Gi Yoon Han(gh336)
Sean McNeil (srm274)
Samuel Dickerman (sld94)

Description of the data:
We used three main data sources for our data as well as three separate json files to help us map and group world data. The three main data sources we used are
- Happiness data- retrieved from http://www.happyplanetindex.org/data/
- GDP data - retrieved from https://www.google.com/fusiontables/DataSource?docid=1N0MaGTWtytF3zW-nj7OGnJFB3lf_n4Zcr54sib4#rows:id=1
- Drug use data - retrieved from: https://data.unodc.org/

We had to pull three more json files to help us map the data:
- World map data - retrieved from: https://github.com/mbostock/topojson/blob/master/examples/world-50m.json
- Country name+ID - retrieved from: http://bl.ocks.org/mbostock/4090846
- Country regions - retrieved from: https://github.com/lukes/ISO-3166-Countries-with-Regional-Codes

The happiness data includes several columns, but we chose to only use three of these columns: country name, adjusted HPI rank and inequality adjusted HPI.

The GDP data also has several columns, but we chose to only use two of these columns: country name and GDP per capita ($US). We also added a third variable to this dataset: rank. We did this in excel by sorting the columns by GDP per capita (ascending order) and adding a third row that has the rank with 1 being the country with the highest GDP value.

To convert these CSV files to JSON we used: https://shancarter.github.io/mr-data-converter/.

Finally, the drug use data was written to a JSON file by hand.

All of the datasets were used together based on matching country name.

We used the three json files to help us with our mapping. The world map data was the same world map data that we used in class. This was used to draw the map. Because the happiness, GDP, and drug use data did not have country IDs that matched the world map data country IDs, we had to use another JSON file that had the country IDs that matched the world map data and the corresponding country name. The country name could then be matched to our data files. There were a few country names that had to be edited by hand (for example: "The United States of America" vs. "The United States").

We also used a final JSON file to color code the countries by region called all.json. This all.json file includes several features about a country, but we chose to use the country name and the sub-region. We wanted to see the sub region (Eastern Europe vs. Western Europe) of the country to check if any region is particularly unhappy or happy. Using the region alone was not a small enough scale to see a trend. For example, some parts of Europe are particularly

happy while other parts are particularly unhappy. Comparing the sub-regions allows us to see these trends while using regions does not.

We also used happiness and GDP data for something we call relative happiness. If a country is very happy but also poor, it is happy relative to its wealth. If the country is rich but not happy, it is unhappy relative to its wealth. For this, we simply subtracted the GDP rank from the happiness rank (Happiness rank - GDP rank). Higher the value, happier the country is relative to its GDP; lower the value, unhappier relative to its GDP.

Description of the mapping from data to visual elements:

There are four key maps included in our visualization. There maps are an attempt to see the correlation between certain variables and happiness.

First graph (Happiness Map): A map of the world that shades in the colors based on the happiness rank of the country with happiest countries being shaded darker and the saddest countries being shaded lighter. Our data scales the colors from #B23Ad4 (purple) to white.

Second graph (GDP Map): A map of the world that shades in the colors based on the GDP per capita of each country. We rank each country, and the country with the highest GDP per capita gets the darkest shade, likewise the country with the lowest GDP per capita gets the lowest shade.Our data scales the colors from #36e24D (green) to white.

Third graph (Happiness vs. GDP): A scatter plot of happiness index versus GDP per capita is plotted on a log scale. Once each country in our data was plotted, we color coded distinct regions of interest. We grouped the 4 different regions into 4 different colors: South and Central america to green, Africa to orange, Eastern Europe to blue, Western Europe to purple, and the rest of the countries to grey. We wanted to highlight these 4 regions because we thought it was interesting that the South American countries seemed to be happier considering their economic level is much lower than Europe's. We also chose Africa because it shows that no money could also be correlated to low levels of happiness. These 4 regions were also geographically very distinguishable from each other, which allowed for a large spread of data.

Fourth graph (Drug Use vs Relative Happiness): A scatter plot of marijuana use versus (Happiness rank - GDP rank). We call the (Happiness rank - GDP rank) Optimism rank, and the the more optimistic a country is the yellower the circles are, conversely the less optimistic they are the more blue the circles are. We describe relative happiness to be the difference between happiness and money. For example a country with very little money and a high happiness rank would have a very high relative happiness rank. We wanted to look if there is a correlation between relative happiness and drug use, especially if some countries that are poor but happy used more drugs. We noticed that there seemed to be no difference between countries that use marijuana and countries that don't.

The Story

Our story is trying to tell the user about the factors that make a country happy. GDP and happiness are frequently discussed and happiness seems to be related to GDP. We first show that this assumption appears to be generally true with our data through our first graph (happiness map) and second graph (GDP map). Many of the same countries that are darker

purple (happier) are also darker green (have a higher GDP), and many of the countries that are white in the happiness map (unhappy) are also white in the GDP map (have a lower GDP). Next, we show this in our third graph by mapping GDP vs happiness on a scatter plot (happiness vs GDP). The dots generally follow an upward trend of the higher the GDP, the higher the happiness level.

This is also where things start to get more interesting. We ask the general question, what makes a country particularly happy despite an average GDP, and what makes a country particularly unhappy with a lower GDP? Is there a trend amongst these outliers? We first look into these questions by adding colors to the third graph (happiness vs GDP) to represent different regions. We color coded the most interesting regions. After this visualization, we immediately noticed a trend in regions - Central and South American countries seem to be particularly happy for their GDP and Easter European countries seem to particularly unhappy for their GDP. In fact, it seems that while Central and South American countries have similar GDPs to Eastern European countries, they are much happier.

After looking at the regional trends, we tried to look for more possible correlations. We decided to look into drug use. Maybe the countries that are happy for their GDP ("Poor but happy countries") simply did more drugs and that is why they are happier than their peers. We represented a country's happiness relative to GDP by subtracting its GDP ranking from the happiness ranking. More positive the value is, happier the country is relative to its GDP. Then we plotted cannabis usage by country with respect to these values. As it turns out, the happier countries demonstrated slightly more usage overall, but there was no clear correlation unlike what we found in the previous graph. So these "poor but happy" countries were not simply optimistic and happy due to more drug use.

References:
Linear gradient tutorial from: http://bl.ocks.org/mbostock/1086421
Map Legend tutorial from: https://dev.socrata.com/consumers/examples/simple-chart-with-d3.html