Deep Learning for Sentiment Analysis on Music Artists

Ibaad Hassan (zxr2ex), Mayna Malhotra (dma2rv),
Elina Liu (bug4my), Emma Ylagan (fwx9pw)

12-15-25

Videos

Full Presentation: https://youtu.be/75aeonsjKK0

Code demos of each part:

- Lyric Sentiment: https://youtu.be/ZwsHOmrfXFE
- Theme analysis: https://youtu.be/l7HYH53H9nw
- Public Perception: https://youtu.be/GlcwNCH_gfM
- Fame over time: https://youtu.be/BruqunChwj8

Motivation

Music is a central part of everyday life, and we were interested in gaining deeper insights into today's most influential artists beyond surface-level popularity metrics.

Currently, streaming platforms tell us who we listen to, but they don't tell us how that music really feels. There is no easy way to see if your favorite artists are overall positive or negative in their lyrics.

By applying natural language processing to song lyrics and related text data, we can uncover emotional and thematic patterns that are not captured by genre or streaming counts alone. Being able to analyze lyrics gives us a new layer of insight into our own music taste, and could eventually lead to discovering new music based on "mood" and emotion rather than just genre.

Background

Computational music analysis applies NLP and ML techniques to understand:
- Artistic expression
- Audience response
- Popularity dynamics

Early approaches to music analytics:
- Rely primarily on **aggregate popularity metrics** (streams, chart rankings)
- **Shallow text representations** (can fail to capture nuance)

Our Approach: combining textual content, public perception, and temporal fame provides a more complete and interpretable representation of an artist’s identity.

Related Work

[Sentiment Analysis of Social Media Data for Predicting Consumer Behavior Trends]

- BERT to classify sentiments and predict consumer trends

[Multimodal Music Popularity Prediction via Adaptive Fusion of Modality Experts and Temporal Engagement Modeling]

- GAMENet deep learning architecture for predicting a song's commercial success

[How Machine Learning Is Transforming Music Marketing and Trend Analysis]

- Industry overview describing how machine learning is used to analyze streaming data, listener behavior, and social trends for music discovery and personalized marketing
- There is industry demand for predictive analytics in music and understanding the patterns to artist success

Claim / Target Task

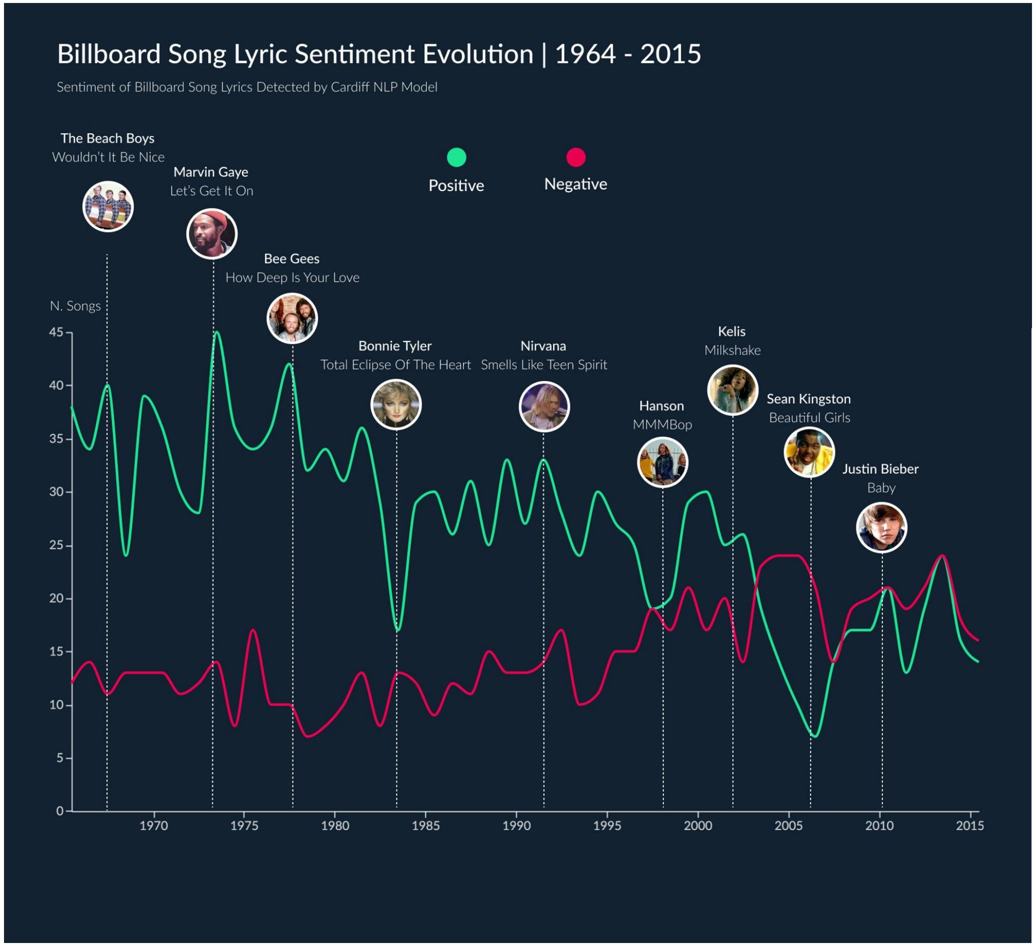
Our goal is to build a comprehensive profile for six of the most popular artists in the world right now: Drake, Kendrick Lamar, The Weeknd, Billie Eilish, Taylor Swift, and Bad Bunny.

Each artist’s profile will consist of:

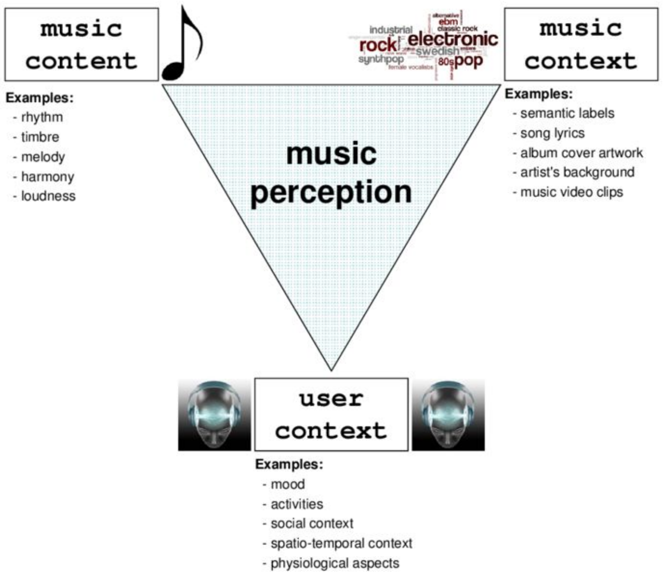
1. Lyric Sentiment in Songs
2. Common Themes/Messages in Songs
3. Artist’s Public Perception via Social Media
4. Artist’s Fame Over Time

We claim that a single metric like lyric sentiment or popularity alone cannot capture an artist’s true identity. Instead, a well-rounded profile requires combining lyrical content, thematic structure, public perception, and temporal fame as it encapsulates the artist in a fuller sense.

An Intuitive Figure Showing WHY Claim



- Music perception is shaped by music content (themes/sentiment), music context (artist fame over time), and user context (public perception)



- Lyric sentiment changes over time & can be subjective
- Positive & negative sentiments are unrelated to how "popular" a song or artist is

Proposed Solution

Lyric Sentiment

- Dataset: Genius Song Lyrics dataset
- Model: distilbert-base-uncased

Common Themes

- Dataset: Genius Song Lyrics dataset
- Models: Sentence-BERT and K-means

Public Perception

- Datasets: Twitter, Reddit, and News
- Models: twitter-roberta-base-sentiment-latest, sentence-transformers/all-mpnet-base-v2

Fame Over Time

- Dataset: Spotify Global Music Dataset (2009-2025)
- Model: SARIMAX

Implementation

1: Lyrics Sentiment

Lyric Sentiment

1. Data Ingestion & Filtering

- Integrated the **Genius Song Lyrics** dataset (~5 million songs) to ensure that there is wide coverage.
- Built a filtration pipeline to isolate discographies for our six target artists, cleaning metadata and standardization tags.

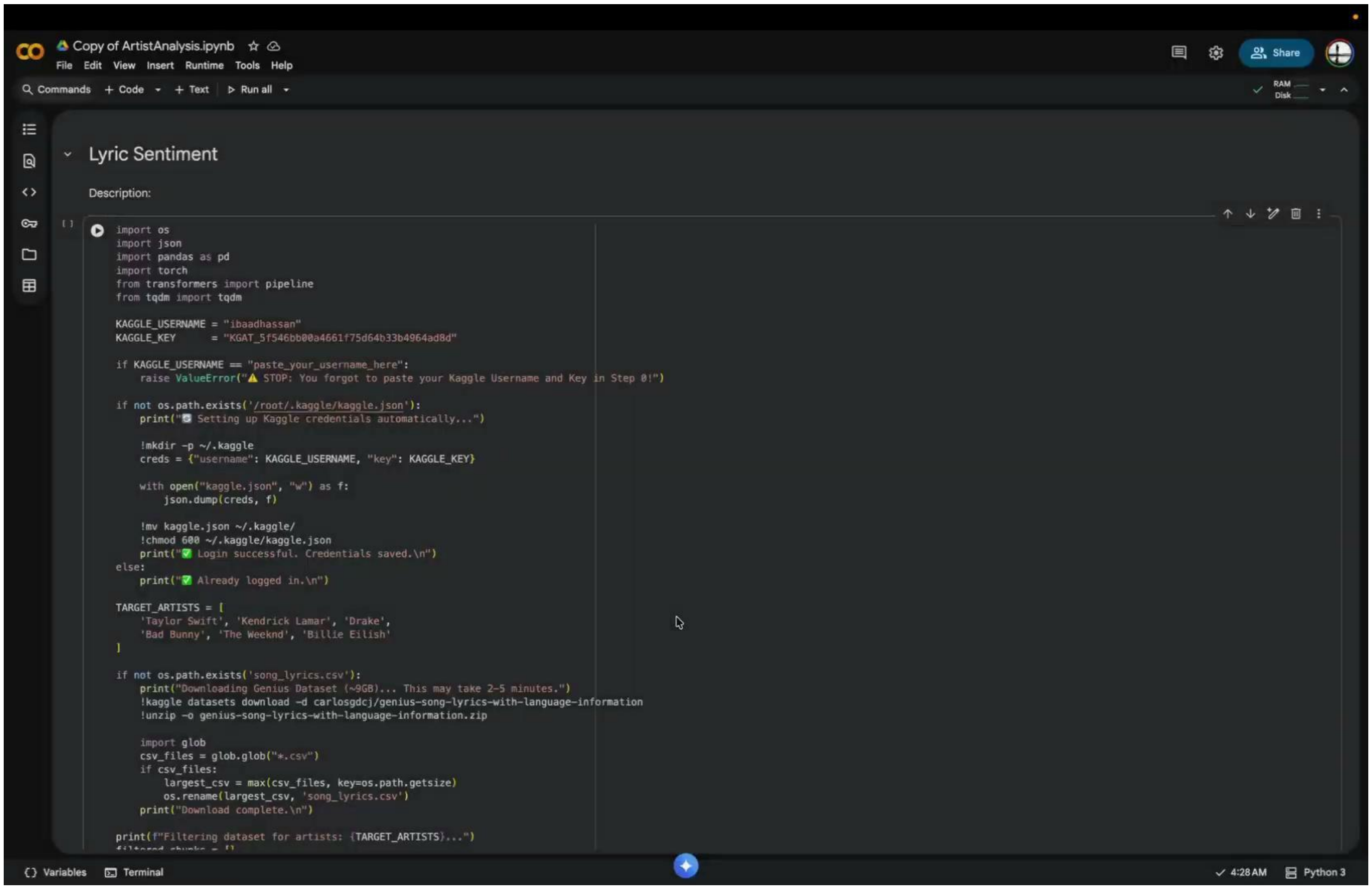
2. Model Architecture: DistilBERT

- Deployed **distilbert-base-uncased**, a distilled version of the BERT transformer.
- I chose it because it retains 97% of BERT's performance while being 40% lighter and 60% faster, allowing for rapid inference across large lyric datasets.
- Leveraged a version fine-tuned on the **SST-2 (Stanford Sentiment Treebank)** to ensure high-accuracy binary classification (Positive/Negative).

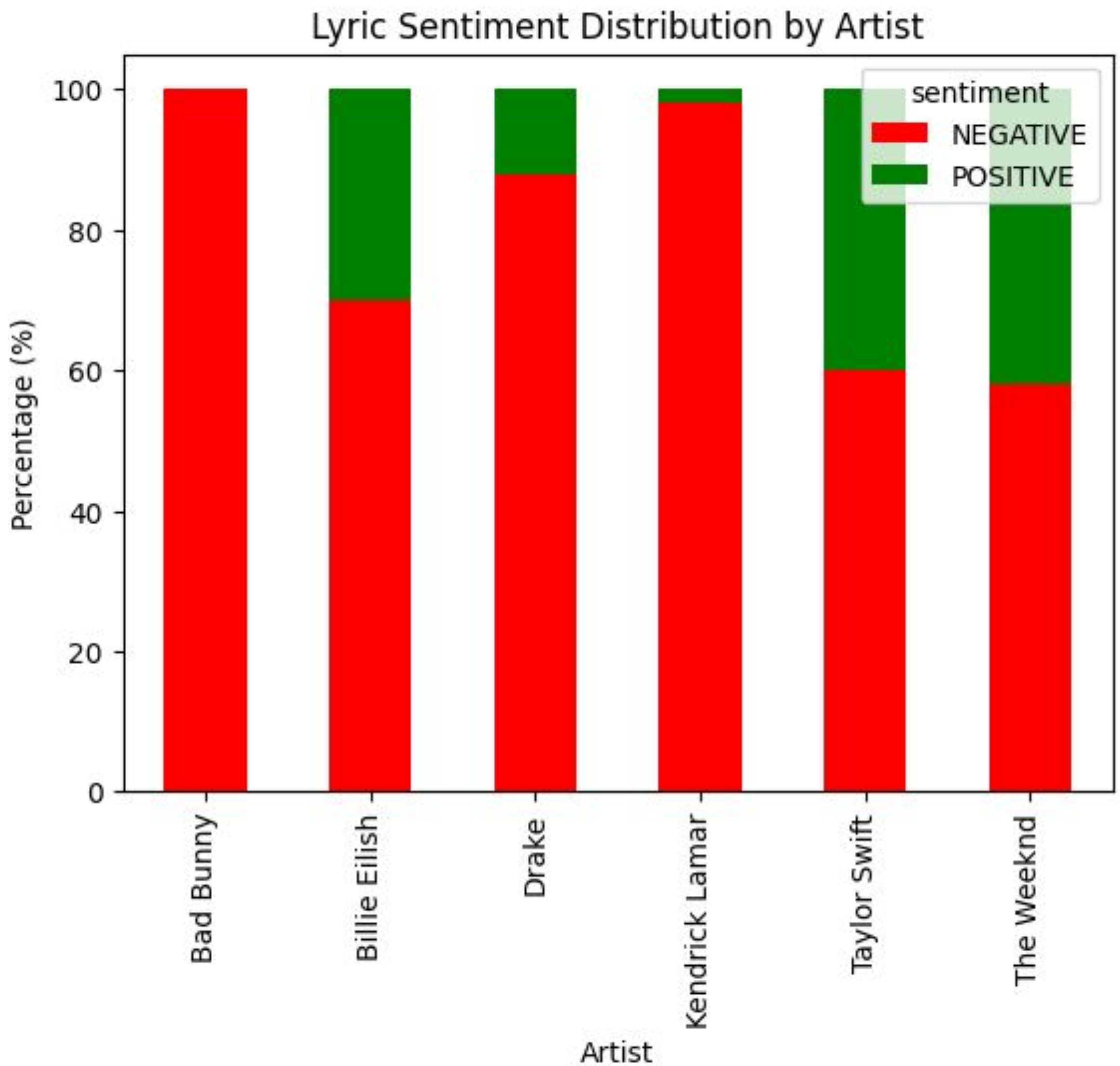
3. Inference & Aggregation Strategy

- Implemented truncation strategies (max 512 tokens) to fit lengthy lyrical compositions within the transformer's context window.
- Aggregated song-level sentiment predictions into normalized distributions, enabling quantitative comparison of "emotional profiles" across different artists.

Lyric Sentiment Per Artist



Lyric Sentiment Per Artist



2: Themes

Commons Themes in Songs by Artist

1. Data Ingestion & Filtering

- Integrated the Genius Song Lyrics dataset (same as Lyric Sentiment)
- Built a chunked filtering pipeline to isolate discographies for our six target artists (Taylor Swift, Drake, Kendrick Lamar, Bad Bunny, Billie Eilish, The Weeknd).
    a. Filtered dataset to include only artist ‘name’ and ‘lyrics’ columns.
- Split lyrics into ‘sentences’ for more granular thematic analysis, discarding extremely short or meaningless lines.

2. Embedding & Clustering

- Used **Sentence-BERT (all-mpnet-base-v2)** to embed sentences into semantic vector space.
- Clustered sentences with K-means (k=12) to group semantically similar lyrics.
    a. Each cluster represents a potential theme.

3. Theme Identification & Labeling

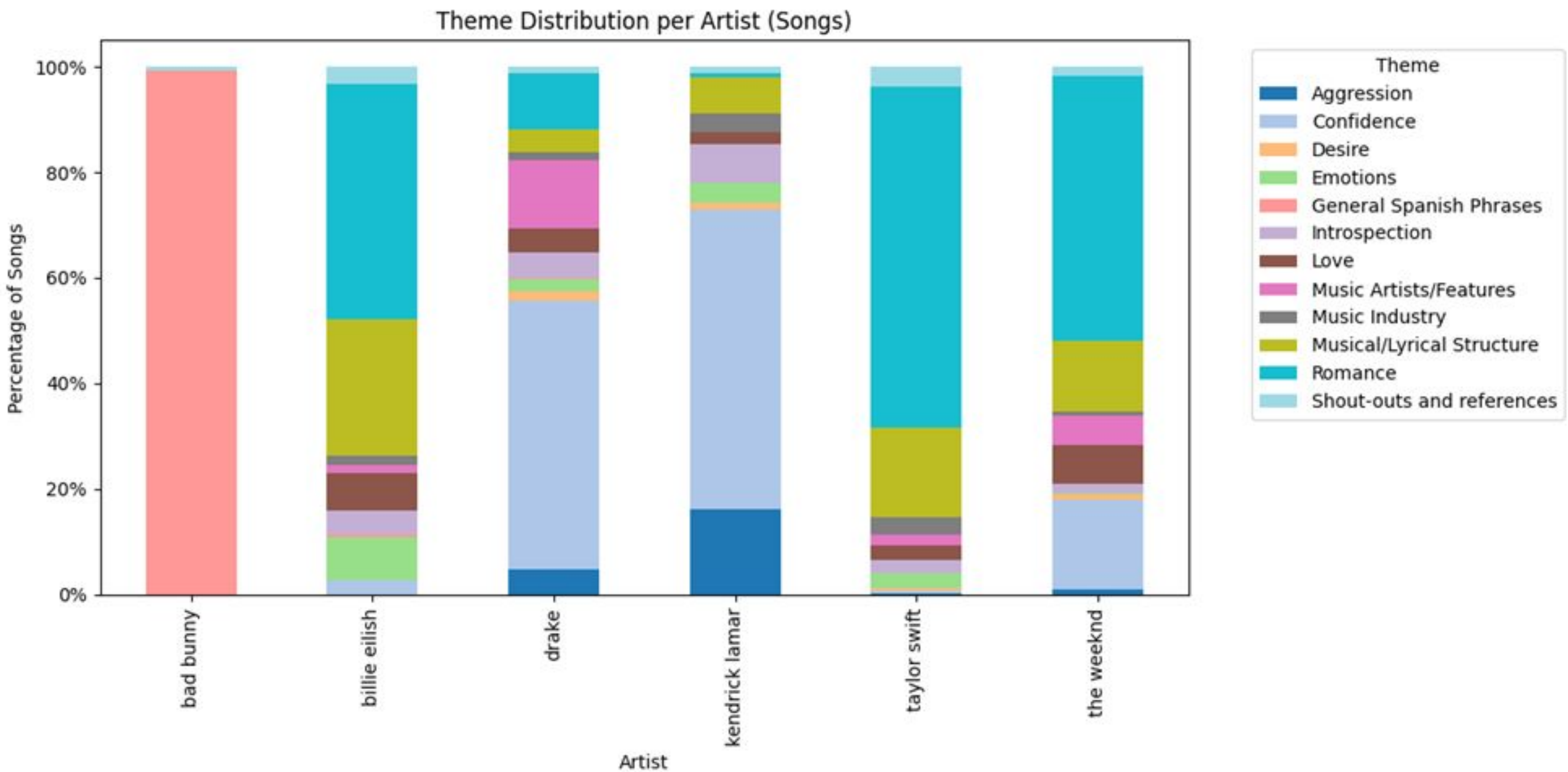
- Inspected **top TF-IDF keywords** per cluster to understand cluster meaning.
- Manually mapped clusters to readable themes:
    a. Love, Romance, Desire, Emotions, Introspection, Confidence, Aggression, Music Industry, Music Artists/Features, Musical/Lyrical Structure, Shout-outs/References, General Spanish Phrases.
- Assigned each sentence and song to a theme for further analysis

4. Aggregation Strategy for Data Visualization

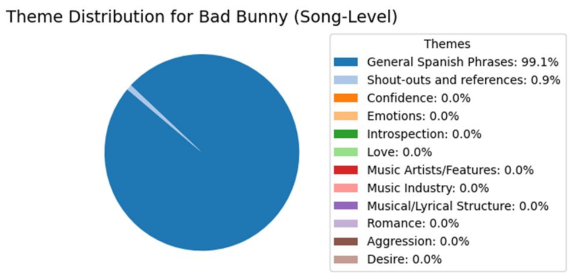
- **Song-level:** Each song was assigned the theme that appeared most frequently in its sentences.
    a. Visualizations show the percentage of *songs* that fall into each cluster per artist
- **Sentence-level:** Visualizations show the percentage of *sentences* that fall into each cluster per artist, across all songs.

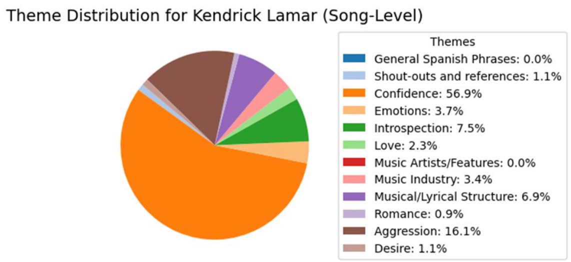
Common Themes in Songs by Artist: Code Demo

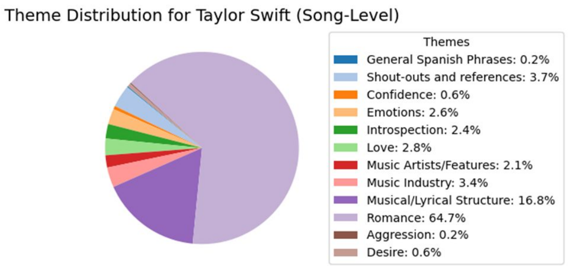
Common Themes in Songs by Artist: Song-Level Results

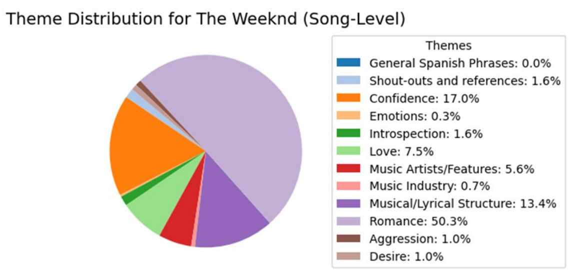


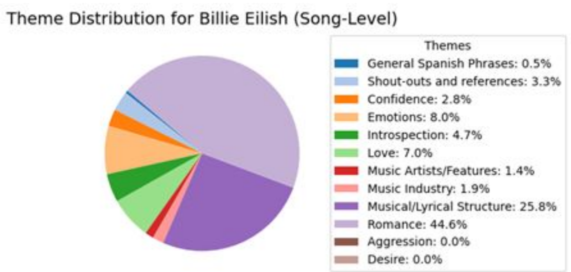
Common Themes in Songs by Artist: Song-Level Results (continued)

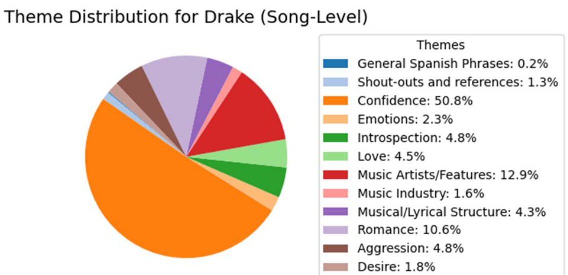




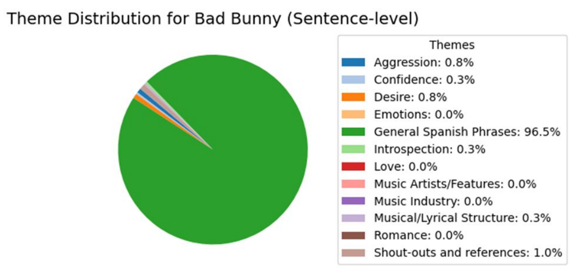


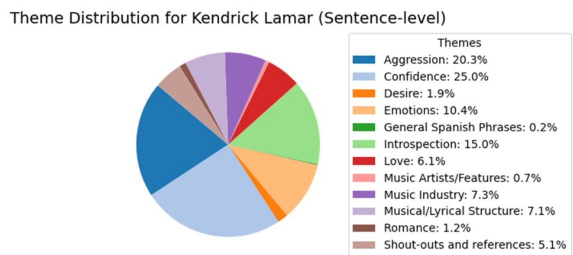


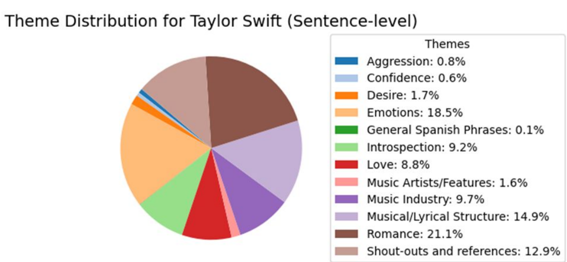


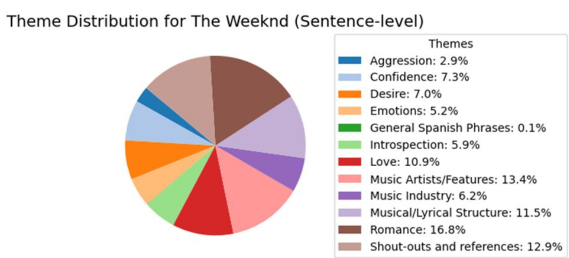


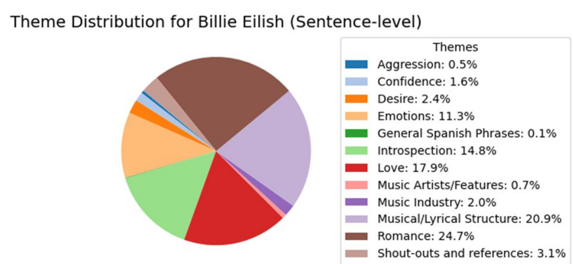
Common Themes in Songs by Artist: Sentence-Level

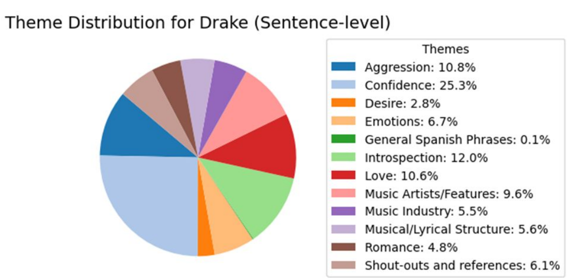




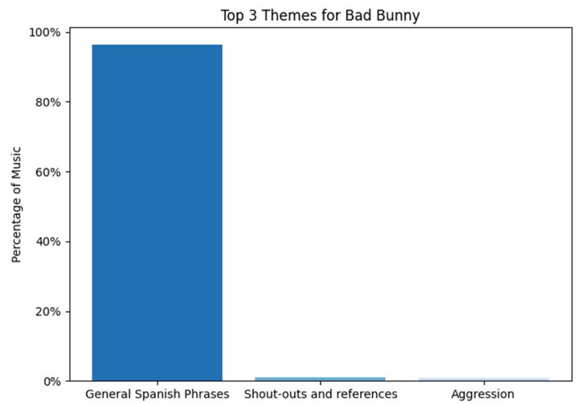


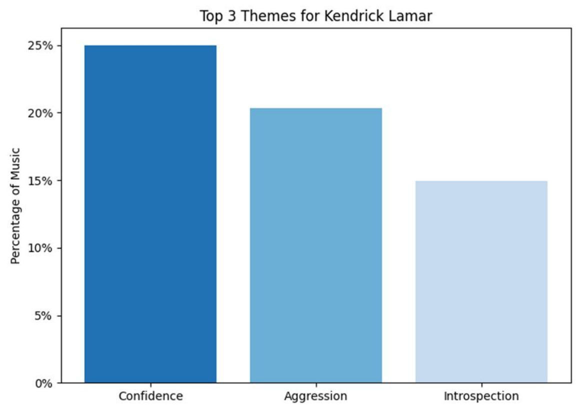


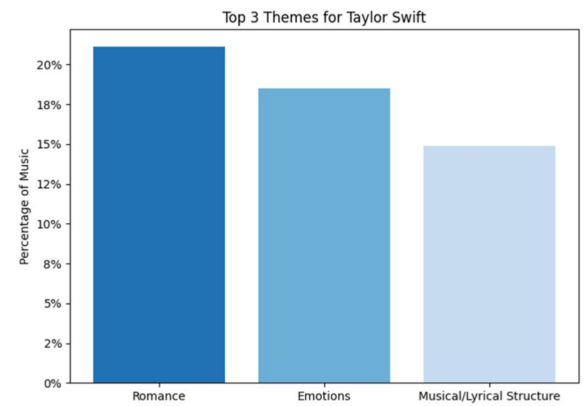


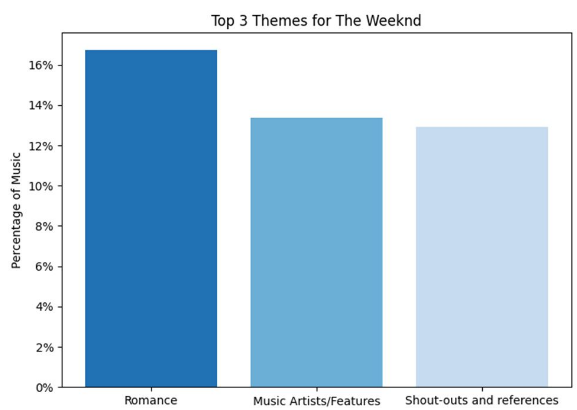


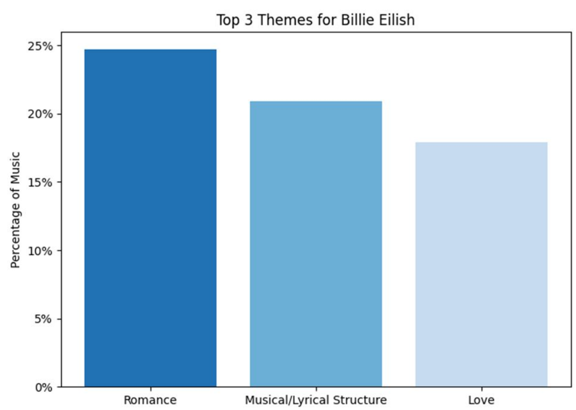
Common Themes in Songs by Artist: Sentence-Level Results (continued)

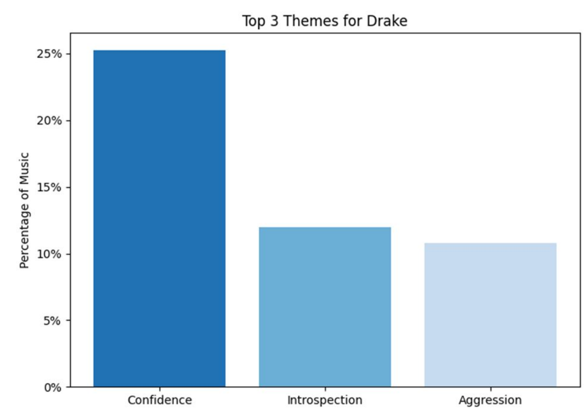












3: Public Perception

Data Sources

Reddit

Link: www.kaggle.com/datasets/alyahmedts13/reddit-sentiment-analysis-dataset-for-nlp-projects

Description: This dataset contains short Reddit posts (≤280 characters) about pop music and pop stars, labeled for sentiment analysis.

Twitter

Link: https://www.kaggle.com/datasets/kazanova/sentiment140/data

Description: This dataset contains 1,600,000 tweets extracted using the Twitter API.

News

Link: https://www.kaggle.com/datasets/rmisra/news-category-dataset

Description: This dataset contains around 210k news headlines from 2012 to 2022 from HuffPost.

Methods

Task Goal: Analyze artists' public perception via media datasets

Pre-processing:

- Filter three datasets (reddit, twitter, news) for target artists (Taylor Swift, Kendrick Lamar, Drake, Billie Eilish, Bad Bunny, The Weeknd)
- Create aggregate dataset

Sentiment:

- Fill in sentiment / add predicted sentiment using twitter-roberta-base-sentiment-latest model
- Create graphs showing these sentiments per artist per media source

RAG:

- Create embeddings using sentence-transformers/all-mpnet-base-v2 and store via FAISS
- Pipeline for evidence retrieval and feeding question + evidence to LLM (Gemini)
- Evaluation using DeepEval for 3 metrics: Answer Relevancy, Faithfulness, and Contextual Relevancy

Code Demo

public_perception.ipynb

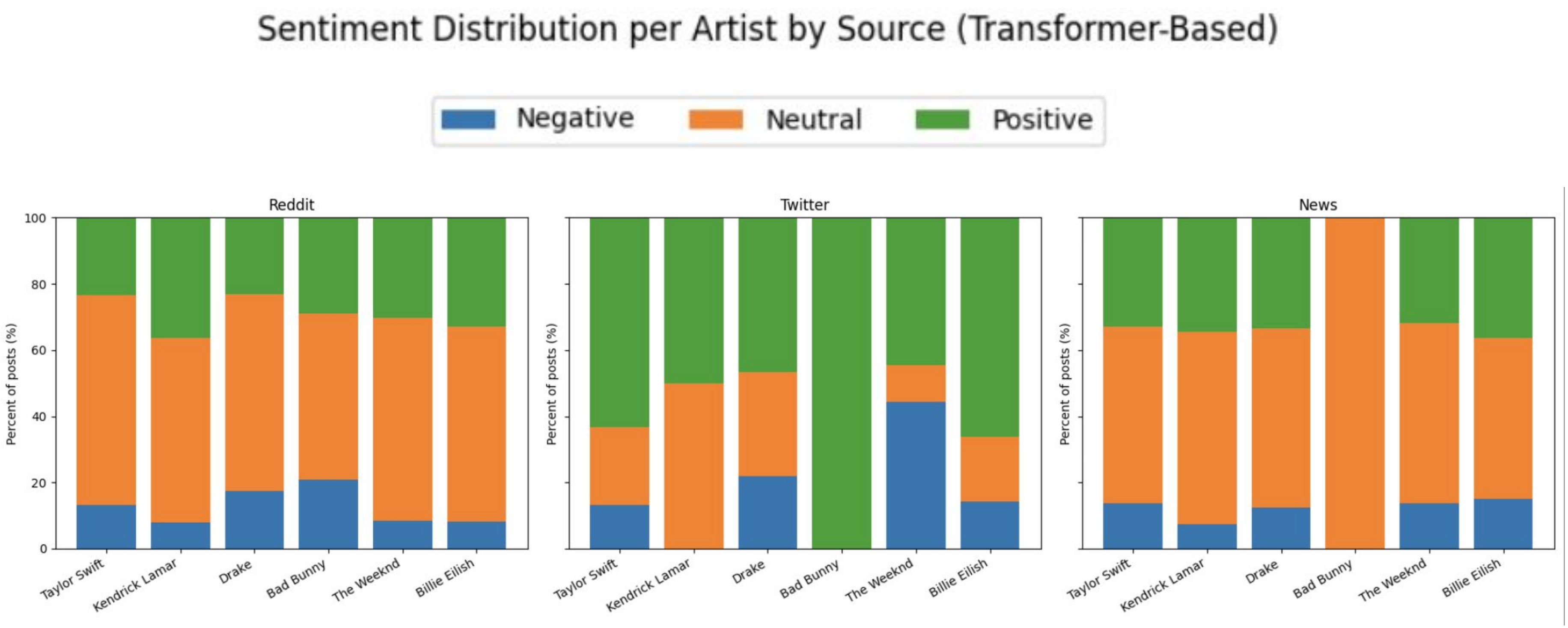
```
!pip install faiss-cpu
!pip -q install sentence-transformers
!pip -q install google-generativeai
!pip -q install deepeval

import os
import re
import pandas as pd

import numpy as np
import torch
from transformers import AutoTokenizer, AutoModelForSequenceClassification
from tqdm.auto import tqdm

import matplotlib.pyplot as plt
import faiss
from sentence_transformers import SentenceTransformer, CrossEncoder
from google.colab import userdata
import google.generativeai as genai
```

```
Requirement already satisfied: faiss-cpu in /usr/local/lib/python3.12/dist-packages (1.13.1)
Requirement already satisfied: numpy<3.0,>=1.25.0 in /usr/local/lib/python3.12/dist-packages (from faiss-cpu) (2.0.2)
Requirement already satisfied: packaging in /usr/local/lib/python3.12/dist-packages (from faiss-cpu) (25.0)
```

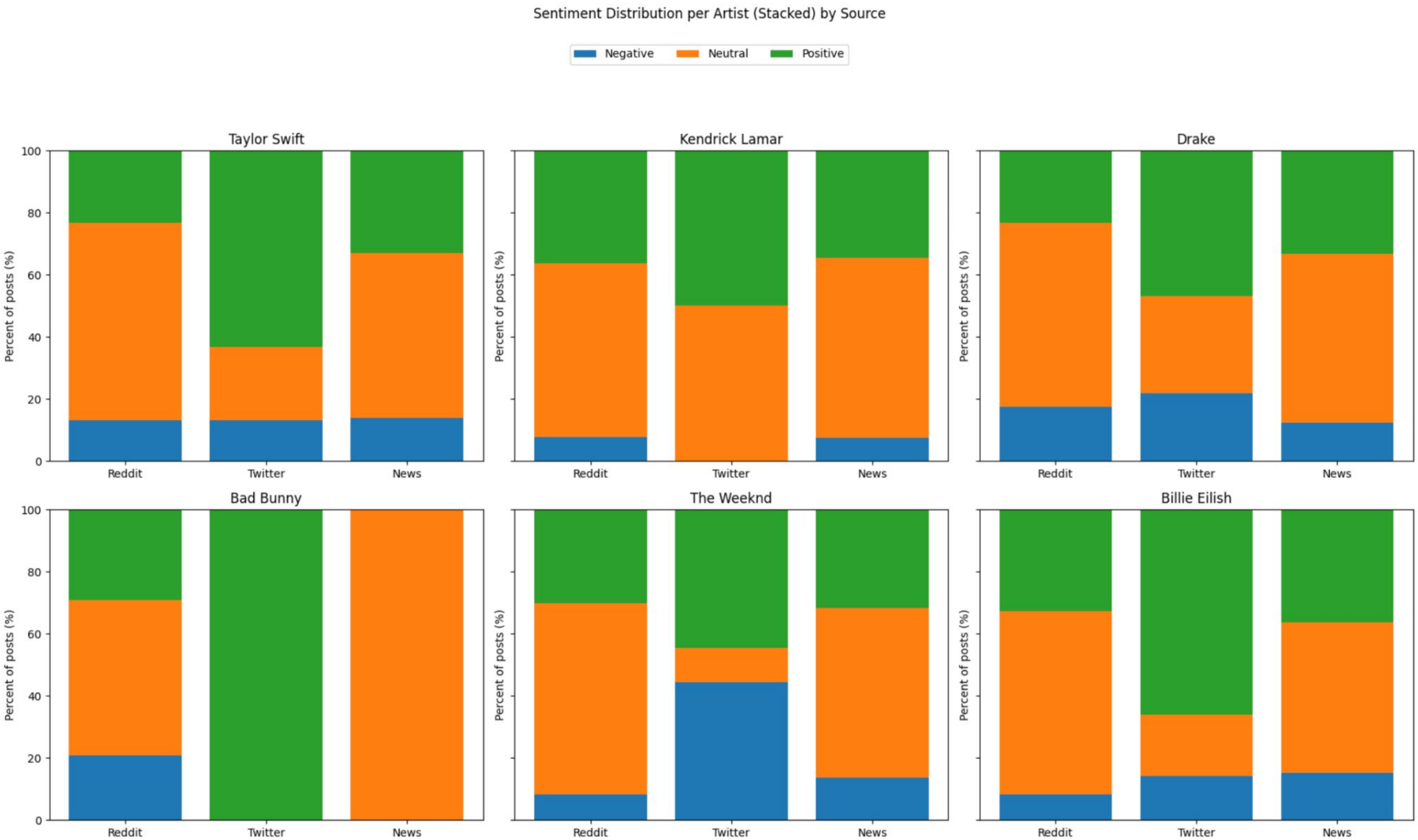
```
kaggle_json_path = "kaggle.json" # import into directory

!mkdir -p ~/.kaggle

if not os.path.exists(kaggle_json_path):
    raise FileNotFoundError(
        "kaggle.json not found in current directory. "
        "Download it from Kaggle → Account → API → Create Token"
    )

!cp {kaggle_json_path} ~/.kaggle/kaggle.json
!chmod 600 ~/.kaggle/kaggle.json

print("Kaggle API setup complete")


!mkdir -p ./data/twitter
!mkdir -p ./data/reddit
!mkdir -p ./data/news
```

Emma Ylagan

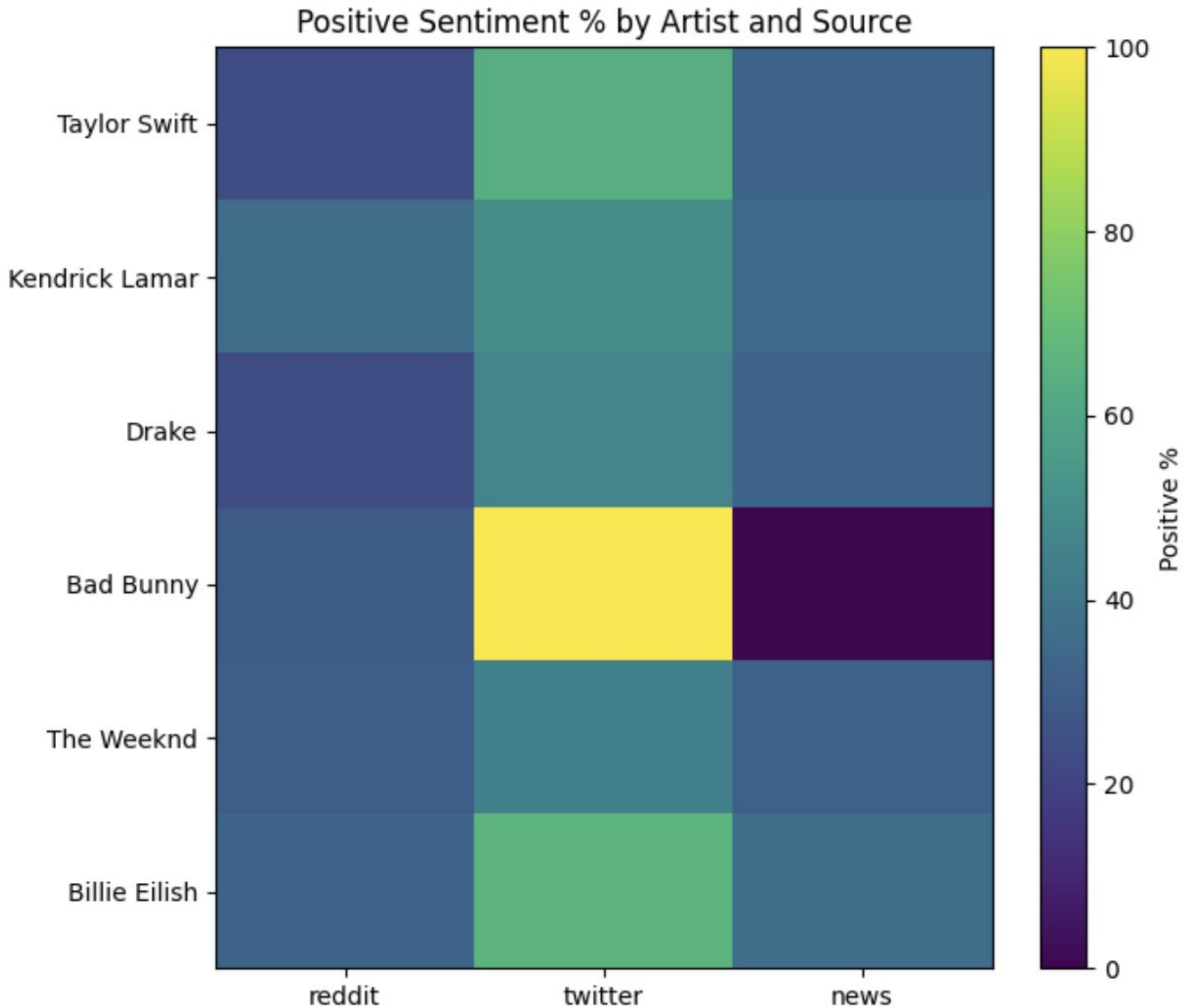
Graph Per Source Per Artist



Graph Per Artist Per Source



Heatmap of Source vs. Artist



Rag Query Pipeline

```
question = "Which artist receives more polarized reactions: Taylor Swift or Billie Eilish?"

answer, evidence, prompt = run_rag_query(
    question=question,
    faiss_index=faiss_index,
    metadata_df=metadata_df,
    embedder=embedder,
    use_rerank=True,
    retrieve_top_k=80,
    rerank_top_k=8,
    k_search=200
)

print(answer)
display(evidence[["source", "title", "sentiment_label_model", "rag_text"]])

print("\n \n")

scores = evaluate_rag_response(
    question=question,
    answer=answer,
    evidence_df=evidence,
)

scores
```

Outputs

[answer] = “Based on the evidence, Taylor Swift appears to receive more polarized reactions than Billie Eilish. The evidence describes an "infamous taylor swift feud" which sparked debate, indicating strong public sentiment and differing opinions surrounding her [3], [5]. Furthermore, there was a "huge 'targeting taylor swift' backlash" [6], suggesting that actions perceived as negative towards Swift can provoke significant public reaction. While Billie Eilish was part of a "debate" [3], [5] and experienced backlash [6], these instances were often framed in relation to Kanye West's actions or a perceived slight against Taylor Swift. The evidence also notes Billie Eilish as "apple music's artist of the year" [7], a positive reception, but lacks explicit mentions of significant opposing negative sentiments comparable to the "infamous feud" associated with Taylor Swift.”

	source	title	sentiment_label_model	rag_text
44	reddit	<NA>	neutral	teens react to billie eilish reacts to teens r...
68	reddit	<NA>	neutral	billie eilish (taylor’s version)
29	reddit	<NA>	negative	kanye west dragging billie eilish sparked a de...
27	reddit	<NA>	neutral	taylor swift vs. billie eilish - you made me d...
14	reddit	<NA>	neutral	kanye west dragging billie eilish sparks debat...
0	reddit	<NA>	neutral	billie eilish sets record straight after huge ...
25	reddit	<NA>	positive	billie eilish is apple music's artist of the y...
1	reddit	<NA>	neutral	billie eilish vs taylor swift? who does what b...

Evaluation Using DeepEval

```
Metrics Summary

  - ✅ Answer Relevancy (score: 0.8888888888888888, threshold: 0.6, evaluation model: gemini-2.5-flash (Gemini), reason: The score is 0.89 because ..., error: None)
  - ✅ Faithfulness (score: 1.0, threshold: 0.6, strict: False, evaluation model: gemini-2.5-flash (Gemini), reason: The score is 1.00 because the actual output is perfectly aligned with the retrieval context! Great job!, error: None)
  - ✅ Contextual Relevancy (score: 0.6666666666666666, threshold: 0.6, strict: False, evaluation model: gemini-2.5-flash (Gemini), reason: The score is 0.67 ..., error: None)
```

4: Fame Over Time

Data Selection

[Kaggle: Spotify Global Music Dataset]

- Data collected via Spotify public API
- Primarily used track_data_final.csv:
  - 7466 unique tracks
  - 2009-2023
- Key variables
  - track popularity: Spotify-defined popularity index (0-100), reflects relative engagement
  - release timing (aggregated to year)
  - artist popularity, artist followers
- Limitations: dataset does not include all tracks for all artists in every year

Spotify Global Music Dataset (2009–2025)

320 Code

Data Card Code (35) Discussion (0) Suggestions (1)

Detail Compact Column 10 of 15 columns

track_id	track_name	track_num...	track_pop...	track_dura...	explicit	artist_name
6pymOcrCnMuCWdgGVTvUgP	3	57	61	213173	False	Britney Spears
21Wc1iJlz2NVcStV5fbtPG	Clouds	1	67	158760	False	BUNT.
1msEuwSBneBKpVCZQcFTsU	Forever & Always (Taylor's Version)	11	63	225328	False	Taylor Swift
7bcy34fBT2ap1L4bfPsl9q	I Didn't Change My Number	2	72	158463	True	Billie Eilish
0GLfodYacy3BJE7AI3A8en	Man Down	7	57	267013	False	Rihanna
7H0ya83CMmgFcOhw0UB6ow	Space Song	3	77	320466	False	Beach House
41zXlQxzTi6cGAjpOXyLYH	idontwannabeyou anymore	2	78	203569	False	Billie Eilish

Methods

Task Goal: Model artist historical popularity and forecast artist's popularity dynamics

Pre-processing:

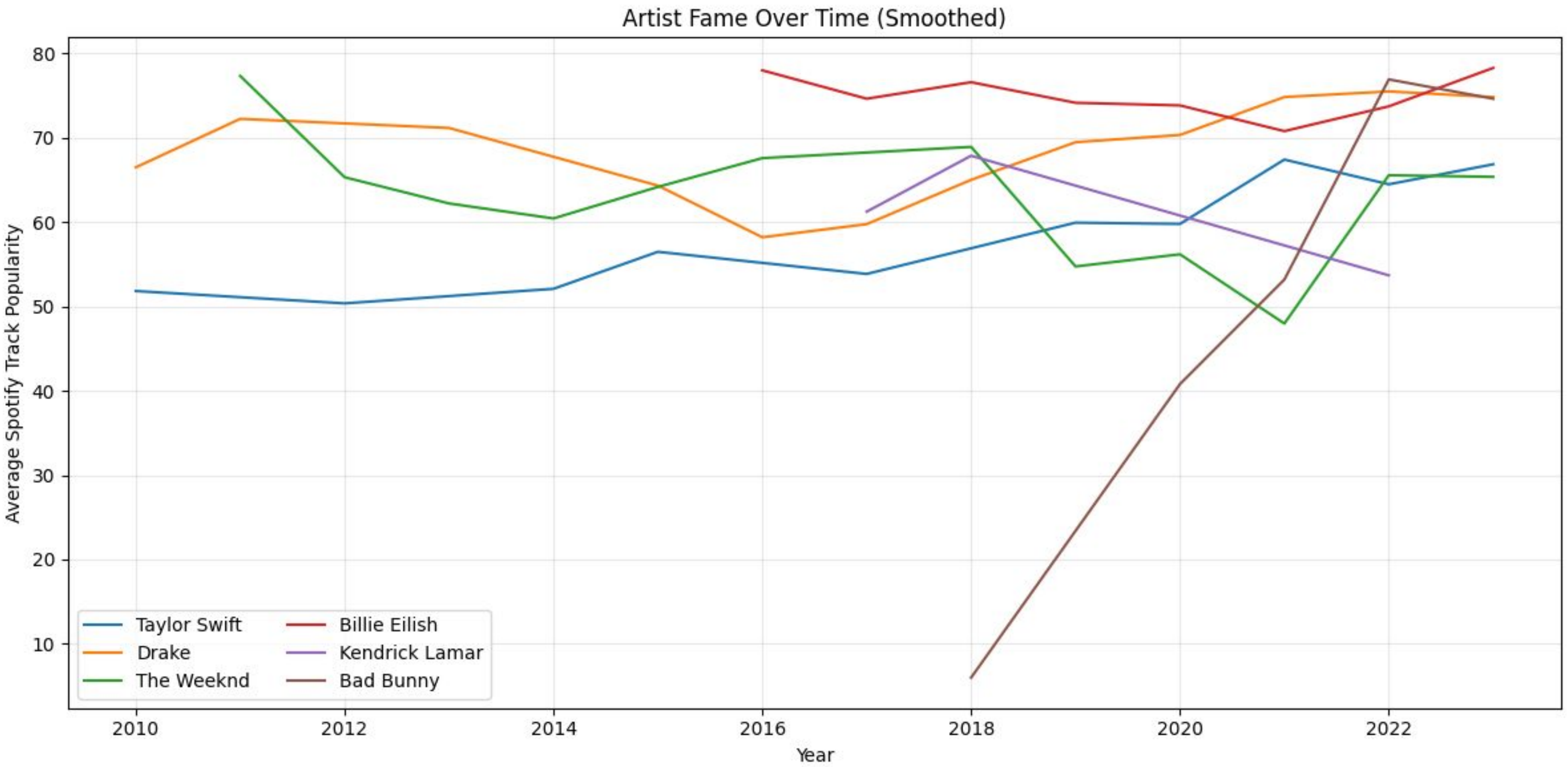
- Filter for only Taylor Swift, Kendrick Lamar, Drake, Billie Eilish, Bad Bunny, The Weeknd
- Converted album release dates to year
- Aggregated tracks to artist-year observations
- Computed rolling averages to smooth short-term noise

Main Process:

1. Estimated trends of fame over time using OLS
2. Identified fame spike years for each artist historically
3. Applied SARIMAX models to artist-level yearly fame series
    a. Forecast horizon: 5 years
    b. Pooled time-series models for artists with sparse histories
    c. Evaluated by out-of-sample hold-out validation using mean absolute error (MAE)

Code Demo

Historical Analysis: Artist Fame Over Time



Results: OLS Baseline - Historical Popularity Trends by Artist

OLS Baseline:

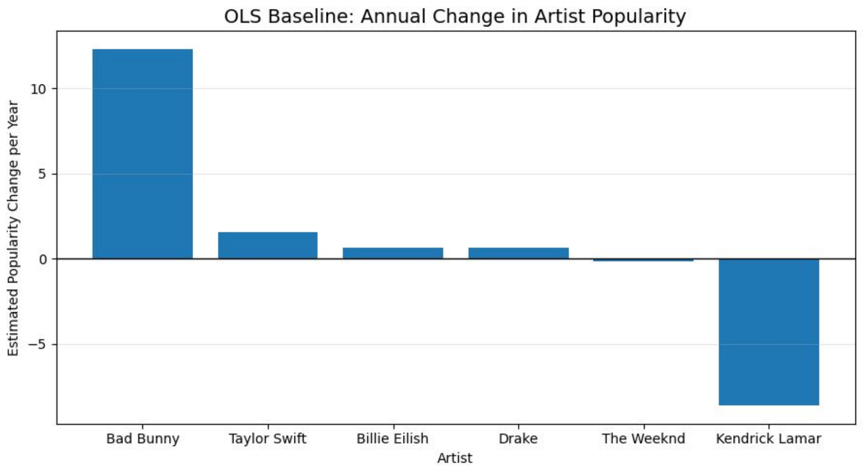
- Avg. annual change in Spotify popularity estimated using OLS
- R^2: How well a linear trend explains each artist's historical popularity

Key Takeaways:

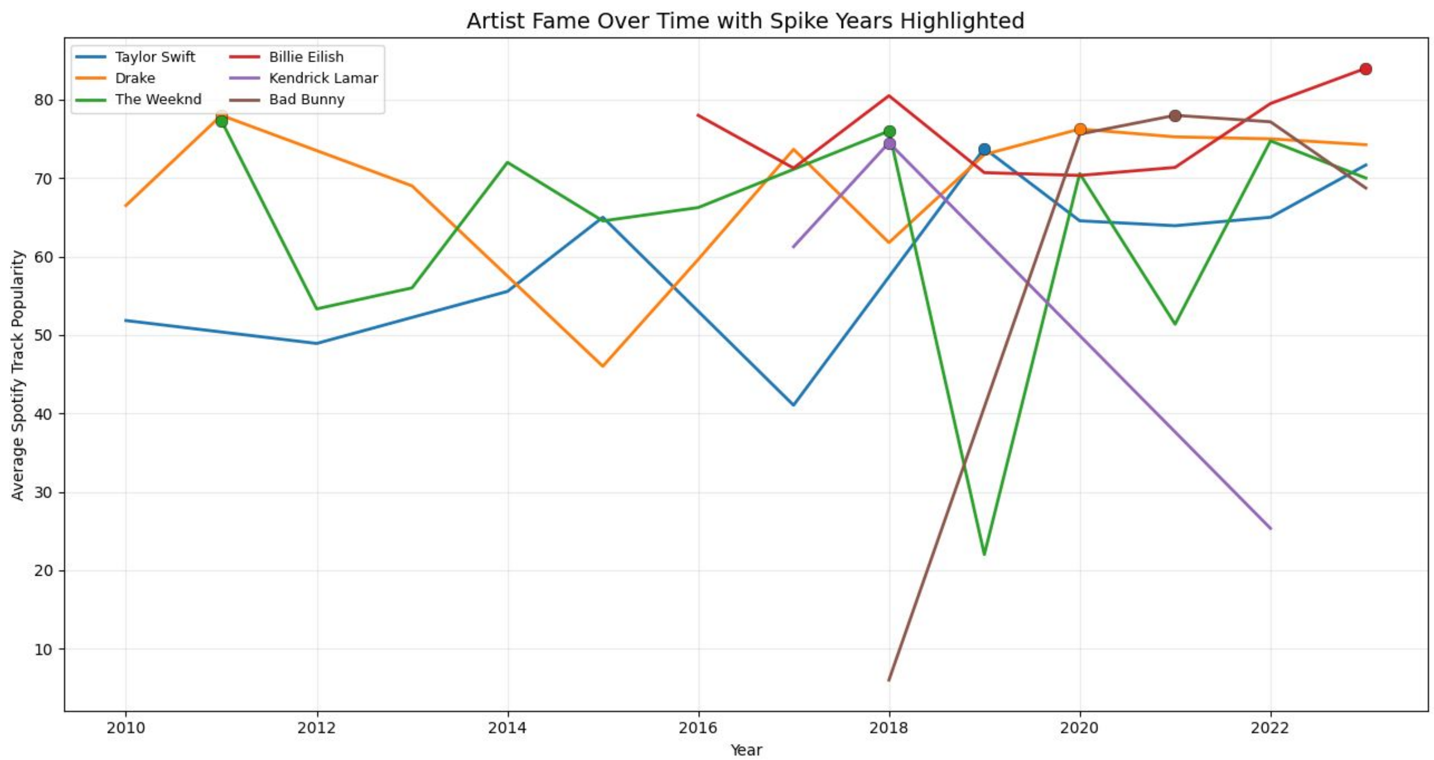
- Largest positive slope: Bad Bunny (+12.3/yr)
  - rapid growth
- Taylor Swift: steady sustained growth (+1.5/yr)
- Drake: low R^2 (0.09)
  - high but volatile popularity, not well captured by a linear trend
- The Weeknd: no consistent linear trend
  - popularity fluctuates with releases
- Kendrick Lamar: sparse and uneven data
  - outlier - dataset insufficient.

OLS shows the need for dynamic time-series models to capture nonlinear popularity patterns

Artist	Estimated Annual Popularity Change	R^2
Taylor Swift	1.52	0.41
Drake	0.65	0.09
The Weeknd	-0.19	0
Billie Eilish	0.66	0.09
Kendrick Lamar	-8.64	0.81
Bad Bunny	12.31	0.58



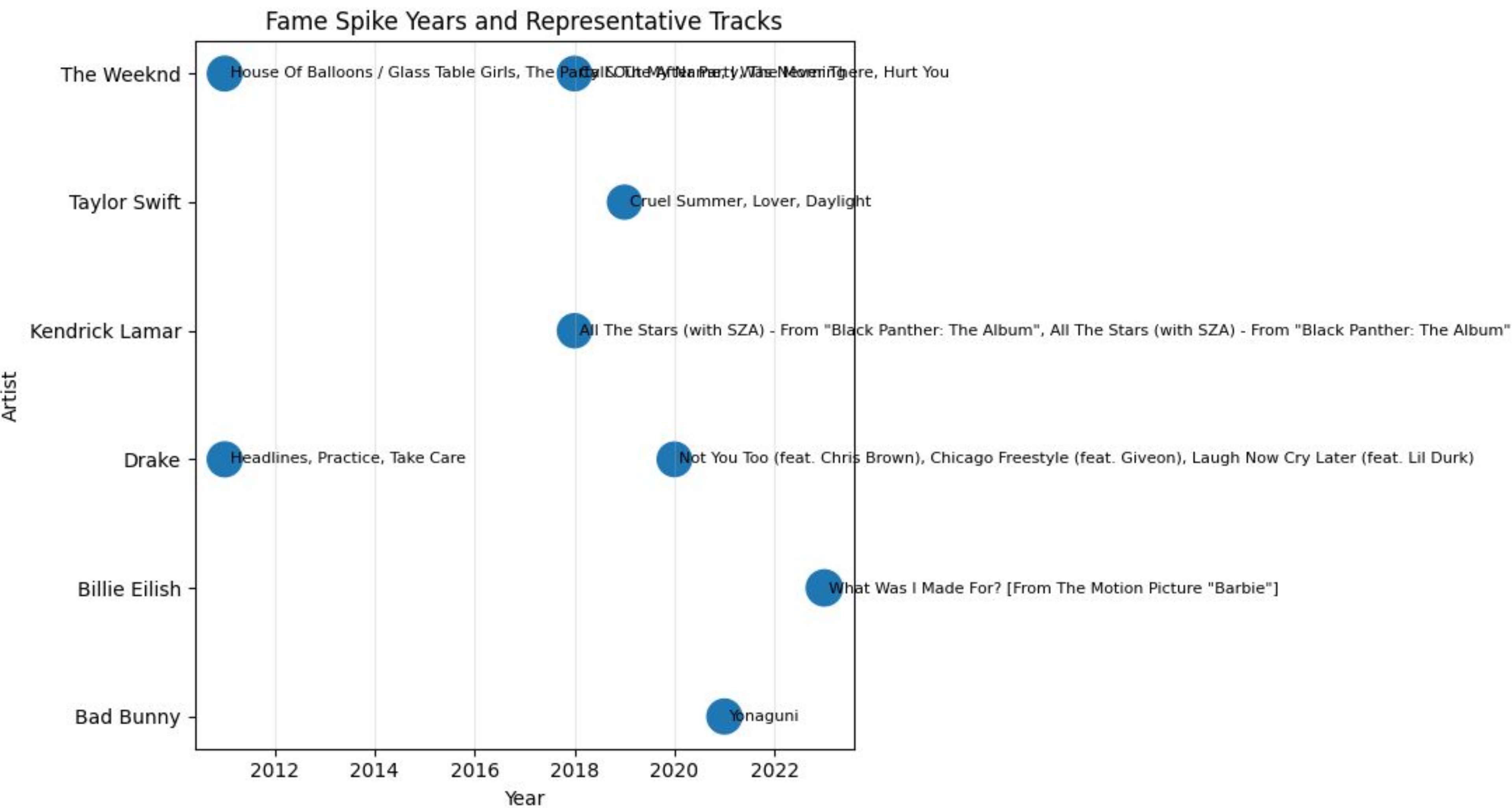
Results: Fame Spike Detection



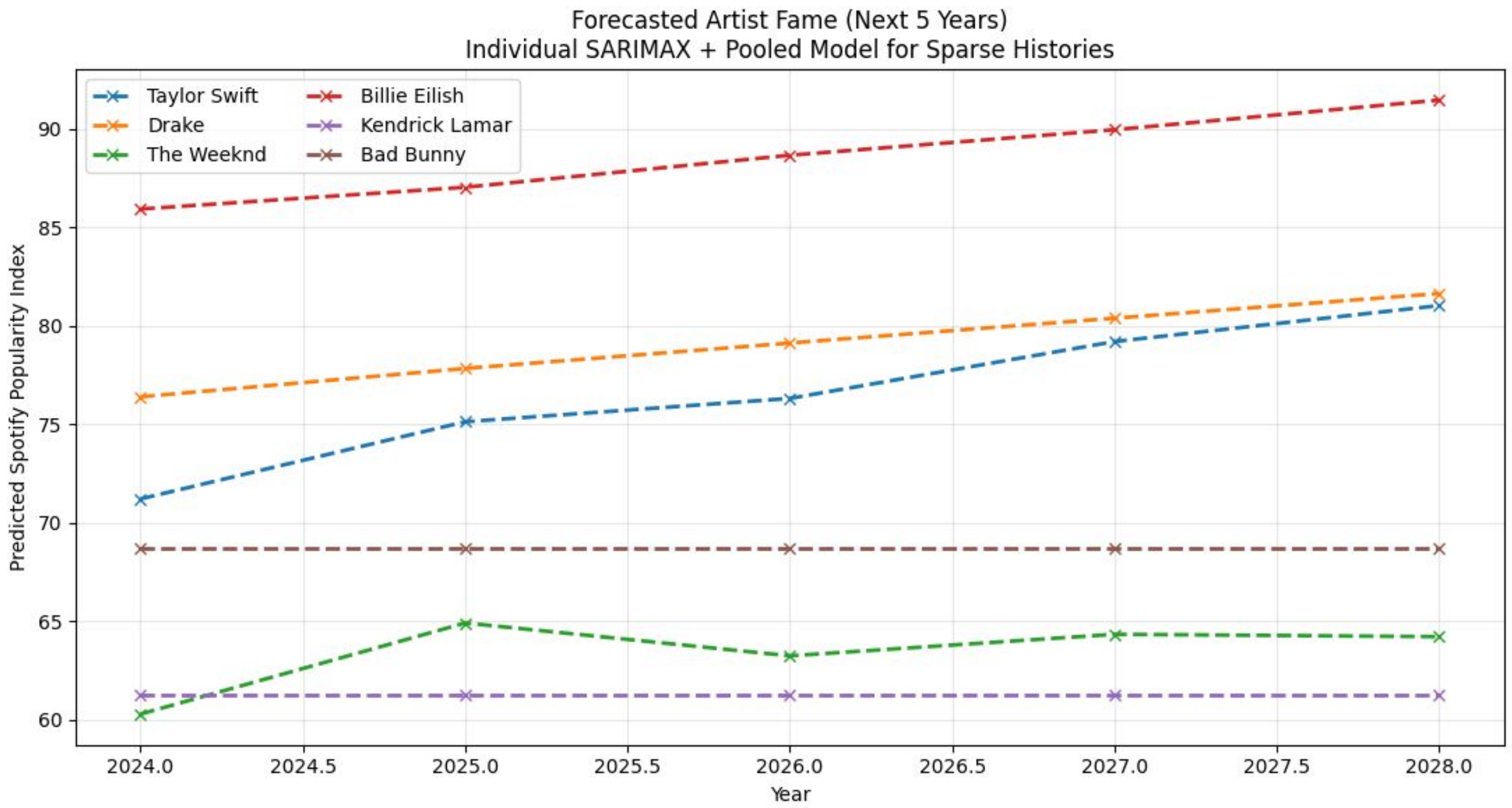
Historical Fame Spikes:

- Spike = a year where the artist reached top 10% of their own popularity

Results: Representative Tracks for Each Spike



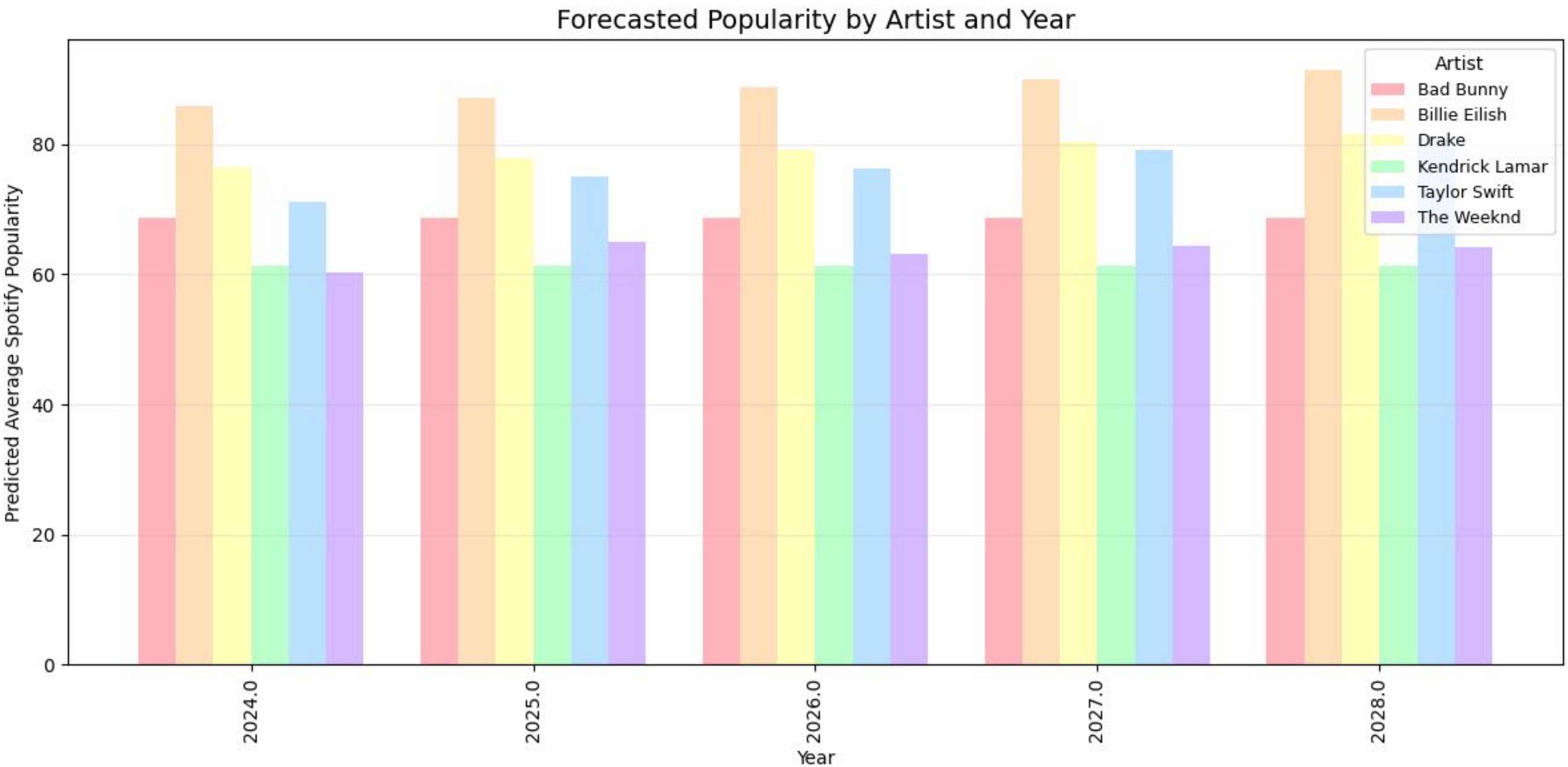
Results: Time-Series Forecasting with SARIMAX



SARIMAX:

- time series model that uses past popularity patterns to forecast future values
- note: pooled model was applied for Bad Bunny and Kendrick Lamar: different baseline popularity levels but shared time-series behavior

Who is predicted to be most popular in the next five years?

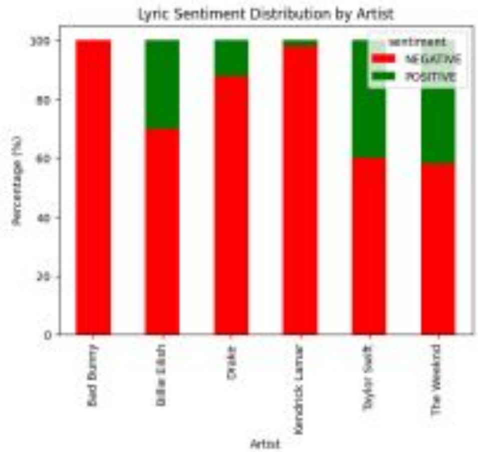


Data Summary: Drake

Drake

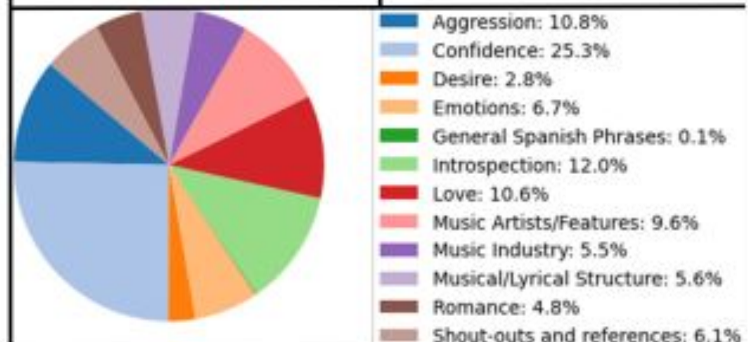
- **generally negative lyrics**
- **primarily energetic, "hype" music with some deeper messages**
- **mixed public perception, leaning positive**
- **fame has steady linear increase over time, projected to 2028**

Lyrics

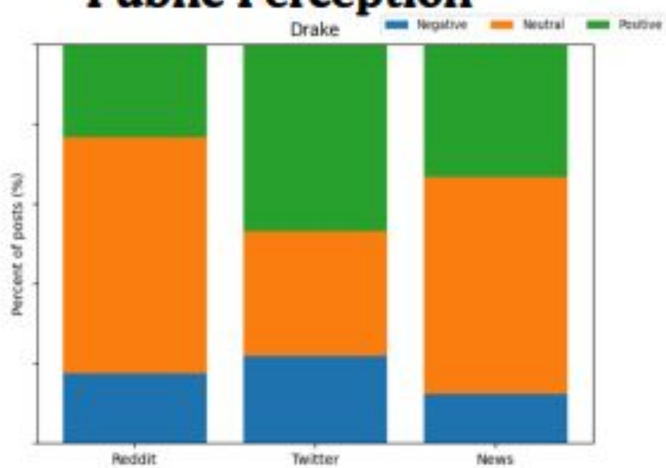


Drake's "Top 3":
- Confidence
- Introspection
- Aggression

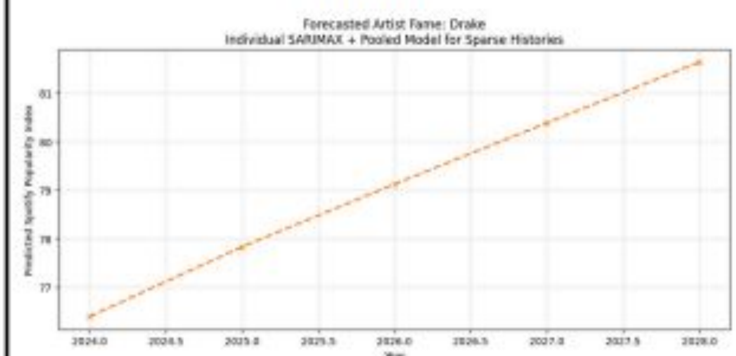
Theme Distribution



Public Perception



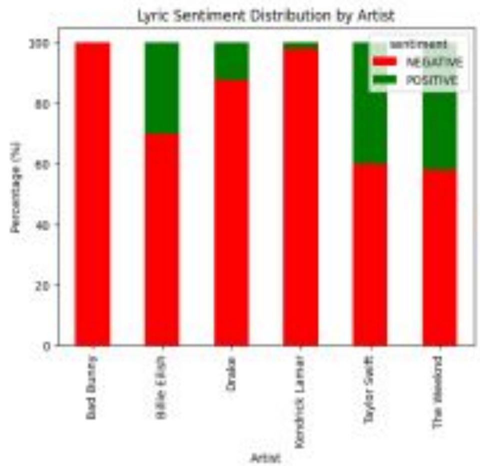
Fame



Data Summary: Kendrick Lamar

Kendrick Lamar

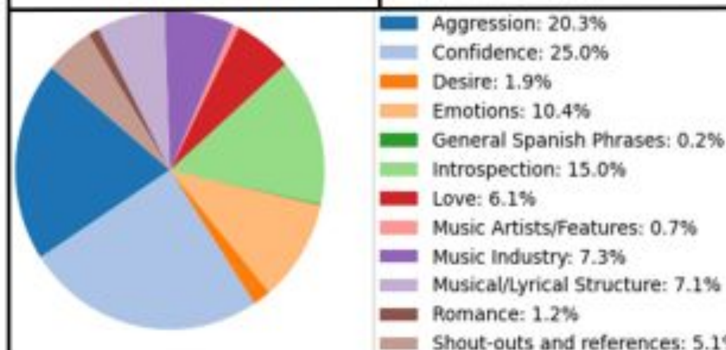
- **mostly negative lyrics**
- **primarily energetic, “hype” music with some deeper messages**
- **mixed public perception**
- **fame has generally stayed at 62 pts over time**



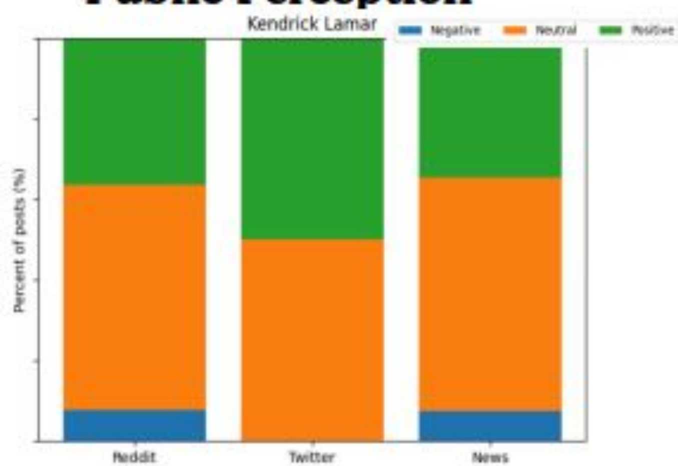
Lyrics

Kendrick's “Top 3”:
- Confidence
- Aggression
- Introspection

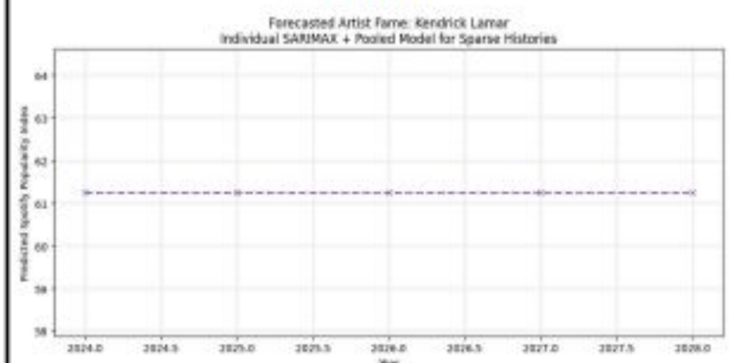
Theme Distribution



Public Perception



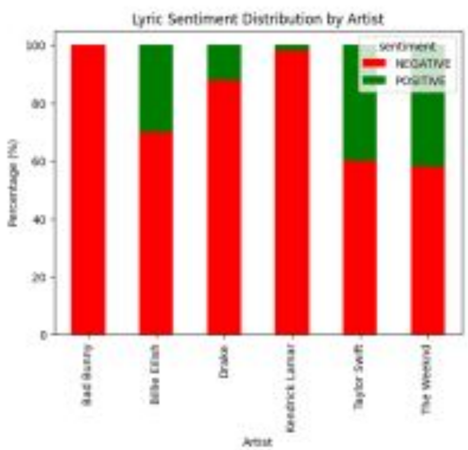
Fame



Data Summary: The Weeknd

The Weeknd

- somewhat negative lyrics
- primarily romantic music, with common references to pop culture
- mixed public perception, leaning negative/neutral
- fame has extreme variation but overall increase (peak in 2025)

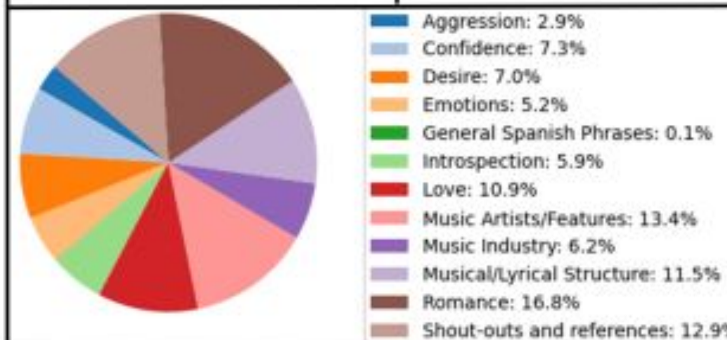


Lyrics

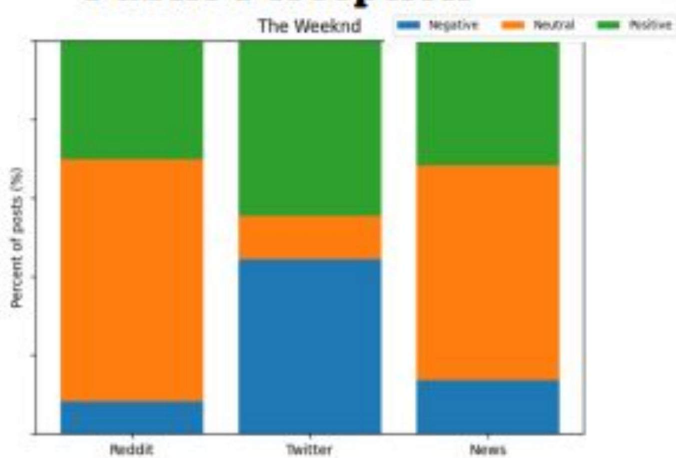
The Weeknd's "Top 3":

- Romance
- Music Artists & Features
- Shout-outs & References

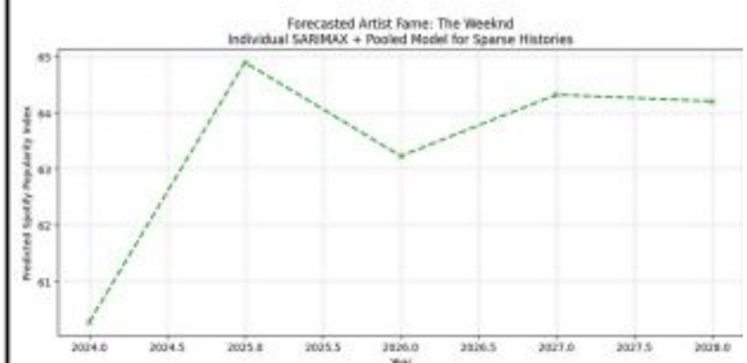
Theme Distribution



Public Perception



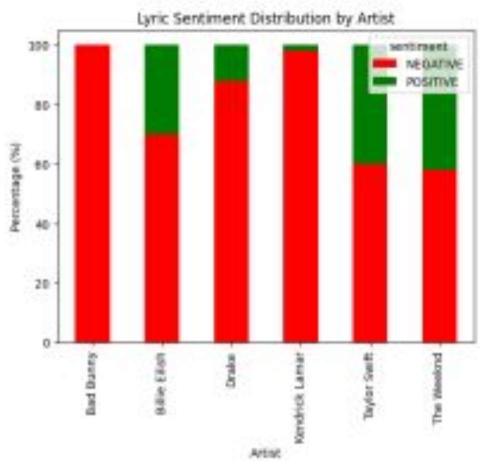
Fame



Data Summary: Billie Eilish

Billie Eilish

- **generally negative lyrics**
- **primarily soft music with themes of love and lyrical structure**
- **generally positive public perception**
- **predicted fame to steadily increase over time (+1pt/year)**

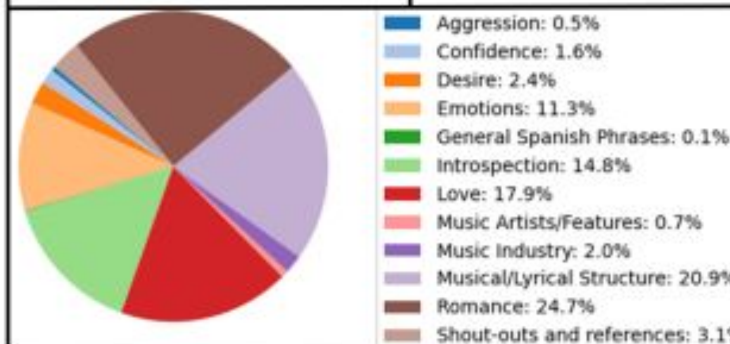


Lyrics

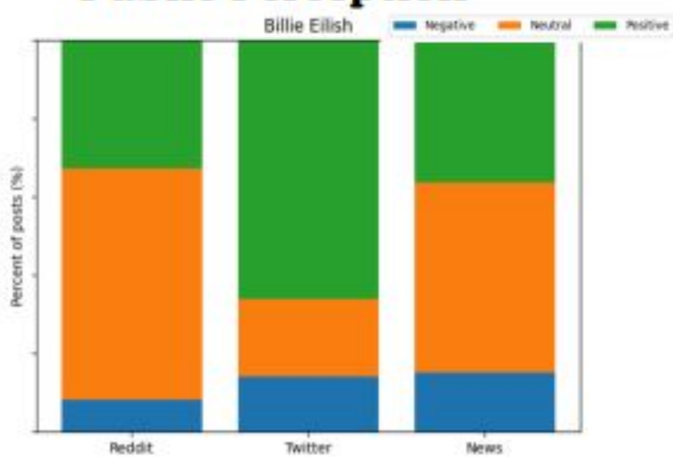
Billie's "Top 3":

- Romance
- Musical/Lyrical Structure
- Love

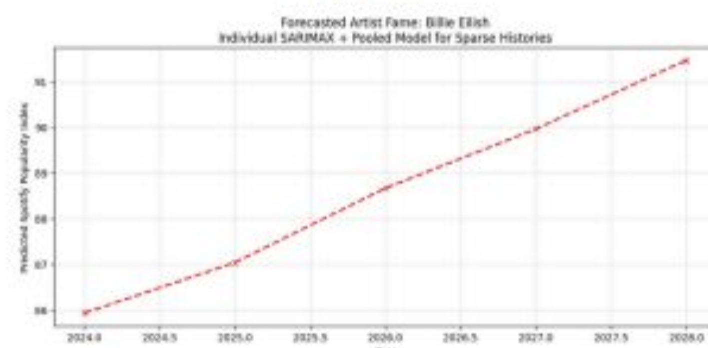
Theme Distribution



Public Perception



Fame

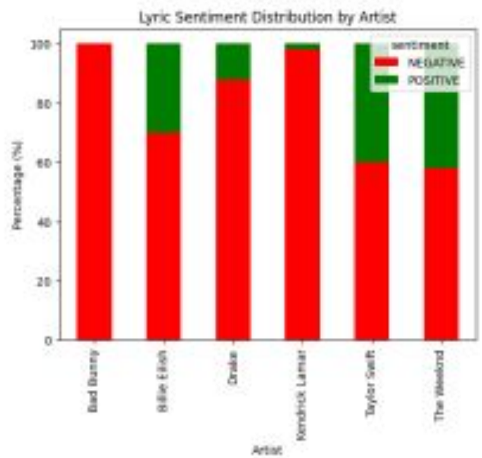


Data Summary: Taylor Swift

Taylor Swift

- generally negative lyrics
- primarily emotional music with themes of romance
- mixed public perception, leaning positive
- fame has some variability but increases +8 pts by 2028

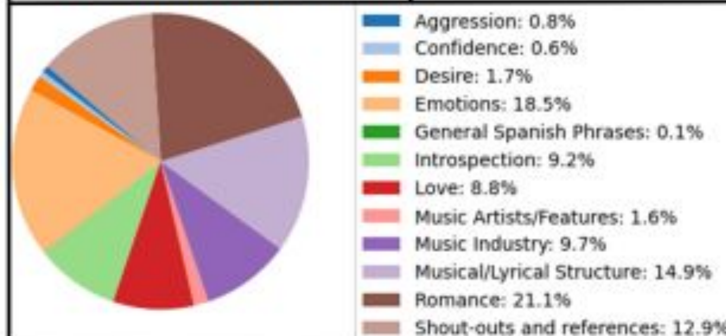
Lyrics



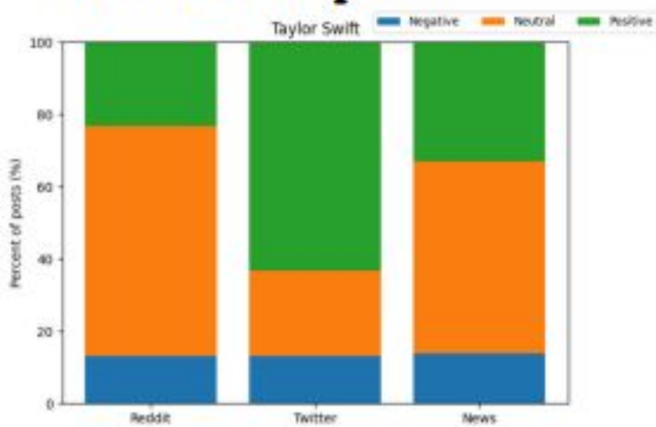
Taylor's "Top 3":

- Romance
- Emotions
- Musical/Lyrical Structure

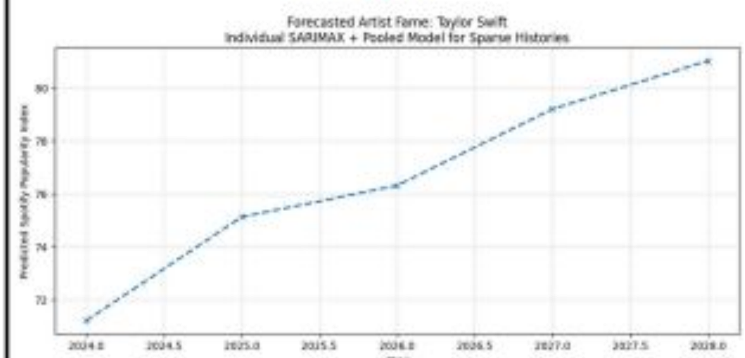
Theme Distribution



Public Perception

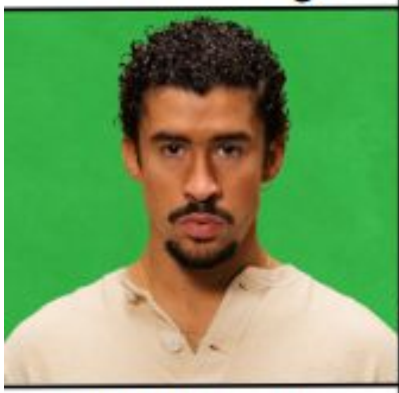


Fame



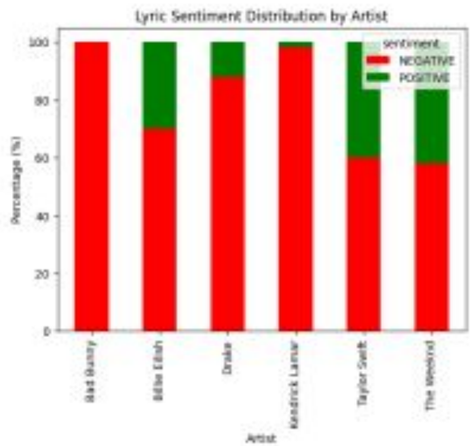
Data Summary: Bad Bunny

Bad Bunny



- generally negative lyrics
- almost entirely Spanish lyrics
- extremely mixed public perception, leaning positive
- fame has steadily maintained over time, around 68.8 pts.

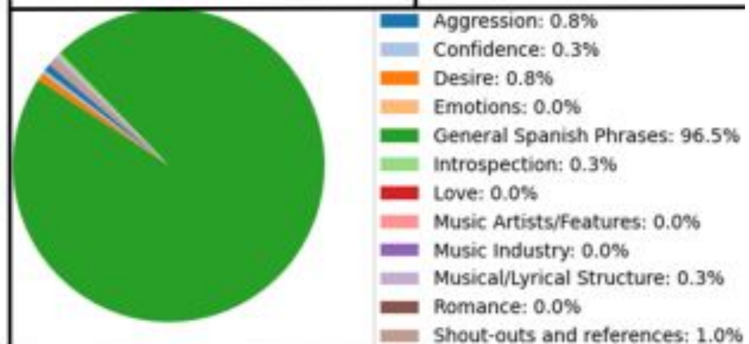
Lyrics



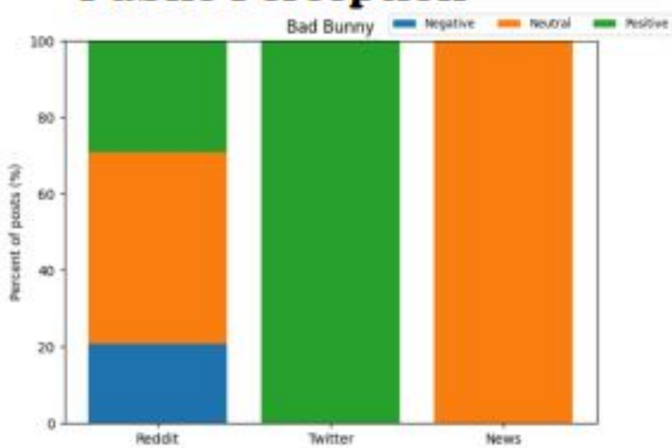
Bad Bunny's "Top 3":

- Spanish Phrases
- Shout-outs & References
- Aggression

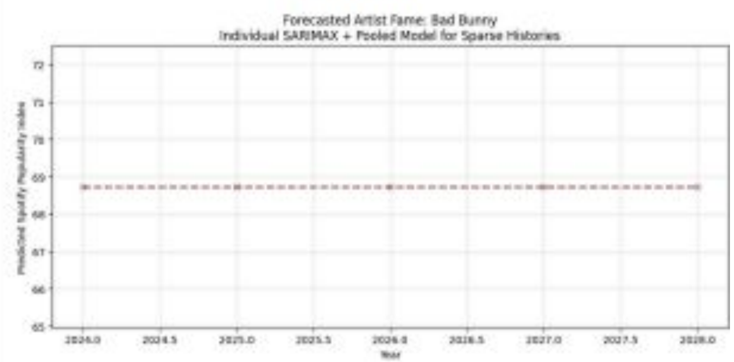
Theme Distribution



Public Perception



Fame



Conclusion & Challenges

What we accomplished:

- Built comprehensive, multi-dimensional artist profiles for six globally popular artists by integrating:
  - **Lyric sentiment** analysis using transformer-based models
  - **Thematic analysis** of lyrics via sentence embeddings and clustering
  - **Public perception** modeling across Reddit, Twitter, and News using sentiment analysis and RAG
  - **Fame-over-time** modeling and forecasting using Spotify popularity data and SARIMAX

Challenges:

- Data limitations & sparsity
- Sentence vs. song-level tradeoffs
- Language mismatch

Future Work

- ➢ Extend the pipeline to a larger and more diverse set of artists
- ➢ Apply sentiment analysis to larger scope (genres instead of artists), or use them to analyze a specific person's taste
- ➢ Apply different models for languages (ex: Bad Bunny's analysis)
- ➢ Incorporate richer time-series data (streaming counts, chart positions, social engagement) to improve trend modeling
- ➢ Build artist-to-artist comparison profiles (e.g., Drake vs. Kendrick) across sentiment, themes, and popularity, or deep dive into specific cultural events such as rap battles or controversy to see the effect of pop culture on artist perception, popularity, and general branding changes

References

Carlos G. D. C. J. (n.d.). *Genius song lyrics with language information* [Data set]. Kaggle. https://www.kaggle.com/datasets/carlosgdcj/genius-song-lyrics-with-language-information

Wardabilal. (n.d.). *Spotify global music dataset (2009–2025)* [Data set]. Kaggle. https://www.kaggle.com/datasets/wardabilal/spotify-global-music-dataset-20092025

Alyahmedts13. (n.d.). *Reddit sentiment analysis dataset for NLP projects* [Data set]. Kaggle. https://www.kaggle.com/datasets/alyahmedts13/reddit-sentiment-analysis-dataset-for-nlp-projects

Go, A., Bhayani, R., & Huang, L. (2009). *Sentiment140* [Data set]. Kaggle. https://www.kaggle.com/datasets/kazanova/sentiment140

Misra, R. (n.d.). *News category dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/rmisra/news-category-dataset

Rakib Ul Karim, S. M., Rasul, R. A., & Sultana, T. (2025). *Sentiment analysis of social media data for predicting consumer behavior trends using machine learning* (arXiv:2510.19656) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2510.19656

Choudhary, Y., Rao, P., & Bhattacharyya, P. (2025). *Who will top the charts? Multimodal music popularity prediction via adaptive fusion of modality experts and temporal engagement modeling* (arXiv:2512.06259) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2512.06259

Martinez, A., Jr. (2024, March 11). *How machine learning is transforming music marketing and trends analysis*. Flourish & Prosper. https://flourishprosper.net/music-resources/how-machine-learning-is-transforming-music-marketing-and-trends-analysis/

Who Did What

Ibaad:

- Section 1 - Lyric Sentiment (code implementation, visualization, slides, demo)
- Motivation/Background

Mayna:

- Section 2 - Themes (code implementations, visualizations, slides, demo)
- Claim/Target Task
- Why Claim
- Data Summary Slides

Emma:

- Section 3 - Public Perception Analysis (code implementations, visualizations, slides, demo)
- Proposed Solution
- Conclusion & Challenges

Elina:

- Section 4 - Fame Over Time (code implementations, visualizations, slides, demo)
- Related Work
- Future Work