HACK CHALLENGE

# Project Title:

Super Predictor of Indian Premier League (IPL)

Team Name: Code Warrior

Name: M V Namitha

# **1 INTRODUCTION:**

As mentioned in the description of the problem statement, IPL has become the highest revenue-generating league cricket. So, Analyzing the data that have been storing the vast data of the IPL till now can help us predict the scores, winning team sometimes even the man of the match. It will be quite interesting to analyze the IPL game.

# **2 LITERATURE SURVEY**

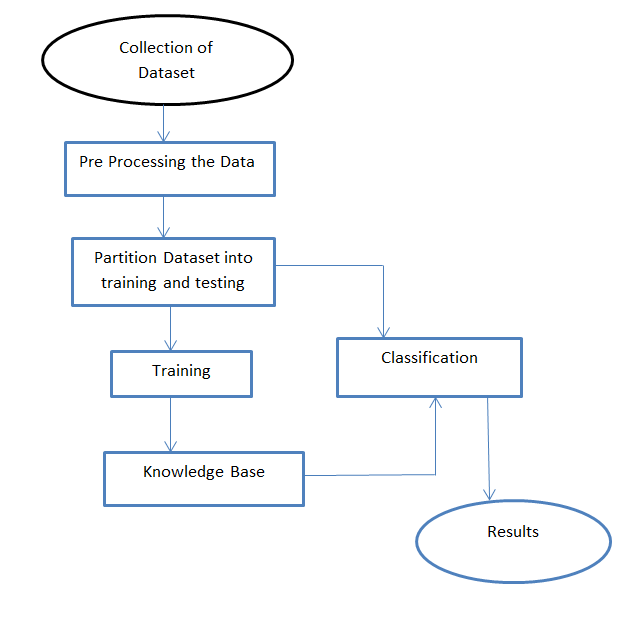
There are many sites for the analyzed data information but not a single page, so here the novelty is all the data in a single dashboard, using the cognos analytics

The objective of this solution is to create a dashboard that visualizes the following capabilities and also forecast the future results

1. To find the team that won the most number of matches in the entire IPL.
2. To find the team that lost the most number of matches in the entire IPL.
3. Does winning a toss increase the chances of victory.
4. To find the player with the most player of the match awards.
5. To find the city that hosted the maximum number of IPL matches.
6. To find the most winning team for each season.
7. To find the on-field umpire with the maximum number of IPL matches.
8. To find the biggest victories in IPL while defending a total and while chasing a total.
9. Which team won the most matches while batting first
10. Which team won the most matches while batting second
11. List of teams which have won matches by most runs cumulatively

# **3 THEORITICAL ANALYSIS**

## **3.1 BLOCK DIAGRAM**



## **3.2 HARDWARE/SOFTWARE DESIGNING**

SOFTWARE REQUIREMENTS:

1. IBM COGNOS ANALYTICS
2. Jupyter Notebook
3. Datasets

# **4 EXPERIMENTAL INVESTIGATION**

While analysing the solution for this problem we have found some unwanted and some duplicate data that need to be processed in a manner such that to get accurate results. So, that we have removed some unwanted data also merged the duplicate data.In our.In our project there are such constraints which can affect the analysis such as

There are some IPL Teams which change their names at some seasons Eg:Delhi Daredevils, Royal Challengers Bangalore,Sunrisers Hyderabad etc...

There are some new IPL Teams which played for only few ipl seasons Eg: Kochi Tuskers Kerala etc..

There are also teams like Pune Warriors played for some season.

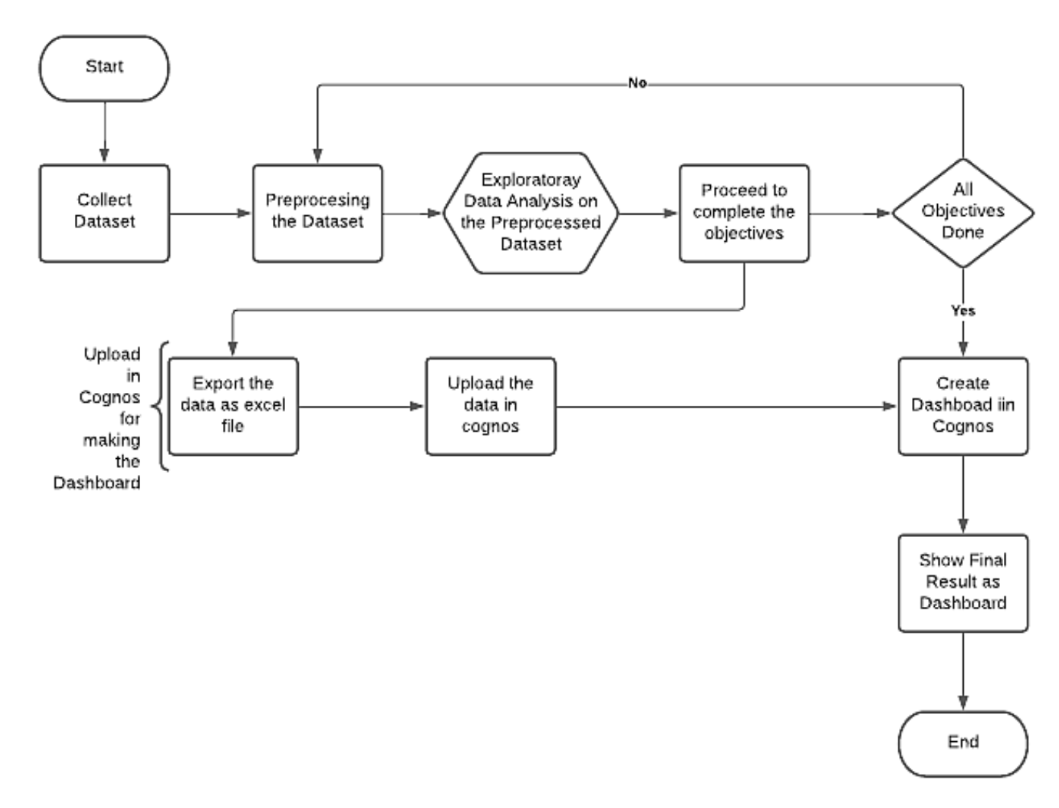
There are also some teams like Chennai and Rajasthan got banned for a couple of years.

There evolved 2 teams Gujarat Lions and Rising Pune Supergiants which were replacing the Chennai and Rajasthan.

These are some factors that may affect our analysis. So, we have sorted those issues extraction using "Pandas" library in python. data

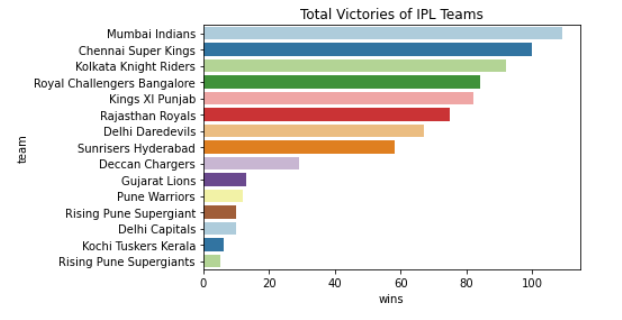
We also created a relationship with the attributes and constraints using the data sets which we obtained from the data extraction process. As every attribute and constraints are related to each other it is easy to figure out the solutions for the required problem statements.

# 5 FLOW CHART

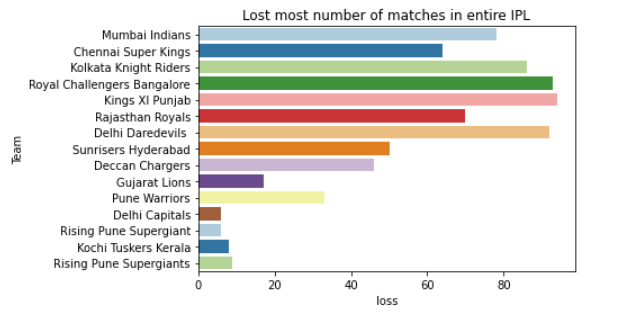
****

# **6 RESULT**

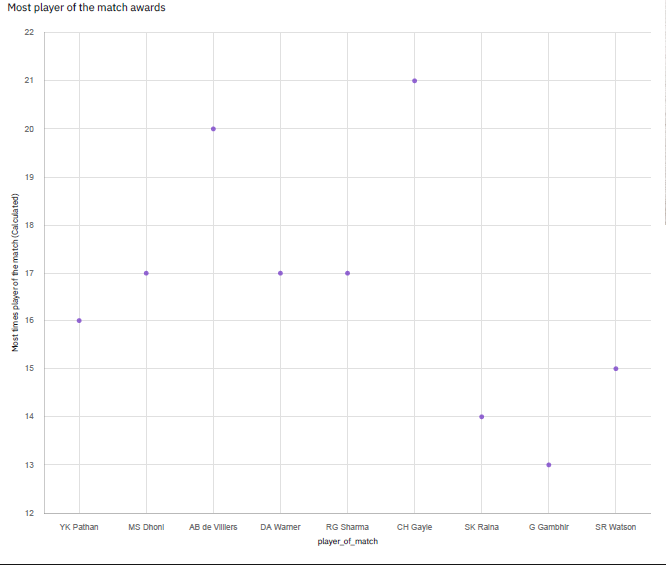
1. To find the team that won the most number of matches in the entire IPL.



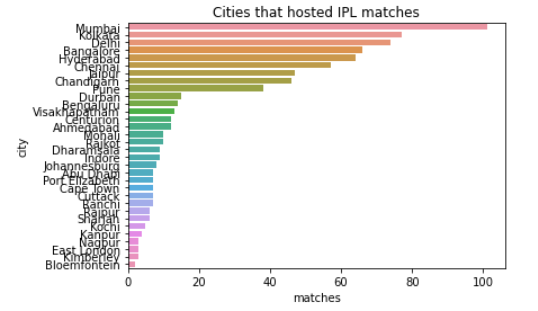
2.To find the team that lost the most number of matches in the entire IPL.



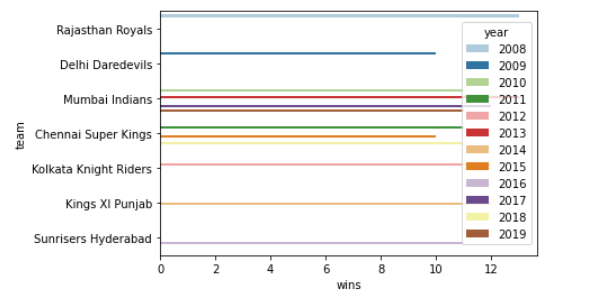
3.To find the player with the most player of the match awards.



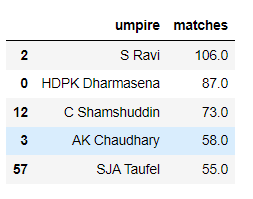
4.To find the city that hosted the maximum number of IPL matches.



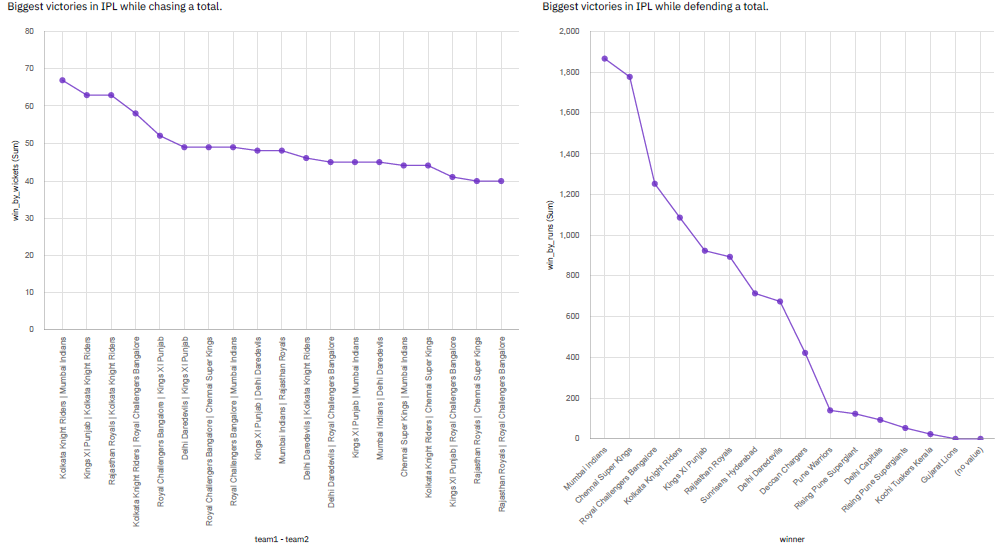
5.To find the most winning team for each season.



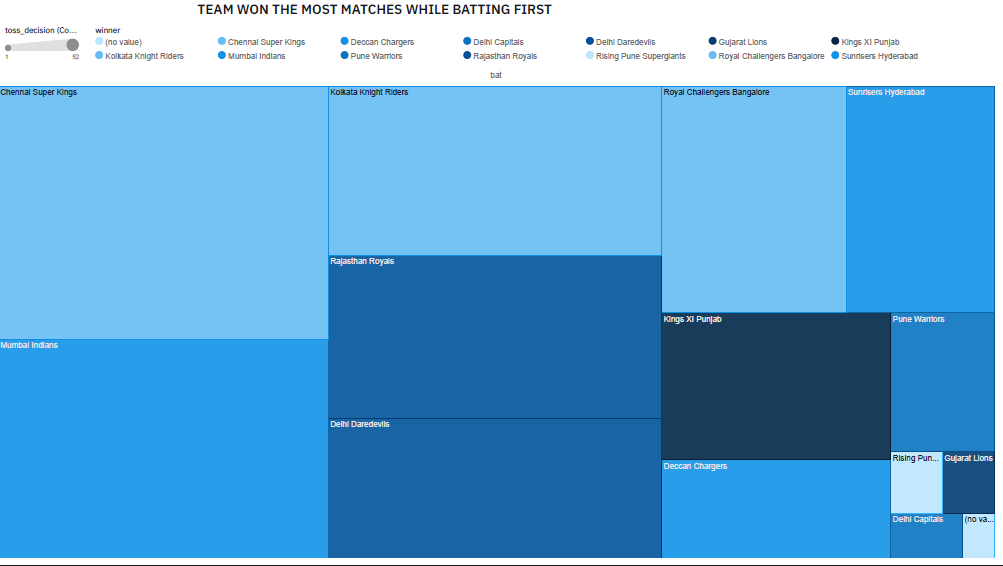
6.To find the on-field umpire with the maximum number of IPL matches.



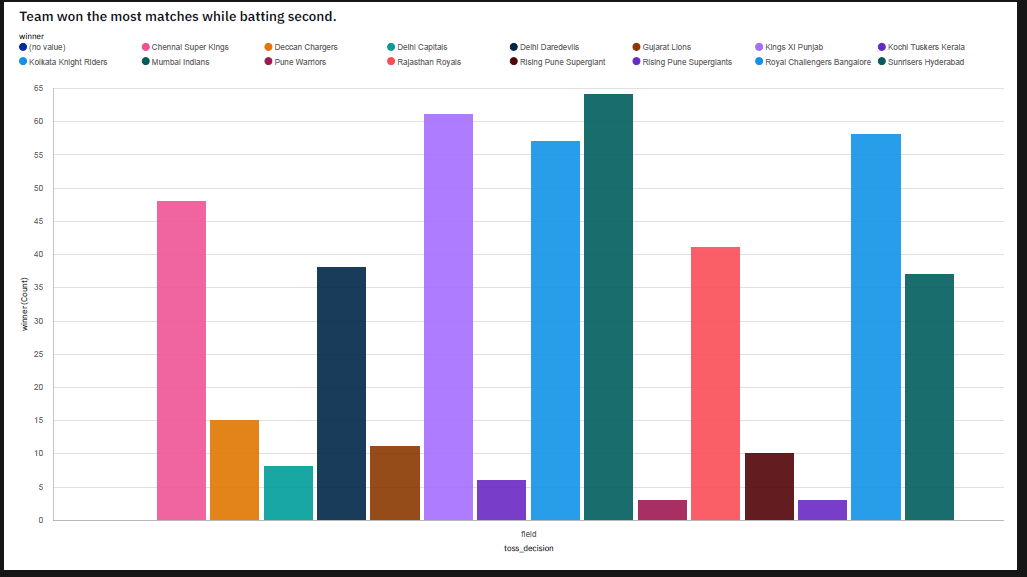
7.To find the biggest victories in IPL while defending a total and while chasing a total.



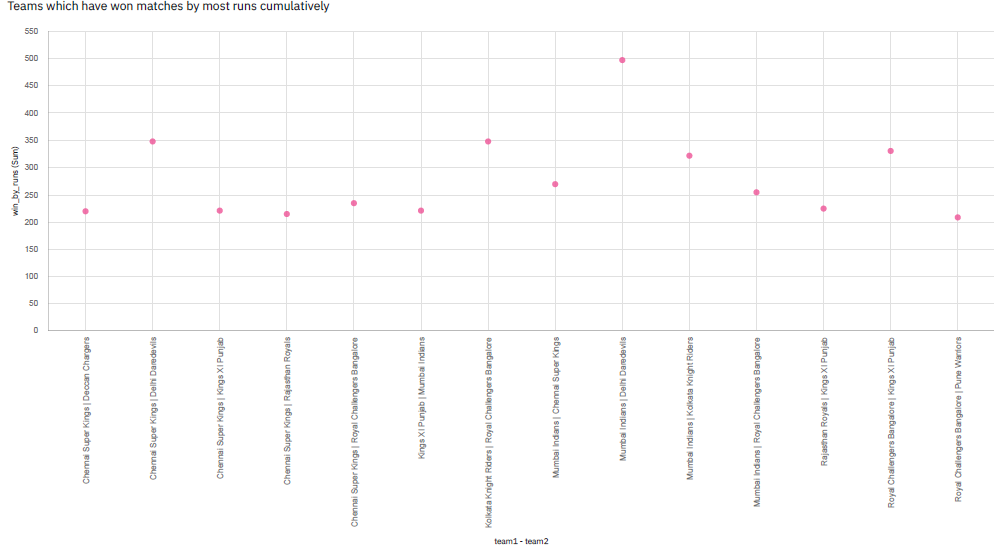
8. Which team won the most matches while batting first.



9.Which team won the most matches while batting second.



10. List of teams which have won matches by most runs cumulatively



# **7 ADVANTAGES AND DISADVANTAGES**

**ADVANTAGES:**

1. The records of upcoming IPL seasons can be managed in a cronological order.

2. Any upcoming problem statements can be sorted out easily.

3. More efficient analysis

4. Time complexity is less as data have been processed multiple times.

**DISADVANTAGES:**

1. Acquires more space because results for different problems store the attributes and constraints separately.

2. Any changes in teams or updates in teams may affect the original datasets.

3. Many duplicate values can affect the results.

4. Cleaning of data sets must be done in each update to avoid false results and to increase the accuracy the results.

# **8 APPLICATIONS**

1. By using the dashboard we can predict the winners of the match before it started, it is called predictive analysis.
2. By using the model we created which can predict the runs
3. We can make use of these predictions in dream 11 or many platforms.

# **9 CONCLUSION**

1. Dashboard created with the IPL History of Indian Premier League(2008-2019)

# **10 FUTURE SCOPE**

In future many more enhancements can be made in this analysis. We may expect a IPL Predictor which is useful to acquire the results with the few known inputs, it can be achieved in several ways like web scraping or machine trained bots which helps to predict the outcome by the analysed data sets. Web scraping is also a technique which fetches the data from different websites and bounded into single data sets and processed. The processed data can be trained using machine language and the chat application bot can be created using the entities and constraints for the problem statements and the bot is trained to respond to the requests of users expectation, This can be achieved in the upcoming days.

# **11 BIBILOGRAPHY**

1. https://www.kaggle.com/nowke9/ipldata?select=matches.csv

2.<https://www.ibm.com/docs/en/cognos-analytics/11.1.0?topic=stories-get-started-dashboards>

3.<https://www.iplt20.com/stats>