Chapter 4 - Practical Data Visualization

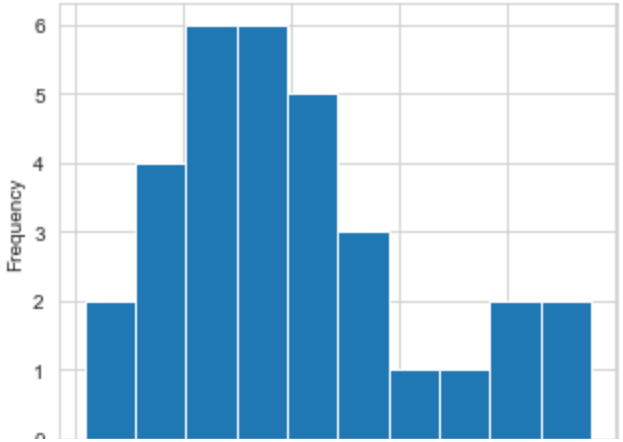
Segment 6 - Creating statistical data graphics

```
import numpy as np
import pandas as pd
from pandas import Series, DataFrame

from pandas.plotting import scatter_matrix

import matplotlib.pyplot as plt
from pylab import rcParams


%matplotlib inline
rcParams['figure.figsize'] = 5, 4


import seaborn as sb
sb.set_style('whitegrid')
```

Eyeballing dataset distributions with histograms

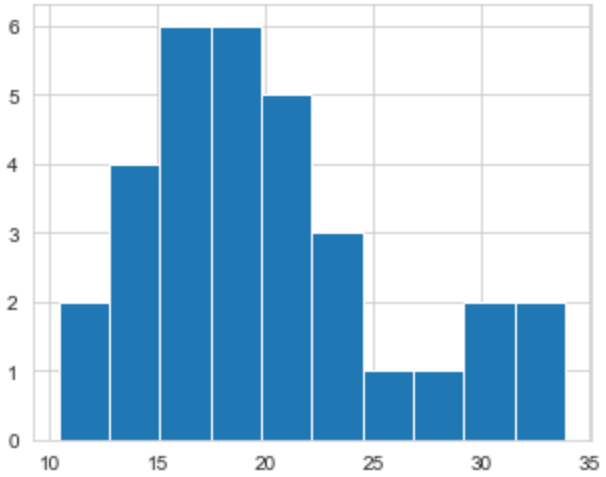
```
address = 'C:/Users/Lillian/Desktop/ExerciseFiles/Data/mtcars.csv'

cars = pd.read_csv(address)

cars.columns = ['car_names','mpg','cyl','disp', 'hp', 'drat', 'wt', 'qsec', 'vs', 'am', 'gear', 'carb']

cars.index = cars.car_names

mpg = cars['mpg']

mpg.plot(kind='hist')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x1e33b645240>
```



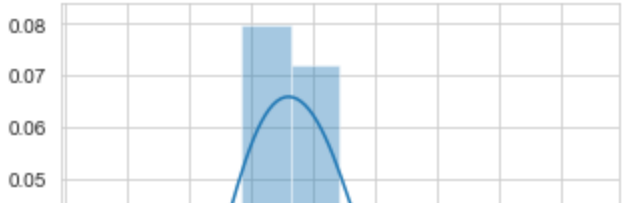

```
plt.hist(mpg)
plt.plot()
```

```
[]
```




```
sb.distplot(mpg)
```

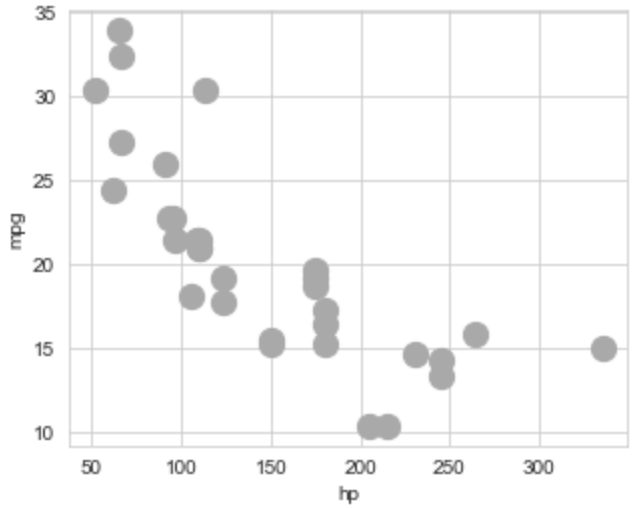
```
<matplotlib.axes._subplots.AxesSubplot at 0x1e33c9ced30>
```



Seeing scatterplots in action

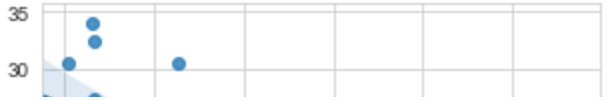
```
cars.plot(kind='scatter', x='hp', y='mpg', c=['darkgray'], s=150)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x1e33ca78fd0>
```




```
sb.regplot(x='hp', y='mpg', data=cars, scatter=True)
```

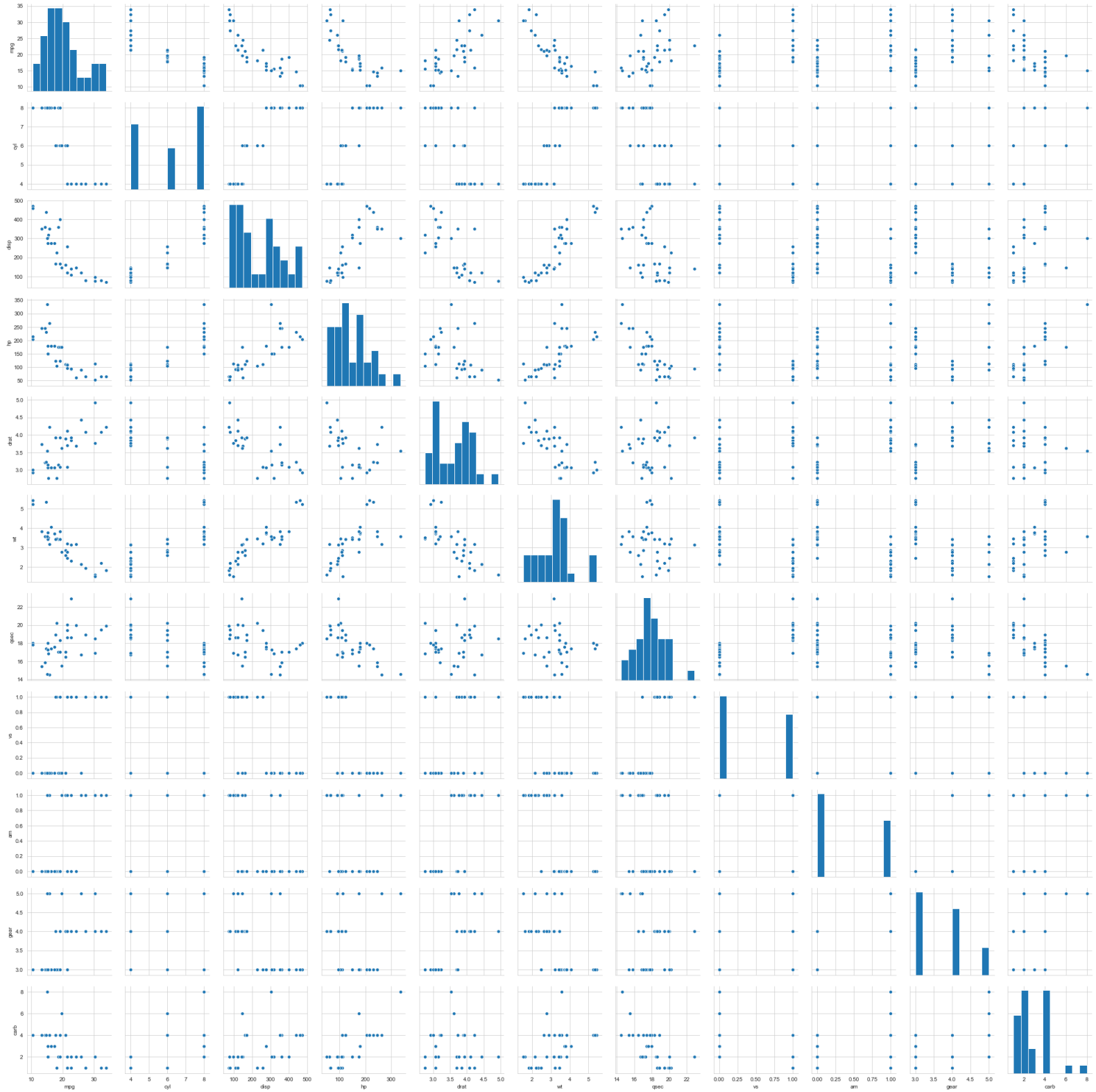
```
<matplotlib.axes._subplots.AxesSubplot at 0x1e33cad2390>
```



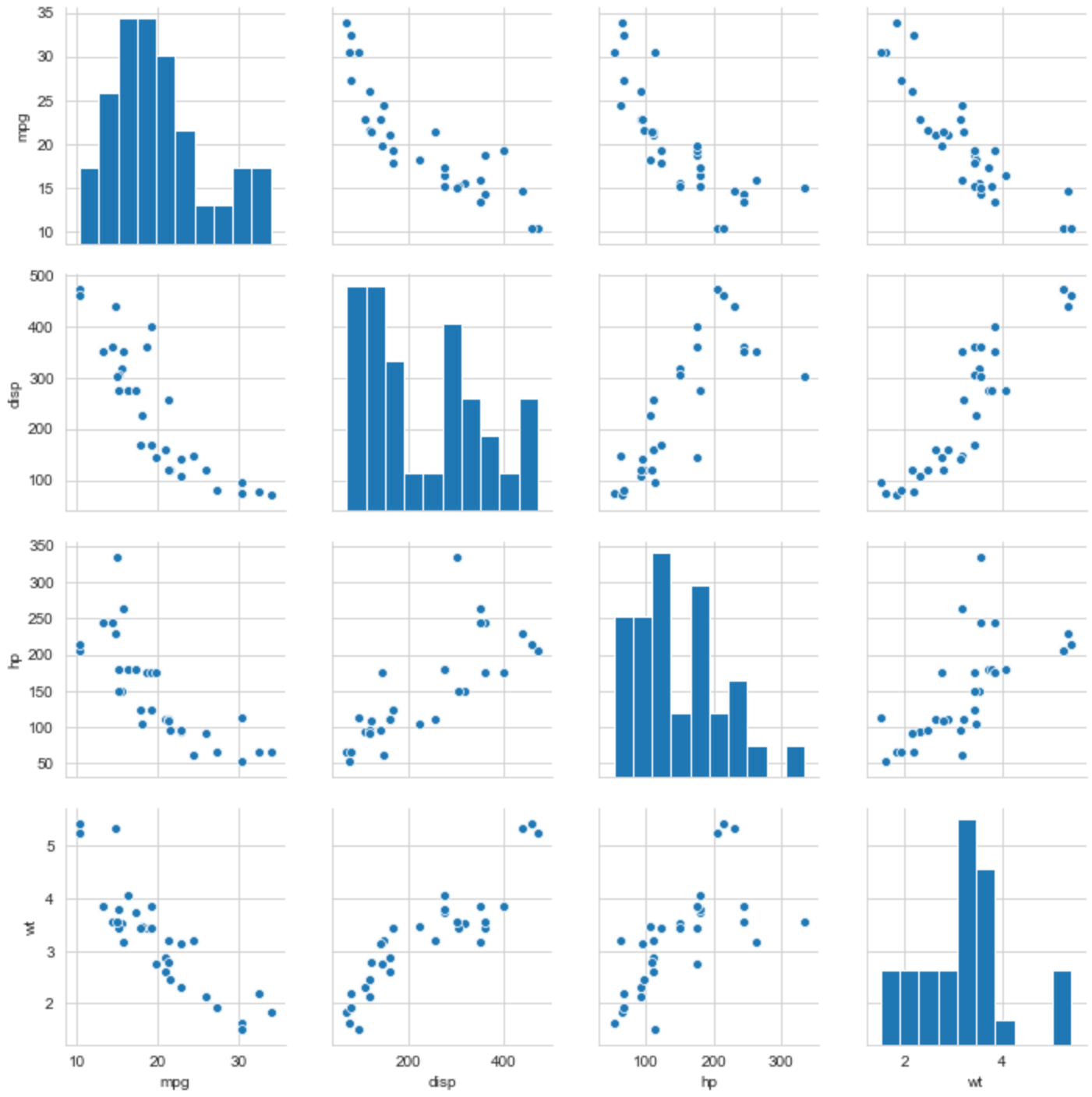
Generating a scatter plot matrix

```
sb.pairplot(cars)
```

```
<seaborn.axisgrid.PairGrid at 0x1e33ca52c50>
```



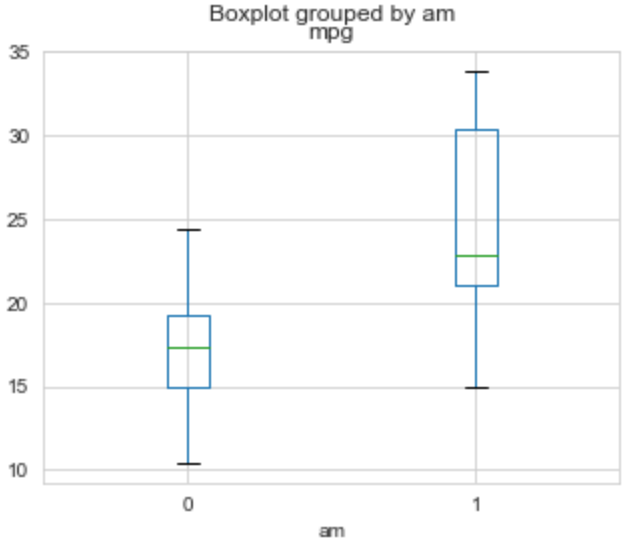
```
cars_subset = cars[['mpg', 'disp', 'hp', 'wt']]
sb.pairplot(cars_subset)
plt.show()
```

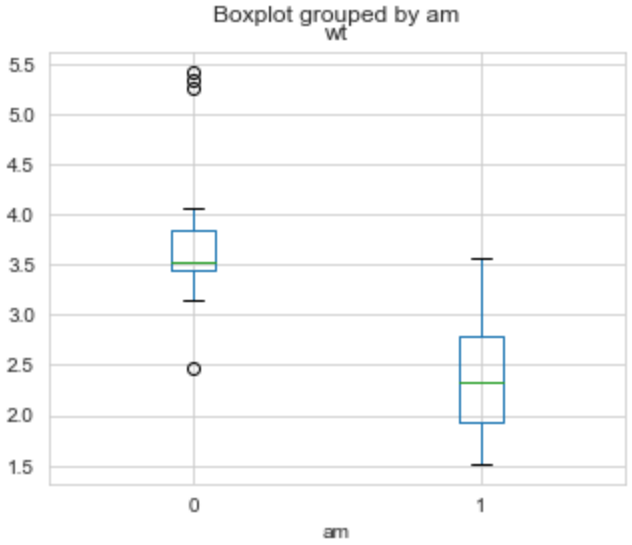


Building boxplots

```
cars.boxplot(column='mpg', by='am')
cars.boxplot(column='wt', by='am')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x1e3425f4cc0>
```






```
sb.boxplot(x='am', y='mpg', data=cars, palette='hls')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x1e342a5cdd8>
```