SUICIDE RATES EDA PROJECT REPORT

Manthan Pandey

EXECUTIVE SUMMARY

This project provides a comprehensive, data-driven assessment of global suicide trends from 1990 to 2022, integrating demographic patterns with macroeconomic indicators to understand the evolving landscape of suicide burden worldwide. Using a multi-layered analytical approach, the study evaluates suicide outcomes across countries, age groups, generations, sex categories, and population-size quantiles. The goal is to identify high-risk segments, uncover structural patterns, and evaluate whether socioeconomic conditions such as GDP, income levels, inflation rates, and employment ratios meaningfully contribute to suicide variations over time.

The analysis reveals substantial heterogeneity across demographic groups and geographies. Consistent with global literature, males exhibit significantly higher suicide rates than females across nearly all time periods and regions. Age-based and generational comparisons show that older individuals tend to experience higher suicide rates per 100,000 population, underscoring the importance of age-sensitive mental health interventions. Population quantile analysis demonstrates that countries with larger populations contribute disproportionately to global suicide counts—not because of higher per-capita risk, but due to the sheer size of their populations. Our ranking of top countries within the highest population quantile further highlights specific nations that account for a large share of the global burden and may benefit from targeted support and policy collaboration.

From a temporal perspective, both suicide counts and rates show clear fluctuations across the 32-year period, with identifiable peaks and troughs. These trends may reflect a combination of global economic shifts, social changes, national prevention efforts, reporting practices, and access to mental health care. The study identifies these inflection points and highlights the years in which the global burden was at its highest and lowest (excluding 2022 to avoid incomplete reporting bias).

To explore the socioeconomic dimension, the project evaluates correlations between DeathRatePer100K and indicators such as GDP, GDP per capita, Gross National Income, inflation rate, and the employment-to-population ratio. Surprisingly, linear correlations are generally weak—suggesting that suicide risk is not strongly or directly associated with the economic metrics included in the dataset. While slight tendencies appear (e.g., lower employment ratios aligning with higher

suicide rates, and higher inflation aligning weakly with increased risk), the magnitudes are insufficient to support strong economic determinism. This highlights the likelihood that suicide is influenced by a more complex interplay of psychological, cultural, environmental, and healthcare-related variables that extend beyond macroeconomic indicators.

The study concludes that while demographic and geographic patterns offer clear insights into risk distribution, macroeconomic indicators alone do not adequately explain variations in suicide rates. This underscores the importance of multi-dimensional prevention strategies—combining demographic targeting, mental health service expansion, community-level interventions, stigma reduction programs, and region-specific policies.

Overall, this analysis provides a deep foundational understanding of global suicide patterns and offers clear direction for data-informed strategy development. The insights derived here can support health agencies, NGOs, governmental bodies, and international organizations in designing more precise, equitable, and context-sensitive suicide prevention initiatives.

PROJECT OBJECTIVES

1. Identify long-term global trends in suicide counts and suicide rates.
2. Compare suicide outcomes across sex, age groups, and generational cohorts.
3. Examine population quantile differences and highlight top-burden countries.
4. Evaluate the correlation between suicide rates and economic indicators.
5. Provide actionable insights to inform mental health policy and prevention programs.

DATA SOURCES AND STRUCTURE

This analysis uses the dataset:

- *suicide_rates_1990-2022.csv*
  (Contains demographic, economic, and suicide indicators for each country-year.)

Key variables include:

- Demographics: Country, Year, Sex, AgeGroup, Generation
- Outcome Metrics: SuicideCount, DeathRatePer100K
- Population Indicators: Population, Population Quantile (Q1–Q4)
- Economic Indicators: GDP, GDP per capita, GNI, GNI per capita, Inflation, Employment Ratio

```
Data columns (total 18 columns):
 #   Column                        Non-Null Count   Dtype
---  ------                        --------------   -----
 0   RegionCode                    118560 non-null  object
 1   RegionName                    118560 non-null  object
 2   CountryCode                   118560 non-null  object
 3   CountryName                   118560 non-null  object
 4   Year                          118560 non-null  int64
 5   Sex                           118560 non-null  object
 6   AgeGroup                      118560 non-null  object
 7   Generation                    118560 non-null  object
 8   SuicideCount                  118096 non-null  float64
 9   CauseSpecificDeathPercentage  114271 non-null  float64
 10  DeathRatePer100K              107896 non-null  float64
 11  Population                    112640 non-null  float64
 12  GDP                           111320 non-null  float64
 13  GDPPerCapita                  111320 non-null  float64
 14  GrossNationalIncome           108600 non-null  float64
 15  GNIPerCapita                  107800 non-null  float64
 16  InflationRate                 104100 non-null  float64
 17  EmploymentPopulationRatio     107440 non-null  float64
dtypes: float64(10), int64(1), object(7)
```

	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

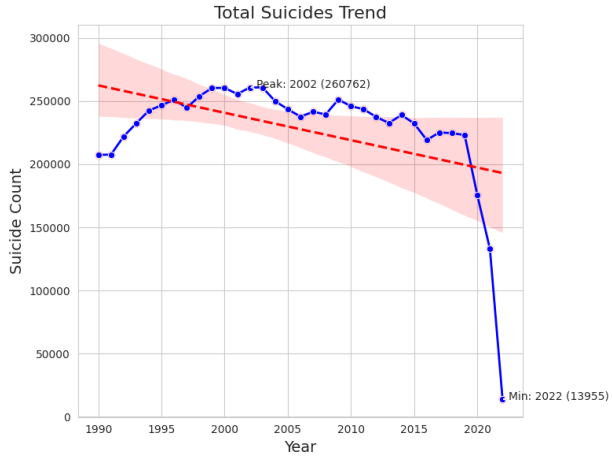
METHODOLOGY

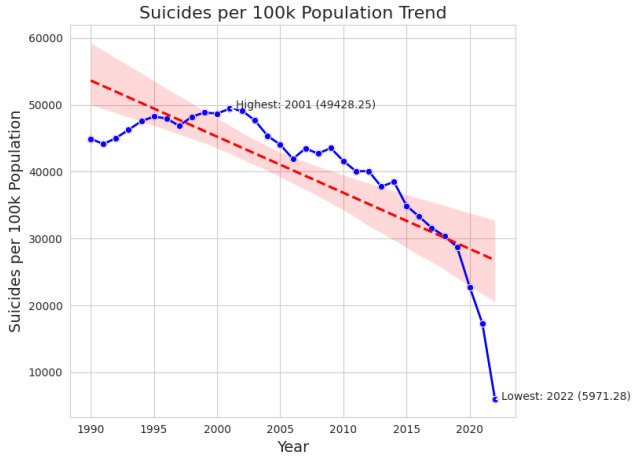
1. ***Data Loading & Cleaning***
   - Loaded CSV files, validated schema, inspected missing values.
2. ***Exploratory Data Analysis (EDA)***
   - Reviewed distribution patterns, demographics, and yearly trends.
3. ***Visual Analytics***
   - Generated line charts, bar charts, stacked visuals, and heatmaps.
4. ***Correlation Assessment***
   - Calculated Pearson correlations with macroeconomic variables.
5. ***Country-Level Burden Assessment***
   - Identified top-10 high-burden nations within Q4.

KEY FINDINGS

1. Global Temporal Trends

Suicide counts and rates show clear peaks and troughs across the 32-year period. These fluctuations may align with global recessions, sociopolitical stressors, changing health policies, or variations in public health infrastructure.

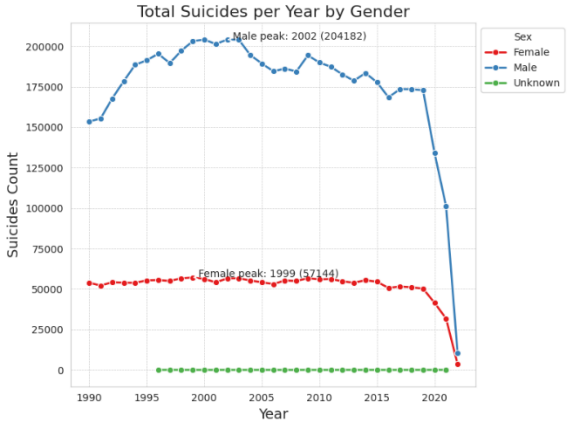


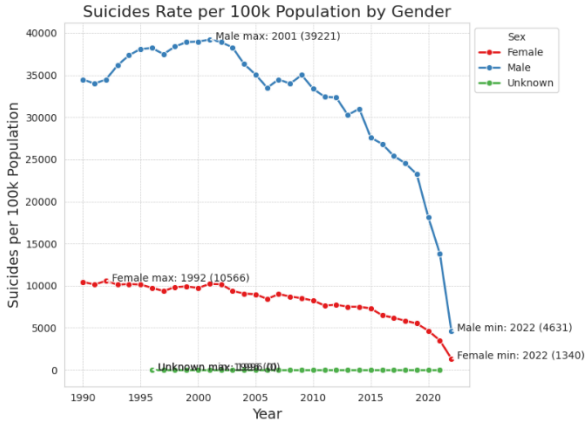


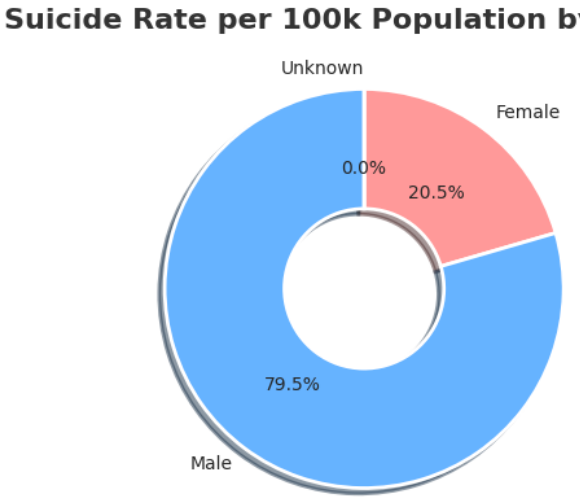
CountryName	Russian Federation	United States of America	Japan	Germany	Republic of Korea	France	Ukraine	Brazil	Poland	Mexico
Year										
1990	39150.0	30906.0	20088.0	13924.0	3159.0	11403.0	10693.0	4843.0	4970.0	1938.0
1991	39388.0	30810.0	19875.0	14010.0	3069.0	11502.0	10743.0	5182.0	5316.0	2114.0
1992	46125.0	30484.0	20893.0	13458.0	3533.0	11644.0	11731.0	5252.0	5713.0	2246.0
1993	56136.0	31102.0	20516.0	12690.0	4124.0	12251.0	12541.0	5547.0	5624.0	2358.0
1994	61886.0	31142.0	20923.0	12718.0	4212.0	12041.0	13907.0	5931.0	5519.0	2603.0

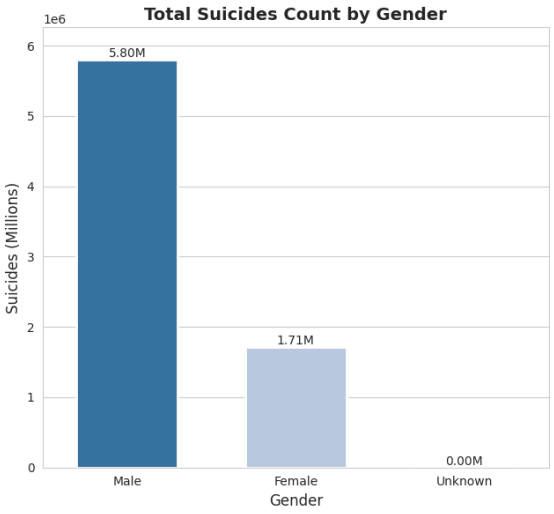
2. Sex-Based Patterns

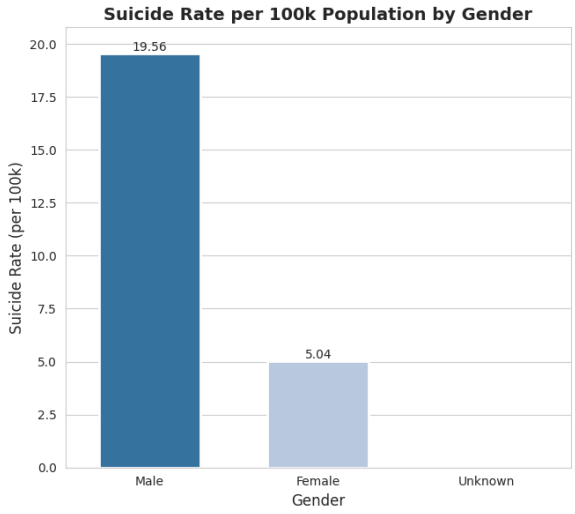
Across the entire dataset, males consistently account for the majority of suicides in both count and rate. The gap remains steady over time.





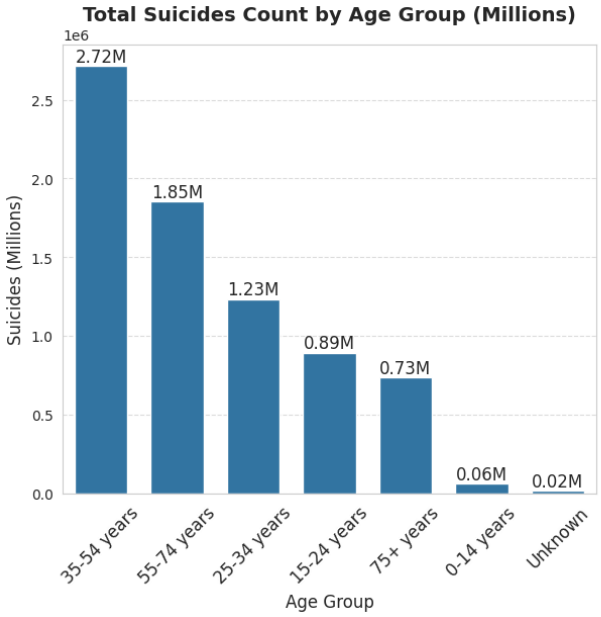


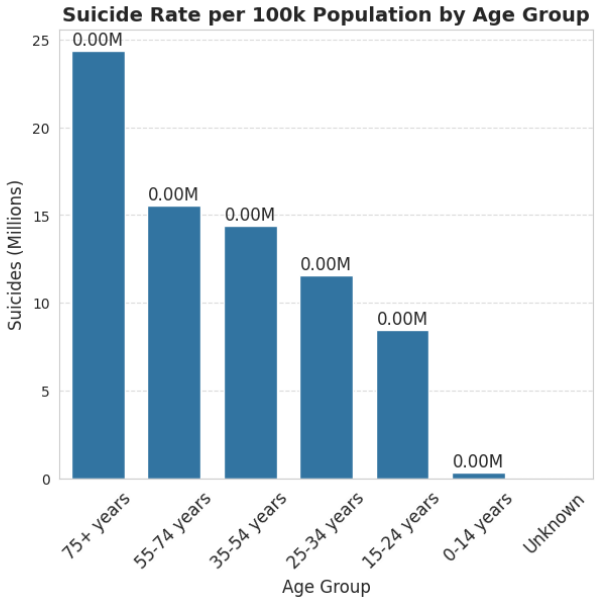


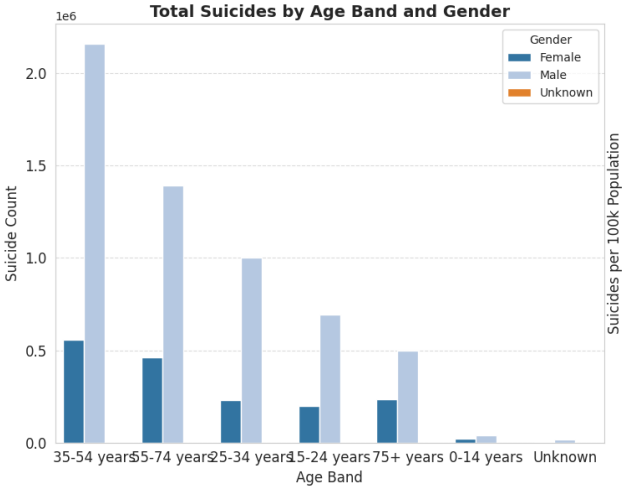


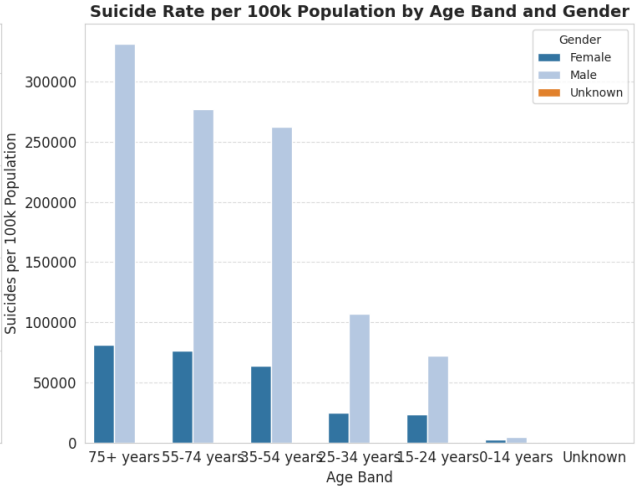
3. Age & Generational Insights

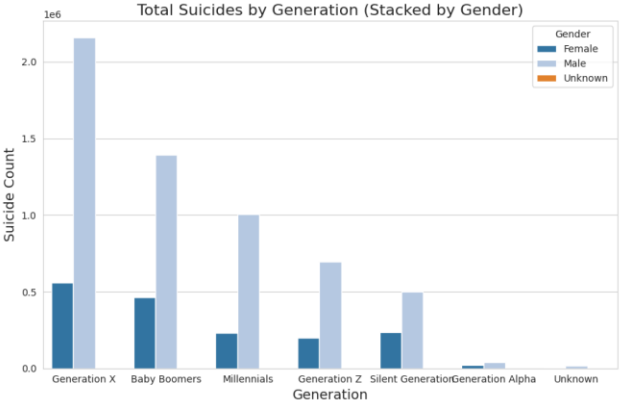
Older age groups exhibit higher suicide rates, while middle-aged groups often represent a large share of total counts.

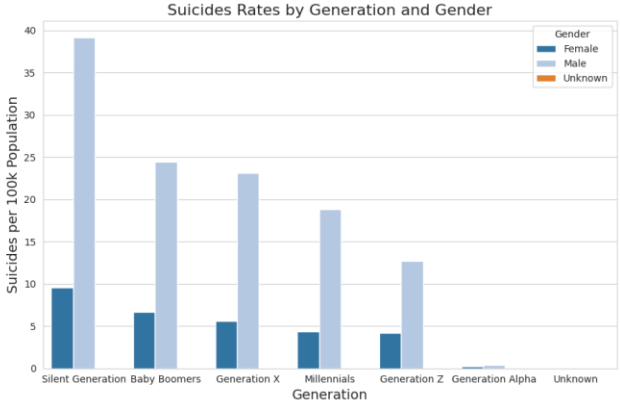






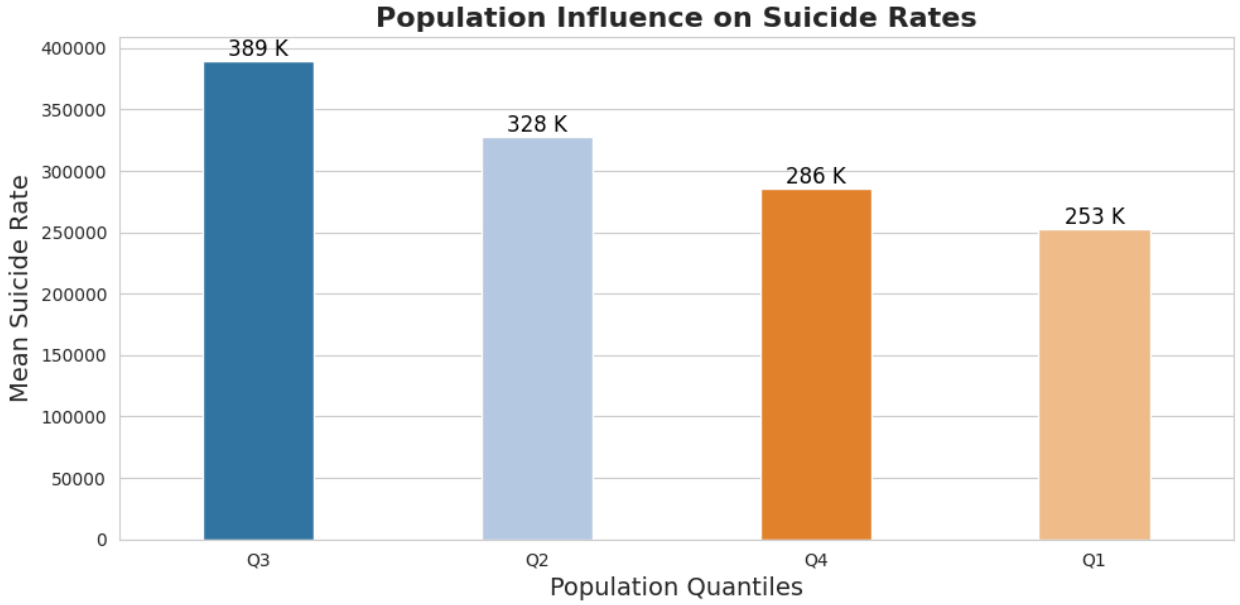






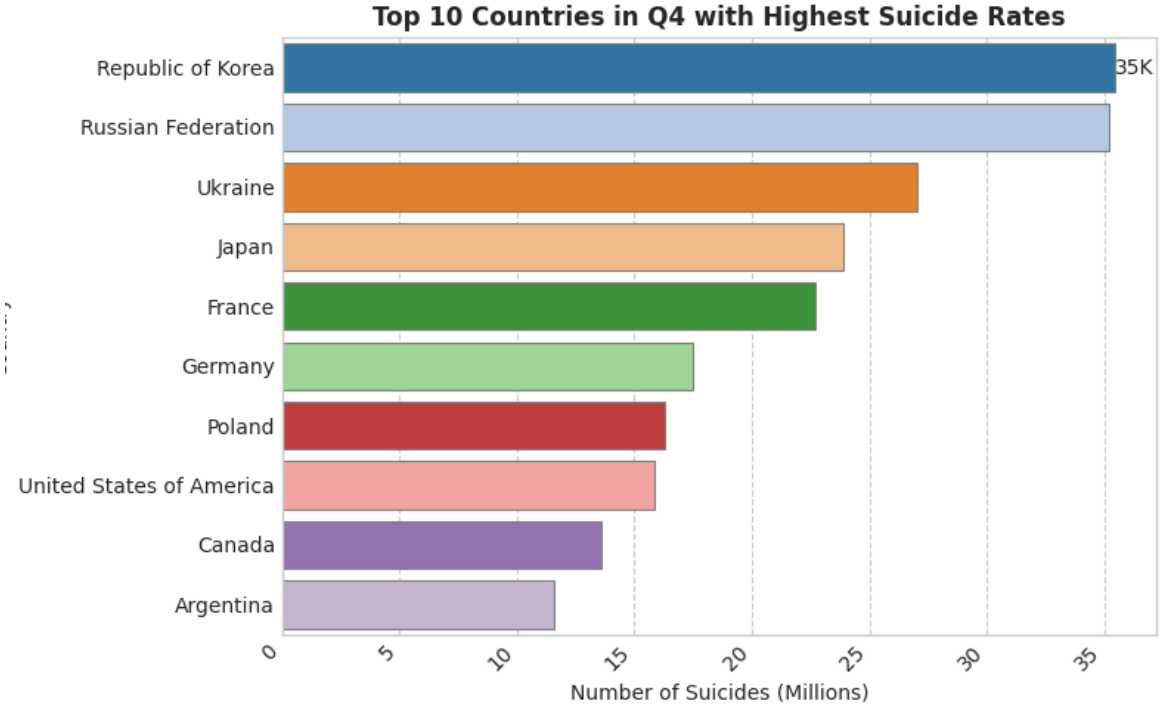
4. Population Quantile Analysis (Q1–Q4)

Countries in Q4 (largest populations) contribute most to the global burden. This is expected due to population size rather than elevated risk



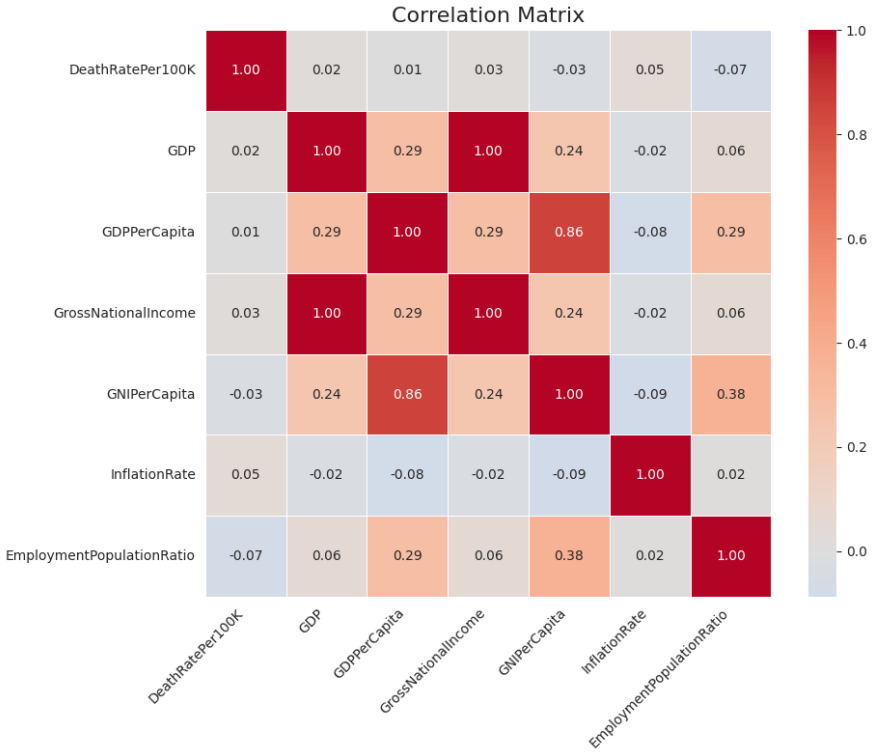
5. Top 10 Countries in Q4

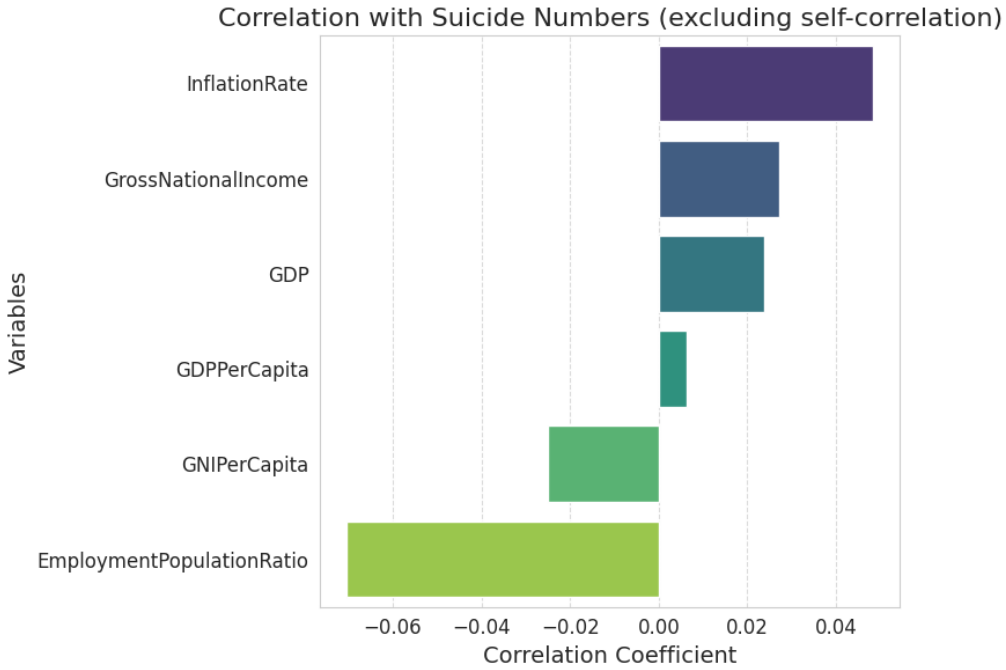
A small group of nations dominate global suicide counts, highlighting key targets for focused policy support.



6. Socioeconomic Correlation Findings

Despite expectations, macroeconomic variables show weak correlations with suicide rates. The lack of strong linear association indicates that suicide risk is driven by complex social and psychological factors beyond economic conditions alone.





INTERPRETATION & IMPLICATIONS

The findings of this study highlight that global suicide patterns are shaped far more by demographic and social structures than by straightforward economic metrics. The consistently higher suicide rates among males reflect long-standing gender differences in mental health expression, coping mechanisms, access to support, and help-seeking behavior. Similarly, the elevated risk observed in older age groups suggests that aging populations may be disproportionately affected by factors such as social isolation, chronic illness, reduced mobility, and limited access to mental health services. The distribution of suicide counts across population quantiles, particularly the concentration of the global burden within the largest countries, emphasizes the need for regionally tailored interventions rather than one-size-fits-all strategies. Furthermore, the weak correlations between suicide rates and macroeconomic indicators indicate that economic conditions—while influential in broader societal well-being—do not fully account for suicide trends. This underscores the importance of examining non-economic drivers such as cultural stigma, mental health literacy, social cohesion, substance use, trauma exposure, and access to timely psychological care. For policymakers and organizations, the implication is clear: effective suicide prevention strategies must combine demographic targeting with improvements in mental health infrastructure, culturally sensitive interventions, and community-based support systems. Investments in crisis prevention, early identification programs, and mental health education are likely to yield more meaningful reductions in suicide rates than interventions focused solely on economic improvement. Overall, these insights support a multi-layered public health approach that prioritizes both population-level strategies and high-risk group interventions.

RECOMMENDATIONS

Based on the insights derived from the analysis, it is recommended that suicide prevention efforts adopt a multi-dimensional and targeted approach rather than relying solely on broad economic interventions. First, demographic-specific mental health programs should be prioritized, with particular attention given to males and older age groups who consistently exhibit elevated suicide rates. This includes developing accessible community-based services, gender-sensitive mental health campaigns, and age-tailored interventions that address the unique challenges faced

by older adults, such as social isolation and chronic health issues. In addition, countries identified as high-burden—particularly those within the largest population quantile—should receive focused resources and support. This may involve collaborating with local governments, NGOs, and health agencies to implement culturally appropriate prevention strategies, strengthen crisis response systems, and expand helpline and counseling availability. To improve analytical precision and long-term planning, it is also essential to enhance the quality and granularity of mental health data by incorporating indicators related to healthcare accessibility, mental health workforce capacity, stigma levels, and reporting reliability. Furthermore, future analyses should move beyond simple linear correlations by applying advanced modeling techniques such as non-linear, lagged, or causal models to better understand the complex socio-psychological and environmental factors influencing suicide trends. Finally, establishing a real-time monitoring and visualization dashboard would allow policymakers and organizations to track shifts in suicide metrics, evaluate the impact of interventions, and respond proactively to emerging risks. Together, these recommendations provide a structured pathway toward evidence-based, high-impact suicide prevention strategies that address both systemic and population-specific needs.

LIMITATIONS

1. Variability in suicide reporting across countries.
2. Economic indicators may lack nuance and require complementary social data.
3. Most analyses are descriptive; causal inference requires advanced modeling.
4. Latest-year (2022) data may be incomplete.

NEXT STEPS

- Integrate geospatial mapping and forecasting.
- Build a client-facing dashboard (Power BI / Looker / Streamlit).
- Conduct deeper modeling with additional determinants.