njitVis | Energy Viz
Building Interactive Web-Based Data Visualization Tool

Visualizing Energy Consumption Patterns in the USA and the World

Parth Merchant | Aritra Dasgupta

Abstract:
In the 21st century, energy consumption is at an all-time high, with individuals possessing multiple electronic devices as well as living in homes filled with electronics. As people run their technological devices all day year long, energy is being used at a rate higher than it has been in previous years. In order to reduce energy consumption, as citizens, we must understand in what sector we're consuming energy and how much we are using. Understanding energy usage on a micro and macro-level is essential to gaining insights on how we utilize energy and how we can potentially reduce our carbon footprint.

Objectives:

1. Analyze energy consumption across different sectors annually between 1949-2019 to visualize energy consumption over time (Interactive Line Chart)
2. Understand how different in the world consumes different energy-related commodities over time in order to understand how each individual nation utilizes different commodities as part of their energy sector (Small-multiple Line Charts)
3. Gain birds-eye-view of how each state in the United States of America consumes different energy-related commodities in order to understand the prevalence of different commodities in different states and regions (USA Heat-map)
4. Visualize correlation between energy consumption and economic activity in the United States in order to gain insight on how economic activity may or may not be impacted by the United States Gross Domestic Product (Interactive Scatterplot)

Dependencies:

The Energy Viz project will be conducted using the following programming languages, libraries, technologies and services:

- Languages: JavaScript, Python, HTML, CSS
- Libraries: React, D3.js, Pandas, Seaborn, NumPy, Matplotlib, Plotly, Pandas-js
- Technologies: Microsoft Visual Studio Code IDE
- Services: AWS Amplify, GitHub

Deliverables:

1. Energy Viz - Interactive Web-based Visualization Tool
    a. Live Link: https://master.d2gp7frmxpzckb.amplifyapp.com/
    b. React, D3.js, Pandas-js, JavaScript, CSS, HTML
2. Research Notebook - Data Visualization Jupyter Notebook
    a. Python, Pandas, NumPy, Pandas-js, Matplotlib, Seaborn
3. Research Statement Report (4-5 pages)

njitVis | Energy Viz
Building Interactive Web-Based Data Visualization Tool

Visualizing Energy Consumption Patterns in the USA and the World

Parth Merchant | Aritra Dasgupta

Datasets:

1. **Energy Sectors (energy_consumption_by_sector_annual.csv)**
   - Year
   - Primary Energy Consumed by the Residential Sector
   - Total Energy Consumed by the Residential Sector
   - Primary Energy Consumed by the Commercial Sector
   - Total Energy Consumed by the Commercial Sector
   - Primary Energy Consumed by the Industrial Sector
   - Total Energy Consumed by the Industrial Sector
   - Primary Energy Consumed by the Transportation Sector
   - Total Energy Consumed by the Transportation Sector
   - Primary Energy Consumed by the Electric Power Sector
   - Energy Consumption Balancing Item
   - Primary Energy Consumption Total
2. **United Nations Energy Census (all_energy_statistics.csv)**
   - country_or_area
   - commodity_transaction
   - year
   - unit
   - quantity
   - quantity_footnotes
   - category
3. **USA Energy (Energy Census and Economic_Data US 2010-2014.csv)**

   Census & Geographic Data:
   - StateCodes: The state 2-letter abbreviations. Note that I added "US" for the United States.
   - Region: The number corresponding to the region the state lies within, according to the 2010 census. (1 = Northeast, 2 = Midwest, 3 = South, 4 = West)
   - Division: The number corresponding to the division the state lies within, according to the 2010 census. (1 = New England, 2 = Middle Atlantic, 3 = East North Central, 4 = West North Central, 5 = South Atlantic, 6 = East South Central, 7 = West South Central, 8 = Mountain, 9 = Pacific)
   - Coast: Whether the state shares a border with an ocean. (1 = Yes, 0 = No)
   - Great Lakes: Whether the state shares a border with a great lake. (1 = Yes, 0 = No
   - CENSUS2010POP: 4/1/2010 resident total Census 2010 population
   - POPESTIMATE{year}: 7/1/{year} resident total population estimate
   - RBIRTH{year}: Birth rate in period 7/1/{year - 1} to 6/30/{year}
   - RDEATH{year}: Death rate in period 7/1/{year - 1} to 6/30/{year}
   - RNATURALINC{year}: Natural increase rate in period 7/1/{year - 1} to 6/30/{year}
   - RINTERNATIONALMIG{year}: Net international migration rate in period 7/1/{year - 1} to 6/30/{year}
   - RDOMESTICMIG{year}: Net domestic migration rate in period 7/1/{year - 1} to 6/30/{year}
   - RNETMIG{year}: Net migration rate in period 7/1/{year - 1} to 6/30/{year}

   Energy Data:
   - TotalC{year}: Total energy consumption in billion BTU in a given year.
   - TotalP{year}: Total energy production in billion BTU in a given year.

- TotalE{year}: Total Energy expenditures in million USD in a given year.
- TotalPrice{year}: Total energy average price in USD/million BTU in a given year.
- TotalC{first year}−{second year}: The first year's total energy consumption divided by the second year's total energy consumption, times 100. (The percent change between years in total energy consumption.)
- TotalP{first year}−{second year}: The first year's total energy production divided by the second year's total energy production, times 100. (The percent change between years in total energy production.)
- TotalE{first year}−{second year}: The first year's total energy expenditure divided by the second year's total energy expenditure, times 100. (The percent change between years in total energy expenditure.)
- TotalPrice{first year}−{second year}: The first year's total energy average price divided by the second year's total energy average price, times 100. (The percent change between years in total energy average price.)
- BiomassC{year}: Biomass total consumption in billion BTU in a given year.
- CoalC{year}: Coal total consumption in billion BTU in a given year.
- CoalP{year}: Coal total production in billion BTU in a given year.
- CoalE{year}: Coal total expenditures in million USD in a given year.
- CoalPrice{year}: Coal average price in USD per million BTU in a given year.
- ElecC{year}: Electricity total consumption in billion BTU in a given year.
- ElecE{year}: Electricity total expenditures in million USD in a given year.
- ElecPrice{year}: Electricity average price in USD per million BTU in a given year.
- FossFuelC{year}: Fossil fuels total consumption in billion BTU in a given year.
- GeoC{year}: Geothermal energy total consumption in billion BTU in a given year.
- GeoP{year}: Geothermal energy net generation in the electric power sector in million kilowatt hours in a given year.
- HydroC{year}: Hydropower total consumption in billion BTU in a given year.
- HydroP{year}: Hydropower total net generation in million kilowatt hours in a given year.
- NatGasC{year}: Natural gas total consumption (including supplemental gaseous fuels) in billion BTU in given year.
- NatGasE{year}: Natural gas total expenditures in million USD in a given year.
- NatGasPrice{year}: Natural gas average price in USD per million BTU in given year.
- LPGC{year}: LPG total consumption in billion BTU in a given year.
- LPGE{year}: LPG total expenditures in million USD in a given year.
- LPGPrice{year}: LPG average price in USD per million BTU in a given year.

Economic Data:

- GDP{year}{quarter}: The GDP in the provided quarter of the given year (in million USD).
- GDP{year}: The average GDP throughout the given year (in million USD)

njitVis | Energy Viz
Building Interactive Web-Based Data Visualization Tool

Visualizing Energy Consumption Patterns in the USA and the World

Parth Merchant | Aritra Dasgupta

Implementation: The implementation process of building the interactive web-based data visualization tool takes into account the features of a single-page application as well as the notebook style layout of the Jupyter Notebook, which allows you to comment on visualizations and informative data and tell the story about what one can see through each interactive visualization. Using the Jupyter Notebook, throughout the course of this project, we will begin to create visualizations using a variety of different charts and plots to measure attributes in order to determine hypotheses about data, such as correlations and predictions. Using a Jupyter Notebook in Python to validate visualizations that would be included in the final interface in React/D3 is essential to this research project, because the iterative process of building visualizations and implementing them by committing the branch for the React Component to GitHub, which automatically conducts the provision, build, deployment and verification of the code onto a live website.

React Interface: I will be creating a web interface built with React library for a smooth and sleek user interface and experience as well as D3 (Data-Driven Documents) to manipulate document objects to create interactive visualizations. The header on the navigation bar has 5 options:

1. The first is a home landing page, where we will explain our hypotheses and why we are visualizing this data as well as what we're trying to learn from our data.
2. The first visualization tab is going to be the visualization of "Energy Sectors" for annual spending of energy per sector. For this, we will generate an interactive line chart, which will allow us to generate a line chart that allows us to view the energy consumption in a particular sector over time. We are measuring numerical/quantitative attributes over time, so the user will be able to view different interactive lines in different color hues in order to interact with the visualization flexibly.
3. The second visualization tab is the "UN Energy Census", where we will generate small multiple line charts for each commodity in the energy-sector used by a particular nation. The interface will feature a dropdown menu to select which country you want to look at. This will allow us to flexibly switch between countries to see small multiple line charts showing the consumption of each commodity over time.
4. The third visualization tab is the "USA Energy" visualization, where I generate a heatmap encompassing all of the data from the dataset in order to show overall energy consumption for different total commodities (Natural Gas, Coal, Biomass, Hydropower, etc.). This will be a USA Map with color weights on each state and there will be an interactive option that allows the user to hover over each state to see it's consumption.
5. The final visualization tab on the interface is the "USA Energy vs. Economy", where I generate a scatter plot to graph the correlation between Total Energy Consumption and the GDP for each state in the United States in order to understand if energy consumption has any influence on the economy. This visualization will allow us to switch between

different years as well as other numerical attributes related to the economy against attributes of energy consumption.