

```
In [2]: from sklearn.datasets import load_boston
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
```

```
In [3]: boston = load_boston()
```

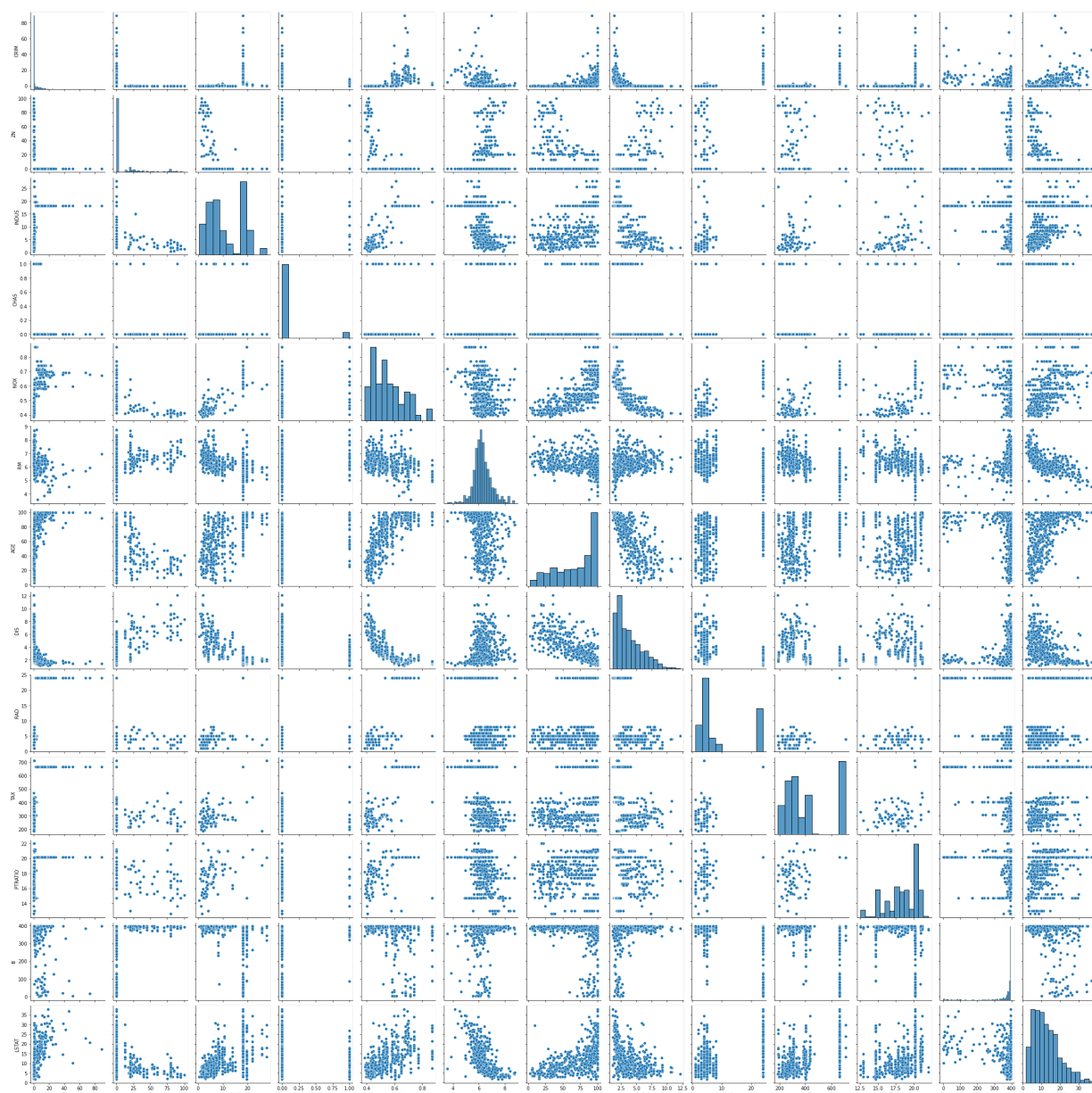
```
In [4]: data = pd.DataFrame(boston.data, columns = boston.feature_names)
```

```
In [5]: data.shape
```

```
Out[5]: (506, 13)
```

```
In [6]: sns.pairplot(data)
```

```
Out[6]: <seaborn.axisgrid.PairGrid at 0x24317421760>
```



```
In [14]: corr_tab = data.corr()
corr_tab.iloc[[0]]
```

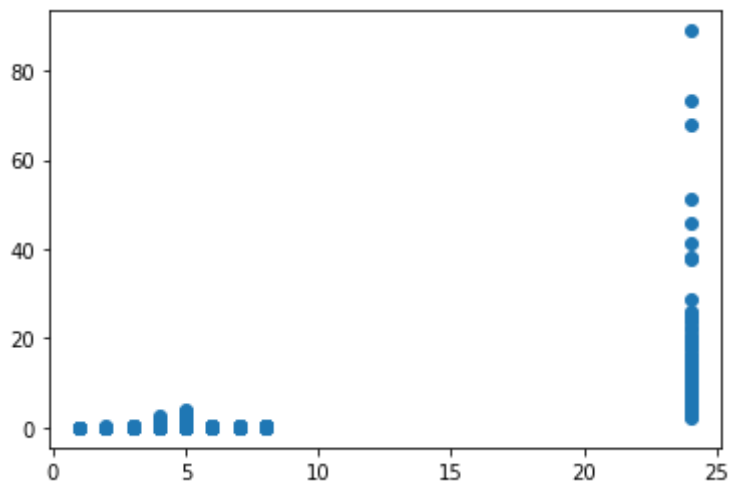
```
Out[14]:
```

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT
CRIM	1.000000	-0.103210	0.006614	0.000000	0.000000	-0.403254	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD
CRIM	1.0	-0.200469	0.406583	-0.055892	0.420972	-0.219247	0.352734	-0.37967	0.625505

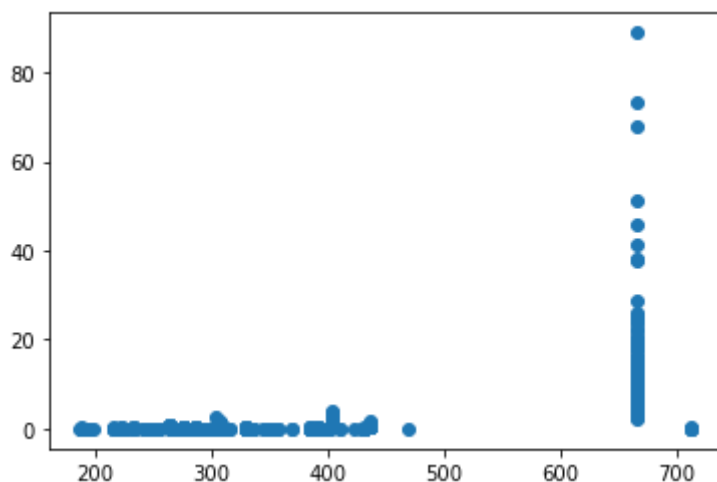
```
In [15]: plt.scatter(data['RAD'],data['CRIM'])
```

```
Out[15]: <matplotlib.collections.PathCollection at 0x24324c736a0>
```



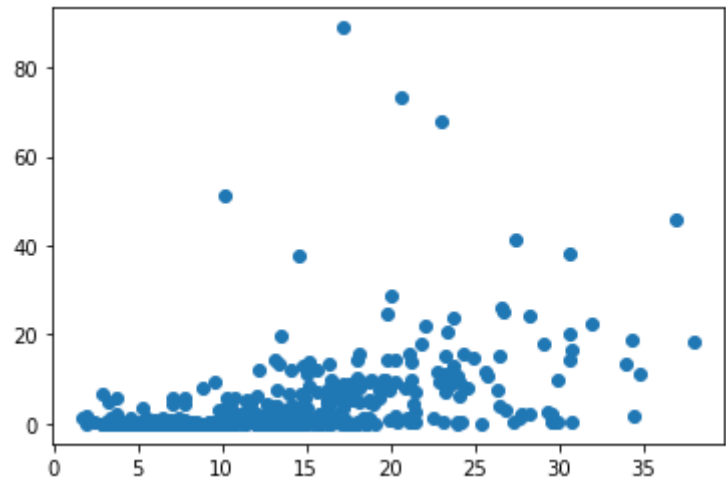
```
In [16]: plt.scatter(data['TAX'],data['CRIM'])
```

```
Out[16]: <matplotlib.collections.PathCollection at 0x243234f6ca0>
```



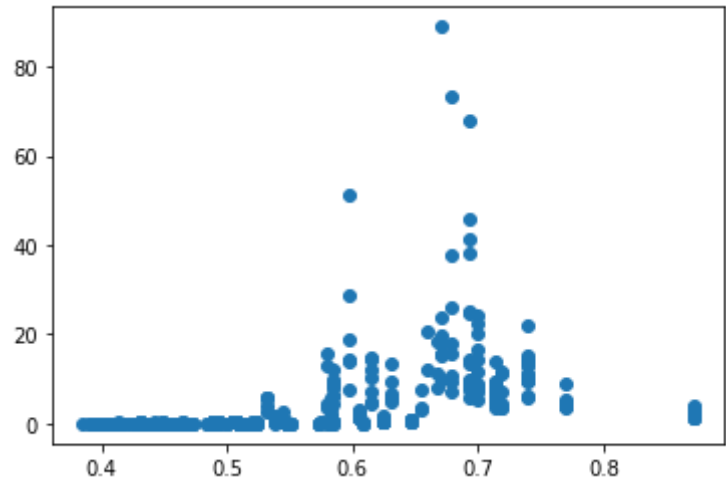
```
In [17]: plt.scatter(data['LSTAT'],data['CRIM'])
```

```
Out[17]: <matplotlib.collections.PathCollection at 0x24324ce7940>
```



```
In [18]: plt.scatter(data['NOX'],data['CRIM'])
```

Out[18]: <matplotlib.collections.PathCollection at 0x24324d32f70>

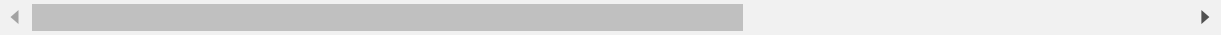


```
In [46]: summary = data[data['CRIM']>0]
len(summary)
```

Out[46]: 506

```
In [47]: summary.describe()
```

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	D
count	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000	506.000000
mean	3.613524	11.363636	11.136779	0.069170	0.554695	6.284634	68.574901	3.795000
std	8.601545	23.322453	6.860353	0.253994	0.115878	0.702617	28.148861	2.105700
min	0.006320	0.000000	0.460000	0.000000	0.385000	3.561000	2.900000	1.129600
25%	0.082045	0.000000	5.190000	0.000000	0.449000	5.885500	45.025000	2.100100
50%	0.256510	0.000000	9.690000	0.000000	0.538000	6.208500	77.500000	3.207400
75%	3.677083	12.500000	18.100000	0.000000	0.624000	6.623500	94.075000	5.188400
max	88.976200	100.000000	27.740000	1.000000	0.871000	8.780000	100.000000	12.126500

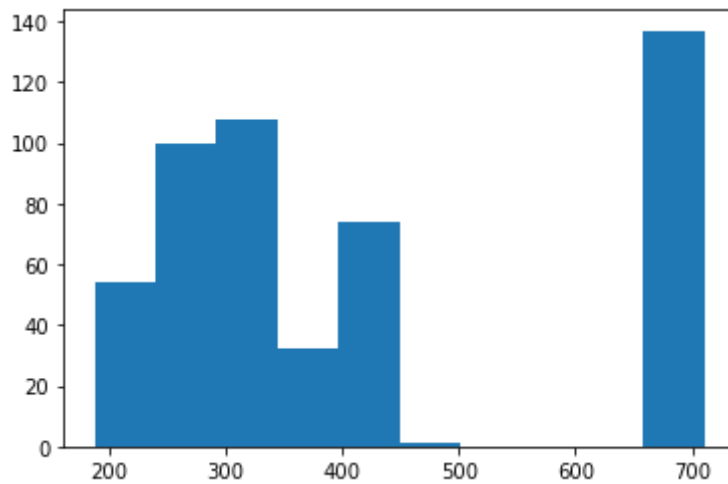


```
In [48]: summary = data[data['CRIM']>20]
len(summary)
```

Out[48]: 18

```
In [53]: plt.hist(data['TAX'])
```

```
Out[53]: (array([ 54., 100., 108.,  32.,  74.,   1.,   0.,   0.,   0., 137.]),  
          array([187. , 239.4, 291.8, 344.2, 396.6, 449. , 501.4, 553.8, 606.2,  
                658.6, 711. ]),  
          <BarContainer object of 10 artists>)
```



```
In [55]: tax = data[data['TAX']>500]  
len(tax)
```

Out[55]: 137

```
In [56]: plt.hist(data['PTRATIO'])
```

```
Out[56]: (array([ 15.,   2.,  58.,  15.,  35.,  69.,  76.,  40., 178.,  18.]),  
          array([12.6 , 13.54, 14.48, 15.42, 16.36, 17.3 , 18.24, 19.18, 20.12,  
                21.06, 22. ]),  
          <BarContainer object of 10 artists>)
```

