Analysis of Indian Premiere League Dataset(2008-19)

Project Report for Indian Institure of Technology, Bombay - DS203: Programming for Data Science (2021)

1ˢᵗ Dhruv Pandey
Dept. of Mechanical Engineering
Indian Institute of Technology, Bombay
Mumbai, Maharashtra
210020048@iitb.ac.in

2ⁿᵈ Prakhar Agarwal
Dept. of Mechanical Engineering
Indian Institute of Technology, Bombay
Mumbai, Maharashtra
210020093@iitb.ac.in

3ʳᵈ Saumya Bhavesh Sheth
Dept. of Mechanical Engineering
Indian Institute of Technology, Bombay
Mumbai, Maharashtra
210020120@iitb.ac.in

4ᵗʰ Shivesh Gupta
Dept. of Mechanical Engineering
Indian Institute of Technology, Bombay
Mumbai, Maharashtra
210020130@iitb.ac.in

Abstract—**The sport of cricket is glorified by the Indian Premiere League hosted every year with great talents getting an opportunity of their lifetime to showcase themselves and finally to be the part of the National Team. We did an extensive analysis of matches held between the seasons 2008 and 2019 and found some interesting insights. We did analysis for varoius teams as well as players on different parameters to get their performance summary through the course of 12 seasons.**

Index Terms—**component, formatting, style, styling, insert**

I. INTRODUCTION

Sports analytics is a field that is becoming widely popular due to the competitive edge that it can give both to sports teams as well as stakeholders involved in the sport. Various data which is available such as the players and team statistics, environment conditions, etc is made use of to predictive models which can help stakeholders make informed decisions on the game. The main objective is to improve the performance of the team and assist in creating strategies which would help the team perfectly counter its opponents. This can be done both prior to a game as well as dynamically as the game progresses. In recent times, it has been observed that the audience themselves are also interested in the data analysis that goes on in the game and hence, sports analysts try to present this data to the audience by making simplifications to it and making use of pictorial elements such as graphs and charts to capture their attention.

A. About Cricket

Cricket is a sport that is played by two teams, each having eleven members. A team consists of batsmen, bowlers, and all rounders. The role of the batsmen is to score as many runs as possible in the limited time/overs available, while the

bowlers try to restrict the score that the batsmen try to make. Allrounders are players that play both roles and have sufficient expertise in both batting and bowling. The performance of a team depends on various factors such as the constitution of the team in terms of types of players, the venue in which the match is being held, the environmental conditions, and the type of opponents that they're playing against. Data analytics can be made use of to help the teams management figure out which players to play in a specific match, the odds of them reaching a specific stage in a tournament, the environmental conditions that they're going to play in, etc. It can also be used during a match to help the team adjust their strategy according to the state at which the match is in, to provide them a competitive edge against their opponent. These days, data science techniques are being made use of by every team that competes in the sport professionally. When used correctly, it can help teams bridge the gap in skill by formulating an effective strategy to counter their opponents.

B. About Indian Premiere League

The Indian Premier League (IPL) is the world's biggest domestic cricket tournament. It is a 20-over format of the game that makes for short, fast-paced games which is one of the reasons for its massive fanbase. It is an annual tournament and has seen 14 such tournaments conducted so far. There are 9 teams involved in the tournament and the teams themselves consist of players from all around the world. The tournament generates a large revenue and has many stakeholders heavily invested in it. So teams will do everything they can to get an edge over their opponents in a game. Data Analysis is now heavily used by all teams to try and gain this edge.

II. Datasets

For the project we took two datasets. One of them containing data about all the matches played between 2008 and 2019 and the other having data of each ball bowled in the corresponding matches. Both the datasets were taken from Kaggle and were verified using the data present on the official IPL site.

In the matches dataset each match is represented by a unique match id which is also present in the deliveries dataset. We have all the details of the matches including between which two teams was the match played, who won the match, who won the toss and what was the decision the team took after winning the toss, who were the umpires in the match, where was the match played etc. We used this data in doing match specific analysis like the team with best winning ratio, or who is the luckiest in terms of winning the toss.

The deliveries dataset is the major one of the two. It has about 180,000 rows having details of each ball bowled in the matches between season 2008 and 2019. There is a ton of detail for each ball, like the batting team, bowling team, who is the batsman on strike, who is the non striker, the bowler balling, no. of runs scored by the batsman, any extra bowled, if yes then of what type - no ball, wide ball, leg bye etc., if wicket taken, then of what type, and if it is a catch who caught it. This dataset was used in the majority of our EDA like the batsman with highest strike rate, best average, best partnership, overwise strike rates of certain batsmen, etc.

III. Data Preprossecing

The datasets matches.csv and deliveries.csv can not be used directly for predictive analysis. There are some problems like missing values, duplicate team names etc. Hence, some steps for data pre-processing are needed.

First, the null values of the 'city' column of matches.csv are filled based on the values of the 'venue' column. The 'city' values were empty where 'venue' was 'Dubai International Cricket Stadium', so the 'city' values are filled as 'Dubai'.

The datasets contain duplicate names for a team. One team has names 'Rising Pune Supergiants' and 'Rising Pune Supergiant'. The name of the team is changed to 'Rising Pune Supergiants' in the columns 'team_1', 'team_2', 'toss_winner', 'winner' of matches.csv and 'batting_team' and 'bowling_team' of deliveries.csv.

The datatype of the 'date' column of matches.csv is changed to 'DateTime' from 'object'. The columns 'umpire3' of matches.csv has a large number of missing values. Hence, it was removed.

IV. EDA Analysis

Various manipulations are performed on the available datasets to extract some insightful information from them

A. Number of matches hosted by various stadiums

We used the matches.csv dataset to calculate the no. of matches that took place at different venues and plotted a pie
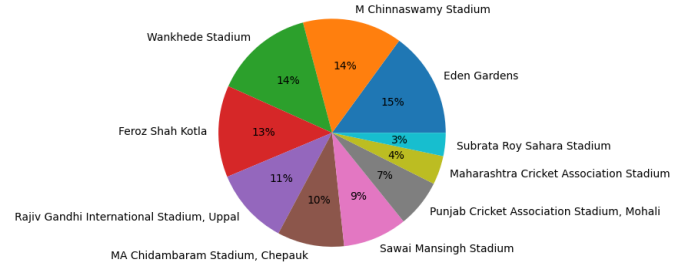

Fig. 1. Pie chart representing no. of matches hosted in different stadiums

chart representing the top ten venues where IPL matches are held.

The pie chart shows that among the top ten venues about 15% matches are held at Eden Gardens, Kolkata, followed by Chinnaswamy, Bengaluru and Wankhede Stadium, Mumbai at 14% and Feroz Shah Kotla, Delhi at 13%.

B. Number of tosses won by each team

We used matches.csv dataset to see which team is luckiest on the basis of how many tosses the team has won in the given seasons. From the plot it is clear that Mumbai Indians has won the most no. of tosses.
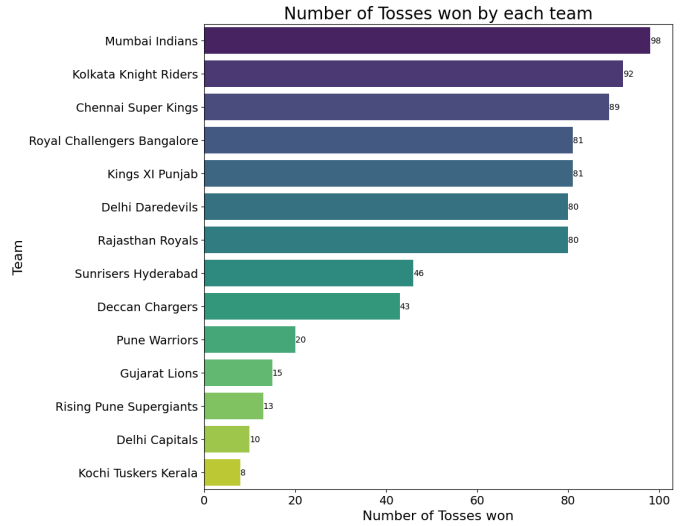

Fig. 2. Bar chart representing no. of tosses won by different teams

After that we also calculated that how fruitful the toss decisions of the team turns out to be i.e. how many times does a team wins after winning the toss. Fig 3 shows that Gujarat Lions lead in this plot with about 66.67% of fruitful toss decisions.

C. Player with most no. of catches

Next we use the deliveries.csv dataset and specifically its dismissal_kind and fielder column to find the player who took most no. of catches in the mentioned seasons.
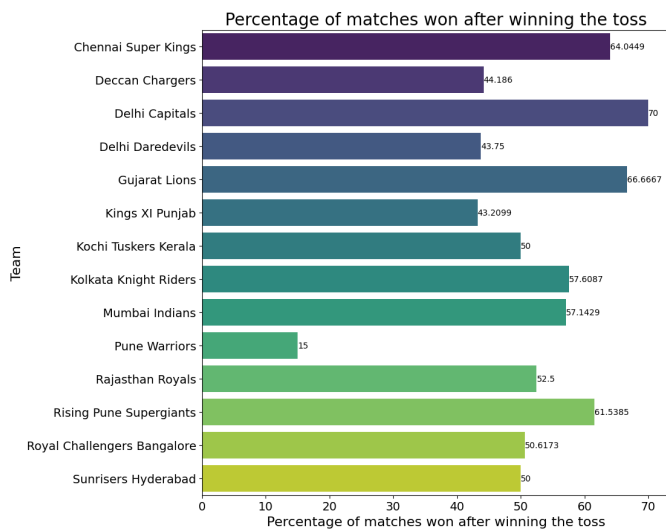
Fig. 3. Bar graph representing percentage of matches won after winning the toss

From the table in the Fig 4 we can see that MS Dhoni is leading player in terms of no. of catches taken. This is quite assuring as he is a wicket keeper who gets most of the catches in a match. This also explains that all top 3 in the list are wicket keepers for their respective teams.

Fig. 4. Table showing list of player with most no. of catches

D. Most common dismissal type

In the game of cricket there can be varios ways for a batsman to get out. Some are the usual ones like caught, bowled or run outs but some though do not occur that often can cost a batsman his innings. So lets see which is the most common way for a batsman to get out.

The pie chart is a great representative of the fact that we discussed before i.e. caught, bowled and run outs are the most usual ones. But there are others as well like stumped, retired hurt and even obstructing the field which happen not that often.
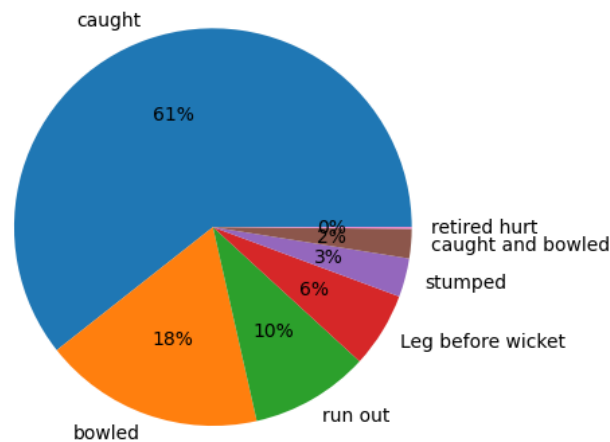

Fig. 5. Pie chart for most common dismissal kind

E. Most Valuable Player

Cricket is a team game and each player's performace is crucial in the victory of a team but there is always one player who stands out in a match and helps the team to get closer to the victory. This player is awarded the man of the match. So lets see which player has won the most man of the matches awards in their career.
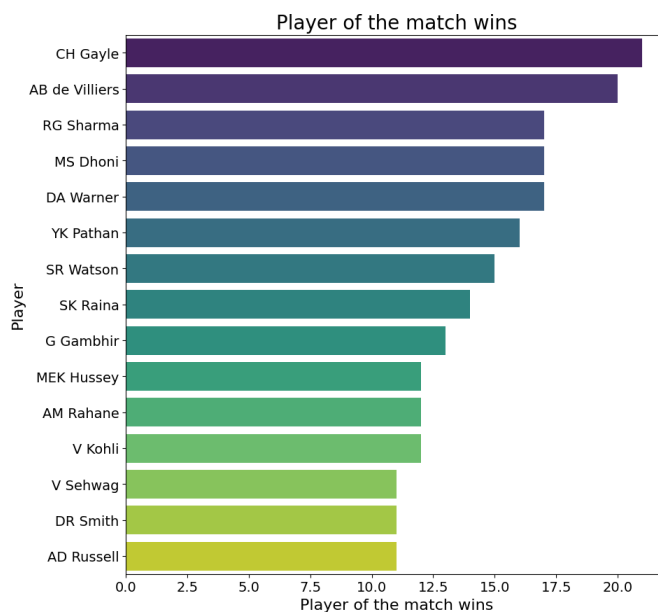

Fig. 6. Bar graph for most man of the match awards

Undoubtedly one of the greatest batsman in IPL, Chris Gayle has the most man of matches followed by AB de Villiers who is also a great batsman.

F. Greatest Victories

Now we will the greatest victories by the teams both bowling and batting first. These victories represent how dominant

they were in certain matches and demolished the opponent team badly. Fig 7 shows clearly that Mumbai Indians has had
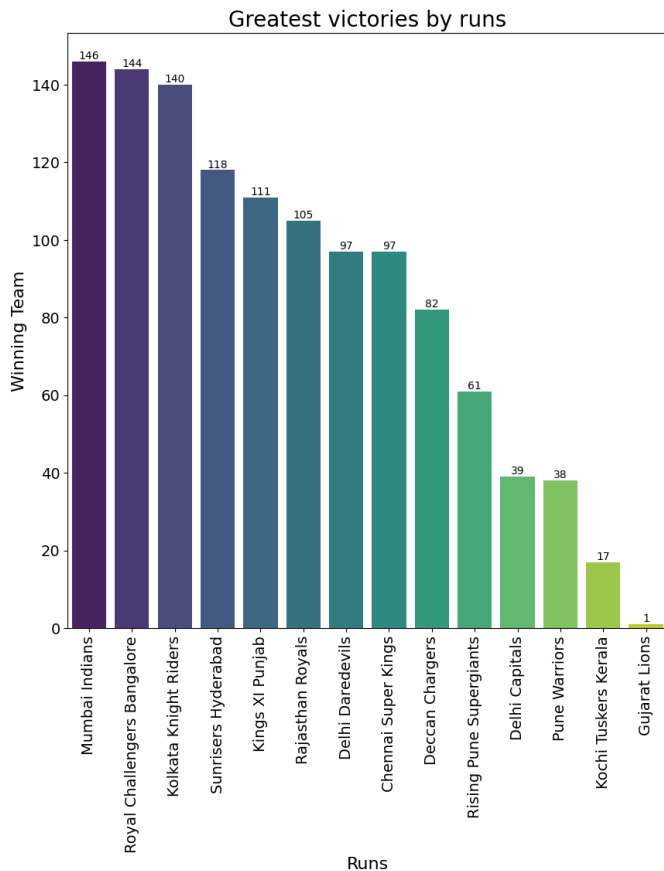
Greatest victories by runs

Fig. 7. Bar graph for most man of the match awards

the greatest victory by runs in the IPL.

For the greatest victory by wickets many teams have won the match by 10 wickets including Royal Challengers Bangalore, Sunrisers Hyderabad, Mumbai Indians, Kolkata Knight Riders, Kings XI Punjab, Delhi Daredevils, Chennai Super Kings, Rajasthan Royals, Deccan Chargers.

G. Most wickets by a player

Being a bowler in this format of the game is not usually a player's first choice because of the ruthless smashing by the batsman. Still the bowler's role is as important as that of a batsman and a team with a good bowler often gets an edge above other teams in the tournament.

From the table it is clear that Lasith Malinga is at the top with a whopping 159 wickets, followe by Amit Mishra and Harbhajan Singh.

H. Most 4ś and 6ś by a batsman

IPL is played in the shortest format thus requires a lot of power hitting including a lot of 4ś and 6ś. Thus a batsman having the potential to hit the big boundaries is very effective in the match and leads the team to a good score.

	Bowler Name	No. of wickets
0	SL Malinga	159
1	A Mishra	143
2	Harbhajan Singh	142
3	DJ Bravo	141
4	PP Chawla	133
5	B Kumar	125
6	R Ashwin	111
7	UT Yadav	110
8	SP Narine	106
9	A Nehra	101

Fig. 8. Most wickets by a bowler

In the 12 seasons mentioned the most no. of 4ś are hit by Shikhar Dhawan(526 fours) and the most no. of sixes are undoubtedly hit by Chris Gayle (376 sixes).

I. Percentage wins batting first and bowling first

Some teams are good and chasing while some teams are good at defending. Other than the team it also depends on the opponent, venue and many other factors. So lets see for what all teams batting first is better.
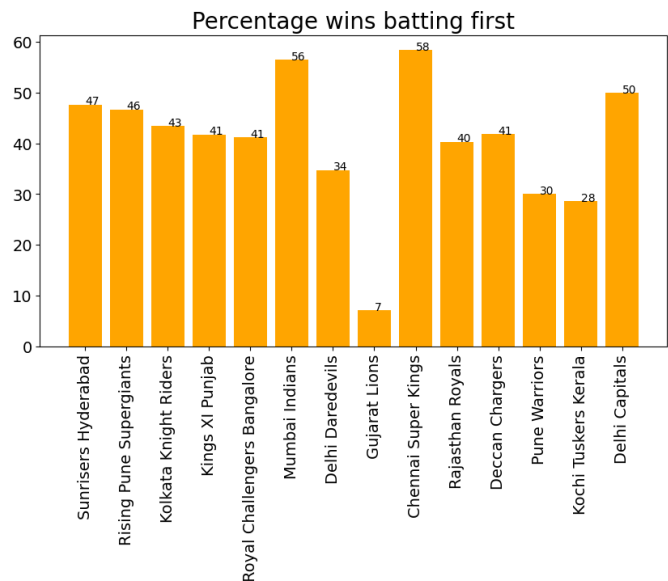
Percentage wins batting first

Fig. 9. Bar graph showing percentage of wins batting first

The above graph show that Chennai Super Kings is the best at batting first closely followed by Mumbai Indians.

The bar graph for percentage wins while batting second shows a great variation with the team with highest percentage being Gujarat Lions with 75% wins followed by Chennai

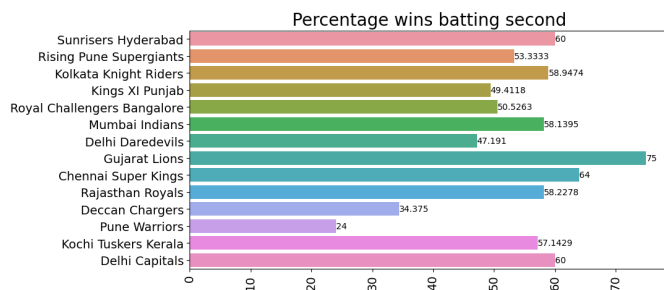Super Kings. The analysis shows that Chennai Super Kings is a good team both at batting and bowling first.

Fig. 10. Bar graph showing percentage of wins batting second

J. Best team in the powerplay overs

The powerplay overs are the first six overs of the match where only 2 fielders are allowed outside the 30 yard circle. It is a crucial part of the game where you set the tone for the rest of the game. A good powerplay gives the players a boost to perform well in the rest of the match. The bowling team looks to restrict the batting team to as low score as possible while taking as many wickets while the batting team looks to score as high as possible while giving away less no. of wickets.
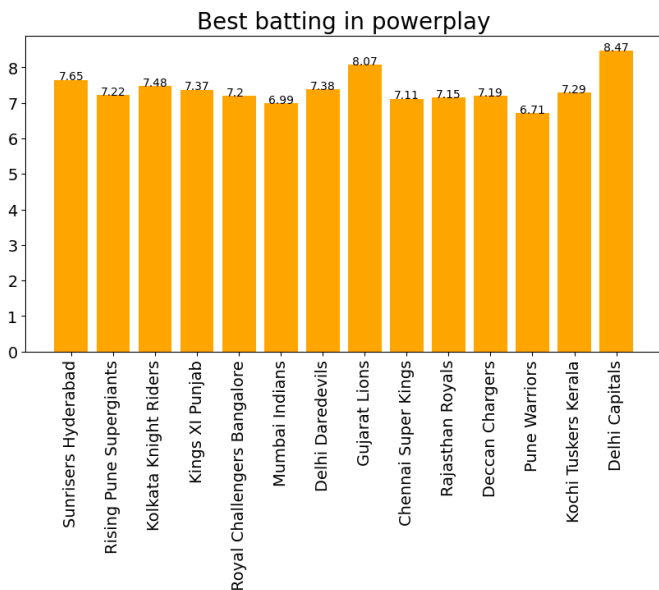

Fig. 11. Bar graph showing best batting in powerplay in terms of run rate per over

In the batting in powerplay graph it shows that Delhi Capital is the best at scoring runs in the initial overs of the match. While the bowling in powerplay graph shows that Kochi Tuskers Kerela is the most effective team in taking wickets in the powerplay.

K. Most Teams in which a Player has played for

The IPL tries to do full justice to all the teams such that after every 2 years it holds a major auction where the teams
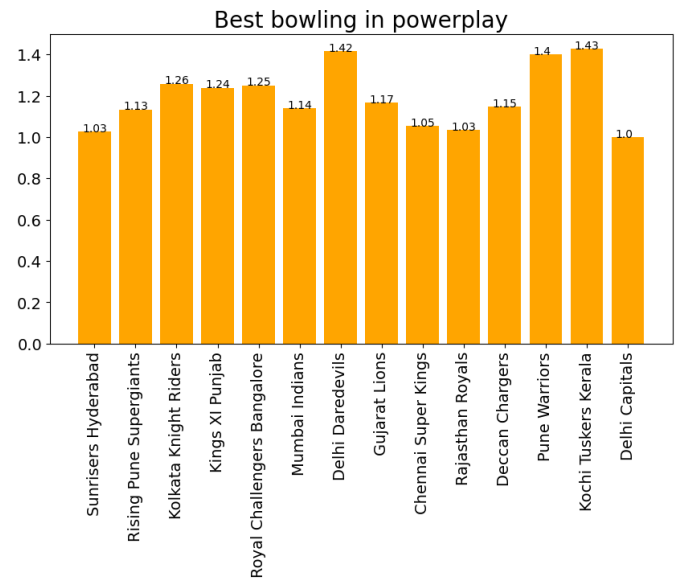

Fig. 12. Bar graph showing best bowling in powerplay in terms of wickets per match

can bid for all the available players and the player goes to the highest bidder. Thus a player might not play for the same team in all the seasons. Now we will see which players have changed their teams the most no. of times.
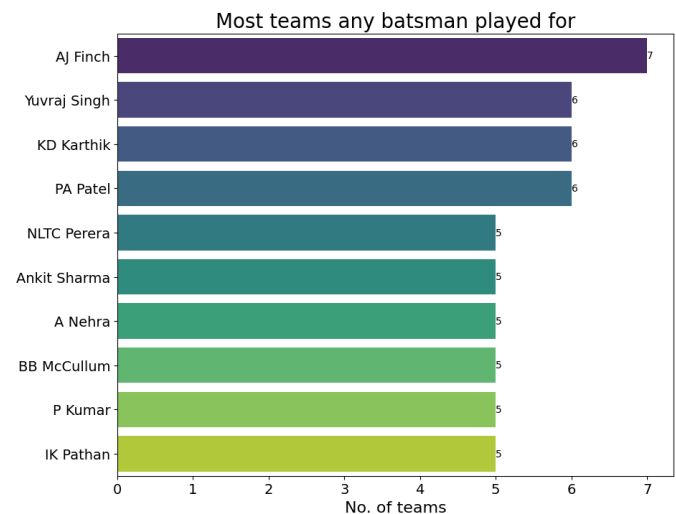
Fig. 13. Bar graph showing no. of teams a player has played for

The graph shows tha Aaron Finch has played for the most no. of franchises i.e. 7 followed by Yuvraj Singh, Dinesh Karthik and Parthiv Patel who all have played for 6 different teams.

L. Most Hundreds and Fifties

In the shortest format of the game, even scoring 30 is a big deal but still players manage to score hundreds and even hundred and fifty runs in a single innings just showing the potential a single international player has in terms of skills.

Fig. 14. Line plot showing no. of fifties by the top 10 players

Fig. 15. Line plot showing no. of hundreds by the top 10 players

M. Average Runs by Different batsmen

Consistency is very important and that is reflected by average of a batsman. Higher the average indicates that the batsman is more consistent and average is calculated by total runs divided by total number of matches he played except the ones where he remains not out. But for 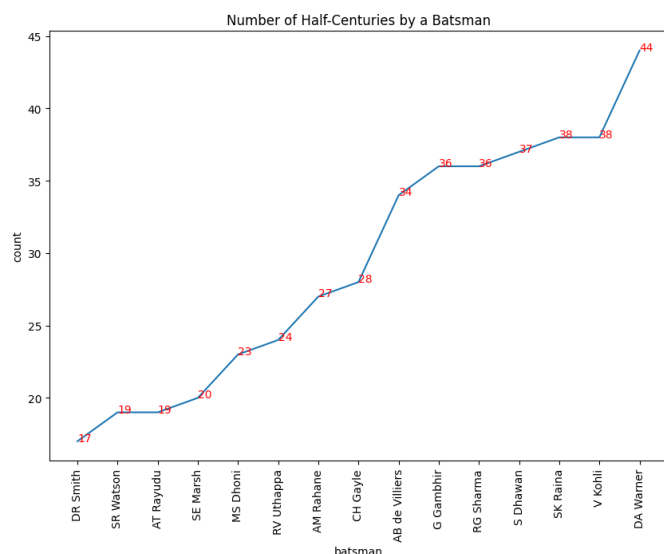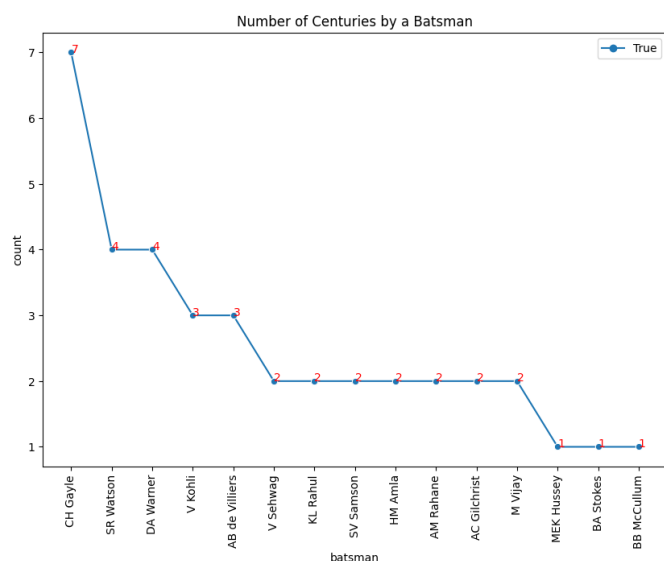this project we have calculated it by dividing total runs by total no of matches. Below table depicts the averages of different batsmen. From the figure we can see that Jonny Bairstow has the highest average followed by David Warner and so on.

N. Overall strike rate of different batsman

Strike rate plays a key role especially when it comes to the T20 Cricket. Strike rate determines how quickly a batsman

Fig. 16. Average of different batsmen

can score runs. It is defined by runs scored divided by balls taken multiplied by 100. Players with good strike rate are very important for every team as they help in accelerating the run rate while batting. Below is the table which shows the overall strike rate of various matches. From the table we can see that Andre Russell has the highest strike rate followed by sunil narine and these can be verified by their hitting capabilities.

O. Most Extras by a Bowler

Though most people ignore or do not pay much attention to extras but they prove to be crucial especially in a close match. Extras in a close match at the end overs can cost the whole match and so the extras cannot be ignored. Bowlers should try to minimize the extras conceded to the batting team. The line and length of the ball is very important. Below table depicts which bowlers concede higher extras.

From the table we can see that Lasith Malinga has given the highest number of runs as extras.

P. Average Runs Scored by teams in first innings

Many teams prefer chasing in a t20 match because you can play according to the target set but every time luck is not in your favour. Scoring a good total while batting first when you are initially examining how the pitch is behaving is not that easy. Let's have a look at how different teams have batted in the first innings when they had to bat first. Setting a target for opponent is not that easy. Below is the bar plot depicting average score made by the teams while batting first.

Batsman	Strike Rate
AD Russell	179.950187
SP Narine	166.943867
RR Pant	162.318841
GJ Maxwell	155.543237
CH Morris	153.392330
HH Pandya	151.902174
JC Buttler	150.000000
V Sehwag	148.827059
AB de Villiers	148.740343
CH Gayle	145.640370
KH Pandya	144.094488
KA Pollard	141.751527
P Shaw	141.609195
DA Warner	139.523249
CA Lynn	139.297125
YK Pathan	138.860326
HM Amla	138.038278
KL Rahul	137.406143
SR Watson	136.945813
JA Morkel	136.938202
DA Miller	136.264535

Fig. 17.   Overall Strike rate of batsmen

	Bowler Name	Extras
0	SL Malinga	293
1	P Kumar	236
2	UT Yadav	219
3	DJ Bravo	201
4	B Kumar	197
5	I Sharma	194
6	RP Singh	181
7	DW Steyn	171
8	SR Watson	171
9	Harbhajan Singh	170

Fig. 18.   Most extras conceded by a bowler

From the graph we can see that delhi capitals has the highest average runs scored in first innings. So we can say that delhi capitals is good at batting first
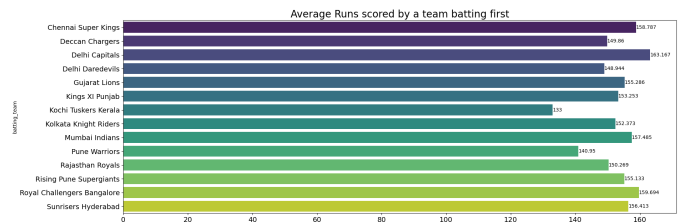

Fig. 19.   Average Runs Scored by different teams in First Innings

Q.   Top Run Scorers for each team

Here We are looking at the players who has performed the best for their team in batting. Every team has a key batsman who plays consistently and on whom team can trust that he can win the match for the team. Below is the table showing highest run scorers of each team.

	Team Name	Batsman
0	Sunrisers Hyderabad	DA Warner
1	Royal Challengers Bangalore	V Kohli
2	Mumbai Indians	RG Sharma
3	Rising Pune Supergiants	AM Rahane
4	Gujarat Lions	SK Raina
5	Kolkata Knight Riders	G Gambhir
6	Kings XI Punjab	SE Marsh
7	Delhi Daredevils	V Sehwag
8	Chennai Super Kings	SK Raina
9	Rajasthan Royals	AM Rahane
10	Deccan Chargers	AC Gilchrist
11	Kochi Tuskers Kerala	BB McCullum
12	Pune Warriors	RV Uthappa
13	Delhi Capitals	S Dhawan

Fig. 20.   Top Run Scorers for each team

From the table we can see that for sunrisers hyderabad, David Warner was their top run scorer and for Royal Challengers Bangalore Virat Kohli is their top run scorer and so on.

R. Top run scorers against a team

By this, we are trying to analyze who has scored highest runs against a team and so that batsman has to play when there is a match against that team as he has best records against that team.

From the table we can see that against Royal Challengers Bangalore, MS Dhoni has played the best; then against Chennai Super Kings Virat Kohli has played the best and so on.

Team Name	Batsman
Royal Challengers Bangalore	MS Dhoni
Sunrisers Hyderabad	SR Watson
Rising Pune Supergiants	V Kohli
Mumbai Indians	SK Raina
Kolkata Knight Riders	DA Warner
Gujarat Lions	DA Warner
Kings XI Punjab	DA Warner
Delhi Daredevils	V Kohli
Chennai Super Kings	V Kohli
Rajasthan Royals	SK Raina
Deccan Chargers	R Dravid
Kochi Tuskers Kerala	SR Tendulkar
Pune Warriors	CH Gayle
Delhi Capitals	AD Russell

Fig. 21. Top run scorers against a team

S. Best Partnerships

Partnerships are really very important for any team and even one big partnership can change the whole game. When a team has some big partnerships then it releases pressure for the other upcoming batsmen and allows them to play their game freely. Partnership also shows us the coordination and understanding between two batsmen who are batting together. So partnerships are really very important for every team. Let's look at some of the best partnerships in ipl. From the table we can see that

Batsman 1	Batsman 2	Partnership Runs
AB de Villiers	V Kohli	229
AB de Villiers	V Kohli	215
AC Gilchrist	SE Marsh	206
CH Gayle	V Kohli	204
DA Warner	NV Ojha	189
J Bairstow	DA Warner	186
RG Sharma	HH Gibbs	167
CH Gayle	TM Dilshan	167
AC Gilchrist	VVS Laxman	155
DA Warner	S Dhawan	139

Fig. 22. Best Partnerships

virat Kohli and ab de villiers have top 2 best partnership.

T. Highest wins by a team in each season

In this we are looking at teams which have dominated in each season by looking at the highest number of wins by a team in each season. Below is the table depicting that.

Season	Team	Number of Wins
2008	Rajasthan Royals	13
2009	Delhi Daredevils	10
2010	Mumbai Indians	11
2011	Chennai Super Kings	11
2012	Kolkata Knight Riders	12
2013	Mumbai Indians	13
2014	Kings XI Punjab	12
2015	Chennai Super Kings	10
2016	Sunrisers Hyderabad	11
2017	Mumbai Indians	12
2018	Chennai Super Kings	11
2019	Mumbai Indians	11

Fig. 23. Highest victories by a team for every season

From this table we can see that Rajasthan Royals dominated the first season by winning 13 matches in that season. Similarly we can see the dominating team for each year

U. Strike rate of different batsmen in each over

Here we are calculating the strike rate of top batsmen with good strike rates in each over using the ball by ball data. We are doing this so that we get to know which batsman is good in that particular over. Below is the table showing over wise strike rates.

over	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
batsman																				
AB de Villiers	166.7	122.5	86.1	118.8	131.5	134.8	113.6	115.6	122.7	134.1	134.5	139.2	152.0	154.8	170.5	196.1	189.2	251.5	271.8	253.6
AM Rahane	89.7	115.9	119.1	137.1	147.2	130.6	102.5	108.5	104.7	136.1	115.8	129.9	129.2	155.5	154.5	144.0	165.9	177.9	111.6	181.2
AT Rayudu	117.9	85.9	103.2	106.0	109.7	127.0	98.7	95.6	121.6	106.0	124.2	135.8	129.3	137.6	142.5	158.1	156.2	176.8	199.1	160.0
BB McCullum	115.0	109.9	143.4	145.2	141.2	162.1	136.3	133.8	123.9	132.8	123.5	151.2	151.4	160.3	148.4	109.1	221.4	123.5	223.5	287.5
CH Gayle	100.5	117.5	162.5	166.1	168.5	152.9	128.6	168.7	181.6	135.6	166.4	169.5	192.2	195.6	214.5	216.5	163.3	272.0	237.9	195.7
DA Warner	102.9	139.9	153.9	147.8	182.4	162.6	119.2	130.8	147.2	136.5	141.9	144.0	139.2	148.1	174.1	188.5	172.6	184.1	255.3	187.0
G Gambhir	98.5	123.7	126.3	133.3	151.0	131.0	115.6	107.7	111.9	138.5	126.8	115.9	149.6	148.3	145.7	156.1	134.3	153.5	131.2	0.0
JH Kallis	89.6	107.8	115.7	116.5	111.0	119.2	87.7	83.8	94.7	115.0	93.9	116.5	115.5	111.2	108.5	126.4	183.6	231.5	150.0	200.0
KA Pollard	300.0	0.0	100.0	44.4	171.4	95.7	78.0	97.7	84.3	107.2	127.5	106.5	127.0	152.9	158.8	140.1	176.3	158.3	186.1	212.9
KD Karthik	87.5	67.6	91.5	106.6	92.3	132.8	111.5	104.0	115.2	104.9	129.7	132.0	131.9	163.1	134.2	128.2	190.9	174.1	181.5	209.0
M Vijay	80.4	97.2	141.3	128.0	149.7	135.8	108.6	108.5	129.4	129.7	154.1	145.8	172.2	169.8	182.9	192.6	136.8	200.0	200.0	180.0
MK Pandey	94.1	101.1	95.7	120.8	131.0	115.0	106.4	110.9	116.2	106.0	133.8	117.2	133.1	128.6	162.4	138.6	152.6	136.6	174.6	209.1
MS Dhoni	0.0	0.0	0.0	0.0	93.3	107.1	81.0	106.9	95.9	92.8	103.4	100.0	107.5	110.7	128.4	138.9	161.7	183.9	199.3	238.2
PA Patel	87.2	122.3	152.7	127.2	145.1	147.9	115.0	106.0	109.5	121.9	143.8	150.0	137.3	104.7	121.6	126.9	175.0	153.8	170.0	85.7
R Dravid	87.3	110.2	114.8	116.7	117.1	135.6	96.2	110.6	102.3	112.3	139.0	121.3	131.1	144.4	130.0	136.7	200.0	135.0	143.8	215.0
RG Sharma	103.6	109.7	108.9	134.6	121.2	115.0	100.5	98.6	104.3	104.2	125.0	129.8	127.1	141.7	149.8	139.3	174.6	185.5	200.0	275.5
RV Uthappa	95.8	131.4	134.6	147.3	130.2	129.2	117.3	120.5	137.7	129.9	125.4	144.3	135.7	142.6	163.8	170.5	173.0	159.2	204.5	215.8
S Dhawan	98.3	122.6	131.4	136.1	132.9	141.9	112.7	112.9	134.6	124.1	123.5	129.3	141.8	170.9	142.7	139.5	162.0	153.8	180.4	166.7
SE Marsh	107.4	107.5	92.0	124.2	106.1	125.8	148.7	155.9	150.8	123.0	145.5	145.5	147.0	171.3	196.5	157.1	175.7	146.9	151.7	200.0

Fig. 24. Over wise strike rate of different players

From the table we gain various insights about the strike rate for example we can see how high strike rate ab de villiers has got in death overs as compared to other batsman

V. Batsman Comparator

We can compare two batsmen on various parameters like runs, strike rate, No. of 4' s and No. of 6's. We can input the

two parameters and then the two batsmen and we get a scatter plot as the output. Below is the example of runs and strike rate comparison of Virat Kohli and Rohit Sharma. From the
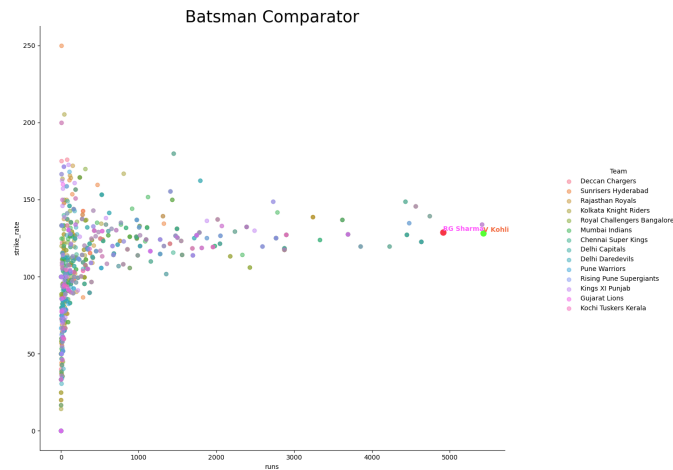

Fig. 25. Scatter plot for comparing Virat Kohli and Rohit Sharma on the basis of striuke rate and runs

plot we can see that Virat Kohli has scored more runs than Rohit Sharma with almost equal strike rate
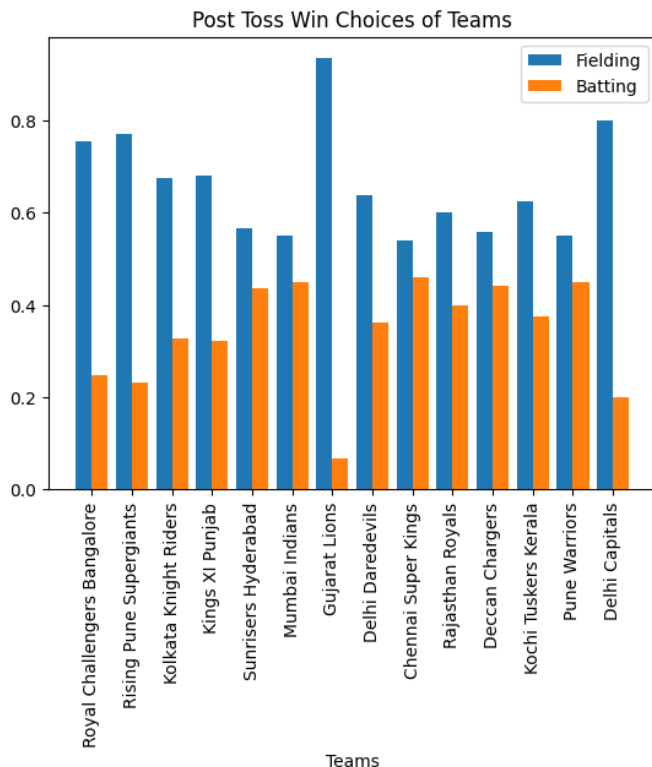
W. Post toss win choices of team

Fig. 26. Graph showing the percentage times teams choose batting or fielding after winning the toss.

Just winning the toss isn't just enough, but to take the right decision after winning is more important. Most teams take a lot of effort in deciding what they will choose if they win the toss, which sometimes also makes a difference in the match.

The graph shows the different trends followed by differnt teams after winning the toss. We can clearly see that almost all the teams prefer fielding over batting. A certain for this can also be that while you are fielding you have a bit more control of the match, like you get the target and you know how hard you need to go to achieve it.

V. DISCUSSION

This project provides useful insights from the IPL dataset about what are the best performing teams and players. Sponsors can focus on which cities host the IPL matches most to analyze the audience in those areas specifically and make their plans accordingly. Best performing players of IPL can be listed with the most MoM awards analysis. Toss decisions have more or less no influence on winning.

Furthermore using the batsman comparator we can compare between two batsman on different given parameters and check which one would be a better fit for a given situation. While auctions teams can use the different player analysis done to build their perfect squad including high strike rate, players with good average, bowlers who take high wickets, bowlers who bowl less extras, the fielder who takes the most no. of catches etc. Much more deeper can be done and used for further analysis.

VI. ACKNOWLEDGEMENT

We would like to thank our fellow wingmates who came up with interesting EDA ideas. We would also like to thank the contributor at kaggle who provided such a great dataset on which we did a ton of analysis.

VII. REFERENCES

1) https://pandas.pydata.org/pandas-docs/stable/user$_guide/visualization.html$
2) https://www.geeksforgeeks.org/python-lambda-anonymous-functions-filter-map-reduce/
3) https://seaborn.pydata.org/generated/seaborn.lineplot.html
4) https://www.kaggle.com/datasets/nowke9/ipldata?select=matches.csv
5) https://www.iplt20.com/stats