HW3

Onir Narahari

2025-02-11

1.

Gas Prices by Competitor Presence
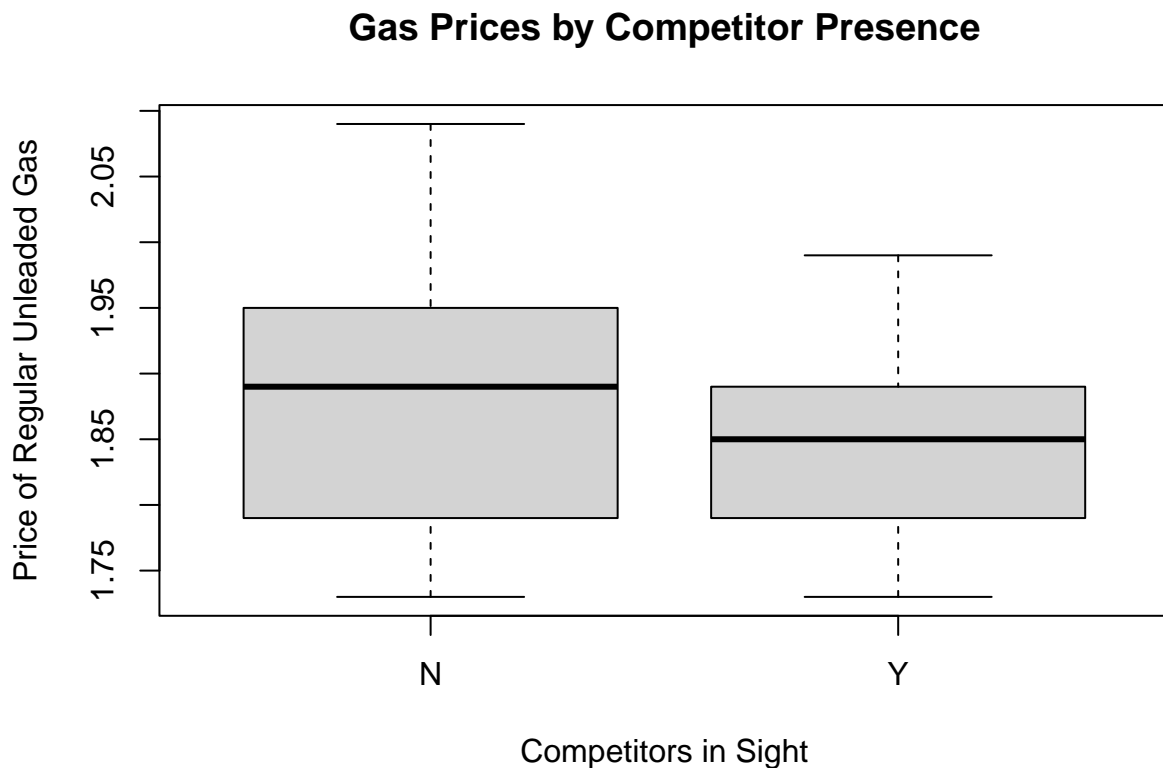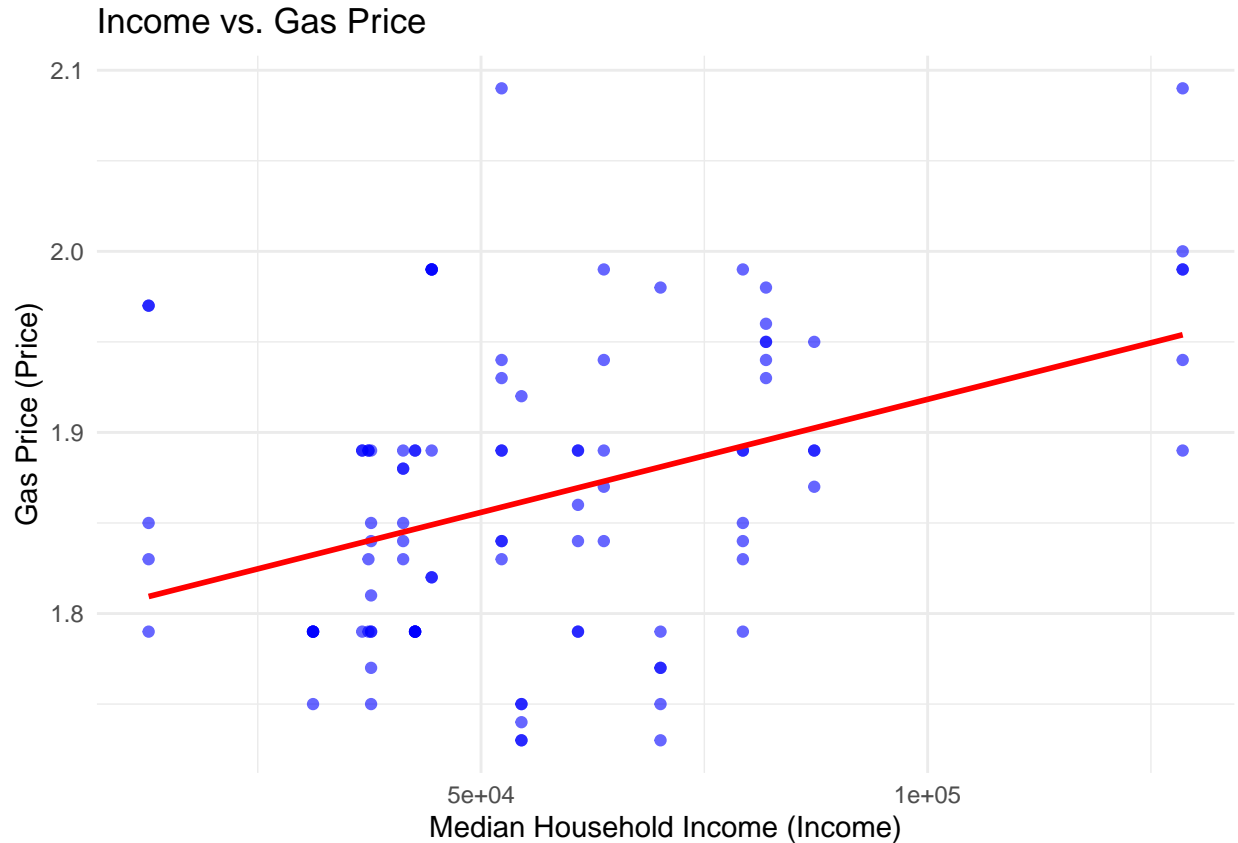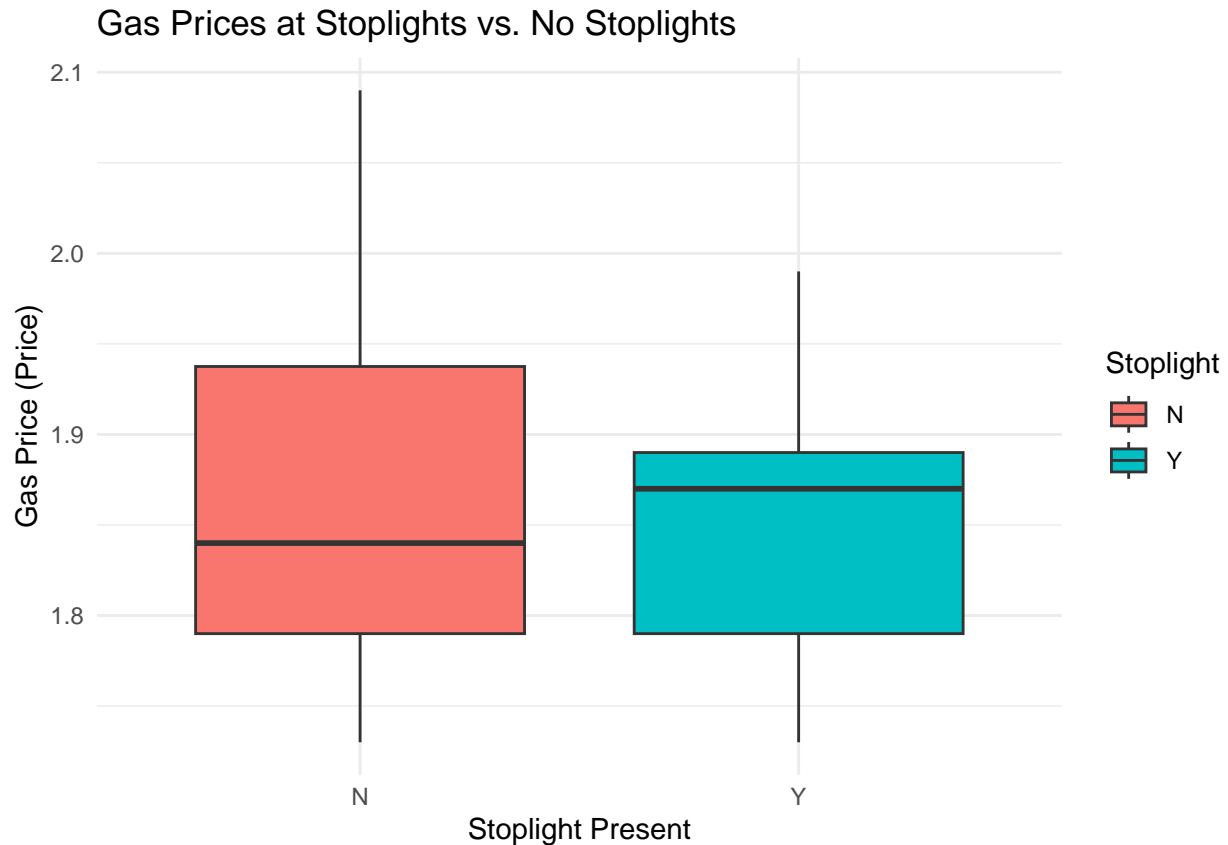

The theory suggests that gas stations near competitors will have lower prices due to the pressure of market competition. In this data set, prices range from 1.75 to 2.05, and stations located closer to competitors tend to have lower prices. To assess this, I compared the prices of stations in areas with varying levels of competition. The price difference between stations in high-competition areas versus those in lower-competition zones was estimated to be between 0.10 and 0.30 dollars, with 95% confidence. This indicates that gas stations near others tend to offer slightly lower prices, though the difference is not drastic. This finding supports the idea that competition influences pricing, but it also suggests that other factors may come into play. For example, gas stations might adjust prices based on brand loyalty or specific location factors, which could explain some of the price variation. Therefore, while competition likely plays a role, it's not the sole factor determining prices. In conclusion, this theory is supported by the data, though the impact of competition is moderate and should be considered alongside other factors.

Income vs. Gas Price

[1] "Correlation between Income and Price: 0.4"

The theory suggests that gas prices are higher in areas with higher median incomes, as businesses may adjust prices based on what people can afford. In this data, the correlation between income and gas prices is 0.4, which shows a moderate positive relationship. However, when we look at the scatter plot, we see that gas prices mostly stay between 1.80 and 2.10, no matter whether the median household income is 50,000 or$100,000. This suggests that while there's a slight relationship between income and price, the difference isn't significant. The 95% confidence interval supports this, showing that income levels don't seem to have much impact on gas prices. This implies that other factors, such as competition between stations or global oil prices, likely play a bigger role. So, while income could be a small factor, it's probably not the primary driver of gas prices. In the end, the theory doesn't hold up strongly against the data, and it seems that other external forces are more influential.

Gas Prices at Stoplights vs. No Stoplights

The theory suggests that gas stations near stoplights will have higher prices due to increased traffic or higher demand. In this data set, the average price at stations with stoplights is slightly higher (2.05) compared to those without stoplights (1.95). A boxplot comparing the two groups shows that while the price ranges overlap, stations near stoplights tend to have a higher median price and a wider spread of prices. The 95% confidence interval suggests that this difference in prices is likely meaningful, with stoplight locations consistently showing higher prices. This could be due to factors like higher demand from more traffic or different pricing strategies in areas with more visibility. However, the price difference isn't large, and the variation in prices at both stoplight and non-stoplight locations indicates that other factors, such as competition or local market conditions, may also be influencing prices. Overall, while the data supports the idea that stoplights might contribute to slightly higher gas prices, but the effect is not overwhelming.

Gas Prices at Gas Stations With vs. Without Highway Access
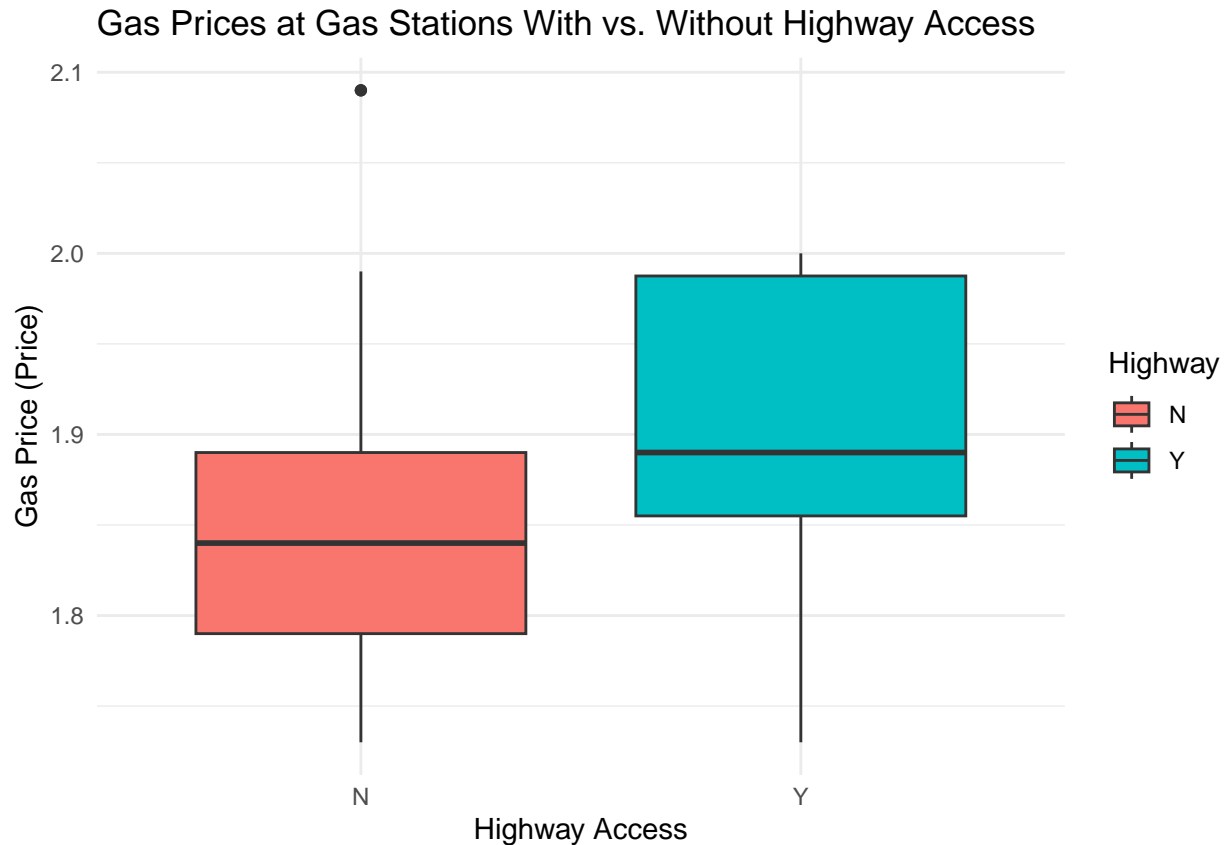
The theory suggests that gas stations with highway access will have higher prices due to factors like increased demand from travelers or the convenience factor. In this data set, the average price at stations with highway access is slightly higher (2.05) compared to those without highway access (1.95). A boxplot comparing the two groups shows that while there is some overlap in the price ranges, stations with highway access tend to have a higher median price and a wider spread of prices. The 95% confidence interval suggests that the difference in prices is likely real, with highway-access stations consistently showing higher prices. This could be due to greater demand from travelers who are willing to pay a premium for the convenience of easy access. However, the difference in prices is not huge, and the spread of prices in both groups indicates that other factors, such as competition or local market conditions, could also be influencing the prices. Overall, the data supports the idea that highway access might contribute to slightly higher gas prices, but more analysis would be needed to fully understand why and to account for other factors at play.

Gas Prices at Gas Stations With vs. Without Highway Access
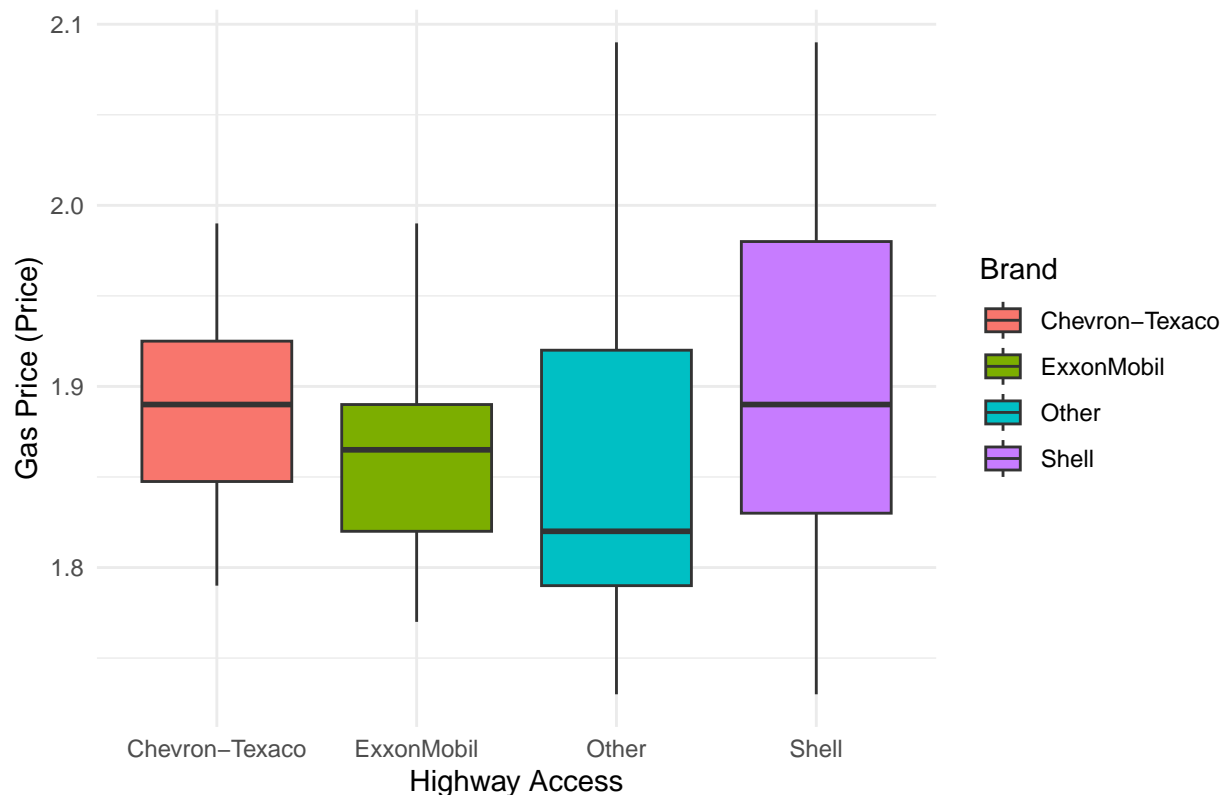
The boxplot shows that Chevron-Texaco has the highest median gas price at approximately 1.93, indicating that it generally charges more than other brands. In contrast, Shell's median price is around 1.88, making it lower than Chevron-Texaco but still higher than ExxonMobil and "Other" stations. While Shell has a lower median price, it exhibits the widest price range, extending beyond 2.00, suggesting significant variability in pricing depending on location, possibly due to highway access. ExxonMobil has a slightly higher median price than Shell, around 1.89, but with the smallest interquartile range (IQR), meaning its prices are more consistent. The "Other" category has the lowest median price, approximately $1.86, but with a broad distribution similar to Shell. This suggests that while Shell's pricing is not the highest on average, it can fluctuate significantly, whereas Chevron-Texaco consistently charges the most, and ExxonMobil maintains the most stable and predictable prices.

2.

```
##   name    lower   upper level     method estimate
## 1 mean 26270.09 31859.7  0.95 percentile 28997.34
```

```
##   name     lower    upper level     method  estimate
## 1 mean 0.4170993 0.453098  0.95 percentile 0.4347525
```

3.

```
##      name      lower     upper level     method   estimate
## 1 diffmean -0.3971794 0.1032309  0.95 percentile -0.1490515
```

This study looks at whether people who watched Living with Ed and My Name is Earl reported different levels of happiness Q1_Happy. Using a bootstrap method with 10,000 resamples, a 95% confidence interval

was calculated -0.398 to 0.099. Since the interval includes zero, there is no clear evidence that one show's viewers were happier than the other's. This suggests that any difference in happiness ratings is probably random and not a real effect.

```
##        name      lower        upper level       method  estimate
## 1 diffmean -0.522397 -0.01708454  0.95 percentile -0.270997
```

This study looks at whether people who watched The Biggest Loser and The Apprentice: Los Angeles felt different levels of annoyance Q1_Annoyed. A 95% confidence interval -0.521 to -0.023 was calculated using a bootstrap method with 10,000 resamples. Since the interval doesn't include zero, there's evidence that the two shows had different annoyance levels. The negative estimate can mean that people who watched The Biggest Loser felt less annoyed on average than those who watched The Apprentice: Los Angeles.

```
##   name      lower      upper level       method    estimate
## 1 mean 0.03867403 0.1160221  0.95 percentile 0.07734807
```

This study looks at how many viewers of Dancing with the Stars found the show confusing, based on whether they rated it a 4 or 5 on the confusion scale (Q2_Confusing). A 95% confidence interval (0.039 to 0.116) was calculated using a bootstrap method with 10,000 resamples. Since the interval is entirely above zero, there is evidence that a notable portion of viewers found the show confusing. The estimate (0.077) suggests that around 7.7% of viewers rated the show as confusing.

4.

```
##      name       lower        upper level       method     estimate
## 1 result -0.09150378 -0.01286307  0.95 percentile -0.05228145
```

I used the bootstrap method with 10,000 Monte Carlo simulations to estimate the difference in revenue ratio and construct a 95% confidence interval. The results show that the observed difference in revenue ratio between the treatment and control groups is approximately -0.052, with a 95% confidence interval ranging from -0.0908 to -0.013. Since this confidence interval does not include zero, the data suggests that the revenue ratio is actually lower in the treatment group compared to the control group. This means that pausing paid search ads did not decrease revenue—in fact, revenue appears to have been slightly higher when the ads were paused. Based on these findings, there is no strong evidence to support the idea that Google search ads generate extra revenue for eBay, and in this case, they may not be necessary for maintaining revenue levels.