

Annexure-I

**DATA SCIENCE MINOR PROJECT REPORT
REPORT**

Submitted in partial fulfilment of the requirements for the award of degree of

B.Tech (CSE)

Submitted to

LOVELY PROFESSIONAL UNIVERSITY

PHAGWARA, PUNJAB

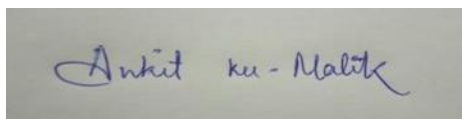


SUBMITTED BY

Name of student: Ankit Kumar Malik

Registration Number: 11904078

Signature of the student:

A photograph of a handwritten signature in blue ink on a light-colored surface. The signature reads 'Ankit Kumar Malik' in a cursive script.

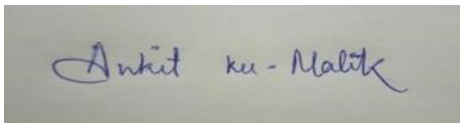
Annexure-II: Student Declaration

To whom so ever it may concern

I, **Ankit Kumar Malik, 11904078**, hereby declare that the work done by me on **“Data Science Minor Project Report”**, is a record of original work for the partial fulfilment of the requirements for the award of the degree, **B.Tech(CSE)**.

Ankit Kumar Malik (11904078)

Signature of the student

A photograph of a handwritten signature in blue ink on a light-colored surface. The signature reads "Ankit ku - Malik" in a cursive script.

Acknowledgement

The success and final outcome of this project required a lot of guidance and assistance from many people. All that I have done is only due to such supervision and assistance and I would not forget to thank them.

I respect and thank” Miss. Sandeep Kaur”, for providing me an opportunity to complete this project by giving me all the support and guidance which made me complete the project duly. I am extremely thankful to her for providing such a nice support and guidance.

I owe my deep gratitude to my project teacher Miss. Sandeep Kaur, who took keen interest on my project work and guided me all along, till the completion of my project work by providing all the necessary information.

I heartily thank her for their guidance and suggestions during this project work. I would not forget to remember their contribution on encouragement and more over for their support and guidance till completion of my project work.

Last but not least I'd rather thanks to Lovely Professional University, and my parent's inspiration, who gave me this golden opportunity to learn many new things, to learn another aspect of life.

Ankit Kumar Malik

CONTANTS:

Sr No.	Title	Page No.
1	Introduction	5
2	Objectives/Scope of the Analysis	7
3	Source of dataset	10
4	ETL Process	15
5	Analysis of dataset	18
6	List of Analysis with results	27
7	Final Dashboard	32
8	Bibliography	35

INTRODUCTION TO DATA SCIENCE MINOR PROJECT

1.1 INTRODUCTION

- Data management is important because the data your organization creates is a very valuable resource.
- The last thing you want to do is spend time and resources collecting data and business intelligence, only to lose or misplace that information.
- In that case, you would then have to spend time and resources again to get that same business intelligence you already had.
- And on that data analysis is carried out which show visualization of our problems in efficient way.
- Data Analysis is a process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information, informing conclusions, and supporting decision- making.
- This project is based on such data analysis on IPL data from 2008 to 2019.

Data science is the study of data to extract knowledge and insights from the data and apply knowledge and actionable insights. In this tutorial, we will work on IPL Data Analysis and Visualization Project using Python where we will explore interesting insights from the data of IPL matches like most run by a player, most wicket taken by a player, and much more from IPL season 2008-2020.

So if you are an IPL cricket fan and love data analysis with Python this project is perfect for you.

OBJECTIVES AND ANALYSIS:

- To find the team that won the most number of matches in a season.
- To find the team that lost the most number of matches in a season.
- To find cities the number of matches played were more.
- To find total match win by the team.
- Best performing team in IPL.
- Respective winning team by making toss decision to bat or ball.
- Total matches played by each team in IPL.
- To find the player with the most player of the match awards.
- To find the city that hosted the maximum number of IPL matches.
- To find the most winning team for each season.
- To find the top 10 "Power Hitter Batsmen" of the IPL.
- To find team scenario in their respective fields.
- Total home win percentage.
- Total away win percentage.

Since usually such tutorials are based on in-built datasets like google trends and tubule database, 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.

The topic I chose for visualization is to show the overview of teams performance and the supporting factor in Indian Premier League(IPL) cricket. I selected this topic because, cricket is like religion in India. By starting IPL, cricket fever has reached the pinnacle as IPL has introduced style and entertainment at the grounds, fans hear loud music, see fireworks everytime batsman hits a boundary or when bowler takes a wicket. Added to that, the name of every team is based on a state. This makes the citizen of that place to encourage and support them. Being an ardent of cricket, I wanted to find interesting factors about the teams performance and how it supports in winning a match.

At the end, I decided to check number of wins by each team over the years which further urge me to check teamwise winning record against all the other teams. Since winning is not only about the performance of players in a team there is an another underlying and important factor which is toss. I also decided to check if the toss is impacting the outcome of the match.

SOURCE OF DATABASE:

I got the data from kaggle.com where the matches.csv dataset contains the details of all the matches that was played from 2008 to 2019. There were totally 836 rows of data where each had the following details.

Source of dataset: <https://www.kaggle.com/ramjidoolla/ipl-data-set>

- ❖ Season number
- ❖ City where it was played
- ❖ Date of match
- ❖ Team1
- ❖ Team2
- ❖ Toss winner
- ❖ Toss decision
- ❖ DL applied
- ❖ Winner of the match
- ❖ Win by runs
- ❖ Win by wickets
- ❖ Player of the match
- ❖ Venue

❖ Umpires

Sample of dataset with data fields is given below:

Kolkata																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	id	Season	city	date	team1	team2	toss_winnr	toss_decis	result	dl_applied	winner	win_by_ru	win_by_wi	player_of	venue	umpire1	umpire2	umpire
2	1	IPL-2017	Hyderabad	#####	Sunrisers	I Royal Chal	Royal Chal	field	normal	0	Sunrisers	35	0	Yuvraj Sing	Rajiv Gand	AY Dandek	NJ Llong	
3	2	IPL-2017	Pune	#####	Mumbai	In Rising Puni	Rising Puni	field	normal	0	Rising Puni	0	7	SPD Smith	Maharashi	A Nand Kis	S Ravi	
4	3	IPL-2017	Rajkot	#####	Gujarat	Lic Kolkata	Kn Kolkata	Kn field	normal	0	Kolkata	Kn	0	10	CA Lynn	Saurashtra	Nitin Men	CK Nandan
5	4	IPL-2017	Indore	#####	Rising Puni	Kings XI Pu	Kings XI Pu	field	normal	0	Kings XI Pu	0	6	GJ Maxwe	Holkar Cric	AK Chaudh	C Shamshuddin	
6	5	IPL-2017	Bangalore	#####	Royal Chal	Delhi Dare	Royal Chal	bat	normal	0	Royal Chal	15	0	KM Jadhav	M Chinnaswamy	Stadium		
7	6	IPL-2017	Hyderabad	#####	Gujarat	Lic Sunrisers	I-Sunrisers	I field	normal	0	Sunrisers	I	0	9	Rashid Khe	Rajiv Gand	A Deshmul	NJ Llong
8	7	IPL-2017	Mumbai	#####	Kolkata	Kn Mumbai	In Mumbai	In field	normal	0	Mumbai	In	0	4	N Rana	Wankhede	Nitin Men	CK Nandan
9	8	IPL-2017	Indore	#####	Royal Chal	Kings XI Pu	Royal Chal	bat	normal	0	Kings XI Pu	0	8	AR Patel	Holkar Cric	AK Chaudh	C Shamshuddin	
10	9	IPL-2017	Pune	#####	Delhi Dare	Rising Puni	Rising Puni	field	normal	0	Delhi Dare	97	0	SV Samsor	Maharashi	AY Dandek	S Ravi	
11	10	IPL-2017	Mumbai	#####	Sunrisers	I Mumbai	In Mumbai	In field	normal	0	Mumbai	In	0	4	JJ Bumrah	Wankhede	Nitin Men	CK Nandan
12	11	IPL-2017	Kolkata	#####	Kings XI Pu	Kolkata	Kn Kolkata	Kn field	normal	0	Kolkata	Kn	0	8	SP Narine	Eden Gard	A Deshmul	NJ Llong
13	12	IPL-2017	Bangalore	#####	Royal Chal	Mumbai	In Mumbai	In field	normal	0	Mumbai	In	0	4	KA Pollard	M Chinnas	KN Ananth	AK Chaudhary
14	13	IPL-2017	Rajkot	#####	Rising Puni	Gujarat	Lic Gujarat	Lic field	normal	0	Gujarat	Lic	0	7	AJ Tye	Saurashtra	A Nand Kis	S Ravi
15	14	IPL-2017	Kolkata	#####	Kolkata	Kn Sunrisers	I-Sunrisers	I field	normal	0	Kolkata	Kn	17	0	RV Uthappa	Eden Gard	AY Dandek	NJ Llong
16	15	IPL-2017	Delhi	#####	Delhi Dare	Kings XI Pu	Delhi Dare	bat	normal	0	Delhi Dare	51	0	CJ Andersc	Feroz Sha	YC Barde	Nitin Menon	
17	16	IPL-2017	Mumbai	#####	Gujarat	Lic Mumbai	In Mumbai	In field	normal	0	Mumbai	In	0	6	N Rana	Wankhede	A Nand Kis	S Ravi
18	17	IPL-2017	Bangalore	#####	Rising Puni	Royal Chal	Royal Chal	field	normal	0	Rising Puni	27	0	BA Stokes	M Chinnas	KN Ananth	C Shamshuddin	
19	18	IPL-2017	Delhi	#####	Delhi Dare	Kolkata	Kn Delhi	Dare bat	normal	0	Kolkata	Kn	0	4	NM Coult	Feroz Sha	Nitin Men	CK Nandan
20	19	IPL-2017	Hyderabad	#####	Sunrisers	I-Kings XI Pu	Kings XI Pu	field	normal	0	Sunrisers	I	5	0	B Kumar	Rajiv Gand	AY Dandek	A Deshmukh
21	20	IPL-2017	Rajkot	#####	Royal Chal	Gujarat	Lic Gujarat	Lic field	normal	0	Royal Chal	21	0	CH Gayle	Saurashtra	S Ravi	VK Sharma	
22	21	IPL-2017	Hyderabad	#####	Sunrisers	I-Delhi Dare	Sunrisers	I-bat	normal	0	Sunrisers	I	15	0	KS William	Rajiv Gand	CB Gaffan	NJ Llong
23	22	IPL-2017	Indore	#####	Kings XI Pu	Mumbai	In Mumbai	In field	normal	0	Mumbai	In	0	8	JC Buttler	Holkar Cric	M Erasmus	C Shamshuddin
24	23	IPL-2017	Kolkata	#####	Kolkata	Kn Gujarat	Lic Gujarat	Lic field	normal	0	Gujarat	Lic	0	4	SK Raina	Eden Gard	CB Gaffan	Nitin Menon
25	24	IPL-2017	Mumbai	#####	Mumbai	In Delhi Dare	Delhi Dare	field	normal	0	Mumbai	In	14	0	MJ McCler	Wankhede	A Nand Kis	S Ravi
26	25	IPL-2017	Pune	#####	Sunrisers	I-Rising Puni	Rising Puni	field	normal	0	Rising Puni	0	6	MS Dhoni	Maharashi	AY Dandek	A Deshmukh	
27	26	IPL-2017	Rajkot	#####	Kings XI Pu	Gujarat	Lic Gujarat	Lic field	normal	0	Kings XI Pu	26	0	HM Amla	Saurashtra	AK Chaudh	M Erasmus	

Fig 1

Since I decided to visualize performance of the team and to find the relation between match outcome and toss outcome, I had to do some data transformations. I created a new excel data based on the matches.csv dataset where I created 4 columns: Team, TW(Tosses won), MW(Matches won based on toss), Count(Total number of wins

irrespective of the toss). After filtering the matches.csv, I got all the necessary data in data. I also created one more excel named TeamwiseInformation.tsv where I created fourteen columns. First two columns were roll and state(representing each team), remaining 12 columns had 12 distinct team names. This dataset contains information about number of wins by each team against all the other teams in IPL. Using TeamwiseInformation.tsv I will be able to visualize the teamwise record for all the teams. Using data.tsv, I will be able to visualize teams performance over the years and I can also find if there is a correlation between matches won and tosses won.

For visualization, I used three charts. Two bar chart and one scatterplot. First bar chart shows the number of wins by each and every team over the years in the history of IPL. Since I used bar chart for this, the exact number will not be known by looking at the chart alone. By hovering every single rectangle bar, we will get to know the number of wins by each team. Second bar chart shows the teamwise record which is basically an annotation of first chart.

Since there are 12 different teams, I decided to make it interactive by placing a dropdown selector where the dropdown will have 12 teams from which we can select any team of our choice to know the stats.

Bar chart will be updated everytime when different team is selected. For

example: if Chennai Super Kings is selected, bar chart will be display multiple bars which represents the number of wins by Chennai Super Kings(CSK) against every other team in IPL. Hovering the each bar will give exact number of wins against that particular team. The third and final scatterplot shows the supporting factor for the previous two charts. If we look at the first and second bar chart, Mumbai Indians(MI) looks like a most dominating team over the years. I plotted matches and tosses won in the scatterplot. Each dot in the scatterplot has different colors and hovering them will provide team name, number of tosses and matches won by them. I also placed a color legend to know the teams name.

Once after this, I found that there is a strong positive correlation between the toss and match outcome. When the number of tosses won by each team increases, number of matches won by them also increases. It is clearly evident from the scatterplot that toss plays a crucial role in teams overall performance as well as the outcome of the match.

	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	id	Season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player
2	1	IPL-2017	Hyderabad	#####	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj S
3	2	IPL-2017	Pune	#####	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPD Sm
4	3	IPL-2017	Rajkot	#####	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	CA Lym
5	4	IPL-2017	Indore	#####	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	GJ Max
6	5	IPL-2017	Bangalore	#####	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Jad
7	6	IPL-2017	Hyderabad	#####	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad	0	9	Rashid
8	7	IPL-2017	Mumbai	#####	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	N Rana
9	8	IPL-2017	Indore	#####	Royal Challengers Bangalore	Kings XI Punjab	Royal Challengers Bangalore	bat	normal	0	Kings XI Punjab	0	8	AR Pat
10	9	IPL-2017	Pune	#####	Delhi Daredevils	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Delhi Daredevils	97	0	SV Sam
11	10	IPL-2017	Mumbai	#####	Sunrisers Hyderabad	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	JJ Bum
12	11	IPL-2017	Kolkata	#####	Kings XI Punjab	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	8	SP Nar
13	12	IPL-2017	Bangalore	#####	Royal Challengers Bangalore	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	KA Poll
14	13	IPL-2017	Rajkot	#####	Rising Pune Supergiant	Gujarat Lions	Gujarat Lions	field	normal	0	Gujarat Lions	0	7	AI Tye
15	14	IPL-2017	Kolkata	#####	Kolkata Knight Riders	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Kolkata Knight Riders	17	0	RV Uth
16	15	IPL-2017	Delhi	#####	Delhi Daredevils	Kings XI Punjab	Delhi Daredevils	bat	normal	0	Delhi Daredevils	51	0	CJ And
17	16	IPL-2017	Mumbai	#####	Gujarat Lions	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	6	N Rana
18	17	IPL-2017	Bangalore	#####	Rising Pune Supergiant	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Rising Pune Supergiant	27	0	BA Sto
19	18	IPL-2017	Delhi	#####	Delhi Daredevils	Kolkata Knight Riders	Delhi Daredevils	bat	normal	0	Kolkata Knight Riders	0	4	NM Co
20	19	IPL-2017	Hyderabad	#####	Sunrisers Hyderabad	Kings XI Punjab	Kings XI Punjab	field	normal	0	Sunrisers Hyderabad	5	0	B Kuma
21	20	IPL-2017	Rajkot	#####	Royal Challengers Bangalore	Gujarat Lions	Gujarat Lions	field	normal	0	Royal Challengers Bangalore	21	0	CH Gay
22	21	IPL-2017	Hyderabad	#####	Sunrisers Hyderabad	Delhi Daredevils	Sunrisers Hyderabad	bat	normal	0	Sunrisers Hyderabad	15	0	KS Will
23	22	IPL-2017	Indore	#####	Kings XI Punjab	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	8	JC Butt
24	23	IPL-2017	Kolkata	#####	Kolkata Knight Riders	Gujarat Lions	Gujarat Lions	field	normal	0	Gujarat Lions	0	4	SK Rain
25	24	IPL-2017	Mumbai	#####	Mumbai Indians	Delhi Daredevils	Delhi Daredevils	field	normal	0	Mumbai Indians	14	0	MJ Mc
26	25	IPL-2017	Pune	#####	Sunrisers Hyderabad	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	6	MS Dh
27	26	IPL-2017	Rajkot	#####	Kings XI Punjab	Gujarat Lions	Gujarat Lions	field	normal	0	Kings XI Punjab	26	0	HM Am

Fig 2.2

ETL PROCESS:

- ETL is a process that extracts the data from different source systems, then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system.
- Full form of ETL is Extract, Transform and Load.
- The triple combination of ETL provides crucial functions that are many times combined into a single application or suite of tools that help in the following areas:
 - ❖ Enhances Business Intelligence solutions for decision making.
 - ❖ Allows verification of data transformation, aggregation and calculations rules.
 - ❖ Allows sample data comparison between source and target system.
 - ❖ Helps to improve productivity as it codifies and reuses without additional technical skills.

❖ 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	R
1	id	Season	city	date	team1	team2	toss_winn	toss_decis	result	dl_applied	winner	win_by_ru	win_by_w	player_of	venue	umpire1	umpire2	umpire
2	1	IPL-2017	Hyderabad	#####	Sunrisers I	Royal Chal	Royal Chal	field	normal	0	Sunrisers I	35	0	Yuvraj Sing	Rajiv Gand	AY Dandek	NJ Llong	
3	2	IPL-2017	Pune	#####	Mumbai In	Rising Puni	Rising Puni	field	normal	0	Rising Puni	0	7	SPD Smith	Maharashtr	A Nand Kis	S Ravi	
4	3	IPL-2017	Rajkot	#####	Gujarat Lic	Kolkata Kn	Kolkata Kn	field	normal	0	Kolkata Kn	0	10	CA Lynn	Saurashtra	Nitin Meni	CK Nandan	
5	4	IPL-2017	Indore	#####	Rising Puni	Kings XI Pu	Kings XI Pu	field	normal	0	Kings XI Pu	0	6	GJ Maxwe	Holkar Cric	AK Chaudh	C Shamshuddin	
6	5	IPL-2017	Bangalore	#####	Royal Chal	Delhi Dare	Royal Chal	bat	normal	0	Royal Chal	15	0	KM Jadhav	M Chinnaswamy	Stadium		
7	6	IPL-2017	Hyderabad	#####	Gujarat Lic	Sunrisers I	Sunrisers I	field	normal	0	Sunrisers I	0	9	Rashid Khe	Rajiv Gand	A Deshmul	NJ Llong	
8	7	IPL-2017	Mumbai	#####	Kolkata Kn	Mumbai In	Mumbai In	field	normal	0	Mumbai In	0	4	N Rana	Wankhede	Nitin Meni	CK Nandan	
9	8	IPL-2017	Indore	#####	Royal Chal	Kings XI Pu	Royal Chal	bat	normal	0	Kings XI Pu	0	8	AR Patel	Holkar Cric	AK Chaudh	C Shamshuddin	
10	9	IPL-2017	Pune	#####	Delhi Dare	Rising Puni	Rising Puni	field	normal	0	Delhi Dare	97	0	SV Samsor	Maharashtr	AY Dandek	S Ravi	
11	10	IPL-2017	Mumbai	#####	Sunrisers I	Mumbai In	Mumbai In	field	normal	0	Mumbai In	0	4	JJ Bumrah	Wankhede	Nitin Meni	CK Nandan	
12	11	IPL-2017	Kolkata	#####	Kings XI Pu	Kolkata Kn	Kolkata Kn	field	normal	0	Kolkata Kn	0	8	SP Narine	Eden Gard	A Deshmul	NJ Llong	
13	12	IPL-2017	Bangalore	#####	Royal Chal	Mumbai In	Mumbai In	field	normal	0	Mumbai In	0	4	KA Pollard	M Chinnas	KN Ananth	AK Chaudhary	
14	13	IPL-2017	Rajkot	#####	Rising Puni	Gujarat Lic	Gujarat Lic	field	normal	0	Gujarat Lic	0	7	AI Tye	Saurashtra	A Nand Kis	S Ravi	
15	14	IPL-2017	Kolkata	#####	Kolkata Kn	Sunrisers I	Sunrisers I	field	normal	0	Kolkata Kn	17	0	RV Uthappa	Eden Gard	AY Dandek	NJ Llong	
16	15	IPL-2017	Delhi	#####	Delhi Dare	Kings XI Pu	Delhi Dare	bat	normal	0	Delhi Dare	51	0	CJ Andersc	Feroz Sha	YC Barde	Nitin Menon	
17	16	IPL-2017	Mumbai	#####	Gujarat Lic	Mumbai In	Mumbai In	field	normal	0	Mumbai In	0	6	N Rana	Wankhede	A Nand Kis	S Ravi	
18	17	IPL-2017	Bangalore	#####	Rising Puni	Royal Chal	Royal Chal	field	normal	0	Rising Puni	27	0	BA Stokes	M Chinnas	KN Ananth	C Shamshuddin	
19	18	IPL-2017	Delhi	#####	Delhi Dare	Kolkata Kn	Delhi Dare	bat	normal	0	Kolkata Kn	0	4	NM Coult	Feroz Sha	Nitin Meni	CK Nandan	
20	19	IPL-2017	Hyderabad	#####	Sunrisers I	Kings XI Pu	Kings XI Pu	field	normal	0	Sunrisers I	5	0	B Kumar	Rajiv Gand	AY Dandek	A Deshmukh	
21	20	IPL-2017	Rajkot	#####	Royal Chal	Gujarat Lic	Gujarat Lic	field	normal	0	Royal Chal	21	0	CH Gayle	Saurashtra	S Ravi	VK Sharma	
22	21	IPL-2017	Hyderabad	#####	Sunrisers I	Delhi Dare	Sunrisers I	bat	normal	0	Sunrisers I	15	0	KS William	Rajiv Gand	CB Gaffan	NJ Llong	
23	22	IPL-2017	Indore	#####	Kings XI Pu	Mumbai In	Mumbai In	field	normal	0	Mumbai In	0	8	JC Buttler	Holkar Cric	M Erasmus	C Shamshuddin	
24	23	IPL-2017	Kolkata	#####	Kolkata Kn	Gujarat Lic	Gujarat Lic	field	normal	0	Gujarat Lic	0	4	SK Raina	Eden Gard	CB Gaffan	Nitin Menon	
25	24	IPL-2017	Mumbai	#####	Mumbai In	Delhi Dare	Delhi Dare	field	normal	0	Mumbai In	14	0	MJ McCler	Wankhede	A Nand Kis	S Ravi	
26	25	IPL-2017	Pune	#####	Sunrisers I	Rising Puni	Rising Puni	field	normal	0	Rising Puni	0	6	MS Dhoni	Maharashtr	AY Dandek	A Deshmukh	
27	26	IPL-2017	Rajkot	#####	Kings XI Pu	Gujarat Lic	Gujarat Lic	field	normal	0	Kings XI Pu	26	0	HM Amla	Saurashtra	AK Chaudh	M Erasmus	

❖ After deleting the unnecessary column and shorting the data sheet

Column which are removed from dataset are given below

- Date
- Results
- Umpire 3

❖ Now to remove null values we first sort the dataset show that null values comes on the top and after that we can delete it through Delete rows from Home tab.

❖ Sort the dataset through filter of rating data fields

Finally, after cleaning the data, the final dataset sample is shown

below:

	A	B	C	D	E	F	G	H	I	J	K	L
2	1 IPL-2017	Hyderabad	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field		0	Sunrisers Hyderabad	35	0	Yuvraj Singh
3	2 IPL-2017	Pune	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field		0	Rising Pune Supergiant	0	7	SPD Smith
4	3 IPL-2017	Rajkot	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field		0	Kolkata Knight Riders	0	10	CA Lynn
5	4 IPL-2017	Indore	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field		0	Kings XI Punjab	0	6	GJ Maxwell
6	5 IPL-2017	Bangalore	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat		0	Royal Challengers Bangalore	15	0	KM Jadhav
7	6 IPL-2017	Hyderabad	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field		0	Sunrisers Hyderabad	0	9	Rashid Khan
8	7 IPL-2017	Mumbai	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field		0	Mumbai Indians	0	4	N Rana
9	8 IPL-2017	Indore	Royal Challengers Bangalore	Kings XI Punjab	Royal Challengers Bangalore	bat		0	Kings XI Punjab	0	8	AR Patel
10	9 IPL-2017	Pune	Delhi Daredevils	Rising Pune Supergiant	Rising Pune Supergiant	field		0	Delhi Daredevils	97	0	SV Samson
11	10 IPL-2017	Mumbai	Sunrisers Hyderabad	Mumbai Indians	Mumbai Indians	field		0	Mumbai Indians	0	4	JJ Bumrah
12	11 IPL-2017	Kolkata	Kings XI Punjab	Kolkata Knight Riders	Kolkata Knight Riders	field		0	Kolkata Knight Riders	0	8	SP Narine
13	12 IPL-2017	Bangalore	Royal Challengers Bangalore	Mumbai Indians	Mumbai Indians	field		0	Mumbai Indians	0	4	KA Pollard
14	13 IPL-2017	Rajkot	Rising Pune Supergiant	Gujarat Lions	Gujarat Lions	field		0	Gujarat Lions	0	7	AJ Tye
15	14 IPL-2017	Kolkata	Kolkata Knight Riders	Sunrisers Hyderabad	Sunrisers Hyderabad	field		0	Kolkata Knight Riders	17	0	RV Uthappa
16	15 IPL-2017	Delhi	Delhi Daredevils	Kings XI Punjab	Delhi Daredevils	bat		0	Delhi Daredevils	51	0	CJ Anderson
17	16 IPL-2017	Mumbai	Gujarat Lions	Mumbai Indians	Mumbai Indians	field		0	Mumbai Indians	0	6	N Rana
18	17 IPL-2017	Bangalore	Rising Pune Supergiant	Royal Challengers Bangalore	Royal Challengers Bangalore	field		0	Rising Pune Supergiant	27	0	BA Stokes
19	18 IPL-2017	Delhi	Delhi Daredevils	Kolkata Knight Riders	Delhi Daredevils	bat		0	Kolkata Knight Riders	0	4	NM Coulter-Nile
20	19 IPL-2017	Hyderabad	Sunrisers Hyderabad	Kings XI Punjab	Kings XI Punjab	field		0	Sunrisers Hyderabad	5	0	B Kumar
21	20 IPL-2017	Rajkot	Royal Challengers Bangalore	Gujarat Lions	Gujarat Lions	field		0	Royal Challengers Bangalore	21	0	CH Gayle
22	21 IPL-2017	Hyderabad	Sunrisers Hyderabad	Delhi Daredevils	Sunrisers Hyderabad	bat		0	Sunrisers Hyderabad	15	0	KS Williamson
23	22 IPL-2017	Indore	Kings XI Punjab	Mumbai Indians	Mumbai Indians	field		0	Mumbai Indians	0	8	JC Buttler
24	23 IPL-2017	Kolkata	Kolkata Knight Riders	Gujarat Lions	Gujarat Lions	field		0	Gujarat Lions	0	4	SK Raina
25	24 IPL-2017	Mumbai	Mumbai Indians	Delhi Daredevils	Delhi Daredevils	field		0	Mumbai Indians	14	0	MJ McClenaghan
26	25 IPL-2017	Pune	Sunrisers Hyderabad	Rising Pune Supergiant	Rising Pune Supergiant	field		0	Rising Pune Supergiant	0	6	MS Dhoni
27	26 IPL-2017	Rajkot	Kings XI Punjab	Gujarat Lions	Gujarat Lions	field		0	Kings XI Punjab	26	0	HM Amla
28	27 IPL-2017	Kolkata	Kolkata Knight Riders	Royal Challengers Bangalore	Royal Challengers Bangalore	field		0	Kolkata Knight Riders	82	0	NM Coulter-Nile
29	28 IPL-2017	Mumbai	Rising Pune Supergiant	Mumbai Indians	Mumbai Indians	field		0	Rising Pune Supergiant	3	0	BA Stokes
30	29 IPL-2017	Pune	Rising Pune Supergiant	Kolkata Knight Riders	Kolkata Knight Riders	field		0	Kolkata Knight Riders	0	7	RV Uthappa
31	30 IPL-2017	Bangalore	Royal Challengers Bangalore	Gujarat Lions	Gujarat Lions	field		0	Gujarat Lions	0	7	AJ Tye
32	31 IPL-2017	Kolkata	Delhi Daredevils	Kolkata Knight Riders	Kolkata Knight Riders	field		0	Kolkata Knight Riders	0	7	G Gambhir

Fig 4

ANALYSIS OF DATA

MATCHES PLAYED

1. The number of IPL matches played in respective cities

- The idiom I used for this is bar chart where x-axis is mapped to different cities which is categorical attribute and y-axis is mapped to count which is quantitative attribute represents the number of matches played by each team. I chose bar chart among all the choices of idiom because bar chart is the best way to compare things between different groups. I sorted the categorical attribute based on most number of matches played by each team from highest to lowest to show the successful teams in IPL history.

Mark - Line

Channel - Vertical spatial position for quantitative and horizontal spatial position for categorical.

Color - Used hex color #065A53 initially for all the bars and #B2EDB7 is used when it is hovered. I selected two different colors to show the difference between the move hover and normal chart.

As a result, I was able to make a neat bar chart that clearly shows the overall winning performance of all the IPL teams from highest to lowest.

Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Count function is used in pivot table for the counting of the number of matches played in particular cities in the pivot table in their respective year.

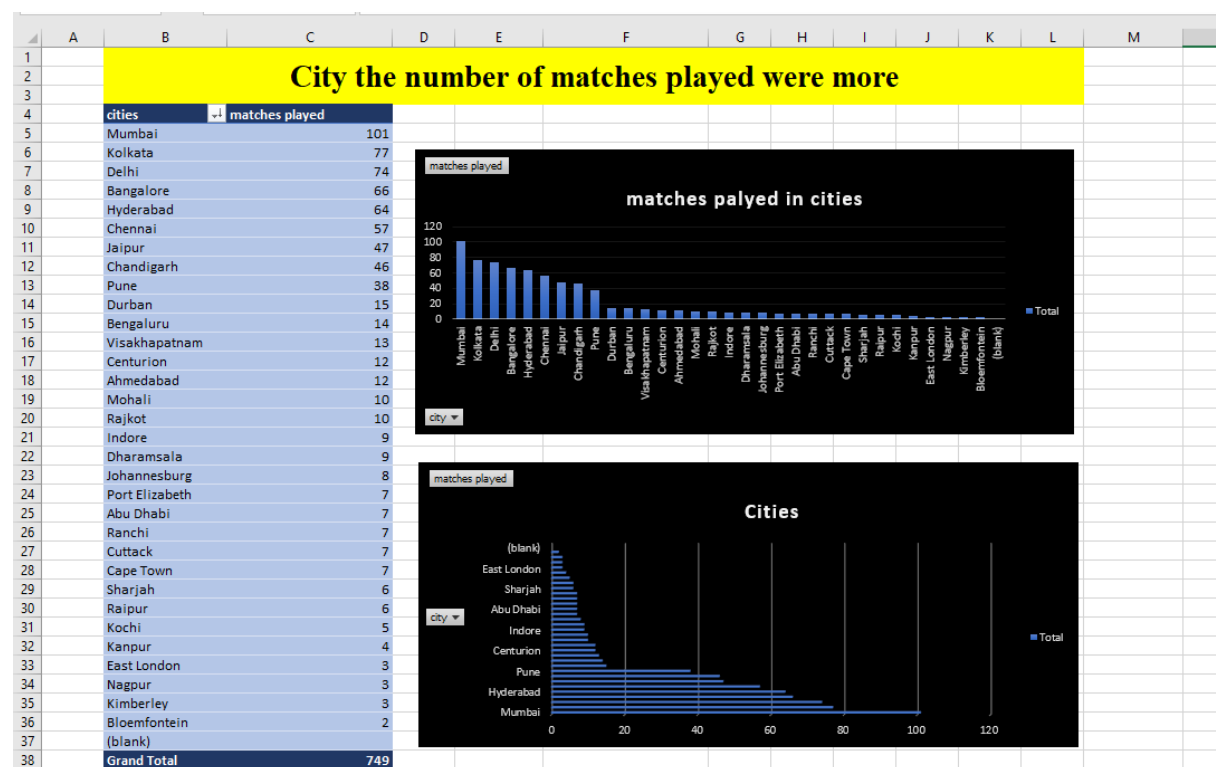


Fig 5

2. The number of IPL matches played in a respective season

- The idiom I used for this is bar chart where x-axis is mapped to different season year wise which is categorical attribute and y-axis is mapped to count which is quantitative attribute represents the number of matches played by each team. I chose line chart among all the choices of idiom because bar chart is the best way to compare things between different groups. I sorted the categorical attribute based on most number of matches played by each team from highest to lowest to show the successful teams in IPL history.
- Mark – points
- Channel - Vertical spatial position for quantitative and horizontal spatial position for categorical.
- Colour - Used hex colour #065A53 initially for all the line and #B2EDB7 is used when it is hovered. I selected two colours to show the difference between the move hover and normal chart.
- As a result, I was able to make a neat bar chart that clearly shows the overall winning performance of all the IPL teams from highest to lowest.

Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Count function is used in pivot table for the counting of the number of matches played per season in the pivot table in their respective year.

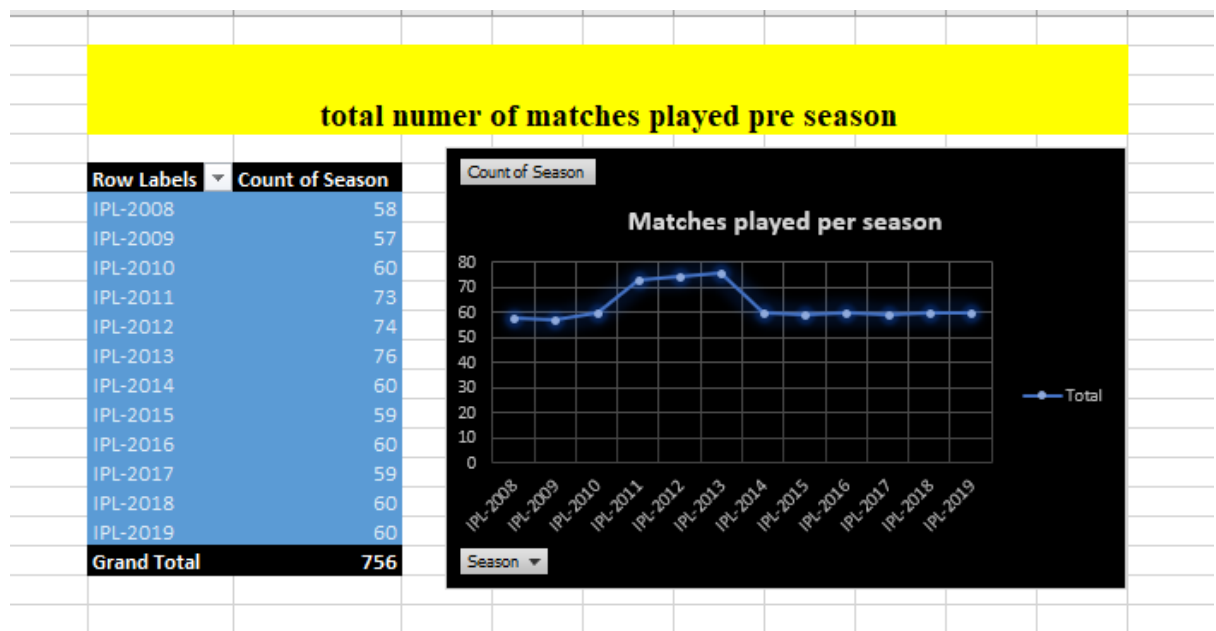


Fig 6

3. Respective winning team by making toss decision to bat or ball

The idiom I used for this is pie chart where the blue part is mapped to tosses won and chose to field which is quantitative attribute and red part is mapped to toss won and chose to bat which is also a quantitative attribute. Used colour legend which is displayed right of

the pie chart representing the team based on the colour choosing to bat or field.

Mark - Points

Channel - Uses spatial position and colour

Colour - Used different colours blue for toss decision to field and colour for toss decision to bat for different teams which can be referred using a colour legend

Using pie chart is the only best way to show the correlation between two different attributes and I did it successfully in my chart

Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Count function is used in pivot table for the counting of the number of toss made for making a decision to field or bat in the pivot table.

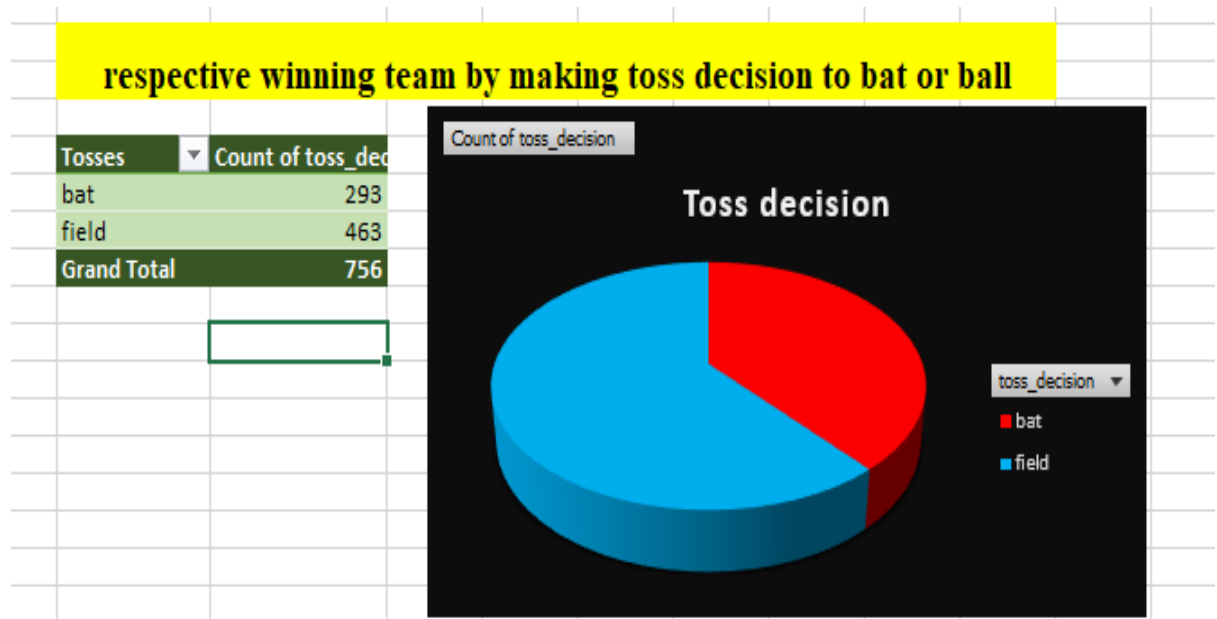


Fig 7

4. Total matches win by the team

The idiom I used for this is pie chart where I have used different colours and chose to field which is quantitative attribute for respective teams which is also a quantitative attribute. Used colour legend which is displayed right of the pie chart representing the team based on the colour shows the matches won by the single team.

Mark - Points

Channel - Uses spatial position and colour

Colour - Used different colours blue for toss decision to field and colour for toss decision to bat for different teams which can be referred using a colour legend

Using pie chart is the only best way to show the correlation between two different attributes and I did it successfully in my char.

Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Count function is used in pivot table for the counting of the number of matches win by the team in the pivot table in their respective year.

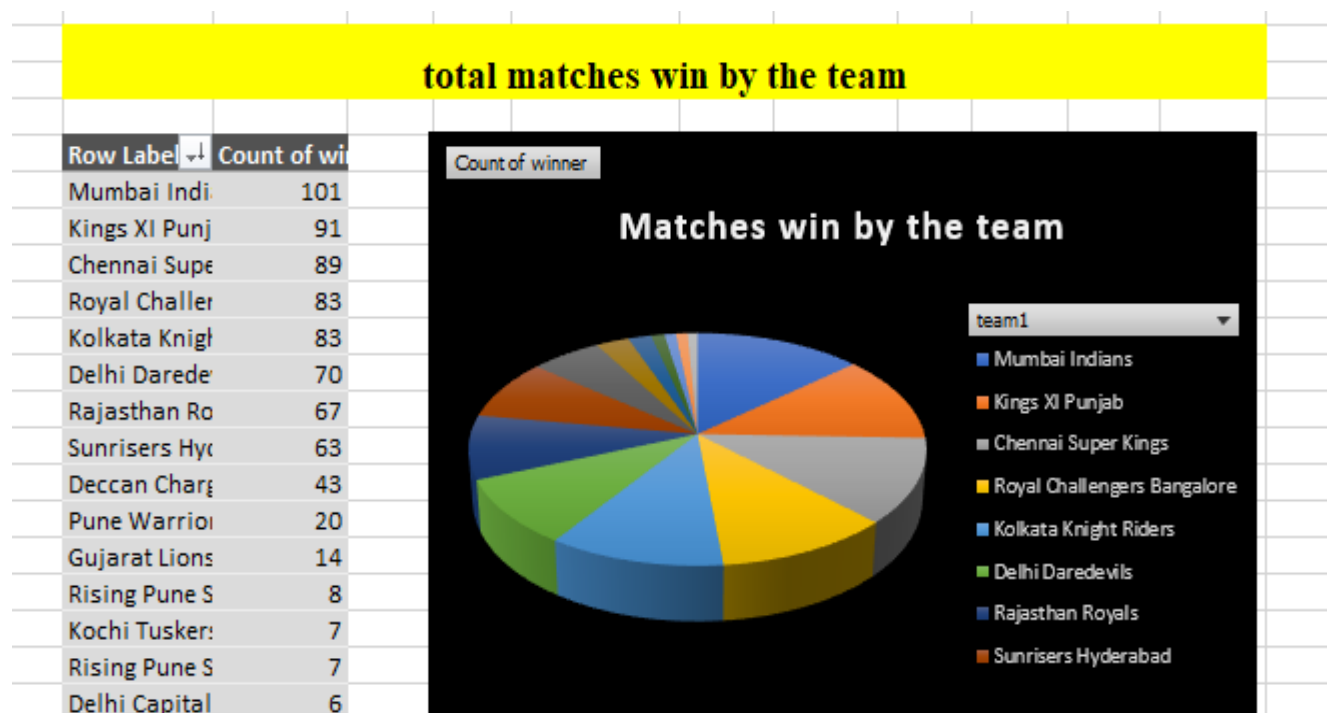


Fig 8

5. Top 10 players in IPL

The idiom I used for this is bar chart where x-axis is mapped to different players name which is categorical attribute and y-axis is mapped to count which is quantitative attribute represents the number award win by respective player of matches played by each team. I chose bar chart among all the choices of idiom because bar chart is the best way to compare things between different groups.

I sorted the categorical attribute based on most number of matches played by each team from highest to lowest to show the successful teams in IPL history.

Mark - Line

Channel - Vertical spatial position for quantitative and horizontal spatial position for categorical.

Colour - Used colour Is purple initially for all the bars and dark blue used when it is hovered. I selected two different colours to show the difference between the move hover and normal chart.

Specific requirements, functions and formulas

- Pivot table is used for the analysis.
- Count function is used in pivot table for the counting of the number of top 10 players of the match and most award winning player in IPL in the pivot table in their respective year.

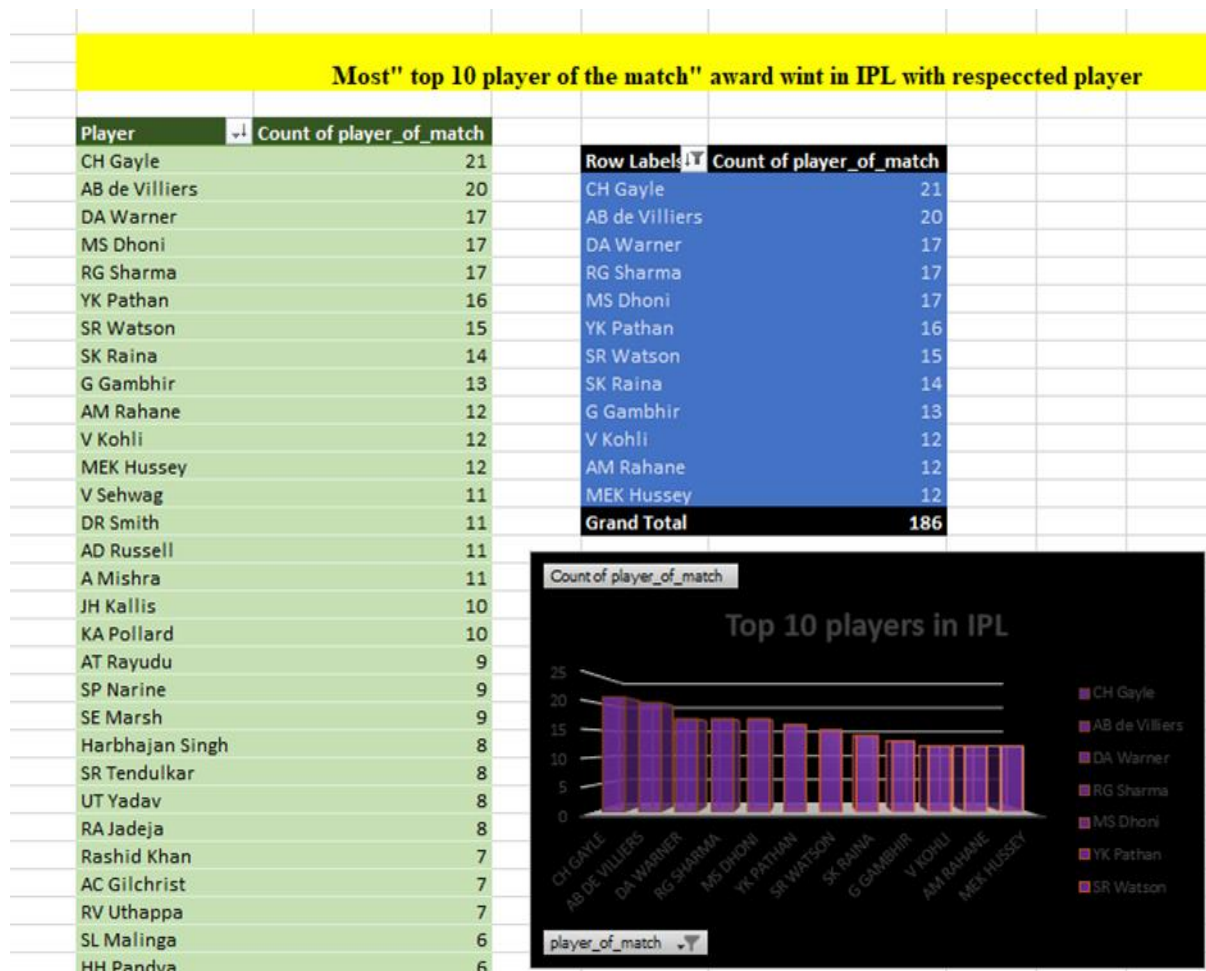


Fig 9

List of Analysis with results

1. Total matches played by the team

The most matches played IPL team is the team that has won most number of times. It can be calculated by the summing up the number of home matches played and the number of away matches played by the team.

Specific requirements, functions and formulas

- ☐ Pivot table is used for the analysis.
- ☐ Sum function is used in pivot table for the summing of the number of home matches and the number of away matches played by the team in the pivot table in their respective year.

Total matches played by the team				
Row Labels	Sum of home_matches	Sum of away_matches	Sum of Total_matches	
Mumbai Indians	101	86	187	
Royal Challengers Bangalore	85	95	180	
Kolkata Knight Riders	83	95	178	
Kings XI Punjab	91	85	176	
Chennai Super Kings	89	75	164	
Delhi Daredevils	72	89	161	
Rajasthan Royals	67	80	147	
Sunrisers Hyderabad	63	45	108	
Deccan Chargers	43	32	75	
Pune Warriors	20	26	46	
Gujarat Lions	14	16	30	
Delhi Capitals	6	10	16	
Rising Pune Supergiant	8	8	16	
Kochi Tuskers Kerala	7	7	14	
Grand Total	749	749	1498	

Fig 10

2. Total matches win by the team

The most matches winning team in IPL can be calculated by the summing up the number of home wins and the number away wins by the team.

Specific requirements, functions and formulas

- Pivot table is used for the analysis.
- Sum function is used in pivot table for the summing of the number of home wins and the number of away wins by the team in the pivot table in their respective year.

Total matches win by the teams			
Row Labels	Sum of home_wins	Sum of away_wins	Sum of Total_win
Mumbai Indians	58	51	109
Chennai Super Kings	51	49	100
Kolkata Knight Riders	34	58	92
Royal Challengers Bangalore	35	49	84
Kings XI Punjab	38	44	82
Rajasthan Royals	29	46	75
Delhi Daredevils	25	42	67
Sunrisers Hyderabad	30	28	58
Deccan Chargers	18	11	29
Gujarat Lions	1	12	13
Pune Warriors	6	6	12
Rising Pune Supergiant	5	5	10
Delhi Capitals	3	7	10
Kochi Tuskers Kerala	2	4	6
Grand Total	335	412	747

Fig 11

3. Total home win percentage

The total home win percentage in IPL teams can be calculated by the finding the percentile of home win and home matches played by the respective teams.

Specific requirements, functions and formulas

- Pivot table is used for the analysis.
- Sum function is used in pivot table for the summing of the number of home wins and the number of away wins by the team in the pivot table in their respective year.

Total home_win percentage				
Row Labels	_home_wins	_Home_matches	_Home win %	
Rising Pune Supergiant	5	8	62.5	
Mumbai Indians	58	101	57.42574257	
Chennai Super Kings	51	89	57.30337079	
Delhi Capitals	3	6	50	
Sunrisers Hyderabad	30	63	47.61904762	
Rajasthan Royals	29	67	43.28358209	
Deccan Chargers	18	43	41.86046512	
Kings XI Punjab	38	91	41.75824176	
Royal Challengers Bangalore	35	85	41.17647059	
Kolkata Knight Riders	34	83	40.96385542	
Delhi Daredevils	25	72	34.72222222	
Pune Warriors	6	20	30	
Kochi Tuskers Kerala	2	7	28.57142857	
Gujarat Lions	1	14	7.142857143	
Grand Total	335	749	44.72630174	

4. Total away win percentage

The total away win percentage in IPL teams can be calculated by the finding the percentile of away win and away matches played by the respective teams.

Specific requirements, functions and formulas

- Pivot table is used for the analysis.
- Sum function is used in pivot table for the summing of the number of away wins and the number of away matches played by the team in the pivot table in their respective year.

Total away_ win percentge			
Row Labels	Sum of away_wins	Sum of away_match	Sum of away win%
Gujarat Lions	12	16	75
Delhi Capitals	7	10	70
Chennai Super Kings	49	75	65.33333333
Rising Pune Supergiant	5	8	62.5
Sunrisers Hyderabad	28	45	62.22222222
Kolkata Knight Riders	58	95	61.05263158
Mumbai Indians	51	86	59.30232558
Rajasthan Royals	46	80	57.5
Kochi Tuskers Kerala	4	7	57.14285714
Kings XI Punjab	44	85	51.76470588
Royal Challengers Bangalore	49	95	51.57894737
Delhi Daredevils	42	89	47.19101124
Deccan Chargers	11	32	34.375
Pune Warriors	6	26	23.07692308
Grand Total	412	749	55.00667557

Fig 12

PLAYER DETAILS

1. Total numbers of player in the country

Total numbers of player in the country can be calculated by counting by the number players from the respective country.

Specific requirements, functions and formulas

- Pivot table is used for the analysis.
- count function is used in pivot table for the counting of the number of player from respective country.
- Bar chart is used representing the number of player from the country.

Total no. of players in country		
	country	players
	India	264
	Australia	72
	South Africa	39
	New Zealand	22
	Sri Lanka	20
	West Indies	19
	England	14
	Pakistan	13
	Bangladesh	5
	Zimbabwe	2
	Netherlands	1
	(blank)	
	Grand Total	471

Fig 13

2. Bowling skill player in respective country

Total numbers of bowling skill player in the country can be calculated by counting by the number players from the respective country.

Specific requirements, functions and formulas

- Pivot table is used for the analysis.
- count function is used in pivot table for the counting of the number of bowling skill player from respective country.
- Bar chart is used representing the number of player from the country.

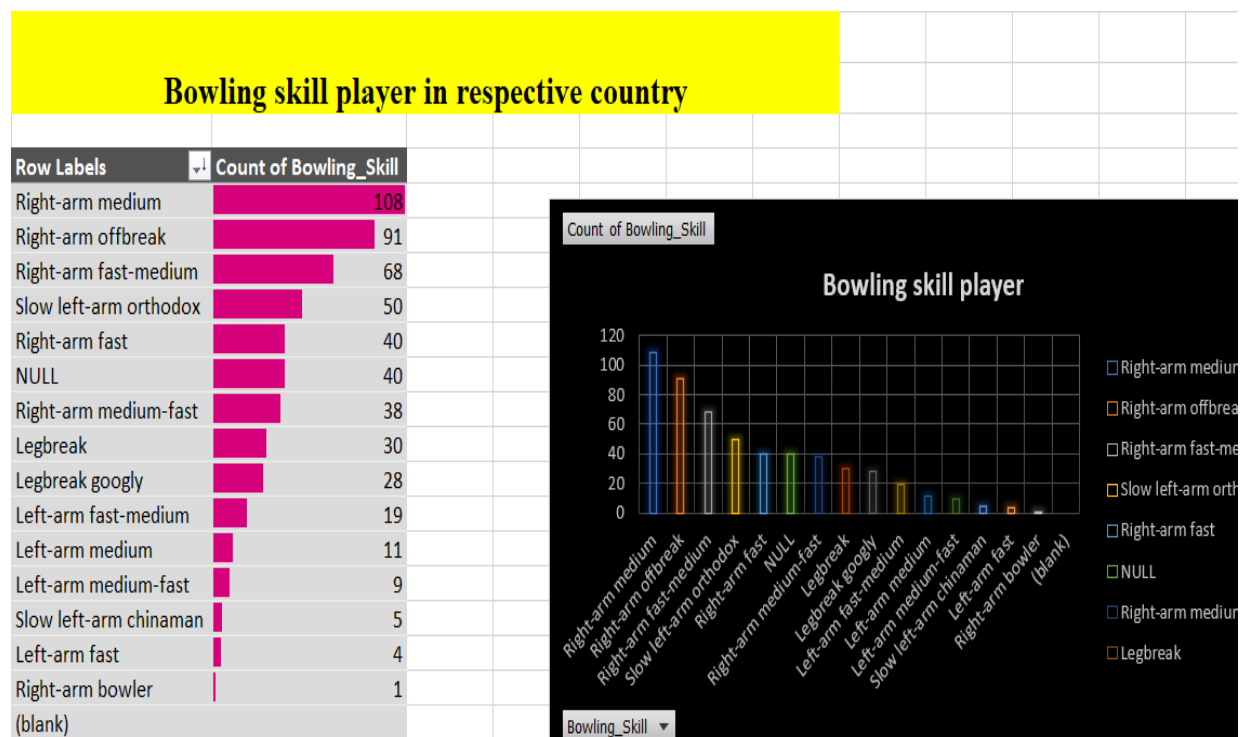
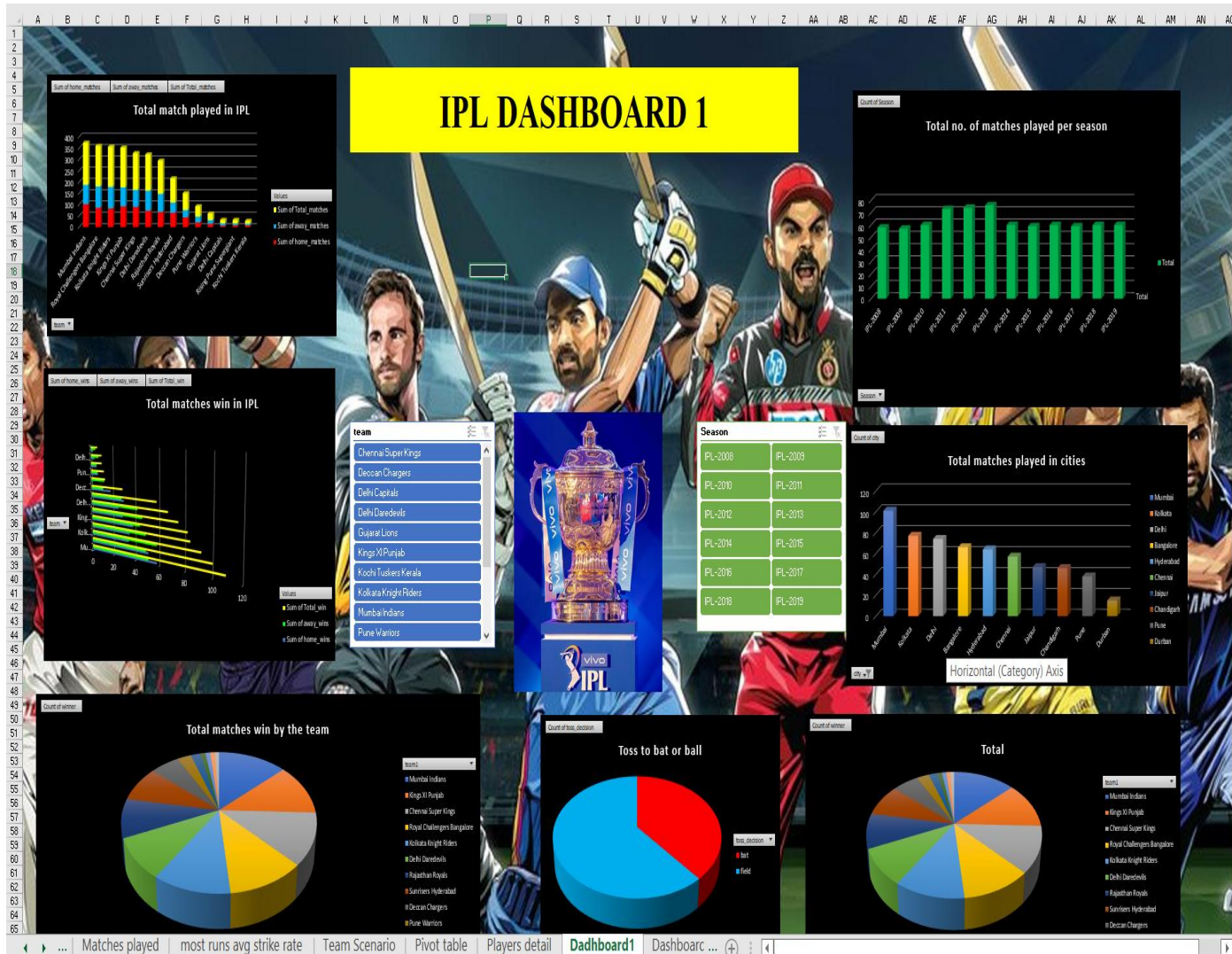
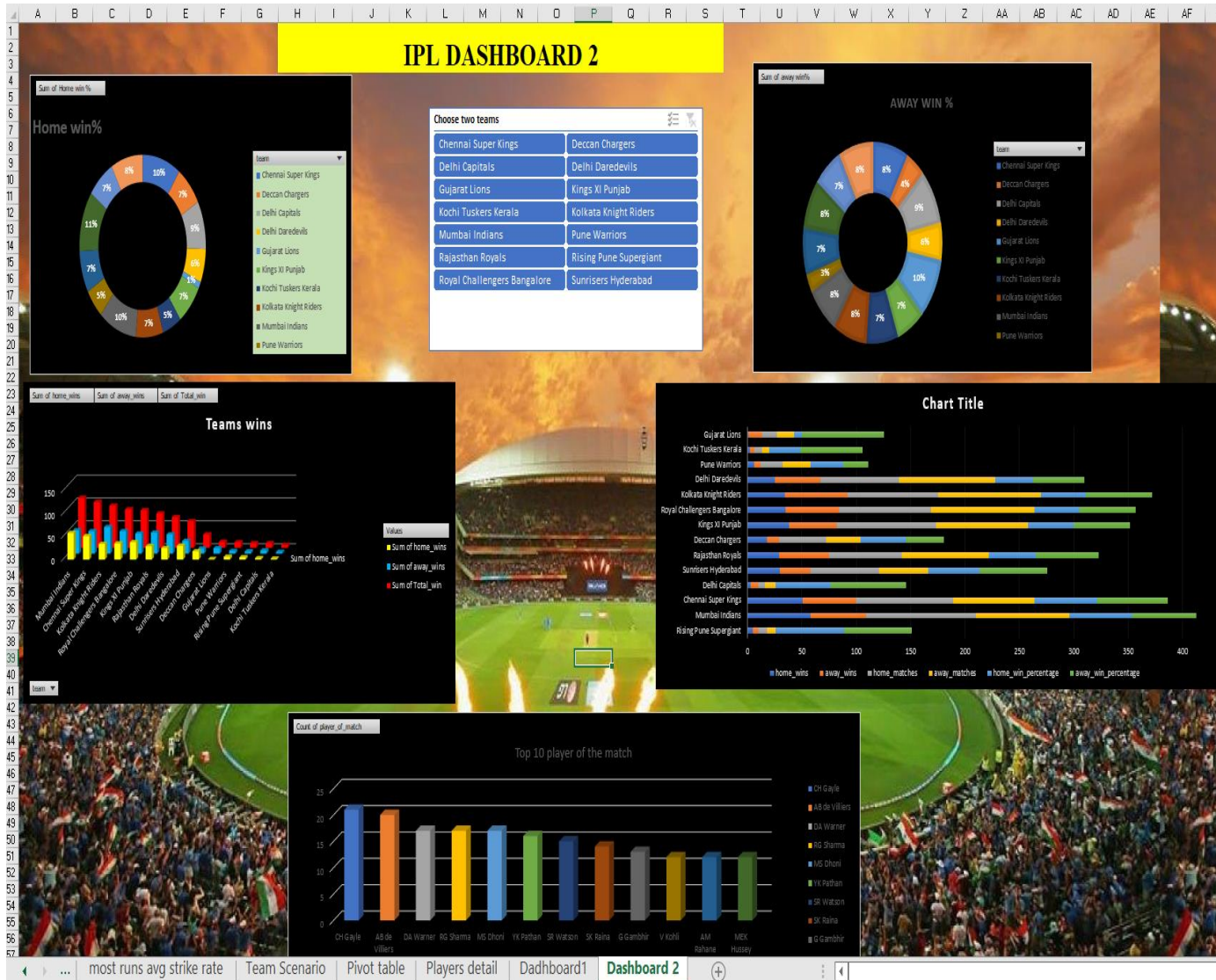


Fig 14

DASHBOARD 1



DASHBOARD 2



BIBLIOGRAPHY:

☐ **Dataset source:**

<https://www.kaggle.com/ramjidoolla/ipl-data-set>

☐ **Dashboard Background Image:**

<https://wallpapercave.com/ipl-t20-wallpapers>

<https://indianexpress.com/article/sports/ipl/ipl-2020-venue-uae-stadiums-match-list-6593316/>

☐ **Information about Data Management:**

<https://towardsdatascience.com/analysing-ipl-data-to-begin-data-analytics-with-python-5d2f610126a>