Data 604 Final Project:

Evolving Energies: A Comprehensive Analysis of Trends and Shifts in U.S. Energy Consumption

Team 1: Jackson Mullin, Annie Shamirian, Gregory Mathurin, Aysha Afreen Althaf, Berenice Morales Silva

December 8, 2023

Table of Contents

Abstract

This study investigates the dynamic landscape of energy consumption in the United States (U.S.), revealing a notable shift towards sustainability, particularly with the ascendance of natural gas and renewable energy over coal and petroleum. Spanning from historical milestones like the Cold War to the transformative impact of events like the COVID-19 pandemic, our analysis highlights the interplay between external influences and evolving energy patterns.

We further examine patterns in energy consumption through seasonal decomposition and analysis on the total energy consumption mix across natural gas, renewable energy, coal, and petroleum. Then, we further analyze the breakdown of consumption for renewable energy subcategories. These patterns which we reveal underscore the connection between socioeconomic conditions and energy consumption in the U.S.. Through this analysis, we are able to identify the patterns in CO^2 emissions, impact of geopolitical events on energy consumption, seasonal patterns in energy consumption, and ultimately understand the country's efforts to move toward a more sustainable energy landscape.

Introduction:

For this project, our group has decided to tackle the important pressing matter of energy production and associated CO^2 emissions. We decided, for this project, to focus on energy production solely within the U.S., as they are not only a global policy leader, but also one of the largest energy producers in the world. As a result, the U.S. plays a pivotal role in establishing precedence and global governing ideas regarding climate change. This issue is important because of the worldwide existential threat that climate change poses in both economic and social security for all nations. Throughout this analysis, we took on a pragmatic approach in comparing the consumption of coal, petroleum, natural gas, and renewable energy, focusing on determining the effects of these sources on CO^2 emissions and the composition of energy production in the U.S. Our guiding questions to shed light on these issues are as follows:

1. Despite the increase in energy consumption, is there a relative decrease in CO^2 emissions, and is there a noticeable change in energy production efficiency by the amount of CO^2 produced?
2. Are we able to see noticeable changes in energy consumption due to significant global events? (For example, collapse of the Soviet Union, wars, 9/11, etc.)
3. Are there noticeable seasonal patterns in energy consumption?
4. How has the contribution of each energy source to the total energy mix changed over time?
5. Are there any subcategories in renewable energy that have experienced rapid adoption?

Prior to diving into the details of the guiding questions and the resulting exploration of the data, it is first necessary to address the composition of the data.

The Datasets:

The data for our project was sourced from the U.S. Energy Information Administration, which regularly publishes recent and historical energy statistics gathered through a monthly energy review process. We utilized five distinct datasets covering the period from January 1973 to September 2023, each focusing on different energy sources as mentioned above. These datasets are publicly accessible, and no specific license is required for their use.

- Petroleum Dataset – This dataset pertains to petroleum energy consumption, containing approximately 609 rows. Our variables of interest in this dataset are the date by month and Petroleum consumption (Thousand Barrels per Day).
- Coal Dataset – This dataset covers coal energy consumption and consists of around 609 rows. Our variables of interest here are the date by month and Coal Consumption (Thousand Short Tons).
- Renewable Energy Dataset – This dataset focuses on renewable energy consumption, including subcategories such as hydroelectric, geothermal and biofuels, comprising approximately 607 rows. Our variables of interest in this dataset are the date by month and Renewable Energy Consumption (Trillion Btu).
- Natural Gas Dataset – This dataset pertains to natural gas consumption, also with around 607 rows. Our variables of interest in this dataset are the date by month and Natural Gas Consumption (Billion Cubic Feet).
- CO^2 Emissions Dataset – This dataset detailing carbon dioxide emissions produced by different energy sources, containing about 607 rows. Our variables of interest in this dataset

include the date by month, petroleum excluding biofuels CO^2 emissions, Coal CO^2 Emissions and Natural Gas CO^2 Emissions (Million Metric Tons of Carbon Dioxide).

Each data set is structured in a tabular format within an xls file and contains roughly 600 rows. We combined these datasets into one unified dataset using the month column as a common identifier for analysis. Using these datasets, we designed a dimensional star model based on Ralph Kimball's methodology. Star schemas consist of fact tables linked to associated dimension tables through primary and foreign key relationships. Our dimension tables represent the entities for energy source, unit, time, and sector. The fact tables contain the main metrics of consumption and carbon dioxide emissions. We established one-to-many relationships between the dimension and fact tables. This model enables us to lower the processing time and reduce the code complexity of sequel queries. The schema is provided in Appendix 1.1.

We list variables of importance in a series of tables which are broken down into the following categories:

1. Coal_by_Sector
2. Natural_Gas_by_Sector
3. Petroleum_Industrial
4. Petroleum_Residential_Commercial
5. Petroleum_Transportation_Electric_Power
6. Renewable_Energy_Industrial
7. Renewable_Energy_Residential_Commercial
8. Renewable_Energy_Transportation_Electric_Power
9. Primary_Energy_Consumption_by_Source

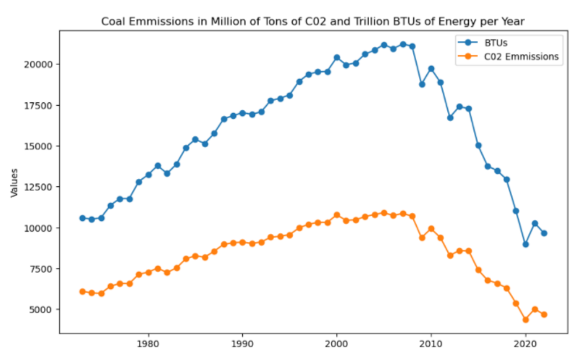
The tables listed above are provided in Appendix 1.2.

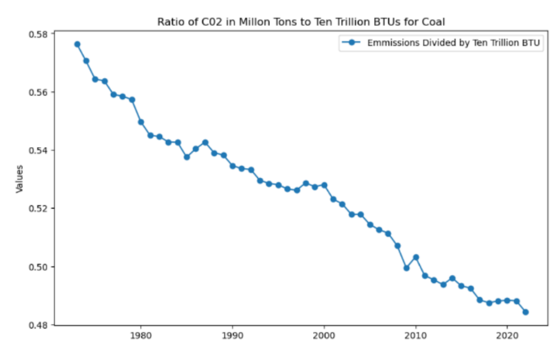
Guiding Question 1 (Exploration):

Despite the increase in energy consumption, is there a relative decrease in CO^2 emissions, and is there a noticeable change in energy production efficiency by the amount of C02 produced?

For this portion of the report, we were most concerned with establishing whether carbon-emitting fuel sources have in their capacity to produce energy, become more efficient due to either advances in technology or by other means. The method by which we undertook this task was to first break down both CO^2 emissions and consumption in BTUs by major unrenewable energy sources, being Natural Gas, Petroleum and Coal. Specific sectors or distinguishing by industry was unessential for this portion of the project as we were concerned simply with demonstrating if more efficient energy production standards have been established within the U.S. For all graphs for efficiency (ratio), the efficiency metric was calculated by the tonnage of CO^2 produced by each fuel source divided by the BTUs of energy produced. The first visualization we created for this was for emissions produced by coal.

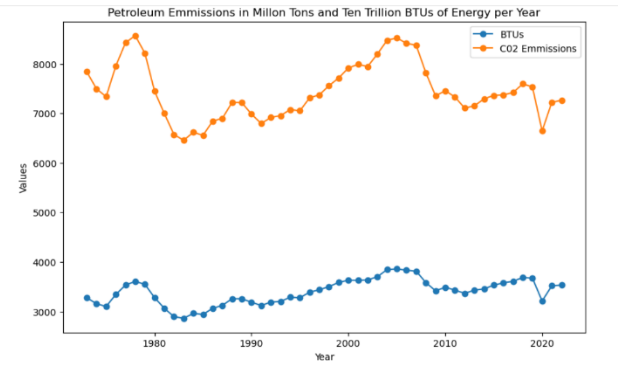
(Figures 2.1, 2.2)

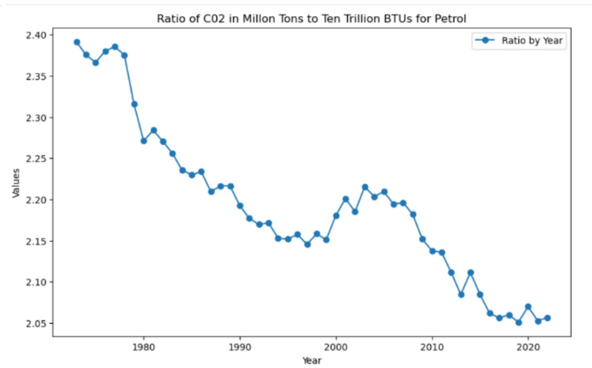




We can see from the above plots that there is a trend in the U.S. towards less coal usage around 2009, yet there is a noticeable increase in efficiency present in the right-hand graph. Next, we can look at the graphs for petroleum.

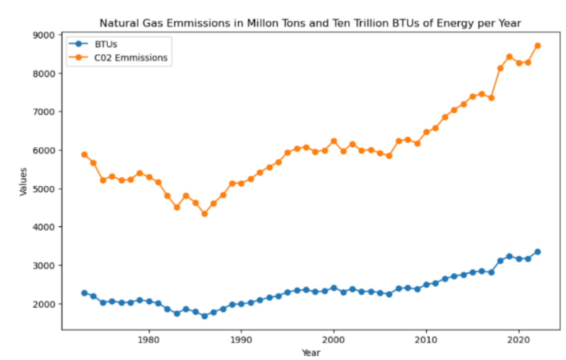
(Figures 2.3, 2.4)

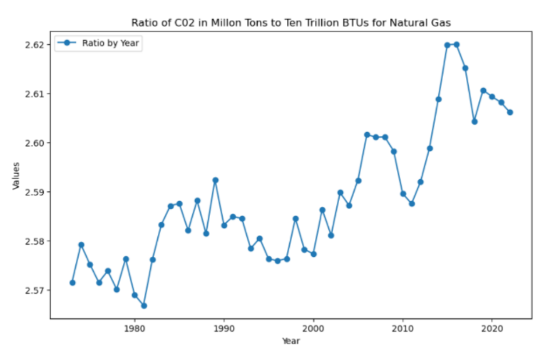




From the above graphs, we can see that the amount of petroleum products consumed over the years remains relatively stagnant, but given the right-hand graph, we can observe that this there has been a massive increase in overall efficiency for energy output by petroleum products. Next, we can look at the same graphs for natural gas consumption.

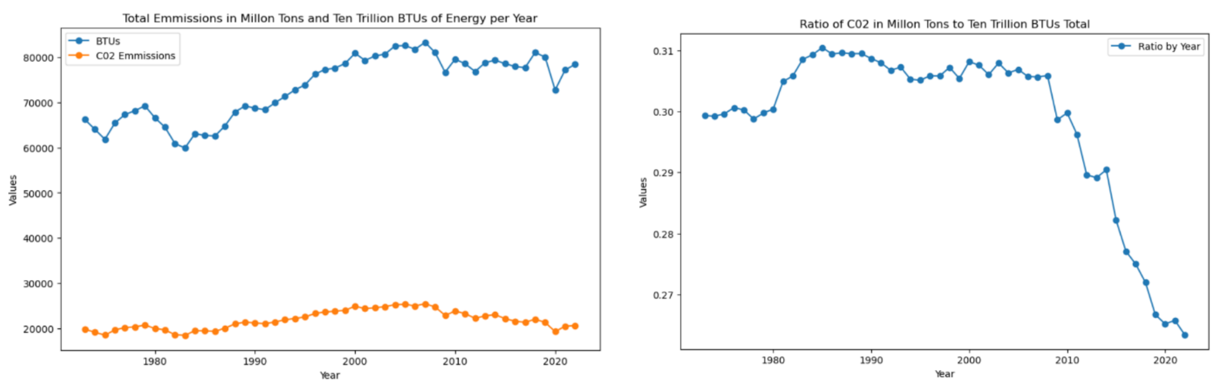
(Figures 2.5, 2.6)





From the above graphs we can see that natural gas has become a much more popular form of energy production, however, from the right-hand side graph we can see that natural gas has largely become less efficient in recent years. This might be because with the phasing out of coal energy production, natural gas has been integral in filling the energy production gap. As natural gas becomes more popular, the technology to make it more efficient might not be as present as it is with other more established fuel sources. Finally, to close out this guiding question, we can look at the total energy production and CO^2 emissions totals.

(Figures 2.7, 2.8)



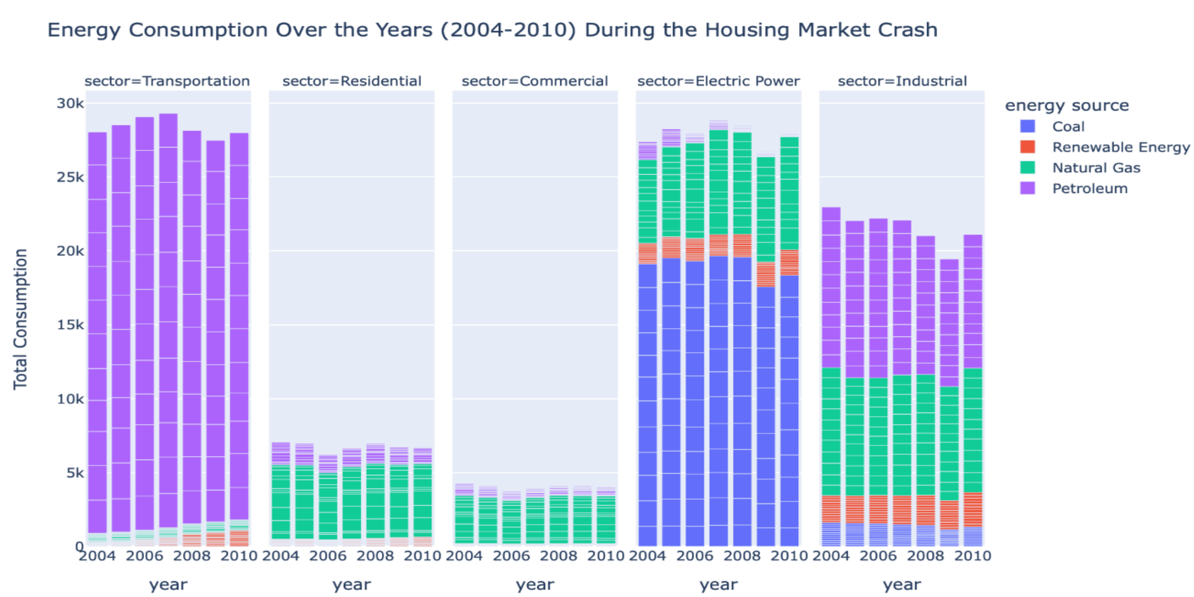
In an era marked by an escalating awareness of climate change and environmental sustainability, it is imperative to understand the environmental consequences of our energy choices. The overarching goal of this guiding question is to assess whether, despite the surge in energy consumption, we are witnessing a relative decrease in CO^2 emissions. This inquiry is rooted in the commitment to hold ourselves accountable for the environmental footprint associated with our energy demands. From the graphs displayed above, we can see that some valuable strides are being made in being accountable for our energy choices, and in the very least, traditional CO^2 emitting energy sources have become more efficient.

Guiding Question 2 (Exploration):

Are we able to see noticeable changes in energy consumption due to significant global events?

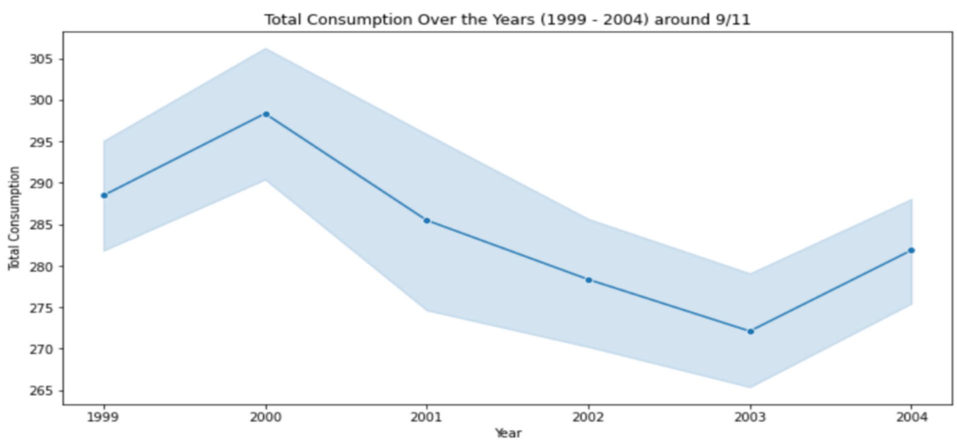
The investigation into changes in energy consumption following significant global events stems from the recognition that the energy sector is deeply intertwined with global socio-economic and geopolitical dynamics. Events such as economic recessions, geopolitical conflicts, or public health crises can exert profound influence on energy demand patterns. By analyzing these shifts, we aim to unveil the energy sector's sensitivity to external influences. Starting with the housing market crash, we query the total consumption by sector and energy source to analyze changes over the period from 2004 to 2010. We are interested in seeing any significant changes from around 2007 to 2008 when the crash took place.

Housing Market Crash: Figure 1



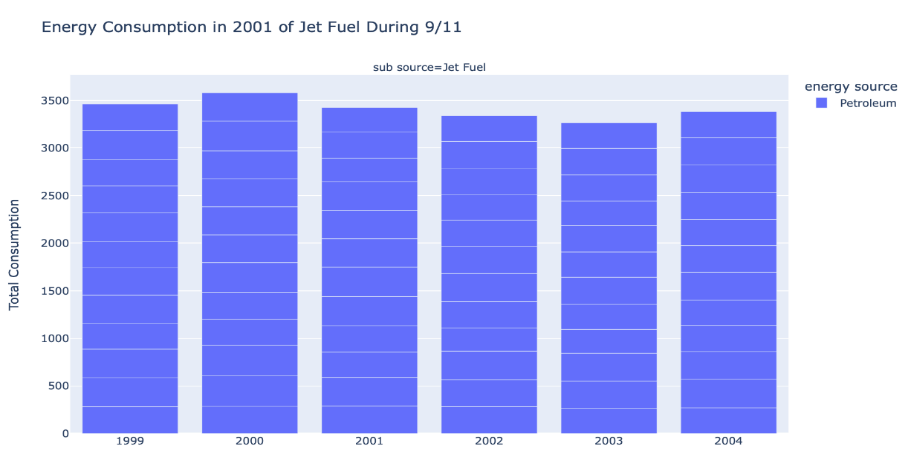
After visualizing the query as shown in Figure 1, we can see there is a slight drop in overall energy consumption during the buildup of the crash. The SQL code for querying and Python code for the above visualization can be found in Appendix 3.1.

Total Consumption During 9/11: Figure 2



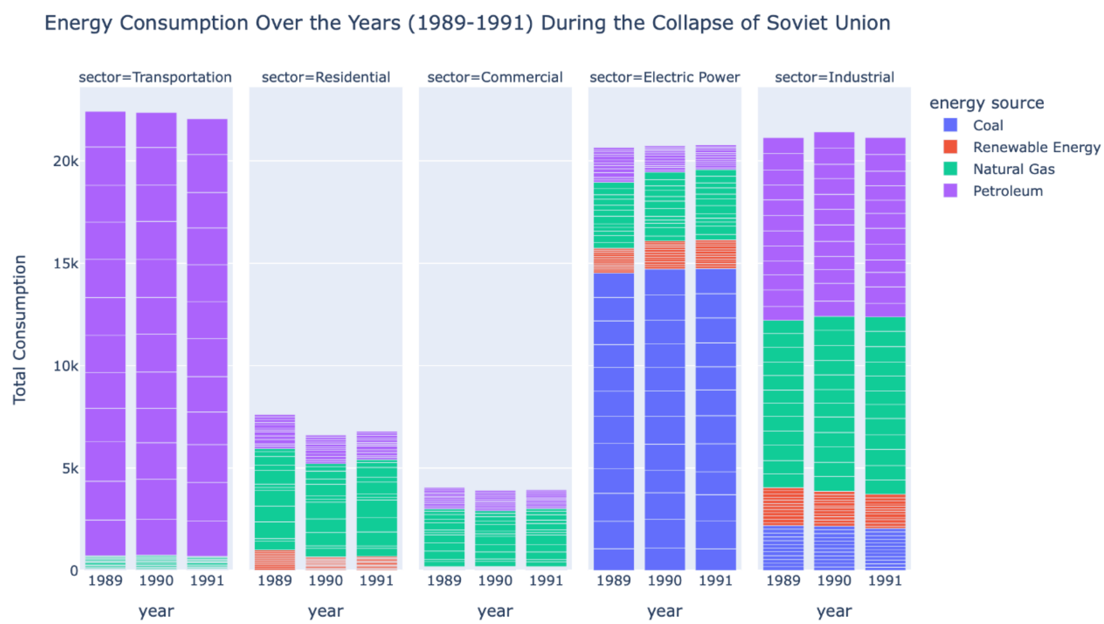
The September 11th attacks are analyzed in the context of changes in travel and transportation patterns. The line graph in Figure 2 displays no significant change in total consumption during the period following 2001.

Jet Fuel Consumption During 9/11: Figure 3



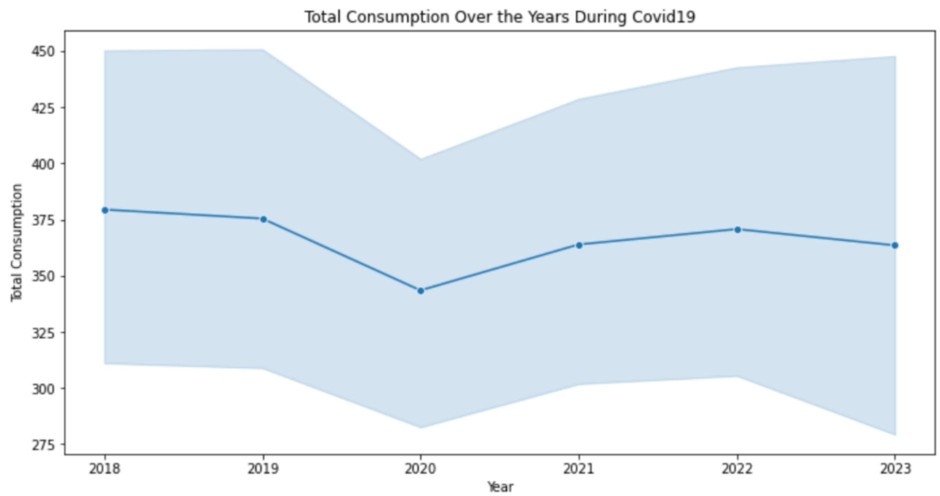
Looking specifically at jet fuel energy consumption during this same period in Figure 3, we can see a slight decrease after the attacks, likely due to reduced travel. The SQL code for queries and Python code for visualizations can be found in Appendix 3.2.

Collapse of Soviet Union: Figure 4



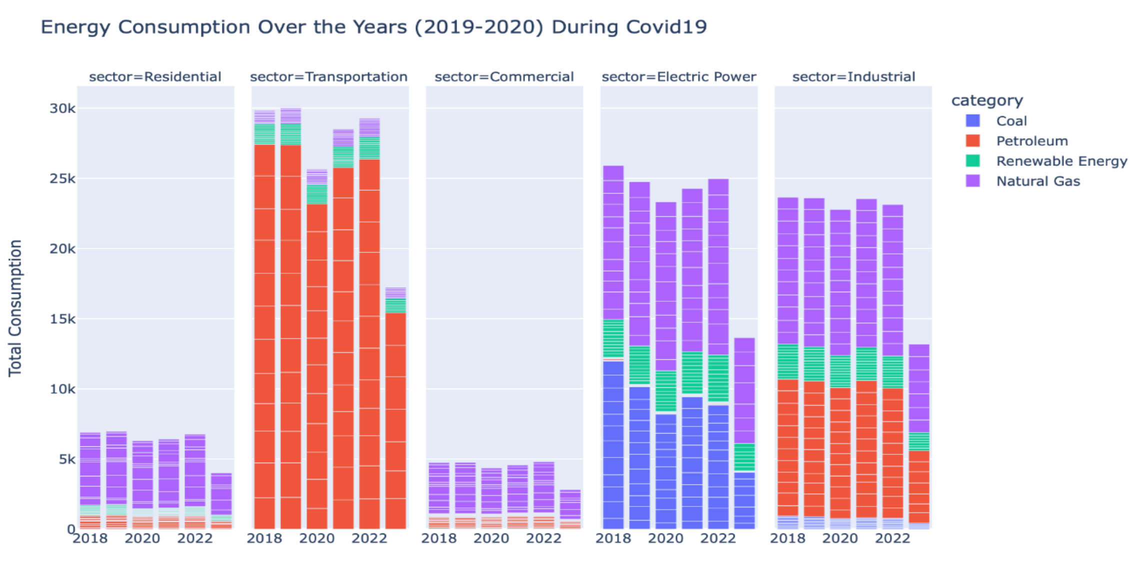
After the dissolution of the Soviet Union, we anticipated shifts in energy consumption as the U.S. no longer allocated energy resources toward Cold War offensive measures. During this period, the Iraq War also could have influence on energy consumption. We can see there are fluctuations across different energy sectors, but there is no consistent pattern which can lead us to make any significant inferences. This is visualized in Figure 4. The SQL code for queries and Python code for visualizations can be found in Appendix 3.3.

Total Consumption During COVID-19: Figure 5



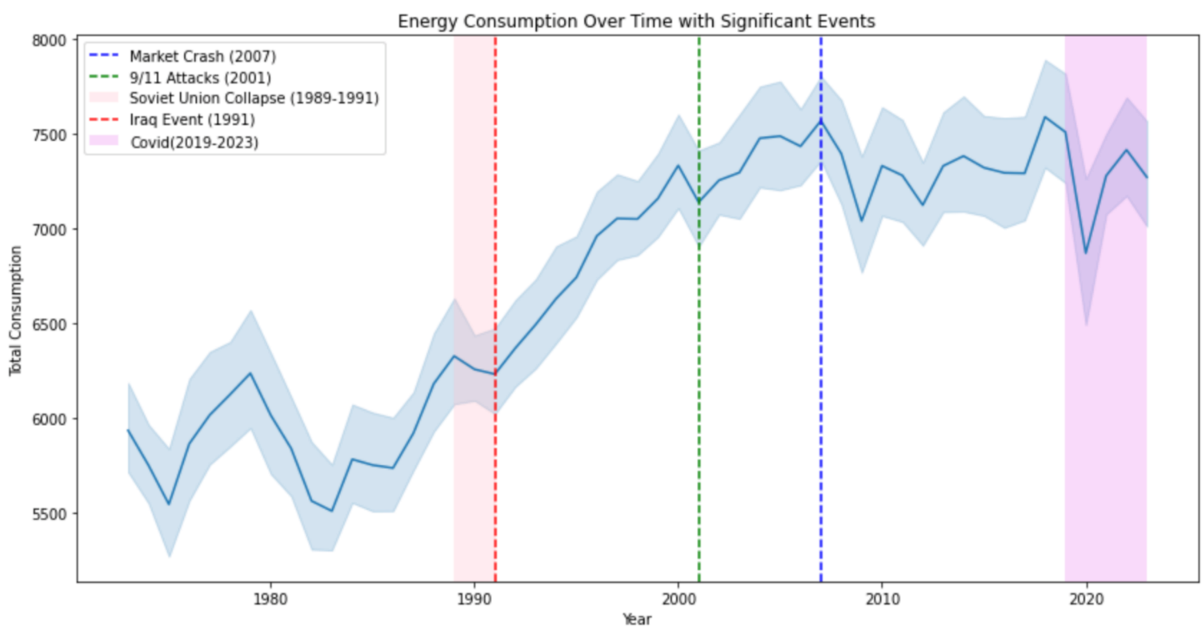
The final event for analysis is the COVID-19 pandemic. Our hypothesis posited a substantial decline in energy consumption across various sectors and sources during and following 2020, attributed to the lockdown measures in the U.S. Figure 5 illustrates this decline in total consumption.

Consumption by Sector During COVID-19: Figure 6



Furthermore, a sector-specific breakdown, as depicted in Figure 6, reveals a decrease across all sectors, notably pronounced in the Transportation sector. This decline can be attributed to altered commuting habits, reduced use of public transportation, and a change in travel patterns compared to the pre-pandemic era. The SQL code for queries and Python code for visualizations is provided in Appendix 3.4.

Total Consumption with Significant Events: Figure 7



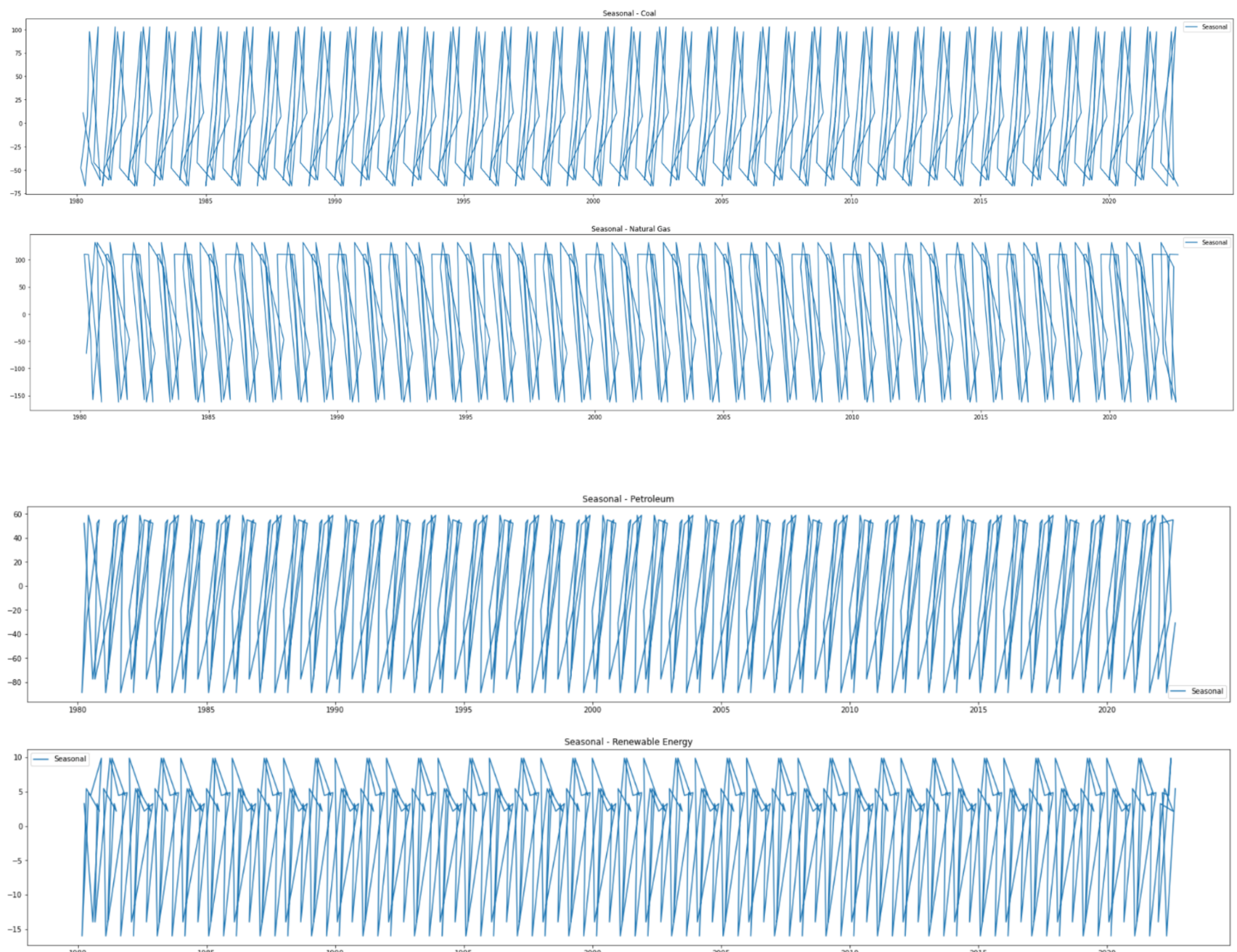
To summarize, Figure 7 displays a timeline of energy consumption and its fluctuations, highlighting relevant events. We see the same patterns which were outlined previously. Clearly, the most significant change is happening during the pandemic and post-pandemic.

Guiding Question 3 (Exploration):

Are there any noticeable seasonal patterns in energy consumption?

To analyze seasonal patterns, we started with conducting a seasonal time decomposition which subtracts general trends and residual noise from the data to unveil any seasonal patterns.

Seasonal Decomposition: Figure 8



The decomposition plots in Figure 8 reveal patterns for each of the energy sources, however, the exact patterns are difficult to interpret.

To enhance readability, a query was executed to rank energy consumption by energy source and season, yielding the output table presented in Figure 9.

Seasonality Patterns: Figure 9

Category	Season	Consumption (TBTU)
Coal	Summer	219636.198518
Coal	Winter	207605.013393
Coal	Fall	193692.686485
Coal	Spring	184214.132070
Natural Gas	Winter	379867.513385
Natural Gas	Spring	292078.144537
Natural Gas	Fall	259995.063646
Natural Gas	Summer	253307.614454
Petroleum	Summer	468957.658931
Petroleum	Winter	463655.026386
Petroleum	Spring	459572.515129
Petroleum	Fall	454626.354410
Renewable Energy	Spring	61633.460044
Renewable Energy	Summer	60632.128963
Renewable Energy	Winter	59334.573019
Renewable Energy	Fall	56650.404941

Examination of the table reveals that winter and summer emerge as the predominant seasons for coal and petroleum consumption, potentially attributed to heightened demand for heating and air conditioning. Notably, winter significantly dominates as the peak season for natural gas, possibly reflecting its prevalent use in U.S. households for heating purposes. In contrast, renewable energy exhibits a peak in consumption during spring and summer, which can be attributed to favorable conditions for production, such as extended daylight hours for solar energy and increased water runoff for hydroelectric energy. The SQL code for queries and Python code for decomposition are provided in Appendix 4.1.

Guiding Question 4 (Exploration):

How has the contribution of each energy source to the total energy mix changed over time?

In this segment of the report, we analyze the dynamics of each energy source's contribution to the overall energy consumption mix and the evolving trends over time. Through this, we aim to unravel the shifting landscape of energy consumption, offering valuable insights into the changing preferences, technological advancements, and environmental considerations that shape the U.S.' energy profile.

Total Energy Consumption Mix: Figure 10

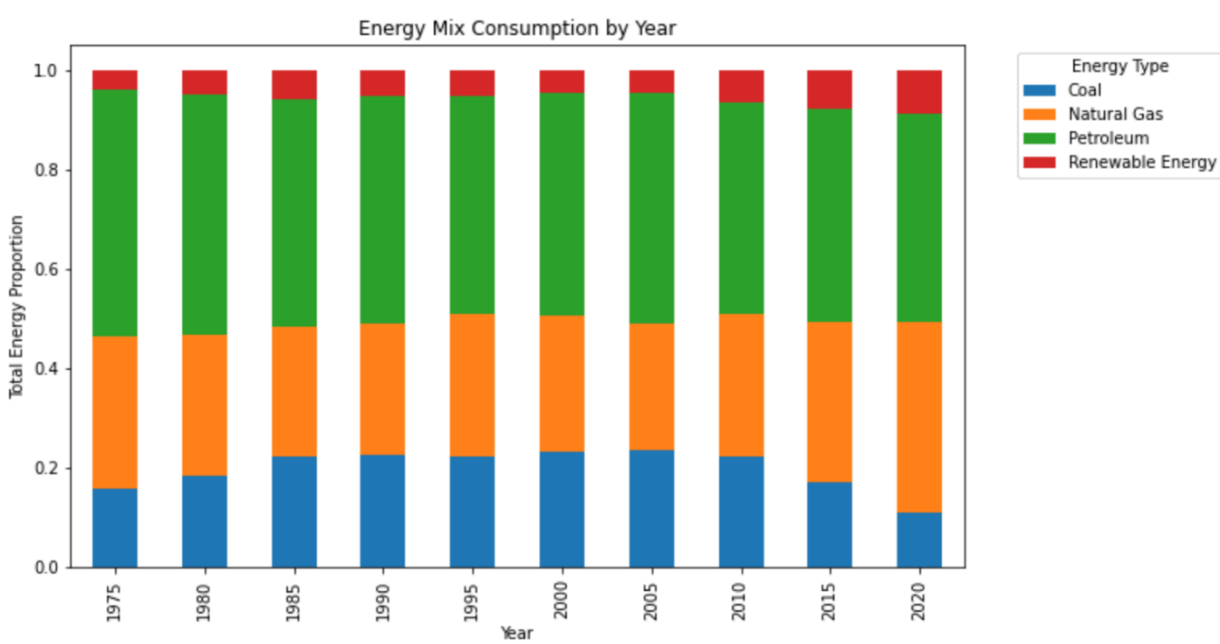


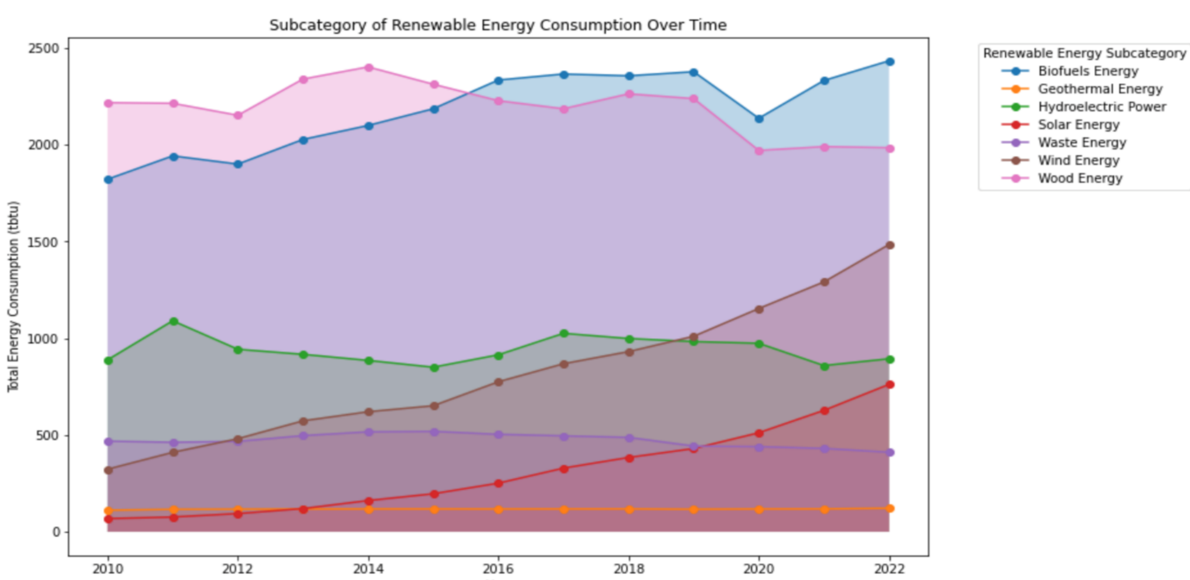
Figure 10 reveals subtle shifts in the energy landscape from 1975 to 2020. The consumption of coal and petroleum exhibits a gradual decline, while natural gas and renewable energy progressively claim a larger share of the overall energy mix. This transformation underscores the U.S.' evolving energy profile, demonstrating a pronounced shift toward more sustainable energy alternatives. Notably, the most substantial changes occurred in the five-year span from 2015 to 2020. The code for the SQL queries and Python code for the visualizations are provided in Appendix 5.1.

Guiding Question 5 (Exploration):

Are there any subcategories in renewable energy that have experienced rapid adoption?

This segment delves into the dynamic landscape of renewable energy, shedding light on key areas where advancements have propelled the U.S. towards a more sustainable and diversified energy portfolio.

Breakdown of Renewable Energy Consumption: Figure 11



In Figure 11, with the increasing trends apparent in Solar, Wind and Biofuels energy, we can infer that the U.S. is slowly attempting to diversify its energy reliance. Biofuel appears to be the primary renewable energy source with rapid adoption. This displays how the U.S. is trying to move towards more reliable energy sources although it may not be at the pace that is expected. Biofuels, which are produced from renewable feedstocks, are primarily used for transportation fuel. Certain policies were put in place to incentivize the use of biofuels over fossil fuels for transportation which explains the spike in biofuel consumption in recent years (EIA, 2022). The SQL code for the queries and Python code for visualizations are provided in Appendix 6.1.

Discussion

In extending this project, a comprehensive examination of energy consumption by state could provide a more granular understanding of regional disparities and variations in energy demands. Additionally, incorporating temperature data would offer valuable insights into the impact of climate on energy consumption patterns, enabling a more nuanced analysis of seasonal variations.

A deeper exploration into the technological developments of each energy source could also contribute to a richer narrative, uncovering the driving forces behind shifts in efficiency and sustainability. Additionally, further research into energy policies would provide context to observe and rationalize trends. Understanding these policies can help us explain the impact of policy decisions on energy consumption patterns, offering valuable information for future projections and policy recommendations.

It is also important to acknowledge the limitations of our study. The absence of certain datasets, such as nuclear energy, hampers a comprehensive understanding of the U.S. energy consumption mix. Our analysis was constrained by the availability of datasets providing total consumption data, limiting the depth of our insights.

Furthermore, energy consumption is inherently tied to economic activities, and our study did not capture the economic factors influencing consumption within the U.S. The intricate relationship between business activities, economic health, and energy demands remains a complex terrain that warrants more in-depth exploration beyond the scope of our current dataset. Recognizing these limitations, future research could strive to integrate additional datasets and incorporate a more nuanced economic perspective to further enhance the robustness of our findings.

Conclusion

In conclusion, our exploration into energy consumption in the U.S. has uncovered a variety of trends, shifts, and significant implications for the country's energy landscape. Our study clearly depicts the shift to more sustainable energy with natural gas and renewable energy gradually assuming a more prominent role. From the historical context of the Cold War to the profound impact of global events like the COVID-19 pandemic, our study has navigated through critical junctures, demonstrating the relationships between external forces and energy consumption patterns.

The analysis of each energy source's contribution to the total energy mix over time underscores a notable path towards sustainability, with natural gas and renewable energy consumption taking over coal and petroleum energy consumption. These findings not only signify a shift in energy preferences but also align with broader global efforts to rely on more environmentally sustainable energy sources.

While our study provides valuable insights, it is crucial to acknowledge its limitations. The absence of certain datasets and a comprehensive economic perspective necessitate continued exploration to refine our understanding of the intricate factors influencing energy consumption. Recognizing this, future iterations of this research could delve deeper into state-level analyses, consider the impact of temperature changes, and explore the technological and policy landscapes in greater detail.

Our navigation through historically significant events and evolving energy profiles serves as a testament to the intricate relationship between socioeconomic conditions and the pursuit of an energy future that balances efficiency and sustainability. Through continued inquiry and adaptive methodologies, we can further illuminate the path towards a more sustainable and conscientious energy landscape for the United States.

References

Star schema olap cube | kimball dimensional modeling techniques. (n.d.). Kimball Group. Retrieved 4 November 2023, from https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/star-schema-olap-cube/

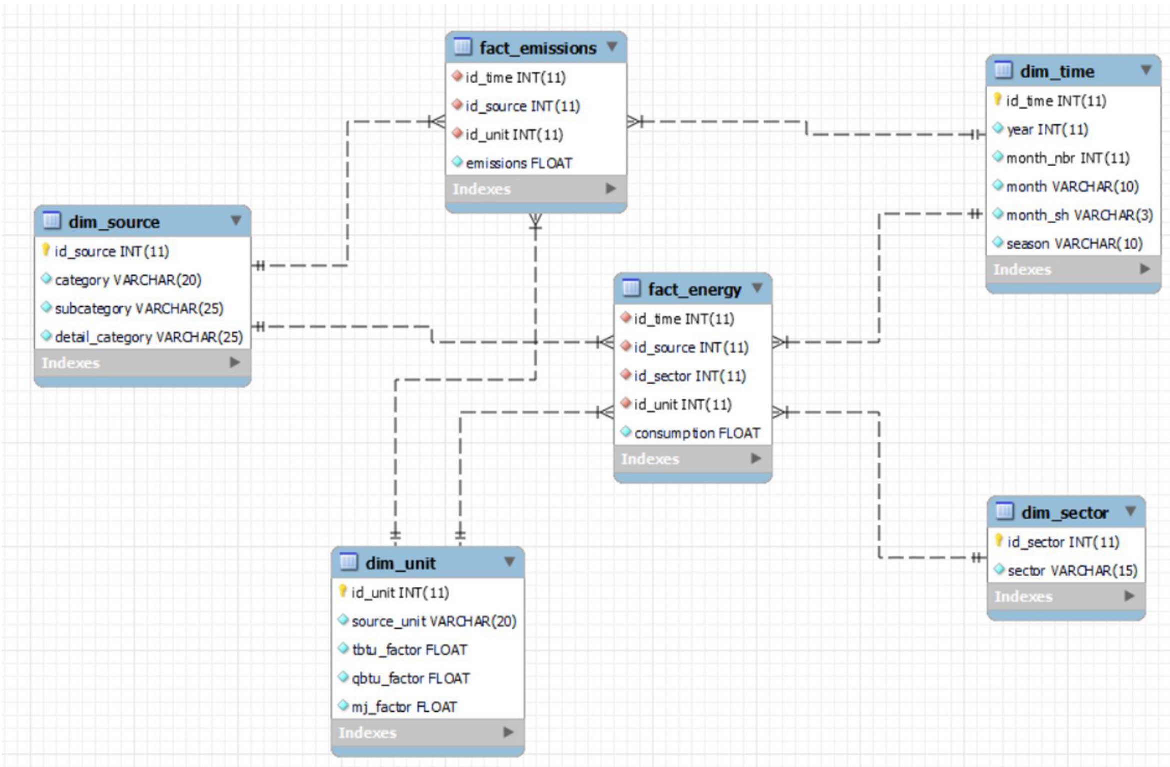
U.S. Energy Information Administration. (2021). U.S. Energy Information Administration – EIA – independent statistics and analysis. Coal and the environment – U.S. Energy Information Information Administration (EIA).

https://www.eia.gov/energyexplained/coal/coal-and-the-environment.php#:~:text=Several%20principal%20emissions%20result%20from,respiratory%20illnesses%20and%20lung%20disease

U.S. Energy Information Administration. (2021). *U.S. Energy Information Administration - EIA - independent statistics and analysis*. International - U.S. Energy Information Administration (EIA). https://www.eia.gov/international/overview/country/USA

Appendix 1

1.



2.

Coal_by_Sector:

The following table breaks down the usage of Coal in Tons by sector of industry, the following table describes its contents:

Column Name/Description	**Unit**
Month	NaN
Coal Consumed by the Residential Sector	Thousand Short Tons
Coal Consumed by the Commercial Sector, CHP	Thousand Short Tons
Coal Consumed by the Commercial Sector, Other	Thousand Short Tons
Coal Consumed by the Commercial Sector, Total	Thousand Short Tons
Coal Consumed by the Industrial Sector, Coke Plants	Thousand Short Tons

Coal Consumed by the Other Industrial Sector, CHP	Thousand Short Tons
Coal Consumed by the Other Industrial Sector, Non-CHP	Thousand Short Tons
Coal Consumed by the Other Industrial Sector, Total	Thousand Short Tons
Coal Consumed by the Industrial Sector, Total	Thousand Short Tons
Coal Consumed by the Transportation Sector	Thousand Short Tons
Coal Consumed by the Electric Power Sector	Thousand Short Tons
Coal Consumption, Total	Thousand Short Tons

Natural_Gas_by_Sector:

The following table breaks down the consumption of natural gas by sector, along with the unit of measurement descriptions:

Column Name/Description	Unit
Month	NaN
Natural gas Consumed by the Residential Sector	Billion Cubic Feet
Natural Gas Consumed by the Commercial Sector	Billion Cubic Feet
Natural Gas Consumed by the Industrial Sector, Lease and Plant Fuel	Billion Cubic Feet
Natural Gas Consumed by the Other Industrial Sector, CHP	Billion Cubic Feet
Natural Gas Consumed by the Other Industrial Sector, Non-CHP	Billion Cubic Feet
Natural Gas Consumed by the Other Industrial Sector, Total	Billion Cubic Feet
Natural Gas Consumed by the Industrial Sector, Total	Billion Cubic Feet
Natural Gas Consumed by the Transportation Sector, Pipelines and Distribution	Billion Cubic Feet
Natural Gas Consumed by the Transportation Sector, Vehicle Fuel	Billion Cubic Feet
Natural Gas Consumed by the Transportation Sector, Total	Billion Cubic Feet

Natural Gas Consumed by the Electric Power Sector	Billion Cubic Feet
Natural Gas Consumption, Total	Billion Cubic Feet

Petroleum_Industrial:

The following table breaks down the consumption of petroleum by type and sector specifically for the industrial sector, along with the unit of measurement descriptions:

Column Name/Description	Unit
Month	NaN
Asphalt and Road Oil Consumed by the Industrial Sector	Trillion Btu
Distillate Fuel Oil Consumed by the Industrial Sector	Trillion Btu
Propane Consumed by the Industrial Sector	Trillion Btu
Propylene Consumed by the Industrial Sector	Trillion Btu
Propane/Propylene Consumed by the Industrial Sector	Trillion Btu
Total Hydrocarbon Gas Liquids Consumed by the Industrial Sector	Trillion Btu
Kerosene Consumed by the Industrial Sector	Trillion Btu
Lubricants Consumed by the Industrial Sector	Trillion Btu
Motor Gasoline Consumed by the Industrial Sector	Trillion Btu
Petroleum Coke Consumed by the Industrial Sector	Trillion Btu
Residual Fuel Oil Consumed by the Industrial Sector	Trillion Btu
Other Petroleum Products Consumed by the Industrial Sector	Trillion Btu
Total Petroleum Consumed by the Industrial Sector	Trillion Btu

Petroleum_Residential_Commercial:

The following table breaks down the consumption of petroleum by type and sector specifically for the residential and commercial sectors, along with the unit of measurement descriptions:

Column Name/Description	Unit
Month	NaN
Distillate Fuel Oil Consumed by the Residential Sector	Trillion Btu
Propane Consumed by the Residential Sector	Trillion Btu
Kerosene Consumed by the Residential Sector	Trillion Btu
Total Petroleum Consumed by the Residential Sector	Trillion Btu
Distillate Fuel Oil Consumed by the Commercial Sector	Trillion Btu
Propane Consumed by the Commercial Sector	Trillion Btu
Kerosene Consumed by the Commercial Sector	Trillion Btu
Motor Gasoline Consumed by the Commercial Sector	Trillion Btu
Petroleum Coke Consumed by the Commercial Sector	Trillion Btu
Residual Fuel Oil Consumed by the Commercial Sector	Trillion Btu
Total Petroleum Consumed by the Commercial Sector	Trillion Btu

Petroleum_Transportation_Electric_Power:

The following table breaks down the consumption of petroleum by type and sector specifically for the electric power and the transportation sector, along with the unit of measurement descriptions:

Column Name/Description	Unit

Month	NaN
Aviation Gasoline Consumed by the Transportation Sector	Trillion Btu
Distillate Fuel Oil Consumed by the Transportation Sector	Trillion Btu
Propane Consumed by the Transportation Sector	Trillion Btu
Jet Fuel Consumed by the Transportation Sector	Trillion Btu
Lubricants Consumed by the Transportation Sector	Trillion Btu
Motor Gasoline Consumed by the Transportation Sector	Trillion Btu
Residual Fuel Oil Consumed by the Transportation Sector	Trillion Btu
Other Products Consumed by the Transportation Sector	Trillion Btu
Total Petroleum Consumed by the Transportation Sector	Trillion Btu
Distillate Fuel Oil Consumed by the Electric Power Sector	Trillion Btu
Petroleum Coke Consumed by the Electric Power Sector	Trillion Btu
Residual Fuel Oil Consumed by the Electric Power Sector	Trillion Btu
Total Petroleum Consumed by the Electric Power Sector	Trillion Btu

Renewable_Energy_Industrial:

The following table breaks down the consumption of renewable energy in the industrial sector:

Column Name/Description	Unit
Month	NaN
Hydroelectric Power Consumed by the Industrial Sector	Trillion Btu
Geothermal Energy Consumed by the Industrial Sector	Trillion Btu
Solar Energy Consumed by the Industrial Sector	Trillion Btu

Wind Energy Consumed by the Industrial Sector	Trillion Btu
Wood Energy Consumed by the Industrial Sector	Trillion Btu
Waste Energy Consumed by the Industrial Sector	Trillion Btu
Fuel Ethanol, Excluding Denaturant, Consumed by the Industrial Sector	Trillion Btu
Biomass Losses and Co-products in the Industrial Sector	Trillion Btu
Biomass Energy Consumed by the Industrial Sector	Trillion Btu
Total Renewable Energy Consumed by the Industrial Sector	Trillion Btu

Renewable_Energy_Residential_Commercial:

The following table breaks down renewable energy consumption by sector specifically for the residential and commercial sectors:

Column Name/Description	Unit
Month	NaN
Geothermal Energy Consumed by the Residential Sector	Trillion Btu
Solar Energy Consumed by the Residential Sector	Trillion Btu
Wood Energy Consumed by the Residential Sector	Trillion Btu
Total Renewable Energy Consumed by the Residential Sector	Trillion Btu
Hydroelectric Power Consumed by the Commercial Sector	Trillion Btu
Geothermal Energy Consumed by the Commercial Sector	Trillion Btu
Solar Energy Consumed by the Commercial Sector	Trillion Btu
Wind Energy Consumed by the Commercial Sector	Trillion Btu
Wood Energy Consumed by the Commercial Sector	Trillion Btu

Waste Energy Consumed by the Commercial Sector	Trillion Btu
Fuel Ethanol, Excluding Denaturant, Consumed by the Commercial Sector	Trillion Btu
Biomass Energy Consumed by the Commercial Sector	Trillion Btu
Total Renewable Energy Consumed by the Commercial Sector	Trillion Btu

Renewable_Energy_Transportation_Electric_Power.:

The following table displays the energy consumed by the transportation and electric power sectors:

Column Name/Description	Unit
Month	NaN
Fuel Ethanol, Excluding Denaturant, Consumed by the Transportation Sector	Trillion Btu
Biodiesel Consumed by the Transportation Sector	Trillion Btu
Renewable Diesel Fuel Consumed by the Transportation Sector	Trillion Btu
Other Biofuels Consumed by the Transportation Sector	Trillion Btu
Biomass Energy Consumed by the Transportation Sector	Trillion Btu
Conventional Hydroelectric Power Consumed by the Electric Power Sector	Trillion Btu
Geothermal Energy Consumed by the Electric Power Sector	Trillion Btu
Solar Energy Consumed by the Electric Power Sector	Trillion Btu
Wind Energy Consumed by the Electric Power Sector	Trillion Btu
Wood Energy Consumed by the Electric Power Sector	Trillion Btu
Waste Energy Consumed by the Electric Power Sector	Trillion Btu
Biomass Energy Consumed by the Electric Power Sector	Trillion Btu

Total Renewable Energy Consumed by the Electric Power Sector	Trillion Btu

Primary_Energy_Consumption_by_Source:

The following table describes the energy consumption by source of energy along with the unit descriptions:

Column Name/Description	Unit
Month	NaN
Coal Consumption	Quadrillion Btu
Natural Gas Consumption (Excluding Supplemental Gaseous Fuels)	Quadrillion Btu
Petroleum Consumption (Excluding Biofuels)	Quadrillion Btu
Total Fossil Fuels Consumption	Quadrillion Btu
Nuclear Electric Power Consumption	Quadrillion Btu
Hydroelectric Power Consumption	Quadrillion Btu
Geothermal Energy Consumption	Quadrillion Btu
Solar Energy Consumption	Quadrillion Btu
Wind Energy Consumption	Quadrillion Btu
Biomass Energy Consumption	Quadrillion Btu
Total Renewable Energy Consumption	Quadrillion Btu
Total Primary Energy Consumption	Quadrillion Btu

Appendix 2

```
import seaborn as sns
import matplotlib.pyplot as plt

#For Coal.

query_table = pd.read_sql_query(
'''
SELECT
    t.year,
    so.category,
    SUM(consumption * u.tbtu_factor)  AS trilbtu,
    SUM(emissions)  AS em
FROM
    fact_energy f
INNER JOIN
    dim_time t ON f.id_time = t.id_time
INNER JOIN
    dim_source so ON f.id_source = so.id_source
INNER JOIN
    dim_unit u ON f.id_unit = u.id_unit
INNER JOIN
    fact_emissions e ON e.id_time = t.id_time AND e.id_source = so.id_source
WHERE
    so.category NOT IN ('Natural Gas', 'Petroleum') AND t.year < 2023
GROUP BY
    t.year,
    so.category;
'''
, engine)
```

The above displays a query for both figures 2.1 and 2.2, it gathers information on the emissions and consumption for the coal sector.

```
plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['trilbtu'], marker='o', label='BTUs')
plt.plot(query_table['year'], query_table['em'], marker='o', label='C02 Emmissions')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Coal Emmissions in Million of Tons of C02 and Trillion BTUs of Energy per Year')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.1.

```
query_table['difference'] = query_table['em'] /query_table['trilbtu']

plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['difference'], marker='o', label='Emmissions Divided by Ten Trillion BTU')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Ratio of C02 in Millon Tons to Ten Trillion BTUs for Petrol')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.2.

```
#For Petrolium.
query_table = pd.read_sql_query(
'''
SELECT
    t.year,
    so.category,
    SUM(consumption * u.tbtu_factor) / 10 AS tentrilbtu,
    SUM(emissions) AS em
FROM
    fact_energy f
INNER JOIN
    dim_time t ON f.id_time = t.id_time
INNER JOIN
    dim_source so ON f.id_source = so.id_source
INNER JOIN
    dim_unit u ON f.id_unit = u.id_unit
INNER JOIN
    fact_emissions e ON e.id_time = t.id_time AND e.id_source = so.id_source
WHERE
    so.category NOT IN ('Coal', 'Natural Gas') AND t.year < 2023
GROUP BY
    t.year,
    so.category;
'''
, engine)
```

The above displays a query for both figures 2.3 and 2.4, it gathers information on the emissions and consumption for the natural gas sector.

```
plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['tentrilbtu'], marker='o', label='BTUs')
plt.plot(query_table['year'], query_table['em'], marker='o', label='C02 Emmissions')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Petroleum Emmissions in Millon Tons and Ten Trillion BTUs of Energy per Year')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.3.

```
query_table['difference'] = query_table['em'] /query_table['tentrilbtu']

plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['difference'], marker='o', label='Ratio by Year')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Ratio of C02 in Millon Tons to Ten Trillion BTUs for Petrol')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.4.

```
#For Natrual Gas.
query_table = pd.read_sql_query(
'''
SELECT
    t.year,
    so.category,
    SUM(consumption * u.tbtu_factor) / 10 AS trilbtu,
    SUM(emissions) AS em
FROM
    fact_energy f
INNER JOIN
    dim_time t ON f.id_time = t.id_time
INNER JOIN
    dim_source so ON f.id_source = so.id_source
INNER JOIN
    dim_unit u ON f.id_unit = u.id_unit
INNER JOIN
    fact_emissions e ON e.id_time = t.id_time AND e.id_source = so.id_source
WHERE
    so.category NOT IN ('Coal', 'Petroleum') AND t.year < 2023
GROUP BY
    t.year,
    so.category;
'''
, engine)
```

The above displays a query for both figures 2.5 and 2.6, it gathers information on the emissions and consumption for the petroleum sector.

```python
plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['trilbtu'], marker='o', label='BTUs')
plt.plot(query_table['year'], query_table['em'], marker='o', label='C02 Emmissions')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Natural Gas Emmissions in Millon Tons and Ten Trillion BTUs of Energy per Year')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.5.

```python
query_table['difference'] = query_table['em'] /query_table['trilbtu']

plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['difference'], marker='o', label='Ratio by Year')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Ratio of C02 in Millon Tons to Ten Trillion BTUs for Natural Gas')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.6.

```python
query_table = pd.read_sql_query(
    '''
SELECT
    t.year,
    SUM(consumption * u.tbtu_factor)  AS trilbtu,
    SUM(emissions)  AS em
FROM
    fact_energy f
INNER JOIN
    dim_time t ON f.id_time = t.id_time
INNER JOIN
    dim_source so ON f.id_source = so.id_source
INNER JOIN
    dim_unit u ON f.id_unit = u.id_unit
INNER JOIN
    fact_emissions e ON e.id_time = t.id_time AND e.id_source = so.id_source
Where t.year < 2023

GROUP BY
    t.year;
'''
,engine)
```

The above displays a query for both figures 2.7 and 2.8, it gathers information on the emissions and consumption totals.

```
plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['trilbtu'], marker='o', label='BTUs')
plt.plot(query_table['year'], query_table['em'], marker='o', label='C02 Emmissions')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Total Emmissions in Millon Tons and Ten Trillion BTUs of Energy per Year')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.7.

```
query_table['difference'] = query_table['em'] /query_table['trilbtu']

plt.figure(figsize=(10, 6))
plt.plot(query_table['year'], query_table['difference'], marker='o', label='Ratio by Year')

plt.xlabel('Year')
plt.ylabel('Values')
plt.title('Ratio of C02 in Millon Tons to Ten Trillion BTUs Total')

plt.legend()

plt.show()
```

The above shows the python code for creating figure 2.8.

Appendix 3

3.1

```
query_marketcrash= pd.read_sql_query(
    '''
select
    t.year,
    t.month,
    so.category as `energy source`,
    s.sector,
    sum(consumption*u.tbtu_factor) as `total consumption`
from
    fact_energy f
inner join
    dim_time t
    on f.id_time = t.id_time
inner join
    dim_source so
    on f.id_source = so.id_source
inner join
    dim_unit u
    on f.id_unit=u.id_unit
inner join
    dim_sector s
    on f.id_sector = s.id_sector
where year between "2004" and "2010"
group by
    t.year,
    t.month,
    so.category,
    s.sector
order by year, month ,`total consumption`
'''
, engine)
print(query_marketcrash)
```

```
import plotly.express as px

# Assuming your DataFrame is named query_marketcrash
fig = px.bar(query_marketcrash,
             x='year',
             y='total consumption',
             color='energy source',
             facet_col='sector',
             labels={'total consumption': 'Total Consumption (TBTU)'},
             title='Energy Consumption Over the Years (2004-2010) During the Housing Market Crash',
             height=600,
            )

fig.update_xaxes(categoryorder='total descending')
fig.update_layout(yaxis=dict(title='Total Consumption'))

fig.show()
```

3.2

```
query_911= pd.read_sql_query(
    '''
select
    t.year,
    t.month,
    so.category as `energy source`,
    so.subcategory as `sub source`,
    s.sector,
    sum(consumption*u.tbtu_factor) as `total consumption`
from
    fact_energy f
inner join
    dim_time t
    on f.id_time = t.id_time
inner join
    dim_source so
    on f.id_source = so.id_source
inner join
    dim_unit u
    on f.id_unit=u.id_unit
inner join
    dim_sector s
    on f.id_sector = s.id_sector
where year between "1999" and "2004" and  so.subcategory in ("Jet Fuel")
group by
    t.year,
    t.month,
    so.category,
    so.subcategory,
    s.sector
order by year, month ,`total consumption`
'''
, engine)
print(query_911)
```

```
import seaborn as sns
import matplotlib.pyplot as plt

# Assuming 'query_marketcrash' is a DataFrame containing the results
# You may need to adjust column names based on your actual DataFrame structure
years = [1999, 2000, 2001, 2002, 2003, 2004]

# Filter the DataFrame to include only the years of interest
filtered_data = query_911[query_911['year'].isin(years)]

plt.figure(figsize=(12, 6))
sns.lineplot(x='year', y='total consumption', data=filtered_data, marker='o', linestyle='-')
plt.xlabel('Year')
plt.ylabel('Total Consumption')
plt.title('Total Consumption Over the Years (1999 - 2004) around 9/11')
plt.show()
```

```
import plotly.express as px

fig = px.bar(query_911,
             x='year',
             y='total consumption',
             color='energy source',
             facet_col='sub source',
             labels={'total consumption': 'Total Consumption (TBTU)'},
             title='Energy Consumption in 2001 of Jet Fuel During 9/11',
             height=600,
            )

fig.update_xaxes(categoryorder='total descending')
fig.update_layout(yaxis=dict(title='Total Consumption'))

fig.show()
```

3.3

```
query_sovietunion= pd.read_sql_query(
    '''
select
    t.year,
    t.month,
    so.category as `energy source`,
    s.sector,
    sum(consumption*u.tbtu_factor) as `total consumption`
from
    fact_energy f
inner join
    dim_time t
    on f.id_time = t.id_time
inner join
    dim_source so
    on f.id_source = so.id_source
inner join
    dim_unit u
    on f.id_unit=u.id_unit
inner join
    dim_sector s
    on f.id_sector = s.id_sector
where year between "1989" and "1991"
group by
    t.year,
    t.month,
    so.category,
    s.sector
order by year, month ,`total consumption`
'''
, engine)
print(query_sovietunion)
```

```python
import plotly.express as px

fig = px.bar(query_sovietunion,
             x='year',
             y='total consumption',
             color='energy source',
             facet_col='sector',
             labels={'total consumption': 'Total Consumption'},
             title='Energy Consumption Over the Years (1989-1991) During the Collapse of Soviet Union',
             height=600,
            )

fig.update_xaxes(categoryorder='total descending')
fig.update_layout(yaxis=dict(title='Total Consumption'))

fig.show()
```

3.4

```python
query_covid= pd.read_sql_query(
    '''
select
    t.year,
    t.month,
    so.category,
    s.sector,
    sum(consumption*u.tbtu_factor) as `total consumption`
from
    fact_energy f
inner join
    dim_time t
    on f.id_time = t.id_time
inner join
    dim_source so
    on f.id_source = so.id_source
inner join
    dim_unit u
    on f.id_unit=u.id_unit
inner join
    dim_sector s
    on f.id_sector = s.id_sector
where year between "2018" and "2023"
group by
    t.year,
    t.month,
    so.category,
    s.sector
order by year, month ,`total consumption`
'''
, engine)
print(query_covid)
```

```python
import seaborn as sns
import matplotlib.pyplot as plt

years = [2018,2019,2020,2021,2022,2023]

filtered_data = query_covid[query_covid['year'].isin(years)]

plt.figure(figsize=(12, 6))
sns.lineplot(x='year', y='total consumption', data=filtered_data, marker='o', linestyle='-')
plt.xlabel('Year')
plt.ylabel('Total Consumption')
plt.title('Total Consumption Over the Years During Covid19')
plt.show()
```

```python
import plotly.express as px

fig = px.bar(query_covid,
             x='year',
             y='total consumption',
             color='category',
             facet_col='sector',
             labels={'total consumption': 'Total Consumption (TBTU)'},
             title='Energy Consumption Over the Years (2019-2020) During Covid19',
             height=600,
            )

fig.update_xaxes(categoryorder='total descending')
fig.update_layout(yaxis=dict(title='Total Consumption'))

fig.show()
```

```python
query_overall = pd.read_sql_query(
    '''
    SELECT
        t.year,
        t.month,
        SUM(consumption * u.tbtu_factor) AS `total consumption`
    FROM
        fact_energy f
    INNER JOIN
        dim_time t ON f.id_time = t.id_time
    INNER JOIN
        dim_unit u ON f.id_unit = u.id_unit
    GROUP BY
        t.year,
        t.month
    ORDER BY
        year,
        month,
        `total consumption`
    '''
, engine)

print(query_overall)
```

```
event_data = query_overall


# Highlight significant events
significant_events = {
    'Market Crash (2007)': 2007,
    '9/11 Attacks (2001)': 2001,
    'Soviet Union Collapse (1989-1991)': (1989, 1991),  # Specify the range
    'Iraq Event (1991)': 1991,
    'Covid(2019-2023)':(2019,2023)
}

# Define a color palette for each event
event_palette = {
    'Market Crash (2007)': 'blue',
    '9/11 Attacks (2001)': 'green',
    'Soviet Union Collapse (1989-1991)': 'pink',
    'Iraq Event (1991)': 'red',
    'Covid(2019-2023)':'violet'
}

# Create line plots with markers for significant events using the specified color palette
plt.figure(figsize=(14,7))
sns.lineplot(x='year', y='total consumption', markers=True, data=event_data)

# Highlight significant events with vertical lines or shaded areas
for event, year in significant_events.items():
    color = event_palette[event]
    if isinstance(year, tuple):  # Check if it's a range
        plt.axvspan(year[0], year[1], facecolor=color, alpha=0.3, label=event)
    else:
        plt.axvline(x=year, color=color, linestyle='--', label=event)

plt.title('Energy Consumption Over Time with Significant Events')
plt.xlabel('Year')
plt.ylabel('Total Consumption')
plt.legend()
plt.show()
```

Appendix 4

4.1

```
start_date = '1980-03-01'
end_date = '2022-09-30'

# Filter the DataFrame for the specified time range
query_table = query_table[(query_table['date'] >= start_date) & (query_table['date'] <= end_date)]

sectors = query_table['category'].unique()
sectors1 = [sectors[0],sectors[1]]


for sector in sectors1:
    sector_data = query_table[query_table['category'] == sector]
    sector_data = sector_data.set_index('date')['tbtu']

    result = seasonal_decompose(sector_data, model='additive', period=8)


    plt.figure(figsize=(30, 5))

    plt.plot(result.seasonal, label='Seasonal')
    plt.title(f'Seasonal - {sector}')
    plt.legend()
```

```
sectors2 = [sectors[2],sectors[3]]


for sector in sectors2:
    sector_data = query_table[query_table['category'] == sector]
    sector_data = sector_data.set_index('date')['tbtu']

    result = seasonal_decompose(sector_data, model='additive', period=8)


    plt.figure(figsize=(30, 5))

    plt.plot(result.seasonal, label='Seasonal')
    plt.title(f'Seasonal - {sector}')
    plt.legend()
```

```
seasonal_query = pd.read_sql_query('''select

    so.category,
    t.season,
    row_number() over (partition by category order by sum(consumption*u.tbtu_factor) desc ) as id_row,
    sum(consumption*u.tbtu_factor) as tbtu
from
    fact_energy f
inner join
    dim_time t
    on f.id_time = t.id_time
inner join
    dim_sector s
    on f.id_sector = s.id_sector
inner join
    dim_source so
    on f.id_source = so.id_source
inner join
    dim_unit u
    on f.id_unit=u.id_unit
group by

    so.category,
    t.season
    order by so.category,id_row

    ''', engine)
print(seasonal_query)
```

Appendix 5

5.1

```
query_BP= pd.read_sql_query(

    '''

select
    t.year,
    so.category,
    sum(f.consumption * u.tbtu_factor) as total_consumption,
    sum(f.consumption * u.tbtu_factor) / (
        select sum(f1.consumption * u1.tbtu_factor)
        from fact_energy f1
        inner join
dim_time t1
on f1.id_time = t1.id_time
        inner join
dim_unit u1
on f1.id_unit = u1.id_unit
        where t1.year = t.year
) as proportion
from
    fact_energy f
inner join
    dim_time t
on f.id_time = t.id_time
inner join
    dim_source so
on f.id_source = so.id_source
inner join
    dim_unit u
on f.id_unit = u.id_unit
where mod(t.year, 5) = 0
group by
    t.year,
    so.category
order by
    t.year, so.category desc LIMIT 64;


'''
```

```
ax = query_BP.pivot_table(index='year', columns='category', values='proportion', aggfunc='sum').plot(kind='bar', stack

ax.set_ylabel('Total Energy Proportion')
ax.set_xlabel('Year')
ax.set_title('Energy Mix Consumption by Year')
plt.legend(title='Energy Type', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.show()
```

Appendix 6

6.1

```
query_lp= pd.read_sql_query(

    '''
select
    t.year,
    so.detail_category,
    sum(consumption*u.tbtu_factor) as `total consumption`
from
    fact_energy f
inner join
    dim_time t
    on f.id_time = t.id_time
inner join
    dim_source so
    on f.id_source = so.id_source
inner join
    dim_unit u
    on f.id_unit=u.id_unit
where so.category = 'Renewable Energy'
group by
    t.year,
    so.category,
    so.detail_category
Having `total consumption` <> 0 AND t.year >= 2010 AND t.year <> 2023
order by year DESC ,`total consumption` limit 100 ;
'''
, engine)

print(query_lp)
```

```
# Pivot the data for plotting
df_pivot = query_lp.pivot_table(index='year', columns='detail_category', values='total consumption', aggfunc='sum')

# Plotting filled stacked lines
fig, ax = plt.subplots(figsize=(12, 8))

# Plot each line separately and fill the area between them
for column in df_pivot.columns:
    ax.plot(df_pivot.index, df_pivot[column], label=column, marker='o')
    ax.fill_between(df_pivot.index, 0, df_pivot[column], alpha=0.3)

ax.set_ylabel('Total Energy Consumption (tbtu)')
ax.set_xlabel('Year')
ax.set_title('Subcategory of Renewable Energy Consumption Over Time')
plt.legend(title='Renewable Energy Subcategory', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.show()
```