The Impact of Gasoline Prices on U.S. Gasoline Consumption

Nikolas Lee-Bishop

Overview:

The research question this project will be asking is, “Does an increase in gasoline prices lead to a decrease in gasoline consumption in the United States?” The analysis will be guided by the principle that as prices increase, quantity demanded tends to decrease assuming other factors remain constant. Gas prices tend to be influenced by global events and holiday periods that change travel patterns. The analysis will only focus on the National US level to ensure data consistency.

Consumption Data1:

All headers of the sheet (rows 1,2) and columns C-G were removed. This left the *Date* column and *Consumption (Thousands of Barrels)*. On the *Date* column, the datatype was converted to date only and set as the first of each month. The data was filtered to start at April 1993 to match the start date of the **Price** data and to end in December 2024. The *Consumption (Thousands of Barrels)* column was then converted from Barrels into Gallons to match the same standard of measurement as the Price data. Next it was transformed into *Gasoline\_Consumption(Millions of Gallons)* to improve readability of the numbers.

Once data cleaning was done, basic statistics were conducted. After that different charts were visualized to view trends in the data including time-series plots, rolling averages, and heatmaps.

Prices Data2:

All headers of the sheet (rows 1,2) and columns C-P were removed. This left the *Date* column and *U.S. All Grade All Formulations Retail Gasoline Prices (Dollars Per Gallon)* which was renamed as *Dollars\_Per\_Gallon*. The *Date* column datatype was converted to date only and set as the first of each month. Then the data was filtered to start at April 1993 as that was when the first data point of that column was and to end in December 2024.

Once data cleaning was done, basic statistics were conducted. After that different charts were visualized to view trends in the data including line plots, heatmaps, and time-series plots.

Merged Data:

Once each dataset had been cleaned and explored, the two sets were merged on *Date* using an inner join. This combined dataset will be what’s used for correlation and regression analysis to test the projects main question.

Additional Data:

* To keep the project simple, I have currently excluded the **Table\_Automobile\_Profile** dataset which contains additional data such as vehicle registrations and vehicle miles traveled from 1993-2021.
* I also plan to find and use additional sources that talks about events that affected gasoline prices and production such as the Arab Spring, COVID, and Russia’s Invasion of Ukraine. This data will be used to help explain inflections in price of gasoline.
* In the **Prices** data section, I included a chart which shows the difference in nominal price vs inflation adjusted price. The data to create that chart was retrieved from the Federal Reserve Economic Data (FRED) database. The Consumer Price Index (CPI) dataset used was the ‘Consumer Price Index for All Urban Consumers: All Items in U.S. City Average (CPIAUCSL)’ 3.

Citations:

1. U.S. Energy Information Administration. (n.d.). *U.S. Product Supplied of Finished Motor Gasoline (Thousand Barrels per Day)* [Data set]. U.S. Department of Energy. Retrieved October 26, 2025, from <https://www.eia.gov/dnav/pet/pet_cons_psup_a_epm0f_vpp_mbbl_m.htm>
2. U.S. Energy Information Administration. (n.d.). *U.S. All Grades All Formulations Retail Gasoline Prices (Dollars per Gallon)* [Data set]. U.S. Department of Energy. Retrieved October 26, 2025, from <https://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_m.htm>
3. U.S. Bureau of Labor Statistics. (n.d.). *Consumer Price Index for All Urban Consumers: All Items (CPIAUCSL)* [Data set]. Federal Reserve Bank of St. Louis. Retrieved October 26, 2025, from <https://fred.stlouisfed.org/series/CPIAUCSL>