Analyzing Musical Preferences and Trends on Spotify: A Data-Driven Exploration of Genre, Attributes, and Popularity

Abstract

This paper analyzes the factors influencing track popularity on Spotify by examining the relationships between audio features and music genre. Moreover, the variations in audio characteristics between the 1900s and 2000s are analyzed. A dataset of over 28,000 unique tracks is used, key musical attributes—including danceability, energy, valence, tempo, loudness, speechiness, acousticness, instrumentalness, and liveness are used for this study. To make a quantifiable association between these features and track popularity, we compute correlation coefficients and fit linear regression models, from which we derive slopes, R-squared values, and p-values. This analysis reveals audio features play a big role in track popularity, several features (e.g., loudness, tempo, danceability, and energy) exhibit significant differences between tracks from the 1900s and the 2000s and a one-way ANOVA indicates that genre is also a significant predictor of track popularity, with genres like Pop and Rap outperforming others such as EDM. The implications of this research extend to music producers, streaming platforms, and academic researchers interested in the evolving trends of digital music consumption. Limitations regarding data scope and measurement constraints are discussed, and directions for future research are proposed.

Keywords: *Spotify, Audio Features, Track Popularity, Genre Analysis, Music Evolution, Digital Music Consumption, Music Production, Regression Analysis, ANOVA, Hit Song Science.*

1. Introduction

Music has always been a big part of human culture and history, with the rise of digital streaming platforms like Spotify the way we experience music has changed a lot. Digital platforms like Spotify not only provide streaming services to consumers but also provide large datasets to study music industrial trends, listener preferences, and evolution of music throughout the time. Examining elements like danceability, energy, and genre helps scientists find trends in the creation, consumption, and evolution of music. Research suggests that certain musical attributes significantly influence song popularity and chart performance. High danceability and low instrumentalness increase the popularity of songs (Al-Beitawi et al., 2020). Less than 25% of a song's popularity can be ascribed to musical elements including danceability, energy, instrumentalism, acousticness, duration, speechiness, and valence, implying that the secret to producing a hit song remains essentially unknown. (Preet D. Singh, 2021)

Recent studies have explored the factors influencing digital music popularity using machine learning techniques. Zhang et al. (2024) examining a sizable Spotify dataset (2024), researchers found that song popularity is much influenced by genre and audio characteristics; effects vary across genres. . Their random forest model outperformed other predictive models in forecasting popularity. Similarly, Sebastian et al. (2024) used a huge dataset which included data from multiple decades and they used machine learning techniques to predict the popularity of songs. They found genre to be the primary influencer of popularity, with EDM being particularly significant. The most effective model for predicting popularity is found to be Random Forest by both (Zhang et al., 2024; Sebastian et al., 2024). However, Sebastian et al. (2024) noted that accurately predicting song success remains challenging due to complex interactions between genre and other factors. These findings offer valuable insights for music industry professionals in understanding and capitalizing on popular songs in the digital streaming era.

Research on popular music trends reveals significant changes over time. Studies have found that contemporary Western popular music has become more homogeneous in timbre and louder over the past 50 years, while pitch transitions have become more restricted (Serrà et al., 2012). Additionally, American popular music has evolved to sound sadder, with an increase in minor mode usage and slower tempos (Schellenberg & von Scheve, 2012). These changes in musical features can be quantitatively analyzed using objective parameters extracted from audio recordings (Dorochowicz & Kostek, 2018). Temporal embedding models have been developed to explore how listening preferences change over time, allowing for visualization and quantification of musicological trends using large-scale data from platforms like Last.fm (Moore et al., 2013). These studies collectively demonstrate that popular music undergoes measurable changes in various aspects, including emotional cues, structural characteristics, and listener preferences, providing insights into the evolution of musical tastes and styles.

This paper explores these audio features to find their relation with the popularity of a track by computing correlation coefficient with popularity score of a track to measure the association, and then calculating the R-squared value as well as p-value to determine how much variability in popularity is explained by each feature and to assess the statistical significance of these relationships.Music has always been a big part of human culture and history, with the rise of digital streaming platforms like Spotify the way we experience music has change a lot. Digital platforms like spotify not only provide streaming services to consumers but also provide large datasets to study music industrial trends, listener preferences, and evolution of music throughout the time. By analyzing attributes like danceability, energy, and genre, researchers can uncover patterns in how music is created, consumed, and evolved. Research suggests that certain musical attributes significantly influence song popularity and chart performance. High danceability and low instrumentalness increase the popularity of songs (Al-Beitawi et al., 2020). Musical attributes such as danceability, energy, instrumentalness, acousticness, duration, speechiness, and valence account for less than 25% of a song's popularity, suggesting that the key to creating a hit song remains largely undiscovered. (Preet D. Singh, 2021)

Recent studies have explored the factors influencing digital music popularity using machine learning techniques. Zhang et al. (2024) analyzed a large Spotify dataset, finding that genre and audio features significantly impact song popularity, with effects varying across genres. Their random forest model outperformed other predictive models in forecasting popularity. Similarly, Sebastian et al. (2024) investigated various machine learning approaches to predict song popularity, using a dataset spanning multiple decades. They found genre to be the primary influencer of popularity, with EDM being particularly significant. Both studies identified Random Forest as the most effective model for predicting popularity (Zhang et al., 2024; Sebastian et al., 2024). However, Sebastian et al. (2024) noted that accurately predicting song success remains challenging due to complex interactions between genre and other factors. These findings offer valuable insights for music industry professionals in understanding and capitalizing on popular songs in the digital streaming era.

Research on popular music trends reveals significant changes over time. Studies have found that contemporary Western popular music has become more homogeneous in timbre and louder over the past 50 years, while pitch transitions have become more restricted (Serrà et al., 2012). Additionally, American popular music has evolved to sound sadder, with an increase in minor mode usage and slower tempos (Schellenberg & von Scheve, 2012). These changes in musical features can be quantitatively analyzed using objective parameters extracted from audio recordings (Dorochowicz & Kostek, 2018). Temporal embedding models have been developed to explore how listening preferences change over time, allowing for visualization and quantification of musicological trends using large-scale data from platforms like Last.fm (Moore et al., 2013). These studies collectively demonstrate that popular music undergoes measurable changes in various aspects, including emotional cues, structural characteristics, and listener preferences, providing insights into the evolution of musical tastes and styles.

This paper explores these audio features to find their relation with the popularity of a track by computing correlation coefficient with popularity score of a track to measure the association, and then calculating the R-squared value as well as p-value to determine how much variability in popularity is explained by each feature and to assess the statistical significance of these relationships.

To determine whether certain musical styles are inherently more popular than others, we conduct a genre analysis by comparing the mean track popularity across genres using ANOVA. Furthermore, we conduct a temporal analysis by comparing tracks released in the 1900s versus those in the 2000s to look at how music has changed over time.

Following questions have been addressed in this study:

1. How do various audio features relate to the popularity of a music track?

2. Whether the genre of a music track has a significant impact on its popularity?
3. How have audio features evolved over time, between the 1900s and the 2000s?

The remainder of this paper is organized as follows: In section 2 the overall framework of this analysis has been outlined. The conceptual background and review of relevant literature and the hypotheses derived from theory and prior research has been provided in section 3. Section 4 describes the empirical study, including data description, methodology, statistical tests and results. Section 5 discusses the findings, implications, and limitations along with a conclusion.

2. Music Analysis

2.1 Overview

A music track consists of various audio features which are computed using signal processing algorithms. For example, danceability shows how suitable a track is for dancing, energy measures the intensity of a track, valence shows how happy or cheerful a track sounds which listening, tempo shows the speed of a track which is measured in beats per minute, liveness detects the presence of an audience in the recording in which higher liveness values represent an increased probability that the track was performed live, instrumentalness shows whether a track contains no vocals in which rap or spoken word tracks are clearly vocal, acousticness measures whether a track is acoustic i.e. without electrical amplification, loudness represents overall loudness of a track in decibels, duration shows the length of a track in milliseconds, and genre represents a specific type or category a music track falls in. These features give us the structure and production style of a music track.

2.2 Metrics to evaluate audio characteristics

We use statistical metrics to assess the role of each audio feature in determining a track's popularity. At first, we compute the correlation coefficient between each audio feature and the popularity score to determine the strength and direction of their relationship in which high positive or negative correlation indicates that the feature plays a significant role in determining the popularity of a music track.

We fit linear regression models whereby track popularity is the dependent variable and each audio feature is an independent variable in order to further estimate the correlation of each audio feature with track popularity. The slope of the regression line indicates how much track popularity changes with a unit change in the feature, while the R-squared value reveals the proportion of variability in popularity that is explained by that one feature. The p-value of the slope helps to determine the statistical significance of the relationship between features and track popularity. This paper explores these audio features by computing the correlation coefficient of each feature with the popularity score, and then calculating the R-squared and p-values to determine the variability in popularity explained by each feature and assess its significance. This systematic approach enables us to identify which features most strongly predict track popularity, thereby shedding light on the underlying factors that contribute to a track's commercial success. Similar methodologies have been adopted in recent studies, such as Sciandra and Spera (2022) and Yee and Raheem (2022), which highlight the importance of audio features in predicting music popularity.

2.3 Aggregating audio feature metrics

The genre of a music track has a significant impact on its popularity in online social networks (REN, Jing; SHEN, Jialie; and KAUFFMAN, Robert John. What makes a music track popular in online social networks?. (2016)). To understand broader trends in music production, we use ANOVA to average popularity scores for different genres so we can determine whether the genre of a music track affects its popularity. We then characterize the overall evolution of musical production i.e. the change in audio features in the 2000s in comparison to the 1900s.

3. Conceptual Background

3.1 Literature Review

One of the earliest studies in "hit song science, was (Salganik et al. 2006). It investigated how random changes in cultural markets might affect music popularity. Later research including Mauch et al. (2015) has produced extensive empirical data connecting particular audio characteristics—including tempo, loudness, and danceability—to track success.

Recent works (Müller et al., 2017) have included machine learning approaches to predict popularity. (Tzanetakis and Cook, 2002) showed that quantitative analysis of audio signals could consistently classify music genres. These studies suggest that intrinsic audio features can serve as strong predictors of commercial success of a music track.

Temporal analyses of music have also garnered attention. Several studies have documented trends such as increasing loudness and energy in modern music, attributing these shifts to advances in digital production (Hamoudi et al., 2015). Meanwhile, research into the evolution of musical sentiment has highlighted shifts in valence and emotional tone over time (Mauch et al., 2015).

Research that has been conducted in recent times is more refined. For example, Sciandra and Spera (2022) proposed a Beta Generalized Linear Mixed Model (Beta GLMM) to analyze Spotify data. It effectively addressed issues like the clustering of songs within albums and the bounded nature of the popularity index. This model-based approach helps identify which audio features—such as loudness, duration, and harmonic simplicity—are most influential in predicting a song's popularity.

Despite these advances, comprehensive analyses that integrate multiple statistical approaches to relate audio features to popularity, compare different eras and genres are hard to find. This study provides a comprehensive view of how audio features and popularity are related along with the evolution of audio attributes over time. This study aims to fill that gap by simultaneously examining correlations, regression models, ANOVA, and t-tests.

3.2 Hypotheses

These hypothesis were developed after carefully considering the findings from previous research and conceptual background:

H1 Audio Feature Hypotheses

There is a significant relationship between each audio feature and track popularity.

H2 Genre Hypotheses

The genre of a music track significantly influences its popularity.

H3 Temporal (Era) Hypotheses

There are significant differences in the audio features of tracks released in the 1900s and those released in the 2000s.

4. Empirical Study

We explored the relationship between popularity of a track and musical attributes, the influence of genre on popularity, and the way musical features have evolved over time. This helps to analyze what makes music popular. This study is focused on finding the key factors that make the track popular on streaming platform - Spotify.

The detailed description of our dataset, method used for data preparation , and statistical analysis techniques used to extract valuable insights from the data are mentioned below.

4.1 Data

The initial dataset used in this study consists of 32,833 rows and 32 columns retrieved from Kaggle. The dataset contains the metrical and nominal variables. Metrical variables are track popularity with value range of 1 to 100 , music attributes. Genre is a nominal variable, indicating the type of music, and playlist information is also included to analyze its impact on song success.

4.2 Data Preparation

The raw dataset initially contained 32,833 records. After removing duplicate tracks based on the unique track_id, the number of unique records was reduced to 28,356, eliminating redundant data to ensure each track was represented only once.

Data cleaning was an important step to make sure the analysis was accurate and consistent. At first, we removed all the null values using the na.omit function, to maintain the completeness. Then we removed the unnecessary columns like album_id, artist_name, album_name, sub_genre because they won't contribute to our research.

The dataset was further prepared for each hypothesis as follows:

Hypothesis 1: For the analysis of the relationship between various audio features and track popularity, the popularity and musical attribute columns were already in a numeric format. As a result, no additional data preparation was required for this hypothesis.

Hypothesis 2: The track_popularity column was converted to numeric using as.numeric(), and the playlist_genre column was converted to a factor using as.factor() which is necessary to perform ANOVA as it requires a categorical independent variable (genre) and a numeric dependent variable (track popularity).

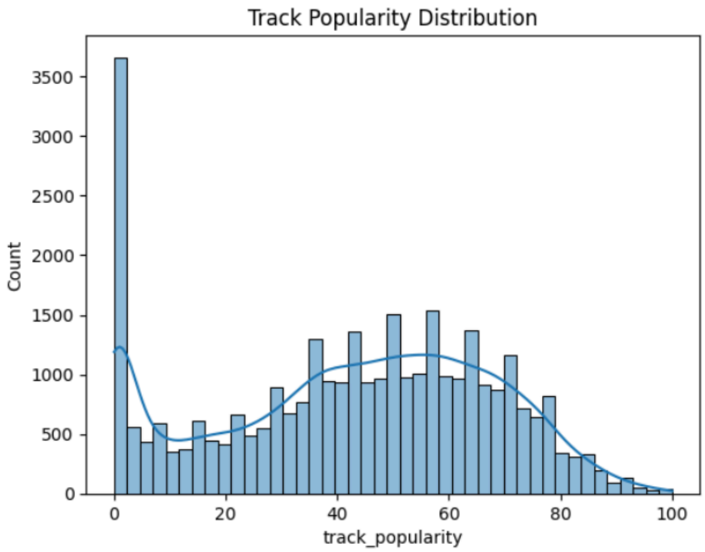
Hypothesis 3: For this analysis, the track_album_release_date was formatted to ensure consistency, and if only the year was provided, the date was adjusted to YYYY-01-01. A new era variable was created to differentiate between tracks released before and after the year 2000. This era variable was then converted into a factor in order to analyze audio features by era (1900s vs 2000s).

The data preparation procedures guaranteed that the dataset was clean, organized, and ready for hypothesis testing.

Table 1: Dataset Description

Feature	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Track Popularity	0.00	21.00	42.00	39.34	58.00	100.00
Danceability	0.00	0.5610	0.6700	0.6534	0.7600	0.9830
Energy	0.000175	0.5790	0.7220	0.6984	0.8430	1.0000
Key	0.00	2.000	6.000	5.367	9.000	11.000
Loudness	-46.448	-8.310	-6.261	-6.818	-4.709	1.275
Mode	0.0000	0.0000	1.0000	0.5655	1.0000	1.000
Speechiness	0.00	0.0410	0.0626	0.1079	0.1330	0.9180
Acousticness	0.00	0.0143	0.0797	0.1772	0.2600	0.9940
Instrumentalness	0.00	0.00	0.0000207	0.0911	0.0065725	0.99400
Liveness	0.00	0.0926	0.1270	0.1910	0.2490	0.9960
Valence	0.00	0.3290	0.5120	0.5104	0.6950	0.9910
Tempo	0.00	99.97	121.99	120.96	134.00	239.44
Duration	4000	187741	216933	226575	254975	517810

Figure 1: Track Popularity Distribution



4.3 Exploratory Data Analysis

In the exploratory data analysis, a heatmap was created to explore the relationships between different musical attributes and track popularity. The findings reported several important relationships: Popularity was found to be negatively correlated with energy, instrumentality, and duration. This suggests that music with more instrumental content, higher energy, or longer duration are often less well-liked. On the other hand, valence and danceability were positively correlated, suggesting that lively tunes tend to be more danceable. Additionally, Energy was negatively correlated with acousticness, but positively correlated with loudness. It indicates that higher energy tracks are louder but less acoustic. Lastly, negatively correlate with loudness, acousticness was found, showing that more acoustic tracks tend to be quieter. These insights provide a better understanding of how different musical features influence a track's popularity.

Table 2 : Correlation Matrix

	track_popularity	valence	energy	mode	key	acousticness	danceability	instrumentalness	loudness	liveness	tempo	speechiness	duration_ms
track_popularity	1.000	0.023	-0.104	0.016	-0.008	0.092	0.047	-0.125	0.037	-0.053	0.004	0.005	-0.140
valence	0.023	1.000	0.150	-0.003	0.022	-0.019	0.334	-0.174	0.050	-0.020	-0.025	0.065	-0.033
energy	-0.104	0.150	1.000	-0.004	0.013	-0.546	-0.081	0.024	0.682	0.164	0.152	-0.029	0.018
mode	0.016	-0.003	-0.004	1.000	-0.176	0.007	-0.055	-0.006	-0.018	-0.000	0.017	-0.060	0.013
key	-0.008	0.022	0.013	-0.176	1.000	0.004	0.007	0.007	-0.001	0.002	-0.010	0.023	0.019
acousticness	0.092	-0.019	-0.546	0.007	0.004	1.000	-0.029	-0.003	-0.372	-0.075	-0.114	0.025	-0.094
danceability	0.047	0.334	-0.081	-0.055	0.007	-0.029	1.000	-0.002	0.015	-0.127	-0.185	0.184	-0.088
instrumentalness	-0.125	-0.174	0.024	-0.006	0.007	-0.003	-0.002	1.000	-0.154	-0.008	0.021	-0.108	0.059
loudness	0.037	0.050	0.682	-0.018	-0.001	-0.372	0.015	-0.154	1.000	0.082	0.097	0.013	-0.105
liveness	-0.053	-0.020	0.164	-0.000	0.002	-0.075	-0.127	-0.008	0.082	1.000	0.022	0.059	0.008
tempo	0.004	-0.025	0.152	0.017	-0.010	-0.114	-0.185	0.021	0.097	0.022	1.000	0.033	-0.002
speechiness	0.005	0.065	-0.029	-0.060	0.023	0.025	0.184	-0.108	0.013	0.059	0.033	1.000	-0.098
duration_ms	-0.140	-0.033	0.018	0.013	0.019	-0.094	-0.088	0.059	-0.105	0.008	-0.002	-0.098	1.000

4.4 Results

The relationship between Spotify track popularity, audio features, genres, and temporal trends are thoroughly investigated in this study. Building on our hypotheses—(1) audio features correlate with popularity, (2) genres differ in popularity, and (3) modern tracks exhibit distinct audio characteristics—we analyzed 28,352 tracks using correlation analysis, linear regression, ANOVA, and t-tests. The results are organized into three subsections: first, we present findings on audio features and their association with popularity; second, we compare popularity across genres; and third, we examine how audio attributes have evolved between the 1900s and 2000s.

4.4.1 Relation between track popularity and musical attributes

For our first hypothesis test, we analyzed various audio attributes to find how they relate to track popularity by calculating the correlation, regression slope, R-squared, and p-values for each attribute.

Table 3: Music Attribute Regression Analysis

Attribute	Correlation	R_squared	Slope	P_value
danceability	0.046574393	2.16917e-03	7.571038e	4.290636e-15
energy	-0.103510773	1.071448e-02	-1.336805e+1	2.206449e-68
valence	0.022594291	5.105020e-04	2.284985e+00	1.419438e-04
tempo	0.004321794	1.867790e-05	3.799889e-03	4.668112e-01
loudness	0.037336843	1.394040e-03	2.914151e-01	3.201098e-10
speechiness	0.005439570	2.958892e-05	1.257126e+00	3.597278e-01
acousticness	0.091624759	8.395096e-03	9.745582e+00	6.492961e-54

instrumentalness	-0.124546651	1.551187e-02	-1.269202e+01	2.177638e-98
liveness	-0.052752799	2.782858e-03	-8.019949e+00	6.202812e-19
duration_ms	-0.139675812	1.950933e-02	-5.419393e-05	1.741323e-13

Danceability:
The correlation with popularity is small but positive (0.0466), and the regression shows that for a one-unit increase in danceability, track popularity increases by about 7.57 points. Although the R-squared value is very low (0.00217, or 0.22% of the variance explained), the relationship is statistically significant ($p \approx 4.29\times10^{-15}$). This suggests that while danceability is linked to popularity, it alone explains only a tiny fraction of the overall variability.

Energy:
Energy has a small negative correlation (-0.1035) with track popularity, and the regression slope indicates that a one-unit increase in energy is associated with a decrease of about 13.37 points in popularity. Despite the low R-squared (1.07% of the variance explained), the relationship is highly significant ($p \approx 2.21\times10^{-68}$). This result implies that higher energy levels in a track might be linked with lower popularity, though the effect size is modest.

Valence:
The correlation for valence is very weak (0.0226) and positive, with a regression slope of approximately 2.28, meaning that as valence increases by one unit, popularity rises by 2.28 points. The R-squared is only 0.000511 (0.05% of the variance), but the relationship is statistically significant ($p \approx 1.42\times10^{-4}$). This indicates a slight positive association between the mood or musical positiveness of a track and its popularity.

Tempo:
Tempo shows an extremely low correlation (0.00432) with track popularity, and the regression analysis does not find a significant relationship ($p \approx 0.467$). The R-squared is

nearly zero, suggesting that tempo does not play a meaningful role in predicting popularity.

Loudness:

Loudness has a modest positive correlation (0.0373) with track popularity, with the regression showing that for each one-unit increase in loudness, popularity increases by about 0.29 points. The R-squared is 0.001394, and the relationship is statistically significant ($p \approx 3.20\times10^{-10}$), indicating that louder tracks (or tracks with less negative dB values) are slightly more popular.

Speechiness:

Speechiness has a very small correlation (0.00544) with popularity, and its relationship is not statistically significant ($p \approx 0.360$). This suggests that the presence of spoken words in a track does not have a clear impact on popularity.

Acousticness:

Acousticness shows a positive correlation (0.0916) with popularity, and the regression slope of about 9.75 indicates that a one-unit increase in acousticness is associated with an increase of 9.75 points in popularity. The R-squared is 0.008395, and the relationship is highly significant ($p \approx 6.49\times10^{-54}$), suggesting that more acoustic tracks tend to be somewhat more popular.

Instrumentalness:

Instrumentalness has a stronger negative correlation (-0.1245) with track popularity. The regression analysis reveals that a one-unit increase in instrumentalness is associated with a decrease of about 12.69 points in popularity, with an R-squared of 0.015512 (1.55% of the variance explained) and a very significant p-value ($\approx 2.18\times10^{-98}$). This indicates that tracks with higher instrumentalness (i.e., those with less vocal content) are less popular.

Liveness:

Liveness has a small negative correlation (-0.05275) with popularity, and the regression slope indicates that an increase in liveness by one unit corresponds to a decrease in popularity of about 8.02 points. Although the R-squared is low (0.002783), the relationship is statistically significant ($p \approx 6.20\times10^{-19}$), suggesting that more "live" sounding tracks tend to be slightly less popular.

Duration (duration_ms):
Duration shows a negative correlation (-0.13968) with popularity. The regression slope is very small (-0.000054 per millisecond), which translates to a decrease of about 0.054 points in popularity for every additional second of track duration. The R-squared is 0.019509 (1.95% of the variance explained), and the relationship is highly significant ($p \approx 1.74\times10^{-123}$), indicating that longer tracks are slightly less popular.

4.4.2 Genre influence on popularity trends

We conducted a one-way Analysis of Variance (ANOVA) to determine whether the average track popularity differs significantly across six music genres in our dataset. In other words, we tested whether any genre's mean popularity score is distinct from the others.

Table 4: ANOVA Result Playlist Genre

	Degree of Freedom	Sum of Squares	Mean of Squares	F Value Pr(>F)
Playlist genre	5	692345	138469	257.7 $< 2\times10^{-16}$
Residuals	28346	15231380	537	

Degrees of Freedom for Playlist Genre = Number of genre groups - 1
Degrees of Freedom for Residuals =Total tracks−Number of genre groups
Sum of Squares (SS):
Between-group variation (due to genre differences):
SSgenre= 692345

Within-group variation (due to variation within genres):
SSresidual= 15231380

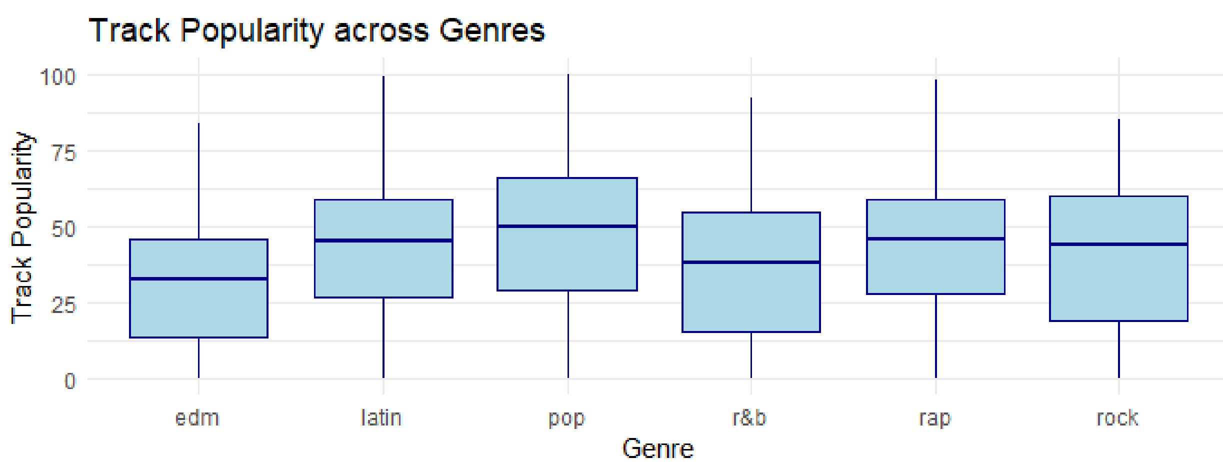
Mean Square for Genre:
MSgenre= SSgenre / Dfgenre = 138469

Mean Square for Residuals:
MSresidual= SSresidual / Dfresidual = 537

FValue = MSgenre / MSresidual = 257.7

Additionally, we used the boxplot to visually compare the distribution of track popularity across the different genres.

Figure 2: Box Plot of Track Popularity across Genres



From the boxplot, you can visually confirm the patterns revealed by the ANOVA results. Each box represents the distribution of track popularity within a genre, showing the median (the thick horizontal line), the interquartile range (the box itself), and any potential outliers (dots beyond the whiskers). Notice that some genres—such as EDM—tend to cluster at lower popularity scores, while others—like Pop—often reach higher values. Latin, Rap, and Rock lie somewhere in between, with R&B showing a distribution that partially overlaps but is generally below Pop. These visual differences align with the ANOVA findings, where the large F-value and very low p-value (< 2e-16) indicate that genre plays a significant role in explaining variations in track popularity. In other words, the patterns you see in the boxplot reflect the statistical conclusion that not all genre means are the same—some genres do indeed have distinctly higher or lower popularity scores.

4.4.3 Evolution of musical features over time

In our analysis comparing audio features between the 1900s and the 2000s, we used Welch's two-sample t-tests. Welch's t-test was chosen because it does not assume that the two groups have equal variances or the same sample sizes.

Table 5: Feature Difference on 1900s and 2000s

Feature	Mean (1900s)	Mean (2000s)	Difference	p-value	Significance
Loudness	-9.0564	-6.4397	2.6167 (louder)	< 2.2e-16	Highly significant (tracks are louder)
Tempo (BPM)	118.94	121.29	2.35 (faster)	5.616e-07	Highly significant (tracks have faster tempo)
Danceability	0.6236	0.6585	0.0349 (more dancey)	< 2.2e-16	Highly significant
Valence	0.6090	0.4937	-0.1153 (less happy)	< 2.2e-16	Highly significant
Instrumentalness	0.0456	0.0988	0.0532 (more instrumental)	< 2.2e-16	Highly significant
Liveness	0.1938	0.1904	-0.0034 (no real change)	0.2173	Not significant
Energy	0.6657	0.7039	0.0382 (higher energy)	< 2.2e-16	Highly significant
Acousticness	0.1719	0.1779	0.0060 (no real change)	0.0927	Not significant
Speechiness	0.0883	0.1113	0.0230 (more speechy)	< 2.2e-16	Highly significant

From the above table, we found that most audio features exhibited p-values below 2.2×10^{-16} which indicates there are significant differences between the two eras. Modern tracks from the 2000s are characterized by higher loudness, tempo, danceability, and energy, suggesting a shift toward more intense, faster-paced, and club-oriented production. In contrast, tracks from the 1900s display higher valence, which implies that older music tends to sound “happier.” Additionally, instrumentalness and speechiness are greater in the 2000s, which might be due to the rise of instrumental genres and the increasing influence of hip-hop and rap elements in recent times.

5. Summary, Implications & Limitations

5.1 Summary

This study provides a comprehensive analysis of Spotify track data by examining how audio features relate to track popularity, the role of genre, and the evolution of music production from the 1900s to the 2000s.The majority of auditory attributes and popularity have strong correlations, according to our statistical testing. Furthermore, the genre analysis confirms significant differences in popularity across musical styles, while the temporal analysis shows that modern music is typically defined by increased loudness, tempo, danceability, energy, instrumentality, and speechiness, along with a decrease in valence, while the genre analysis validates notable variations in popularity across musical styles.

5.2 Implications

For Music Producers and Recording Engineers:
The results show a distinct trend in the 2000s toward songs that were louder, more upbeat, and more danceable. Advances in digital production methods, such as equalization and dynamic range compression, may be the cause of this change. (Sexton, 2008; Irava & Georgescu, 2024). Machine learning models can accurately predict music popularity on streaming platforms, allowing producers to tailor music features to consumer preferences.(Gao, 2021)

For Streaming Services and Record Labels:

By leveraging insights from user data, businesses can better understand user behavior, personalize experiences, and optimize interactions. Techniques such as A/B testing, gamification, and social proof can be employed to drive engagement and foster a sense of community. (Mayokun Daniel Adegbola 2024)

For Academic Research:

The usefulness of quantitative techniques in music analysis is illustrated by our integrative approach, which combines regression, ANOVA, correlation analysis, and t-tests. This study adds to the expanding body of research on digital music trends and music information retrieval (MIR), and it emphasizes the necessity of multifaceted analyses that take into account both more general cultural changes and inherent musical qualities.

5.3 Limitations

Data Limitations:

The study is based solely on Spotify data, which may not fully represent the global diversity of music, particularly for older tracks. The dataset's inherent biases—such as overrepresentation of certain genres or regions—could affect the generalizability of the findings.(Zeighami & Shahabi, 2024)

Measurement Constraints:

The audio characteristics are generated using specialized algorithms, which, though effective, may not capture every subtle detail of musical expression. Moreover, the popularity scores on Spotify are impacted by external elements, such as marketing and playlist placements, which are beyond the scope of this study.

Temporal Categorization:

The binary division into "1900s" and "2000s" is a simplification of a continuous evolution. While useful for detecting broad trends, this categorization may mask subtler shifts that occur within each era. Future studies could adopt a more granular temporal analysis, such as by decade.

Causality:

The analyses conducted are correlational. Although significant associations are identified, we do not provide evidence for causality—i.e., we cannot definitively state that variations in track popularity are caused by changes in audio characteristics.

5.4 Discussion

The results of this study highlight important shifts in music production over the past century. Modern tracks are notably louder, faster, and more energetic than those produced in the 1900s, which is likely a reflection of technological advances in recording and production. Moreover, the increased danceability and speechiness suggest that listener tastes are changing and that genres like hip-hop and EDM are becoming more and more popular.

The observed decrease in valence for modern tracks is particularly interesting and may suggest that while modern production techniques emphasize energy and rhythm, they may come at the expense of a perceived positive emotional tone. This might affect how people listen to music in various social settings (Ziv, 2004).

Genre analysis further supports the notion that stylistic differences are critical determinants of popularity. The significant differences in popularity across genres indicate that certain styles resonate more with contemporary audiences. Both practical applications in music marketing and academic models of cultural consumption can benefit from this discovery.

Overall, the study highlights the importance of using diverse analytical techniques to capture a variety of analytical methods in order to fully convey the complex character of music. The combination of regression, t-tests, ANOVA, and visualization not only validates our hypotheses but also provides a comprehensive picture of the evolving landscape of music production and consumption.

Future research should look into adding other variables, such as social media metrics, listener demographics, and cultural factors, to gain deeper insights into the drivers of these trends. Longitudinal studies tracking changes in audio characteristics over time, along with experimental designs that isolate the effects of production techniques on listener behavior, would further enhance our understanding of this complex phenomenon.

References

Yee, Yap & Raheem, Mafas. (2022). Predicting Music Popularity Using Spotify and YouTube Features. Indian Journal Of Science And Technology. 15. 1786-1799. 10.17485/IJST/v15i36.2332.

Jing Ren, Jialie Shen, and Robert J. Kauffman. 2016. What Makes a Music Track Popular in Online Social Networks? In Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16 Companion). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 95–96. https://doi.org/10.1145/2872518.2889402

Serra`, J., Corral, A´., Bogun˜a´, M., Haro, M. & Arcos, J.L. Measuring the Evolution of Contemporary Western Popular Music. Sci. Rep. 2, 521; DOI:10.1038/ srep00521 (2012).

Moore, J.L., Chen, S., Turnbull, D., & Joachims, T. (2013). Taste Over Time: The Temporal Dynamics of User Preferences. *International Society for Music Information Retrieval Conference*.

Zhang, J., Yu, S., Liu, R., Xie, G., & Zurawicki, L. (2024). Unveiling the melodic matrix: exploring genre-and-audio dynamics in the digital music popularity using machine learning techniques. *Marketing Intelligence & Planning*.

Schellenberg, E.G., & von Scheve, C. (2012). Emotional Cues in American Popular Music: Five Decades of the Top 40. *Psychology of Aesthetics, Creativity, and the Arts, 6*, 196-203.

Sebastian, N., Jung, & Mayer, F. (2024). Beyond Beats: A Recipe to Song Popularity? A machine learning approach. *ArXiv, abs/2403.12079*.

Ioannis Karydis, Aggelos Gkiokas, Vassilis Katsouros. Musical Track Popularity Mining Dataset. 12th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Sep 2016, Thessaloniki, Greece. pp.562-572, ff10.1007/978-3-319-44944-9_50ff. Ffhal-01557622

Zayd Al-Beitawi, Mohammad Salehan, Sonya Zhang (2020) Cluster Analysis of Musical Attributes for Top Trending Songs . Hawaii International Conference on System Sciences·

Singh, P.D. (2021). How to make a hit song: Invest in a good studio. *SSRN Electronic Journal*.

Salganik, M.J., Dodds, P.S. and Watts, D.J. (2006) ‘Experimental study of inequality and unpredictability in an artificial cultural market’, Science, 311(5762), pp. 854–856. doi:10.1126/science.1121066.

Léveillé Gauvin, Hubert. (2016). Changing Use of Seventh Chords: A Replication of Mauch et al. (2015). Empirical Musicology Review. 11. 103. 10.18061/emr.v11i1.4949.

G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," in IEEE Transactions on Speech and Audio Processing, vol. 10, no. 5, pp. 293-302, July 2002, doi: 10.1109/TSA.2002.800560.

Martiny, Nadège. (2018). Mueller etal 2017.

Hamoudi, A., Murray, D. W., Sorensen, L., & Fontaine, A. (2015). Self-regulation and toxic stress: A review of ecological, biological, and developmental studies of self-regulation and stress (OPRE Report 2015-30). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Mariangela Sciandra and Irene Carola Spera, (2022), A model-based approach to Spotify data analysis: a Beta GLMM, Journal of Applied Statistics, 49, (1), 214-229

Sexton, J. (2008). Digital music: production, distribution and consumption.

Chioma Susan Nwaimo, Ayodeji Enoch Adegbola and Mayokun Daniel Adegbola (2024) ‘Data-driven strategies for enhancing user engagement in digital platforms’, *International Journal of Management & Entrepreneurship Research*, 6(6), pp. 1854–1868. doi:10.51594/ijmer.v6i6.1170.

Gao, A. (2021). Catching the Earworm: Understanding Streaming Music Popularity Using Machine Learning Models. *E3S Web of Conferences*.

Zeighami, S., & Shahabi, C. (2024). BiasBuster: a Neural Approach for Accurate Estimation of Population Statistics using Biased Location Data. *2024 25th IEEE International Conference on Mobile Data Management (MDM)*, 1-10.

Ziv, N. (2004). THE RELATIONSHIP BETWEEN SOCIAL CIRCUMSTANCES AND RECOUNTED MUSICAL EXPERIENCES.

Appendix

```
# Load necessary libraries

library(dplyr)

library(ggplot2)

library(tidyr)

library(reshape2)

# Read the dataset

spotify <- read.csv("spotify_songs.csv", stringsAsFactors = FALSE)

# View the dataset

View(spotify)

print(spotify)

# Check for duplicates in track_id

cat("Number of unique track IDs:", length(unique(spotify$track_id)), "\n")
```

```
cat("Total number of track IDs:", length(spotify$track_id), "\n")

# Remove duplicate tracks

spotify_no_duplicates <- spotify %>%

  distinct(track_id, .keep_all = TRUE)

# Remove missing values and drop the track_id column

spotify_cleaned <- spotify_no_duplicates %>%

  na.omit() %>%

  select(-track_id)

# Define the music attributes to analyze

attributes <- c("danceability", "energy", "valence", "tempo", "loudness",

                "speechiness", "acousticness", "instrumentalness", "liveness", "duration_ms")

# Define columns to convert to numeric (track_popularity plus your attributes)

cols_to_convert <- c("track_popularity", attributes)

spotify_cleaned <- spotify_cleaned %>%

  mutate(across(all_of(cols_to_convert), ~ as.numeric(as.character(.))))

# Create an empty data frame to store analysis results in tabular format

analysis_results <- data.frame(

  Attribute = character(),

  Correlation = numeric(),

  R_squared = numeric(),

  Slope = numeric(),

  P_value = numeric(),

  stringsAsFactors = FALSE

)
```

```
# Loop over each attribute to perform the analysis and collect results
for(attr in attributes) {
  # Calculate the correlation coefficient between track_popularity and the attribute
     corr_coeff <- cor(spotify_cleaned[[attr]], spotify_cleaned$track_popularity, use = "complete.obs")
  # Create a formula for the linear regression (track_popularity ~ attribute)
  formula_str <- paste("track_popularity ~", attr)
  # Fit the linear model
  model <- lm(as.formula(formula_str), data = spotify_cleaned)
  mod_summary <- summary(model)
  # Extract the R-squared, slope (coefficient estimate), and p-value for the attribute
  r_sq <- mod_summary$r.squared
  slope <- mod_summary$coefficients[2, "Estimate"]
  p_val <- mod_summary$coefficients[2, "Pr(>|t|)"]
  # Append the results to the analysis_results data frame
  analysis_results <- rbind(analysis_results, data.frame(
    Attribute = attr,
    Correlation = corr_coeff,
    R_squared = r_sq,
    Slope = slope,
    P_value = p_val,
    stringsAsFactors = FALSE
  ))
  # Generate a scatter plot with a regression line for visualization
```

```
  p <- ggplot(spotify_cleaned, aes_string(x = attr, y = "track_popularity")) +
    geom_point(alpha = 0.3) +              # Plot points with some transparency
    geom_smooth(method = "lm", color = "red") +# Add a linear regression line
    ggtitle(paste("Track Popularity vs", attr)) +
    theme_minimal()
  # Display the plot
  print(p)
}
# Display the analysis results in a tabular format
print(analysis_results)
#H2
# Step 1: Load Necessary Libraries
library(dplyr)
library(ggplot2)
# Step 2: Read the Dataset
# (Assuming your CSV file "spotify_songs.csv" contains columns such as track_id, track_popularity, and genre)
spotify <- read.csv("spotify_songs.csv", stringsAsFactors = FALSE)
# Optional: View the dataset
View(spotify)
print(spotify)
# Step 3: Preprocess the Data
# Remove duplicate tracks (using track_id) and rows with missing values.
spotify_cleaned <- spotify %>%
```

```
  distinct(track_id, .keep_all = TRUE) %>%  # Remove duplicate tracks
  na.omit()                          # Remove rows with missing data
# Step 4: Convert Variables to Appropriate Types
# Convert track_popularity to numeric (if not already) and genre to a factor.
spotify_cleaned <- spotify_cleaned %>%
  mutate(track_popularity = as.numeric(as.character(track_popularity)),
         genre = as.factor(playlist_genre))
# Step 5: Stating Hypotheses
# Null Hypothesis (H0): There is no significant difference in mean track popularity across genres.
# Alternative Hypothesis (H1): At least one genre has a significantly different mean track popularity.
# Step 6: Perform ANOVA
anova_model <- aov(track_popularity ~ playlist_genre, data = spotify_cleaned)
anova_summary <- summary(anova_model)
print(anova_summary)
# Step 7: Visualize the Results with a Boxplot
ggplot(spotify_cleaned, aes(x = playlist_genre, y = track_popularity)) +
  geom_boxplot(fill = "lightblue", color = "darkblue") +
  theme_minimal() +
  labs(title = "Track Popularity across Genres",
       x = "Genre",
       y = "Track Popularity")


#H3
```

```
# Load necessary libraries

library(dplyr)

library(ggplot2)

library(lubridate)

# Step 1: Read and Preprocess the Data

# Read the dataset (adjust the file path as needed)

spotify <- read.csv("spotify_songs.csv", stringsAsFactors = FALSE)

# Remove duplicate tracks based on track_id and remove rows with missing values

spotify_cleaned <- spotify %>%

  distinct(track_id, .keep_all = TRUE) %>%

  na.omit()

# Handle release date: if only year is provided, append "-01-01"

spotify_cleaned <- spotify_cleaned %>%

  mutate(track_album_release_date = as.character(track_album_release_date),

         track_album_release_date = ifelse(nchar(track_album_release_date) == 4,

                                           paste0(track_album_release_date, "-01-01"),

                                           track_album_release_date),

         track_album_release_date = as.Date(track_album_release_date))

# Extract the release year from track_album_release_date

spotify_cleaned <- spotify_cleaned %>%

  mutate(release_year = year(track_album_release_date))

# Create an 'era' variable: if release_year is before 2000 label as "1900s", otherwise "2000s"

spotify_cleaned <- spotify_cleaned %>%

  mutate(era = ifelse(release_year < 2000, "1900s", "2000s"))
```

```
spotify_cleaned$era <- as.factor(spotify_cleaned$era)

# Step 2: Define the Audio Features to Analyze

# Features: loudness, tempo, danceability, valence, instrumentalness, liveness, energy, acousticness, speechiness

features <- c("loudness", "tempo", "danceability", "valence",

              "instrumentalness", "liveness", "energy", "acousticness", "speechiness")

# Step 3: Loop Through Each Feature and Perform Statistical Tests & Visualization

for (f in features) {

  cat("========================================\n")

  cat("Analyzing feature:", f, "\n")

  # Build the formula for the t-test (e.g., loudness ~ era)

  formula_str <- as.formula(paste(f, "~ era"))

  # Perform a two-sample t-test (Welch's t-test by default)

  t_test_result <- t.test(formula_str, data = spotify_cleaned)

  print(t_test_result)

  # Create a boxplot to visualize the distribution of the feature by era

  p <- ggplot(spotify_cleaned, aes_string(x = "era", y = f, fill = "era")) +

    geom_boxplot() +

    theme_minimal() +

    labs(title = paste("Boxplot of", f, "by Era (1900s vs 2000s)"),

         x = "Era",

         y = f) +

    scale_fill_manual(values = c("1900s" = "lightblue", "2000s" = "salmon"))

  print(p) }
```