GAS PRICE ANALYSIS AND PREDICTION PROJECT

Thu V Nguyen

Dataset: U.S. Regular All Formulations Gas Price (GASREGW)

- **Source**: FED - GASREGW
- **Description**: Weekly U.S. average price for regular gasoline, all formulations.
- **Frequency**: Weekly (ending Monday)
- **Units**: Dollars per gallon
- **Coverage**: August 20, 1990 - Present

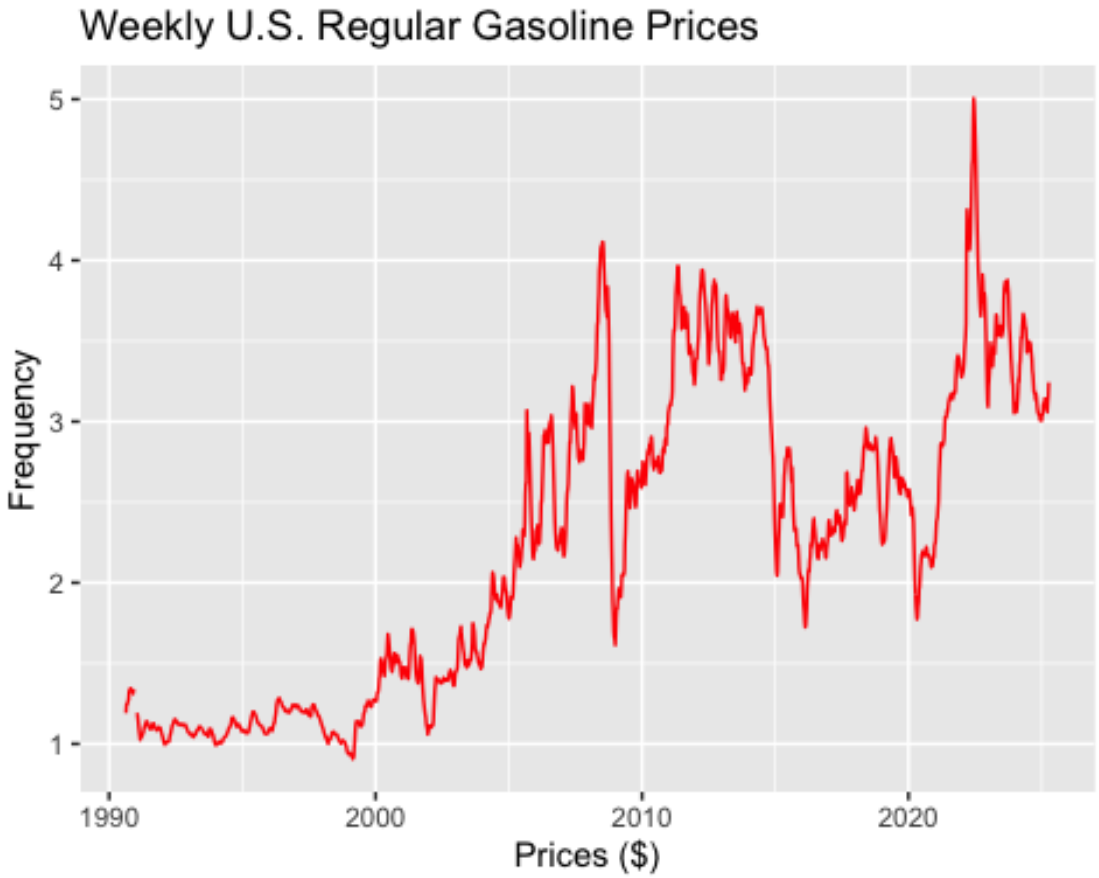
Gas Price Trend of All Time

Is the Price at the Pump Predictable?

In a time of economic uncertainty and volatile energy markets, predicting gas prices isn't just a numbers game; it's a powerful tool for anticipating consumer behavior, optimizing logistics, and informing policy. This project explores the intersection of historical data and economic indicators to build a predictive model for U.S. gas prices, combining statistical rigor with real world relevance.

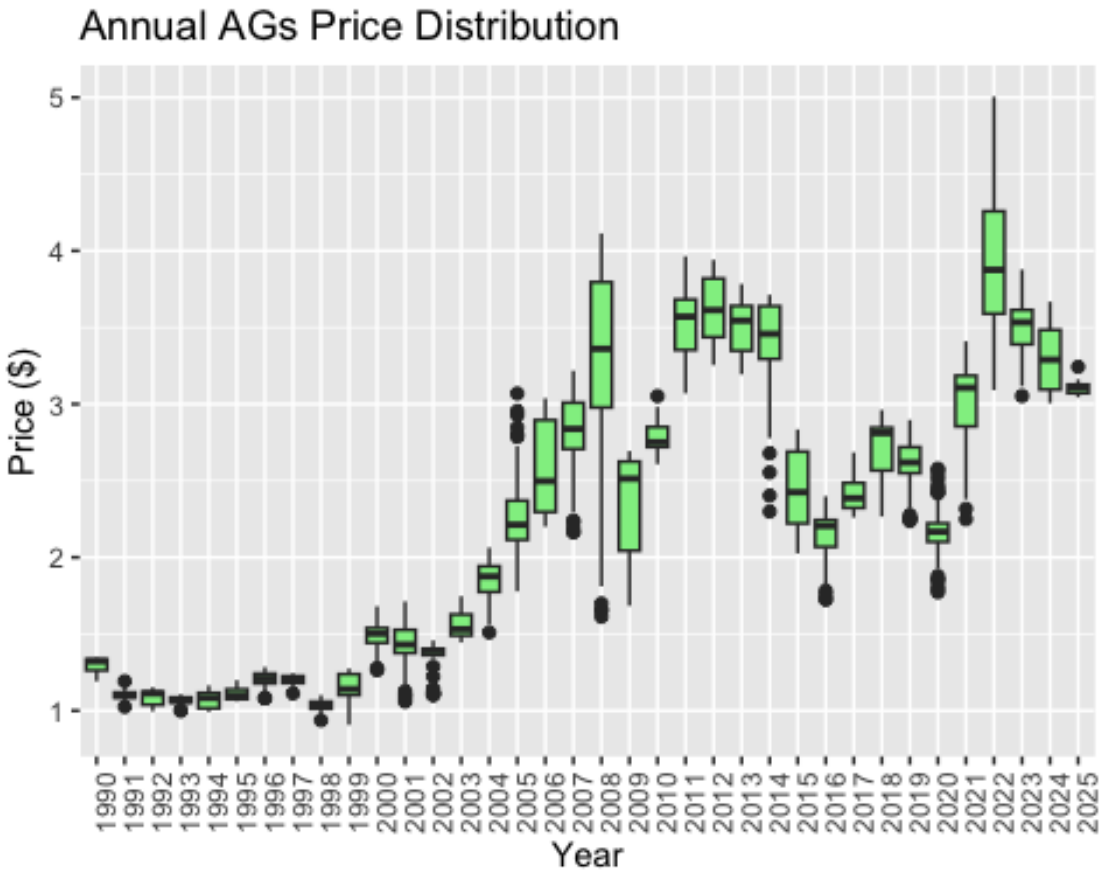
Under the Hood: Exploring What Drives Gas Prices

Gas Prices Over Time



At a glance, you can see that gas prices have experienced significant ups and downs over the past few decades. From the early 1990s to the early 2000s, prices were relatively stable and low, mostly hovering around $1–$2 per gallon. However, starting around 2005, there's a noticeable surge—prices began climbing sharply, peaking just before the 2008 financial crisis. After that, they dipped briefly but rose again through the early 2010s, reaching another peak close to $4 per gallon. Following 2014, there's a sharp decline, and prices fluctuated more moderately for a few years. Then, in the post-2020 period, there's another dramatic spike—likely tied to global supply chain issues and energy market disruptions following the COVID-19 pandemic and geopolitical tensions.

Prices by Year



- Early 1990s to early 2000s: Prices were relatively low and stable, with very little variation within each year. Most boxplots are tight, close to the $1–$1.50 range.

- 2005–2008: This period marks a sharp increase in gas prices and volatility. The boxes are taller and whiskers longer, showing wide price swings within a year. The 2008 box, in particular, reflects the peak just before the financial crisis.

- 2011–2014: Prices remained high, typically around $3.5–$4 per gallon, but with less year-to-year volatility compared to 2008.

- 2015–2020: A clear decline in prices, along with more stability. However, 2020 starts to show fluctuation again—likely linked to the COVID-19 pandemic's market disruptions.

- 2022 spike: One of the tallest and highest-positioned boxes, indicating both high average prices and extreme volatility, likely tied to post-pandemic recovery and geopolitical events (e.g., war in Ukraine).

- 2023–2025: Prices appear to be cooling off and stabilizing, although they remain higher than pre-2020 averages.

Testing the Trends: Hypothesis Testing and ANOVA on Gas Price Variations

The t-test compares the average gas prices between 2020 and 2022. The results show:

- 2020 mean: $2.17

- 2022 mean: $3.95

- p-value: <2.2e-16 (extremely small)

There's a highly significant difference in average gas prices between these two years. Prices in 2022 were substantially higher than in 2020, and the result is statistically robust. The confidence interval for the mean difference (about $1.63–$1.93) supports that the increase wasn't due to random variation.

An ANOVA test to compare gas prices across five years: 2019 to 2023.

- The price range goes from about $1.77 (min) to $5.01 (max)

- 2022 stands out with higher quartiles and extreme highs

- There's visible variation in means between years:

+ 2020 average ~ $2.5

+ 2021 average ~ $3.1

+ 2022 average ~ $3.95

+ 2023 average ~ $3.6

Binary Boom: Visualizing the Likelihood of Gas Price Surges

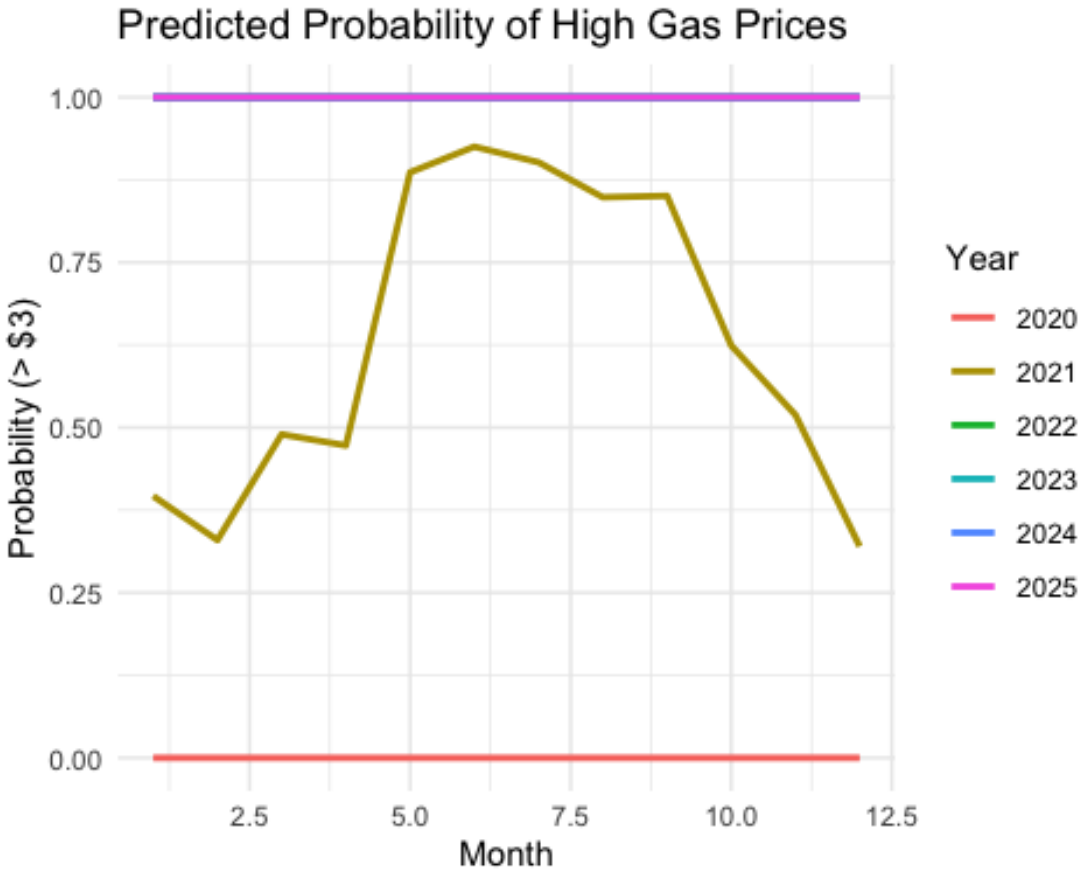
- From May to September, these months have postive and significant coefficients, indicating higher chances of gas prices exceeding $3. Especially, the model suggests a summer spike in gas prices

- Most older years (1991–2019) have coefficients near zero with p > 0.99 — meaning they're statistically insignificant.

- Recent years (2021–2025) have very large positive coefficients — e.g., factor(year)2022 = 45.8. These indicate a very high likelihood of high gas prices in recent years.

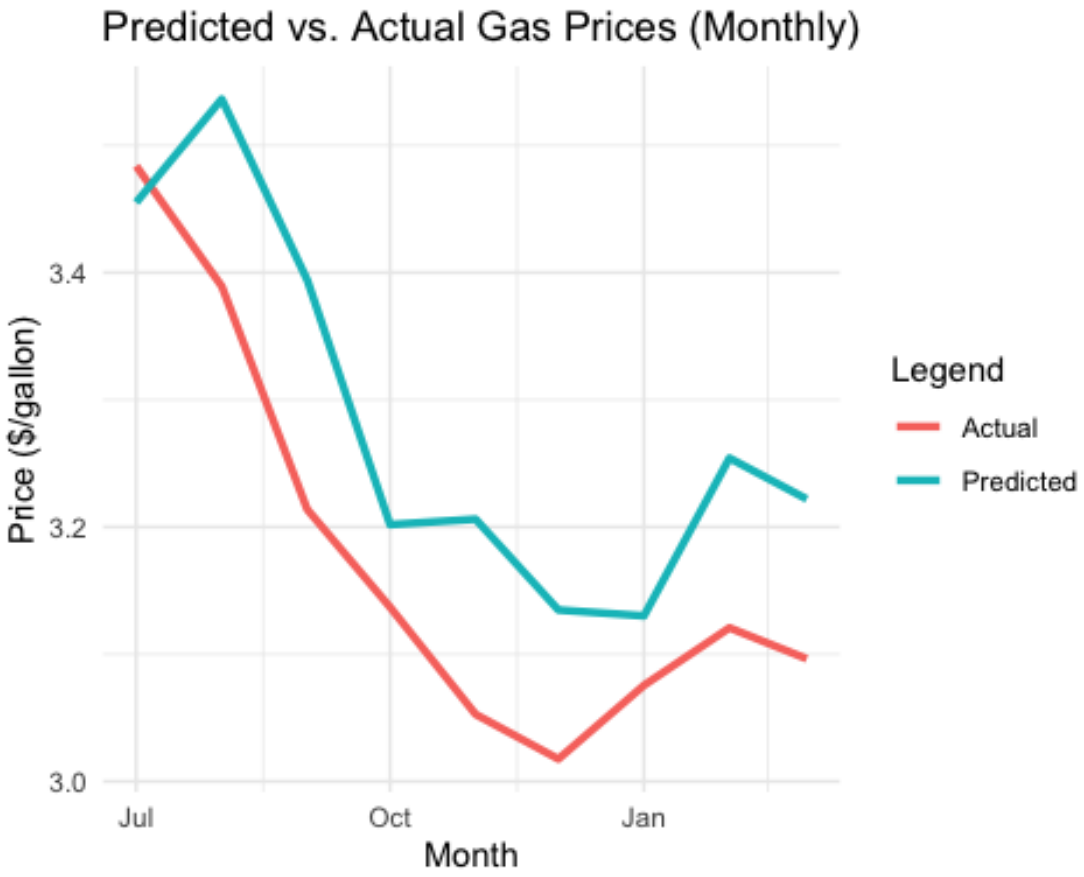
- The prediction gives 1, means there is a ~100% predicted probability that gas price exceeded $3 in July 2022.

Visualizing the Odds: Predicted Probabilities from VAR Forecasts



Drive or Detour? Evaluating the Roadmap of Model Accuracy

MAPE: 3.54 %



- The graph illustrates a comparison between predicted and actual U.S. monthly gas prices, supported by a strong model performance with a MAPE of 3.54%. From July to September, the model closely aligns with real-world prices, accurately capturing the initial downward trend. However, from October to January, the model consistently overpredicts, anticipating price stabilization earlier than observed. While the model correctly identifies the upward shift beginning in February, it slightly overestimates the strength of the recovery. Overall, the model demonstrates high accuracy in tracking long-term trends but shows room for improvement in adapting to short-term market fluctuations and shocks.

-THE END-

Thank you for taking the time to explore this project. I hope it offered valuable insights into the power of data-driven forecasting and the challenges of modeling real-world volatility. I look forward to any feedback, questions, or opportunities to further develop and apply these insights.