Ephemeral Music Comprehension

March 29, 2024

1 Introduction of Data

Dataset 1: FinalDatabase_updated.csv

The dataset provided contains information related to music tracks available on Spotify. It includes various attributes such as country, URI (Uniform Resource Identifier), popularity, title, artist, album/single designation, genre, artist followers, explicit content indication, release date, track number, danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, liveliness, valence, tempo, duration in milliseconds, time signature, and several sentiment analysis metrics.

Additionally, the dataset features newly categorized genres, days since release, explicit content indicators (both false and true), and various sentiment analysis scores normalized across different methodologies. It also includes country-specific indicators for sentiment and genre distribution, along with flags indicating whether the track is part of an album, compilation, or a single release.

This comprehensive dataset aims to provide insights into the popularity, characteristics, and sentiment analysis of music tracks across different regions and genres, facilitating further analysis and exploration in the realm of music streaming platforms like Spotify.

Dataset 2

This dataset was used for the Country to Popularity Map . It is available here. The data was preprocessed to extract the required data.
The main focus for this phase of the project was on the columns of **Nationality** and **Points (Ind for each Artist/Nat)** along with **Continent** and **Date**, as the artists were treated individually and per day data was required. First the data was processed by calculating the total sum of points per country and also average score for each country.

2 Important Parts of Data

In this phase we shall majorly be focusing on the following colums:

- Country Name: This column represents the name of the country where the music track is popular or where it is being analyzed. It serves as a geographical identifier for regional analysis.
- Popularity: This column indicates the popularity of the music track on Spotify. Popularity values typically range from 0 to 100, with higher values indicating greater popularity among Spotify users.
- Title: This column contains the title or name of the music track. It serves as a unique identifier for each track and is essential for reference and visualization purposes.
- Explicit (Binary Value): This column is a binary indicator (0 or 1) that denotes whether the music track contains explicit content. A value of 1 typically indicates that the track contains explicit language or themes, while a value of 0 indicates that it does not.
- Genre: This column specifies the genre or category to which the music track belongs. Genres categorize music based on stylistic or thematic characteristics, providing insights into the musical style and audience preferences.

These columns provide crucial information for our analysis and visualization tasks. Using them, we have created heat map and choropleth to visualize the popularity, explicit content, and genre distribution of music tracks across different countries, thus, identifying trends, preferences, and patterns in music consumption behavior worldwide.

3 Visual Design

3.1 Country to Popularity Map

This visualization is designed for emerging artists and music enthusiasts. It features two independent toggles for year and continent, providing users with precise control over the data. The use of country flags to represent data points simplifies the user's experience with the visualization, making it more accessible and effective.

Throughout most years, an evident upward trend can be observed, representing a common pattern across all countries: achieving a higher rating is often a result of an increase in the number of better-performing artists, indicating that artists from these countries are in tune with industry trends. However, outliers exist, primarily in the bottom-right and top-left quadrants of the plot. Countries located in the bottom right owe their position to a select few artists performing exceptionally well, raising the country's average performance even though their total score remains modest. On the other hand, countries in the top left have a significant number of artists who might not chart as highly but collectively contribute to a substantial total score due to their numbers.

This detailed breakdown highlights the diversity of musical success across different regions and periods, providing valuable insights for artists planning their journey toward recognition and success in the global music scene.

3.2 World Choropleth of Explicit Content

This is a choropleth which assigns color ranging from blue to deep red for percentage of explicit songs out of the total songs of the country which were in the top 200 lists of spotify in that month (it changes every year). User can use the arrow keys to switch between the months and years to see the changing trends at a glance. On hovering over any country, users can see a tooltip which shows details like the country's name, the percentage of explicit songs, and also the actual number of explicit songs and number of total songs. On hovering, a heatmap is also shown which is the heatmap of ditribution of genres of that month and year of that country. You can click on the country to solidify this heatmap. You can solidify multiple heatmaps across countries and different months and years to compare genre trends between countries or of the same country across years according to your preference. Double click on that same country in the same month and year to remove this solidified heatmap.

3.3 Heatmap of genre distribution of countries across years

This is a compiled heatmap, which shows same content as the tooltip heatmaps of above graph, but in a well formated and compiled way. Users which only want to see the genre trends etc can use this heatmap.

4 Aim of Visualisation

4.1 Country to Popularity Map

This visualisation is aimed to give upcoming artists an image of how artists in their country and also their region perform , this along with the data on genre will help them plan out their strategies for growth. It visualises the past popularity and trends of a country. It helps artists discover untapped markets or find inspiration in the success stories of others, leading to innovative music projects that cross traditional boundaries of genre and geography.

4.2 Choropleth of percentage of explicit songs out of top 200 songs of each country across the years

This graph displays the entire world map with the countries that the database has data for being shown in various colors ranging from a light blue for low explicit content to deep red for high explicit content.
This graph is needed because it tackles a few problems which either SQL and traditional plots cannot solve or it is difficult for them to solve.
The first advantage of a choropleth: when people see many country names along

with percentage data associated with each of them, it is confusing to recognize each country, its position, etc while trying to analyse trends in explicitness.
Choropleth removes those worries by clearly showing each country in its proper place with proper coloring to indicate, with the degree of darkness of red color, how much is the percentage of explicit songs. The user can see at one glance how the trends in explicit songs shift around the world, and how different countries in different geographical groups display different trends in explicit percentage (such as the east and the west; this is important as the west tends to be more liberal with explicit language).
Another advantage of choropleth is that instead of comprehending a long string of numbers and iterating through them to find hotspots, the piercing red of countries with high explicit content pop out at the user and the calming blue of low explicit countries can be seen at a glance.
Hovering over a country also shows a heatmap of the country of the percentage of songs of each genre of the selected month and year which can be used to identify the most popular genres at a glance so that users can make a correlation between explicitness percentage of a country and the famous songs in that country in the selected month (as genres like rap and hip-hop tend to have more explicit lyrics so such genres being popular would lead to a trend of more explicit songs in that country).
Users can solidify the heatmap tooltip by clicking on the country and thus they can solidify multiple heatmaps of genres across the months of the same country (one below the other) to see the trends or even compare heatmaps of multiple countries. You can create custom couplings of different heatmaps according to requirements. This offers much more versatility and easier comprehension than a bar graph or other traditional data visualizations, or pure data output by programs.

4.3 Heatmap of genre distributions of songs across the years of each country

The same heatmaps, parts of which are shown in above choropleth as tooltips are presented in a well formatted fashion here, so users who want to analyse only the genre data can easily identify the most popular amongst the many genres according to the month and year selected by them.