Data visualisation: Milestone 2

Jeanne Chaverot - Etienne Caquot - Aslam Cader

Due in May 1, 2020

Our interest in the Covid-19 data-set arose when we realized individuals were 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 off, we would like to compare the date of the first lockdown (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 would like 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.
- Finally, we want to take a deeper look into Switzerland's reality. Switzerland is divided into 26 cantons, and each of these cantons has its own statistics. From a statistician's point of view, we can describe this epidemic as a simple SIR model. The SIR model (Suspected-Infected-Recovered) is the simplest compartmental model in the epidemic analysis. It is represented by three categories: an individual is either suspected (he may become infected), infected, or recovered. For each of these states, there exists a transition probability from one state to the other. We call β the probability $P(s(t) = I|s(t-1) = S)$. That is to be 'infected' on day t given that our state s on day $t-1$ was 'suspected'. The same way, we define γ as $\gamma = P(s(t) = R|s(t-1) = I)$. Back to our data visualization protocol: in this last step, we want to analyze different elements from Swiss hospitals: 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. If this number increases although lock-down measures have been implemented, it may be an indicator that individuals are not respecting such measures, yielding a relatively large β. The fault may be retraced on individuals. However, if we see that the number of hospitalizations, individuals in ICUs, or under respirators amplifies, it illustrates the canton's inability to face the epidemic yielding a relatively large γ. On the other hand, if we see that the number of released keeps increasing, and the number of complications decreases, we can suggest a good ability from a canton to deal with the crisis, and estimate a reducing γ.

This analysis is as we see it innovative as we want to create a link between these visualizations and our country's situation. We want to learn something from this data, as we use it as a messag

1 Visualisations

1.1 Switzerland vs the world

It's interesting to compare how Switzerland differed from the other countries when facing this pandemic. There were different political choices and we hope to show them with this visualization. We want to build a dynamic graph where we can select countries in order to see the new cases per day and compare between the countries. We will add a vertical bar with the decision taken by the authorities for those countries.

For this, we're going to use D3.js paths and the lecture 4.2 on D3.js introduction.

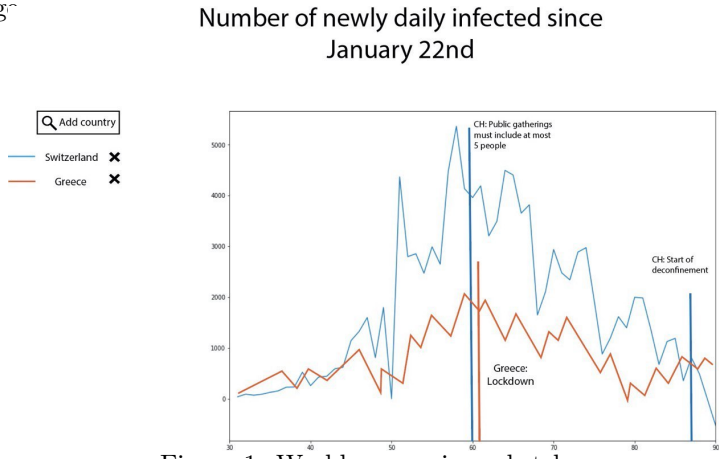

Figure 1: World comparison sketch

1.2 Symptoms

Using the textual symptoms in our dataset, we want to build a word cloud to identify the most common ones. The user would be able to interact with the word cloud by putting the mouse over some words to see more information.

For this, we're going to use D3.js word cloud, and the lecture 9 on Text.

Figure 2: Word Cloud

1.3 Comparison with seasonal flu

In this visualization, we will compare statistics from COVID-19 with flu, to do so we will construct a pyramid to compare data across age category. We will have to get the data for flu then creates the bars and we could add animations, clicking on an age category to have the number to make more interactive for the user.

For this, we're going to use D3.js paths, and the lecture 4.2 on D3.js.

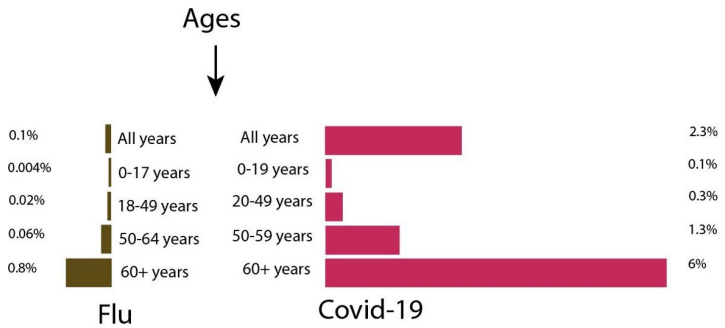

Figure 3: Death depending age of people

1.4 Graphs for Switzerland (Hospitalisation stats)

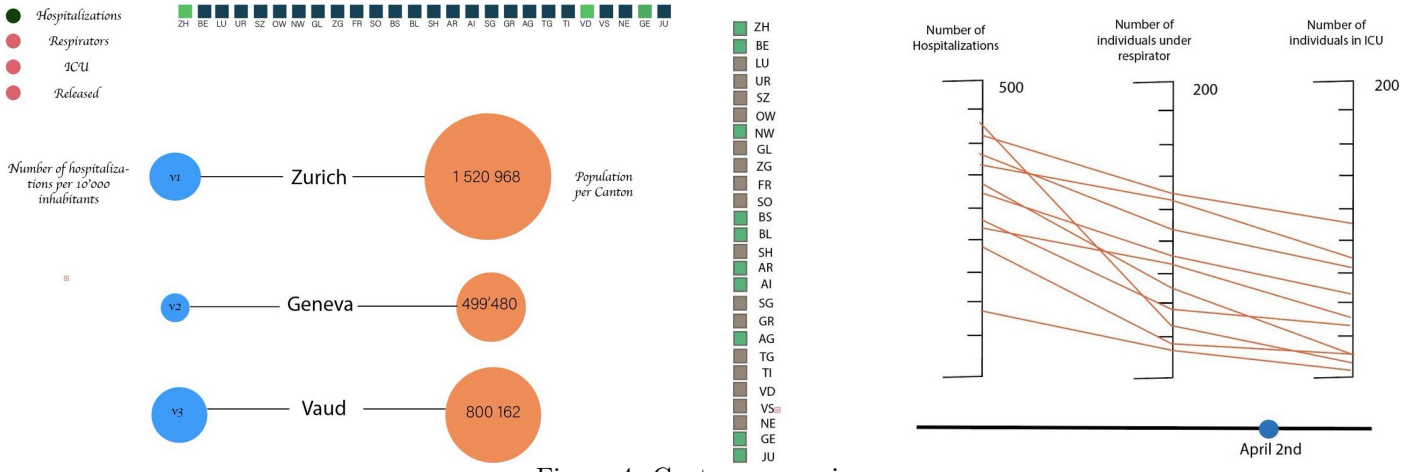

Figure 4: Canton comparison

We want to create graphs in order to be able to compare the statistics canton by canton. This approach aims to see how each canton is suffering from this crisis.

For the following graphs, we want to have interactive graphs where you can select cantons to display and observe the differences between cantons regarding the number of hospitalized people. We will plot the number of people concerned depending on the date.

We will show the pressure that the hospital system per canton is handling and has handled so far by having the number of hospitalized people per canton in ICU, normal unit, and people with a ventilator. We want the user to see if some cantons were able to better handle the crisis and the workload in hospitals.

For this, we're going to use D3.js paths, and the lecture 4.2 on D3.js.

1.5 Switzerland Map

In this visualization, we want to show the propagation of the virus over time in each canton. So that we can see how the disease propagated around Switzerland.

The idea is to have a map with a different color for each Canton representing the intensity of the statistic. There will be a scale, and a slider in order to select a date. There will also be a play/pause button to activate/deactivate the time-traveling animation. By putting your mouse over a canton you could have the numbers appearing. First, we draw the map, then link the data and finally will handle the slider.

To improve this visualization, we could let the user choose between different statistics (cases, death, recovered, in absolute value, or relative to the density of the Canton).

For this, we're going to use D3.js maps with TopoJSON, the lecture 8.1 on Maps and exercise 8.

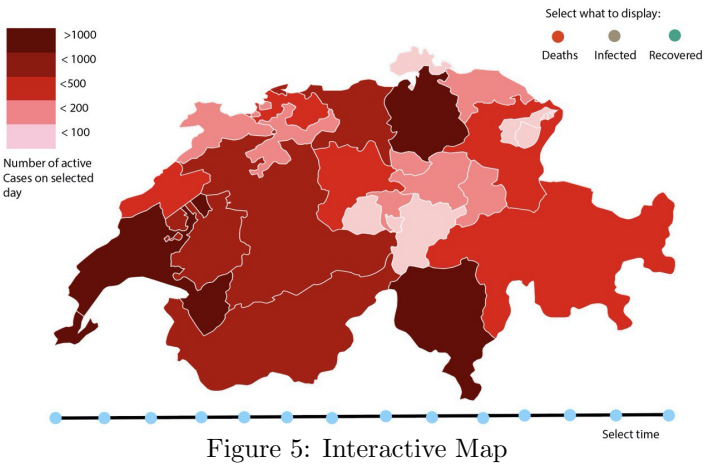

Figure 5: Interactive Map

2 Further idea

If time permits, we want to make visualization, where we could see the impact of not contaminating one person. This would look like a tree of person where everyone is contaminated and you could cut edge to act like there was no contamination and then all children won't be contaminated anymore. This visualization could help, understand the huge impact of lock-down.