Suicide Rates Data Visualization

João Silva 88813, Bernardo Rodrigues 88835
Information Visualization, 2020 University of Aveiro

Abstract

This project was created for the Information Visualization course of the University of Aveiro.

The objective of this assignment is to create a visualization website using d3.js.

For the dataset, we picked a suicide rate overview dataset from Kaggle.

Motivation and objectives

Discover how the number of suicide cases has changed through the years for various countries of the world, and discover what factors influence said cases.

Users and the Questions

Users

Our userbase would consist of people interested in studying and/or knowing more about the behavior of suicide victims in recent years. Namely:

- Analysists of a given country or institutions that want to know more about the suicide rates of a country and how it compares to others.
- Historians that may want to know about how an important historical event has affected the suicide rates (for example see if a market crash has caused suicides to increase)
- Lastly students or anyone trying to retrieve information about the topic.

Questions to Answer

With our program we expect our users to be able questions like:

- If age and/or gender has any impact on the likelihood of a person committing suicide
- How the number of cases of a country has changed through the years
- If factors like GPD or HDI influence the number of cases of a country
- Have a general overview of the situation per continent/country.

Dataset

Dataset from:
https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016

The data consists of the number of suicides cases of several countries through the years along with information regarding the age and gender of the deceased, along with it, also contains the total population of a country and how the population is divided (age and gender). Additionally, it contains data regarding the GDP, GDP per capita and HDI of the countries through the years.

Visualization Solution

Since we want for our users to see the suicide cases on a country-by-country basis, we decided that the best way to present our data would be by creating a chloroplast map of the world with countries with high suicide rates represented with a red color while countries with low rates represented with a more toned-down variant. On total about 4 to 5 different colors would be enough to demonstrate the difference in case for the countries.

Countries where no data is available would be represented as greyed out on the map.

The suicide cases of a country represented on the map would be on a per capita scale to prevent countries with a high population to

dominate the high cases, although we think an option to view the total cases of a country would be useful and might give some additional information to the user.

Along with the map, 2 additional pie-charts would be used to represent the suicide cases one to represent the cases by gender and the other for age.

Since viewing the evolution of a country through the years was difficult using the map. Our initial idea was then to create various line-charts to show the evolution of the suicide cases and the GDP / GDP per capita. But after the responses from the prototype. We decided that it would be best to use a line-chart with 2 lines to compare the evolution of the GDP per capita and the suicides per capita.

All the visualizations would be subjected to several filters such as per year, sex, age and country.

Low fidelity prototype and user feedback

To test our idea, we used Adobe XD to create our prototype as test to see if it was easy to understand and able to answer the questions, we intend it to answer.
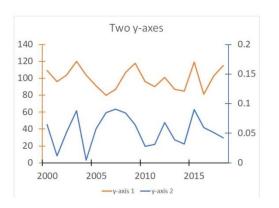

Figure 1: Prototype

Our prototype was static, so to do the user tests we created various copies with small changes to simulate the progress of the users during the tests.

Evaluation and changes in the prototype

Overall, the feedback we got from our evaluation was very positive users were able to understand the prototype and no major flaws were detected, however there were a few tweaks to be made and some helpful tips to improve the design like:

- A way to find smaller countries (that may be hard to click on the map) we decided to add a zoom feature to the map.
- Some users didn't immediately notice that the map was interactive and suggested that if it was bigger it would be easier to notice
- Easier way to filter the maps, ex: a checkbox to select/ deselect all, along with checkbox to select/deselect all from a specific region. We decided it would be better to remove the selection by country with checkboxes and implement only selection by continent. The selection by country would be done clicking on the map.
- As said before our prototype included the one-line line charts and during the texts we asked if they would prefer an alternative to how the graphics were shown and suggested the following:

Since the responses where positive we opted to show a

version similar to this one and remove the others to prevent redundancy.

Implementation challenges

For our implementation we used the d3js software, our main challenge was that this was a unknow technology for us prior to this course and as such we initially struggled on how we should approach the problem.

Also, the problem of finding and patching bugs is increased when one is unfamiliar with the language.

Some other concerns include how we could represent the territorial changes of countries through the years on the map (our dataset shows data from 1985 to 2016 and though those years some countries changed borders our ceased to exist, for example prior to 1990 Germany was divided into East and West Germany although our map shows a constant reunited Germany). Since our dataset also didn't make the distinctions of these countries. We decided to trade historical accuracy for a simpler and easier to implement design.

One change we made to our database was to join the suicide cases of people between "5-14 years" and "15-24 years" into "24 years or less". We did so because even if the topic of our work is pretty morbid, we don't want to make our users feel depressed. Plus, the number of cases of the younger ones was very low and when shown on the pie chart, it was hard to see, literally and emotionally.

Visualizations

Filters

The filters allow our visualization to be more dynamic. Our project only has 4 visualizations, but they all are affected by the filters and can produced several outputs according to what is selected.

We can filter the data by year, sex, gender and continent (we can also filter by country clicking the map). We also have the option to see the number per 100k population or by total cases.

The filters selected will affect all the visualizations, the map, pie charts and line chart will only show the data of suicides and population for the selected filters.

Map

Our chloroplast map shows the information we have for the selected filters. Countries that have no data will be presented in black.

The map supports mouse hover which will display information about the country given the filters selected, ex: if all filters are selected, we will have the total numbers, but if only male, 75+ is selected the number of suicides and population will only be the number of males with 75+ years old on that country.

We also can zoom in using mouse wheel select smaller countries.

We can click on a given country to select it, this way the rest of the visualizations will only display the data for that given country. We can re-select all countries clicking again on the same country.

We can view the data by cases per 100k population or by total cases and the color labels will change accordingly.
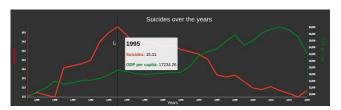
Pie charts

The pie charts allow the user to see the distribution of suicided per sex and age group for the selected filters, ex (if only 35-44 years old is selected the sex pie chart will only display the relation of male and females in the 35-44 age group).

The charts implement mouse hover which will display the number of suicides of the given group or the suicides of the group per 100k population of that same group.

Line chart

The line chart displays the evolution of cases (total number or per 100k population) and GDP per cap. The goal of this visualization is to understand if the economic evolution of a given country/continent is affecting the number of suicides on that region.

The red line (suicided) has its y axis on left and the green line (GDP per cap) has its y axis on the right, both axes have the same color of the respective line.

The chart implements a tooltip along with a vertical dark line that shows the value of the lines for a given year. The value of the number of suicides is affected by the filters selected, however the GDP per cap. Line is not affected by gender/sex filter since it would not make sense.

Usability test

We asked some friends who did not participate in this project and asked them to use the application for a while.

The biggest common problem was one that we already had noticed which is the positioning of the filters. When the user is in the charts page, he has to go to the map page to change the filters and although it is only one click away, it would be

better to just have the filters stay in the page when visualizing the pie and line charts. We knew already this was going to be a problem but since it was not really the scope of the subject and time is running low for this and other projects, we opted to just keep it this way.

Another problem pointed out was that It would be nice to be able to see the number of suicide cases in the pie charts or the percentage of a group. This was before we implemented the mouse hover so this is no longer a problem.

Things we learnt with the visualizations

During the testing of the website, we couldn´t help noticing some things we didn't knew before in our own visualizations. Some examples are:

- The age group of 75+ years is the most affected by this issue. Its numbers are not the biggest since they are a small part of the population but if we consider just a population of the 75+ age group they have the higher suicide rate.
- In the case of Portugal if all the age groups (-24, 25-34, 35-54, 55-74, 75+) had the same population, around of 60% of national suicide cases would be from people with more than 75 years old.
- Male suicide rate is much higher than female suicide rate (we already knew this).
- China is the only country we found to have larger female suicide rate over male suicide rate although it is becoming less apparent over the years (this is probably due to the son preference).
- Most of the countries show an inverse relation between number of suicides and GDP per cap, meaning that when the country is in a better economical phase, the number of suicides diminishes. This is not true in the EUA, where the number of suicides increases at the same rate as the GDP per cap line.
- Russia has really high male suicide rates (this is probably due to known problem Russia has with alcohol).

Conclusion and Future Work

In the end we were glad to have been able to create a website that present the topic in what we think is a comprehensible and respectful way.

We implemented several useful visualizations while also furthering our knowledge on how to create websites and how to use the D3 tools.

If we were to continue this work, the main points we would improve would be:

- The position of the filters (no need to go back up)
- New visualization to compare all countries with the data from all the years (the only visualization we have for seeing the cases within all the years is the line chart, it would be great to have a way of comparing all the countries.)
- Refactor the code. This project was done in a hurry with no real background knowledge of d3.js or even html and JavaScript (we are not really used to it) so a great deal of great practices were left behind.

The project is also available here: https://github.com/joaonpsilva/VI_suicideRates

Open index.html in a live server (if opened directly won't work) ex: VsCode, extension Live server.

References

chloroplast map - https://www.d3-graph-gallery.com/graph/choropleth_basic.html

pie charts - https://www.d3-graph-gallery.com/graph/pie_annotation.html

https://www.d3-graph-gallery.com/graph/pie_changeData.html

line chart annotations - http://bl.ocks.org/wdickerson/64535aff478e8a9fd9d9facccfef8929

basic page layout - https://codepen.io/bradtraversy/pen/xBdyzr

https://www.youtube.com/watch?v=y9nlfqT4s9s&t=528s

dataset - https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016