

A photograph of a cricket bat and a white cricket ball resting on a green grassy field. The bat is positioned vertically, with its handle pointing upwards. The ball is a standard white cricket ball, slightly worn, and is placed near the base of the bat. The background is a soft-focus green field.

Cricket Analytics using Python and PowerBI

Team Lead
Ashwini Kumar

Scrum Master
Aditya Shah

Team Members
Neel Shah
Hardik Pithadiya
Divyesh Pithadiya
Radha Sawant



OUTLINE

01

Problem Statement

Introduction, Problem Statement, Workflow

02

Preprocessing

Data Preprocessing, Exploratory Data Analysis

03

Developing KPIs

Understanding stats and developing KPIs for ranking.

04

Dashboard Creation

Move the Data to Power BI & Creating Dashboard



Problem Statement





ABSTRACT

In this project we want to mimic how computer analyst associated with IPL teams help them find best fit batsman and bowler for their team. Data analysis helps IPL team to answer following questions:

1. “Should they spend on a particular player or not?”
2. “How valuable is the player going to be for the team?”
3. “Which player should they buy and which one they shouldn’t it?”
4. “How much money should be spent on which player?”
5. “What are the values of the different players?”



PROBLEM STATEMENT

The owners and the management of Mumbai Indians want to identify the best possible players suitable for them for the next auction. They have hired you as an Analyst to help them identify the best IPL team for the 2022 auction. You need to analyze the IPL data from 2008 to 2020, and perform the following tasks:

1. Rank batsmen by year and overall based on ability
2. Rank bowlers by year and overall based on ability
3. Rank the most valuable player in IPL over all the years
4. Visualize this in an intuitive Power BI Dashboards
5. Find Similar Batsmen and Bowler



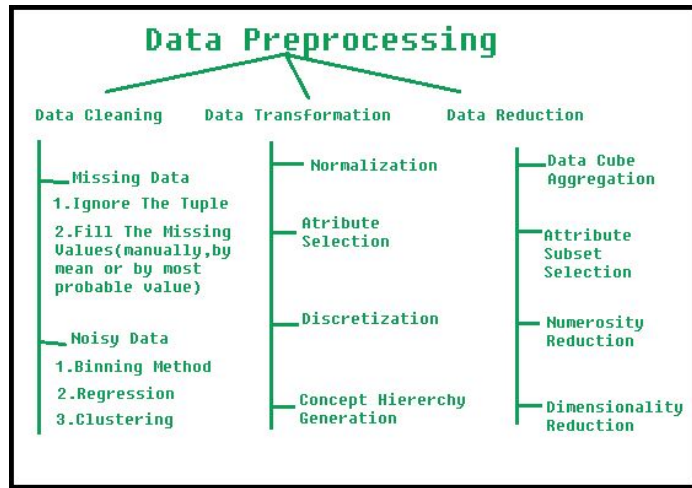
Preprocessing

Data Preprocessing

Data preprocessing is a data mining technique which is used to transform the raw data in a useful and efficient format.

Steps Involved in Data Preprocessing:

1. Data Cleaning
2. Data Transformation
3. Data Reduction





Available Data - Ball by Ball

id: Unique number for Matches	extra_runs: Extra runs scored in that ball
inning: Inning attribute (1 or 2)	total_runs: runs scored by batsman + extra runs
over: Over count (0 to 19)	non_boundary: Boundaries (4s or 6s)
batsman: Batsman on the Strike	is_wicket : 0 - not out / 1 - out
non_striker: Batsman on the Non-strike	dismissal_kind : Type of dismissal (Wicket)
bowler: Bowler bowling that particular over	Player_dismissed: Name of Dismissed Player
batsman_runs : runs scored by batsman (0 to 6)	fielder : fielder's name



Available Data - Matches

id: Unique number for Matches	toss_winner: Team who won the toss
city: City in which the match is played	toss_decision: Decision taken by the team after winning the toss
player_of_match: Name of the player who won Player of the match	winner: Team who won the match
venue: Venue in which the match is played	result: How the team won the match
neutral_venu: If it is a neutral venue or not	result_margin: By how much margin did the team won
team1: Team playing match 1	eliminator: 0: not elimination, 1: eliminator
team2: Team playing match 2	method: Method using which the team won
umpire1: Umpire 1	umpire2: Umpire 2



Preprocessing

Previous Name	Changed Name
Rising Pune Supergiant	Rising Pune Supergiants
Delhi Daredevils	Delhi Capitals
Deccan Chargers	Sunrisers Hyderabad
Pune Warriors	Rising Pune Supergiants
Kochi Tuskers	-
Gujarat Lions	-
Rising Pune Supergiants	-
City: NA	Dubai



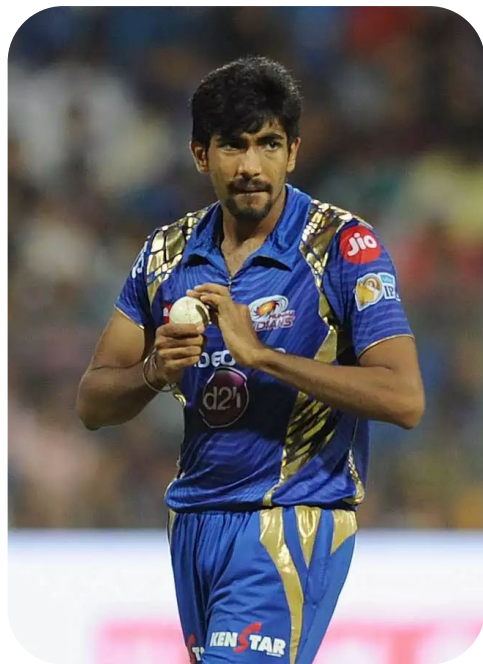
Developing KPIs

Batsman



1. Strike rate
2. Strike rate in power play and death over
3. Percentage of not out
4. Percentage of dot balls
5. Percentage of boundaries
6. Average runs per match
7. Percentage of 30+ runs scored in matches
8. Man of the match count

Bowler



1. Economy
2. Economy for Powerplay, Middle over and Death overs
3. Strike Rate
4. Average Wickets per match
5. Percentage of Dot balls
6. Percentage Boundaries
7. Bowling Average
8. 4 Wicket Haul

Team



1. Percentage wins
2. Percentage wins by chasing
3. Percentage wins by defending
4. Percentage wins by tie
5. Percentage toss wins
6. Percentage home wins
7. Number of times qualified
8. Number of times toss and match winner

Match



1. Percentage wins by chasing
2. Percentage wins by defending
3. Percentage of tie matches
4. Percentage won the toss as well as match

How Ranking Works?

Step 1: We counted the individual ranking from all the KPIs

Step 2: We counted the average rank of all the KPIs to get a final rank

Step 3: The all-rounder rank is the mean of batsman and bowler rank






Dashboard Creation

Designing

Batsman and Bowler Dashboard



BATSMAN

BOWLER

TEAM

INDIVIDUAL

Statistics

Sorting

Strike rate

%age times not dismissed

%age of dot balls

%age of boundaries

Man of the match count

%age of 30+ runs

Power play stats

Death over stats

Filtering

Strike rate

Batting average

No of matches

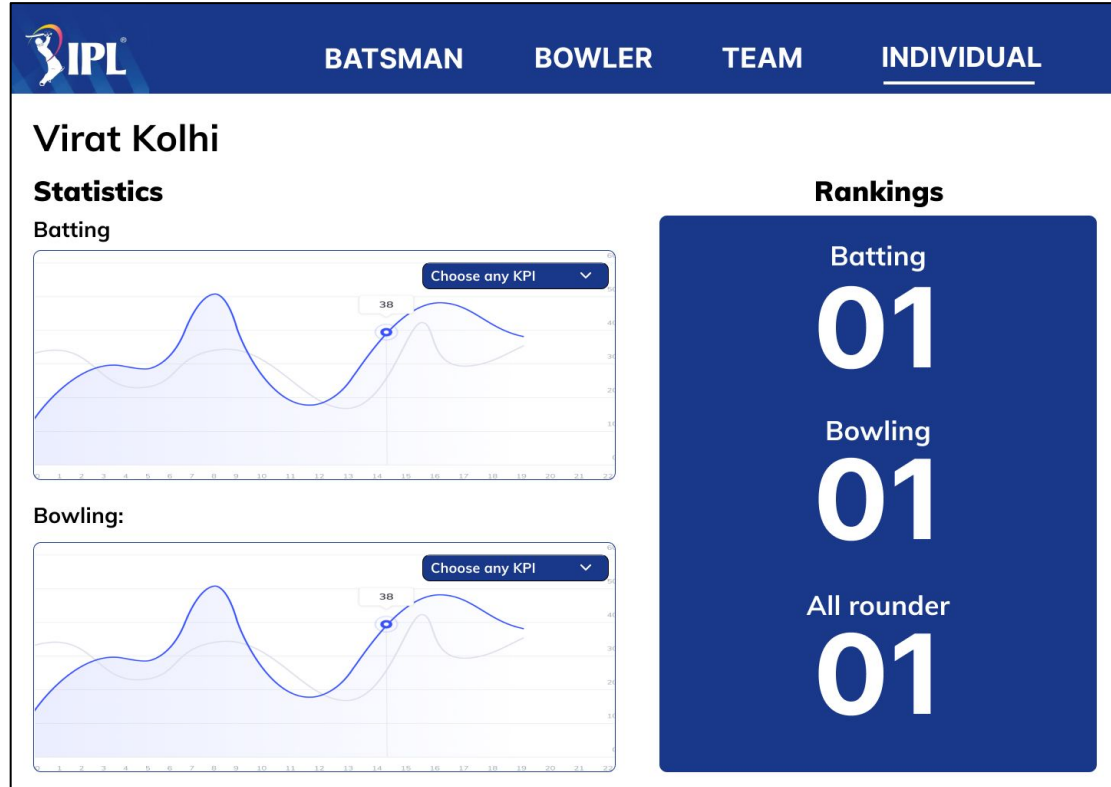
Name	Strike Rate	Not Dismissed	Boundaries	Man of the match	30+ runs scored
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%
Virat Kohli	105.94	89.99%	50.20%	50.20%	50.20%

Rows per page: 10 1-8 of 1240 < >




Designing

Individual Dashboard



Teams Dashboard




BATSMAN


BOWLER

TEAM


INDIVIDUAL




Total matches	200
Win percentage	56%
Number of times qualified	20
Chase percentage	30%
Defend percentage	30%
Home ground wins	100



Total matches	200
Win percentage	56%
Number of times qualified	20
Chase percentage	30%
Defend percentage	30%
Home ground wins	100



Total matches	200
Win percentage	56%
Number of times qualified	20
Chase percentage	30%
Defend percentage	30%
Home ground wins	100

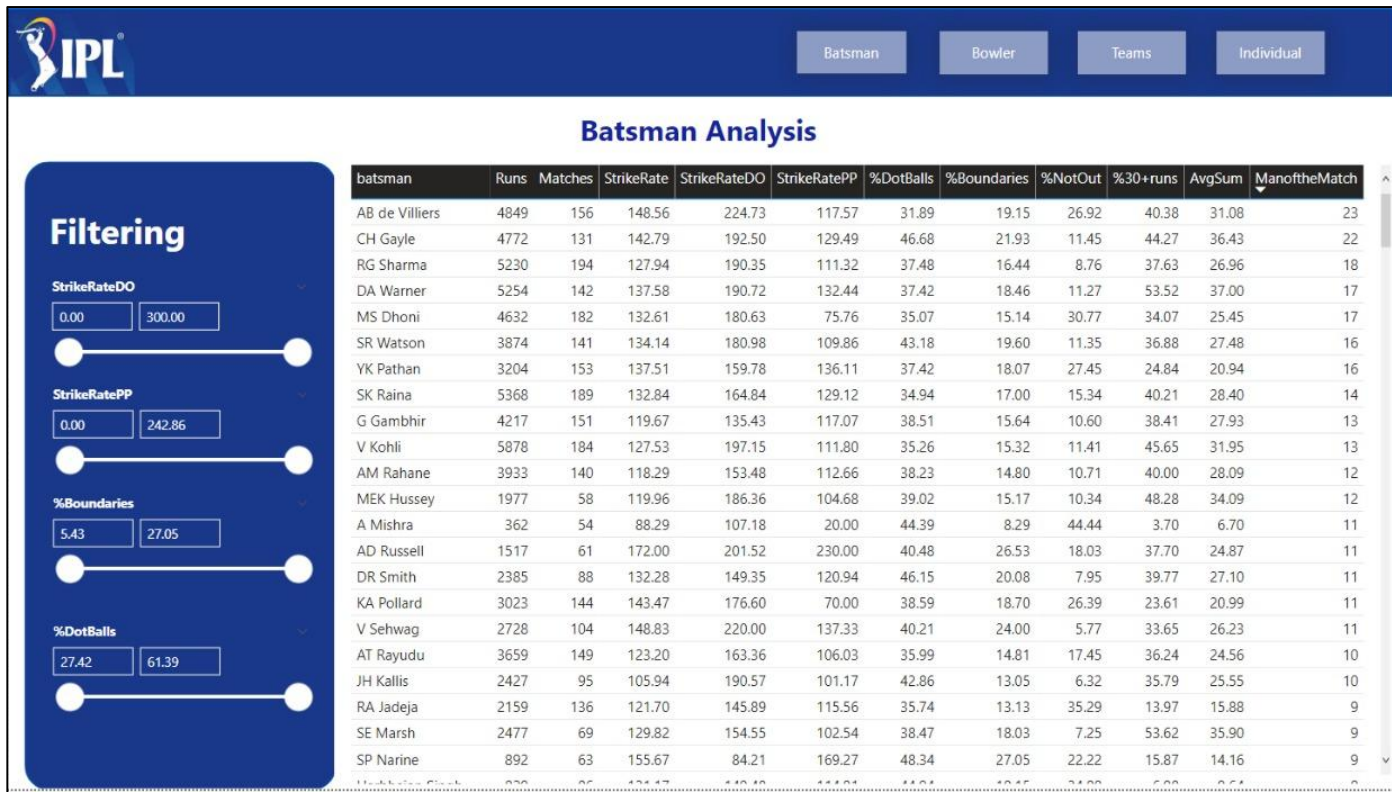


Total matches	200
Win percentage	56%
Number of times qualified	20
Chase percentage	30%
Defend percentage	30%
Home ground wins	100

	CSK	Mi	RCB	SRH	DC	KKR
CSK	0	45	44	52	55	55
Mi	0	45	44	52	55	55
SRH	0	45	44	52	55	55
RR	0	45	44	52	55	55
RPS	0	45	44	52	55	55
DC	0	45	44	52	55	55
KKR	0	45	44	52	55	55
RCB	0	45	44	52	55	55

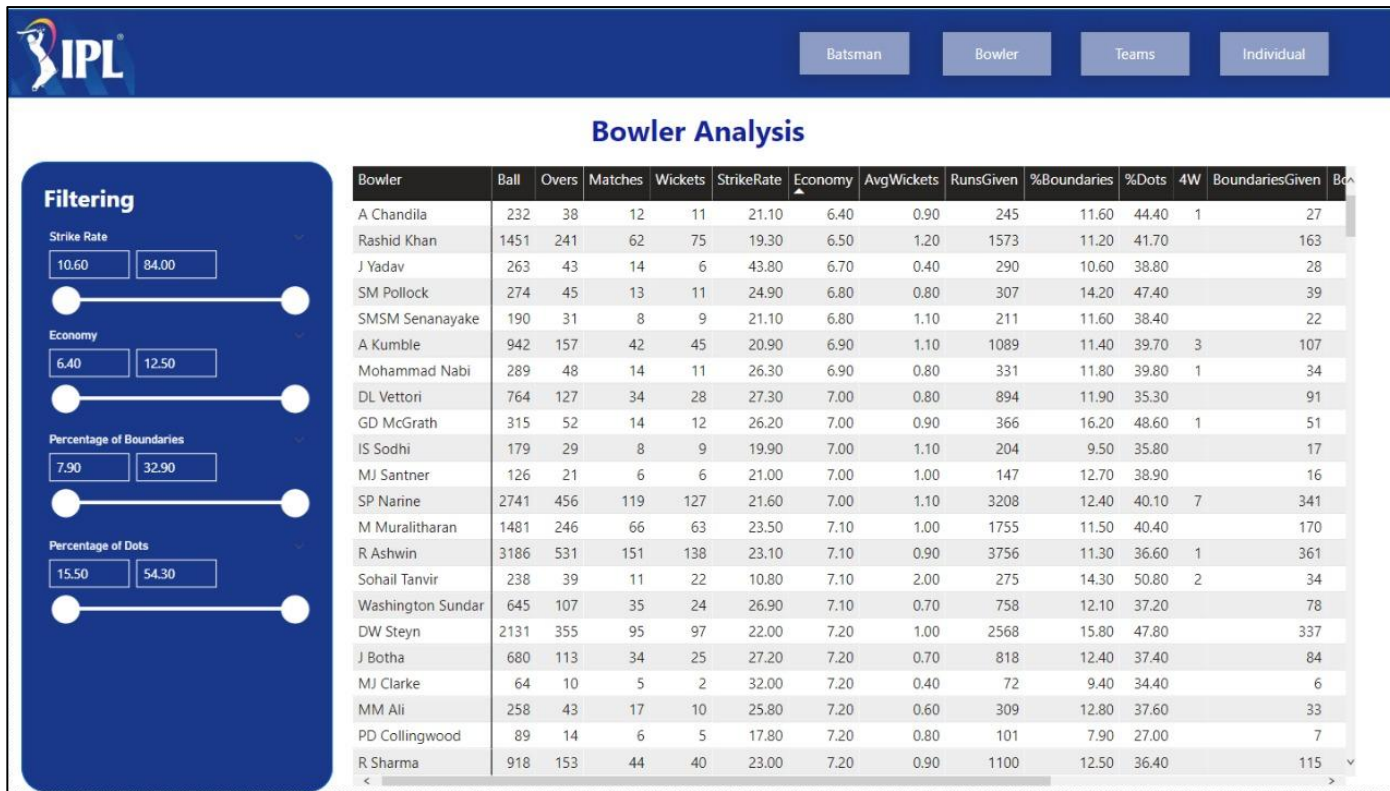
Power BI Dashboard

Batsman Analysis



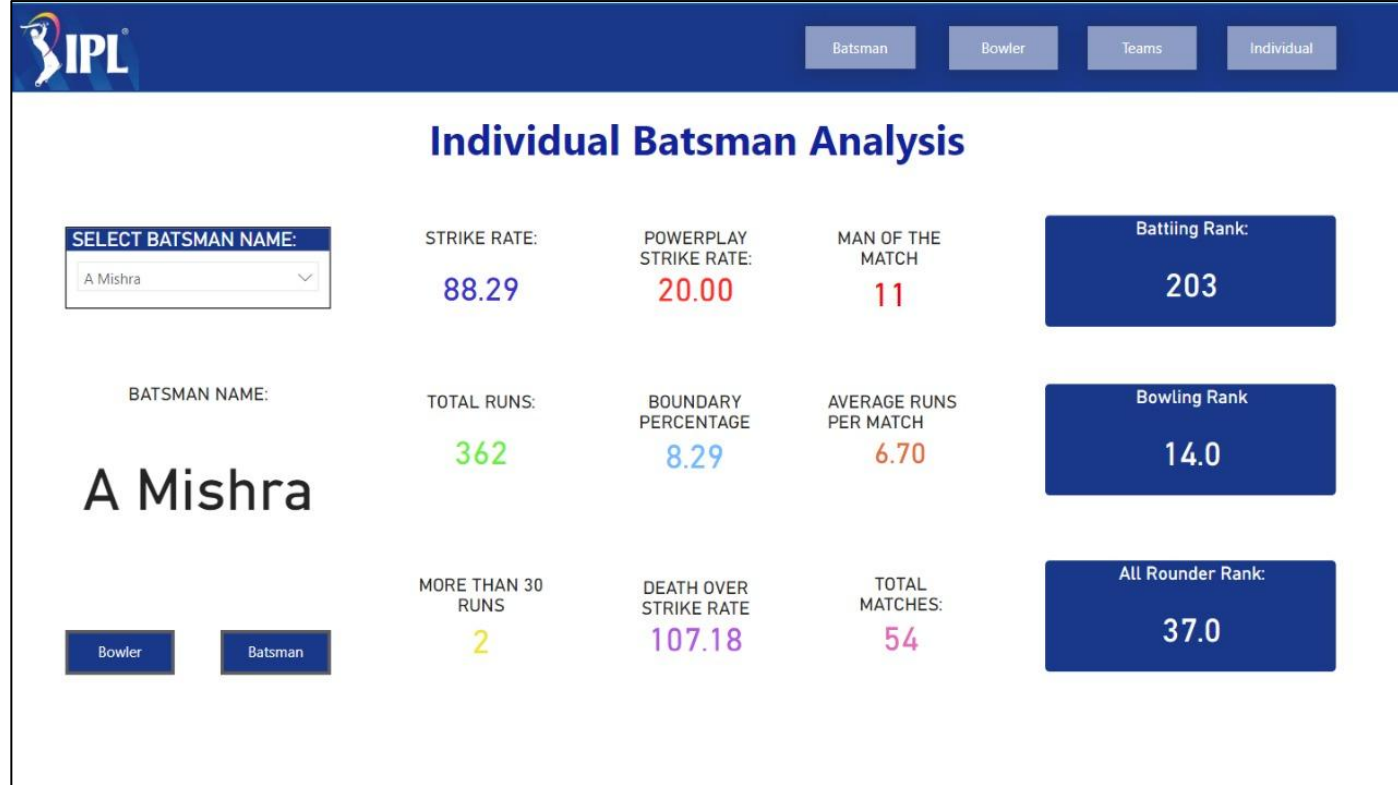
Power BI Dashboard

Bowler Analysis



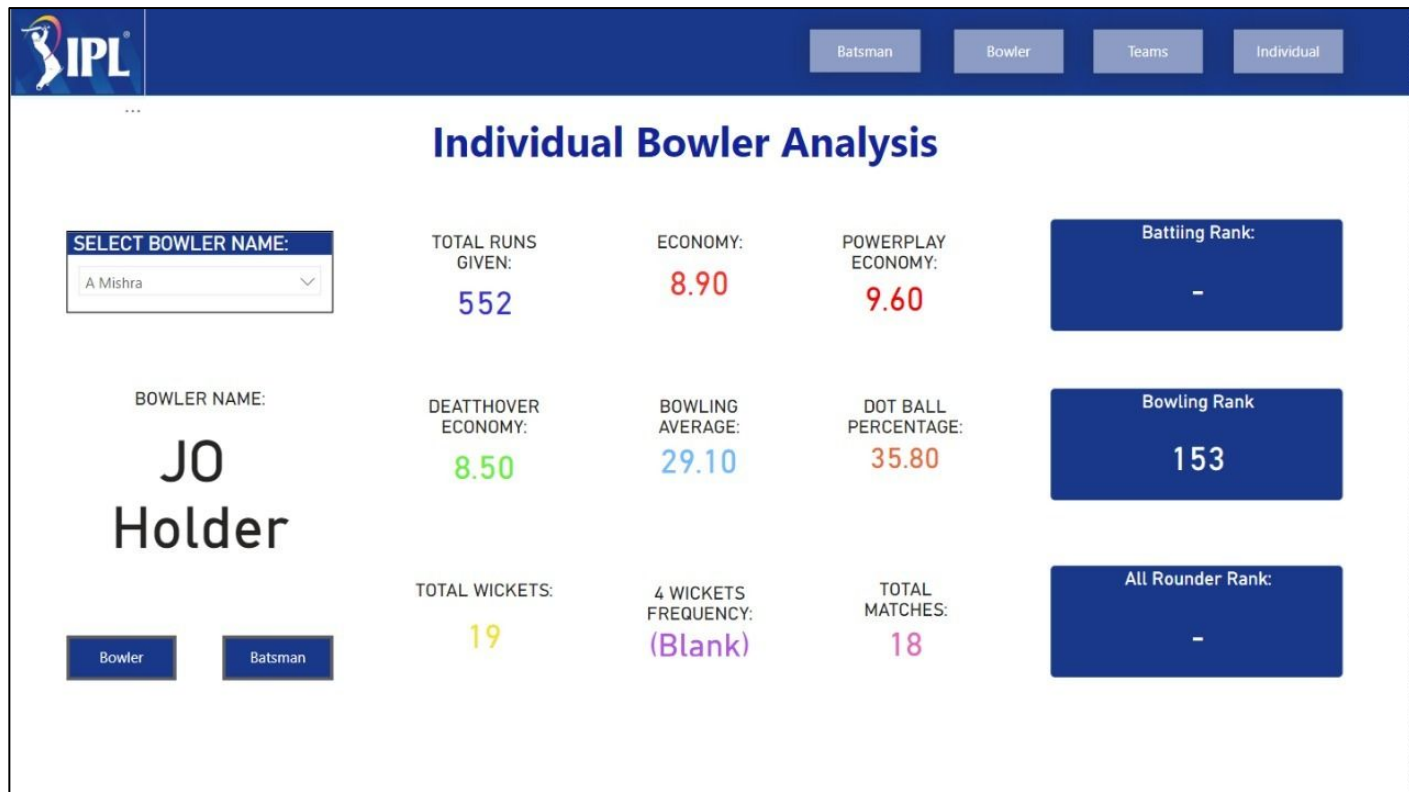
Power BI Dashboard

Individual Batsman Analysis



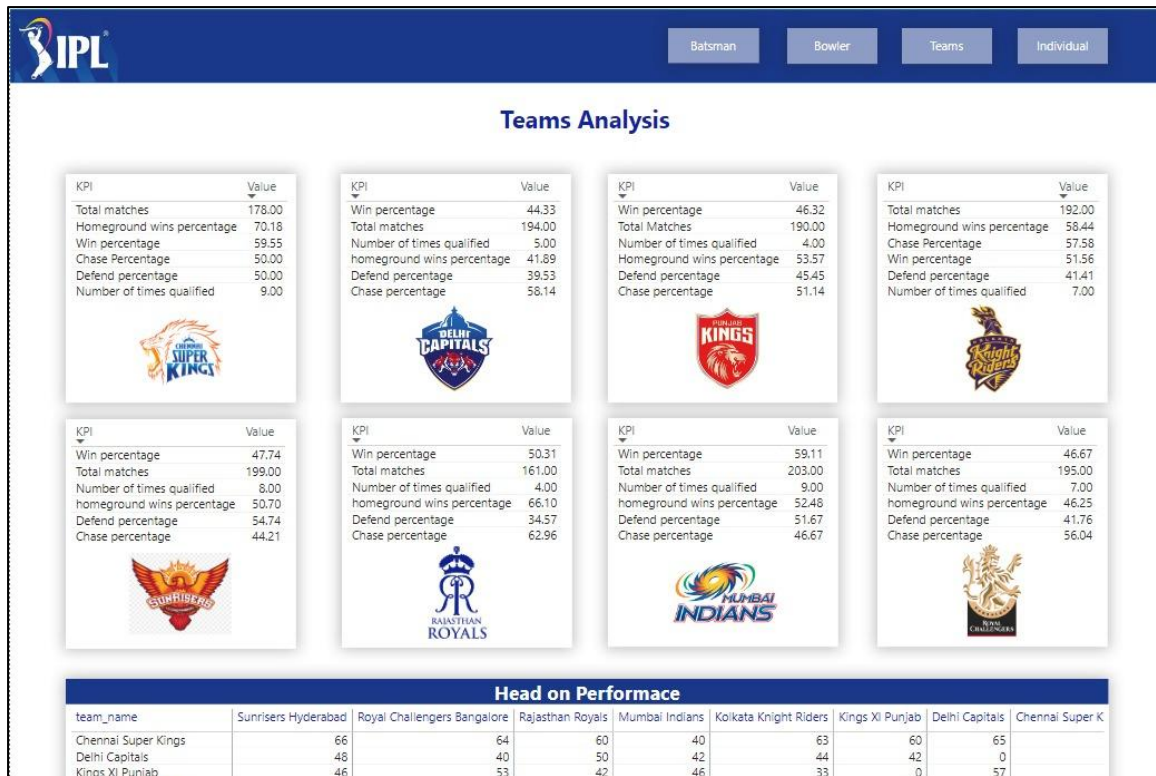
Power BI Dashboard

Individual Bowler Analysis



Power BI Dashboard

Team Analysis





Conclusion

- As we know that in cricket, player's performance depends upon the multiple factors known as Key Performance Indicator (KPIs).
- KPIs help to analyze the performance of the player and in-turn generate the Rank
- We developed the analysis based on the those KPIs which will help the management and owners to buy the best and suitable player.



Future Scope

- We would be adding different models which will help the user to identify which player is suitable for which position.

Eg. Opener, Finisher etc

- We would also be adding a player comparison and player head on statistics
- We are also thinking about adding the Value of that particular player to the team

A cricket bat and a cricket ball are lying on a green lawn. The bat is made of wood and has a dark handle. The ball is red and white. A white cloth is partially visible in the top left corner. The background is a light green gradient.

**We will now welcome any
doubts or queries**