In [1]:
```
#loading the required libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
#loading the ipl matches dataset
ipl=pd.read_csv('matches.csv')
```

In [3]:
```
#having a glance at first five records of the dataset
ipl.head()
```

Out[3]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	venue	umpire1	um
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj Singh	Rajiv Gandhi International Stadium, Uppal	AY Dandekar	NJ
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 Cricket Association Stadium	A Nand Kishore	S
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 Cricket Association Stadium	Nitin Menon	CK Na
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	Holkar Cricket Stadium	AK Chaudhary	Sharma
4	5	2017	Bangalore	2017-04-09	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Jadhav	M Chinnaswamy Stadium	NaN	N

In [5]:
```
#lookin at number of rows and columns in the dataset
ipl.shape
```

Out[5]:
```
(636, 18)
```

In [6]:
```
#getting the frequency of most man of the match awards
ipl['player_of_match'].value_counts()
```

Out[6]:
```
CH Gayle        18
AB de Villiers  15
DA Warner       15
AB de Villiers  15
SK Raina        14
                 ..
AD Mascarenhas   1
PA Patel         1
MEK Hussey       1
SA Yadav         1
KJ Harrison      1
Name: player_of_match, Length: 201, dtype: int64
```

In [7]:
```
#getting the top 10 players with most man of the match awards
ipl['player_of_match'].value_counts()[0:10]
```

Out[7]:
```
CH Gayle        18
AB de Villiers  15
DA Warner       15
AB de Villiers  15
SK Raina        14
MS Dhoni        13
RG Sharma       14
G Gambhir       13
MS Dhoni        12
YK Pathan       12
MEK Hussey      12
Name: player_of_match, dtype: int64
```

In [8]:
```
#getting the top 5 players with most man of the match awards
ipl['player_of_match'].value_counts()[0:5]
```

Out[8]:
```
CH Gayle        18
YK Pathan       16
DA Warner       15
AB de Villiers  15
SK Raina        14
Name: player_of_match, dtype: int64
```

In [9]:
```
list(ipl['player_of_match'].value_counts[0:5].keys())
```

Out[9]:
```
['CH Gayle', 'YK Pathan', 'DA Warner', 'AB de Villiers', 'SK Raina']
```

In [10]:
```
#making a bar plot for top 5 players with most man of the match awards
plt.figure(figsize=(8,5))
plt.bar(list(ipl['player_of_match'].value_counts[0:5].keys()),list(ipl['player_of_match'].value_counts[0:5]),color='g')
plt.show()
```

In [11]:
```
#getting the frequency of result column
ipl['result'].value_counts()
```

Out[11]:
```
normal    626
tie         7
no result   3
Name: result, dtype: int64
```

In [12]:
```
#finding out the numbers of toss wins by each team
ipl['toss_winner'].value_counts()
```

Out[12]:
```
Mumbai Indians                 85
Kolkata Knight Riders          78
Kings XI Punjab                75
Royal Challengers Bangalore    76
Kings XI Punjab                68
Chennai Super Kings            66
Rajasthan Royals               66
Delhi Daredevils               63
Sunrisers Hyderabad            35
Pune Warriors                  20
Kochi Tuskers Kerala           14
Kochi Tuskers Kerala            8
Rising Pune Supergiants         7
Rising Pune Supergiant          6
Name: toss_winner, dtype: int64
```

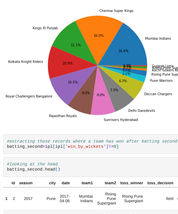
In [13]:
```
#extracting the records where a team won batting first
batting_first=ipl[ipl['win_by_runs']!=0]
```

In [14]:
```
#looking at the head
batting_first.head()
```

Out[14]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	venue	umpire1	umpire
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj Singh	Rajiv Gandhi International Stadium, Uppal	AY Dandekar	N
4	5	2017	Bangalore	2017-04-09	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Jadhav	M Chinnaswamy Stadium	NaN	Na
8	9	2017	Pune	2017-04-12	Delhi Daredevils	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Delhi Daredevils	97	0	SV Samson	Maharashtra Cricket Association Stadium	AY Dandekar	S Ra
10	11	2017	Kolkata	2017-04-15	Kolkata Knight Riders	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Kolkata Knight Riders	17	0	RV Uthappa	Eden Gardens	AY Dandekar	S
14	15	2017	Delhi	2017-04-19	Delhi Daredevils	Kings XI Punjab	Delhi Daredevils	bat	normal	0	Delhi Daredevils	51	0	CJ Anderson	Feroz Shah Kotla	YC Barde	Ne Meno

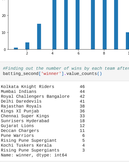
In [15]:
```
# making a HISTOGRAM
plt.figure(figsize=(5,5))
plt.hist(batting_first['win_by_runs'])
plt.title('Distribution of runs')
plt.xlabel('Runs')
plt.show()
```



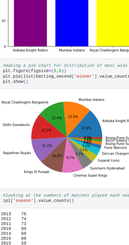
In [16]:
```
#finding out the number of wins by each team after batting first
batting_first['winner'].value_counts()
```

Out[16]:
```
Mumbai Indians                 47
Chennai Super Kings            38
Kings XI Punjab                32
Kolkata Knight Riders          31
Royal Challengers Bangalore    29
Rajasthan Royals               23
Sunrisers Hyderabad            23
Delhi Daredevils               19
Deccan Chargers                18
Pune Warriors                   6
Kochi Tuskers Kerala            6
Kochi Tuskers Kerala            2
Rising Pune Supergiants         2
Rising Pune Supergiant          1
Name: winner, dtype: int64
```

In [17]:
```
#making a bar plot for top 3 teams with most wins after batting first
plt.figure(figsize=(6,6))
plt.bar(list(batting_first['winner'].value_counts[0:3].keys()),list(batting_first['winner'].value_counts[0:3]),color=['yellow','orange','red'])
plt.show()
```



In [18]:
```
#making a pie chart
plt.figure(figsize=(7,7))
plt.pie(list(batting_first['winner'].value_counts()),labels=list(batting_first['winner'].value_counts().keys()),autopct='%0.1f%%')
plt.show()
```



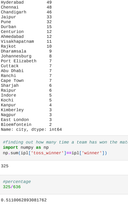
In [19]:
```
#extracting those records where a team has won after batting second
batting_second=ipl[ipl['win_by_wickets']!=0]
```

In [20]:
```
#looking at the head
batting_second.head()
```

Out[20]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	venue	umpire1	umpire
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 Cricket Association Stadium	A Nand Kishore	S Ra
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 Cricket Association Stadium	Nitin Menon	CK Nandan
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	Holkar Cricket Stadium	AK Chaudhary	Dharmshuda
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, Uppal	A Deshmukh	NJ Llong
6	7	2017	Mumbai	2017-04-09	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	N Rana	Wankhede Stadium	Nitin Menon	CK Nandan

In [21]:
```
#making a histogram for frequency of wins by numbers of wickets
plt.figure(figsize=(7,7))
plt.hist(batting_second['win_by_wickets'],bins=30)
plt.show()
```



In [22]:
```
#finding out the number of wins by each team after batting second
batting_second['winner'].value_counts()
```

Out[22]:
```
Kolkata Knight Riders  48
Mumbai Indians         44
Royal Challengers Bangalore  42
Delhi Daredevils       41
Kings XI Punjab        36
Kings XI Punjab        36
Chennai Super Kings    33
Rajasthan Royals       28
Sunrisers Hyderabad    12
Deccan Chargers        11
Gujarat Lions          11
Rising Pune Supergiant  5
Kochi Tuskers Kerala   4
Rising Pune Supergiants  4
Name: winner, dtype: int64
```

In [23]:
```
#making a bar plot for top 3 teams with most wins after batting second
plt.figure(figsize=(7,7))
plt.bar(list(batting_second['winner'].value_counts[0:3].keys()),list(batting_second['winner'].value_counts[0:3]),color=['purple','blue','yellow'])
plt.show()
```


In [24]:
```
#making a pie chart for distribution of most wins after batting second
plt.figure(figsize=(5,5))
plt.pie(list(batting_second['winner'].value_counts()),labels=list(batting_second['winner'].value_counts().keys()),autopct='%0.1f%%')
plt.show()
```


In [25]:
```
#looking at the numbers of matches played each season
ipl['season'].value_counts()
```

Out[25]:
```
2013  76
2012  74
2011  73
2017  72
2014  60
2016  60
2015  59
2010  60
2008  58
2009  57
Name: season, dtype: int64
```

In [26]:
```
#looking at the numbers of matches played each city
ipl['city'].value_counts()
```

Out[26]:
```
Mumbai          85
Bangalore       66
Kolkata         61
Delhi           60
Hyderabad       49
Chennai         48
Chandigarh      46
Jaipur          33
Pune            32
Durban          15
Centurion       12
Ahmedabad       12
Visakhapatnam   11
Rajkot          10
Dharamsala       9
Johannesburg     8
Port Elizabeth   7
Cuttack          7
East London      7
Kanpur           7
Cape Town        7
Raipur           6
Ranchi           6
Indore           5
Kochi            5
Bloemfontein     4
Kimberley        3
Nagpur           3
East London      2
Bangalore        1
Name: city, dtype: int64
```

In [27]:
```
#finding out how many time a team has won the match after winning the toss
import numpy as np
np.sum(ipl['toss_winner']==ipl['winner'])
```

Out[27]:
```
325
```

In [28]:
```
#percentage
325/636
```

Out[28]:
```
0.5110062893081762
```