

ipl dataset analysis by "swapan chetri"

In [1]:

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

In [2]:

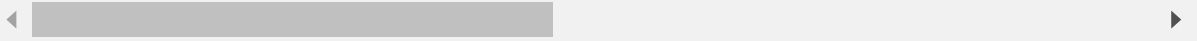
```
ipl=pd.read_csv("IPL Matches 2008-2020.csv")
```

In [3]:

```
ipl.head()
```

Out[3]:

| | id | city | date | player_of_match | venue | neutral_venue | team1 | t |
|---|--------|------------|------------|-----------------|--|---------------|-----------------------------|-------------|
| 0 | 335982 | Bangalore | 2008-04-18 | BB McCullum | Chinnaswamy Stadium | 0 | Royal Challengers Bangalore | Ki |
| 1 | 335983 | Chandigarh | 2008-04-19 | MEK Hussey | Punjab Cricket Association Stadium, Mohali | 0 | Kings XI Punjab | Ch |
| 2 | 335984 | Delhi | 2008-04-19 | MF Maharoo | Feroz Shah Kotla | 0 | Delhi Daredevils | Raja R |
| 3 | 335985 | Mumbai | 2008-04-20 | MV Boucher | Wankhede Stadium | 0 | Mumbai Indians | Challe Banç |
| 4 | 335986 | Kolkata | 2008-04-20 | DJ Hussey | Eden Gardens | 0 | Kolkata Knight Riders | De Cha |



In [4]:

```
ipl.tail()
```

Out[4]:

| | id | city | date | player_of_match | venue | neutral_venue | team1 | team2 |
|-----|---------|-----------|------------|-----------------|-------------------------------------|---------------|-----------------------------|---------------------|
| 811 | 1216547 | Dubai | 2020-09-28 | AB de Villiers | Dubai International Cricket Stadium | 0 | Royal Challengers Bangalore | Mumbai Indians |
| 812 | 1237177 | Dubai | 2020-11-05 | JJ Bumrah | Dubai International Cricket Stadium | 0 | Mumbai Indians | Delhi Capitals |
| 813 | 1237178 | Abu Dhabi | 2020-11-06 | KS Williamson | Sheikh Zayed Stadium | 0 | Royal Challengers Bangalore | Sunrisers Hyderabad |
| 814 | 1237180 | Abu Dhabi | 2020-11-08 | MP Stoinis | Sheikh Zayed Stadium | 0 | Delhi Capitals | Sunrisers Hyderabad |
| 815 | 1237181 | Dubai | 2020-11-10 | TA Boult | Dubai International Cricket Stadium | 0 | Delhi Capitals | Mumbai Indians |

In [5]:

```
ipl.shape
```

Out[5]:

(816, 17)

In [6]:

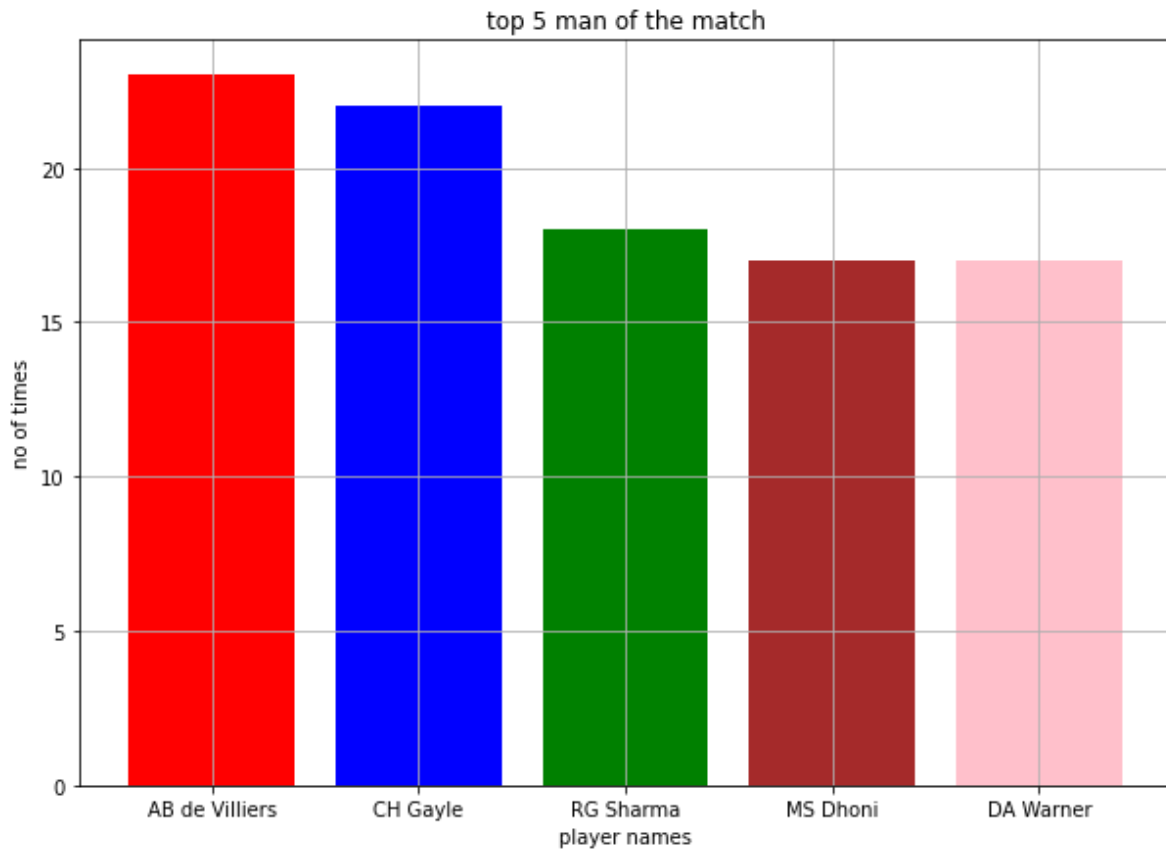
```
# top 5 man of the match  
ipl["player_of_match"].value_counts()[:5]
```

Out[6]:

```
AB de Villiers    23  
CH Gayle          22  
RG Sharma         18  
MS Dhoni          17  
DA Warner         17  
Name: player_of_match, dtype: int64
```

In [7]:

```
# bar plot for top 5 man of the match
plt.figure(figsize=(10,7))
plt.bar(ipl["player_of_match"].value_counts()[:5].keys(),ipl["player_of_match"].value_count
plt.title("top 5 man of the match")
plt.xlabel("player names")
plt.ylabel("no of times")
plt.grid(True)
plt.show()
```



In [8]:

```
# getting the frequency of result column  
ipl["winner"].value_counts()
```

Out[8]:

| | |
|-----------------------------|-----|
| Mumbai Indians | 120 |
| Chennai Super Kings | 106 |
| Kolkata Knight Riders | 99 |
| Royal Challengers Bangalore | 91 |
| Kings XI Punjab | 88 |
| Rajasthan Royals | 81 |
| Delhi Daredevils | 67 |
| Sunrisers Hyderabad | 66 |
| Deccan Chargers | 29 |
| Delhi Capitals | 19 |
| Gujarat Lions | 13 |
| Pune Warriors | 12 |
| Rising Pune Supergiant | 10 |
| Kochi Tuskers Kerala | 6 |
| Rising Pune Supergiants | 5 |

Name: winner, dtype: int64

In [9]:

```
# getting the frequency of result top 5 column  
ipl["winner"].value_counts()[:5]
```

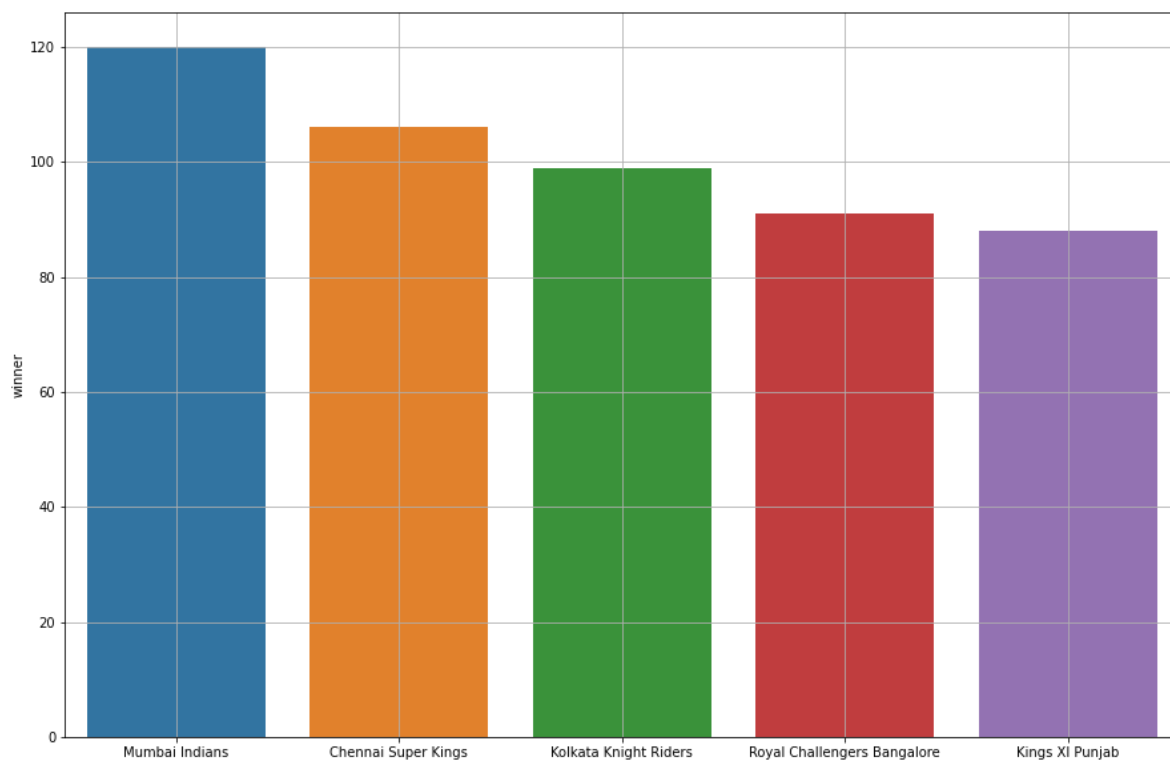
Out[9]:

| | |
|-----------------------------|-----|
| Mumbai Indians | 120 |
| Chennai Super Kings | 106 |
| Kolkata Knight Riders | 99 |
| Royal Challengers Bangalore | 91 |
| Kings XI Punjab | 88 |

Name: winner, dtype: int64

In [10]:

```
# bar plot for max winners  
plt.figure(figsize=(15,10))  
sns.barplot(x=ipl["winner"].value_counts()[:5].keys(),y=ipl["winner"].value_counts()[:5],da  
plt.grid(True)  
plt.show()
```



In [11]:

```
ipl.head()
```

Out[11]:

| | id | city | date | player_of_match | venue | neutral_venue | team1 | team2 |
|---|--------|------------|------------|-----------------|--|---------------|-----------------------------|-----------------------|
| 0 | 335982 | Bangalore | 2008-04-18 | BB McCullum | M Chinnaswamy Stadium | 0 | Royal Challengers Bangalore | Kolkata Knight Riders |
| 1 | 335983 | Chandigarh | 2008-04-19 | MEK Hussey | Punjab Cricket Association Stadium, Mohali | 0 | Kings XI Punjab | Chennai Super Kings |
| 2 | 335984 | Delhi | 2008-04-19 | MF Maharoo | Feroz Shah Kotla | 0 | Delhi Daredevils | Rajasthan Royals |
| 3 | 335985 | Mumbai | 2008-04-20 | MV Boucher | Wankhede Stadium | 0 | Mumbai Indians | Chennai Super Kings |
| 4 | 335986 | Kolkata | 2008-04-20 | DJ Hussey | Eden Gardens | 0 | Kolkata Knight Riders | Delhi Daredevils |

In [12]:

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

Out[12]:

```
wickets    435
runs       364
tie         13
Name: result, dtype: int64
```

In [13]:

```
# toss winner
ipl["toss_winner"].value_counts()
```

Out[13]:

| | |
|-----------------------------|-----|
| Mumbai Indians | 106 |
| Kolkata Knight Riders | 98 |
| Chennai Super Kings | 97 |
| Rajasthan Royals | 87 |
| Royal Challengers Bangalore | 87 |
| Kings XI Punjab | 85 |
| Delhi Daredevils | 80 |
| Sunrisers Hyderabad | 57 |
| Deccan Chargers | 43 |
| Pune Warriors | 20 |
| Delhi Capitals | 20 |
| Gujarat Lions | 15 |
| Kochi Tuskers Kerala | 8 |
| Rising Pune Supergiants | 7 |
| Rising Pune Supergiant | 6 |

Name: toss_winner, dtype: int64

In [14]:

```
#toss winner decision
ipl["toss_decision"].value_counts()
```

Out[14]:

| | |
|-------|-----|
| field | 496 |
| bat | 320 |

Name: toss_decision, dtype: int64

In []:

In [15]:

```
# extracting value for batting first
bat_first=ipl[ipl["toss_decision"]=="bat"]
bat_first.head()
```

Out[15]:

| | id | city | date | player_of_match | venue | neutral_venue | team1 | team2 |
|---|--------|------------|------------|-----------------|--|---------------|-----------------------|-----------------------------|
| 1 | 335983 | Chandigarh | 2008-04-19 | MEK Hussey | Punjab Cricket Association Stadium, Mohali | 0 | Kings XI Punjab | Chennai Super Kings |
| 2 | 335984 | Delhi | 2008-04-19 | MF Maharoo | Feroz Shah Kotla | 0 | Delhi Daredevils | Rajasthan Royals |
| 3 | 335985 | Mumbai | 2008-04-20 | MV Boucher | Wankhede Stadium | 0 | Mumbai Indians | Royal Challengers Bangalore |
| 4 | 335986 | Kolkata | 2008-04-20 | DJ Hussey | Eden Gardens | 0 | Kolkata Knight Riders | Deccan Chargers |
| 5 | 335987 | Jaipur | 2008-04-21 | SR Watson | Sawai Mansingh Stadium | 0 | Rajasthan Royals | Kings XI Punjab |

In [16]:

```
# team which select to choose bat first maximum times
bat_first["toss_winner"].value_counts()
```

Out[16]:

| | |
|-----------------------------|----|
| Chennai Super Kings | 51 |
| Mumbai Indians | 48 |
| Kolkata Knight Riders | 34 |
| Rajasthan Royals | 34 |
| Delhi Daredevils | 29 |
| Kings XI Punjab | 27 |
| Deccan Chargers | 24 |
| Royal Challengers Bangalore | 24 |
| Sunrisers Hyderabad | 24 |
| Pune Warriors | 11 |
| Delhi Capitals | 7 |
| Rising Pune Supergiants | 3 |
| Kochi Tuskers Kerala | 3 |
| Gujarat Lions | 1 |

Name: toss_winner, dtype: int64

In [17]:

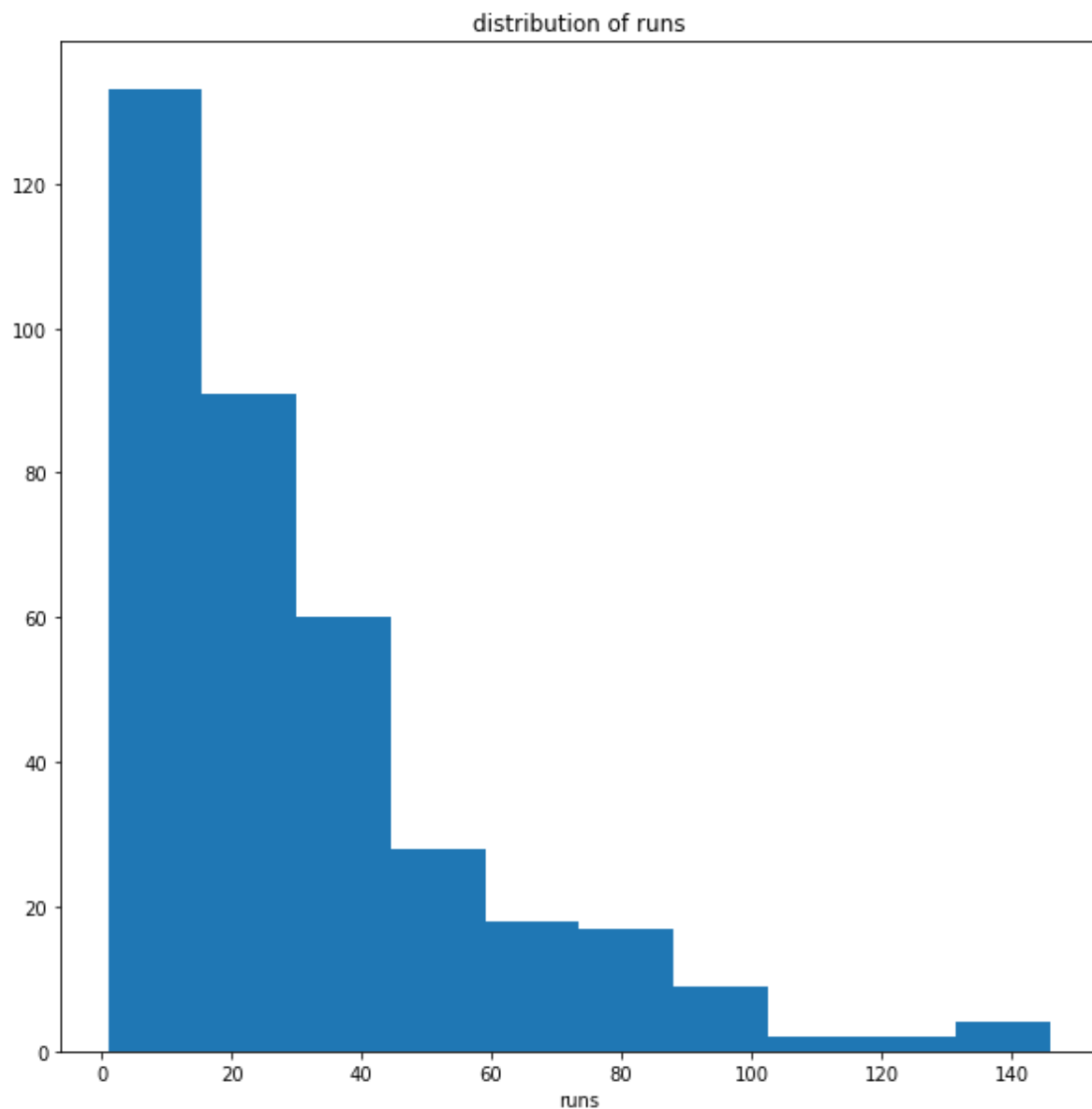
```
# team won by bat first
bat_first_winner=ipl[ipl["result"]=="runs"]
bat_first_winner.head()
```

Out[17]:

| | id | city | date | player_of_match | venue | neutral_venue | team1 | tea |
|----|--------|------------|------------|-----------------|--|---------------|-----------------------------|--------------------|
| 0 | 335982 | Bangalore | 2008-04-18 | BB McCullum | M Chinnaswamy Stadium | 0 | Royal Challengers Bangalore | Kolk Kni Rid |
| 1 | 335983 | Chandigarh | 2008-04-19 | MEK Hussey | Punjab Cricket Association Stadium, Mohali | 0 | Kings XI Punjab | Chen Su Kii |
| 7 | 335989 | Chennai | 2008-04-23 | ML Hayden | MA Chidambaram Stadium, Chepauk | 0 | Chennai Super Kings | Mum Indi |
| 9 | 335991 | Chandigarh | 2008-04-25 | KC Sangakkara | Punjab Cricket Association Stadium, Mohali | 0 | Kings XI Punjab | Mum Indi |
| 14 | 335996 | Bangalore | 2008-04-28 | MS Dhoni | M Chinnaswamy Stadium | 0 | Royal Challengers Bangalore | Chen Su Kii |

In [18]:

```
#making a histogram  
plt.figure(figsize=(10,10))  
plt.hist(bat_first_winner["result_margin"])  
plt.title("distribution of runs")  
plt.xlabel("runs")  
plt.show()
```



In [19]:

```
#finding out the number of wins w.r.t each team after bat first  
bat_first_winner["winner"].value_counts()
```

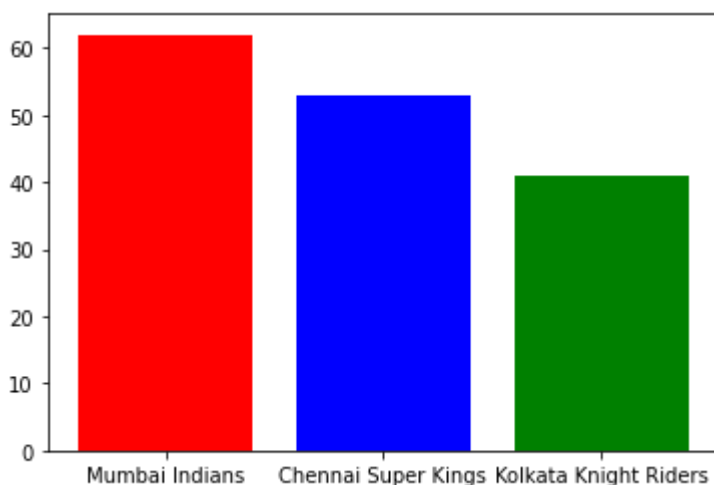
Out[19]:

| | |
|-----------------------------|----|
| Mumbai Indians | 62 |
| Chennai Super Kings | 53 |
| Kolkata Knight Riders | 41 |
| Kings XI Punjab | 40 |
| Royal Challengers Bangalore | 38 |
| Sunrisers Hyderabad | 34 |
| Rajasthan Royals | 28 |
| Delhi Daredevils | 25 |
| Deccan Chargers | 18 |
| Delhi Capitals | 9 |
| Pune Warriors | 6 |
| Rising Pune Supergiant | 5 |
| Rising Pune Supergiants | 2 |
| Kochi Tuskers Kerala | 2 |
| Gujarat Lions | 1 |

Name: winner, dtype: int64

In [20]:

```
#making a plot for top 3 teams who wins with bat first  
plt.bar(list(bat_first_winner["winner"].value_counts()[:3].keys()),list(bat_first_winner["w  
plt.show()
```



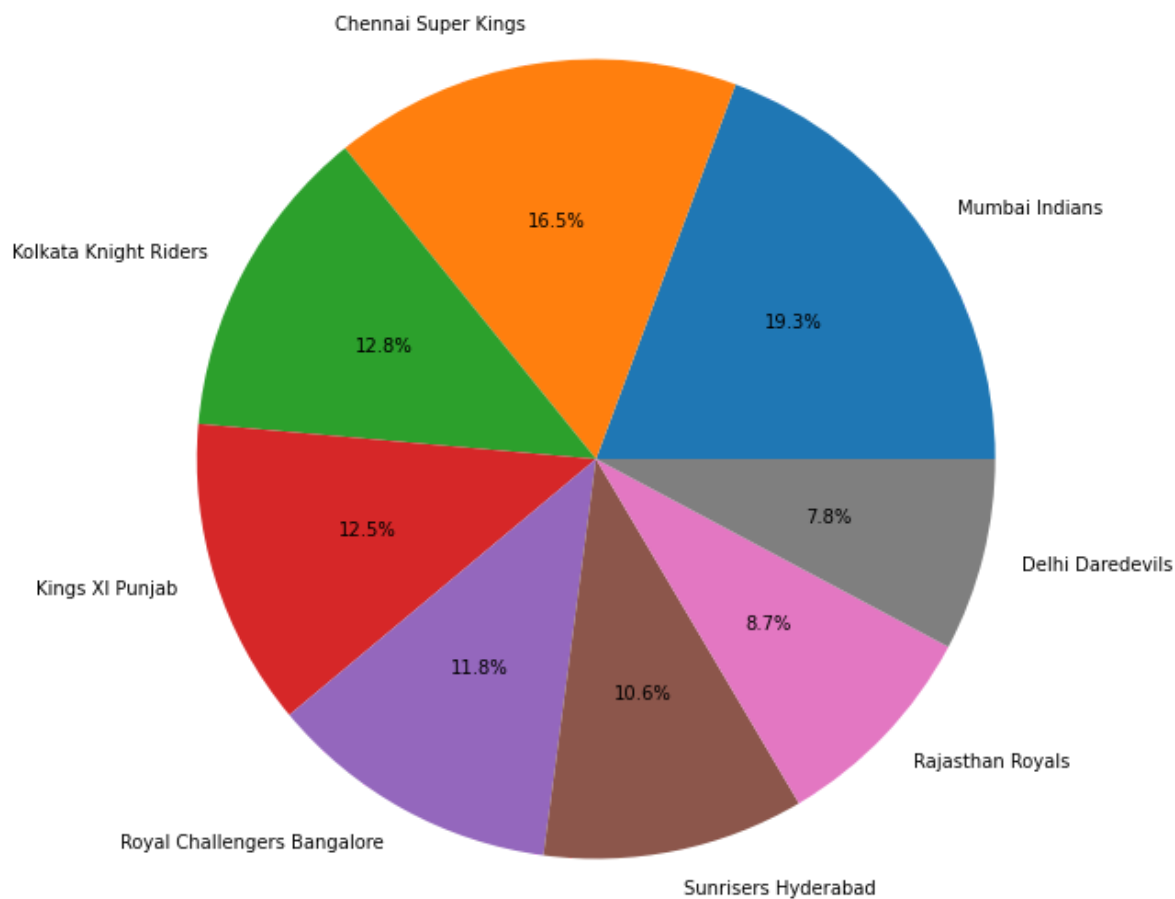
In [21]:

```
#making a pie chart for top 8 teams
```

```
plt.figure(figsize=(10,10))
```

```
plt.pie(list(bat_first_winner["winner"].value_counts()[:8]), labels=list(bat_first_winner["
```

```
plt.show()
```



In [22]:

```
#team won after batting second
batting_second=ipl[ipl["result"]=="wickets"]
batting_second.head()
```

Out[22]:

| | id | city | date | player_of_match | venue | neutral_venue | team1 | team2 |
|---|--------|-----------|------------|-----------------|---|---------------|-----------------------|-----------------------------|
| 2 | 335984 | Delhi | 2008-04-19 | MF Maharoof | Feroz Shah Kotla | 0 | Delhi Daredevils | Rajasthan Royals |
| 3 | 335985 | Mumbai | 2008-04-20 | MV Boucher | Wankhede Stadium | 0 | Mumbai Indians | Royal Challengers Bangalore |
| 4 | 335986 | Kolkata | 2008-04-20 | DJ Hussey | Eden Gardens | 0 | Kolkata Knight Riders | Deccan Chargers |
| 5 | 335987 | Jaipur | 2008-04-21 | SR Watson | Sawai Mansingh Stadium | 0 | Rajasthan Royals | Kings XI Punjab |
| 6 | 335988 | Hyderabad | 2008-04-22 | V Sehwag | Rajiv Gandhi International Stadium, Uppal | 0 | Deccan Chargers | Delhi Daredevils |

In [23]:

```
batting_second.shape
```

Out[23]:

(435, 17)

In [24]:

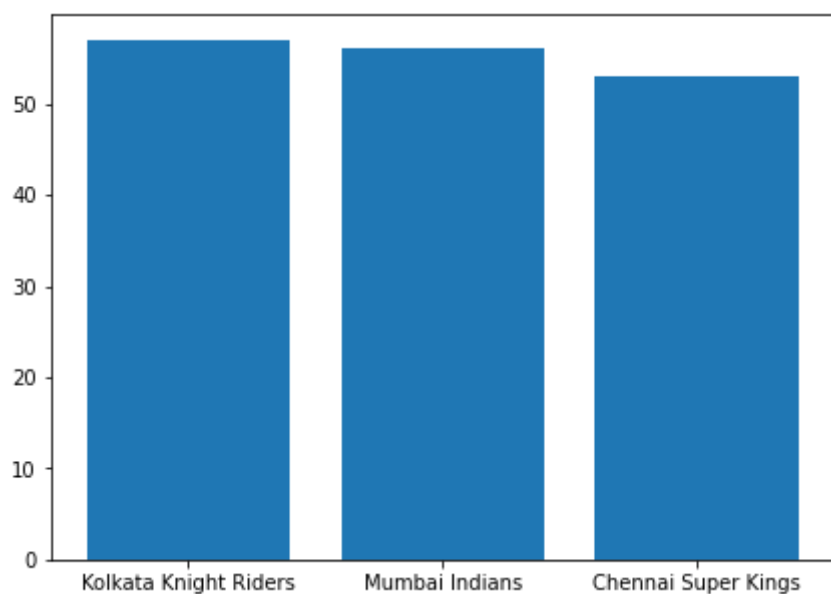
```
batting_second["winner"].value_counts()[ :10]
```

Out[24]:

```
Kolkata Knight Riders      57
Mumbai Indians              56
Chennai Super Kings        53
Royal Challengers Bangalore 51
Rajasthan Royals           51
Kings XI Punjab            45
Delhi Daredevils           42
Sunrisers Hyderabad        31
Gujarat Lions              12
Deccan Chargers            11
Name: winner, dtype: int64
```

In [25]:

```
# bar plot for top 3 team who did bat second  
plt.figure(figsize=(7,5))  
plt.bar(list(batting_second["winner"].value_counts()[:3].keys()),batting_second["winner"].v  
plt.show()
```

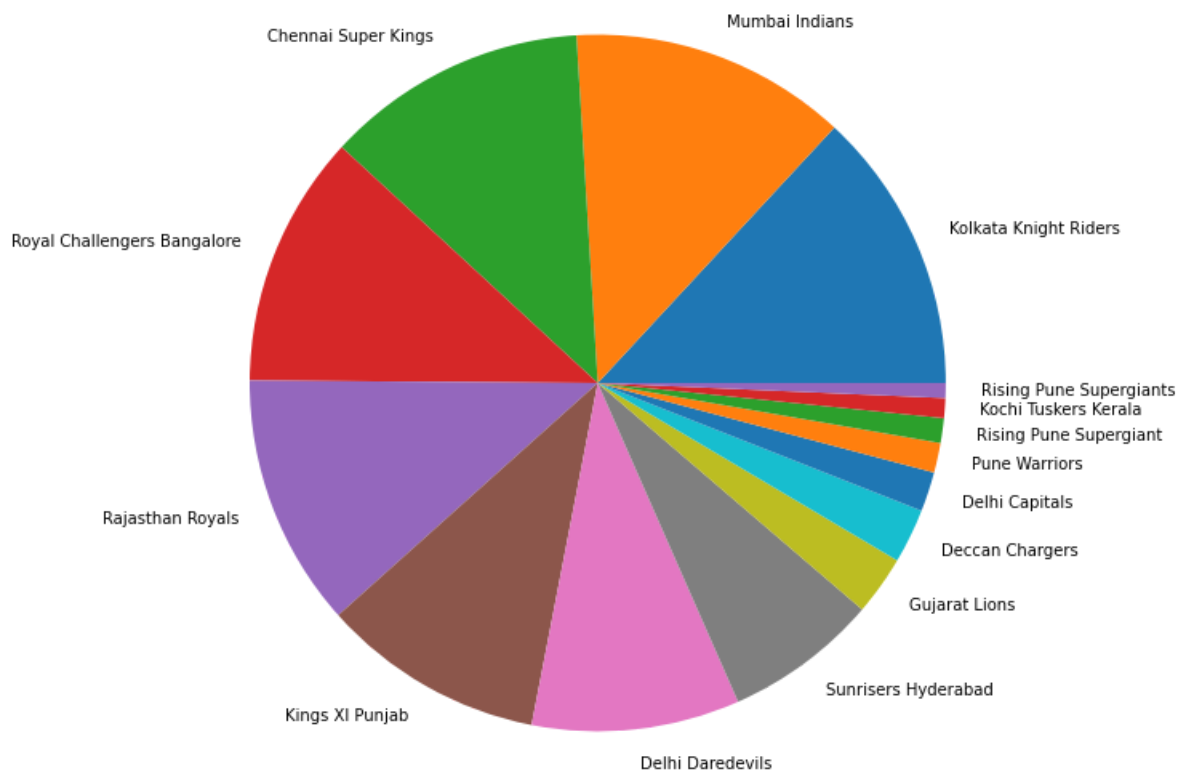


In [26]:

```
# pie plot for most winners who choosed bowling first
plt.figure(figsize=(10,10))
plt.pie(batting_second["winner"].value_counts(),labels=list(batting_second["winner"].value_
plt.plot()
```

Out[26]:

[]



In [27]:

```
# no of matches played in each city  
ipl["city"].value_counts()
```

Out[27]:

| | |
|----------------|-----|
| Mumbai | 101 |
| Kolkata | 77 |
| Delhi | 74 |
| Bangalore | 65 |
| Hyderabad | 64 |
| Chennai | 57 |
| Chandigarh | 56 |
| Jaipur | 47 |
| Pune | 38 |
| Abu Dhabi | 29 |
| Dubai | 26 |
| Durban | 15 |
| Bengaluru | 15 |
| Visakhapatnam | 13 |
| Sharjah | 12 |
| Ahmedabad | 12 |
| Centurion | 12 |
| Rajkot | 10 |
| Indore | 9 |
| Dharamsala | 9 |
| Johannesburg | 8 |
| Ranchi | 7 |
| Cape Town | 7 |
| Cuttack | 7 |
| Port Elizabeth | 7 |
| Raipur | 6 |
| Kochi | 5 |
| Kanpur | 4 |
| East London | 3 |
| Nagpur | 3 |
| Kimberley | 3 |
| Bloemfontein | 2 |

Name: city, dtype: int64

In [28]:

```
# no of times a team won after winning the toss  
np.sum(ipl["toss_winner"]==ipl["winner"])
```

Out[28]:

418

In [29]:

```
#importing a new dataset
deliveries=pd.read_csv("IPL Ball-by-Ball 2008-2020.csv")
deliveries.head()
```

Out[29]:

| | id | inning | over | ball | batsman | non_striker | bowler | batsman_runs | extra_runs | total_r |
|---|--------|--------|------|------|-------------|-------------|-----------|--------------|------------|---------|
| 0 | 335982 | 1 | 6 | 5 | RT Ponting | BB McCullum | AA Noffke | 1 | 0 | |
| 1 | 335982 | 1 | 6 | 6 | BB McCullum | RT Ponting | AA Noffke | 1 | 0 | |
| 2 | 335982 | 1 | 7 | 1 | BB McCullum | RT Ponting | Z Khan | 0 | 0 | |
| 3 | 335982 | 1 | 7 | 2 | BB McCullum | RT Ponting | Z Khan | 1 | 0 | |
| 4 | 335982 | 1 | 7 | 3 | RT Ponting | BB McCullum | Z Khan | 1 | 0 | |

In [30]:

```
# total number of matches
t=deliveries["id"].unique()
len(t)
```

Out[30]:

816

In [31]:

```
match_1=deliveries[deliveries["id"]==335982]
```

In [32]:

```
match_1.head()
```

Out[32]:

| player | bowler | batsman_runs | extra_runs | total_runs | non_boundary | is_wicket | dismissal_kind | player |
|-------------|--------------|--------------|------------|------------|--------------|-----------|----------------|--------|
| BB Illum | AA Noffke | 1 | 0 | 1 | 0 | 0 | NaN | |
| ating | AA Noffke | 1 | 0 | 1 | 0 | 0 | NaN | |
| ating | Z Khan | 0 | 0 | 0 | 0 | 0 | NaN | |
| ating | Z Khan | 1 | 0 | 1 | 0 | 0 | NaN | |
| BB Illum | Z Khan | 1 | 0 | 1 | 0 | 0 | NaN | |

In [33]:

```
# we have 240 total balls in t20 but here we have 225 that means 15 have not played before  
match_1.shape
```

Out[33]:

(225, 18)

In [35]:

```
# analysing the team kolkata knight riders in the first match  
kkr=match_1[match_1["inning"]==1]  
kkr["batsman_runs"].value_counts()
```

Out[35]:

```
0    45  
1    39  
4    15  
6    14  
2    11  
Name: batsman_runs, dtype: int64
```

In [42]:

```
kkrr["dismissal_kind"].value_counts()
```

Out[42]:

```
caught      3
Name: dismissal_kind, dtype: int64
```

In [40]:

```
# analysing the team Royal Challengers Bangalore in the first match
rcb=match_1[match_1["inning"]==2]
rcb["batsman_runs"].value_counts()
```

Out[40]:

```
0      65
1      27
6       3
4       3
2       3
Name: batsman_runs, dtype: int64
```

In [43]:

```
# as we can see that rcb had been all out at the first match
rcb["dismissal_kind"].value_counts()
```

Out[43]:

```
caught      6
bowled      3
run out     1
Name: dismissal_kind, dtype: int64
```

In []: