

1. In Lab 3, you wrangled data from Essentia, Essentia models and LIWC. Rework your solution to Lab 3 using **tidyverse** (Wickham et al., 2019) instead of base R. Specifically, rewrite your code for steps 1-4 of task 2 using **tidyverse** (Wickham et al., 2019). Make sure to address any issues I noted in your code file, and ensure that your code runs in the directory as it is set up.

```
library(tidyverse)
library(jsonlite)
##### Part 1 #####
current_filename <- "The Front Bottoms-Talon Of The Hawk-Au Revoir (Adios).json"
json_data <- fromJSON(current_filename)
file_parts <- str_split_fixed(current_filename, "-", 3) |>
str_replace("\\\\.json$", "")
artist <- file_parts[1]
album <- file_parts[2]
song <- file_parts[3]
overall_loudness = pluck(json_data, "lowlevel", "loudness_ebu128", "integrated")
spectral_energy = pluck(json_data, "lowlevel", "spectral_energy")
dissonance = pluck(json_data, "lowlevel", "dissonance")
pitch_salience = pluck(json_data, "lowlevel", "pitch_salience")
bpm = pluck(json_data, "rhythm", "bpm")
beats_loudness = pluck(json_data, "rhythm", "beats_loudness")
danceability = pluck(json_data, "rhythm", "danceability")
tuning_frequency = pluck(json_data, "tonal", "tuning_frequency")

#####Part 2#####

library(tidyverse)
library(jsonlite)

# Get JSON files
json_files <- list.files(path = "EssentiaOutput/", pattern = "\\\\.json$", full.names = TRUE)

# Function to extract data from a single JSON file
extract_essentia_data <- function(file_path) {
  file_name <- basename(file_path)
  file_parts <- str_split(file_name, "-", simplify = TRUE)

  artist <- file_parts[1]
  album <- file_parts[2]
  song <- str_replace(file_parts[3], ".json$", "")

  json_data <- fromJSON(file_path)

  tibble(
    artist = artist,
    album = album,
    song = song,
    overall_loudness = pluck(json_data, "lowlevel", "loudness_ebu128", "integrated", .default = NA),
    spectral_energy = pluck(json_data, "lowlevel", "spectral_energy", .default = NA),
    dissonance = pluck(json_data, "lowlevel", "dissonance", .default = NA),
    pitch_salience = pluck(json_data, "lowlevel", "pitch_salience", .default = NA),
    bpm = pluck(json_data, "rhythm", "bpm", .default = NA),
    beats_loudness = pluck(json_data, "rhythm", "beats_loudness", .default = NA),
    danceability = pluck(json_data, "rhythm", "danceability", .default = NA),
    tuning_frequency = pluck(json_data, "tonal", "tuning_frequency", .default = NA)
  )
}

# Process all JSON files and combine into a single tibble
essentia_df <- map_dfr(json_files, extract_essentia_data)

#####
# Step Three
#####
essentia_model_output <- read_csv("EssentiaOutput/EssentiaModelOutput.csv")

# Compute averaged columns
essentia_model_output <- essentia_model_output |>
mutate(
  valence = (deam_valence + emo_valence + muse_valence) / 3,
  arousal = (emo_arousal + muse_arousal + deam_arousal) / 3,
  aggressive = (nn_aggressive + eff_aggressive) / 2,
  happy = (nn_happy + eff_happy) / 2,
  party = (nn_party + eff_party) / 2,
  relaxed = (nn_relax + eff_relax) / 2,
  sad = (nn_sad + eff_sad) / 2,
  acoustic = (nn_acoustic + eff_acoustic) / 2,
```

```

    electric = (nn_electronic + eff_electronic) / 2,
    instrumental = (nn_instrumental + eff_instrumental) / 2
  )|>
  rename(timbreBright = eff_timbre_bright) |>
  select(artist, album, track, valence, arousal, aggressive, happy,
         party, relaxed, sad, acoustic, electric, instrumental, timbreBright)

# View the cleaned data
print(essentia_model_output)

## # A tibble: 181 x 14
##   artist      album track valence arousal aggressive   happy party relaxed   sad
##   <chr>      <chr> <chr>   <dbl>   <dbl>         <dbl>   <dbl> <dbl>   <dbl> <dbl>
## 1 All Get 0~ All ~ 06 W~   4.34   4.17       0.554 0.0106 0.979 0.275 0.710
## 2 All Get 0~ All ~ 01 C~   4.60   4.28       0.657 0.00714 0.973 0.350 0.780
## 3 All Get 0~ All ~ 07 T~   3.90   3.80       0.122 0.00713 0.993 0.0286 0.520
## 4 All Get 0~ All ~ 03 W~   4.61   4.36       0.522 0.0114 0.941 0.413 0.689
## 5 All Get 0~ All ~ 02 C~   4.41   4.22       0.528 0.0107 0.985 0.296 0.662
## 6 All Get 0~ All ~ 05 T~   4.52   4.33       0.679 0.00889 0.962 0.435 0.771
## 7 All Get 0~ All ~ 04 L~   4.59   4.33       0.534 0.0126 0.966 0.370 0.764
## 8 All Get 0~ Nobo~ Room~   4.42   4.35       0.428 0.00811 0.976 0.264 0.649
## 9 All Get 0~ Nobo~ Home~   4.70   4.38       0.597 0.0103 0.970 0.436 0.791
## 10 All Get 0~ Nobo~ Wait~  4.14   4.01       0.206 0.0112 0.993 0.0689 0.404
## # i 171 more rows
## # i 4 more variables: acoustic <dbl>, electric <dbl>, instrumental <dbl>,
## #   timbreBright <dbl>

#####
#Step Four
json_files <- list.files(path = "EssentiaOutput/", pattern = "\\\\.json$", full.names = TRUE)

extract_essentia_data <- function(file_path_2) { #Extracts file name without it's ending
  file_name_2 <- basename(file_path_2)
  file_parts_2 <- str_split(file_name_2, "-", simplify = TRUE)

  json_data <- fromJSON(file_path_2)

  tibble(
    artist = file_parts_2[1],
    album = file_parts_2[2],
    song = str_replace(file_parts_2[3], ".json$", ""),
    overall_loudness = pluck(json_data, "lowlevel", "loudness_ebu128", "integrated", .default = NA),
    spectral_energy = pluck(json_data, "lowlevel", "spectral_energy", .default = NA),
    dissonance = pluck(json_data, "lowlevel", "dissonance", .default = NA),
    pitch_salience = pluck(json_data, "lowlevel", "pitch_salience", .default = NA),
    bpm = pluck(json_data, "rhythm", "bpm", .default = NA),
    beats_loudness = pluck(json_data, "rhythm", "beats_loudness", .default = NA),
    danceability = pluck(json_data, "rhythm", "danceability", .default = NA),
    tuning_frequency = pluck(json_data, "tonal", "tuning_frequency", .default = NA)
  )
}

# Process all JSON files and combine into a single tibble
essentia_df_2 <- map_dfr(json_files, extract_essentia_data)

# View the cleaned data
print(essentia_df)

## # A tibble: 181 x 11
##   artist album song overall_loudness spectral_energy dissonance pitch_salience
##   <chr>   <chr> <chr>         <dbl> <named list>   <named li> <named list>
## 1 All G~ All ~ 01 C~   -7.04 <dbl [1]>   <dbl [1]> <dbl [1]>
## 2 All G~ All ~ 02 C~   -6.47 <dbl [1]>   <dbl [1]> <dbl [1]>
## 3 All G~ All ~ 03 W~   -6.29 <dbl [1]>   <dbl [1]> <dbl [1]>
## 4 All G~ All ~ 04 L~   -6.16 <dbl [1]>   <dbl [1]> <dbl [1]>
## 5 All G~ All ~ 05 T~   -6.73 <dbl [1]>   <dbl [1]> <dbl [1]>
## 6 All G~ All ~ 06 W~   -7.39 <dbl [1]>   <dbl [1]> <dbl [1]>
## 7 All G~ All ~ 07 T~   -9.20 <dbl [1]>   <dbl [1]> <dbl [1]>
## 8 All G~ Move~ All ~   -6.95 <dbl [1]>   <dbl [1]> <dbl [1]>
## 9 All G~ Move~ Bala~   -7.06 <dbl [1]>   <dbl [1]> <dbl [1]>
## 10 All G~ Move~ Move~  -6.52 <dbl [1]>   <dbl [1]> <dbl [1]>
## # i 171 more rows
## # i 4 more variables: bpm <dbl>, beats_loudness <named list>,
## #   danceability <dbl>, tuning_frequency <dbl>

```

References

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.