



Increasing Audio Streaming Engagement via Music Recommendation Systems

Team Land'o'Datalakes:

Alistair Marsden, Eduard Stalmakov, Olivier Rochaix, Vanessa Gleason

Introduction

Why song recommenders?

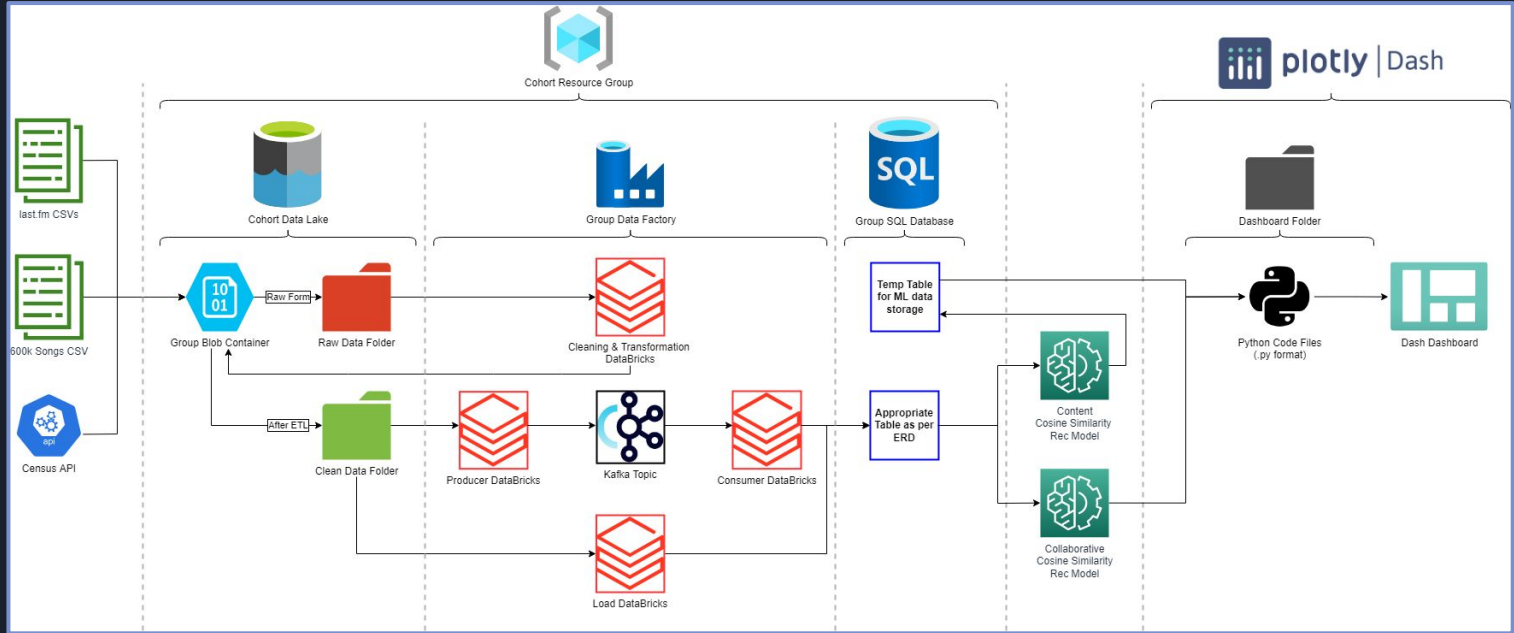
Exploratory questions:

- Who are the people who use audio streaming services?
- How might we build a database that allows us to make recommendations based on music?
- How would songs trend over time in this database?
- What trends or patterns are there in the user metrics of our application?
- How does our database compare to actual music streaming services?
- Are recommendation models important to user retention on a music platform?



Source: <https://www.npr.org/sections/allsongs/2013/07/10/200788243/the-good-listener-when-you-make-someone-a-mix-what-do-they-owe-you>

Data Platform



ETL Pipeline



Audio Streaming Market

Age:

23-30 years old

Location:

Major cities within highly populated states

Income Level:

Middle and Upper Middle Class

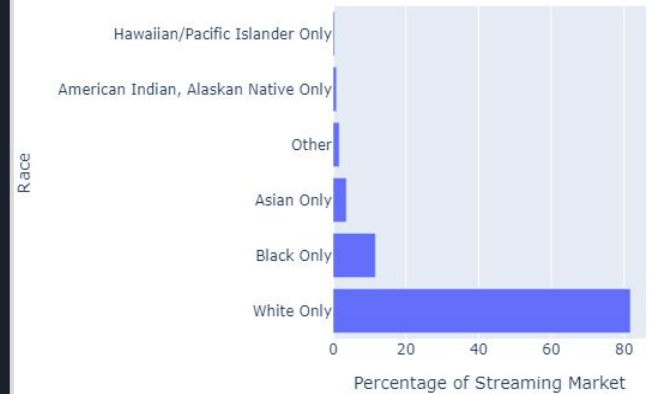
Race:

White

Sex:

Males and females evenly

Music Streaming by Race

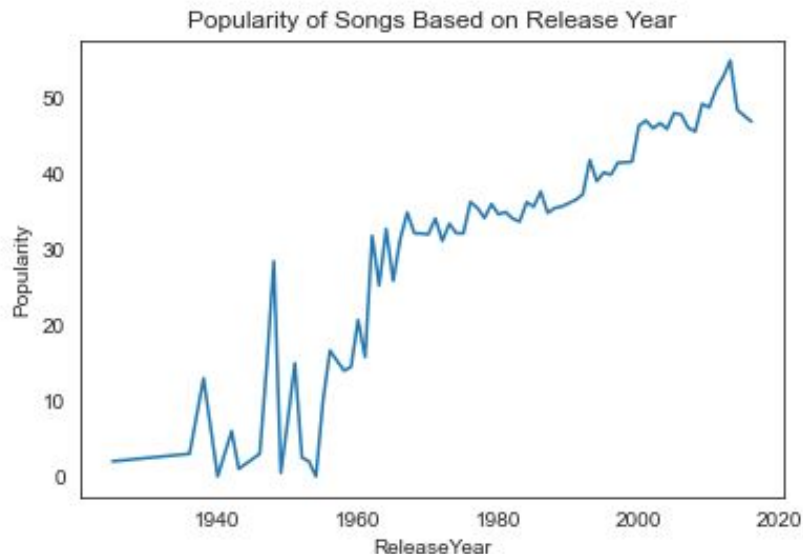


Trends of Songs in our Library

Popularity calculated based on the number total number of track plays and how recent those plays were

General increases with release year

Spike in popularity for songs released around 1950



What trends or patterns are there in the user metrics of music applications?

8000+ Tracks

3000+ Artists

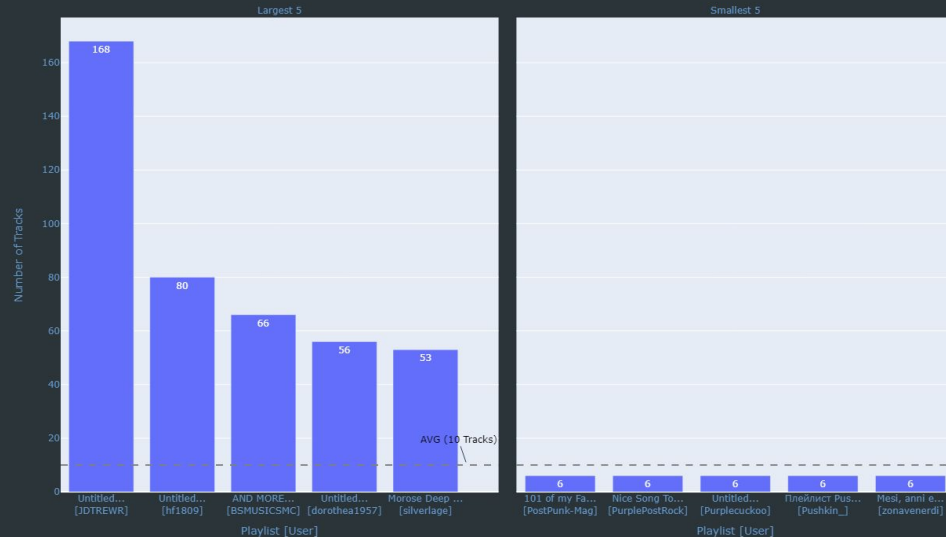
1-67 Tracks to an Artist

1400+ Users

2600+ Playlists

1-73 Playlists to a User

PLAYLIST SIZES BY NUMBER OF TRACKS



How does our database compare to actual music streaming services?

In the year 2022, Spotify has “82 million songs” and “11 million artists” (Ruby, 2022).

Without proprietary information, there is no way to ensure that our database is a representative sample.



This means that despite using songs and tracks from Spotify, we cannot make substantiated claims about their platform off of our data.

We can, however, make a realistic recommender using that data that can give us insight into how they work and can be properly scaled should new data be introduced.

Importance of Recommenders in User Retention

Maximize user engagement

“Consumers will get themselves more engaged in a music streaming service when **personalized music recommendations** are made to them.”
(Culture Next 2022)

Easy music discoverability

53% said they've **sought more content** from more diverse creators and podcasts in the past year (Culture Next 2021)

Increased customer satisfaction

“69% of American Gen Zs feel ‘more centered and **generally happier**’ when listening to their **favorite music** on a daily basis” (Culture Next 2022)



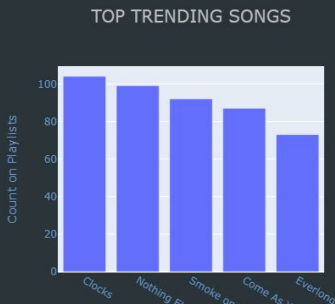
Source: <https://happinesson.com/how-to-be-happy-by-listening-to-music/>

Types of Recommenders

Non-personalized

Trending songs, playlists, artists

Does not require streaming platform to know anything specific about the user



Personalized

Need to know listening history of the user but more personalized

Content based

What are the attributes of the user's songs?
What songs have similar attributes?

Collaborative based

When users listen to a song, the model searches for playlists that have that song. The songs that appear most often in playlists with the user's song are returned as recommendations

Content Based Recommender

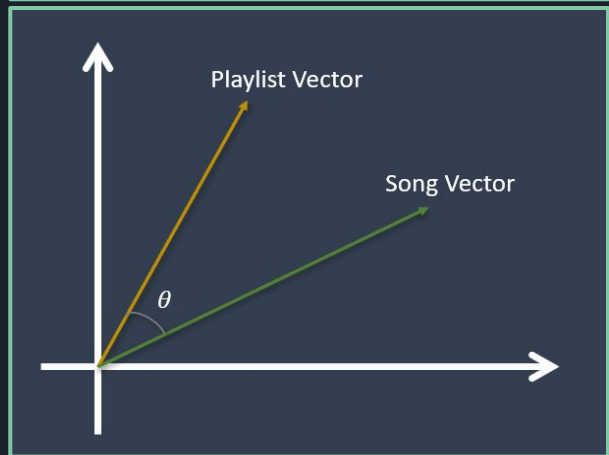
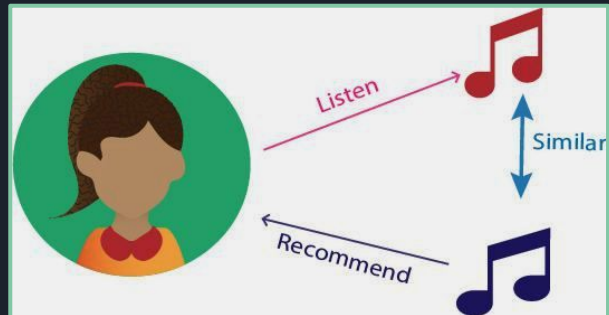
Create a vector for each song within our database

- Create dummy variables for categorical columns
- Scale numerical columns to a range between 0-1
- Further scale our features based on what attributes were more important

Create a vector for a selected playlist

Calculate the cosine similarity

Recommend the songs with the greatest cosine similarity



Collaborative Recommender

Pivot table with the index being the TrackID, and the columns being PlaylistID

- 1 if TrackID is in PlaylistID
- 0 if not

Convert to a sparse matrix (bottom right)

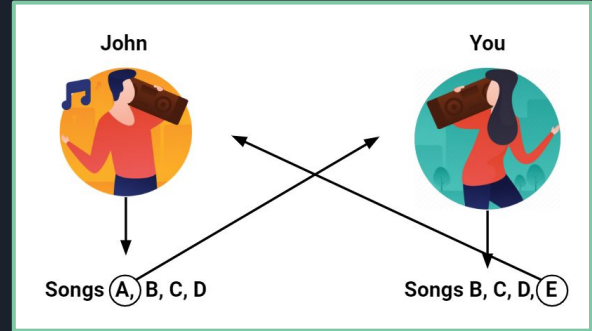
Calculate the cosine distance (-1 to 1)

Convert to pairwise distance (0 to 1)

- 0 is most similar

Returns a distance matrix

When user searches a song, the songs closest to 0 are returned



Source: <https://www.univ.ai/post/spotify-recommendations>

Sparse Matrix sample

(0, 2643)	1.0
(1, 438)	1.0
(1, 2034)	1.0

Dashboard

Music Streaming Song Recommender

User InterfaceTarget MarketAboutBehind the Scenes

Welcome LOVES-DESIRE

Choose a playlist to get song recommendations in real time!

BEYONCE

TrackName	ArtistName
Sirena	Sin Bandera
Burning	Green Day
Get Free	The Vines
V.I.P.	Shaun Baker
Truth	Dwele
Rigia	Sin Bandera
Nacy's Day Parade	Green Day
Superhéroe	Alexis y Fido
Edovexa	Lumen
It's Over Now	112

Get recommendations based on what other users saved to their playlists:

Type a Song Name

Deeper and Deeper

Submit

Enter a value and press submit

If you like Deeper and Deeper (Madonna), you should try:

Song_Artist	Percent_Similar
Rescue Me (Madonna)	55%
Fever (Madonna)	48%
Dress You Up (Madonna)	45%
Bad Girl (Madonna)	42%
To Have and Not to Hold (Madonna)	39%

TOP TRENDING SONGS

Song	Count on Playlists
Chick	100
Nothing But Sadness	95
Shine on Me Sister	90
Love to the Max	85
Waiting	80

TOP TRENDING GENRES

Genre	Count on Playlists
Shoebiz Rock	3000
Alternative Metal	2500
Rock Pop	1800
Alternative Rock	1800
Hard Alternative	1200

TOP TRENDING ARTISTS

Artist	Count on Playlists
Madonna	450
The Cure	420
Goldie	400
Red Hot Chili Peppers	380
Green Day	350

Select User

Pick one user to view your personalized dashboard

LOVES-DESIRE --

Content Recommender

Collaborative Recommender

Non Personalized Recommender



Datasets

Boland, D. (n.d.). Streamable Playlists with User Data. Retrieved October 4, 2022, from <http://www.dcs.gla.ac.uk/~daniel/spud/>

Eren Ay, Y. (2021, April). Spotify Dataset 1921-2020, 600k+ Tracks. Kaggle. Retrieved September 28, 2022, from <https://www.kaggle.com/datasets/yamaerenay/spotify-dataset-19212020-600k-tracks>

United States Census Bureau. (2020, June 22). Current Population Survey: Computer and Internet Use Supplement 2019. Retrieved September 22, 2022, from <http://api.census.gov/data/2019/cps/internet/nov>

References

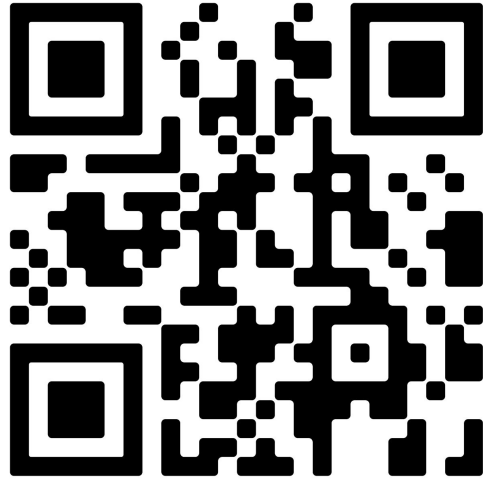
Jena, A. (2022, April 18). Music Recommendation System – Revolutionizing the Music Streaming Industry. Muvi. Retrieved October 5, 2022, from <https://www.muvi.com/blogs/music-recommendation-system.html>

Spotify. (n.d.). Culture Next: 2021 | Spotify Advertising. Retrieved October 5, 2022, from <https://culturenext.byspotify.com/en-US>

Spotify. (n.d.). Culture Next Global Trends Report 2022. Culture Next: 2022 | Spotify Advertising. Retrieved October 4, 2022, from <https://culturenext2022.byspotify.com/en-US>



Learn more about our project



Github

Thank you for your time!

Questions?

