

Data analytics IPL completed in Lab.

(Assignment 3)

Dhanraj Tungar TY-BCA (A)

2022018100094671

- IPL DATA ANALYSIS

```
import pandas as pd
```

```
from matplotlib import pyplot as plt
```

```
df = pd.read_csv('C:\\\\Users\\\\Sujal Amit Kadam\\\\Downloads\\\\matches.csv')
```

```
#displaying first 5 rows
```

```
df.head()
```

#displaying first 5 rows df.head()															
	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj Singh	Raj Int
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPD Smith	Ma A:
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	CA Lynn	S A:
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	GJ Maxwell	
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Jadhav	Chir

```
#no. of rows and columns  
df.shape
```

```
(756, 18)
```

```
#getting most man of the man awards
```

```
df['player_of_match'].value_counts()
```

```
#getting most man of the man awards  
df['player_of_match'].value_counts()
```

player_of_match	
CH Gayle	21
AB de Villiers	20
RG Sharma	17
MS Dhoni	17
DA Warner	17
	..
PD Collingwood	1
NV Ojha	1
AC Voges	1
J Theron	1
S Hetmyer	1

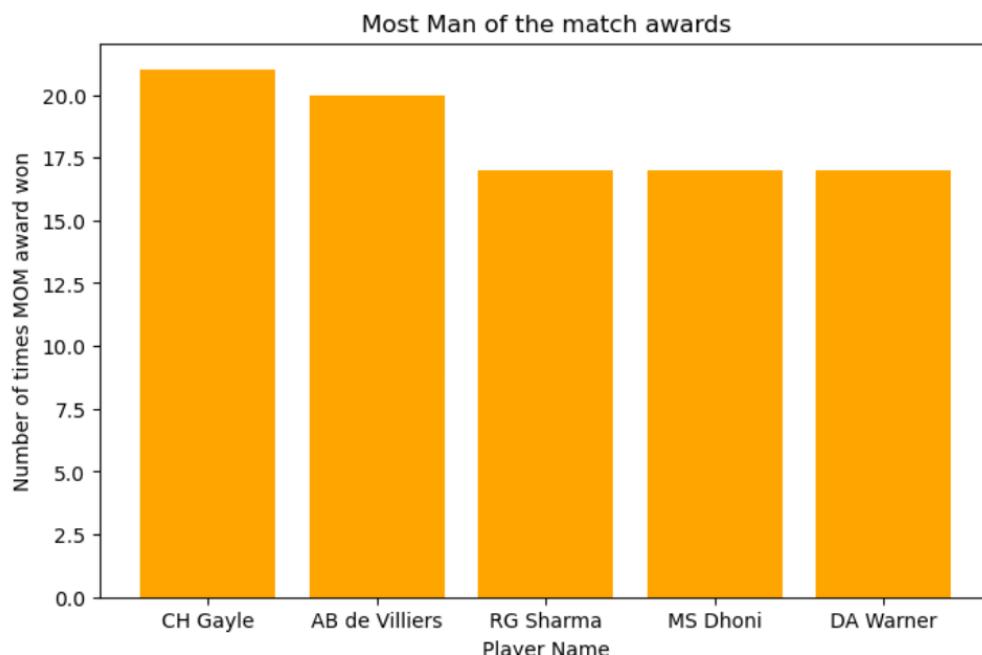
```
Name: count, Length: 226, dtype: int64
```

```
#top 10 players with man of the match awards  
df['player_of_match'].value_counts()[0:10]
```

```
player_of_match  
CH Gayle          21  
AB de Villiers    20  
RG Sharma         17  
MS Dhoni          17  
DA Warner         17  
YK Pathan         16  
SR Watson         15  
SK Raina          14  
G Gambhir          13  
MEK Hussey        12  
Name: count, dtype: int64
```

```
#plotting bar graph to see top 5 man of the match award winners
```

```
plt.figure(figsize=(8,5))  
  
plt.bar(list(df['player_of_match'].value_counts()[0:5].keys()),list(df['player_of_match'].value_  
counts()[0:5]),color='orange')  
  
plt.title('Most Man of the match awards')  
  
plt.xlabel('Player Name')  
  
plt.ylabel('Number of times MOM award won')  
  
plt.show()
```



Dhanraj Tungar

```
#frequency of match status
df['result'].value_counts()
```

```
result
normal      743
tie          9
no result    4
Name: count, dtype: int64
```

```
#checking top 5 toss winners
df['toss_winner'].value_counts().head(5)
```

```
toss_winner
Mumbai Indians        98
Kolkata Knight Riders  92
Chennai Super Kings   89
Royal Challengers Bangalore 81
Kings XI Punjab       81
Name: count, dtype: int64
```

#getting records where teams win while batting first

```
batting_first_df = df[df['win_by_runs'] != 0]
```

```
batting_first_df.head()
```

```
batting_first_df = df[df['win_by_runs'] != 0]
batting_first_df.head()
```

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj Singh
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Jadhav
8	9	2017	Pune	2017-04-11	Delhi Daredevils	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Delhi Daredevils	97	0	SV Samson
13	14	2017	Kolkata	2017-04-15	Kolkata Knight Riders	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Kolkata Knight Riders	17	0	RV Uthappa
14	15	2017	Delhi	2017-04-15	Delhi Daredevils	Kings XI Punjab	Delhi Daredevils	bat	normal	0	Delhi Daredevils	51	0	CJ Anderson

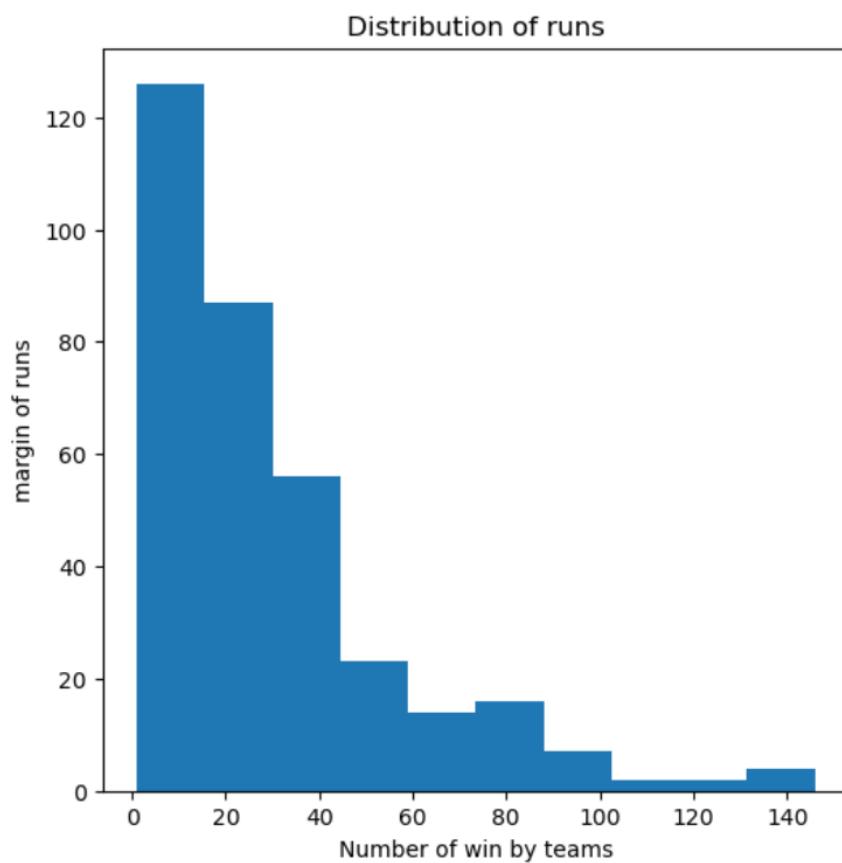
Dhanraj Tungar

```
#drawing histogram of teams won batting first
plt.figure(figsize=(6,6))

plt.hist(batting_first_df['win_by_runs'])

plt.xlabel('Number of win by teams')
plt.ylabel('margin of runs')
plt.title('Distribution of runs')

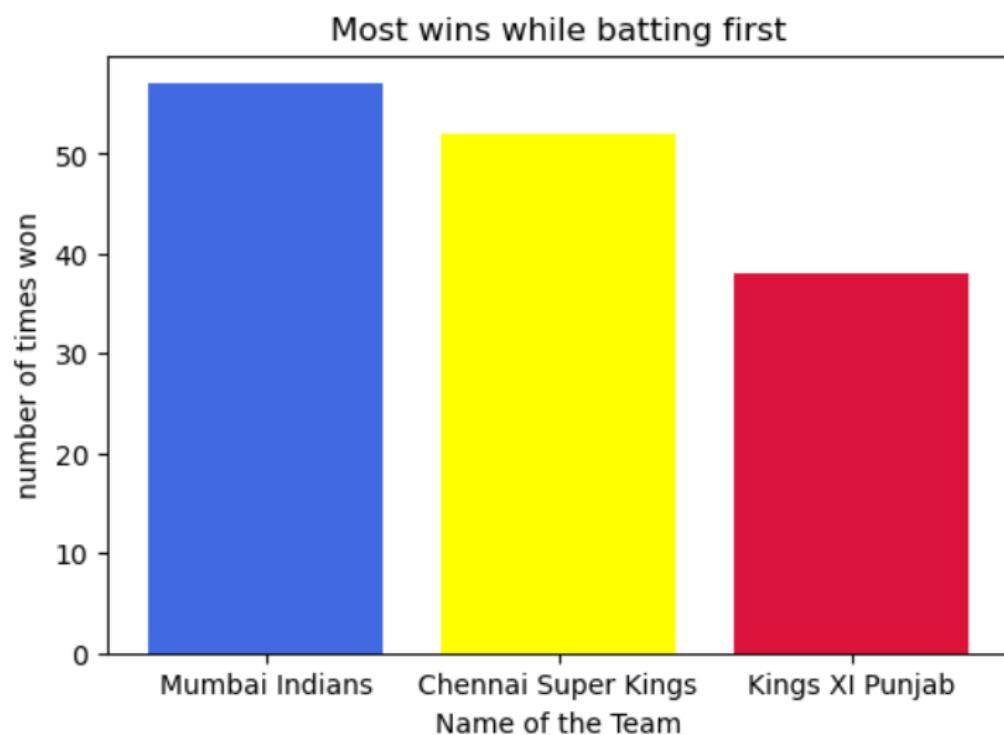
plt.show()
```



```
#teams win while batting first
batting_first_df['winner'].value_counts().head(5)
```

```
winner
Mumbai Indians      57
Chennai Super Kings 52
Kings XI Punjab     38
Kolkata Knight Riders 36
Royal Challengers Bangalore 35
Name: count, dtype: int64
```

```
#representing teams win while batting first in bar graph  
plt.figure(figsize=(6,4))  
  
plt.bar(list(batting_first_df['winner'].value_counts()[0:3].keys()),list(batting_first_df['winner'].value_counts()[0:3]),color=['royalblue','yellow','crimson'])  
  
plt.xlabel('Name of the Team')  
  
plt.ylabel('number of times won')  
  
plt.title('Most wins while batting first')  
  
plt.show()
```



```
#checking the year of seasons data available
df['season'].unique()
```

```
array([2017, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2018,
       2019], dtype=int64)
```

```
#teams qualified for playoffs in season 2019
season_2019Qualified = df[df['season'] == 2019]['winner'].value_counts().head(4)
season_2019Qualified
```

```
winner
Mumbai Indians      11
Chennai Super Kings 10
Delhi Capitals       10
Kolkata Knight Riders 6
Name: count, dtype: int64
```

```
season_2019Qualified = df[df['season'] == 2019]['winner']
```

```
season_2019Qualified.value_counts()[0:4]
```

```
winner
Mumbai Indians      11
Chennai Super Kings 10
Delhi Capitals       10
Kolkata Knight Riders 6
Name: count, dtype: int64
```

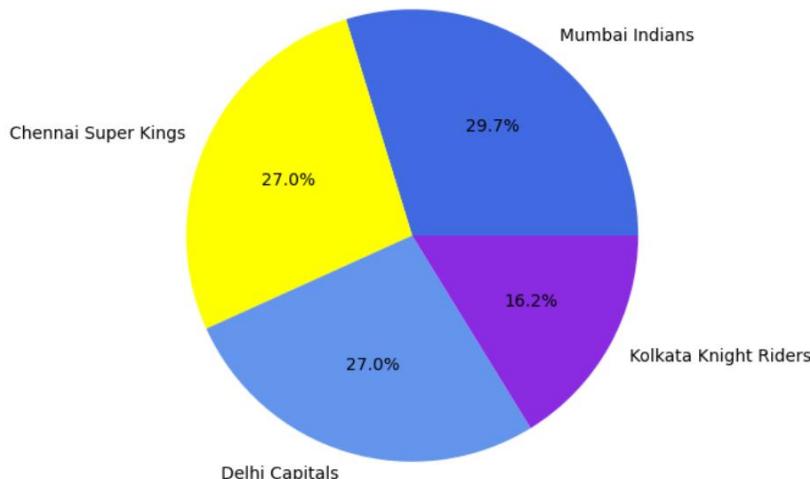
```
plt.figure(figsize=(6,6))
```

```
plt.pie(list(season_2019Qualified.value_counts()[0:4]),labels=(list(season_2019Qualified.value_counts()[0:4].keys())),autopct='%.0f%%',colors=['royalblue','yellow','cornflowerblue','blueviolet'])
```

```
plt.title('Season 2019 Qualified teams with win percentage %')
```

```
plt.show
```

Season 2019 Qualified teams with win percentage %



Dhanraj Tungar

#extracting those records where a team has won after batting second

```
batting_second=df[df['win_by_wickets']!=0]
```

#looking at the head

```
batting_second.head()
```

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPD Smith	Maharashtra Assn St
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	CA Lynn	Saurashtra Assn St
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	GJ Maxwell	Gujarat Assn St
5	6	2017	Hyderabad	2017-04-09	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad	0	9	Rashid Khan	Rajiv Gandhi International Stadium
6	7	2017	Mumbai	2017-04-09	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	N Rana	Waranasi St

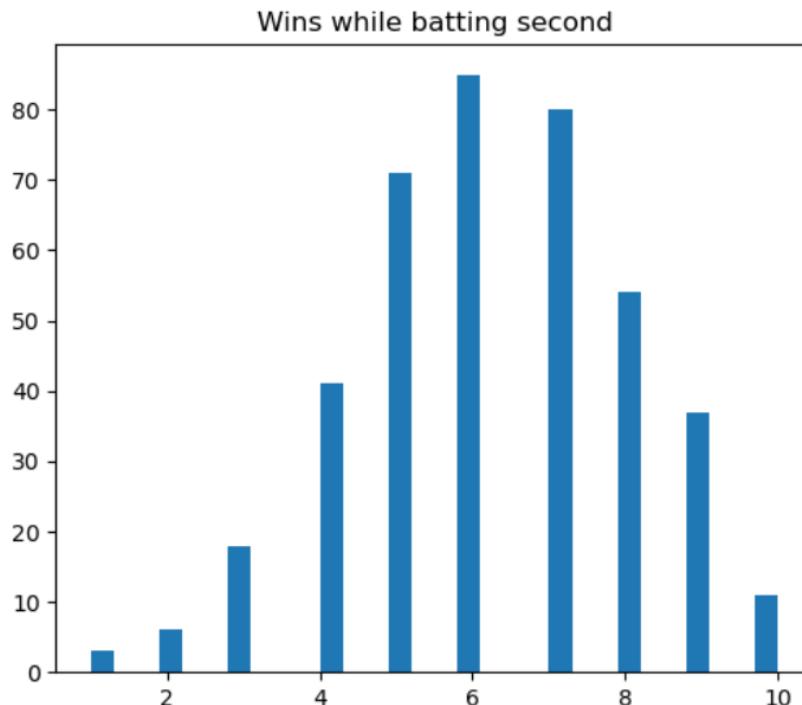
#Making a histogram for frequency of wins w.r.t number of wickets

```
plt.figure(figsize=(6,5))
```

```
plt.hist(batting_second['win_by_wickets'],bins=30)
```

```
plt.title('Wins while batting second')
```

```
plt.show()
```



#Looking at the number of matches played in each city

```
df['city'].value_counts()
```

city	
Mumbai	101
Kolkata	77
Delhi	74
Bangalore	66
Hyderabad	64
Chennai	57
Jaipur	47
Chandigarh	46
Pune	38
Durban	15
Bengaluru	14
Visakhapatnam	13
Centurion	12
Ahmedabad	12
Rajkot	10
Mohali	10
Indore	9
Dharamsala	9
Johannesburg	8
Cuttack	7
Ranchi	7
Port Elizabeth	7
Cape Town	7
Abu Dhabi	7
Sharjah	6
Raipur	6
Kochi	5
Kanpur	4
Nagpur	3
Kimberley	3
East London	3
Bloemfontein	2
Name: count, dtype: int64	