What is the problem you want to solve?

A lot of talk these days is centered around energy, specifically consumption of energy and the pervasiveness of fossil fuels. Given the electricity consumption data set, as well as the data I found on intrastate generation spread, I want to model the relationship between the various states, their energy consumption, and their move either toward or away from conventional fossil fuel resources. With this, hopefully we can predict the future of the country’s energy spread, and see if we are potentially heading in the right direction, and at the right rate.

Who is your client and why do they care about this problem? In other words, what will your client DO or DECIDE based on your analysis that they wouldn’t have otherwise?

My client would be anyone in the energy space. Generators have the need to know if they’re going to be able to stay in the market, consumers want to know how expensive being in the market will be as well as their options, and policy makers want to keep their constituents happy.

What data are you going to use for this? How will you acquire this data?

The EIA has an abundance of data, which is where I got the energy mix by state (last 30 years). Additionally, Kaggle has a data set mapping energy consumption by state for the past ~60 years (could extrapolate consumption and percentage based on the data from Kaggle to fill in what the mix might have looked like for the 30 missing years, or just run analysis on 30 years of data).

In brief, outline your approach to solving this problem (knowing that this might change later).

My plan is to combined the data, set up a dataframe that outlines the various kinds of generation (renewable, semi-renewable, fossil), by state and year, and figure out trends in generation. Once this is established, attempt to make a prediction on what the future of the country’s energy mix will look like, both with overall level of energy consumption, and the spread between the various forms of generation.

What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck.

My deliverables will be a slide deck and the code necessary to run this analysis. The project will involve combining/cleaning the data, visualizing the data into an easy to understand set of figures, and then running ML algorithms to determine what the mix will look like in the coming years.