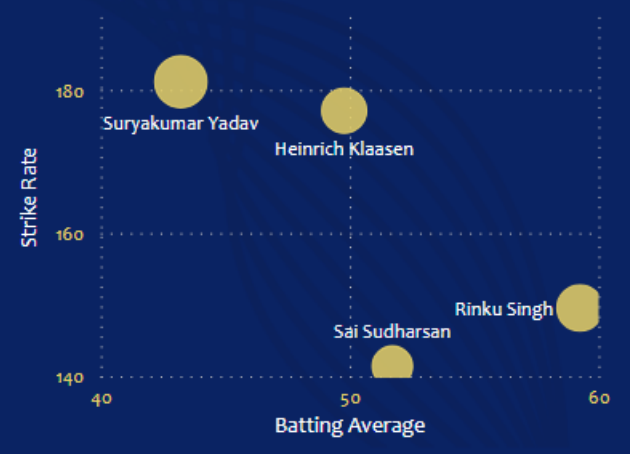
Anchors / Middle Order

Select Player(s) by clicking the player's name to see their individual or combined strength.

Name	Team	Batting Style	Innings Batted	Batting Position	Runs	Balls Faced	Batting S/R	Batting Average	Boundary %
Suryakumar Yadav	Mumbai Indians	Right hand Bat	16	4	605	334	181.14	43.21	70.74
Rinku Singh	Kolkata Knight Riders	Left hand Bat	14	6	474	317	149.53	59.25	62.87
Heinrich Klaasen	Sunrisers Hyderabad	Right hand Bat	11	6	448	253	177.08	49.78	62.05
Sai Sudharsan	Gujarat Titans	Left hand Bat	8	3	362	256	141.41	51.71	56.35







Player Analysis | Final 12

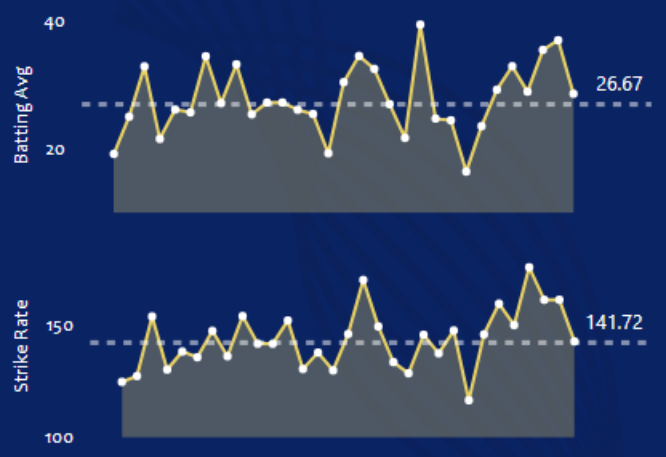
INDIAN PREMIER LEAGUE

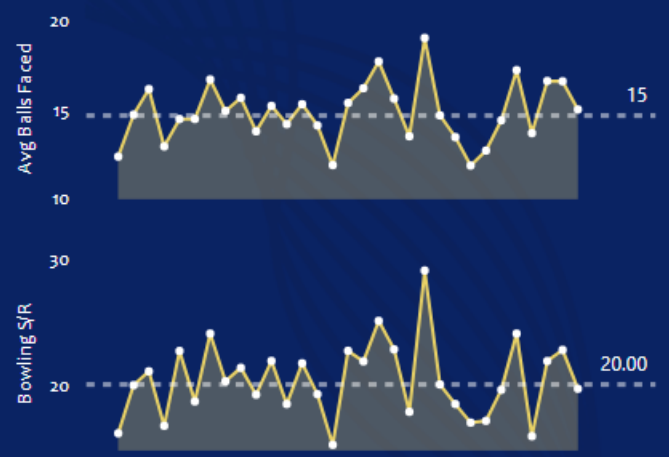
Power Hitters | Anchors | Finisher | All Rounder | Fast Bowlers

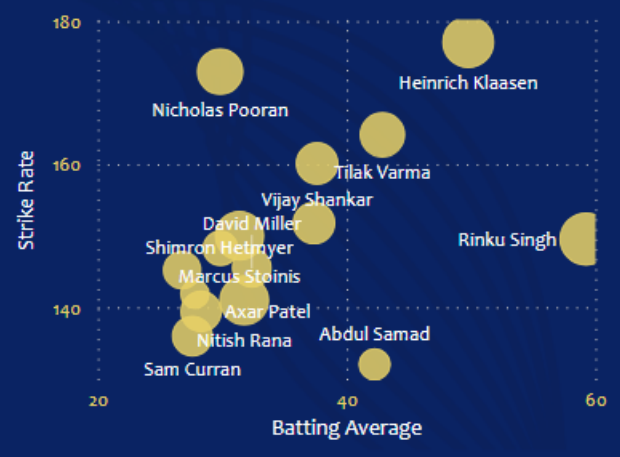
Finisher / Lower Order

Select Player(s) by clicking the player's name to see their individual or combined strength.

Name	Team	Batting Style	Innings Batted	Batting Position	Runs	Balls Faced	Batting S/R	Batting Average	Boundary %	Bowling Economy	Bowling S/R
Heinrich Klaasen	Sunrisers Hyderabad	Right hand Bat	11	6	448	253	177.08	49.78	62.05		
Nicholas Pooran	Lucknow Super Giants	Left hand Bat	15	6	358	207	172.95	29.83	72.63		
Tilak Varma	Mumbai Indians	Left hand Bat	11	5	343	209	164.11	42.88	70.55	7.0	0.00
Vijay Shankar	Gujarat Titans	Right hand Bat	10	5	301	188	160.11	37.63	65.78		
Shimron Hetmyer	Rajasthan Royals	Left hand Bat	13	6	299	197	151.78	37.38	62.21		
Marcus Stoinis	Lucknow Super Giants	Right hand Bat	15	5	408	272	150.00	31.38	67.16	9.2	13.00







Player Analysis | Final 12

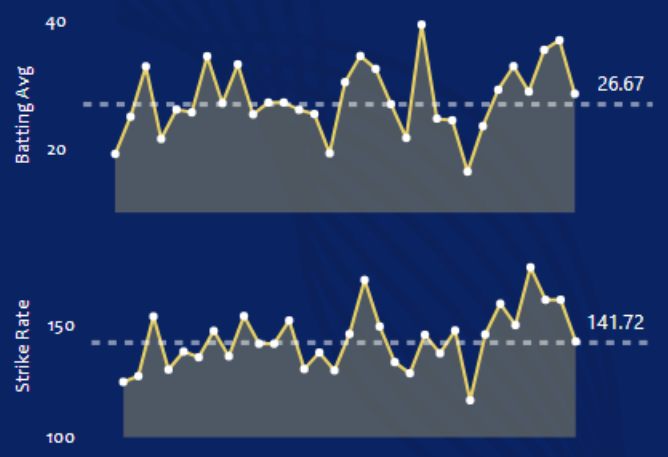
INDIAN PREMIER LEAGUE

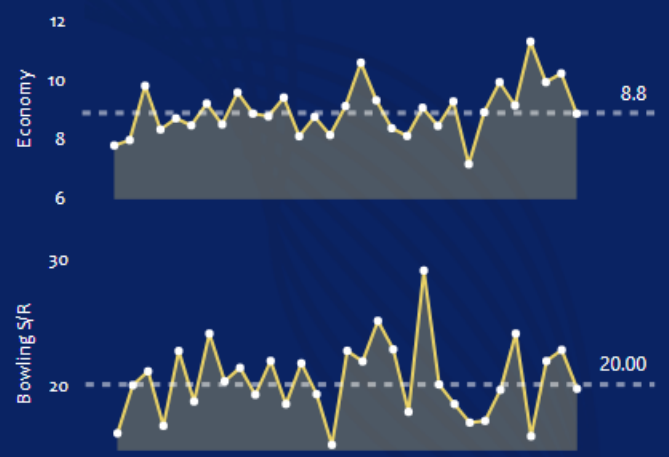
Power Hitters | Anchors | Finisher | All Rounder | Fast Bowlers

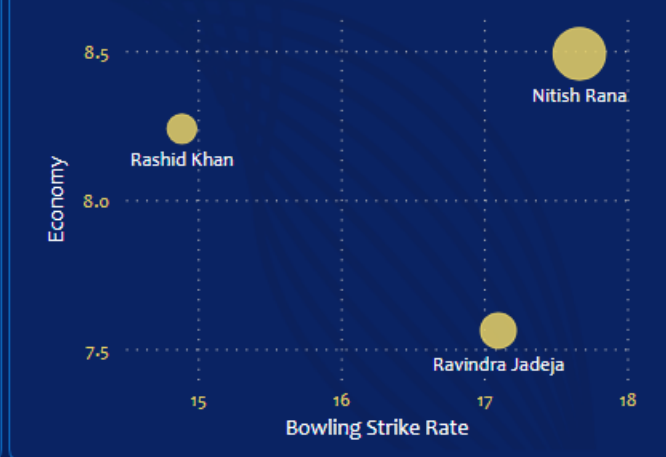
All Rounder

Select Player(s) by clicking the player's name to see their individual or combined strength.

Name	Team	Batting Style	bowling_style	Innings Batted	Runs	Batting Average	Batting S/R	Innings Bowled	Wickets	Bowling Economy	Ball Bowled	Bowling S/R
Rashid Khan	Gujarat Titans	Right hand Bat	Legbreak Googly	9	130	32.50	216.67	17	27	8.2	402	14.89
Ravindra Jadeja	Chennai Super Kings	Left hand Bat	Slow Left arm Orthodox	12	190	23.75	142.86	16	20	7.6	342	17.10
Nitish Rana	Kolkata Knight Riders	Left hand Bat	Right arm Offbreak	14	413	31.77	140.96	6	3	8.5	53	17.67







Player Analysis | Final 12

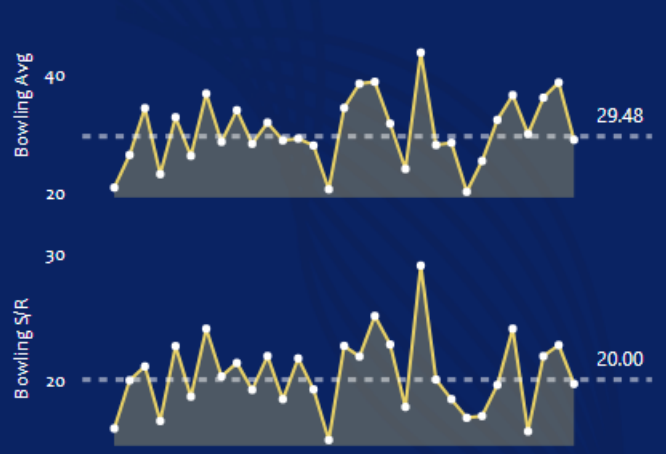
INDIAN PREMIER LEAGUE

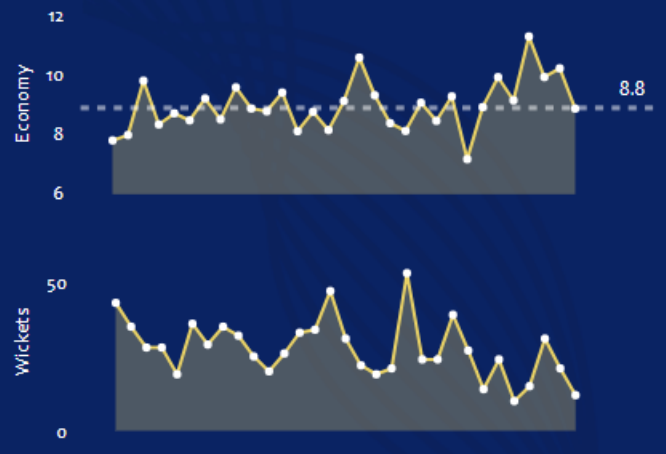
Power Hitters | Anchors | Finisher | All Rounder | Fast Bowlers

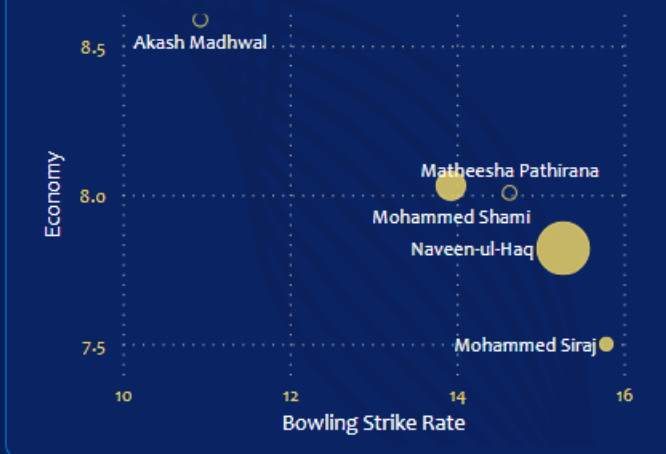
Fast Bowlers

Select Player(s) by clicking the player's name to see their individual or combined strength.

Name	Team	Bowling Style	Innings Bowled	Ball Bowled	Runs Conceded	Wickets	Bowling Economy	Bowling S/R	Bowling Average	Dot Ball %	Maidens
Mohammed Shami	Gujarat Titans	Right arm Fast	17	390	522	28	8.0	13.93	18.64	49%	2
Mohammed Siraj	Royal Challengers Bangalore	Right arm Fast	14	300	375	19	7.5	15.79	19.74	54%	0
Matheesha Pathirana	Chennai Super Kings	Right arm Fast	12	278	371	19	8.0	14.63	19.53	37%	0
Naveen-ul-Haq	Lucknow Super Giants	Right arm Medium fast	7	168	219	11	7.8	15.27	19.91	40%	0
Akash Madhwal	Mumbai Indians	Right arm Medium fast	8	153	219	14	8.6	10.93	15.64	35%	0







Player Analysis | Final 12

INDIAN PREMIER LEAGUE

Select your Final 12

Search

- [ ] Aarya Desai
- [ ] Abdul Basith
- [ ] Abdul Samad
- [ ] Abhinav Manohar
- [ ] Abhishek Sharma
- [ ] Abishek Porel
- [ ] Adam Zampa
- [ ] Adil Rashid
- [ ] Aiden Markram
- [ ] Ajay Mandal
- [ ] Ajinkya Rahane
- [ ] Akash Deep
- [x] Akash Madhwal
- [ ] Akash Singh
- [ ] Akash Vasisht
- [ ] Akeal Hosein
- [ ] Alzarri Joseph
- [ ] Aman Hakim Khan

Name	Image	Team	Custom Batting Order	Role	Batting Style	Batting Average	Batting S/R	Bowling Style	Bowling Economy	Bowling S/R	Bowling Average
Faf du Plessis		Royal Challengers Bangalore	1	Middle order Batter	Right hand Bat	56.15	153.68	Legbreak			
Shubman Gill		Gujarat Titans	2	Opening Batter	Right hand Bat	59.33	157.80	Right arm Offbreak			
Virat Kohli		Royal Challengers Bangalore	2	Top order Batter	Right hand Bat	53.25	139.82	Right arm Medium			
Sai Sudharsan		Gujarat Titans	3	Top order Batter	Left hand Bat	51.71	141.41	Legbreak			
Suryakumar Yadav		Mumbai Indians	4	Batter	Right hand Bat	43.21	181.14	Right arm Offbreak			
Heinrich Klaasen		Sunrisers Hyderabad	5	Wicketkeeper Batter	Right hand Bat	49.78	177.08	Right arm Offbreak			

This all are the screenshots I provided that how the our project is looking after designing for hours and hardworking towards the our goal.

Conclusion

- In conclusion, through in-depth analysis of IPL 2023 data, we have successfully identified key players who excel in various roles on the cricket field. By leveraging advanced data analytics techniques, we were able to pinpoint power hitters, consistent performers, effective bowlers, and versatile all-rounders. This analysis not only provides a comprehensive understanding of player strengths and contributions but also aids in the creation of an optimized and unbeatable playing 11. With a data-driven approach, we have unveiled insights that contribute to strategic decision-making for team formation, match planning, and player selection, enhancing the overall competitiveness of the team.
- I tried my best to teach you how to design this dashboard, and I will give every file to you for better understanding of project.

Thank you to all
Hopefully I teach you in better way..