

Energy Generation trends from 2001-2021

Covering the Contiguous United States

February 20, 2024

Dan Rausch
Data Visualization 44-613
2024 Spring OP Block 1
Dr. Case



Image created by ChatGPT – Asked to generate a collage based on the primary energy generation types in the United States in 2021.

<u>SECTION 1. INTRODUCTION</u>	3
<u>SECTION 2. DATA DESCRIPTION</u>	3
<u>SECTION 3. DATA CLEANING STRATEGIES</u>	4
<u>SECTION 4. CLEAN DATASET</u>	5
<u>SECTION 5. VISUALIZATION TOOLS</u>	6
<u>SECTION 6. VISUALIZATIONS AND STORIES</u>	7
<i>GOAL 1: WHAT IS THE DISTRIBUTION OF ENERGY GENERATION TYPES IN THE US FROM 2001 TO 2021?</i>	7
<i>GOAL 2: WHAT IS THE PRIMARY GENERATION TYPE BY STATE IN 2001 AND HOW DOES THAT CHANGE IN 2021?</i>	8
<i>GOAL 3: HOW DOES POPULATION IMPACT TOTAL MWH'S GENERATED IN 2021?</i>	10
<i>GOAL 4: WHAT IS THE DISTRIBUTION OF ENERGY GENERATION BY US REGION IN 2021?</i>	11
<i>GOAL 5: WHAT IS THE DISTRIBUTION OF AVERAGE RETAIL PRICE FOR A KWH BY REGION?</i>	12
<i>GOAL 6: BASED ON 20 YEAR HISTORY, WHAT IS THE ENERGY GENERATION FORECAST FROM 2021 TO 2024?</i>	13
<u>SECTION 7. CONCLUSIONS</u>	14

Section 1. Introduction

This project looks at historical energy generation statistics by state and by type of power in order to help the reader draw conclusions about the overall progress of energy generation and which fuels are most prevalent. The concept of cost of generation and retail cost for consumer by state is also explored to help create a comprehensive view of US energy generation and consumption both relatively currently as well as historically. This report would be useful for state and federal lawmakers, energy investors, or certain energy advocates as it provides a comprehensive view of state and national trends in types of fuel consumed, forecasted generation trends, state populations, and regional retail pricing.

Section 2. Data Description

Data Domain	Description
Description	Energy and Population; Fuel generation, retail, population, and consumption by State for the contiguous United States
Data Source	Description
Source 1: Kaggle	https://www.kaggle.com/datasets/kevilmorgado/us-energy-generation-2001-2022
Source 2: Census.gov	https://www2.census.gov/programs-surveys/popest/tables/2020-2022/state/totals/NST-EST2022-POP.xlsx
Source 3: EIA State Archives	https://www.eia.gov/electricity/state/archive/2021/
	Combined these three sources into two Tableau data sources (below)
Column/Row Details	Description
Final Project Data	12 fields, 49 rows
Raw_Generation_Data	8 fields, 107,292 rows
	Data relationship established via State field
Column / Field Details	Description
Final Project Data	
State	Location - State within the United States
State Abbrv	Abbreviated Location used for combining data
Geographic Region	Associated Region per Census.gov
2021 Population (Est.)	Estimated Population in the State per Census.gov
2021 Primary	Type of Primary power generation in 2021
2021%	% of Primary power generation versus total in 2021
2021 Primary	2021 Primary power generation in MWh
2001%	2021 Percentage of Total Power generation
2001 Primary	2021 Type of Primary power generation

Retail Price Range	2021 Primary power generation in MWh
Avg Retail Price (cents/kWh)	Average Retail price of a kWh of electricity in cents in 2021
Net generation (MWh)	Total Net Generation (MWh) in 2021
Total Retail Sales (MWh)	Total Retail Sales of Electricity (MWh) in 2021
Raw_Generation_Data	
Year	Year of Energy Production
Month	Month of Energy Production
Date	Calculated field derived from Year and Month
State1	State where energy was produced
Type of Producer	Total Electric Power Industry
Energy Source	Type of Fuel used (e.g. Coal, Nuclear, Hydro, Wind)
Generation (Megawatthours)	Amount of Megawatthours produced (referred to as MWh)
Energy Source (group)	Grouping to combine low energy source production

Section 3. Data Cleaning Strategies

FinalProjectData Source:

- Normalize the State field. Some used abbreviations while others used state names.
- Add geographic area from Census.gov (https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf)
- Removed Hawaii and Alaska as this report focuses only on Contiguous United States

Raw_Generation_Data Source:

- Created Energy Source (Group) to combine little used power generation fuel methods
- Created Calculated field Date combining Year and Month
- Removed Negative Generation Values as this reflects facility use of generation per EIA
- Removed 2022 data as it only contained data through May 2022
- Removed Hawaii and Alaska as this report focuses only on Contiguous United States
- Removed US-TOTAL from State Field as it caused duplication of data
- Removed TOTAL from Energy Source Field as it caused duplication of data
- Retained only TYPE OF PRODUCER = “Total Electric Power Industry” as it represented the total volume of production for the state and type of producer.

Section 4. Clean Dataset

This is representative of the data used in the Tableau project. It reflects the above referenced data fields as well as the steps taken to clean the data.

FinalProjectData:

FinalProjectData State	FinalProjectData State Abrv	Abs FinalProjectData Geographic Region	# FinalProjectData 2021 Population (Est.)	Abs FinalProjectData 2021 Primary	# FinalProjectData 2021% 2001 Primary	Abs FinalProjectData 2001% 2001 Primary	# FinalProjectData Retail Price Range	Abs FinalProjectData Avg Retail Price (cents/...)	# FinalProjectData Net generation (MWh)	# FinalProjectData Total Retail Sales (MWh)	
Alabama	AL	South	5,049,846	Natural Gas	0.376249	Coal	0.575995	Medium	10.1800	142,733,330	85,585,166
Arizona	AZ	West	7,264,877	Natural Gas	0.443671	Coal	0.445567	Medium	10.7300	108,604,620	81,219,990
Arkansas	AR	South	3,028,122	Coal	0.355617	Coal	0.524760	Low	9.1000	61,100,068	48,663,142
California	CA	West	39,142,991	Natural Gas	0.490477	Natural Gas	0.563618	Medium	19.6500	197,165,106	247,249,865
Colorado	CO	West	5,811,297	Coal	0.419585	Coal	0.766939	Medium	10.9000	56,838,472	56,351,209
Connecticut	CT	Northeast	3,623,355	Natural Gas	0.586282	Nuclear	0.505984	Medium	18.3200	44,079,943	27,737,606
Delaware	DE	South	1,004,807	Natural Gas	0.87542	Coal	0.949809	Medium	10.5000	4,305,126	11,479,655
District of Columbia	DC	South	668,791	Natural Gas	0.625465	Petroleum	1,000000	Medium	12.8100	211,067	10,083,372
Florida	FL	South	21,828,069	Natural Gas	0.738958	Coal	0.357216	Medium	10.6700	246,450,375	241,562,082
Georgia	GA	South	10,788,029	Natural Gas	0.462849	Coal	0.630183	Medium	10.4300	124,200,528	137,363,952
Idaho	ID	West	1,904,214	Natural Gas	0.810472	Natural Gas	0.777780	Low	9.1700	16,836,475	25,286,616

Raw_Generation_Data

# RawGenerationData Year	# RawGenerationData Month	# Calculation Date	RawGenerationData State1	Abc RawGenerationData Type Of Producer	Abc RawGenerationData Energy Source	# RawGenerationData GENERATION (Megawat...	Group Energy Source (group)
2001	1	1/1/2001	AL	Total Electric Power Industry	Coal	6,557,913	Coal
2001	1	1/1/2001	AL	Total Electric Power Industry	Petroleum	107,497	Other
2001	1	1/1/2001	AL	Total Electric Power Industry	Natural Gas	566,478	Natural Gas
2001	1	1/1/2001	AL	Total Electric Power Industry	Other Gases	25,283	Other
2001	1	1/1/2001	AL	Total Electric Power Industry	Nuclear	2,940,300	Nuclear
2001	1	1/1/2001	AL	Total Electric Power Industry	Hydroelectric Conventional	727,118	Hydroelectric
2001	1	1/1/2001	AL	Total Electric Power Industry	Wood and Wood Derived Fuels	399,474	Other
2001	1	1/1/2001	AL	Total Electric Power Industry	Other Biomass	1,694	Other
2001	1	1/1/2001	AL	Total Electric Power Industry	Other	885	Other
2001	1	1/1/2001	AR	Total Electric Power Industry	Coal	2,149,808	Coal
2001	1	1/1/2001	AR	Total Electric Power Industry	Petroleum	124,389	Other

Section 5. Visualization Tools

Tableau for visualizations. My primary focus was to learn about how to build visualizations in Tableau. It was the best for the geographic information, interactivity, and use of the analytics module.

Excel for data consolidation. Moving from three files to two and adding in state geographical data was much easier to accomplish in Excel.

Link to Github with Working Files: <https://github.com/s545773/DATA-VIZ-808182SP24/wiki>

Link to Tableau Public:

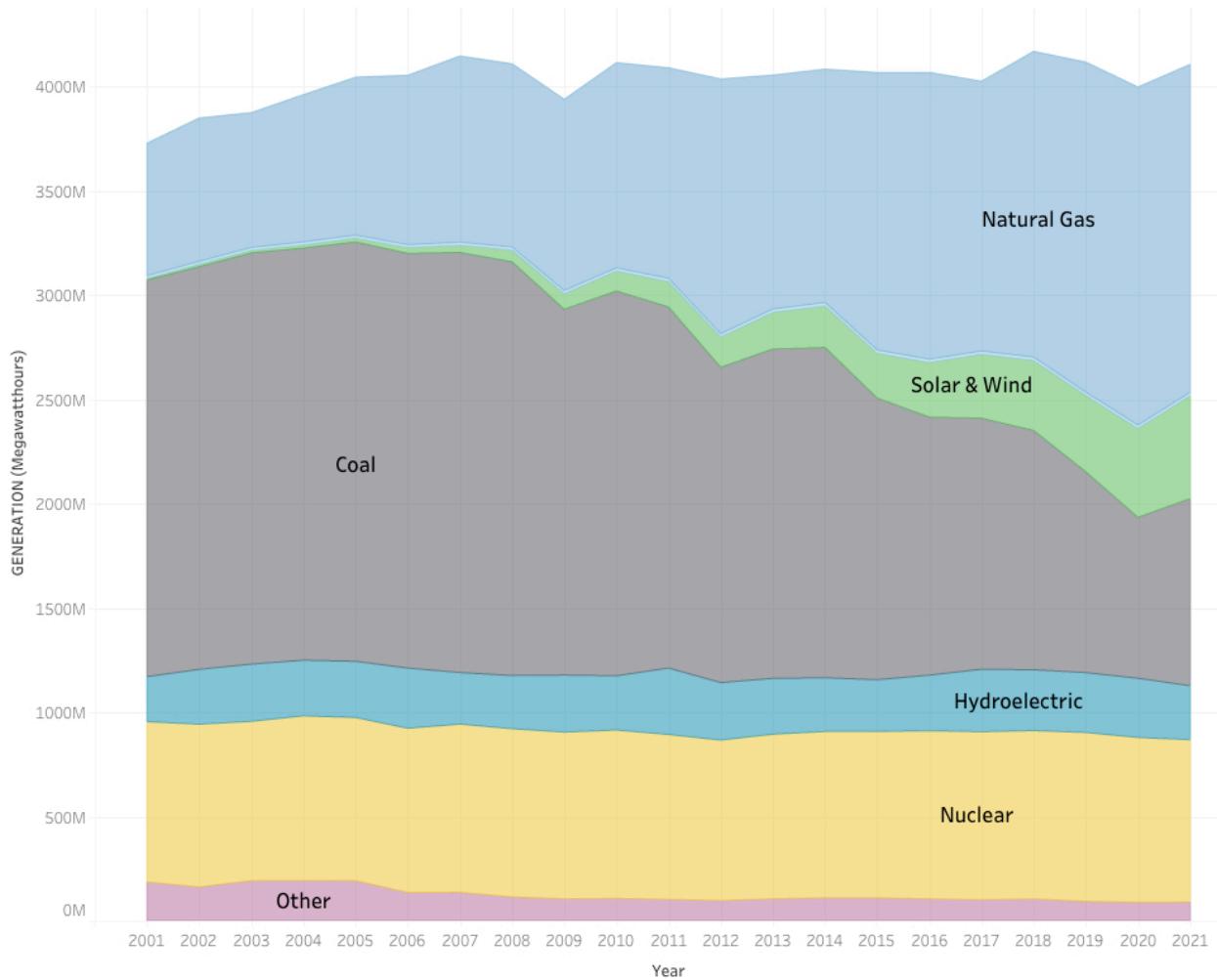
https://public.tableau.com/views/Rausch_Mod7FinalProject/Goal1WhatisthedistributionofenergygenerationtypesintheUSfrom2001to2021?:language=en-US&:sid=&:display_count=n&:origin=viz_share_link

Section 6. Visualizations and Stories

Goal 1: What is the distribution of energy generation types in the US from 2001 to 2021?

Goal 1: What is the distribution of energy generation types in the US from 2001 to 2021?

Throughout the last 20 years, Coal has mostly been replaced by Natural Gas. Solar and Wind have gained popularity since 2005. Coal has shown a slight uptick from 2020-2021. Hydroelectric and Nuclear have remained relatively stable in the past 20 years. The overall power generation in total for the US has only risen slightly in 20 years. Power consumption must either be stable or impacts on higher electricity efficient machinery is keeping the generation needs stable. More investigation would be required.



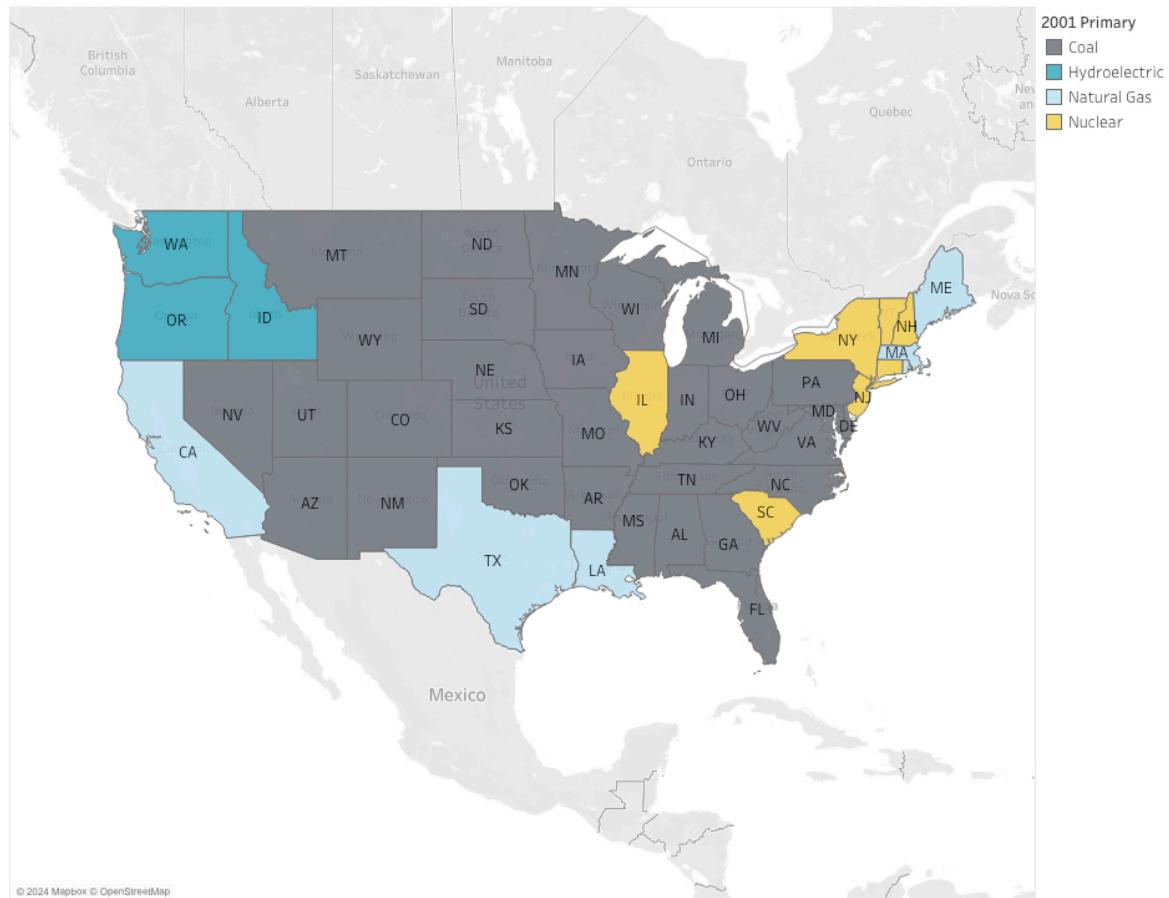
Goal 2: What is the primary generation type by state in 2001 and how does that change in 2021?

There are two pages for this goal.

Goal 2: What is the primary generation type by state in 2001 and how does that change in 2021?

In 2001, Most states predominately used Coal as their primary generation method. The Northeast contained a high degree of Nuclear. The Pacific Northwest is fueled by water, most likely the Columbia River, as it's primary generation type is Hydroelectric

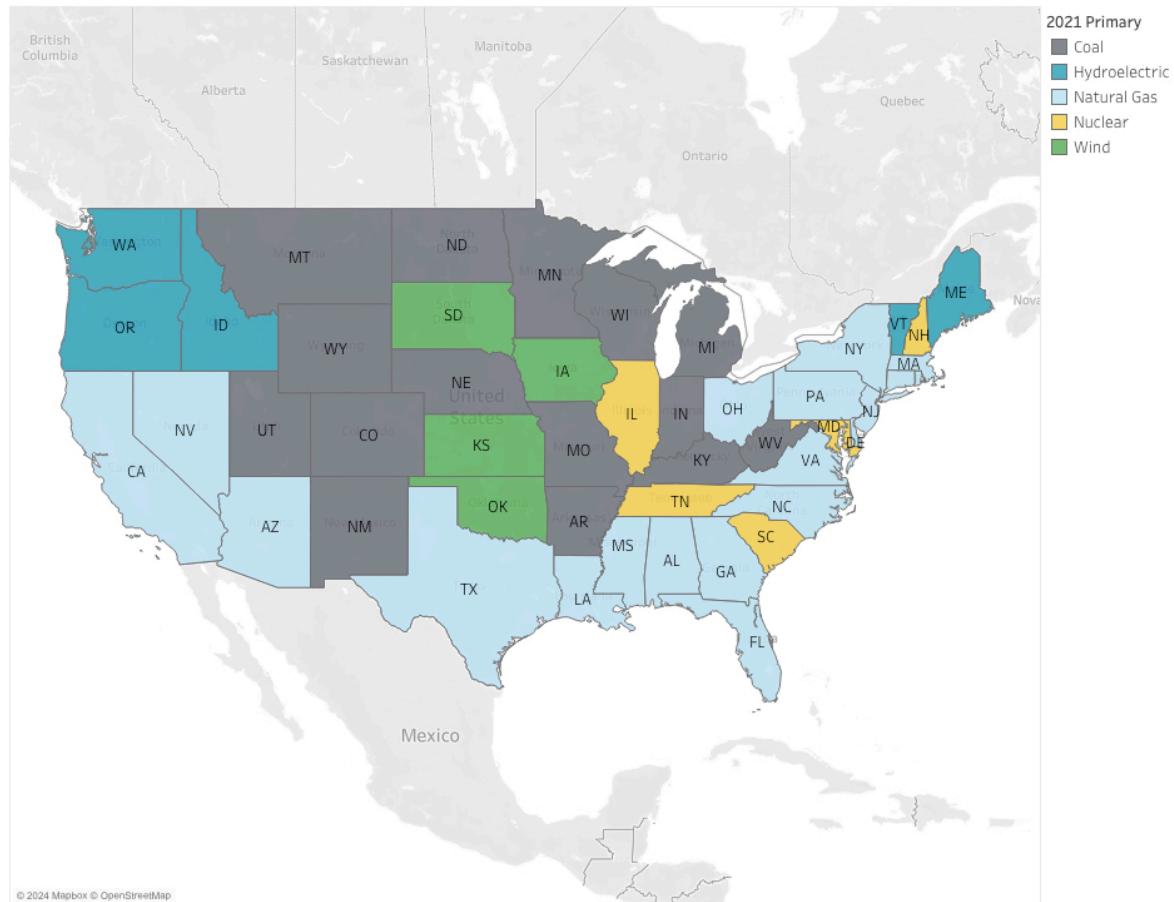
In 2021, a number of states have replaced coal with Natural Gas. The Pacific Northwest continues to leverage hydroelectric as it's primary method. The Northeast has moved away from Nuclear toward Natural Gas and Hydroelectric. The gulf coast is predominately Natural gas due to the abundance of resources of oil and gas available in this region. There is a high concentration of wind generation in the central plains states most likely due to the availability of land and possibly due to government subsidies. More investigation is needed.



Goal 2: What is the primary generation type by state in 2001 and how does that change in 2021?

States predominately used Coal as their primary generation method. The Northeast has a high degree of Nuclear. The Pacific Northwest is fueled by hydroelectric power from the Columbia River, as it's primary generation type is

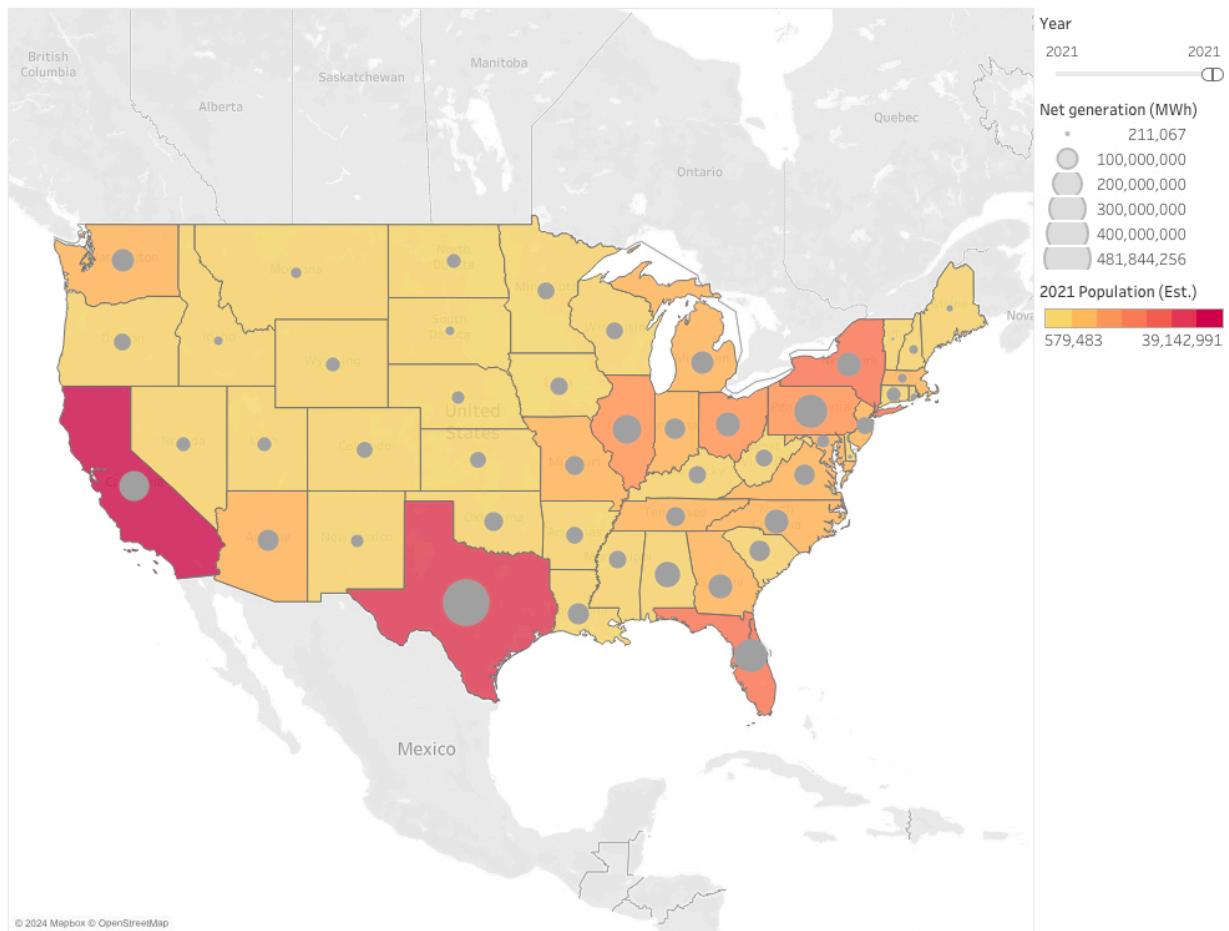
In 2021, a number of states have replaced coal with Natural Gas and Wind. The Pacific Northwest continues to leverage hydroelectric as its primary generation method. The Northeast has moved away from Nuclear toward Natural Gas and Hydroelectric. The gulf coast is predominately Natural gas due to the natural resources of oil and gas available in this region. There is a high concentration of wind generation in the central plains states most likely due to terrain but also possibly due to government subsidies. More investigation is needed.



Goal 3: How does population impact total MWh's generated in 2021?

Goal 3: How does population impact total MWh's generated in 2021?

A state's population size should generally predict how much power generation it will produce. However, looking at this dual layer map of the US, you can see that it isn't always the case. A notable comparison is between Texas and California. Texas is a major generator of energy and yet has less population than California. California most likely relies on other states such as Arizona for its power needs. A similar relationship seems to exist between Pennsylvania and New York. The population density of the South and Northeast create more energy demand.

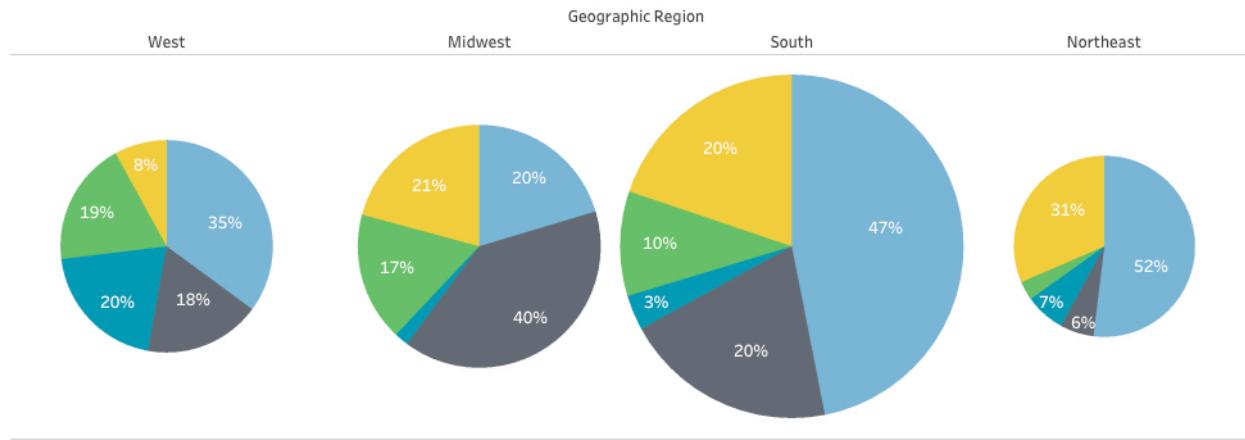


Goal 4: What is the distribution of energy generation by US region in 2021?

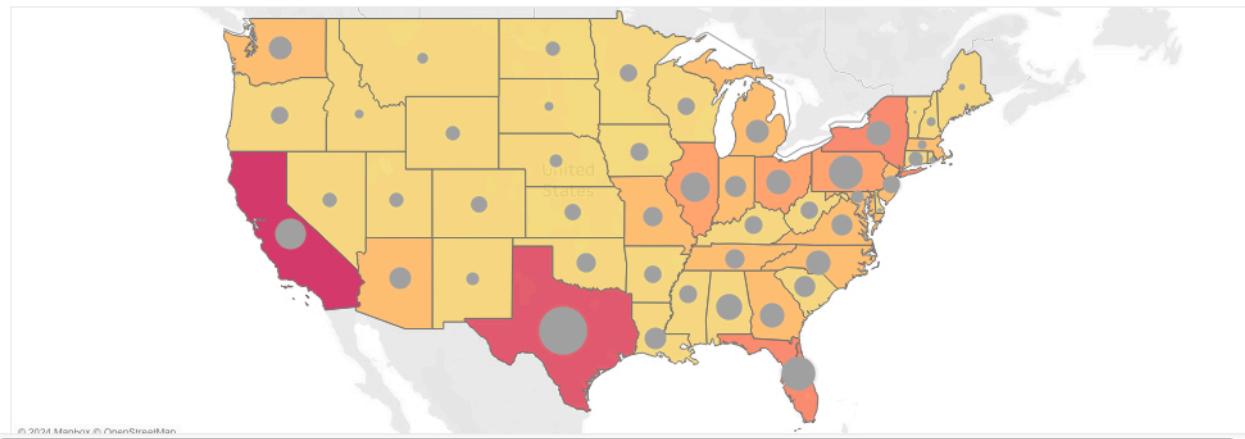
Goal 4: What is the distribution of energy generation by US region in 2021?

These two charts are best viewed together for this goal. As you can see from the pie charts at the top, the majority of power generation in the contiguous US originates from the Midwest and the South. The predominate method in the South is Natural Gas and the Midwest remains Coal. Most of the power generation in the Northeast comes from Pennsylvania and New York in the form of Natural Gas. The West region contains the highest balance across all types between Coal, Natural Gas, Wind/Solar, and Hydroelectric. The total amount of Coal generation in the Northeast has dropped to just 6%.

Distribution of Energy Generation by Region



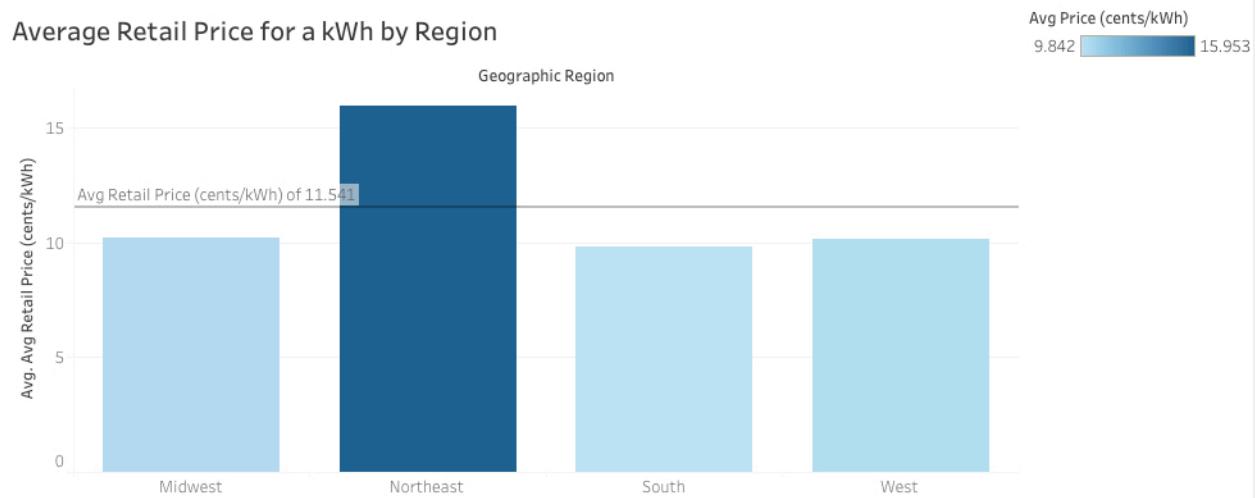
Comparing Population Size to Energy Generation



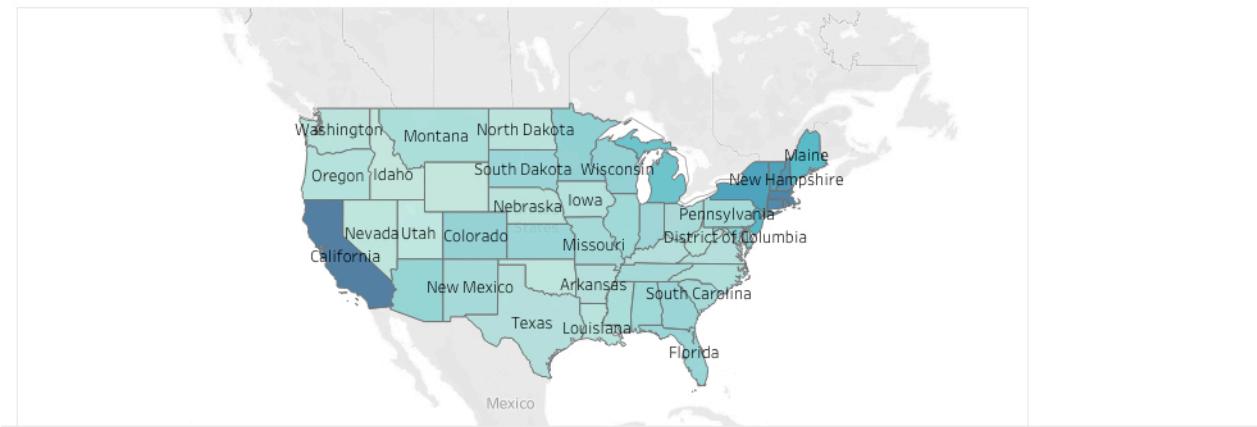
Goal 5: What is the distribution of average Retail Price for a kWh by region?

Goal 5: What is the distribution of average Retail Price for a kWh by region?

Based on this graphic, the consumer in the Northeast United States can expect to pay well above the national average of 11.54 cents per kWh. Based on prior charts, I would attribute this to the fact that the Northeast does not do much power generation, and would rely on purchasing additional power to meet its demands from most likely states in the Midwest. California, another state with high population and not as much generation capability, also pays a much higher rate for energy than other states in the region.



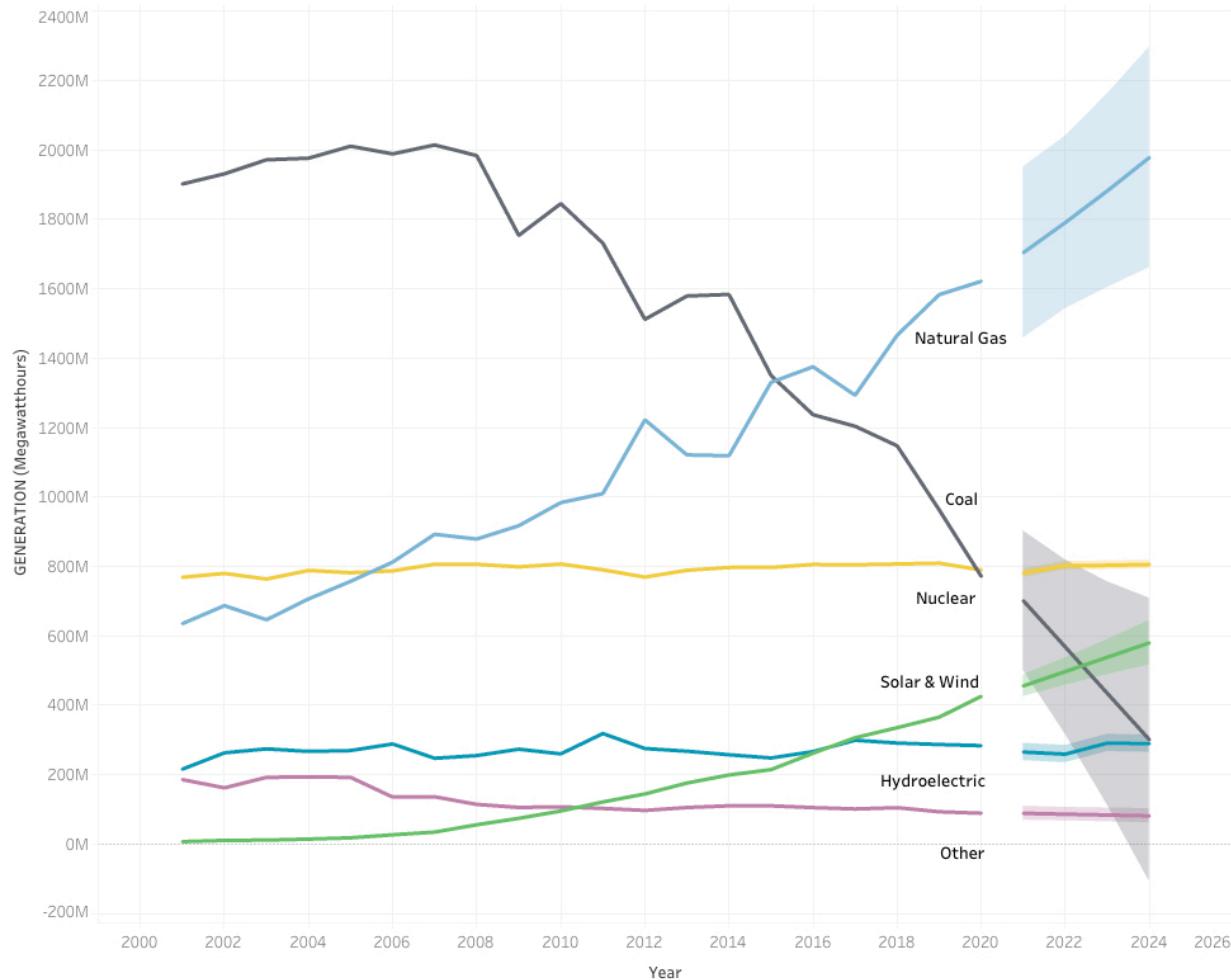
Choropleth Map of Retail price by State



Goal 6: Based on 20 year history, what is the energy generation forecast from 2021 to 2024?

Goal 6: Based on 20 year history, what is the energy generation forecast from 2021 to 2024?

Based on 20 years of history. Natural gas will continue to increase it's overall market share. Solar and Wind will grow. Coal will be a potential loser. However, it is noteworthy that this forecast model does not consider 2021 data where Coal showed an uptick in generation. Both Coal and Natural Gas show the highest volatility in forecasted totals. Hydroelectric and Nuclear are projected to remain stable.



Section 7. Conclusions

This analysis reveals that the United States is undergoing a significant transformation in how it's overall energy is generated. While Coal is making a comeback in 2021, the prevailing direction is moving towards Natural Gas and Wind/Solar. The energy grids of the United States are moving towards more efficient forms of energy production. Coal is clearly the big loser over the last 10 years with more and more states moving towards Natural gas. My assessment is that is due to energy policy and favoring the oil and gas industry, especially in the Gulf Coast states. The South is also clearly the highest generation region which has a lot to do with population, but is also impacted by the hot summer and air conditioning usage.

The move towards alternative energy is fueled not only by governmental regulations and subsides which help drive down US greenhouse gas emissions, it's also a reflection on local policy and consumer choice. Now more than ever, energy consumers have options to "vote with their wallet" on their choice for energy generation. While this report references primary generation methods, looking at a state by state assessment, Wind and Solar are growing in popularity.

When there are not as much generation in the state, then the cost for energy is higher. That is evidenced by the Northeast region and California being dependent on others for their energy production.

Some key additional areas for investigation were a review of the total amount of kilowatt hours generated in the US – it hasn't significantly grown in the last 20 years. Is that a problem with the data set or is this indicative of reality. What is the rational for this? More energy efficiency in consumer, commercial, and industrial products?

Finally, more research should be done on impact of public policy on some of these trends. For example, tax incentives to purchase more energy efficient equipment, additional taxes on usage in high cost states (like CA), impact of deregulation on pricing, and any subsidies which favor alternative energy fuels.