Project Proposal: ISyE 6414 (Dr. Yajun Mei) - Spring 2020

Samuel Garcia, Michael Szostak, Osman Ghandour, and Peter Williams

2020-03-11

1 Project Description

According to the *World Health Organization*, ("Suicide: One Person Dies Every 40 Seconds" 2019) one person dies every 40 seconds from suicide. It is the second leading cause of death among teenagers and adults aged 15-29 years. Despite the staggering number of suicides happening worldwide, just 38 governments worldwide have a national suicide prevention strategy.

Each of these deaths are tragic, and sadly also preventable. For every suicide, there are many more attempts, and previous suicide attempts are the single most important predictor or risk factor for future suicide attempts. ("Suicide: Key Facts" 2019) The possibility of prevention and the scale of the problem highlight the need for policy makers, at the national level, to understand the factors that contribute to suicide not only in their own nations but also globally.

The objective of our project is to analyze country-level and gender specific data related to suicide rates over the past ten years, and augment it with other country level data and metadata to gain an understanding and better describe rising suicide rates worldwide. Our intentions are two-fold:

1. To describe and identify measures and indicators that impact suicide rates at the country level for both males and females, in order to provide high-level decision making support for leaders and authors of public policy related to mental health and suicide. This includes identifying, and monitoring over time, meaningful factors and measures related to suicide prevention, for those who manage health related planning activities and priorities.

2. To discuss the sources, and process of collection, of data related to suicide at the country level; describe ancillary datasets that were employed to augment and enrich insights related to international differences in suicide rates. The goal is to make further research and data analysis on this topic more accessible to other interested researchers and data analysts in the future.

2 Dataset Description: How and From Where

The core dataset we plan to rely on comes directly from the *World Health Organization* and is referenced below. The key measure of interest for our study is the *the age-standardized suicide rate* by country, which is defined as *a weighted average of the age-specific mortality rates per 100,000 persons, where the weights are the proportions of persons in the corresponding age groups of the WHO standard population.* ("Suicide Rate Estimates, Age-Standardized Estimates by Country" 2019) The data consists of country-level measures of suicides rates, as defined above, for 183 nations, broken out by gender, and standardized as describe

above. These estimates of age-standardized suicide rates were taken in the year 2000, 2010, 2015, and 2016. This enables the data analyst to observe differences in rate estimates over time, allowing more context and benchmarking for the purposes of this analysis. In addition to relying on the core suicide rate statistics provided above, we also intend to append country-level data, for corresponding time periods, from ancillary data sources. Currently, we are considering the following data:

1. GDP per capita, as a high level measure of wealth, sourced from the *World Bank* (https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?view=map)
2. Adult education level, as made available from the *OECD*, the *Organization for Economic Co-operation and Development* (https://data.oecd.org/eduatt/adult-education-level.htm)
3. The female labor force participation rate, from the *United Nations Development Programme* (http://hdr.undp.org/en/content/labour-force-participation-rate-female-male-ratio)
4. Whether or not the government of a country has a suicide prevention strategy, according to the *World Health Organization* (https://apps.who.int/iris/rest/bitstreams/1174021/retrieve)

3 Scientific Research Questions To Address

Integrating data sourced from various programs, and data collection agencies detailed above, we plan to tackle our objectives by focusing research questions in the following areas:

1. Can we statistically confirm and identify which measures/indicators are best associated with male and female suicide rates at the country level?

2. Based on the data we obtain, are there any inferences we can make about how a country may be able to reduce the suicide rate (male, female, or both)?

3. Are there recommendations we can make to future researchers or policy makers beginning to study this topic?

As researchers we understand that our project is limited by our ability to collect complete and accurate data at the country level, and our research design and methodology. We will attempt to strike a balance between addressing all research questions listed above while also communicating potential limitations to our study.

4 Proposed Statistical Methods and Models

We plan to rely on multiple linear regression as a mechanism to discover what variables contribute to describing the outcome of interest, along with their statistical significance. Our purpose in using this tool will be to describe the dataset intuitively, and identify statistical significance of particular explanatory variables on our outcome of interest.

We also plan to explore using another generalized linear model, the poisson regression model, as our key response variable consists of count data (the number of suicides per 100,000 persons). Our plan is to use both of these models as a means to describe and test various statistical hypotheses related to variables that may contribute to suicide rates country by country across the globe.

References

"Suicide Rate Estimates, Age-Standardized Estimates by Country." 2019. http://apps.who.int/gho/data/node.main.MHSUICIDEASDR?lang=en.

"Suicide: Key Facts." 2019. https://www.who.int/news-room/fact-sheets/detail/suicide.

"Suicide: One Person Dies Every 40 Seconds." 2019. https://www.who.int/news-room/detail/09-09-2019-suicide-one-person-dies-every-40-seconds.