Data Visualization: LCELO Process Book

Aslam Cader, Etienne Caquot, Jeanne Chaverot

May 2020

1 Introduction

The goal of this project is to display the impact of the Covid-19 virus on the swiss population. The data set is a daily time series containing the number of contaminated/deaths/recovered, as well as information about the number of individuals hospitalized in a normal ward, ICU, or in need of respiratory assistance. As we have access to this data for each Swiss canton, we have decided to use this project in order to enlighten the situation of this health crisis in each canton. The output from this project will help have a better understanding of the situation in each canton, as well as seeing which one(s) are better able to handle this crisis.

2 Path to obtain the final result

2.1 Abandoned goals

The interest in the Covid-19 data-set arose upon realization that as individuals we are surrounded by misconceptions around the reality of this pandemic. As we know, many websites have already used Covid-19 data for general visualizations from a worldwide perspective. We are interested in taking a closer look at Switzerland's situation. First, we wanted to compare the date of the first lock-down (and other security measures) between Switzerland and the rest of the world. We find it interesting to see which Government was the promptest and acted the quickest when faced with this pandemic. Then, we wanted to work on a more human scale. We want to raise awareness on how this virus not only presents rather basic symptoms but is more dangerous than the basic flu to mankind. To do so, we will implement a word-cloud based on symptoms positively-tested individuals have presented at the hospital. We also want to plot the statistics comparing the Covid-19 to the seasonal flu.
However, these two first ideas were rejected after careful consideration and feed-back from our supervisor. It has been suggested that it would be a better idea to focus on a single country in order to extract and display more useful information. Finally, we were told that what was expected were lesser visualizations with deeper meaning and complexity. A visualization such as a dynamic word-cloud, one of our ideas to compare the Covid-19 to the seasonal flu, would be considered too simple and more of a decor than an actual source of information.

2.2 Final goals

Therefore, we decided to take a deeper look into Switzerland's reality. Switzerland is divided into 26 cantons, and each of these cantons has its own statistics. In the data visualization protocol, we want to analyze different elements from Swiss hospitals: the number of hospitalizations, number of individuals under respirators, number of individuals in Intensive Care Unit (ICU), and number of released individuals. We would like to compare these values with the magnitude of the Canton in question, as well as being able to find similarities between cantons. With an interactive map, we can see the number of infected individuals per Canton.

2.3 Finding the right visualizations

Once we knew that we wanted to work on Switzerland's we needed to define which visualizations to draw. This has been a challenging part as currently many parties are working on the same subject. We knew that by focusing on Switzerland we had a certain differentiation, however, we needed our ideas to be original. In order to validate our visualizations hypothesis, we conducted a poll to gather some feedback about which information individuals would like to have access to. This has started an iterative process between brain-storming and finding the appropriate visualization to answer a certain need. We decided to focus on two aspects of our data:

- A first visualization using the number of infected/recovered/deaths per canton
- A second visualization using the data on the situation in hospitals per canton

3 Challenges and design decisions

In this part, we explain our choices regarding the website design, as well as detailed explanation of our visualizations.

First of all, we needed to have a dynamic skeleton with a highly responsive website.

Working on this subject was interesting and challenging at the same time. Many companies, universities and research departments have been working on the same data set. Therefore, producing something creative required a lot of brainstorming. That is why we decided to focus on Switzerland: in order to still have a broad view, but with the opportunity to compare the situation between each canton.

3.1 Map overview

Choosing the map was a strategic choice: it's an easy way to compare the evolution of the situation canton by canton.It allows the viewer to understand the data without raw numbers in a more illustrated way. We have chosen blue for cases, red for deaths and green for recovered, as blue represents a rather neutral color, while red and green are more attache to negative and positive sentiments. This map is interactive with its slider. The user can see the evolution of the situation in time (starting on March the 6th until now). The user can decide which information to display: the number of cases, deaths or recovered, as well as the numbers format: either absolute or relative. The absolute represents the total number of affected individuals per canton, while the relative value is as a percentage of the canton's population. We believe the relative data is more informative as absolute data from highly populated cantons will always be higher in comparison to smaller cantons. Relative numbers are more representative of a canton's situation.

Below our map, following a discussion with our supervisor, we decided to add a time series displaying the evolution of the number of cases/deaths/recovered in Switzerland. This plot allows the viewer to have a general view of the virus' evolution in the country. As before, this plot can be shown in terms of absolute or relative numbers.

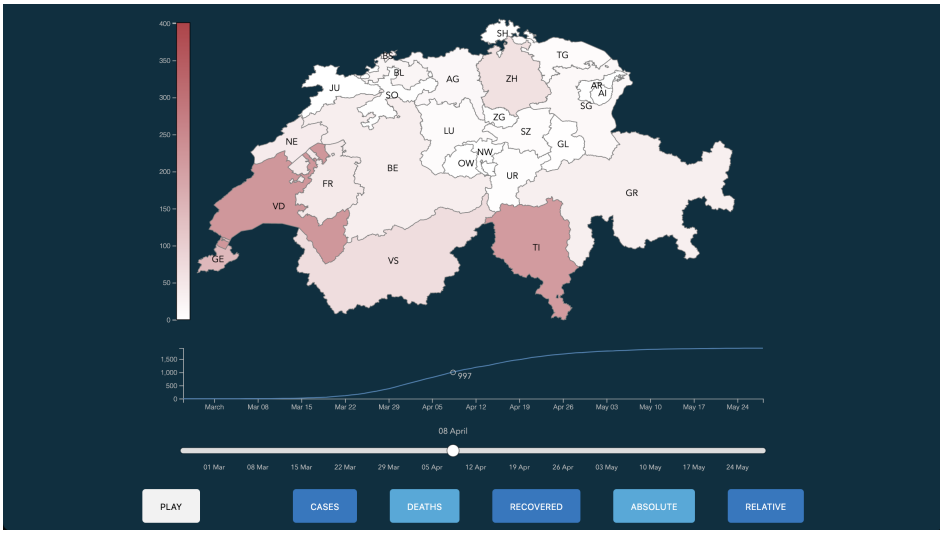

Figure 1: Switzerland map with deaths, absolute count

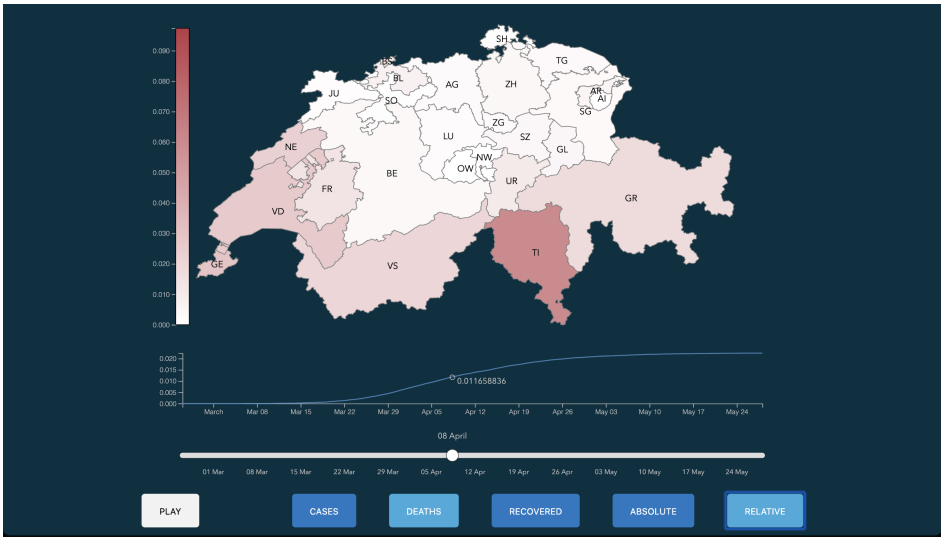

Figure 2: Switzerland map with deaths, relative count

3.2 Graph for hospitalisation

This graph shows the distribution of the hospitalized covid-19 cases: normal ward, ICU ward, or under respiratory assistance. This allows individuals to compare the pressure that each canton's hospitals unit has handled so far. We believe that this visualization is a good indicator of a canton's ratio between its density and its health system support.

The viewer can select the cantons he would like to see displayed in the plot, and can travel through time with the slider. We decided to let the user decide which cantons he would like to see so that he can compare two or several cantons over time (thanks to the slider). The graph, called a Marimekko chart 3, allows us to have an easy to read illustration, in order to compare the hospitalization situation among cantons. This slice and dice visualisation permits an hierarchical view which helps to compare the cases between cantons. An additional feature has been added in case the viewer wishes to see the data displayed for all the cantons 4. By clicking a toggle button, you'll be able able to simply display the situation for all cantons. We chose to work with a Marimekko chart as it is more dynamic and responsive, providing a more visual classification of the cantons depending on their hospitalization situations. However, for a more general view, we think it remains interesting to keep a stack bar chart with all cantons.

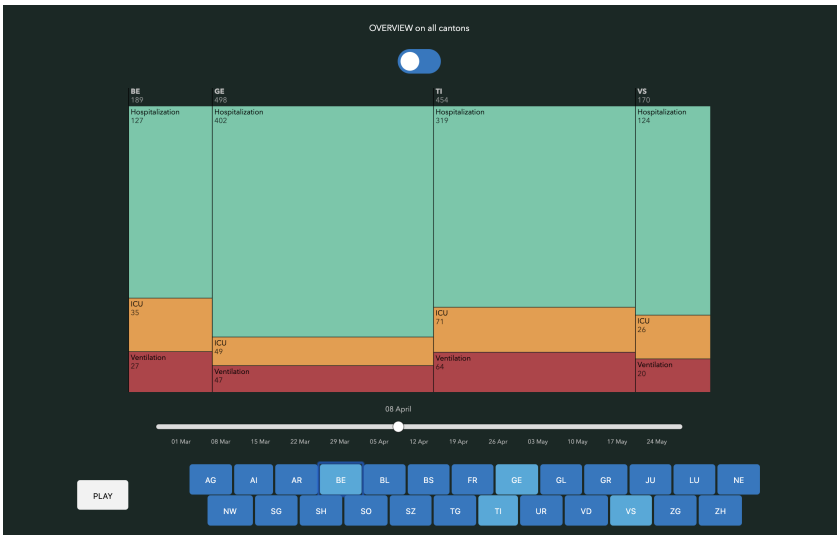

Figure 3: Comparative view for cantons: Marimekko slice and dice view

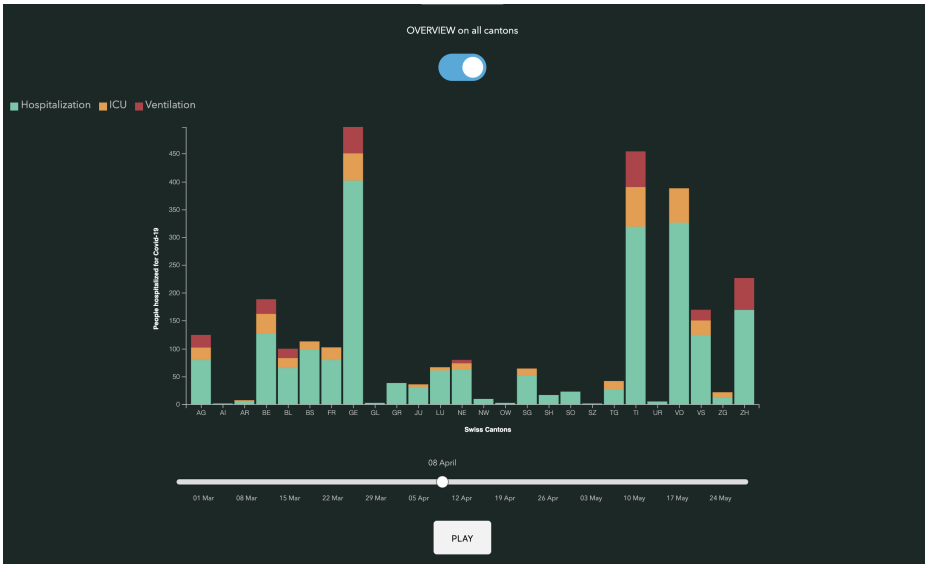

Figure 4: Overview on all cantons: Stack bar chart

4 Changes since milestone I

In milestone 1, our focus wasn't only around Switzerland, we wanted to compare Switzerland versus the rest of the world, as well as looking into the differences and similarities between the Covid-19 and the flu. However, we were correctly indicated that we were going into too many directions at the same time. That is why we decided to focus on Switzerland, in order to dedicate our visualizations around one subject and be more consistent in our work.

5 Peer assessment

In the context of this project, we decided to continuously brainstorm our ideas in order to agree on the story telling and direction of our visualizations. Etienne focused on the D3.js implementations of our visualizations. He worked on the JavaScript code in order to implement the heat map and the Marimekko chart. Jeanne worked in the creation of the visual sketches in order to define a first view of the website. Then, she worked on the data cleaning to give a less biased view of the data. We switched from quantitative data to relative data. Aslam worked on an interactive bubble word-cloud implementation with D3.js in order to detect symptom tendencies in Covid-19 positive patients. However, we later decided not to keep this visualization as we considered it didn't bring enough information and was more of a decorative asset. He worked on the skeleton of the website and did the stacked bar chart.

6 Conclusion

References

StackOverflow