



# Context of the project

The idea of the project initially started within an EPFL startup (Unki) working on a solution to make group organisation for restaurant outings and holidays easier. As such, in the scope of the startup, a dataset of all the restaurants present on TripAdvisor was extracted and processed. The dataset was used in the scope of the group organisation, but the startup is still at an early stage and there was the feeling of potential for other use cases with the same dataset and that is how we came up with the idea for this project.

## The idea

The startup is user-centric and in this project we wanted to explore a business-centric solution. Restaurant owners are often very busy people handling the supply, service and team management of their restaurants and lack the time and expertise necessary to analyse the trends of their neighbours, region or country and what people are consuming outside of the restaurant. We figured we could help them by providing insightful data visualisations that would empower restaurant owners to take business decisions based on local, regional or country-level knowledge. At this stage we started searching and brainstorming about the kind of insights that would be useful for this purpose.

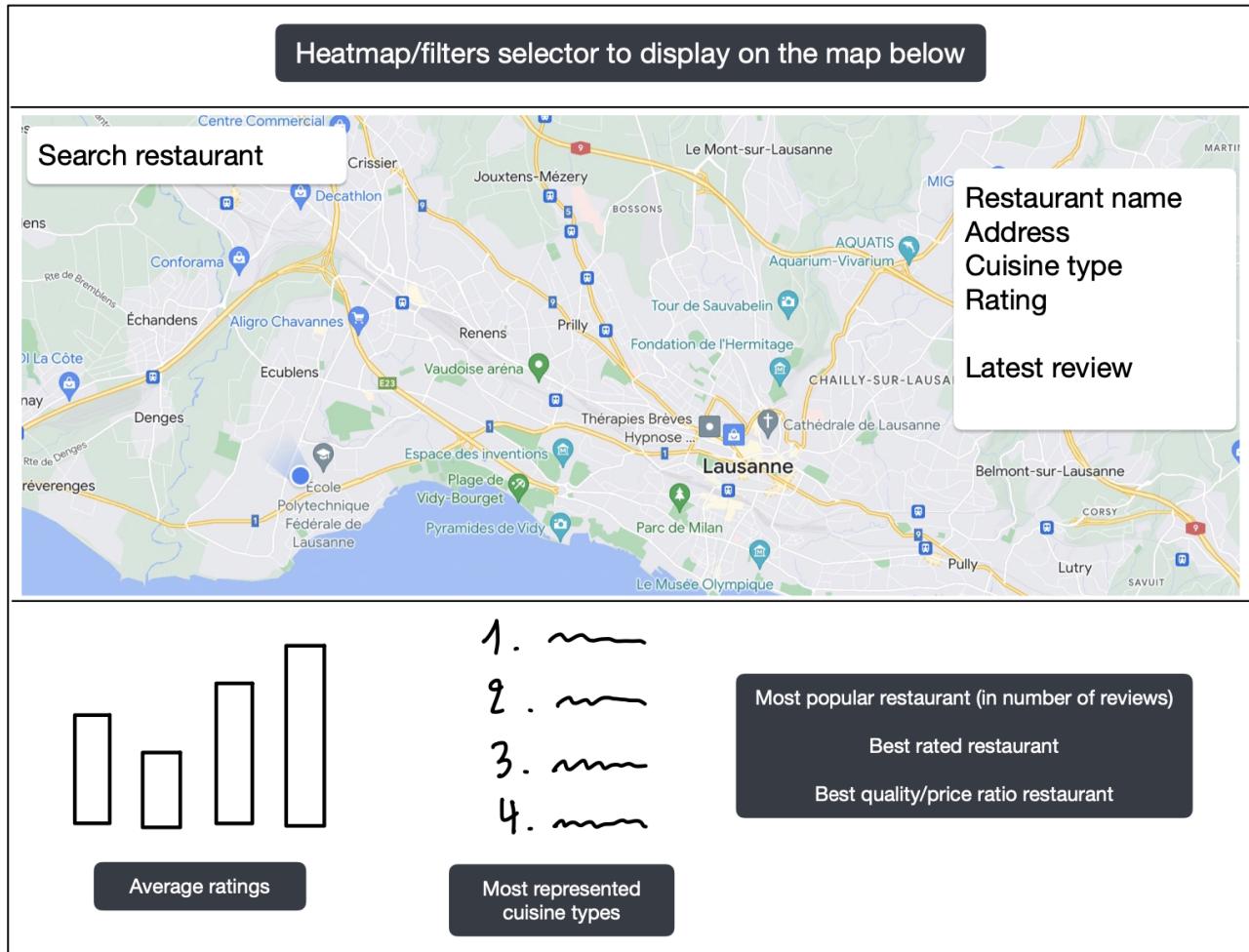
## Structuring the idea

### What we had

In milestone 1, we started off by analysing further the kind of data available and the most present attributes. We came to the conclusion that since mood rating, age and sex were attributes that were not present enough in order to have meaningful visualisations, we would not consider those in our work. The rest of the data was present enough for it to be representative of each region considered for analysis. Knowing the data available, we needed to look for ideas of the kind of visualisations that could make it insightful and we loved the different ideas available in [Tableau Public](#) as well as the approach taken by Will Su with "[nyc foodiverse](#)", which is a data visualisation experiment about inspection data and foursquare ratings for restaurants in New York City.

# What we wanted to achieve

With this in mind, we took our exploration further by trying out different sketch ideas and ended up with two sketches for milestone 2, that would become our webpage visualisation.



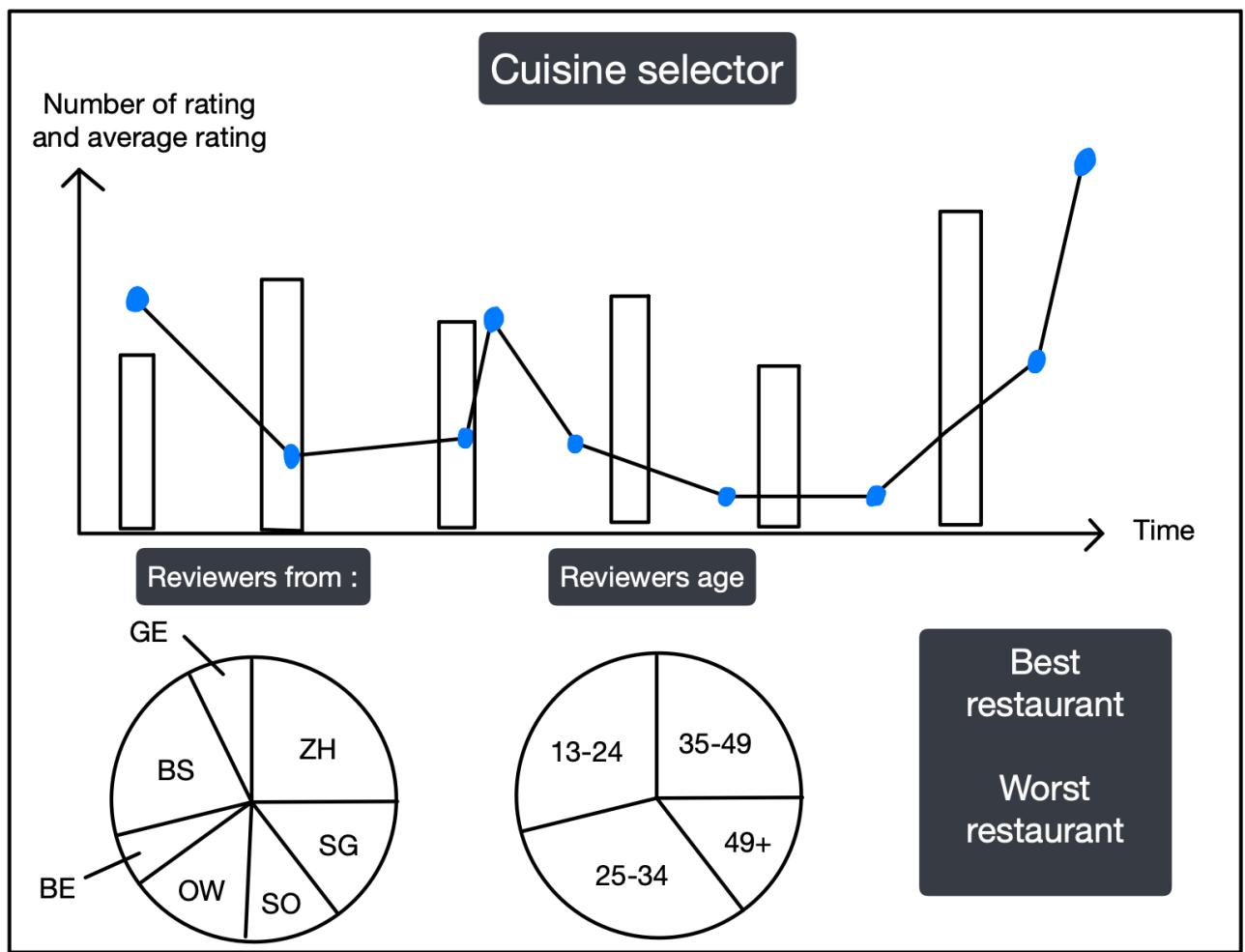
Sketch 1: Main page structure

The goal here was to have a global overview of the key statistics for the displayed region. Looking at the bottom first from left to right, we wanted to understand how well people were rating and considering the restaurants present in the region in order for a business to understand how hard it is to stand out from the competitors. This for us, was a way of assessing the quality of the restaurants in the displayed region. Moving to the second visualisation, we wanted to understand what types of cuisine were popular within the displayed region. This would empower restaurant owners to potentially adapt their menu based on the popular cuisine types in the region they are present in or could be one day.

Finally, we want to display the top of the charts for the displayed region, so that restaurant owners could have a reference of what's working and what's not. We wanted to base the analysis for this on the popularity (defined by the number of reviews), the restaurant with the best rating, which is not always the most popular and finally seeing what restaurants are focusing their strategy on having the best quality / price ratio.

Moving on to the map, middle sector, we want to display the different restaurants considered and empower our users to get more specific information on the restaurants displayed by clicking on them and to search for specific places if they knew what they were looking for. We figured heat maps are a good way to display more information in the same place, without creating a completely different visualisation tool. So, we allow users to choose on top what they want to see displayed on the map and this includes for example price range, so that they can understand the prices being practised for each region and adapt their own prices accordingly, but also to display different ratings such as the overall rating or cuisine rating in order to understand how a restaurant is performing according to region it is active in.

By zooming in or out, we wanted to group restaurants so that one could understand the number of competitors per region and know where it is harder to make it work, or, on the other hand, where there could be interesting opportunities.



Sketch 2 : Deeper analysis

With the second sketch the goal was to add a time dimension to the geographical visualisations we already had in mind. Statistics evolved over time and we wanted to understand the trends behind it, specifically how cuisine preferences were evolving over time. Users could focus on a particular cuisine they like and are interested in and this

page would help them get information about the audience of those restaurants and even see if there exists potential trends for this cuisine.

## What has changed from the initial idea

Concerning the first sketch, not much changed from the initial idea we had. Concerning the bottom statistics, we simplified the third visualisation, since many restaurants had the same overall and quality price ratings. We could have sorted them by number of reviews, but it did not make sense to us since we were already displaying popularity on top. We also gave up on the search functionality, as we did not feel it was bringing anything really relevant to the visualisations and potentially only making the interface more complex. The design of the top bar evolved a bit in order to leave more space for the map and the rest of the display.

Another feature that we decided to not implement was the map filters to decide which restaurants to display on the map based on different parameters (not to be confused with heat maps) : it would overload the interface and we decided to keep it simple so that no user would be lost on our website.

Another problem we got was for the panel describing the origin of the reviewers : our original idea was to plot the distribution by canton but after careful examination, we were not able to compute this data. Therefore, we changed this plot and decided to go with the number of comments distribution for the reviewers. This could help understand if they are communities that get more involved than others and know the commitment of users. Apart from that, the overall page is faithful to the idea we had in mind in the first place.

## Challenges

We have all coded a lot with Python in the past, but had little experience with HTML, CSS and Javascript. So the first challenge was understanding how they worked together and how to debug them. Then came the issue of the libraries, which ones to use and when. The choices are vast and for our use case some were more complicated to implement than others. Mastering the basics of graphs and maps was okay, but things got complicated with heat maps. We struggled to display them and realised “PapaParse.js” was having a problem loading one of the rows of our dataset, but we lost quite some time until we figured this out. Once this was done, things started to run smoothly again.

Another challenge was responsiveness. Not so much on the technical side, but rather on the ideas, it was not always so straightforward to know how to represent information on mobile devices.

Regarding the statistics panel, one major challenge was to compute the required data for the visualisations : since we wanted to plot aggregated data, it required additional steps to get the suitable metrics we were looking for. We also took advantage of that pre-processing

step to format the data in the correct layout in order to facilitate its import later on with "amCharts", the library used for the statistics panel.

## Campus analysis

To give you an example, of how to use our website, we decided to focus on our campus. So, looking at EPFL / UNIL campus and surroundings, there are around 20 places where to have food. People tend to give an overall rating of 4 to restaurants they've been to, which means there is margin for improvement. Types of cuisine available are rather European centric, Italian restaurants topping the charts with 5 places where you can have Italian dishes. The most popular restaurant of the whole campus is Gina with a total of 245 reviews. Cuisines tend to be better rated on EPFL's side of the campus rather than UNIL's. Prices are rather low overall.

## Peer assessment

Alexandre	Eloi	Mike
<ul style="list-style-type: none"><li>• Set up website structure</li><li>• Set up map, pinpoints and restaurant info display on click</li><li>• Dynamic visualisations of the bottom of the page</li><li>• Report writing</li><li>• Website debugging, configuration and upload</li></ul>	<ul style="list-style-type: none"><li>• Dataset analysis and preprocessing</li><li>• Main visualisation ideas</li><li>• Overall statistics page for cuisines</li><li>• Report writing</li><li>• Website debugging, configuration and upload</li></ul>	<ul style="list-style-type: none"><li>• Set up the heat map</li><li>• Improved features of heat map</li><li>• Improved overall aesthetic details of the web page</li></ul>