Part 1

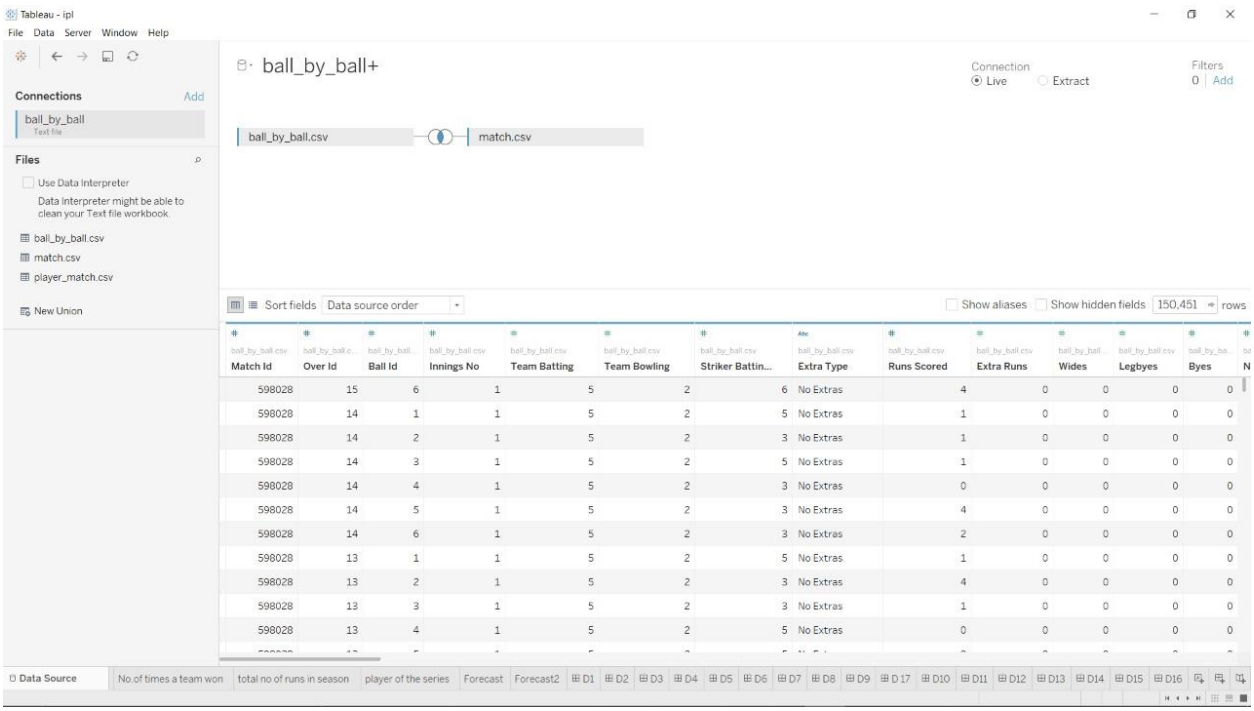
Exploratory Analytics and Visualization of Data via Information Dashboard

Introduction:

The datasets which are used in this exploratory analytics are IPL datasets that are taken from the data.world.gov and we took the data from 2009-2017 and also used another excel sheet that took from article for listing of all the IPL cup winners from 2008-2019.I used inner joins for some sheets and also blended many sheets and made relationships in between those sheets.

Link: https://data.world/raghu543/ipl-data-till-2017

Attribute	Explanation
Match_id	Unique id for every match
Extra_type	Shows what type of extra run
Runs_scored	How many the bowler gave in each ball
Extra_runs	How many extra run gave for that ball
wides	Shows no of runs given by wide
byes	Shows no of given by bye runs
Noball	Shows whether noball or not
Out_type	Shows type of wickets
Caught	Whether out type is caught or not
Bowled	Whether out type is bowled or not
Run_out	Whether out type is runout or not
Lbw	Whether out type is lbw or not
Retired_hurt	Whether out type is retired hurt or not
Stumped	Whether out type is stumped by keeper or not
Hit_wicket	Whether out type hit wicket or not
Season	Displays which season it was
Striker	Shows striker unique number
bowler	Shows unique number
Player_name	Shows player name
Player_team	Shows which team the player belongs
Player captain	Displays caption of team
Player_keeper	Displays keeper name
City	In which city the match hosted
toss_winner	Which team win the toss
toss_decision	What the toss winner choose
win_by_runs	With how much runs the team wins
win_by_wickets	By how many wickets the team wins
umpire1	Umpire 1 name
Umpire3	Umpire 3 name
Winner	Season winner team name
Runner up	Season runner up team name

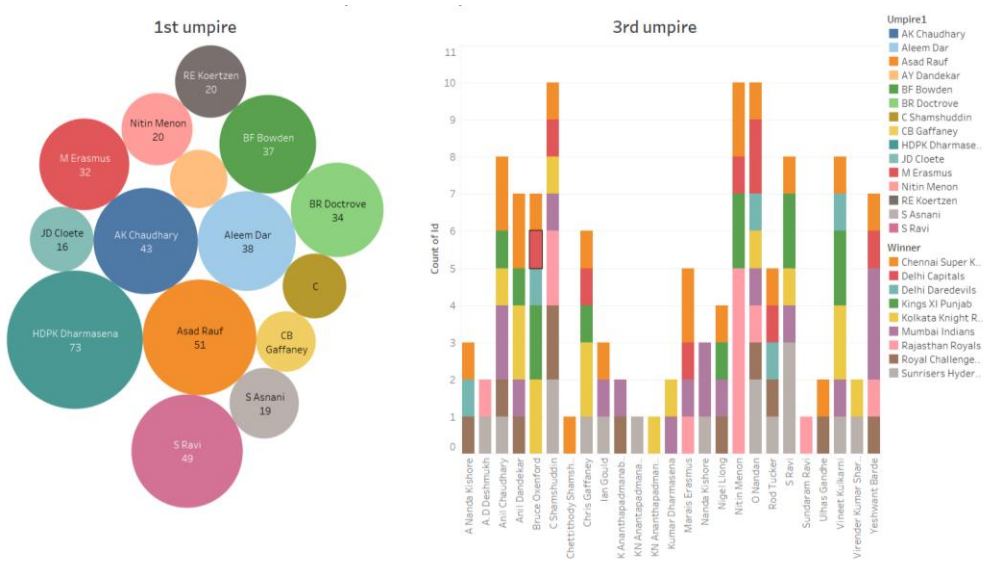


I used inner join to club these two datsets to obtain the information required to visualize the data as there is a match id which is unique for all the data and same for the data sets I used that variable to merge the data.

What Questions users might ask?

1.) Which team won many cups in IPL?
2.) Which player did more runs in whole data?
3.) What is the name of the bowler who bowled more no balls?
4.) What is the name of the captain who played as a captain for more winning team?

Dashboard 1: 1st umpire & 3rd empire with no of matches



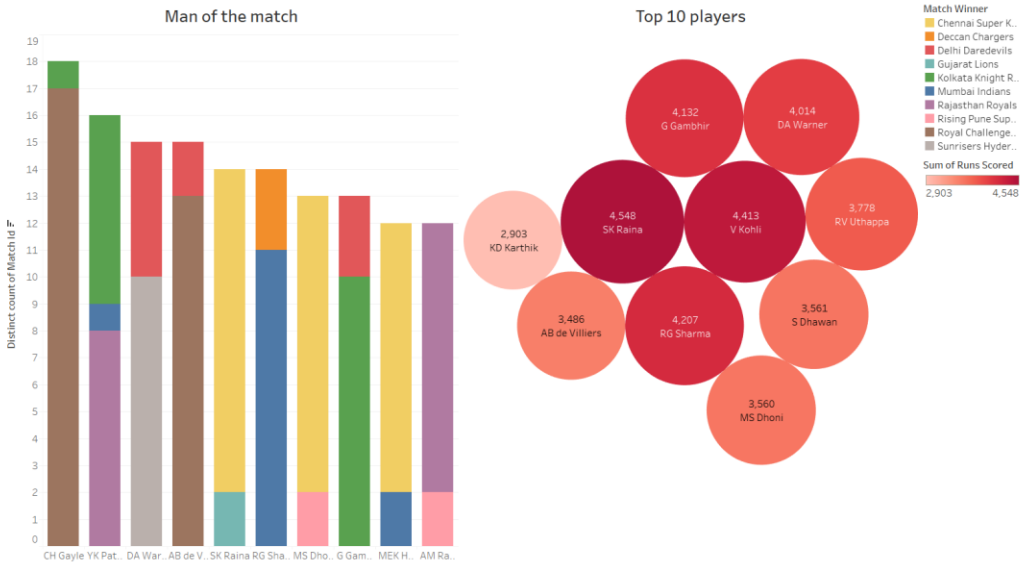
In the IPL, umpires plays a key role for match. In the above analysis it shows which umpires are mostly doing umpiring in the IPL from 2008 to 2017. From this analysis we can know that HDPK Dharmasena has done umpirirng for 73 matches for 1st umpiring and for 3rd umpiring C shamshuddin, Nitin menon, o Nandan made umpiring for 10 matches.

Outcome:

For 1st umpire: HDPK Dharmasena most favorable umpire

For 3rd umpire: C shamshuddin, Nitin menon, o Nandan are most favorable 3rd umpires.

Dashboard 2: Man of the Match and Top 10 players.



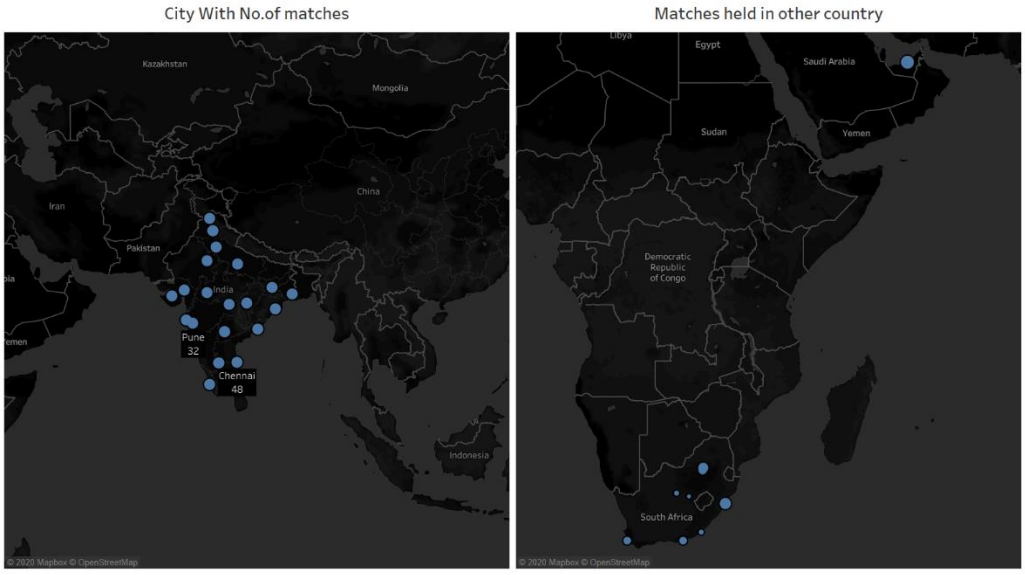
From the above analysis we can clearly see that CH Gayle is continuously going for man of the match for 17 times .Most of the matches are played under Royal challengers Bangalore . And for Top 10 players Raina and kohli are close each other with highest number of runs.

Outcome:

Most man of the matchs: CH Gayle for 18 times

Top 10 players: SK Raina is at top with 4548 runs.

Dashboard 3: City with no of matches&matches held in other countries



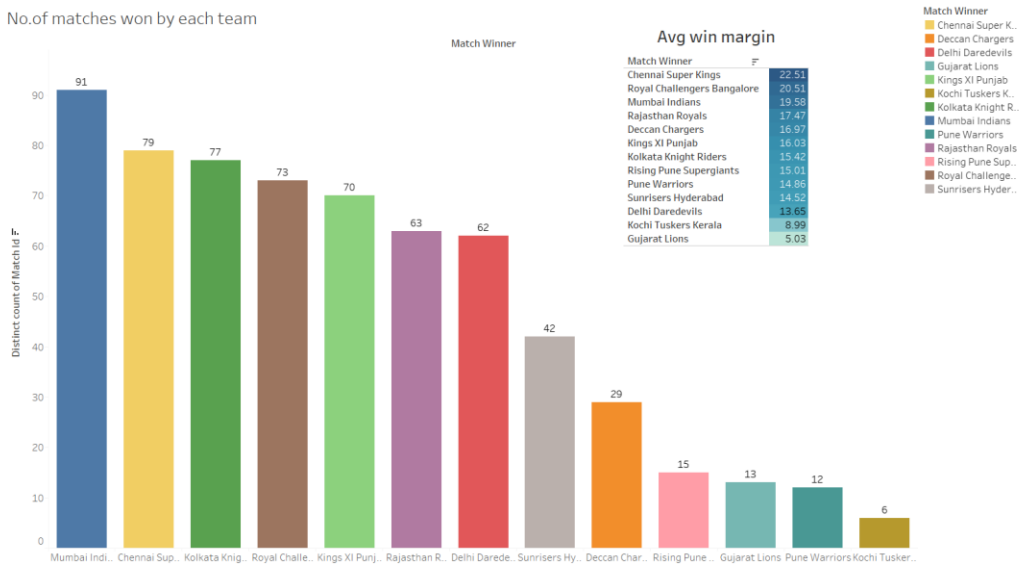
In the above map it points that city with no of matches are held by IPL and it shows the city and no of matches .In this map Mumbai has venue to 81 no of matches place top for held IPL. Next one goes to Kolkata and beside that some matches are held in other countries most of the matches are held in south Africa.

Outcome:

Mumbai places top for held IPL venue for most number of matches.

South Africa held most number of matches in other countries.

Dashboard 4: No of matches won by each team and Avg win margin



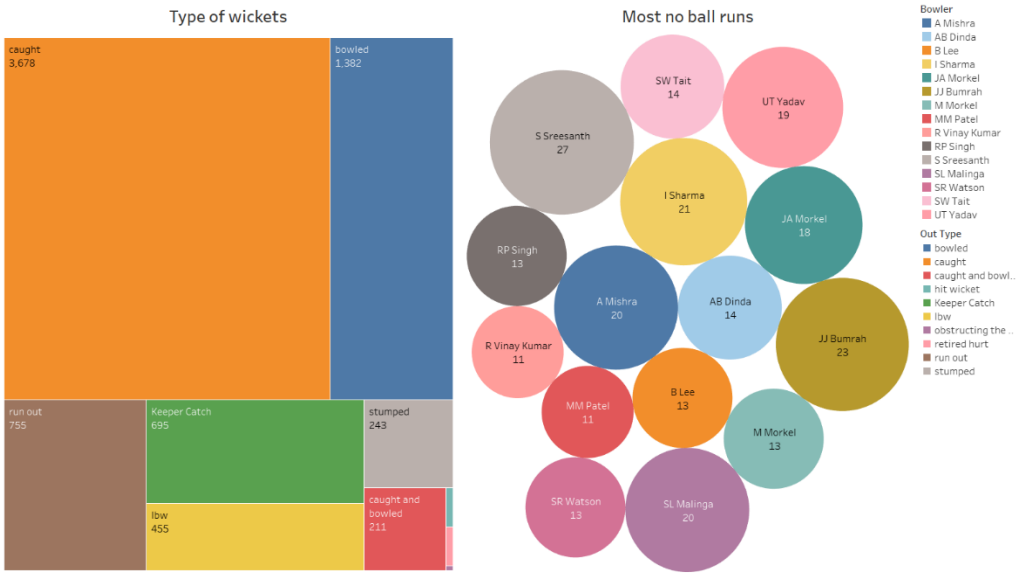
We already know that in ipl Mumbai Indians and Chennai super kings are always competitor for each other. While coming to most number of win matches Mumbai Indians wins 91 matches followed to this Chennai super kings won 79 matches. While coming to average margin win Chennai super kings places top with 22.51 runs to win against other.

Outcome:

Mumbai places top by most winning matches

Chennai tops by avg win margin by 22.51 runs

Dashboard 5: Type of wickets& no ball runs



In the IPL the wickets took by bowlers mostly by catches, after to the catches the wickets mostly fallen by bowled. But there is large margin between the two, it is around 1500 wickets . Least wicket taken by hit wicket and retire hurt.

This dashboard gives the details of bowlers who bowled most no balls in IPL and gave runs, we can see S.sreeshant is the player who bowled 27 no balls and gave 27 runs to opposite team and we can also see the only top bowlers who gave most number of no ball runs.

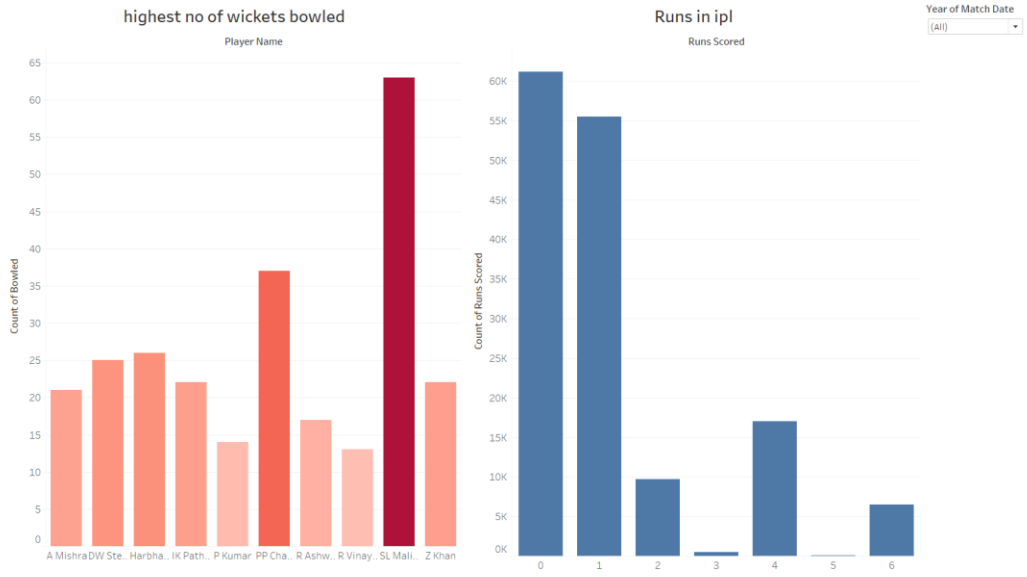
Outcome:

Most number of wickets are taken by bowled that are 1382.

Most no ball runs given: S.sreeshant

Most no ball runs bowled by bowlers and no of bowls bowled.

Dashboard 6: Highest number of wickets bowled by bowler and runs in IPL.



In the above analysis it shows the top 10 bowlers by most wicket takers and runs in ipl.

By coming to bowlers malinga took 91 wickets places top position . Next to him pp chawla places in second with 37 wickets.

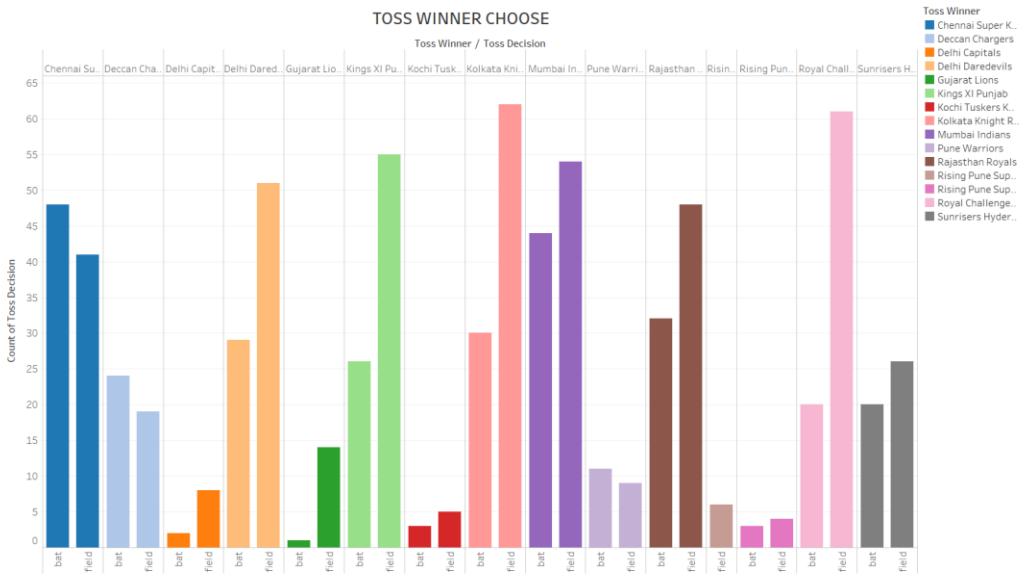
While coming to runs mostly of runs took by batsmens are singles and fours.

Outcome:

SL malinga places tops with most number of wicket takers.

Singles are the runs most batsmens taken.

Dashboard 7: Toss winner chosen

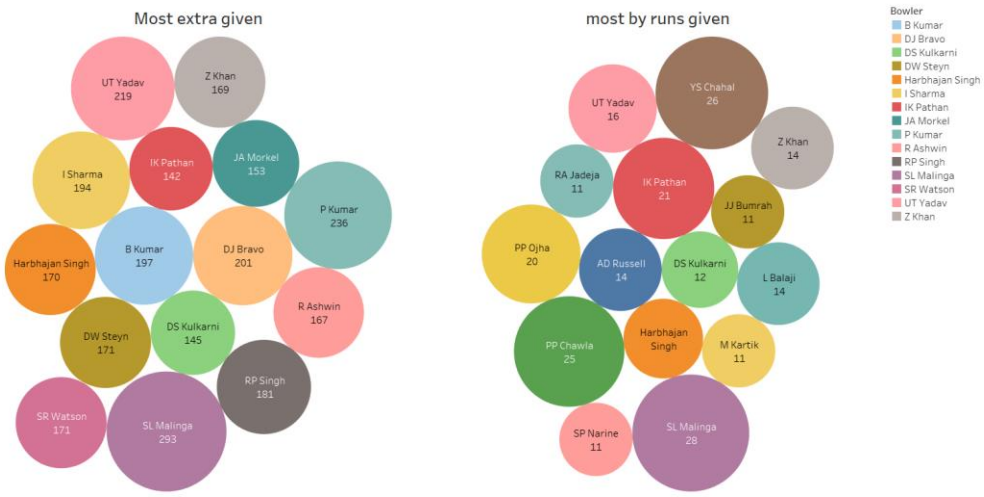


In the above analysis it shows toss winner choose bat or field most of the times Kolkata won the toss with more number times and choose too field. Next to that Royal challengers banglore won many times and choose field .It shows that most of teams want field first.

Outcome:

Most of time Kolkata knight riders won the match and choose to field

Dashboard 8: Most extra runs and most by runs given by bowler



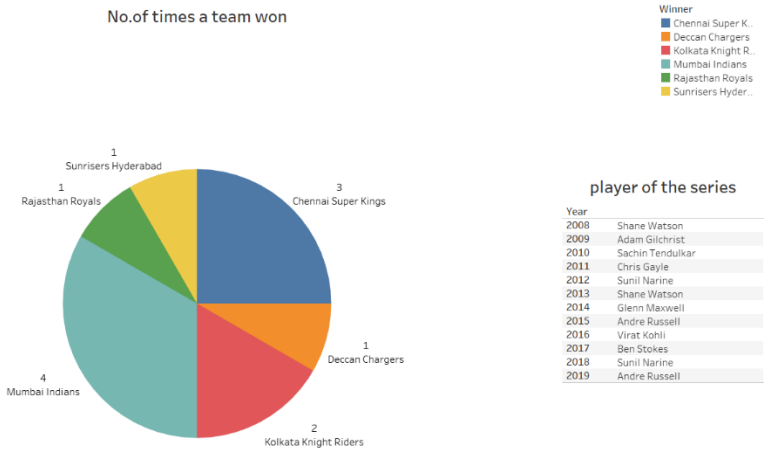
In the above analysis it shows that Most extra runs given by bowler and most by runs given by bowler.

In both of that SL malinga places top with 293 extra runs and 28 by runs.

Outcome:

Malinga gives most extra runs and most by runs given and also he tooks most number of wickets.

Dashboard 9: Most Number of IPL Cups Won



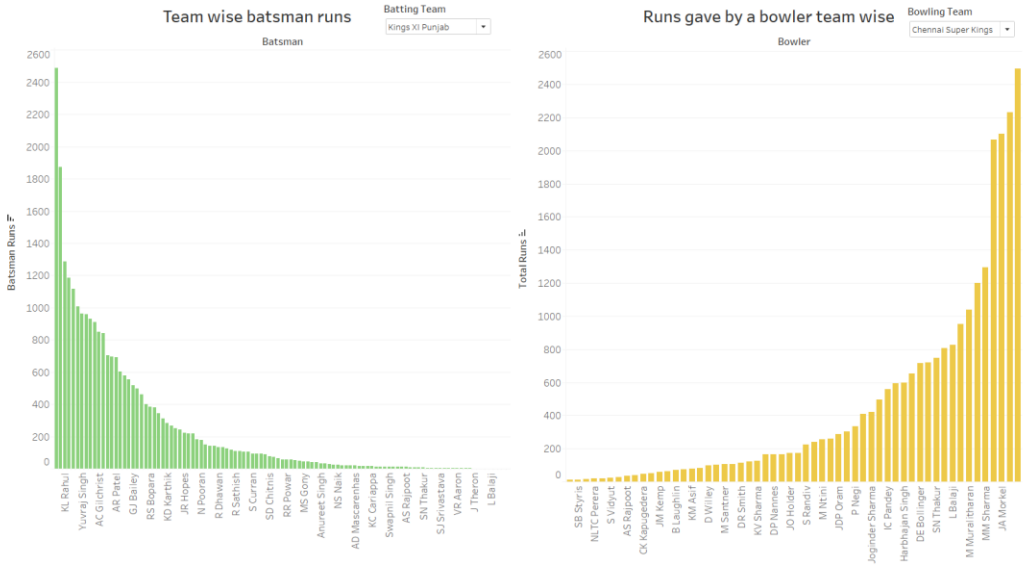
In this dashboard we can find which team won more IPL cups and Mumbai Indians is the team which one cup 4 times from 2008 to 2019 and also we have a table of player of series which gives us the year and name of the player.

Outcome:

Most no. of Cups won: Mumbai Indians

Player of the series from 2008-2019.

Dashboard 10: No of runs taken and given by players



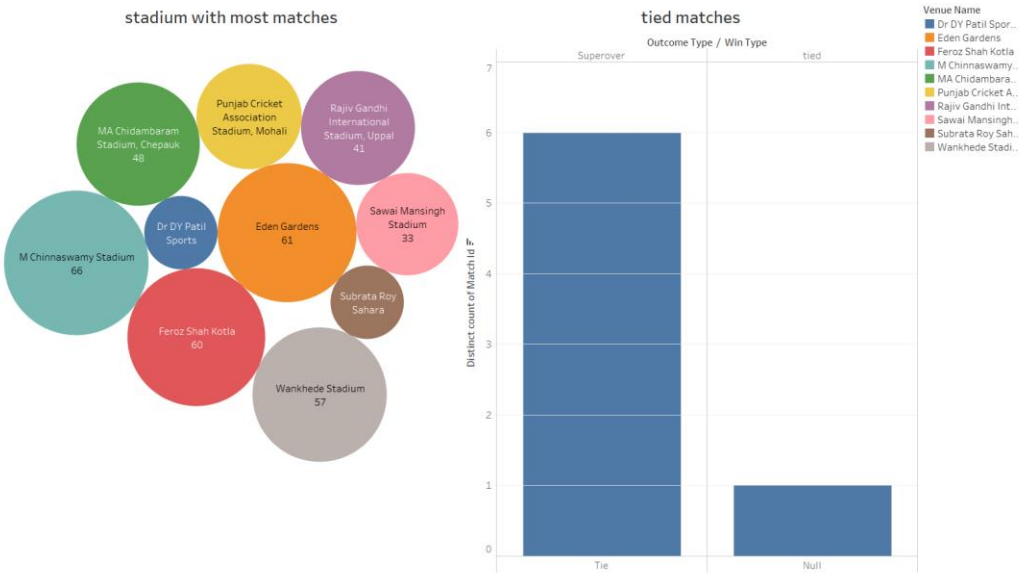
This dashboard tell us that which player had taken more runs and which player had given less runs and more runs by bowling and by the filter u can select and team wise performance of batting and bowling and list of players in that team.

Outcome:

No. of runs taken by batsman team wise

No. of runs given by bowler team wise.

Dashboard 11: Stadium hosted most matches and no. of matches got tied



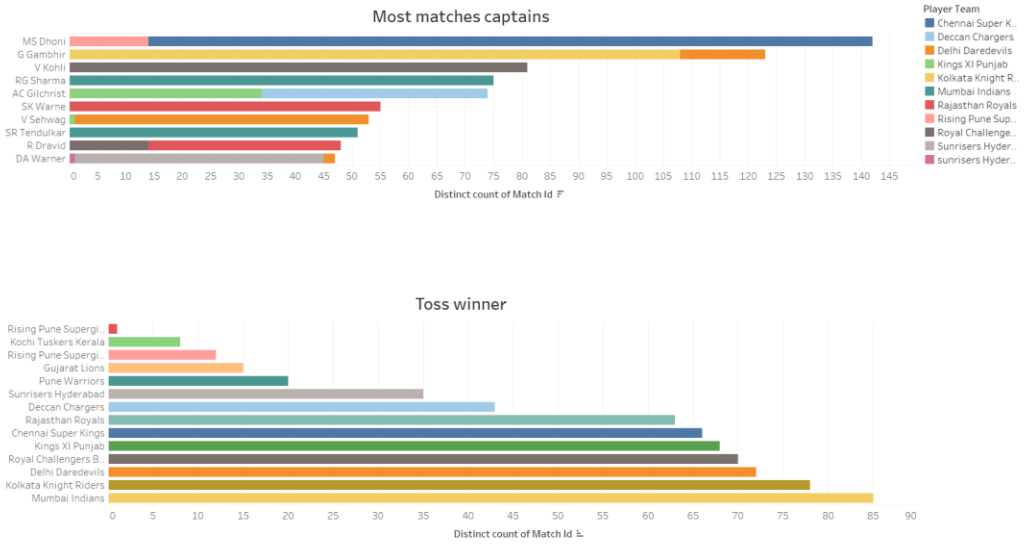
This dashboard gives the data which stadium in India hosted more no of matches for IPL from 2008-2017. We can see that M.Chinnaswamy stadium hosted 66 matches and we can also that the matches are tied and played super yet for some matches the results are not announced.

Outcome:

Stadium hosted more number of matches: M.Chinnaswamy Stadium

Number of matches tied and got result from super over.

Dashboard 12: Most matches captain& Toss winner



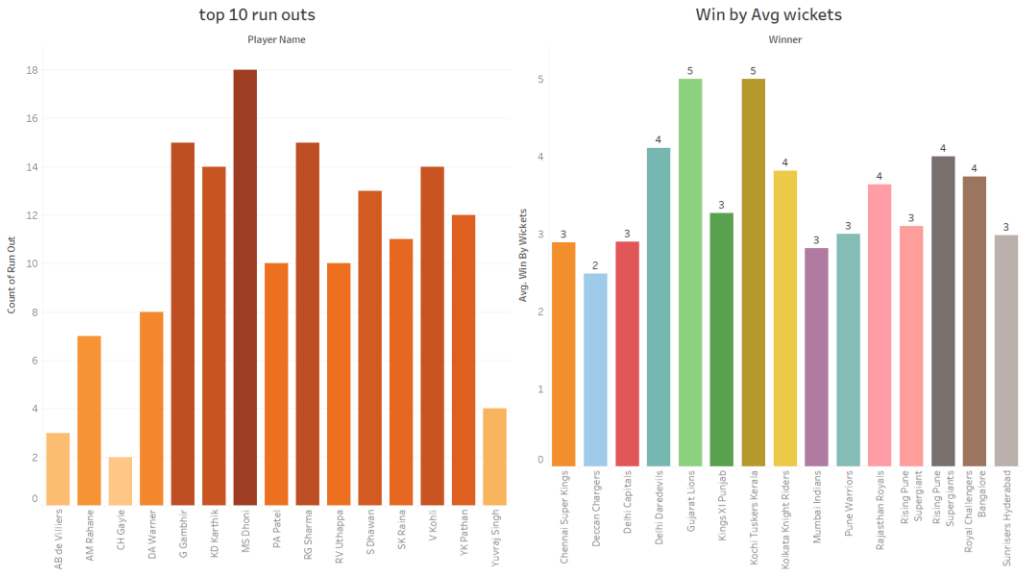
This dashboard shows which player is captain for most number of matches and which team won the toss. From the above figure we can see that MS Dhoni is the senior captain and done captaincy for over 143 matches and Mumbai Indians is the team which won toss most number of times from 2008-2017.

Outcome:

Most Number of matches captain: MS Dhoni

Most Tosses Won: Mumbai Indians

Dashboard 13: Top 10 run outs& Win by Avg Wickets.



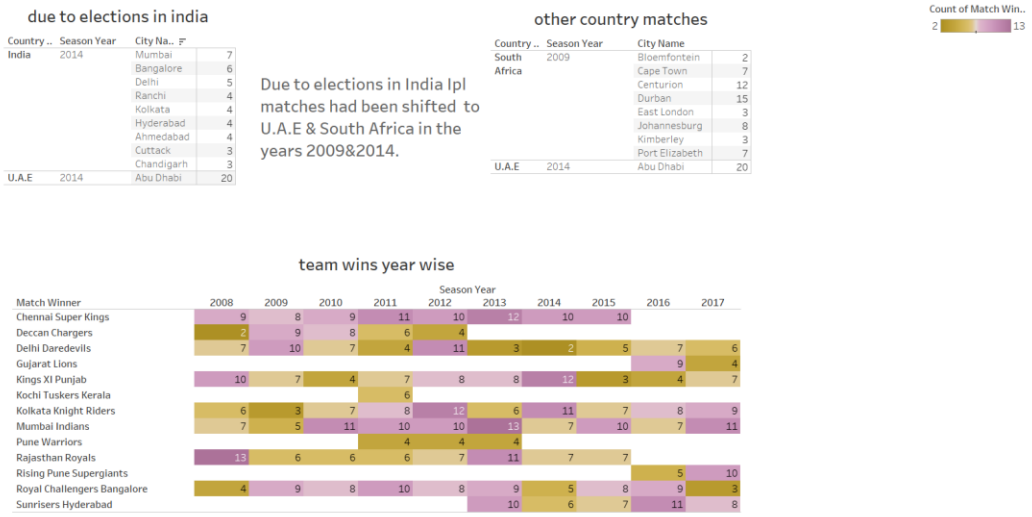
This dashboard shows us the data of top 10 players who are run outs in the IPL from 2008-2017. We can see MS Dhoni is the player who have most number of run outs and we used color intensity to represent which players have more number of run outs and in this dashboard most no of teams won by 5 wickets while other teams won with 3 or 4 wickets against other teams.

Outcome:

Most Number of runouts: MS Dhoni

Win by avg wickets.

Dashboard 14:



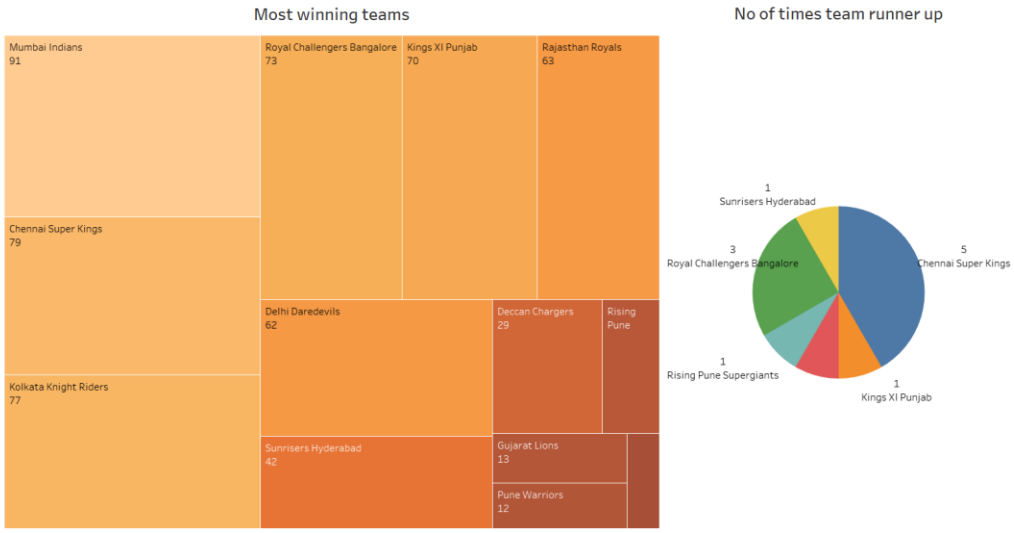
In this dashboard we can see some matches are played in other countries that is U.A.E& South Africa in the years 2009 and 2014 that is due to elections in India the games that should be played in India Had changed the venue to South Africa and U.A.E and the rest of the matches had happened there. We can also see team wins year wise that is which team won how many matches in that year regardless of final, semi-final and normal matches.

Outcome:

Matches Played in Other countries

Team wins year wise.

Dashboard 15: Most winning team



From the above dashboard we can see that which team won most number of times in the IPL from 2008-2017 and it is Mumbai Indians team which won 91 times and is the most winning team in whole IPL series from 2008-2017 and in this dashboard we can also which team has lost most number of final matches in IPL and Chennai super kings is the team which lost 5 times in final and yet Chennai is in the second most winning teams.

Outcome:

Most winning team: Mumbai Indians.

Most number of times lost final: Chennai Super kings.

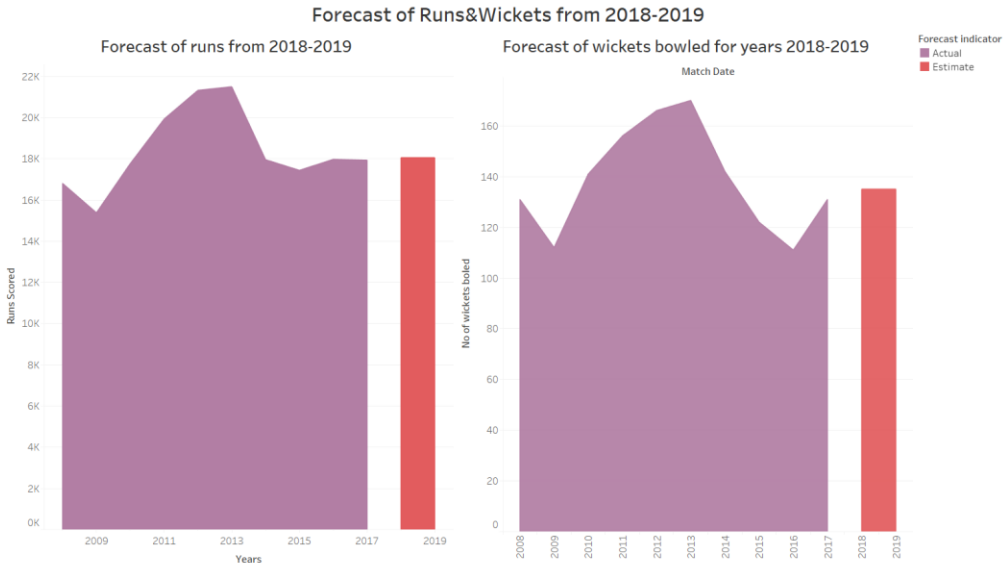
Part 2

Predictive Analytics and Visualization insights

What questions users might ask?

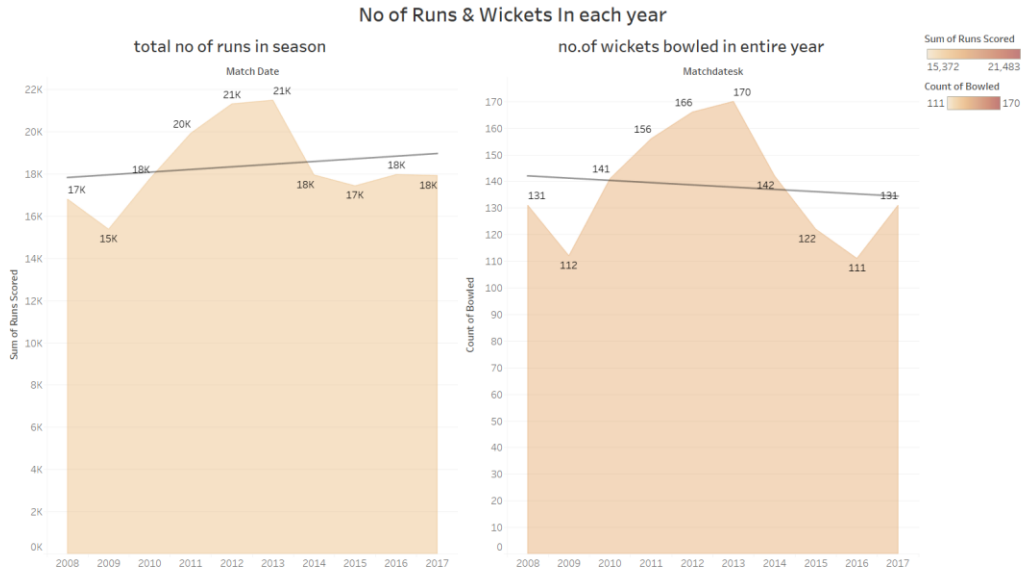
1.) What are the runs scored in future?
2.) How many wickets may be bowled in the future?
3.) What is the trend line/pattern over runs in whole data from 2009-2017?
4.) What is the trend line/pattern over Wickets Bowled in whole data from 2009-2017?

1.) Forecast of runs & wickets in Future



To answer the first 2 questions we are seeing the insights that I created from 2009-2017 and we can see that the curve had increases few years and then decreased over few years and when I used forecast function for 2 periods I saw that the runs that will be scored in the year 2018 and 2019 are similar and equal to 1.8k and the results are accurate when we checked the original result. And even no of wickets bowled in the year will be equal to 135.

2.) Trend lines/Patterns of Runs & Wickets from 2009-2017



We can observe the trend patterns from the above graph that the line is not linear it is some-times increasing and some-times decreasing that means the runs are not consistent in every year and even bowling is not consistent and we can also learn that there is no pattern in the above two graphs but we can find an insight that tells about the how many runs and how many wickets have been bowled during each year.