#EXPLORATORY DATA ANALYSIS (EDA) FOR IPL DATASET USING PYTHON PROGRAMMING @27/07/2021

#NAME : SEKHAR REDDY
#EMAIL : sekharreddy717.niper@gmail.com
#EMAIL : 8331995717

#IMPORTING THE DATASET

#matches.csv

#importing the python libraries


```python
import pandas as pd #stores the data as dataframes
import numpy as np #numerical computations using data strctures
import matplotlib as mpl #data visualization
import matplotlib.pyplot as plt #data visualization
import seaborn as sns #data visualization
%matplotlib inline
```


```python
data = pd.read_csv("/content/matches.csv") #reading the data by giving data path
```


```python
data.head() #top 5
```

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


```python
data.tail() #last 5
```

	id	Season	city	date	team1	team2	toss_winner	toss_decisi
751	11347	IPL-2019	Mumbai	05-05-2019	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	fi
752	11412	IPL-2019	Chennai	07-05-2019	Chennai Super Kings	Mumbai Indians	Chennai Super Kings	
753	11413	IPL-2019	Visakhapatnam	08-05-2019	Sunrisers Hyderabad	Delhi Capitals	Delhi Capitals	fi
754	11414	IPL-2019	Visakhapatnam	10-05-2019	Delhi Capitals	Chennai Super Kings	Chennai Super Kings	fi
755	11415	IPL-2019	Hyderabad	12-05-2019	Mumbai Indians	Chennai Super Kings	Mumbai Indians	

```
data.head(4) #top 4
```

	id	Season	city	date	team1	team2	toss_winner	toss_decision	r
0	1	IPL-2017	Hyderabad	05-04-2017	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	n
1	2	IPL-2017	Pune	06-04-2017	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	n
2	3	IPL-2017	Rajkot	07-04-2017	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	n
3	4	IPL-2017	Indore	08-04-2017	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	n

```
data.tail(3) #last 3
```

	id	Season	city	date	team1	team2	toss_winner	toss_decisi
753	11413	IPL-2019	Visakhapatnam	08-05-2019	Sunrisers Hyderabad	Delhi Capitals	Delhi Capitals	fi
754	11414	IPL-2019	Visakhapatnam	10-05-	Delhi Capitals	Chennai Super	Chennai Super Kings	fi

```python
data.shape #data size - rows and colums
```

```
(756, 18)
```

| | | 2019 | | 2019 | Indians | Kings | Indians | |

```python
data.info #data more precisely
```

```
<bound method DataFrame.info of        id    Season  ...             umpire2
0        1  IPL-2017  ...        NJ Llong                NaN
1        2  IPL-2017  ...          S Ravi                NaN
2        3  IPL-2017  ...       CK Nandan                NaN
3        4  IPL-2017  ...   C Shamshuddin                NaN
4        5  IPL-2017  ...             NaN                NaN
..     ...       ...  ...             ...                ...
751  11347  IPL-2019  ...        O Nandan             S Ravi
752  11412  IPL-2019  ...     Nitin Menon          Ian Gould
753  11413  IPL-2019  ...             NaN                NaN
754  11414  IPL-2019  ...   Bruce Oxenford  Chettithody Shamshuddin
755  11415  IPL-2019  ...       Ian Gould        Nigel Llong

[756 rows x 18 columns]>
```

```python
data.columns #all columns
```

```
Index(['id', 'Season', 'city', 'date', 'team1', 'team2', 'toss_winner',
       'toss_decision', 'result', 'dl_applied', 'winner', 'win_by_runs',
       'win_by_wickets', 'player_of_match', 'venue', 'umpire1', 'umpire2',
       'umpire3'],
      dtype='object')
```

```python
data.isna().any()
```

```
id                 False
Season             False
city                True
date               False
team1              False
team2              False
toss_winner        False
toss_decision      False
result             False
dl_applied         False
winner              True
win_by_runs        False
win_by_wickets     False
player_of_match     True
venue              False
umpire1             True
```

```
        umpire2           True
        umpire3           True
        dtype: bool
```

data.describe() #stastical description

	id	dl_applied	win_by_runs	win_by_wickets
count	756.000000	756.000000	756.000000	756.000000
mean	1792.178571	0.025132	13.283069	3.350529
std	3464.478148	0.156630	23.471144	3.387963
min	1.000000	0.000000	0.000000	0.000000
25%	189.750000	0.000000	0.000000	0.000000
50%	378.500000	0.000000	0.000000	4.000000
75%	567.250000	0.000000	19.000000	6.000000
max	11415.000000	1.000000	146.000000	10.000000

data.count()

```
        id                756
        Season            756
        city              749
        date              756
        team1             756
        team2             756
        toss_winner       756
        toss_decision     756
        result            756
        dl_applied        756
        winner            752
        win_by_runs       756
        win_by_wickets    756
        player_of_match   752
        venue             756
        umpire1           754
        umpire2           754
        umpire3           119
        dtype: int64
```

data["id"].count() #total number of matches played

```
        756
```

data["Season"].unique() #how many ipl seasons we are analyzing

```
        array(['IPL-2017', 'IPL-2008', 'IPL-2009', 'IPL-2010', 'IPL-2011',
               'IPL-2012', 'IPL-2013', 'IPL-2014', 'IPL-2015', 'IPL-2016',
               'IPL-2018', 'IPL-2019'], dtype=object)
```

data.iloc[data["win_by_runs"].idxmax()] #won by highest number of wickets

```
id                              44
Season                     IPL-2017
city                          Delhi
date                     06-05-2017
team1              Mumbai Indians
team2            Delhi Daredevils
toss_winner      Delhi Daredevils
toss_decision                field
result                      normal
dl_applied                       0
winner             Mumbai Indians
win_by_runs                    146
win_by_wickets                   0
player_of_match        LMP Simmons
venue            Feroz Shah Kotla
umpire1                 Nitin Menon
umpire2                   CK Nandan
umpire3                        NaN
Name: 43, dtype: object
```

```
data.iloc[data['win_by_wickets'].idxmax()] #won by maxium wickets
```

```
id                                                3
Season                                     IPL-2017
city                                         Rajkot
date                                     07-04-2017
team1                                  Gujarat Lions
team2                          Kolkata Knight Riders
toss_winner                    Kolkata Knight Riders
toss_decision                                  field
result                                        normal
dl_applied                                         0
winner                         Kolkata Knight Riders
win_by_runs                                        0
win_by_wickets                                    10
player_of_match                             CA Lynn
venue            Saurashtra Cricket Association Stadium
umpire1                                   Nitin Menon
umpire2                                     CK Nandan
umpire3                                          NaN
Name: 2, dtype: object
```

```
data.iloc[data["win_by_wickets"].idxmin()] #won by taking minimum wickets
```

```
id                                                1
Season                                     IPL-2017
city                                      Hyderabad
date                                     05-04-2017
team1                              Sunrisers Hyderabad
team2                      Royal Challengers Bangalore
toss_winner                Royal Challengers Bangalore
toss_decision                                  field
result                                        normal
dl_applied                                         0
winner                             Sunrisers Hyderabad
win_by_runs                                       35
win_by_wickets                                     0
player_of_match                         Yuvraj Singh
```

```
    venue               Rajiv Gandhi International Stadium, Uppal
    umpire1                                         AY Dandekar
    umpire2                                           NJ Llong
    umpire3                                                NaN
    Name: 0, dtype: object
```

data.iloc[data["win_by_runs"].idxmin()] #won by minimum number of runs

```
    id                                                       2
    Season                                            IPL-2017
    city                                                  Pune
    date                                            06-04-2017
    team1                                        Mumbai Indians
    team2                                  Rising Pune Supergiant
    toss_winner                            Rising Pune Supergiant
    toss_decision                                        field
    result                                              normal
    dl_applied                                               0
    winner                                 Rising Pune Supergiant
    win_by_runs                                              0
    win_by_wickets                                           7
    player_of_match                                  SPD Smith
    venue               Maharashtra Cricket Association Stadium
    umpire1                                     A Nand Kishore
    umpire2                                             S Ravi
    umpire3                                                NaN
    Name: 1, dtype: object
```
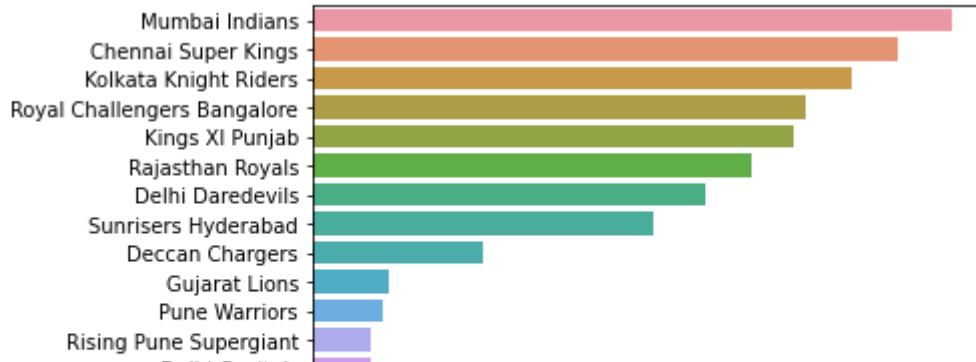
```
fig_dims=(20,4)
fig , ax = plt.subplots(figsize=fig_dims)
sns.countplot(x='Season',ax = ax, data = data)
plt.show() #Highest number of matches played in which season
```



```
data1=data.winner.value_counts() #sucessful IPL team
sns.barplot(y = data1.index , x = data1)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f2f6225b550>
```
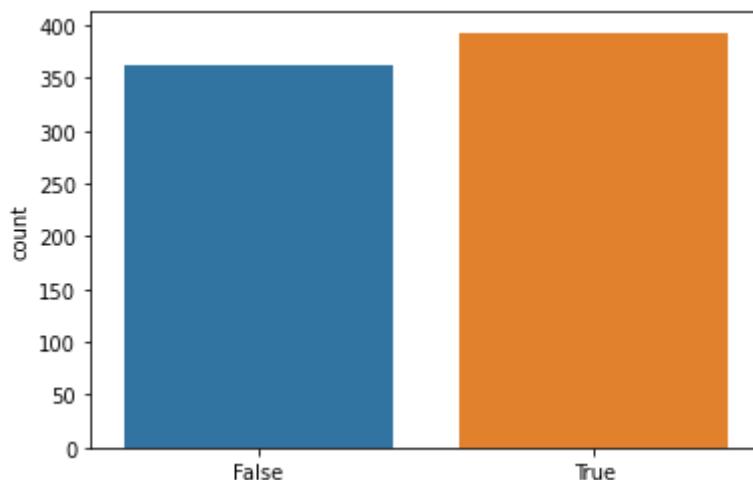


```
probability_of_win = data['toss_winner']== data['winner'] #probability of win if toss is w
probability_of_win.groupby(probability_of_win).size()
```

```
False    363
True     393
dtype: int64
```

```
sns.countplot(probability_of_win)
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pas
  FutureWarning
<matplotlib.axes._subplots.AxesSubplot at 0x7f2f62149ad0>
```



```
pd.set_option('max_rows',99999)
pd.set_option('max_colwidth',400)
pd.describe_option('max_colwidth')
```

```
display.max_colwidth : int or None
    The maximum width in characters of a column in the repr of
    a pandas data structure. When the column overflows, a "..."
    placeholder is embedded in the output. A 'None' value means unlimited.
    [default: 50] [currently: 400]
```

```
data.groupby('Season')['winner'].value_counts()
```

```
          Pune Warriors           4
IPL-2013  Mumbai Indians          13
          Chennai Super Kings     12
          Rajasthan Royals        11
          Sunrisers Hyderabad     10
```

```
              Royal Challengers Bangalore      9
              Kings XI Punjab                  8
              Kolkata Knight Riders            6
              Pune Warriors                    4
              Delhi Daredevils                 3
    IPL-2014  Kings XI Punjab                 12
              Kolkata Knight Riders           11
              Chennai Super Kings             10
              Mumbai Indians                   7
              Rajasthan Royals                 7
              Sunrisers Hyderabad              6
              Royal Challengers Bangalore      5
              Delhi Daredevils                 2
    IPL-2015  Chennai Super Kings             10
              Mumbai Indians                  10
              Royal Challengers Bangalore      8
              Kolkata Knight Riders            7

              Rajasthan Royals                 7
              Sunrisers Hyderabad              7
              Delhi Daredevils                 5
              Kings XI Punjab                  3
    IPL-2016  Sunrisers Hyderabad             11
              Gujarat Lions                    9
              Royal Challengers Bangalore      9
              Kolkata Knight Riders            8
              Delhi Daredevils                 7
              Mumbai Indians                   7
              Rising Pune Supergiants          5
              Kings XI Punjab                  4
    IPL-2017  Mumbai Indians                  12
              Rising Pune Supergiant          10
              Kolkata Knight Riders            9
              Sunrisers Hyderabad              8
              Kings XI Punjab                  7
              Delhi Daredevils                 6
              Gujarat Lions                    4
              Royal Challengers Bangalore      3
    IPL-2018  Chennai Super Kings             11
              Sunrisers Hyderabad             10
              Kolkata Knight Riders            9
              Rajasthan Royals                 7
              Kings XI Punjab                  6
              Mumbai Indians                   6
              Royal Challengers Bangalore      6
              Delhi Daredevils                 5
    IPL-2019  Mumbai Indians                  11
              Chennai Super Kings             10
              Delhi Capitals                  10
              Kings XI Punjab                  6
              Kolkata Knight Riders            6
              Sunrisers Hyderabad              6
              Rajasthan Royals                 5
              Royal Challengers Bangalore      5
    Name: winner, dtype: int64
```

```
data['toss_decision'].value_counts()
```

```
    field    463
    bat      293
```

```
       Name: toss_decision, dtype: int64

data['player_of_match'].value_counts() #man of the match highest to lowest

       CH Gayle             21
       AB de Villiers       20
       RG Sharma            17
       DA Warner            17
       MS Dhoni             17
       YK Pathan            16
       SR Watson            15
       SK Raina             14
       G Gambhir            13
       V Kohli              12
       MEK Hussey           12
       AM Rahane            12
       V Sehwag             11
       A Mishra             11
       AD Russell           11
       DR Smith             11
       KA Pollard           10
       JH Kallis            10
       AT Rayudu             9
       SP Narine             9
       SE Marsh              9
       Harbhajan Singh       8
       UT Yadav              8
       SR Tendulkar          8
       RA Jadeja             8
       RV Uthappa            7
       AC Gilchrist          7
       Rashid Khan           7
       RR Pant               6
       HH Pandya             6
       SL Malinga            6
       S Dhawan              6
       BJ Hodge              6
       A Nehra               6
       M Vijay               6
       JP Faulkner           5
       PA Patel              5
       Yuvraj Singh          5
       SV Samson             5
       GJ Maxwell            5
       Sandeep Sharma        5
       JJ Bumrah             5
       JD Unadkat            5
       B Kumar               5
       SPD Smith             5
       AJ Finch              5
       KC Sangakkara         5
       BB McCullum           5
       SC Ganguly            5
       DW Steyn              5
       DPMD Jayawardene      5
       JC Buttler            5
       ML Hayden             4
       F du Plessis          4
       JP Duminy             4
       NM Coulter-Nile       4
       SK Warne              4
```

```
    MK Pandey              4
    KL Rahul               4
```

```
data['city'].value_counts() #maximum number of matches played
```

```
    Mumbai            101
    Kolkata            77
    Delhi              74
    Bangalore          66
    Hyderabad          64
    Chennai            57
    Jaipur             47
    Chandigarh         46
    Pune               38
    Durban             15
    Bengaluru          14
    Visakhapatnam      13
    Ahmedabad          12
    Centurion          12
    Mohali             10
    Rajkot             10
    Dharamsala          9
    Indore              9
    Johannesburg        8
    Cuttack             7
    Abu Dhabi           7
    Cape Town           7
    Port Elizabeth      7
    Ranchi              7
    Sharjah             6
    Raipur              6
    Kochi               5
    Kanpur              4
    Nagpur              3
    Kimberley           3
    East London         3
    Bloemfontein        2
    Name: city, dtype: int64
```