DATA VISUALIZATION

Indian Premier League

Prepared by:

RANU PARATE
X17161452
NATIONAL COLLEGE OF IRELAND

INTRODUCTION

Cricket is one of the famous sports played all over the world. It is the bat and ball game played between two teams, each team having eleven players at the center on the field. Cricket is considered as a multi-faceted game. It has multiple formats that depend on the time available for the game, desired level of formality and the playing standard (Wikipedia, 2019). One of the format is Twenty20, in this format the game is played for several hours given 20 Overs to each team for batting. “The Indian Premier League (IPL), officially Vivo Indian Premier League for sponsorship reasons, is a professional Twenty20 cricket league in India contested during April and May of every year” (Wikipedia, 2007). In IPL every team represents the cities and some states of India.

DATA SOURCE

The data for the visualizations of “Indian Premier League (IPL Cricket) till 2016” is taken from the website of data.world (Raghunath, Karupakala, S. and Raik, S., 2017). The data from this link was the collections of 23 excel files. Most of the files were combined in one sheet of excel by creating unique key. In this data the modifications were made to get the visualization more precise and clear. This data contains all the information of matches from 2008 to 2016, like the runs scored in each ball for particular over in particular inning. The data was published on website on 22 May 2017. Figure 1, shows the screenshot of data.

B	C	D	E	F	G	H	I	J	K
Date	Date1	over_id	ball_id	innings_no	Team_Batting	Team_Bowling	striker_batting_position	Striker	Non_Striker
4/18/2008	2008	1	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	SC Ganguly	BB McCullum
4/18/2008	2008	1	1	2	Royal Challengers Bangalore	Kolkata Knight Riders	1	R Dravid	W Jaffer
4/18/2008	2008	1	2	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	1	2	2	Royal Challengers Bangalore	Kolkata Knight Riders	2	W Jaffer	R Dravid
4/18/2008	2008	1	3	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	1	3	2	Royal Challengers Bangalore	Kolkata Knight Riders	2	W Jaffer	R Dravid
4/18/2008	2008	1	4	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	1	4	2	Royal Challengers Bangalore	Kolkata Knight Riders	2	W Jaffer	R Dravid
4/18/2008	2008	1	5	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	1	5	2	Royal Challengers Bangalore	Kolkata Knight Riders	1	R Dravid	W Jaffer
4/18/2008	2008	1	6	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	1	6	2	Royal Challengers Bangalore	Kolkata Knight Riders	2	W Jaffer	R Dravid
4/18/2008	2008	1	7	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	1	7	2	Royal Challengers Bangalore	Kolkata Knight Riders	2	W Jaffer	R Dravid
4/18/2008	2008	2	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	2	1	2	Royal Challengers Bangalore	Kolkata Knight Riders	1	R Dravid	W Jaffer
4/18/2008	2008	2	2	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	2	2	2	Royal Challengers Bangalore	Kolkata Knight Riders	3	V Kohli	W Jaffer
4/18/2008	2008	2	3	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	2	3	2	Royal Challengers Bangalore	Kolkata Knight Riders	3	V Kohli	W Jaffer
4/18/2008	2008	2	4	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	2	4	2	Royal Challengers Bangalore	Kolkata Knight Riders	3	V Kohli	W Jaffer
4/18/2008	2008	2	5	1	Kolkata Knight Riders	Royal Challengers Bangalore	2	BB McCullum	SC Ganguly
4/18/2008	2008	2	5	2	Royal Challengers Bangalore	Kolkata Knight Riders	2	W Jaffer	V Kohli

Data | Extra runs | batsman_scored | wicket taken | batting style | Wicket Type | Wicket Take by Purple Cap

Figure 1: Data Set

The analysis is done using desktop version of Tableau Professional Edition and Excel. Both the tools provide varies number of graphs and chart for the easy analysis. In tableau it is easy to plot the cities using coordinates of latitude and longitude which makes visualization more accurate and easy.

In this data set there are number of fields such as the venue name, city name, country name where the match held. Also the data is having the date on which the match was played.

VISUALIZATION:

Number of Matches Played Year wise:

The number of matches played varies from 2008 to 2016. This is shown in figure 2 line chart is used here. In year 2008 to 2010 the number of matches played were less as compared to the next three years. This was because the match played by particular team was less for the year 2008 to 2010. From 2011 to 2013 the numbers of matches played are 72, 74 and 76 respectively. This is because the team got increased more new teams were added in the IPL. After 2013, there was change in the rules of IPL so the graph of number of matches played tends to down to 60 in the year 2016.

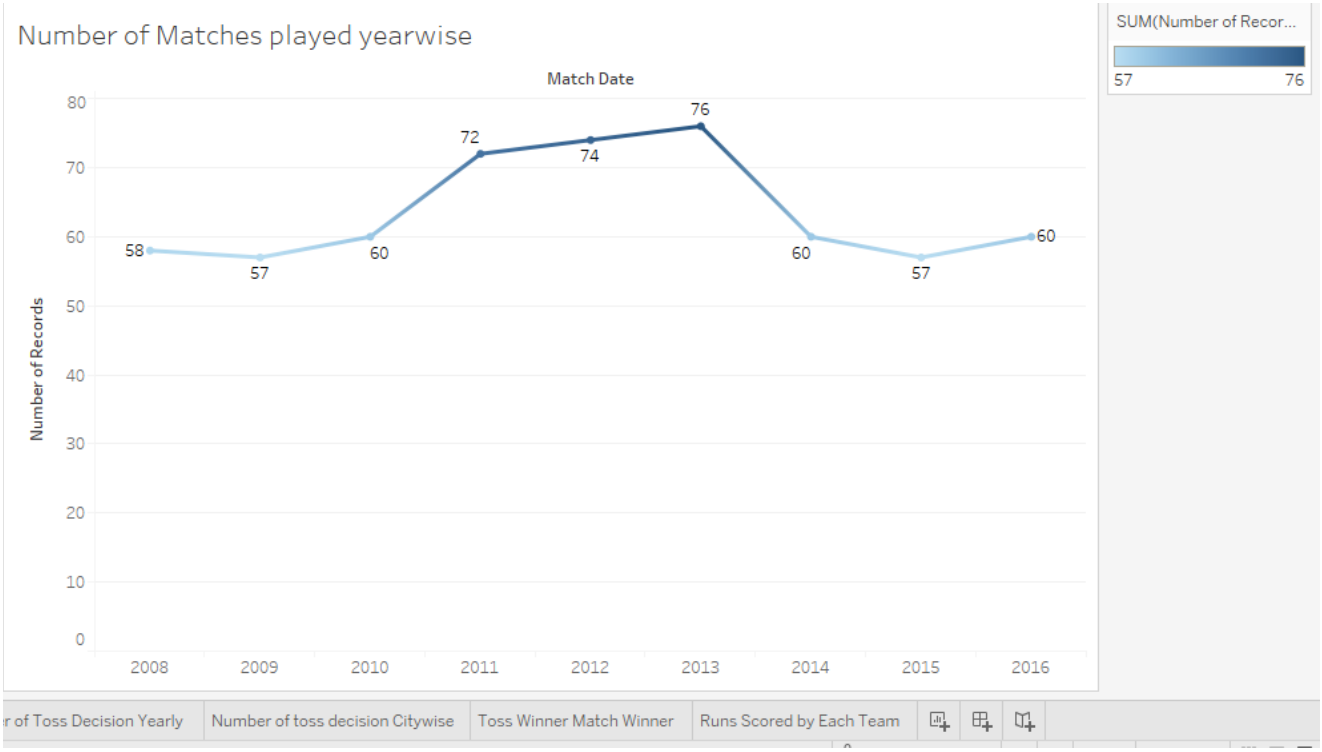

Figure 2: Number of Matches Played Year wise

Number of Toss Decision City wise:

Before starting the match the toss is done. This outcome of toss can be to choose to bat or to field. In Figure 3 symbol map is used, here the city wise number of toss decision made is shown. If the outcome of toss is bat it is shown in blue color and if it is to field it is shown in orange color. For example, in city Kolkata which is in country India the chosen toss decision is to bat and the number of records is 28 and similarly in city Durban which is in country South Africa the toss decision chosen is to field having number of record as 5. So, in this way the visualization is done city and country wise according to the toss decision.

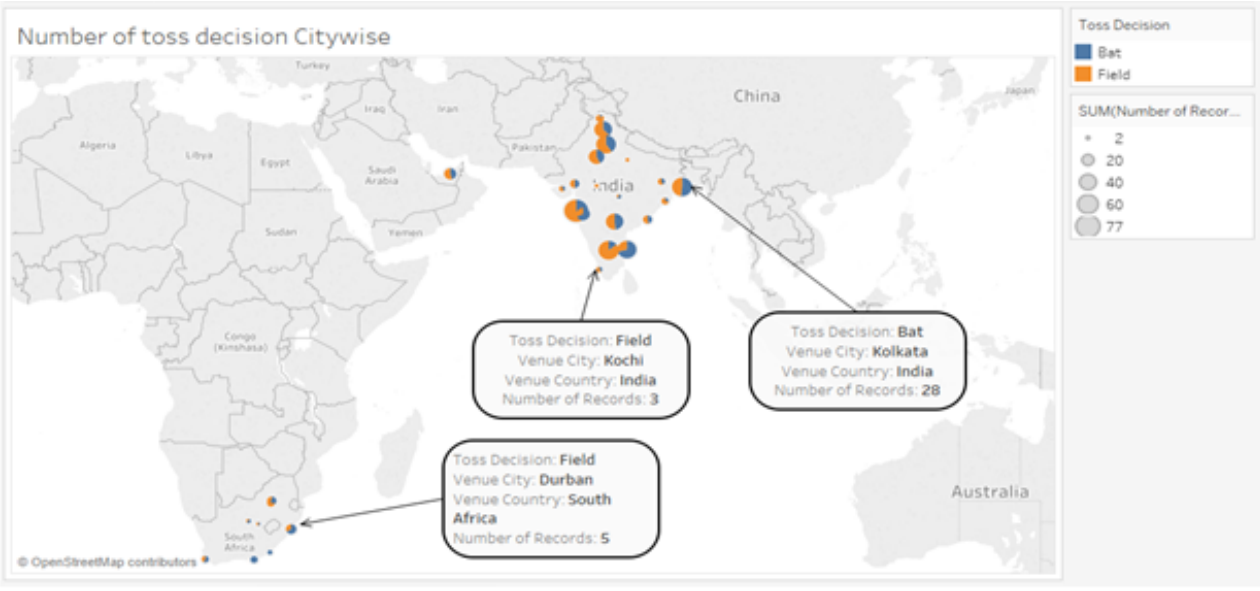

Figure 3: Number of Toss Decision City wise

Number of Toss Decision Yearly:

This toss decision can be analyzed yearly. Figure 4 shows the number of toss decision made yearly. To visualize this side-by-side bars chart is used. For 9 different years the 9 different colors are used which makes the visualization more clear and precise. Here, it can be seen that for year 2013, toss decision made against bat is 90 which is the highest among all the years and for field in same year it is 31.

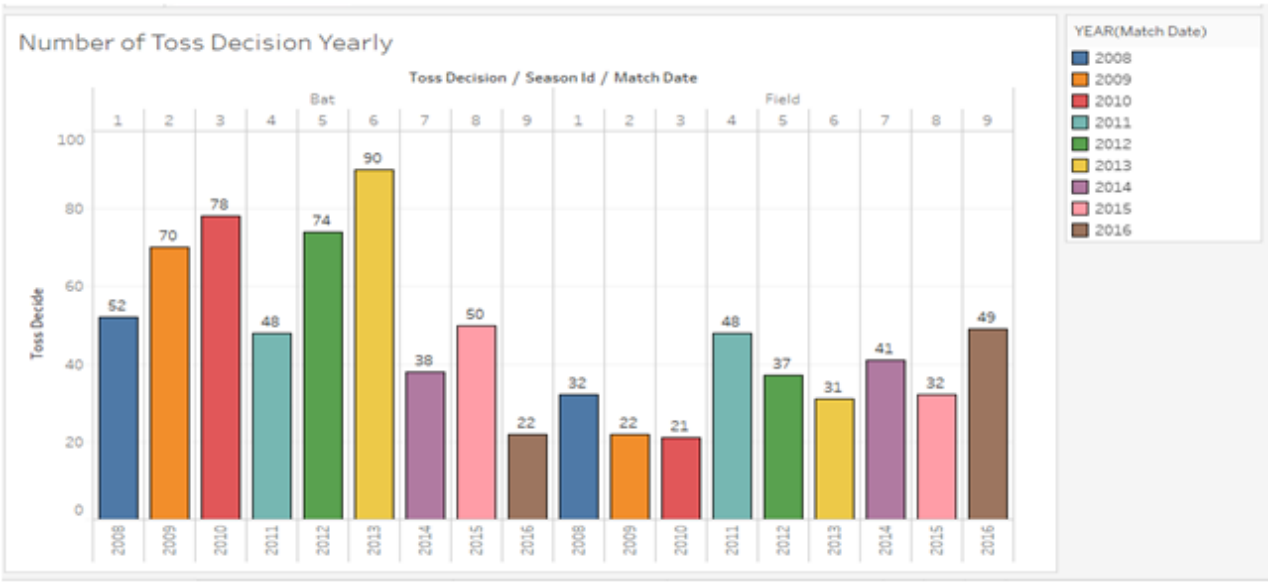

Figure 4: Number of Toss Decision Yearly

Toss Winner Match Winner Matrix:

The visualization can be made on the toss winner and the percentage of match winner. In the figure 5 highlight table is used. Here it can be visualized that the team who win the toss for the same team what is the percentage to win the match in comparison with other teams. For example, it can be seen that if the Chennai Super Kings win the toss than then percentage of winning the match of Chennai Super Kings is 63.64% whereas the percentage of winning other team like Mumbai Indians is 6.06%

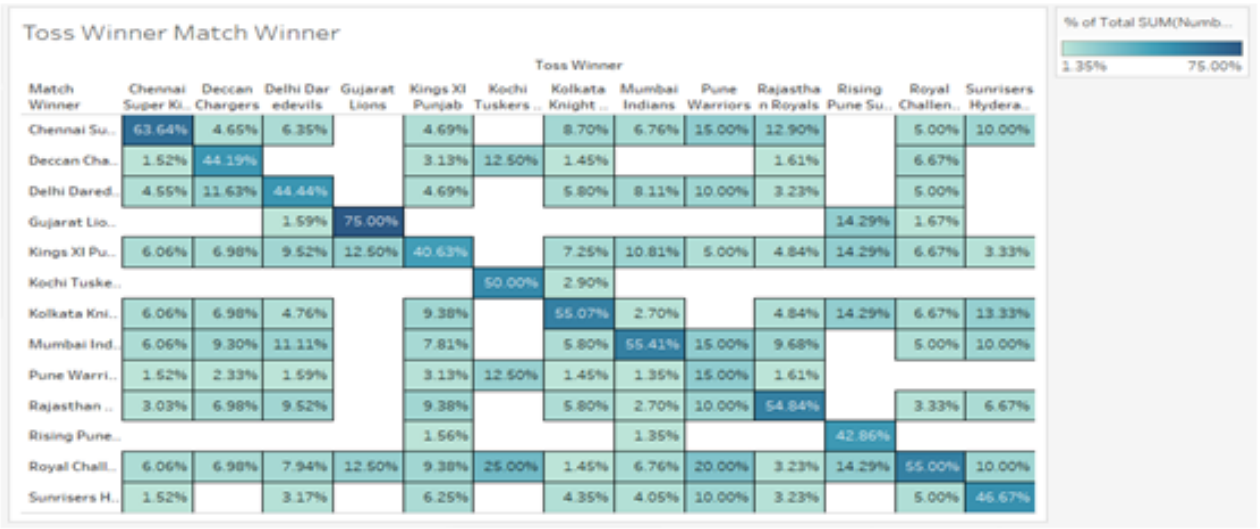

Figure 5: Toss Winner Match Winner

Number of Matches Won according to Batting:

The analysis and visualization is done for the number of matches won by particular team if it goes for batting first or batting second. In figure 6, 3D cylinder column chart is used. Here it can be visualized that the team which goes for batting first won how many number of matches. For example, the Chennai Super Kings won maximum number of matches if it goes for batting first.

Figure 6: Number of Matches Won

Number of boundaries scored by the Team:

The runs scored by the team can be in boundaries, these boundaries are 6's or 4's. In the figure 7, clustered bar graph is used. Here it is visualized that the number of 6's is shown by red color and number of 4's is shown by blue color. The more number of 6's is 1926 which is scored by Mumbai Indian team. And the maximum number of 4's is 859 which is scored by Royal Challengers Bangalore.

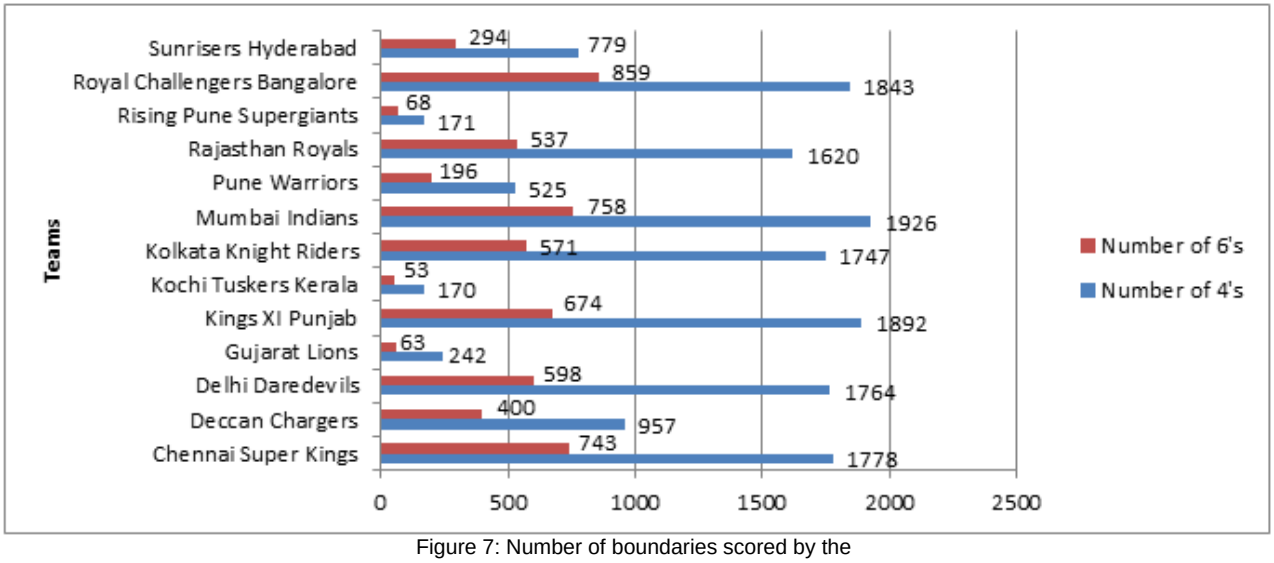

Figure 7: Number of boundaries scored by the Team

Percentage of Wicket Type:

In the cricket match the wicket can be taken by various types. In figure 8, 3D pie chart is show with the percentage of wicket taken in different type. Here different colors are used for different type of wicket, by which the visualization becomes more clear and specific. The more percentage of wicket taken is by type caught (58.78%).

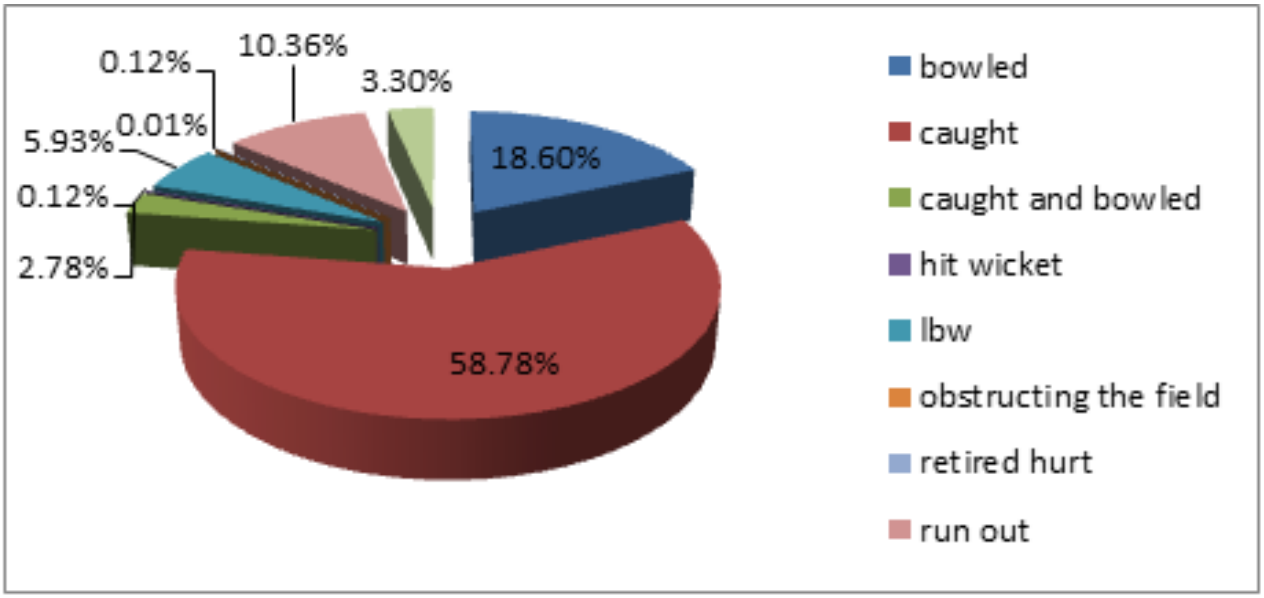

Figure 8: Wicket Type (in percentage)

Run Scored by Each Team:

The total run scored by each team in all the seasons from year 2008 to 2016 can be visualized. In the figure 9, a packed bubble is used to visualize the number of runs scored by each team. Here different colors are used for different teams. And the sizes of the bubble symbolize the number of runs scored. The bigger the size of bubble the more the number of runs scored. Here the highest run scored is 16,988 by the team Mumbai Indians.

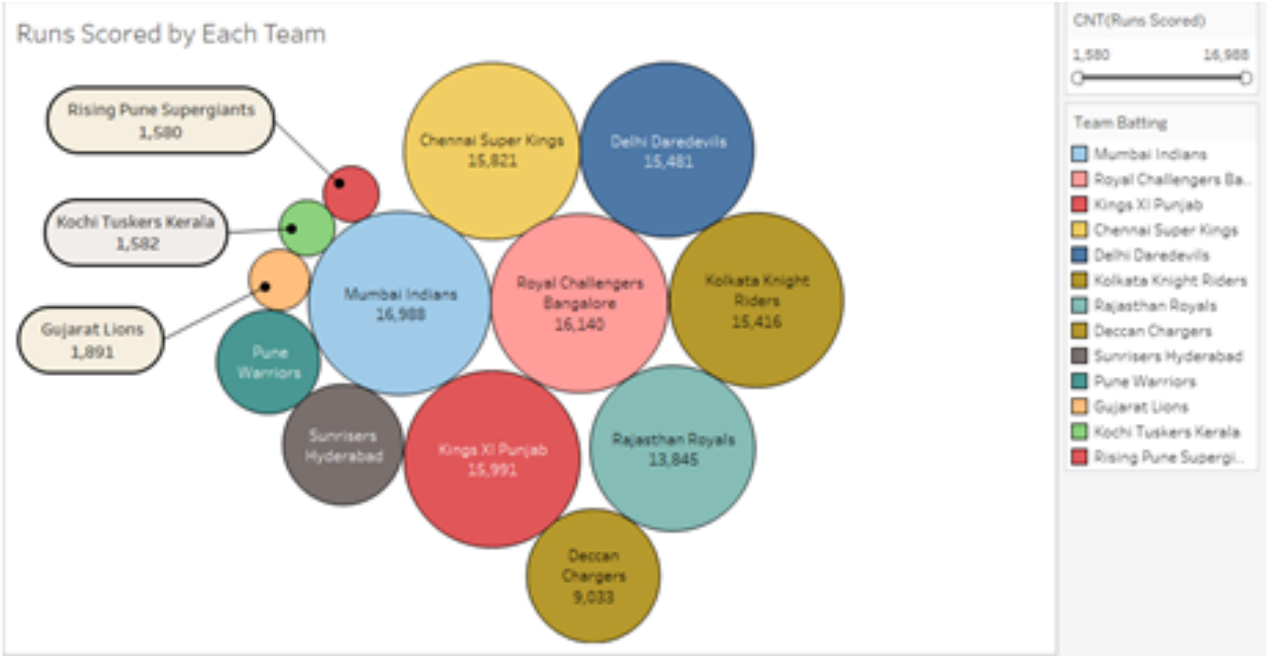

Figure 9: Runs Scored by Each Team

Runs Scored by Top 10 Players:

For scoring the runs the important role is played by the batters. In all the matches from 2008 to 2016 top 10 players with their scores is visualized. In figure 10, exploded doughnut is used. Here the different colors are used for 10 different players. The highest run is scored by SK Raina, he is the 1st player scoring 4106 runs. Similarly, the lowest run is scored by S Dhawan.

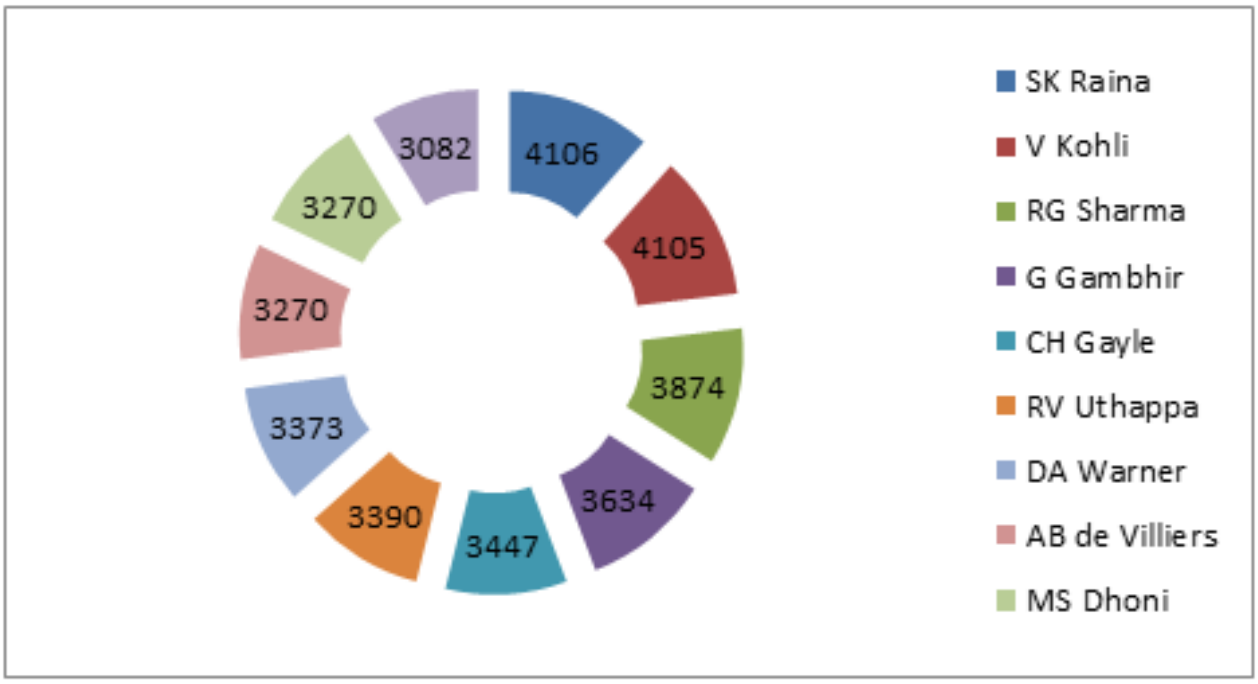

Figure 10: Runs scored by top 10 players

Runs Scored by Orange Cap Holder:

In IPL the annual awards are presented. One of these awards the Orange Cap. This award is given to the player who is the highest run scorer. The player or the bats man who scored the highest runs in the tournament would wear the Orange Cap during fielding. In figure 11, radar chart with markers is used. Here, the highest scored is 969 which was in the year 2016.

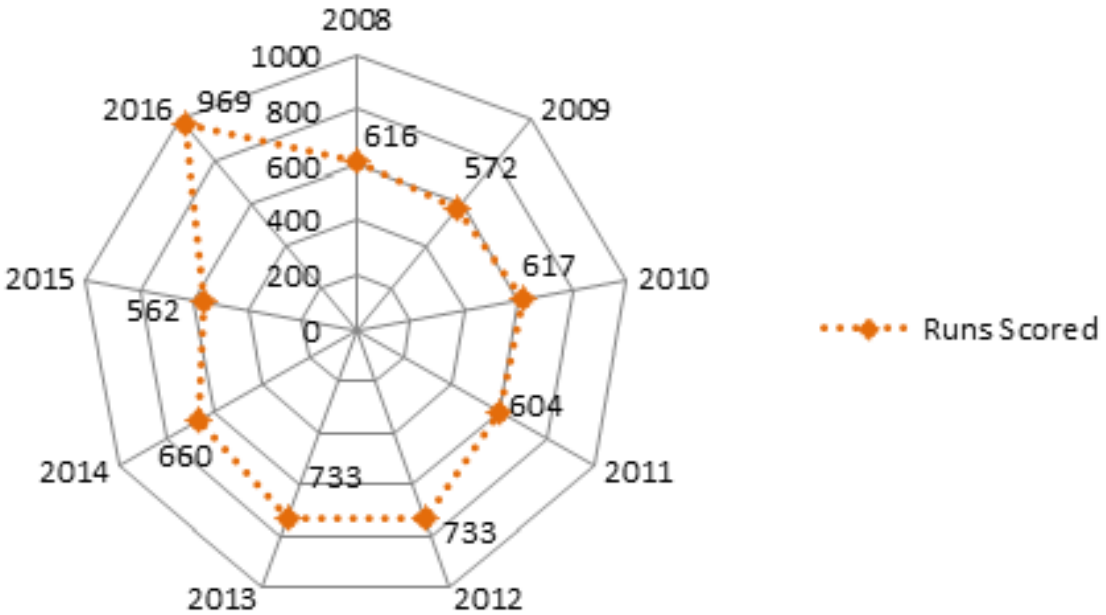

Figure 11: Runs Scored by Orange Cap Holder

Wicket Taken by Purple Cap Holder:

Purple Cap is another type of award. This award is given to the highest wicket-taker in the tournament. If the result of match is tie, then the bowler having high economy rate will be given the Purple Cap. In figure 12, radar chart with markers is used. Here, the highest number of wicket is 32 which were in the year 2013.

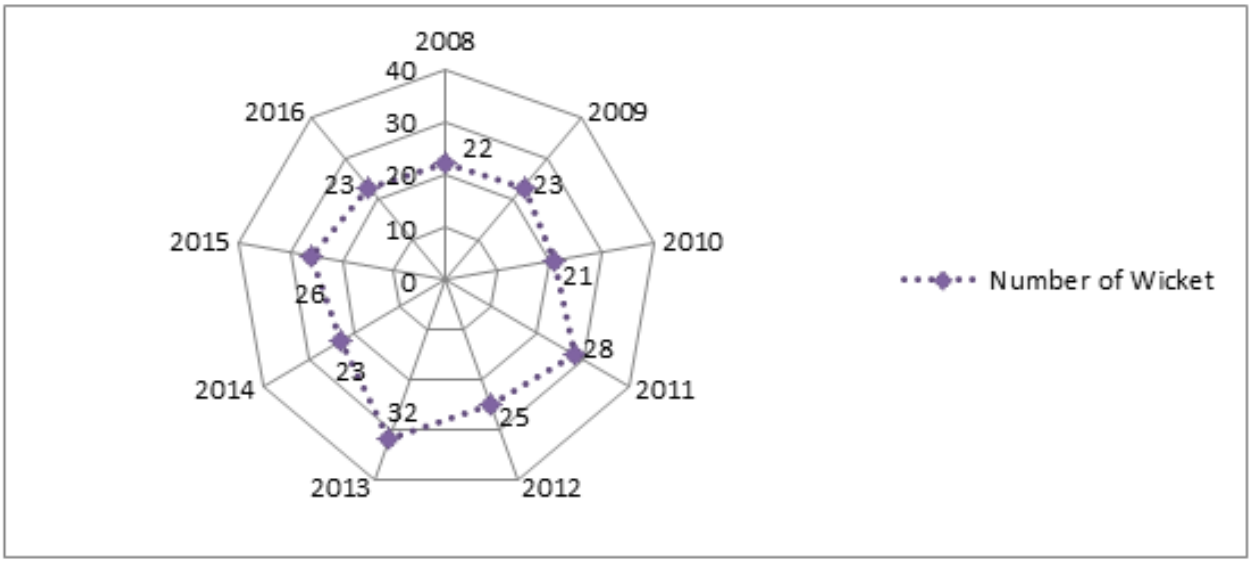

Figure 12: Wicket Taken by Purple Cap holder

CONCLUSION:

Thus, after the analysis and visualization of IPL data the comparisons between the teams are made, the runs scored in all the matches are compared, total number of matches held in particular year is seen, the top 10 players who scored more in the matches are visualized. By this visualization the performance of individual team can be increased. The wicket type can be seen by which type of wickets the more wicket take place. The visualization also includes the awards which are held in the tournament.

For this analysis the tableau and excel software is used. Both the software is user friendly and the prediction and result was accurate.

Reference:

1. Wikipedia (2019) *Forms of Cricket* [Online] Available at: https://en.wikipedia.org/wiki/Forms_of_cricket [Accessed 21 February 2019]

2. Wikipedia (2007) *Indian Premier League* [Online] Available at: https://en.wikipedia.org/wiki/Indian_Premier_League [Accessed 21 February 2019]

3. Raghunath, Karupakala, S. and Raik, S. (2017) *Indian Premier league(IPL Cricket) till 2016* [Online] Available at: https://data.world/raghu543/ipl-data-till-2016-set-of-csv-files [Accessed 2 February 2019]

4. https://www.canva.com/create-a-design [Accessed 21 February 2019]

Appendix:

Figure No:	Description
1.	Data Set: The screenshot provided is of dataset.
2.	Graph: Line Chart (Number of Matches Played Year wise). Color: Shades of blue. Tool: Tableau
3.	Graph: Symbol Map (Number of Toss Decision City wise). Color: Two colors are used orange and blue to identify between bat and field. Tool: Tableau
4.	Graph: Side-by-side bar chart (Number of Toss Decision Yearly). Color: Nine colors are used to identify 9 different years Tool: Tableau
5.	Graph: Highlight Table (Toss Winner Match Winner). Color: Blue color shades are used Tool: Tableau
6.	Graph: 3D Cylinder Column (Number of Matches Won). Color: Three different colors are used green, blue and red. Tool: Excel
7.	Graph: Clustered bar graph (Number of Matches Won Number of boundaries scored by the Team) Color: Two different colors are used blue and red to identify number of 6's and 4's. Tool: Excel
8.	Graph: 3D Pie chart (Wicket Type (in percentage)) Color: Eight different colors are used to identify different types of wicket. Tool: Excel
9.	Graph: Packed Bubble (Runs Scored by Each Team) Color: Thirteen different colors are used to identify different teams. Tool: Tableau
10.	Graph: Exploded Doughnut (Runs scored by top 10 players) Color: Ten different colors are used to identify top ten players. Tool: Excel

11.	Graph: Radar Chart (Runs Scored by Orange Cap Holder) Color: An orange color is used to identify run scored by orange cap holder. Tool: Excel
12.	Graph: Radar Chart (wicket Taken by Purple Cap holder) Color: An Purple color is used to identify wicket taken by purple cap holder. Tool: Excel