Global Suicide Indicators – Springboard Data Science Track Proposal

Contents

What is the problem you want to solve?

The whole world has been under pressure since the pandemic started; this has prompted many to pay attention to people's well-being. However, mental health issues have been an ongoing public health issue before COVID. It's been estimated that in the United States alone, between 1999 and 2018, there has been an increase of 35% in suicide rates. From 10.5 per 100,000 population in 1999 to 14.2% in 2018. It has also been estimated that in 2018, the suicide rate in men was 3.7 times higher compared to women. Men committed suicide at a rate of 22.8 per 100,000 whereas women committed suicide at a rate of 6.2 per 100,000.[1]

Additionally, the World Health Organization (WHO) estimates that 800,000 people die to suicide per year. This equates to a person dying every 40 seconds. Per WHO, "79% of suicides occurred in low- and middle-income countries in 2016. Suicide accounted for 1.4% of all deaths worldwide."[2] WHO marked suicide as the 18th leading cause of death in 2016 worldwide.

In this project, I would like to find factors that can help lower suicide rates. There are several disparate data from the Centers for Disease Control (CDC) and WHO regarding current population mental health status, mental health policies, mental health infrastructures, and different economic factors across the globe that are potentially driving the suicide epidemic. The key indicators found in this study can help shape and build on policies to reinforce programs to help people struggling with their mental health and are contemplating suicide.

---

[1] Increase in Suicide Mortality in t he United States, 1998 – 2018.
https://www.cdc.gov/nchs/data/databriefs/db362-h.pdf

[2] Mental Health and Substance Abuse - Suicide data. https://www.who.int/teams/mental-health-and-substance-use/suicide-data

Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis?

Mental health policymakers, the government, and nonprofits focusing on community mental health betterment would be a good target market for this project. The information obtained from the project would be able to advise them on what would be a good resource to focus on.

Policymakers can use it to build new policies or build on existing ones to establish stronger campaign against suicide. With policymakers/government intervention, new and expanded resources can become available to the public. Narrowing the focus of the policies through the result of the project model can highly inform and direct funding to the important key factors that can help prevent suicide.

Nonprofit companies can also benefit in this information as they can additionally focus their efforts and funding to recognized key factors. They can reassess their current pursuits and see if the new factors detected in this study, if any, can help them with their missions.

What data are you using? How will you acquire the data?

The data that will be used in this study will be from the global suicide indicator dataset available on Kaggle. However, this dataset contains a lot of missing data when it comes to the different mental health infrastructures available across the world. Therefore, as a supplement, WHO's database regarding mental health governance across the globe will also be utilized.[3]

Briefly outline how you'll solve this problem.

I am thinking of utilizing a multivariate linear regression to see which factors are driving the suicide rates across the globe. Similarly, I would like to use an ensemble learning model, Random Forest, in particular. I am going to use this model to conduct a feature importance test. In addition, the random forest model has more capabilities of hyperparameter tuning.

What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck.

For this project, I will be a writing a report going in a more in-depth explanation of the processes and modeling. In addition, I will be creating a slide deck presenting high level explanation along with the results and recommendations from the study. If time allows, I want to create an interactive dashboard in either PowerBI or Tableau to provide a more interactive visualization of the project.

---

[3] https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/mental-health