

## **INT217 PROJECT REPORT**

(Project Semester August-December 2021)

### ***Analysing IPL dataset(2008-2020)***

Submitted by

**Pappur Harsha Vardhan Reddy**

Programme and Section **Bachelor of Technology(CSE)**

### **CONTENTS:**

## Contents

## Page No

Introduction	3
Objectives	4
Source of dataset	5
ETL Process	6
Analysis of dataset and Visualization	12
List of Analysis with Results	25
Dashboard	26
Conclusion	27
Bibliography	28

## **Introduction**

Data management is an administrative process that includes acquiring, validating, storing protecting, and processing required data to ensure the accessibility, reliability, and timeliness of the data for its users

Organizations and enterprises are making use of Big Data more than ever before to inform business decisions and gain deep insights into customer behavior, trends, and opportunities for creating extraordinary customer experiences.

To make sense of the vast quantities of data that enterprises are gathering analyzing and storing today, companies turn to data management solutions and platforms. Data management solutions make processing validation, and other essential functions simpler and less time intensive.

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps. data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

There are many options for Excel focused on formulas which help you to automate your spreadsheets. Charts, graphs, and models can all be used to add a visual component to your data With Excel Charts, users will be able to find the best ways to display their data that could drive more effective decision making. Not only can you optimize the way which your data is presented through courses in Excel Charts, but you can also learn customization techniques so you can coordinate your Excel figures with associated documents in Microsoft Office.

In this project I chose IPL as my project theme. Because IPL was started in 2008 and if we analyze the data we can get the idea about teams, players, winners, places(stadiums) etc. I made a dashboard using pivot charts and pivot tables.

An Excel dashboard is one-pager that helps managers and business leaders tracking key or metrics and take a decision based on it. It contains charts/tables/views that are backed by data. A dashboard is often called a report, however, not all reports are dashboards.

When making a dashboard using dataset on excel objectives are taking main place. When we are handling big data set we need to make proper objectives to visualize the data. In this project I used

5 objectives to visualize data on the dashboard.

## **Objectives**

1. Matches won by team 1) By Batting first 2) By Bowling first
2. More times player of the match
3. (A) Winning Toss and Winning match (B) Winning Toss and Losing Match (C) Losing both Toss and Match (D) Losing Toss and Winning Match
4. Number of IPL tournaments won by a team
5. Matches won by Team by Huge Margin (number of runs Scored).

In this dataset I used pivot tables and pivot charts to visualize the data on my dataset. In dashboard year wise, team wise, city wise data are visualized.

## Source of Dataset

- A data set is a collection of data. In the case of tabular data, a data set corresponds to one or more database tables, where every column of a table represents a particular variable, and each row corresponds to a given record of the data set in question.
- The data set lists values for each of the variables, such as height and weight of an object, for each member of the data set.
- Each value is known as a datum. Data sets can also consist of a collection of documents or files. In this project I used a dataset which is including the details of IPL (2008-2020). And I found out this data set from Kaggle.com.
- Kaggle allows users to find and publish data sets, explore, and build models in a web-based data science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.
- Inside Kaggle you'll find all the code & data you need to do your data science work. Use over 19,000 public datasets and 200,000 public notebooks to conquer any analysis in no time.
- Apparently, they have a community of around 3 million people on the platform. In Kaggle.com there are lot of data sets available. Among them I chose the dataset IPL (2008-2020).

## **ETL PROCESS**

ETL is a process that extracts the data from different source systems, then transforms the data like applying calculations, concatenations and finally loads the data into the Data Warehouse system. Full form of ETL is Extract, Transform and Load.

It's tempting to think a creating a Data warehouse is simply extracting data from multiple sources and loading into database of a Data warehouse. This is far from the truth and requires a complex ETL process. The ETL process requires active inputs from various stakeholders including developers, analysts, testers, top executives and is technically challenging.

- Initially, the raw dataset was arranged as shown in given picture:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	id	city	date	player_of	venue	neutral_ve	team1	team2	toss_winn	toss_decis	winner	result	result_mai	eliminator	method	umpire1	umpire2
2	335982	Bangalore	4/18/2008	BB McCull	M Chinnas	0	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Chal	field	Kolkata Knight Riders	runs	140	N	NA	Asad Rauf	RE Koertzen
3	335983	Chandigarh	4/19/2008	MEK Husse	Punjab Cri	0	Kings XI Punjab	Chennai Super Kings	Chennai Su	bat	Chennai Super Kings	runs	33	N	NA	MR Benso	SL Shastri
4	335984	Delhi	4/19/2008	MF Mahar	Feroz Sha	0	Delhi Daredevils	Rajasthan Royals	Rajasthan	bat	Delhi Daredevils	wickets	9	N	NA	Aleem Dar	GA Pratapkuma
5	335985	Mumbai	4/20/2008	MV Bouch	Wankhede	0	Mumbai Indians	Royal Challengers Bangalore	Mumbai In	bat	Royal Challengers Bangalore	wickets	5	N	NA	SJ Davis	DJ Harper
6	335986	Kolkata	4/20/2008	DJ Hussey	Eden Gard	0	Kolkata Knight Riders	Deccan Chargers	Deccan Ch	bat	Kolkata Knight Riders	wickets	5	N	NA	BF Bowde	K Hariharan
7	335987	Jaipur	4/21/2008	SR Watson	Sawai Mar	0	Rajasthan Royals	Kings XI Punjab	Kings XI Pu	bat	Rajasthan Royals	wickets	6	N	NA	Aleem Dar	RB Tiffin
8	335988	Hyderabad	4/22/2008	V Sehwag	Rajiv Gand	0	Deccan Chargers	Delhi Daredevils	Deccan Ch	bat	Delhi Daredevils	wickets	9	N	NA	IL Howell	AM Saheba
9	335989	Chennai	4/23/2008	ML Hayden	MA Chidar	0	Chennai Super Kings	Mumbai Indians	Mumbai In	field	Chennai Super Kings	runs	6	N	NA	DJ Harper	GA Pratapkuma
10	335990	Hyderabad	4/24/2008	YK Pathan	Rajiv Gand	0	Deccan Chargers	Rajasthan Royals	Rajasthan	field	Rajasthan Royals	wickets	3	N	NA	Asad Rauf	MR Benson
11	335991	Chandigarh	4/25/2008	KC Sangak	Punjab Cri	0	Kings XI Punjab	Mumbai Indians	Mumbai In	field	Kings XI Punjab	runs	66	N	NA	Aleem Dar	AM Saheba
12	335992	Bangalore	4/26/2008	SR Watson	M Chinnas	0	Royal Challengers Bangalore	Rajasthan Royals	Rajasthan	field	Rajasthan Royals	wickets	7	N	NA	MR Benso	IL Howell
13	335993	Chennai	4/26/2008	JDP Oram	MA Chidar	0	Chennai Super Kings	Kolkata Knight Riders	Kolkata Kn	bat	Chennai Super Kings	wickets	9	N	NA	BF Bowde	AV Jayaprakash
14	335994	Mumbai	4/27/2008	AC Gilchris	Dr DY Patil	0	Mumbai Indians	Deccan Chargers	Deccan Ch	field	Deccan Chargers	wickets	10	N	NA	Asad Rauf	SL Shastri
15	335995	Chandigarh	4/27/2008	SM Katich	Punjab Cri	0	Kings XI Punjab	Delhi Daredevils	Delhi Dare	bat	Kings XI Punjab	wickets	4	N	NA	RE Koertze	I Shivram
16	335996	Bangalore	4/28/2008	MS Dhoni	M Chinnas	0	Royal Challengers Bangalore	Chennai Super Kings	Chennai Su	bat	Chennai Super Kings	runs	13	N	NA	BR Doctro	RB Tiffin
17	335997	Kolkata	4/29/2008	ST Jayasur	Eden Gard	0	Kolkata Knight Riders	Mumbai Indians	Kolkata Kn	bat	Mumbai Indians	wickets	7	N	NA	BF Bowde	AV Jayaprakash
18	335998	Delhi	4/30/2008	GD McGr	Feroz Sha	0	Delhi Daredevils	Royal Challengers Bangalore	Royal Chal	field	Delhi Daredevils	runs	10	N	NA	Aleem Dar	I Shivram
19	335999	Hyderabad	5/1/2008	SE Marsh	Rajiv Gand	0	Deccan Chargers	Kings XI Punjab	Kings XI Pu	field	Kings XI Punjab	wickets	7	N	NA	BR Doctro	RB Tiffin
20	336000	Jaipur	5/1/2008	SA Asnodk	Sawai Mar	0	Rajasthan Royals	Kolkata Knight Riders	Rajasthan	bat	Rajasthan Royals	runs	45	N	NA	RE Koertze	GA Pratapkuma
21	336001	Chennai	5/2/2008	V Sehwag	MA Chidar	0	Chennai Super Kings	Delhi Daredevils	Chennai Su	bat	Delhi Daredevils	wickets	8	N	NA	BF Bowde	K Hariharan
22	336002	Hyderabad	5/25/2008	R Vinay Ku	Rajiv Gand	0	Deccan Chargers	Royal Challengers Bangalore	Deccan Ch	bat	Royal Challengers Bangalore	wickets	5	N	NA	Asad Rauf	RE Koertzen
23	336003	Chandigarh	5/3/2008	IK Pathan	Punjab Cri	0	Kings XI Punjab	Kolkata Knight Riders	Kings XI Pu	bat	Kings XI Punjab	runs	9	N	NA	DJ Harper	I Shivram
24	336004	Mumbai	5/4/2008	SM Pollock	Dr DY Patil	0	Mumbai Indians	Delhi Daredevils	Delhi Dare	field	Mumbai Indians	runs	29	N	NA	IL Howell	RE Koertzen
25	336005	Jaipur	5/11/2008	Sohail Tan	Sawai Mar	0	Rajasthan Royals	Chennai Super Kings	Chennai Su	bat	Rajasthan Royals	wickets	8	N	NA	Asad Rauf	AV Jayaprakash

## ➤ Extraction

In this step, data is extracted from the source system into the staging area. Transformations if any are done in staging area so that performance of source system is not degraded. Also, if corrupted data is copied directly from the source into Data warehouse database, rollback

will be a challenge. Staging area gives an opportunity to validate extracted data before warehouse. moves into the Data.

**Step1:** To extract a csv file go to data tab and select From CSV/text in Get&Transform data Which takes us to the Query editor and then to clean the data select transform.

IPL Matches 2008-2020 (2).csv

File Origin

Delimiter

Data Type Detection

1252: Western European (Windows)

Comma

Based on first 200 rows

id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner	result	result_margin	eliminat
335982	Bangalore	4/18/2008	BB McCullum	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight Riders	runs	140	N
335983	Chandigarh	4/19/2008	MEK Hussey	Punjab Cricket Association Stadium, Mohali	0	Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs	33	N
335984	Delhi	4/19/2008	MF Maharoof	Feroz Shah Kotla	0	Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	bat	Delhi Daredevils	wickets	9	N
335985	Mumbai	4/20/2008	MV Boucher	Wankhede Stadium	0	Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore	wickets	5	N
335986	Kolkata	4/20/2008	DJ Hussey	Eden Gardens	0	Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders	wickets	5	N
335987	Jaipur	4/21/2008	SR Watson	Sawai Mansingh Stadium	0	Rajasthan Royals	Kings XI Punjab	Kings XI Punjab	bat	Rajasthan Royals	wickets	6	N
335988	Hyderabad	4/22/2008	V Sehwag	Rajiv Gandhi International Stadium, Uppal	0	Deccan Chargers	Delhi Daredevils	Deccan Chargers	bat	Delhi Daredevils	wickets	9	N
335989	Chennai	4/23/2008	ML Hayden	MA Chidambaram Stadium, Chepauk	0	Chennai Super Kings	Mumbai Indians	Mumbai Indians	field	Chennai Super Kings	runs	6	N
335990	Hyderabad	4/24/2008	YK Pathan	Rajiv Gandhi International Stadium, Uppal	0	Deccan Chargers	Rajasthan Royals	Rajasthan Royals	field	Rajasthan Royals	wickets	3	N
335991	Chandigarh	4/25/2008	KC Sangakkara	Punjab Cricket Association Stadium, Mohali	0	Kings XI Punjab	Mumbai Indians	Mumbai Indians	field	Kings XI Punjab	runs	66	N
335992	Bangalore	4/26/2008	SR Watson	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Rajasthan Royals	Rajasthan Royals	field	Rajasthan Royals	wickets	7	N
335993	Chennai	4/26/2008	JDP Oram	MA Chidambaram Stadium, Chepauk	0	Chennai Super Kings	Kolkata Knight Riders	Kolkata Knight Riders	bat	Chennai Super Kings	wickets	9	N
335994	Mumbai	4/27/2008	AC Gilchrist	Dr DY Patil Sports Academy	0	Mumbai Indians	Deccan Chargers	Deccan Chargers	field	Deccan Chargers	wickets	10	N
335995	Chandigarh	4/27/2008	SM Katich	Punjab Cricket Association Stadium, Mohali	0	Kings XI Punjab	Delhi Daredevils	Delhi Daredevils	bat	Kings XI Punjab	wickets	4	N
335996	Bangalore	4/28/2008	MS Dhoni	M Chinnaswamy Stadium	0	Royal Challengers Bangalore	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs	13	N
335997	Kolkata	4/29/2008	ST Jayasuriya	Eden Gardens	0	Kolkata Knight Riders	Mumbai Indians	Kolkata Knight Riders	bat	Mumbai Indians	wickets	7	N
335998	Delhi	4/30/2008	GD McGrath	Feroz Shah Kotla	0	Delhi Daredevils	Royal Challengers Bangalore	Royal Challengers Bangalore	field	Delhi Daredevils	runs	10	N
335999	Hyderabad	5/1/2008	SE Marsh	Rajiv Gandhi International Stadium, Uppal	0	Deccan Chargers	Kings XI Punjab	Kings XI Punjab	field	Kings XI Punjab	wickets	7	N
336000	Jaipur	5/1/2008	SA Asnodkar	Sawai Mansingh Stadium	0	Rajasthan Royals	Kolkata Knight Riders	Rajasthan Royals	bat	Rajasthan Royals	runs	45	N
336001	Chennai	5/2/2008	V Sehwag	MA Chidambaram Stadium, Chepauk	0	Chennai Super Kings	Delhi Daredevils	Chennai Super Kings	bat	Delhi Daredevils	wickets	8	N

The data in the preview has been truncated due to size limits.

<

>

Load

Transform Data

Cancel



## ➤ Transformation

Data extracted from source server is raw and not usable in its original form. Therefore it needs to be cleansed, mapped and transformed. In fact, this is the key step where ETL process adds value and changes data. In this step, you apply a set of functions on extracted data. Data that does not require any transformation is called as direct move or pass-through data.

**Step 2:** Transform the data by removing the unneeded columns, changing datatypes, removing null values and adding a new column known as season toss loss match loss and some combinations of previous columns using **If then Else** with as shown in the figure.

The screenshot displays the Power Query Editor interface for a query named "IPL Matches 2008-2020". The main data view shows a table with the following columns: Season, id, city, date, player\_of\_match, and venue. The data is organized into 28 rows, representing individual matches. The right-hand pane, titled "Query Settings", shows the "APPLIED STEPS" for this query, which include: Source, Promoted Headers, Changed Type, Removed Columns, Renamed Columns, Removed Columns1, Added Custom, Added Custom1, Added Custom2, Added Custom3, Added Custom4, Added Custom5, Added Custom6, Reordered Columns, Replaced Value, and Changed Type1. The bottom status bar indicates that the data consists of 22 columns and 816 rows, with column profiling based on the top 1000 rows.

Season	id	city	date	player_of_match	venue
2008	335982	Bangalore	4/18/2008	BB McCullum	M Chinnaswamy Stadium
2008	335983	Chandigarh	4/19/2008	MEK Hussey	Punjab Cricket Association Stadium, Mohali
2008	335984	Delhi	4/19/2008	MF Maharoof	Feroz Shah Kotla
2008	335985	Mumbai	4/20/2008	MV Boucher	Wankhede Stadium
2008	335986	Kolkata	4/20/2008	DI Hussey	Eden Gardens
2008	335987	Jaipur	4/21/2008	SR Watson	Sawai Mansingh Stadium
2008	335988	Hyderabad	4/22/2008	V Sehwag	Rajiv Gandhi International Stadium, Uppal
2008	335989	Chennai	4/23/2008	ML Hayden	MA Chidambaram Stadium, Chepauk
2008	335990	Hyderabad	4/24/2008	YK Pathan	Rajiv Gandhi International Stadium, Uppal
2008	335991	Chandigarh	4/25/2008	KC Sangakkara	Punjab Cricket Association Stadium, Mohali
2008	335992	Bangalore	4/26/2008	SR Watson	M Chinnaswamy Stadium
2008	335993	Chennai	4/26/2008	IDP Oram	MA Chidambaram Stadium, Chepauk
2008	335994	Mumbai	4/27/2008	AC Gilchrist	Dr DY Patil Sports Academy
2008	335995	Chandigarh	4/27/2008	SM Katich	Punjab Cricket Association Stadium, Mohali
2008	335996	Bangalore	4/28/2008	MS Dhoni	M Chinnaswamy Stadium
2008	335997	Kolkata	4/28/2008	ST Jayasuriya	Eden Gardens
2008	335998	Delhi	4/30/2008	GD McGrath	Feroz Shah Kotla
2008	335999	Hyderabad	5/1/2008	SE Marsh	Rajiv Gandhi International Stadium, Uppal
2008	336000	Jaipur	5/1/2008	SA Asnodkar	Sawai Mansingh Stadium
2008	336001	Chennai	5/2/2008	V Sehwag	MA Chidambaram Stadium, Chepauk
2008	336002	Hyderabad	5/25/2008	R Vinay Kumar	Rajiv Gandhi International Stadium, Uppal
2008	336003	Chandigarh	5/4/2008	IK Pathan	Punjab Cricket Association Stadium, Mohali
2008	336004	Mumbai	5/4/2008	SM Pollock	Dr DY Patil Sports Academy
2008	336005	Jaipur	5/4/2008	Sohail Tanvir	Sawai Mansingh Stadium
2008	336006	Bangalore	5/5/2008	S Sreesanth	M Chinnaswamy Stadium
2008	336007	Chennai	5/6/2008	AC Gilchrist	MA Chidambaram Stadium, Chepauk
2008	336008	Mumbai	5/7/2008	A Nehra	Dr DY Patil Sports Academy
2008	336009	Delhi	5/8/2008	MS Dhoni	Feroz Shah Kotla

- If then else function is used in Transformation to form new columns using toss winner, toss loser, game winner and game loser. This function was used as follows.

The screenshot displays the Microsoft Power Query Editor interface. A 'Custom Column' dialog box is open, allowing the user to create a new column based on existing data. The formula entered is `= if [team1] = [toss_winner] then [team2] else [team1]`, which is designed to identify the team that won the toss. The new column is named 'Toss\_loser'. The background shows a table of IPL matches from 2008 to 2020, with columns for Season, id, city, date, player\_of\_match, and venue. The status bar at the bottom indicates that there are 17 columns and 816 rows of data, with column profiling based on the top 1000 rows.

## ➤ Loading

Loading data into the target data warehouse database is the last step of the ETL process. In a typical Data warehouse, huge volume of data needs to be loaded in

a relatively short period (nights). Hence, load process should be optimized for performance.

**Step 3:** After pressing load in the power query editor the output will be as follows

The screenshot displays the Microsoft Excel interface with a data table of IPL matches. The table has columns for Team1, Team2, Toss, Toss Decider, Game Winner, Game Loser, Toss Result, and various player statistics. The Queries & Connections pane on the right shows a single query named 'IPL Matches 2008-2020' with 816 rows loaded. The status bar at the bottom indicates the average, count, and sum of the data.

Team1	Team2	Toss	Toss Decider	Game Winner	Game Loser	Toss Result	Player1	Player2	Toss and Game Winner	Toss
Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	Kolkata Knight Riders	field	Kolkata Knight Riders	Royal Challengers Bangalore runs	140 N	Asad Raul	RE Koertzen	NA
Kings XI Punjab	Chennai Super Kings	Kings XI Punjab	Chennai Super Kings	bat	Chennai Super Kings	Kings XI Punjab runs	33 N	MR Benson	SL Shanti	Chennai Super Kings
Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	Delhi Daredevils	bat	Delhi Daredevils	Rajasthan Royals	violates	5 N	Aleem Dar	GA Prasad Kumar
Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	Royal Challengers Bangalore	bat	Royal Challengers Bangalore	Mumbai Indians	violates	5 N	S/Davis	NA
Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	Kolkata Knight Riders	bat	Kolkata Knight Riders	Deccan Chargers	violates	5 N	BF Bowden	K Harsharan
Rajasthan Royals	Kings XI Punjab	Rajasthan Royals	Kings XI Punjab	bat	Rajasthan Royals	Kings XI Punjab	violates	5 N	Aleem Dar	RE Tiffin
Deccan Chargers	Delhi Daredevils	Deccan Chargers	Delhi Daredevils	bat	Delhi Daredevils	Deccan Chargers	violates	3 N	L Howell	AN Saheba
Chennai Super Kings	Mumbai Indians	Chennai Super Kings	Chennai Super Kings	field	Chennai Super Kings	Mumbai Indians	runs	5 N	D/Harper	GA Prasad Kumar
Deccan Chargers	Rajasthan Royals	Deccan Chargers	Rajasthan Royals	field	Rajasthan Royals	Deccan Chargers	violates	3 N	Asad Raul	MR Benson
Kings XI Punjab	Mumbai Indians	Mumbai Indians	Kings XI Punjab	field	Kings XI Punjab	Mumbai Indians	runs	66 N	Aleem Dar	AN Saheba
Royal Challengers Bangalore	Rajasthan Royals	Rajasthan Royals	Royal Challengers Bangalore	field	Rajasthan Royals	Royal Challengers Bangalore	violates	7 N	MR Benson	L Howell
Chennai Super Kings	Kolkata Knight Riders	Kolkata Knight Riders	Chennai Super Kings	bat	Chennai Super Kings	Kolkata Knight Riders	violates	9 N	BF Bowden	AV Jayaprakash
Mumbai Indians	Deccan Chargers	Deccan Chargers	Mumbai Indians	field	Deccan Chargers	Mumbai Indians	violates	10 N	Asad Raul	SL Shanti
Kings XI Punjab	Delhi Daredevils	Delhi Daredevils	Kings XI Punjab	bat	Kings XI Punjab	Delhi Daredevils	violates	4 N	RE Koertzen	I Shivan
Royal Challengers Bangalore	Chennai Super Kings	Chennai Super Kings	Royal Challengers Bangalore	bat	Chennai Super Kings	Royal Challengers Bangalore	runs	13 N	BR Doctrove	RB Tiffin
Kolkata Knight Riders	Mumbai Indians	Kolkata Knight Riders	Mumbai Indians	bat	Mumbai Indians	Kolkata Knight Riders	violates	7 N	BF Bowden	AV Jayaprakash
Delhi Daredevils	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	field	Delhi Daredevils	Royal Challengers Bangalore	runs	10 N	Aleem Dar	I Shivan
Deccan Chargers	Kings XI Punjab	Kings XI Punjab	Deccan Chargers	field	Kings XI Punjab	Deccan Chargers	violates	7 N	BR Doctrove	RE Tiffin
Rajasthan Royals	Kolkata Knight Riders	Rajasthan Royals	Kolkata Knight Riders	bat	Rajasthan Royals	Kolkata Knight Riders	runs	45 N	RE Koertzen	GA Prasad Kumar
Chennai Super Kings	Delhi Daredevils	Chennai Super Kings	Delhi Daredevils	bat	Delhi Daredevils	Chennai Super Kings	violates	6 N	BF Bowden	K Harsharan
Deccan Chargers	Royal Challengers Bangalore	Deccan Chargers	Royal Challengers Bangalore	bat	Royal Challengers Bangalore	Deccan Chargers	violates	5 N	Asad Raul	RE Koertzen
Kings XI Punjab	Kolkata Knight Riders	Kings XI Punjab	Kolkata Knight Riders	bat	Kings XI Punjab	Kolkata Knight Riders	runs	9 N	D/Harper	I Shivan
Mumbai Indians	Delhi Daredevils	Delhi Daredevils	Mumbai Indians	field	Mumbai Indians	Delhi Daredevils	runs	29 N	L Howell	RE Koertzen
Rajasthan Royals	Chennai Super Kings	Chennai Super Kings	Rajasthan Royals	bat	Rajasthan Royals	Chennai Super Kings	violates	8 N	Asad Raul	AV Jayaprakash
Royal Challengers Bangalore	Kings XI Punjab	Kings XI Punjab	Royal Challengers Bangalore	field	Kings XI Punjab	Royal Challengers Bangalore	violates	6 N	S/Davis	BR Doctrove
Chennai Super Kings	Deccan Chargers	Deccan Chargers	Chennai Super Kings	field	Deccan Chargers	Chennai Super Kings	violates	7 N	MR Benson	RB Tiffin
Mumbai Indians	Rajasthan Royals	Mumbai Indians	Rajasthan Royals	field	Mumbai Indians	Rajasthan Royals	violates	7 N	D/Harper	RE Koertzen
Delhi Daredevils	Chennai Super Kings	Chennai Super Kings	Delhi Daredevils	field	Chennai Super Kings	Delhi Daredevils	violates	4 N	Aleem Dar	RE Tiffin
Kolkata Knight Riders	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	bat	Kolkata Knight Riders	Royal Challengers Bangalore	runs	5 N	Asad Raul	L Howell
Rajasthan Royals	Deccan Chargers	Rajasthan Royals	Deccan Chargers	field	Rajasthan Royals	Deccan Chargers	violates	6 N	MR Benson	AN Saheba
Royal Challengers Bangalore	Mumbai Indians	Mumbai Indians	Royal Challengers Bangalore	field	Mumbai Indians	Royal Challengers Bangalore	violates	9 N	BF Bowden	AV Jayaprakash
Chennai Super Kings	Kings XI Punjab	Kings XI Punjab	Chennai Super Kings	bat	Chennai Super Kings	Kings XI Punjab	runs	18 N	AV Jayaprakash	BG Jelting
Deccan Chargers	Kolkata Knight Riders	Kolkata Knight Riders	Deccan Chargers	bat	Kolkata Knight Riders	Deccan Chargers	runs	23 N	L Howell	AN Saheba
Rajasthan Royals	Delhi Daredevils	Rajasthan Royals	Delhi Daredevils	field	Rajasthan Royals	Delhi Daredevils	violates	3 N	S/Davis	RE Koertzen
Kings XI Punjab	Royal Challengers Bangalore	Kings XI Punjab	Royal Challengers Bangalore	bat	Kings XI Punjab	Royal Challengers Bangalore	violates	9 N	BR Doctrove	I Shivan
Kolkata Knight Riders	Delhi Daredevils	Kolkata Knight Riders	Delhi Daredevils	bat	Kolkata Knight Riders	Delhi Daredevils	runs	23 N	Asad Raul	L Howell
Chennai Super Kings	Mumbai Indians	Chennai Super Kings	Mumbai Indians	field	Chennai Super Kings	Mumbai Indians	violates	3 N	BR Doctrove	AN Saheba
Kings XI Punjab	Rajasthan Royals	Rajasthan Royals	Kings XI Punjab	field	Rajasthan Royals	Kings XI Punjab	runs	41 N	S/Davis	K Harsharan
Delhi Daredevils	Deccan Chargers	Deccan Chargers	Delhi Daredevils	field	Deccan Chargers	Delhi Daredevils	runs	12 N	BG Jelting	GA Prasad Kumar
Mumbai Indians	Kolkata Knight Riders	Mumbai Indians	Kolkata Knight Riders	field	Mumbai Indians	Kolkata Knight Riders	violates	6 N	BR Doctrove	D/Harper
Delhi Daredevils	Kings XI Punjab	Delhi Daredevils	Kings XI Punjab	bat	Kings XI Punjab	Delhi Daredevils	runs	6 N	AV Jayaprakash	RE Koertzen
Rajasthan Royals	Royal Challengers Bangalore	Rajasthan Royals	Royal Challengers Bangalore	bat	Rajasthan Royals	Royal Challengers Bangalore	runs	65 N	BF Bowden	SL Shanti
Deccan Chargers	Mumbai Indians	Deccan Chargers	Mumbai Indians	field	Mumbai Indians	Deccan Chargers	runs	25 N	BR Doctrove	D/Harper
Kolkata Knight Riders	Chennai Super Kings	Chennai Super Kings	Kolkata Knight Riders	bat	Chennai Super Kings	Kolkata Knight Riders	runs	3 N	Asad Raul	K Harsharan
Royal Challengers Bangalore	Delhi Daredevils	Delhi Daredevils	Royal Challengers Bangalore	field	Delhi Daredevils	Royal Challengers Bangalore	violates	5 N	S/Davis	GA Prasad Kumar

## **Analysis of Data Set**

### **i. Introduction**

A dashboard is a visual interface that provides at-a-glance views into key measures relevant to a particular objective or business process. Excel dashboards make it easy to perform quick overviews of data reports rather than going through large volumes of data. Overviews help in making quick and urgent decisions since one can skim through a lot of information at once and within a short time.

Excel dashboards include various elements such as charts, tables, figures, and gauges that help in presenting the data. They can handle any type of data from different market and purposes, and the information can be used for marketing, financial, or other business projects. The dashboard is most applicable to large volumes of data since it would otherwise be hectic to go through such large volumes of data, especially with limited time.

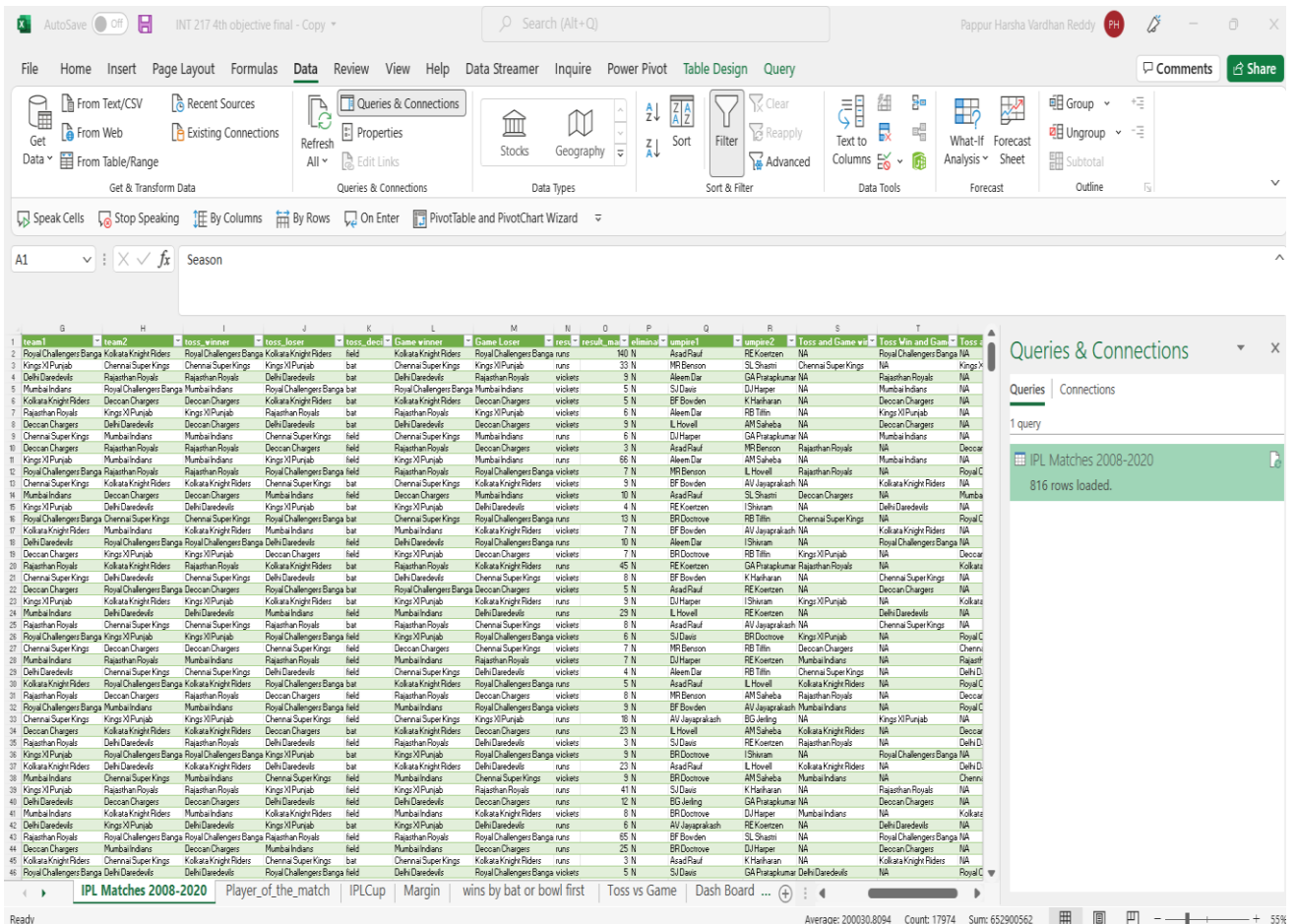
Dashboards have three main attributes:

- ❖ Dashboards are typically graphical in nature, providing visualizations that help focus attention on key trends, comparisons, and exceptions
- ❖ Dashboards often display only data that are relevant to the goal of the dashboard

Because dashboards are designed with a specific purpose or goal, they inherently contain predefined conclusions that relieve the end user from performing his own analysis

## ii. General Description

In this project I used a data set IPL(2008-2020) from the kaggle.com. In the data set there are 17 columns & 817 rows Data set is containing all data for 2008-2020.



The screenshot displays the Microsoft Excel interface with the 'IPL Matches 2008-2020' data set loaded. The ribbon at the top includes 'File', 'Home', 'Insert', 'Page Layout', 'Formulas', 'Data', 'Review', 'View', 'Help', 'Data Streamer', 'Inquire', 'Power Pivot', 'Table Design', and 'Query'. The 'Data' tab is active, showing the 'Queries & Connections' pane on the right, which indicates that 816 rows have been loaded from the 'IPL Matches 2008-2020' data source. The spreadsheet grid shows columns for match details, including 'Season', 'Team', 'Venue', 'Toss', 'Game', 'Result', 'Player', and 'Score'. The data is organized into rows, with the first row showing a match between Royal Challengers Bangalore and Kolkata Knight Riders in the 2008 season.

Season	Team	Venue	Toss	Game	Result	Player	Score
2008	Royal Challengers Bangalore	Kolkata Knight Riders	field	Kolkata Knight Riders	Royal Challengers Bangalore	nuns	40
2008	Chennai Super Kings	Chennai Super Kings	field	Chennai Super Kings	Kings XI Punjab	nuns	33
2008	Dehi Daredevils	Rajasthan Royals	bat	Dehi Daredevils	Rajasthan Royals	victors	9
2008	Mumbai Indians	Royal Challengers Bangalore	bat	Royal Challengers Bangalore	Mumbai Indians	victors	5
2008	Kolkata Knight Riders	Deccan Chargers	bat	Kolkata Knight Riders	Deccan Chargers	victors	5
2008	Rajasthan Royals	Kings XI Punjab	bat	Rajasthan Royals	Kings XI Punjab	victors	6
2008	Deccan Chargers	Dehi Daredevils	field	Deccan Chargers	Deccan Chargers	victors	9
2008	Chennai Super Kings	Mumbai Indians	field	Chennai Super Kings	Chennai Super Kings	nuns	6
2008	Deccan Chargers	Rajasthan Royals	field	Rajasthan Royals	Deccan Chargers	victors	3
2008	Kings XI Punjab	Mumbai Indians	field	Kings XI Punjab	Mumbai Indians	nuns	66
2008	Royal Challengers Bangalore	Rajasthan Royals	field	Rajasthan Royals	Royal Challengers Bangalore	victors	7
2008	Chennai Super Kings	Kolkata Knight Riders	bat	Chennai Super Kings	Kolkata Knight Riders	victors	9
2008	Mumbai Indians	Deccan Chargers	field	Mumbai Indians	Deccan Chargers	victors	10
2008	Kings XI Punjab	Dehi Daredevils	bat	Kings XI Punjab	Dehi Daredevils	victors	4
2008	Royal Challengers Bangalore	Chennai Super Kings	field	Chennai Super Kings	Royal Challengers Bangalore	nuns	13
2008	Kolkata Knight Riders	Mumbai Indians	bat	Mumbai Indians	Kolkata Knight Riders	victors	7
2008	Dehi Daredevils	Royal Challengers Bangalore	field	Dehi Daredevils	Royal Challengers Bangalore	nuns	10
2008	Deccan Chargers	Kings XI Punjab	field	Kings XI Punjab	Deccan Chargers	victors	7
2008	Rajasthan Royals	Kolkata Knight Riders	bat	Rajasthan Royals	Kolkata Knight Riders	nuns	45
2008	Chennai Super Kings	Dehi Daredevils	field	Dehi Daredevils	Chennai Super Kings	victors	8
2008	Deccan Chargers	Royal Challengers Bangalore	bat	Royal Challengers Bangalore	Deccan Chargers	victors	5
2008	Kings XI Punjab	Kolkata Knight Riders	bat	Kings XI Punjab	Kolkata Knight Riders	nuns	9
2008	Mumbai Indians	Dehi Daredevils	field	Mumbai Indians	Dehi Daredevils	nuns	28
2008	Rajasthan Royals	Chennai Super Kings	bat	Rajasthan Royals	Chennai Super Kings	victors	8
2008	Royal Challengers Bangalore	Kings XI Punjab	field	Kings XI Punjab	Royal Challengers Bangalore	victors	6
2008	Chennai Super Kings	Deccan Chargers	field	Deccan Chargers	Chennai Super Kings	victors	7
2008	Mumbai Indians	Rajasthan Royals	field	Rajasthan Royals	Mumbai Indians	victors	7
2008	Dehi Daredevils	Chennai Super Kings	field	Chennai Super Kings	Dehi Daredevils	victors	4
2008	Royal Challengers Bangalore	Rajasthan Royals	field	Rajasthan Royals	Royal Challengers Bangalore	victors	5
2008	Deccan Chargers	Deccan Chargers	field	Deccan Chargers	Deccan Chargers	victors	8
2008	Royal Challengers Bangalore	Mumbai Indians	field	Mumbai Indians	Royal Challengers Bangalore	victors	8
2008	Chennai Super Kings	Kings XI Punjab	field	Chennai Super Kings	Kings XI Punjab	nuns	18
2008	Kolkata Knight Riders	Deccan Chargers	bat	Kolkata Knight Riders	Deccan Chargers	nuns	23
2008	Rajasthan Royals	Dehi Daredevils	field	Rajasthan Royals	Dehi Daredevils	victors	3
2008	Kings XI Punjab	Royal Challengers Bangalore	bat	Kings XI Punjab	Royal Challengers Bangalore	victors	9
2008	Kolkata Knight Riders	Dehi Daredevils	bat	Kolkata Knight Riders	Dehi Daredevils	nuns	23
2008	Mumbai Indians	Chennai Super Kings	field	Mumbai Indians	Chennai Super Kings	victors	9
2008	Kings XI Punjab	Rajasthan Royals	field	Rajasthan Royals	Kings XI Punjab	nuns	41
2008	Dehi Daredevils	Deccan Chargers	field	Dehi Daredevils	Deccan Chargers	nuns	12
2008	Mumbai Indians	Kolkata Knight Riders	field	Mumbai Indians	Kolkata Knight Riders	victors	8
2008	Dehi Daredevils	Kings XI Punjab	bat	Dehi Daredevils	Kings XI Punjab	nuns	6
2008	Rajasthan Royals	Royal Challengers Bangalore	field	Rajasthan Royals	Royal Challengers Bangalore	nuns	65
2008	Deccan Chargers	Mumbai Indians	field	Deccan Chargers	Mumbai Indians	victors	25
2008	Kolkata Knight Riders	Chennai Super Kings	bat	Chennai Super Kings	Kolkata Knight Riders	nuns	3
2008	Royal Challengers Bangalore	Dehi Daredevils	field	Royal Challengers Bangalore	Dehi Daredevils	victors	5

In this data set there are 22 columns and 817 rows columns named as id, city, date, team 1, team 2, toss winner, toss decision, result, al applied winner, win by runs, win by wickets, player of the match,venue, umpire 1. umpire 2. umpire

In dashboard mainly I used 5 charts to Visualize the data and they change according to year.

### **(iii)Specific Requirements-Functions and Formulas**

Making dashboard I used pivot tables and pivot charts to analyze the data. With the help of them we can analyse the data and visualize the data very well. There are lot of advantages of using pivot tables and pivot charts. This is effective way of representing data in a pictorial manner, makes the process of data visualization of data easier. And I used 1 slicer to visualize data more attractive. And for the Toss vs Match objective I used the Index, Match , count, count if functions and a Dropbox.

#### **✓Pivot Tables**

Pivot tables are one of Excel's most powerful features. A pivot table allows you to extract the significance from a large, detailed data set. A pivot table is a statistics tool that summarizes and reorganizes selected columns and rows of data in a spreadsheet or database table to obtain a desired report. The tool does not actually change the spreadsheet or database itself, it simply "pivots" or turns the data to view it from different perspectives.

Pivot tables are especially useful with large amounts of data that would be time-consuming to calculate by hand. A few data processing functions a pivot table can perform include identifying sums, averages, ranges, and others. The table then arranges this information in a simple, meaningful layout that draws attention to key values. Pivot table is a generic term. This refers to a tool specific to Excel for creating pivot tables.

#### ✓Pivot Charts

Pivot Charts complement PivotTables by adding visualizations to the summary data in a Pivot Table, and allow you to easily see comparisons, patterns, and trends. A pivot chart is especially useful when dealing with tremendous amounts of data.

For example, a society having many employees is maintaining the working hours of each pupil through Excel, such that at the end of each month, the employee with the highest number of working hours, would be provided a bonus due to the sincerity and devotion to the society.

### **(iv) Analysis Results**

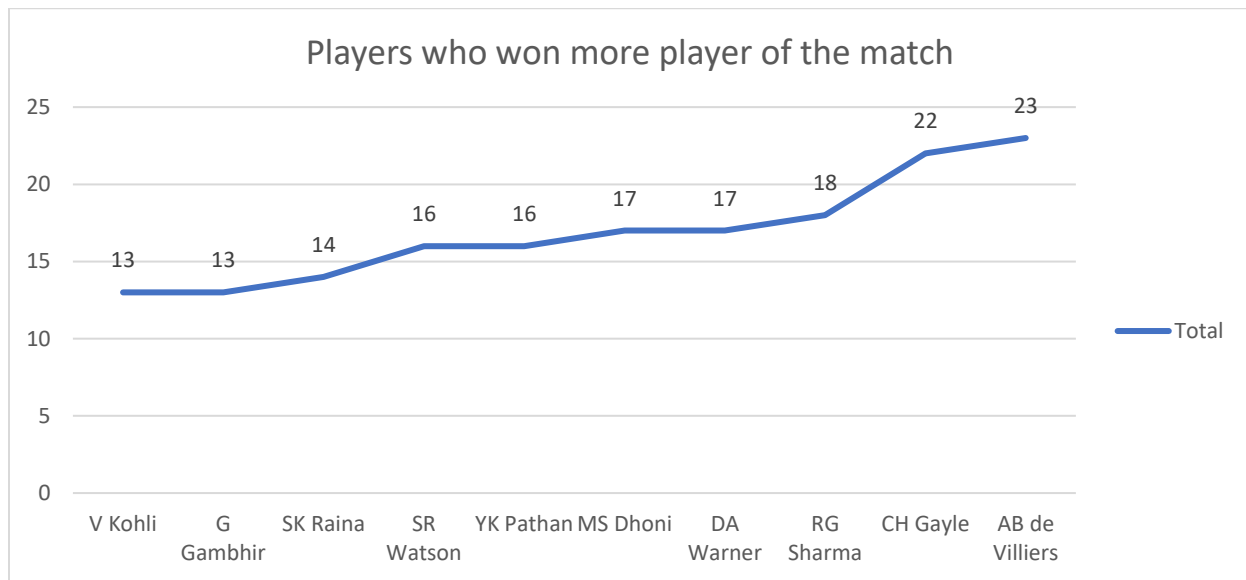
#### **1.Top 10 players who won Man of the Match:**

**a)** Introduction: In this objective we have listed the top 10 players who won more number of times the Man of the Match player.

**b)** Specific Requirements:

- i. Inserted a Pivot table from the data and filtered the Top10 players who won more number of Man of the Match.
- ii. With the help of data Column chart is plotted.
- iii. With the help of slicer of season, we can know which player won more man of the matches in a particular season

c) Analysis Results: AB de Villiers is the player who won a greater number of times Man of the match in all seasons



## 2.Total Number of matches won by each team

a) Introduction: In this objective I had showed that the count of IPL tournament winnings of each team

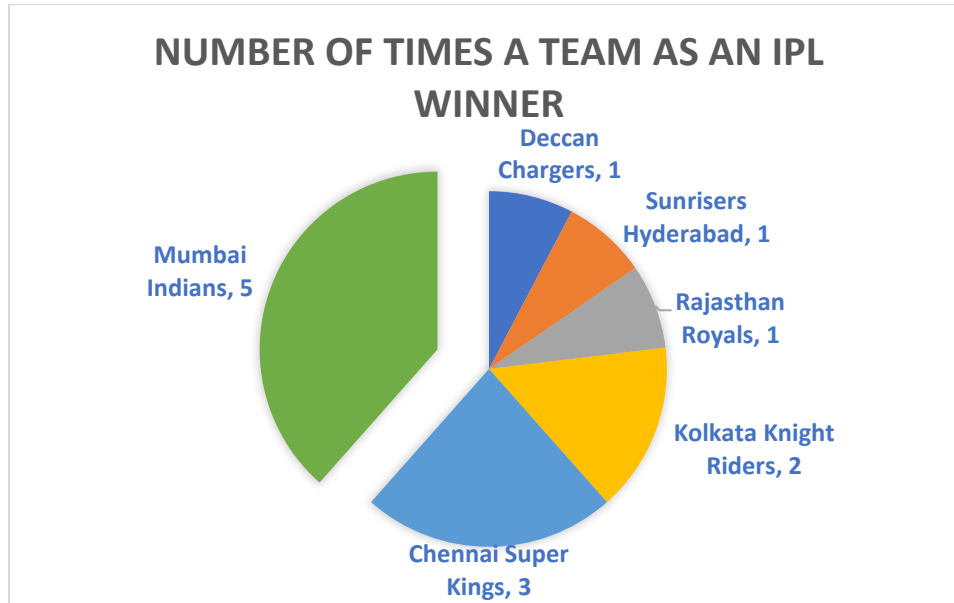
b) Specific Requirements:

- i. Inserted a Pivot table from the data and filtered the finals of each season and got the team who won the IPL season of that particular season.
- ii. With the help of data line chart is plotted.
- iii. With the help of slicer of season we can know the total winning of each IPL season.



**c) Analysis Results:**

MI is the team who won most number of IPL tournaments .



**3.Total Number of matches won by each team**

**a) Introduction:** In this objective I had showed that the count of winnings of each team by choosing bowling first and by batting first.

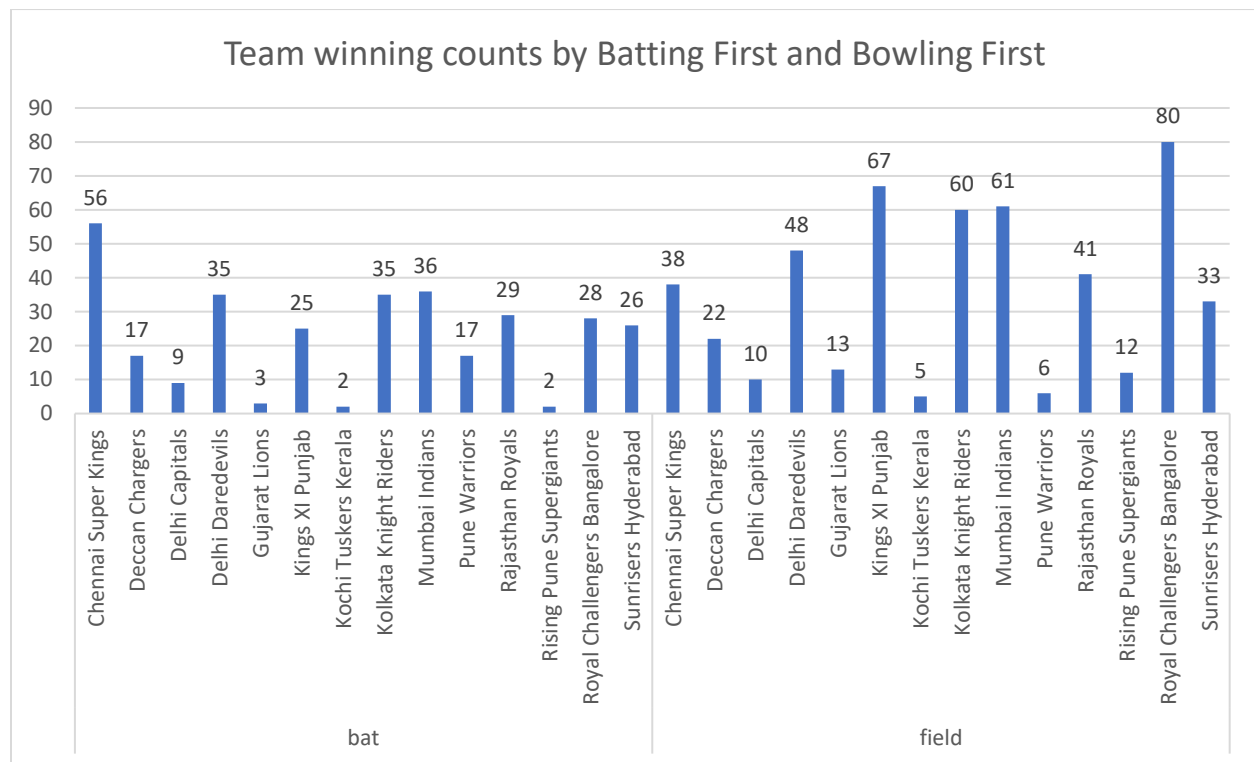
**b) Specific Requirements:**

- i. Inserted a Pivot table from the data and filtered the Top10 players who won a greater number of Man of the Match.
- ii. With the help of data Bar chart is plotted.
- iii. With the help of slicer of city and season we can know the total winning of each team by bowling first and batting first.

**c) Analysis Results:**

RCB is the team who won more number of matches by choosing bowling first

CSK is the team who won more number of matches by choosing batting first.



#### 4.Toss Vs Match:

**a)Introduction:** In this Objective, We can see the Number of Matches :

1)Won Toss and Won Match

2)Won Toss and Lost Match

3)Lost Toss and Won Match

4)Lost Toss and Won Match by respective Teams.

**b) Specific Requirements/Functions and Formulas:**

- ❖ If then else functions on **ETL** are used.
- ❖ I have created two columns of Toss Looser and Game Looser with the help of ETL.

- ❖ And after creating those columns in ETL I have used again IF then else function in ETL to create a combinations of toss and match wins and losses.

- ❖ After that I have created the four pivot tables for the

- 1)won toss and won match

- 2)lost toss and won match

- 3)lost toss and lost match

- 4) won toss and lost match.

And then created a dynamic regular table which takes input from the Drop Down list as I created.

- ❖ I have created a Dropdown List in which we can select the above four listed categories and with the help of **Index** and **Match** functions I created a formula which will display the count of matches played by respective teams who

- ❖ 1)won toss and won match

- ❖ 2)lost toss and won match

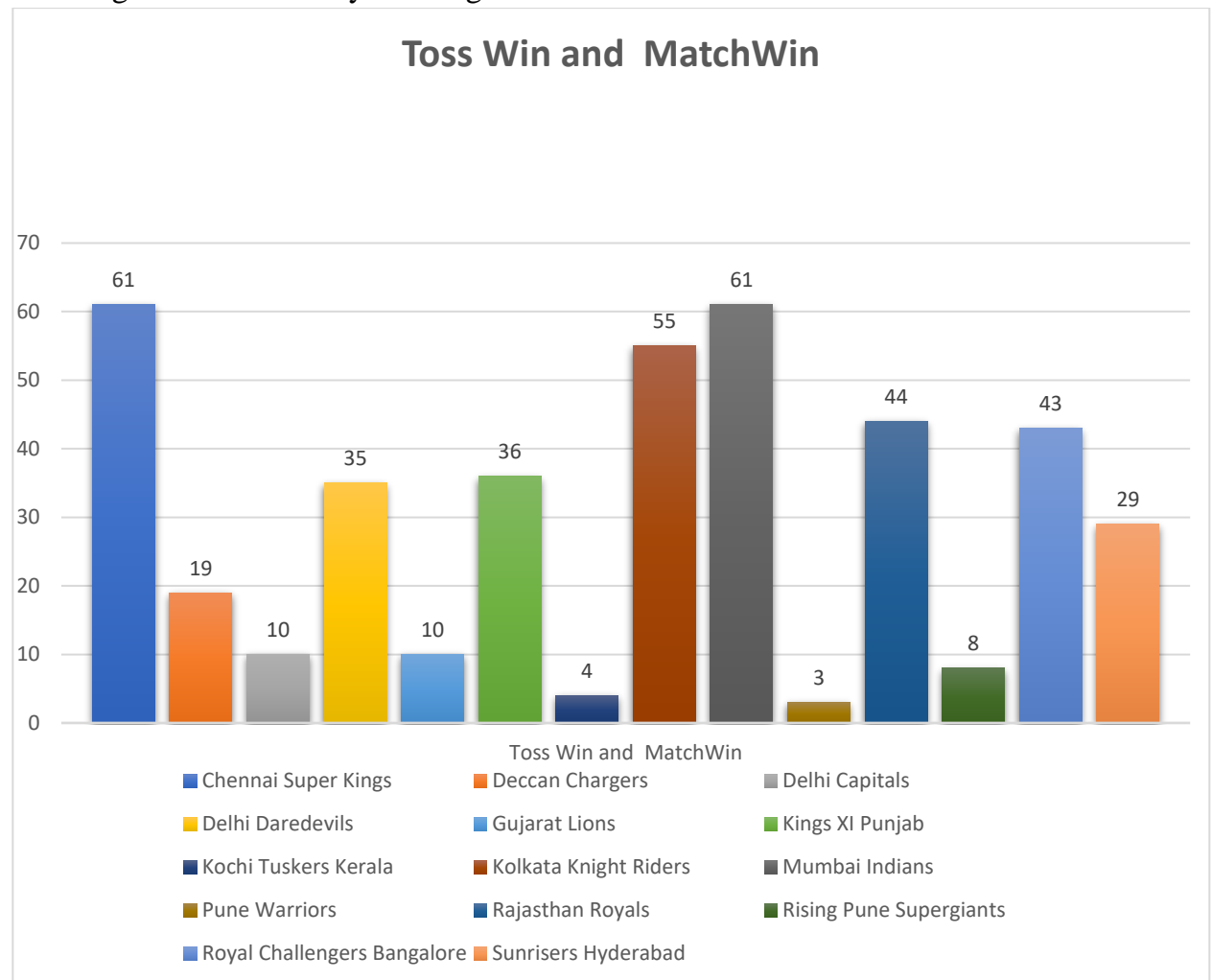
- ❖ 3)lost toss and lost match

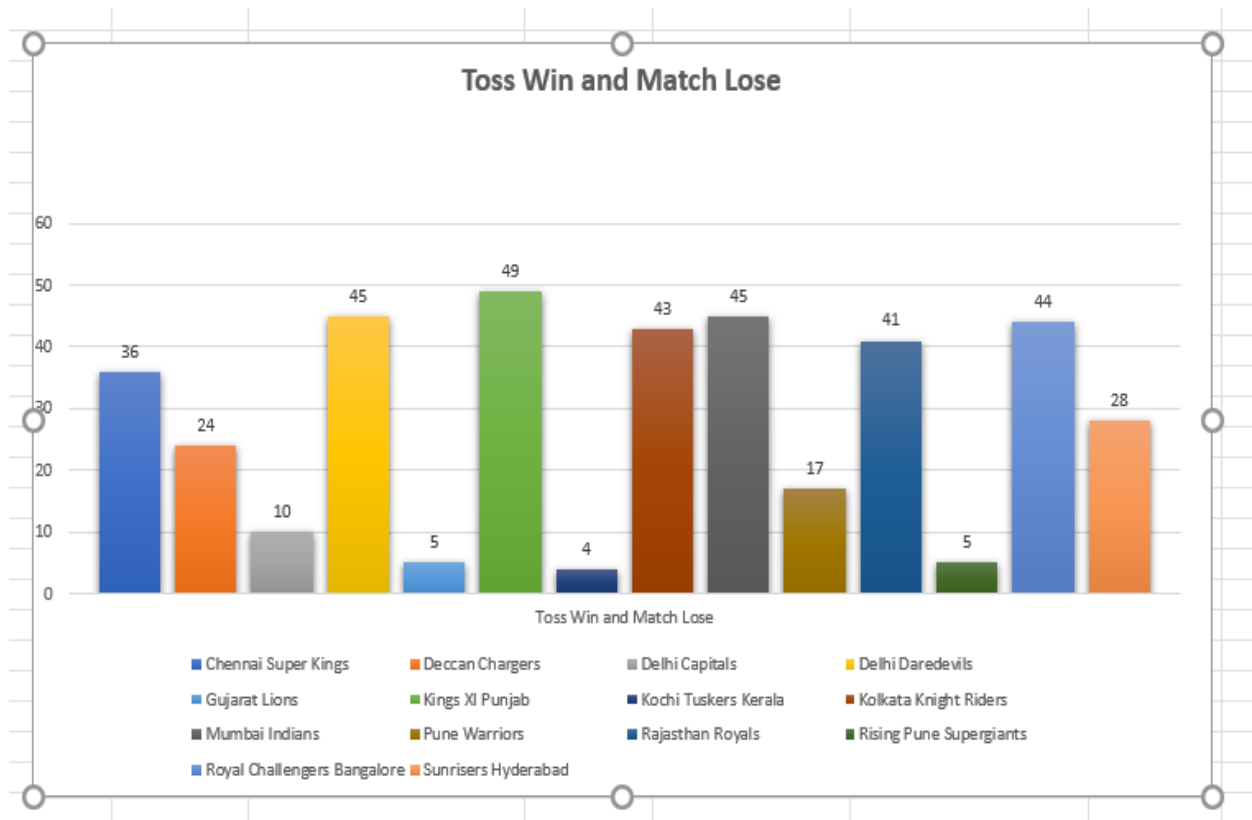
- ❖ 4) won toss and lost match.

And when we select the an item from the drop down it is taken reference in the regular table and the Index match functions take place from there and gives us the dynamic values.

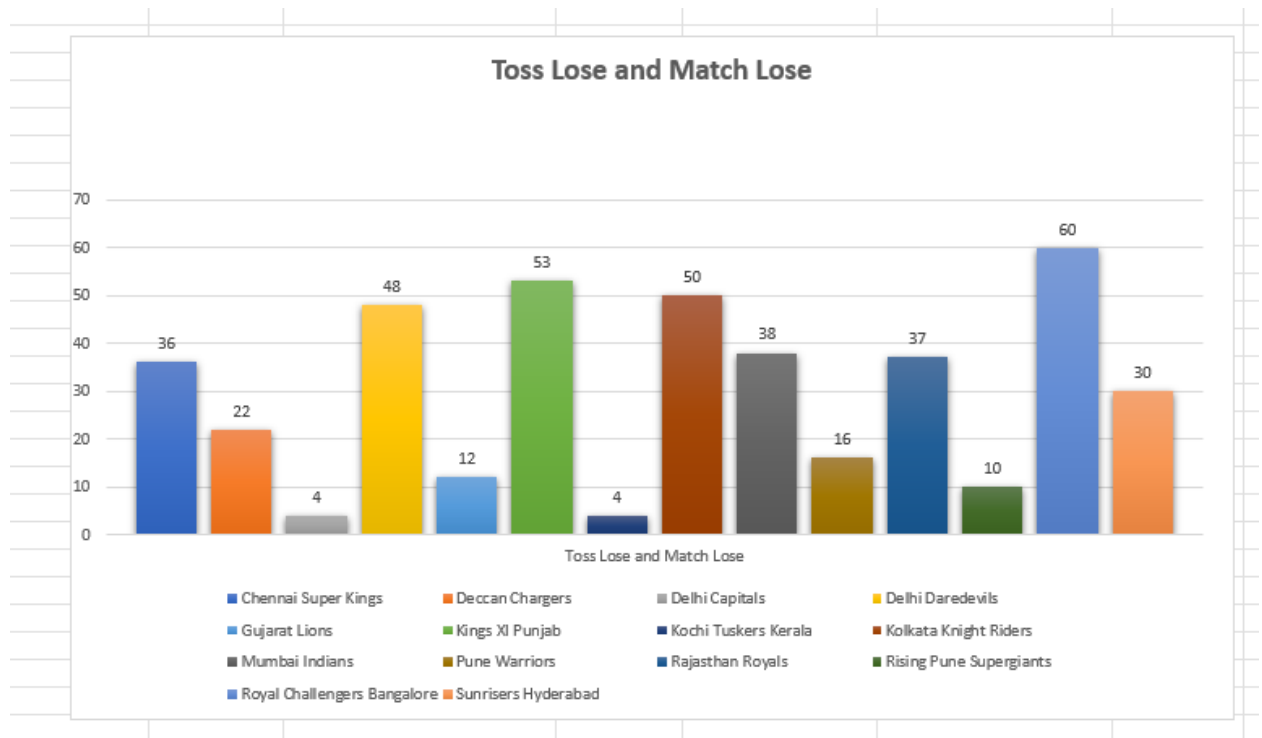
### **c) Analysis Results:**

Mumbai Indians and Chennai Super Kings had won more Number of Matches  
Winning Toss and also By Loosing Toss.

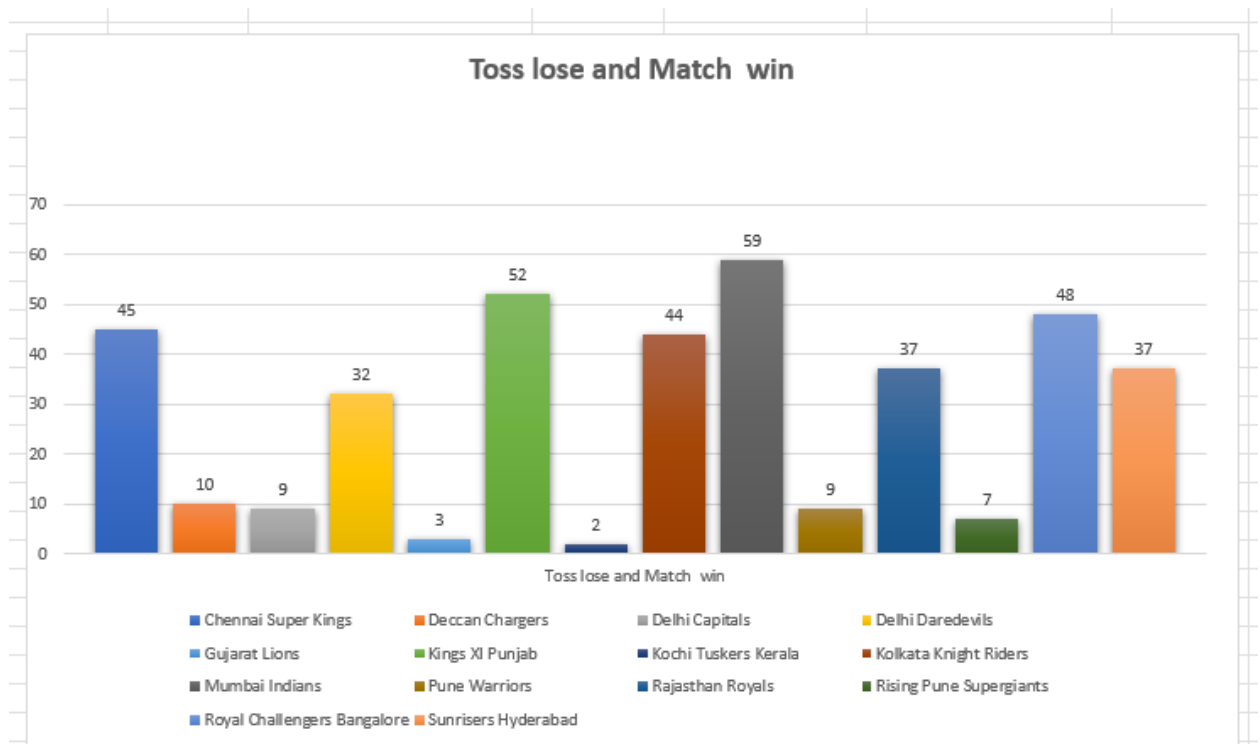




**Kings XI Punjab has greater effect by Toss winning and Match lose.**



**RCB lost more matches when toss is lost.**



**Delhi Capitals won more matches when toss is lost.**

### **5.Matches won by Team by Huge Margin :** (By runs and By wickets)

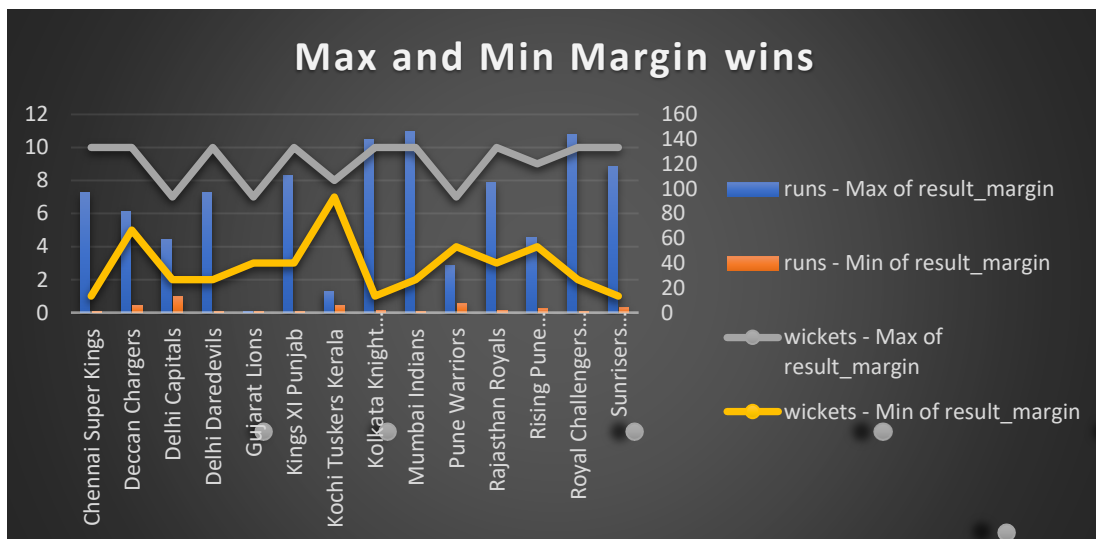
**a) Introduction:** The analysis shows the Teams who won matches By huge margins (i.e, by more runs and wickets) and the Teams who won matches by Low margin (i.e., by least runs) while defending. And the Teams who won matches by more wickets and a smaller number of wickets while chasing.

### **b) Specific Requirements/Functions and Formulas:**

- i. Pivot table of the data of Import
- ii. With the help of data Combined Chart (i.e., column and line chart) is plotted.

### **c) Analysis Results:**

Chennai Super Kings is the team who won by huge margin of 97 runs while defending, while chasing also it has won the match without losing a single wicket.

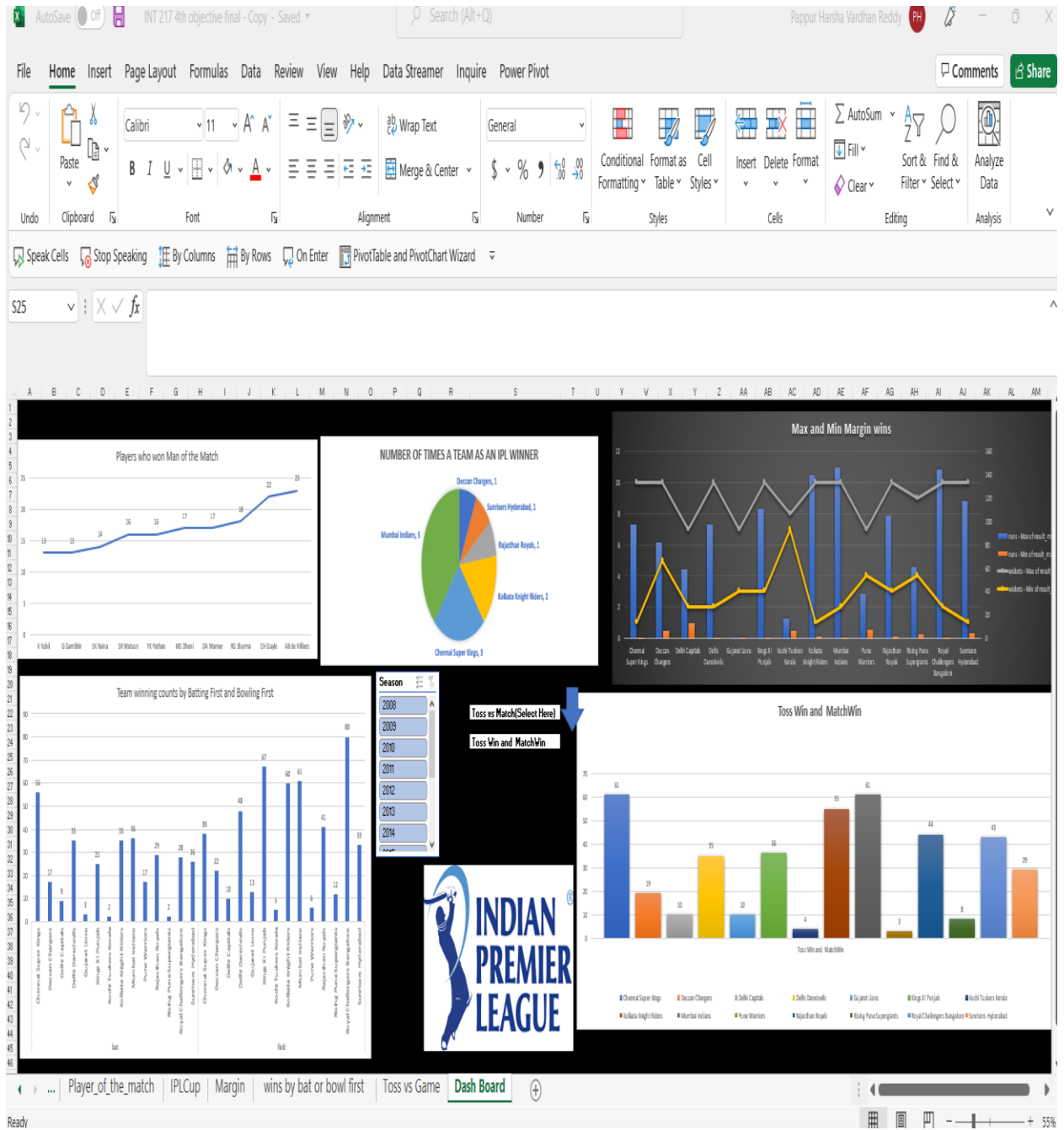




## **List of Analysis with results**

- AB de Villiers is the player who is a greater number of times Man of the match in all seasons and in all the cities.
- RCB is the team who won the greatest number of matches in all seasons and cities.
- Chennai Super Kings is the team who won by huge margin of 97 runs while defending, while chasing also it has won the match without losing a single wicket.
- Mumbai Indians had won a greater Number of Matches Winning Toss and By Loosing Toss.
- Royal Challengers Bangalore had Lost a greater number of matches By Loosing Toss.
- Kings XI Punjab had Lost many Matches although By Winning Toss.

# Dashboard



## **Conclusion**

From this subject I learned lot of things such that how to manage data, visualizing data, preparing data etc. Using excel, tableau how we can handle big data in these platforms.

This whole subject is containing lot of things such that ETL process in data warehouse, cleaning data, preparing data, visualizing data, formulas. Using these methods, we can manage; handle big data and visualize the data.

In this project I came to know that how to make attractive dashboard using pivot tables and pivot charts. Using these methods and functions we can do data management easily in near future.

## **BIBLIOGRAPHY:**

- Dataset source:

<https://www.kaggle.com/patrickb1912/ipl-complete-dataset-20082020>

- Information about Data Management:

<https://www.blue-pencil.ca/what-is-data-management-and-why-it-is-important/>

### **Excel formula references**

[https://www.youtube.com/watch?v=-](https://www.youtube.com/watch?v=-ujVQzTtxSg&list=PLWPirh4EWFpEpO6NjjWLbKSCb-wx3hMql)

[ujVQzTtxSg&list=PLWPirh4EWFpEpO6NjjWLbKSCb-wx3hMql](https://www.youtube.com/watch?v=-ujVQzTtxSg&list=PLWPirh4EWFpEpO6NjjWLbKSCb-wx3hMql)