

# Capstone Project: Album Recommendation System



Luis Fernando Silva

# Understanding the problem

## Problem Statement

How can a streaming platform recommend a series of albums based on your favorite albums accurately.

## Context

The streaming platform wants to make sure it can recommend music based on an album you like, even for first-time users.

## Criteria for Success

The model correctly groups together albums that might be from the same artist or genre to predict music you might like.

# The Data: Top 5000 Albums of All Time

## Categorical Features

- Artists Name: 'ars\_name'
- Album Name: 'album'
- Release Date: 'rel\_date'
- Genres: 'gens',
- Descriptions: 'descs',

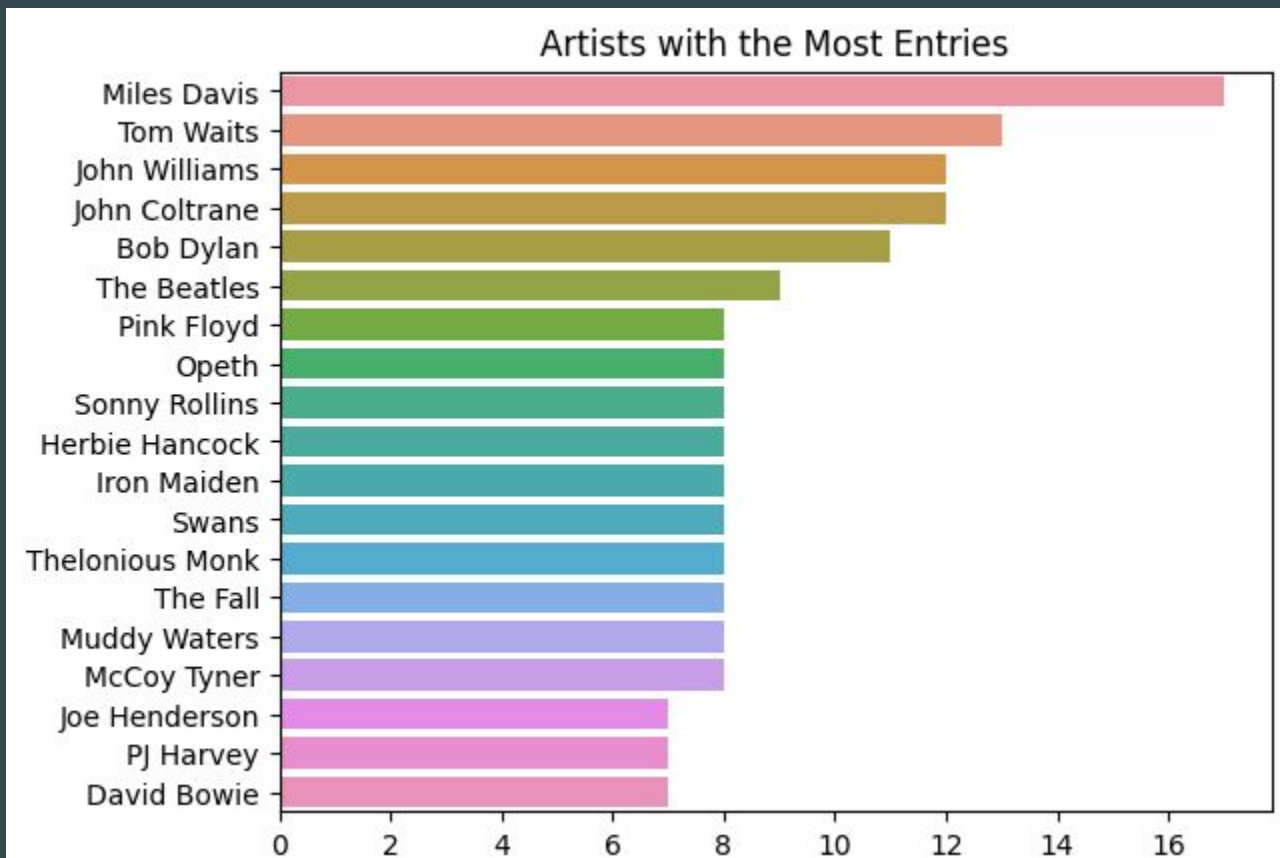
---

# The Data: Top 5000 Albums of All Time

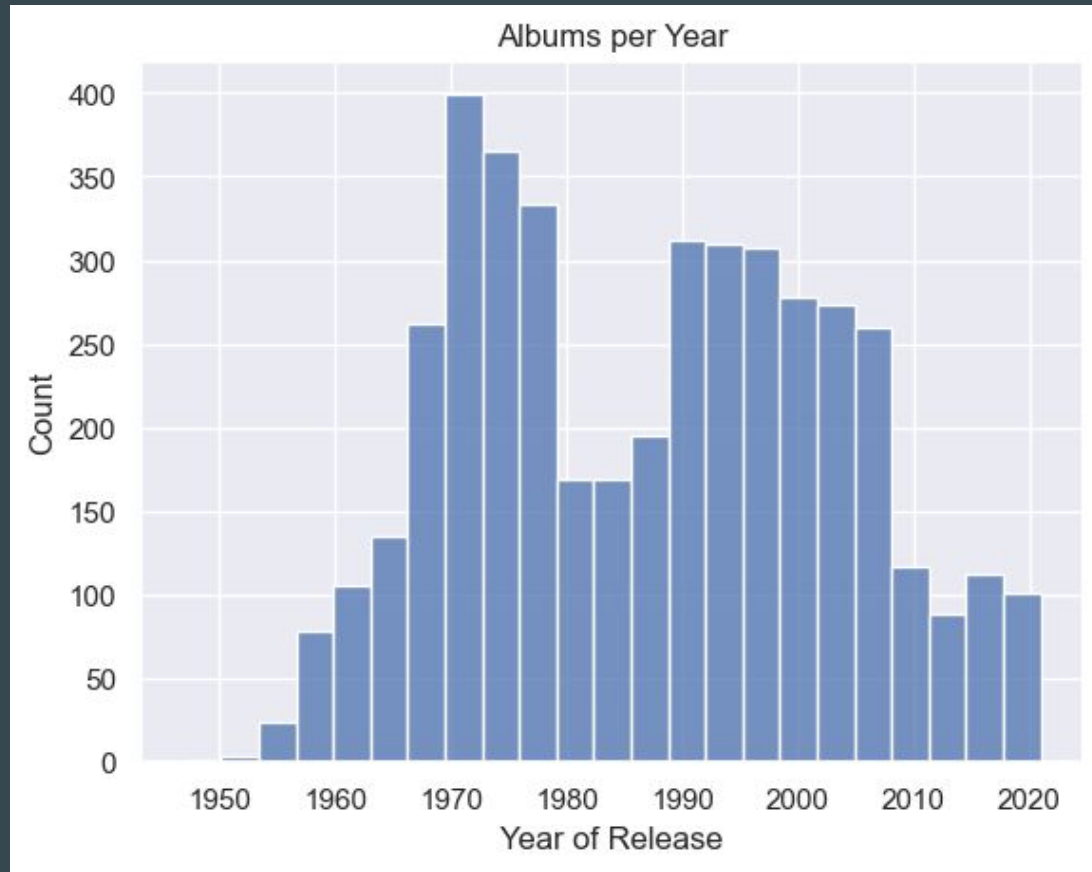
## Numerical Features

- 'avg\_rat'
  - 'num\_rat'
  - 'num\_revs'
  - 'acousticness'
  - 'danceability'
  - 'energy'
  - 'instrumentalness',
  - 'liveness',
  - 'loudness',
  - 'speechiness',
  - 'tempo',
  - 'valence',
  - 'duration\_ms',
  - 'time\_signature',
-

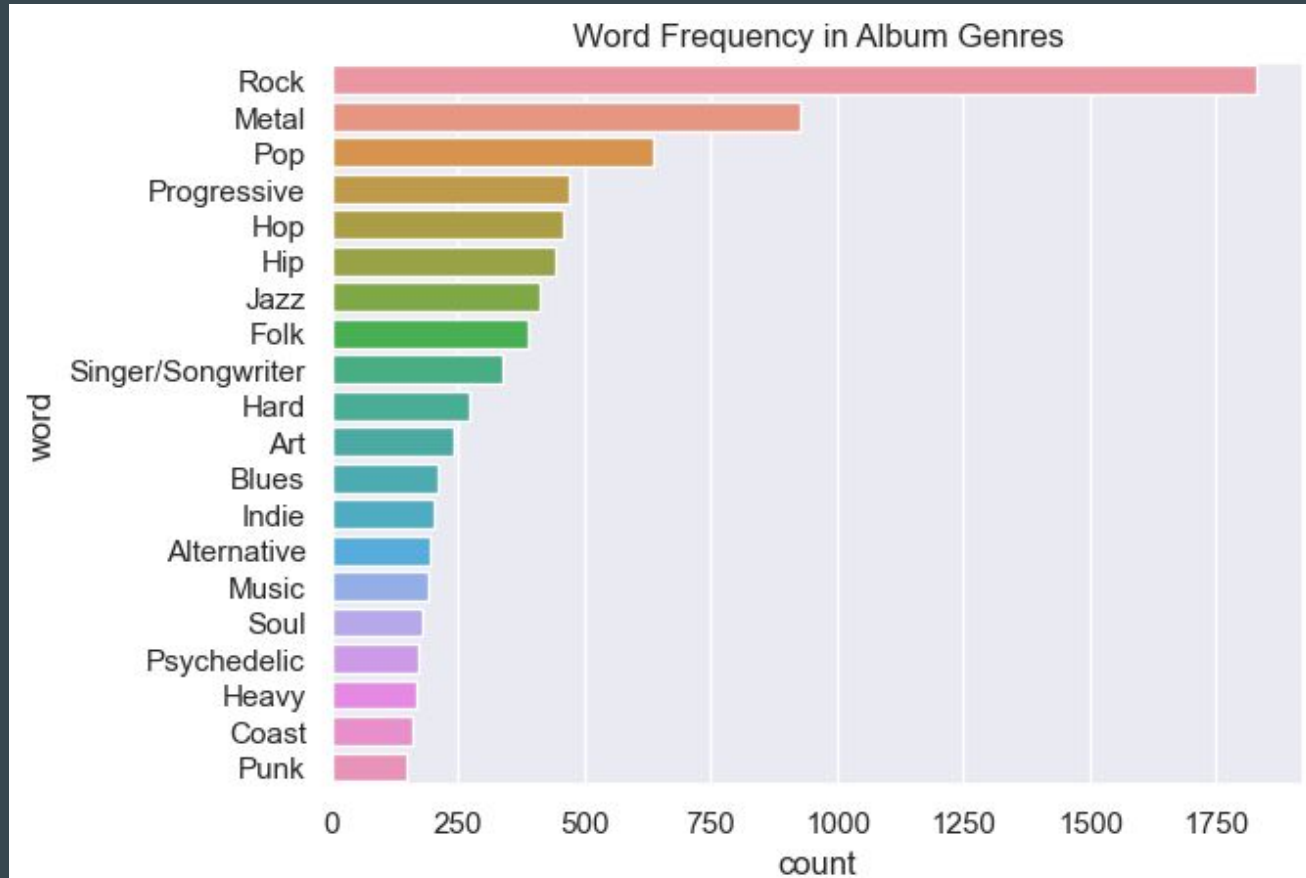
# Artists



# Release Date

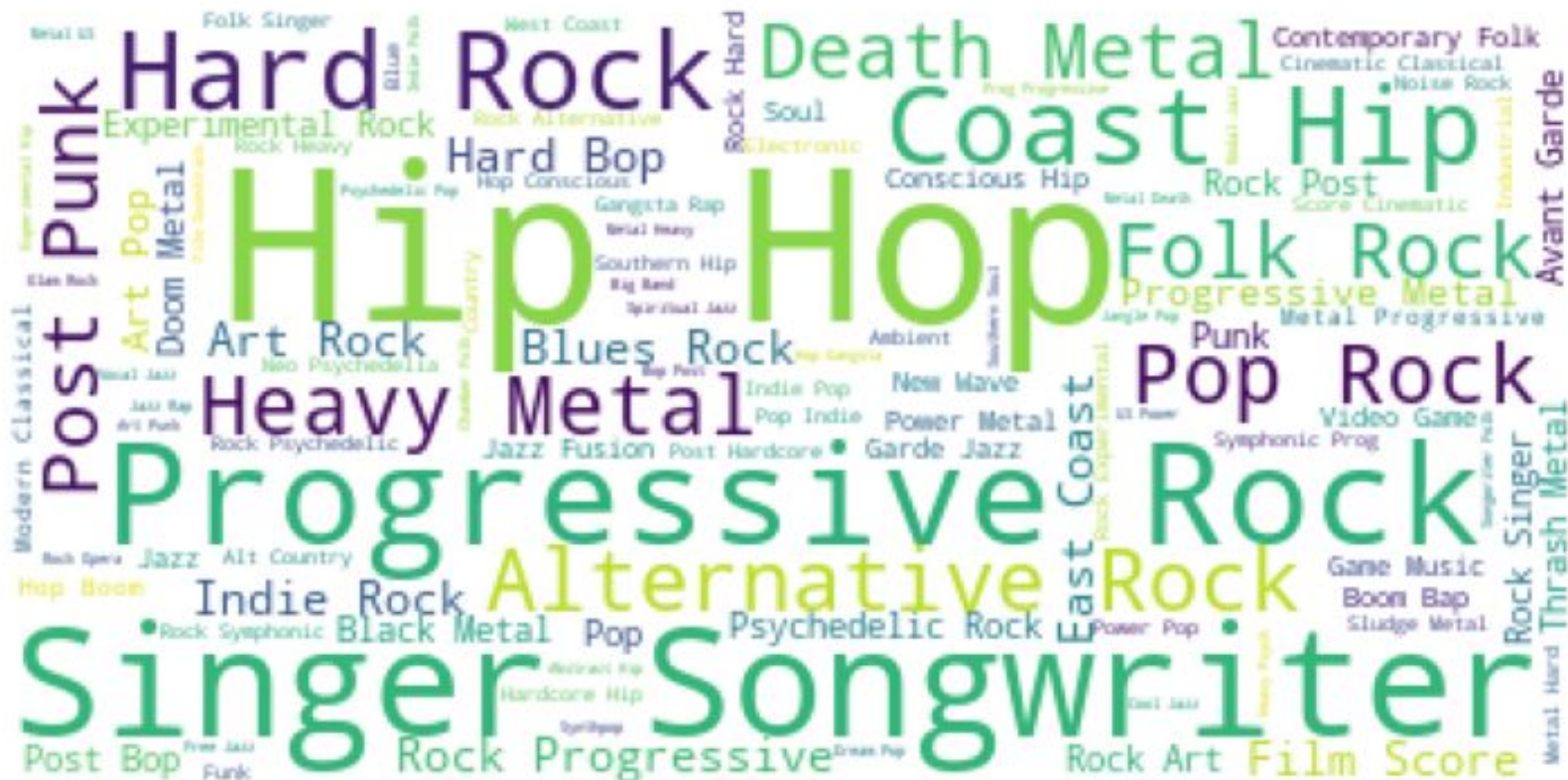


# Genres



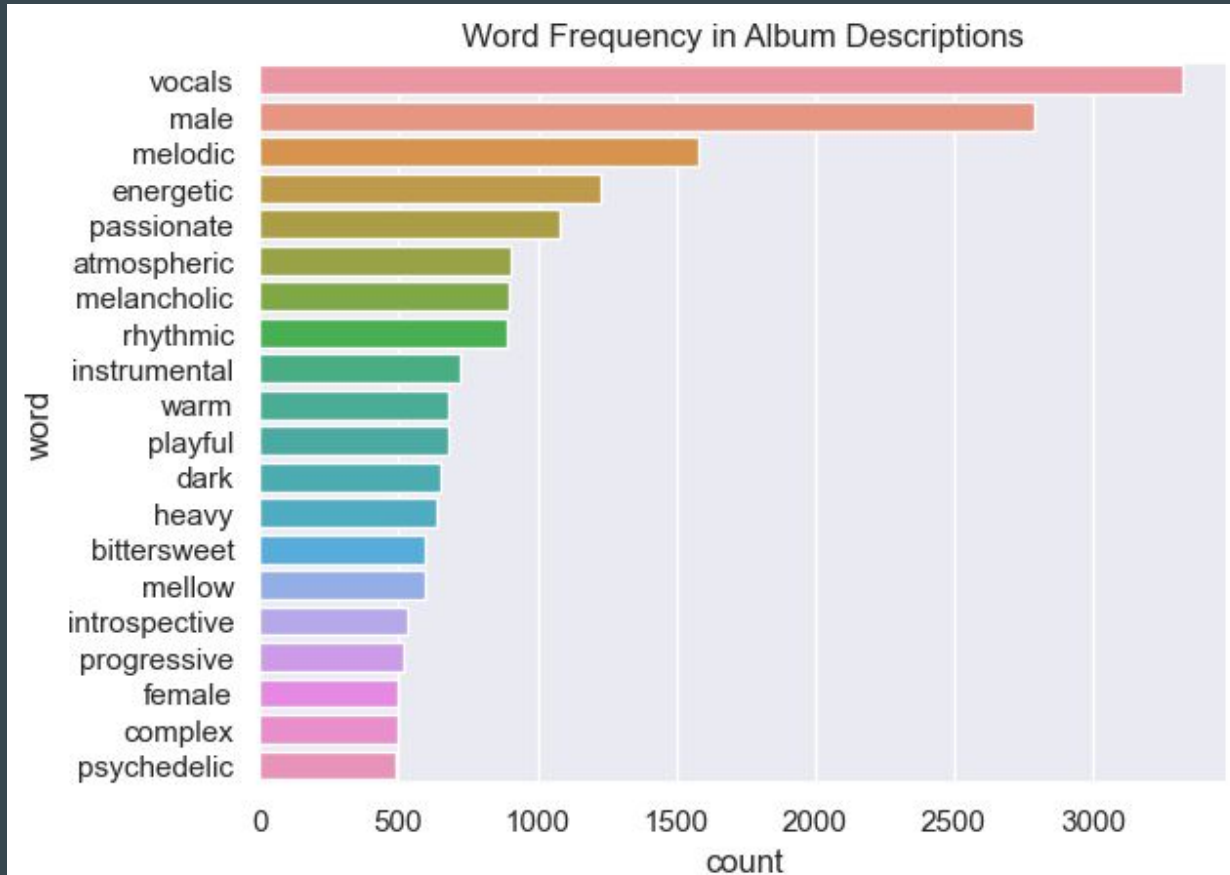
# Genres

### Album Genres Word Cloud





# Descriptions



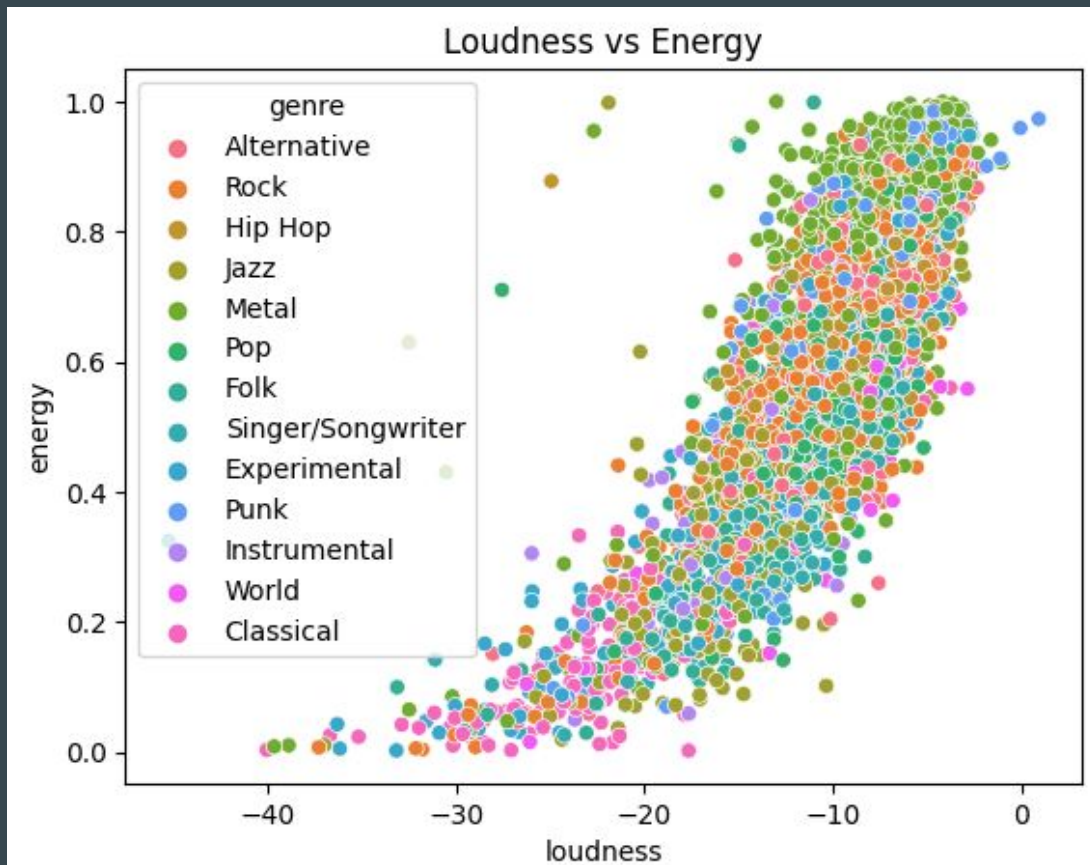
# Descriptions

### Album Descriptions Word Cloud

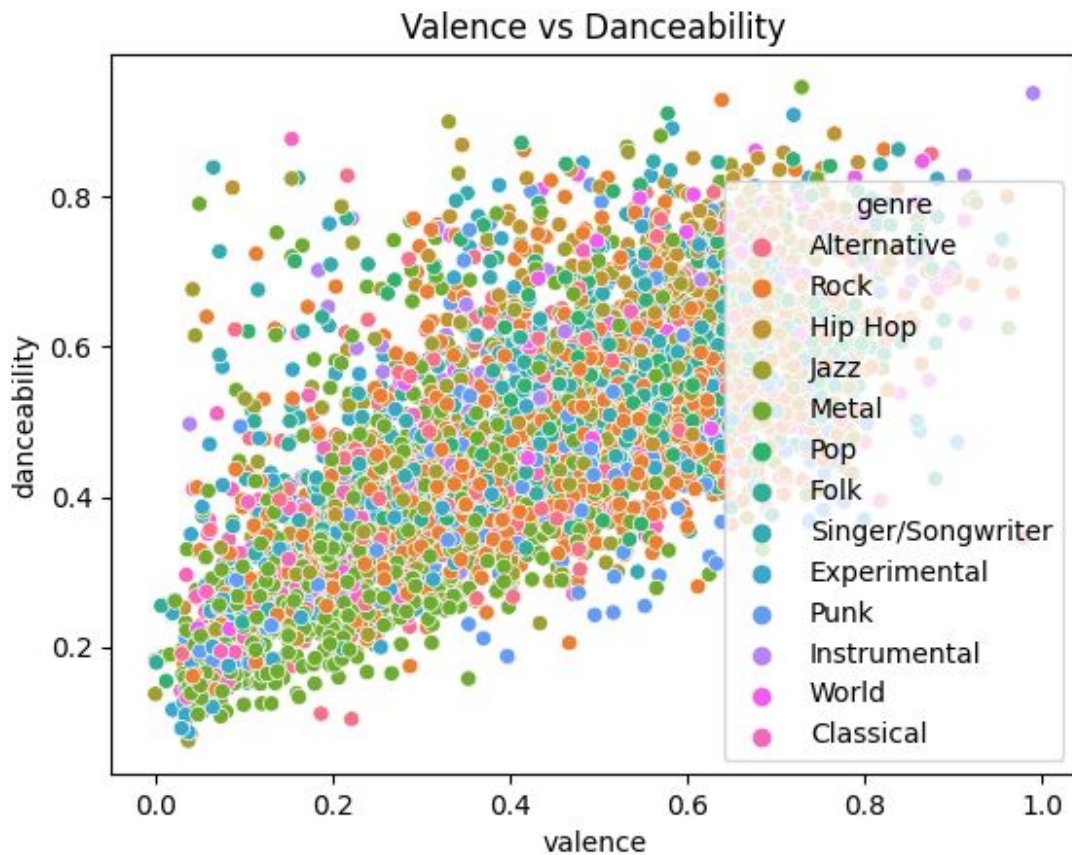


Heatmap showing the correlation matrix for 14 music features. The color scale ranges from -0.75 (dark purple) to 1.00 (light yellow). The diagonal is all 1.00. Notable correlations include a strong positive correlation between rel\_date and avg\_rat, and a strong negative correlation between acousticness and energy.

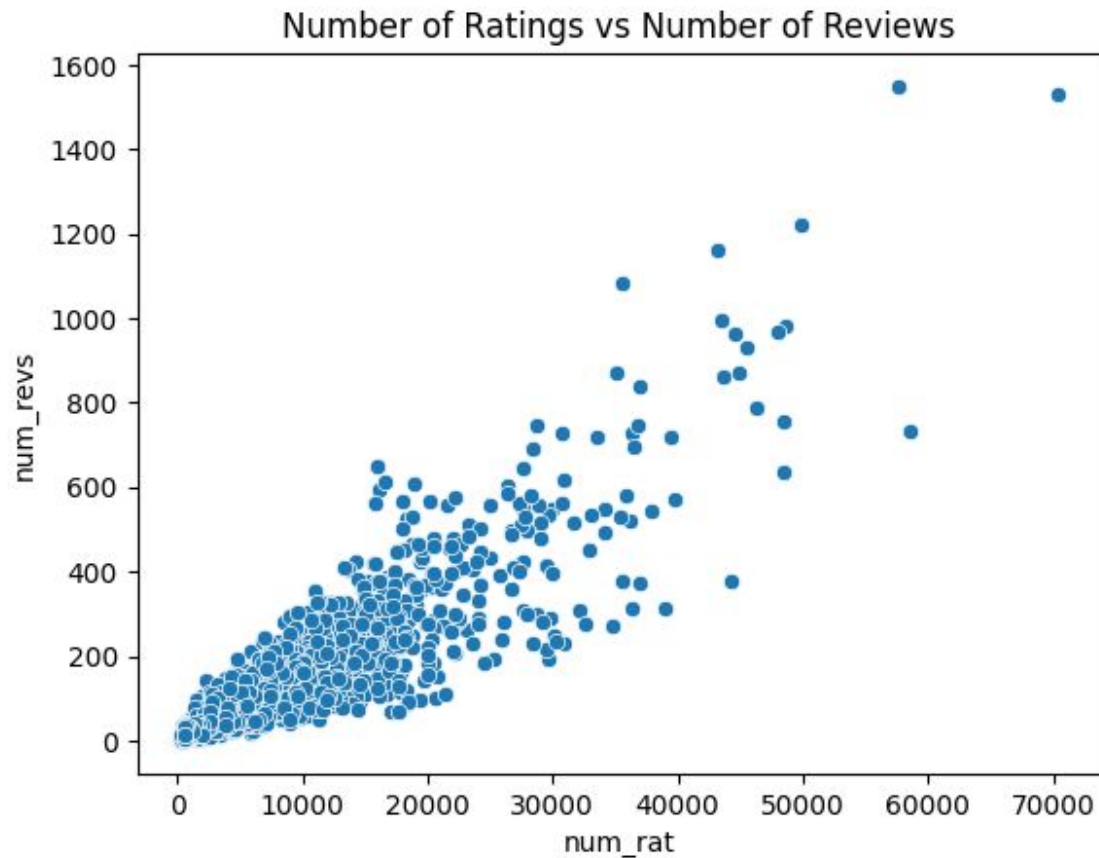
# Loudness and Energy



# Valence and Danceability



# Number of Ratings and Reviews



# The Models

## Model 1

Cosine Similarity with  
Linear Kernel and  
TfidfVectorizer

## Model 2

Non-negative Matrix  
Factorization (NMF) with  
TfidfVectorizer

## Model 3

Cosine Similarity with  
Word2Vec

# Metrics

## Artist

- Albums from same artist
- Albums from similar artists

## Genre

- Genre variety
- Out of place entries



# Bob Dylan - Blonde on Blonde

## Model 1

name	genres
Bob Dylan - Blonde on Blonde	Folk Rock, Singer/Songwriter
Bob Dylan - Highway 61 Revisited	Folk Rock, Singer/Songwriter
Nacho Vegas - Desaparezca aquí	Folk Rock, Singer/Songwriter
Bob Dylan - Bringing It All Back Home	Folk Rock, Singer/Songwriter, Contemporary Folk
Bob Dylan - Desire	Folk Rock, Singer/Songwriter
Van Morrison - Into the Music	Singer/Songwriter, Folk Rock
Nacho Vegas - Cajas de música difíciles de parar	Folk Rock, Singer/Songwriter
Richard Thompson - Rumor and Sigh	Folk Rock, Singer/Songwriter
Tim Buckley - Happy Sad	Singer/Songwriter, Psychedelic Folk
Smog - Knock Knock	Singer/Songwriter
Fiona Apple - When the Pawn	Singer/Songwriter, Piano Rock

## Model 2

name	genres
Bob Dylan - Blonde on Blonde	Folk Rock, Singer/Songwriter
Bob Dylan - Highway 61 Revisited	Folk Rock, Singer/Songwriter
Smog - Knock Knock	Singer/Songwriter
Bob Dylan - Bringing It All Back Home	Folk Rock, Singer/Songwriter, Contemporary Folk
Radiohead - Kid A	Art Rock, Experimental Rock, Electronic
R.A.P. Ferreira - Purple Moonlight Pages	Jazz Rap, Abstract Hip Hop
Bob Dylan - John Wesley Harding	Folk Rock, Singer/Songwriter, Contemporary Folk
Robert Wyatt - Rock Bottom	Canterbury Scene, Art Rock
Гражданская оборона [Grazhdanskaya Oborona] - Лунный переворот	Psychedelic Rock, Post-Punk
Pavement - Brighten the Corners	Indie Rock
Neutral Milk Hotel - In the Aeroplane Over the Sea	Indie Folk, Indie Rock

## Model 3

name	genres
Magnetic Fields - Holiday	Indie Pop, Synthpop
Raul Seixas - Krig-ha, bandolo!	Singer/Songwriter, Folk Rock, MPB
Raul Seixas - Gita	Singer/Songwriter, Rock, Folk Rock
Randy Newman - Bad Love	Singer/Songwriter, Pop Rock, Piano Rock
Millie Jackson - Caught Up	Southern Soul, Deep Soul
Guided by Voices - Under the Bushes Under the Stars	Indie Rock, Lo-Fi / Slacker Rock
Billy Bragg - Talking With the Taxman About Poetry	Folk Rock, Singer/Songwriter
Tom Waits - Small Change	Singer/Songwriter, Piano Blues
Donald Fagen - The Nightfly	Sophisti-Pop, Jazz Pop, Pop Rock
Dick Gaughan - Handful of Earth	Scottish Folk Music

# Bob Marley and The Wailers - Catch a Fire

## Model 1

name	genres
Bob Marley and The Wailers - Catch a Fire	Roots Reggae
Bob Marley & The Wailers - Exodus	Roots Reggae
Bob Marley & The Wailers - Kaya	Roots Reggae
Bob Marley and The Wailers - Soul Revolution	Roots Reggae
Culture - Two Sevens Clash	Roots Reggae
Black Uhuru - Showcase	Roots Reggae
Toots & The Maytals - In the Dark	Roots Reggae
Bob Marley & The Wailers - Survival	Roots Reggae
Dennis Brown - Visions of Dennis Brown	Roots Reggae
The Congos - Heart of The Congos	Roots Reggae
Bob Marley & The Wailers - Rastaman Vibration	Roots Reggae

## Model 2

name	genres
Bob Marley and The Wailers - Catch a Fire	Roots Reggae
Bob Marley & The Wailers - Exodus	Roots Reggae
Toots & The Maytals - Funky Kingston	Reggae, Rocksteady
Ini Kamoze - Ini Kamoze	Dancehall, Reggae
Fela Ransome-Kuti and The Africa 70 - He Miss Road	Afrobeat
Jorge Ben - África Brasil	Samba-rock, MPB, Funk
Buena Vista Social Club - Buena Vista Social Club	Son cubano, Bolero
Gilberto Gil - Gilberto Gil	Tropicália
Bob Marley & The Wailers - Rastaman Vibration	Roots Reggae
Santana - Abraxas	Latin Rock
João Donato - Quem é quem	Bossa nova, MPB

## Model 3

name	genres
Bob Marley & The Wailers - Exodus	Roots Reggae
Wilson Pickett - The Exciting Wilson Pickett	Southern Soul, Deep Soul
O'Jays - Back Stabbers	Philly Soul
Lee Dorsey - Yes We Can	New Orleans R&B, Funk, Soul
Curtis Mayfield - Curtis	Chicago Soul
Stevie Wonder - Songs in the Key of Life	Soul
Nina Simone - I Put a Spell on You	Vocal Jazz, Soul
Gil Scott-Heron/Brian Jackson - Winter in America	Soul, Jazz-Funk, Soul Jazz
Gil Scott-Heron & Brian Jackson - Bridges	Soul, Jazz-Funk, Funk
Michael Jackson - Bad	Dance-Pop, Contemporary R&B, Pop

# Results

## Model 1

Demonstrated the best results with consistent genre recommendations and repeated suggestions of albums by the same artist.

## Model 2

Offered more genre variety while still predicting albums by the same artists.

## Model 3

Performed poorly, yielding inconsistent genres and sometimes failing to predict albums by the same artist.

# Conclusions

Overall, the TfidfVectorizer model with Cosine Similarity (Model 1) proved to be the most reliable and accurate for the album recommendation system.

---

# Further Research

- Incorporate user feedback and ratings into the recommendation system to personalize suggestions based on individual preferences.
- Explore the use of collaborative filtering techniques to enhance the accuracy and relevance of album recommendations.

---

# Capstone Project: Album Recommendation System



Luis Fernando Silva