D214: Data Analytics Graduate Capstone

Executive Summary

Performance Assessment

Task 3

Logan Rosemeyer

Western Governor's University

D214: Data Analytics Graduate Capstone

Dr. Daniel Smith

May 9, 2023

Executive Summary: Gas vs. Grocery

The purpose of this report is to find a relationship between gas prices and prices of common grocery items that can be used to predict what grocery prices will look like going forward. This relationship can be beneficial to both grocers and distributors.

Problem and Hypothesis

The question that needs to be answered is: Can we use gas prices to predict grocery prices? My null hypothesis is that there is not a statistically significant relationship between gas and grocery prices. My alternate hypothesis is that there is a statistically significant relationship between gas and grocery prices. I will use the Granger Causality Test to test the relationship. If our p-value is lower than 0.05, we will reject the null hypothesis.

Data-Analysis Process

I collected monthly data on grocery and diesel prices from 1980 to 2023 from the US Department of Labor. Because of some discrepancies in the time periods for the data, I decided to start all of my data at the year 1998. The lettuce data set only had data until 2020, so I had to make a separate data frame with diesel prices and lettuce prices from 1998 to 2020. There were also some missing prices in a few of the months. To counter this, I imputed the average of the previous month and the next month for the missing observation. Once all of the data was clean, I ran the Granger Causality Test on all of the grocery items and diesel prices. Every time I ran the test with diesel being the predictor for the grocery item, I received a p-value of less than 0.05. Every time I ran the test with the grocery item being the predictor for diesel, I received a p-value of more than 0.05.

Findings

My findings indicate that there is a statistically significant relationship between gas and grocery prices that allow us to use gas prices to predict grocery prices.

Limitations

One limitation of this study is that I did not account for general inflation. Taking inflation out would give us an even better idea of how gas prices influence grocery prices. Another limitation is that the data did have some missing values. I imputed the average of the previous month and the next month, but there is no way to know if these are accurate or not.

Action Steps

The next step in this study would be to take inflation into account. It is possible that this would give us some different numbers. Another step to take would be to look at the relationship between gas prices and grocery items that are made in a factory. Most of my grocery items come from farms. Harvesting these items uses a lot of gas, where products made in factories do not necessarily require gas to produce. This difference in production could mean that the farm products that I studied are influenced more by gas prices than the factory products.

Benefits

Grocers and distributors alike can benefit from my study. Grocers can use this study to predict what to set their prices at, as well as what they might have to pay the distributors. Distributors will also find this information valuable because this will allow them to determine their shipping expenses and set their prices accordingly.