Data Visualization - Process Book

KIRELL BENZI

Ajaloeian Ahmad
Yaghini Mohammed
Köse Süha Kağan

Fall 2017

Last Updated: 26 November 2017

1. Abstract

The popular media is an (imperfect) mirror of the society. In this project, we want to study the effect of social trends and events of the past century on TV/Radio broadcasts. The 20th century saw the invention and ubiquity of the broadcast media: first radio and then television. At the same time, many international events (World Wars, the rise and the fall of Communism, the formation of European Union, etc) as well as national incidents (sexual revolution of 1960's, the establishment of women voting right) took place in this period. All of these events had profound impact on the society, and as a result on broadcast media such as TV and Radio. Having access to a wealth of information about TV/Radio emissions all the way from 1930's until today, we believe we can track this events and by means of a data story and a convincing visualization, highlight the events of the decades past.

2. Project Overview

2.1. Motivation

Our main motivation is the sheer amount and scope of the structured data provided by the Swiss public broadcasting organization, RTS. Although the majority of the data is understandably recorded in recent years, it covers an expansive temporal span (from 1900's until today).
More importantly, given that in March 2017 the Swiss will decide by popular vote whether to abolish the "Billag" (which also supports RTS), such a study is very much relevant and timely. An interactive and easy-to-understand visualization has the potential to showcase the added value of this public organization over the course of the past century.

2.2. Target Audience

Given the last point, our target audience is the general public. Ideally, we want our results to be accessible to the electorate, while being accurate and meaningful for the informed citizen.

2.3. Inspirations

This is an inspiring visualization/data story made for a campaign. But since, we have to dealth with geographical and temporal data, something like this would be ideal example.

2.4. Visualization

We would like to have an interactive map with a timeline (to keep track of evolution of the map through time) that shows the how broadcasts connect different places

together. In order to emphasise the importance of RTS/SRG/SSR in bringing the Swiss together, we have thought of a map where edges (a field map) connect the areas of interest together. (Note that such visualization is only possible for the records that have more than one “geographical theme” (which is a feature in the dataset) associated with them.)

Also, we are planing to demonstrate with which topics that cantons are connected, by displaying the key words related to the topics for each of the connection. By doing that, we also plan to demonstrate which topics are mentioned in broadcasts through years.

3. Dataset and Processing Steps

As we mentioned earlier, we are going to mainly use the RTS dataset. However, to complement our analyses, we will potentially make use of Spinn3r, other Swiss Open datasets, and Wikimedia data as well.

Highly convenient API of the RTS dataset can be found in here. The page also contains information about how to use the API.

The required data for our project is that cantons related to each of the broadcast (from 1900's to 2000's) and the content (or category) of these broadcasts both of which are provided within RTS dataset. For data accusation we are working with another team (one of our members is also part of this team) from the class Applied Data Analysis. They agreed on providing us the clean data of broadcasts. There might be some missing or ambiguous fields related to broadcasts so, it requires cleaning and filling the missing by data analysis techniques before the data can become useful for our visualization purposes. (Since the Data Visualization class doesn't require us to do these kind of analysis, it is a win-win for both teams.) Collaborating team's Github repository can be found here.

There couple more steps to mention about data processing. First, although the name of the cantons related to each broadcast is provided, we require to retrieve the corresponding geographical locations to use in our map. It is also required to have a .geojson file of to retrieve the border of Switzerland and cantons.

4. Development Processes

We've organized the project (and the process book) such that it consist of three major parts which are data acquisition, map and timeline construction.

4.1 Map Construction

The first step to decide was which tool to use for building the map. We decide to use Leaflet since it is one of the leading open-source JavaScript library for interactive

maps (and also the one we've talked about in the class). Creating the initial map with Leaflet was highly convenient. We initially, created a layer and appended to the map along with the coordinates of Switzerland.

```
var initialLayer =
  L.tileLayer('https://api.tiles.mapbox.com/v4/{id}/{z}/{x}/{y}.png?access_token=' + mapboxAccessToken, {
    id: 'mapbox.light'});
var map = L.map('map', {center: [46.818, 8.227], zoom: 8, layers: [initialLayer]});
```

Then first step was to build plain map of Switzerland which only displays the borders of its cantons as shown in Figure 1. Utilized geojson data can be found here and the code was fairly simple.

```
var geojson = L.geoJson(geojsonData).addTo(map);
```

Figure 1: Plain Switzerland Map

The next step is to color the map since color schemes might be a good representation of how well-connected a cantons is connected with others (via broadcasts). In this stage, we didn't have the real data so we created an arbitrary data and color the map accordingly. The resulting map is in Figure 2. We coded a function (called `getColor`) to map values into corresponding colors. In order to find a suitable color scheme, we used a pretty helpful website called Carto. Then we added the style function to the map.

```
function styleFunction(feature) {
  return {
      fillColor: getColor(feature.properties.density),
      weight: 1,
      opacity: 1,
      color: 'black',
      fillOpacity: 1
  };
}
geojson = L.geoJson(geojsonData, {style: styleFunction}).addTo(map);
```

One of the main property of our visualization will be the lines connecting cantons according to broadcasts. But again, since we didn't have the data yet, we randomly connected each canton to five others. We used `PolyLine` function of Leaflet to create the lines and added them to the map (as in Figure 3). We though that the most convenient way to show the connection lines is when a canton clicked on.

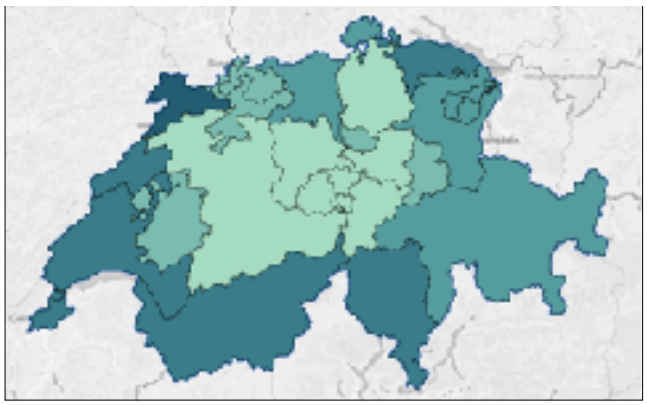

Figure 2: Colored Switzerland Map

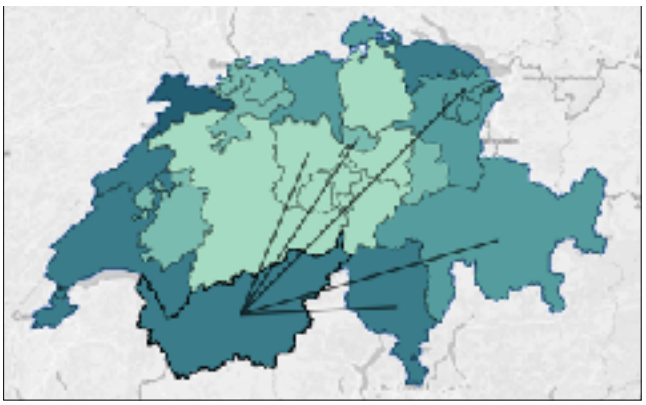

Figure 3: Lines Connecting Cantons

We are planing to serve information about cantons and broadcasts that connect them. Our plan is to display this information inside the line connecting cantons but we made a design decision yet about how exactly we do that. As a prototype, we created a information box on top-right of the map for information display purposes. In addition to that, we though that it might be useful to create a legend for color mapping. The code was relatively long and it can be reached in here. The initial prototype of the map resulted as demonstrated in Figure 4.

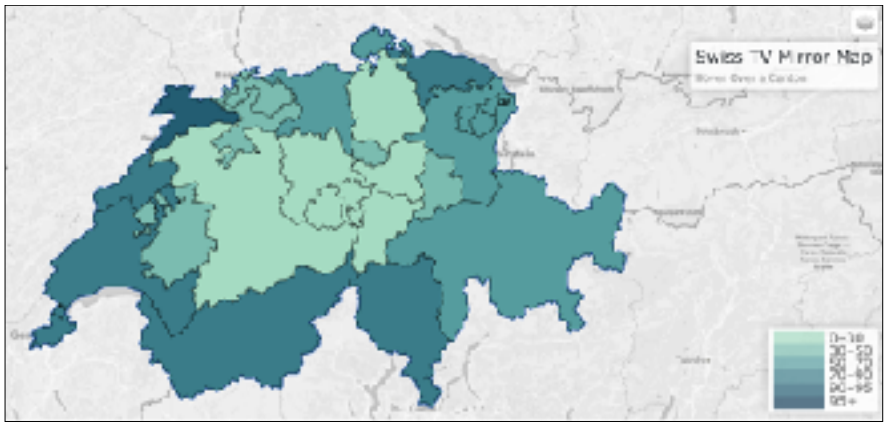

Figure 4: Initial Version of the Map Prototype

We mentioned that in order to emphasise the importance of RTS the Swiss together, we want to draw a map where edges (a field map) connect the areas of interest together. So far, we've only created a map with basic edges connection cantons so as the next step, we wanted to create polygons connecting cantons. First, we've formed a Javascript object that maps location to their coordinates. And created a listener so that when a canton is clicked on 2 other random locations are randomly chosen then, these 3 locations are connected together using polygons. Moreover, in real data there will be more than one connections related to each canton and we wanted to simulate by drawing 2 polygons are when a canton is clicked (with clicked canton itself and 2 randomly chosen corner points for each polygon).The resulting map is demonstrated in Figure 5. Note that since a broadcast in RTS dataset can be related to not only cantons in Switzerland but other places on earth an therefore, our initial polygon drawings are restricted to Switzerland. For the purposes of our visualization (that is to illustrate the connection map of Switzerland), we'r planing to prune the dataset so that it only contains broadcast that are connecting Swiss cantons.

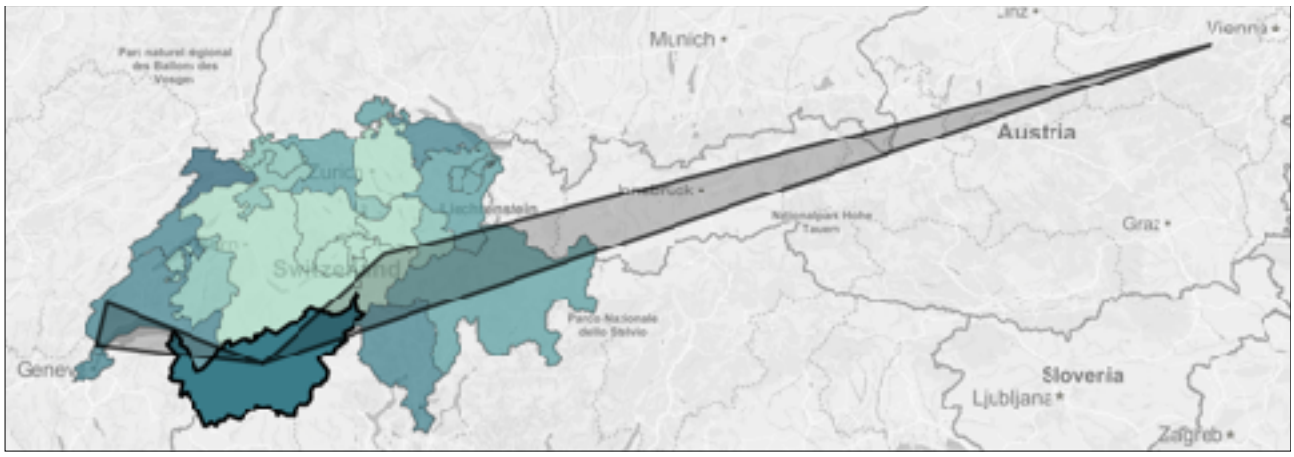

Figure 5: Polygons Showing Connection between Places