Spotify Data Analysis

# **Myrna Acuna Semester** <https://github.com/oblea/MusicTasteAnalysis>

# Proposal

Music seems to be a very popular and integral part of everyone’s lives. I would like to use Spotify’s API to do some data analysis on my playlist and determine what kind of music taste I have.

# Which Domain?

What domain is this data going to come from? Please list 10 references (with a brief annotation) to use to make sense of what you’re doing with these data.

|  |  |
| --- | --- |
| **Source** | **Description** |
| <https://developer.spotify.com/documentation/web-api/reference/tracks/> | Tools and reporting for analytics from Spotify |
| Seaborn, matplotlib, pandas, numpy, sklearn | Seaborn and matplotlib for data visualization  Pandas and numpy for analysis  Sklearn for machine learning |
| spotipy | Python module imported for spotify music analysis |
| <https://www.northeastern.edu/graduate/blog/spotify-big-data/> | Understand Spotify’s current data analysis on how they predict the music that you’ll most likely enjoy |
| <https://developer.spotify.com/documentation/web-api/reference/tracks/get-several-audio-features/> | Understanding audio features categorizations that spotify assigns to each song such as acousticness, danceability, and energy |
| <https://medium.com/@ericboam/i-decoded-the-spotify-recommendation-algorithm-heres-what-i-found-4b0f3654035b> | Understand the inner workings of Spotify’s recommendation algorithm |
| <https://blog.salesflare.com/spotify-iconic-product> | Determine what set spotify apart from its competitors |
| <https://labs.spotify.com/2015/03/09/understanding-spotify-web-api/> | Details on understanding spotify’s API |
| <https://www.wired.com/story/spotify-tips-and-tricks/> | How am I able to customize my spotify listening patterns |
| <https://www.boldanalytix.com/getting-data-from-spotify/> | Testing spotify’s API |
| <https://www.kaggle.com/theoverman/the-spotify-hit-predictor-dataset> | Spotify hit predictor dataset depending on the criteria of the artist |

# Which Data?

What is the dataset you’ll be examining? Please provide a codebook if there is one or a link to the dataset as well as a detailed description.

There is currently no codebook as I will be using my Spotify playlist to understand my listening patterns and tastes, Also, I would like to see if I can determine how unique my music taste is to the general population – I will be using Spotify’s top 50 songs to compare this to.

# Research Questions? Benefits? Why analyze these data?

How are you proposing to analyze this dataset? This is about your approach. Here, you’ll be proposing your research questions as well as justifications for why you’d offer these data in this way.

How different is my music taste from the general public?

What are my music tastes?

Are my spotify findings varied, instrumental, or boring?

# What Method?

What methods will you be using? What will those methods provide in terms of analysis? How is this useful?

I will use Python, Spotify’s API, Seaborn and matplotlib for data visualization, Pandas and numpy for analysis, Sklearn for machine learning

# Potential Issues?

What challenges do you anticipate having? What could cause this project to go off schedule?

I am concerned with what I will compare my music habits to, maybe I can compare my music to my friend’s music.

# Concluding Remarks

The purpose of this analysis is to determine whether my music taste is really that unique or does it cater mostly to what is popular according to Spotify’s top tracks.