

ECO395M: Data Mining — Exercise 01

Scott Cohn

2021-01-20

Data visualization: Gas Prices

First, we import the data and geocode the addresses. The gas stations in Austin, grouped by brand, appear in Figure 1.

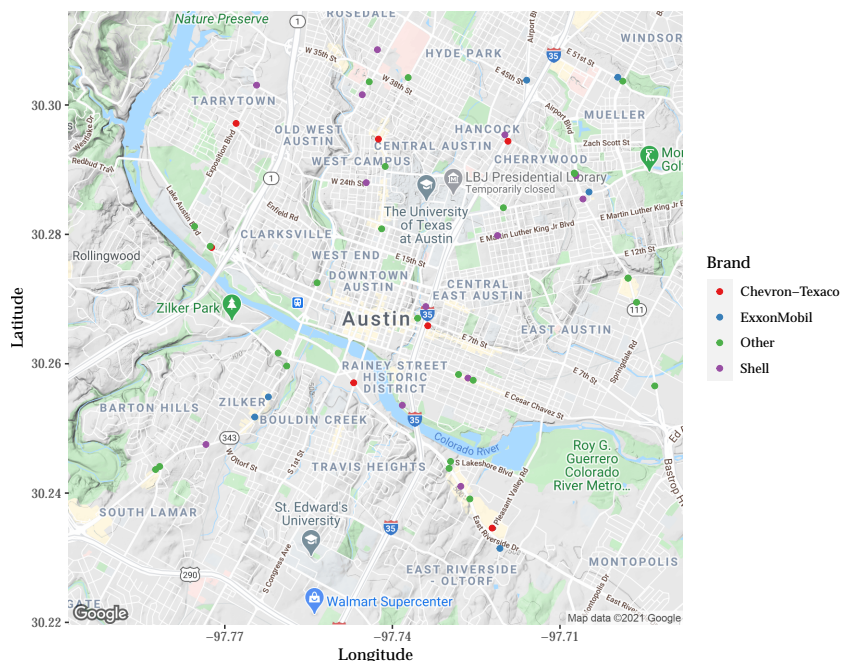


Figure 1: Map of Austin gas stations

The Theories

THEORY A: Gas stations charge more if they lack direct competition in sight.

CONCLUSION: The cheapest gas stations seem to charge similarly. Gas stations that have no competition have higher prices at the 50th and 75th percentile. There are several gas stations, again facing no competition, pricing 10 cents per gallon higher than all of the rest.

THEORY B: The richer the area, the higher the gas price.

CONCLUSION: There is a lot of variation within brand across the

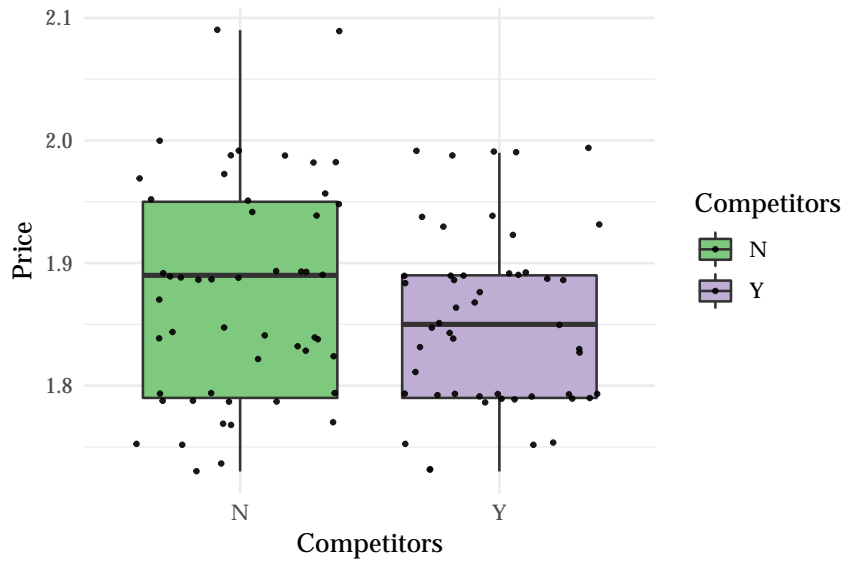


Figure 2: Price with competition

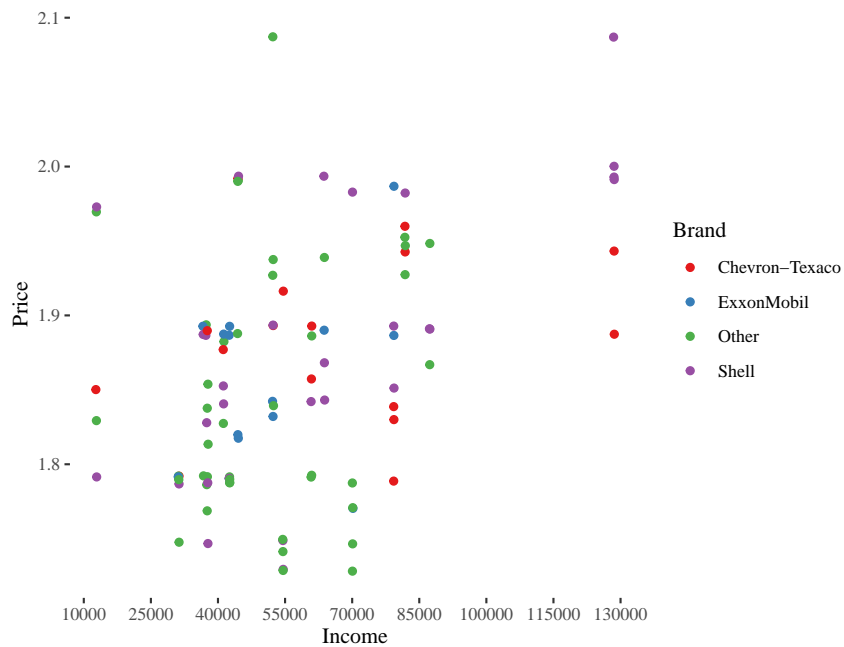
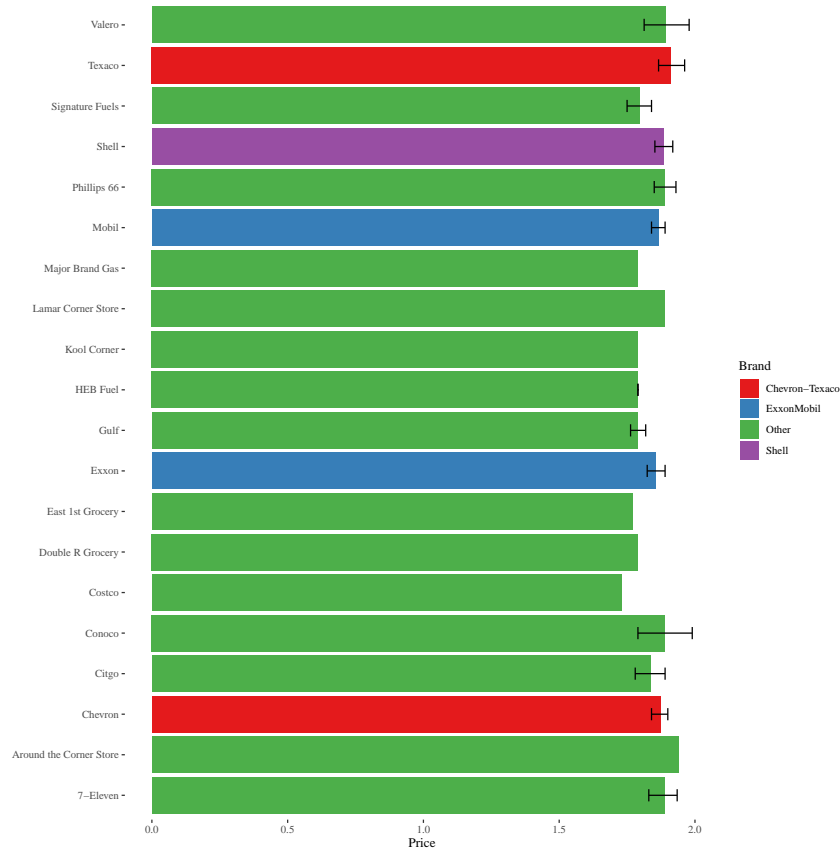


Figure 3: Price with income

income spectrum. There does seem to be a trend that higher income areas do have higher prices per gallon at the gas station.

THEORY C: Shell charges more than other brands.

CONCLUSION: Visually, the averages are very close; Shell seems to price similarly to other brands. Broken down by franchise, this trend



continues. Other gas stations, like Costco, that price well below the mean are able to take advantage (I assume) of other factors.

THEORY D: Gas stations at stoplights charge more.

CONCLUSION: The distribution of price not at a stoplight is right-skewed whereas the distribution of price at a stoplight appears to be more normally distributed (maybe even left-skewed). Limited data makes this a bit difficult to parse. They seem to be similarly centered around 1.8ish. It seems that gas stations as stoplights price a bit higher than their non-stoplight counterparts.

THEORY E: Gas stations with direct highway access charge more.

CONCLUSION: The distribution of prices without direct highway access is more left-skewed than the price distribution with direct access. This is clear to see in the density plot on the right, but also when comparing the medians in the boxplot on the left.

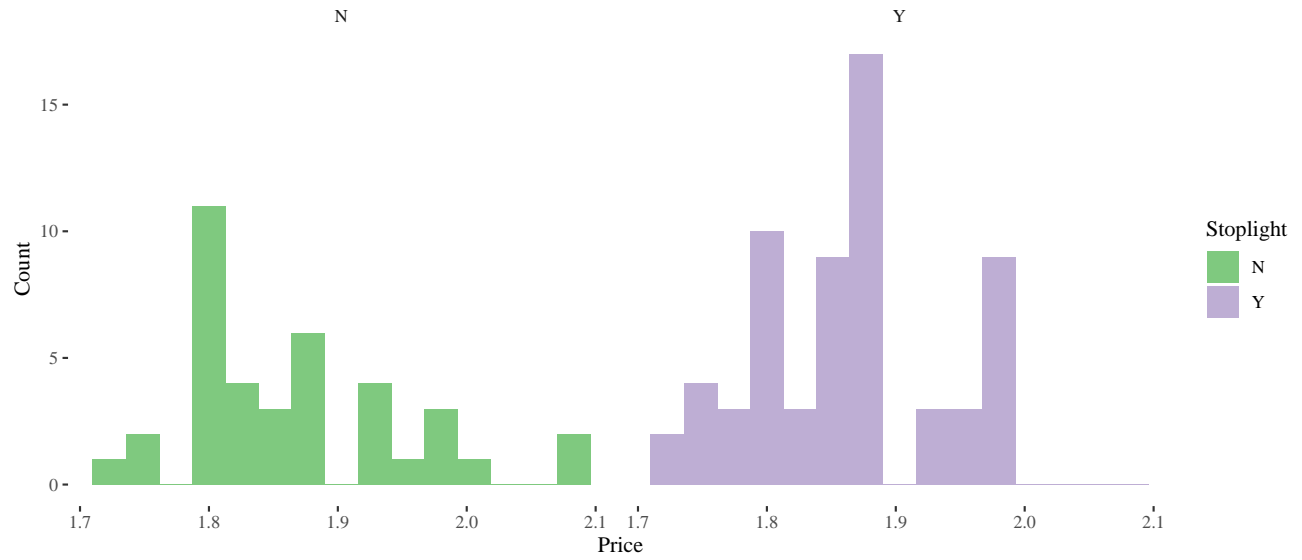


Figure 5: Price with stoplights

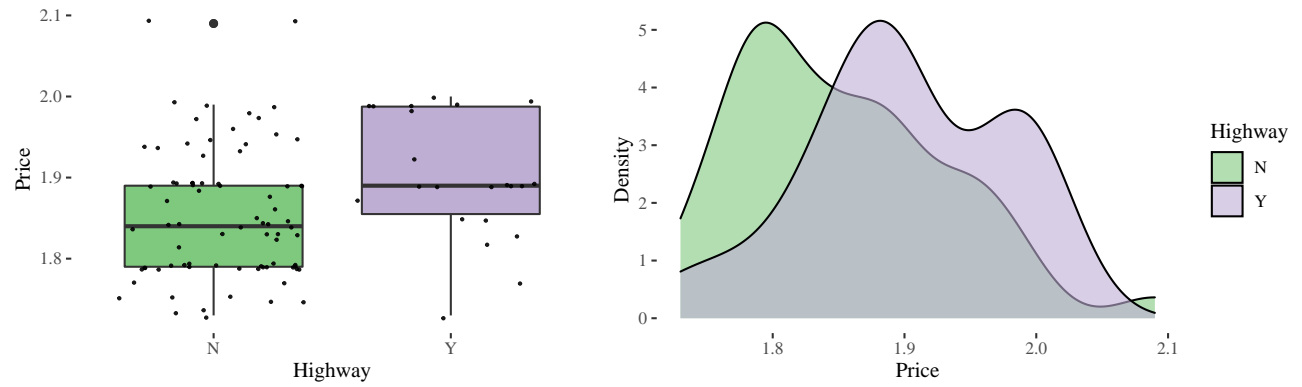


Figure 6: Price with highway access

Data Visualization: A Bike Share Network

Data visualization: Flights at ABIA