Homework 1

January 25, 2021

[2]:
```
import os
path = r'C:\Users\srsid\Documents\Machine Learning\Boston'
os.chdir(path)
os.listdir()
```

[2]:
```
['.ipynb_checkpoints',
 'Boston data_description.txt',
 'Boston.csv',
 'Homework 1 Problems.docx',
 'Homework 1.ipynb',
 '~$mework 1 Problems.docx']
```

[3]:
```
import pandas as pd
Boston_Housing = pd.read_csv('Boston.csv')
Boston_Housing
```

[3]:

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	\
0	0.00632	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1.0	296.0	
1	0.02731	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2.0	242.0	
2	0.02729	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2.0	242.0	
3	0.03237	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3.0	222.0	
4	0.06905	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3.0	222.0	
..	…	…	…	…	…	…	…	…	…	…	
501	0.06263	0.0	11.93	0.0	0.573	6.593	69.1	2.4786	1.0	273.0	
502	0.04527	0.0	11.93	0.0	0.573	6.120	76.7	2.2875	1.0	273.0	
503	0.06076	0.0	11.93	0.0	0.573	6.976	91.0	2.1675	1.0	273.0	
504	0.10959	0.0	11.93	0.0	0.573	6.794	89.3	2.3889	1.0	273.0	
505	0.04741	0.0	11.93	0.0	0.573	6.030	80.8	2.5050	1.0	273.0	

	PTRATIO	B	LSTAT	MDEV
0	15.3	396.90	4.98	24.0
1	17.8	396.90	9.14	21.6
2	17.8	392.83	4.03	34.7
3	18.7	394.63	2.94	33.4
4	18.7	396.90	5.33	36.2
..	…	…	…	…
501	21.0	391.99	9.67	22.4
502	21.0	396.90	9.08	20.6

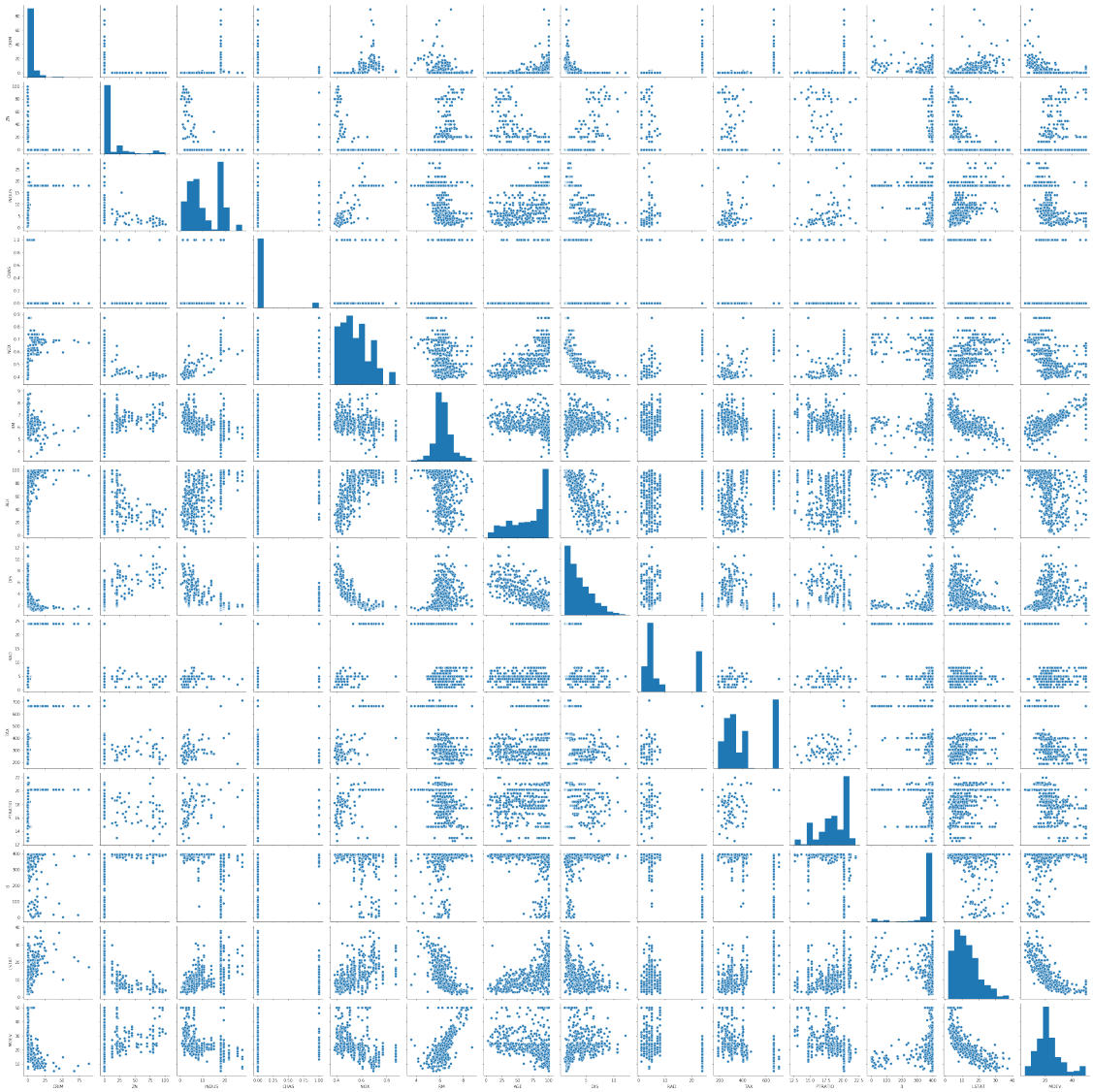
```
503     21.0  396.90   5.64  23.9
504     21.0  393.45   6.48  22.0
505     21.0  396.90   7.88  11.9

[506 rows x 14 columns]
```

[4]:
```
Boston_Housing.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 506 entries, 0 to 505
Data columns (total 14 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   CRIM     506 non-null    float64
 1   ZN       506 non-null    float64
 2   INDUS    506 non-null    float64
 3   CHAS     506 non-null    float64
 4   NOX      506 non-null    float64
 5   RM       506 non-null    float64
 6   AGE      506 non-null    float64
 7   DIS      506 non-null    float64
 8   RAD      506 non-null    float64
 9   TAX      506 non-null    float64
 10  PTRATIO  506 non-null    float64
 11  B        506 non-null    float64
 12  LSTAT    506 non-null    float64
 13  MDEV     506 non-null    float64
dtypes: float64(14)
memory usage: 55.5 KB
```

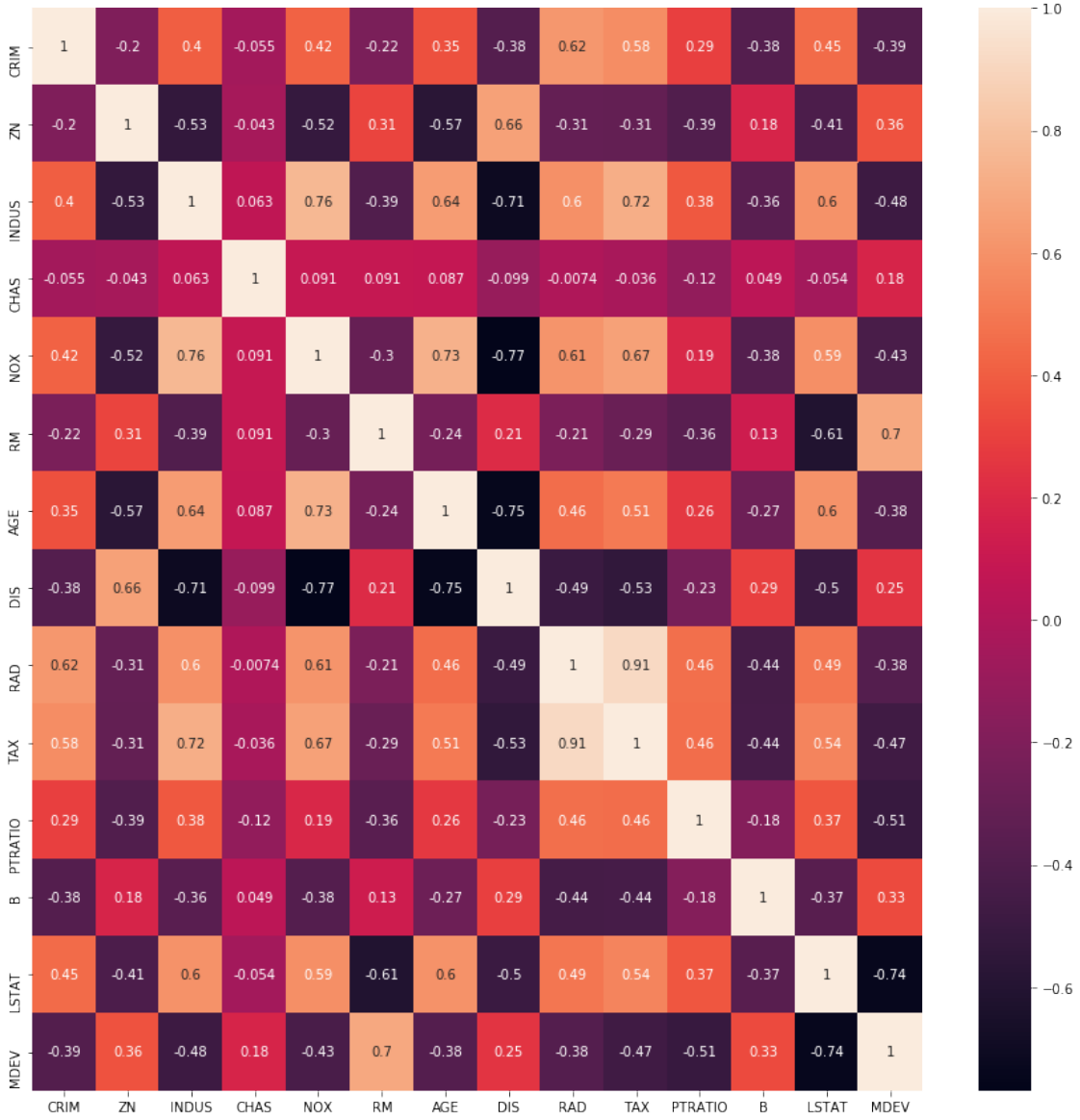
[5]:
```
import seaborn as sns

sns.pairplot(Boston_Housing, vars =␣
 ↪['CRIM','ZN','INDUS','CHAS','NOX','RM','AGE','DIS','RAD','TAX','PTRATIO','B','LSTAT','MDEV']
```

[5]:
```
<seaborn.axisgrid.PairGrid at 0x299f1a29808>
```

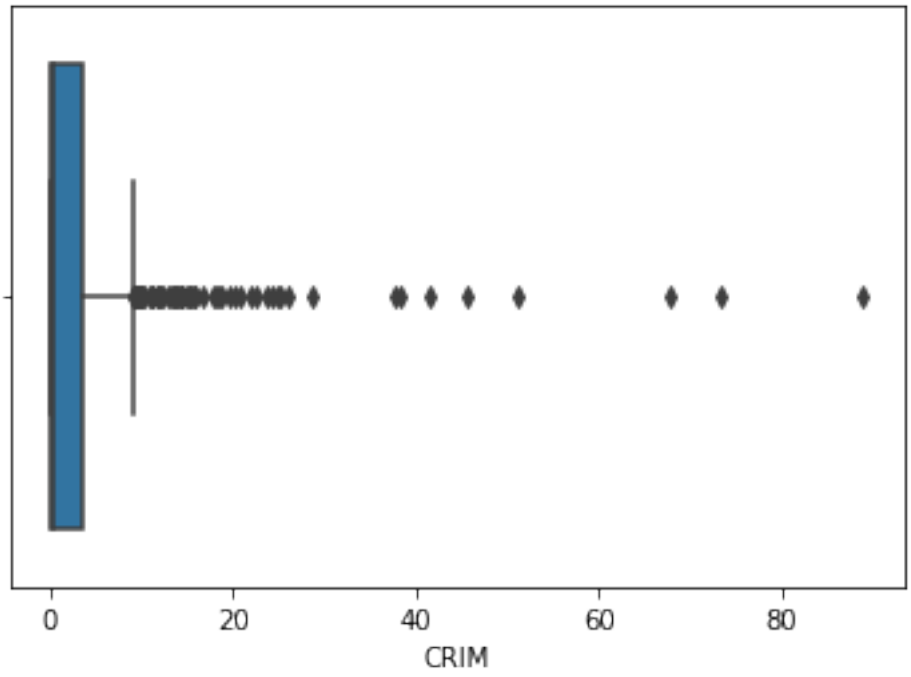


[6]:
```
import matplotlib.pyplot as plt
plt.figure(figsize=(15,15))
sns.heatmap(Boston_Housing.corr(), annot=True)
```

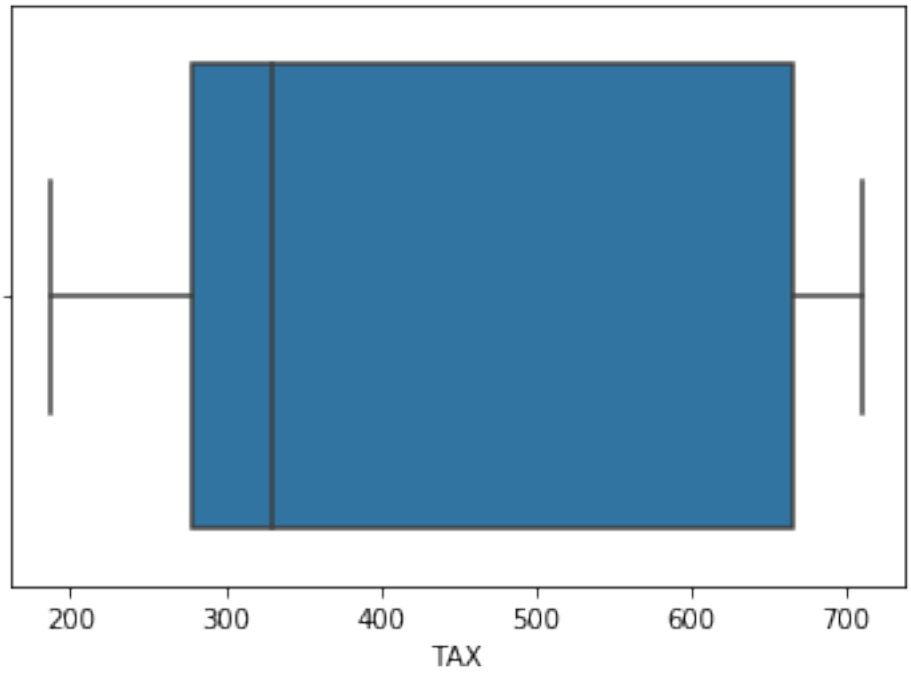
[6]: <matplotlib.axes._subplots.AxesSubplot at 0x299ecc7c488>



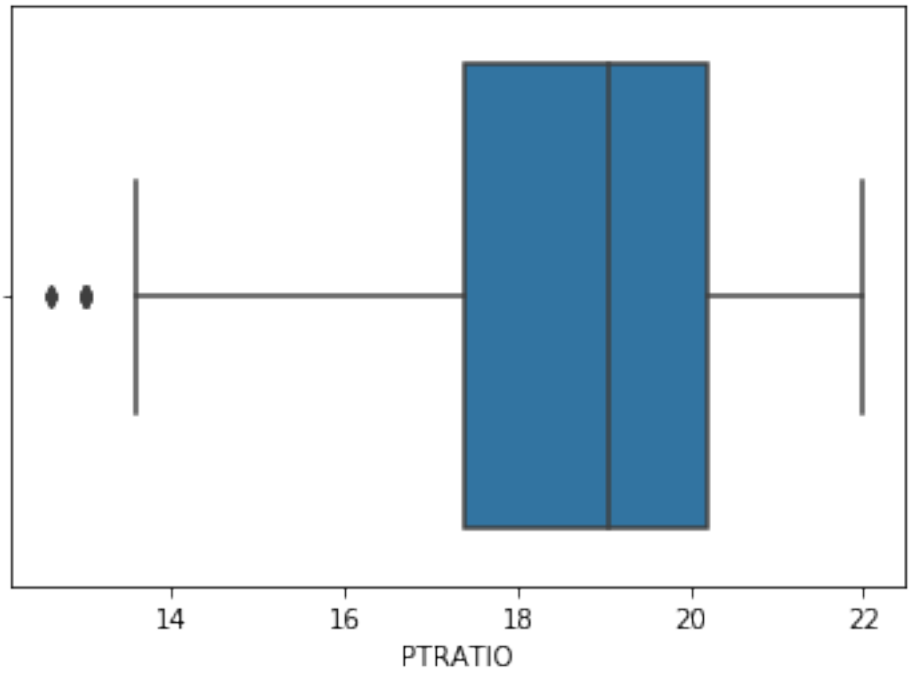
[7]:
```
ax = sns.boxplot(x="CRIM", data=Boston_Housing)
```



[8]:
```
ax = sns.boxplot(x="TAX", data=Boston_Housing)
```



[9]:
```
ax = sns.boxplot(x="PTRATIO", data=Boston_Housing)
```



[10]:
```
Chas = Boston_Housing[Boston_Housing['CHAS']==1]
Chas['CHAS'].count()
```

[10]:
```
35
```

[11]:
```
Median = Boston_Housing['MDEV']
Median.describe()
```

[11]:
```
count    506.000000
mean      22.532806
std        9.197104
min        5.000000
25%       17.025000
50%       21.200000
75%       25.000000
max       50.000000
Name: MDEV, dtype: float64
```

[12]:
```
Boston_Housing.describe()
```

[12]:
```
              CRIM          ZN       INDUS        CHAS         NOX          RM  \
count   506.000000  506.000000  506.000000  506.000000  506.000000  506.000000
mean      3.593761   11.363636   11.136779    0.069170    0.554695    6.284634
```

```
std      8.596783   23.322453    6.860353    0.253994    0.115878    0.702617
min      0.006320    0.000000    0.460000    0.000000    0.385000    3.561000
25%      0.082045    0.000000    5.190000    0.000000    0.449000    5.885500
50%      0.256510    0.000000    9.690000    0.000000    0.538000    6.208500
75%      3.647422   12.500000   18.100000    0.000000    0.624000    6.623500
max     88.976200  100.000000   27.740000    1.000000    0.871000    8.780000

              AGE         DIS         RAD         TAX     PTRATIO           B  \
count  506.000000  506.000000  506.000000  506.000000  506.000000  506.000000
mean    68.574901    3.795043    9.549407  408.237154   18.455534  356.674032
std     28.148861    2.105710    8.707259  168.537116    2.164946   91.294864
min      2.900000    1.129600    1.000000  187.000000   12.600000    0.320000
25%     45.025000    2.100175    4.000000  279.000000   17.400000  375.377500
50%     77.500000    3.207450    5.000000  330.000000   19.050000  391.440000
75%     94.075000    5.188425   24.000000  666.000000   20.200000  396.225000
max    100.000000   12.126500   24.000000  711.000000   22.000000  396.900000

            LSTAT        MDEV
count  506.000000  506.000000
mean    12.653063   22.532806
std      7.141062    9.197104
min      1.730000    5.000000
25%      6.950000   17.025000
50%     11.360000   21.200000
75%     16.955000   25.000000
max     37.970000   50.000000
```

[13]:
```
Boston_min = Boston_Housing.describe().iloc[3]
Boston_max = Boston_Housing.describe().iloc[7]
Boston_range = Boston_max-Boston_min
print(Boston_range)
```

```
CRIM        88.96988
ZN         100.00000
INDUS       27.28000
CHAS         1.00000
NOX          0.48600
RM           5.21900
AGE         97.10000
DIS         10.99690
RAD         23.00000
TAX        524.00000
PTRATIO      9.40000
B          396.58000
LSTAT       36.24000
MDEV        45.00000
dtype: float64
```

[14]:
```
from statistics import median
Lowest = Boston_Housing[Boston_Housing['MDEV'] < Boston_Housing['MDEV'].
 ↪median()]
Lowest.min()
```

[14]:
```
CRIM         0.0136
ZN           0.0000
INDUS        1.6900
CHAS         0.0000
NOX          0.3850
RM           4.1380
AGE         18.5000
DIS          1.1370
RAD          1.0000
TAX        188.0000
PTRATIO     14.7000
B            0.3200
LSTAT        5.5700
MDEV         5.0000
dtype: float64
```

[15]:
```
Seven =  Boston_Housing[Boston_Housing['RM'] >= 7]
Seven['RM'].count()
```

[15]:
```
64
```

[16]:
```
Eight =  Boston_Housing[Boston_Housing['RM'] >= 8]
print(Eight['RM'].count())
print(Eight['AGE'].mean())
Eight
```

```
13
71.53846153846153
```

[16]:

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	\
97	0.12083	0.0	2.89	0.0	0.4450	8.069	76.0	3.4952	2.0	276.0	
163	1.51902	0.0	19.58	1.0	0.6050	8.375	93.9	2.1620	5.0	403.0	
204	0.02009	95.0	2.68	0.0	0.4161	8.034	31.9	5.1180	4.0	224.0	
224	0.31533	0.0	6.20	0.0	0.5040	8.266	78.3	2.8944	8.0	307.0	
225	0.52693	0.0	6.20	0.0	0.5040	8.725	83.0	2.8944	8.0	307.0	
226	0.38214	0.0	6.20	0.0	0.5040	8.040	86.5	3.2157	8.0	307.0	
232	0.57529	0.0	6.20	0.0	0.5070	8.337	73.3	3.8384	8.0	307.0	
233	0.33147	0.0	6.20	0.0	0.5070	8.247	70.4	3.6519	8.0	307.0	
253	0.36894	22.0	5.86	0.0	0.4310	8.259	8.4	8.9067	7.0	330.0	
257	0.61154	20.0	3.97	0.0	0.6470	8.704	86.9	1.8010	5.0	264.0	
262	0.52014	20.0	3.97	0.0	0.6470	8.398	91.5	2.2885	5.0	264.0	
267	0.57834	20.0	3.97	0.0	0.5750	8.297	67.0	2.4216	5.0	264.0	
364	3.47428	0.0	18.10	1.0	0.7180	8.780	82.9	1.9047	24.0	666.0	

```
     PTRATIO       B  LSTAT  MDEV
97      18.0  396.90   4.21  38.7
163     14.7  388.45   3.32  50.0
204     14.7  390.55   2.88  50.0
224     17.4  385.05   4.14  44.8
225     17.4  382.00   4.63  50.0
226     17.4  387.38   3.13  37.6
232     17.4  385.91   2.47  41.7
233     17.4  378.95   3.95  48.3
253     19.1  396.90   3.54  42.8
257     13.0  389.70   5.12  50.0
262     13.0  386.86   5.91  48.8
267     13.0  384.54   7.44  50.0
364     20.2  354.55   5.29  21.9
```

[17]:
```
import statsmodels.formula.api as smf
ZN = smf.ols('CRIM ~ ZN', data = Boston_Housing).fit()
ZN.summary()
```

[17]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.040
Model:                            OLS   Adj. R-squared:                  0.038
Method:                 Least Squares   F-statistic:                     20.88
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           6.15e-06
Time:                        21:44:21   Log-Likelihood:                -1795.8
No. Observations:                 506   AIC:                             3596.
Df Residuals:                     504   BIC:                             3604.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      4.4292      0.417     10.620      0.000       3.610       5.249
ZN            -0.0735      0.016     -4.570      0.000      -0.105      -0.042
==============================================================================
Omnibus:                      568.366   Durbin-Watson:                   0.862
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            32952.356
Skew:                           5.270   Prob(JB):                         0.00
Kurtosis:                      41.103   Cond. No.                         28.8
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
```

```
specified.
"""
```

[18]:
```
INDUS = smf.ols('CRIM ~ INDUS', data = Boston_Housing).fit()
INDUS.summary()
```

[18]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.164
Model:                            OLS   Adj. R-squared:                  0.162
Method:                 Least Squares   F-statistic:                     98.58
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           2.44e-21
Time:                        21:44:21   Log-Likelihood:                -1760.9
No. Observations:                 506   AIC:                             3526.
Df Residuals:                     504   BIC:                             3534.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -2.0509      0.668     -3.072      0.002      -3.362      -0.739
INDUS          0.5068      0.051      9.929      0.000       0.407       0.607
==============================================================================
Omnibus:                      585.528   Durbin-Watson:                   0.990
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            41469.710
Skew:                           5.456   Prob(JB):                         0.00
Kurtosis:                      45.987   Cond. No.                         25.1
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[19]:
```
CHAS = smf.ols('CRIM ~ CHAS', data = Boston_Housing).fit()
CHAS.summary()
```

[19]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.003
Model:                            OLS   Adj. R-squared:                  0.001
Method:                 Least Squares   F-statistic:                     1.546
Date:                Sat, 23 Jan 2021   Prob (F-statistic):              0.214
```

```
Time:                        21:44:21   Log-Likelihood:                -1805.3
No. Observations:                 506   AIC:                             3615.
Df Residuals:                     504   BIC:                             3623.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      3.7232      0.396      9.404      0.000       2.945       4.501
CHAS          -1.8715      1.505     -1.243      0.214      -4.829       1.086
==============================================================================
Omnibus:                      562.698   Durbin-Watson:                   0.822
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            30864.755
Skew:                           5.205   Prob(JB):                         0.00
Kurtosis:                      39.818   Cond. No.                         3.96
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[20]:
```
NOX = smf.ols('CRIM ~ NOX', data = Boston_Housing).fit()
NOX.summary()
```

[20]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.174
Model:                            OLS   Adj. R-squared:                  0.173
Method:                 Least Squares   F-statistic:                     106.4
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           9.16e-23
Time:                        21:44:21   Log-Likelihood:                -1757.6
No. Observations:                 506   AIC:                             3519.
Df Residuals:                     504   BIC:                             3528.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept    -13.5881      1.702     -7.986      0.000     -16.931     -10.245
NOX           30.9753      3.003     10.315      0.000      25.076      36.875
==============================================================================
Omnibus:                      591.496   Durbin-Watson:                   0.994
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            42994.381
Skew:                           5.544   Prob(JB):                         0.00
```

```
Kurtosis:                       46.776   Cond. No.                         11.3
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[21]:
```
RM = smf.ols('CRIM ~ RM', data = Boston_Housing).fit()
RM.summary()
```

[21]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.048
Model:                            OLS   Adj. R-squared:                  0.046
Method:                 Least Squares   F-statistic:                     25.62
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           5.84e-07
Time:                        21:44:21   Log-Likelihood:                -1793.5
No. Observations:                 506   AIC:                             3591.
Df Residuals:                     504   BIC:                             3600.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     20.5060      3.362      6.099      0.000      13.901      27.111
RM            -2.6910      0.532     -5.062      0.000      -3.736      -1.646
==============================================================================
Omnibus:                      576.890   Durbin-Watson:                   0.883
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            36966.825
Skew:                           5.361   Prob(JB):                         0.00
Kurtosis:                      43.477   Cond. No.                         58.4
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[22]:
```
AGE = smf.ols('CRIM ~ AGE', data = Boston_Housing).fit()
AGE.summary()
```

[22]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
```

```
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.123
Model:                            OLS   Adj. R-squared:                  0.121
Method:                 Least Squares   F-statistic:                     70.72
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           4.26e-16
Time:                        21:44:21   Log-Likelihood:                -1772.9
No. Observations:                 506   AIC:                             3550.
Df Residuals:                     504   BIC:                             3558.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -3.7527      0.944     -3.974      0.000      -5.608      -1.898
AGE            0.1071      0.013      8.409      0.000       0.082       0.132
==============================================================================
Omnibus:                      575.090   Durbin-Watson:                   0.960
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            36851.412
Skew:                           5.331   Prob(JB):                         0.00
Kurtosis:                      43.426   Cond. No.                         195.
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[23]:
```
DIS = smf.ols('CRIM ~ DIS', data = Boston_Housing).fit()
DIS.summary()
```

[23]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.143
Model:                            OLS   Adj. R-squared:                  0.141
Method:                 Least Squares   F-statistic:                     83.97
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           1.27e-18
Time:                        21:44:21   Log-Likelihood:                -1767.1
No. Observations:                 506   AIC:                             3538.
Df Residuals:                     504   BIC:                             3547.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      9.4489      0.731     12.934      0.000       8.014      10.884
```

```
DIS           -1.5428      0.168     -9.163      0.000      -1.874      -1.212
==============================================================================
Omnibus:                      577.090   Durbin-Watson:                   0.957
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            37542.100
Skew:                           5.357   Prob(JB):                         0.00
Kurtosis:                      43.815   Cond. No.                         9.32
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[24]:
```
RAD = smf.ols('CRIM ~ RAD', data = Boston_Housing).fit()
RAD.summary()
```

[24]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.387
Model:                            OLS   Adj. R-squared:                  0.386
Method:                 Least Squares   F-statistic:                     318.1
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           1.62e-55
Time:                        21:44:21   Log-Likelihood:                -1682.3
No. Observations:                 506   AIC:                             3369.
Df Residuals:                     504   BIC:                             3377.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -2.2709      0.445     -5.105      0.000      -3.145      -1.397
RAD            0.6141      0.034     17.835      0.000       0.546       0.682
==============================================================================
Omnibus:                      654.232   Durbin-Watson:                   1.336
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            74327.568
Skew:                           6.441   Prob(JB):                         0.00
Kurtosis:                      60.961   Cond. No.                         19.2
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[25]:
```
TAX = smf.ols('CRIM ~ TAX', data = Boston_Housing).fit()
TAX.summary()
```

[25]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.336
Model:                            OLS   Adj. R-squared:                  0.335
Method:                 Least Squares   F-statistic:                     254.9
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           9.76e-47
Time:                        21:44:21   Log-Likelihood:                -1702.5
No. Observations:                 506   AIC:                             3409.
Df Residuals:                     504   BIC:                             3418.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -8.4748      0.818    -10.365      0.000     -10.081      -6.868
TAX            0.0296      0.002     15.966      0.000       0.026       0.033
==============================================================================
Omnibus:                      634.003   Durbin-Watson:                   1.252
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            63141.063
Skew:                           6.134   Prob(JB):                         0.00
Kurtosis:                      56.332   Cond. No.                     1.16e+03
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 1.16e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[26]:
```
PTRATIO = smf.ols('CRIM ~ PTRATIO', data = Boston_Housing).fit()
PTRATIO.summary()
```

[26]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.083
Model:                            OLS   Adj. R-squared:                  0.081
Method:                 Least Squares   F-statistic:                     45.67
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           3.88e-11
Time:                        21:44:21   Log-Likelihood:                -1784.1
```

```
No. Observations:                 506   AIC:                             3572.
Df Residuals:                     504   BIC:                             3581.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept    -17.5307      3.147     -5.570      0.000     -23.714     -11.347
PTRATIO        1.1446      0.169      6.758      0.000       0.812       1.477
==============================================================================
Omnibus:                      568.808   Durbin-Watson:                   0.909
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            34373.378
Skew:                           5.256   Prob(JB):                         0.00
Kurtosis:                      41.985   Cond. No.                         160.
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[27]:
```
B = smf.ols('CRIM ~ B', data = Boston_Housing).fit()
B.summary()
```

[27]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.142
Model:                            OLS   Adj. R-squared:                  0.141
Method:                 Least Squares   F-statistic:                     83.69
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           1.43e-18
Time:                        21:44:21   Log-Likelihood:                -1767.2
No. Observations:                 506   AIC:                             3538.
Df Residuals:                     504   BIC:                             3547.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     16.2680      1.430     11.376      0.000      13.458      19.078
B             -0.0355      0.004     -9.148      0.000      -0.043      -0.028
==============================================================================
Omnibus:                      591.626   Durbin-Watson:                   1.001
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            43282.465
Skew:                           5.543   Prob(JB):                         0.00
Kurtosis:                      46.932   Cond. No.                     1.49e+03
```

```
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 1.49e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[28]:
```
LSTAT = smf.ols('CRIM ~ LSTAT', data = Boston_Housing).fit()
LSTAT.summary()
```

[28]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.205
Model:                            OLS   Adj. R-squared:                  0.203
Method:                 Least Squares   F-statistic:                     129.6
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           7.12e-27
Time:                        21:44:21   Log-Likelihood:                -1748.2
No. Observations:                 506   AIC:                             3500.
Df Residuals:                     504   BIC:                             3509.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -3.2946      0.695     -4.742      0.000      -4.660      -1.930
LSTAT          0.5444      0.048     11.383      0.000       0.450       0.638
==============================================================================
Omnibus:                      600.766   Durbin-Watson:                   1.184
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            49637.173
Skew:                           5.638   Prob(JB):                         0.00
Kurtosis:                      50.193   Cond. No.                         29.7
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[29]:
```
MDEV = smf.ols('CRIM ~ MDEV', data = Boston_Housing).fit()
MDEV.summary()
```

[29]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.149
Model:                            OLS   Adj. R-squared:                  0.147
Method:                 Least Squares   F-statistic:                     88.15
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           2.08e-19
Time:                        21:44:21   Log-Likelihood:                -1765.3
No. Observations:                 506   AIC:                             3535.
Df Residuals:                     504   BIC:                             3543.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     11.7202      0.935     12.539      0.000       9.884      13.557
MDEV          -0.3606      0.038     -9.389      0.000      -0.436      -0.285
==============================================================================
Omnibus:                      559.282   Durbin-Watson:                   1.000
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            32809.507
Skew:                           5.114   Prob(JB):                         0.00
Kurtosis:                      41.099   Cond. No.                         64.5
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
"""
```

[30]:
```
Boston_Housing.iloc[:,1:]
```

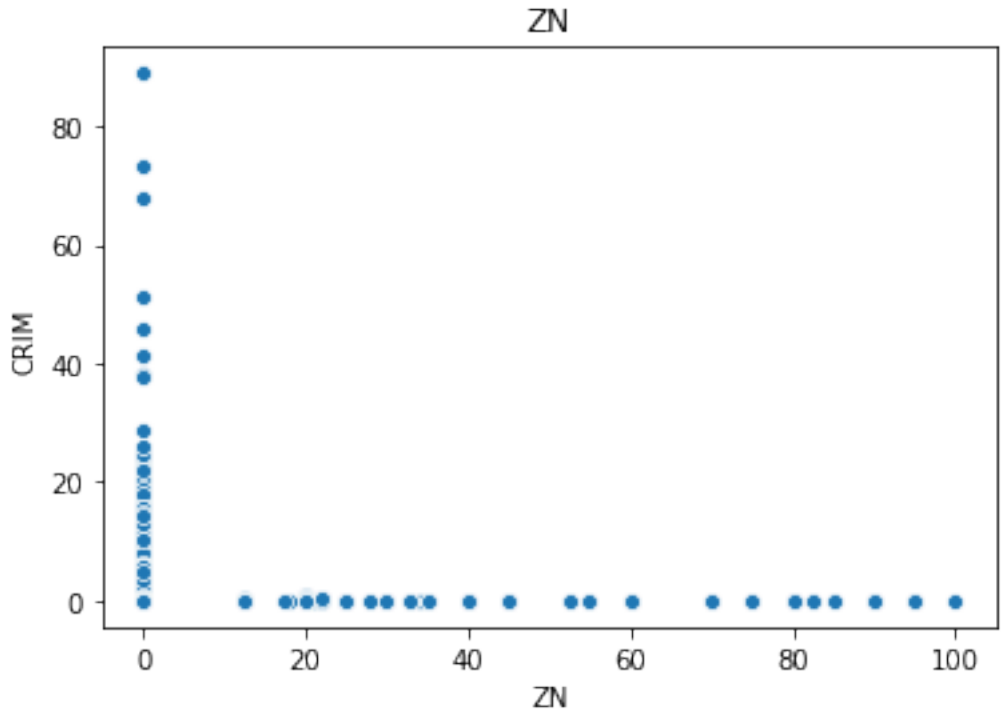
[30]:

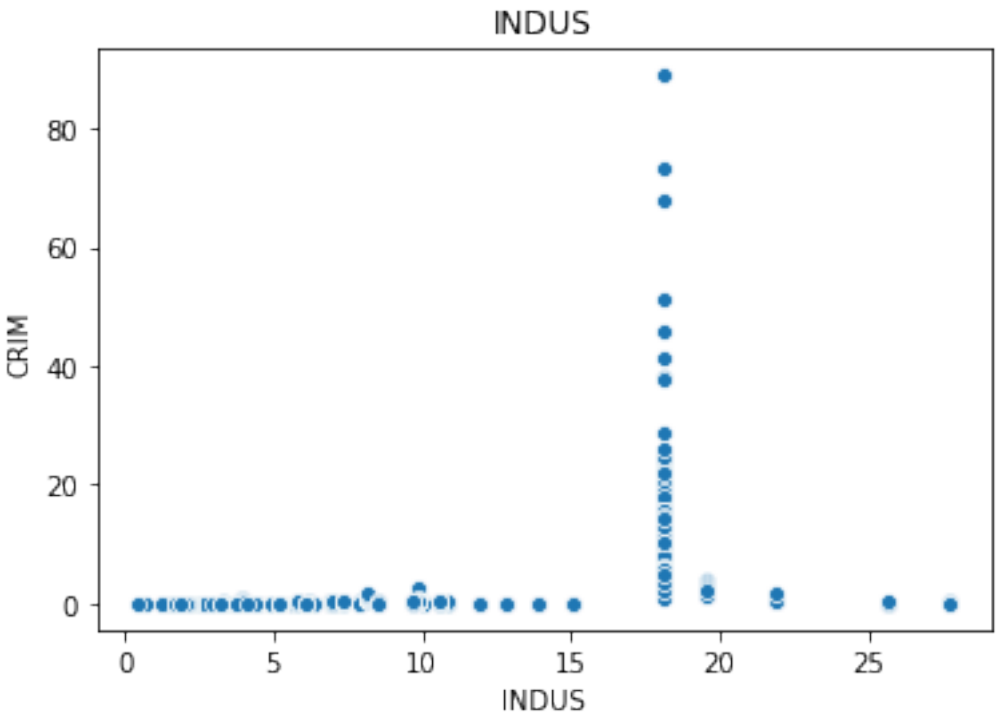
	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	\
0	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1.0	296.0	15.3	
1	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2.0	242.0	17.8	
2	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2.0	242.0	17.8	
3	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3.0	222.0	18.7	
4	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3.0	222.0	18.7	
..	…	…	…	…	…	…	…	…	…	…	
501	0.0	11.93	0.0	0.573	6.593	69.1	2.4786	1.0	273.0	21.0	
502	0.0	11.93	0.0	0.573	6.120	76.7	2.2875	1.0	273.0	21.0	
503	0.0	11.93	0.0	0.573	6.976	91.0	2.1675	1.0	273.0	21.0	
504	0.0	11.93	0.0	0.573	6.794	89.3	2.3889	1.0	273.0	21.0	
505	0.0	11.93	0.0	0.573	6.030	80.8	2.5050	1.0	273.0	21.0	

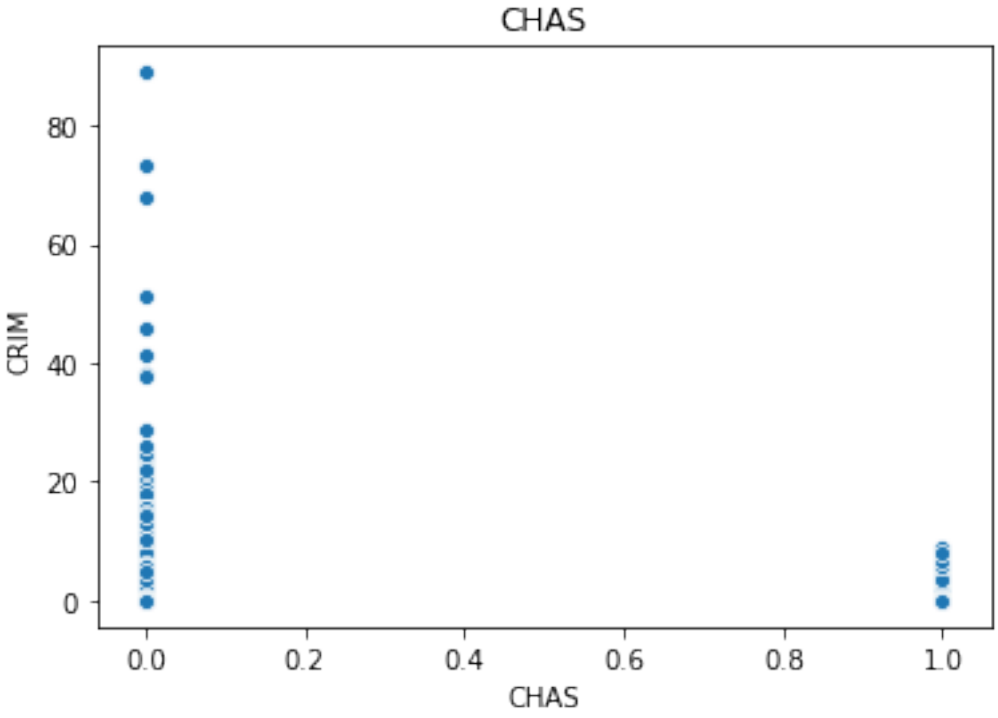
	B	LSTAT	MDEV
0	396.90	4.98	24.0
1	396.90	9.14	21.6
2	392.83	4.03	34.7

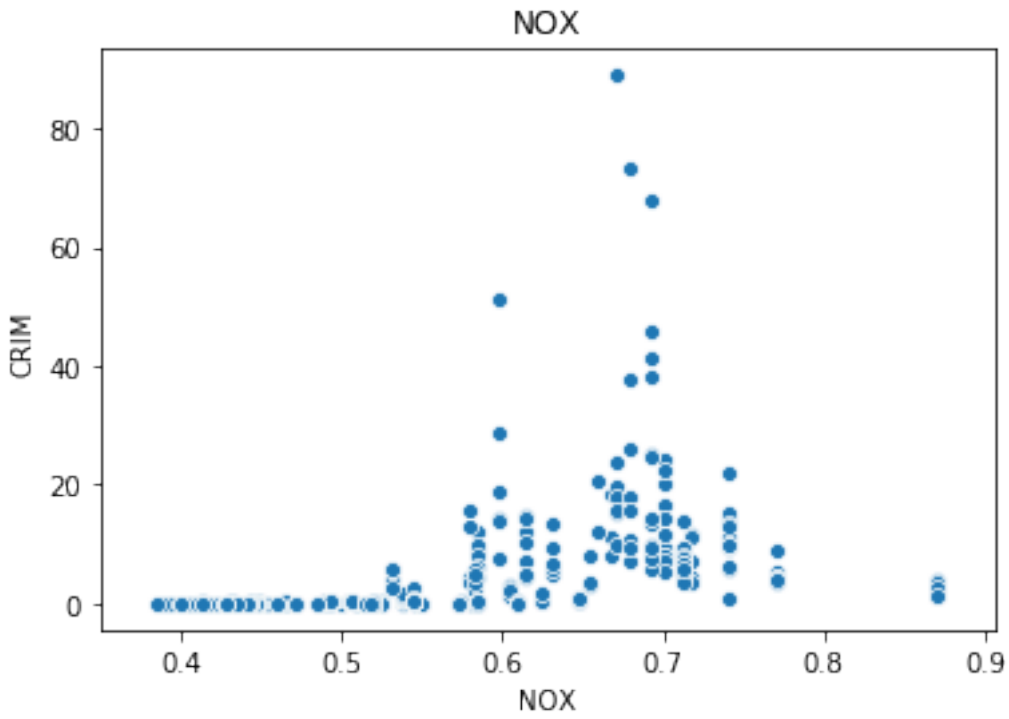
```
3    394.63   2.94  33.4
4    396.90   5.33  36.2
..      …     …   …
501  391.99   9.67  22.4
502  396.90   9.08  20.6
503  396.90   5.64  23.9
504  393.45   6.48  22.0
505  396.90   7.88  11.9

[506 rows x 13 columns]
```

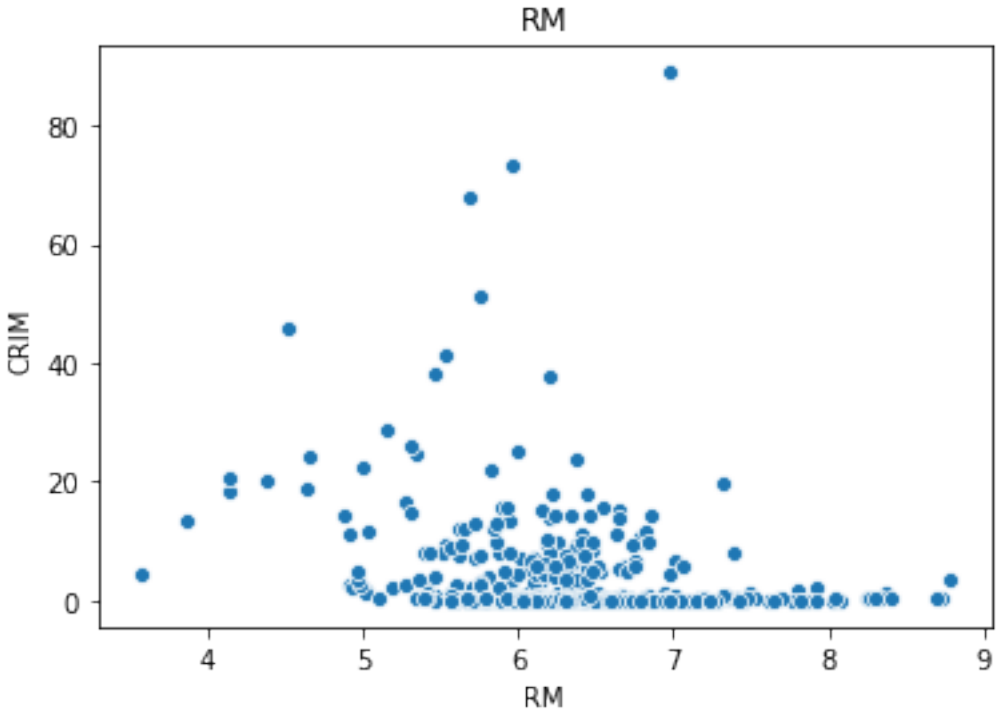
[31]:
```
import matplotlib.pyplot as plt
import seaborn as sns
for col in Boston_Housing.iloc[:,1:].columns:
    sns.scatterplot(Boston_Housing[col],Boston_Housing['CRIM'])
    plt.title(col)
    plt.xlabel(col)
    plt.ylabel('CRIM')
    plt.show()
```

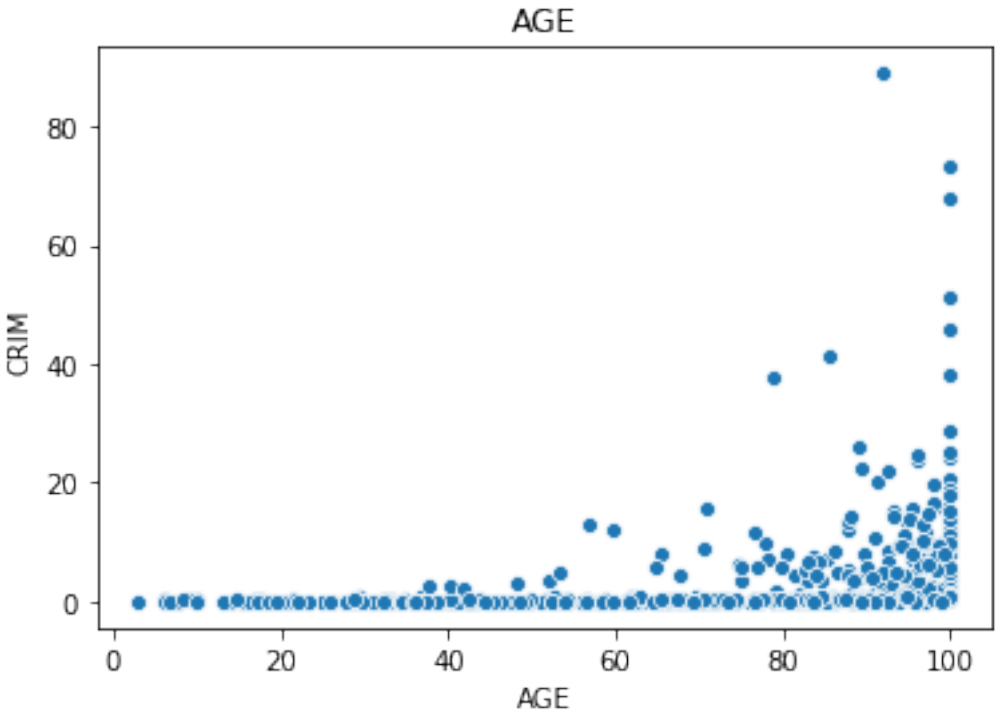


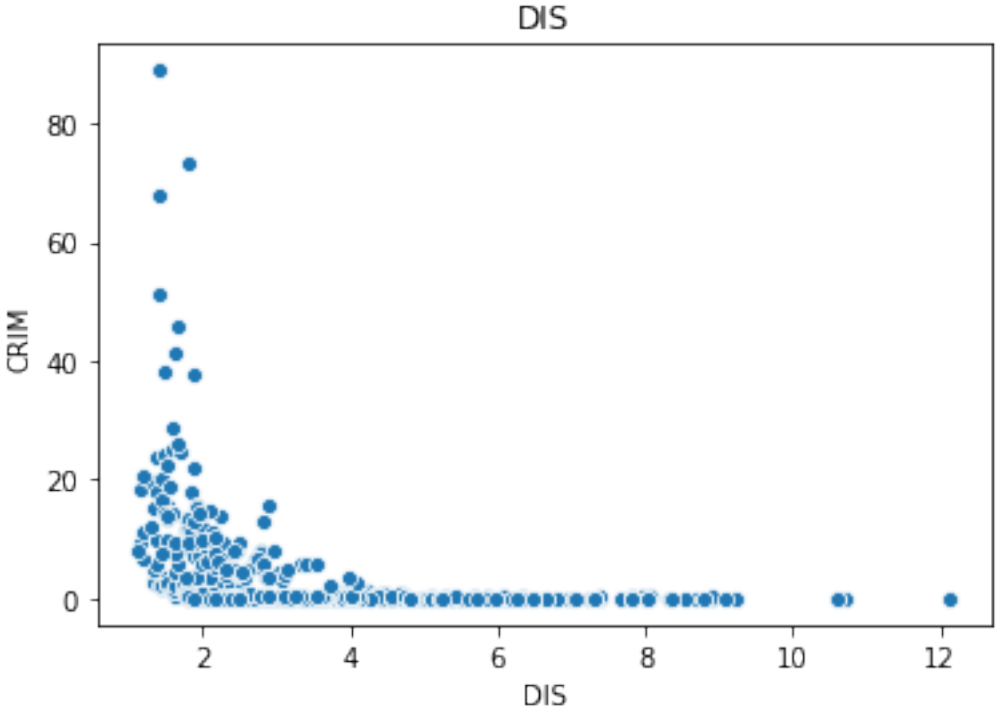

INDUS
CRIM
INDUS
0
20
40
60
80
0
5
10
15
20
25

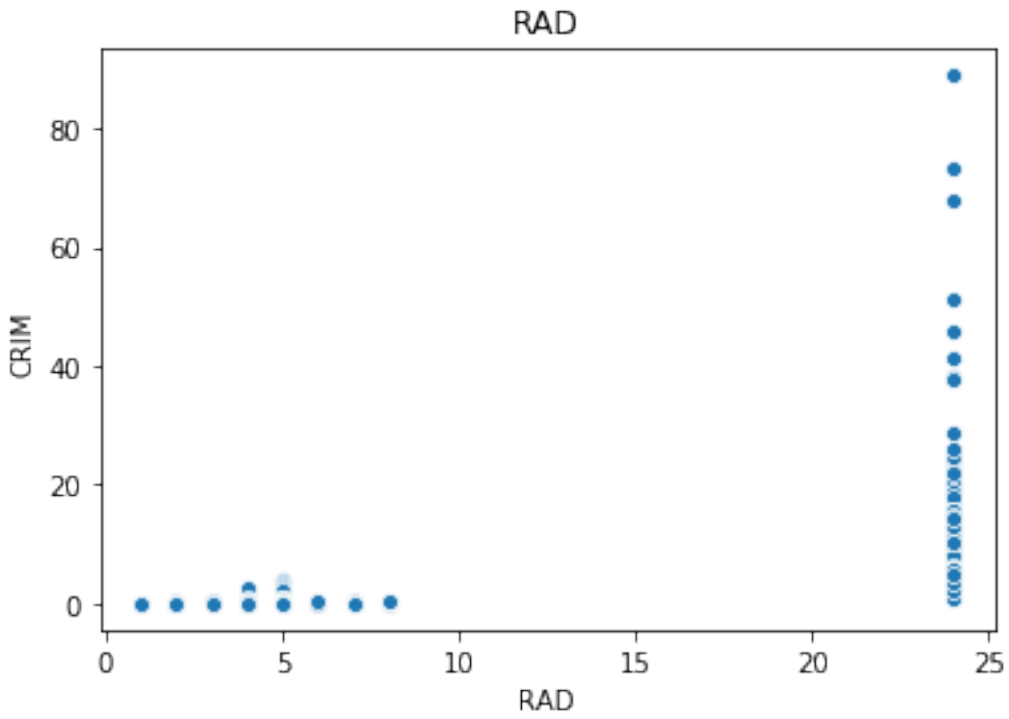

CHAS
CRIM
CHAS
0
20
40
60
80
0.0
0.2
0.4
0.6
0.8
1.0

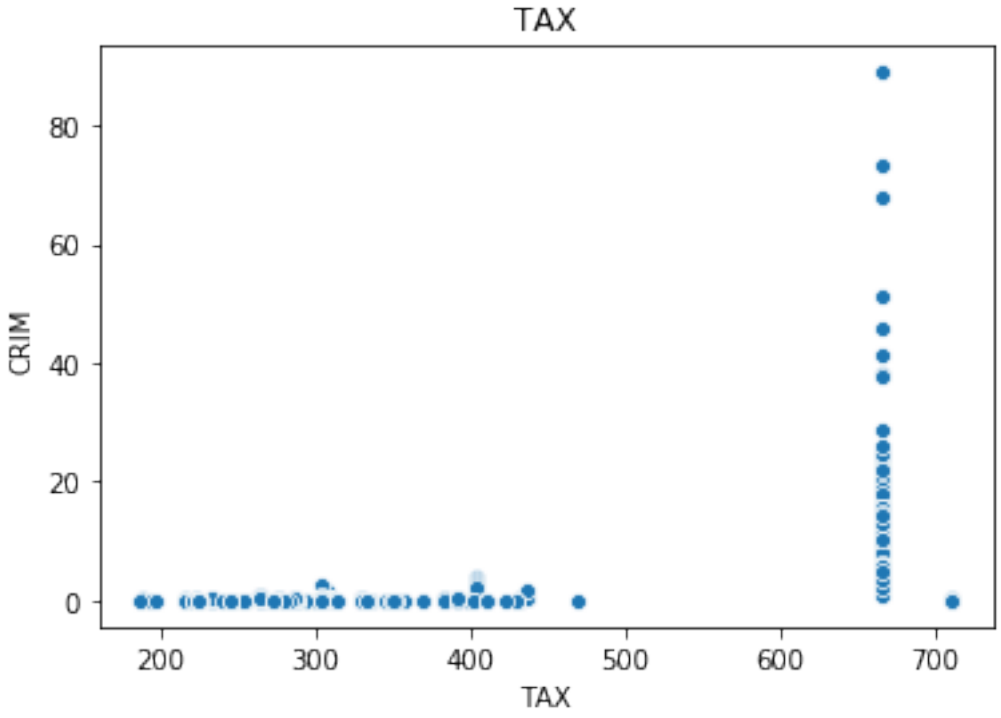

NOX
CRIM
NOX
0
20
40
60
80
0.4
0.5
0.6
0.7
0.8
0.9

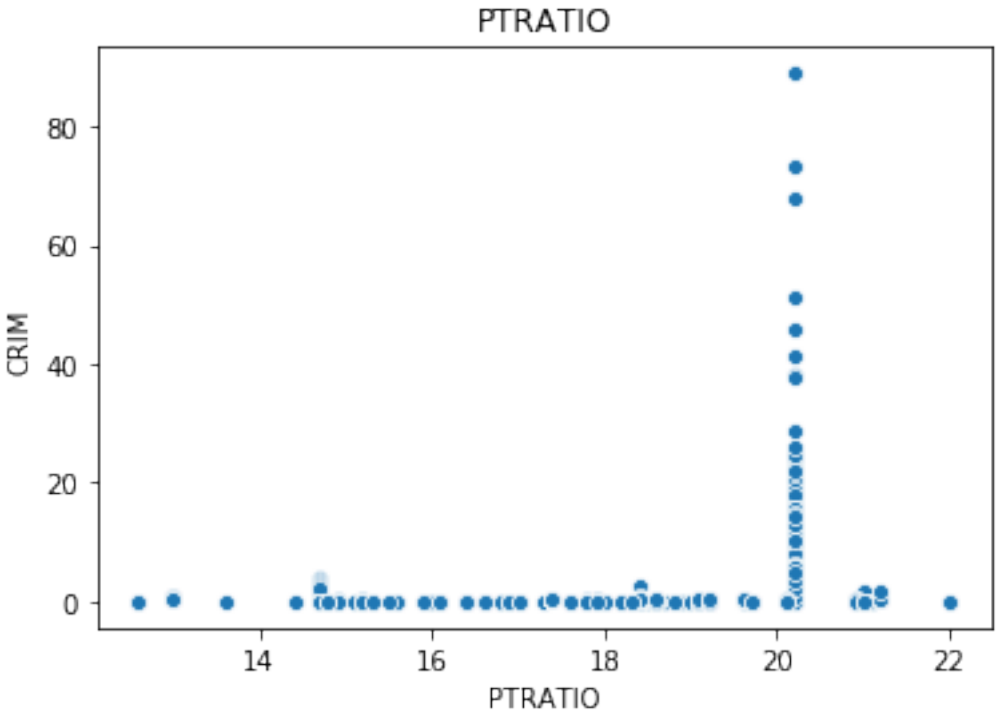

RM
CRIM
RM
0
20
40
60
80
4
5
6
7
8
9

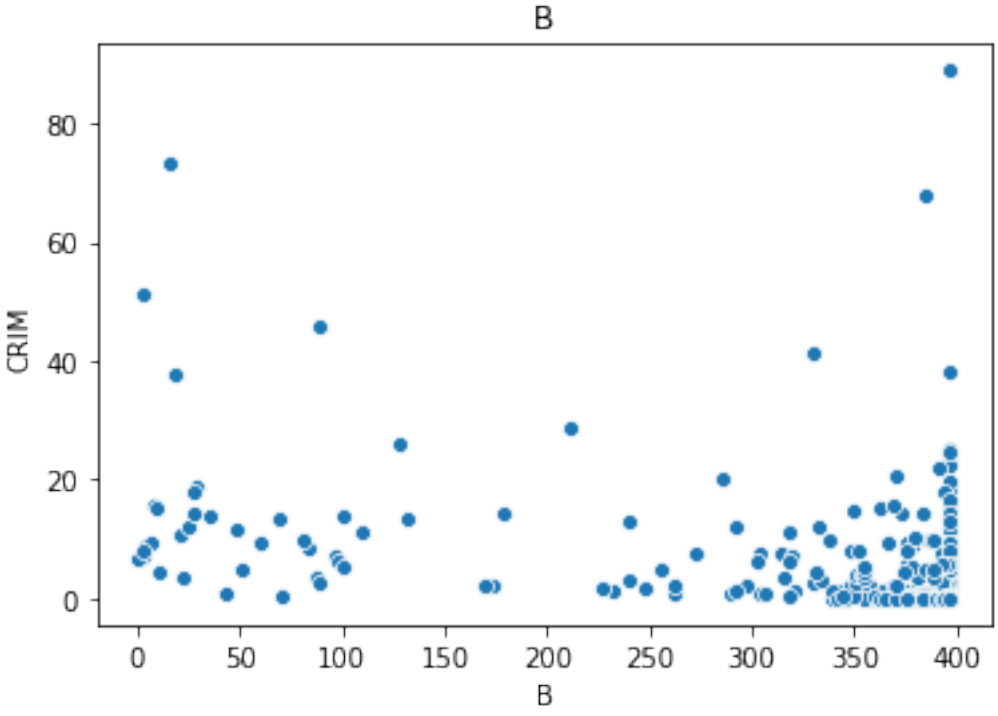

AGE
CRIM
AGE
0
20
40
60
80
100

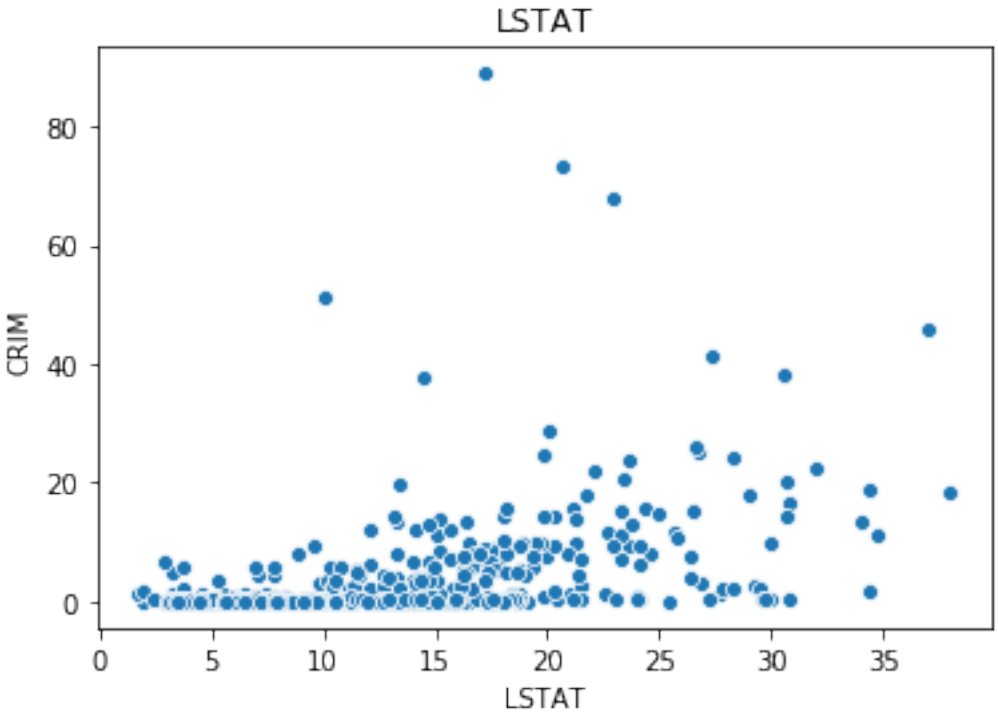

DIS
CRIM
DIS
0
20
40
60
80
2
4
6
8
10
12

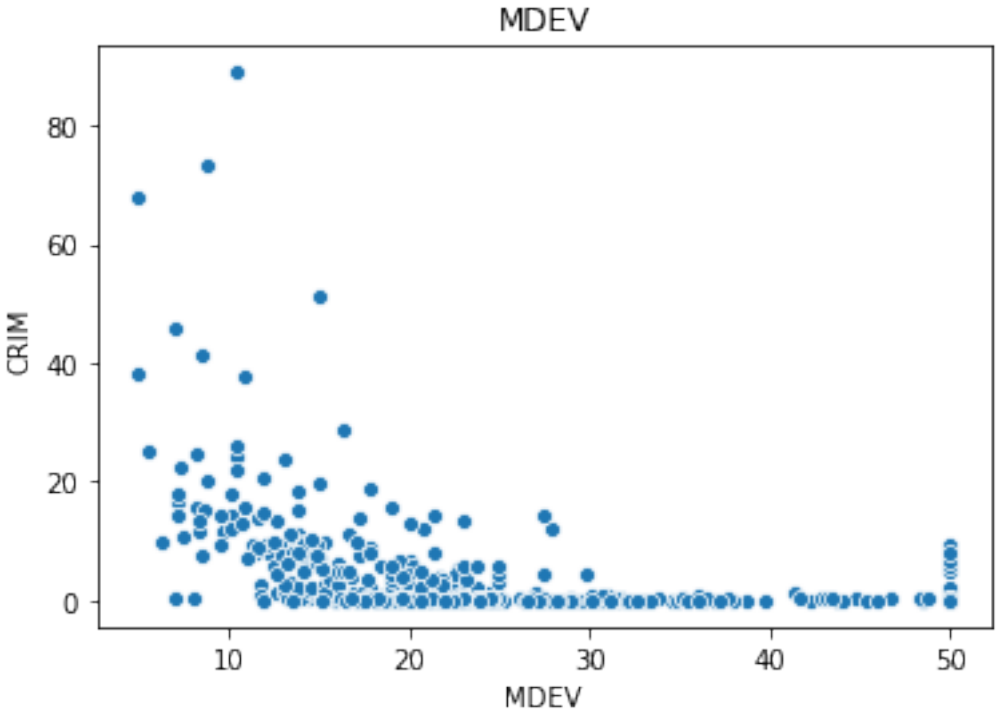

RAD
CRIM
RAD
0
20
40
60
80
0
5
10
15
20
25


TAX
CRIM
TAX
0
20
40
60
80
200
300
400
500
600
700


PTRATIO
CRIM
PTRATIO
0
20
40
60
80
14
16
18
20
22


B
CRIM
B
0
20
40
60
80
0
50
100
150
200
250
300
350
400


LSTAT
CRIM
LSTAT
0
20
40
60
80
0
5
10
15
20
25
30
35


MDEV
CRIM
MDEV
0
20
40
60
80
10
20
30
40
50

[32]:
```
predictors = ' + '.join(Boston_Housing.columns.difference(['CRIM']))
Multiple_Regression = smf.ols('CRIM ~ {}'.format(predictors),data =␣
 ↪Boston_Housing).fit()
print(Multiple_Regression.summary())
Multiple_Regression.params
```

OLS Regression Results

Dep. Variable:	CRIM	R-squared:	0.448
Model:	OLS	Adj. R-squared:	0.434
Method:	Least Squares	F-statistic:	30.73
Date:	Sat, 23 Jan 2021	Prob (F-statistic):	2.04e-55
Time:	21:44:23	Log-Likelihood:	-1655.7
No. Observations:	506	AIC:	3339.
Df Residuals:	492	BIC:	3399.
Df Model:	13		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	17.4184	7.270	2.396	0.017	3.135	31.702
AGE	0.0020	0.018	0.112	0.911	-0.033	0.037
B	-0.0069	0.004	-1.857	0.064	-0.014	0.000
CHAS	-0.7414	1.186	-0.625	0.532	-3.071	1.588
DIS	-0.9950	0.283	-3.514	0.000	-1.551	-0.439
INDUS	-0.0616	0.084	-0.735	0.463	-0.226	0.103
LSTAT	0.1213	0.076	1.594	0.112	-0.028	0.271
MDEV	-0.1992	0.061	-3.276	0.001	-0.319	-0.080
NOX	-10.6455	5.301	-2.008	0.045	-21.061	-0.230
PTRATIO	-0.2787	0.187	-1.488	0.137	-0.647	0.089
RAD	0.5888	0.088	6.656	0.000	0.415	0.763
RM	0.3811	0.616	0.619	0.536	-0.829	1.591
TAX	-0.0037	0.005	-0.723	0.470	-0.014	0.006
ZN	0.0449	0.019	2.386	0.017	0.008	0.082

Omnibus:	662.271	Durbin-Watson:	1.515
Prob(Omnibus):	0.000	Jarque-Bera (JB):	82701.666
Skew:	6.544	Prob(JB):	0.00
Kurtosis:	64.248	Cond. No.	1.58e+04

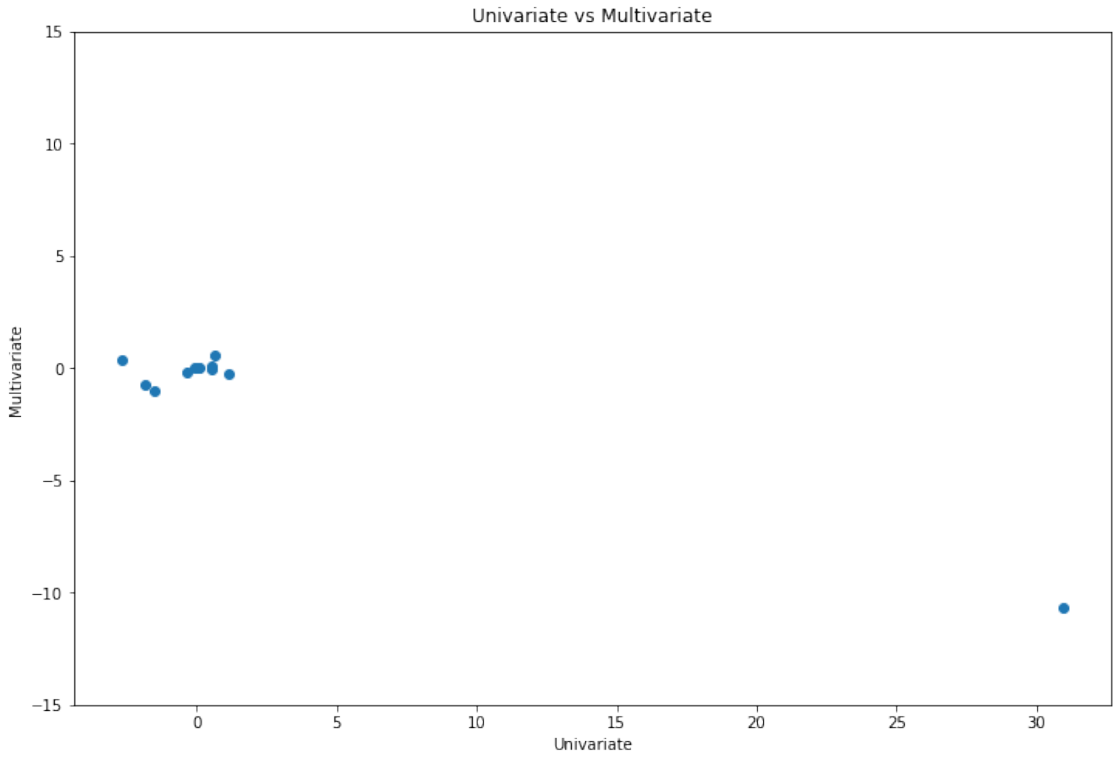
Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.58e+04. This might indicate that there are strong multicollinearity or other numerical problems.

[32]:
```
Intercept    17.418362
AGE           0.002011
B            -0.006855
CHAS         -0.741435
DIS          -0.994992
INDUS        -0.061576
LSTAT         0.121269
MDEV         -0.199218
NOX         -10.645500
PTRATIO      -0.278731
RAD           0.588838
RM            0.381070
TAX          -0.003746
ZN            0.044919
dtype: float64
```

[33]:
```
import matplotlib.pyplot as plt
X = [AGE.params[1], B.params[1], CHAS.params[1],DIS.params[1],INDUS.params[1],␣
 ↪LSTAT.params[1], MDEV.params[1], NOX.params[1], PTRATIO.params[1], RAD.
 ↪params[1], RM.params[1], TAX.params[1], ZN.params[1]]
y = Multiple_Regression.params[1::1]

#Source:https://stackoverflow.com/questions/52340435/
 ↪extract-regressions-coefficient-from-statsmodels
```

[34]:
```
plt.figure(figsize = (12,8))
plt.ylim(-15,15)
plt.scatter(X,y)
plt.title('Univariate vs Multivariate')
plt.xlabel("Univariate")
plt.ylabel("Multivariate")
```

[34]:
```
Text(0, 0.5, 'Multivariate')
```



[35]:
```
Boston_Housing
```

[35]:
```
        CRIM    ZN  INDUS  CHAS    NOX     RM   AGE     DIS  RAD    TAX  \
0    0.00632  18.0   2.31   0.0  0.538  6.575  65.2  4.0900  1.0  296.0
1    0.02731   0.0   7.07   0.0  0.469  6.421  78.9  4.9671  2.0  242.0
2    0.02729   0.0   7.07   0.0  0.469  7.185  61.1  4.9671  2.0  242.0
3    0.03237   0.0   2.18   0.0  0.458  6.998  45.8  6.0622  3.0  222.0
4    0.06905   0.0   2.18   0.0  0.458  7.147  54.2  6.0622  3.0  222.0
..       ...   ...    ...   ...    ...    ...   ...     ...  ...    ...
501  0.06263   0.0  11.93   0.0  0.573  6.593  69.1  2.4786  1.0  273.0
502  0.04527   0.0  11.93   0.0  0.573  6.120  76.7  2.2875  1.0  273.0
503  0.06076   0.0  11.93   0.0  0.573  6.976  91.0  2.1675  1.0  273.0
504  0.10959   0.0  11.93   0.0  0.573  6.794  89.3  2.3889  1.0  273.0
505  0.04741   0.0  11.93   0.0  0.573  6.030  80.8  2.5050  1.0  273.0

     PTRATIO       B  LSTAT  MDEV
0       15.3  396.90   4.98  24.0
1       17.8  396.90   9.14  21.6
2       17.8  392.83   4.03  34.7
3       18.7  394.63   2.94  33.4
4       18.7  396.90   5.33  36.2
..       ...     ...    ...   ...
501     21.0  391.99   9.67  22.4
```

```
502     21.0  396.90   9.08  20.6
503     21.0  396.90   5.64  23.9
504     21.0  393.45   6.48  22.0
505     21.0  396.90   7.88  11.9

[506 rows x 14 columns]
```

[37]:
```
import numpy as np
ZN2 = smf.ols(formula = 'CRIM ~ ZN + np.power(ZN,2) + np.power(ZN,3)', data =␣
 ↪Boston_Housing).fit()
INDUS2 = smf.ols(formula = 'CRIM ~ INDUS + np.power(INDUS,2) + np.
 ↪power(INDUS,3)', data = Boston_Housing).fit()
CHAS2 = smf.ols(formula = 'CRIM ~ CHAS + np.power(CHAS,2) + np.power(CHAS,3)',␣
 ↪data = Boston_Housing).fit()
NOX2 = smf.ols(formula = 'CRIM ~ NOX + np.power(NOX ,2) + np.power(NOX ,3)',␣
 ↪data = Boston_Housing).fit()
RM2 = smf.ols(formula = 'CRIM ~ RM + np.power(RM ,2) + np.power(RM ,3)', data =␣
 ↪Boston_Housing).fit()
AGE2 = smf.ols(formula = 'CRIM ~ AGE + np.power(AGE ,2) + np.power(AGE ,3)',␣
 ↪data = Boston_Housing).fit()
DIS2 = smf.ols(formula = 'CRIM ~ DIS + np.power(DIS,2) + np.power(DIS ,3)',␣
 ↪data = Boston_Housing).fit()
RAD2 = smf.ols(formula = 'CRIM ~ RAD + np.power(RAD ,2) + np.power(RAD ,3)',␣
 ↪data = Boston_Housing).fit()
TAX2 = smf.ols(formula = 'CRIM ~ TAX + np.power(TAX ,2) + np.power(TAX ,3)',␣
 ↪data = Boston_Housing).fit()
PTRATIO2 = smf.ols(formula = 'CRIM ~ PTRATIO + np.power(PTRATIO ,2) + np.
 ↪power(PTRATIO ,3)', data = Boston_Housing).fit()
B2 = smf.ols(formula = 'CRIM ~ B + np.power(B ,2) + np.power(B ,3)', data =␣
 ↪Boston_Housing).fit()
LSTAT2 = smf.ols(formula = 'CRIM ~ LSTAT + np.power(LSTAT ,2) + np.power(LSTAT␣
 ↪,3)', data = Boston_Housing).fit()
MDEV2 = smf.ols(formula = 'CRIM ~ MDEV + np.power(MDEV ,2) + np.power(MDEV␣
 ↪,3)', data = Boston_Housing).fit()
```

[38]:
```
ZN2.summary()
```

[38]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.058
Model:                            OLS   Adj. R-squared:                  0.052
Method:                 Least Squares   F-statistic:                     10.24
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           1.49e-06
Time:                        21:45:41   Log-Likelihood:                -1791.1
```

```
No. Observations:                 506   AIC:                             3590.
Df Residuals:                     502   BIC:                             3607.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
===
                     coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
---
Intercept           4.8193      0.433     11.133      0.000       3.969
5.670
ZN                 -0.3303      0.110     -3.008      0.003      -0.546
-0.115
np.power(ZN, 2)     0.0064      0.004      1.670      0.096      -0.001
0.014
np.power(ZN, 3) -3.753e-05   3.14e-05     -1.196      0.232   -9.92e-05
2.41e-05
==============================================================================
Omnibus:                      570.003   Durbin-Watson:                   0.879
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            33886.468
Skew:                           5.285   Prob(JB):                         0.00
Kurtosis:                      41.672   Cond. No.                     1.89e+05
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 1.89e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[39]:
```
INDUS2.summary()
```

[39]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.257
Model:                            OLS   Adj. R-squared:                  0.252
Method:                 Least Squares   F-statistic:                     57.86
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           3.88e-32
Time:                        21:46:17   Log-Likelihood:                -1731.0
No. Observations:                 506   AIC:                             3470.
Df Residuals:                     502   BIC:                             3487.
Df Model:                           3
Covariance Type:            nonrobust
```

```
==============================================================================
======
                 coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
------
Intercept      3.6410      1.576      2.310      0.021       0.545
6.737
INDUS         -1.9533      0.483     -4.047      0.000      -2.901
-1.005
np.power(INDUS, 2)     0.2504      0.039      6.361      0.000       0.173
0.328
np.power(INDUS, 3)    -0.0069      0.001     -7.239      0.000      -0.009
-0.005
==============================================================================
Omnibus:                      611.416   Durbin-Watson:                   1.118
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            51547.097
Skew:                           5.815   Prob(JB):                         0.00
Kurtosis:                      51.059   Cond. No.                     2.47e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 2.47e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[40]:
```
CHAS2.summary()
```

[40]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.003
Model:                            OLS   Adj. R-squared:                  0.001
Method:                 Least Squares   F-statistic:                     1.546
Date:                Sat, 23 Jan 2021   Prob (F-statistic):              0.214
Time:                        21:46:33   Log-Likelihood:                -1805.3
No. Observations:                 506   AIC:                             3615.
Df Residuals:                     504   BIC:                             3623.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
=====
                 coef    std err          t      P>|t|      [0.025
0.975]
```

```
------------------------------------------------------------------------------
-----
Intercept             3.7232      0.396      9.404      0.000       2.945
4.501
CHAS                 -0.6238      0.502     -1.243      0.214      -1.610
0.362
np.power(CHAS, 2)    -0.6238      0.502     -1.243      0.214      -1.610
0.362
np.power(CHAS, 3)    -0.6238      0.502     -1.243      0.214      -1.610
0.362
==============================================================================
Omnibus:                      562.698   Durbin-Watson:                   0.822
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            30864.755
Skew:                           5.205   Prob(JB):                         0.00
Kurtosis:                      39.818   Cond. No.                     3.42e+32
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The smallest eigenvalue is 4.39e-63. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
"""
```

[41]:
```
NOX2.summary()
```

[41]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.292
Model:                            OLS   Adj. R-squared:                  0.288
Method:                 Least Squares   F-statistic:                     69.14
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           1.94e-37
Time:                        21:46:36   Log-Likelihood:                -1718.6
No. Observations:                 506   AIC:                             3445.
Df Residuals:                     502   BIC:                             3462.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
====
                     coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
----
Intercept          230.1421     33.734      6.822      0.000     163.864
296.420
```

```
NOX              -1264.1021    170.860     -7.398      0.000   -1599.791
-928.414
np.power(NOX, 2)  2223.2265    280.659      7.921      0.000    1671.816
2774.637
np.power(NOX, 3) -1232.3894    149.687     -8.233      0.000   -1526.479
-938.300
==============================================================================
Omnibus:                      612.604   Durbin-Watson:                   1.159
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            52872.508
Skew:                           5.824   Prob(JB):                         0.00
Kurtosis:                      51.705   Cond. No.                     1.36e+03
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 1.36e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[42]:
```
RM2.summary()
```

[42]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.068
Model:                            OLS   Adj. R-squared:                  0.063
Method:                 Least Squares   F-statistic:                     12.29
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           9.06e-08
Time:                        21:46:42   Log-Likelihood:                -1788.2
No. Observations:                 506   AIC:                             3584.
Df Residuals:                     502   BIC:                             3601.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
===
                     coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
---
Intercept         111.9002     64.460      1.736      0.083     -14.744
238.545
RM                -38.7040     31.284     -1.237      0.217    -100.167
22.759
np.power(RM, 2)     4.4655      5.005      0.892      0.373      -5.369
14.300
```

```
np.power(RM, 3)    -0.1694      0.264     -0.643      0.521      -0.687
0.348
==============================================================================
Omnibus:                      586.445   Durbin-Watson:                   0.919
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            40548.719
Skew:                           5.484   Prob(JB):                         0.00
Kurtosis:                      45.461   Cond. No.                     5.36e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 5.36e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[43]:
```
AGE2.summary()
```

[43]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.172
Model:                            OLS   Adj. R-squared:                  0.167
Method:                 Least Squares   F-statistic:                     34.86
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           1.76e-20
Time:                        21:46:44   Log-Likelihood:                -1758.2
No. Observations:                 506   AIC:                             3524.
Df Residuals:                     502   BIC:                             3541.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
====
                     coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
----
Intercept          -2.5592      2.771     -0.924      0.356      -8.003
2.884
AGE                 0.2743      0.186      1.471      0.142      -0.092
0.641
np.power(AGE, 2)   -0.0072      0.004     -1.987      0.047      -0.014
-8.25e-05
np.power(AGE, 3)  5.737e-05   2.11e-05      2.719      0.007    1.59e-05
9.88e-05
==============================================================================
Omnibus:                      577.859   Durbin-Watson:                   1.027
```

```
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            39629.126
Skew:                           5.342   Prob(JB):                         0.00
Kurtosis:                      45.018   Cond. No.                     4.74e+06
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 4.74e+06. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[44]:
```
DIS2.summary()
```

[44]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.276
Model:                            OLS   Adj. R-squared:                  0.272
Method:                 Least Squares   F-statistic:                     63.74
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           6.20e-35
Time:                        21:46:51   Log-Likelihood:                -1724.4
No. Observations:                 506   AIC:                             3457.
Df Residuals:                     502   BIC:                             3474.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
====
                     coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
----
Intercept          29.9496      2.448     12.235      0.000      25.140
34.759
DIS               -15.5172      1.737     -8.931      0.000     -18.931
-12.104
np.power(DIS, 2)    2.4479      0.347      7.061      0.000       1.767
3.129
np.power(DIS, 3)   -0.1185      0.020     -5.802      0.000      -0.159
-0.078
==============================================================================
Omnibus:                      577.986   Durbin-Watson:                   1.133
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            42441.952
Skew:                           5.310   Prob(JB):                         0.00
Kurtosis:                      46.592   Cond. No.                     2.10e+03
==============================================================================
```

```
Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 2.1e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[45]:
```
RAD2.summary()
```

[45]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.396
Model:                            OLS   Adj. R-squared:                  0.392
Method:                 Least Squares   F-statistic:                     109.5
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           1.47e-54
Time:                        21:46:58   Log-Likelihood:                -1678.7
No. Observations:                 506   AIC:                             3365.
Df Residuals:                     502   BIC:                             3382.
Df Model:                           3
Covariance Type:            nonrobust
================================================================================
====
                     coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------------
----
Intercept           -0.6050      2.057     -0.294      0.769      -4.645
3.435
RAD                  0.5122      1.047      0.489      0.625      -1.545
2.569
np.power(RAD, 2)    -0.0750      0.149     -0.504      0.615      -0.368
0.218
np.power(RAD, 3)     0.0032      0.005      0.699      0.485      -0.006
0.012
==============================================================================
Omnibus:                      657.375   Durbin-Watson:                   1.349
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            76643.757
Skew:                           6.487   Prob(JB):                         0.00
Kurtosis:                      61.881   Cond. No.                     5.43e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

```
[2] The condition number is large, 5.43e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[46]:
```
TAX2.summary()
```

[46]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.365
Model:                            OLS   Adj. R-squared:                  0.361
Method:                 Least Squares   F-statistic:                     96.10
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           3.69e-49
Time:                        21:47:03   Log-Likelihood:                -1691.3
No. Observations:                 506   AIC:                             3391.
Df Residuals:                     502   BIC:                             3407.
Df Model:                           3
Covariance Type:            nonrobust
================================================================================
====
                     coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------------
----
Intercept           19.0705     11.827      1.612      0.107      -4.166
42.307
TAX                 -0.1524      0.096     -1.589      0.113      -0.341
0.036
np.power(TAX, 2)     0.0004      0.000      1.476      0.141      -0.000
0.001
np.power(TAX, 3) -2.193e-07   1.89e-07     -1.158      0.247   -5.91e-07
1.53e-07
==============================================================================
Omnibus:                      642.369   Durbin-Watson:                   1.292
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            68905.900
Skew:                           6.249   Prob(JB):                         0.00
Kurtosis:                      58.786   Cond. No.                     6.16e+09
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 6.16e+09. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[47]:
```
PTRATIO2.summary()
```

[47]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.112
Model:                            OLS   Adj. R-squared:                  0.107
Method:                 Least Squares   F-statistic:                     21.21
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           5.99e-13
Time:                        21:47:07   Log-Likelihood:                -1775.9
No. Observations:                 506   AIC:                             3560.
Df Residuals:                     502   BIC:                             3577.
Df Model:                           3
Covariance Type:            nonrobust
================================================================================
========
                           coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------------
--------
Intercept              474.0255    156.823      3.023      0.003     165.915
782.135
PTRATIO                -81.8089     27.649     -2.959      0.003    -136.131
-27.487
np.power(PTRATIO, 2)     4.6039      1.609      2.862      0.004       1.444
7.764
np.power(PTRATIO, 3)    -0.0842      0.031     -2.724      0.007      -0.145
-0.023
==============================================================================
Omnibus:                      572.978   Durbin-Watson:                   0.949
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            36189.609
Skew:                           5.303   Prob(JB):                         0.00
Kurtosis:                      43.050   Cond. No.                     3.02e+06
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 3.02e+06. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[48]:
```
B2.summary()
```

[48]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.144
Model:                            OLS   Adj. R-squared:                  0.139
Method:                 Least Squares   F-statistic:                     28.14
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           7.83e-17
Time:                        21:47:08   Log-Likelihood:                -1766.8
No. Observations:                 506   AIC:                             3542.
Df Residuals:                     502   BIC:                             3558.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
==
                     coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
--
Intercept         17.9898      2.312      7.782      0.000      13.448
22.531
B                 -0.0845      0.056     -1.497      0.135      -0.196
0.026
np.power(B, 2)     0.0002      0.000      0.760      0.447      -0.000
0.001
np.power(B, 3) -2.895e-07   4.38e-07     -0.661      0.509   -1.15e-06
5.7e-07
==============================================================================
Omnibus:                      589.534   Durbin-Watson:                   0.990
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            42752.655
Skew:                           5.512   Prob(JB):                         0.00
Kurtosis:                      46.661   Cond. No.                     3.59e+08
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 3.59e+08. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[49]:
```
LSTAT2.summary()
```

[49]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.214
Model:                            OLS   Adj. R-squared:                  0.210
```

```
Method:                 Least Squares   F-statistic:                     45.67
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           4.13e-26
Time:                        21:47:12   Log-Likelihood:                -1745.0
No. Observations:                 506   AIC:                             3498.
Df Residuals:                     502   BIC:                             3515.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      1.0836      2.032      0.533      0.594      -2.909       5.076
LSTAT         -0.4133      0.466     -0.887      0.375      -1.328       0.502
np.power(LSTAT, 2)     0.0530      0.030      1.758      0.079      -0.006       0.112
np.power(LSTAT, 3)    -0.0008      0.001     -1.423      0.155      -0.002       0.000
==============================================================================
Omnibus:                      607.032   Durbin-Watson:                   1.239
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            53255.699
Skew:                           5.717   Prob(JB):                         0.00
Kurtosis:                      51.941   Cond. No.                     5.20e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.2e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```

[50]:
```
MDEV2.summary()
```

[50]:
```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:                   CRIM   R-squared:                       0.416
Model:                            OLS   Adj. R-squared:                  0.413
Method:                 Least Squares   F-statistic:                     119.2
Date:                Sat, 23 Jan 2021   Prob (F-statistic):           2.65e-58
Time:                        21:47:15   Log-Likelihood:                -1670.0
No. Observations:                 506   AIC:                             3348.
```

```
Df Residuals:                      502   BIC:                             3365.
Df Model:                            3
Covariance Type:            nonrobust
==============================================================================
=====
                         coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
-----
Intercept            52.9386      3.366     15.725      0.000      46.325
59.553
MDEV                 -5.0774      0.435    -11.668      0.000      -5.932
-4.222
np.power(MDEV, 2)     0.1551      0.017      8.995      0.000       0.121
0.189
np.power(MDEV, 3)    -0.0015      0.000     -7.277      0.000      -0.002
-0.001
==============================================================================
Omnibus:                      568.100   Durbin-Watson:                   1.360
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            47296.533
Skew:                           5.084   Prob(JB):                         0.00
Kurtosis:                      49.259   Cond. No.                     3.67e+05
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 3.67e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
"""
```