**Classify Song Genres from Audio Data**

Rock or rap? Apply machine learning methods in Python to classify songs into genres.

#### Project Description

Using a dataset comprised of songs of two music genres (Hip-Hop and Rock), you will train a classifier to distinguish between the two genres based only on track information derived from [Echonest](http://the.echonest.com) (now part of Spotify). You will first make use of pandas and seaborn packages in Python for subsetting the data, aggregating information, and creating plots when exploring the data for obvious trends or factors you should be aware of when doing machine learning. Next, you will use the scikit-learn package to predict whether you can correctly classify a song's genre based on features such as danceability, energy, acousticness, tempo, etc. You will go over implementations of common algorithms such as PCA, logistic regression, decision trees, and so forth.

This project lets you apply what you learned in [Supervised Learning with scikit-learn](https://www.datacamp.com/courses/supervised-learning-with-scikit-learn), plus data preprocessing, dimensionality reduction, and machine learning using the scikit-learn package.

#### Project Tasks

* 1 Preparing our dataset
* 2 Pairwise relationships between continuous variables
* 3 Normalizing the feature data
* 4 Principal Component Analysis on our scaled data
* 5 Further visualization of PCA
* 6 Train a decision tree to classify genre
* 7 Compare our decision tree to a logistic regression
* 8 Balance our data for greater performance
* 9 Does balancing our dataset improve model bias?
* 10 Using cross-validation to evaluate our models