Data Science Group - The Fantastic 4

Proposal due 4/7/19

The question that our group has decided to tackle is, “What factors are correlated to life expectancy of females as a percentage of males?” in hope to discover ways the lower life expectancies could be increased. We will do this by graphing to see the other country outliers (both hi and low) and looking at those. We have decided that it would be helpful to examine outliers of “Survival rate to the last grade of primary schools: females as a % of males” as well. We will then determine why the higher ones may exist and why the lower ones may as well, based on the other variables. Its relevance to a societal need is that, if it turns out that the proportionately lower or higher life expectancy for women has a strong correlation with another variable, that could lead us to determining how the life expectancy could be increased by either increasing or decreasing other factors. For example, if there is a higher mortality rate for women who go through childbirth, that would most likely lead to a lower life expectancy of women as a percentage of men for the region. Perhaps that region needs to address its healthcare system. This is just an example, but these are the sorts of correlations that we are interested in exploring, leading to being able to predict female life expectancies based on those correlations. If we had the means, we could then use our information on what causes the lower life expectancies to help those in at-risk regions work their way up to having longer life expectancies.

The UNICEF data we are using has data in regards to the state of people around the world in the following topics (more detail can be found about these topics on the data set pdf under the section “Notes of specific tables” on the page marked 151 (6th page in the pdf)):

Mortality Rates

Nutrition

Health

HIV/AIDS

Women

Child Protection

Early Childhood Development

Economic Indicators

We plan on comparing all of the UNICEF data in order to find correlations between different aspects of life and life expectancy between men and women.

We also have found a paper that is somewhat closely related to this topic and are interested in using its datasets in our research. These datasets show the difference in life expectancy between men and women, as well as the rates per cause of death of men and women. The link for that paper can be found here: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6015620/>

UNICEF dataset available on our github: <https://github.com/bigbbv/COSC_481-Major_Project/blob/master/SOWC-2017-statistical-tables.pdf>

It can also be downloaded from the UNICEF website:

<https://data.unicef.org/resources/state-worlds-children-2017-statistical-tables/>