Exploring Factors Contributing to the Suicide Rate

Principal Investigators: Tina Cao, Yumeng Zhang (tcao32@wisc.edu)

Introduction

Suicide is a very serious and complex issue. Around 1.4% of global deaths in 2017 were due to suicide. According to the World Health Organization, more than 700,000 people commit suicide each year, with many more attempting suicides. Suicide is not just a significant public health issue but also a tragedy for individuals, communities, and nations, with long-term consequences for those leftovers.

For this project, we are going to investigate the suicide mortality rate. The suicide mortality rate is calculated by dividing the number of suicide deaths in a year by the population and multiplying by 100,000. We want to see how the suicide mortality rate changes over time and how the suicide mortality rate is composed of people of different sex and age groups. Moreover, we want to investigate how different factors (GDP per capita, unemployment rate, life satisfaction, and the number of depressions) affect the suicide mortality rate by performing linear and multiple regressions. We think that there should be a significant correlation between these factors and the suicide mortality rate.

The main result is that there is a downward trend in the world's average suicide mortality rate over time. Moreover, GDP per capita alone does not have a significant effect on the suicide mortality rate, but GDP per capita and unemployment together have a significant effect on the suicide mortality rate.

Data

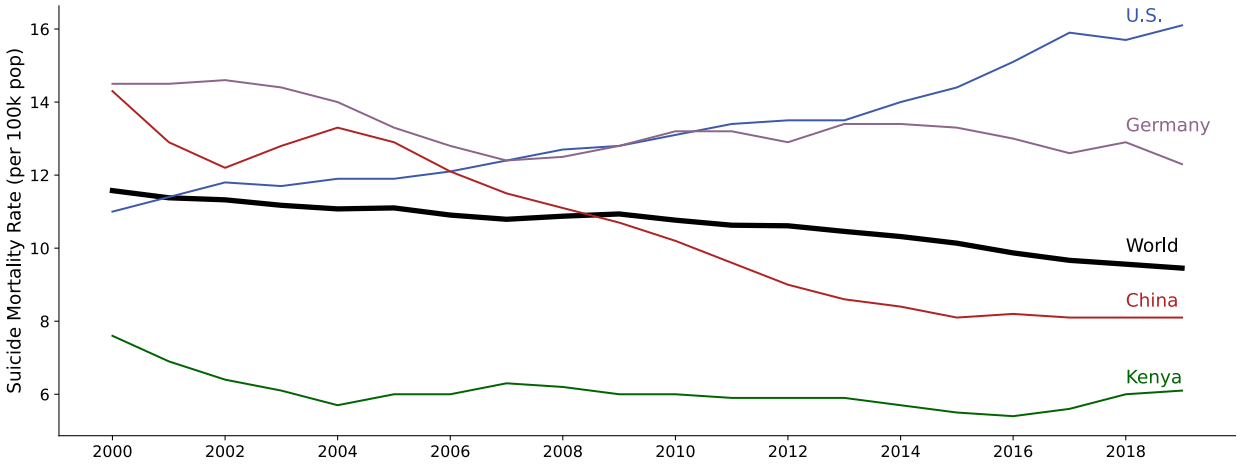
All data comes from the website Our World in Data. We used eight datasets in total. For all datasets, each row represents data from a particular country and a particular year. The eight datasets include three datasets about the suicide mortality rate, one dataset about the population, one dataset about GDP per capita, one dataset about the unemployment rate, one dataset about life satisfaction, and one dataset about the number of depressions. The main dataset is the one about the suicide mortality rate from 2000 to 2019. In the end when we are to examine the relationship between different variables and the suicide mortality rate, we only used 2019 data because 2019 is the latest year in our data. We merged the 2019 suicide mortality rate dataset with other datasets by country, country code, and year to create a combined dataset.

Methods and Results

Time-series Analysis

To investigate how the suicide mortality rate changed from 2000 to 2019, we plotted the world average suicide mortality rate and the average suicide mortality rates for representative countries from 2000 to 2019 in a time-series graph shown in Figure 1. We decided to show the average suicide mortality rates of the U.S., Germany, China, and Kenya to see how their average suicide rates compared to the world's average suicide rate. Overall, there is a downward trend in the global suicide mortality rate from 2000 to 2019, meaning there is a decrease in the number of people who die from suicide. For the representative countries, we see that there is an upward trend in the U.S. suicide mortality rate. One of the possible reasons for this upward trend might be that it is easier to find a gun in the U.S. There is a decreasing trend in the suicide mortality rates for both China and Germany and there is a very steep decrease for China. Kenya remains a flat trend and it has the lowest average suicide mortality rate among all the countries we discussed.

Figure 1: Suicide Mortality Rate from 2000-2019

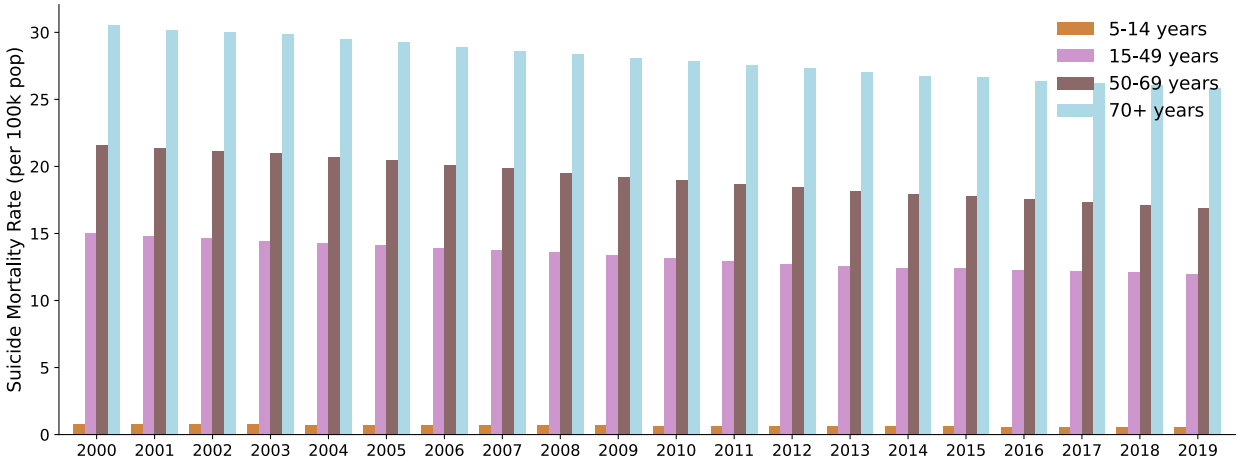


To investigate how the worldwide suicide mortality rate is composed of people of different sex and age groups, we plotted two bar graphs. Figure 2 shows the average suicide mortality rate by sex and Figure 3 shows the average suicide mortality rate by age group. According to Figure 2, we can clearly see that the suicide mortality rate among males is much higher than the suicide mortality rate among females. One of the possible reasons for this is that men are subjected to more social pressure than women such as higher work and life pressures.

Figure 2: Average Suicide Mortality Rate by Sex from 2000-2017

According to Figure 3, we can see that the suicide mortality rate is the highest among all age groups for those over 70 years old. In the accelerated aging environment, the social security system and socio-economic development can hardly meet the material and spiritual needs of the elderly, and the multiple contradictions have created a situation where the elderly are under more serious survival and spiritual pressure than the young.

Figure 3: Average Suicide Mortality Rate by Age Groups from 2000-2019



Suicide Mortality Rate around the World in 2019

From here, our analysis shifts from analyzing the suicide mortality rate over a period to focusing on the suicide mortality rate around the world in the year 2019 only. The reason why we want to focus on the year 2019 is that it is the latest data we have in our dataset. Figure 4 shows a world map with the color indicating the suicide mortality rates. From Figure 4, we see that Lesotho, an enclaved country surrounded by South Africa seems to have the highest suicide mortality rate in 2019 because the color of the country is yellow. To reassure this finding, we created a table

(Table 1) to show the top 5 countries with the highest suicide mortality rate in 2019. The number one country with the highest suicide mortality rate in Lesotho with a 72.4 suicide mortality rate in 2019. According to the world map and the table, we find that the 2019 suicide mortality rates in South American and Asian countries are on the high side of the world. The high suicide rates in South American countries may be because South American countries like Guyana are poor and have high alcohol consumption and family violence (Rawlins & Bishop, 2018). We also think that low GDP per capita, high unemployment rate, low life satisfaction, a high number of people with depression, and so many other variables could also be possible explanations for a high suicide mortality rate.

Figure 4: Suicide Mortality Rate around the World in 2019

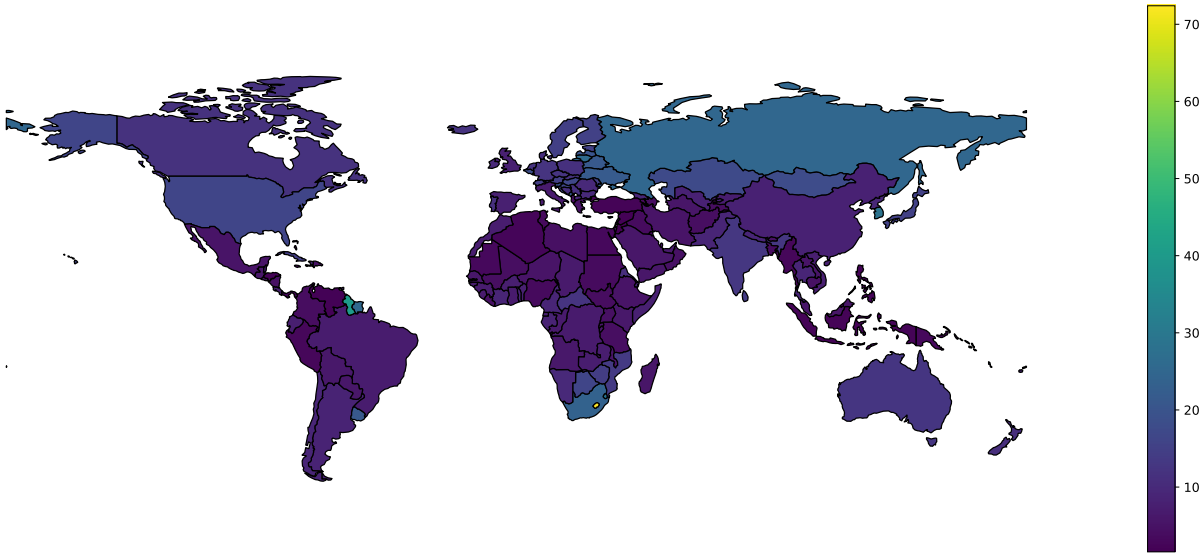


Table 1: Top 5 Countries with the Highest Suicide Mortality Rate

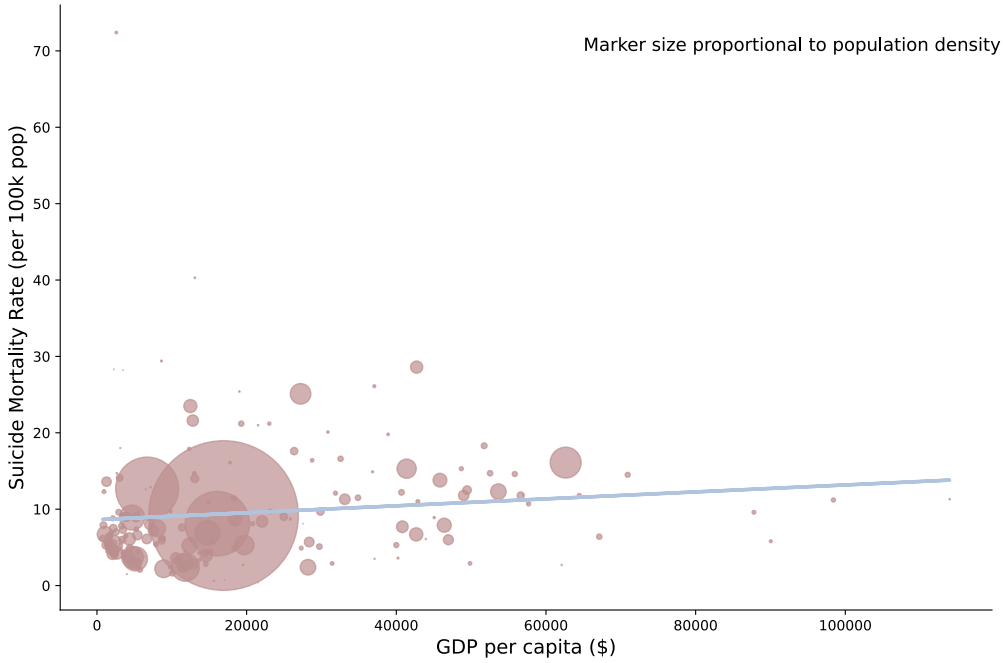
	Entity	Code	Year	Suicide mortality rate (per 100,000 population)
2399	Lesotho	LSO	2019-01-01	72.4
1699	Guyana	GUY	2019-01-01	40.3
1279	Eswatini	SWZ	2019-01-01	29.4
3899	South Korea	KOR	2019-01-01	28.6
2179	Kiribati	KIR	2019-01-01	28.3

Linear Regression: GDP per capita & Suicide Mortality Rate

To investigate what possible factors contribute to the suicide mortality rate, we first thought about GDP per capita. We think that GDP per capita and the suicide mortality rate have a strong connection as people who are poorer may be threatened by making a living and hence are more likely to commit suicides and die from suicides. To examine whether this is true, we first

graphed a scatterplot with GDP per capita on the x-axis and the suicide mortality rate on the y-axis shown in Figure 5. Every circle represents a country, and the size of the circle represents the population density of the country. We also fitted a linear line to the data points, and it seems that there is a positive correlation between GDP per capita and the suicide mortality rate, however, the positive trend is not very obvious.

Figure 5: Relationship between GDP and Suicide Mortality in 2019



To further investigate whether the GDP per capita has a significant effect on the suicide mortality rate is significant, we regressed GDP per capita on the suicide mortality rate (Table 2). The p-value of GDP per capita is 0.119 which is greater than 0.05 (using a 95% confidence interval), meaning that there is an 11.9% chance that GDP per capita has no effect on the suicide mortality rate. The R-squared is 0.014 which means that only 1.4% of observed suicide mortality rate data variation can be explained by the model. This shows that this linear regression model is not very good.

Table 2: Linear Regression Result

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                Suicide   R-squared:                       0.014
Model:                            OLS   Adj. R-squared:                  0.008
Method:                 Least Squares   F-statistic:                     2.449
Date:                Fri, 06 May 2022   Prob (F-statistic):              0.119
Time:                        22:23:26   Log-Likelihood:                -618.81
No. Observations:                 177   AIC:                             1242.
Df Residuals:                     175   BIC:                             1248.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      8.6213      0.848     10.171      0.000       6.948      10.294
GDP         4.549e-05   2.91e-05      1.565      0.119   -1.19e-05       0.000
==============================================================================
Omnibus:                      174.978   Durbin-Watson:                   0.127
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             4153.640
Skew:                           3.674   Prob(JB):                         0.00
Kurtosis:                      25.566   Cond. No.                     4.10e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.1e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Multiple Regression: Four Factors & the Suicide Mortality Rate

Since the previous linear regression perform poorly, we think that it may be due to there are multiple factors together affect the suicide mortality rate. As a result, we planned to do a multiple regression. We chose four variables that we considered relevant, and the four variables include GDP per capita, unemployment rate, life satisfaction, and the number of people with depression. Table 3 shows the result of the multiple regression. We see that GDP per capita and unemployment rate have an effect on the suicide mortality rate while life satisfaction and the number of people with depression seem to have no effect on the suicide mortality rate. Nevertheless, our multiple regression is still not working very well as the R-Squared (goodness of fit) is only 0.131. This means that a lot more factors have also caused the suicide mortality rate.

Table 3: Multiple Regression Result

OLS Regression Results

Dep. Variable:	Suicide	R-squared:	0.131
Model:	OLS	Adj. R-squared:	0.105
Method:	Least Squares	F-statistic:	4.969
Date:	Fri, 06 May 2022	Prob (F-statistic):	0.000919
Time:	22:23:26	Log-Likelihood:	-468.33
No. Observations:	137	AIC:	946.7
Df Residuals:	132	BIC:	961.3
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	8.5059	4.540	1.874	0.063	-0.474	17.486
GDP	0.0001	4.45e-05	2.555	0.012	2.57e-05	0.000
Unemployment	0.4472	0.129	3.474	0.001	0.193	0.702
Satisfaction	-0.7849	0.870	-0.902	0.368	-2.505	0.936
Depression	1.789e-08	1.06e-07	0.169	0.866	-1.91e-07	2.27e-07

Omnibus:	135.169	Durbin-Watson:	0.487
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2978.473
Skew:	3.356	Prob(JB):	0.00
Kurtosis:	24.834	Cond. No.	4.65e+07

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.65e+07. This might indicate that there are strong multicollinearity or other numerical problems.

Conclusions and Directions for Future Research

Based on our study, we found that the world's average suicide mortality rate has decreased from 2000 to 2019. Secondly, the suicide mortality rate among males is much higher than the suicide mortality rate among females. Thirdly, the suicide mortality rate is the highest for the 70+ years age group. Moreover, we discovered that GDP per capita does not have a significant impact on the suicide mortality rate which is contrary to our expectations. However, when we added more variables into the regression, the link between GDP and unemployment and suicide rates became substantial, showing that suicide mortality rates are not determined by one or two factors alone, but rather by the interaction of several. For our future research, more factors such as gender, alcohol consumption, literacy rate, etc. should be included in the multiple regression. It is more likely to see a better-fitted multiple regression if more variables are included.

Reference

Rawlins, W. C., & Bishop, M. (2018, June 29). Trying to stop suicide: Guyana aims to bring down its high rate. *National Public Radio*. Retrieved from https://www.npr.org/sections/goatsandsoda/2018/06/29/622615518/trying-to-stop-suicide-guyana-aims-to-bring-down-its-high-rate