Design study project: Suicide Rates

Christian Rauch* Daniel Hanzer† Roman Schneglberger‡

University of Vienna
Faculty of Computer Science

1 Motivation

1.1 Background information

Our Application deals with the suicide rate of a total of 41 countries. The goal of this project was to create a visualization that show a variety of graphs that are changeable by user interactions. The graphs are able to filter to get more precisely data out of it. It is possible to compare different countries with each other. Furthermore, it is possible to compare the suicide rates with unemployment rates, poverty rates as well as Gross domestic product (gdp) rates for each country. These comparisons should help to find possible parallel between the data in order to find any conclusions about the reasons for suicidal ones.

1.2 Tasks

Our tasks was to visualize the existing records as good as possible. One of our tasks was creating a heatmap. In addition, a timeline should be available to filter for certain periods of time. In general, our the main task was to install meaningful filters or groupings to give the graphs as much meaning as possible. Another task was to find an explanation for the reason for the increased suicide rates in certain countries or regions, therefore we compared them with gdp, unemployment and poverty rates and visualized these through a matrix scatterplot and a line chart over time. Another task was comparing different countries or regions with one another over certain years.

1.3 Users

On the one hand, our users are generally interested in our project, as well as users who conduct research in the field of suicide. Our users should be able to better understand the records by filtering them. For example, users might question why certain countries have a higher suicide rate than others and whether it involves economic or geographic causes. In particular, the course of suicide rates over time are relevant for this. Potential users may include journalists, researchers and psychologists.

1.4 Data

The Project focused mainly on the Dataset ”Suicide rates” from OECD Data. The Data are from the year 1960 until to the year 2015 and include the amount of suicides per 100.000 in percentage. Altogether, the data of 41 countries are presented in this dataset. Furthermore, we used ”Gross domestic product (GDP)”, ”Poverty rate”, as well as ”Unemployment rate” from OECD Data for comparing these datasets with the suicide rates. The Data from GDP are from the year 1960 until to the year 2015 and include 40 countries. The Data from Poverty rate are from the year 1976 until to the year 2015 and include 40 countries. The Data from Unemployment rate are from the year 1953 until to the year 2015 and include 40 countries.

*e-mail: a01202875@unet.univie.ac.at
†e-mail: a01349699@unet.univie.ac.at
‡e-mail: a01127050@unet.univie.ac.at

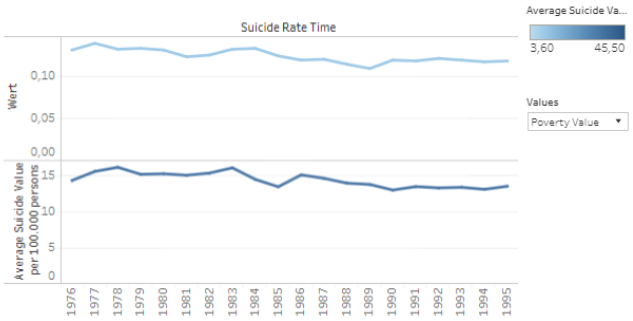

Figure 1: Line chart showing a dip in the value of suicide rates in Canada of 1985

One problem with our Data was to merge them together. Another one was that not for each country are Data available for each year, so to speak there are a lot of null values in the existing records.

2 Related work

3 Approach

4 Implementation

5 Results

5.1 Scenarios of use examples

- Researching the impact of events on suicide rates

  A researcher wants to view the impact on suicide rates in a country of a certain event that happened in 1985. For this, the researcher selects the country in the map view (Figure 3) and selects a appropriate time scale (Figure 2).

  In the line chart (Figure 1) a slightly lower suicide rate in Canada in the year 1985 is visible.

- Researching how much / if economical factors influence the suicide rates in different countries

  We select multiple European countries with CTRL + click on the map view as shown in Figure 4. In the color coded scatter plot matrix (shown in Figure 5) you can compare the correlation between various economical factors and suicide rates in the selected European countries.

- Simply comparing suicide rates in different countries

  Without selecting anything, per default our visualization shows the suicide rates of all countries (with available data) in a bar chart. As you can see in Figure 6, the bar chart features color coding (identical to scatter plot matrix) and a mouse-over hint.

  This default view should show the large differences in suicide rates in the various countries and encourage interest to the question ”Why is that?”. The user can now play around with the other charts and try to find possible explanations in economical factors.

Figure 2: Selecting a time scale for a more detailed view

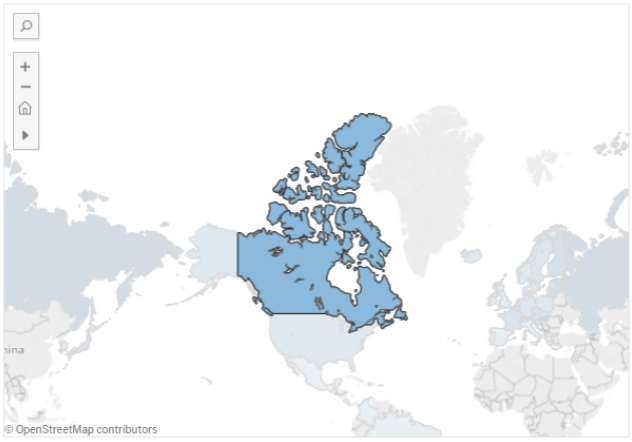

Figure 3: Map view

5.2 Performance of the system

We designed, tested and improved our visualization to perform well with the intended use case scenarios, including but not limited to the ones detailed in 5.1.

The performance in speed is limited by our visualization tool of choice (tableau) and the size of the dataset. We discussed limiting the dataset, but that would possibly lower the functionality for a rather small usability gain and could also violate the design principle to show context.

5.3 Feedback from evaluations

User feedback was a vital element in the design and improvement of our visualization. We improved the color coding (figures 5 and 6), implemented a scatter plot matrix instead of multiple scatter plots (or scatter plots selectable by a drop-down menu) and made a split line graph (figure 1) all based on user feedback and evaluations.

With our final visualization the only user complaint remaining is the loading time when selecting multiple countries one after another and initial difficulties to interpret the scatter plot chart.

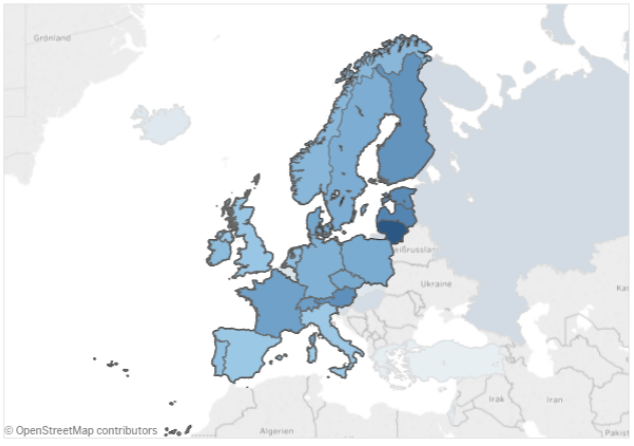

Figure 4: Multiple selection on map view

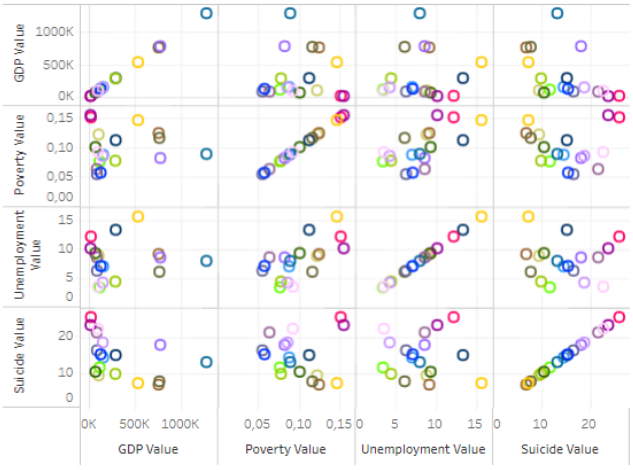

Figure 5: Scatter plot matrix

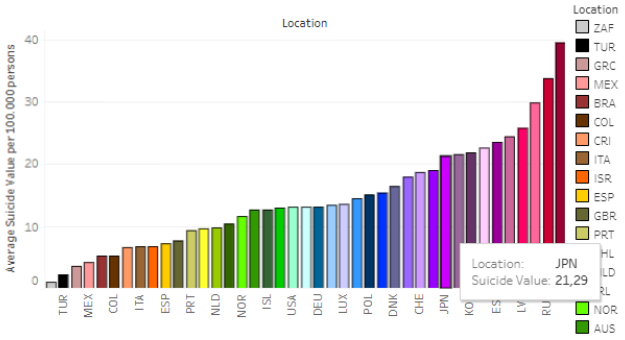

Figure 6: Default view of bar chart with mouse-over hint

6 Discussion

6.1 Strengths and weaknesses of our visualization

Strengths Every graph is necessary to visualize the intended concepts - no ”chart clutter”. Our visualization provides a quick overview of the different suicide rates in many OECD nations per default and makes it easy to compare them, which would probably be the most common use case. For a more detailed analysis it provides the necessary tools to find possible correlations to other factors (economical and developmental).

The scatter plot matrix makes it easier, especially for advanced users, to spot correlation between the visualized data on a grand scale - the line chart on a more time based / historical scale. It can be simplified by restricting the number of countries shown.

(Possible) Weaknesses There are only so many factors we could incorporate into our visualization and it is not guaranteed there are correlations between these factors and suicide rates. If we included more it could lead to a too complex visualization and possibly make it even less likely for users to spot correlations as a result.

Selecting countries in the map view leads to loading times. It is not a very significant problem, but if the user wants to select many countries (for example European countries) it becomes annoying pretty fast. We tried to improve the loading times, but it seems to be a problem with tableau and the size of the datasets.

The scatter plot matrix can be slightly difficult to read and interpret correctly because of its high data density. However we are

confident this is the best way to visualize possible correlations between the various datasets and suicide rate among all countries and it enables the user to spot correlations on a grand scale in one graph.

6.2 Lessons learned

- Visualization can be a time consuming process and good time planning is necessary.
- User feedback is a vital element for visualization. The developers are very likely to overlook possible flaws and shortcomings.
- Tableau is an easy to use tool if you want your visualization to work and look exactly as the developers intended it to. If you want to do something ”non-standard” with this tool, the process can become quite complicated.

7 Separation of Tasks: Milestone 4

Christian Rauch:

- ”Results” section
- ”Discussion” section
- ”Separation of Tasks” section
- LaTeX document setup and coding

Daniel Hanzer:

- Improvements and implementation of feedback in visualization (correct color coding and scatter plot matrix)
- ”Motivation” section

Roman Schneglberger:

- ”Related Work” section
- ”Approach” section
- ”Implementation” section