# Project 1: Music Recommender

**Executive Summary**

**Nik Baldis**

## Methodology

This project aimed to make a recommender system that recommends songs to the user based on their play count. The project only successfully used an SVD Model to make recommendations. Still, an attempt at a cosine-similarity model was also created to try to solve the cold-start problem that all recommender systems face in the beginning stages of a user’s history. An exploratory analysis was done to understand the data and get fundamental insights about the users and the sampled song database. A baseline model was created to test the effectiveness of the SVD Model tuning, and some recommendations were made based on specific constraints.

## Key Findings

The Global baseline model produced an RMSE of 2.5584 and an MAE of 1.9063 with the hyperparameters of 5 epochs, and a regularization factor of 5 for users and items. The Tuned SVD Model created an RMSE of 2.5547 and an MAE of 1.9039 with the hyperparameters of 75 factors, 15 epochs, a learning rate of 0.005, and a regularization of 0.05. With the best-fit hyperparameters, the SVD model was slightly improved over baseline, based on error. However, the recommendations that I got for myself and my peer were not highly accurate to what we know are our music tastes. There were some songs in our top ten recommendations that we enjoyed, but many of them were random. To improve recommendations for a user and expand their relevant song recommendation pool, I would use more content-based filtering to introduce content that may not follow the latent factors that the model learns on its own. However, the model will provide recommendations similar to the user's usual listening.  Although content-based filtering may not have the randomness a user is accustomed to, it would be much more relevant to their current listening habits. The easiest and most common way to answer the cold-start problem for a user is to recommend the most popular songs. An attempt was made to create an item-based model with the data we have, but without actual categorical variables to connect with, the content-based model was not accurate. I suggest introducing a hybrid model between collaborative filtering and content-based filtering to improve recommendations- such as the SVD I created and a KNN model.  The SVD is very good at finding some latent factors that are not clearly defined and connecting them to new recommendations.  On the other hand, the KNN or content-based recommendation models are much better at connecting similar songs based on the specified characteristics.