**Project 1: Clustering Music Into Predictive Subgenres Using Audio Features**

Victor Ramirez

**Problem**

Some music recommendation engines rely on user-specified genres, but oftentimes music can’t be pigeonholed into a broad spectrum within a genre. Moreover, relying on users requires enough listeners to classify a song and each listener’s classification is subjective. This project aims to classify songs into descriptive subgenres based purely on audio features.

**Deeper Insights Into Genres**

Genre classification is the backbone of music recommendation engines (MRE). Netflix can recommend users, say “visually-striking thrillers from the 80’s”, but such a recommendation engine isn’t present in music platforms like Pandora or Spotify. Having a robust music recommendation engine will allow users to discover music that best suits their taste or current mood.

**Data**

The data used is from the [Free Music Archive](https://github.com/mdeff/fma). This data contains audio files and metadata for each track.

**Approach**

The canonical approach to this problem would be to classify each song in a supervised manner and to check against the given genres. A further analysis would use k-means clustering to find common features within genres and separate them into descriptive subgenres.

**Deliverables**

Tentatively, a code and a paper.