IPL 2017 DATASET ANALYSIS

In [1]:

```python
#loading a required library
import numpy as np
import pandas as pd
import seaborn as sbn
from matplotlib import pyplot as plt
```

In [2]:

```python
#loading ipl matches dataset
ipl=pd.read_csv('matches.csv')
```

In [3]:

```python
#view top 5 records in the dataset
ipl.head()
```

Out[3]:

	id	Season	city	date	team1	team2	toss_winner	toss_decision	result
0	1	IPL-2017	Hyderabad	05-04-2017	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal
1	2	IPL-2017	Pune	06-04-2017	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal
2	3	IPL-2017	Rajkot	07-04-2017	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal
3	4	IPL-2017	Indore	08-04-2017	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal
4	5	IPL-2017	Bangalore	08-04-2017	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal

In [5]:

```python
#view total rows and tables in matches dataset
ipl.shape
```

Out[5]:

```
(756, 18)
```

In [7]:

```python
#getting the frequency of most man of the match award
ipl['player_of_match'].value_counts()
```

Out[7]:

```
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: player_of_match, Length: 226, dtype: int64
```

In [9]:

```python
#getting the top 10 players with most man of the match award
ipl['player_of_match'].value_counts()[0:10]
```

Out[9]:

```
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: player_of_match, dtype: int64
```

In [10]:

```python
#getting the top 5 players with most man of the match award
ipl['player_of_match'].value_counts()[0:5]
```

Out[10]:

```
CH Gayle          21
AB de Villiers    20
RG Sharma         17
MS Dhoni          17
DA Warner         17
Name: player_of_match, dtype: int64
```

In [11]:

```python
# getting a top5 name of the players in man of the matche award
list(ipl['player_of_match'].value_counts()[0:5].keys())
```
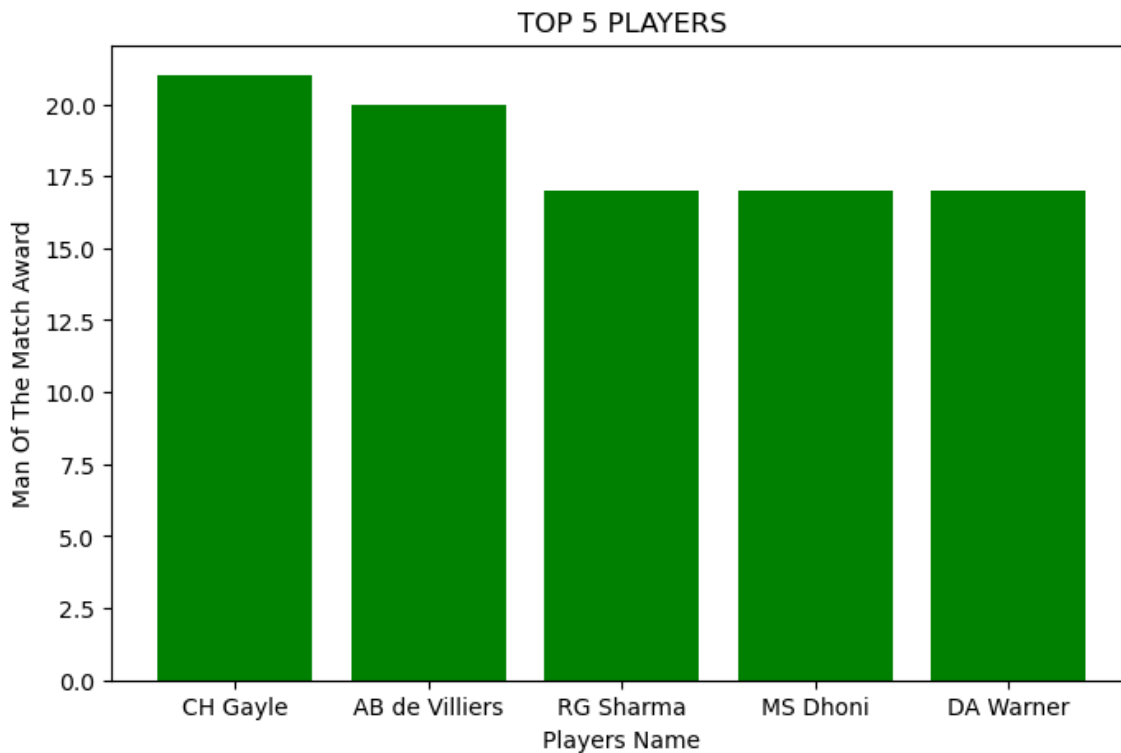
Out[11]:

```
['CH Gayle', 'AB de Villiers', 'RG Sharma', 'MS Dhoni', 'DA Warner']
```

In [13]:

```python
#here we creat a bar chart of top 5 players who awarded in man of the match awarded
plt.figure(figsize=(8,5))
plt.bar(list(ipl['player_of_match'].value_counts()[0:5].keys()),list(ipl['player_of_match
plt.title('TOP 5 PLAYERS')
plt.xlabel('Players Name')
plt.ylabel('Man Of The Match Award')
```

Out[13]:

```
Text(0, 0.5, 'Man Of The Match Award')
```


In [15]:

```python
#geting a frequency of result column
ipl['result'].value_counts()
```

Out[15]:

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

In [16]:

```python
#finding out the records where who teams wins a toss
ipl['toss_winner'].value_counts()
```

Out[16]:

```
Mumbai Indians              98
Kolkata Knight Riders       92
Chennai Super Kings         89
Royal Challengers Bangalore 81
Kings XI Punjab             81
Delhi Daredevils            80
Rajasthan Royals            80
Sunrisers Hyderabad         46
Deccan Chargers             43
Pune Warriors               20
Gujarat Lions               15
Delhi Capitals              10
Kochi Tuskers Kerala         8
Rising Pune Supergiants      7
Rising Pune Supergiant       6
Name: toss_winner, dtype: int64
```

In [18]:

```python
# Extracing the records where a team won bating first
batting_first=ipl[ipl['win_by_runs']!=0]
```

In [19]:

```python
#view top batting_first winners head
batting_first.head()
```

Out[19]:

|	id	Season	city	date	team1	team2	toss_winner	toss_decision	result
0	1	IPL-2017	Hyderabad	05-04-2017	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	norma
4	5	IPL-2017	Bangalore	08-04-2017	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	norma
8	9	IPL-2017	Pune	11-04-2017	Delhi Daredevils	Rising Pune Supergiant	Rising Pune Supergiant	field	norma
13	14	IPL-2017	Kolkata	15-04-2017	Kolkata Knight Riders	Sunrisers Hyderabad	Sunrisers Hyderabad	field	norma
14	15	IPL-2017	Delhi	15-04-2017	Delhi Daredevils	Kings XI Punjab	Delhi Daredevils	bat	norma

In [21]:

```python
#count of total wins matches by first bating
batting_first.shape
```

Out[21]:

(337, 18)

In [30]:

```python
#making a histogram of win by batting_first matche
plt.figure(figsize=(5,5))
plt.hist(batting_first['win_by_runs'],color='blue')
plt.title('Distribution of Runs')
plt.xlabel('Runs')
plt.show()
```
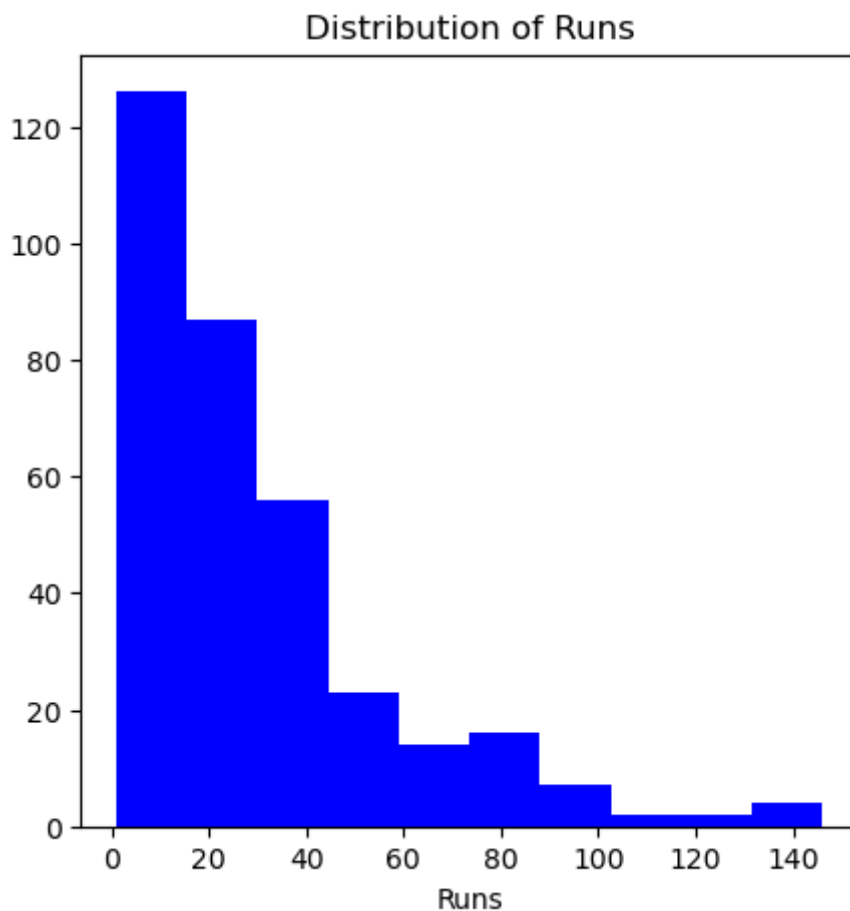
In [31]:

```python
#finding number of wins match after batting first
batting_first['winner'].value_counts()
```

Out[31]:

```
Mumbai Indians               57
Chennai Super Kings          52
Kings XI Punjab              38
Kolkata Knight Riders        36
Royal Challengers Bangalore  35
Sunrisers Hyderabad          30
Rajasthan Royals             27
Delhi Daredevils             25
Deccan Chargers              18
Pune Warriors                 6
Rising Pune Supergiant        5
Delhi Capitals                3
Kochi Tuskers Kerala          2
Rising Pune Supergiants       2
Gujarat Lions                 1
Name: winner, dtype: int64
```
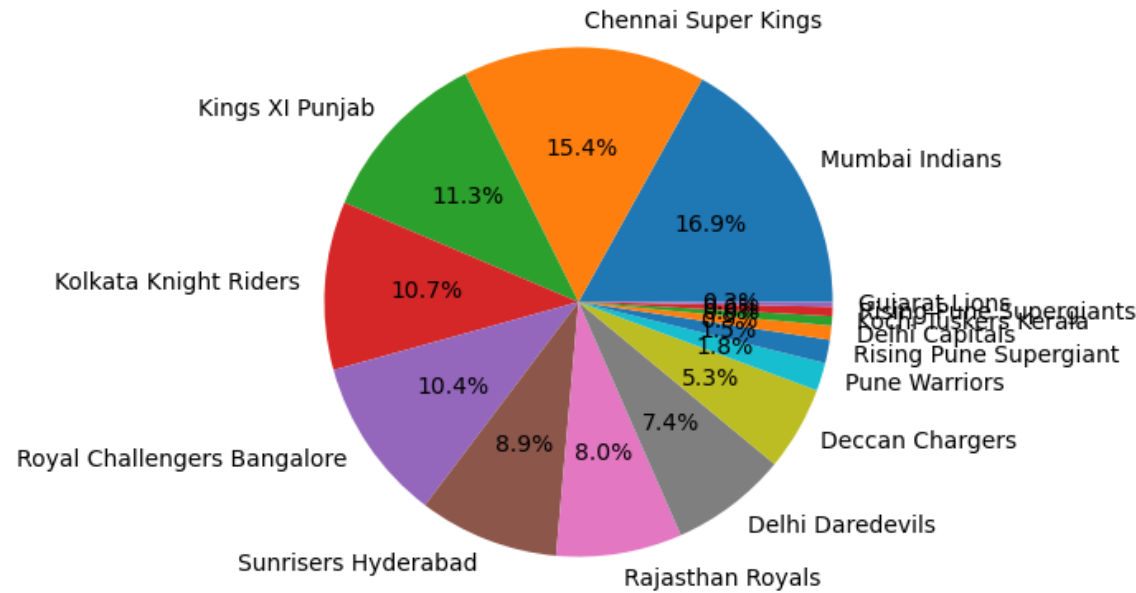
In [34]:

```python
#making a bar plot for top3 teams who wins batting-first
plt.figure(figsize=(6,6))
plt.bar(list(batting_first['winner'].value_counts()[0:3].keys()),list(batting_first['winn
plt.show()
```
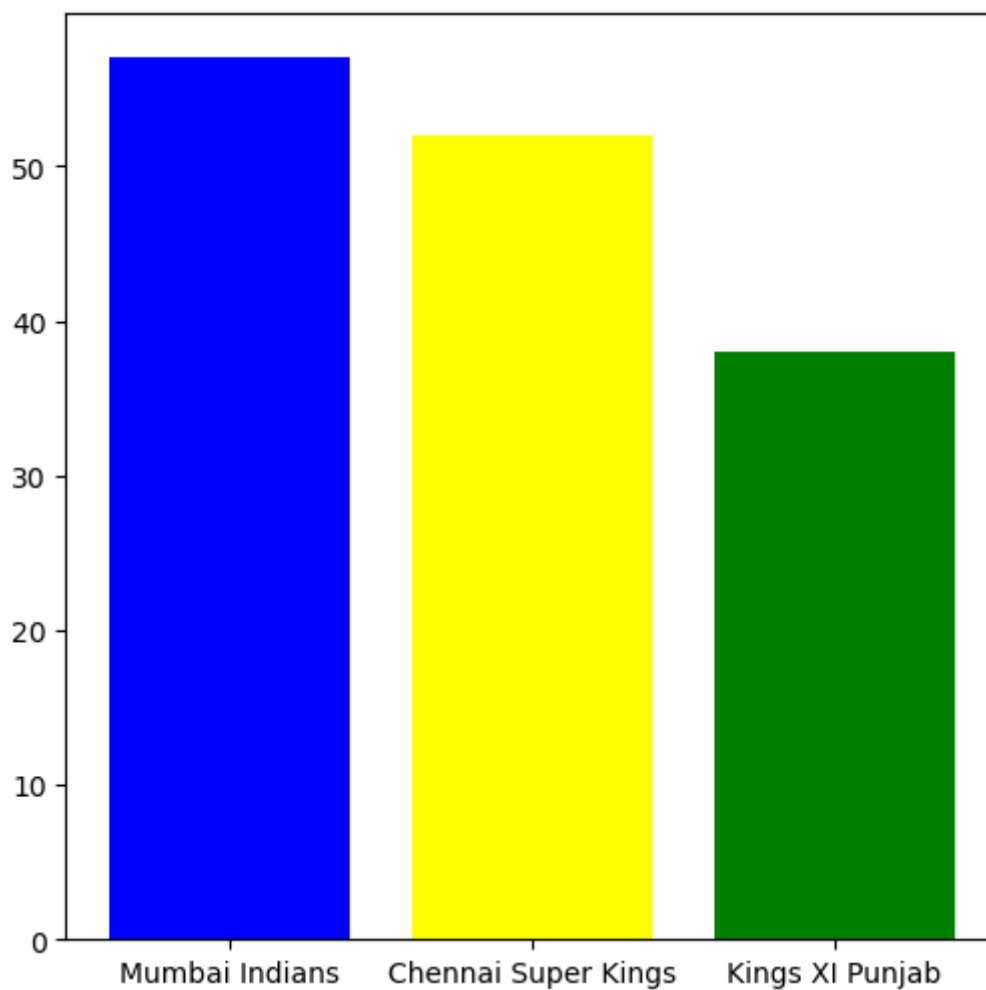
In [38]:

```python
#making a pie chart of top wins after batting-first
plt.figure(figsize=(5,5))
plt.pie(list(batting_first['winner'].value_counts()),labels=list(batting_first['winner'].
plt.show()
```


In [39]:

```python
# Extracting those recordes where a team has won after batting secccond
batting_second=ipl[ipl['win_by_wickets']!=0]
batting_second.head()
```

Out[39]:

	id	Season	city	date	team1	team2	toss_winner	toss_decision	result	d
1	2	IPL-2017	Pune	06-04-2017	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	
2	3	IPL-2017	Rajkot	07-04-2017	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	
3	4	IPL-2017	Indore	08-04-2017	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	
5	6	IPL-2017	Hyderabad	09-04-2017	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	
6	7	IPL-2017	Mumbai	09-04-2017	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	

In [45]:

```python
# Making a histogram for frequencyof wins mattch numbers of wicket
plt.figure(figsize=(5,5))
plt.hist(batting_second['win_by_wickets'],bins=30)
plt.show()
```
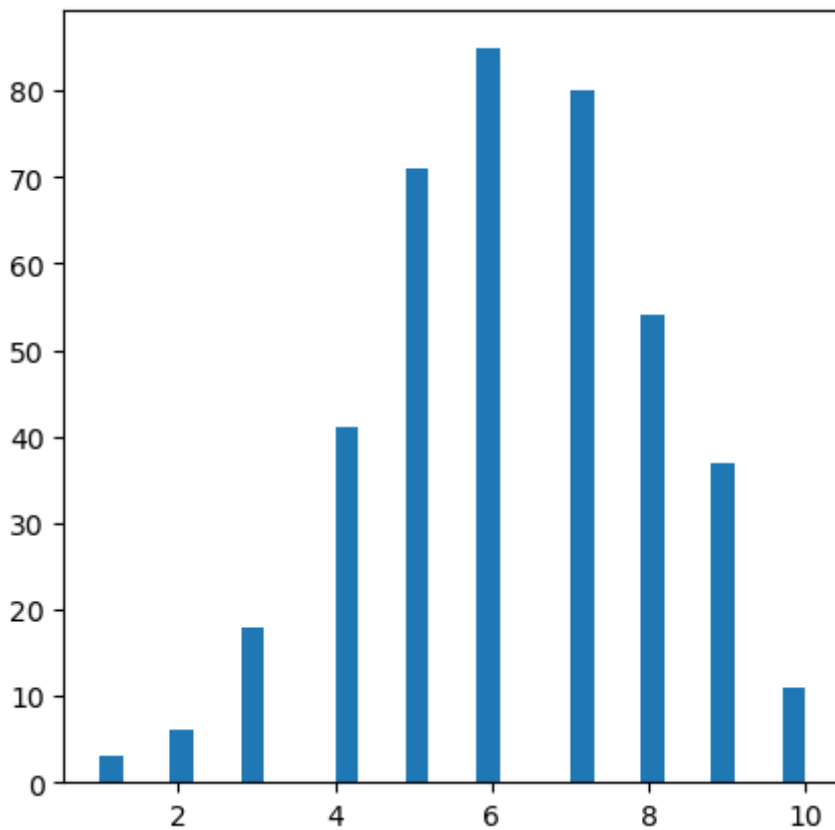

In [47]:

```python
# Find a numbers of teams frequency who wins batting seccond
batting_second['winner'].value_counts()
```

Out[47]:

```
Kolkata Knight Riders          56
Mumbai Indians                 50
Royal Challengers Bangalore    48
Chennai Super Kings            48
Rajasthan Royals               46
Kings XI Punjab                42
Delhi Daredevils               42
Sunrisers Hyderabad            27
Gujarat Lions                  12
Deccan Chargers                11
Pune Warriors                   6
Delhi Capitals                  6
Rising Pune Supergiant          5
Kochi Tuskers Kerala            4
Rising Pune Supergiants         3
Name: winner, dtype: int64
```
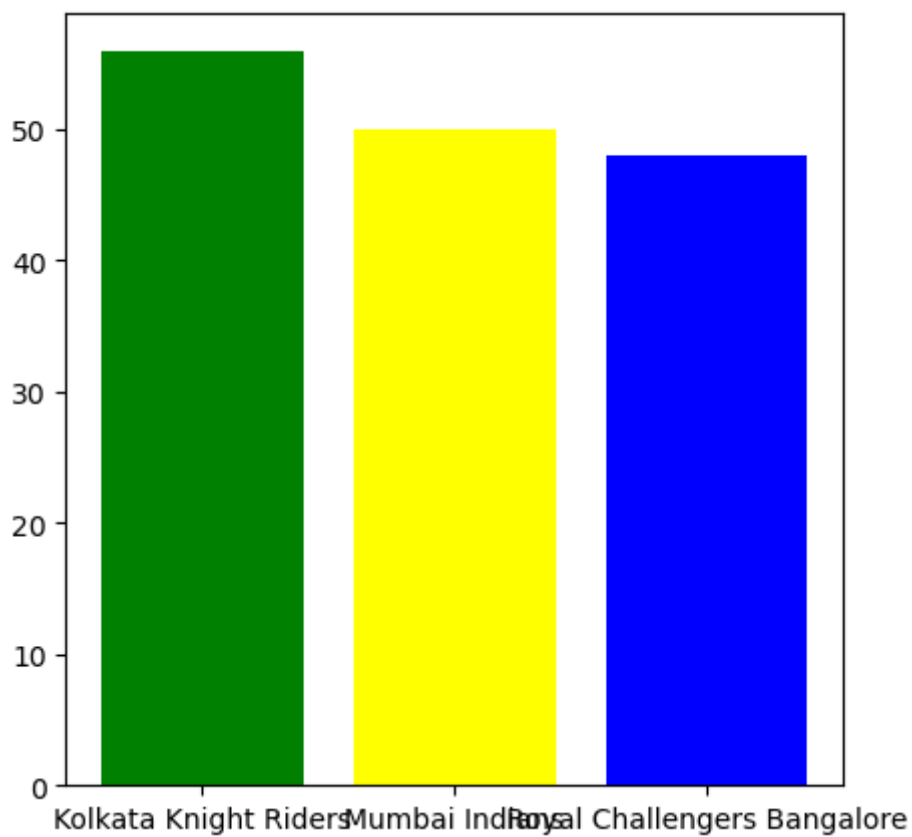
In [50]:

```python
#creating a bar plot for top 3 team who wins after batting_seccond
plt.figure(figsize=(5,5))
plt.bar(list(batting_second['winner'].value_counts()[0:3].keys()),list(batting_second['wi
plt.show()
```

In [52]:

```python
# Creating a pie chart of winer teams who batting seccond
plt.figure(figsize=(5,5))
plt.pie(list(batting_second['winner'].value_counts()),labels=list(batting_second['winner'
plt.show()
```
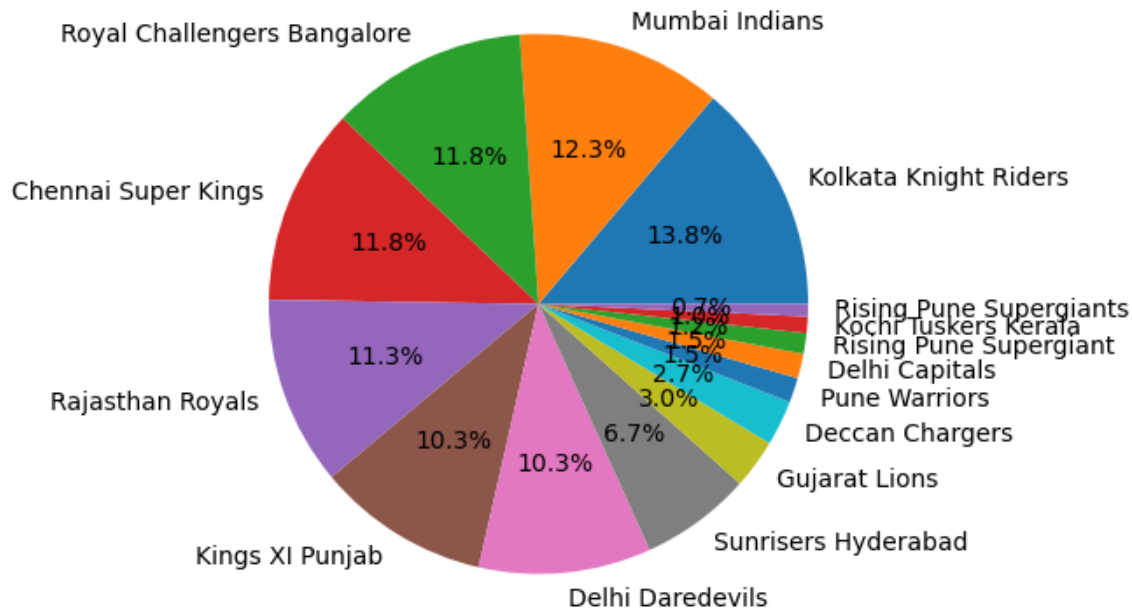

In [54]:

```python
# Find a numbers of matches playes in those seasion
ipl['Season'].value_counts()
```

Out[54]:

```
IPL-2013    76
IPL-2012    74
IPL-2011    73
IPL-2010    60
IPL-2014    60
IPL-2016    60
IPL-2018    60
IPL-2019    60
IPL-2017    59
IPL-2015    59
IPL-2008    58
IPL-2009    57
Name: Season, dtype: int64
```

In [57]:

```python
# Find a numbers of top 5 matches playes in those seasion
ipl['Season'].value_counts()[0:5]
```

Out[57]:

```
IPL-2013    76
IPL-2012    74
IPL-2011    73
IPL-2010    60
IPL-2014    60
Name: Season, dtype: int64
```

In [55]:

```python
# Find a numbers of city where matches are played
ipl['city'].value_counts()
```

Out[55]:

```
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: city, dtype: int64
```

In [56]:

```python
# Find a numbers of top 5 city where matches are played
ipl['city'].value_counts()[0:5]
```

Out[56]:

```
Mumbai        101
Kolkata        77
Delhi          74
Bangalore      66
Hyderabad      64
Name: city, dtype: int64
```

In [58]:

```python
# find a how many matches are win after winning toss
import numpy as np
np.sum(ipl['toss_winner']==ipl['winner'])
```

Out[58]:

```
393
```

In [59]:

```python
# Finf a what ratio of if we win toss then we win matche
#first we found a number of matches
#and we divided into match win after toss win
ipl.shape
```

Out[59]:

```
(756, 18)
```

In [61]:

```python
# persont of win matches where we win a toss
393/756
```

Out[61]:

```
0.5198412698412699
```

In [ ]: