Global Suicide Statistics: An Analysis of Critical Variables

By: Jack McNeilly, Shannon Houser, Brian Linder

Duke UNIVERSITY

Introduction

Why Suicide? After the recent loss of two Duke students to suicide, we wanted to analyze global suicide rates and the prevalent factors that might contribute to people taking their own lives. The dataset we analyzed was built to find signals correlated to increased suicide rates among different cohorts globally, across the socio-economic spectrum. We sought to find which factors caused higher rates of these tragedies in order to bring light to this topic, and increase awareness about the issue across the world.

Variables: country, year, sex, age group, count of suicides, population, suicide rate, country-year composite key, HDI for year, GDP for year, GDP per capita, and generation.

Data Source: World Health Organization, the World Bank, the United Nations Development Program (1985-2015)

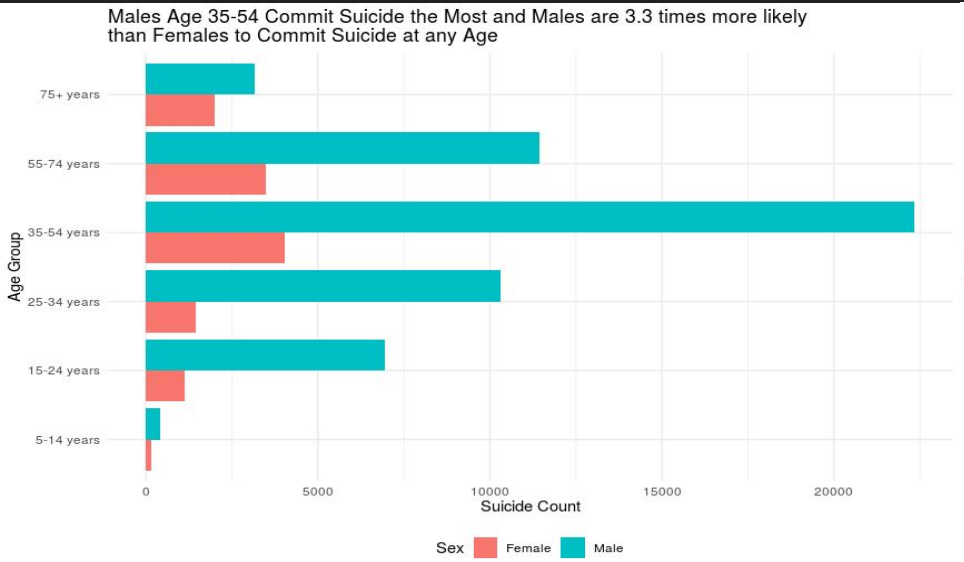
Objectives

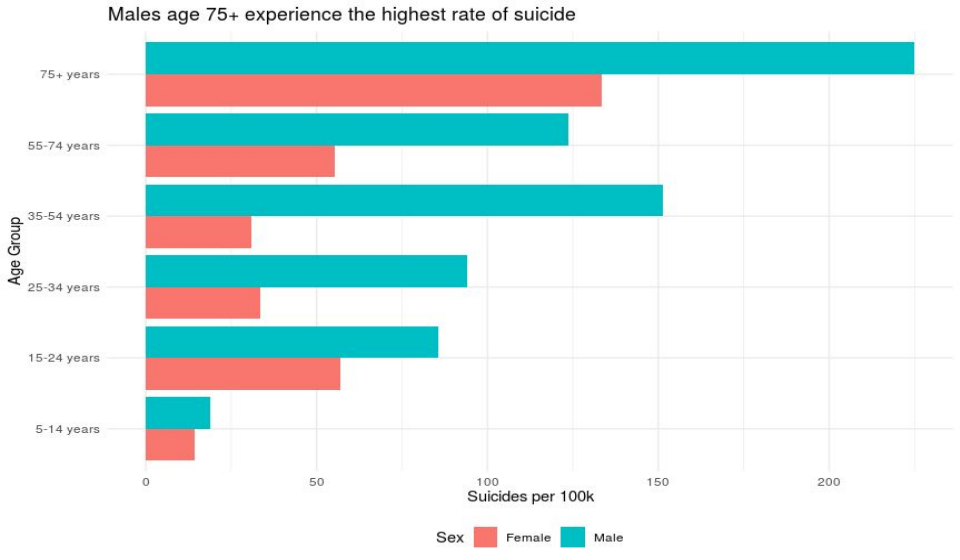
- What are the most prevalent factors correlated to those that take their own lives?
- Can we predict suicide rates based on certain social, political, economic, geographical, or personal factors?
- Which age bracket, sex and countries have the highest rates of suicides over the duration of the study

Countries with Highest Average Suicide Rate are Mostly ex-Soviet Union

Exploring Variables Correlated to Higher Suicide Rates

Variables: Sex and Age

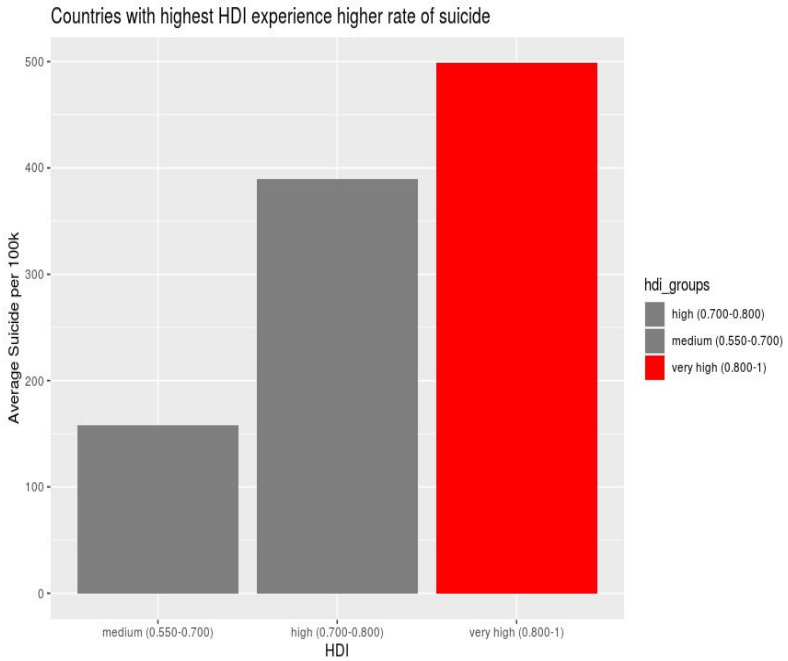




- Males committed the largest number of suicides in the dataset, being 3.3 times more likely to commit suicide than females.
- Middle aged men (35-54) represent the group with the highest total suicides, while men in the 75+ age range have a higher suicide rate.
- These visualizations highlight that from the population tested in this dataset, men seem more susceptible to suicide, with middle-aged and elderly males being the most affected.
- The difference between these visualizations highlights that the greatest loss by suicide is from middle-aged men, while the group with the largest proportion of suicides in its population are men aged 75+.

Linear Model to Predict Suicide Rate

Variables: HDI, GDP per capita, sex, age, and year



Conclusions

- Surprisingly, countries with higher Human Development Index (HDI) scores also displayed the highest suicides rates. This suggests that perhaps more developed countries have higher suicide rates than their less developed counterparts.

$$\widehat{log(suicides/100k)} = sex + agegroup + year + hdi + log(gdp/capita)$$

- A linear model we created to predict the log of the suicide rate based on the variables of sex, age, year, hdi, and log(gdp per capita), gave us an R^2 of .64. This means that our model accounts for approximately 64% of the variation in the log suicide rates.