In [4]:
```
#loading the required libraries
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
```

In [21]:
```
#loading dataset
ipl=pd.read_csv('data.csv')
```

In [22]:
```
#having glance of first five matches of the dataset
ipl.head()
```

Out[22]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets
0	1	2008	Bangalore	2008-04-18	Kolkata Knight Riders	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Kolkata Knight Riders	140	0
1	2	2008	Chandigarh	2008-04-19	Chennai Super Kings	Kings XI Punjab	Chennai Super Kings	bat	normal	0	Chennai Super Kings	33	0
2	3	2008	Delhi	2008-04-19	Rajasthan Royals	Delhi Daredevils	Rajasthan Royals	bat	normal	0	Delhi Daredevils	0	9
3	4	2008	Mumbai	2008-04-20	Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	normal	0	Royal Challengers Bangalore	0	5
4	5	2008	Kolkata	2008-04-20	Deccan Chargers	Kolkata Knight Riders	Deccan Chargers	bat	normal	0	Kolkata Knight Riders	0	5

In [23]:
```
#looking at the number of rows and columns of the dataset
ipl.shape
```

Out[23]: (577, 18)

In [24]:
```
ipl['season'].value_counts()
```

Out[24]:
```
2013    76
2012    74
2011    73
2010    60
2014    60
2016    60
2015    59
2008    58
2009    57
Name: season, dtype: int64
```

In [6]:
```
#checking frequency of most player of the match
ipl['player_of_match'].value_counts()
```

Out[6]:
```
CH Gayle          17
YK Pathan         16
AB de Villiers    15
DA Warner         14
RG Sharma         13
                  ..
MN Samuels         1
S Badrinath        1
R Bhatia           1
RE Levi            1
BCJ Cutting        1
Name: player_of_match, Length: 187, dtype: int64
```

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

Out[7]:
```
CH Gayle          17
YK Pathan         16
AB de Villiers    15
DA Warner         14
RG Sharma         13
SK Raina          13
AM Rahane         12
MEK Hussey        12
G Gambhir         12
MS Dhoni          12
Name: player_of_match, dtype: int64
```

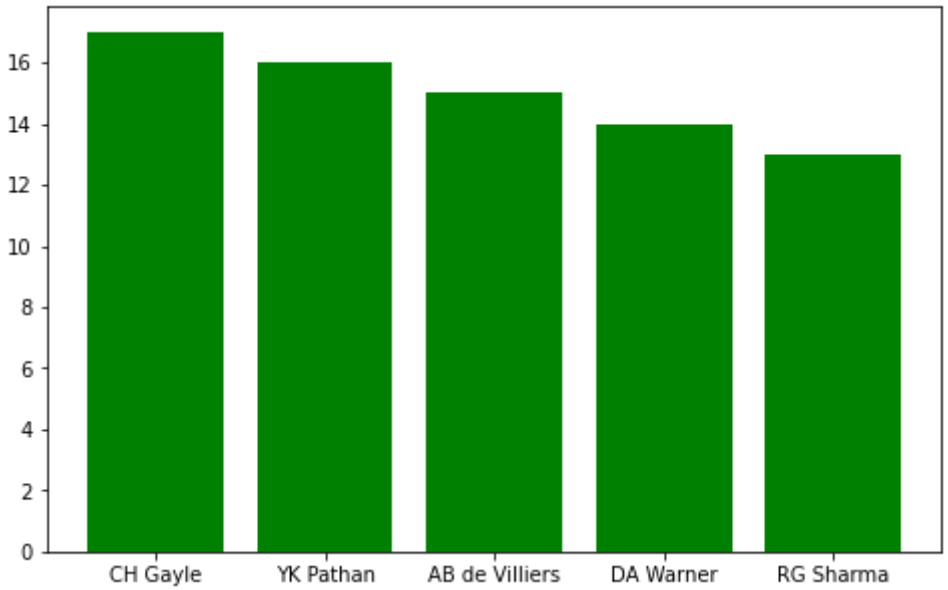
In [8]:
```
#getting the top 5 players with most player of the match
ipl['player_of_match'].value_counts()[0:5]
```

Out[8]:
```
CH Gayle          17
YK Pathan         16
AB de Villiers    15
DA Warner         14
RG Sharma         13
Name: player_of_match, dtype: int64
```

In [10]:
```
list(ipl['player_of_match'].value_counts()[0:5].keys())
```

Out[10]:
```
['CH Gayle', 'YK Pathan', 'AB de Villiers', 'DA Warner', 'RG Sharma']
```

In [15]:
```
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="
plt.show()
```



In [16]:
```
#frequency of result column
ipl['result'].value_counts()
```

Out[16]:
```
normal       568
tie            6
no result      3
Name: result, dtype: int64
```

In [17]:
```
#frequency of toss winner column
ipl['toss_winner'].value_counts()
```

Out[17]:
```
Mumbai Indians                74
Kolkata Knight Riders         69
Chennai Super Kings           66
Kings XI Punjab               64
Delhi Daredevils              64
Rajasthan Royals              63
Royal Challengers Bangalore   61
Deccan Chargers               43
Sunrisers Hyderabad           30
Pune Warriors                 20
Kochi Tuskers Kerala           8
Gujarat Lions                  8
Rising Pune Supergiants        7
Name: toss_winner, dtype: int64
```

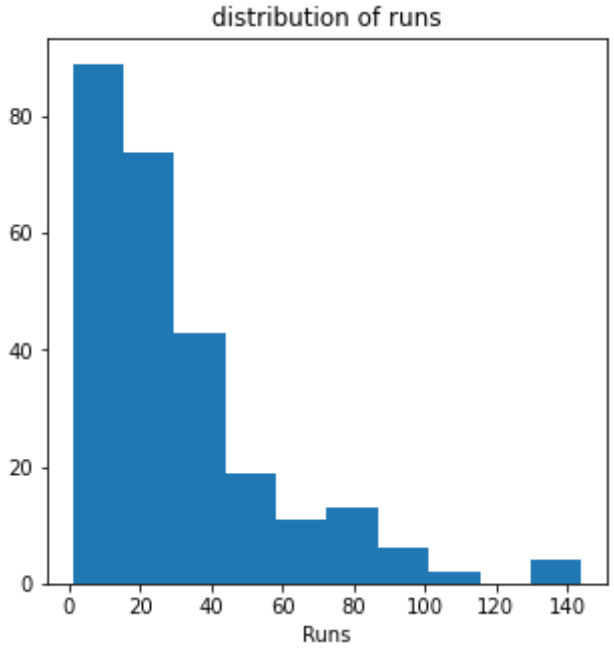
In [18]:
```
#extracting a data which team won batting first
batting_first=ipl[ipl['win_by_runs']!=0]
```

In [19]:
```
#Looking at the head
batting_first.head()
```

Out[19]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	pla
0	1	2008	Bangalore	2008-04-18	Kolkata Knight Riders	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Kolkata Knight Riders	140	0	
1	2	2008	Chandigarh	2008-04-19	Chennai Super Kings	Kings XI Punjab	Chennai Super Kings	bat	normal	0	Chennai Super Kings	33	0	
7	8	2008	Chennai	2008-04-23	Chennai Super Kings	Mumbai Indians	Mumbai Indians	field	normal	0	Chennai Super Kings	6	0	
9	10	2008	Chandigarh	2008-04-25	Kings XI Punjab	Mumbai Indians	Mumbai Indians	field	normal	0	Kings XI Punjab	66	0	K
14	15	2008	Bangalore	2008-04-28	Chennai Super Kings	Royal Challengers Bangalore	Chennai Super Kings	bat	normal	0	Chennai Super Kings	13	0	

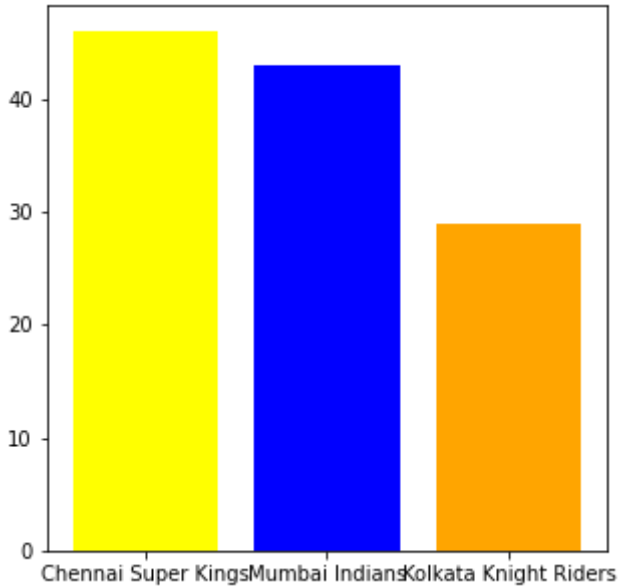
In [26]:
```
#making a histogram
plt.figure(figsize=(5,5))
plt.hist(batting_first['win_by_runs'])
plt.title("distribution of runs")
plt.xlabel("Runs")
plt.show()
```



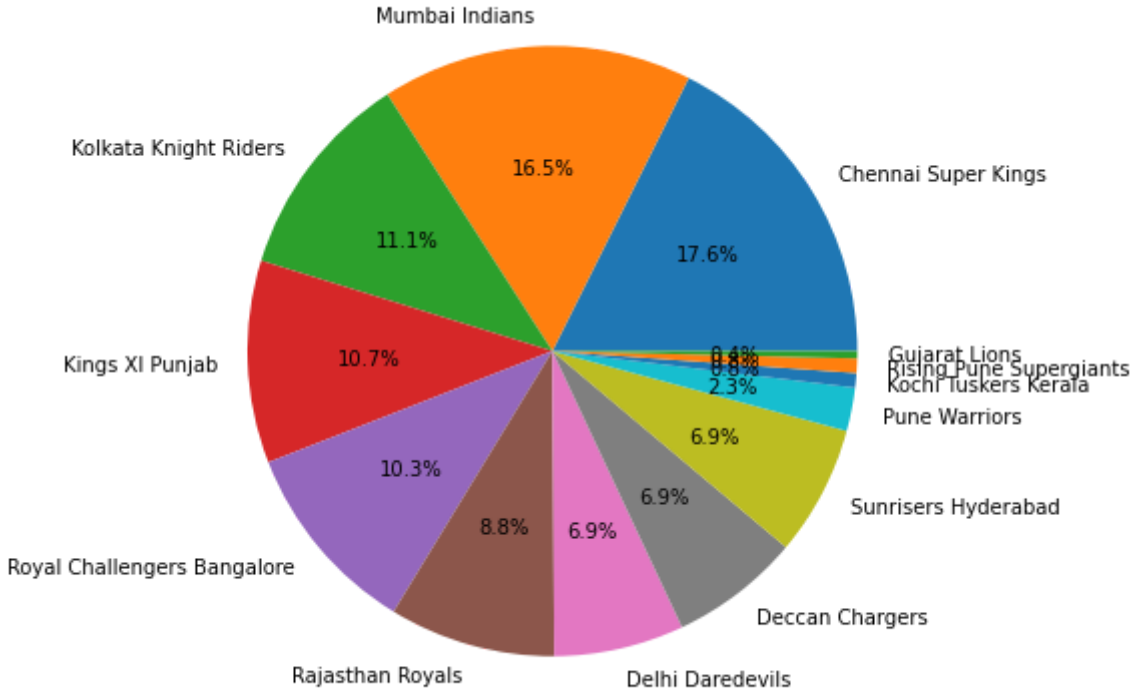
In [28]:
```
#finding number of wins w.r.t each team after batting first
batting_first['winner'].value_counts()
```

Out[28]:
```
Chennai Super Kings            46
Mumbai Indians                 43
Kolkata Knight Riders          29
Kings XI Punjab                28
Royal Challengers Bangalore    27
Rajasthan Royals               23
Delhi Daredevils               18
Deccan Chargers                18
Sunrisers Hyderabad            18
Pune Warriors                   6
Kochi Tuskers Kerala            2
Rising Pune Supergiants         2
Gujarat Lions                   1
Name: winner, dtype: int64
```

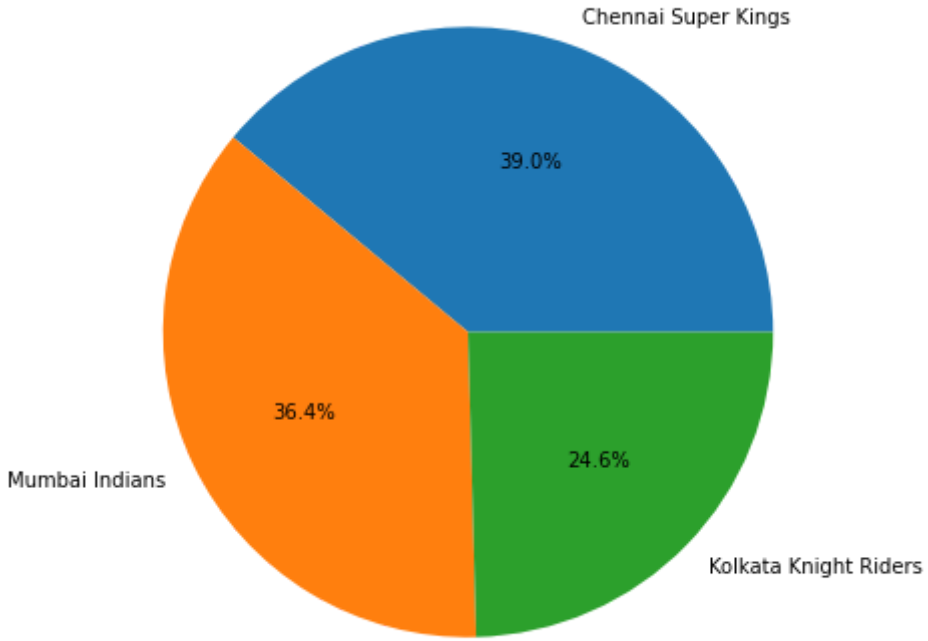
In [35]:
```
#making histogram
plt.figure(figsize=(5,5))
plt.bar(list(batting_first['winner'].value_counts()[0:3].keys()),list(batting_first['winner'].value_counts()[0:3]),color
plt.show()
```



In [40]:
```
plt.figure(figsize=(7,7))
plt.pie(list(batting_first['winner'].value_counts()),labels=list(batting_first['winner'].value_counts().keys()),autopct=
plt.show()
```



In [42]:
```
#pie chart for first 3 teams who are winner after batting first
plt.figure(figsize=(7,7))
plt.pie(list(batting_first['winner'].value_counts()[0:3]),labels=list(batting_first['winner'].value_counts()[0:3].keys()
plt.show()
```



In [43]:
```
#extracting those records where team won batting second
batting_second=ipl[ipl['win_by_wickets']!=0]
```

In [55]:
```
batting_second.head()
```

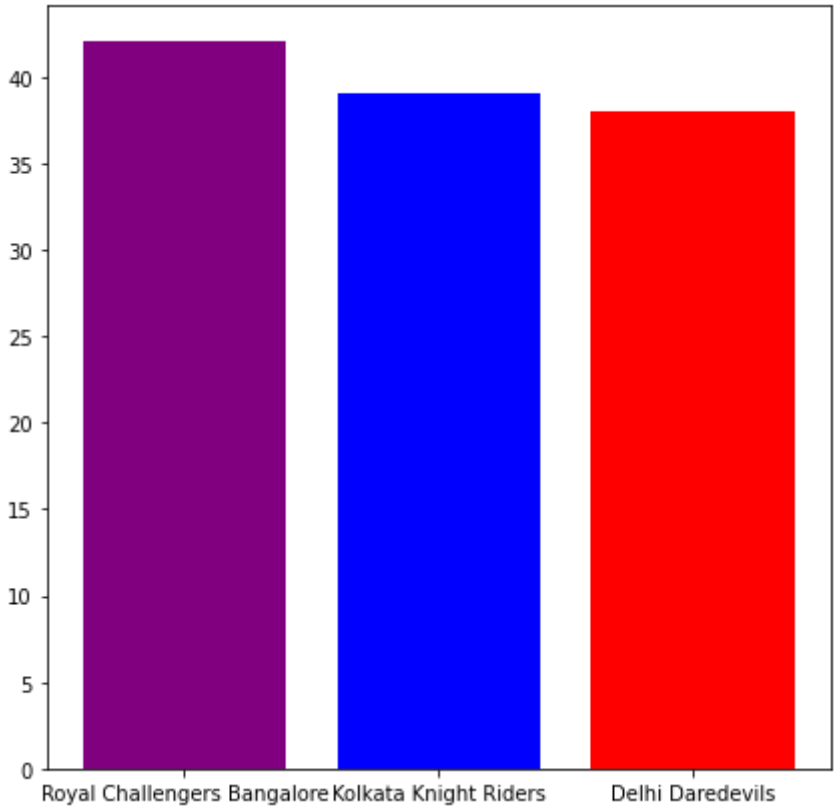
Out[55]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets
2	3	2008	Delhi	2008-04-19	Rajasthan Royals	Delhi Daredevils	Rajasthan Royals	bat	normal	0	Delhi Daredevils	0	9
3	4	2008	Mumbai	2008-04-20	Mumbai Indians	Royal Challengers Bangalore	Mumbai Indians	bat	normal	0	Royal Challengers Bangalore	0	5
4	5	2008	Kolkata	2008-04-20	Deccan Chargers	Kolkata Knight Riders	Deccan Chargers	bat	normal	0	Kolkata Knight Riders	0	5
5	6	2008	Jaipur	2008-04-21	Kings XI Punjab	Rajasthan Royals	Kings XI Punjab	bat	normal	0	Rajasthan Royals	0	6
6	7	2008	Hyderabad	2008-04-22	Deccan Chargers	Delhi Daredevils	Deccan Chargers	bat	normal	0	Delhi Daredevils	0	9

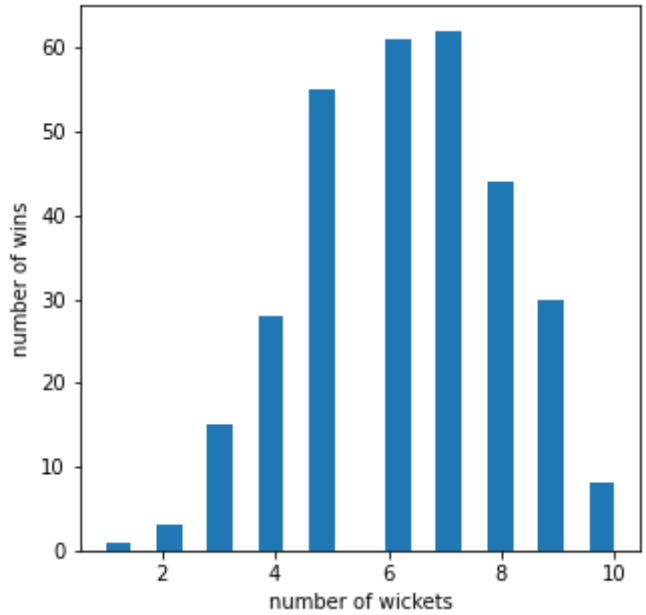
In [56]:
```
#finding out frequency of number of wins w.r.t everytime batting second
batting_second['winner'].value_counts()
```

Out[56]:
```
Royal Challengers Bangalore    42
Kolkata Knight Riders          39
Delhi Daredevils               38
Rajasthan Royals               38
Mumbai Indians                 37
Chennai Super Kings            33
Kings XI Punjab                33
Sunrisers Hyderabad            15
Deccan Chargers                11
Gujarat Lions                   8
Pune Warriors                   6
Kochi Tuskers Kerala            4
Rising Pune Supergiants         3
Name: winner, dtype: int64
```

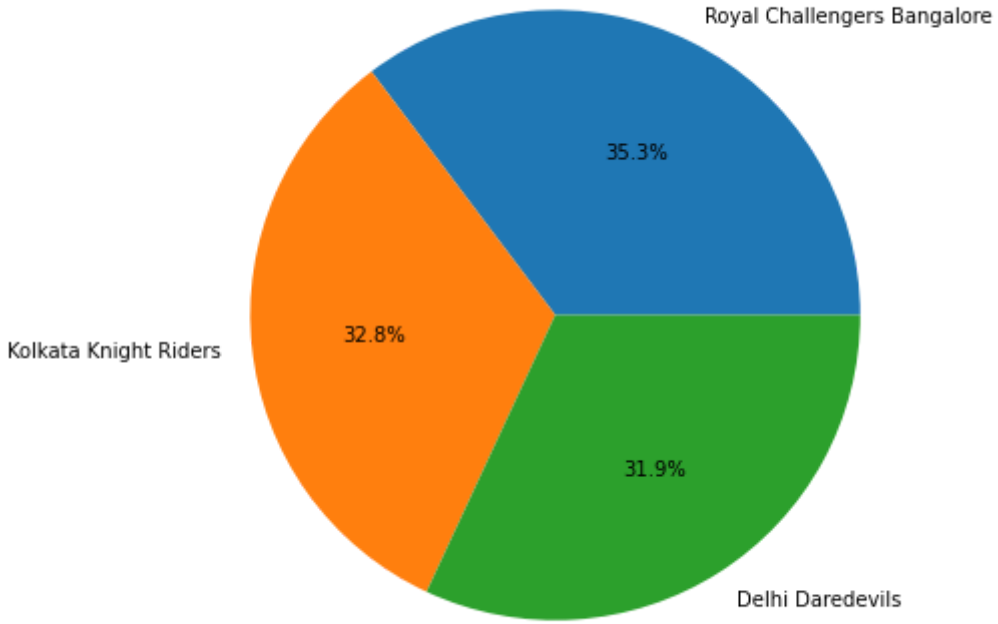
In [61]:
```
#making a bar plot for top 3 teams with most wins after batting second
plt.figure(figsize=(7,7))
plt.bar(list(batting_second['winner'].value_counts()[0:3].keys()),list(batting_second['winner'].value_counts()[0:3]),col
plt.show()
```



In [59]:
```
#making a histogram for frequency of wins w.r.t number of wickets
plt.figure(figsize=(5,5))
plt.hist(batting_second['win_by_wickets'],bins=20)
plt.ylabel("number of wins")
plt.xlabel("number of wickets")
plt.show()
```



In [64]:
```
plt.figure(figsize=(7,7))
plt.pie(list(batting_second['winner'].value_counts()[0:3]),labels=list(batting_second['winner'].value_counts()[0:3].keys
plt.show()
```



In [65]:
```
#number of matches played each season
ipl['season'].value_counts()
```

Out[65]:
```
2013    76
2012    74
2011    73
2010    60
2014    60
2016    60
2015    59
2008    58
2009    57
Name: season, dtype: int64
```

In [67]:
```
#number of matches played in each city
ipl['city'].value_counts()
```

Out[67]:
```
Mumbai           77
Bangalore        58
Kolkata          54
Delhi            53
Chennai          48
Chandigarh       42
Hyderabad        41
Jaipur           33
Pune             25
Durban           15
Centurion        12
Ahmedabad        12
Visakhapatnam    11
Dharamsala        9
Johannesburg      8
Cuttack           7
Port Elizabeth    7
Cape Town         7
Ranchi            7
Abu Dhabi         7
Sharjah           6
Raipur            6
Kochi             5
Rajkot            5
Kimberley         3
Nagpur            3
East London       3
Bloemfontein      2
Indore            2
Kanpur            2
Name: city, dtype: int64
```

In [69]:
```
#finding out how many times team has won after winning toss
np.sum(ipl['toss_winner']==ipl['winner'])
```

Out[69]:
```
291
```

In [70]:
```
291/577
```

Out[70]:
```
0.5043327556325823
```

In [5]:
```
deliveries=pd.read_csv('delivery.csv')
```

In [6]:
```
deliveries.head()
```

Out[6]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	...	bye_runs	legbye_runs	noball_runs
0	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	1	SC Ganguly	BB McCullum	P Kumar	0	...	0	1	0
1	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	2	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0
2	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	3	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0
3	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	4	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0
4	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	5	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0

5 rows × 21 columns

In [7]:
```
deliveries['match_id'].unique()
```

Out[7]:
```
array([1, 2, 3, 4, 5], dtype=int64)
```

In [8]:
```
match_1=deliveries[deliveries['match_id']==1]
```

In [9]:
```
match_1.head()
```

Out[9]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_super_over	...	bye_runs	legbye_runs	noball_runs
0	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	1	SC Ganguly	BB McCullum	P Kumar	0	...	0	1	0
1	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	2	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0
2	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	3	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0
3	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	4	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0
4	1	1	Kolkata Knight Riders	Royal Challengers Bangalore	1	5	BB McCullum	SC Ganguly	P Kumar	0	...	0	0	0

5 rows × 21 columns

In [12]:
```
match_1.shape
```

Out[12]:
```
(225, 21)
```

In [14]:
```
srh=match_1[match_1['inning']==1]
```

In [15]:
```
srh['batsman_runs'].value_counts()
```

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

In [16]:
```
srh['dismissal_kind'].value_counts()
```

Out[16]:
```
caught    3
Name: dismissal_kind, dtype: int64
```

In [17]:
```
rcb=match_1[match_1['inning']==2]
```

In [18]:
```
rcb['batsman_runs'].value_counts()
```

Out[18]:
```
0    65
1    27
2     3
6     3
4     3
Name: batsman_runs, dtype: int64
```

In [19]:
```
rcb['dismissal_kind'].value_counts()
```

Out[19]:
```
caught     6
bowled     3
run out    1
Name: dismissal_kind, dtype: int64
```