

# Suicide Rates Dataset

## Description

This spatio-temporal dataset is a combination of four different datasets to find the signals correlated to increased suicide rates among different cohorts globally, across socio-economic spectrum. The dataset contains for each (country, year, sex, age range) tuple number of suicides, population size of the demographic, suicides per 100k population, human development index for the year (HDI), and GDP.

Statistic	Value
Size	2.7 MB
# of instances (or rows)	27820
# of fields	12

## Fields

- **country:** the target country for which the data is being presented. The number of countries represented in the dataset is 101.
- **year:** the target year for which the data is being presented. The range of years in the dataset is from 1985 to 2016 with possible missing values.
- **sex:** the target sex
- **age:** the target age range. There are 6 cohorts being represented: 5-14 years, 15-24 years, 25-34 years, 35-54 years, 55-74 years, and 75+ years.
- **suicides\_no:** the number of suicides for a given country in a given year for the given sex and age group
- **suicides/100k pop:** rate of suicides per 100k population
- **country-year:** field representing the target country and year
- **HDI for year:** the Human Development Index (HDI) is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable, and having a decent standard of living. The HDI is the geometric mean of normalized indices for each of the three dimensions. There are possible missing values.
- **gdp\_for\_year (\$):** the Gross domestic product (GDP) that represents in dollars the total monetary or market value of all the finished goods and services produced within a country's borders in a specific time period
- **gdp\_per\_capita (\$):** the GDP per capita i.e. a country's economic output divided by its population
- **generation:** a field for the age range as it contains the names of the generations from the following list: Boomers, G.I Generation, Generation X, Generation Z, Millennials, and Silent.

## Example Hypotheses:

- Is GDP of a country in a year predictive of the suicide rates? How does this vary across different age groups (cohorts), and different sexes?
- Is HDP of a country predictive of its suicide rates?

## License:

**World Bank Dataset Terms of Use:** <https://www.worldbank.org/en/about/legal/terms-of-use-for-datasets>

## Data files:

*master.csv:* This is the only file in this dataset, and contains the entire dataset of 27820 rows and 12 fields.

**References:**

1. United Nations Development Program. (2018). Human development index (HDI). Retrieved from <http://hdr.undp.org/en/indicators/137506>
2. World Bank. (2018). World development indicators: GDP (current US\$) by country:1985 to 2016. Retrieved from <http://databank.worldbank.org/data/source/world-development-indicators#>
3. [Szamil]. (2017). Suicide in the Twenty-First Century [dataset]. Retrieved from <https://www.kaggle.com/szamil/suicide-in-the-twenty-first-century/notebook>
4. World Health Organization. (2018). Suicide prevention. Retrieved from [http://www.who.int/mental\\_health/suicide-prevention/en/](http://www.who.int/mental_health/suicide-prevention/en/)