# Suicides Prevention Analysis

Based on GDP and HDI Indices per Country
for the Period from 1985 to 2015

Garegin Manvelyan | Gurkiran Kaur

September 3rd, 2023

# Goal

The goal of this research is to analyze the relationship between suicide rates and socio-economic indicators, specifically GDP and HDI indices, for various countries over the period from 1985 to 2015. We aim to identify potential correlations and patterns to inform suicide prevention strategies.

# Glossary

**GDP (Gross Domestic Product):**

GDP is a way to measure the total value of everything a country makes and does in a year, like goods, services, and more.

**HDI (Human Development Index):**

HDI is a measure that looks at how healthy, educated, and well-off people are in a country to see how developed it is.

# Dataset Selection

The "Suicide Rates Overview 1985 to 2016" dataset was chosen because it intrigued us due to its interesting nature.

Upon initial examination, the data appeared to be effectively compiled, revealing interesting patterns.

Furthermore, the prospect of mitigating the increasing trend in suicide rates was found to be motivating.

Dataset was downloaded from kaggle online platform.

# Dataset Description: 12 Fields / 27820 Rows

1. **Country:** The name of the country where the data was recorded.
2. **Year:** The specific year when the data was collected (e.g. 1985)
3. **Sex:** The gender of the individuals being studied (e.g. male)
4. **Age:** Age groups or ranges used for grouping individuals (e.g.5-14)
5. **Population:** The total number of individuals within a particular country, age group, and sex category.
6. **Country-year:** A field that combines the country name and the year, providing a unique identifier for each data entry.
7. **Suicides_no:** Indicates the number of suicides within a particular country, age group, and sex category.
8. **Suicides/100k pop:** The number of suicides per 100,000 population.
9. **GDP per capita:** Indicates the average economic well-being of an individual in the specified country.
10. **GDP for year:** The Gross Domestic Product value for a specific year and country.
11. **HDI for year:** The Human Development Index value for a specific year and country.
12. **Generation:** A categorical field that represents the generation to which an individual belongs (e.g.Generation X)

# EDA: Cleaning

**HDI for year:** 19’456 nulls were replaced by mean values

**Country-year:** Removed, correlated with country and year fields

**GDP for year:** Removed, correlated with GDP per capita

**Duplicates:** Not detected

**GDP per capita:** range of data from 251 to 126352, was normalized

**Age and Sex:** pivoted, created columns for each age and sex categories with suicides rate as values

| sex | age | suicides/100k pop |
|---|---|---|
| male | 15-24 years | 6.71 |
| male | 35-54 years | 5.19 |
| female | 15-24 years | 4.83 |

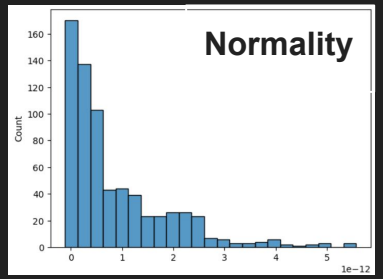
| 75+ years | 55-74 years | 35-54 years | 25-34 years | 15-24 years | 5-14 years | male | female |
|---|---|---|---|---|---|---|---|
| 7.40 | 0.73 | 7.34 | 4.84 | 11.54 | 0.00 | 20.50 | 11.35 |
| 9.97 | 4.88 | 5.87 | 3.70 | 8.04 | 0.00 | 18.91 | 13.55 |

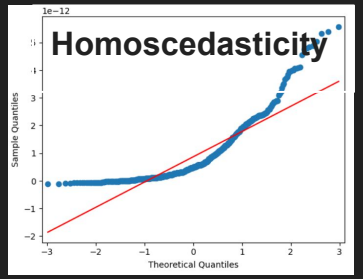
# Regression Models : 6 Attempts

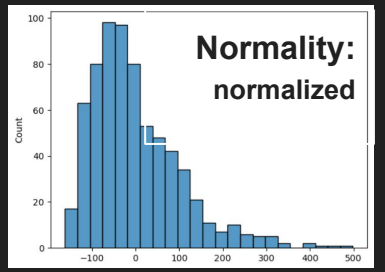
**Creating a predictive model that can help understand relationship between suicides rate and dependant values.**

1. **Linear model with all dependable variables : failed**
2. **Linear model with GDP and HDI : failed**
3. **Linear model with GDP only : failed**
4. **Linear model with GDP and dropped zeros for number of suicides: failed**
5. **Linear model with normalized GDP and HDI : failed**
6. **Logistic model with GDP only : failed**

**Visuals - linear regression assumptions:**

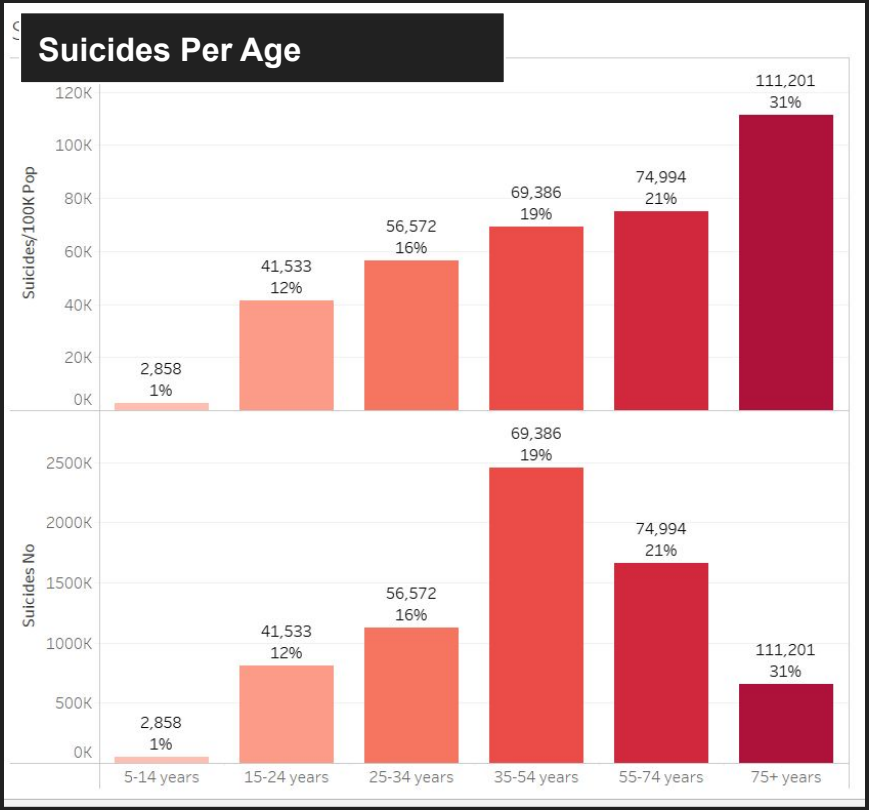


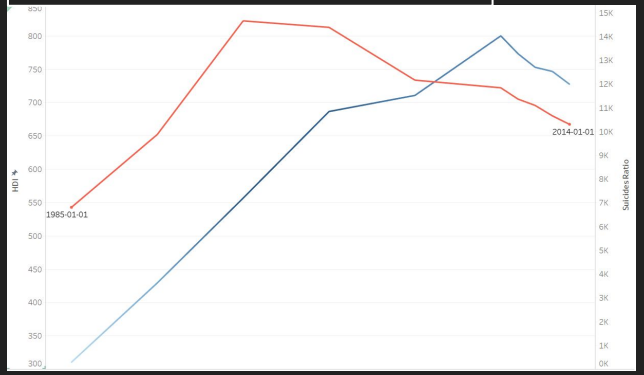



# Data Visualizations

## Suicides Per Age



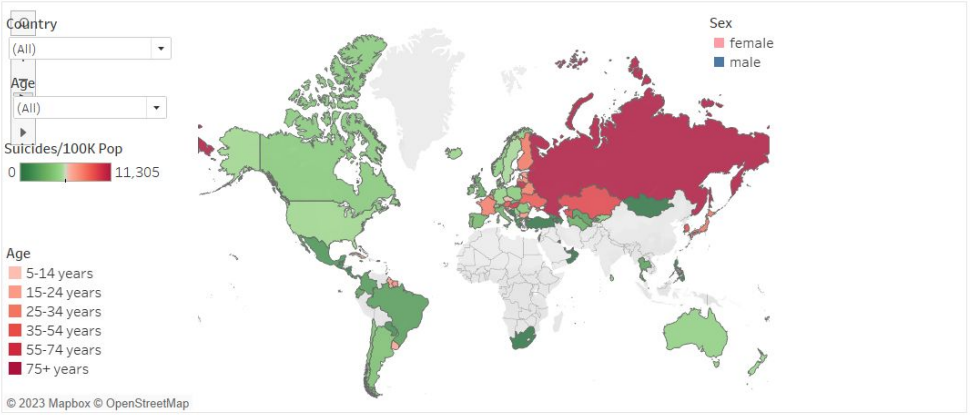

## Suicides Rate VS HDI

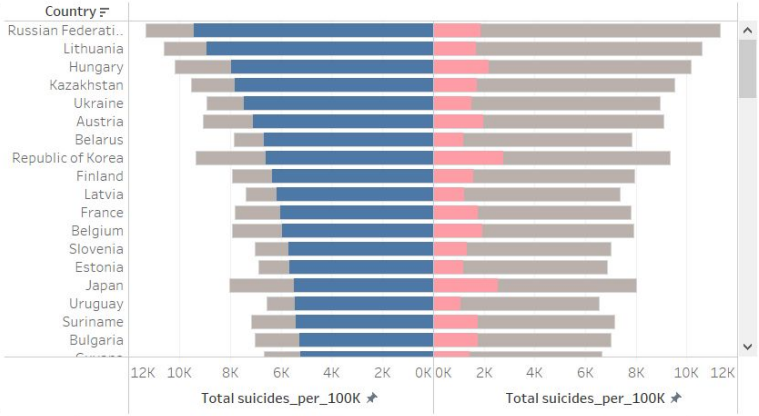


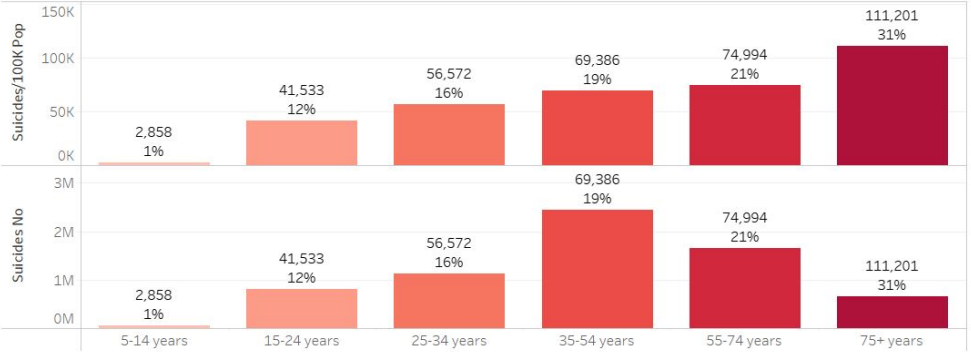

## Suicides Rate VS GDP

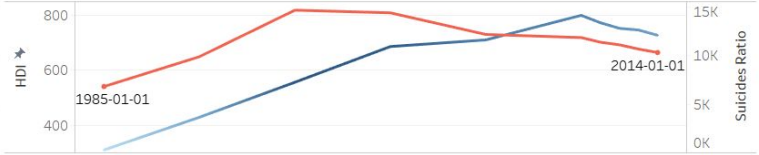
# SUICIDES OVERVIEW FOR PERIOD FROM 1985 TO 2015

| Average Suicides Per Year | Number of Suicides Over Years | Number of Countries |
|---|---|---|
| 596 | 6,748,420 | 101 |



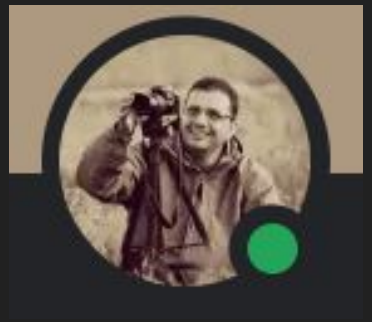



# Conclusion:

Despite building six distinct models, none of them provided conclusive evidence of a relationship between the number of suicides and our chosen variables.

More in-depth analysis is required, and we should consider incorporating additional parameters, such as ethnicity, religion, or cultural factors, into our research to gain a deeper understanding of this complex issue.

# Thank you!



Garegin Manvelyan

Gurkiran Kaur