

Data Analysis of Spotify*

Dawson Li, Betty Liu

October 10, 2024

For this analysis we used the following packages: R Core Team (2023), Wickham et al. (2019), Müller (2020), Wickham et al. (2024), Thompson et al. (2022), and Wickham (2016). The dataset is downloaded from: Spotify (n.d.). The code were used from Alexander (2023).

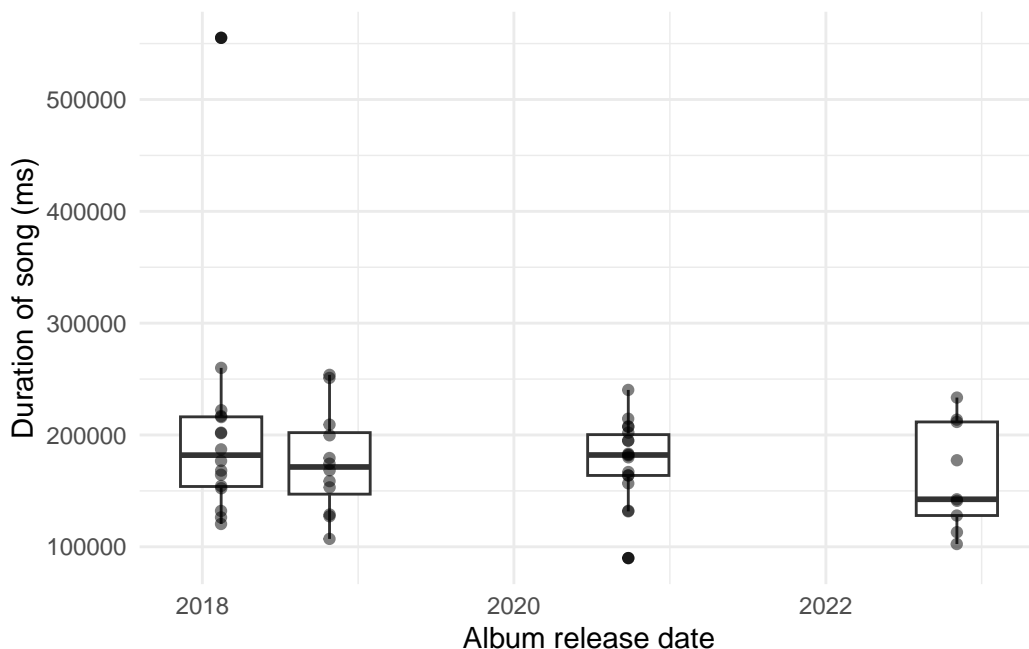


Figure 1: Duration of the Song vs Album Release Date

Boxplot showing the distribution of song durations in milliseconds for album release years 2018, 2020, and 2022. On the y-axis, song durations range from approximately 100,000 to over 500,000 milliseconds, with most songs between 100,000 and 300,000 milliseconds. Each year is represented by a boxplot summarizing the median, interquartile range (IQR), and outliers for

*Code and data are available at: https://github.com/betty-tGM2/Data_analysis_spotify.git

that year. Around 2018, the median song duration was approximately 200,000 milliseconds, with one extreme outlier at over 500,000 milliseconds. The IQR shows that most songs fall within a narrow range. In 2020, the median is slightly lower, with less variation in song duration and no extreme outliers. By 2022, the median is similar to 2018, but the IQR is wider, indicating greater variation in song length, as well as some outliers for shorter songs.

What's particularly interesting about this graph is that the variation in song duration has increased over time, as shown by the widening IQR in 2022. Despite this, the median song duration has remained relatively consistent over the years. These patterns suggest that while there may be an overall trend in song duration consistency, individual deviations are becoming more pronounced, particularly around 2022.

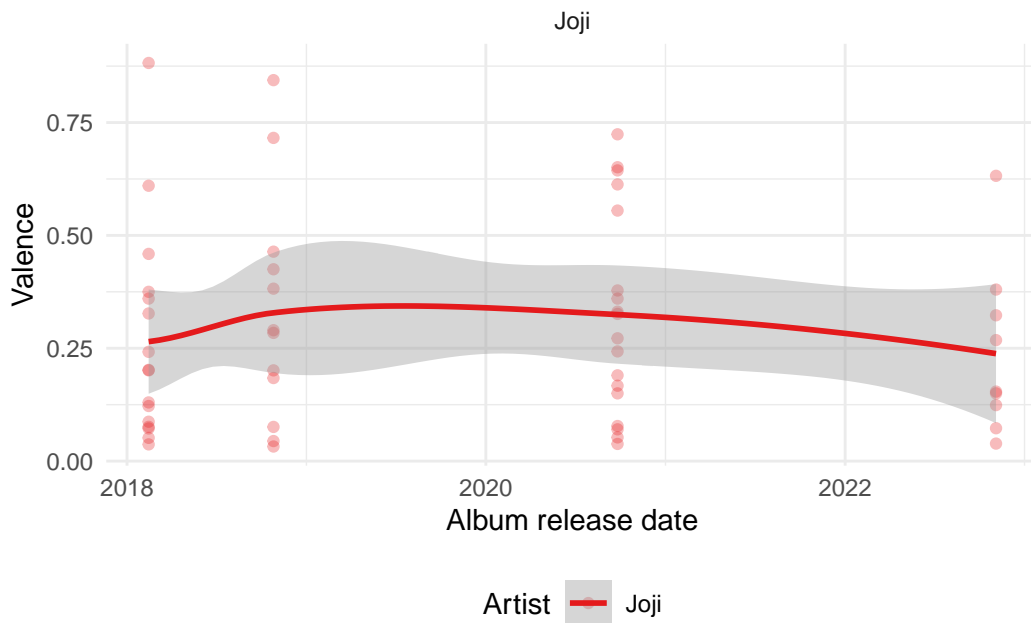


Figure 2: Valencevs Album Release Date

The line graph shows the relationship between valence and album release date for the artist Joji from 2018 to 2022. The y-axis represents valence, ranging from 0.00 to 1.00, while the x-axis categorizes the data by album release year. Joji's music in 2018 had a wide range of valence scores, with most songs clustered between 0.00 and 0.30, indicating a melancholic tone, and relatively narrow confidence intervals indicating consistent emotional themes. By 2020, the average valence had risen slightly, although songs still mostly fell between 0.25 and 0.50, with increased variability. In 2022, the valence scores showed a slightly downward trend, indicating that Joji's music has become more melancholic over time, with wider confidence intervals, implying greater variability in emotional tone. What's particularly interesting about this graph is that the valence has gradually decreased from 2020 to 2022, reflecting a shift in musical style toward a more melancholic tone as Joji's albums progress.

Appendix

Data were downloaded from Spotify, and method were used from Alexander (2023).

References

- 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/>.
- Spotify. n.d. *Joji: Spotify Artist Page*. https://open.spotify.com/artist/3MZsBdqDrRTJihTHQrO6Dq?si=-QF6t_TtQTqd8aJkyrz30g&nd=1&dlsi=cc5e58aca6f5462b.
- 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. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolmund, 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>.