According to WHO, suicide accounts for 1.4% of all deaths worldwide. To see the reasons behind this, we analyzed two main datasets: worldwide and US domestic data.

The worldwide data aims to describe the general suicide rates among different countries, time, gender and age. The gif shows the suicide rate change over time. The line plot shows the total suicide cases change over time. We also described the suicide cases in male and female separately. It’s worth noticing that suicide happens much more in male than in female. Then, we can see suicide happens differently among age groups. Generally, the 75+ group had highest suicide rate, while 5-14 age group had lowest rate all the time. The overall suicide rate is decreasing in the recent years. However, in US, the suicide rate is increasing year by year.

Data from several sources were joined together into a merged dataset. We constructed a shiny app to see the distribution of suicide cases in different states of US. Users can select year and gender. Through the shiny app, users can also select the risk factors and the top N states to see the distribution over states.

We used a linear regression to build the model. Main outcome is suicide rate for each state, candidate predictors are gun, alcohol, temperature, precipitation, marijuana, education, GDP and gender for each state. We used stepwise approach to select model. We made scatter plot to intuitively see the linear relationships. And we made correlation plot to compute the correlation coefficient of each two variables. The result is significant. According to the results, suicide rate is higher in states where there is a higher gun ownership rate, higher marijuana usage, higher ratio of males to females, lower temperature and lower educational attainment.

For more details, please go to our website or contact us.