

# Explanatory Analysis of MPG Data Set

Samuel Venzi Lima Monteiro de Oliveira

915635895

In this document we'll be analyzing a data set regarding cars manufactured between 1970 and 1982 and their attributes. This type of analysis is interesting because we are able to see trends and relate them to actual happenings of the time, such as the oil crisis of the 70s.

In this data set we are dealing with the name of the cars as unique “attributes”, that is, there aren't any entries with the same name. Each one of the cars has the following attributes: MPG, Cylinders, Displacement, Horsepower, Weight, Acceleration, Model and Origin.

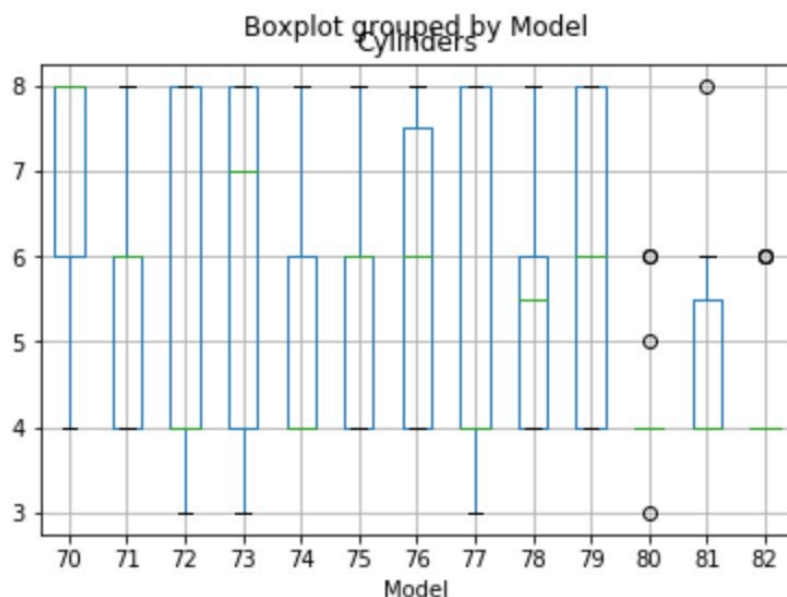


Figure 1

In the box plot above, we can see the range of the number of cylinders per year. It's interesting to notice that when the oil crisis hit in 1973, car manufacturers had to sell more economic cars, so the number of cylinders decreased.

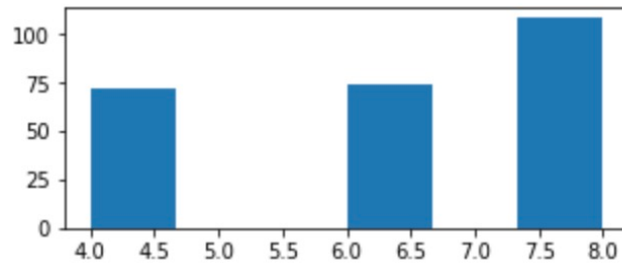


Figure 2: cylinders in the US

The plot shown in Figure 2 is a histogram of Cylinders of US cars only. It's possible to see that, despite the crisis, cars with 8 cylinders outnumber the rest. In the following plots it is shown that Europe and Japan did not follow that trend, and had always manufactured cars with less than 8 cylinders.

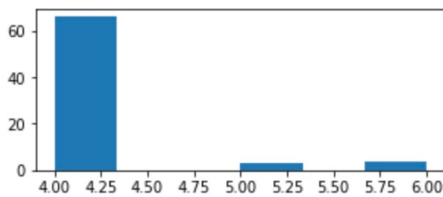


Figure 3: cylinders in Europe

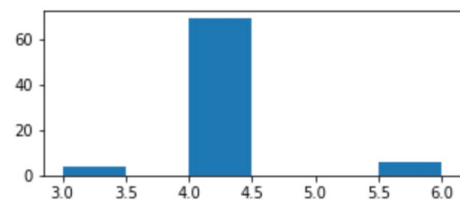


Figure 4: cylinders in Japan