Moosic

SPOTIFY PLAYLIST CREATION

By Group 1
Jan Tittel
Marcus Hoffmann
Michael Los
Sudharshana Ramalingam

Answering Business Questions

How did you create your prototype?

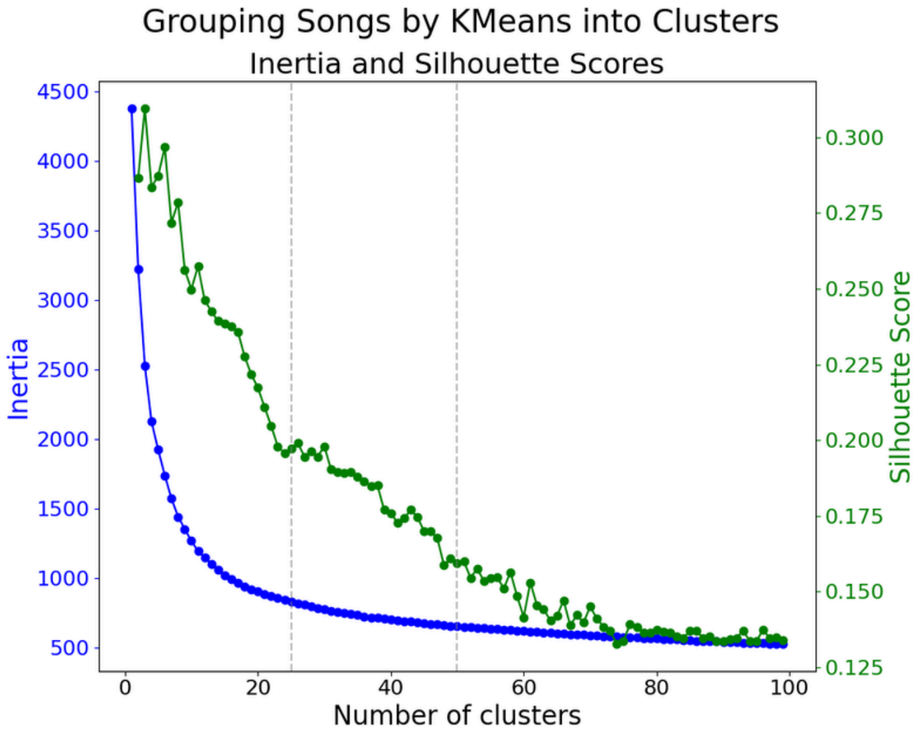
Clustering algorithms, KMeans Clustering and Reclustering

What audio features did you use and what did you drop? Why?

Dropped: type, key, duration_ms, time_signature

How many playlists (clusters) are there?

Mathematical approach --> ~12
Business approach --> 25 / 50



Answering Business Questions

Is the prototype effective at creating cohesive playlists?

The prototype creates some smaller playlists, that are somewhat cohesive (up to 50 tracks)
The larger playlists (100+ tracks) are very mixed and incohesive

Are Spotify's audio features capable of identifying "similar songs" as defined by humanly detectable criteria?

Yes they are, because Spotify Playlists are very cohesive. Given the approach that we used, it did not work so well. They apparently use more sophisticated clustering methods - which is to be expected.

What kind of data might help us create better playlists?

Genre, Language, Region, Year of Release, Listener Behavoir, Appearance in Playlists

Answering Business Questions

Is K-Means a good method for creating playlists?

Since we use a rather “subjective” categorization, utilizing euclidian distances seems to be limited in it’s precision. It might be more useful for numeric values like salaries, prices and actual distances

Pros

- Simple and Efficient
- Scalable

Cons

- Subjective Choice of clusters
- Disproportionate Effect of Outliers --> Threshold reclustering

What would be your next steps if you continued with this project

- Try other clustering methods, DBSCAN, Gaussian Mixture
- Use ML with Labeled data, like genre, popularity or ratings

Small Cluster

Playlist

Sad Sundays

Playlist obtained via KMeans clustering

WBS Allstars • 26 songs, about 1 hr 45 min

#	Title	Album	Date added
1	Se Eu Quiser Falar Com Deus Gilberto Gil	Eletracústico	8 minutes ago
2	Rebel Rebel Seu Jorge	The Life Aquatic Featuring Seu Jorge	8 minutes ago
3	Rebel Rebel Seu Jorge	The Life Aquatic with Steve Zissou (Original...	8 minutes ago
4	Samba Triste Stan Getz, Charlie Byrd	Stan Getz: The Bossa Nova Albums	8 minutes ago
5	Fotografia Vanessa Da Mata	Vanessa da Mata canta Tom Jobim (Deluxe ...	8 minutes ago
6	Diz Que Eu Fui Por Aí Eva Fernández, David Pastor, Josep Lluís Guart, Miquel À...	That Darkness	8 minutes ago
7	Nega Maluca / Billie Jean / Eleanor Rigby Caetano Veloso	Caetano Veloso	8 minutes ago
8	Se eu quiser falar com Deus - Ao vivo Gilberto Gil	Eletracústico (Ao vivo)	8 minutes ago
9	Koop Island Blues Koop, Ane Brun	Koop Islands	8 minutes ago
10	Mort d'un heroi romàntic Manel	Atletes, baixin de l'escenari	8 minutes ago
11	Limit To Your Love James Blake	Limit To Your Love	8 minutes ago

Large Cluster

Playlist

Ultimate Mood Swings

Playlist obtained via KMeans clustering

WBS Allstars • 246 songs, about 18 hr

#	Title	Album	Date added
1	La Nueva Reconquista de Graná Grupo de expertos Solynieve	El Eje de la Tierra	6 minutes ago
2	Just Like Honey The Jesus and Mary Chain	Psychocandy (Expanded Version)	6 minutes ago
3	Nothing To Be Done The Pastels	Truckload Of Trouble	6 minutes ago
4	Waterfall The Fresh & Onlys	Play It Strange	6 minutes ago
5	La Pedregada Univers	La Pedregada	6 minutes ago
6	Turn Up The Speakers - Original Mix AFROJACK, Martin Garrix	Turn Up The Speakers (Original Mix)	6 minutes ago
7	Sending My Love - Afrojack Edit R3HAB, Swanky Tunes, Hard Rock Sofa, AFROJACK	Sending My Love (The Remixes)	6 minutes ago
200	Where the Rivers of the Madness Stream Cemetary	An Evil Shade of Grey	35 minutes ago
201	Storm Of Stress Terrorizer	World Downfall	35 minutes ago
202	Forty Seconds Bloodbath E Vomitory	Redemption (Bonus Edition)	35 minutes ago

Using Song Genres for Clustering

1. Obtain genres (of artists) via Spotify API:

['album rock', 'classic rock', **'rock'**, 'soft rock', 'yacht rock']

['adult standards', 'easy listening', 'lounge', 'vocal **jazz**']

2. Determine most likely genre from given list of genres.

3. Perform Clustering (on audio features) for each Genre separately

genre	count
death metal	690
brutal death metal	556
pop	548
metal	442
deathgrind	384
edm	372
technical death metal	357
pop dance	342
dance pop	319
black metal	303
thrash metal	263
reggaeton	261
rock	234
electro house	233
urbano latino	231
classical	229
soft rock	199
progressive electro house	196
florida death metal	196
german pop	193

Tfidf-Vectorization

+ KMeans Clustering

genre_cluster_label	count
Metal	1054
Others	1033
Pop	570
Classical	349
Electronic dance music	344
Rock	333
Hip Hop	260
Jazz	257
Raggeaton	256
Rock'N'Roll	143
Soul	117
New Wave Pop	116
Música popular brasileira	91
Reggae	80

KMeans Clustering

Rock Vol. 1

Rock Vol. 2

Rock Vol. 3

Soul