

# HW 3

Genavieve Middaugh

2025-02-07

## Problem 1 - Gas Station Prices

In Problem 1 we are assessing multiple theory given the prices of gas stations with multiple outside factors. The goal is to try to understand the variation in gas station prices and what causes that variation. The data given by GasPrices.csv a dataset made by students in spring of 2016 on a project.

The important variables are below:

- ID: Order in which gas stations were visited
- Name: Name of gas station
- Price: Price of regular unleaded gasoline, gathered on Sunday, April 3rd, 2016
- Highway: Is the gas station accessible from either a highway or a highway access road?
- Stoplight: Is there a stoplight in front of the gas station?
- Competitors: Are there any other gas stations in sight?
- Zipcode: Zip code in which gas station is located
- Income: Median Household Income of the ZIP code where the gas station is located based on 2014 data from the U.S. Census Bureau
- Brand: ExxonMobil, ChevronTexaco, Shell, or Other.

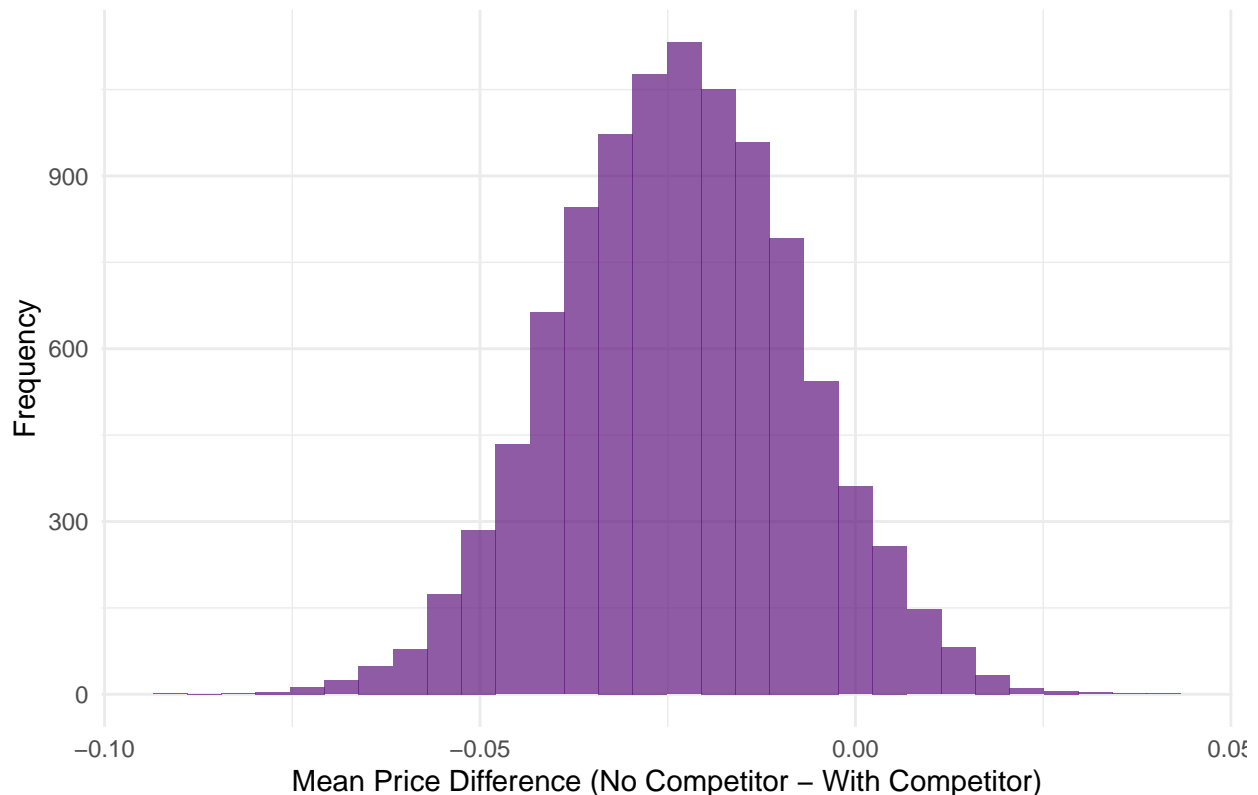
The theoris we are trying to either prove or disprove are below:

- A) Gas stations charge more if they lack direct competition in sight.
- B) The richer the area, the higher the gas prices.
- C) Gas stations at stoplights charge more.
- D) Gas stations with direct highway access charge more.
- E) Shell charges more than all other non-Shell brands.

### 1. Gas stations charge more if they lack direct competition in sight.

## 'stat\_bin()' using 'bins = 30'. Pick better value with 'binwidth'.

Bootstrap sampling distribution for difference in mean gas prices based on



If you have ever driven around, anywhere in America, I'm sure you've seen your fair share of gas stations, the fluctuation of prices varies from place to place, but one may wonder if having competitors nearby leads to gas stations having lower prices. The Competitors and Price variables from the gasprices data were used to understand if there were competitors around or not and what the prices were for each gas station. The above graph shows a bootstrap sampling distribution of the difference in the means of these two. The graph skews to the left of 0 only slightly, so we can say with 95% confidence that the average difference in gas price between gas stations with no competitors and with competitors is between -0.055 and 0.008 cents. Since the graph is so close to 0, this means it is a common occurrence for there to be no difference, meaning the average prices were the same or at least almost the same. It seems that there is very little correlation, especially considering the mean is -0.023, a very low number, and when you think about the difference in gas prices, this looks like 2.34 and 2.32; the difference is quite small.

To conclude, the data does not present any real evidence of the idea that gas stations' prices are correlated to whether there are competitors nearby. But we must remember that this data was only done over the course of 1 day, in the spring of 2016, in Austin. This data does not cover enough time for us to draw conclusions on if over time this is the same conclusion.

2. The richer the area, the higher the gas prices.
3. Gas stations at stoplights charge more.
4. Gas stations with direct highway access charge more.
5. Shell charges more than all other non-Shell brands.