

I have IPL DataSet of matches.csv & deliveries.csv. I'll Analyst data set using EDA; Pandas, Matplotlib and Seaborn libraries are used to analyze and visualize dataset.

```
In [92]: #Import Required Libraries
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
```

```
In [93]: #Loading the Dataset
ipl=pd.read_csv('matches.csv')
```

```
In [94]: #First Five Records of Dataset
ipl.head()
```

```
Out[94]:
```

|   | id | season | city      | date       | team1                       | team2                       | toss_winner                 | toss_decision | result |
|---|----|--------|-----------|------------|-----------------------------|-----------------------------|-----------------------------|---------------|--------|
| 0 | 1  | 2017   | Hyderabad | 2017-04-05 | Sunrisers Hyderabad         | Royal Challengers Bangalore | Royal Challengers Bangalore | field         | no     |
| 1 | 2  | 2017   | Pune      | 2017-04-06 | Mumbai Indians              | Rising Pune Supergiant      | Rising Pune Supergiant      | field         | no     |
| 2 | 3  | 2017   | Rajkot    | 2017-04-07 | Gujarat Lions               | Kolkata Knight Riders       | Kolkata Knight Riders       | field         | no     |
| 3 | 4  | 2017   | Indore    | 2017-04-08 | Rising Pune Supergiant      | Kings XI Punjab             | Kings XI Punjab             | field         | no     |
| 4 | 5  | 2017   | Bangalore | 2017-04-08 | Royal Challengers Bangalore | Delhi Daredevils            | Royal Challengers Bangalore | bat           | no     |

```
In [95]: #No of Columns and Rows in Dataset
ipl.shape
```

```
Out[95]: (756, 18)
```

```
In [96]: #most player of match awards
ipl['player_of_match'].value_counts()
```

```
Out[96]: 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
         AM Rahane        12
         MEK Hussey       12
         V Kohli          12
         A Mishra         11
         AD Russell       11
         DR Smith         11
         V Sehwag         11
         JH Kallis        10
         KA Pollard       10
         AT Rayudu        9
         SP Narine        9
         SE Marsh         9
         Harbhajan Singh  8
         RA Jadeja        8
         SR Tendulkar     8
         UT Yadav        8
         AC Gilchrist     7
         RV Uthappa       7
         Rashid Khan      7
         SL Malinga       6
         RR Pant          6
         ..
         DL Chahar        1
         Z Khan           1
         KK Cooper        1
         S Gill           1
         TL Suman         1
         Q de Kock        1
         KMDN Kulasekara  1
         NV Ojha          1
         JJ Roy           1
         RE Levi          1
         DL Vettori       1
         Imran Tahir      1
         EJG Morgan       1
         A Joseph         1
         AC Voges         1
         S Sreesanth      1
         JDP Oram         1
         J Archer         1
         MJ Lumb          1
         H Gurney         1
         CR Brathwaite    1
         M Ur Rahman      1
         GD McGrath       1
         SB Jakati        1
         HH Gibbs        1
         K Rabada         1
         Mohammed Shami   1
         CRD Fernando     1
         M Kartik         1
```

A Singh 1  
Name: player\_of\_match, Length: 226, dtype: int64

```
In [97]: #Top 10 Player with Most Player of Match  
ipl['player_of_match'].value_counts()[0:10]
```

```
Out[97]: 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  
AM Rahane              12  
Name: player_of_match, dtype: int64
```

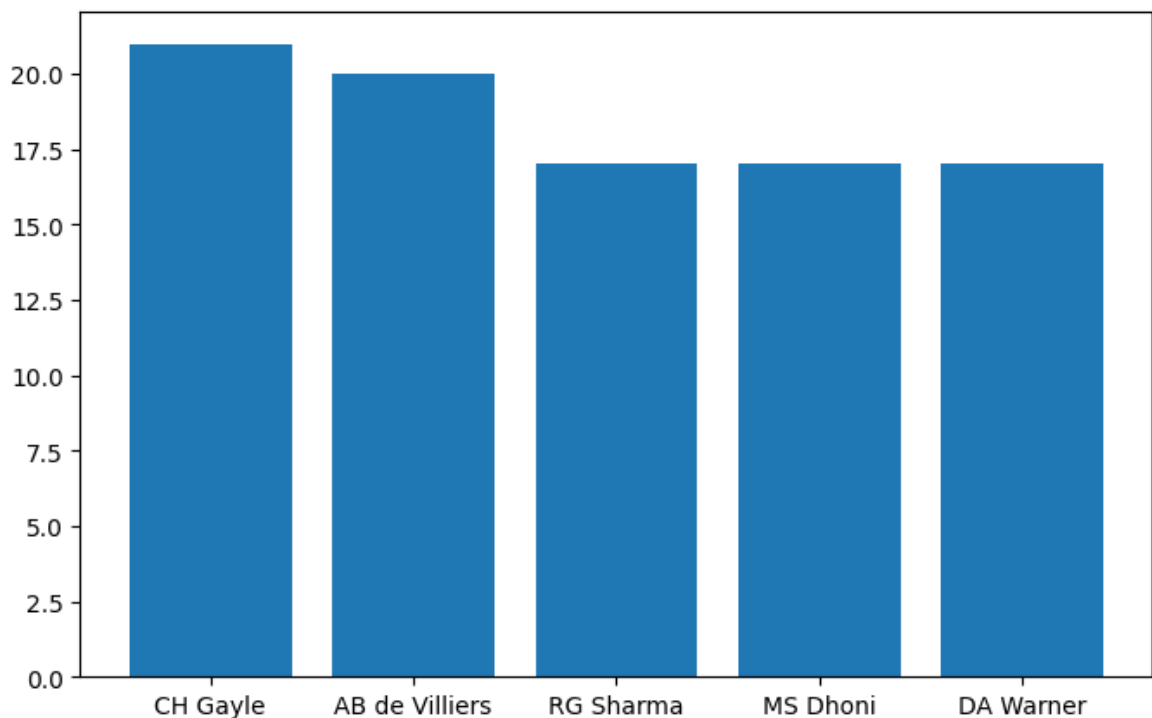
```
In [98]: #Top 5 Player of Match  
ipl['player_of_match'].value_counts()[0:5]
```

```
Out[98]: CH Gayle          21  
AB de Villiers         20  
RG Sharma              17  
MS Dhoni               17  
DA Warner              17  
Name: player_of_match, dtype: int64
```

```
In [99]: list(ipl['player_of_match'].value_counts()[0:5].keys())
```

```
Out[99]: ['CH Gayle', 'AB de Villiers', 'RG Sharma', 'MS Dhoni', 'DA Warner']
```

```
In [100... #Barplot of Top 5 Player of Match  
plt.figure(figsize=(8,5))  
plt.bar(list(ipl['player_of_match'].value_counts()[0:5].keys()),list(ipl['player  
plt.show()
```



```
In [101... #Getting the frequency pf result columns
ipl['result'].value_counts()
```

```
Out[101]: normal      743
tie          9
no result    4
Name: result, dtype: int64
```

```
In [102... #Find out No of Toss Wins w.r.t. Each Team
ipl['toss_winner'].value_counts()
```

```
Out[102]: 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 [103... #Extracting the records where a team won batting first
batting_first=ipl[ipl['win_by_runs']!=0]
```

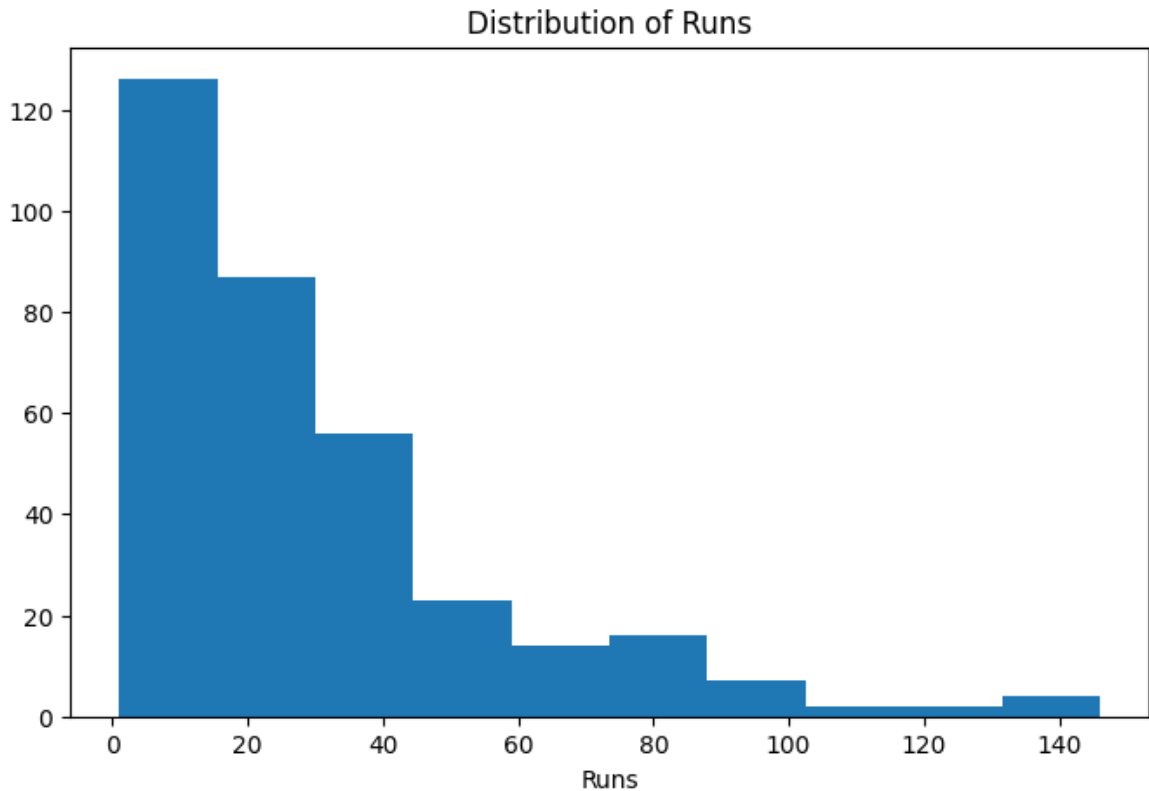
```
In [104... batting_first.head()
```

```
Out[104]:
```

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

```
In [105... #creating plot of Win_by_runs
plt.figure(figsize=(8,5))
```

```
plt.hist(batting_first['win_by_runs'])
plt.title('Distribution of Runs')
plt.xlabel('Runs')
plt.show()
```



In [106... *#No of Wins w.r.t. Rach Team After Batting First*

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

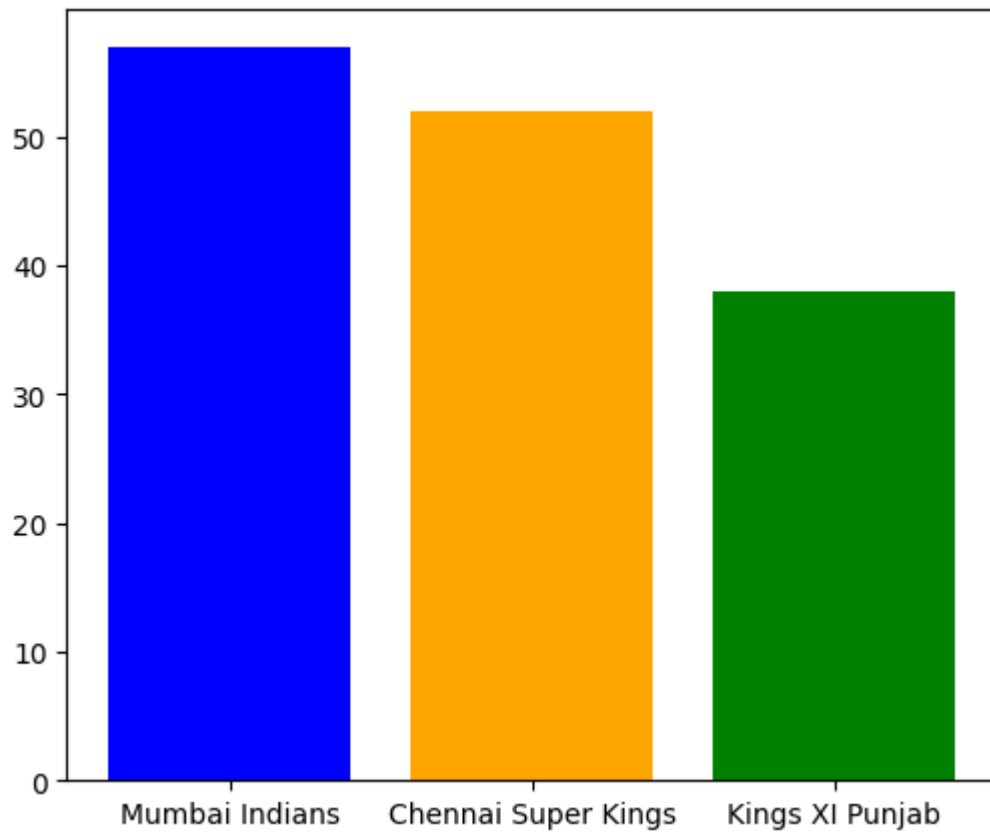
Out[106]:

|                             |    |
|-----------------------------|----|
| 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  |

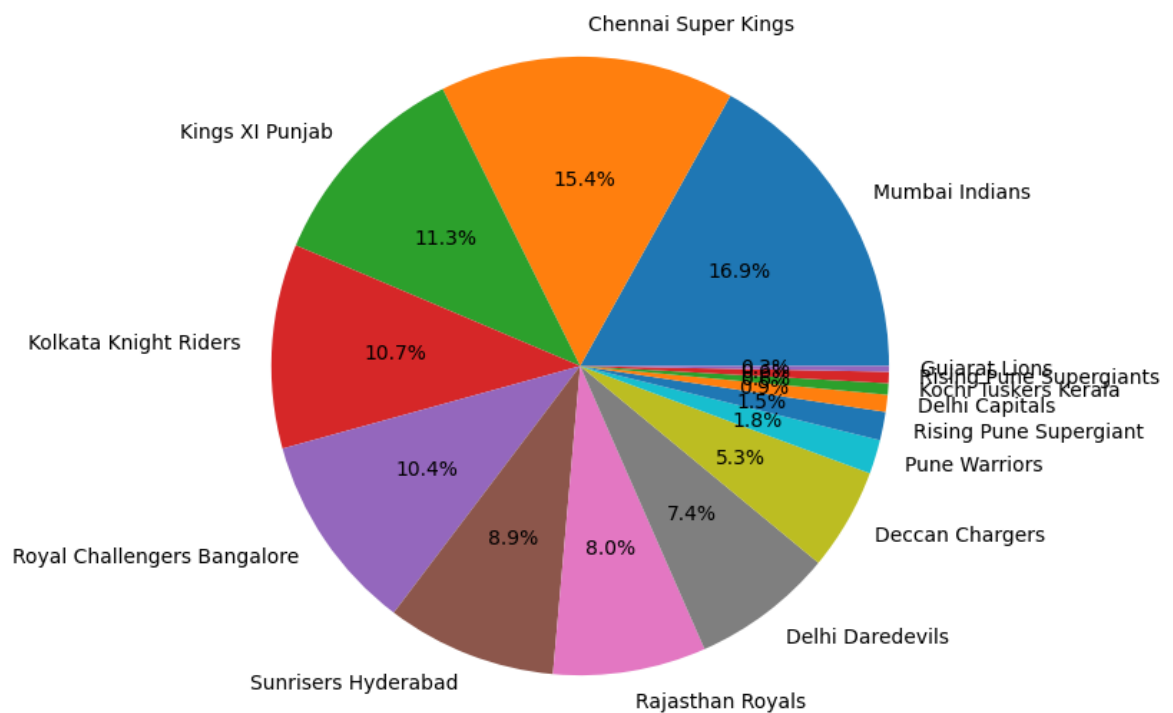
Name: winner, dtype: int64

In [107... *#Creating barplot for Top 3 Teams Wins After Batting First*

```
plt.figure(figsize=(6,5))
plt.bar(list(batting_first['winner'].value_counts()[0:3].keys()),list(batting_fi
plt.show()
```



```
In [108... #Making Pie Chart
plt.figure(figsize=(7,7))
plt.pie(list(batting_first['winner'].value_counts()),labels=list(batting_first['
plt.show()
```



```
In [109... #Extracting those Records Where A Team has Won After Batting Second
batting_second=ipl[ipl['win_by_wickets']!=0]
```

```
In [110... batting_second.head()
```

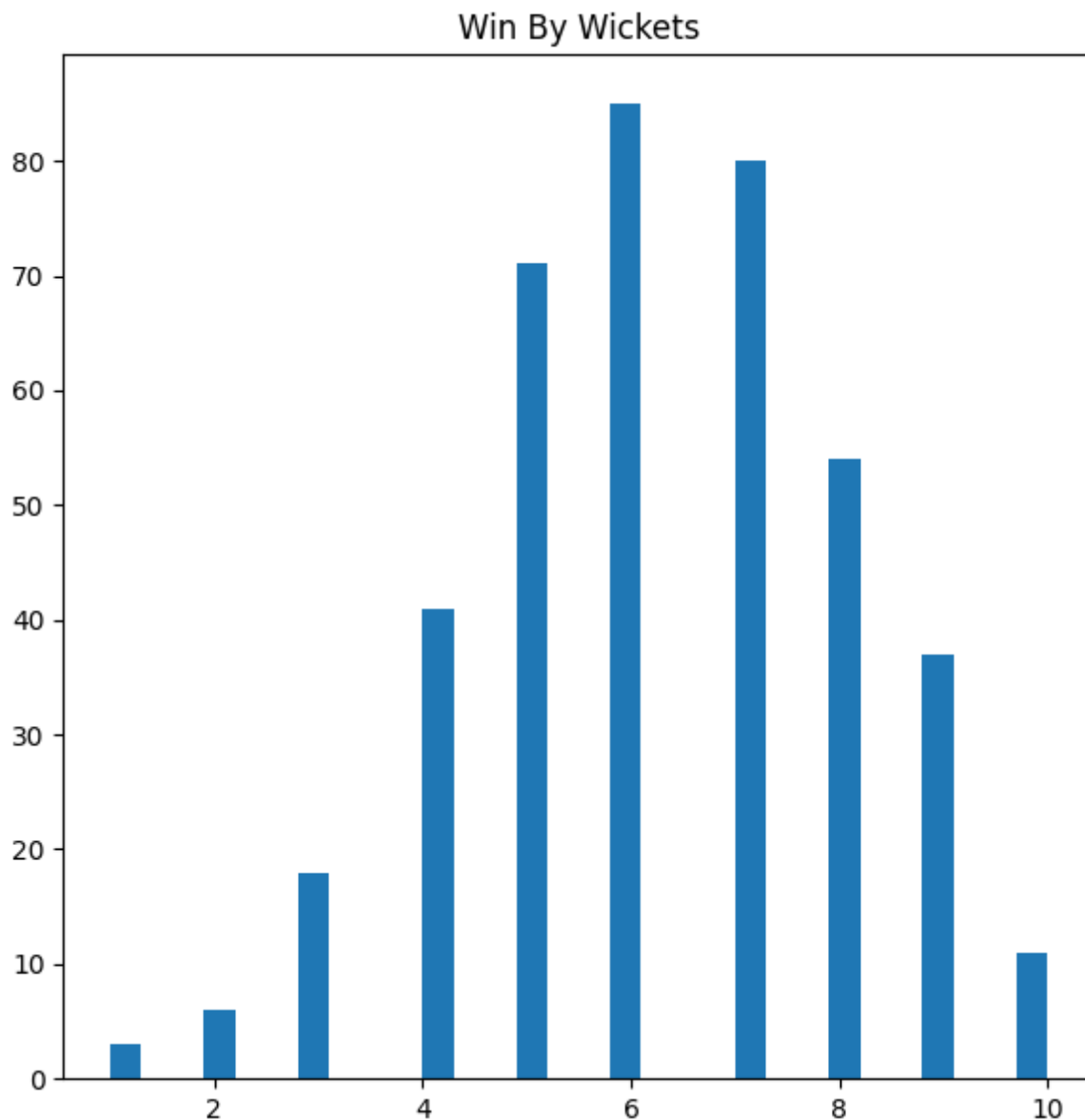
Out[110]:

|  | id | season | city | date | team1 | team2 | toss_winner | toss_decision | re |
|--|----|--------|------|------|-------|-------|-------------|---------------|----|
|--|----|--------|------|------|-------|-------|-------------|---------------|----|

|   |   |      |           |            |                        |                        |                        |       |     |
|---|---|------|-----------|------------|------------------------|------------------------|------------------------|-------|-----|
| 1 | 2 | 2017 | Pune      | 2017-04-06 | Mumbai Indians         | Rising Pune Supergiant | Rising Pune Supergiant | field | nor |
| 2 | 3 | 2017 | Rajkot    | 2017-04-07 | Gujarat Lions          | Kolkata Knight Riders  | Kolkata Knight Riders  | field | nor |
| 3 | 4 | 2017 | Indore    | 2017-04-08 | Rising Pune Supergiant | Kings XI Punjab        | Kings XI Punjab        | field | nor |
| 5 | 6 | 2017 | Hyderabad | 2017-04-09 | Gujarat Lions          | Sunrisers Hyderabad    | Sunrisers Hyderabad    | field | nor |
| 6 | 7 | 2017 | Mumbai    | 2017-04-09 | Kolkata Knight Riders  | Mumbai Indians         | Mumbai Indians         | field | nor |

In [111...

```
#Making Histogram for Frequency of wins w.r.t. No. of Wickets
plt.figure(figsize=(7,7))
plt.hist(batting_second['win_by_wickets'],bins=30)
plt.title('Win By Wickets')
plt.show()
```



In [112... *#finding out the frequency of no of wins w.r.t. each time after batting second*  
`batting_second['winner'].value_counts()`

Out[112]:

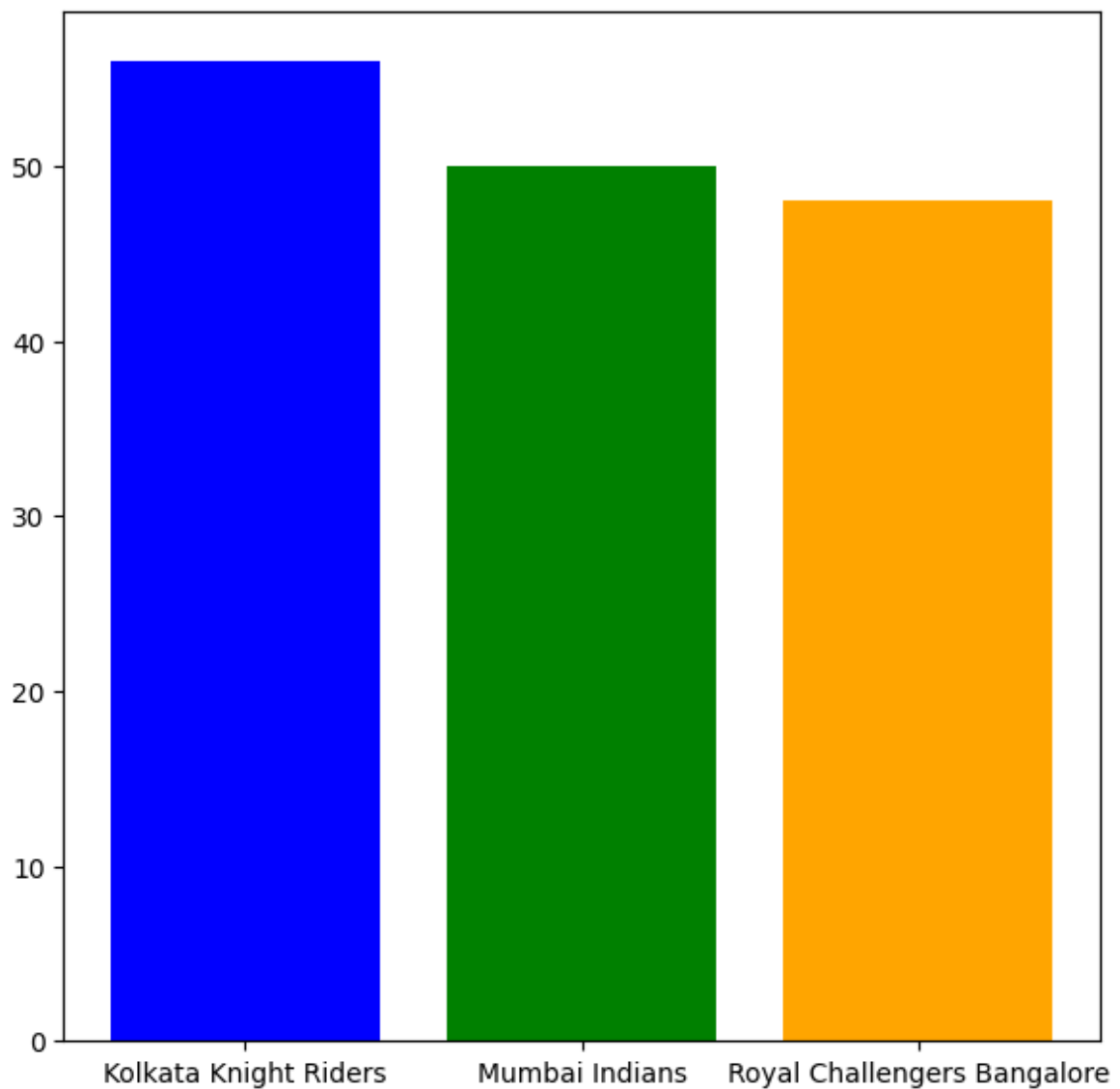
|                             |    |
|-----------------------------|----|
| Kolkata Knight Riders       | 56 |
| Mumbai Indians              | 50 |
| Royal Challengers Bangalore | 48 |
| Chennai Super Kings         | 48 |
| Rajasthan Royals            | 46 |
| Delhi Daredevils            | 42 |
| Kings XI Punjab             | 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

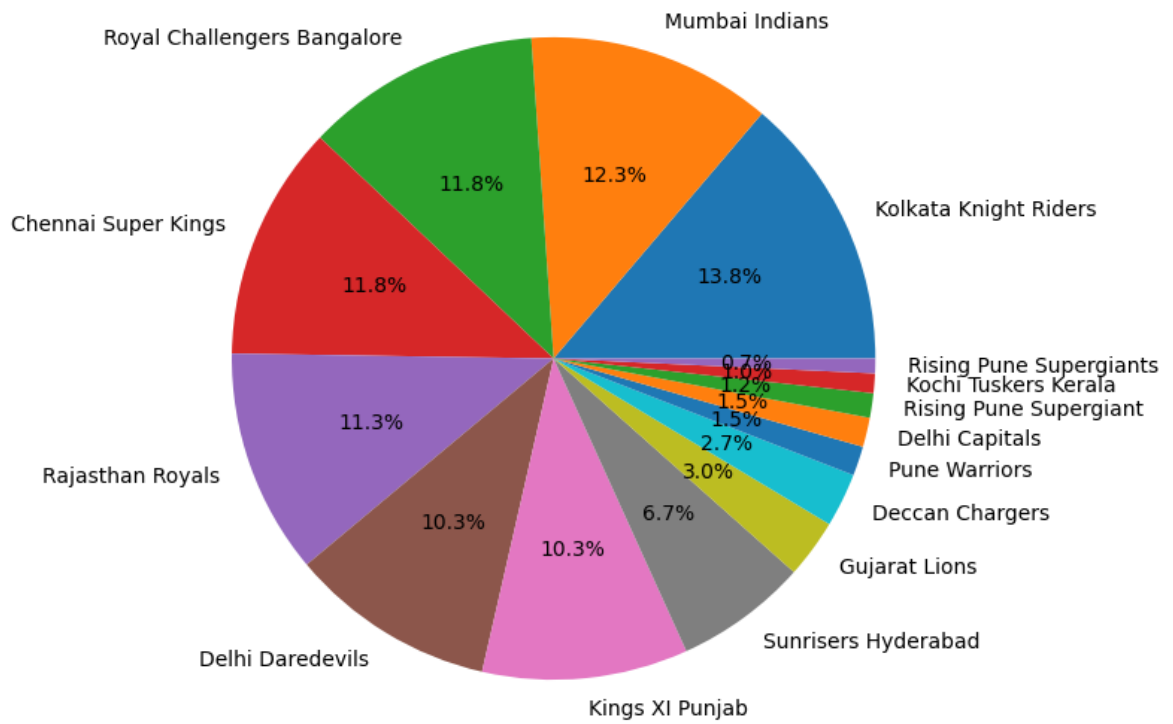
In [113... *#Making a bar plot for Top 3 teamsn with wins after batting second*



```
In [114... plt.figure(figsize=(7,7))  
plt.bar(list(batting_second['winner'].value_counts()[0:3].keys()),list(batting_s  
plt.show()
```



```
In [115... #Making Pie chart for wins after batting second  
plt.figure(figsize=(7,7))  
plt.pie(list(batting_second['winner'].value_counts()),labels=list(batting_secon  
plt.show()
```



```
In [116... #No of matches played each season
ipl['season'].value_counts()
```

```
Out[116]: 2013    76
          2012    74
          2011    73
          2019    60
          2018    60
          2016    60
          2014    60
          2010    60
          2017    59
          2015    59
          2008    58
          2009    57
          Name: season, dtype: int64
```

```
In [117... #No of matches played in each city
ipl['city'].value_counts()
```

```
Out[117]: 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
          Mohali            10
          Rajkot            10
          Indore            9
          Dharamsala        9
          Johannesburg      8
          Cuttack           7
          Cape Town         7
          Port Elizabeth    7
          Ranchi            7
          Abu Dhabi         7
          Raipur            6
          Sharjah           6
          Kochi             5
          Kanpur            4
          Nagpur            3
          East London       3
          Kimberley         3
          Bloemfontein      2
          Name: city, dtype: int64
```

```
In [118... #Find out how many times a team has won the match after wining toss
import numpy as np
np.sum(ipl['toss_winner']==ipl['winner'])
```

```
Out[118]: 393
```

```
In [119... deliveries=pd.read_csv('deliveries.csv')
```

```
In [120... deliveries.head()
```

Out[120]:

|   | match_id | inning | batting_team        | bowling_team                | over | ball | batsman   | non_striker | bc |
|---|----------|--------|---------------------|-----------------------------|------|------|-----------|-------------|----|
| 0 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 1    | DA Warner | S Dhawan    |    |
| 1 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 2    | DA Warner | S Dhawan    |    |
| 2 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 3    | DA Warner | S Dhawan    |    |
| 3 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 4    | DA Warner | S Dhawan    |    |
| 4 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 5    | DA Warner | S Dhawan    |    |

5 rows × 21 columns

◀

▶

In [121... deliveries['match\_id'].unique()

```
Out[121]: 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 [122... match_1=deliveries[deliveries['match_id']==1]
```

```
In [123... match_1.head()
```

Out[123]:

|   | match_id | inning | batting_team        | bowling_team                | over | ball | batsman   | non_striker | bc |
|---|----------|--------|---------------------|-----------------------------|------|------|-----------|-------------|----|
| 0 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 1    | DA Warner | S Dhawan    |    |
| 1 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 2    | DA Warner | S Dhawan    |    |
| 2 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 3    | DA Warner | S Dhawan    |    |
| 3 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 4    | DA Warner | S Dhawan    |    |
| 4 | 1        | 1      | Sunrisers Hyderabad | Royal Challengers Bangalore | 1    | 5    | DA Warner | S Dhawan    |    |

5 rows × 21 columns

In [124... match\_1.shape

Out[124]: (248, 21)

In [125... srh=match\_1[match\_1['inning']==1]

In [126... srh['batsman\_runs'].value\_counts()

Out[126]:

|   |    |
|---|----|
| 1 | 57 |
| 0 | 32 |
| 4 | 17 |
| 6 | 9  |
| 2 | 9  |
| 3 | 1  |

Name: batsman\_runs, dtype: int64

In [127... srh['dismissal\_kind'].value\_counts()

Out[127]:

|        |   |
|--------|---|
| caught | 3 |
| bowled | 1 |

Name: dismissal\_kind, dtype: int64

In [128... rcb=match\_1[match\_1['inning']==2]

In [129... rcb['batsman\_runs'].value\_counts()

Out[129]:

|   |    |
|---|----|
| 0 | 49 |
| 1 | 44 |
| 4 | 15 |
| 6 | 8  |
| 2 | 7  |

Name: batsman\_runs, dtype: int64

```
In [130... rcb['dismissal_kind'].value_counts()
```

```
Out[130]: caught      6  
run out    2  
bowled     2  
Name: dismissal_kind, dtype: int64
```