```
Chapter 5 - Basic Math and Statistics

Segement 5 - Starting with parametric methods in pandas an dscipy
```

In [1]:
```
import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
import seaborn as sb
from pylab import rcParams

import scipy
from scipy.stats.stats import pearsonr
```

In [7]:
```
%matplotlib inline
rcParams['figure.figsize'] = 8,4
plt.style.use('seaborn-whitegrid')
```

Type *Markdown* and LaTeX: α^2

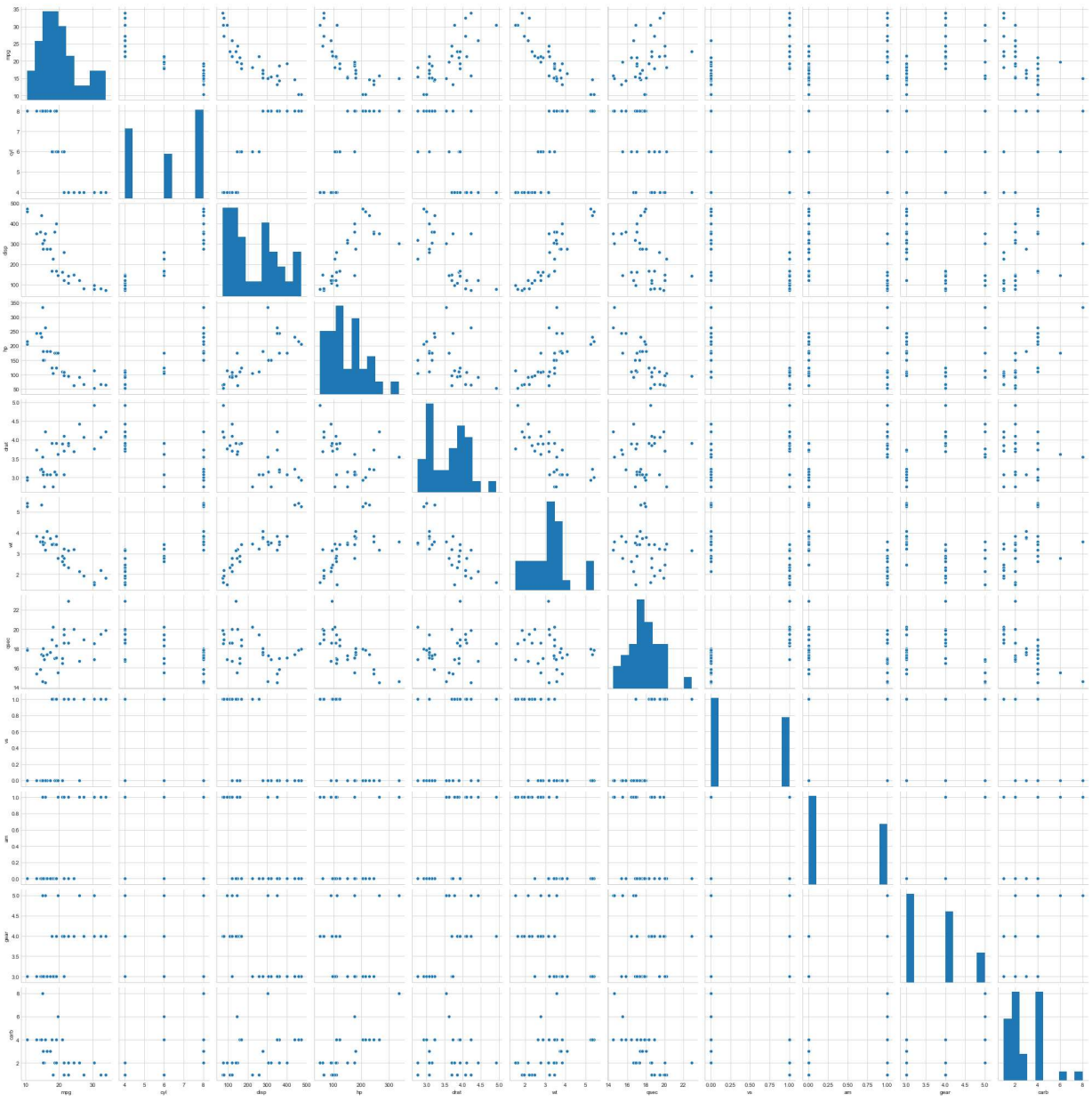
```
The Pearson Correlation
```

In [8]:
```
address = 'C:/Users/danal/Desktop/ExerciseFiles/Data/mtcars.csv'

cars = pd.read_csv(address)
cars.columns = ['car_names', 'mpg', 'cyl', 'disp', 'hp', 'drat', 'wt', 'qsec', 'v
```

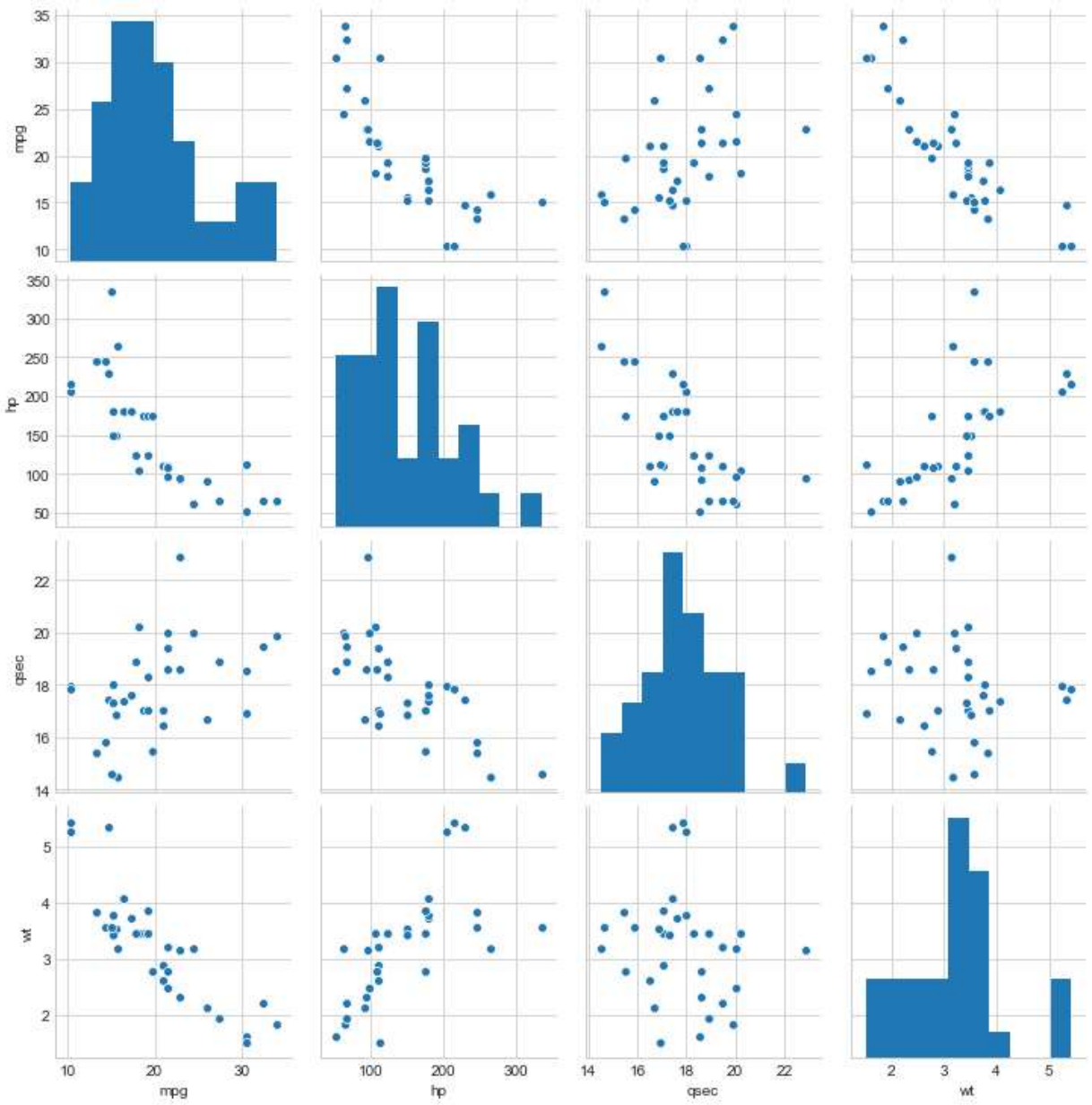
In [9]:
```
sb.pairplot(cars)
```

Out[9]:
```
<seaborn.axisgrid.PairGrid at 0x1fd378f04c8>
```



In [11]:
```
x = cars[['mpg', 'hp', 'qsec', 'wt']]
sb.pairplot(x)
```

Out[11]:
```
<seaborn.axisgrid.PairGrid at 0x1fd3ed59e88>
```



In [ ]:
```
Using scipy to calculate the Pearson correlation coefficient
```

In [18]:
```
mpg = cars['mpg']
hp = cars['hp']
qsec = cars['qsec']
wt = cars['wt']

pearsonr_coefficient, p_value = pearsonr(mpg, hp)
print('PeasonR Correlation Coefficient %0.3f'% (pearsonr_coefficient))
```

```
PeasonR Correlation Coefficient -0.776
```

In [19]:
```
pearsonr_coefficient, p_value = pearsonr(mpg, qsec)
print('PeasonR Correlation Coefficient %0.3f'% (pearsonr_coefficient))
```

```
PeasonR Correlation Coefficient 0.419
```

In [20]:
```
pearsonr_coefficient, p_value = pearsonr(mpg, wt)
print('PeasonR Correlation Coefficient %0.3f'% (pearsonr_coefficient))
```

```
PeasonR Correlation Coefficient -0.868
```

```
Using pandas to calculate the Pearson correlation coefficient
```

In [22]:
```
corr = x.corr()
corr
```

Out[22]:

	mpg	hp	qsec	wt
mpg	1.000000	-0.776168	0.418684	-0.867659
hp	-0.776168	1.000000	-0.708223	0.658748
qsec	0.418684	-0.708223	1.000000	-0.174716
wt	-0.867659	0.658748	-0.174716	1.000000

```
Using Seaborn to visualize the Pearson correlation coefficient
```

In [24]:
```
sb.heatmap(corr, xticklabels=corr.columns.values, yticklabels=corr.columns.values
```

Out[24]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1fd3f93f708>
```