

April | 2022



SPOTIFY RECOMMENDATION SYSTEM

with Machine Learning



BDS 516 | Group 3: Mangara Paul Alexander Hutagalung, Shruthi Ramakrishnan, Seiya Kato, Sabrina Morawej, and Michael Adjei-Poku

OBJECTIVE

Methods Spotify currently uses to recommend music



User Modeling

Spotify tethers users' moods and time of day (the algorithm will recommend different music on a Monday night vs. Friday night)



Approximate Nearest-Neighbor Search Algorithm

Spotify also uses an approximate nearest-neighbor SEARCH algorithm to group songs and users together to recommend Discover Weekly playlists



"Enhance" feature

They have a feature called "enhance" which recommends songs within the playlist based on the similarity to the songs in the playlist using clustering analysis



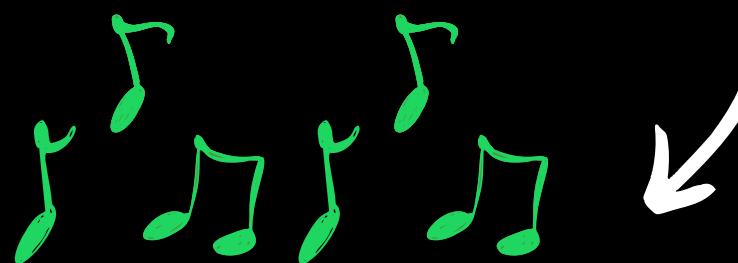
Top Tracks in Your Network

Since 2011, Spotify has been fine-tuning a feature called "Top Tracks in Your Network" that aggregates the top songs of the people you follow and creates a playlist for you.

OBJECTIVE

Filling a Gap

Our Algorithm

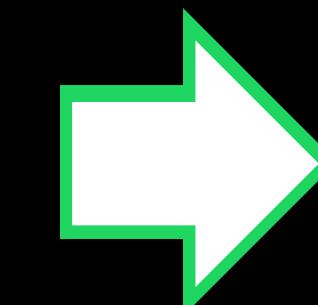


Retention



Personalization

Spotify Free



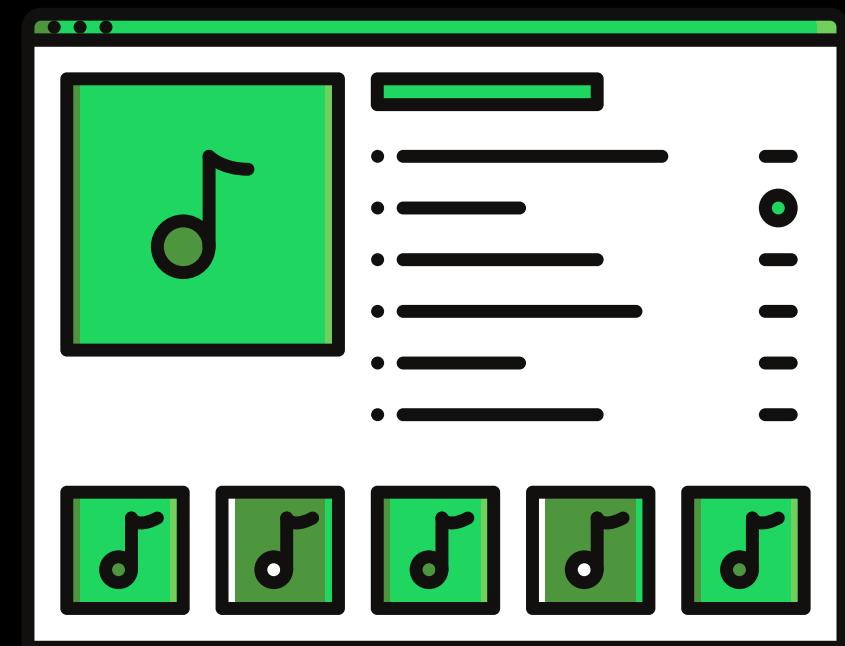
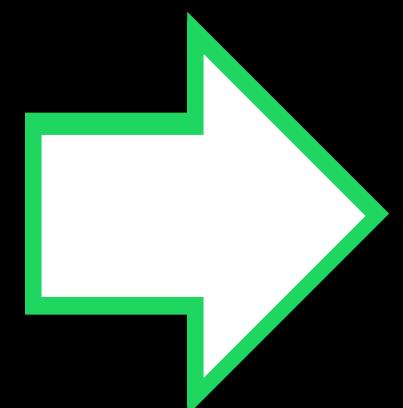
Spotify Premium



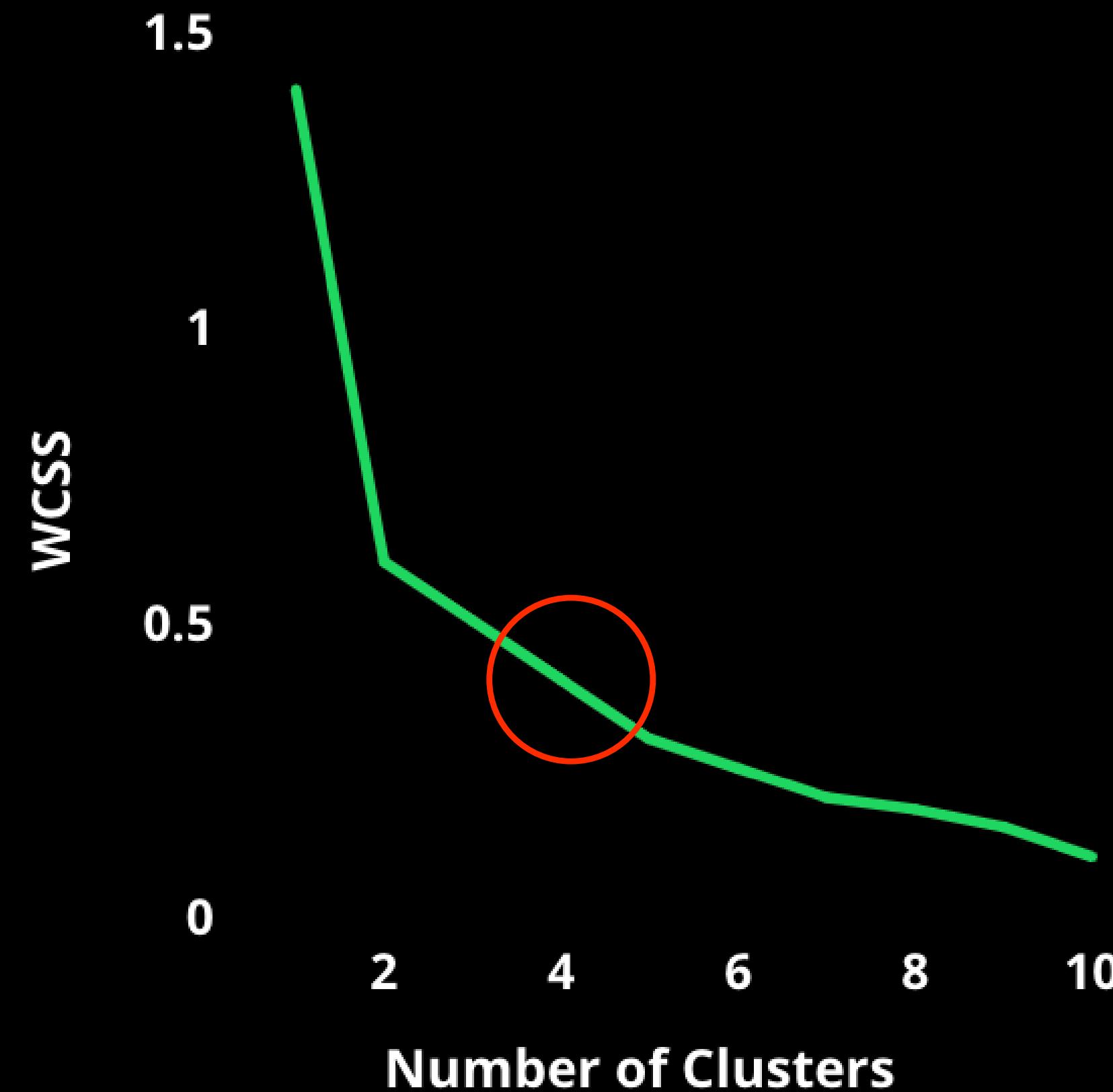
METHODS

- 1 Scraped data from Spotify's API
- 2 Collected playlists from 20 users
- 3 Used data from 16 users whose playlists were public
- 4 Normalized the data

Deliverable



SCREE PLOT



CLUSTERING

Cluster 0

Alyssa	Ravenn
Arvind	Rebecca
Loizos	Sabrina
Mariela	Sita

Cluster 1

Sam

Cluster 2

Alex	Mangara
Elina	Michael
Hanning	Sara
Lennart	

Cluster 3

Emmanuella

CLUSTERING

- Danceability
- Energy
- Loudness
- Speechiness
- Acousticness
- Instrumentalness
- Liveness
- Valence
- Tempo
- Popularity

Cluster 0	
Alyssa	Ravenn
Arvind	Rebecca
Loizos	Sabrina
Mariela	Sita

Cluster 1	
Sam	

Cluster 2	
Alex	Mangara
Elina	Michael
Hanning	Sara
Lennart	

Cluster 3	
Emmanuella	

MUSIC FEATURES



Loudness : Its values are averaged across the entire track. It is the quality of a song. It ranges from -60 to 0 DB. Higher the value, the louder the song.



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.



Energy : It represents a perceptual measure of intensity and activity. Energetic tracks feel fast, loud, and noisy. Higher the value (from 0.0 to 1.0) more energetic the song.



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.

MUSIC FEATURES



Danceability : It describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. The value ranges from 0 to 1. Higher the value more suitable the song is for dancing.



Speechiness : It 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.



Instrumentalness : Predicts whether a track contains no vocals. The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content.



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.

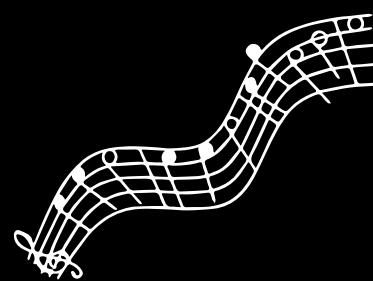
MUSIC FEATURES



Popularity : It is calculated by algorithm and is based, in the most part, on the total number of plays the track has had and how recent those plays are. Generally speaking, songs that are being played a lot now will have a higher popularity than songs that were played a lot in the past.



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).

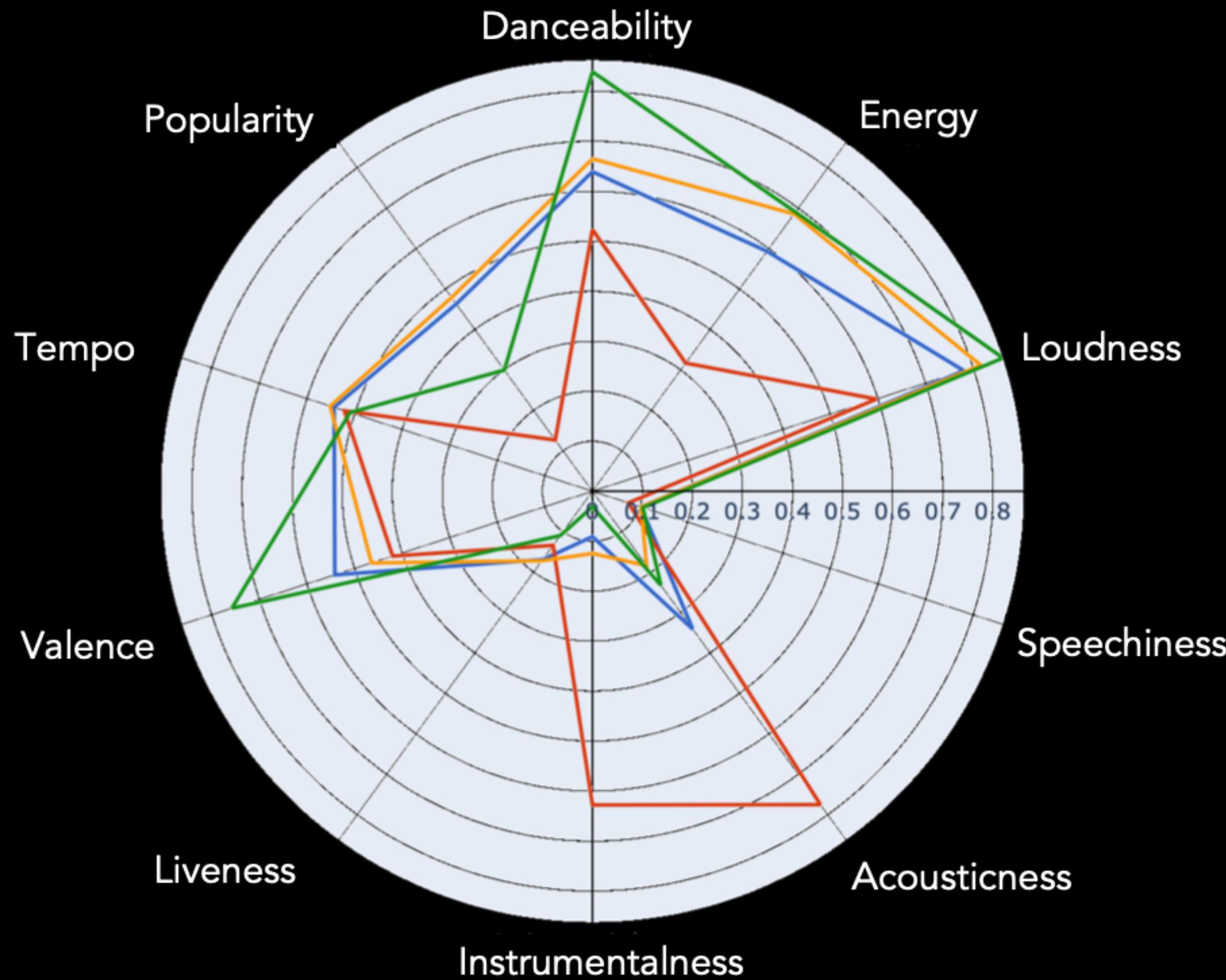


Mode : Songs can be classified as major and minor. 1.0 represents major mode and 0 represents minor.



Key : It is the pitch, notes or scale of song that forms the basis of a song. 12 keys are ranging from 0 to 11.

CLUSTERING



Cluster 0

Alyssa	Ravenn
Arvind	Rebecca
Loizos	Sabrina
Mariela	Sita

Cluster 1

Sam

Cluster 2

Alex	Mangara
Elina	Michael
Hanning	Sara
Lennart	

Cluster 3

Emmanuella

CLUSTERING - OVERVIEW

Cluster 0

Alyssa
Arvind
Loizos
Mariela

Cluster 1

Sam

Cluster 2

Alex
Elina
Hanning
Lennart

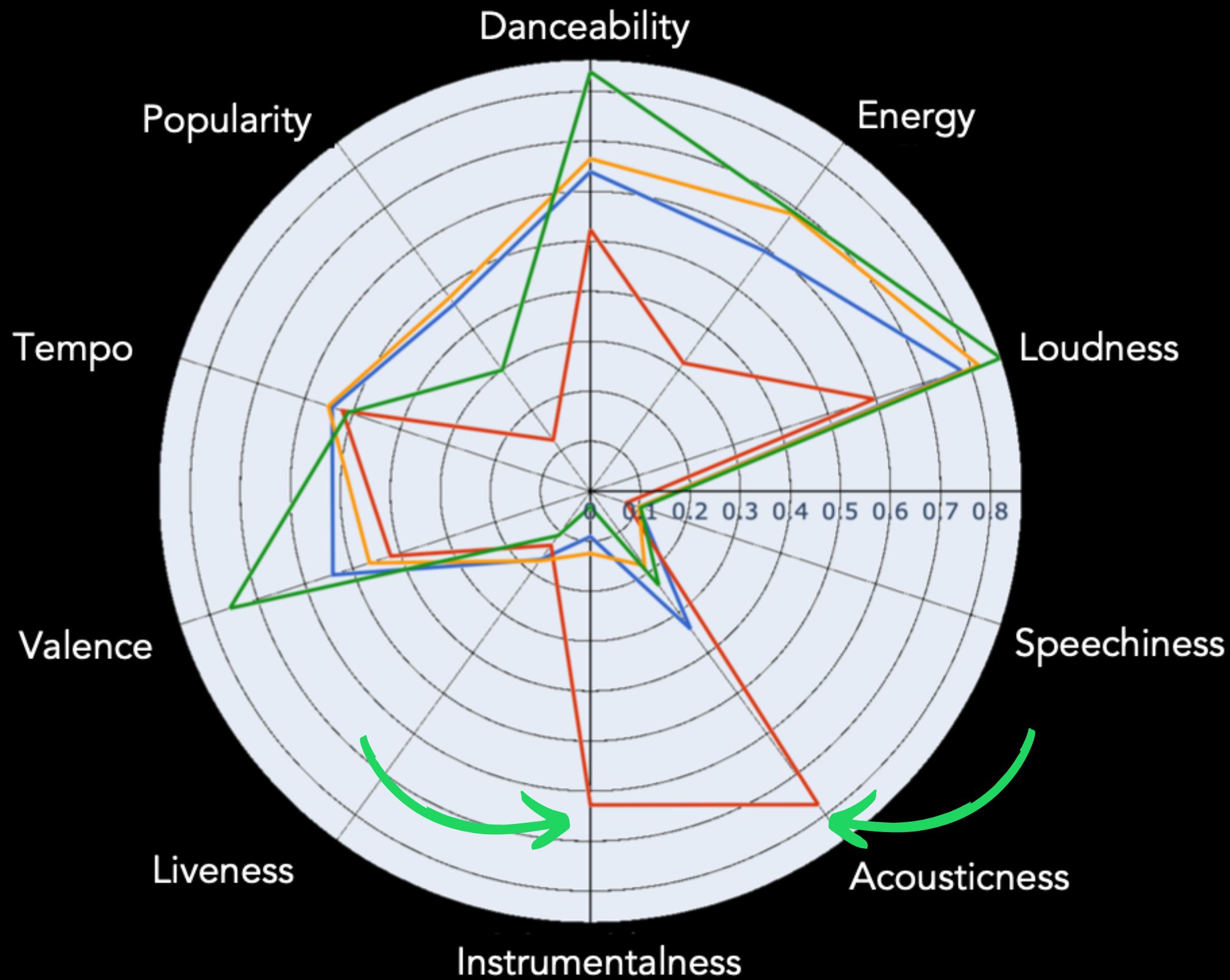
Mangara
Michael
Sara

Cluster 3

Emmanuella

- Cluster 1 distinguishes itself by high acousticness and instrumentalness
- Cluster 3 is characterized by its high score in valence and danceability
- Cluster 0 and Cluster 2 are relatively similar in terms of relatively high danceability and moderate levels of all other music elements
- None of the clusters listened to music that could be categorized as "speechiness"

CLUSTER 1



Cluster 0

Alyssa
Arvind
Loizos
Mariela
Ravenn
Rebecca
Sabrina
Sita

Cluster 1

Sam

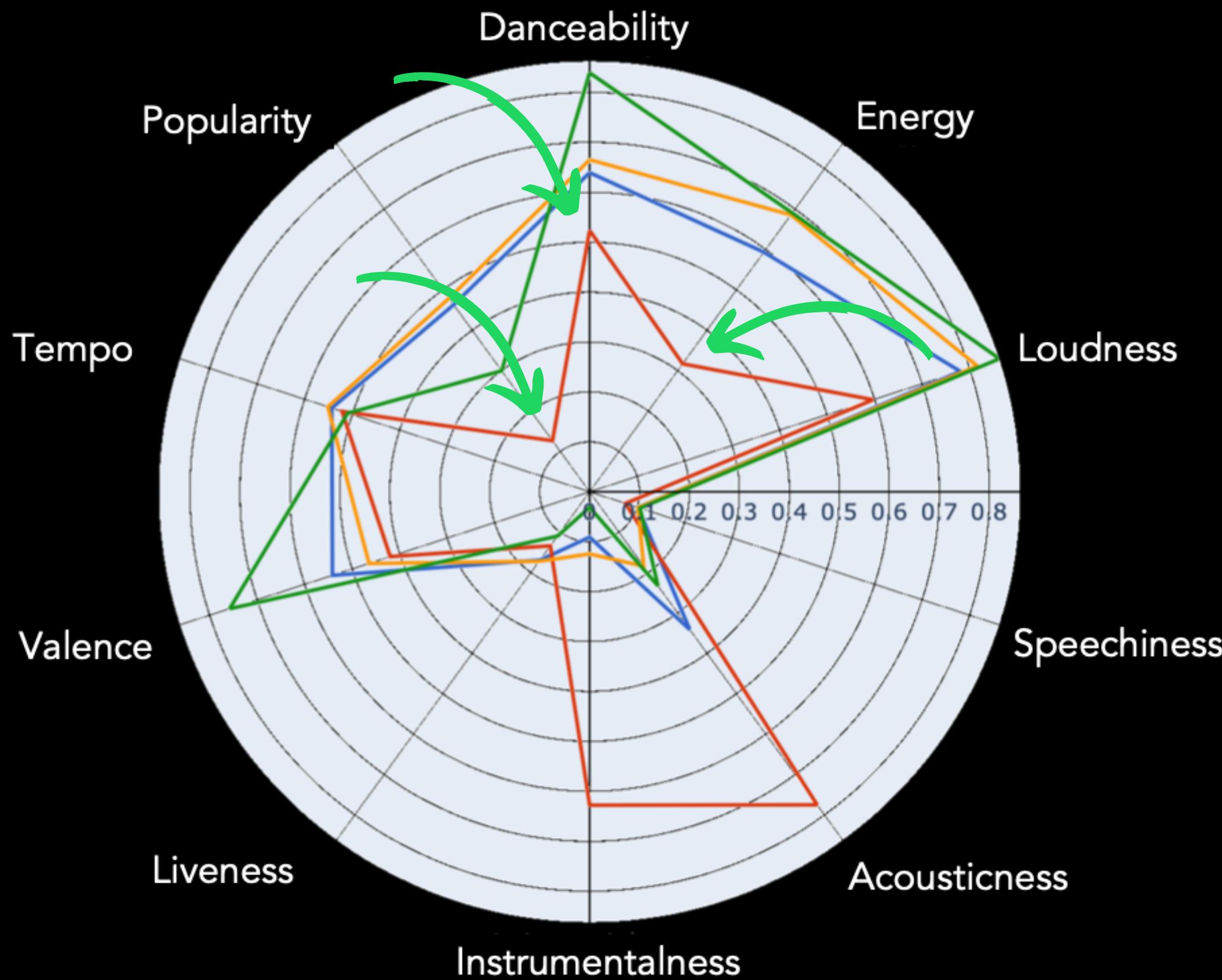
Cluster 2

Alex
Elina
Hanning
Lennart
Mangara
Michael
Sara

Cluster 3

Emmanuella

CLUSTER 1



Cluster 0

Alyssa
Arvind
Loizos
Mariela
Ravenn
Rebecca
Sabrina
Sita

Cluster 1

Sam

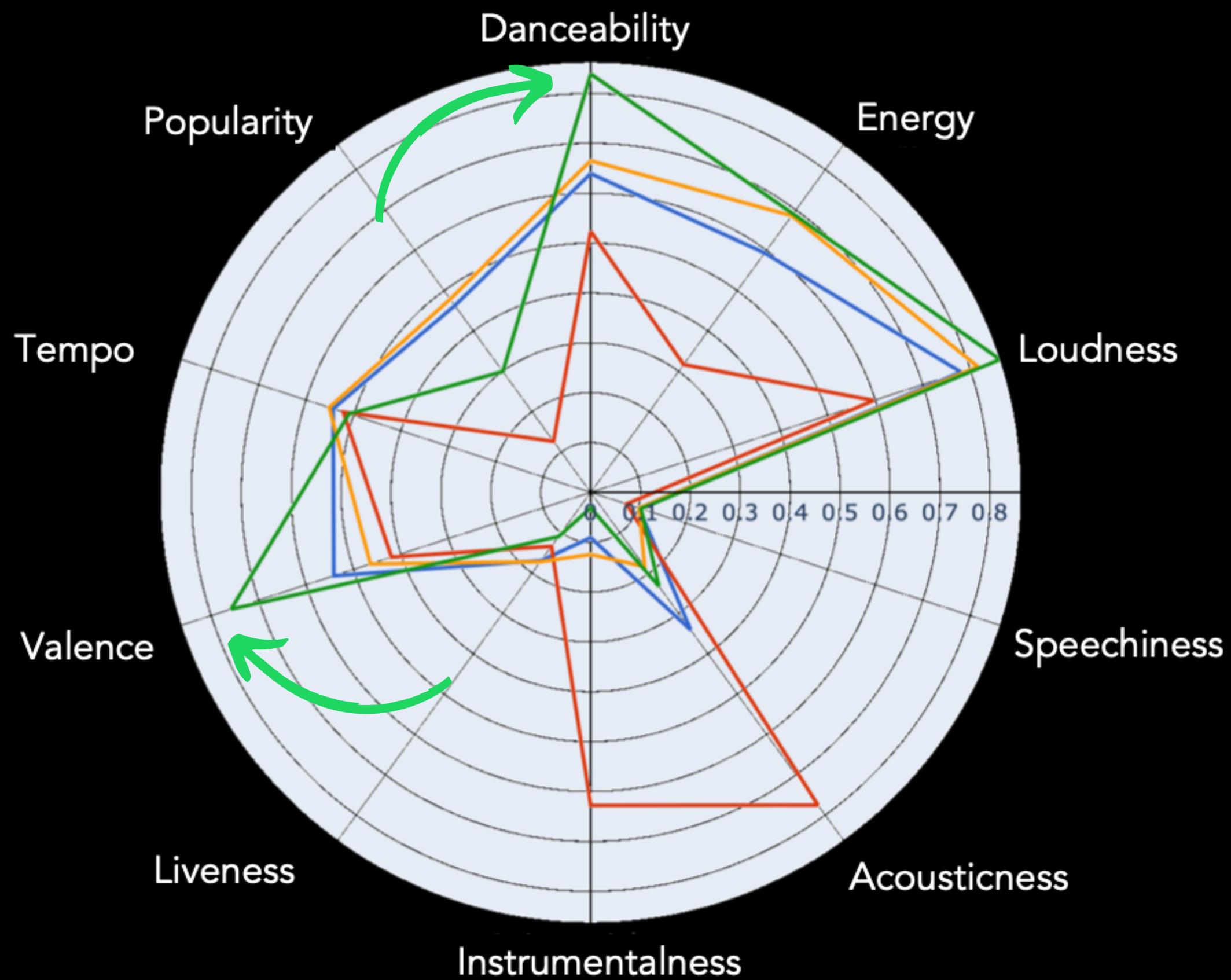
Cluster 2

Alex
Elina
Hanning
Lennart
Mangara
Michael
Sara

Cluster 3

Emmanuella

CLUSTER 3



Cluster 0

Alyssa
Arvind
Loizos
Mariela

Ravenn
Rebecca
Sabrina
Sita

Cluster 1

Sam

Cluster 2

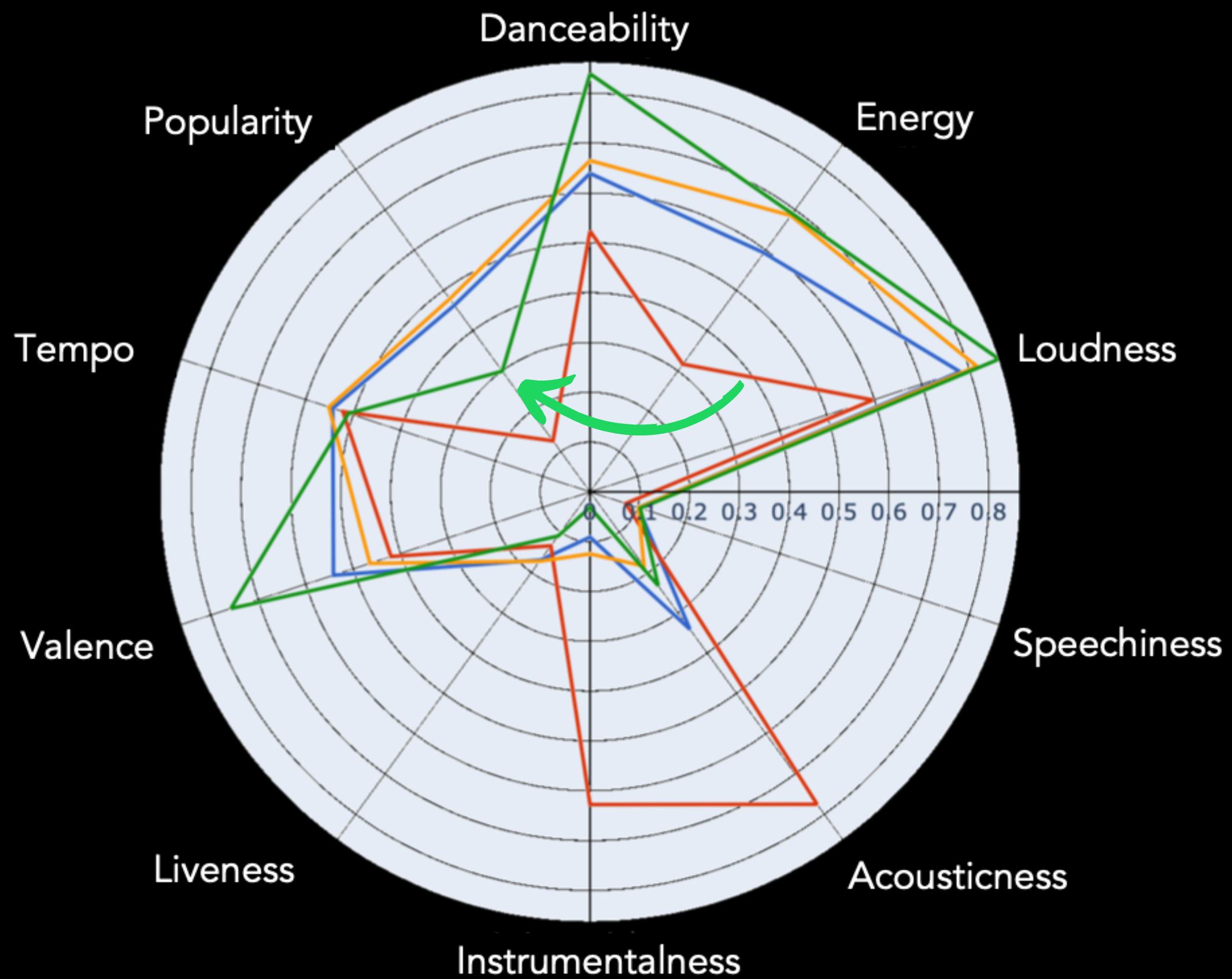
Alex
Elina
Hanning
Lennart

Mangara
Michael
Sara

Cluster 3

Emmanuella

CLUSTER 3



Cluster 0

Alyssa
Arvind
Loizos
Mariela

Cluster 1

Sam

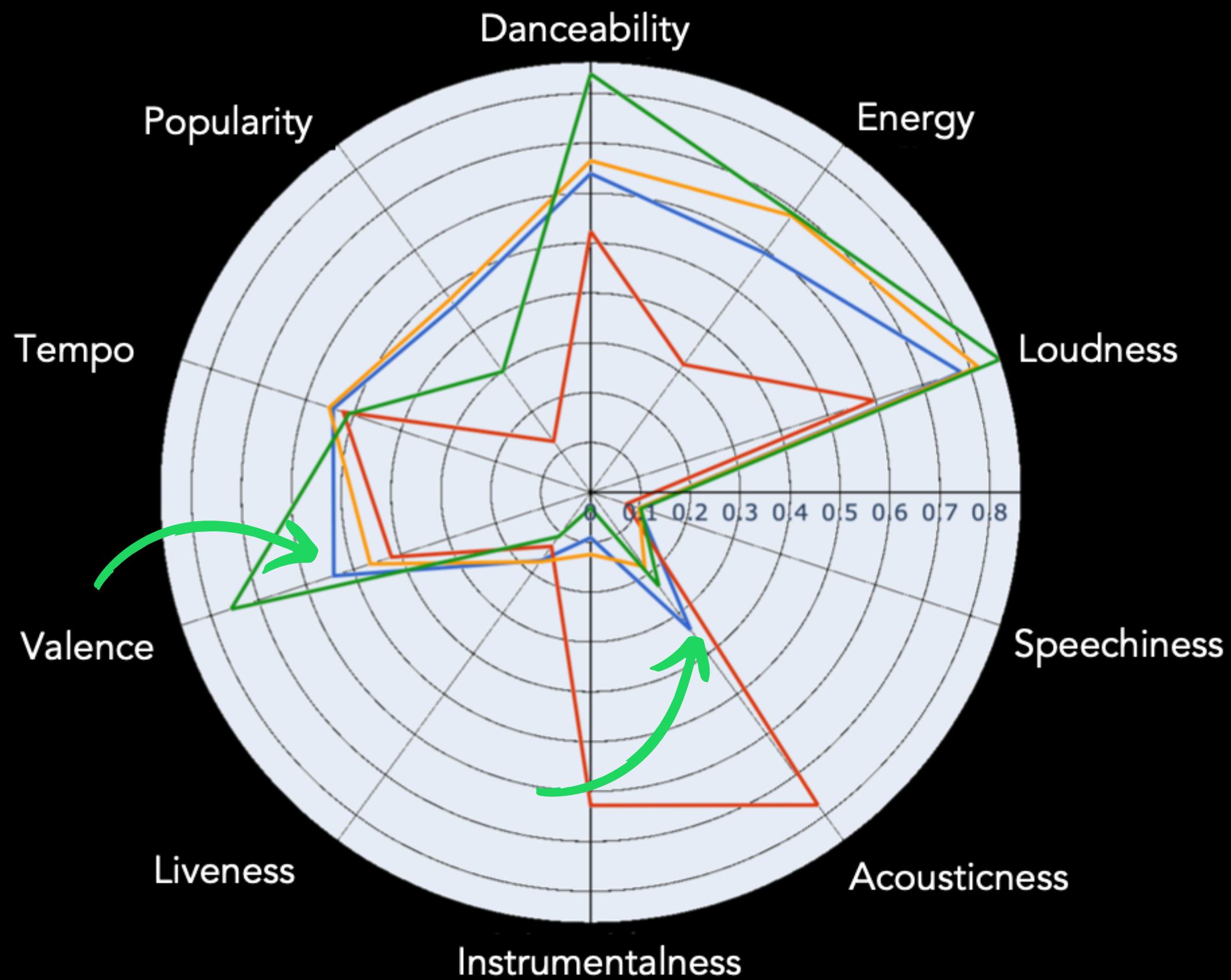
Cluster 2

Alex
Elina
Hanning
Lennart

Cluster 3

Emmanuella

CLUSTER 0



Cluster 0

Alyssa
Arvind
Loizos
Mariela
Sabrina
Sita

Cluster 1

Sam

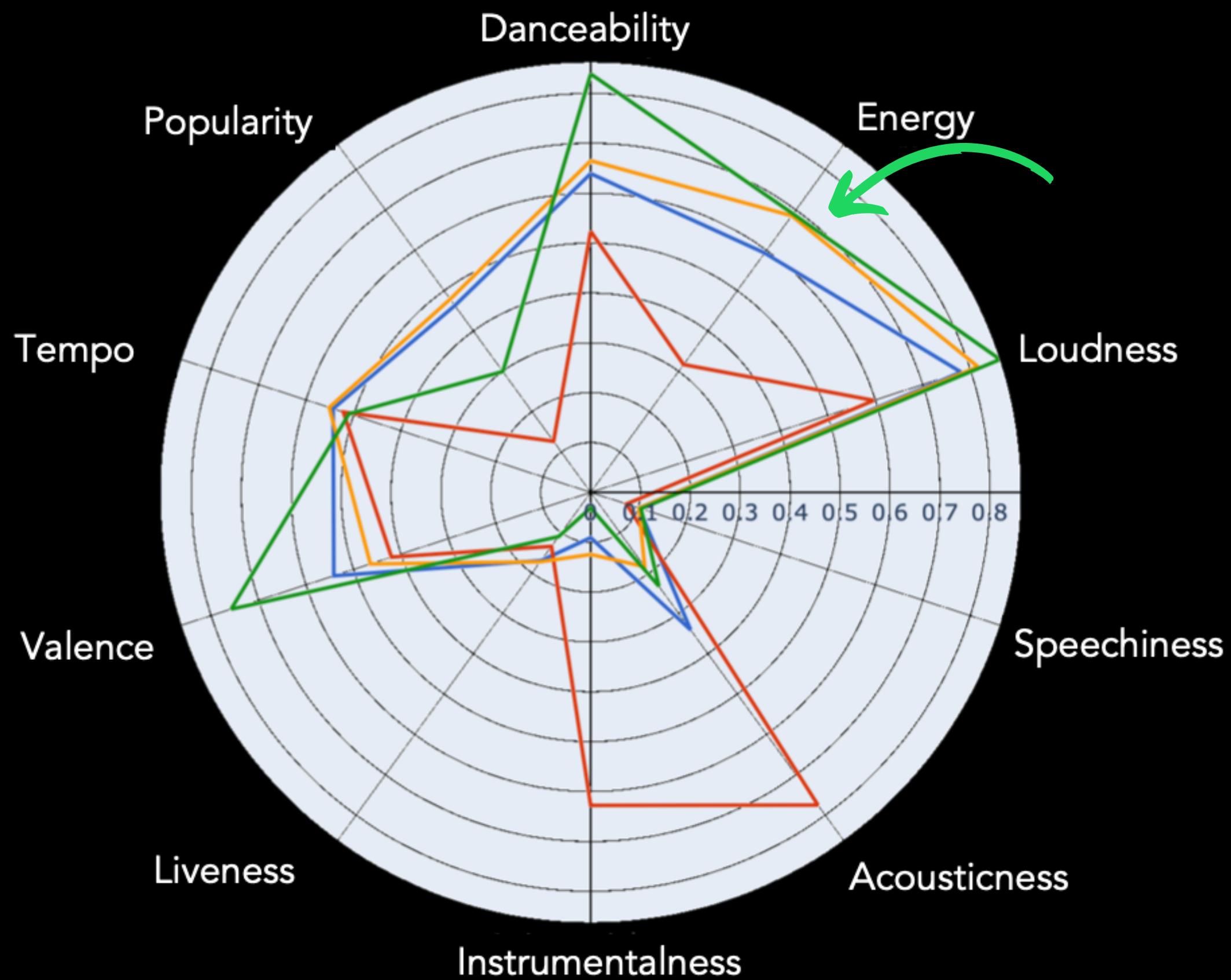
Cluster 2

Alex
Elina
Hanning
Lennart
Mangara
Michael
Sara

Cluster 3

Emmanuella

CLUSTER 2



Cluster 0

Alyssa
Arvind
Loizos
Mariela
Sabrina
Sita

Cluster 1

Sam

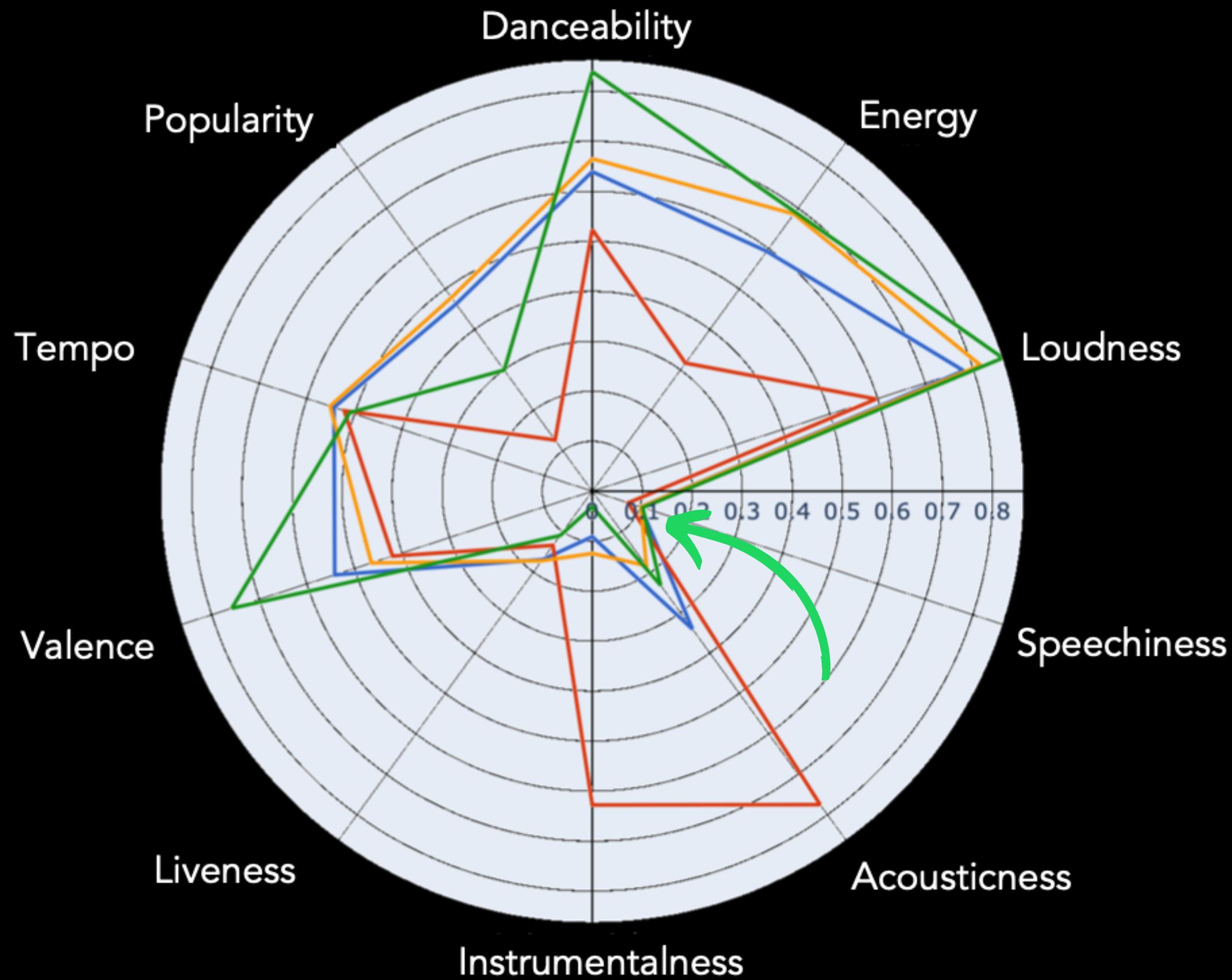
Cluster 2

Alex
Elina
Hanning
Lennart
Mangara
Michael
Sara

Cluster 3

Emmanuella

CLUSTERING



Cluster 0

Alyssa
Arvind
Loizos
Mariela
Ravenn
Rebecca
Sabrina
Sita

Cluster 1

Sam

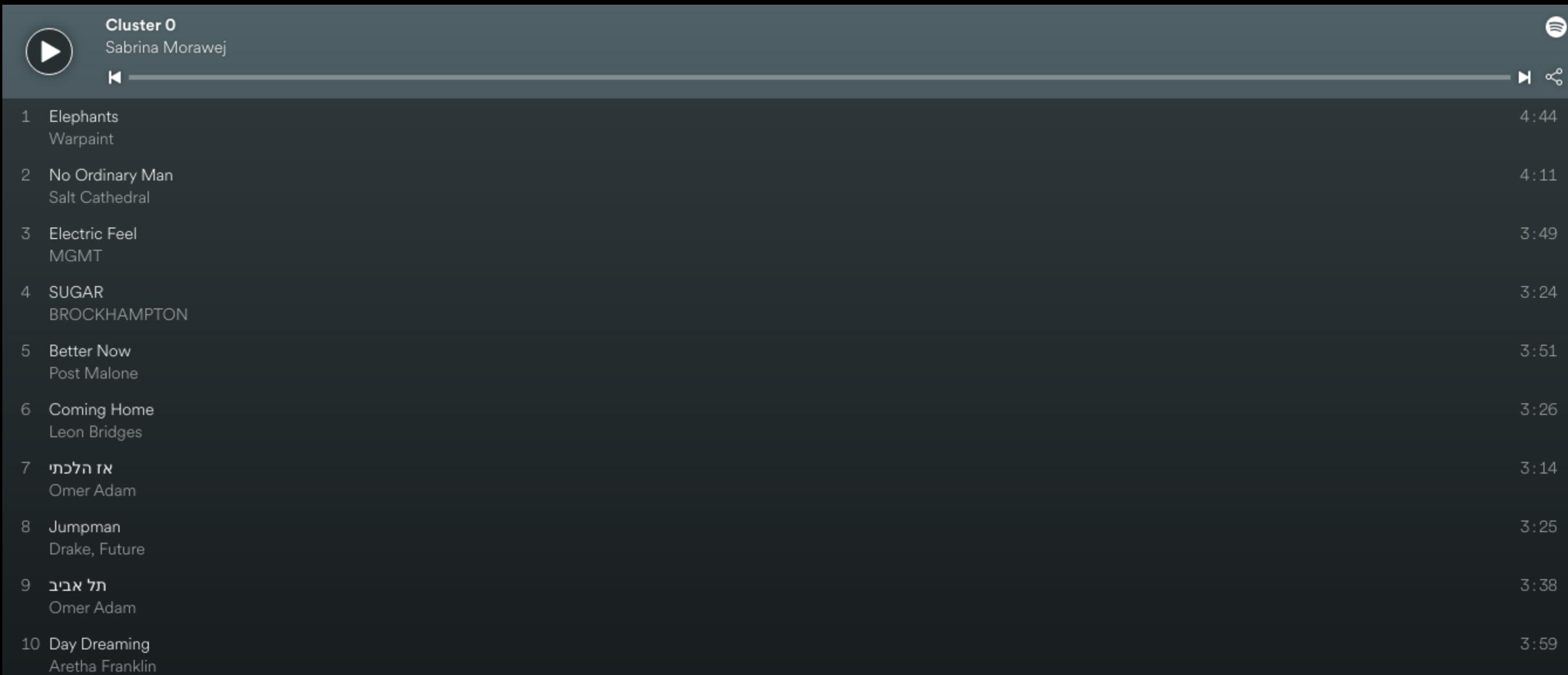
Cluster 2

Alex
Elina
Hanning
Lennart
Mangara
Michael
Sara

Cluster 3

Emmanuella

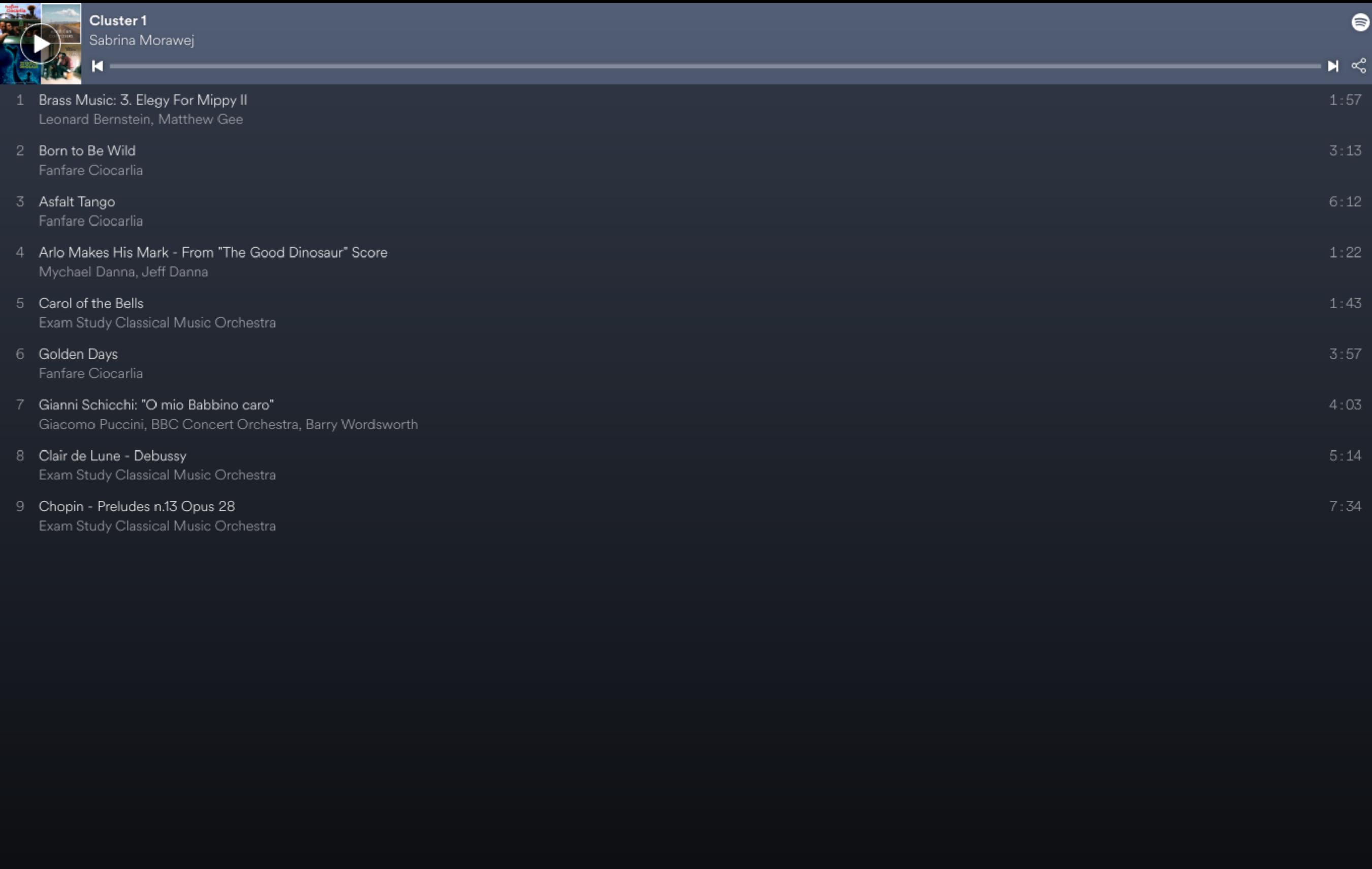
CLUSTER PLAYLISTS - 0



A screenshot of a Spotify player interface. The top bar is dark grey with white text. On the left is a play button icon. In the center, it says "Cluster 0" above "Sabrina Morawej". On the right is a Spotify logo and a shuffle button. Below the bar is a horizontal grey progress bar with arrows at both ends. The main area is a white table with a dark grey header row. The header contains columns for the song number (1-10), the song title, the artist, and the duration. The body of the table lists ten songs. The first song is "Elephants" by Warpaint (4:44). The second is "No Ordinary Man" by Salt Cathedral (4:11). The third is "Electric Feel" by MGMT (3:49). The fourth is "SUGAR" by BROCKHAMPTON (3:24). The fifth is "Better Now" by Post Malone (3:51). The sixth is "Coming Home" by Leon Bridges (3:26). The seventh is "از הלקוח" by Omer Adam (3:14). The eighth is "Jumpman" by Drake, Future (3:25). The ninth is "תל אביב" by Omer Adam (3:38). The tenth is "Day Dreaming" by Aretha Franklin (3:59).

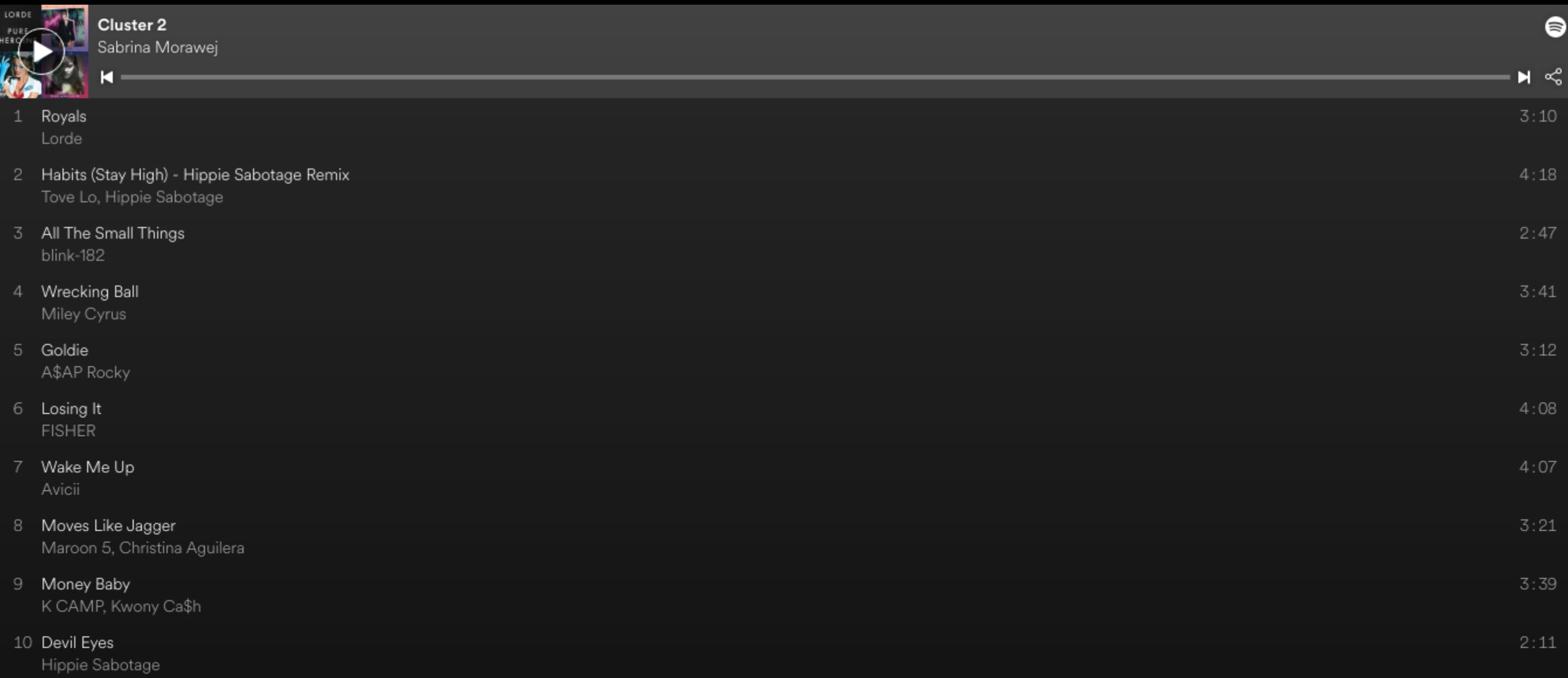
1	Elephants	Warpaint	4:44
2	No Ordinary Man	Salt Cathedral	4:11
3	Electric Feel	MGMT	3:49
4	SUGAR	BROCKHAMPTON	3:24
5	Better Now	Post Malone	3:51
6	Coming Home	Leon Bridges	3:26
7	از הלקוח	Omer Adam	3:14
8	Jumpman	Drake, Future	3:25
9	תל אביב	Omer Adam	3:38
10	Day Dreaming	Aretha Franklin	3:59

CLUSTER PLAYLISTS - 1

A screenshot of a Spotify playlist titled "Cluster 1" by Sabrina Morawej. The interface shows a dark blue header with the title and artist name. Below the header is a play bar with a play button, a back arrow, a progress slider, and a forward arrow. To the right of the play bar are a share icon and a Spotify logo. The main area displays a list of nine tracks with their titles, artists, and durations.

Rank	Track	Artist	Duration
1	Brass Music: 3. Elegy For Mippy II	Leonard Bernstein, Matthew Gee	1:57
2	Born to Be Wild	Fanfare Ciocarlia	3:13
3	Asfalt Tango	Fanfare Ciocarlia	6:12
4	Arlo Makes His Mark - From "The Good Dinosaur" Score	Mychael Danna, Jeff Danna	1:22
5	Carol of the Bells	Exam Study Classical Music Orchestra	1:43
6	Golden Days	Fanfare Ciocarlia	3:57
7	Gianni Schicchi: "O mio Babbino caro"	Giacomo Puccini, BBC Concert Orchestra, Barry Wordsworth	4:03
8	Clair de Lune - Debussy	Exam Study Classical Music Orchestra	5:14
9	Chopin - Preludes n.13 Opus 28	Exam Study Classical Music Orchestra	7:34

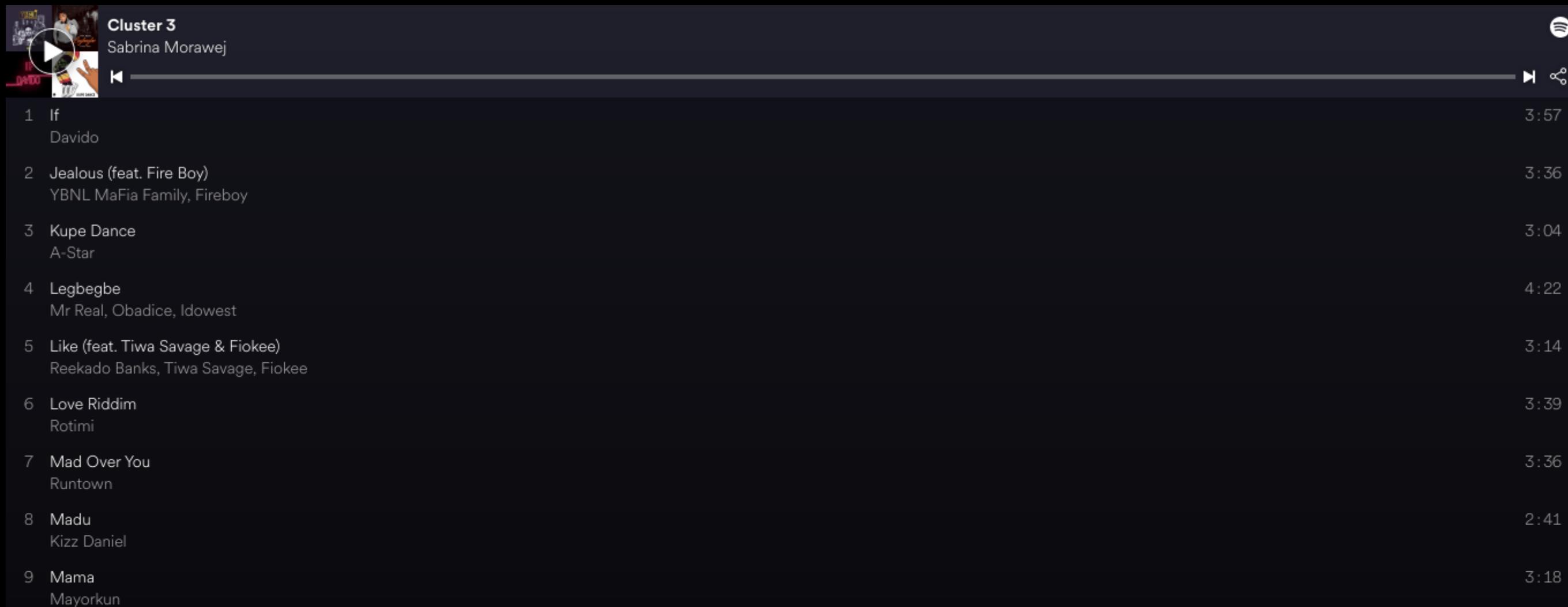
CLUSTER PLAYLISTS - 2



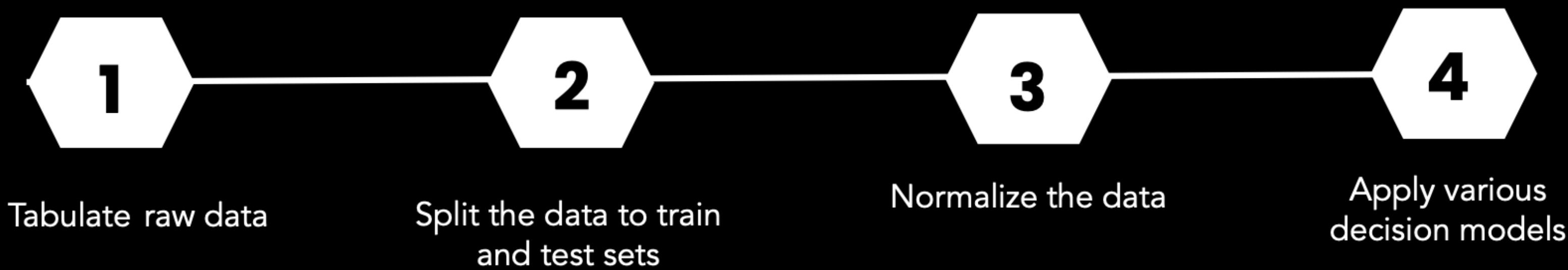
A screenshot of a Spotify playlist titled "Cluster 2" by Sabrina Morawej. The interface shows a dark-themed player with a play button, a progress bar, and a share icon. The playlist contains 10 tracks, each with a thumbnail, title, artist, and duration.

Rank	Song	Artist	Duration
1	Royals	Lorde	3:10
2	Habits (Stay High) - Hippie Sabotage Remix	Tove Lo, Hippie Sabotage	4:18
3	All The Small Things	blink-182	2:47
4	Wrecking Ball	Miley Cyrus	3:41
5	Goldie	A\$AP Rocky	3:12
6	Losing It	FISHER	4:08
7	Wake Me Up	Avicii	4:07
8	Moves Like Jagger	Maroon 5, Christina Aguilera	3:21
9	Money Baby	K CAMP, Kwony Ca\$h	3:39
10	Devil Eyes	Hippie Sabotage	2:11

CLUSTER PLAYLISTS - 3



PROCESSING FOR PREDICTION



Note : Clusters 1 and 3 were used as training data for songs that Alex would not like since those are the most different clusters from cluster 2 (Alex's cluster).

Cluster 0 is not used as the list of songs to predict Alex's playlist since we first want to make the prediction accuracy as high as possible. We can then slowly add more songs into the pool to continuously train the model.

RECOMMENDED PLAYLISTS

Using Boosting



Boosting



Logistic
Regression



SVM



Random
Forest



Bagging

ALEX'S CURRENT PREFERENCE



Loudness
0.817



Energy
0.676



Danceability
0.618



Popularity
0.463



Valence
0.459



Acousticness
0.182



Liveness
0.172



Speechiness
0.075



Tempo
0.075



Instrumentalness
0.032

RECOMMENDED PLAYLISTS

Top 10 Songs Alex Would Add to His Playlist



Rank	Song Title	Artist(s)	Length
1	Through to You	Tobiah	3:27
2	Do You Think About Me	Captain Cuts, Zooképer, Georgia Ku	3:36
3	Runaway (U & I)	Galantis	3:47
4	Rave	Sam Paganini	6:45
5	Losing It	FISHER	4:08
6	Fingerprint	Lane 8	6:44
7	Habits (Stay High) - Hippie Sabotage Remix	Tove Lo, Hippie Sabotage	4:18
8	Conjure Dreams	Maceo Plex	7:21
9	Stella - Kölisch Remix	Jam & Spoon, Kölisch	5:01
10	Mad at Love	Gioli & Assia	4:11
11	I Wait for You - Ben Pearce Remix	SOULS, Ben Pearce	6:44
12	Upperground	ARTBAT	6:45
13	Skywalking	Mathame	6:14
14	Eye Of The Swarm	Spektre	7:15
15	Kamikaze		6:26

RECOMMENDED PLAYLISTS

Top 10 Songs Alex Would NOT Add to His Playlist



Rank	Song Title	Artist(s)	Length
1	Problems	Petit Biscuit, Lido	3:39
2	Body	Loud Luxury, Brando	2:43
3	Feelin' - Interlude	Hayden James	2:43
4	Better Together	Hayden James, Running Touch	3:56
5	Just Friends	Hayden James, Boy Matthews	3:58
6	Favours	Hayden James, Nat Dunn	3:54
7	Sad in Scandinavia	Seeb, Zak Abel	3:24
8	LA & The Parties	Lost Kings, Sevyn Streeter, Luh Kel	2:45
9	Secrets	Regard, RAYE	2:56
10	Act of God - Robotaki Remix	Tyler Touché, Jason Gaffner, Robotaki	4:38
11	Regardless	RAYE, Rudimental	3:17
12	Love Is	Sons Of Maria	3:25
13	History	Cosmo's Midnight	3:28
14	Something About You	Hayden James	3:45

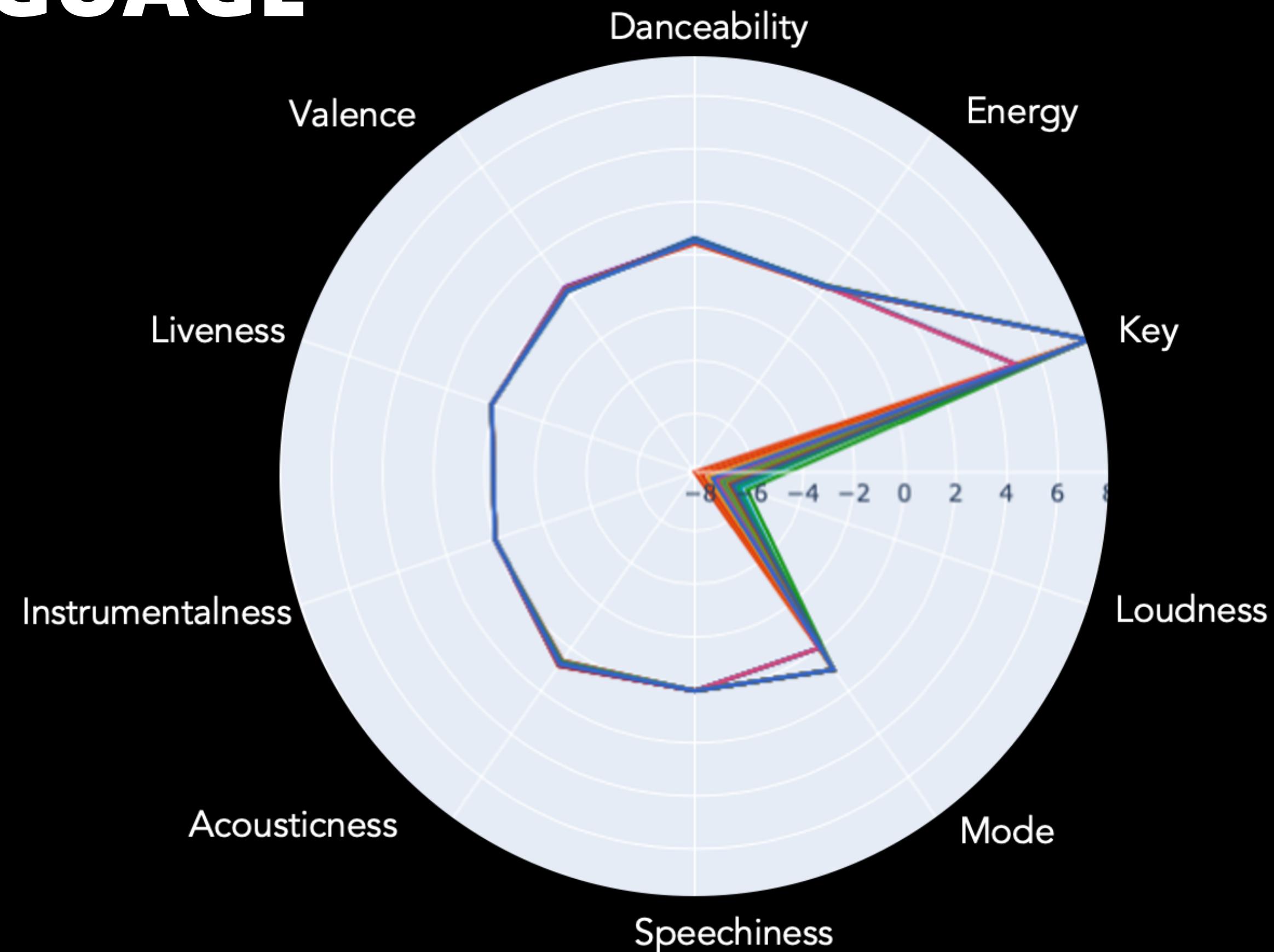
HOW WE CONSTRUCTED THE ALGORITHM TO PREDICT SONGS FOR ALEX

- In coding :
 - 0 = Clusters 1 & 3 - songs opposite of what Alex is already listening
 - 1 = Songs in Alex's playlist
- We used the all the songs from users in Cluster 2 to predict which songs Alex would like and would not like.
- So even though Alex is in Cluster 2, there are still songs Alex may not be familiar with that other users had on their own playlists.
- We used Cluster 2 because those songs are the ones that Alex is most likely to enjoy and add to his own playlist. Once the algorithm has a high enough accuracy, we can then add more songs into the pool and further improve the accuracy.

EFFECT OF LANGUAGE ON MUSIC FEATURES



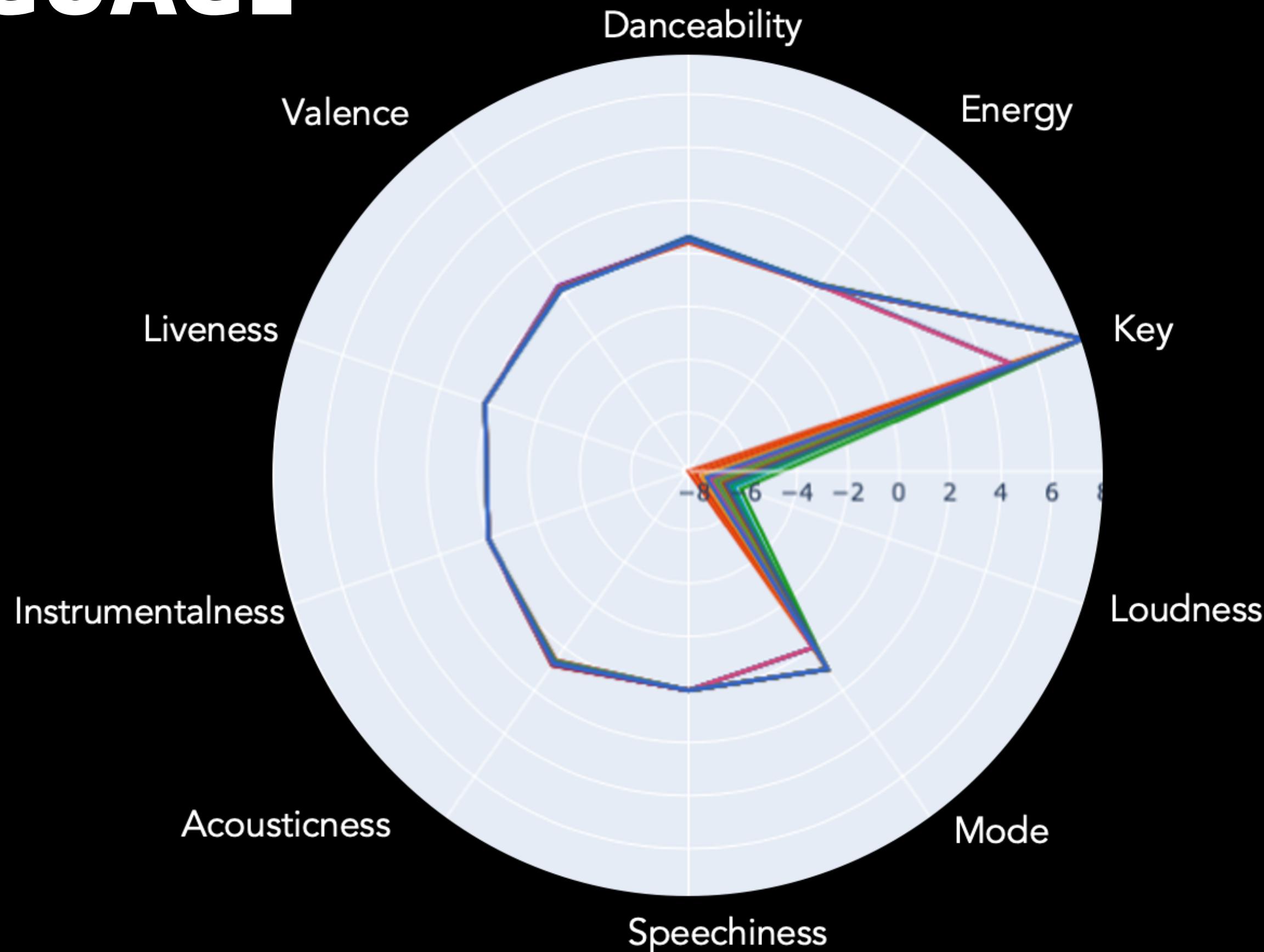
"Let It Go"
in
21 languages



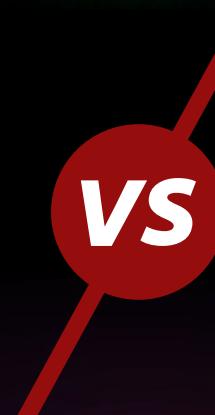
EFFECT OF LANGUAGE ON MUSIC FEATURES

"Let It Go" in 21 languages

- Highest value of Key is in Latvian
- Loudness was highest in Czech
- Loudness was least in Cantonese
- Minor to no difference in the rest of the categories indicating that there is almost minimal impact of the song's language on it's features and this would therefore not substantially impact recommendations.



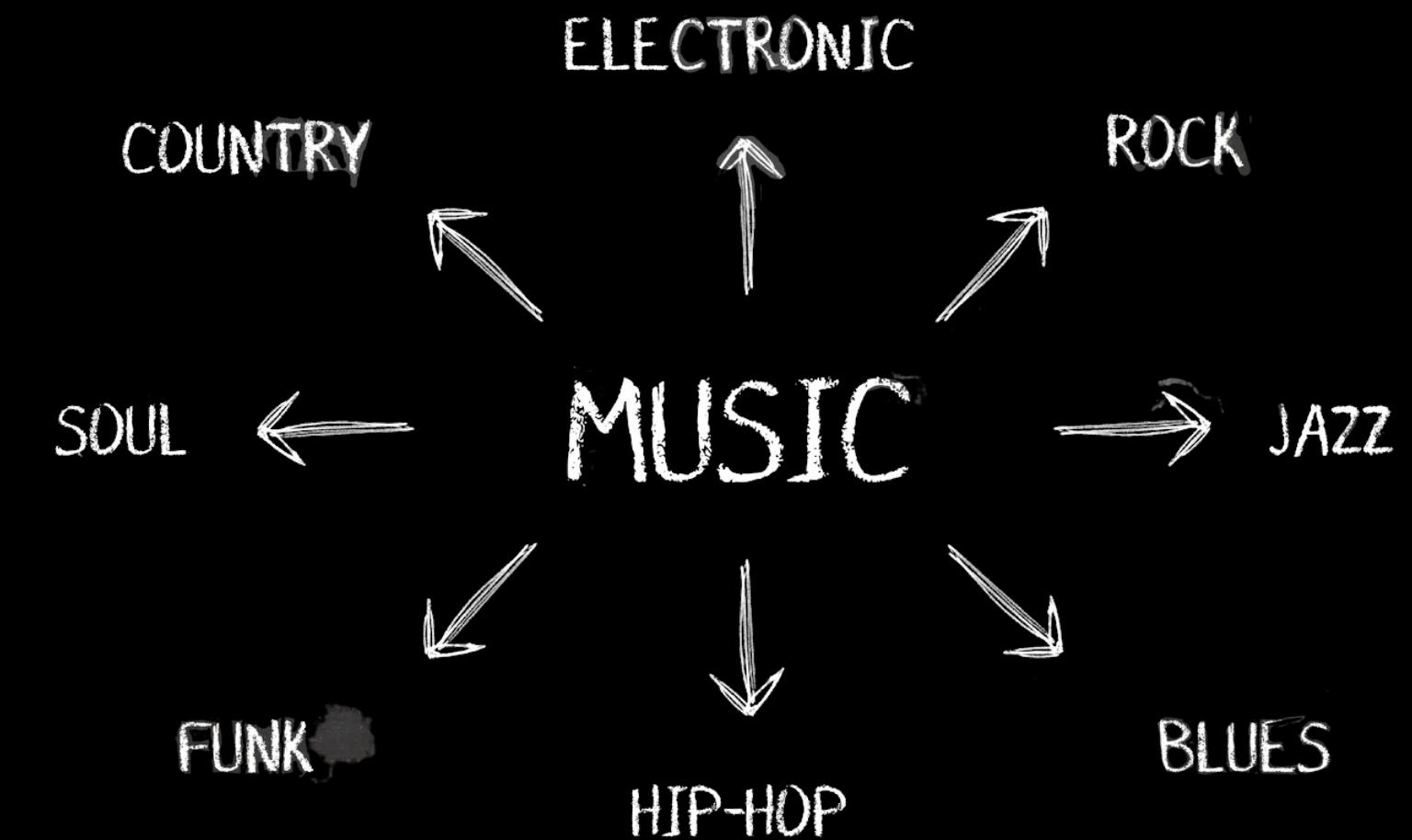
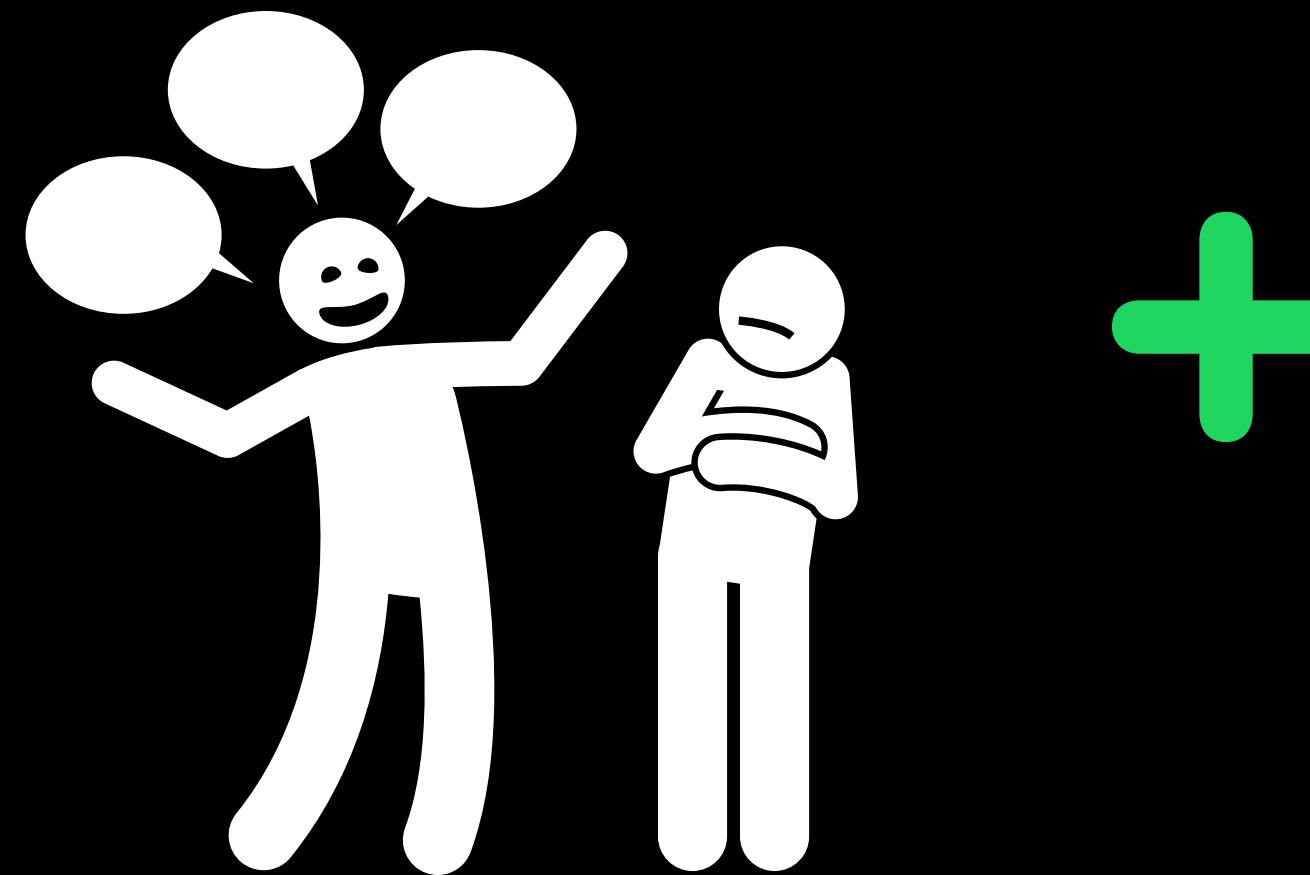
Conclusion



Spotify

- Has 130 million paid subscribers in 2020 (60 million more than Apple Music)
- Has 31% of global market share of music and podcast listeners (largest level)
- This is due to many factors including Spotify's:
 - (a) Robust podcast library
 - (b) Individualised song and playlist recommendations
 - (c) End of year summaries and statistics
- It is therefore imperative that we continue to develop features that prioritise and individualise the user experience to promote user acquisition and retention

FUTURE DIRECTIONS



**Link personality data to music preferences
to contribute to personalization**

**THANK YOU FOR
LISTENING!**



ANNEXURE

CLUSTER PLAYLISTS



Cluster 0

Elephants

از הלחתי

No Ordinary Man

Coming Home

Electric Feel

Jumpman

Sugar

Day Dreaming

Better Now

تل אביב

CLUSTER PLAYLISTS



Cluster 1

Brass Music: 3. Elegy For Mippy II

Born To Be Wild

Asfalt Tango

Arlo Makes His Mark

Carol of The Bells

Golden Days

Gianni Schicchi: O Mio Babbino Caro•

Clair de Lune – Debussy•

Amadeus•

Chopin - Preludes n.13 Opus 28•

CLUSTER PLAYLISTS



Cluster 2

Royals

Losing It

Habits (Stay High)

Wake Me Up

All The Small Things

Moves Like Jagger

Wrecking Ball

Money Baby

Goldie

Devil Eyes

CLUSTER PLAYLISTS



Cluster 3

If

Like (feat. Tiwa Savage & Fiokee

Jealous (feat Fire Boy)

Love Riddim

Joromi

Mad over Love

Kupe Dance

Madu

Legbegbe

Mama



RECOMMENDED PLAYLISTS

Top 5 Songs Alex Would **Love**

#	Title	Artist
1	Through to You	Tobiah's
2	Do You Think About Me	Captain Cuts
3	Runaway (U & I)	Galantis
4	Rave	Sam Paganini
5	Losing It	FISHER



RECOMMENDED PLAYLISTS

Top 5 Songs Alex Would Hate

#	Title	Artist
1	Problems	Petit Biscuit
2	Body	Loud Luxury
3	Feelin' - Interlude	Hayden James
4	Better Together	Hayden James
5	Just Friends	Hayden James