INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Analysis of the User Behavior and Music Preferences: Suggestions for Refining Spotify's Algorithms in Various Contexts

Monika Pacut

Master in Marketing

Supervisor:
PhD Maria de Fátima Ramalho Fernandes Salgueiro,
Full Professor, Department of Quantitative Methods for Management & Economics at
ISCTE Business School

September, 2024

BUSINESS
SCHOOL

Department of Marketing, Operations & General Management

Analysis of the User Behavior and Music Preferences: Suggestions for Refining Spotify's Algorithms in Various Contexts

Monika Pacut

Master in Marketing

Supervisor:
PhD Maria de Fátima Ramalho Fernandes Salgueiro,
Full Professor, Department of Quantitative Methods for Management & Economics at
ISCTE Business School

September, 2024

Resumo

Esta dissertação investiga as preferências dos utilizadores do Spotify, focando o papel da plataforma na formação dos gostos musicais, a evolução das preferências musicais ao longo da última década e a influência de Taylor Swift nas escolhas dos utilizadores. O estudo procura contribuir para a compreensão do comportamento dos utilizadores no contexto dos serviços de streaming de música, recorrendo a metodologias que incluem a análise SPSS - principalmente com análise de regressão de conjuntos de dados segmentados por geografia, clima e tempo, complementada por uma revisão abrangente da literatura de artigos académicos e relatórios da indústria.

Os resultados da análise SPSS, combinados com a pesquisa de relatórios, revelaram várias informações importantes. Em primeiro lugar, o Spotify desempenha um papel significativo na formação das preferências dos utilizadores, com o país e as condições meteorológicas a actuarem como factores influentes na popularidade das músicas e, consequentemente, nas preferências dos utilizadores. Em segundo lugar, as preferências musicais sofreram alterações notáveis nos anos 1920-1950, causadas por muitas mudanças culturais e tecnológicas. A caraterística que apresentou a diminuição mais substancial foi a acústica, reflectindo a evolução da música pop e eletrónica. Por último, a discografia de Taylor Swift oferece informações valiosas sobre a dinâmica de mudança das preferências dos utilizadores, demonstrando o potencial da influência dos artistas para moldar o comportamento de audição e sugerir modelos de recomendação da potencial popularidade das canções.

Estas conclusões sublinham a importância de integrar a segmentação geográfica e as considerações meteorológicas nos algoritmos de transmissão em contínuo para melhorar a participação dos utilizadores. Além disso, a análise do desenvolvimento da carreira dos artistas fornece um quadro para compreender as mudanças nos gostos dos consumidores, o que pode ajudar a aperfeiçoar os sistemas de recomendação. As implicações desta investigação estendem-se à conceção de experiências de utilizador personalizadas e ao desenvolvimento de modelos preditivos, oferecendo lacunas para investigação futura na recuperação de informação musical e estratégias de marketing.

JEL: M31 Marketing

Abstract

This dissertation investigates the user preferences on Spotify, focusing on the platform's role in forming musical tastes, the evolution of music preferences over the past decade, and the influence of Taylor Swift on users' choices. The study seeks to contribute to the understanding of user behavior within the context of music streaming services, using methodologies that include SPSS analysis - mainly with regression analysis of datasets segmented by geography, weather, and time, complemented by a comprehensive literature review of scholarly articles and industry reports.

The results of the SPSS analysis combined with report research showed several key insights. First, Spotify plays a significant role in forming user preferences, with country and weather conditions acting as influential factors in song popularity, therefore user preferences. Second, music preferences have experienced notable changes in the years 1920-1950, caused by many cultural and technological shifts. The trait that showed the most substantial decrease was accousticness, reflecting the evolution of Pop and Electronic music. Finally, Taylor Swift's discography offer valuable insights into the changing dynamics of user preferences, demonstrating the potential of artists' influence to shape listening behavior and suggest recommendation models of potential song popularity.

These findings underscore the importance of integrating geographic segmentation and weather considerations into streaming algorithms to enhance user engagement. Moreover, the analysis of artists' career development provides a framework for understanding the shifts in consumer tastes, which could help in refining recommender systems. The implications of this research extend to the design of personalized user experiences and the development of predictive models, offering gaps for future research in music information retrieval and marketing strategies.

JEL: M31 Marketing

Table of contents

Introduction

Following the dynamic growth of the music streaming industry and personalization trends, this dissertation aims to contribute to the research in the area of user behavior by investigating the strategies and technologies used by a music streaming platform, Spotify. The primary objective of this work is conducting a comprehensive, quantitative analysis of the publicly available datasets of Spotify song variables and their impact on user behaviors, experiences, and on the recommendation algorithms.

To justify this research, it is crucial to bring the attention to the increasing significance of personalization in the music streaming industry and in the overall global market, noticing simultaneously the limited academic exploration of Spotify's specific strategies and their implications (Sletten, 2021). This dissertation seeks to fill this gap to some extent by examining user preferences and user behavior on Spotify, distinguishing it from other platforms like YouTube and Apple Music (Beuscart et al.,2023).

User behavior analysis takes center stage in the structure of this dissertation, aiming at identifying strategies for retaining users and reducing churn rates that contributes significantly to the platform's long-term success. The dissertation will explore emerging trends and technologies in music personalization, providing insights into their influence on the future of the music streaming industry. In the current technological era, everything is focused on AI and machine learning, thus the described industry should follow the advancements to create even more personalized and engaging music experiences on platforms like Spotify for sustained competitiveness and innovation.

Alongside digital development, all industries went and still must go through major strategical shifts. Interesting case in this matter is the situation of music industry that changed completely due to the emerge of music streaming platforms. In the results from IFPI Global Music Report 2023, it was stated that music streaming platforms grew by over 10% in 2022 and the annual revenue surpassed $17 billion (IFPI, 2023). Further details on the annual revenues of the music streaming industry can be found in Chart 7 in the annex. On the global scale, IFPI identified that 67% of total music industry revenue comes directly from music streaming platforms, comparing to 16,1% coming from physical sales. As of number of subscribers, the Global Music Report stated that the number has increase almost 10 times from 68 million in Q4 (period from

October to December) of 2015 to 616,2 million in Q2 (period from April to June) of 2022. Nowadays, an average listener spends over 20 hours on listening to music.

According to recent findings, Forbes stated that music streaming accounts for 84% of total music industry revenue in the United States (Durrani, 2023). The industry experiences the rapid development during pandemic, when music streaming subscriptions increased by 13% from Q4 (period from October to December) of 2019 to Q1 (period from January to March) of 2020, with elimination of ads as the reason for their paid subscription. The researchers also established that 32% of music streamers use Spotify, situating this platform on the top position in this regard. All this explains why this dissertation focus on this company: Spotify.

As there are significant disruptive changes in the consumption of music, understanding how Spotify's personalization techniques influence user behavior and vice versa, becomes a crucial matter (Fricke et al., 2021). This study focuses on Spotify's platform user data, especially from the past decade, available in different context on a commonly used website named Kaggle. The following datasets will be used in this dissertation: Spotify dataset (Spotify Dataset, 2021), Top Spotify Songs in 73 Countries (Top Spotify Songs in 73 Countries (Daily Updated), 2024), 6k Weather Labeled Spotify (6K Weather Labeled Spotify Songs, 2023) and Taylor Swift Spotify Dataset (Taylor Swift Spotify Dataset, 2024).

The choice of word “user” instead of any other can be justified by the essence of what this dissertation aims to measure – actual behavior of individuals ‘using’ the Spotify music streaming platform.

The research involves analysis of user datasets, consecutive marketing insights, user preferences and explores potential qualitative factors.

The analysis will consider not only the global based data but also datasets in the local context, emphasizing the need for country-specific perspective. By such focus, the research tries to uncover unique country-based patterns in user behavior on Spotify. This study will address potential challenges and opportunities in enhancing personalization on Spotify, finally suggesting actions to optimize algorithms for improved user engagement and loyalty.

Furthermore, this dissertation will analyze the evolution of music preferences on Spotify over the past, seeking to identify the contributing factors to these shifts, such as cultural events, emerging artists, and technological developments. It will evaluate how Spotify's recommendation system utilizes Artificial Intelligence (AI) to provide personalized music suggestions and assess its impact on user engagement and satisfaction. It will examine the impact of music streaming platforms on traditional music consumption methods, including CD

sales, radio, and music charts. An additional role will be taken by personalized playlists and features like "Discover Weekly", contributing to user engagement and loyalty on Spotify.
Three main research questions will be addressed in this dissertation:

- What is the role of Spotify in forming preferences from the users' perspective? – answered in the analysis based on dataset by country and weather.
- How have music preferences on Spotify evolved over the past decade? – answered in the analysis conducted on dataset by year.
- What is the scope of influence of Taylor Swift on user musical preferences? – answered in the analysis of her discography on Spotify.

Before that, Chapter 1 gives a brief overview of the music industry and Spotify, using a user approach, while Chapter e revises the literature on musical user preferences. After addressing the research questions in Section 3, this dissertation ends with key conclusions and recommendations.

1. Music Industry and Spotify, a user approach

1.1 User-based growth of Spotify

In the dynamic years of the mid-2000s, a Swedish web designer and entrepreneur named Daniel Ek decided to create a new startup in the industry of digital music (Development, 2022). Ek, having already engaged in entrepreneurial ventures since the age of fourteen, envisioned a platform that would revolutionize how people experienced music. His career began with the creation of Advertigo in 2005, a successful enterprise that he was a manager of until its acquisition by Tradedoubler in 2006. Later, Ek and his business partner Martin Lorentzon, the founder of Tradedoubler, created a venture that would soon become the main player in music streaming - Spotify. The name is a combination of "spot" and "identify," and aims to illustrate the essence of the platform— a place to identify with the music the user loves (*Spotify &Mdash; About Spotify*, 2023).

In 2006, Ek and Lorentzon registered the brand and began to work on Spotify, relying on the expertise of a group of talented designers and engineers, including Ludvig Strigeus, the developer of uTorrent. The development of Spotify took a few months, but its launch was delayed by negotiations with record labels and licensing agreements. Finally, on October 7, 2008, Spotify was launched, initially appearing in the European countries. Spotify's first offering was a streaming music playback application for computers, offering users a vast music library without the need for downloads. The platform featured the freemium model, where a free account was available to anyone but included advertisements between songs, while a paid subscription provided unrestricted access.

With the technology development, smartphones entered the markets, creating a need for Spotify to expand into different devices. In 2009, the platform introduced its mobile application, along with Premium and Unlimited modes.

Spotify's popularity grew rapidly over the years, leading to its entry into the United States of America market in 2011. The platform achieved contracts with major record labels, competing with established music applications like iTunes, Amazon Music, Pandora, and the developing Tidal, Spotify quickly surpassed others in the number of active, loyal users, indicating how immersive the experience on this app was and is.

In 2011, Spotify undertook transformation and diversification, positioning itself as a platform hub for various music-related applications. Ek envisioned Spotify as more than just a music platform, proclaiming that "Music goes beyond the music." The platform evolved into a multi-platform ecosystem, incorporating among others, third-party applications such as TuneWiki

with lyrics, CrowdAlbum, providing photographs into the platform, and Billboard with musical lists and charts. This underlines the user focus of Spotify and can be considered one of the rationales behind this dissertation.

The expansion of Spotify continued, integrating specialized channels and magazines like Rolling Stone and Moi. It eliminated the listening hours limit for free users in 2014, making the experience even more accessible. The platform integrated marketing, sales, and user experience seamlessly together by partnering with for example Topspin and Starbucks, advertising itself, for example in cafes. Spotify positioned itself as one of the most commonly used apps, securing partnerships with social networks like Facebook, Twitter, Instagram, and Snapchat. These partnerships allowed users to seamlessly access and share content among different channels, enhancing the overall musical experience.

By the year of 2016, Spotify-made lists like Discover Weekly, Billboard Hot 100, and Hipster International became user favorites. In 2018, Spotify entered the stock market with a market value of $26.5 billion. With over 574 million users in 184 countries across the whole world in 2023, Spotify has become a significant part of the global digital music composition (Development, 2022).

Crucial matter to be mentioned is the fact that Spotify's keeps on balancing conflicting markets for music, listeners, advertisers and finance, as it tries to ensure a growing consumer base for music producers, as well as industry advertisers, while simultaneously attracting continuous investment. Ad-supported part of this platform is crucial for its finances, as it drives 60% of its total gross added Premium Subscribers, and the advertising revenue is growing faster than premium subscription revenue (Prey, 2020). Spotify brands itself as a platform that connects engaged listeners with their preferred music creators, using native advertising, which is a strategy of presenting consumers with advertisements that resemble the non-advertising content published on the same platform (Wojdynski, B.W., 2016). In the case of Spotify, this is done especially through playlists with, for example, Sponsored Playlists or Branded Moments. Here, playlists are a way to provide brands that advertise on this platform with opportunities to target users in specific moments and moods, depending on what music they are currently listening to. Spotify reported its first profitable quarter in 2018, and its business model relies on finance capital and investor funding. They developed two main strategies for profitability, namely increasing subscription rates and advertising revenues, as well as seeking disintermediation by reducing its reliance on record labels by creating direct relationships with artists. Recently, there has been some controversies on Spotify's use of "fake artists", the artificial music creators, to reduce royalty payments to record labels. Thus, the company is faced with a challenge of

delicate interplay between the execution of its current strategy to catch the attention of investors and still maintaining record label support.

1.2 Music Industry Landscape

As previously mentioned, the music industry has undergone grand transformations, for example with the creation of music streaming platforms like Spotify in the years of 2000s. However, YouTube, initially a general video platform, has become the most utilized platform for streaming music. Recent industry data reveals that 47% of music streaming in Western Europe and North America occurs on video platforms, particularly YouTube, surpassing both paid and free audio platforms.

This leads to a statement that digital technologies have supplemented traditional means of music consumption. Users engage in digital listening through various devices and services, indicating a mix of technological generations. Moreover, the digital era has led to abundance, emphasizing large catalogs and the need for personalized recommendations by streaming platforms (Fricke, K. R., et al., 2019).

The traditional music industry, which is dominated by record labels, experienced a decline in revenues due to the shift towards digital consumption. Streaming services with freemium and premium models, emerged as key players, where platforms like Spotify continuously note a substantial increase in users (Barata et al., 2021).

The evolving landscape of the music industry and digital music consumption indicates a complex interplay of technological advancements and user preferences. As users navigate a world of abundant digital options, platforms like Spotify continue to shape how music fits into everyday life, challenging traditional norms and user behavior.

The music industry went through major transition, going from physical formats like vinyl records, cassette tapes, and compact discs (CDs) to the digitalized solutions such as MP3 and streaming services like Spotify and Apple Music (Xiaorui Guo, 2023). Additionally, the described industry started integrating artificial intelligence (AI), blockchain technology, and non-fungible tokens (NFTs). Before, artists success was defined by physical sales in the record stores and chart rankings. Everything was changed by the digital revolution with music encoding and storage, music sharing, and user-friendly platforms. The shift in revenue models to amount per stream creates questions about fair compensation for artists. It shows a transformative nature of the music industry's evolution, emphasizing the need for analysis and ongoing adaptation and innovation advancements in constantly evolving digital music industry.

Moreover, and for the purpose of this dissertation, it is crucial mentioning how music has a profound impact on cognition and emotions. The existing literature highlights certain limitations of the research conducted within the evolving music industry. Namely, many studies focus on specific consumer groups, such as music fans or young adults, overlooking the diverse relationships people have with music. Additionally, there is a lack of research on platform choice in online practices, with a predominant focus on audio platforms. Studies on musical preferences reveal that these preferences are influenced by various factors, including age, personality and values. Research has identified underlying dimensions such as Arousal, Valence, and Depth. Arousal captures the level of energy or intensity of the music track, whether the music causes excitement or calmness. Valence represents the emotional quality of music from positive to negative, indicating whether the sound is happy and uplifting or sad and melancholic. Depth on the other hand describes the complexity of musical experience, explaining how engaging the sound is, based on lyrics, structure and instrumentation, thus music with high depth provokes deeper reflection, while low depth is lighter regarding listening behavior. All three are discussed to provide insights into the psychological aspects of music listening behavior. With these gaps and opportunities in research there is a room for this dissertation and many more works in this area.

2. Literature on musical user preferences

2.1 The interplay between music traits and user preferences

There is a broad spectrum of musical engagement where people's lives are influenced across various contexts and situations (Duman et al., 2022). Music serves as a powerful mean for entertainment, relaxation, mood regulation, as well as social bonding. In this regard there is a crucial correlation between music features, such as tempo, energy, and danceability, and the reasons people listen to music, namely reasons for listening (RL). Moreover, the connection between musical features that induce movement, such as rhythm or tempo, and the RL has long been recognized, with dance often considered an integral part of musical expression. All these perspectives help gather insights into mechanisms driving music choices and consumption patterns.

Duman et al., 2022 have investigated how the complex interplay works between music features, RL, and movement, through a combination of surveys, audio feature extraction, and statistical analysis. They conducted two online surveys via the webropol.com platform with participants recruited via online platforms and university emailing lists. There were 105 respondents, including 61 women, 41 men, and 3 individuals identifying as other genders. The age demographics ranged between 16 and 54 years old, with a mean age of 27.07 and a standard deviation of 6.46 years old. Participants came from 19 different countries, with the majority being Finnish (N = 56), followed by Turkish (N = 23). Participants were asked to complete an online questionnaire indicating their general music listening habits with items such as background listening, emotional regulation, and social bonding, rating their agreement with each item on a 5-point Likert scale. Moreover, participants indicated a song that made them dance and the RL for this track. In the second survey participants chose 3 danceable songs indicating alongside their demographics. The analyzed Spotify audio traits considered energy, acousticness, danceability, valence, loudness, instrumentalness, speechiness, liveness, and tempo. The analysis conducted two-sample t-tests, seeking insights into Spotify audio features correlation with the baseline and dance music datasets.

As the results of the t-tests revealed, dance music tracks had significantly higher levels of energy, valence, and loudness, and lower levels of speechiness, instrumentalness, and acousticness compared to the other tracks. A further analysis, using k-means clustering, checked the homogeneity of dance music characteristics, and determined an optimal 5-cluster solution. The clusters were named based on their characteristic features: "happy-energy," "sad-energy," "sad-instrumental," "fast-lyrical," and "soft-acoustic.", with unique combinations of Spotify

audio features, reflecting the diverse nature of dance music preferences among participants. For example, the "sad energy" subgroup, characterized by low valence and high energy, included rock and metal songs, possibly reflecting the musical preferences of participants from regions with a strong metal culture. The "soft-acoustic" subgroup, featuring high acousticness and low energy, may be preferred for relaxation or intimate settings.

With the intention of identifying the reasons why people listen to danceable tracks, participants' RL general and dance ratings were compared, using data from Survey 1 with 100 dance songs and RL ratings. Additionally, they reduced the dimensionality of the data with factor analysis and identified underlying factors associated with RL general and dance. Researchers plotted RL dance item ratings for each subgroup of dance music as a heatmap to visualize the associations between RL and dance music subgroups, suggesting that specific subgenres may be more suited to certain RL. For instance, songs with the "happy energy" were associated with pleasure, mood enhancement, and social interaction, reflecting their upbeat and danceable nature, while songs in the "sad-energy" subgroup received lower ratings for dance-related activities, possibly due to their lower valence, tempo, and mood characteristics. These findings show the importance of considering both musical features and contextual factors when analyzing music preferences and user consumption behavior.

The most commonly indicated reasons for both RL general and dance were "For pleasure/entertainment" and "To improve my mood/raise energy." Moreover, RL dance ratings prioritized reasons such as "To dance/move along" and "To sing/play along." Factor analysis further clarified the three underlying factors associated with RL general and dance ratings: "Achievement of self-awareness," "Regulation of Arousal and Mood," and "Expression of Social Relatedness."

The t-tests gave many insights regarding the musical traits. There was a medium influence of energy, acousticness, danceability, valence, and loudness; a low effect was observed for speechiness. As it comes to liveness and tempo, there were no statistical differences, however visual inspection suggested potential differences. Notably, both datasets indicated average tempo around 120 bpm, while the dance music dataset showed less deviation from this average, indicating tighter clustering around 120 bpm, reflecting the optimal tempo for human motion.

There were 30 songs with the highest danceability scores indicated, providing insight into the types of songs participants chose to move to. These songs included various genres and eras, presenting diversity in musical preferences among participants.

By identifying subgroups of dance music and exploring their relationship with RL, the researchers contributed to a deeper understanding of how and why individuals engage with

music on Spotify. The authors suggested future research could further investigate the role of cultural, social, and individual factors in shaping music preferences and consumption behavior, ultimately enhancing our understanding of the psychological and sociocultural dynamics of music preferences.

2.2 The influence of music traits on its popularity

To measure musical song traits, Spotify undertakes a comprehensive set of audio features, categorizing the songs. The developers seek understanding of a track's composition, influencing the appeal to listeners. One of the traits is acousticness where high score suggests that the track has mainly acoustic instruments or arrangements (*Web API Reference | Spotify for Developers*, n.d.). Danceability assesses a track's suitability for dancing based on tempo, rhythm stability, beat strength, and overall regularity. Energy measures how fast-paced, loud, and dynamic the song is and Instrumentalness predicts whether a track contains instrumental elements. Furthermore, Loudness quantifies the overall volume level of a track in decibels (dB), Speechiness detects the presence of spoken words in a track, and Valence measures the musical positivity conveyed by a track.

The analysis of aforementioned traits through Web API Reference leads to a crucial understanding of the impact of the features on song's popularity. Spotify provides vast data on each track, including all the attributes mentioned above. These traits give insight into the musical characteristics that resonate with listeners and lead to a song popularity.

The analysis of the attributes and popularity correlation lead to a first conclusion: the year a song was released showed a high correlation with its popularity. This aligns with Spotify's algorithm that considers recently released streams when calculating a song's popularity score. Thus, the newer the song, the higher its popularity rating. Moreover, songs with high energy levels reflected a moderate correlation with popularity, suggesting that energy alone does not guarantee popularity. Other factors, such as lyrics, melody, and artist appeal, contribute to the success of the song. Acousticness showed a significant negative correlation with popularity, indicating the fewer acoustic elements, the more popularity the song will achieve. Such trend is clearly visible in the prevalence of electronic and produced music in today's mainstream culture. Further analysis through regression modeling could help developers to assess the predictive power of music traits on popularity. While energy emerged as a significant predictor, its influence on popularity was lower than anticipated. This suggests that while energy contributes to a song's appeal, additional factors may drive its popularity. While certain attributes like energy and release year correlate with popularity, they do not serve as the only

determinants. While music traits like energy, acousticness, and release year may influence a song's popularity, they represent only a fraction of the complex interplay of factors at play and therefore further analysis in this area is needed.

2.3 Spotify *versus* YouTube user traffic

It can be considered beneficial to undertake a comprehensive comparison of Spotify with the other commonly used music streaming platform - YouTube. As essential players in the digital music landscape, these platforms adjust their offerings to diverse user preferences with defined features that shape the music consumption of the consumers. While Spotify emerges as a default choice for casual music enthusiasts and a dedicated space for streaming, YouTube extends its influence beyond casual listening.

Considering the cross-platform usage of both Spotify and YouTube, users commonly combine YouTube with audio platforms like Spotify or Deezer, leveraging YouTube on computers and audio platforms on phones. The nature of open access of YouTube, without the need for login, contributes to its frequent use and sharing, even among those who do not use it frequently.

Users on YouTube trust the platform with the music selection for them, which is done by its recommendation algorithm, curating continuous, personalized playlists based on selected videos. Thus, a pivotal aspect of the comparison lies in the examination of recommendation algorithms employed by both platforms. While Spotify focuses on personalized playlists and recommendations within the platform, YouTube's algorithm takes center stage in soundtracking configurations. It suggests videos based on similarity to the selected video and user history, fostering a seamless and continuous listening experience.

Furthermore, user interactions play a crucial role in shaping the music consumption landscape. Spotify users actively engage in playlist creation, organization, and exploration of recommendations within the platform. On the contrary, YouTube users only appreciate the platform's ability to understand their taste, delegating the choice of music and enjoying the homogeneity of recommendations, creating an ambient background for various activities, not engaging in the organization. However, since there is a much easier process of uploading content on YouTube than on Spotify, YouTube brings more creative engagement.

Going through a comparative analysis, one should explore the catalog and content offerings of Spotify and YouTube. Spotify is known for its extensive catalog, organizational features, and offline capabilities, making it a preferred choice for music enthusiasts. Meanwhile, YouTube's appeal lies on its vast catalog, allowing users to explore niche and emerging artists or genres not easily available on licensed audio platforms, like Spotify. Thus, regarding platforms' roles

in music discovery and search functionalities, Spotify users rely on the platform for curated recommendations and diverse discovery experiences. In contrast, YouTube users praise its quality search engine, facilitating the discovery of specific genres, moods, or keywords. YouTube's associated videos contribute to a unique discovery mechanism.

While analyzing the sharing and accessibility features of Spotify and YouTube, there can be a crucial differentiation observed. While Spotify users predominantly share music within the platform, YouTube's open access and immediate availability without login requirements make it a popular choice for sharing music (Beuscart et al., 2022). Concluding the initial comparison, one should highlight the contextual preferences, user habits, and implications that shape the choices between Spotify and YouTube in diverse music consumption scenarios.

An important part of the comparison would be to see different traffic and engagement on the two platforms. A digital intelligence provider Similarweb, which traffic estimations are considered the most accurate and closely aligned to users' Google Analytics data, issued such comparison visible below. Regarding the global rank Spotify places itself far below YouTube, as the former ranks at 66th in January 2024 and the later as high as the 2nd website in the global rank (Similarweb, 2024). Tables 1 and 2, extracted from the forementioned website, represent the traffic and user engagement of both platforms over the period from November 2023 till January 2024.

Table 1: Spotify traffic

Total Visits	Bounce Rate	Pages per Visit	Avg Visit Duration
491.5M	36.03%	6.45	00:09:01

Source: spotify.com traffic analysis | Similarweb

Table 2: Youtube only music listening traffic.

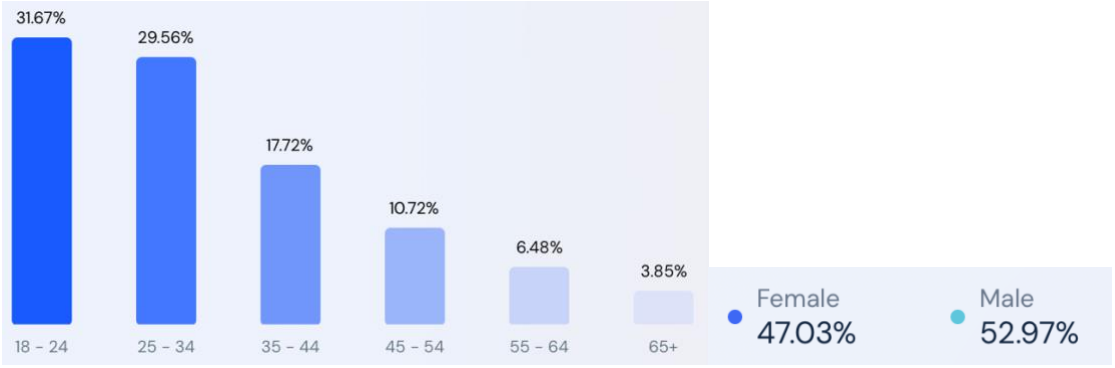
Total Visits	Bounce Rate	Pages per Visit	Avg Visit Duration
256.8M	28.16%	11.89	00:27:16

Source: music.youtube.com traffic analysis | Similarweb

It is possible to conclude that the musical part of YouTube has a much smaller audience, when compared to Spotify. Its user's engagement, however, is far of better quality, with average visit duration positioned more than double as much as in the case of Spotify and the percentage of bounce rate being far lower – Spotify with 36% and YouTube 28%.

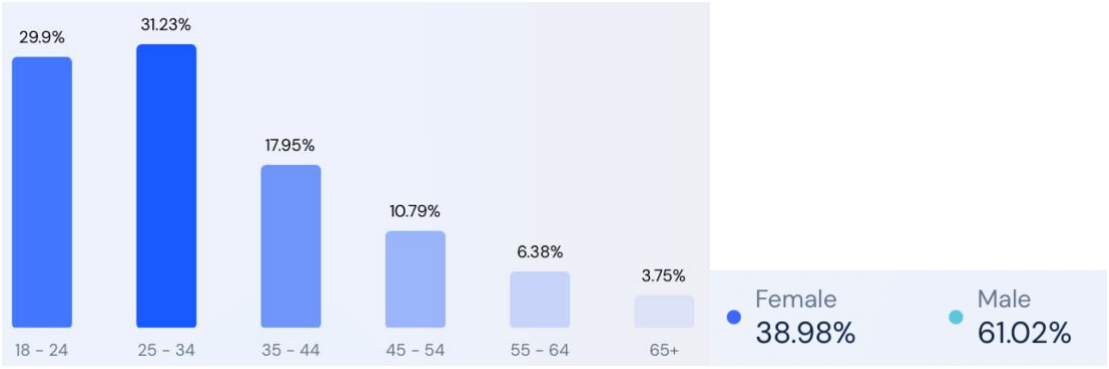
Furthermore, Similarweb analyzes the audience demographics, as displayed in graphs 1 and 2.

Graph 1 – Spotify gender and age distribution.



Source: spotify.com Traffic Comparison | Similarweb

Graph 2 – YouTube music listening gender and age distribution.



Source: music.youtube.com Traffic Comparison | Similarweb

Looking at the right side of the two graphs it is possible to conclude that Spotify is more balanced regarding gender distribution in comparison with YouTube, which has a higher percentage of male audience. When it comes to the age structure, Spotify mainly reaches younger generations with the18-24 age group as the largest part of the audience; While YouTube mainly targets the age group 25 to 34 years old and has a higher share of responses from all groups above 35 compared to the percentages of Spotify. To compare the two platforms' different marketing channels, Similarweb created graphs, 3 and 4.

Graph 3 – Spotify marketing channels distribution.

Source: spotify.com Traffic Comparison | Similarweb

Graph 4 – YouTube music marketing channels distribution.

Source: music.youtube.com Traffic Comparison | Similarweb

It is possible to identify a clear dominance of direct marketing on both platforms; And another clear trend is that there is higher activity of Spotify organic search and social marketing than in the case of YouTube.

In brief: the comparisons made above lead to the conclusion that the two platforms answer distinctive needs of their consumers and there is a need for a deep understanding of their music preferences. In this dissertation, the focus will be on Spotify as it holds a dominant role in shaping users' music consumption patterns. Unlike YouTube, where we notice wide range of type of content, Spotify tailors for music discovery and personalized listening experience. It's much stronger focus on organic search and social marketing channels signals how innovative, data driven and up to date this platform is, ensuring that the choice of Spotify will provide more valuable insights into evolving dynamics of user preferences.

2.4 The role of Artificial Intelligence and recommendation algorithms in user experience

Digitalization and online networks have created an abundance of choices for users, making recommender systems (RS) increasingly crucial (Wang, H., Wang, N., & Yeung, D., 2014). Companies like Amazon and Netflix employ RS to enhance user experience with personalized recommendations. There are three main methods: content-based, collaborative filtering (CF), and hybrid, which combines both approaches. While content-based methods face challenges in collecting user profiles due to privacy concerns, CF is based on past activities and thus has limitations like reduced accuracy with limited ratings and lack of ability to recommend new items. Because of these barriers, hybrid methods that incorporate supplementary information have gained popularity. The extent of interaction between rating information and additional information differentiates the hybrid methods as loosely or tightly coupled. Tightly coupled methods allow two-way interaction, automatically learning features and balancing the influence of ratings and content information. Tightly coupled methods demonstrate higher performance

in comparison to the loosely coupled. In the case of a recommender system, it is crucial to mention the deep learning models that have shown promise in learning effective recommendations. However, existing attempts of incorporating Deep Learning (DL), like using restricted Boltzmann machines or incorporating user-user and item-item correlations, lack the integration of content information crucial for accurate solutions. Collaborative Deep Learning (CDL) stands as a pioneering approach, emphasizing the importance of CDL for enhancing recommender systems in the age of online services. Considering the importance of accurate categorization of songs and recommending them to the right target audience, Artificial Intelligence (AI) is considered to become a powerful tool in this matter.

The development of AI has transformed music streaming, as the rest of the industries, shifting the way users engage with music. It leads to a need of exploring the impact of AI-based and music variables on customers' cognitive and emotional states that influence their behavioral responses. There are psychological dimensions of music that have been linked to customer engagement, and subsequent consumption behavior. Incorporating AI-driven technologies, such as facial and music recognition, may be used to enhance the customer experience. With its diverse solutions like data mining, machine learning, and sentiment analysis, AI truly has the potential to revolutionize customer analysis (Rodgers et al., 2021) and therefore holds a relevant part in this dissertation.

Rodgers et al., 2021 emphasizes the use of machine learning and deep learning in distinguishing patterns, like human cognitive processes, to capture customers' preferences and tailor their music-streaming experience accordingly. To understand the digital transformation effects of music on customers' perceptions and decisions, the authors used the throughput (TP) model that can illustrate the multifaceted impact of music. The likability and tempo of the music seem to have influence on cognition; AI-based facial and music biometrics show to mediate emotions, thereby shaping behavioral intentions. The authors enforced the statement of significant influences of music on customers' perceptions by conducting a study in a Chinese optical retail store. This topic can be easily connected to Spotify, as the analyzed platform utilizes many AI solutions.

One of the brilliant ideas that Spotify implemented is an issue of annual Wrapped that creates a summary of the music that the user listened to during this year. It reveals a small but significant amount of data like what genres, which artists did one listened to or how much time was spent on Spotify. This solution became so well-known, now Spotify users wait impatiently for the end of the year to receive their summary and moreover, share it with the world through different digital platforms like Instagram or Facebook. Thus, Spotify creates a community of listeners

that actively engage in sharing their experiences on the app. Spotify members use the yearly Wrapped to show their musical taste and distinguish themselves among their friends. This social phenomenon was described by Pierre Bourdieu (1987) who identified that people communicate to others who they are with the use of their cultural tastes (Bonini & Magaudda, 2024). As Bonini and Magaudda stated in their book „Platformed! How Streaming, Algorithms and Artificial Intelligence are Shaping Music Cultures", music consumption can be compared to fashion, as individuals seek acknowledgment of others through their clothing choices as well as their music preferences. The Wrapped feature underscores the value of datafication in Spotify. This concept describes the fact that music platforms, as any other digital platforms, track users' behaviors and turns them into data relevant for further functioning of the company. Wrapped shows that Spotify takes notice of any user choice, from the title and genre of the track, to the very second the user stops listening or changes to another song. Based on such analysis, platform builds detailed profiles and predicts with increasing accuracy future user consumption behaviors.

In the case of playlist curation, one of the tools Spotify utilizes is called PUMA that stands for Playlist Usage Monitoring and Analysis, and can track performance metrics, listener demographics, and engagement of the users (Prey, 2021). Such tool helps with recommendations and even with ordering the tracks in the playlist, as its flow is crucial in the user perception.

Spotify invests on advertisement of platform-curated playlists as they are more dynamic and easily modified so that they include user feedback and engagement data. The company holds great influence on artists ability to reach the audience as every decision in building recommendation algorithms affects it.

A European Commission study cited by Prey (2021) indicated that appearing on Spotify playlist like „Today's Top Hits" significantly boosts an artist revenue. Thus, musicians and record labels seek such an exposure, trying to create music in line with current trends and signing contracts with Spotify itself. This leads to a statement that musicians lost their control over how their work is presented and the platforms are the ones that create music experience, causing actual music creators to take a secondary role in the industry. Moreover, the playlist era diminishes the knowledge about specific artists, as playlists combine different creators, focusing more on the flow rather than the whole discography of one artist. This shift, from albums to playlists, indicates changes in music consumption and creates risks of competition for the musicians, where their work is only a component of a musical experience.

This data-driven approach to music extends into behaviorism and user conditioning as playlists become dynamic tools for shaping musical behavior and consumer engagement.

As the music industry expands, it becomes more and more difficult to categorize songs by genre, which greatly influences how customers create playlists and perceive music recommendations. Here AI can emerge as a solution as it can detect songs' bandwidth, frequency, or other traits, and based on that classify songs into genres (Ghildiyal & Sharma, 2021). There is a visible trend of utilization of such technology in recommendation algorithms where researchers test approaches like residual neural networks, convolutional neural networks (CNN), and Mel-Frequency Cepstral Coefficient (MFCC) analysis. A common dataset used for such tests is the GTZAN dataset that contains audio features from 10 different genres. Ghildiyal & Sharma (2021) in the pre-processing phase extract relevant audio features with time as well as frequency domain features. Then they used different statistical methods to evaluate predictor significance, to assess different feature combinations, and to automatically perform feature selection during model fitting. In genre classification they use models such as Artificial Neural Network (ANN), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Tree, and Logistic Regression. Authors measure the performance on a basis of confusion matrix, and calculate parameters such as sensitivity, specificity, and accuracy. The results lead to a statement that data filtering has a great impact on the overall accuracy of the genre classification model, thus concluding the role of AI in this task. The effectiveness can be even greater if the customer behavior data is constantly updated and immediately analyzed by AI.

2.5 Recent trends in Spotify and music streaming

The music streaming industry has gone through multiple transformations. Spotify has been, and continues to be, playing a transformative role in the music industry. With subscription-based payment models and pioneering algorithmic recommendation systems, Spotify has been leading technological innovations. Thus, analyzing the trends on this streaming platform can be considered a great reference for the whole music industry. It is crucial to mention a phenomenon discussed in previous chapters, namely platformization that was defined by Nieborg and Poell (Nieborg & Poell, 2018). These authors turned the world attention to the increasing cultural and social significance of online platforms with the rise of Spotify, along with other platforms, questioning their dominance and power. A significant change in Spotify was the ongoing focus on podcasts and other non-musical audio content, including audiobooks, live audio, and news (Håvard Kiberg & Hendrik Spilker, 2023). Spotify, as well as many streaming services, is becoming a comprehensive audiovisual platform. This strategic transformation began in 2015

with the podcasts that lead to exclusive distribution agreements with well-known podcasters like Joe Rogan and Michelle Obama in 2018. By February 2019, Spotify's CEO, Daniel Ek, declared the ambitious goal of becoming the "world's number one audio platform." The fact that usually there are neither managers nor record companies in the podcast industry, enables Spotify to directly collaborate with creators and maintain full control and ownership of the content it distributes, creating „Spotify Originals". Therefore, Spotify and many other platforms seek the transition into a role of content producer, as well as rightsholder, distributor, and promoter. This dynamic, in turn, raises concerns about the centralization of creativity and artistic production. Despite the skepticism, the forementioned strategy has proven financially successful, with podcast revenues alone rising by over 300%, generating nearly $200 million in 2021.

Spotify has also started to include not only human but also machine-generated audio content, causing controversies around "fake artists," royalty-free music, and AI-generated tracks. This expansion challenges traditional music creation and reinvents the role of technology in creative expression. As these artists have no presence outside Spotify, it is a cost-effective way for the platform to create on-demand mood-based playlists. There is also rise of AI-generated music to be noted. Hologram artists like Hatsune-Miku could challenge the traditional role of human creators, potentially minimizing their importance.

Moreover, Spotify has invested on live streaming, making concerts and various audiovisual content available. This shows a broader trend in the industry incorporating live experiences, blurring the lines between the online and offline world.

There is a widespread allegation that video content will be the next transformational point in the streaming industry, as recent developments suggest industry shift towards audiovisual content. Spotify featured virtual concerts during the pandemic and then corona concert series in the summer of 2021, featuring headliners like the Black Keys, Leon Bridges, and Girl in Red. There are also "vodcasts" (video-on-demand casts), video versions of podcasts. Some of the vodcasts are evolving into traditional talk shows, while others are transformed into film and television genres. Collaborations with major players in the film and television industries, such as Netflix or HBO suggest Spotify's potential entry into the broader entertainment market. All this means that in the streaming industry the boundaries between audio and visual content become thinner and thinner. Based on Spotify's shifts, one can detect a trend of moving from professionally produced music to its current emphasis on exclusive content, potentially overshadowing independent content creators. The evolving user preferences can be seen in the

implementation of live modes, particularly demonstrated by live-streaming concerts and applications like Spotify Live.

Here one can also examine how the company generates and repurposes user data for targeted marketing, framing it as a form of 'biopolitical marketing' (Till, 2023) - this will be one of the investigations of Spotify's current innovative strategy. Year 2023 brought increasing interest in health and wellness, with a development of both generic and personalized playlists designed to capture the attention of users that engage in physical activities. It may be relevant for advertisers to understand how Spotify transforms users' health-related activities into 'audience commodities'.

Explaining the concept of 'biopolitical marketing,' Spotify actively engages in shaping users' health, exercise, and wellness practices, including the 'deep mediatization' of users' experiences, aligning with the broader trend of datafication and financialization of everyday life through digital media. This phenomenon means that social media and search platforms commodify everyday activities, framing them as natural resources for data extraction. Such activity is emphasized by Spotify's claim that it can harness the affective experiences of health, exercise, and wellness practices in real-time to generate valuable data. While technologies often promote neoliberal subjectivities through 'chrono normative' orientations, Spotify's case represents a unique form of 'governing through rhythm.' The platform utilizes digital health and exercise monitoring to modulate controls over users' periods of time, particularly non-work time, sleep, and exercise, which are used for advertising purposes.

Another trend to be noticed is Spotify's tracking of emotions and moods, leading to sentiment analysis and its potential for generating psychological insights. The platform's emphasis on intimacy and subjective investment in neoliberal ideology, framed through playlists as tools for 'mood management'. Here we can identify a trend that individuals' music preferences are influenced by emotions, personality, geographical location, weather and many more factors. Moreover, Spotify's business model has evolved to centralize user engagement through a focus on advertising and integration with 'financialized' capitalism. Users are construed as 'scalable subjects,' whose experiences are transformed into 'affective publics'—groups organized based on emotional states. These groups, subject to modulating controls of affective states, become 'living multiplicities' and contribute to the 'financialization' of users' experiences, which are presented as compelling narratives for potential investors.

Continuing the topic of trends, Spotify stands out from other digital platforms by adopting a unique 'ubiquity strategy' instead of converging into various areas. The platform prioritizes depth, integrating seamlessly into users' lives to offer a seamless experience. This deep

mediatization aligns with the overarching trend in digital capitalism to encourage 'chronomormativity' and functional use of music in daily activities.

Lately, as the subscription model struggles to produce profits, Spotify increasingly relies on programmatic advertising. This approach involves auctioning individual user interactions to advertisers in real-time. Spotify's advertising infrastructure targets 'Addressable Moments,' such as the popular 'Running' moment, aligning with users' health and wellness activities. These moments may be more than valuable opportunities for advertisers to engage users with relevant and personalized content.

Spotify's 'Moments' strategy focuses on users' rhythms and tempos, leveraging their activities and emotions for targeted advertising. The platform uses patent applications to automate the identification of these moments, incorporating data on repetitive motion activity, cadence, and heart rate. This enhances the overall user experience and making advertising more effective.

Spotify's interest in speech recognition to investigate emotional states further underscores its commitment to understanding and utilizing users' moods for targeted advertising.

Another recent application of Spotify as a streaming service is as a support of social causes. An example worth mentioning is a situation of 31-year-old American singer that in 2021 was diagnosed with sarcoma tumor (Today, 2024). Cat Janice won few awards over the years and even while fighting the disease, she kept creating upbeat and full of spirit songs. In the beginning of the year 2024 she shared on her social media accounts that her cancer has progressed and made it impossible for her to speak. In response, Cat undertook formalities to ensure that proceeds of her newly released songs streams will go directly to her son Loren. The track is called “Dance You Outta My Head” and few days after release the track charted on iTunes on the 7th spot worldwide with number 1st in Romania, Ireland and the Czech Republic and number 12th in the USA.

Concluding this chapter, it is crucial to state that by commodifying moments, emotions, and personality traits, Spotify positions itself at the forefront of the evolving landscape of digital capitalism, where user data becomes a valuable resource and leads to more and more accurate targeting.

3. Identifying recent relevant trends in Spotify user behavior

3.1 The Spotify datasets and variables under analysis

Spotify datasets publicly available on a platform called Kaggle are used in this dissertation. Four separate datasets are considered: by year (Spotify dataset), by country (Top Spotify Songs in 73 Countries (Daily Updated)), by weather (6k Weather Labeled Spotify Songs) and focused only on Taylor Swift songs (Taylor Swift Spotify Dataset). Several variables were common to the four datasets. The main variables used represent a musical attribute, ranging from the technical aspects like tempo and key – a musical note of the track - to more nuanced elements such as danceability and valence. These variables were used in different perspectives, allowing to understand the diverse reasons for a user to choose a music track or to discover layers of information that contribute to the holistic understanding of user preferences. From the identification of artists and track specifics to the subtle nuances of accousticness and liveness, the below set of variables explanation is the basis for understanding subsequent analyses (Maharshi Pandya, 2022).

- **track_id:** The unique Spotify identifier for the track, facilitating easy retrieval and reference.
- **artists:** The names of the artists who performed the track, separated by semicolons in case of multiple contributors.
- **album_name:** The title of the album in which the track is featured, providing contextual information about its release.
- **track_name:** The specific name of the track, identifying it within the artist's discography.
- **popularity:** A numerical value between 0 and 100, calculated algorithmically based on the track's total plays and recency. It serves as a measure of the track's current popularity and will be used in this dissertation as a reference for user preference.
- **duration_ms:** The length of the track in milliseconds, indicating its temporal extent.
- **explicit:** A Boolean variable indicating whether the track contains explicit lyrics. 'True' signifies explicit content, 'False' indicates no explicit content, or 'Unknown' if information is not available.
- **danceability:** A metric ranging from 0 (least danceable) to 1 (most danceable), assessing the track's suitability for dancing based on various musical elements.

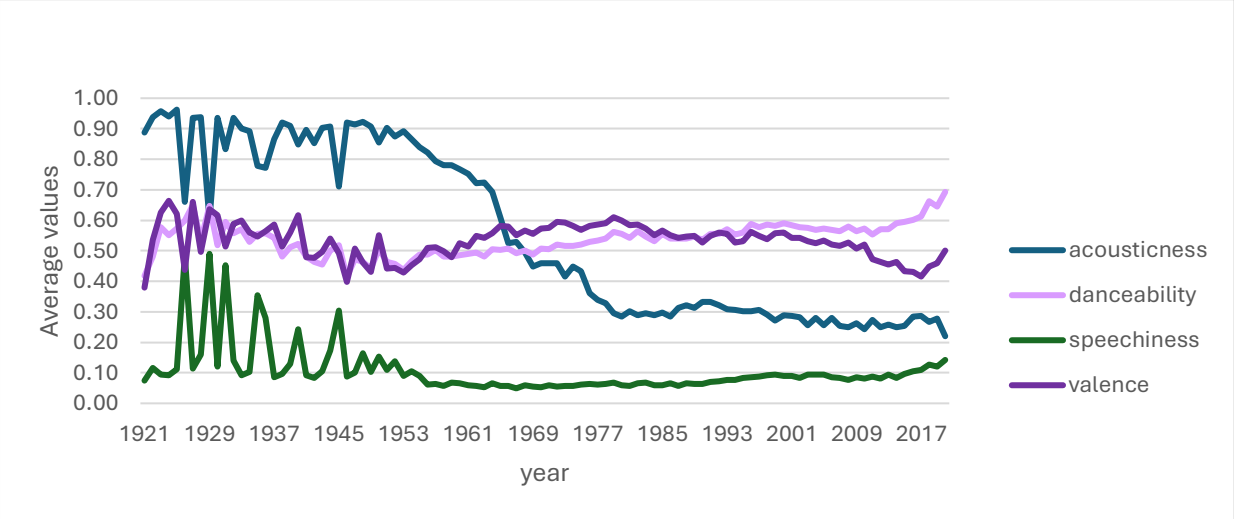
- **energy:** A measure from 0 to 1 reflecting the track's perceived intensity and activity. High energy values suggest a lively and dynamic sound.
- **key:** An integer representing the key to the track, mapping to standard Pitch Class notation. A value of -1 indicates no detected key.
- **loudness:** The overall loudness of the track in decibels (dB), influencing its perceived volume and intensity.
- **mode:** Indicates whether the track is in a major (1) or minor (0) modality, revealing the type of scale from which its melodic content is derived.
- **speechiness:** A metric ranging from 0 to 1, identifying the presence of spoken words in the track. Higher values suggest a more speech-like composition.
- **acousticness:** A confidence measure from 0 to 1, indicating the likelihood of the track being acoustic. Higher values signify a higher confidence in the acoustic nature of the track.
- **instrumentalness:** A measure predicting the presence of vocals in the track. Higher values (closer to 1) suggest a greater likelihood of the track being instrumental.
- **liveness:** Detects the presence of an audience in the recording. Higher values indicate a higher probability that the track was performed live.
- **valence:** A measure from 0 to 1 describing the musical positiveness conveyed by the track. Higher values indicate a more positive emotional tone.
- **tempo:** The estimated tempo of the track in beats per minute (BPM), reflecting its speed or pace.
- **time_signature:** An estimated time signature, indicating the number of beats in each bar or measure.
- **track_genre:** The genre in which the track belongs, providing information about its stylistic characteristics.

3.2 Identifying user trends over the decades using Spotify dataset by year

Seeking identification of possible trends in the user musical preferences over the years, the publicly available dataset (Spotify dataset) extracted from Spotify with all the variables mentioned before, sorted by year when the song was released was used. The variables include the level of each song characteristic summarized from the songs of a particular year. There are 100 inputs available, ranging from year 1921 to 2020. As generations change, so do music tastes, which becomes evident from the statistical data analysis. Chart 1 displays the average

values of four variables, namely acousticness, danceability, speechiness and valence, over the years 1921 to 2020. These four dimensions were selected to be represented in the same chart, since they share a common measurement scale. Thus, it is possible to observe changes over time of some of the most shifting variables that describe the proportion of spoken words, acoustic sounds, track's suitability for dancing and overall positiveness of the sound.

Chart 1 - Line chart, with average values of accousticness, danceability, speechiness and valence, by year.



Source: Authors' analysis based on Kaggle dataset "Spotify dataset"

A first look at the chart gives a direct observation that all four chosen music characteristics differentiate more year by year in the period between 1921 and 1950, as there are many visible spikes to be noted. Various historical events, societal changes and cultural shifts may be perceived as an explanation to this observation, as such life-changing events naturally lead to different preferences of consumers, including their music consumption. Moreover, it was a time of relevant developments regarding musical instruments, such as electric guitar in the 1930s and recording techniques, namely multitrack recording in the late 1940s or magnetic tape in the 1930s (Martin, R., 2023, November 28). Continuing the analysis, it can be noticed that the most substantial decrease is in the case of acousticness, 0.9 in 1925 to 0.2 in 2020. This phenomenon may indicate a shift from traditional instrumentation to a greater use of electronic and synthesized sounds in music production. Such a trend can be easily explained by technological advances in music recording over the years, reducing the reliance on purely acoustic instruments. Moreover, this decrease might be caused by changes in the popularity of genres over time, meaning that electronic and pop genres gained popularity, while classical or folk genres lost their relevance.

Concerning danceability of the tracks, it is noticeable that there was a preference for songs suitable for dancing in the 1920s with a decrease until the year 1953. After that, the average

value of danceability increases steadily, achieving its peak in 2020, suggesting future growth. The year 1920 is commonly known as the swing and jazz era with its upbeat and rhythmic nature, where dancehalls were highly popular and the demand for danceable tracks was high, explaining the peak of danceability. Then, the post-war transition led to less upbeat track, introducing early rock and blues that are not characterized by their danceability as much as swing in the previous era. However, after 1953 one can identify a steady increase, which can be explained by the rise of rock and roll era, a genre known for its energetic and danceable qualities with artists like Elvis Presley and Chuck Berry (Kot, G., 2023, December 1). Technological advances had their influence since the music production technology was, and still is, gaining new possibilities to create more danceable beats. The peak in 2020 could be attributed to the diversity of dance genres, caused by globalization, including pop and hip-hop. The emphasis on danceability in music consumption can also reflect a global desire for joy and self-expression of the consumers.

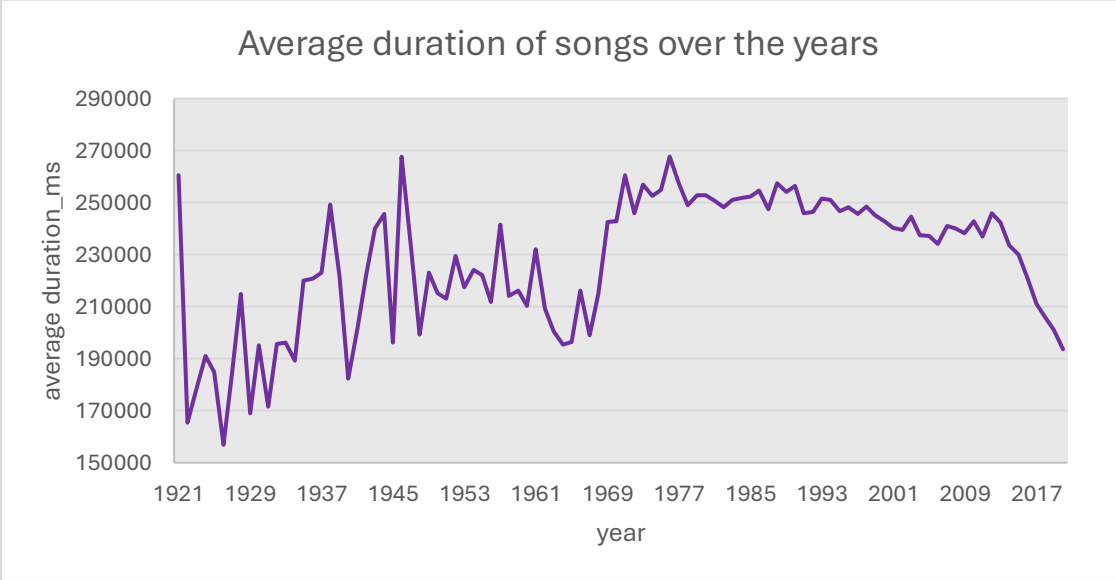
Analysis of the speechiness variable doesn't show any incremental changes over past decades, but this characteristic is visibly lower, on average values, in recent years than in the years before 1950s. The observation may be explained by the fact that before, music style involved clear enunciation and prominence of lyrics, where vocalists emphasized storytelling. The lyrical content of the songs was often a pivotal point of the song creation. After the 1950s, however, the increasing use of synthesizers, electronic instruments and advanced recording techniques lead to a decrease in the relevance of actual words in the tracks.

Lastly, the average valence of the tracks visibly fit within the values 0.4 to 0.6, indicating no big differentiation over the years. The variable had the lowest average values in the years 1920, 1945 and then around 2017-2020. This range suggests a moderate level of positiveness without extreme emotional highs or lows of the tracks, which can be caused by the fact that listeners find various types of valences relatable, flattening the curve. The noticeable decrease in the average valence values in 1920 and 1945 are understandable in the context of post war emotions. This period could be characterized by complex mix of relief, reflection, and rebuilding, leading to a lower valence in song production and preferences. On the other hand, 2017-2020 are the years of many global political, social, and public health issues that can be conveyed in musical emotional tone. Musicians often respond to the social and political climate and listeners expect them to do so.

Furthermore, in the matter of the analysis by year, it may be interesting to see how the duration of the tracks changed over the decades. The song duration is not merely a technical attribute, but often a well thought through choice made by musicians and producers. Analyzing this

dimension may provide information on the impact of technological developments on music production and consumption. In the past, the physical formats, such as vinyl records and cassette tapes, limited the possible duration of songs. With time, streaming platforms, like Spotify democratized music distribution, extending the ability to create longer tracks. Moreover, such analysis provides insights into customer preferences and consumption habits, considering changing attention spans that influence the duration of songs. Chart 2 illustrates the changes of average song duration over the years.

Chart 2 – Line chart with average values of song duration over the years, in milliseconds (ms).



Source: Authors' analysis based on Kaggle dataset "Spotify dataset"

There were visible variations in average track duration in milliseconds (ms) over the period from 1920 till 1969, ranging from as low as 155000ms in 1926 to as high as 270000ms in 1948. Such observation may be influenced by experimentation and diversity in musical styles and formats in these years, indicating a dynamic and evolving industry. There is a trend of longer average durations of songs in the 1970s, with visible peaks in 1970 and 1977, depicting the rise of rock, concept albums, and extended musical compositions. There is a moment of stable average song duration from the late 1970s to the early 2010s, around 240000ms. The stabilization may indicate the standardization of music formats like vinyl records or CDs and radio broadcasting practices. After that there is a clear shift in 2013, with a relevant decrease in track duration. By 2020, the average track duration has decreased to 190000ms. Such phenomenon may be explained by various factors with changes in audience attention spans, development of streaming platforms favoring shorter songs, and with evolving production and distribution methods. With all this knowledge, the music streaming industry can better forecast the audience preferences and market demands, and therefore adjust their services accordingly.

3.3 Identifying user geographical trends using Spotify dataset by country

This section of the dissertation uses the publicly available dataset - Spotify Dataset, 2021 - depicting the top 50 songs for each country, considering 73 different countries (Top Spotify Songs in 73 Countries (Daily Updated)). This dataset contains 320 675 inputs of Spotify songs. An analysis by country may hold significant relevance in the context of core-periphery dynamics and geographical inequality within the music industry (Tófalvy et al., 2021). The existence of a core-periphery structure implies that in the spatial dimension (space and place), the socioeconomic development is usually uneven (Klimczuk & Klimczuk-Kochańska, 2019). It has been widely established that location is considered an influential factor in the music industry's economic and its marketing approaches. Thus, global music consumption is shaped by geographical inequalities. One of the most influential centers are global cities, that play a pivotal role in creative and cultural industries. From these centers operate major record labels, such as Sony, Universal, and Warner, controlling a substantial share of music records revenues. The flow of music distribution tends to be directed in only one direction, with content coming from these centers, reaching and influencing other parts. Therefore, a hierarchical structure is created and music production companies not within the centers may struggle to gain visibility. Each genre exhibits its own set of inequalities and center-periphery dynamics. For instance, Metal, originating in the United Kingdom in the 1960s, reflects both Anglo-Saxon dominance and unique socio-cultural trajectories. Nordic countries like Finland, Norway and Sweden have emerged as global centers for metal music, significantly influencing this genre's evolution (DeHart, 2018; Maguire, 2015). While digital technologies were expected to help in the diffusion of music, researchers question this hypothesis, stating that online networks can in fact reinforce geographical inequalities. One of the goals of this dissertation is to examine whether the Spotify's platform ecosystem, powered by algorithms and recommendation systems, follows or challenges existing core-periphery structures in the music industry. Understanding these dynamics is essential for the future existence of streaming platforms in the evolving digitalized landscape of the music industry.

Moreover, in the long tail economy of streaming, which on assumption offers limitless options, users still meet with what is known as the "tyranny of choice" (Tófalvy et al., 2021). Rather than facilitating the discovery of niche and less popular content, the long tail appears to favor the fame of superstars in the music industry, creating a „superstar effect". This effect has been a defining characteristic of the music industry, where a small percentage of popular performers commands a disproportionate share of attention and revenue. For reference, in 2007, 1% of

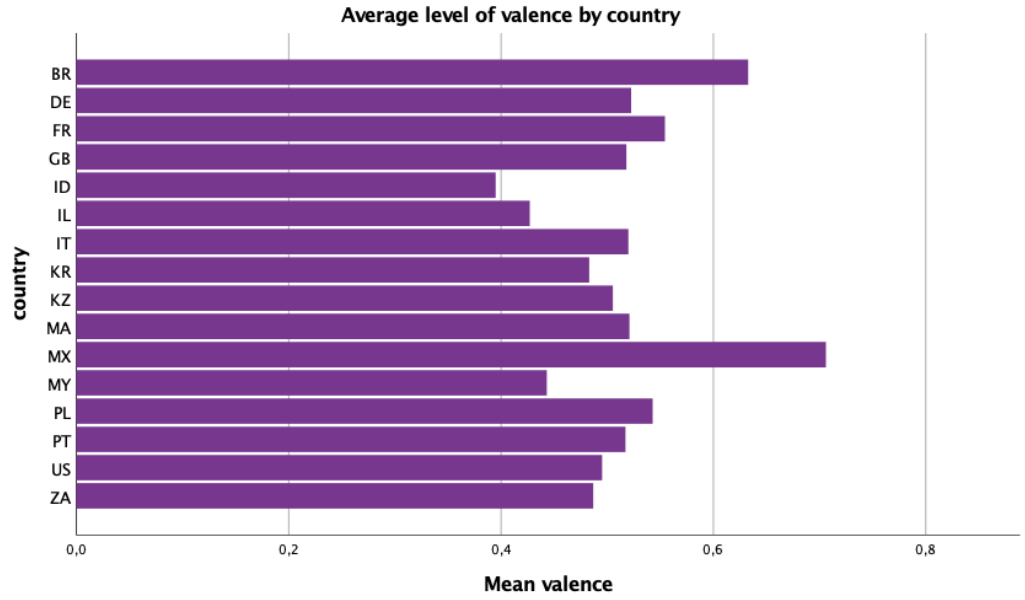
digital tracks accounted for 80% of all track sales, and 1,000 albums comprised 50% of album sales (Celma, 2010). And the superstar effect was not decreased and was possibly elevated by the transition to online music distribution.

Examining data available from Spotify reveals similar trends, where approximately 10% of music streams belong to the top 1% of artists on the platform, indicating a concentration of listenership towards a select few (Blake, 2020). Moreover, Spotify's playlists exhibit a bias towards major label content, aligning with the broader superstar effect in the industry. Another important issue in the superstar effect is the algorithmic bias, where popular items tend to receive better positioning due to the feedback loop in collaborative filtering that led to recommendations being influenced by user interactions. This creates a Matthew-effect, favoring already popular creators and hindering those with less visibility. These mechanisms reinforce existing offline core-periphery inequalities within the music industry.

To identify cultural music preferences and to compare the different top songs in different countries, one could start with selecting cases with daily rank below or equal to 5, so that the analysis focuses only on the top 5 songs in the country. After that, to visualize the top songs it would be crucial to create a clustered bar of daily rank by country with a filter of artists. However, different configurations of top songs analysis showed too big of a discrepancy within countries, making it difficult to identify any correlations.

Another interesting case would be to compare various song characteristics differentiating among countries based on charts. Let us for example consider valence, the musical positiveness of the track, which can indicate country's population expectations from music in the matter of overall emotional tone. Thus, it may be a reference for marketers and music streaming platforms, concerning the tone on which to convey their messages and music recommendations. Chart 3 displays average values of valence, focusing on a few chosen countries: Brazil, Germany, France, United Kingdom, Indonesia, Israel, Italy, South Korea, Kazakhstan, Morocco, Mexico, Malaysia, Poland, Portugal, United States, and South Africa. The reasoning behind such a selection is to include countries with most outstanding values, the ones with big impact on the world like United States, United Kingdom or Germany and countries that can be considered interesting for the analysis.

Chart 3 – Using bars to display average values of valence, by country.

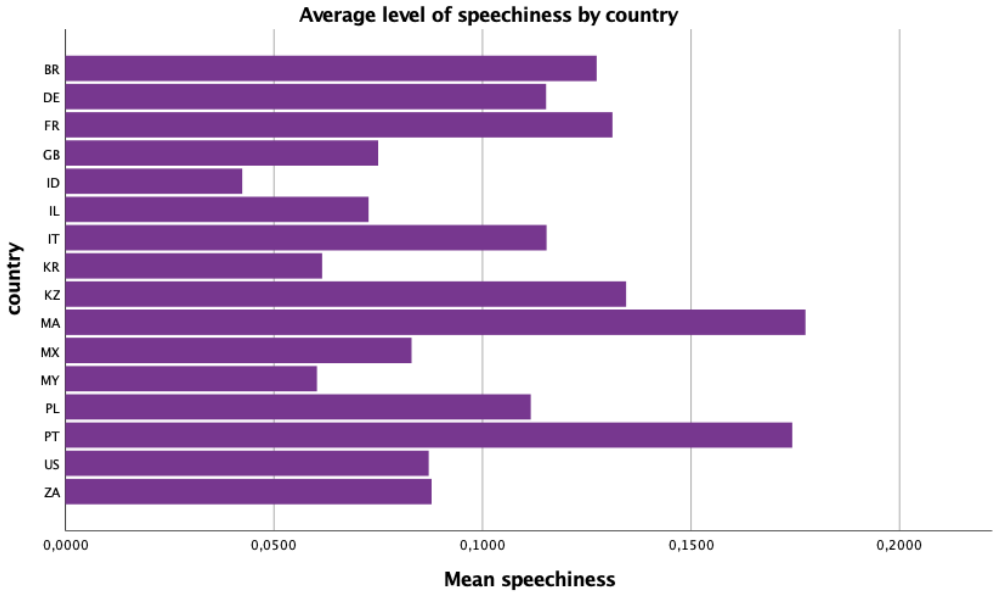


Source: Authors' analysis based on Kaggle dataset "Top Spotify Songs in 73 Countries"

The data analyst attention goes first to Brazil (BR) and Mexico (MX), with the highest average values of valence in the songs listened by their Spotify users. This can lead to the statement that Mexicans, even more that Brazilians, tend to favor positive and cheerful tracks in their music usage, which aligns with the vibrant and lively nature often associated with these countries. On the opposite side are Indonesia (ID), Israel (IL) and Malaysia (MY), showing lower average valence, suggesting that their music consumption focuses on more nostalgic and slow rhythms. An intuitive factor for Israel could be the recent horrific events of the ongoing war there that surely has relevant influence on music choices in the accounts registered in this country. Regarding Indonesia and Malaysia, communal activities and events may favor music with a more contemplative or emotionally nuanced atmosphere. Moreover, Malaysian traditional gamelan music may contribute to a preference for lower-valence music. The other countries reflect preferences for more balanced emotional content with average valence values around 0.5. South Africa (ZA), Morocco (MA), Kazakhstan (KZ) with their various cultural incoming influences, causing wide range of values that flatten the curve. Polish (PL) and French (FR) music consumers appear to prefer slightly more positive emotional content than other European countries included in this analysis.

Chart 4 shows the average speechiness of the most-listened tracks across the same countries, reflecting the number of spoken words in the music.

Chart 4 – using bars to represent average values of speechiness, by country.



Source: Authors' analysis based on Kaggle dataset "Top Spotify Songs in 73 Countries"

The highest average values are indicated by Morocco (MA) and Portugal (PT). This higher preference for songs containing more words can be caused by cultural factors, such as favoring traditional music styles with spoken words. In Morocco, genres like Raï or Gnawa, or in Portugal fado, highly focus on storytelling in their songs. France (FR) and Kazakhstan (KZ) also have relatively high average speechiness values, with possible influence of *chanson française* or French song tradition, which often features narrative elements and may contribute to the higher speechiness, similarly in Kazakhstan through folk songs or oral poetry. Intermediate speechiness is displayed in the case of Germany (DE), Israel (IL), United States (US), Italy (IT) and Poland (PL) who focus on poetic lyrics and storytelling with folk music traditions. In Germany, genres like spoken word, rap, or hip-hop may contribute to a moderate level of speechiness. In Israel on the other hand, a diverse musical landscape influenced by various cultures and in the United States, the prevalence of rap and hip-hop.

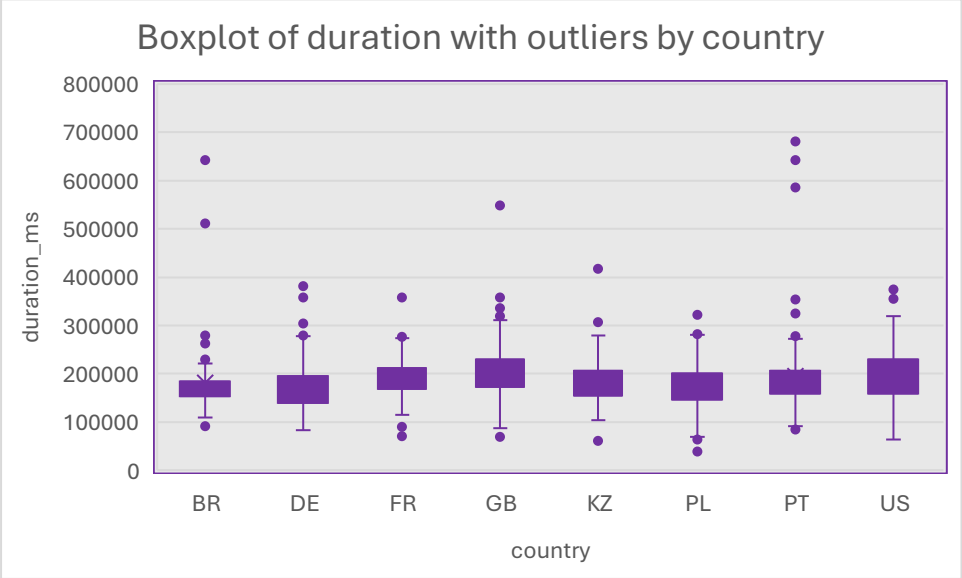
The countries that show lower speechiness are Brazil (BR), the United Kingdom (GB), Indonesia (ID), South Korea (KR), Malaysia (MY), Mexico (MX), and South Africa (ZA) with a greater emphasis on instrumental or melodic aspects. This diversity in the values of speechiness among countries reflects musical preferences present in each region and the fact that it's crucial to take such dimension into consideration when constructing algorithms.

When analyzing Spotify's consumer behavior in Poland in 2023, one could observe dynamic changes that reveal new light on music preferences and trends dominating the local music industry (Cieślak, 2023). The popularity of music by local artists in Poland increased in the year

under review. Data shows that Poles streamed 18% more songs by local artists compared to the previous year. Mateusz Smółka, Music Team Lead at Spotify, highlights an important indicator for the development of the home music scene. Furthermore, the Polish market has seen a consistently high popularity of rap with artist Taco Hemingway as the most streamed artist in the country, tied with the club2020 project and his solo album '1-800-ŚWIECENIE'. The seamlessness of musical genres and diversity in music production and reception can be observed. An interesting phenomenon is the inspiration of Polish artists with global trends, such as the Jersey Club sound, which is reflected in songs such as 'Malibu Barbie' from club2020. Additionally, there is a revival of pop punk guitar sounds, which are increasingly used in Polish rap and pop music. There is also a noticeable trend of an increasing number of tracks representing different music genres and artists, both local and global. Polish listeners are choosing music more and more depending on their mood or moment of the year, which is reflected in the popularity of themed playlists, such as May Day or Halloween. Polish Spotify users are selecting music depending on the occasion, such as the elections, when the song 'I Love Freedom' by Chłopcy z Placu Broni hit the playlist the day after the elections, which may indicate the expanding role of music in shaping local identity and community participation. All this suggests that the algorithm should assume moments of the year and important events in the country to recommend relevant songs at given moments.

Another fundamental characteristic of a song and the one that can vary greatly between different countries is the duration of a song. Analyzing this variable can offer relevant insights into national musical preferences and trends of listeners' engagement and artistic expression. Chart 5 uses boxplots to display the distribution of song duration, by country, suggesting regional variations of song duration, as different countries have different musical genres, and preferences regarding song structure. For instance, some cultures may favor longer, more elaborate compositions, whilst other countries may prefer shorter, more concise tracks. Such analysis can be of use for artists and music producers, while tailoring their content to specific geographical segments. Moreover, streaming platforms such as Spotify can leverage insights from this observation in their recommendation algorithms and programming strategies to customize services according to the country they target.

Chart 5 – Boxplots with the distribution of song duration (in ms), by countries.



Source: Authors' analysis based on Kaggle dataset "Top Spotify Songs in 73 Countries"

From chart 5, and concerning the distribution of song durations by country, it is possible to conclude that Brazil (BR) has a relatively wide interquartile range (IQR) of song duration, with the 50% central observations ranging from approximately 180000 to 190000 milliseconds (ms); outliers at 500000 ms and 650000 ms suggest Brazilians favor occasional longer, extended tracks. This phenomenon can be explained by Brazil's diverse genres such as samba, bossa nova, and MPB (*Música Popular Brasileira*), expanding the range of song duration. A similar pattern can be observed in the case of Portugal (PT) with the 50% central observations ranging from 150000 to 200000 ms, and even further away outliers' values of 600000 ms, 650000 ms, and 700000 ms, depicting exceptionally long tracks. Both countries share similar music admiration, and this observation can help algorithms, suggesting possible demand for tracks longer in duration. On the other hand, Germany (DE) shows a narrower IQR, spanning from 120,000 to 200,000 ms with trending their electronic and techno songs that exhibit similar lengths of tracks. France (FR) shows a tendency towards consistency in the duration as well, possibly reflecting cultural norms or industry standards. The United Kingdom (GB) has a wider IQR and visible outliers at 550,000 ms, which may be due to experimental musical projects and musical innovation happening in this country. Kazakhstan (KZ) and Poland (PL) display a common range of song durations. United States (US), with a wide IQR, exhibit diverse musical preferences in the country.

All of the above gives insight into different geographical user preferences and shows evidence that such findings are crucial in targeting Spotify users.

3.4 Identifying user preferences on various weather conditions using Spotify weather dataset

Another relevant investigation is the relationship between weather conditions and music preferences. In their analysis, Anglada-Tort defined music consumption patterns in relation to weather variables (temperature, sunshine, rainfall) (Anglada-Tort et al., 2023). This research explored the non-musical domains that link weather conditions to mood and behavior, providing a foundation for the hypothesis that weather holds an influence on music consumption. The research analyzed music features using music information retrieval techniques and machine learning models and weather variables like temperature, sunshine, or rainfall to investigate the associations between them.

There is a potential limitation of such study: the lack of commercial factors or recommendation algorithms. However, it concludes there is a significant correlation between weather conditions and consumer music choices.

3.4.1 Weather conditions and songs' popularity

From the publicly available dataset that presents songs listened by Spotify users on specific weather conditions (write here the exact name of the dataset used), paired with the dataset of song traits influencing song popularity, it is possible to observe which traits users seek when considering specific weather conditions. After merging the two previously mentioned datasets, 2100 observations were obtained. Described conditions are as follows: clear, clouds, drizzle, fog, mist, rain, snow and thunderstorm. The frequencies of the weather conditions were considered; due to the small number of observations in some of them, the decision was to combine similar weather conditions – drizzle, fog, and mist. Thus, 6 weather groups were analyzed by traits, with the objective of identifying the traits that characterize the most commonly listened songs on specific weather occasions.

In order to understand how weather conditions affect music preferences, and which song characteristics become more or less important depending on the weather, a multiple regression was conducted using SPSS, to examine which factors related to the characteristics of songs available on Spotify have a statistically significant effect on their popularity, under different weather conditions. The 9 factors analyzed are: danceability, energy, volume, word count, acoustic, instrumentality, longevity, value, and tempo.

The estimated regression coefficients are presented in the last column of Table 4 (in the annex). Table 3 indicates the song traits that have statistically significant influence on song popularity – positive (+) or negative (-).

Table 3 – Song traits that have statistically significant influence on popularity

SIGNIFICANT INFLUENCE ON POPULARITY	WEATHER CONDITION DIFFERENTIATION						NO WEATHER CONDITION DIFFEREN-TIATION
FEATURES	CLEAR	CLOUDS	DRIZZLE/ FOG/ MIST	RAIN	SNOW	THUNDER STORMS	
DANCEABILITY		+		-			
ENERGY			-	-		-	-
LOUDNESS			+	+		+	+
SPEECHINESS			-			+	-
ACOUSTICNESS							
INSTRUMENTALNESS	-	-	-	-	-	-	-
LIVENESS							
VALENCE		-					
TEMPO						-	

Source: Authors' analysis based on Kaggle dataset "6k Weather Labeled Spotify Songs"

From Table 3, it is possible to conclude that several traits have significant influence on the popularity, in the case of weather conditioned music preferences. Firstly, looking at the case without weather differentiation (Last column of the table), one can observe that energy, speechiness and instrumentalness have negative correlation, meaning that in general, when not taking into account weather conditions, when these traits rise, the popularity decreases. Following this observation, it can be assumed that users overall prefer less energetic and less instrumental songs with fewer spoken words. As it comes to loudness, it is a statistically significant trait that has a positive influence on popularity, suggesting users favor louder songs. To understand how different weather conditions influence the impact of song traits on popularity, separate regression analysis was conducted, having popularity as the dependent variable and all the 9 traits (danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence, and tempo) as explanatory variables. Estimated regression coefficients are displayed in columns 2 to 7 of Table 4 (in the annex). Table 3 (columns 2 to 7) highlights the song characteristics that are relevant in determining the popularity of the track, under specific weather condition.

From the regression analyses it is possible to conclude that clear and snowy weather conditions and song traits - acousticness and liveness – do not influence popularity. On the other hand, instrumentalness influences popularity in all weather conditions.

One interesting observation is that rainy days had the most observations, possibly suggesting that users are most likely to listen to music under these conditions. A hypothesis could be that rainy days are conducive to staying indoors and seeking indoor activities, such as listening to music. Rain may affect people's moods, causing them to seek out music that enhances their mood or aligns with their emotions.

Under clear sky conditions, the only factor that had a statistically significant effect on song popularity was instrumentality. A possible explanation is that when the weather is clear, people prefer calmer, less engaging music that can act as background music during various outdoor activities.

On a cloudy day, the traits that influence the popularity of a song are: danceability, instrumentalness and valence. The increase in danceability suggests that people may be looking for music that will improve their mood. Instrumentalness continues to play a role, and valence, which refers to the positive nature of the songs, may be sought as a contrast to the gray surroundings. Valence appears only here, and not in any other weather condition. Under neutral or positive conditions (such as clear skies), the need to lift the mood with music may be lower, hence value does not become a factor in popularity. On rainy days, on the other hand, more dynamic musical qualities (such as danceability, energy) may be more important in boosting mood, which explains the lower importance of valence under these weather conditions.

When drizzle, fog or mist appears, weather affects a greater number of factors: energy, loudness, speechiness and instrumentalness. The increased energy and loudness may be due to the need to stimulate and sustain attention in less favorable conditions. The speechiness may suggest that people prefer more narrative songs that can distract from adverse weather.

Rainy days show the influence of danceability, energy, loudness and instrumentality. Danceability may suggest a need for mood enhancement, energy and loudness are responsible for more dynamic songs that can help overcome melancholy. Instrumentalness, while less pronounced, is still important, suggesting that people may also be looking for more relaxing music.

When it’s snowing, only the songs’ instrumentalness significantly influences songs popularity. Instrumental music may be preferred because of its tranquil nature, which may correspond with the silence and serenity of the snowy landscape.

During thunderstorms, the popularity of the songs is most significantly influenced by energy, loudness, speechiness, instrumentalness and tempo. The high energy and volume may be a result of the need to drown out the noise of the storm, as well as to sustain a high level of indoor activity. The speechiness again indicates a preference for more narrative tracks that can engage the listener for longer. The tempo trait appears only in the thunderstorm, suggesting that only in these weather conditions, users seek such high speed of the song, resonating with the intensity of the environment.

The trait that does not reveal statistically significant in any weather condition - acousticness - refers to the degree to which a song is based on acoustic instruments as opposed to electronic ones. Nowadays, music preferences are often dominated by electronic music, pop, hip-hop, and other genres that do not always rely on acoustic sounds. Therefore, acousticness may not be an important factor for a wide range of listeners regardless of the weather conditions. It is also possible that acoustic is a more niche preference that does not significantly affect the overall popularity of songs under the conditions analyzed. Another trait that does not reveal significant in the analysis is liveness, often associated with more lively and energetic compositions. While such trait may be subjectively important to some users, its effect on popularity may be difficult to capture in the overall regression analyses, especially if other factors, such as danceability and energy, have stronger correlations with popularity. Liveliness may also co-depend on other characteristics that were more important, which may have diminished its independent effect in the analysis.

3.4.2 Weather conditions and artists

Also considered relevant to be investigated is the relationship between weather conditions and artists, in order to identify which artists can be associated with specific weather conditions. Since the dataset is very large, the decision was made to focus on songs with popularity above the value of 85. Table 5 in the Annex, presents the relationship between the artists of those songs and weather conditions, allowing to identify in which weather condition each artist is most trending.

The first row of Table 5 in the Annex displays the number of observations per weather condition. The fewest observations were recorded in the snowy conditions, suggesting that people listen to music at that time the least often. There are also no high numbers of songs (>10) in neither artist case, having over 91% observations in the range of 1-3 songs. Further, cloudy sky influenced few observations, where there were also no high numbers of observations with 88% of observations being 1-3 songs per artist. The thunderstorm caused little bit more observations

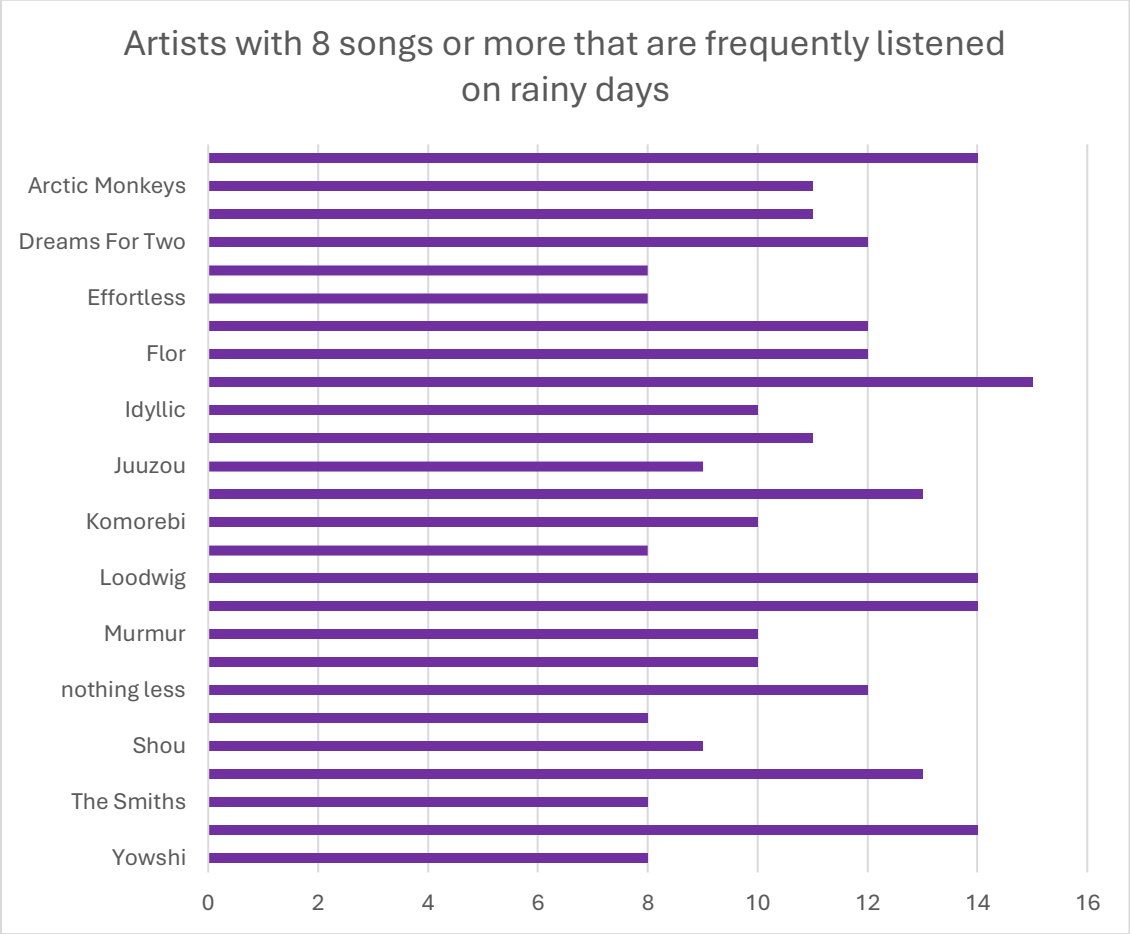
and had only 3 artists with higher numbers of songs registered. There is an artist – Lil Peep – that had 35 songs registered in this weather condition, constituting 15% of the observations. There were 80% out of 241 observations that had 1-3 songs. Clear conditions registered 279 observations with 89% of artists with 1-3 songs. Rain has the biggest number of observations on its own, as mentioned in previous analysis. There were many artists with high number of songs with only 33,5% of observation of range 1-3 songs. Another case is the combined weather conditions that together have the biggest number of observations, suggesting that when the weather is drizzle/fog/mist, so overall misty, people tend to listen to music the most often. However, there were 83% of observations within the range of 1-3 songs per artist.

One of the most listened artists is Taylor Swift, who will be analyzed in more detail in Section 3.5 of this dissertation. She is most popular when the weather is characterized as Clouds, then when it is foggy or raining. Therefore, algorithms should recommend her songs in the case of such weather conditions. Another widely listened artist is the Weeknd, and his songs are mostly trending during days with clear skies or when it is cloudy. There are many artists that are trending only in specific weather moments. For instance, the most popular songs (>85 value of popularity) of Harry Styles are only listened on days with clear skies, which indicates that it should be the main artist in recommendations during good weather conditions. The same applies to ROSALIA, Marshmello, Sabrina Carpenter or Tiesto, all known for their upbeat, dance music.

By looking at the table that can be found in the annex under Table 5, we can notice that specific artists receive higher popularity when the sky is clear, clouded or other.

Following, the insight that Spotify users mostly listen to music during rainy weather, chart 6 represents artists that have 8 or more songs that are popular in such conditions. Through this chart we can identify and prioritize the artists that should be included in the algorithms of music recommendations when weather forecasts predict rain. Some of the artists most listened to are: Akasha, her.story or Whimsical, all of whom are associated with nostalgic mood. People might be more receptive to emotionally rich or reflective songs when the sky is like that, perhaps seeking comfort or resonance in the music and therefore it is crucial for algorithm to consider such correlation.

Chart 6 – Artists most frequently listened on rainy days.



Source: Authors' analysis based on Kaggle dataset "6k Weather Labeled Spotify Songs"

In brief: results show that different weather conditions affect musical preferences, which is in line with psychological and social theories on how the surroundings affect human mood and behavior. Weather conditions can affect mood, energy levels and general activity, which in turn translates into musical preferences and the need to adjust marketing on streaming platforms accordingly.

3.5 Taylor Swift discography in Spotify

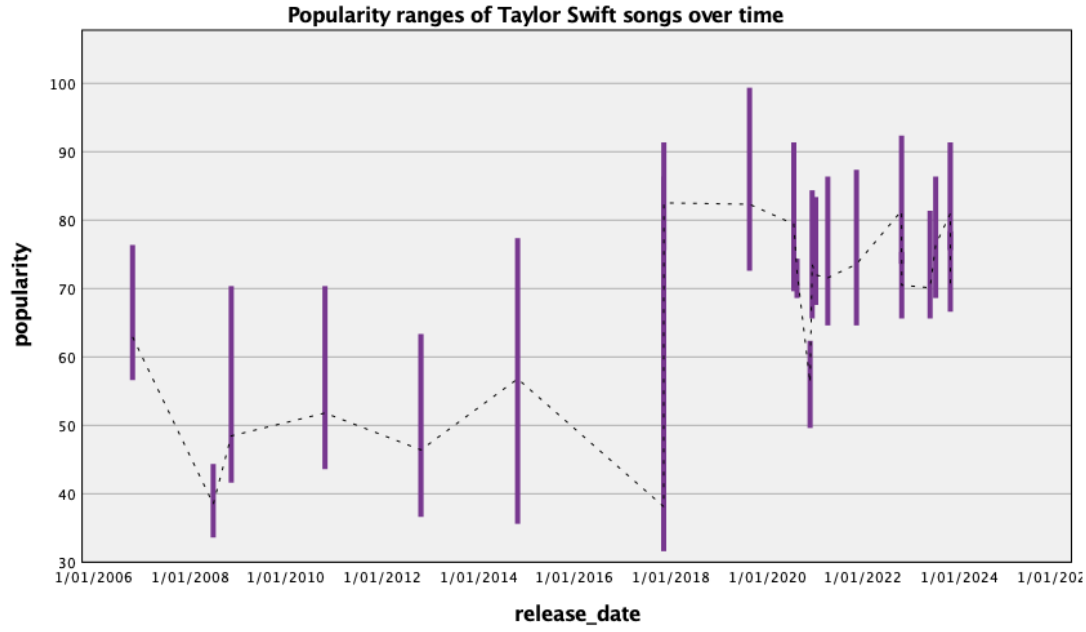
Analyzing Taylor Swift's Spotify songs can be considered relevant to marketing and understanding customer behavior due to her wide range of audience, as well as song traits. The whole world knows of her and people from all group ages consider themselves fans. As it comes to her songs, some of them have more positive notes, some include more spoken words and so on, thus providing a variety of cases to be analyzed. Moreover, research show that celebrity culture significantly influences the construction of individuals' identities, thus music choices (Junes, E. K. M., 2023). Taylor Swift, being a prominent figure in pop culture, has a significant influence that extends beyond her immediate fan base, impacting broader societal attitudes.

Identifying trends among her songs can reveal patterns in music consumption that reflect user preferences, which would be beneficial knowledge for marketers. By analyzing the variables of her Spotify songs, one can gain insights into the themes and narratives that resonate with Swift's audience. The Kaggle Spotify dataset (Taylor Swift Spotify Dataset, 2024) has 530 inputs, the number of songs Taylor Swift has released. Analyzing the dataset helps understanding the lyrical content, musical style shifts, and the overall tone of her songs. Marketers can use these insights to create campaigns that resonate with the themes associated with Swift, making their messaging more relevant to her audience.

3.5.1 Traits and Popularity of Taylor Swift songs, over time

In order to examine the variation in popularity of Taylor Swift's songs over the years, a high-low plot was created (Chart 7), where the Y-axis shows the popularity of the songs (on a scale of 0 to 100), and the X-axis shows the release dates. The analysis covers songs released from 2007 to 2024, with data divided into popularity ranges for each year.

Chart 7 – Taylor Swift songs' popularity over time



Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

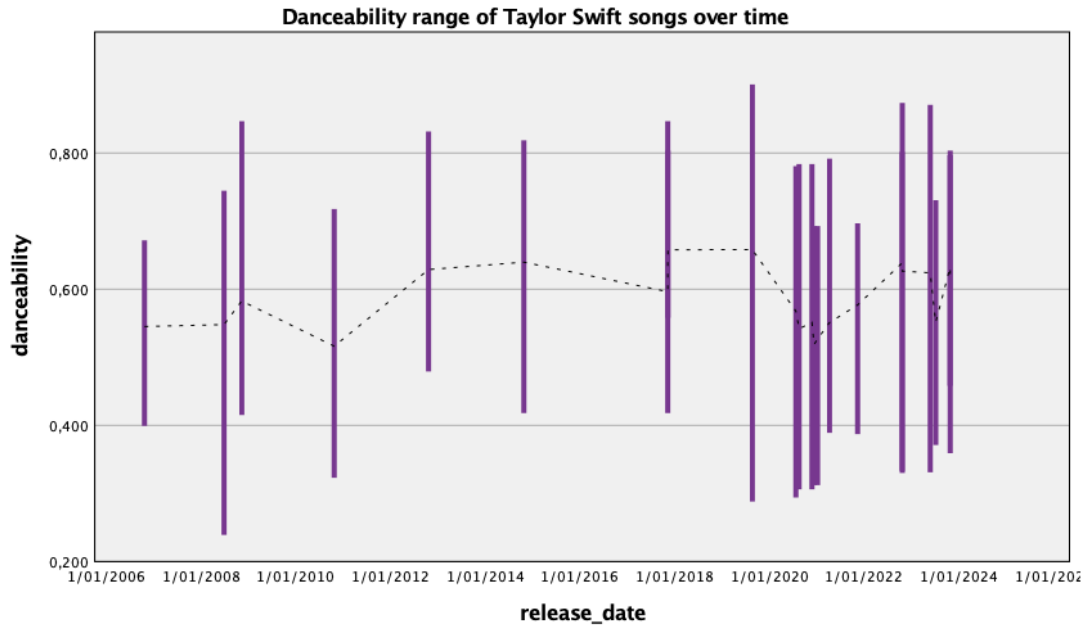
In 2007, Taylor Swift made her debut on the music scene, which may have contributed to the fluctuating popularity of her songs. The range of popularity during this period is relatively narrow, suggesting a stable but moderate interest in her work. Afterwards in 2008, the popularity of a few singles is low, but then she released a whole album that captured higher and wider range of popularity. The year 2011 shows a consistent range of popularity, which may indicate a stabilization of Taylor Swift's position in the music market. This may be the result of

an already built fan base and recognition from critics. The end of 2012 was marked by the release of album "Red" – considered a slight change of her previously more country sounds to more pop genres - that apparently didn't capture the musical preferences of her fans, as it lowered popularity. The year 2014 brought an expansion in popularity. This may reflect greater diversity in the quality and reception of songs, but also a growing diversity of fans who may have different music preferences. This year brought the public very popular songs like "Shake It Off" and "Blank Space". The end of 2017 shows the largest range of popularity, which may suggest broad musical experimentation and diverse responses from users. This year the most controversial album was released called "reputation", in which Taylor addressed her bad reputation, scandals and overall music industry. In 2020, Taylor Swift reached one of the highest levels of popularity for her songs, with a range of 75-100, which may be the result of a long career, high production quality and fan loyalty. The later 70-90 range may indicate some stabilization, but still at a very high level of popularity.

The years 2022-2024 are characterized by the stable high popularity of Taylor Swift's songs, suggesting that the artist has achieved a stable position in the music industry, and her work is regularly positively received by a wide audience. Indeed, in the last decade her songs gained more popularity, especially after the 24th of October 2017 release.

Similar reasoning was followed for the remaining song traits and high-low plots were created for each of the other traits, over time. Chart 8 displays the danceability range of Taylor Swift songs over time.

Chart 8 – Taylor Swift songs' danceability over time

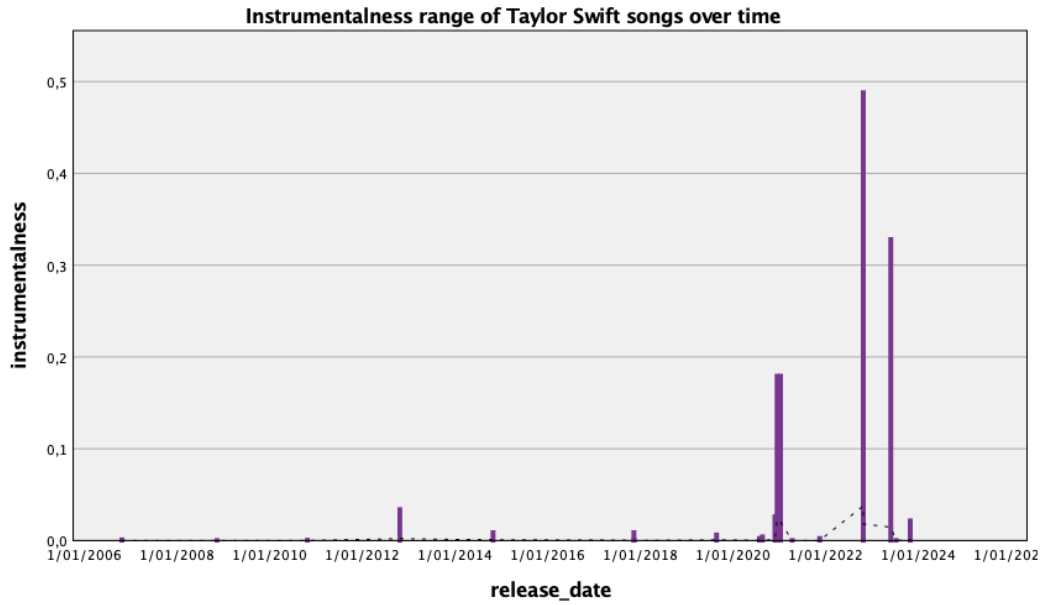


Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

The danceability of Taylor Swift's songs shows various fluctuations over time, reflecting the evolution of her musical style, as well as changing musical preferences and trends. In 2008 the wide range of danceability indicates a diverse style of songs. The increase in the range by the end of the year can be due to the release of the album "Fearless," which featured a mix of songs ranging from more balladic to danceable, such as "You Belong with Me." Year 2010 reveals a continuing diversity, but with a slightly smaller range than at the end of 2008, which may be related to the "Speak Now" album, with a more cohesive musical character. Years 2012, 2014, 2018 have similar ranges indicate stabilization in danceability, with possible releases such as "Red" (2012) and "1989" (2014), which were more experimental and pop oriented. Year 2019 brought a range in danceability that suggests a return to greater diversity in dance styles, which could be the result of the "Lover" album, which included both ballads and more energetic songs. In 2020 there was a drop in average danceability, and range may be related to the albums "Folklore" and "Evermore," which have a more folk and reflective feel, with fewer dance tracks. After 2020, despite similar ranges, the average danceability varies, which may be due to Taylor Swift's musical diversity in post-2020 releases, as well as to the influence of changing musical trends and artistic experimentation.

Chart 9 displays the instrumentalness range of Taylor Swift songs over time.

Chart 9 – Taylor Swift songs' instrumentalness over time



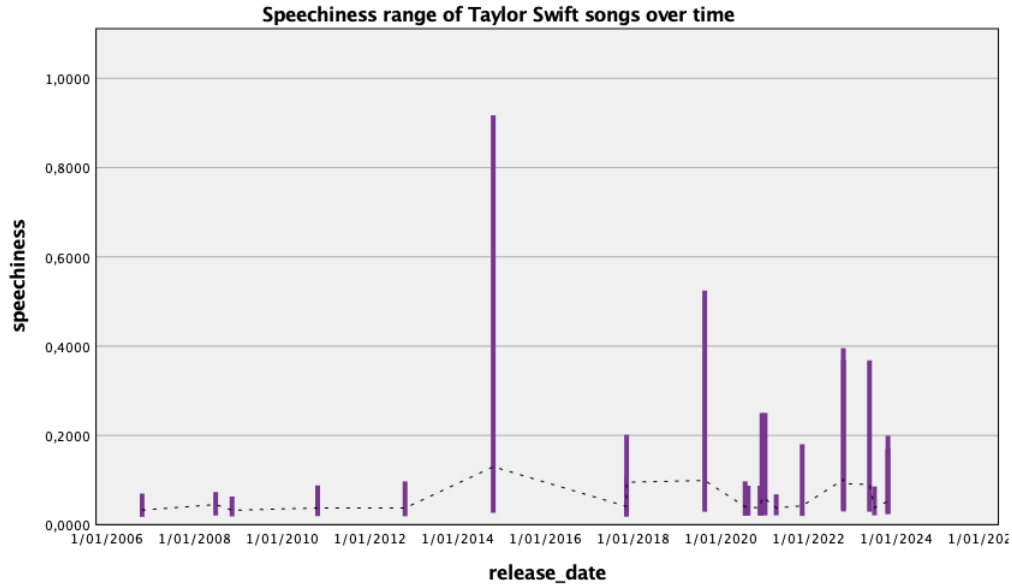
Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

The instrumentalness of Taylor Swift's songs was extremely low, almost zero, until 2020, indicating the dominance of vocal compositions in her work. Instrumentality near zero may

suggests that her songs were mainly vocal, which is typical of pop and country music, which is dominated by lyrics and vocals. In 2020 however, the range of instrumentality indicated a higher level of inclusion of instrumental elements in Taylor Swift's music, which may be related to the more experimental nature of the “Folklore” and “Evermore” albums. Year 2023 showed two ranges which may suggest that this year Taylor Swift released music with greater instrumental diversity, possibly the result of further musical experimentation and expansion of her style.

Chart 10 displays the speechiness range of Taylor Swift songs over time.

Chart 10 – Taylor Swift songs’ speechiness over time



Source: Authors’ analysis based on Kaggle dataset “Taylor Swift Spotify Dataset”

The speechiness of Taylor Swift's songs has also evolved over time. By 2012, low word content suggests that her earlier songs had more condensed lyrics, typical of country and pop music. In 2015 the range of speechiness indicates a drastic increase in word count, which may be the result of the “1989” album and subsequent releases, where lyrics become more narrative and complex. However, in 2018 there was a return to the lower range, suggesting a period of return to simpler lyrics or a reduction in word count in favor of more melodic elements. Then, in 2020, a noticeable rise to the range may be due to the more narrative nature of the “Folklore” and “Evermore” albums, where the lyrics became a storytelling. In 2023, the range of speechiness may suggest further adaptation of varied approaches to lyrics, balancing narrative and more concise songs.

Charts for the remaining 6 traits – acousticness, energy, liveness, loudness, tempo and valence – are displayed in the annex, Charts 15 to 20. From these 6 charts it is possible to conclude that the corresponding traits showed more stable trends, less interesting from a dynamic change perspective. The stability of the other factors may be due to their primary role in defining Taylor Swift's overall style of music, which remains recognizable and consistent despite its evolution. Stable acoustic indicates that Taylor Swift maintains a balance between acoustic and electronic elements in her music. Energy suggests that despite the diversity of styles, the dynamics and intensity of the songs remain at a similar level.

Stable liveliness shows that the tempo and dynamics of songs are consistent, which is important to her recognizable sound. Volume reflects consistent song production and mastering, which is standard in modern pop music. A stable tempo indicates consistency in the rhythmic structures of her music. Value suggests that the emotional tone of songs remains similar, which is key to her brand of music.

Further analysis could lead to examining the impact of various features of Taylor Swift's songs on their popularity. To better understand which song traits, have the greatest impact on the number of plays and overall user interest, separate regressions were run for each of the selected variables. The results present standardized beta coefficients that show the impact of each feature on a song's popularity. Standardized beta coefficients allow a comparison of the strength of the influence of different independent variables on the dependent variable, in this case popularity. Table 6 shows the beta and standardized beta coefficients for the chosen variables: energy, liveness and loudness.

Table 6 – Taylor Swift songs' traits and their effects on popularity

COEFFICIENTS WITH POPULARITY	B VALUE	STANDARDIZED BETA	SIG VALUE
ENERGY	17,883	,225	,007
LIVENESS	-17,238	-,161	<,001
LOUDNESS	-1,779	-,343	<,001

Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

Examining the beta coefficients between traits and song popularity, there are three main significant effects. A high positive beta coefficient (17.883) suggests a significant positive effect of energy on popularity. This effect is statistically significant at the 0.01 level, indicating a relatively strong relationship (standardized beta equals 0.225). In view of Taylor Swift's songs,

this means that more vibrant and dynamic songs tend to gain more popularity. Such a result suggests that platforms such as Spotify should recommend more Taylor Swift songs with high energy levels. Moreover, as mentioned earlier, since Taylor Swift attracts the attention of a significant portion of the music market, these findings can be applied to a broader context.

Moving onto the effect of liveness on popularity, the beta coefficient is -17.3238 and is statistically significant at the 0.01 level. A negative standardized beta coefficient of -0.161 indicates that songs with lower levels of liveness (less likelihood of audience attendance during a recording) are more likely to be popular. For music marketers and producers, understanding this negative relationship is crucial. It means that songs clearly recorded in the studio, without "live" elements, are more likely to gain popularity.

Another negative effect (standardized beta equals –0.343) is observed for loudness. It suggests that the louder a song is, the less popular it becomes. This means that Taylor Swift's fans, and by extension the wider public, prefer quieter, less noisy songs.

Noting the magnitude of the standardized beta coefficients, which show the strength of the influence of each variable on popularity, it can be seen that loudness has the most influence on popularity (it is a negative influence), followed by positive influence of energy. The least influence (negative) is liveness.

3.5.2 Traits and popularity of Taylor Swift songs, by album

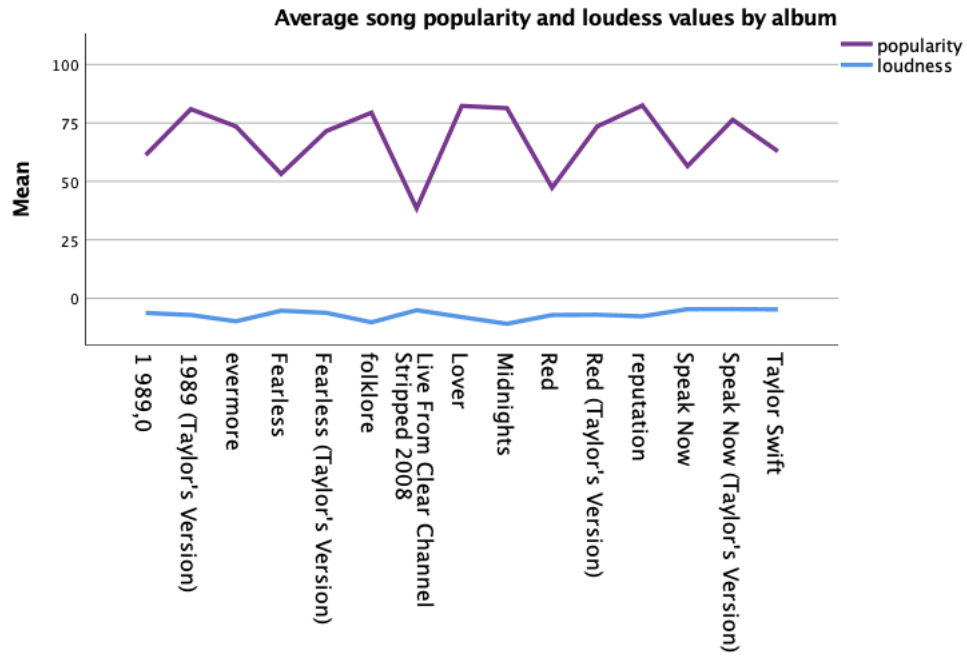
Taylor Swift's fans often identify with one of her albums, creating a custom in which they choose the one that fits them best. For example, a statement: "I'm a folklore person," would mean that this is the album that resonates with them the most and consider themselves nostalgic and introvert person. Each album is associated with certain characteristics that allow a person to be quickly assigned to a certain type. Further analysis could look at what characteristics of the songs on each album affect their popularity. To do this, the dataset was divided into individual albums and a regression analysis conducted.

Following up on the previous result, showing that songs with lower volume are more likely to be popular, one can go a step further and examine which Taylor Swift albums have the greatest potential to reach listeners in terms of this variable. It is worth remembering, however, that there are different versions of most of her albums, as Taylor re-recorded them, adding "Taylor's Version" to the titles in order to retain the copyright to her lyrics. This was the result of a famous case involving accusations of plagiarism, when other artists claimed that Taylor had allegedly

copied their lyrics. To protect herself from such allegations, she created new album versions with the same lyrics but making it clear that they were written by her.

Chart 11 displays average popularity and loudness of Taylor Swift songs, by album.

Chart 11 – Taylor Swift songs' average popularity and loudness values by album



Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

In the situation of rerecording, she had an opportunity to adjust the traits of her songs, increasing volume, changing beat or instrumentalness slightly. Based on loudness and on Chart 11, we can see that in the case of 1989 and Fearless album, she adjusted her TV (Taylor's Version) in line with the insight from the previous analysis, meaning that she decreased the loudness of songs in the rerecorded version. Then looking at the popularity line, one can identify that in fact the lowered loudness led to increased popularity of the same album but rerecorded.

As the frequencies of each album are low, it was decided to just analyze the correlations (and corresponding p-values) between Taylor Swift song traits and their popularity, as shown in Table 7, and not to conduct regression analysis.

Table 7 – Taylor Swift songs' traits correlations with popularity, by album

STANDARDIZED BETA COEFFICIENTS	1989	EVER MORE	FEARLESS	LOVER	RED	REPUTATION	SPEAK NOW	TAYLOR SWIFT
DANCEABILITY		**-,472** ,006			**-,380** ,001	**,315** ,013		**,538** ,038
ENERGY	**,249** ,031			**,560** ,016				
LOUDNESS	**,274** ,017			**,561** ,015		**-,411** <,001	**,295** ,012	
SPEECHINESS	**-,493** <,001		**,263** ,046			**,625** <,001		
ACOUSTICNESS	**-,423** <,001							
INSTRUMENTALNESS								
LIVENESS	**-,275** ,017						**-,448** <,001	
VALENCE						**-,386** ,002		
TEMPO								

Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

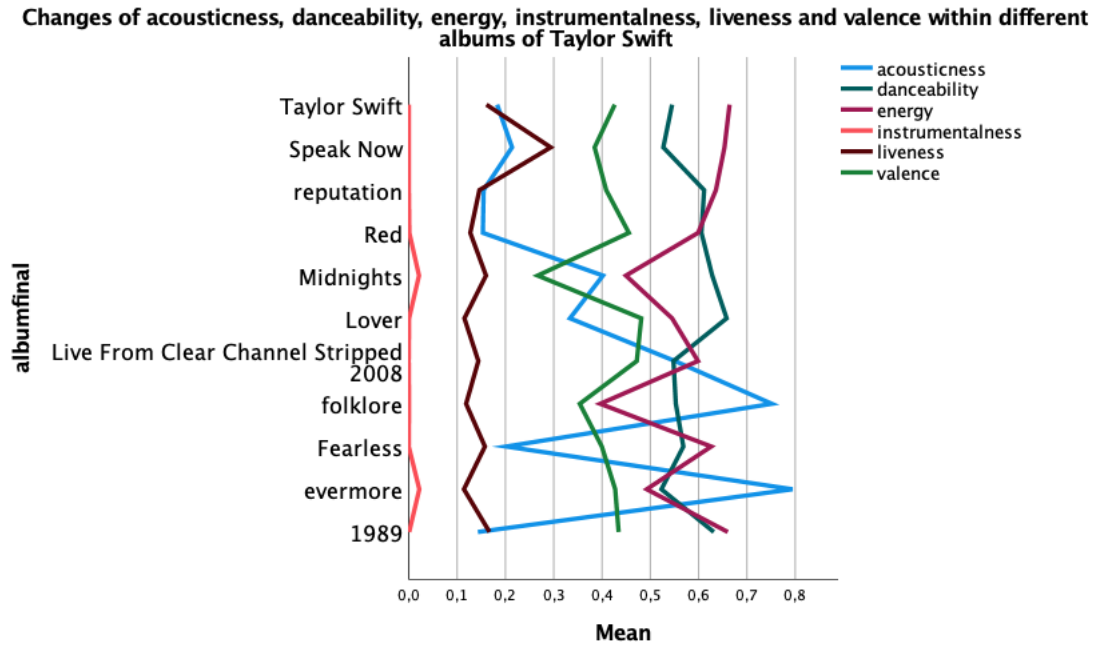
Considering only statistically significant correlations, it can be seen that those who listen to the 1989 album most often prefer songs with lower acoustics, lower spoken word content (speechiness) and lower volume. Based on this, it can be assumed that fans of the 1989 album value songs with less spoken word and less acoustic instrumentation, which may help in creating recommendation algorithms. Fans of the Red album, on the other hand, are mainly attracted to higher acoustics and lower danceability and liveness. It can be seen here that these people value acoustic instruments and slower, more nostalgic sounds more, which provides direct guidance for future recommendations.

In contrast, listeners of the reputation album prefer songs with high danceability and spoken word content (speechiness). Based on this, marketers may conclude that these fans value songs that can be danced to and those that contain storytelling through lyrics or other speech-like elements, suggesting that they would also appreciate other artists who offer similar elements in their music.

In the case of the Speak Now album, an important factor influencing popularity is liveness. Marketers can deduce that frequent users of this album prefer songs with lower liveness, that is, songs that are less dynamic and more nostalgic, with a slower tempo.

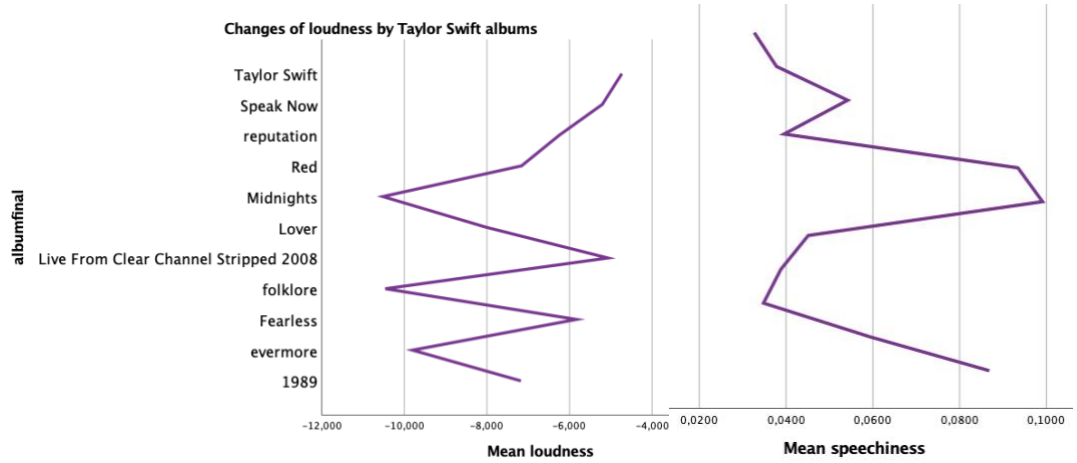
Further analyses include the creation of line graphs that shows how these qualities have changed over the course of Taylor Swift's various albums, thus helping to illustrate the evolution of her music in terms of key traits (Charts 12 to 14).

Chart 12 – Taylor Swift songs' average popularity and loudness values, by album



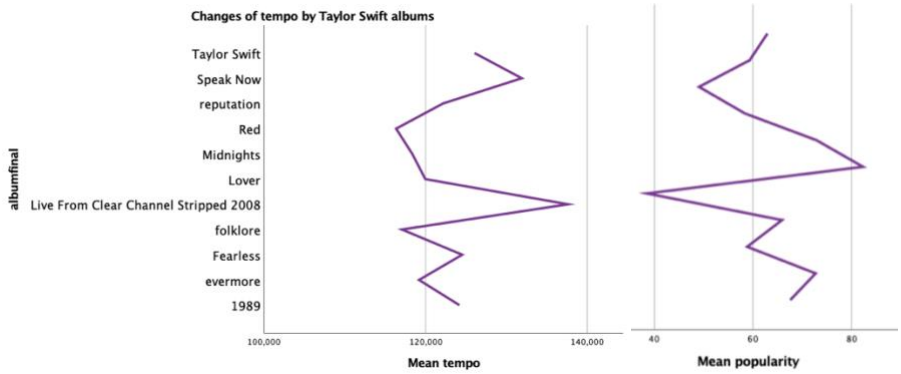
Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

Chart 13 – Taylor Swift songs' average popularity and loudness values, by album



Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

Chart 14 – Taylor Swift songs' average popularity and loudness values, by album



Source: Authors' analysis based on Kaggle dataset "Taylor Swift Spotify Dataset"

The most varied characteristics are clearly found in chart 13, namely loudness and speechiness (i.e., the content of spoken parts in songs). The albums Fearless and Live have the highest loudness, which, in line with previous results, suggests that these two albums will achieve the lowest popularity. These results confirm that higher volume negatively affects the reception of songs, as seen in the analyzed data. On the other hand, in chart 14 we can see that tempo and popularity varies greatly as well, having Speak Now and Live recording as the most significant outliers.

In terms of speechiness (chart 13), albums such as 1989, Lover and Midnights contain the most spoken parts. This may mean that fans of these albums are more likely to appreciate songs in which lyrics play an important role, and narrative or spoken word elements appear in the music. On the other hand, danceability (i.e., the danceability of the songs) is quite similar in all Taylor Swift albums, indicating that in terms of rhythm and tempo, the songs from different albums have a similar structure. It is also worth noting that liveness (chart 12) only differs in the case of the Speak Now album, which may suggest that the songs from this album are less "lively," more studio-like, which may affect their popularity among fans who prefer quieter sounds.

Concluding this section, one can notice that analysis of Taylor Swift songs provides insights into a broad range of users and can constitute a basis for analysis of overall musical preferences. It would be fairly easy to group users into fans and non-fans of this artist and as proven above it gives a great intake on user preferences, thus such segmentation could contribute greatly to the marketing of music streaming platform, such as Spotify.

4. Conclusion

Gathering all the insights from conducted research and statements in previous chapters, in the authors' view the analysis of data by year, country and weather conditions, combined with the Taylor Swift case study and comprehensive literature review creates a complex overview of the topic at hand. All the elements bring a distinctive perspective into this dissertation, making sure the results are reliable and well thought through. By gathering the results from literature review, it was shown musical traits have significant influence on popularity of the song, and popularity translates directly into user preferences. Thus, it was stated that further analysis in this area is relevant and responds to the needs for consequent user analysis of the music streaming platforms. The comparison of Spotify with YouTube has shown the dynamics of this industry and explanation of recent trends noted the movement of data driven approaches, enforcing the relevance of the further statistical analysis.

Moving to the results from the statistical analysis that were conducted, and to the insights from them, the whole picture was created starting off with year-based observations, then geographical factors, weather and ending with study of Taylor Swift vast songs portfolio. As observed, some of the musical traits changed significantly over time. A clear insight was that in the years before 1950 there were far more abrupt changes, with many increases and decreases in the level of for example accousticness, which started off with the highest values and showed the largest decrease over the years. Then in the geographical sense, there are visible differences between countries, with some outstanding observation for locations. Some of the most interesting insights were the Indonesia (ID), Israel (IL) and Malaysia (MY) preference for lower valence in the songs, suggesting more melancholic tracks; users in Morocco (MA) and Portugal (PT) greatly favor high speechiness, putting great importance on spoken words in the songs. These and the others mentioned in the corresponding chapter, emphasize the need for geographical segmentation and personalization for specific locations. Moving onto the weather conditioned analysis, the most obvious conclusion was that rain had the largest frequency of observation, thus suggesting that during such weather, users listen to music most often, giving clear guidance for streaming platforms and their algorithm creation to detect such weather condition and use it to their advantage. There were also traits that influenced the popularity in the drizzle/mist/fog and thunderstorms, namely speechiness, suggesting usage of such and the other mentioned in the chapter in the algorithms. Further notice of artists listened on specific weather conditions, like Lil Peep when there is Thunderstorm, leads to a direct linkage of them to the weather, showing the preference for their music in such conditions. In the part of the

analysis, where Taylor Swift songs were analyzed, it was noted that her albums vary greatly in the sense of which traits influence the popularity. The most influential was the trait called loudness and had the greatest influence in the album named Lover. Analysis over time helped understanding of the development of this artist and the respective changes of popularity, showing user preferences through the reactions of her fans. All the above has shown how complex the topic of musical user preferences is, and this dissertation aimed at expanding the view of it even further.

Based on the above results, one can recommend further research toward geographic segmentation and personalization for local preferences. It also would appear important to incorporate weather conditions into the algorithms of streaming platforms, which could increase user engagement. In addition, by analyzing the career development of artists such as Taylor Swift, it is possible to better understand users' changing preferences, which can contribute to the creation of more accurate models predicting the popularity of songs in the future.

Bibliography

6K Weather labeled Spotify Songs. (2023, June 20). Kaggle. https://www.kaggle.com/datasets/receplyasolu/6k-weather-labeled-spotify-songs

Aguiar, L., & Waldfogel, J. (2018). *Platforms, promotion, and product discovery: Evidence from Spotify playlists* (No. w24713). National Bureau of Economic Research.

Anderson, I., Gil, S., Gibson, C., Wolf, S. T., Shapiro, W., Semerci, O., & Greenberg, D. M. (2020). "Just The Way You Are": linking music listening on Spotify and personality. Social Psychological and Personality Science, 12(4), 561–572. https://doi.org/10.1177/1948550620923228

Anglada-Tort, M., Lee, H., Krause, A. E., & North, A. C. (2023). Here comes the sun: music features of popular songs reflect prevailing weather conditions. *Royal Society Open Science, 10*(5). https://doi.org/10.1098/rsos.221443

Barata, M. L., & Coelho, P. S. (2021). Music streaming services: understanding the drivers of customer purchase and intention to recommend. Heliyon, 7(8), e07783. https://doi.org/10.1016/j.heliyon.2021.e07783

Bello, P., & García, D. (2021). Cultural Divergence in popular music: the increasing diversity of music consumption on Spotify across countries. *Humanities and Social Sciences Communications*, *8*(1). https://doi.org/10.1057/s41599-021-00855-1

Beuscart, J., Coavoux, S., & Garrocq, J. (2022). Listening to music videos on YouTube. Digital consumption practices and the environmental impact of streaming. Journal of Consumer Culture, 23(3), 654–671. https://doi.org/10.1177/14695405221133266

Blake E (2020) Data shows 90 percent of streams go to the top 1 percent of artists. Rolling Stone. Available at: https://www.rollingstone.com/pro/news/top-1-percent-streaming-1055005/

Bonini, T., & Magaudda, P. (2024). Platformed! How Streaming, Algorithms and Artificial Intelligence are Shaping Music Cultures. In *Pop music, culture, and identity*. https://doi.org/10.1007/978-3-031-43965-0

Celebrity, music, and personal persona: A case study of Taylor Swift [master's thesis, Minnesota State University, Mankato]. Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato. https://cornerstone.lib.mnsu.edu/etds/1296/

Celma Ó (2010) Music Recommendation and Discovery Revisited: The Long Tail, Long Fail, and Long Play in the Digital Music Space. New York: Springer.

Cieślak, J. (2023, November 29). Spotify'23: Polacy są muzycznymi patriotami - rp.pl. *Rzeczpospolita*. https://www.rp.pl/muzyka-popularna/art39486751-spotify-23-polacy-sa-muzycznymi-patriotami

DeHart C (2018) Metal by numbers: revisiting the uneven distribution of heavy metal music. Metal Music Studies 4(3): 559–571.

De Fleurian, R., & Pearce, M. T. (2021). The Relationship between Valence and chills in Music: A Corpus analysis. I-perception, 12(4), 204166952110246. https://doi.org/10.1177/20416695211024680

Development, O.-. W. &. M. (2022, February 24). *Daniel Ek: A brief history of Spotify and its founder*. https://www.linkedin.com/pulse/daniel-ek-brief-history-spotify-its-founder-oqtacore/

Durrani, A. (2023, March 27). The Average American Spends Over 13 Hours A Day Using Digital Media—Here's What They're Streaming. *Forbes Home*. https://www.forbes.com/home-improvement/internet/streaming-stats/

Duman, D., Neto, P., Mavrolampados, A., Toiviainen, P., & Luck, G. (2022). Music we move to: Spotify audio features and reasons for listening. *PLOS ONE, 17*(9), e0275228. https://doi.org/10.1371/journal.pone.0275228
Fricke, K. R., Greenberg, D. M., Rentfrow, P. J., & Herzberg, P. Y. (2019). Measuring musical preferences from listening behavior: Data from one million people and 200,000 songs.
Ghildiyal, A., & Sharma, S. (2021). Music Genre Classification using Data Filtering Algorithm: An Artificial Intelligence approach. *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*. https://doi.org/10.1109/icirca51532.2021.9544592
Håvard Kiberg & Hendrik Spilker (2023) One More Turn after the Algorithmic Turn? Spotify's Colonization of the Online Audio Space, Popular Music, and Society, 46:2, 151-171, DOI: 10.1080/03007766.2023.2184160
IFPI GLOBAL MUSIC REPORT 2021. (n.d.). https://gmr2021.ifpi.org/report
Junes, E. K. M. (2023).
Kot, G. (2023, December 1). rock and roll. Encyclopedia Britannica. https://www.britannica.com/art/rock-and-roll-early-style-of-rock-music
Klimczuk, A., & Klimczuk-Kochańska, M. (2019). Core-Periphery model. In *Springer eBooks* (pp. 1–8). https://doi.org/10.1007/978-3-319-74336-3_320-1
Martin, R. (2023, November 28). electric guitar. Encyclopedia Britannica. https://www.britannica.com/art/electric-guitar
McFee, B., Bertin-Mahieux, T., Ellis, D. P. W., & Lanckriet, G. R. G. (2012). The million-song dataset challenge. WWW '12 Companion, 909–916. https://doi.org/10.1145/2187980.2188222
Mulligan, M. (2019, April 3). 10 Trends that will Reshape the music industry. https://midiaresearch.com/blog/10-trends-that-will-reshape-the-music-industry
Maharshi Pandya. (2022). 🎹 Spotify Tracks Dataset [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/4372070
Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, *20*(11), 4275–4292. https://doi.org/10.1177/1461444818769694
Prey, R. (2020). Locating Power in Platformization: Music Streaming Playlists and Curatorial Power. *SAGE*, *6*(2), 205630512093329. https://doi.org/10.1177/2056305120933291
Prey, R. (2021). Playlists and the datafication of music formatting. *ResearchGate*. https://www.researchgate.net/publication/349324390_Playlists_and_the_Datafication_of_Music_Formatting
Psychology of Music, 49(3), 371–381. https://doi.org/10.1177/0305735619868280
Greenberg, D. M., Kosinski, M., Stillwell, D., Monteiro, B., Levitin, D. J., & Rentfrow, P. J. (2016). The song is you. Social Psychological and Personality Science, 7(6), 597–605. https://doi.org/10.1177/1948550616641473
Rodgers, W., Yeung, F., Odindo, C., & Degbey, W. Y. (2021). Artificial intelligence-driven music biometrics influencing customers' retail buying behavior. *Journal of Business Research*, *126*, 401–414. https://doi.org/10.1016/j.jbusres.2020.12.039
Sletten, Riley T., "Spotify: Strategic Plan and Analysis" (2021). Undergraduate Theses, Professional Papers, and Capstone Artifacts. 325. https://scholarworks.umt.edu/utpp/325
Spotify dataset. (2021, December 17). Kaggle. https://www.kaggle.com/datasets/vatsalmavani/spotify-dataset
Spotify — About Spotify. (2023, October 24). Spotify. https://newsroom.spotify.com/company-info/

Similarweb. (2024, January). *spotify.com Traffic Comparison | Similarweb*. https://www.similarweb.com/website/spotify.com/#competitors
Similarweb. (2024, January). *music.youtube.com Traffic Comparison | Similarweb*. https://www.similarweb.com/website/music.youtube.com/#traffic-sources
Taylor Swift Spotify Dataset. (2024, August 3). Kaggle. https://www.kaggle.com/datasets/jarredpriester/taylor-swift-spotify-dataset
Till, C. (2023). Spotify as a technology for integrating health, exercise, and wellness practices into financialised capitalism. *Big Data & Society*, *10*(2). https://doi.org/10.1177/20539517231210278
Today, U. (2024, January 30). "Motherly love": Singer Cat Janice enters hospice; last song proceeds go to 7-year-old son. *USA TODAY*. https://eu.usatoday.com/story/life/humankind/2024/01/26/cat-janice-cancer-hospice-dance-you-outta-my-head/72371623007/
Tófalvy, T., & Koltai, J. (2021). "Splendid Isolation": The reproduction of music industry inequalities in Spotify's recommendation system. New Media & Society, 25(7), 1580–1604. https://doi.org/10.1177/14614448211022161
Top Spotify songs in 73 countries (Daily updated). (2024, September 21). Kaggle. https://www.kaggle.com/datasets/asaniczka/top-spotify-songs-in-73-countries-daily-updated
Wang, H., Wang, N., & Yeung, D. (2014). Collaborative deep learning for recommender systems. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1409.2944\
Wojdynski, B.W. (2016). Native advertising: Engagement, deception, and implications for theory. In R. Brown, V. K. Jones, and B. M. Wang (Eds.), The New Advertising: Branding, Content and Consumer Relationships in a Data-Driven social media Era (pp. 203-236). Santa Barbara, CA: Praeger/ABC Clio.
Wojdynski, B.W. (2016). Native advertising: Engagement, deception, and implications for theory. In R. Brown, V. K. Jones, and B. M. Wang (Eds.), The New Advertising: Branding, Content and Consumer Relationships in a Data-Driven social media Era (pp. 203-236). Santa Barbara, CA: Praeger/ABC Clio.
Wojdynski, B.W. (2016). Native advertising: Engagement, deception, and implications for theory. In R. Brown, V. K. Jones, and B. M. Wang (Eds.), The New Advertising: Branding, Content and Consumer Relationships in a Data-Driven social media Era (pp. 203-236). Santa Barbara, CA: Praeger/ABC Clio.
Web API Reference | Spotify for Developers. (n.d.). https://developer.spotify.com/documentation/web-api/reference/get-several-audio-features

Wojdynski, B.W. (2016). Native advertising: Engagement, deception, and implications for theory. In R. Brown, V. K. Jones, and B. M. Wang (Eds.), *The New Advertising: Branding, Content and Consumer Relationships in a Data-Driven social media Era* (pp. 203-236). Santa Barbara, CA: Praeger/ABC Clio.

Xiaorui Guo, The Evolution of the Music Industry in the Digital Age: From Records to Streaming. Journal of Sociology and Ethnology (2023) Vol. 5: 7-12. DOI: http://dx.doi.org/10.23977/jsoce.2023.051002.

Annex A

Graph 7- GLOBAL RECORDED MUSIC INDUSTRY REVENUES 1999 - 2022 (US$ BILLIONS) - IFPI

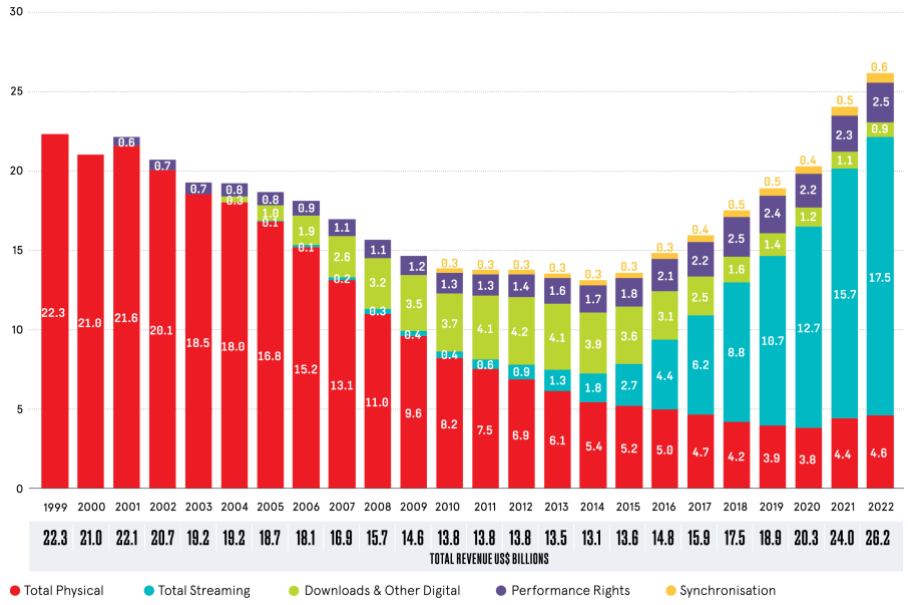


Table 4 – Standardized beta coefficients of the traits significantly influence popularity

STANDARDIZED BETA COEFFICIENTS	WEATHER CONDITION DIFFERENTIATION						NO WEATHER CONDITION DIFFEREN-TIATION
FEATURES	CLEAR	CLOUDS	DRIZZLE/ FOG/ MIST	RAIN	SNOW	THUNDER STORMS	
DANCEABILITY		**,203** ,005		**-,111** ,027			
ENERGY			**-,300** <,001	**-,312** <,001		**-,341** ,007	**-,238** <,001
LOUDNESS			**,269** <,001	**,247** ,002		**,230** ,033	**,196** <,001
SPEECHINESS			**-,102** ,019			**,160** ,012	**-,069** ,002
INSTRUMENTALNESS	**-,149** ,033	**-,253** <,001	**-,106** ,014	**-,203** <,001	**-,306** <,001	**-,171** ,014	**-,192** <,001
VALENCE		**-,249** ,001					
TEMPO						**-,136** ,036	

Table 5 – Relationship between Artist and Weather conditions

	279	233	617	525	205	241
	Clear	**Cloud**	**Drizzle/Fog/Mist**	**Rain**	**Snow**	**Thunderstorm**
Akasha				14		
Amy Winehouse						6
Arctic Monkeys			5	11		
Avicci	5					
Beabadoobee			5			
Billie Eilish		2	12			5
Bon Iver			4			
Calvin Harris	11	2				
Camila Cabello	2	4				
Cashmere				11		
Cavetown			7			
Cigarettes After Sex		2	19	2		
Cody Fry			5	1		
Conan Gray			4			
David Guetta	4	2				
Dreams For Two				12		
Dua Lipa	6	1				
Dulcet				4		
Ed Sheeran		3		8		
Effortless				8		
eiji				12		
Eminem		4				
Flor				12		
Frank Ocean		1			4	
Girl in red			4			
Good Morning					4	
Gracie Abrams			4			
Halsey		7				
her.story				15		
Hozier			8	1		
Idyllic				10		
James Arthur			1	4		
Jeremy Zucker			5			
just johnny				11		
Justin Bieber	1			4		
Juuzou				9		
Klein				13		
Komorebi				10		
Labrinth			4			
Lana Del Rey		1	6	5		
Lil Peep			3			35

Lonely Boy				8		
Loodwig				14		
meow meow				14		
Murmur				10		
Natsu				10		
Nirvana			1	6		
Nivla				7		
nothing less				12		
Radiohead			3	8		
Rascal				7		
Shawn Mendes	1	5		5		
Shou				9		
softhead				13		
Tashi				6		
The 1975	1	1	3	7		
The Avett Brothers					4	
The Neighborhood		6	9			
The Smiths		1	1	8		
Vacations					5	
Wells				5		
Whimsical				14		
Wolly				5		
XXXTENTACION			1		4	5
Yowshi				8		

Chart 15 – Taylor Swift songs' acousticness over time

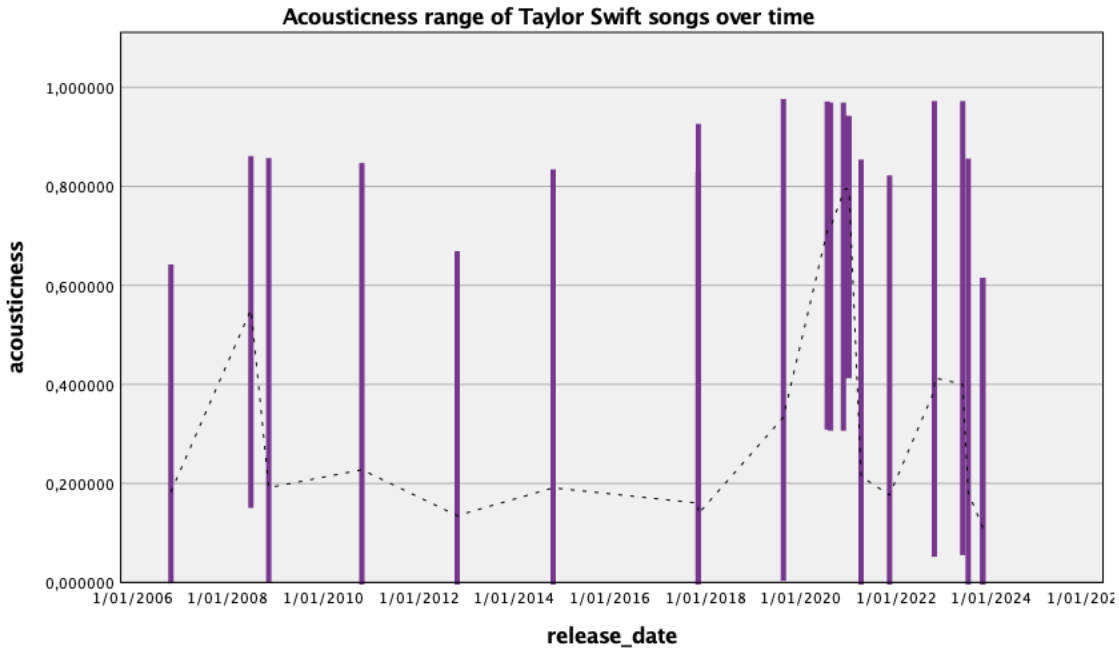


Chart 16 – Taylor Swift songs' energy over time

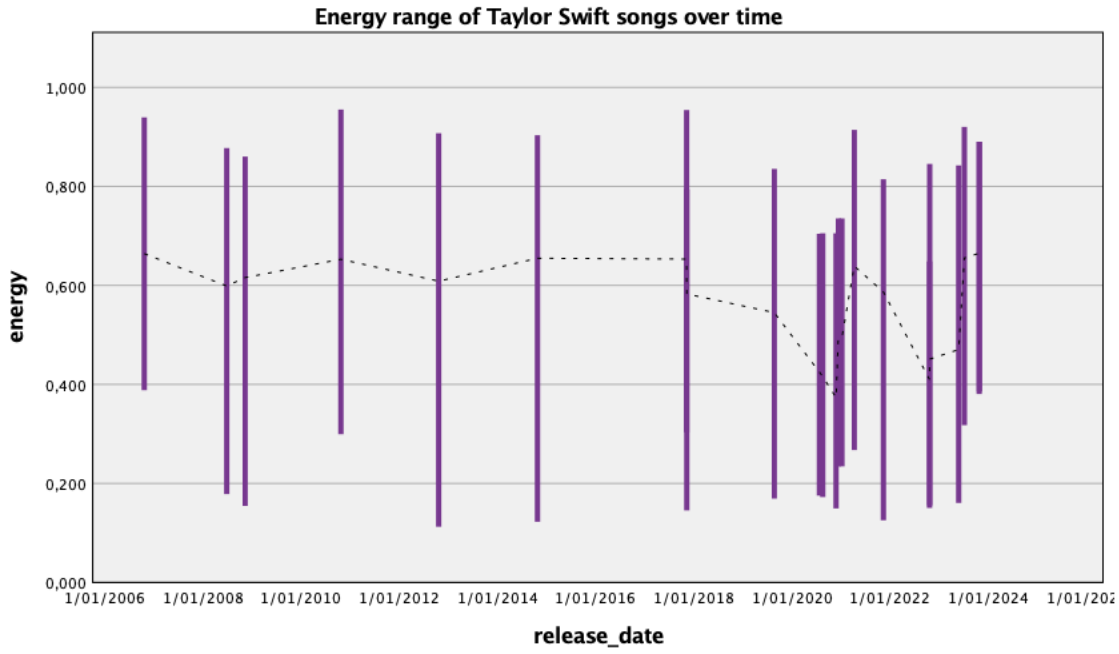


Chart 17 – Taylor Swift songs' liveness over time

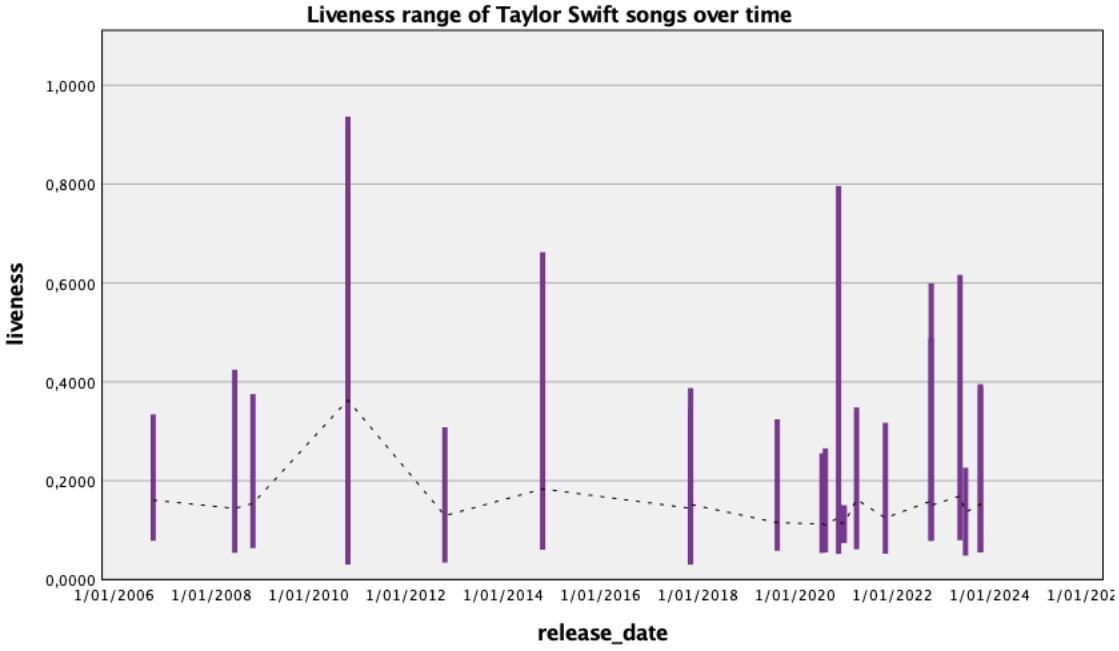


Chart 18 – Taylor Swift songs' loudness over time

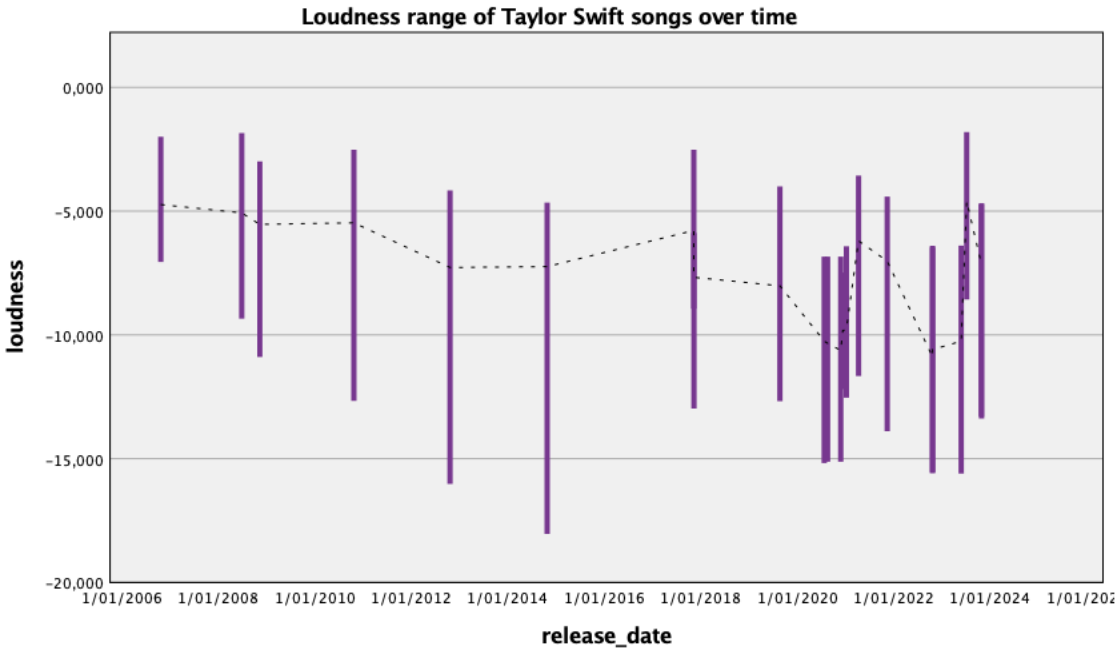


Chart 19 – Taylor Swift songs' tempo over time

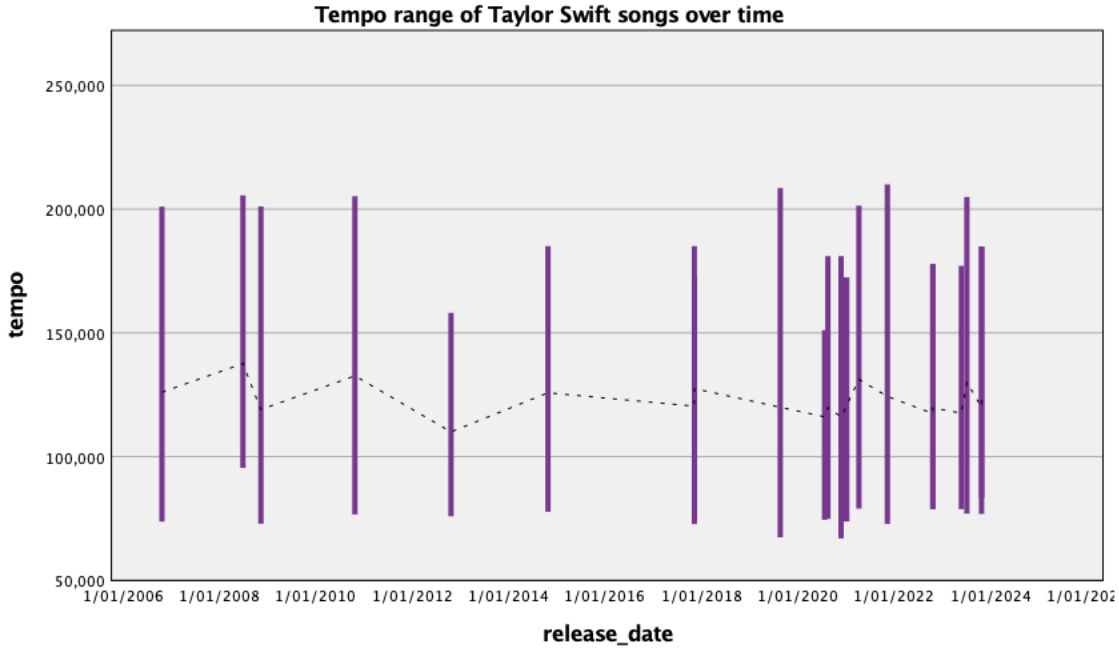


Chart 20 – Taylor Swift songs' valence over time