Untitled

May 16, 2022

[272]:
```
import os
import folium
import requests
import pandas
import numpy

epl_standings = pandas.read_csv("notebooks/Final/EPL_Standings.csv")

print(epl_standings)
```

```
      Season  Pos                     Club  Pld   W   D   L  GF  GA  GD  Pts  \
0    2010-11    1       Manchester United   38  23  11   4  78  37  41   80
1    2010-11    2                 Chelsea   38  21   8   9  69  33  36   71
2    2010-11    3         Manchester City   38  21   8   9  60  33  27   71
3    2010-11    4                 Arsenal   38  19  11   8  72  43  29   68
4    2010-11    5       Tottenham Hotspur   38  16  14   8  55  46   9   62
..       …  …                       …  …  ..  ..  ..  ..  ..  ..  …
215  2020-21   16  Brighton & Hove Albion   38   9  14  15  40  46  -6   41
216  2020-21   17                 Burnley   38  10   9  19  33  55 -22   39
217  2020-21   18                  Fulham   38   5  13  20  27  53 -26   28
218  2020-21   19    West Bromwich Albion   38   5  11  22  35  76 -41   26
219  2020-21   20        Sheffield United   38   7   2  29  20  63 -43   23

                          Qualification or relegation
0    Qualification for the Champions League group s…
1    Qualification for the Champions League group s…
2    Qualification for the Champions League group s…
3    Qualification for the Champions League play-of…
4    Qualification for the Europa League play-off r…
..                                                 …
215                                  Not Applicable
216                                  Not Applicable
217               Relegation to the EFL Championship
218               Relegation to the EFL Championship
219               Relegation to the EFL Championship

[220 rows x 12 columns]
```

[273]:
```
transfer_table = pandas.read_csv("notebooks/Final/Premier_league_transfer.csv")

print(transfer_table)
```

```
                    Club  2021-22  2020-21  2019-20  2018-19  2017-18  \
0                Arsenal  -136.02   -66.85  -107.15   -71.05     9.15
1         Newcastle United  -131.50   -38.73   -37.26    -8.70   -25.28
2        Manchester United  -109.30   -64.30  -153.62   -52.15  -152.90
3           Crystal Palace   -85.62    -2.40    47.78   -11.50   -45.95
4          West Ham United   -70.27    -9.29   -64.32   -87.14    12.22
..                     …       …        …        …        …        …
15             Aston Villa    -2.82   -98.58  -156.50    -2.95    15.03
16                 Chelsea     1.95  -189.80   112.27  -125.55   -65.90
17  Brighton & Hove Albion     4.80    -7.90   -59.90   -73.50   -66.10
18                 Everton     6.50   -68.95   -33.20   -71.15   -76.82
19             Southampton    17.27   -11.00   -34.20   -36.15    37.10

    2016-17  2015-16  2014-15  2013-14  2012-13    Total
0   -102.69   -24.00   -91.18   -37.10     9.85  -617.04
1     36.63  -102.28   -21.15    22.07   -17.17  -323.36
2   -137.75   -55.33  -148.65   -75.33   -66.80 -1016.13
3    -51.00   -23.40   -28.35   -33.00    14.67  -218.77
4    -42.50   -34.19   -30.75   -23.47   -18.85  -368.54
..      …        …        …        …        …
15   -39.70    -1.85   -12.14   -11.74   -24.63  -335.88
16   -23.90    -9.01     5.11   -52.42   -84.25  -431.50
17    -8.75   -13.47     9.42     3.20    -0.67  -212.88
18   -25.20   -37.90   -38.26    14.30    -2.90  -333.58
19    16.15    -7.40    27.83   -35.40   -41.50   -67.30

[20 rows x 12 columns]
```

[286]:
```
ntransfer_table = transfer_table.drop(columns="2021-22") #Delete 21-22 season
ntransfer_table= ntransfer_table.drop(10) #Delete Brentford
clublist= ntransfer_table['Club'].tolist()

nepl_standings = epl_standings.drop(range(0,40)) ##Delete the 2010-11 and␣
 ↪2011-12 seasons

nepl_standings = nepl_standings[nepl_standings['Club'].isin(clublist)] #Take␣
 ↪out clubs that aren't in PL

nepl_standings = nepl_standings.drop(columns="W") #Delete Information not␣
 ↪needed for analysis
nepl_standings = nepl_standings.drop(columns="D")
nepl_standings = nepl_standings.drop(columns="L")
nepl_standings = nepl_standings.drop(columns="Qualification or relegation")
```

```
nepl_standings = nepl_standings.drop(columns="Pld")

nepl_standings.loc[nepl_standings['Club'] == "Arsenal", 'Club_num'] = 0##Create␣
 ↪this row to check seasons will delete later
nepl_standings.loc[nepl_standings['Club'] == "Newcastle United", 'Club_num'] = 1
nepl_standings.loc[nepl_standings['Club'] == "Manchester United", 'Club_num'] =␣
 ↪2
nepl_standings.loc[nepl_standings['Club'] == "Crystal Palace", 'Club_num'] = 3
nepl_standings.loc[nepl_standings['Club'] == "West Ham United", 'Club_num'] = 4
nepl_standings.loc[nepl_standings['Club'] == "Leicester City", 'Club_num'] = 5
nepl_standings.loc[nepl_standings['Club'] == "Tottenham Hotspur", 'Club_num'] =␣
 ↪6
nepl_standings.loc[nepl_standings['Club'] == "Leeds United", 'Club_num'] = 7
nepl_standings.loc[nepl_standings['Club'] == "Liverpool", 'Club_num'] = 8
nepl_standings.loc[nepl_standings['Club'] == "Manchester City", 'Club_num'] = 9
nepl_standings.loc[nepl_standings['Club'] == "Watford", 'Club_num'] = 11
nepl_standings.loc[nepl_standings['Club'] == "Norwich City", 'Club_num'] = 12
nepl_standings.loc[nepl_standings['Club'] == "Wolverhampton Wanderers",␣
 ↪'Club_num'] = 13
nepl_standings.loc[nepl_standings['Club'] == "Burnley", 'Club_num'] = 14
nepl_standings.loc[nepl_standings['Club'] == "Aston Villa", 'Club_num'] = 15
nepl_standings.loc[nepl_standings['Club'] == "Chelsea", 'Club_num'] = 16
nepl_standings.loc[nepl_standings['Club'] == "Brighton & Hove Albion",␣
 ↪'Club_num'] = 17
nepl_standings.loc[nepl_standings['Club'] == "Everton", 'Club_num'] = 18
nepl_standings.loc[nepl_standings['Club'] == "Southampton", 'Club_num'] = 18

nepl_standings['Club_num'] = nepl_standings['Club_num'].fillna(0).astype(int)
df= pandas.DataFrame()

for i, val in nepl_standings.iterrows():  ##adding the netspend to the dataframe
    seas= nepl_standings.at[i,'Season']
    cid = (nepl_standings.at[i,'Club_num'])
    df.at[i,'Net_Spend'] = ntransfer_table.iloc[cid][seas]

result = pandas.concat([nepl_standings, df], axis=1, join='inner')
nr = result.drop(columns= 'Club_num')
print(nr)
```

```
      Season  Pos                    Club  GF  GA  GD  Pts  Net_Spend
40   2012-13    1       Manchester United  86  43  43   89     -66.80
41   2012-13    2         Manchester City  66  34  32   78     -17.65
42   2012-13    3                 Chelsea  75  39  36   75      -0.67
43   2012-13    4                 Arsenal  72  37  35   73       9.85
44   2012-13    5       Tottenham Hotspur  66  46  20   72      -0.47
..       …  …                       …  ..  ..  ..  …        …
212  2020-21   13  Wolverhampton Wanderers  36  52 -16   45       1.20
```

```
213  2020-21   14           Crystal Palace  41  66 -25   44      -2.40
214  2020-21   15              Southampton  47  68 -21   43     -11.00
215  2020-21   16   Brighton & Hove Albion  40  46  -6   41     -68.95
216  2020-21   17                  Burnley  33  55 -22   39     -98.58

[133 rows x 8 columns]
```

[317]:
```
N = 4
# Drop first N columns of dataframe
ntr= ntransfer_table
for i, val in ntransfer_table.iterrows():
    ntr.at[i,'Total Last 5']= (ntransfer_table.at[i, '2016-17'] +␣
 ↪ntransfer_table.at[i, '2017-18'] + ntransfer_table.at[i, '2018-19'] +␣
 ↪ntransfer_table.at[i, '2019-20'] + ntransfer_table.at[i, '2020-21'])

for i, val in ntransfer_table.iterrows():
    club= ntransfer_table.at[i, 'Club']
    temp= nr[nr['Club']== club]
    ntr.at[i, 'Avg. Pos. Last 10'] = temp['Pos'].mean() ##Average position last␣
 ↪10 years
    ntr.at[i, 'Avg. Pts. Last 10'] = temp['Pts'].mean()
    ntr.at[i, 'Avg. GF Last 10'] = temp['GF'].mean()
    ntr.at[i, 'Avg. GA Last 10'] = temp['GA'].mean()
    temp = temp[temp.Season != '2012-13']
    temp = temp[temp.Season != '2013-14']
    temp = temp[temp.Season != '2014-15']
    temp = temp[temp.Season != '2015-16']
    ntr.at[i, 'Avg. Pos. Last 5'] = temp['Pos'].mean() ##Average position last␣
 ↪5 years
    ntr.at[i, 'Avg. Pts. Last 5'] = temp['Pts'].mean()
    ntr.at[i, 'Avg. GF Last 5'] = temp['GF'].mean()
    ntr.at[i, 'Avg. GA Last 5'] = temp['GA'].mean()

pandas.set_option("display.max_rows", 10, "display.max_columns", None)

ntr['Total abs 10']= ntr['Total']*-1
ntr['Total abs 5']= ntr['Total Last 5']*-1
print(ntr)
```

```
                    Club  2020-21  2019-20  2018-19  2017-18  2016-17  \
0                Arsenal   -66.85  -107.15   -71.05     9.15  -102.69
1         Newcastle United   -38.73   -37.26    -8.70   -25.28    36.63
2        Manchester United   -64.30  -153.62   -52.15  -152.90  -137.75
3           Crystal Palace    -2.40    47.78   -11.50   -45.95   -51.00
4          West Ham United    -9.29   -64.32   -87.14    12.22   -42.50
..                     …       …        …        …        …        …
15            Aston Villa   -98.58  -156.50    -2.95    15.03   -39.70
```

```
16                 Chelsea  -189.80   112.27  -125.55   -65.90   -23.90
17  Brighton & Hove Albion    -7.90   -59.90   -73.50   -66.10    -8.75
18                 Everton   -68.95   -33.20   -71.15   -76.82   -25.20
19              Southampton   -11.00   -34.20   -36.15    37.10    16.15

    2015-16  2014-15  2013-14  2012-13    Total  Total Last 5  \
0    -24.00   -91.18   -37.10     9.85  -617.04       -338.59
1   -102.28   -21.15    22.07   -17.17  -323.36        -73.34
2    -55.33  -148.65   -75.33   -66.80 -1016.13       -560.72
3    -23.40   -28.35   -33.00    14.67  -218.77        -63.07
4    -34.19   -30.75   -23.47   -18.85  -368.54       -191.03
..      …        …        …        …        …            …
15    -1.85   -12.14   -11.74   -24.63  -335.88       -282.70
16    -9.01     5.11   -52.42   -84.25  -431.50       -292.88
17   -13.47     9.42     3.20    -0.67  -212.88       -216.15
18   -37.90   -38.26    14.30    -2.90  -333.58       -275.32
19    -7.40    27.83   -35.40   -41.50   -67.30        -28.10

    Avg. Pos. Last 10  Avg. Pts. Last 10  Avg. GF Last 10  Avg. GA Last 10  \
0            5.000000          69.222222        67.888889        42.555556
1           13.375000          43.000000        42.125000        58.750000
2            4.000000          71.666667        65.222222        38.777778
3           12.625000          44.500000        42.125000        54.625000
4           10.888889          48.666667        50.222222        55.333333
..                ...                ...              ...              ...
15          15.833333          37.333333        40.000000        62.666667
16           3.777778          73.555556        68.333333        39.000000
17          15.750000          39.500000        37.000000        53.500000
18           8.666667          55.666667        52.666667        48.444444
19          11.333333          48.444444        48.555556        53.000000

    Avg. Pos. Last 5  Avg. Pts. Last 5  Avg. GF Last 5  Avg. GA Last 5  \
0               6.40              65.0           67.00           46.60
1              12.00              44.5           41.25           53.75
2               3.80              71.2           65.20           38.20
3              13.00              44.2           43.60           57.40
4              11.20              48.6           51.60           59.20
..               …                 …               …               …
15             14.00              45.0           48.00           56.50
16              3.40              73.6           67.40           40.00
17             15.75              39.5           37.00           53.50
18              9.00              54.4           50.20           50.40
19             13.40              43.2           44.20           59.40

    Total abs 10  Total abs 5
0         617.04       338.59
1         323.36        73.34
2        1016.13       560.72
```

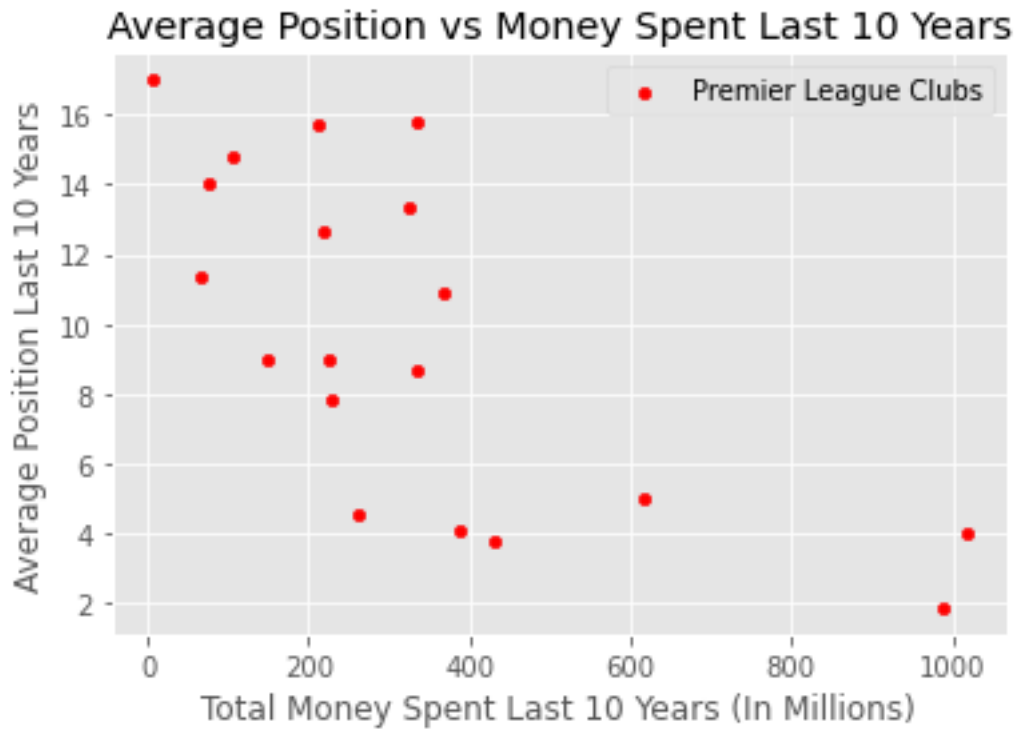
```
3         218.77         63.07
4         368.54        191.03
..           …            …
15        335.88        282.70
16        431.50        292.88
17        212.88        216.15
18        333.58        275.32
19         67.30         28.10

[19 rows x 22 columns]
```

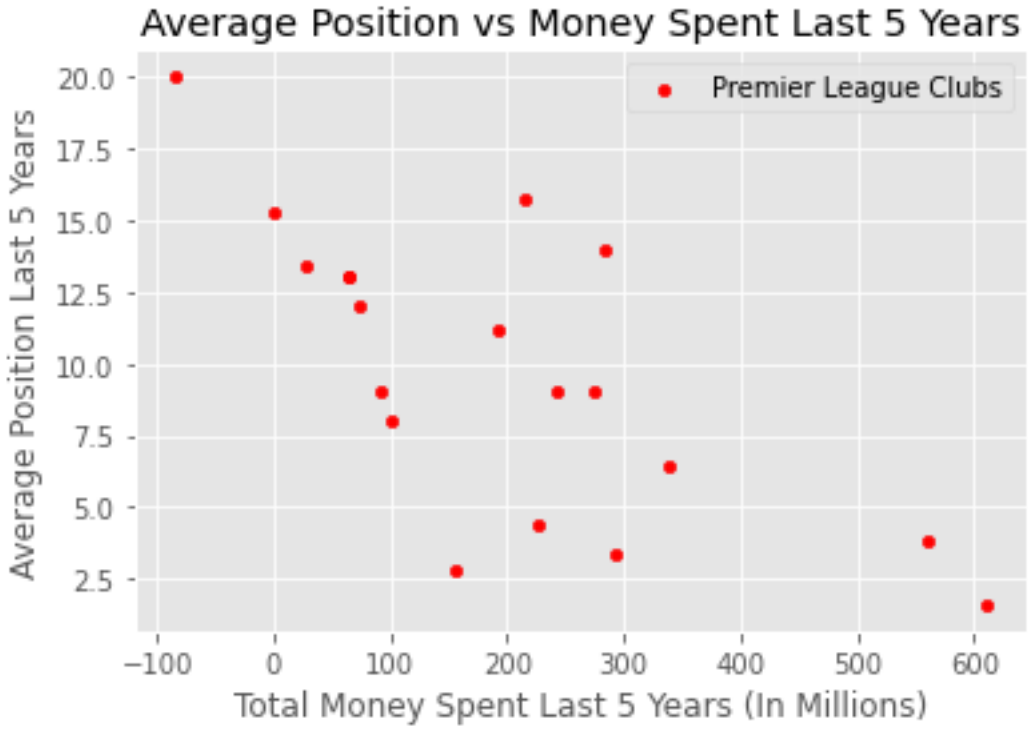
[316]:
```
analysis= pandas.DataFrame()
analysis['Club']= ntr['Club']
analysis['PPMS5']= ntr['Avg. Pos. Last 5']/ntr['Total Last 5']
analysis['PPMS10']= ntr['Avg. Pos. Last 10']/ntr['Total']
analysis['GPMS5'] =ntr['Avg. GF Last 5']/ntr['Total Last 5']
analysis['GPMS10'] =ntr['Avg. GF Last 10']/ntr['Total']
analysis['GAPMS5'] =ntr['Avg. GA Last 5']/ntr['Total Last 5']
analysis['GAPMS10'] =ntr['Avg. GA Last 10']/ntr['Total']
analysis['PtPMS5'] =ntr['Avg. Pts. Last 5']/ntr['Total Last 5']
analysis['PtPMS10'] =ntr['Avg. Pts. Last 10']/ntr['Total']
pandas.set_option("display.max_rows", None, "display.max_columns", None)

print(analysis)
```

```
                       Club      PPMS5    PPMS10      GPMS5    GPMS10  \
0                   Arsenal  -0.018902 -0.008103  -0.197879 -0.110023
1            Newcastle United  -0.163621 -0.041363  -0.562449 -0.130273
2          Manchester United  -0.006777 -0.003937  -0.116279 -0.064187
3             Crystal Palace  -0.206120 -0.057709  -0.691295 -0.192554
4            West Ham United  -0.058630 -0.029546  -0.270115 -0.136273
5             Leicester City  -0.079976 -0.034461  -0.579826 -0.253133
6          Tottenham Hotspur  -0.019405 -0.017439  -0.314002 -0.256904
7               Leeds United  -0.097371 -0.060419  -0.670778 -0.416219
8                  Liverpool  -0.017934 -0.010610  -0.517517 -0.198141
9            Manchester City  -0.002619 -0.001914  -0.152547 -0.088705
11                   Watford -17.528736 -0.138369 -49.425287 -0.396410
12              Norwich City   0.235239 -2.603369   0.305810 -5.130168
13  Wolverhampton Wanderers  -0.037182 -0.039991  -0.184535 -0.198474
14                   Burnley  -0.202366 -0.182268  -0.610212 -0.486048
15              Aston Villa  -0.049522 -0.047140  -0.169791 -0.119090
16                   Chelsea  -0.011609 -0.008755  -0.230128 -0.158362
17   Brighton & Hove Albion  -0.072866 -0.073985  -0.171177 -0.173807
18                   Everton  -0.032689 -0.025981  -0.182333 -0.157883
19               Southampton  -0.476868 -0.168400  -1.572954 -0.721479

       GAPMS5    GAPMS10     PtPMS5   PtPMS10
0   -0.137630  -0.068967  -0.191973 -0.112184
```

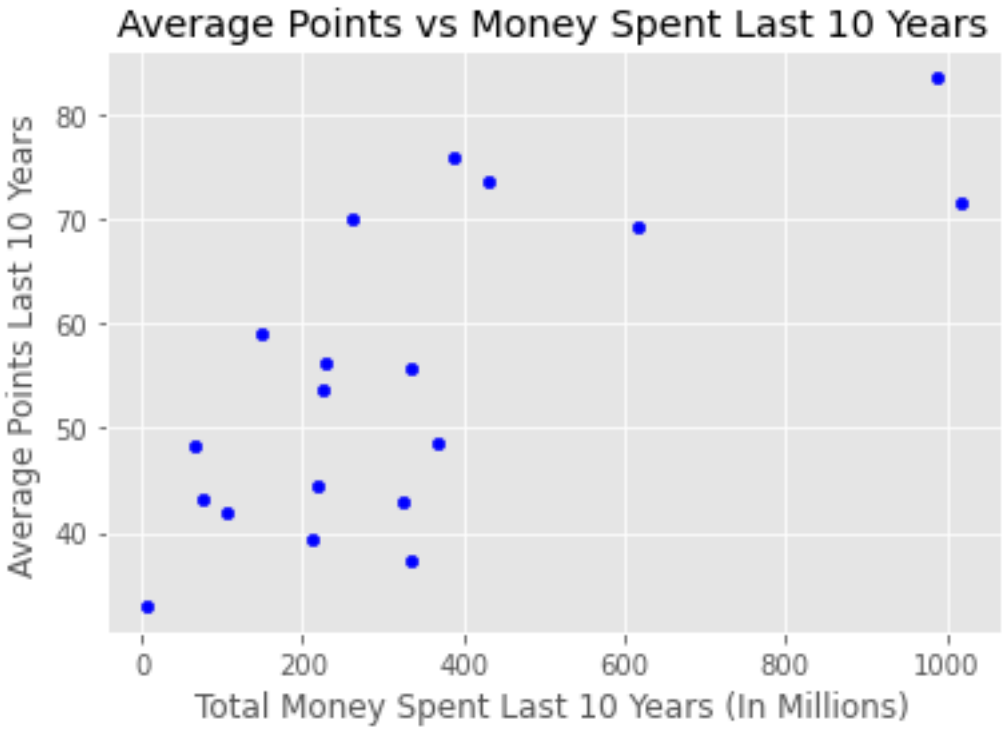
```
1   -0.732888  -0.181686  -0.606763 -0.132979
2   -0.068127  -0.038162  -0.126980 -0.070529
3   -0.910100  -0.249691  -0.700809 -0.203410
4   -0.309899  -0.150142  -0.254410 -0.132053
5   -0.523843  -0.221178  -0.541837 -0.246241
6   -0.170232  -0.160778  -0.313120 -0.267963
7   -0.584226  -0.362513  -0.638321 -0.396079
8   -0.226734  -0.105522  -0.532889 -0.195847
9   -0.051067  -0.034446  -0.145018 -0.084653
11 -73.275862  -0.570307 -47.413793 -0.392670
12   0.882145 -10.030628   0.247001 -5.053599
13  -0.190043  -0.204399  -0.221717 -0.238466
14  -0.831258  -0.694354  -0.706725 -0.564163
15  -0.199859  -0.186575  -0.159179 -0.111151
16  -0.136575  -0.090382  -0.251297 -0.170465
17  -0.247513  -0.251315  -0.182743 -0.185551
18  -0.183060  -0.145226  -0.197588 -0.166877
19  -2.113879  -0.787519  -1.537367 -0.719828
```

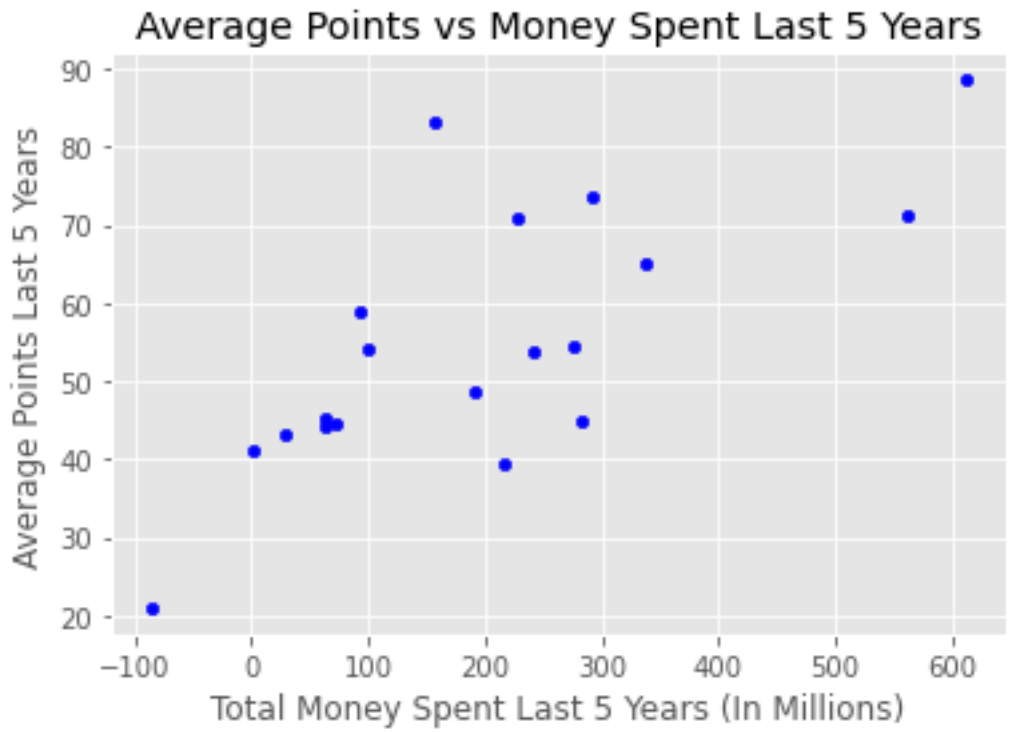
[309]:
```
import matplotlib.pyplot as plt
import matplotlib
import seaborn
matplotlib.style.use('ggplot')

ax = ntr.plot('Total abs 10', 'Avg. Pos. Last 10', kind='scatter', c='r',␣
 ↪label='Premier League Clubs')
plt.title('Average Position vs Money Spent Last 10 Years')
plt.xlabel('Total Money Spent Last 10 Years (In Millions)')
plt.ylabel('Average Position Last 10 Years')
plt.show()
bx = ntr.plot('Total abs 5', 'Avg. Pos. Last 5', kind='scatter', c='r',␣
 ↪label='Premier League Clubs')
plt.title('Average Position vs Money Spent Last 5 Years')
plt.xlabel('Total Money Spent Last 5 Years (In Millions)')
plt.ylabel('Average Position Last 5 Years')
plt.show()
```


Average Position vs Money Spent Last 10 Years
Premier League Clubs
Average Position Last 10 Years
Total Money Spent Last 10 Years (In Millions)
16
14
12
10
8
6
4
2
0
200
400
600
800
1000


Average Position vs Money Spent Last 5 Years
Premier League Clubs
Average Position Last 5 Years
Total Money Spent Last 5 Years (In Millions)
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
−100
0
100
200
300
400
500
600

[314]:
```
cx = ntr.plot('Total abs 10', 'Avg. Pts. Last 10', kind='scatter', c='b')
plt.title('Average Points vs Money Spent Last 10 Years')
plt.xlabel('Total Money Spent Last 10 Years (In Millions)')
plt.ylabel('Average Points Last 10 Years')
plt.show()
dx = ntr.plot('Total abs 5', 'Avg. Pts. Last 5', kind='scatter', c='b')
plt.title('Average Points vs Money Spent Last 5 Years')
plt.xlabel('Total Money Spent Last 5 Years (In Millions)')
plt.ylabel('Average Points Last 5 Years')
plt.show()
```




Average Points vs Money Spent Last 5 Years
Average Points Last 5 Years
Total Money Spent Last 5 Years (In Millions)
90
80
70
60
50
40
30
20
−100
0
100
200
300
400
500
600

[ ]: