IPL Data Analysis of Dataset 2008 - 2019

In [ ]:
```
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
import numpy as np
```

In [6]:
```
#loading the dataset
ipl = pd.read_csv("A:\Data set\IPL DataSet\matches.csv")
```

In [15]:
```
ipl.head()
```

Out[15]:

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

In [16]:
```
ipl.shape
```

Out[16]:
```
(756, 18)
```

In [17]:
```
#ipl['city'].value_counts()
#most player of match awards
ipl['player_of_match'].value_counts()
```

Out[17]:
```
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 [19]:
```
#Top 10 player with most player of the match
ipl['player_of_match'].value_counts()[0:10]
```

Out[19]:
```
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 [20]:
```
y=ipl['player_of_match'].value_counts()[0:5]
y
```

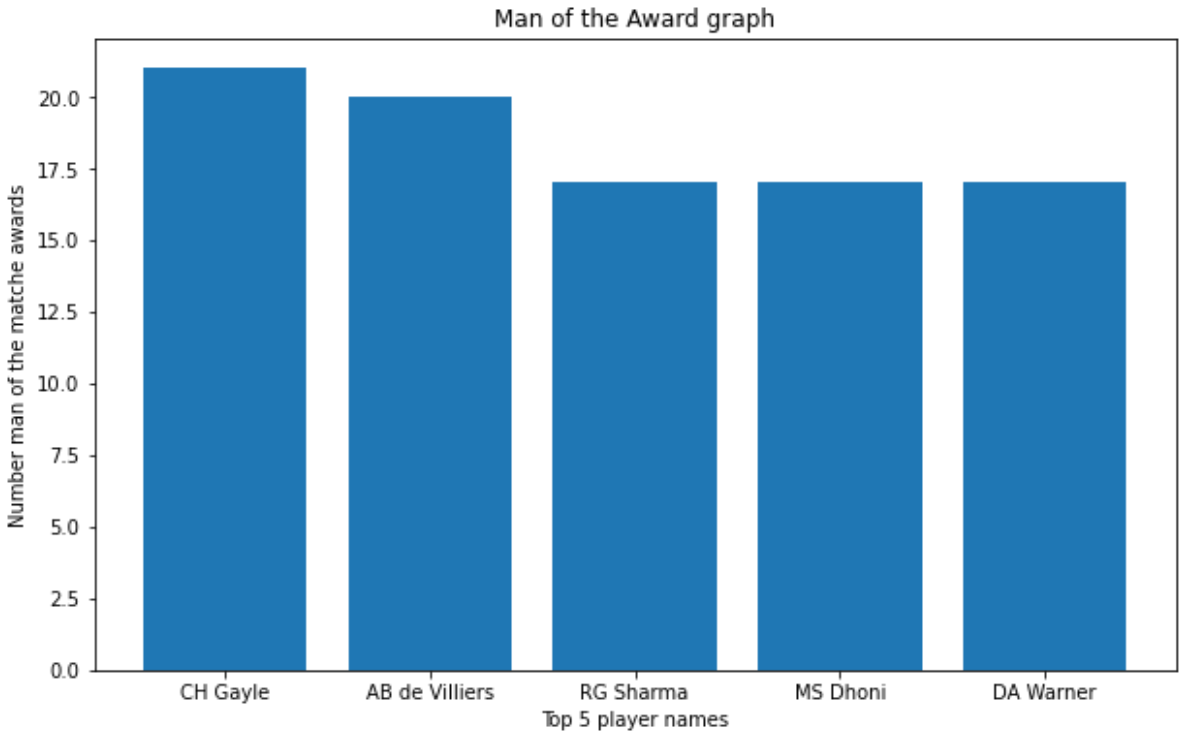
Out[20]:
```
CH Gayle          21
AB de Villiers    20
RG Sharma         17
MS Dhoni          17
DA Warner         17
Name: player_of_match, dtype: int64
```

In [21]:
```
x=list(ipl['player_of_match'].value_counts()[0:5].keys())
x
```

Out[21]:
```
['CH Gayle', 'AB de Villiers', 'RG Sharma', 'MS Dhoni', 'DA Warner']
```

In [22]:
```
#barplot of top 5 player of the match
plt.figure(figsize=(10,6))
plt.bar(x,y)
plt.xlabel("Top 5 player names")
plt.ylabel("Number man of the matche awards")
plt.title('Man of the Award graph')
plt.show
```

Out[22]:
```
<function matplotlib.pyplot.show(close=None, block=None)>
```



In [23]:
```
#Getting the frequency pf result columns
x=ipl['result'].value_counts()
x
```

Out[23]:
```
normal       743
tie            9
no result      4
Name: result, dtype: int64
```

In [24]:
```
#Find out No of Toss Wins w.r.t. Each Team
ipl['toss_winner'].value_counts()
```

Out[24]:
```
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
```

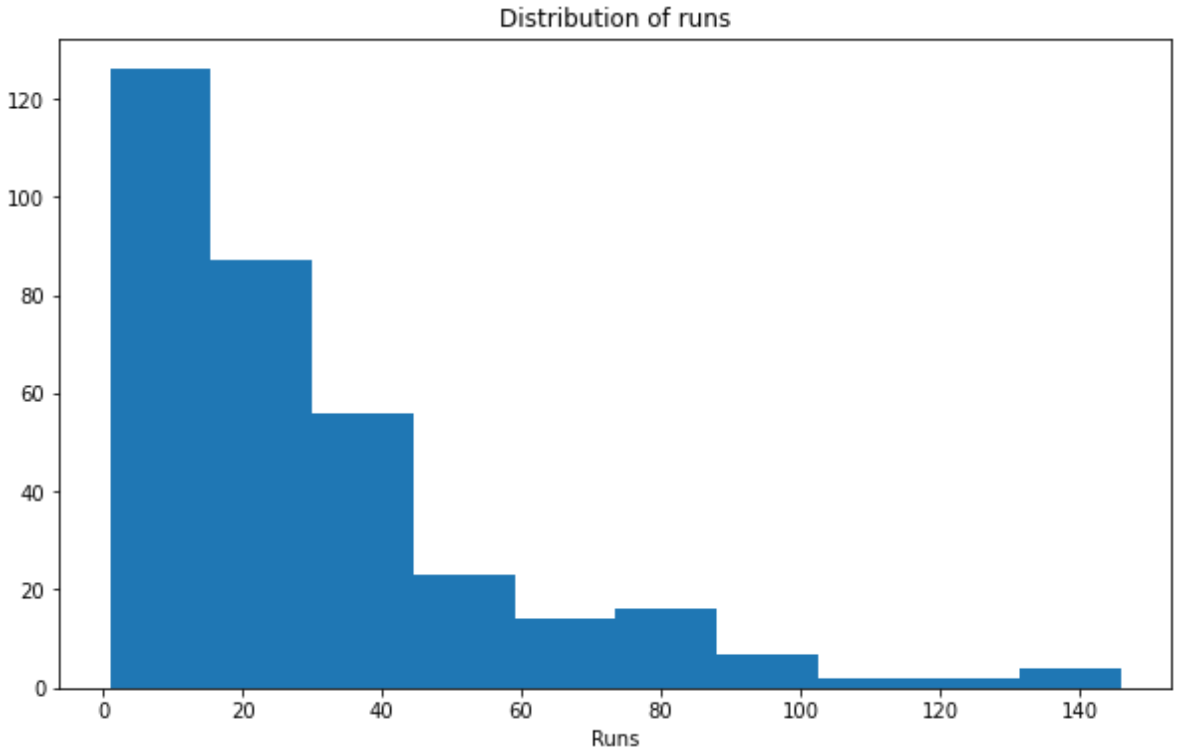
In [60]:
```
#Extracting the records where a team won batting first

batting_first=ipl[ipl['win_by_runs']!=0]
batting_first.head()
```

Out[60]:

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

In [61]:
```
#creating plot of toss decision
plt.figure(figsize=(10,6))
plt.hist(batting_first['win_by_runs'])
plt.title('Distribution of runs')
plt.xlabel('Runs')
plt.show()
```



In [62]:
```
#No of WIns w.r.t. Rach Team After Batting First
batting_first['winner'].value_counts()
```

Out[62]:
```
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 [66]:
```
list(batting_first['winner'].value_counts().keys())
```

Out[66]:
```
['Mumbai Indians',
 'Chennai Super Kings',
 'Kings XI Punjab',
 'Kolkata Knight Riders',
 'Royal Challengers Bangalore',
 'Sunrisers Hyderabad',
 'Rajasthan Royals',
 'Delhi Daredevils',
 'Deccan Chargers',
 'Pune Warriors',
 'Rising Pune Supergiant',
 'Delhi Capitals',
 'Kochi Tuskers Kerala',
 'Rising Pune Supergiants',
 'Gujarat Lions']
```

In [37]:
```
#Top 3 team who wins after batting first
a =batting_first['winner'].value_counts()[0:3]
a
```

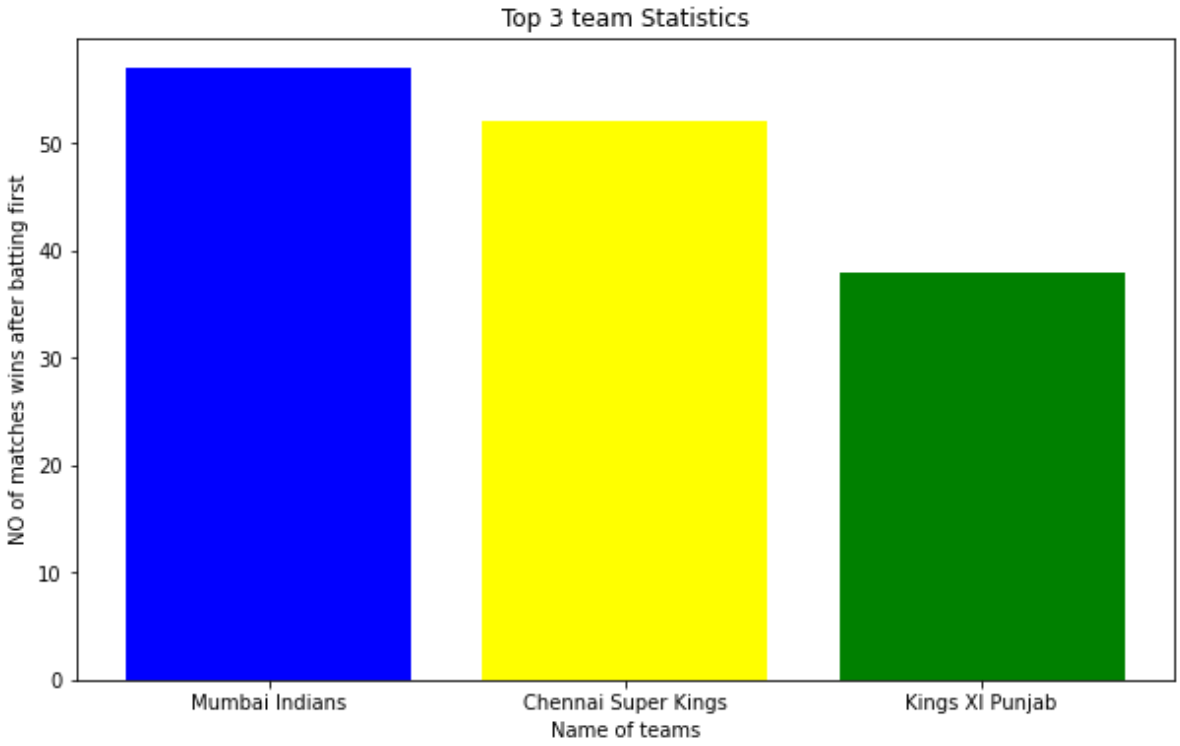
Out[37]:
```
Mumbai Indians         57
Chennai Super Kings    52
Kings XI Punjab        38
Name: winner, dtype: int64
```

In [40]:
```
#Extract name of top 3 team who wins after batting first
b =list(batting_first['winner'].value_counts().keys()[0:3])
b
```

Out[40]:
```
['Mumbai Indians', 'Chennai Super Kings', 'Kings XI Punjab']
```

In [119…
```
#Creating barplot for Top 3 Teamsn Wins After Batting First
plt.figure(figsize=(10,6))
c = ['blue','yellow','green']
plt.bar(b,a,color=c)
plt.xlabel("Name of teams")
plt.ylabel("NO of matches wins after batting first")
plt.title("Top 3 team Statistics")

plt.show()
```



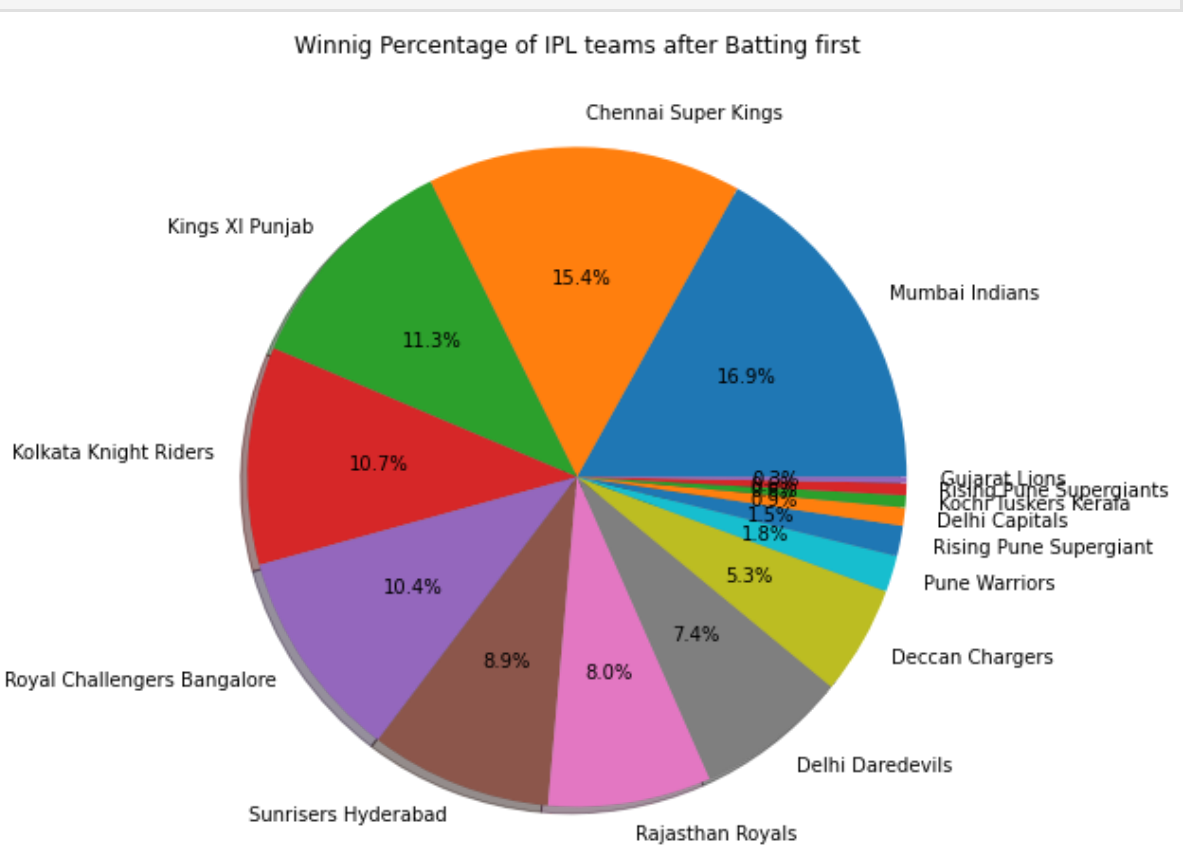
In [78]:
```
data=list(batting_first['winner'].value_counts())
data
```

Out[78]:
```
[57, 52, 38, 36, 35, 30, 27, 25, 18, 6, 5, 3, 2, 2, 1]
```

In [82]:
```
Label=list(batting_first['winner'].value_counts().keys())
Label
```

Out[82]:
```
['Mumbai Indians',
 'Chennai Super Kings',
 'Kings XI Punjab',
 'Kolkata Knight Riders',
 'Royal Challengers Bangalore',
 'Sunrisers Hyderabad',
 'Rajasthan Royals',
 'Delhi Daredevils',
 'Deccan Chargers',
 'Pune Warriors',
 'Rising Pune Supergiant',
 'Delhi Capitals',
 'Kochi Tuskers Kerala',
 'Rising Pune Supergiants',
 'Gujarat Lions']
```

In [100…
```
#Making a pie chart
plt.figure(figsize=(8,8))
plt.pie(data ,labels =Label,autopct ='%1.1f%%',shadow = True)
plt.title('Winnig Percentage of IPL teams after Batting first')
plt.show()
```



In [102…

```
#Extracting those  Records where a Team has won after Batting secnd
batting_second = ipl[ipl['win_by_wickets']!=0]
batting_second.head()
```

Out[102]:

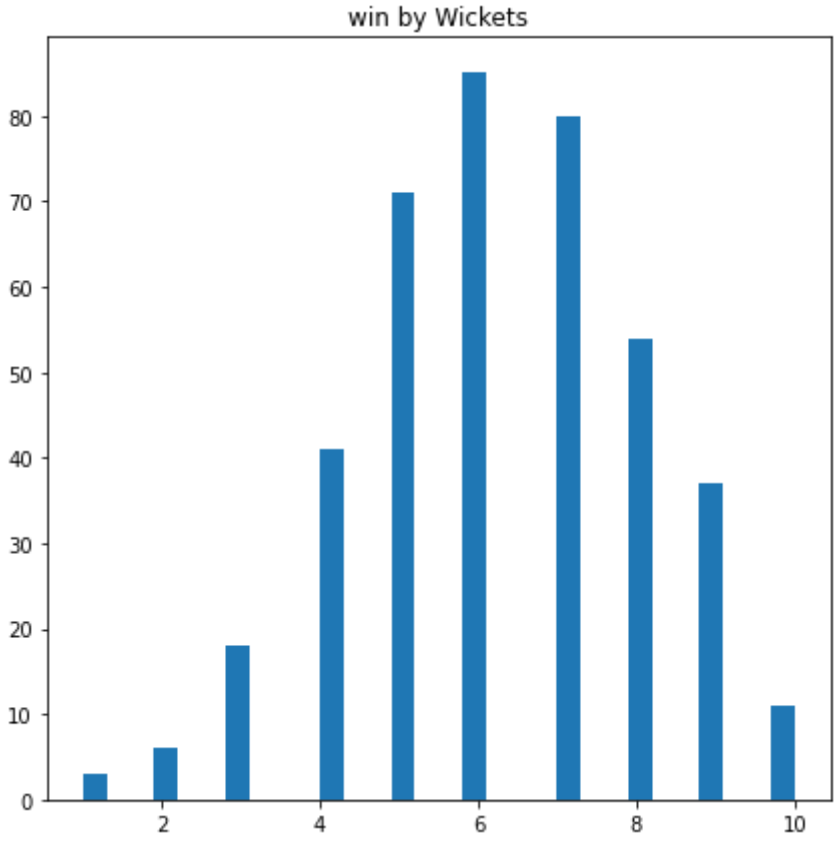
	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_ap
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	
3	4	2017	Indore	2017-04-08	Rising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

In [108…

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

Out[108]:
```
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 [111…
```
plt.figure(figsize=(7,7))
plt.hist(batting_second['win_by_wickets'],bins=30)
plt.title('win by Wickets')
plt.show()
```



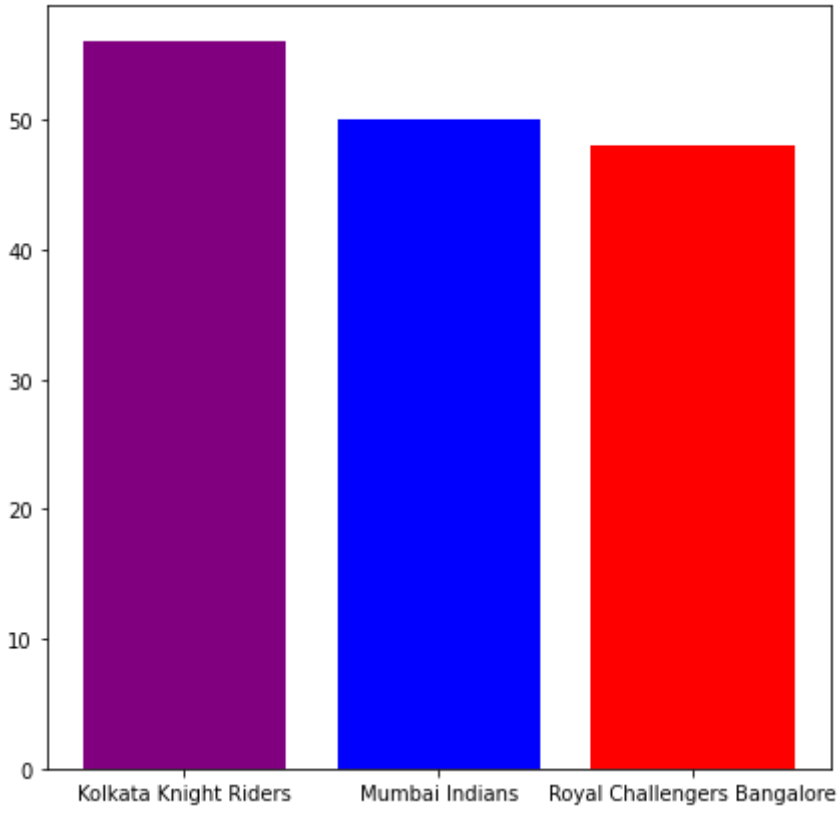
In [116…
```
#extract top 3 team who wins after second batting
f =list(batting_second['winner'].value_counts()[0:3])
f
```

Out[116]:
```
[56, 50, 48]
```

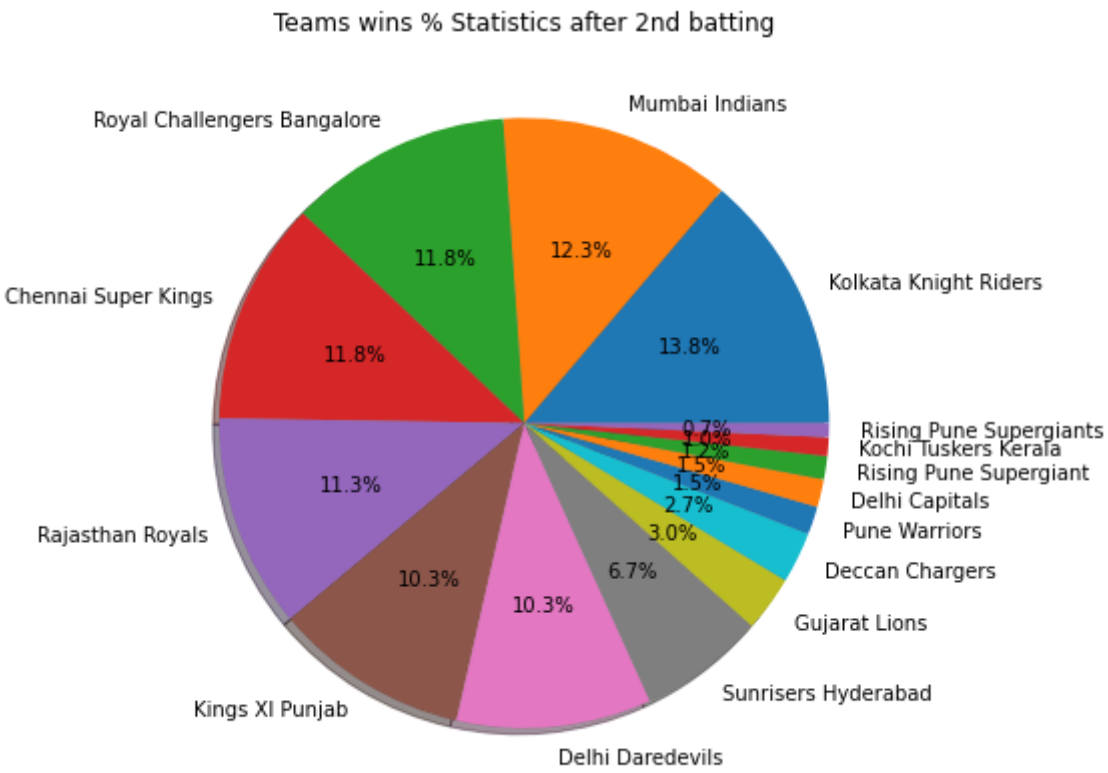
In [115…
```
g =list(batting_second['winner'].value_counts()[0:3].keys())
g
```

Out[115]:
```
['Kolkata Knight Riders', 'Mumbai Indians', 'Royal Challengers Bangalore']
```

In [123…
```
#Making a bar plot for top 3  teams with win after batting second
plt.figure(figsize=(7,7))
c = ['purple','blue','red']
plt.bar(g,f,color=c)
plt.show()
```



In [128…
```
#Making Pie chart for wins after batting second
data = list(batting_second['winner'].value_counts())
label= list(batting_second['winner'].value_counts().keys())
plt.figure(figsize=(7,7))
plt.pie(data,labels =label,autopct ='%1.1f%%',shadow =True)
plt.title('Teams wins % Statistics after 2nd batting')
plt.show()
```



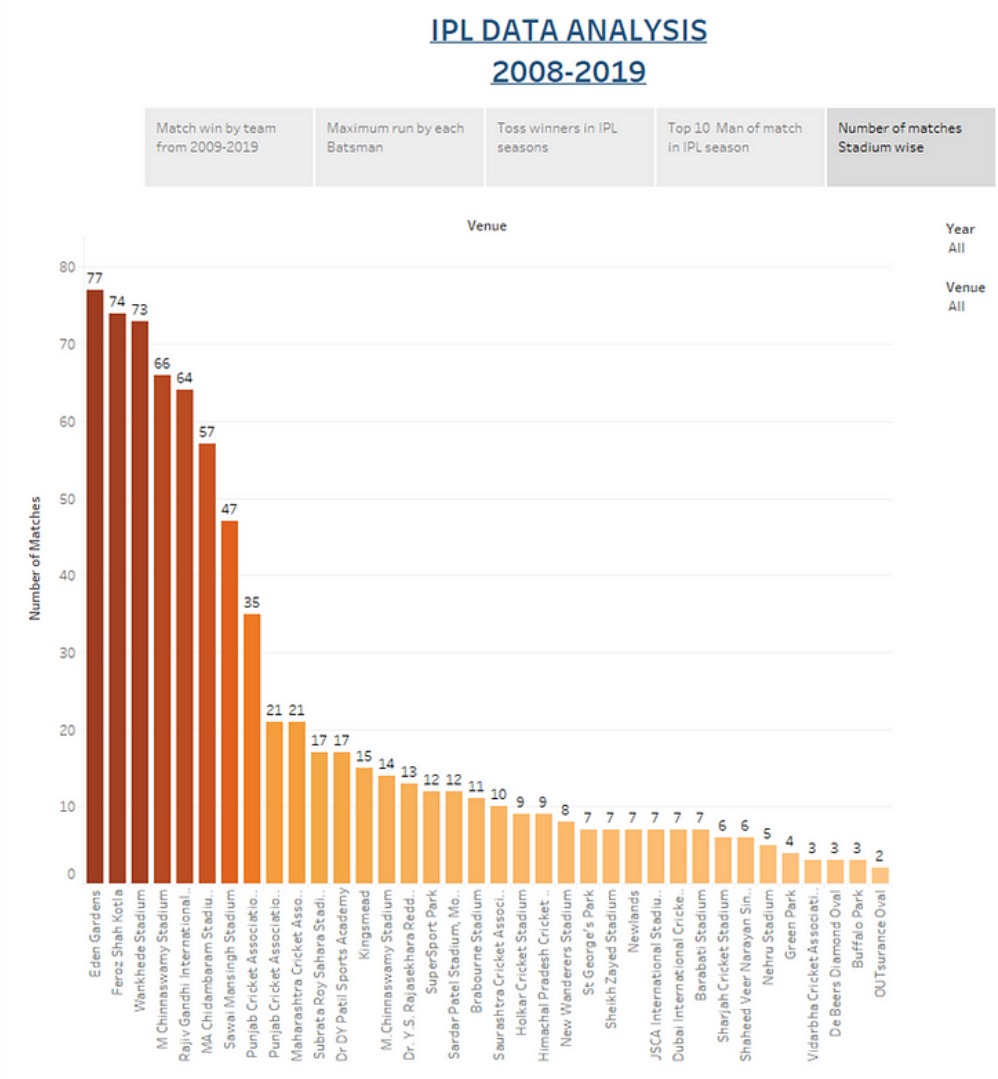
In [132…
```
#No of matches played each season
ipl['season'].value_counts()
```

Out[132]:
```
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: season, dtype: int64
```

In [133…
```
#No of matches played in each City
ipl['city'].value_counts()
```

Out[133]:

```
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 [134…
```
#Find out how many times a team has won the match after wining toss
np.sum(ipl['toss_winner']==ipl['winner'])
```

Out[134]:
```
393
```

In [136…
```
deliveries = pd.read_csv("A:\Data set\IPL DataSet\deliveries.csv")
```

In [151…
```
deliveries.tail()
deliveries.head()
```

Out[151]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is_su
0	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan	TS Mills	
1	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan	TS Mills	
2	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan	TS Mills	
3	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan	TS Mills	
4	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan	TS Mills	

5 rows × 21 columns

In [139…
```
deliveries.shape
```

Out[139]:
```
(150460, 21)
```

In [141…
```
deliveries['match_id'].unique()
```

Out[141]:
```
array([  1,   2,   3,   4,   5,   6,   7,   8,   9,  10,  11,  12,  13,
        14,  15,  16,  17,  18,  19,  20,  21,  22,  23,  24,  25,  26,
        27,  28,  29,  30,  31,  32,  33,  34,  35,  36,  37,  38,  39,
        40,  41,  42,  43,  44,  45,  46,  47,  48,  49,  50,  51,  52,
        53,  54,  55,  56,  57,  58,  59,  60,  61,  62,  63,  64,  65,
        66,  67,  68,  69,  70,  71,  72,  73,  74,  75,  76,  77,  78,
        79,  80,  81,  82,  83,  84,  85,  86,  87,  88,  89,  90,  91,
        92,  93,  94,  95,  96,  97,  98,  99, 100, 101, 102, 103, 104,
       105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117,
       118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130,
       131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143,
       144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156,
       157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169,
       170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182,
       183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195,
       196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208,
       209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221,
       222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234,
       235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247,
       248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260,
       261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273,
       274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286,
       287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299,
       300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312,
       313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325,
       326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338,
       339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351,
       352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364,
       365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377,
       378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390,
       391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403,
       404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416,
       417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429,
       430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442,
       443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455,
       456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468,
       469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481,
       482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494,
       495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507,
       508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520,
       521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533,
       534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546,
       547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559,
       560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572,
       573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585,
       586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598,
       599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611,
       612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624,
       625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636],
      dtype=int64)
```

In [142…
```
match_1=deliveries[deliveries['match_id']==1]
```

In [144…
```
match_1.head(10)
```

Out[144]:

	match_id	inning	batting_team	bowling_team	over	ball	batsman	non_striker	bowler	is
0	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	1	DA Warner	S Dhawan	TS Mills	
1	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	2	DA Warner	S Dhawan	TS Mills	
2	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	3	DA Warner	S Dhawan	TS Mills	
3	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	4	DA Warner	S Dhawan	TS Mills	
4	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	5	DA Warner	S Dhawan	TS Mills	
5	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	6	S Dhawan	DA Warner	TS Mills	
6	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	1	7	S Dhawan	DA Warner	TS Mills	
7	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	2	1	S Dhawan	DA Warner	A Choudhary	
8	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	2	2	DA Warner	S Dhawan	A Choudhary	
9	1	1	Sunrisers Hyderabad	Royal Challengers Bangalore	2	3	DA Warner	S Dhawan	A Choudhary	

10 rows × 21 columns

In [147…
```
match_1.shape
```

Out[147]:
```
(248, 21)
```

In [148…
```
srh = match_1[match_1['inning']==1]
```

In [149…
```
srh['batsman_runs'].value_counts()
```

Out[149]:
```
1    57
0    32
4    17
6     9
2     9
3     1
Name: batsman_runs, dtype: int64
```

In [152…
```
srh['dismissal_kind'].value_counts()
```

Out[152]:
```
caught    3
bowled    1
Name: dismissal_kind, dtype: int64
```

In [153…
```
rcb=match_1[match_1['inning']==2]
```

In [154…
```
rcb['batsman_runs'].value_counts()
```

Out[154]:
```
0    49
1    44
4    15
6     8
2     7
Name: batsman_runs, dtype: int64
```

In [155…
```
rcb['dismissal_kind'].value_counts()
```

Out[155]:
```
caught     6
bowled     2
run out    2
Name: dismissal_kind, dtype: int64
```

Project by - SADASHISH KUMAR & VAISHNAVI BHUSHAN