Lab 02 – MATH 240 – Computational Statistics

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

02/04/2025

Abstract

In this lab, we analyzed musical characteristics of JSON files and started considering automated processing of audio. We want to analyze bands' contributions to the song "'Allen town', so we created a batch file containing 181 commands to avoid repetitive calling of an executable file. We traversed structured directory of music files to extract relevant data for creating commands. We also practiced traversing a JSON file to extract key features of a song in preparation to analyze bands' contributions to "'Allen town'.

Keywords: Installing and using libraries; data extraction; looping structure; vectors and lists operations.

1 Introduction

Allen Town 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. To achieve correct analysis, we purchased all releases before Allen Town, which consists of 180 tracks, or 181 including Allen Town itself.

We use Essentia to analyze, synthesize, and describe data about 181 songs to determine the musician who makes the most contribution to Allen Town. We analyze the style and characteristics of tracks belonging to each band to determine stylistic contribution.

We want to automate the process of generating the command-line prompts to speed up the process of executing the data. Therefore, we create a batch file that contains multiple commands for analysis of all given tracks.

1.1 Lab02 Specifics

During Lab 02, we practices creating a batch file with command lines of un-copyrighted songs to practice task automation. Moreover, we learned how to analyse one JSON file for track characteristics, such as tempo in beats and average loudness. This knowledge will aid us in processing larger set of tracks by The Front Bottoms and Manchester Orchestra.

1.2 Paper Structure

This paper is structured as follows: We start with Introduction, focusing on background about Allen Town and the question of main band contributor. In Methods, we summarize data collection process. The Results section presents the outcome of this lab. In Discussion section, we interpret the findings from this lab.

2 Methods

For the purpose of data collection, we were given uncopyrighted MUSIC directory that contained artists and albums. We used `stringr` package, including its functions `str_count()` and `str_sub()` to subset `.WAV` files in the sub directories and obtain wanted naming format for command line (Wickham, 2023). We also extracted track's name, artist, and album to create a batch file that creates command line prompts for each track. Creation of the batch file was automated with the use of a for loop that processed files in each album sub directory and storing desired output for each file name in a vector.

2.1 Processing JSON Output

Using `jsonlite` package for `R`, we practiced analyzing stylistic features of a song (Ooms, 2014). Our song example was Au Revoir (Adios) by The Front Bottoms. We used `fromJSON()` to load JSON file of the song 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.

3 Results

As a result of this lab, we learned how to create a batch file for analysis of multiple tracks at the same time. Using an uncopyrighted directory provided us a chance to navigate sub directories and extract needed names for the creation of the command line. As a result of processing JSON file, we were able to analyze key features of Au Revoir (Adios) by The Front Bottoms.

4 Discussion

It is easier to analyze a bulk of data if we have a batch file that runs command lines together for all data. A for loop and libraries are helpful for simplifying creation of the batch file and data collection. Additionally, the processing of JSON file of the song Au Revoir (Adios) showed that its average loudness is 0.55, its spectral energy is 0.21, dancebility is 0.97, bpm is 140, and its length is 108 seconds. All of this data could be used when analyzing contributions of The Front Bottoms to Allen Town.

References

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.