**The Relation of Gas Price to**

**Proximity to Other Fueling Stations**

The University of Texas at El Paso

**Spring 2017 - STAT 3320 - Professor Koomson**

**Final Project**

Stephanie Loya sloya4@miners.utep.edu

Stephanie Figueroa sfigueroa2@miners.utep.edu

Victor Fernandez vmfernandez4@miners.utep.edu

Joseph Gutierrez jagutierrez21@miners.utep.edu

**The Relation of Gas Price to**

**Proximity to Other Fueling Stations**

*Introduction:*

Today’s escalating gas prices have become a growing problem for many people who commute daily, whether it is for work, school, errands, etcetera. In order to save money, many people try and look for the cheapest gas prices available in their area. When deciding where to put gas, consumers care about the cost of gas more than any other factor, and they are willing to drive out of their way to find the cheapest prices in order to save a few cents. How can we as a society make more efficient choices when it comes to our car gas and the amount of money that goes towards it?

This study conducts an experiment focused on finding the association between a gas station’s prices and the proximity of that station to other surrounding gas stations. It is hypothesized that the smaller the distance to other stations the lower the prices. This hypothesis is based on the idea that the less distance there is between gas stations, the more that gas stations will compete with one another and consequently drive the gas price lower in more populous areas. In other words, the threat of competition will force the gas stations to provide better prices.

*Data:*

The importance of this research is to understand through data collecting if there is an association between gas prices being lower or higher when there is less distance between gas stations. This research is important as it would help many people to further save money on their fuel expenses. The sample population is focused on gas stations that are within different zip code areas in the state of Texas. Our research helped us understand if gas stations were more susceptible to lower their prices if they had other gas stations surrounding them. Upon completion of this study, we were able to make a generalization on the correlation of Texas gas stations and their prices. The data was collected with custom software[[1]](#footnote-1) made by the team. The software would track the prices in real time using gasbuddy.com as the database. The statistical method used for the data analysis is the scatter plot.

The data collected was also used to determine if time of day has an association with gas prices being cheaper in the morning or evening. This data will help to either support or reject or second null hypothesis, claiming time of day affects gas prices.

*Statistical Methods:*

In this investigation, we decided to use the Scatter Plot (Fig. 1) test because two quantitative variables were being compared. Texas gas stations are being used as the sample population to test if the gas prices are lower when one gas stations is surrounded by more gas stations within certain distance. The purpose was to gather data from a sufficient number of gas stations from all of the state of Texas in order to avoid bias. Texas was chosen as our sample space because it has plenty of urban and rural areas. This mix would allow us to get plenty of data from both isolated and more incorporated stations

Using Minitab 17 we calculated the “Degrees of Freedom”, “P-Value”, the significance level used is 0.05 in order to determine if the results are statistically significant or not so the null hypothesis may be rejected or supported. One lurking factor that can impact the prices of gasoline was the gas station brand, depending on the brand they rise their prices because they claim to have additives in their gas to improve quality.

*Results:*

Results were obtained using Minitab 17 by creating a scatter plot(Fig 1). By interpretation of the scatter plot we can say that there is no real evidence that suggests gas prices are lower the less distance there is between gas stations. The cluster found on the scatter plot shows that there is no real change on prices according to the distances between gas stations.

*Discussion:*

The findings attained as a result of this project’s outcomes are incredibly beneficial and applicable to real life use. As a group, we are able to better understand the fluctuation between gas prices in accordance to the location, time, and distance between each gas station. This information is valuable to the Texas population who commutes to work, is going to plan a road trip across Texas, or simply likes to save money on gas. We believe that one of the strongest factors in our study is the large amount of data we were able to collect and analyze, using our custom built web scraping program, as it gives this investigation solid evidence for our outcomes. One weak point that we found was that many gas station’s prices are not updated as often as others, meaning that we were potentially getting repeated data from them. In order to receive more accurate results and expand on our hypothesis, this investigation may be improved in the future by using data from other states, as the results of this project generalize the Texas population and landscape specifically. Another major advantage would be to be able to run the study for as long a period as possible. Given enough time, other factors like season, political climate, and economic stressors could also be factored in and discerned from the data. Furthermore, another way to improve this investigation may be by comparing the gas stations by brands, in order to examine the most economically friendly brands according to area.

*Appendix:*

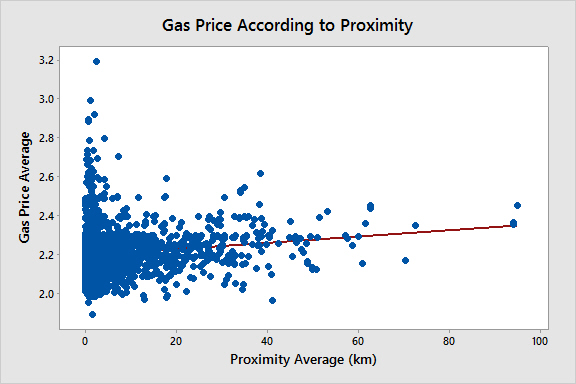


Figure 1. Scatterplot

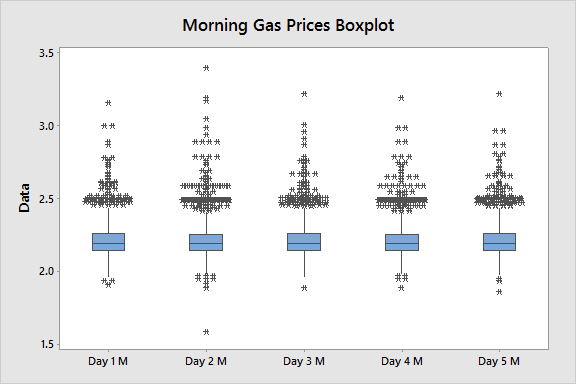


Figure 2. Morning Prices Boxplot

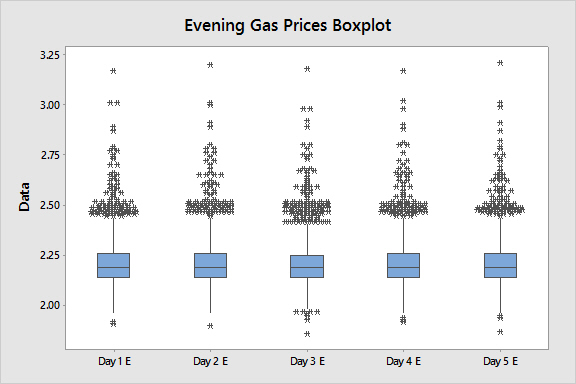


Figure 3. Evening Prices Boxplot

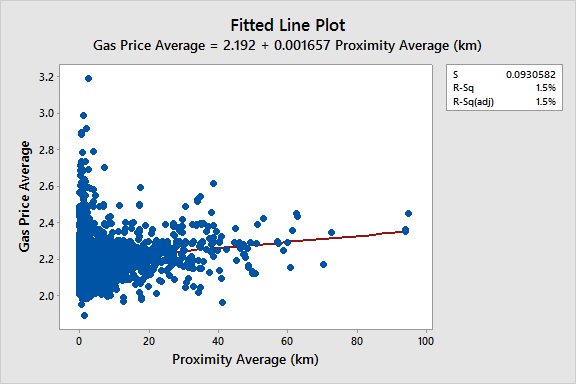


Figure 4. Fitted Line Plot

1. <https://github.com/victormfer/Stats-Project/blob/master/priceWebScraper.java> [↑](#footnote-ref-1)