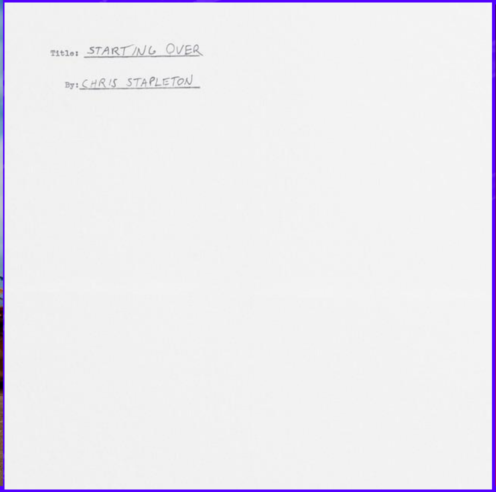
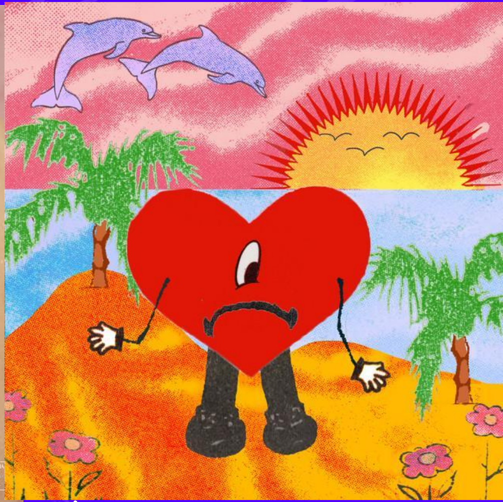

PARENTAL
ADVISORY
EXPLICIT CONTENT


Title: STARTING OVER
By: CHRIS STAPLETON


Something For Thee Hotties
FROM THEE ARCHIVES



# Face the Music: Can an Album's Cover Art Predict its Popularity?

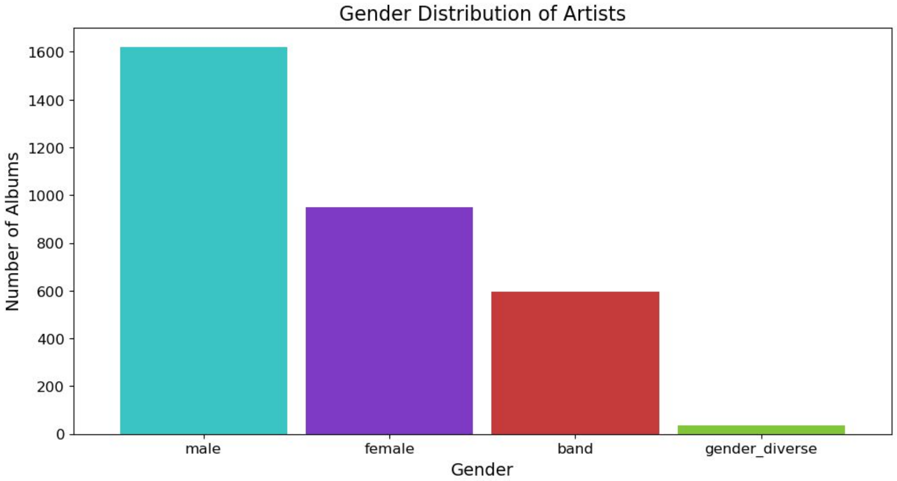
JESSICA JOY, RACHNA RAWALPALLY, SAMYU VAKKALANKA, MARIE VAUGHAN

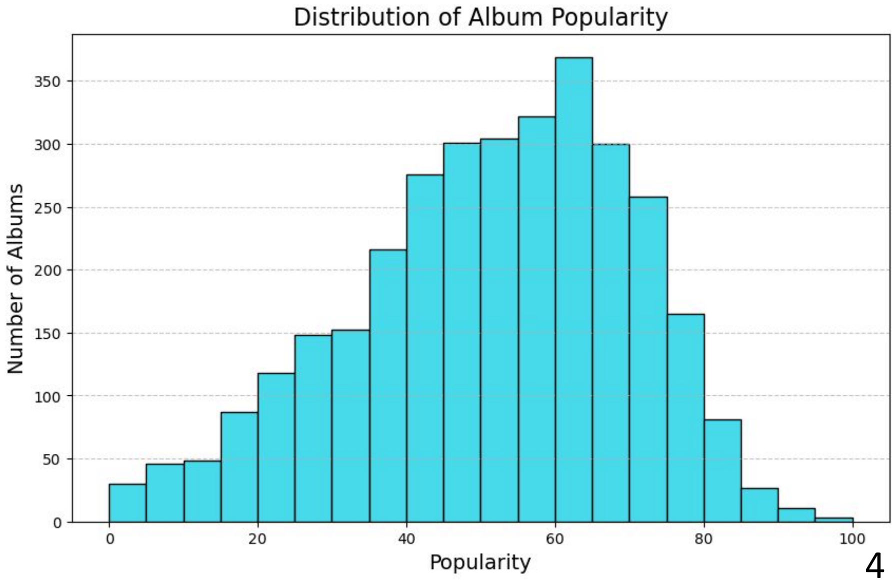
# Introduction

- Album covers are the first visual representation of an artist's work, shaping public perception and genre identity.
- Do female artists display more sexualized images or show more skin on their albums covers compared to covers from male artists? Do sexualized covers attract more popularity?
- How does the composition of an album cover correlate with popularity?
- We will analyze image data using a CNN and then extract features using pretrained models to make an ANN to compare the visual content of album covers.

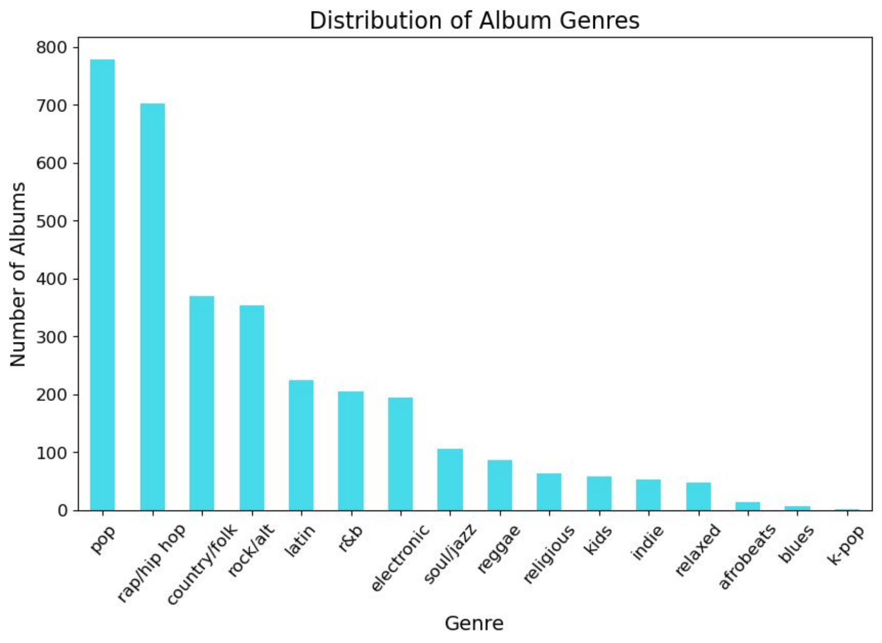
# The Data

- Spotify API:
  - Features pulled: artist name, artist id, artist genre, album id, album name, release date, total tracks, popularity, image url, artist popularity, followers
- Genre API (last.FM and Wikipedia API):
  - Features pulled: album genre
- Create/standardize genre categories
- Pretrained Models
  - YOLO v8 - You Only Look Once
  - NudeNet - Nudity detection
- Python libraries
  - Color detection
  - Skin percentage
  - Canny Edge Detection
- Dataset:
  - 3290 different albums
  - 16 different genres
  - 100+ features

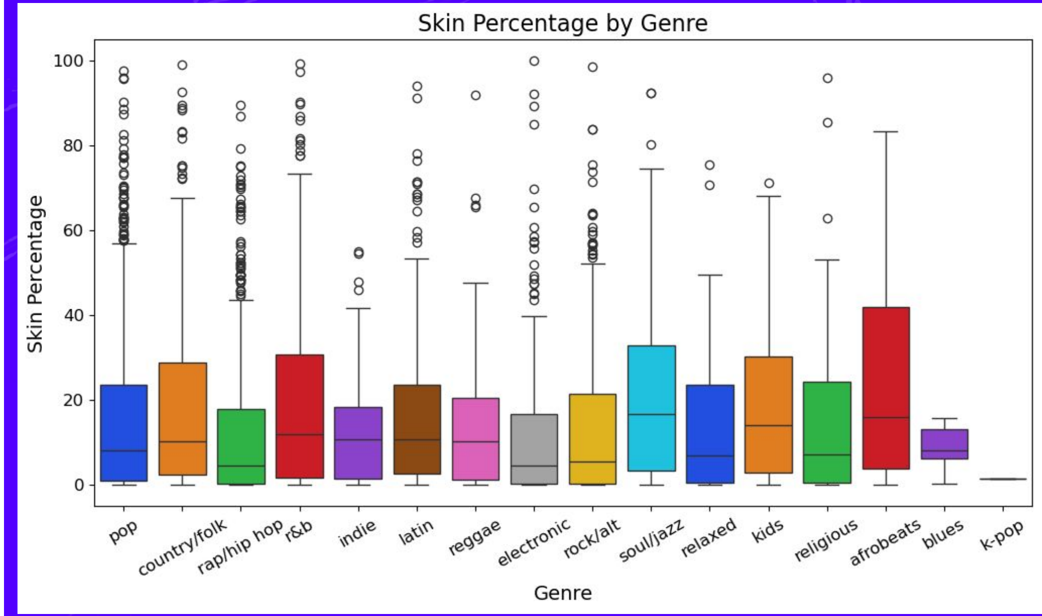


# Genre

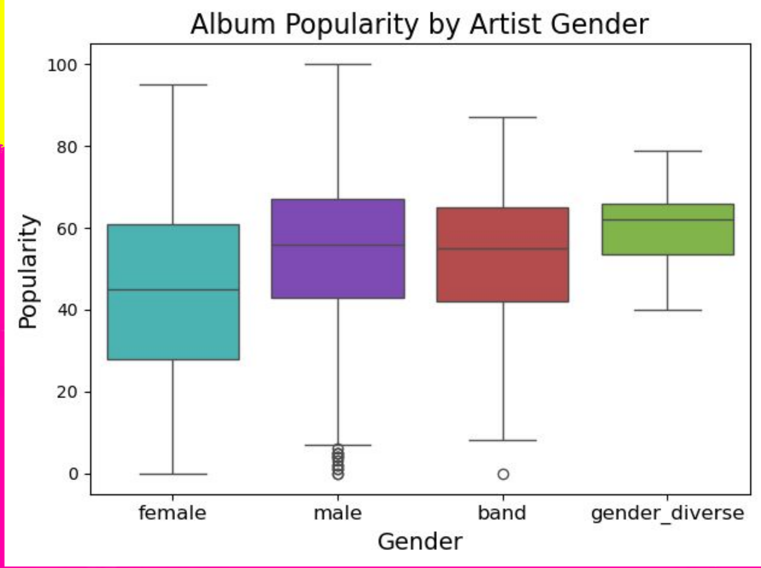


Distribution of Album Genres
Number of Albums
800
700
600
500
400
300
200
100
0
pop
rap/hip hop
country/folk
rock/alt
latin
r&b
electronic
soul/jazz
reggae
religious
kids
indie
relaxed
afrobeats
blues
k-pop
Genre

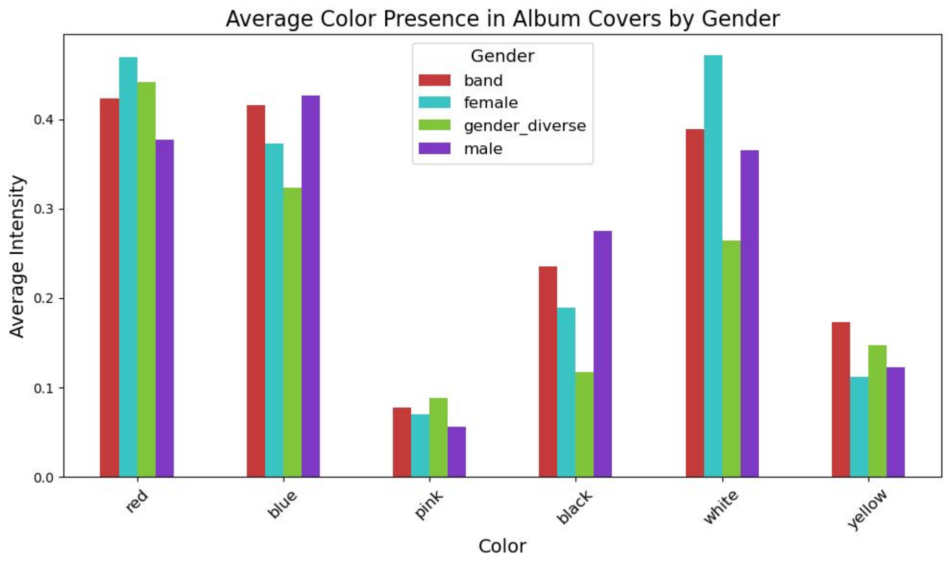



Skin Percentage by Genre
Skin Percentage
100
80
60
40
20
0
pop
country/folk
rap/hip hop
r&b
indie
latin
reggae
electronic
rock/alt
soul/jazz
relaxed
kids
religious
afrobeats
blues
k-pop
Genre

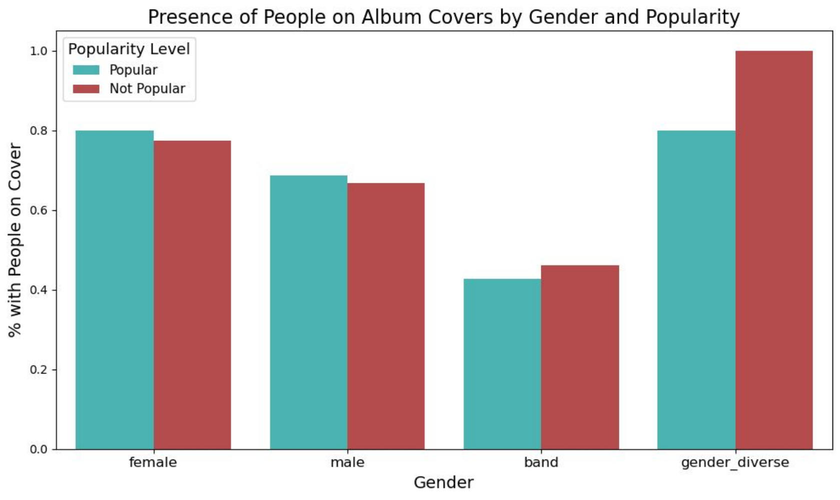
# Gender



Album Popularity by Artist Gender
Popularity
100
80
60
40
20
0
female
male
band
gender_diverse
Gender

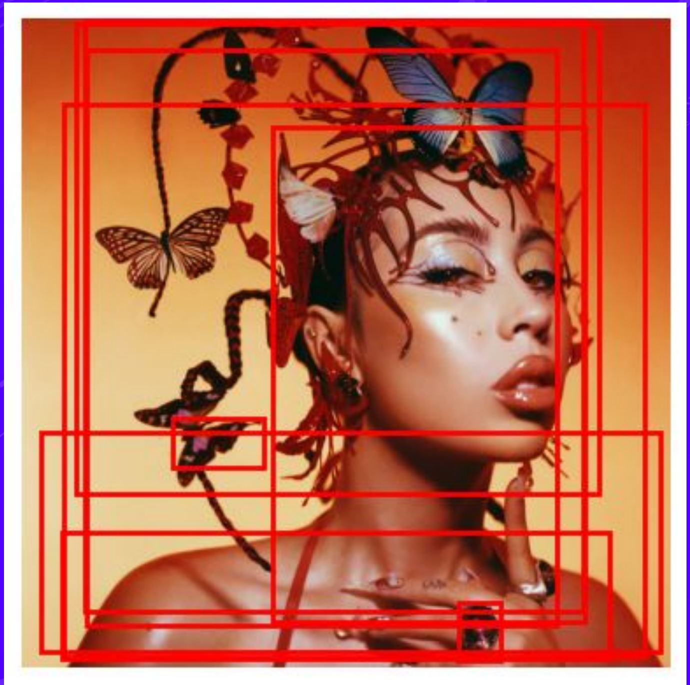
Average Color Presence in Album Covers by Gender
Gender
band
female
gender_diverse
male
Average Intensity
0.4
0.3
0.2
0.1
0.0
red
blue
pink
black
white
yellow
Color

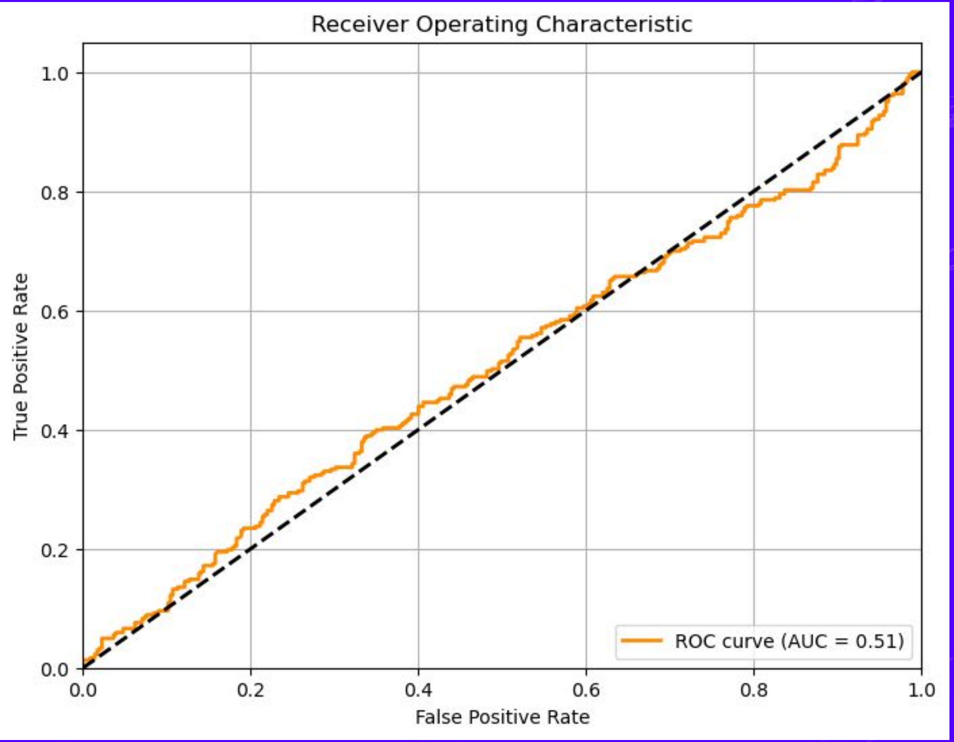
Presence of People on Album Covers by Gender and Popularity
Popularity Level
Popular
Not Popular
% with People on Cover
1.0
0.8
0.6
0.4
0.2
0.0
female
male
band
gender_diverse
Gender

# The Models

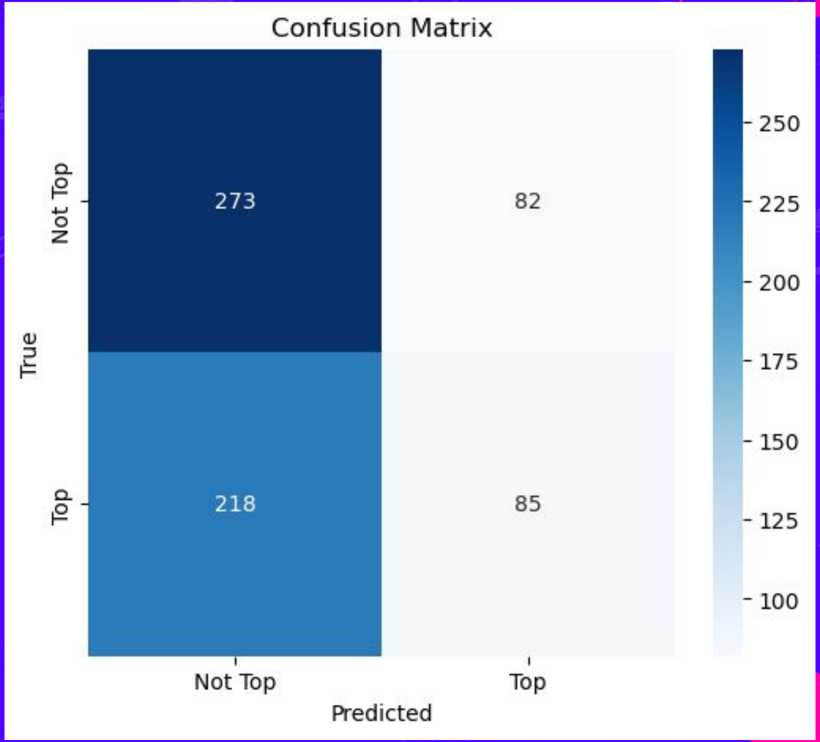
- CNN Model
  - Classifying popularity from album cover images
  - Architecture: Augmentation, 3 convolutional blocks, Flatten, 2 Fully Connected, Dropout
- ANN Models
  - Main model: classifying popularity from album cover features
  - Gender models: important features for male vs female artists
  - Genre models: important features in popular genres



# CNN Model



Receiver Operating Characteristic
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC curve (AUC = 0.51)

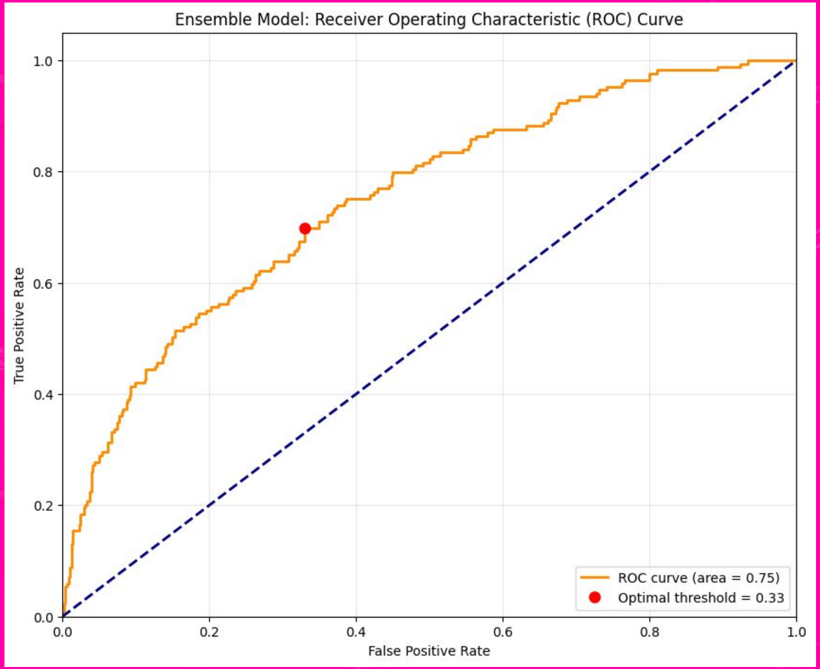



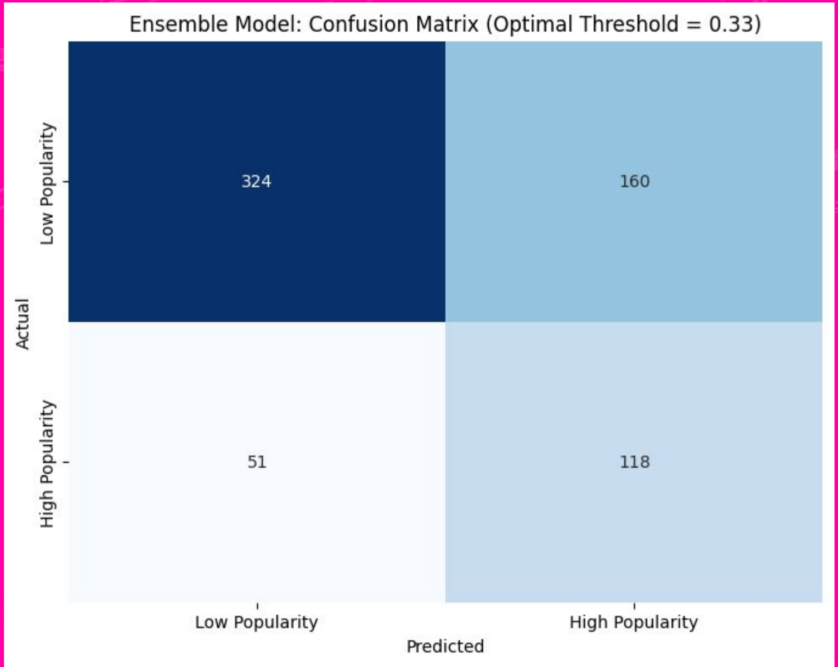
Confusion Matrix
True
Predicted
Not Top
Top
273
82
218
85
250
225
200
175
150
125
100

# ANN Model Architecture

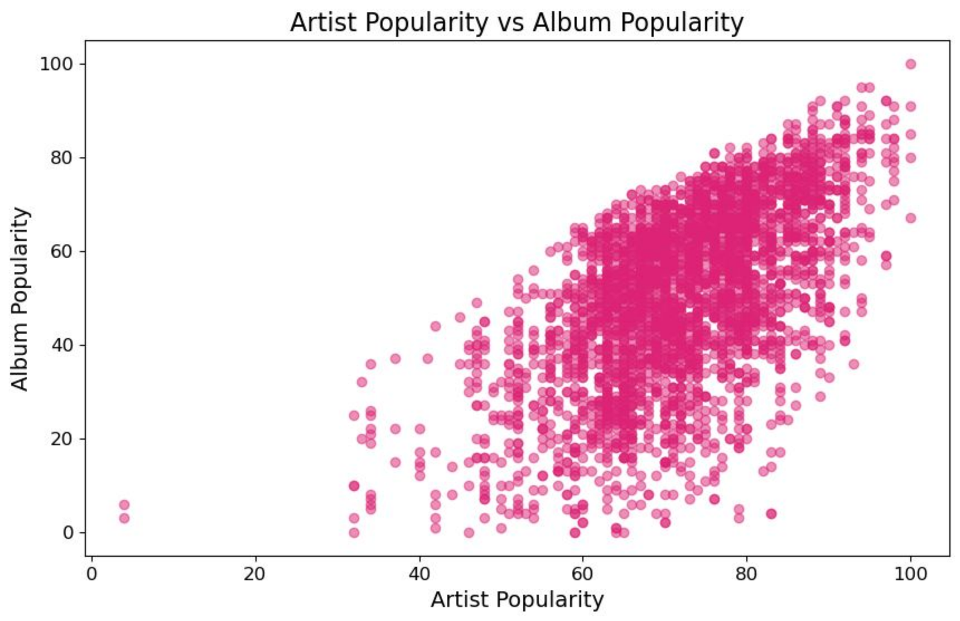
- **Feature Engineering: Standardization**, **SMOTE**, **Feature Selection** of 100+ features using a Random Forest (features with importance above the median retained)
- **ANN Architecture**
  - **Input Layer**: Selected album features
  - **Regularization Layer**: GaussianNoise
  - **First Hidden Layer**: 112 neurons
    - ReLU activation, L1 regularization, Batch normalization, Dropout
  - **Second Hidden Layer**: 64 neurons - same regularization structure
  - **Output Layer**: sigmoid activation
  - **Loss Function**: Focal loss (good for class imbalance)
  - **Optimizer**: Adam
- **Ensemble Approach - combined ANN predictions with tuned XGBoost model:** Simple average of predictions from both models.

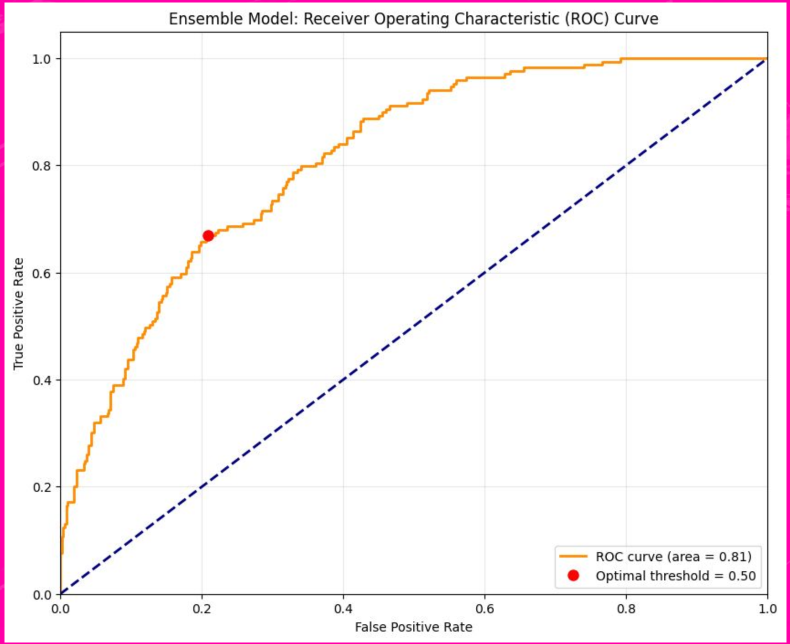
# Ensemble Model v1



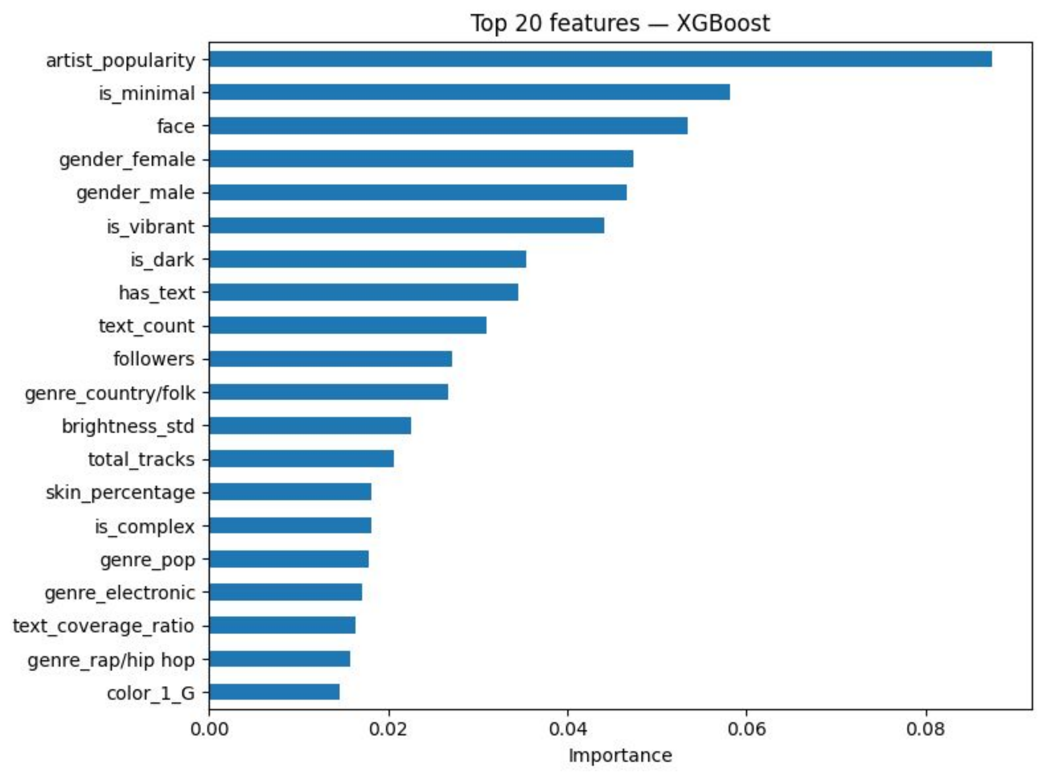



# Ensemble Model v2



# Ensemble Feature Importance



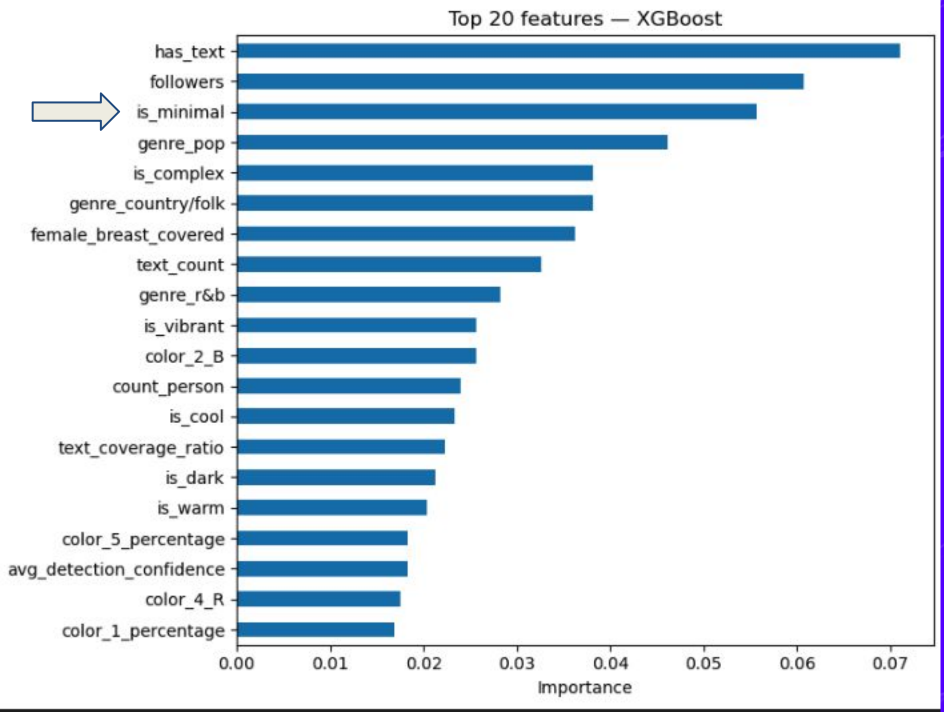
# Ensemble Feature Importance

Face

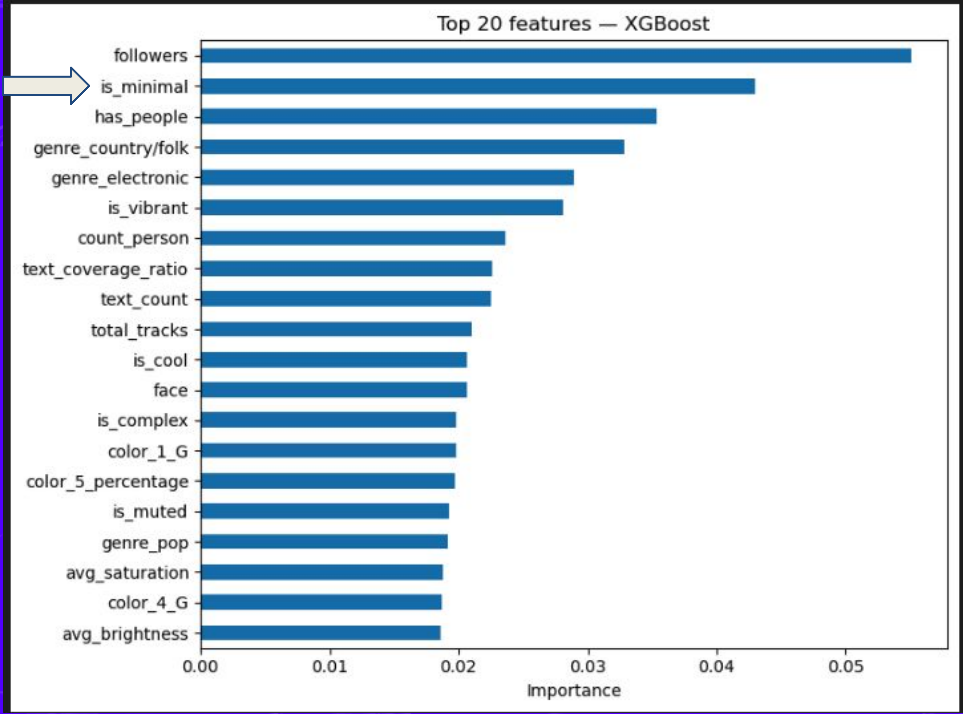
No Face

# Model by Gender

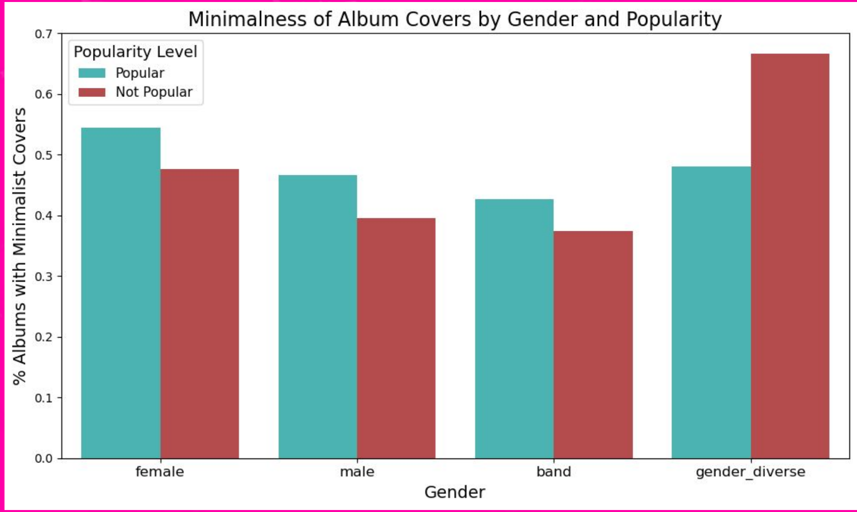
Female Artists




Male Artists



# Minimalism

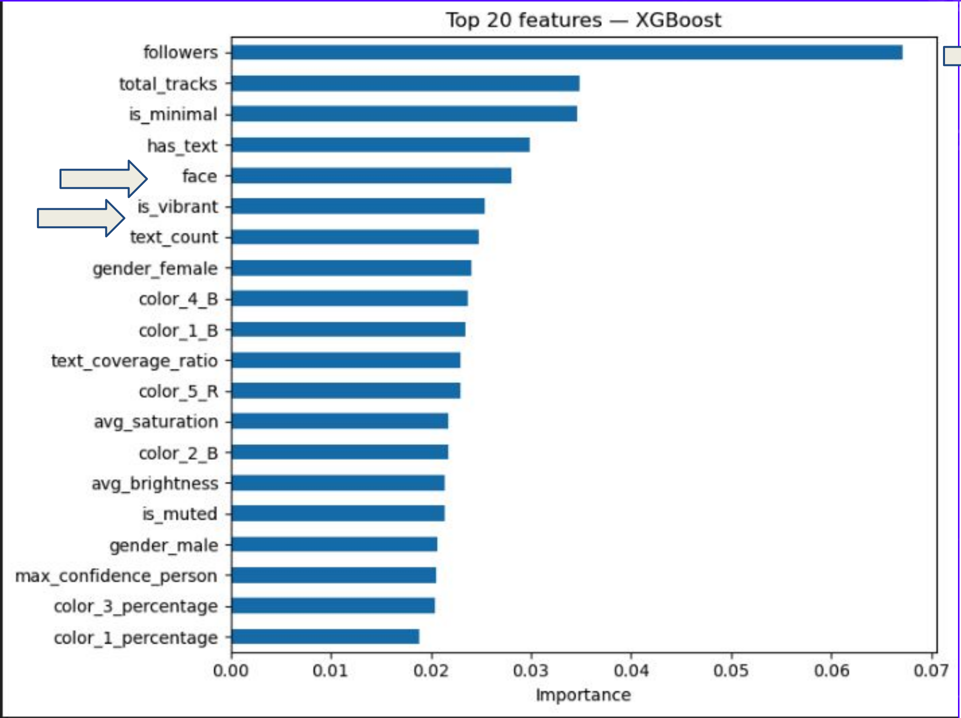



Minimal

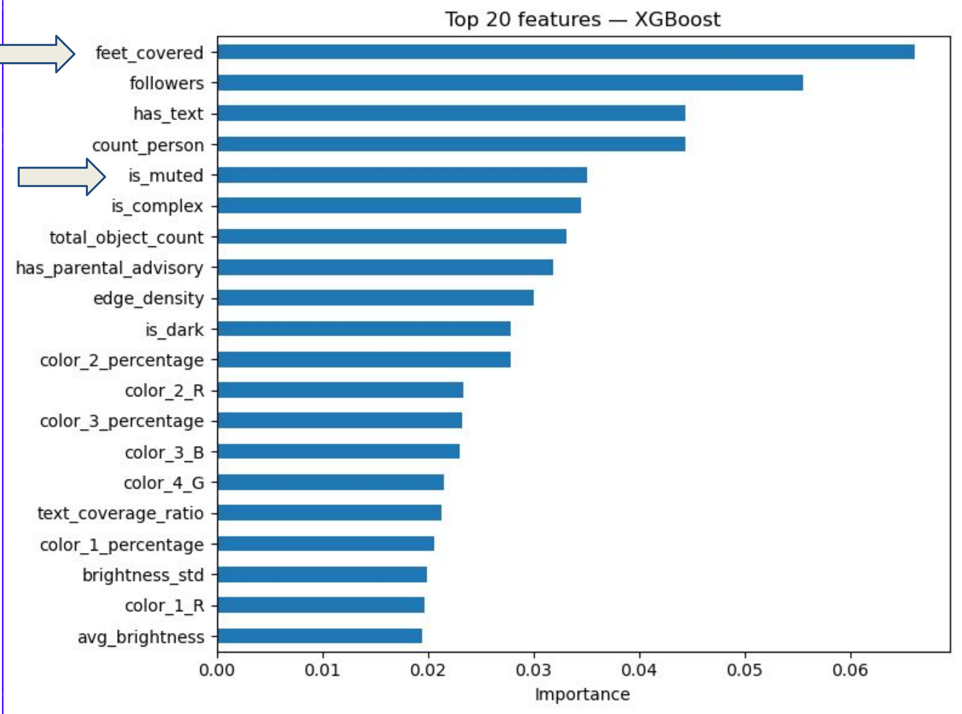

Complex

# Model by Genre

Pop




Rap/Hip Hop



# Model by Genre

Pop: face/text


Rap/Hip Hop: feet covered

# Model by Genre

Pop: vibrant


Rap/Hip Hop: muted

# Results

- Pre-trained models to analyze images works better than a CNN from scratch
- Across the board, albums with minimal features do better
- Artist popularity has one of the biggest impacts on album popularity
- Gender
    - Both: bodies in different ways (presence of people)
    - Female albums: covered female body parts present
    - Male albums: faces, people
- Genre
    - Pop: minimal, vibrant colors
    - Rap/hip-hop: muted colors, more complex designs
    - Both: has text

# Limitations

- Reduced Data Transparency: Spotify's API is limiting access to detailed track and album data like genre
  - Restricts transparency for analysts hurting artists, music producers, and researchers
  - Rate limits decreased the amount of data we could collect
- Streams vs record sales
  - Predicting based off Spotify's popularity metric - likely doesn't emphasize sales of physical records where album covers probably come into play more
  - Better metric of popularity would have been sales the week of release for physical albums
- When separating by different genres, amount of data used in model significantly decreased
  - Pop: 779
  - Rap/ Hip Hop: 702
  - Rock/Alt: 354
  - Country/ Folk: 307
- Many more male artists than women
  - Male: 1622
  - Female: 951
- ANN too complex for small amounts of data - XGBoost Model tended to perform better

# Conclusion

- Initially testing to see if women/ men show more skin on album covers and how that compares with popularity
- Artist's popularity is the most significant predictor of album success
- Gender differences do exist in Album Art as well as genre-specific designs
- Call to action
    - Artists can learn about the prominent features of popular albums in their genre to inform their own album cover design
    - Marketers should actively work to promote and give visibility to less popular or emerging artists
    - We need to identify and avoid gender bias in how albums are marketed and promoted
- Future Research
    - Understand how social media presence/ branding/ post interactions/ marketing can influence album popularity
    - Investigate how streaming has changed album art
        - Differences in impact of album on art on Spotify popularity vs. retail sales

# THANK YOU

# Citations


Anthropic. "Conversation with Claude AI Assistant." 22 Apr. 2025.

Duncan, L. (2020). *wikipedia-api* (version 0.5.4), from https://pypi.org/project/wikipedia-api/. Accessed 20 Apr 2025.

Last.fm. *Last.fm API*. Last.fm,  https://www.last.fm/api. Accessed 02 Apr 2025.

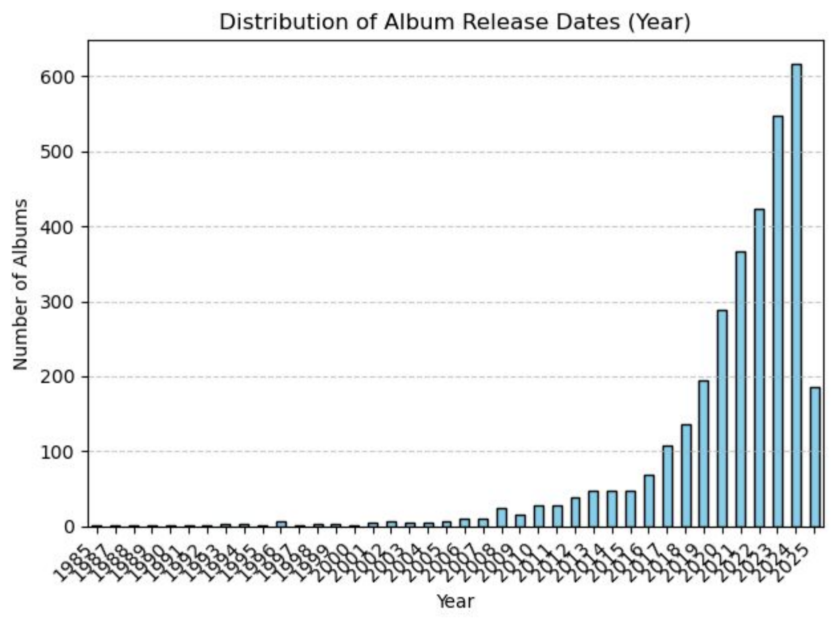
OpenAI. "ChatGPT (GPT-4)." *OpenAI*, https://openai.com/chatgpt. 22 Apr. 2025.

Spotify. Spotify Web API. Spotify for Developers, https://developer.spotify.com/documentation/web-api/. Accessed 04 Apr. 2025.

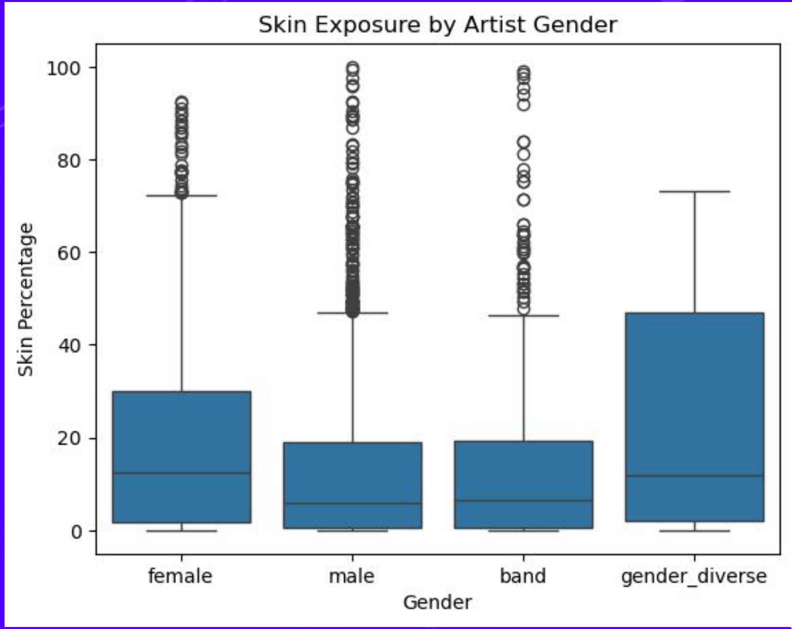
notAI-tech. *NudeNet*. GitHub, https://github.com/notAI-tech/NudeNet. Accessed 12 Apr. 2025.

Ultralytics. "YOLOv8." GitHub, 2023, github.com/ultralytics/ultralytics.

# Appendix



Distribution of Album Release Dates (Year)
Number of Albums
0
100
200
300
400
500
600
Year

Skin Exposure by Artist Gender
Skin Percentage
0
20
40
60
80
100
female
male
band
gender_diverse
Gender