Analyzing the data of rapper: Pitbull's songs from 2004 to 2024*

Analyzing the data of rapper: Pitbull's songs from 2004 to 2024 to find the relationship between the year of album released and valence, valence and energy, energy and danceability

Yuanchen Miao Ziqi Zhu

October 10, 2024

In this paper, the relationship of year of album released and valence, valence and energy, energy and danceability of rapper Pitbull's song from 2004 to 2024 are analyzed. Pitbull is one of the most famous rapper in the world, lots of people loves rap since rap has its rhythm and beats make people get a feeling of swing. Also, rappers write their lyrics about their feelings or things they wanna talk about. Thus, we are finding the relationship between the year album released and valence to see the emotion change of pitbull, the feeling the song brings and the energy of the song are connected so the data of these two variables are analyzed. Since rap songs are easily making people want to dance, so the energy and danceablity of the song are also analyzed.

Table of contents

*Code and data are available at: https://github.com/NevaeH-9/vincentspotify

1 Introduction

The R programming language (R Core Team (2023)) and tidyverse (Wickham et al. (2019)) were used to clean data, generate tables and create graphs, package knitr (Xie (2023)) and (Zhu (2024)) is used for adding caption to table. The data were gathered and downloaded form spotify by using package spotifyr (Thompson et al. (2022)) and devtools (Wickham et al. (2022)).

2 Data

The data is obtained from the Spotify (Thompson et al. 2022), all of the songs from 2004 to 2024 and "album_release_date", "danceability", "valence" and "energy" are chosen to analyze. This table Table 1 shows the first several rows of the data we use.

Table 1: 2004 - 2024 data of pitbull's song

album_release_date	danceability	valence	energy
2024-02-16	0.798	0.696	0.956
2024-02-16	0.842	0.911	0.857
2024-02-16	0.869	0.935	0.832

3 Result

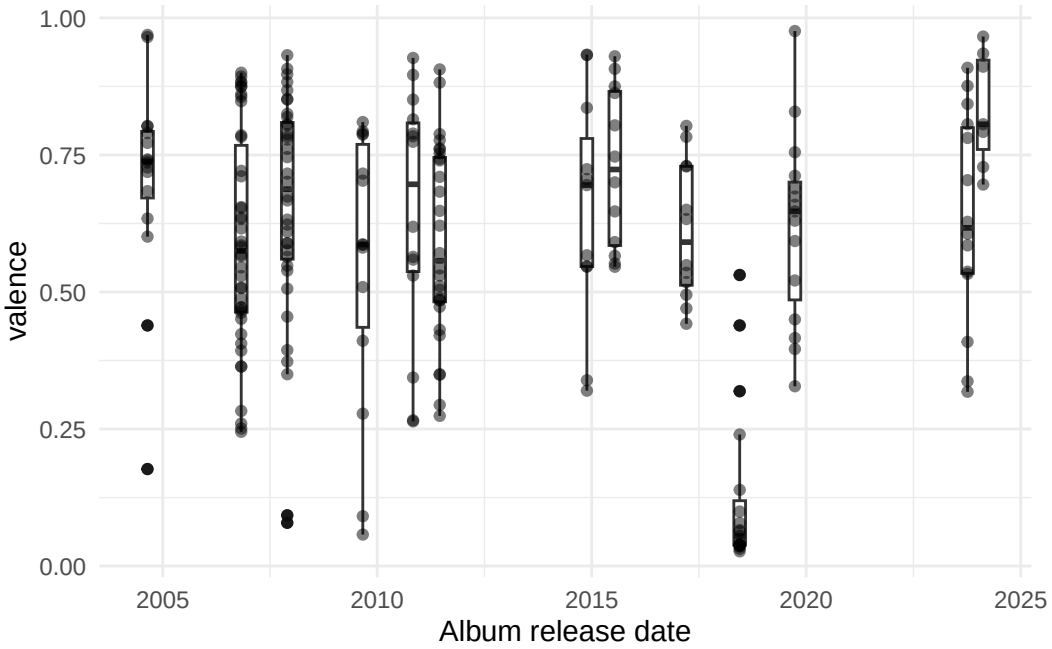

Figure 1: Relationship between year album released and valence

a box plot shows the valence (musical positivity) of Pitbull's songs over time, positive in general but become extremely neagtive around 2018

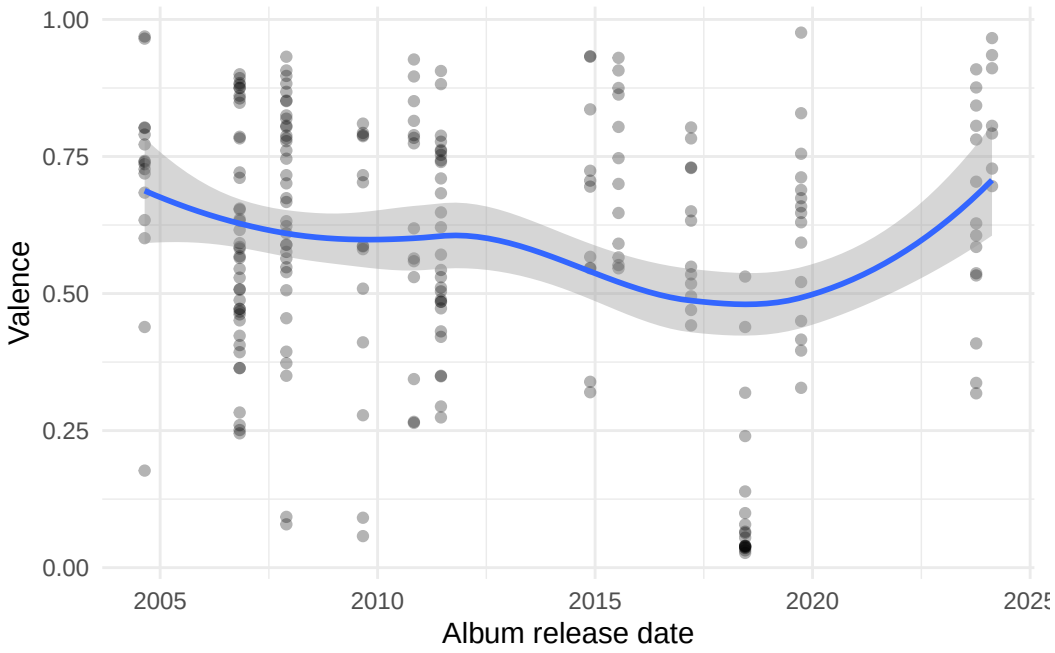

Figure 2: Relationship between year album released and valence

scatter plot of album released and valence, showing a decrease from 2020 to 2018, and an increase after 2018.

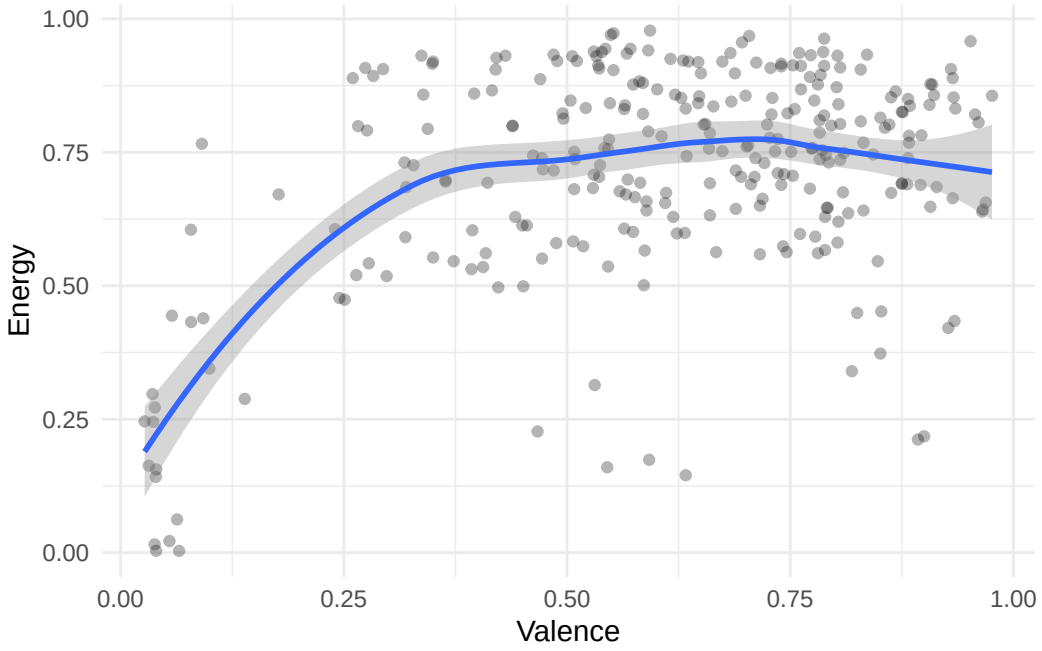

Figure 3: Relationship between valence and energy

scatter plot of relationships between valence and energy, showing a positive correlation when Valence < 0.5.

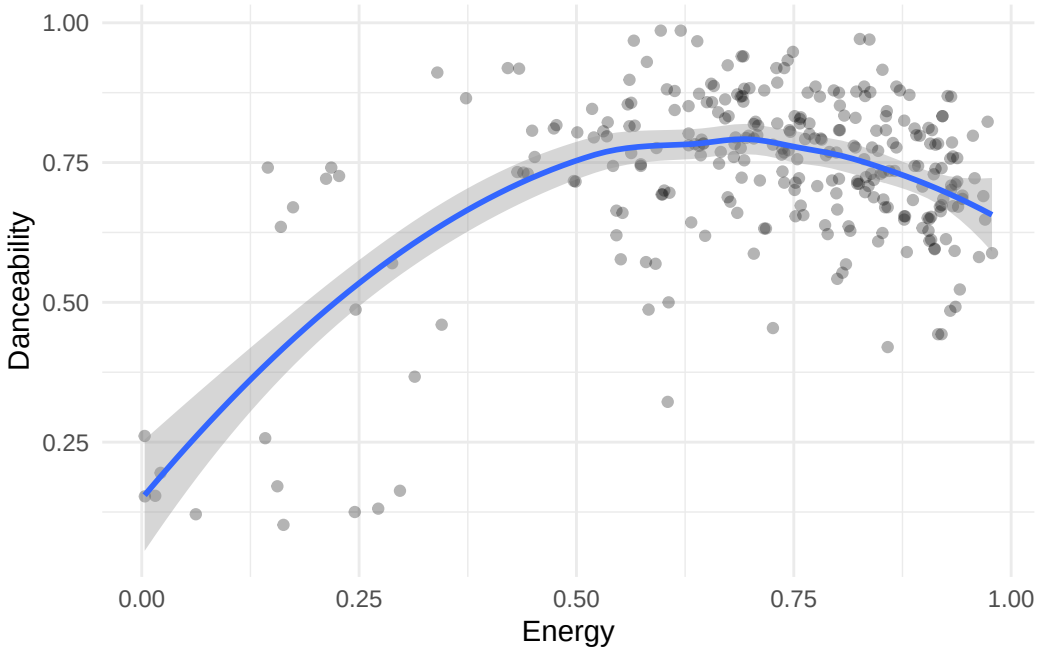

Figure 4: Relationship between energy and danceability

scatter plot of relationships between Energy and Danceability, showing a positive correlation when Energy < 0.5.

4 Discussion

From the graph Figure 1, we can see that In the early years (pre-2010), there is a wide range of valence values, with some songs having very high positivity (close to 1) and others showing lower values, suggesting a mix of upbeat and moody tracks. As time progresses (between 2010 and 2020), the valence remains fairly consistent, though the range in some albums narrows, indicating a tendency toward maintaining a certain upbeat style. Around 2020, there seems to be a drop in valence for a few songs, as seen by some outliers with lower valence scores. This could suggest Pitbull released some less positive songs during this period, though most of his music remained relatively upbeat.

When we fit a smoothed line to the graph as shown in Figure 2, the trend is more obvious, suggests that Pitbull's music started off with a relatively high valence (around 0.75) in the early 2000s, and then gradually declined, reaching a low point around 2018, where the valence dipped below 0.5. After that valence increases again, climbing back toward 0.75 until recent years.

Other two graphs that visualize relationships between valence, energy, and danceability of Pitbull's songs Figure 3 and Figure 4 have similar patterns. Both of them shows there is a strong positive correlation between two variables when one varaible is smaller than 0.5. Figure 3 shows that positive songs tend to have higher energy, but beyond a certain point(around 0.75), increasing valence doesn't necessarily result in more energetic. Figure 4 demonstrates that moderately high energy songs(with energy range from 0.5 to 0.8) are the most danceable, but extremely high energy might reduce danceability.

References

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Thompson, Charlie, Daniel Antal, Josiah Parry, Donal Phipps, and Tom Wolff. 2022. *Spotifyr: R Wrapper for the 'Spotify' Web API*. https://github.com/charlie86/spotifyr.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Jim Hester, Winston Chang, and Jennifer Bryan. 2022. *Devtools: Tools to Make Developing r Packages Easier*. https://CRAN.R-project.org/package=devtools.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://CRAN.R-project.org/package=knitr.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.