Introduction:

Understanding the relationship between Natural Gas Prices and Heating Degree Days (HDD) holds significance in forecasting and analyzing real-world metrics, such as in predicting oil future prices. This project aims to investigate the relationship, if any, that exists directly between HDD and Natural Gas Prices.

Background and Data Sources:

HDD days describes a metric used in meteorology and energy consumption analysis to quantify how much and for how long the outdoor temperature is below a certain level, in this case 65 degrees Fahrenheit. Natural Gas Prices are “priced per million British thermal units (BTUs) - one BTU is the amount of energy needed to change one pound of water by one degree Fahrenheit,”(Nasdaq). Finally, futures indicate that a buyer will buy a specific amount of natural gas at a fixed date and price in the future. The data sources for this project can be found at the end of the document. All data sources are reporting on national averages form the United States.

HDD and Natural Gas Price

A graph of gas prices

Description automatically generated

As can be viewed on the graph, the HDD follows a mostly consistent sinusoidal pattern, while natural gas price follows a much more erratic pattern. This can be explained easily. Heating Degree days are a simple environmental measure, where consistency can be expected. The HDD for a given month in winter will consistently reflect the colder temperatures that the US experiences. Winters have variation in temperature, causing the different peaks in HDD over the years.

Natural Gas Prices, however, are much more tied to the state of the world and the people in it. According to an energy outlook report conducted by the US Energy Information Administration, the spikes in in December of 2000, and February 2014 can be explained by historically cold weather. The spike in 2005 is a direct result of hurricanes Katrina and Rita. Furthermore, the 2008 economic collapse led to the peak in the corresponding year.

Directly correlating data produces no meaningful results, producing r2= 0.007, as to be expected. To analyze the data more thoroughly, the data was broken into the trend line (underlying pattern or tendency of the data) , the seasonal pattern(regular, recurring fluctuations or patterns within the data) and the residual (data that remains after removing the trend and seasonal components).

A graph of a computer

Description automatically generated with medium confidence

A screenshot of a graph

Description automatically generated

The seasonal trend lines, of the Gas prices are of particular note here. As can be seen, there is a regular yearly large spike, and another regular smaller spike. This corresponds to a seaonal increase in price in both the winter in the summer. “Residential and commercial sector consumption of natural gas typically triples between the summer (April – September) and winter (October – March) because of demand for space heating” (US Energy Information Administration). However this smaller spike in the summer can be explained by uptick in electricity demand (where electricity is generated by natural gas) for air conditioning.

The following is the graph comparing the seasonal lines of HDD vs Natural Gas Price

A graph showing a graph of a graph

Description automatically generated with medium confidence

This graph shows a stronger correlation of r2=0.311. In addition to HDD, the corresponding metric for cooling degree days (CDD) is accounted for in the Natural Gas Price seasonal trend, while it is not accounted for in the HDD line. By filtering out months April-September, we can see a higher level of correlation between seasonal trend lines where r2= 0.417

A graph of a comparison of seasonal data

Description automatically generated

Even the with the just the seasonal trends extrapolated and the warmer months removed, there is still not a incredibly strong correlation between HDD and Natural Gas Price. Furthermore, no obvious correlation or trend line exists when looking at a scatter plot between HDD and Natural Gas Price:

A graph of blue dots

Description automatically generated

As can be seen, no clear trend exists where strictly higher HDD leads to strictly higher Gas Price. The conclusion can therefore be drawn that Natural Gas Price and HDD are not directly correlated.

Oil Futures:

The price of predicting Natural Gas Oil Futures was of interest. If a correlation between HDD and Natural Gas could be found, a model could be built to potentially predict oil future prices. Different types of oil futures exists for different types of oil (crude, natural gas, etc.). Natural Gas Futures are heavily dependent on the price of natural gas as can be seen below (r2=0.93)

A graph of gas prices

Description automatically generated

Because Oil Futures are so heavily based on the price of natural gas, it reasonably follows that they are strongly correlated. And furthermore the prediction/correlation of these futures with HDD has the same issues.

When examining the seasonal increases in Natural Gas Prices, the obvious suspected reason that the price increases seasonally is the increase in consumption. It was therefore intriguing to research the correlation between HDD and the consumption of natural gas. While the price of natural gas has many causes for its fluctuation, it is logical to suspect that an increase in demand/consumption for natural gas would lead to a higher price. If a model can be accurately created between HDD and natural gas consumption, then this could potentially be of use for a larger model taking into account more factors for the prediction of natural gas price (and by extension oil future prices).

The correlation between HDD and natural gas consumption is logical and well documented. Unfortunately, the publicly available datasets were not able to show any correlation:

A graph with blue dots and a red line

Description automatically generated

Even when removing the summer months, no meaningful correlation above r2>=0.1 is achieved.

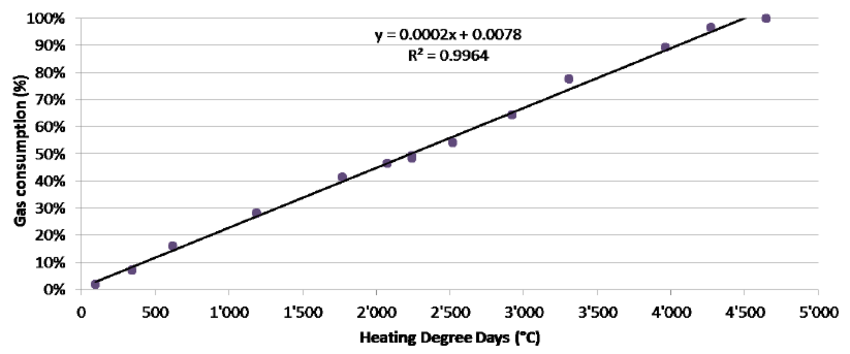
A graph of blue dots and red line

Description automatically generated

With other datasets in a more limited time period the following figures had been produced:

A graph of a graph showing the average temperature of a natural gas company

Description automatically generated with medium confidence



Conclusion:

A clear seasonal pattern exists among HDD, Natural Gas Price, and Natural Gas Futures. Due to falling temperatures in the Winter months, all three metrics have a tendency to increase during these times. However, using a monthly dataset over many years, the correlation between HDD and Natural Gas Price cannot be directly observed. The varied uses of natural gas for heating, cooling and other energy uses, as well as the many worldly factors that affect the supply and demand of natural gas makes the metric have many more components factor into it’s price than just HDD. Although it was not evident with publicly available datasets, the correlation between HDD and natural gas consumption may hold more insight for predicting the supply, demand, and price of natural gas price.

Link to open-source repository: <https://github.com/johnCavatelli/CS-5990-Gas-Price>

Sources [Data was sourced from freely available online resources]

Natural Gas Prices (reported daily and averaged): <https://datahub.io/core/natural-gas#resource-daily>

Monthly HDD: <https://www.eia.gov/totalenergy/data/browser/index.php?tbl=T01.10#/?f=M&start=197301&end=202307&charted=32-10>

Natural Gas Future Prices:<https://www.investing.com/commodities/natural-gas-historical-data>

Natural Gas Consumption: <https://www.eia.gov/dnav/ng/hist/n9140us2m.htm>

Correlation between the Heating Degree Days and the gas consumption (from 2011 to 2013): <https://www.researchgate.net/figure/Correlation-between-the-Heating-Degree-Days-and-the-gas-consumption-from-2011-to-2013_fig3_290019133>

Report On Natural Gas Prices: <https://www.eia.gov/outlooks/steo/special/pdf/2014_sp_03.pdf>