

# Exploratory Analysis of MPG Data Set

Zhijia Chen

2020-09-10

On October 19, 1973, the Organization of Arab Petroleum Exporting Countries (OPEC) imposed an oil embargo targeted at nations perceived as supporting Israel during the Yom Kippur War [1]. That embargo caused a global oil crisis and was later called the "first oil shock". During the oil crisis, the oil price had risen almost 300% from \$3 per barrel to nearly \$12 globally [Oil Embargo, 1973–1974]. As a measure to reduce energy consumption and improve nation's energy security, the congress passed Corporate Average Fuel Economy (CAFE) Standards in 1975 [2]. The CAFE standards regulates the minimum mileage per gallon (MPG) that must be achieved by all automakers for their cars and trucks each year since 1978. The regulation pushes automakers to make more efficient cars year by year. Between 1975 and 1985, average passenger vehicle mileage doubled from about 13.5 mpg to 27.5 [3]. But an interesting question is, how do automakers achieve that? By auto technology advancement alone, or by making performance compromise? In this exploratory analysis, we will examine the MPG dataset and try to get some insight about this question.

We first check how the MPG changed before and after the enforcement of CAFE. Figure 1 shows the boxplot of MPG grouped by model year between 1970 and 1982, and the yellow line shows the average MPG while the red line with star markers shows the CAFE standards since 1978. Clearly, the average MPG has increased steadily before the "first oil shock", without the regulation mandate. However, the gap between the minimum value and the first quartile is increasing, which implies the regulation does drive some automakers that lag behind average to make more efficient cars.

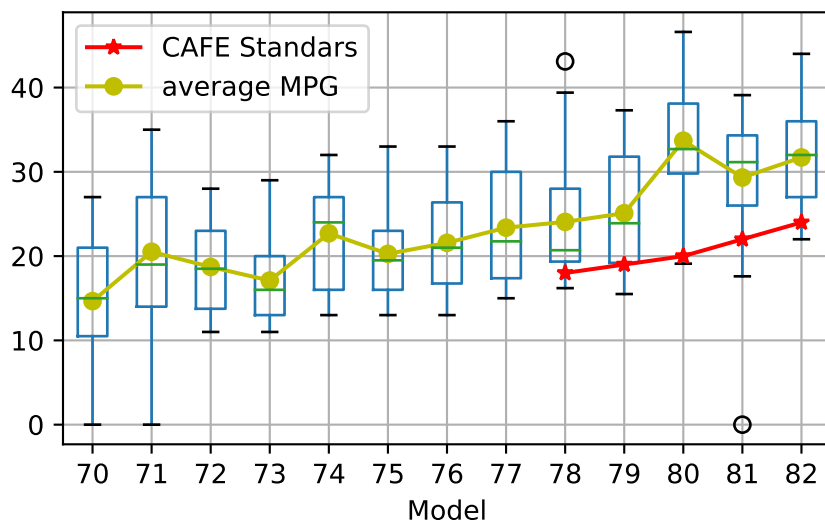


Figure 1: MPG boxplot grouped by model.

To study if automakers make any performance compromise to improve MPG, we will look into two key parameters: horsepower and acceleration. Figure 2 shows the MPG scatter plot with horsepower (left) and acceleration (right) respectively. Interestingly, horsepower tends to decrease as MPG increases, which could imply that automakers sacrifice engine powers in favor of higher MPG. However, The acceleration actually positively correlated with MPG, meaning that cars run faster with higher MPG.

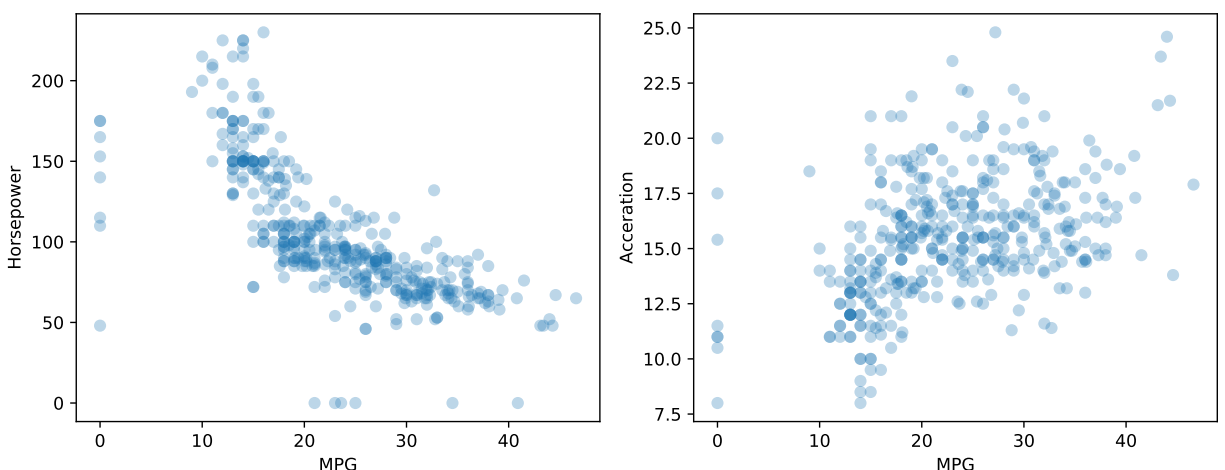


Figure 2: MPG scatter plot with Horsepower and Acceleration.

The fact seems somewhat counter intuitive, as faster acceleration usually requires higher power. But still, this could be explained by two reasons:

1. Automakers are able to reduce car weight to compensate the horsepower reduction and even achieved faster acceleration.
2. The technology advancement enable automakers to produce more efficient car so that acceleration is increased even with lower horsepower.

Note that  $a = \sqrt{\frac{P\eta}{2mt}}$ , where  $a$ ,  $P$ ,  $m$  and  $t$  are acceleration, horsepower, weight and time respectively. And  $\eta$  is the overall efficiency factor.

We will first study the change of car weight in this period. Figure 3 shows the average car weight and average acceleration of each year. We see that, overall speaking, the car weight dose decrease overtime, and the two parameters present a clear opposite trend except for the year of 75, 76 and 77 when they both decrease. This proves that car weight is a key parameters targeted by automakers to improve MPG and acceleration.

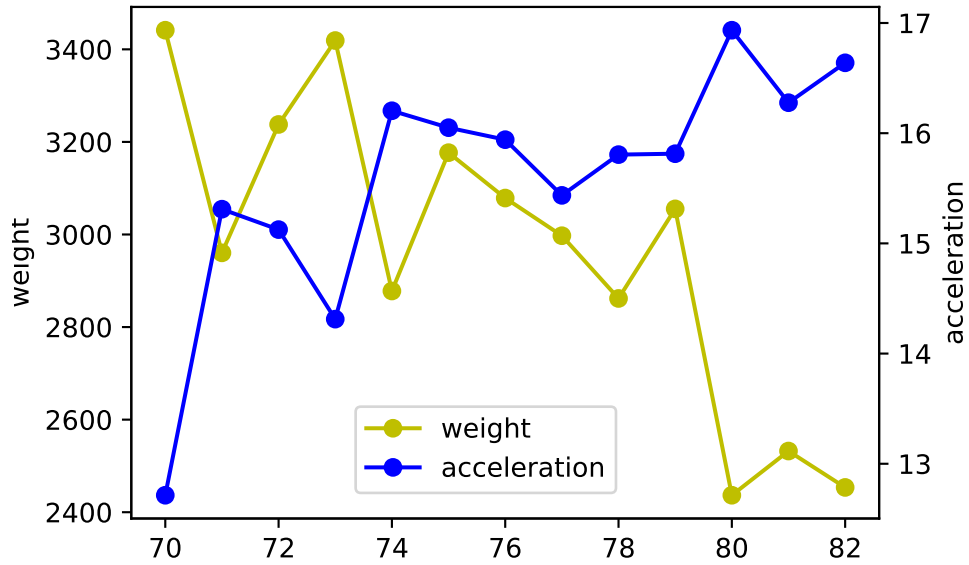


Figure 3: Weight boxplot grouped by model.

There is no parameters to show the efficiency factor directly, so we try to exame if weight is a dominant factor that improves MPG. We group cars by the number of cylinders, and for each pair of cars  $c_i, c_j$  in a group , we compute a tuple  $(x, y)$  where

$$\begin{aligned}
 weight(c_i) &\geq weight(c_j) \\
 x &= weight(c_i) - weight(c_j) \\
 y &= MPG(c_j) - MPG(c_i)
 \end{aligned}$$

If weight is a dominant factor in improving MPG, then most of the points should be located in the first quadrant. Figure 4 presents the distribution of the points. We see that most of points are located in the first quadrant, so we get the conclusion that the main factor that contributes to higher MPG is car weight.

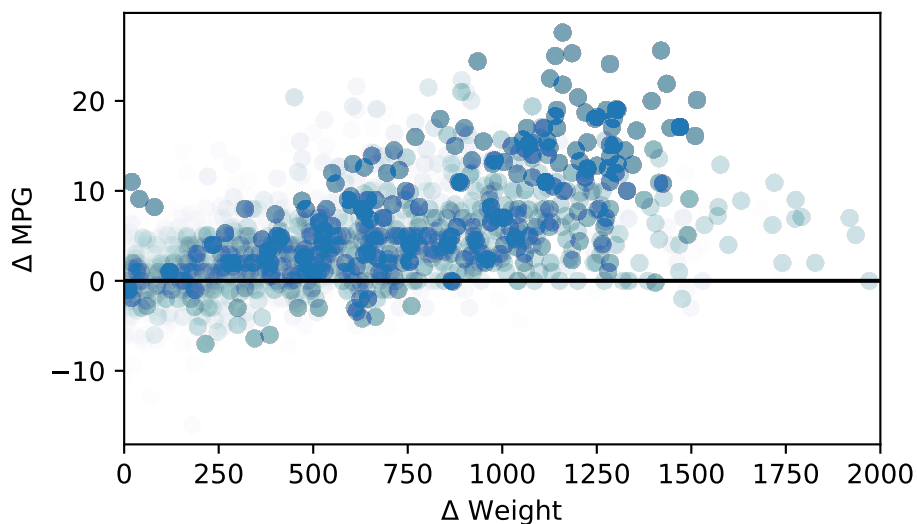


Figure 4:  $\Delta$  Weight vs  $\Delta$  MPG

## References

- [1] C. D. Smith, *Palestine and the Arab-Israeli conflict*. Bedford/St. Martin's, 2004.
- [2] U. D. of Transportation, *Corporate Average Fuel Economy (CAFE) Standards*, 2014 (accessed Sep. 9, 2020).
- [3] T. Swanson, *Driving to 54.5 MPG*, 2011 (accessed Sep. 9, 2020).