**Adam Curry - Term Project Summary**

This was a fun project. I got to utilize what I’ve learned in the data wrangling course in conjunction with my Spotify playlist to develop a dataset of the various artists that are in my playlists. Not only that, I was able to develop ideas for future projects that could also offer a fun look into Spotify user’s listening habits. But for this project, I stuck with my own listening habits. The following is a summary of the steps I took to get my final dataset and some of the trials and tribulations I went through.

To begin, I went to Spotify’s developer website to obtain and access token so I could obtain data through their API. This was an easy step and I obtained a key instantly. Next, I read their user agreement to see the limits of my calls to the API. I noticed I could get access revoked, so I wanted to remain diligent in my approach. In order to do that, I established metrics I wanted to wrangle based on the documentation provided by Spotify. I was able to establish my own metadata based on Spotify’s documentation to the variables I would produce in my dataset, but not before traveling down several rabbit holes to obtain my variables. There is an abundance of data available through Spotify’s API, which made my mind fire off several possible scenarios for a final dataset. Should I produce all albums by a few artists, the audio features of an artists, the audio features of an album, the playlists available for multiple users, etc.? I had several ideas, but ultimately landed on my own listening habits. I would produce audio analysis for the tracks within my own playlist.

Once I established my end goal, I begin extracting data. This required several “dry runs” on single values, as I didn’t want to get stuck in and endless loop and lose access to my API token. This was also a challenge as I had to iterate over several items within a single Spotify object. For example, I have several Spotify playlists and within each playlist is a track. Therefore, I had to iterate over the playlists and then within that iteration, applied an additional iteration over the tracks. This will make future pulls significantly easier and optimal for frequent use, as the key values are already parsed and ready for extraction.

Finally, I came away with two tables, one with my playlist data and basic data on the tracks and the other containing the audio analysis for each track. I was able to easily load these to a Pandas data frame and merge on the uri track key from both datasets for a total of 20 variables. The exploratory analysis was conducted for my own knowledge. I simply wanted to see the distribution of some of the variables within my dataset.