

# Assignment 3

## Working with IPL Data

### 1. Importing Libraries

```
import pandas as pd
import matplotlib.pyplot as plt
```

### 2. Loading csv file

```
ipl=pd.read_csv('matches.csv')
```

### 3. First five records of the dataset

```
ipl.head()
```

season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match
2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj Singh
2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPD Smith
2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	CA Lynn
2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	GJ Maxwell
2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Jadhav

### 4. Number of rows and columns in the dataset

```
ipl.shape
```

```
(756, 18)
```

### 5. Frequency of most man of the match awards

```
ipl['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
```

## 6. Top 10 players with most man of the match awards

```
ipl['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
```

## 7. Top 5 players with most man of the match awards

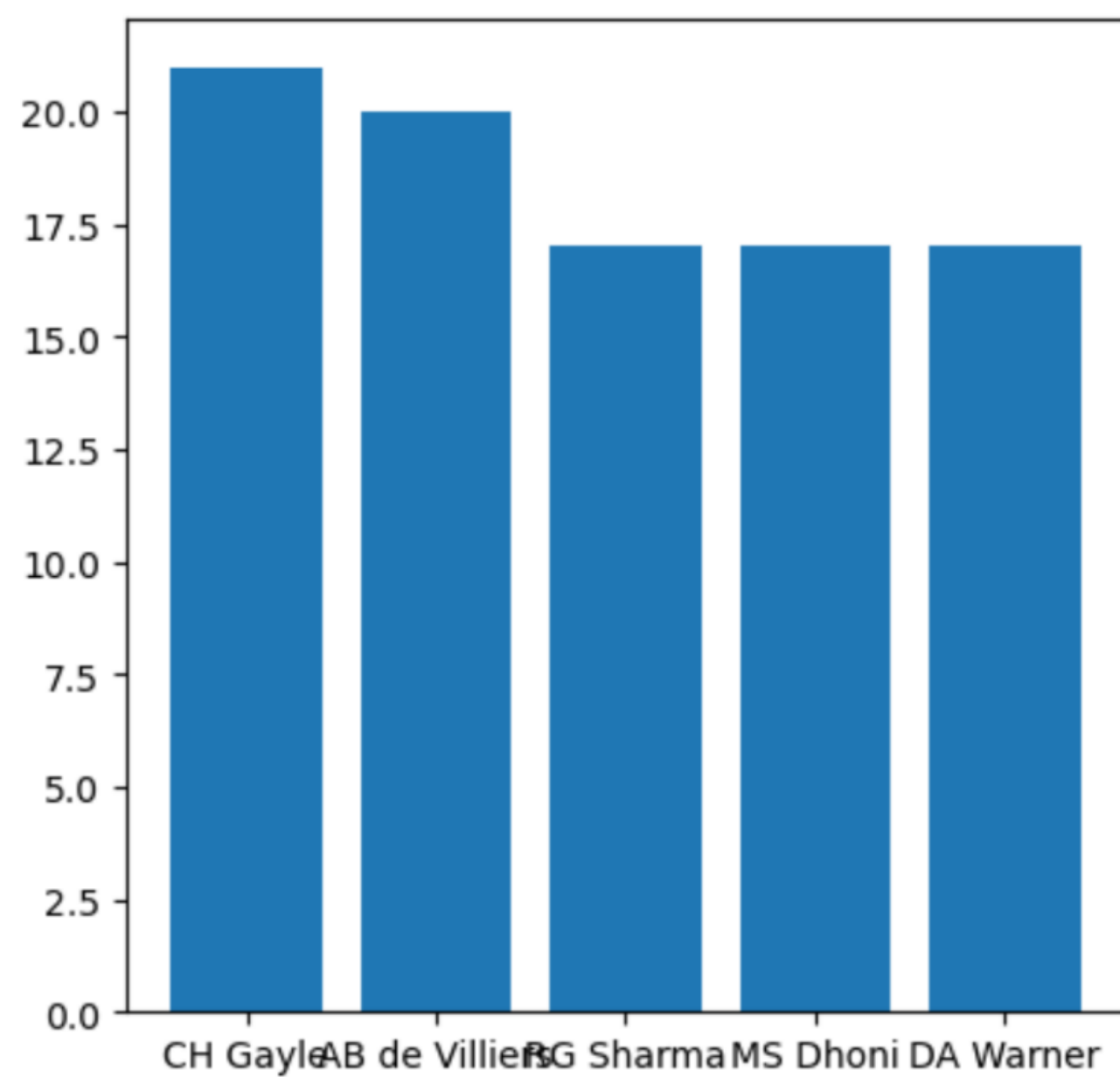
```
ipl['player_of_match'].value_counts()[0:5]
```

---

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

## 8. Bar-plot for the top 5 players with most man of the match awards

```
plt.figure(figsize=(5,5))
plt.bar(list(ipl['player_of_match'].value_counts()[0:5].keys()),list(ipl['player_of_match'].value_counts()[0:5]))
plt.show()
```



## 9. Frequency of result column

```
ipl['result'].value_counts()
```

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

## 10. Number of toss wins w.r.t each team

```
ipl['toss_winner'].value_counts()
```

```
toss_winner
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: count, dtype: int64
```

11. Records where a team won batting first

```
batting_first=ipl[ipl['win_by_runs']!=0]
```

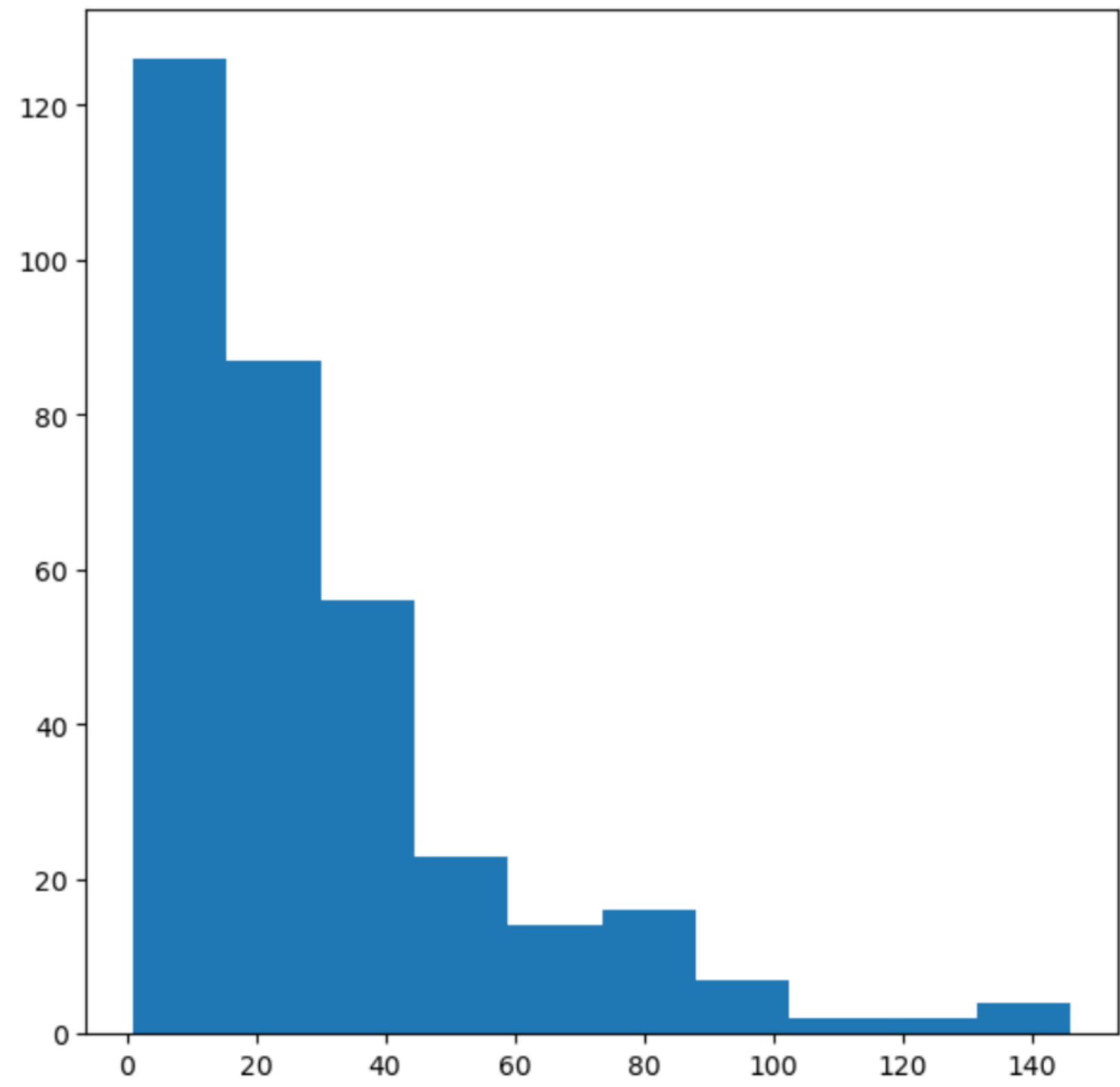
12. First five records of the dataset

```
batting_first.head()
```

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_n
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj : 100
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM Ja 100
8	9	2017	Pune	2017-04-11	Delhi Daredevils	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Delhi Daredevils	97	0	SV Sa 100
13	14	2017	Kolkata	2017-04-15	Kolkata Knight Riders	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Kolkata Knight Riders	17	0	RV Uth 100
14	15	2017	Delhi	2017-04-15	Delhi Daredevils	Kings XI Punjab	Delhi Daredevils	bat	normal	0	Delhi Daredevils	51	0	CJ And 100

13. Making a histogram

```
plt.figure(figsize=(7,7))
plt.hist(batting_first['win_by_runs'])
plt.show()
```



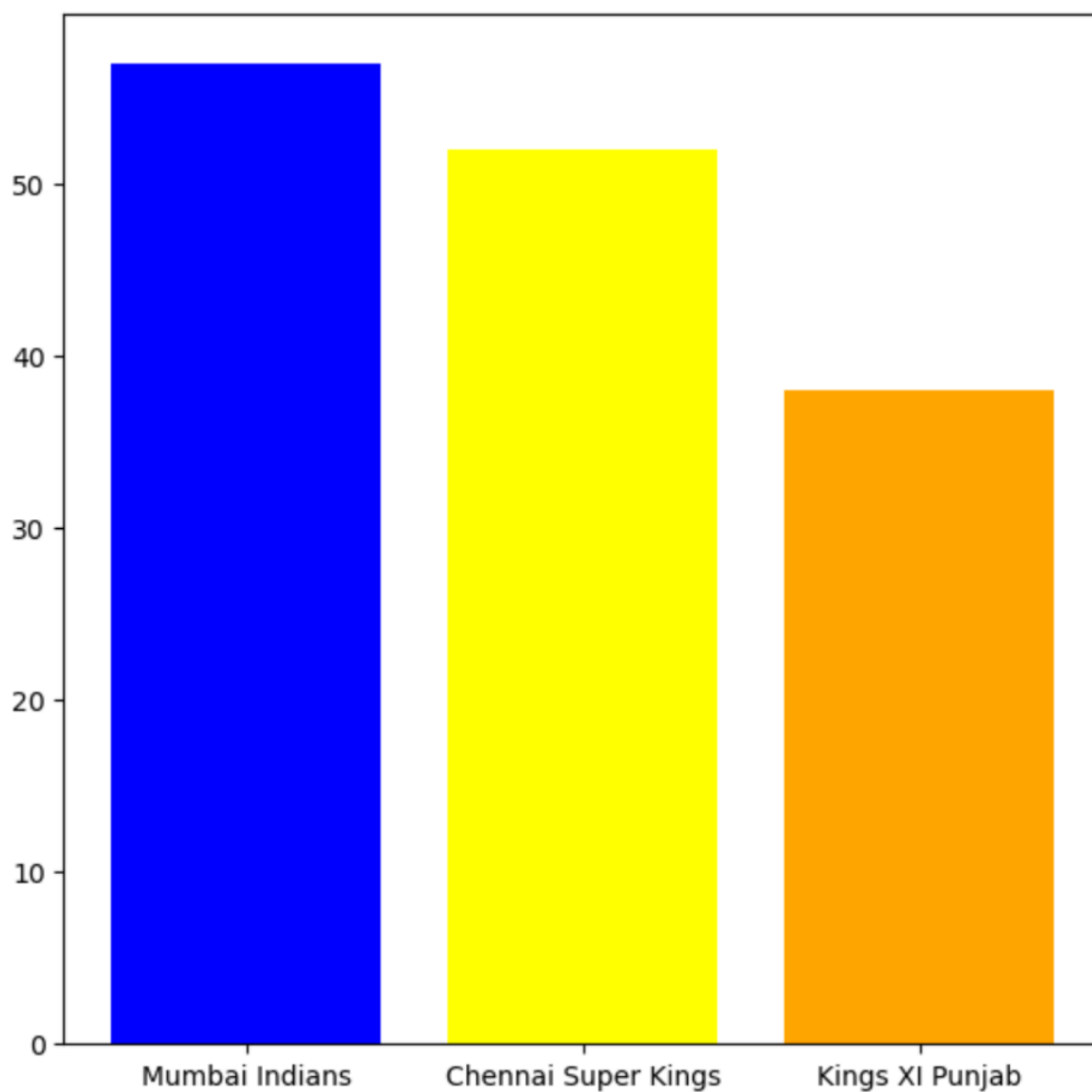
## 14. Number of wins w.r.t each team after batting first

```
batting_first['winner'].value_counts()
```

```
winner
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: count, dtype: int64
```

## 15. Bar-plot for top 3 teams with most wins after batting first

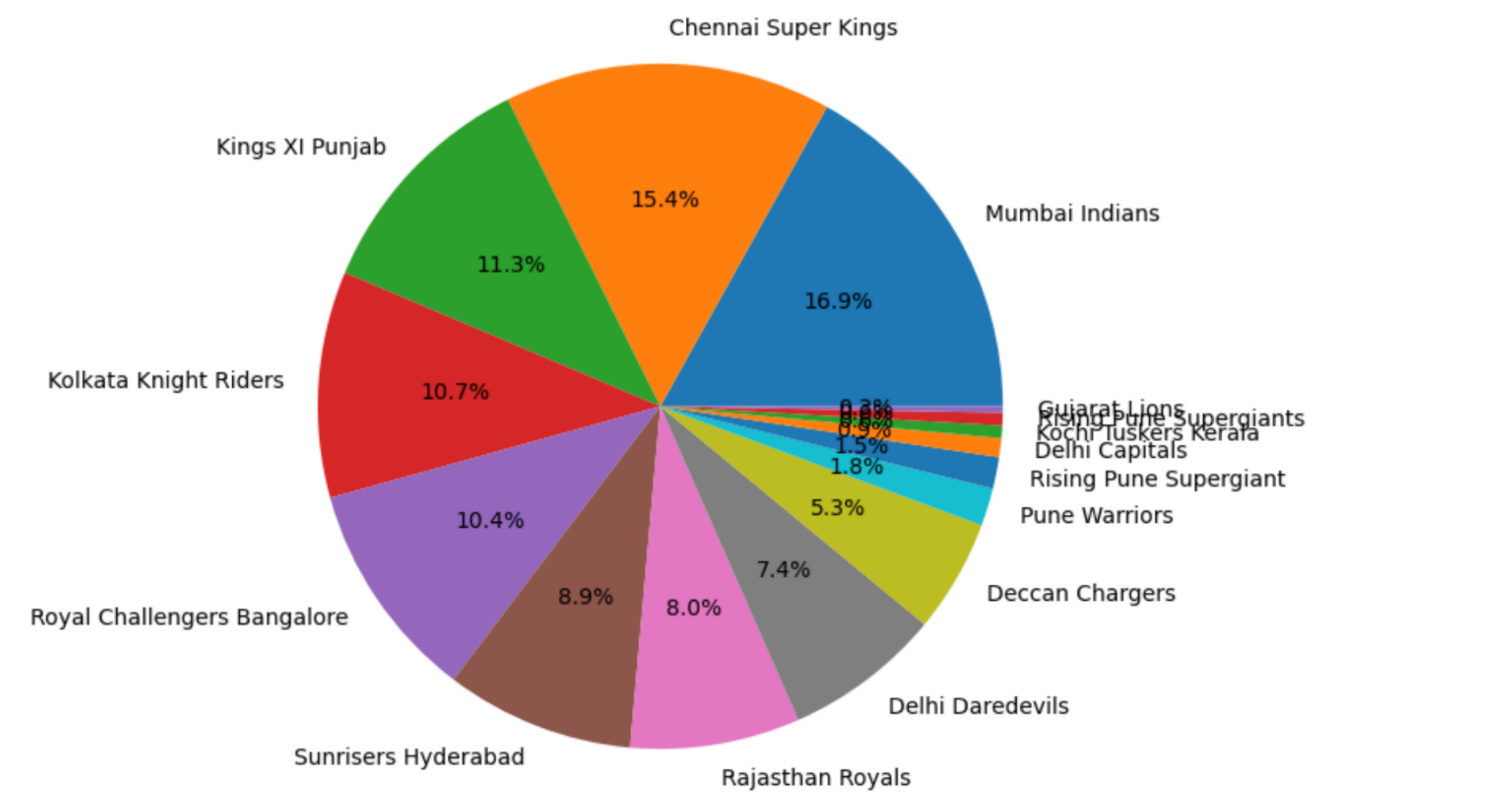
```
plt.figure(figsize=(7,7))
plt.bar(list(batting_first['winner'].value_counts()[0:3].keys()),list(batting_first['winner'].value_counts
()[0:3]),color=["blue","yellow","orange"])
plt.show()
```





16. 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()
```



17. Records where a team has won after batting second

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

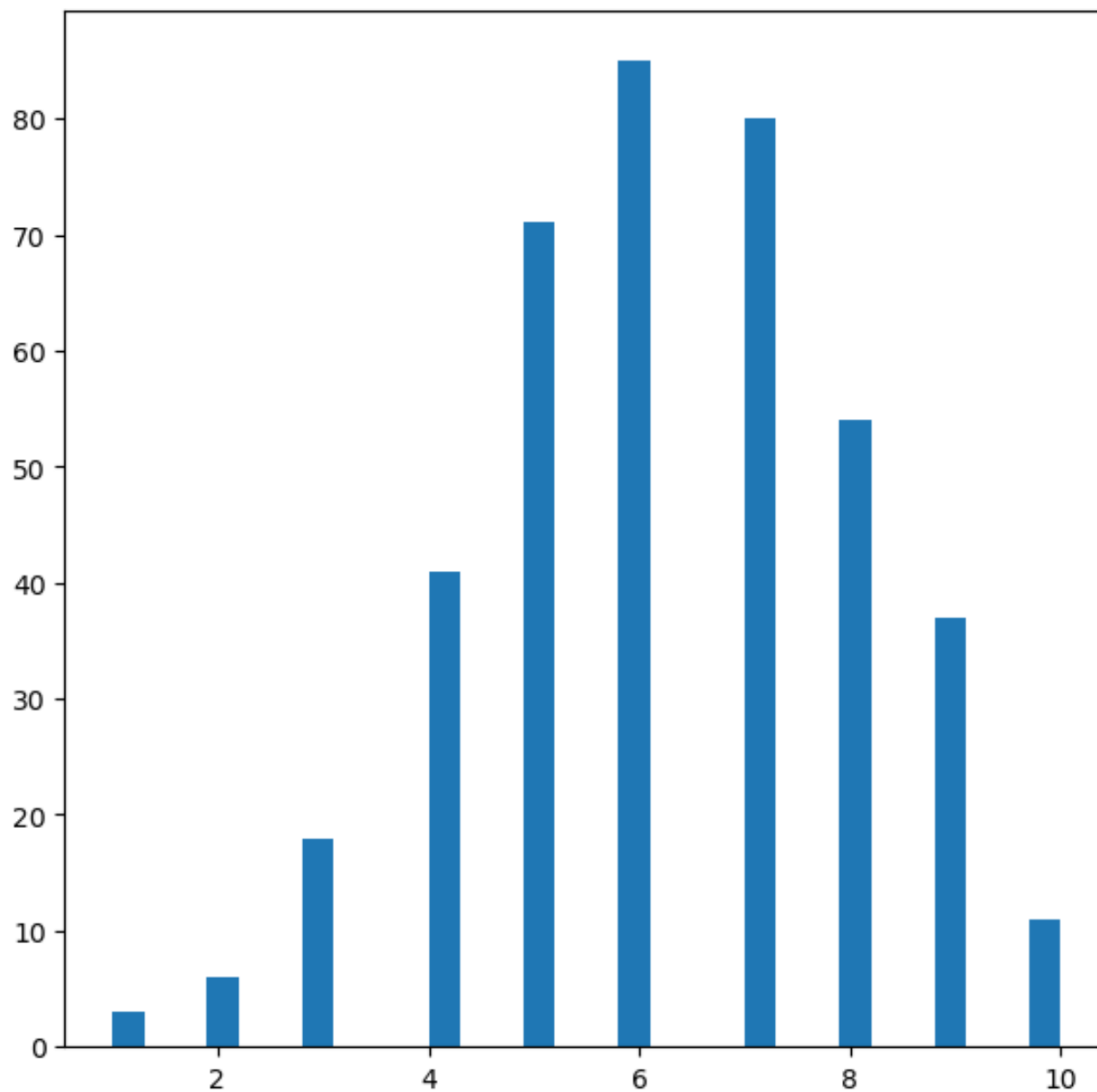
18. Looking at the head

```
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	Maha Asso 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	Sau Asso S
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	S
5	6	2017	Hyderabad	2017-04-09	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad	0	9	Rashid Khan	Rajiv Intern St
6	7	2017	Mumbai	2017-04-09	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians	0	4	N Rana	War S

## 19. Histogram for frequency of wins w.r.t number of wickets

```
plt.figure(figsize=(7,7))
plt.hist(batting_second['win_by_wickets'],bins=30)
plt.show()
```



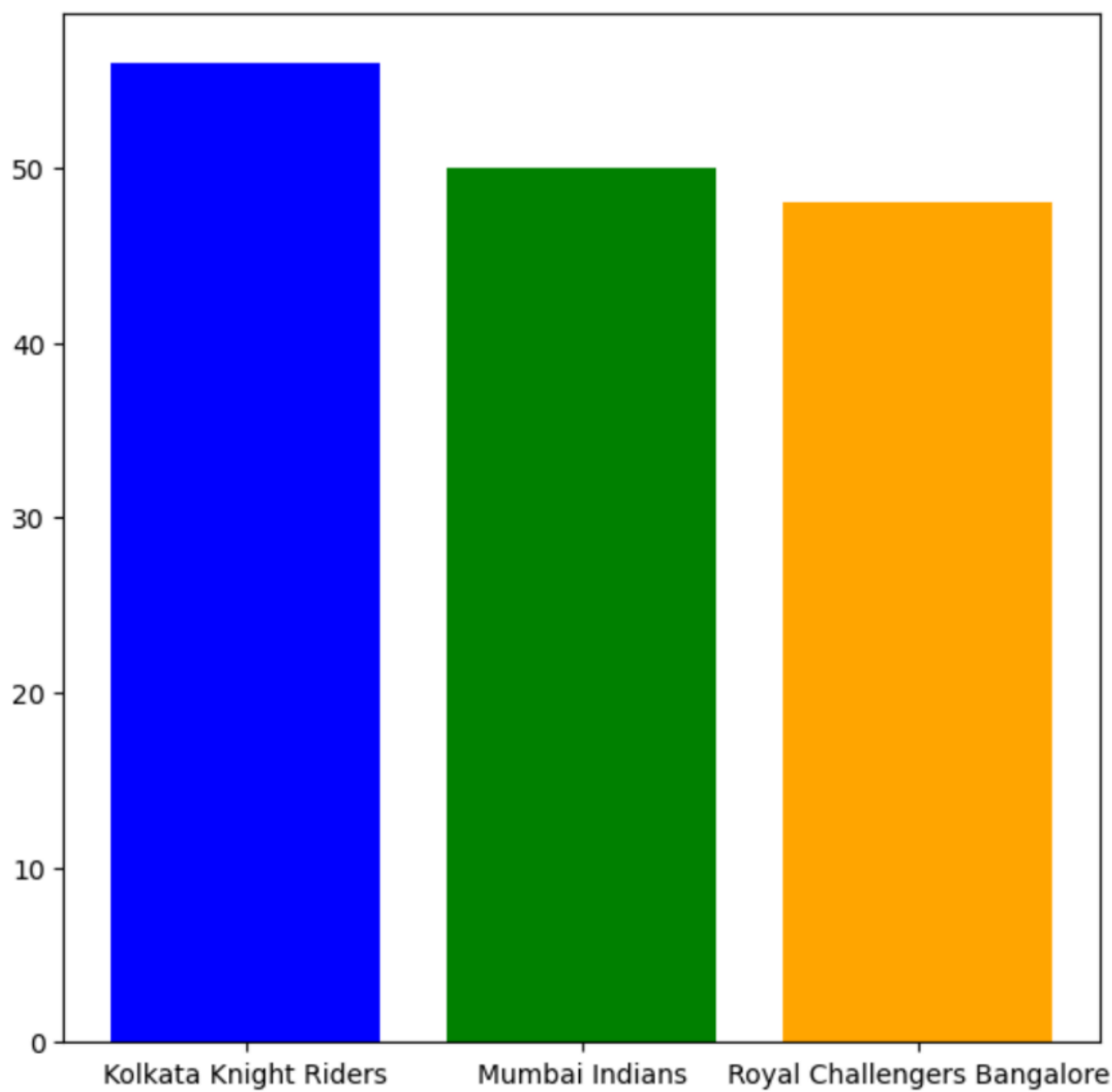
## 20. Frequency of number of wins w.r.t each time after batting second

```
batting_second['winner'].value_counts()
```

```
winner
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: count, dtype: int64
```

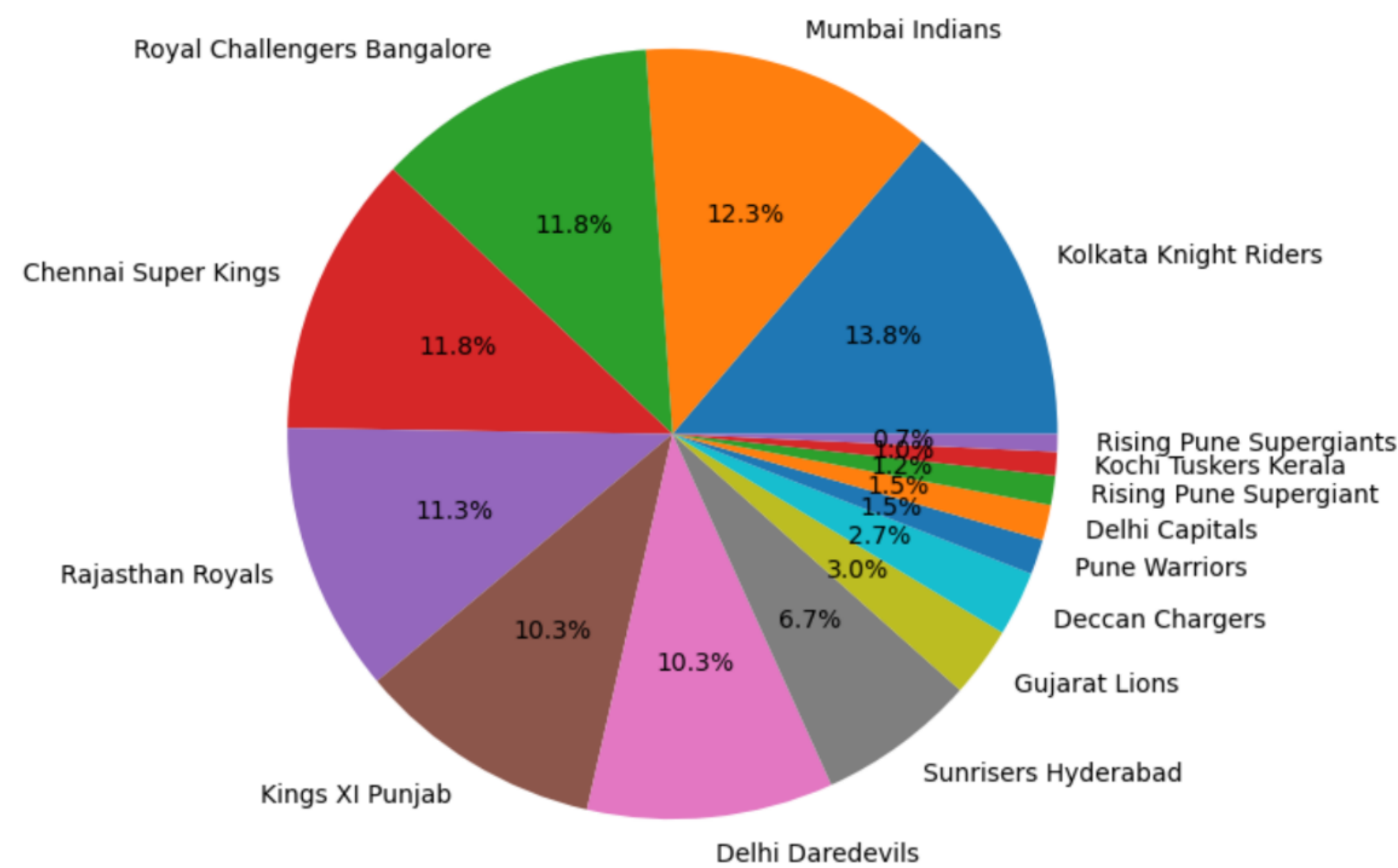
21. Bar plot for top-3 teams with most wins after batting second

```
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=["blue","green","orange"])
plt.show()
```



22. Pie chart for distribution of most wins after batting second

```
plt.figure(figsize=(7,7))
plt.pie(list(batting_second['winner'].value_counts()),labels=list(batting_second['winner'].value_counts().keys()),autopct='%0.1f%%')
plt.show()
```





23. Number of matches played each season

```
ipl['season'].value_counts()
```

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

24. Number of matches played in each city

```
ipl['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	

25. How many times a team has won the match after winning the toss

```
import numpy as np

np.sum(ipl['toss_winner']==ipl['winner'])
```

393

26.

```
deliveries=pd.read_csv('deliveries.csv')
deliveries.head()
```

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	bye_runs	legbye_runs	noball_runs	penalty_runs	batsman
0	2	1	Mumbai Indians	Chennai Super Kings	15	1	Virat Kohli	David Warner	Deepak Chahar	0	1	0	1	0	
1	5	2	Chennai Super Kings	Kolkata Knight Riders	12	5	KL Rahul	Faf du Plessis	Pat Cummins	0	1	0	1	0	
2	9	2	Kolkata Knight Riders	Kolkata Knight Riders	7	4	KL Rahul	Ruturaj Gaikwad	Deepak Chahar	0	1	1	1	0	
3	8	1	Chennai Super Kings	Kolkata Knight Riders	18	3	KL Rahul	David Warner	Jasprit Bumrah	0	0	1	0	0	
4	4	2	Chennai Super Kings	Mumbai Indians	17	2	KL Rahul	Ruturaj Gaikwad	Jasprit Bumrah	1	1	1	0	0	

27.

```
deliveries['match_id'].unique()
```

array([ 2, 5, 9, 8, 4, 1, 7, 6, 10, 3], dtype=int64)

28.

```
match_1=deliveries[deliveries['match_id']==1]
match_1.head()
```

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	bye_runs	legbye_runs	noball_runs	penalty_runs	batsman
8	1	2	Mumbai Indians	Chennai Super Kings	5	4	Virat Kohli	Faf du Plessis	Jasprit Bumrah	0	0	0	1	0	
13	1	2	Mumbai Indians	Royal Challengers Bangalore	19	3	Rohit Sharma	Ruturaj Gaikwad	Rashid Khan	0	0	0	1	0	
14	1	2	Royal Challengers Bangalore	Kolkata Knight Riders	11	3	Rohit Sharma	David Warner	Jasprit Bumrah	0	0	0	0	0	
27	1	2	Royal Challengers Bangalore	Royal Challengers Bangalore	4	4	Rohit Sharma	Faf du Plessis	Rashid Khan	1	1	0	0	0	
46	1	2	Mumbai Indians	Kolkata Knight Riders	10	2	Virat Kohli	David Warner	Rashid Khan	0	0	0	1	0	

**29.**

```
match_1.shape
```

---

```
(23, 16)
```

**30.**

```
srh=match_1[match_1['inning']==1]
srh['batsman_runs'].value_counts()
```

---

```
batsman_runs
3      3
0      2
2      1
4      1
5      1
Name: count, dtype: int64
```

**31.**

```
srh['dismissal_kind'].value_counts()
```

---

```
dismissal_kind
Caught      3
Bowled      2
LBW         1
Name: count, dtype: int64
```

**32.**

```
rcb=match_1[match_1['inning']==2]
rcb['batsman_runs'].value_counts()
```

```
batsman_runs
2      3
5      3
6      2
4      2
1      2
0      2
3      1
Name: count, dtype: int64
```

**33.**

```
rcb['dismissal_kind'].value_counts()
```

```
dismissal_kind
LBW      4
Run Out   3
Bowled    2
Caught    2
Name: count, dtype: int64
```

---