

## Song Recommender Executive Summary

### Goal:

- Create a machine learning model to suggest songs to continue a user's playlist.
- Create a web app to allow users to interact with our model.

### Data:

- The primary dataset the "[Million Playlist Dataset](#)" from Spotify, which contained 1,000,000 user generated playlists.
- We scraped audio features and genre data using the Spotify API and lyrics using the Genius API.

### Preprocessing:

- Lyrics and genre data were vectorized via TF-IDF (Term Frequency - Inverse Document Frequency)
- Performed sentiment analysis on lyrics using NLTK's Naive Bayes Analyzer.

### Models:

- We took an unsupervised learning approach.
- Our first model uses K-Means clustering. We trained one model using all audio features. In addition, we trained several models using only one audio feature.
- We refined this model with fuzzy clustering, allowing data points to fit into several clusters.
- Given a playlist, our model ranks the relevant clusters and recommends songs in those clusters.