Challenge: Storytelling With Data & Sustainable Development Goals

ADS-B

Kristina Krasteva | 3472892 | 13-05-2020

Table of Contents

INTRODUCTION

The purpose of this document is to provide information about the personal challenge "Storytelling with data & SDG" of ADS-B. Below, the datasets used, the target audience, the goal of the challenge and the whole process of creation of the visualizations can be found.

For the challenge, I wanted to create visualizations that have some kind of relation with my team's SDG (Goal 3 – Health life & well-being). When I think of a healthy lifestyle, I most of the time consider only my physical health, but we are more than this. Our mental health is equally as important and even more sometimes and paying attention to it is something that I believe we tend to struggle with.

Mental health and wellbeing were defined as a considerable part of health by the WHO since 1978. Nevertheless, it was only in the last few years that mental health has been included on the unified global agenda. In the Sustainable Development Goals (SDGs) in 2015, it was also mentioned as a target to prioritize "prevention and treatment of non-communicable diseases, including behavioral development and neurological disorders, which constitute a major challenge to sustainable development". Thus, I decided that I will be in the hunt for a dataset which will provide me with the opportunity to dig a bit deeper in that area.

My dashboard can be found on: https://krisxtinaa.github.io/KristinaPersonalChallenge/

DATASETS

The dataset that I have found and chosen for the challenge is from Kaggle and is called **"Suicide Rates Overview 1985 to 2016"**. (https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016)

In it there is the comparison between socio-economic information (e.g. GDP per country, HDI for year) and the suicide rate for each country per year. I was really stuck by the fact that "Suicide remains the second-highest cause of death among people aged 15 to 29 globally, with 79 per cent of suicides found in low- and middle-income countries in 2016" and I was once more certain that I want to explore this rather sensitive and complicated topic.

This dataset is compiled from 4 other datasets which are linked by time and place. For reference:

1. United Nations Development Program. (2018). Human development index (HDI). Retrieved from http://hdr.undp.org/en/indicators/137506
2. World Bank. (2018). World development indicators: GDP (current US$) by country:1985 to 2016. Retrieved from http://databank.worldbank.org/data/source/world-development-indicators#

3. [Szamil]. (2017). Suicide in the Twenty-First Century [dataset]. Retrieved from https://www.kaggle.com/szamil/suicide-in-the-twenty-first-century/notebook
4. World Health Organization. (2018). Suicide prevention. Retrieved from http://www.who.int/mental_health/suicide-prevention/en/

The dataset contains of 12 columns and 27820 rows. I have also added some other fields to it, but I will elaborate on them in the chapter "Process". Below, the description of the columns can be found:

Column name	Column data type	Column description
Country	Country	The country's name
Year	Integer	The year for each country
Sex	String	The gender (male/female)
Age	String	The range of age (e.g. 5-14, 15-24, 25-34, etc.)
Suicides_no	Integer	The number of suicides based on country, year, gender, and age range
Population	Integer	The number of people based on country, year, gender, and age range
Suicides/100k pop	Decimal	The suicide rate
Country-year	String	A composite key of the country & year
HDI for year	Decimal	The human development index for country & year
GDP_for_year	Decimal	The gross domestic product for country & year
GDP_for_capita	Integer	The gross domestic product for country & year based on the population
Generation	String	The generation based on age grouping average (e.g. Generation X, Boomers, Millennials etc.)

TARGET AUDIENCE

My target audience are people of all ages who want to be informed about this serious topic and to see what might be the causes which lead to it. I have an idea to explore different countries and what the differences are (does GDP has influence on that?), also examine if there are difference between the 2 genders (why do rates of men committing suicide are significantly larger than rates of women).

GOAL

The goal that I want to achieve through this challenge is to raise awareness for a topic that is sometimes overlooked and not talked about as it is "sensitive". According to the World Health Organization and the Global Burden of Disease study estimate that almost 800,000 people die from suicide every year. That is one person every 40 seconds. And, indeed, it is a complex issue that causes pain to thousands of people every day around the world.

The objective of creating a data visualization is to contribute to make the society more well-informed. Health is not only about medicine and medical resources, but it is important to acknowledge that in order to achieve universal health coverage and sustainable financing for health, we have to address the growing numbers of non-communicable diseases, including mental health.

EXPLORING THE DATASET

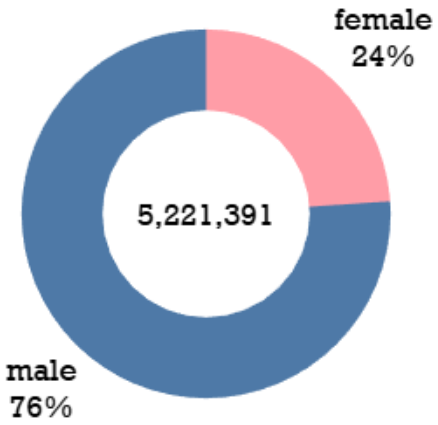
The dataset that I used is overall good and appropriate for making visualizations. A downside of this dataset is the fact that there is no data available for most of the African and South-East Asian countries, which is a pity as it would have made a huge difference for the visualizations. This also means that it will not support the above mentioned statement that "Suicide remains the second-highest cause of death among people aged 15 to 29 globally, with 79 per cent of suicides found in low- and middle-income countries in 2016". However, realistically looking at it, such information is difficult to track in those countries and it is normal that we still have a gap there. Additionally, the dataset is lacking the last 5 years and thus I am only focusing on information till 2015.

Nevertheless, the dataset is particularly nice to work with when it comes to comparison between the 2 genders as many visualizations and conclusions can be taken from it. Another thing that I like about the dataset is that it provides the information not only of age gaps (e.g. 25-34, etc.) but also the different generations. Again, we can create beautiful and meaningful statistics with this data. The addition of GDP and HDI is also interesting and we can examine if they have correlation with the suicide rates.

DESIGN CHOICES

For the design of visualizations, I chose to focus on two main colors – blue and pink. I think I made the right decision as these colors are usually not connected to such a serious topic and I think it might be not so stressful when a user observes them.

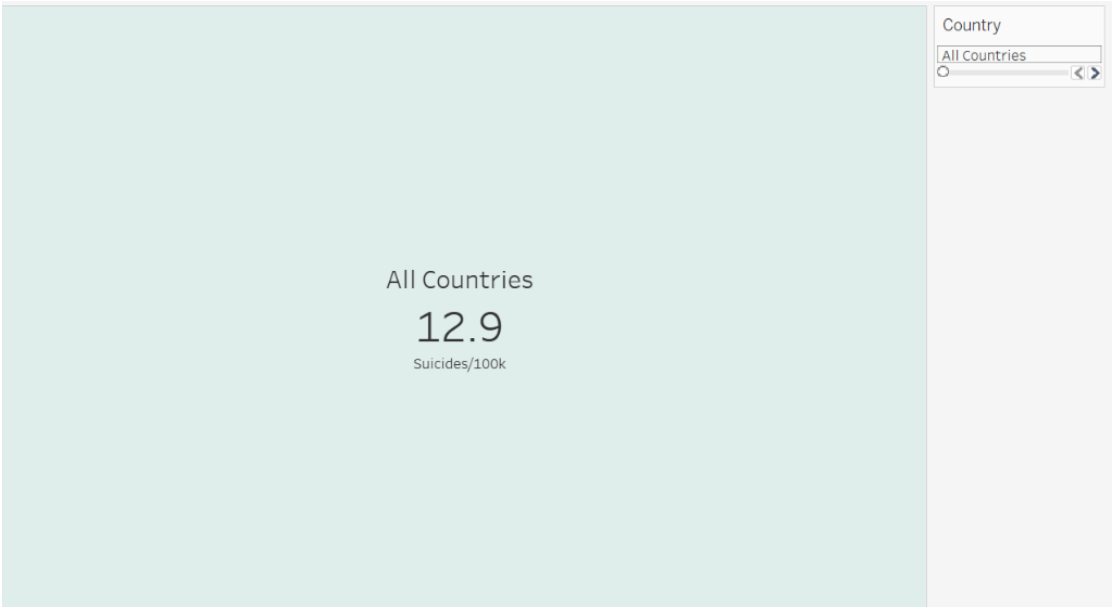
Another thing that I took into consideration were the colors of the pie chart representing the 2 gender and the statistics about the amount of suicide cases. As generally blue is color related to males and pink to females, I chose to keep that social construct. I believe in this way by just looking at the chart, the person is familiar with the situation that I am comparing statistics for the 2 genders.



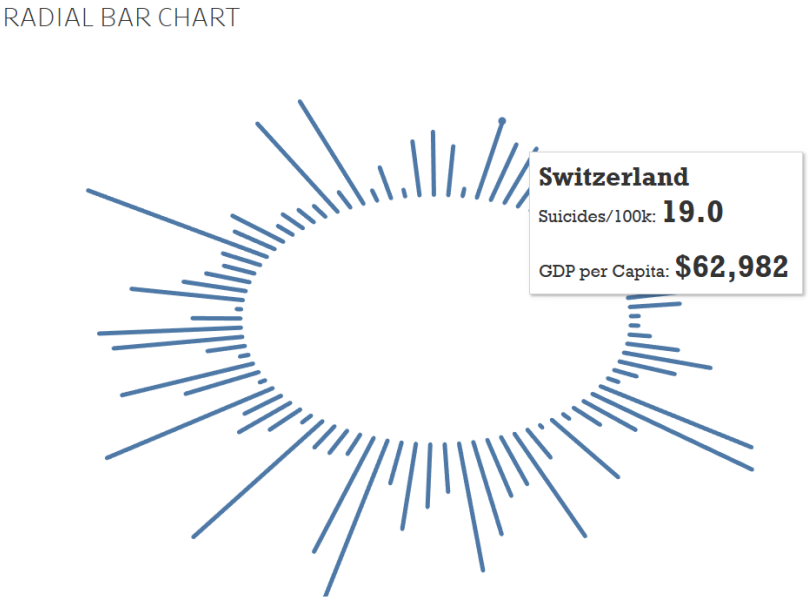
What I think was a nice touch to add, was the quote from Victor Hugo's book "Les Misérables" – "Even the darkest night will end, and the sun will rise." As I am visualizing a topic that is rather dark and negative and with the title "Suicide Rates from 1985-2015" is not optimistic as well, I wanted to bring some light and hope by adding this subtitle.

VISUALIZATIONS

The first visualization I made was to show the **average suicide rate**. It turned out to be 12.9 for all countries which implies that approximately 13 per 100k people take their own life. After calculating the average for all countries, I added a filter so that this can be calculated individually for every country.



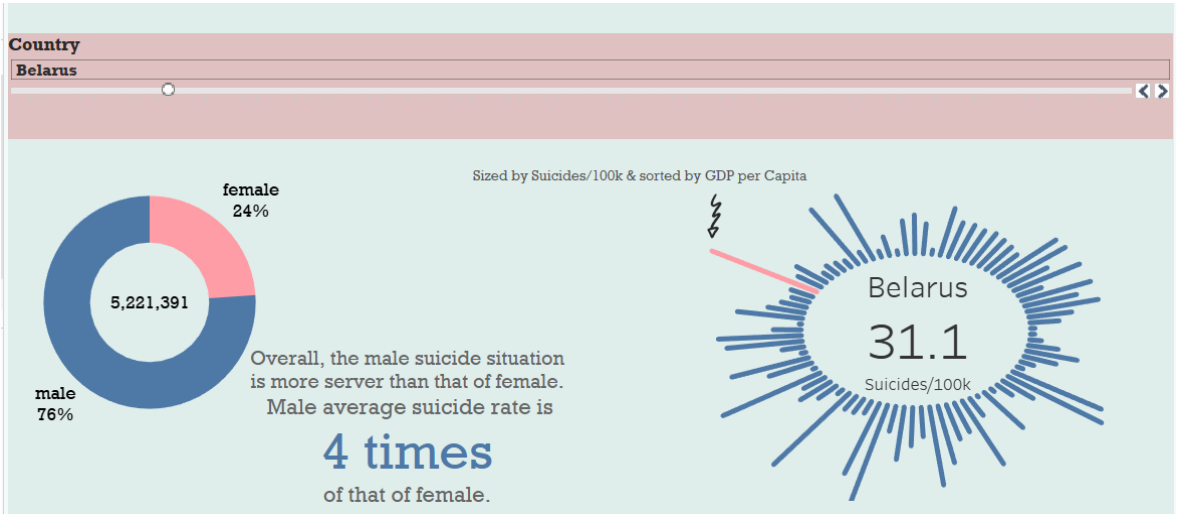
I put the information from above in the empty space of the **Radial bar chart**. This chart was one of the most interesting to make as it is something that I do not see much in people's dashboards. It shows the suicides per 100k people and the GDP per Capita; the length of the bar lines is corresponding the suicides/100k.



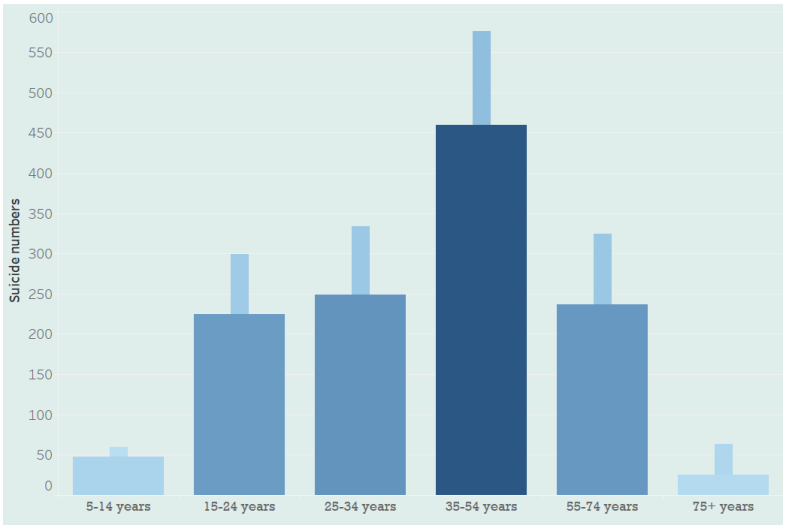
For this visualization, I started off by creating a new field in my dataset called "path order" which will help me draw this bar lines. The formula for it is, where 0 is the beginning of the line and 1 is the end of the line:

I have also added 2 more fields to help me calculate the angle where the line should be drawn, but I will not dive into more details as the formulas were just copied from a YouTube tutorial. That is the link: https://www.youtube.com/watch?v=d6-aptKLvgg

I added a slider to the dashboard where all countries are listed and the corresponding country in the slider is highlighted in the radial bar chart:

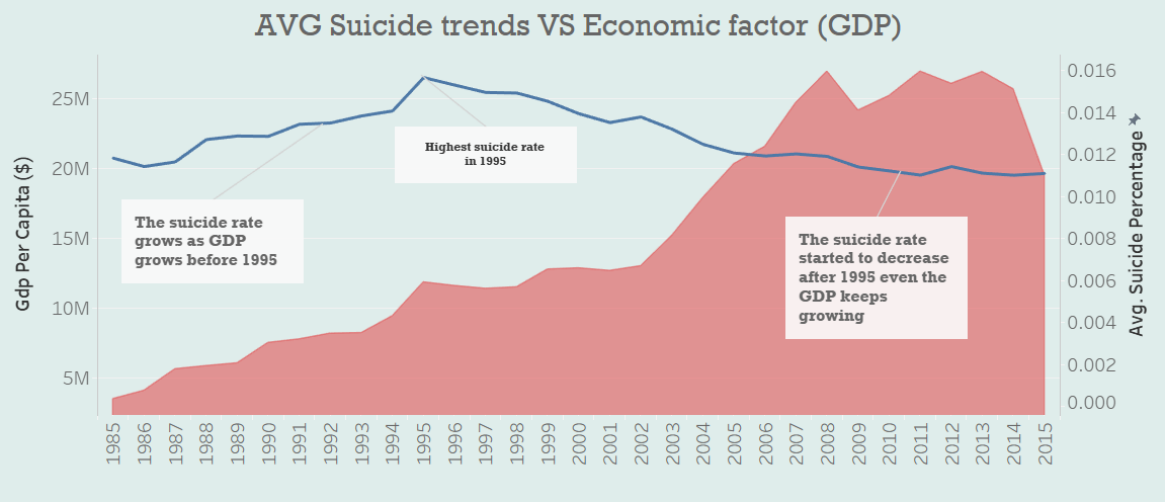


To finish this part of my visualization, I added a **third bar chart** to show the amount of suicides for each country for both men & women for the different age groups. This graph also works with the slider from above.

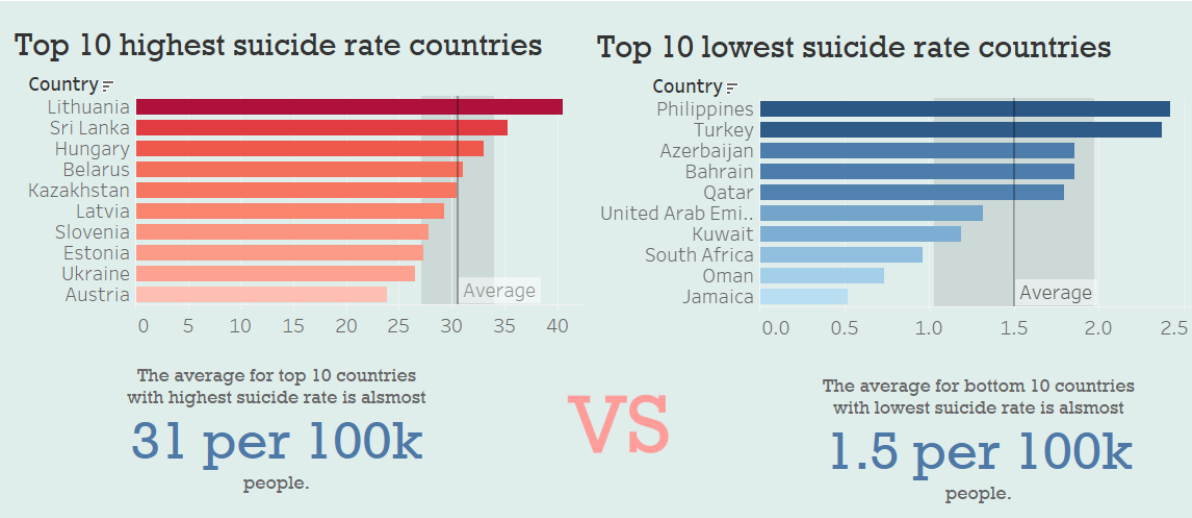


The **donut pie chart** represents the total number of people who have taken their life and the ratio between males and females. Based on the statistics, we understand that the male suicide situation is more server than that of female as the average suicide rate for males is 4 times of that of females.

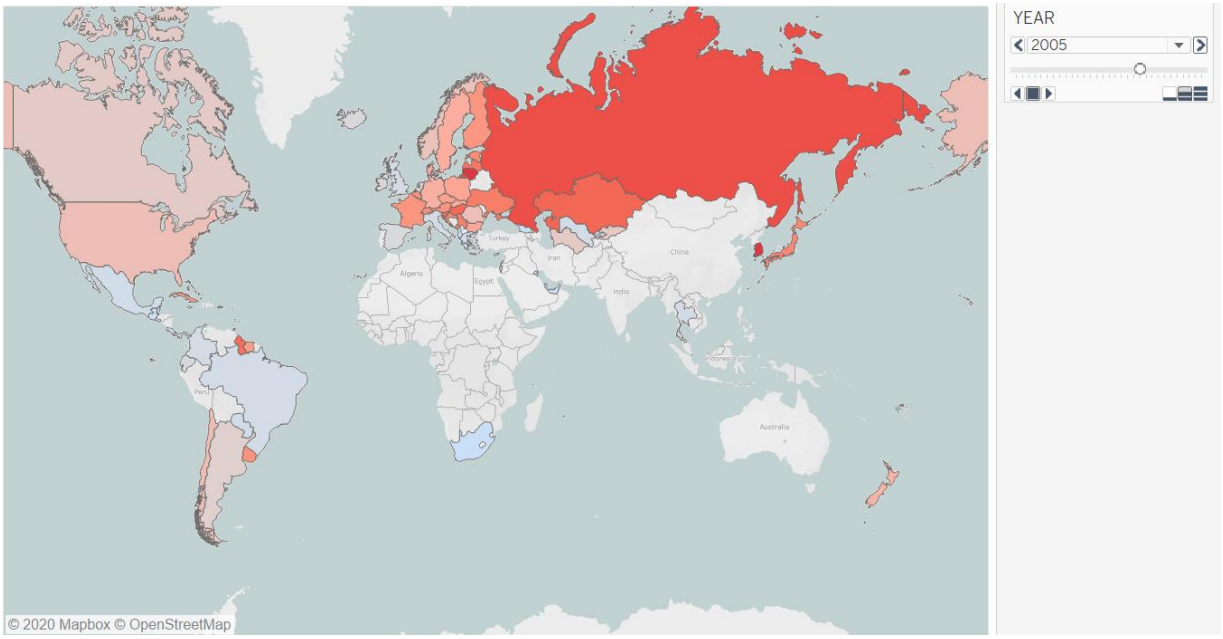
It was really interesting to me to see if the economic factor plays an important role in people taking their own lives. Therefore, I created a **dual-axis graph** where on one hand there is a line with the average suicide trends over the years and on the other hand there is the GDP per capita.



I have also created a visualizations to see which countries are the **top 10** with the highest and lowest suicide rate.



To wrap my data story, I made a map as Danny suggested in the beginning of this challenge. And what is a data story without a map? I decided to add one with a slider to see how new data is added to the map, and how the average number of suicides per 100k people increases/decreases for each country.



THE WEBSITE

After finishing with the second draft of my visualizations, I decided to start building my website. It is developed by using HTML, CSS and JavaScript, and the Tableau dashboard was first published on Tableau Public and afterwards embedded into the website. As a software student, it did not take much to lay the basics as I already have some code that I can reuse and fit for the information I am displaying on the website. I generally prefer to host it on GitHub instead of Hera which is provided by Fontys as I have more control and I can easily apply changes. For the design I again kept similar color scheme. In the main page, I have a description of the project so that people are familiar with what this project is about.

CONCLUSION

Overall, I believe that it was an interesting topic which is getting more and more recognition especially thanks to social media (many movies, tv series, songs, books). I hope that my visualizations also contribute to the distribution of how important it is to do not overlook mental health. My final thought and hypothesis based on the visualizations are:

1) There is a general decline in the average suicide rates globally. This paints a positive picture for the world. It could be attributed to the anti-depressants available in the recent years. It could also be due to the increase in awareness and help available over the past few years and the recognition of it on social media.
2) We also see that the top 10 countries with lowest suicide rate per 100k people are warm countries with higher temperatures relative to most countries which have a high suicide rate. So perhaps there could be a relation between temperatures and suicide rates.
3) The number of suicides for the age group 35-54 years could be highest due to the 'mid-life crisis' that most people go through.
4) The high difference in suicide rate by gender could imply that women express and seek help more often. It could also be due to the substance abuse or possession of guns which is more common comparatively among men (e.g. in the USA).
5) It is likely that males are facing more pressure from work than female around the world.
6) Some of my suggestions are:
    - Mental health programs should be designed for male between age 35 and 54.
    - Set up a friendly working environments for female such that they can take care of the kids as well as the family.
    - Free mental health programs should be promoted to as many as communities.