Craft Beers in the World

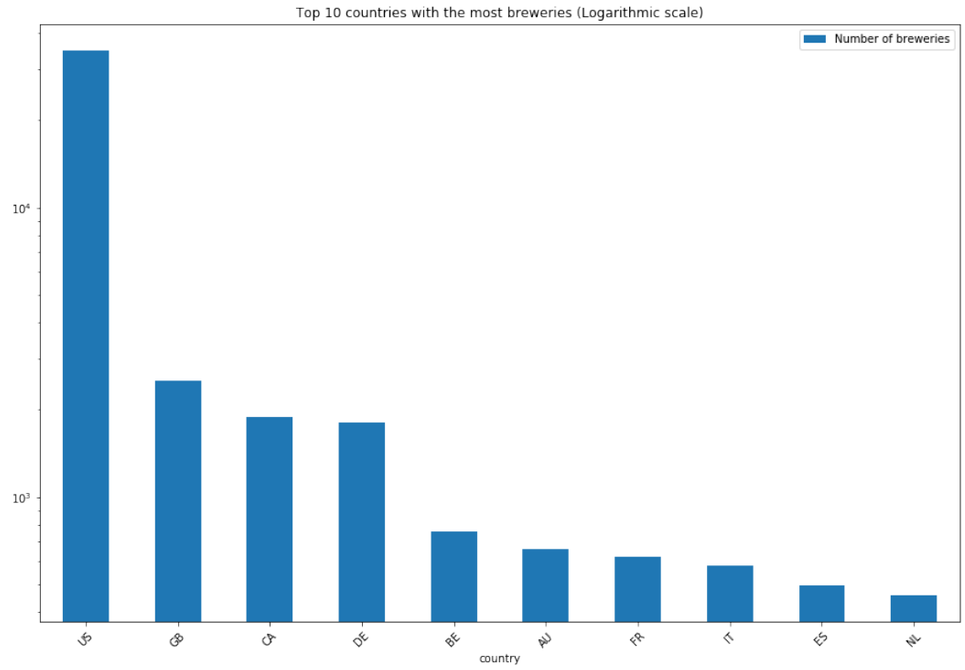
Alexis Dewaele, Robin Leurent, Loïc Vandenberghe

1 How it all started

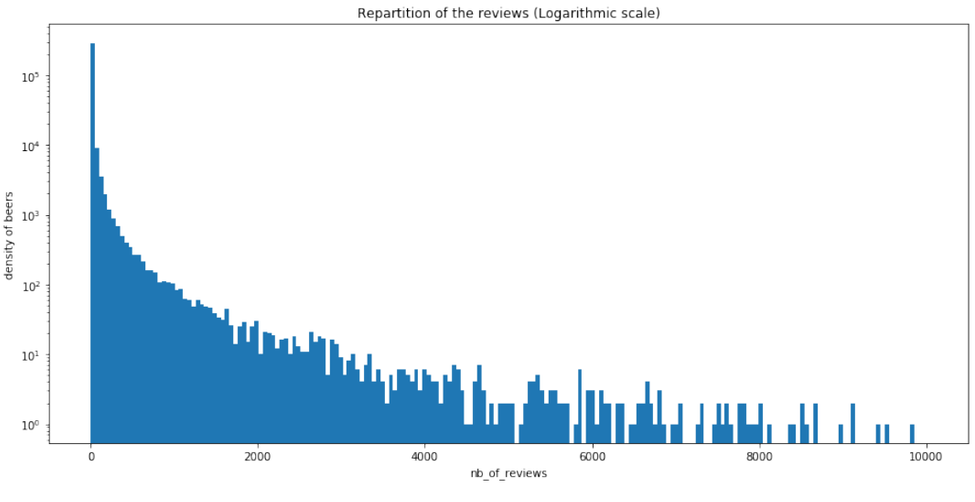
As students in EPFL, we are often led to the campus bar Satellite. This is something that is shared by many students and people working at EPFL and thus we felt that craft beers would be an interesting visualization subject.
This is why we wanted to build a visualization tool that will help fellow craft beer enthusiasts discover locally crafted beers. To that effect, thanks to the recommendations of our supervising teaching assistant Jean-Baptiste Cordonnier, we contacted Satellite to obtain their data but because of client confidentiality, we received a negative answer. We found a dataset from BeerAdvocate.com, a website from beer lovers for beer lovers.

2 Data exploration

Our dataset contains information from 358 000 beers and 50 000 breweries. The breweries are specified via their name, their country, their city and their type. The beers are described with their name, their brewery, their alcohol by volume and their type and some notes about it. A substantial part of the dataset is 9 000 000 reviews where beer amateurs rated the smell, look, taste and the feel of a given beer, optionally giving a textual description of the rated beverage.



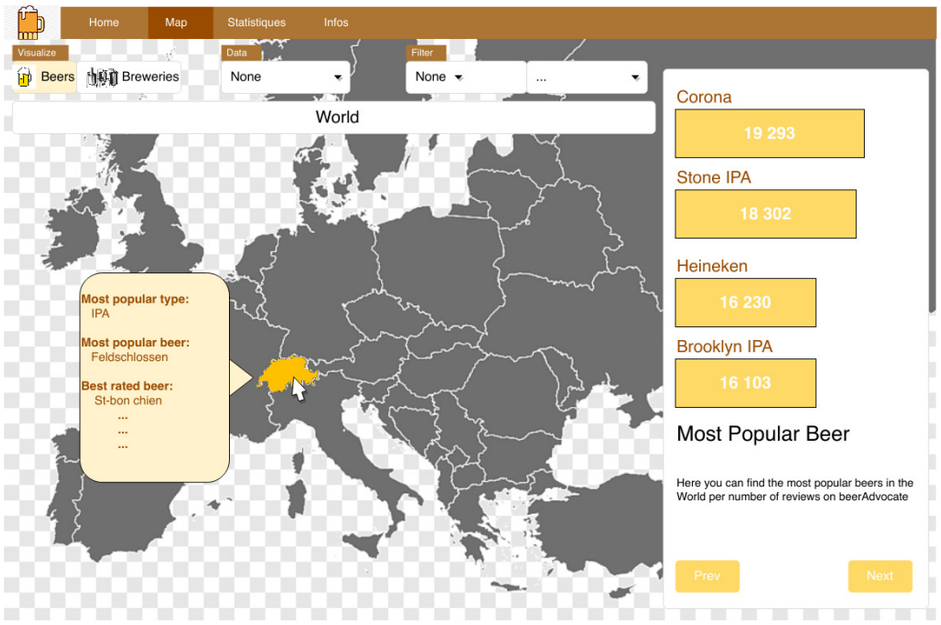
As seen in the previous figure, we quickly realised that the country beer distribution was very unbalanced and that could give us problems later.



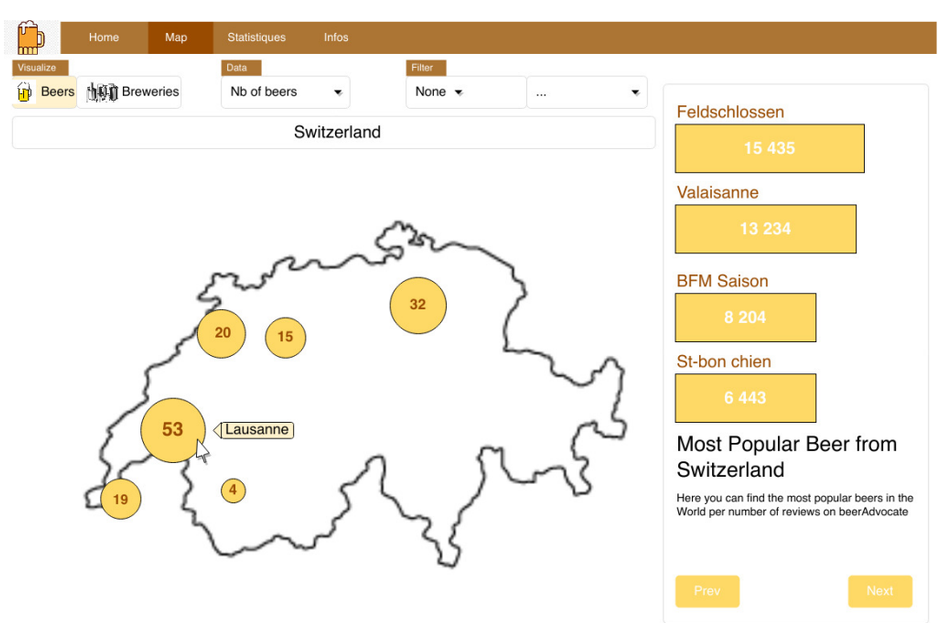
We also realised that most of the beers only have really few number of reviews (between 0 and 50) since only a few beers have more than 2000 reviews. We would need to also take into account that beers having few reviews would not be representative of the quality of the beer.

Lastly, with almost 112 different beer types, with several very close to each other such as the Double IPA or the IPA, we also needed to create super classes for the beer types. For example, a class regrouping all the different kinds of Ales, the Stouts, the IPAs (Indian Pale Ales), etc. We managed to get 11 types regrouping 60% of our data.

3 First desired outcome



On a tab Map, we wanted to have several possible filters, such as displaying stats about beers or breweries, displaying only stats about a chosen type of beer (Pale Ale on this example), what kind of data we want to display on each country (on this example it is the number of beer), etc... On the right side are general world statistics for the chosen filter. For this example we have only Europe, but as we have data for the whole world we want to be able to see the whole world and zoom in to region of interest. By hovering over a country, we want to display the top rated beers as well as the most popular beer and other information...

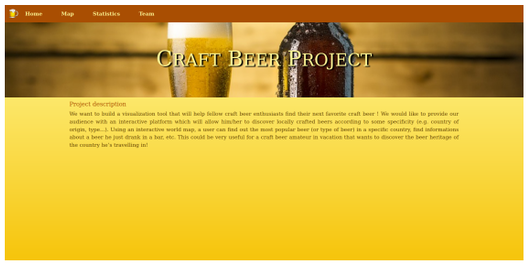


By clicking on a country, we wanted to display the most popular or top rated beers per city. On the right side, the world's general statistics would be modified to the country's general statistics. We can still modify the filters on the country's map. We also would like in the Statistics tab to be able to search for a particular beer and get all the stats about it (number of reviews, alcohol by volume (ABV),max ABV, ...) We wanted to have several pages, with a home page, the map page, the statistics page, and one describing our team.

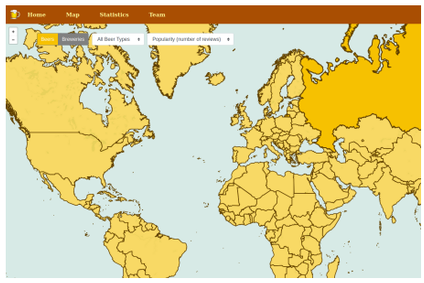
4 The Journey

4.1 Multiple pages, Map visualization and Filters

We started by creating multiples pages with a navigation bar, and a map that will serve our visualization. For the map, we used the Leaflet library combined with the Mapbox API as well a GeoJson file of country borders *(see section 7)*. We didn't face any problems during this task. We also created dropdown menus using the Bootstrap library to create our first filters.



(a) The first implementation of the navigation bar



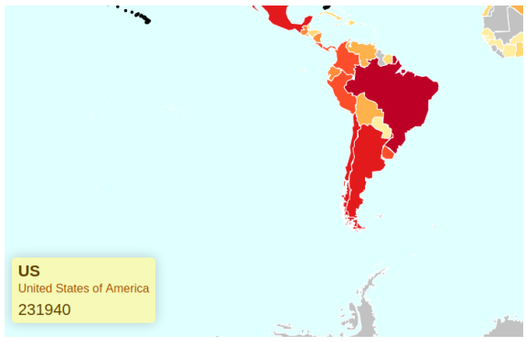
(b) The first map and filters visualization

Later, we decided to remove all the pages except the map and the team page, since our project is mainly this map. The new design for the search bar will be discussed in Section 4.6.

4.2 Hovertool and info card

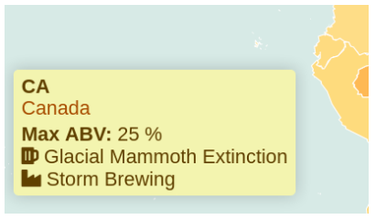
We wanted the hovertool to be displayed on a country when hovering over it. We thus created this first hovertool as seen on the left side of the following image.

(a) The first implementation



(b) The second implementation

The problem with this first implementation was that the popup was following the mouse. Therefore, if the user was moving his mouse upward, the mouse was getting over the popup, then the mouseover listener for the country would think the mouse was off the country, then the popup would flash and lag. Hence we designed a second hovertool that is located on the left down corner. Finally we improved this second one by adding more information about the data when possible.



(a) Latest Hovertool

(b) Information Panel

For the information card on the right side of the website, we wanted at first to only display data about either the world or a specific country. But thank to our TA, we shifted to writing some facts about beer in the world view, and fun facts about beer in specific countries. Since we collected them by hand, we couldn't do it for every single country in the world, but we did it for the most significant ones, i.e. Switzerland, France, Belgium, Great Britain and Germany. For the other countries, we only display the beer with the highest ABV, the number of beers and the best rated beer in the country. For example a fun fact would be : in 2001 in Belgium, in an effort to reduce childhood obesity, a proposal was made to serve low-alcohol beers (between 1.5% and 2.5%) to school children instead of sugary soft drinks.

4.3 Data visualization on the world map

When we first displayed our data on the map, we created a color scale that followed the distribution of the data. In spite of having a nice color variation on the map and after discussing with our TA, we understood that this was misleading and not accurate, as we had colors representing intervals having extremely different sizes.

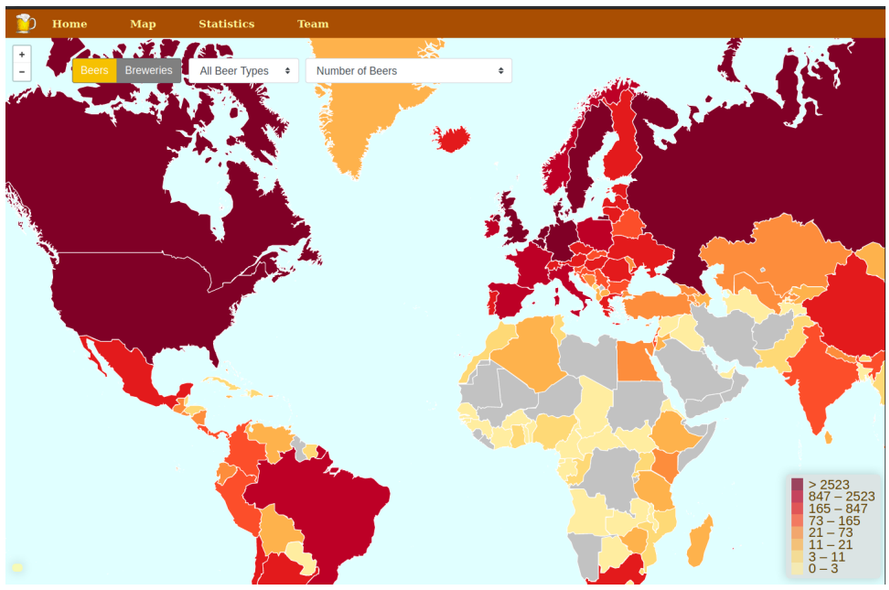


Figure 4: The misleading color scale

For example, we can see on figure 4 intervals of size 3 and 1500. We hence decided to go for a linear color scale going from the minimum value on the map to the maximum one, giving result in fig 5. We also changed the colors to make it looks more like beers.

4.4 Word cloud

We wanted to help our users understanding the characteristics of all different beer types. Moreover, some beer types were unknown even to us, such as the

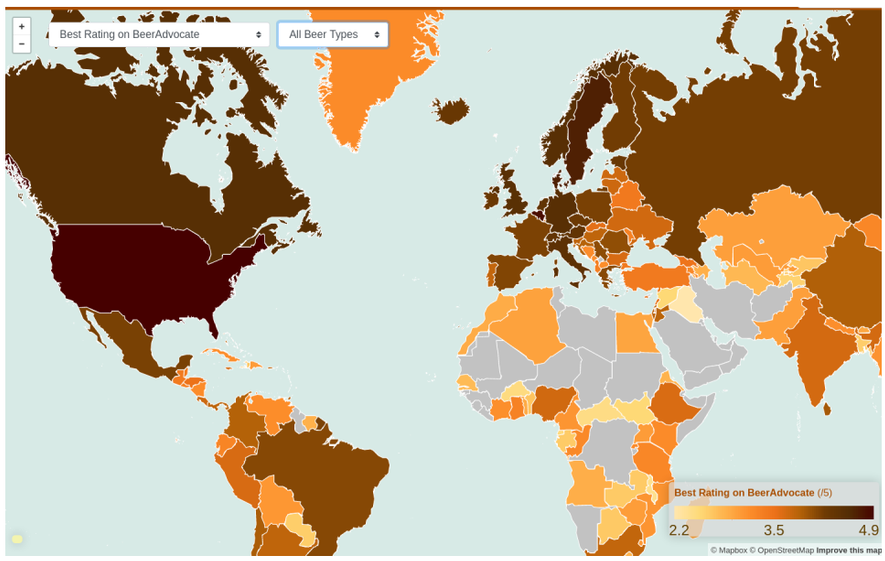


Figure 5: The new color scale

Barleywines. Therefore, we selected the textual reviews for the beers in each beer type we have, and created a word cloud of the most frequent words. The title of the word cloud was not highlighted enough to clearly understand that the words were ones to describe the selected beer type. We hence created a button linked to the one on the map, to be able to modify the beer type also on the word cloud. We used the AnyChart library to create the cloud.

(a) The word cloud with text title

(b) The second with button title

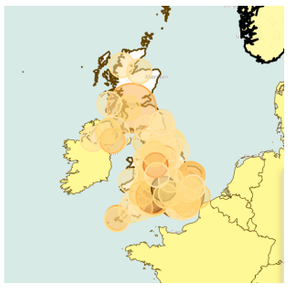
4.5 Data visualization on each country

When a user clicks on a country, the view is zoomed on this country and new information appear. Indeed, we want to display information about each brewery in the country in a fashion that is easily understandable for the user.To that effect we plotted a circle for each brewery onto their location where the color depends on the current filter. The locations for each breweries were collected with GeoNames and OpenStreetMap (when the city couldn't be find on GeoNames). If multiples breweries are on the same city we scaled the circle accordingly.

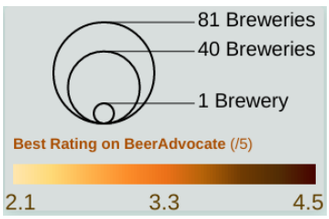
Unfortunately, some country have an extremely high density of brewery and thus the information is unclear (fig. 7a). This is why we started using clusters to regroup breweries, depending on the zoom level as you can see in figure 7c.

As you can see from the figure above, the colors of the clusters are not representative of the data thus we decided to change that. Unfortunately, this was no easy task : to properly color a cluster we collect recursively the data corresponding to the current filter (e.g. Max ABV for Stouts) among the cluster's "children" and combine it into the "parent" cluster. You can see the final result in figure 8.

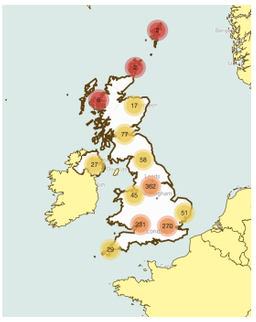
You can also notice that there is a warning at the bottom of the screen : this is because computing the data is too slow if the clusters have more than 1000 breweries.



(a) Country without clusters

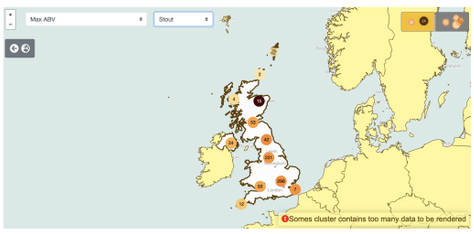


(b) legend for a country

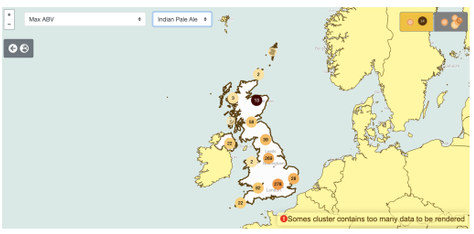


(c) Country with clusters

Figure 7: Cluster Implementation



(a) Clusters for Max ABV and Stouts



(b) Clusters for Max ABV and IPAs

Figure 8: Cluster result

Hence the cluster will be displayed in red instead of the color representing its data. This is only the case for relatively few countries and at a wide zoom level so we thought it was not worth slowing down the browser.

At last, a button on the top right allows to disable clustering, this is particularly useful for countries that does not have a lot of data.

4.6 Search bar

Originally, we had decided to develop a separate page that would allow users to research information about the data we collected. The prototype for this page can be seen figure 9.

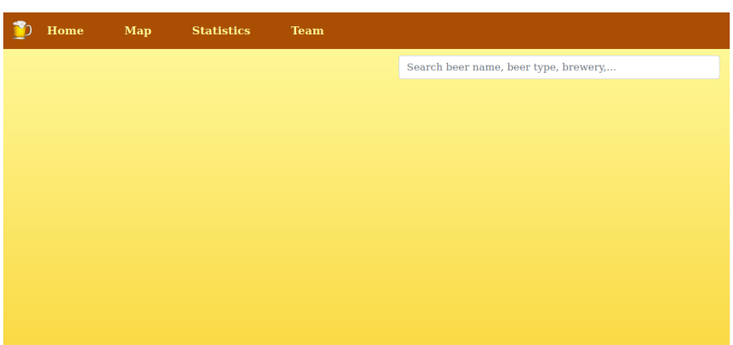


Figure 9: First prototype of a search bar

Upon discussions with our supervising TA, we deleted this page and placed our search bar on the page containing the map. The biggest challenge coming from implementing the search bar was to design an auto-complete function. Indeed, the amount of beers and breweries we collected proved to be counter-productive : so much data was too much to handle for the webpage and it would crash. To solve this problem, we built an associative array as such : the key is composed of the first three letters of the beer's/brewery's name and it points to the set of corresponding beer/brewery. Here is a subset of this table :

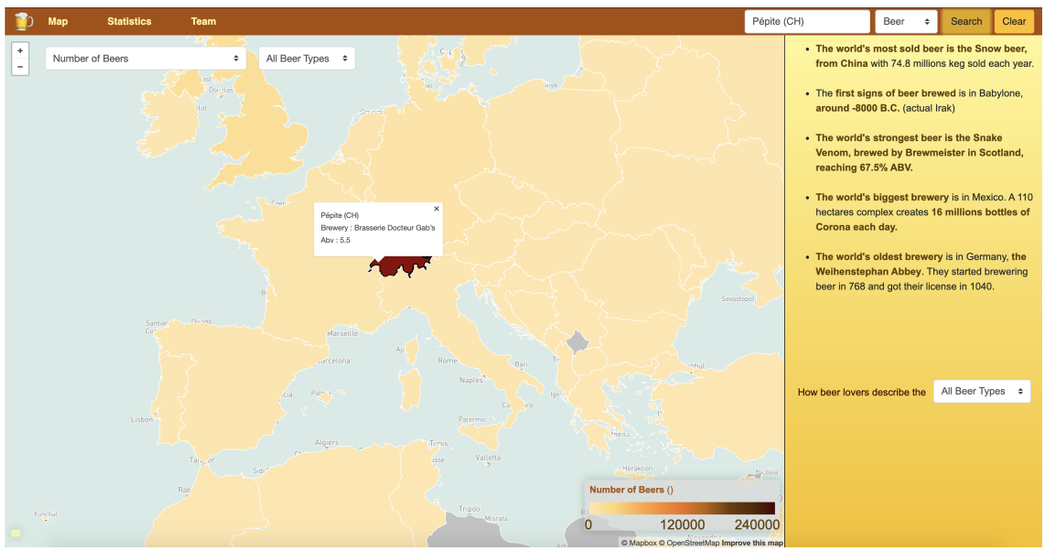
Here is what the auto-complete function looks like :

Beer Key	Beer List
Old	Old Ruffian - Barrel-Aged, Old Red Eye, Old City Is Burning,...
Cor	Corby Ale, Cornish Knocker, Cordova,...
Fel	Feldschlosschen Dunkel Perle, Felstar Good Knight,...

Table 1: Example of our associative array

Figure 10: Auto-complete example

Upon completing a search, a popup appears on the map at the coordinates of where the beer/brewery is located. This popup contains some basic information on the searched beer/brewery. Furthermore the country where the popup is located is highlighted. Here is how it looks :



5 Peer Assessment

Alexis worked on the implementation of the first version of the hovertool. He also worked on collecting fun facts and implementing the search bar with its auto-complete function.

Robin worked on collecting the data for the different types of beer, on the word cloud visualization, on the first implementation of the hovertool and the filters, as well as collecting the fun facts, and on coloring the clusters according to the legend.

Loïc worked on displaying the data on the map and on clustering the markers.

Everyone also worked on this process book and on the general design of the website.

6 Final Product

Here, at the end of the road, we have our final visualization :

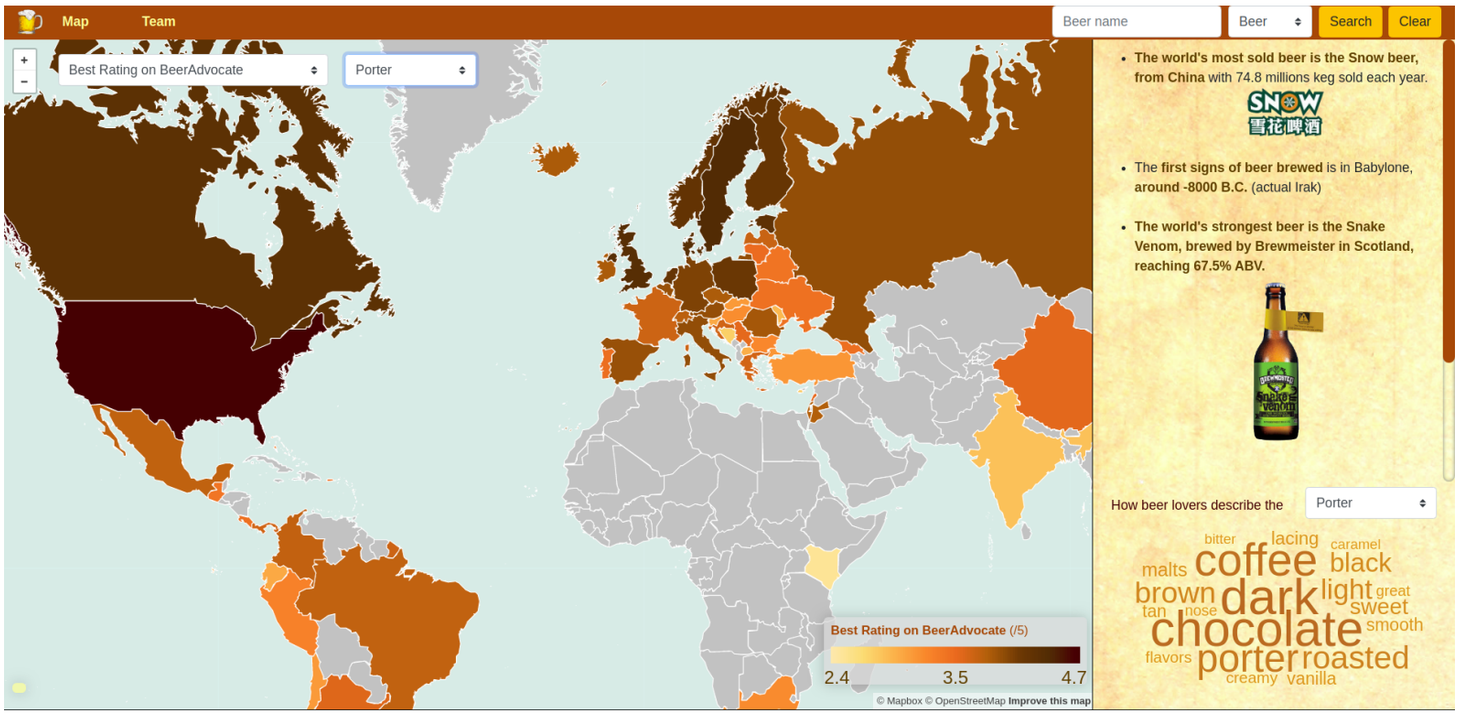


Figure 11: Final world view

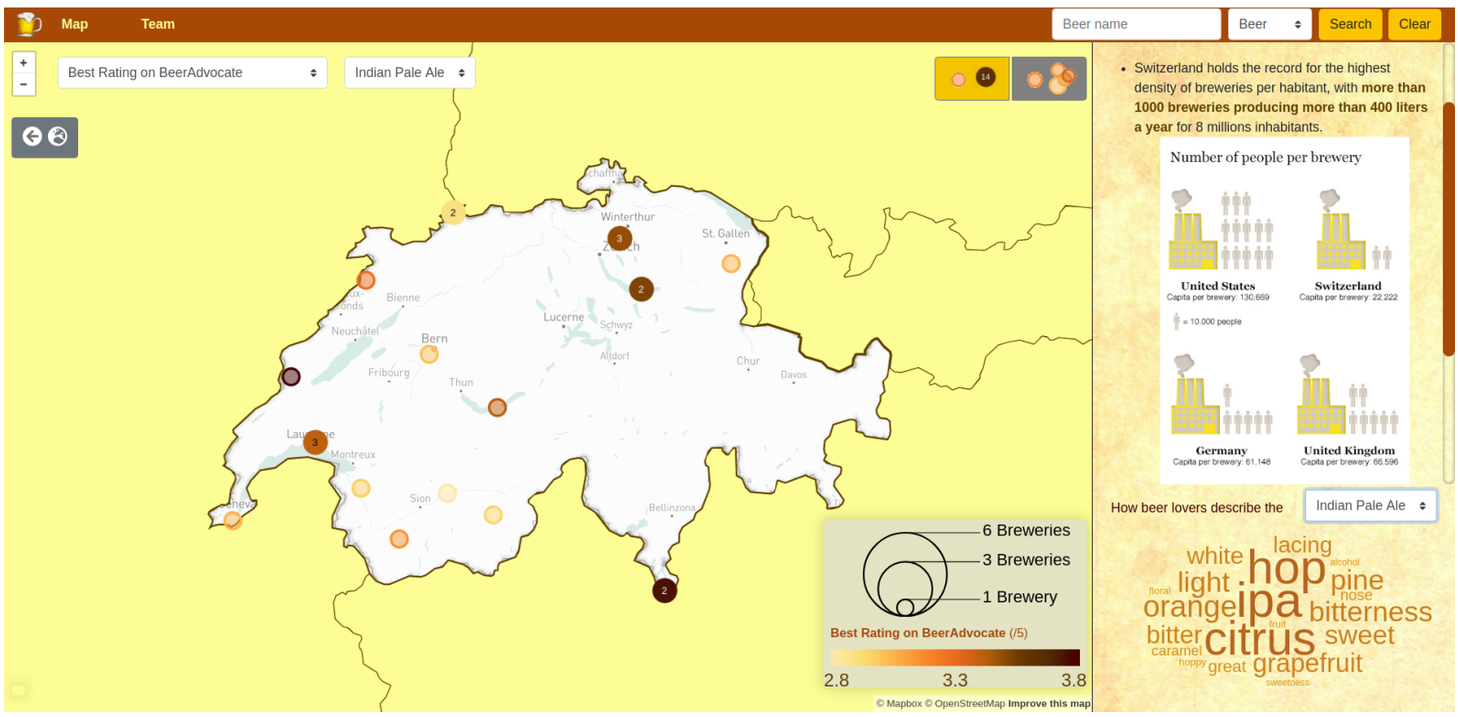


Figure 12: Final country view

7 References

- **Bootstrap :**

  getbootstrap.com

  Bootstrap is the most popular front-end open source toolkit.

- **AnyChart :**

  anychart.com

  AnyChart is a lightweight and robust JavaScript charting solution:

- **Leaflet :**

  leafletjs.com

  Leaflet is an open-source JavaScript library for interactive maps. It has all the mapping features most developers ever need.

- **MapBox :**

  mapbox.com

  An open source mapping platform for custom designed maps.

- **GeoNames :**

  geonames.org

  GeoNames is a geographical database available and accessible through various web services, under a Creative Commons attribution license. The GeoNames database covers all countries and contains over eleven million placenames that are available for download free of charge.

- **OpenStreetMap :**

  www.openstreetmap.org

  OpenStreetMap is a map of the world free to use under an open license.