

# Analysis Of IPL Dataset

By -

Revanth Babu P N, PES1201800042

Navyadhara G, PES1201800230

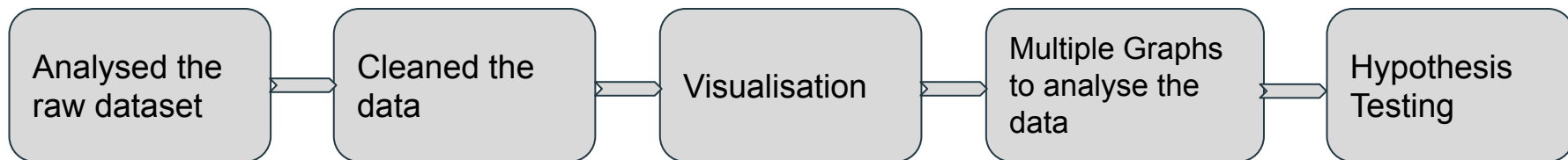
# What is our dataset?

This dataset contains the following :

- ❑ All Indian Premier League Cricket matches between 2008 and 2017.
- ❑ This is the ball by ball data of all the IPL cricket matches till season 9.
- ❑ The dataset contains 2 files: deliveries.csv and matches.csv. Matches.csv contains details related to the match such as season, location, contesting teams, umpires, winner, etc.
- ❑ deliveries.csv is the ball-by-ball data of all the IPL matches including data of the batting team, batsman, bowler, non-striker, runs scored, dismissal kind, etc.

Cleaning :

Removed umpire3 column as all the values were NaN. Replaced Rising Pune Supergiant with Rising Pune Supergiants as it's the same team



# ANALYSIS

season		winner
0	2008	Rajasthan Royals
1	2009	Deccan Chargers
2	2010	Chennai Super Kings
3	2011	Chennai Super Kings
4	2012	Kolkata Knight Riders
5	2013	Mumbai Indians
6	2014	Kolkata Knight Riders
7	2015	Mumbai Indians
8	2016	Sunrisers Hyderabad
9	2017	Mumbai Indians

	bowler	economy
0	SP Narine	6.395706
1	R Ashwin	6.493639
2	DW Steyn	6.615599
3	A Kumble	6.680982
4	M Muralitharan	6.711027

Most Economical bowlers(min 150 overs)

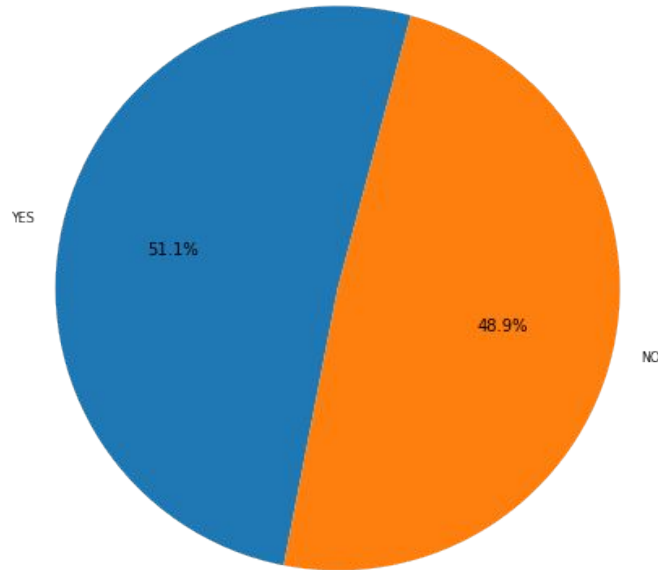
- Sunil Narine has the best economy

	batsman	bowler	No_of_Dismissals
0	V Kohli	A Nehra	6
1	CH Gayle	R Ashwin	4
2	RG Sharma	R Vinay Kumar	6
3	AB de Villiers	KH Pandya	4
4	MS Dhoni	Z Khan	7

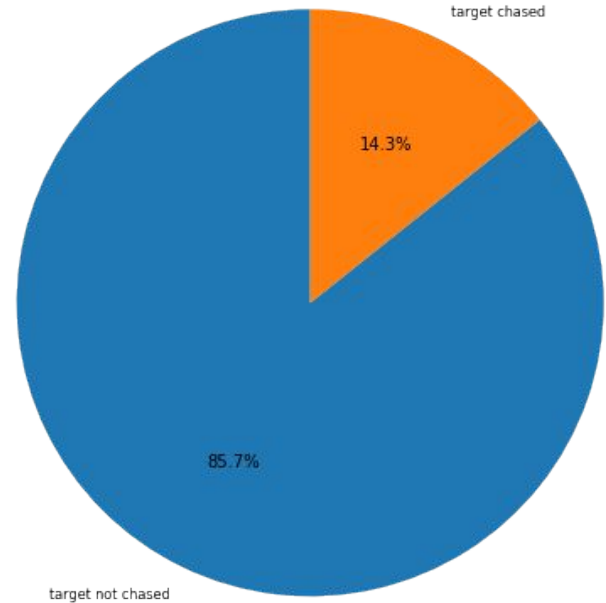
- Mumbai Indians have won the most number of times

# VISUALISATION

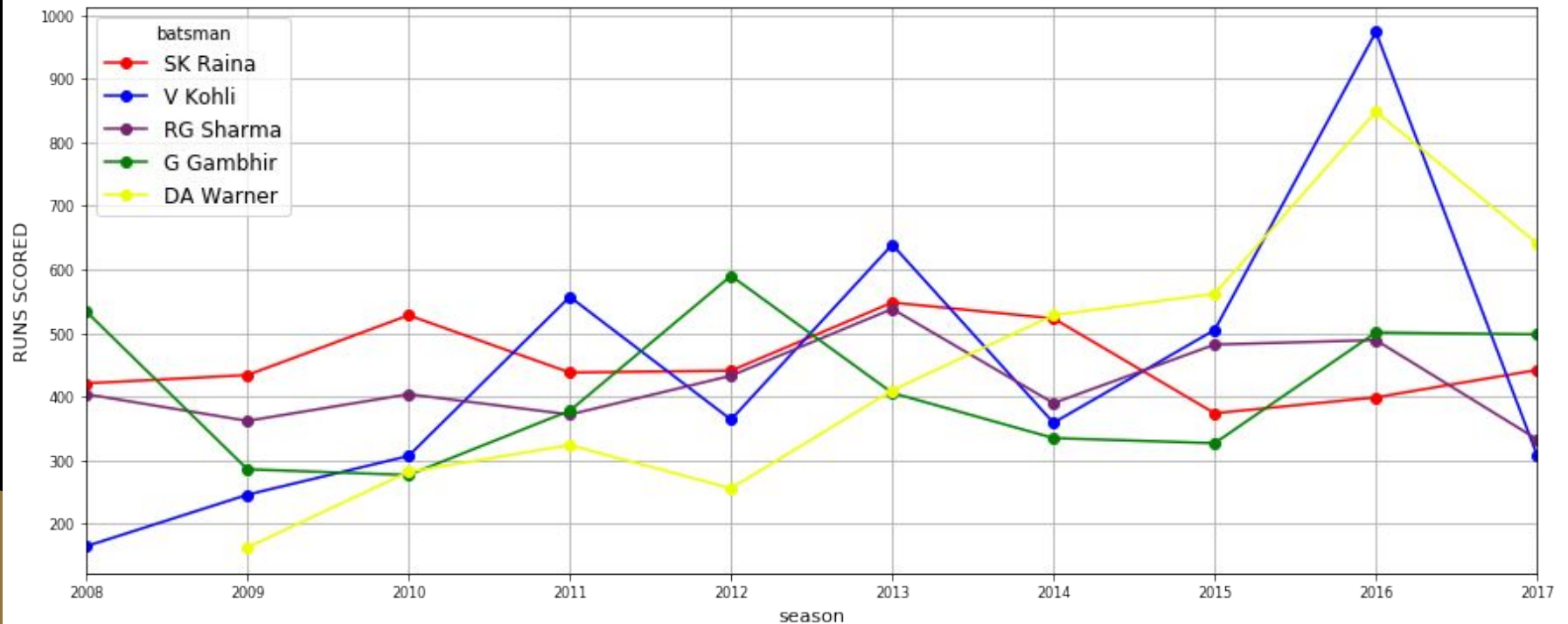
IS TOSS WINNER THE MATCH WINNER ?



CHANCES OF CHASING 200+ TARGET

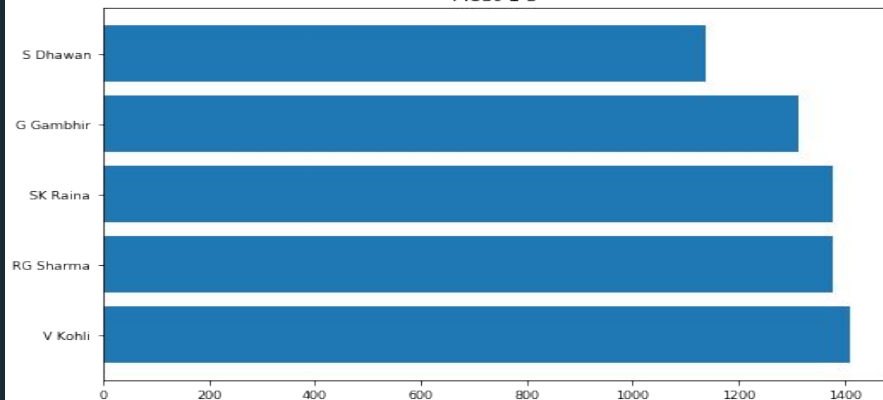


1. It is not true that toss winners win the match
2. There is high chance that the team scoring 200+ wins the match

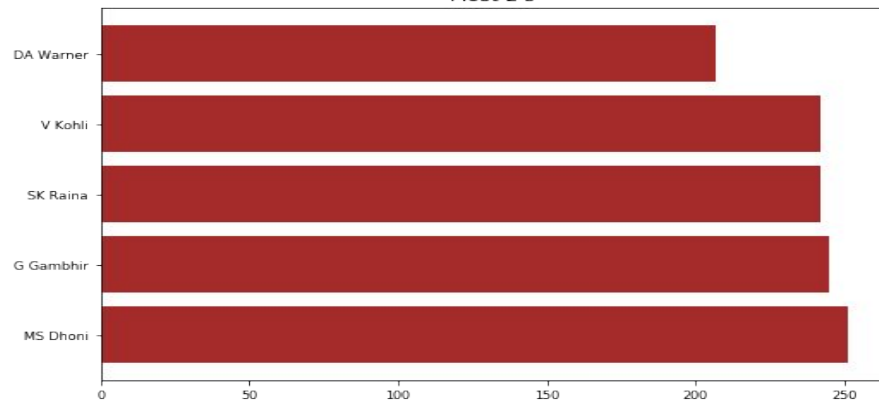


1. David Warner's form looks to be improving season by season.
2. Virat Kohli holds the record of most runs in a season 973. There has been a sharp decline in Kohli's Runs from 2016 to 2017.
3. Raina has consistently scored 300+ runs in every season

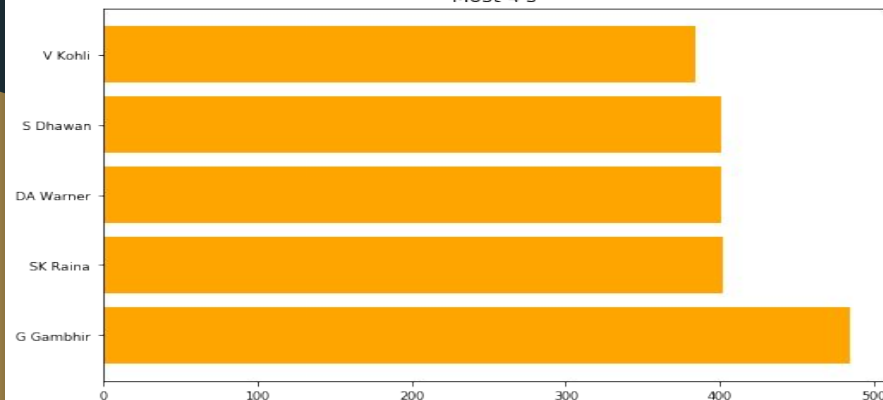
Most 1's



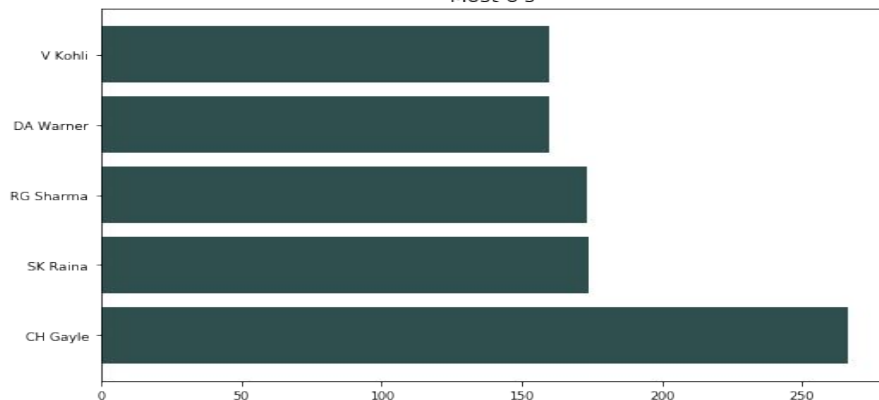
Most 2's



Most 4's

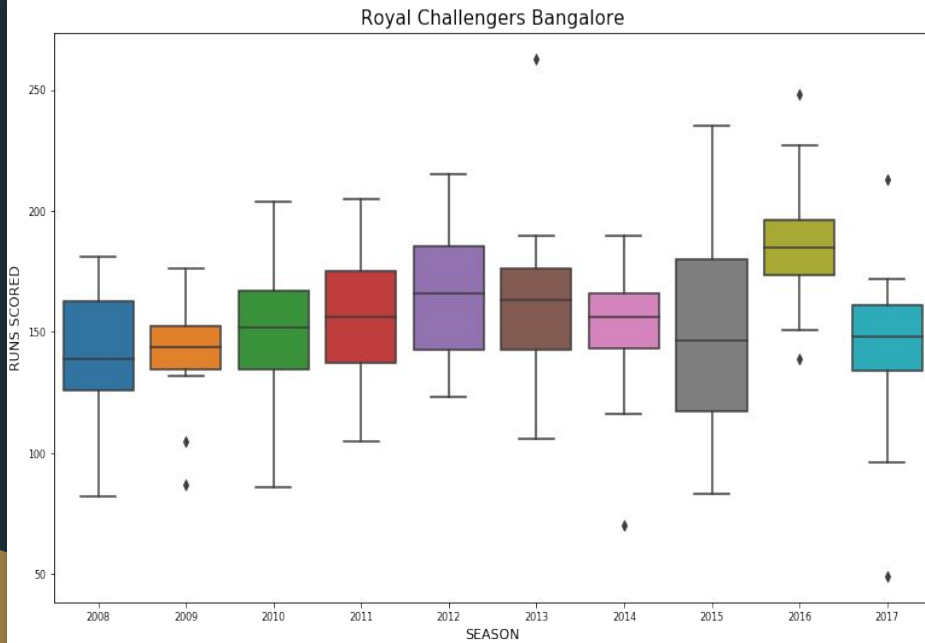


Most 6's



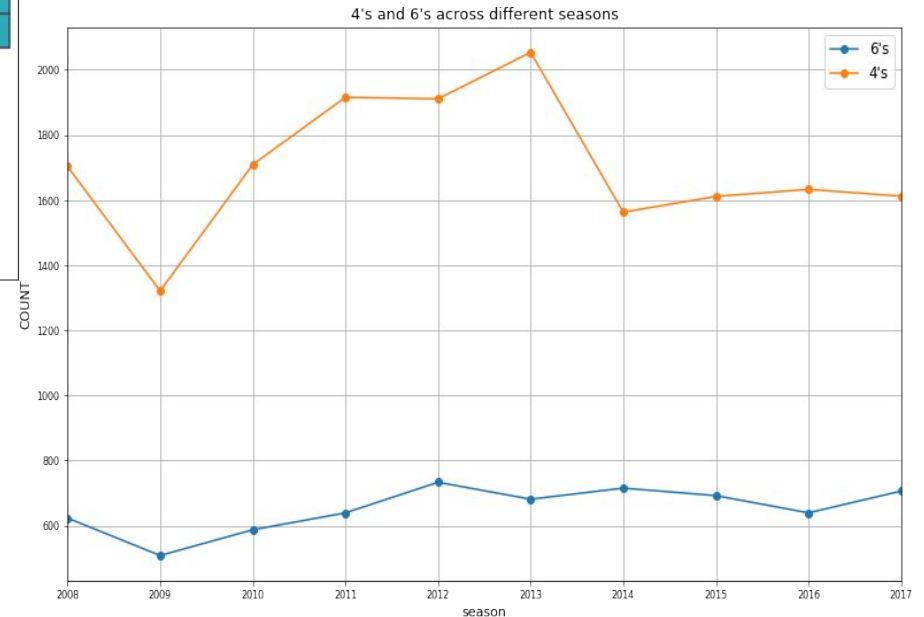
Virat has most number of 1s and is among the top 5 batsman with most 1s, 2s, 4s and 6s.

Gayle has most number of 6s. Gambhir has most number of 4s. Dhoni has most number of 2s

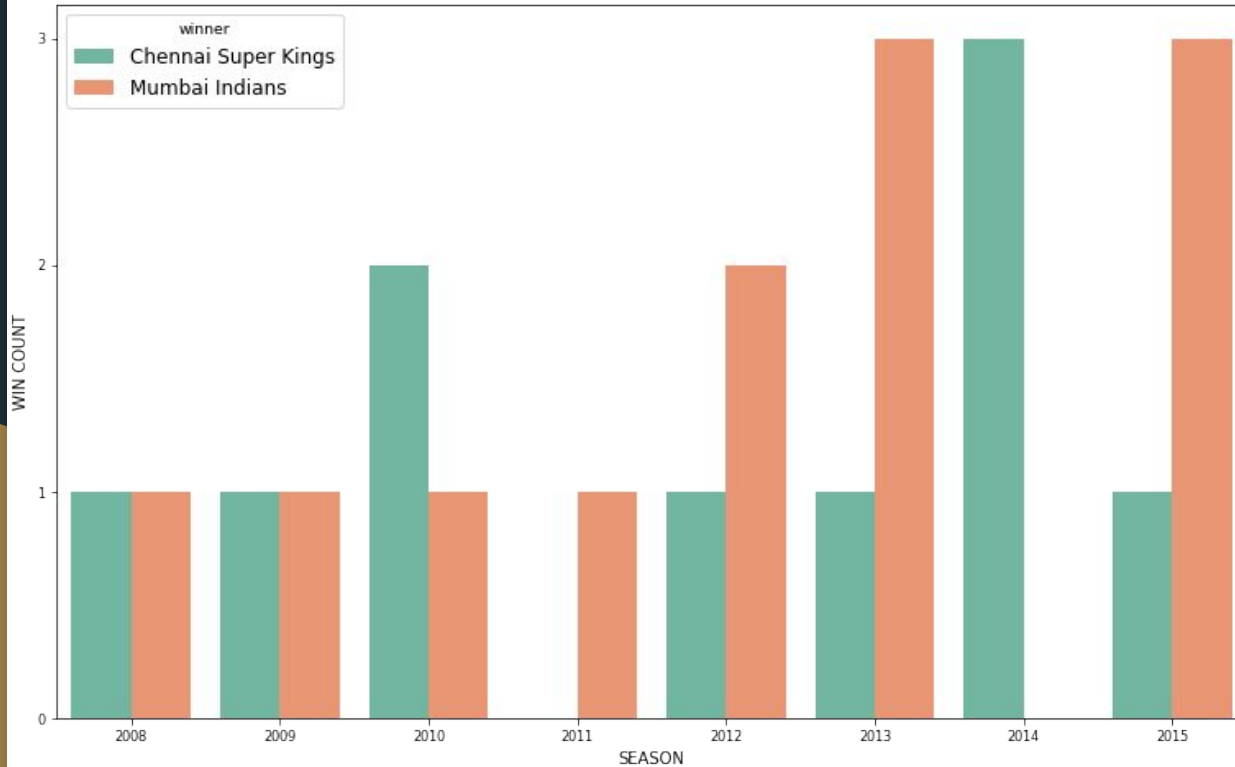


The number of boundaries was least in the year 2009

Average runs scored per match by RCB was highest in 2016.  
RCB has highest team total 263 runs (in 2013) and the lowest team total 49 runs



# CSK vs MI



MOST WINS: MI - 12 wins

2008 - MI:1 CSK:1

2009 - MI:1 CSK:1

2010 - MI:1 CSK:2

2011 - MI:1 CSK:0

2012 - MI:2 CSK:1

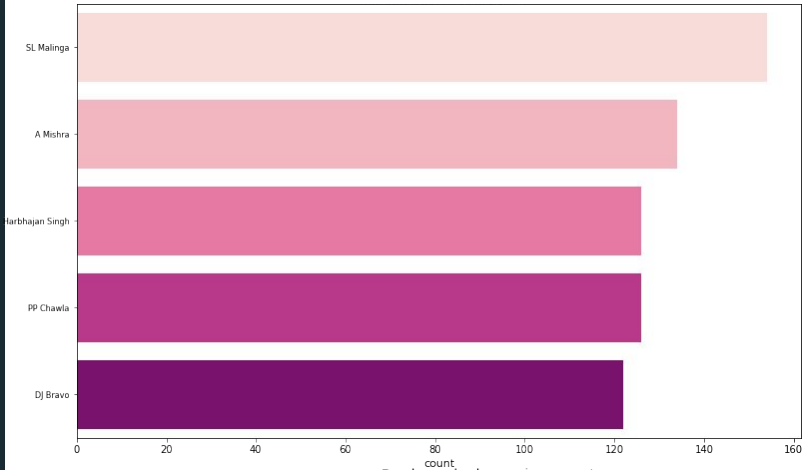
2013 - MI:3 CSK:1

2014 - MI:0 CSK:3

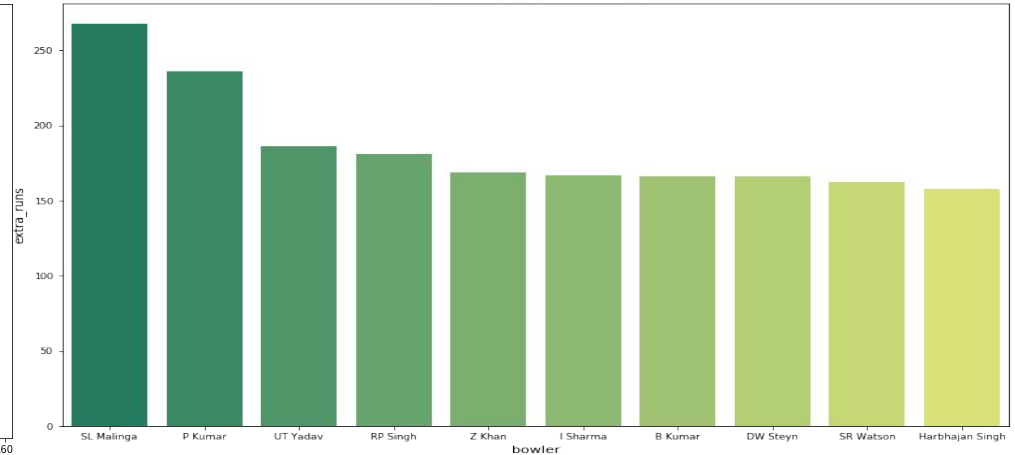
2015 - MI:3 CSK:1



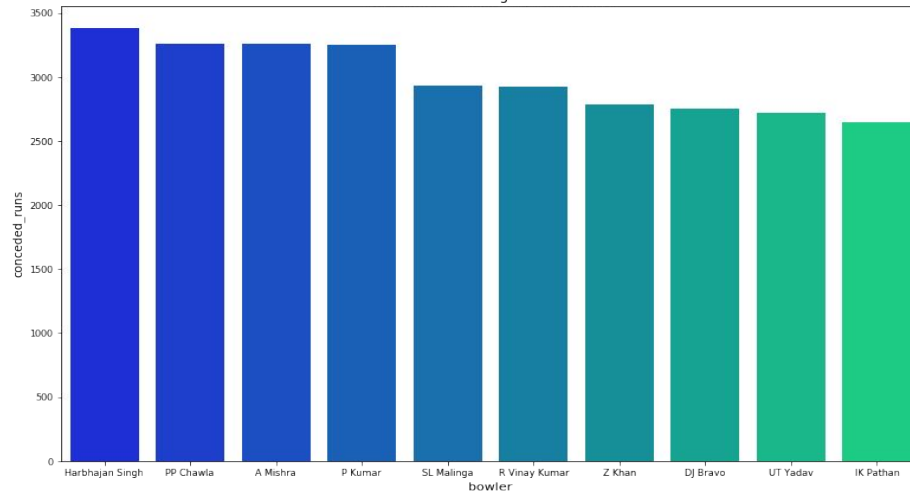
Bowlers with most Wickets



Bowlers with more extras

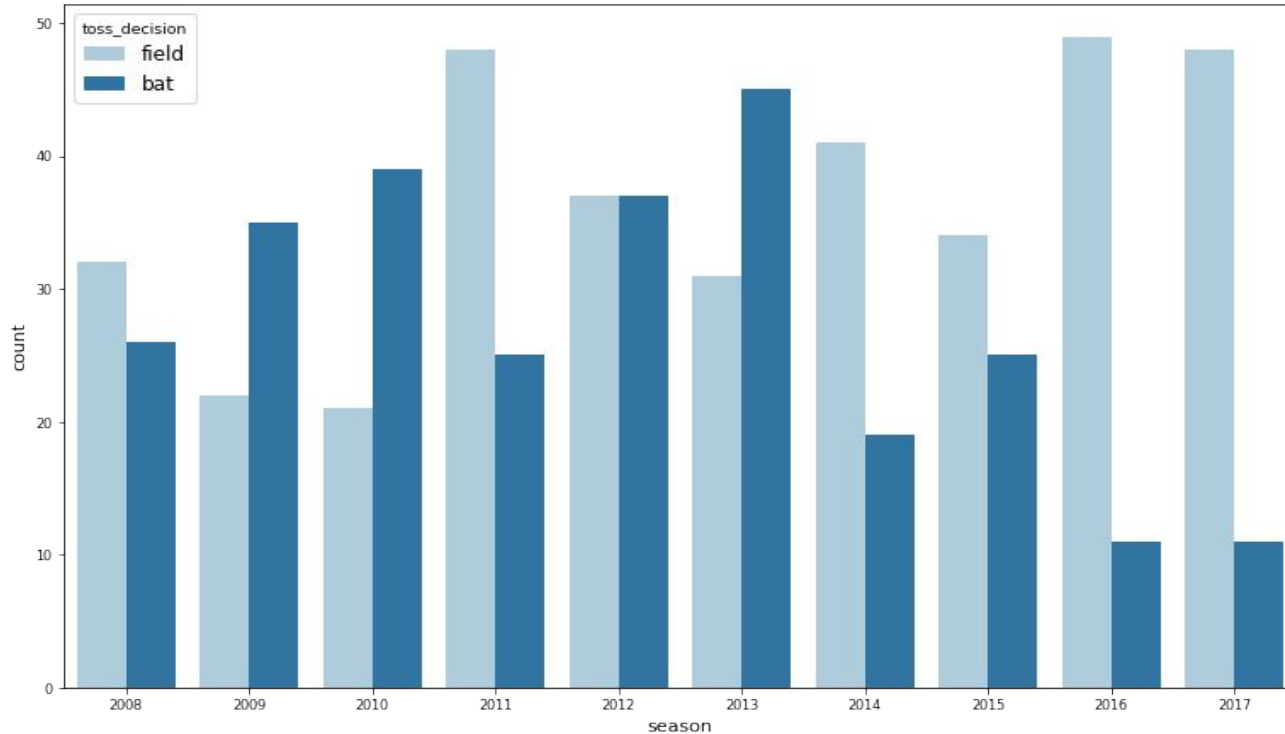


Bowlers who have given most runs



## Bowlers Analysis

1. Most Wickets - Malinga
2. Most Extras - Malinga
3. Bowler who has given most runs - Harbhajan Singh
4. Most Economical bowler(min 150 overs) - Sunil Narine



In the initial seasons of IPL, there is no much difference in number of times toss winners chose to bat and field. In seasons 2009, 2010, 2013 batting first was the preferred choice. Whereas in the last few seasons it is clearly seen that the trend is to bowl first.

# HYPOTHESIS TESTING

We have also framed hypothesis and tested using the data.

We have tested the case of average scores in 2009 vs that of in 2016.

x: denotes year 2016 , y: denotes year 2009

$H_0 : u_x - u_y \leq 0$  ,  $H_a : u_x - u_y > 0$

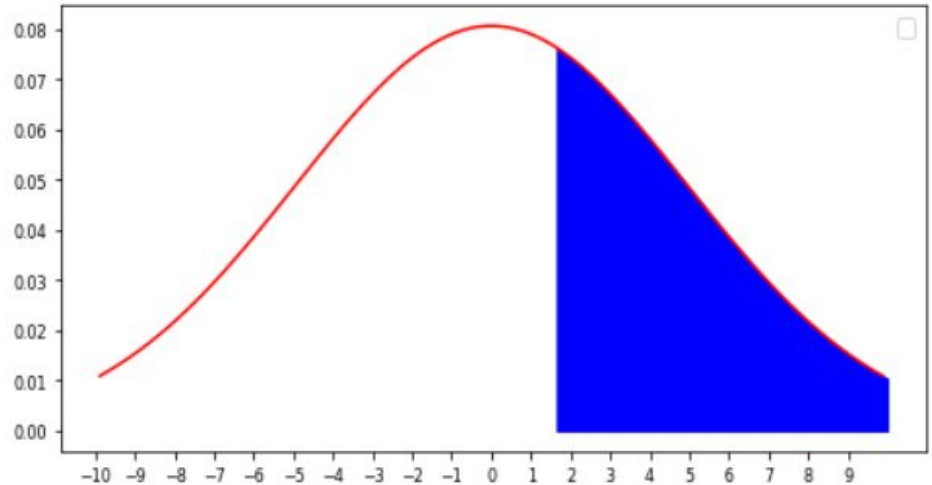
Since  $p\_value = 0.0063 < 0.05$

We reject the null hypothesis and conclude average scores in 2016 was greater than that in 2009.

IPL 2009 was hosted by South Africa.

Thus we can conclude the pitches in South Africa are less batsmen supportive. This is also evident from the line graph, which indicated that the number of boundaries was least in 2009

$u_x = 162.6$   $u_y = 150.26315789473685$   
 $sd_x = 29.35$   $sd_y = 23.96$   
 $\sigma = 4.94$   $z = 2.5$   
 $p = 0.0063$



# CONCLUSIONS

1. It is not true that toss winners win the match
2. There is a high chance that the team scoring 200+ wins the match
3. David Warner's form looks to be improving season by season.
4. Virat Kohli holds the record of most runs in a season 973. There has been a sharp decline in Kohli's Runs from 2016 to 2017.
5. Raina has consistently scored 300+ runs in every season
6. Virat has most number of 1s and is among the top 5 batsman with most 1s, 2s, 4s and 6s.
7. Gayle has most number of 6s. Gambhir has most number of 4s. Dhoni has most number of 2s.
8. Suresh Raina has scored the most runs in IPL followed by Virat Kohli and Rohit sharma
9. RCB has scored most number of sixes and MI have scored most number of 4s
10. Average runs scored per match by RCB was highest in 2016. RCB has highest team total 263 runs.
11. The most common dismissal type in IPL is caught followed by bowled. There are very few instances of hit wicket as well.
12. Chris Gayle has won the most number of man of the match awards
13. In the initial seasons of IPL, there is no much difference in number of times toss winners chose to bat and field. In seasons 2009, 2010, 2013 batting first was the preferred choice. Whereas in the last few seasons it is clearly seen that the trend is to bowl first.

By hypothesis Testing,

We have come to a conclusion that the pitches in South Africa are less batsmen supportive than those of India. Alongside this we have arrived at multiple other conclusions about the performances of players.