**Analyzing Global Suicide Trends (1995 to 2014)**

by

**Arvin Rene P. Caabay**

Submitted to

**Coding School / Diretcho Trabaho (2024)**

**as Final Capstone Project**

**April 2024**

**Background**

Suicide is a complex phenomenon with profound implications worldwide, especially to those who are directly involved. Understanding the trends of suicide rates is crucial for developing effective prevention strategies and interventions.

The examination of global suicide trends provides valuable insights into the prevalence and distribution of suicide rates across different countries and time periods. From socioeconomic disparities to cultural norms and mental health policies, a multitude of factors can influence suicide rates and patterns, making it imperative to conduct comprehensive analyses.

The years 1985 to 2016 represent a significant timeframe for studying global suicide trends. This period encompasses substantial societal changes, economic fluctuations, and advancements in mental health research and interventions. Understanding how suicide rates have evolved during these years can offer critical insights into the impact of various social, economic, and healthcare factors on suicidal behavior.

Analyzing data on suicide rates is crucial for informing suicide prevention efforts, policy development, and mental health advocacy initiatives. By unraveling the dynamics underlying suicide trends, we strive to contribute to the collective effort to reduce the burden of suicide and promote mental well-being worldwide.

**Objectives**

* Explore the correlation between GDP per capita and suicide rates to ascertain whether nations with higher GDP per capita exhibit lower suicide rates;
* Assess the impact of age and gender on suicide rates across diverse countries;
* Assess the impact of gender on suicide rates across diverse countries.

**Methodology:**

1. **Data Loading and Inspection:**

* Import necessary libraries: pandas, numpy, matplotlib, seaborn, pycountry, Basemap, and linregress.
* Load the dataset using pd.read\_csv().
* Check the shape of the dataset to ensure it meets the project requirements.
* Inspect the first few rows of the dataset using data.head().

1. **Items Formatting / Cleaning:**

* Improve column titles for better readability and clarity.
* Remove the "HDI for year" column due to mostly incomplete data.
* Remove the "country-year" column as it is redundant.
* Change the data type of the "gdp\_for\_year ($)" column from object to numeric.
* Correct invalid country names using the pycountry library.
* Removed Year 1985-1994 and 2015-2016 due to majority of data are missing or not available
* Removed Countries with incomplete data after checking with original Source (WHO) that data is unavailable, (morality clause in making up number of deaths)

1. **Exploratory Data Analysis (EDA):**

* Explore the distribution of suicide rates across different countries, years, sexes, and age groups.

1. **Data Visualization:**

* Visualize correlations between numerical columns using a heatmap.
* Plot a map chart showing average suicide rates per country.
* Perform linear regression analysis between GDP per capita and suicide rates.
* Visualize the relationship between GDP per capita and suicide rates using a scatter plot.
* Standardize the data of GDP per capita and suicide rates and visualize the correlation matrix.

1. **Correlation Analysis:**

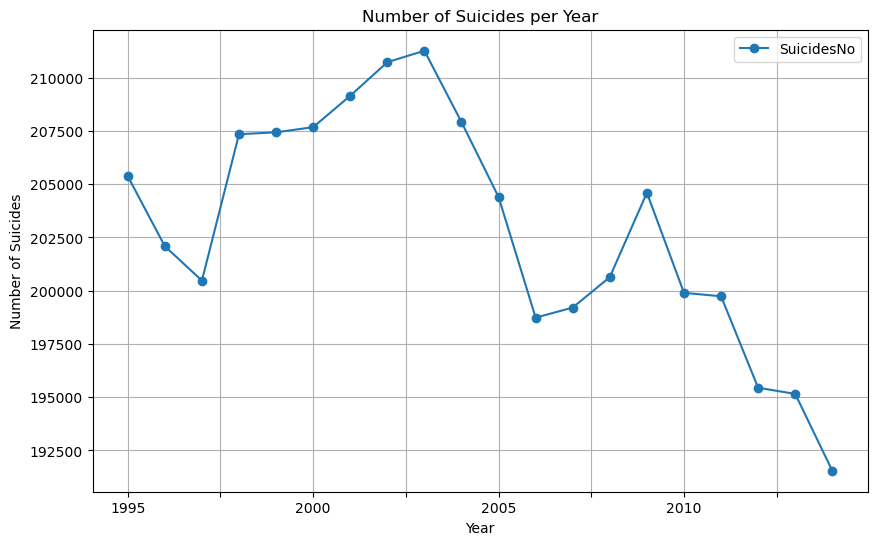
* Calculate correlation coefficients between GDP per capita and suicide rates per country.
* Investigate correlations between GDP per capita and suicide rates across different age groups.

1. **Gender Comparison:**

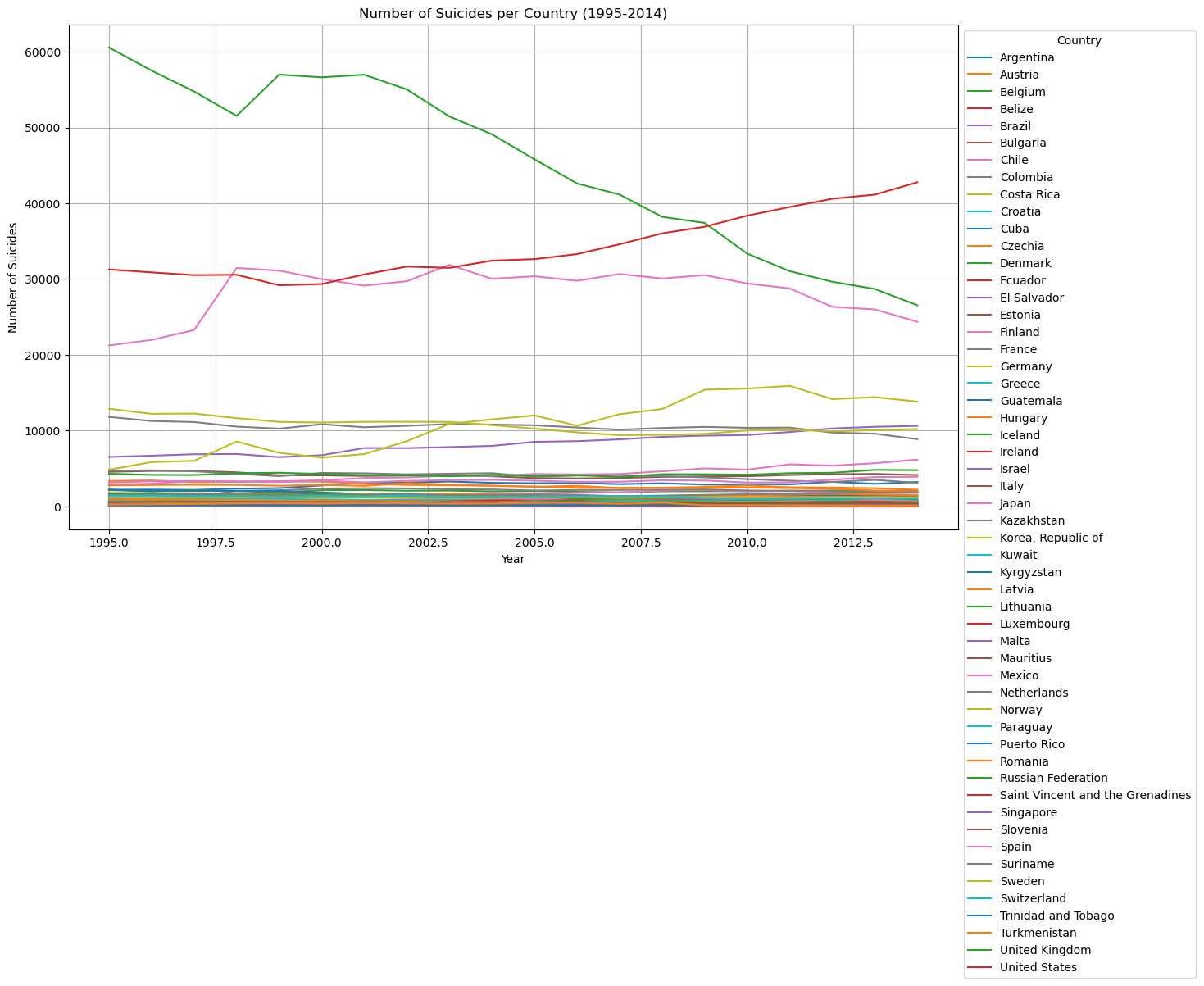
* Compare suicide rates between genders using a T-test.
* Visualize the distribution of suicide rates by gender using histograms.
* Compare mean suicide rates between genders using a bar plot.

**Data Visualization:**

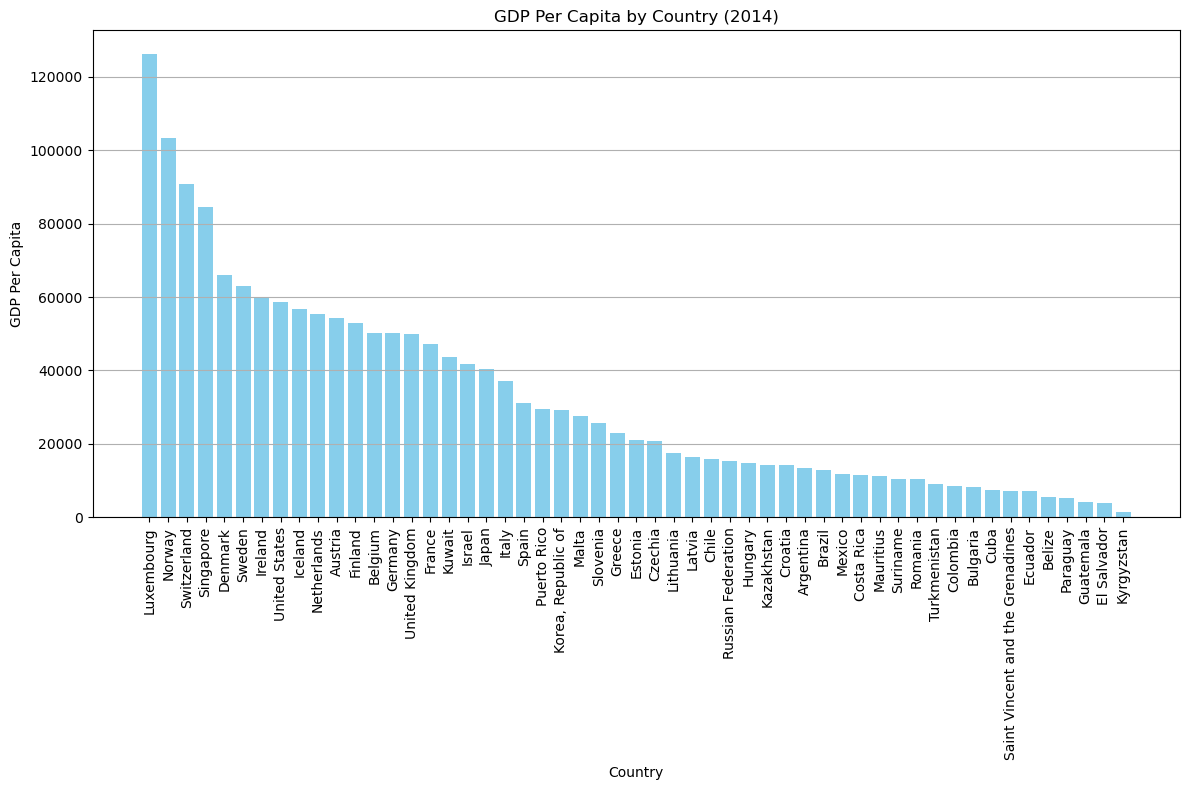
1. **Line Graph of Total Suicides (World) from 1995 to 2014**



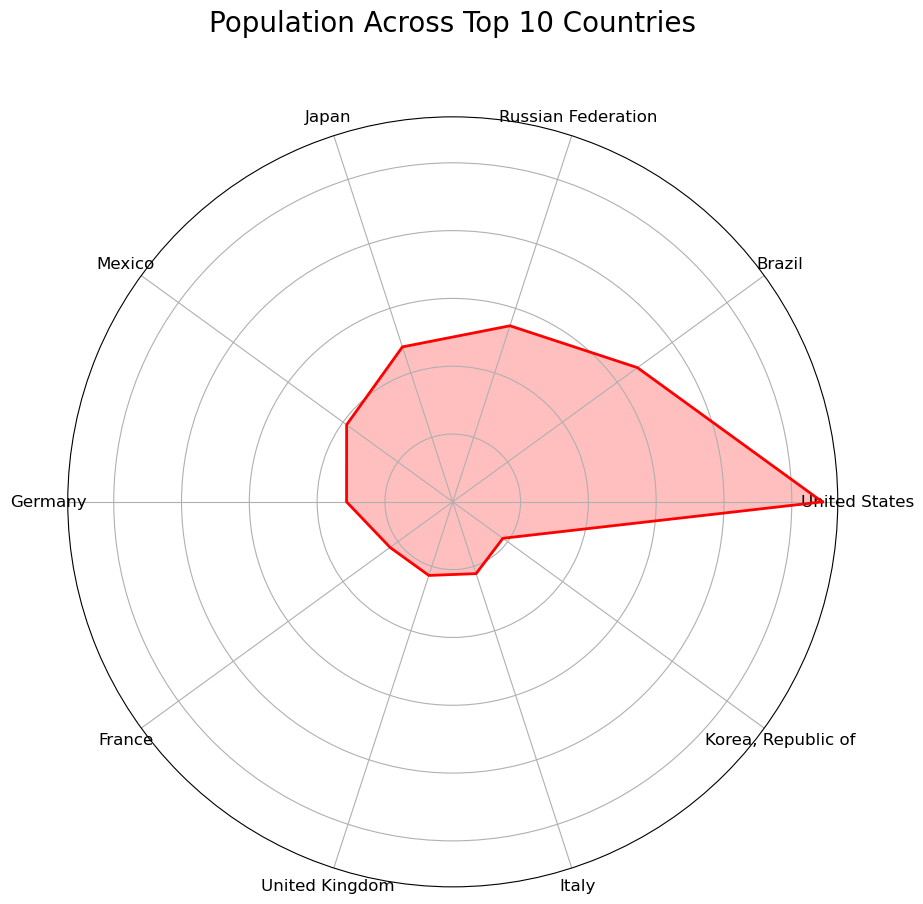
1. Line Graph of No. of Suicides per country



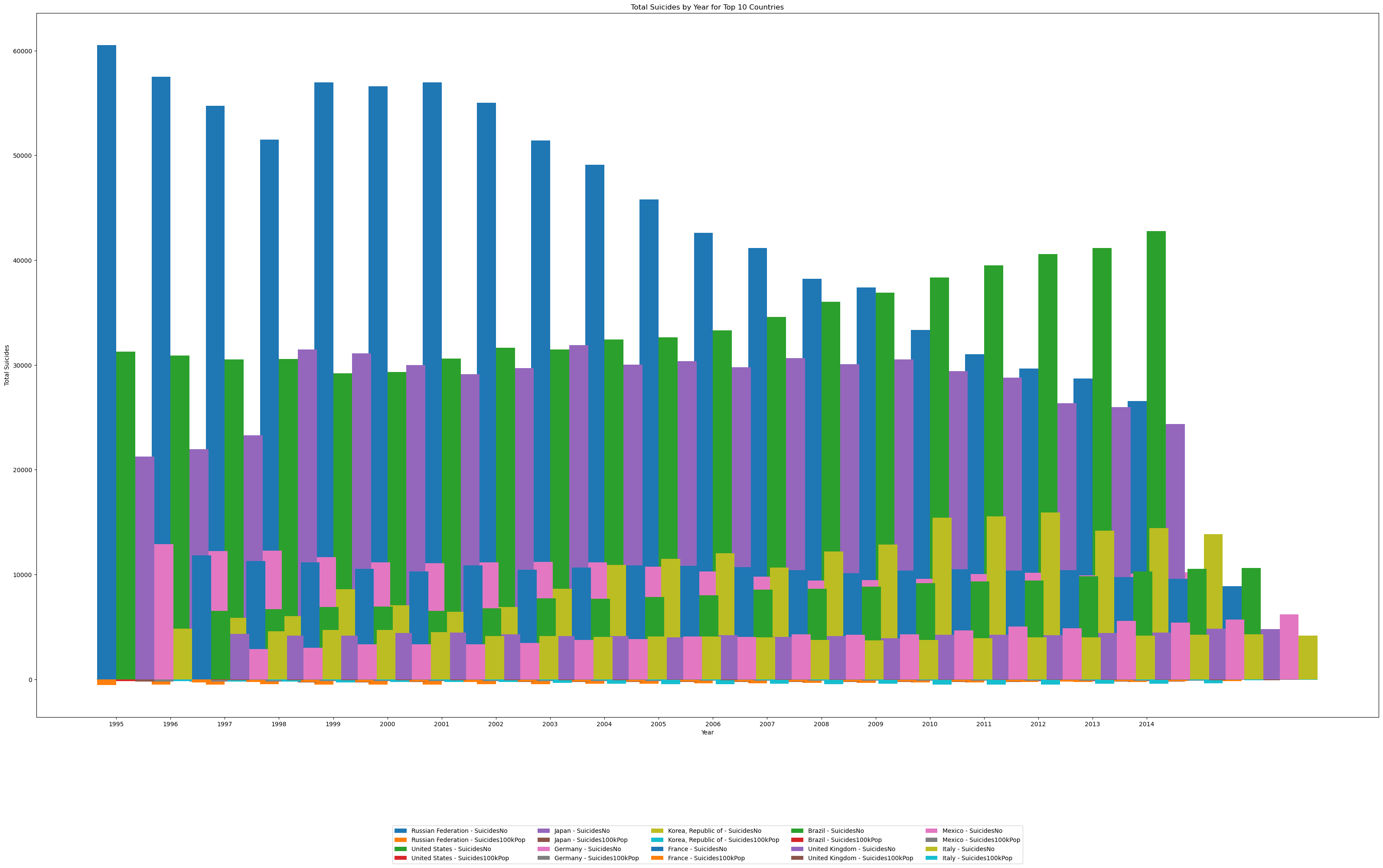
1. **Bar Graph of GDP (2014) per country**



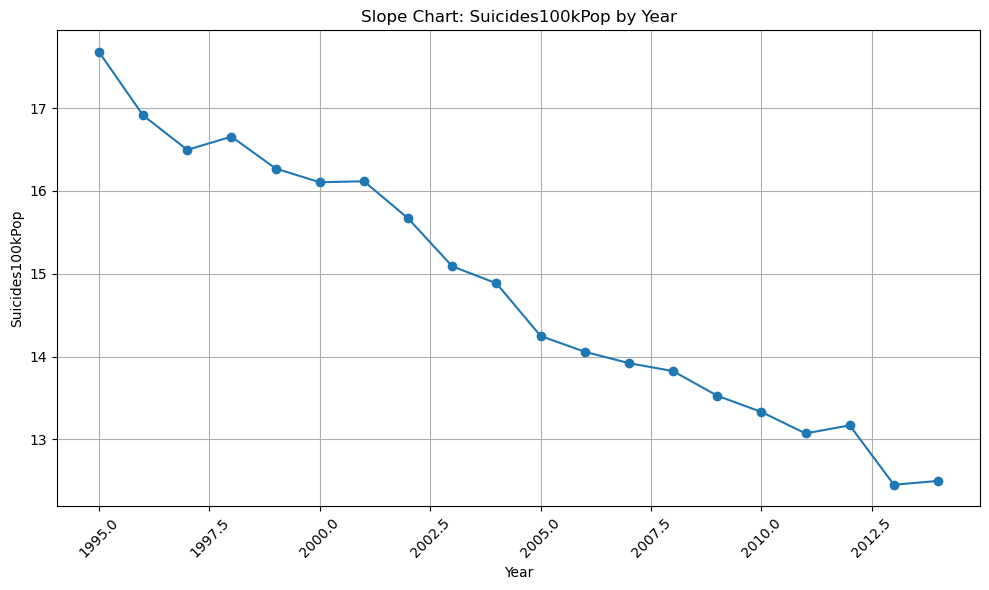
1. **Radial Graph of Population**



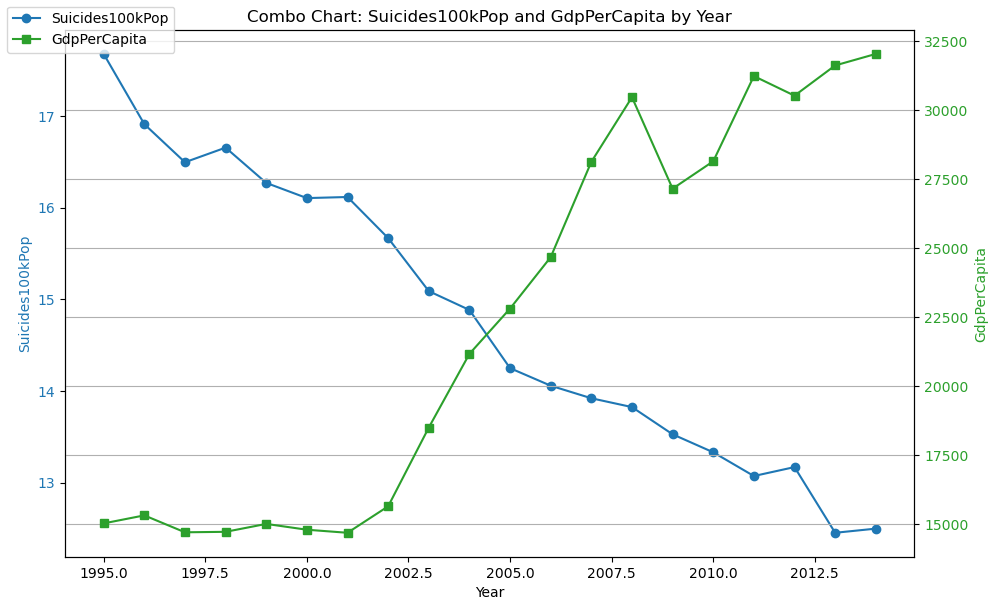
1. **Butterfly Graph: Total No. of Suicides vs Suicide per 100k Population**



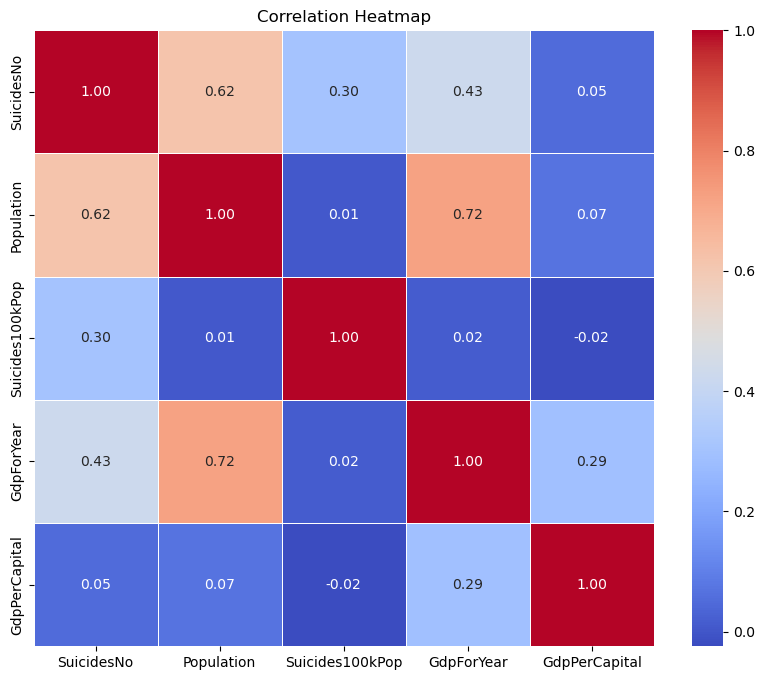
1. **Slope Chart Suicide per 100k population over the inclusive years**



1. **Combo Chart**

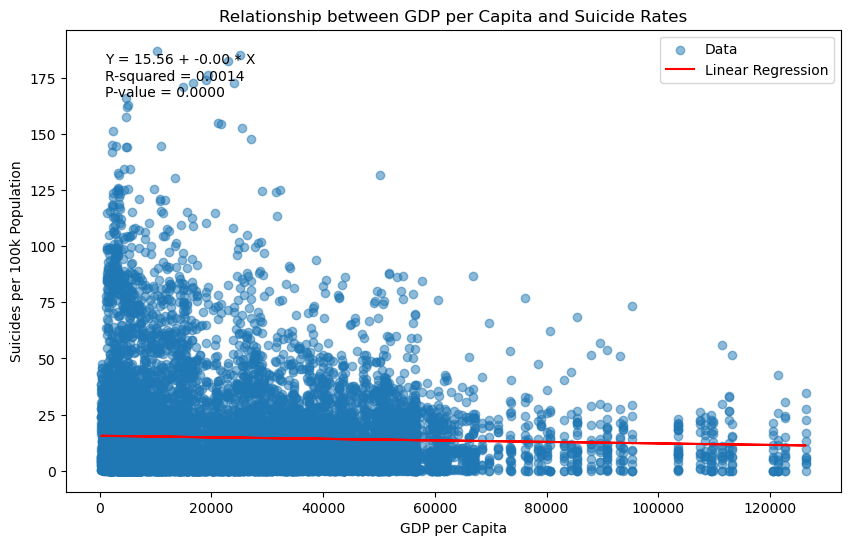


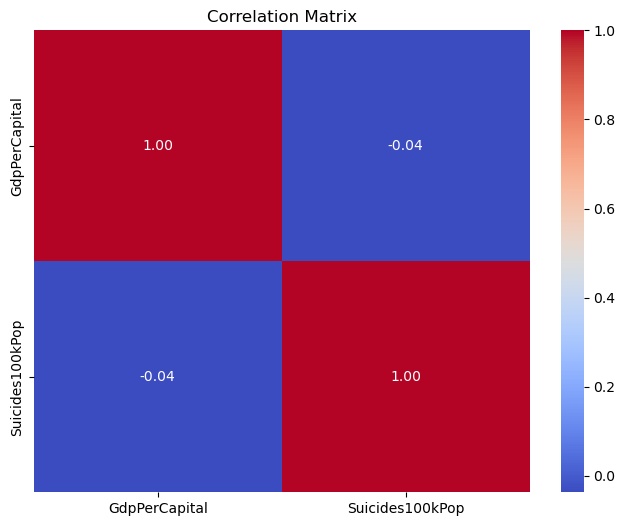
1. **Heat Map**



**Insights**

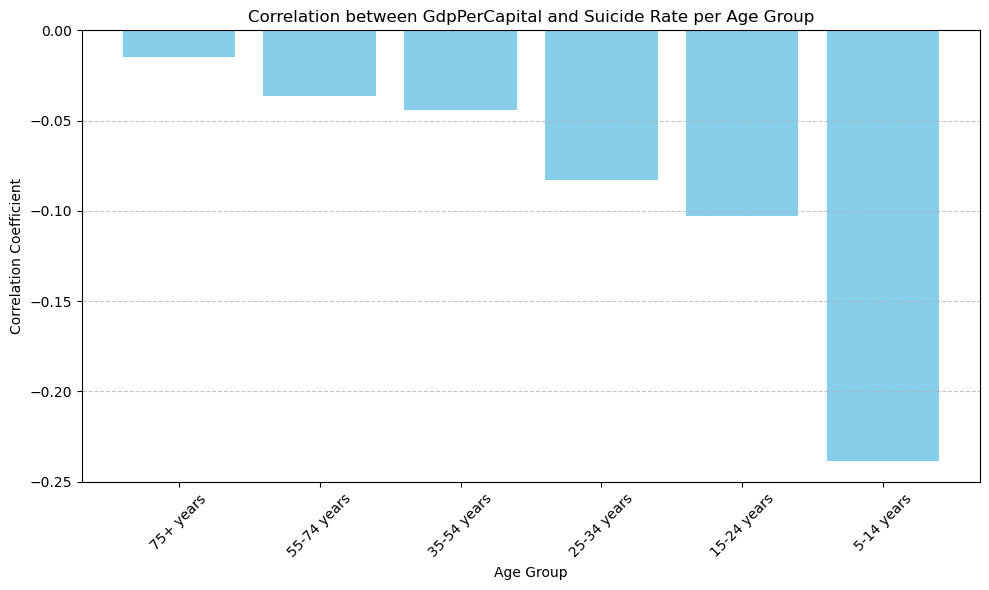
1. **GDP per capita vs. Suicide Rate per 100k Population:**





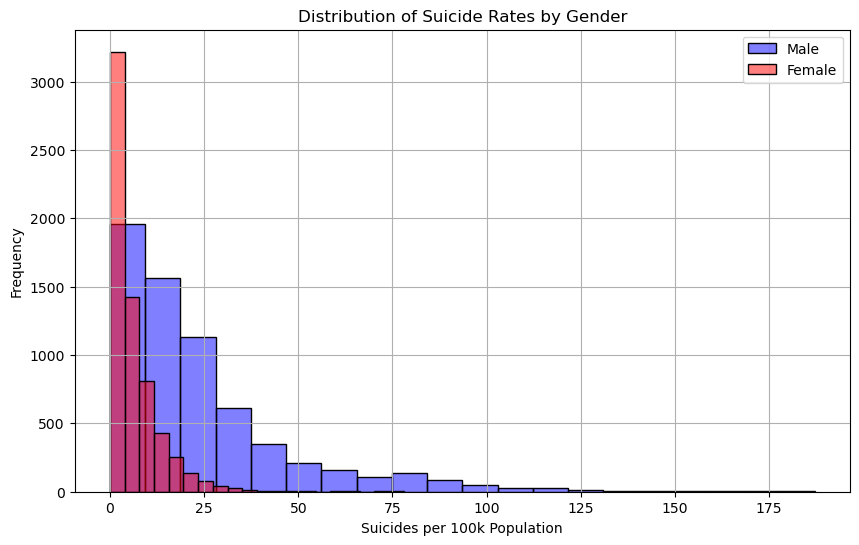
* The linear regression analysis indicates a weak negative relationship between GDP per capita and suicide rate per 100k population, with a negative slope and low R-squared value.
* The correlation coefficient between GDP per capita and suicide rate per 100k population is also very low (-0.036751), suggesting a weak inverse relationship.

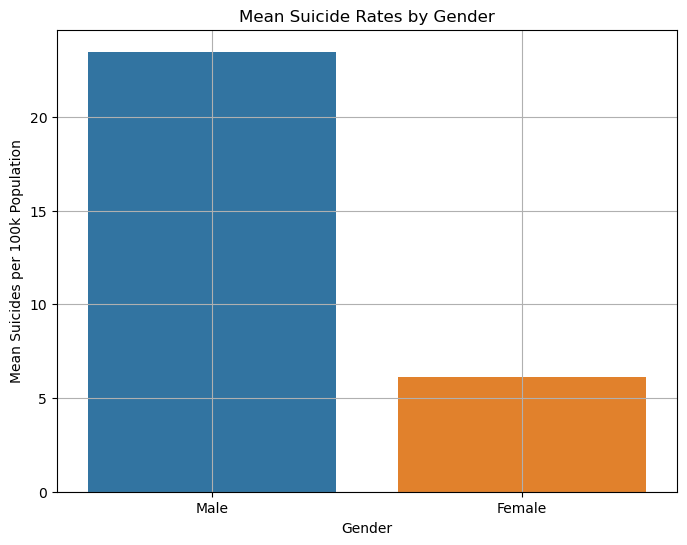
1. **GDP per Capital per Age Group**



* There is a negative correlation between GDP per capita and suicide rate across different age groups, with stronger negative correlations observed among younger age groups. This suggests that nations with higher GDP per capita tend to have lower suicide rates across all age groups, with the strongest effect observed in the 5-14 years age group.
* But in its entirety, the correlation is irrelevant.

1. **Suicide Rate Based on Gender**





* The analysis shows that males have a significantly higher suicide rate compared to females, as indicated by the total suicides per 100k population.
* The T-test results confirm that the difference in suicide rates between genders is statistically significant, with a very low p-value.

**Recommendations:**

* While the analysis indicates a weak relationship between GDP per capita and suicide rates, it's essential to address underlying socioeconomic factors that may contribute to suicide risk. Policies aimed at improving economic opportunities, reducing income inequality, and providing social support systems could help mitigate suicide risk.
* Given the stronger negative correlation between GDP per capita and suicide rates among younger age groups, targeted interventions focusing on mental health support, education, and social inclusion for adolescents and young adults may be beneficial.
* Develop gender-specific suicide prevention strategies that address the unique risk factors and challenges faced by males and females. This could involve promoting mental health awareness, improving access to support services, and challenging societal norms related to masculinity and help-seeking behavior.
* Continued Research and Monitoring: Further research is needed to better understand the complex interplay between socioeconomic factors, mental health, and suicide risk. Continual monitoring of suicide trends and risk factors can inform evidence-based interventions and policy decisions aimed at reducing suicide rates.

**Reflections:**

* **Challenges Encountered**
* One of the primary challenges was dealing with incomplete or missing data. This required careful consideration when cleaning the dataset and making decisions about which data points to include in the analysis.
* Understanding the limitations of correlation analysis was challenging. While correlations provide insights into relationships between variables, they do not imply causation.
* **Lessons Learned**
* Data Preparation is always Key!
* Visualizing data using various techniques such as heatmaps, scatter plots, and histograms proved invaluable in uncovering patterns and relationships within the data.
* **Opportunities for Future Research:**
* Update the Paper once complete data is available
* Longitudinal Analysis
* Recommendations for Policy Evaluation

**Conclusions:**

* The statistical analysis revealed a weak negative relationship between GDP per capita and suicide rates. While the correlation was low and the R-squared value indicated limited explanatory power, the negative slope suggests that nations with higher GDP per capita tend to have slightly lower suicide rates. However, this relationship is not strong enough to draw definitive conclusions about causation.
* Across different age groups, there was a negative correlation between GDP per capita and suicide rates, with stronger negative correlations observed among younger age groups. This indicates that nations with higher GDP per capita may have lower suicide rates across all age groups, particularly among adolescents and young adults.
* Gender-based analysis revealed significant disparities in suicide rates, with males exhibiting significantly higher rates compared to females. This finding underscores the importance of gender-specific approaches in suicide prevention efforts and highlights the need for targeted interventions addressing the unique risk factors faced by males.

**References:**

<https://www.kaggle.com/datasets/russellyates88/suicide-rates-overview-1985-to-2016>

<https://www.kaggle.com/code/szamil/suicide-in-the-twenty-first-century/notebook>

<https://databank.worldbank.org/source/world-development-indicators>

<https://hdr.undp.org/data-center/country-insights#/ranks>

<https://www.who.int/campaigns/world-suicide-prevention-day/2022>

<https://crisis-clinic.org/suicide/>

<https://healthygrandcounty.org/resource-directory/2122/suicide-prevention-resources/>

<https://platform.who.int/mortality/themes/theme-details/topics/indicator-groups/indicator-group-details/MDB/self-inflicted-injuries>