Spotify Song Genre Comparison

Data Analysis and Clustering

Main Objective of Analysis

Spotify is the world's most popular streaming platform for music and as of 2024, it had 574 million monthly active listeners. The content includes a multitude of genres and each artist and song are associated with various data attributes which are used to create recommendations and playlists that a user could like to continue listening. An important set of data is the acoustic attributes for songs, which are quantitative aspects of a song that could lead to a user enjoying a recommendation of a similar one they might like. In this analysis, data included songs of five different genres: hip hop, rock, jazz, pop, classical. These were then analyzed and compared for their various acoustic attributes. Finally, various unsupervised learning models were created to cluster the songs together to observe how well the genre tags of songs aligned with their acoustic attributes.

Description of Dataset

Data was retrieved from Spotify's database utilizing Spotipy, a Python library to access the Spotify API. 100 songs were pulled from each of the five genres of hip hop, rock, jazz, pop, and classical. More than 100 songs from each genre would have been more desirable, but accessing the database for this analysis using Spotipy only allowed 100 songs. The dataset contained the following attributes and contained no null values:

- Track (object)
- Album (object)
- Artist (object)
- Genre (object)
- Popularity (int64)
- Explicit (bool)
- Duration (int64)
- Danceability (float64)
- Energy (float64)
- Key (int64)
- Loudness (float64)
- Mode (int64)
- Speechiness (float64)
- Acousticness (float64)
- Instrumentalness (float64)
- Liveness (float64)
- Valence (float64)
- Tempo (float64)

Exploratory Data Analysis, Data Cleaning, and Feature Engineering

It is crucial to understand the definitions of the acoustic attributes of each song and ChatGPT provided a helpful description:

"Spotify provides several acoustic attributes for tracks, which are derived from its analysis of audio features. These attributes help in understanding the musical and acoustic characteristics of a song. Here are the primary acoustic attributes used by Spotify:

1. **Danceability:**

- **Definition:** Describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity.

- **Range:** 0.0 to 1.0 (higher value indicates higher suitability for dancing).

2. **Energy:**

- **Definition:** Measures the intensity and activity of a track. High energy tracks feel fast, loud, and noisy, while low energy tracks feel slow, quiet, and calm.

- **Range:** 0.0 to 1.0 (higher value indicates higher energy).

3. **Speechiness:**

- **Definition:** Detects the presence of spoken words in a track. Tracks with more spoken words have higher speechiness.

- **Range:**

  - Values below 0.33: Likely music and other non-speech elements.

  - Values between 0.33 and 0.66: Mix of speech and music.

  - Values above 0.66: Predominantly spoken words, like podcasts or talk shows.

4. **Acousticness:**

- **Definition:** Measures the likelihood of a track being acoustic.

- **Range:** 0.0 to 1.0 (higher value indicates higher likelihood of being acoustic).

5. **Instrumentalness:**

- **Definition:** Predicts whether a track contains no vocals. Higher values indicate fewer or no vocal elements.

- **Range:** 0.0 to 1.0 (higher value indicates more instrumental content).

- **Note:** Scores above 0.5 are intended to represent instrumental tracks.

6. **Liveness:**

- **Definition:** Detects the presence of a live audience in the recording. Higher liveness values indicate an increased probability that the track was performed live.

- **Range:** 0.0 to 1.0 (higher value indicates higher liveness).

7. **Valence:**

- **Definition:** Describes the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g., happy, cheerful), while tracks with low valence sound more negative (e.g., sad, angry).

- **Range:** 0.0 to 1.0 (higher value indicates more positive valence).

8. **Tempo:**

- **Definition:** The speed or pace of a track, measured in beats per minute (BPM).

- **Range:** BPM values typically range from 0 to 250 BPM.

9. **Key:**

- **Definition:** The key in which the track is played. Values correspond to pitches using standard Pitch Class notation (e.g., 0 = C, 1 = C♯/D♭, 2 = D, etc.).

10. **Mode:**

- **Definition:** Indicates the modality (major or minor) of a track, where major is represented by 1 and minor is 0.

- **Values:** 0 (Minor) or 1 (Major).

11. **Time Signature:**

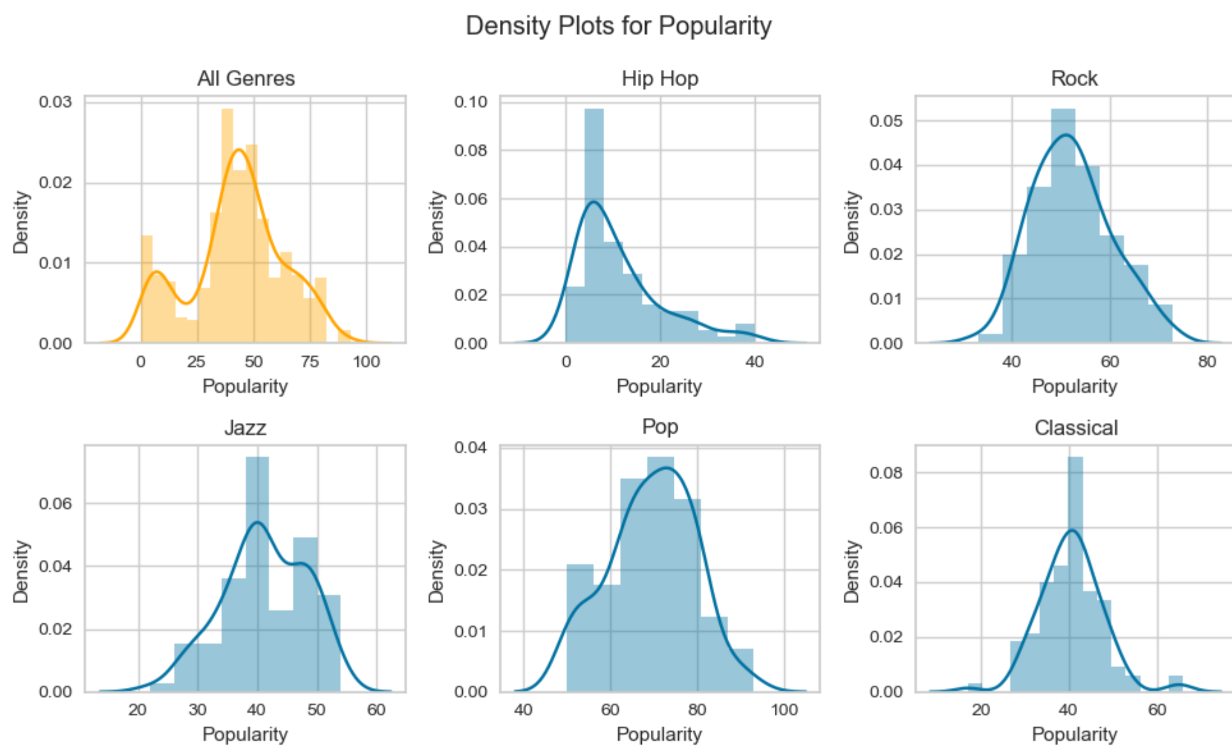
- **Definition:** An estimated overall time signature of a track. The time signature (meter) is a notational convention to specify how many beats are in each bar (or measure).

- **Common Values:** Typically 3, 4, or 5 (representing 3/4, 4/4, or 5/4 time, respectively).

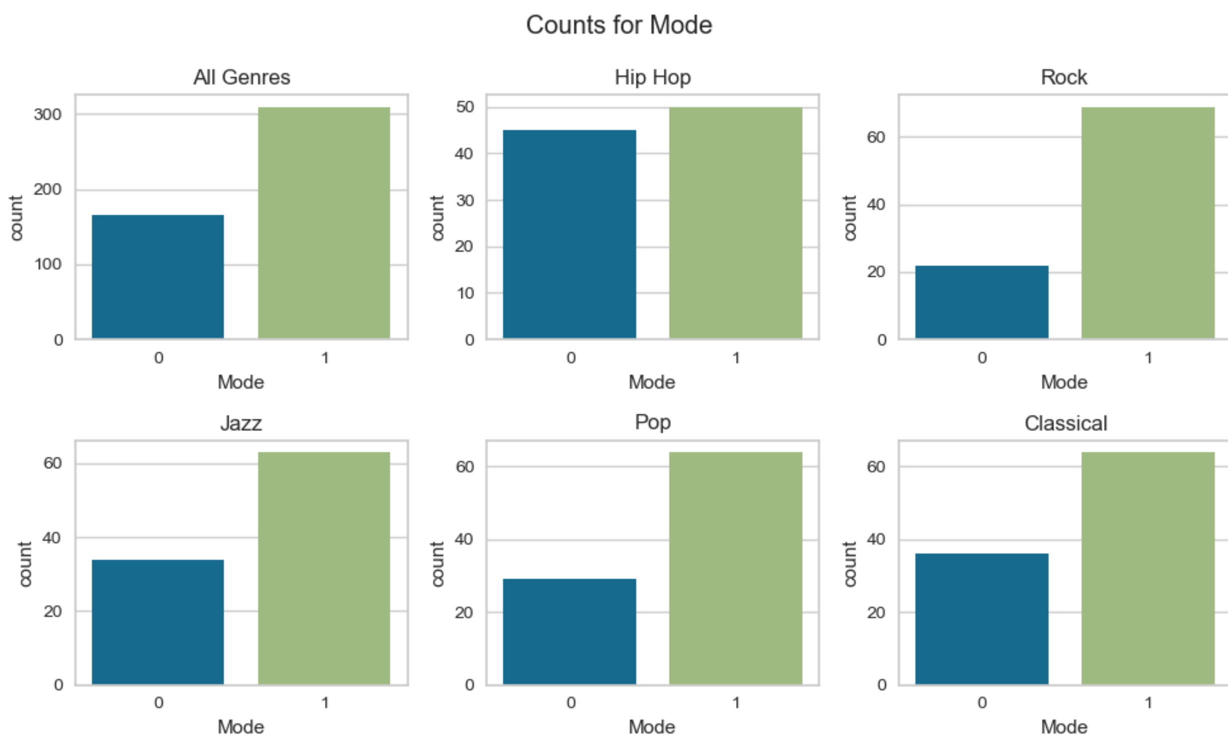
These attributes provide a comprehensive overview of the sonic characteristics of a track and are widely used in recommendation systems, playlist generation, and music analysis."

The genres and acoustic attributes in relation to popularity were compared and contrasted for the analysis, as well as how the acoustic attribute of mode plays on popularity within each genre.

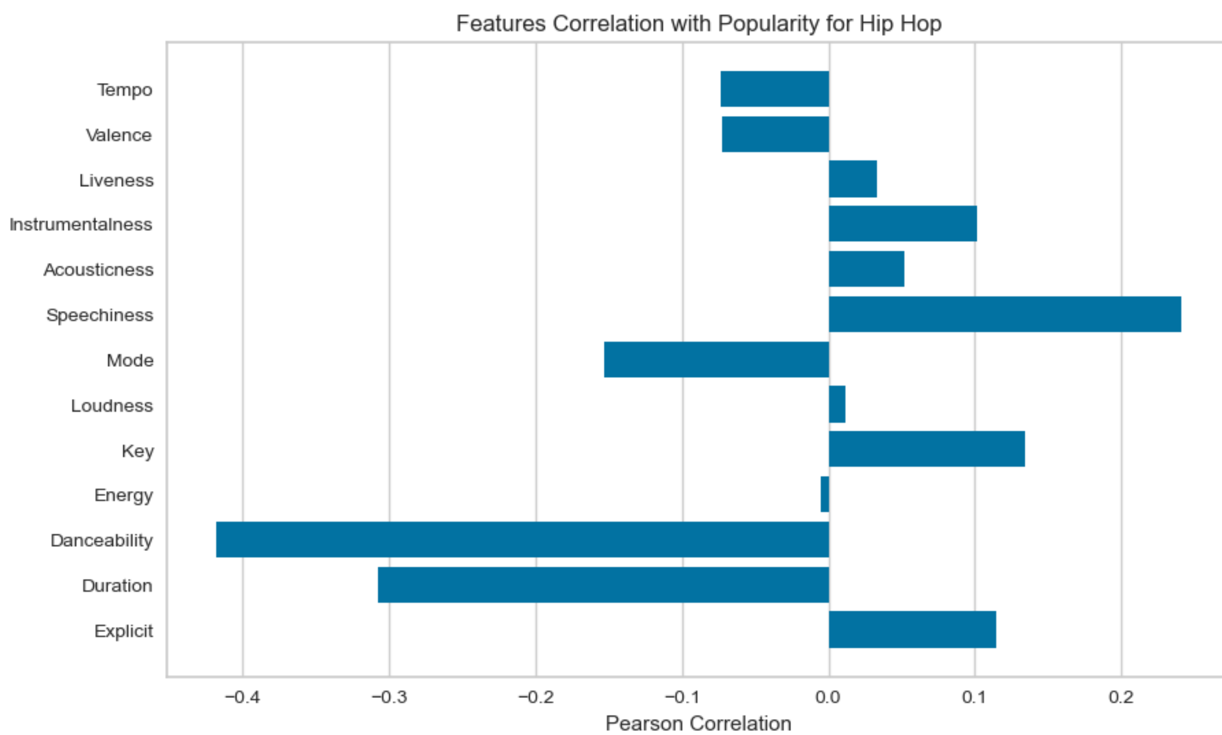
In the distribution plots for popularity, hip hop is skewed and less popular songs are more numerous, perhaps due to the fact that there is an entire plethora of underground hip hop geared towards smaller audiences. Rock, pop, and classical followed a more normal distribution, as relatively few songs in top 10 lists can turn the focus on listeners to a more limited variety of songs and genres such as classical contain a set of very popular and familiar pieces. Jazz is a genre that encompasses many distinct styles, anything from swing to more dissonant free jazz and popularity varies between these subgenres, as well as the type of listener.

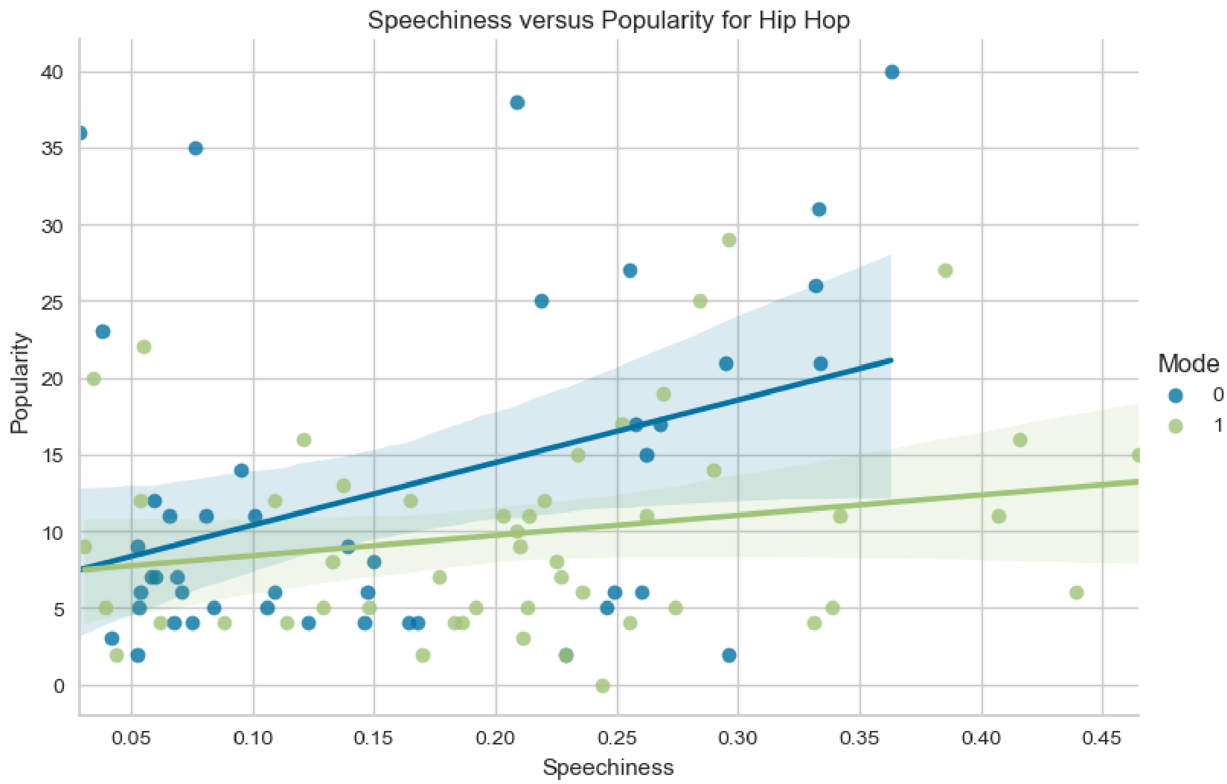


A mode of 1 denotes that the song was in a major scale and 0 a minor one. Major scales have a happier, brighter sound while a minor scale has a sadder, darker sound. In every genre included in this analysis, songs with mode 1 were more numerous, but by how much varied across genres. In hip hop, the two modes were more even while in rock, they were the most distinct. Jazz is a style of music that incorporates both the blues scale akin to mode 0 as well as cadences, which are mode 1. Classical pieces also vary between mode 0 and mode 1, but overall, songs with the mode 1 acoustic attribute are more numerous in both genres.

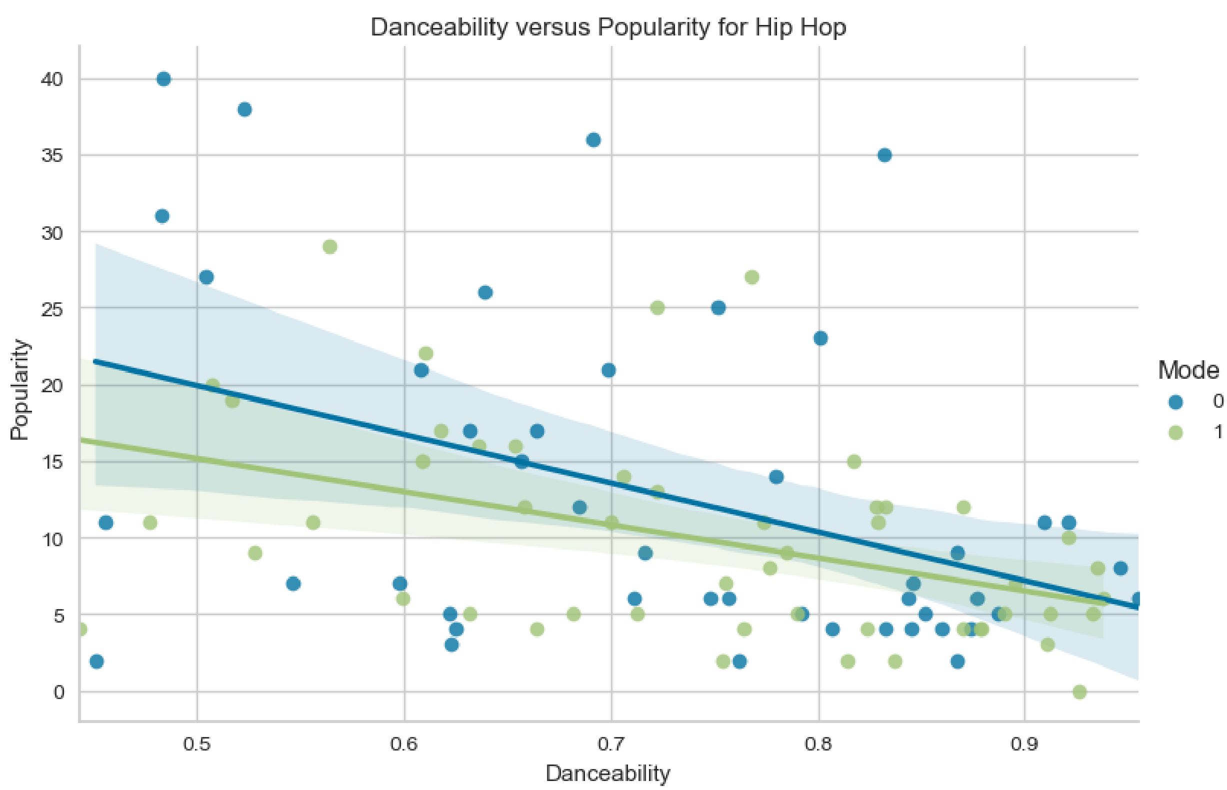


In the hip hop genre, the largest positive correlation in acoustic attribute to popularity is speechiness and this makes sense, as it is a largely lyrically driven genre. Mode has a negative correlation, indicating songs in a minor key drive popularity more. An odd negative correlation to popularity is danceability, how the genre is often associated with dance parties and social events. The danceability acoustic attribute is partly based on faster tempo and rhythm stability and hip hop songs typically play at a slower tempo and the rhythms are syncopated, so while people dance to it, the nature of the songs do not fit neatly into Spotify's definition. Perhaps this is where Spotify can make improvements as far as how to measure danceability to improve the function of the recommendation engine for hip hop.



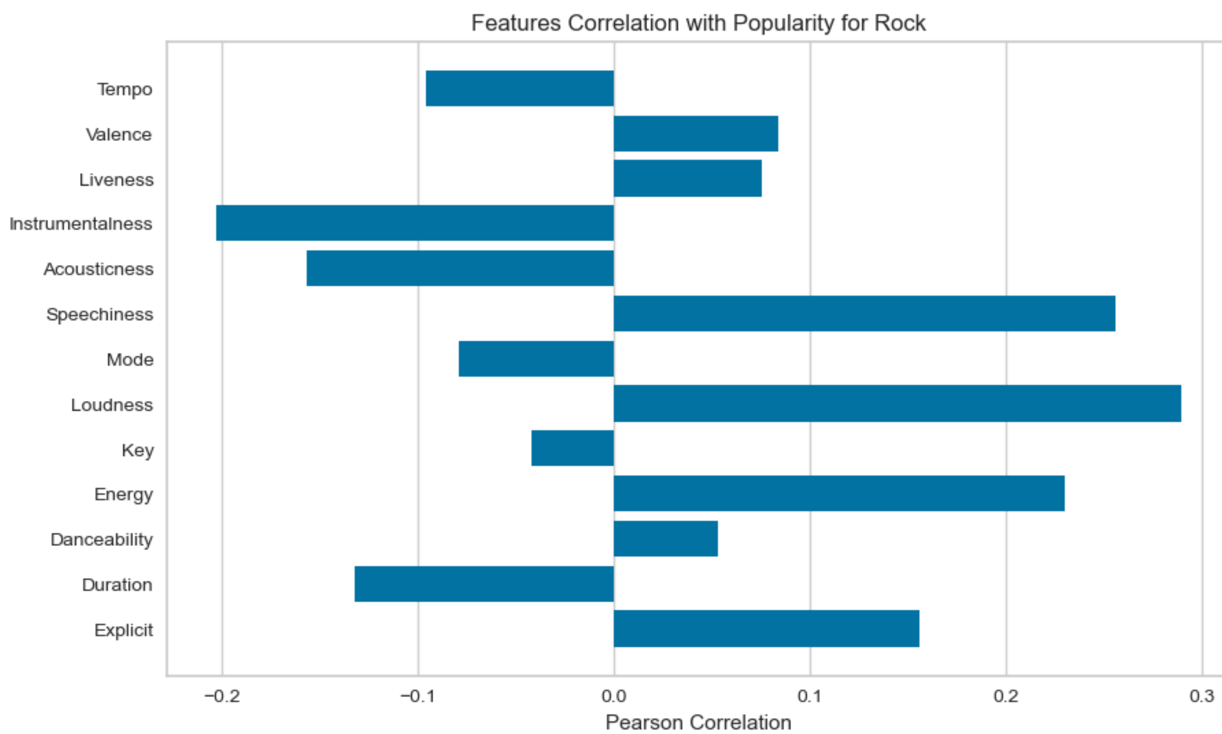


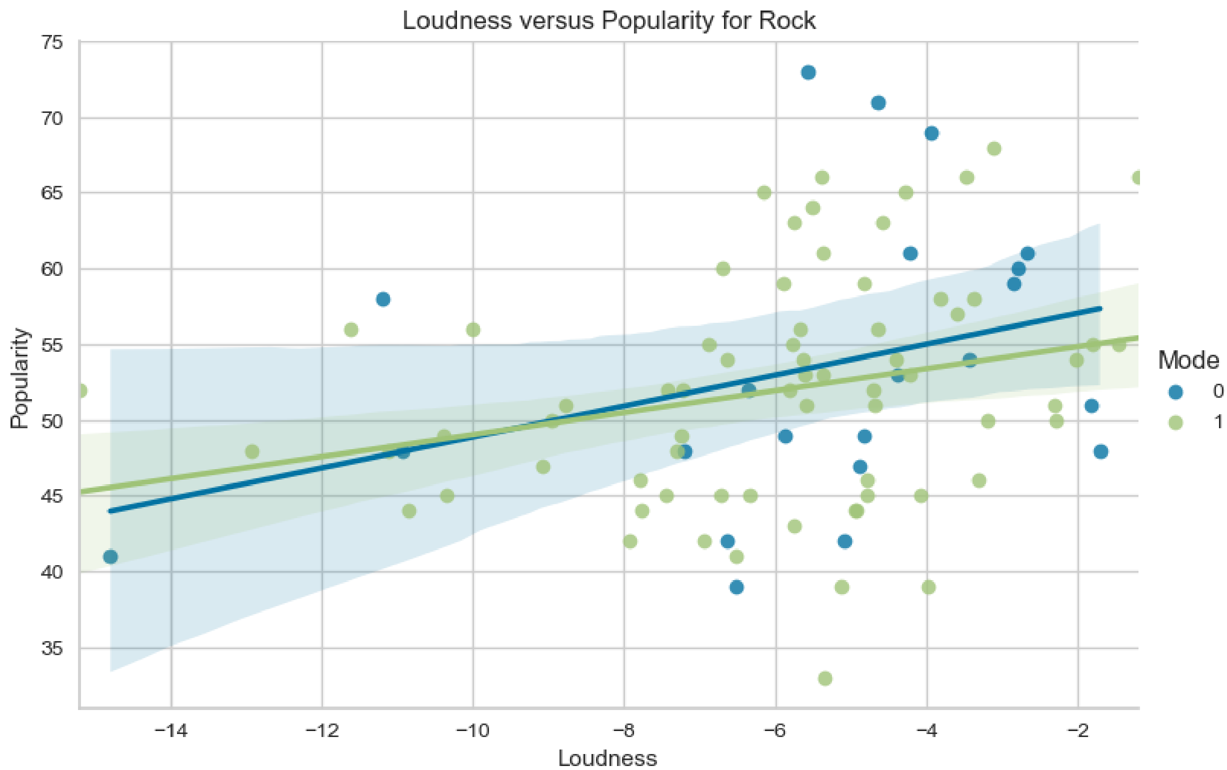
Speechiness versus Popularity for Hip Hop
Popularity
Speechiness
0
5
10
15
20
25
30
35
40
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
Mode
0
1

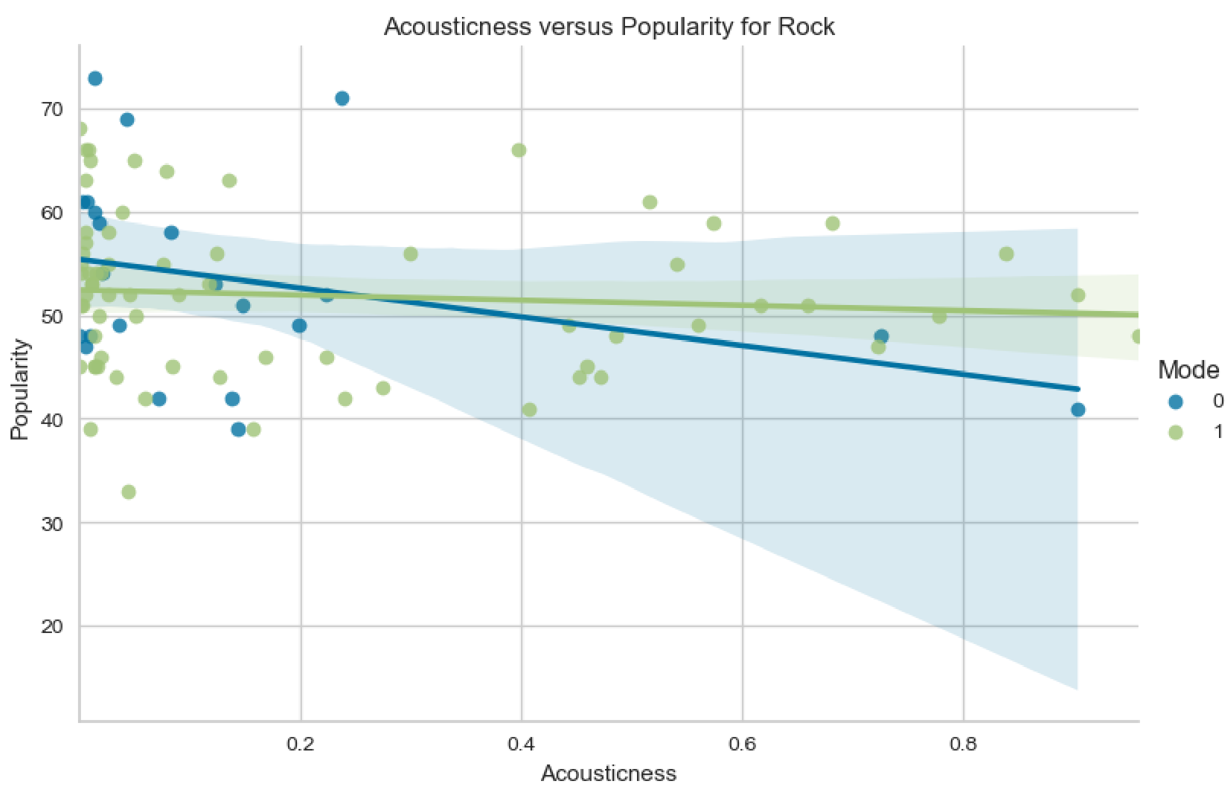


Danceability versus Popularity for Hip Hop
Popularity
Danceability
0
5
10
15
20
25
30
35
40
0.5
0.6
0.7
0.8
0.9
Mode
0
1

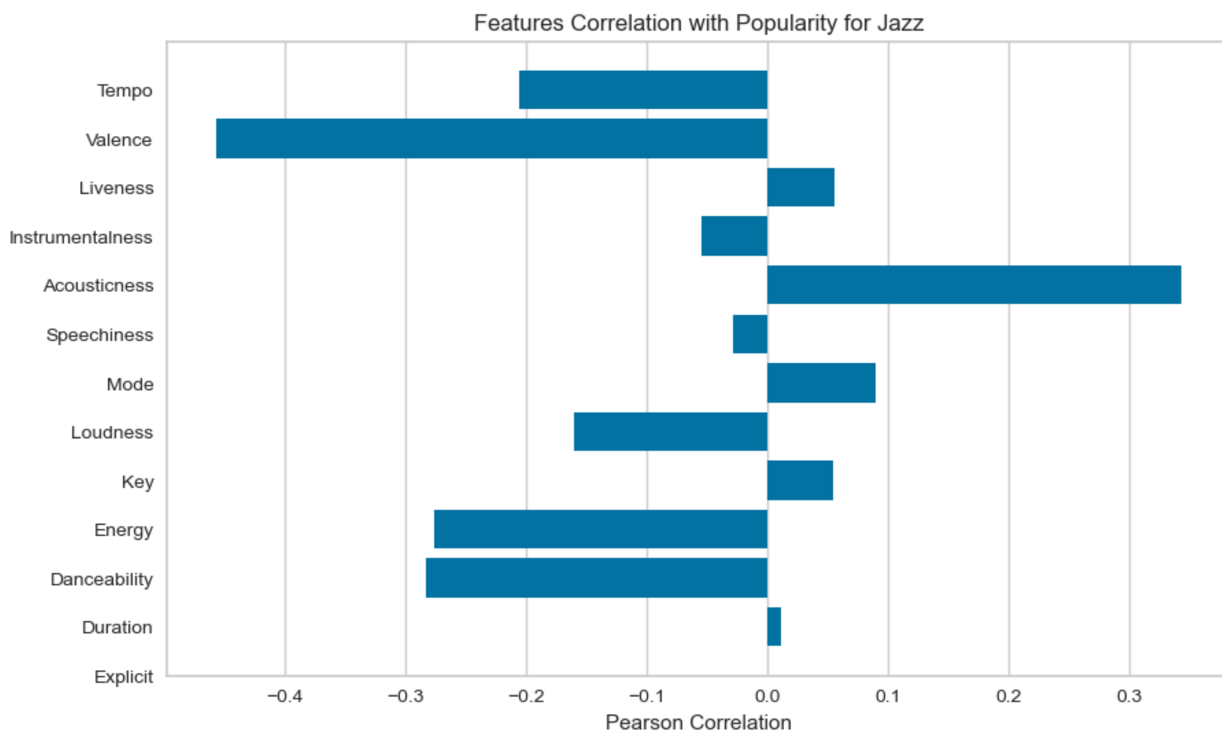
In the rock genre, loudness has the most positive correlation to popularity and energy has a high positive correlation as well. One associates rock concerts in large venues with large, loud sound systems as well as the instrumentation itself, with amped electric guitars and big drum kits. The mode for popularity leans towards 0 as a lot of popular rock songs make heavy use of the minor pentatonic scale. Acousticness is the second to most negative correlation to popularity and while notable famous acoustic rock songs exist, popularity leans more towards songs with electric instruments.

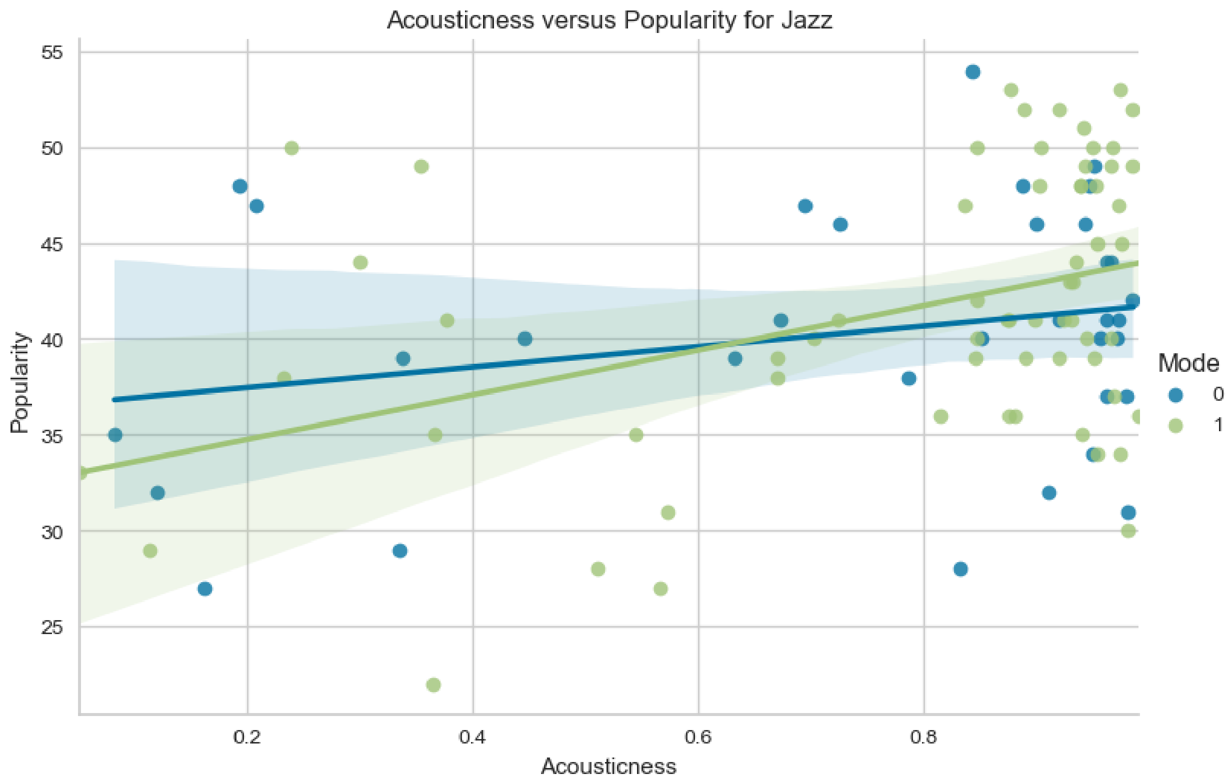


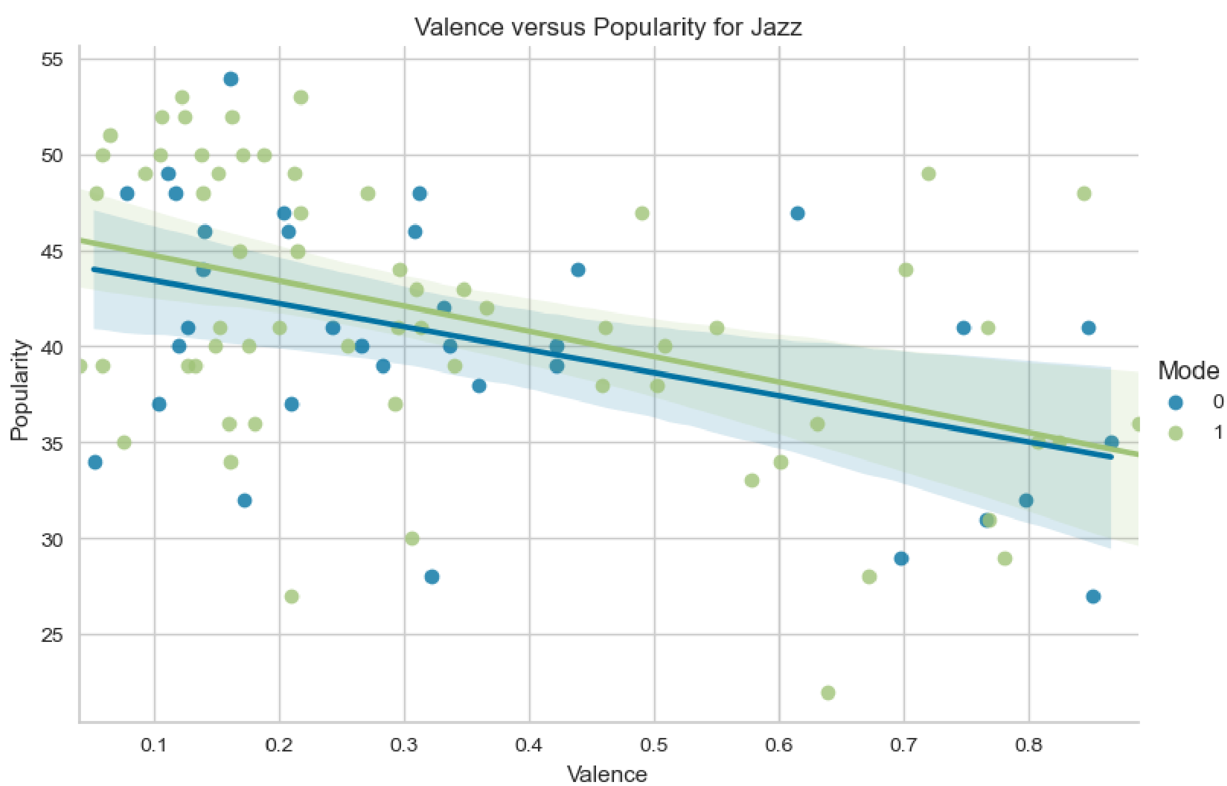

Loudness versus Popularity for Rock
Popularity
Loudness
75
70
65
60
55
50
45
40
35
−14
−12
−10
−8
−6
−4
−2
Mode
0
1


Acousticness versus Popularity for Rock
Popularity
Acousticness
70
60
50
40
30
20
0.2
0.4
0.6
0.8
Mode
0
1

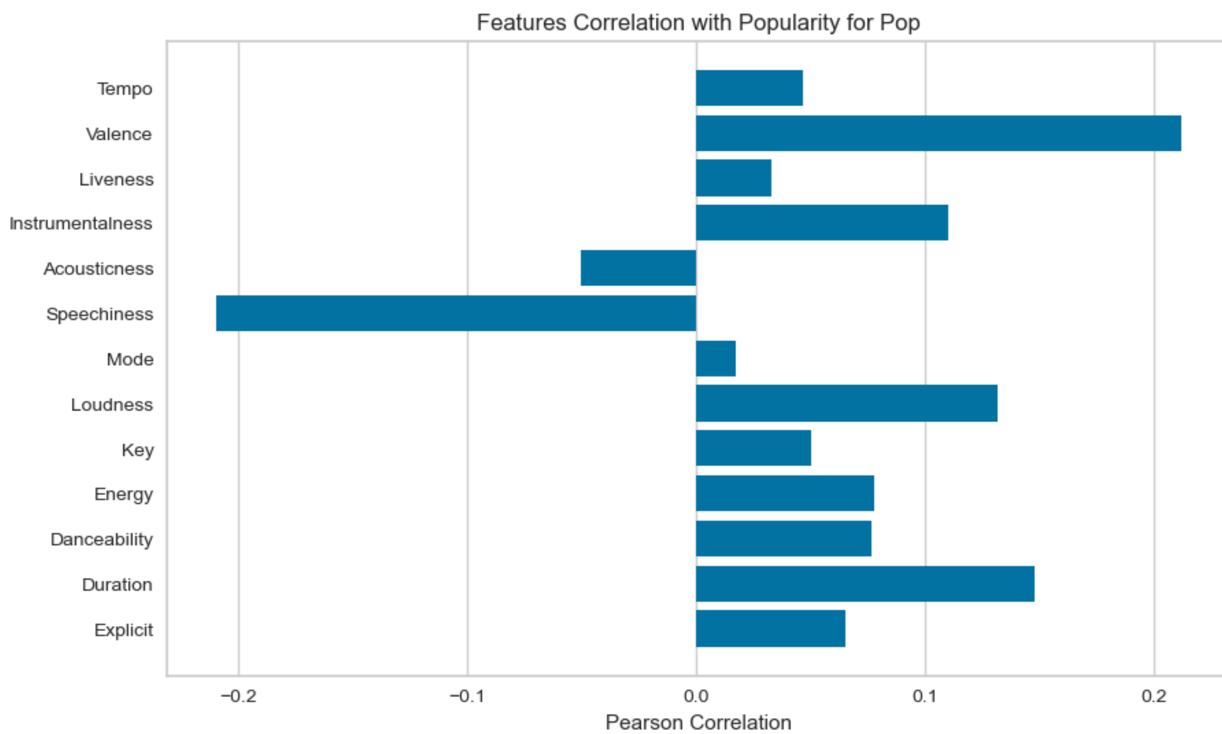
In the jazz genre, acousticness has the largest positive correlation to popularity. The vast majority of jazz songs aside from smaller subgenres like fusion are comprised of acoustic instruments like piano, double bass, and horns. More popular jazz songs are in major keys. Despite the mode, valence has the most negative correlation to popularity. Jazz songs tend to sound more somber, especially in smaller settings such as a quartet or piano trio. And not surprisingly, acoustic attributes such as energy and danceability have negative correlations to popularity, jazz is a much more distinct musical genre compared to rock or pop.

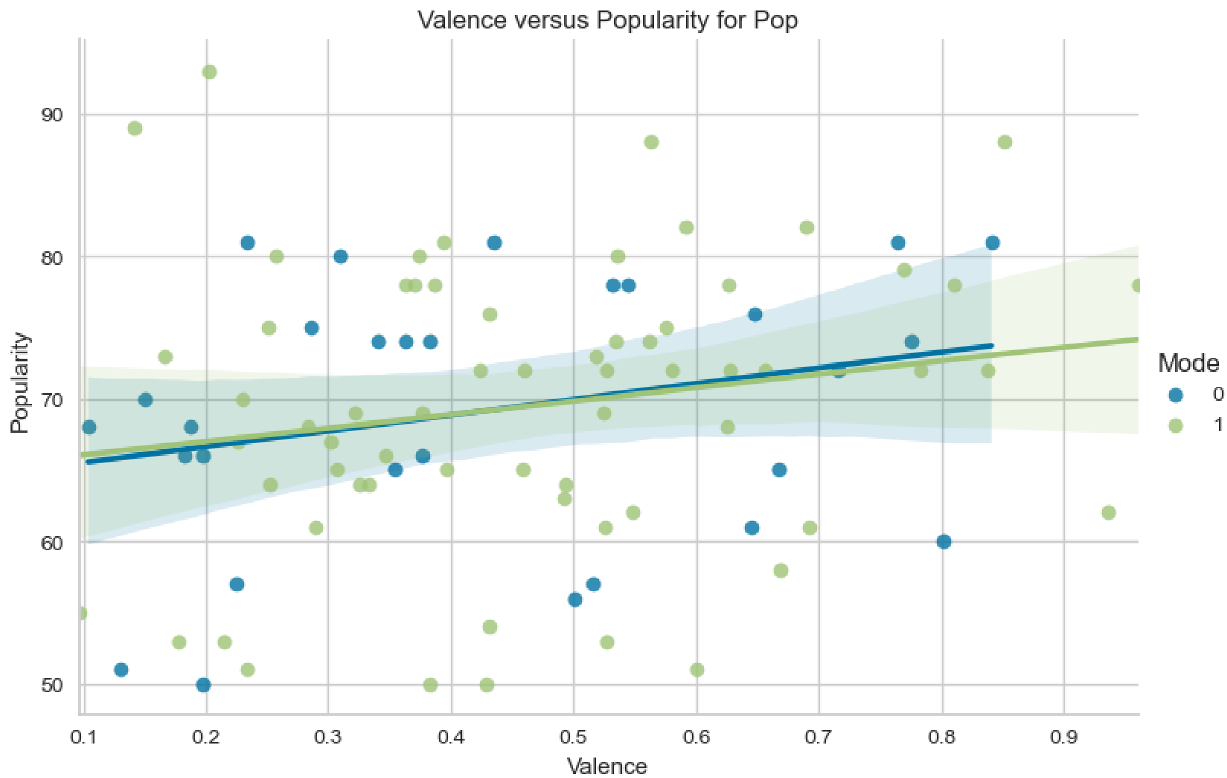


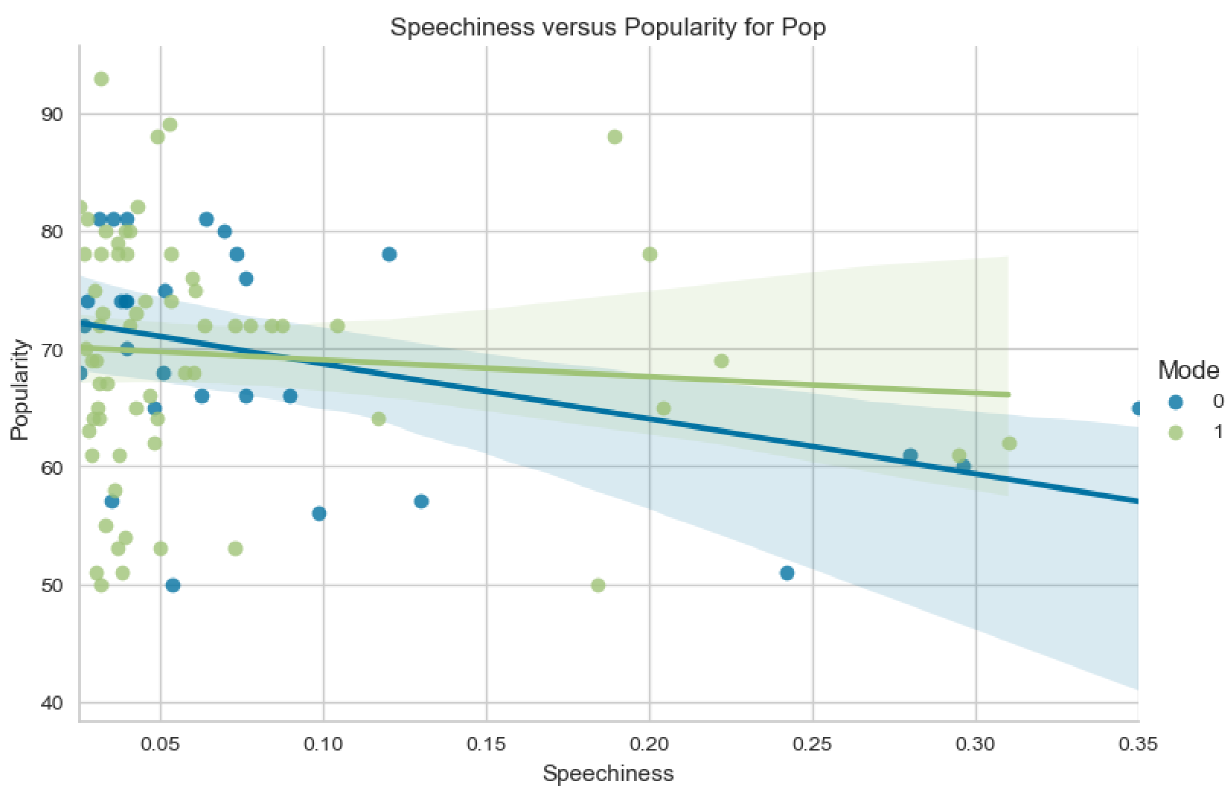

Acousticness versus Popularity for Jazz
Popularity
Acousticness
Mode
0
1


Valence versus Popularity for Jazz
Popularity
Valence
Mode
0
1

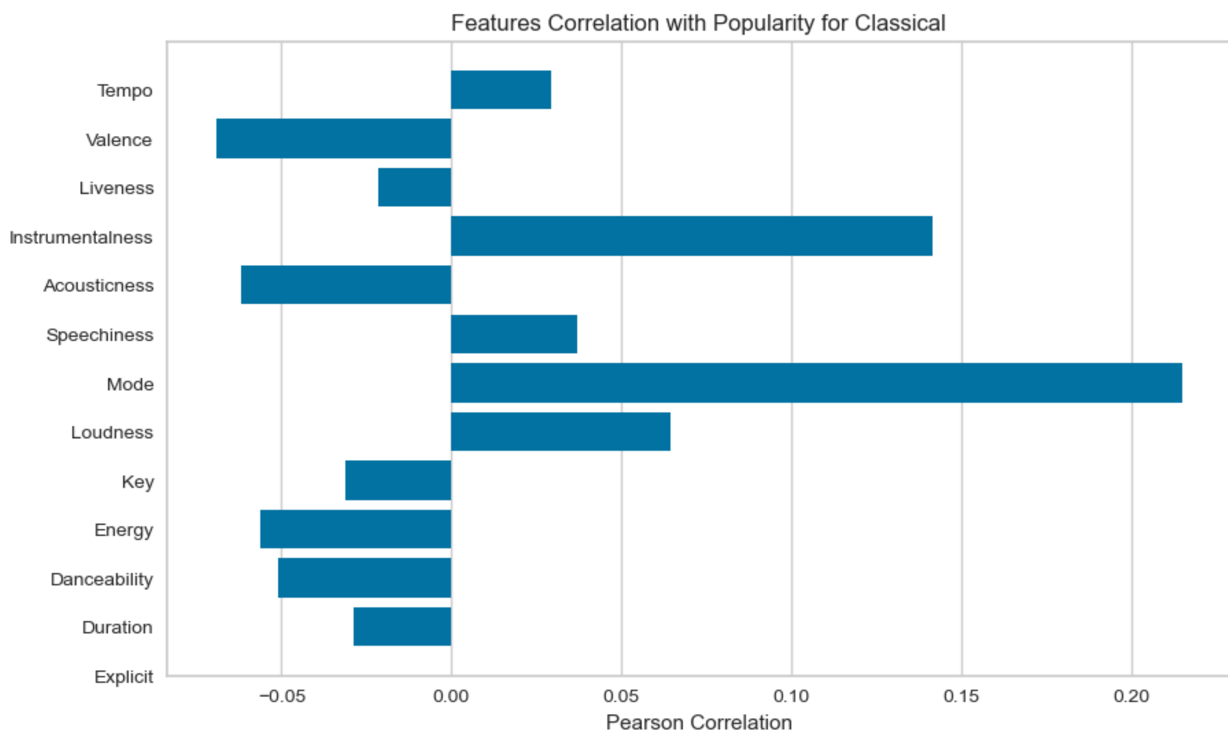
In the pop genre, valence is what correlates most positively with popularity. The most prominent pop songs sound happy, bright, and universally appealing. And as with the rock genre, a lot of pop artists play in stadiums where the volume is loud and songs streaming on Spotify would be designed to replicate that effect. The acoustic attribute with the most negative correlation to popularity is speechiness. Pop songs tend to have simple, repetitive lyrics with less variety, especially compared to a genre like hip hop. Interestingly, even with valence being the chief driver for popularity, mode only has a relatively slight effect, meaning whether a song is in a major or minor key is not a key determiner for popularity.

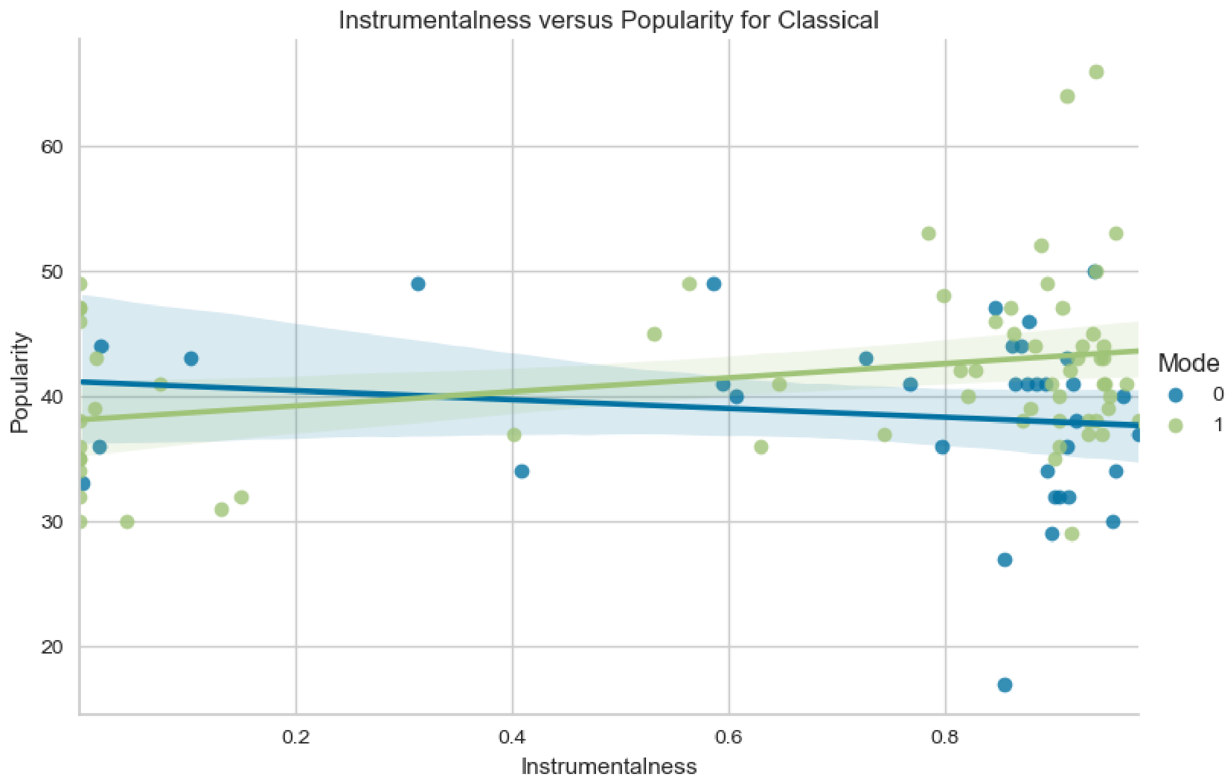


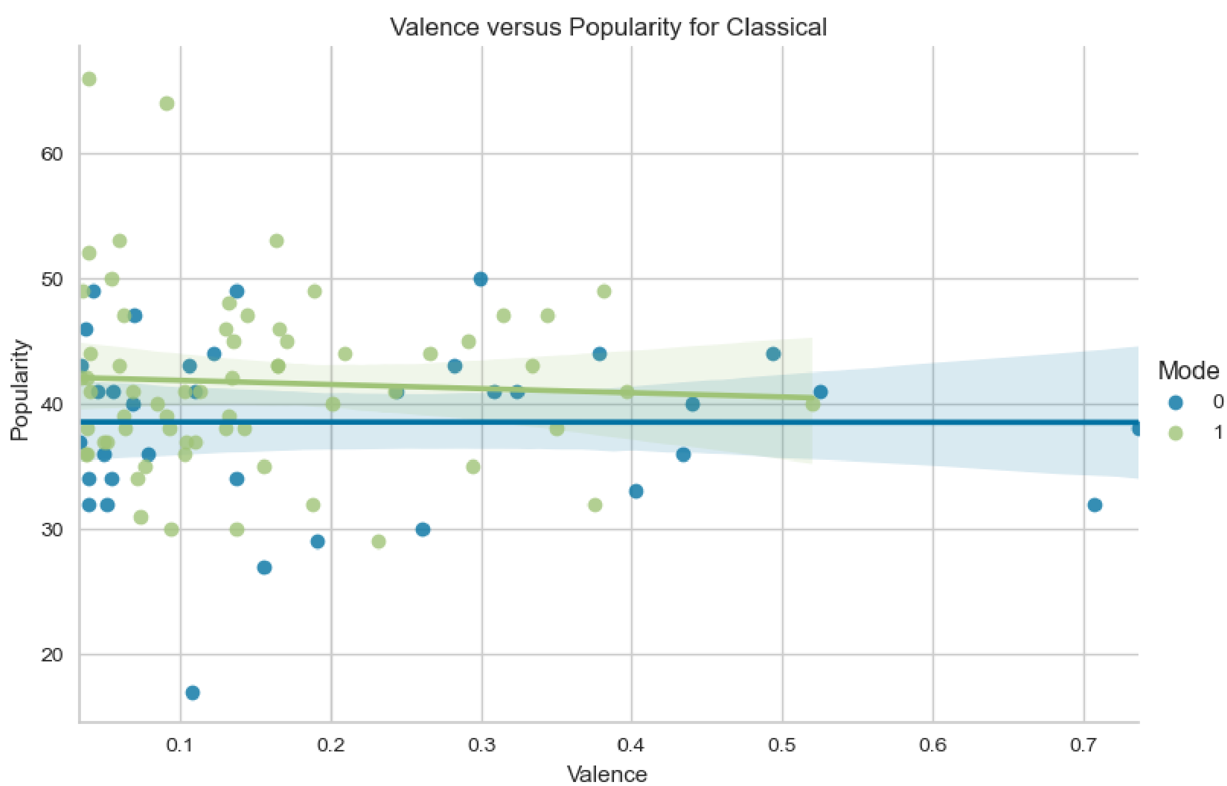

Valence versus Popularity for Pop
Popularity
Valence
Mode
0
1
90
80
70
60
50
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


Speechiness versus Popularity for Pop
Popularity
Speechiness
Mode
0
1
90
80
70
60
50
40
0.05
0.10
0.15
0.20
0.25
0.30
0.35

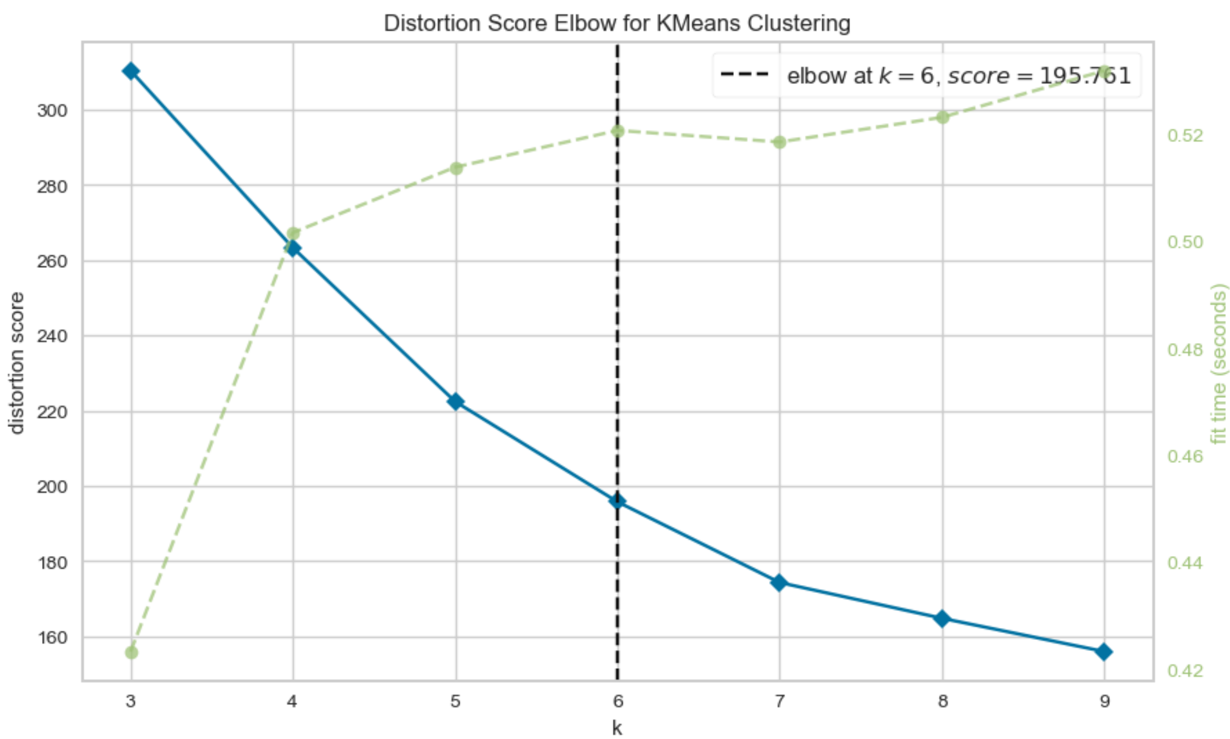
In the classical genre, mode has the highest positive correlation with popularity. Songs in major keys drive popularity the most. Instrumentalness has the second most positive correlation to popularity, as aside from opera, the vast majority of classical songs solely consist of instruments. Valence has the most negative correlation to popularity. The classical genre seems like the antithesis of the pop genre and very bright, electronic sounds are absent. Instead, instrumentation such as bowed acoustic string instruments dominate, which leans to genre to sound darker.



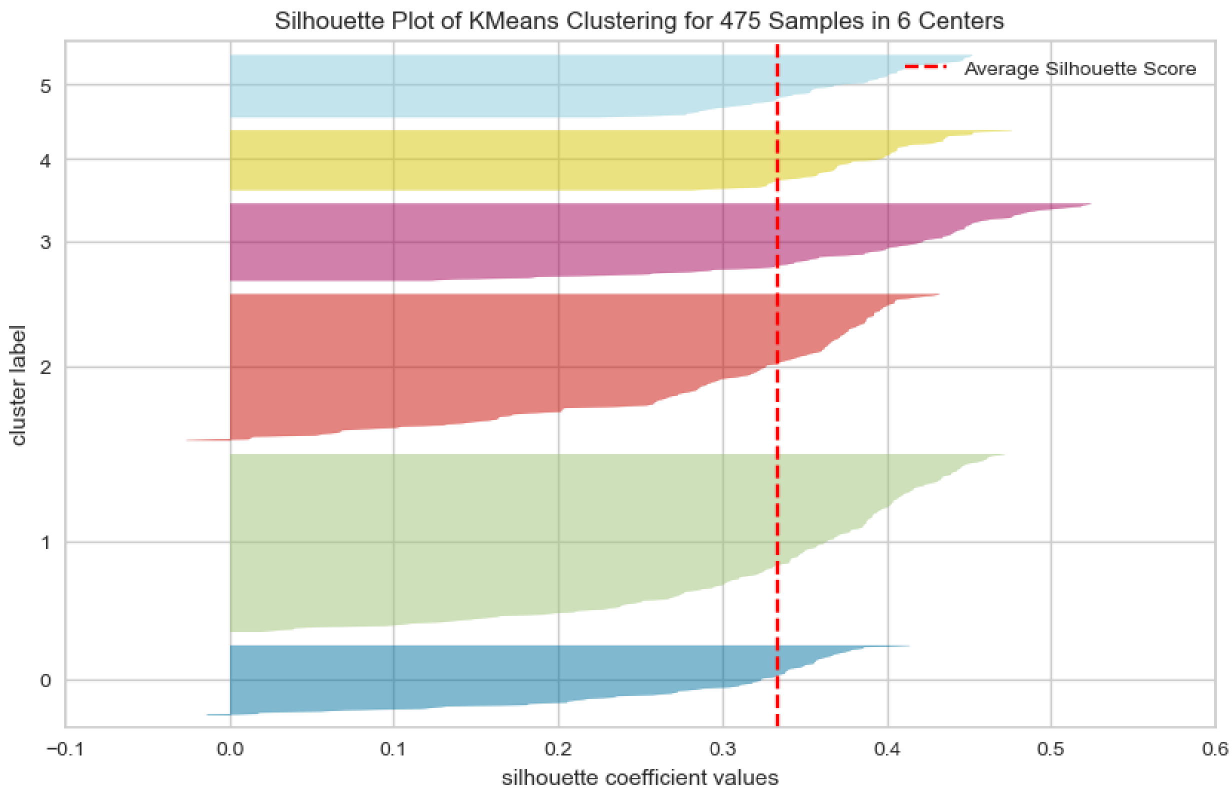

Instrumentalness versus Popularity for Classical
Popularity
60
50
40
30
20
0.2
0.4
0.6
0.8
Instrumentalness
Mode
0
1


Valence versus Popularity for Classical
Popularity
60
50
40
30
20
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Valence
Mode
0
1

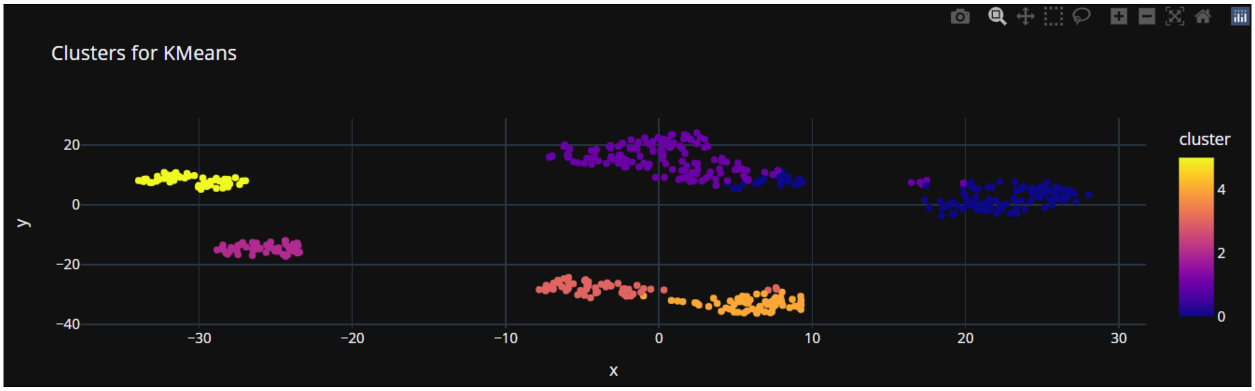
Unsupervised learning clustering methods were used to group together data points based on acoustic attributes to compare them to how Spotify defines songs by genre. KMeans, DBSCAN, and Mean Shift algorithms were utilized to perform the clustering. Out of the three, DBSCAN and Mean Shift are true clustering algorithms while KMeans requires the number of centroids to be defined. The Yellowbrick library was used to determine the optimal number of centroids to use for KMeans. An elbow curve was plotted and the cluster number with the best distortion score was determined to be 6.

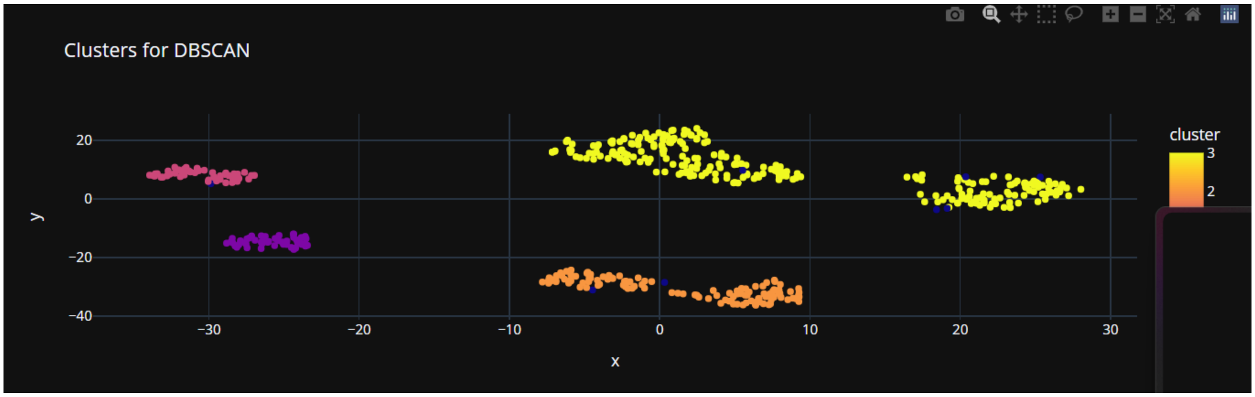


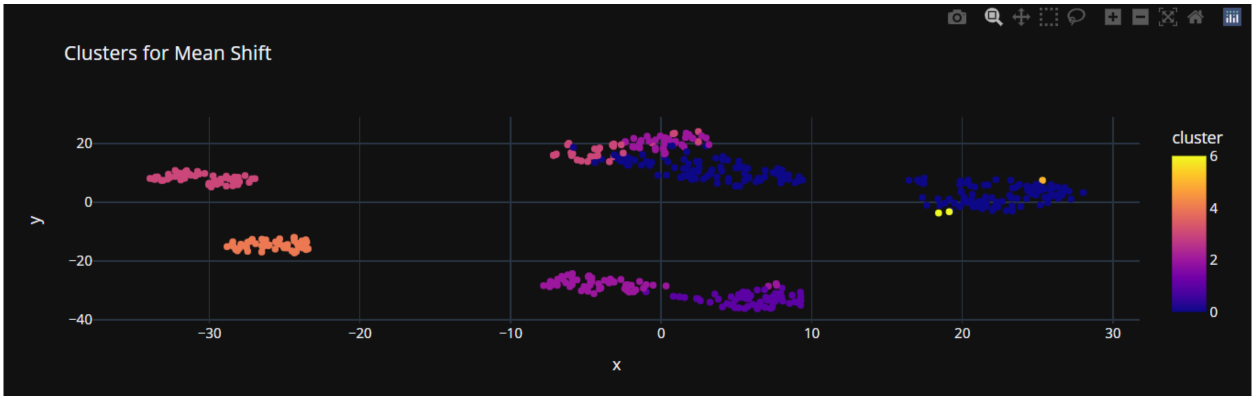
Yellowbrick can only make silhouette plots for centroid-based algorithms and with 6 centroids, the silhouette score for the KMeans model was 0.33. A grid search was performed to determine the best hyperparameters for DBSCAN and Mean Shift and for the optimal models, the silhouette scores were 0.21 for DBSCAN and 0.20 for Mean Shift. Out of the three unsupervised learning models, KMeans had the best silhouette score to form clusters based on acoustic attributes of this sample of songs from the Spotify database.



Interestingly, while the songs in the analysis were from 5 different genres, KMeans had 6 clusters, DBSCAN 4 clusters, and Mean Shift 7 clusters upon visualization.







Key Findings and Insights

Comparing different genres from the Spotify database revealed how distinct each of these categories are, based on acoustic attributes the app uses to make recommendations and create playlists. Overlaps existed between genres with smaller audiences such as jazz and classical, as well as more popular genres such as rock and pop. What was notable was that mode, whether the song is in a major or minor key, didn't necessarily correspond to valence, or the general positive or negative mood of the song in certain genres. Some unexpected correlations for popularity arose as well for acoustic attributes for the different genres. However, the sample size was much smaller than desired. Clustering algorithms produced various results when visualized and even for the best performing, KMeans, a strange anomaly is clearly visible.

Next Steps

The data from the Spotify API was clean and reliable, but only 100 songs were pulled from each of the 5 genres. This was the biggest hurdle in doing a better analysis, the small sample size. While a centroid-based clustering algorithm like KMeans can perform fairly on a relatively smaller dataset, true clustering algorithms like DBSCAN and Mean Shift require a bigger sample size to perform adequately. The most obvious improvement would be to pull substantially more data from the database. This would also improve analysis of correlations between various acoustic attributes and popularity and provide a much more robust and refined understanding. Other genres can be compared and contrasted as well. Finally, in addition to the clustering algorithms used in this analysis, others can be utilized as well, such as hierarchical clustering.