drawing_conclusions

October 23, 2017

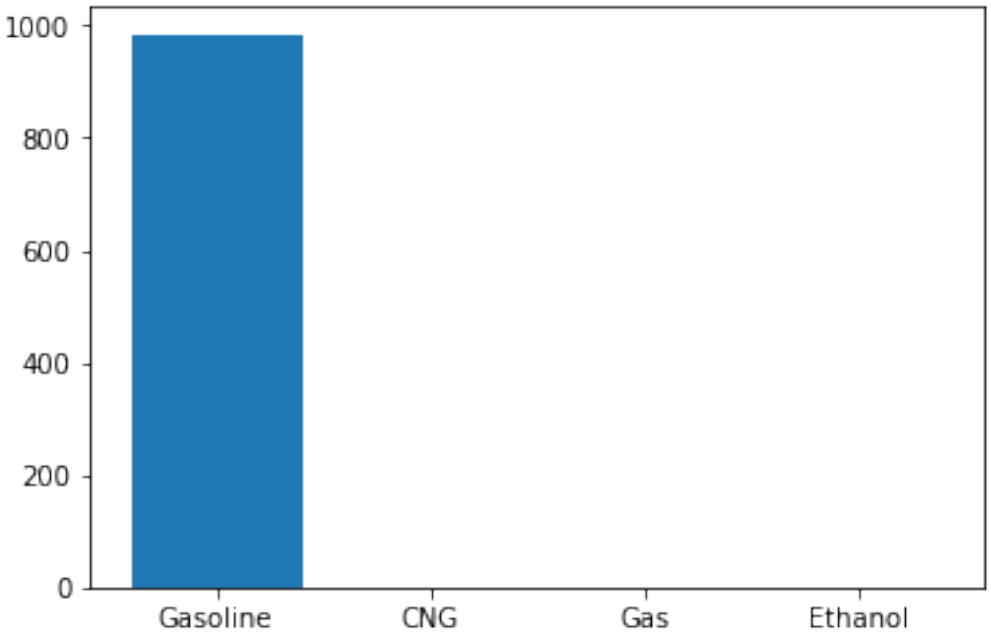
1 Drawing Conclusions

Use the space below to address questions on datasets `clean_08.csv` and `clean_18.csv`

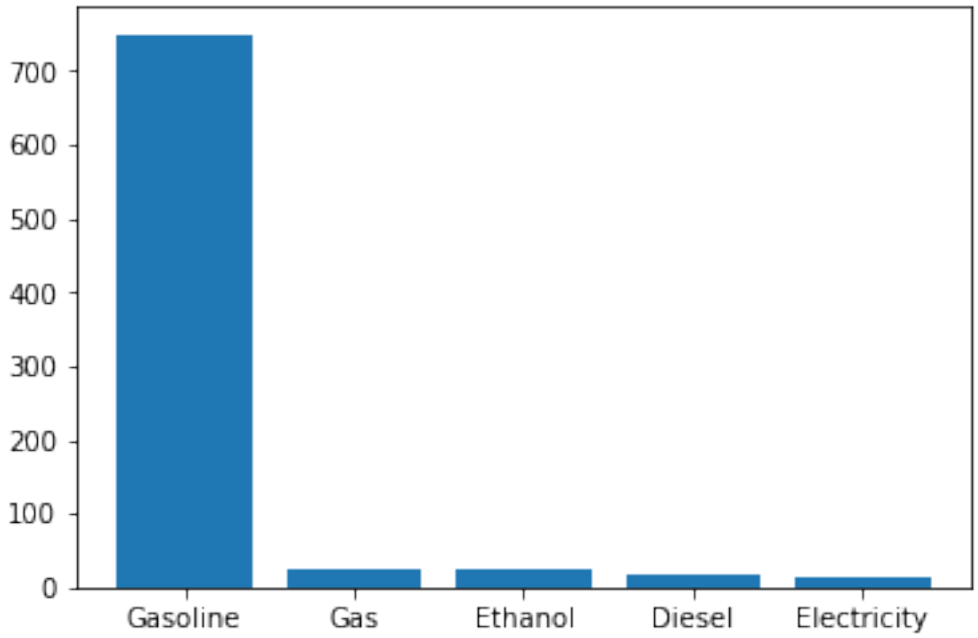
In [3]:
```
# load datasets
import pandas as pd
import matplotlib.pyplot as plt
% matplotlib inline
df_08 = pd.read_csv('clean_08.csv')
df_18 = pd.read_csv('clean_18.csv')
```

1.0.1 Q1: Are more unique models using alternative sources of fuel? By how much?

In [4]:
```
df_08['fuel'].value_counts()
gasoline = df_08['fuel'].value_counts()[0]
cng = df_08['fuel'].value_counts()[1]
gas = df_08['fuel'].value_counts()[2]
ethanol = df_08['fuel'].value_counts()[3]

x_labels = ['Gasoline', 'CNG', 'Gas', 'Ethanol']
y_values = [gasoline, cng, gas, ethanol]
locations = [1, 2, 3, 4]
plt.bar(locations, y_values, tick_label = x_labels);
```




```
In [5]: df_18['fuel'].value_counts()
        gasoline = df_18['fuel'].value_counts()[0]
        gas = df_18['fuel'].value_counts()[1]
        ethanol = df_18['fuel'].value_counts()[2]
        diesel = df_18['fuel'].value_counts()[3]
        electricity = df_18['fuel'].value_counts()[4]

        x_labels = ['Gasoline', 'Gas', 'Ethanol', 'Diesel', 'Electricity']
        y_values = [gasoline, gas, ethanol, diesel, electricity]
        locations = [1, 2, 3, 4, 5]
        plt.bar(locations, y_values, tick_label = x_labels);
```



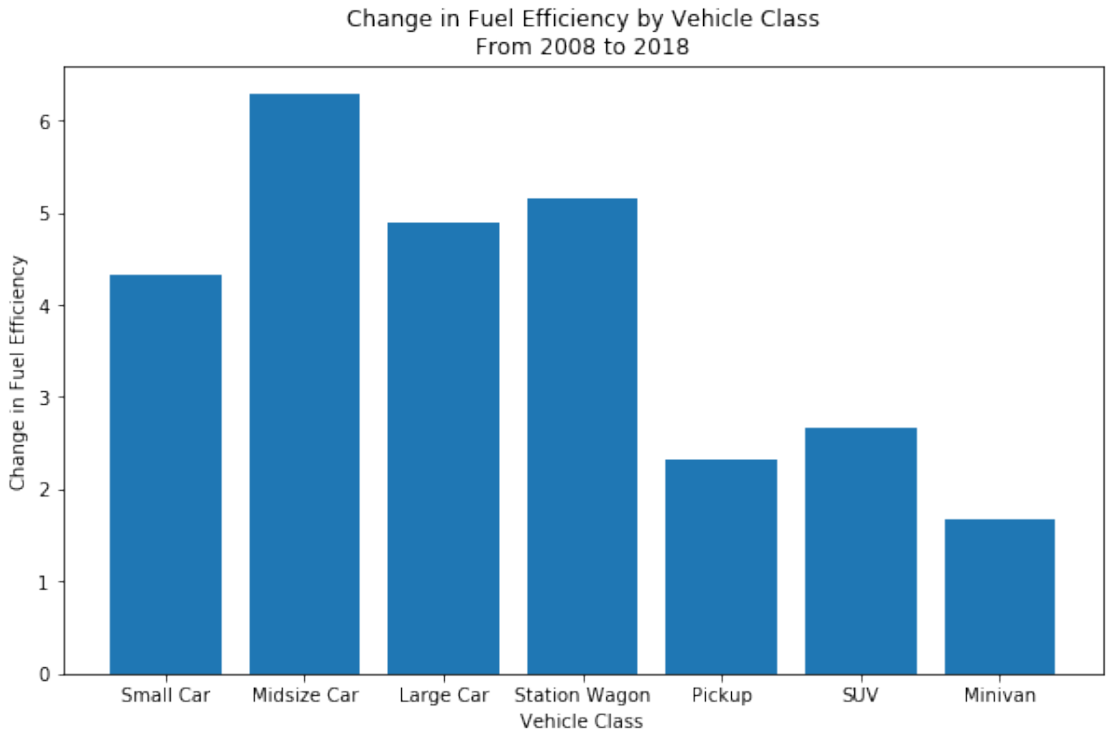
1.0.2 Q2: How much have vehicle classes improved in fuel economy?

```
In [6]: #df_08.head()
        #df_08.groupby('veh_class')['cmb_mpg'].describe()
        df_08.groupby('veh_class')['cmb_mpg'].mean()

Out[6]: veh_class
        SUV              18.471429
        large car        18.509091
        midsize car      21.601449
        minivan          19.117647
        pickup           16.277108
        small car        21.105105
        station wagon    22.366667
        van              14.952381
        Name: cmb_mpg, dtype: float64

In [7]: #df_18.groupby('veh_class')['cmb_mpg'].describe()
        df_18.groupby('veh_class')['cmb_mpg'].mean()

Out[7]: veh_class
        large car          23.409091
        midsize car        27.884058
        minivan            20.800000
        pickup             18.589744
        small SUV          24.074074
```

```
small car           25.421053
special purpose     18.500000
standard SUV        18.197674
station wagon       27.529412
Name: cmb_mpg, dtype: float64
```

In [8]:
```
station_wagon_dx = df_18.groupby('veh_class')['cmb_mpg'].mean()[8] - df_08.groupby('veh_
large_car_dx = df_18.groupby('veh_class')['cmb_mpg'].mean()[0] - df_08.groupby('veh_clas
midsize_car_dx = df_18.groupby('veh_class')['cmb_mpg'].mean()[1] - df_08.groupby('veh_cl
small_car_dx = df_18.groupby('veh_class')['cmb_mpg'].mean()[5] - df_08.groupby('veh_clas
minivan_dx = df_18.groupby('veh_class')['cmb_mpg'].mean()[2] - df_08.groupby('veh_class'
pickup_dx = df_18.groupby('veh_class')['cmb_mpg'].mean()[3] - df_08.groupby('veh_class')
suv_dx = (df_18.groupby('veh_class')['cmb_mpg'].mean()[4] + df_18.groupby('veh_class')['
```

In [9]:
```
location = [1, 2, 3, 4, 5, 6, 7]
labels = ['Small Car', 'Midsize Car', 'Large Car', 'Station Wagon', 'Pickup', 'SUV', 'Mi
values = [small_car_dx, midsize_car_dx, large_car_dx, station_wagon_dx, pickup_dx, suv_d
plt.figure(figsize=(10,6));
plt.bar(location, values, tick_label=labels);
plt.title('Change in Fuel Efficiency by Vehicle Class\nFrom 2008 to 2018');
plt.xlabel('Vehicle Class');
plt.ylabel('Change in Fuel Efficiency');
```



1.0.3 Q3: What are the characteristics of SmartWay vehicles? Have they changed over time?

```
In [10]: df_08.query('smartway == "yes"').describe()
```

```
Out[10]:              displ         cyl  air_pollution_score    city_mpg     hwy_mpg  \
         count  380.000000  380.000000           380.000000  380.000000  380.000000
         mean     2.602895    4.826316             7.268421   20.984211   28.413158
         std      0.623436    1.002025             0.970027    3.442672    3.075194
         min      1.300000    4.000000             6.000000   17.000000   22.000000
         25%      2.275000    4.000000             7.000000   19.000000   26.000000
         50%      2.400000    4.000000             7.000000   20.000000   28.000000
         75%      3.000000    6.000000             7.000000   22.000000   30.000000
         max      5.000000    8.000000             9.000000   48.000000   45.000000

                   cmb_mpg  greenhouse_gas_score
         count  380.000000            380.000000
         mean    23.736842              6.868421
         std      3.060379              0.827338
         min     20.000000              6.000000
         25%     22.000000              6.000000
         50%     23.000000              7.000000
         75%     25.000000              7.000000
         max     46.000000             10.000000
```

```
In [11]: print("Combined MPG Comparison")
         print('2008:', df_08.query('smartway == "yes"')['cmb_mpg'].mean())
         print('2018:', df_18.query('smartway == "Yes"')['cmb_mpg'].mean())
```

```
Combined MPG Comparison
2008: 23.7368421053
2018: 34.4395604396
```

```
In [12]: print("Greenhouse Gas Score Comparison")
         print('2008:', df_08.query('smartway == "yes"')['greenhouse_gas_score'].mean())
         print('2018:', df_18.query('smartway == "Yes"')['greenhouse_gas_score'].mean())
```

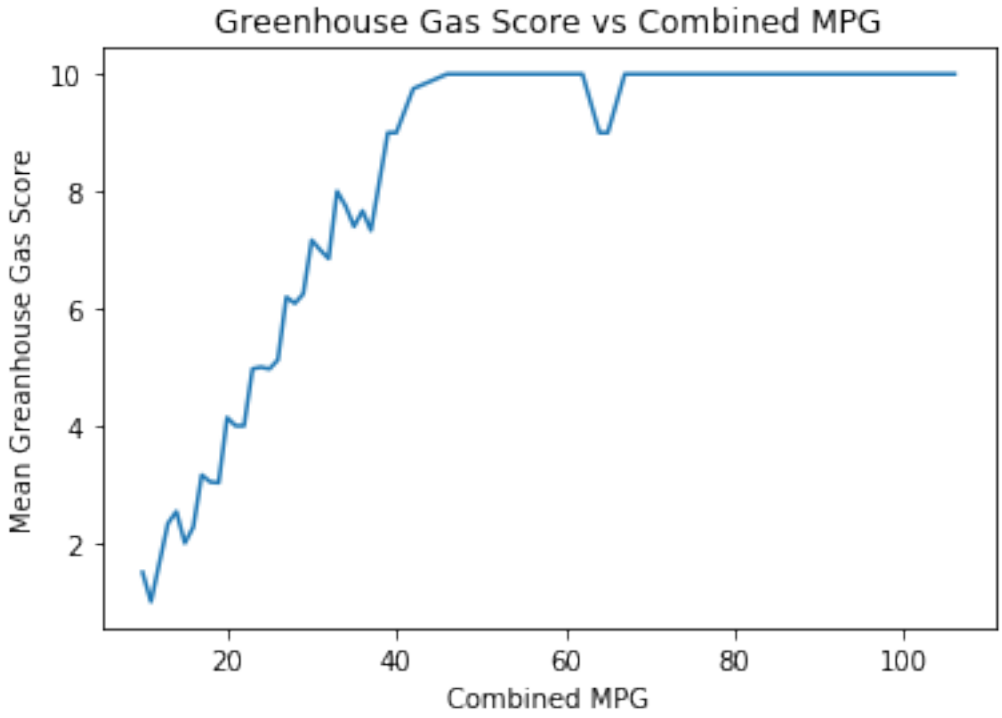
```
Greenhouse Gas Score Comparison
2008: 6.86842105263
2018: 7.53846153846
```

```
In [13]: print("Air Pollution Score")
         print('2008:', df_08.query('smartway == "yes"')['air_pollution_score'].mean())
         print('2018:', df_18.query('smartway == "Yes"')['air_pollution_score'].mean())
```

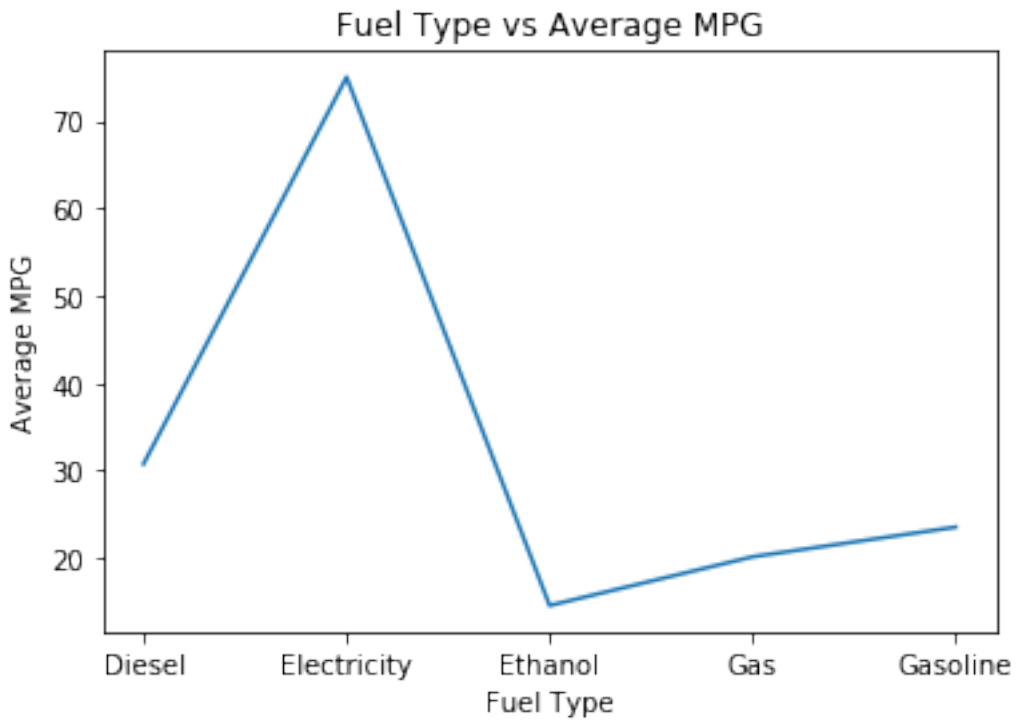
```
Air Pollution Score
2008: 7.26842105263
2018: 4.87912087912
```

1.0.4 Q4: What features are associated with better fuel economy?

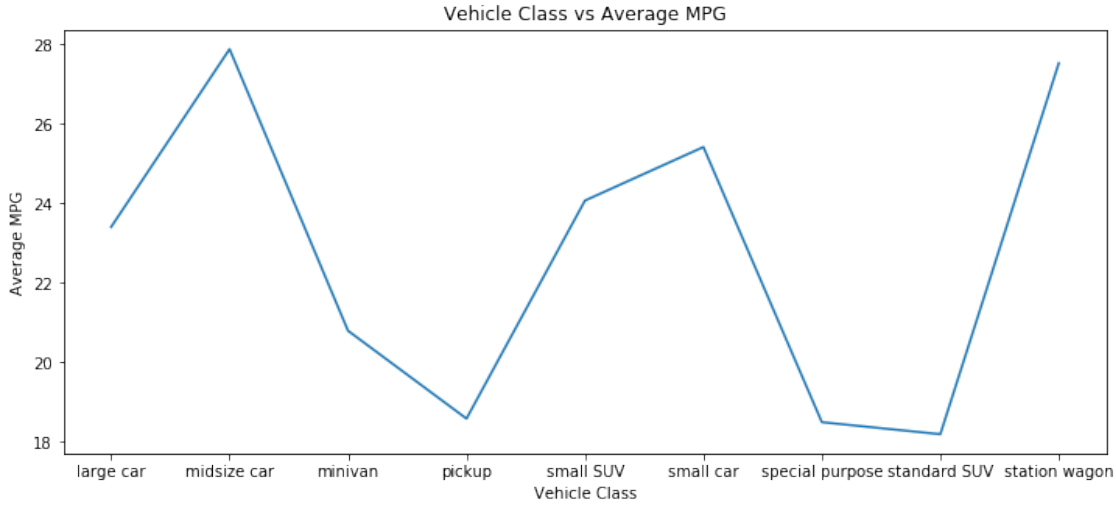
```
In [80]: x = df_18.groupby('cmb_mpg')['cmb_mpg'].mean()
         y = df_18.groupby('cmb_mpg')['greenhouse_gas_score'].mean()
         plt.plot(x,y);
         plt.title('Greenhouse Gas Score vs Combined MPG');
         plt.xlabel('Combined MPG');
         plt.ylabel('Mean Greanhouse Gas Score');
```




```
In [167]: x = df_18.groupby('fuel')
          y = df_18.groupby('fuel')['cmb_mpg'].mean()
          plt.plot(y);
          plt.title('Fuel Type vs Average MPG');
          plt.xlabel('Fuel Type');
          plt.ylabel('Average MPG');
```




```
In [115]: y = df_18.groupby('veh_class')['cmb_mpg'].mean()
          plt.figure(figsize=(12,5));
          plt.plot(y);
          plt.title('Vehicle Class vs Average MPG');
          plt.xlabel('Vehicle Class');
          plt.ylabel('Average MPG');
```



In [ ]: