# PROJECT REPORT ON

## "SUPER PREDICTOR IPL"


Submitted by:

**Adarsh Aditya**

**Swasti Shree**

Under the guidance of-

**Shivam Shivhare**


JULY,2021


**IBM Hack Challenge - 2021**

Powered by

**SMARTINTERNZ**

In collaboration with

**IBM University Relations & AICTE**

# ACKNOWLEDGEMENT

We deem it a pleasure to acknowledge our sense of gratitude to our project mentor Mr.Shivam Shivhare under whom we have carried out the project work. His objective guidance and timely advice encouraged us to complete the project successfully.

We wish to reciprocate in full measure the kindness of Mr. Hemant who has always helped in overcoming the hurdles and technical issues in the journey.

We shall of course remain grateful to Smartbridge, Smartinternz and IBM community who provided all the necessary platforms and tools and opportunity for this amazing experience.

Finally we would like to show our gratitude to our parents and friends, who encouraged us throughout the journey. They provided whatever the help they can. We heartily thanks each and everyone, who have made our dreams into reality.

Thanks!!

# ABSTRACT

IPL is most awaited time for its fans. It attracts many people, companies and sponsors over the years. As the time passes, IPL business and revenues are increasing rapidly.Its popularity and fan base is all over the world. So, it is generating enormous amount of data which need to be analyzed for future benefits.

In this project we have-

- Gathered the data
- Filtered data
- Visualized the data
- Analyzed the data

We have done all the process of our project using IBM Cognos Analytics with its amazing features and options. We designed a Dashboard using Cognos Analytics.

A dashboard is a type of graphical user interface which often provides at-a-glance views of key performance indicators (KPIs) relevant to a particular objective or business process. In other usage, "dashboard" is another name for "progress report" or "report" and considered a form of data visualization.

It is used for information management and business intelligence.

# CONTENTS

# INTRODUCTION

## 1.1.OVERVIEW

Basically we have built a dashboard that will display the visualized data of IPL from 2008-2020. So, let us first understand what a dashboard is.
Dashboard in simple words is end to end data visualization capabilities.
A dashboard is a tool that provides a centralized, interactive means of monitoring, measuring, analyzing, and extracting relevant business insights from different datasets in key areas while displaying information in an interactive, intuitive, and visual way.

Process to create a dashboard-

1. Upload data
2. Create template
3. Exploration
4. Visualization
5. Dashboard Presentation

Importance of Dashboard-

- Enhanced visibility
- Time saving efficiency
- Better forecasting
- Key performance indicator
- Real time customer analytics
- Better decision making
- Enhances productivity

I hope now you have understood all about dashboard.
In making a good dashboard, **exploration** is an important part. If you learn it properly, then it become very easy to go for dashboard. In exploration, we explore the data set according to our requirements. We see our questions to be displayed on dashboard, then think which graph will be well suited for that and what will be its parameters to be taken. Now coming to dashboard we have to just take care of better formats, styling, colors and layout, which would be user friendly.

## 1.2.PURPOSE

Our project's aim is to build a user-friendly dashboard of IPL matches for better visualization and forecasting the data.

# LITERATURE SURVEY

## 2.1. EXISTING PROBLEM

Most of the data available on the internet today about the IPL are in the form of simple graphs, excel sheets, written data or in tabular format.

We don't have dynamic visualizations in dashboards.

Much useful and accurate data is lacking.

## 2.2 PROPOSED SOLUTION

We have build dynamic visualizations in dashboards with very interactive interface, user-friendly dashboard, data and graphs are very easily understandable.

# SOFTWARE REQUIREMENTS

## 1. IBM CLOUD

## 2. IBM COGNOS ANALYTICS


## EXPERIMENTAL INVESTIGATION

First we read and understood the data set for the problem. Then we started understanding the questions and looking for their answers. We analyzed each graphs and figures to have proper insights about them.Below are the question wise brief explanation:-

1.To find team that won most number of matches:- Here we took parameters like name of each teams and number of wins in a column graph. MI with 120 wins is the required team.

2.To find team that lost most number of matches:- Here we took parameters like name of each teams and no. of loose in a column graph. RCB with 103 loose is required team.

3.Does winning toss increases chances of victory:-Here we took pie chart, where 51.2% is the required figure.

4.To find most player of the match:- Here we took no. of matches and player of the match as parameter in column graph.A B de Villiers player of match 23 times.

5.City that hosted maximum matches:-Here we took no. of matches and names of cities as parameter in column graph. Mumbai hosted maximum matches ie 101.

6.Most winning team in each season:-Here we took no. of wins and seasons as parameter in line graph. 2008-RR,2009-DD,2010-MI,2011-CSK,2012-KKR, 2013-MI,2014-KXIP,2015-CSK,2016-SRH,2017-MI, 2018-CSK,2019-MI,2020-MI are required results.

7.On field umpires with maximum matches:-Here we took names of umpires and no. of matches as parameters in column graph.S Ravi with 118 matches is required result.

8.Biggest victory in defending & chasing:-We took line & column graph, where winners,win by wickets & win by runs are the parameters.MI got biggest victory in defending

(146 Runs) & MI,CSK,KKR,RCB,DC,RR,DD,KXIP,SRH got biggest victory in chasing(10 Wickets).

9.Most wins by batting first:- Here we took no.of matches won and teams&toss decision as parameters in bar graph. CSK is required team with 51 wins.

10.most wins by batting second:-Here we took no.of matches won and teams&toss decision as parameters in bar graph.KKR is required result with 64 wins.

11.Teams won matches with most runs cumulatively:- Here we took name of teams& win by runs as parameters in column graph.MI is result with 2111 win by runs.

## ADVANTAGES & DIADVANTAGES

Plus points:- The major positive side of our project is that our all graphs and visualization are dynamic, not static. It means when you click or roll mouse over any values it will shows results related to it. Also, our graphs are easy to understand and doesn't require any coding.

Minus points:- Some of the lacking were more and useful data was needed. We couldn't do prediction in all visualization.

## APPLICATIONS

Our solution can be applied in IPL matches, for better understanding of large data easily through visualizing them on different required graphs. Also, it will help to predict future results based on current data.

## FUTURE SCOPE

Users and viewers can see these on dashboard of IPL too, or can be displayed on screens during matches.More designed and visualization and predictions can help in IPL business and revenue generations. Or we can make our dashboard as a live website or in form of app, where it will automatically update with live data.

## SUMMARY

First we want to tell about how did we get this project. We got the mail about IBM Hack Challenge from our college. We got excited as it was having the brand name IBM and also was free of cost. So, thus we register for it and our journey started.

Our plan was to develop a dashboard which will visualize and forecast the data of IPL.

By the help of this project, we can have the insights in forecasting match results, scores, wicket takers, best players etc.

Many years ago, there was no such vast amount of data and real time prediction were made. No much viewers were associated with it.So, it was difficult to analyze data. But as time passes and technology enhances, more and more data, more analysis, more predictions and forecasting came into light.

Data cleaning is another crucial aspect. We have tried to make our data valuable and informative, as represented in different graphs. We have tried to add easy and accurate graphs to understand well and user friendly interface.

## BIBILIOGRAPHY

1.https://www.kaggle.com/nulldata/begin-your-data-analysis-in-python-with-ipl-data?select=deliveries.csv

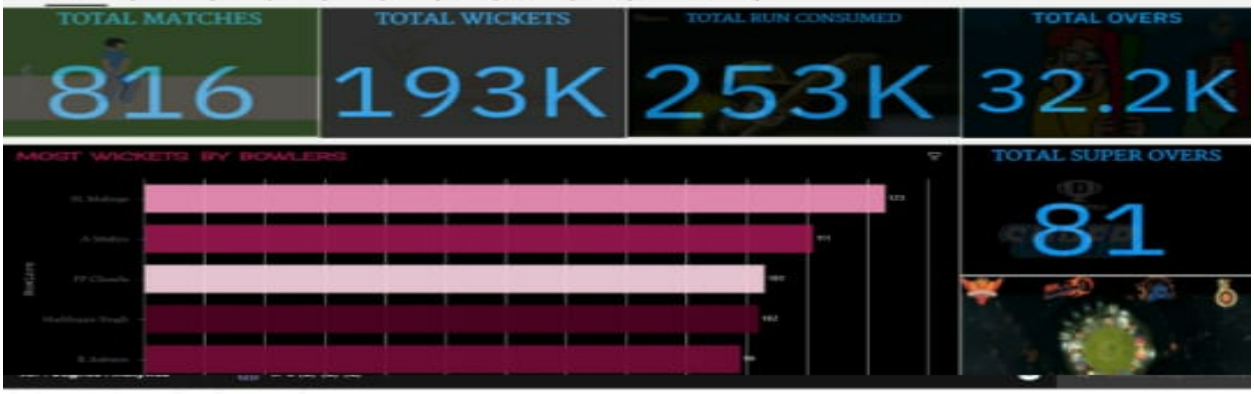
2.https://jovian.ai/bipinp2013/ipl-analysis/v/41?utm_source=embed

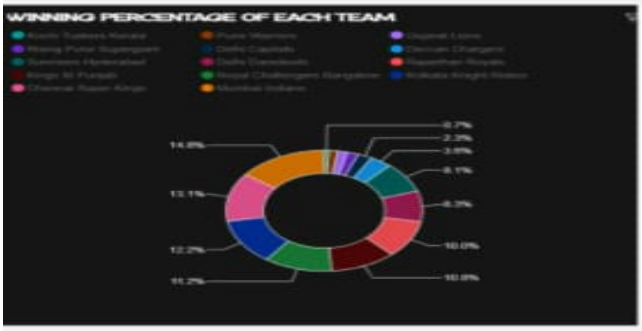
3.https://towardsdatascience.com/exploratory-data-analysis-of-ipl-matches-part-1-c3555b15edbb

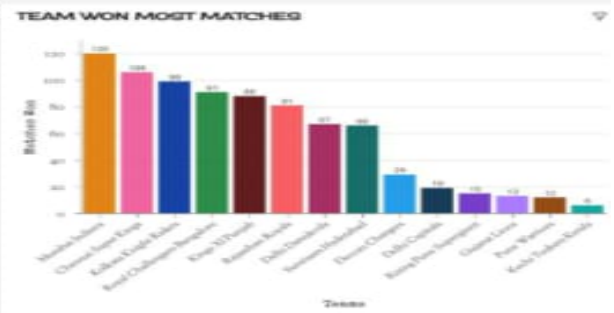
**VIDEO DEMONSTRATION--**

https://youtu.be/pnWbtu8a5RU

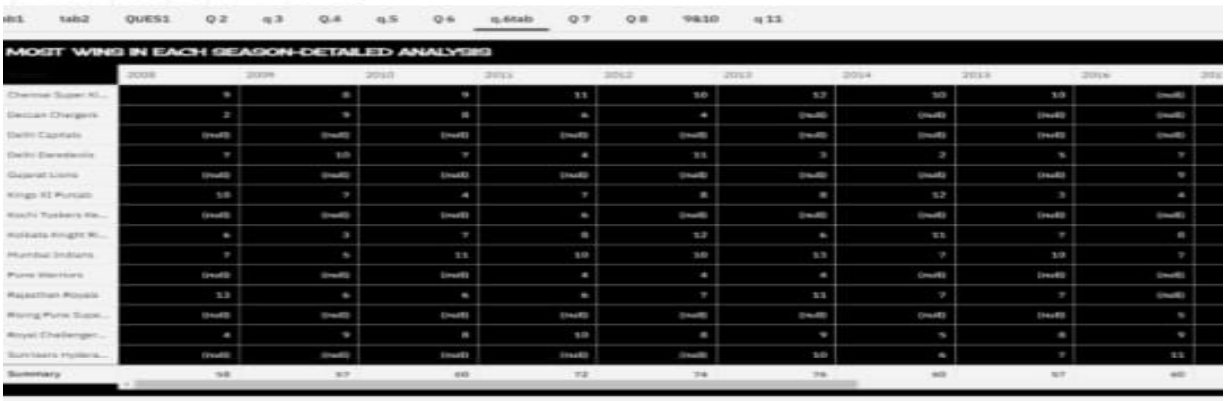

TOTAL MATCHES
816
TOTAL BATSMEN RUNS
240K
TOTAL CATCHES
5.74K
TOTAL OVERS
32.2K
MOST RUNS BY BATSMEN
IPL

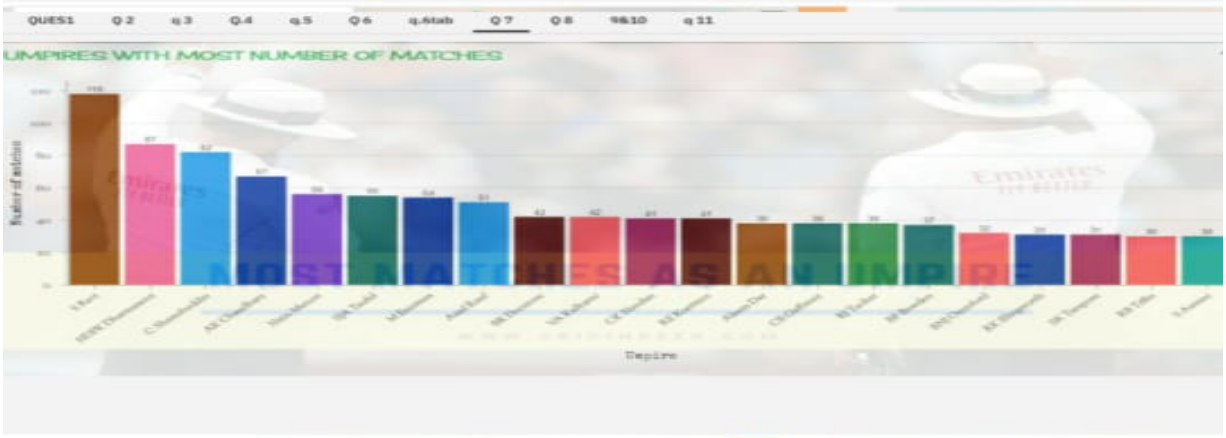

TOTAL MATCHES
816
TOTAL WICKETS
193K
TOTAL RUN CONSUMED
253K
TOTAL OVERS
32.2K
MOST WICKETS BY BOWLERS
TOTAL SUPER OVERS
81

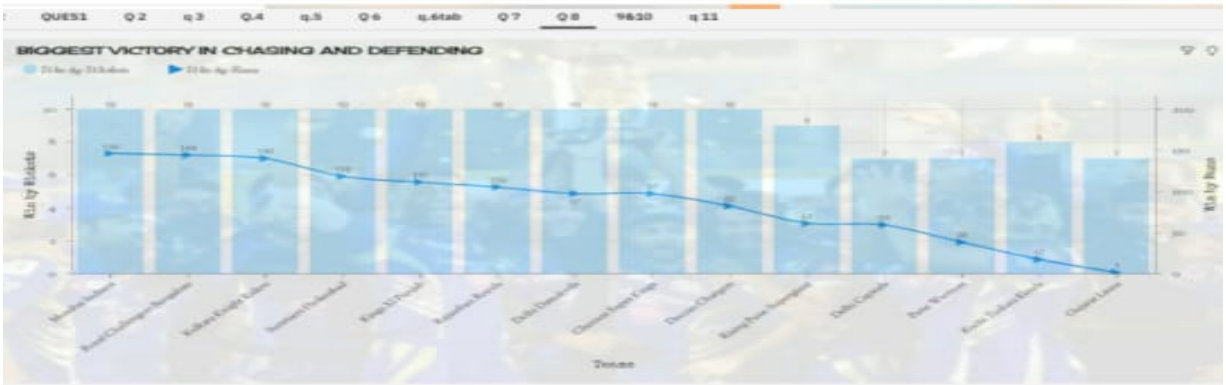

WINNING PERCENTAGE OF EACH TEAM


TEAM WON MOST MATCHES


IBM Cognos Analytics
TEAM LOST MOST MATCHES
LOSING PERCENTAGE OF EACH TEAM


tab1 tab2 QUES1 Q 2 q 3 Q 4 q 5 Q 6 q.6tab Q 7 Q 8 9&10 q 11
MOST WING IN EACH SEASON-DETAILED ANALYSIS


QUES1 Q 2 q 3 Q 4 q 5 Q 6 q.6tab Q 7 Q 8 9&10 q 11
UMPIRES WITH MOST NUMBER OF MATCHES
Umpire


QUES1 Q 2 q 3 Q 4 q 5 Q 6 q.6tab Q 7 Q 8 9&10 q 11
BIGGEST VICTORY IN CHASING AND DEFENDING
Team


QUES1 Q 2 q 3 Q 4 q 5 Q 6 q.6tab Q 7 Q 8 9&10 q 11
MOST WIN BY BATTING FIRST
MOST WIN BY BATTING SECOND


DEPENDENCE OF WINNING MATCH ON WINNING TOSS
GOOD TO LOSE THE TOSS?


PLAYER WITH MOST "PLAYER OF THE MATCH TROPHY"


5. CITY HOSTED MOST NUMBER OF MATCHES
HOW MANY CITIES, HOW MANY MATCHES?
VENUES HOSTED MAXIMUM MATCHES
MATCHES HELD IN DIFFERENT LOCATIONS


6. MOST WIN IN EACH SEASON


TEAM WON BY MOST RUN CUMULATIVELY