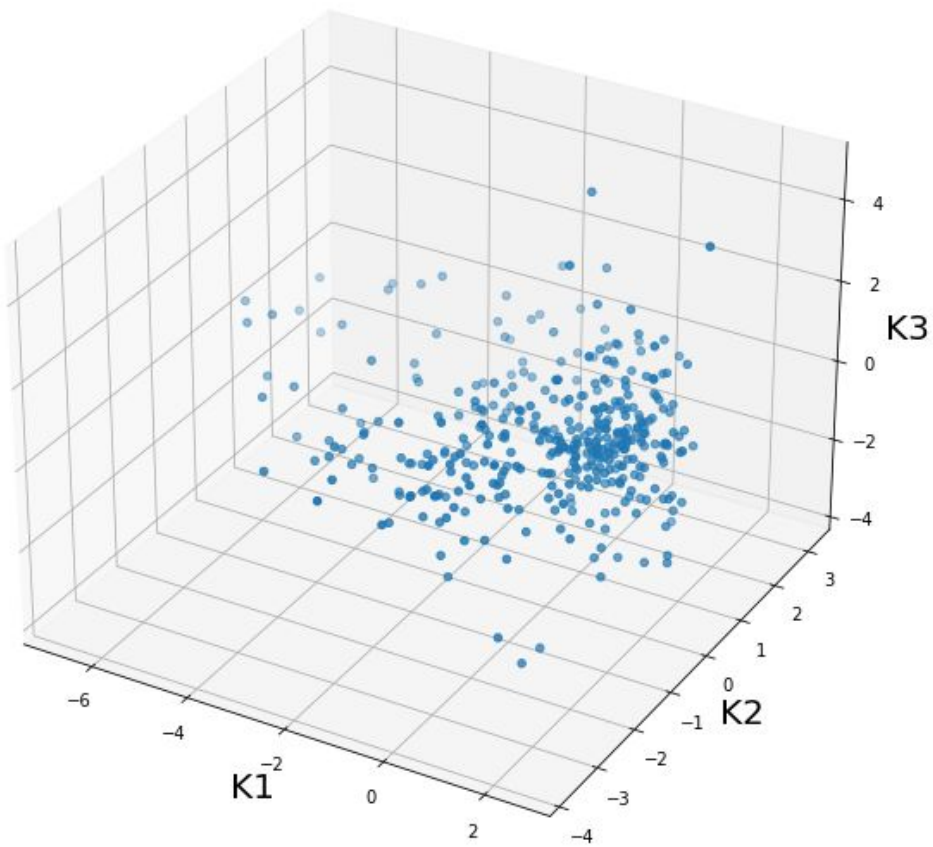




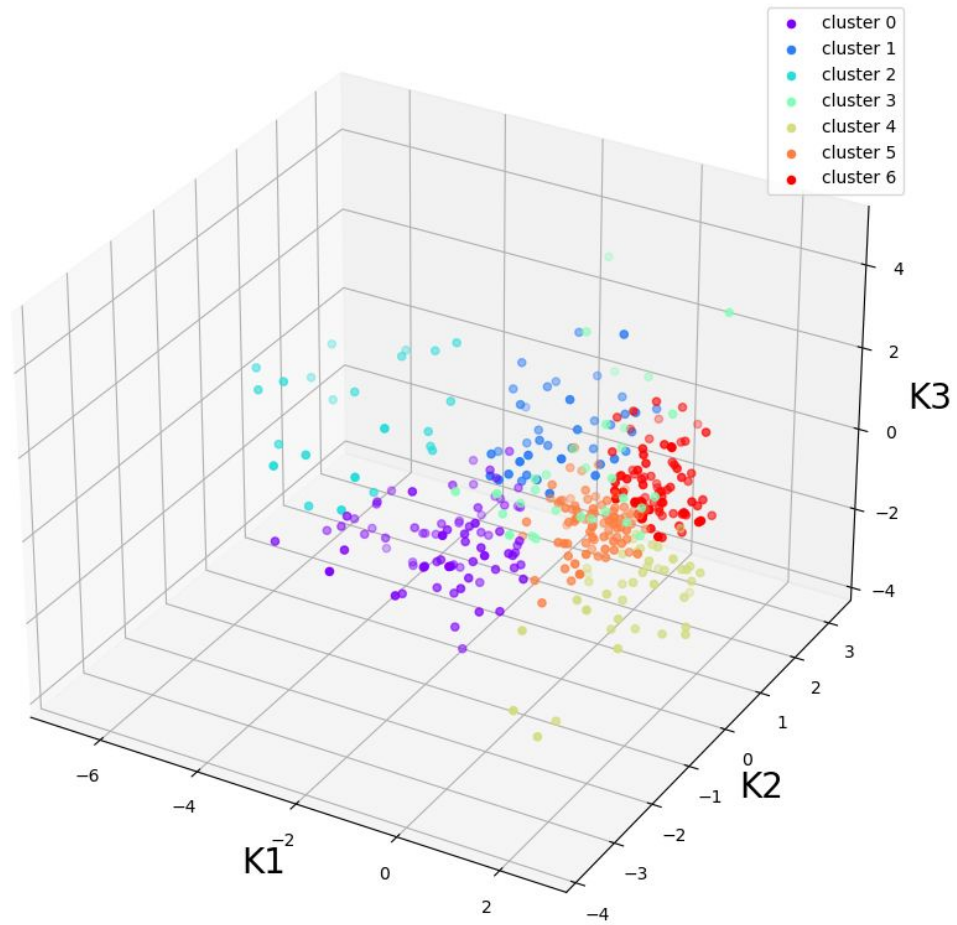
Spotify K-Means and PCA

Joshua Truong

PCA Dimensionality Reduction



Clusters





Spotify Audio Features

Feature	Description	Measure
Acousticness	A confidence measure if a track is acoustic.	0.0: low confidence; 1.0: high confidence
Danceability	Is the track suitable for dancing?	0.0: least danceable; 1.0: most danceable
Energy	A perceptual measure of intensity and activity. High entropy.	0.0: least energetic; 1.0: most energetic
Instrumentalness	Predicts whether a track contains no vocals.	1.0: no vocal content; 0.5: instrumental tracks
Liveness	Detects the presence of an audience in the recording.	1.0: likely performed live
Loudness	The overall loudness of a track in decibels (dB).	Values typically range between -60 and 0 db.
Speechiness	Speechiness detects the presence of spoken words in a track.	1.0: exclusively speech-like; 0.33-0.66: tracks that are probably made entirely of spoken words; <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).	
Valence	Describes the musical positiveness conveyed by a track.	1.0: Very positive; 0.0: Very negative



Sample of the DataSet

	track_name	artist_name	id	url	acousticness	danceability	energy	instrumentalness	liveness	loudness	speechiness	tempo	valence
0	Victory	Two Steps from Hell	2F9xBxKbx2M0pbgtSu8fLf	https://api.spotify.com/v1/tracks/2F9xBxKbx2M0...	0.00903	0.503	0.917	0.326	0.1620	-6.111	0.0649	119.993	0.266
1	Heart of Courage	Thomas Bergersen	4JtyWkWQTPVcroZf8Jlkp	https://api.spotify.com/v1/tracks/4JtyWkWQTPV...	0.51100	0.403	0.591	0.896	0.1080	-8.126	0.0514	135.047	0.192
2	Strength of a Thousand Men	Thomas Bergersen	3GJZLvGXaVszYdSBLMtJFX	https://api.spotify.com/v1/tracks/3GJZLvGXaVsz...	0.09620	0.478	0.906	0.919	0.1150	-6.781	0.0419	140.021	0.590
3	Blackheart	Thomas Bergersen	2jn89xTHIVsy167uOmAewt	https://api.spotify.com/v1/tracks/2jn89xTHIVsy...	0.09450	0.136	0.462	0.936	0.1190	-10.982	0.0367	161.073	0.112
4	Star Sky	Two Steps from Hell	06AMpcajziFnEKniV25fiU	https://api.spotify.com/v1/tracks/06AMpcajziFn...	0.00997	0.443	0.822	0.107	0.0889	-5.656	0.0508	130.055	0.361

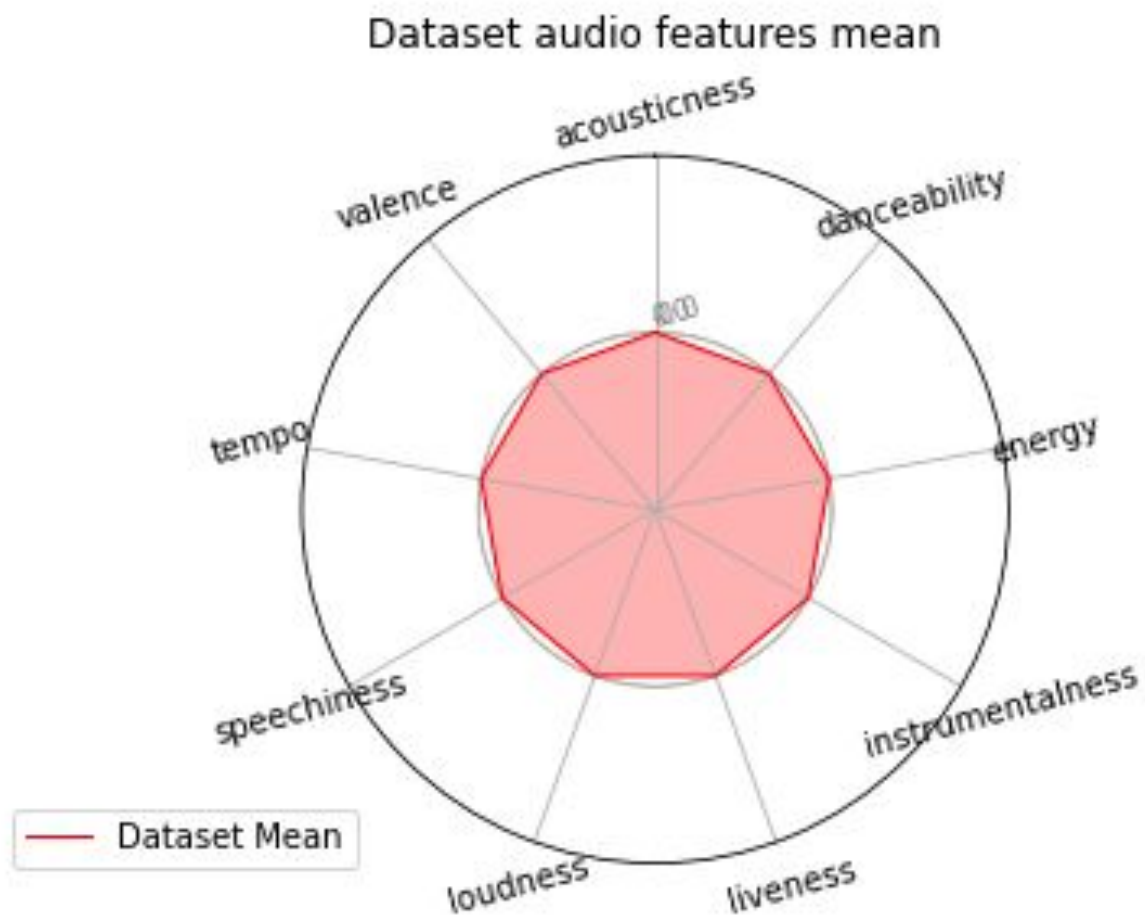


Normalization w/ Z-Score

Standardizes all features to a common scale. Especially, loudness and tempo.

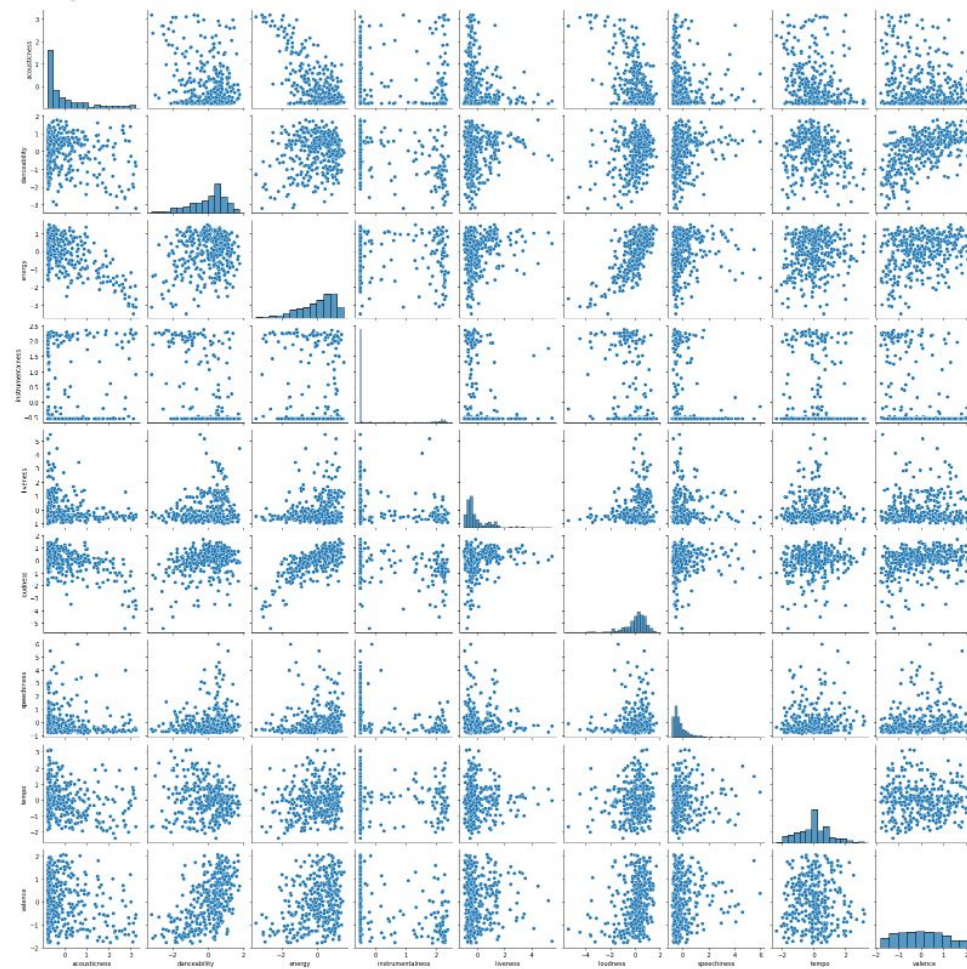
	acousticness	danceability	energy	instrumentalness	liveness	loudness	speechiness	tempo	valence
0	-0.739725	-0.687412	1.074172	0.442650	-0.071653	0.043965	-0.030835	-0.118852	-0.848106
1	1.315929	-1.322203	-0.625894	2.133579	-0.460824	-0.547213	-0.291775	0.490870	-1.152380
2	-0.382749	-0.846110	1.016808	2.201809	-0.410376	-0.152605	-0.475400	0.692329	0.484120
3	-0.389711	-3.017096	-1.298620	2.252241	-0.381549	-1.385132	-0.575910	1.544984	-1.481325
4	-0.735876	-1.068287	0.578754	-0.207023	-0.598475	0.177457	-0.303372	0.288682	-0.457484

Normalization w/ Z-Score

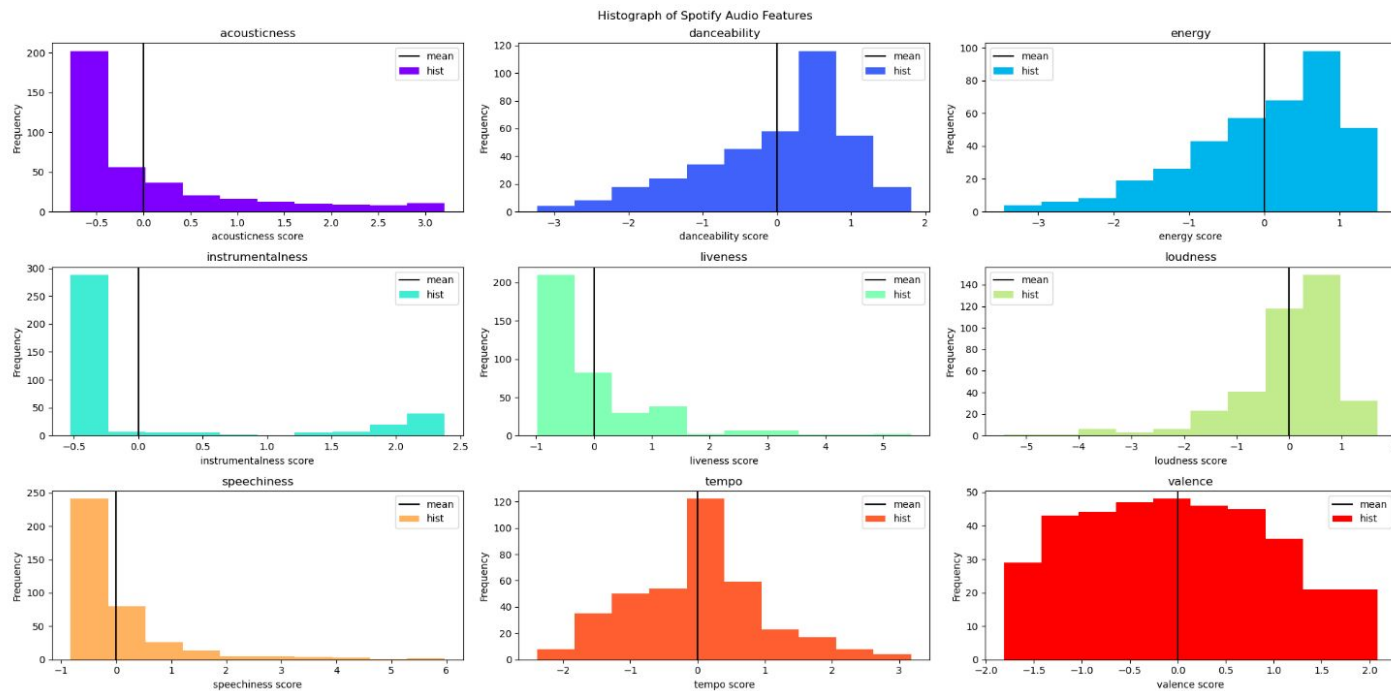




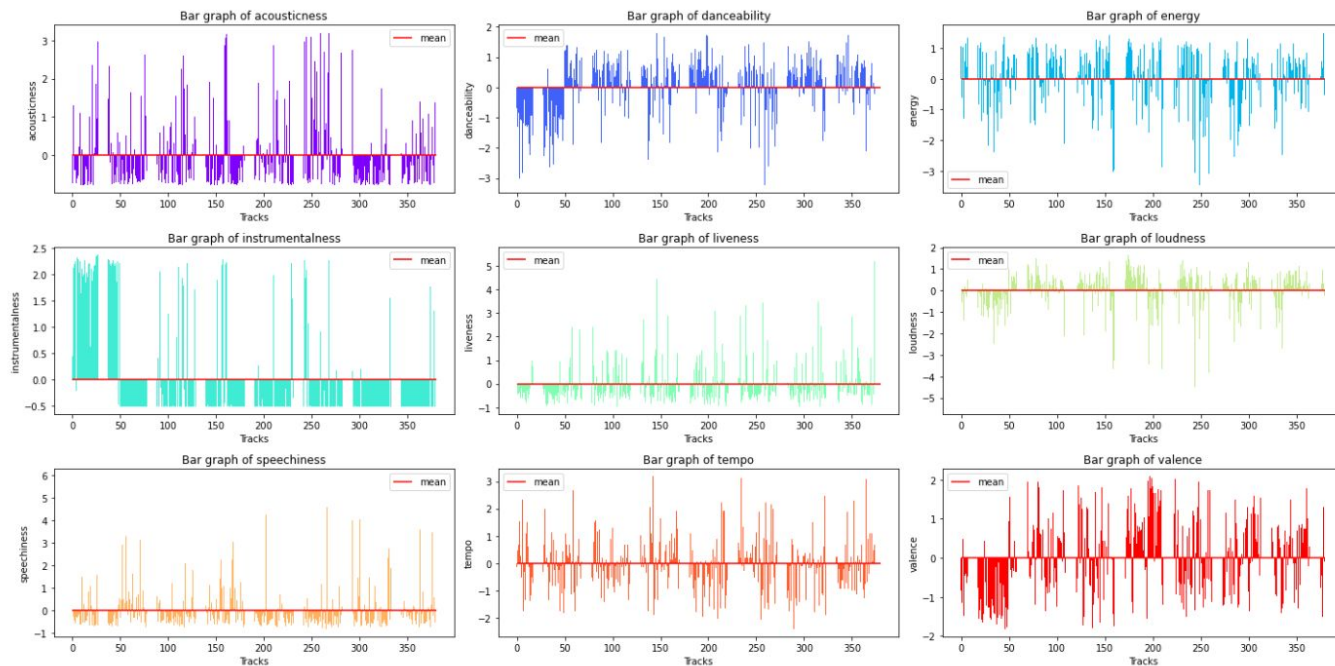
Preliminary Analysis



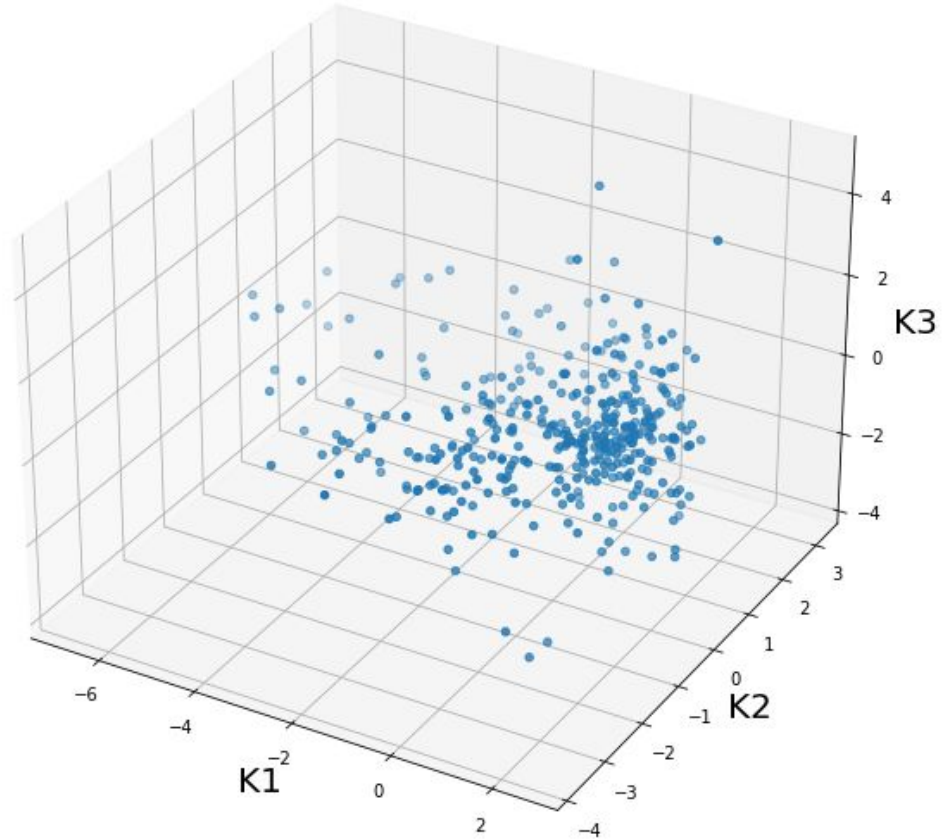
Audio Features Frequency Distribution



Data Visualization

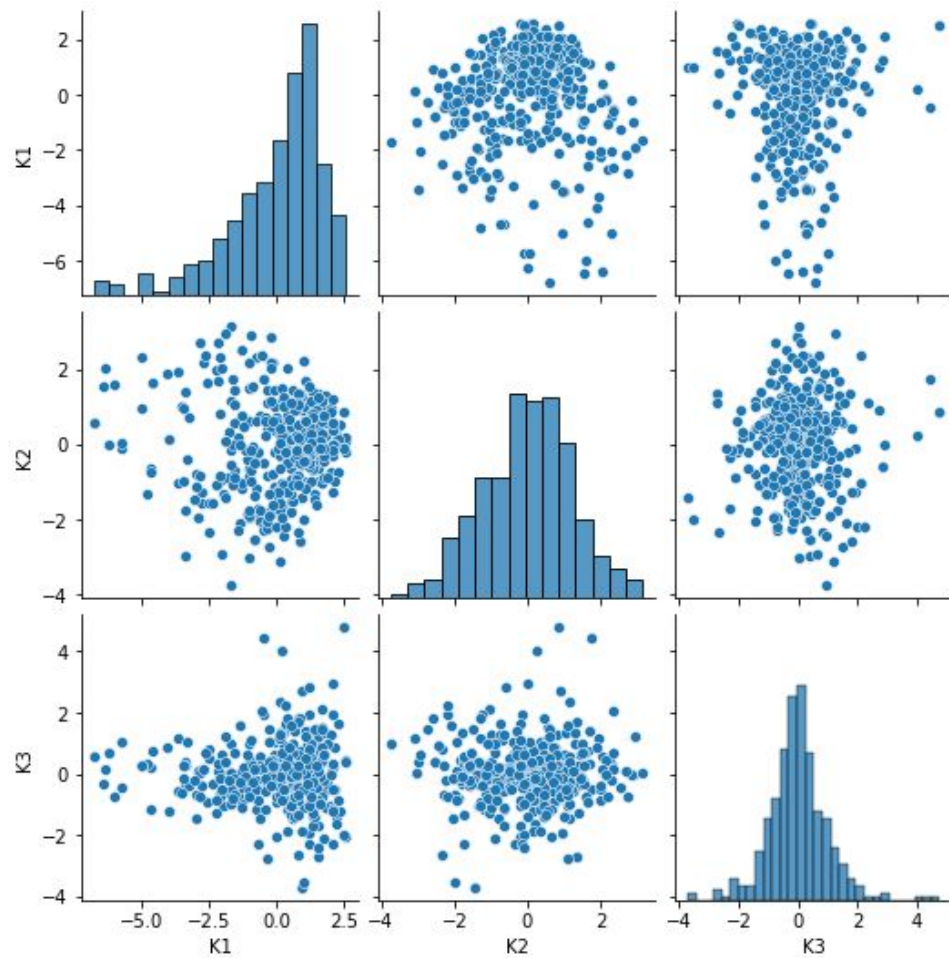


PCA Dimensionality Reduction

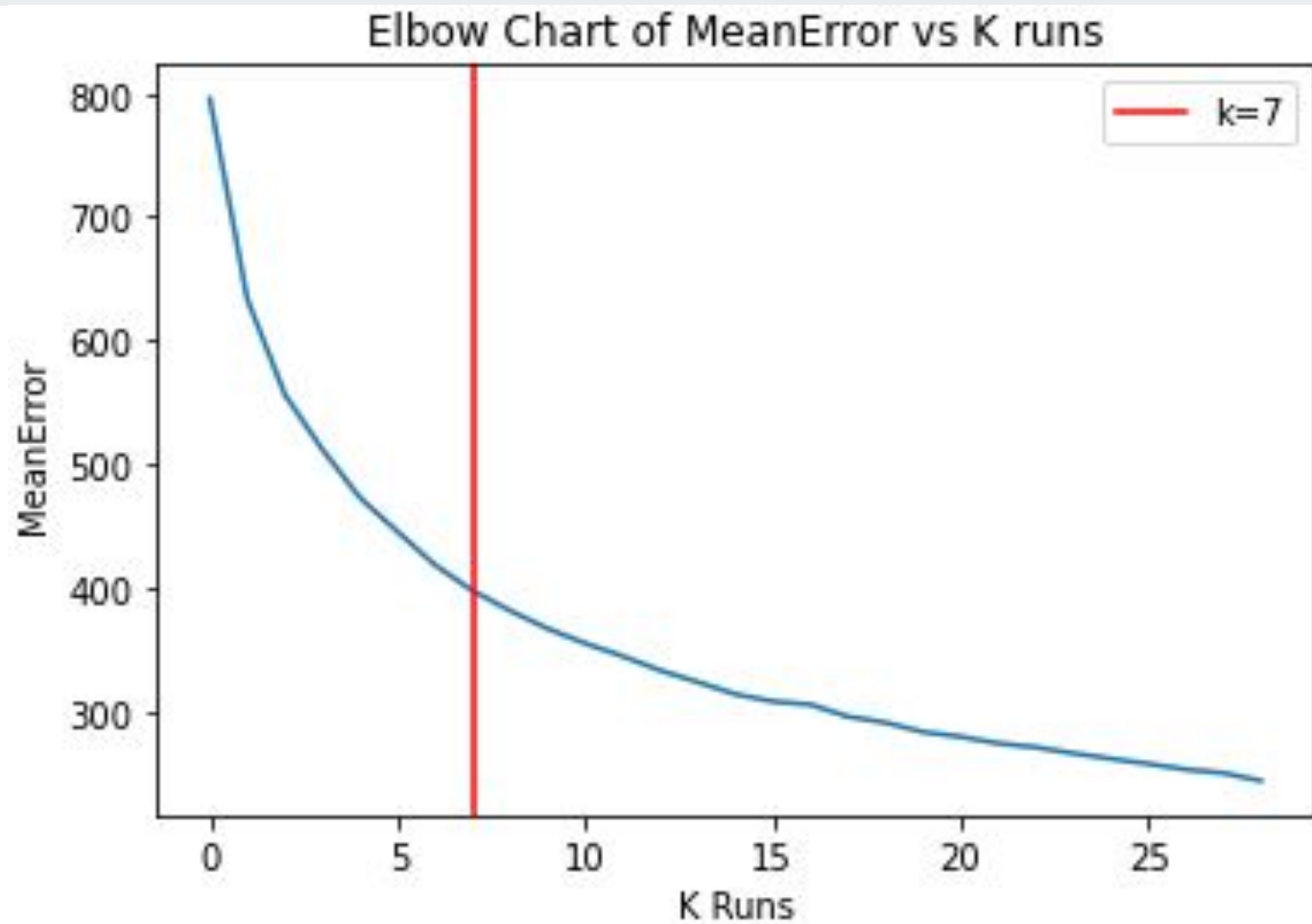




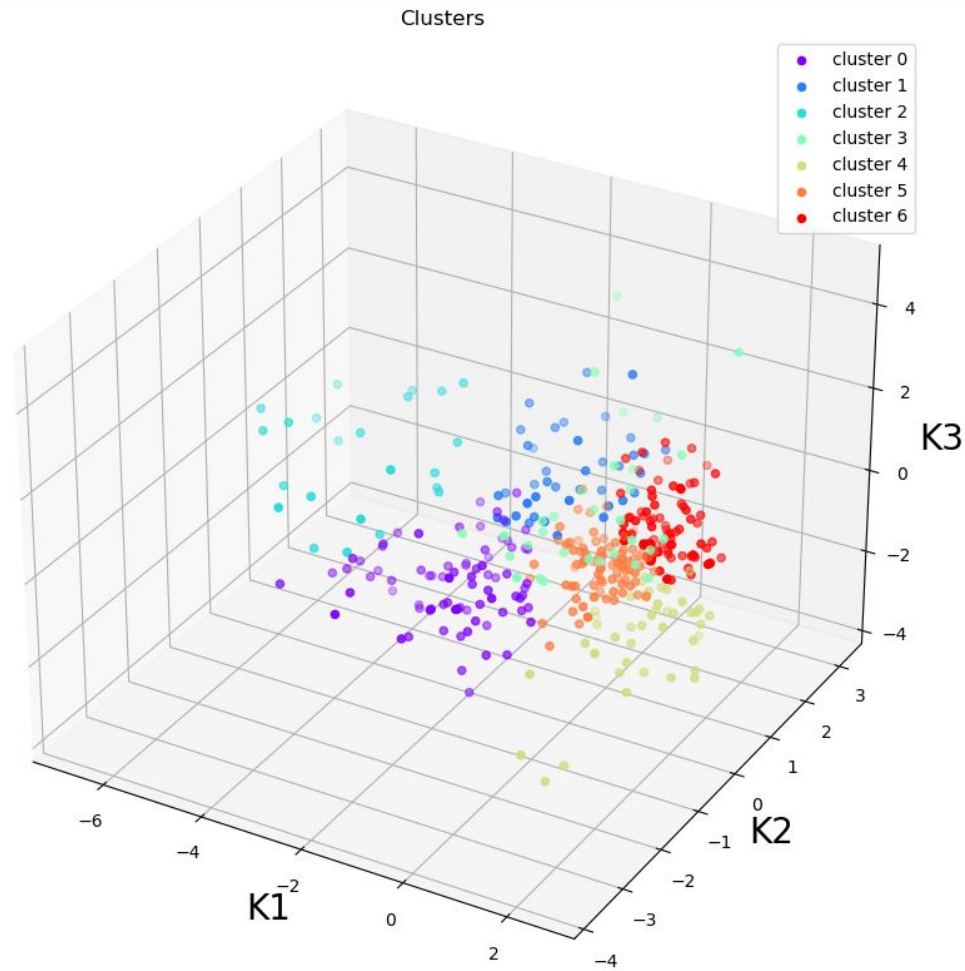
PCA Dimensionality Reduction



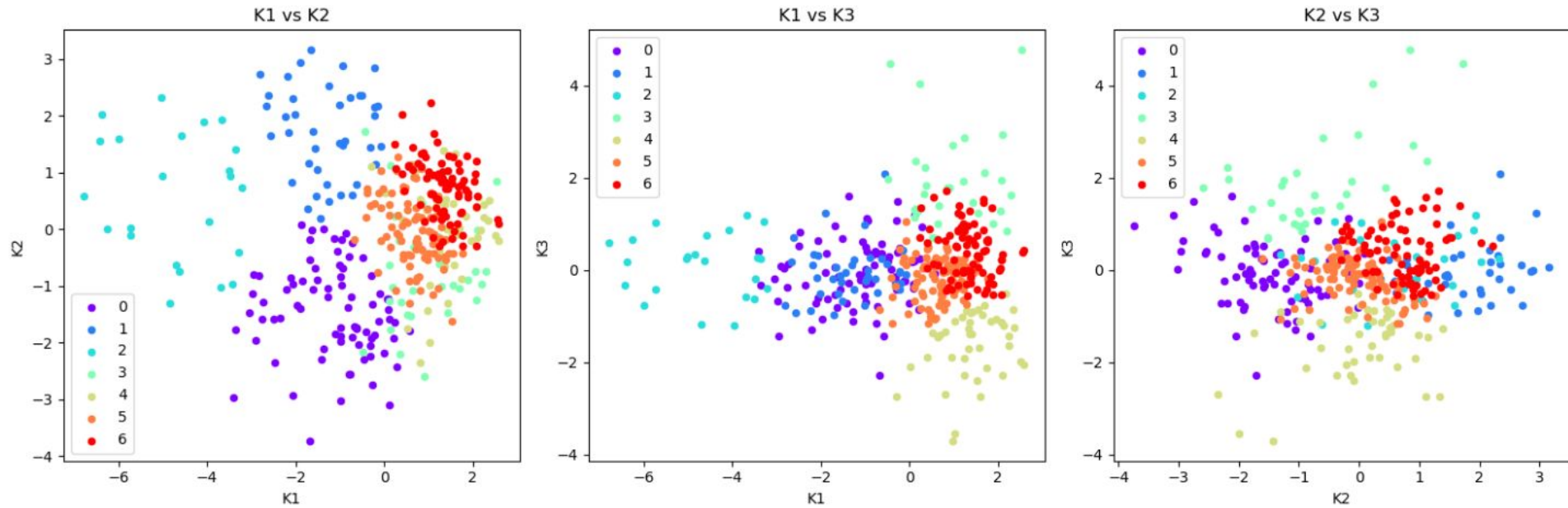
K-Means



PCA - Clusters in 3D

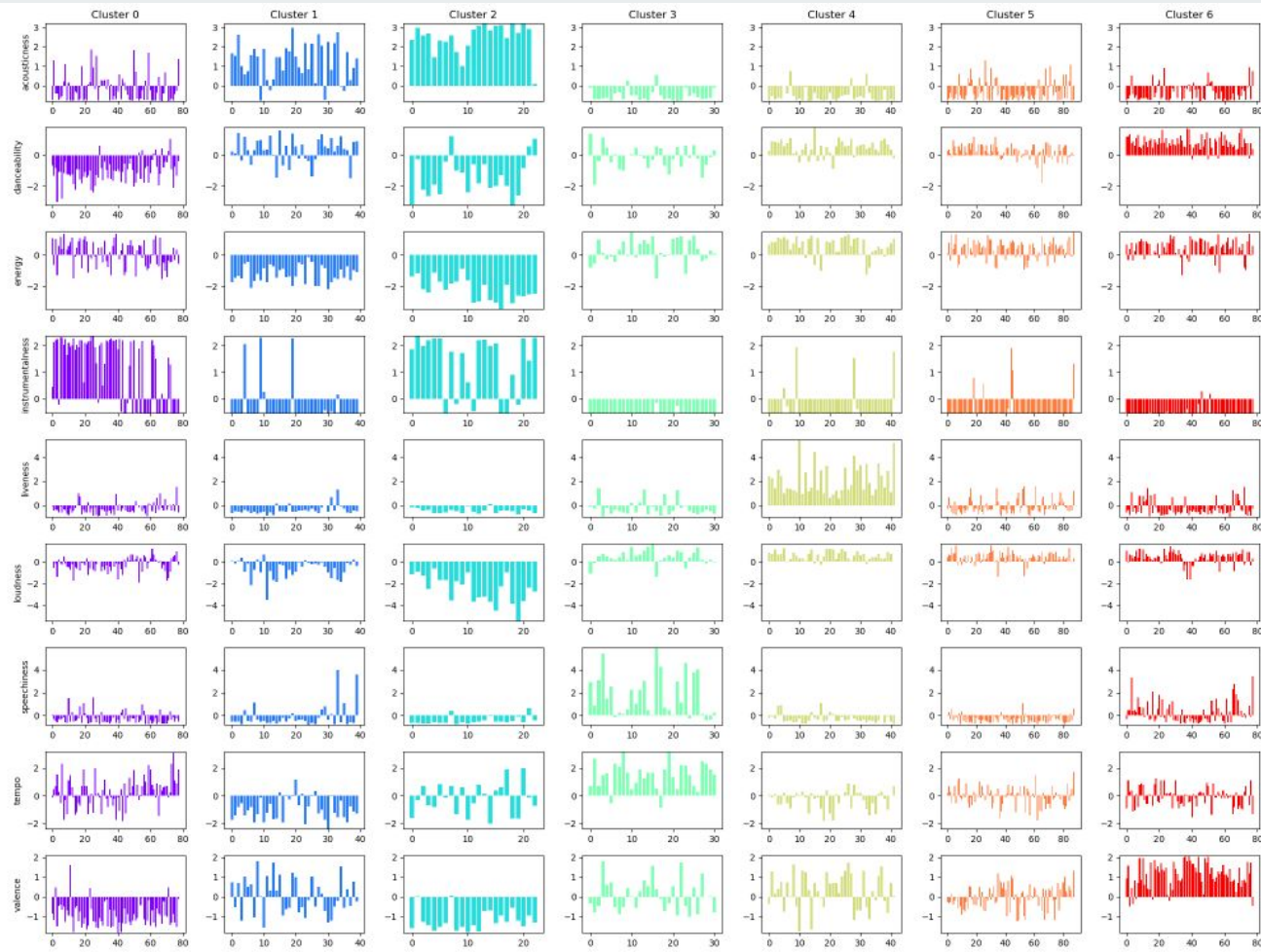


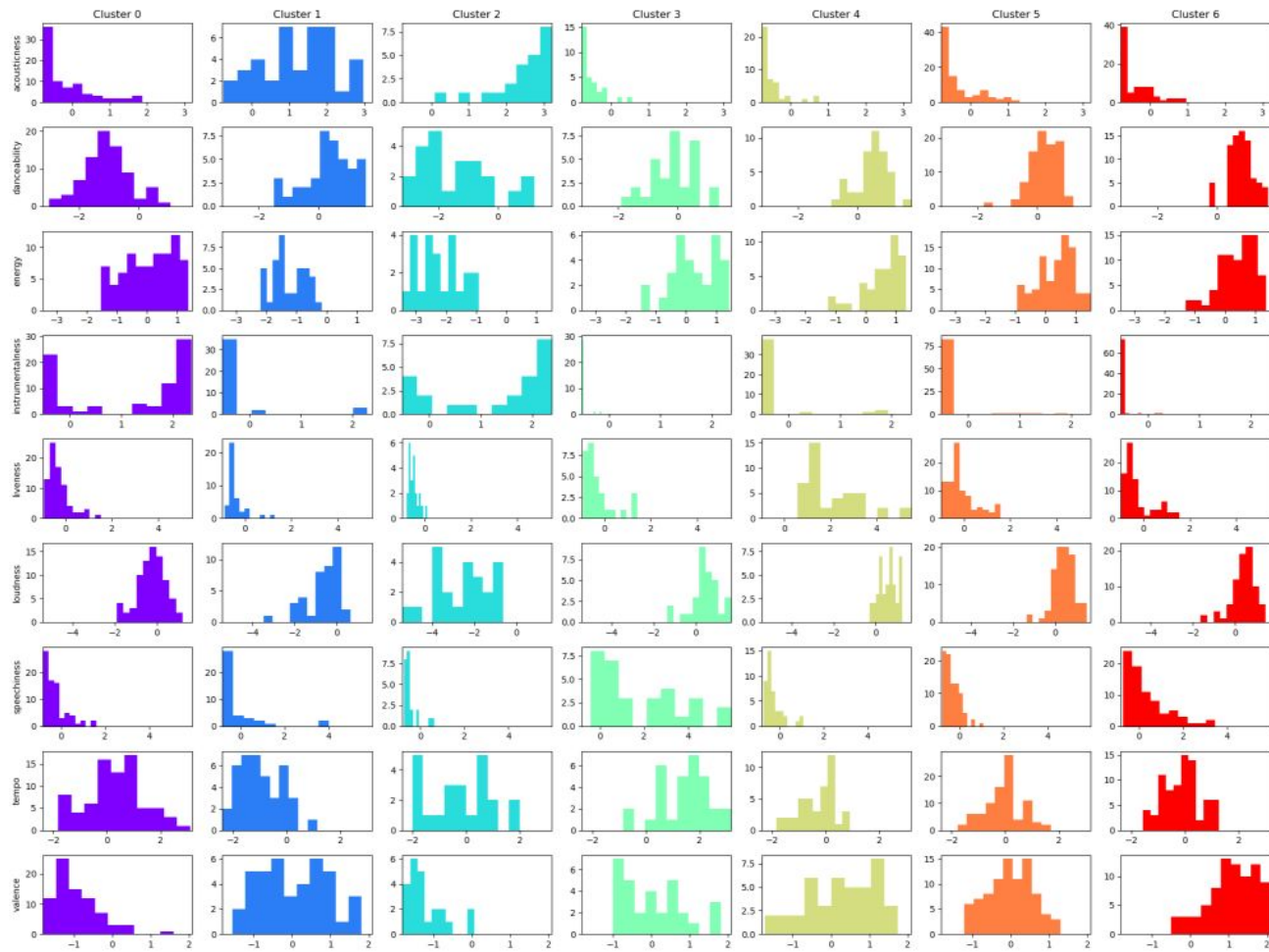
PCA - Clusters in 2D





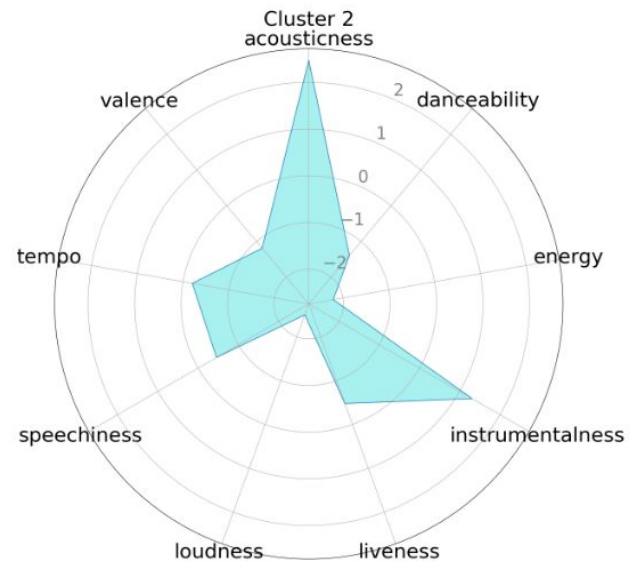
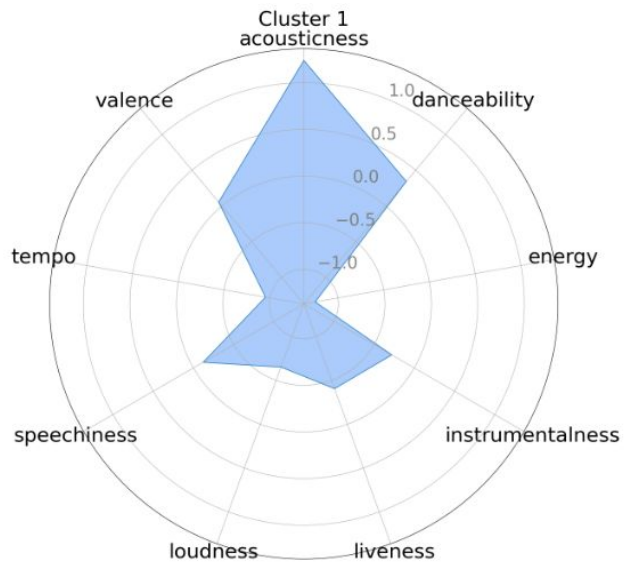
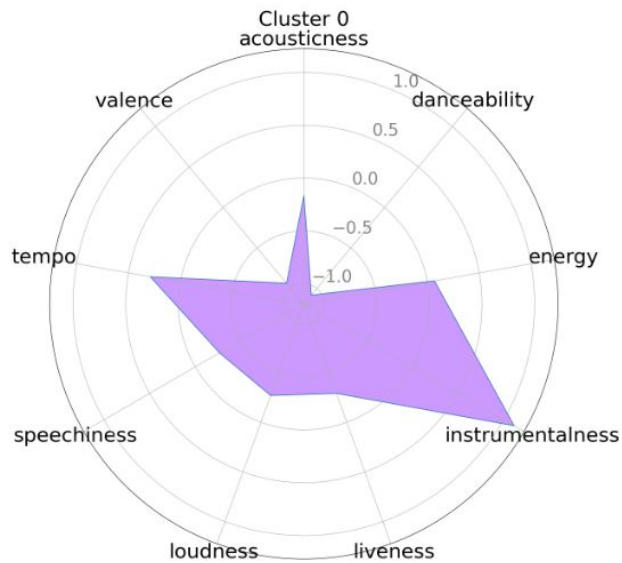
BarChart on Clusters



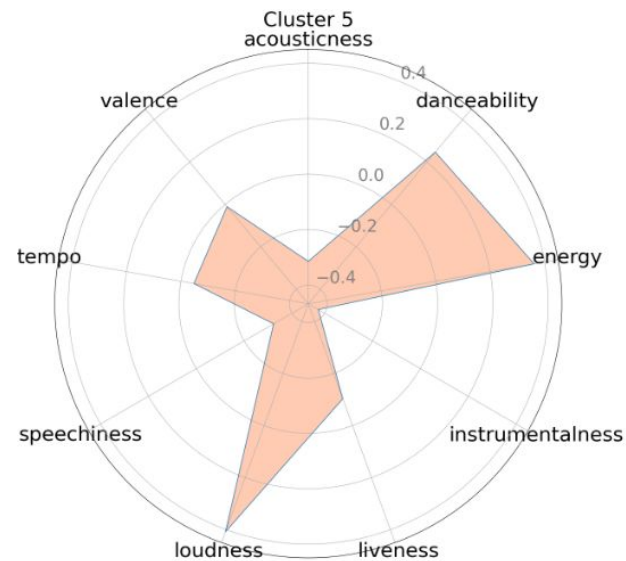
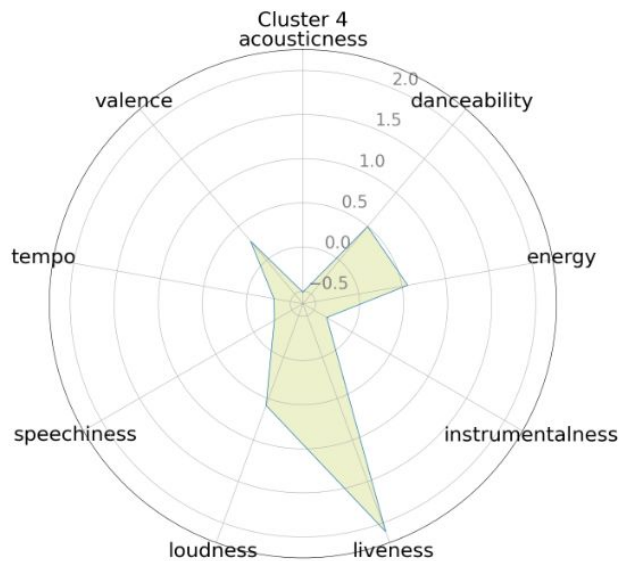
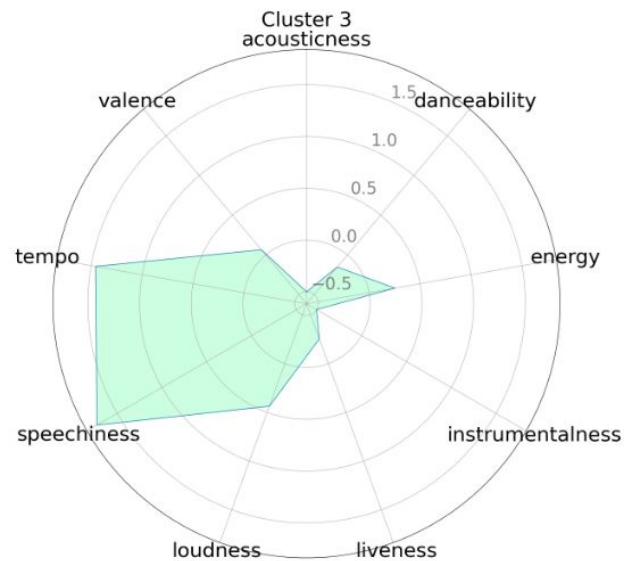




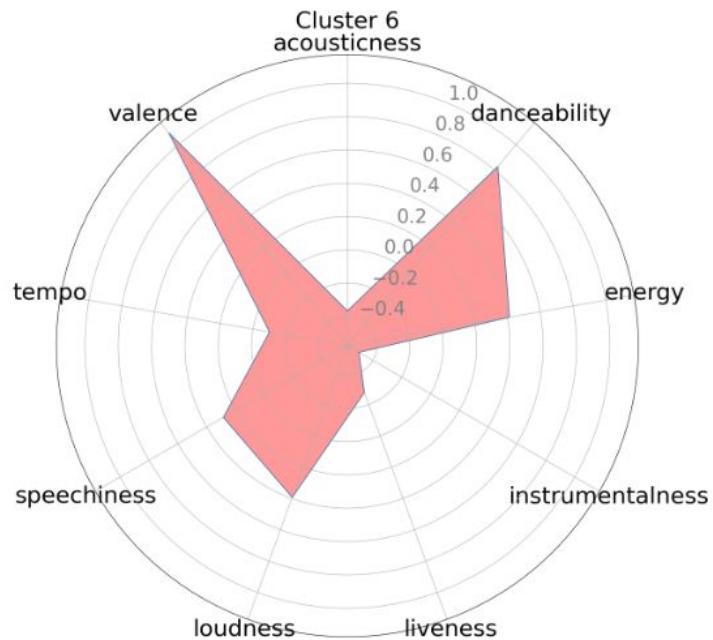
Radar Chart of Clusters



Radar Chart of Clusters



Radar Chart of Clusters





Spotify Playlist

Cluster 0 - <https://open.spotify.com/playlist/3C6QUaAMJQwwkBPZSxYMR7>

Cluster 1 - <https://open.spotify.com/playlist/0lewEz46HxvvsrqrMqvxZ0>

Cluster 2 - <https://open.spotify.com/playlist/15XASDIJAv1biPeHVqMZ8y>

Cluster 3 - <https://open.spotify.com/playlist/5cSwYfTJmdHrJtDZGrt4PQ>

Cluster 4 - <https://open.spotify.com/playlist/5RjAESsCLVwY7qE2h0gQuE>

Cluster 5 - <https://open.spotify.com/playlist/70MkJHozNry5wQM9ah5hMu>

Cluster 6 - <https://open.spotify.com/playlist/2FTjPcYvq7k4HBhz61BKkt>