Data Visualization - Milestone 2

Inès Kahlaoui (310587), Romain Berquet (316122) & Antoine Munier (314500)

5th May 2023

Project's goal

In this project, we want to explore the different relationships between a music's success and its audio representation by creating data visualizations that showcase insights into the most popular music on Spotify in 2023.

You will find in annex a wireframe of the general aspect of our website (iPad mock-up). Each section is detailed below.

Artist Exploration

We wanted to present the top artist's on Spotify in a more in-depth view. For this reason, this part of the project represents the majority of our website, as it is the most interactive visualisation we intend to create.

To this date, the Spotify app doesn't display (at all) the top artists on the platform. We wanted the users to be able to explore today's trends by displaying the artists in a more melodic way (lecture 'Sound viz' might help us to achieve that). This is why we wanted to do clustering on top artists' genres and displaying them on a 2D plane (cf. wheel of genres). User interaction is the main focus for this task, so the user is able to click on a 'point' on the wheel, which represents an artist, and get more insights about his albums and tracks. Links between artists will be added, with the strength / thickness and number of links signaling a deeper collaboration between two artists (a large number of tracks in common).

To achieve a good representation on Spotify top artists' albums and tracks, we plan on focusing on slicing and elision : on the left a description of the artist (and more data about him/her, e.g. popularity score, number of tracks in top 500, image), then two circular arcs, with the first one being for an artist's albums, and the second for an album's tracks.

The last layer of this representation is an added track information section (upon clicking on the track's point) with various details given by Spotify's API (featuring artists, image, audio features, genres, popularity score). We still have yet to determine a concise way to represent together an audio's features and to evaluate whether or not this task is relevant for the global scope of our project.

For this section, we took inspiration from the Reuters 'Connected China' website presented during class. As this task the major axis of our project (thus takes time to achieve), we preferred not to include it in our initial website functioning prototype and undertake it at a later time. We notably will make great use of the 'Interactions' lectures to achieve a good representation, and p5.js mentioned in lecture 'Beyond visualization' might be of use.

Audio Analysis

'Compare track'

The audio analysis section of our website gives a more 'abstract' representation of the top 500 tracks on Spotify. It allows the user to select 2 tracks (we may add an option to randomly select 2 tracks) to compare their audio features. According to the distribution we found during our exploratory data analysis decided to make use of the following features available for a given track : acousticness, danceability, liveliness, energy, tempo and valence (cf. Spotify API documentation for a detailed explanation). As some of these terms can be non-intuitive to understand for non-music experts, and to further our goal to reach a larger audience than music enthousiasts, we intend to add textual explanation for each feature. We also might add a visual explanation if the textual one is not sufficient (and if time allows, as this can be a pretty time-consuming task). Part of this visualisation will consist in a violin graph of each attribute, using simple D3.js functions. This would offer an alternative representation to the EDA distribution we plotted in the first milestone.

As we have a large number of attributes, we decided to represent them using a 6-leaf radar chart. We are aware that radar charts may not be the best way to represent numerical data in a clear way, but we feel like this representation serves more to visualize abstract musical concepts, which is the direction we want this visualization to follow as cited above.

'Audio analysis of popularity'

The subsection offers an insight on what might impact the popularity of tracks. We keep this visualisation as optional for the moment, depending on our advancements on the other sections of our website. Nonetheless, we intend to use a simple lineplot using D3.js. We intend to put a maximum of four features displayable at once, not to overcharge the user's view ('Do and dont in viz' lecture gave us insights to set these conditions).

Surprise me visualisation

In accordance to the Principal Analysis Components k-means clustering we performed during our initial exploratory data analysis, we wanted to take 'abstract' to the next level and bring together AI & Music. We thus considered amongst all 2-principal components graph representation to selection the one with the largest distances between point. The reason for this is that (as seen in the EDA notebook), a large number of graphs have points too close together, which impacts in a great way its visualisation. So to add clarity, we plan on using D3.js to create a 2D Network graph with Skeleton-based Edge Bundling as shown in the course. The number of links and their intensity, and the links in themselves will be chosen according to a threshold on the correlation of attributes of audio features we will set depending on the density of the graph (we don't want to lose in clarity if the graph is too dense, as learned in the 'Do and dont in viz' lecture). We took inspiration of Kirell Benzi's work on Jazz Luminaries for this representation (as mentioned on our first report).

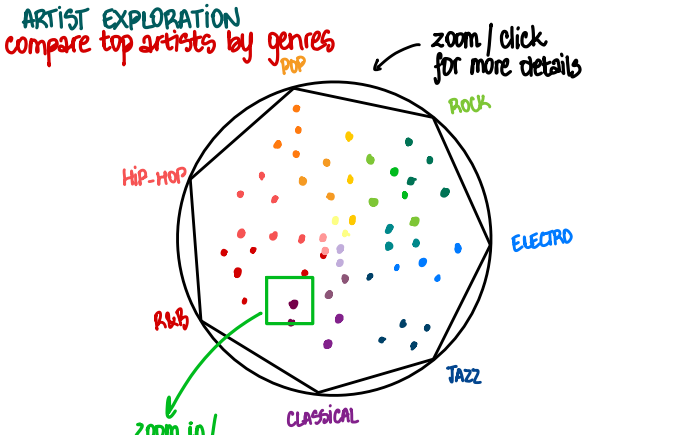
INTRODUCTION

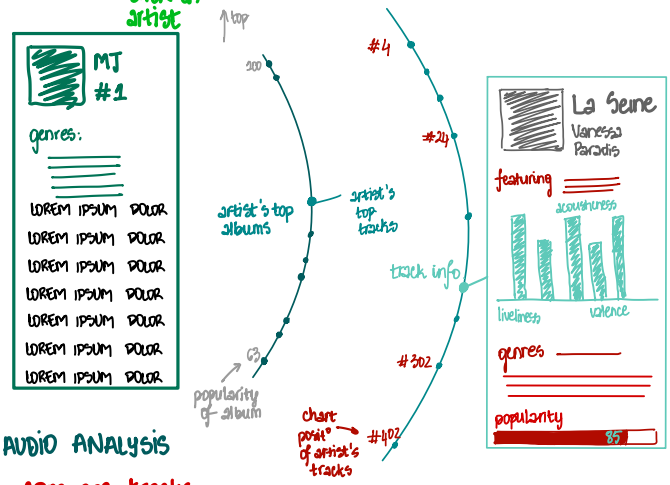
LOREM IPSUM DOLOR SIT AMET LOREM IPSUM DOLOR SIT AMET

LOREM IPSUM DOLOR SIT AMET LOREM IPSUM DOLOR SIT AMET

ARTIST EXPLORATION

compare top artists by genres

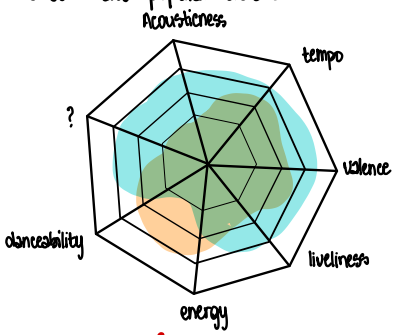




AUDIO ANALYSIS

compare tracks

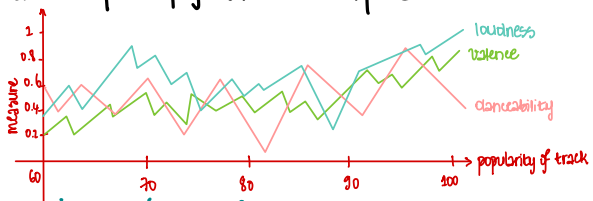
Choose two popular tracks TRACK 1 ▼ TRACK 2 ▼



LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET

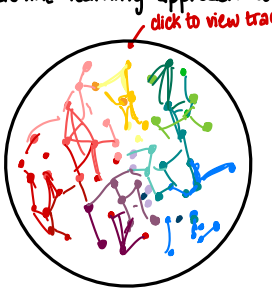
popularity of top tracks

choose up to 4 features to compare



SURPRISE ME VISUALIZATION

Machine learning approach to music visualization (clustering)



LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET
LOREM IPSUM DOLOR SIT AMET