

The Chainsmokers' Musical Evolution*

From Pop-Infused EDM to a Consistent Electronic Core

Jerry Xia

Peter Fan

Jason Yang

October 10, 2024

This paper analyzes the evolution of The Chainsmokers' music by examining key audio features such as acousticness, energy, valence, and danceability using data from the Spotify API. The study focuses on the band's most popular albums, showcasing their transition through various pop-influenced styles while retaining a consistent electronic production core. By analyzing these attributes across their discography, this study provides insights into how The Chainsmokers adapted their music while maintaining high energy and rhythmic accessibility.

Table of contents

1	Introduction	2
2	Data	2
2.1	Source	2
2.2	Measurement	3
2.3	Results	3
2.3.1	Energy Across Albums	3
3	Discussion	7
3.1	Weaknesses and Next Steps	7
A	Appendix	9
A.1	Raw Data Processing	9
A.2	Cleaned Data	9
A.3	Data Processing Overview	9
	References	11

*Code and data are available at: https://github.com/ycfan0991/analyzing_the_chainsmokers_streaming_data

1 Introduction

The Chainsmokers are a globally recognized DJ duo known for their electronic dance music (EDM) blended with pop influences. Since their breakout in 2014 with “Selfie,” they have consistently released chart-topping singles and albums that blend infectious beats with mainstream pop appeal. This fusion has allowed them to carve out a unique space in the music industry, where they collaborate with various vocalists to create a signature sound that resonates across diverse audiences.

This paper seeks to explore the evolution of The Chainsmokers’ sound by analyzing key audio features such as **acousticness**, **energy**, **valence**, **danceability**, and **speechiness** using data sourced from the Spotify API. By examining these attributes across their most popular albums, this study aims to uncover trends in their musical style and provide insights into how they have adapted to changes in the music landscape while maintaining core elements of their production.

Focusing on albums such as *Memories...Do Not Open*, *Sick Boy*, and *World War Joy*, this paper will analyze how The Chainsmokers have balanced pop trends with their electronic roots. The analysis will highlight their consistent use of high-energy tracks and their exploration of emotional tone (valence) and danceability.

This paper is structured as follows: Section 2 presents the data sources, methodologies, and visualization of the results. The findings are then discussed in Section 3, and additional detailed information is provided in the appendix at Section A.

2 Data

This project is guided by Rohan Alexander’s work and his book on data analysis (Alexander 2022). The data used in this paper were downloaded, cleaned, and analyzed using the programming language R (Team 2024) and the following R packages: **dplyr** (Wickham 2024a), **usethis** (Bryan and Wickham 2024), **magrittr** (Bache and Wickham 2024), **spotifyr** (Thompson 2024), **tidyverse** (Wickham 2024b), **lubridate** (Grolemund and Wickham 2024), **here** (Muller 2024), and **readr** (Wickham and Hester 2024).

2.1 Source

Data for this paper was obtained from the Spotify API (AB 2024). The Spotify API provides comprehensive track-level data, including details about acoustic features, popularity metrics, and track metadata for artists. Using the **spotifyr** package (Thompson 2024), we were able to query The Chainsmokers’ discography and retrieve key information such as track-level audio features and album release dates.

The dataset includes all albums released by The Chainsmokers, such as *Memories...Do Not Open*, *Sick Boy*, and *World War Joy*. Each track in the dataset is associated with its album release date, enabling a temporal analysis of their musical evolution.

2.2 Measurement

Spotify provides several audio features that are key to understanding the musical characteristics of tracks. These features, while not exhaustively described by Spotify, are commonly interpreted as follows:

- **Acousticness:** A measure of the likelihood of a track being acoustic. Higher values indicate that the track is more acoustic and uses non-electronic instruments.
- **Energy:** A measure of intensity and activity, capturing the dynamic and energetic feel of a track. High energy typically represents fast-paced, intense music.
- **Valence:** A measure of musical positiveness. Tracks with higher valence are considered more positive and cheerful, while lower valence suggests a more somber or melancholic tone.
- **Danceability:** A measure of how suitable a track is for dancing, based on its rhythm stability, tempo, and beat strength.
- **Speechiness:** Speechiness quantifies the presence of spoken words in a track. Higher values indicate a greater presence of speech, as found in rap or spoken word tracks.

Further details about the data processing and cleaning steps can be found in the appendix at [Section A](#).

2.3 Results

This section presents the analysis of key audio features from The Chainsmokers' albums, focusing on **energy**, **valence**, **danceability**, and **acousticness**. The data visualizations illustrate how these features have evolved over time and provide insight into the musical trends that define The Chainsmokers' sound.

2.3.1 Energy Across Albums

Energy is a crucial feature of The Chainsmokers' music, reflecting the intensity and dynamic nature of their tracks. As shown in [Figure 1](#), The Chainsmokers have consistently maintained high energy levels across their albums. Tracks from albums such as *Memories...Do Not Open* and *World War Joy* exhibit energy values that typically remain above 0.7, indicating the fast-paced, upbeat nature of their music. This trend is reflective of their electronic dance music (EDM) roots, where high energy is a signature element.

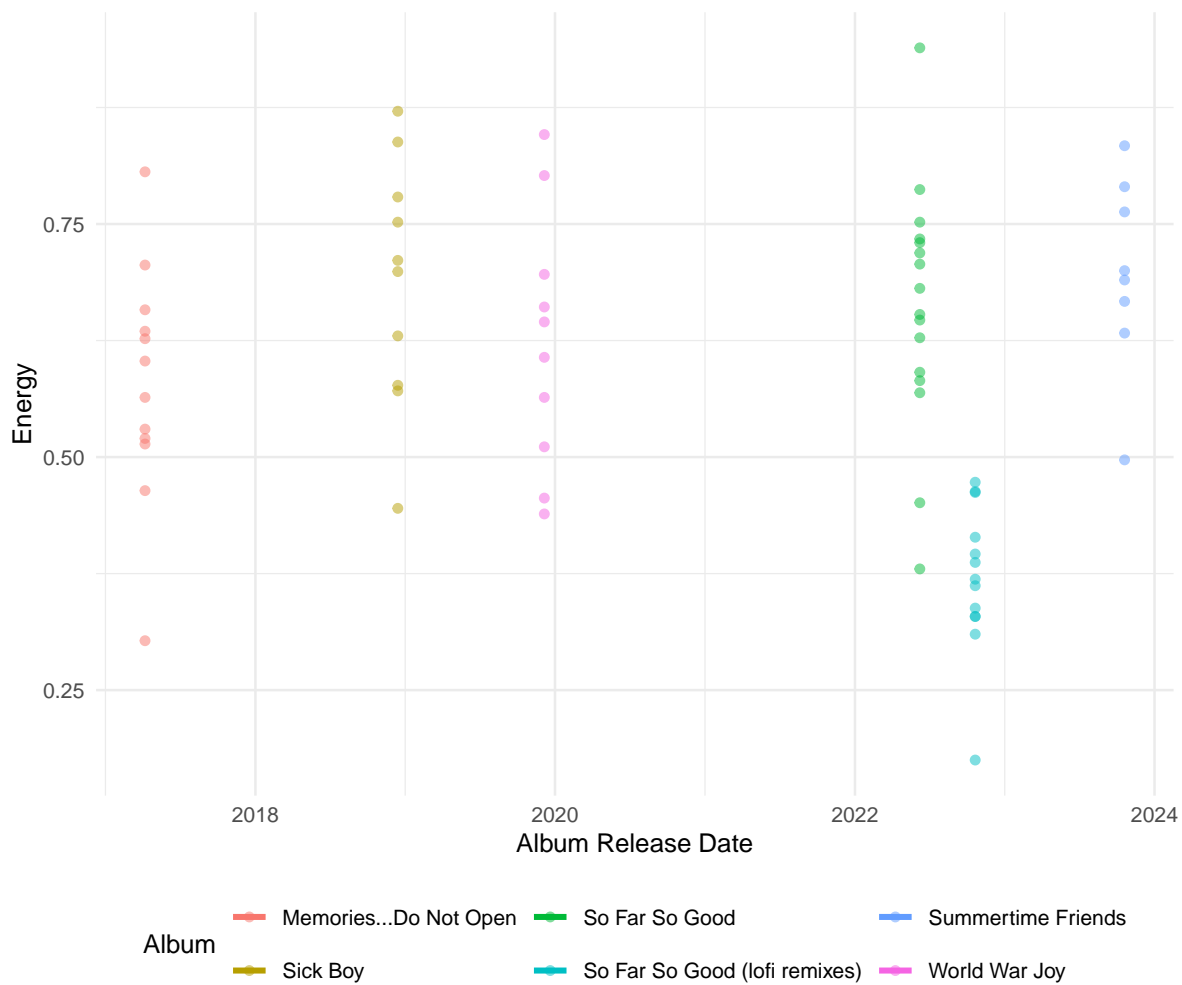


Figure 1: Energy Levels of The Chainsmokers' Tracks Over Time (2014-2023)

Figure 1 illustrates the consistency in energy across The Chainsmokers' discography, reinforcing their reputation for creating music with high intensity and dynamic beats.

Valence measures the emotional positivity of a track. A higher valence score suggests a more cheerful, upbeat tone, while lower scores indicate a more somber or melancholic mood. As shown in Figure 2, there is a slight upward trend in valence in The Chainsmokers' music, especially in more recent albums like *So Far So Good* and *Summertime Friends*. This gradual increase may reflect a subtle shift toward more positive or varied emotional tones in their music, signaling an evolution in their thematic and production choices.

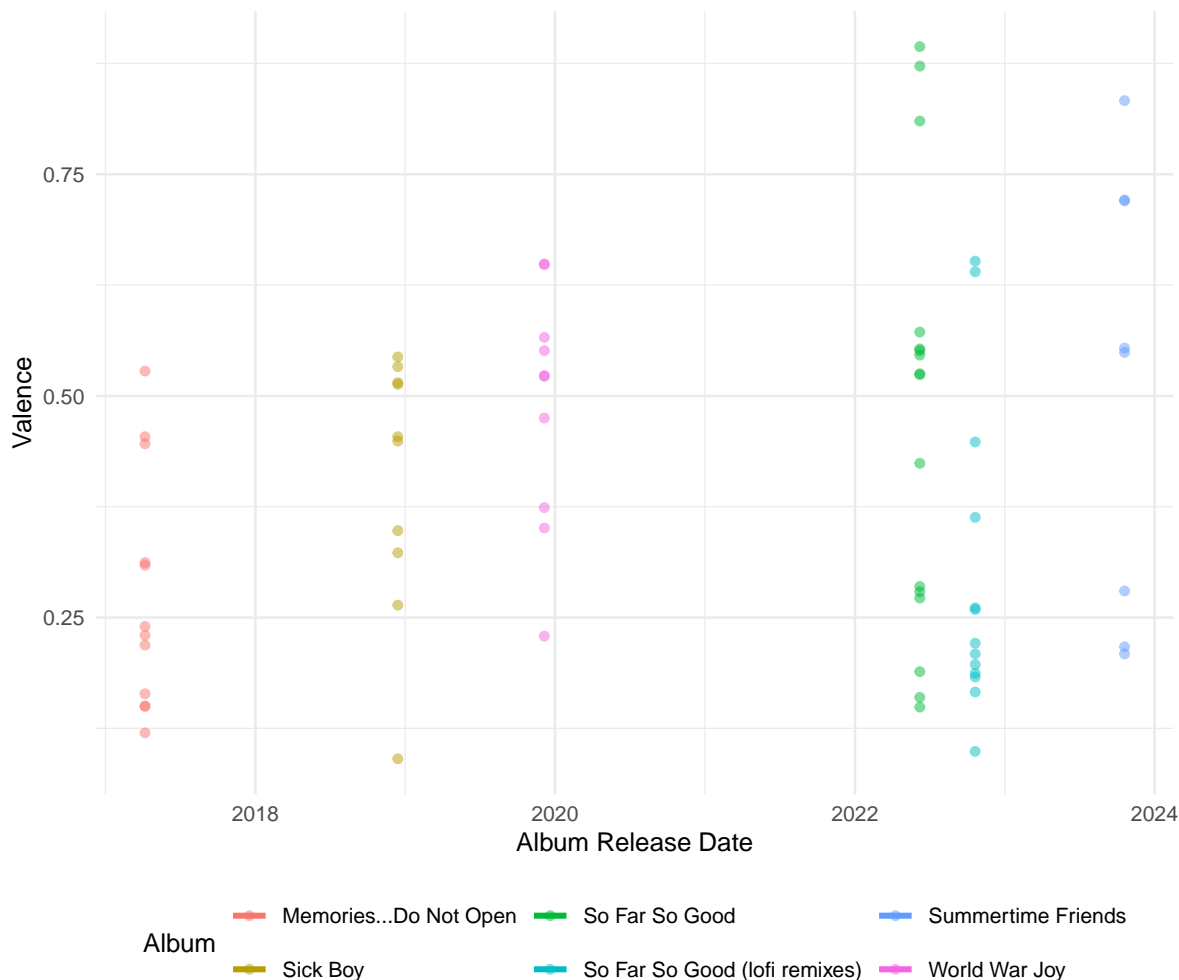


Figure 2: Valence of The Chainsmokers' Tracks Across Albums

Figure 2 illustrates how The Chainsmokers' emotional tone has evolved slightly, showcasing a more positive shift over time across their discography.

The Chainsmokers' music is typically characterized by its suitability for dancing, with danceability being a key feature in their tracks. As shown in Figure 3, their tracks maintain high danceability scores, with values above 0.6 across most albums. This reflects their commitment to creating music that is rhythmically engaging and suited for dance floors.

On the other hand, acousticness remains low across their discography, indicating a consistent use of electronic production and minimal reliance on acoustic instruments. This reinforces their electronic music identity, as seen in albums like Sick Boy and Memories...Do Not Open, where acousticness scores remain close to 0.

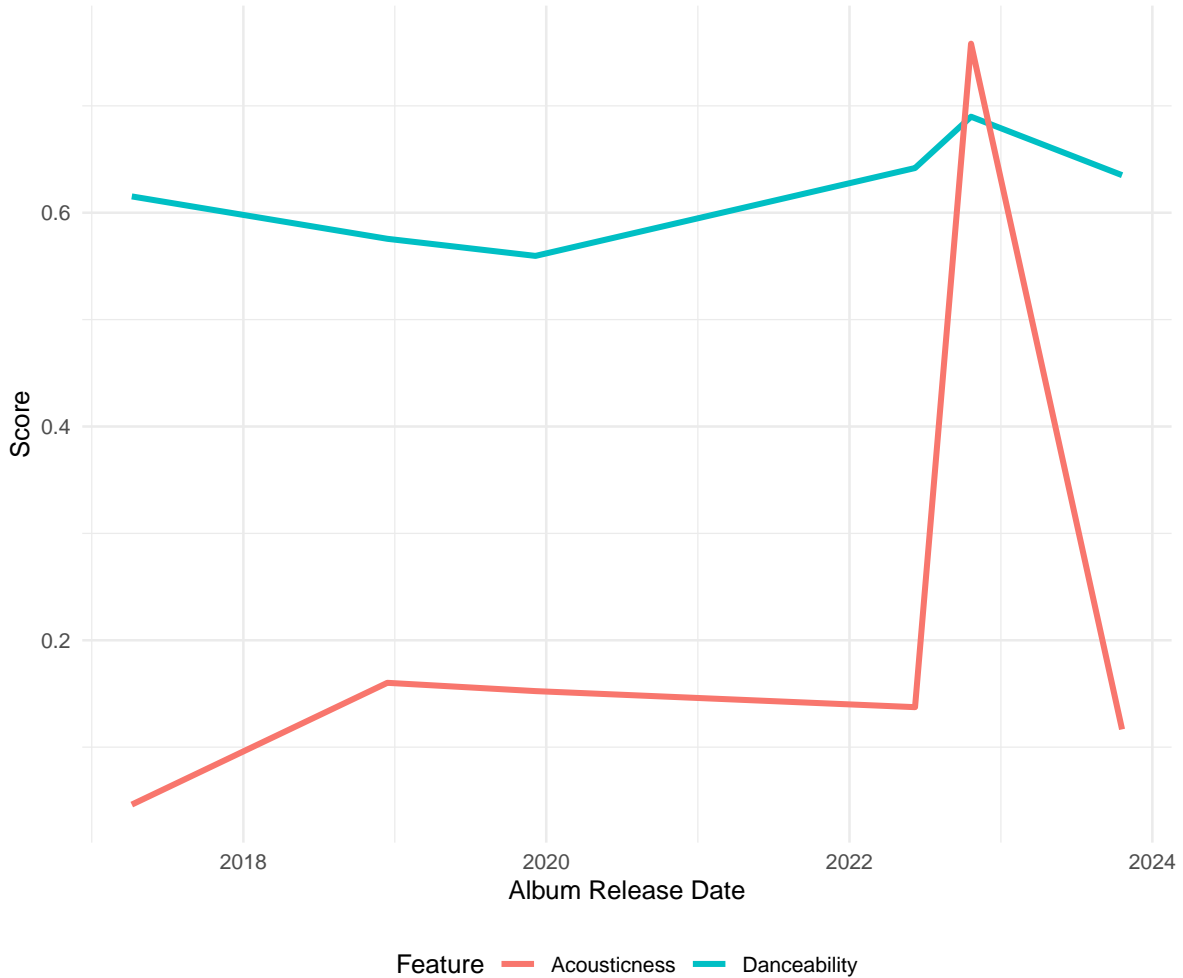


Figure 3: Danceability and Acousticness of The Chainsmokers' Tracks Over Time

Figure 3 shows that while The Chainsmokers' music is highly danceable, their reliance on electronic elements remains central, with low acousticness scores across their albums.

3 Discussion

The analysis of The Chainsmokers’ discography reveals several key trends in their musical evolution. Consistent with their identity as electronic dance music (EDM) producers, The Chainsmokers maintain high **energy** levels across their tracks. As seen in Figure 1, the energetic intensity of their music is a defining characteristic, particularly in albums like *Memories...Do Not Open* and *World War Joy*. This sustained high energy supports their reputation for creating tracks suited for both mainstream pop audiences and electronic music festivals.

In contrast, the **valence** analysis shows a slight downward trend in recent albums, such as *Sick Boy* and *World War Joy*, suggesting a shift toward more complex and potentially melancholic emotional tones (Figure 2). This evolution might reflect broader trends in popular music, where artists often shift from bright, upbeat themes to more introspective and emotionally nuanced topics over time. The Chainsmokers have managed to balance this shift in tone while still maintaining their electronic roots.

The comparison of **danceability** and **acousticness** further emphasizes their electronic foundation. As demonstrated in Figure 3, The Chainsmokers consistently produce tracks with high danceability, ensuring their music remains accessible to club and festival settings. At the same time, **acousticness** remains low, indicating their continued reliance on electronic instruments and production techniques rather than incorporating acoustic elements. This distinction between danceability and acousticness highlights their ability to stay true to their EDM origins while adapting to trends in popular music.

Overall, these findings suggest that while The Chainsmokers have experimented with emotional tone (valence), they have largely preserved the key elements of their signature sound—high energy, danceable beats, and minimal acoustic production. These elements have helped them remain relevant in a constantly evolving music industry.

3.1 Weaknesses and Next Steps

While this analysis provides valuable insights into The Chainsmokers’ musical evolution, there are several limitations that could be addressed in future studies. One notable weakness is the reliance on a relatively small set of audio features (energy, valence, danceability, acousticness, and speechiness). Although these features offer a useful framework for understanding their music, they do not fully capture other important aspects such as **lyrical content**, **instrumentation intricacies**, and **production techniques**. Future research could expand the analysis by incorporating sentiment analysis or lyrical theme extraction to assess how their thematic content has evolved alongside their sound.

Another limitation is the focus on **studio-recorded tracks**. While these tracks represent the core of The Chainsmokers’ discography, they do not account for **live performances** or **remixes**, which often play a crucial role in electronic dance music culture. Including data from

live recordings or remix albums could provide a more holistic understanding of their artistic expression and how their tracks are adapted for different contexts.

An interesting observation in the data is the **consistency in tempo** across their albums, which was not a primary focus of this study. Analyzing tempo could offer additional insights into their pacing and how they maintain listener engagement. Similarly, **loudness** could be another feature worth investigating, especially in relation to how their tracks are mastered for different platforms (e.g., streaming services vs. live performances). These additional features could help to paint a more complete picture of their production choices.

Finally, **collaborations with featured artists** could be explored in more depth. The Chainsmokers frequently collaborate with other artists, which can significantly influence the sound and direction of individual tracks. An analysis of how **collaborations** impact key features like **valence** or **energy** could reveal whether these partnerships have contributed to the emotional shifts observed in their more recent works.

In future work, expanding the dataset to include these additional dimensions—lyrics, live performances, tempo, loudness, and collaborations—could yield a more nuanced understanding of The Chainsmokers’ music. This would provide a richer context for their evolution as artists and offer new perspectives on how they continue to navigate and influence the ever-changing music industry.

Table 1: Raw The Chainsmokers Spotify Data

album_name	album_release_date	danceability	energy	acousticness	valence
Summertime Friends	2023-10-20	0.52	0.69	0.01	0.28
Summertime Friends	2023-10-20	0.75	0.70	0.49	0.72
Summertime Friends	2023-10-20	0.48	0.50	0.08	0.21
Summertime Friends	2023-10-20	0.60	0.67	0.05	0.22
Summertime Friends	2023-10-20	0.62	0.63	0.16	0.55
Summertime Friends	2023-10-20	0.73	0.79	0.01	0.83

Table 2: Processed The Chainsmokers Spotify Data

album_name	album_release_date	danceability	energy	acousticness	valence
Summertime Friends	2023-10-20	0.52	0.69	0.01	0.28
Summertime Friends	2023-10-20	0.75	0.70	0.49	0.72
Summertime Friends	2023-10-20	0.48	0.50	0.08	0.21
Summertime Friends	2023-10-20	0.60	0.67	0.05	0.22
Summertime Friends	2023-10-20	0.62	0.63	0.16	0.55
Summertime Friends	2023-10-20	0.73	0.79	0.01	0.83

A Appendix

A.1 Raw Data Processing

The raw data for this analysis was sourced from the Spotify API using the `spotifyr` (Thompson 2024) package. The initial data includes audio features for The Chainsmokers' tracks, such as danceability, energy, acousticness, valence, speechiness, and tempo. The cleaning process focused on ensuring temporal accuracy by filtering and formatting the album release dates.

A.2 Cleaned Data

To clean the data, tracks missing release dates were removed, and the `album_release_date` was formatted into a consistent year-month-day format using `lubridate::ymd()`. Tracks with missing or invalid audio features were also excluded to maintain data integrity. The cleaned data was saved for further analysis.

A.3 Data Processing Overview

The cleaning process began by loading the raw data from a `.rds` file. To ensure that all tracks included in the analysis were associated with valid albums, any tracks with missing release dates were excluded. The date formatting was standardized using `lubridate::ymd()`,

ensuring consistency across all entries. Additionally, audio features critical to the analysis, such as danceability, energy, acousticness, and valence, were retained while others were excluded for relevance. Finally, the cleaned data was saved in the **analysis_data** folder.

The difference between the raw and cleaned datasets can be observed by comparing Table 1 and Table 2. The final dataset contains only complete and reliable records, ensuring that subsequent analyses are based on valid data.

References

- AB, Spotify. 2024. *Spotify Web API Documentation*. <https://developer.spotify.com/documentation/web-api/>.
- Alexander, Rohan. 2022. *Telling Stories with Data*. <https://tellingstorieswithdata.com/>.
- Bache, Stefan Milton, and Hadley Wickham. 2024. *Magrittr: A Forward-Pipe Operator for r*. <https://magrittr.tidyverse.org/>.
- Bryan, Jenny, and Hadley Wickham. 2024. *Usethis: Automating Package and Project Setup*. <https://usethis.r-lib.org/>.
- Grolemund, Garrett, and Hadley Wickham. 2024. *Lubridate: Make Dealing with Dates a Little Easier*. <https://lubridate.tidyverse.org/>.
- Muller, Kirill. 2024. *Here: A Simpler Way to Find Your Files*. <https://here.r-lib.org/>.
- Team, R Development Core. 2024. *R Programming Language Documentation*. <https://www.r-project.org/>.
- Thompson, Charlie. 2024. *Spotifyr: R Wrapper for Spotify's Web API*. <https://github.com/charlie86/spotifyr>.
- Wickham, Hadley. 2024a. *Dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org/>.
- . 2024b. *Tidyverse: Easily Install and Load the 'Tidyverse'*. <https://www.tidyverse.org/>.
- Wickham, Hadley, and Jim Hester. 2024. *Readr: Read Rectangular Text Data*. <https://readr.tidyverse.org/>.