In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:
```
df = pd.read_csv('Auto.csv')
df.head()
```

Out[2]:

	mpg	cylinders	displacement	horsepower	weight	acceleration	year	origin	name
0	18.0	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165	3693	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150	3436	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150	3433	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140	3449	10.5	70	1	ford torino

In [17]:
```
# Problem 2
pd.set_option('display.max_rows', None)
desiredCols = ['mpg', 'cylinders', 'displacement', 'horsepower', 'weight', 'acceleration', 'year', 'origin']
df1 = df.copy()

df1 = df1[desiredCols]

df1['horsepower'] = pd.to_numeric(df1['horsepower'], errors='coerce')
# horsePowerAvg = df1.mean()
# df1['horsepower'] = df1['horsepower'].fillna(horsePowerAvg)

df1.dropna()

df1.max() - df1.min()
```

Out[17]:
```
mpg              37.6
cylinders         5.0
displacement    387.0
horsepower      184.0
weight         3527.0
acceleration     16.8
year             12.0
origin            2.0
dtype: float64
```

In [18]:
```
# Problem 3


df1.mean()
```

Out[18]:
```
mpg              23.515869
cylinders         5.458438
displacement    193.532746
horsepower      104.469388
weight         2970.261965
acceleration     15.555668
year             75.994962
origin            1.574307
dtype: float64
```

In [19]:
```
df1.std()
```

Out[19]:
```
mpg              7.825804
cylinders        1.701577
displacement   104.379583
horsepower      38.491160
weight         847.904119
acceleration     2.749995
year             3.690005
origin           0.802549
dtype: float64
```

In [20]:
```
# Problem 4
df1.drop(df1.loc[10:85].index, inplace=True)
```

In [21]:
```
df1.max() - df1.min()
```

Out[21]:
```
mpg              35.6
cylinders         5.0
displacement    387.0
horsepower      184.0
weight         3348.0
acceleration     16.3
year             12.0
origin            2.0
dtype: float64
```

In [22]:
```
df1.mean()
```

Out[22]:
```
mpg              24.444860
cylinders         5.370717
displacement    187.174455
horsepower      101.003155
weight         2933.183801
acceleration     15.709034
year             77.143302
origin            1.598131
dtype: float64
```

In [23]:
```
df1.std()
```

Out[23]:
```
mpg              7.899928
cylinders        1.653486
displacement    99.864568
horsepower      36.003208
weight         809.638650
acceleration     2.706441
year             3.128202
origin           0.816163
dtype: float64
```

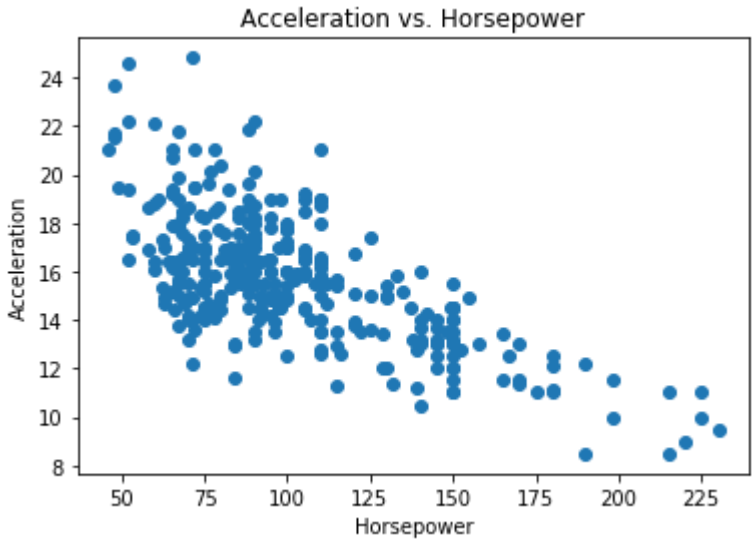
In [24]:
```
df1.corr()
```

Out[24]:

	mpg	cylinders	displacement	horsepower	weight	acceleration	year
mpg	1.000000	-0.764225	-0.797656	-0.765203	-0.833738	0.381023	0.617479
cylinders	-0.764225	1.000000	0.945427	0.831907	0.886452	-0.453747	-0.418113
displacement	-0.797656	0.945427	1.000000	0.895207	0.936849	-0.485930	-0.441536
horsepower	-0.765203	0.831907	0.895207	1.000000	0.859582	-0.680755	-0.461444
weight	-0.833738	0.886452	0.936849	0.859582	1.000000	-0.361513	-0.389327
acceleration	0.381023	-0.453747	-0.485930	-0.680755	-0.361513	1.000000	0.305802
year	0.617479	-0.418113	-0.441536	-0.461444	-0.389327	0.305802	1.000000
origin	0.552135	-0.542273	-0.600552	-0.446462	-0.579747	0.179905	0.207449

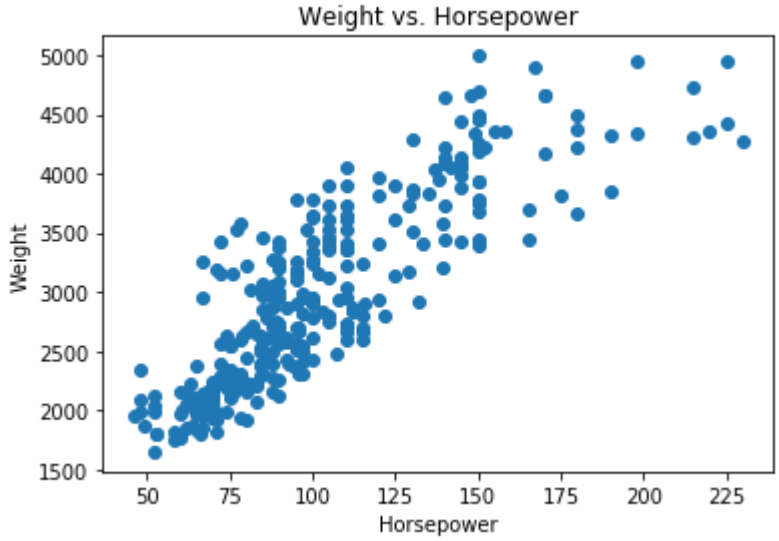
In [25]:
```
plt.scatter(df1['horsepower'], df1['acceleration'])
plt.title("Acceleration vs. Horsepower")
plt.xlabel("Horsepower")
plt.ylabel("Acceleration")
```

Out[25]:
```
Text(0,0.5,'Acceleration')
```



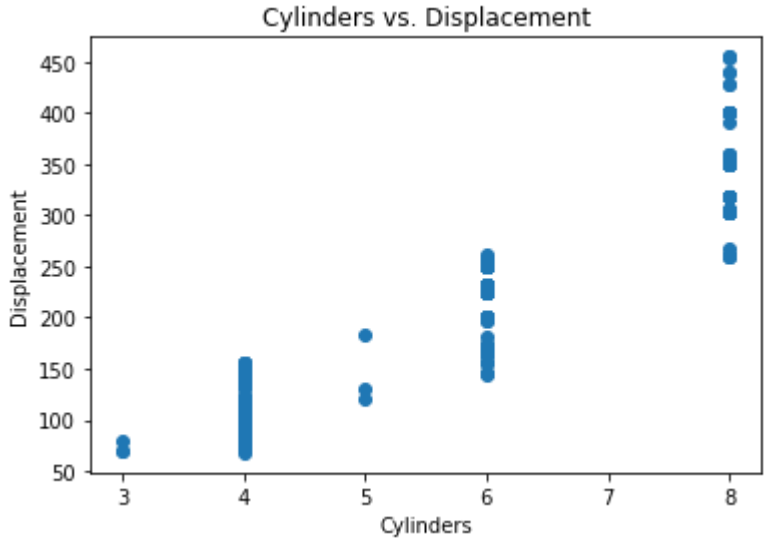
In [26]:
```
plt.scatter(df1['horsepower'], df1['weight'])
plt.title("Weight vs. Horsepower")
plt.xlabel("Horsepower")
plt.ylabel("Weight")
```

Out[26]: Text(0,0.5,'Weight')

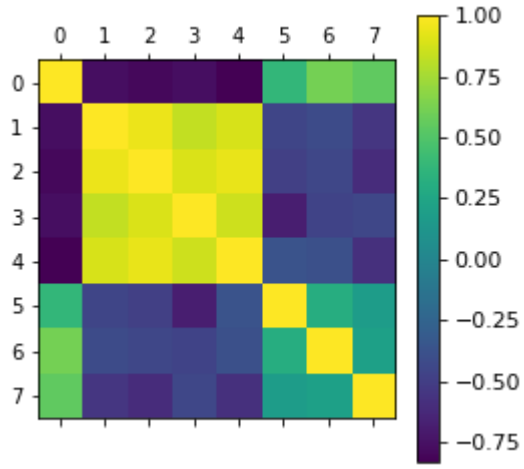


In [27]:
```
plt.scatter(df1['cylinders'], df1['displacement'])
plt.title("Cylinders vs. Displacement")
plt.xlabel("Cylinders")
plt.ylabel("Displacement")
```

Out[27]: Text(0,0.5,'Displacement')



In [28]:
```
plt.matshow(df1.corr())
cb = plt.colorbar()
cb.ax.tick_params(labelsize=11)
plt.show()
```



In [ ]:

In [ ]:

In [ ]: