Brief introduction to the problem/research question

The phenomenon I wanted to investigate is the manner in which people connect to the internet on a global scale. The main focus of my project is to see how this differs between regions and especially the relation between developing and developed countries. Therefore, my research question is: In what way does the manner people access the internet differ between different regions in the period 1990-2017?

Description of your data set(s)

The data for the project was downloaded from the world bank website and consist of three datasets containing data for total internet as a percentage of the population, broadband subscriptions per 100 people and mobile subscriptions per 100 people. The data is formatted in 4 columns country name, country code, year, value.

Another data set used is a shapefile dataset. This shapefile dataset has a larger number of columns such as country name, country code, latitude, longitude, population, area, and geometry. The geometry columns contain polygon and multi-polygon object. Polygons are a collection of x and y data points that make up the shape of the countries. Multi-polygons are collections of polygons for countries consisting of multiple land masses.

Description of the transformations applied to this data set

The transformation performed on the country data mainly consisted of flattening the data. Due to the large amounts of columns, the datasets put together would become difficult to work with. Therefore, I wanted to flatted out the data where each country occupied a single row and the columns represented the data for each year. This would reduce the length of the data from 10.000 rows to 207. This was done by loading the data in a dictionary, adding that data back to the fame, and then normalising the data with a pandas function.

The operation performed on the country dataset was mainly working with the polygon objects so that they could be inputted into a columndatasource. First, I extracted all the polygons from the multi-polygon objects. The isistance() function proved to be very effective in determining the type of polygon, the multi-polygons were then dissected with a for loop and the individual polygons were append to a data frame together with their country code. When all the multi-polygons were changed to polygons I could extract the x and y coordinates from them so bokeh could plot them. I used the .apply() function to create two new columns with respectively x and y values for every polygon. These coordinates can be accessed by calling the .exterior.coords.xy[] methods on the polygons. When I got all the coordinates I merged the two data frame containing the geographical data together with the data frame containing the country data so that every row contained a country name, data and the x and y points for that country.

A description of the visualization.

When looking at the map before the year 2000 the thing that stands out is that almost no one in the world is connected in one form or another. When moving the slider to 2001 another image appears. Here the world is clearly divided in developed and developed countries. Where developed countries are highly connected both by mobile and broadband connection and developing countries are not connected at all.

When moving the slider to 2010 a new image appears. Here you see that the divergence in broadband connections have largely remained the same. South America has started to become more connected to broadband and Russia and China have also joined up. Nevertheless, in 2010 you already start seeing the impact of the mobile revolution on the world as many countries start getting connected through mobile. The most remarkable one is Libya which has the largest connectivity in Africa. After the attack on Libya and the end of the Khadafy regime the number dropped substantially.

When fast forwarding to 2016, the difference in broadband subscriptions largely remains intact while the mobile revolution had great effect all around the world. Only in sub-Sahara countries there is some lagging behind, but for most of the world the difference has completely vanished. One of the reasons for this is the relatively cheap investments needed for mobile infrastructure versus broadband cables. The west had a head start in internet usage and invested heavily in placing broadband connection, while the rest of the world skipped this step and went mobile immediately. Another reason is the emergence of cheap phones from China flooding the African markets. While in the west people pay large amounts of money for American or South-Korean phones the phones in Africa are sometimes a tenfold cheaper. This has led to widespread usage of phones forever changing the future of the continent.