# Exploratory Data Analysis:
UCI Auto MPG Data Set

The UCI Auto MPG data set is a data set of popular cars models from 1970 to 1982 with information about their MPG, engine displacement, acceleration, weight, country of manufacture, horsepower, and number of cylinders. After completing some exploratory Data Analysis on the Auto MPG data set, I have come to some interesting conclusions.

First, we can explore the relationship between engine characteristics—MPG, displacement, horsepower, and acceleration. In the graph below, you can see the relationship between displacement, weight, MPG, and acceleration. The size of the dot represents the MPG and the color represents the acceleration speed. As expected, as engines increase in weight and size, they become slower and less efficient.
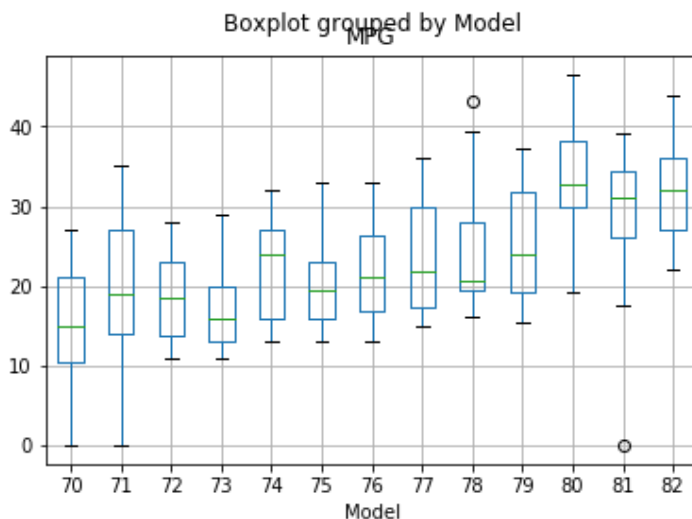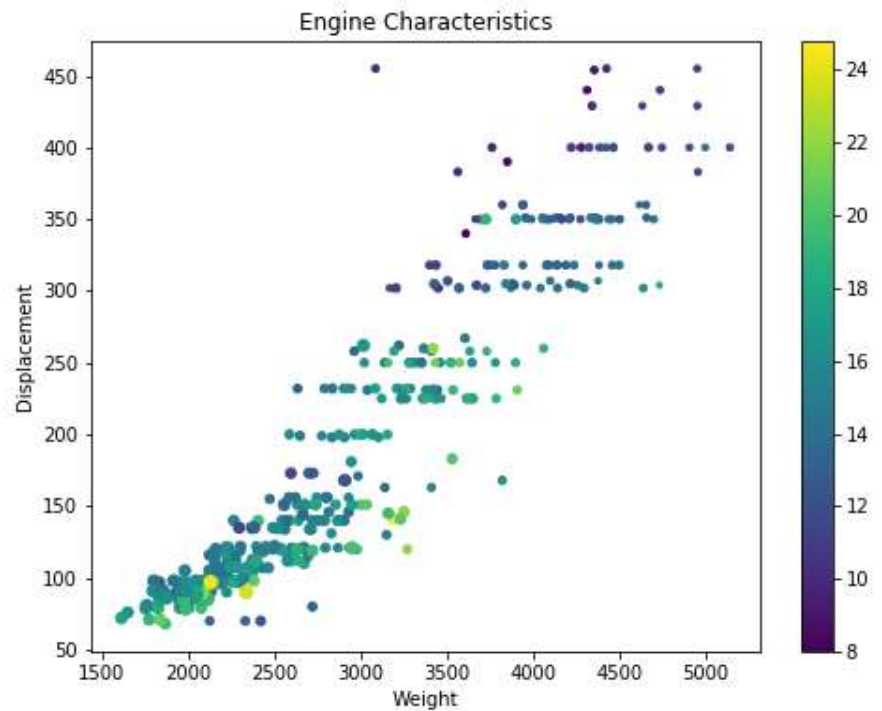
Something interesting to see is the trend in MPG over time. Since the data set contains a twelve-year span of cars, we can see the progression of engine efficiency during the OPEC crisis. See graph on left for details.

Here, to the right, one can see the average MPG of each manufacturer listed in the data set. This analysis was created by parsing the manufacturer from the car title, and then calculating the average MPG for that specific make.

It is interesting that some cars known today for their efficiency were efficient in the 70s—Honda, Nissan, Subaru, and Volkswagen.

Manufacturer vs. Average MPG
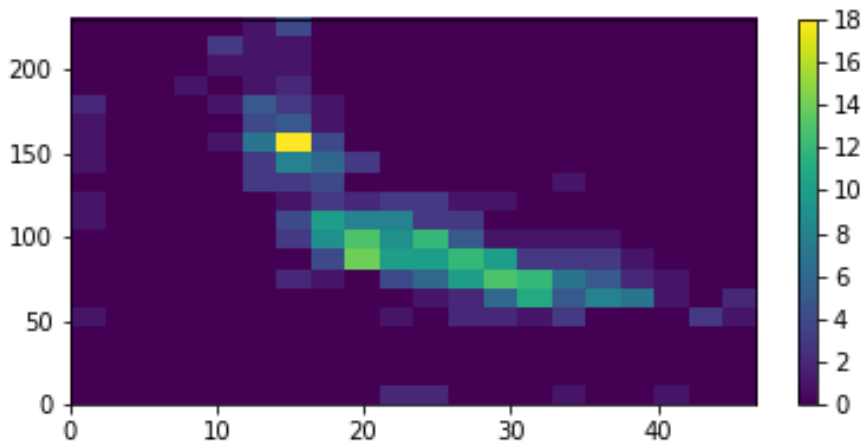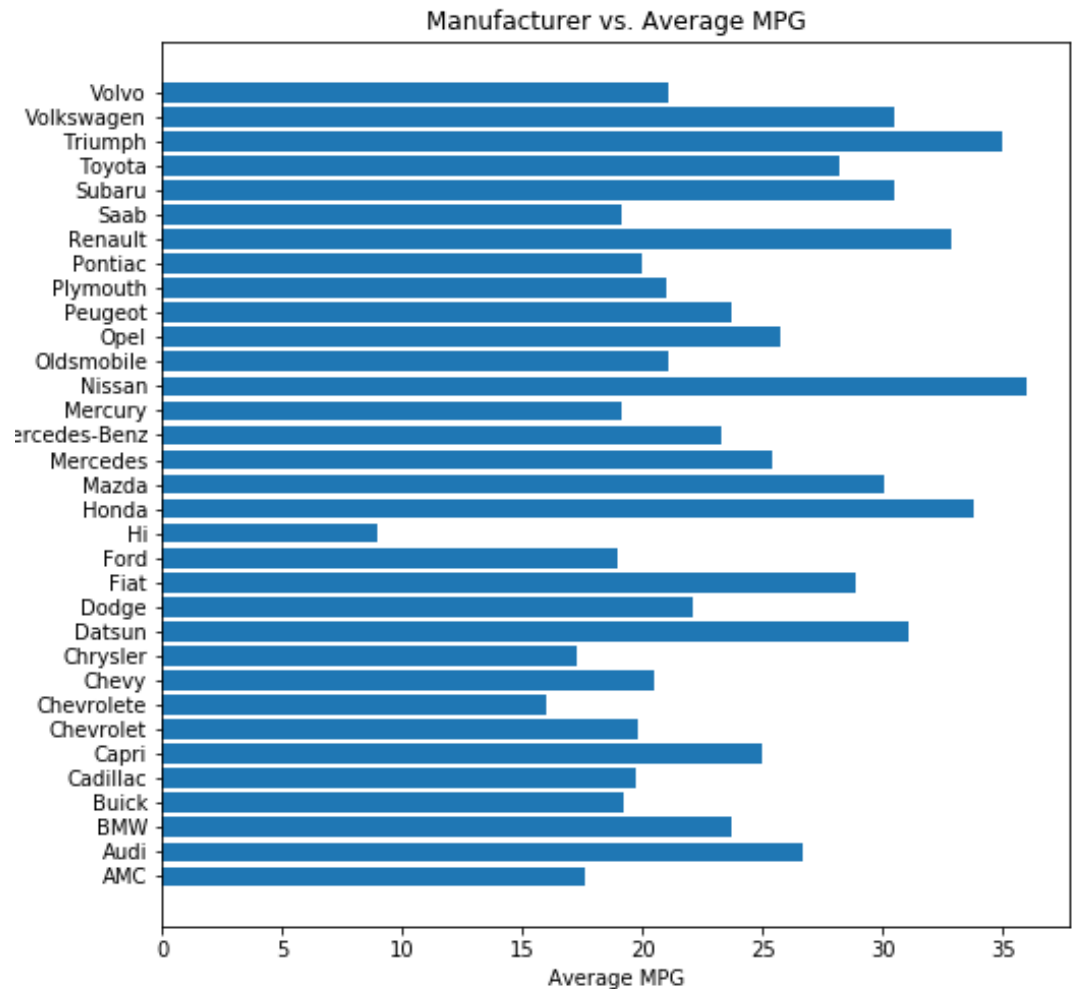
In this last graph, we are plotting a 2D histogram of MPG vs Horsepower. Not surprisingly, there is an exponential decay curve. We can see that most cars are located from 15 to 35 MPG, with the highest occurance at around 15 MPG.