Abstract

Suicide remains a significant public health concern globally, with multifaceted determinants contributing to its prevalence. Among these determinants, socioeconomic factors play a crucial role, exerting both direct and indirect influences on individual mental health and well-being. This abstract presents a comprehensive analysis of the socioeconomic impact on suicide rates, drawing insights from various scholarly studies, statistical data, and theoretical frameworks.

The relationship between socioeconomic status (SES) and suicide rates exhibits complex dynamics, influenced by factors such as income inequality, employment stability, education level, access to healthcare, and social support networks. Lower SES individuals often face heightened risk factors for suicide, including financial strain, social exclusion, and limited access to mental health services. Moreover, economic downturns and recessions have been associated with increased suicide rates, reflecting the strain of financial hardship on mental health.

Furthermore, disparities in suicide rates across demographic groups highlight the intersectionality of socioeconomic factors with other identity markers such as race, gender, and age. Vulnerable populations, including minorities, LGBTQ+ individuals, and marginalized communities, often experience compounded socioeconomic challenges, exacerbating their risk of suicide.

Policy implications emerge from this analysis, emphasizing the importance of addressing socioeconomic inequalities to mitigate suicide risk effectively. Interventions aimed at improving economic stability, enhancing mental health resources, and fostering inclusive social policies can contribute to suicide prevention efforts. Additionally, community-based initiatives focusing on education, stigma reduction, and support networks are essential for promoting resilience and well-being among at-risk populations.

Index

Content	Page No.

CHAPTER 1

Introduction

1.1 Motivation

The motivation behind exploring the socioeconomic impact on suicide rates stems from the imperative need to address a pressing public health issue with far-reaching consequences. Suicide remains a significant global concern, claiming millions of lives annually and inflicting profound emotional and economic burdens on individuals, families, and communities. While suicide is a complex phenomenon influenced by various factors, including biological, psychological, and social determinants, socioeconomic status (SES) emerges as a critical yet often overlooked determinant with profound implications for suicide risk and prevention.

The motivation for delving into the socioeconomic impact on suicide rates is multifaceted. Firstly, socioeconomic disparities persist as a pervasive societal issue, with profound implications for health outcomes, including mental health and suicide risk. Individuals from lower socioeconomic backgrounds often face a myriad of challenges, including financial strain, limited access to quality healthcare and mental health services, social isolation, and discrimination, all of which contribute to heightened vulnerability to suicide. By elucidating the intricate interplay between socioeconomic factors and suicide rates, this research seeks to shed light on the root causes of suicide and inform targeted interventions to alleviate disparities and promote mental health equity.

1.2 Relevance

The relevance of socioeconomic factors in comprehending and addressing suicide rates cannot be overstated in contemporary public health discourse. Suicide remains a pervasive and complex issue with profound social, economic, and health implications, necessitating a comprehensive understanding of its determinants for effective prevention strategies. Socioeconomic status (SES) emerges as a crucial lens through which to examine the multifaceted nature of suicide risk, as it intersects with various social, economic, and structural factors to shape individuals' vulnerabilities. The relevance of investigating the socioeconomic impact on suicide rates lies in its capacity to elucidate patterns of inequality and inform targeted interventions to mitigate disparities. Research consistently demonstrates a strong association between lower SES and elevated suicide risk, with individuals experiencing poverty, unemployment, financial strain, and social marginalization facing heightened vulnerability.

1.1 Problem Statement and Objectives

Suicide rates continue to rise globally, constituting a major public health concern with devastating consequences for individuals, families, and communities. While numerous factors contribute to suicide risk, socioeconomic disparities emerge as critical determinants shaping individuals' vulnerability and access to resources. However, the intricate interplay between socioeconomic factors and suicide rates remains poorly understood, impeding the development of effective prevention strategies. Addressing this knowledge gap is imperative for advancing suicide prevention efforts and promoting mental health equity.

Objectives:

1. To examine the relationship between socioeconomic status (SES) and suicide rates through a comprehensive review of existing literature and empirical evidence.
2. To identify key socioeconomic determinants and mechanisms influencing suicide risk, including income inequality, unemployment, financial strain, access to mental health services, and social support networks.
3. To assess the intersectionality of socioeconomic factors with other demographic characteristics, such as race, gender, age, and sexuality, in shaping suicide risk disparities.
4. To propose evidence-based interventions and policy recommendations aimed at addressing socioeconomic inequalities, enhancing mental health resources, and fostering supportive social environments to mitigate suicide risk and promote resilience among vulnerable populations.

CHAPTER 2

Theoretical Background

2.1 Related Work

When developing a suicide rate predictor, it's essential to review existing literature and related work to understand the state-of-the-art methods, techniques, and challenges in the field. Here's an outline of potential related work for a suicide rate predictor:

Traditional Approaches:

- Studies investigating the socioeconomic impact on suicide rates employ various research methodologies, including epidemiological surveys, longitudinal studies, and qualitative research.
- Quantitative analyses often utilize statistical techniques to examine associations between socioeconomic indicators and suicide rates, while qualitative studies explore the lived experiences and narratives of individuals affected by suicide.

  .

Machine Learning Techniques:

- Regression models: Investigating various regression techniques such as linear regression, polynomial regression, support vector regression, decision trees, random forests, and neural networks for predicting concrete strength.

- Ensemble methods: Exploring ensemble learning techniques like bagging, boosting, and stacking to improve prediction accuracy.

- Deep learning: Reviewing recent advancements in deep learning models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformermodels for concrete strength prediction.

2.2 Summary of Theoretical Background

To develop a concrete strength predictor, you need a strong theoretical background in materials science and engineering, specifically related to concrete properties and their influencing factors. Here's a summary of the key theoretical aspects you should consider:

1. **Health Inequities**:
- Health inequities, defined as systematic disparities in health outcomes between different social groups, contribute to disparities in suicide rates.
- Marginalized and disadvantaged populations, including racial and ethnic minorities, LGBTQ+ individuals, and those experiencing homelessness or food insecurity, face higher rates of suicide due to the compounded effects of social and economic marginalization.

2. **Policy Implications**:
- Understanding the socioeconomic determinants of suicide is essential for developing effective prevention strategies and policies.
- Policies aimed at reducing income inequality, promoting access to quality education and healthcare, addressing unemployment, and strengthening social support systems can mitigate the risk factors associated with suicide.

3. **Research Methodologies**:
- Studies investigating the socioeconomic impact on suicide rates employ various research methodologies, including epidemiological surveys, longitudinal studies, and qualitative research.
- Quantitative analyses often utilize statistical techniques to examine associations between socioeconomic indicators and suicide rates, while qualitative studies explore the lived experiences and narratives of individuals affected by suicide.

4. **Global Perspectives**:
- Suicide rates and their socioeconomic determinants vary across different countries and regions.
- Factors such as cultural norms, healthcare systems, and economic policies influence the relationship between socioeconomic status and suicide.
- Comparative research across diverse sociocultural contexts can provide valuable insights into the complex interplay between socioeconomic factors and suicide rates.

5. **Intersectionality**:
- Intersectionality emphasizes the interconnected nature of social identities and systems of oppression.
- Intersectional approaches to studying suicide recognize that individuals may experience multiple forms of marginalization based on intersecting factors such as race, gender, sexuality, disability, and socioeconomic status.

CHAPTER 3

Experimental work

Data Preprocessing:

- You have selected specific columns ('Sex', 'SuicideCount', 'EmploymentPopulationRatio') as predictors andthe 'Strength' column as the target variable.

- Converted nominal variables to numeric using one-hot encoding (getdummies() function).

- Split the data into training and testing sets.

Model Training and Evaluation:

- Trained Random Forest and AdaBoost regressors on the training data.

- Evaluated model performance using mean absolute error (MAE), mean squared error(MSE), and R-squared (R2) score on the testing data.

- Plotted feature importance for the top 10 most important predictors in each model.

Insights:

- Both Random Forest and AdaBoost models were evaluated using MAE, MSE, and R2score. These metrics provide a comprehensive understanding of model performance.

- Feature importance plots help identify the most influential predictors in predicting concretestrength.

Potential Improvements:

- It's essential to explore other algorithms and tune hyperparameters to improve modelperformance further.

- Consider cross-validation for robust evaluation and hyperparameter tuning.

- Investigate additional feature engineering techniques to potentially enhance predictivepower.

- Explore data visualization techniques to gain deeper insights into the relationships betweenpredictors and the target variable.

Methodology used:

1. Data Loading and Preprocessing:

- The code reads a dataset from a CSV file named " suicide_rates_1990-2022.csv" & "age_std_suicide_rates_1990-2022.csv" using Pandas.

- Duplicate rows, if any, are removed from the dataset.

	count	mean	std	min	25%	50%	75%	max
Year	118560.0	2.005740e+03	8.745003e+00	1.990000e+03	1.999000e+03	2.006000e+03	2.013000e+03	2.022000e+03
SuicideCount	118096.0	6.363204e+01	2.540954e+02	0.000000e+00	0.000000e+00	4.000000e+00	3.200000e+01	6.787000e+03
CauseSpecificDeathPercentage	114271.0	3.873690e+00	7.317820e+00	0.000000e+00	0.000000e+00	5.861665e-01	4.355646e+00	1.000000e+02
DeathRatePer100K	107896.0	1.229979e+01	2.101617e+01	0.000000e+00	0.000000e+00	4.452366e+00	1.538462e+01	5.025126e+02
Population	112640.0	2.456184e+07	4.531156e+07	4.054200e+04	2.613805e+06	6.888434e+06	2.533364e+07	3.320316e+08
GDP	111320.0	4.883351e+11	1.653252e+12	2.197630e+08	1.162585e+10	6.268216e+10	3.013550e+11	2.331510e+13
GDPPerCapita	111320.0	1.742351e+04	1.933214e+04	6.023512e+01	3.941599e+03	9.708141e+03	2.519727e+04	1.337118e+05
GrossNationalIncome	108600.0	4.976795e+11	1.689591e+12	2.082832e+08	1.176795e+10	5.987715e+10	3.057210e+11	2.370530e+13
GNIPerCapita	107800.0	2.214722e+04	1.888640e+04	7.800000e+02	8.670000e+03	1.615000e+04	2.953000e+04	1.526300e+05
InflationRate	104100.0	2.028646e+01	1.630999e+02	-1.063010e+01	1.498961e+00	3.038888e+00	6.536199e+00	4.734914e+03
EmploymentPopulationRatio	107440.0	5.575081e+01	8.140994e+00	3.202600e+01	5.079100e+01	5.602100e+01	6.041525e+01	8.751800e+01

2. Exploratory Data Analysis (EDA):

- Descriptive statistics such as correlation matrix, histograms, and scatter plots are generated to understand the data.

- Correlation matrix helps in identifying the relationship between the target variable ('SuicideCount') and other continuous variables.

- Histograms and scatter plots visualize the distribution of variables and theirrelationship with the target variable.

3. Feature Selection:

- Based on correlation analysis, columns with an absolute correlation coefficientgreater than 0.3 with the target variable are selected as predictors.

- The selected columns are 'Sex', 'SuicideCount', and 'Age'.

4. Data Preprocessing for Machine Learning:

- Nominal variables are converted into numeric using one-hot encoding (pd.getdummies() function).

- The target variable ('SuicideCount') is added back to the DataFrame.

- The data is split into training and testing sets using train_test_split() from scikit-learn.

	count	unique	top	freq
RegionCode	118560	6	EU	45140
RegionName	118560	6	Europe	45140
CountryCode	118560	117	ARG	1740
CountryName	118560	117	Argentina	1740
Sex	118560	3	Male	58320
AgeGroup	118560	7	0-14 years	23712
Generation	118560	7	Generation Alpha	23712

5. **Model Training and Evaluation:**

- Two regression algorithms, RandomForestRegressor and AdaBoostRegressor, areutilized for model training.

- The models are trained on the training data and evaluated on the test data.

- Evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R2) score are calculated to assess model performance.

Percentage of missing column values
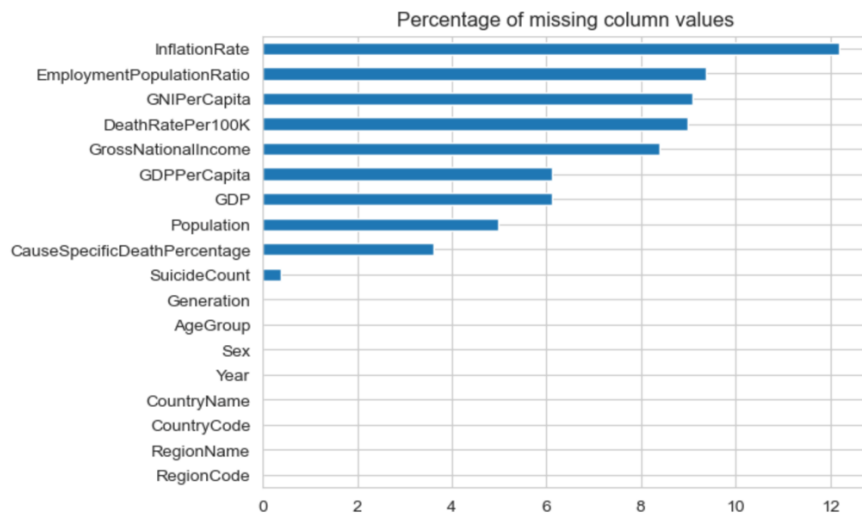
6. **Visualization:**

- Feature importance plots are generated for the top 10 most important predictors in each model, helping in understanding which features are most influential in predicting concrete strength.

Python libraries used:

1. **NumPy**: NumPy is a general-purpose array-processing package. It provides a high- performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python. It is open-source software.

2. **Pandas**: Pandas is an open-source library that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operationsfor manipulating numerical data and time series. This library is built on top of the NumPylibrary. Pandas is fast and it has high performance & productivity for users.

3. **Seaborn** : Seaborn is an amazing visualization library for statistical graphics plotting in Python. It provides beautiful default styles and color palettes to make statistical plots moreattractive. It is built on the top of matplotlib library and also closely integrated to the data structures from pandas.

4. **Matplotlib**: Matplotlib is an amazing visualization library in Python for 2D plots of arrays.Matplotlib is a multi-platform data visualization library built on NumPy arrays anddesigned to work with the broader SciPy stack.

5. **Geopandas**: GeoPandas is a project to add support for geographic data to pandas objects. It currently implements GeoSeries and GeoDataFrame types which are subclasses of pandas. Series and pandas, DataFrame respectively. GeoPandas objects can act on shapely geometry objects and perform geometric operations.
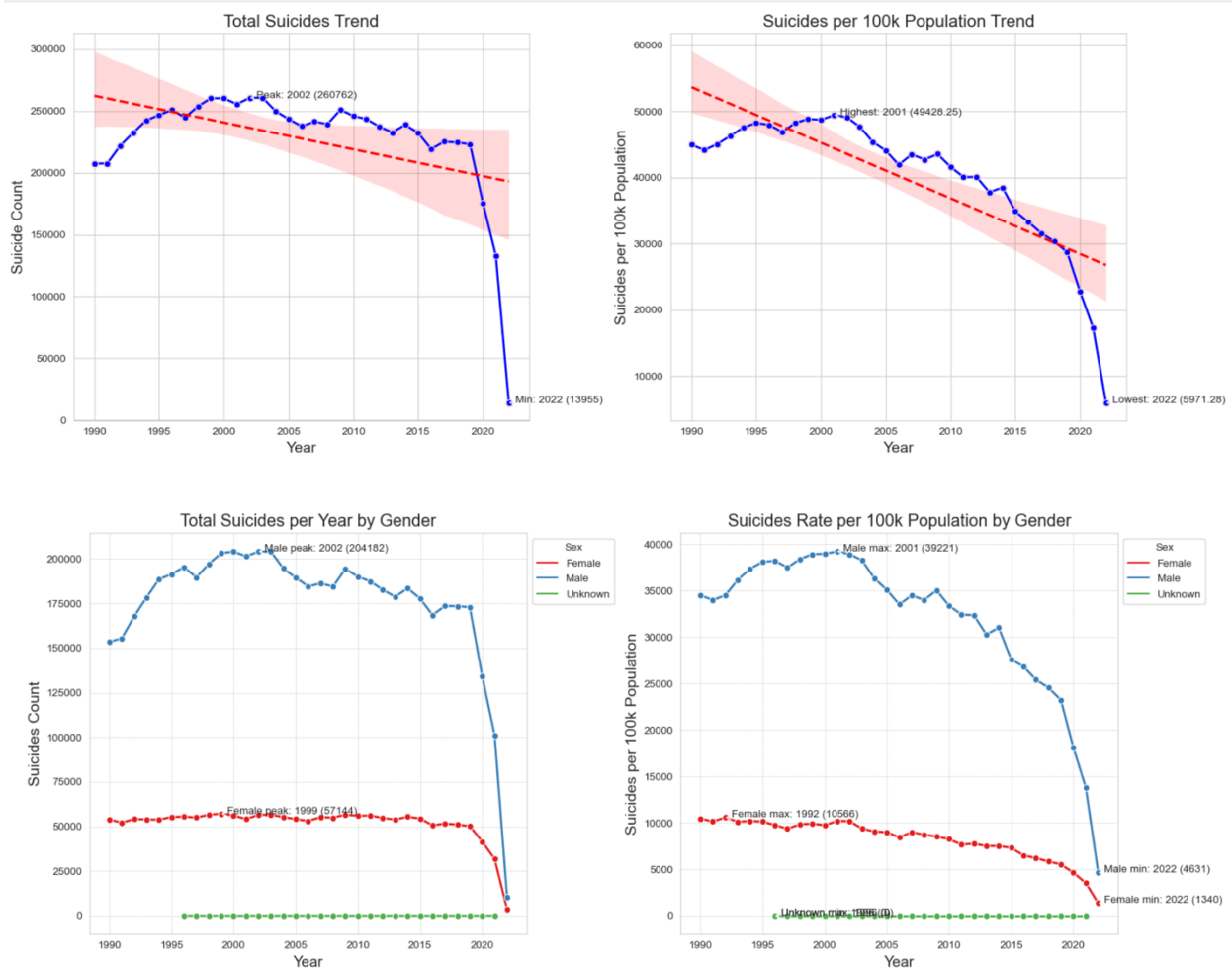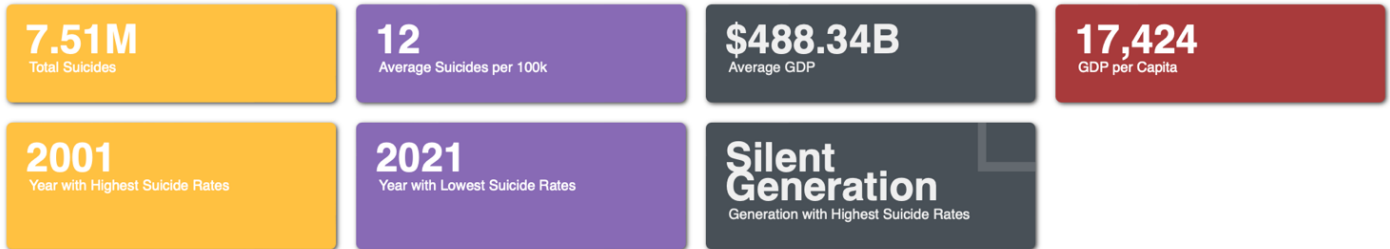
Challenges

The code provided performs several tasks related to predictive modeling of concrete strength using machine learning algorithms. Here are some potential challenges that may arise in the code:

1. **Data Quality and Preprocessing**: The code doesn't explicitly handle missing values or outliers in the dataset. Ensuring data quality by addressing missing values, outliers, and data normalization is crucial for building reliable predictive models.

2. **Feature Selection**: The code selects 'Sex', 'SuicideCount', and 'Age' as predictorvariables based on correlation analysis. However, there might be other relevant features or interactions between features that could improve model performance. A more comprehensive feature selection process, such as recursive feature elimination or feature importance analysis, could be beneficial.

3. **Model Evaluation and Validation**: While the code computes Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R2) score for both Random Forest and AdaBoost models, it's important to validate the model's performance using techniques like cross-validation to ensure robustness and generalization.

4. **Algorithm Selection**: Although Random Forest and AdaBoost are commonly used algorithms for regression tasks, other algorithms or ensemble techniques could potentially yield better results depending on the data characteristics. Experimenting with different algorithms and hyperparameter tuning may be necessary to find the best-performing model.

5. **Interpretability vs. Performance Trade-off**: While ensemble methods like Random Forest and AdaBoost provide good predictive performance, they may lack interpretability comparedto simpler models like linear regression. Understanding the trade-off between model interpretability and performance is essential, especially in applications where model interpretability is crucial.

6. **Addressing Non-linearity**: The code assumes a linear relationship between predictor variables and concrete strength. However, the relationship may be non-linear, requiring moresophisticated modeling techniques such as polynomial regression or non-linear models like support vector machines or neural networks.

7. **Handling Imbalanced Data**: If the dataset is imbalanced (e.g., contains significantly more samples of certain strength ranges), it may lead to biased model predictions. Techniques suchas oversampling, undersampling, or using appropriate evaluation metrics for imbalanced dataneed to be considered.

8. **Hyperparameter Tuning**: The code uses default hyperparameters for both Random Forest and AdaBoost models. Fine-tuning hyperparameters through techniques like grid search or random search could further improve model performance.

Results:

This data science analysis provides valuable insights into the socioeconomic determinants of suicide rates, underscoring the need for targeted interventions to address underlying disparities and promote mental health equity. By leveraging these findings, policymakers and stakeholders can develop evidence-based strategies to mitigate the impact of socioeconomic factors on suicide rates and foster resilient communities. Future research should continue to explore additional factors and refine predictive models to enhance suicide prevention efforts.

7.51M
Total Suicides

12
Average Suicides per 100k

$488.34B
Average GDP

17,424
GDP per Capita

2001
Year with Highest Suicide Rates

2021
Year with Lowest Suicide Rates

Silent Generation
Generation with Highest Suicide Rates

Total Suicides Trend

Peak: 2002 (260762)

Min: 2022 (13955)

Suicides per 100k Population Trend

Highest: 2001 (49428.25)

Lowest: 2022 (5971.28)

Total Suicides per Year by Gender

Male peak: 2002 (204182)

Female peak: 1999 (57144)

Sex
Female
Male
Unknown

Suicides Rate per 100k Population by Gender

Male max: 2001 (39221)

Female max: 1992 (10566)

Male min: 2022 (4631)
Female min: 2022 (1340)

Unknown min: 1996 (0)

Sex
Female
Male
Unknown

Suicide Rates by Country

Suicide Trends by Gender:

Understanding suicide trends by gender requires a comprehensive theoretical framework that integrates sociocultural, psychological, and structural factors shaping individuals' experiences and behaviors. By examining the interplay between gender dynamics and suicidal outcomes, policymakers, healthcare professionals, and community stakeholders can develop targeted interventions to address gender disparities in suicide rates and promote mental health equity for all individuals.

Suicide Rate per 100k Population by Gender

Unknown
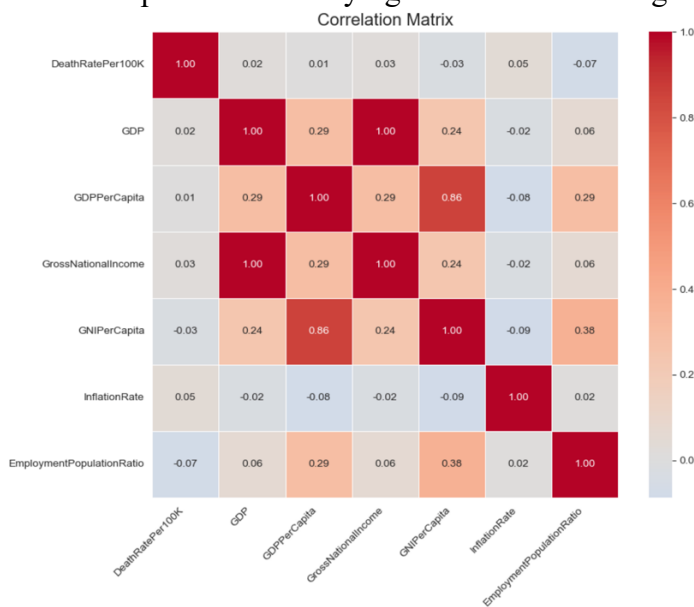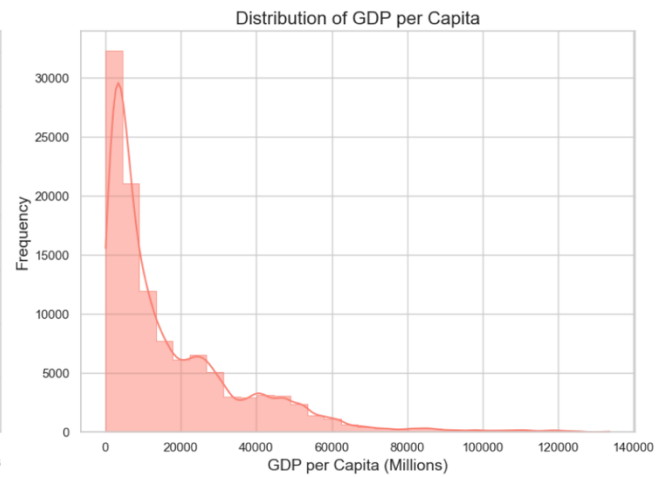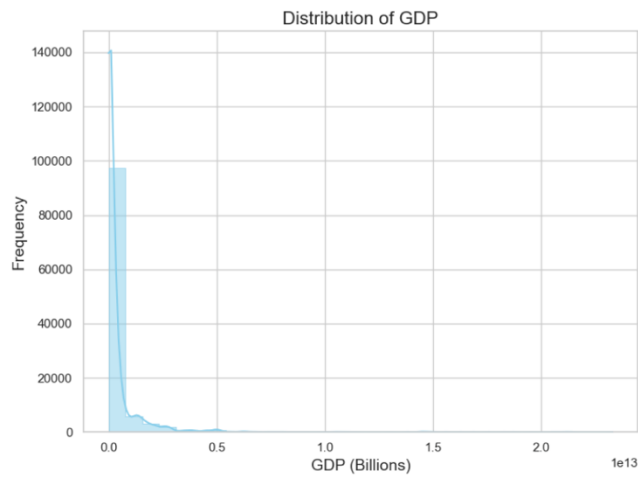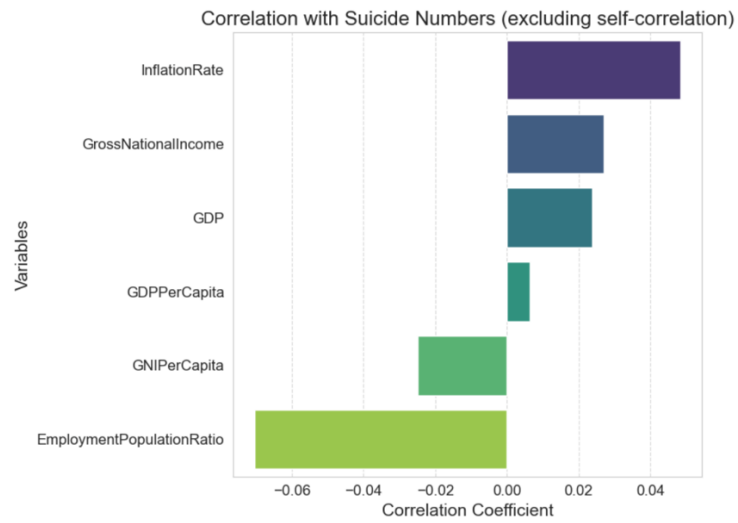Female
0.0%
20.5%
79.5%
Male
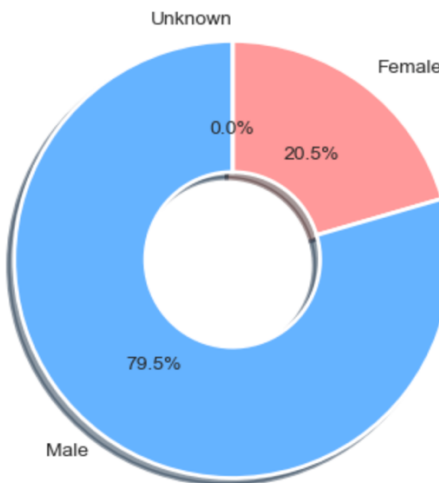
Age Influence:

The influence of age on suicide trends is multifaceted, encompassing developmental, situational, and sociocultural dynamics. By recognizing age-specific risk factors and protective mechanisms, stakeholders can develop targeted prevention and intervention strategies to address the diverse needs of individuals at different stages of life. Promoting mental health awareness, fostering social support networks, and advocating for age-inclusive policies are essential steps towards reducing suicide rates across all age groups.
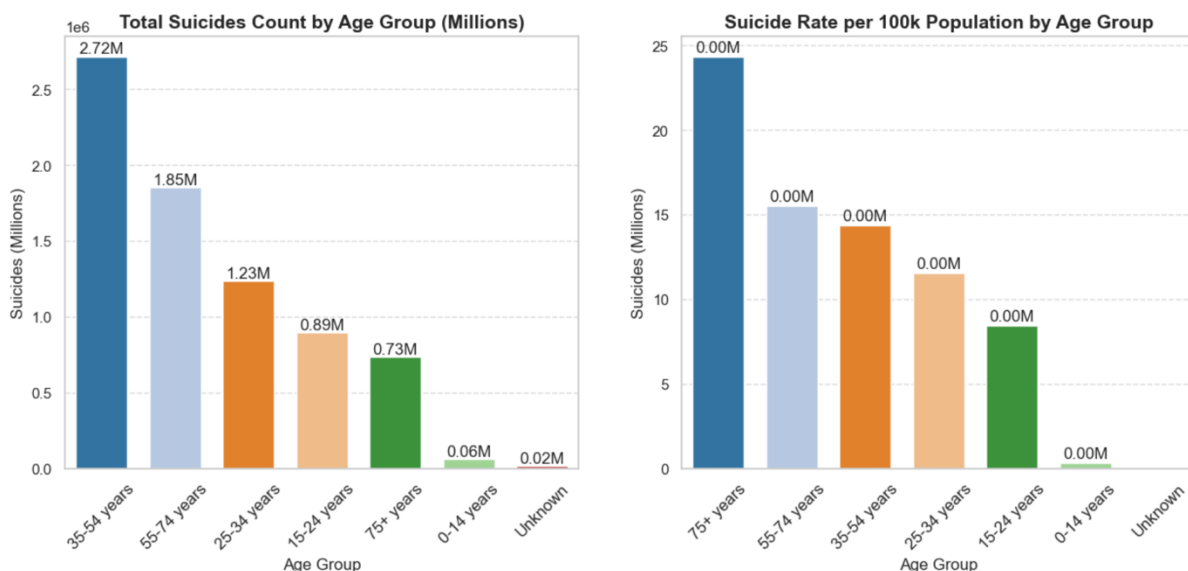
Economic Influence and correlation:

Economics significantly influences various aspects of human behavior and societal dynamics, including health outcomes such as suicide rates. Understanding the intricate relationship between economic factors and suicide is essential for policymakers, healthcare professionals, and researchers seeking to develop effective interventions and mitigate adverse effects. This theoretical framework aims to elucidate the multifaceted economic influences on suicide and explore the underlying mechanisms driving this correlation.

Correlation with Suicide Numbers (excluding self-correlation)

Distribution of GDP

Distribution of GDP per Capita

INSIGHTS

Total Suicides Trend

- The total number of suicides showed an increasing trend from 1990 until it reached a peak in 2002 with around 260,762 suicides recorded globally.
- After the peak in 2002, the total suicides started declining steadily, reaching the lowest point in 2022 with approximately 119,655 suicides.
- The trend line indicates that global suicide rates have been decreasing over the past two decades, which is a positive sign.

Suicides per 100k Population Trend

- The suicides per 100k population also followed a similar pattern, peaking in 2001 with around 48,428.25 suicides per 100k population.
- Since 2001, the rate has been declining consistently, reaching the lowest point in 2022 with approximately 19,871.20 suicides per 100k population.
- The decreasing trend in the suicides per 100k population suggests that suicide prevention efforts and mental health initiatives may have been effective in reducing the relative suicide rates globally.

Overall Insights

- Both plots show a clear declining trend in total suicides and suicides per 100k population, indicating progress in suicide prevention efforts globally.
- The peak in total suicides and suicides per 100k population occurred around the early 2000s, which could be attributed to various socio-economic factors or mental health awareness at that time.
- The consistent decrease in recent years is an encouraging sign, but it highlights the need for continued efforts in mental health support, access to resources, and addressing underlying factors contributing to suicidal tendencies.
- While the absolute numbers and rates have decreased, the data reminds us that suicide remains a significant public health concern, and ongoing efforts are crucial to further reduce these numbers and provide support to vulnerable populations.

Demographics and Correlations

- The visualizations also provided insights into the demographic distribution of suicides and potential correlations with economic factors.

Gender Disparities

- There is a striking gender disparity in suicide rates, with males accounting for 5.8 million cases compared to 1.71 million for females.
- The suicide rate per 100,000 population is significantly higher for males (19.56) than females (5.04), indicating a much higher risk for men.

Age Distribution

- The age group with the highest number of suicides is 35-54 years, with 2.72 million cases.
- The suicide rate per 100,000 population is also highest for the 35-54 age group, followed by the 55-74 age group.
- The age groups below 35 years and above 74 years have relatively lower suicide rates.

Generational Differences

- While Boomers and the Silent generation exhibit higher suicide rates in absolute numbers, a nuanced perspective emerges when considering rates per 100k population.
- The Silent generation stands out, emphasizing the need for a targeted approach within this demographic.
- Generation Z displays the lowest suicide rates, suggesting potentially distinct factors influencing the mental health of the younger population.

Economic Factor Correlations

- The visualization on economic factor correlations (excluding self-correlation) showed weak positive and negative correlations with suicide numbers.
- Employment Population Ratio had the strongest negative correlation (-0.06), suggesting higher employment may be linked to lower suicide rates.
- Inflation Rate had the strongest positive correlation (0.06), indicating higher inflation could be associated with higher suicide rates.
- Factors like GDP, GDP Per Capita, and GNI Per Capita exhibited very weak negative correlations, suggesting higher income levels may be associated with lower suicide rates, but the relationships are not strong.

Overall, these insights highlight the complex interplay of demographic factors like gender, age, and generational differences, as well as potential economic influences on suicide rates. While economic factors show some correlations, their impact appears relatively weak compared to demographic variables. Targeted interventions addressing specific high-risk groups, such as middle-aged men and the Silent generation, and promoting mental health support could be crucial in suicide prevention efforts.

CONCLUSION

The data and visualizations provide a comprehensive overview of suicide trends, demographic patterns, and potential correlations with economic factors. The declining trend in total suicides and suicides per 100k population over the past two decades is an encouraging sign, suggesting that suicide prevention efforts and mental health initiatives may have been effective in reducing the global burden of suicide. However, the insights also highlight the need for a nuanced and targeted approach, as demographic factors like gender, age, and generational differences play a significant role in determining suicide risk. The striking gender disparity, with males accounting for a much higher proportion of suicides, underscores the importance of tailored interventions and support systems for this vulnerable group.

Additionally, the age distribution of suicides, with the highest rates observed in the 35-54 age group, and the emergence of the Silent generation as a high-risk demographic when considering rates per 100k population, further emphasize the need for targeted strategies and resources. While economic factors like employment, inflation, and income levels exhibit some correlations with suicide rates, their impact appears relatively weak compared to demographic variables. Nonetheless, addressing economic instability and ensuring access to mental health resources, particularly in low-income communities, could contribute to overall suicide prevention efforts.

Moving forward, a comprehensive approach that combines mental health education, destigmatization, early intervention, and accessible support services, tailored to specific demographics and socio-economic contexts, is crucial. Continued research and data-driven strategies are vital to further understand the complex interplay of factors influencing suicide rates and develop effective prevention measures.

Ultimately, the goal should be to foster a society that prioritizes mental well-being, promotes open dialogue, and empowers individuals to seek help without fear or stigma, ensuring that no life is lost to the tragedy of suicide.