Logan Lomonaco

# Search trends and CPI

# Data

Monthly data, from 2004-01-01 to 2025-03-01.

Google trend data explanation: “Numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means there was not enough data for this term.”

**From FRED:**

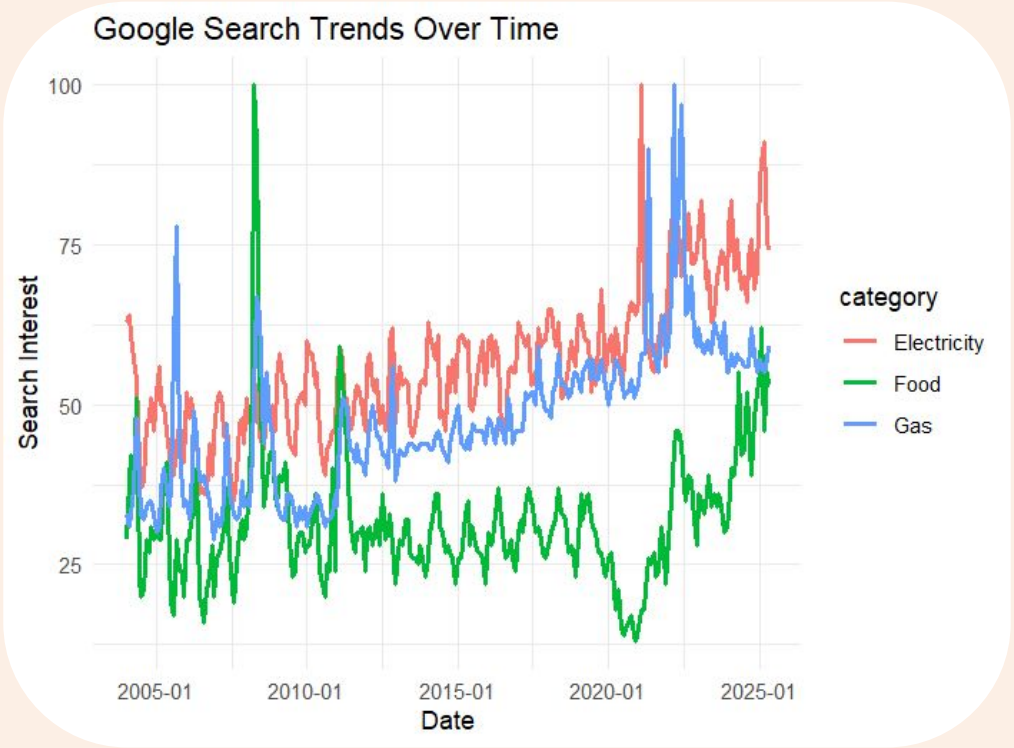
- Consumer Price Index for All Urban Consumers: All Items in U.S. City Average (CPIAUCSL)
- University of Michigan: Consumer Sentiment (UMCSENT)
- Unemployment Rate (UNRATE)
- Real Disposable Personal Income (DSPIC96)
- Producer Price Index by Commodity: All Commodities (PPIACO)

**From Google Trends:**

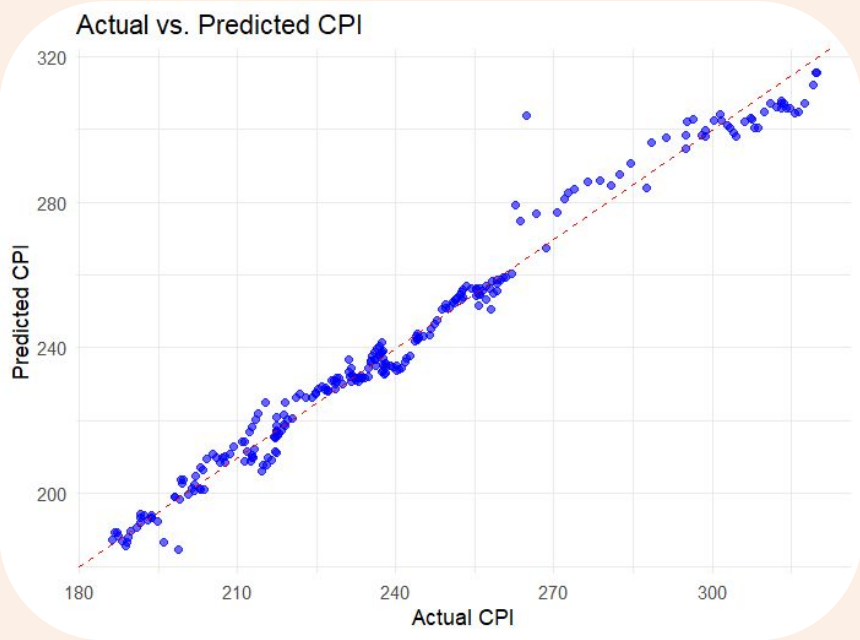
Focused on search trends about food, electricity and gasoline

This can easily be expanded, but both GtrendR and GtrendsR packages for R are having difficulty web scraping data, so manual data download is currently necessary.

# Google Trends data visualized:



# Results of simple linear regression



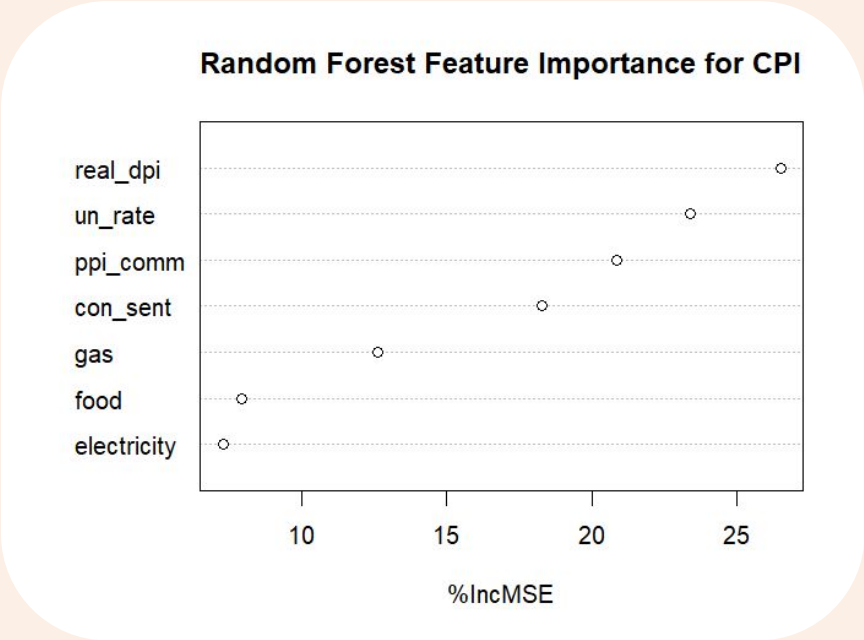

**Model Fit:**

- **High Predictive Power ($R^2$ = 97.98%)** – The model effectively explains CPI fluctuations.
- **Significant F-statistic (p < 2.2e-16)** – Strong statistical confidence in predictor relationships.

**Important Predictors:**

- **PPI Commodities (+) & Real Disposable Income (+)** → Major drivers of CPI inflation.
- **Unemployment Rate (-)** → Higher unemployment reduces CPI, aligning with Phillips Curve Theory.
- **Electricity Trends (+)** → Search spikes correlate with inflationary pressures.
- **Gas Trends (-)** → Unexpected negative correlation, possibly reflecting consumer distress.

# Results of random forest




**Most Important Predictors (Higher %IncMSE & IncNodePurity):**

- **Real Disposable Income (real_dpi) – Top predictor for CPI. Higher disposable income boosts inflation, making it a key driver.**
- **Producer Price Index (ppi_comm) – Strong influence on CPI, reinforcing wholesale cost pressure effects.**
- **Unemployment Rate (un_rate) – CPI drops when unemployment rises, consistent with economic theory.**

**Moderate Importance Variables:**

- **Consumer Sentiment (con_sent) – Moderate effect but weaker than economic fundamentals.**
- **Gas Trends (gas_t) – Some impact, but less predictive than disposable income or wholesale costs.**

**Least Important Variables:**

- **Food Trends (food_t) & Electricity Trends (electricity_t) – Lower importance suggests search trends may not directly drive CPI. They could be more indirect signals.**

# Next steps:

- **Experiment with different regression models to find the one with the best fit**
- **See if there more data and variables that could be added to make the regressions and finding more robust**
- **Implement sentiment analysis of news sources to supplement the data and regressions (as time allows)**
- **Visualize the data and regressions to better show relationships between CPI and other variables as well as other insights**
- **Find more ways to implement labs and workshops into the project where they fit my use case (and as time allows)**

# Thank You!

## Questions?

Is there anything I can clarify, go over, or expand upon?

## Thoughts?

What are your impressions of the project?

## Suggestions?

Any ideas, constructive criticism, or improvements that you would like to suggest?