Trend Analysis of Texas Educational Data

January 6, 2021

```python
[1]: #Import libraries
     import numpy as np
     import pandas as pd
     import matplotlib
     import matplotlib.pyplot as plt
     import seaborn as sns
     import requests
     import xlrd
     import scipy
```

```python
[2]: #Read data into dataframe
     df = pd.read_excel('Safal_Performance Exercise.xlsx')
```

```python
[3]: #get list of column names in file
     df.columns
```

```
[3]: Index(['CAMPUS', 'CAD_GAP', 'CAD_MATH', 'CAD_POST', 'CAD_PROGRESS', 'CAD_READ',
            'CAD_SCIE', 'CAD_SOCI', 'CAMPNAME', 'CFLAEATYPE', 'CFLAEC', 'CFLALTED',
            'CFLCHART', 'CFLDAEP', 'CFLEEK', 'CFLJJ', 'CFLNEWCAMP', 'CFLPAIR',
            'CFLPAIRD', 'CFLPAIRO', 'CFLRTF', 'CI1', 'CI1_CUT', 'CI1_MAXPTS',
            'CI1_MET', 'CI1_TOTPTS', 'CI2', 'CI2_CUT', 'CI2_MAXPTS', 'CI2_MET',
            'CI2_TOTPTS', 'CI3', 'CI3_CUT', 'CI3_MAXPTS', 'CI3_MET', 'CI3_TOTPTS',
            'CI4', 'CI4_CUT', 'CI4_GRD_WGT', 'CI4_MET', 'CI4_PSG_WGT',
            'CI4_RHS_WGT', 'CI4_STR_WGT', 'CNTYNAME', 'COUNTY', 'CPEMALLC',
            'CPEMALLP', 'CPEMALLT', 'CPETALLC', 'CPETECHC', 'CPETECHP', 'CPETECOC',
            'CPETECOP', 'CPETLEPC', 'CPETLEPP', 'CPETSPEC', 'CPETSPEP', 'C_RATING',
            'C_UPDATE', 'C_YRS_IR', 'DISTNAME', 'DISTRICT', 'GRDHIGH', 'GRDLOW',
            'GRDSPAN', 'GRDTYPE', 'PAIRCAMP', 'PAIRNAME', 'REGION', 'REGNNAME'],
           dtype='object')
```

```python
[4]: #Evaluate correlation coefficients among variables
     df.corr()
```

```
[4]:                CAD_GAP  CAD_MATH  CAD_POST  CAD_PROGRESS  CAD_READ  CAD_SCIE  \
     CAD_GAP       1.000000  0.430179  0.549991      0.364545  0.443659  0.389964
     CAD_MATH      0.430179  1.000000  0.508179      0.444649  0.452456  0.369247
     CAD_POST      0.549991  0.508179  1.000000      0.334827  0.541109  0.407885
     CAD_PROGRESS  0.364545  0.444649  0.334827      1.000000  0.357205  0.246878
```

```
CAD_READ       0.443659  0.452456  0.541109     0.357205  1.000000  0.373586
CAD_SCIE       0.389964  0.369247  0.407885     0.246878  0.373586  1.000000
CAD_SOCI       0.394360  0.377682  0.406969     0.226243  0.395936  0.406456
CI1            0.466216  0.373169  0.445864     0.292001  0.390446  0.327546
CI1_CUT             NaN       NaN       NaN          NaN       NaN       NaN
CI1_MAXPTS    -0.033025  0.017356  0.000051     0.003312 -0.021094  0.012280
CI1_TOTPTS     0.039233  0.074402  0.071051     0.051516  0.038239  0.065449
CI2            0.315042  0.265568  0.239309     0.502002  0.229800  0.157719
CI2_CUT        0.000597 -0.137582 -0.116619    -0.006861 -0.113747 -0.099342
CI2_MAXPTS    -0.043151 -0.020371 -0.032770    -0.001695 -0.048439 -0.003559
CI2_TOTPTS     0.123676  0.112444  0.097587     0.248305  0.080822  0.082263
CI3            0.587935  0.432862  0.488337     0.343922  0.429607  0.388671
CI3_CUT       -0.001227  0.104916  0.087364    -0.024242  0.078641  0.080540
CI3_MAXPTS    -0.049758 -0.038531 -0.028077     0.018712 -0.049258 -0.004984
CI3_TOTPTS     0.214422  0.152175  0.192677     0.172027  0.141760  0.169513
CI4            0.268866  0.338470  0.406555     0.187557  0.328234  0.284216
CI4_CUT       -0.002742  0.132694  0.121957     0.001779  0.104067  0.099632
CI4_GRD_WGT    0.117764  0.077654  0.225860     0.075981  0.138494  0.127514
CI4_PSG_WGT    0.139799  0.129692  0.314707     0.087102  0.182352  0.131403
CI4_RHS_WGT    0.145469  0.164437  0.307243     0.096264  0.116397  0.105914
CI4_STR_WGT    0.324289  0.206839  0.287182     0.221099  0.236280  0.186519
CPEMALLC      -0.085180 -0.032960 -0.072709    -0.046625 -0.079189 -0.037279
CPEMALLP      -0.110299 -0.143553 -0.142276    -0.070451 -0.149937 -0.122671
CPEMALLT      -0.032244  0.047531  0.006925    -0.010753 -0.013700  0.037394
CPETALLC      -0.020469  0.051748  0.014588    -0.010077 -0.008921  0.045540
CPETECHC      -0.034831 -0.011189 -0.011812    -0.024934 -0.018815 -0.035358
CPETECHP      -0.025212  0.010188  0.034820    -0.001872  0.007159 -0.031942
CPETECOC      -0.043999  0.000170 -0.039717    -0.033524 -0.041247 -0.007976
CPETECOP      -0.026243 -0.057639 -0.058955    -0.023423 -0.051093 -0.064986
CPETLEPC      -0.013540 -0.011729 -0.031371    -0.012190 -0.027697 -0.016321
CPETLEPP       0.005423 -0.027375 -0.024566     0.001738 -0.023132 -0.022358
CPETSPEC      -0.067163 -0.008326 -0.047963    -0.041394 -0.065288  0.001691
CPETSPEP      -0.097341 -0.096370 -0.107589    -0.052537 -0.104255 -0.085056
C_UPDATE            NaN  0.421637  0.293972     0.421637       NaN       NaN
C_YRS_IR            NaN       NaN       NaN          NaN       NaN       NaN


              CAD_SOCI       CI1   CI1_CUT  CI1_MAXPTS  …   CPETECHC  \
CAD_GAP       0.394360  0.466216       NaN   -0.033025  … -0.034831
CAD_MATH      0.377682  0.373169       NaN    0.017356  … -0.011189
CAD_POST      0.406969  0.445864       NaN    0.000051  … -0.011812
CAD_PROGRESS  0.226243  0.292001       NaN    0.003312  … -0.024934
CAD_READ      0.395936  0.390446       NaN   -0.021094  … -0.018815
CAD_SCIE      0.406456  0.327546       NaN    0.012280  … -0.035358
CAD_SOCI      1.000000  0.317104       NaN    0.006959  … -0.056512
CI1           0.317104  1.000000  0.364239    0.099350  …  0.036822
CI1_CUT            NaN  0.364239  1.000000    0.206166  …  0.024259
CI1_MAXPTS    0.006959  0.099350  0.206166    1.000000  …  0.114474
```

```
CI1_TOTPTS      0.066276  0.243296  0.209333   0.975803  …   0.102859
CI2             0.188808  0.467652  0.259676  -0.171802  …  -0.089145
CI2_CUT        -0.073270  0.145608  0.575493  -0.014468  …  -0.126454
CI2_MAXPTS      0.027503  0.195151  0.374267   0.669144  …  -0.010790
CI2_TOTPTS      0.114141  0.334227  0.289355   0.414391  …  -0.084942
CI3             0.377867  0.896483  0.254598   0.072553  …   0.097579
CI3_CUT         0.072755  0.364406  0.927167   0.120879  …   0.077029
CI3_MAXPTS      0.006209  0.122972  0.246205   0.641544  …  -0.042789
CI3_TOTPTS      0.179769  0.468409  0.243623   0.585283  …  -0.040040
CI4             0.223644  0.467078 -0.335003   0.101330  …   0.156337
CI4_CUT         0.074376  0.047381 -0.118542   0.134684  …   0.182199
CI4_GRD_WGT     0.117509 -0.385665 -0.868781  -0.331246  …  -0.069329
CI4_PSG_WGT     0.154860  0.342464       NaN  -0.227219  …   0.071979
CI4_RHS_WGT     0.153260  0.273072       NaN   0.079482  …   0.118117
CI4_STR_WGT     0.133395  0.649071  0.279464   0.058542  …  -0.087885
CPEMALLC       -0.048486 -0.255850 -0.173401   0.478228  …   0.138531
CPEMALLP       -0.165258 -0.541135 -0.564154  -0.274826  …  -0.037583
CPEMALLT        0.037003  0.068061  0.142936   0.852678  …   0.173765
CPETALLC        0.046243  0.123462  0.207886   0.836145  …   0.166556
CPETECHC       -0.056512  0.036822  0.024259   0.114474  …   1.000000
CPETECHP       -0.045033  0.111539  0.009145  -0.015998  …   0.674811
CPETECOC       -0.005536 -0.208774  0.179032   0.643495  …   0.268583
CPETECOP       -0.059715 -0.590719 -0.069908  -0.135568  …   0.072634
CPETLEPC        0.015553 -0.153731  0.135557   0.278779  …   0.095747
CPETLEPP        0.005383 -0.199611  0.051922  -0.049429  …  -0.017952
CPETSPEC       -0.026028  0.023629  0.183737   0.799126  …   0.164935
CPETSPEP       -0.144394 -0.265353 -0.102771  -0.070658  …  -0.039639
C_UPDATE        0.540062  0.093621  0.012214   0.165726  …  -0.089803
C_YRS_IR             NaN -0.176347  0.152072   0.065754  …  -0.066400


              CPETECHP  CPETECOC  CPETECOP  CPETLEPC  CPETLEPP  CPETSPEC  \
CAD_GAP       -0.025212 -0.043999 -0.026243 -0.013540  0.005423 -0.067163
CAD_MATH       0.010188  0.000170 -0.057639 -0.011729 -0.027375 -0.008326
CAD_POST       0.034820 -0.039717 -0.058955 -0.031371 -0.024566 -0.047963
CAD_PROGRESS -0.001872 -0.033524 -0.023423 -0.012190  0.001738 -0.041394
CAD_READ       0.007159 -0.041247 -0.051093 -0.027697 -0.023132 -0.065288
CAD_SCIE      -0.031942 -0.007976 -0.064986 -0.016321 -0.022358  0.001691
CAD_SOCI      -0.045033 -0.005536 -0.059715  0.015553  0.005383 -0.026028
CI1            0.111539 -0.208774 -0.590719 -0.153731 -0.199611  0.023629
CI1_CUT        0.009145  0.179032 -0.069908  0.135557  0.051922  0.183737
CI1_MAXPTS    -0.015998  0.643495 -0.135568  0.278779 -0.049429  0.799126
CI1_TOTPTS    -0.012284  0.534737 -0.248455  0.206378 -0.096509  0.756560
CI2           -0.094981 -0.221828 -0.151694  0.084096  0.192534 -0.212890
CI2_CUT       -0.179291  0.047997  0.070534  0.242942  0.294588 -0.019581
CI2_MAXPTS    -0.085254  0.392645 -0.191441  0.268615  0.064559  0.560840
CI2_TOTPTS    -0.156380  0.085699 -0.274118  0.186015  0.112974  0.278785
CI3            0.179211 -0.121494 -0.462100 -0.043953 -0.083399  0.036711
```

```
CI3_CUT         0.080387   0.179637  -0.092321   0.095865  -0.000029   0.181095
CI3_MAXPTS     -0.083851   0.304528  -0.220575   0.104287  -0.101107   0.492709
CI3_TOTPTS     -0.063537   0.113280  -0.454132  -0.026321  -0.198964   0.411477
CI4             0.224543  -0.100387  -0.457501  -0.222325  -0.330912   0.107435
CI4_CUT         0.199738   0.142975  -0.134804  -0.156094  -0.305963   0.224118
CI4_GRD_WGT    -0.044029  -0.276365   0.179567  -0.146874   0.194376  -0.301689
CI4_PSG_WGT     0.181102  -0.235603  -0.108560  -0.175263  -0.121460  -0.268328
CI4_RHS_WGT     0.212186   0.056423  -0.024760   0.101488   0.108701  -0.027694
CI4_STR_WGT    -0.048412  -0.228678  -0.415821   0.009382   0.041799  -0.066642
CPEMALLC        0.007440   0.568627   0.117427   0.257961   0.053285   0.554560
CPEMALLP       -0.047022  -0.220287   0.161378  -0.170242  -0.057448  -0.269532
CPEMALLT        0.027269   0.752922  -0.108530   0.355786   0.009515   0.891258
CPETALLC        0.018645   0.767316  -0.126789   0.413420   0.052473   0.889020
CPETECHC        0.674811   0.268583   0.072634   0.095747  -0.017952   0.164935
CPETECHP        1.000000   0.076786   0.057099  -0.017047  -0.049833  -0.021609
CPETECOC        0.076786   1.000000   0.383976   0.680087   0.339636   0.724147
CPETECOP        0.057099   0.383976   1.000000   0.411119   0.521903  -0.092774
CPETLEPC       -0.017047   0.680087   0.411119   1.000000   0.804078   0.322672
CPETLEPP       -0.049833   0.339636   0.521903   0.804078   1.000000  -0.009974
CPETSPEC       -0.021609   0.724147  -0.092774   0.322672  -0.009974   1.000000
CPETSPEP       -0.095553  -0.130754   0.028996  -0.140665  -0.127509   0.020284
C_UPDATE       -0.089803   0.135888   0.201566  -0.077523  -0.113758   0.098858
C_YRS_IR       -0.050839   0.133043   0.253465  -0.000240  -0.011193   0.096592


                CPETSPEP   C_UPDATE   C_YRS_IR
CAD_GAP        -0.097341        NaN        NaN
CAD_MATH       -0.096370   0.421637        NaN
CAD_POST       -0.107589   0.293972        NaN
CAD_PROGRESS   -0.052537   0.421637        NaN
CAD_READ       -0.104255        NaN        NaN
CAD_SCIE       -0.085056        NaN        NaN
CAD_SOCI       -0.144394   0.540062        NaN
CI1            -0.265353   0.093621  -0.176347
CI1_CUT        -0.102771   0.012214   0.152072
CI1_MAXPTS     -0.070658   0.165726   0.065754
CI1_TOTPTS     -0.087810   0.196713   0.033471
CI2            -0.133290   0.203099   0.336039
CI2_CUT        -0.133715   0.116000   0.155625
CI2_MAXPTS     -0.185507   0.001765   0.181449
CI2_TOTPTS     -0.043116   0.273927   0.293766
CI3            -0.249585   0.009864  -0.169164
CI3_CUT        -0.085617  -0.014614   0.123470
CI3_MAXPTS     -0.164790   0.016600   0.131368
CI3_TOTPTS     -0.073998   0.107007   0.050065
CI4            -0.070042   0.164953  -0.183598
CI4_CUT         0.018145  -0.093222  -0.073003
CI4_GRD_WGT     0.086234   0.439372  -0.093038
```

```
CI4_PSG_WGT  -0.271803  0.205398 -0.236144
CI4_RHS_WGT  -0.440854  0.351925 -0.476254
CI4_STR_WGT  -0.123256  0.103198 -0.145999
CPEMALLC     -0.010579 -0.239798  0.058903
CPEMALLP      0.300029  0.009527  0.036730
CPEMALLT     -0.118786 -0.134136  0.057728
CPETALLC     -0.151917  0.129779  0.080073
CPETECHC     -0.039639 -0.089803 -0.066400
CPETECHP     -0.095553 -0.089803 -0.050839
CPETECOC     -0.130754  0.135888  0.133043
CPETECOP      0.028996  0.201566  0.253465
CPETLEPC     -0.140665 -0.077523 -0.000240
CPETLEPP     -0.127509 -0.113758 -0.011193
CPETSPEC      0.020284  0.098858  0.096592
CPETSPEP      1.000000 -0.155405  0.026386
C_UPDATE     -0.155405  1.000000       NaN
C_YRS_IR      0.026386       NaN  1.000000

[39 rows x 39 columns]
```

By evaluating the correlation table above, I found that the following sets of variables have relatively strong correlations. Both the correlation value and a visual aid are provided.

```
[5]: #CPETECOP:% Econ Disadv
     #CI4_STR_WGT: Campus 2017 Index 4: STAAR weighted value

     df1 = df[['CPETECOP', 'CI4_STR_WGT']].dropna()
     df1.corr()
```

```
[5]:               CPETECOP  CI4_STR_WGT
     CPETECOP      1.000000    -0.415821
     CI4_STR_WGT  -0.415821     1.000000
```

```
[6]: x = df1.CPETECOP
     y = df1.CI4_STR_WGT

     fig, ax = plt.subplots(figsize=(10,8))
     ax.scatter(x, y, color='g')

     m, b = np.polyfit(x, y, 1)
     ax.plot(x, m*x + b, 'orange')

     ax.set_title('Campus STAAR weighted value vs. Percent Economically Disadvanted␣
      ↪Students')
     ax.set_ylabel('Campus STAAR weighted value')
     ax.set_xlabel('% of students who are economically disadvantaged')
```
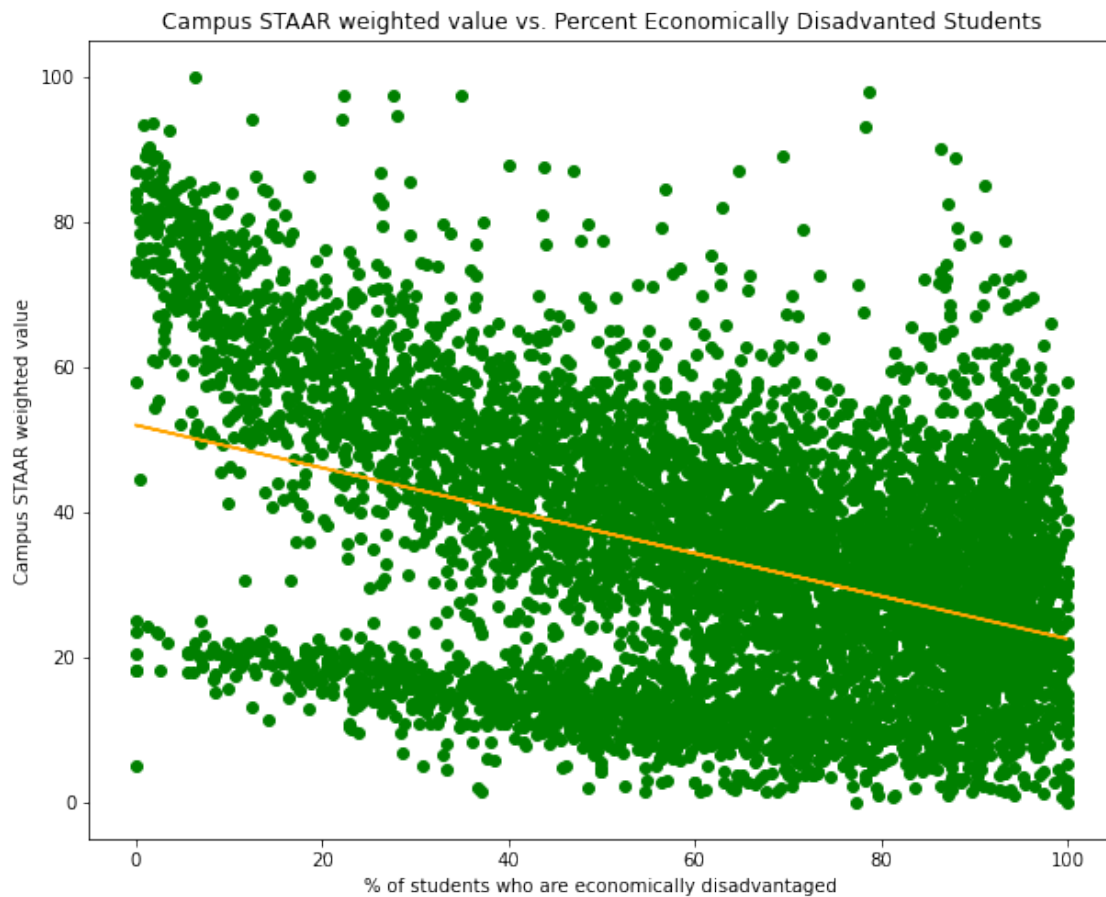
```
plt.show()
```

Campus STAAR weighted value vs. Percent Economically Disadvanted Students


```
[7]: #CPETECOP:% Econ Disadv
     #CPETLEPP: % LEP Students

     df2 = df[['CPETECOP','CPETLEPP']].dropna()
     df2.corr()
```

```
[7]:           CPETECOP   CPETLEPP
     CPETECOP  1.000000   0.521903
     CPETLEPP  0.521903   1.000000
```

```
[8]: x = df2.CPETECOP
     y = df2.CPETLEPP

     fig, ax = plt.subplots(figsize=(10,8))

     ax.scatter(x, y, color='g')
```
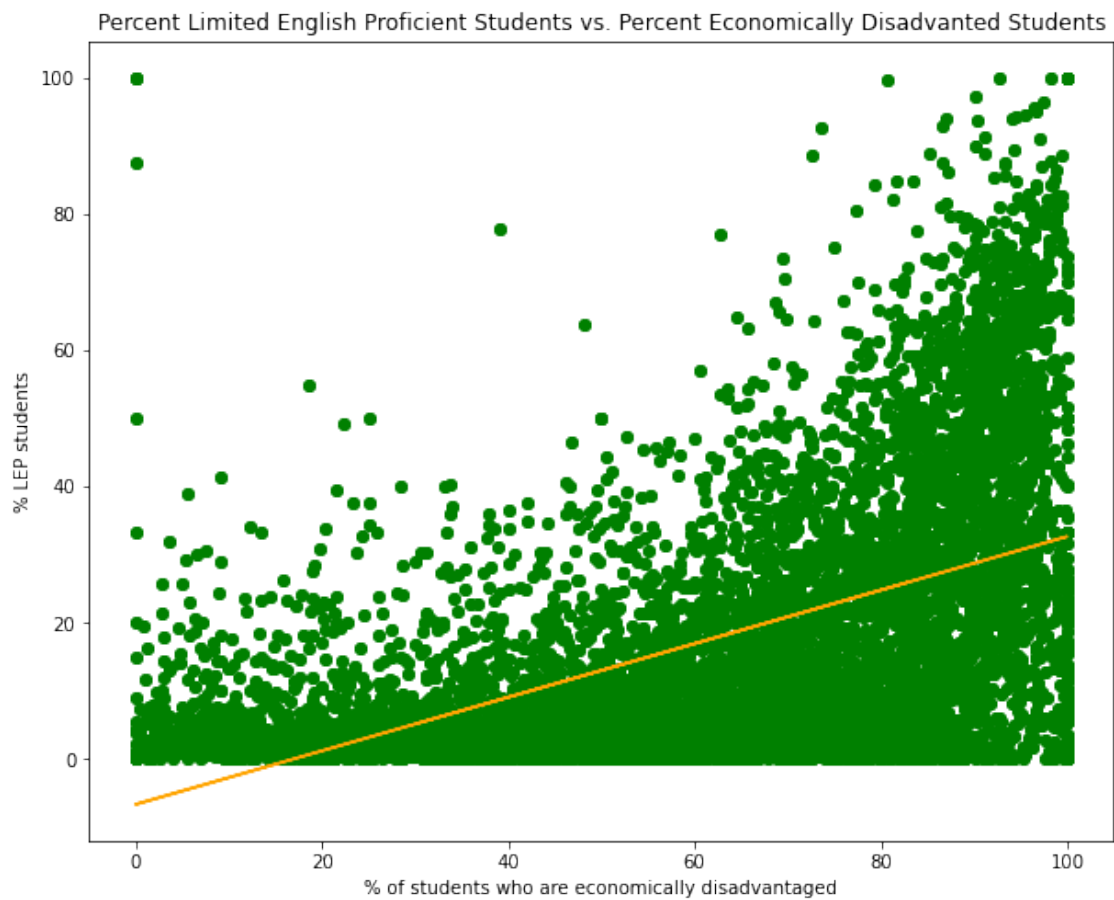
```
m, b = np.polyfit(x, y, 1)
ax.plot(x, m*x + b, 'orange')
ax.scatter(x, y, color='g')

ax.set_title('Percent Limited English Proficient Students vs. Percent␣
↪Economically Disadvanted Students')
ax.set_ylabel('% LEP students')
ax.set_xlabel('% of students who are economically disadvantaged')

plt.show()
```



```
[9]: #CPETALLC: Total Number of Students
     #CPETECOC: # Econ Disadv

     df3 = df[['CPETALLC','CPETECOC']]
     df3.corr()
```

```
[9]:            CPETALLC   CPETECOC
     CPETALLC   1.000000   0.767316
     CPETECOC   0.767316   1.000000
```

```
[10]: x = df3.CPETALLC
      y = df3.CPETECOC

      fig, ax = plt.subplots(figsize=(10,8))

      ax.scatter(x, y, color='g')

      m, b = np.polyfit(x, y, 1)
      ax.plot(x, m*x + b, 'orange')
      ax.scatter(x, y, color='g')

      ax.set_title('Number of Economically Disadvanted Students vs. Total Number of␣
       ↪Students')
      ax.set_ylabel('Number of economically disadvanted students')
      ax.set_xlabel('Total number of students')
      plt.show()
```