Do Suicide Rates Differ Between Men and Women Around the World?

Introduction

Suicide has been a serious public health concern for the past decade. From 2000 to 2018, suicide rates increased by a whopping 37% (1). There are, on average, about 132 suicides per day in the U.S (2). The total deaths in 2021 caused by suicide was 48,183 which is approximately 1 death by suicide every 11 minutes (1). In 2021, the suicide rate after adjusting for age was 14.04 per 100,000 individuals. In the same year, men were 3.9 times more likely to die from suicide than women (2). People of all ages can be affected by suicides with 10-14 and 20-34 being the most prone age groups to suicide (1).

Suicide ideation occurs when a person starts to develop thoughts about ending their life to escape from the emotional pain that could be caused by their mental health or past traumatic events. These negative thoughts become hard to ignore for the individual which can make it hard for them to deal with. According to the American National studies, many individuals passively go through suicide ideation but do not plan or act upon it, but about 3.3% of Americans “actively consider suicide every year”. On top of this, about 20% of individuals with suicide ideation attempt suicide, 34% go on to develop a plan for suicide, and 72% of those who make a plan attempts suicide (3).

Many different groups of people are more prone to suicidal behavior like marginalized groups. Suicide rates can also differ based on age, race, economic status, and many more factors. One factor that people have focused on is sex as there seems to be a higher rate of suicide in men than in women. This phenomenon is also known as the "gender paradox". It is believed that men have a higher risk of suicide due to gender roles imposed from society, substance abuse, financial burdens, and many more reasons (3). Additionally, men tend to use more lethal methods of suicide, have higher impulsivity, and higher rates of substance abuse. Therefore, the higher suicide rates in men seem to be due to psychological differences rather than biological differences from women.

Suicide is a serious problem that affects many individuals around the world. It not only affects the individual themselves, but also the family and friends that care for them. Therefore, it is imperative to understand suicide patterns and causal reasons to tackle this problem. The gender paradox is a phenomenon that is greatly researched in the U.S., but not as researched in other countries. With more information about this phenomenon, more initiatives to help those with suicidal thoughts can be made and hopefully reduce the suicidal rates and gender paradox around the world.

The main question I want to address with this data set is whether men have a higher suicide rate than women in different regions of the world as seen in the U.S. I hypothesize that men will have a higher rate of suicide than women in every parent location.

Methods

The data set containing suicide rates was taken from the World Health Organization and downloaded as a CSV file. The suicide data was collected and calculated from the WHO Global Health Estimates under the WHO Mortality database. The data for suicide rates were collected from vital registration data from statistics reported to WHO annually. The data was loaded into R

using the "data.table" library. The original data set had 10980 rows/observations and 34 columns/variables. The variables are either integer, character, or numeric types. The columns with only "NA" values were removed, so the total number of variables after removal was 20. There are 6 different areas/parent locations included in the World Health Organization's suicide rate data set: the Americas, Eastern Mediterranean, Western Pacific, South-East Asia, Europe, and Africa. From these parent locations, a total of 183 cities/locations were included. The suicide rates were taken from 2000 to 2019. There were no missing values for the key variables (rate, sex, confidence intervals). There were some outlier suicide rates but they were kept in because those values were still plausible. The suicide rate values refer to suicide incidences per 100,000 people. Negative values of suicide rates that could not have been interpretable were checked and none were found. The names of columns were changed to be more descriptive of the variable.

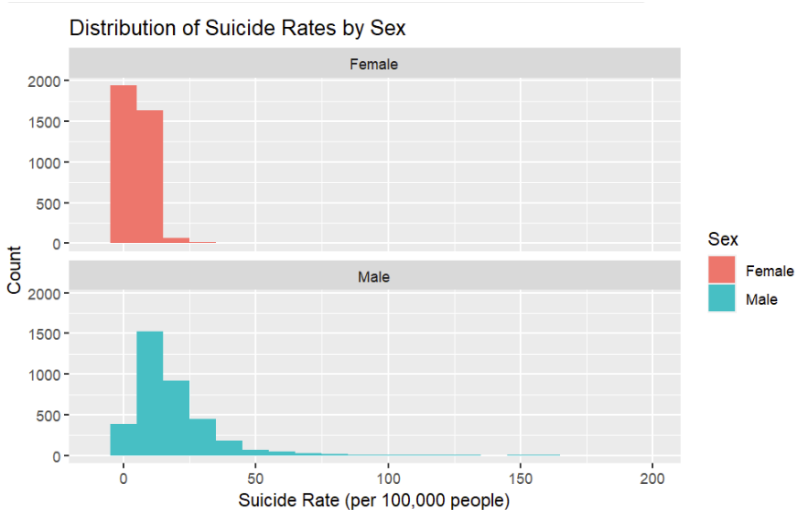
The data set containing the populations of each parent location was exported from the United Nations Population Division Data Portal as a CSV file. The global data was collected from the UN's Populations Division. The population data was collected from a census of the country. The census may be from usual residents of the country or the total number of people present in the country when the census was collected. The portal allows the user to choose the type of value (population, mortality, fertility, etc.), time period, and countries to include in the exported data set. The data set used in this project included the total population of both sexes, the time period from 2000 to 2019, and the countries included in each parent location.

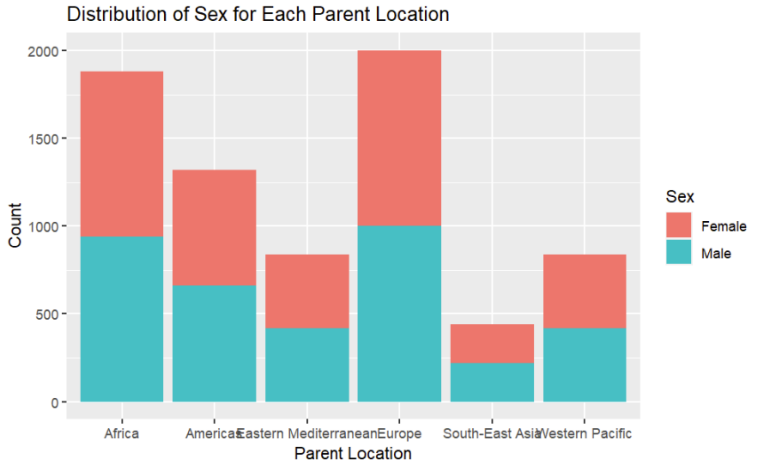
The CSV files were loaded into R using the "data.table" library. The population data set for each parent location varied in the number of rows depending on how many countries were in each parent location. However, all data sets contained 27 columns which included variables that were either character or integer variables. Many columns in the data set were unnecessary to the project such as "IndicatorId" and "IndicatorName" and were removed in the next step. Since the original population data set contained the population of each country under the parent location, a new data set was created with the sum of the populations of each country for each year to obtain the population of the parent location. A new column was created in the new data set called "ParentLocation" to label which location the population referred to. The new data set for each parent location After obtaining the population of each parent location for each year between 2000 to 2019, the data sets were combined together using the "rbind" function. The combined population data set was then merged with the suicide rates by parent location and year.

Data exploration and statistical analysis was carried out using the "ggplot2", "dplyr", and "plotly" packages.

Results

Exploratory Plots





The histograms of all suicide rates for each sex in the data set shows a positively skewed distribution. The suicide rates for females peaks at the 0-10 range. The highest suicide rate for females is around 50. For males, the histogram peaks at around 20-30 incidences per 100,000 people. The highest suicide rate that can be seen on the graph is at around 100. The bar plot shows a relatively even proportion of females to males in the data set for each parent location.

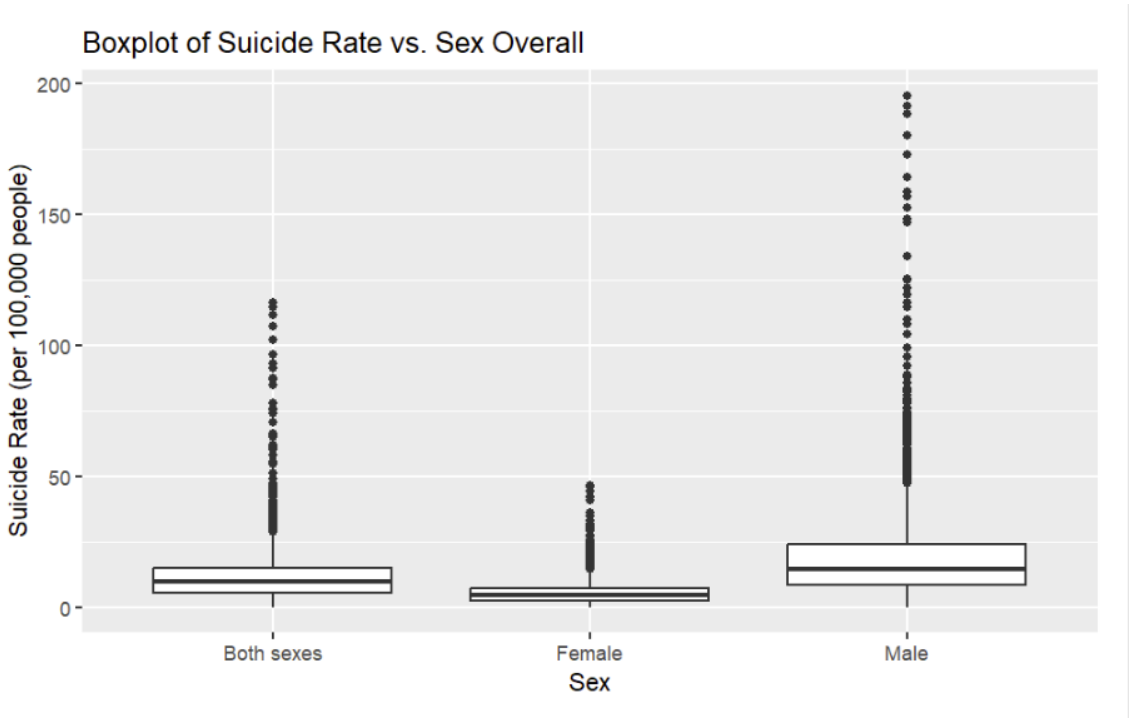
Summary Statistics

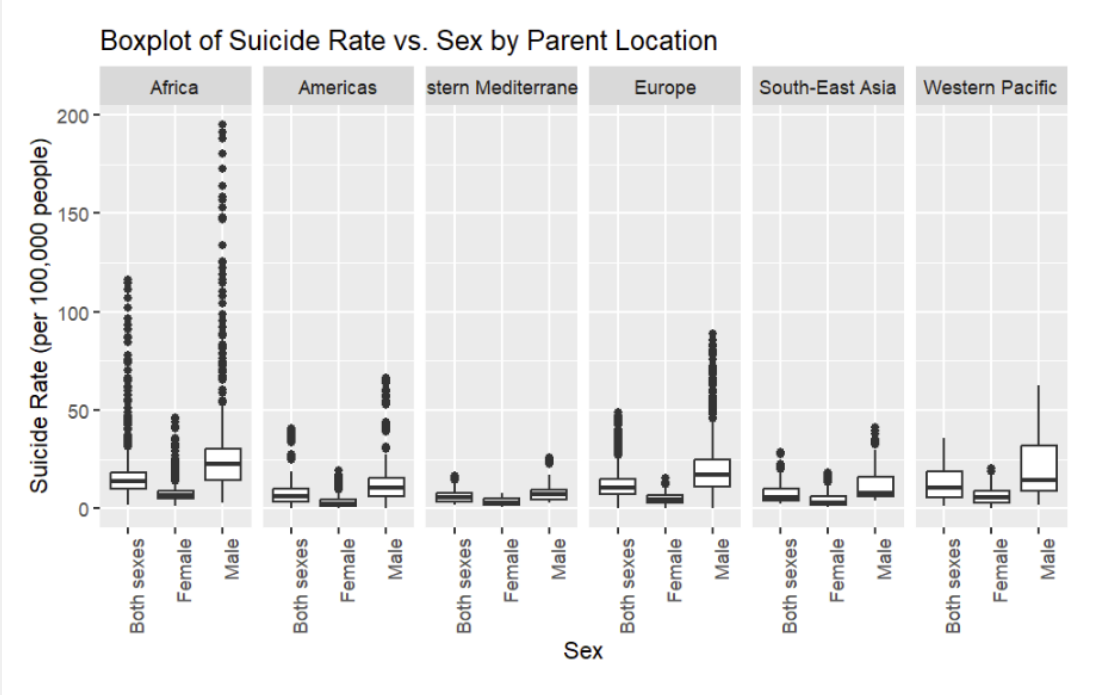
Sex	**Frequency**	**Mean (95% CI)**	**Median**	**Standard Deviation**	**Minimum**	**Maximum**
Both	3660	11.74 (8.03, 17.09)	9.76	9.85	0	116.20
Female	3660	5.40 (3.62, 8.17)	4.75	4.20	0	46.34
Male	3660	18.74 (12.87, 26.91)	14.48	17.00	0	195.2

Parent Location	Sex	Frequency	Mean (95% CI)	Median	Standard Deviation	Min	Max
Africa	Both	940	16.64 (9.44, 26.73)	13.81	13.39	1.99	116.2
	Female	940	7.88 (4.37, 13.28)	6.78	5.56	1.06	46.34
	Male	940	26.90 (15.36, 42.51)	22.43	23.15	3.05	195.2
Americas	Both	660	8.12 (6.19, 11.77)	6.42	7.30	0	40.85
	Female	660	3.50 (2.58, 5.35)	2.32	3.41	0	19.47
	Male	660	13.06 (10.04, 18.56)	10.39	11.62	0	66.42
Eastern Mediterra-nean	Both	420	6.24 (3.51, 10.24)	5.58	3.39	1.75	16.83
	Female	420	3.53 (2.07, 5.69)	2.93	2.13	0.67	8.02
	Male	420	8.55 (4.68, 14.21)	7.41	5.07	2.78	26.04
Europe	Both	1000	12.30 (10.36, 15.02)	10.77	7.91	0.33	48.89
	Female	1000	5.01 (4.10, 6.33)	4.65	2.60	0.37	15.31
	Male	1000	20.36 (17.27, 24.6)	16.91	14.58	0.26	88.81
South-Eas t Asia	Both	220	7.92 (5.25, 11.80)	5.83	5.22	2.55	28.58
	Female	220	4.48 (3.13, 6.65)	2.97	3.79	0.84	18.37
	Male	220	11.59 (7.50, 17.30)	8.07	7.73	3.93	41.46
Western Pacific	Both	420	12.68 (8.21, 18.41)	10.36	8.48	1.14	35.64
	Female	420	6.09 (4.25, 8.84)	5.58	3.72	0	20.2
	Male	420	19.50 (12.29, 28.30)	14.32	14.39	1.7	62.59

From the summary table of the overall sex-specific suicide rates, it is apparent that men have a higher suicide rate per 100,000 people than females. The 95% confidence interval for mean suicide rate for males and females do not overlap. This suggests a statistically significant difference in suicide rates between sex. When looking at the suicide rates separately in each parent location, the same trend can be seen where the males have higher suicides rates. In all parent locations except the Eastern Mediterranean, the average 95% confidence interval of suicide rates between men and women do not overlap suggesting a statistically significant difference in mean suicide rates.

Box Plots of Suicide Rates vs. Sex



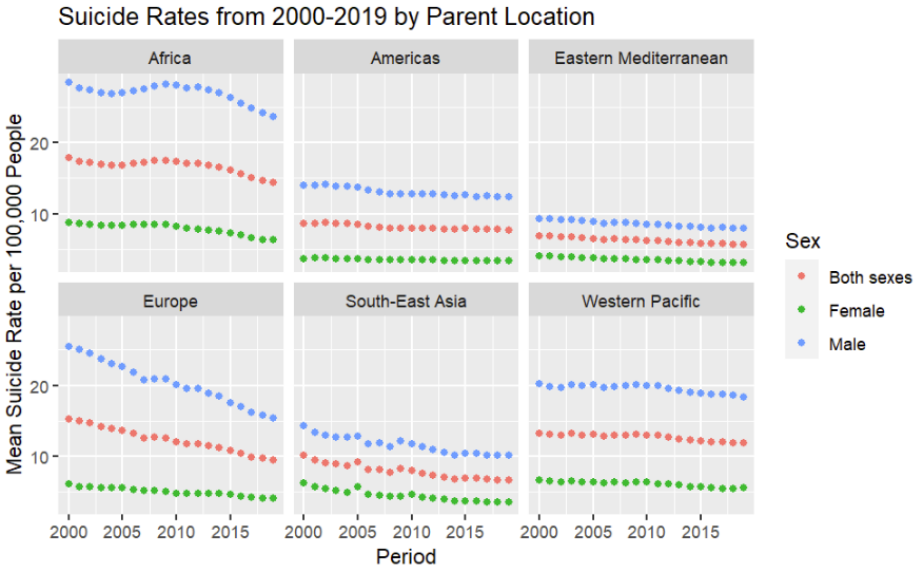


The difference in suicide rates between sex can also be seen with the box plots. The interquartile range for males and females do not overlap. The males have a lot of outliers.

When the boxplots are grouped by parent location, a similar pattern can be seen. In all 6 parent locations, the median suicide rate for males is higher than females. The boxplots also show that the men in Africa have the highest median suicide rate than any other area, but also has the most outliers in their data set. The people in Eastern Mediterranean have the lowest

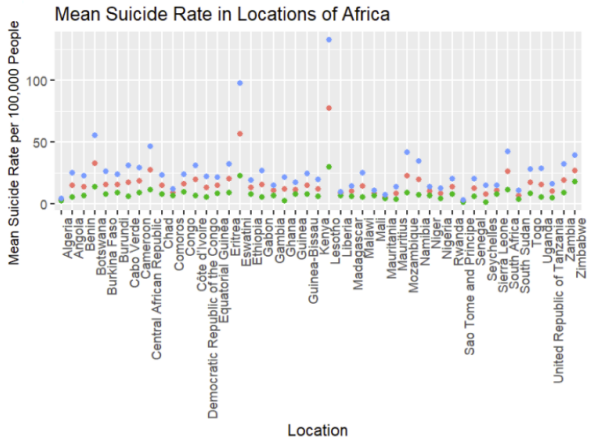
suicide rates and the narrowest quartiles which means that the suicide rates do not vary that much from each person.

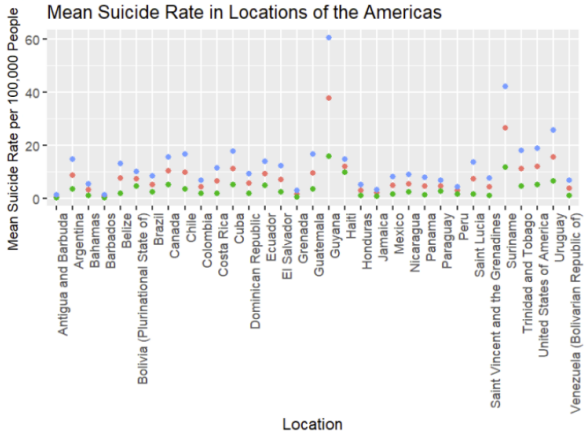
Suicide Trend from 2000-2019 by Parent Location

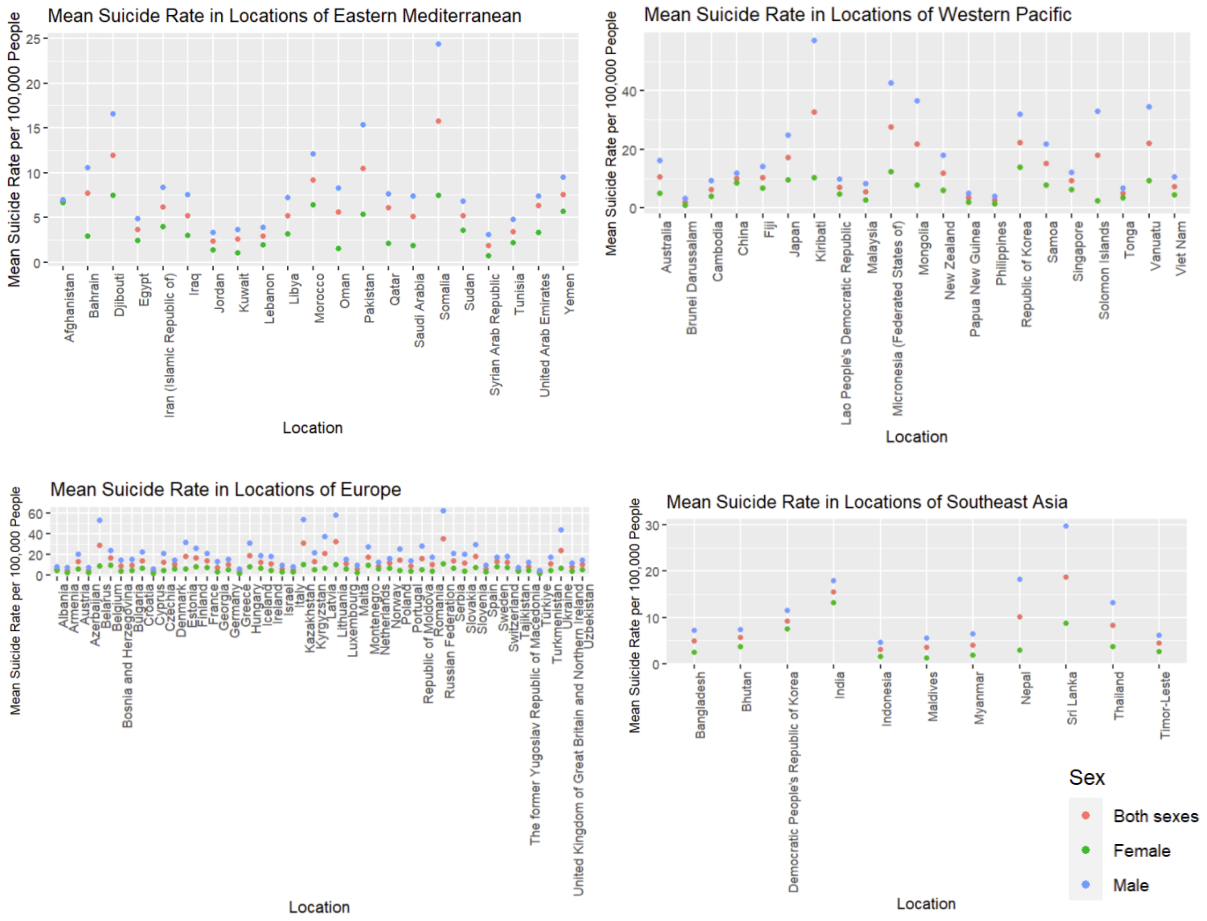


The scatter plot of mean suicide rates over the period of 2000-2019 illustrates an overall decreasing trend in suicide rates. The suicide rates in Europe have the steepest negative slope. The slope of the lines in the Americas seem to have little to no decrease over the years. The scatter plot also portrays that men have the highest suicide rates each year even when the rates are decreasing.

Suicide Rates Between Sex for Each Location Under Parent Location







The scatter plots of mean suicide rates between sex based on locations within the parent location also show similar results as the other graphs. In every location within each parent location, the average male suicide rate is higher than the average female suicide rate. The location with the highest male suicide rate is in Kenya.'

Conclusion and Summary

Overall, the data emphasizes how males have higher suicide rates than women all around the world. This pattern is in tangent with the literature about the disparity in suicide rates between men and women, the gender paradox. However, the data does not show what exactly is causing males to have higher suicide rates than females.

Future investigations can look at what specific reason or reasons causes the disparity in suicide rate. Figuring out the reasons can allow people to make an effort to address what is causing men to be at higher risk of suicide and hopefully decrease suicide rates. Future investigations could also analyze why suicide rates differ between parent locations. For example, the trend that Africa has the highest suicide rates can be looked into further and why there was a slight increase in suicide rates between 2005 and 2010. Cultural differences between parent locations can play a role in why suicide rates differ and could give insight to how people can help decrease suicide rates based on each parent location's differences.

References

1. https://www.cdc.gov/suicide/suicide-data-statistics.html
2. https://afsp.org/suicide-statistics/#:~:text=Additional%20facts%20about%20suicide%20in,are%20132%20suicides%20per%20day.
3. https://journals.sagepub.com/doi/full/10.1177/15579883221123853