Data Visualization: Milestone 2

Florian Singer, Alessandro Tempia Calvino, Alexandre Hutter

7 May 2021

1 Introduction

Our goal is to depict different trends in music history, from 1922 to 2021 based on the contemporary taste. In this visualization, we aim to show: 1) the trends of music production showing the most popular tracks per year and investigating how their audio features affect the popularity; 2) the evolution of genres in history, their main audio features, and popularity. The website will be structure on a single page containing at least three interactive visualizations that all compare the evolution of a particular aspect over the years.

The main visualizations are discussed in Section 2. Additionally, we introduce further ideas in Section 3.

2 Minimal Viable Product

In this section, we introduce the main structure of our visualization defining what we want to show and how.

2.1 Viz 1

2.2 The most popular tracks per year

This visualization shows the most popular track per year, from 1922 to 2021, based on the current popularity. The popularity score is between 0 and 100. The visualization is on an interactive chart. Zooming and panning is possible using (D3 brush+zoom) and tracks and artists data is visualized using D3 tooltips. The visualization takes the initial inspiration from exercise 5.
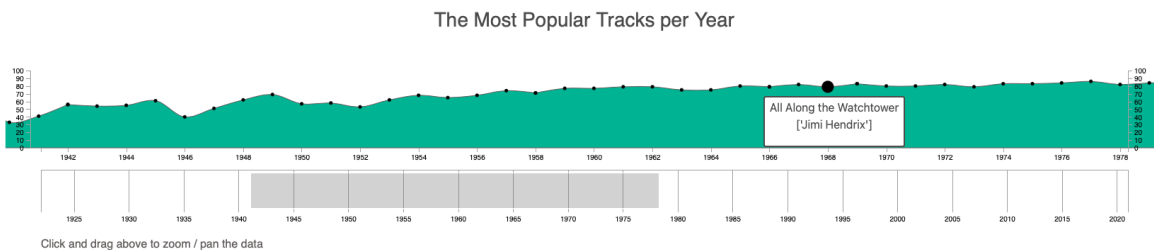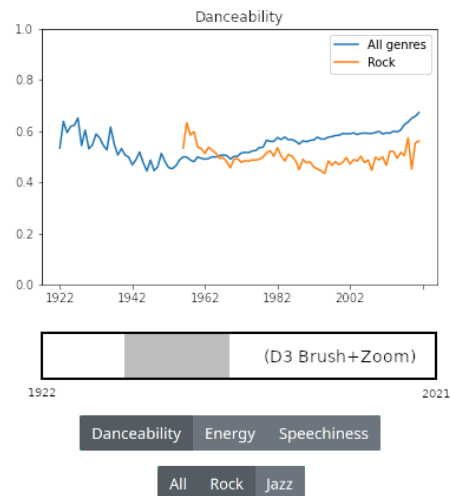

2.3 Audio features by genre

The Spotify API provides scores between 0 and 1 for multiple audio features such as danceability, energy, or speechiness. These features are available in our dataset for each of the tracks. The idea of the visualization is to plot the evolution of these features for each genre, and compare the different genres.

The data has been processed to compute the mean value of each audio feature, for each year and each genre.

The main part of the visualization consists of a plot, with the x axis describing the years and the y axis describing the feature (between 0 and 1). The user can select the desired audio feature using radio buttons. Then, the user selects the genres to be included in the graph. This is implemented using check-boxes, so any number of genres can be shown simultaneously. Finally, the user can zoom on specific periods with a D3 brush (as seen in exercise 5).

The image below illustrates this idea, but more audio features and genres will be available in the final product.

3   Extra Ideas

To further improve the visualization, depending on the available time, we have the following ideas:

- Cluster songs base on audio features in a 2- or 3-dimension fashion with different colors based on genres to generate a principal component analysis (PCA)

- Add a play mode in the "the most popular tracks per year" visualization that allows to reproduce the most popular track per year using the Spotify API. Lecture on sound visualization will be useful to this purpose

- Include a lyrics analysis. Some keywords (e.g. love, baby, never, etc.) are selected and visualized over time. The goal would be to select some words and visualize the trend of usage over time to answer a question: do old songs have deeper, more romantic, more socially involved lyrics?

4   Website

The initial website is available at the following link.