KIBRAM :
The Spotify visualiser

Process Book : deep dive into Spotify top Charts!

By

**Inès Kahlaoui
Romain Berquet
Antoine Munier**

For

**COM-480 :
Data Visualization
Spring 2023**

Path

From Music to AI

We are a group of music enthusiasts and all did our bachelor in Communication Systems at EPFL with a strong background in mathematics. We started brainstorming about combining the two in a fun and engaging way, while still maintaining the website easy to explore for anyone.

We decided that exploring the statistics of the music app Spotify was the perfect way to do so.

A big part of this project was the back-and-forth communications with Spotify Web API. A lot of data was available to us, and we had to deep dive into the API to get a sense of the most important (and engaging) aspect of that data we wanted to share.

Once we got a clear view about what we wanted to communication and present on our website, we came up with three main axes, which defines the current structure of our website : Spotify top artists, Spotify top songs, and AI & Music.

We started by collecting the minimum necessary data we needed for our visualisations, and keeping metadata for probable future usage. This approach turned out to be the right one to take, since we had to collect more data (compared to the data extracted in Milestone 1) and decided to add add extra visualisations.

We also at the same time were able to define the AI&Music section during the data exploration phase, which needed us to use machine learning techniques to compute a visual representation of music clustering based on audio signal processing features of Spotify's songs.

From there, we were able to define the content of each of our website's section, by sketching and designing on iPad the visualisations we intended to create. This gave us a clear wireframe of the website and insights on how to separate the tasks in hand for the coding of the visualisations (each team member was assigned a section)

We worked on the skeleton of the website, and started the coding process! As mentioned above, we extracted further data in this step.

After coding of our visualisations, we focused on redesigning the website, by defining a clear visual identity, which we didn't take the time to do when creating the skeleton of the website. This allowed to have a consistent and 'mordern-looking' website.

Finally, we focused on redacting the texts for each section, and telling a story. We wanted to create a transition between 'concrete' numbers / facts, given in the Artist and Music sections of our website, and the 'abstract' numbers / facts, by diving into a more machine-learning oriented way of thinking.

KIBRAM Visualiser
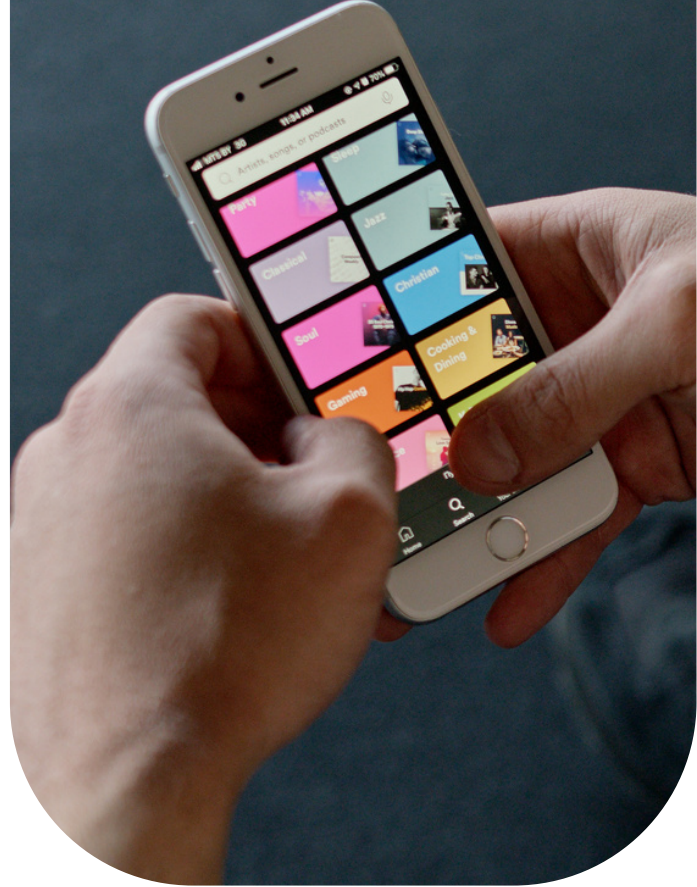
Challenges

What worked and what didn't work for us

The very first challenge was, for some of us, to build the website itself having no prior knowledge in html, css and javascript. The first few weeks allowed us to navigate through the libraries and frameworks available for a data visualisation task like this one.

One challenge though was dealing with non-static data. Calls to Spotify Web API data is updated daily, and we had to rerun our algorithms to fetch and save data while discarding the old one, and ensuring a correct formatting so that the updates would not impact the visualisations.
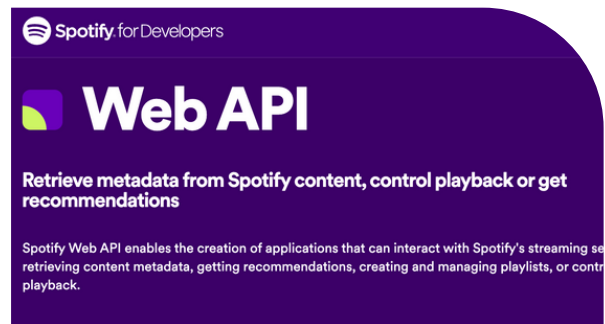
As we wanted this website to be easy to use and for everyday Spotify users, we had to define a visual identity that resembled the Spotify one. We reviewed the whole design of the website since Milestone 2 to achieve a more modern-looking website. However, the main issue arose when trying to match Spotify visual identity : fonts and other elements in the identity were not available on the Internet, and we spent some time searching for similar looking fonts and elements to create this feeling of 'déjà-vu' on the site (for daily Spotify users).

Finally, the last problem for us, and we cited it before, was really making sense of all the data in hands. Upon designing and sketching the Artist exploration section, we weren't sure of how to link in the data artists, albums and songs, and wondered if we had every element to make the visualisation. We would have found tedious to extract for each artist every album he/she/they made on the platform and then extract for each album every song and their corresponding features. We instead extract a large number of top artists, albums and songs and computed our own features linking these together. For example, an album doesn't have a popularity score. We calculated this score of a track as the average popularity score of all tracks in the album. Similar data problem occurred we weren't expecting came up during the coding process. We preferred focusing on getting the data correctly ready for use than the actual visual design of the website.
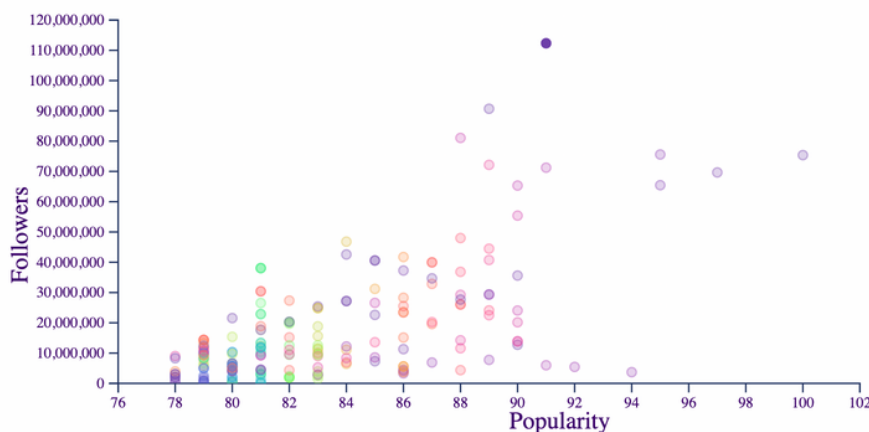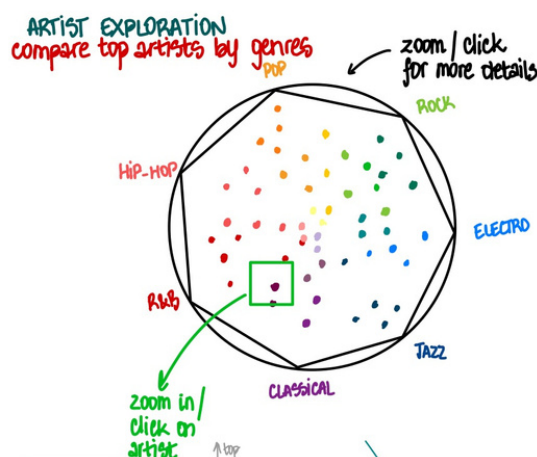
Sketches

Artists exploration

This section is divided into 2 connected visualisations.

We initially mentioned in our mock-up that we intended to create a wheel of colours based on machine learning clustering on music genres, but with over 85 genres in the dataset, clustering was near impossible. We instead showcased the relationship between an Artist' main genre, its popularity and its follower count on Spotify. Hovering over a point allows you to see information about the Artist on Spotify.


It turns out this shows that pop artists dominate the charts, which in reality makes Spotify a pop-oriented platform rather than a fully musical one. Although our initial wheel of colour idea was quickly thrown out, we decided to keep the connection between the genre-graph and the Artist graph below. An integration of Spotify embed snippets allows you, upon clicking on a point on the graph above, to listen to the Artist's top tracks.

We initially had some doubts about this visualisation, complex in formatting and data-extraction. We had to combine a multitude of datasets extracted from the Spotify Web API, and create links between them, kind of like a mini database. We ended up fully implementing it, and you can hover on an artist's top albums and tracks to take a look at their popularity on the platform. You can also get insights on a specific song or track of an artist, and visualise the relationship between this track and others in the same album. However, we noticed some problems with the implementation : some artists, even big in popularity / follower count, did not have any album or songs in the top charts. This impacted the resulting visualisation.
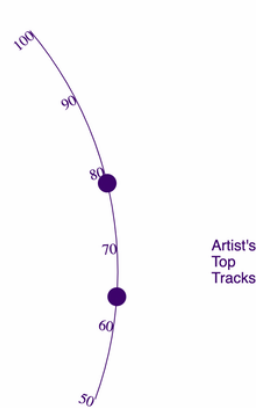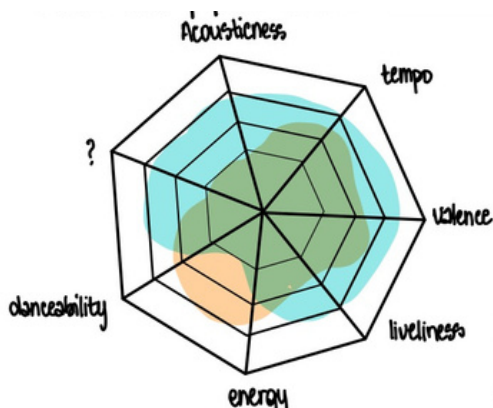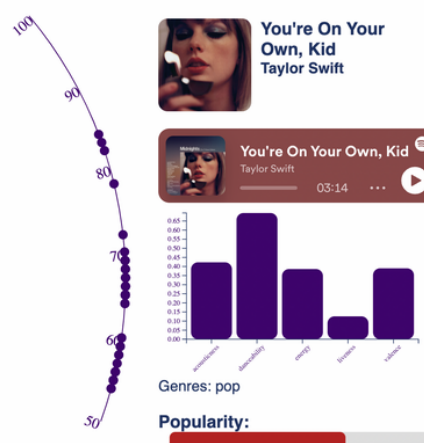
Sketches

Artists exploration (cont.) and Audio Analysis exploration
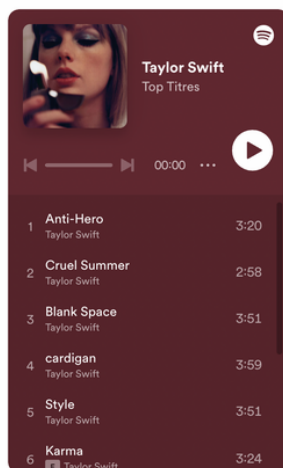
Spotify Web API allows you to extract 'musical' features from basically any song on the platform.
Sometimes, we like the most random song ever. Did you ever wonder why? Well us too. We wondered wether musical (audio signal processing) features could explain our taste in music. You can select any two songs of your liking from the top 1000 songs on Spotify to compare its musical features on a spider chart. See a pattern?
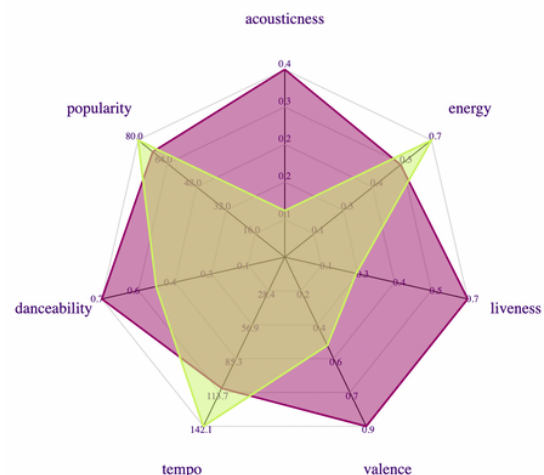
Sketches

AI & Music

This section was by far the most resemblant to the initial mock-up we had designed. It showcases the correlation between features and popularity of top tracks on the platform. Since the number of features available to us is quite large, we simplified the visualisation by only allowing a maximum of 4 features to be displayed at once.

The curves were drawn by regression: we had a discrete set of vectors for popularity and audio features, and approximated the curve to get a clear picture of the correlation between popularity of a track and its audio features.
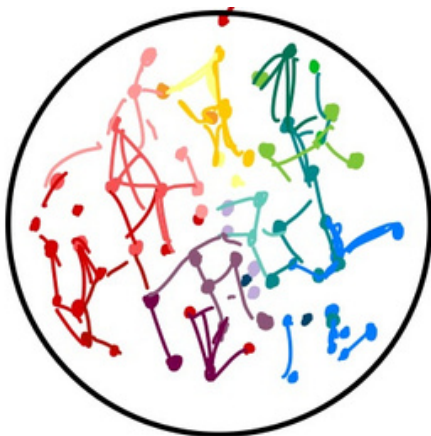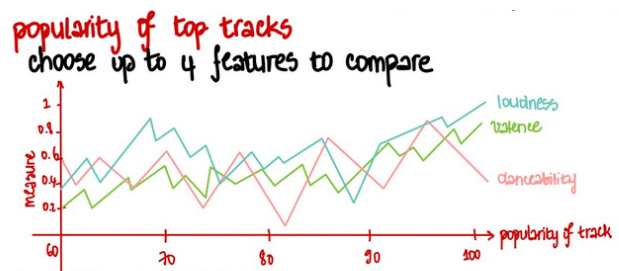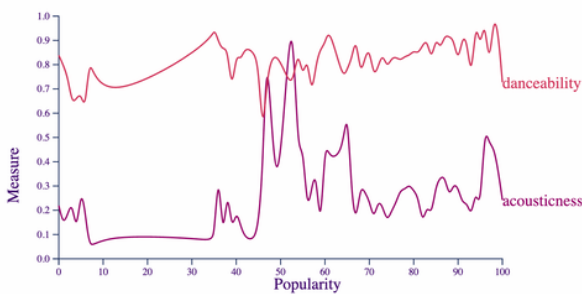
We wanted to showcase a new way of thinking about music. We initially wanted to implement a 2D scatter plot of the 500 top tracks on Spotify to do so, but found that the randomness of the clusters (see Milestone 1) impacted the visualisation. We opted for an alternative view by cluster, and implemented an option for listening to the track in question upon clicking on a point.
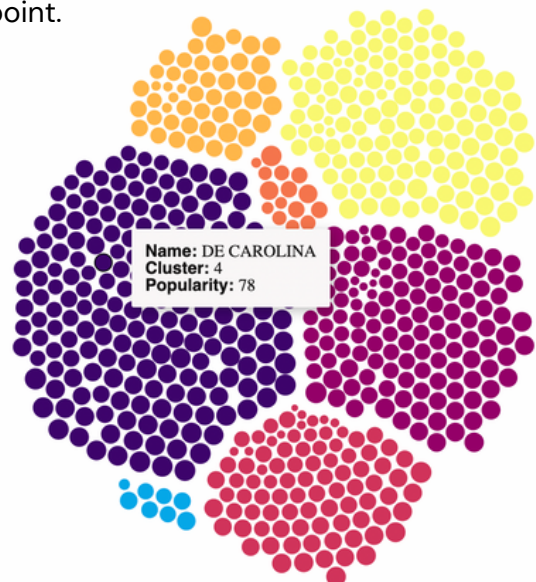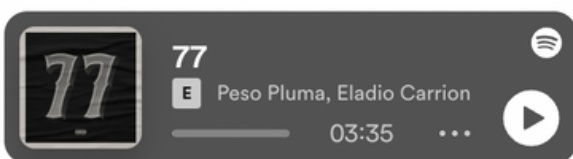
Peer Assessment

Throughout the entirety of the project, we made it a point to consider and make all significant decisions as a group. Every member came up with fresh ideas, which was incredibly beneficial in maintaining our forward momentum and was the key to our continuous progression.

It is difficult to clearly estimate the task division on the website. Each of us was responsible for designing a section of the website, like cited below, but we all did some tweaking on other team members' sections, which really allowed us to reach the final design and charts you see on the website. Here's however a detailed distribution:

Antoine

I worked on the overall re-design of the site (we decided to take another route from Milestone2). I also contributed to writing the process book.

I was also responsible for the AI&Music section.

I finally worked on cleaning up the code, which lacked modularisation at the start of the project.

Inès

I was a bit familiar with Spotify's Web API so I did the exploratory data analysis, the continuous data extraction and ML algorithms for all the visualisations of our website.

I worked on designing sketches of Milestone 2 and contributed to the process book.

Finally, I was responsible for the Audio analysis exploration section.

Romain

Having some experience in HTML, I first worked on creating the skeleton of the website.

I was also responsible for the Artists exploration section.

I was also in charge of filming the screencast video, with the contribution of the rest of the team.

I also contributed to writing the process book.

KIBRAM

Where to?

Access to the site : com-480-data-visualization.github.io/project-2023-kibram/website/index.html
Access to the GitHub : github.com/com-480-data-visualization/project-2023-kibram
Spotify Web API : developer.spotify.com/documentation/web-api