

Correlation of Location, Year, Economic Status, and Demographic Information with Suicide Rate

STAT 420, Summer 2020, Kun Hu, Madeline Old, Yuwen Xiang, Patrick Willhalm

07/18/2020

Description of data file

This file compares the number of suicides per 100,000 across categories stratified by age group, sex, country, and year. There are 27820 observations across 12 variables. Variables include country, year, sex, age group, number of suicides, suicide rate, Human Development Index (HDI) and Gross Domestic Product (GDP).

Background on the data set

This dataset is a compilation of data from the UN Development Program, World Bank, and the World Health Organization. <https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016>

Statement of Interest

Suicide is unique as a public health issue in that it is 100% preventable. Stopping deaths from cancer requires developing new cancer drugs, but suicide is a social and economic rather than technical problem. Here we intend to find the correlation among different social status (includes year, sex and age) and the economic situation with the suicide rate with an eye toward suicide prevention. Depending on social status, we may find some causal links with suicide risk.

```
suicide <- read.csv('master.csv') %>%  
  as_tibble(.)  
summary(suicide)
```

```
##          ountry          year          sex          age  
## Austria   : 382  Min.   :1985  female:13910  15-24 years:4642  
## Iceland   : 382  1st Qu.:1995  male   :13910  25-34 years:4642  
## Mauritius  : 382  Median :2002                35-54 years:4642  
## Netherlands: 382  Mean   :2001                5-14 years :4610  
## Argentina  : 372  3rd Qu.:2008                55-74 years:4642  
## Belgium    : 372  Max.   :2016                75+ years  :4642  
## (Other)    :25548  
## suicides_no  population  suicides.100k.pop  country.year  
## Min.      :    0  Min.      :    278  Min.      : 0.00  Albania1987: 12
```

```

## 1st Qu.: 3 1st Qu.: 97498 1st Qu.: 0.92 Albania1988: 12
## Median : 25 Median : 430150 Median : 5.99 Albania1989: 12
## Mean : 243 Mean : 1844794 Mean : 12.82 Albania1992: 12
## 3rd Qu.: 131 3rd Qu.: 1486143 3rd Qu.: 16.62 Albania1993: 12
## Max. :22338 Max. :43805214 Max. :224.97 Albania1994: 12
## (Other) :27748
## HDI.for.year gdp_for_year.... gdp_per_capita....
## Min. :0 1,002,219,052,968: 12 Min. : 251
## 1st Qu.:1 1,011,797,457,139: 12 1st Qu.: 3447
## Median :1 1,016,418,229 : 12 Median : 9372
## Mean :1 1,018,847,043,277: 12 Mean : 16866
## 3rd Qu.:1 1,022,191,296 : 12 3rd Qu.: 24874
## Max. :1 1,023,196,003,075: 12 Max. :126352
## NA's :19456 (Other) :27748
## generation
## Boomers :4990
## G.I. Generation:2744
## Generation X :6408
## Generation Z :1470
## Millenials :5844
## Silent :6364
##

```

suicide

```

## # A tibble: 27,820 x 12
##   ountry year sex age suicides_no population suicides.100k.p~
##   <fct>   <int> <fct> <fct>   <int>   <int>   <dbl>
## 1 Albania 1987 male 15-2~    21    312900    6.71
## 2 Albania 1987 male 35-5~    16    308000    5.19
## 3 Albania 1987 fema~ 15-2~    14    289700    4.83
## 4 Albania 1987 male 75+ ~     1     21800    4.59
## 5 Albania 1987 male 25-3~     9    274300    3.28
## 6 Albania 1987 fema~ 75+ ~     1     35600    2.81
## 7 Albania 1987 fema~ 35-5~     6    278800    2.15
## 8 Albania 1987 fema~ 25-3~     4    257200    1.56
## 9 Albania 1987 male 55-7~     1    137500    0.73
## 10 Albania 1987 fema~ 5-14~     0    311000     0
## # ... with 27,810 more rows, and 5 more variables: country.year <fct>,
## # HDI.for.year <dbl>, gdp_for_year.... <fct>, gdp_per_capita.... <int>,
## # generation <fct>

```