Due to human energy consumption, atmospheric carbon dioxide concentration recently surpassed 400 ppm. The trajectory of atmospheric CO2 has been positive since the industrial revolution and 400 ppm amounts to over a 30% increase in atmospheric CO2 since the 1950s. Although mitigation efforts have been devoted primarily to reducing carbon and methane emissions, there is another important factor that has arguably not garnered sufficient attention. In addition to the release of carbon dioxide, humans are also busy decimating carbon sinks that could otherwise be actively sequestering atmospheric carbon; namely plants. It’s not uncommon to hear alarming statistics about rain forest destruction in Brazil, but deforestation is in fact a global phenomenon. The question I’ll attempt to answer is: What are the most important factors impacting global deforestation? The motivation for working on the project is to determine if there are factors which could be influenced by grass roots movements, governments, or companies to reduce the rate of deforestation.

The World Bank provides indicators on Agriculture & Rural Development which can be accessed here (“<https://data.worldbank.org/topic/agriculture-and-rural-development?view=chart>”). In addition to historic data on deforestation they also offer co-temporal historic data on potentially related factors such as (land area, poverty, and agricultural development). Before analyzing any of these, let’s take a look at global deforestation data.

#mean deforestation figure goes here

To put the data into perspective, the decline in forested land (~1.4 million square kilometers) is approximately twice the area of the state of Texas. It is important to note that countries with any missing data between 1990 and 2015 were excluded from the analysis, so the figure is almost certainly an underestimate.

Now let’s see if we can work out what factors are important for determining the rate of decline in forests. Since we’re looking at deforestation let’s use a random forest model to investigate. Sklearn RandomForestRegressor class can return a feature importances metric. For time series, I’ve simply calculated a slope for each country. The target variable in the model is the forestation slope calculated by (%forestation-2015 - %forestation-1990)/25.