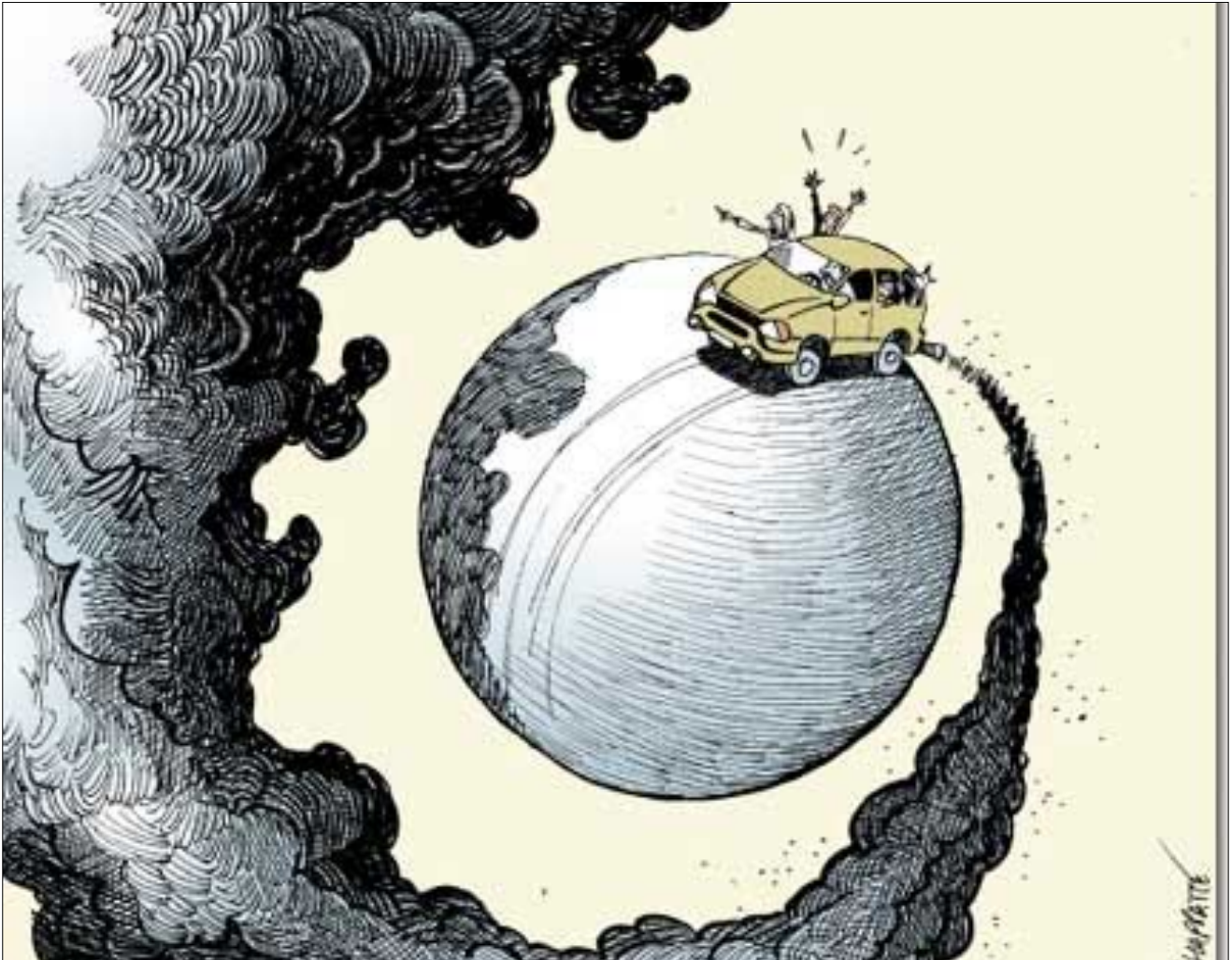# Process Book

Prepared for: Data Visualization
By: Brandon Tobin, Matt Howa, Matt Willden
November 2, 2016

CLIMATE CHANGE PLAYERS

Climate Change Players is more than a simple visualization, it combines a massive datable of socioeconomic data for nearly every major country in the world. It allows its viewers to see correlations between these socioeconomic properties and the countries producing them. What country produces the most greenhouse gasses?, Is there a relation between lang mass and emission output?, How does this data change year to year? These are only a few of endless questions which this graphic can help to answer.

OVERVIEW & MOTIVATION

The Question

The aim of this project, amongst other things, is to show correlations between countries and a wide arrange of known emissions data for those countries as well as other socioeconomic data for countries, such as GDP, land size, coast size etc.. The hope of these correlations is that specific questions about climate change can be answered. The intent is not to disprove or prove climate change, but to let the user decide how to interpret the data through who is contributing to world wide emissions. After finding large amounts of data, but not finding a visualization which gave the user a way to compare data between countries, we decided that we could take the data which is more than plentiful and answer the question ourselves.

Problems

We were able to find data, but there was a lot of data merging that had to be done. One problem we faced right away was trying to find a base year to go with. We knew from the beginning that we wanted a time element in our

An example word cloud that sparked our desire for a world cloud

visualization, so it could show how countries emissions changed year to year. The problem was that data wasn't available at the same time period and we wanted to present the data in most consistent way possible, finally we found a database that had data going back to 1990[1]. We then created a few other tables to help parse this data.

Another problem, though not as big as the first was with the word cloud. We wanted the word cloud data to be unique from the usual array of "array of words and their size". We thought[2] it would be interesting to correlate the word frequency based how much that word had been searched. Then when a country was selected it would change the cloud to show data corresponding to how much that country had searched for words. This presented many challenges, not only in data acquisition, but also in the fact the domain of countries that were represented by google trends was incredibly small. Further, we realized that

Working prototype of the design we have working.

---

[1] "World DataBank." http://databank.worldbank.org/data/home.aspx

[2] "Word Cloud." https://thumbs.dreamstime.com/z/climate-change-word-cloud-shape-concept-...jpg

between clicks the random nature of the word cloud would make it hard to draw any actual vague conclusion from anyway, so we decided to find the total searches for a large array of words display that instead. This put the word cloud somewhat behind our schedule.

The third problem we found was with how heat map. The heat map came together nicely, however finding a mapping of colors that allow a user to quickly and more importantly accurately, see the how each country's emission stacks up is challenging. We have tried mapping multiple colors and scales of one color and combinations of the 2, however, it still seems to come up short. We are looking into more techniques to understand how we can better solve this problem. Fortunately our tool tip for displaying detail of a country's data is coming along nicely and will help this problem.

PROJECT PROPOSAL
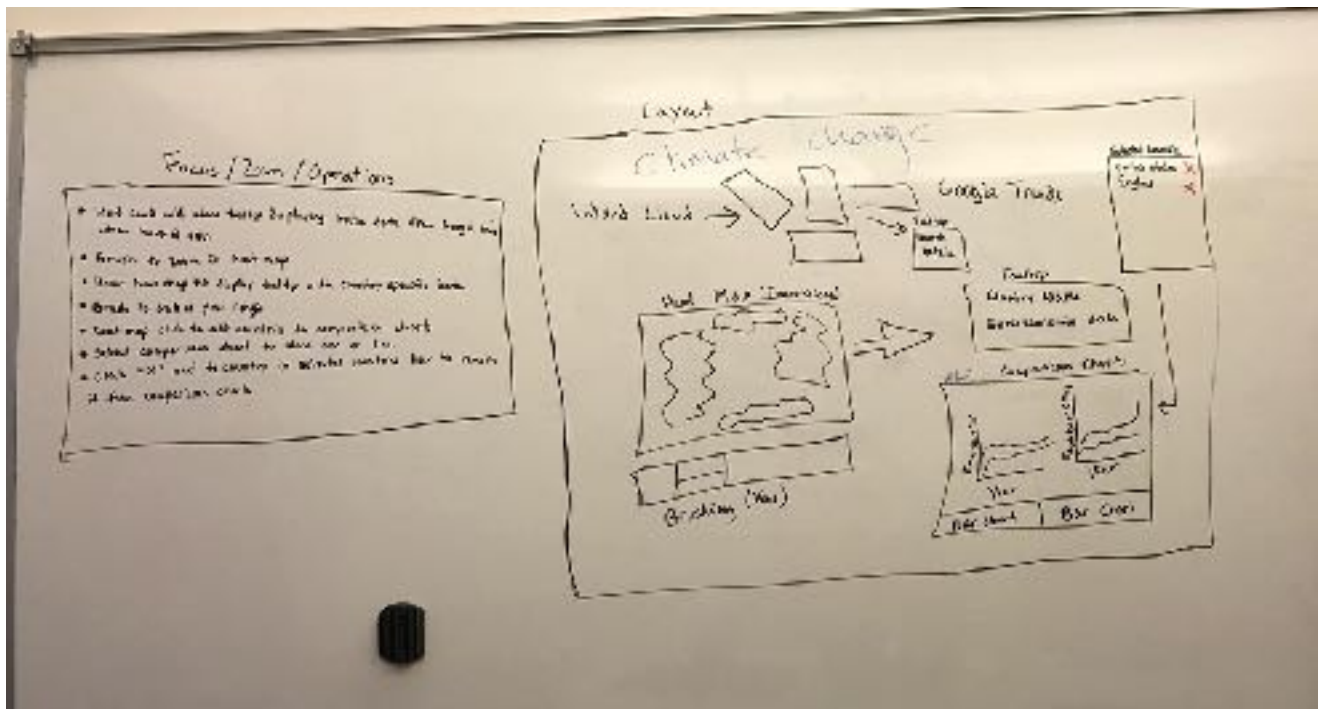
Overview

Our initial project consisted of 3 main moving parts: a word cloud formed from the search trends for countries. A heat map which displayed an overview of all countries specific emission output. Finally the last major part was a tool-tip which interacted with the heat map to provide detail of that country's emissions.

Of course there are many more little parts that will be incorporated as well, these include a looping animation button that will let the heat map cycle through the years so that the user can see how it changes over the 26 year data period. Brushing so that a user can see more fine grain details of a country's specific output by county/region. Finally making sure all of the graphs were interconnected so that a change in one would update everywhere else.

PROJECT REVISAL, MILESTONE
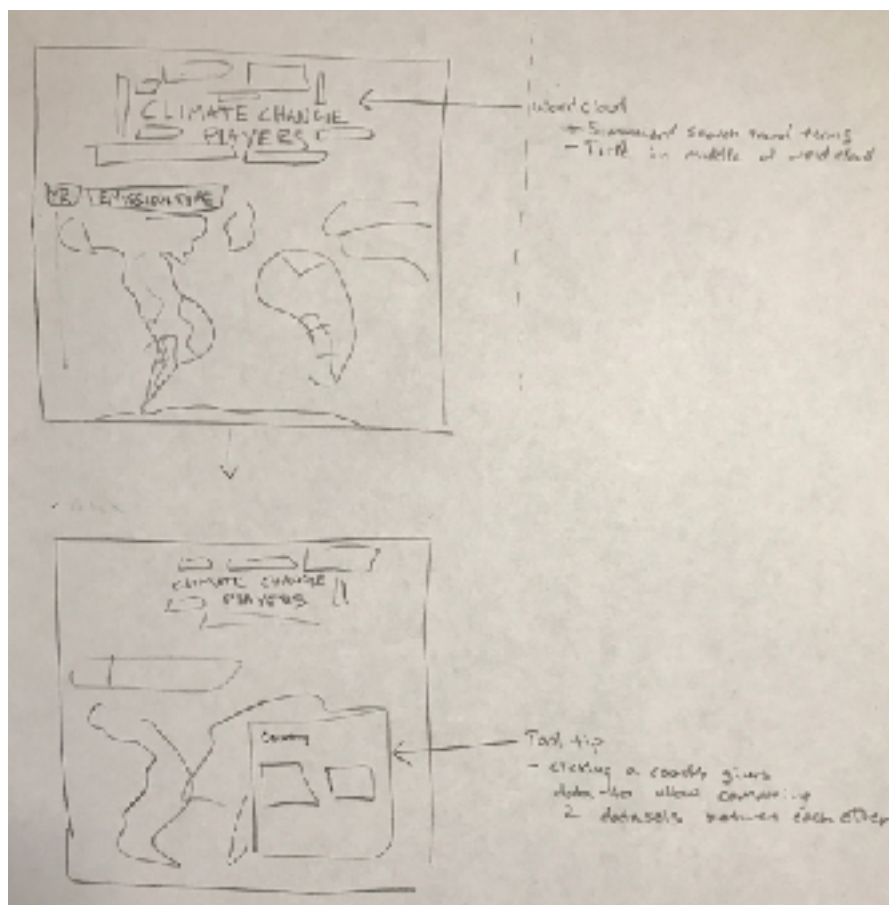
Data

The data for the heat map contains every major country, and for most countries, it also contains their underlying region as well. For each country entry there is a breakdown of several type of different emissions as well as socioeconomic data such as land size, GDP, population etc… For this project thus far, we have not been able to determine if brushing into the country regions is a viable option, simply because not every country has these data points. Moreover, we do not have the geographic data plots to actually paint these. This could prove an insurmountable task.

Tasks

The main tasks were:

- Finding the Actual Data: this was a large task and took some time, though as previously mentioned, a database which allows for excellent data gathering was found, and with a little fiddling it is possible to generate dataset very quickly. Though it was not necessarily trivial to write the data model. It required planning for the future in how to incorporate all of the moving parts, with not knowing the exact implementation. Ultimately this meant writing helper code to help with accessing the data.

- Creating a Heat Map: Once the data and data model was establish the heat map did come together. We wrote it in such a way to help incorporate the tooltip later. When a country is clicked it prints that country's info to the console. We hope this will make integration of the tool tip easier. The Heat Map did take several iterations to figure out a color scheme, something we are still working on:

As can be seen, we tried several different but still would like to investigate what color

color schemes, we found the last one to be the best, options could help represent the data better, and more


accurately.

- Word Cloud: This also proved to be more of a challenge than expected, the main problem was trying to figure out how to actually map the data to the word cloud. More specific, we realized that our interconnectivity of

clicking on a country and having that country's search trends for climate change topics was limited to only a few countries. This took time out of the development schedule, but we were able to still create what we think is a good visualization.

- Finally there are 3 tasks that remain: implementation of a comparison chart, tooltips, brushing, and animations. Our schedule is ahead on tooltips and comparison charts. The main idea here is upon clicking on a country the user will have the ability to interact further with the data. The illustration below would be an example of clicking on a country where a tool tip / comparison chart would be populated. Here we see that the 2 graphs show types of emissions compared against 3 socioeconomic properties.
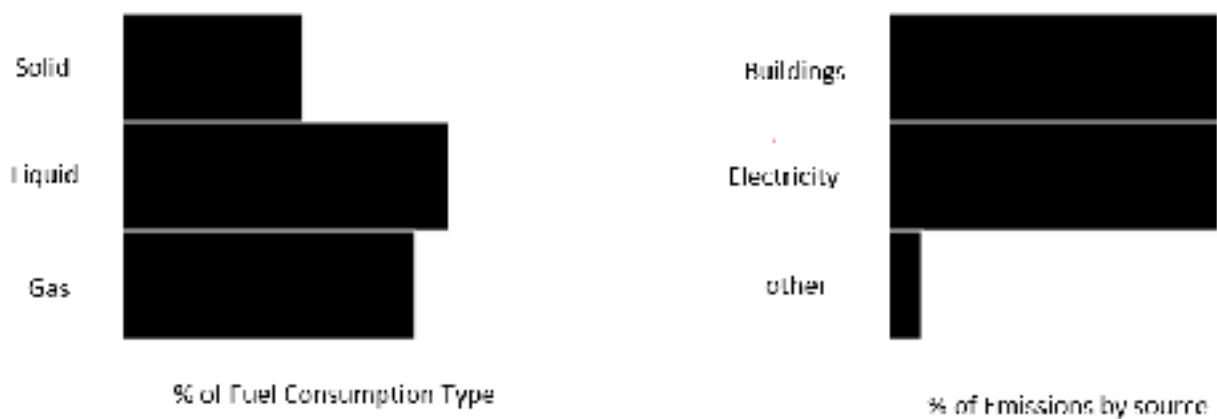
Users

In our vision, the target audience would be something of a New York Times / online newspaper reader. One where the author would want to show as little bias as possible to tell his narrative. The user would then be able to form their own opinions. It isn't meant to be a stance piece for or against, simple an informative visualization which can help a person who has an average understanding of climate change problems and general economic theories, answer their own questions.

Conclusion

We still have a ways to go, but we are making progress. We have used the time up to this milestone to gather data and weed out some problems that could stump our design in the future. We hope that it has been apparent through our previous project proposal and this process book that it is never the first implementation that sticks, but through each iterative implantation and sketch or design we learn a bit more about what works and what doesn't.