

# Describe the data

The data which we will working on is a Suicide Rates Overview 1985 to 2016. It consist on 12 features/columns which are:

1. country
2. year
3. sex
4. age
5. suicides no
6. population
7. suicides/100k pop
8. country-year
9. HDI for year
10. gdp for year (\$)
11. gdp per capita (\$)
12. generation

For this study we have 27,800 different samples. Therefore our data will be represented in a dataframe of size  $27800 \times 12$ .

For our study and due to the big range on years. We will group the data by decades. After that we will analysis the rate of suicide by age, sex and population. We will also, study the overall count of suicide by country. Finally we will visualize the countries with an increasing rate of suicide and the countries with a decreasing rate of suicide.

## Reference

For reference a link of the raw data is below.

*[https : //raw.githubusercontent.com/sitronco/Capstone/master/master.csv](https://raw.githubusercontent.com/sitronco/Capstone/master/master.csv)*