Individual Project – Week 2

CSC324

Will Chapin

02.02.22

Data Set Selection Process ~ 2hrs

I started searching for a dataset by vaguely searching on various platforms for Climate Change data. I found that many of these datasets were too vague, too large to handle easily (more than 15 MB), or had to little information to make interesting visualizations of.

One dataset I spent particularly long examining was on endangered species, and it included information on the number and name of the species as well as the amount of protected land/sea in each country by year and the number of endangered species in the country. I was thinking there may be an interesting coloration to see between expansion of protected land in a country and a decrease in endangered species. I ended up picking a dataset which interests me a bit more.

The dataset that I am hesitantly selecting is about energy generation and consumption in the United States. It breaks down information by state, by per capita levels, **by carbon intensity**, by types of consumption and sources of generation. The data is partially split across dozens of files on the same website, so I will need to spend a significant amount of time exploring the various files and determining the information I want to include. Helpfully, the files are all fairly well labeled, and the site has a few rudimentary visualizations I can use to help get a basic understanding of the information which I might include.

**Note**: The time estimate listed is my best guess. I did not record the time I spent searching, but 2 hours seems like a fair guess.