SoundScapes

Bernardo Vargas Cristiano Larréa Paloma Borges

***Abstract*—This project showcases an interactive web application called SoundScapes: A deep dive into Spotify's tracks, which combines data visualization and scrollytelling, allowing users to explore the characteristics of the most popular songs on Spotify from 2000 to 2019, providing insights into the audio features utilized by the platform. It also enables users to better understand the musical landscape during this period.**

***Index Terms*—music, Spotify, interactive data visualization, scrollytelling**

I. Introduction

In the last years, Spotify has revolutionized how we consume and discover music. With its huge library of songs and powerful recommendation algorithms, Spotify has become a reference platform for music lovers worldwide. As the popularity of the platform has grown, so has the curiosity about the trends and defining attributes of the most popular songs on Spotify, so we decided to investigate and visualize this question.

The connection between waves (sound waves) and music inspired the idea of making the project resemble a dive into the ocean, starting with a shallow, superficial section, and adding depth as you dive into the sea, i.e. starting with a more generic visualization and adding complexity and interactivity as you scroll down the page. With this, the concept of deep dive into Spotify's tracks was born.

With this in mind, we created the Soundscapes project aiming to provide an immersive and informative experience for users to explore and learn about the fascinating world of music on Spotify, discovering the patterns and features shaping the musical landscape of the past two decades. Through scrollytelling and interactive data visualization, users can analyze diverse aspects of Spotify's music and understand a song before they even hear it, learning more about what each music feature represents.

II. Related Work

Several prior research works have explored different aspects of music analysis, visualization, and discovery, providing valuable insights that contributed to the Soundscapes project.

"Music Discovery in Spotify with RAMA" [1] presents an application for visualizing and interacting with networks of music artists. This work emphasizes the importance of personalized recommendations in enhancing the user's music exploration experience.

Another relevant work is "Music Circles" [2], which focuses on visualizing music based on audio features. It offers an interactive interface where users can explore and compare different tracks and their characteristics. "Music Visualization and Discovery using Audio Features" [3] presents an interactive way to visualize, find and compare music rankings using different techniques, including the display of music attributes. These projects provided inspiration for the interactive visualization components implemented in Soundscapes and a foundation for the understanding of musical characteristics, their relevance, and how they could be visualized.

III. Methods

We utilized a dataset that contains audio statistics of the top 2000 tracks on Spotify from 2000-2019 [4]. To begin, we did a brief exploratory data analysis using Python to understand the metrics used to describe the tracks. From this, we decided to work only with the following metrics:

- **Tempo**: Beats per minute (BPM); how fast, medium, or slow a piece of music is played or sung.
- **Danceability (0-1)**: Describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity.
- **Energy (0-1)**: Measures the intensity and activity level of a track. Typically, energetic tracks feel fast, loud, and noisy.
- **Speechiness (0-1)**: Detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value.
- **Valence (0-1)**: Describes the musical positiveness conveyed by a track. Higher valence values indicate more positive and uplifting music, while lower values indicate more negative or melancholic music.
- **Accousticness (0-1)**: A confidence measure of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

Then, the project was developed in javascript and HTML, heavily relying on the D3.js library to build all the visualizations displayed. The interactive potential that the D3 library allows to achieve in the visualization fits with the project's intention since its conception, facilitating its choice.

The project employs a combination of charts to represent the data in an intuitive and visually appealing way. The scrollytelling technique is incorporated to guide users through the narrative, revealing information and visualizations as they scroll down the page. The search bar functionality allows users to explore specific songs and obtain detailed information about their features.

Since the identity of the project was defined, it was clear that the idea of a deep dive should guide the construction of the page, being one of the main focuses in the design and construction of the visualizations. The page structure needed to be in harmony with one of the most essential concepts of the project.

Besides that, another key concept was the color palette. Since the dataset was from Spotify, prioritizing the use of green and black was very important to stay connected to the basis of the project. Using Adobe Color's Color Wheel tool, the following analogous color palette was chosen for the last visualization, centered on Spotify's main color, but defining different colors for each feature:

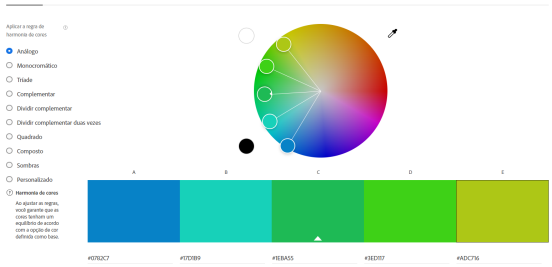

Fig. 1. Choosing the color pallete.

IV. Results

A. First visualization: a time series of musical features

In the first visualization, we get a set of line charts that show the evolution over the years of the average of the different features explore. The lines that appear in a left-to-right motion are the representation of a wave, a fluid and moving design that refers to the concept of sound waves.

In addition to allow consumers to broadly analyze the change in music through each metric, this visualization set the stage for the materialization of the other concepts. The ease and straightforwardness of the line chart connect with the idea of a first step in an ocean dive, starting at the shallowest part of the sea.

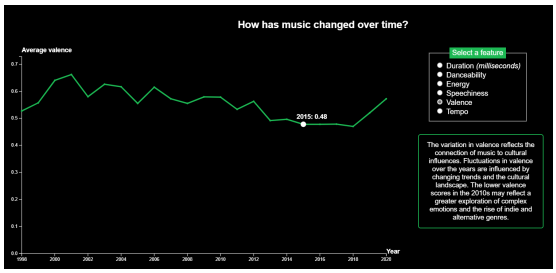

Fig. 2. First visualization.

B. Second visualization: a genre overview

Next, we have a bar chart of the most popular genres, according to the number of songs in the top hits.

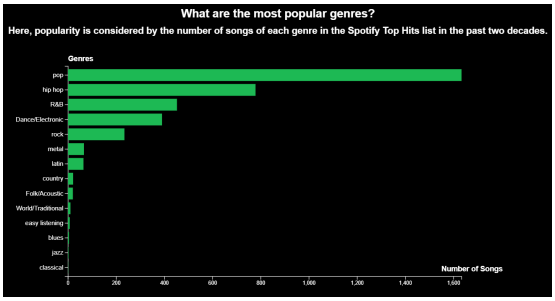

Fig. 3. Bar chart.

There is also the possibility to click on each genre and explore how the songs in that genre perform, according to their popularity and energy, making the dive even deeper, adding another layer of interactivity, as it allows the user to select their favorite genre, get more information about it, and even view the songs individually.

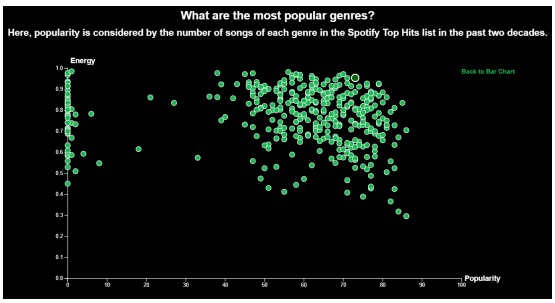

Fig. 4. Visualization after the click in one bar from the previous chart.

Since some genres have a huge number of songs, it is sometimes difficult to find a specific track, limiting the user's ability to fully interact with the data in front of them. From this we set the stage for the final dive, fulfilling its role of being the middle of the ocean, with the user not yet fully emerged.

C. Final visualization: a specific music view

For the final visualization, a search bar has been added to allow the user to search for their favorite songs, receiving all the scores the song got on each analyzed metric (energy, valence, danceability, acousticness, speechiness) in a circular graph, with its popularity value in the center.

With all features (besides popularity) being rated from 0 to 1, the circle chart was chosen because of the way it allows the user to quickly identify whether songs scored well or poorly based on filling in the circles, with a tooltip that allows them to see the exact value. In addition, the arcs reference the Spotify logo, again bringing that visual connection to the user.

This technique is linked to Gestalt's Principles of Continuity (which suggests that we tend to perceive continuous lines or shapes, even when they are interrupted by other elements) and closure (which suggests that we tend to perceive incomplete objects or complex figures as complete and recognizable shapes). These principles attempt to explain how the human

mind can interpret visualizations, finding patterns, recognizing shapes, and are fundamental principles to keep in mind when creating an appealing visualization that aims to be informative at the same time.

Besides that, the results of the songs are compared to the other artists' tracks, giving the user full information about the songs/artists of interest, helping the visualization to be the final tie that completes the deep dive concept, creating the most personal and interactive visualization of all.

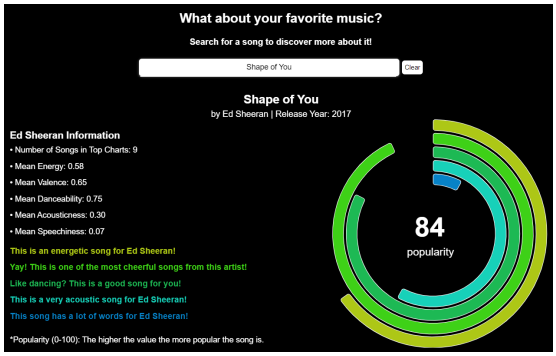

Fig. 5. A specific music visualization.

D. Broad Results

The system produces a range of visualizations that facilitate the exploration of Spotify's music data.

The first visualization showcases the average values of different metrics, such as duration, allowing users to observe trends and patterns in the musical landscape over the years. The genre popularity view provides insights on the prevalence of different genres in the top charts, with the opportunity to explore individual genres further through scatter plots.

The arc chart visualization allows users to examine the features of specific songs and compare them to the mean values of the respective artists. The interactive nature of the visualizations encourages users to actively interact with the data and better appreciate the intricacies of Spotify's music collection.

From the color palette to the idea of deep dives to the choice of specific graphics and visualizations, every design decision was thoroughly considered to provide a set of informative yet fun and complete visualizations, allowing users to see the big picture while at the same time learning about specific features.

V. Discussions

Through "Soundscapes," the audience has the opportunity to gain valuable insights into the musical landscape of the past two decades. By exploring the characteristics of popular songs on Spotify, users can develop a greater comprehension of how music has evolved over time and discover new trends and patterns. The interactive nature of the visualizations improves the user experience by allowing personalized exploration and fostering a sense of discovery.

Although a full user study was not performed, informal observations suggest that the project successfully immerses users in exploring the data and provides a platform for meaningful interactions. The scrollytelling narrative and interactive visualizations encourage users to actively participate in the exploration process, resulting in a more immersive and informative experience.

VI. Conclusion

Since its conception, the project aimed to reach people with very different musical tastes, appealing to an almost universal pleasure, music. With the right data set, the page could bring all kinds of information to users interested in visiting it.

Beyond future ideas, the identity and concepts given to the project joined the broad appeal to help build an interesting storytelling and personal stamp to the page, an uniqueness that is essential to stand out from the pack.

This uniqueness reflects the uniqueness that surrounds music inherently, with each individual having their own unique taste, favorite songs, favorite artists and playlists. And the project offers everyone the opportunity to discover more about their uniqueness.

VII. Future Work

In the future, there are several opportunities to extend and refine the "Soundscapes" application.

Spotify's API integration could be incorporated to provide a personalized analysis of users' playlists. Another idea is the inclusion of user-generated playlists and recommendations that could enhance the personalization aspect of the project, allowing users to discover new music based on their preferences. These user-generated playlists could then serve as the basis for recommendations. The system could analyze the characteristics and patterns of the songs within the playlists and suggest similar tracks or artists that users may enjoy. This recommendation functionality would consider the audio features of the songs, such as valence and energy.

This personalization aspects increase user engagement and satisfaction, as it allows them to actively participate in shaping their musical journey within the "Soundscapes" project.

References

[1] J. Bateira, F. Gouyon, M. Davies, and M. Caetano, "Music Discovery in Spotify with RAMA", in Proceedings of the International Society for Music Information Retrieval Conference, 2014.

[2] S. Kim, et al., "Music-Circles: Can Music Be Represented With Numbers?", 2021. Available at: ¡ https://musiccircles.netlify.app/ ¿. Accessed: June 2023.

[3] L. Guedes, "MusicVis: Interactive Visualization Tool For Exploring Music Rankings," Federal University of Rio Grande do Sul, 2017.

[4] "Top Hits Spotify from 2000-2019," Kaggle.