

# **Exploring the Relationship Between Suicide Rates and Global Happiness: A Multi-Dimensional Analysis**

**IS6335 Data Visualization**

**Final Report**

**Semester A 23/24**

Li Xiaoxue 58166334

<b>Description of Dataset .....</b>	<b>3</b>
<b>Research objectives and research background .....</b>	<b>3</b>
<b>Visualization techniques .....</b>	<b>3</b>
<b>Data Visualization and Analysis .....</b>	<b>5</b>
<i>Part I: Basic global situation of suicide .....</i>	<i>5</i>
<i>Part II: Correlated Variables Visualization .....</i>	<i>8</i>
<i>Part III: Visualization of Korea .....</i>	<i>11</i>
<b>Research Implications .....</b>	<b>13</b>
<i>Part I: Basic Situation Analysis .....</i>	<i>13</i>
<i>Part II: Analysis of Relevant Variables with Suicide Rate .....</i>	<i>13</i>
<i>Part III: Situation analysis of Korea .....</i>	<i>14</i>
<b>Future Extension .....</b>	<b>15</b>
<b>References .....</b>	<b>17</b>

## Description of Dataset

To visualize the suicide situation around the world and find out the possible variables relative to the global suicide rate. Two Datasets are quoted: *Suicide Rates Overview 1985 to 2016* and *World Happiness Report*<sup>1</sup>. The first one is a compilation of information from four different datasets that are linked based on time and location, which includes 12 columns and 27,000+ ranks. Its purpose is to identify correlations between various factors and the increase in suicide rates among different global cohorts across the socio-economic spectrum. The other one is some indexes related to happiness all over the world. The focus of the analysis will be solely on the data from 2015, allowing for a more comprehensive examination of the dependent variables that impact suicide rates.

## Research objectives and research background

The goal of the research is to analyze the relationship between suicide rates as well as happiness indices and try to identify the factors that have a greater impact on suicide rates. By integrating these two datasets, the association between suicide rates and different social, economic, political, and cultural factors can be explored. Studies like this can help improve the understanding of suicide and may provide clues and policy recommendations on how to reduce suicide rates.

Political Dimensions: Trust (Government Corruption), Freedom, Country, Region;

Economic Dimensions: GDP;

Cultural Dimensions: Dystopia Residual, Generosity;

Social Dimensions: Health (Life Expectancy), Family, Population, Generation, Age, Sex.

## Visualization techniques

---

<sup>1</sup> Helliwell, J. F., Layard, R., & Sachs, J. D. (2015). World Happiness Report 2015 (Report No. ISBN 978-0-9968513-2-9). Sustainable Development Solutions Network. Retrieved from <https://worldhappiness.report/ed/2015/>

MySQL: In order to present the link between the suicide rate and the indicators that affect the happiness index, the first step would be to link the two tables in MySQL. This can be achieved by identifying the columns that are consistent across both tables, such as the country and year columns. Confirming their one-to-one relationships will allow for the linking of primary keys and foreign keys, facilitating the integration of the data.

Python: Python is well suited for data consolidation and cleaning, and can produce the denormalized relational schema, the merged data csv. Besides, python offers various libraries such as pandas, matplotlib to visualize radar graph, scatter plots, etc. These visualizations can effectively demonstrate the relationship between suicide rates and the indicators influencing the happiness index.

Tableau: Various visualizations can be created using drag-and-drop functionality. Tableau offers a wide range of visualization options, including maps, charts, and dashboards. Interactive dashboards can be designed to explore the relationship between suicide rates and different indicators by applying filters and drilling down into the data.

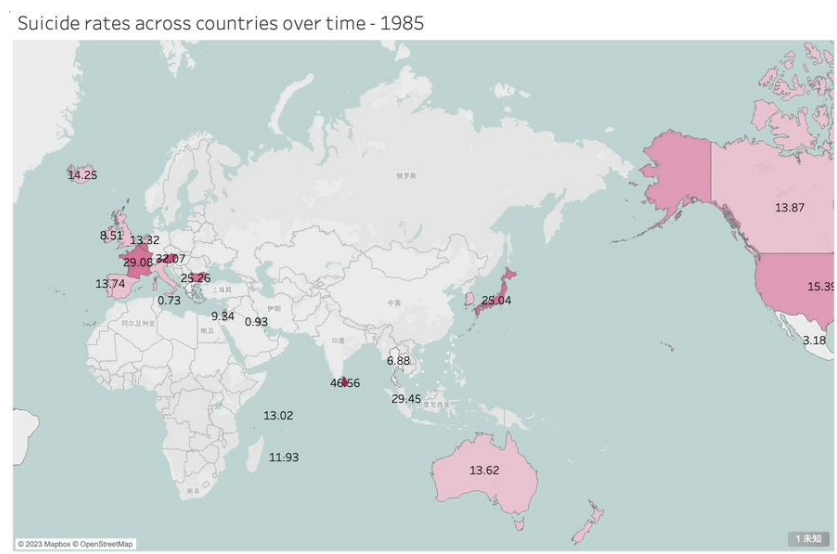
R: R is a popular advanced programming language for statistical computing and graphics, offers powerful visualization capabilities. In this study, R will help to gain insights into the patterns of suicide rates across age groups and uncover the interdependencies between various factors. Thus, two specific visualization graphs will be employed using R.

- A streamgraph can be created to visualize the suicide rates in different age groups using the "Suicide Rates Overview 1985 to 2016" dataset. This streamgraph will provide an overview of how suicide rates have varied across different age groups over time.
- A correlation heatmap can be generated using the two datasets to explore the relationships between different variables influencing suicide rates and happiness. This heatmap will display the correlation coefficients between pairs of variables, allowing for the identification of significant correlations and potential drivers of suicide rates.

# Data Visualization and Analysis

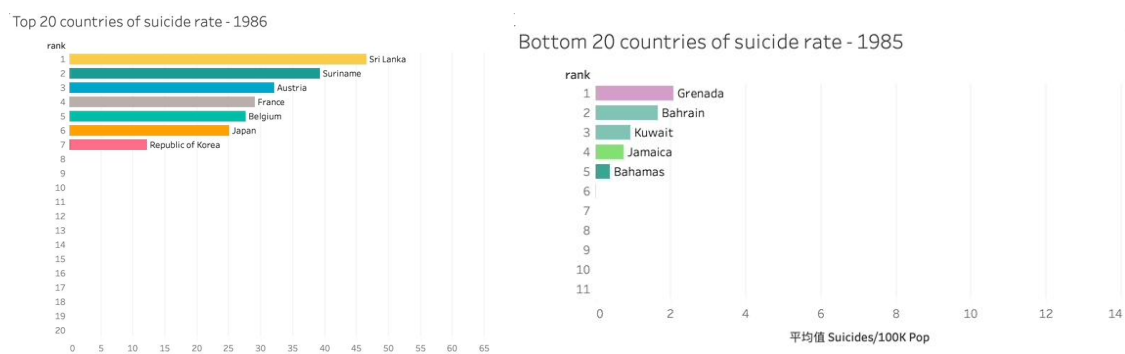
## Part I: Basic global situation of suicide

Graph 1: Suicide rates across countries over time



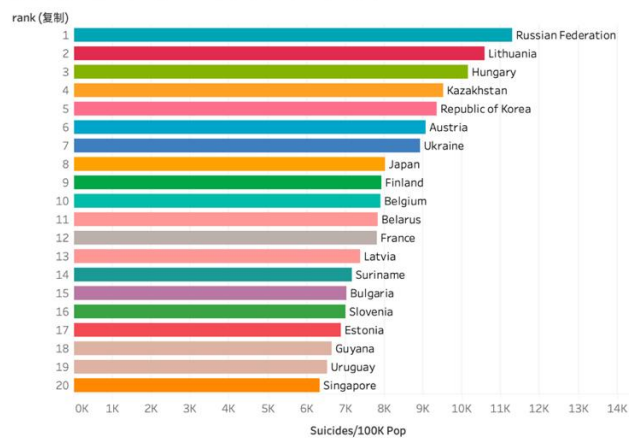
Graph 1 shows the dynamics of annual suicide rates from 1990 to 2020 in countries around the world. Due to the incomplete nature of the database, some areas without data will not be colored, and areas with data will be darkened as the suicide rate increases.

Graph 2&3: Top 20 and Bottom 20 countries of suicide rate (dynamic graph)



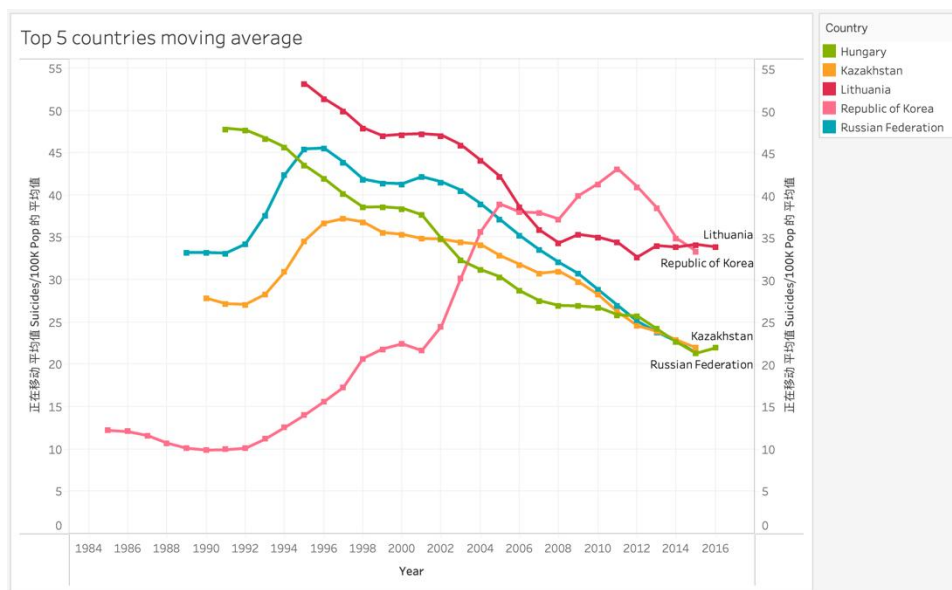
Graph 4: Top 20 Average Suicide Rate Countries Among 1985-2016

(total) Top 20 countries of suicide rate (2)



In addition to a dynamic bar chart of the rankings that moves with the year, A ranking of the 20 countries with the highest 30-year total suicide rates was also created using Tableau, providing a more macroscopic view of suicide rates in individual countries. In this regard, the countries with high suicide rates are mostly concentrated in the Soviet Middle East, followed by South Korea and Japan in East Asia.

Graph5: Top 5 countries moving average of suicide rate

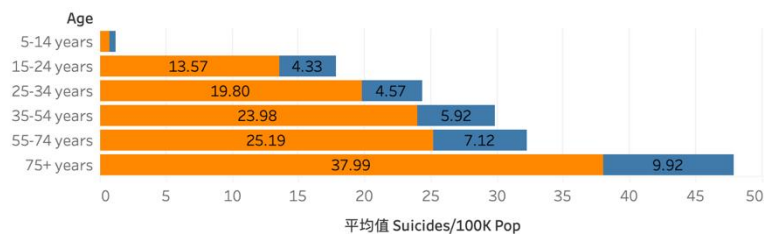


Based on graph 5, the top 5 countries were selected to visually illustrate the trend of suicide rates over a 30-year period. Except for South Korea, all of them are on a downward trend, and it is noteworthy that all three former Soviet Union countries saw an increase in suicide rates after the dissolution of the USSR and

began to see a gradual decline in suicide rates after peaking in 1997. Thus, perhaps political factors caused the suicide rates in these three countries to peak around the time of the collapse of the Soviet Union and to slow down afterward. However, Part III of the report will delve into a comprehensive analysis of the reasons behind Korea's significant increase in suicide rates and its status as the country with the highest suicide rate in the world.

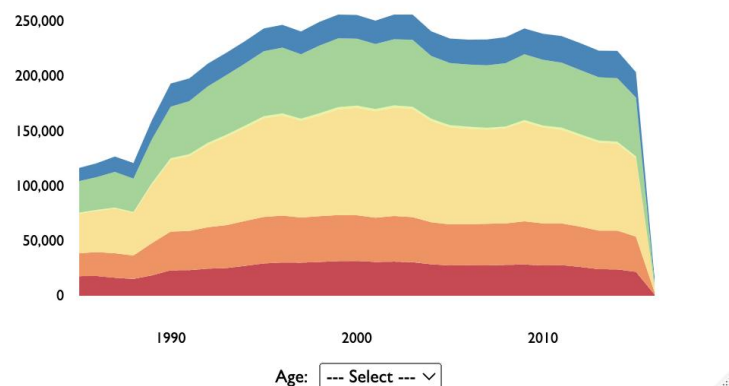
*Graph6: The relationship between suicide rate, age, and gender*

The relationship between suicide rate and gender



In order to show the suicide rate in relation to age and gender, the bar stacked graphs were created. The analysis revealed that the older the age, the higher the suicide rate, and that regardless of age, the male suicide rate is about three times higher than the female rate.

**Graph 7: Streamgraph of Suicide rate in different Age Group**



In addition, to show the trend of suicide rates more visually for the 6 age groups for all years, the R language ‘streamgraph\_plot’ is used to create a streamgraph, which clearly shows that the 75+ year old group contributes the most to the overall trend, i.e., has the highest suicide rate.

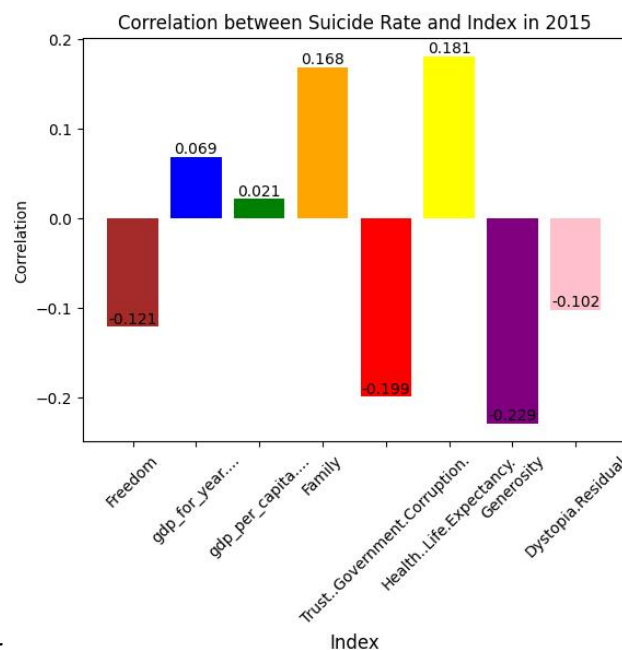
## Part II: Correlated Variables Visualization

Create Merged CSV: After visualizing the basic situation of global suicide rate, a merged CSV file (HappySuicide\_num.csv) is created to analyze the potential variables correlated with the suicide rate.

Besides, since a country has different suicide rates for different age groups and different genders, calculate the average suicide rate for each country using the ‘groupby’ function is also needed, and finally add the column of the average suicide rate to the merged csv.

Furthermore, all the strings in the HappySuicide\_num.csv table were converted into numeric form. For instance, "male" was replaced with 0, "female" with 1, and each age group was substituted with its median age. For example, "35-55 years old" was replaced by 45 years old. Subsequently, based on the modified table, the following visualization tasks were performed.

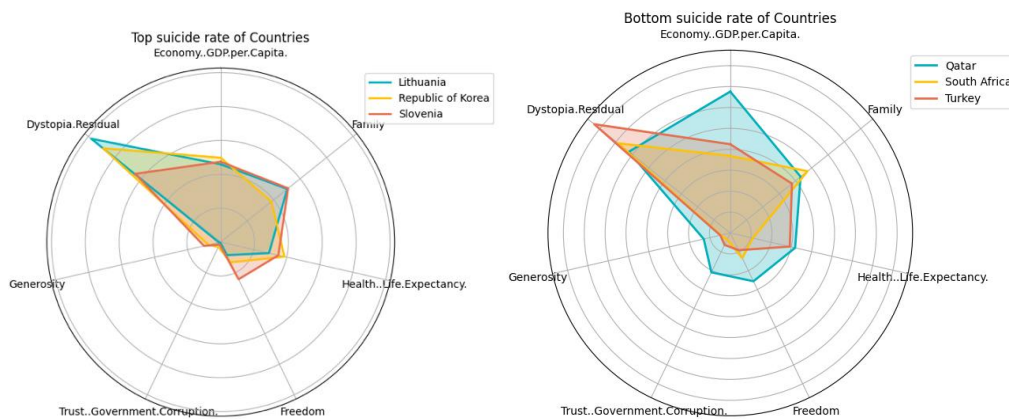
*Graph 7: Correlation between Suicide Rate and various Index in*





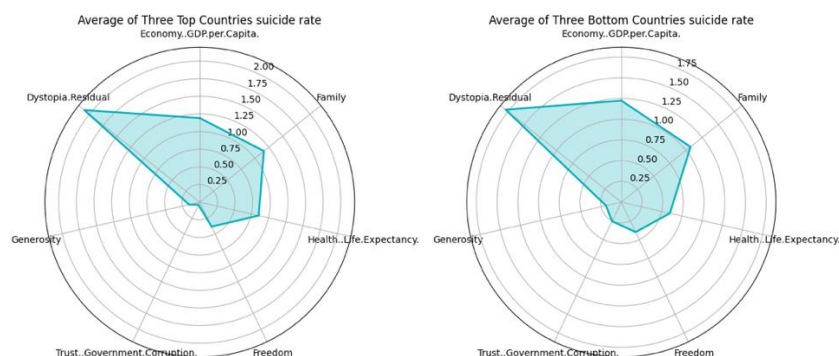
A total of eight indicators measuring the happiness of a country were analyzed. Using Python, the correlation between these indicators and the average suicide rate was visualized. The findings revealed that Freedom, Trust in Government, Generosity, and Dystopia residual exhibited a negative correlation with the average suicide rate, while the remaining indicators displayed a positive correlation.

*Graph 8&9: Radar Graph of happiness indicators for countries with top 3 and bottom 3 suicide rates*



For the purpose of studying the differences in happiness indicators between countries with the highest and lowest suicide rates, radar charts were employed. Based on the bar chart generated by Tableau in Part I, three countries with the highest suicide rates and three countries with the lowest suicide rates in 2015 were identified. To facilitate the analysis, two CSV files were created in R: Top3\_rate.csv (including Lithuania, Republic of Korea, and Slovenia) and Bottom3\_rate.csv (including Qatar, South Africa, and Turkey). Subsequently, radar plots were generated using Python for each category.

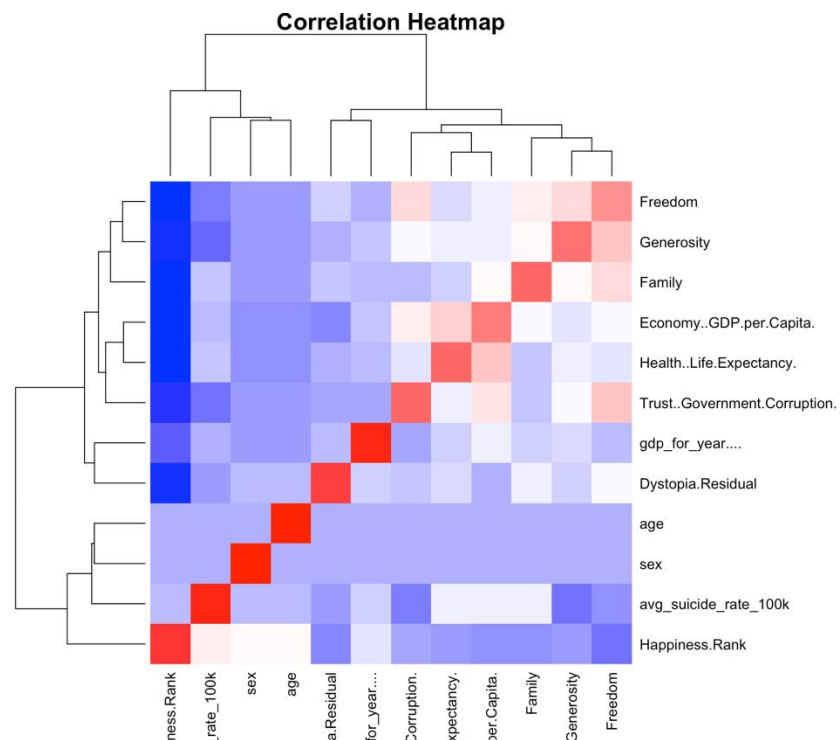
*Graph 10& 11: Radar Graph of averages of happiness indicators for Top 8 and Bottom 8 Suicide Rates*



In order to provide more comprehensive insights, two additional CSV files were created in R:

Top10\_rate.csv and Bottom8\_rate.csv. These files were used to compare and analyze the average index of each variable between countries with the lowest suicide rates and those with the highest suicide rates. The analysis revealed that Trust in Government and Family are particularly noteworthy variables that warrant attention.

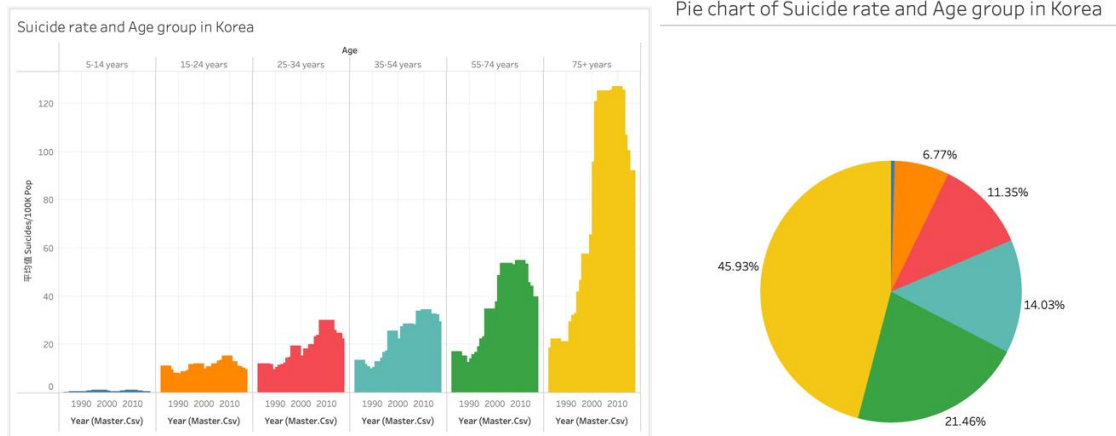
*Graph 12: Correlation Heatmap*



Graph 12 shows the positive and negative correlation between all the variables, where blue color indicates negative correlation and red color indicates positive correlation. Besides focusing on the variables that affect the suicide rate, the correlation between the variables is also of great interest, which helps in understanding the instrumental variables that contribute to the suicide rate.

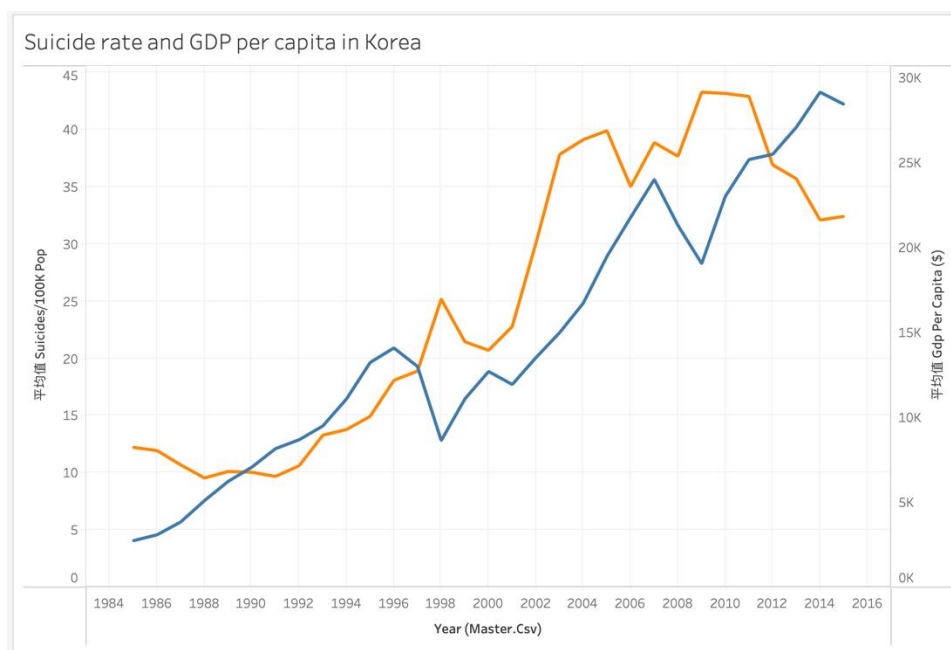
### Part III: Visualization of Korea

*Graph 13&14: The Age distribution of Korea Suicide Rate*



To analyze Korea's dramatically rising suicide rate, the analysis has been conducted across multiple dimensions. The first was to use a Histogram in Tableau to visualize the Age distribution of Korea Suicide Rate. While for the globe, the suicide rate is the highest for people over 75+ years of age, this phenomenon is particularly acute in Korea.

*Graph 15: The Line Chart of Korea Suicide Rate and GDP Per Capita*

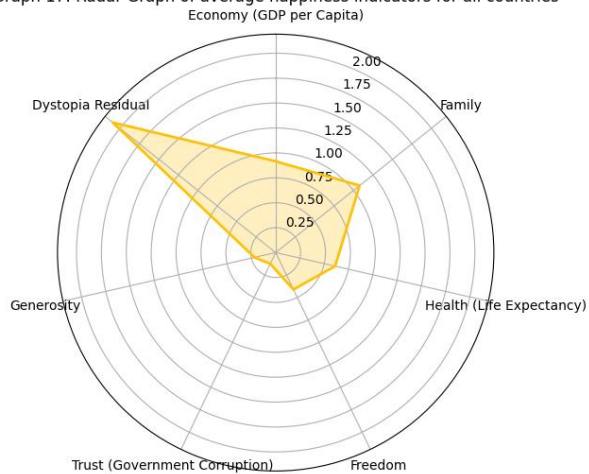


From Graph 15, the suicide rate in Korea is climbing along with the GDP per capita, for which it can be hypothesized that the fast-growing economic situation may have brought some life and psychological pressure to the Korean people.

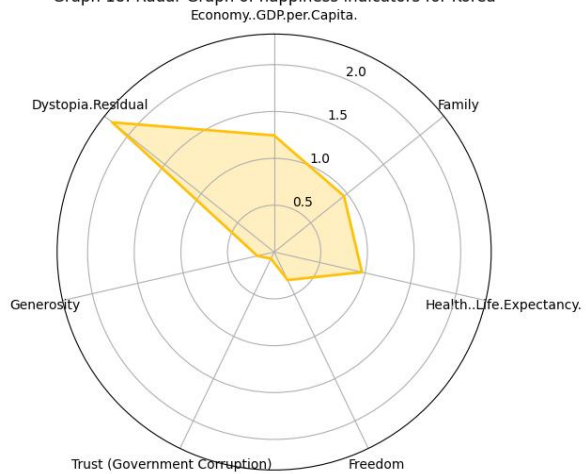
Graph 16: Radar Graph of average happiness indicators for all countries

Graph 17: Radar Graph of happiness indicators for Korea

Graph 17: Radar Graph of average happiness indicators for all countries



Graph 18: Radar Graph of happiness indicators for Korea



Since the larger the amount of data, the smaller the error. Thus, for graph 16, a merged data CSV (2015-2019.csv) of the World Happiness Index for a total of 6 years from 2015-2019 is created. To see if Korea is lacking in political, economic, social and cultural aspects compared to the world average. The figure finds that Korea's trust in government and family harmony are relatively below average.

# Research Implications

## Part I: Basic Situation Analysis

Gender: men have a higher suicide rate than women at all ages. This phenomenon may be related to the fact that men are less likely to have help-seeking intentions<sup>2</sup> for psychological problems, under more economic pressure (higher societal expectation) than women, and access to emotional support is not as high as it is for women. In workplace, men-centric support services and implement workplace mental health programs that cater to the specific needs of male employees can be created. In society, there is a need to revisit the issue of gender parity so that men and women have equal access to emotional support and emotional outlets.

Age: age is positively correlated with suicide rates. According to graph 7, state welfare for the elderly affects the suicide rate. Thus, young people should care more about the elderly. The state should increase welfare for the elderly, and society needs to increase the perceived usefulness of the elderly and provide more social support to mitigate the cumulative risk of the elderly.

Country & Politic: Generally speaking, over time, the suicide rate in the former Soviet Union countries is getting lower and lower. However, the suicide rate in some East Asian countries such as South Korea and Japan, and some South American countries is getting higher and higher. And from graph 5, after experiencing the great changes caused by the disintegration of the Soviet Union, the suicide rate in the former Soviet Union countries has gradually declined. This means that politics will have a greater impact on the suicide rate. In this regard, it is very important to improve national stability and social happiness.

## Part II: Analysis of Relevant Variables with Suicide Rate

---

<sup>2</sup> Reynders, A., Kerkhof, A. J. F. M., Molenberghs, G., & Audenhove, C. (2014). Attitudes and stigma in relation to help-seeking intentions for psychological problems in low and high suicide rate regions. *Social Psychiatry and Psychiatric Epidemiology*, 49(2), 231-239. DOI: 10.1007/s00127-013-0736-1

Economic Dimensions: From graph 8, a strong correlation between the suicide rate and the overall GDP of many countries. This suggests that the level of prosperity in a country has a certain mitigating effect on the suicide rate. However, excessive focus on economic growth can have unintended consequences, such as overlooking other aspects of political, cultural, and social development. In light of this, while continuously stimulating economic growth, it is important to maintain the government's authority and enhance public trust in the government.

Political Dimensions: from graph 13, suicide rate have negative correlation with freedom and trust to government.

Cultural Dimensions: Graph 7 showed that suicide rate has a negative correlation with Dystopia Residual and Generosity. Dystopia Residual refers to the degree to which a country deviates from its least happy state (i.e. dystopia). And A higher generosity score indicates that the residents of a country are more inclined to help and support others, demonstrating more acts of kindness and generosity.

### **Part III: Situation analysis of Korea**

The suicide rate in Korea is climbing along with the GDP per capita, for which it can be hypothesized that the fast-growing economic situation may have brought some life and psychological pressure to the Korean people. Besides, the suicide rate among the elderly (over 75+) in South Korea is quite high, mostly 59%. Moreover, Korea's trust in government and family harmony are relatively below average.

Based on these, there are following implications:

Society: it is crucial to prioritize the well-being and support of elderly people, considering the high suicide rates within this age group. Policies and programs should be implemented to address their specific needs and provide them with adequate care and support systems. For example, establishing community-based senior centers that offer social activities, healthcare services, and mental health support can help combat social isolation and promote the well-being of the elderly.

Politic: It is important to address the potential for corruption and to increase people's trust in government and the level of governmental authority so as to avoid a Tahitian effect. Measures should be taken to promote transparency, accountability and good governance, which can help to improve the overall well-being of the population and potentially reduce suicide rates. This could include enforcing anti-corruption laws and regulations, strengthening oversight mechanisms and promoting a culture of transparency and ethical behavior within government institutions.

Economic: Attention should be paid to considering the potential negative impact on mental health behind rapid economic growth. While economic development is desirable, it is equally important to ensure that the rate of economic growth is sustainable and does not place an undue burden on individuals or lead to social inequality. Measures such as promoting work-life balance through the implementation of flexible working hours and encouraging employers to provide mental health support services for their employees can help mitigate the negative impact of economic stress on mental health. In addition, creating a supportive and inclusive work environment that values the well-being of employees and promotes a healthy work-life balance can also help reduce stress and improve mental health.

## **Future Extension**

Simply correlating the mental indicator of happiness with the suicide rate is only a small aspect. Therefore, in future studies, it would be valuable to explore additional mental and physical indicators in relation to the suicide rate, beyond just happiness. Given the significance of the Internet era, it would be particularly important to investigate the correlation between cyberbullying and suicide rates. To conduct such an analysis, the JiSoKe (GooSeeker) tool can be utilized, which is a powerful tool for conducting opinion analyses of the Internet. Additionally, Python can be employed to create WordCloud visualizations to analyze emotional indices.

By using the JiSoKe tool, researchers can gather data from online platforms and social media to examine the prevalence and impact of cyberbullying on mental health and suicide rates. This tool can help identify

patterns, sentiments, and trends related to cyberbullying incidents and their association with suicidal behaviors. By analyzing the emotional content of online discussions, researchers can gain insights into the psychological impact of cyberbullying and its potential contribution to the suicide rate.

Furthermore, Python can be employed to create WordCloud visualizations, which can provide a visual representation of the most frequently used words or phrases related to cyberbullying and suicide. These WordClouds can help identify key themes and emotions associated with these topics, allowing researchers to better understand the underlying factors contributing to the suicide rate.

By combining the use of the JiSoKe tool for opinion analysis and Python for WordCloud visualization, future research can gain a comprehensive understanding of the relationship between cyberbullying and suicide rates. This analysis can provide valuable insights for the development of preventive measures, intervention strategies, and support systems to address the detrimental effects of cyberbullying on mental health and reduce the incidence of suicide.



## References

1. United Nations Development Program. (2018). Human development index (HDI).
2. World Bank. (2018). World development indicators: GDP (current US\$) by country:1985 to 2016.
3. [Szamil]. (2017). Suicide in the Twenty-First Century [dataset].
4. World Health Organization. (2018). Suicide prevention.
5. Helliwell, J. F., Layard, R., & Sachs, J. D. (2015). World Happiness Report 2015 (Report No. ISBN 978-0-9968513-2-9). Sustainable Development Solutions Network. Retrieved from <https://worldhappiness.report/ed/2015/>