Natural Gas Consumption & Air Quality Analysis

—

By Sadah Hakim, Leo Crowal, Jennifer Harris
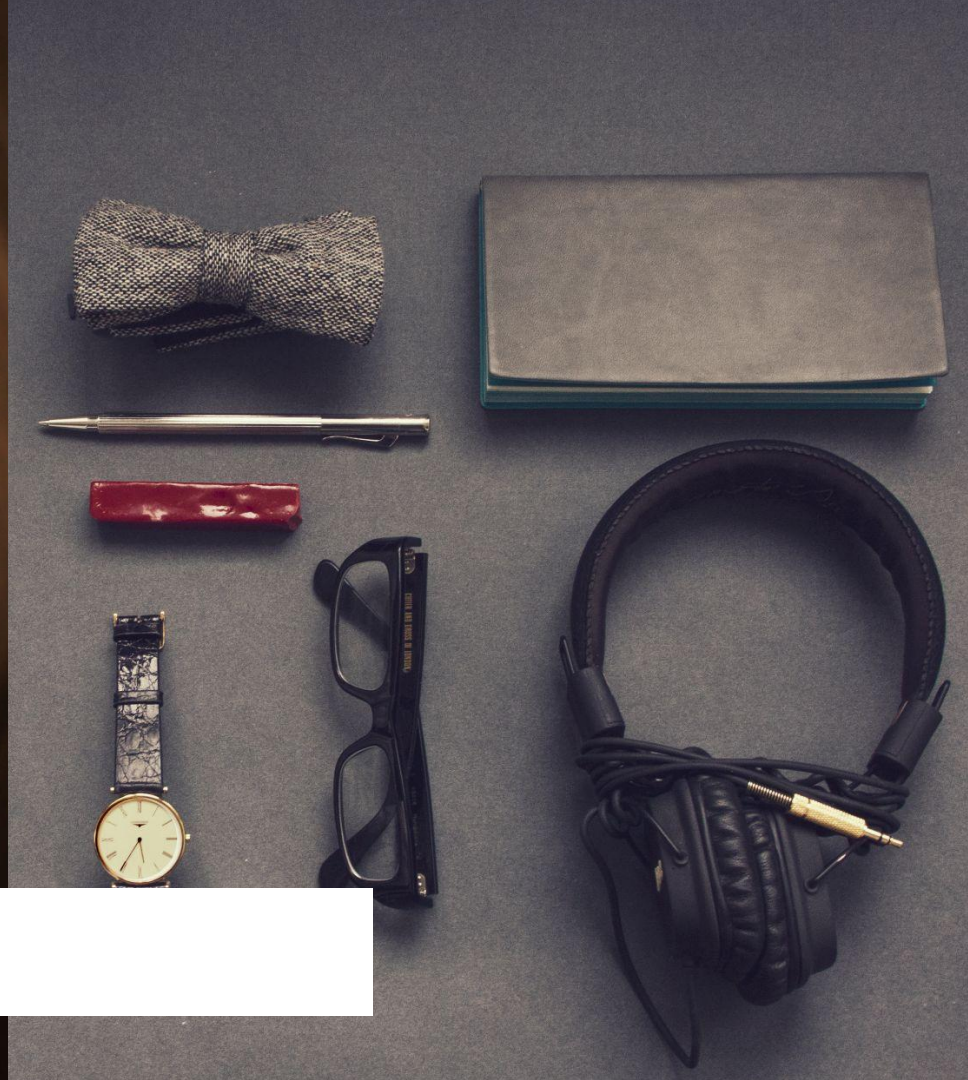
How does natural gas consumption impact air quality?

Hypothesis

Hypotheses: Higher Natural Gas Consumption Leads to Poorer Air Quality. We assume that states with higher natural gas consumption will have poorer air quality due to increased emissions.

Our Analysis

Analysis

1. **Geographical view of Natural Gas Consumption and Air Quality Index:** A State-Level visualization (State level consumption and air quality data by month)

2. **Impact of Seasonal Natural Gas Consumption on Air Quality**: Analyze how seasonal variations in natural gas consumption influence air quality, narrowing it down to top-10 States.

3. **Correlation Between Natural Gas Consumption and Air Quality**: To find out whether there is a significant correlation between natural gas consumption and air quality levels across different regions.

Materials

Datasets used in our project are found on kaggle:

1. Natural Gas Consumption (The data comes from the U.S. Energy Information Administration (EIA))

2. US Air Quality (The data comes from the US Environmental Protection Agency (EPA))

- Jupyter Notebook (for data cleanup and data analysis)

- SQLite for data storage

- Python Pandas, Numpy and sqlalchemy libraries for analysis, Bokeh library and Leaflet maps for visuals

- JavaScript and D3 to create the leaflet map and design

- Flask backend with interactive API routes

Procedure

1. **Data Merging:** Cleaned and merged two key datasets—Air Quality Index (AQI) and Natural Gas Consumption—based on the state and month columns(for the purpose of this project to get a complete data we filtered the datasets for the year 2021 only). This resulted in a consolidated, clean dataset saved as combined_df.csv.

2. **Database Creation:** Established a SQLite database using the cleaned dataset. All data for subsequent analysis is extracted from this database.

3. **Interactive Visuals:** Leveraged JavaScript, Leaflet, D3 and Bokeh libraries to create dynamic and interactive visualizations.

4. **API Development:** Created Flask app to build API routes, enabling seamless interaction and integration of our data
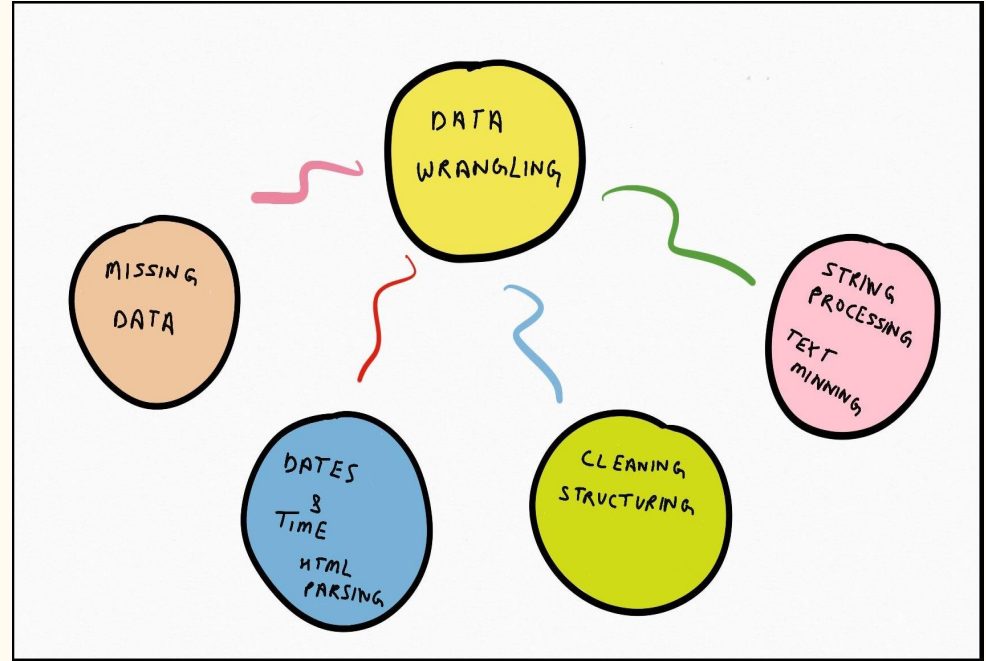
Geographical view of analysis

1. Data cleaning:
    a. Aggregated daily values to get monthly view and combine data sets
    b. Calculated average values by state and by month
    c. Removed cities to view state level data and aggregated values accordingly
2. Created Flask app to route API with Natural Gas consumption and AQI data by state and by month
3. Utilized Leaflet library to visualize data geographically and D3 library to create interactive visualizations

Key observations:

- While Natural Gas is not the only factor impacting Air Quality Index, there is a correlation between consumption and air quality

- States with overall higher population count have higher natural gas consumption
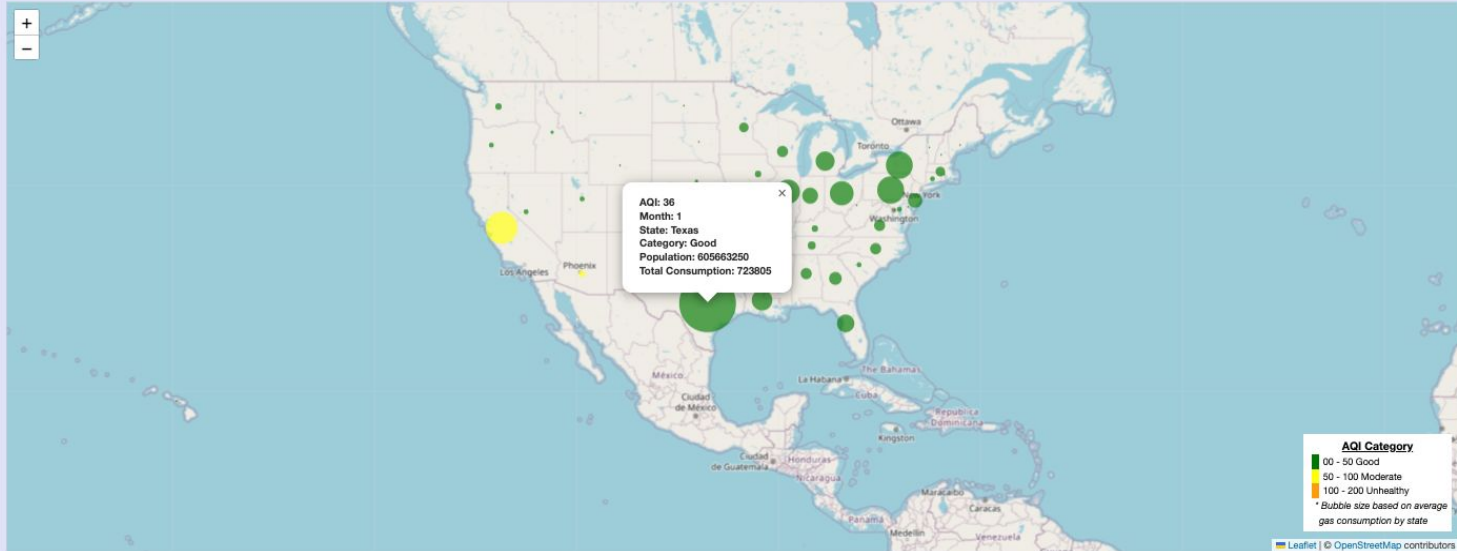
Demo

Natural Gas Consumption and Air Quality Analysis

Seasonality   Correlation

Consumption And Air Quality by Month

January ▾

AQI: 36
Month: 1
State: Texas
Category: Good
Population: 605663250
Total Consumption: 723805

AQI Category
00 - 50 Good
50 - 100 Moderate
100 - 200 Unhealthy
* Bubble size based on average
gas consumption by state

Leaflet | © OpenStreetMap contributors

Air Quality and Natural Gas Correlation Coefficient analysis

- We observed a mild correlation between natural gas consumption and AQI. This suggests other factors are also influencing air quality.

Regional Differences:
- Here we analyze if certain regions exhibit stronger or weaker correlations.
- Our data was sourced from the year 2021, with monthly averages providing a broad overview of how natural gas consumption impacts AQI.
- We assumed that more densely populated and industrially active regions tend to consume more natural gas. (we will take a closer look at this later when looking at seasonality)
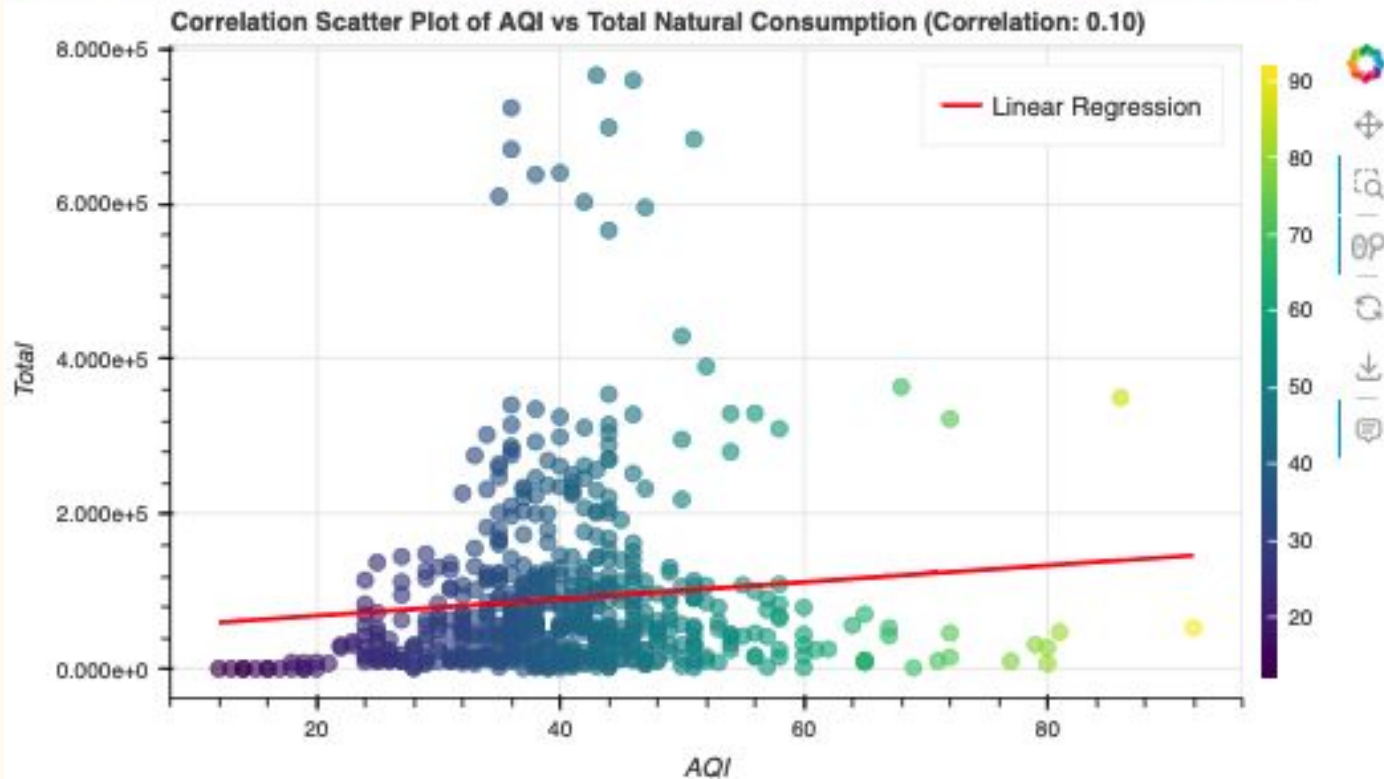
Seasonal Variations:
- As mentioned previously we will explore seasonal implications. However when you hover over each point on the scatter plot you can reference the location and month relating to its AQI vs Natural Gas Consumption.

Impact of Outliers:
- Outliers may indicate specific events that significantly impacted the data.
  - For example, during the COVID-19 pandemic, NO_2 levels dropped by approximately 65%, primarily due to reduced transportation, a major source of NO_2 emissions. However, other sectors, such as residential, commercial, industrial, and electrical power generation, continued to use natural gas, contributing to NO_2 emissions. (NIH)
  - NO_2 is primarily produced during the combustion of fossil fuels, including natural gas, when nitrogen in the air reacts with oxygen at high temperatures. This reaction occurs in engines, power plants, and other combustion processes.
- This highlight a specific areas of focus for future projects to better understand the impact of transportation on the total consumption of natural gas and its impact to Air Quality.

Low Correlation Between Air Quality Index and Natural Gas Consumption

Correlation Scatter Plot of AQI vs Total Natural Consumption (Correlation: 0.10)

Seasonal analysis

Seasonal variations play a significant role in both energy consumption patterns and environmental quality. This analysis aims to explore how these seasonal changes in natural gas consumption influence air quality, for the sake of time in this project we narrowed this part of our analysis for 10 different states that are ranked as top-10 natural gas consumers per U.S. Energy Information Administration (EIA)https://www.eia.gov/state/rankings/#/series/12 .

I categorized our combined(AQI &NG_consumption) data for the months into seasons, then I grouped by state and seasons and averaged the AQI and NG_consuption value. I visualized our findings using Python's Bokeh library and combined multiple plots together and saved it as html file and then render it on our flask app by adding as a hyperlink in our index html file and give the source to our bokeh plots html file .

Impact of Seasonal Natural Gas Consumption on Air Quality

Comparison of Air Quality Index against Natural Gas consumption over different seasons for top 10 States

Seasonal Trends of Air Quality Index

Seasonal Trends of Natural Gas Consumption
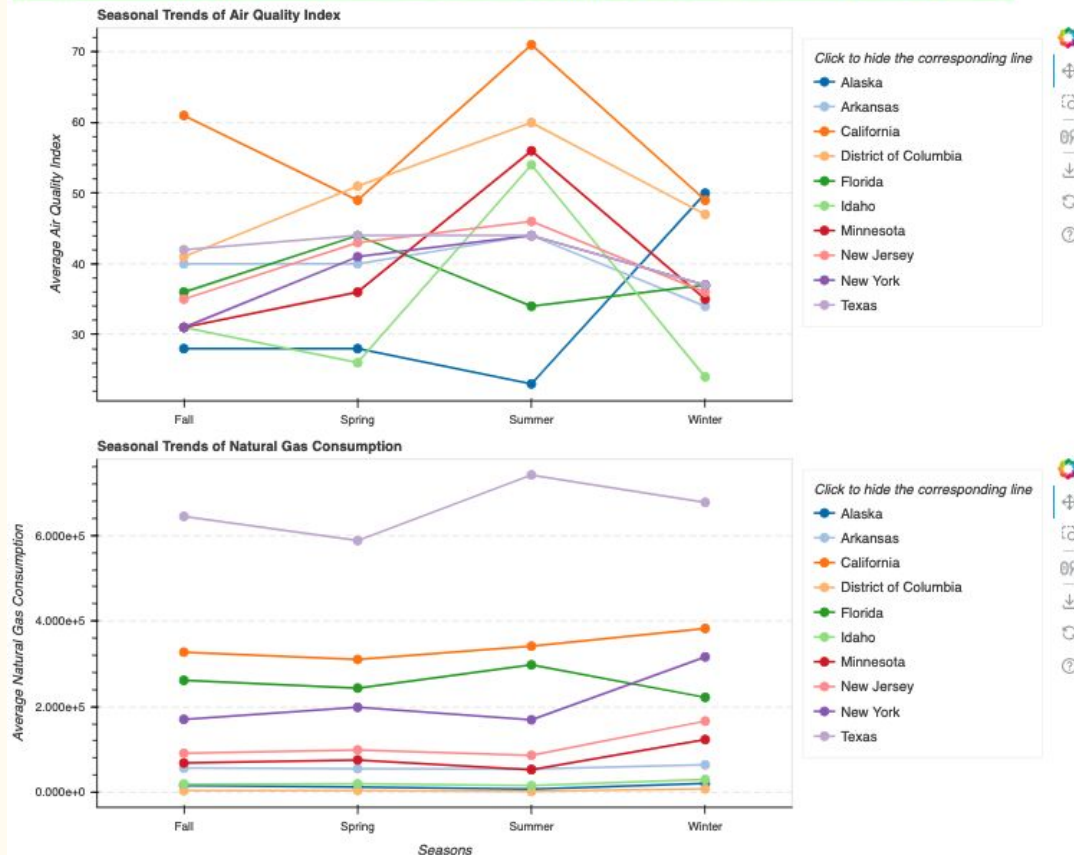
Observations: Although through this analysis natural gas consumption may not be high enough to significantly impact air quality,**causes** could be other environmental conditions may dilute or spread pollutants more effectively. One thing to point out is also due to the aggregated AQI, which might mask local variations(Industrial Zones vs Residentials) where the consumptions have a more noticeable effect on air quality.

Conclusion

We have created a interactive data tool where we analyzed potential correlation and seasonality between Air Quality and Natural Gas Consumption:

Takeaway/Next Steps:
- Due to the observed correlation, our analysis could potentially serve as a valuable tool for policymakers in regulating energy use and emissions to enhance public health..
- While correlation does not imply causation, the subtle correlation we noticed suggests underlying factors worth investigating:
  - Taking a closer look at the outliers from our correlation scatter plot.
  - Our seasonality analysis suggests that conducting smaller-scale studies, such as city-level assessments using daily AQI values, might uncover impacts that remain hidden in broader datasets.
  - Expand out data variables against AQI:by examining other potential factors that may negatively impact air quality.
  - Some factors to analyze alongside Air Quality across the US overtime:
    - Vehicle Emissions: diesel / gasoline
    - Wildfire
    - Manufacturing
    - Agriculture
    - Construction/Demolition
    - Natural Sources: Volcanoes, Dust Storm
    - Chemical Reactions in the Atmosphere
    - Transboundary Pollution