Global Suicide Trends and Analysis

# Introduction

According to the [WHO Suicides organization](https://www.who.int/news-room/fact-sheets/detail/suicide), there are close to 800 000 people die due to suicide every year, which is one person in every 40 seconds. Many more attempt suicide. Suicide occurs throughout the lifespan and is the second leading cause of death among 15-29 year old globally. It is a global phenomenon.

The outcome variable of concern is number of suicides. I’d like to analyze the data to see whether the suicide rate is decreasing or increasing among different countries and to find any correlation between number of suicides and other variables of the dataset.

I will use R to analyze the global trend of suicides, analyze the rate of suicide based on gender, age, continent and investigate to see if there is a relationship between numbers of suicides and the gross domestic product (GDP) per capitaof the countries.

# Literature Review

According to an article published in Cambridge university website [1],the relationship between suicide rates and age was examined in both sexes in different age groups, from the World Health Organization website for all the 62 listed countries and there was a significant increase in suicide rates with increasing age in males and females in 25 and 27 countries respectively.

A US suicide rate analysis [2] shows that there has been a sharp and steady increase in suicide rates in the US since 2000. According to this analysis, from the average suicide rates by genders, suicide rate is much higher among males than females. From the average suicide rates by age groups (20–44 years, 45–64 years, and 65 years and above) in general in the US, the average suicide rate is higher for older adults than for the other two age groups.

The use of regression methods for the investigation of trends in suicide rates in Hungary between 1963 and 2011 [3] shows that suicide trend had an overall peak in 1982 and remained constant after 2006. There was a significant increase until 1982 (an increment of \*20 suicides per 100,000 persons) during two decades. This was followed by a significant decrease during the next about two decades, when suicide rates decreased approximately to the same level where they started in 1963, and with relative constancy after 2006 overall. Similar trends were observed in both genders. Risk estimates of suicide rates were calculated and was two times more in males than in females overall.

It seems that for most of the European countries, the suicide rates correlated strongly with GDP per capita and rather weakly with unemployment and the other economic indices. This correlation was strong within each country and less strong across those countries. [4]

According to a research on worldwide impact of economic cycles on suicide trends, to investigate the trends and correlations of gross domestic product (GDP) adjusted for purchasing power parity (PPP) per capita on suicide rates, it’s strongly correlates to suicide rates worldwide, and the direction and magnitude of the correlation differs between developing and developed countries. [5]

Impact of 2008 global economic crisis on suicide in 54 countries was studied [6].The study shows that suicide rates in several European countries increased after the 2008 global recession. Similar cases were seen in American countries as well, specifically among men, and in countries experiencing greater increase in unemployment.

About suicide rates among adolescents, a study shows that late adolescence (15 to 19 years old) and male gender is universal risk factors for adolescent suicide. Male adolescent suicide rates were about 2–3 times higher than female rates in every country in this study.[7]

A study was done on suicide rates in older adults in Queensland, Australia, during the years 2000–2012 [8]. It showed that the average suicide rate for 2000–2012 increased with age for males, whereby the highest suicide rate was observed for oldest old males (85 years and over) but, in females it was relatively stable.

Based on a study to analyze the association between generous unemployment benefit programs and suicide rates among different states in US, total suicide rates increased as unemployment rates rose. However, the positive association between unemployment rates and suicide was greater for states and years with maximum unemployment benefits below the sample mean than for states and years with more generous unemployment benefits. [9]

An analysis was done about Suicide rates and the average income per household in Japan.It showed that the average income per household was related to the suicide rate among the total population and among men. However, the average income per household was not related to the suicide rate among women.[10]Discussion Based on Analysis of the Suicide Rate and the Average Disposable Income per Household in Japan

# Dataset and Data Description

This compiled dataset is downloaded from Kaggle website which is pulled from four other datasets linked by time and place. Those four datasets are mentioned at the end of this document. [11]

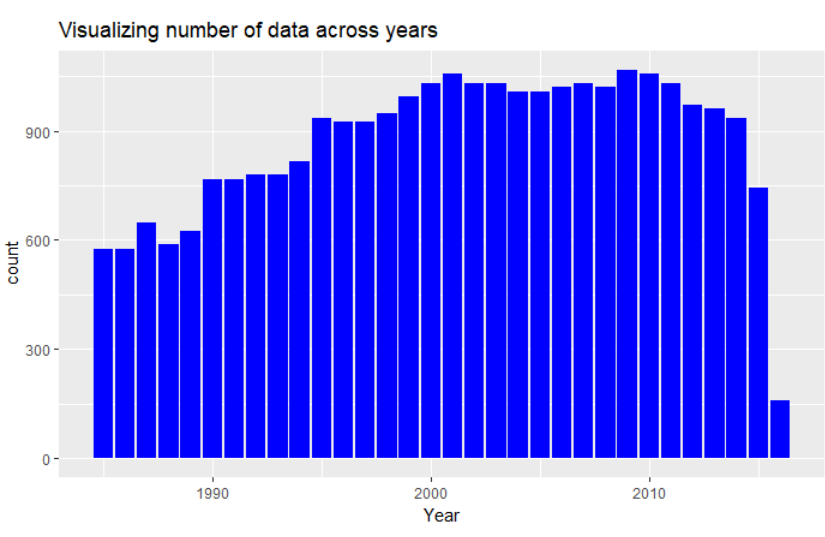
The data set is about the Suicide rates across the world from 1985 to 2016 which has been provided by the World Health Organization(WHO). It consists of the suicide count across various countries which is categorized based on the gender and age groups. The data explores the factors that seem to have some correlation with suicide rates among different groups across the world. The population of the countries is mentioned. The intention is to visualize and analyze how the factors affect the suicide rates.

The outcome variable of concern is suicide\_no, i.e., number of suicides. For this purpose, data from WHO has been chosen to identify the past trends and patterns to predict which groups are most affected. I expect HDI for Year, GDP Per Capita, Country, Sex and Ageto be associated with the outcome variable :suicides/100k population(Suicide Rate).

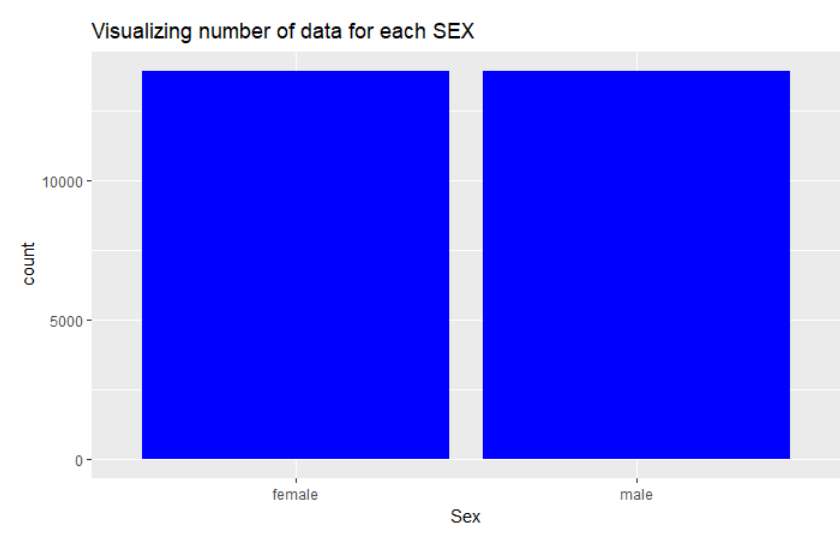
HDI or the human development index of a country measures the life expectancy, education and living standards of a country. Lower HDI is a reflection of unhappy population. Similarly, for GDP of a country: The studies have shown a strong association between economic activities with suicide rates.

Dataset contains of below attributes:

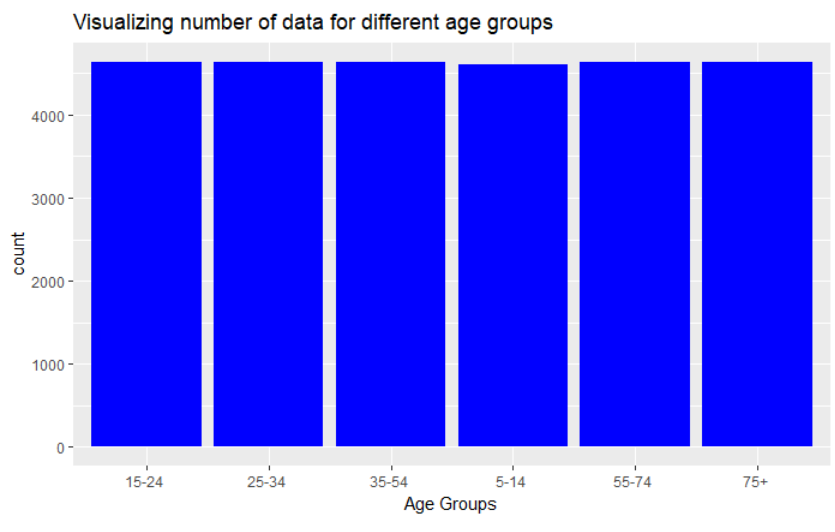
1. Country: 93 Countries in Dataset [Categorical]
2. Year :from 1985 to 2016 [Numerical]



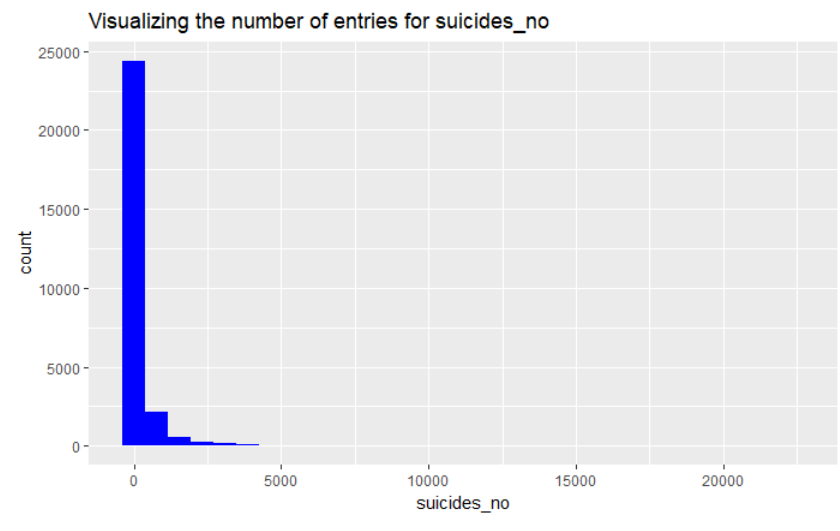
1. sex :male/female [Categorical]



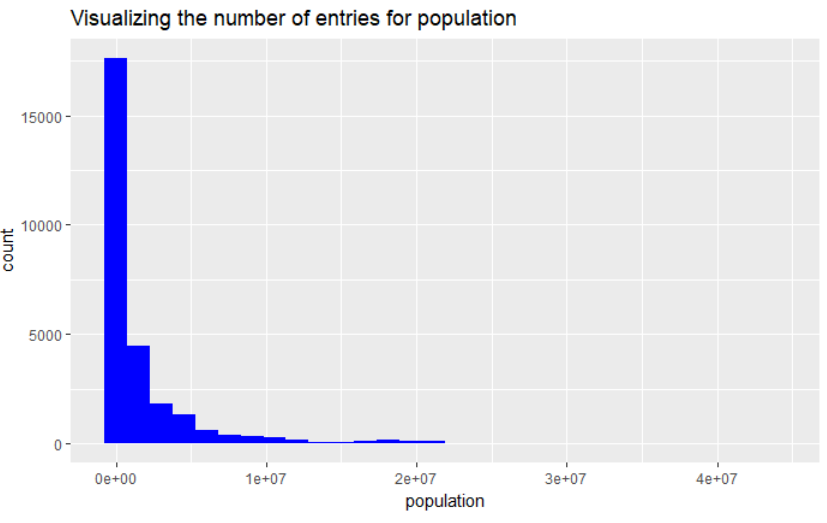
1. Age :Agehas been summarizedwithin specific range [Categorical]



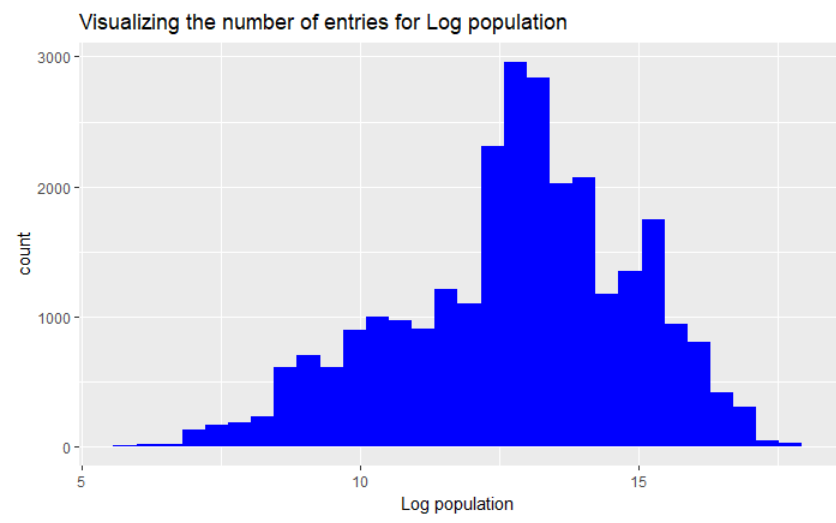
1. suicides\_no: number of suicides [Numerical]

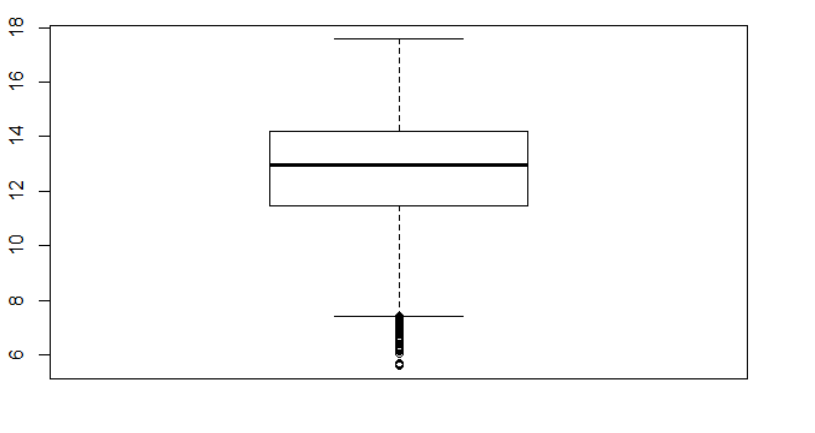


1. Population: shows the population of the given country [Numerical]



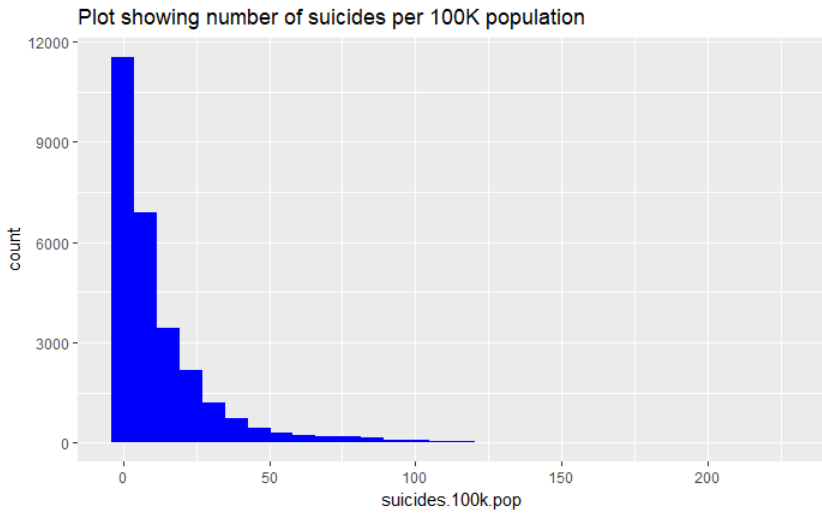
It’s skewed to the right. I applied Log on the values of that variable.



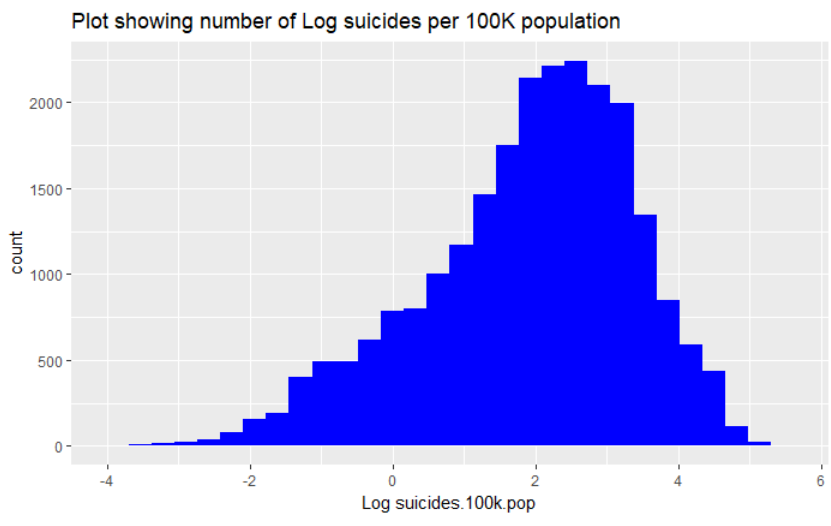


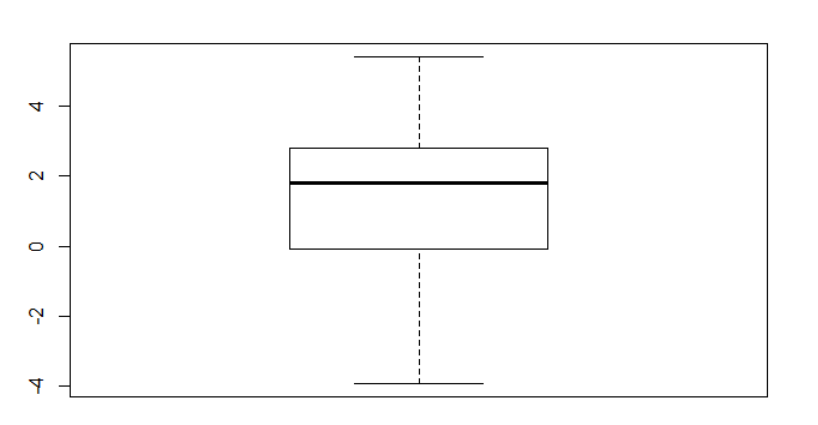
Boxplot shows that there are some outliers in the Log of population data.

1. suicides/100k pop : Suicide by population of 100,000 people [Numerical]



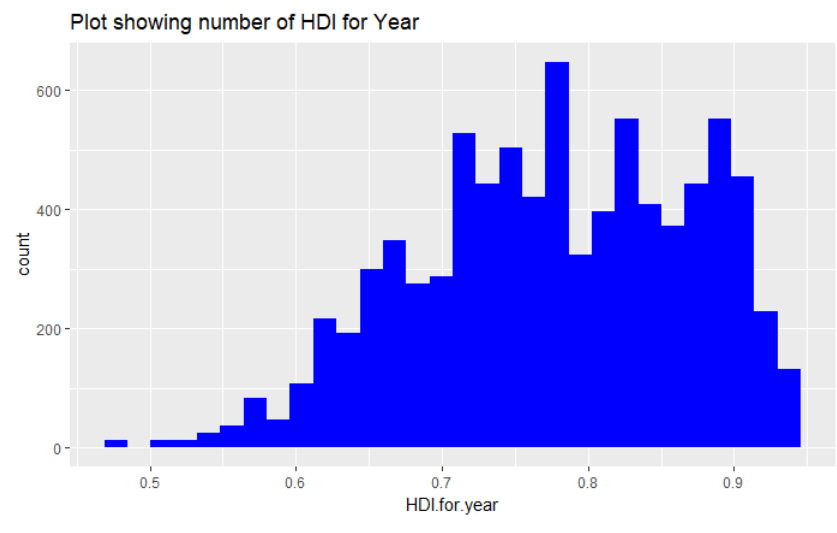
It’s skewed to the right. I applied Log on the values of that variable.



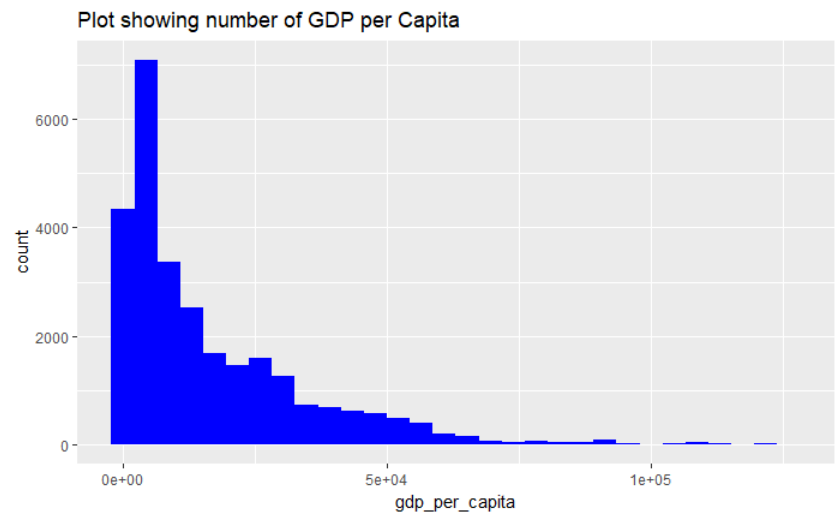


There isn’t any outlier in the Log of suicides/100k pop values.

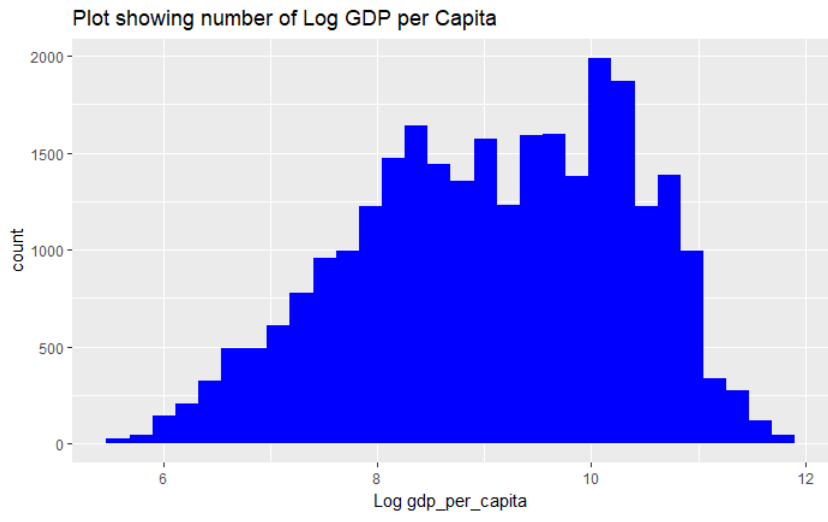
1. country-year : combination of Country and year [Categorical]
2. HDI for year: The Human Development Index is a simple factor for measuring the average degree of accomplishment in a country regarding three aspects of human development: Health, Education and Income [Numerical]

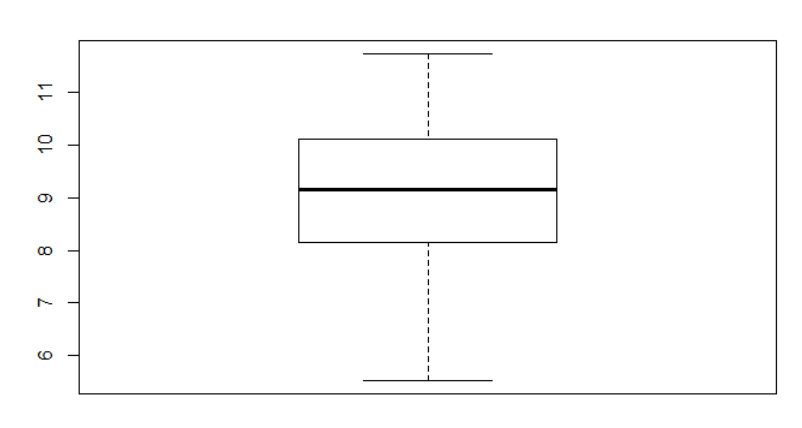


1. gdp\_per\_capita: The GDP per capita was defined as GDP divided by mid-year population and is in Dollars [Numerical]



It’s skewed to the right. I applied Log on the values of that variable.





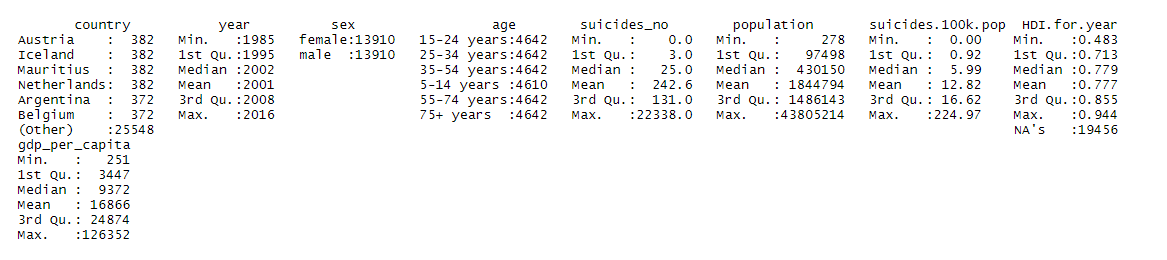
There isn’t any outlier in the Log of gdp\_Per\_Capita values.

# Approach

## Step 1: Data Cleaning

* “HDI for year” column is removed because more than 2/3 of the values are NA
* Data for 2016 was removed because only a few countries had data for that year
* Continent data is added to the dataset by using the “countrycode” package

Below is the summary of cleaned data:



## Step 2: Exploratory Data Analysis

I checked the global trend of suicide rates over time, the global trend by continent, by sex, by age and by country.

Also I analyzed the top 12 most increasing and decreasing trends among countries, gender difference trend by continent and age difference trend by continent.

## Step 3: Linear Regression Analysis

I did a regression analysis and found a week but significant positive linear relationship between GDP\_per\_capita and Suicide rates.

There are significant positive linear relationships between number of suicides and being a "Male" , living in “Asia”and the "Population“ of the country.

Here is a Github link to the code:

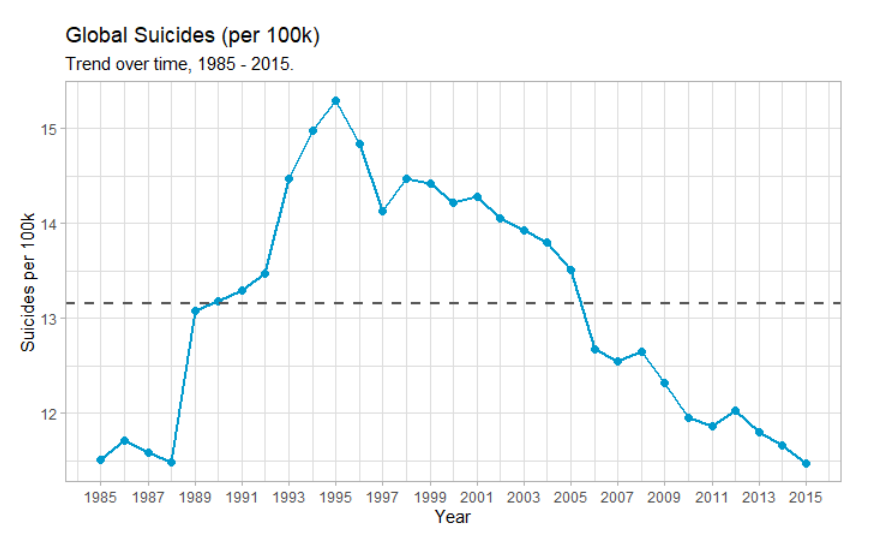
# <https://github.com/neda-noor/Capstone_NedaNoorbakhsh/blob/master/Capstone_NedaNoorbakhsh_FinalCode_2.Rmd>

# Results

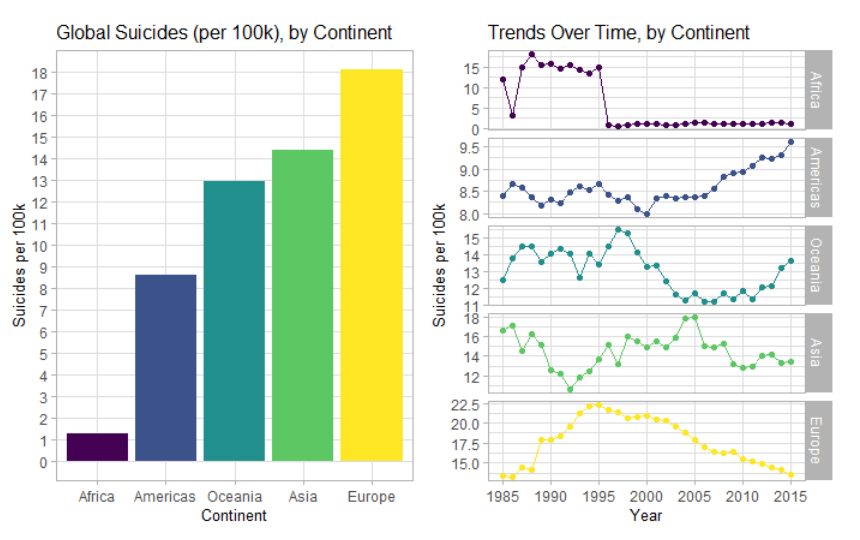
1.**Global Trend**:

As it’s shown in below diagram, the suicide rate is decreasing globally.

Peak suicide rate was about 15 deaths per 100K in 1995 but since then it decreased steadily to 11.5 deaths per 100K in 2015.



2.**By Continent:**

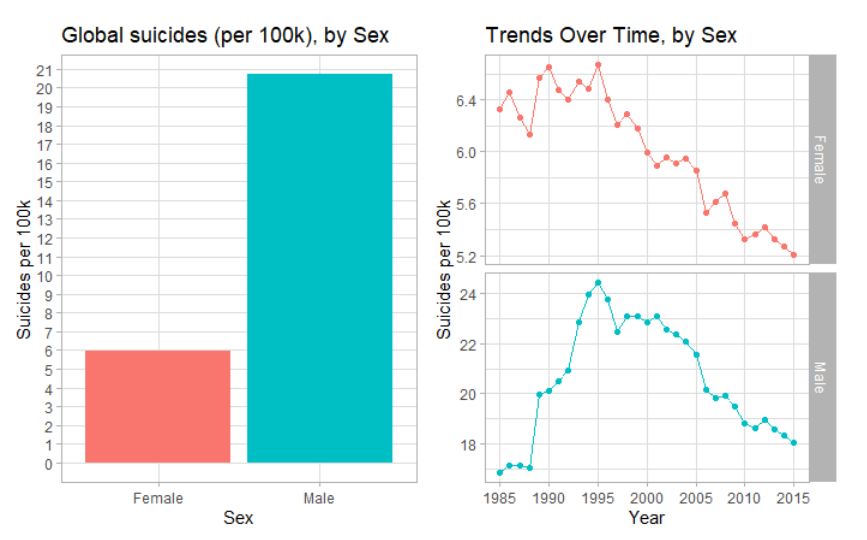
****

European countries have the highest suicide rate in overall, but it has steadily decreased since 1995 and its rate in 2015 is almost similar to Asia and Oceania.

The trend line for Africa is due to poor quality of data as only 3 countries have provided data.

Among all continents, the trends for Oceania and Americas are more concerning because they are increasing.

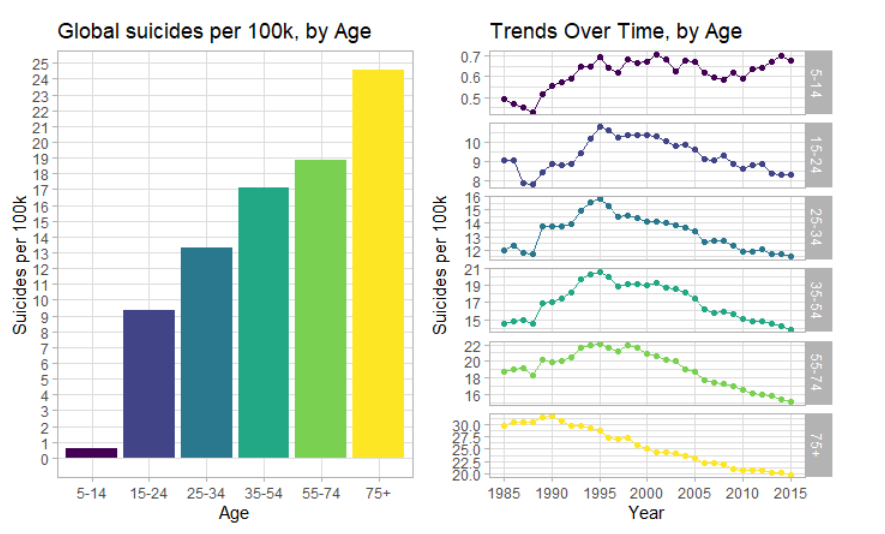
3.**By Sex:**

****

Globally, the suicide rate in men has been about 3.5 times higher than women.

The peak amount for both men and women was in 1995 and it has been declined for both genders since then.

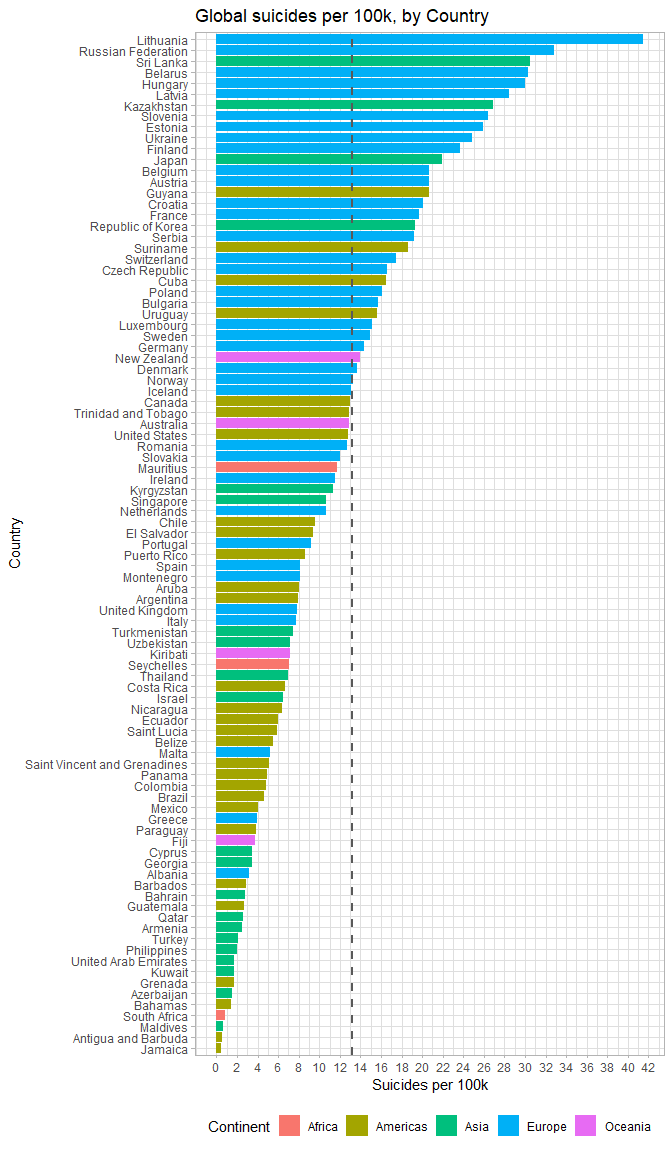
3.**By Age:**

****

As we can see in above charts, the number of suicides increases with age globally.

Since 1995, suicide rate for those of 15 years old and up, has been linearly decreasing but it remained almost steady and small for 5-14 years old category as about 1 person per 100K per year.

4.**By Country:**

****

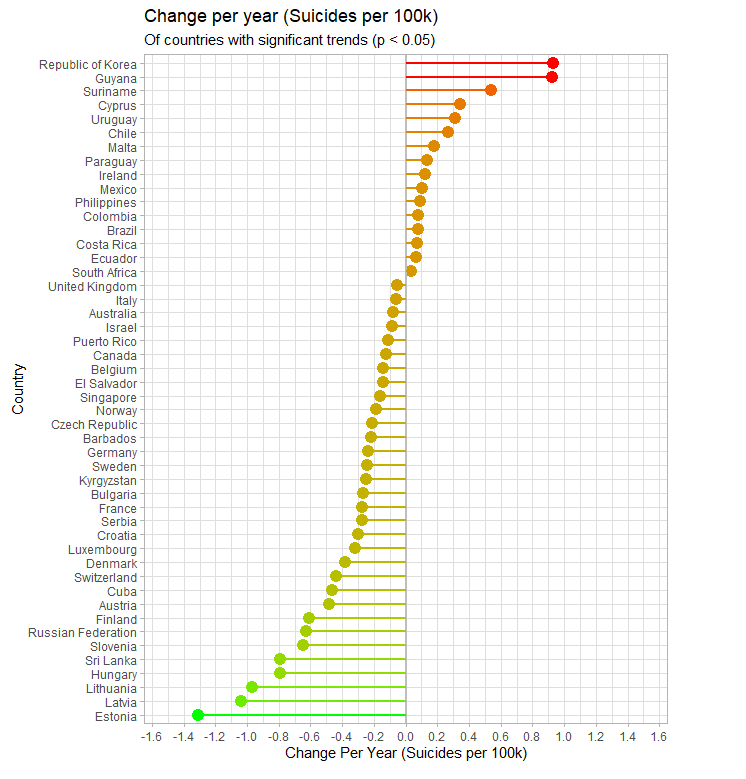


Lithuania has the highest suicide rate among all other countries of our dataset.

There are lots of European countries with high suicide rates and a few with low rates among all countries.

5.**Linear Trends:**

**5.1: Trends within each country over years:**

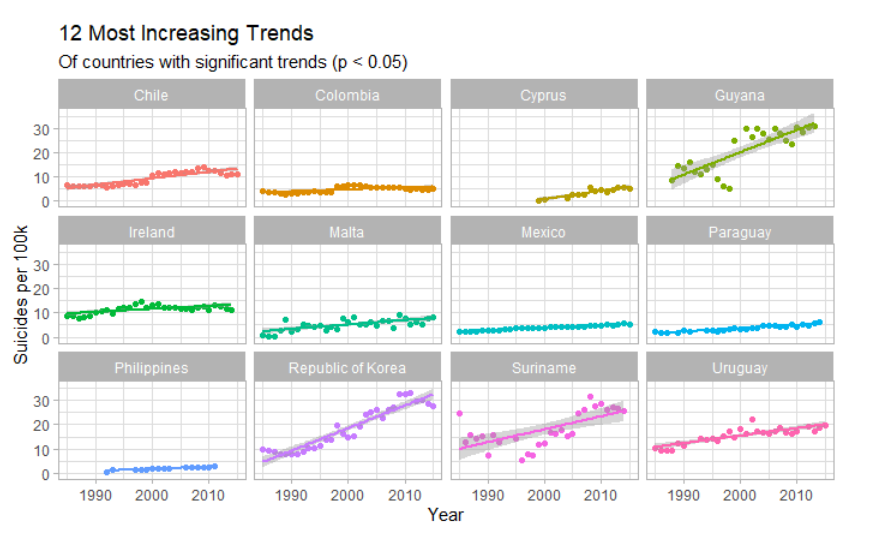
****

To check how the suicide rate is changing over time within each country, I fitted a simple linear regression to every countries data and extracted those which have “year” P-Value < 0.05

As we can see in the diagram, suicide rate in 48 countries is changing linearly as time passes.

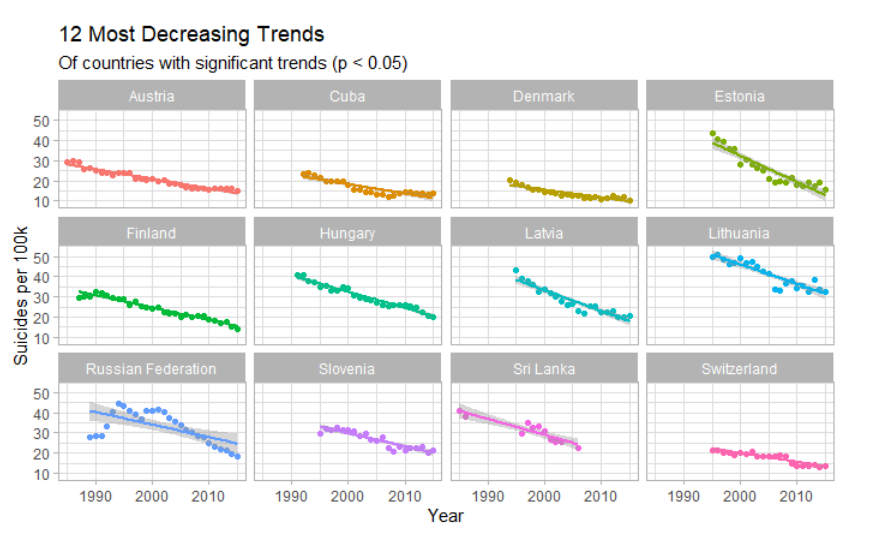
It’s decreasing in 32 countries out of 48 countries over time.

**5.2. Most Increasing Trends:**

****

South Korea and Guyana have the steepest increase among other countries.

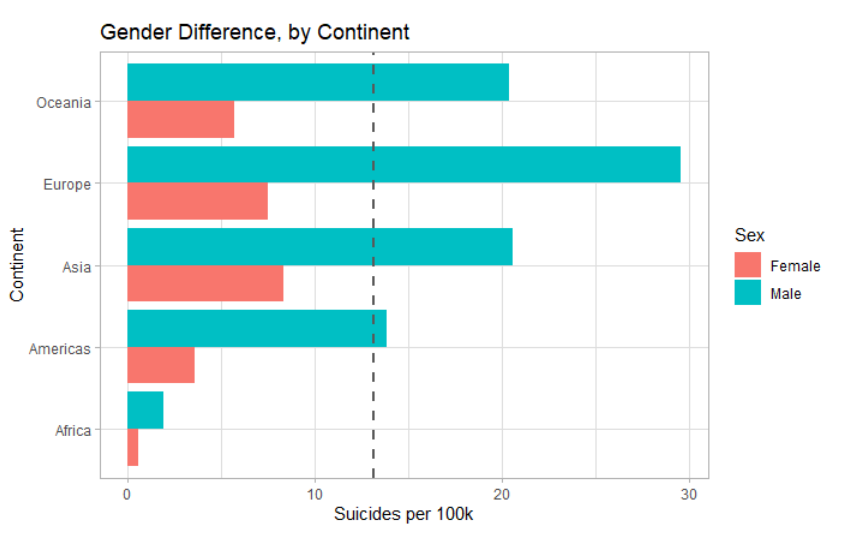
5.3.**Most Decreasing Trends:**



Estonia has steepest decrease in suicide rate.

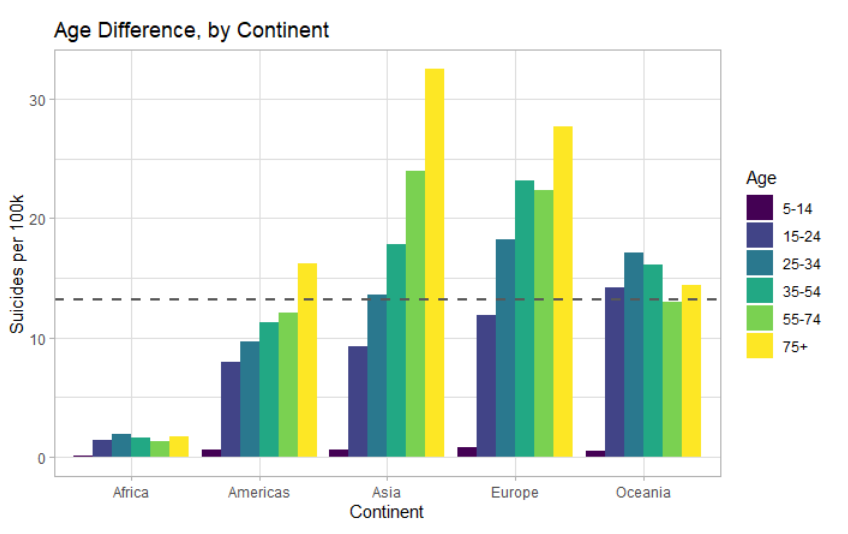
For Russia, it only began to decrease in 2002. Since then it has decreased about 50%

6.**Gender difference by Continent:**

****

European men had the highest suicide ratebetween 1985 – 2015, with about 30 suicides per 100K, per year and it was about 4 times more than women.

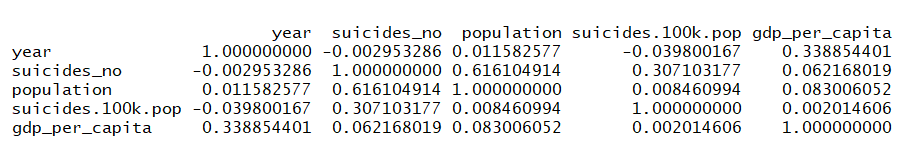
7.**Age difference by Continent:**

****

In America, Asia and Europe, suicide rate increases with age but in Oceania and Africa, the highest rate is for those with 25-34 years old.

# 8. correlation between numerical attributes of the dataset:

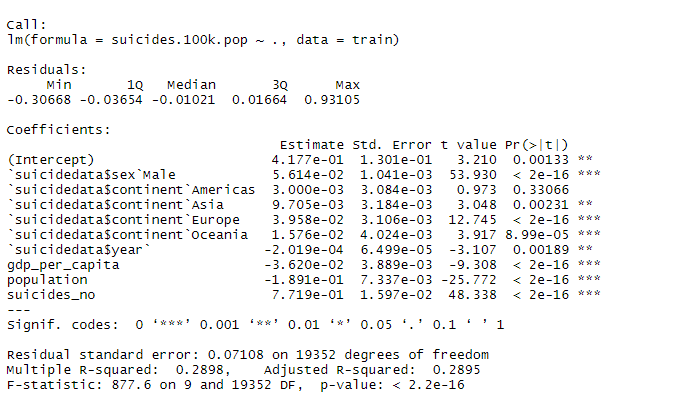
I checked the correlation between all the numerical attributes of the dataset:



It shows that there is a positive correlation between “suicides.100k.pop” with “population” and “gdp\_per\_capita” and a negative correlation with “year”. But the values are so small and close to zero.

**9. multiple liner regression:**

Below is the summary of multiple linear regression between suicides.100k.pop and other attributes of the dataset:

****

# gdp\_per\_capita,population,suicides\_no,suicides.100k.pop are normalized.

# The \*\*p-value\*\* of the model is < 2.2e-16. This means that, at least, one of the predictor variables is significantly related to the outcome variable.

# Except for America, the p-value for other attributes of the model is < 0.05 which means we can reject H0 that there are no association between those attributes and the number of suicide.

# Conclusions

I did an Exploratory Data Analysis on a suicide dataset originated from WHO Suicides organization.

It required some data cleaning such as removing attributes which had lots of missing values. Also I added Continent attribute to the dataset in order to be able to do some analysis by the continent value.

Based on my analysis, the peak suicide rate was in 1995 but it has been decreasing steadily since then globally. However, the analysis shows that the suicide rate for Oceania and Americas is increasing in 21st century.

Also the suicide rate for men has been almost 3.5 times higher than for women globally.

The analysis of suicide and age showed that on average, suicide rate increases with age.

**References:**

[1] Shah, A. (2007).

The relationship between suicide rates and age: An analysis of multinational data from the World Health Organization.

*International Psychogeriatrics*

[*https://doi.org/10.1017/S1041610207005285*](https://doi.org/10.1017/S1041610207005285)

[2] Erick Kitenge , Md RafayetAlam, SediqSameem (2019)

Convergence of U.S. suicide rates: What does it imply?

Economic Analysis and Policy Journal

<https://www.journals.elsevier.com/economic-analysis-and-policy>

[3] Anna M. Laszlo • Adam Hulman• JozsefCsicsman• Ferenc Bari • Tibor A. Nyari (2014)

The use of regression methods for the investigation of trends in suicide rates in Hungary between

1963 and 2011

Soc Psychiatry Psychiatr Epidemiol (2015) 50:249–256

DOI 10.1007/s00127-014-0926-9

[4] Konstantinos N. Fountoulakis, Wolfram Kawohl,… (2014)

Relationship of suicide rates to economic variables in Europe: 2000–2011

The British Journal of Psychiatry (2014)

205, 486–496. doi: 10.1192/bjp.bp.114.147454

[5] Blasco-Fontecilla H, Perez-Rodriguez MM,Garcia-Nieto R, et al.

Worldwide impact of economic cycles on suicide trends over 3 decades:

differences according to level of development. A mixed effect model study. BMJ

Open 2012;2:e000785.doi:10.1136/bmjopen-2011-000785

[6] Shu-Sen Chang, David Stuckter, Paul Yip

Impact of 2008 global economic crisis on suicide: time trend study in 54 countries

BMJ 2013;347;f5239 doi:10.1136/bmj.f5239

[7] Beop-Rae Roh, EunHee Jung, Hyun Ju Hong

A Comparative Study of Suicide Rates among 10-19-Year-Olds in 29 OECD Countries

Psychiatry Investigation 2018;15(4):376-383.

Published online: February 28, 2018

DOI: <https://doi.org/10.30773/pi.2017.08.02>

[8] Yu Wen Koo, KairiKõlves and Diego De Leo

Suicide in older adults: differences between the young-old, middle-old, and oldest old

*International Psychogeriatrics* (2017), 29:8, 1297–1306 © International Psychogeriatric Association 2017

doi:10.1017/S1041610217000618

[9] Jonathan CylusM. , Maria Glymour, Mauricio Avendano

Do Generous Unemployment Benefit Programs Reduce Suicide Rates? A State Fixed-Effect Analysis Covering 1968–2008

American Journal of Epidemiology

DOI: 10.1093/aje/kwu106

[10] K Inoue, Y Nishimura, Y Okazazi, and T Fukunaga

Discussion Based on Analysis of the Suicide Rate and the Average Disposable Income per Household in Japan**Discussion Based on Analysis of the Suicide Rate and the Average Disposable Income per Household in Japan**

# Discussion Based on Analysis of the Suicide Rate and the Average Disposable Income per Household in Japan

West Indian Med J. 2014 Aug; 63(4): 340–343.

Published online 2015 Mar 17. doi: 10.7727/wimj.2012.298

[11] Data References:

United Nations Development Program. (2018). Human development index (HDI). Retrieved from http://hdr.undp.org/en/indicators/137506

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#

[Szamil]. (2017). Suicide in the Twenty-First Century [dataset]. Retrieved from https://www.kaggle.com/szamil/suicide-in-the-twenty-first-century/notebook

World Health Organization. (2018). Suicide prevention. Retrieved from http://www.who.int/mental\_health/suicide-prevention/en/