IJAN OCTOBER 2023 DATA ANALYSIS PHYSICAL CLASS

DATA ANALYTICS PROJECT

PROJECT OBJECTIVE: PERFORM AN EXPLORATORY DATA ANALYSIS ON THE DATASET GIVEN TO DISCOVER IF;

- CRUDE DEATH RATES CAN BE INFLUENCED BY POPULATION AGE STRUCTURES
- CORRELATION BETWEEN SUICIDE RATES AND COUNTRY GDP/GNI

COURSE: BEGINNER DATA ANALYSIS
LECTURER: MR. AYOMIDE SAMUEL

PREPARED BY: M. F. MIKAIL

- CRUDE DEATH RATES CAN BE INFLUENCED BY POPULATION AGE STRUCTURES
- CORRELATION BETWEEN SUICIDE RATES AND COUNTRY GDP/GNI

BREIF OUTLINE

This Dataset given to us to work on was sourced on Kaggle, it contains two excel files named Suicide rate 1990-2022 and Age Std Suicide Rate 1990-2022. For the purpose of this project am going to be working mainly with the Suicide Rate 1990-2022, I would like to state that the other Dataset was cleaned, validated and also the visualization was done. Before we dive into the exploratory analysis of the dataset, I would like to define each heading of the dataset.

SUICIDE RATE

The suicide rate is a measure that quantifies the frequency of suicide within a specific population over a specified period, usually expressed as the number of suicides per a standard unit of population (often per 100,000 individuals) within a given timeframe, such as a year. It is a critical indicator used in public health to assess the prevalence and impact of suicide within different demographic groups, regions, or countries.

AGE STD

Age STD" refers to "Age Standardization" in the context of population data analysis. Age standardization is a statistical method used to adjust for differences in age distributions when comparing populations with different age structures.

When analyzing data related to populations, especially in studies comparing different regions or time periods, it's essential to consider the age distribution of the populations being studied. Age standardization allows for a fair comparison by removing the effect of differences in age structures.

For example, if you were comparing the suicide rates of two countries with different age distributions, you might age-standardize the rates to ensure that any differences observed are not simply due to variations in the age demographics of the populations.

Age standardization involves applying age-specific rates (such as suicide rates) from a standard population to the age distribution of the populations being compared. This process produces age-standardized rates, which represent what the rates would be if the populations had the same age distribution as the standard population.

In summary, "Age STD" in relation to population likely refers to the process of age standardization, which is a statistical technique used to adjust for differences in age distributions when comparing populations.

Insights Report on Suicide Rates, Causes of Death, and Socioeconomic Indicators Across Regions and Countries

    1. **Introduction:**

Understanding the dynamics of suicide rates across different regions and countries is crucial for public health interventions and policy-making. This analysis delves into a comprehensive examination of suicide data, encompassing various demographic and socio-economic variables. By scrutinizing factors such as age, sex, economic indicators, and cause-specific death percentages.

- This report analyzes a dataset containing information on suicide rates, cause-specific death percentages, socioeconomic indicators, and demographic data across various regions and countries.
- The dataset covers multiple years and includes variables such as region code, country code, suicide count, suicide rate, GDP, GNI, population, and more.

2. Overview of Data:

- The dataset comprises entries for different regions and countries, spanning multiple years from 1990-2022.
- Variables include demographic indicators, suicide counts and rates, cause-specific death percentages, and socioeconomic metrics such as GDP, GNI, inflation rate, and employment-population ratio.
- The dataset encompasses a wide array of variables including region code, region name, country code, country name, year, sex, age group, generation, suicide count, suicide rate (%), cause-specific death percentage (%), death rate per 100K (%), population, GDP ($), GDP per capita ($), gross national income (GNI) ($), GNI per capita, inflation rate, and employment population ratio
- The size of the data is 13,120KB, Row Count is 95244, the number of Rows 16, the number of Column is 81,893, the suicide count across region is 7514684, the total population is 2trilion, suicide count for male is 47,774, suicide count for female is 39,390.

3. Key Insights:

3.1 Suicide Rates and Demographics:

- According to this Analsis suicide counts and rates across different regions and countries is 2434134588944 Approximately 2 trillion and the sum of suicide count is 7514684, in. this pattern it means Suicide Rate is Approximately 0.0003% of the entire population.
- The Examination of this data suggest that suicide rates is prevalent among the Male than the female, with suicide rate in male is 47,774 count and in Female 39,390 count. To understand the demographic vulnerabilities Age Group 35-54 years has the

highest suicide rate with 2716428 count, followed by 55-74 years with 1854582, next is 23-34years with 1234665 and the next is 75+ years with 734862 count, the Age Group with the lowest suicide rate is age group 0-14years with 62986 count of suicide.

- The result of Investigating correlations between suicide rates and population density, GDP per capita, and other demographic factors shows that At 780166512606, Europe had the highest Sum of Population and was 2,710.44% higher than Oceania, which had the lowest Sum of Population at 27759618474. Europe accounted for 32.05% of Sum of Population. Across all 7 Region Name, Sum of Suicide Count ranged from 15592 to 7514684. Sum of Inflation Rate and total Sum of Year are positively correlated with each other. Sum of Year and Sum of Inflation Rate diverged the most when the Region Name was Europe, when Sum of Year were 79,284,702.49 higher than Sum of Inflation Rate.

3.2 Cause-Specific Death Percentages:
- The mortality rate of a population is determined by this factors: Age, Healthcare Access and Quality, Lifestyle and Behavior, Environmental Factors, Genetics, Socioeconomic Status but the most prevalent factor is Age- Older populations tend to have higher mortality rates due to age-related health issues and decreased immunity which means the cause-specific death percentage should be higher than suicide rate but in this data cause specific death rate is lower than suicide rate.

3.3 Socioeconomic Indicators:
- In evaluating the relationship between socioeconomic factors (GDP, GNI, inflation rate, employment-population ratio) and suicide rates this Dataset studies have found weak or no significant correlation between suicide rates and GDP or GNI. This may be due to the complex interplay of multiple factors influencing suicidal behavior, including cultural norms, social cohesion, access to lethal means, mental health stigma, and healthcare infrastructure.

- While there may be a correlation between suicide rates and country GDP/GNI, the relationship is complex and influenced by various factors. Some research suggests a negative correlation between suicide rates and GDP or GNI, indicating that higher economic prosperity is associated with lower suicide rates. This could be attributed to improved access to mental healthcare, better social support systems, and reduced financial stress in wealthier countries.

- According to this Dataset, regions with higher Gross National Income (GNI) and Gross Domestic Product (GDP), such as North America and The Caribbean (NAC) and Asia (AS), tend to have higher suicide rates, indicating a correlation between economic indicators and suicide occurrences.

- Across all 2,767 Population, Sum of Employment Population Ratio ranged from 0.00% to 4189.15%. Sum of Employment Population Ratio and Sum of Suicide Rate (%) diverged the most when the Year was 2009, when Sum of Employment Population Ratio were 176785.15% higher than Sum of Suicide Rate (%).

4. Regional Specific Analysis:
- Europe has the highest rate of suicide with **3624527 count while Africa has 15592 count, this might be due to population because Europe accounts for 32% of the world population.**

- THIS PROJECTS ASKS TWO CRUCIAL QUESTION
    1. DISCOVER IF CRUDE DEATH RATES CAN BE INFLUENCED BY POPULATION AGE STRUCTURE

To discover if crude death rate can be influenced by the population age structure, you would need to analyze the relationship between the age distribution of a population and the observed death rates. Here's how you could approach this analysis using the data headings you've provided:

The crude death rate is a demographic measure that represents the number of deaths occurring in a given population during a specified period, usually per 1,000 individuals per year.

Population age structure refers to the distribution of individuals across different age groups within a population.

Here's how population age structure can influence the crude death rate:

1. **Age-specific mortality rates**: Different age groups within a population experience varying levels of mortality. For instance, infants and elderly individuals typically have higher mortality rates compared to young adults. Therefore, the proportion of these age groups within the population will affect the overall crude death rate. A population with a higher proportion of elderly individuals will likely have a higher crude death rate compared to a population with a younger age structure. According to the dataset Age Group 35-54 years has the highest suicide rate with 2716428 count, followed by 55-74 years with 1854582, next is 23-34years with 1234665 and the next is 75+ years

with 734862 count, the Age Group with the lowest suicide rate is age group 0-14years with 62986 count of suicide.

2. **Life expectancy**: Population age structure can provide insights into the life expectancy of individuals within that population. A population with a larger proportion of older individuals may have a lower life expectancy due to higher mortality rates associated with advanced age. Consequently, such a population may exhibit a higher crude death rate.

3. **Healthcare needs and services**: The age structure of a population influences its healthcare needs and the provision of healthcare services. A population with a larger proportion of elderly individuals may require more extensive healthcare services to address age-related illnesses and conditions. In contrast, a population with a higher proportion of young individuals may have different healthcare needs. Disparities in healthcare access and quality across age groups can impact mortality rates and, consequently, the crude death rate.

4. **Epidemiological factors**: Certain diseases and health conditions may disproportionately affect specific age groups within a population. For instance, infectious diseases such as influenza and pneumonia might have a more significant impact on the elderly, while conditions like childhood pneumonia may affect younger age groups more prominently. The prevalence of these diseases within different age cohorts can influence the overall crude death rate.

5. **Socioeconomic factors**: Population age structure is often correlated with socioeconomic factors such as education, income, and access to healthcare. These factors can influence mortality rates through their effects on lifestyle choices, healthcare utilization, and exposure to health risks. For example, socioeconomically disadvantaged groups may experience higher mortality rates regardless of age due to limited access to healthcare and higher exposure to risk factors such as poor nutrition and environmental hazards.

In summary, population age structure plays a crucial role in shaping the crude death rate by influencing age-specific mortality rates, healthcare needs, epidemiological patterns, and socioeconomic disparities within a population. Understanding these relationships is essential for policymakers and public health officials to effectively address health challenges and improve overall population health outcomes.

- THIS PROJECTS ASKS TWO CRUCIAL QUESTION
  2. **DISCOVER THE CORRELATION BETWEEN SUICIDE RATES AND COUNTRY GDP/GNI**

Analyzing the correlation between suicide rates and country GDP (Gross Domestic Product) or GNI (Gross National Income) provides valuable insights into the relationship between

economic factors and mental health outcomes. Here's a detailed discussion on this correlation:

1. Economic Conditions and Mental Health:
   - Economic conditions can significantly impact mental health outcomes, including suicide rates. Unemployment, poverty, income inequality, and economic instability are known risk factors for mental health disorders and suicidal behavior.
   - Higher levels of economic development and prosperity are generally associated with better mental health outcomes due to increased access to resources, social support networks, and healthcare services.

2. Correlation Analysis:
   - Conducting a correlation analysis between suicide rates and country GDP or GNI helps quantify the relationship between economic indicators and suicide risk.
   - The correlation coefficient (usually denoted as "r") measures the strength and direction of the linear relationship between two variables, ranging from -1 to +1. A positive correlation indicates that as one variable increases, the other variable also tends to increase, while a negative correlation indicates an inverse relationship.

3. Research Findings:
   - Studies have reported mixed findings regarding the correlation between suicide rates and economic indicators.
   - Some research suggests a negative correlation between suicide rates and GDP or GNI, indicating that higher economic prosperity is associated with lower suicide rates. This could be attributed to improved access to mental healthcare, better social support systems, and reduced financial stress in wealthier countries.
   - However, other studies have found either weak or no significant correlation between suicide rates and GDP or GNI. This may be due to the complex interplay of multiple factors influencing suicidal behavior, including cultural norms, social cohesion, access to lethal means, mental health stigma, and healthcare infrastructure.

4. Moderating Factors:
   - The correlation between suicide rates and economic indicators may vary depending on several moderating factors:
     - Cultural Factors: Cultural attitudes towards mental health, suicide, and help-seeking behaviors can influence the relationship between economic conditions and suicide rates. For example, in some cultures, stigma surrounding mental illness may deter individuals from seeking help despite economic prosperity.
     - Social Support Systems: Strong social support networks and community cohesion can buffer the adverse effects of economic hardship on mental health, potentially weakening the correlation between economic indicators and suicide rates.

- Healthcare Infrastructure: The availability and quality of mental healthcare services play a crucial role in mitigating suicide risk. Even in countries with high GDP or GNI, inadequate mental health resources can contribute to higher suicide rates.

5. Policy Implications:
  - Understanding the correlation between suicide rates and country GDP/GNI has important policy implications:
    - Policymakers should prioritize investments in mental health services, suicide prevention programs, and social safety nets, regardless of a country's economic status.
    - Economic development policies should be complemented by measures to address social determinants of mental health, such as income inequality, unemployment, and social isolation.
    - Cross-national comparisons can inform evidence-based strategies for suicide prevention and mental health promotion, taking into account both economic and sociocultural factors.

While there may be a correlation between suicide rates and country GDP/GNI, the relationship is complex and influenced by various factors. Understanding these nuances is essential for designing effective policies and interventions to promote mental health and prevent suicide.

5. Policy Implications and Recommendations:
Based on the insights derived from the data analysis, although this is an explorative data analysis but the dataset analysed is a global problem and the analysis wont be complete without providing recommendations for this menace called suicide.

To addressing Mental Health Issues and Suicide Prevention:

- Increase Mental Health Awareness: Launch campaigns to reduce stigma, raise awareness, and promote mental health education in communities.

- Enhance Access to Mental Health Services: Expand mental health services, including hotlines, counseling, and therapy, making them more accessible and affordable.

- Implement Suicide Prevention Programs: Develop and implement programs targeting at-risk populations, providing support, counseling, and resources for those struggling with suicidal thoughts.

- Training for Professionals: Provide training for healthcare providers, educators, and community workers on recognizing signs of mental health issues and suicide risk.

- Initiatives for Economic Development and Socioeconomic Disparities Reduction:

- Job Creation Programs: Implement initiatives to create job opportunities, especially in areas with high unemployment rates, to improve economic stability and mental well-being.

- Skills Development Training: Offer training programs to enhance skills and employability, helping individuals secure sustainable employment.

- Promote Entrepreneurship: Support entrepreneurship and small business development to stimulate economic growth and create employment opportunities.

- Social Welfare Programs: Introduce social welfare programs to provide financial assistance, healthcare, and support to vulnerable populations, reducing socioeconomic disparities.

- Education and Training: Enhance educational opportunities and vocational training programs to empower individuals and improve their socioeconomic status.

- By combining efforts to address mental health issues, prevent suicides, and improve access to mental health services with initiatives promoting economic development, reducing unemployment, and mitigating socioeconomic disparities, communities can work towards creating a healthier and more supportive environment for all individuals.

6. Conclusion:

This analysis highlights the multifaceted nature of suicide and its relationship with various demographic, socio-economic, and cultural factors. Key findings from the data analysis presented Data Visualization:

1. Males consistently exhibit higher suicide rates compared to females over the years 1990 to 2022, with a notable difference in average numbers.
2. Female suicide rates demonstrate more stability over time, in contrast to the fluctuating trend seen in male suicide rates.
3. A decline in suicide rates is observed from 2019 onwards.
4. The age group with the highest number of suicide cases is individuals aged 35–54, followed by 55–74 and 25–34 age groups.
5. Regions with higher Gross National Income (GNI) and Gross Domestic Product (GDP), such as North America and The Caribbean (NAC) and Asia (AS), tend to

have higher suicide rates, indicating a correlation between economic indicators and suicide occurrences.

6. There is a plausible inference that individuals in their most productive age groups are more susceptible to suicide based on the data analysis.

In summary, the data reveals a concerning trend of higher suicide rates among males, a correlation between economic indicators and suicide rates, and a concentration of suicide cases within certain age groups, particularly those in their peak career years.

7. Further Research:
- We find that while certain trends may be consistent across regions and countries, there are also significant disparities and complexities that warrant further investigation.

- Further research is needed to explore the nuanced relationship between economic indicators and suicide rates, considering regional variations, time trends, and specific population subgroups.
- Longitudinal studies and multilevel analyses can provide insights into the dynamic interactions between economic conditions, mental health outcomes, and contextual factors.

- further research, such as exploring additional factors that may influence suicide rates, and investigating the effectiveness of interventions in reducing suicide risk.

References

World Health Organization
https://chat.openai.com/
https://www.popai.pro/chat/5f82fbae-9291-4d19-9aeb-e7a78a963d63
https://www.datacamp.com/blog/unique-ways-to-use-ai-in-data-analytics