https://doi.org/10.22398/2525-2828.1028112-134

How to make a hit: factors associated with music consumption on Spotify

Como fazer um hit: fatores associados ao consumo musical no Spotify

Gabriel Henrique Galvão[I], Ana Flavia Machado[II], Lucas Resende de Carvalho[II]

ABSTRACT

Digitalization has transformed the cultural consumption. The analyses sought to enhance the understanding of consumers' preferences in the music market, which has been significantly transformed by the rise of streaming in recent years. Thus, in this study we aimed at comprehending the characteristics of the most listened-to songs on digital platforms and explore the factors that lead a song to become a hit on Spotify. Data were directly obtained from the platform using the Web API tool provided by the streaming service itself. The used dataset comprises 562,453 songs released between 1922 and 2021, considering data from listeners and artists worldwide. Analyses were conducted using a zero-inflated negative binomial model to investigate interactions between different indicators available on Spotify (audio features) and the popularity of tracks on the platform. On average, songs with higher values for the explicit, danceability, and energy variables demonstrated greater popularity on Spotify.

Keywords: Culture. Consumption. Spotify. Audio features.

RESUMO

A digitalização transformou o consumo cultural. As análises realizadas buscaram aprofundar o entendimento das preferências dos consumidores no mercado musical, que tem passado por transformações significativas com o advento do streaming nos últimos anos. Dessa forma, o objetivo deste estudo foi investigar as características das músicas mais ouvidas nas plataformas digitais e explorar os fatores que contribuem para que uma música se torne um sucesso no Spotify. Os dados foram obtidos diretamente da plataforma por meio da ferramenta Web API fornecida pelo próprio serviço de streaming. O conjunto de dados utilizado abrange 562.453 músicas lançadas entre 1922 e 2021, considerando informações de ouvintes e artistas em escala global. As análises foram conduzidas por meio de um modelo binomial negativo inflado de zeros para examinar as interações entre diferentes indicadores disponíveis no Spotify (características sonoras) e a popularidade das faixas na plataforma. Em média, músicas com valores mais elevados para as variáveis de conteúdo explícito, dançabilidade e energia apresentaram maior popularidade no Spotify.

Palavras-chave: *Cultura. Consumo. Spotify. Características sonoras.*

[I]Universidade Federal de Minas Gerais. Belo Horizonte (MG), Brasil. E-mail: galvaohenriquesgabriel@gmail.com
[II]Universidade Federal de Minas Gerais, Programa de Pós-graduação em Economia. Belo Horizonte (MG), Brasil. E-mail: afmachad@cedeplar.ufmg.br; lucasrc@cedeplar.ufmg.br
Received on: 02/06/2024. **Accepted on:** 02/06/2025

INTRODUCTION

Digitalization has transformed the arts and culture economy in many ways, changing how it produces, consumes, and distributes goods and services (Kübler; Seifert; Kandziora, 2021). Technology has demanded the creation and evolution of existing business models, making it necessary for the cultural goods industry to adapt to market and consumer behavior (Peukert, 2019). Among the various areas that are being transformed by digitalization, the music sector is one of the most prominent. With the emergence of streaming platforms and the facilitation of content sharing via the Internet, the music industry has had to rethink its business model and find new ways of generating revenue (Datta; Knox; Bronnenberg, 2018).

Streaming is an online service that allows users to consume a wide range of audiovisual content on demand. Famous examples of streaming platforms are Spotify, Deezer, and Amazon Music. These market players have found new ways to generate profits by offering users monthly plans that give them access to a vast library of digital music (Datta; Knox; Bronnenberg, 2018). For this reason, streaming has increased its dominance over the music industry in recent years. This service has displaced other markets, such as downloads and physical media, and generated more than 7 billion dollars in revenue in the first half of 2023 only in the USA (Riaa, 2023). However, the changes brought about by this technology are more comprehensive than just the music industry's sources of revenue. Streaming has also transformed how we consume music, enabling the access to users' data that opens doors for developing new studies and research.

Studies on the consumption of cultural goods generally approach demand and consumption from the perspective of the consumer choice theory and primarily aim at calculating elasticities and investigating how sociodemographic factors interfere with the demand for artistic and cultural goods (Conrnig; Levy, 2002). In this sense, few studies are centered on the product, i.e., seeking to understand how music's characteristics affect listeners' consumption. However, this reality has changed with digitization and the development of the field of MIR (Music Information Retrieval). MIR is an emerging area of research that seeks to extract and analyze information from music using audio analysis algorithms (Pérez-Verdejo et al., 2021). One of these algorithms is Spotify's Web API, which allows for the extraction of various information (audio features) about the songs available in the platform's library. This information includes the artist's name, track's BPM, duration, and release date. In addition, Spotify's algorithm calculates indicators that reflect more subjective characteristics of a song. For example, the valence marker describes how much a song conveys positive feelings, and the instrumentalness marker represents the predominance of instrumental elements in a track (Spotify, s.d.).

In this context, in this study we aimed at understanding the characteristics of the most listened-to songs in the world, by observing, through a quantitative method, whether there are regularities among the major global hits on Spotify and identifying these regularities as potential characteristics of a 21st-century pop song. To this end, this study is divided into five sections. First, we provide an Introduction, which aims at giving an overview on the study's subject. Then, in Section 1, we carried out a literature review to understand the effects of digitalization on the creative and cultural industries. In addition, we sought to review existing knowledge of music consumption, addressing studies and research whose authors have identified variables and factors that affect this behavior. In Section 2, we present the methodological aspects, explaining the data and methods used for the analysis. In Section 3, we performed descriptive analyses and evaluated the results found in the model in the light of the literature reviewed. Finally, in Section 4, we summarize the main findings of the research, compiling the most relevant factors related to music consumption on digital platforms. We also address the limitations of the methodology and raise possibilities for future research in the Final Considerations.

DIGITAL CONSUMPTION OF MUSIC: A BRIEF REVIEW OF THE LITERATURE

The impacts of digitalization on cultural consumption

The digitalization of cultural markets has brought significant changes in recent years.

Smartphones and mobile devices have played a significant role in the digitization of the symbolic dimension. These devices have made connecting to the Internet and many applications and services more accessible, allowing people to consume cultural content conveniently on various platforms such as Spotify and Netflix (Alves, 2019).

Recent studies show that users who subscribe to platforms, such as Spotify, see a significant increase in the amount and diversity of music they consume. Personalized recommendations and playlists positively impact the discovery of new tracks. Streaming has introduced new variables into the study on music consumption, and researchers have increasingly turned their attention to the impact of streaming on aggregate demand and consumption of culture (Datta; Knox; Bronnenberg, 2018).

In this regard, the popularization of smartphones and Web 2.0, streaming, and other phenomena have facilitated the process of producing, distributing, and consuming content, resulting in a change in the operating logic of cultural markets. Institutions operating in these markets have had to incorporate new technologies, looking for innovative ways of generating revenue and delivering value to their consumers. Today's creative and entertainment players have reinvented how they "make art" through digitalization.

What is known about music consumption and other cultural goods?

Studies in the field of Cultural Economics are typically guided by the theory of consumer choice and focus on estimating elasticities and analyzing how sociodemographic factors, such as income, age, and education level, affect the demand for artistic and cultural goods (Conrnig; Levy, 2002). Early research in this area investigated whether the price and income elasticities of the demand curve for cultural goods followed the axioms established by the neoclassical theory. However, economists have discovered that certain particularities of cultural goods call for specific studies into the nature of consumption of these products (McKenzie; Shin, 2020).

Moving away from a more general dimension of the consumption of cultural goods, researchers have focused on analyzing the factors that impact music consumption. Many of these studies aim to understand how streaming and the Internet affect consumption and the more "traditional" music industry such as CDs and live concerts. Nguyen, Dejean, and Moreau (2014) applied questionnaires in France and identified that streaming positively affects the music industry in terms of promoting new artists and business models. The authors point out that streaming does not negatively impact physical media sales and even contributes to the success of live performances. However, the authors warn that this result is only valid when consumers cannot download the music and mention that file sharing can be a counterpoint to consolidating these more positive effects. Within this more pessimistic logic, Borja and Dieringer (2016) point out that streaming can play a complementary role in piracy, helping consumers discover new tracks that can later be accessed illegally through other sources.

METHODOLOGICAL ASPECTS

Data used

The analysis of this study was based on data obtained directly from Spotify using the Web API tool provided by the platform. An API (Application Programming Interface) is a set of standards and protocols that allow communication between different pieces of software. It acts as a bridge and helps develop codes to exchange data between two systems[1]. Packages have been developed in Python and R programming languages to simplify work and integration with APIs. For this study, the Spotify and Spotify R packages were chosen, which have features ready to work with Spotify's Web API. The primary objective is to extract relevant data from the songs available in the streaming service's library, including the author's name, release date, and track length. In addition, Spotify calculates indicators that reflect specific characteristics of the tracks, which are fundamental to this study. The selected parameters (audio features) and their respective definitions are based on the official Spotify API documentation and is explained in Chart 1.

In total, the database contains 562,453 songs released between 1922 and 2021.

1 The database does not come from a survey, but from Web API tool provided by the streaming service itself. The authors state that they agree with and have respected all rules defined by the Spotify platform for the use of its API, as described in the Spotify Developer Policy dated May 8, 2024. These rules can be accessed at: https://developer.spotify.com/policy. For this reason, the database is available on the Internet and there are no ethical issues that prevent it from being used.

Chart 1. Variables description.

Variables	Definition	Range
Popularity	Calculated by the platform's algorithm based mainly on the number of plays a song has.	Between zero and one hundred
Acousticness	Acoustic music emphasizes the use of non-electronically processed sounds.	Between zero and one
Danceability	This parameter is estimated based on various other musical elements of a track such as tempo, rhythm, beat intensity, and regularity.	Between zero and one
Duration	Track's duration in minutes.	In minutes
Energy	It is a measure that seeks to capture how much a track conveys an intense, high-activity feeling.	Between zero and one
Explicit	An indicator that shows if a music track contains explicit content, such as profanity, references to violence, drugs, or sexual themes.	Dummy zero if no reference to violence, drugs, or sexual themes
Instrumentalness	A parameter that assesses how instrumental a track is.	Between zero and one
Liveness	It is a marker that detects the presence of an audience at the recording of a track.	Between zero and one
Loudness	It indicates how intense the sounds are in a piece of music, measured in decibels (dB).	From -60 to zero dB.
Mode	Variable that determines whether a track is composed in a major or minor key.	Dummy zero if the track is minor, one if it is major
Speechiness	An indicator that assesses the presence of spoken words in a track.	Between zero and one
Tempo	It measures the pace of a song in beats per minute (BPM).	In BPM
Valence	It indicates how much a track conveys feelings of positivity.	Between zero and one

Model and Techniques Used

Tests were conducted on the database to select the most suitable econometric model to assess the parameters' impact on popularity. A linear model was estimated, showing non-normality and heteroscedasticity in the residuals. To confirm the presence of heteroscedasticity, the Breusch-Pagan test was performed, and the results are shown in Table 1. The test produced a p-value of less than 0.001, indicating evidence to reject homoscedastic residuals (H0) (Breusch; Pagan, 1979).

In Graph 1 we show the Normal Probability Plot, known as a QQ-plot, for the residuals of a linear model. This plot is used to determine whether the analyzed data follows a normal distribution. If the observations are linear, the variable can be considered normally distributed. However, a deviation from linearity is observed

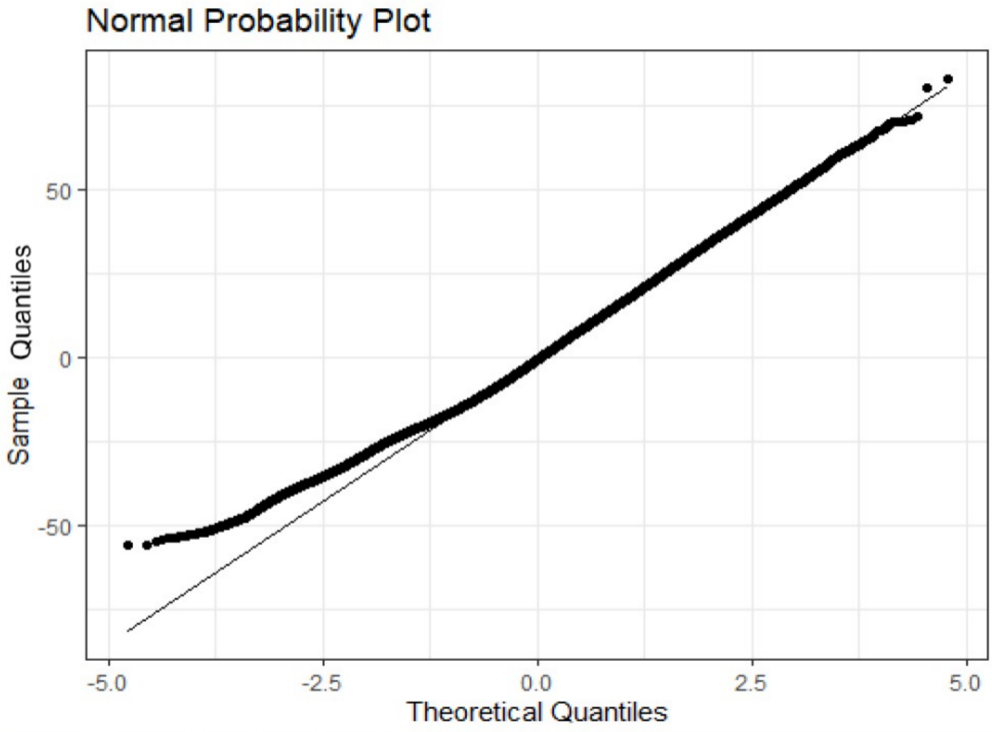
Table 1. Results of the Breusch-Pagan test.

Breusch-Pagan	df	p-value
10665	12	<0.001 ***
		H0: Residuals are **homoscedastic**

df: degrees of freedom.
Source: Prepared by the authors, 2023.

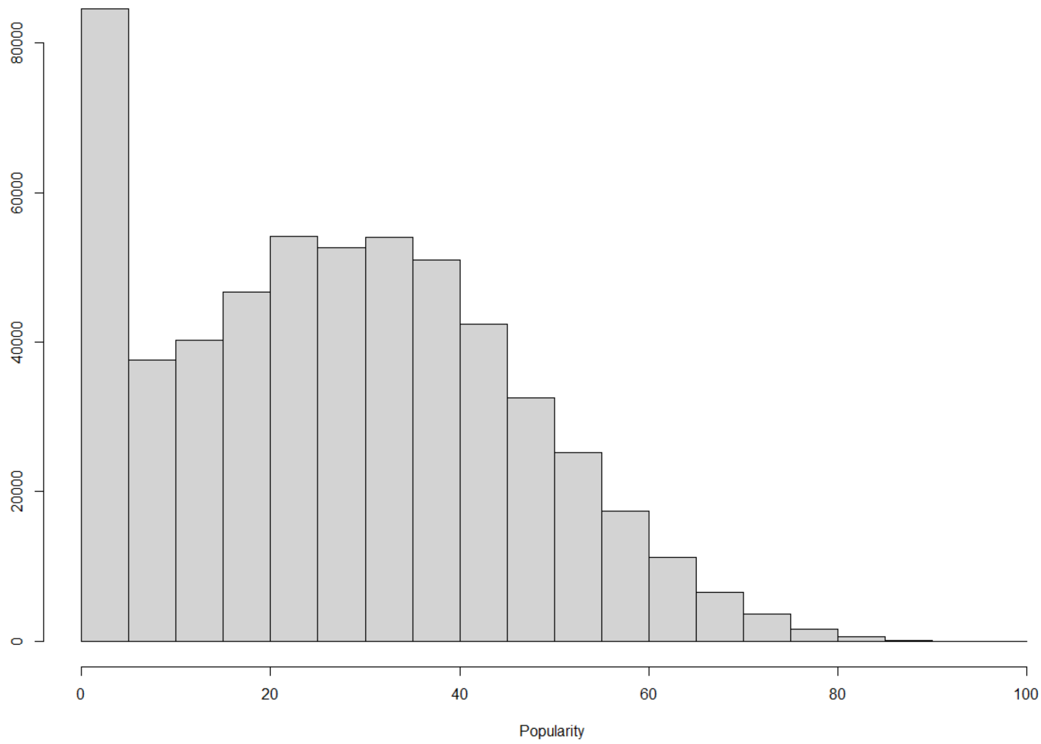
in the first quantiles, indicating that the residuals may not be normally distributed. This deviation is caused by a significant number of songs with a zero value for the response variable, popularity (Graph 2).

Graph 1. Normal Probability Plot.



Source: Prepared by the authors, 2023.

Graph 2. Histogram of Popularity.



Source: Prepared by the authors, 2023.

The findings indicate that the assumptions of Gauss-Markov, required to apply a multiple linear regression model, were not met. To address these concerns, a zero-inflated negative binomial (ZINB) distribution model was used. According to Cameron and Trivedi (2013), a zero-inflated negative binomial distribution is present in a random variable when its probability function is defined by Equation 1.0.

$$P(yi = j) = \{\frac{\pi i + (1-\pi i)g(yi=0), if j=0}{(1-\pi i)g(yi), if j>0}\} \quad (1.0)$$

Essentially, the ZINB model assumes that the considerable volume of zeros in the database arises from a process different from the count values. For this reason, two independent models are estimated: a negative binomial model to deal with the count values (count model) and a logit model to deal with the group of zeros (Long, 1997; Long; Freese, 2001; Long; Freese, 2014).

Therefore, the model that will be used can be written as represented in Equation 1.1:

$$y_i = \beta_0 + \beta_1\chi_1 + \beta_2\chi_2 + \beta_k\chi_k + u \quad (1.1)$$

where:
y_i is the popularity-dependent variable for each song i;
χ_1, ..., χ_k are the independent variables Acousticness, Danceability, Energy, etc;
β_0 is the intercept;
β_1, ..., β_k are the slope coefficients;
u is the error term.

EXPLORATORY ANALYSIS OF THE RESULTS

Descriptive analysis

To understand the impact of different variables on the popularity of Spotify songs[2], it is essential to conduct some descriptive analysis. This includes plotting scatter plots and calculating statistics such as maximum and minimum values, first and third quantile values, mean, and median. In Table 2 we provide a summary of some of the statistics obtained.

Firstly, the average popularity value of Spotify songs is 27.57. In addition, when considering the third quantile, it becomes evident that up to 75% of the songs have a popularity score of 41.00 or less. This implies that most of the songs on Spotify are not very popular.

Another noteworthy finding is related to the explicit and mode dummy variables. The averages of 0.66 for mode and 0.04 for explicit indicate that songs in major tones and without explicit content are more common on Spotify.

2 Examples of songs with high values in the selected variables can be found in the Appendix.

Table 2. Descriptive statistics.

Est.	Populariy	Duration	Explicit	Danceability	Energy	Loudness	Liveness
Minimum	0.00	0.06	0.00	0.00	0.00	-60.00	0.00
Maximum	100.00	93.69	1.00	0.99	1.00	5.38	1.00
1st Quantile	13.00	2.92	0.00	0.45	0.34	-12.89	0.10
3rd Quantile	41.00	4.40	0.00	0.69	0.75	-6.48	0.28
Mean	27.57	3.83	0.04	0.56	0.54	-10.21	0.21
Median	27.00	3.58	0.00	0.58	0.55	-9.24	0.14
Variance	337.48	4.45	0.04	0.03	0.06	25.90	0.03
Stdev	18.37	2.11	0.21	0.17	0.25	5.09	0.18
Est.	**Mode**	**Speechiness**	**Acousticness**	**Instrumentalness**		**Valence**	**Tempo**
Minimum	0.00	0.00	0.00	0.00		0.00	0.00
Maximum	1.00	0.97	1.00	1.00		1.00	246.38
1st Quantile	0.00	0.03	0.10	0.00		0.35	95.60
3rd Quantile	1.00	0.08	0.79	0.01		0.77	136.32
Mean	0.66	0.10	0.45	0.11		0.55	118.46
Median	1.00	0.04	0.42	0.01		0.56	117.38
Variance	0.22	0.03	0.12	0.07		0.07	885.90
Stdev	0.47	0.18	0.35	0.27		0.26	29.76

Est.: statistics; Stdev: standard deviation.
Source: Prepared by the authors, 2023.

Similarly, the averages for the duration and tempo variables provide some interesting information. According to Table 2, Spotify songs generally have an average tempo of 118 beats per minute and an average duration of 3.83 minutes, equivalent to roughly 3 minutes and 49 seconds.

Finally, there are a few additional points worth highlighting. The value of 0.01 for the third quantile of instrumentalness suggests that only a tiny percentage of the songs on Spotify are instrumental. Most of the songs on the platform contain passages of sung lyrics. Similarly, 0.08 in the third quantile of speeches implies that most tracks contain a mix of instrumental elements and sung lyrics. However, there are a few entirely instrumental and "spoken" tracks, such as jazz and classical music, and audiobooks and podcasts, respectively.

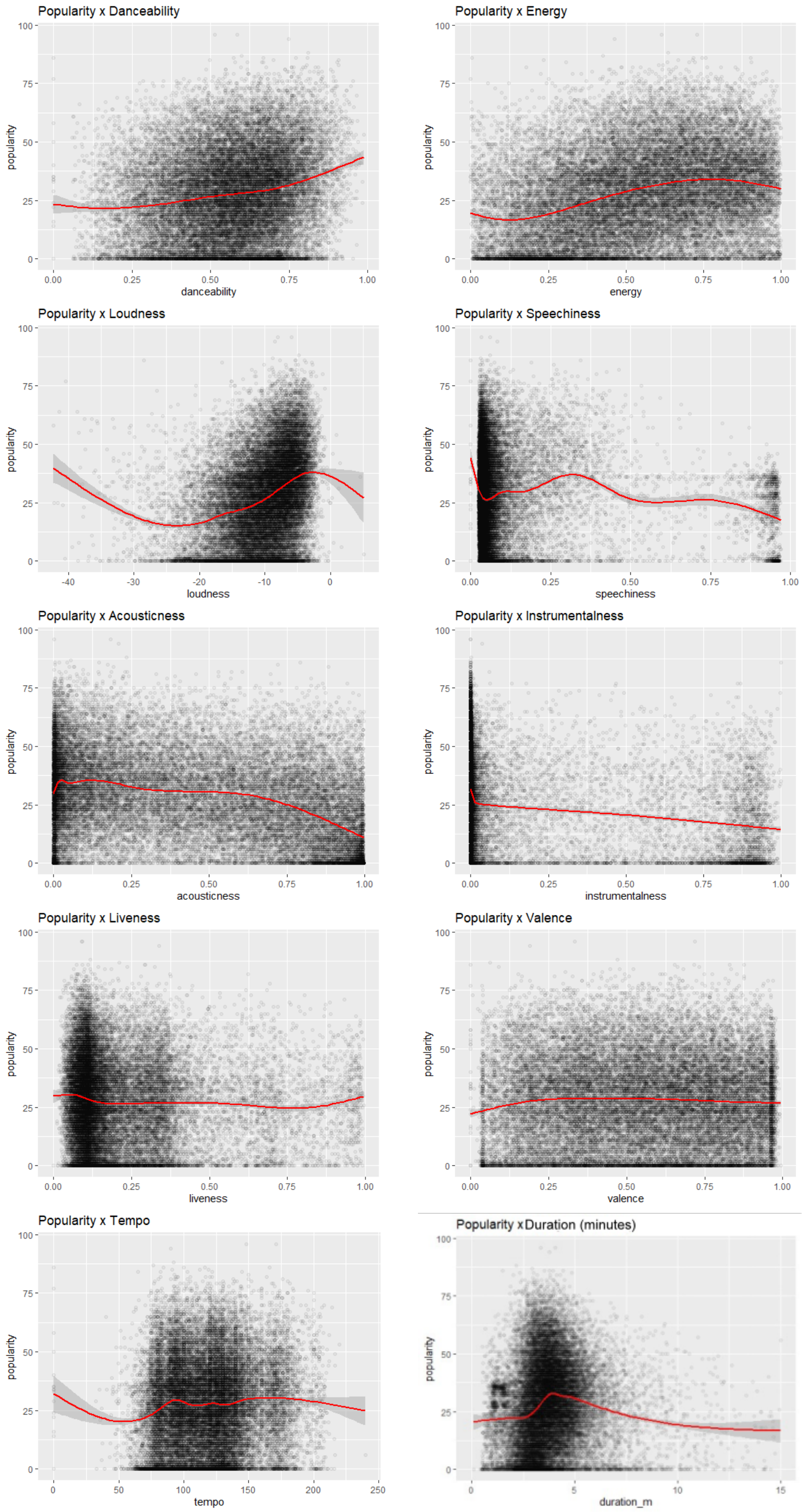
In Graph 3, we display scatter plots for the parameters used, except for mode and explicit, as they are dummy variables and do not display the data well.

Danceability, energy, and loudness tend to correlate positively with popularity, increasing as popularity increases. Conversely, speechiness, acousticness, and instrumentalness variables are observed to have an inverse relationship with popularity, in which popularity decreases with an increase in these parameters. The liveness variable also follows this trend, with most observations between 0 and 0.20 liveness showing a negative trend.

The tempo and valence variables do not show any readily observable trend, with the line remaining stable as the values vary. However, the tempo variable shows some peaks, indicating a multimodal distribution.

Lastly, the duration variable shows a unimodal distribution, with a single peak near the 3-minute mark. This may indicate that the most popular songs have an "optimal" duration in this range.

Graph 3. Scatter Plot with trendline.



Source: Prepared by the authors, 2023.

To better understand the possible correlations and effects of the variables, quantiles were calculated for each parameter and divided into four ranges: 0 to 25 popularity; 26 to 50; 51 to 75; and 75 to 100 popularity.

Regarding the duration variable (Table 3), there is no clear correlation trend between popularity and the length of a song. Although the average length of songs increases from the 0 to 25 popularity group to the 26 to 50 popularity group, the opposite is observed when comparing the 26 to 50 group with the 51 to 75 group. This suggests that the hits orbit around an "optimum duration." Empirically, this idea makes sense, considering that listeners may need more patience to listen to a 5-minute song, while a 30-second may seem brief and have less content to enjoy. The explicit variable shows a positive correlation with popularity, meaning that the higher popularity quantiles have relatively more songs with explicit content than the lower popularity ones.

Table 3. Average of variables by popularity range.

Popularity	Duration	Explicit	Danceability	Energy
0 to 25	3.77 (2.51)	0.01 (0.11)	0.54 (0.17)	0.47 (0.26)
26 to 50	3.91 (1.76)	0.05 (0.21)	0.58 (0.16)	0.60 (0.24)
51 to 75	3.83 (1.32)	0.16 (0.36)	0.62 (0.16)	0.63 (0.21)
76 to 100	3.52 (0.84)	0.32 (0.46)	0.66 (0.15)	0.64 (0.18)
Popularity	**Loudness**	**Liveness**	**Mode**	**Speechiness**
0 to 25	-11.69 (5.29)	0.219 (0.18)	0.67 (0.47)	0.106 (0.19)
26 to 50	-9.28 (4.57)	0.217 (0.19)	0.65 (0.48)	0.108 (0.18)
51 to 75	-7.58 (4.04)	0.189 (0.16)	0.63 (0.48)	0.088 (0.09)
76 to 100	-6.50 (2.91)	0.170 (0.12)	0.60 (0.49)	0.100 (0.09)
Popularity	**Acousticness**	**Instrumentalness**	**Valence**	**Tempo**
0 to 25	0.569 (0.35)	0.167 (0.31)	0.550 (0.26)	116.64 (29.76)
26 to 50	0.359 (0.31)	0.071 (0.21)	0.561 (0.25)	119.80 (29.69)
51 to 75	0.293 (0.28)	0.051 (0.18)	0.531 (0.24)	121.05 (29.59)
76 to 100	0.235 (0.25)	0.019 (0.10)	0.516 (0.23)	121.81 (29.69)

Note: Values in brackets represent the standard deviation.
Source: Prepared by the authors, 2023.

For the danceability variable, there is evidence of a positive correlation as popularity increases, indicating that more danceable songs are, on average, more popular. The same is true of the energy variable, suggesting that more energetic songs are, on average, more popular. The loudness variable also shows evidence of a positive correlation with popularity, as the average loudness increases as popularity increases for all popularity ranges.

Conversely, the mode variable behaves oppositely to the variable of interest, considering that its average decreases as popularity increases. The mode is a dummy variable that indicates whether a song uses a major or minor scale. Major songs are assigned a value of 1, and minor songs, of 0. As the average value of mode decreases with increasing popularity, this indicates that songs in minor keys, which convey more melancholic and introspective sensations, may be, on average, more popular.

Regarding the speechiness variable, there is no clear correlation trend with popularity, although there is a slight movement toward lower speechiness values in the two most popular song ranges. In terms of the acousticness variable, there is evidence of a negative correlation, as the acousticness average decreases as popularity increases. The same is true for the instrumentalness variable, which shows that the instrumentality of songs decreases as popularity increases. This suggests that the most popular songs are less instrumental on average.

The averages of the valence variable slight varies, but there is a more significant downward trend as popularity increases. However, the average difference between the quantiles is slight in absolute terms. Finally, according to the statistics in Table 3, there is a suggestion of an upward movement in tempo as popularity grows. It is worth remembering that tempo, in this case, is different from the duration of a song. While tempo is related to the speed with which a song progresses, duration refers to the length of a song.

Model results

In Tables 4 and 5 we show the estimated outcomes of the model. Overall, the results are consistent with the descriptive analyses and findings of the literature.

Table 4. Estimation results of the zero-inflated negative binomial model (Count Model).

Variable	Exp(B) - Odds Ratio	sd	p-value	
Intercept	35.3584	0.0092	<0.001	***
Duration	1.0050	0.0005	<0.001	***
Explicit	1.3142	0.0042	<0.001	***
Danceability	1.7878	0.0069	<0.001	***
Energy	1.0543	0.0073	<0.001	***
Loudness	1.0183	0.0003	<0.001	***
Mode	1.0068	0.0018	<0.001	***
Speechiness	0.9285	0.0056	<0.001	***
Acousticness	0.7076	0.0038	<0.001	***
Instrumentalness	0.8381	0.0041	<0.001	***
Liveness	0.8838	0.0049	<0.001	***
Valence	0.6339	0.0045	<0.001	***
Tempo	1.0003	0.0000	<0.001	***
Log(theta)	2.7985	0.0023	<0.001	***

sd: standard deviation; Exp(B): odds ratio.
Source: Prepared by the authors, 2023.

Table 5. Estimation results of the zero-inflated negative binomial model (Logit Model).

Variable	Exp(B) - Odds Ratio	sd	p-value	
Intercept	0.0339	0.0580	<0.001	***
Duration	1.0217	0.0019	<0.001	***
Explicit	0.8387	0.0475	<0.001	***
Danceability	0.8501	0.0455	<0.001	***
Energy	0.1074	0.0494	<0.001	***
Loudness	1.0827	0.0017	<0.001	***
Mode	1.0871	0.0122	<0.001	***
Speechiness	5.5224	0.0267	<0.001	***
Acousticness	13.9544	0.0291	<0.001	***
Instrumentalness	7.6955	0.0174	<0.001	***
Liveness	1.1496	0.0336	<0.001	***
Valence	2.5099	0.0299	<0.001	***
Tempo	0.9997	0.0002	0.0879	.

sd: standard deviation; Exp(B): odds ratio.
Source: Prepared by the authors, 2023.

The odds ratio for the duration variable is close to 1, indicating no significant correlation (positive or negative) between the length of a track and its popularity. The odds ratio of 1.005 suggests that if the duration of a track increases by one unit, its popularity increases by 1.005 times or 0.5%. In other words, the impact of varying the length of a song on its popularity is negligible.

However, it is worth noting that this result does not imply that the length of songs is irrelevant to artists and the music industry. On the contrary, it indicates that deliberately manipulating the length of a song, increasing or decreasing its minutes, is not a determining factor in creating a hit in the music charts. Nevertheless, the length of a song can still play a crucial role in the music industry, particularly after the emergence of Spotify, where this characteristic began to directly impact the music sector's turnover. According to Cohen (2023), artists can only receive royalties from streaming a song on Spotify if it is played for at least 30 seconds. Consequently, creators need more financial motivation to make very long songs, because once the playing time limit is exceeded, the revenue generated will be the same.

As a matter of fact, in recent years, there has been a decrease in the average length of the main songs in the market. However, it is worth noting that there is a limit to this reduction. Tracks must be long enough to hold the listener's attention for at least 30 seconds to generate revenue. Additionally, Spotify has already removed songs from artists who tried to benefit from excessively short lengths, such as the band Vulfpeck (Cohen, 2023). Finally, songs that are too short can hinder the artists' ability to express themselves and affect the listener's judgment, as they may need to provide more space to build passages and choruses that connect with the audience.

Artists are likely searching for an "optimum point" of duration to balance their expression needs with the audience's experience and commercial financial strategies. This idea needs further investigation, but it aligns with the descriptive analyses conducted in this study, especially in the graphs in which a unimodal distribution was observed for duration (Graph 3).

The positive coefficient of the explicit variable indicates that, on average, songs with explicit language are more popular. In recent decades, lyrics with swear words or apologies for sexual or violent content have become increasingly prevalent among the hits of the music industry. This trend is evident in the database, in which only 4.36% of the 562,453 analyzed songs have lyrics with some explicit content. However, when considering only the one hundred most famous songs, 42% have explicit content.

To some extent, this phenomenon has been observed in other studies. Primack *et al.* (2008) analyzed the tracks that appeared on Billboard's most popular music charts in 2005 and found that over a third referred to sexual activity, often in a "degrading" way. Furthermore, the authors found that it was common for these songs to address other themes of risky behavior such as drug use, carrying weapons, and acts of violence. Similarly, Hobbs and Gallup (2011) analyzed the lyrics of 174 songs that made the Billboard Top 10 in 2009 and found that approximately 92% alluded to some sexual/reproductive content. They also found that hit songs had a significantly higher number of passages with references to sex compared to less popular songs.

It is noteworthy that the increase in explicit content in song lyrics is directly related to the rise of specific styles such as hip-hop and rap. Madanikia and Bartholomew (2014) showed that the amount of explicit content varies between musical genres. Hip-hop is a genre that adopts an aesthetic of explicit themes and uncensored lyrics more frequently (Primack *et al.*, 2008; Aubrey; Frisby, 2011). According to Spotify (2021), almost a quarter of the total global streams of 2023 were hip-hop songs. The platform also states that, in the last three years, artists from the hip-hop genre, such as Drake, Nicki Minaj, and 21 Savage, have occupied almost half of the spots in the top 50 most listened-to chart. The popularization of this style has contributed to the spread of uncensored lyrics, which helps explain the results found in the model. Nonetheless, it is paramount to observe that the greater permeability of explicit lyrics is not exclusive to rap and hip-hop. Several examples of songs from other genres, such as pop, have been very successful with explicit content such as Justin Bieber's *Peaches* and Olivia Rodrigo's *Good 4 U.*

The variables danceability and energy also showed statistically significant positive coefficients. This suggests that more danceable and energetic songs tend to be more popular on Spotify. Indeed, this result is consistent with findings from other studies (Gao, 2021). Music and dance are closely intertwined practices. It is expected that when listening to music, there is an impulse to tap feet, clap hands, and sway the head (Duman *et al.*, 2022), as music activates certain parts of the brain related to movement (Grahn; Brett, 2007). However, scientists have discovered that along with motor regions, areas related to pleasure and reward are also activated (Duman *et al.*, 2022). For example, Menon and Levitin (2005) found a strong correlation between the enjoyment of music and the release of dopamine, a neurotransmitter associated with movement and pleasure. Thus, more energetic danceable songs may create positive emotional connections with the listeners, as they can promote interactions between reward systems and movement. Ultimately, this helps explain the greater popularity of these tracks, as people are naturally attracted to content that elicits feelings of pleasure.

Another hypothesis concerns the importance of "word of mouth" (WOM) dissemination. It is worth considering that musical preferences vary depending on the context. North and Hargreaves (1996) asked a group of individuals to associate musical characteristics with specific situations. The results showed that pop, energetic, and danceable songs were more associated with integration contexts such as going to a nightclub or a party. Thus, tracks carrying these characteristics are more susceptible to WOM dissemination, as they are frequently preferred in socialization contexts (North; Hargreaves, 1996). Therefore, it is likely that these songs benefit from situations such as concerts and parties, facilitating their popularization among a larger number of listeners.

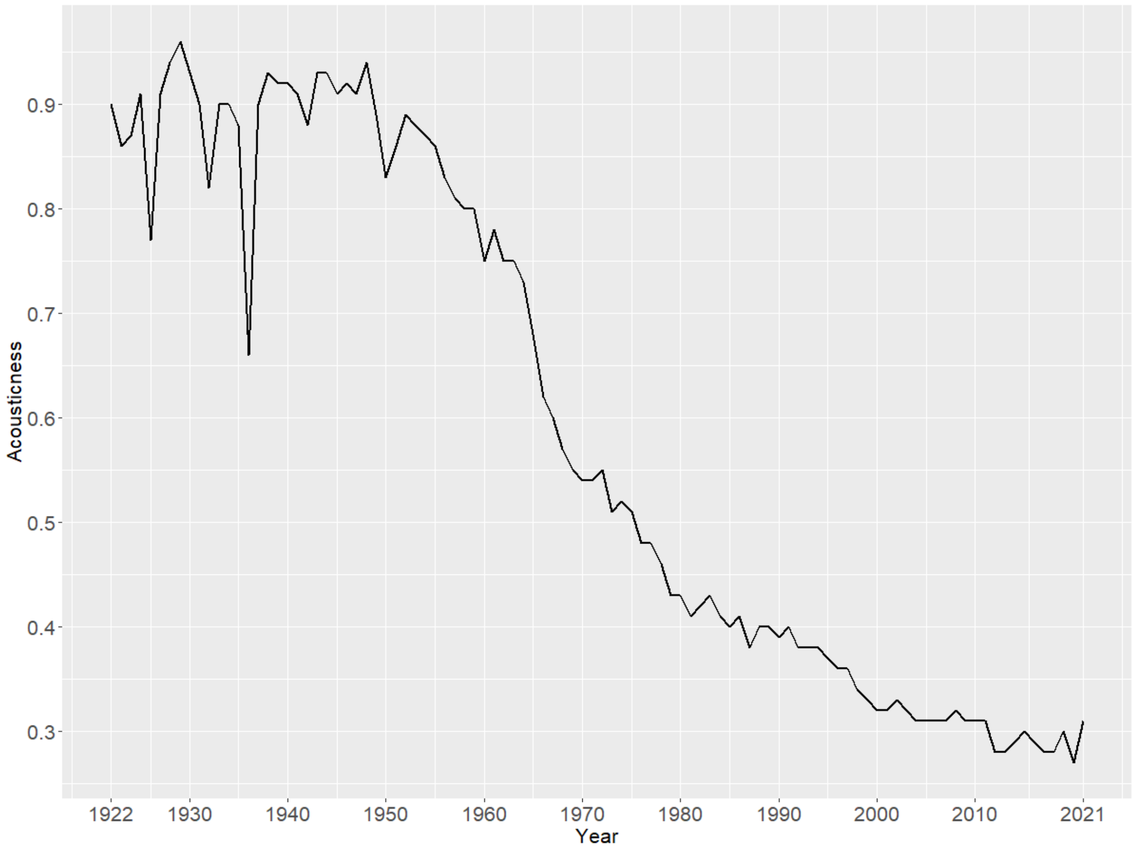
The odds ratio of the speechiness variable indicates that tracks with a higher presence of speech in the recording tend to be, on average, less popular on Spotify. This result is consistent with descriptive analyses. However, for a deeper understanding, it is necessary to evaluate some observations with a high value for this indicator.

It should be considered that although Spotify was initially designed as a music streaming service, it is now also used by artists to promote other audio content. Nevertheless, this poses a challenge when analyzing data, as Spotify's API currently needs to provide a way to separate music from other types of audios. As a result, the variable measuring speechiness may be biased, as non-musical content is likely to have lower average popularity due to how it is consumed.

Spotify determines a track's popularity by counting the times it has been played. Music tends to have significantly more plays than other types of audio content, such as audiobooks or podcasts, due to its accessibility, shorter duration, and repeatability. Spotify's algorithm even automatically creates personalized playlists with songs the user has repeated the most in the last month.

The variable measuring acousticness shows that songs with non-electronically processed sounds and instruments that have few digital effects tend to be, on average, less popular on Spotify. This may be because Spotify's algorithm considers not only the total number of plays a song receives, but also when those plays occurred, when determining its popularity. Songs with many recent plays have higher popularity than songs with the same number of plays in the past.

In Graph 4, we show that the average value of the acousticness variable has considerably decreased over time, indicating that songs released today use more digital elements and effects than those released in the past. This trend is in line with what has been observed in the literature. New technologies have directly impacted innovation in the arts, particularly music, as digital tools — such as synthesizers, samplers, and audio effects — have enabled new means of musical expression. This has expanded creative boundaries and even led to new musical genres. Ultimately, Graph 4. Acousticness over the years (1922 to 2021).



Source: Prepared by the authors, 2023.

modern music reflects contemporary culture, so artists tend to gravitate towards more digital aesthetics.

Nowadays, artists have access to various digital tools that were not popularized in the early 20th century, such as synthesizers, samplers, audio effects (for example, AutoTune), as well as other VSTs[3]. This opens space for the creation of new sounds and musical textures. Furthermore, there is an expansion of creative boundaries, enabling the emergence of new musical genres, such as Dance Music and Hip Hop, but also transforming the sound of existing genres as well as modes of performance (Bakhshi; Throsby, 2012). Ultimately, today's music reflects contemporaneity. For this reason, artists tend to gravitate toward more digital aesthetics, at the expense of the more acoustic compositions of past decades.

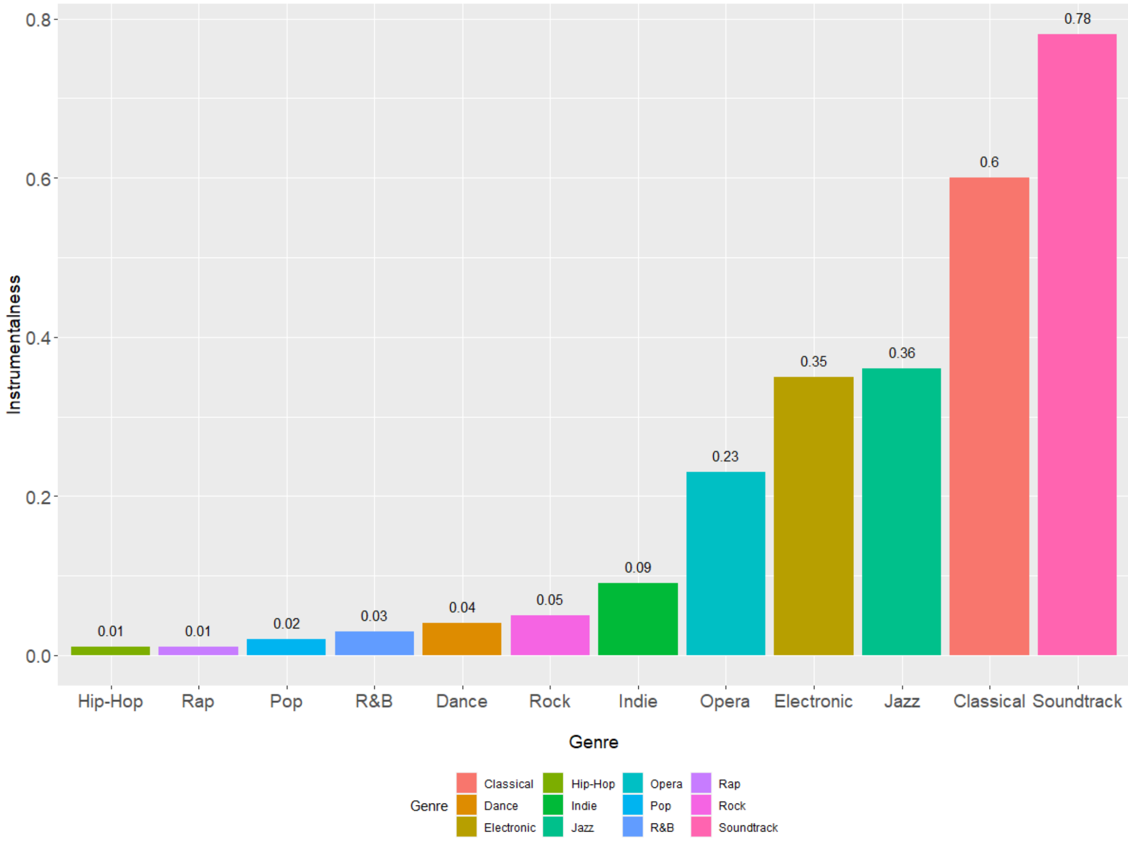
Starting from the hypothesis that, in general, newly released songs receive more recent plays than songs that were released many years ago, there is an explanation for the model's outcome. This is because Spotify's algorithm assigns greater weight to recent plays in determining popularity. Consequently, contemporary songs with a lower degree of acousticness are naturally more popular on average. Conversely, older songs with a high degree of acousticness and more plays "from

3 The acronym stands for Virtual Studio Technology. VSTs are software plugins that enable virtual instruments, audio effects, and signal processors in computer programs for music production.

the past" tend to be less popular. As a result, the variable of acousticness becomes negatively correlated with popularity, which explains the coefficients found.

The odds ratio of the instrumentalness variable indicates that instrumental tracks, on average, are less popular on Spotify. To understand this phenomenon, it is necessary to delve deeper into the songs that possess this characteristic. For this reason, we calculated the average instrumentalness for some musical genres available on Spotify. The result can be observed in Graph 5.

Graph 5. Instrumentalness by genre.



Source: Prepared by the authors, 2023.

We can observe that music genres, such as Soundtrack, Classical, and Jazz, have a high average of instrumentalness. In contrast, genres such as Hip-Hop, Rap, and Pop have a considerably lower average for this variable (Graph 5). It is important to understand that these music genres with low instrumentalness are precisely the most popular among Spotify users. According to Spotify's 2021 Wrapped, the most-listened-to tracks worldwide are associated with low-instrumentality genres.

On the one hand, music genres like classical and jazz, which have a more niche audience, often have a high degree of instrumentalness. Such music is generally not designed for casual listening and lacks characteristics — such as high danceability and easily reproducible choruses — that make it easy to popularize a song. On the other hand, mainstream genres incorporate vocal passages and other elements that allow them to appeal to a broader audience. Therefore, it is comprehensible that the instrumentalness variable has a negative coefficient in the estimated model.

The liveness variable also showed a negative coefficient in the model, indicating that "live" songs are less popular than those recorded in the studio. As studies using Spotify's audio features are relatively recent, we only identified a few of them whose authors have explored the effects of the liveness variable on popularity. In the limited literature found, we observed that there is no consensus. Some researchers support the model's estimates (Febirautami; Surjandari; Laoh, 2018; Nijkamp, 2018), but there are also others that do not (Gao, 2021). However, regardless of the divergences, these studies have one thing in common: the discussions on liveness have been limited to merely mentioning the effect, without delving into more profound explanations of why listeners prefer studio recordings or "live" ones. One hypothesis for the evidence that "live" songs are less popular on Spotify is that of "re-recordings." It has been observed that a common practice in the music market is to release several versions of the same track, mainly when a song has achieved great success. This gives rise to re-recordings such as remixes, covers, and, particularly relevant to this study, "live" versions.

To investigate this hypothesis, we compared the popularity value of songs that have both live and acoustic versions. Overall, the original versions are more popular. In Chart 2 we present a selection of tracks that fall into this category. This can be attributed to several factors such as novelty, familiarity, and recording quality. Listeners may have different impetus to listen to something they have been exposed to several times over life, or they may have built bonds with recordings and the original structure of a song, making them less open to new versions. Moreover, live music has elements — such as applauses and shouts from the audience —, and performers are much more susceptible to mistakes in a live version than in a studio one. For this reason, studio recordings tend to have superior technical quality compared to live versions (Nijkamp, 2018). This can impact listeners' preferences, making them opt for the original versions with lower liveness value.

The valence variable in Spotify's documentation measures the degree of positivity a track conveys. Songs with higher valence values are considered more cheerful and positive, while those with lower values are considered more negative and sadder.

Chart 2. Comparison of liveness and popularity variables.

Track	Artist	Version	Liveness	Popularity
Chan Liveles	Buena Vista Social Club	Studio (1997)	0.1120	68
Chan Buen	Buena Vista Social Club	Live – Live at Carnegie Hall	0.9810	46
The Times They Are A-Changin'	Bob Dylan	Studio (1964)	0.0828	70
The Times They Are A-Changin	Bob Dylan	Live – Live at Sony Music Studios, NY	0.9860	59
Hotel California	Eagles	2013 Remaster – Studio (1976)	0.0575	83
Hotel California	Eagles	Live – Live at the Los Angeles Forum	0.6990	69
Hotel California	Eagles	Live – Live Version	0.8550	50

Source: Prepared by the authors, 2023.

The model shows that songs with lower valence values are more popular on the Spotify platform. However, this result should be critically analyzed, considering that the definition of happiness or sadness is subjective and varies from listener to listener. The algorithms Spotify uses to calculate these variables are not available in the documentation, making it difficult to determine objective criteria for defining a track's positivity level.

Although there is no consensus on this issue, some researchers suggest that listeners are increasingly drawn to sadder and more negative music. For instance, Schellenberg and von Scheve (2012) found that popular music has become more melancholic over the past five decades, mainly due to an increase in songs in minor keys and a decreased average tempo. Another study by DeWall et al. (2011) analyzed the lyrics of popular songs between 1980 and 2007 and found that they increasingly used words related to individualism and antisocial behavior, while words related to social interactions and positive emotions decreased. Finally, a data science study conducted by Napier and Shamir (2018) showed that popular music lyrics have expressed more anger, disgust, fear, and sadness over time, with fewer feelings of joy, trust, and receptivity. There is evidence suggesting that people consume music with a lower valence value. Nevertheless, it is crucial to understand the actual reasons behind this preference for less cheerful music.

Bennet (2008) states that popular music functions as a reflection of society, exposing the needs of individuals in a cultural and social context. Throughout history, we have seen the rise of various musical styles and aesthetics that sought to translate the desires of certain social groups. Punk, for example, emerged as the soundtrack of a youth revolt against class inequalities and the dominant values of the time. For this reason, punk songs sought to translate revolt, rebellion, and antiauthoritarianism into their lyrics and melodies (Hebdige, 1979). Conversely, rap became popular among young African Americans from the suburbs who were marginalized and disgusted by social injustices and problems (Rose, 1994). Not coincidentally, many rap tracks deal with issues such as police violence, sex, and drugs. Hence, part of rap's success lies in its ability to express the desires of young people from disadvantaged urban areas.

Similar to what happened with punk and rap, we could ponder that the rise of less cheerful music reflects a contemporary society made up of more isolated and reflective individuals in which social relationships are marked by uncertainty and fragility (Elias, 2010).

Another factor that should be mentioned is the influence of the new coronavirus (COVID-19) pandemic. As the data were collected in 2021, the global context of crisis and social isolation may have influenced the result of a preference for less cheerful music. Authors of some studies already indicate that individuals who have contracted COVID-19 have had negative impacts on their mental health (Vindegaard; Benros, 2020). Other researchers are beginning to find evidence that the lockdown has somehow altered music consumption (Yeung, 2020). However, more in-depth analysis is still needed to state that listeners' preferences have changed during the pandemic, especially regarding the demand for sadder music as a reflection of the global situation. This topic is relevant to new research agendas and could be explored in future studies.

Lastly, it is worth noting that the Loudness, Mode, and Tempo markers had coefficients close to zero and, therefore, were not considered in the analysis. All three variables have significant p-values; however, this is due to the large number of observations in the data and does not necessarily mean that these parameters significantly affect popularity. The odds ratios indicate the opposite. For instance, the odds ratio of 1.0068 for the Mode variable (Table 4) indicates that songs in lower tones are only 0.68% more popular. In other words, according to the estimated model, the effect of Mode on popularity is negligible. The same goes for Loudness and Tempo.

FINAL CONSIDERATIONS

In this study we provided a comprehensive discussion on the consumption of cultural goods, focusing on the changes brought about by digitization in music consumption. Digitization has made it easier for new artists to enter the music industry, providing a broader and more diverse music catalog. Streaming has played a significant role in this process, introducing a new way of listening to music and overthrowing current models. Digitization has also led to changes in the most famous music's aesthetics and styles. Furthermore, the music market's funding logic has also been affected by streaming, leading to new ways of generating revenue.

We sought to innovate in this article by offering a new approach to analyzing the quantitative factors that influence consumer behavior. We focused on the product and usage markers that reflect the characteristics of Spotify's songs (Audio Features), looking for evidence of factors interfering with listeners' behavior on the platform.

According to the results, Spotify's music consumption is mainly in line with what has been observed in the literature. Overall, songs with higher values for explicit parameters, danceability, and energy were more prevalent on Spotify. The opposite was observed for speechiness, acousticness, instrumentalness, liveness, and valence. Nonetheless, there are still limitations to the research carried out, mainly related to the fact that Spotify restricts access to the platform's data.

Furthermore, it is currently impossible to conduct studies using anonymized listeners' data made available by the platform. If this rule is changed, it will provide much scope for future research agendas such as assessing how some sociodemographic factors impact the preferences of the platform's listeners. It would also be relevant to assess regional differences in music consumption in future studies.

REFERENCES

ALVES, Elder P. Maia. A digitalização do simbólico e o capitalismo cultural-digital: a expansão dos serviços culturais-digitais no Brasil. **Revista Sociedade e Estado**, Brasília, v. 34, n.1, p. 129-157, jan. 2019.

AUBREY Jennifer Stevens; FRISBY Cynthia M. Sexual objectification in music videos: A content analysis comparing gender and genre. **Mass Communication and Society**, v. 14, n. 4, p. 475-501, jul. 2011.

BAKHSHI, Hasan; THROSBY, David. New technologies in cultural institutions: theory, evidence, and policy implications. International Journal of Cultural Policy, Inglaterra, v. 18, n. 2, p. 205-222, mar. 2012.

BENNETT, Andy. Towards a cultural sociology of popular music. **Journal of Sociology**, Austrália, v. 44, n. 4, p. 419-432, dez. 2008.

BORJA, Karla; DIERINGER, Suzanne. Streaming or stealing? The complementary features between music streaming and music piracy. **Journal of Retailing and Consumer Services**, Holanda, v. 32, p. 86-95, set. 2016.

BREUSCH, Trevor S.; PAGAN, Adrian R. A Simple Test for Heteroscedasticity and Random Coefficient Variation. **Econometrica**, v. 47, n. 5, p. 1287-1294, set. 1979.

CAMERON, Adrian Colin; TRIVEDI, Pravin K. **Microeconometrics:** Methods and Applications. Cambridge: Cambridge University Press, 2013.

COHEN, Daniel. To Monopolise Our Ears. **London Review of Books**, v. 45, n. 9, 2023.

CORNING, Jonathan; LEVY, Armando. Demand for Live Theater with Market Segmentation and Seasonality. **Journal of Cultural Economics**, Nova York, v. 26, n. 3, p. 217-35, 2002.

DATTA, Hannes; KNOX, George; BRONNENBERG, Bart J. Changing Their Tune: How Consumers' Adoption of Online Streaming Affects Music Consumption and Discovery. **Marketing Science**, Catonsville, v. 37, n. 1, p. 5-21, jan. 2018.

DEWALL, C. Nathan; POND, Richard S.; CAMPBELL, W. Keith; TWENGE, Jean M. Tuning in to psychological change: Linguistic markers of psychological traits and emotions over time in popular U.S. song lyrics. **Psychology of Aesthetics, Creativity, and the Arts**, Washington, v. 5, n. 3, p. 200-207, ago. 2011.

DUMAN, Denis; NETO, Pedro; MAVROLAMPADOS, Anastasios; TOIVIAINEN, Petri; LUCK, Geoff. Music we move to: Spotify audio features and reasons for listening. **Plos One**, v. 17, n. 9, e0275228, set. 2022.

ELIAS, Norbert. **The civilizing process**: sociogenetic and psychogenetic investigations. Malden, Massachusetts: Blackwell Publishing, 2010.

FEBIRAUTAMI, Limisgy Ramadhina; SURJANDARI, Isti; LAOH, Enrico. Determining Characteristics of Popular Local Songs in Indonesia's Music Market. *In*: INTERNATIONAL CONFERENCE ON INFORMATION SCIENCE AND CONTROL ENGINEERING (ICISCE), 5., 2018. p. 197-201.

GAO, Andrea. Catching the Earworm: Understanding Streaming Music Popularity Using Machine Learning Models. **E3S Web of Conferences**, v. 253, p. 03024, 2021.

GRAHN, Jessica A.; BRETT, Matthew. Rhythm and beat perception in motor areas of the brain. **Journal of Cognitive Neuroscience**, v. 19, n. 5, p. 893-906, 2007.

HEBDIGE, Dick. **Subculture**: The Meaning of Style. Londres: Routledge, 1979.

HOBBS, Dawn R.; GALLUP, Gordon G. Songs as a medium for embedded reproductive messages. **Evolutionary Psychology**, v. 9, n. 3, p. 390-416, jul. 2011.

KÜBLER, Raoul; SEIFERT, Rouven; KANDZIORA, Michael. Content valuation strategies for digital subscription platforms. **Journal of Cultural Economics**, v. 45, n. 2, p. 295-326, 2021.

LONG, J. Scott. Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage, 1997.

LONG, J. Scott; FREESE, Jeremy. Predicted Probabilities for Count Models. **The Stata Journal**, v. 1, n. 1, p. 51-57, nov. 2001.

LONG J. Scott; FREESE, Jeremy. Regression models for categorical dependent variables using Stata. 3rd. ed. College Station, Texas: Stata Press Publication, 2014.

MADANIKIA, Yasaman; BARTHOLOMEW, Kim. Themes of Lust and Love in Popular Music Lyrics from 1971 to 2011. **Sage Open**, Nova York, v. 4, n. 3, p. 1-8, 2014.

MCKENZIE, Jordi; SHIN, Sunny Y. Demand. In: TOWSE, Ruth. **Handbook of Cultural Economics**. Cheltenham: Edward Elgar, 2020. p. 216-227.

MENON, V.; LEVITIN, D. J. The rewards of music listening: Response and physiological connectivity of the mesolimbic system. **NeuroImage**, v. 28, n. 1, p. 175-184, out. 2005.

NAPIER, Katleen; SHAMIR, Lior. Quantitative Sentiment Analysis of Lyrics in Popular Music. **Journal of Popular Music Studies**, v. 30, n. 4, p. 161-176, dez. 2018.

NGUYEN, Godefroy Dang; DEJEAN, Sylvain; MOREAU, François. On the complementarity between online and offline music consumption: the case of free streaming. **Journal of Cultural Economics**, v. 38, n. 4, p. 315-330, 2014.

NIJKAMP, Rutger. **Prediction of product success**: explaining song popularity by audio features from Spotify data. 2018 Tese - Faculty of Behavioural, Management and Social sciences, University of Twente, Enschede, 2018.

NORTH, A. C.; HARGREAVES, David J. Situational influences on reported musical preference. **Psychomusicology: A Journal of Research in Music Cognition**, v. 15, n. 1-2, p. 30-45, 1996.

PÉREZ-VERDEJO, J. Manuel; PIÑA-GARCÍA, C. A.; OJEDA, Mario Miguel; RIVERA-LARA, A.; MÉNDEZ-MORALES, L. The rhythm of Mexico: an exploratory data analysis of Spotify's top 50. **Journal of Computational Social Science**, Singapura, v. 4, n. 1, p. 147-161, 2021.

PEUKERT, Christian. The next wave of digital technological change and the cultural industries. **Journal of Cultural Economics**, v. 43, n. 2, p. 189-210, 2019.

PRIMACK, Brian A.; DALTON, Madeline A.; CARROL, Mary V.; AGARWAL, Aaron A.; FINE, Michael J. Content analysis of tobacco, alcohol, and other drugs in popular music. **Archives of Pediatrics & Adolescent Medicine**, Pittsburgh, v. 162, n. 2, p. 169-175, fev. 2008.

RIAA - Recording Industry Association of America. RIAA Mid-Year 2023 Revenue Report. Washington, 2023.

ROSE, T. Black noise: Rap music and Black Culture in Contemporary America. Hanover, New Hampshire: Weskeyan University Press, 1994.

SCHELLENBERG, E. Glenn; VON SCHEVE, Christian. Emotional cues in American popular music: Five decades of the Top 40. **Psychology of Aesthetics, Creativity, and the Arts**, v. 6, n. 3, p. 196-203, ago. 2012.

SPOTIFY. Nearly a Quarter of All Streams on Spotify Are Hip-Hop. Spotify's Global Editors Reflect on the Genre's Growth. Spotify, 2023. Available at: https://newsroom.spotify.com/2023-08-10/hip-hop-50-murals-new-york-atlanta-miami-los-angeles/. Access on: Oct. 5, 2023.

SPOTIFY. Spotify: Web API. Spotify, s.d. Available at: https://developer.spotify.com/documentation/web-api. Access on: Apr. 23, 2023.

SPOTIFY. What The World Streamed Most in 2021. For the Record, Estocolmo, 1 de dez. de 2021. Available at: https://newsroom.spotify.com/2021-12-01/what-the-world-streamed-most-in-2021//. Access on: Oct. 26, 2023.

VINDEGAARD, Nina; BENROS, Michael Eriksen. COVID-19 pandemic and mental health consequences: Systematic review of the current evidence. **Brain, Behavior, and Immunity**, v. 89, p. 531-542, 2020.

YEUNG, Timothy Yu-Cheong. Did the COVID-19 Pandemic Trigger Nostalgia? Evidence of Music Consumption on Spotify. SSRN Electronic Journal, 2020.

About the authors

Gabriel Henique Galvão: Economist from Universidade Federal de Minas Gerais.
Ana Flavia Machado: Full Professor of the Department of Economic Sciences from Universidade Federal de Minas Gerais, PhD in Economy from Universidade Federal do Rio de Janeiro.
Lucas Resende de Carvalho: Adjunct Professor of the Department of Economic Sciences from Universidade Federal de Minas Gerais, PhD in Economy from Universidade Federal de Minas Gerais.

Conflicts of interests: nothing to declare – **Financial support:** none.
Authors' contributions: Galvão, G. H.: Conceptualization, Data Curation, Methodology, Writing – Original Draft. Machado, A.F.: Conceptualization, Supervision, Validation, Writing – Review & Editing. de Carvalho, L.R.: Methodology, Supervision, Validation, Writing – Review & Editing.

APPENDIX A – TRACKS THAT EXEMPLIFY THE INDEPENDENT VARIABLES OF THE MODEL

Track	Artist	Variable	Value	Description
Click on the song name to listen to it on Spotify.				
Don Giovanni, K. 527 (Excerpts Sung in German)	Mozart, Richard Mayr	Acousticness	0.996	Music with many acoustic characteristics
SICKO MODE	Travis Scott	Acousticness	0.005	Music with few acoustic characteristics
Billie Jean	Michael Jackson	Danceability	0.920	Music with high danceability
Anyways	Arctic Monkeys	Danceability	0.280	Music with low danceability
A Love Supreme, Pt. III – Pursuance	John Coltrane	Duration	10.70	Music with long duration
Her Majesty	The Beatles	Duration	0.420	Music with short duration
Chop Suey!	System of a Down	Energy	0.934	Energetic music
Espelho	Alexandre Galvão	Energy	0.292	Not very energetic music
Fuck Tha Police	N.W.A	Explicit	1	Explicit music
Mistério do Planeta	Novos Baianos	Explicit	0	Non-explicit music
Mia & Sebastian's Theme – From "La La Land" Soundtrack	Justin Hurwitz	Instrumentalness	0.928	Very instrumental music
Venom – Music from The Motion Picture	Eminem	Instrumentalness	0.001	Not very instrumental music
Liberdade para Dentro da Cabeça – Ao Vivo	Natiruts	Liveness	0.713	Music with many live recording characteristics
Chega de Saudade	Toquinho, Vinicius de Moraes	Liveness	0.127	Music with few live recording characteristics

Source: Prepared by the authors, 2023.