Super Predictor Of Indian Premier League (IPL)

Team Quad

1. Introduction

1.1 Overview

Though there are many sports in the world, cricket is one such sport which has a huge fan base. When it comes to Indian premier league, it has gained immense popularity not only in India but has led to remarkable status in international cricket. To address these sentiments of people, there is need to have a proper platform where one can see the trade off between the various factors which are effecting the victory of the match. The super predictor of Indian premier league is a tool where one can check the overall trends that took place in the IPL over past 10 years. It gives the insights about the dependent and independent factors that are lying under the victory of a match. The dash board itself gives the overall performance achieved by the individual teams and also helps to predict the probability of winning a match by estimating the passive elements such as winning a toss. The whole project is done with the help of IBM Cognos Analytics.

1.2 Purpose

People love to predict. When it comes to their favorite sport like cricket then this curiosity of predicting becomes more. Not everyone is aware of the strategies of the game. To help these set of fans, super predictor plays a major role. The main purpose of the project is to let people have better understandings over the IPL analytics. Even though the person has no prior knowledge of cricket analytics, he /she must be able to understand the overall performance that took place over the years. The project also focuses on the ease of understanding of the person. Even if the person is not good with numbers can understand, it is because of its graphical representation. which summarizes the entire tabular data into graphical form. so one can get through the data at a glance. It also helps the industries which relay on the prediction of cricket analytics.

2. Literature Survey

2.1 Existing Methods

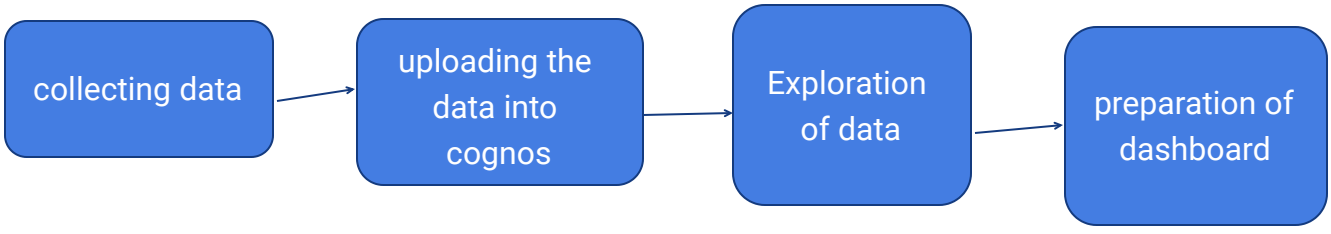
There are many websites which provide data sets of the past matches. These data sets are hard to understand. The data in the table can't be understood by simply looking at it. We need some special tools to get insights from them. Some of the websites such as towards data science, Kaggle.com etc provide us with huge amount of data. The only problem we face with these is the unordered organization of data. These websites provide data in tabular form. By looking at tables we can say how many columns are present but can't answer underlying facts.
Cricket apps such as cricbuzz provide us with information such as news , match schedule , point table. Some apps provide basic prediction. They won't sum up the entire insights.

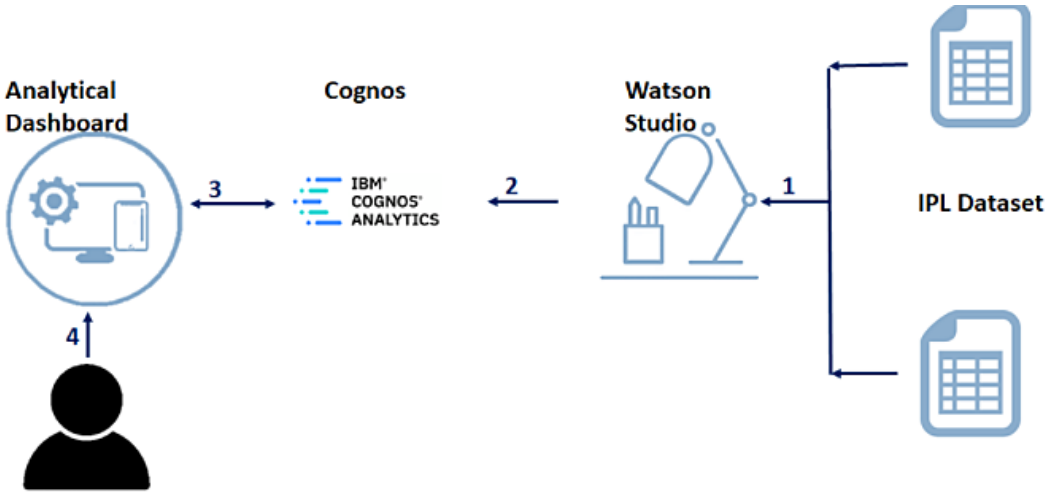
2.2 Proposed Methods

we are having the sites which provide data, the sites which provide only specific insight. There is a need for a application which provide overall insights, which will be easy to understand. In the project we made sure that all the interpretations which were made to be in graphical form. Customized tables can be found while interpreting the data. Has given the brief dependency relationship of passive components such as winning a toss and winning a match. Not only about teams the dashboard also gives information about the cities which high number of matches. It also provide information about the on field umpires. It means it covers overall aspects of Indian Premier League.

3. Theoretical Analysis

3.1 Block Diagram





3.2 Software/Hardware designing

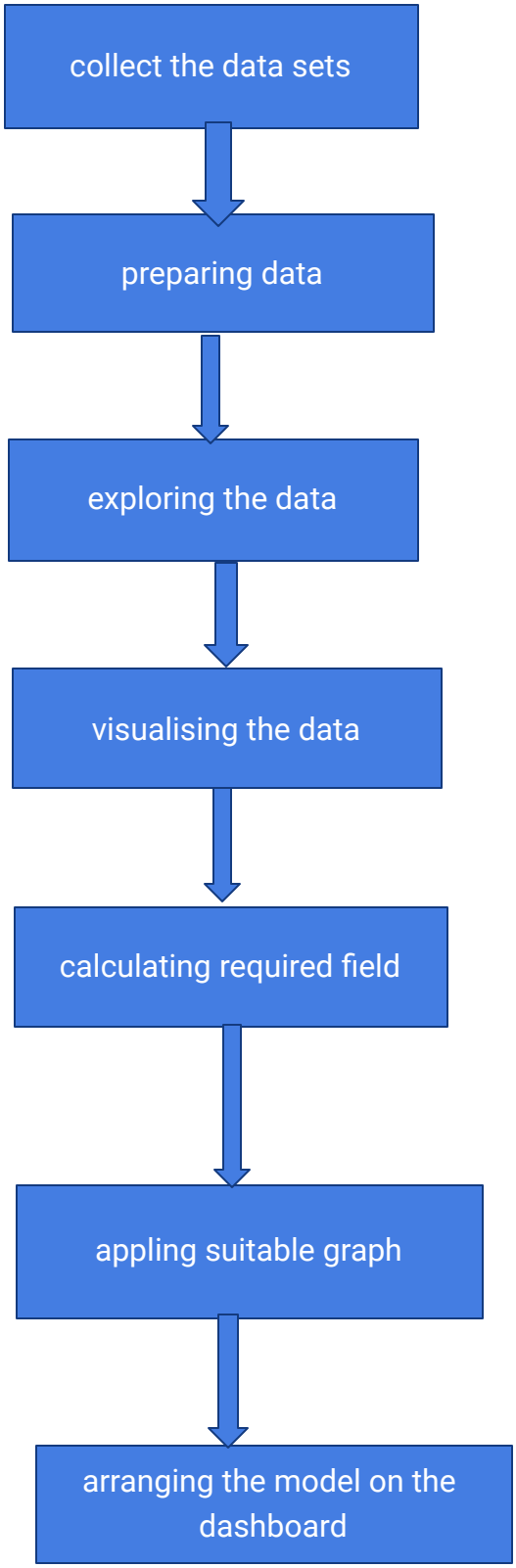
- A system installed with IBM Cognos
- System installed with Watson Studios
- IBM initiative account
- Download data sets
- Upload and prepare data
- Exploration of data

4. Experimental Investigation

Once the data sets are uploaded into IBM cognos, the fields which are required are explored. Cognos provide a feature called prepare data where on can customize the tables which ever are required. Number of data sets are taken to process the data. On successful preparation of data we need to explore the data to get insights of data. Cognos provide a feature which is known as exploration. This feature is use full to find the insights. It gives information about the dependencies of fields in the data.

The exploration feature in cognos provide the strength of dependencies of each field which directed us to explore more. This eventually provided underlying insights of the tournament. With the help of various graphs, the data exploration was made easy. The exploration which are use full in dashboard preparation are pinned. The total number of matches, total runs are calculated with help of datasets.

5. Flowchart


collect the data sets
preparing data
exploring the data
visualising the data
calculating required field
appling suitable graph
arranging the model on the dashboard

6. Result

6.1 Finding total Number of runs

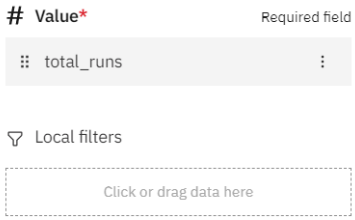


fig1: calculation for finding total number of matches

6.2 Total overs

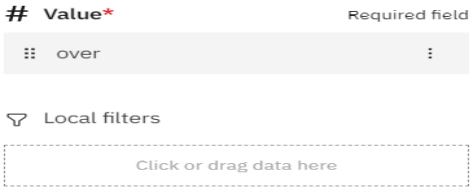


fig2: Calculation total number of overs

6.3 Total wickets

fig3: calculation for total wickets

6.4 Total matches

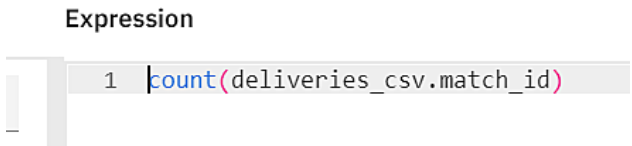


fig4: calculation for finding total number of matches

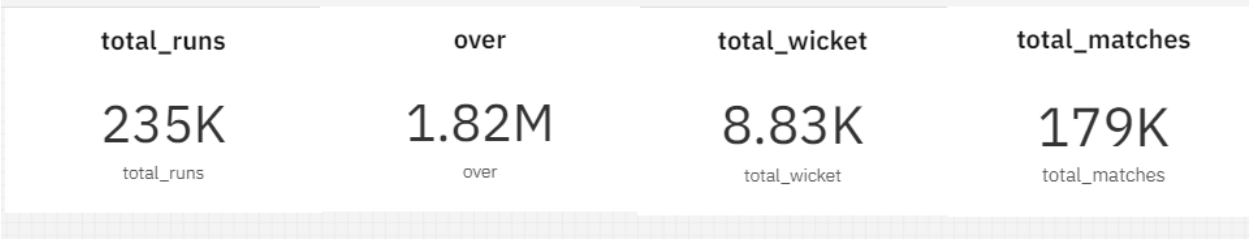


fig5: results of 6.1 - 6.4

6.5 Most winning teams

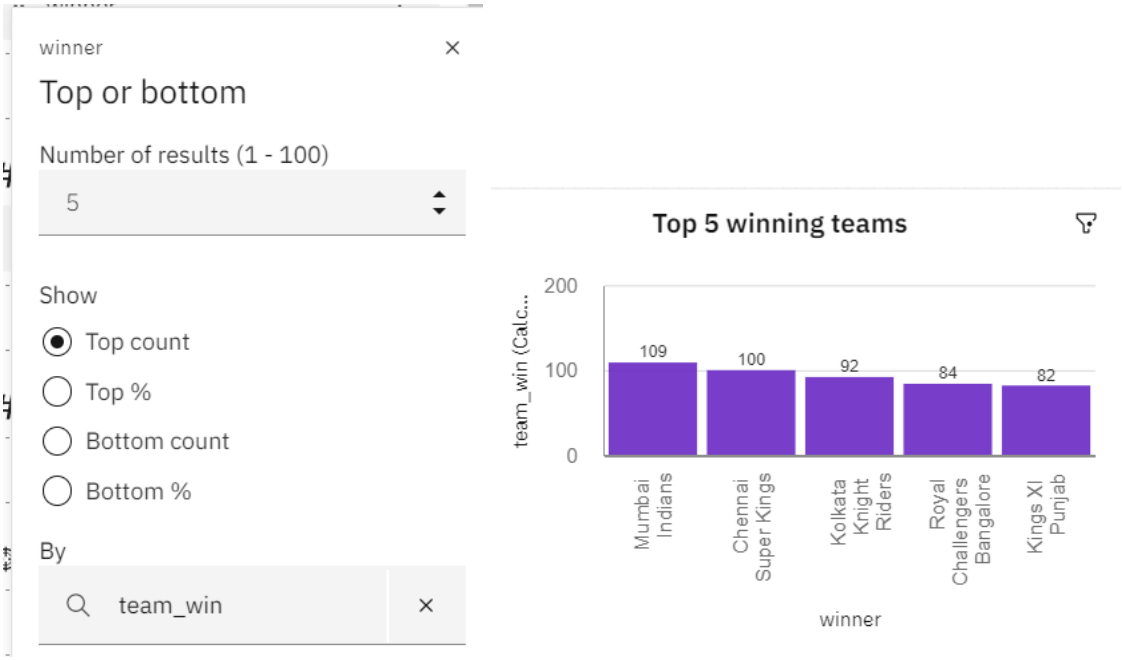


Fig6: Showing the teams with maximum number of victories

6.6 Least winning teams

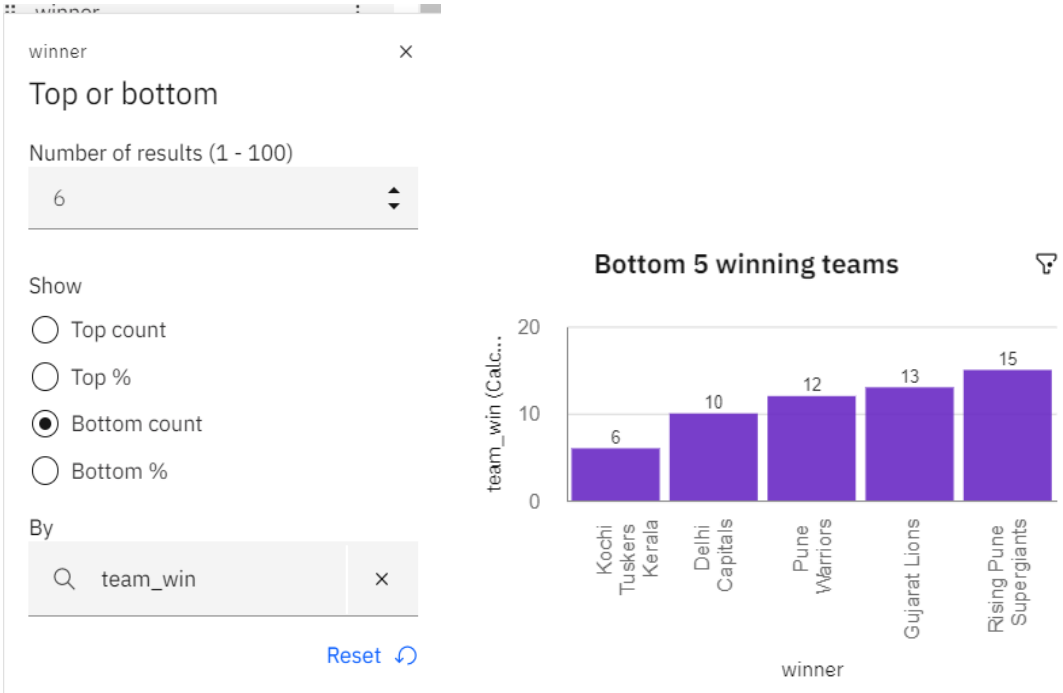


Fig7: Showing the teams with less number of victories. Since we are having no values so there is need to take plus 1 number of teams

6.7 Players with maximum number of match of the man awards and cities that hosted max number of matches

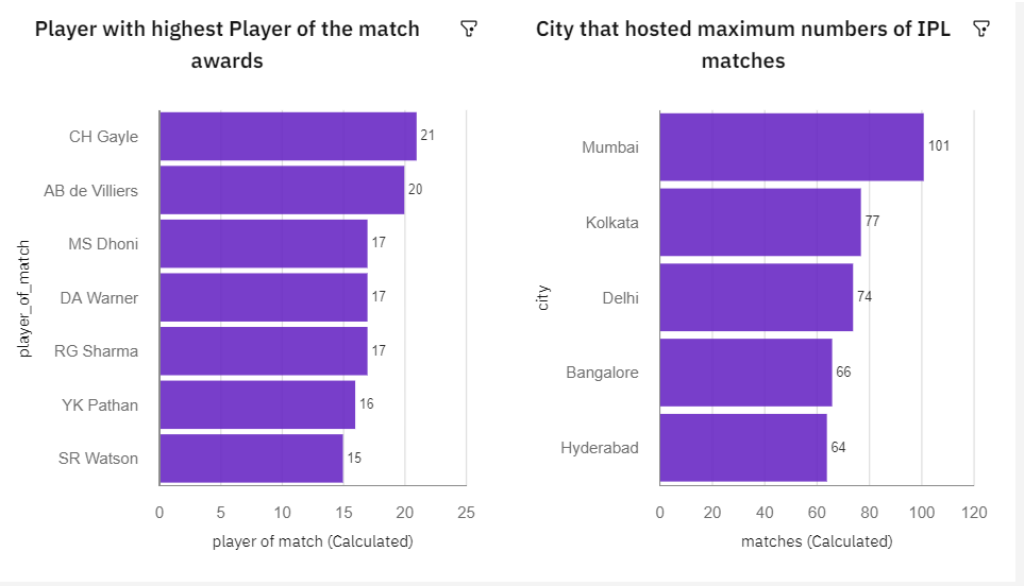


Fig8: players with highest number of man of the match award and cities that hosted max number of matches

6.8 On field umpires with maximum number of IPL matches

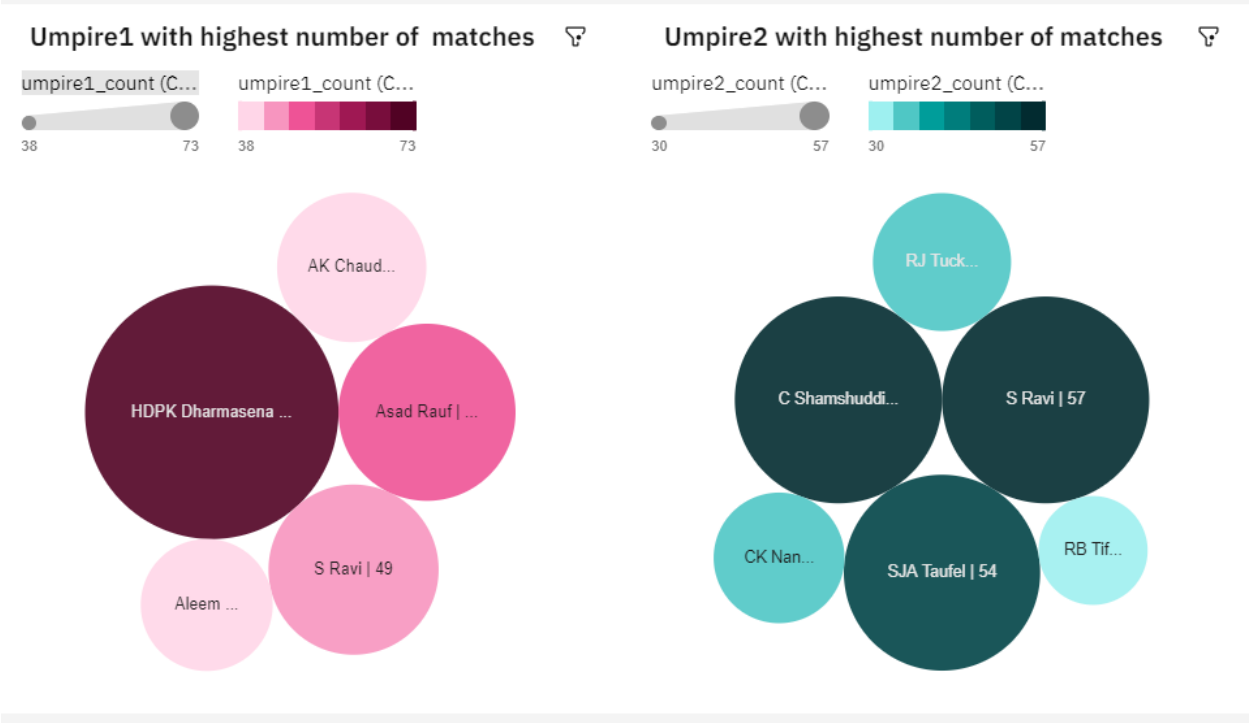


Fig9: On field umpires with maximum numbers of IPL matches

6.9 Teams with the their corresponding wins in each season

Fig10:Teams with the their corresponding wins in each season

6.10 Top teams which won while batting first and fielding first

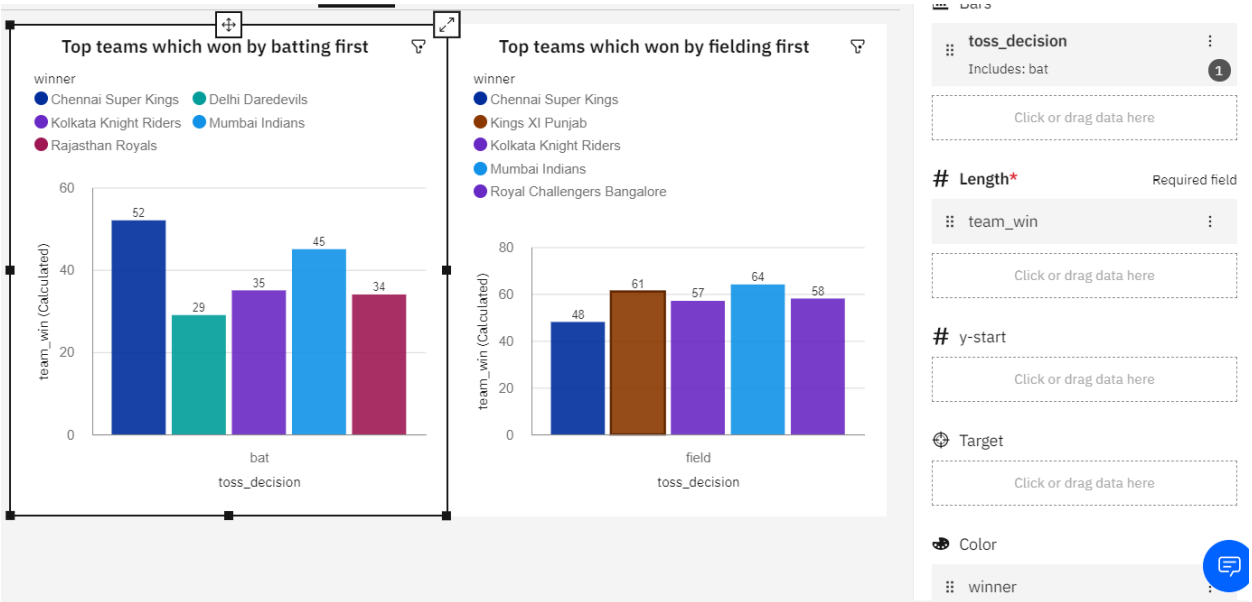


Fig 11:Teams with maximum victories while choosing either batting or fielding

6.11 teams which have won matches by most runs cumulatively

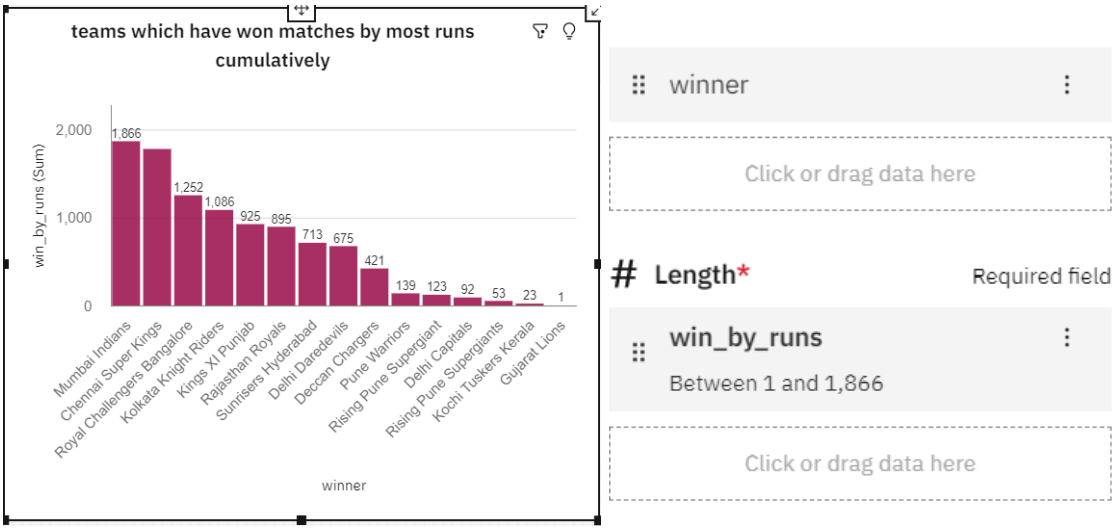


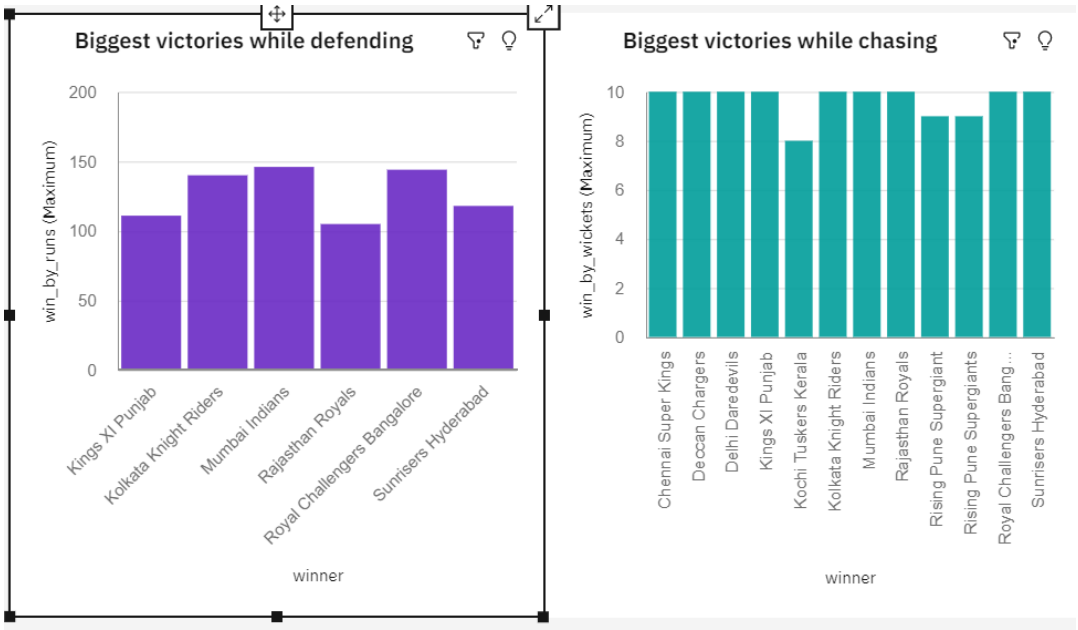
Fig12: teams which have won matches by most runs cumulatively

6.12 Probability of victory if toss won

Created a new table where a team won both toss and match and then calculated the total count of the match.

Fig13: Probability of victory if toss won

6.13 To find the biggest victories in IPL while defending a total and while chasing a total



7. Advantages

- Brief understanding of entire IPL trade off.
- Can understand underlying factors of victory.
- Graphical interpretation of huge data.
- Easy to understand .
- No Technical knowledge is required to understand the analysis.
- Filters helps to give some extra information on specifying certain value in the graph.
- Detailed information not only about the present IPL but also about the past matches.
- It also gives idea about geographical locations where the maximum matches are being hosted.
- No data base knowledge is required to get the required query output
- Easy to implement various queries

8. Applications

- Mostly preferred for IPL fans
- Can be used to predict the team which has higher probability of winning a match
- Can be used by cricket analytic professionals
- Non regular cricket follower can also use this dashboard to understand overall strengths of each team
- Can be used to identify the players with high scores, which helps organizers to get an understanding to invest on players.
- To clear myths and facts about the match
- Can be easily included in any commercial cricket applications such as crickbuzz

9. Conclusion

There many applications which are providing the data sets of IPL. But the data is not organized and hence the users are facing difficulty while analyzing the data. These tabulated data can be aggregated by only professionals. If common people wants to access such data they need to have technical knowledge. Which is limiting the common people to analyze cricket outcomes.

To overcome such issue, dash boards are provided with graphical interfaces. Even the people with no tech knowledge can understand by just having a glance. The huge amount of the data is processed and use full information is represented in graphs.

10. Future Scope

Since the data is growing exponentially we can have better accuracy in future for predicting the victory. Some of the aspects such as how the location is effecting the victory of the match is not covered. In this area, we have scope to improve the dashboard. Including this dashboard in the commercial apps, it helps to cover more aspects of the match. Which ultimately help app usage.

11. Bibliography

1. https://www.kaggle.com/nowke9/ipldata?select=matches.csv
2. https://towardsdatascience.com/exploratory-data-analysis-of-ipl-matches-part-1-c3555b15edbb
3. http://www.ijesrt.com/issues%20pdf%20file/Archive-2019/November-2019/5.pdf
4. https://keyskill-clms.comprehend.ibm.com/course/view.php?id=233
5. https://keyskill-clms.comprehend.ibm.com/course/view.php?id=237
6. https://keyskill-clms.comprehend.ibm.com/course/view.php?id=236