## Self Harm and Substance Abuse Deaths Worldwide

## Data Source Information:

### Obtained CSV Files:

### (renamed from world\_wide\_self\_harm\_and\_stubstance\_deaths) for readability)

### Covers Data from Year/s:

Years 2017 to 2021

### Sourced From:

Kaggle –

<https://www.kaggle.com/datasets/thomaseltonau/self-harm-and-substance-abuse-deaths-worldwide>

### Original Source Data/Authors:

Original data is sourced from the World Health Organization (WHO) and was collected through the WHO Mortality Database:

World Health Organization (2023). WHO Mortality Database. <https://www.who.int/data/data-collection-tools/who-mortality-database>

\*Per Kaggle Author, Thomas Elton;

The data for the deaths by country were taken from the World Health Organization Mortality Database (2023)here. From the database, the “Country codes”, “Population and live births”, and “Mortality, ICD-10 (part 5/5)” files were downloaded and used in this analysis.

### License:

CC0: Public Domain

### Important Notes:

Files were cleaned and merged by Kaggle Author prior to publication of the dataset. These were the notes from the author on steps he took to create this dataset:

(Additionally this link leads to his portfolio detailed his code and rational from the cleaning process <https://rpubs.com/tjelton/WHO-Mortality-DB-Suicide-and-Substance-Abuse-Deaths> )

1. The goal of this report is to create a cleaned data set of the number of deaths from suicide and substance abuse per country.
2. There is lots of data cleaning here… Data is split into separate files, and within the Morticd10\_part5.csv, there are lots of columns and codes that are hard to interpret without looking at the documentation document.
3. The following data cleaning has been split into three parts.
4. The first part looks at the cleaning for the WHO Mortality data.
5. The second part looks at the cleaning for the WHO Population data.
6. The third part looks at stitching all the data together into one data set.
7. The fourth part uses an additional data set to get the three letter ISO code for each country in the stitched together data set.
8. The fifth part merges groups together all sex values (male, female, unspecified) into a separate “All” category.
9. The sixth part includes gathering the population data from “Our World in Data” to fill in for where WHO has missing population data.
10. The seventh part saves the merged data set into a single csv.