

Song Lyrics Sentiment Analysis

Adam Nowicki

Table of contents

01

Introduction

02

Data

03

Data cleaning

04

**Database &
API**

05

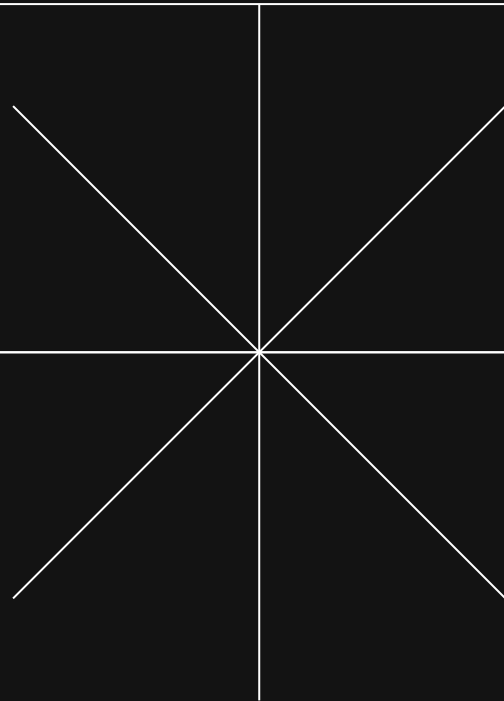
**Sentiment
Analysis**

06

**Data
Visualization**

01

Introduction



Introduction

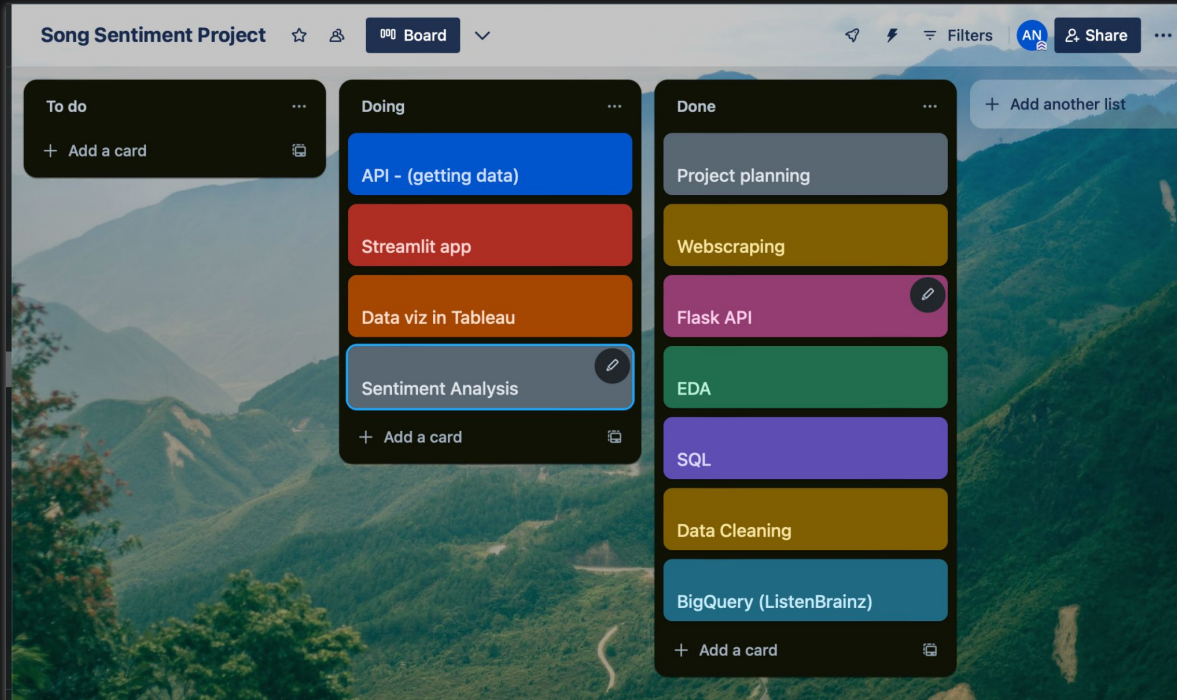
Goal:

- Conduct sentiment analysis of song lyrics using (NLP)
- Understand emotional tones of songs and identify genre and time trends.

Business Case:

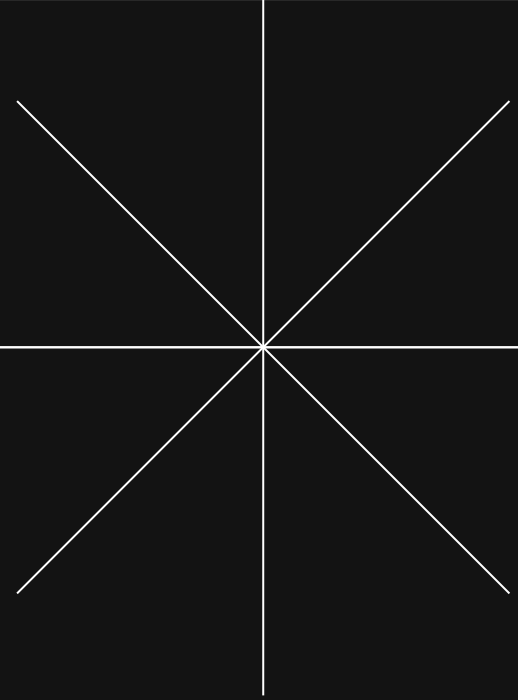
- Understanding societal changes
- Informing business strategies
- Understanding the emotional content of music can support mental health initiatives

Project Planning



02

Data



Data Gathering

Flat File

Genius Song Lyrics from Kaggle

5 134 856 x 11

Big Data System

ListenBrainz DataBase

8 480 062 x 3

Web Scraping

Billboard Top 100 Charts

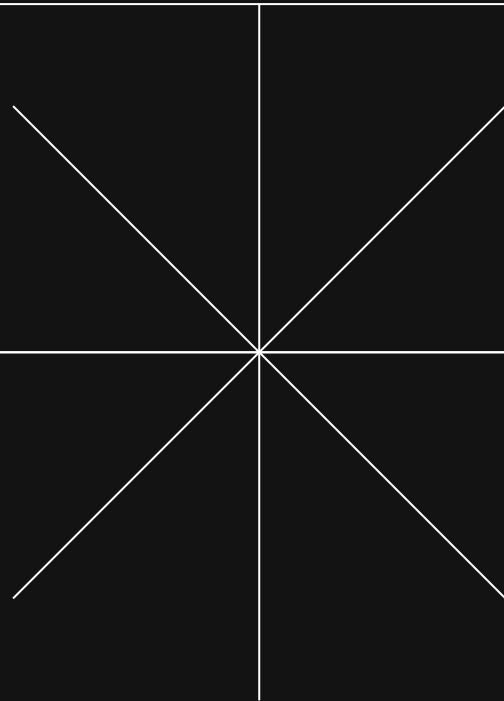
387 287 x 3

API

Genius API for retrieving individual song lyrics.

03

Data cleaning & wrangling



Data cleaning & wrangling



Step 1

Keeping only songs in English



Step 2

Removing null values and duplicates



Step 3

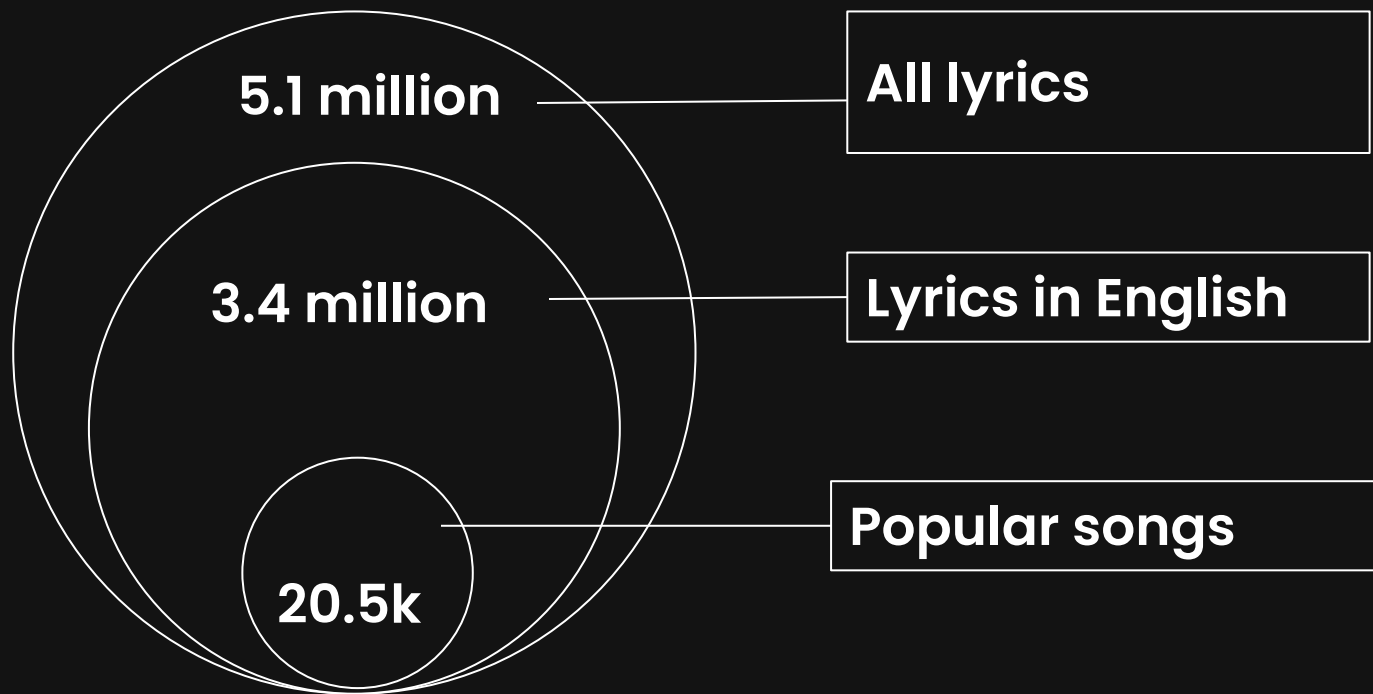
Filtering on song_id



Step 4

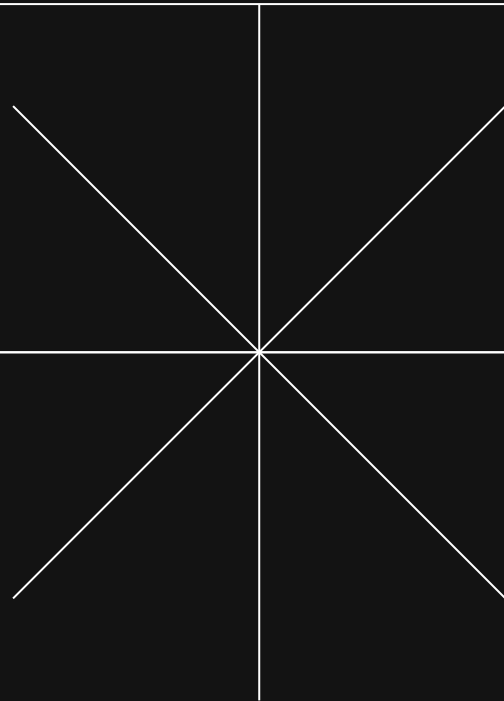
Removing duplicates and finalizing datasets

Size of the main dataset

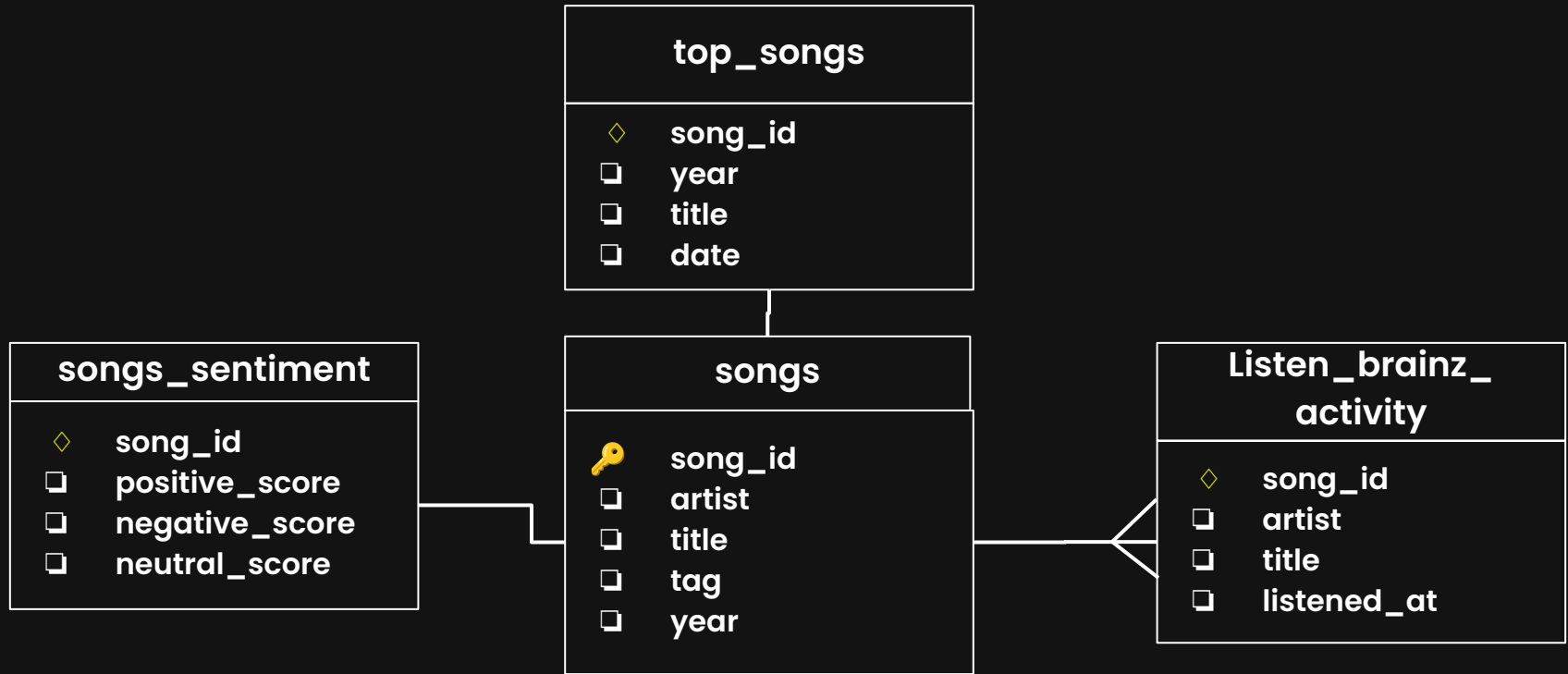


04

Database & API



Entity Relationship Diagram



SQL Query Example

```
36  -- Decades ranked by average positive score
37
38 • SELECT 10*FLOOR((`year`)/ 10) AS decade, avg(positive_score) as avg_positive_score
39 from songs as s
40 left join songs_sentiment as sent on s.song_id = sent.song_id
41 WHERE 10*FLOOR((`year`)/ 10) > 1940
42 group by decade
43 order by avg_positive_score DESC;
```

decade	avg_positive_score	
1950	0.47208783839753754	
1960	0.422362970422499	
1970	0.41350318835388955	
1990	0.3691134343563501	
1980	0.3520046154298824	
2000	0.3303761079619206	
2010	0.3287113055895929	
2020	0.23218435749359376	

Flask API

"/artists/artist"

This endpoint retrieves all songs of a specified artist from the SQL database and returns the songs along with their sentiment scores, categorized into positive, neutral, and negative.

"/genres/genre"

This endpoint fetches all songs belonging to a given genre from the SQL database and provides the songs along with their sentiment scores.

GET

/artists/{artist}

Get artist information and their songs' sentiment scores

Retrieves information about a specific artist and the sentiment scores of their songs.

Parameters

Cancel

Name	Description
artist required	The name of the artist
string (path)	
	<input type="text" value="the Rolling Stones"/>

Execute

Clear

Responses

Curl

```
curl -X 'GET' \
  'http://localhost:8080/artists/the%20Rolling%20Stones' \
  -H 'accept: application/json'
```

Request URL

```
http://localhost:8080/artists/the%20Rolling%20Stones
```

Server response

Code

Details

200

Response body

```
{
  "artist": "The Rolling Stones",
  "songs": [
    {
      "negative_score": 0.8398,
      "neutral_score": 0.1443,
      "positive_score": 0.0159,
      "title": "19th Nervous Breakdown"
    },
    {
      "negative_score": 0.3401,
      "neutral_score": 0.419,
      "positive_score": 0.2409,
      "title": "Aint Too Proud to Beg"
    },
    {

```

Download

Flask API

"/artists/artist"

This endpoint retrieves all songs of a specified artist from the SQL

GET

/artists/{artist}

Get artist information and their songs' sentiment scores

↗

Retrieves information about a specific artist and the sentiment scores of their songs.

Parameters

Cancel

Name	Description
artist * required string (path)	The name of the artist
	<input type="text" value="the Rolling Stones"/>

Execute

Clear

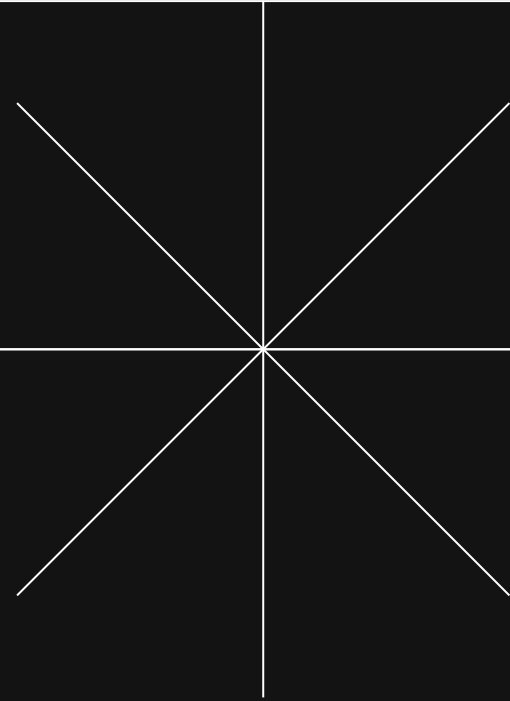
Response body

```
{
  "artist": "The Rolling Stones",
  "songs": [
    {
      "negative_score": 0.8398,
      "neutral_score": 0.1443,
      "positive_score": 0.0159,
      "title": "19th Nervous Breakdown"
    },
    {
      "negative_score": 0.3401,
      "neutral_score": 0.419,
      "positive_score": 0.2409,
      "title": "Aint Too Proud to Beg"
    },
    {
```

Download

05

Sentiment Analysis



Model

Model chosen for the sentiment analysis:

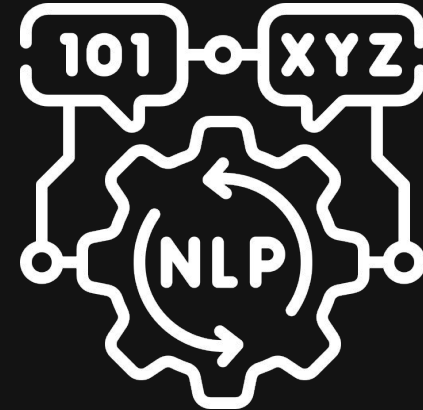
“Twitter-roBERTa-base for Sentiment Analysis – UPDATED (2022)”

Advantages:

- Informal Language Understanding
- Large-scale pretraining ~124 million tweets.

Shortcomings:

- Model trained on twitters posts
- Length of the song lyrics
- No training dataset for verification



Process



Step 1

Data gathering &
cleaning

Step 2

Text preprocessing

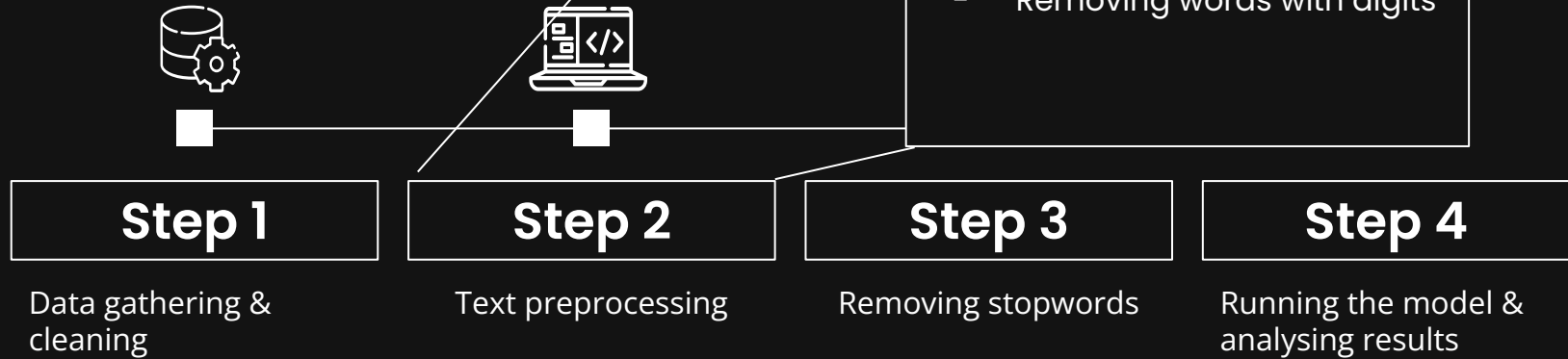
Step 3

Removing stop
words

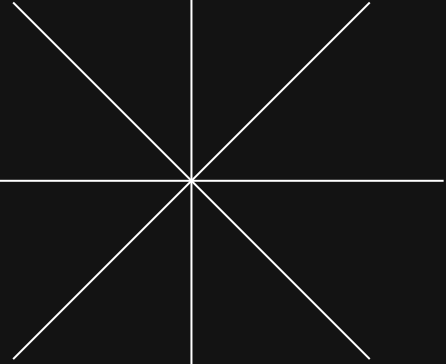
Step 4

Running the model &
analysing results

Process



Examples





*"I see trees of green, red roses too. I see them bloom
for me and you. And I think to myself, what a
wonderful world"*

Positive 0.95 Negative: 0.002 Neutral 0.03



It's the end of the world as we know it
It's the end of the world as we know it
It's the end of the world as we know it, and I feel fine

Positive 0.7 Negative: 0.03 Neutral 0.26



*Hello darkness, my old friend
I've come to talk with you again*

Positive 0.07 Negative: 0.25 Neutral 0.68

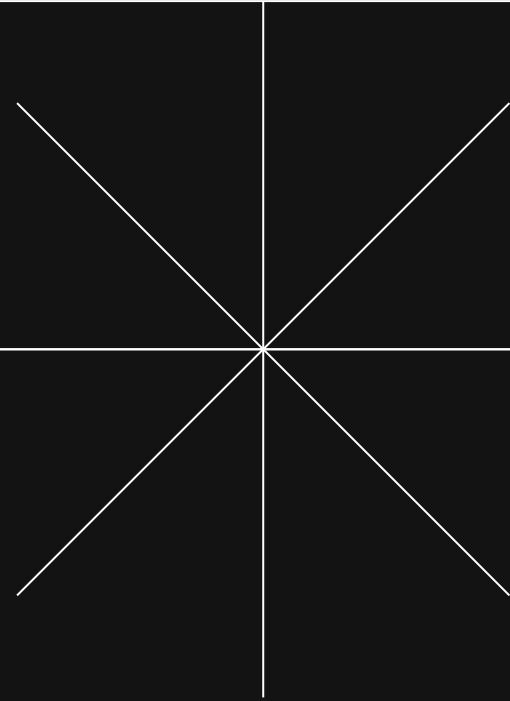


*When I die, f*** it, I wanna go to hell
'Cause I'm a piece of shit, it ain't hard to f*** tell
It don't make sense, goin' to heaven with the goodie-goodies
Dressed in white, I like black Timbs and black hoodies*

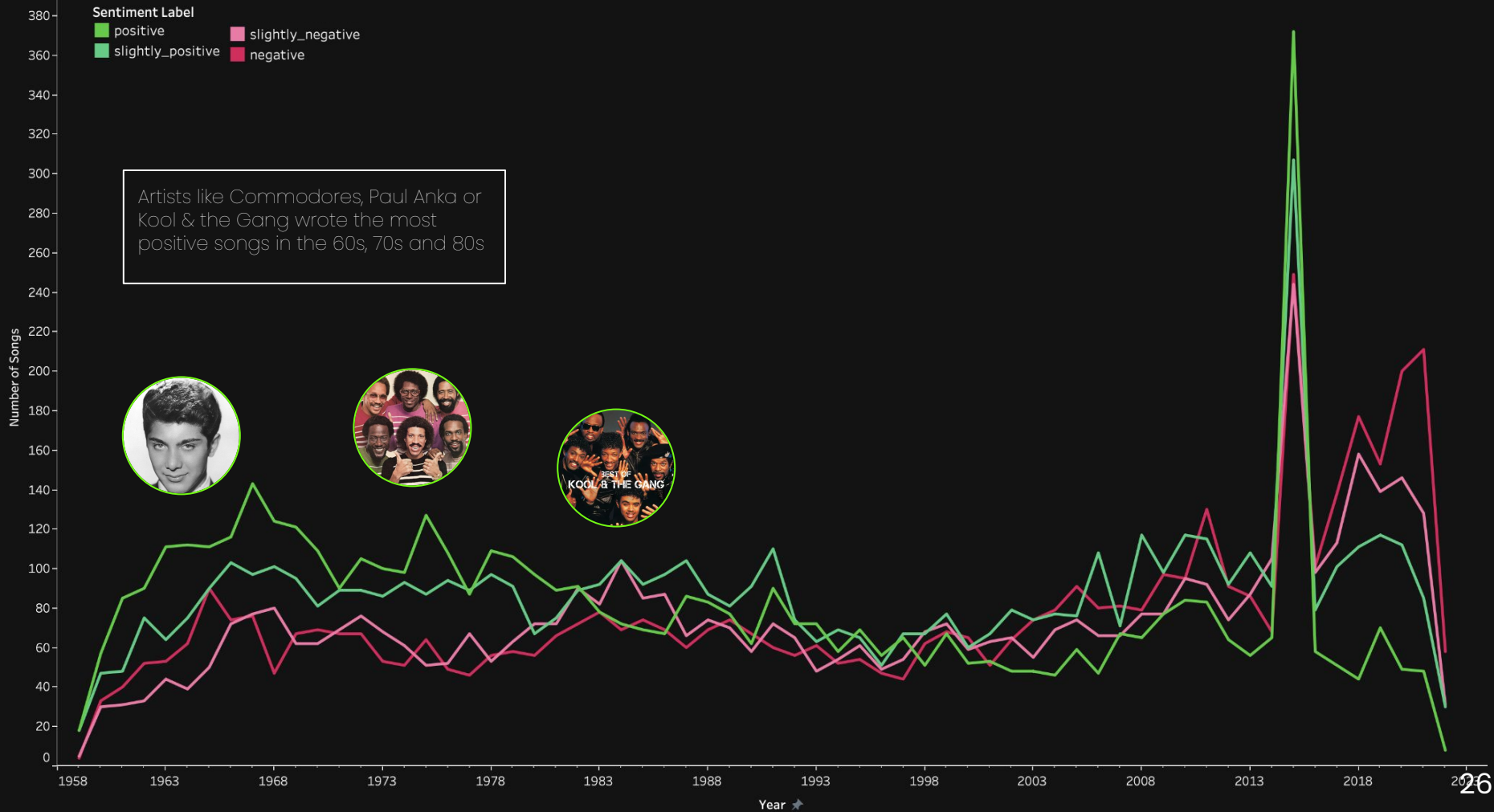
Positive 0.03 Negative: 0.8 Neutral 0.16

05

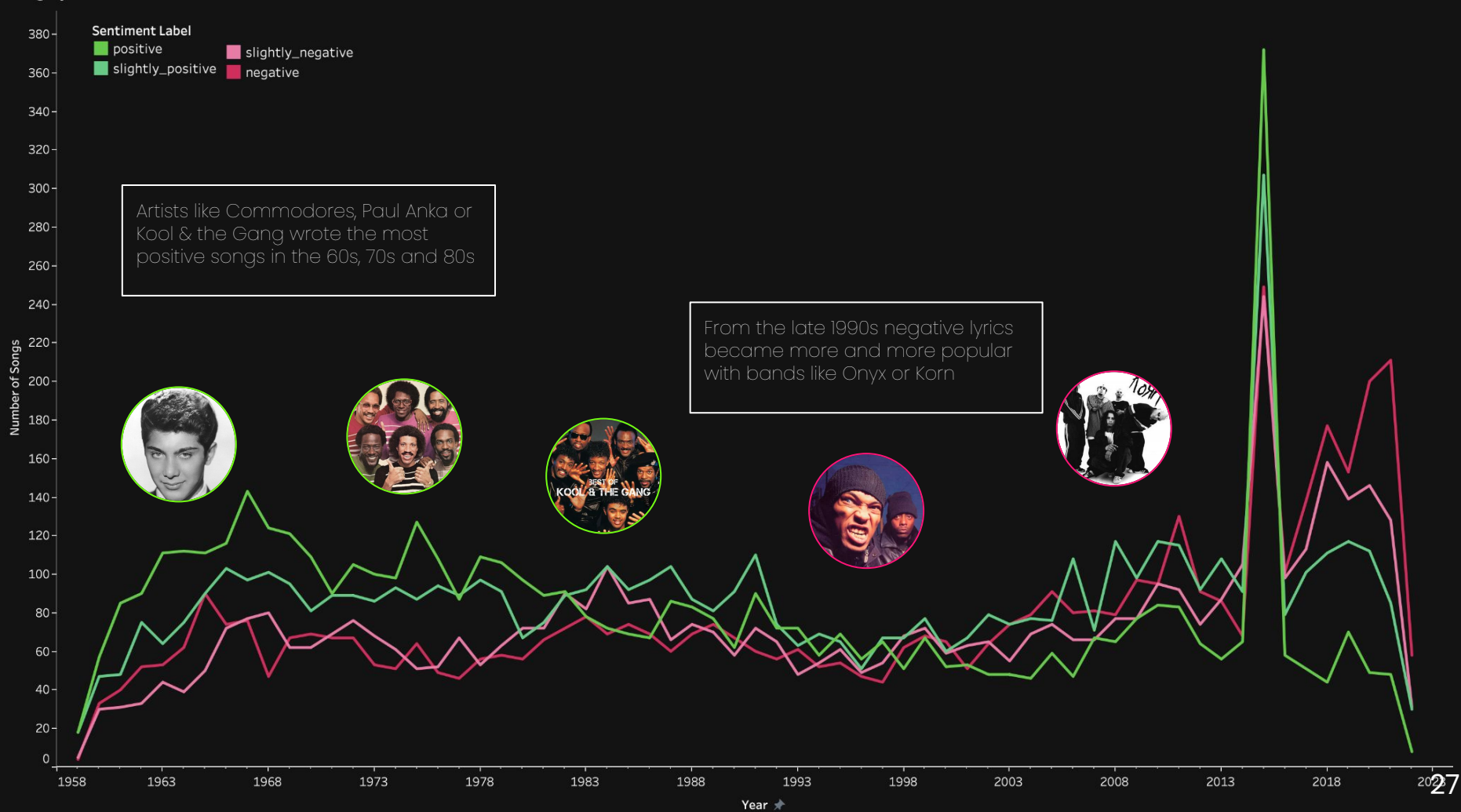
Data Visualization



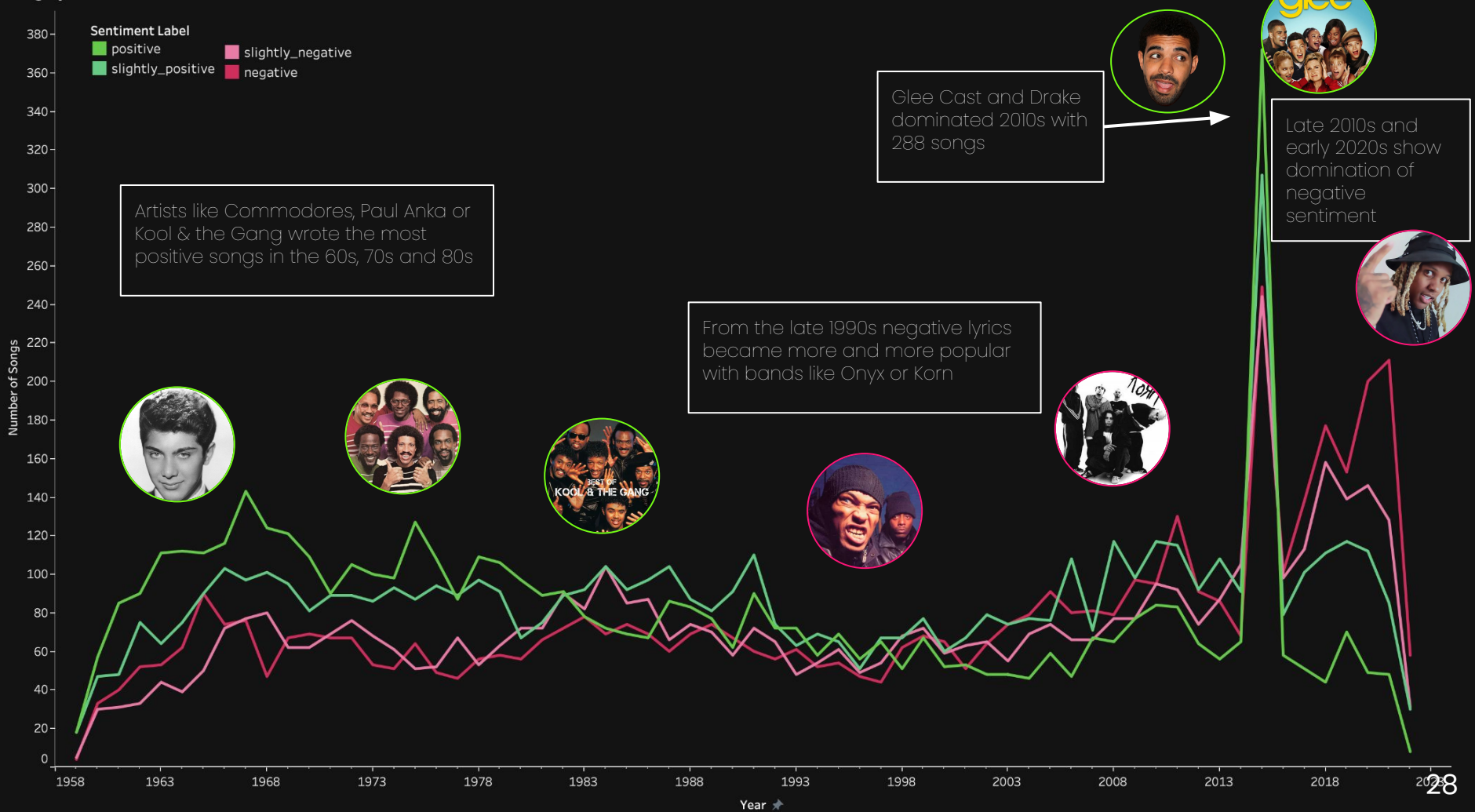
Song Lyrics Sentiment Over Time



Song Lyrics Sentiment Over Time



Song Lyrics Sentiment Over Time



Word Cloud Charts (decades)

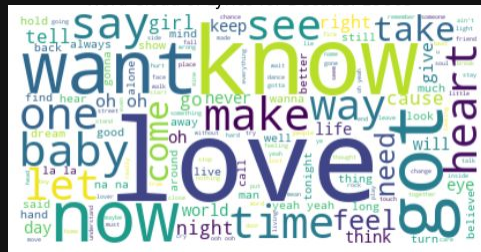
1960s



1970s



1980s



1990s



2000s



2010s



Tableau Story

Streamlit Demo

Next Steps

Expand the scope

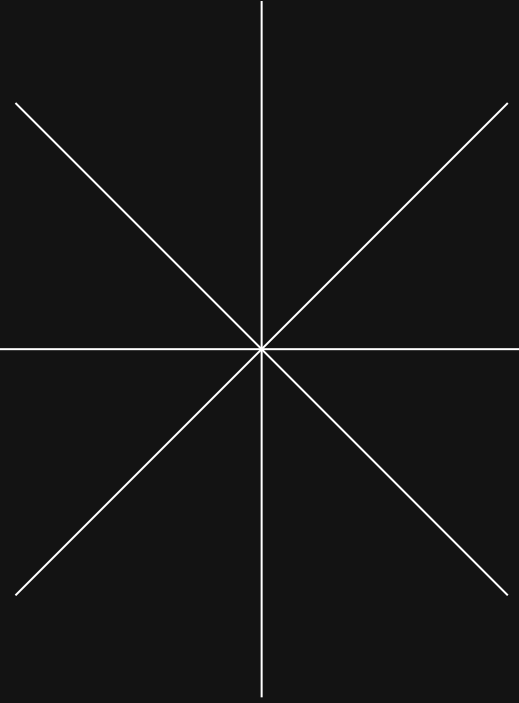
- Expand song selection
- Incorporate mental health
- Multilingual

Model refinement

- Tailor for lyrics sentiment analysis
- Enhance accuracy

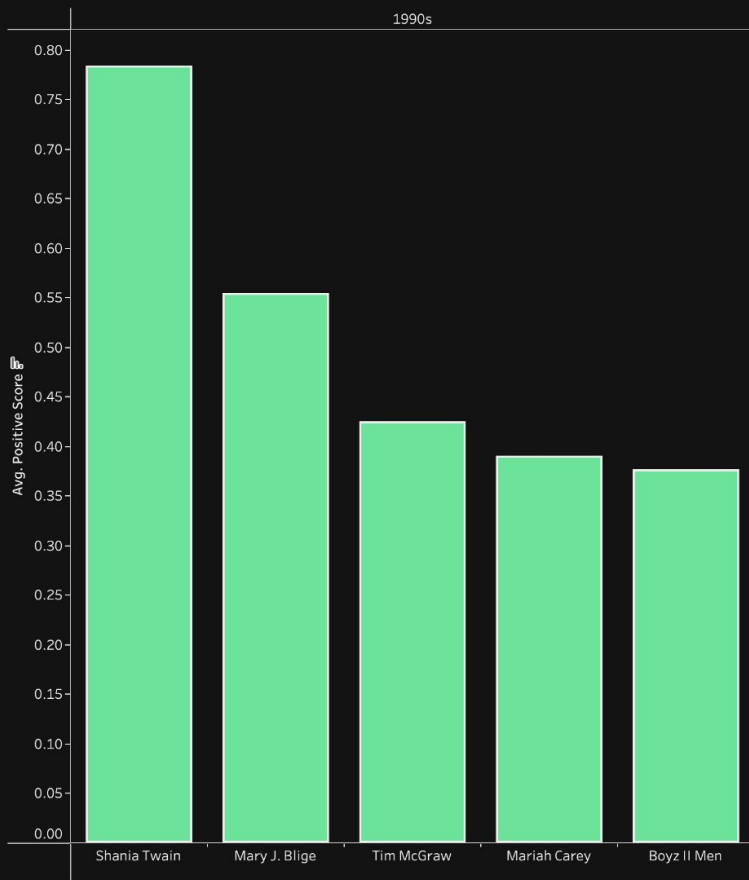
Thanks!

CREDITS: This presentation template was created by Slidesgo, and includes icons by Flaticon, and infographics & images by Freepik

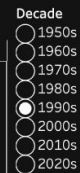
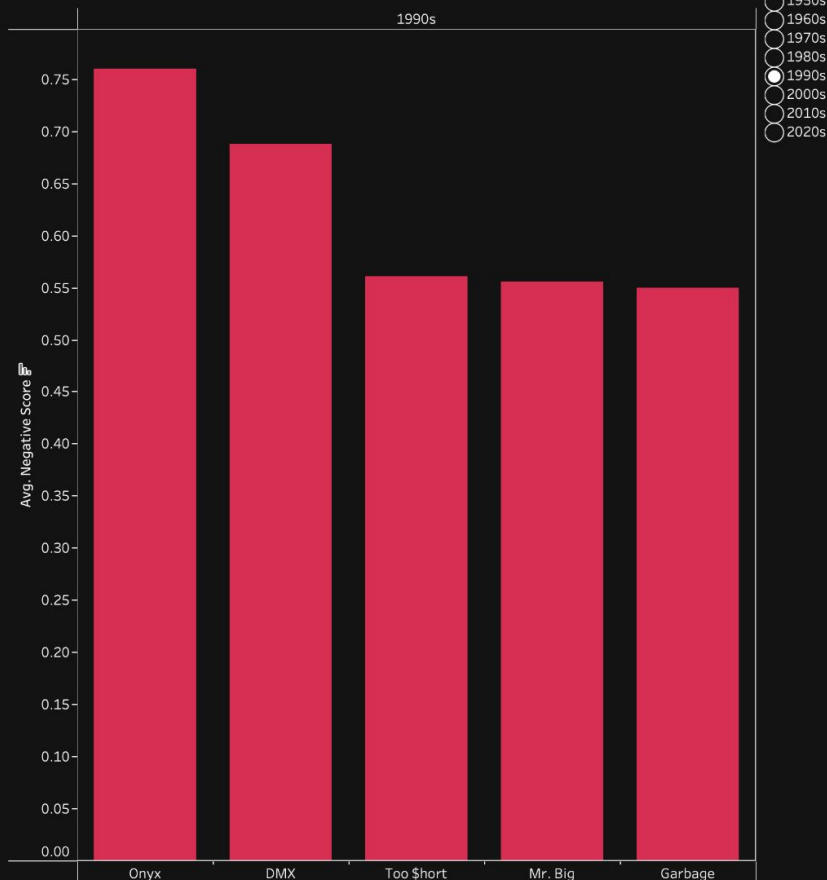


Artists

Most positive artists per decade

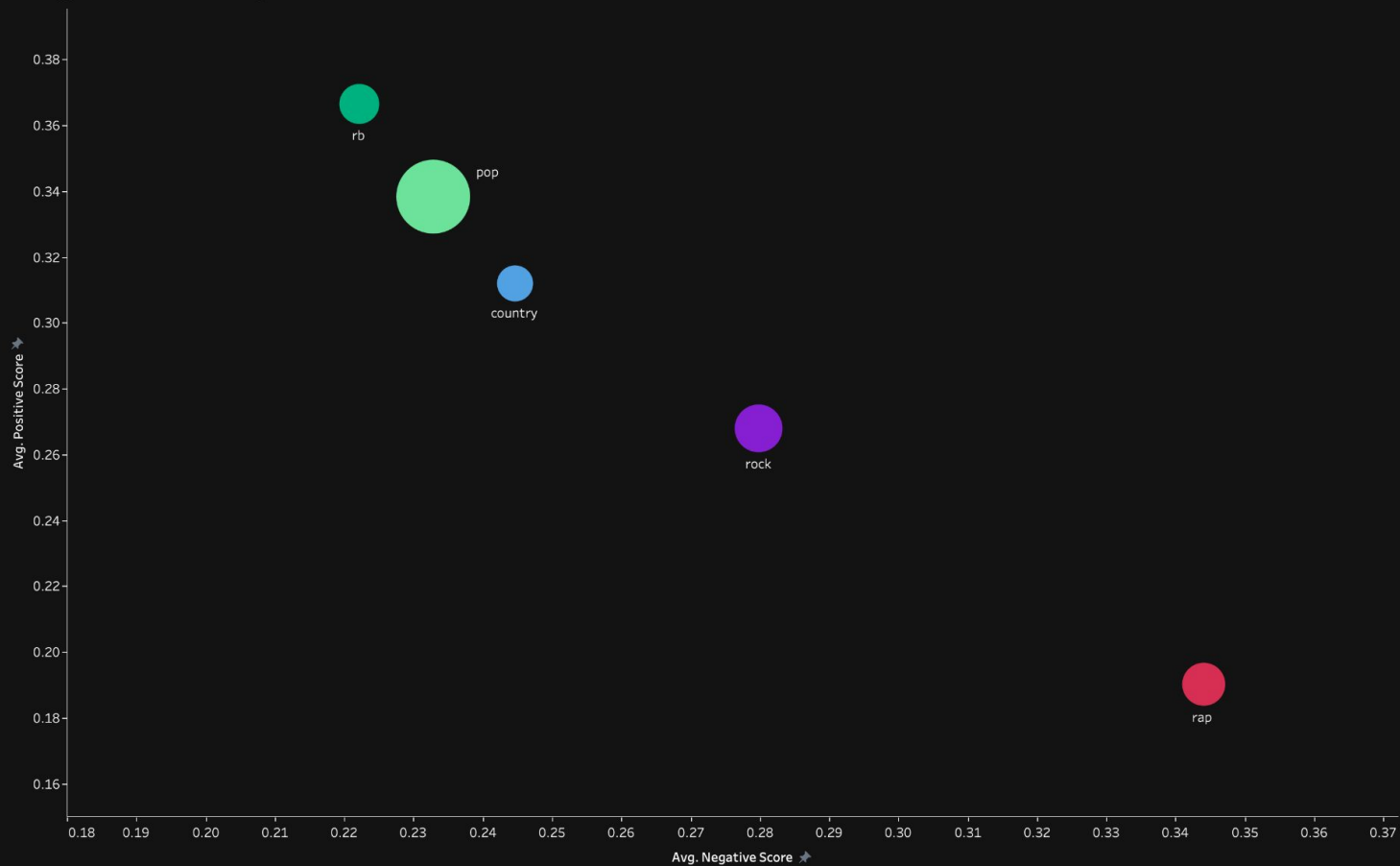


Most negative artists per decade



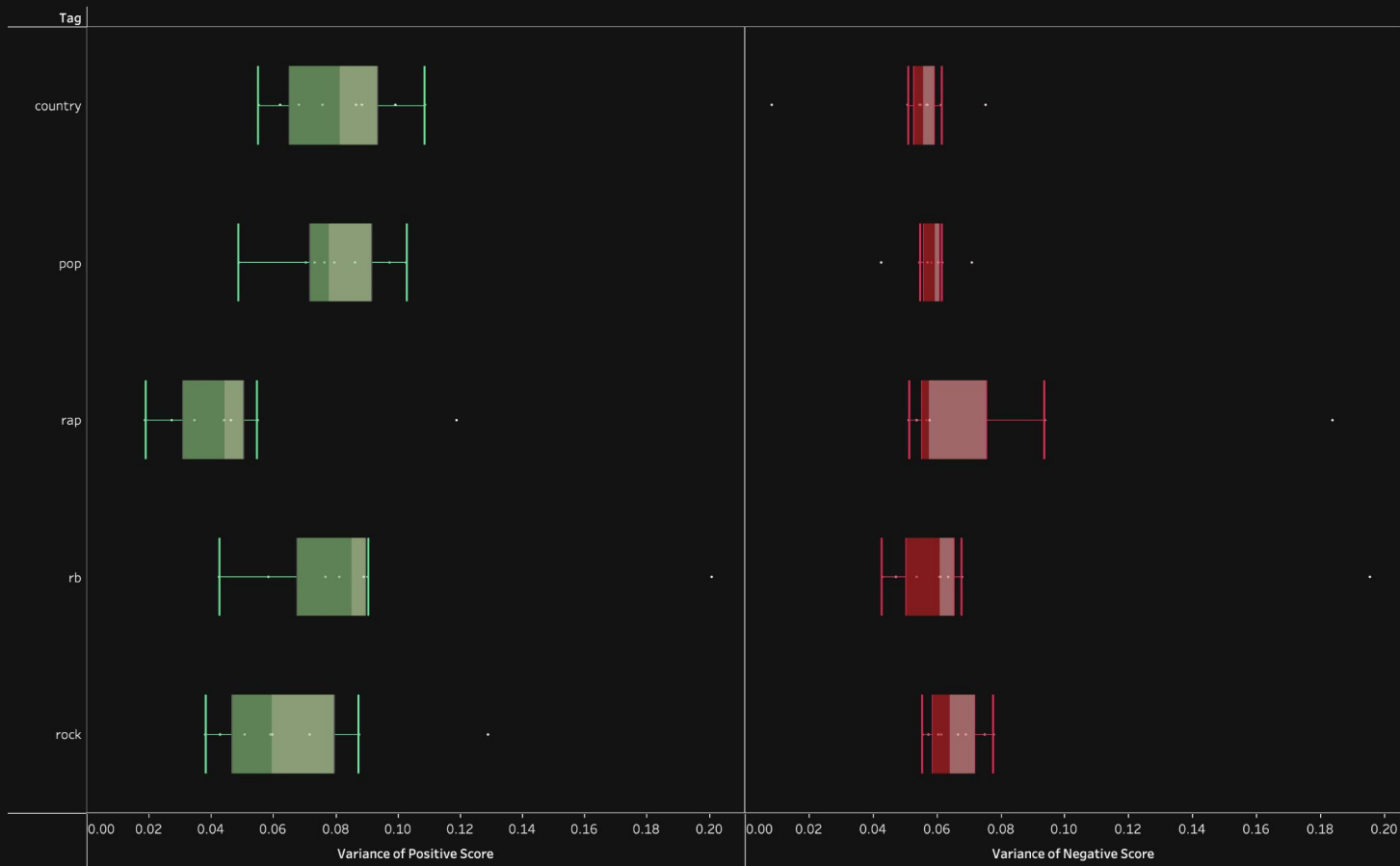
Genres

Average Sentiment Score by Genre

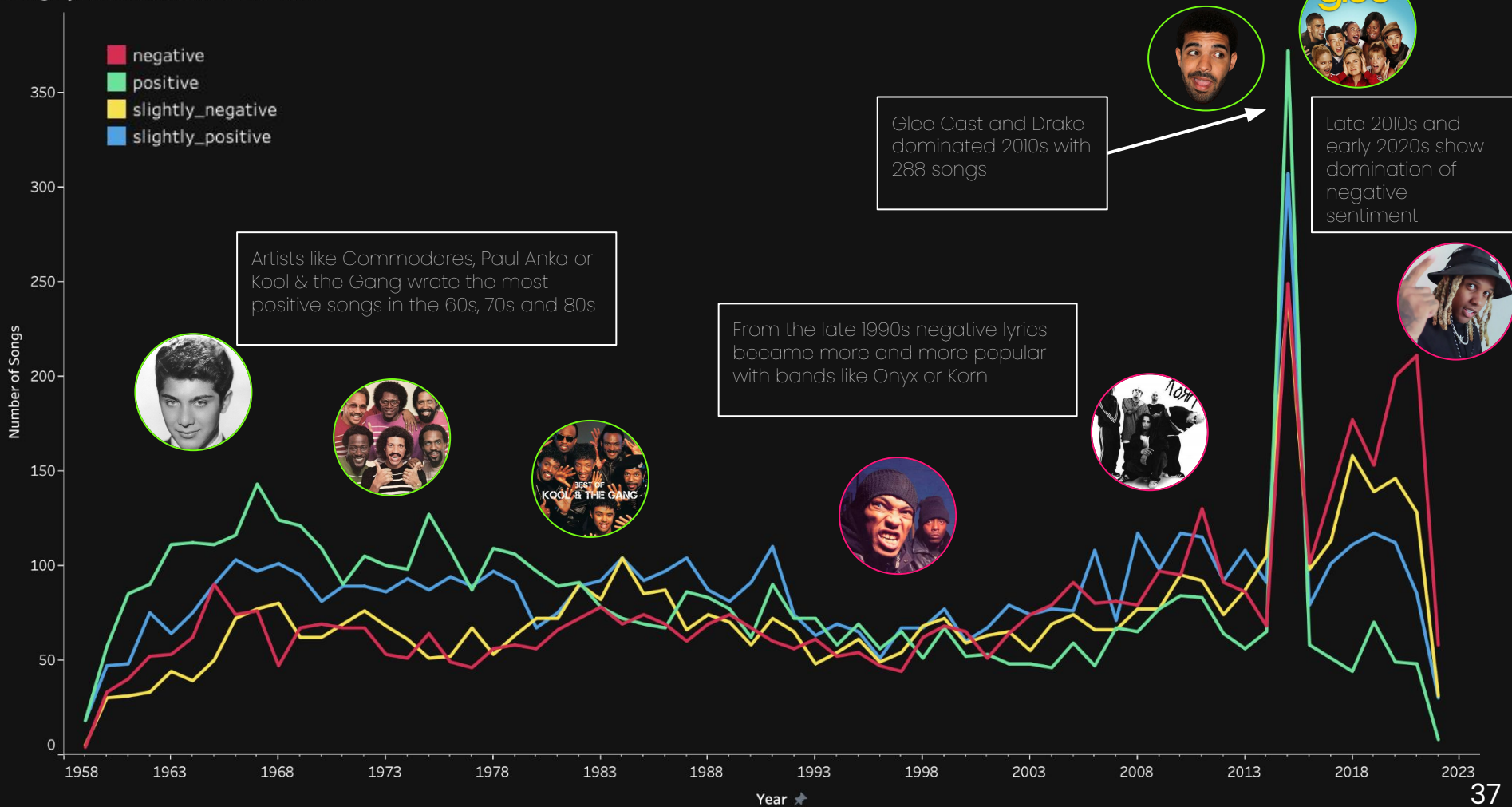


Genres

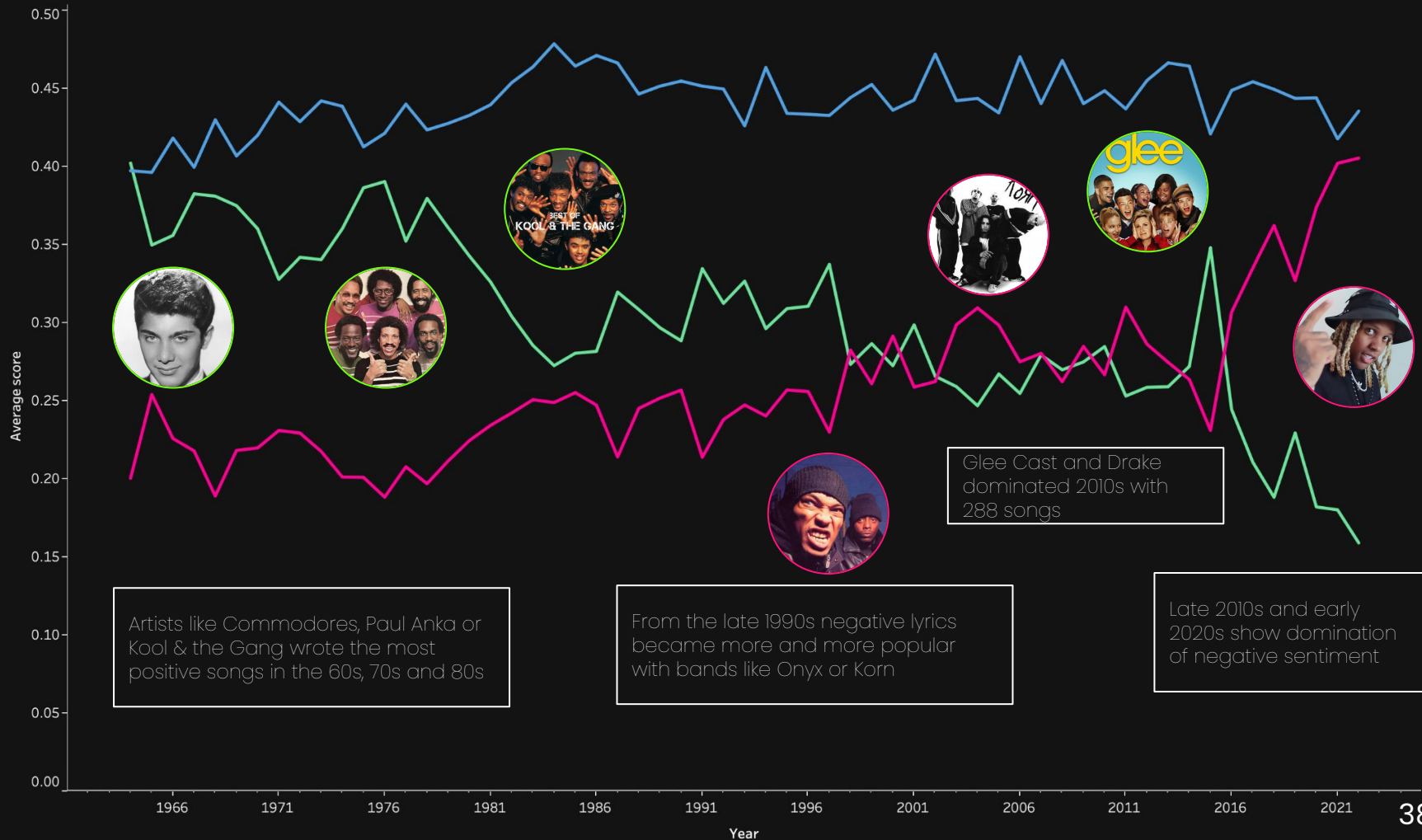
Positive and Negative Sentiment Variance per Genre



Song Lyric Sentiment Over Time



Average sentiment scores



Resources

Did you like the resources on this template? Get them for free at our other websites:

Vectors

- Business concept instagram posts

Photos

- Black man giving presentation on a meeting
- People taking part of business event
- Woman giving report to her boss
- People taking part of business event I
- People taking part of business event II
- People taking part of business event III
- Businesswoman posing outdoors
- Smiley man holding project documents