

**“STATISTICAL ANALYSIS  
ON INDIAN PREMIER LEAGUE (IPL)”**



# Contents

Chapter no.	Page no.
<b>Chapter1</b>	<b>Introduction.....5</b>
	1.1 Motivation
	1.2 Introduction
	1.3 IPL Teams
	1.4 History
	1.5 Objective
	1.6 Scope of Project
<b>Chapter2</b>	<b>Data Description ..... 12</b>
	2.1 Parameter
	2.2 Rules
<b>Chapter3</b>	<b>Graphical Analysis.....14</b>
<b>Chapter4</b>	<b>Statistical Technique ..... 22</b>
	4.1 Chi-Square test Analysis
	<b>Conclusion ..... 26</b>
	<b>References.....26</b>

# Chapter1. INTRODUCTION

## 1.1 Motivation

Being students of M.Sc. (Statistics) with specialization in Industrial Statistics, we were interested in knowing sports Statistics. As IPL is quite popular in India and lot of data is available. Also, we got opportunity to analyse cricket data and how to apply statistical tool on the cricket data.

## 1.2 Introduction to India Premier League(IPL)

The Indian Premier League (IPL) is a professional league for Twenty 20 cricket championship in India. It was initiated by the Board of Control for Cricket in India (BCCI), head quartered in Mumbai and is supervised by BCCI Vice President Rajeev Shukla, who serves as the league's Chairman and Commissioner. It is currently contested by nine teams, consisting of players from around the world. It was started after an altercation between the BCCI and the Indian Cricket League. In 2010, IPL became the first sporting event ever to be broadcast live on YouTube in association with India times. Its brand value is estimated to be around US\$2.99 billion in fifth season. The IPL works on a franchise-system based on the American style of hiring players and transfers. These franchises were put for auction, where the highest bidder won the rights to own the team, representing each city. The auction for the same took place on January 24, 2008 and the total base price for the auction was \$400 million. The auction went on to fetch \$723.59 million. The Mumbai franchise owned by Mukesh Ambani's Reliance Industries Limited (RIL) was the most expensive franchise - fetching \$111.9 million closely followed by Vijay Mallya's United Breweries which paid \$111.6 million for the Bangalore franchise. Media house Deccan Chronicle won the Hyderabad chapter of the IPL for \$107 million, while India Cements Chennai franchise cost \$91 million. Bollywood also made its presence felt with two of its leading stars bagging the ownership of their respective teams - ShahRukh Khan and Juhi Chawla's Red Chillies Entertainment buying out Kolkata for \$75.09, while Preity Zinta and her beau Ness Wadia bought the Mohali team for \$76 million.

GMR, the infrastructure development group which who are involved in a project for revamping the Delhi airport, bagged the ownership of the Delhi team for \$84 million and the Emerging Media, consisting of its CEO Fraser Castellino, Manoj Badale and Lachlan Murdoch and other investors won the rights for the Jaipur franchise for \$67 million. On 21 March 2010, Pune and Kochi were unveiled as the two new franchises for the fourth edition of the Indian Premier League. The base price was \$225 million. While Pune was bought by Sahara Adventure Sports Group for \$370 million, the Kochi franchise was bought by Rendezvous Sports World Limited for \$333.3 million. The process was to have been completed on 7 March but was postponed by two weeks after many bidders and the BCCI objected to stiff financial clauses. The second franchise auction fetched total \$703 million. Over 200 million Indian viewers, 10 million international viewers, 4 million live spectators

:the Indian Premier League (IPL) is a sports and entertainment revolution in the making, surpassing all records of viewership on ground and on media. Advertising revenue and ticket sales have exceeded all expectations, making IPL highly profitable for the organizers, broadcasters and successful team owners. Zealous fan following—even hostility for visiting teams—shows local loyalties are building up faster than any one expected.

### 1.3 IPL Teams

#### CHENNAI SUPER KINGS:

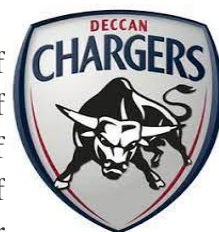
Chennai Super Kings (often abbreviated as CSK) is a franchise cricket team based in Chennai, Tamil Nadu that plays in the Indian Premier League. Founded in 2008, the team is currently captained by **Mahendra Singh Dhoni** and coached by Stephen Fleming, a former New Zealand cricketer. The team's home ground is the M.A.



Chidambaram Stadium (often referred to as Chepauk) located in Chennai. Chennai Super Kings is the most successful team in the Indian Premier League team so far. The team has won the IPL title twice in succession (2010 and 2011) and reached the play-offs every season. They were also the first Indian team to have won the Champions League Twenty20.

#### DECCAN CHARGERS:

Deccan Chargers known in short as DC is a cricket franchise based in the city of Hyderabad in the Indian Premier League. They, after finishing last in the first season of the IPL, won the second season held in South Africa in 2009, under the captaincy of former Australian wicketkeeper batsman Adam Gilchrist. Gilchrist was the captain of the team for the first three seasons of the IPL. From the fourth season, Kumar Sangakkara has been leading the team and Cameron White has been playing as his deputy. The team is coached by Darren Lehmann, former Australian cricketer. The Deccan Chargers franchise is owned by the Deccan Chronicle Holdings Limited.



#### DELHI DAREDEVILS:

Delhi Daredevils is the Delhi franchise of the Indian Premier League in cricket. The franchise is owned by the GMR Group. Founded in 2008, the team captained by Virender Sehwag and coached by former South African cricketer Eric Simons. They play all their home matches at the historic Feroz Shah Kotla Ground. Virender Sehwag has been accorded the icon player status in the Delhi Daredevils team and was also the captain of the side during the first two seasons. However, he resigned and passed on the leadership to his opening partner Gautam Gambhir for the 2010 season. But after Gambhir left the team for Kolkata Knight Riders in the fourth edition, Sehwag was once again given the duty to captain the team. Delhi Daredevils renamed as Delhi Capitals, Shreyas Iyer was captain in IPL 2019.



### **KINGS XI PUNJAB:**

Kings XI Punjab is a cricket franchise based in Mohali, Punjab in the Indian Premier League. It is Captained by Ravichandran Ashwin. KXIP finished 6th in the IPL and did not qualify for the playoffs. The team plays its home matches at the PCA Stadium, Mohali. Since 2010 IPL, they have been playing some of their home games at Dharamsala. The first owners of the franchise included Preity Zinta, Karan Paul (Apeejay Surendera Group) and Mohit Burman (Dabur).



### **KOLKATA KNIGHT RIDERS:**

Kolkata Knight Riders (often abbreviated as KKR) is a cricket franchise representing Kolkata in the Indian Premier League and owned by Bollywood actor Shah Rukh Khan's company Red Chillies Entertainment in partnership with actress Juhi Chawla and her husband Jay Mehta from the Mehta Group. The team was captained by Dinesh Karthik and coached by Jacques Kallis. The bowling legend, Wasim Akram is the team's bowling consultant and mentor. Sourav Ganguly, who was the team's Icon Player for the first three seasons, captained the side in the first and third seasons of the tournament while Brendon McCullum led the team in the intervening period. The team won its first title in 2012 by defeating defending champions Chennai Super Kings by 5 wickets at their home ground in Chennai. The official theme of the team is Korbo, Lorbo, Jeetbo Re (We will do it, Fight for it, win it) and the official colours are purple and gold.



### **MUMBAI INDIANS:**

Mumbai Indians is a franchise cricket team representing the city of Mumbai in the Indian Premier League. The team is one of the eight founding members of the IPL in 2008. Mumbai Indians is currently led by Rohit Sharma and coached by Mahela Jayawardene. They are one of the main teams in the city, together with clubs like the Mumbai cricket team and football club Mumbai FC. The team is owned by India's biggest conglomerate, Reliance Industries, through its 100% subsidiary IndiaWin Sports.



### **PUNE WARRIORS INDIA:**

Pune Warriors India is a franchise cricket team that plays in the Indian Premier League (IPL) representing the city of Pune, Maharashtra. The team is one of two new franchisees added to the Indian Premier League (IPL) for the 2011 season, along side the Kochi Tuskers Kerala. On 21 March 2010, Sahara Adventure Sports Limited made the winning bid of \$370 million (approximately Rs 1900 crore) for the Pune franchise. The bid is the highest bid by any company in the short history of IPL. The Videocon Group lost the bid for the Pune IPL Team. Sourav Ganguly, the former captain of Pune Warriors India, serves as the mentor of the team. Former South African pacer Allan Donald is the bowling coach of the team. The team's home ground is the newly-constructed Subrata Roy Sahara Stadium in Pune.



### **RAJASTHAN ROYALS:**

Rajasthan Royals is an Indian Premier League franchise based in the city of the Jaipur. It was captained by former captain Steve Smith. They won the inaugural edition of the Indian Premier League under the captaincy of former Australian spin-legend Shane Warne. In 2010, they were terminated by the BCCI for a brief period for violations interms of agreement but were allowed to play in the league soon after the Board lost the case in the Supreme Court. Rajasthan Royals were the least active franchise during the players' auction. The team didn't have a designated icon player, which saved a lot of cap space for the franchise.



### **ROYAL CHALLENGERS BANGALORE:**

Royal Challengers Bangalore (often abbreviated as RCB) is a cricket team based in Bangalore that plays in the Indian Premier League. The team is lead by Virat Kohli and coached by Gary Kristen. The team is owned by Vijay Mallya, through his flagship firm UB Group. The director of the team is Siddharth Mallya. The team won only 4 matches in the inaugural season, losing 10 matches and finishing second from the bottom in the table. RCB plays all its home matches at the M. Chinnaswamy Stadium.



### **KOCHI TUSKERS KERALA (NOWDEFUNCT):**

Kochi Tuskers Kerala was a franchise cricket team that played in the Indian Premier League (IPL) representing the city of Kochi, Kerala. The team was one of two new franchisees added to the Indian Premier League (IPL) for the 2011 season, along side the Pune Warriors India. The team franchise was owned by Kochi Cricket Pvt Ltd., which was a consortium of multiple companies. On September 19, 2011, the BCCI announced that the Kochi Tuskers Kerala IPL franchise was terminated for braching its terms of agreement.



### **Rising Pune Supergiant:**

Rising Pune Supergiant was a franchise cricket team based in Pune, Maharashtra, that played in the IPL in 2016 and 2017. It was the second team representing Pune after Pune Warriors. The team along with Gujrat Lions featured as two-season replacements for Chennai Super kings and Rajasthan Royals, who were suspended due to the irinvolvement in illegall betting by their respective owners. The Rising Pune Supergiant franchise is owned by RP-Sanjiv Goenka Kapoor controlled by Sanjiv Goenka The team name was announced (as *Rising Pune Supergiant*) on 18 January 2016 by Goenkain Kolkata and Raghu Iyer was appointed CEO. The owners changed the team name to *Rising Pune Supergiant* on 26 March 2017. The team lost the 2017 IPL final to Mumbai Indians by 1 run , which was the team's last game in the IPL





### **SunrisersHyderabad:**

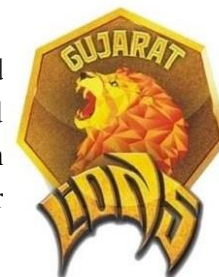


Sunrisers Hyderabad are a franchise cricket team based in Hyderabad, Telangana, India, that plays in the IPL. The franchise is owned by Kalanithi Maran of the SUN Group and was founded in 2012 after the Hyderabad-based Deccan Chargers were terminated by the IPL. The

Team is currently captained by Kane Williamson and coached by Trevor Bayliss. Their primary home ground is the Rajiv Gandhi International Stadium, Hyderabad which has capacity of

55,000. The team made their first IPL appearance in 2013, where they reached the playoffs, eventually finishing in fourth place. The Sunrisers won their maiden IPL title in the 2016 season, defeating the Royal Challengers Bangalore by 8 runs in the final. The team has qualified for the play-off stage of the tournament in every season since 2016. In 2018, the team reached the finals of the Indian Premier League, but lost to Chennai Super Kings.

### **Gujarat Lions:**



Gujarat Lions was a franchise cricket team based in the city of Rajkot, that represented Indian state Gujarat in the IPL. The team played in the IPL for 2 years (the 2016 and 2017 seasons) as one of the replacements for Chennai Super Kings and Rajasthan Royals, who were both suspended for two seasons due to illegal betting by their respective owners. The franchise was owned by Intex Technologies.

## **1.4 History and Background of Formation of the IPL**

Kerry Francis Bullmore Packer, AC (17 December 1937 – 26 December 2005), was an Australian media tycoon whose family company owned controlling interests in both the Nine television network and leading Australian publishing company Australian Consolidated Press. Packer was best known for founding World Series Cricket. In 1977 the Nine Network cricket rights deal led to a confrontation with the cricket authorities, as top players from several countries rushed to join him at the expense of their international sides. Packer's aim was to secure broadcasting rights for Australian cricket, and he was largely successful. Many of the well-known cricketers of that period left their national team to play in Kerry Packer's World Series cricket. Some of our legendary cricketers also contacted to play in that series. But due to some controversies, mainly with Australian board due to television rights, this league could not be successful. On the background of this idea, Zee Entertainment Enterprises organized a league called Indian cricket league. The Indian Cricket League (ICL) was a private cricket league funded by Zee Entertainment Enterprises that operated between 2007 and 2009 in India. In its two seasons included tournaments between four international teams (The World XI, India, Pakistan and Bangladesh) and nine domestic teams notionally located in major Indian cities as well as the champions Lahore Badshah's who were based in Lahore, Pakistan. The matches were played in the Twenty20 format. Zee T.V owner Subhash Chandra founded this league in the response of BCCI's some of working style. Many times, Subhash Chandra give bid for television rights but every time he was rejected even if his bid was the highest. This results in formation of Indian Cricket League (ICL). But due to some reasons this league could not be successful. Some of reasons are commercial factors, lacking of the support of the BCCI and ICC.

The ICL was set up with a billion-dollar Indian Rupee corpus, and was to initially comprise six teams playing Twenty20 cricket, with plans to expand to sixteen teams within three years and to eventually

move to 50-over matches. These plans, if they had been realised, would have made the ICL the richest professional league in India. On 24 July 2007, some famous international names were announced to have signed to play in the ICL, including highest innings record-holder Brian Lara. The BCCI refused to recognise the ICL as a cricket league, and criticised Kiran More and Kapil Dev for joining the ICL. On August 21, 2007 Kapil Dev was sacked from his NCA post. The International Cricket Council gave a statement through its chief executive, Malcolm Speed, that the ICC would not recognize the ICL unless the BCCI chooses to recognise it. Faced with the threat of young players joining the ICL, the BCCI jacked up prize money for winners, runners-up and losing semi-finalists across all tournaments. An average domestic cricketer can hope to make around Rs 35,000 per match day from the season of 2007-08: more than double the Rs 16,000 they got in 2005-06. The BCCI has also planned to do away with honorary selectors, who will be paid professionals from September 2008 onwards. Then BCCI started its own international Twenty20 league. The official league, which was launched in April 2008, was called the Indian Premier League Twenty20.

## 1.5 Objectives

**The project study tends to follow and achieve following objective :**

- i. To give ranking to all batsman and bowlers of IPL.
- ii. Season wise comparison of IPL seasons.
- iii. Impact of toss decision / Home ground on match result.

## 1.6 Scope of project

The project was conceived with the objective of understanding to check the performance of players and teams and to choose player according to ranking based on previous performance with the help of this information teams choose any one player in next season.



## Chapter 2. Data Description

### 2.1 Introduction

We have collected the information with the secondary information source information of the IPL ([www.kaggle.com/datasets](http://www.kaggle.com/datasets)) and from [www.iplt20.com](http://www.iplt20.com). We have tried best level for doing various analyses for making the project effective with available data. Extensive research was carried out for successful completion of this project. Secondary data were collected from various sources. The secondary data in this study is collected from various websites. Analysis and others parts are done on the basis of these secondary data and knowledge collected from the Internet. In this project, past 12-year data (2008-2019) of IPL. The data containing the players information such as match details, team details, record details etc.

### 2.2 Parameters

#### For Batsman

1. **Runs:** The total number of scored by a player in the IPL 2008 to 2019 seasons. Higher values indicate stronger performance

2. **Batting Average:** The total number of runs a batsman has scored divided by the total number of times he has been called out in the IPL 2008 to 2019 seasons. Higher values indicate stronger performance. However, for a batsman with several “not out” cases, this number over rates the batsman, which is weakness in this measure, and this is why it should not be used as the only variable for batting performance analysis.

3. **Batting Strike rate (SR=(Runs/Balls) \*100):** The batting strike rate is defined as the number of runs scored per 100 balls faced by a batsman in the IPL 2008 to 2019 seasons. Again, higher values indicate stronger performance. An aggressive batting style is always helpful in shorter versions of limited –overs cricket matches like Twenty 20. However, a high strike rate accompanying a low batting average is not desirable.

4. **Economy rate (Econ=Runs/(overs bowled)):** The average number of runs conceded per over. Lower values are preferred since this is the run – rate against a specific bowler for a batting team. Therefore, the bowler’s aim is to keep this measure as small as possible.

5. **Wickets:** The number of wickets taken by bowler. There are ten possible wickets for an inning and there should be at least five bowlers, each of whom can bowl a maximum of four overs. A bowler’s goal is to take the maximum number of wickets from the overs that he bowls, so taking a large number of wickets from batsmen is one performance measure for bowlers. However, like the total number of runs statistic for a batsman, the number of wickets taken is not sufficient to measure the quality of bowler. The goal of bowler is to get the maximum number of wickets by using a minimum number of balls while simultaneously conceding a minimum number of runs.

**Some of the team composition rules are**

- Minimum squad strength of 16 players plus one physio and a coach.
- No more than 11 foreign players in the squad and maximum 4 foreign players should be in the playing 11.
- Minimum of 14 Indian players must be included in each squad.
- A minimum of 6 players from the BCCI under-22 pool in each squad.
- At least seven bowlers will be there for completing the 20 overs of match.
- Of the seven bowlers in the game, six bowl 3 overs and seventh one will bowl the remaining 2 overs.
- For a match to be declared complete, full 40 overs should have been bowled and batted on.
- The batting team can call all of their extra players to bat, till the end of 20<sup>th</sup> over.

However, only the first 11 players have the right to bowl and field.

The batting team can continue to bat till end of the over, even if it has reached the Winning target.

**IPL games utilize television timeouts, hence there is no time limit for teams to complete their innings. However, there may be a penalty if the umpires find teams misusing this privilege at their own choice. Additionally, each team is awarded two strategic timeouts of 2:30 each per innings—the teams can take the timeout when instructed, but is necessary to take it from 6th to 9th and 13th to 16th over.**

## Chapter3.GraphicalAnalysis

### 3.1 Introduction

Exploratory analysis and others parts are done on the basis of theses secondary data.

- **Import Libraries**

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

- **Load the data**

```
match=pd.read_csv('/content/match.csv') match
delivery=pd.read_csv( /content/delivery.csv ) delivery
```

	id	city	date	player_of_match	venue	neutral_venue	team1	team2	toss_winner	toss_decision	winner	result	result_margin	eliminator	meth
0	335982	Bangalore	2008-04-18	BB McCullum	Chinnaswamy Stadium	0	Royal Challengers Bangalore	Kolkata Knight Riders	Royal Challengers Bangalore	field	Kolkata Knight Riders	runs	140.0	N	Na
1	335983	Chandigarh	2008-04-19	MEK Hussey	Punjab Cricket Association Stadium, Mohali	0	Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	bat	Chennai Super Kings	runs	33.0	N	Na
2	335984	Delhi	2008-04-19	MF Maharoof	Feroz Shah Kotla	0	Delhi Daredevils	Rajasthan Royals	Rajasthan Royals	bat	Delhi Daredevils	wickets	9.0	N	Na
3	335985	Mumbai	2008-04-20	MV Boucher	Wankhede Stadium	0	Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	Royal Challengers Bangalore	wickets	5.0	N	Na
4	335986	Kolkata	2008-04-20	DJ Hussey	Eden Gardens	0	Kolkata Knight Riders	Deccan Chargers	Deccan Chargers	bat	Kolkata Knight Riders	wickets	5.0	N	Na



- **List of the Participating Teams**

```
all_teams = match['team1'].tolist() + match['team2'].tolist()
all_teams = list(set(all_teams))
all_teams
```

### OUTPUT

['Kochi Tuskers Kerala',

'Kings XI Punjab',  
 'Rising Pune Supergiant',  
 'Chennai Super Kings',  
 'Rising Pune Supergiants',  
 'Delhi Capitals',  
 'Rajasthan Royals',  
 'Mumbai Indians',  
 'Gujarat Lions',  
 'Pune Warriors',  
 'Sunrisers Hyderabad',  
 'Royal Challengers Bangalore',  
 'Delhi Daredevils',  
 'Deccan Chargers',  
 'Kolkata Knight Riders']

- **Number of Matches per Venue**  
`sns.countplot('venue', data=match)`  
`plt.xticks(rotation='vertical')`

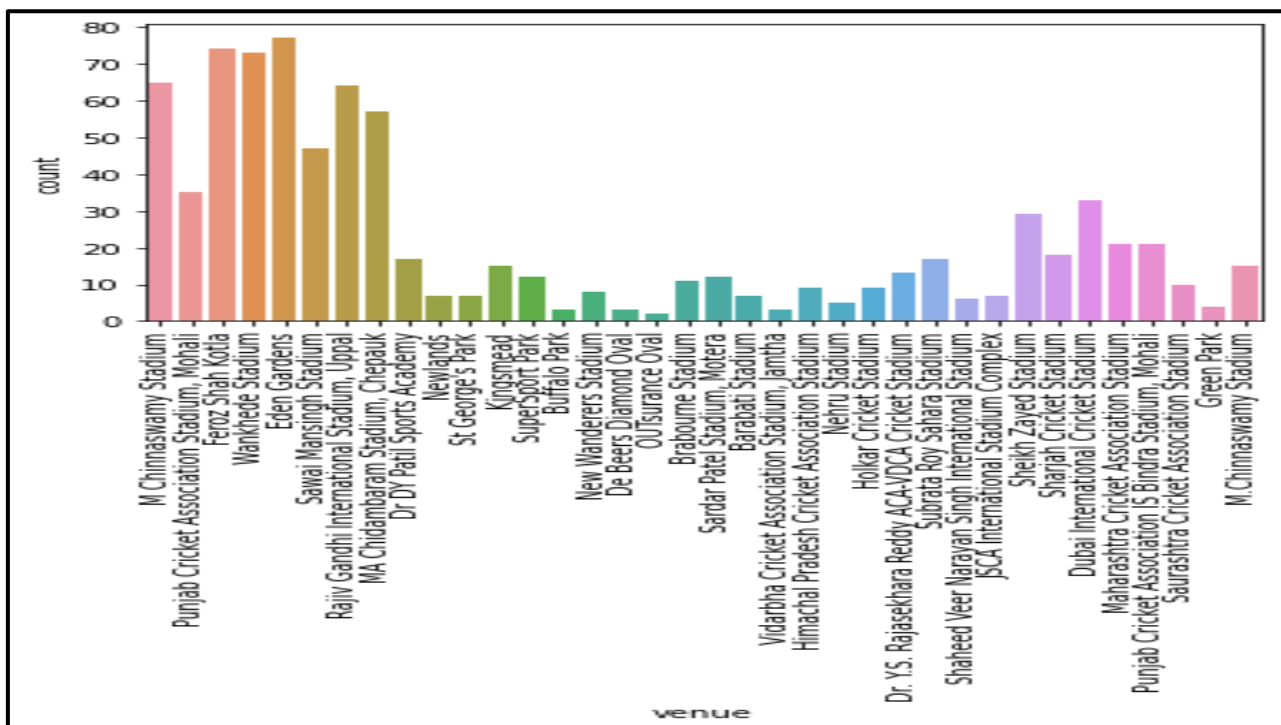
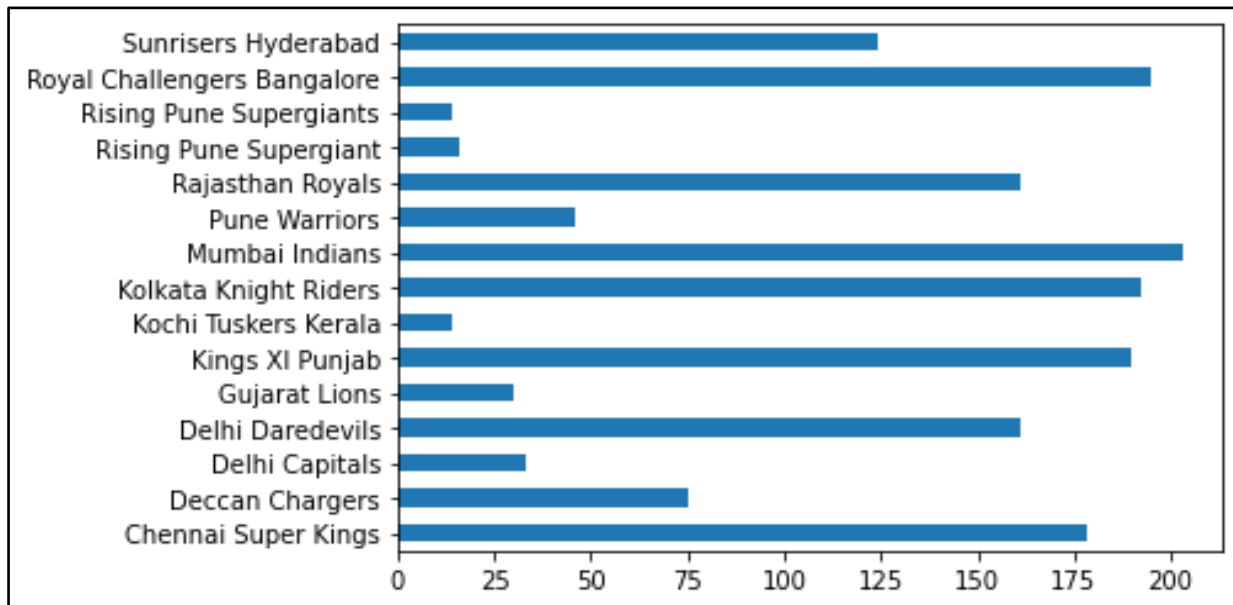


Fig.3.1

**Conclusion:** From Fig 3.1 as we see that Eden Gardens is the fan-favourite ground of IPL, nearly 80 matches have been hosted there.

- **Matches Played by Each Team**

```
x = match['team1'].value_counts()
y = match['team2'].value_counts()
(x+y).plot(kind='barh')
```



**Fig.3.2**

**Conclusion:** From Fig.3.2 We count the value of each team playing in column one and add to the count of each team from team two to get the desired output. For example, if CSK played 90 times from team one and 85 times from team 2 then the total of 175 matches are shown in the graph. You can see that Mumbai Indians played the highest number of matches in the IPL.

- **Matches Won by Each Team**

```
x=pd.DataFrame({"Winner":match['winner']}).value_counts()
print(x)
```

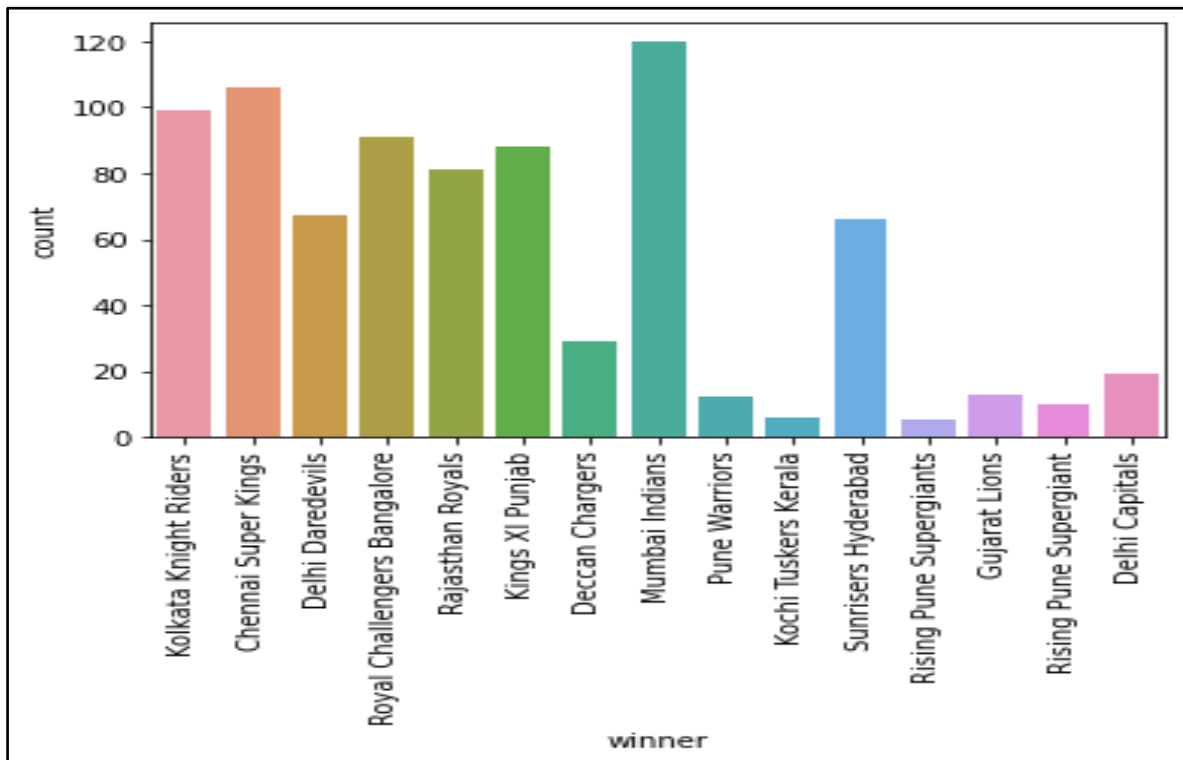
**Table3.3 teams and their total wins**

Team	Wins
MI	109
CSK	98
KKR	92
RCB	83
KXIP	81
RR	75
DD	66

SH	56
DC	29
RPS	15
GL	13
PW	12
DC	10
KTK	6
<b>TotalMatches</b>	<b>745</b>

- **Number of Matches per Venue**

```
sns.countplot('winner', data=match)
plt.xticks(rotation='vertical')
```



**Fig.3.3**

**Conclusion:** From Fig 3.3 Mumbai Indians win the highest number of matches followed by CSK and other teams.

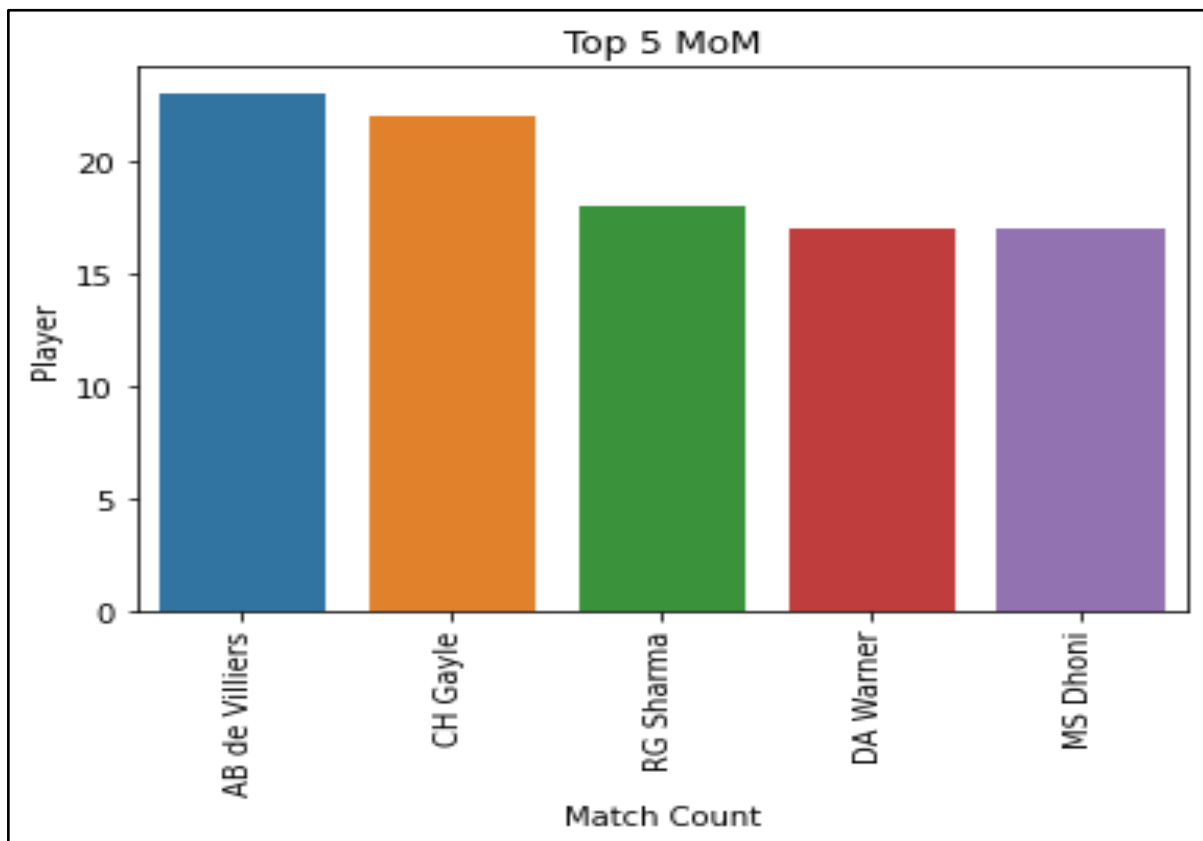


- **Top 5 Players with the Highest Number of Man of the Match Awards**

```
temp_data=match['player_of_match'].value_counts().head()
print(temp_data)
sns.barplot(x=temp_data.index,y=temp_data.values,data=match)plt.title("Top 5 MoM")
plt.xticks(rotation=90)
plt.xlabel("Match Count")
plt.ylabel("Player")
plt.show()
```

AB de Villiers	23
CH Gayle	22
RG Sharma	18
DA Warner	17
MS Dhoni	17

Name: player\_of\_match, dtype: int64

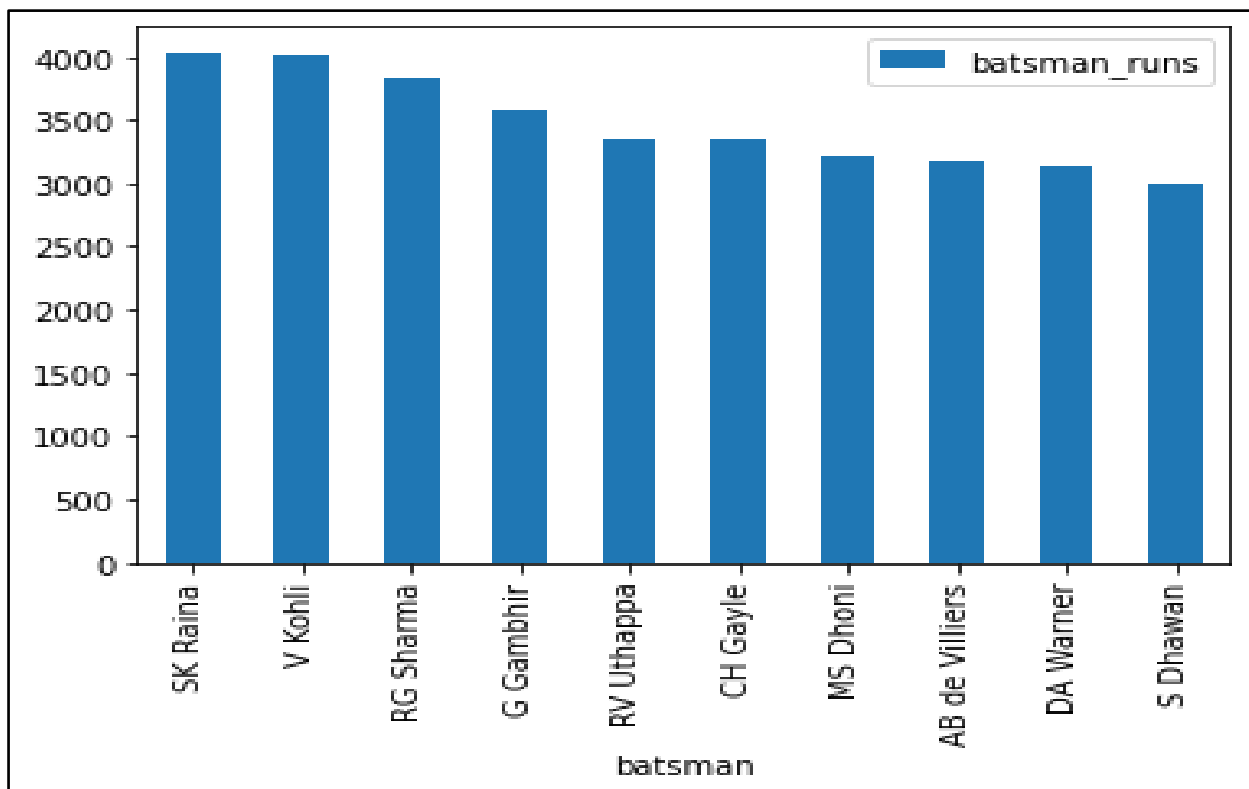


**Fig.3.4**

- **The Top Batsman in the IPL**

```
top_batsman=delivery.groupby('batsman')  
['batsman_runs'].agg('sum').reset_index().sort_values('batsman_runs',  
ascending=False).head(10)
```

```
top_batsman.set_index('batsman', inplace=True)  
top_batsman.plot(kind='bar')
```



**Fig.3.5**

**Conclusion:** From Fig 3.5 We grouped the top 10 batsmen from the delivery dataset and summed up their runs. After this, King Kohli is at the top followed by Suresh Raina and other batsmen.

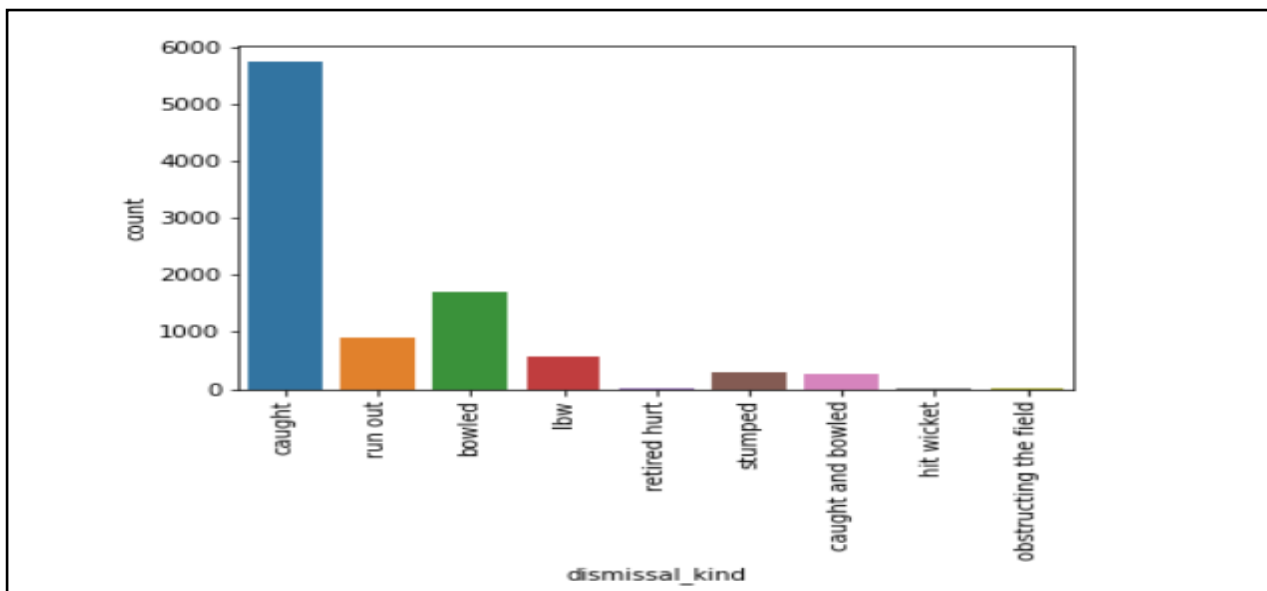
- **The Bowler Who Has Given the Highest Number of Runs**

```
delivery.groupby('bowler')['total_runs'].agg('sum').reset_index().sort_values('total_runs', ascending=False).head(10)
```

Index	Bowler	Total_runs
108	Harbhajan Singh	3148
213	PP Chawla	3136
202	P Kumar	3097
4	A Mishra	2914
228	R Vinay Kumar	2907
79	DJ Bravo	2728
114	IK Pathan	2648
275	SL Malinga	2646
218	R Ashwin	2518
330	Z Khan	2507

- **Dismissal Kind**

```
sns.countplot('dismissal_kind', data=delivery)
plt.xticks(rotation='vertical')
```

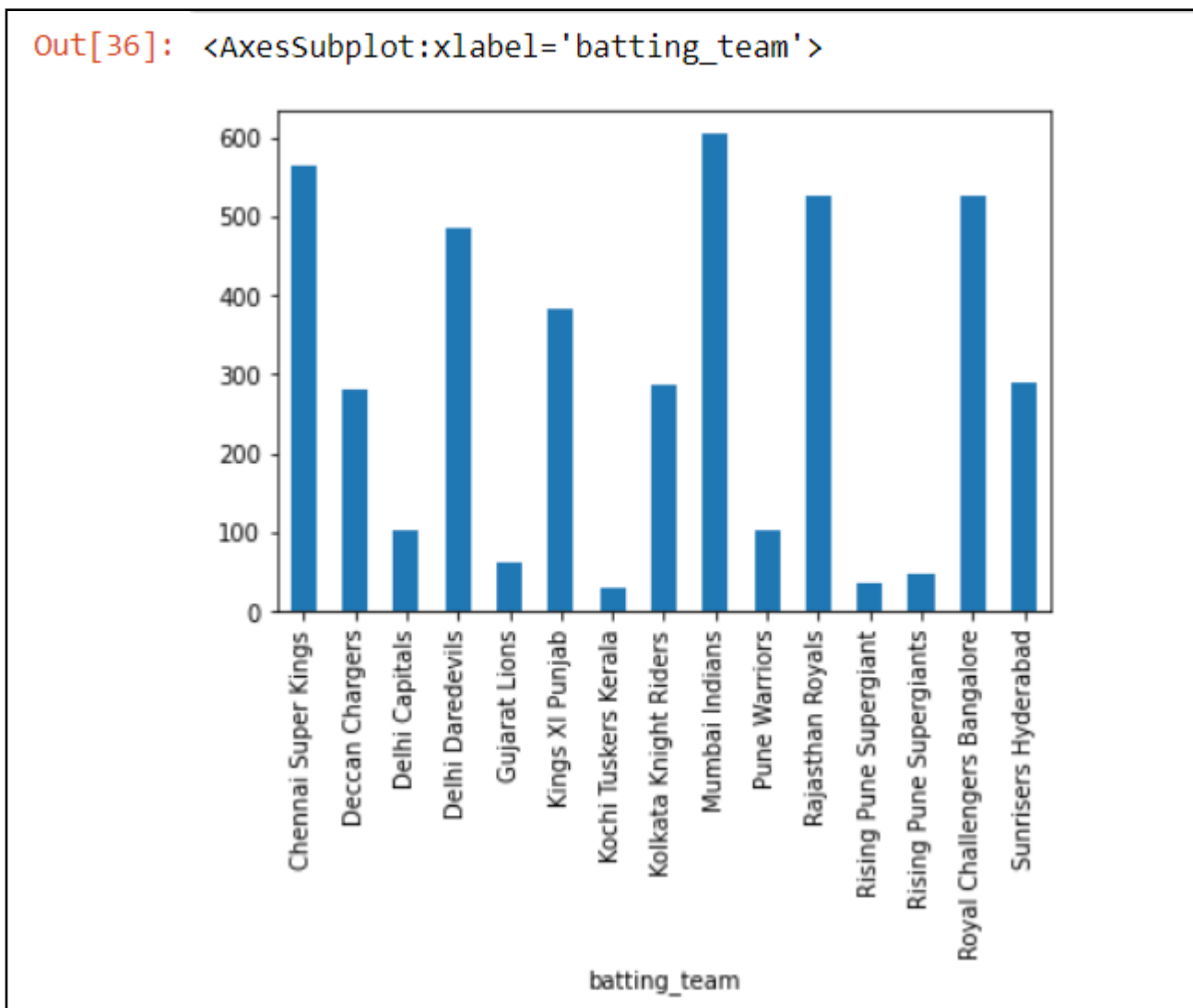


**Fig.3.6**

**Conclusion:** From Fig 3.6 we conclude that most of the player out by caught .

- **The Bowler with Team-wise Performance**

```
mask=delivery['bowler']=='PP Chawla'  
delivery[mask].groupby('batting_team')['total_runs'].agg('sum').plot(kind='bar')
```



**Fig.3.7**

We are taking the example of PP Chawla. This bowler has given the highest number of runs in the IPL history till 2020. We summed up the total runs given by PP Chawla to the opponent team.

**Conclusion:** From Fig 3.7 it's clear that if you have PP Chawla in your team then don't let him play against MI, CSK, RCB, RR, and DC.

## Chapter4 . Statistical Technique

### Chi-SquareTest

#### Introduction

The Chi-square test for independence is non-parametric test also called Pearson's Chi-square test or chi-square test for association is used to discover if there is a relationship between two categorical variables.

#### Assumption

- The two variables should be measured at an ordinal and nominal level.
- The two variables should consist of two or more categorical independent group.

#### Objectives

To check the association between two variables

- To check the association between Match winning and Toss decision.
- To check the association between Match winning and Home Ground

#### Data Description

**Below Table 4.43** is about toss win by corresponding team and taking decision about Bat or Field and match result. And other data is about home ground of corresponding teams and their Match results. The below data is for two years IPL 2018,201

$$Toss\ Decision = \begin{cases} 1, & \text{if Team choose Batting} \\ 0, & \text{if Team choose Field} \end{cases}$$

$$Match\ Result = \begin{cases} 1, & \text{if Team choose Batting} \\ 0, & \text{if Team choose Field} \end{cases}$$

$$Home\ Ground = \begin{cases} 1, & \text{if match is played at home ground of corresponding team} \\ 0, & \text{if match is not played at home ground of corresponding team} \end{cases}$$

Match	Toss	Toss	Match	Home	Match	Toss	Toss	Match	Home
ID	Toss	Decision	Result	Ground	ID	Toss	Decision	Result	Ground
7894	Field	0	1	0	7915	Field	0	0	0
7895	Field	0	1	1	7916	Field	0	0	0
7896	Field	0	1	1	7917	Field	0	1	0
7897	Field	0	1	1	7918	Field	0	0	1
7898	Field	0	1	1	7919	Field	0	0	1
7899	Field	0	0	1	7920	Field	0	1	0
7900	Field	0	1	1	7921	bat	1	1	0
7901	Field	0	1	1	7922	Field	0	1	0
7902	Field	0	1	0	7923	Field	0	0	1
7903	Field	0	1	0	7924	Field	0	0	1
7904	Field	0	0	0	7925	Field	0	0	1
7905	Field	0	0	1	7926	Field	0	1	1
7906	Field	0	0	1	7927	Field	0	1	0
7907	Field	0	0	1	7928	Field	0	1	1
7908	bat	0	1	0	7929	bat	1	0	1
7909	Field	1	1	1	7930	Field	0	0	1
7910	Field	0	0	1	7931	Field	0	1	1
7911	Field	0	1	0	7932	Field	0	0	1
7912	Field	0	1	1	7933	bat	1	1	1
7913	Field	0	0	0	7934	Field	0	0	0
7914	bat	1	0	1	7935	bat	1	0	0

Match	Toss	Toss	Match	Home	Match	Toss	Toss	Match	Home
ID	Toss	Decision	Result	Ground	ID	Toss	Decision	Result	Ground
7936	bat	1	0	1	11313	Field	0	1	1
7937	Field	0	0	0	11314	Field	0	1	1
7938	Field	0	1	0	11315	Field	0	1	1
7939	Field	0	1	1	11316	Field	0	1	0
7940	Field	0	1	0	11317	Field	0	1	0
7941	Field	0	1	0	11318	Field	0	1	0
7942	Field	0	1	1	11319	Field	0	1	0
7943	Field	0	0	1	11320	Field	0	1	0
7944	Field	0	0	1	11321	Field	0	0	0
7945	Field	0	0	1	11322	Field	0	1	1
7946	bat	1	1	1	11323	Field	0	0	1
7947	bat	1	0	0	11324	Bat	1	0	1
7948	bat	1	1	1	11325	Bat	1	1	0
7949	Field	0	1	1	11326	Field	0	0	0
7750	Field	0	1	0	11327	Field	0	1	1
7951	Field	0	0	1	11328	Field	0	1	1
7952	Field	0	0	0	11329	Field	0	1	1
7953	Field	0	1	0	11330	Field	0	0	1

11137	field	0	1	1	11331	field	0	1	0
11138	field	0	1	1	11332	field	0	1	1
11139	field	0	0	0	11333	field	0	0	1
11140	field	0	0	0	11334	field	0	1	0



11141	bat	1	0	0	11335	field	0	0	0
11142	field	0	0	1	11336	field	0	1	1
11143	field	0	0	0	11337	Bat	1	1	1
11144	bat	1	0	0	11338	field	0	0	1
11145	field	0	1	1	11339	field	0	0	1
11146	field	0	1	1	11341	field	0	0	1
11147	field	0	0	0	11342	Bat	1	1	1
11148	field	0	0	0	11343	field	0	1	0
11149	field	0	0	0	11344	Bat	1	0	1
11150	field	0	1	1	11345	field	0	1	1
11151	field	0	0	0	11346	field	0	1	1
11152	field	0	1	1	11347	field	0	1	1
11153	field	0	1	1	11412	Bat	1	0	0
11309	bat	1	1	1	11413	field	0	1	0
11310	field	0	0	0	11414	field	0	0	0
11311	field	0	1	1	11415	bat	1	1	0
11312	field	0	1	1					

## Chi-Square Test for Independence

**Chi-Square Test for Association :** Toss Decision, Match

Result Rows : Toss decision Columns : Match Result

$H_0$ : The impact of Toss Decision is independent on Match result

$H_1$ :The impact of Toss Decision is dependent on Match result

### Observed values

	Toss Decision	Match Result	Total
Field	42	57	99
Bat	10	10	20
Total	52	67	119

### Expected values

	Expected Values		Total
Field	43.26	55.74	99
Bat	8.739	11.261	20
Total	52	67	119

### Squared difference values

	Toss Decision	Match Result	Total
Field	0.037	0.029	0.066
Bat	0.182	0.141	0.323
Total	0.219	0.17	0.389

**Chi-Square= 0.389, DF= 1p-Value = 0.533**  
p-Value= 0.535 > 0.05

**Therefore, we fail to reject  $H_0$**

**Conclusion:** It is conclude that null hypothesis  $H_0$  is not rejected.  
Therefore, the impact of toss decision is independent on match result

**Chi-Square Test for Association : Home Condition, Match Result**

**Rows :** Home Condition **Columns :** Match Result

$H_0$  :The impact of home condition is independent of match result

$H_1$  : The impact of home condition is dependent of match result

**Observed values**

	<b>Toss Decision</b>	<b>Match Result</b>	<b>Total</b>
<b>Field</b>	19	30	49
<b>Bat</b>	33	37	70
<b>Total</b>	52	67	119

**Expected values**

	<b>Expected values</b>		<b>Total</b>
<b>Field</b>	21.41	27.59	49
<b>Bat</b>	30.59	39.41	70
<b>Field</b>	52	67	119

**Squared difference values**

	<b>Toss Decision</b>	<b>Match Result</b>	<b>Total</b>
<b>Field</b>	0.272	0.211	0.483
<b>Bat</b>	0.19	0.148	0.338
<b>Total</b>	0.462	0.359	0.821

**Chi-Square= 0.821, DF= 1, p-Value = 0.365**

**Here, p-Value =0.365 > 0.05**

**Therefore, we fail to reject  $H_0$ .**

**Conclusion:** Here  $H_0$  is not rejected so, the impact of home condition is independent on match result

## Conclusions

The approach as brought out analysis and visualization of various aspects of IPL Matches and gives useful results to the user.

1. It could be great help to team owners who purchase players for their teams in auction every year, on the basis of overall batting and bowling parameters for corresponding batsman and bowlers.
2. It also could be great help to coaches and captain to make right selection of players (i.e., batsman, bowlers etc) or playing 11 team on the basis of **performance score** of batsman or bowlers.
3. It also gives information about the helpful decision on various venues.

## References

1. [www.iplt20.com](http://www.iplt20.com)
2. [www.howstat.com](http://www.howstat.com), [www.espncriinfo.com](http://www.espncriinfo.com)
3. [www.kaggle.com](http://www.kaggle.com)
4. [www.wikipedia.org](http://www.wikipedia.org)

## Softwares

1. Python
2. MS-Excel
3. MS-Word

