Lab 02 – MATH 240 – Computational Statistics

Cristian Palmer
Student
Mathematics
cpalmer@colgate.edu

Abstract

In this lab, we first created a batch file that we will use in our next lab to process data about specific songs. To achieve this, we first downloaded the directory of songs and subsequently altered the output, finally saving what we desired into our final batch-file. In the second section of the lab, we used the `jsonlite` package to extract important data about a specific song. This process will become very important in the next installment of this lab.

Keywords: Batch File : JSON : Loops

1 Introduction

This lab originated from the question of which band, *The Front Bottoms*, or *Manchester Orchestra* contributed more to their collaboratory song Allen Town. One way to approach this question is to use a program to analyze all of each band's songs. The first step to this process is to create a batch file with all of these songs, which prompted the idea for the first part of this lab. We were tasked with downloading a fake directory of songs, and then subsequently formatting the list and processing it to a batch file. The second task of this lab sees us utilizing `jsonlite` to extract data about a specific *Front Bottoms* song. This part of the lab is directly applicable to the subsequent lab which we will be completing next week.

2 Methods

For this lab we were provided with seven fake "albums", each of with containing two fake "songs". Although the songs and albums are not necessarily real, the process we followed is exactly the same one which could be used on real music. The "songs" were provided to us in a multiply embedded folder. Specifically, the outer folder we were provided was called Music, inside that folder were two more titled OfficeStuff and PeopleStuff. Subsequently, these folders held four and three albums respectively, and each of these albums held two songs, bringing us to our total of 7 albums and 14 songs. Our first step in turning these songs into a batch file was to assign an object to a vector containing the file name of each song. To achieve this step we focused on using the `\str_count()` function. Then, we focused on using the `\paste()` function, the `\str_sub()` function, and the `\str_split()` to turn each "song" title into the output we desired for our batch file. Finally, we saved our desired output to a vector named `code.to.process` and utilized the `\writelines()` function to create our final batch file named batfile.txt. To complete this process for each of our 14 songs we utilized two loops in our code. For the second task of this lab, we used the `jsonlite` package to extract data about the *Front Bottoms* song Au Revoir (Adios). We first used the `\str_split()` function to break the file name down into its components and then used the `\fromJSON()` function to load the JSON file for this song into `R`. Finally, we utilized the Essentia documentation to aid us in extracting the average loudness, the mean of spectral energy, danceability, bpm (tempo in beats per minute), key key (musical key), key scale (musical mode), and the length (duration in seconds) of the song.

3 Results

For the first task of this lab, the resulting output was a batch file consisting of the 14 "songs", in our desired format, which we hope to later analyze. A resulting output in the file looks like this: `streaming_extractor_music.exe` "01-OfficeStuff-Zip.wav" "OfficeStuff-Backpack-Zip.json". Using our loops, we were able to create one of these for every "song". In the second section of the lab, we successfully were able to extract the JSON data for the *Front Bottoms* Song Au Revoir (Adios). For example, running our code gives us the data that the BPM of the song is 140.8809 beats per minute. The processes conducted in this lab will prove extremely useful for the subsequent labs to follow.