Moodify

Exploratory Data Science Project

Maya Nesen

July - August 2023

Project Goals

This project uses the "Emotions Labeled Spotify Songs" dataset that I found on Kaggle to explore different patterns of how different songs impact people's mood. I applied various data science techniques, such as K-means clustering, correlation, Linear Regression, statistical tests, PCA, and other machine learning algorithms to identify any trends between musical and emotional features. All the code is done using Python, packages including numpy, pandas, scipy, matplotlib, and more.
There are two datasets being used: one has 278k songs/rows and the other has 1200 songs/rows. Since I am not using a workstation or a computer which can handle large amounts of data, I primarily used the one with 1200 songs.
The goal of this project is to both explore the dataset and also practice key data science techniques to prepare for a career in Data Science. Anything that is marked as "review" in the cell title of the code is not being used towards the project goals and more for personal practice. (See code in GitHub "spotify_mood_project.py".)

The Dataset

The dataset is comprised of 17 columns, 9 of which are the musical features of the songs (listed below) and one being the emotions labels: happy, sad, energetic, and calm.

Danceability

Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.

Energy

Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.

Loudness

The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing the relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.

Speechiness

Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audiobook, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.

Acousticness

A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

Instrumentalness

Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.

Liveness

Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides a strong likelihood that the track is live.

Valence

A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

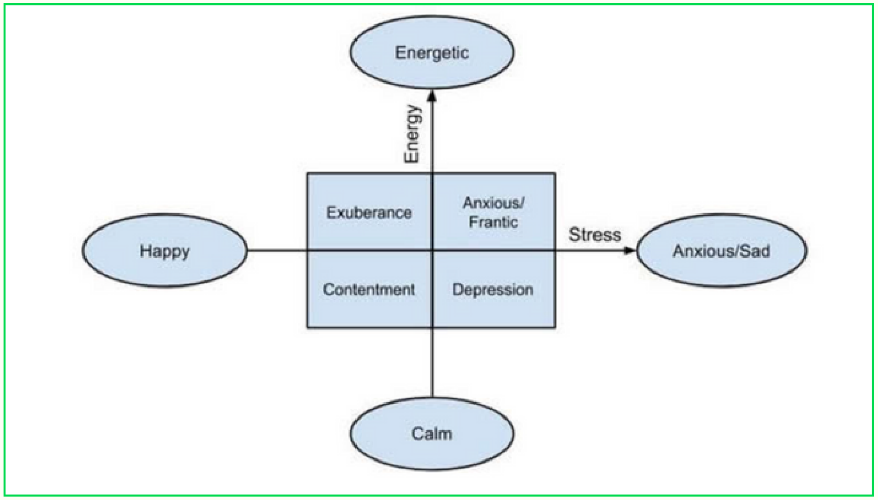
Tempo

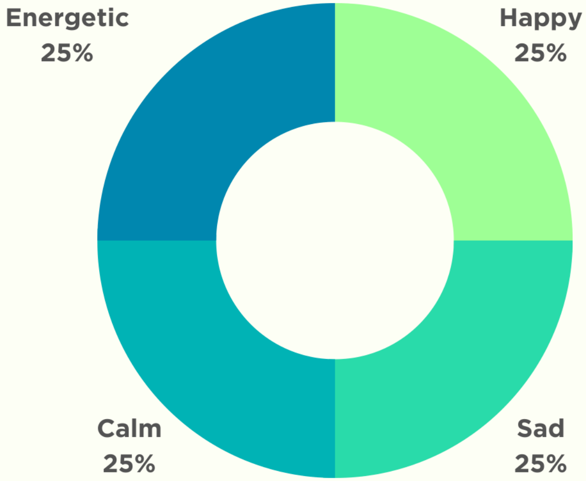
The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, the tempo is the speed or pace of a given piece and derives directly from the average beat duration.

Music qualities from Moodify Dataset Creators' GitHub: https://github.com/orzanai/Moodify/blob/main/Moodify.pdf

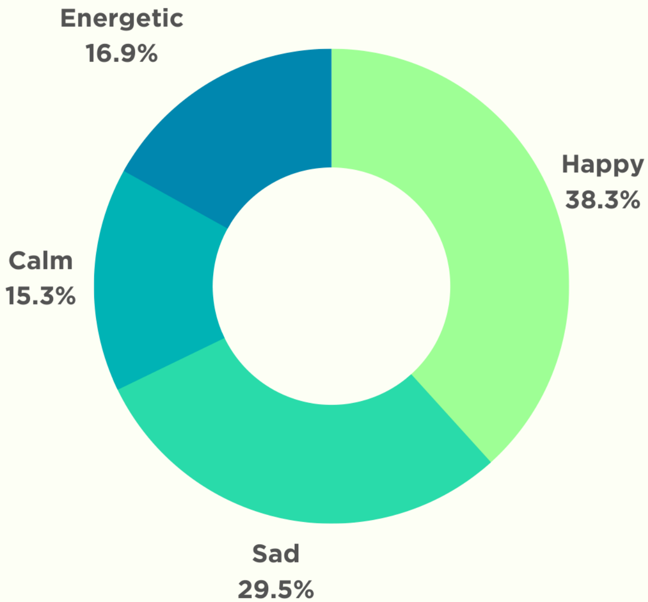
Mood Labels

The mood/emotions labeled for each song in the Moodify dataset are based off of the "Music Mood Classification" research paper by Michael Nuzzolo at Tufts University. Emotional reactions to music of course vary from person to person. However, for simplication, Nuzzolo and the dataset creators focus on three moods: "happy", "sad", "calm", and "energetic".





For 1200 songs dataset

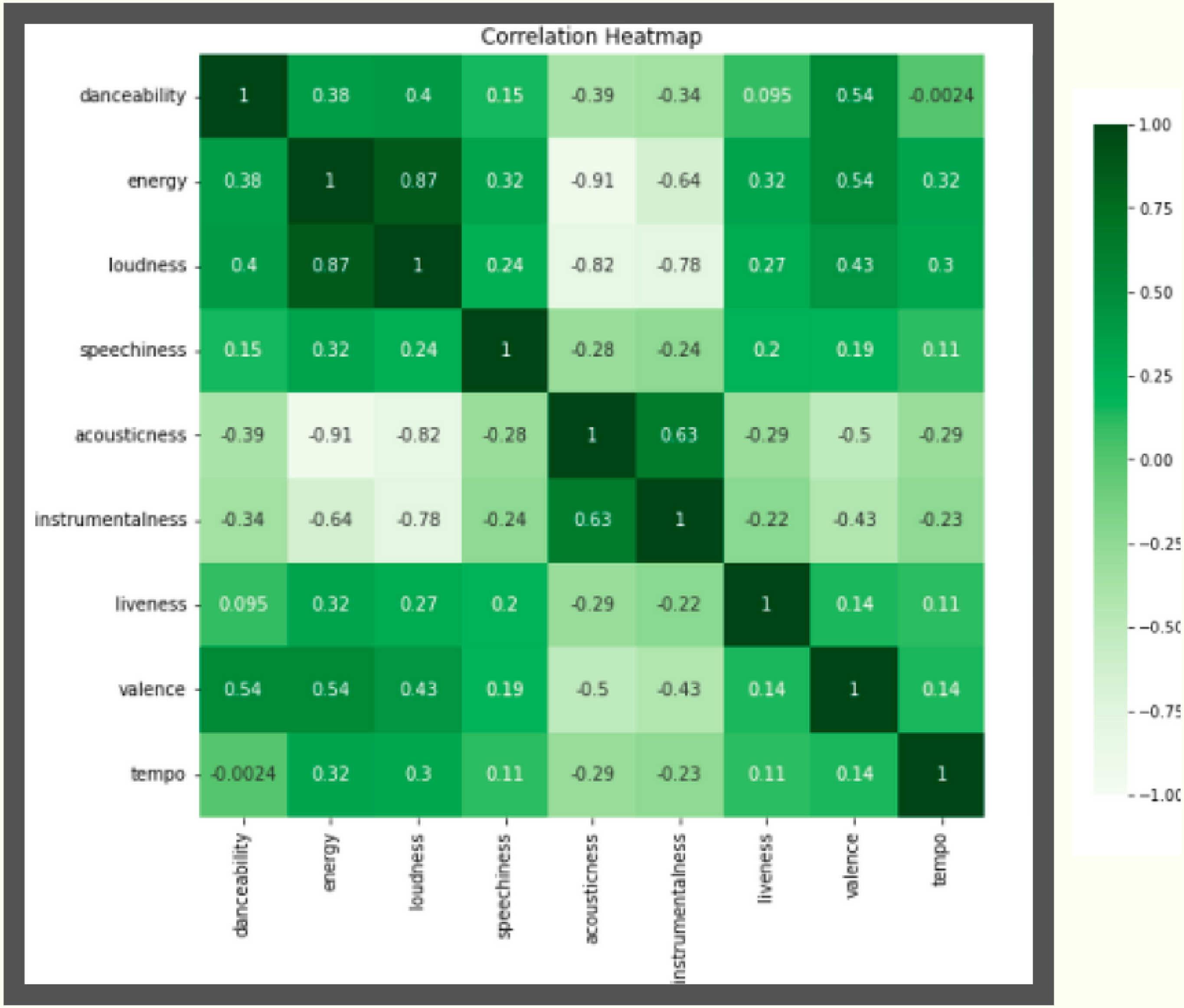


For 278k songs dataset

Using a pie graph method in Python, I created the following pie charts for each dataset, showing the proportion of songs by mood to determine how well-balanced the dataset is. The proportion of songs by mood in the two dataset are fairly similar, but are clearly equally balanced in the smaller dataset. The larger 278k dataset has a higher proportion of songs labeled as "sad" or "happy". This could perhaps be because most songs made fall into these two categories, more so than "energetic" or "calm".

Diagram from Tufts Mood Research Study: https://sites.tufts.edu/eeseniordesignhandbook/2015/music-mood-classification/

Correlation



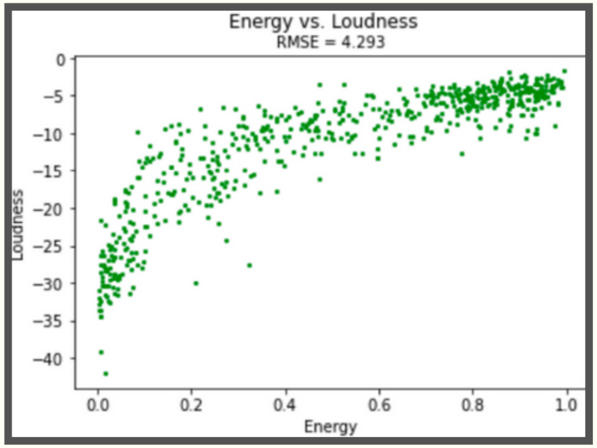
Notable correlations (darker or very light squares):

1. Energy and loudness = 0.87 (extremely positive)
2. Danceability and valence = 0.54 (moderate)
3. Instrumentalness and acousticness = 0.63 (moderate)
4. Energy and valence = 0.54 (moderate)
5. Loudness and instrumentalness = -0.78 (fairly negative)
6. Energy and acousticness = -0.91 (extremely negative)

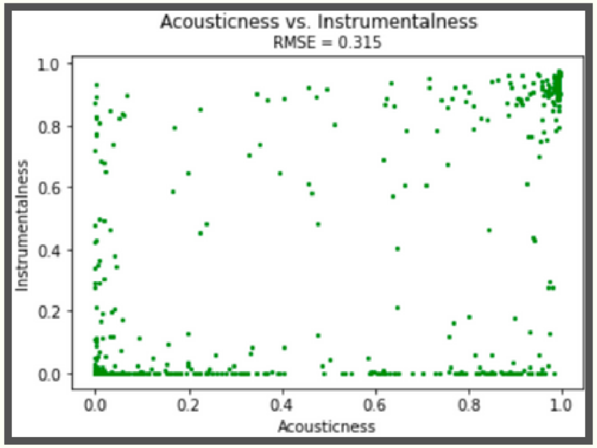
These correlations are the motivation for plotting Linear Regression between these variables.

This was plotted using the seaborn package. Code taken from Dhruv Choudhary on Kaggle, and edited to function with the 1200 song dataset and to change the color scheme.

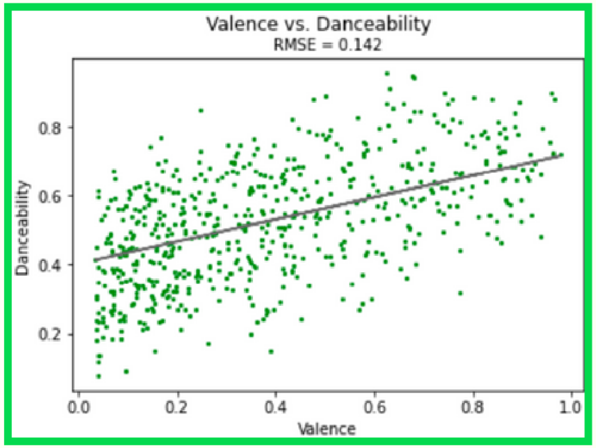
Linear Regression Analysis



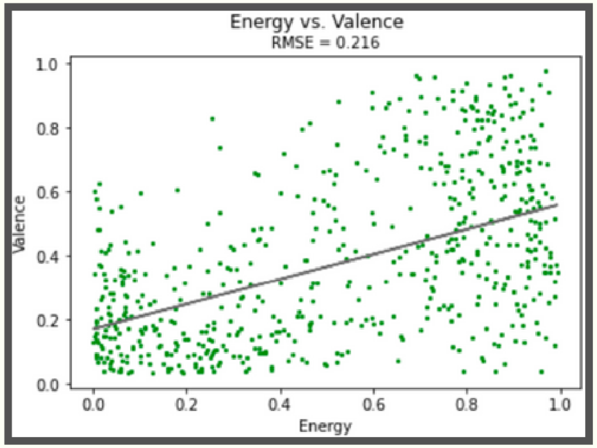
Pearson's r = 0.87 and Spearman's ρ = -0.897. The correlations are similar, but our plot clearly gives an exponential or logarithmic relationship. The RMSE is very high, so the relationship is not accurate. As energy increases, loudness slowly increases.



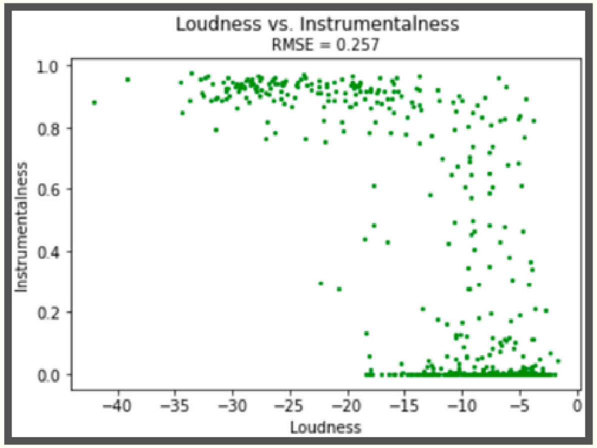
Pearson's r = 0.63 and Spearman's ρ = 0.527. Our plot clearly does not give a linear relationship. However, most points are found at the extremities at 0 or 1. Perhaps logistic regression would be a better fit.



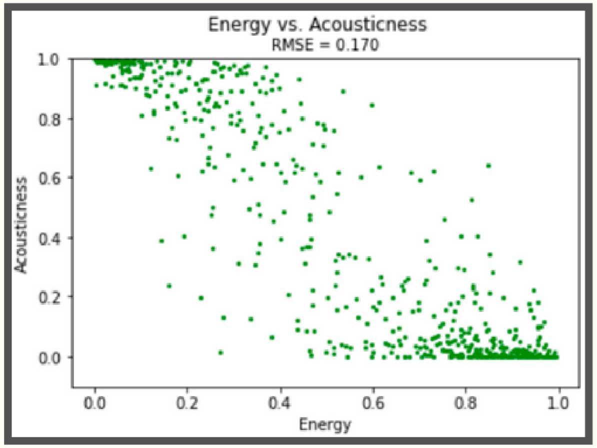
Pearson's r = 0.54 and Spearman's ρ = 0.527. The correlations are very similar and our plot gives a clear linear relationship. The RMSE is low, signaling a fairly strong prediction. As valence increases, danceability increases.



Pearson's r = 0.54 and Spearman's ρ = 0.524. While our correlation values are similar, our plot does not show a linear relationship, with most points found at energy values of 0 or 1.

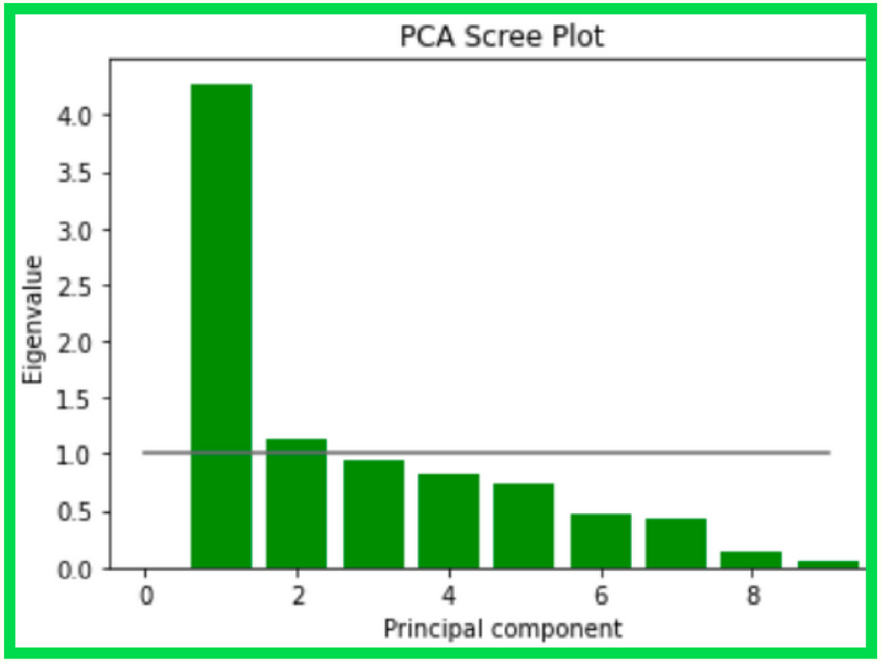


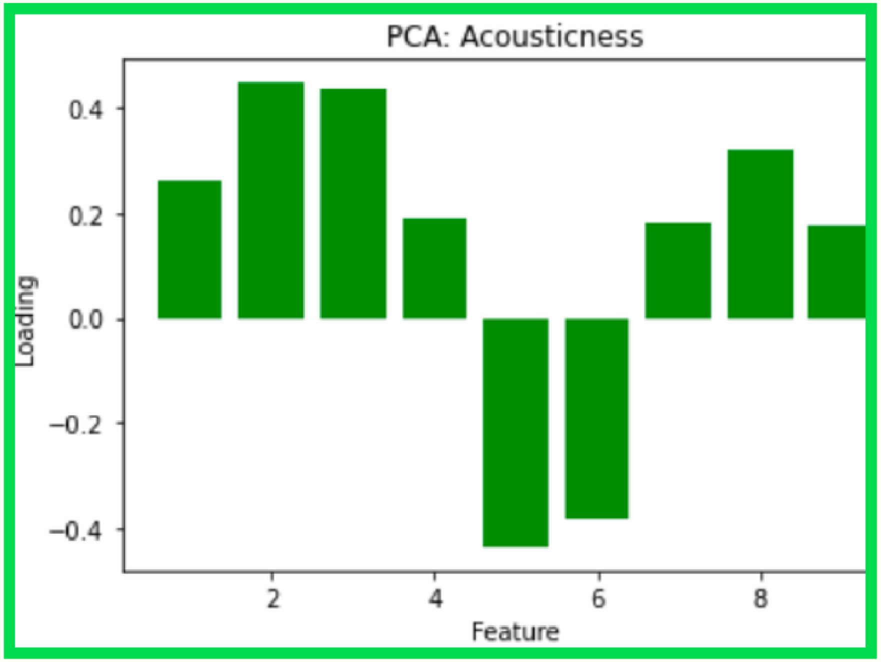
Pearson's r = -0.78 and Spearman's ρ = 0.527. Not only are the correlation values very different, our plot is clearly not linear. Most points are found at instrumentalness of 0 or 1. Perhaps logistic regression would be a better fit.



Pearson's r = -0.91 and Spearman's ρ = -0.88. While the correlation values are similar, our plot is clearly not linear. Most points are found at acousticness of 0 or 1. Perhaps logistic regression would be a better fit.

Principal Component Analysis





Principal Component Analysis (PCA) is an unsupervised machine learning algorithm which reduces the dimensionality of a large dataset to a smaller dataset with variables/components that provide the most or strongest information. I performed a PCA on our dataset, focusing on the columns with our 9 musical qualities. In the PCA Scree Plot, I used the Kaiser Criterion, signifying that components with an eigenvalue above 1.0 are strongest and explain the most variance in our data. The first component with an eigenvalue of 4.0 corresponds to the Acousticness feature. From the bar plot on the right, we see that Acousticness describes best all the features except two: liveness and loudness. This makes some sense since acoustic music is generally not loud. Think of an acoustic guitar version of a pop song; it would have a much softer sound. In addition, music that is not heavily acoustic may not necessarily be performed live.