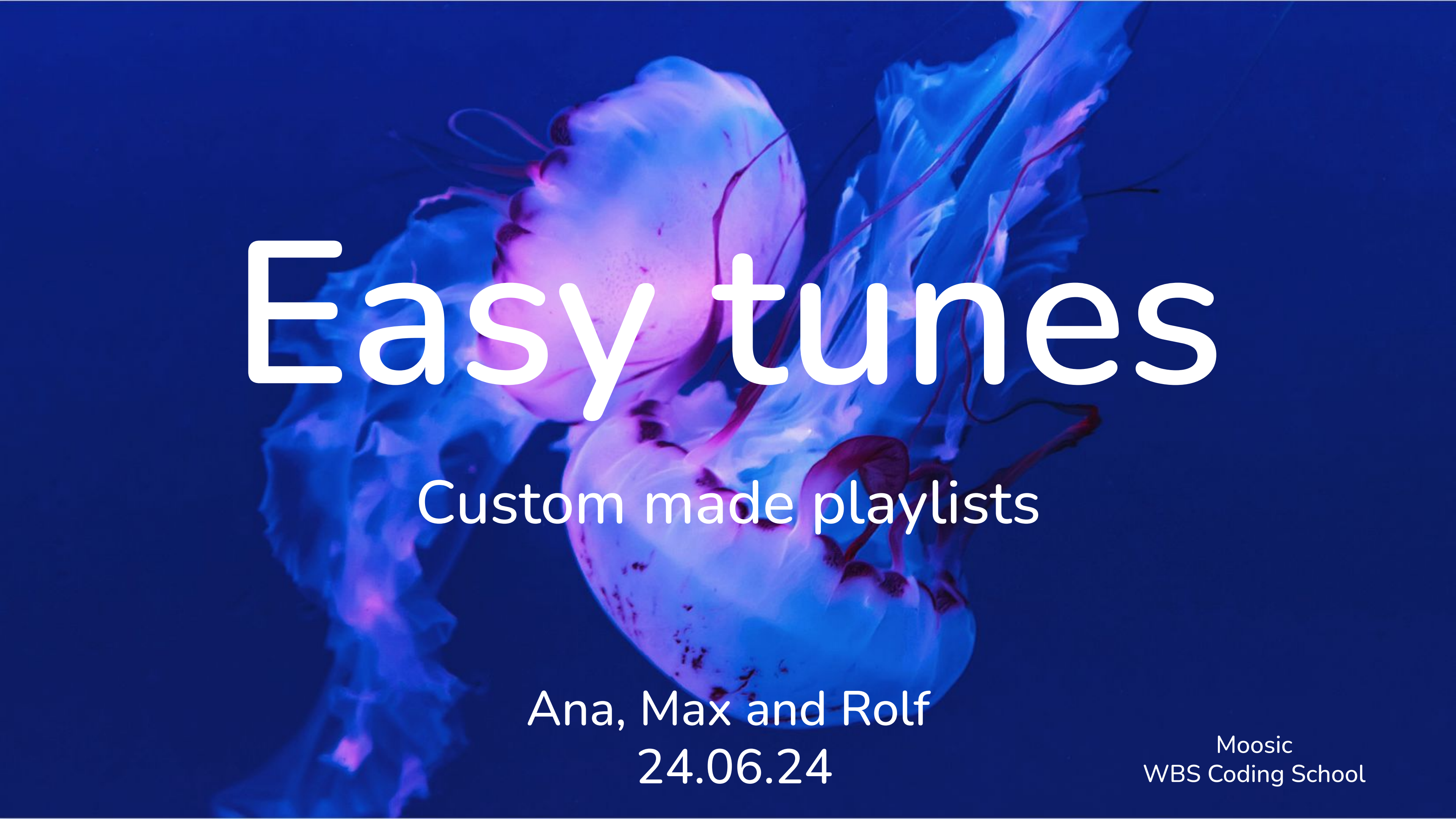


Easy tunes

Custom made playlists

Ana, Max and Rolf

24.06.24

Moosic
WBS Coding School

The Prototype

Dataset

Spotify list of 5000 songs with 13 features

Features

- danceability
- energy
- key
- loudness
- mode
- speechiness
- acousticness
- instrumentalness
- liveness
- valence
- tempo
- duration_ms
- time_signature

Workflow

- Data Cleaning
- Dropping features (categorical)
- Scaling features to value range [0, 1]
- Reduce dimensionality from #13 to #5 "features" with PCA
- Determined number of Playlists: 30-50
- Created Playlists with K-Means Clustering

Swinging with you

Cohesive playlists can be created

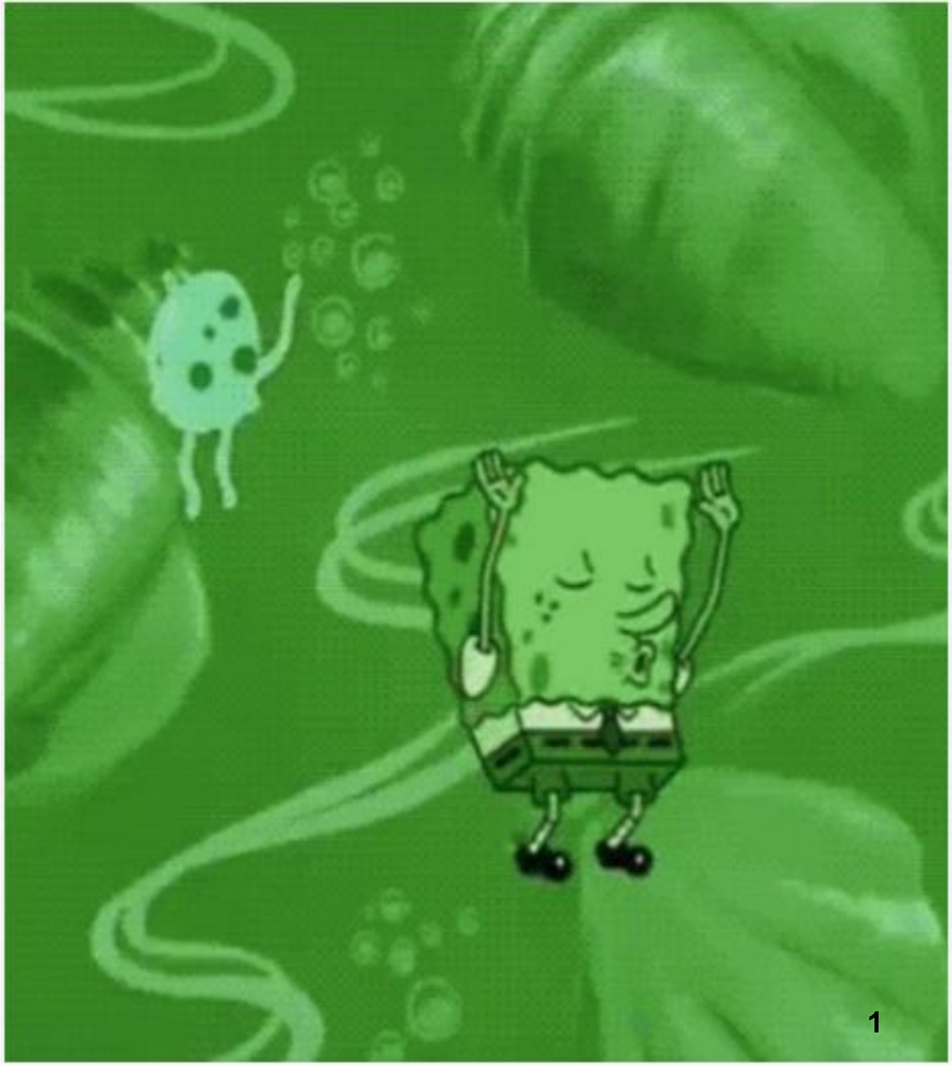
Check out our jams!

Heavy Metal, K = 30

Name	Artist
Raining Blood	Slayer
L'enfant sauvage	Gojira
Pull the Plug	Death
Chapel of Ghouls	Morbid Angel
Dawn of Eternity	Massacre
Hypochristianity	Altar
Psycho Damn	Altar
Black Embrace	Amorphis
Of Pure Unholiness - Studio	Vital Remains
The Way of All Flesh	Morpheus Descends

Playlist for a coffe, K = 30

Mulher Eu Sei	Chico César
Chega De Saudade	João Gilberto
Rebel Rebel	Seu Jorge
Brigas, Nunca Mais	João Gilberto
Kaipuusamba	Maria Gasolina
Water Music, Suite No. 2, HWV 349: Water Music, Suite No. 2, HWV 349: XII. [Alla Hornpipe]	George Frideric Handel
My Baby Just Cares for Me - 2013 Remastered Version	Nina Simone
Splanky	Count Basie
Freddie Freeloader (feat. John Coltrane, Cannonball Adderley, Wynton Kelly & Paul Chambers)	Miles Davis
Bags' Groove - Rudy Van Gelder 2001 - Remaster	Milt Jackson



Sorry about that!

Failed Playlist, K = 30

Name	Artist
Criminal	Natti Natasha
Otra vez (feat. J Balvin)	Zion & Lennox
Bonita	J Balvin
Felices los 4	Maluma
He Did It - Live	The Brown Boyz
Aquele Abraço	Gilberto Gil
Come Saturday	The Pains Of Being Pure At Heart
The Day That Thatcher Dies	Hefner
Falling Out Of Love (With You)	The 6ths
Summer Teeth	Wilco

The numbers in music and feelings

Data collection from music

- Perceptive data works as a bridge between factual information and feelings.

- Additional data like producer, songwriter, (associated) genre, language, lyrics and user interactions would improve the results.

K-Means

Pro

- Easy to Use
- Efficient
- Able to handle large datasets
- Easy to understand through centroids
- Flexible

Con

- Strict assignments to clusters
- Choosing the right number of clusters

Improvements

- Trying K-Means++, the improved version of the K-Means
- Trying the Gaussian Mixture Models (GMM) for softer clustering
- Using web scraping to find out the genre and the language of the songs
- Hierarchical Clustering
- Doing a second round of clustering for better results

Next Steps

- After deciding on the playlists we can implement a system for updating the playlist
- Possibly integrating with Spotify's API to create and manage the dataset on the platform


Thank you.
Any questions?

Sources

[1] https://soundcloud.com/sarveproductions/spongebob-sqaurepants-jellyfish-jam-stadium-rave-a-recreation

[2] https://www.reddit.com/r/spongebob/comments/12hbrf5/why_squidward_cant_play_the_clarinet/

[3] https://en.wikipedia.org/wiki/Hierarchical_clustering

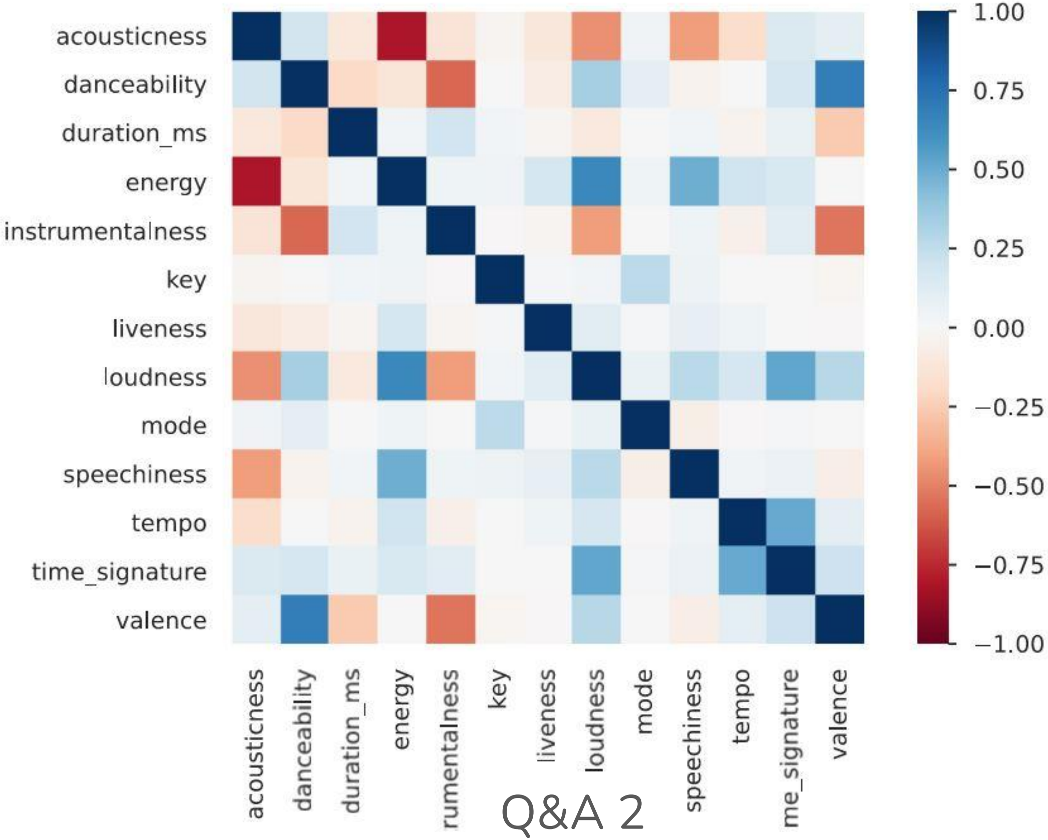
[4] https://scikit-learn.org/stable/modules/clustering.html#hierarchical-clustering

[5]https://learn.wbscodingschool.com/courses/data-science/lessons/%f0%9f%8f%97-2-k-means-clustering/topic/%f0%9f%93%9a-understanding-audio-features/

Data Cleaning

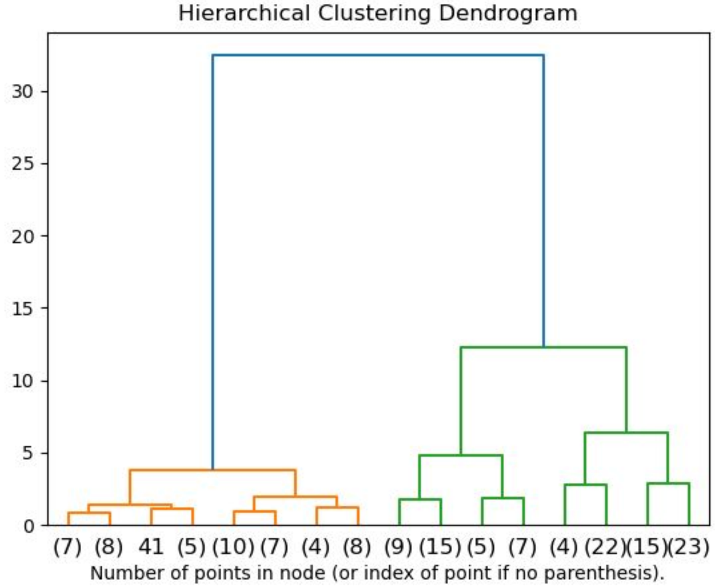
- Column names had weird blank spaces
- Some Columns had no information -> dropped
- Html links were sometimes incorrect
- Html links were checked with track-ID
- track-IDs & html-links were checked for duplicates
- Wrong song titles and artist names (numbers, times, symbols-,#,+,?) -> dropped

Correlation between features - use PCA to reduce dimensionality



Hierarchical Clustering - possible alternative to K-Means

Pairwise grouping of datapoints, bottom-up algorithm.
References [3], [4]



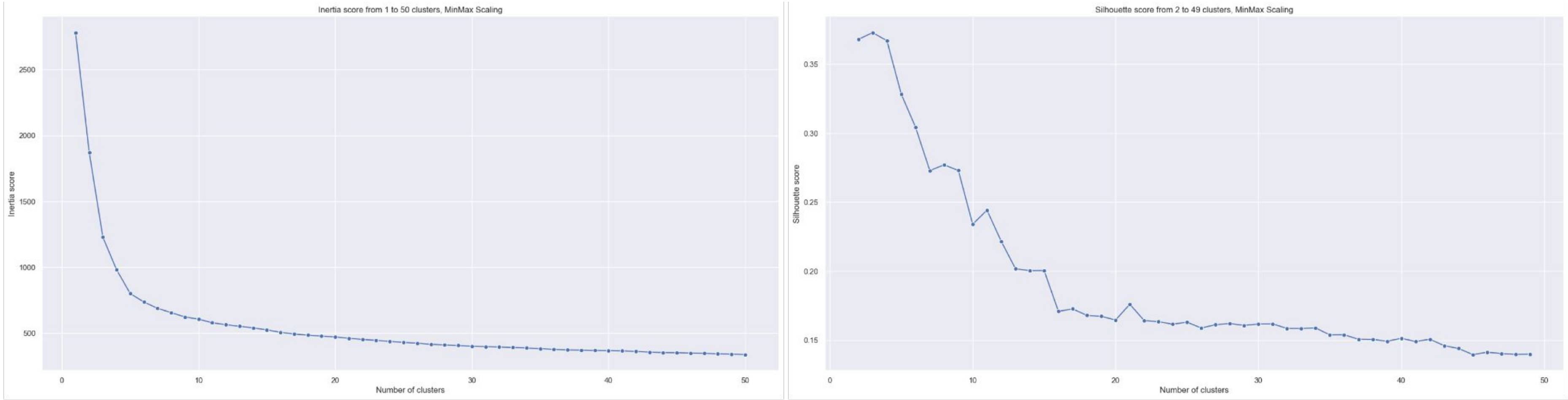


Figure 1: Inertia and Silhouette Scores for MinMax Scaling. Data used: 'danceability', 'energy', 'loudness', 'speechiness', ‘acousticness', 'instrumentalness', 'liveness', 'valence', 'tempo', ‘duration_ms’ from around 5000 songs

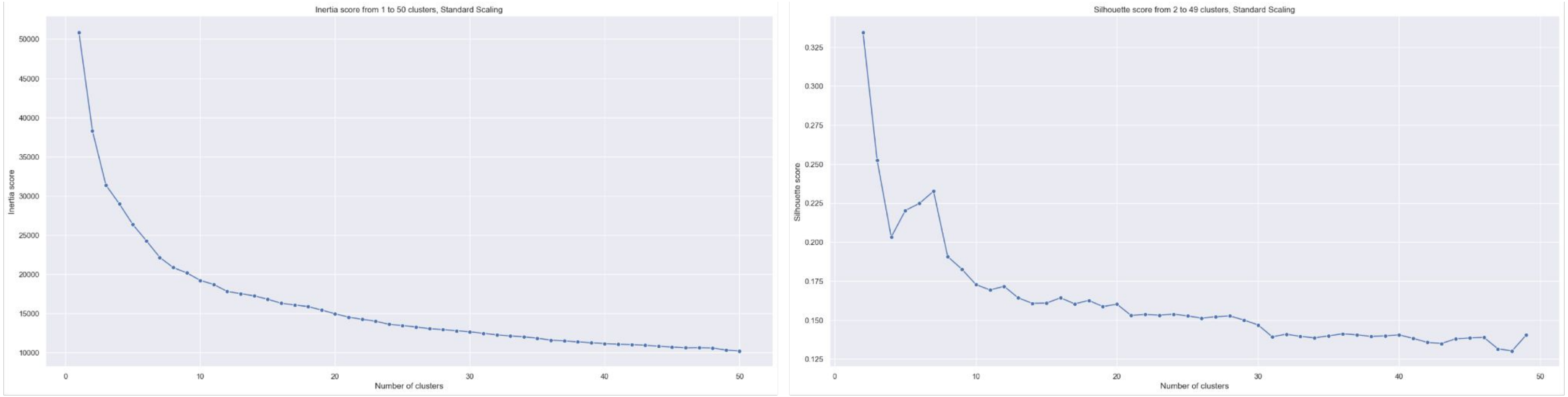


Figure 2: Inertia and Silhouette Scores for Standard Scaling. Data used: 'danceability', 'energy', 'loudness', 'speechiness', ‘acousticness', 'instrumentalness', 'liveness', 'valence', 'tempo', ‘duration_ms’ from around 5000 songs

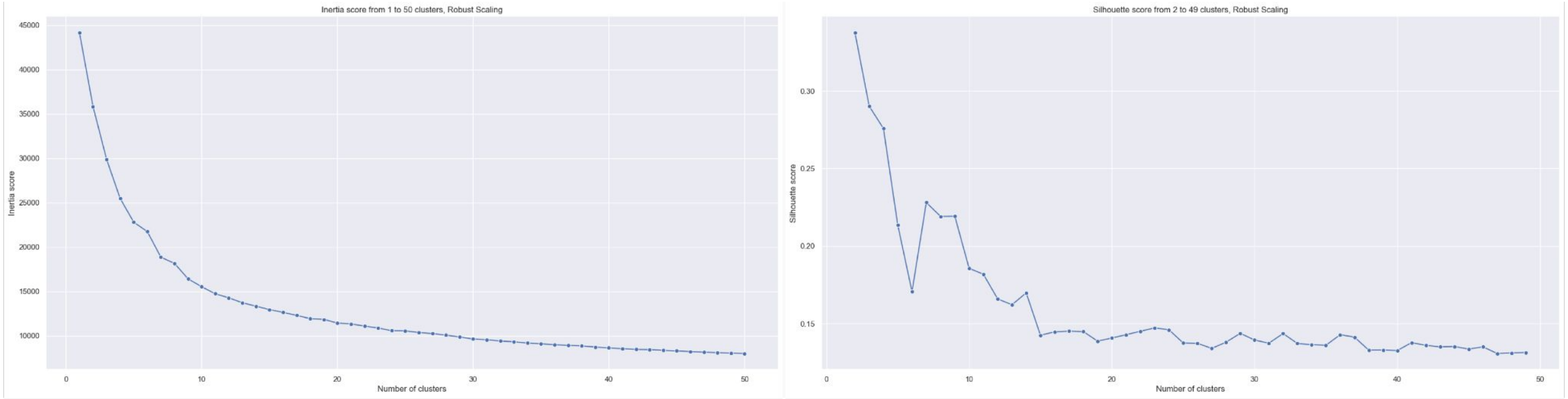


Figure 3: Inertia and Silhouette Scores for Robust Scaling. Data used: 'danceability', 'energy', 'loudness', 'speechiness', ‘acousticness', 'instrumentalness', 'liveness', 'valence', 'tempo', ‘duration_ms’ from around 5000 songs

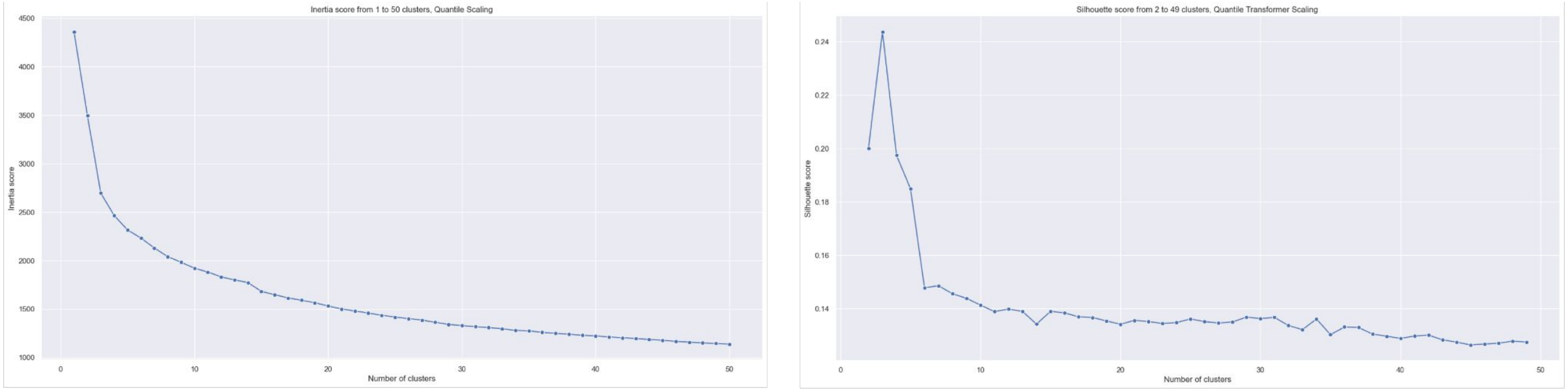


Figure 4: Inertia and Silhouette Scores for Quantile Transformer Scaling. Data used: 'danceability', 'energy', 'loudness', 'speechiness', ‘acousticness', 'instrumentalness', 'liveness', 'valence', 'tempo', ‘duration_ms’ from around 5000 songs

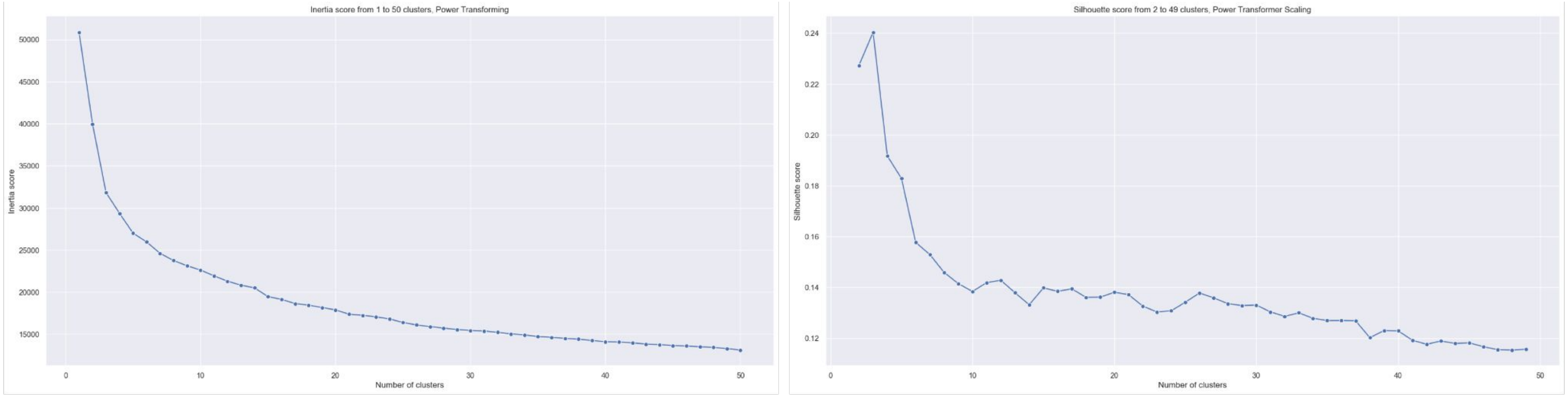


Figure 5: Inertia and Silhouette Scores for Power Transformer Scaling. Data used: 'danceability', 'energy', 'loudness', 'speechiness', ‘acousticness', 'instrumentalness', 'liveness', 'valence', 'tempo', ‘duration_ms’ from around 5000 songs

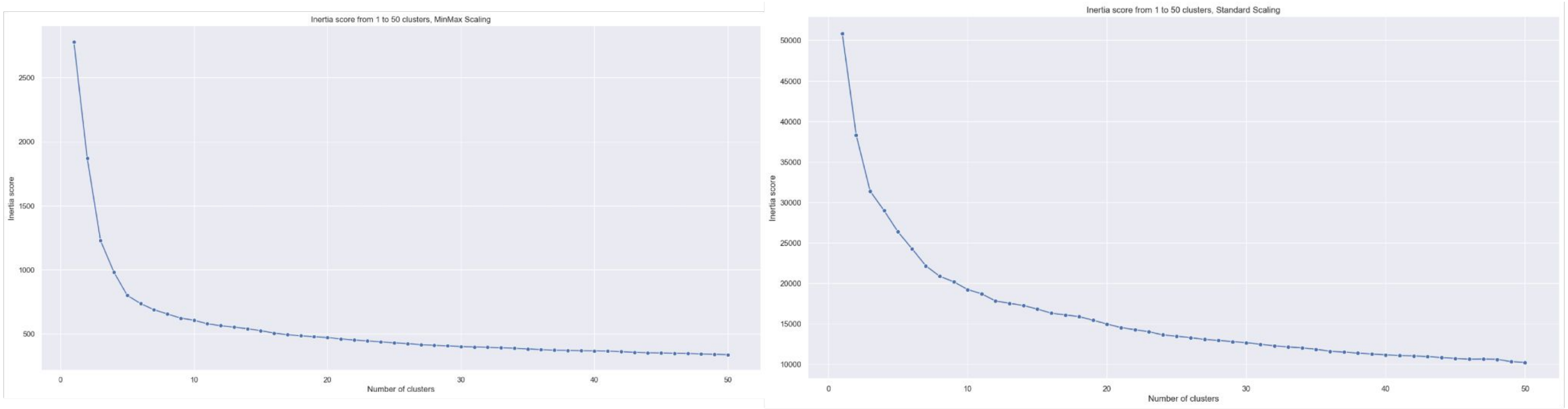
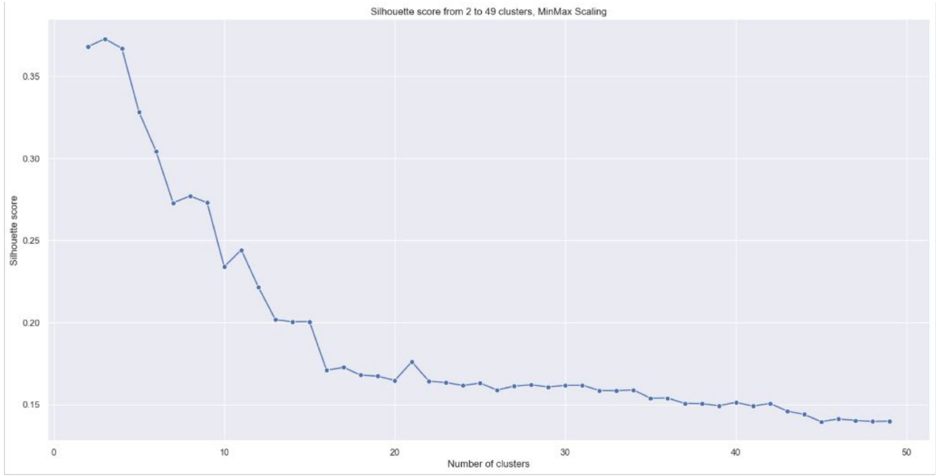


Figure 6: Comparison of Inertia between MinMax Scaling and Standard Scaling. Lower values in the inertia score tells us that the values are in a closer distance to the centroid. MinMax Scaling shows an overall tighter group compared to Standard Scaling and is therefore better suited for scaling.



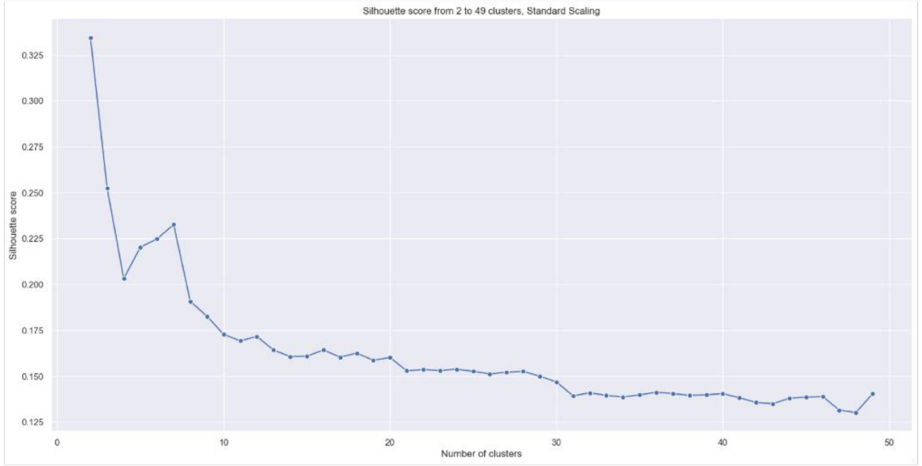


Figure 7: Comparison of Silhouette Scores between MinMax Scaling and Standard Scaling. Higher values in the silhouette score tells us that the groups are homogene and the overall inhomogeneity between different groups bigger. MinMax Scaling shows higher values compared to Standard Scaling. We can interpret that the songs in a playlist have many similarities, i.e. are tighter at the centroid. Songs from other playlists are different to foreign playlists, i.e. are further from foreign centroids.

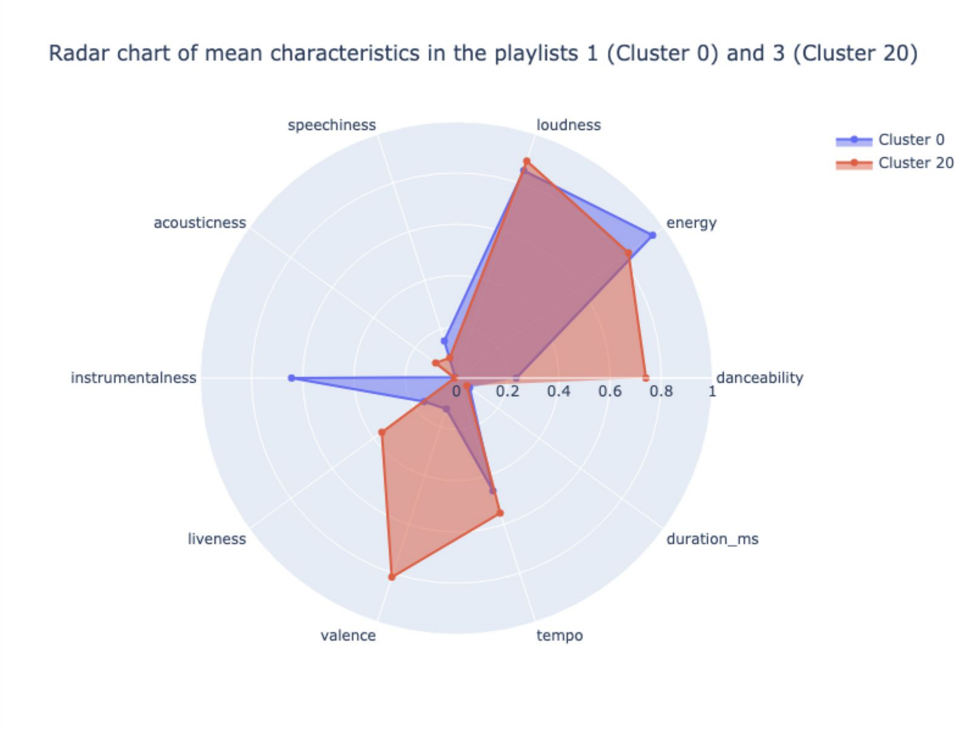


Figure 8: Radar chart from characteristic scores in playlist 1 and playlist 3. Instrumentalness, valence and danceability are the main components distinguishing the successful playlist 1 and the unsuccessful playlist 3, k = 30.

Q&A 11

`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.
`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.
`duration_ms`
The duration of the track in milliseconds.
`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.
`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.
`key`
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.

Information from [5]

`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 strong likelihood that the track is live.
`loudness`
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 typically range between -60 and 0 db.
`mode`
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`
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`
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`
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`
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).

Information from [5]