

SUICIDE DATA ANALYSIS PROJECT REPORT

Introduction

Close to 800 000 people die due to suicide every year, which is one person every 40 seconds. Suicide is a global phenomenon and occurs throughout the lifespan. Effective and evidence-based interventions can be implemented at population, sub-population and individual levels to prevent suicide and suicide attempts. There are indications that for each adult who died by suicide there may have been more than 20 others attempting suicide.

Suicide is a global phenomenon; in fact, 79% of suicides occurred in low- and middle-income countries in 2016. Suicide accounted for 1.4% of all deaths worldwide, making it the 18th leading cause of death in 2016.

In 2017, a man whom I considered one of the greatest musicians of all time and author, or co-author, of a number of songs that had become the soundtrack of my life, died in suicide. Many were shocked, and many miss him today. The same tragic story repeated with three of these heroes—Kurt Cobain, Robin Williams, and Anthony Bourdain.

The Problem

WHO has created this data over the years to see any kind of pattern among people who commit suicides. WHO runs various campaigns and programs across different countries. It needs to know where campaigns are more needed. If it doesn't know about where, which age group, to focus then WHO will allocate its resources equally everywhere. This will not be efficient allocation of resources. So, The main goal of this analysis is to give clear understanding of patterns in suicide data over the years to help WHO focus on it accordingly.

About Data

The Data that i am going to use is available on WHO's website. Looking at the data, you can observe long-time trends and differences between countries, as well as within countries across a few demographic groups - in both cases you will see that these differences may be very large.

Data has 12 columns. Following are those 12 columns.

1. country
2. year - Christian era.
3. age - The age is summarized within a certain range. 5-14 years 15-24 years 25-34 years 35-54 years 55-74 years 75+ years
4. Sex
5. Population of the country at that time.
6. suicides_no - Number of suicides.
7. suicides/100k pop - Suicide by population of 100,000 people.
8. country-year - Combination of country and year.
9. HDI for year - The Human Development Index (HDI) is a simple indicator for measuring the average degree of accomplishment in a country regarding three aspects of human development: health, education, and income.
10. generation - Generations in America.

G.I. Generation - 1900 - 1924

Silent - 1925 - 1942

Boomers - 1946 - 1964

Generation X - 1965 - 1980

Millennials - 1981 - 2000

Generation Z - 2001 - 2010

11. GDP of the year

12. GDP per capita

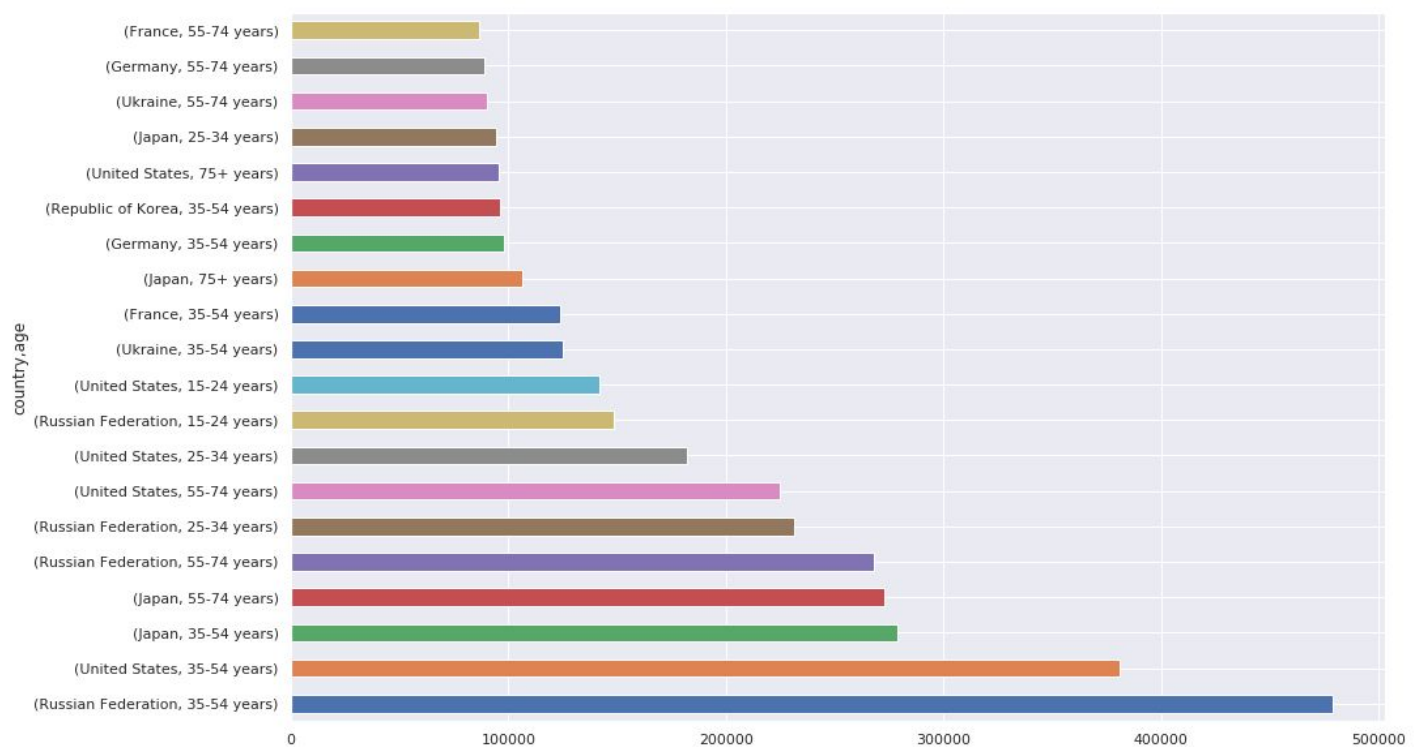
Methodology

1. Cleaning Data

There is a space in the name of column, which makes it difficult to handle. Change the name of the column to make it easier. GDP of year is string. Let's change it to int. The data is mostly clean.

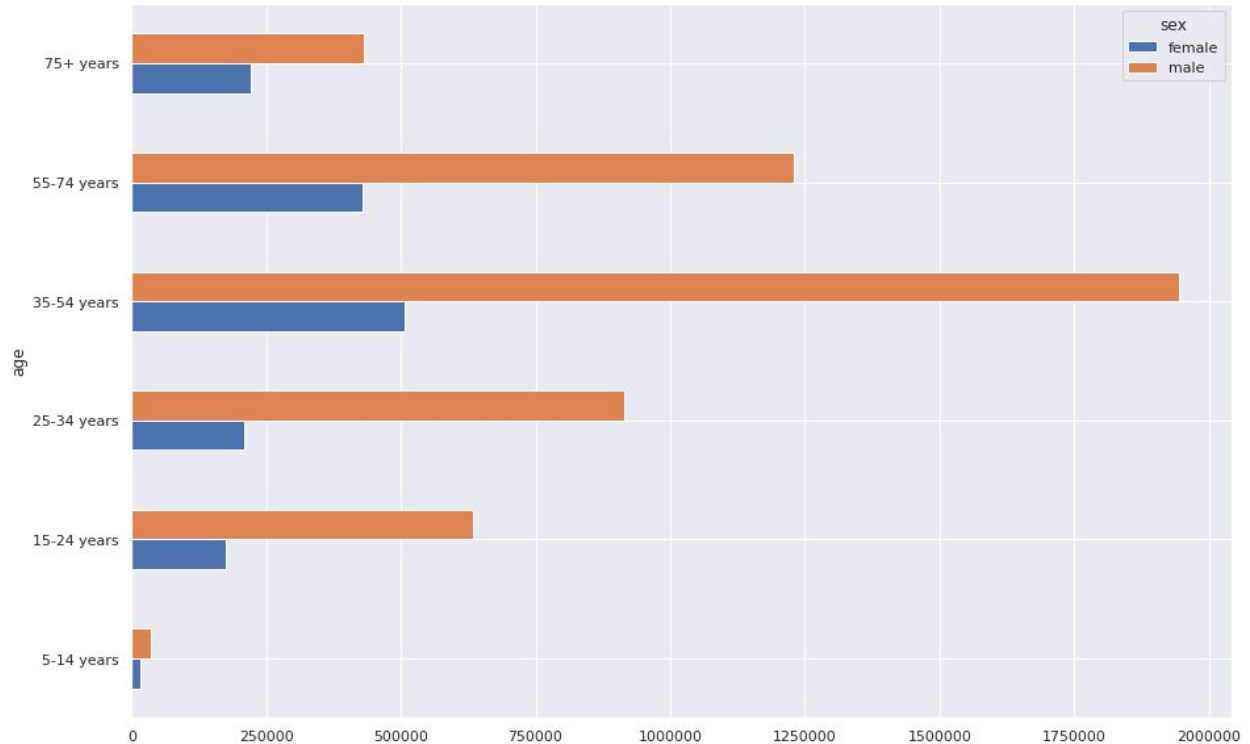
2. Exploratory Analysis

1. Country, Age Group vs Number of suicides



This graph can contain every age group for every country. This helps us find that Russian Federation has maximum cases of suicides from age group 35-54 over the years.

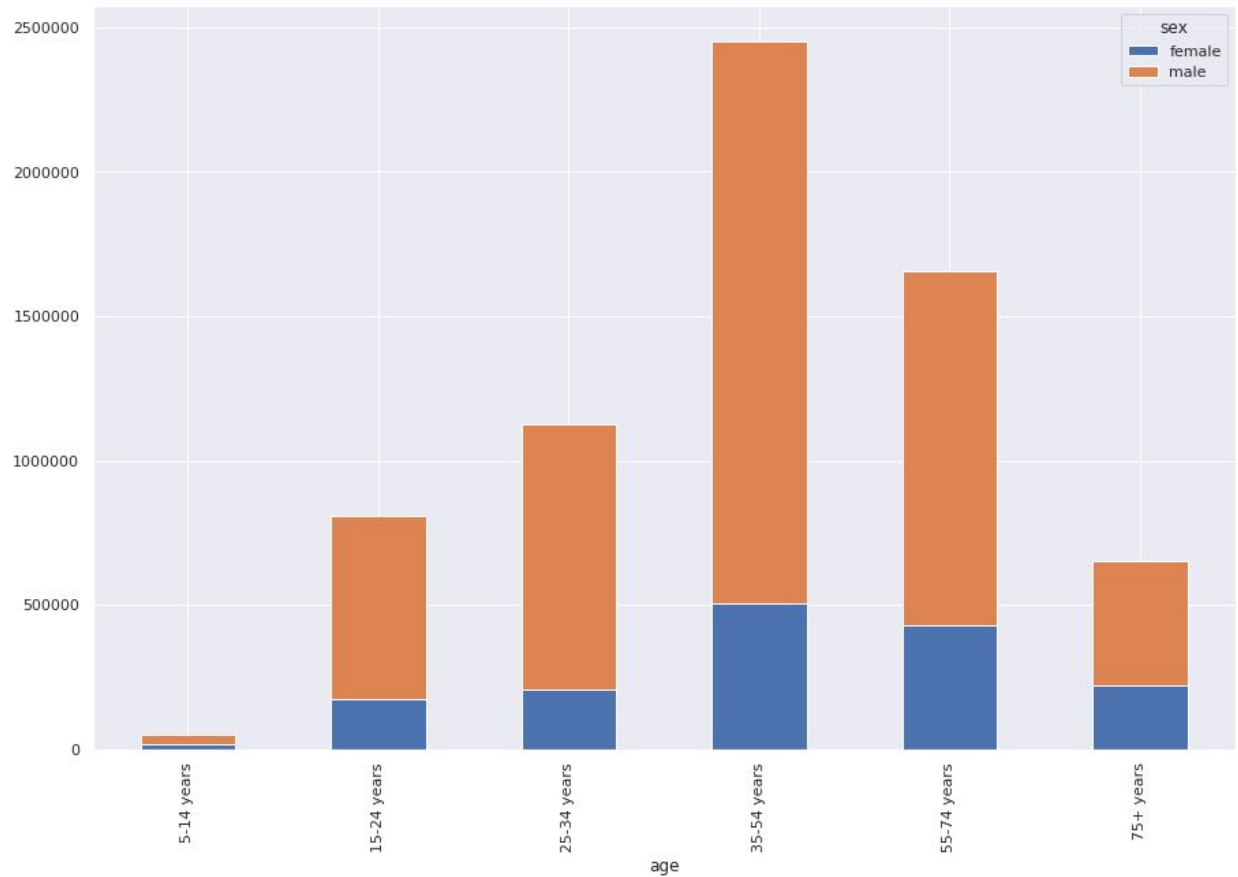
2. Age groups, sex vs number of suicides



It is evident that people in their 30-50s are most stressed. As the graph suggests, The people from the age group 35-54 are more prone to suicides especially males.

3. Age groups, sex vs number of suicide cases (stacked view)

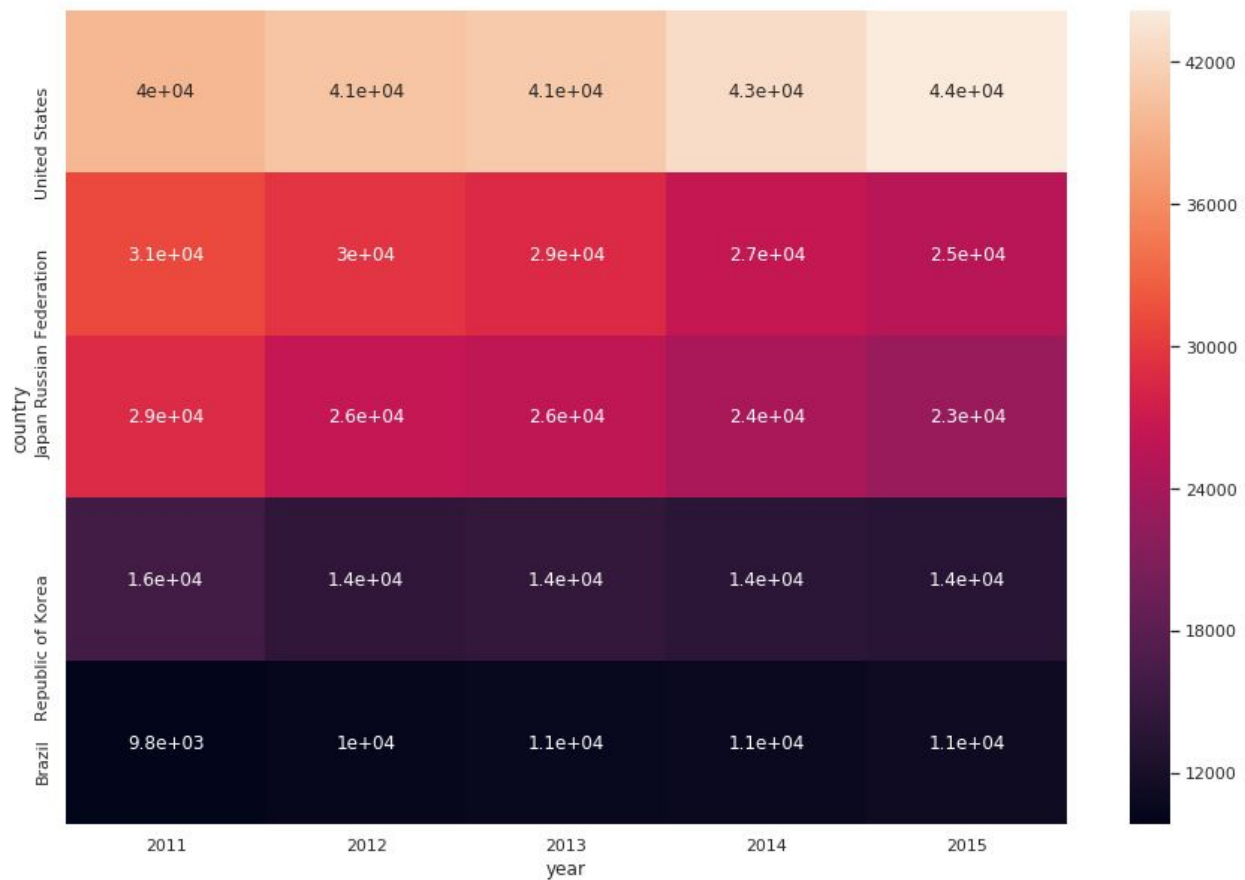
Here is stacked Bar Graph of number of suicide cases based on age groups. The graph suggests that males of age group 35-50 are more prone to suicides.



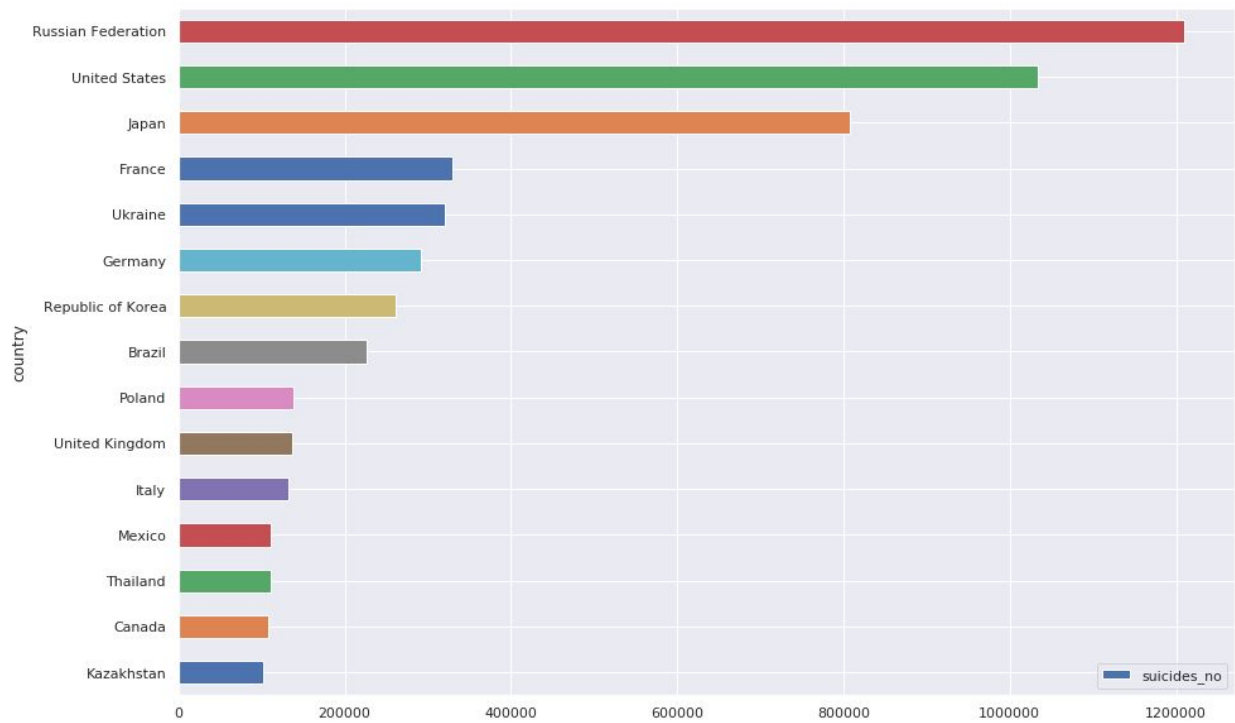
4. Top countries in recent years based on suicidal cases.

The graph below is the heatmap of suicidal cases in top 5 countries in recent years[2011, 2012, 2013, 2014, 2015]

United states has most cases of suicides in recent years and it is increasing every year. Rest of the top countries has decreasing number of suicides over these years which is a good thing. WHO needs to give special attention to its campaigns in USA.

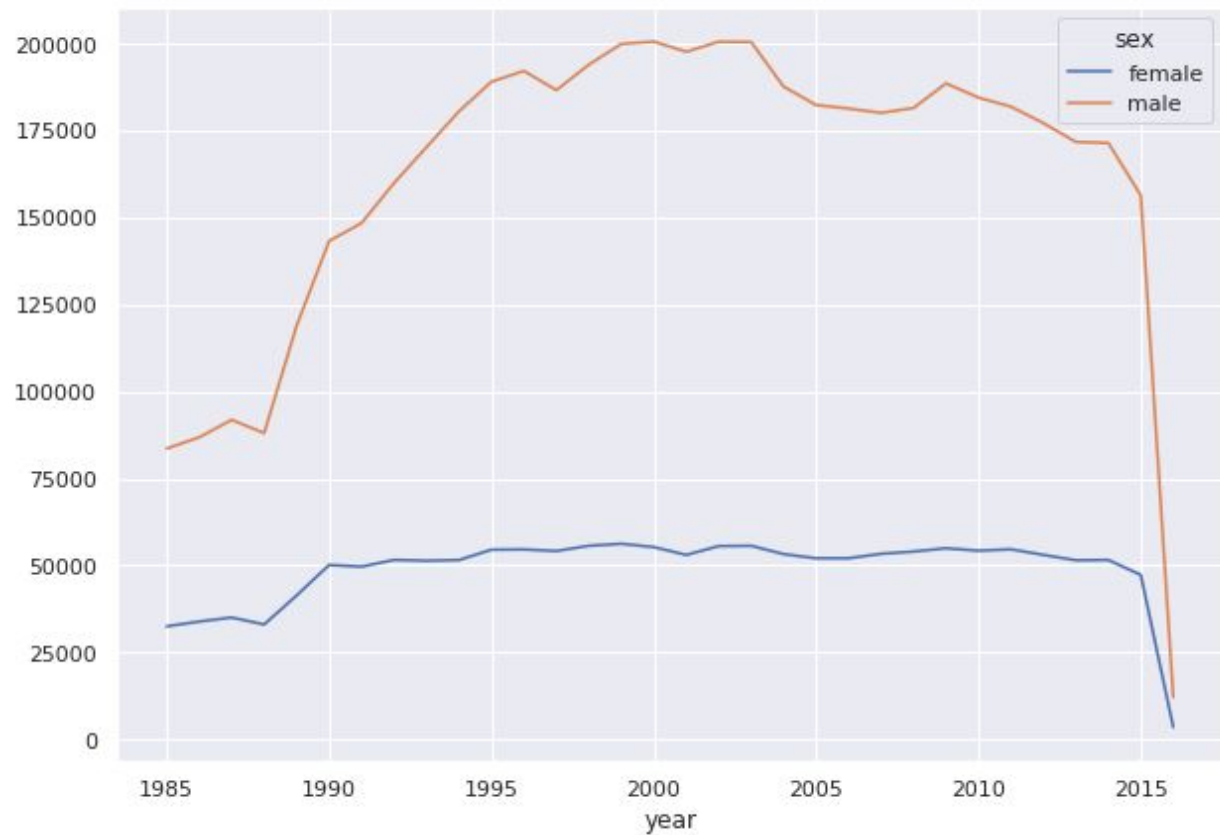


5. Countries vs Number of suicides



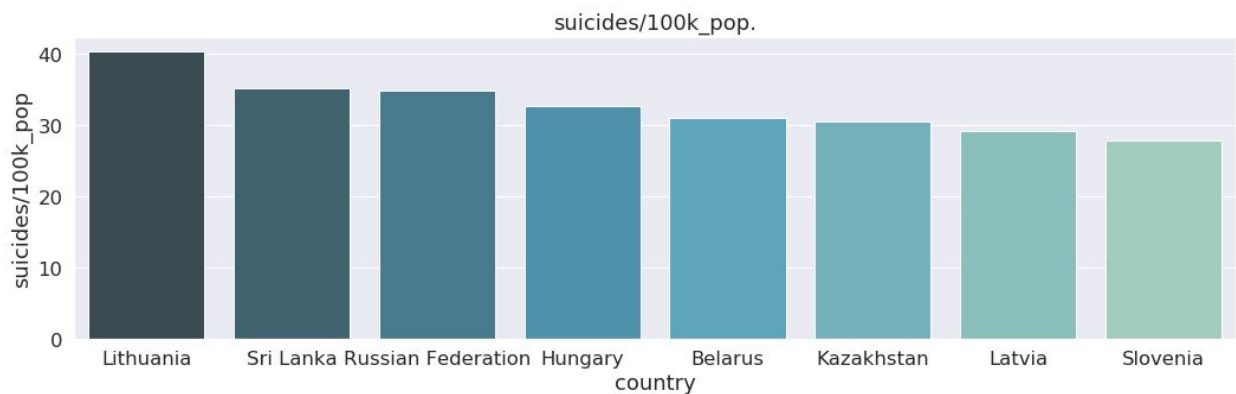
Russian Federation has maximum cases of suicides followed by USA. But from the previous graph we can say that USA has more suicides than Russian federation in recent years.

6. Male suicides vs Female suicides



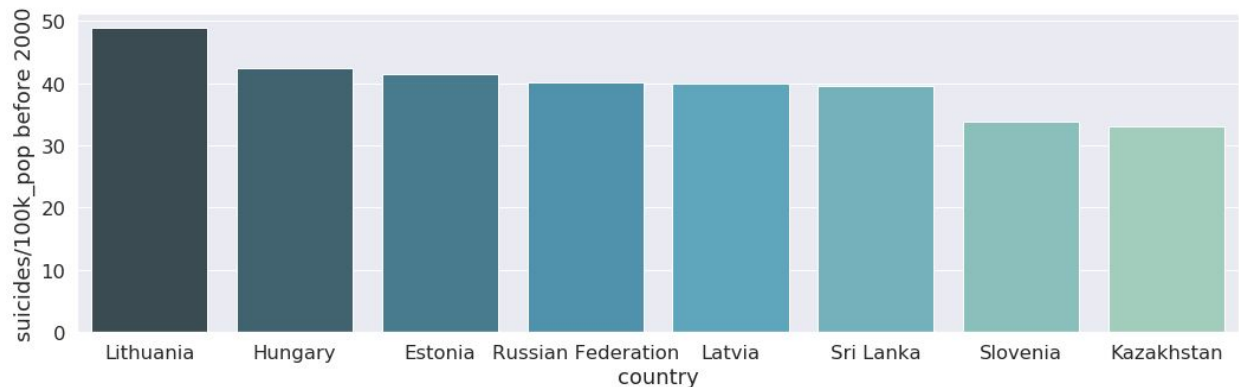
Number of suicides in females has been constant throughout the time period, whereas males have increasing number of cases.

7. suicides / 100k population vs country



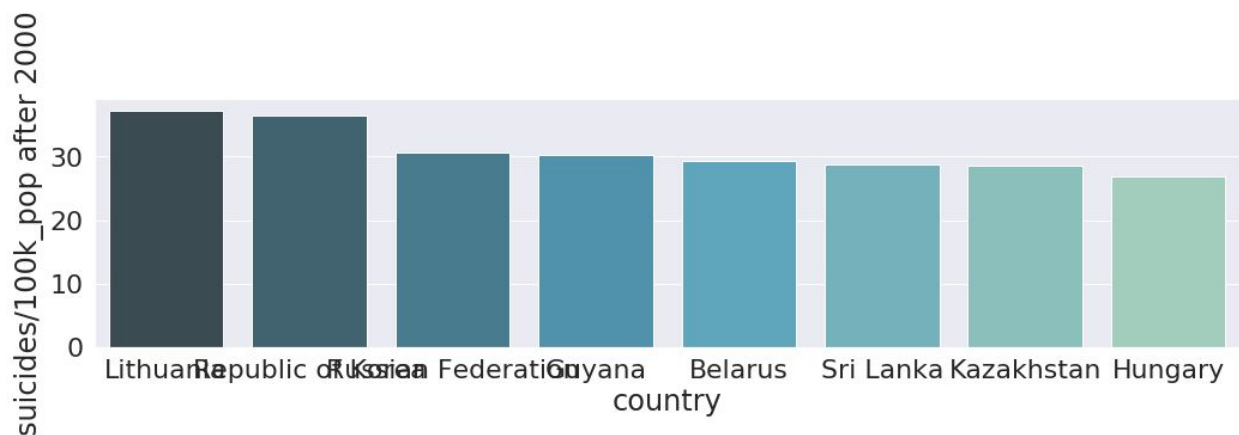
Lithuania has maximum suicides per 100k people followed by Sri Lanka. Most of the countries in the above graph has lesser population but they have more cases of suicides per 100k people. This is a key insight. These countries are not in top countries based on total suicides but the percentage of people committing suicides is higher.

8. suicides/100k population before 2000 vs Countries

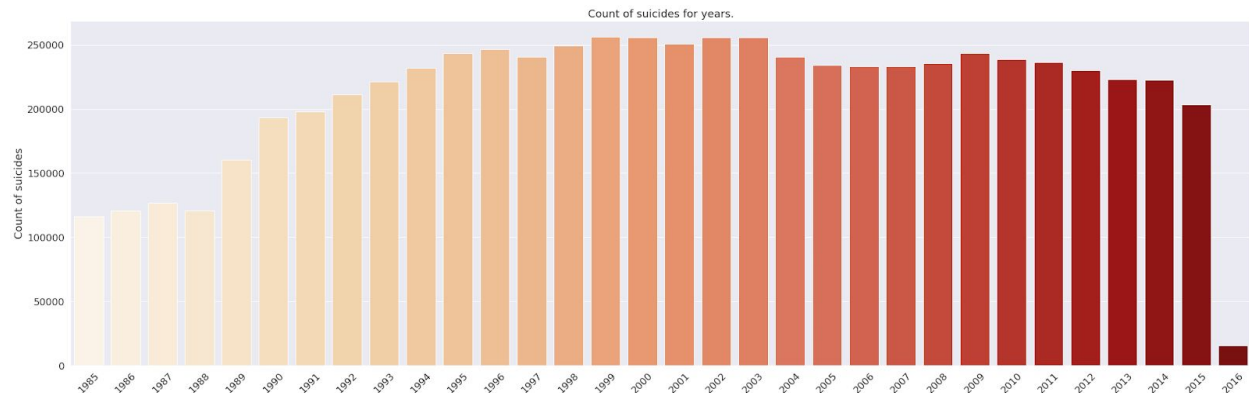


Lithuania was still at top before 2000.

9. suicides/100k population after 2000 vs Countries

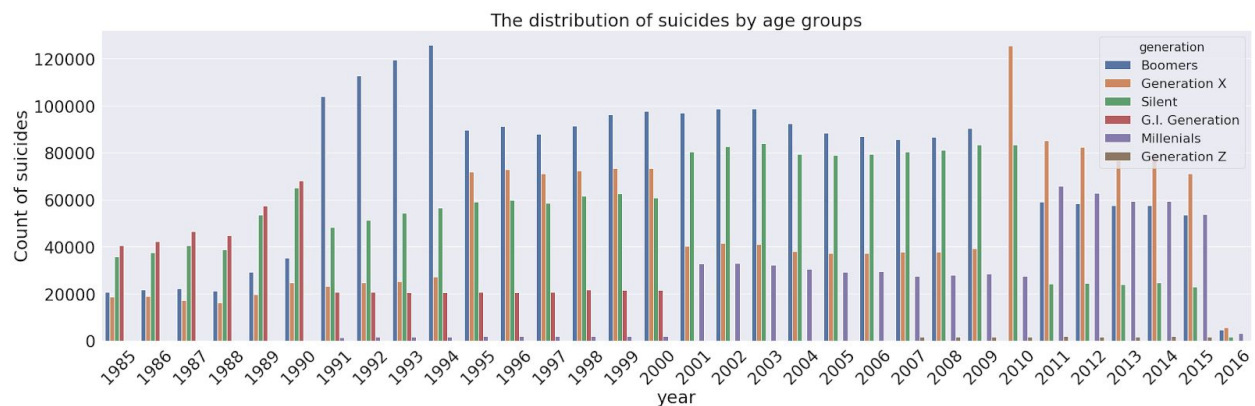


10. Total suicides in all countries over years



Suicides have been increasing over the years and has been constant for most of the years.

11. Most vulnerable age groups over the years



This graph serves a lot of insights. We can see that from 1985 to 1990 people from silent and G.I. generation had most suicide cases. From 1991 to 2009 Boomers were suiciding most. Now, In the recent years generation X has most cases of suicides.

Results

- Russian Federation has most cases of suicide over the years.
- Age Group of 35-54 has most number of suicide cases.
- USA has increasing number of cases in past few years. Also, USA has maximum number of suicide cases in these years. Rest of the countries have decreasing number of cases.
- Suicide cases among males have been increasing over the period. Female cases have been constant.
- When it comes to percentage of people committing suicide, Lithuania has maximum percentage followed by sri lanka.
- Number of suicide cases across all countries had increased in start, remained constant in the middle and slightly decreased in past few years.
- From 1985 to 1990 people from silent and G.I. generation had most suicide cases. From 1991 to 2009 Boomers were suiciding most. Now, In the recent years generation X has most cases of suicides.
- Males are more likely to commit suicide than females.

Discussion

The most prone age group for suicidal activities is 35-54 years age group. WHO should give special attention to people from that age group. Russian Federation has most cases of suicides over the entire period but US has experienced increasing number of suicide cases in the recent years. Although Russian Federation and US has most number of cases in these years. Lithuania and Sri Lanka has maximum percentage of people committing suicides. Males have more cases of suicide than females. In recent years, Boomers have been committing more suicides than any other age groups.

Conclusion

The insights gained from this analysis can be used by WHO to create their campaigns and use their resources more efficiently. WHO can focus on countries which have increasing trends of suicides over the past few years. It also needs to focus on 35-54 age group.

The insights gained from tracking everyday activities, jobs, family, friends would be more powerful in predicting whether a person is trying to suicide in recent future or not. But this analysis provided lots of insights about the age group, countries, sex, etc of suicidal people.