19CSE304
Foundations of Data Science

Assignment

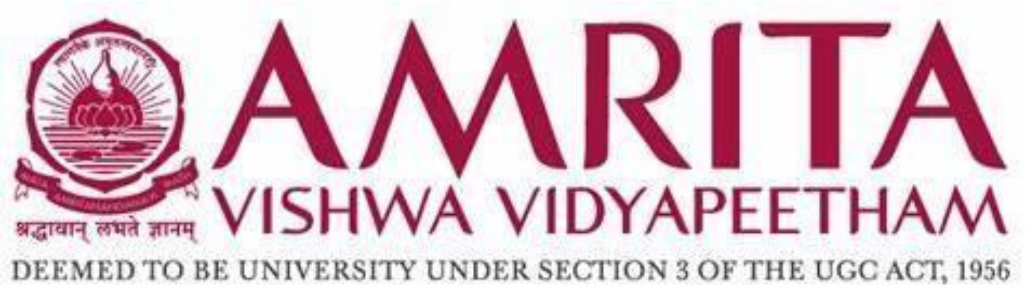
Exploratory Data Analysis of Renewable Energy Production and Consumption by Source

Submitted by:

K. VIGNESWARA REDDY

BL.EN.U4CSE22030

CSE-A



AMRITA SCHOOL OF ENGINEERING, BANGALORE

AMRITA VISHWA VIDYAPEETHAM

BENGALURU-560035

October -2024

INTRODUCTION:

The dataset titled **"Renewable Energy Production and Consumption by Source"** provides comprehensive data on the production and consumption of various renewable energy sources in the United States. It covers key renewable energy categories such as **solar energy**, **wind energy**, **hydroelectric power**, **geothermal energy**, **wood energy**, and **biofuels**, detailing their monthly and annual production and consumption over a specific period. This dataset is crucial for understanding the trends and contributions of renewable energy sources in the energy mix and for tracking the shift toward cleaner energy alternatives.

Renewable energy is at the forefront of global efforts to combat climate change, reduce greenhouse gas emissions, and decrease dependence on fossil fuels. The dataset provides insights into how different renewable energy sources have performed over time, their seasonal fluctuations, and their contribution to the overall energy consumption in the country. It allows for the analysis of growth trends, comparisons between production and consumption, and the identification of which renewable sources are becoming more prominent in the energy sector.

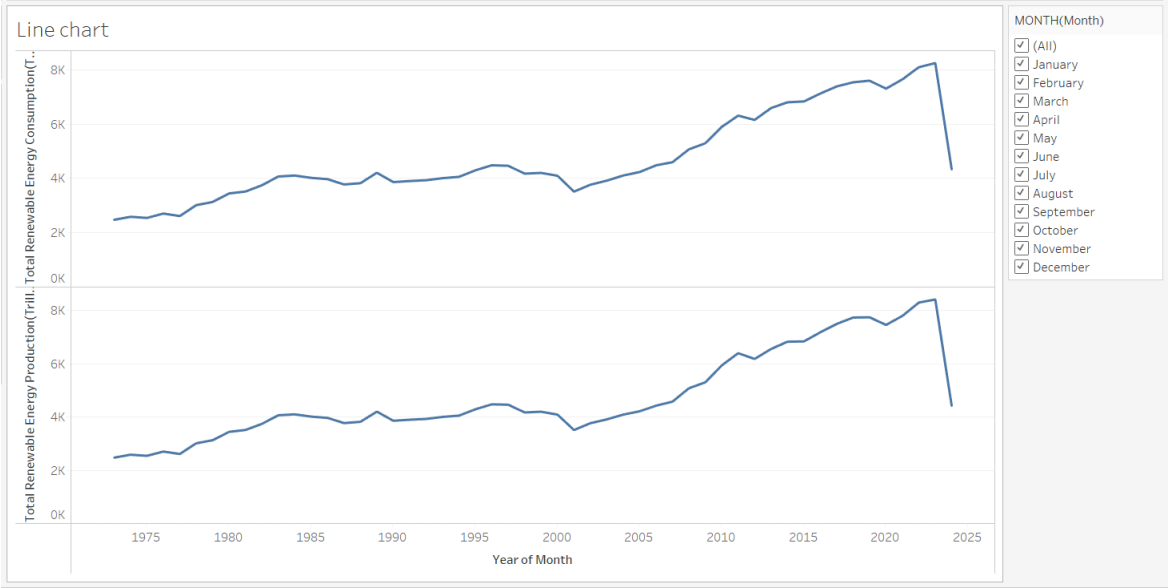
This data can be utilized to track progress toward energy sustainability goals, assess the impact of renewable energy policies, and make data-driven decisions regarding future investments in the renewable energy industry. By analyzing this dataset, we can explore key questions about energy trends, understand the dynamics of energy production and consumption, and gain a deeper appreciation of the role renewable energy plays in the broader energy landscape.

Questions:

1. How have the total renewable energy production and consumption changed over time from the start of the dataset?
2. What are the contributions of different energy sources (wood, biofuels, solar, etc.) to the total renewable energy production, and how have these contributions changed over time?
3. How do the yearly production and consumption values for each energy source compare?
4. During which months do different renewable energy sources have the highest and lowest consumption?
5. What is the distribution of renewable energy production values for different sources, and which sources have the most variation?
6. Is there a relationship between solar and wind energy consumption over time?
7. How do geothermal and hydroelectric energy consumption trends compare over time?
8. In a specific year, what was the contribution of each renewable energy source to the total renewable energy production?
9. Which energy sources have the highest and lowest total consumption across the dataset?
10. How has the total biomass energy consumption changed on a yearly basis?

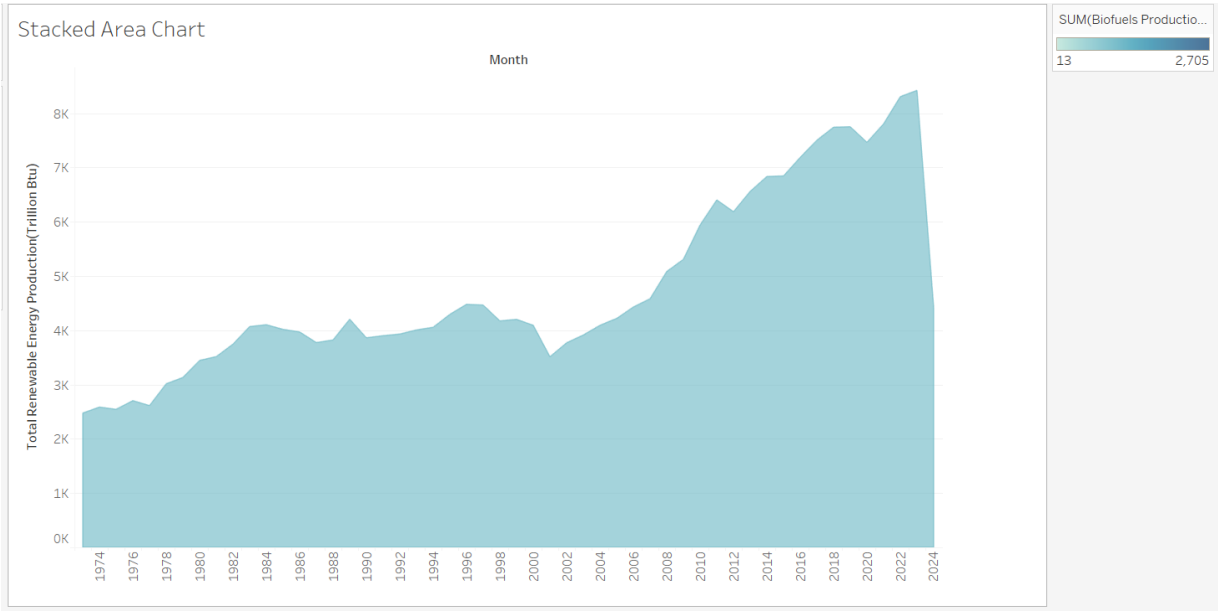
1. **Line Chart**: Monthly Trend of Total Renewable Energy Production and Consumption

   This shows the dynamics of renewable energy production and consumption. You could easily see the trend, peaks, and difference in production and consumption on the same graph while comparing the same over the years.



The graphical representation includes both production and consumption trends over time for total renewable energy consumption and production from the early 1970s to the mid-2020s. Smooth positive trends are observable for both production and consumption. However, it can be noted that the production of renewables takes off strongly after the 1990s. There is a sharp dip in the production and consumption, it seems, following 2020, perhaps related to exogenous factors such as disruptions globally or changing economies. Although renewable energy appears to have slightly dipped in earlier years, the overall trend has been upward, indicating a long-term shift toward more use of renewable sources. The dip at the tail end of the timeline thus suggests a need for further investigation into the causes of decline in renewable energy activity post-2020.

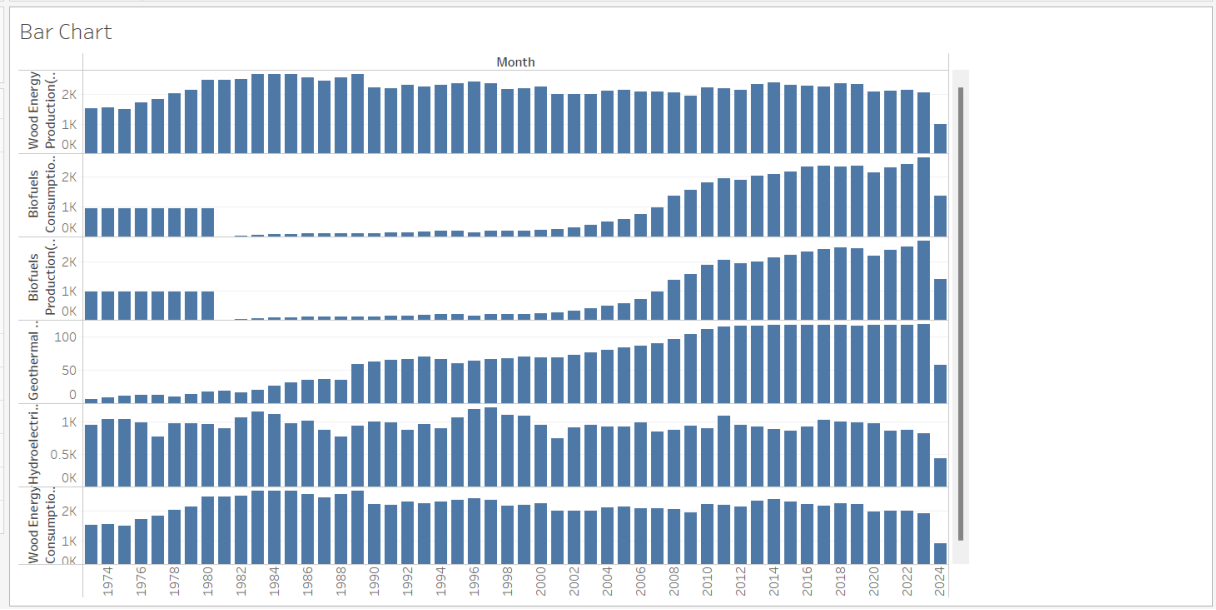
2. **Stacked Area Chart**: Breakdown of Energy Sources Over Time
   This shows for each of the renewables - wood, biofuels, solar and the like - how much comes in on each of those. The areas are stacked so that the reader can see how each
   source has come into and out of relative importance over time.



The stacked area chart illustrates the renewable energy production from 1974 to 2024, with a perfectly positive trend for all the years. At first, the production was moderately on the rise until the early 2000s, whereas after that there was heavy production in renewable energy, peaking around 2020. Global interest is building up in renewable energy source production and use. However, a drastic fall off after 2020 could be an external factor or disturbance in energy production. The long-term trend is committed to more expansion of renewable energy, though the drop requires research into its causes.

3. **Bar Chart**: Yearly Comparison of Production and Consumption for Each Energy Source

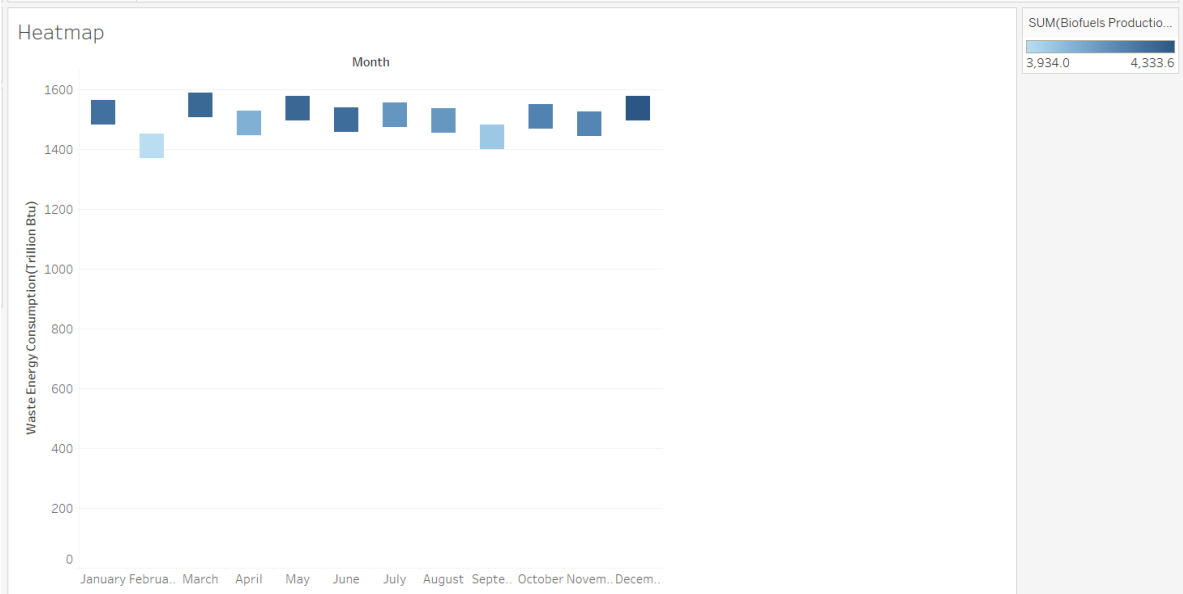
   A bar chart would also allow you to compare how much is being produced and consumed from each source on an annual basis, which gives insight to the increase or decrease of specific sources.



   Several metrics over time, by month are shown in a bar chart. In different parts of the figure, different variables are depicted - namely, Wood Energy, Electricity, Renewables, and Coal/Natural Gas. This would mean some fluctuations in the consumptions or productions pattern of those energy resources year-over-year, with the electricity consumption and renewables having impressive growth. Coal and natural gas look rather volatile. Wood energy is stable but with slight fluctuations. This shows a gradual swing towards more sustainable energies, particularly renewables, over time.

4. **Heatmap**: Monthly Energy Consumption of Each Source:

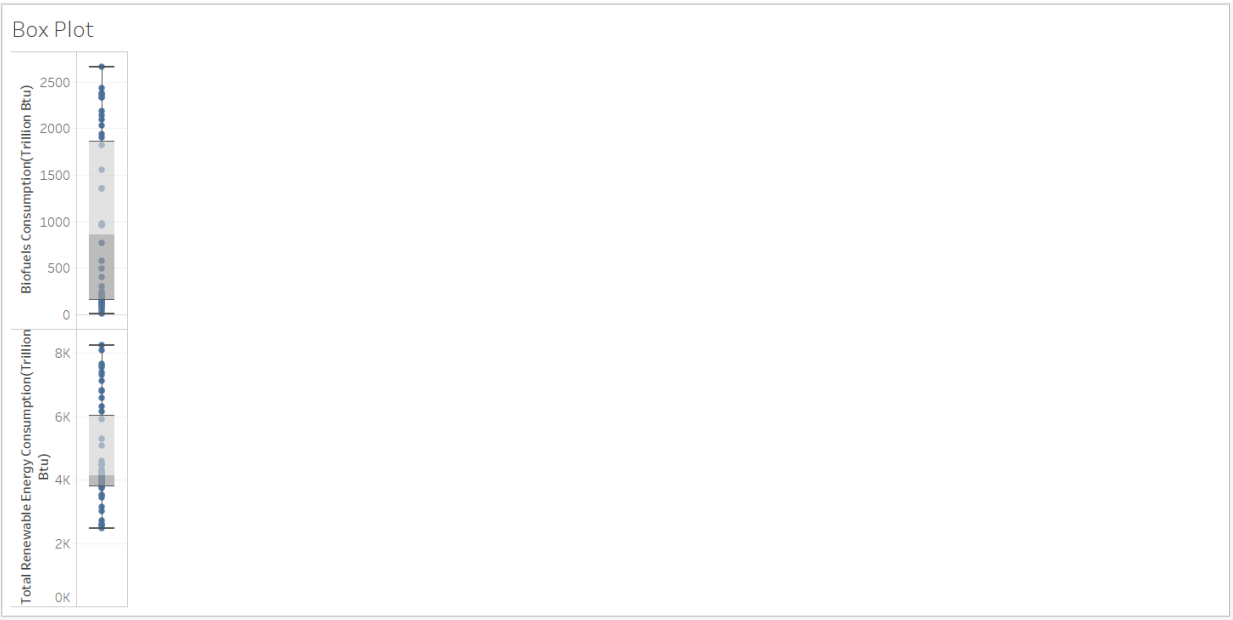
This visualization uses color intensity to show how different sources of energy consumption vary monthly. Darker colors may represent higher consumption, helping you quickly see high-consumption periods



This represents a heatmap of the fluctuation of energy consumption in million Btus over the months. More darker colors mean more and lighter colors indicate low values of energy consumption. From the graph, one can see that there is uniform distribution in consumption. Then again, this uniformity is slightly variant. Overall, the months near the end of the year appear to be higher in consumption, indicated by the darker blue squares, whereas the earlier months are relatively low to moderately sized in energy use. This would suggest seasonal trends possibly tied to heating and/or cooling needs.

5. **Box Plot**: Distribution of Renewable Energy Production by Source

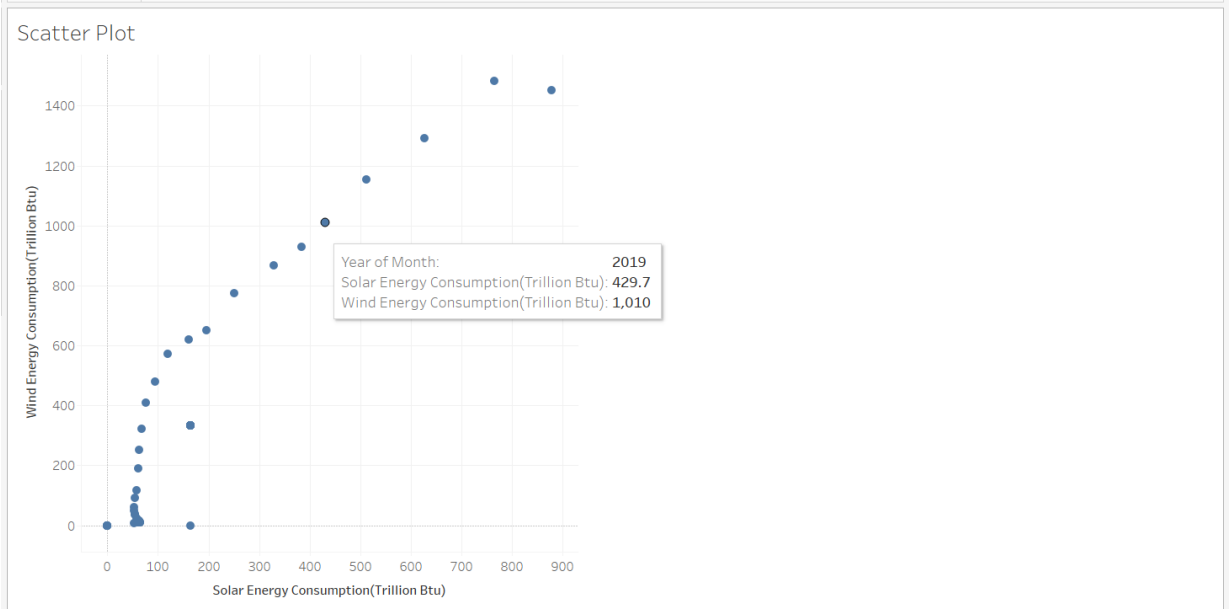
The box plot in the figure provides the distribution of renewable energy production from the sources above, which gives the median, quartiles, and outliers. This is helpful in understanding the variability of the production among the sources.



The box plot represents the energy use distribution from biomass consumption into total renewable energy consumption. Under biomass, the data is largely spread with different outliers showing that most of the usage types are around a variety of values, but the most values seem to be located around the median level. IQR seems narrow compared with the total range. Therefore, it is indicated that the average usage is stable, but there are occasionally extreme values. The spread in total renewable energy consumed is larger with higher median values, and its IQR is wider in comparison, meaning this shows more variability in the use of renewable energy. This too has some outliers that occur because of occasional consumption peaks. The box plots shown above reflect variability and trends in consumption: generally, renewables tend to show a wider range.

6. **Scatter Plot**: Relationship Between Solar and Wind Energy Consumption

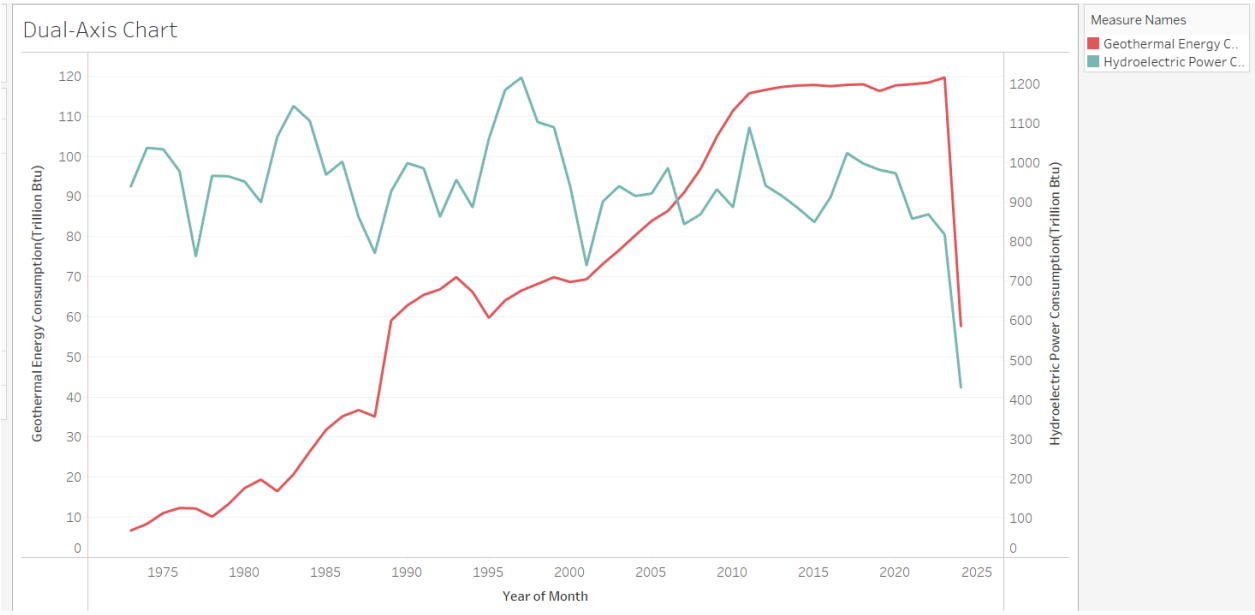
This scatter plot shows correlation between consumption of solar and wind energy. Each point is the value for a particular month and the trends may indicate that increases in consumption of one correlate with increases in the consumption of the other, or perhaps no patterns exist.



The scatter plot very clearly shows that there is a high positive correlation between consuming solar and wind energy. This simply means that as the consumption of solar energy rises, so does the consumption of wind energy; it illustrates a trend towards consuming cleaner energy sources. Both solar and wind energy consumptions trend upwards with time. The same is the conclusion for the data points on 2019, with 429.7 trillion Btu of solar energy consumption and 1,010 trillion Btu of wind energy consumption, marking an enormous growth in utilizing renewable sources.

7. **Dual-Axis Chart**: Geothermal vs Hydroelectric Consumption Over Time

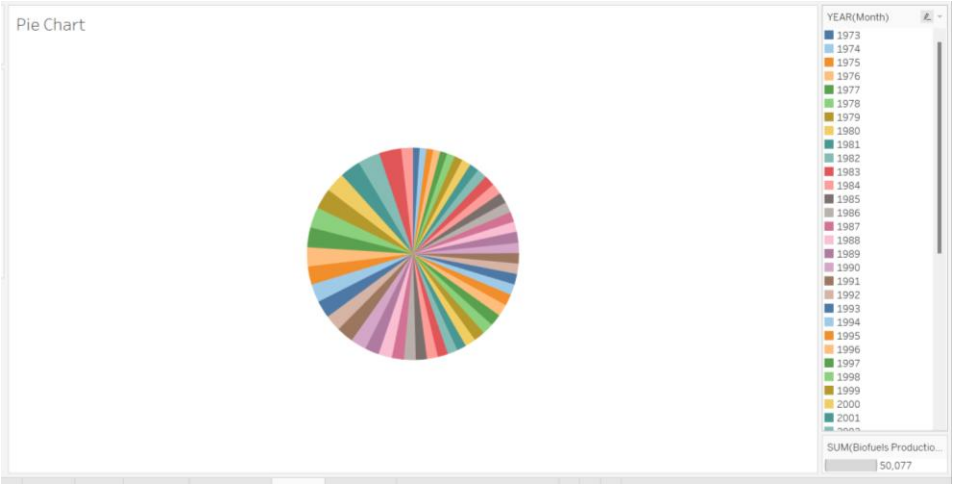
This chart allows for a trend of geothermal energy consumption over time by two axes similar to hydroelectric energy; it helps you check whether both run in line or diverge in a certain direction



The bar chart illustrates the trend for consumptions of geothermal and hydroelectric power in the United States from 1975 to 2025. On closer observation, it is evident that generally the hydroelectric power consumptions are higher as compared to the consumption of geothermal power in the same period. Hydroelectric power consumptions had seasonal movements; however, the overall level of hydroelectric consumption was pretty stable during the period. From the graph above, it has been seen that utilization of geothermal power has tended more sharply upward, depicting an increased usage of the power generation facilities. There is a swing toward the use of geothermal energy, as depicted by the chart. The chart depicts the U.S. energy situation, increasingly but steadily relying on geothermal energy.

8. **Pie Chart**: Contribution of Energy Sources to Total Renewable Energy in a Year

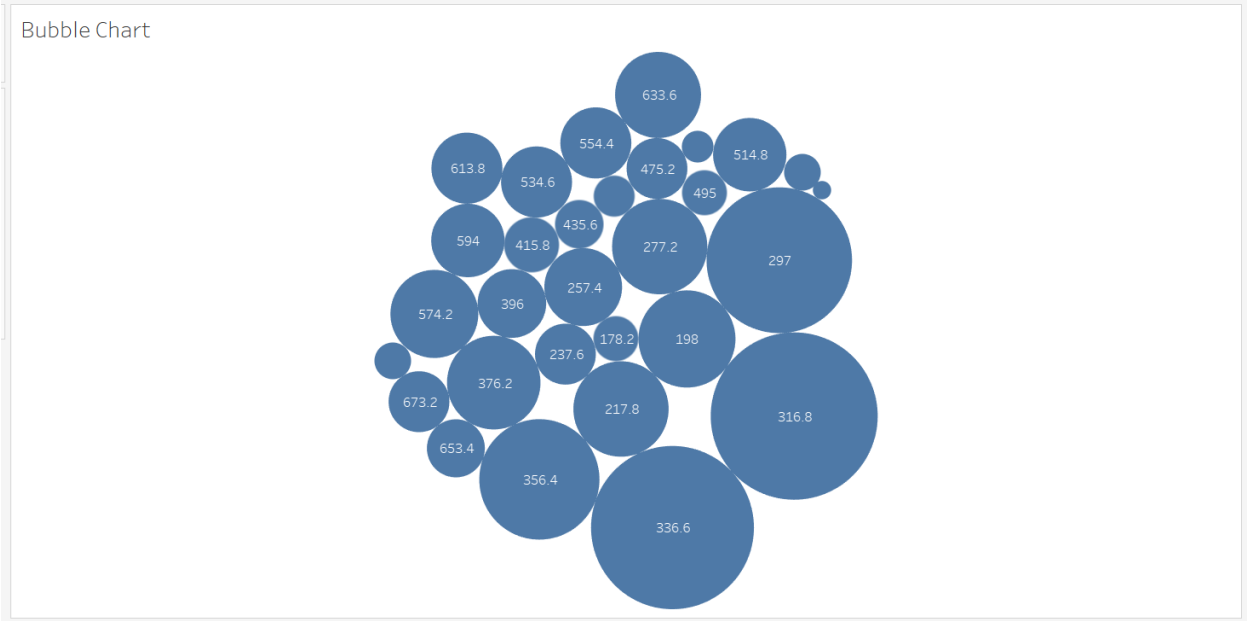
This pie chart is good for illustrating the share of various energy resources (wood, biofuels, etc.) that make up aggregate renewable energy generated in a given year. It explains which types of energy resources tend to dominate



Pie chart The production of biofuels in distribution per year: each bar is a different year and the height of the bar, its value represents. The biofuels production growth is observable from the chart; it is generally upward sloping with a major peak mid the 1990s and a gradual increase from the early 200. This suggests a growing trend to accept and utilize biofuels as an alternative source of energy.

9. **Bubble Chart**: Energy Sources by Total Consumption

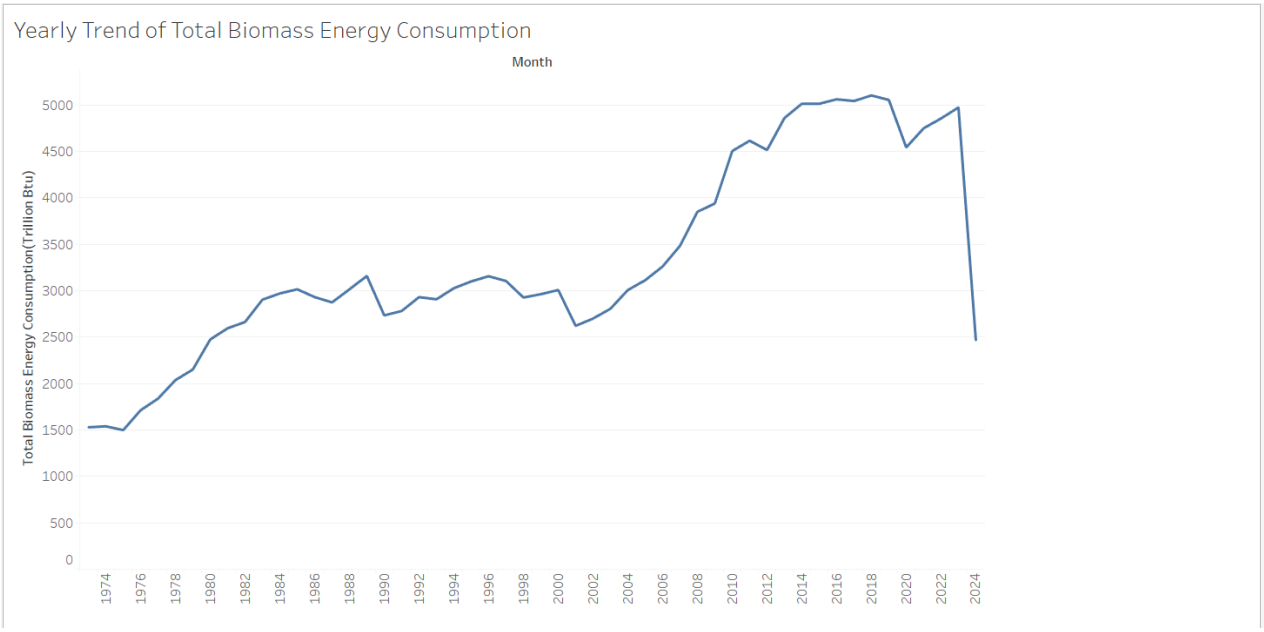
Each little bubble represents the overall consumption of a source of energy. Greater means more consumption, so it's pretty easy to tell which are more important.



The size of each bubble is directly proportional to the value that it is representing. For example, the largest bubble represents a value of 633.6 and the smallest represents a value of 198. This type of graph is nice for visualizing data that spans an enormous range since it clearly shows which values are the largest and smallest. In the chart, it has been shown that 633.6 was at its maximum value, and 198 at the minimum.

10.Yearly Trend of Total Biomass Energy Consumption

This visualization shows the yearly trend of biomass energy consumption. It highlights long-term growth or decline, helping to understand the role of biomass in renewable energy



The graph provided is a line graph of yearly totals of the consumption of biomass energy in the United States from 1974 to 2024. The graph shows that in the recent past, the consumption of biomass energy was in a sharp positive trend. It starts quite low in the 1970s and steadily increases up to the 1980s and 1990s, then more sharply up to the 2000s and 2010s. Despite a slight sag in consumption around the year 2000, the overall message here is one of steady and significant growth in biomass energy use over the past half-century.