```
# Installing seaborn library
%pip install seaborn

# Importing required library
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sns

# Loading csv file
ipl = pd.read_csv('ipl2024 Matches.csv')
ipl.head()
```

```
   id           date      team1      team2 toss_winner decision
first_score  \
0   1  March 22,2024   Banglore    Chennai    Banglore      Bat
173
1   2  March 23,2024      Delhi     Punjab      Punjab    Field
174
2   3  March 23,2024    Kolkata  Hyderabad   Hyderabad    Field
208
3   4  March 24,2024  Rajasthan    Lucknow   Rajasthan      Bat
193
4   5  March 24,2024    Gujarat     Mumbai      Mumbai    Field
168

   first_wkts  second_score  second_wkts     winner
player_of_the_match  \
0           6           176            4    Chennai   Mustafizur
Rahman
1           9           177            6     Punjab          Sam
Curran
2           7           204            7    Kolkata       Andre
Russell
3           4           173            6  Rajasthan        Sanju
Samson
4           6           162            9    Gujarat       Sai
Sudharsan

       most_runs          most_wkts
0     Anuj Rawat  Mustafizur Rahman
1     Sam Curran      Kuldeep Yadav
2  Andre Russell        T Natarajan
3   Sanju Samson        Trent Boult
4  Dewald Brevis     Jasprit Bumrah
```

```
# First 5 rows of datasets
ipl.tail()
```

```
    id         date      team1      team2 toss_winner decision
first_score  \
```

```
69  70  May 19,2024  Rajasthan    Kolkata     Kolkata      NaN
0
70  71  May 21,2024  Hyderabad    Kolkata   Hyderabad      Bat
159
71  72  May 22,2024   Banglore  Rajasthan   Rajasthan    Field
172
72  73  May 24,2024  Hyderabad  Rajasthan   Rajasthan    Field
175
73  74  May 26,2024  Hyderabad    Kolkata   Hyderabad      Bat
113

    first_wkts  second_score  second_wkts     winner
player_of_the_match  \
69           0             0            0  Abandoned
NaN
70          10           164            2    Kolkata       Mitchell
Starc
71           8           174            6  Rajasthan  Ravichandran
Ashwin
72           9           139            7  Hyderabad        Shahbaz
Ahmed
73          10           114            2    Kolkata       Mitchell
Starc

           most_runs       most_wkts
69               NaN             NaN
70      Shreyas Iyer  Mitchell Starc
71  Yashasvi Jaiswal      Avesh Khan
72      Dhruv Jurel    Shahbaz Ahmed
73    Venkatesh Iyer   Andre Russell
```

```
# used to describe the dataset
ipl.describe()
```

```
              id  first_score  first_wkts  second_score  second_wkts
count  74.000000    74.000000   74.000000     74.000000    74.000000
mean   37.500000   180.554054    6.148649    169.054054     5.783784
std    21.505813    51.855474    2.469998     47.651386     2.934305
min     1.000000     0.000000    0.000000      0.000000     0.000000
25%    19.250000   162.250000    4.250000    145.250000     3.250000
50%    37.500000   182.500000    6.000000    173.500000     6.000000
75%    55.750000   208.000000    8.000000    198.250000     8.000000
max    74.000000   277.000000   10.000000    262.000000    10.000000
```

```
# used to describe the dataset
ipl.describe()
```

```
              id  first_score  first_wkts  second_score  second_wkts
count  74.000000    74.000000   74.000000     74.000000    74.000000
mean   37.500000   180.554054    6.148649    169.054054     5.783784
```

```
std    21.505813    51.855474    2.469998     47.651386     2.934305
min     1.000000     0.000000    0.000000      0.000000     0.000000
25%    19.250000   162.250000    4.250000    145.250000     3.250000
50%    37.500000   182.500000    6.000000    173.500000     6.000000
75%    55.750000   208.000000    8.000000    198.250000     8.000000
max    74.000000   277.000000   10.000000    262.000000    10.000000
```

```
# Used to check null value in data set
ipl.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 74 entries, 0 to 73
Data columns (total 14 columns):
 #   Column               Non-Null Count  Dtype 
---  ------               --------------  ----- 
 0   id                   74 non-null     int64 
 1   date                 74 non-null     object
 2   team1                74 non-null     object
 3   team2                74 non-null     object
 4   toss_winner          74 non-null     object
 5   decision             71 non-null     object
 6   first_score          74 non-null     int64 
 7   first_wkts           74 non-null     int64 
 8   second_score         74 non-null     int64 
 9   second_wkts          74 non-null     int64 
 10  winner               74 non-null     object
 11  player_of_the_match  71 non-null     object
 12  most_runs            71 non-null     object
 13  most_wkts            71 non-null     object
dtypes: int64(5), object(9)
memory usage: 5.6+ KB
```

```
# Used to count total number of rows and column
ipl.shape
```

```
(74, 14)
```

```
ipl['most_runs'].head()
```

```
0       Anuj Rawat
1       Sam Curran
2    Andre Russell
3     Sanju Samson
4    Dewald Brevis
Name: most_runs, dtype: object
```

```
# Used for finding out number of POM won by players
ipl['player_of_the_match'].value_counts()[0:8]
```

```
player_of_the_match
Travis Head           3
```

```
Abhishek Sharma       3
Sunil Narine          3
Jos Buttler           2
Rishabh Pant          2
Sam Curran            2
Varun Chakravarthy    2
Kuldeep Yadav         2
Name: count, dtype: int64
```

```
# Finding out how many times player score highest runs
ipl['most_runs'].value_counts()[0:10]
```

```
most_runs
Virat Kohli        4
Venkatesh Iyer     3
Travis Head        3
Ruturaj Gaikwad    3
Marcus Stoinis     3
Riyan Parag        3
Riyan Parag        2
Phil Salt          2
Sam Curran         2
Shubman Gill       2
Name: count, dtype: int64
```

```
# Finding out how many times player get most wickets in a match
ipl['most_wkts'].value_counts()[0:10]
```

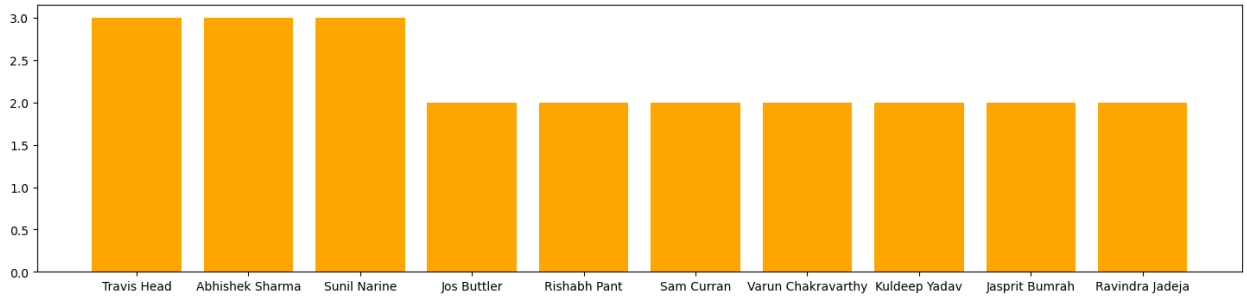
```
most_wkts
Mitchell Starc         3
T Natarajan            3
Jasprit Bumrah         3
Andre Russell          3
Mukesh Kumar           2
Ravindra Jadeja        2
Bhuvneshwar Kumar      2
Kuldeep Yadav          2
Matheesha Pathirana    2
Rahul Chahar           2
Name: count, dtype: int64
```

```
ipl['player_of_the_match'].value_counts()[0:4].keys()
```

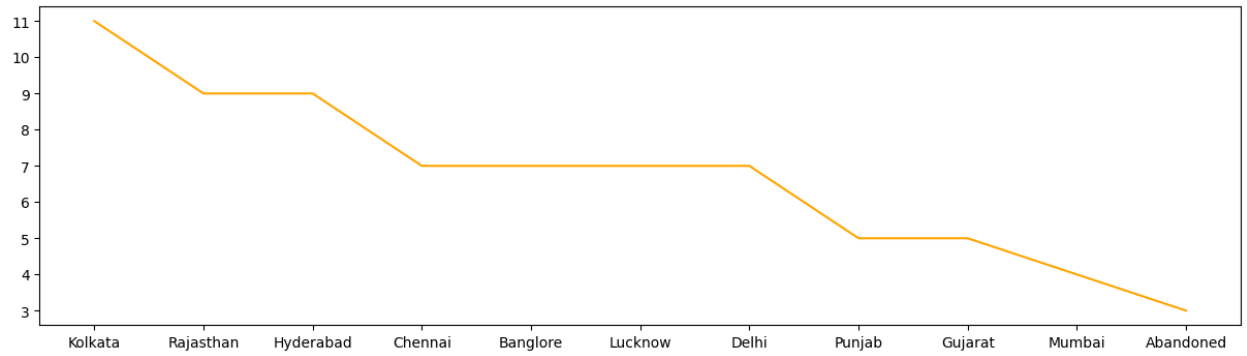
```
Index(['Travis Head', 'Abhishek Sharma', 'Sunil Narine', 'Jos
Buttler'], dtype='object', name='player_of_the_match')
```

```
# Bar plot of most player of the match award winner
plt.figure(figsize=(18,4))
plt.bar(list(ipl['player_of_the_match'].value_counts()
[0:10].keys()),list(ipl['player_of_the_match'].value_counts()[0:10]),
```

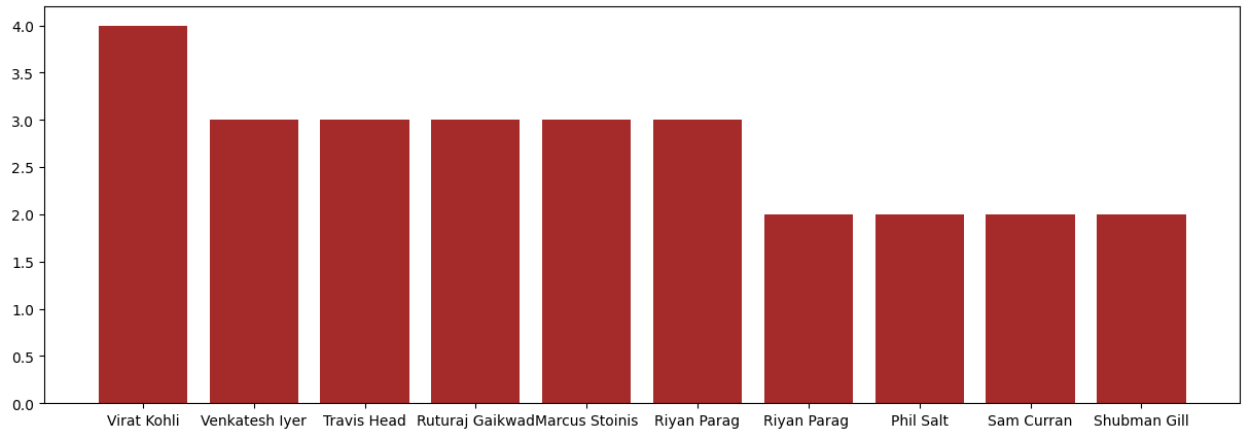
```
color='orange')
plt.show

<function matplotlib.pyplot.show(close=None, block=None)>
```




```
ipl['winner'].value_counts()[0:10]

winner
Kolkata      11
Rajasthan     9
Hyderabad     9
Chennai       7
Banglore      7
Lucknow       7
Delhi         7
Punjab        5
Gujarat       5
Mumbai        4
Name: count, dtype: int64

# plot of most match winning team
plt.figure(figsize=(15,4))
plt.plot(list(ipl['winner'].value_counts().keys()),list(ipl['winner'].
value_counts()), color='orange')
plt.show

<function matplotlib.pyplot.show(close=None, block=None)>
```

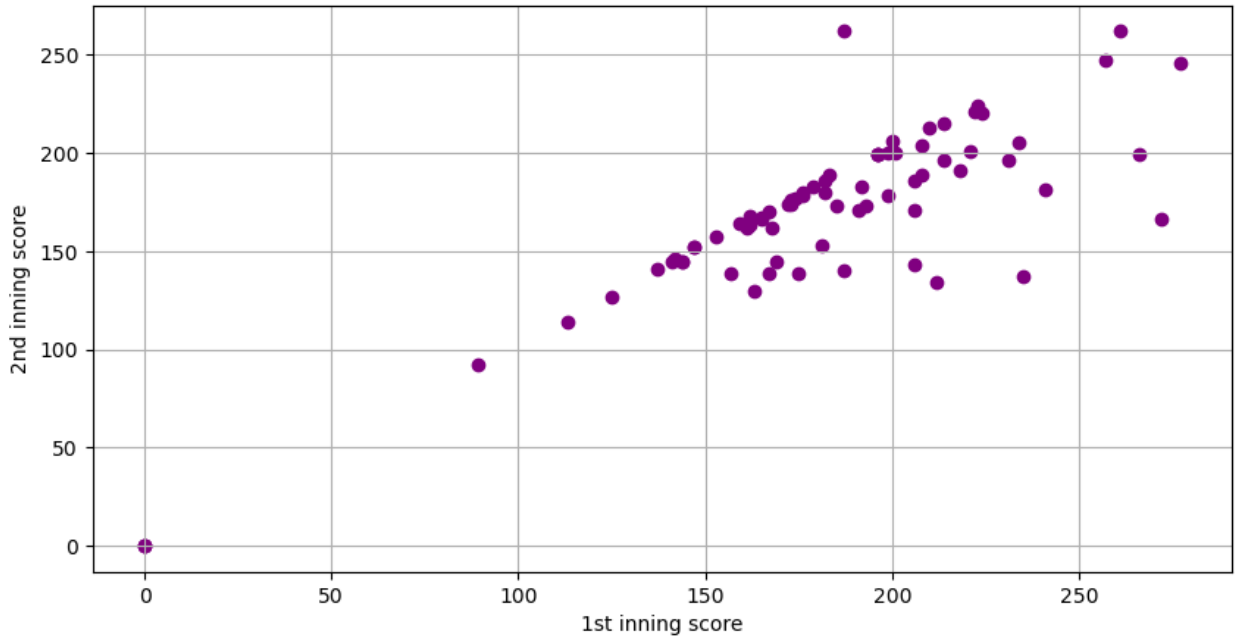


```
# Bar plot of most time highest run scorer for the team
plt.figure(figsize=(15,5))
plt.bar(list(ipl['most_runs'].value_counts()
[0:10].keys()),list(ipl['most_runs'].value_counts()[0:10]),
color='brown')
plt.show()
```



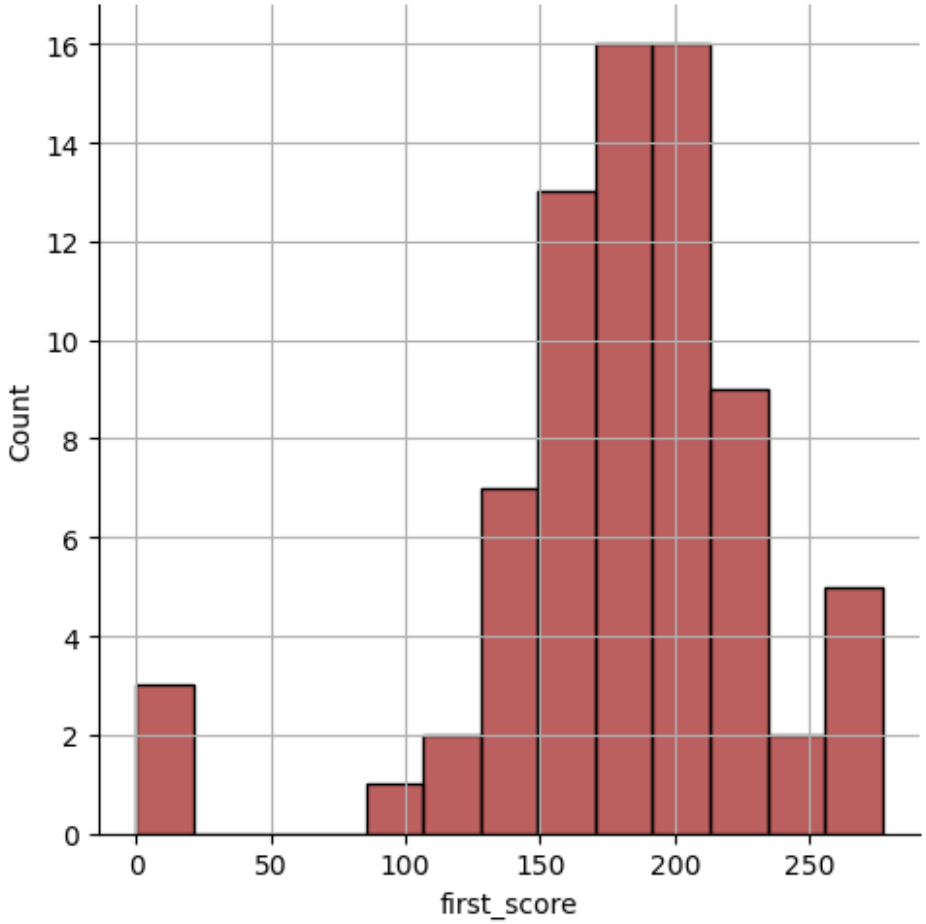

```
# Scatter plot between 1st inning score and 2nd inning score in ipl
2024
plt.figure(figsize=(10,5))
plt.scatter(x='first_score', y='second_score', data=ipl,
color='purple')
plt.ylabel('2nd inning score')
plt.xlabel('1st inning score')
plt.grid(True)
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
# Distribution plot of 1st inning scores
sns.displot(ipl['first_score'], color='brown')
plt.grid(True)
plt.show
```

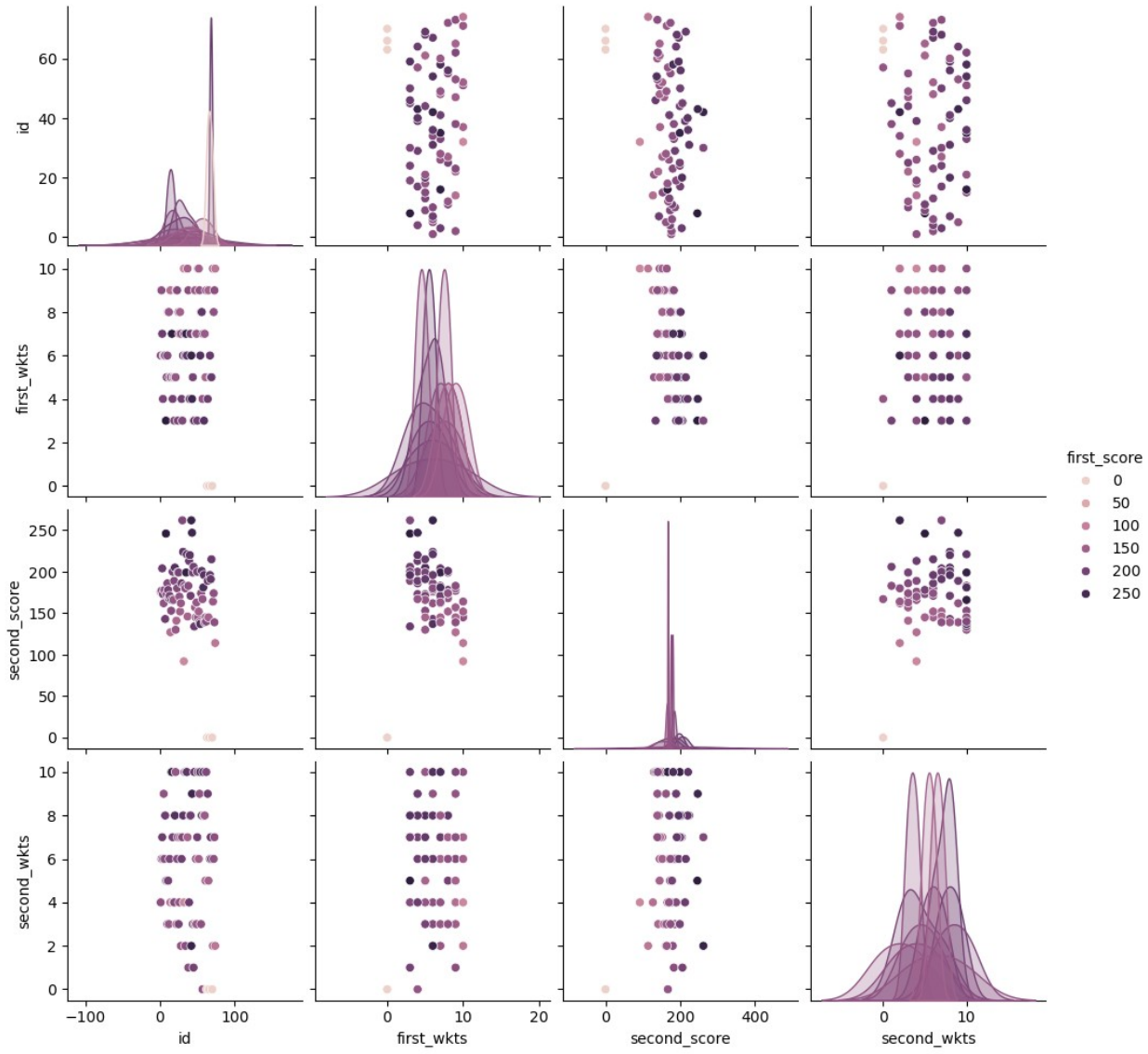
```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
# pairplot is used for visualizing relationships between multiple
variables in a dataset like 1st inning score
plt.figure(figsize=(10,5))
sns.pairplot(ipl, hue='first_score')
plt.show
```

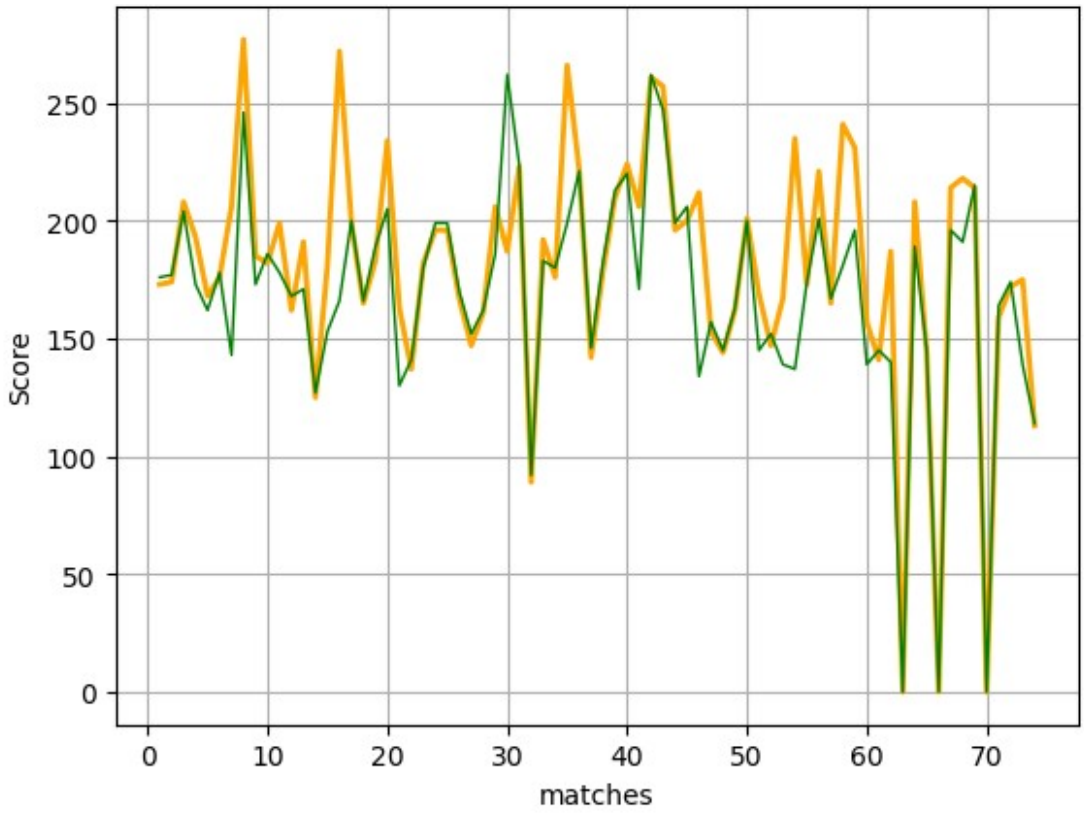
```
<function matplotlib.pyplot.show(close=None, block=None)>
```

```
<Figure size 1000x500 with 0 Axes>
```



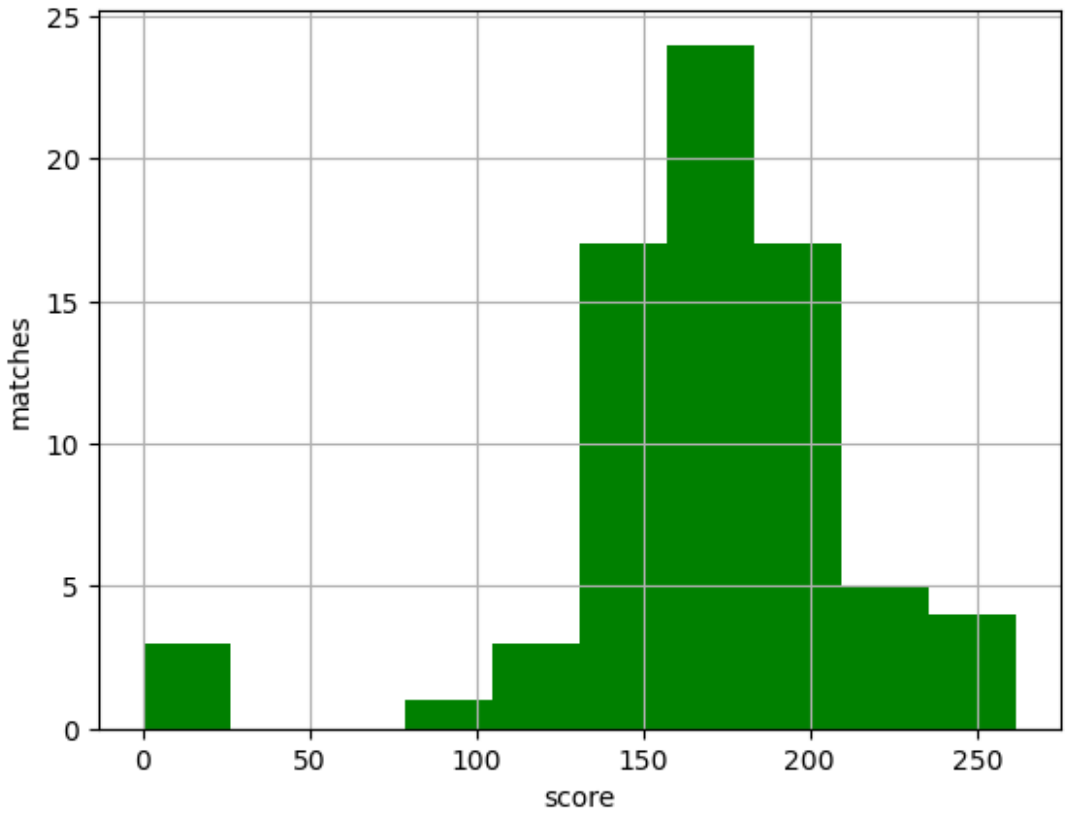

```
# Line plot between 1st inning score and 2nd inning score
plt.plot('id','first_score', data=ipl,color='orange', linewidth=2)
plt.plot('id', 'second_score',data=ipl, color='g', linewidth =1)
plt.ylabel('Score')
plt.xlabel('matches')
plt.grid(True)
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
# Histogram chart of 2nd inning score
plt.hist('second_score', data=ipl, color='g')
plt.xlabel('score')
plt.ylabel('matches')
plt.grid(True)
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
# Anaysing data when toss winner is also match winner
win=ipl[ipl['toss_winner']==ipl['winner']]
win.head()
```

```
    id           date      team1     team2 toss_winner decision
first_score  \
1    2  March 23,2024      Delhi    Punjab      Punjab    Field
174
3    4  March 24,2024  Rajasthan   Lucknow   Rajasthan      Bat
193
5    6  March 25,2024     Punjab  Banglore    Banglore    Field
176
9   10  March 29,2024   Banglore   Kolkata     Kolkata    Field
182
10  11  March 30,2024    Lucknow    Punjab     Lucknow      Bat
199

    first_wkts  second_score  second_wkts     winner
player_of_the_match  \
1            9           177            6     Punjab          Sam
Curran
3            4           173            6  Rajasthan        Sanju
Samson
5            6           178            6   Banglore         Virat
```

```
Kohli
9            6           186             3    Kolkata        Sunil
Narine
10           8           178             5    Lucknow        Mayank
Yadav

         most_runs      most_wkts
1       Sam Curran  Kuldeep Yadav
3     Sanju Samson    Trent Boult
5      Virat Kohli  Harpreet Brar
9     Virat Kohli   Andre Russell
10  Shikhar Dhawan   Mayank Yadav

win.tail()

    id         date      team1      team2 toss_winner decision
first_score  \
48  49   May 1,2024    Chennai     Punjab      Punjab    Field
162
49  50   May 2,2024  Hyderabad  Rajasthan   Hyderabad      Bat
201
51  52   May 4,2024    Gujarat   Banglore    Banglore    Field
147
54  55   May 6,2024  Hyderabad     Mumbai      Mumbai    Field
173
71  72  May 22,2024   Banglore  Rajasthan   Rajasthan    Field
172

    first_wkts  second_score  second_wkts     winner
player_of_the_match  \
48           7           163            3     Punjab        Harpreet
Brar
49           3           200            7  Hyderabad    Bhuvneshwar
Kumar
51          10           152            6   Banglore       Mohammed
Siraj
54           8           174            3     Mumbai     Suryakumar
Yadav
71           8           174            6  Rajasthan  Ravichandran
Ashwin

           most_runs          most_wkts
48   Ruturaj Gaikwad       Rahul Chahar
49      Riyan Parag   Bhuvneshwar Kumar
51    Faf du Plessis        Josh Little
54  Suryakumar Yadav      Hardik Pandya
71  Yashasvi Jaiswal         Avesh Khan

win.shape
```

```
(31, 14)
```

```
win.describe()
```

```
              id  first_score  first_wkts  second_score  second_wkts
count  31.000000    31.000000   31.000000     31.000000    31.000000
mean   29.387097   181.258065    6.322581    175.774194     5.032258
std    17.248531    36.075816    2.103760     31.907898     2.330721
min     2.000000    89.000000    3.000000     92.000000     1.000000
25%    16.500000   162.500000    5.000000    162.500000     3.000000
50%    27.000000   182.000000    6.000000    174.000000     5.000000
75%    43.000000   199.000000    8.000000    199.000000     6.500000
max    72.000000   272.000000   10.000000    262.000000    10.000000
```

```
win.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 31 entries, 1 to 71
Data columns (total 14 columns):
 #   Column               Non-Null Count  Dtype 
---  ------               --------------  ----- 
 0   id                   31 non-null     int64 
 1   date                 31 non-null     object
 2   team1                31 non-null     object
 3   team2                31 non-null     object
 4   toss_winner          31 non-null     object
 5   decision             31 non-null     object
 6   first_score          31 non-null     int64 
 7   first_wkts           31 non-null     int64 
 8   second_score         31 non-null     int64 
 9   second_wkts          31 non-null     int64 
 10  winner               31 non-null     object
 11  player_of_the_match  31 non-null     object
 12  most_runs            31 non-null     object
 13  most_wkts            31 non-null     object
dtypes: int64(5), object(9)
memory usage: 2.5+ KB
```

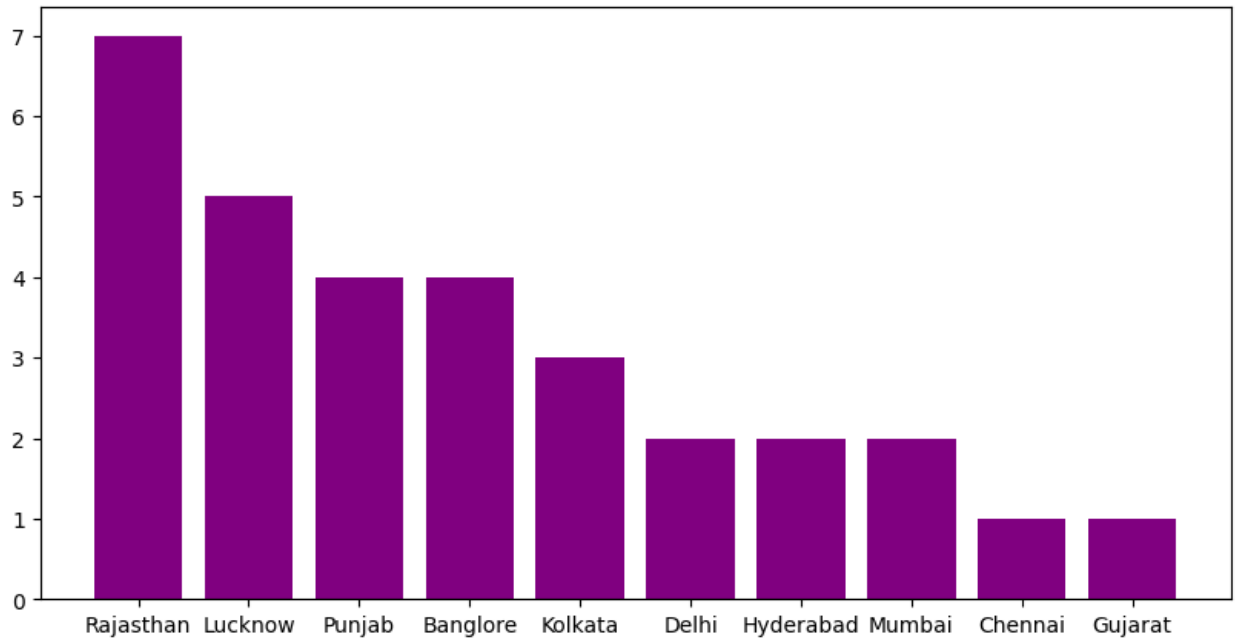
```
# Finding out number of times when a team win the toss also win the match
win['winner'].value_counts()
```

```
winner
Rajasthan    7
Lucknow      5
Punjab       4
Banglore     4
Kolkata      3
Delhi        2
Hyderabad    2
Mumbai       2
```

```
Chennai      1
Gujarat      1
Name: count, dtype: int64
```

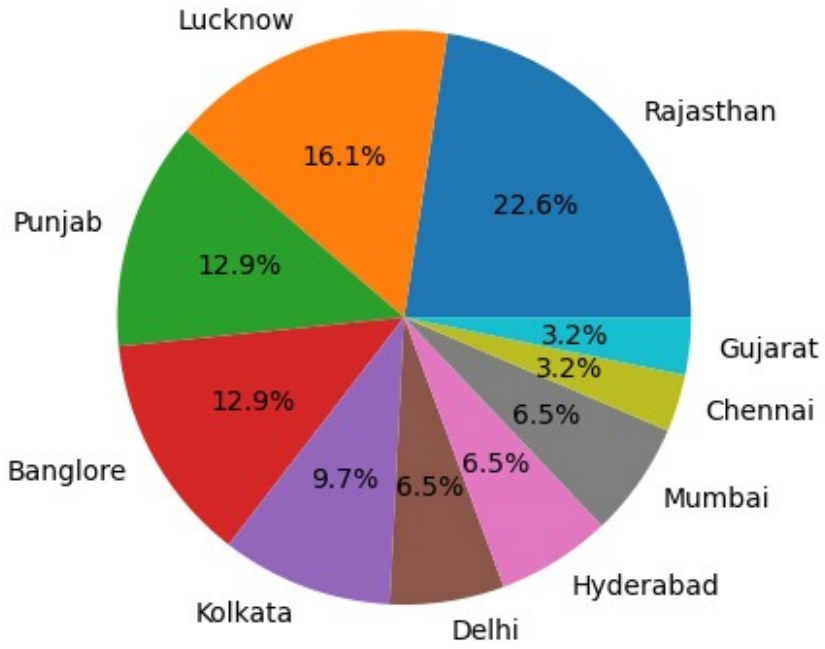
```
#  Bar graph of teams who win toss as well as match
plt.figure(figsize=(10,5))
plt.bar(list(win['winner'].value_counts().keys()),
list(win['winner'].value_counts()), color='purple')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```



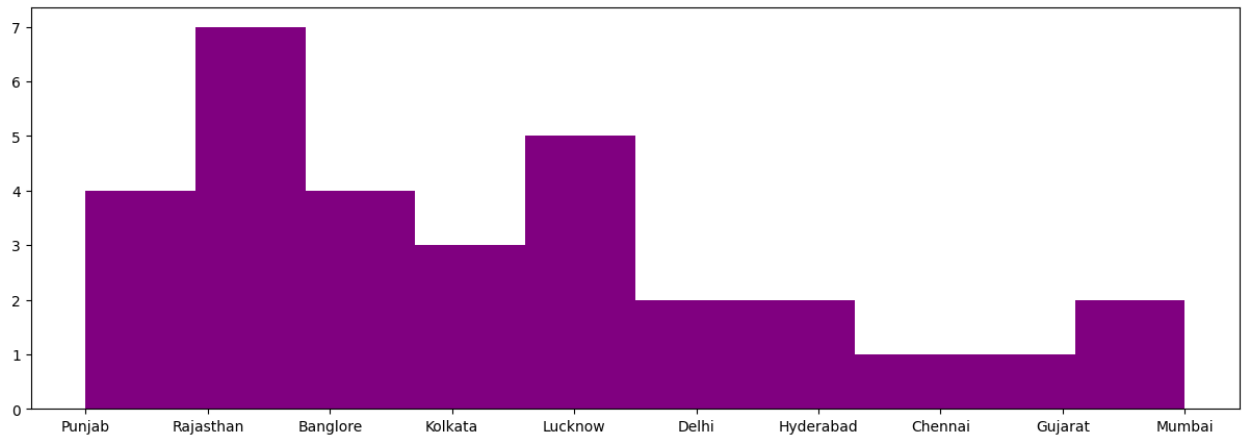

```
# Pie chart of teams along with win percentage of winner of toss and
match
plt.pie(list(win['winner'].value_counts()),labels =
list(win['winner'].value_counts().keys()), autopct='%0.1f%%')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
#  histogram plot of teams who win toss as well as match
plt.figure(figsize=(15,5))
plt.hist(win['winner'], color='purple')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```



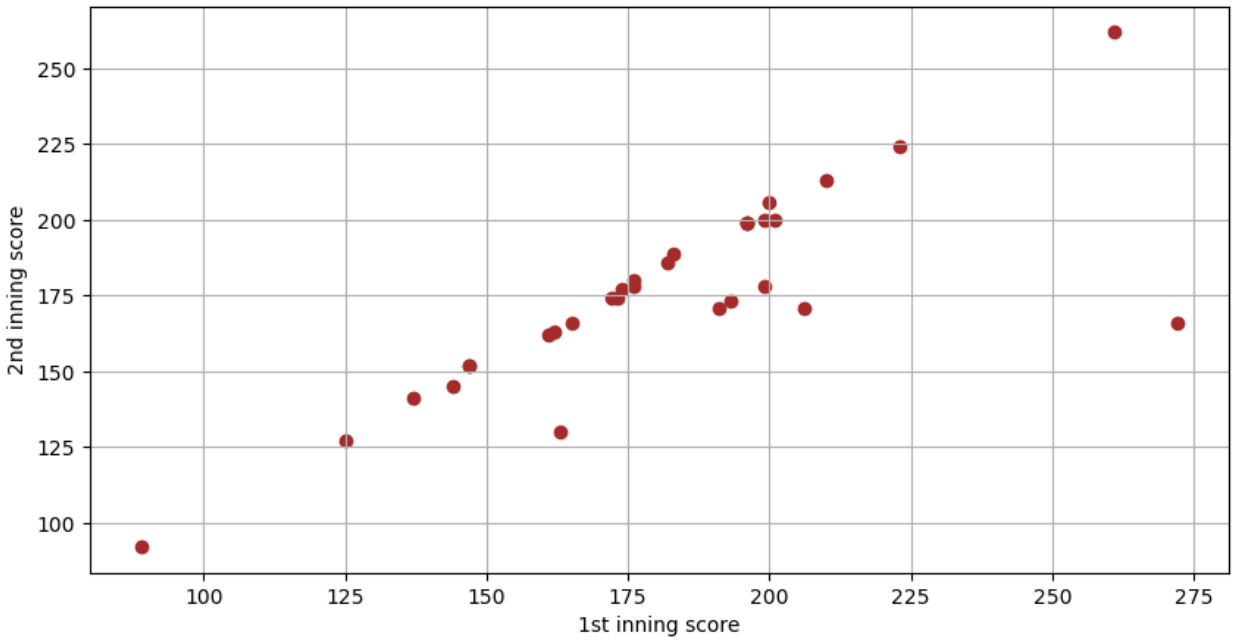

```
# Finding out most time run scorer in winning the toss and winning the match
win['most_runs'].value_counts()[0:5]
```

```
most_runs
Virat Kohli        3
```

```
Marcus Stoinis     3
Riyan Parag        2
KL Rahul           2
Ruturaj Gaikwad    2
Name: count, dtype: int64
```

```
# Scatter plot between 1st inning score and 2nd inning score of
matches in which teams win toss along with also win match
plt.figure(figsize=(10,5))
plt.scatter('first_score', 'second_score', data=win, color='brown')
plt.grid(True)
plt.ylabel('2nd inning score')
plt.xlabel('1st inning score')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
# Finding out most time wicket taker in winning the toss and winning
the match
win['most_wkts'].value_counts()[0:5]
```

```
most_wkts
Mukesh Kumar     2
Hardik Pandya    2
Kuldeep Yadav    1
Kagiso Rabada    1
Josh Little      1
Name: count, dtype: int64
```

```
# Analysing the data when toss winner in not the winner of the match
loss=ipl[ipl['toss_winner']!=ipl['winner']]
loss.head()
```

```
   id           date      team1      team2 toss_winner decision
first_score  \
0   1  March 22,2024   Banglore    Chennai    Banglore      Bat
173
2   3  March 23,2024    Kolkata  Hyderabad   Hyderabad    Field
208
4   5  March 24,2024    Gujarat     Mumbai      Mumbai    Field
168
6   7  March 26,2024    Chennai    Gujarat     Gujarat    Field
206
7   8  March 27,2024  Hyderabad     Mumbai      Mumbai    Field
277

   first_wkts  second_score  second_wkts     winner
player_of_the_match  \
0           6           176            4    Chennai   Mustafizur
Rahman
2           7           204            7    Kolkata       Andre
Russell
4           6           162            9    Gujarat       Sai
Sudharsan
6           6           143            8    Chennai         Shivam
Dube
7           3           246            5  Hyderabad     Abhishek
Sharma

          most_runs          most_wkts
0        Anuj Rawat  Mustafizur Rahman
2     Andre Russell        T Natarajan
4     Dewald Brevis     Jasprit Bumrah
6       Shivam Dube   Tushar Deshpande
7  Heinrich Klaasen        Pat Cummins
```

```
loss.tail()
```

```
    id         date      team1      team2 toss_winner decision
first_score  \
68  69  May 19,2024     Punjab  Hyderabad      Punjab      Bat
214
69  70  May 19,2024  Rajasthan    Kolkata     Kolkata      NaN
0
70  71  May 21,2024  Hyderabad    Kolkata   Hyderabad      Bat
159
72  73  May 24,2024  Hyderabad  Rajasthan   Rajasthan    Field
175
73  74  May 26,2024  Hyderabad    Kolkata   Hyderabad      Bat
```

```
113

    first_wkts  second_score  second_wkts     winner
player_of_the_match  \
68           5           215            6  Hyderabad     Abhishek
Sharma
69           0             0            0  Abandoned
NaN
70          10           164            2    Kolkata      Mitchell
Starc
72           9           139            7  Hyderabad       Shahbaz
Ahmed
73          10           114            2    Kolkata      Mitchell
Starc

            most_runs       most_wkts
68  Prabhsimran Singh     T Natarajan
69                NaN             NaN
70       Shreyas Iyer  Mitchell Starc
72       Dhruv Jurel    Shahbaz Ahmed
73     Venkatesh Iyer   Andre Russell

loss.shape

(43, 14)

loss.describe()

               id  first_score  first_wkts  second_score  second_wkts
count  43.000000    43.000000   43.000000     43.000000    43.000000
mean   43.348837   180.046512    6.023256    164.209302     6.325581
std    22.533450    61.183936    2.721067     56.231900     3.219886
min     1.000000     0.000000    0.000000      0.000000     0.000000
25%    27.500000   163.500000    4.000000    144.000000     4.500000
50%    47.000000   185.000000    6.000000    170.000000     7.000000
75%    62.500000   216.000000    8.500000    196.000000     9.000000
max    74.000000   277.000000   10.000000    262.000000    10.000000

loss.info()

<class 'pandas.core.frame.DataFrame'>
Index: 43 entries, 0 to 73
Data columns (total 14 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   id                   43 non-null     int64
 1   date                 43 non-null     object
 2   team1                43 non-null     object
 3   team2                43 non-null     object
 4   toss_winner          43 non-null     object
 5   decision             40 non-null     object
```

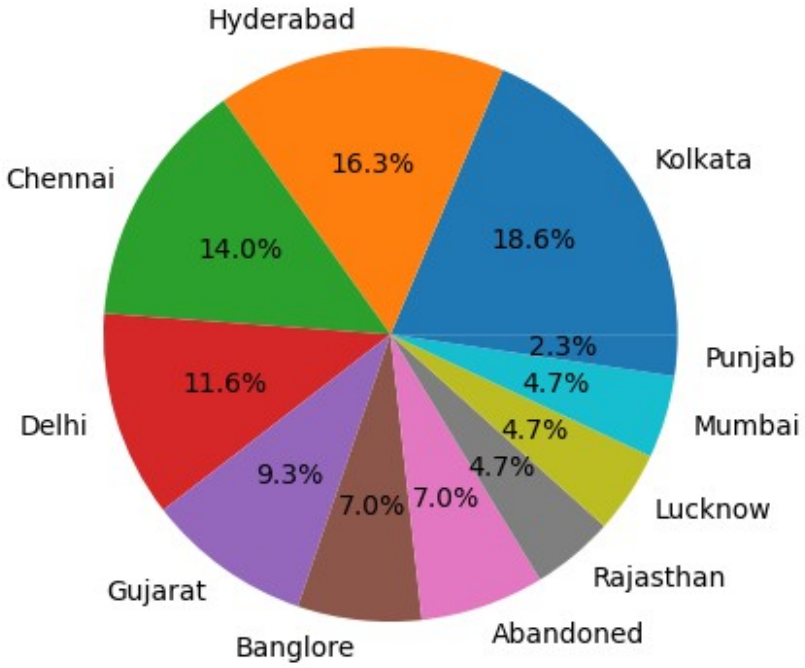
```
 6   first_score          43 non-null     int64
 7   first_wkts           43 non-null     int64
 8   second_score         43 non-null     int64
 9   second_wkts          43 non-null     int64
 10  winner               43 non-null     object
 11  player_of_the_match  40 non-null     object
 12  most_runs            40 non-null     object
 13  most_wkts            40 non-null     object
dtypes: int64(5), object(9)
memory usage: 3.5+ KB
```

```
# Finding out number times when teams win the match but not win the
toss
loss['winner'].value_counts()
```

```
winner
Kolkata      8
Hyderabad    7
Chennai      6
Delhi        5
Gujarat      4
Banglore     3
Abandoned    3
Rajasthan    2
Lucknow      2
Mumbai       2
Punjab       1
Name: count, dtype: int64
```

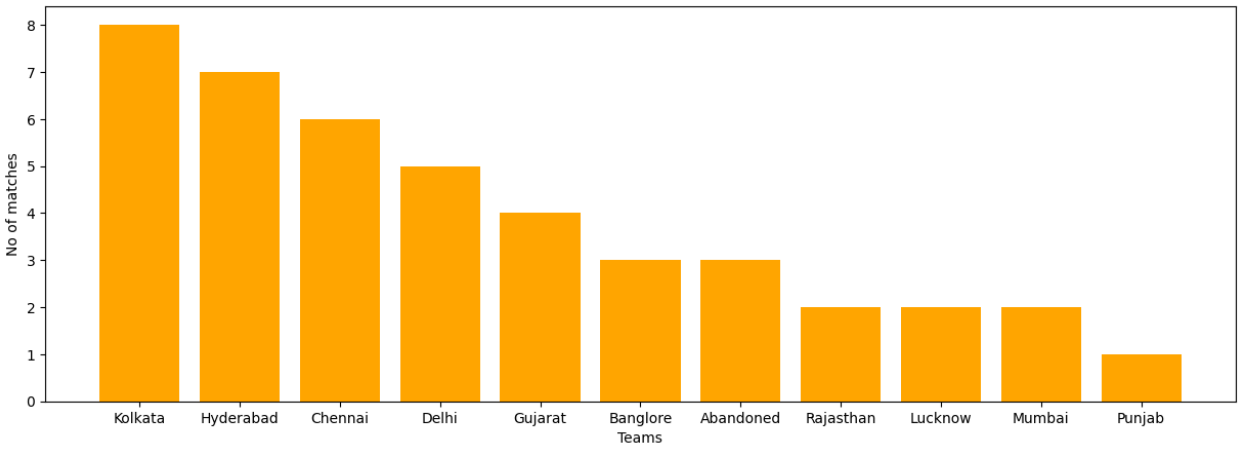
```
# Pie chart of teams who win the match but not win toss along with
percentage
plt.pie(list(loss['winner'].value_counts()),
labels=list(loss['winner'].value_counts().keys()), autopct='%0.1f%%')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```



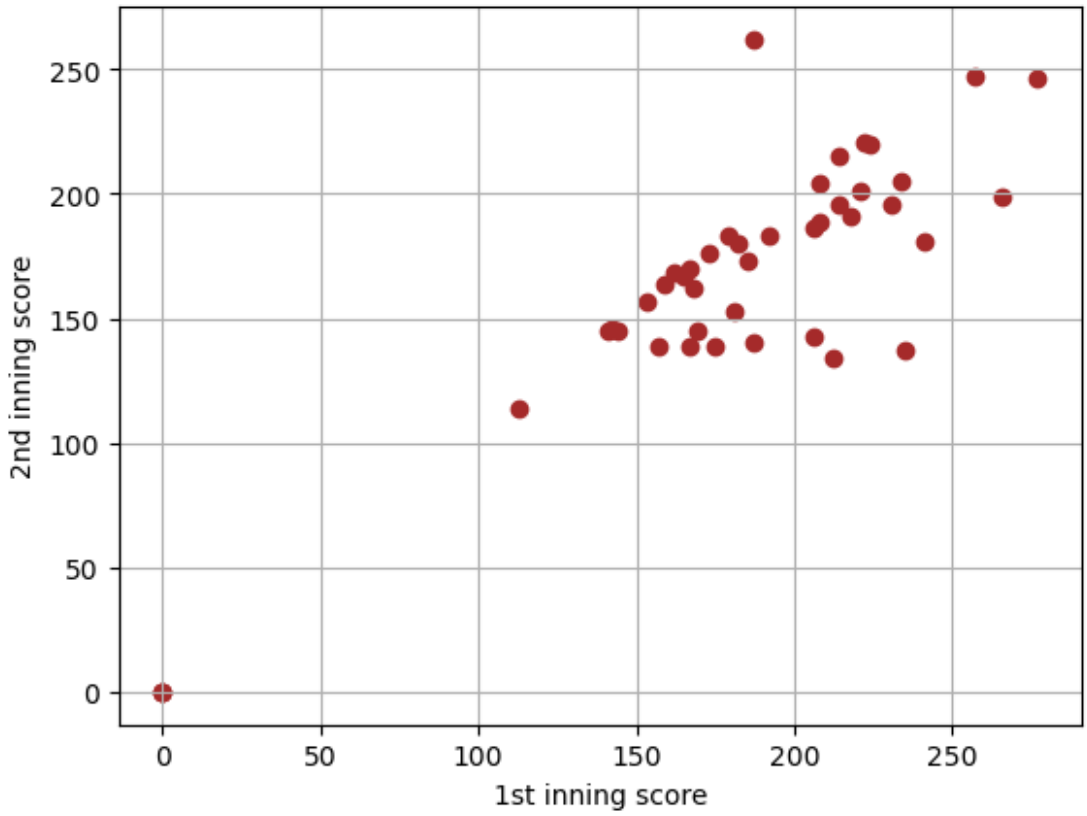

```
# Bar plot of teams winning matches who not win toss
plt.figure(figsize=(15,5))
plt.bar(list(loss['winner'].value_counts().keys()),
list(loss['winner'].value_counts()), color='orange')
# plt.grid(True)
plt.xlabel('Teams')
plt.ylabel('No of matches')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```



```
# Scatter plot between 1st inning score and 2nd inning score of teams
who win the match but not toss
plt.scatter('first_score','second_score', data=loss, color='brown')
plt.grid(True)
plt.ylabel('2nd inning score')
plt.xlabel('1st inning score')
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```




```
loss['most_runs'].value_counts()[0:5]
```

```
most_runs
Venkatesh Iyer     3
Travis Head        3
Riyan Parag        2
Nicholas Pooran    2
Anuj Rawat         1
Name: count, dtype: int64
```

```
# Total runs in 1st inning
np.sum(ipl['first_score'])
```

```
np.int64(13361)
```

```
# Total runs of 2nd inning
np.sum(ipl['second_score'])
```

```
np.int64(12510)
```

```
# Total number matches in ipl 2024
win1=np.sum(ipl['winner'].value_counts())
win1
```

```
np.int64(74)
```

```
# Total number of matches win while winning the toss
win2=np.sum(ipl['winner']==ipl['toss_winner'])
win2
```

```
np.int64(31)
```

```
# Total number of matches win while lossing toss
win3=np.sum(ipl['winner']!=ipl['toss_winner'])
win3
```

```
np.int64(43)
```

```
# Win percentage of match win while winning the toss
round((win2/win1)*100)
```

```
42
```

```
# Win percentage of match win while lossing the toss
round((win3/win1)*100)
```

```
58
```