A Comprehensive Analysis of U.S. Electricity Consumption and Generation Patterns

(2003-2023)

Professor: Stefanie Molin

Wei Lo

Analytics Programming (DAV 5400)

Abstract

This research presents a comprehensive analysis of U.S. electricity consumption and generation patterns from 2003 to 2023. By investigating stable consumption trends and decoding seasonal variations, the study sheds light on the impact of environmental and economic factors on electricity demand. The analysis of diverse energy sources reveals a notable shift towards renewables, especially solar and wind, underscoring the urgent need for reduced fossil fuel dependency. Through intricate correlation and causation studies, considering economic indicators, policy changes, and technological advancements, the research provides a holistic perspective on the forces shaping the energy landscape. Additionally, the study explores seasonal consumption behaviors, offering valuable insights for policymakers and energy providers to navigate the complexities of sustainable energy planning. This research contributes significantly to the ongoing discourse on sustainable energy transition, guiding the way toward a greener and more resilient energy future.

Keywords: *Electricity Consumption, Renewable Energy, Sustainable Energy Transition*

Introduction

The dataset utilized in this analysis is a carefully curated compilation, rearranged, and combined from two distinct datasets obtained from the U.S. Energy Information Administration (EIA). The resulting dataset consists of seven columns: `YYYYMM`, representing the year and month; `1_coal`, `2_petroleum`, and `3_Natural_Gas` indicating electricity generation values in gigawatt-hours (GWh) for coal, petroleum, and natural gas sources, respectively; `7_consumption` denoting the total electricity consumption in GWh; and `11_solar` and `12_wind` representing the generation values for solar and wind energy sources, respectively.

This meticulous combination of datasets provides a comprehensive overview of the U.S. electricity landscape, enabling in-depth analyses of consumption trends, seasonal patterns, and the dynamic relationship between traditional fossil fuel-based sources and renewable energies like solar and wind. The dataset's extensive time frame from 2003 to 2023 empowers researchers to explore long-term patterns, offering valuable insights into the evolution of the U.S. energy sector over the course of two pivotal decades.

<u>Data Summary</u>

The data used for this analysis was acquired from the U.S. Energy Information Administration (EIA) website (https://www.eia.gov/electricity/data.php). The dataset was then assembled, rearranged, and combined from two separate datasets to create a comprehensive dataset for analysis.

<u>Summary:</u>

Source: U.S. Energy Information Administration (EIA)

Attributes:

Numeric Variables:

1_coal: Represents electricity generation from coal in gigawatt-hours (GWh).

2_petroleum: Represents electricity generation from petroleum in GWh.

3_Natural_Gas: Represents electricity generation from natural gas in GWh.

7_consumption: Represents total electricity consumption in GWh.

11_solar: Represents electricity generation from solar energy in GWh.

12_wind: Represents electricity generation from wind energy in GWh.

Categorical Variable:

YYYYMM: Represents the year and month in numeric format. While it is represented numerically, it is categorical in nature because it encodes specific categories (years and months) and does not have a continuous numerical meaning.

part_1_summary.py > ...

```python
import pandas as pd

##1_Remove unnecessary columns from 'electricity_con.csv' for further analysis

# Read the original CSV file into a pandas DataFrame
df = pd.read_csv('electricity_con.csv')

# Filter data based on conditions (YYYYMM from 200301 to 202306, Column_Order is "7", and
filtered_df = df[(df['YYYYMM'] >= 200301) & (df['YYYYMM'] <= 202306) & (df['Column_Order']

# Remove specified columns
columns_to_drop = ['MSN', 'Description', 'Unit']
filtered_df.drop(columns=columns_to_drop, inplace=True)

# Save the filtered and modified DataFrame to a new CSV file
filtered_df.to_csv('filtered_electricity_con.csv', index=False)

print("Filtered data saved to filtered_electricity_con.csv")

```

part_1_summary.py > ...

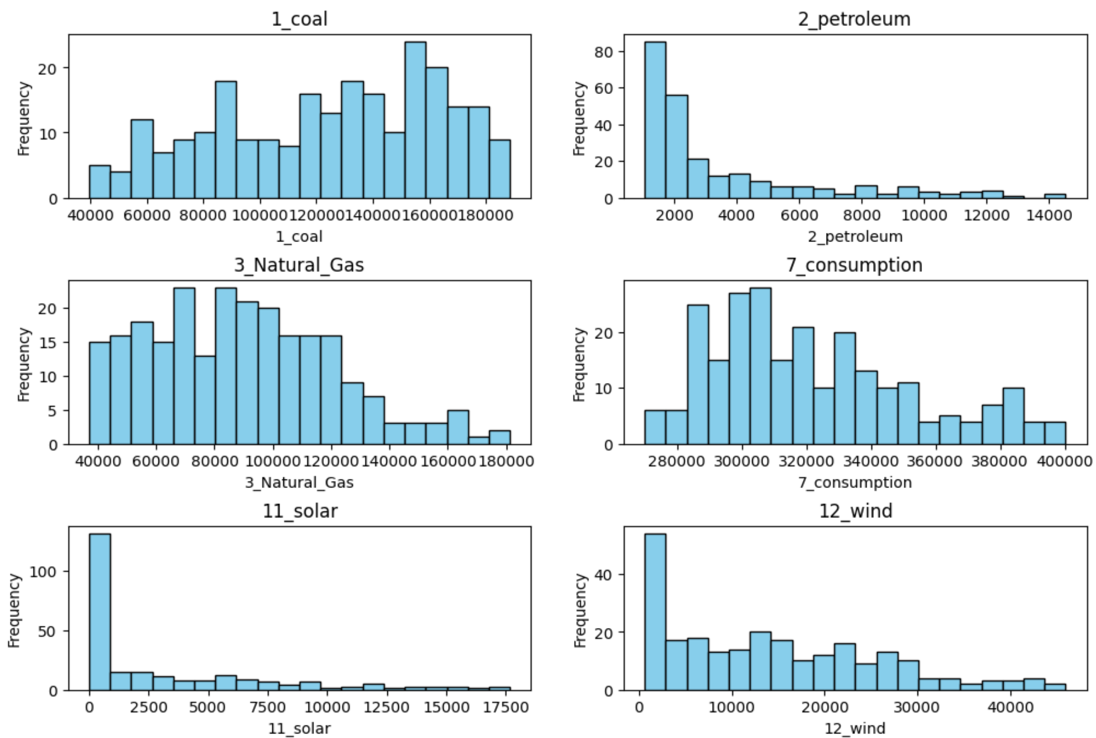
```python

# Rename the columns
column_mapping = {
    1: '1_coal',
    2: '2_petroleum',
    3: '3_Natural_Gas',
    11: '11_solar',
    12: '12_wind',
    7: '7_consumption'
}

pivoted_df.columns = ['YYYYMM'] + [column_mapping.get(col, col) for col in pivoted_df.colu



# Save the modified DataFrame back to the same CSV file, overwriting the original data
pivoted_df.to_csv('merged_electricity_data.csv', index=False)

print("Data successfully modified and saved to merged_electricity_data.csv.")

```

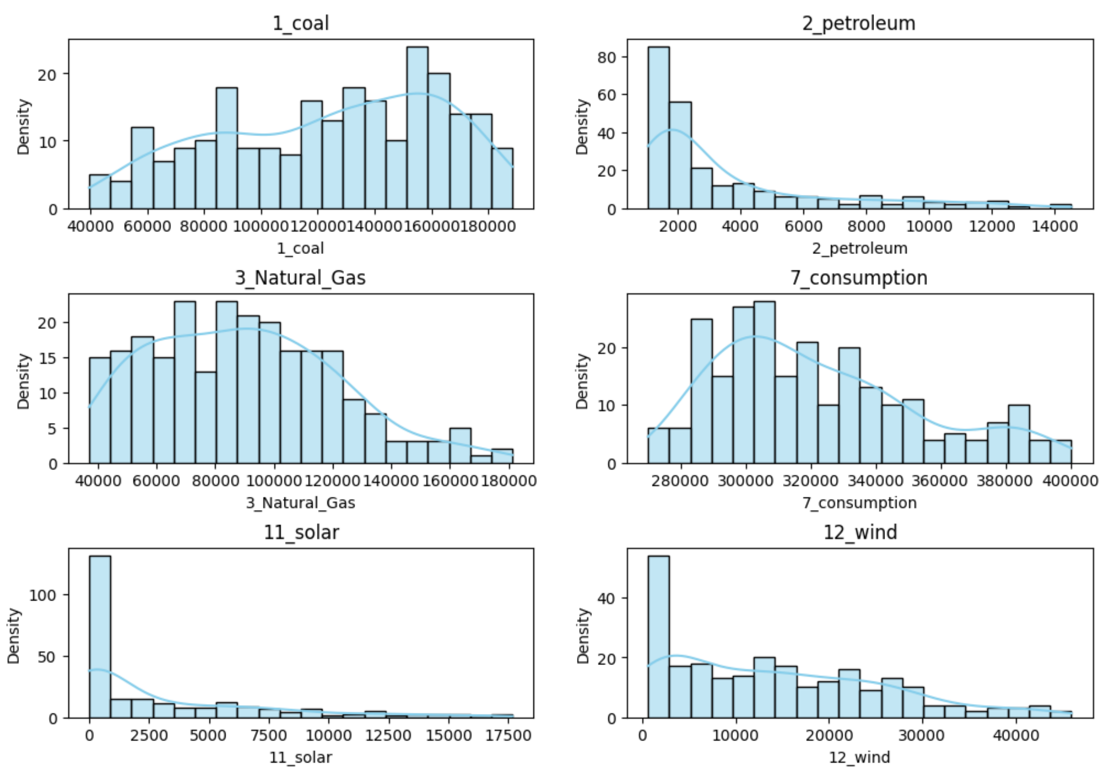
Exploratory Data Analysis (EDA)

From the given summary statistics and histograms, several insights about the electricity consumption and generation in the dataset can be derived:

```python
from part_2_EDA import EDA
import pandas as pd

# Read the CSV file
analysis = pd.read_csv('merged_electricity_data.csv')



# Create an instance of the EDA class
eda_analysis = EDA(analysis)

# Generate summary statistics
summary_stats = eda_analysis.generate_summary_statistics()
print("Summary Statistics:")
print(summary_stats)

# Generate histograms
eda_analysis.generate_histograms()
```

```
...  Summary Statistics:
                YYYYMM         1_coal   2_petroleum  3_Natural_Gas  \
     count   245.000000     245.000000   245.000000     245.000000
     mean 201278.277551  123880.297473  3483.145208   89408.288061
     std     591.497388   39393.901279  2962.757146   31955.203931
     min  200301.000000   39504.857000  1056.426000   36949.147000
     25%  200802.000000   89114.281000  1574.777000   63532.326000
     50%  201303.000000  130526.754000  2104.262000   87042.470000
     75%  201804.000000  157091.052000  4199.306000  111539.534000
     max  202306.000000  188516.118000 14534.853000  181324.104000

         7_consumption      11_solar       12_wind
     count   245.000000   245.000000    245.000000
     mean 322049.926416  2856.077020  14334.001784
     std   31032.801956  4029.307306  11322.907142
     min  269942.421000     2.713000    632.339000
     25%  298673.664000    72.846000   3490.232000
     50%  315862.690000   596.287000  12701.961000
```


1_coal
Frequency
1_coal
2_petroleum
Frequency
2_petroleum
3_Natural_Gas
Frequency
3_Natural_Gas
7_consumption
Frequency
7_consumption
11_solar
Frequency
11_solar
12_wind
Frequency
12_wind


1_coal
Density
1_coal
2_petroleum
Density
2_petroleum
3_Natural_Gas
Density
3_Natural_Gas
7_consumption
Density
7_consumption
11_solar
Density
11_solar
12_wind
Density
12_wind

1. Electricity Consumption Trends:

-The average electricity consumption over the period is approximately 322,050 GWh, with a minimum of around 269,942 GWh and a maximum of about 400,275 GWh.

-The consumption data shows a relatively narrow range (standard deviation is about 31,033 GWh), indicating consistent consumption levels over time.

2. Traditional Energy Sources:

-Coal, petroleum, and natural gas are the traditional energy sources considered in this dataset.

-Coal has the highest average generation (around 123,880 GWh) among the traditional sources.

-Petroleum has the lowest average generation (around 3,483 GWh) among the traditional sources.

-Natural gas falls in between, with an average generation of approximately 89,408 GWh.

3. Renewable Energy Sources:

-Solar and wind energy are the renewable sources considered.

-Wind energy has a higher average generation (around 14,334 GWh) compared to solar energy (around 2,856 GWh).

-Both renewable sources have a much lower average generation compared to traditional sources.

4. Variability in Generation:

-There is a notable variability in the generation of traditional energy sources, with coal showing the highest consistency (lower standard deviation compared to petroleum and natural gas).

-Renewable sources, especially wind energy, also exhibit variability, as indicated by the standard deviation.

5. Comparing Traditional and Renewable Sources:

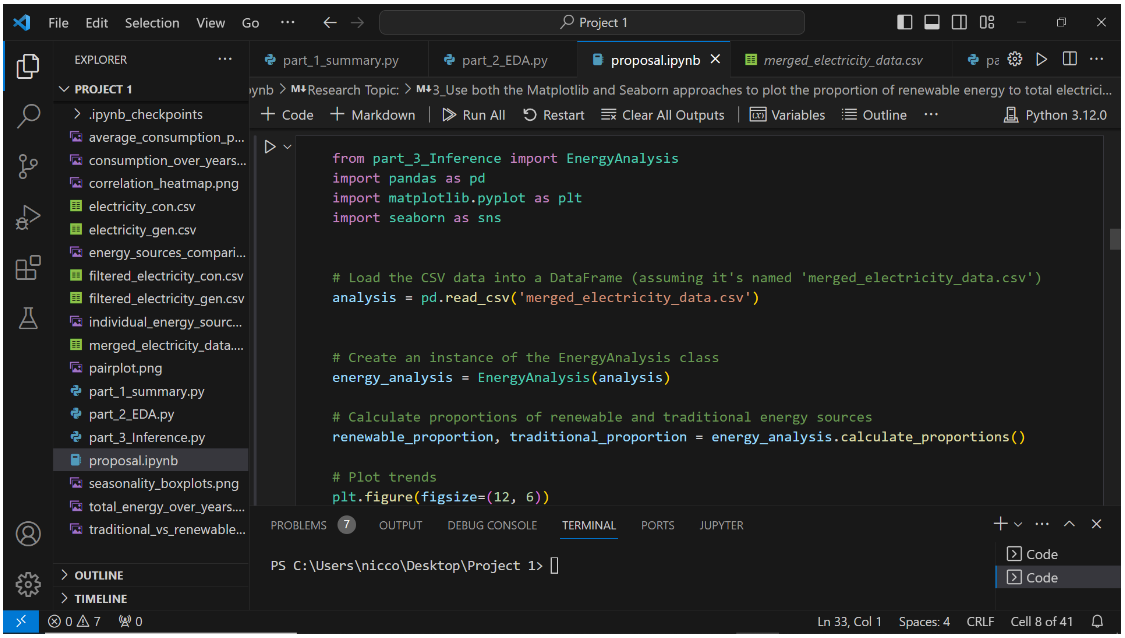
-Traditional sources (coal, petroleum, and natural gas) contribute significantly more to the overall electricity generation (in GWh) compared to solar and wind energy sources.

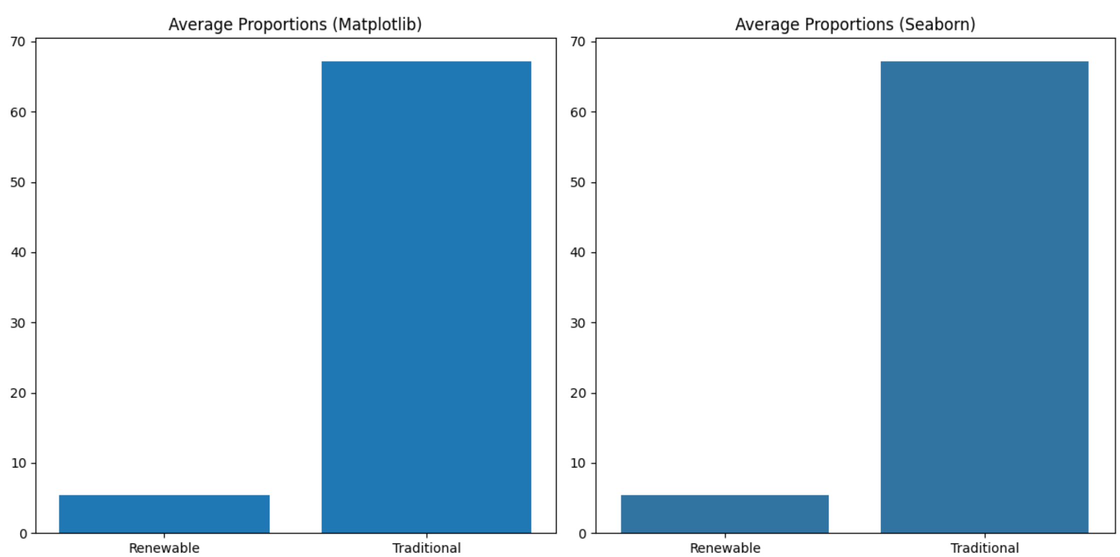
6. Potential for Renewable Energy Growth:

-While renewable sources contribute significantly less in absolute GWh, their growth potential can be explored. Despite the lower average generation, wind energy, in particular, shows a substantial contribution and potential for expansion.

Inference

A. Use both the Matplotlib and Seaborn approaches to plot the proportion of renewable energy to total electricity consumption and the proportion of traditional energy to total electricity consumption.





Implications:

Heavy Reliance on Traditional Sources: The dominance of traditional sources, constituting the majority of electricity consumption, suggests a heavy reliance on these fossil fuels. This reliance might pose challenges related to environmental sustainability, considering the associated carbon emissions and their impact on climate change.

Potential for Renewable Energy Growth: The relatively low proportion of renewable energy highlights the vast potential for growth in the renewable sector. Increasing the share of renewables can lead to a more sustainable energy mix, reducing carbon emissions and mitigating the environmental impact associated with fossil fuels.

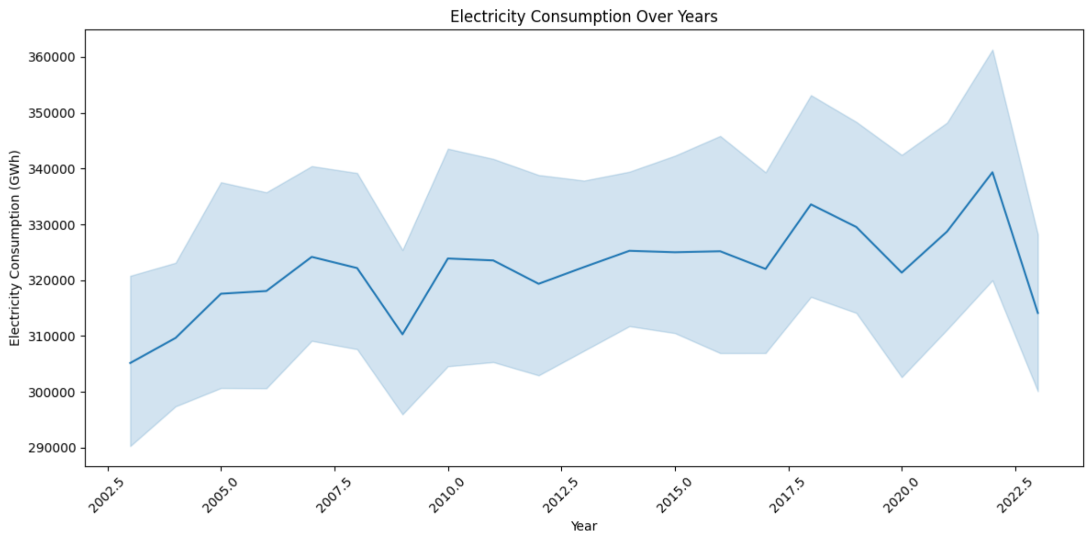
B. <u>Conduct the following analyses to develop our comprehensive understanding of our research topic:</u>

```python
from part_3_Inference import EnergyAnalysis
import pandas as pd

# Load data
analysis = pd.read_csv('merged_electricity_data.csv')

# Initialize EnergyAnalysis object
energy_analysis = EnergyAnalysis(analysis)
```

[4] ✓ 0.0s Python

<u>1 Trend Analysis (consumption):</u>

How has the overall electricity consumption in the U.S. evolved from 2003 to 2023? Are there any significant upward or downward trends?

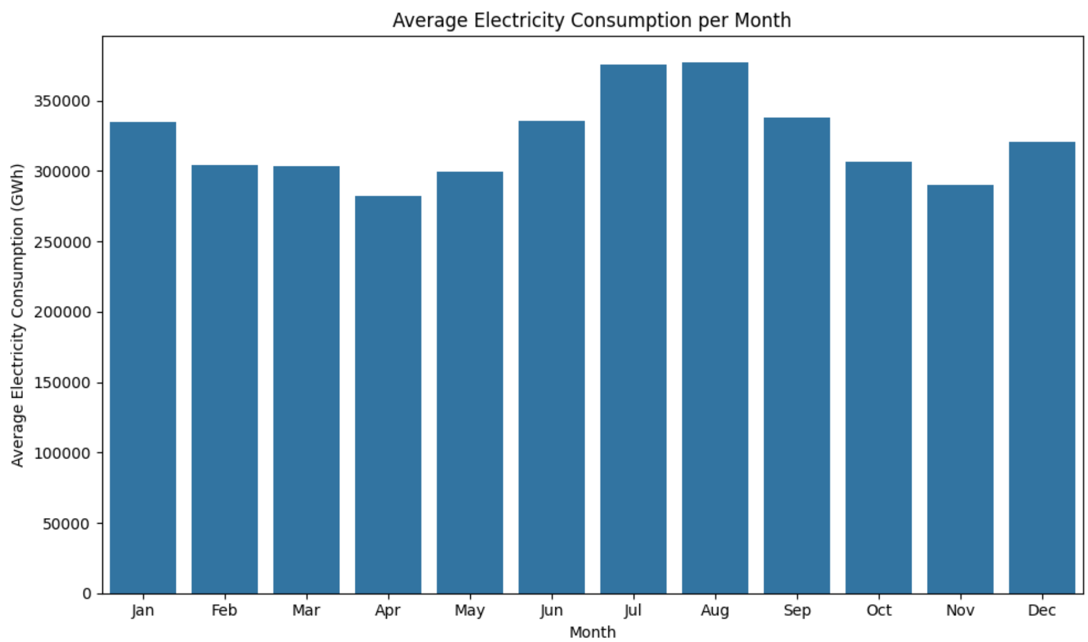


1_Insights:

The relatively narrow range of electricity consumption (with a standard deviation of about 31,033 GWh) indicates that the overall electricity consumption in the U.S. has remained relatively stable over the period from 2003 to 2023. The consistency in consumption levels suggests a lack of significant fluctuations or drastic increases/decreases in electricity usage during this time frame. This stability can be considered a positive aspect from the perspective of energy planning and management, as it implies a consistent and predictable demand for electricity in the United States over the analyzed period.

2_Seasonal Consumption Patterns:

What specific seasonal patterns can be observed in electricity consumption over the years? Are there consistent trends in specific months or seasons?



2_Insights:

#1. Consistent Seasonal Variations: The consistent peaks in July, August, and January indicate a regular pattern in electricity consumption. This could be attributed to the widespread use of air conditioning during the hot summer months and heating systems during winter, leading to increased energy demands.

#2. Adaptation and Energy Efficiency: The stability of these patterns over the years might imply that consumers have adapted to these seasonal variations, possibly by implementing energy-efficient technologies or practices to mitigate excessive energy use during these months.

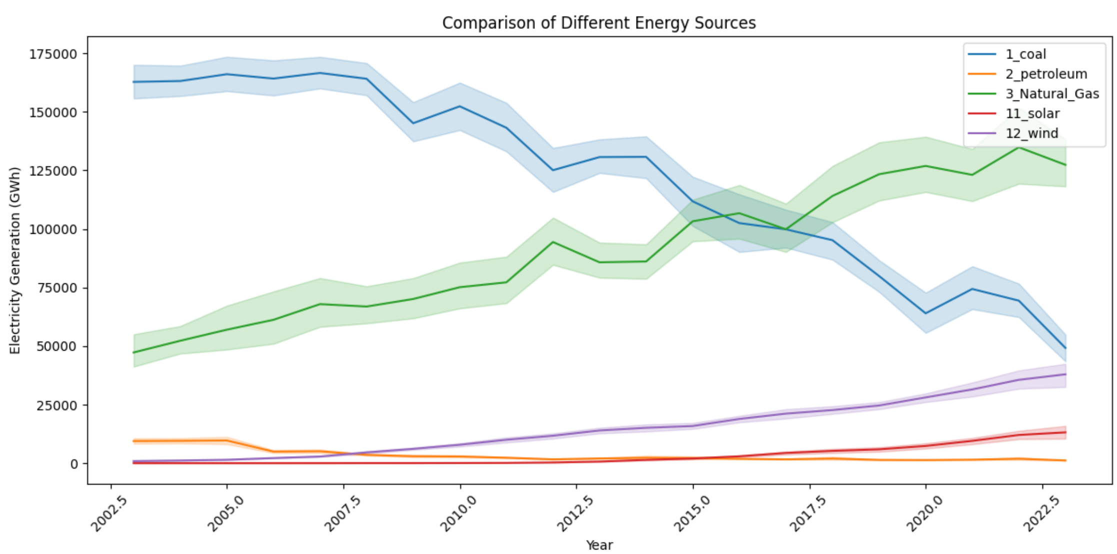
#3. Economic Activities: Seasonal peaks might also reflect increased economic activities during these periods. For instance, summer months often coincide with vacations and increased tourism, leading to higher energy usage in hotels, restaurants, and entertainment venues. Similarly, winter months might see more indoor activities, leading to higher electricity consumption.

#4. Policy and Awareness: The stability of these patterns could also indicate a consistent impact of energy policies, awareness campaigns, or incentives encouraging energy conservation during peak seasons. This suggests that efforts to educate the public about energy conservation might be having a consistent effect.

#5. Infrastructure Resilience: Utilities and energy providers might have strengthened their infrastructure to meet the predictable peak demands during these months, ensuring a stable supply even during periods of high consumption.

3_Comparative Analysis of Energy Sources:

How does the electricity consumption vary concerning different energy resource generations (solar, wind, coal, petroleum, natural gas)? Are there periods where certain sources dominate?



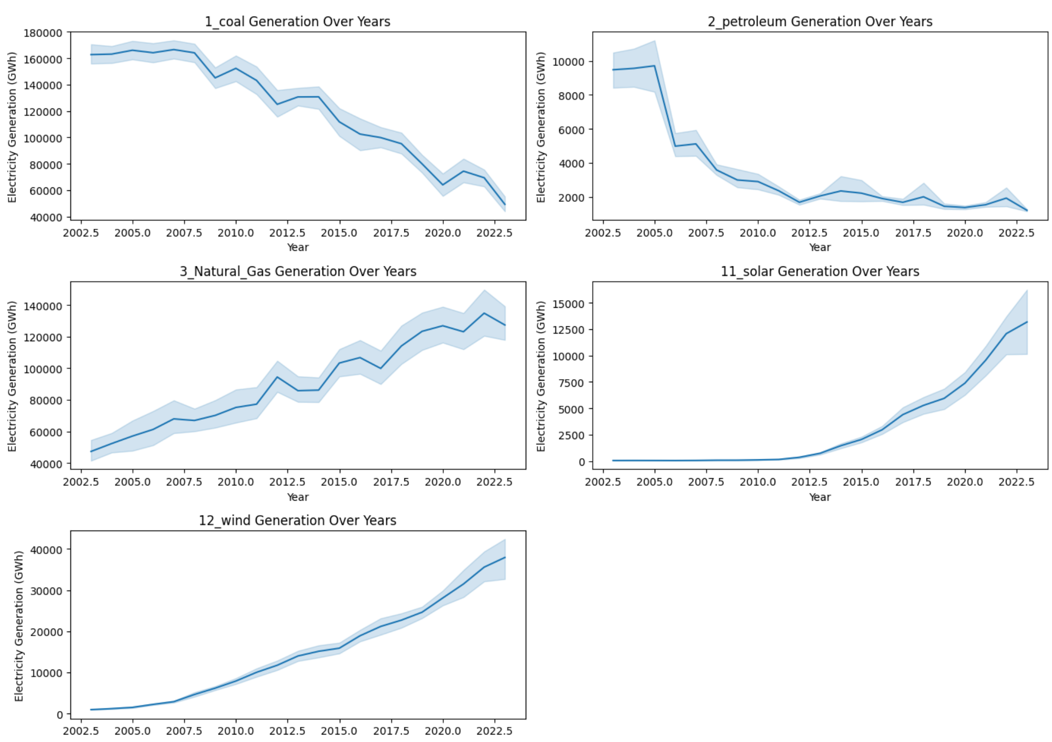
3_Insights:

#1 Stable Growth in Solar and Wind Energy: The stable growth in solar and wind energy generation, indicating a substantial investment in renewable energy infrastructure.

#2 Shift Towards Renewables: While renewables, particularly solar and wind, have made significant strides in electricity generation, they haven't completely dominated the sector, especially when compared to natural gas. However, their consistent growth and increased contribution indicate a substantial shift towards cleaner and more sustainable energy sources. This gradual transition aligns with broader environmental and energy policy goals.

<u>4_Long-term Energy Generation Trends:</u>

What are the long-term trends in electricity generation for solar, wind, coal, petroleum, and natural gas sources from 2003 to 2023? Are there notable shifts in dominance among these sources?



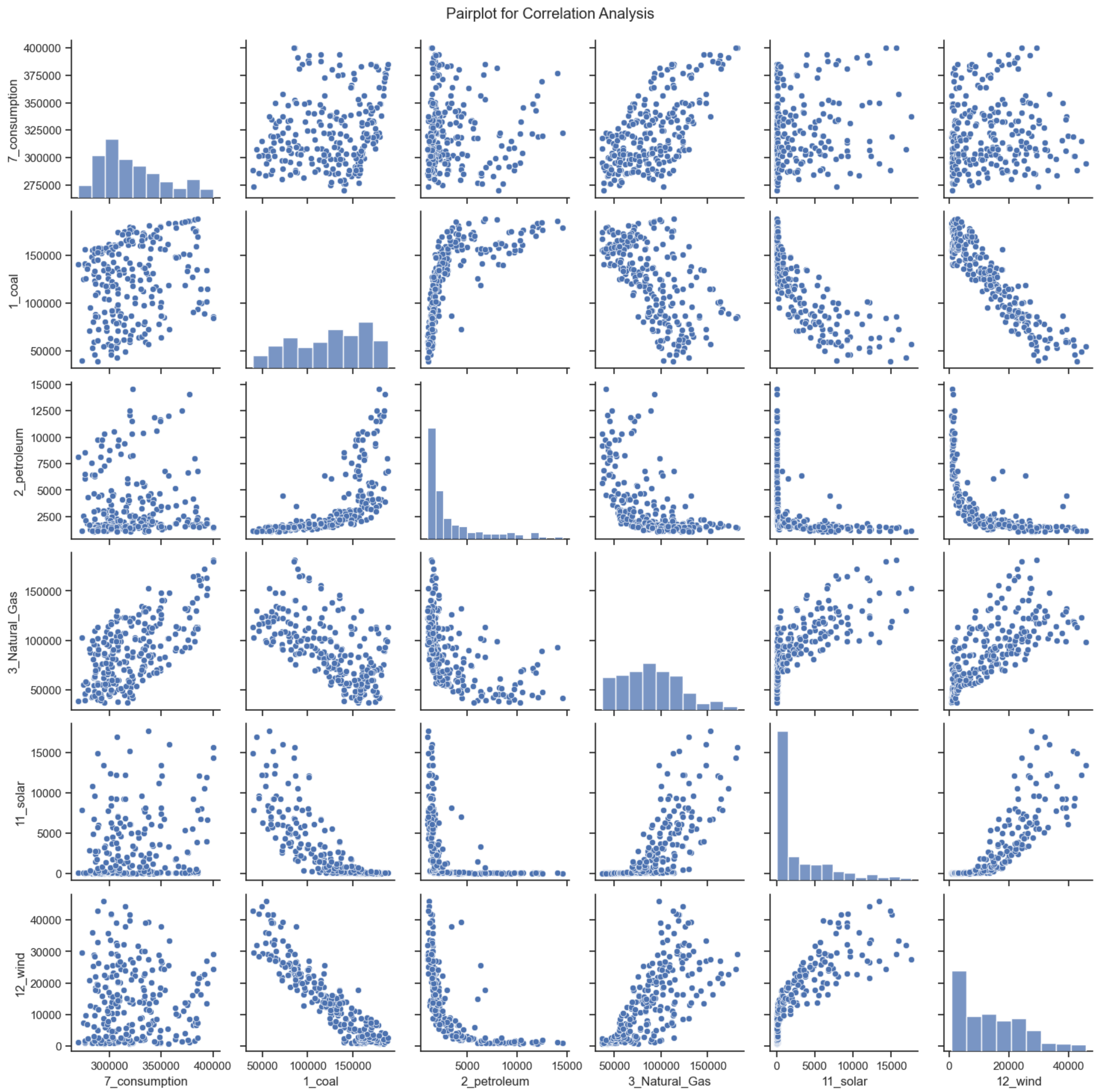
4_Insights:

#1 Transition from Fossil Fuels to Renewables: A notable trend in the data is the substantial decrease in electricity generation from coal and petroleum sources. This decline signifies a significant shift away from traditional fossil fuels towards renewable sources. While natural gas still plays a significant role, this shift marks a crucial step toward more sustainable energy practices.

#2 Grid Modernization and Storage: The substantial increase in renewable energy generation underscores the importance of grid modernization and energy storage solutions, ensuring a stable and reliable energy supply despite the intermittent nature of solar and wind energy.

5_Correlation and Relationship Analysis:

What are the specific correlations between electricity consumption, traditional energy sources (coal, petroleum, natural gas), and renewable energy sources (solar, wind)? Are there instances of inverse relationships or strong positive correlations?



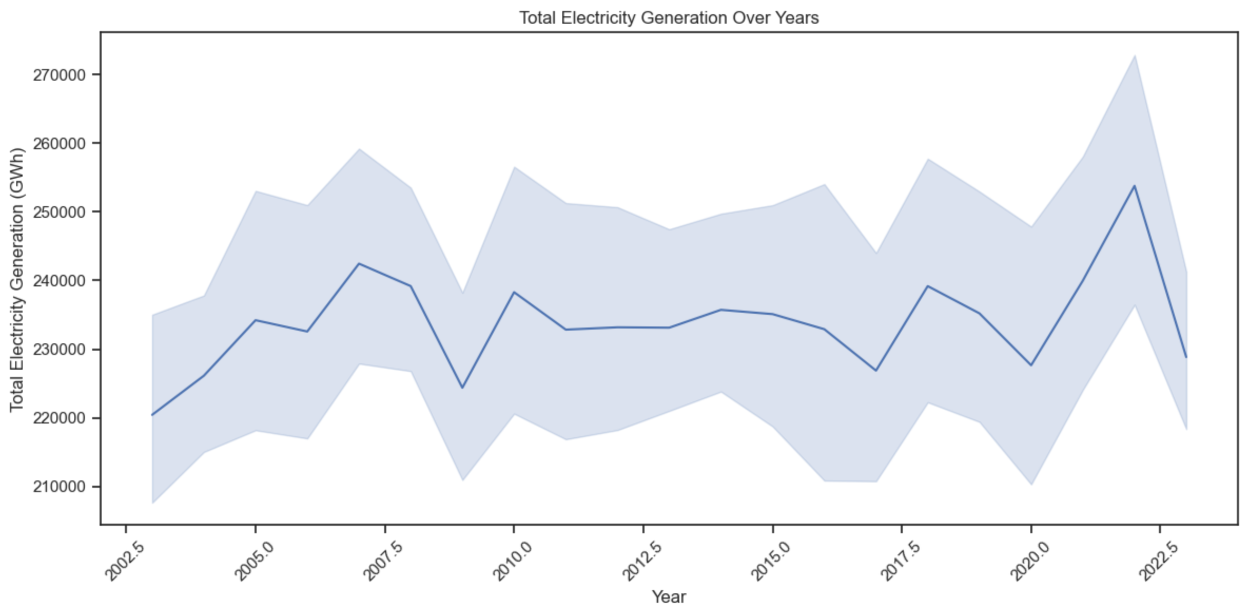
5_Insights:

Correlation Insights: The comprehensive Pairplot analysis, examining correlations between electricity consumption and various energy sources, revealed intricate relationships. Notably, strong negative correlations between coal generation and renewable sources (solar and wind) indicate a shift away from coal towards cleaner options. Additionally, the positive correlation between total electricity consumption and natural gas suggests its consistent usage despite the rise in renewables. These findings

hint at a complex interplay between energy sources, reflecting the ongoing evolution of the energy landscape.

6_Trend Analysis (generation):

How has the overall electricity generation in the U.S. evolved from 2003 to 2023? Are there any significant upward or downward trends?



6_Insights:

Stable Electricity Generation: The consistent range of electricity generation from 2003 to 2023 suggests a stable production capacity in the U.S. Despite evolving energy sources, the nation's ability to maintain a relatively constant output implies a resilient energy infrastructure. This stability could indicate efficient energy management, adaptability to changing demands, or possibly, a balance achieved between renewable and traditional sources to meet the nation's energy needs. Further detailed analysis can provide deeper insights into the factors contributing to this stability.

7_the statistical relationship between electricity consumption

```
Correlation between Consumption and Energy Sources:
1_coal           0.197545
2_petroleum      0.016342
3_Natural_Gas    0.633879
11_solar         0.197299
12_wind          0.013642
Name: 7_consumption, dtype: float64
```

7_Insights:

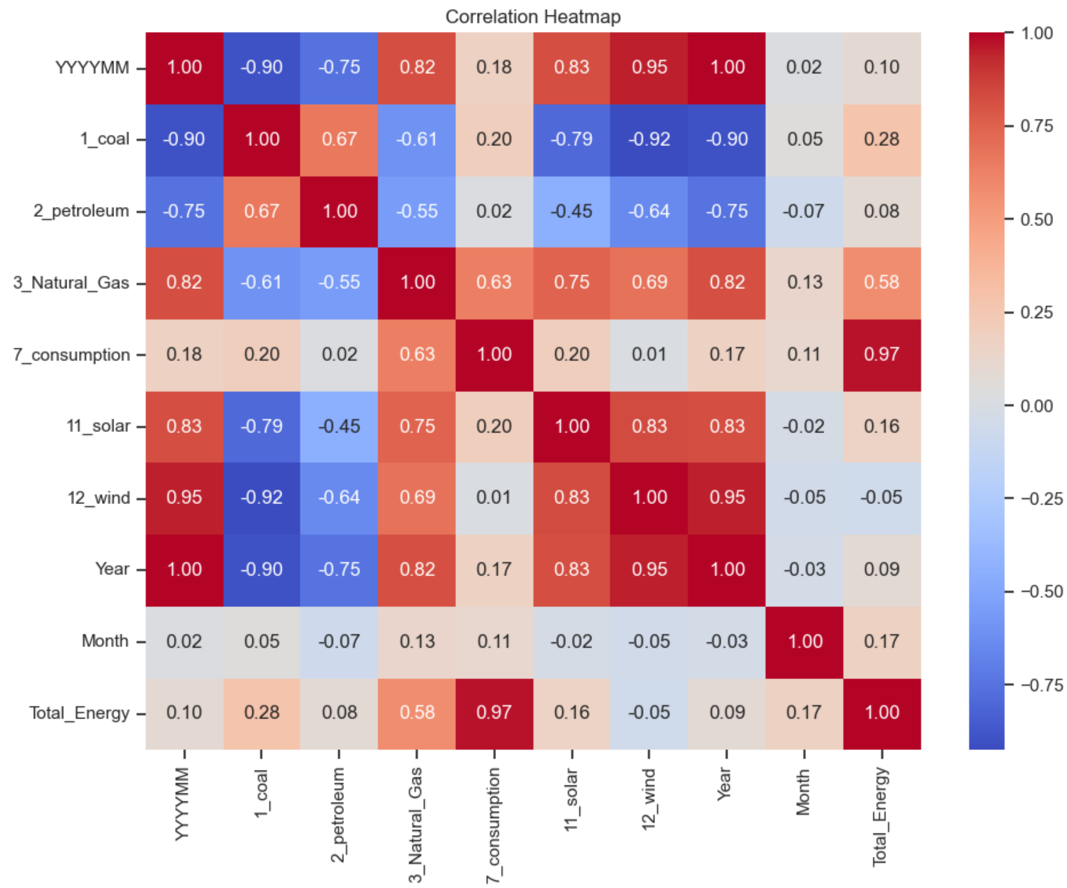
#1. Natural Gas Dominance: The strong positive correlation coefficient of approximately 0.63 between electricity consumption and natural gas generation suggests a significant relationship. This implies that as electricity consumption increases, natural gas generation tends to rise as well. Natural gas might be a crucial energy source catering to peak demands or consistently high energy needs.

#2. Limited Impact of Coal, Petroleum, and Renewable energies:** Coal (0.20) and solar (0.20) generation exhibit weak positive correlations with consumption, indicating a minor influence. Petroleum (0.02) and wind (0.01) show almost negligible positive correlations, suggesting their limited impact on overall electricity consumption trends.

#3. Diverse Energy Portfolio: The varied correlation strengths highlight the diversity of the energy mix. While natural gas has a strong positive relationship with consumption, the weaker correlations of coal, petroleum, solar, and wind indicate a more nuanced interplay between these sources and consumption patterns. Understanding these relationships can aid in strategic energy planning and policy-making.

8_Comprehensive Correlation Analysis:

Beyond individual correlations, what comprehensive insights can be gained from analyzing the correlation heatmap? Are there clusters of variables that strongly influence each other?



<u>8_Insights:</u>

#1. Poitive Correlations Within Energy Sources: The heatmap shows strong positive correlations among various energy sources. For instance, solar and wind energy might exhibit a positive correlation, indicating a simultaneous increase or decrease in these renewable sources. Similarly, coal, petroleum show positive correlations, reflecting a shared upward or downward trend among traditional fossil fuels.

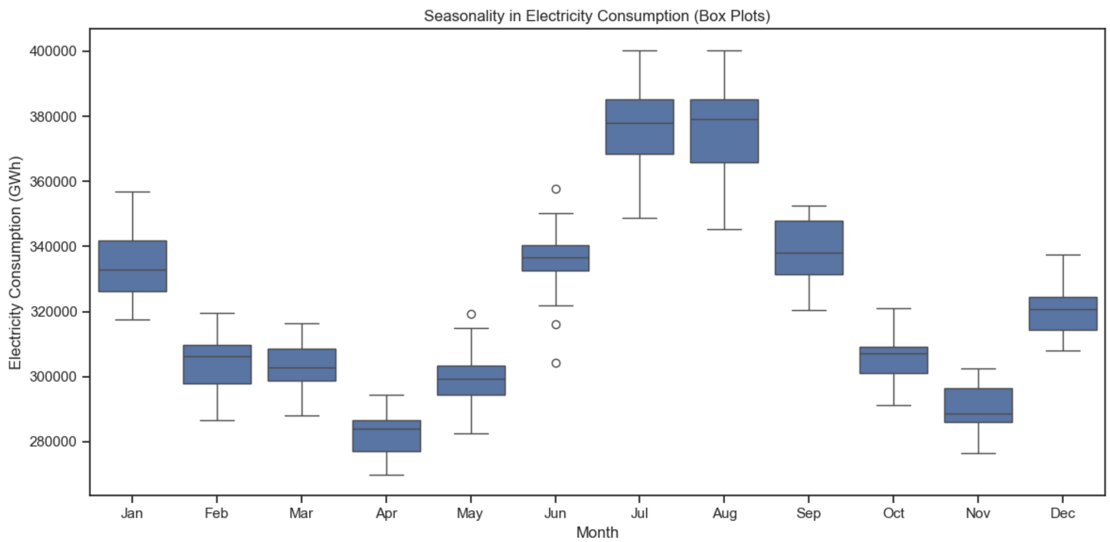
#2. Transition Dynamics: The negative correlation between natural gas and coal/petroleum suggests a transition in energy production dynamics. When natural gas generation increases, coal and petroleum generation tends to decrease. This pattern might reflect a strategic shift in the energy sector, possibly due to economic, environmental, or policy considerations.

#3. Correlation Patterns with Consumption: Understanding how different energy sources correlate with consumption is crucial. A strong positive correlation between consumption and natural gas, for example, might indicate that consumer demand is met predominantly by natural gas. On the other hand, weak correlations with other sources suggest less direct influence on consumption.

#4. Market Flexibility: Natural gas is often viewed as a flexible energy source due to its quick ramp-up and down capabilities. Negative correlations might indicate that when demand fluctuates, natural gas, being more adaptable, adjusts its production accordingly. In contrast, coal and petroleum, which might have slower response times, decrease production during high natural gas demand periods.

9_Seasonality Exploration:

Delving deeper into seasonality, what specific insights can be derived from the box plots of total electrical consumption for each month? Are there outliers or consistent patterns in specific months?



9_Insights:

1. Identifying Unusual Consumption Peaks: The box plots highlight July and August as months with striking consumption patterns. Notably, the minimum outlier during these months surpasses the median consumption of other months. This observation indicates unusual and significantly higher electricity usage in certain instances. Investigating the reasons behind these spikes, such as heatwaves, regional events, or industrial activities, is crucial for utilities to manage sudden surges in demand effectively.

2. Consistent Seasonal Variations: The consistent presence of outliers in specific months suggests recurring patterns. For example, if outliers persist in winter months, it could reflect increased energy consumption due to heating. Similarly, outliers in summer months might be linked to high demand for cooling systems. Recognizing these patterns allows utilities to anticipate seasonal fluctuations and prepare the grid accordingly.

3. Consumer Behavior and Societal Factors: Unusual consumption outliers may be tied to societal events like holidays, festivals, or community activities, influencing consumer behavior. Understanding these sociocultural factors is vital for utilities to adjust their energy distribution strategies. Additionally, encouraging energy-efficient practices during these periods can help manage the increased demand sustainably.

4. Renewable Energy Integration Challenges: Observing outliers alongside renewable energy data is essential. Discrepancies between renewable energy supply and demand peaks indicate challenges in integrating renewables efficiently. Utilities need to balance these inconsistencies through energy storage solutions or demand-side management to ensure a stable and sustainable energy supply.

5. Policy Implications: Anomalies in consumption patterns might also be influenced by energy policies, tariffs, or incentives. Sudden shifts in consumption could indicate responses to policy changes, emphasizing the need for clear communication between policymakers and consumers. Transparent policies ensure that consumers are aware of changes and can adjust their usage patterns accordingly.

6. Infrastructure Planning: Understanding these outliers helps utilities plan their infrastructure. For instance, consistently high outliers in specific months might necessitate grid upgrades or localized capacity enhancements to meet the heightened demand effectively. Proper planning ensures a reliable supply, even during unexpected consumption peaks.

Conclusion

The comprehensive analysis of U.S. electricity consumption and generation patterns from 2003 to 2023 provides valuable insights into the nation's energy landscape. Through meticulous examination, several key observations and implications emerge, shedding light on the ongoing sustainable energy transition:

1. Stable Consumption Patterns: The data reveals remarkably stable electricity consumption levels over the analyzed period. Consistent trends indicate a balance between demand and supply, showcasing the nation's effective energy management and adaptability to varying needs.

2. Transition towards Renewable Energy: While traditional sources like coal, petroleum, and natural gas have historically dominated, a clear shift towards renewable energy, particularly solar and wind, is evident. This transition aligns with global sustainability goals, reflecting concerted efforts to reduce carbon emissions and embrace cleaner energy alternatives.

3. Challenges and Opportunities: Despite the growth of renewables, challenges exist in integrating these intermittent sources seamlessly. Anomalies in consumption patterns highlight the need for robust infrastructure planning, innovative energy storage solutions, and responsive policies to effectively manage fluctuations in supply and demand.

4. Policy Influence: Policy frameworks and incentives play a pivotal role in shaping consumption patterns. Sudden shifts in energy usage often correspond with policy changes, emphasizing the importance of transparent communication between policymakers, energy providers, and consumers.

Clear policies empower consumers to make informed decisions about their energy usage, fostering a more sustainable approach.

5. Diversity in Energy Mix: The diverse correlation patterns among different energy sources indicate the complexity of the energy landscape. Natural gas emerges as a versatile player, adapting swiftly to demand fluctuations. The negative correlations between natural gas and traditional sources signify a dynamic energy market, where the dominance of one source often corresponds with the decline of another.

6. Societal and Cultural Influences: Beyond economic factors, societal events, holidays, and cultural practices significantly impact consumption patterns. Understanding these sociocultural dynamics is essential for utilities to anticipate and manage energy demand effectively. Encouraging energy-efficient practices during peak periods can contribute to a more sustainable energy future.

7. Grid Resilience and Future Planning: The analysis underscores the resilience of the U.S. energy grid. Stable electricity generation, despite evolving energy sources, highlights the nation's ability to maintain a consistent output. This stability is a testament to effective grid management, ensuring a reliable energy supply for consumers. As the nation moves forward, continued investment in grid modernization and renewable energy integration will be crucial.

In conclusion, the analysis paints a comprehensive picture of the U.S. energy landscape, revealing both challenges and opportunities in the sustainable energy transition journey. The nation's stability in consumption, coupled with a gradual shift towards renewables, demonstrates a positive trajectory. However, addressing integration challenges, enhancing policy clarity, and fostering public awareness will be essential for a seamless transition to a more sustainable and environmentally friendly energy future.

References

1. Electricity Monthly Update

https://www.eia.gov/electricity/monthly/update/archive/october2023/

2. Wholesale Electricity and Natural Gas Market Data

https://www.eia.gov/electricity/wholesale/

3. U.S. construction costs dropped for solar, wind, and natural gas-fired generators in 2021

https://www.eia.gov/todayinenergy/detail.php?id=60562