APRIL 2023

DATA VISUALIZATION

MINI PROJECT
SUICIDE DASHBOARD

Group Members :

Shubham Kumar

B20AI039

Tanmay

B20AI047

Deployment Link :

https://dashboard-suicide-analysis-dv.onrender.com/

GitHub Link :

https://github.com/shubhamByte/B20AI039_B20AI047_DV_project

02 INTRODUCTION

Nearly every year, suicide is one of the top 15 causes of death. It now affects 800,000 individuals worldwide yearly, or one person every 40 seconds, and reaches epidemic proportions. That's the second-leading cause of mortality for those between the ages of 15 and 24.The survival and mental health of both our present and future generations throughout the world are seriously at risk due to this.

Insufficient research is being done to prevent and treat cardiovascular disease and cancer, which are the two leading causes of mortality worldwide. Suicide is not publicly discussed as much as it should be due to the stigma associated with it in many cultures. With efforts being made to lessen suicide attempts, there has been a shift in recent years, and it is now more often discussed. Research is being done to learn what motivates someone to commit this crime and how to stop it. In particular, the World Health Organisation has conducted in-depth study on suicide prevention on a worldwide scale.

03

PROBLEM DESCRIPTION AND TASK ANALYSIS

There are several factors, not just one, that frequently cause people to take their own lives. A few of the causes are drug and alcohol addiction, stress from daily life and finances, a lack of social support, etc. We could better comprehend the scenario if we looked at certain elements. Therefore, we are interested in determining if there are any meaningful connections between certain characteristics that influence suicide rates globally.

DATA

We mainly make use of the dataset "Suicide Rates Overview 1985-2016" [6], which was compiled using information made available by The World Bank, World Health Organisation, and United Nations Development Programme. The main dataset includes statistics on suicide for 101 nations from 1985 to 2016. There are 27,820 observations with the characteristics listed in Table 1.

Feature	Value Range
Country	101 countries e.g Albania
Year	1985-2016
Gender	Male, Female
Age (in years)	5-14, 15-24, 25-34, 35-54, 55-74, 75+
Number of suicides	0 - 22338
Human Development Index	0.483 - 0.944
Gross Domestic Profit ($)	251 - 126352
Generation	Generation X, Silent, G.I., Generation, Boomers, Millenials, Generation Z

Table 1. Overview of Primary Dataset

04

We extracted additional information from the repository known as "World Development Indicators" and added it to our main dataset. We utilise a few carefully chosen elements from this source that interest us to look into the suicide data in the table below.

Feature	Value Range
Unemployment Rate	0.15 - 33.47
Life Expectancy	51.06 - 86.99
Labor Rate (Female to Male)	32.76 - 90.99
Region	7 continents

Table 2. Overview of Additional Data

The continuous variables "Unemployment Rate," "Life Expectancy," and "Labour Rate (Female to Male")" in the generated dataset have some missing data. Aggregation operations can still be applied with this data. The dataset is missing 95 nations. The data-set includes the most prominent nations, except China and India,
However, there is enough information to look at nations, their characteristics, trends, and connections to suicide rates, among other things.
The global suicide rates are aggregated by year, nation, age, and gender in the final enriched dataset.This dataset is also intriguing because of the extra characteristics, which allow us to determine whether any of them exhibit any surprising correlations with or patterns with respect to the suicide rate. Therefore, by changing the condition of those characteristics in the various nations, it may have a significant influence on avoiding suicide if we could discover such unanticipated trends. They are all included in the dataset to be examined to see whether they truly have an influence and to quantify that impact. For each of them, it is not improbable that they have an impact on the suicide rate.

05

ANALYSIS

The enhanced data in Table 2 begins in 1991, therefore the data utilised spans from 1991 to 2016. We look at any connections, trends, and correlations between the characteristics in Table 1 and the suicide rate.

We'll talk about the following issues:
Which area has seen the greatest annual suicide rate?
- Which nations in the area have the greatest suicide rate?
- Which age groups and genders in these nations have the greatest suicide rate?

Interacting with the plots, choosing data from the plots, and applying multi-select filters for age, gender, income, and region will be among the activities completed. The plots may also assess the relationship between other characteristics, such as the unemployment rate, life expectancy, and GDP in connection to the global suicide rate.

06

VISUALIZATION DESIGN

To develop the visualisation dashboard, we utilise "Dash by Plotly" (insert citation here). The questions outlined in analysis served as the foundation for our decision-making about the visual encoding and interface design methodologies. In order to find general trends, patterns, and relationships between qualities and the overall number of suicides, we employ point and line-based approaches for the visual encodings.
We employ the "Overview + Detail" method for interface design.
Navigation methods like item reduction (zoom, restricted) and attribute reduction (slice, cut) can be used to adjust the plots.

VISUAL ENCODINGS AND INTERACTION

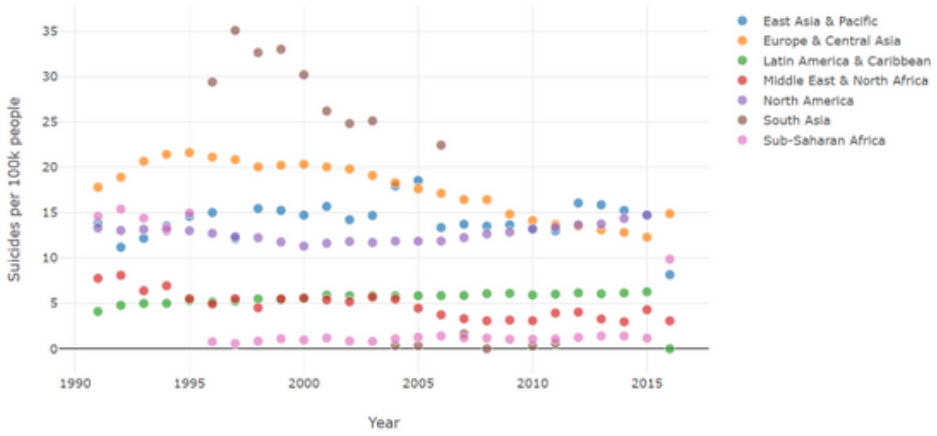
Maps and statistical distributions such the choropleth map, box-and-whisker, parallel coordinate, and scatter plot are used as a result. The graphs in this section depict the suicide rates per 100,000 persons from 1991 to 2016.
We typically employ the overview first, zoom and filter, and details on demand interaction method. When the dashboard is loaded, the overview is displayed with all filters, including all years, age groups, income groups, and geographies, chosen. Filters may be removed, plots can be zoomed into according to your needs, and click, box pick, and lasso functions are provided for plots to lookup, explore, compare, and summarise the data for all plots on the dashboard.

07

PLOTS

Scatter Plot

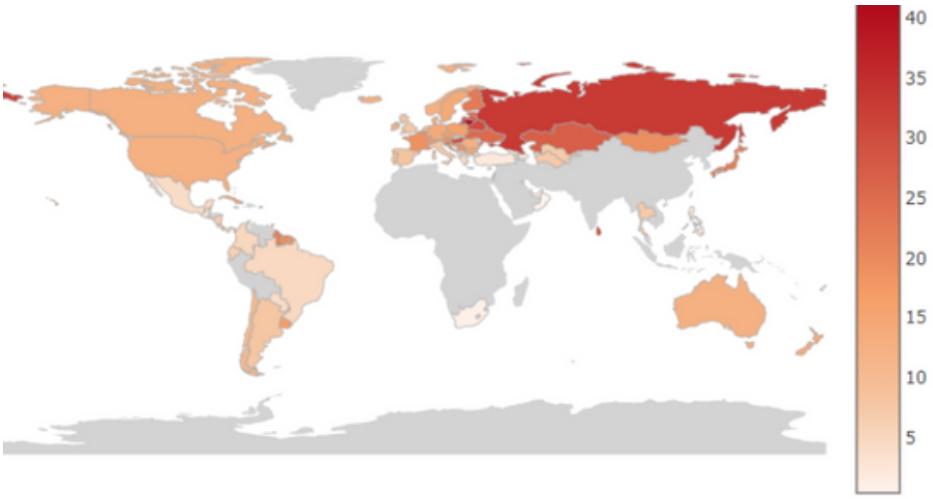
The scatter plot is used to find patterns, clusters, correlations, and extreme values for the overall number of suicides in connection to various characteristics, such as nations, regions, age, etc. It covers a wide range and is updated in response to use of the filters and/or the global map. It is also used to compare characteristics between nations, for instance, the total number of suicides for the chosen nations over time with regard to age categories.



Choropleth Plot

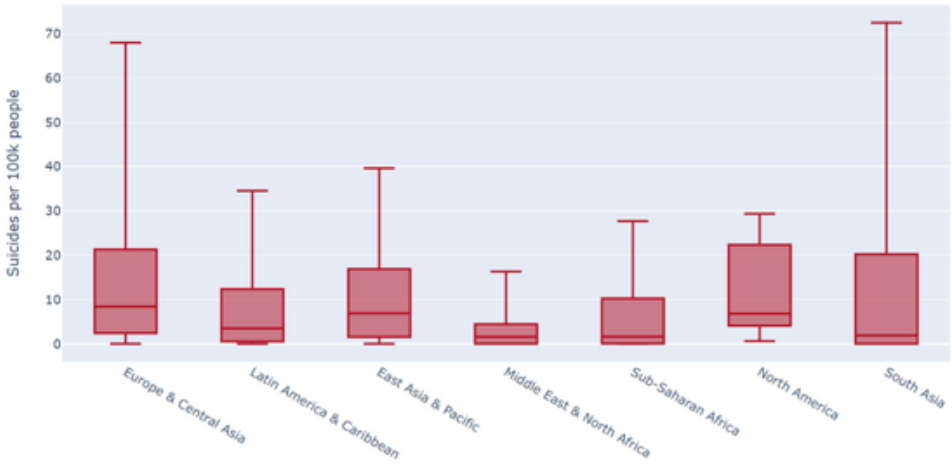
We may visualise our geographic data based on various nations and areas using choropleth maps. The intensity of the suicide rate across the chosen years is depicted in Fig. using a colour coding of the nations. Red is used to symbolise intensity through differences in luminosity, with darker reds denoting higher suicide rates.

08



Box-and-Whisker Plot

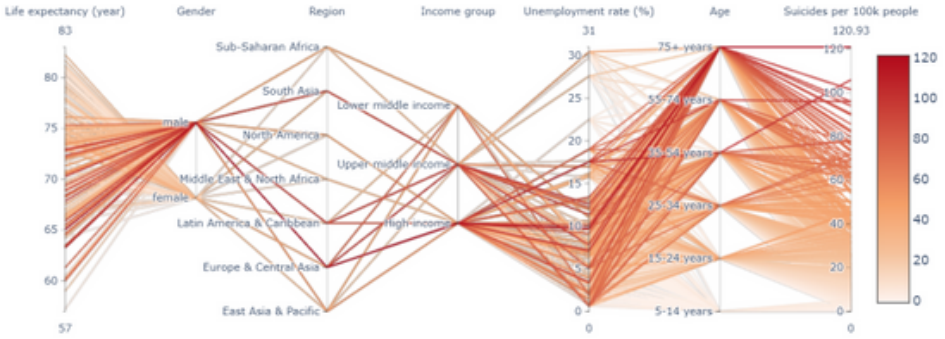
The median, quantiles, and 5% upper and lower range of the suicide rates are shown in a box plot. Understanding how the various nations and regions are distributed generally aids in identifying any extreme numbers that could be pushing the averages in a particular way. The data initially displays a box plot for the areas. The globe map may be used to interact with the story. On the globe map, when a country is selected, it will appear in a separate box so that it can be contrasted with the areas and shown how distinct it is from the region as a whole.



09

Parallel Coordinate Plot

We chose to utilise at least one graphic that could display multiple variables together while being somewhat compact because the data set contains numerous aspects pertaining to the suicide rate. We can see trends among the characteristics and how they connect to one another in terms of the suicide rate thanks to the parallel coordinate plot.



USE CASES

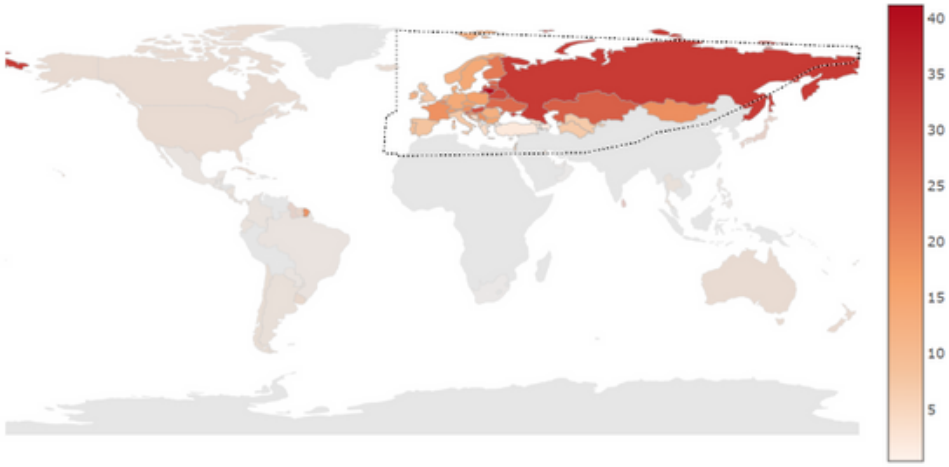
Which region has had the highest suicide rate over the years?

Europe consistently seems to have the highest suicide rate apart from a few years. This is further affirmed by looking at the world map, Lithuania, Russia, and the surrounding countries seem to have the highest suicide rate over the past 25 years. These countries belong to the biggest region, namely: “Europe & Central Asia”.

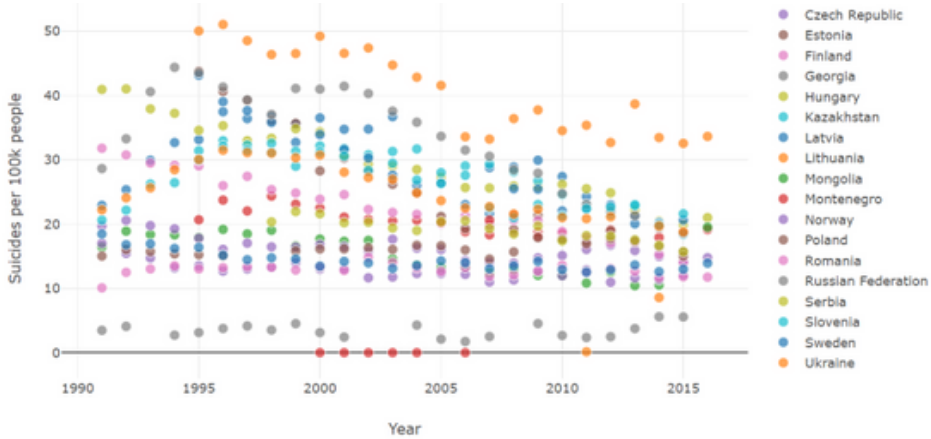
10

Which countries have the highest suicide rate within this region?

We find this out by selecting the 'Europe & Central Asia' region on the world map using box select or lasso



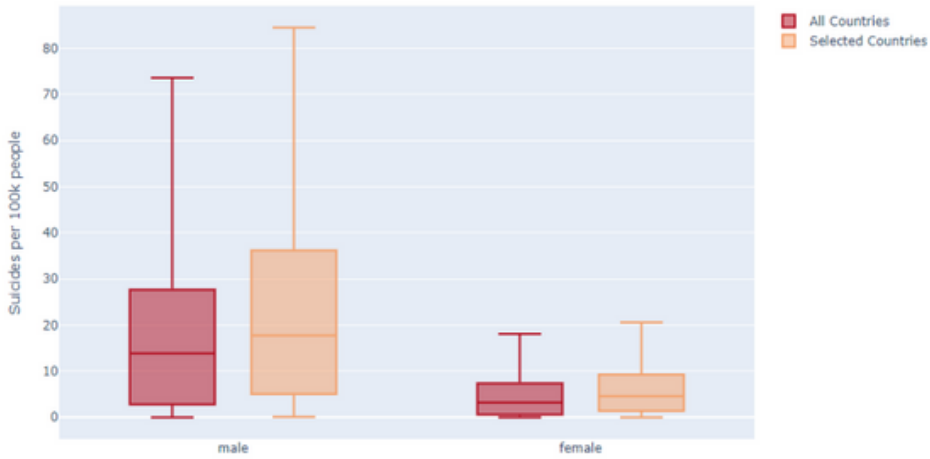
We obtain the scatter plot (after updating) by selecting the group of countries on the world map. This plot can be used to compare the countries within the region to identify their suicide trend over the past 25 years. This shows over the years that the suicide rate is steady or mostly decreasing with Lithuania and Russia contributing to the highest values. While Georgia and Bosnia contribute to the lowest rates, potentially because we have less data available on them



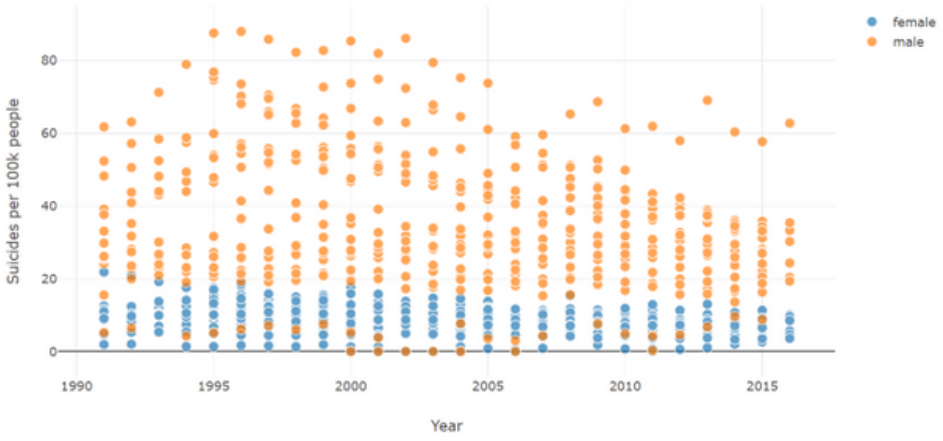
11

Which age groups and gender contribute to the highest suicide rate within these countries

We find this out by selecting the 'Europe & Central Asia' region on the world map using box select or lasso



We see that men in the European region dominate the suicide rates as opposed to women. Women seem to nearly have one-third the suicide rate of men approximately over the years.



12

CONCLUSION

The significant contributions were in determining the suicide rate's trajectory during the previous 25 years. Over this time, we observed a continuous upward trend in suicide rates in the European area. Over the years, Sub-Saharan Africa and the Middle East region had the lowest rate.
Additionally, we observed that, on average, during this time, Russia and Lithuania had the highest suicide rates; nevertheless, the trend showed that the suicide rates were declining year over year. These and other European nations also revealed that the majority of suicides were committed by persons in their 75s and between the ages of 35 and 54. Because some nations with greater unemployment rates exhibited reduced suicide rates while others did not, we were unable to identify a meaningful correlation between the unemployment rate and the suicide rate. This was not definitive. Similarly, no notable trends were found for life expectancy.
We will be better able to comprehend the historical context of suicides throughout the world and the amount of effort that needs to be done to assist organisations and society get ready to prevent suicides thanks to this study and visualisation.