Data Analysis For Artist: Ed Sheeran*

Exploring Song Durations Across Ed Sheeran's Albums Using Spotify API

Qingyang Feng

October 10, 2024

This paper analyzes the trends in song durations across Ed Sheeran's discography using data from the Spotify API. By visualizing album release dates and song lengths, we explore the variability in his music over time, highlighting shifts in his artistic style.

Introduction

Music data analysis provides insightful perspectives on artists and their creative trajectories. In this analysis, we focus on Ed Sheeran, a renowned singer-songwriter whose discography spans multiple albums and years. We use R (R Core Team (2023)) and several key packages to analyze the relationship between the release dates of Ed Sheeran's albums and the duration of his songs.

We retrieved Ed Sheeran's audio data from Spotify's Web API using the spotifyr package (Thompson et al. (2022)), which provides a convenient wrapper for interacting with Spotify's extensive dataset. To manage file paths efficiently, we used the here package (Müller (2020)), and for overall project management, usethis (Wickham et al. (2024)) was employed. The analysis is supported by the tidyverse ecosystem (Wickham et al. (2019)), enabling the data transformation and visualization required to gain insights into the dataset.

This paper aims to explore whether any patterns in song duration can be detected across Ed Sheeran's discography and whether these patterns reveal any interesting trends about his musical evolution.(Alexander (2023))

---

*Code and data are available at: [https://github.com/Fqy10987/Spotify-Data-Analysis.git]

Result

The scatter box The scatter plot (see in Appendix-A) shows that Ed Sheeran's song durations have generally become more varied over time. While his earlier songs around 2015 tend to cluster around a similar duration, more recent releases show a broader range. Notably, some tracks released after 2020 have durations significantly longer than 400,000 milliseconds, suggesting a shift towards longer compositions in his recent work. This could reflect changes in his creative process, experimentation with different genres, or adaptations to trends in the music industry.

Discussion

Duration of Songs Across Ed Sheeran's Discography The boxplot in Figure 1 illustrates the song duration (in milliseconds) for Ed Sheeran's tracks across his albums. Each point represents a track, and the jitter helps to visualize the spread of the song lengths. The x-axis shows the release dates of albums, while the y-axis represents the duration of the songs in milliseconds.

From the plot, it's noticeable that song durations tend to cluster between 200,000 ms and 300,000 ms, which corresponds to 3-5 minutes, a typical duration for popular music tracks. However, there are notable outliers, especially for albums released post-2020, where song durations have increased. These might suggest experimental tracks or collaborations with other artists, potentially extending the standard pop song format.

Interestingly, earlier albums like + and x (multiply) have tightly clustered durations, suggesting a consistent style in terms of song length. In contrast, more recent albums like = (Equals) show a wider variation in song durations, indicating a broader artistic range.

Impact of Release Dates on Song Duration There seems to be no significant trend that links the album release date directly to the song length. However, post-2020 albums display more variability in song durations, possibly influenced by changing industry standards or Ed Sheeran's experimentation as he matured as an artist.

The analysis could be expanded by comparing these trends to other artists of the same genre or even extending the analysis to additional audio features like tempo, energy, and danceability, which would provide a fuller picture of the evolution of Ed Sheeran's music.

Appendix

Appendix-A
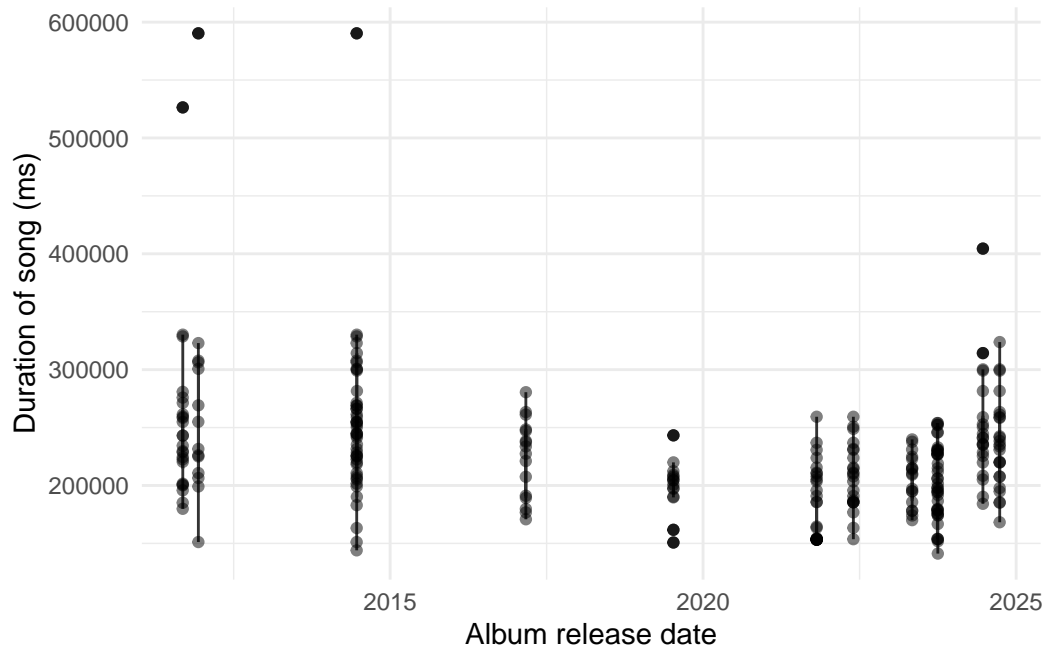
Reference

Alexander, Rohan. 2023. *Telling Stories with Data: With Applications in r.* Chapman; Hall/CRC.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files.* https://CRAN.R-project.org/package=here.

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, Jennifer Bryan, Malcolm Barrett, and Andy Teucher. 2024. *Usethis: Automate Package and Project Setup.* https://CRAN.R-project.org/package=usethis.