Lab 03 – 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 for audio features. The batch file contains 181 commands referring to 77 tracks by Manchester Orchestra, 61 by The Front Bottoms, 42 by All Get Out, and one collaborative track "Allentown". We analyzed musical characteristics of .WAV files from three bands, including valence, instrumental and acoustic features, emotional impact, loudness, and lyrics. We loaded and cleaned the data to perform analysis regarding characteristics of each band's tracks and their similarity with "Allentown".

Keywords: Installing and using libraries; creating, cleaning, and merging data frames; looping structure; vectors and lists operations.

1 Introduction

"Allentown" is a song released in 2018 by collaboration of The Front Bottoms and Manchester Orchestra. 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. We use Essentia (Bogdanov et al., 2013) to analyze, synthesize, and describe data about 181 songs to determine the musician who makes the most contribution to "Allentown". We analyze 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".

This paper is structured as follows: In Section 2, we summarize the process of collecting, loading, and cleaning of data. Section 3 presents the outcome of data manipulation and analysis. In Section 4, we interpret the findings obtained by analyzing audio data given from 180 tracks of The Front Bottoms, Manchester Orchestra, and All Get Out compared to "Allentown".

2 Methods

For the purpose of data collection, we were given Essentia models (Alonso-Jiménez et al., 2020) of 181 songs of interest. 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 analyzed .WAV files for track characteristics, such as tempo in beats and average loudness.

Using `jsonlite` package for `R` (Ooms, 2014), we extracted stylistic features of provided songs. We used `fromJSON()` to load JSON file containing music analysis into `R` and analyze average loudness, mean of spectral energy, dancebility, tempo in beats per minute, musical key, musical mode and duration of the track in seconds. We aslo used LIWC text analysis tool (Boyd et al., 2022) to extract features that describe thoughts, feelings, and personality traits.

We loaded the data frame given from Essentia Model for 181 songs, including songs' valence, arousal, aggressiveness, acoustic and instrumental features. At the end, we merged three data frames and saved them to .csv files to combine output about lyrics and Essentia 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 feature for All Get Out, Manchester Orchestra, and The Front Bottoms to determine if "Allentown"'s feature is within range or outlying for each band. Based of the statistical summaries, we computed visual aids of a table using `xtable` library (Dahl et al., 2019) and column plots using `tidyverse` to determine how "Allentown" differs of 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

By creating a batch file, we significantly decreased the amount of time for data collection from multiple tracks at the same time using Essentia command line prompts. "Allentown" and its data was written separately into a different .csv file for data testing and determining the main band's contributions based on audio and lyrics analysis. The resulting .csv files contain data regarding spectral energy, tempo, dissonance, danceability, valence, happiness, tone, pronouns, etc. to derive conclusions of stylistic dominant features of each band in comparison of style with "Allentown". Three side-by-side column plots that demonstrate the count of features for each band that are outlying or that lie within range or out of range of "Allentown" demonstrate that Manchester Orchestra has the largest contribution to the collaborative track. Manch-

ester Orchestra has the most musical features that lie within range of "Allentown" and the fewest features that are out of range.

4 Discussion

We performed preliminary data analysis by creating `boxplot` that represent relation of data from "Allentown" in comparison to median values of the same data collected from three artists who participated in the song creation. Comparing the vertical line that represented "Allentown" in each plot, we cannot come to a conclusion of which artist has the closest audio style to "Allentown". In terms of overall loudness, "Allentown" is most similar to style of Manchester Orchestra, but "Allentown" has the overall word count of the median of The Front Bottoms tracks. In regard to aggressiveness, "Allentown" is the most similar to tracks by All Get Out. It is evident that three bands stylistically contribute to different audio feaitres of "Allentown", but more data analysis has to be performed to determine which band had the most contribution to creating the collaborative song.
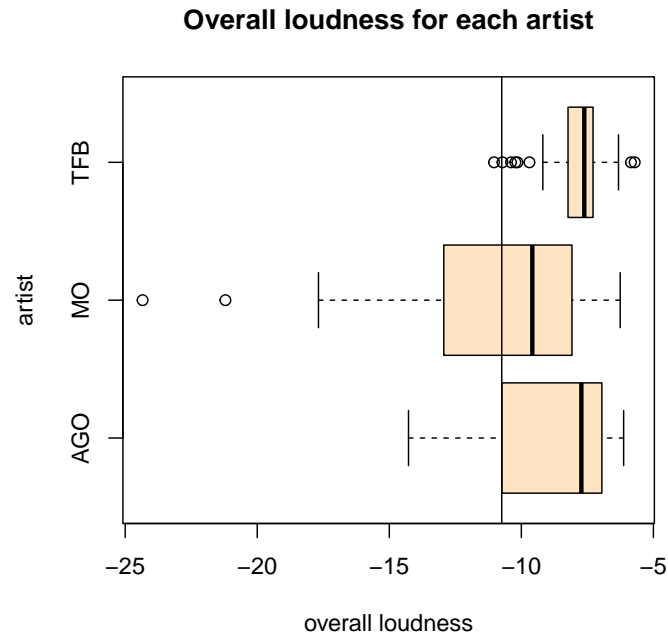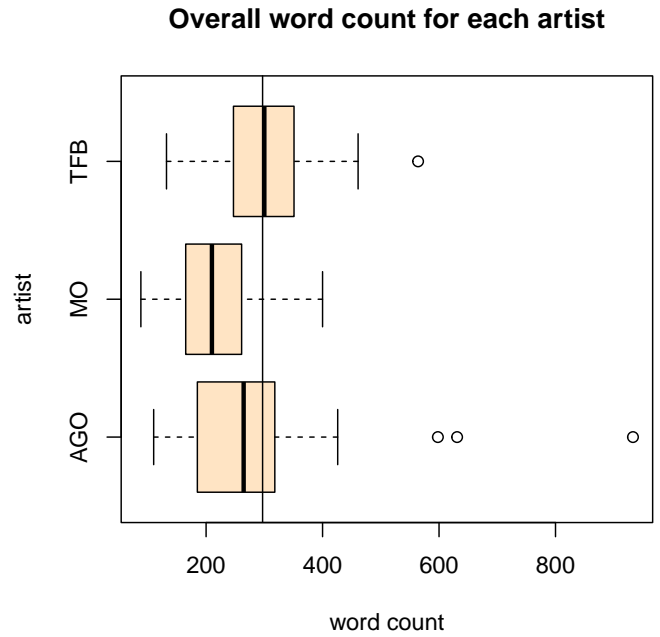
Overall word count for each artist

Figure 2: Vertical line represents "Allentown"

Overall agressivness for each artist

Figure 3: Vertical line represents "Allentown"

After analysis of all given features, it is evident that Manchester Orchestra has the most similar audio features to "Allentown". 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 (4). However, Manchester Orchestra has significantly fewer out of range features than the other two bands, which leads us to

Overall loudness for each artist

Figure 1: Vertical line represents "Allentown"

conclude that Manchester Orchestra has the largest contributions to "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

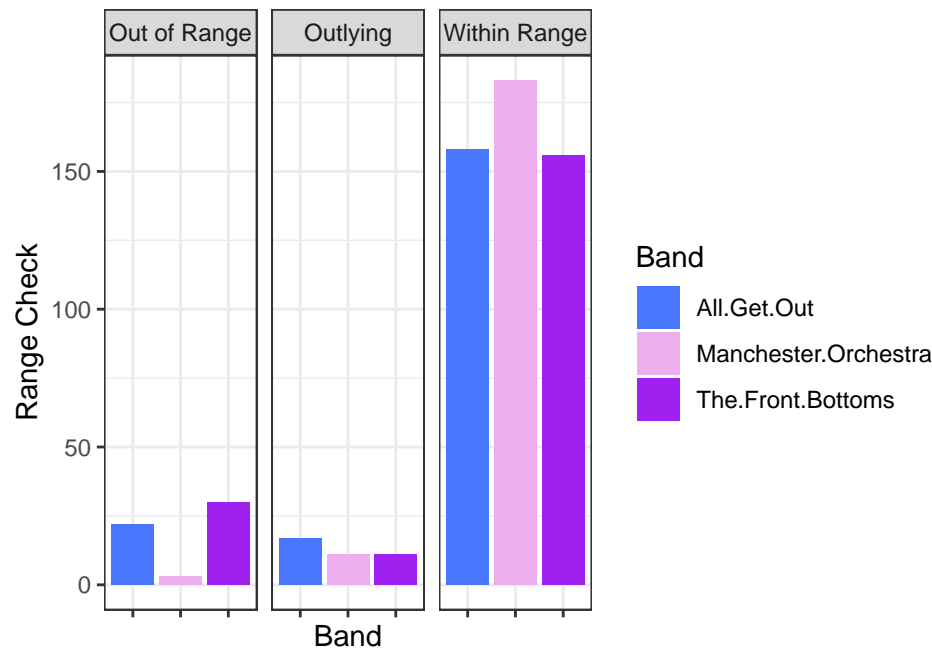Table 1: Comparison of Allentown's audio features with the range of band's features

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