DS 2500 Final Project

Analyzing the Economic Impact of Clean Energy

Presented by Hang Hang, Harshini Dinesh, Ryan Jiang, Hersh Joshi

Table of Contents

01 Problem Statement and Background

Problem Statement and Background

Problem Statement

Analyze the economic impact of the renewable energy across U.S. states

Motivation

- Understand how renewable energy is related with various economic metrics like GDP, energy consumption, and EV adoption rates
- Impact on policy-making and investment strategies

02 Introducing and Explaining the Data

Introducing and Explaining the Data

1. Net Generation for All Sectors by State

- Energy production data by state.
- Source: U.S. Energy Information Administration (EIA), 2000-2023.
- Key Info: Includes renewable and non-renewable energy sources for all 50 states and D.C.

2. Costal and Inland Status by State

- Classification of states as coastal or inland.
- Source: Office For Coastal Management
- Manually compiled the Excel file from a provided list of coastal states (states that have coastal residents), as most sources were in map format.

3. US State Shapefile

- Simplified state boundaries for thematic mapping.
- Source: U.S. Census Bureau
- Used to visualize energy generation percentages across U.S. states.

4. Levelized Cost of Charging Electric Vehicles (LCOC)

- LCOC and lifetime fuel cost savings for battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs) at state & national levels.
- Source: National Renewable Energy Laboratory

Introducing and Explaining the Data

5. Production by State 1960-2022

- Renewable energy production data by state from 1960-2022.
- Source: U.S. Energy Information Administration (EIA),

6. Consumption by State 1960-2022

- Renewable energy consumption data by state from 1960-2022.
- Source: U.S. Energy Information Administration (EIA),

7. GDP by State 1997-2022

- Real GDP (in millions) per state from 1997-2022
- Source: U.S. Energy Information Administration (EIA),

8. Consumption by Different Sectors Over Time 1949 - 2023

- Various energy sources (in Trillions Btu) for different sectors from 1949 - 2023
- Source: U.S. Department of Energy
  - Energy Consumed by Electric Power Sector
  - Energy Consumed by Residential/Commercial Sectors
  - Energy Consumed by Industrial Sector

Introducing and Explaining the Data

9. Alternative Fuel Stations by State in 2021

- Alternative fuel stations data by state in 2021
- Source: U.S. Department of Energy

10. Gas Consumed in Various states by Hybrid/Clean Energy Vehicles in the last 15 years

- **10. Gas Consumed in Various states by Hybrid/Clean Energy Vehicles in the last 15 years**
- Source: EIA.gov

11. Gas Consumed in Various states by Disel Vehicles in the last 15 years

- **10. Gas Consumed in Various states by Diseel Energy Vehicles in the last 15 years**
- Source: EIA.gov

12. Trends with Different modes of Transportation

- **Shows various modes of transportation and associated emmisions**
- Source: data.gov

03 Data Science Approaches & Insights

Renewable Energy Consumption and Production over Time

- Which states have had the greatest changes in renewable energy consumption since 1970?
- Which states have had the greatest changes in renewable energy production since 1970?

Sample Code

```
def calculate_change(df, present_year, past_year):
    """
    Given a dataframe, the most recent year and a past year, this function
    calculates the total change in energy production (or consumption)
    for each state during the given time period.
    """
    # Convert year to string
    present_year = str(present_year)
    past_year = str(past_year)

    # Remove commas and convert the columns to numeric
    df[present_year] = pd.to_numeric(df[present_year].str.replace(",", ""))
    df[past_year] = pd.to_numeric(df[past_year].str.replace(",", ""))

    # Calculate the change
    df["Change"] = df[present_year] - df[past_year]

    # Return the relevant columns
    return df[["State", "Change"]]
```

```
prod_change_df = calculate_change(prod_df, 2022, 1970)
top_5_states_prod = prod_change_df.nlargest(5, 'Change')
print("Top 5 states with greatest changes in production:")
print(top_5_states_prod)
```

Results

Top 5 States by Changes in Consumption:

1. Texas
2. California
3. Iowa
4. Florida
5. Georgia

```
Top 5 states with greatest changes in consumption:
   State  Change
43    TX  687122
4     CA  633415
12    IA  408455
9     FL  251724
10    GA  206390
```

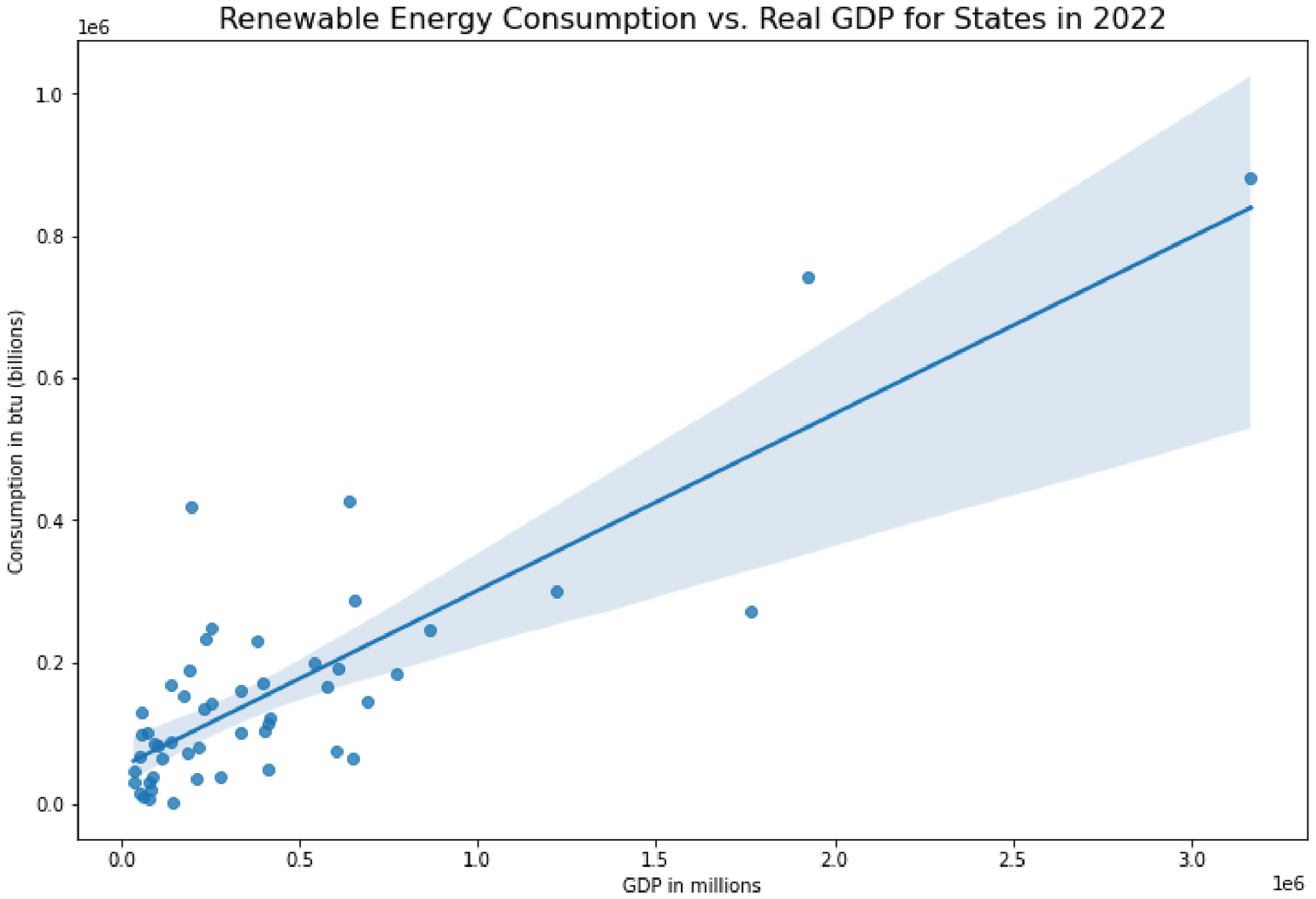
Top 5 States by Changes in Production:

1. Iowa
2. Texas
3. Nebraska
4. California
5. Illinois

```
Top 5 states with greatest changes in production:
   State  Change
12    IA  761528
43    TX  599284
29    NE  321086
4     CA  294666
14    IL  293807
```

Energy Consumption/Production vs. GDP

- Correlation between renewable energy consumption and Real GDP (in 2022): **r = 0.8333**
- Correlation between renewable energy production and Real GDP (in 2022): **r = 0.5530**



Observations:

- Both production and consumption have a strong positive correlations with GDP, but there is a slightly stronger relationship between GDP and consumption than GDP and production.

Renewable Energy Consumption and Production vs. GDP

- Which states have the most energy consumption and production compared to their GDP?

Results

Top 5 States by Consumption:

State
1. Washington D.C
2. Delaware
3. New Jersey
4. Maryland
5. Massachusetts

Top 5 States By Production:

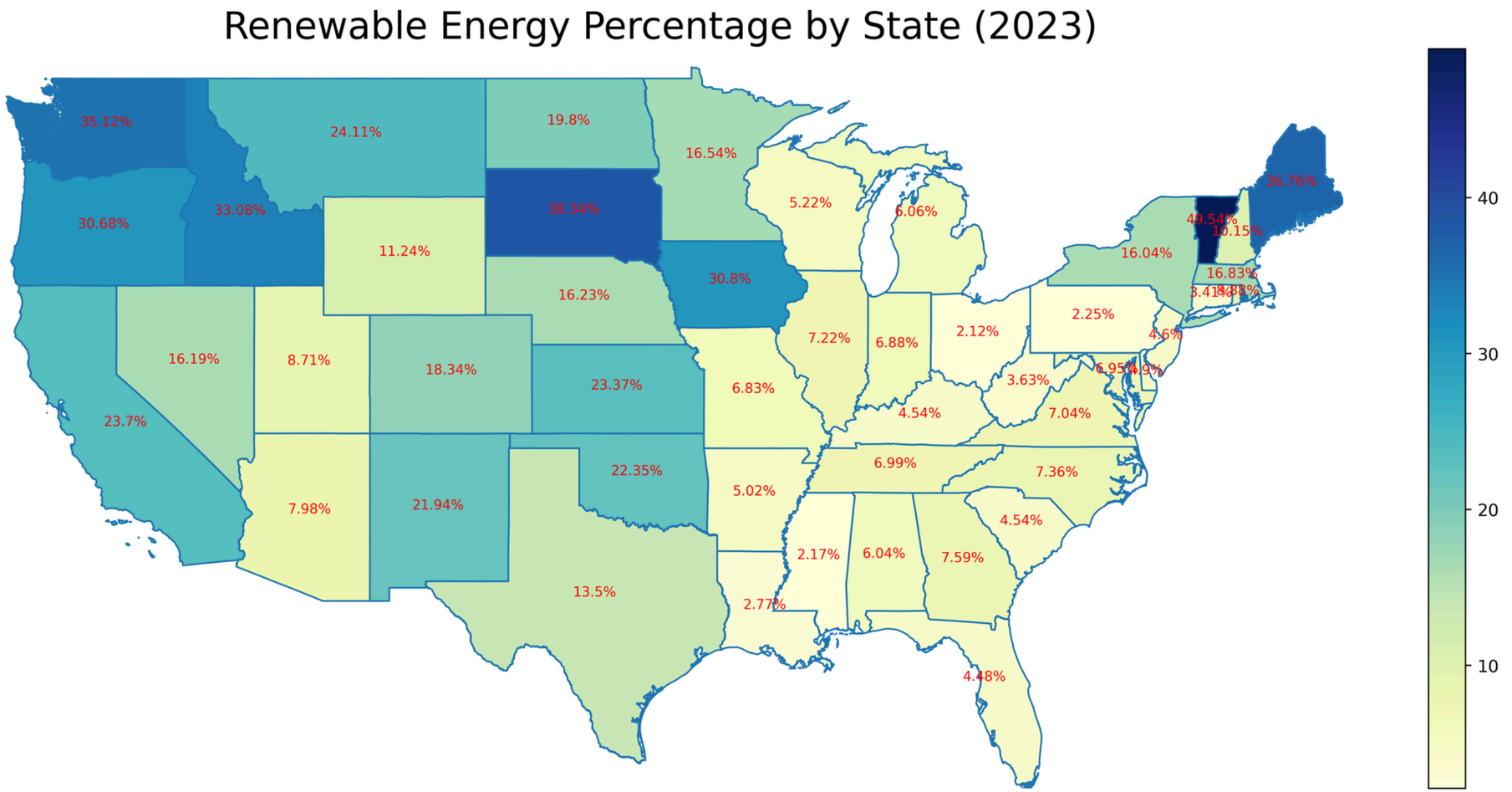
State
1. Washington D.C
2. Delaware
3. New Jersey
4. Maryland
5. Massachusetts

```
Top 5 states by renewable energy consumption (divided by GDP) in 2022:
[('DC', 55.48135593220339), ('DE', 10.437739516800889), ('NJ', 9.950018462106527),
('MD', 8.588684040580796), ('MA', 8.052412295311314)]
Top 5 states by renewable energy production (divided by GDP) in 2022:
[('DC', 84.6734861845973), ('DE', 26.572145634499826), ('NJ', 19.445302023512433),
('MD', 15.026518934285818), ('MA', 11.569976069685076)]
```

Exploring Renewable Energy Generation by State...

- Which states lead in renewable energy generation?
- Do coastal states outperform inland states in renewable energy production?

Geospatial Distribution

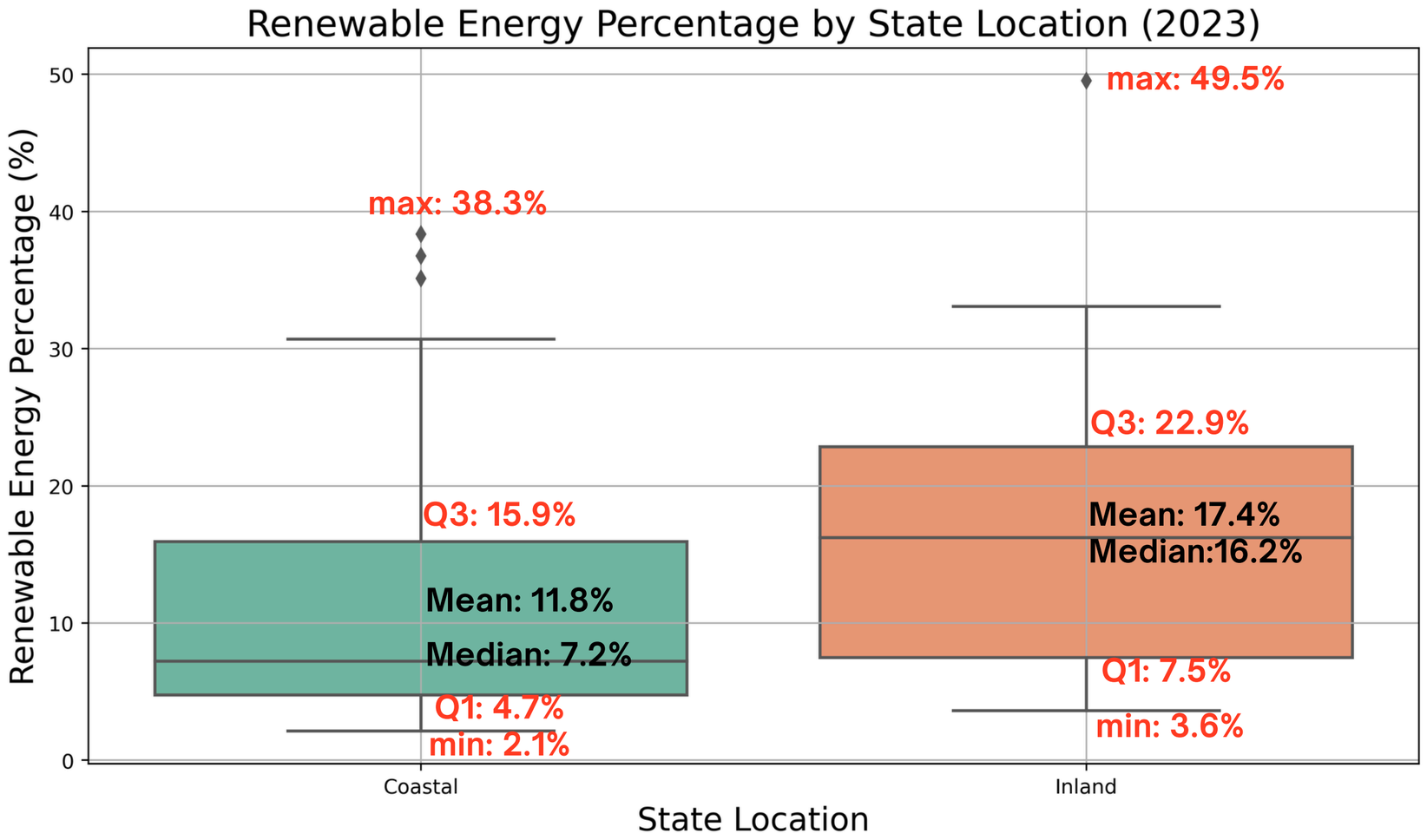


Top 5
Vermont (Inland): 49.54%
South Dakota (Coastal): 38.34%
Maine (Coastal): 36.76%
Washington (Coastal): 35.12%
Idaho (Inland): 33.08%

- **Percentage calculation: extracted renewable generation percentage by state using net generation data**
- **Shapefile Integration: Merged U.S. state boundaries with renewable energy percentage data, focusing on the contiguous U.S.**
- **Visualization:** use GeoPandas and Matplotlib to show renewable energy generation across states

Correlation & Box Plot

- Correlation between state location & renewable energy percentage: **r = -0.2434**
- Due to the weak correlation, a box plot is used to better illustrate the relationship



Observations:

- **Median & Mean**: Inland > Coastal
- **Min & Max**: both are similar but Inland has slightly higher values
- **IQR**: Inland displays greater diversity
  - Influences: varying inland climates compared to coastal areas, diversity in natural resources (Vermont's rivers and forests)

Diving into the cost of charging EV data...

- How does the levelized cost of charging EVs vary by state?
  - Which states incur the highest and lowest costs?

- Are there regional patterns or clusters in the cost of charging EVs?
  - Which regions have the most favorable costs?

Residental vs. Workplace Charging

- How does the levelized cost of charging battery electric vehicles (BEVs) in residential and workplace setting vary across different states?
  - Which states have the highest and lowest costs?

Residential Charging	Workplace Charging
◦ Mean: $0.146/kWh ◦ Median: $0.141/kWh ◦ Range: $0.085 - $0.303/kWh ◦ Maine (Lowest: $0.085/kWh) & Hawaii (Highest: $0.303/kWh)	◦ Mean: $0.150/kWh ◦ Median: $0.141/kWh ◦ Range: $0.119 - $0.307/kWh ◦ Nevada (Lowest: $0.119/kWh) & Hawaii (Highest: $0.307/kWh)

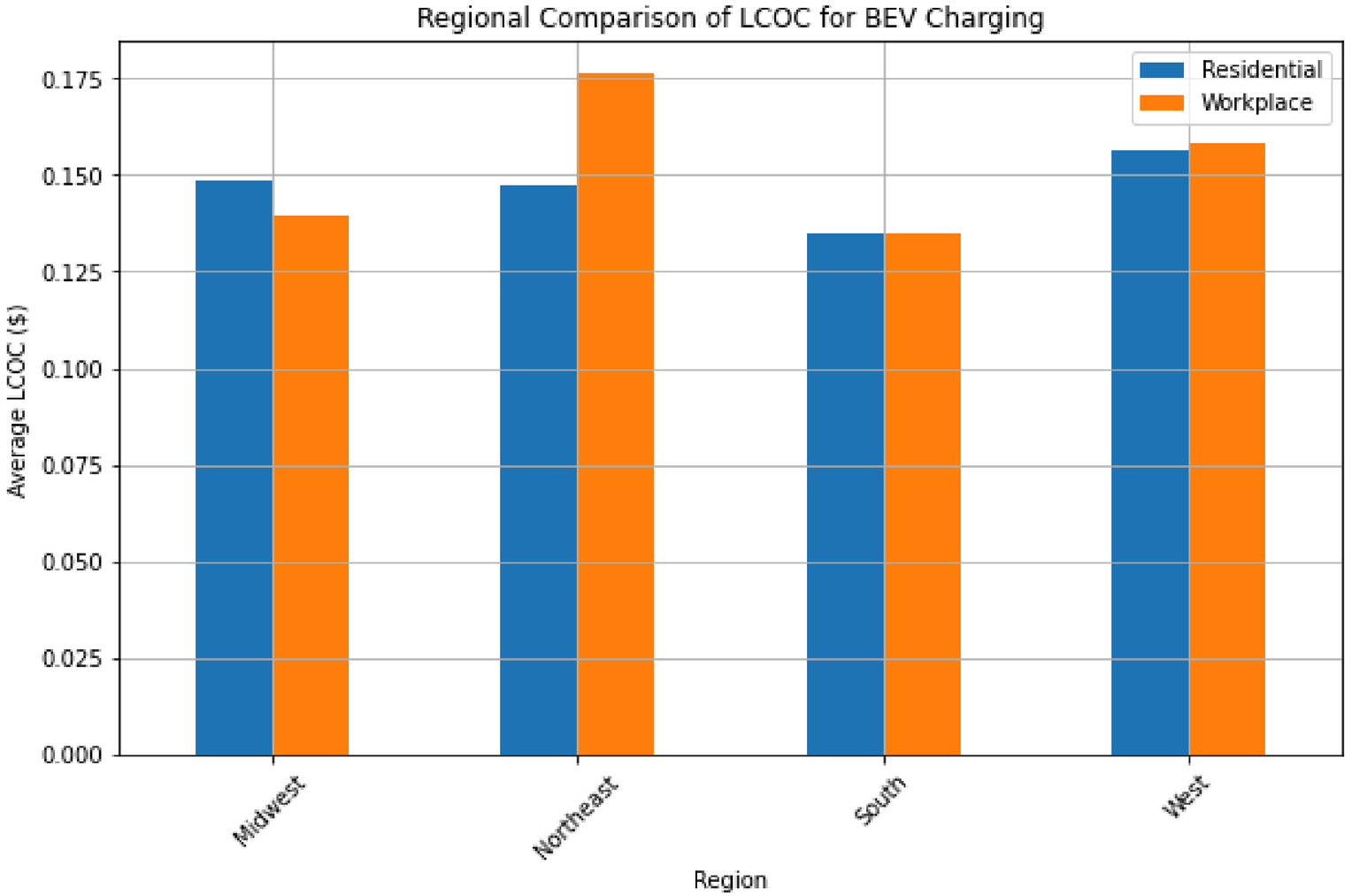
Observations:

- **Hawaii** has the **highest costs in both residential and workplace** settings, likely due to high electricity costs, limited infrastructure, and reliance on imported energy.
- **Workplace charging cost is slightly higher**, often due to higher infrastructure and operational costs (i.e. installing Level 2 or DC fast charging stations)
  - Source: https://afdc.energy.gov/fuels/electricity-charging-workplace

Regional Cost Analysis for BEV Charging

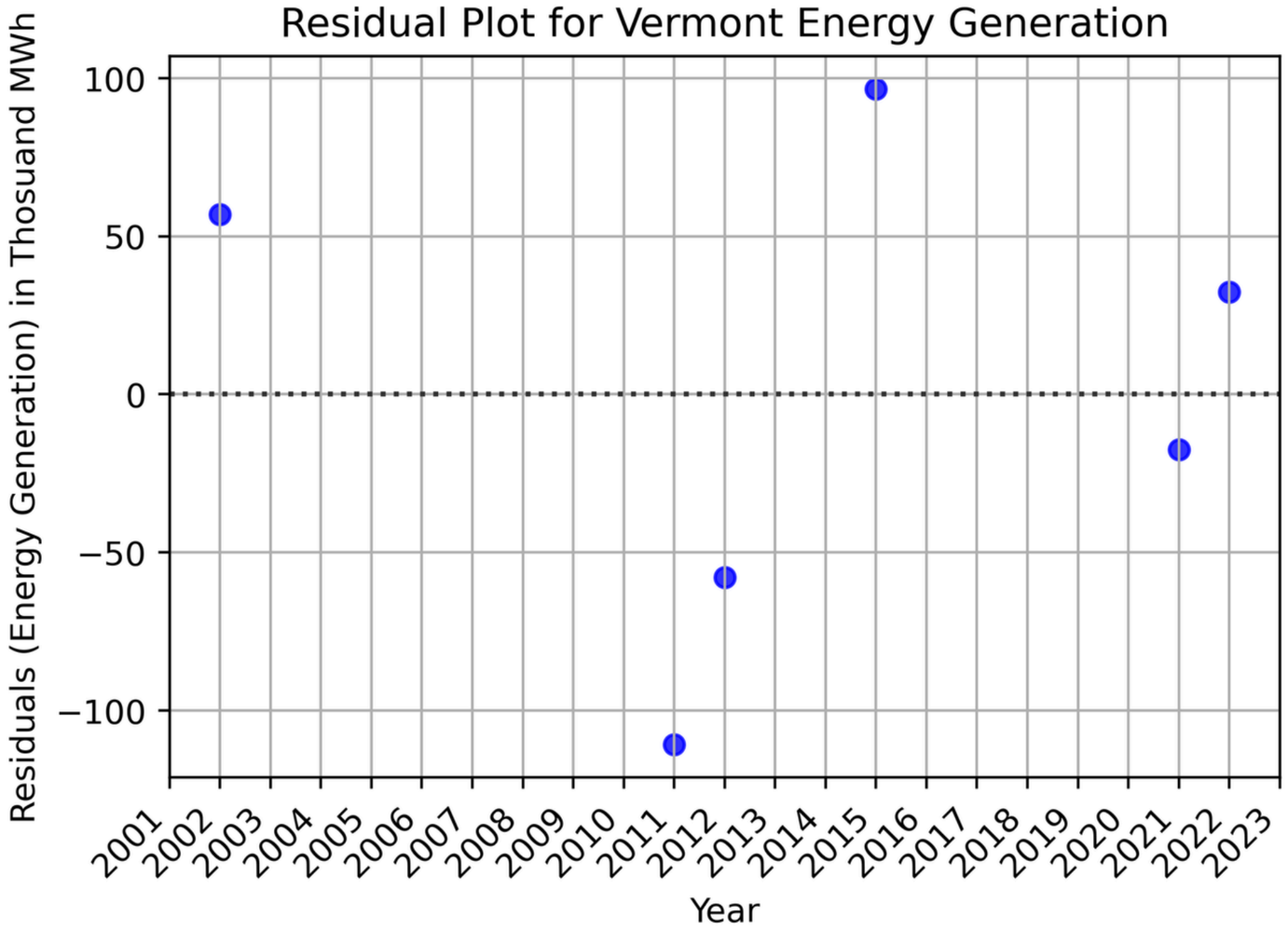
- Are there regional patterns in the cost of charging BEVs?
  - Which regions have the most favorable costs?

- Observations:
- **Northeast & West region**: the LCOC is the highest, especially for the workplace
  - a potentially higher premium for workplace infrastructure
- Midwest & South: offer more consistent and favorable costs across both settings



Linear Regression Analysis On Vermont's Wind Energy

- Vermont's wind energy has seen an rising trend over the years
- X (Independent variable): Year (ranging from 2001 to 2013)
- Y (Dependent variable): Wind energy generation in Vermont (measured in thousand megawatt-hours)
  - RMSE for Vermont wind energy: 72.36

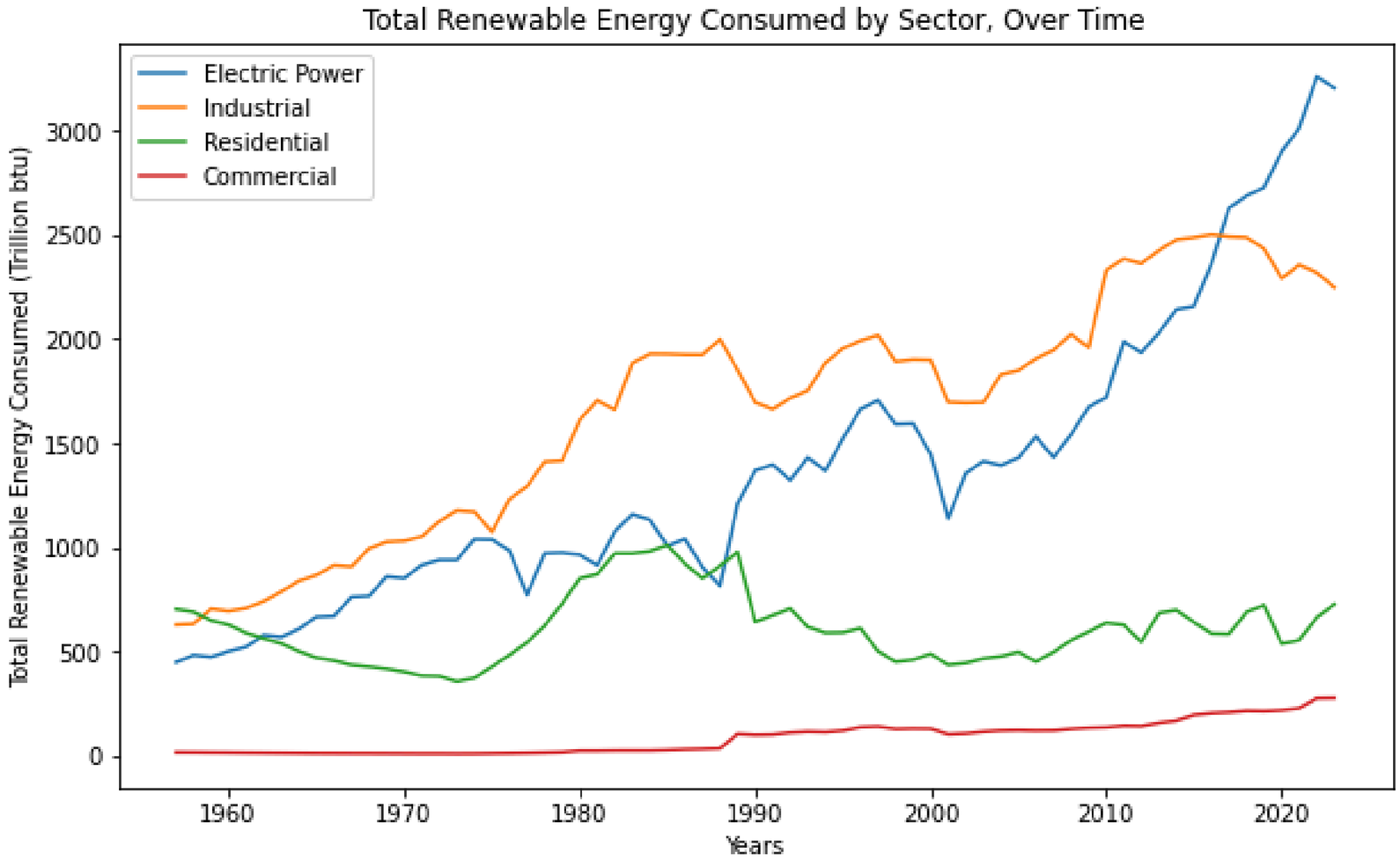


Renewable Energy Consumption Analysis by Sector

- Given the growth of the EV industry, has the electric power sector consumed more renewable energy than the industrial, commercial, and residential sectors?

Observations:

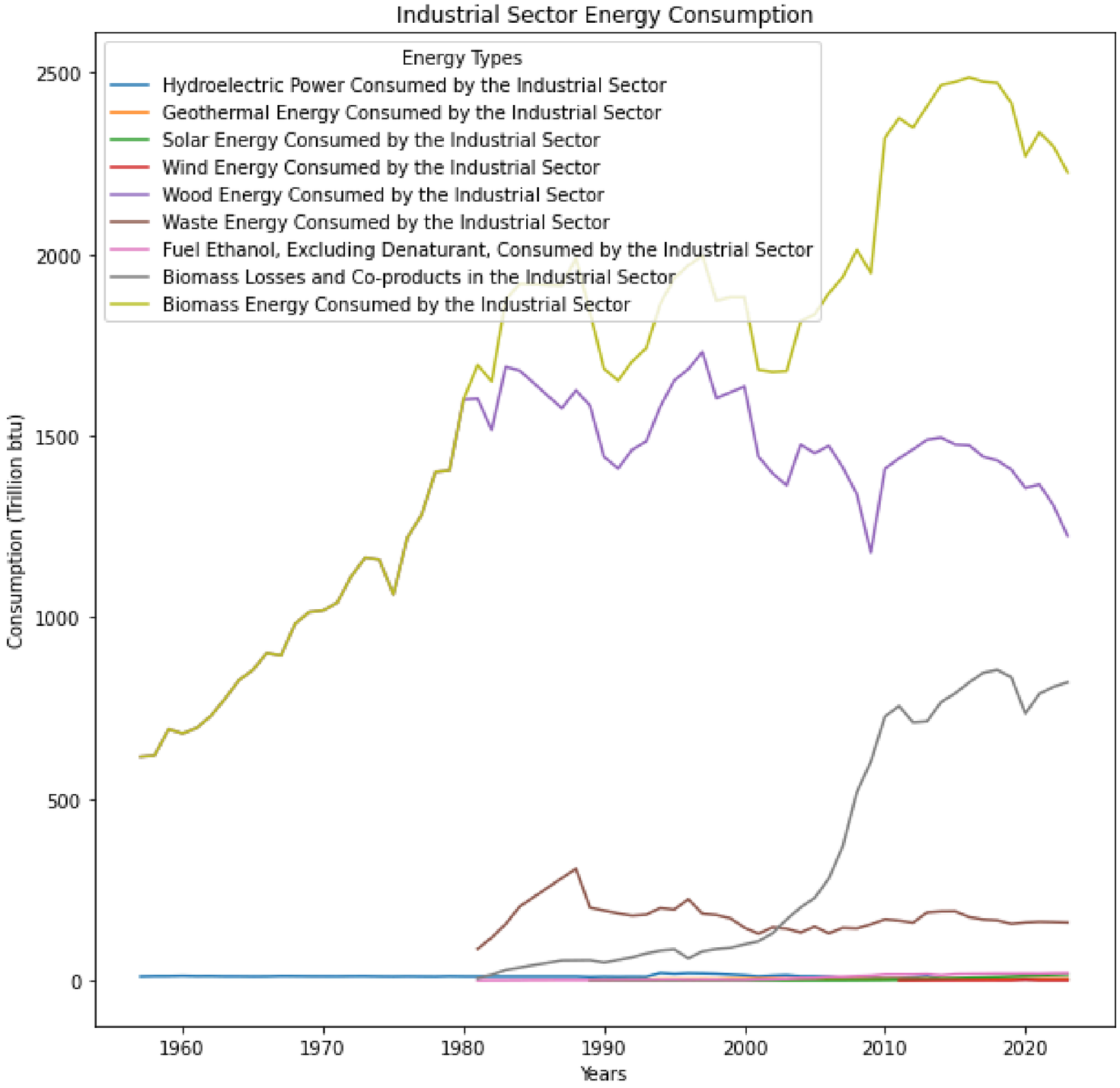
- **Electric Power Sector:** has steadily increased its total renewable energy consumption from 1949 to 2023
- **Holistically:** industrial sector has consumed most renewable energy over the years
  - Electric power sector: 91,597.39 Trillion btu
  - **Industrial sector: 111,415.737 Trillion btu**
  - Residential sector: 40,630.32 Trillion btu
  - Commercial sector: 5,643.39 Trillion btu



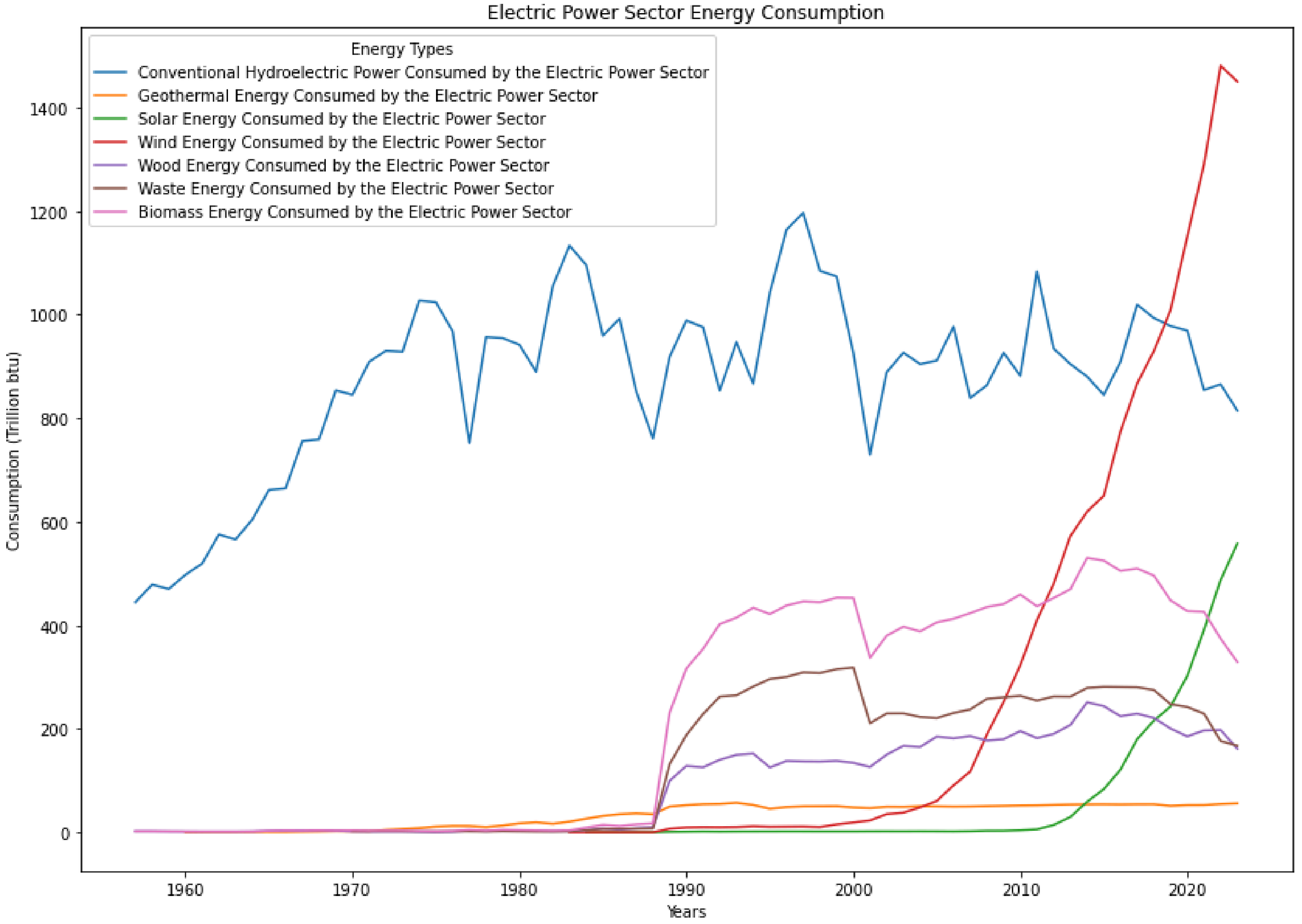
Diving into the most consumed energy source by sector...

- How has energy consumption changed for different energy sources over the years (1949 to present) in each sector?

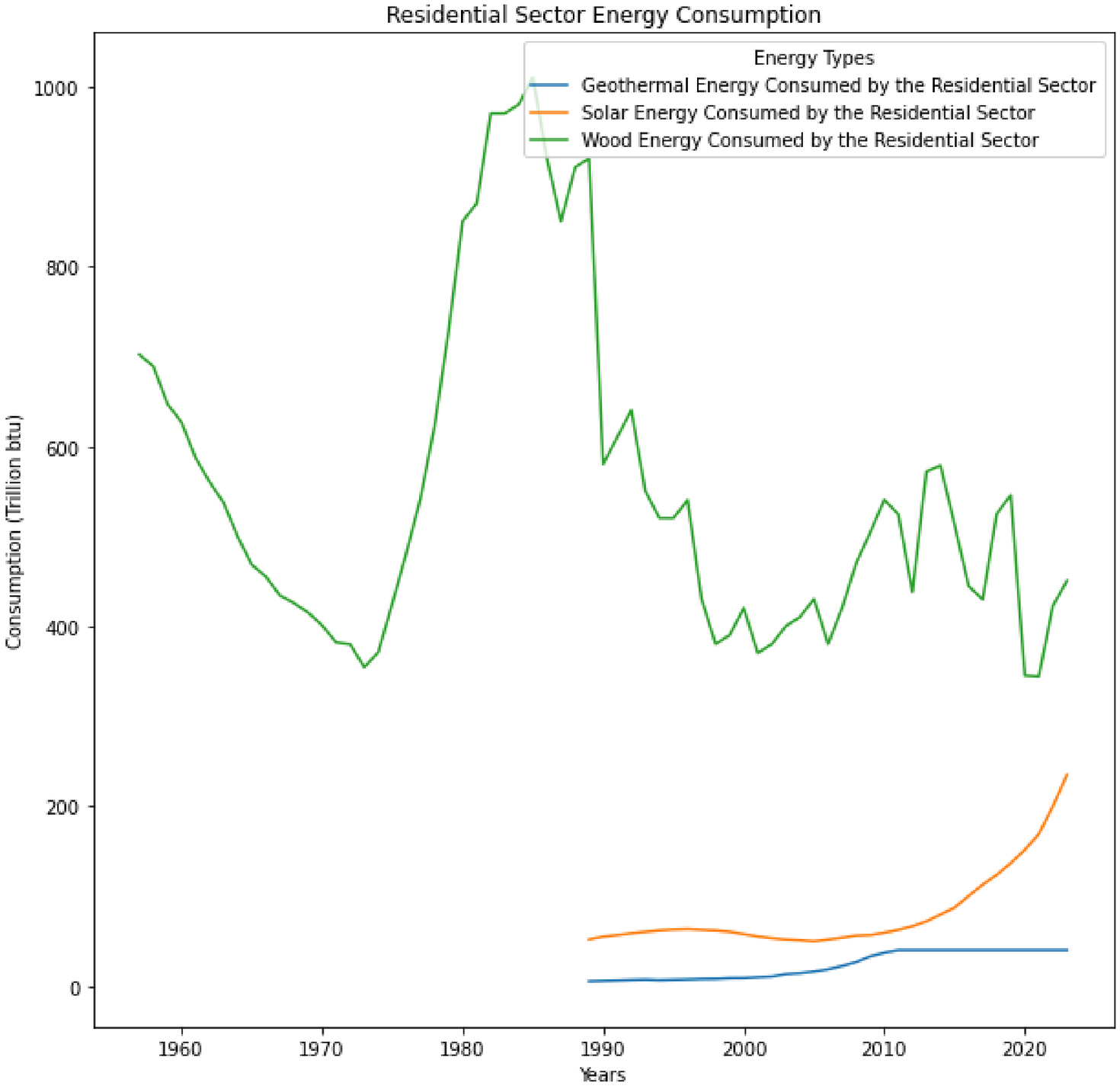
- Are there patterns?



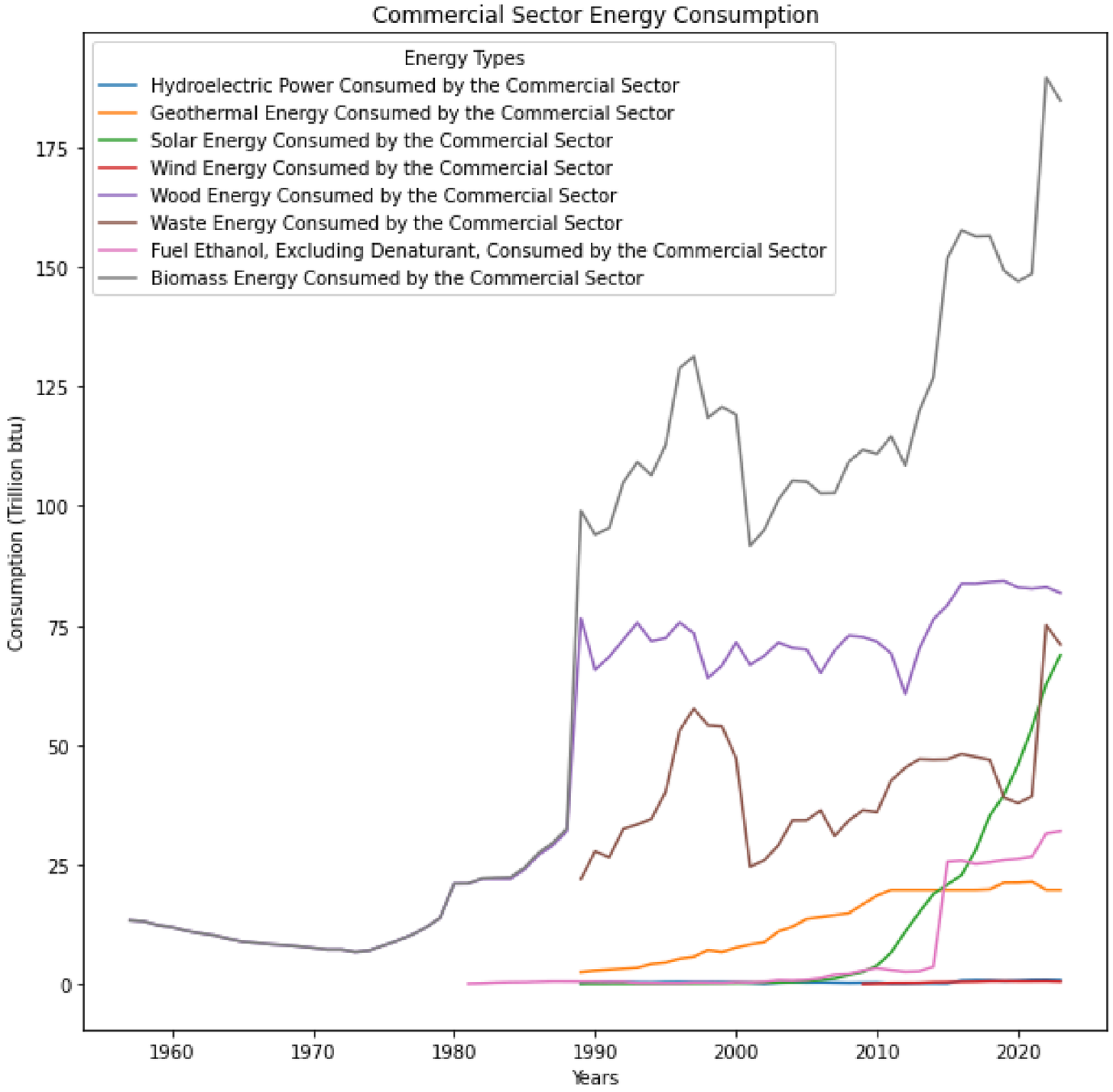
Biomass Energy: 110,480.94 Trillion btu



Conventional Hydroelectric Power: 58,757.67 Trillion Btu



Wood Energy: 36,997.09 Trillion btu

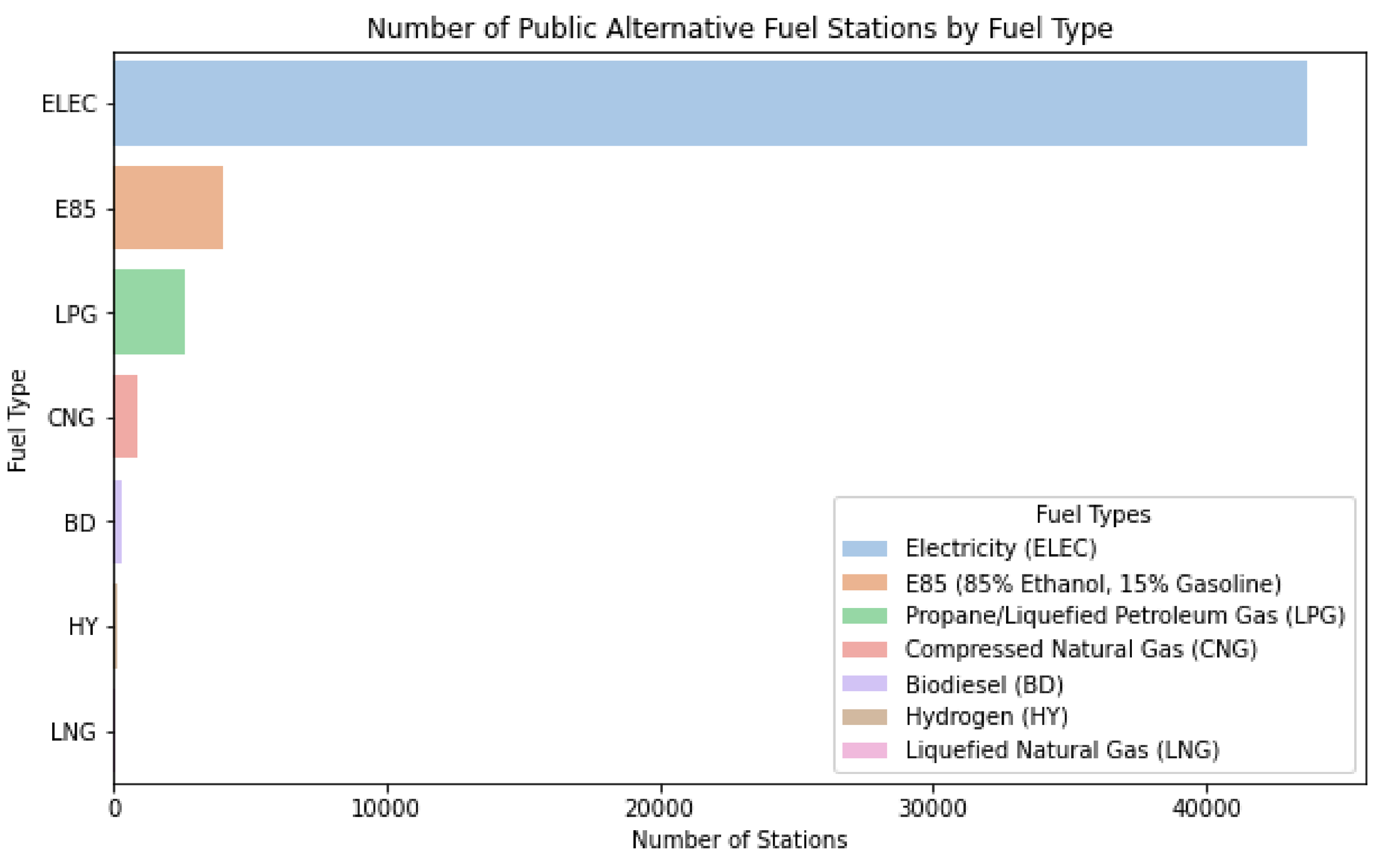


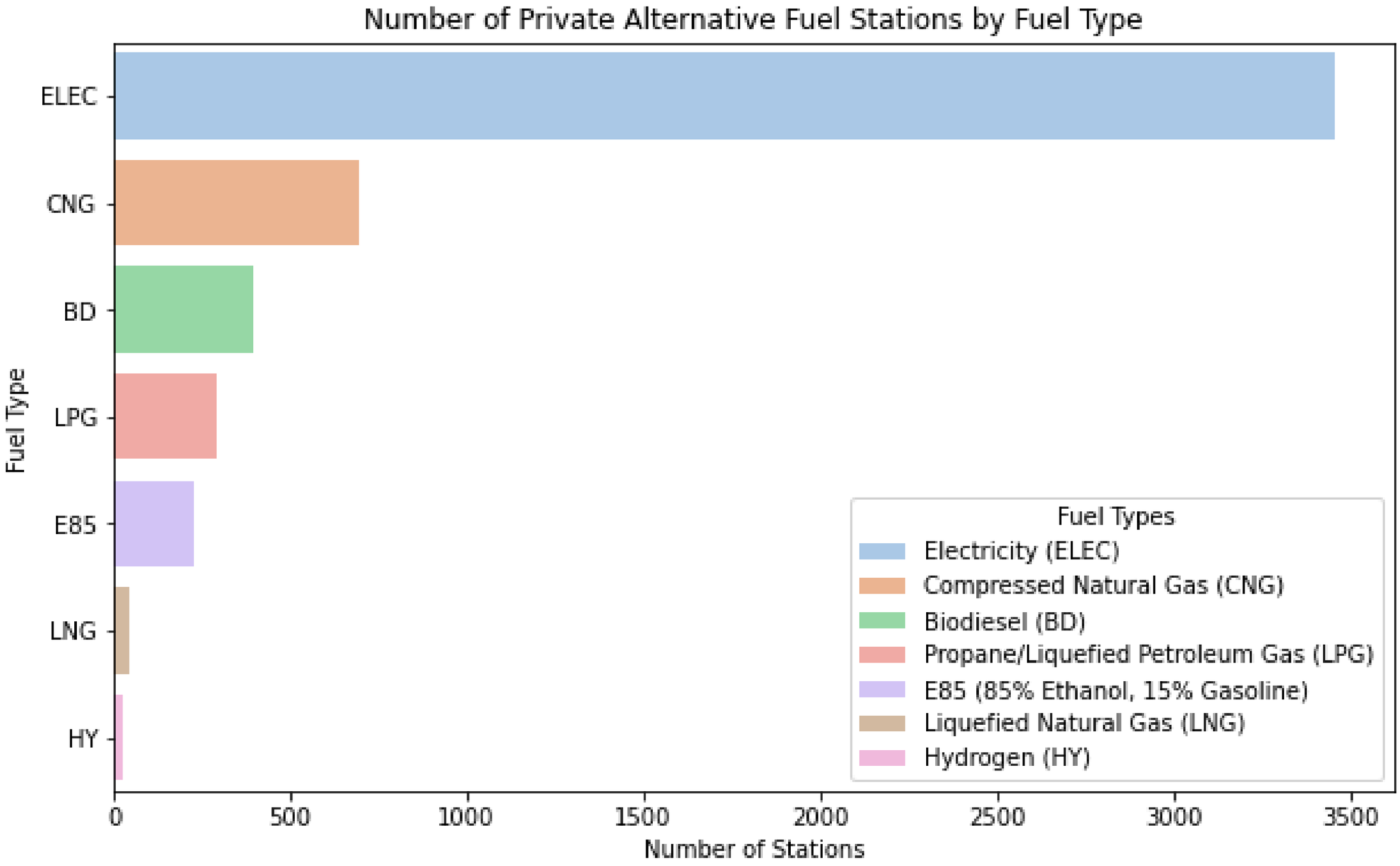
Biomass Energy: 4,732.62 Trillion btu

Diving into the accessibility of Alternative Fuel Stations (AFS)...

- Are there more alternative fuel stations available to the public for electric cars than for stations providing other fuels?

- Do private stations provide more electricity than other fuels?


Number of Public Alternative Fuel Stations by Fuel Type
ELEC
E85
LPG
CNG
BD
HY
LNG
Fuel Type
0
10000
20000
30000
40000
Number of Stations
Fuel Types
Electricity (ELEC)
E85 (85% Ethanol, 15% Gasoline)
Propane/Liquefied Petroleum Gas (LPG)
Compressed Natural Gas (CNG)
Biodiesel (BD)
Hydrogen (HY)
Liquefied Natural Gas (LNG)



There are more public alternative fuel stations (43,632) which provide electricity than private alternative fuel stations (3,454).

Now lets look at correlations with climate of different regions

- Do states with more landmasses and extreme temperatures consume more energy?

```
def corr_fuel_vehicles_gas_consumed(data_fuel_vehicles):
    #using the mean finish time for women for all the years to get a correlation value
    #initialize empty list
    total_gas = emissions_vehicles(data_fuel_vehicles)
    #correlation across all years, are the values increasing or decreasing
    #partcipation criteria is also changing, with respectice to soemthing how
    #is another variable changing
    yr_lst = []
    for year in vehicle_types:
        yr_lst.append(int(year))
    gas_consumed = statistics.correlation(yr_lst, total_gas)
    return gas_consumed

def corr_fuel_vehicles_electirc_consumed(data_fuel_vehicles):
    #using the mean finish time for women for all the years to get a correlation value
    #initialize empty list
    total_gas = emissions_vehicles(data_fuel_vehicles)
    #correlation across all years, are the values increasing or decreasing
    #partcipation criteria is also changing, with respectice to soemthing how
    #is another variable changing
    yr_lst = []
    for year in vehicle_types:
        yr_lst.append(int(year))
    gas_consumed = statistics.correlation(yr_lst, total_gas)
    return gas_consumed
```

Categorizing Into Different Lists

```
    Parameters:
    directory (str): The directory where the data files are located.

    Returns:
    dict: A dictionary with data organized by weight category and year.
    """
    # Load the datasets from the specified directory
    # electricity_data = pd.read_csv(os.path.join(directory, 'Electricity_data.csv'))
    # fuel_data = pd.read_csv(os.path.join(directory, 'Fuel_data.csv'))

    # Combine both datasets for consistent processing
    combined_data = pd.concat([electricity_data, fuel_data])

    # Define weight class categories
    weight_class_mapping = {
        'Light Duty': 'Light',
        'Medium Duty': 'Medium',
        'Heavy Duty': 'Heavy'
    }

    # Apply the weight class mapping
    combined_data['Weight Category'] = combined_data['Weight Class'].map(lambda wc: weight_class_mapping.get(wc, 'Other'))

    # Create a dictionary to store data by weight category and then by year
    data_by_weight_and_year = {}

    # Group the data by Weight Category and Year
    grouped_data = combined_data.groupby(['Weight Category', 'Year'])

    # Store the data in the dictionary
    for (weight_category, year), group in grouped_data:
        if weight_category not in data_by_weight_and_year:
            data_by_weight_and_year[weight_category] = {}
        data_by_weight_and_year[weight_category][year] = group
```

04 Future Work

Next steps

Deep Dive into State-Level Incentives	• Analyze the effectiveness of specific state incentives, like MassEVIP, in reducing LCOC and their impact on EV adoption rates
Infrastructure Expansion Analysis	• Explore how expanding advanced charging infrastructure (e.g., Level 2, DC fast charging) influences both residential and workplace charging costs
Cost-Benefit Assessment	• Conduct a detailed cost-benefit analysis of EV charging infrastructure investments, considering long-term savings for the charging station users
Link to EV Adoption	• Look into the relationships between EV incentives and EV adoption rate
Forecast Energy Consumption & AFS accessibility	• Forecast future energy consumption levels by sector and future AFS accessibility using correlation and regression analysis • Look into how consumption of each energy source and how increases in AFS impact climate change

Need further data to address these questions:

- Do states with stations with longer access times have less alternative fuel stations (electric)?
  - Which states need more alternative fuel stations to be built?
- Does climate cause more people to switch to different vehicle types or is it just a correlation?
- How could different environments affect the modes of transportation sought after?

Questions?