Energy Empires

Fiona Heaney, Rachel Murphy,
and
Shivangi Pandey

Submitted 03 March 2017

Name	GitHub ID	Email
Fiona Heaney	fhheaney	fhheaney@wpi.edu
Rachel Murphy	rlmurphy2018	rlmurphy@wpi.edu
Shivangi Pandey	shivangipandey11	spandey@wpi.edu

Overview

Demonstrate the knowledge acquired pertaining to D3 and basic data visualization principles over the duration of our seven week CS 4802 course

Goal & Motivation

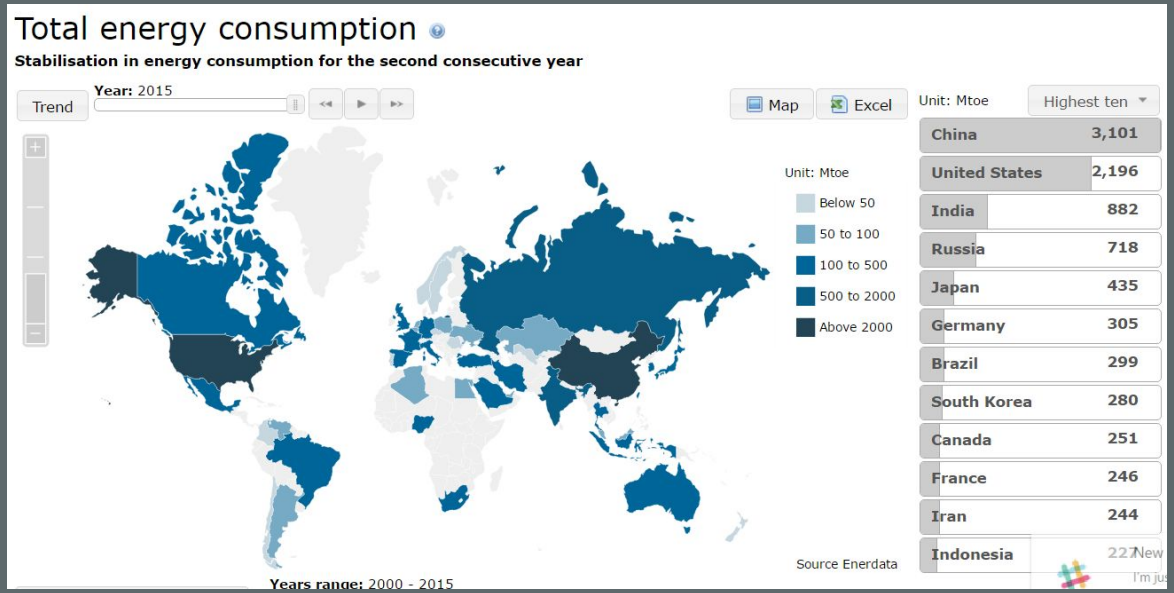
GOAL

Tie together political history of twenty countries along with energy consumption data from 1980 - 2014 to highlight major shifts in trends

MOTIVATION

Vast number of energy consumption data sets but inability to find visualizations that looked beyond annual country vs country usage.

Related Work - Enerdata



- Year by year choropleth with country specific sidebar
- Great for single year comparisons but no longitudinal trends

Related Work - OECD Bar Charts

- Wanted to have similar display for global overview small multiple displays
- On hover, grey out all other bars except the given country on each resource bar chart

Questions

Initially, with the potential to analyze not only global energy data sets, but climate and other potential energy affecting variables our questions were focused on detecting variables that influenced energy trends.

As the data sets were narrowed down the visualization was redesigned to focus on cohesively showing the longitudinal energy usage trends with a country to country comparison. As the design changed, the questions evolved to focus on the change of energy trends over time, specifically looking to draw attention to how individual countries have utilized various energy types over time.

Questions the Final Design Focuses On

- How have energy consumption trends changed over time?
- Globally, what are the trends for energy consumption?
- Does the type of energy consumed vary by country development?
- Has the type of energy consumed changed on a longitudinal scale?

Data

The data used for this project came from two sources. The countries were selected by human development index based off of data provided by the United Nations. The energy consumption data was provided by The Shift Project. The website allowed for energy consumption for individual countries to be downloaded as csv's. Two csv's per country were downloaded, one had data for the main energy types (coal, oil, gas etc) and the other provided a specific breakdown of the renewable energy resources that the country utilized (solar, geothermal, biodiesel, etc).

Links to the Data Sources

- Human Development Index

  http://hdr.undp.org/en/content/human-development-index-hdi

- Energy Consumption by Type

  http://www.tsp-data-portal.org/Energy-Consumption-Statistics#tspQvChart

Data

Countries Selected For Analysis by HDI

- Very High Development
  - Norway
  - Netherlands
  - Australia
  - Saudi Arabia
  - USA
- High Development
  - Uruguay
  - Belarus
  - Romania
  - Russian Federation
  - China
- Medium Development
  - Moldova
  - Turkmenistan
  - Egypt
  - Indonesia
  - India
- Low Development
  - Burundi
  - Sierra Leone
  - Niger
  - Guinea
  - Mozambique

Data

Countries Selected For Analysis by HDI

- Common Energy Types
  - Coal
  - Gas
  - Hydro
  - Nuclear
  - Oil
- Breakdown of Other Renewable Sources
  - Biodiesel
  - Biomass
  - Ethanol
  - Geothermal
  - Solar / Tide / Wave
  - Wind

Data Cleanup

Because The Shift Project allowed for data to be downloaded for individual countries, the data cleanup process was fairly simple.

CSVs with data for the common energy types of each country were downloaded and using excel, the totals for each type were calculated and saved into a separate CSV. This separate CSV contained the total energy consumption per type for each country. This file was used for visualizations involving the total energy consumption over the 34 year span.

Additional CSV's with the breakdown of the other renewable energy sources were downloaded for each country. The two CSV's for each country were then combined, allowing for these CSV's to provide data for the yearly breakdown of energy consumption for each type (including the renewables).

To make data processing easier, CSV's for each country were combined to make the small multiples implementation easier.

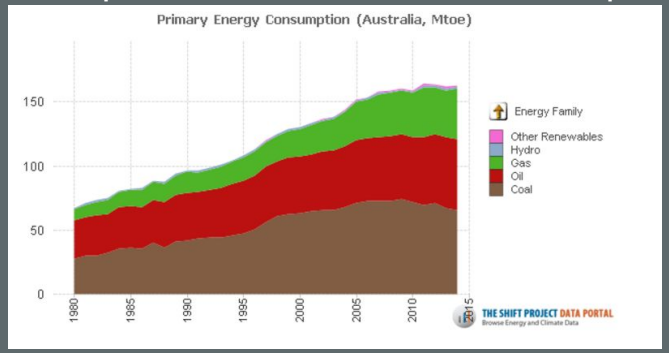
Exploratory Data Analysis

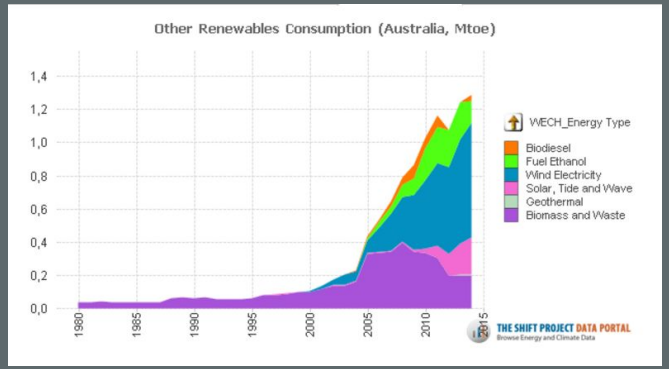
The data from The Shift Project was visualized as stacked line graphs. We used the provided graphs to initially analyze the data.

Some initial insights provided from this analysis were:

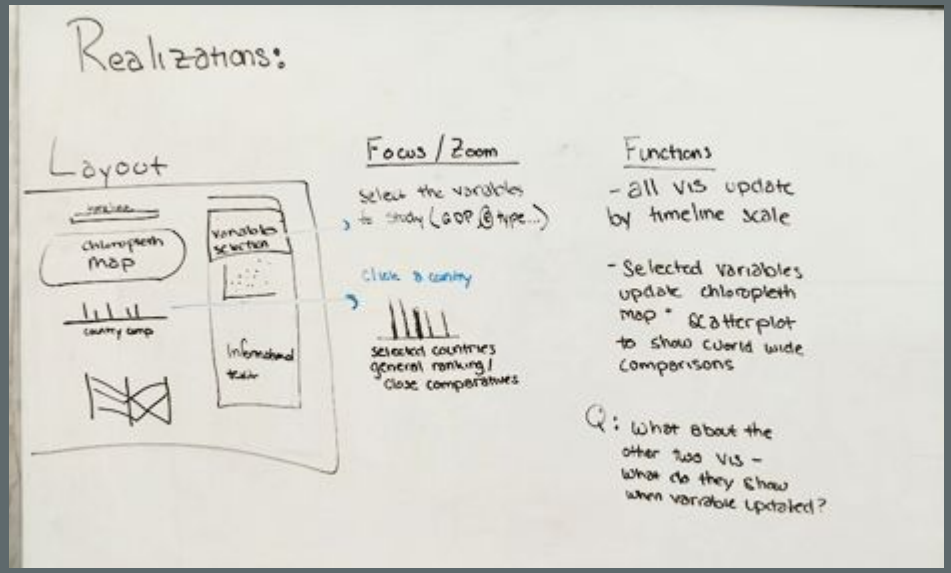
- There was a jump in the consumption of renewable sources between 2000 and 2005 in many countries
- The lower developed countries relied almost solely on one energy type
- In all cases, energy consumption has increased over the time span

Example: Australia Stacked Line Graph





Design Evolution - Proposal



The initial proposal was designed to have multiple visualizations that would all be updated by a timeline slider. This would allow the viewer to interact and discover how the energy trends change over time.

The left side would have a timeline slider, choropleth map, zoomable bar graph, and parallel coordinates. The right side would have a drop down menu for variables, scatter plot, and text box to provide facts about the data presented.

The choropleth map and scatterplot would also be updateable by selecting a variable to analyze from a dropdown menu.

Design Evolution

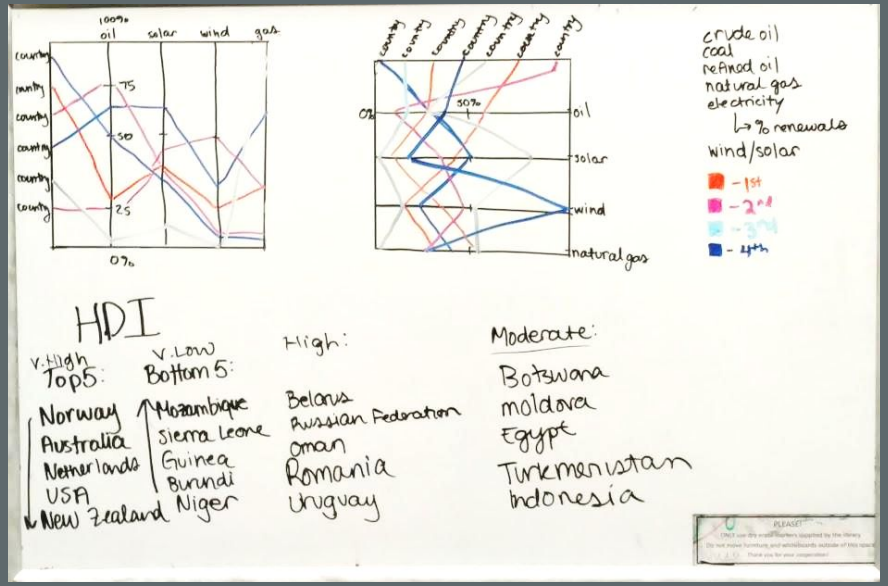
Between the initial proposal and the prototype presentation, the following design changes were made:

- Countries should be ranked by development
  - Originally, countries were going to be ranked as 1st world, 2nd and 3rd world. However, this ranking system does not accurately represent a country's development. This ranking system was created after WWII and groups many countries together. The Human Development Index accounts for a country's GDP, birth rate, death rate and education levels to provide a more accurate categorization of a country's development level.
- Variables to be analyzed should be more focused
  - Just energy consumption by type should be analyzed instead of looking at climate and seasonal variables. By focusing the data, we hope to provide more of a goal or story to the visualization instead of just pure exploration of different variables and their effects.

Design Evolution - Prototyping

The prototype presented to class had the following visualizations:

- Stacked bar chart
  - To show the breakdown of each energy type per country
- Choropleth map
  - To show the countries colored by HDI ranking and/or color by energy consumption levels
- Timeline slider
  - To update the data shown by year
- Parallel coordinates
  - To show the breakdown of energy type consumed for each country. This would possibly allow for a scrolly story implementation to point out major energy trends.



Design Evolution

Between the prototype presentation and the final design, the following design changes were made:

- Data Source Changed
  - Originally, the data was going to be collected from EnerData, however, data for only 45 countries was provided for free. Additionally, the data provided by EnerData was not consistent in formatting, some values were measured in Mtoe while others were given as percentages, this would have to be fixed by a great deal of data cleanup. Instead, data collection was switched to The Shift Project.
- Number of Countries Minimized
  - Originally, the data was going to cover every country. But because of the fees for EnerData's data and the time consumption that it would take to individually download each country's data from The Shift Project, we decided to limit the countries to 20. We picked 5 countries from each of the HDI categories to provide a global representation of many different global development levels.
- Small Multiple Views
  - Feedback from the prototype presentation lead us to the idea of using small multiples to allow all the countries data to be compared to each other at once.

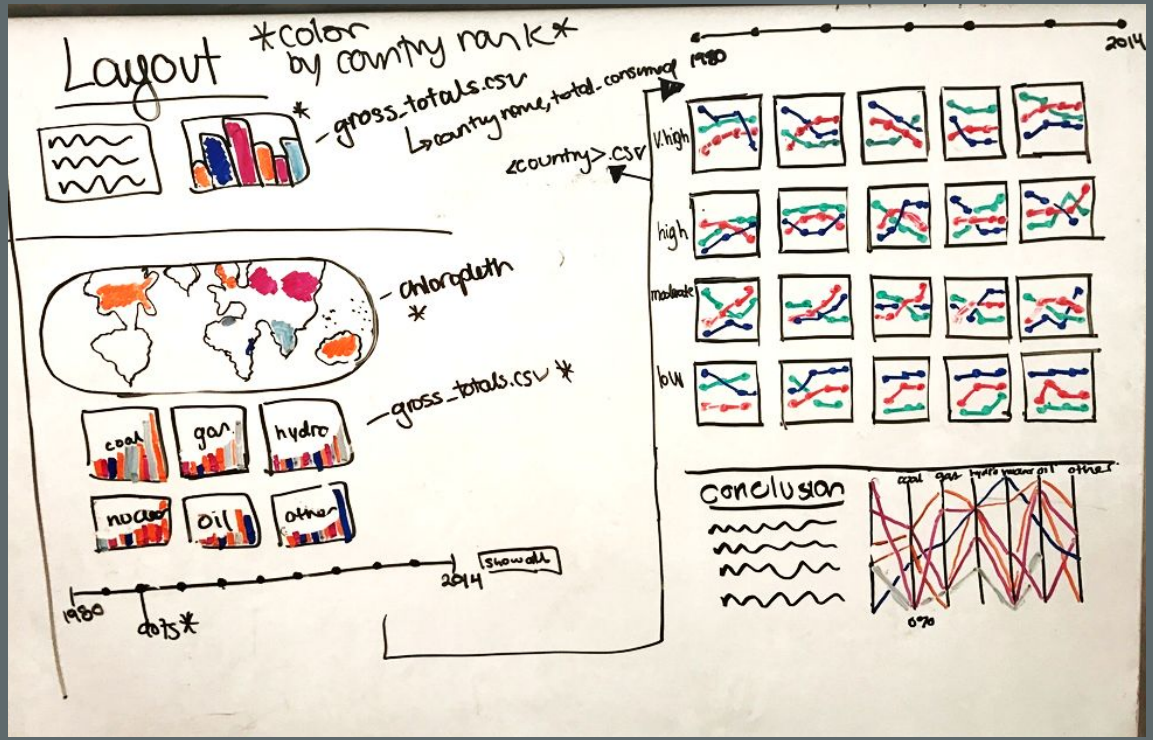
Design Evolution

After the prototype presentation and change in data source, the final design was changed.

The final design would include:

- Bar graph
  - To show total energy consumption by country.
- Choropleth map
  - To show country location and help orient the user.
- Linkes small multiples bar charts
  - To show the breakdown of the bar graph by energy type.
- Timeline
- Linked small multiples line charts
  - To show the energy trends by type for each country over the years
- Parallel coordinates
  - To provide a summary of how each country depends on different energy types.

Design Evolution - Final Design

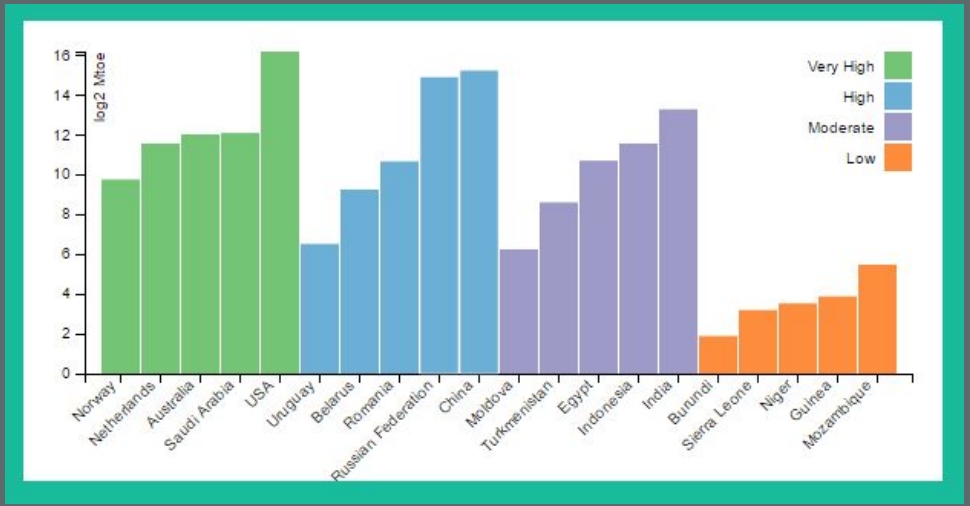


Implementation

The first visualization is a bar chart, grouped and color coordinated by HDI (as described in the legend).

This bar chart shows the total energy consumed by each country over the 34 year span. This data was chosen to initially be shown to help introduce users to the concept that the more developed a country is, the more energy they consume.

The user can interact with this visualization by highlighting over a bar to get the specific Mtoe value for the country.



Implementation

The second interactive visualization is the choropleth map. Each of the 20 countries from our data set are colored. The 4 colors used correspond the the 4 development categories:

- Very High (green)
- High (blue)
- Medium (purple)
- Low (orange)

These colors are consistent with the ones presented in the initial bar chart.

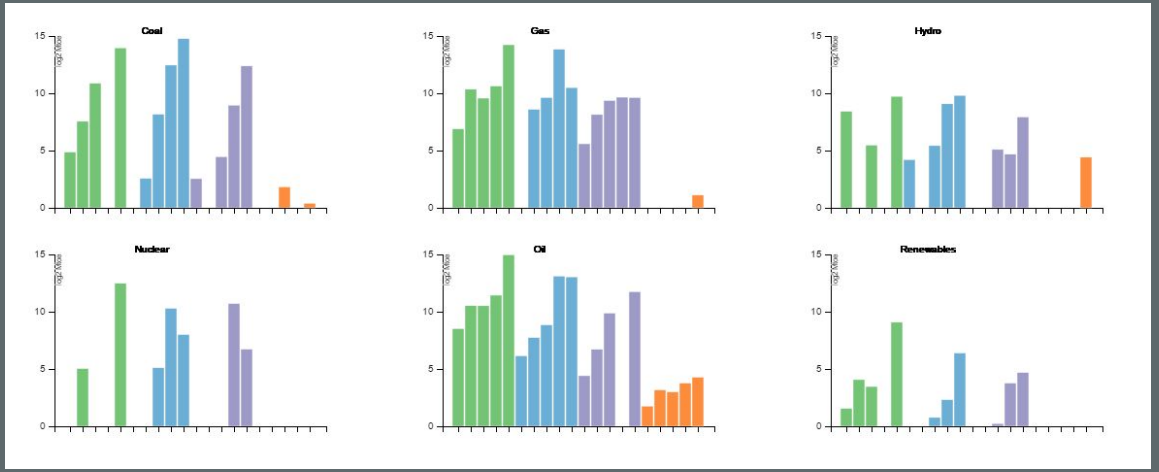
When the user hovers the mouse over a country, the name appears. This helps users orient the selected countries to their global location.

Implementation

The third visualization is bar graphs showing the breakdown of the total energy consumption by energy type. This was implemented as a small multiple view visualization. When one bar is highlighted over, the country name and the consumption amount (in log2 Mtoe) pop up as a tooltip.

Additionally, because this is a small multiple view, when one bar is highlighted, all corresponding country bars in the other views will highlight as well.

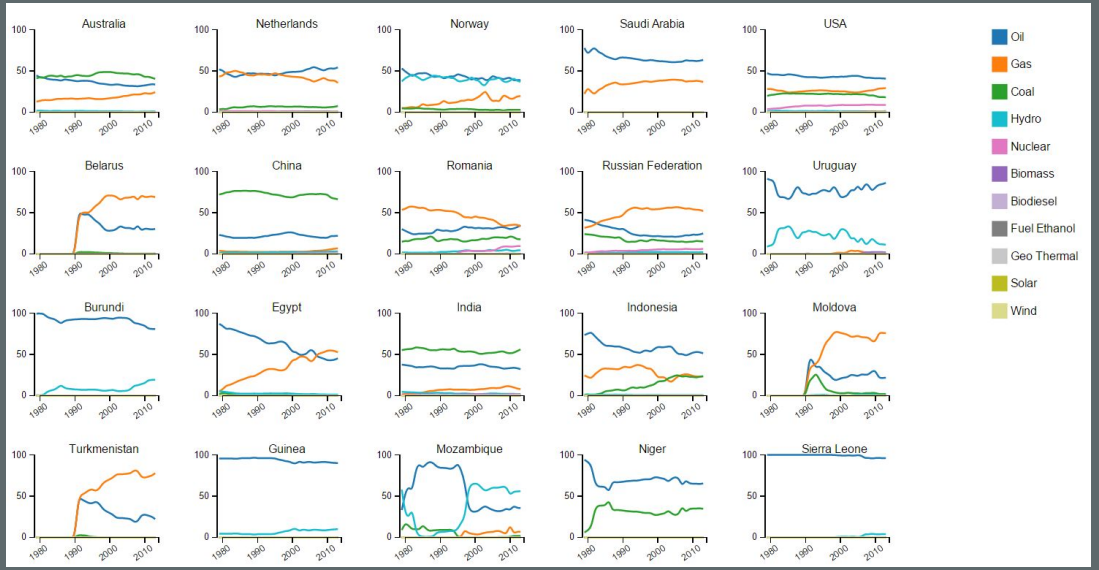
This visualization provides further analysis of the initial bar chart, showing how the total energy consumption can be broken down by type. The intent is to show the user that some energy types are relied on more than others.



Implementation

The third visualization is a multi-line chart showing energy consumption year to year, with each line representing a different resource type. We had hoped to implement a hover event that would show resources and their percentages on a hover with a bisector to the x-axis but we were unable to do so.

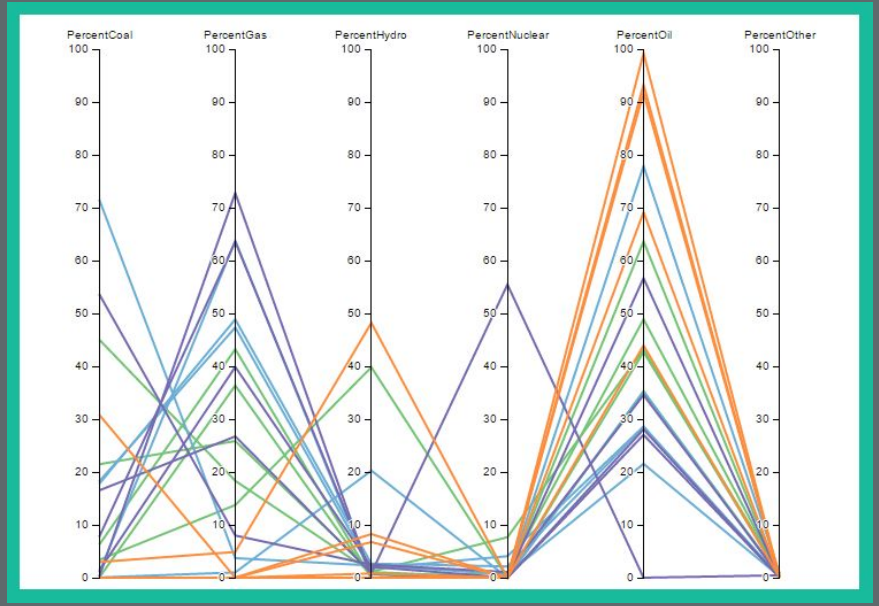
This visualization is useful because we can see increases or decreases of energy type usage over time. We had hoped to implement a timeline so that hover events would highlight related events on the timeline, but implementation was not ready in time. This visualization still enables users to see when renewable energies begin to be implemented or decrease.



Implementation

The conclusion section contains a parallel coordinates visualization. Each country (color coded to match its development index) is represented by a line. Each axes represents the percentage of the country's total energy that is consumed for each energy type.

The user can hover over a line to get the country name as well as its HDI ranking. Additionally, the axes can be rearranged to make comparison of two energy types easier. A brush feature is included, so the user can select specific values of one axis and only the countries that meet those values will be shown.



This visualization provides another way to look at how individual countries consume different types of energy. Unlike the small multiple bars, this visualization was implemented to demonstrate trends in countries dependencies on various energy types.

Evaluation & Analysis

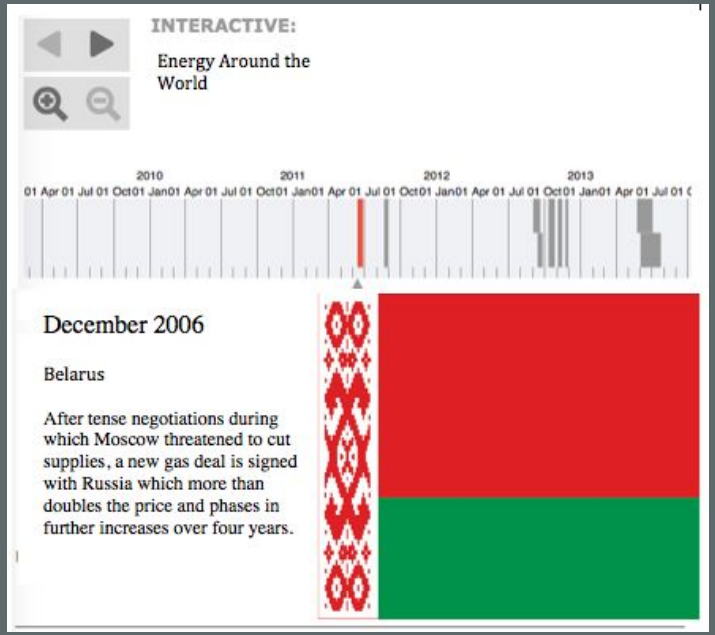
From the data in our visualizations, we learned that the highly developed countries have continued their use of oil and gas for the past thirty-four years; consumption rates are high in comparison to the high, moderate, and low development countries. We also found that few countries have taken the charge to implement more clean energy sources; this may be due to many countries already having defined infrastructures and reluctance to adopt new resources.

Our visualizations allowed for us to answer questions about our data. We used the small value multiples line charts to see how energy consumption had changed in the thirty-four year span for the 20 countries. The bar charts of specific energy renewables allowed us to explore the rates of the energy consumption for each country specifically. And it allowed us to explore how each country's development rate correlated to the energy consumption rate. The parallel coordinates graph allowed us to further visualize the percentage of energy consumption in relation to specific resources.

We felt that our visualizations worked fairly well, considering the amount of time we had to work on the project. Had more troubleshooting time been available, we would have been able to add more to the map, such as highlighting data from other countries throughout the visualizations when a single country is selected. If we had more time, we would have finished implementing our timeline to the visualizations to add to the political, social, and economic factors of energy consumption to draw more conclusions from our story.

Future Work

- Implement a Timeline
  - A timeline would include events and government policies that align with the energy data trends. The user could interact with the timeline to look at specific years in which they notice trends to discover potential causes. Another option for the timeline would be to implement it as a scrolly story to help guide the user through the analysis and linkage of trends to events.
  - Examples of what would be included in the timeline:
    - In 2008, the US increased activity in the Wind Energy Technology Office (WETO), focusing on increasing energy consumption from wind and hydropower.
    - https://shivangipandey11.github.io/DataVisFinal/index.html



Future Work

- Add in data for more countries
  - Include data for all countries, however it is important to keep the visualization clean and easy to understand, so some of the visualizations (such as the small multiples) may be better suited to only show a selected few countries.
- Link the visualizations to each other
  - Having the visualizations work together, so that a user's interaction in one visualization updates the others would be a possible future application.
- Use different ways to show energy consumption
  - Instead of measuring energy consumption in Mtoe, have at least some visualizations that take into account a country's population. These visualizations could show a country's energy consumption per person, making it easier to compare countries of varying sizes.

References

- Data set - http://www.tsp-data-portal.org/Energy-Consumption-Statistics#tspQvChart
- Small Multiples - http://flowingdata.com/2014/10/15/linked-small-multiples/
- Small Multiples- by Mike Bostock - https://bl.ocks.org/mbostock/1157787
- Nesting Data Structures- by Learn JS Data - http://learnjsdata.com/group_data.html
- Nesting Data Structures - by CLHenrick - http://clhenrick.io/d3-v4-general-update-pattern-punchcard-chart/
- World Map - https://datamaps.github.io/
- Energy usage by resource type & country energy usage - http://flowingdata.com/2014/10/15/linked-small-multiples/
- Parallel Coordinates in D3 - by Jason Davies - https://bl.ocks.org/jasondavies/1341281
- Timeline in D3- by Chris-Creditdesign - http://codepen.io/chris-creditdesign/pen/yuFjr?editors=1111
- BBC Country Facts- http://www.bbc.com/news/world-africa-13087604 (insert continent and specific country ID)