Lab 5:Global Health Exploratory Data Analysis

Daniel Ocampo and Jesse Del Greco

October 26, 2017

1 Objectives

To learn how to ask exploratory analytics questions which may be answered using an investigation of global health data from the World Health Organization (WHO). To determine how global health data can be manipulated to address specific types of questions. To create meaningful modifications to the data for the production of plots to visually answer the questions.

2 Questions For part 1

1. Does spending more on health equate to longer expected lifetimes?
2. What country has the best ratio of health expenditure to expected lifetime?
3. Are there differences in the trends between males and females?
4. What is the correlation between health expenditure and suicide rate?
5. Have these trends changed significantly over the last ten years?

3 Part 2

Please note that Jesse Del Greco has the code because my got erased somehow

Your Health Data Analysis: In addition to the R code that provides your solutions, you are to write a report that includes the following:

- The link and name of your WHO data set used for your exploratory questions.
  Health expenditure per capita, by country
  Life expectancy Data by country
  Suicide rates, age-standardized Data by country

- Description of your data set (describe what each column contributes to the data set). For this a simple sentence would be sufficient to introduce each column. **Health expenditure per capita, by country**: Per capita total expenditure on health at average exchange rate (US$) This data set contributes the amount of money a person spends Per capita total expenditure on health at average exchange rate (US$). This data set allows people to see how much they spent health expenses. **Age-standardized mortality rate (per 100 000 population** This column gives of the years of all the suicide rate per 100,000 people.

  Life expectancy at birth (years): Life expectancy at birth reflects the overall mortality level of a population. It summarizes the mortality pattern that prevails across all age groups - children and adolescents, adults and the elderly.

  Life expectancy at age 60 (years): Life expectancy at age 60 reflects the overall mortality level of a population over 60 years. It summarizes the mortality pattern that prevails across all age groups above 60 years.

  * Does spending more on health equate to longer expected lifetimes? * This is our main question that we're interested in, and the others stem from this one. We have heard that the United States spends more on healthcare than many other countries, yet has one of the least healthy populations, and we have heard that Japan spends the least and has one of the longest life expectancies in the world, and we're curious if the data supports those statements and what differences may lie in comparisons between countries.

  * What country has the best ratio of health expenditure to lifetime expectancy? * As economics majors and minors, we're also interested in looking into which countries are getting the best "bang for the buck" when it comes to lifetime expectancies versus healthcare expenditures. Once those countries are identified, numerous other research questions open up in the realm of what can the less efficient countries learn from the more efficient ones.

  * What is the correlation between health expenditure and suicide rates? * An interesting side question we came up with, is the question of higher healthcare expenditures being potentially correlated with higher or lower average suicide rates. This is a question that we don't have an expected outcome for, and the answer may very well be that they aren't related in any consistent way. If there is a positive correlation, then perhaps this is evidence that more of that funding should be spent towards mental health.

  * Are there differences in these trends between males and females? * Besides anatomy differences, males and females have historically tended to have very different occupations and workplaces, which might tilt the trends in one direction or another; looking at them separately might show some correlations where looking at them averaged out hid those correlations from recognition.

* Have these trends changed significantly over the last decade? * It is worth considering not only what the trends are with the most current data available, but also how they have changed, to see whether the trends stay the same, reverse, or shift. Various factors such as advancements in medical treatments and changes in workplaces might have some bearing on the trends we're wanting to look at, and these factors take time to be reflected in the datasets.

4 The Talk

The talk was about data collection and how we should be careful with biases such as women lying about there weight, compared to male who are usually more open to them. How survey try to figure out if a person is lying by asking smart and effective question. Asking question that imply something but may not mean explicitly say something. Well also talked about data clustering, and how that is useful, and diff rent data sets. explanations why some people were getting more sick than others. Which was quiet interesting, cause one of her case studies was about how women in a coal industry were getting sick more than the men, well turned out that the women were washing the guys clothing and therefore just get sick. That was one of explanation for the talk. It might have been a little diffrent but the idea is there.