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 realised individuals were surrounded by misconceptions around the reality of this epidemic. As we know, many websites have already used Covid-19 data for general visualisations from a worldwide perspective. We are interested in taking a closer look onto Switzerland' situation.

- First off, we would like to to compare the date of first lockdown (and other security measures) between Switzerland and the rest of the world. We find it interesting to see which Government were the promptest and acted the quickest when faced to 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 actually 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 flu.
- Finally, we want to take a deeper look into Switzerland's reality. Switzerland is divided into 26 cantons, and each on these cantons has their own statistics. On 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 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 visualisation protocol: in this last step, we want to analyse different elements from Swiss hospitals: number of hospitalisations, number of individuals under respirators, number of individuals in 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 canton. With an interactive map, we can see the number of infected individuals per Cantons. 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, it illustrates the canton's inhability ability 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 visualisation and the situation in our country. We want to learn something from this data, as use it as a message.

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 choice and we hope to show them with this visualisation. 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.

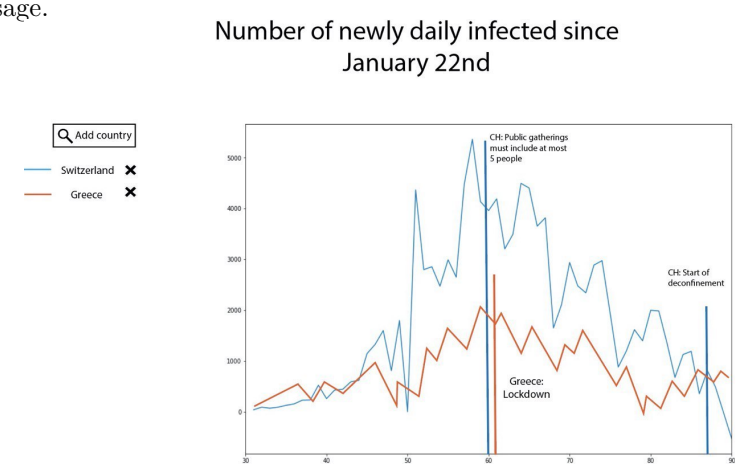
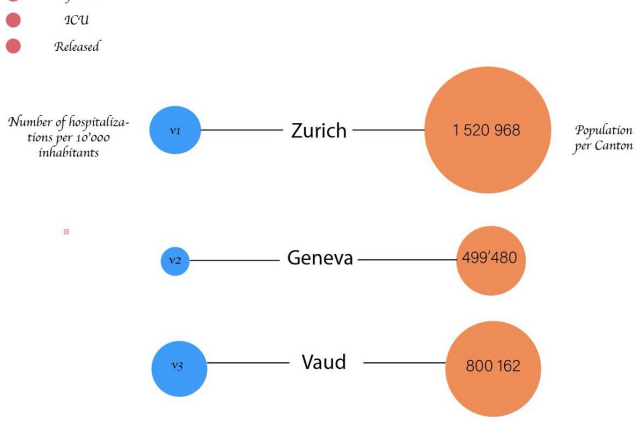


Figure 1: World comparison sketch

1.2 Switzerland Map



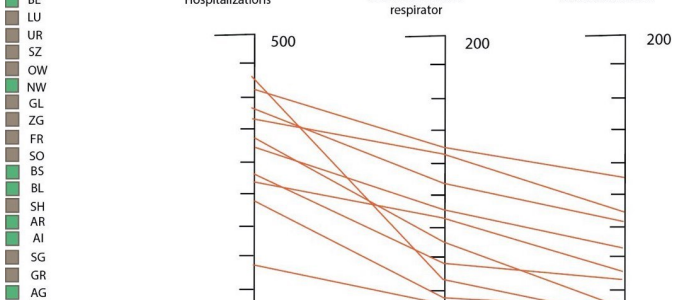


Figure 3: Canto

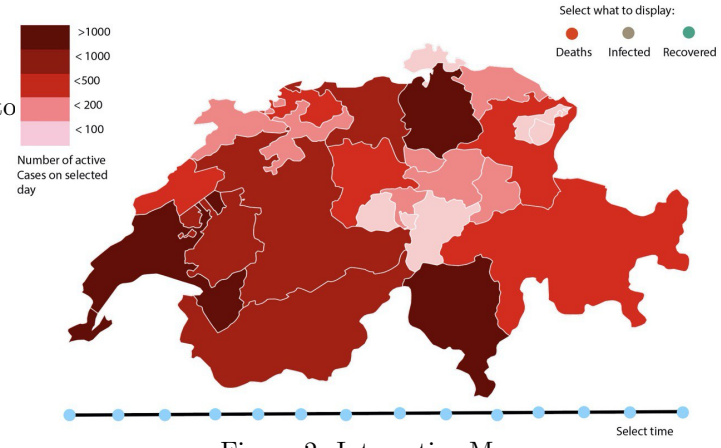


Figure 2: Interactive Map

In this visualisation, we want to show the propagation of the virus over time. We want the user to be able to see how the cases has evolve with time, this will help us to explain the difference (if any) that we will encounter when studying Cantons separately.

The idea is to have a map with different colour 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 play/pause button to activate/deactivate the time travelling 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 visualisation, we could let the user choose between different statistics (cases, death, recovered, in absolute value or relative to 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.

1.3 Graphs for Switzerland (Hospitalisation stats)

We want to create graphs in order to be able to compare the evolution canton by canton. The aim of this approach is 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 hospitalised people. We will plot the amount 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 number of hospitalised people per canton in ICU ward, normal ward and people with ventilator.

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

1.4 Comparison with seasonal flu

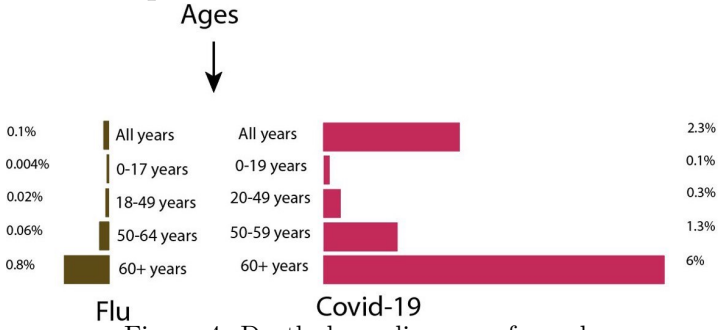


Figure 4: Death depending age of people

At the beginning of the crisis, several people (political and scientific) were saying that this virus will not lead to more death than a normal seasonal flu. We want to do a graph to show the difference for Switzerland to determine whether they were right or not. In this visualisation, we will show the number of detected cases for both diseases in order to compare them. We will have the values for the seasonal flu in one side and for Covid-19 in the other one.

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

1.5 Symptoms

Using the textual symptoms in our dataset, we want to build a word cloud to identify the most common ones. We hope that this visualisation raises awareness on the principal symptoms and show the user if he/she should take the ones he has seriously.

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

Figure 5: Word Cloud

2 Further idea

If time permits, we want to make visualisation, 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 child won't contaminated anymore. This visualisation could help, understand the huge impact of lock-down.