



HOME



YOUR TOP SONGS



PLAYLIST

01

Risk

02

Borderline

03

*Save Your
Tears*

04

*Bombay
Rhapsody*



Spotify Streaming Analytics Dashboard

By: Aanvi Goel



Data, Data, Data
Hans Zimmer



2:54

3:49

Overview

- **Spotify** is a Swedish audio streaming company that has taken over globally, with 33 million monthly active users, including 188 million paying subscribers, as of June 2022
- Building a **dashboard** to visualize the audio trends on Spotify would help music enthusiasts, audiophiles and music creators gain insights to latest trends in the music industry



Data, Data, Data
Hans Zimmer

2:54



3:49



Table of contents

01

**Data
Ingestion**

03

Processing

02

Data Storage

04

Deployment



Data, Data, Data
Hans Zimmer

2:54



3:49



Data Pipeline

Data Ingestion

Spotify Audio
Features dataset
with 1.2M+ song
data.
File format: .csv



Data Storage

Load data into a
MongoDB
collection



Processing

Clean data using
Python and
pandas. Perform
exploratory data
analysis to identify
global streaming
trends



Deployment

Visualize the data
and create a
user-friendly
dashboard to be
hosted on the web
via Streamlit +
StreamlitCloud



Data, Data, Data
Hans Zimmer

2:54



3:49



HOME



TABLE OF CONTENTS



PLAYLIST

01

**Data
Ingestion**

02

Data
Storage

03

Processing

04

Deployment



Data Ingestion

- ❖ **Dataset: Spotify 1.2+M Songs** dataset from Kaggle
 - Format : **.csv**
 - 1.2+ M data points for Spotify streamed songs
- ❖ Web API: Spotify Web API through **spotipy** library



Data, Data, Data
Hans Zimmer

2:54



3:49

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1173210 entries, 0 to 1204024
Data columns (total 20 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   id                    1173210 non-null object
 1   name                  1173210 non-null object
 2   album                 1173210 non-null object
 3   artists               1173210 non-null object
 4   artist_ids            1173210 non-null object
 5   explicit              1173210 non-null object
 6   danceability          1173210 non-null float64
 7   energy                1173210 non-null float64
 8   key                   1173210 non-null int64
 9   loudness              1173210 non-null float64
10   mode                  1173210 non-null int64
11   speechiness           1173210 non-null float64
12   acousticness          1173210 non-null float64
13   instrumentalness      1173210 non-null float64
14   liveness              1173210 non-null float64
15   valence                1173210 non-null float64
16   tempo                 1173210 non-null float64
17   duration_ms           1173210 non-null int64
18   year                  1173210 non-null int64
19   release_date          1173210 non-null object
dtypes: float64(9), int64(4), object(7)
memory usage: 188.0+ MB
```

Data Set



Data, Data, Data
Hans Zimmer

2:54



3:49

ENDPOINTS

Albums

Artists

Shows

Episodes

Tracks

Get Track

Get Several Tracks

Get User's Saved Tracks

Save Tracks for Current User

Remove Tracks for Current User

Check User's Saved Tracks

Get Tracks' Audio Features

Get Track's Audio Features

Get Track's Audio Analysis

Get Recommendations

Search

Users

Playlists

Categories

Genres

Player

Markets

Get Track

OAuth 2.0

Get Spotify catalog information for a single track identified by its unique Spotify ID.

Request

GET /tracks/{id}

id string

The Spotify ID for the track.

Example value: "7lDfghXANMIKmxXsNCbNI"

Query

market string

An ISO 3166-1 alpha-2 country code. If a country code is specified, only content that is available in that market will be returned.

If a valid user access token is specified in the request header, the country associated with the user account will take priority over this parameter.

Note: If neither market or user country are provided, the content is considered unavailable for the client.

Users can view the country that is associated with their account in the account settings.

Example value: "ES"

Responses

200 401 403 429

A track

Body

application/json

album allOf

The album on which the track appears. The album object includes a link in href to full information about the album.

object

album_type string

The type of the album.

Allowed values: "album" "single" "compilation"

total_tracks integer

The number of tracks in the album.

Request Sample: Shell / cURL

```
curl --request GET \
  --url https://api.spotify.com/v1/tracks/{id} \
  --header 'Authorization: ' \
  --header 'Content-Type: application/json'
```

Response Example

```
1 {
2   "album": {
3     "album_type": "compilation",
4     "total_tracks": 9,
5     "available_markets": [
6       "CA",
7       "BR",
8       "IT"
9     ],
10    "external_urls": {
11      "spotify": "string"
12    },
13    "href": "string",
14    "id": "2up30Pmp9Tb4dAKM2erWQ",
15    "images": [
16      {
17        "url": "https://i.scdn.co/image/ab67616d00001
18        "height": 300,
19        "width": 300
20      }
21    ],
22    "name": "string",
23    "release_date": "1981-12",
24    "release_date_precision": "year",
25    "restrictions": {
26      "reason": "market"
27    },
28    "type": "album",
29    "uri": "spotify:album:2up30Pmp9Tb4dAKM2erWQ",
30    "album_group": "compilation",
31    "artists": [
32      {
33        "external_urls": {
34          "spotify": "string"
35        },
36        "href": "string",
37        "id": "string",
38        "name": "string",
```

Spotify Web API

Data, Data, Data
Hans Zimmer

2:54

3:49



HOME



TABLE OF CONTENTS



PLAYLIST

01

*Data
Ingestion*

02

***Data
Storage***

03

Processing

04

Deployment



Data Storage



- ❖ MongoDB is used to create a database from the .csv file
- ❖ After EDA, the pandas dataframe is saved as a .pkl file to be used by the Streamlit App



Data, Data, Data
Hans Zimmer

2:54



3:49



HOME



TABLE OF CONTENTS



PLAYLIST

01

*Data
Ingestion*

02

*Data
Storage*

03

Processing

04

Deployment



Processing



- ❖ Pandas is used to clean the data, drop redundant columns and handle Null data



Data, Data, Data
Hans Zimmer

2:54



3:49



HOME



TABLE OF CONTENTS



PLAYLIST

01

*Data
Ingestion*

02

*Data
Storage*

03

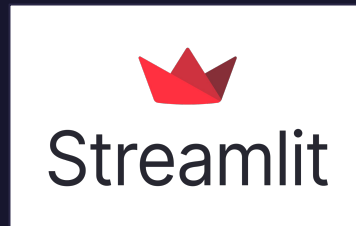
Processing

04

Deployment



Deployment



Streamlit Web
Application
Developed



Hosted on
Streamlit Cloud



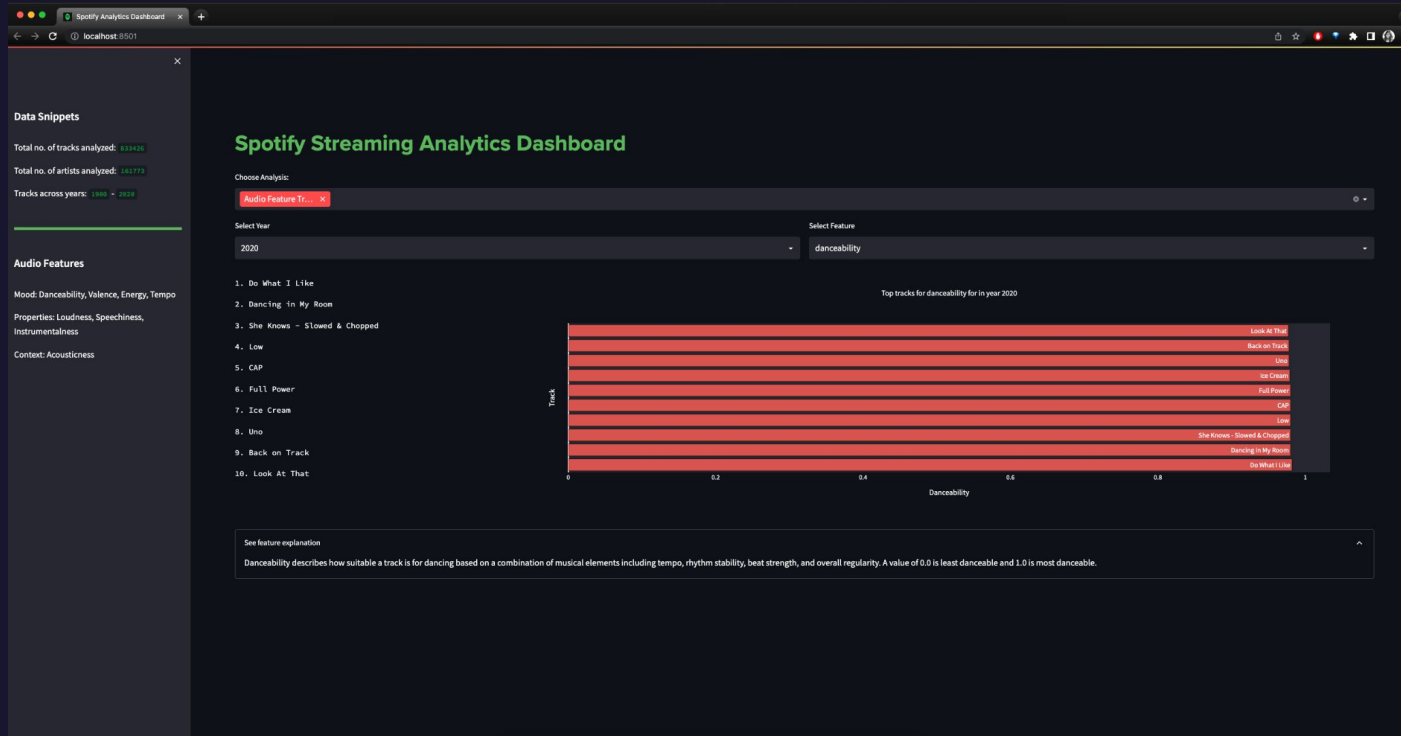
Data, Data, Data
Hans Zimmer

2:54



3:49

Streamlit App



Data, Data, Data
Hans Zimmer

2:54



3:49

Spotify Streaming Analytics Dashboard

Analysis Option 1:
Audio Features
Trends

Choose Analysis:

Audio Feature Tr... X

Select Year

2018

Select Feature

loudness

danceability

energy

loudness

speechiness

acousticness

instrumentalness

valence

tempo

1. No Damin' Up the Mighty Mississippi

2. Baseless Recursions

3. Rope in a Yankee

4. Kiss the Dog

5. Raw

6. Landed Me Montana

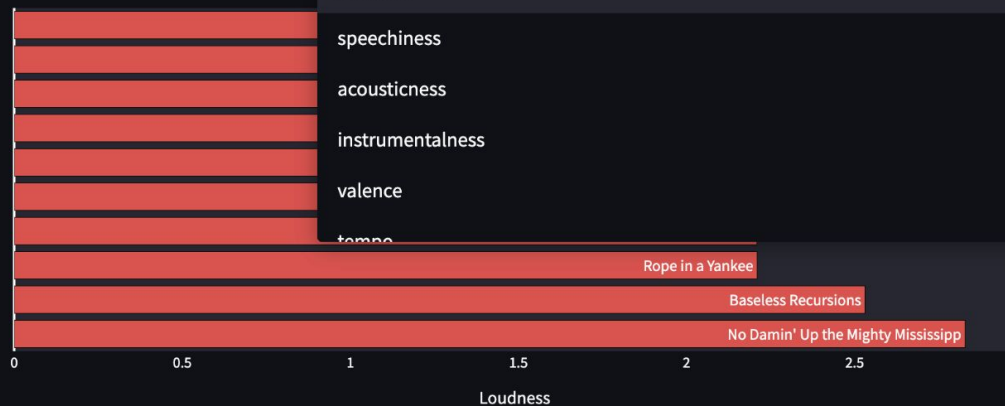
7. Doubts Are Just Detours

8. Last Words

9. Halleluyah

10. Everything Again

Track



Data, Data, Data
Hans Zimmer

2:54



3:49

Spotify Streaming Analytics Dashboard

Analysis Option 2:
Data Statistic

Choose Analysis:

Data Statistics ×

Select Year

2005

Data Statistics by Year

	stats	danceability	energy	loudness	speechiness	acousticness	instrumentalness	valence	tempo
	range	0-1	0-1	-60db - 0db	0-1	0-1	0-1	0-1	-
mean	mean	0.498537	0.495046	-11.672475	0.081099	0.480754	0.275145	0.459240	117.595742
min	min	0.000000	0.000000	-60.000000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	75%	0.632000	0.741000	-6.857000	0.067500	0.872000	0.696000	0.686000	136.608000
max	max	0.990000	1.000000	4.651000	0.965000	0.996000	1.000000	1.000000	236.062000



Data, Data, Data
Hans Zimmer

2:54



3:49

Sidebar

Data Snippets

Total no. of tracks analyzed: 833426

Total no. of artists analyzed: 161773

Tracks across years: 1980 - 2020

Audio Features

Mood: Danceability, Valence, Energy, Tempo

Properties: Loudness, Speechiness,
Instrumentalness

Context: Acousticness

Top Songs + Artists

1. OH
2. Shinkansen High Tension
3. Juggle Bubble
4. The World Is Not A Drag
5. Flight Of the Donkeys
6. San Francisco
7. White Gloves
8. Garbage In Garbage Out
9. Lifepreserver
10. Untitled



See feature explanation

The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

Feature Explanation Expander



Data, Data, Data
Hans Zimmer

2:54



3:49

Future Improvements

- **Web Application:**
 - Add more Analysis options: Genre Breakdown, Feature Comparison
 - Improve UI and make the application more user friendly: add discography for top charts
- **Data acquisition:**
 - Update Global data from Spotify Charts on a bi-weekly cadence using CRON job automation
- **Unit Testing:**
 - Add unit tests to check integrity of the data after every data update cycle
- **Machine Learning:**
 - Add a Recommender system based on Global trends and user genre preferences



Data, Data, Data
Hans Zimmer

2:54



3:49