

# Problem Statement - Suicide Rate Analysis

## Context

Close to 800 000 people die due to suicide every year, which is one person every 40 seconds. Suicide is a global phenomenon and occurs throughout the lifespan. Effective and evidence-based interventions can be implemented at population, sub-population, and individual levels to prevent suicide and suicide attempts. There are indications that for each adult who died by suicide there may have been more than 20 others attempting suicide.

## Objective

The objective of this case study is to find the patterns for increased suicide rates among different cohorts globally, across the socioeconomic spectrum by using exploratory data analysis.

## Data Dictionary

We will be using the dataset about suicide rates from 1985 to 2016. This dataset has the following attributes:

- country: Country
- year: Year
- sex: Sex (male or female)
- age: Suicide age range, ages divided into six categories
- suicides\_no: number of suicides
- population: population of that sex, in that age range, in that country, and in that year
- suicides/100k pop: Number of suicides per 100k population
- gdp\_for\_year(\$): GDP of the country in that year in dollars
- gdp\_per\_capita(\$): Ratio of the country's GDP and its population
- generation: Generation of the suicides in question, being possible 6 different categories.

## Questions to explore

- Is the suicide rate more prominent in some age categories than others?
- Which countries have the most and the least number of suicides?
- What is the effect of the population on suicide rates?
- What is the effect of the GDP of a country on suicide rates?
- What is the trend of suicide rates across all the years?
- Is there a difference between the suicide rates of men and women?