Global CO_2 Emissions and Energy Usage

CS171 Final Project Proposal

Rachael Smith '16 and Anna Zhong '16

Background and Motivation

Global warming is a reality of the world we live in. Multiple lines of scientific evidence show that sea levels and temperatures are rising and, as a result, making environments uninhabitable for an increasing number of the world's organisms. Global warming also impacts humans by loss of habitat from inundation and by threatening food security by decreasing crop yields. While there is skepticism over the human contribution to global warming, with some believing the current situation of global warming to be a part of natural historical changes of climate, the Intergovernmental Panel on Climate Change (IPCC) reported in 2014 that scientists were more than 95% certain that most of global warming is caused by increasing concentrations of greenhouse gases and other human activities. To that end, we decided to explore the correlations between human activities and the volume of global CO2 emissions over the years. We intend to span over various indicators, such as population, GDP, energy usage (fossil fuels and alternative), energy production and deforestation (by forest area), in order to seek a comprehensive understanding of the various human factors contributing to carbon emissions that in turn might contribute to global warming. We want to clarify that this project is not about the scientific aspect of climate change—we will not be comparing temperature, sea level or ocean acidification data. We also do not claim that any correlations shown in the visualization necessarily point to direct causal relationships. We are more interested in exploring how humans might have contributed to global warming, as well as visualizing the pace and scale with which we have done so. As such, we are expecting to visualize some interesting, if not alarming, data, and we hope that this will educate viewers on the current global carbon emissions situation.

Project Objectives

Ideally, our visualization will consist of three "chapters." We consider it essential that we implement the first "chapter," we are hopeful that we will implement the second "chapter" and it would be outstanding if we could accomplish the third.

The first chapter tells the story of the *history of global carbon emissions*. Our hope is to educate users on how emissions have changed over time, and how these changes have correlated with economic growth, population growth and deforestation. In this chapter, we are mainly interested in seeing exponential growth in all of these factors within the past twenty – thirty years, although we are aware that this doesn't necessarily prove any causation between the factors.

The second chapter tells the story of *global carbon emissions today*. Here, we hope to answer questions about how the income of each country, as well as its geographical region, affects its current emissions. We intend to look at the relative amounts of total emissions per country, as well as per capita emissions. Users will be able to visualize both the geographic distribution of emissions today, as well as the economic distribution of emissions.

The third chapter will tell story of *carbon emissions in the USA today*. For instance, we think it would be interesting to visualize the breakdown of emissions by industry. More research will be necessary to determine the precise nature of this chapter, as this will depend on what (if any) relevant data is available (See "Data").

The ideology is that each chapter will progress to an increasingly more specific set of data to visualize. Each previous chapter serves to contextualize the restricted focus of the subsequent chapters.

Data

We are using data from the World Bank, which can be found at the links listed below. Note that all data sets have data for 248 countries from 1960 – 2012, as well as metadata including each country's geographical region and income group.

For the first chapter, we will use the following data sets. Refer to "Visualization" for how we will use them:

- CO2 Emissions (kt): http://data.worldbank.org/indicator/EN.ATM.CO2E.KT
- Forest Area (% of land area): http://data.worldbank.org/indicator/AG.LND.FRST.ZS
- Energy Use (kt of oil equivalent): http://data.worldbank.org/indicator/EG.USE.COMM.KT.OE
- Alternative and Nuclear Energy (% of total energy use): http://data.worldbank.org/indicator/EG.USE.COMM.CL.ZS
- Population, total: http://data.worldbank.org/indicator/SP.POP.TOTL
    - Here, we could break this into other possible data sets from the World Bank that show population in rural and urban areas.
- GDP (current $US): http://data.worldbank.org/indicator/NY.GDP.MKTP.CD

For the second chapter, we will use the following data sets:

- CO2 Emissions (kt): http://data.worldbank.org/indicator/EN.ATM.CO2E.KT
- CO2 Emissions (metric tons per capita): http://data.worldbank.org/indicator/EN.ATM.CO2E.PC

For the third chapter, the World Bank provides separate data sets for a number of contributing factors to carbon emissions. These include electricity and heat production, gaseous fuel

consumption, liquid fuel consumption, manufacturing industries and construction, residential buildings and commercial and public services, solid fuel consumption, transport and "other". To specifically visualize which industries or contributing factors the United States' carbon emissions are coming from, we will just scrape data for the United States from each of these data sets.

We also want to find relevant data that contains relationships between countries (such as the trade relationships in homework 2), because we like the circular layout with links, and the use of hierarchical bundling. However, this is proving quite challenging in the context of global carbon emissions since there is no reasonable "link" to make between the emissions or energy usage of individual countries. Any suggestions for other relevant correlations are welcome in the feedback for this proposal.

Data Processing

We hope to keep this step simple, and we are fortunate that the World Bank provides all of its global indicators in separate data sets, such that in order for us to include the data for any indicator, we would only need to load that corresponding data set. The data sets we are using generally only provide the information for the indicator in question, so we do not have to clean out much of the data. However, the data is only available in Excel, CSV or XML formats. While it won't be difficult to load an XML file using Javascript and transfer and reorganize all of its information into Javascript variables within our code, we might want to consider whether this will slow down the visualization. If we have time during implementation and if we are finding that our visualization is too slow, we could write a Python program to convert the XML data and write it into a JSON file of our desired format, which we would just load straight into our visualization. Either way, we will have to parse the XML data into custom objects, which gives us a lot of freedom in how we want to organize the information. In particular, it would be helpful if we compiled all of the indicators together under their corresponding countries and years, since all data sets have the same 248 countries and span the same years (1960 – 2012). The metadata for each country is also the same for all data sets, which is convenient. An example of this would be:

```
{
    “country”: “Australia”,
    “income”: “high”,
    “geographical region”: “Pacific”,
    “years”: [
        “1960”: [
            “CO2_total”: 1923844709,
            “CO2_per_capita”: 273,
            “forest”: 60,
            “energy”: 17283019201191,
            “alternative_energy”: 17
            ...
        ]
    ]
}
```

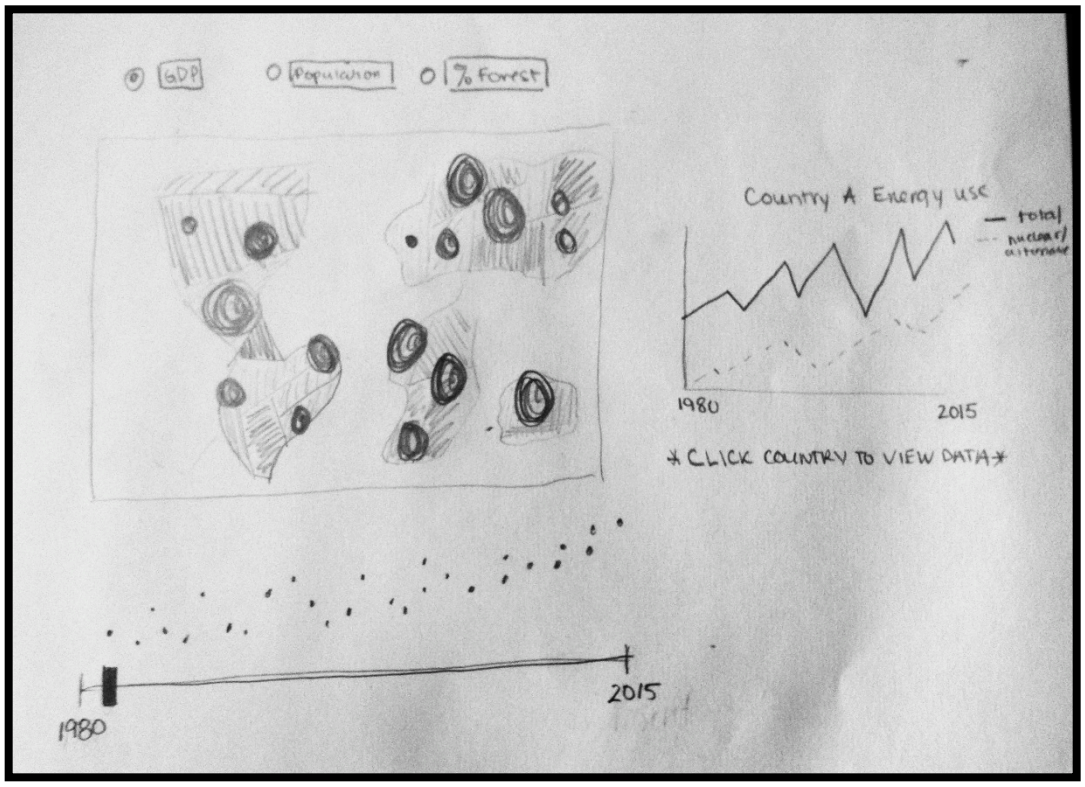
Using such a format would make it easy for us to visualize changes of multiple indicators together over time, i.e. for any given year and country, we could easily access all of its information. However, if we were visualizing a certain topic, such as forest area over all countries and all years, then it would be easier to keep the data in separate variables according to topic.

Visualization

As we have mentioned, our visualization will break down into 1-3 different chapters, each of which tells a different story about Carbon Emissions. At the core of the first chapter will be a map-based visualization which can represent discrepancies between countries in GDP, Population, and Forest Area. There will be an animated slider component that will show each country’s emissions per year through circles that will “pop up” on the map to emulate carbon emissions. The amount of carbon emissions will be represented by the size of the circle and the

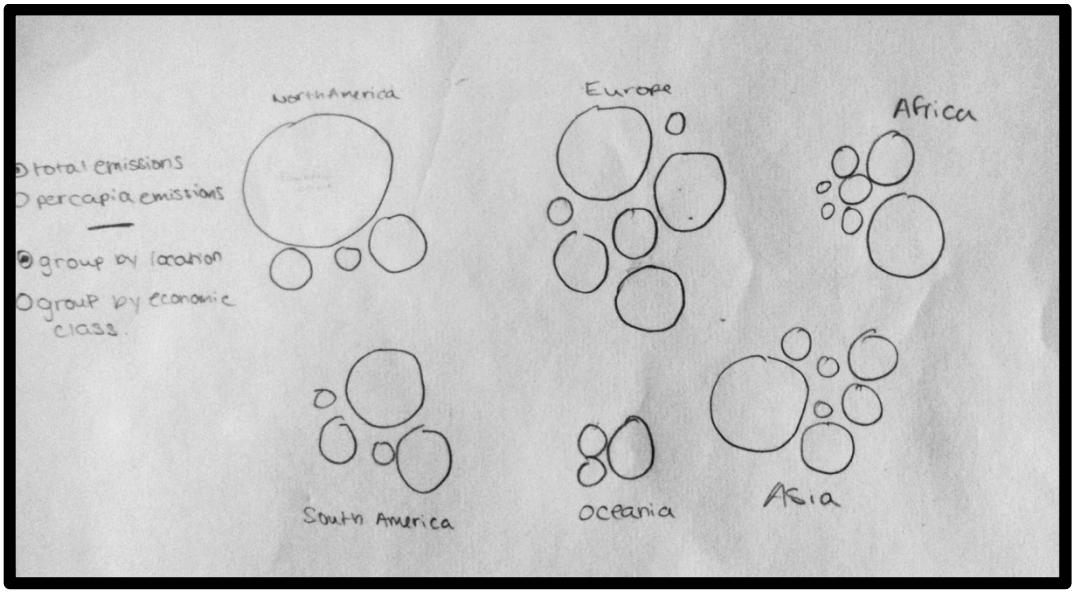
frequency of the circles appearing. Users will also have the ability to click on a country on the map to view country-specific energy usage stats on a line chart to the right of the map. These include data for total energy usage over time, with a percentage of that colored for total alternative energy usage.

Sketch of Our First Chapter



Our second chapter will represent countries as nodes in a force diagram rather than as a map. The size of the nodes will represent either 2014 emissions or 2014 emissions per-capita, depending on the user's selection. Nodes can either be grouped geographically or grouped by economic class (via rankings provided by the metadata in our World Bank CO_2 dataset). Users will be able to toggle between chapters via buttons at the bottom left hand of the page.

Sketch of Second Chapter



Our visualization plan for a third chapter is slightly more nebulous. We hope to tell some kind of US-specific story regarding Carbon Emissions. Ideally we would like to visualize the breakdown of US emissions by industry, ideally with a bar chart.

Must-Have Features

- Animated geographic visualization of global carbon emissions through time.
    - **Yearly carbon emissions per country** are represented by the size of circles which 'pop' up and fade out over time. These circles are overlaid on a map.
    - A slider will represent the **year** (from 1960-present).
    - Button above the map will allow users to toggle between **GDP, Population,** or **% Forest Area** for each country. This will be encoded by the gradient of the color of the country.
    - **Total yearly carbon emissions** will be represented in a scatterplot above the slider.
- Country specific energy usage data will be represented by an area chart to the right of the global visualization
    - Users can click on a country in the global visualization to see country-specific data isolated in a second view.
    - The chart will plot **energy usage** and **alternative/nuclear energy usage** from 1960 to present.

Optional Features

(These are listed from highest priority to lowest priority).

- Second Chapter: Force diagram representing 2014 global carbon emissions
    - Each node represents one country. Users will be able to toggle between two options for node size: representing **total emissions** or representing **per capita emissions.**
    - Users will be able to select between groupings of nodes by **geographic location** or by **economic class.**
    - Users will be able to mouse-over nodes to view more specific data regarding each country (rate of population growth, rate of GDP growth, % change in emissions from previous year).
- Third Chapter: US-specific emission visualization
    - Pie Chart will represent Carbon Emissions by industry in the US.
    - Slider will allow Users to interact with the pie chart to see changes in this data over time.
    - Line chart will show relative size of each industry in the US over time.

Project Schedule

April 3-10: Main map structure for the first chapter: incorporating total CO2 emissions as circles that "pop up" and GDP, population and forest area as color gradients on a map of the world. Anna to work on the circles and Rachael will to work the color gradient map. Update process book.

April 11 – 17: Animate the first chapter and include interactivity for clicking on certain countries on the map and viewing line graph visualizations of total energy consumption and percentage of alternative energy consumption on the side. Anna to work on animation and Rachael to work on interactivity. Update process book.

April 17: Submit Milestone. We aim to have a decent version of our first chapter done by the milestone.

April 18-25: Finish first chapter implementation and fix bugs (this might take longer than we expect).

April 25-May 1: Build the second chapter. It will be an interactive force diagram with mouse-over tooltip data. If time permits, build the third chapter, honing in on carbon emission for the United States using a pie chart or other circular data visualization. Both team members will work together. Update process book.

May 1-May 5: Clean up the storyline, boost up the graphic design and aesthetic elements, add seamless animations and make the video. Update process book.

May 5: Submit! Prepare to be blown away by our visualization**.**