DATA SCIENCE PROJECT REPORT

(Project Semester August-December 2021)

IPL DATA 2008 - 2019

Submitted by

GOTTIMUKKULA KISHORE

Registration No: 11907817

Computer Science and Engineering

Section: KM004

Course Code INT217

Under the Guidance of

Sandeep Kaur

Discipline of CSE/IT

Lovely School of Computer Science and Engineering

Lovely Professional University, Phagwara

CERTIFICATE

This is to certify that Kishore bearing Registration no. 11907817 has completed Data Science project titled, **"IPL DATA 2008 - 2019"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

Signature and Name of the Supervisor

Designation of the Supervisor

School of Computer Science

Lovely Professional University

Phagwara, Punjab.

DECLARATION

I, GOTTIMUKKULA KISHORE, student of Lovely Professional University under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this assignment is based onmy own intensive work and is genuine.

Name of the student: G . KISHORE

Date:

Signature

Registration No. 11907817

ACKNOWLEDGEMENT

A project work is a combination of views, ideas, suggestions and contribution of many people. Thus, one of the pleasant parts of writing the report is to thank those who have contributed towards its fulfilment.

I consider it as great privilege to have esteemed Lecturer Ms. Sndeep Kaur as my project guide. I take this opportunity to express my sincere gratitude to her through constant advice and constructive criticism nourished my interest in the subject and provided a free and pleasant atmosphere to work against all odd situations. I avail this opportunity to extend my heart full thanks and deep respect to faculty member for their able guidance during this project.

My gratitude to all those, who **responded to my questionnaire** in a well-defined manner and helped me acquiring knowledge.

I would like to communicate a deep sense of gratitude to all these people without whom my project would not have been such a great learning experience.

G . KISHORE

KM004

Reg no: 11907817

Lovely Professional University

Table of Contents

Sr.No.	Particulars
1.	Introduction
2.	Scope of The Analysis
3.	Existing System
4.	ETL Process
5.	Analysis on Dataset
6.	List of Analysis with Results
7.	Future Scope
8.	References
9.	Bibliography

Introduction

The Indian Cricket League (ICL) was founded in 2007, with funding provided by Zee Entertainment Enterprises. The ICL was not recognised by the Board of Control for Cricket in India (BCCI) or the International Cricket Council (ICC) and the BCCI were not pleased with its committee members joining the ICL executive board. To prevent players from joining the ICL, the BCCI increased the prize money in their own domestic tournaments and also imposed lifetime bans on players joining the ICL, which was considered a rebel league by the board.

On 13 September 2007,on the back of India's victory at the 2007 T20 World Cup, BCCI announced a franchise-based Twenty20 cricket competition called Indian Premier League. The first season was slated to start in April 2008, in a "high-profile ceremony" in New Delhi. BCCI vice-president Lalit Modi, who spearheaded the IPL effort, spelled out the details of the tournament including its format, the prize money, franchise revenue system and squad composition rules. It was also revealed that the IPL would be run by a seven-man governing council composed of former India players and BCCI officials and that the top two teams of the IPL would qualify for that year's Champions League Twenty20. Modi also clarified that they had been working on the idea for two years and that the IPL was not started as a "knee-jerk reaction" to the ICL. The league's format was similar to that of the Premier League of England and the NBA in the United States.

In order to decide the owners for the new league, an auction was held on 24 January 2008 with the total base prices of the franchises costing around $400 million. At the end of the auction, the winning bidders were announced, as well as the cities the teams would be based in: Bangalore, Chennai, Delhi, Hyderabad, Jaipur, Kolkata, Mohali, and Mumbai. In the end, the franchises were all sold for a total of $723.59 million. The Indian Cricket League soon folded in 2008.

Scope of The Analysis

Data Science / Analytics is all about finding valuable insights from the given dataset. In short, Finding answers that could help business.

Since usually such tutorials are based on in-built datasets like iris, It becomes harder for the learner to connect with the analysis and hence learning becomes difficult. To overcome this, The dataset that we use in this notebook is IPL (Indian Premier League) Dataset posted on Kaggle Datasets sourced from cricsheet. IPL is one of the most popular cricket tournaments in the world, thus the problems we try to solve and the questions that we try to answer should be familiar to anyone who knows Cricket.

Questions:

- How many matches we've got in the dataset?
- How many seasons we've got in the dataset?
- Which Team had won by maximum runs?
- Which Team had won by maximum wicket?
- Which Team had won by closest Margin (minimum runs)?
- Which Team had won by minimum wicket?
- Which Season had most number of matches?
- Which IPL Team is more successful?
- Has Toss-winning helped in winning matches?

Existing System

Before existence of Data Science, analyzing data used to be hectic task and existing system didn't used to analyses the data with perfection.
Without existence of current cutting-edge technology of data science, we can get actionable insights in the dataset of the Indian Trade.
Following are the benefits which weren't present in the existing system of data analyzing:

1. Making Better Decision with The Help of Data
2. Directing actions based on trends- which later defines the goals required for profit.
3. Doing challenging stuffs with the help of prediction which is done by data.
4. Identifying various opportunities to increase the profit,
5. Making decision with Quantifiable, data driven evidence so that loss doesn't happens.
6. Testing the decisions taken by the data and watching and analyzing the trend.

Source of The Dataset

- The dataset is taken from the Kaggle with the name 'IPL DATA 2008-2019'.
- https://www.kaggle.com/ramjidoolla/ipl-data-set
- Author of the Dataset
  Lakshya Agrawal
- Data last updated
  August 2019

Analysis of Dataset

HOME PAGE OF THE PROJECT:

1. **Year Wise No of Matches Held in Each Stadium :**

a) Introduction: The analysis shows the year wise data of the Matches held in each stadium.

b) Specific Requirements/Functions and Formulas:

i)Pivot table of Import Data

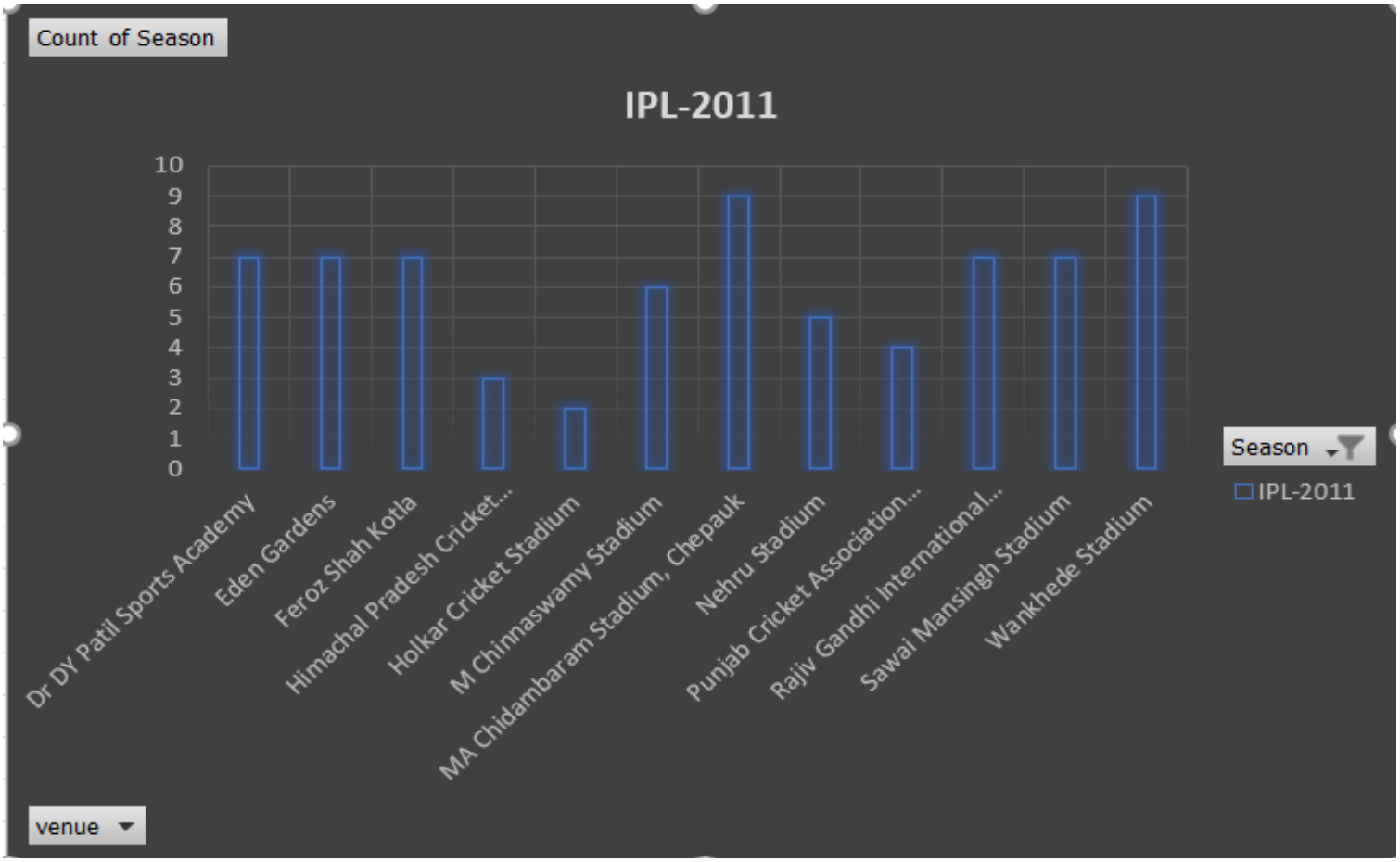
ii)Pivot table of Export Data

iii)SLICERS AND HYPER LINKS ARE USED

iv)Clustered Column Chart

c) Analysis Results:

we have observed that in every year mostly played ground is Wankhede Stadium


Count of Season
IPL-2011
10
9
8
7
6
5
4
3
2
1
0
Dr DY Patil Sports Academy
Eden Gardens
Feroz Shah Kotla
Himachal Pradesh Cricket...
Holkar Cricket Stadium
M Chinnaswamy Stadium
MA Chidambaram Stadium, Chepauk
Nehru Stadium
Punjab Cricket Association...
Rajiv Gandhi International...
Sawai Mansingh Stadium
Wankhede Stadium
Season
IPL-2011
venue

2. Year Wise Team Performance:

a)Introduction: The analysis shows the year wise Team performance in IPL .

b)Specific Requirements/Functions and Formulas:

i)Pivot table of the data of Import

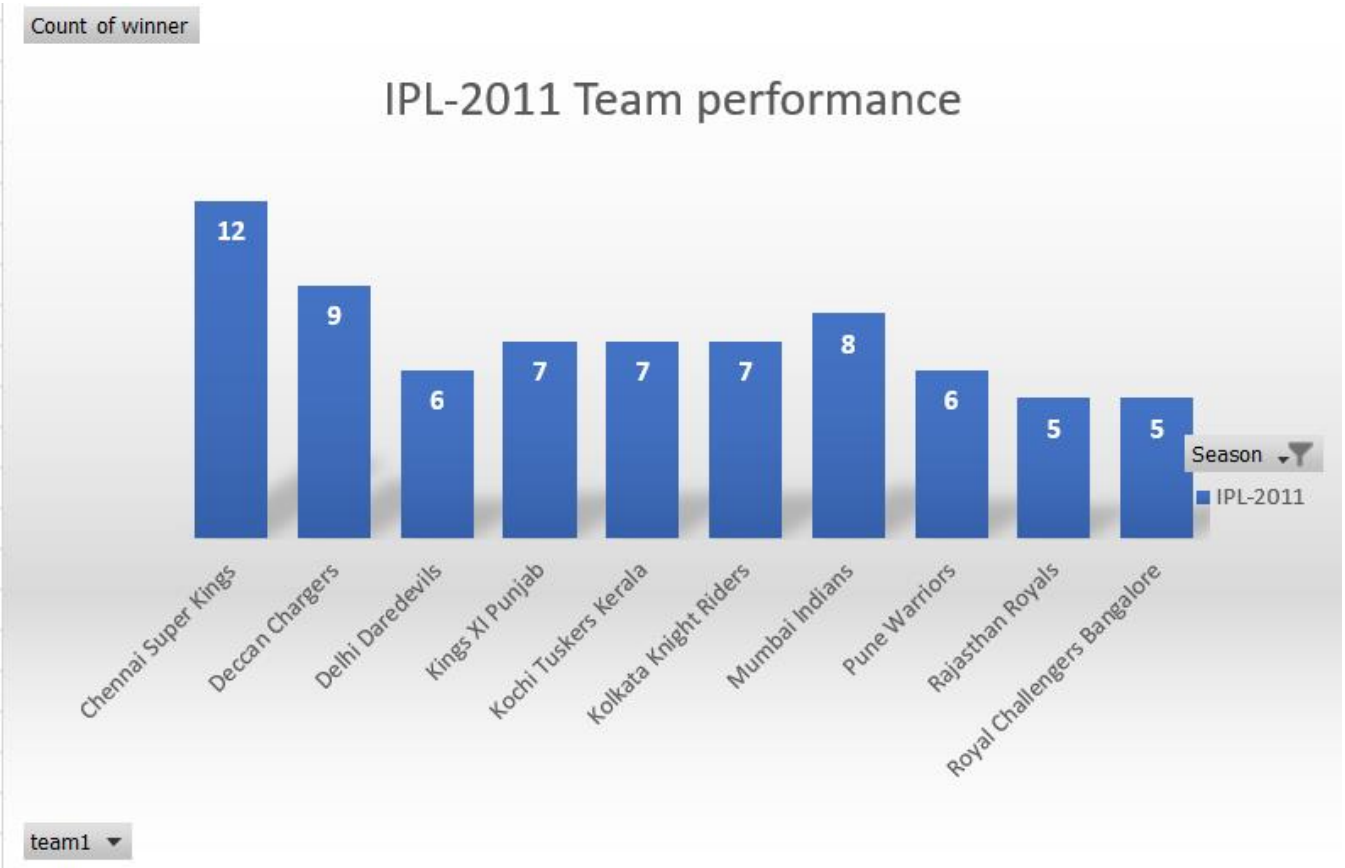
ii)With the help of data Clustered Bar chart is plotted.

iii)Slicers are used to change the year

c)Analysis Results:

By change the years we can observe that Chennai super kings and Mumbai Indians were Performing consistently

d)Visualization:



3. **Year wise Team win by runs:**

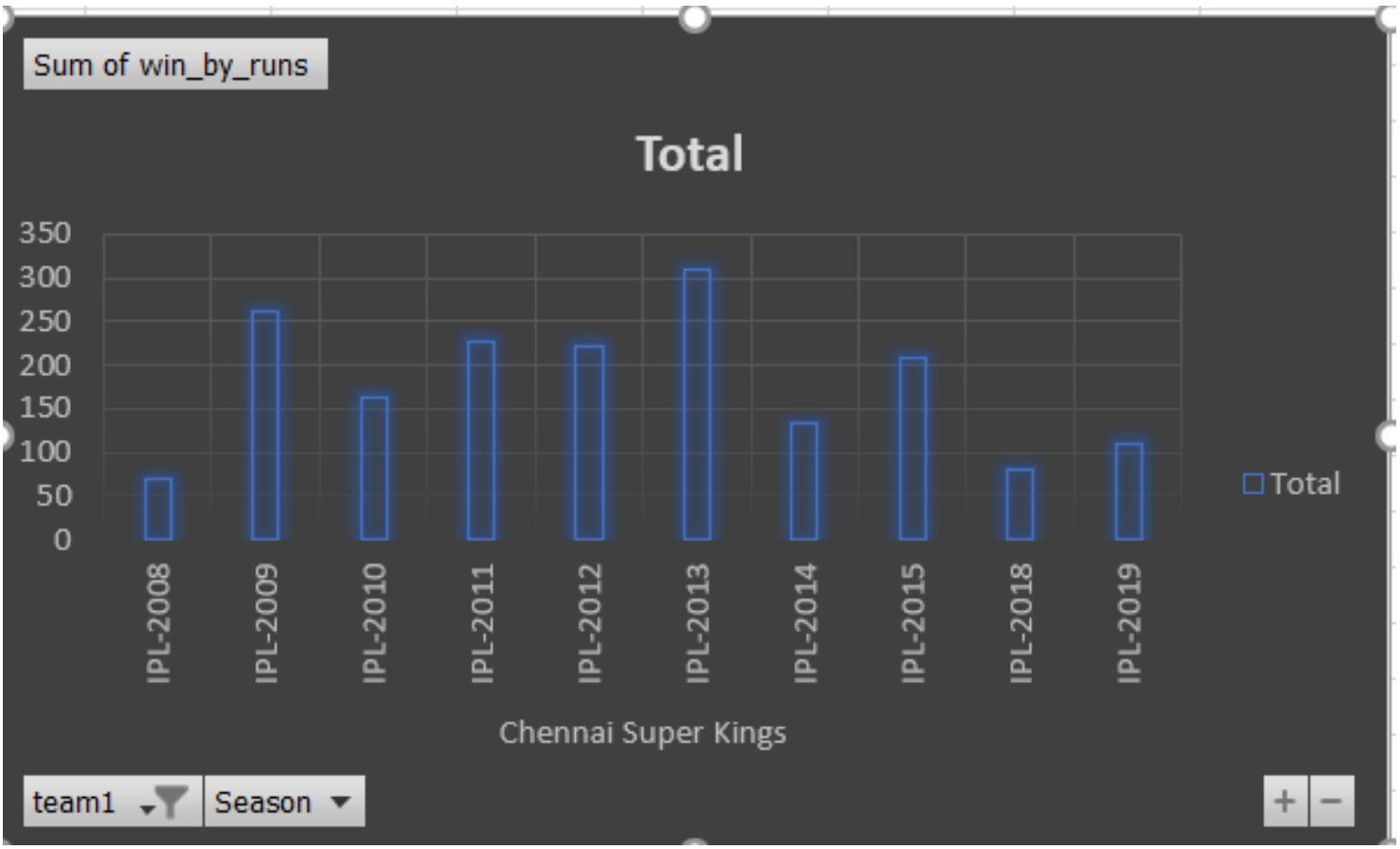
a)Introduction: The analysis shows the Year wise team win by runs in the overall year match

b) Specific Requirements/Functions and Formulas:

- Pivot table of the data of RUNS
- With the help of data Clustered Bar chart is plotted.
- With the help of Slicers we can change the year

c)Analysis Results:

- By bar graph we can observe that Chennai super kings have won many matches in different years with highest runs.



4. Year wise teams win by Wickets:

a)Introduction: The analysis shows the Team wise win by wickets in the Indian premier league

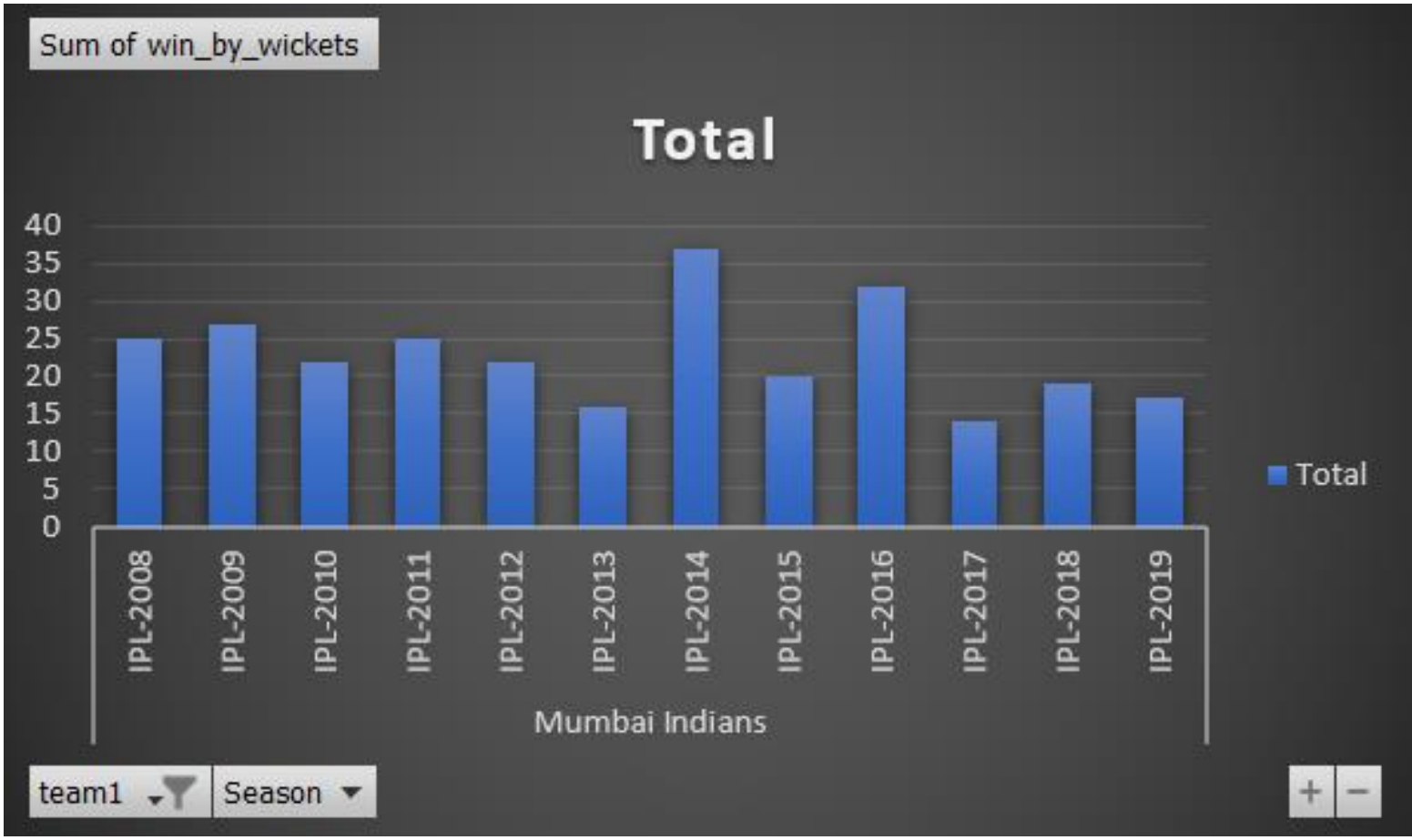
b)Specific Requirements/Functions and Formulas:

i)Pivot table of the data of wickets and year

ii)With the help of data Clustered Bar chart is plotted.

c)Analysis Results:

- Mumbai Indians is the teams which wins the all the matches with highest no of wicket taking in different years.



5. Toss decision of the teams after winning the toss:

a)Introduction: The analysis shows the Team wise overall toss decision in IPL.

b)Specific Requirements/Functions and Formulas:

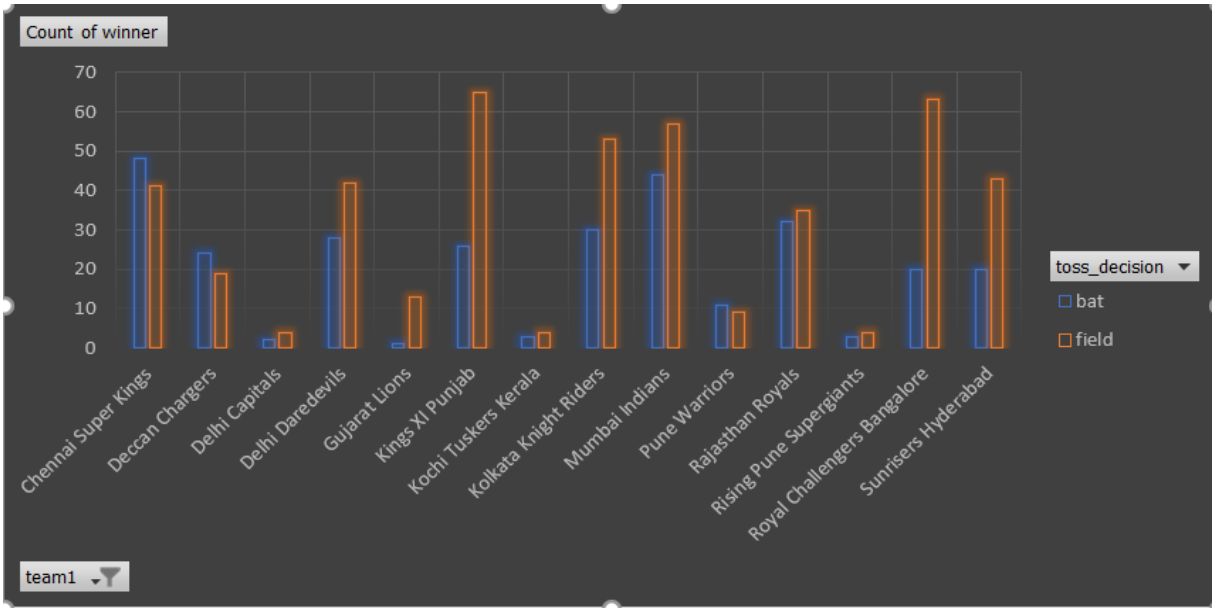
i)Pivot table of the data of Import.

ii)With the help of data Clustered Bar chart is plotted.

c)Analysis Results:

From the clustered graph we can observe that most of

teams are choosing field option after wining the toss



6. PERCENTAGE OF TOSS DECISION:

a)Introduction: The analysis shows the overall percentage of toss decision.

b)Specific Requirements/Functions and Formulas:

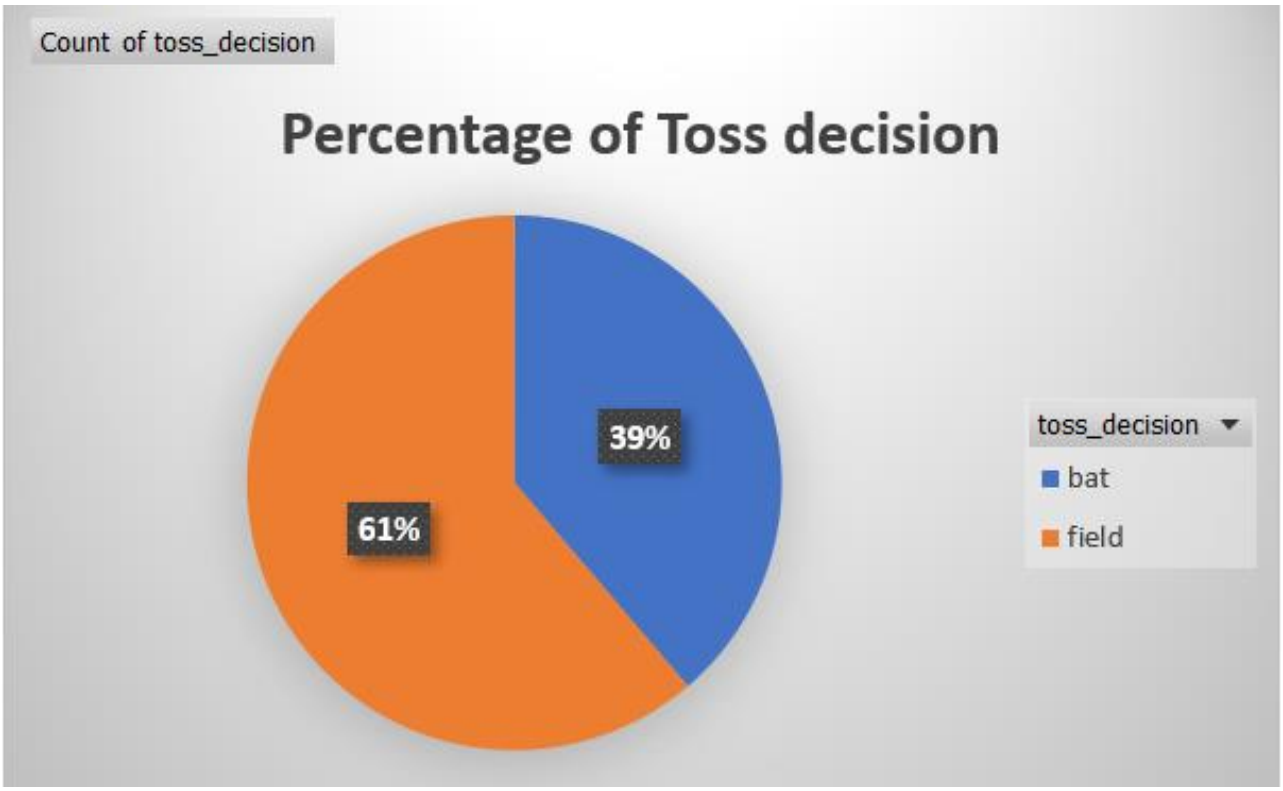
i)Pivot table of the data of toss decision

ii)With the help of data pie chart is plotted.

c)Analysis Results:

we can observe that most of teams prefer to take field

option after wining the toss



7. **Home win percentage of each team:**

a)Introduction: The analysis shows the Winning percentage of each team in their home ground

b)Specific Requirements/Functions and Formulas:

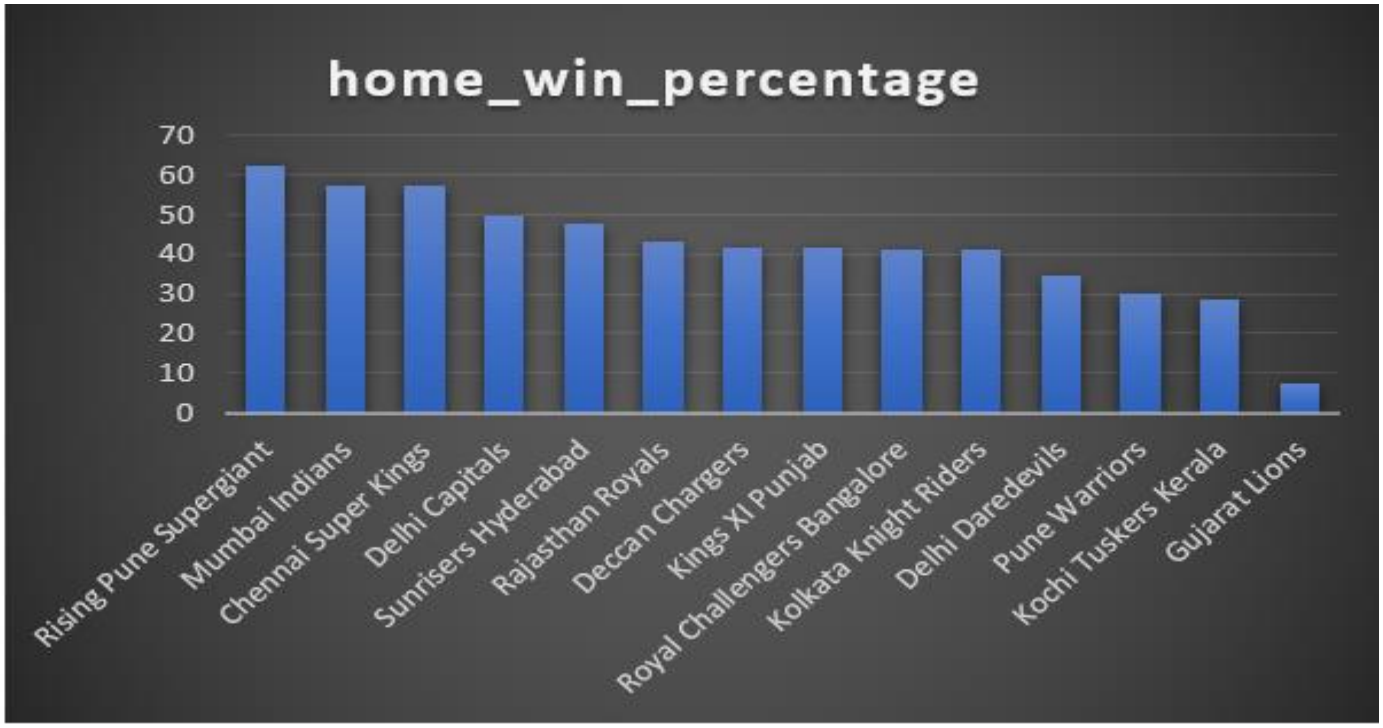
i)Pivot table of the data of Import

ii)With the help of data Clustered Bar chart is plotted.

c)Analysis Results:

- We can observe that the winning percentage of team is more in their home ground than other grounds

. d)Visualisation:



8. Away win percentage of teams:

a)Introduction: The analysis shows the Away win percentage of teams .

b)Specific Requirements/Functions and Formulas:

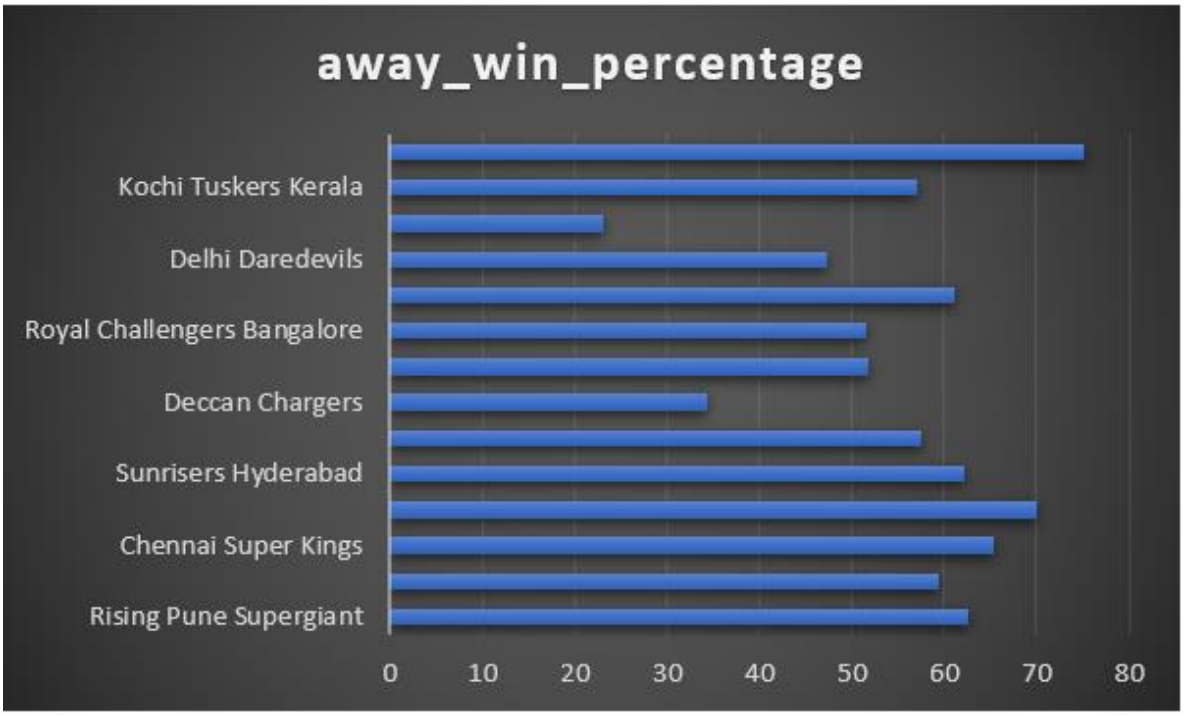
i)Pivot table of the data of Import

ii)With the help of data Clustered Column chart is plotted.

c)Analysis Results:

We can observe that away win percentage of every team

is low compared to home ground



9. Total DL'S Taken by all the Teams:

a)Introduction: The analysis shows the Team wise DL'S taken

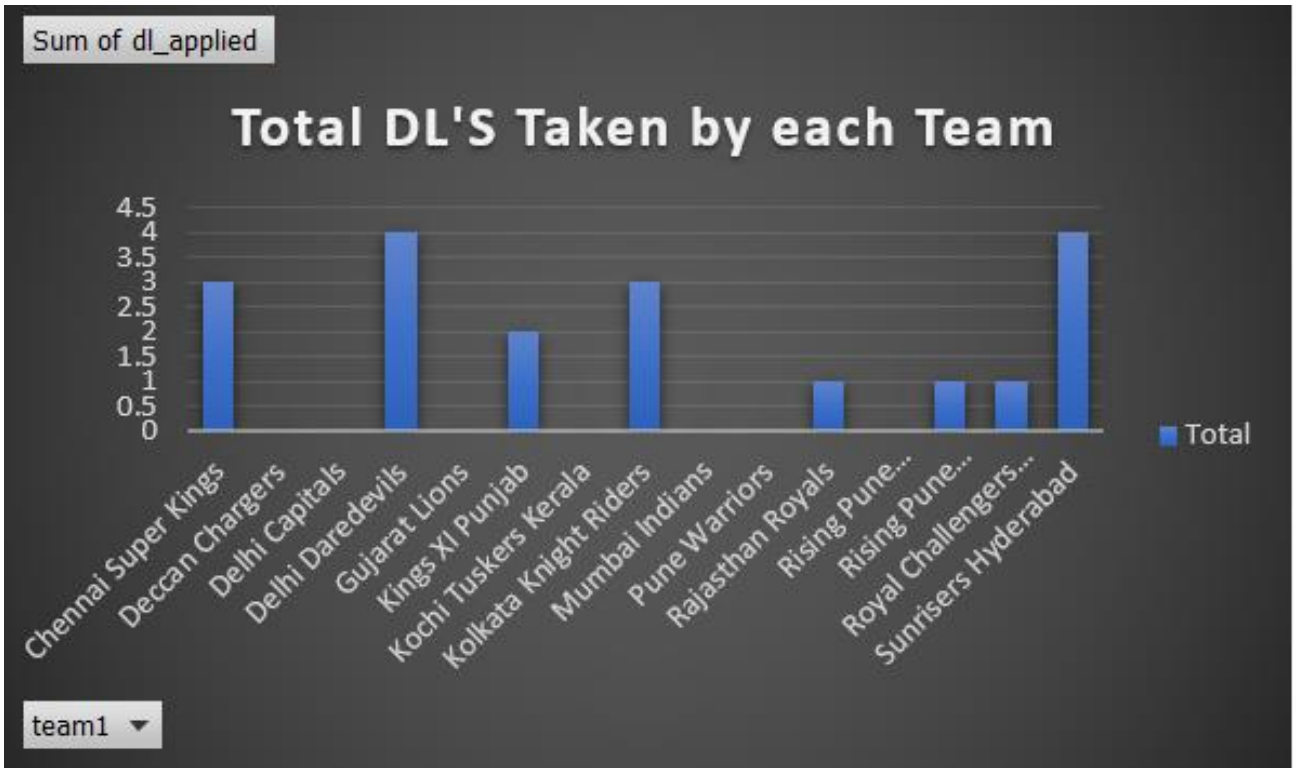
b)Specific Requirements/Functions and Formulas:

i)Pivot table of the data of Import

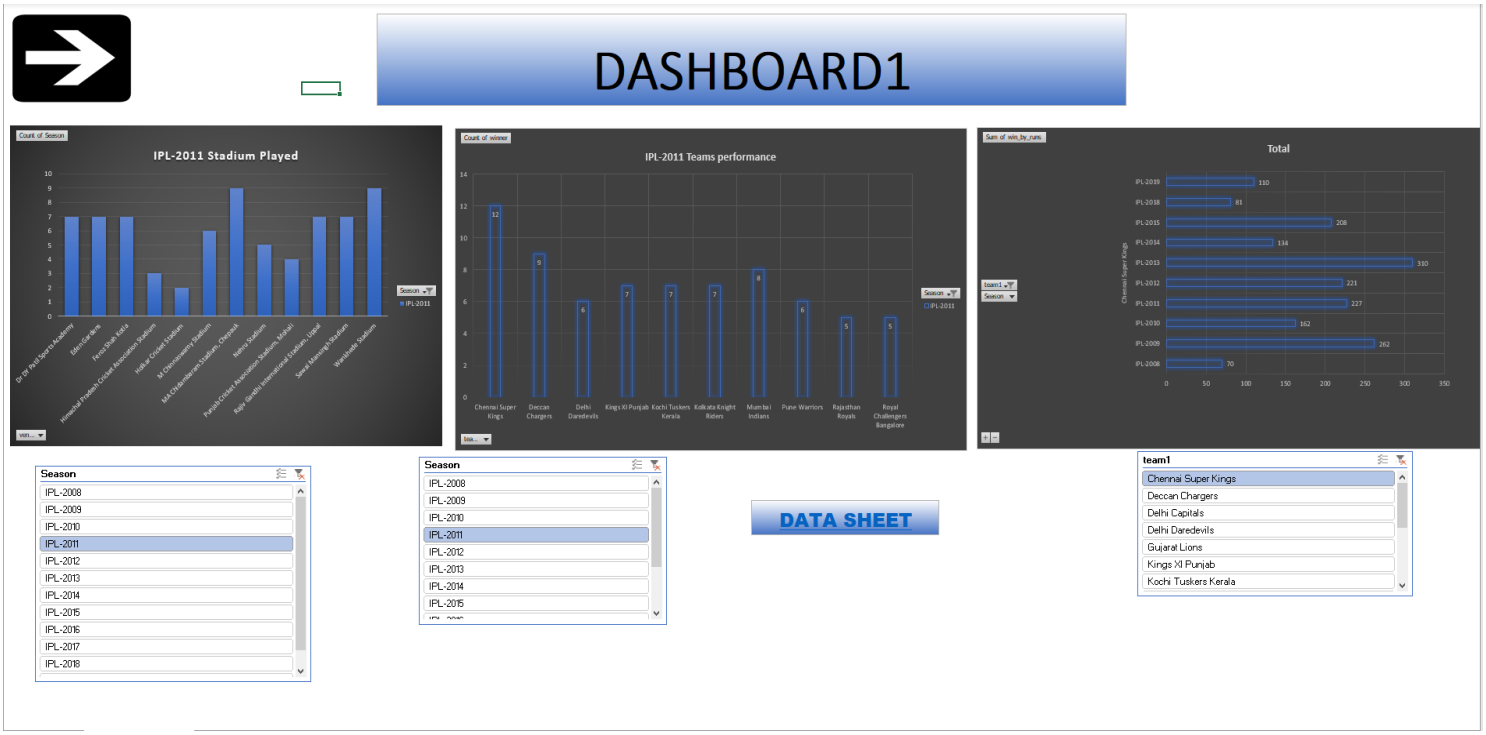
ii)With the clustered bar plotted.

c)Analysis Results:

We can observe total DL's taken by the team



DASHBOARDS OF THE PROJECT



References

1. www.kaggle.com
2. www.youtube.com
3. www.google.com
4. www.stackoverflow.com
5. www.github.com

Bibliography

1. Microsoft Excel 2016 Bible: The Comprehensive Tutorial Resource by John Walkenbach, Wiley
2. Fundamentals of Business Analytics by R.N. Prasad, Seema Acharya, Wiley