U.S. Energy Use Trends and Predictions by Source and State 1960 – 2013

Lorelyn Kilby, Jennifer Laws, Rosario Delgado, Annie Ho

December 3rd, 2015

1. Introduction

Currently there is no interactive visual available to see a complete view of energy use by source trends and the resulting carbon footprint of those sources in the United States. Using a dataset from the United States Energy Information Administration, our group has created an interactive map that allows the user to view past, current and predicted energy use measures by source, state and year, in both per capita and total pounds CO_2 emitted. We also included a timeline to control the visualization while viewing the environmental policies enacted throughout. Our goal is to provide and clear window into our nation's energy trends incorporating a wide breadth of dynamic data in a single user-friendly browser.

1.1 Heilmeier Questions

1. Using a public dataset, we have presented energy use by source by state from the 1960 to 2013 and resulting carbon emissions and BTU's per capita, while also displaying energy policies and predicting future trends. By showing the breakdown of various sources (solar, petroleum, etc), we are able to see which states have been shifting towards "greener" energy consumption.
2. Currently energy use is often presented as a snapshot using a single map, using simple graphics like pie charts to show source, as shown by first page of google image results for energy use by source.

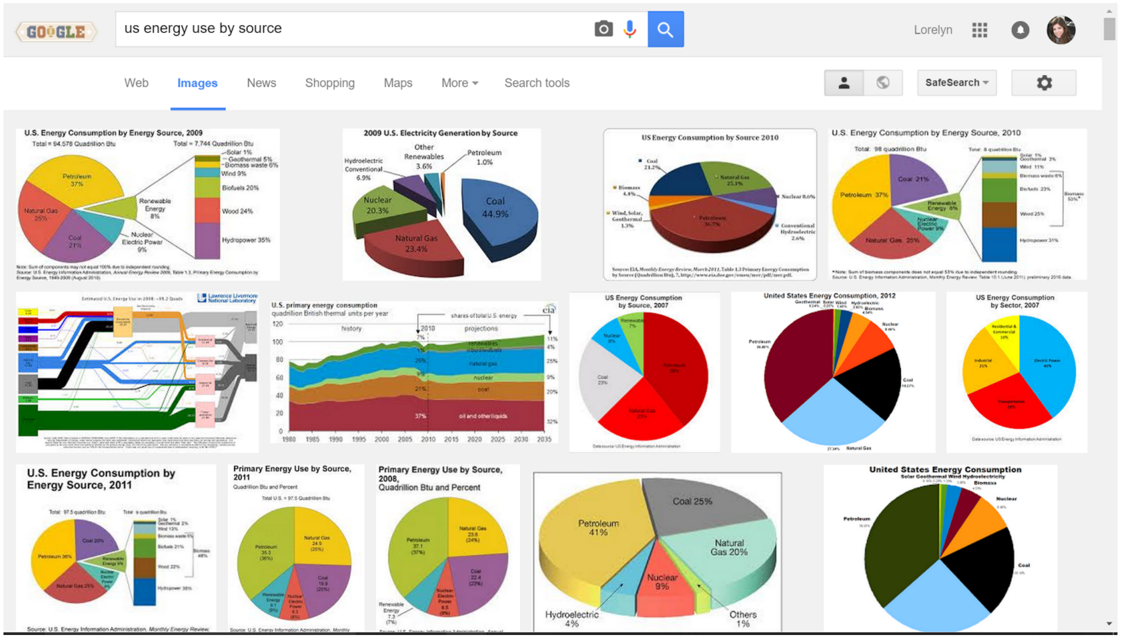

Figure 1: Results of Google Search on energy use by source

3. Little has been done to show the progression of energy use by source over time and the resulting carbon emissions. Our approach presents how these factors have shifted over time and predicts the future direction of energy trends. Our approach is successful because it takes a very complete data picture and combines it with an interactive visual map to allow people to immediately identify trends.
4. Energy consumption has many stakeholders such as the Environmental Working Group (EWG) or the EPA that impose policy on its use.
5. By observing energy trends from a macro view-point, we predict users will be able to see the effects of different policies and the future direction of energy usage.
6. A risk of a map visualization is that it may be overwhelming to the user with so much information. On the other hand, a significant payoff of a well-organized map will be the ability to view so much data at once.
7. While our data is freely available, our costs were associated with the time required to create an appropriate visualization and learn the javascript and specific d3 functions needed to power transitions and data filters, as we are all entirely new to the languages required.

8. The project will take 8 weeks.
9. We initially focused our efforts on data cleaning, then created the functioning prototype with basic display capabilities up to the midterm. We then spent remaining project duration improving our visuals and interactivity.

1.2 Scope and Innovations

By modeling this data, we will highlight which states are moving in a more environmentally sustainable direction, which states need adjust policies, as well as which states are likely to have heavy emissions in the future. A few of our best innovations exhibited in our approach include:

1. Forecasting of energy use
2. Easy to use interactive visual representation of complex data
3. Methodology to analyze the impact (or lack thereof) of energy policies

1.3 Literature Survey

Energy Trends

The sources from which people use energy have changed greatly over the past 50 years. For example, while traditional sources such as oil have increased, entirely new sources of energy such as wind have emerged [Cleveland, C]. However when looking at predicting future use from that source in the short term, energy usage by source tends in increase linearly, thus allowing us to use the last 5 years of data to predict short term future trends[“History”]. However as energy sources have changed over time, the EPA shows emissions have gone up by 7.4% from 1990 to 2013 [“Greenhouse”]. Some sources of energy, such as coal and petroleum emit much more carbon dioxide than others such as wind and solar, and if they want to decrease carbon emissions, moving towards these sources is ideal [Rosenbloom, E]. Society must examine states where energy consumption has grown the least to analyze effective policy measures to control its source and use [Brown, R]. Our visualization will help spot those states exhibiting similar trends.

Politics

The politics regarding energy policy are complex and varied greatly by governmental level and region [Doris, E]. Energy policy typically plays role in shaping which sources of energy increase or decrease in use[Wise,Marshall]. For example, when the Renewable Electricity Production Tax Credit began in 1992, the number of wind energy projects began increase drastically [Sissine, Fred]. Additionally, California has enacted many efficiency programs since the 1970s and as a result, California's per capita energy use has remained flat, while the per capita energy use has increased by about 33% for the rest of the U.S["Energy Efficiency"]. The growing concern of CO2 emissions as a public interest puts pressure on law making bodies to make environmentally sound policies[Bierbaum,R]. While our project will be making no policy recommendations, by displaying policies the year they were enacted and then displaying the resulting energy trends make it easier to identify effective policies.

2. Method

2.1 Data Preparation

To prepare the data we downloaded the energy use source dataset freely provided by the United States Energy Information Administration, and then used Google refine and Excel to clean the data. The data eventually become a clean csv file that clearly shows by state and year the energy in BTU's from each source. The following shows filtered results of a few sources from Alaska and Alabama:

State	MSN	1960	1961	1962	1963	1964	1965
AK	Biomass	3681	4145	4246	4383	4728	4863
AK	Coal	7189	11552	13559	11999	12029	9888
AK	Hydropower	3120	3168	3207	3410	3374	3655
AL	Biomass	45681	46171	45959	45764	46341	47641
AL	Coal	395400	382455	427177	442465	459000	533142
AL	Hydropower	67128	71851	79605	66918	91651	74247

Figure 2: Filtered portion of energy use by source by state by year dataset

We then gathered a dataset of population by state per year, a dataset showing the pounds of CO_2 emitted per source and a dataset showing enacted environmental policies using EPA internet resources.

2.2 Calculations

Several computational steps on the original dataset were required in order to show the desired information. The energy use predictions were calculated before activating the visual, while the energy use per capita and pounds of CO_2 emitted are calculated dynamically within the visualization.

To calculate the predicted energy use by source and state, we used linear regression using the panda library in python looking at the previous 10 years of data to predict the energy use sources for the years 2014-2020. The pandas library implements the following formula:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j}{\sum_{i=1}^{n}(x_i^2) - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2}$$

$$= \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\operatorname{Cov}[x, y]}{\operatorname{Var}[x]} = r_{xy}\frac{s_y}{s_x},$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\,\bar{x},$$

This python script creates two csvs, one storing the regression coefficients for reference (A total of 550 regressions, one for each state and source combination), and another with the actual predicted values. We then merged the predicted data points with the original data and read that file into the javascript. We performed a similar procedure to estimate population. We chose a linear regression because it is most accurate for a wide range of data and the datasets we are using followed a semi-linear trend. In future works we would consider a closer examination of various regression fits.

2.3 D3/Topojson Visualization

To visualize our calculated information, we used a geomap library called datamaps that reference topojson. We hosted our project in a repository on GitHub, allowing all members to access and collaborate on its development. The main deliverable of our project is built in the index file where the javascript powers our d3 visualization, incorporating bootstrap css for page styling. After opening all required CSV's, the script grabs the current values for the timeline object , select object, and toggle switch designating whether to view data as pounds of CO_2 emitted per BTU or BTU usage per capita. The script then builds the necessary data type required for the datamap. During this process, the per capita energy use is calculated by dividing the BTU value by each state's population in the given year, returning an array of the {'state', 'choropleth color' }, where the color is calculated from a linear d3 scale. If the user has selected to view as co2 emissions, it multiplies the BTU value by that energy source's factor (x pounds C02 per BTU), loaded from a dictionary defined on page load. When a dom element is changed the script rebuilds a new set of data and the map is updated. Each source is presented in a different color for easy distinction. The figures below show current screen shots of our product. The darker the shading of the state, the higher the usage of that energy source, or in the case of the carbon emissions, the more carbon emissions. Within the timeline, the user may click the policy title and be routed to another web-page giving further details of the policy or act. The title is dynamically updated to summarize the selected inputs.

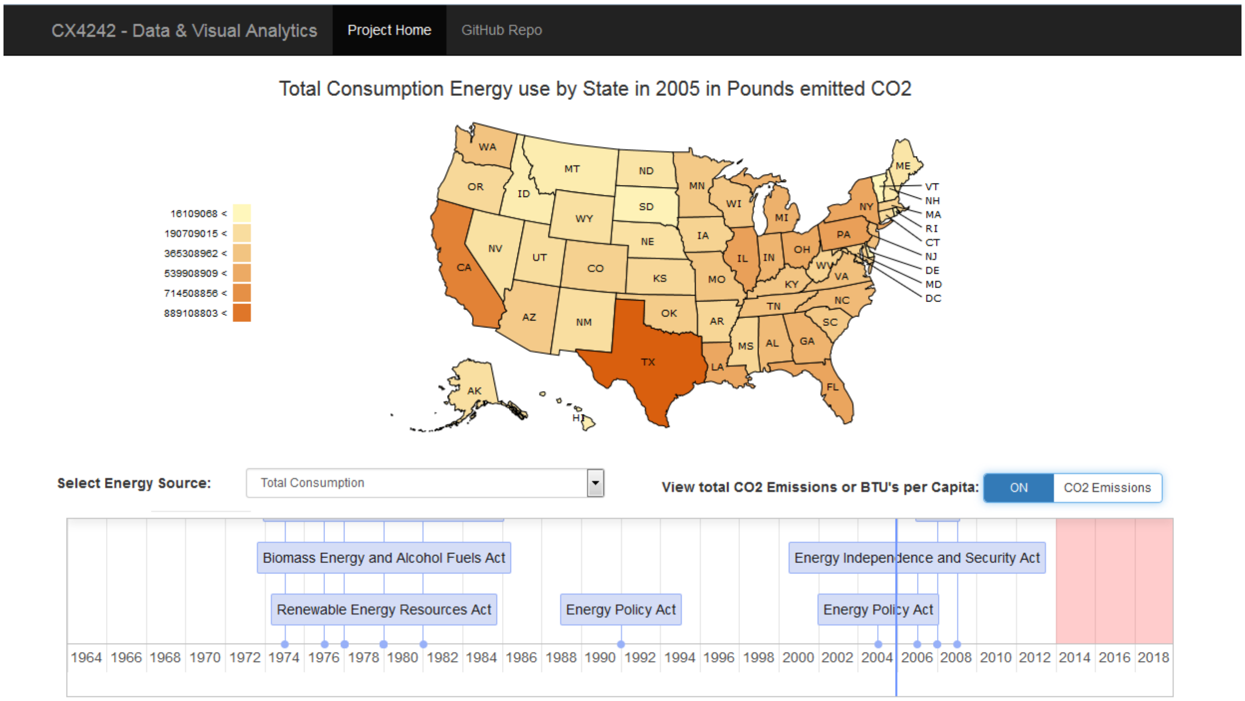

Figure 3: View of visualization map of Total Consumption Energy Use and total resulting lbs of CO2. Darker state indicates higher CO2 emissions.

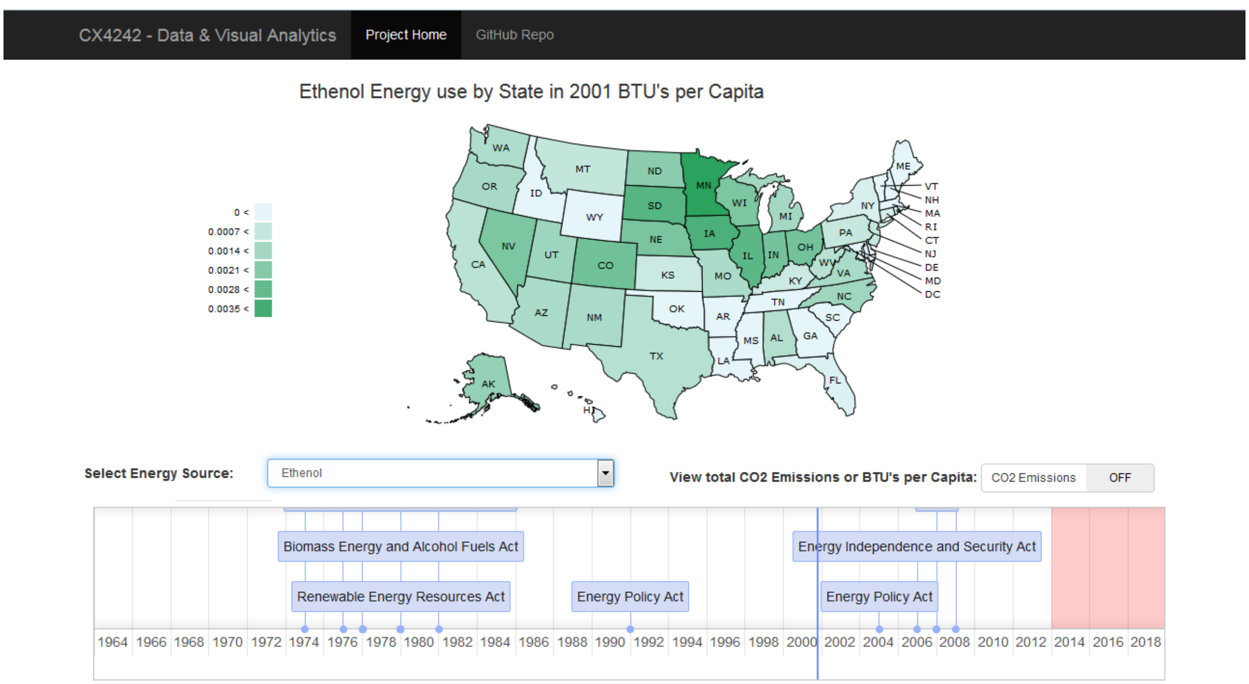

Figure 4: View of visualization map of Ethanol use by state. Darker color indicates higher usage

3. Experiments and Observations

To validate our approach, we wanted to answer the following questions:

1. How valid is our prediction model?
2. What sorts of observations can we make about energy trends from our analytic visualizations.

3.1 Experiments

In order to test the accuracy of our predictions we ran our regression prediction on 1970, 1980, 1990 and 2000 using the 10 years prior to the calculated year to predict energy usage. After running these regressions, the predicted results were on average within 21% of the actual values.

Year	1970	1980	1990	2000	**Total**
Result Delta	.072	.220	.321	.231	**.211**

3.2 Observations

One of the main objectives of this visual was to be able to see the trends in energy use: which states have been using more of particular sources, and which appear to be adopting methods of alternative energy.

Using our visualization we were able to make the following observations:

1.

Energy Source	**Observations**
Biomass	In the Midwest, we observed an influx use, changing from 0.026 BTU person/year to 0.0625.

Coal	The choropleth changes only minorly showing that its usage per capita has plateaued.
Ethenol	Usage has dramatically increased across the entire US since the 1980’s and almost all states use above .003 BTU person/year currently.
Geothermal	Usage is heavily weighted on the west coast due to tectonic plate movements
Hydropower	Across the U.S. and heavily in the Northwest we see a reduction in per capita use of hydropower due to increased populations with little increase in production.
Natural Gas	We do not see major changes per capita, however it is clear though co2 emissions texas heavily relies on natural gas
Nuclear	Once most US nuclear plants were established and running (1970-1990), use per capita stays relatively constant due to difficulty of approval process
Petroleum	Follows same trends of Natural Gas
Solar	We don’t observe any recorded solar energy use until the 1990s, following the implementation of the Energy Policy Act.

2. From these observations we gathered that the states that have moved in a more sustainable, renewable direction were typically restricted to the western part of the U.S. with a few exceptions. States with high usage per capita of solar energy are California, Arizona, Florida, Nevada, New Mexico and Colorado and Florida. States with high geothermal usage are California and Nevada. Hydropower was typically concentrated in the Northwest in Washington, Oregon, Montana, and Idaho. These states typically had much lower values for total consumption energy use per capita than other states without high usage of these sources.

4. Conclusion and Next Steps

4.1 Future Work

While we were very satisfied with our end product and visualizations, the impact of environmental policies on energy use was not as apparent as we predicted. In future research we would need to closely examine each policy to quantitatively determine which policies achieve their desired impact most effectively. We were able to see overall energy use trends over time as well as which states were heading in a more green direction. In the future, it could be helpful to visualize more granular data such as use by county.

There are some aspects of our visualization we would want to expand on going forward, such as ensuring better cross browser and device compatibility, better tooltip placement as well as better integrated policy information.

4.2 Summary

Using freely available data, javascript, d3 and topojson map data we created an effective map visualization tool that displays both the historical and predicted energy usage and resulting carbon emissions trends. This approach is unique in its complexity and ability to present a wide breadth of information, both static and dynamic in a single user interface. We hope that having such a clear visual on the effects of a state's energy use will influence policy makers and individuals to move towards greener policies and will open the door to further investigation of this topic and expansion of the current model we have created.

Contribution Statement: All team members contributed a similar amount of effort

Appendices

Appendix A. Timeline - From Progress Report

Name | Deadline	11/20	11/27	12/3
Jennifer Laws	Final Report Draft	Improvements to Report	Report Due
Lorelyn Kilby	Finish adding all javascript functionality	Apply CSS to product	No changes to product after 11/27, help with report and poster
Rosario Delgado	Poster Template	Finalize poster	Poster Presentation
Annie Ho	Poster Template	Finalize poster	Poster Presentation

Appendix B. Heilmeier Questions

1. What are you trying to do? Articulate your objectives using absolutely no jargon.

2. How is it done today, and what are the limits of current practice?

3. What's new in your approach and why do you think it will be successful?

4. Who cares?

5. If you're successful, what difference will it make?

6. What are the risks and the payoffs?

7. How much will it cost?

8. How long will it take?

9. What are the midterm and final ”exams” to check for success?

References

1. Bierbaum, Rosina M., William K. Reilly, Richard L. Revesz and Richard A. Meserve. Bulletin of the American Academy of Arts and Sciences. Vol. 61, No. 2 (Winter, 2008), pp. 29-41
2. Brown, R., & Koomey, J. (2002). Electricity use in California: Past trends and present usage patterns. Energy Policy, 849-864. http://enduse.lbl.gov/Info/LBNL-47992.pdf Retrieved October 6, 2015.
3. Cleveland, C. "Net Energy from the Extraction of Oil and Gas in the United States." Energy (2005): 769-82. Web.
4. Doris, E., J. Cochran, and M. Vorum. Energy Efficiency Policy in the United States: Overview of Trends at Different Levels of Government. Tech. NREL/TP-6A2-46532. National Renewable Energy Laboratory, Dec. 2009. Web. 4 Oct. 2015.
5. Energy Efficiency. (n.d.). Retrieved November 13, 2015, from http://www.cpuc.ca.gov/PUC/energy/Energy Efficiency/
6. "Greenhouse Gas Inventory Data Explorer." *US EPA United States Environmental Protection Agency*. US EPA, 4 Apr. 2015. Web. 13 Nov. 2015.
7. "History of Energy Consumption in the United States" Today in Energy. U.S. Energy Information Administration - EIA - Independent Statistics and Analysis, 9 Feb. 2011. Web. 13 Nov. 2015.
8. Ilmakunnas, P. and Törmä, H. Journal of Applied Econometrics. Vol. 9, No. 3 (Jul. - Sep., 1994), pp. 305-320
9. Rosenbloom, E. (n.d.). Wind and the Mitigation of CO2 Emissions -- The Global Picture. Retrieved November 13, 2015, from http://www.aweo.org/windIPCC.html
10. Sissine, Fred. "Renewable Energy Policy: Tax Credit, Budget, and Regulatory Issues" (PDF). *CRS Report*. Retrieved 24 March 2011.
11. Sherwood Larry. "U.S. Solar Markets Trends." Interstate Renewable Energy Council and Sherwood Associates (2010). Web. 4 Oct. 2015.

12. “U.S. Energy Information Administration – EIA – Independent Statics and Analysis.” Energy Sources Have Changed Throughout History of the United States. Web. 5 Oct. 2015.