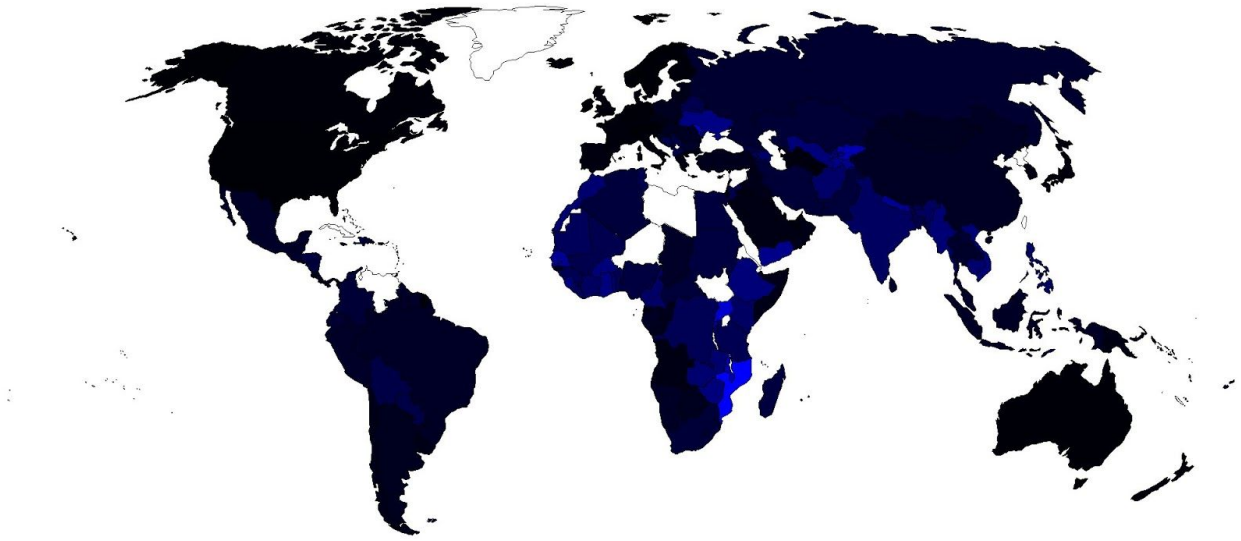
Interactive Data Visualization



Software Design Spring ‘18

Jordan Crawford-O’Banner

Juan Carlos del Rio

3.15.18

Project Overview

Easy access to information across the world through the internet has changed our lives in ways that not even science fiction could've predicted. The democratization of information and increased mobile connectivity have helped the world economy blossom. Therefore we decided to use interactive programing along with data visualization to produce a map of the world that showed the relationship between economic growth and internet access from 2001-2016. The idea behind this visualization would be to get a sense of whether more internet access really did mean greater with economic growth.

Results

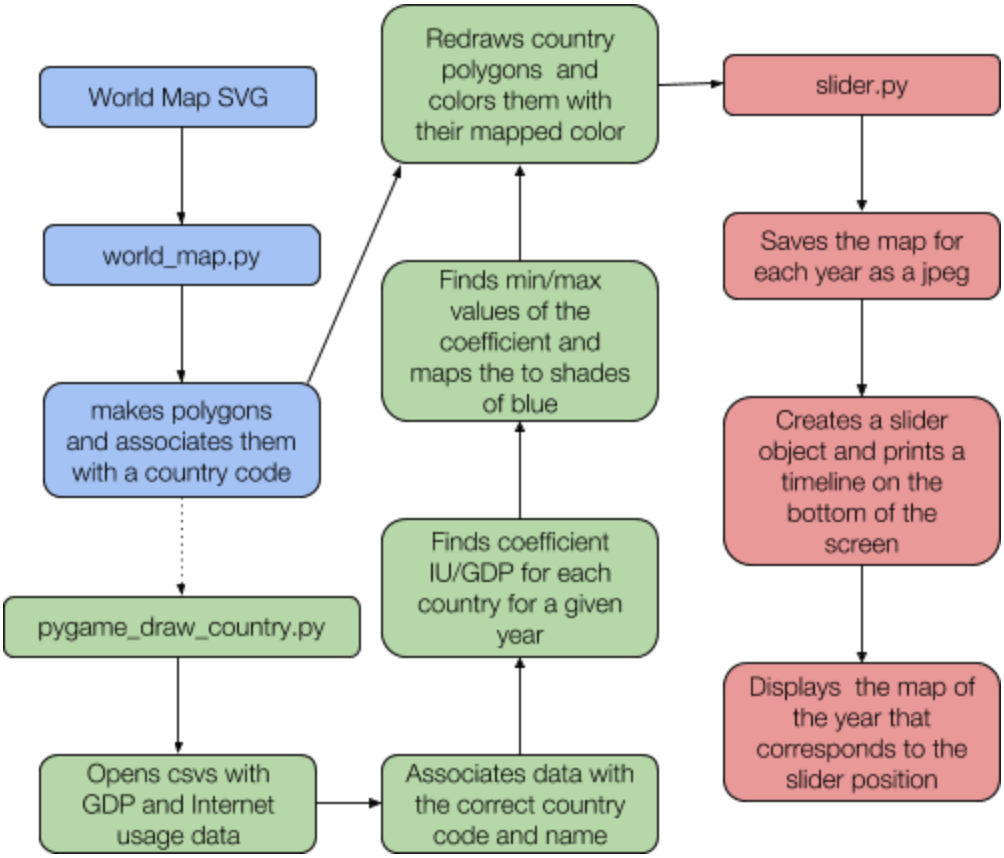
The amount of people with access to the internet within a country does not appear to have an effect on the gross domestic product of that country. Our maps showed that a majority of countries stayed dark shade of blue which suggests that the increase of internet users did little to affect the gross domestic product per capita of any of the countries. Most of the changes in color showed countries becoming a lighter shade of blue because the gross domestic product did not increase nearly as fast as the number of internet users.

Our results may be somewhat skewed because our metric was not as accurate or as clear a representation of the data as we thought it would be. The normalization of the number of internet users per GDP per capita to RGB values did not change the color of individual countries as much because small changes in RGB values are not very visible to the human eye. A metric which changed more drastically or a different coloring method that was affected more by smaller changes would have represented our data more clearly.

Implementation

Our interactive map is comprised of three main components: a script that reads a world map in an SVG, a script that reads our data and colors countries according to a color mapping for a given year, and a pygame script that allows the user to scroll through the maps for every year from 2001-2016. The svg reading and coloring modules were adapted from a 2016 Soft Des in class example and modified for our purposes. We decided to go this route to plot our colored

maps because of the poor documentation on more comprehensive libraries such as geoplotlib and gmap which we had identified in our project proposal.



Interactive data UML diagram

The diagram above shows the broad strokes of how the three scripts allow users to visualize our data. After adapting the code from the old class example for our purposed in world_map.py we turned our attention to the data. The data was all available in the World Bank database and we were able to download them as csvs that could be easily parse by our second script. Once we read the data we had to account for empty fields in the csv by filling them with zeros before calculating our mapping coefficients ($\frac{\%\ of\ the\ population\ with\ Internet\ access}{GDP\ per\ capita}$). The color mapping then takes the highest and lowest values of the coefficients and sets them to 0 and 255 accordingly to assign them to the blue RGB value. The map is then drawn for the year requested by slider.py and shows the countries with a low coefficient darker blue and the ones with a high coefficient a lighter shade.

We chose to save each individual map for each year as an image and then display the images because doing the calculations for each year caused the program to slow down too much for the slider bar to be functional. We attempted to go beyond our minimal viable product and include the ability to view the gross domestic product and internet access of the countries that the user hovered over, but because we ran into issues with the speed of our calculations, we were unable to include it.

Reflection

A positive aspect of our project was our scope. We scaled it in such a way that we had enough time to work on it outside of class without putting excess stress on ourselves while being able to accomplish our learning goals and gain a greater idea of the usefulness and flexibility of classes. The scoping of his project also allowed us to slowly increment the functionality of our final deliverable to ensure we would have at least a minimum viable product if we encountered severe difficulties. Another good representation of our teamwork was our ability to split the code into two distinct parts, which allowed us to work on the project if we were not able to meet each other.

In future projects, we could budget our time better so that we could progress past the minimal viable product and obtain a result that showcased even more data. We could also communicate more about pushing and pulling from github because we had multiple issues with our repo not letting one or either of us push to the repo.