Lab 05 – MATH 240 – Computational Statistics

Yuliia Heleveria
MATH 240 Lab A
Mathematics
`yheleveria@colgate.edu`

02/25/2025

Abstract

We want to analyze the contributions of The Front Bottoms, Manchester Orchestra, and All Get Out to their collaborative song "Allentown" using audio analysis. Therefore, we created a batch file to speed up data collection of audio features. We extracted musical characteristics of .WAV files from three bands, including valence, instrumental and acoustic features, emotional impact, loudness, and lyrics analysis. We loaded and cleaned the data to perform statistical analysis for features of each band's tracks and their similarity with "Allentown".

Keywords: Installing and using libraries; creating, cleaning, and merging data frames; statistical analysis; data visualization.

1 Introduction

"Allentown" is a song released in 2018 by collaboration of The Front Bottoms and Manchester Orchestra and contributions from All Get Out. We would like to inspect which band made the most contribution to this song. We purchased all pre-"Allentown" releases, totaling 181 tracks, including Allentown itself. Using Essentia (Bogdanov et al., 2013), we extracted music data about given songs to determine the primary contributor to "Allentown". We summarized the style and characteristics of tracks, such as valence, arousal, emotions, and absence of voice or acoustic features, belonging to each band to determine stylistic features that resonate the most with "Allentown".

Section 2 covers data collection and processing. Section 3 presents data analysis, and Section 4 discusses findings from comparing the stylistic features of three band's tracks with "Allentown".

2 Methods

We obtained Essentia models (Alonso-Jiménez et al., 2020) of 181 songs for analysis. We used `stringr` package (Wickham, 2023) to create a batch file that creates command line prompts for each track. With the use of the batch file, we extracted data from .WAV files and then Essentia model for each track about features such as tempo in beats, average loudness, valence, arousal, aggressiveness, acoustic and instrumental sound.

Using `jsonlite` package for `R` (Ooms, 2014), we extracted stylistic features of provided songs. We loaded JSON files containing music analysis into `R` and extracted mean of spectral energy, dancebility, tempo in beats per minute, musical key, musical mode and duration of the track in seconds. We also used LIWC text analysis tool (Boyd et al., 2022) to extract features that describe thoughts, feelings, and personality traits in the tracks. We merged the data frames and saved them as two .csv files, as training and testing data, combining lyrics and audio analysis.

Using `tidyverse` package (Wickham et al., 2019), we summarized data numerically for each feature of interest. We computed minimum, maximum, lower and upper fences for each tracks's feature for All Get Out, Manchester Orchestra, and The Front Bottoms to determine if "Allentown"'s features are within range or outlying for each band's songs. Based of the statistical summaries, we computed visual aids by creating a table using `xtable` library (Dahl et al., 2019) and column plots using `tidyverse` to determine how "Allentown" differs from music of each band based on the count of how many features of "Allentown" are within range and out of range for each band.

3 Results

We analyzed stylistic features of "Allentown" in comparison to 180 tracks from the contributing bands. Out statistical summaries and visualizations (see Figure 4) show that Manchester Orchestra has the most similar musical features to "Allentown", depicting the band's major contribution to creation of the track. Manchester Orchestra has the largest number of features that "Allentown" is within the range of. Manchester Orchestra also has the fewest features that are out of range of "Allentown". The result suggest that while each band contributes unique stylistic features, Manchester Orchestra has influenced creation of "Allentown" the most.

4 Discussion

We performed preliminary data analysis by creating `boxplots` that represent relation of data from "Allentown" in comparison to median values of the same data collected from the three contributing artists. `Boxplots` comparing "Allentown" to media values of each band suggest artists' mixed influence

on track's features, with Manchester Orchestra being most similar to "Allentown" in overall loudness. The Front Bottoms tracks have the median word count as "Allentown"'s word count . In regard to aggressiveness, "Allentown" is the most similar to tracks by All Get Out. Each band adds distinct stylistic elements to "Allentown", but further analysis is needed to quantify their contributions.
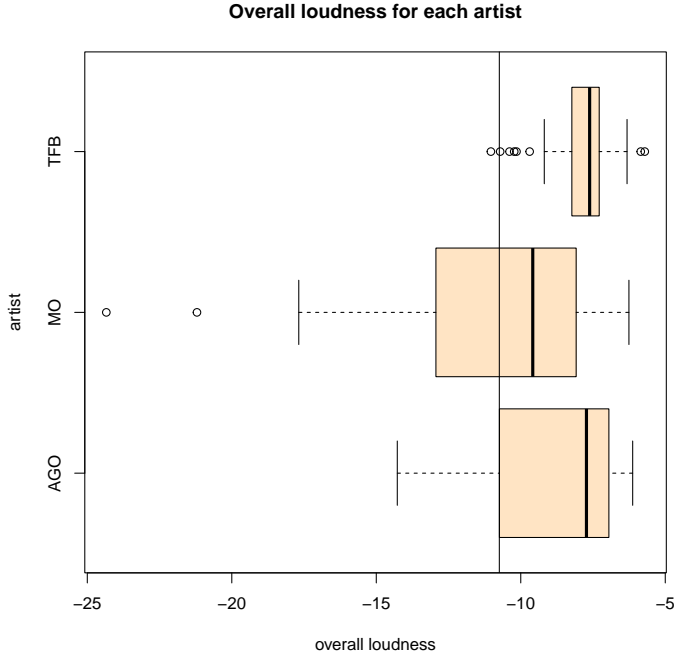

Overall loudness for each artist

Figure 1: Vertical line represents "Allentown"

Overall word count for each artist
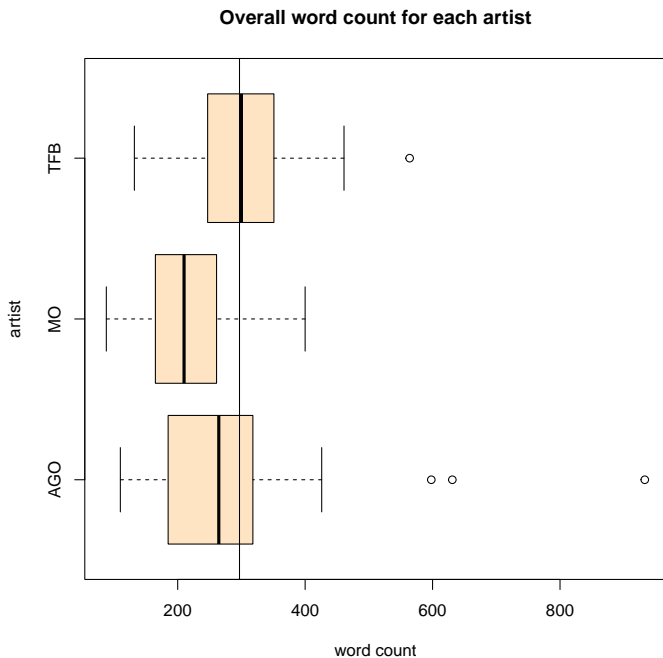
Figure 2: Vertical line represents "Allentown"
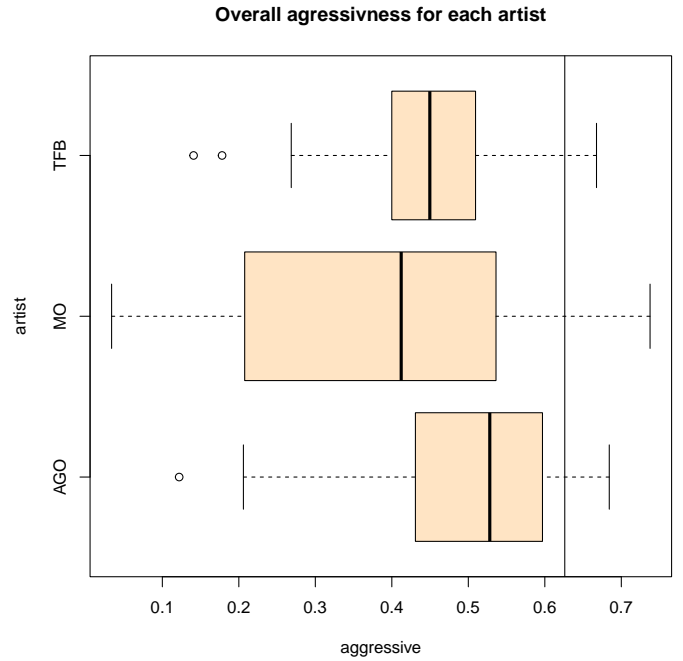

Overall agressivness for each artist

Figure 3: Vertical line represents "Allentown"

Manchester Orchestra's audio features are the most similar to "Allentown" (see Table 1). All Get Out has the second most similar audio features to "Allentown" because it has fewer out of range and more in range features than The Front Bottoms has (see Figure 4). Manchester Orchestra has the fewest out of range and the most within range features, indicating that this band contributed the most to creating "Allentown".

References

Alonso-Jiménez, P., Bogdanov, D., Pons, J., and Serra, X. (2020). Tensorflow audio models in essentia. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 266–270. IEEE.

Bogdanov, D., Wack, N., Gómez Gutiérrez, E., Gulati, S., Boyer, H., Mayor, O., Roma Trepat, G., Salamon, J., Zapata González, J. R., Serra, X., et al. (2013). Essentia: An audio analysis library for music information retrieval. In *Britto A, Gouyon F, Dixon S, editors. 14th Conference of the International Society for Music Information Retrieval (ISMIR); 2013 Nov 4-8; Curitiba, Brazil.[place unknown]: ISMIR; 2013. p. 493-8*. International Society for Music Information Retrieval (ISMIR).

Boyd, R. L., Ashokkumar, A., Seraj, S., and Pennebaker, J. W. (2022). The development and psychometric properties of liwc-22. *Austin, TX: University of Texas at Austin*, 10:1–47.

Dahl, D. B., Scott, D., Roosen, C., Magnusson, A., and Swinton, J. (2019). *xtable: Export Tables to LaTeX or HTML*. R package version 1.8-4.

Ooms, J. (2014). The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv:1403.2805 [stat.CO]*.

Wickham, H. (2023). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.5.1.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686.

5 Appendix

Category	All.Get.Out	Manchester.Orchestra	The.Front.Bottoms
Out of Range	22	3	30
Outlying	17	11	11
Within Range	158	183	156

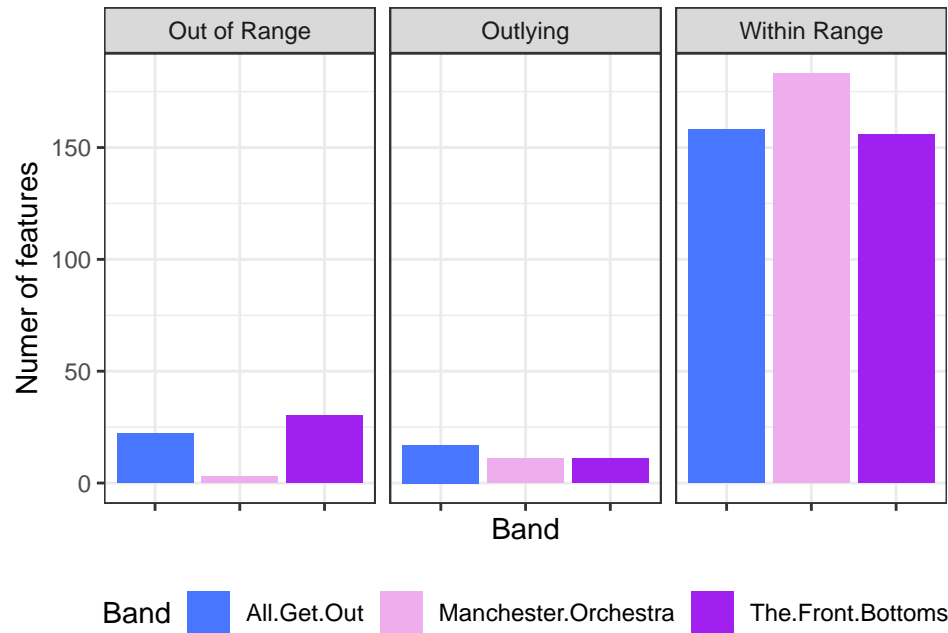Table 1: Comparison of "Allentown"'s audio features with the range of bands' features

Figure 4: Comparison of "Allentown"'s audio features with the range of bands' features