



Spotify Playlist Generator

a more immersive way to generate playlists
MARK YUNG

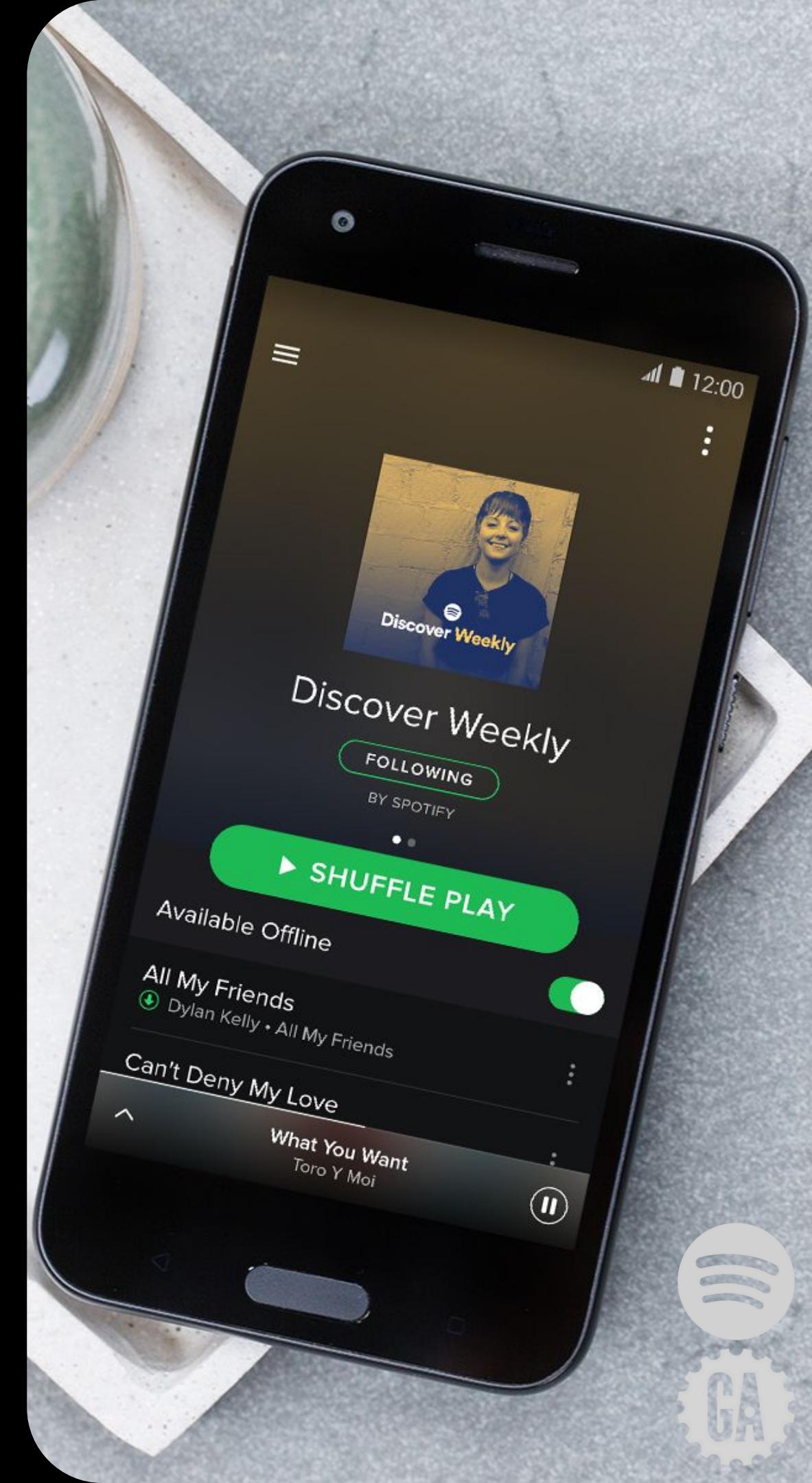


The Spotify Issue

Spotify already has a robust recommendation system, and this usually comes in the form of procedurally generated, user-tailored playlists to a person's taste.

But this comes with a catch.

Spotify likes to recommend you songs that are new, and keep you listening and discovering new music. While this is a great way to find 'new digs', it does not allow you to intuitively create playlists from a *known* pool of songs - say for example from your saved songs.



A black and white photograph of a person from behind, wearing a textured hat and a dark jacket, playing an acoustic guitar. The person is positioned on the left side of the frame, with the guitar's headstock visible.

A little about why this **matters**



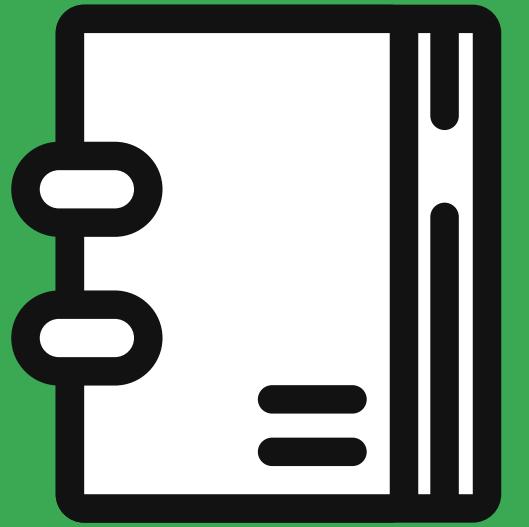
Table of Contents



POINTS FOR DISCUSSION:

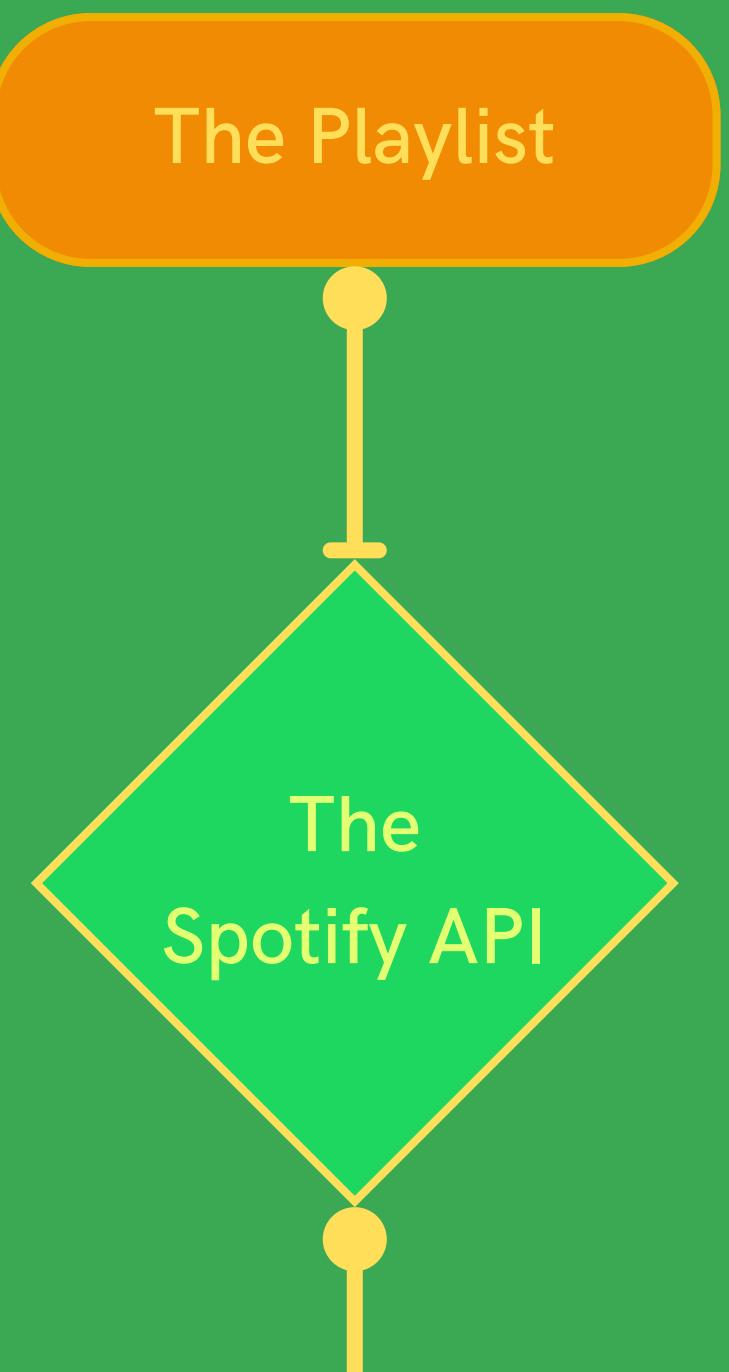
- The Data - The Spotify API and Playlist Scraping
- EDA and Data Processing Concerns
- Feature Engineering
- Modeling and Considerations
- Deployment
- Limitations and Future Developments





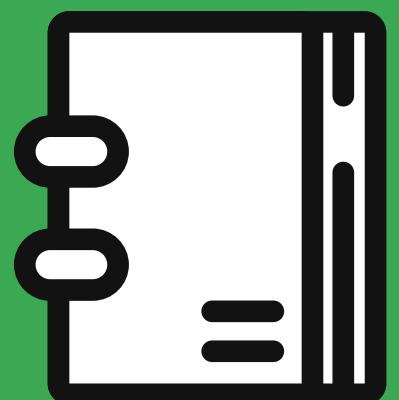
Section 1: The Data

Scraping Tracks through Spotify's API



Dataset Information

Information pulled from Spotify API endpoints



Track Details

Name, Album, Artists, Popularity Index, URIs, etc.

Artist Information

Related Genres

Audio Meta Features

Mood Vectors, Liveness, Loudness, Speechiness, Acousticness, etc.

Song Features

Key, Mode, Time Signature, Duration, etc.

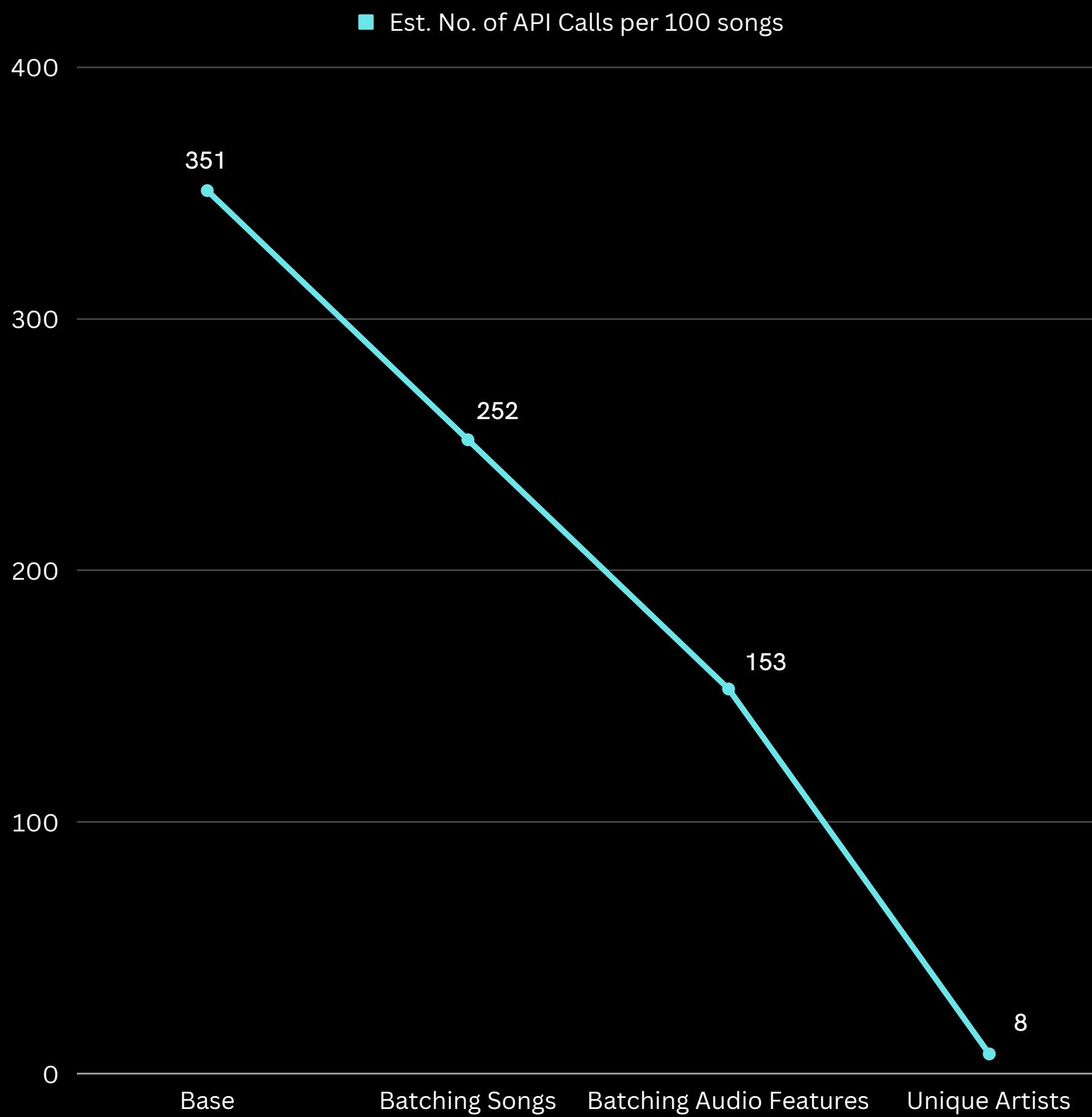


API Scraping Optimisations

Limited by Spotify's request limits and rate limits

Prioritize batch requests to optimize both speed and usability of API

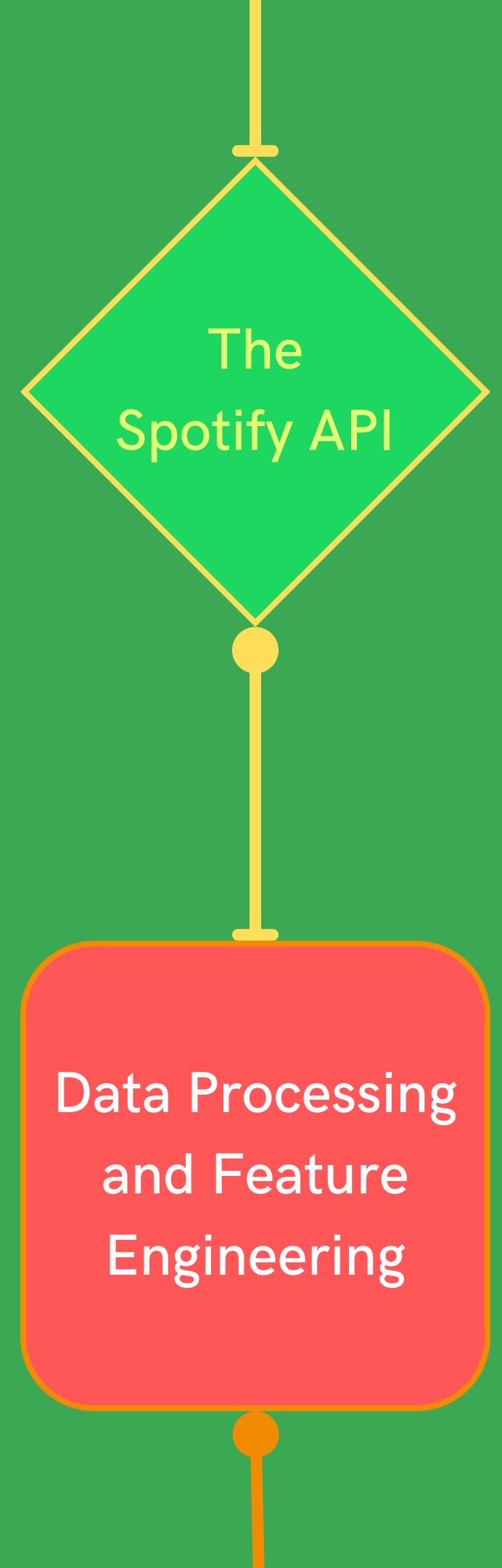
1. Batch-pulling tracks
2. Batch-pulling audio features
3. Batch-pulling Artists by using list of unique artists





Section 2: EDA and Data Processing

Looking at the Meta Features



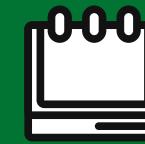


Spotify Meta Features



DATA PROCESSING

Formatting the playlist
metadata for similarity solving



1

Playlist



2

Remove True
Duplicates



5

To Feature
Engineering



4

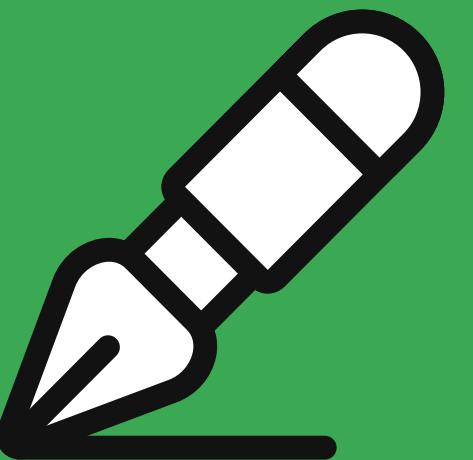
Standard Scaling
Numerical
Features



3

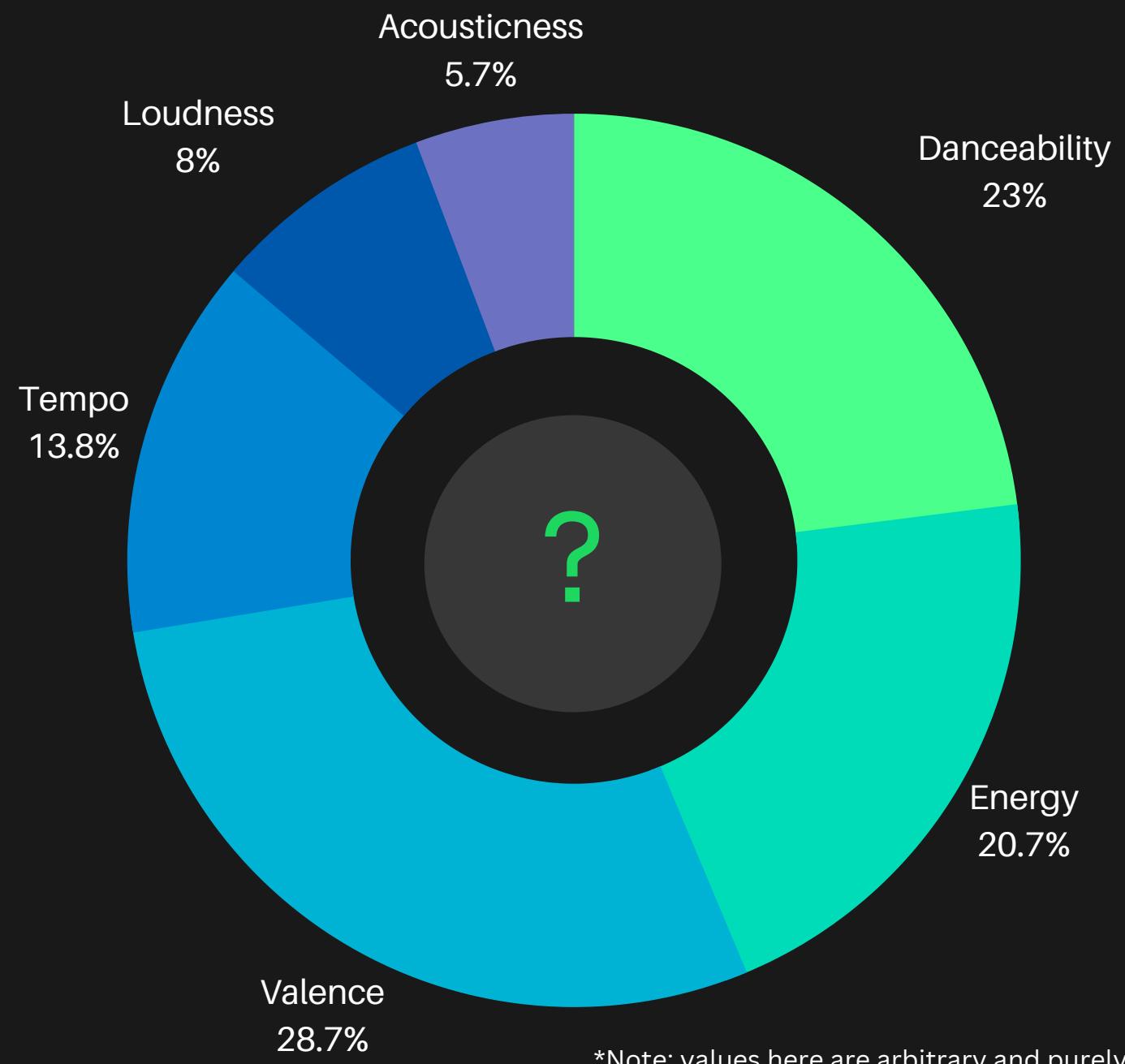
Removing
Unnecessary
Columns





Section 3:
**Feature
Engineering**
Creating Custom Mood Vectors

Data Processing
and Feature
Engineering



Meta Features

The metrics given to us by Spotify describe rather abstractly singular attributes about each track, but do little to describe the subjective 'characteristics' of a song.



Creating Mood Vectors

SOMETIMES WE NEED SUBJECTIVITY TO BE ACCURATE



'Oomph' or Impact

Tries to describe the 'punch' of a song.

Combines energy, acousticness, loudness, and tempo.



Hype

Attempts to be a more holistic measure of energy.

Incorporates danceability, energy, valence, tempo, and loudness.



Vibes

How much you feel the music's energy.

Mostly danceability, also considers energy and valence, and an inverse relationship to loudness





Section 4: **Modeling**

Building a Content-Based Filtering
System

Output

CREATING A TARGET VECTOR



Target Vector from User Inputs



Normalize vectors using means and st.dev
of respective columns in scaled dataframe



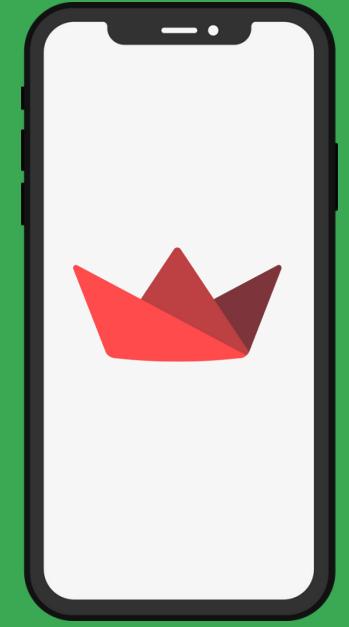
Modeling Considerations
Distance or Similarity Metrics



Cosine Similarity

Using Cosine Similarity allows us to measure
the angular similarity between points to
determine what to recommend





Section 5: **Deployment**

**Creating a Streamlit App
(Cloud Deployment Coming Soon)**

THE SPOTIFY PLAYLIST GENERATOR

An Intuitive Deployment Flow

Select Playlist or
Use Default

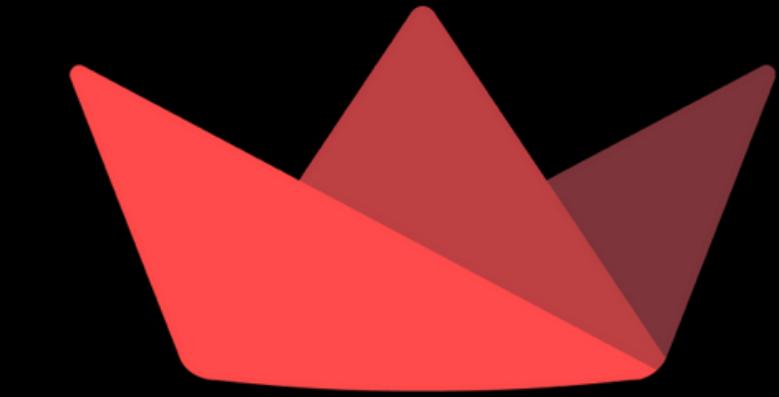
Choose Number of
Songs to Pool From

Adjust Meta-Features,
and Select Number of
Songs in New Playlist

Run the Recommender
to Get New Playlist

Upload Your Brand New
Playlist!





App Demo





Section 6: **Limitations** And Future Considerations



The Limitations of Music Recommendations

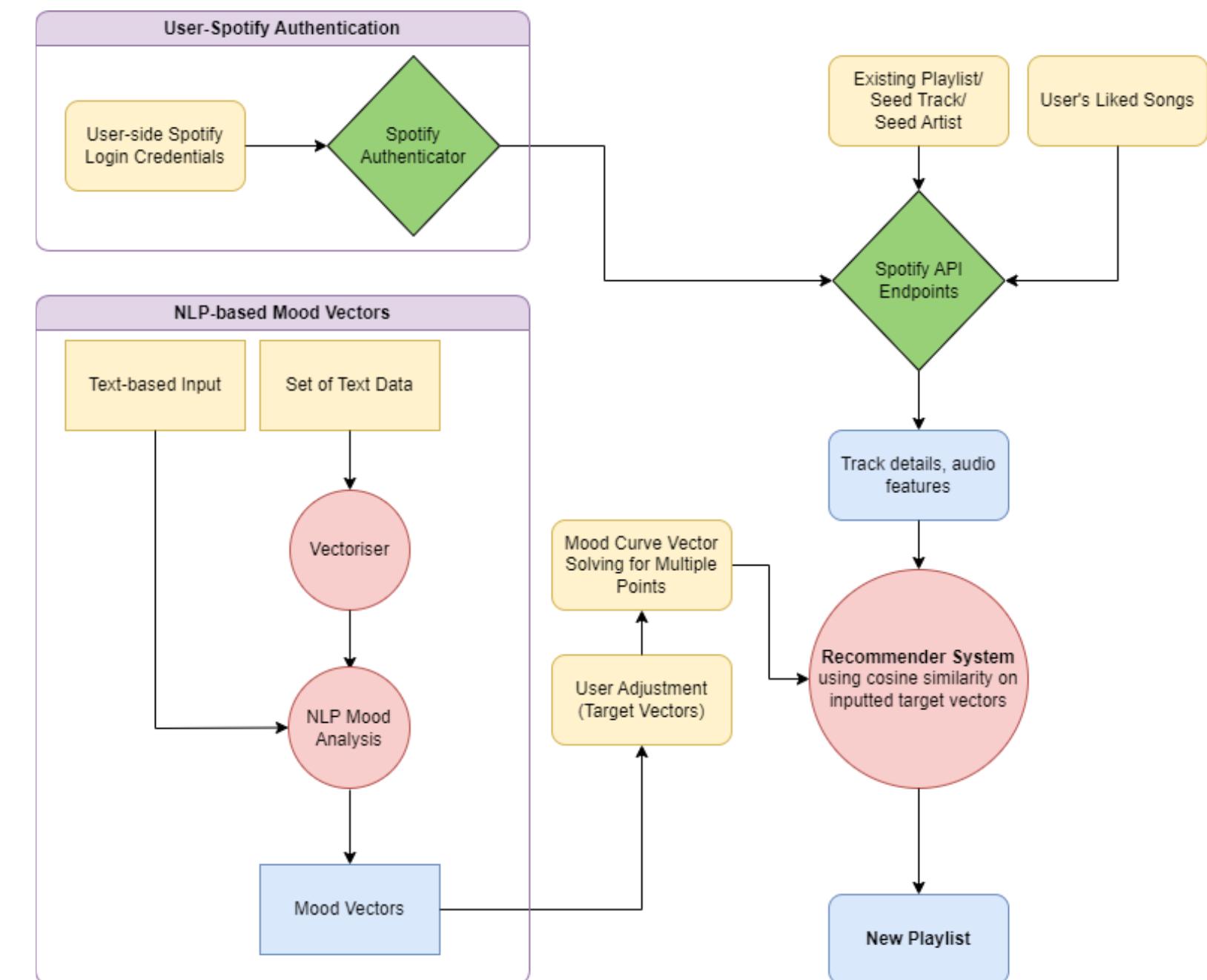
Music is inherently a personal experience, and everyone's definition of good music varies to some degree.

What makes the result 'good' is highly subjective, and may need some additional user-based tuning to get it just right for recommendations



Future Upgrades and Considerations...

- Expanding Seed Source to Spotify's in-built seeding system, and from user's liked songs
- Mood-based setlists with multi-point similarity solving, and inclusion of genres for interpolation of intermediate songs
- Incorporating other ML models to factor in distance-based clustering/similarity
- Text-based Sentiment Solving for multipoint similarity solving





Thank You!

If you have any questions
please feel free to ask.



: <https://github.com/mywk94>



: <https://www.linkedin.com/in/mywk>

