

# Can You Predict a Song's Commercial Success?

## ORIE 4741 Final Project Proposal

Darren Ting, Mario Medina, Joshua Kwasi Agyei-Gyamfi

October 4th, 2020

- <https://www.kaggle.com/yamaerenay/spotify-dataset-19212020-160k-tracks>
- <https://www.kaggle.com/daniel2255/data-on-songs-from-billboard-19992019>

Music has been a staple in society that has become immensely diverse through the evolution of different cultures. Ultimately, it is an expression of one's self and in today's society, it is a way to gain financial capital to improve one's life situation. Being able to predict the success of a song can further estimate its potential to bring revenue for the artist/group as well as its ability to go mainstream and make an influence in today's society. In addition, it shows the trends of current musical interests in mainstream culture which can allow musicians to tailor their music to. For this project, we explore this further by asking, "Can you predict a song's commercial success based on characteristics of the song?".

To answer this question using supervised machine learning, we are looking to build a predictive model using classification and regression. The model will first predict whether a song will be successful or not, then, if successful, the duration of the song in the Billboard Top 100 Chart. We will use audio features of songs requested from Spotify's API, which include 13 numerical-valued features and two boolean-valued features (See first reference for details), as our input variables. This data will be used in conjunction with Billboard Hot 100 data that has been shared on Kaggle which will be used as a benchmark for mainstream musical success. This data set will serve as the response variables for both the classification and regression parts of our model. We believe that these data sets will allow us to successfully answer our question and give us more insight on what makes a song popular in society. The data from Spotify in particular covers a wide range of audio features that we believe are good indicators of whether a song will perform well in the mainstream or not.