Sport Analysis: IPL 2008-2017 Matches

Introduction:

The scope of professional sports has changed over the years. I remember watching every minute of the 2007 Cricket World Cup and spending every waking minute tracking statistics, like the total runs scored, highest run-scorer, highest run rate, and so on. It was fairly rudimentary stuff but enough to keep me glued to the screen. How times have changed since then!

Sports analytics is quickly becoming mainstream. Media outlets and leading sports websites regularly curate statistics, produce deep technical insights, and add a whole new level of analysis we haven't seen before.

We can now answer questions like what is the probability of the team winning a match while batting first and second?

And so on. Honestly, the sky is the limit when it comes to sports analytics use cases. I'm a sports lover and I'm always looking out for applications where I can apply my analytics knowledge to improve the team strategy as well as fan experience.

What is Sports Analytics?

Sports Analytics is all about analyzing and extracting useful insights from sports data.

I would broadly dive Sports Analytics into 2 categories:

1. Descriptive Sports Analytics
2. Predictive Sports Analytics

Descriptive Sports Analytics

Descriptive Sports Analytics is about **summarizing the sports data** in the form of numbers. In other words, to come up with important statistics. This might sound like a simple concept but it's a very powerful one.

The thought behind descriptive sports analytics plays a crucial role in team tactics.

In cricket we can analyze how frequently a batsman gets out to a specific bowler. This number will decide the bowling strategy of a team.

Another interesting use case in cricket is to **analyze the team's probability of winning a match while batting first as well as second.** This influences the captain who wins the toss and has to make a decision – bat or bowl first.

Predictive Sports Analytics

Predictive Sports Analytics is about **making predictions using sports data**. One such use case in cricket is to **predict the number of runs a batsman scores** against

an opponent in a particular match. This would help the team management and captain select the best team for every match.

In a sport like football, predictive sports analytics helps to understand the chances of scoring a goal from any location on the pitch.

You can think of similar use cases for your favorite sport and let me know in the comments section below the article.

About problem question

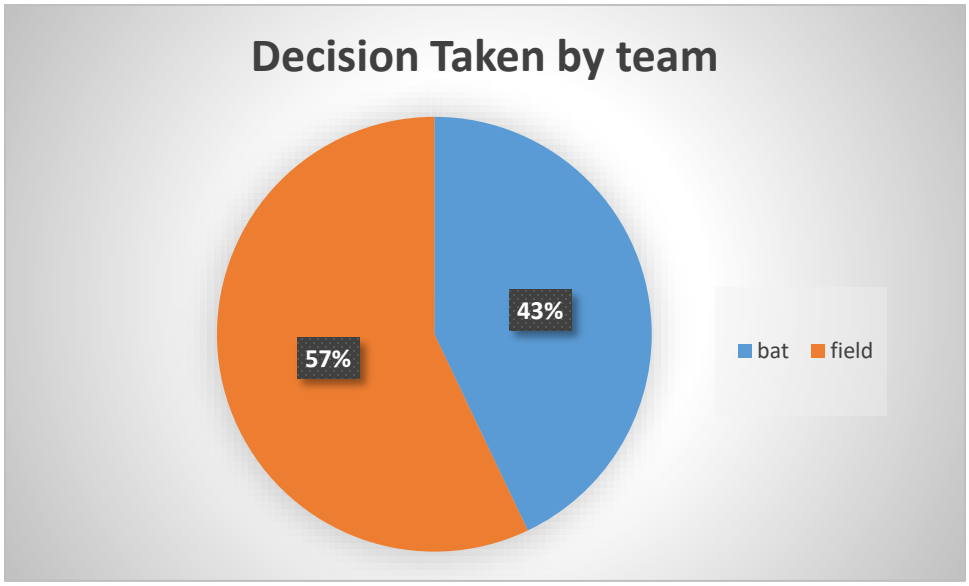
- What are the most decisions taken by teams after winning the toss?
- Which player has the most player of the match award?
- Which team had a good winning % in Home and Away matches?
- Which pair of umpire had used most of the time?
- Which player had a most player of the match awards?
- How many matches played in different grounds?
- How many matches win by chasing and defending runs?

About the Dataset for Sports Analytics

I have collected the data of IPL from 2008 to 2017 matches. Download a sample dataset from here then I make it clear and ready to use. It's time to analyze the matches and find some appealing insights. Let's do it!

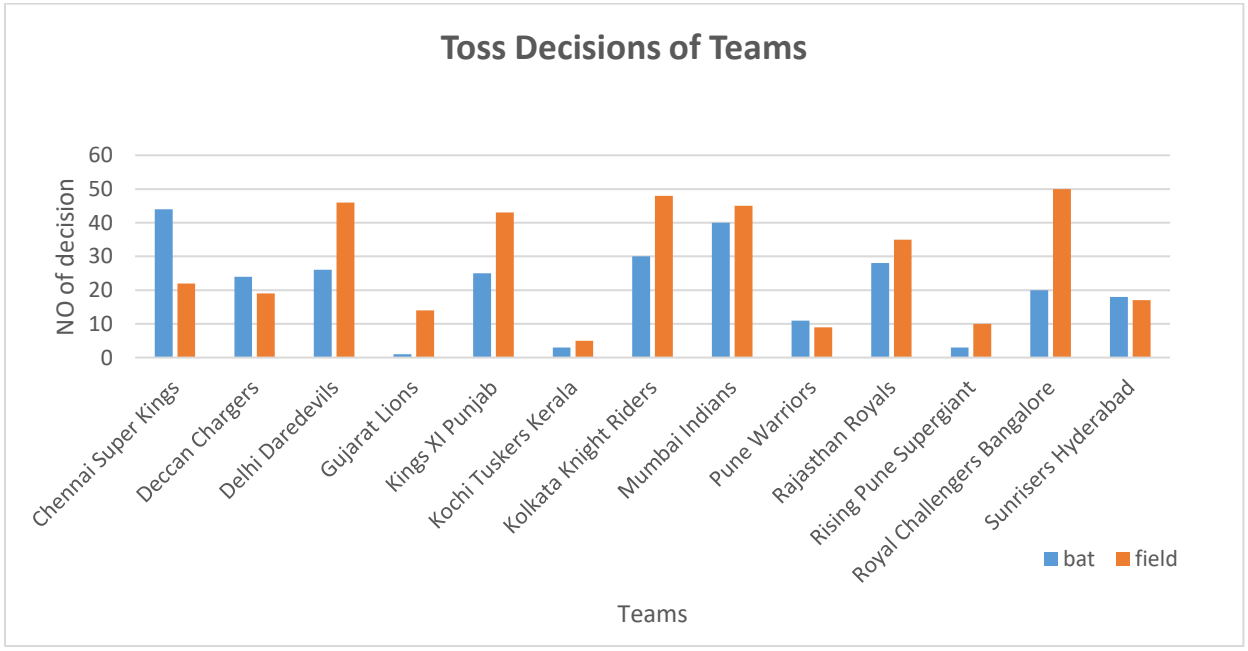
Since the data is not large, I have made analysis in excel file only.

1) Decisions taken by team



It is clear here that whenever the teams have won the toss, they have decided to field more than bat first.

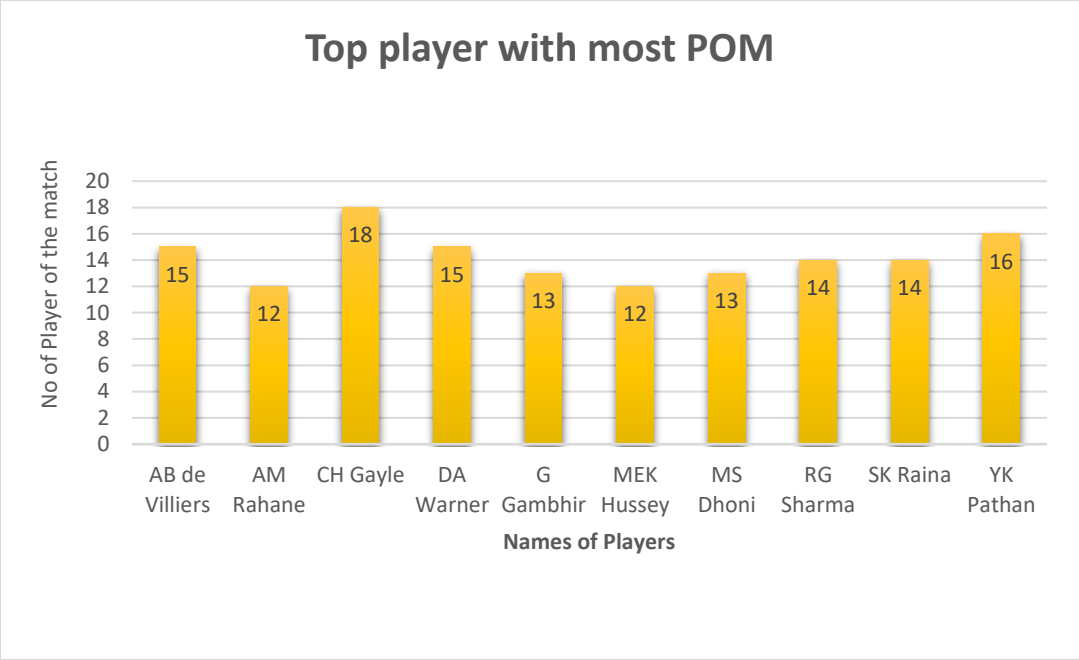
2) Toss decision by each team



It is clearly visible here that **Royal Challengers Bangalore** has taken the decision of field most of the time (50), and **Chennai Super Kings** has taken the decision of Bat most of the time (44).

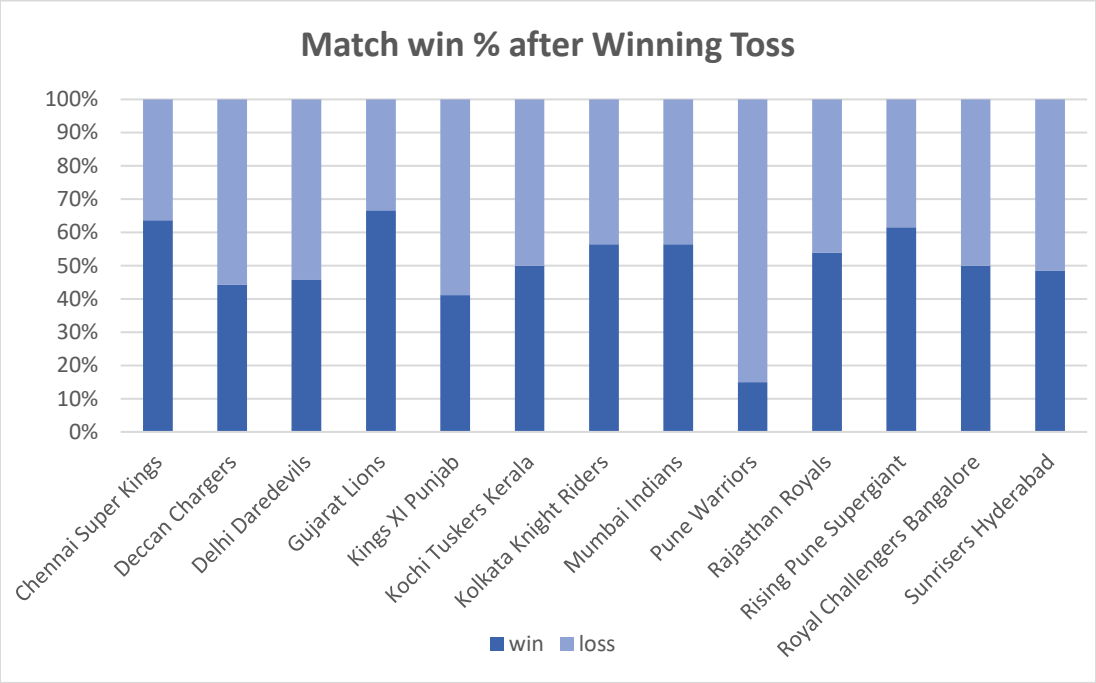
Here are some teams like **Gujarat Lions, Rising Pune Supergiant, Pune Warriors** and **Kochi Tuskers Kerala** played for one or two years only, that's why all of them has less data.

3) Tops players with most player of the match (POM)



Here is the list of top 10 player with most player of the match awards. It is clearly visible that **CH Gayle** has 18 player of the match which is most among all other players followed by **YK Pathan** and **AB de Villiers.**

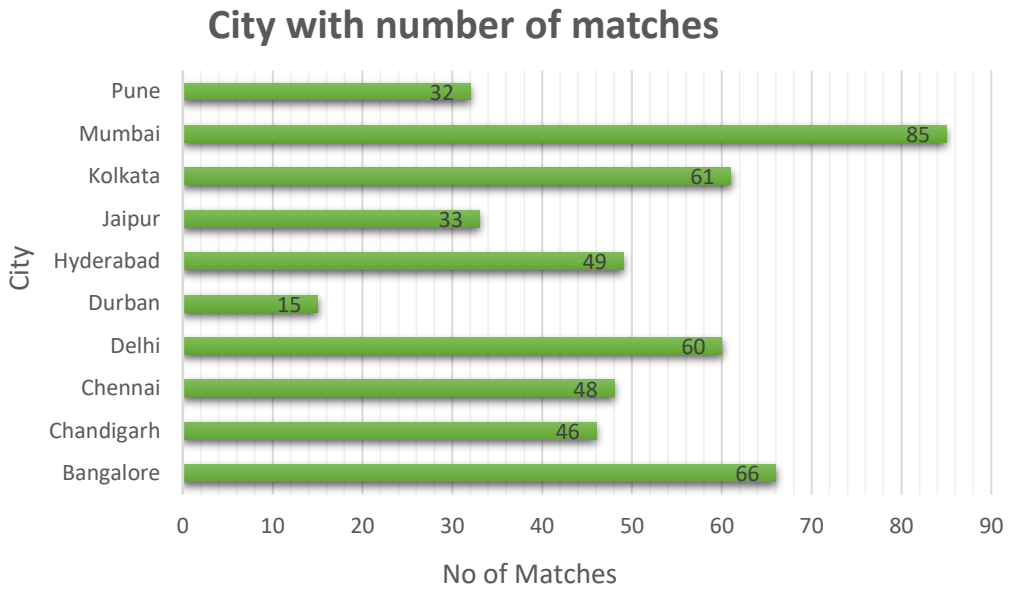
4) Match win percentage of team after winning toss



It is shown in the above graph that **Gujarat Lions** (67%) has highest win % after winning toss but this team played for two years that's why it is high, it has played less than 25 matches. if we see that the team which has played more than 25 matches, among all of them **Chennai Super Kings** (64%) has highest win % after winning toss.

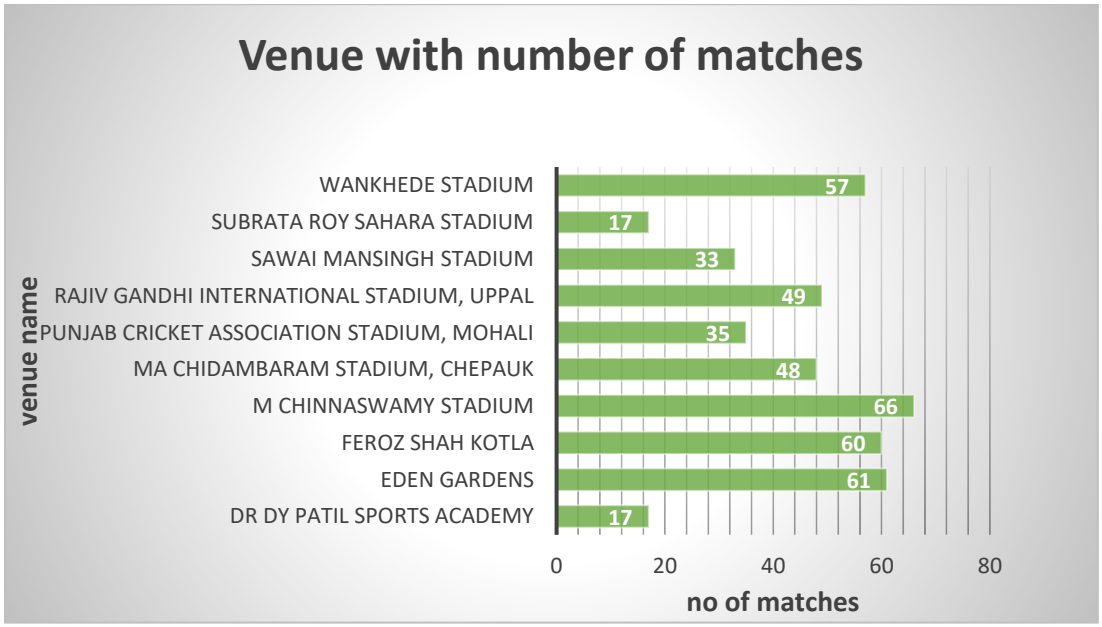
In other side we can see that lowest win % after winning toss is 15% which has **Pune Warriors,** but this also played less than 25 matches. The team which played more than 25 matches and has lowest win % after winning toss is **Kings XI Punjab** (41%).

5) City with Highest number of matches



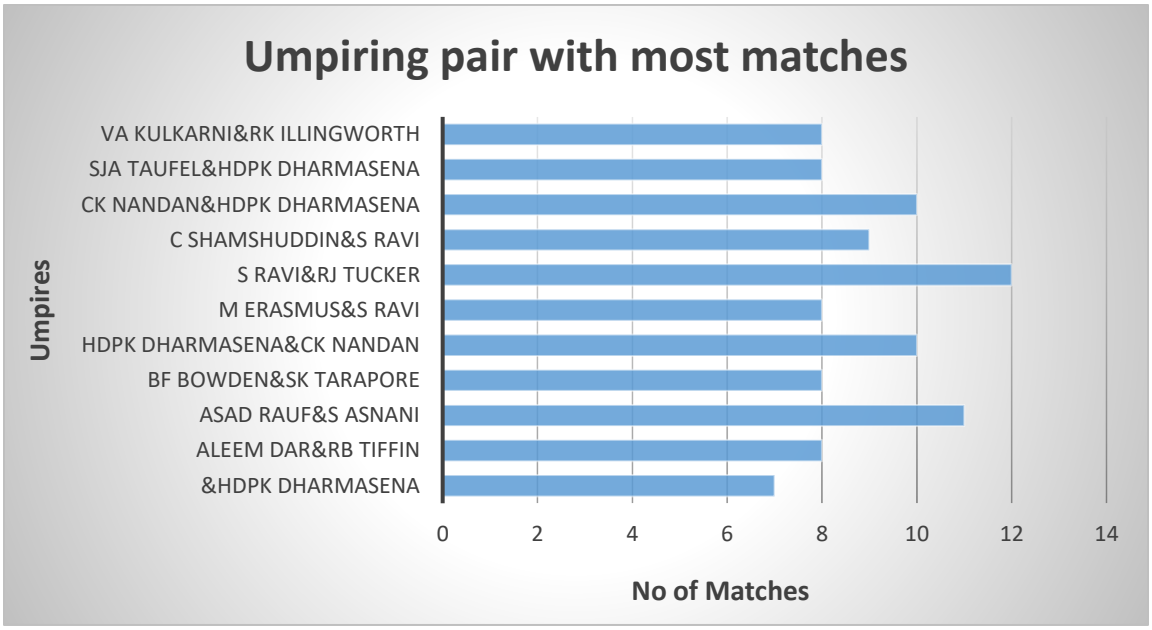
From above graph it is visible that city with most number of matches is **Mumbai** with **85** matches. Mumbai city has three venue Wankhede Stadium, Dr DY Patil Sports Academy and Brabourne Stadium

6) Venue with highest number of matches.



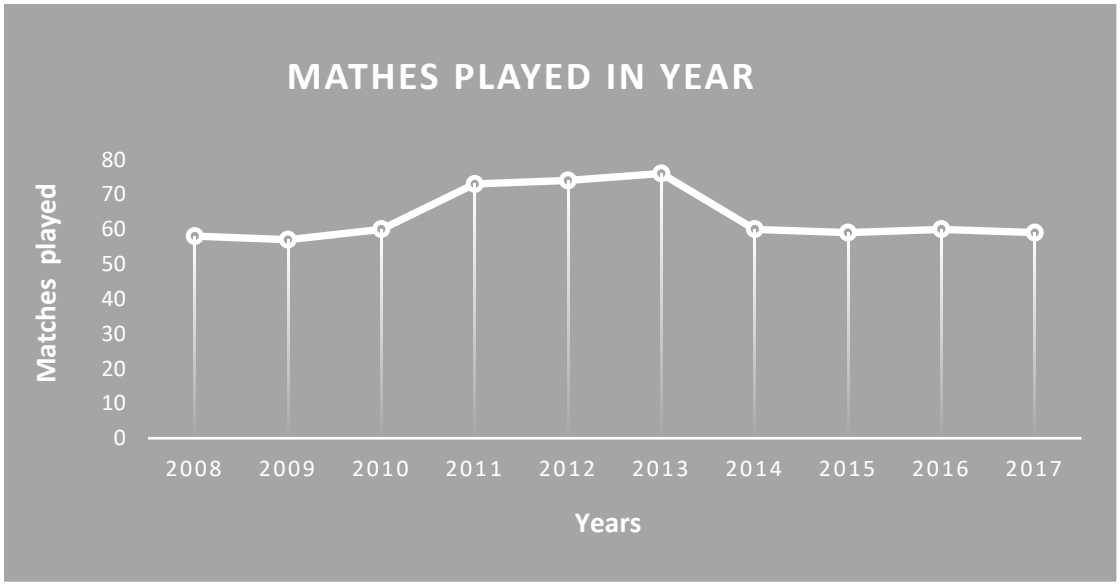
Most matches played in **M Chinnaswamy Stadium** (66 matches) which is located in Bangalore city. Other venues like **Eden Garden** (Kolkata), **Feroz Shah Kotla** (Delhi) and **Wankhede Stadium** (Mumbai) has average of 60 matches played.

7) Umpire pair umpiring in most matches



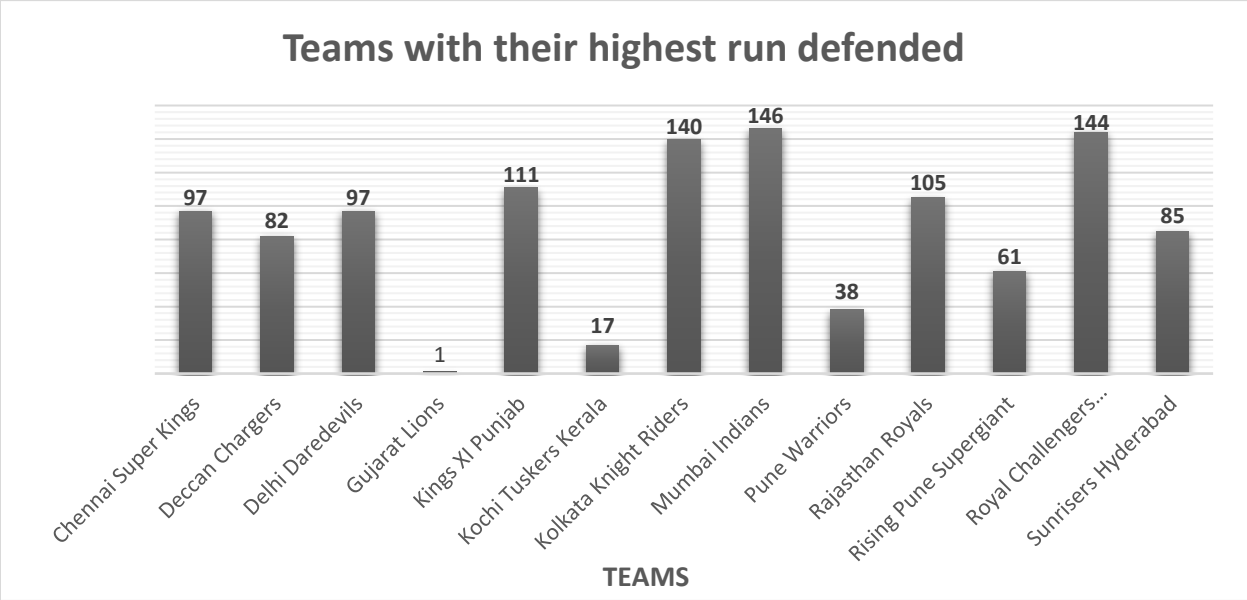
From above graph we can say that **RJ Tucker & S Ravi** pair have umpiring in most matches (12) and there are many pair have umpiring in 8 matches.

8) Number of Matches played in year



In above graph it's seen that graph increase then after some year it decrease because of number of team increase and decrease in year. 8 teams in the year **2008 to 2010** and 10 teams in the year **2011** and 9 teams in the year **2012-2013** and 8 teams in the year **2014 to 2017.** Most number of matches played in the year **2013** (76).

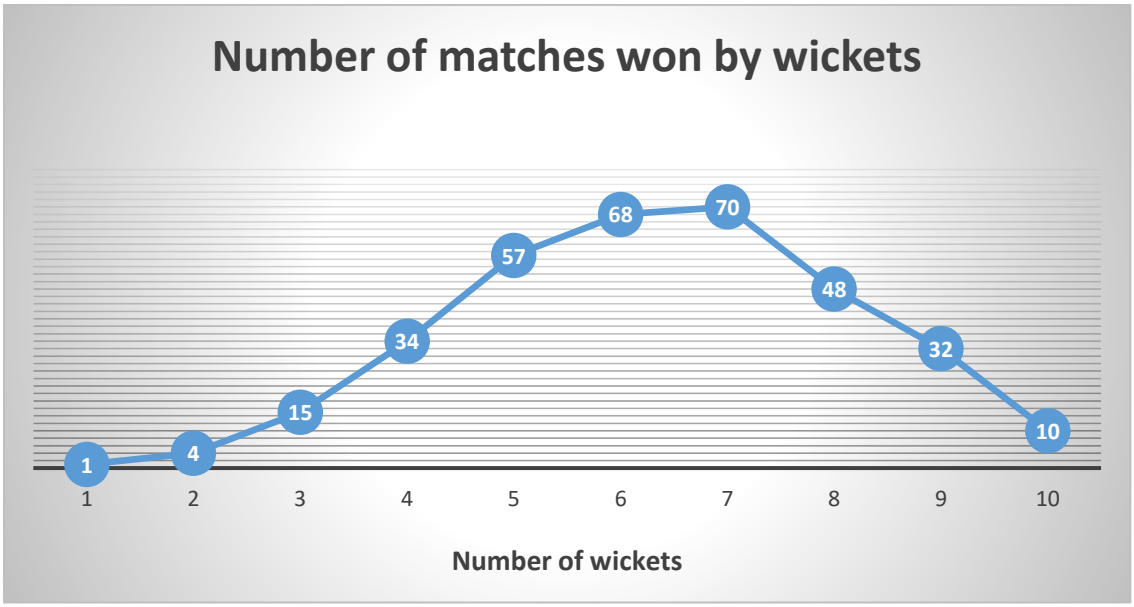
9) Each Team with Their Highest Run Defended



49	52	54	56	64	68	69	70	73	79	81	83	84	88	89	90	91	94	95	96	99	100

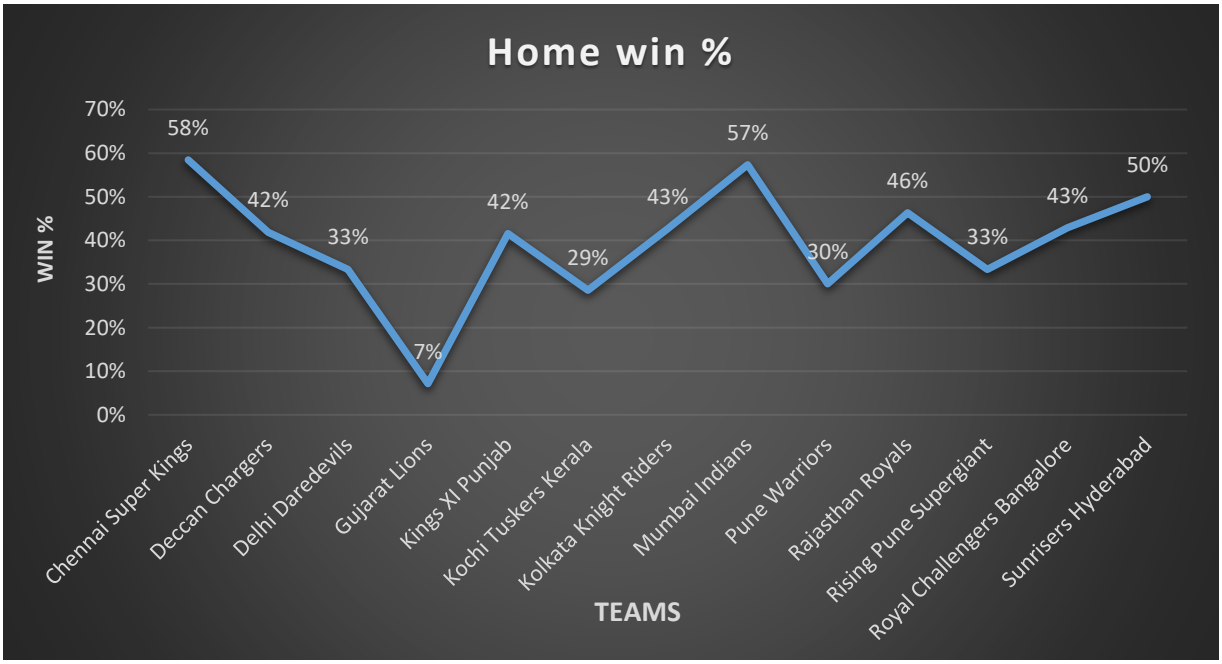
This are the runs numbers by which never a single team won a game in the range of 1 to 100 runs. There are 7 time teams won by more than 100 runs, in which **Royal Challengers Bangalore** is only team who won 3 game with more than 100 runs. The team that won the match with the most runs is **Mumbai Indians** against Delhi Daredevils with **146 runs**.

10) Number of Matches won by number of wickets

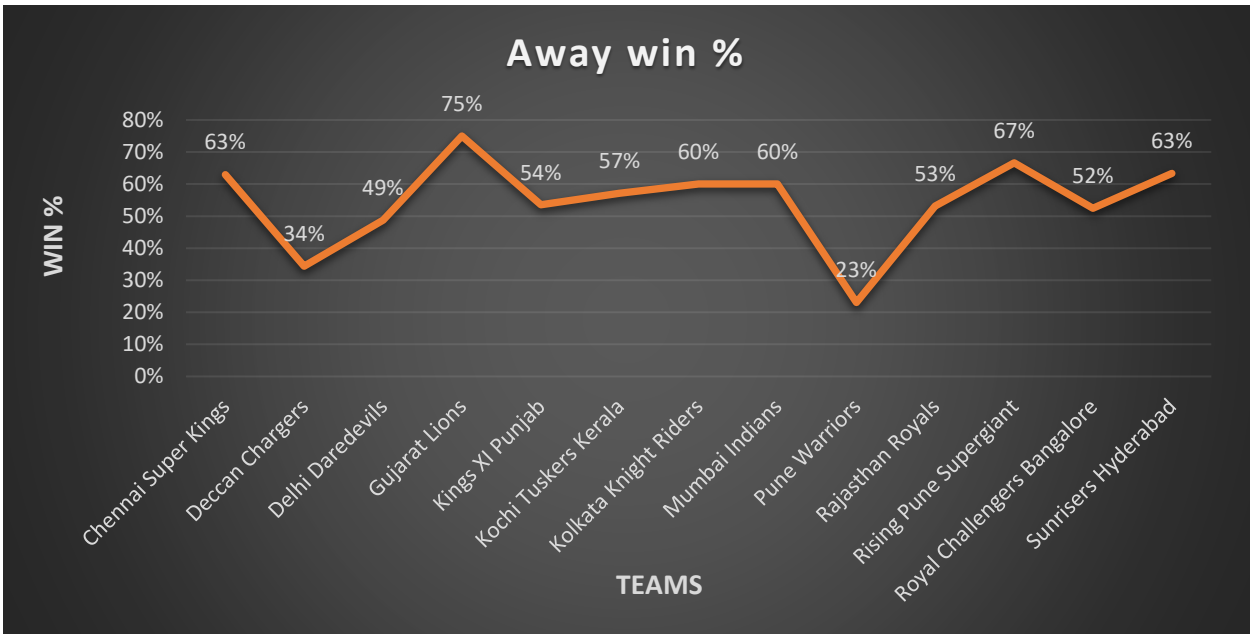


It is show that most number of matches won by **7** wickets and there are 10 matches only which won without giving single wicket. Average of matches won by wicket is **6-7**.

11) Each Teams Home and Away Win %



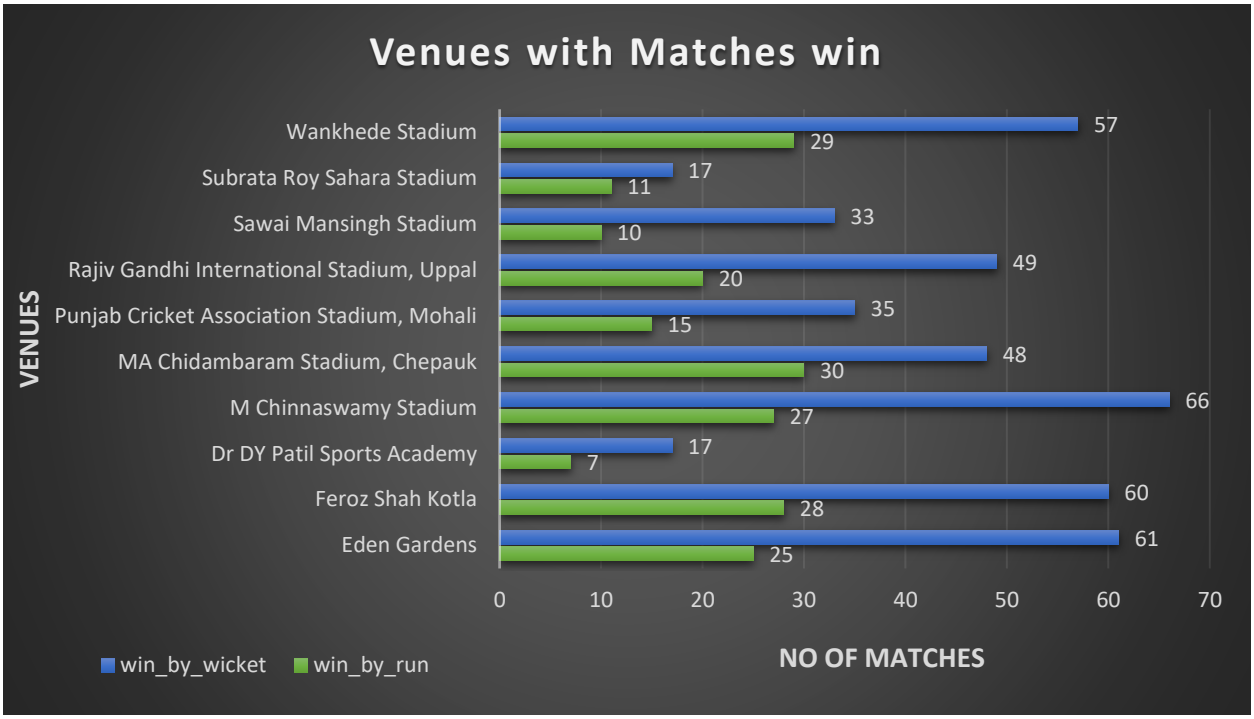
It is shown that highest home win % of **Chennai Super Kings** and **Mumbai Indians** is almost same **58%** and **57%** respectively. **Gujarat Lions** has the lowest home win % with 7% only.



It is shown that **Gujarat Lions** has the highest away win % with **75%** but this team played less than 25 matches. So, we consider the team with highest away win % with played more than 25 matches is **Chennai Super Kings** with **63%** win %.

If we conclude with both Home and Away win % than Chennai Super Kings is the best team with 58%(home) and 63%(away) win.

12) Venus with number of matches win by Runs and Wickets



In above graph it is shown that in **M Chinnaswamy Stadium** most matches are won by chasing a target with number of **66** matches. And In **MA Chidambaram Stadium, Chepauk** most number of matches won by defending the target with a number of **30** matches. We can say that it is easy to chase a run in stadiums like **M. Chinnaswamy Stadium, Feroz Shah Kotla, Eden Garden and Wankhede Stadium** by difference in win_by_wicket and win_by_run ratio.

Descriptive Analytics tells us what happened in the past and Predictive Analytics predicts what is most likely to happen in the future. So, we did a Descriptive Analysis of Matches played in IPL 2008 to 2017.