Music Metadata Analysis: Exploring Patterns in Song Attributes

I. Objective

To conduct an in-depth analysis of song metadata to identify and explore patterns related to song popularity, genre distinctions, and artist trends, with the ultimate goal of developing a song recommendation system.

II. Literature review

In the digital age of streaming services, music is at our fingertips, complete with extensive music metadata. Despite the vast array of musical genres and the constant emergence of new artists, the key to a song's or artist's popularity often lies in the music's attributes, like tempo and key.

Listed below are some previous research efforts that have aimed to analyze various aspects of music data to identify prevailing trends within the music industry.

The work of Rentfrow et al. [5] replicates and extends prior work on individual differences in music-genre preferences, which suggested four to five robust music-preference factors. It encompasses a broad spectrum of musical styles and assesses preferences for multiple music pieces. The findings from three independent studies consistently converge, revealing five distinct music-preference dimensions: a "Mellow" factor, characterized by smooth and relaxing musical styles; an "Urban" factor defined primarily by rhythmic and percussive music; a "Sophisticated" factor composed of complex, intelligent, and inspiring music; an "Intense" factor representing loud, forceful, and energetic music; and a "Campestral" factor encompassing various styles of country and singer-songwriter music.
The authors of this study [3] address the challenges of managing music-related data within the Music Information Retrieval (MIR) framework. MIR, a research area gaining prominence, plays a pivotal role in supporting audio streaming platforms like Spotify, Lastfm, and Billboard. It empowers these platforms with algorithms and services such as recommender systems, music similarity analysis, and in-depth popularity assessment, ultimately enhancing the quality of user experience. The paper introduces two core contributions Unified Multimodal Database and a HitMusicNet Architecture. The Unified Multimodal Database process includes the creation of a comprehensive database by merging data from Spotify and Genius, encompassing over 100,000 tracks from 26 countries. The database includes high-level and low-level audio features, text features extracted from lyrics, artist information, track popularity, genres, audio previews, and complete lyrics. The HitMusicNet Architecture is a novel end-to-end Deep Learning (DL) architecture, named HitMusicNet, which is presented for music popularity prediction. HitMusicNet outperforms previous approaches and addresses both regression and classification perspectives. It surpasses existing Machine Learning (ML) techniques that have limitations in learning complex functions. The proposed architecture adopts a soft-decision approach, estimating popularity values along with corresponding errors, rather than relying on subjective popularity thresholds. This approach enhances the precision of popularity prediction.

The study [4] focuses on "Hit Song Science" (HSS), a task involving the prediction of a song's commercial success based on audio features before its release. It challenges the notion that popularity prediction is infeasible and presents positive results using state-of-the-art machine learning algorithms. The investigation, spanning five decades of the UK top 40 singles chart, aims to distinguish top 5 hits from those ranked 30-40. It organizes the dataset chronologically and employs the Shifting Perceptron to account for evolving musical preferences over time. Key findings highlight trends in song characteristics: slower songs, such as ballads, were favored from the 1980s to the 1990s, evolving to a preference for faster songs in the new century. Simple songs with clear time signatures and harmonic simplicity remained hits across decades. Song duration increased over time, and louder songs became more popular. Compared to previous studies, this research offers more optimistic results, likely due to variations in the task description, the inclusion of novel audio features, and the application of the time-shifting perceptron. Future investigations will explore different study designs and tasks related to popularity label prediction, enhancing our understanding of this dynamic field.

III. Dataset

In our study, we aim to employ two publicly available datasets to evaluate and answer the questions we have surrounding songs and how musical features, when analyzed properly, can help the artists and the music industry at large. However, due to the reasons mentioned below, we have decided to proceed with the best dataset out of the two analyzed.

A. *Million Song Dataset*

Million Song Dataset [2] is a comprehensive collection of contemporary popular music tracks, designed to facilitate research in music information retrieval (MIR) and related fields. Created through a collaboration between The Echo Nest and LabROSA at Columbia University, the dataset comprises a rich array of metadata and audio features for a substantial number of songs, although the exact count may vary. The dataset includes information such as artist details, genre labels, and a diverse set of audio features, offering a valuable resource for tasks ranging from music recommendation systems to genre classification and artist similarity analysis.

B. Spotify Web API [1]

Spotify requires users to create a developer account to have access to the portal and make API calls. Upon successfully creating an account, users are required to create an app the user will be given a client ID and client secret to make calls to the different API endpoints. This report utilized three endpoints; artist, tracks [6], and audio features endpoints. Due to the limitation on how many calls to make to an endpoint at a time, we queried and cached artists for the year 2023 up to 1000 records in chunks of 50 records per API.

For this project, we have chosen to use the Spotify Web API for several reasons. Firstly, the Spotify Web API contains up-to-date data that is refreshed daily, whereas the Million Song Dataset does not include recent-year data. Furthermore, the Spotify Web API allows us to query specific years and look up songs by song ID. Additionally, it enables us to utilize daily updated playlists, categorized by genre, artist, popularity, etc.

We created a function that searches for artists for the year 2023 and stored the cached results in a JSON file. This was done to prevent re-calling the API for the same records the next time we run our code. This concept was also applied in generating the dataset for tracks. We queried the track endpoint for all tracks in 2023, retrieving up to 1000 records, and stored the cached results in a JSON file.

To obtain the audio features crucial to our project, we extracted the track IDs from the track dataset into a list. This list of track IDs was enumerated in batches of 50, as per the API call requirement. The audio features were then retrieved for each batch and concatenated. The results of the track ID and its corresponding audio features were stored in a data frame and finally converted to a CSV to avoid re-querying the API.

For context, Table I shows the features in the artist data as acquired from the Spotify Web API.

Table II illustrates the artist data structures as outlined in the Spotify API documentation. The artist data incorporates nested attributes such as 'album' and 'artists.' Notably, certain attributes were excluded from our final track data as they were deemed irrelevant to our study.

In the context of our track data, we gather the audio features for each individual track. Table III details the attributes associated with each track, and the utilization of these features serves as the cornerstone of our work. Moving forward to Table IV, we showcase the data frame structure crafted from information extracted from the Spotify Web API, covering artist, track, and audio features data. The IDs within these datasets share relationships, streamlining the merging process. This ultimate dataset will be instrumental in our exploratory data analysis (EDA) and machine learning endeavors.

TABLE I
ARTIST DATA STRUCTURE

Attribute	Meaning
external_urls	Known external URLs for the artist
followers	Information about the followers of the artist (nested structure)
followers.href	Link to the Web API endpoint for followers details
followers.total	Total number of followers
genres	List of genres associated with the artist
href	Link to the Web API endpoint for full artist details
id	Spotify ID for the artist
images	Images of the artist in various sizes
name	Name of the artist
popularity	Popularity score of the artist (0 to 100)
type	Object type (usually "artist")
uri	Spotify URI for the artist

TABLE II
TRACK DATA STRUCTURE

Attribute	Meaning
album	Object containing details about the album (nested structure)
album_type	Type of the album (e.g., "album", "single", "compilation")
total_tracks	Number of tracks in the album
available_markets	Markets where the album is available (ISO 3166-1 alpha-2 country codes)
external_urls	Known external URLs for the album
href	Link to the Web API endpoint for full album details
id	Spotify ID for the album
images	Cover art for the album in various sizes
name	Name of the album
release_date	Date the album was first released
release_date_precision	Precision with which release_date value is known ("year", "month", "day")
restrictions	Content restriction information (if applied)
artists	Array of artists who performed the album (nested structure)
artists.external_urls	Known external URLs for the artist
artists.href	Link to the Web API endpoint for full artist details
artists.id	Spotify ID for the artist
artists.images	Images of the artist in various sizes
artists.name	Name of the artist
artists.popularity	Popularity score of the artist (0 to 100)
artists.type	Object type (usually "artist")
artists.uri	Spotify URI for the artist
available_markets (in track)	List of countries in which the track can be played (ISO 3166-1 alpha-2 codes)
disc_number	Disc number of the track (usually 1 unless multi-disc album)
duration_ms	Duration of the track in milliseconds
explicit	Indicates whether the track has explicit lyrics (True/False)
external_ids	Known external IDs for the track
external_urls (in track)	Known external URLs for the track
href (in track)	Link to the Web API endpoint for full track details
id (in track)	Spotify ID for the track
is_playable	Indicates whether the track is playable in the given market
linked_from	Information about the original track (when Track Relinking is applied)
restrictions (in track)	Content restriction information (if applied)
name (in track)	Name of the track
popularity (in track)	Popularity score of the track (0 to 100)
preview_url (in track)	Link to a 30-second preview of the track (MP3 format, nullable)
track_number	Number of the track within the album
type (in track)	Object type (usually "track")
uri (in track)	Spotify URI for the track
is_local (in track)	Indicates whether the track is from a local file

IV. EXPLORATORY ANALYSIS

We were tasked with addressing specific questions using the dataset at our disposal, with the overarching goal of understanding their implications for the music industry, artists, and listeners.

In our preliminary analysis, we explored potential correlations among the features identified for our Exploratory Data

TABLE III
AUDIO FEATURES DATA STRUCTURE

Attribute	Meaning
danceability	Danceability describes how suitable a track is for dancing
energy	Energy is a measure from 0 to 1 representing a perceptual measure of intensity and activity
key	The key the track is in. Integers map to pitches using standard Pitch Class notation
loudness	The overall loudness of a track in decibels (dB)
mode	Modality of the track. Major is represented by 1 and minor is 0
speechiness	Speechiness detects the presence of spoken words in a track
acousticness	Acousticness is a confidence measure from 0 to 1 of whether the track is acoustic
instrumentalness	Predicts whether a track contains no vocals. 0 represents high confidence that the track is not instrumental
liveness	Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live
valence	A measure from 0 to 1 describing the musical positiveness conveyed by a track
tempo	The overall estimated tempo of a track in beats per minute (BPM)
duration_ms	Duration of the track in milliseconds
time_signature	An estimated overall time signature of a track. The time signature (meter) is a notational convention to specify how many beats are in each bar or measure

TABLE IV
OUR FINAL DATA STRUCTURE FOR ANALYSIS

Attribute	Meaning	Source
artist_id	Spotify ID of the artist	Artist Endpoint
artist_name	Name of the artist	Artist Endpoint
artist_popularity	Popularity of the artist (0 to 100)	Artist Endpoint
artist_followers	Number of followers of the artist	Artist Endpoint
artist_genres	Genres associated with the artist	Artist Endpoint
song_id	Spotify ID of the song	Track Endpoint
song_name	Name of the song	Track Endpoint
album_id	Spotify ID of the album	Track Endpoint
album_name	Name of the album	Track Endpoint
track_popularity	Popularity of the track (0 to 100)	Track Endpoint
track_available_markets	List of countries where the track is available	Track Endpoint
duration_ms	Duration of the track in milliseconds	Track Endpoint
track_number	Number of the track in the album	Track Endpoint
album_release_date	Release date of the album	Track Endpoint (Extracted from Album data)
album_total_tracks	Total number of tracks in the album	Track Endpoint (Extracted from Album data)
album_type_y	Type of the album (e.g., album, single, compilation)	Track Endpoint (Extracted from Album data)
danceability	Danceability measure of the track	Audio Features Endpoint
energy	Energy measure of the track	Audio Features Endpoint
key	Key of the track	Audio Features Endpoint
loudness	Loudness of the track in decibels (dB)	Audio Features Endpoint
mode	Modality of the track (major or minor)	Audio Features Endpoint
speechiness	Speechiness measure of the track	Audio Features Endpoint
acousticness	Acousticness measure of the track	Audio Features Endpoint
instrumentalness	Instrumentalness measure of the track	Audio Features Endpoint
liveness	Liveness measure of the track	Audio Features Endpoint
valence	Valence measure of the track	Audio Features Endpoint
tempo	Tempo of the track in beats per minute (BPM)	Audio Features Endpoint
time_signature	Estimated time signature of the track	Audio Features Endpoint
flattened_genres	Flattened list of genres associated with the artist	Artist Endpoint (Genres flattened for analysis)

Analysis and subsequent training of the Machine Learning model, utilizing the Spotify Web API.

In Figure 1, we present the Pearson Correlation of features within our dataset. The figure highlights a robust correlation of 0.8 between artist popularity and artist followers. Similarly, a substantial correlation of 0.75 is observed between audio features, such as loudness and energy. It's noteworthy that songs with high energy levels often tend to be louder.

Additionally, there's a notable correlation of 0.70 between track features, like track number and album total tracks. This correlation is expected, as songs are typically either standalone singles or part of an album, explaining the connection between track number and the total tracks on an album.

Figure 2 depicts as per our dataframe the number of songs in each genre. The pop genre appears to dominate the other genres. Closely following "pop" by a margin of 10 is "rap".

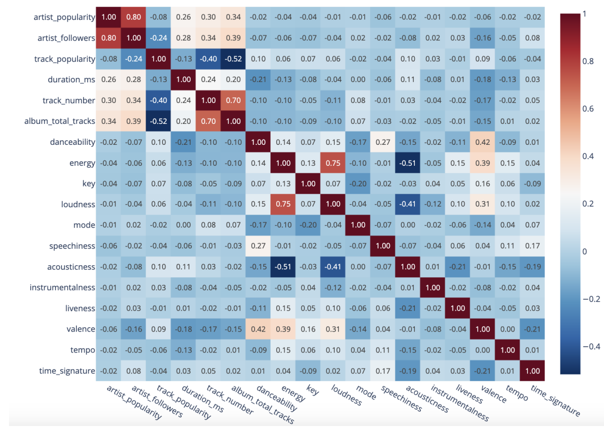

Fig. 1. Pearson Correlation of features

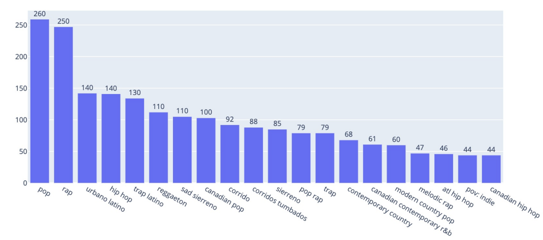

Fig. 2. Top genres and the number of songs

There is a considerable difference in the number of songs in all the genres after pop and rap.

1) Genre Distinctions and Attributes: **How do song attributes like tempo, key, and loudness vary across genres? What attributes make a genre unique or distinguishable from others?**

To answer this question we filtered the top 25 genres and worked with it. Figure 3, 4 5 provides a visual summary of the distribution and central tendency of tempo, key, and loudness across different musical genres.

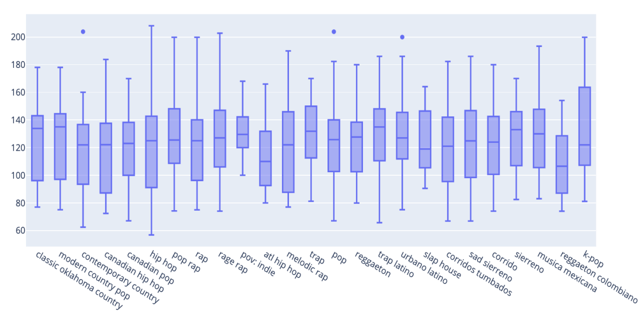

Fig. 3. Variation of Tempo on the top 25 genres

Loudness exhibits a higher potential for outliers compared to attributes like key and tempo, particularly in genres such as rap, trap, Latino, etc. There is a noticeable variation in tempo, key, and loudness across different genres, as depicted

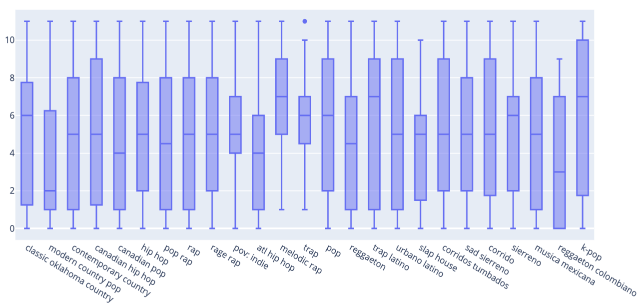

Fig. 4. Variation of Key

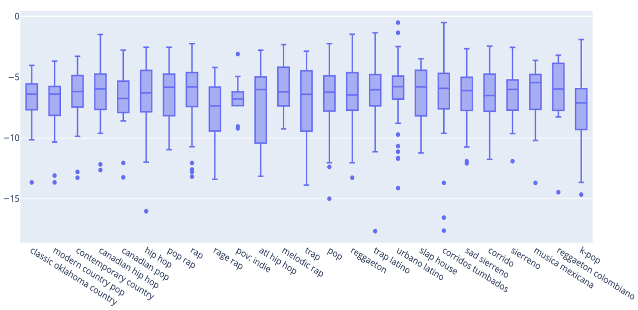

Fig. 5. Variation of Loudness

by different heights in the figure 3, with each box representing the interquartile range (IQR) containing the middle 50

Examining the first three genres in the figure 3, the line representing the median (50th percentile) appears to indicate a left-skewed distribution, given its proximity to the upper part of the box. This suggests that the median is greater than the mean, indicating a longer tail to the left. Conversely, the 'key' attribute in the genre 'modern country pop' appears right-skewed, with the median closer to the lower part of the box, suggesting a longer tail to the right.

Moving on to the discussion of popularity and trends, we explore the attributes most strongly correlated with a song's popularity in figure 6. While anticipating that features like artist popularity and artist followers would show a high correlation, the figure surprisingly reveals that the number of songs on an album is the most highly correlated feature with a song's popularity. Furthermore, audio features like danceability, valence, and acousticness, associated with genres like pop, also demonstrate correlations with songs that tend to be popular.

How have song attributes evolved over the decades? For example, has song duration or loudness changed over time? Can we predict the potential success of a song based on its attributes?

Utilizing our dataset, we applied the RandomForestRegressor to predict a track's popularity. The model was trained on 80% of our data, reserving the remaining 20% for testing. The results include Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) values of 67.64, 4.87, and 8.22, respectively.

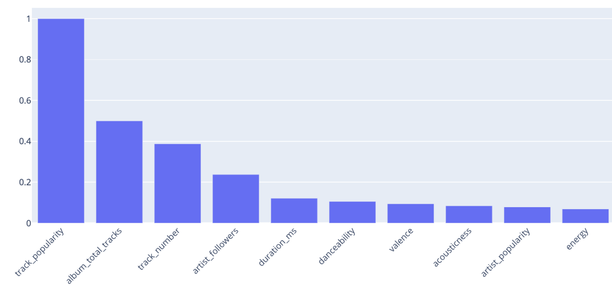

Fig. 6. Top Correlated Attributes with Popularity

The RMSE of 8.22 indicates that, on average, the model's predictions deviate by approximately 8.22 units from the actual values. The notable improvement in accuracy is evident when comparing the model's MSE to the baseline MSE. The low values of MSE, MAE, and RMSE collectively highlight the model's strong performance.

In conclusion, our model effectively predicts the target variable, demonstrating its efficacy in the given context.

2) Artists and Their Signature Styles: **Do established artists have a consistent pattern or "signature" in their song attributes?** In Figure 7, we observe artists whose popularity surpasses 85, along with their distribution across various genres. Notably, Drake stands out with affiliations in four distinct genres: "hip hop," "rap," "pop rap," "Canadian pop," and "Canadian hip hop." In contrast, an artist like Taylor Swift is primarily associated with a single genre, namely "pop." This analysis aimed to investigate whether there exists any correlation between the number of genres an artist engages in and their followers and popularity.

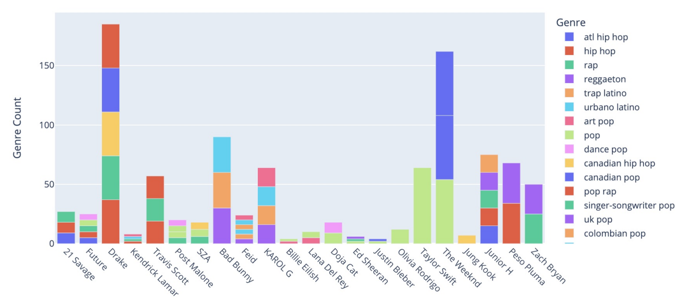

Fig. 7. Genre Distribution Across Artists

In Figure 8, the correlation between the number of genres and the popularity and followers of artists is illustrated. Notably, Taylor Swift, recognized for her association with the "pop" genre exclusively, boasts a popularity score of 100 and a follower count exceeding 90 million.

In contrast, Ed Sheeran, affiliated with three genres, has a popularity rating of 86 and an impressive following of 115 million. Artists like Drake, Bad Bunny, and The Weeknd, spanning 5, 3, and 3 genres respectively, have substantial followings, with follower counts of 81 million, 76 million,

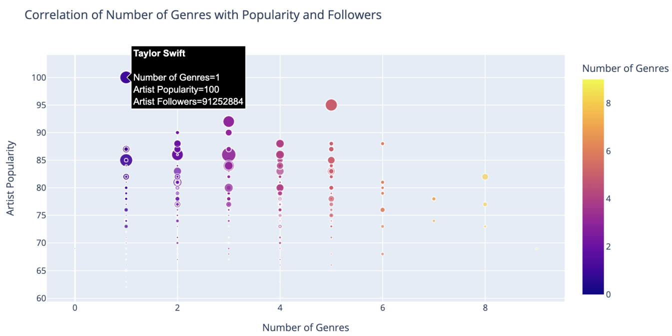

Fig. 8. Correlation of Number of Genres with Popularity and Followers

and 73 million, coupled with popularity scores of 95, 96, and 92, in that order.

3) Clustering and Recommendation: **Can we develop clusters of songs that might cater to similar audiences based on their attributes?**

We aim to use K-means clustering for this task. To find an optimal K, we used the heuristic "elbow method". Figure 9 visually helps us decide on the value of k by using an elbow plot where the y-axis is a measure of goodness of fit and the x-axis is the value of k. Hence, k = 5 is probably the best we can do without overfitting.

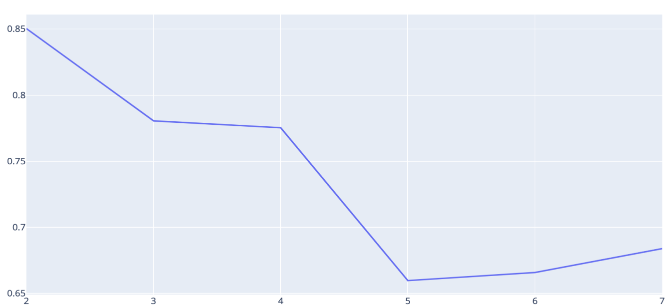

Fig. 9. Elbow Method for Optimal k

Fig. 10. K-means Clustering of Songs

How do song recommendations work in terms of song metadata? Can we devise a basic recommendation algorithm? A basic recommendation algorithm can be developed based on the characteristics of songs and users' preferences. We created a user profile by assigning values to the song attributes. To recommend a song, a calculation of the similarity between the user profile and each song in the dataset is performed. Cosine similarity is employed in our task. In Figure 11, we assigned the following values to these attributes danceability=0.8, energy=0.8, key=5, loudness=-5, mode=1, speechiness=0.1, acousticness=0.2, instrumentalness=0.05, liveness=0.1, valence=0.7, tempo=120. The user profile created generated the figure in 11 where most of the songs recommended appear to belong to the pop genre.

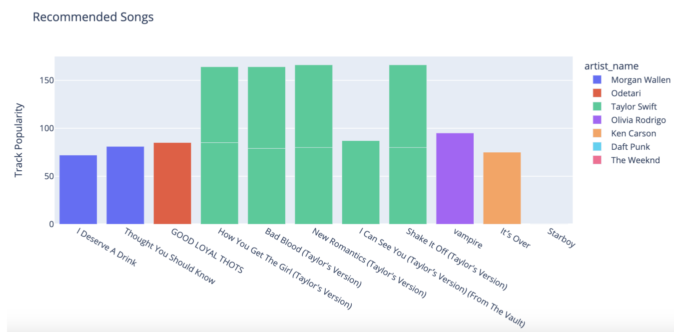

Fig. 11. Recommended Songs

We then created a much similar user profile danceability=0.8, energy=0.7, valence=0.9. Figure 12 was the result of the user profile. The recommended songs appear to be the corrido and sierreno, Latino genres. The user profile and the songs recommended by our recommendation system appear to be songs unique to Latinos. There are outliers like Tyler the Creator (rap), Selena Gomez (pop), and Kelsea Ballerini (country) appears to be an outlier in this case.

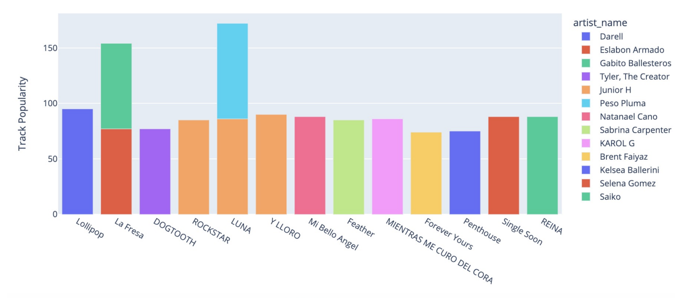

Fig. 12. Recommended Songs

V. Next Steps

In the upcoming project phase, we aim to address the following inquiries:

- How do characteristics of song metadata compare between new/emerging artists and seasoned veterans?
- What distinctions exist in attributes between instrumental songs and those with heavy lyrics?
- Does the song's language or origin (e.g., English vs. Spanish, Western vs. Eastern) influence its characteristics?

- We plan to explore additional datasets such as the Million Song dataset and the Billboard dataset.

Moreover, we offer a comprehensive explanation of the preliminary analyses conducted, encompassing:

- The recommendation system
- The clustering algorithm

REFERENCES

[1] Spotify web api documentation. https://developer.spotify.com/documentation/web-api. Accessed on: November 16, 2023.

[2] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 2011.

[3] David Martín-Gutiérrez, Gustavo Hernández Peñaloza, Alberto Belmonte-Hernández, and Federico Álvarez García. A multimodal end-to-end deep learning architecture for music popularity prediction. *IEEE Access*, 8:39361–39374, 2020.

[4] Yizhao Ni, Raúl Santos-Rodríguez, Matt McVicar, and Tijl De Bie. Hit song science once again a science? 2011.

[5] Peter J. Rentfrow, Lewis R. Goldberg, and Daniel J. Levitin. The structure of musical preferences: a five-factor model. *Journal of Personality and Social Psychology*, 100(6):1139–1157, 2011.

[6] Spotify. Get a track. https://developer.spotify.com/documentation/web-api/reference/get-track, 2023. Accessed on: November 10, 2023.