Data Visualization (COM-480)

Project Datart

Milestone 2

Abdeljalil MOUSSA
Jeremy BENSOUSSAN
Xiao ZHOU

30th April 2020

Introduction and Project Goals

Following the last milestone's feedback, our project has been modified and we now focus exclusively on France. The purpose of this project is to explore the relationship between the population's features and their votes on elections. Our study begins in 2002 and ends with the latest French presidential election, 2017. Our visualizations explore this topic through multiple lens to obtain the most accurate and pertinent view possible.

Visualizations and Extra Ideas

Events

The basis of a population's ecosystem is the events which affect this population at the national level. Therefore, we wanted to show the different events that occurred in France since 2002. Those events are important enough to change a French citizen's life and therefore are pertinent to our study. This chronological visualization can be seen in figure [1], plotting the most important events on the same line with a histogram plot below in order to show the number of events on the same time period. We also provide a small description of these events as a complement (by hovering the mouse).

Elections

Now that we know a little bit more about France since 2002, it remains necessary to inform about the French elections results. For this part, we decided to create a map of France (figure [2]) with the results of the presidential elections (the most important election in France). This is the core visualization of the project.
Since each region/department in France votes differently and because of that, the results are rarely homogeneous, we decided to show the differences in votes between the regions. In this sense, the map shows the results of the first and second round of the election, specifying the results for the two best candidates. It also shows the degree of support to the winning candidate during the second round.

Population

Now that we discovered how each part of France votes, it's time to show the population's ecosystem. A first pertinent visualization (figure [4] is a responsive scatter-plot which combines several features in the same plot in order to give as much information as possible to the viewer. The minimum viable product of this plot is the number of voters (x-axis) depending on the social contribution of the population (y-axis). However, we added other metrics such as the size of the bubbles (total population), color for the type of election (presidential, legislative or cantonale) and a box with the name of the region, and the year of course. Additionally, an extra idea for this plot, would be either to select the regions we want to compare via a button, or to have a "zoom" feature via an area selection tool such as "brush".

In parallel, The figure [3] displays a quadrant visualization plotting the correlation between the profile of the departments and their votes to a political party. Since the statistical analysis and the preprocessing for this visualisation is heavy, the MVP could only be a scatter-plot where the point are the parties mapped on two axes (revenues and unemployment rate). After this MVP completed, we could offer to choose more features (crime rate, gender...) or to hover the points (ie. the political parties) to have more information on those metrics.
We could even think about a "tooltip" describing the three departments supporting the most this party. This would allow the reader to seek for the revenues breakdown of these departments on the precedent visualization.

Lastly, we can introduce an easy-going visualization (figure [5] which plots the profile of selected regions in terms of taxes (or revenues) over a selected number of years. The most valuable product would be a bubble chart going from 2002 to 2015 showing the taxes' evolution for some selected regions, and the selected colors represent the year. Also, the bubbles have a "collision effect" which gives a fun insight about the data. An extra idea would be to select direclty the years and regions via 2 separate buttons.

Visualization Tools

For these first visualizations, we mainly focused on using **D3.js** in its versions 3.5.5 or version 5. However, we used many visualization tools in combination with D3 in order to have more control on our visualization. You will find them below :

- bootstrap and bootstrap-select
- Jquery
- "layout.timeline" package from D3
- "tip" package from D3
- "Iwanthue" color palette picker
- Libraries from Python for data preprocessing : pandas, numpy, networkx, glob, os

Moreover, for this project we used several lectures in order to build our basic visualization and then enhance them. For instance we can list the lectures that were truly useful for us :

- Lectures about "Interactions" "Perception colors" because we want to make our visualizations as interactive and easy-to-use as possible with pleasant color palettes.
- Lecture about "Map" since one of our main visualizations is an interactive map of France
- We are also looking forward to knowing about the lectures on "Graphs" and "Tabular Data" because we want to exploit other innovative ways to make visualizations (such as stacked bar charts, steamgraphs or even sunburst graphs)

Functional Project Prototype

We also created an initial website running with the basic skeleton of the visualization in the following link : `https://jeremybensoussan.github.io/`

A Appendix : Examples

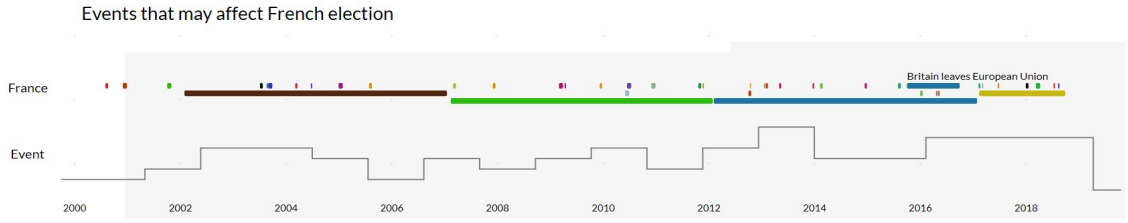

Figure 1: Interactive Chronological Plot

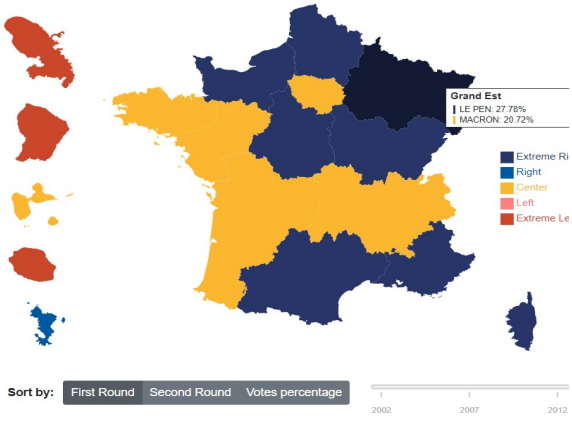

Figure 2: Responsive Voting Map of France

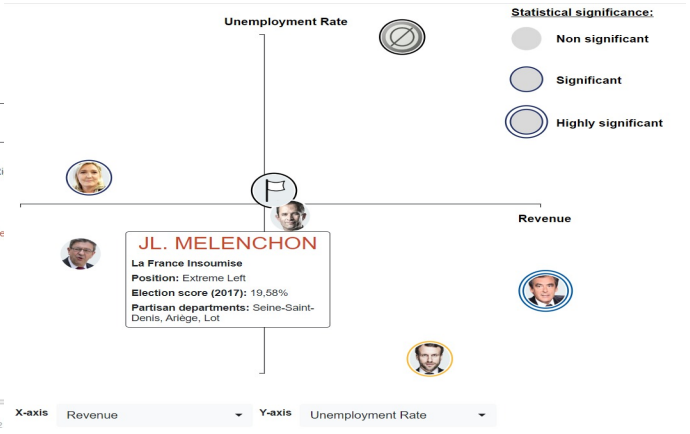

Figure 3: Responsive Statistical Scatter-plot

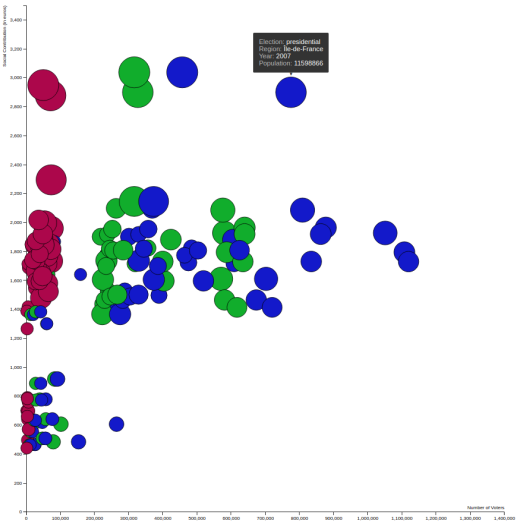

Figure 4: Responsive 2D scatter-plot

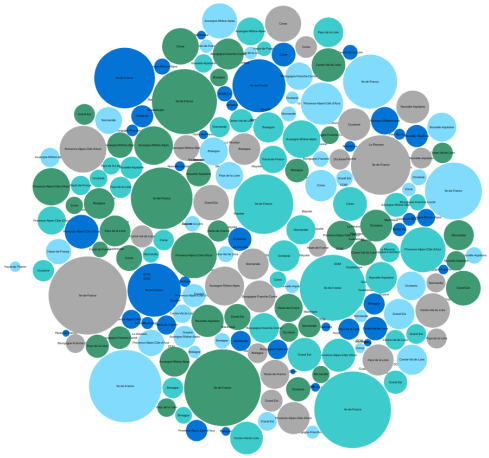

Figure 5: Animated Bubble chart with collision