Kim’s script: 8-9 mins

Hello everyone!

We are Two Gassy Girls!

My name is Kim and my teammate’s name is Phoebe.

We have worked on Gas Price Analysis, I hope you understand well and hope you guys enjoy!

This is the deliverables for our presentation and I’m going to speak on the introduction and some graphs and I’m going to also show you our website that we created.

Our gas price analysis uses Python, ETL, Web Scraping, and Machine Learning to figure out the patterns of the past Gas Price and forecast future gas price in the U.S. We performed Regression modeling based on various factors that we thought could affect Gas Price. Our goal was to set to provide users the access to the information on the future gas price.

Our Project objective: to Analyze gas price from the data from 2 decades and find the best regression model to predict gas price for the future.

Target audience can be anyone who wants to know our gas price analysis 😊

First of all, I want us to understand about gas price. We actually focused on Natural gas for our analysis.

There is Gas Imports and Exports and prices from both.

There is Supply, Demand, and storage. Supply should be the gas production and Demand should be gas consumption. We wanted to also include gas storage from underground to know how it could affect gas price.

When we did research on Gas, there were actually a lot of things that we had to learn and the first thing was the types of natural gas. There are All Grades all formulation retail, regular, midgrade, premium, and Diesel.

We were going to stop here for our analysis, after talking to our TA’s, we wanted to think of more factors that could be related to gas price.

And here we are! We were wondering maybe gas price could change by consumer types, for example, nowadays there are more people who drive electric cars so there are more in use of electric power, so we thought it could affect gas price. Also, as we think of our daily life, we are experiencing inflation in everything, so we thought inflation could affect gas price as well.

We found the right data csv files and worked in Python. But before that, we wanted to see what columns we have and how they can relate to each other. This is how our ERD looks like, and as we can see, the “Year” is the connection for each data. This helped because we knew what to do with our raw data.

Here is our target variable which is U.S. All Grades All Formulations Retail Gasoline Prices.

We wanted to put in graph so we can see how the prices have been changing since 1990s. After 2020, it looks like it’s increasing.

This is a graph showing some trend among 4 factors I picked for our all grades gasoline price. We wanted to see how these 5 variables would look like in the past and see if there is any similar trend.

It looks like the line graph for all grades gasoline price does not really look similar to other factors but we wanted to see more in our machine learning models. You will see the results in the later slides.

* LATER

We created a website from web scraping for latest gas price news articles.