Interactive Swiss Sustainability Map - Milestone 3

Website: https://com-480-data-visualization.github.io/datavis-project-2022-mng/

Problematic:

The aim of our visualization is to provide Swiss citizens with a better view of how different geographical areas of Switzerland are performing at improving their sustainability. Our aim is to encourage viewers to consider switching to electric transportation, installing solar power on roofs, or integrating greener heating solutions. Hopefully, we can motivate citizens to further improve the sustainability of their communities to bring Switzerland to a greener future.

On our website we provide the following visualizations:

- Display on **Bar Charts** sustainability statistics and metrics of Switzerland alongside an **interactive map**. Users are able to click on a canton and/or a commune to zoom in and see statistics specific to the selected region. A comparison of the statistics of the selected region and Switzerland allows viewers to see how the selected region is performing compared to the rest
- 3 **Time Series** plots showing the progress of each canton in the past year on each sustainability metric. The user can show and hide the time series of each canton, which helps him compare the performance of each region.
- A **Heatmap** showing how cantons compare amongst each other on each sustainability metric

Path to Obtaining Final Results:

1) Extracting and Preparing the Data for the website:

The first task was to prepare our 3 datasets separately:

- **SwissBoundaries3D:** The swiss boundaries dataset contains the borders of all swiss cantons and municipalities. The data is in an ESRI Shapefile format. Therefore, the preprocessing is to reorganize the dataset in a GeoJSON format. The GeoJSON format is more suitable for drawing the map with d3. We successfully preprocessed the data in the "swissBoundaries3D.ipynb" notebook. By the end of the processing, we have 2 dictionaries in a GeoJSON format: 1 containing the borders of all the cantons and the other containing the borders of all the communes

- **EnergyReporter:** The energy reporter dataset was made by geoimpact AG, WWF Switzerland, and EnergySwiss using data from swisstopo (Federal Office of Topography). It contains government data that describes how advanced each swiss canton and commune is in terms of sustainability. The dataset has 3 indicators:

  a. ***Renewable_heating_share***: share of buildings with renewable heat technologies
  b. ***Electric_car_share***: share of electric cars
  c. ***Solar_potential_usage***: share of roof area that is economically and technologically exploitable that already has solar panels installed

  The data is very clean. It's in a CSV format so we simply loaded it into a pandas dataframe

- **Swiss Population:** This dataset is provided by the Swiss government (link). Preprocessing has to be done for this dataset because the layout of the file is made to be read in an excel spreadsheet. This preprocessing is done in the 'swiss_population.ipynb' notebook. By the end of the processing, we have a dictionary containing the population of each Swiss canton and commune

Once all 3 datasets were preprocessed, it was time to merge them together in a GeoJSON format so that our interactive map could easily access all the data. Our swiss boundaries data has already been prepared in a GeoJSON format so the only thing left to do is to add the 2 other datasets to it by using the canton_id and the commune_id. In the end, we are left, with two GeoJSON files: 1 for the data of communes and one for the data of cantons. The challenge in merging the dataset was that the datasets didn't always have the same unique identifiers (e.g. the energy reporter dataset uses the canton's abbreviation as a unique identifier, the population dataset uses the canton's name and the swiss boundaries dataset uses a canton id).

Finally, the last challenge was to reduce the size of the GeoJSON file to improve the interactive map's performance. In fact, the description of the swiss borders was way too detailed. To reduce the number of coordinates of each border we used the TopoJSON library and format. This allowed us to quantize our borders. The creator of TopoJSON describes quantization as such: " Quantization removes information by reducing the precision of each coordinate, effectively snapping each point to a regular grid. This reduces the size of the generated TopoJSON file because each coordinate is represented as an integer (such as between 0 and 9,999) with fewer digits". Once we did this, we reconverted our data in GeoJSON. Our data is finally ready for our website!

2) D3 Interactive Map:

After all the data for the map was ready (geometries, features), we were able to inject it into a D3 SVG component, which allowed us to have the basic outline of the map. Of course, interaction with the map is a huge benefit. Because of this we decided to allow 'mouseover' and click interaction with the map which allows the user to zoom into every single canton and commune.

The first challenge we encountered at this stage was to find a simple and intuitive way for the user to zoom out of a region. We decided that the best solution was to wrap the entire map in a rectangle. This way, the user must simply click outside of the selected region to go back to the default map's default view (Figure 2).

Regarding performance, a second challenge we encountered is that the GeoJSON of the cantons and of the communes were overlapping each other. Showing both the cantons and the communes from the start would have been a very disorganized design solution, other than being very slow. Because of this, we found a quick solution in which at the start only the canton data is shown (Figure 2). Then when clicking on a canton, it gets hidden and the communes of that canton are displayed (Figure 4). Then, when clicking back to the default view the communes are hidden and the canton is reshown. This simplified our map and increased its performance.

A third challenge is related to zooming. Zooming in d3js is rather simple to implement, but when it comes to double zooming (one for the canton, and then deeper for each commune) it ends up being very tricky and messy. We decided to simplify the user experience by only allowing a first zoom on the cantons but allowing the user to mouse over each commune and when he selects a commune, it turns red (Figure 4). Moreover, the centroid of each canton challenged us in finding the correct translation for a zoom to be straight and smooth.

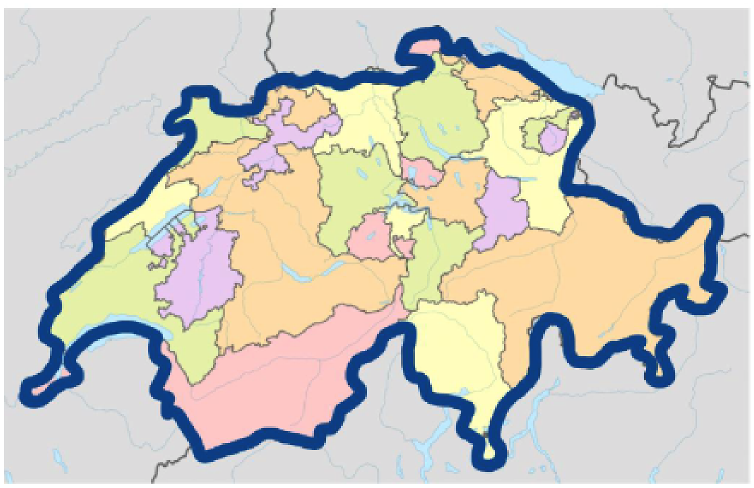

Figure 1. Milestone 1 Sketch of Interactive Map (default view)

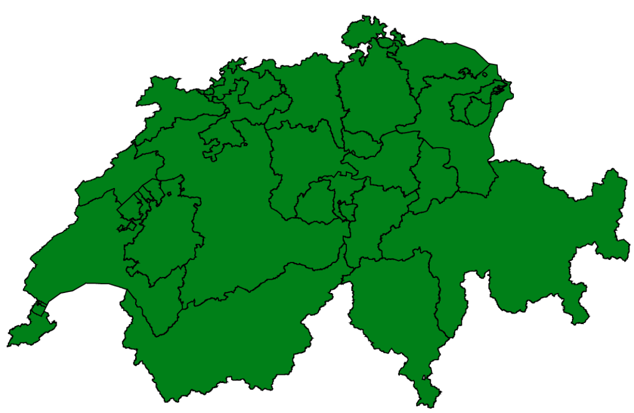

Figure 2. Map On Our Website (default view)

Figure 3. Milestone 1 Sketch of Interactive Map (view when a commune is selected)

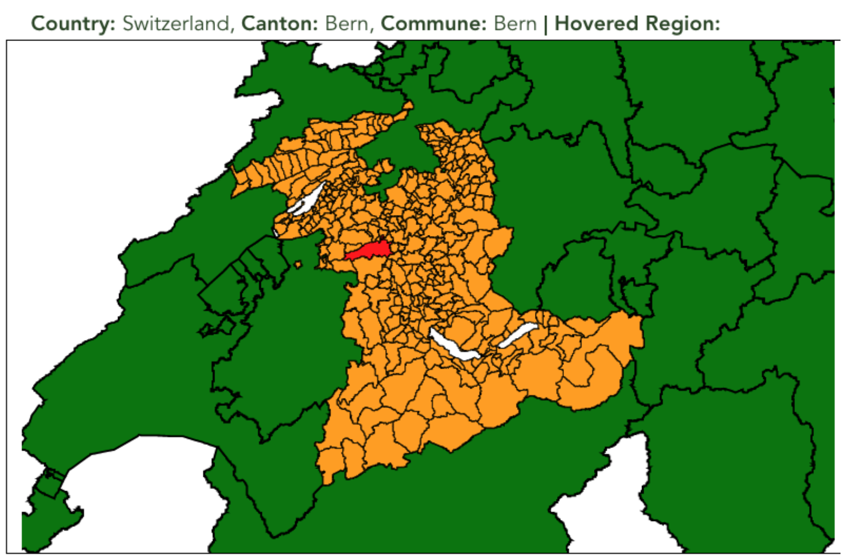

Figure 4. Map On Our Website (view when a commune is selected)

3) Bar Charts

In our initial plans, we wanted to have pie charts to show the sustainability metrics of the region selected with the d3 interactive map (Figure 5). However, we quickly realized that this was not a great idea. Firstly because the metrics in some regions are very small so the user wouldn't even be able to see its value on the pie chart. On the other hand, bar charts allow us to have logarithmic scales which are great for visualizing and comparing data with different orders of magnitude. Secondly, we believe it's easier for a user to compare values on a bar chart than values on a pie chart.

There are 6 bar charts (Figure 6):

- 3 charts are simply the raw sustainability metrics of Switzerland, the canton selected and the commune selected
- The 3 other are the sustainability metrics rescaled in function of the region's population. More specifically each chart displays the following 2 metrics:
    a. How much the canton contributes to the sustainability metric of switzerland. For example, for the electric car share metric, the formula is the following:

$$swiss\%_{electric-car-share} = \frac{electric-car-share_{canton} * population_{canton}}{electric-car-share_{switzerland} * population_{switzerland}}$$

    b. How much the commune contributes to the the sustainability metric of its canton. For example, for the electric car share metric, the formula is the following:

$$canton\%_{electric-car-share} = \frac{electric-car-share_{commune} * population_{commune}}{electric-car-share_{canton} * population_{canton}}$$

The bar charts are implemented using ApexCharts.js. Initially, the charts contain sustainability metrics of Switzerland as a whole. Then when a user selects a region on the map, the sustainability metrics of that region appears on the bar chart too (Figure 6). Once the region is unselected, the region's metrics disappear. The challenge in this part was to link the map to the bar charts and update the chart's data in real time. To implement this, we created an object that had the barchart's data. Then once a region was selected on the map, we changed the values in this object. Finally we added a listener to this object which triggered the update of the charts whenever the object's data was modified.

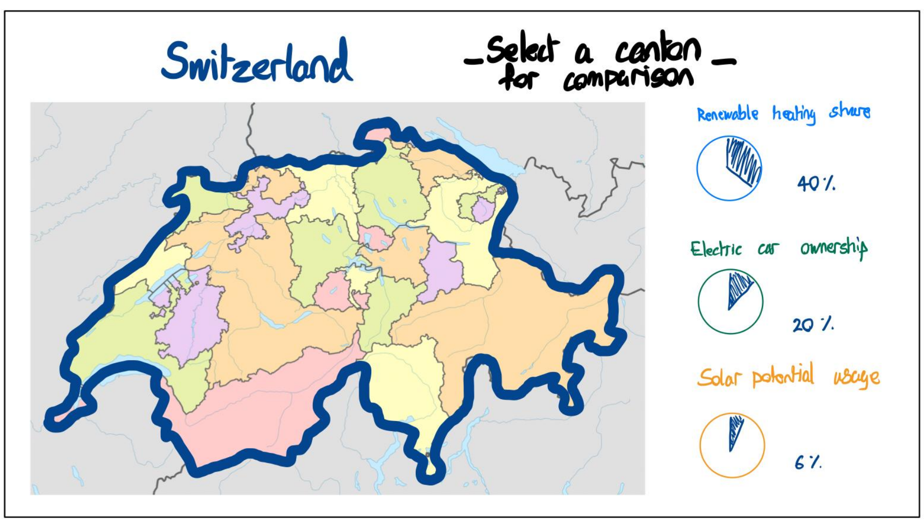

Figure 5. Milestone 1 Sketch of Interactive Map with Pie Charts (default view)

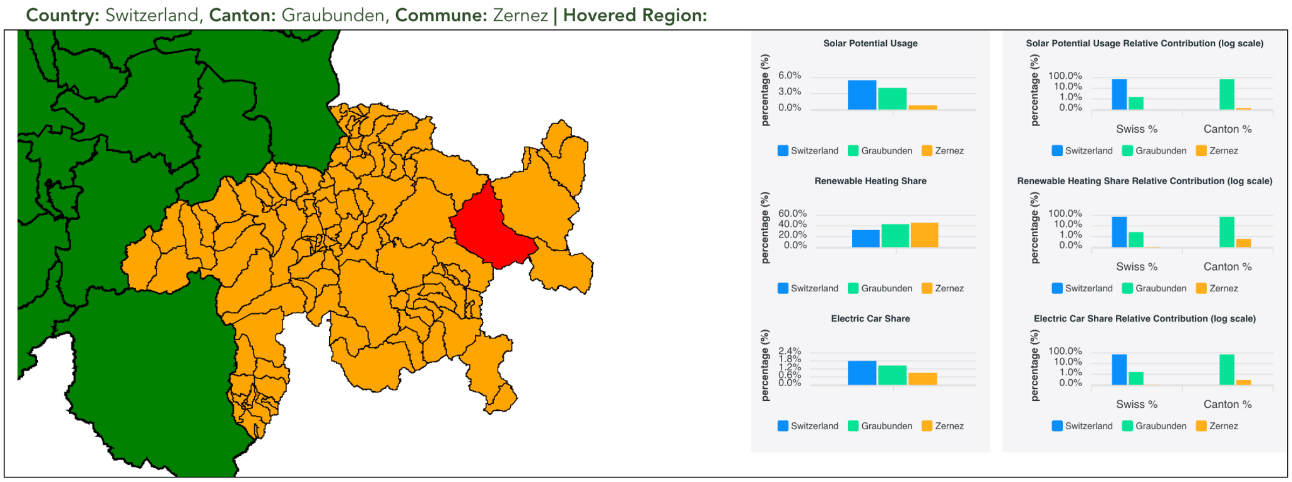

Figure 6. Map On Our Website with Bar Charts (view when a commune is selected)

4) Timeseries

Alongside the bar charts and interactive map, we found it useful to show the progress of the three main indicator percentages in the last year (because our dataset only has the data for this last year). We have implemented it in a way that allows the user to compare different cantons depending on the needs. Initially, we weren't sure how many cantons we should have included in the default view. After reflection, we decided to display one canton per region (French, German and Italian regions) (Figure 8). Additionally, we decided to add a selection button to allow the user to show the time series of all cantons in a single click (Figure 8), and the possibility to hide and show cantons for easier comparison. We ended up with three-time series, one for each sustainability indicator, with on the x-axis the percentage for that indicator.

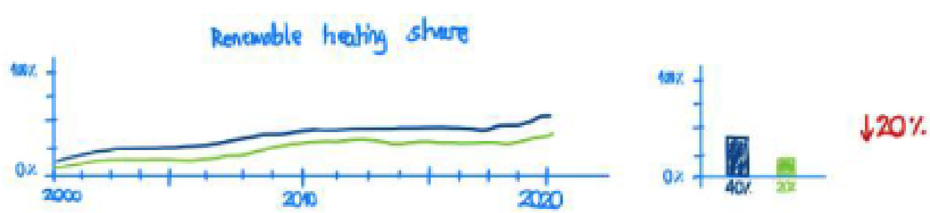

Figure 7. Time Series Plots Initial Sketch

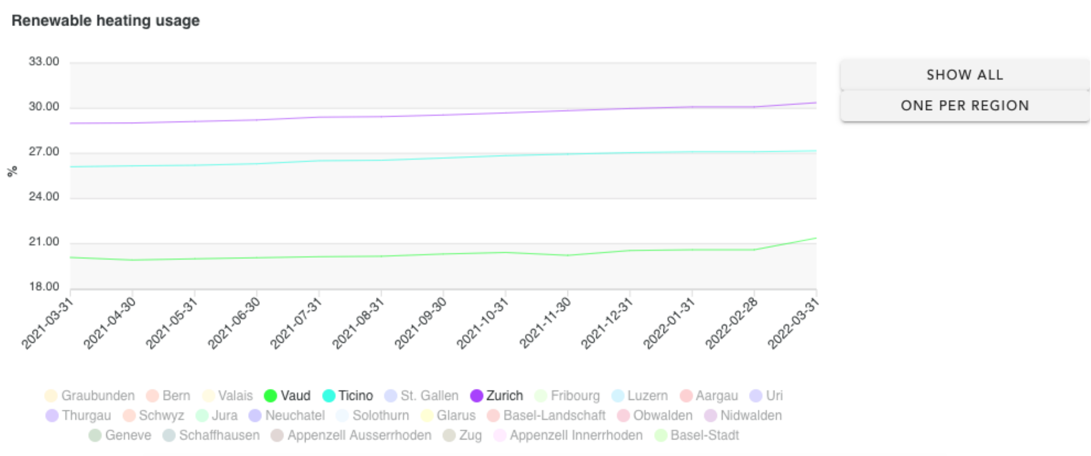

Figure 8. Time Series Plots on Our Website

5) Heatmap

Having a general overview of each indicator is also very helpful when you want to quickly compare cantons and see where each one stands in terms of sustainability. Our heatmap (one for each indicator) allows the user to reach exactly this. We have developed it in a similar way we developed the interactive map, with the difference that we made it fully static and used only an interpolation color scale for the whole map, alongside a legend (Figure 9).
In regard to challenges, we encountered the first one in the decision of an intuitive color selection. We started having a simple scale from white to dark green but then found out that a color selection from red to green would suit more the topic (red represents bad and green represents good in the context of climate change).
Secondly, we min-max scaled our sustainability indicators to cover the full range of the color scale. This ended up being very helpful for the comparison.
Lastly, by having a static map in which it is not possible to mouse over the cantons, we realized the possible misunderstanding for the user in the canton location. Because of this, we decided to alter the initial data, add the abbreviation for each canton, and include it in its centroid.

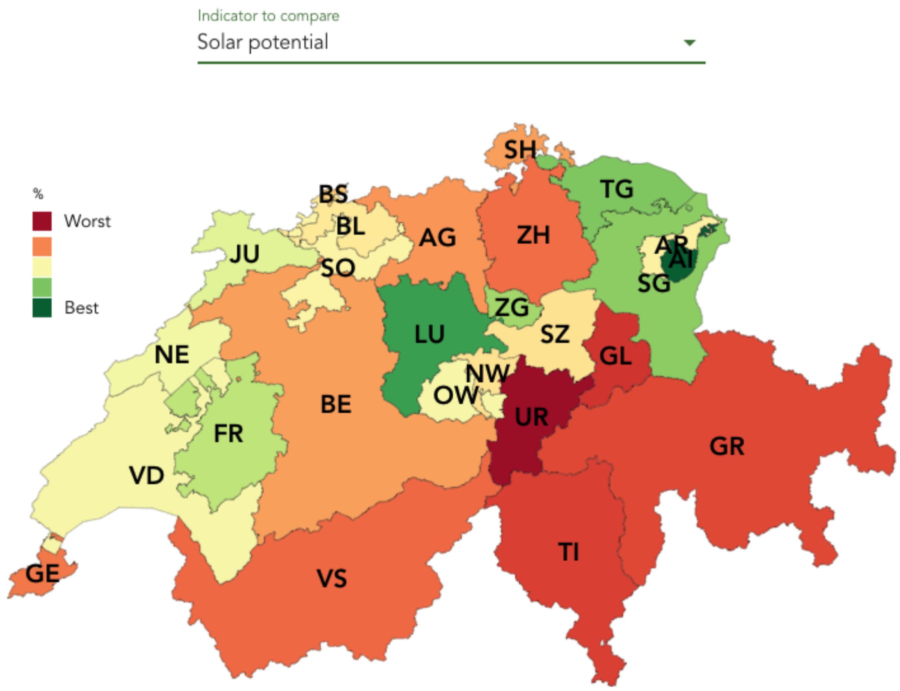

Figure 9. Website's Heatmap of Switzerland For Each Sustainability Metric

Peer assessment:

To better understand the contribution of each member to the project we highlight the contributions for each of the tasks:

1. Idea for the project:
    a. Related work (Gioele)
    b. Datasets (Nicolas)
    c. Feasibility (Team)
2. Design:
    a. Interactive map (Micheal)
    b. Barcharts (Micheal)
    c. Timeseries (Gioele)
    d. Heatmap (Gioele)
3. Data extraction and manipulation
    a. Extraction (Nicolas)
    b. Conversion (Nicolas)
    c. Putting information together (Nicolas)
4. Implementation
    a. Website setup (Gioele)
    b. Interactive map (Nicolas and Gioele)
    c. Barcharts (Nicolas and Micheal)

d. Timeseries (Micheal)
e. Heatmap (Gioele)
f. Improvements in design - CSS,HTML (Gioele)