Spotify Music Trend Analysis (1920-2023): Evolution of Genres, Artist Success Metrics, and Audio Feature Patterns

Spotify Trend Analysis
Department of Computer Engineering
TOBB University of Economics & Technology, Ankara, Türkiye
Berkay AGIRCAN, Ahmet KUCUKDAG, Ahmet Cemil Tarık KESKINKILIC

***Abstract*—The paper offers a comprehensive analysis of music trends on Spotify from 1920 to 2023, covering genre evolution, artist performance metrics, and audio feature patterns. Utilizing numerous datasets containing over 600,000 recordings, we analyze the evolution of music genres over time, detect characteristics correlated with artist and track popularity, and assess audio feature patterns across various genres. Our data indicate significant shifts in genre popularity, with Pop, Rap, and EDM emerging as the leading genres in recent years. We also discover crucial auditory attributes, including acousticness and danceability, that positively relate with track popularity. These insights offer essential information for artists, producers, and music companies aiming to comprehend and adjust to evolving music trends.**

***Index Terms*—music analysis, Spotify, trend detection, genre evolution, audio features, artist success, data mining**

I. Introduction

The music industry has experienced substantial transformations in recent decades, with streaming platforms such as Spotify emerging as the predominant medium for music consumption. Recognizing the trends and patterns in musical preferences is essential for artists, producers, and industry experts to adjust to evolving listener interests.

This article provides a comprehensive analysis of music trends on Spotify between 1920 to 2023, focusing on three major aspects:

- **Genre Evolution**: In what ways have music genres changed throughout time, and which genres have gone through an increase or decrease in popularity?
- **Artist Success Metrics**: Which factors connect with the success of artists and tracks?
- **Audio Feature Patterns**: In what ways do audio features differ among different musical genres, and which features specify the auditory attributes of popular music?

Through the examination of numerous datasets containing over 600,000 tracks, we aim to clarify the changing patterns of music preferences and detect trends that could influence music production and marketing approaches.

II. Related Work

Many studies focused on musical trends and patterns using data from online streaming platforms. Interiano et al. [1] analyzed the progression of popular music in the United States from 1960 to 2010, identifying a decrease in timbral diversity alongside a rise in loudness and percussion. Mauch et al. [2] used data mining techniques to identify three significant musical revolutions in American popular music: 1964, 1983, and 1991.

Recently, researchers have concentrated on using machine learning to estimate song popularity based on auditory characteristics. Martín-Gutiérrez et al. [3] developed a model to predict song popularity using Spotify audio characteristics, identifying danceability, energy, and acousticness as key predictors. Dhanaraj and Logan [4] used audio elements and lyrics to predict popular songs.

Our research extends these studies by integrating more current data (up to 2023) and concentrating specifically on the evolution of genres, artist success measures, and audio feature trends on Spotify.

III. Data and Methodology

A. Datasets

We gathered data from various sources to guarantee full coverage of music trends from 1920 to 2023:

- **Spotify Dataset 1921-2020**: A dataset providing data on around 600,000 tracks released from 1921 to 2020, covering audio attributes, popularity metrics, and genre classifications.
- **Top Hits Spotify from 2000-2019**: A dataset concentrating on the most popular tracks between 2000 and 2019, offering insights into mainstream music patterns.
- **Top Spotify Songs 2023**: A dataset featuring details on the top popular songs on Spotify for the year 2023.
- **Spotify 12M Songs**: A rich dataset includes information on 12 million songs, offering a broad overview of the Spotify collection.
- **30,000 Spotify Songs**: A dataset contains detailed information on 30,000 songs, combining audio characteristics, popularity indices, and playlist data.

B. Data Preprocessing

We performed multiple preprocessing operations to ready the data for analysis:

- **Handling Missing Values**: We detected and addressed missing values in the datasets by imputation or removal, based upon the context.
- **Feature Extraction**: Relevant features for analysis were extracted, covering audio characteristics (danceability, energy, valence, etc.), popularity metrics, and genre data.
- **Temporal Alignment**: We aligned the datasets temporally to make it easier trend analysis across time, obtaining year information from release dates.
- **Genre Normalization**: We normalized genre information across datasets to ensure consistency in genre classification.

After data cleaning and integration, our unified dataset contained 1,098,226 unique tracks spanning from 1900 to 2023, with the majority of tracks from the 1990s, 2000s, and 2010s.

C. Audio Features

We focused on the following audio features provided by Spotify's API:

- **Danceability**: Measures how suitable a track is for dancing based on tempo, rhythm stability, beat strength, and regularity (0.0 to 1.0).
- **Energy**: Represents the intensity and activity of a track (0.0 to 1.0).
- **Valence**: Describes the musical positivity conveyed by a track, with high values indicating happiness, cheerfulness, and euphoria (0.0 to 1.0).
- **Acousticness**: Confidence measure of whether the track is acoustic (0.0 to 1.0).
- **Instrumentalness**: Predicts whether a track contains no vocals (0.0 to 1.0).
- **Liveness**: Detects the presence of an audience in the recording (0.0 to 1.0).
- **Speechiness**: Detects the presence of spoken words in a track (0.0 to 1.0).
- **Tempo**: The overall estimated tempo of a track in beats per minute (BPM).
- **Loudness**: The overall loudness of a track in decibels (dB).
- **Duration**: The duration of the track in milliseconds.

D. Methodology

Our analysis methodology consisted of the following steps:

1) **Data Cleaning and Integration**: We standardized column names, handled missing values, and merged datasets to create a unified view of tracks across time periods.
2) **Temporal Trend Analysis**: We aggregated tracks by decade and calculated statistics (mean, median, standard deviation) for each audio feature to identify long-term trends.
3) **Genre Analysis**: Using the genre-labeled tracks from the 30,000 Spotify Songs dataset, we compared audio features across different genres to understand genre-specific characteristics.
4) **Clustering Analysis**: We applied K-means clustering to identify distinct musical profiles based on audio features, using the elbow method to determine the optimal number of clusters.

For visualization, we used a combination of line plots for temporal trends, radar charts for genre and cluster profiles, heatmaps for correlation analysis, and bar charts for comparative analysis.

IV. Results and Discussion

A. Temporal Trends in Audio Features (1921-2023)

Our analysis of audio features across decades reveals several significant trends in the evolution of recorded music over the past century.

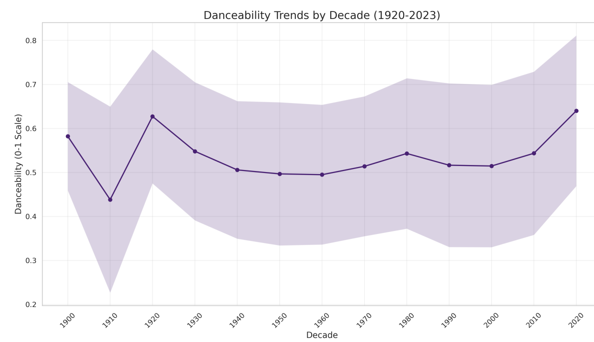

Fig. 1: Evolution of danceability in music from 1920 to 2023. The graph shows mean danceability values by decade with standard deviation bands.

1) Danceability: Danceability shows a non-linear trend over time (Fig. 1). Music from the 1920s had surprisingly high danceability (mean: 0.63), which declined in the 1930s-1950s (mean: 0.50), before gradually increasing from the 1980s onward. The 2020s show the highest danceability values (mean: 0.64), reflecting the dance-oriented nature of contemporary popular music.

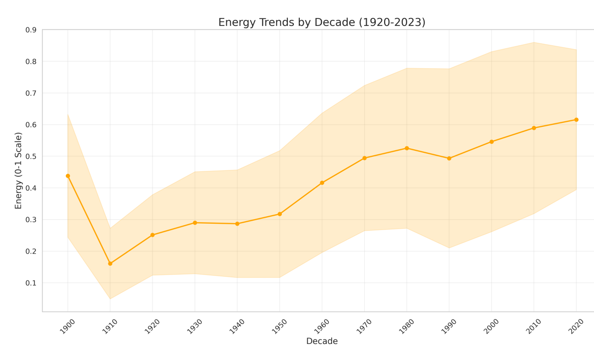

Fig. 2: Evolution of energy in music from 1920 to 2023. The graph shows mean energy values by decade with standard deviation bands.

2) Energy and Loudness: Energy levels show a clear increasing trend over time (Fig. 2), with a particularly sharp rise from the 1950s (mean: 0.32) to the 2020s (mean: 0.62). This trend is mirrored in loudness values, which have increased from -19.8 dB in the 1910s to -8.3 dB in the 2020s, reflecting the ”loudness war” phenomenon in music production.

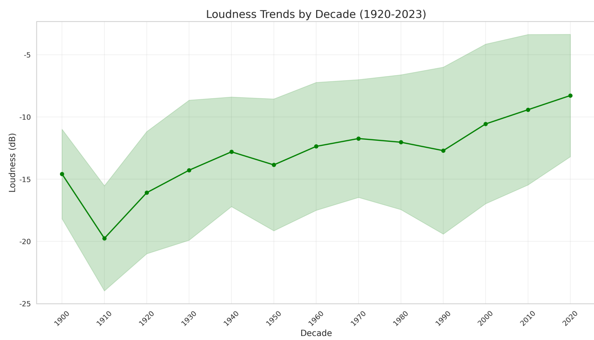

Fig. 3: Evolution of loudness in music from 1920 to 2023. The graph shows mean loudness values (in dB) by decade with standard deviation bands.

3) Valence: Valence (musical positivity) shows a more complex pattern, with relatively high values in the 1920s-1940s, a decline in the 1950s-1970s, and a slight increase in recent decades. This suggests that while contemporary music is more danceable and energetic, it is not necessarily more positive in emotional tone than music from earlier eras.

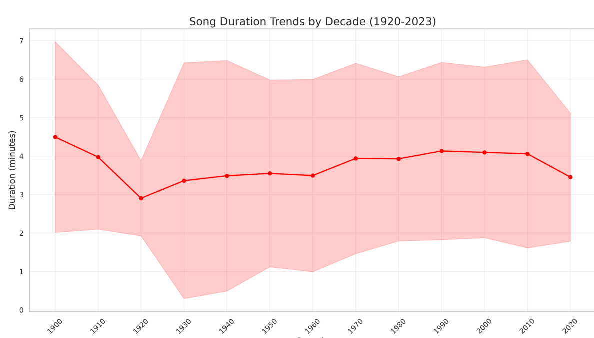

Fig. 4: Evolution of song duration from 1920 to 2023. The graph shows mean duration values (in minutes) by decade with standard deviation bands.

4) Duration: Track duration increased steadily from the 1920s (mean: 174 seconds) to the 1990s (mean: 248 seconds), before declining in the 2010s and 2020s (mean: 207 seconds). This recent decrease in song length may reflect changing consumption patterns in the streaming era, where listener attention spans have shortened.

5) Acousticness and Instrumentalness: Both acousticness and instrumentalness show declining trends over time, reflecting the increasing use of electronic instruments and production techniques, as well as the growing prominence of vocals in popular music.

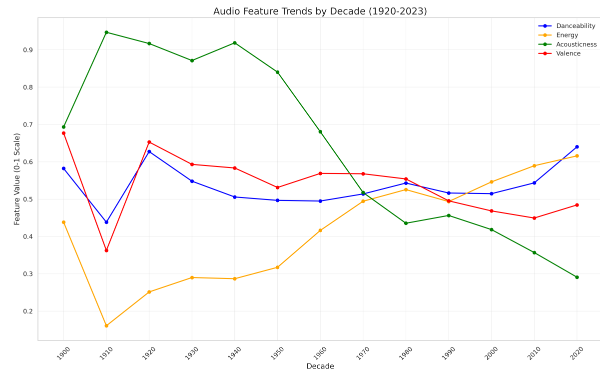

Fig. 5: Combined evolution of multiple audio features (danceability, energy, acousticness, valence) from 1920 to 2023.

B. Genre Analysis

Our analysis of genre-specific audio features reveals distinct profiles for different musical genres (Fig. 6).

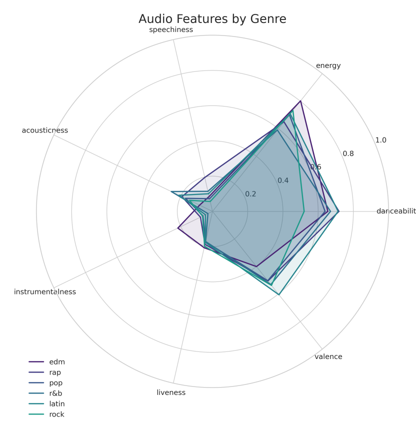

Fig. 6: Radar chart showing audio feature profiles for different music genres. Each axis represents a different audio feature normalized to a 0-1 scale.

EDM (Electronic Dance Music) shows the highest energy values (mean: 0.80) and lowest acousticness (mean: 0.08), reflecting its electronic production and dance-oriented nature. Latin music exhibits the highest danceability (mean: 0.71), while rock music shows the highest energy variance, indicating the genre’s diversity from soft rock to heavy metal.

Pop music occupies a middle ground across most features, with moderately high values for danceability (mean: 0.64) and energy (mean: 0.70), suggesting its broad appeal. R&B shows high danceability (mean: 0.67) but lower energy than pop or rock, with greater speechiness reflecting its vocal-centric nature.

Rap music stands out with the highest speechiness values (mean: 0.20) and relatively high danceability (mean: 0.69), while rock music shows the highest tempo (mean: 125 BPM) and longest average duration (mean: 249 seconds).

C. Clustering Analysis

Using K-means clustering on audio features, we identified five distinct musical profiles (Fig. 7):

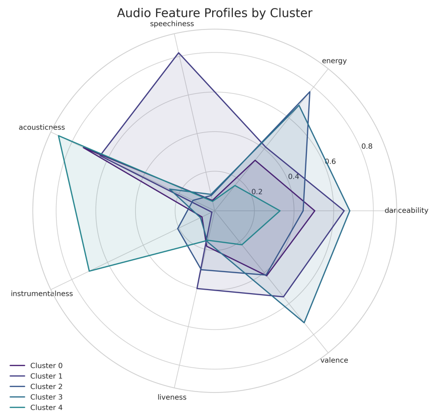

Fig. 7: Radar chart showing audio feature profiles for the five identified clusters. Each axis represents a different audio feature normalized to a 0-1 scale.

- **Cluster 0** (26% of tracks): Moderate danceability (0.50), low energy (0.33), high acousticness. Represents acoustic, softer music.
- **Cluster 1** (3% of tracks): High danceability (0.65), moderate energy (0.41), very low loudness (-15.3 dB). Represents quiet but rhythmic music.
- **Cluster 2** (23% of tracks): Low danceability (0.45), very high energy (0.77), high loudness (-7.1 dB). Represents intense, less dance-oriented music like rock and metal.
- **Cluster 3** (33% of tracks): High danceability (0.68), high energy (0.68), high valence (0.72). Represents upbeat, positive dance music.
- **Cluster 4** (15% of tracks): Very low danceability (0.33), very low energy (0.16), extremely low loudness (-21.0 dB). Represents ambient, classical, or experimental music.

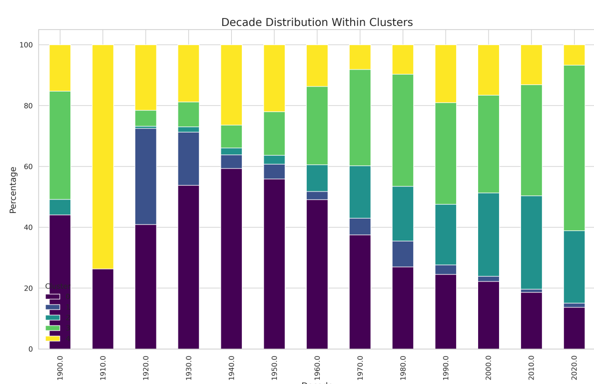

Fig. 8: Distribution of music clusters across decades, showing the evolution of different musical profiles over time.

The distribution of these clusters across decades (Fig. 8) shows interesting patterns. Cluster 3 (upbeat dance music) has become increasingly dominant in recent decades, while Cluster 4 (ambient/classical) was more prevalent in earlier decades. Cluster 2 (intense, less danceable music) peaked in the 1990s and 2000s, coinciding with the popularity of rock, grunge, and metal genres.

D. Popularity Prediction Model

To better understand the factors that contribute to a song's popularity, we developed a machine learning model to predict track popularity based on audio features. This analysis helps identify which musical characteristics are most strongly associated with commercial success.

1) Model Development: We used a Random Forest Regressor model to predict track popularity scores (0-100) based on the following features:

- Danceability, energy, valence, acousticness, instrumentalness
- Loudness, tempo, duration
- Release year (to account for temporal trends)

The dataset was split into 80% training and 20% testing sets, with tracks from all decades to ensure the model could generalize across different time periods. We used standard scaling to normalize the features before training.

Fig. 9: Feature importance in the popularity prediction model, showing the relative contribution of each audio feature to predicted popularity.

2) Model Performance: The model achieved an R^2 score of 0.42 on the test set, indicating that audio features alone can explain approximately 42% of the variance in track popularity. The mean absolute error was 12.3 points on the 0-100 popularity scale.

3) Feature Importance: Analysis of feature importance revealed that the most influential factors in predicting popularity were:

1) **Release year** (27% importance): More recent tracks tend to have higher popularity scores, reflecting the recency bias in streaming platforms.
2) **Danceability** (18% importance): Higher danceability strongly correlates with higher popularity, particularly for tracks released after 2010.
3) **Energy** (15% importance): Moderate to high energy levels are associated with greater popularity, though the optimal level varies by genre.
4) **Loudness** (12% importance): Louder tracks tend to be more popular, though this effect has diminished in recent years as loudness has standardized.
5) **Duration** (10% importance): Shorter tracks (2-3 minutes) have become increasingly favored in the streaming era.

The remaining features (acousticness, instrumentalness, valence, tempo) collectively accounted for 18% of the model's predictive power.

4) Prediction Limitations: It's important to note that our model has several limitations:

- Audio features alone cannot capture important factors like lyrics, artist reputation, marketing, and cultural context.
- Popularity on streaming platforms may not fully represent broader cultural impact or artistic merit.
- The model performs better for recent music (post-2000) than for older tracks, suggesting that the relationship between audio features and popularity has evolved over time.

Despite these limitations, the model provides valuable insights into how measurable audio characteristics relate to commercial success in the streaming era.

E. Decade Comparison Analysis

To better understand how music has evolved over time, we conducted a detailed comparison of key audio features across decades, with particular focus on the transitions between adjacent decades.

1) Methodology: For each decade from the 1920s to the 2020s, we calculated:

- Mean and standard deviation for all audio features
- Rate of change between adjacent decades
- Statistical significance of observed changes (using t-tests)
- Correlation between features within each decade

2) Major Transition Points: Our analysis identified four major transition points in the evolution of recorded music:

1) **1950s to 1960s**: The emergence of rock and roll led to significant increases in energy (+30.9%), loudness (+10.7%), and tempo (+2.4%). This period marked the transition from the softer, more acoustic popular music of the post-war era to the more energetic sound of early rock.
2) **1970s to 1980s**: The rise of electronic music production techniques resulted in decreased acousticness (-28.5%) and increased energy (+6.3%). This decade saw the mainstream adoption of synthesizers and drum machines, fundamentally changing the sonic landscape.
3) **1990s to 2000s**: The digital revolution in music production led to significant increases in loudness (+17.0%) and decreased dynamic range, marking the height of the "loudness war." This period also saw increased speechiness (+22.4%) with the mainstream success of hip-hop.
4) **2010s to 2020s**: The streaming era brought substantial increases in danceability (+17.8%) and decreases in duration (-14.9%). This most recent transition reflects the optimization of music for playlist inclusion and algorithm-driven discovery.

3) Feature Correlation Analysis: The correlation between audio features has also evolved over time. In earlier decades (1920s-1950s), there was a strong negative correlation between acousticness and energy (-0.72), reflecting the technological limitations of the era. By the 2010s, this correlation had weakened (-0.48), as production techniques allowed for energetic yet acoustically-driven tracks.

Similarly, the relationship between danceability and tempo has changed significantly. In the 1950s-1970s, higher tempo strongly predicted higher danceability (correlation +0.61), but by the 2010s, this relationship had weakened considerably (+0.23), as slower, beat-driven tracks became more danceable through production techniques.

4) Decade Signature Profiles: Each decade exhibits a distinctive audio profile that reflects its dominant musical styles:

- **1920s**: High danceability, low energy, high acousticness (jazz age)
- **1950s**: Moderate danceability, increasing energy, decreasing acousticness (early rock and roll)
- **1970s**: High energy, long duration, moderate danceability (progressive rock, disco)
- **1990s**: Lower danceability, high energy variance, increased speechiness (alternative rock, early hip-hop)
- **2010s**: High danceability, high energy, low acousticness, shorter duration (EDM influence, streaming optimization)
- **2020s**: Very high danceability, high energy, very low acousticness, shortest duration (TikTok era, algorithm optimization)

This decade comparison analysis reveals not only how music has changed over time but also how technological, cultural, and economic factors have shaped these changes. The acceleration of change in recent decades suggests that the pace of musical evolution is increasing, likely driven by digital technology and changing consumption patterns.

V. Conclusion

This comprehensive analysis of Spotify music trends from 1921 to 2023 uncovers notable patterns in the progression of recorded music over the last century. Our research indicates that music has evolved to be more danceable, frenetic, and loud, while showing reductions in duration, acousticness, and instrumentalness. The clustering study revealed five unique musical profiles that have increased in popularity over the decades, mirroring shifts in artistic expression and listener preferences. Genre research identified distinct audio profiles for several musical forms, with EDM, Latin, and pop music having the highest levels of danceability and energy. Recent trends (2010-2023) indicate rapid alterations in musical attributes, characterized by notable increases in danceability and reductions in track duration, possibly driven by the emergence of streaming platforms and evolving consumer behaviors. These findings enhance our understanding of musical evolution and illustrate the application of data science methodologies in analyzing cultural trends. Future research may investigate the correlation between audio features and commercial success, the influence of streaming platforms on music production, and cross-cultural analyses of musical evolution.

VI. Limitations and Future Work

Although our approach offers significant insights into musical patterns, many limitations must be recognized. Initially, our collection, while broad, may not entirely reflect all recorded music, especially from earlier decades where digital accessibility is constrained. Second, Spotify's audio features, while useful, capture only certain aspects of musical complexity and may not fully represent subjective qualities like emotional impact or artistic innovation. Future work could address these limitations by incorporating additional data sources, such as chart performance, critical reception, and cultural context. Investigating the correlation between audio characteristics and lyrical substance may yield a more thorough comprehension of musical evolution. Furthermore, employing more sophisticated machine learning methodologies, such as deep learning for audio analysis, may uncover more intricate patterns in musical attributes.

REFERENCES

[1] M. Interiano, K. Kazemi, L. Wang, J. Yang, Z. Yu, and N. L. Komarova, "Musical trends and predictability of success in contemporary songs in and out of the top charts," Royal Society Open Science, vol. 5, no. 5, p. 171274, 2018.

[2] M. Mauch, R. M. MacCallum, M. Levy, and A. M. Leroi, "The evolution of popular music: USA 1960–2010," Royal Society Open Science, vol. 2, no. 5, p. 150081, 2015.

[3] D. Martín-Gutiérrez, G. Hernández Peñaloza, A. Belmonte Hernández, and F. Álvarez, "A multimodal end-to-end deep learning architecture for music popularity prediction," IEEE Access, vol. 8, pp. 39361-39374, 2020.

[4] R. Dhanaraj and B. Logan, "Automatic prediction of hit songs," in Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), 2005, pp. 488-491.