MUSIC RECOMMENDATION USING CLUSTERING

ADVAIT KULKARNI
SUHIT DATTA
VARAD TUPE

University at Buffalo The State University of New York

Agenda

- General Idea & Data Acquisition
- Exploratory Data Analysis
- Clustering
- Applications and Future Scope

General Idea

- Extract the quantitative nuances of a song.
- Can we group them together based on the data?
- Clustering based on similarity of songs (Audio Features) and not just on Genres.
- Build a Recommendation system that would pick songs from these clusters : Such songs would have similar harmonics/characteristics.

Data Acquisition

- Data selected from Spotify Top charts
- Genre:
  - Pop
  - EDM
  - Metal
  - Acoustic
  - Country
  - Rap

Data Acquisition

- spotify-api for fetching data from Spotify Web-API
- pygn for fetch data from Gracenote API
- Get Playlist's Tracks API for fetching track ids of songs
- Get Audio Features for Several Tracks API for fetching audio features

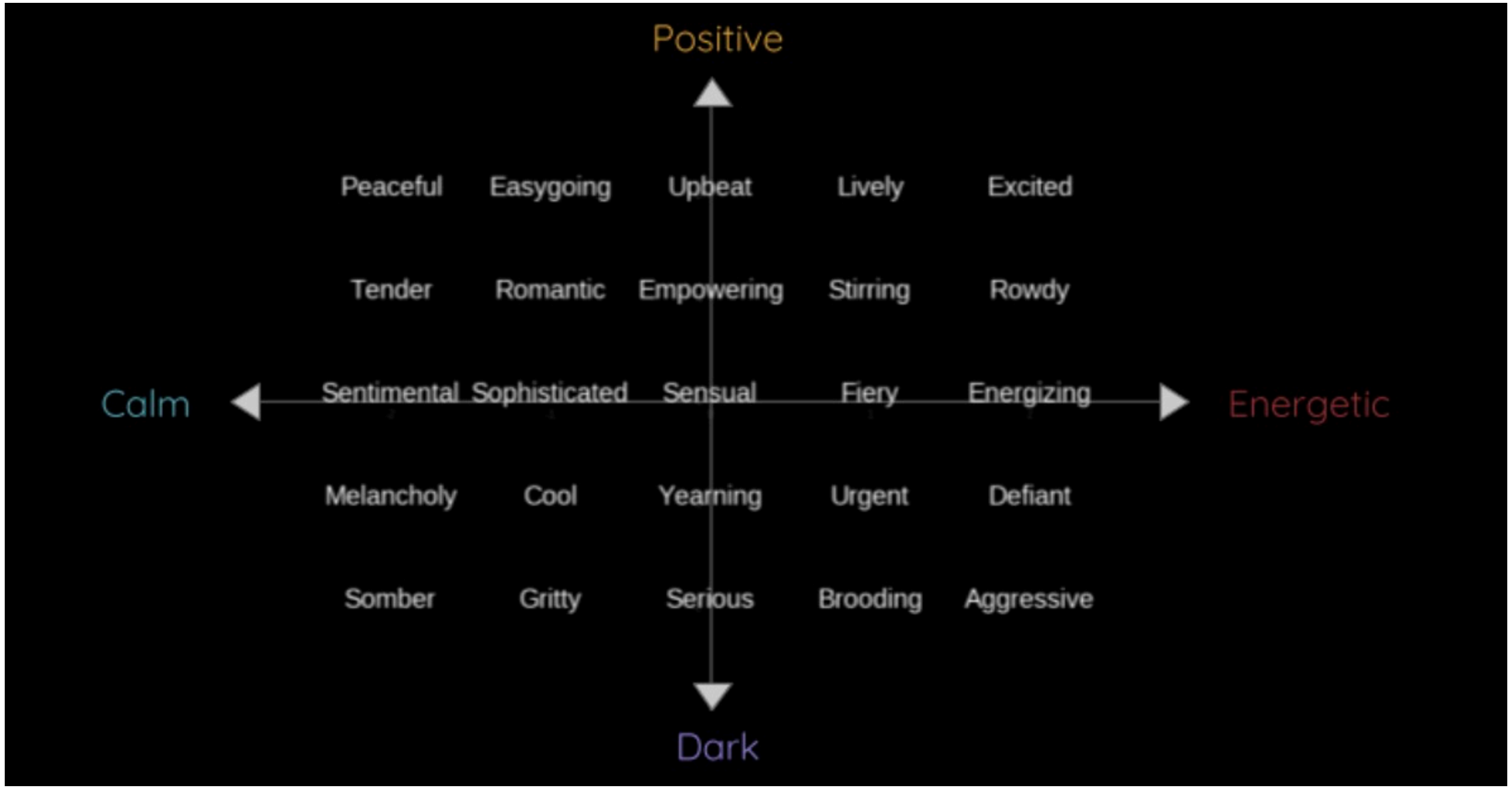
Features

Feature	Source	Description
artist	Spotify	Artist Name
album	Spotify	Album Name
track_name	Spotify	Track Name
genre	Manually Tagged	Genre of playlist
acousticness	Spotify	A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
danceability	Spotify	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.
duration_ms	Spotify	The duration of the track in milliseconds.
energy	Spotify	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.
instrumentalness	Spotify	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.

Features

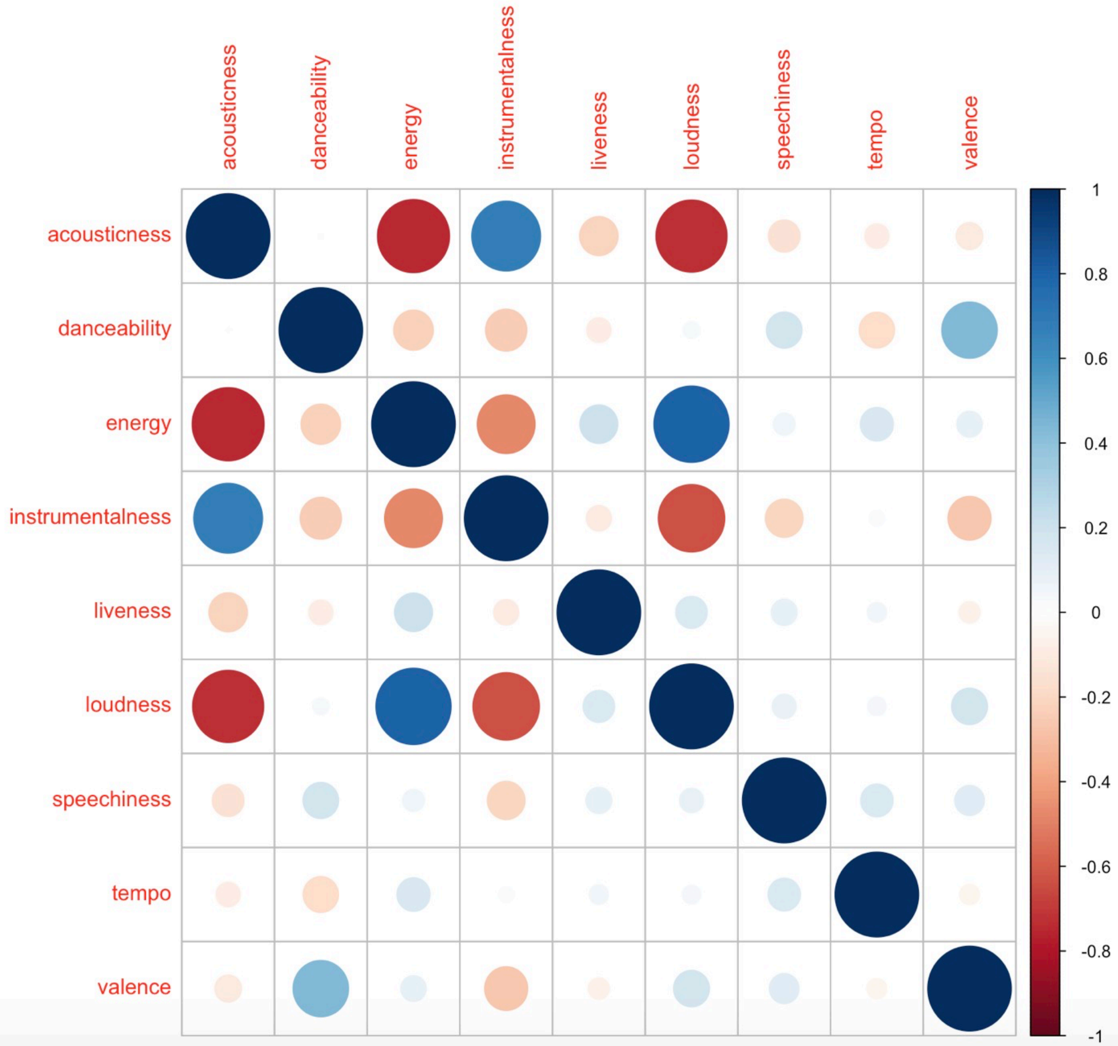
Feature	Source	Description
key	Spotify	The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on.
loudness	Spotify	The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typical range between -60 and 0 db.
mode	Spotify	Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.
speechiness	Spotify	Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, 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.
tempo	Spotify	The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
time_signature	Spotify	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).
valence	Spotify	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).
mood	Gracenote	Mood of a song

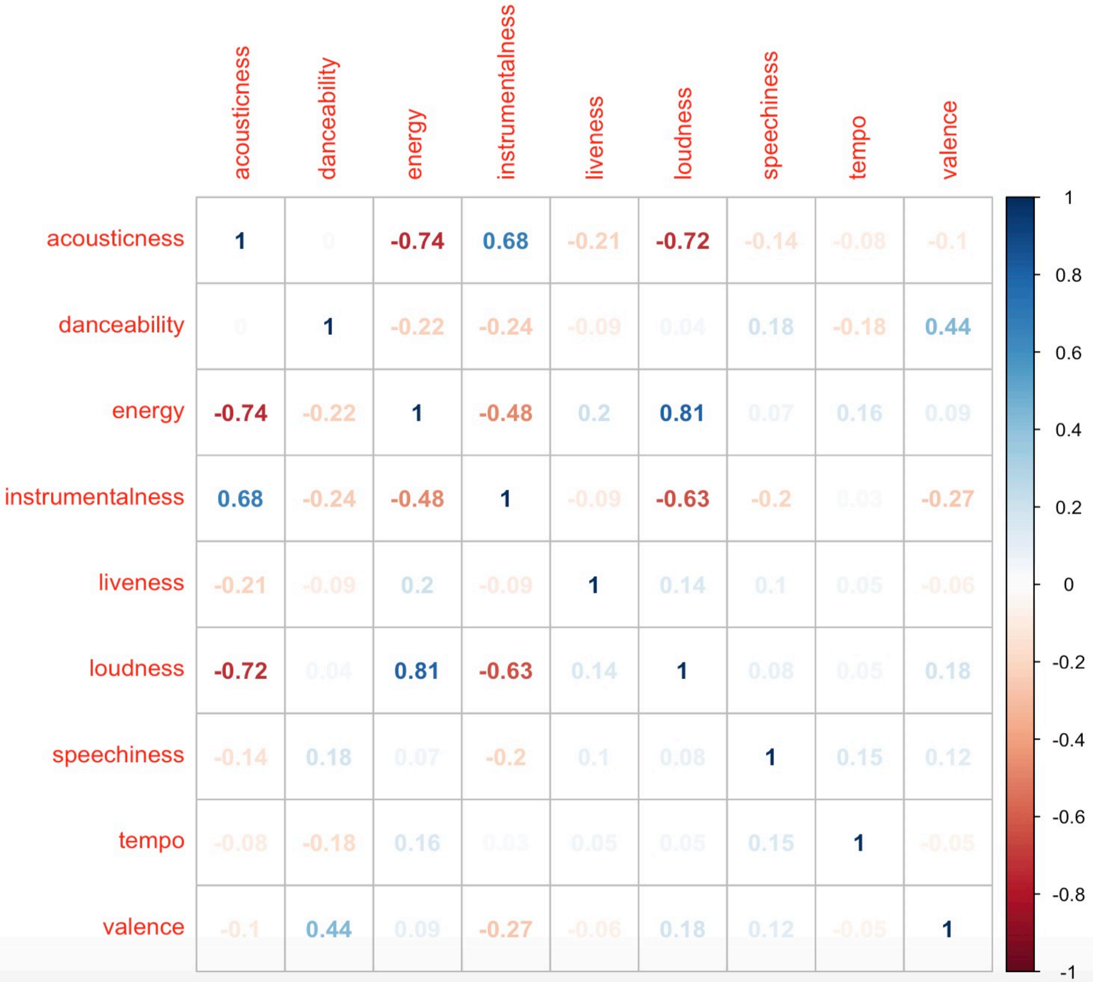
Mood Taxonomy



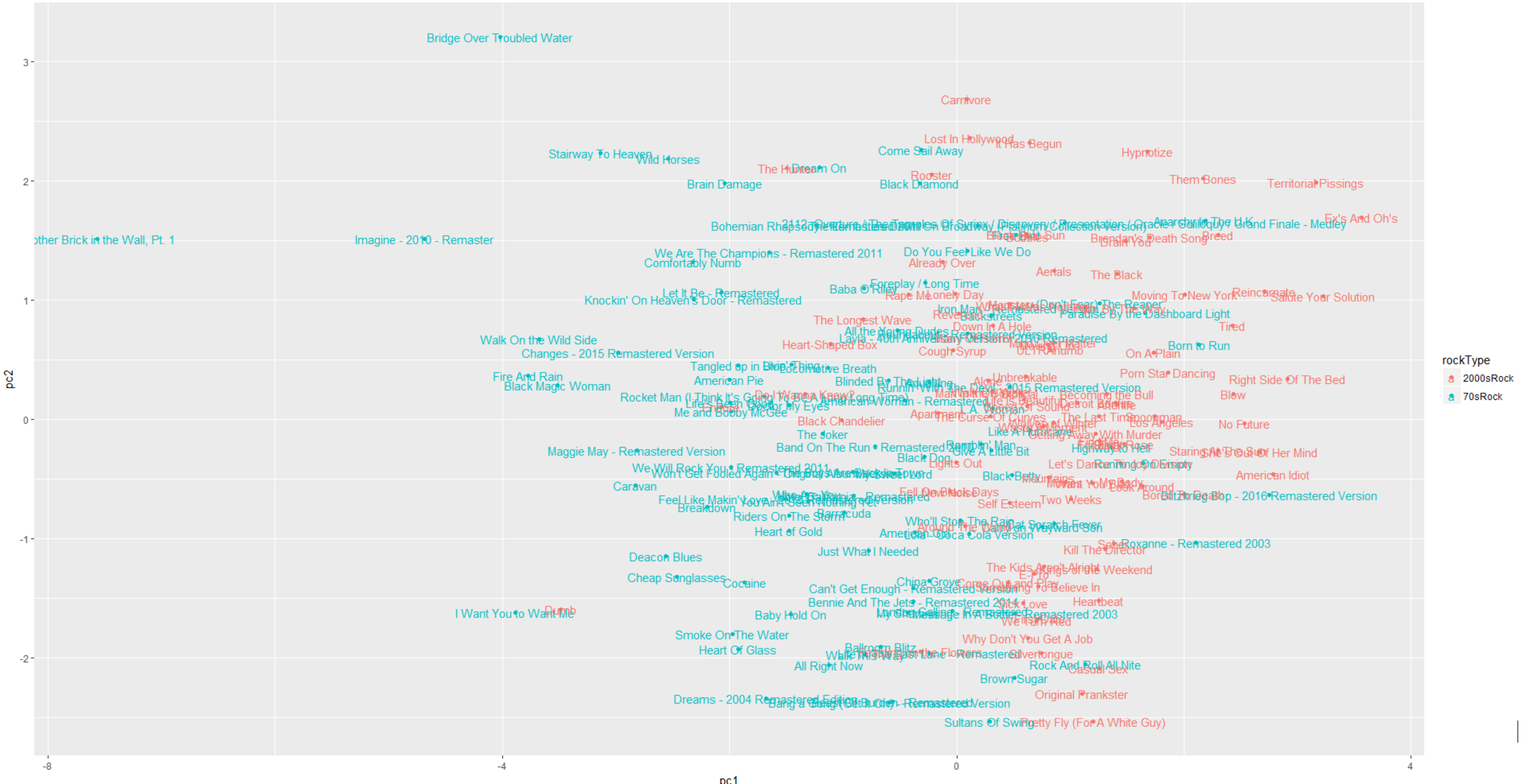
https://neokt.github.io/projects/audio-music-mood-classification/

Correlation Plot between the variables :

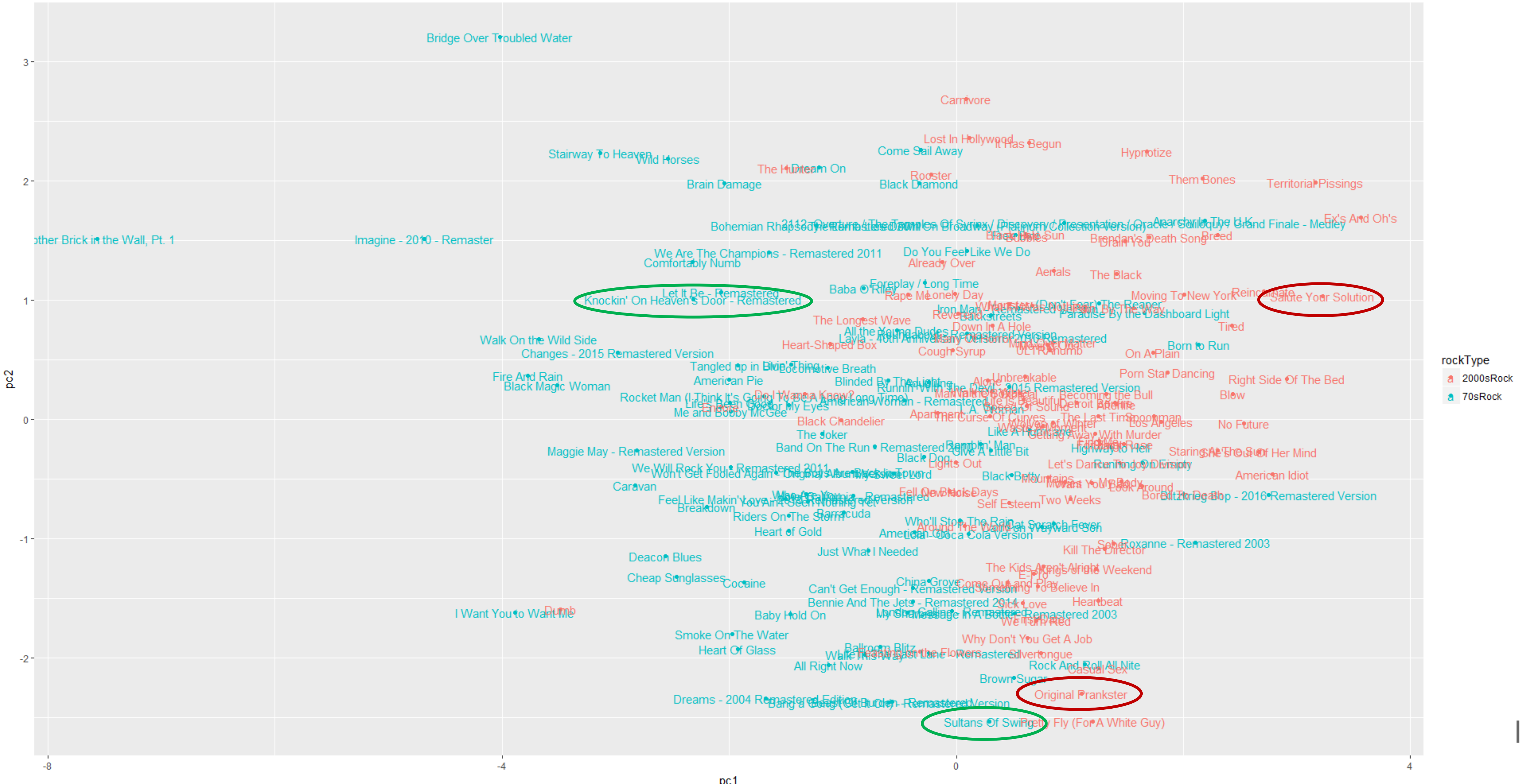




Appetizer for Rock Enthusiasts!



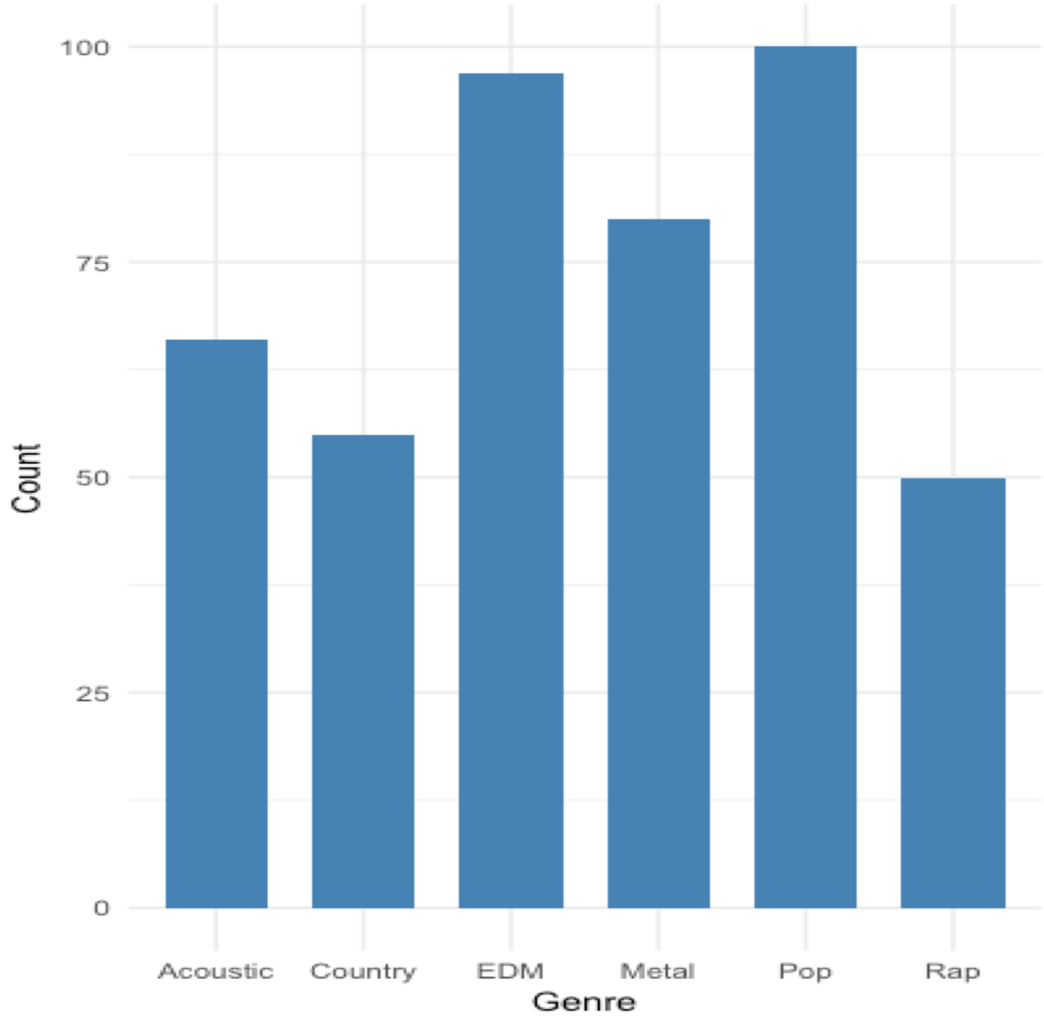
Appetizer for Rock Enthusiasts!



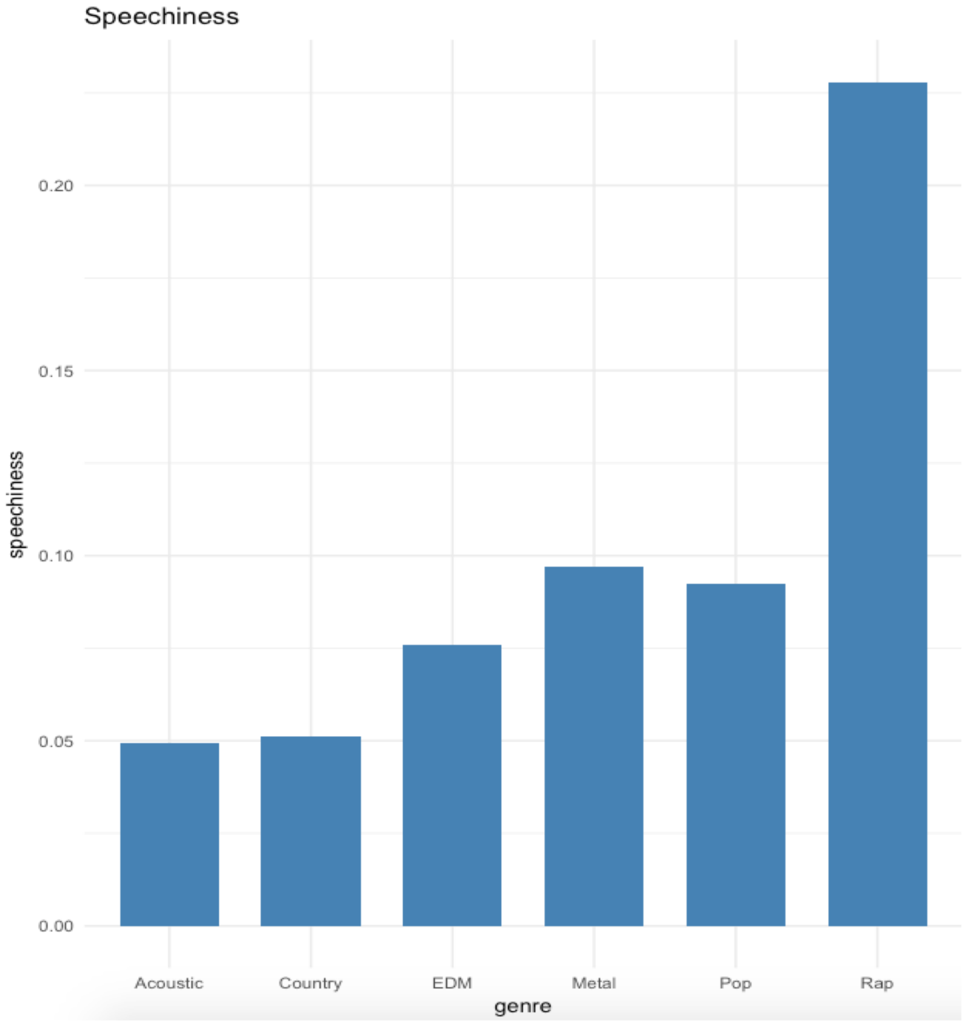
Appetizer for Rock Enthusiasts!

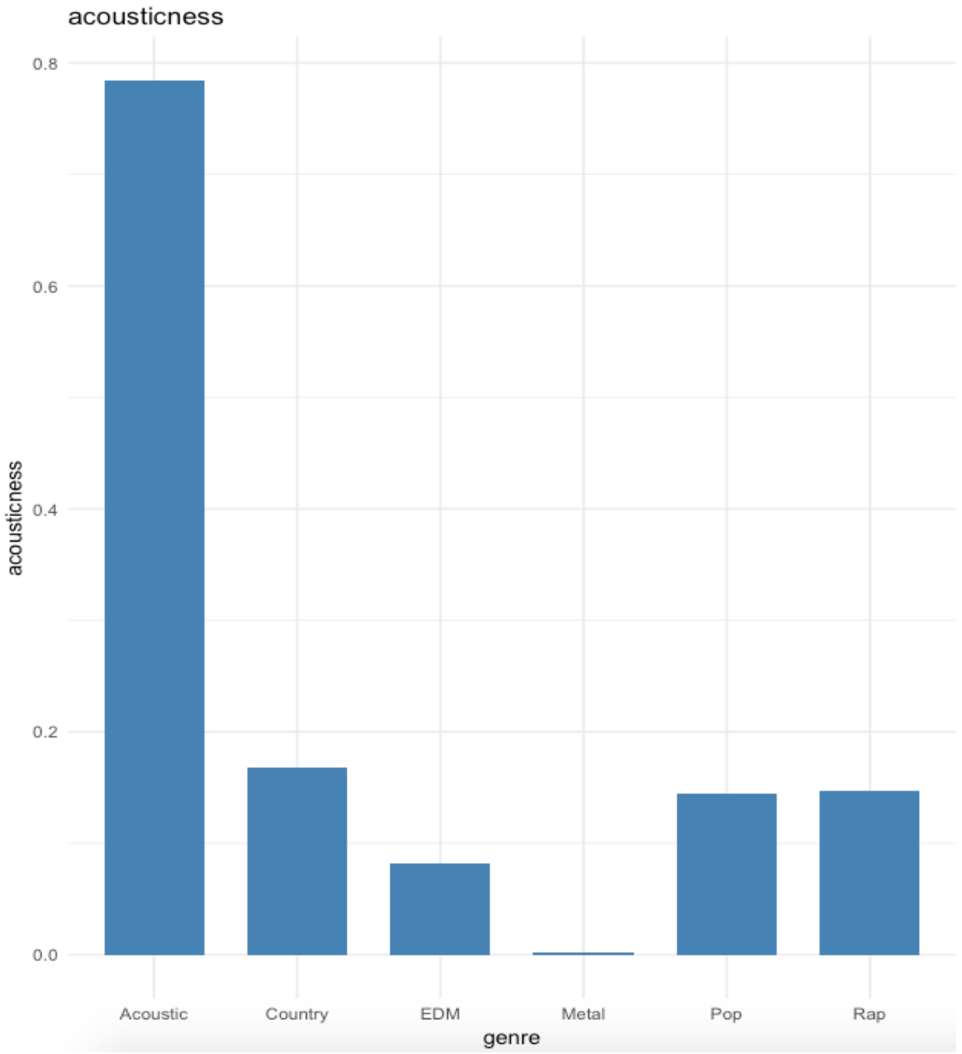
Name	Loudness	Energy	Acousticness	Valence
Knocking on Heaven's Door (1973)	-13.061	0.396	0.251	0.229
Salute Your Solution (2008)	-2.855	0.946	0.00781	0.498
Original Prankster (2000)	-4.149	0.886	0.00061	0.942
Sultans of Swing (1978)	-7.325	0.868	0.0632	0.923

Data distribution by genres

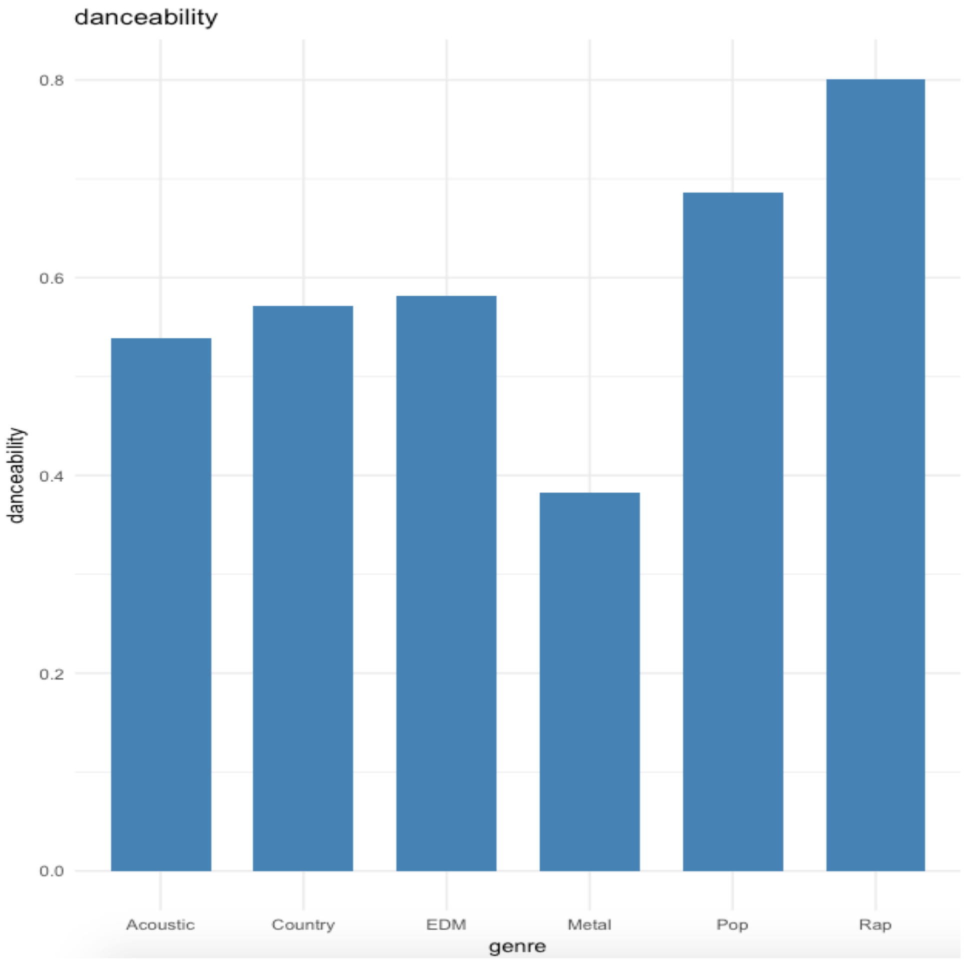


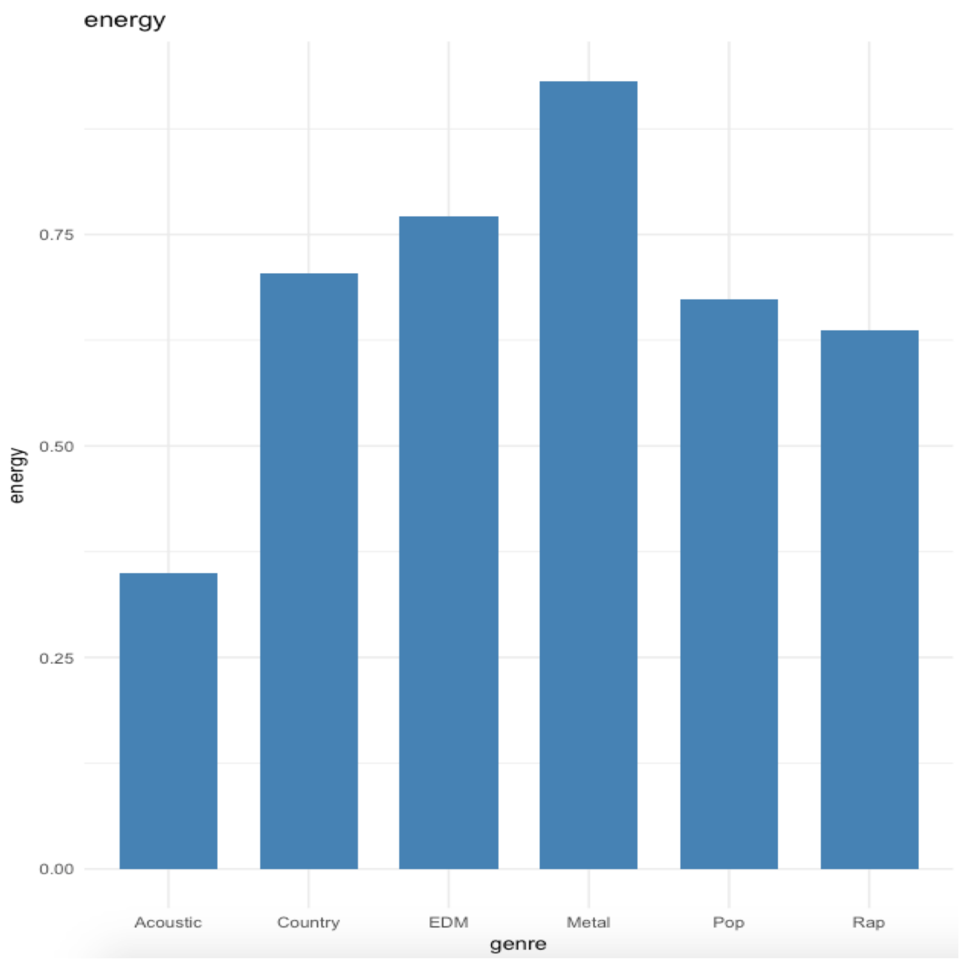
Feature Characteristics by Genre



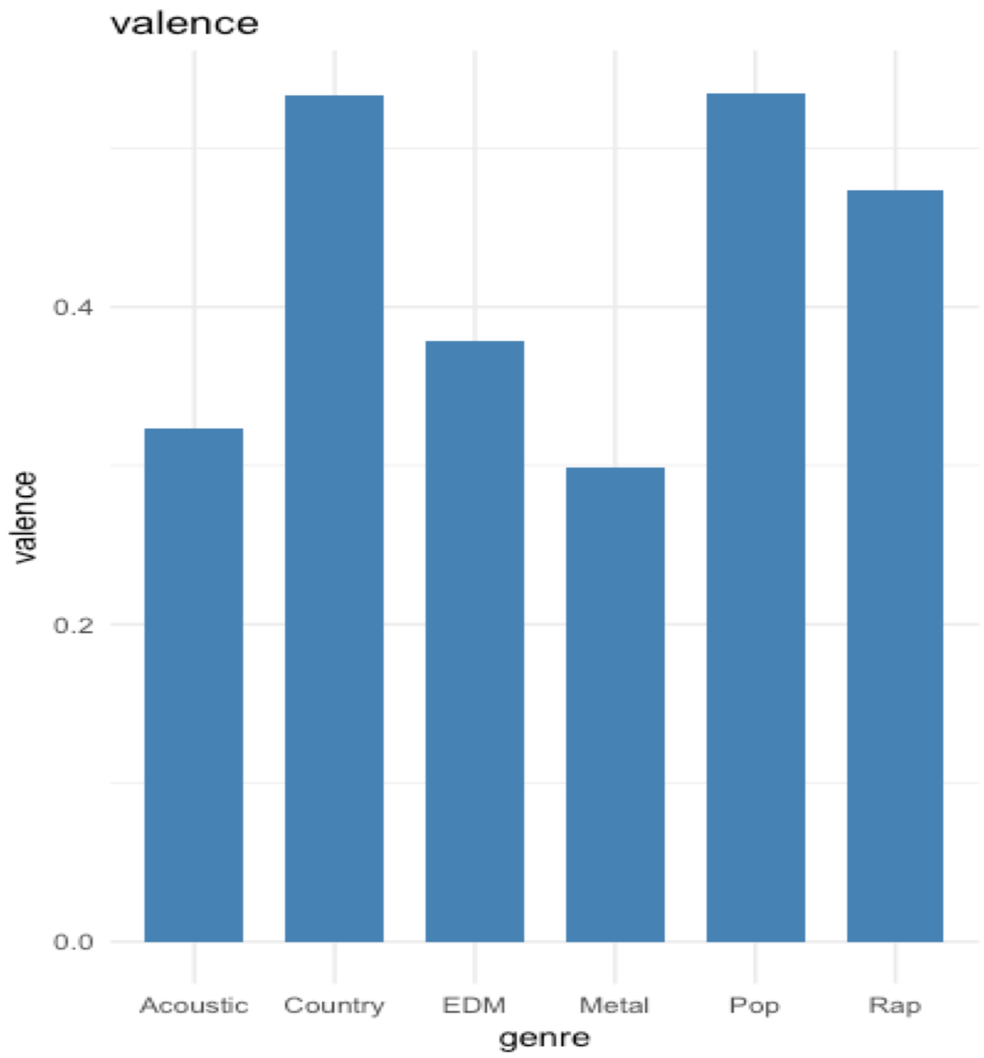


Feature Characteristics by Genre





Feature Characteristics by Genre



Data Pre-Processing

- Feature Selection
  - Acousticness
  - Danceability
  - Energy
  - Instrumentalness
  - Liveness
  - Speechiness
  - Loudness
  - Tempo
  - Valence

Data Pre-Processing

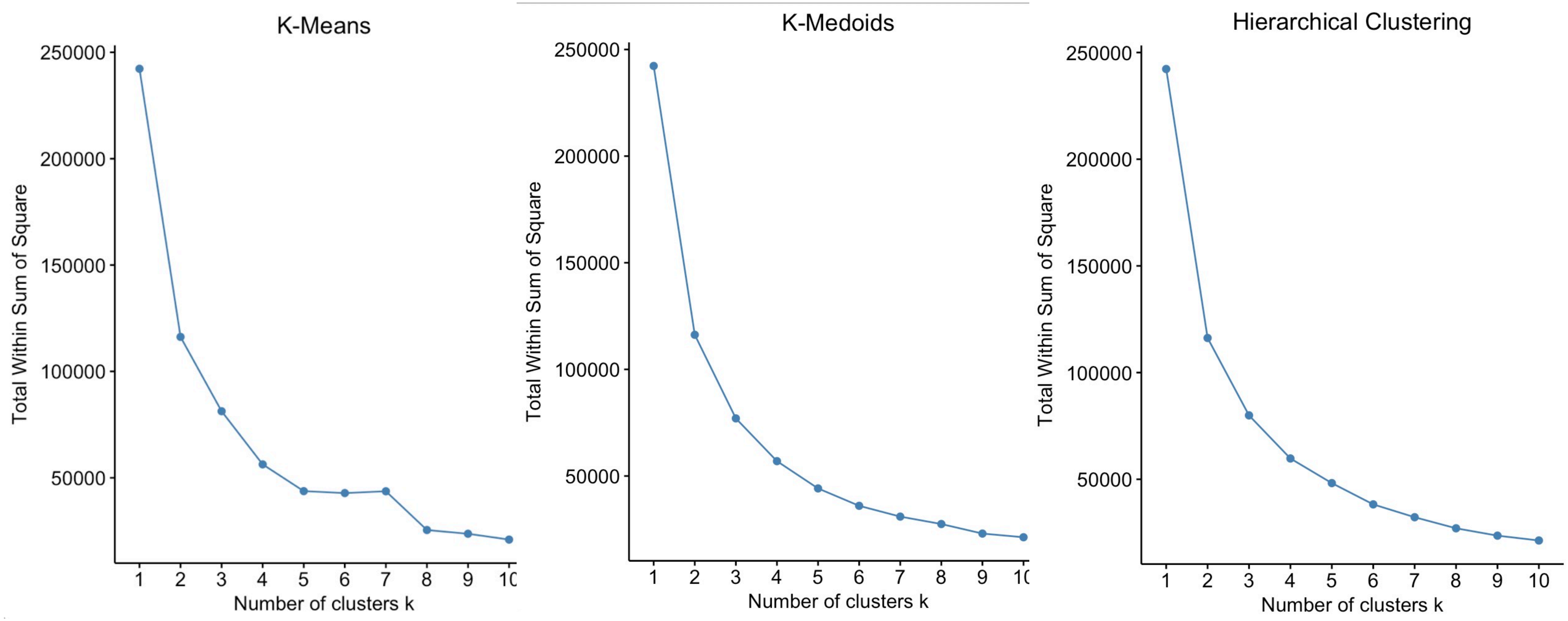
- T-Distributed Stochastic Neighboring Entities (t-SNE)
  - Dimension reduction technique.
  - A probabilistic approach.
  - Uses T-Distribution.
  - Retains the structure in lower dimensions.

Clustering

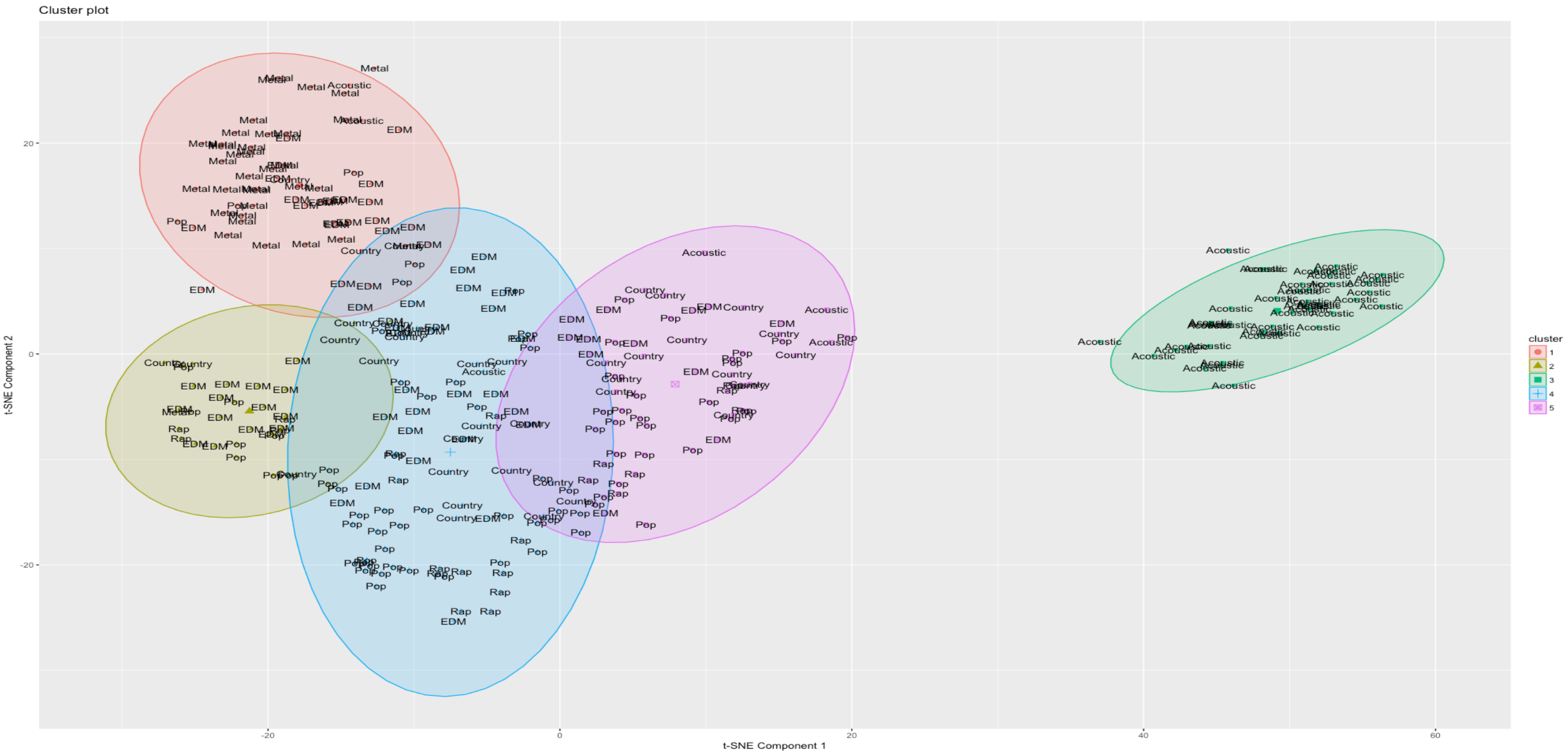
Clustering algorithms used:

- Hierarchical Clustering
- K Means Clustering
- K Medoids Clustering

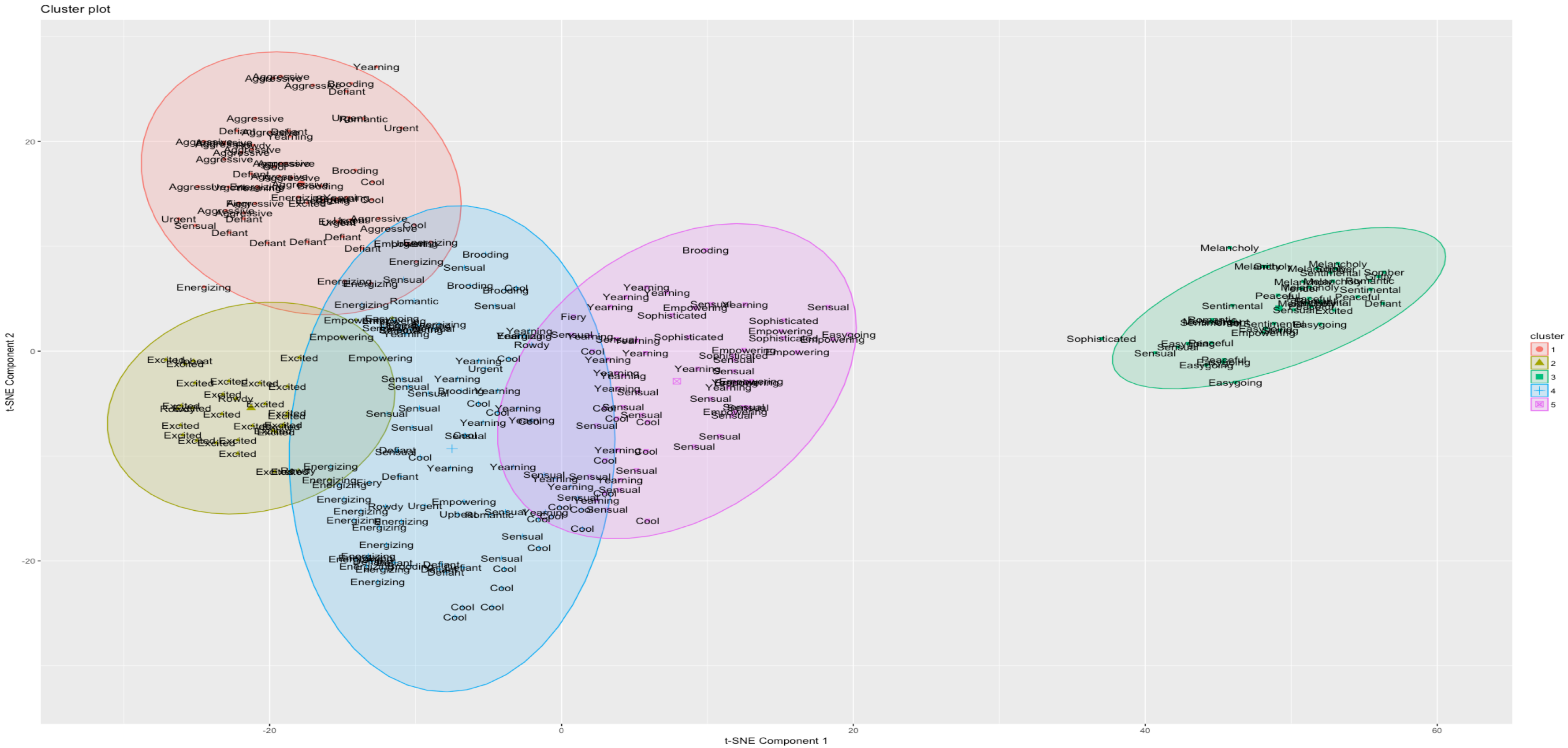
Finding best number of cluster



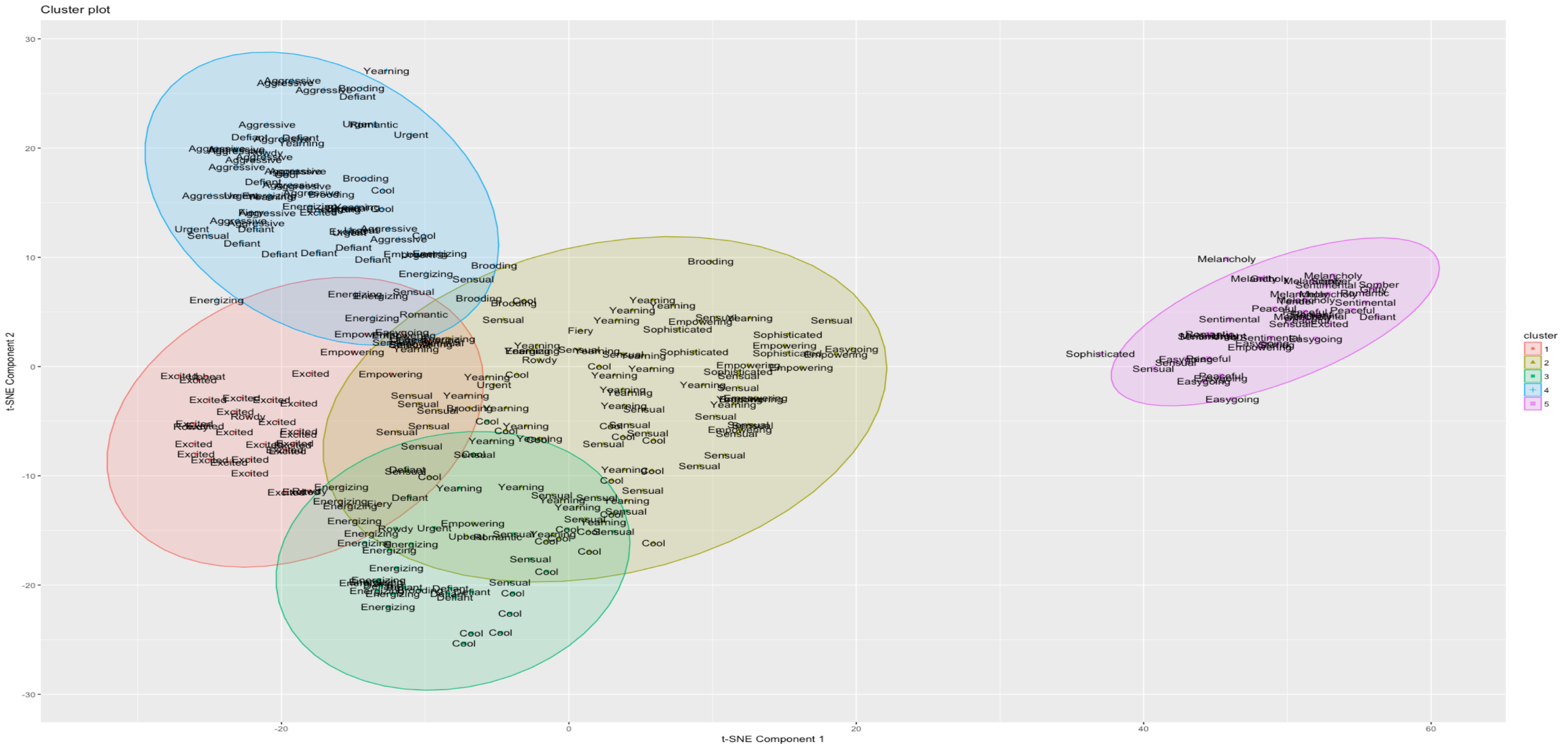
K-Means Clustering (By Genre)



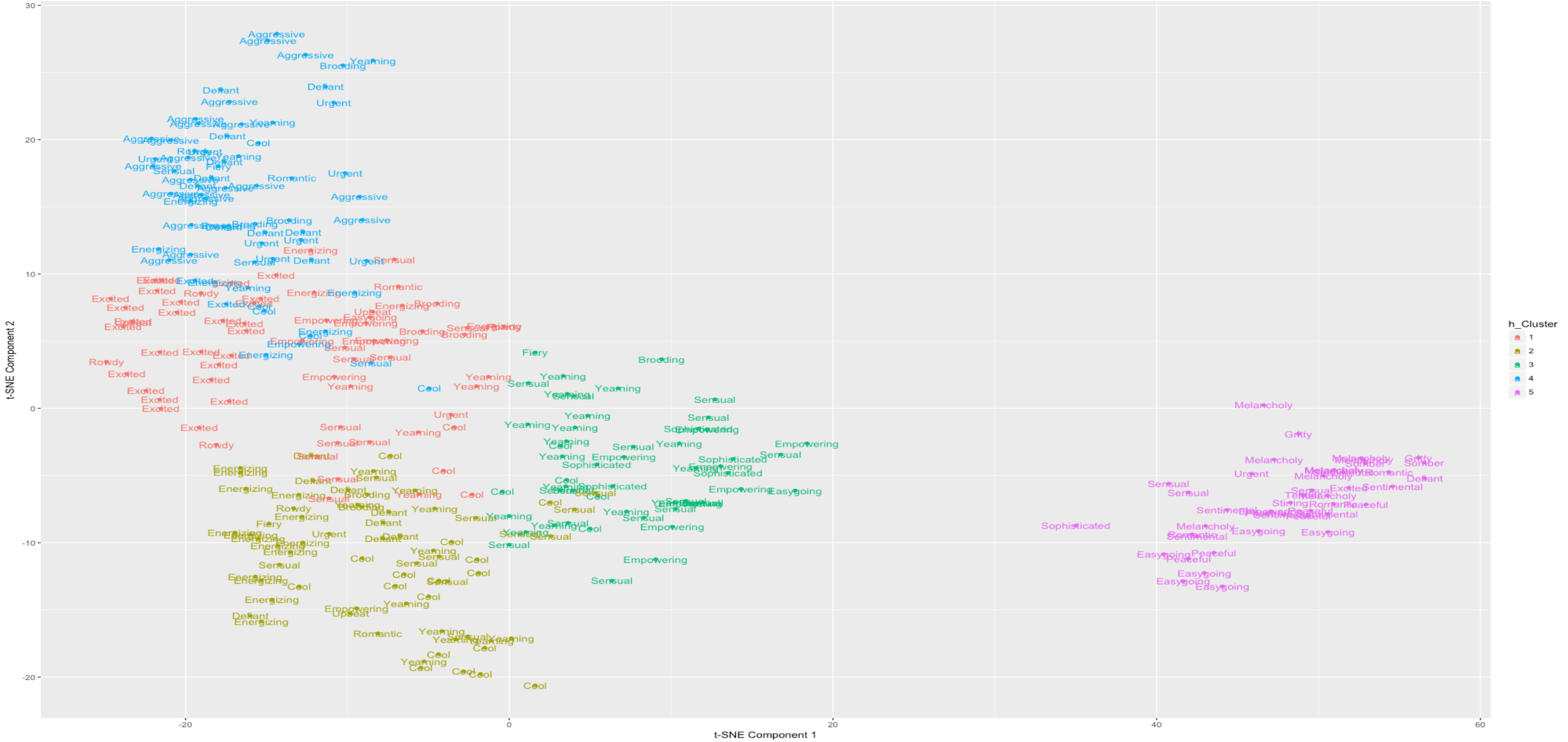
K-Means Clustering (By Mood)



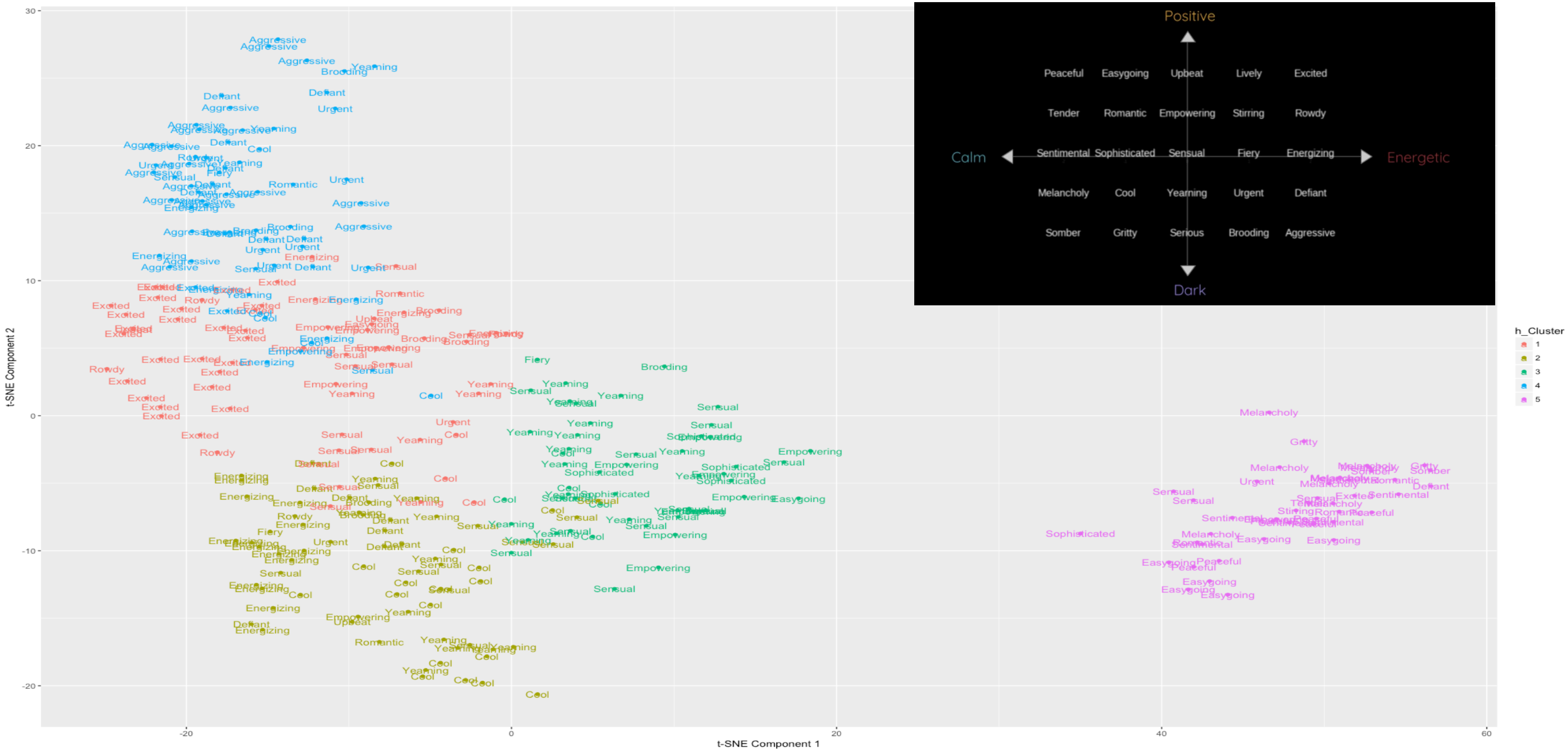
K-Medoids Clustering (By Mood)



Hierarchical Clustering (By Mood)

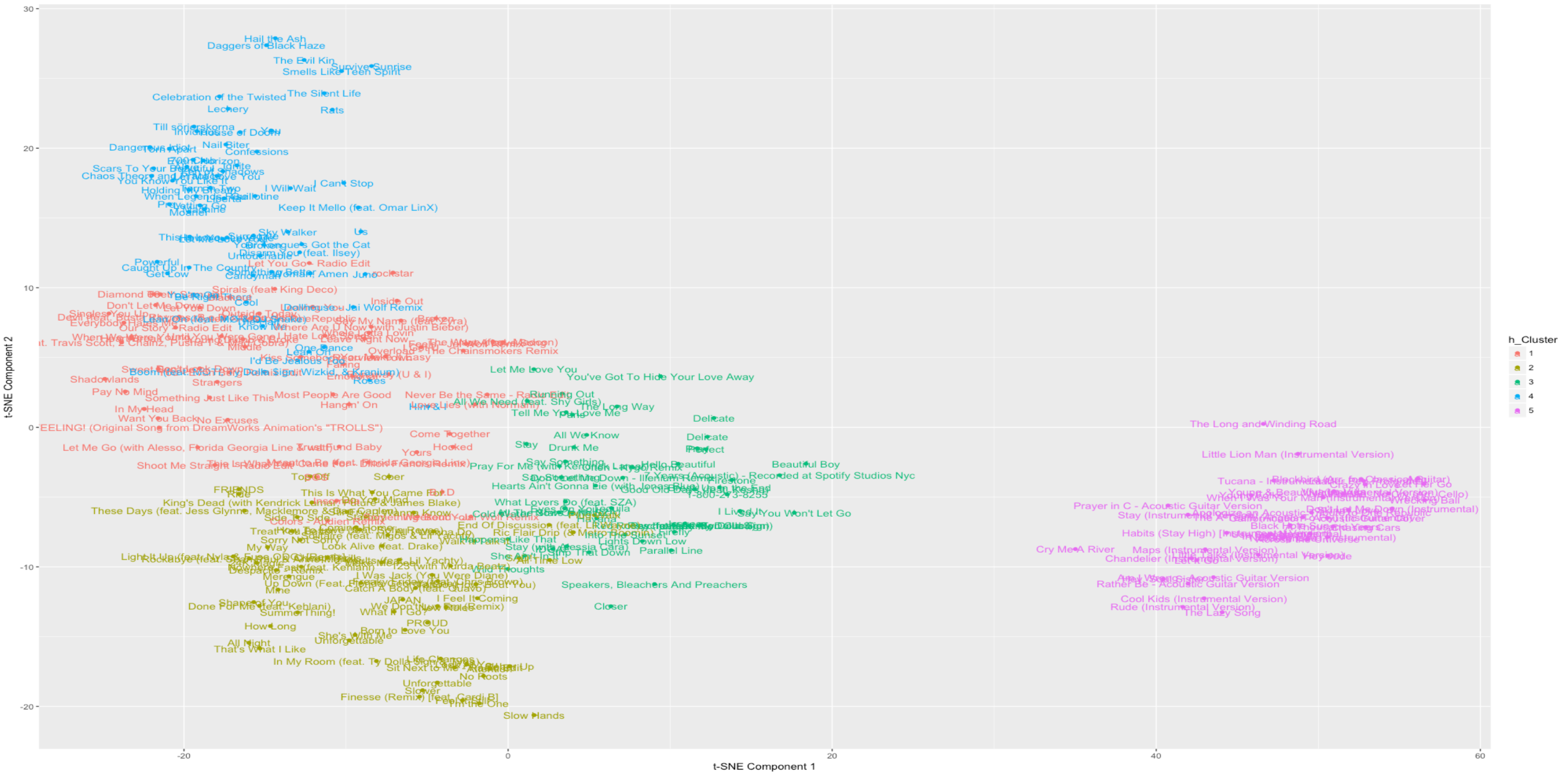


Hierarchical Clustering (By Mood)

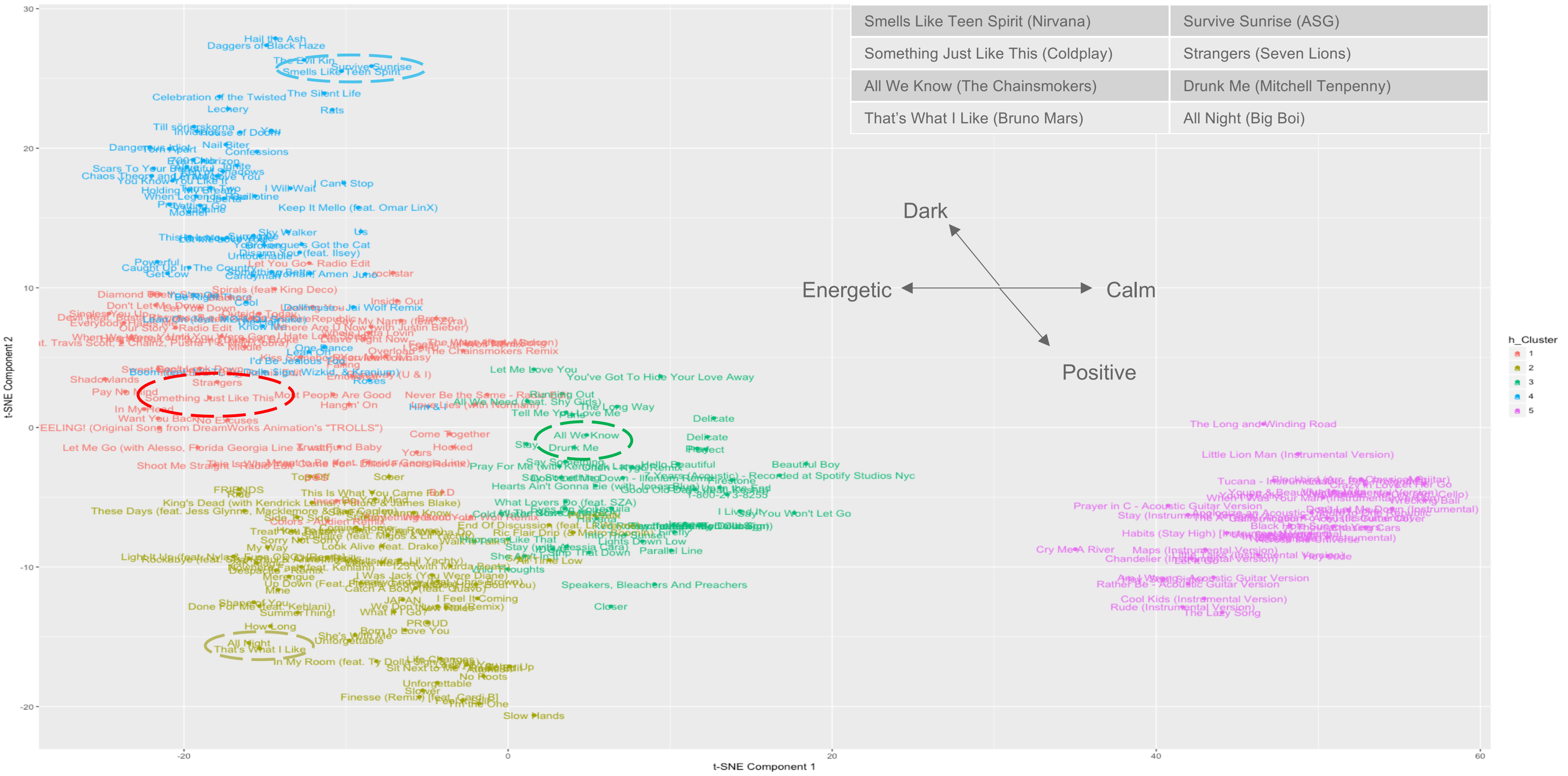


University at Buffalo The State University of New York

Hierarchical Clustering (By track)



Hierarchical Clustering (By track)



Smells Like Teen Spirit (Nirvana)	Survive Sunrise (ASG)
Something Just Like This (Coldplay)	Strangers (Seven Lions)
All We Know (The Chainsmokers)	Drunk Me (Mitchell Tenpenny)
That's What I Like (Bruno Mars)	All Night (Big Boi)

Applications and Future Scope

- Grid Based Music Interface for playing songs which sound similar.
- Integration with Spotify and other Music Applications.
- Incorporating additional features to scale up the model for efficient clustering of songs.
- Labeling the cluster based on characteristics.
- Creating a Similarity Index between two songs : Quantitative representation of how close a song is to another song based on these features. Can modify the radius of Similarity.
- Using Lyrics in addition to audio features for better recommendations! Natural Language Processing for lyrics text

AUDIO FEATURES + LYRICS CONTENT = BETTER RECOMMENDATION


I DONT CARE WHAT SONG YOU ARE
I WILL FIND YOU AND I WILL CLUSTER YOU