



Data and EDA

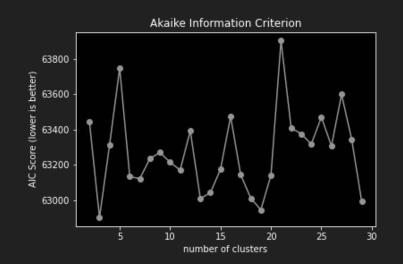




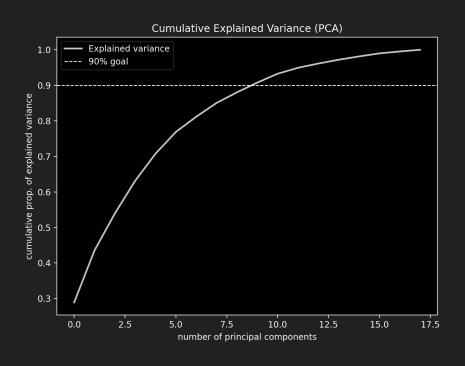
- Playlist Data 1M Playlists from Kagge (.JSON)
- Spotify API request for song analysis
 - Danceability, loudness, mode, speechiness, acousticness, instrumentalness, liveness,
 valance, tempo
 - Goal cluster playlists using only these features
 - Learn more about song analysis
- Feature Selection
 - Take mean, std of song features in each playlist
 - Initially used min, max values but removed later on due to non-normal distribution
 - Method I ultimately chose works well with normally distributed data
 - Compromise between describing playlists better, and improving cluster shapes
 - 18 total features

Selecting Best Methods

- MinMaxScaler vs StandardScaler
- PCA vs NMF
- Kmeans
- DBSCAN
- Gaussian Mixture
- Bayesian Gaussian Mixture



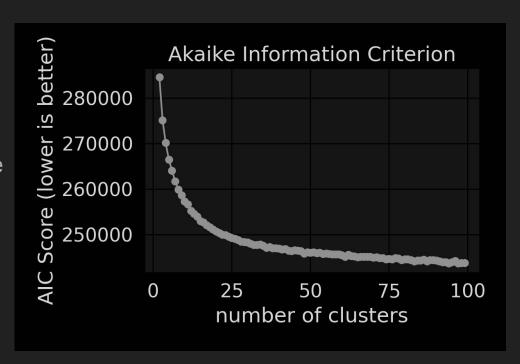




Finalization

- Gaussian Mixture Model
- Use full dataset: 70,000 playlist,6.3M songs
 - More than 5 followers, more than 20 songs
- Run at many K values and choose best (AIC Score)
- Scalability concerns O(NKD³)
- Try reducing complexity in the future
 - Fast Incremental Gaussian





Evaluating Labels

- Train at K=30, append labels
- Heatmap to evaluate labels by their features (training on all 18 features)
- Fit KNN to identify nearest neighbors to cluster centers, grouped by artist count

											_							
										egory								
0	0.18	0.7	-0.24	0.24	0.52	-0.39	1.2	1.2	-0.29	-0.35	-0.094	-0.3	0.18	0.35	0.1	-0.23	0.14	0.026
-	0.043	-0.55	0.25	-0.055	-0.74	1	-0.72	-0.77	0.92	0.33	-0.0053	-0.24	-0.24	-0.39	0.2	-0.16	0.22	-0.14
7	0.36	-0.18	0.21	-0.15	0.3	0.058	-0.68	-0.64	-0.6	-0.62	0.37	-0.072	0.3	0.035	0.31	0.98	-0.47	0.25
m	0.21	0.091	-0.16	0.28	0.62	-0.9	0.17	-0.13	-0.1	-0.4	1.2	0.47	0.067	-0.02	0.32	0.45	-0.53	0.065
4	-0.15	-1	0.36	-0.71	-1.7	1.4	-1.3	-1	1.1	1.5	-0.18	-0.28	-0.38	-0.58	-0.7	-1	0.22	-0.23
'n	0.25	0.12	-0.27	0.43	0.16	0.24	0.018	-0.27	-0.086	-0.41	-0.38	-0.37	-0.11	-0.14	0.27	0.46	0.13	-0.09
9	1.6	-1.3	1.3	-1.4	-1.1	1.3	-0.71	-0.85	1.2	2.2	0.37	-0.17	0.039	0.00092	1.7	0.0039	1.1	-0.52
7	0.55	-0.39	0.39	-0.096	0.27	0.09	-0.24	-0.5	0.9	0.19	0.88	0.18	-0.22	-0.33	0.45	-0.3	0.12	-0.16
œ	0.28	0.035	0.31	-0.62	0.59	-1.1	0.083	-0.22	0.46	0.08	2.2	2.7	-0.19	-0.29	0.45	-0.67	-0.38	-0.27
o	5.6	-1.6	11	-3.4	0.6	-0.44	-0.093	-0.43	1.6	1.1	1.9	2.1	-0.17	-0.82	1.7	-1.2	3.8	-2.6
9	0.37	-0.63	0.52	-0.67	-0.92	1.1	-0.58	-0.66	0.63	0.95	-0.52	-0.38	0.55	0.32	0.19	0.4	0.37	-0.1
=	-1.4	1.2	-0.51	0.61	-0.63	0.46	-0.91	-0.72	-0.2	0.96	0.93	-0.48	-0.59	-0.72	-2	2.6	-1.2	1.1
12	-0.5	-0.28	-0.48	0.57	0.47	-0.65	-0.18	-0.13	-1.7	-1.1	1.3	0.85	0.5	0.71	-0.2	-0.11	-1.1	0.93
g	-0.87	-1.7	-0.48	0.28	-2.3	1.7	-1.3	-1	0.25	-0.32	-0.84	-0.47	1.3	0.98	-1	-1.4	0.17	0.66
17	0.53	-3.4	3.8	-5.9	0.17	-0.44	-1.3	-0.9	-0.69		1.5	4.8	-0.98	-1.1	-2.6	-2.8	0.74	-2.4
15.	-0.46	0.68	0.26	-0.22	0.36	-0.19	1.1	1.3	0.049	0.13	-0.2	-0.28	0.76	0.82	-0.47	0.8	0.072	-0.63
16	-0.91	-0.36	-0.15	0.15	-2.1	1.7	-1.2	-0.94	0.13	0.064	-0.9	-0.47	-0.28	-0.26	-0.077	0.68	0.54	0.55
17	-0.068	1.2	-0.45	0.32	0.58	-0.59	1.3	1.6	-0.53	-0.49	-0.78	-0.46	-0.03	0.13	0.037	-0.3	0.096	0.38
8	-0.72	1.3	-0.11	0.22	0.62	-0.84	0.57	1.1	-0.71	-0.47	-0.6	-0.42	0.18	0.11	-0.54	1.3	0.29	-1.1
13	-0.06	-0.24	0.36	-0.59	-1.9	0.37	-0.28	-0.31	0.24	1.2	0.31	0.1	0.51	0.92	-0.53	0.76	0.032	-0.016
20	-0.002	0.31	-0.28	0.14	0.57	-1.5	0.4	0.22	0.036	-0.078	-0.15	-0.26	-0.28	-0.28	-0.11	-0.23	-0.092	-0.45
21	1.4	-1.6	2.3	-2.3	-0.0047	0.16	-0.8	-0.75	1.2	1.8	2.2		-0.57	-0.66	0.14	-1.3	0.34	-0.54
22	0.58	-0.6	0.45	-0.32	-0.085	0.53	-0.46	-0.6	0.65	0.087	1.4	0.72	0.055	-0.054	0.62	0.1	0.013	0.015
23	0.65	-1.1	1.3	-1.2	-0.87	1.1	-1.1	-0.92	1.3	1.8	1.3	0.4	-0.49	-0.54	0.32	-0.57	0.33	-0.43
54	-0.1	0.5	-0.63	0.64	0.59	-0.6	0.35	0.1	-0.48	-0.55	-0.47	-0.4	-0.077	-0.079	0.087	0.21	-0.29	-0.2
52	-0.31	0.55	-0.62	0.51	0.43	-0.29	0.45	0.54	-0.36	-0.4	-1	-0.49	-0.076	0.083	-0.15	0.17	0.055	0.045
56	0.59	-0.53	0.96	-0.72	0.56	-0.4	-0.64	-0.66	1.1	0.94	0.71	0.25	-0.076	-0.22	0.36	-0.15	0.12	-0.28
27	-0.38	-0.77	-0.68	0.82	-0.33	0.58	-0.56	-0.49	-1	-0.95	0.48	-0.39	0.098	0.28	-0.073	0.2	0.25	0.8
78	0.39	-0.54	0.37	-0.47	-0.44	0.83	-0.76	-0.76	-0.22	-0.33	0.063	-0.2	0.21	-0.035	0.42	1	-0.14	0.18
59	0.57	-2.4	2.2	-4.8	-0.23	0.55	-1.3	-0.87	-1.2	3.9	1.4	5.1	-1.3	-1.1	0.71	-1.5	0.47	-1.8
	('danceability', 'std')	('danceability', 'mean')	('loudness', 'std')	('loudness', 'mean')	('mode', 'std')	('mode', 'mean')	('speechiness', 'std')	('speechiness', 'mean')	('acousticness', 'std')	('acousticness', 'mean')	('instrumentalness', 'std')	'instrumentainess', 'mean')	('liveness', 'std')	('liveness', 'mean')	('valence', 'std')	('valence', 'mean')	('tempo', 'std')	('tempo', 'mean')

Mike Birbiglia John Mulaney Nick Swardson Patrice O'Neal Ron White Dave Attell Joe Rogan Nine Inch Nails Todd Snider Matthew Broussard Christopher Titus

Kanve West Rae Sremmurd Big Sean Kodak Black A\$AP Rocky Lil Wayne Travis Scot Fetty Wap Mac Miller The Weeknd A\$AP Ferq Chris Brown Kid Ink MadeinTY0

Five Finger Death Punch Papa Roach blink-182 Rise Against Jimi Hendrix Rage Against The Machine Black Sabbath The White Stripes

Hillsong Worship tobyMac Tenth Avenue North Jeremy Camp Hillsong Young & Free for KING & COUNTRY David Crowder Band Matt Maher Big Daddy Weave Hawk Nelson The Afters

The Black Eved Peas Hannah Montana Lil Wavne Lady Gaga Trey Songz JAY Z

Italian Mandolin Torna A Surriento Jazz Band Piano Blues Italian Restaurant Music of Italy Otis Redding Sam Cooke Four Tops Frank Sinatra Kav Starr Young Thug The Supremes

Ben Rector

Lana Del Rey





- Input Spotify Playlist URL and locate nearest neighbors
- Discover new music!

neighbors= new playlist('https://open.spotify.com/playlist/6TeyryiZ2UEf3CbLXyztFA', playlist analysis, playlist master)



album_name	track_name	artist_name
	Stranglehold	
		Guns N' Roses



Dilink-18/2
The Boors
A Boogle Wit da Hoo
Lynyrd Skynyrd
Usher
Ozzy Osbourne
Free Hold Hold Hold
Free Hold



Dad's Playlist Life Dad playlist In the Zone psych

Questions?