An Analysis on Electrical Power Generation in the United States

Taylor Byers

MCIS 550: Dashboards for Data Analysis

Dr. Daniel Wu

Final Project

December 16, 2020

Contents

Introduction

The Energy Information Administration (EIA) collects data pertaining to energy information in the United States. The administration analyzes this data and provides it, along with their findings to the public. On EIA's website, they provide open data sources for the public to have access and download. With EIA's analysis and discoveries, in addition to the public's, EIA promotes sound policymaking, efficient markets and public understanding of energy usage and generation in the United States. With this information, EIA wants to provide the public with knowledge of energy and its interaction with the economy and environment (U.S. Energy Information Administration, n.d.).

There are multiple different sources and technologies that are available in the United States to generate electricity. The three main categories for sources of energy for electricity generation are fossil fuels, nuclear energy and renewable energy sources. EIA's website provided an open, public dataset that contains information regarding electricity generation in the United States (U.S. Energy Information Administration, n.d.).

This dataset was obtained from data.world, which used in this analysis to create data visualizations and is intended for an audience of electricity producers and their executive, strategic teams to gain insights on past electricity generation. These unique trends and relationships can help executive teams of electrical producers determine and enhance their business strategies for the present day and for the future. In today's fast-paced technological advancements, the need for strong and stable electrical power is widely popular and has become a necessity for households and businesses. This analysis can help answer questions involving locations within the United States that are best for generating electrical power with specific energy sources. Other questions involving areas of the United States that generate the most and

the least electrical power can also be helped to answer through the visualizations produced. Trends of monthly generation and relationships amongst electrical generation and variables measured will be additional questions that are evaluated in this analysis as well. This can help the executive team determine potential growth areas within their company and gain strategic leads. The producer who can best adhere to the ever-growing demand of electrical power, in both efficiency and effectiveness, has chances of being the leading producer in the marketplace.

The Dataset

The datasets used was obtained from the data.world website, in which the original sources were provided by EIA. The two datasets contained data from the years 2008 and 2009 – 2010, respectively. The 2009 – 2010 dataset only contained monthly data up through April 2010. So, it was decided to exclude 2010, as this would skew our data, making it seem as though 2010 had a very large decrease in electrical generation from previous years, which may not have been the case. The data from 2008 and 2009 was then combined to create a single dataset. This final dataset contained United States' monthly electrical power generation figures from January 2008 – December 2009. The dataset contains 37,959 observations with 6 variables measured. The variables measured include Year, Month, State, Type of Producer, Energy Source and Generation (measured in Megawatt Hours). Measures within Type of Producer include:

1. Total Electric Power Industry
2. Electric Generators, Electric Utilities
3. Electric Generators, Independent Power Producers
4. Combined Heat and Power, Electric Power
5. Combined Heat and Power, Commercial Power

6. Combined Heat and Power, Industrial Power

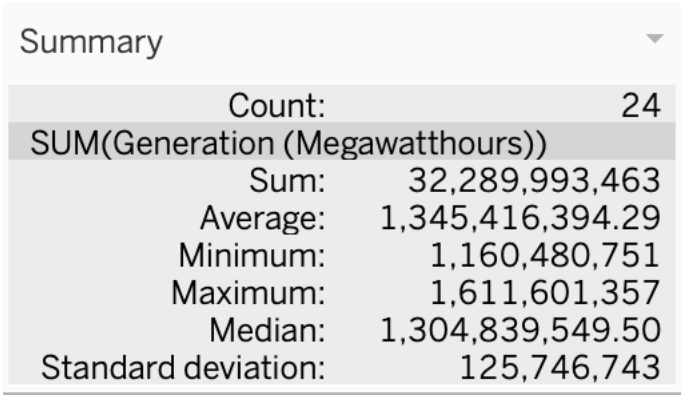
Measures within Energy Source include:

1. Coal
2. Geothermal
3. Hydroelectric Conventional
4. Natural Gas
5. Nuclear
6. Other (batteries, chemicals, hydrogen)
7. Other Biomass (landfill gas, agricultural byproducts, biomass gases)
8. Other Gases (propane, furnace gas)
9. Petroleum
10. Solar Thermal and Photovoltaic
11. Wind
12. Wood and Wood Derived Fuels

Data Exploration

Before producing any final data visualizations, an initial exploration was performed to obtain knowledge on the types of data included in the dataset. First, an overview was performed on our dataset, looking at the data in the Excel file obtained from our data source. Next, the data types of our data were determined. Of the 6 variables measured and recorded within the data, the numeric values included Year, Month and Generation. The string values included Type of Producer and Energy Source. Lastly, the State variable was set to be a geographic location

variable of type state. Descriptive analytics were then performed to improve our understanding of the data. These summary descriptive statistics are shown in Figure 1.



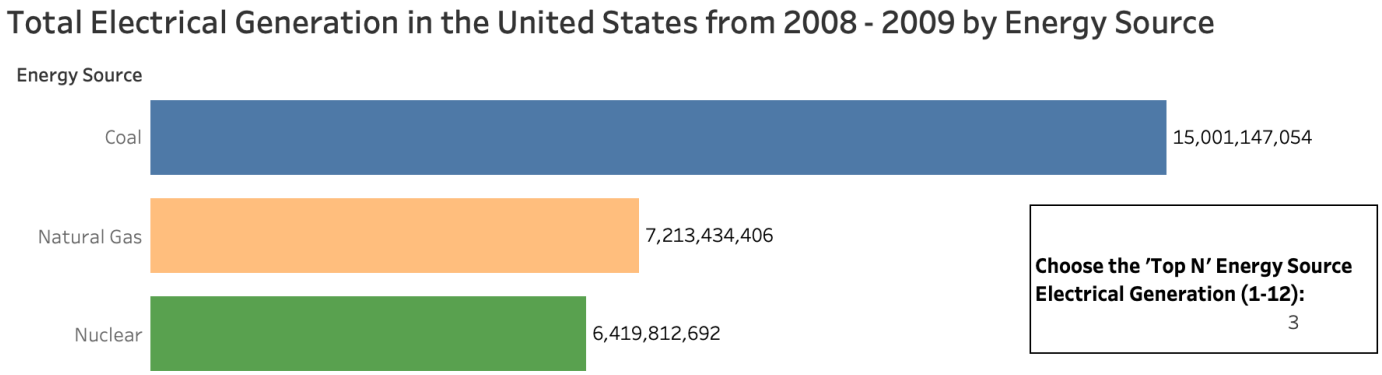
Taylor Byers | Data Source: data.world/doe/electric-power-monthly-data (2016)
Figure 1 Descriptive Statistics for Total Electrical Generation in the United States from January 2008 – December 2009.

As we can see, we have a total count of 24 represented in the statistics. This is because our data is broken down by 24 months, providing a summation of total electrical generation for each month along the two-year time span of our data. The total electrical generation for the entire two years can be seen by the sum in the statistics, as well as the average, minimum and maximum, median, and standard deviation. From these statistics, we can see that the minimum and maximum values do not drastically differ between each other, indicating that perhaps electrical generation for most months are relatively similar. This can also be seen through the standard deviation value. To test our initial data exploration findings, we then produced data visualizations.

Data Visualizations and Analysis

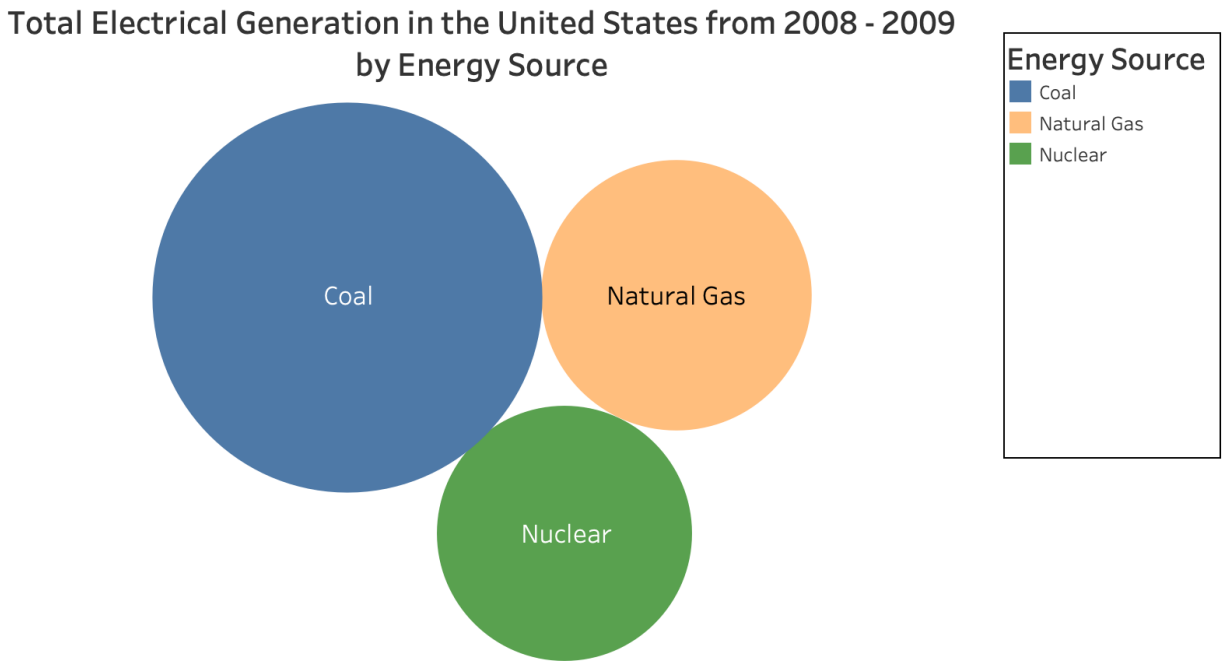
The first two visualizations produced correlate with one another, and so together we populated them to a Tableau Dashboard. These visualizations can be seen in Figures 2 and 3. Figure 2 is a horizontal bar chart, representing the total electrical generation in the United States

for January 2008 through December 2009, being displayed by energy source. Figure 3 is a bubble chart, representing the same data as in Figure 2, but with a more visual view. Together, both the data visualizations can be filtered by the user-decided 'Top N' energy sources.



Taylor Byers | Data Source: data.world/doe/electric-power-monthly-data (2016)

Figure 2 The total electrical generation (measured in Megawatt Hours) in the United States for January 2008 - December 2009, being displayed by energy source.



Taylor Byers | Data Source: data.world/doe/electric-power-monthly-data (2016)

Figure 3 The 'Top N' total electrical generation (measured in Megawatt Hours) from each energy source in the United States for January 2008 - December 2009.

In the produced dashboard visualization, we can see that the top three energy sources that produced the most electricity in the United States from 2008 to 2009 was coal, natural gas and nuclear energy. Coal is the first energy source generator by almost 8,000,000,000 Megawatt

Hours. Coal is a cheap source of energy, so it makes sense as to why it would be the leading. Natural gas is also a cheap source of energy, but nuclear is more expensive than the preceding two. Nuclear gas could be a potential avenue that electrical companies could research and look into to enhance their business strategies.

The next two data visualizations produced were a horizontal bar chart and a bubble chart representing the total electrical generation in the United States for January 2008 through December 2009, being displayed by type of producer. The bar chart and bubble chart by

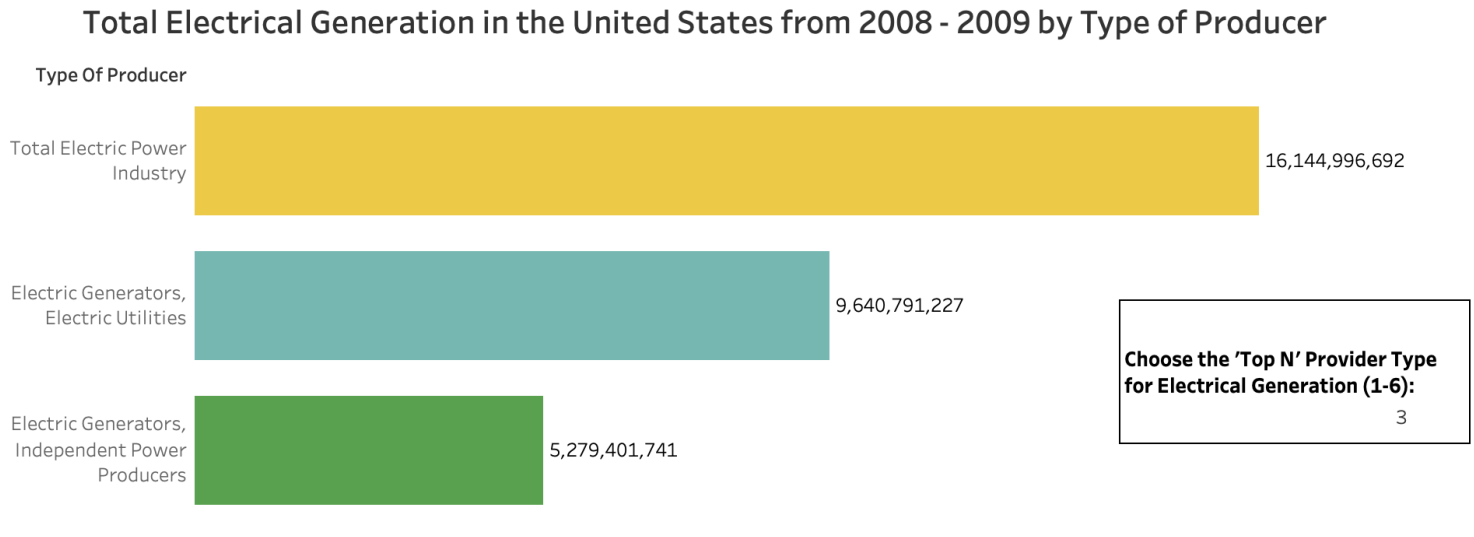


Figure 4 The total electrical generation (measured in Megawatt Hours) in the United States for January 2008 - December 2009, being displayed by type of producer.

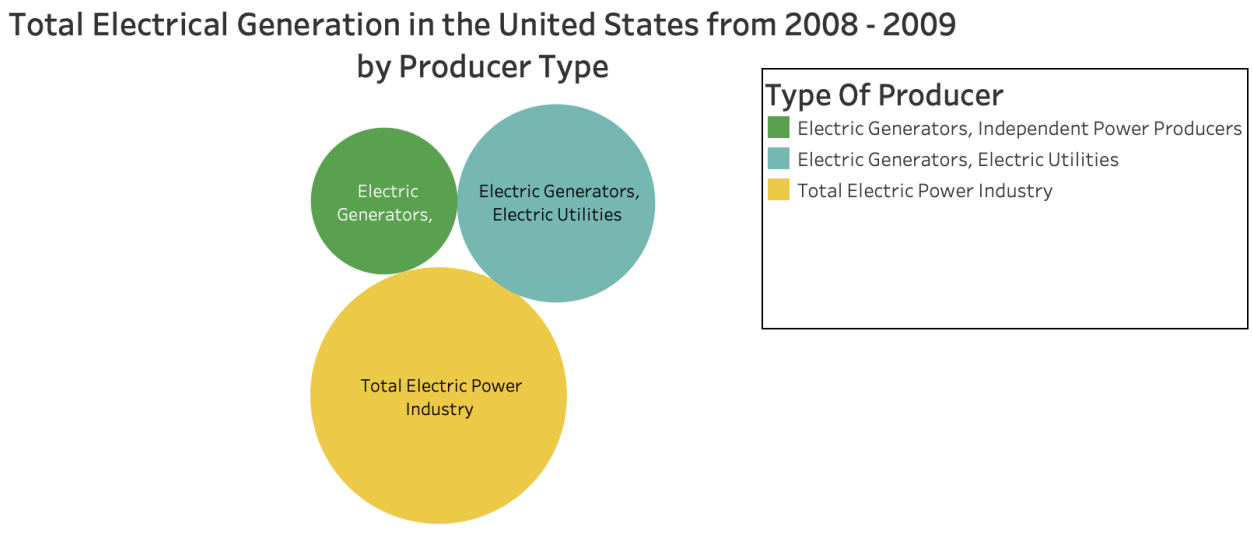


Figure 5 The 'Top N' total electrical generation (measured in Megawatt Hours) from each producer in the United States for January 2008 - December 2009.

producer type can be seen in Figures 4 and 5, respectively, in a Tableau Dashboard format. Again, the viewer can interact with this dashboard visualization by filtering the 'Top N' producer type for electrical generation in the United States. The visualizations shown were filtered to the top three producers of electrical generation.

Here we can see that the producers in the industry for total electric power is in the top spot for producing electricity. This makes sense since these producers are utilizing every energy generating strategy to produce electricity.

The second producer are electric utilities using electric generators. These are public producers that own their own plants and sell the electrical power they generate to the public. A public producer would have a lot of resources to be able to produce large amounts of electric power.

Lastly, the third most producer of electricity are independent power producers using electric generators. Independent power producers are not a public entity and own their own facilities to generate electric power. These producers sell this power to utilities and other end users. It is interesting that privately owned producers made it within the top three in 2008 through 2009. Becoming private could be a potential option for electrical companies ("Independent power producer," 2020).

It is also interesting to note that electric generators produce more electricity than the other producers, such as combined heat and power producers. Companies who do produce using heat and power may want to look into obtaining electric generators as well, since these strategies seem to be more effective.

The next visualization produced with our dataset is shown in Figure 6 below. This data visualization is a map of the United States, showing total electrical generation by state and energy source. Geographic location is an important variable to visualize in our data, as current

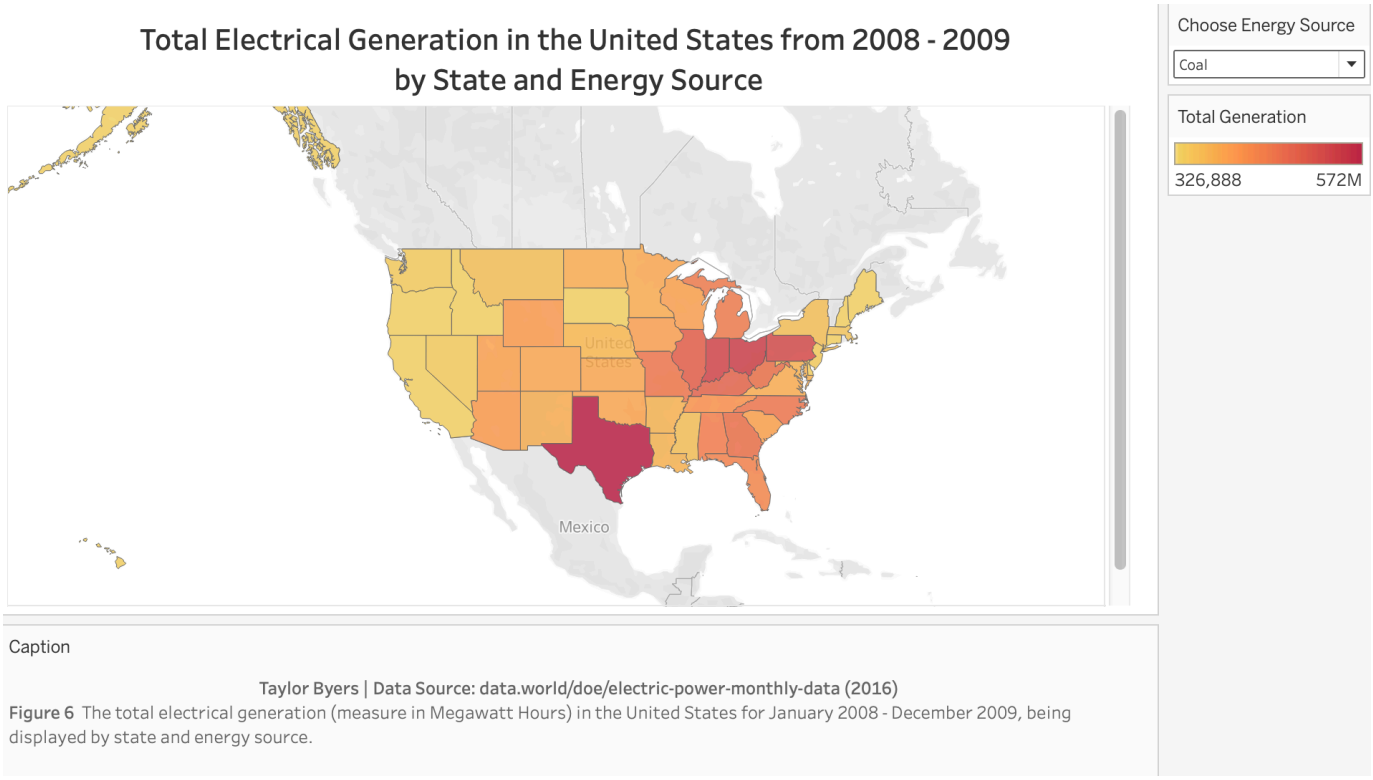


Figure 6 The total electrical generation (measure in Megawatt Hours) in the United States for January 2008 - December 2009, being displayed by state and energy source.

producers should look into what other states they could build plants in to further develop their business. Here we can see that this visualization lets the user filter what energy source is being displayed on the map. Then, the map populates by the color scale with what state produces the most electrical power from the specified energy source. This visualization is a great tool for executive teams at electrical companies. They can see what states within the United States have the most resources available for the energy source as a way to strategize where they could build their next plant and what energy source they could use to create electrical power. In Figure 6, we can see that the coal distribution is higher in the darker colored states, which include Texas,

Pennsylvania, Ohio, Indiana and Illinois. An electrical company that does not use coal as an energy source may want to look into these states to diversify their company.

Figure 7 below shows another map of the United States, with the same color scale and filtering features as in Figure 6. Figure 7 is showing the total electrical generation for each state

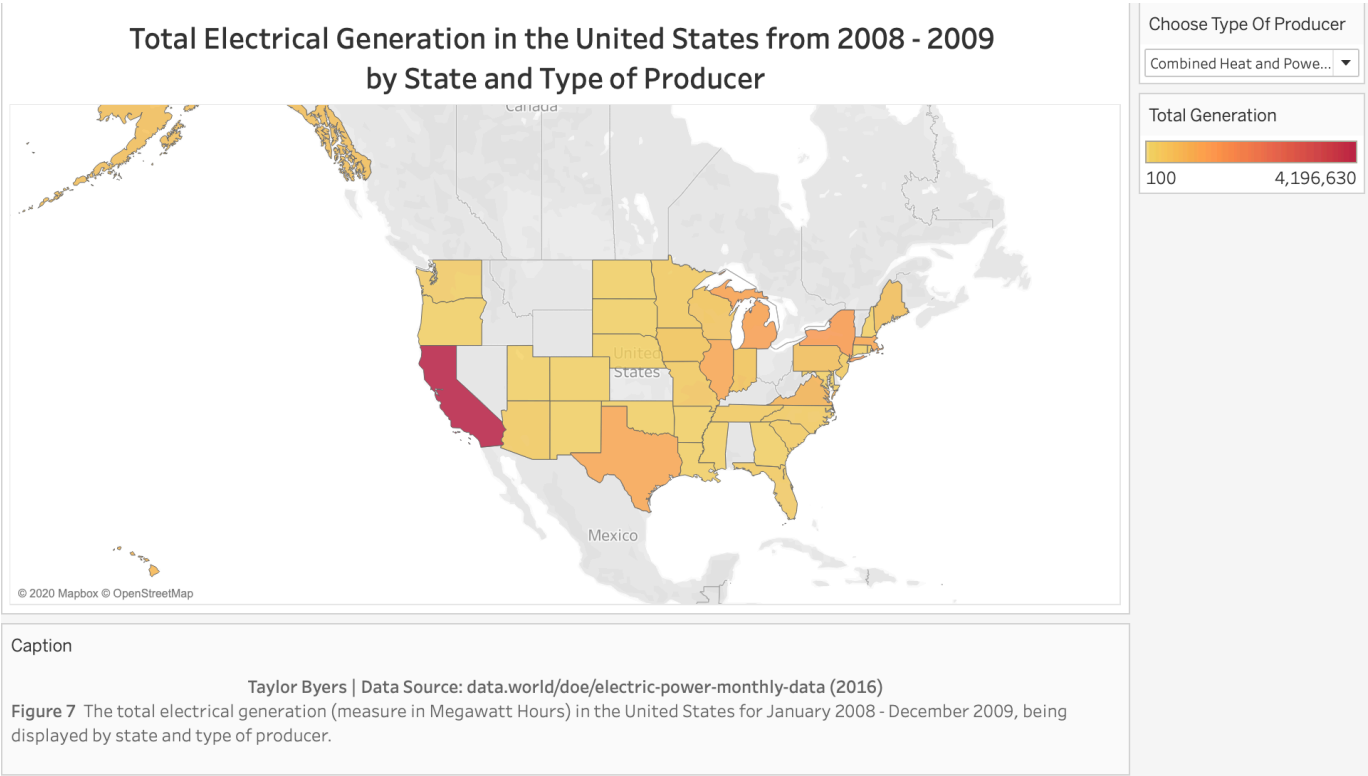


Figure 7 The total electrical generation (measure in Megawatt Hours) in the United States for January 2008 - December 2009, being displayed by state and type of producer.

in the United States for the years 2008 through 2009 by type of producer. In this visual, combined heat and power for commercial power producers was selected. Here we can see that there is a high distribution for this type of electricity producer in California. This is an insight that electrical companies could look into as to why this is the case. Perhaps it is the size and population of California that makes these producers leaders in the electricity industry in California.

It is important to explore possible trends within the 24 months of data we are analyzing. Figure 8 represents the trend of total electrical generation in the United States by month. A

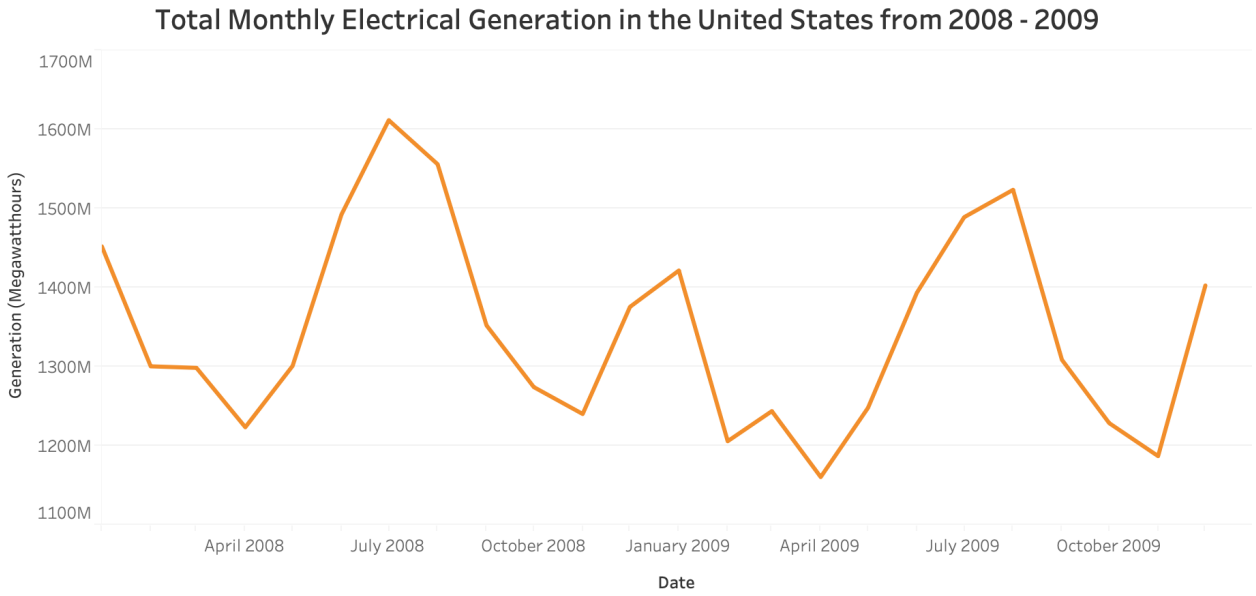


Figure 8 The trend of total electrical generation (measured in Megawatt Hours) in the United States by month for the timespan of January 2008 - December 2009.

trend can be seen in this line graph. The summer months (June, July, August) and winter months (December, January February) are showing spikes in the electrical generation produced in the United States. Electrical companies are generating more electricity due to cooling and heating in homes, offices, other buildings, etc. to adhere to the outside weather. We can then see, the preceding and following months after these spikes have lower total generation. By knowing that there is, in fact, a seasonality trend amongst electrical generation in the United States, executives at electrical companies can strategize their generation distribution to make sure they are keeping up with the demand of electricity.

Along with visualizing trends of the overall data, two additional line graphs were produced to show trends amongst the energy source and type of producer variables that we are also analyzing. Figure 9 below shows total monthly electrical generation in the United States by energy source. This visualization has an added filter that the viewer can interact with and choose what energy source they would like to represent on the graph. Figure 9 shows trends of the

energy sources geothermal, hydroelectric conventional, natural gas, nuclear, wind and wood derived fuels. Amongst all the chosen variables, natural gas does have higher generation over the

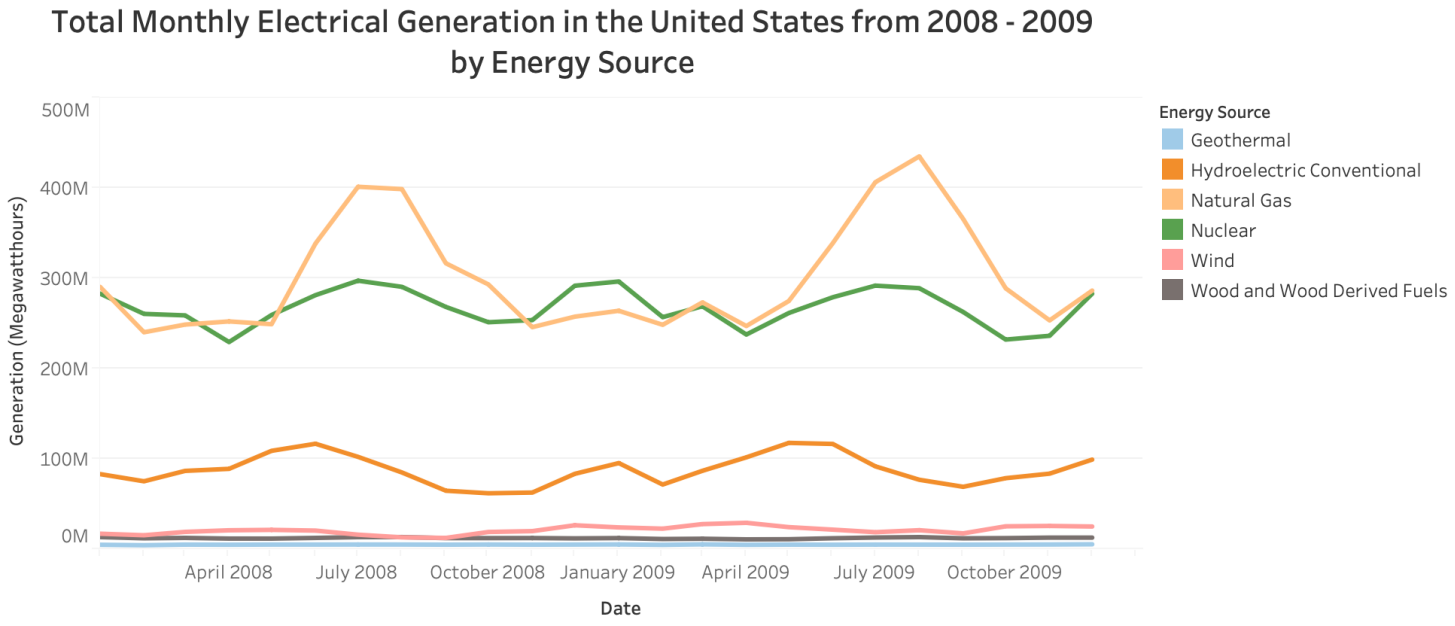
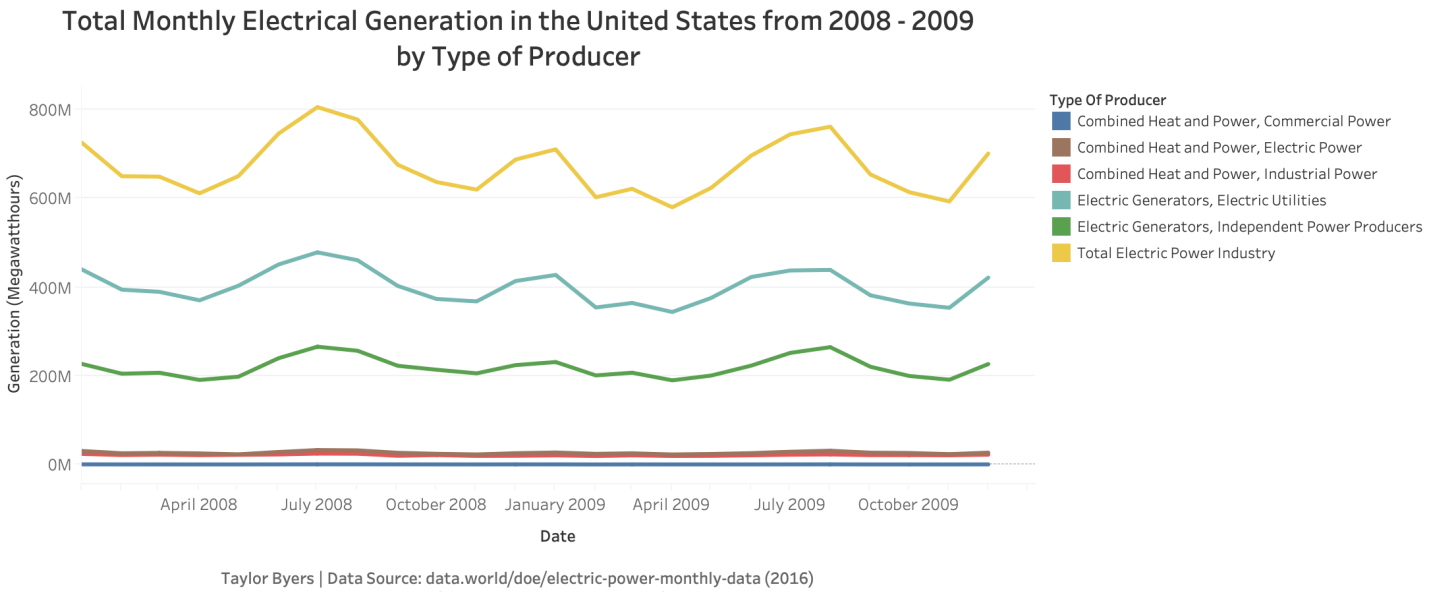


Figure 9 The trend of total monthly electrical generation (measured in Megawatt Hours) in the United States by energy source for the timespan of January 2008 - December 2009.

two-year span. However, around January 2008 through May 2008 and then around November 2008 through May 2009, natural gas and nuclear energy are close in generation. It is important to evaluate why natural gas had huge spikes in electrical generation, but outside of those spikes, it is similar to nuclear energy. These two energy sources rise above the other selected choices, so investing an electrical company in natural gas or nuclear energy sources could lead to a strategic move for companies to increase their overall electrical generation.

Figure 10 shows the trend of our data for monthly electrical generation by type of producer over the timespan of 2008 through 2009. In this visualization line graph, the filter option is available for users to select what generation trend they want to compare based off of type of producer. In the visualization shown in Figure 10, all of the types of producers were

chosen to be represented on the graph. Here we can see that, as expected, the total electric power industry was the leader in electrical generation between 2008 and 2009. We can see



Taylor Byers | Data Source: data.world/doe/electric-power-monthly-data (2016)

Figure 10 The trend of total monthly electrical generation (measured in Megawatt Hours) by producer type in the United States for the timespan of January 2008 - December 2009.

the seasonality trends amongst our data as we saw in Figures 8 and 9, however, the trends are mainly seen through the total electric power industry and the electric generators, electric utilities producers. Independent power producers with electric generators have a slight seasonality trend. Less of a fluctuation in generation amongst a year of time could be more beneficial to some companies. Independent producers may not have the funding that public producers have, making it crucial to maintain a steady generation flow throughout each year.

Conclusion

Electricity is an important secondary energy source, being converted from and generated by multiple primary energy sources, as seen in the data used in this analysis. There are also a variety of producers and tactical strategies used in order to convert and generate electricity from these sources, as we also have seen throughout this analysis. Overall, electricity has dramatically

changed the daily lives of people worldwide, becoming a necessity that is rarely thought about and often taken for granted (U.S. Energy Information Administration, 2020).

Electrical companies need to constantly evaluate their generation amounts as well as their current status of strategies used to generate electricity for their customers. With today's society being heavily driven by technology, new advancements in the technological field are constantly being developed, some requiring more or less electrical power. Electrical companies need to ensure their business is keeping up with this high demand.

As seen through this analysis, coal, natural gas and nuclear energy sources are the most popular and leaders of electrical power generation. However, a question executive teams need to constantly evaluate is: "Is this best for our company?" Evaluating the type of producer they currently are and the type of energy source they use against trends and relationships in the data we provided is the next steps that executive teams at electrical companies need to consider in order for them to take their business to the next level.

References

Data.world. (2016). *Electric Power Monthly Data* [Data file]. Available from data.world website: from https://data.world/doe/electric-power-monthly-data

Independent power producer. (2020, June 25). Retrieved December 10, 2020, from https://en.wikipedia.org/wiki/Independent_power_producer

U.S. Energy Information Administration - EIA - Independent Statistics and Analysis. (n.d.). Retrieved December 7, 2020, from https://www.eia.gov/about/

U.S. Energy Information Administration - EIA - Independent Statistics and Analysis. (2020, March 20). Retrieved December 10, 2020, from https://www.eia.gov/energyexplained/electricity/