SpotifyTop200

PROCESS BOOK

SPOTIFY
Top 200 from 2020 to 2021

With the launch of streaming platforms such as Spotify, a new way of discovering and listening to music has emerged. A huge change for both listeners and performers. It is easy to find the songs from a group of friends playing in the small concert halls of our town as it is for the ones of an award-winning artist. Every performer has a chance to reach and touch people from around the world.

Since the accessibility to music is not a big issue anymore, one can therefore wonder what makes or breaks the popularity of an artist nowadays.

PATH

We first shared ideas about what could be of interest to show about our dataset since there are so many features and interesting elements about Spotify songs. We wanted to have a storyline and not go in every direction possible.
We agreed on the fact that it would be very interesting to study what makes a song popular. Indeed, the dataset is a compilation of all songs that were highly popular between 2020 and 2021, so we would definitely find common points between them and we were curious about what interesting elements we would find and present in our final visualizations. We also found with this axis that it could be fun to have an interactive game to test the user's understanding of the subject and let them guess which songs are more popular than one another.

We then started thinking about the actual visualizations as well as the visual identity of our website. We wanted to keep a look visually similar to Spotify's with dark background, the logo and soft gradient backgrounds. We sketched the graphs and the overall website.

Then came the coding part. We started with what we had learned in class about Javascript, CSS and HTML and did a very basic skeleton of the website, then with CSS we created the backgrounds and we used the Fullpage.js library to get the effect of one page at a time. Once every page was assigned to a future graph and the main structure and visual identity was present, we separated each graph from the rest, worked on it until we got the result we wanted and finally integrated it to the website when it was ready.

We decided to follow the storytelling model of the data scientist, since we wanted to construct visualizations which could be interacted with by the user and where they could explore the data and more information by interacting with the visualization. We found this was the best model for us also because we followed a gradually complex approach, where the user will first see basic information about the data and the more the user advances the more complex the information gets. Our visualizations are quite analytic (histograms, dependency wheel, …).

Finally, we had another brainstorming session to share our opinions on the current state of the website and discuss some specific details. Once it was modified, we sent our website to a few people who could give us honest feedback about its readability and overall quality. With their feedback we were able to make some final modifications to get the best result we could get.

SpotifyTop200

CHALLENGES

We first had to find a complete dataset that gave us ideas for a good visualization. Once we chose the Spotify dataset, some cleaning had to be done and some elements highlighted. Choosing the graphs that would best reflect our objective was not easy as well, as we had to choose which features to put forward and how to best present the dataset without making it too complex or simple. We also needed to find a storyline for our different visualizations, a sequence that would make sense throughout the whole website.

The next challenges came with the implementation of the website. We first took some time to link the dataset to our different visualizations, as it was not so easy to use local files with D3. We found a working solution by locally opening the website when editing under a local server which allowed us to remove the access restrictions.

The graphs also brought some difficulties on how to deal with the libraries needed, even if we had briefly used them in the exercises for some of them. There was quite some trials and errors regarding the layout of the graphs and text on the pages, playing with the CSS attributes. We especially struggled with the histograms of the music features, as we had trouble with showing the axis and with the right layout for the page.

Lastly, probably the biggest challenge was the interactive game. In the beginning we struggled to find a way to randomly pick two songs directly in our datasets. We solved this issue by creating a json file from our csv file. Then we realized that keeping the score was not an easy task. Since we had no experience using HTML and javascript it tooks some time and many researches to learn how to do it without errors. From there, implementing the highscore and the rest of the game became way easier.

 SpotifyTop200

FROM THE SKETCHES TO THE WEBSITE

Overall we followed our main ideas closely and kept our storyline and visualizations similar to the sketches.

INTRODUCTION

We started the website with the Spotify logo as planned and a quick introduction on the dataset and subject of the visualization. We wanted to give a quick explanation of our starting point and some precisions about the data we used and what ideas emerged from it. The only thing that we did not implement was the play button to get into the website. We kept it for another purpose: the start of the guessing game. We used instead simple arrows to access the rest of the website.
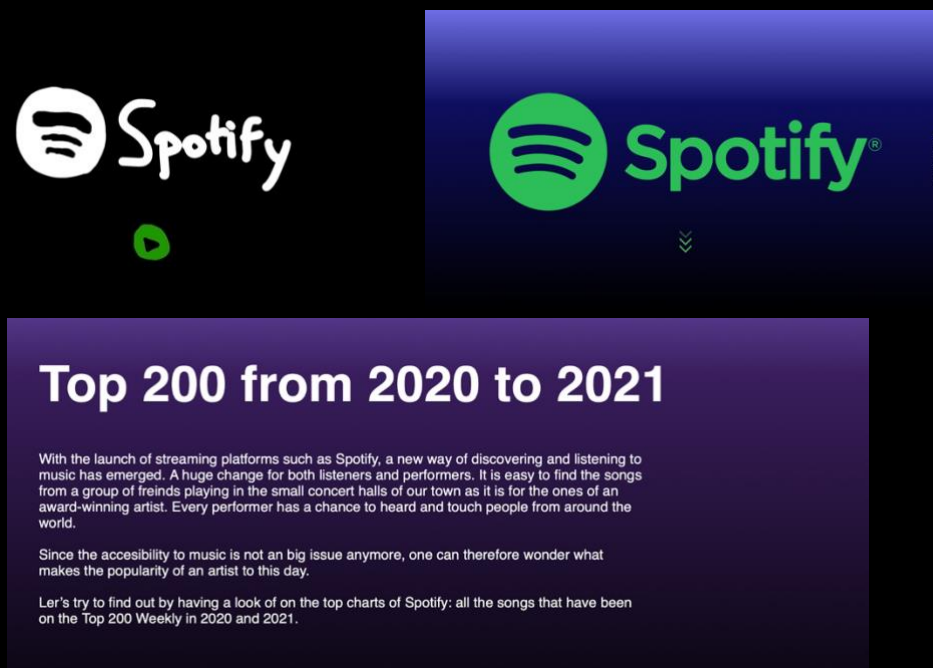

MENU SELECTION

The menu selection allows the reader to get directly to the different visualizations without scrolling. The appearance differs a little from the original one that was similar to Spotify but the functionalities are as we wanted. This displays an overview of the website and its accessibilities.
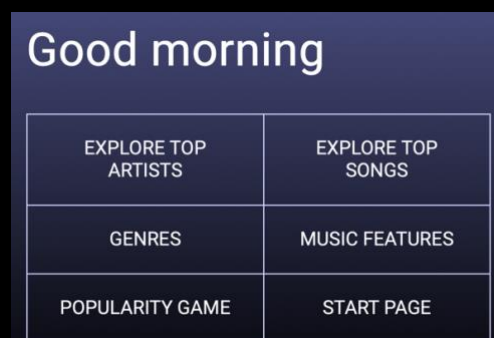
TOP ARTISTS

The top artists' visualization gives a good representation of the popularity trend. The 10 artists with the highest ranking were directly selected from the data. The inspiration was taken from the Spotify display with a few alterations caused by the different tools we could use. By hovering on the play button on the artist profile it is possible to read further information on them: their rank, their top song and the genres they belong to. We added a direct link to the Spotify playlist of the top artists of 2020.
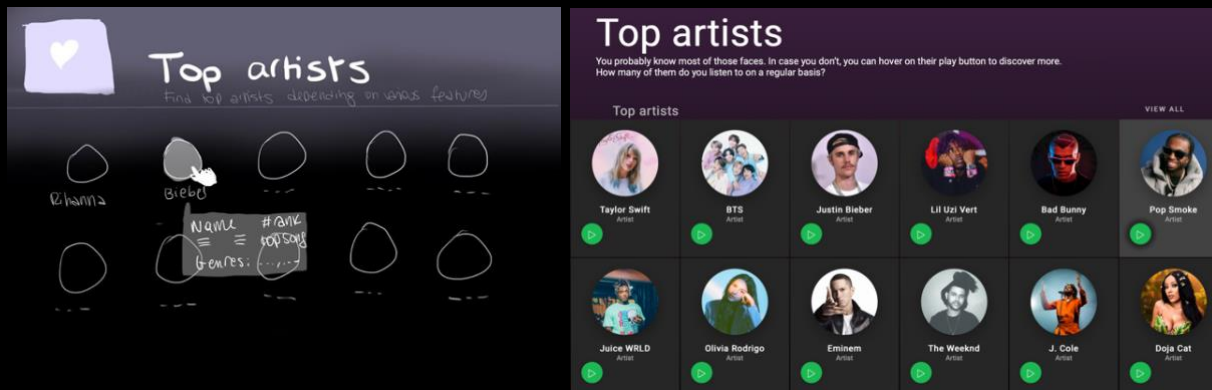
TOP SONGS

The top songs visualization was one of the easiest implementations with a display coming directly from Spotify. Each song can be played and is ordered by rank. We simply created a playlist on one of our Spotify accounts. With the information retrieved from the dataset, we were able to know the 10 songs with the more streams and could add them to the playlist.

For the top artists and top songs visualizations, we were really focused on giving a similar experience to the reader as the one we could find directly on Spotify. We wanted to incorporate the direct access to music as well as the details on the artists including their picture.
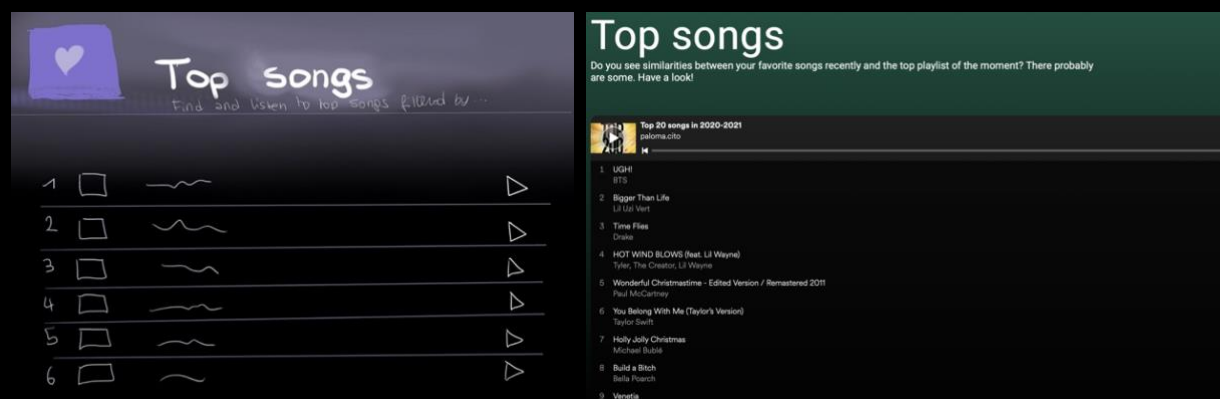
GENRES DEPENDENCIES

The visualization of the genres turned out exactly as we wanted. The dependency wheel is an interesting and interactive type of graph. This displays many information in addition to the links between categories, such as their weights.

We wanted to depict where dependencies exist between genres. The links were created if two genres appeared on a same song. We can assume that most popular songs have similarities with their genre of music. It is easier for a song or an artist to be famous if the genre of music they make is also popular at his time. As we can see from the graph, pop music, as its name describes it, has been since its creation in the 50s the main genre of popular music. We can also see that actually, latin music and its derivatives have gained in popularity and are now the second main genre of popular music. We can see overall a clear connection between the genres that appear the most in the Top 200 chart.
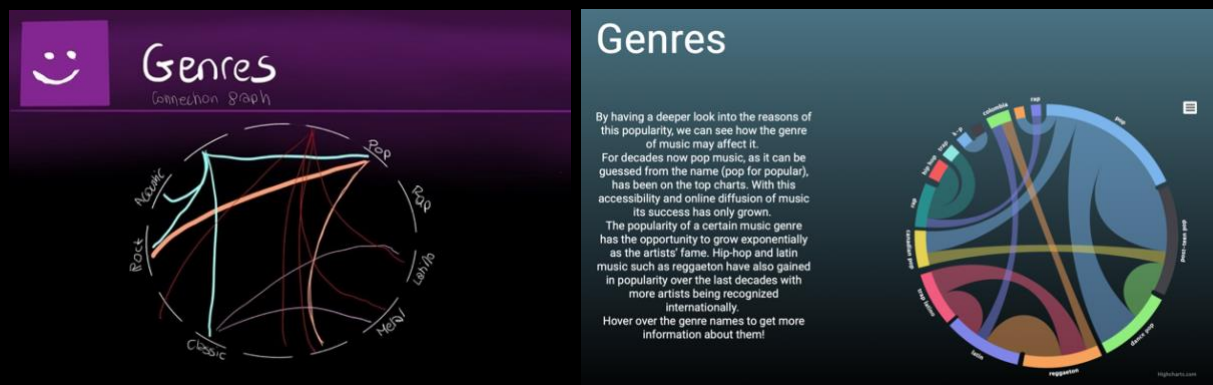

MUSIC FEATURES

As for the genres' dependencies, the music features visualization gives a deeper insight on the Spotify data. This is information that we usually cannot get directly from the Spotify platform. We wanted to give a more advanced analysis on the popularity of the songs by displaying an histogram for each music feature.

We see that the features that give a positive feeling such as danceability, energy, and loudness usually have high scores. This depicts music that we would usually dance on. (They also don't have a lot of lyrics as the speechiness is apparently small.) Whereas more melancholic traits such as acousticness seem not so important.

Unfortunately we did not have enough time to implement the song selection as explained in milestone 2. We only gave a general overview on the top 200 songs, which we find sufficient to have a good understanding on tendencies regarding the music features. However, we could implement the different definition as an interactive element.
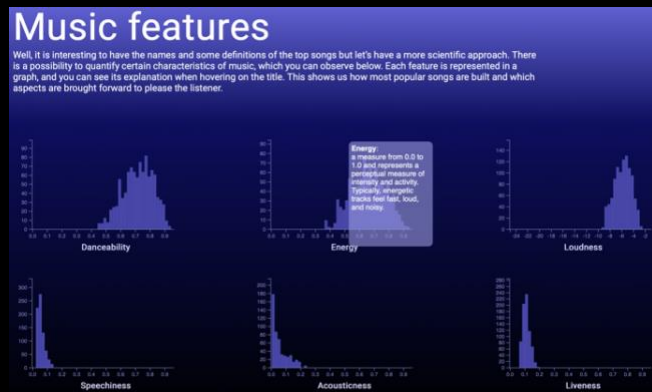
GUESSING GAME

For the final visualization, we managed to implement the guessing game, to finish on a fun note and to test the understanding of the reader. The idea was to simply guess between two songs, which one has the most streams on spotify. If you guess correctly, you earn one point on every round and the game continues until you fail. We also added an highscore that keeps your longest run of good guesses. Compared to the original plan we just didn't put an image for the songs, because we didn't have images linked to the hits in the dataset. Moreover, sometimes when the songs are featuring between multiple artists, the name of the second artist appears in the "song name" part of the button, this comes from the fact that we just print from the dataset and it is not homogenous.
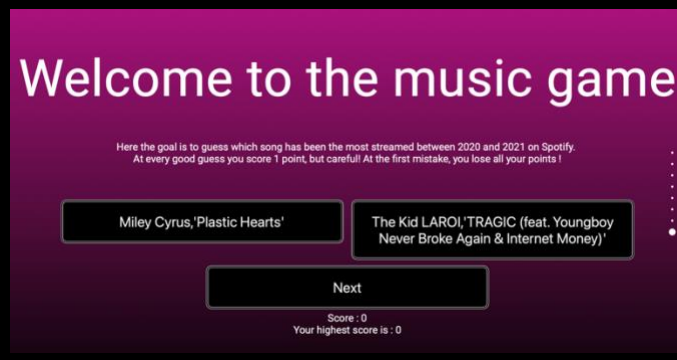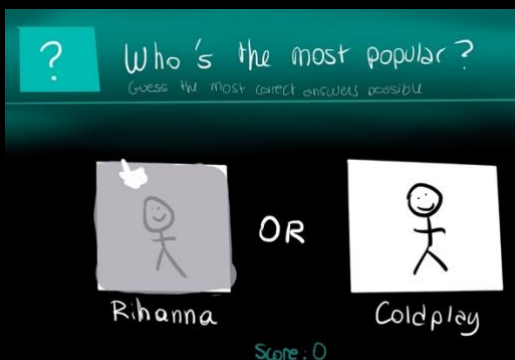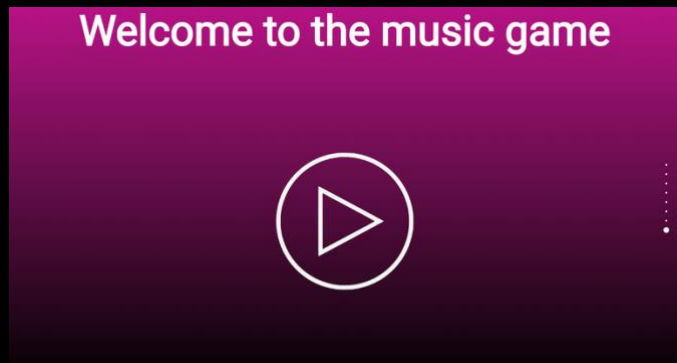
PEER ASSESSMENT

MILESTONE 1

After finding the dataset together, we separated the first milestone between its three categories. After, Marc took care of the question on the problematic, Paloma produced the EDA on Jupyter Notebook and Marie did some research on the related work of the dataset chosen.

MILESTONE 2

For the second milestone, Paloma created the final sketches of the different visualizations we wanted to achieve and listed the tools and lectures we would use. Marie redacted the milestone and the different ideas we agreed on. Marc put in place the initial website with the basic skeleton of the visualizations.

MILESTONE 3

For the final and longest milestone, we first tried to split the work but ended up doing as we could according to our time schedules to progress as fast as we could.

Paloma took care of the linking between the dataset and the website in order to use it in our graphs. She implemented the visualization of the top artists and figured out how to produce the charts for the music features, as well as the basics of the genres graph. She also took care of the aesthetic of the website. Finally, she assembled the storyline video and put it online.

Marc created the guessing game which gave the most difficulties because of the direct interactions with the website reader. He also took care of the storyline for the explicative video of the website

Marie embedded the Spotify playlist into the website, continued and refined the genres graph. She redacted the process book. She also took care of the explicative texts and the narrative of the website.

GENERAL ASSESSMENT

Overall we maintained a very good organization and were able to progress fast. We managed to meet once a week to report on what we did and which steps we still had to complete. Even if we did some different type of work, everybody could do a little bit of everything and we were able to hand out a complete project that we liked.