Ø6/Ø6/2Ø25

STAT 155

Ipsita Bisht

Spotify Music Feature Exploration

What can we say about music genres and their distinct features?

Music Features Help Define Genres

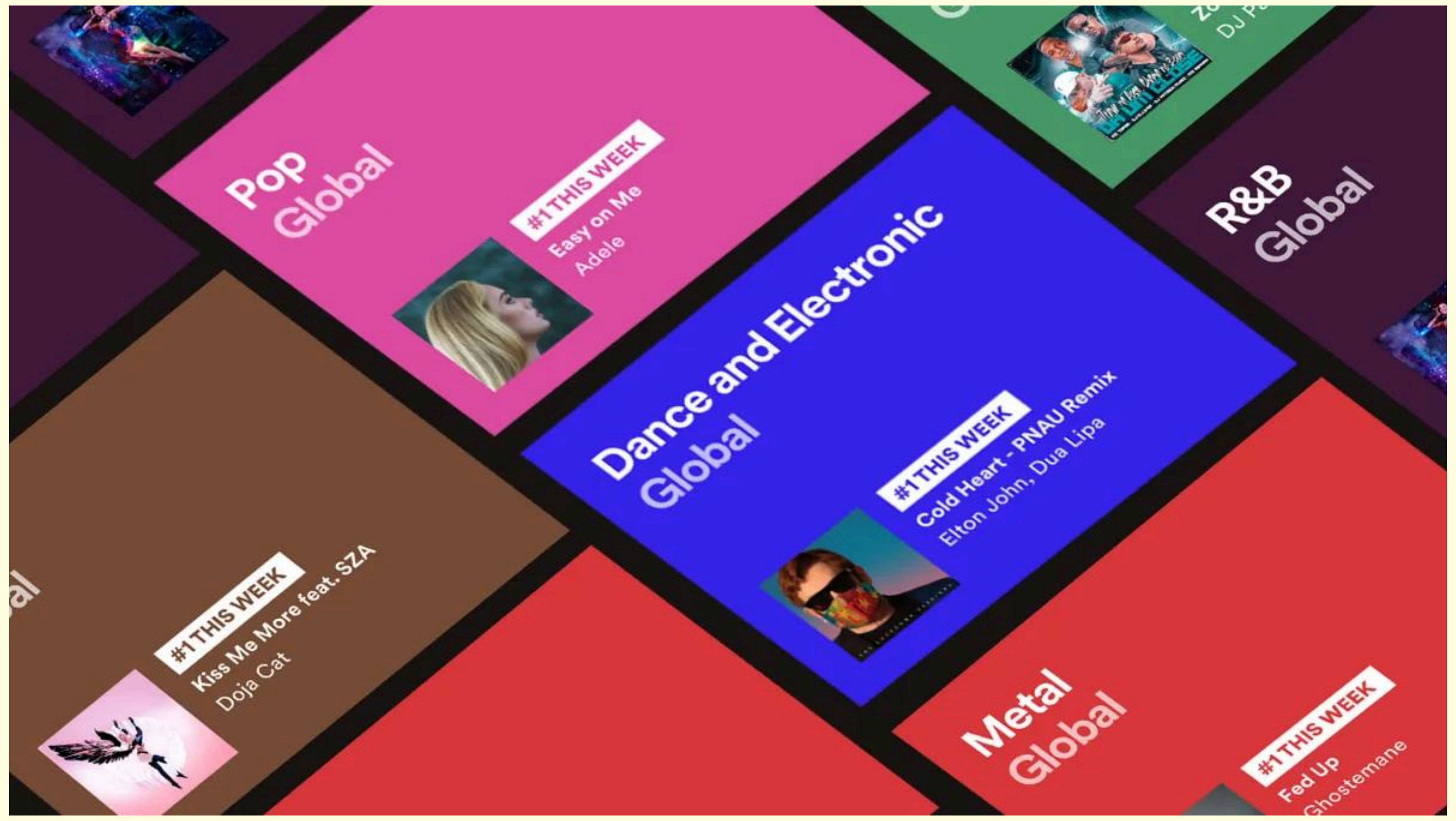
What audio features best predict a track's popularity within specific genres

How do these success factors differ across genres?

Could track lists that deviate from the typical features of their genre still be successful?

Context:

- understand core differences between genres and why certain songs blow up within a genre and others do not
- explore multi-feature differences through data



Source

Dataset + Kaggle

- Dataset containing 140,000 tracks, 125 genres
- Collected using Spotify API Http requests
- Important features: ‘popularity’, ‘track_genre’, ‘danceability’, ‘loudness’, etc.

Tools

Python Libraries + Docs

- Pandas to create the dataframes, NumPy for mathematical functions
- Sklearn for modeling
- Quarto files for running cells and documentation

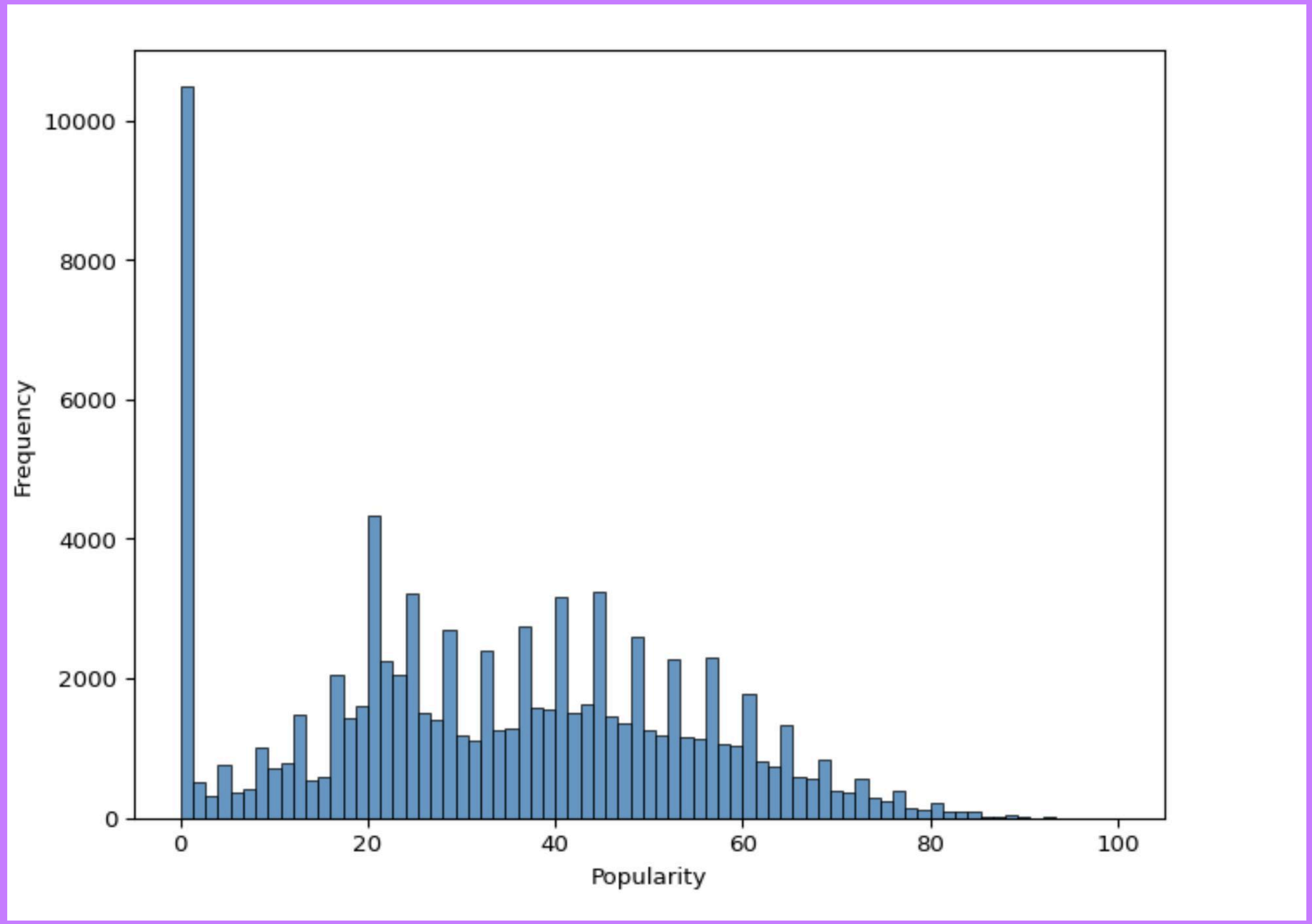
Methods

Preprocessing

- Challenges: skewed distribution of popular songs
  - defined a new column ‘is_hit’ with a threshold of popularity >= 75
- Convert boolean columns to binary value
- Removing duplicates

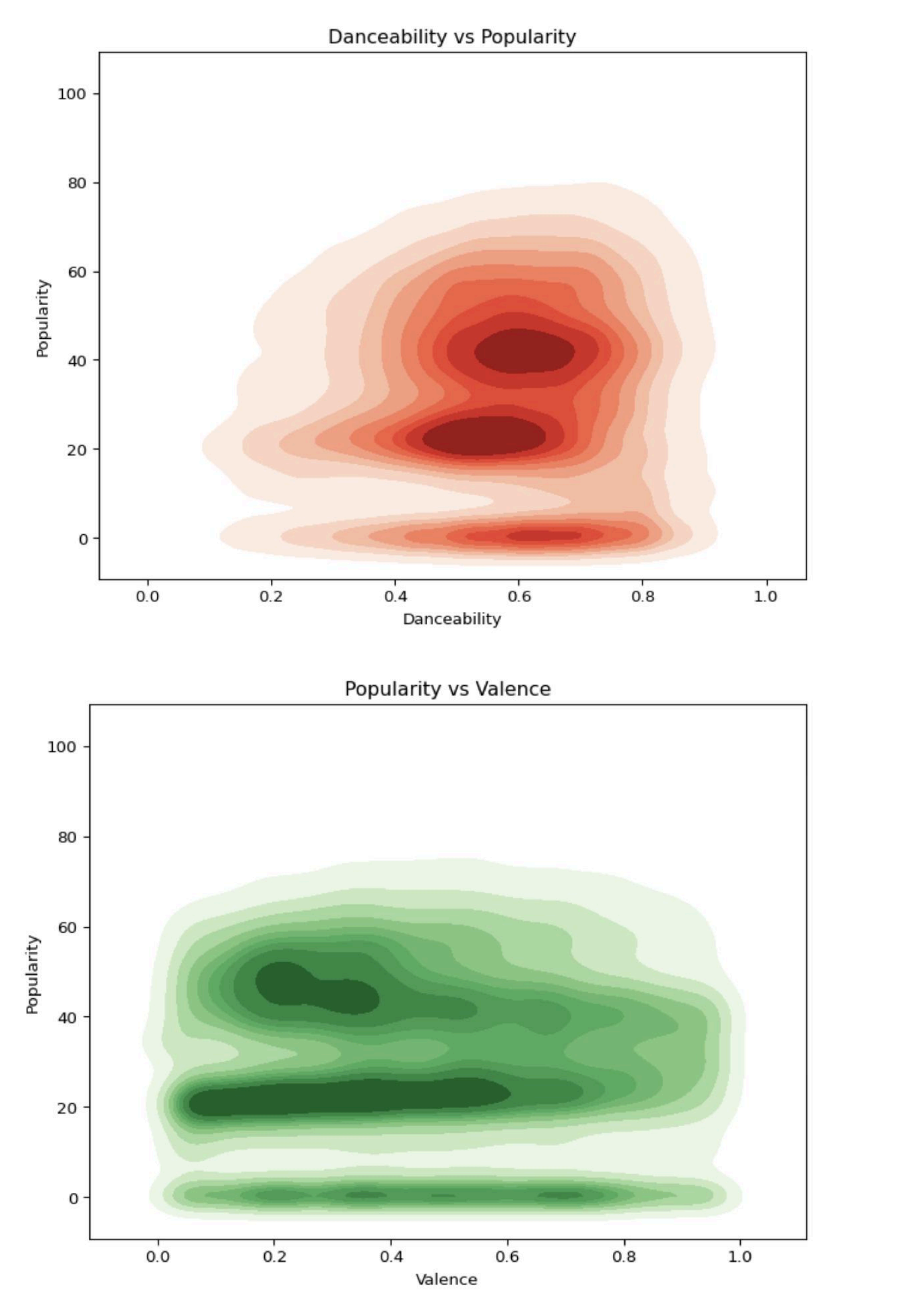
Univariate

- Popularity is rare in this data set
- To define a song as a hit, must use a specific threshold



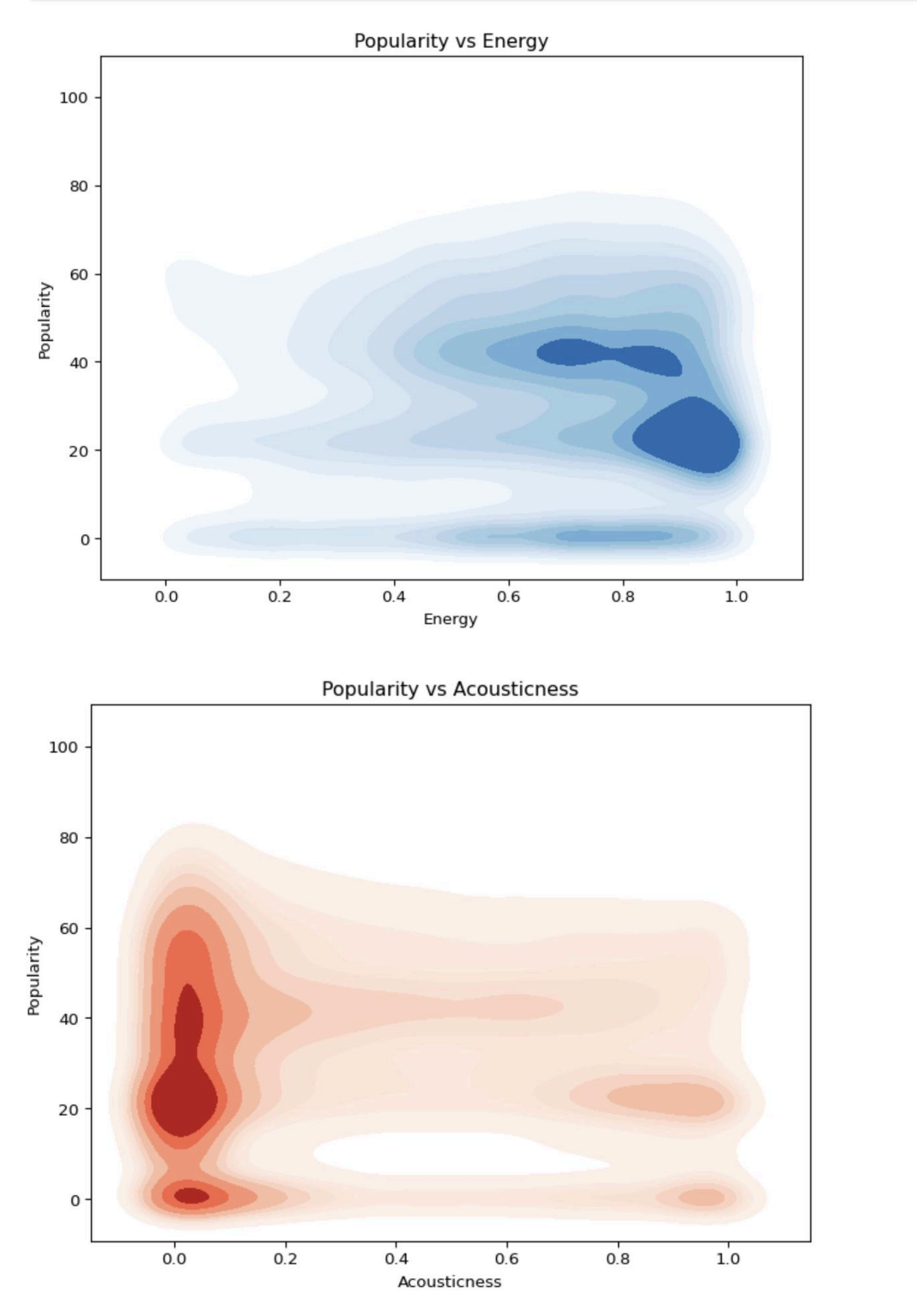
Bivariate

- increasing danceability and valence from 0.0-1.0 minimal impact on popularity
  - no strong relationship between the two variables



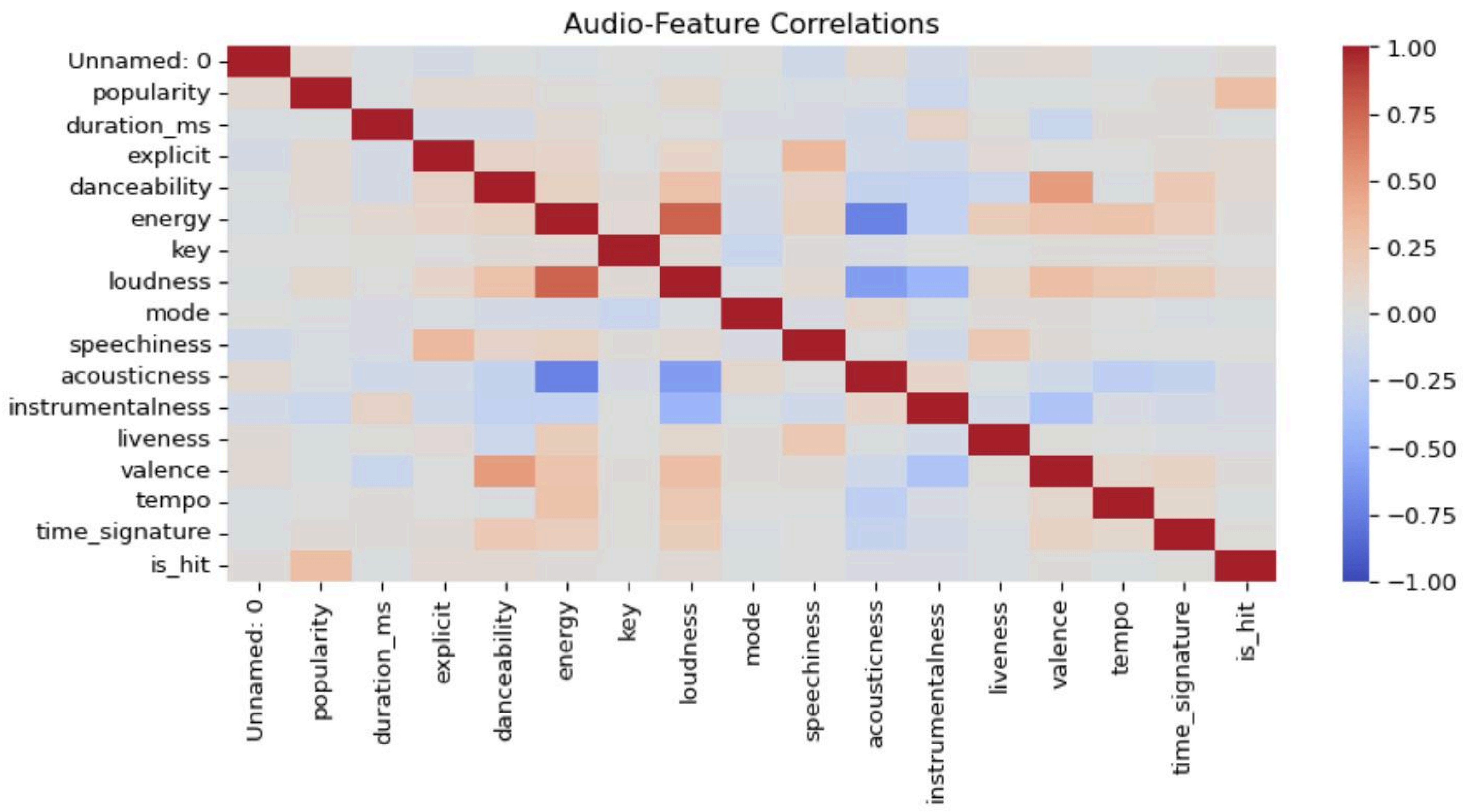
Bivariate

- Same for acousticness and energy, no strong relationship between the two variables
- Distribution leans more heavily on 0.6-1.0 for valence and 0.0-0.2 for acousticness



Multivariate

- energy + loudness have a slightly strong positive correlation
- energy + acousticness have slightly strong negative correlation



Running the Model

Artificial Neural Network

- **Goal**: classify if a song is a hit (1) or not (0) based on musical features within each genre
  - 'danceability', 'energy', 'valence', 'acousticness', 'speechiness', 'instrumentalness', 'tempo', 'loudness'
  - 'is_hit' and 'track_genre_{genre}'
- **Method**:
  - training testing split (90/10)
  - used one hot encoding for genres
- **Results**
  - faster and produces decent Precision and Recall scores

Dataset class imbalances: Binary classification → hard to detect songs that have 'is_hit' = 1

- use Over Sampling method → increases volume of minority class ('is_hit' = 1)

Model Prediction Performance

Accuracy is *misleading*, must look at Precision, Recall, F1-Score

Classification Report:

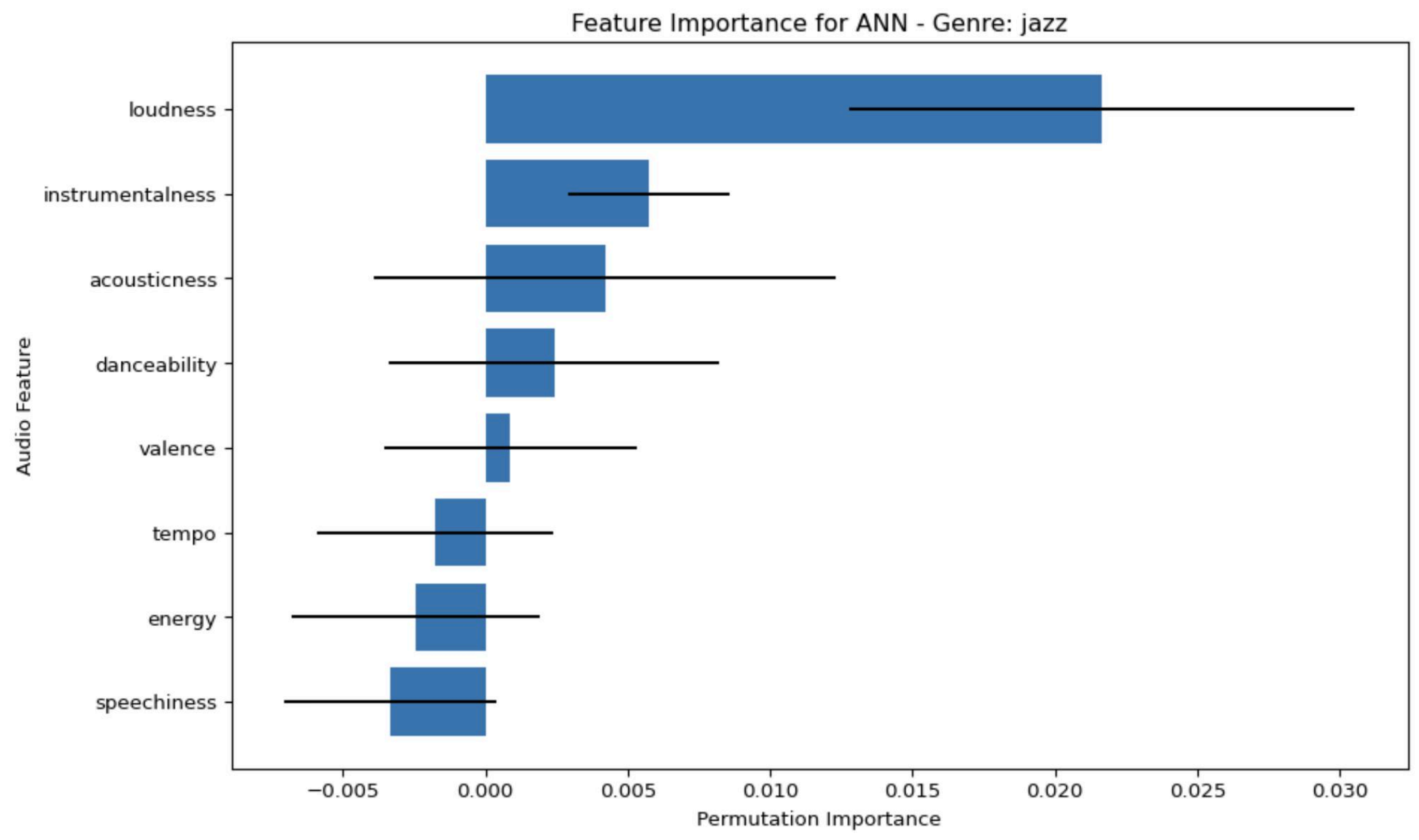
	precision	recall	f1-score	support
0.0	0.95	0.93	0.94	2036
1.0	0.23	0.27	0.25	147
accuracy			0.89	2183
macro avg	0.59	0.60	0.59	2183
weighted avg	0.90	0.89	0.89	2183

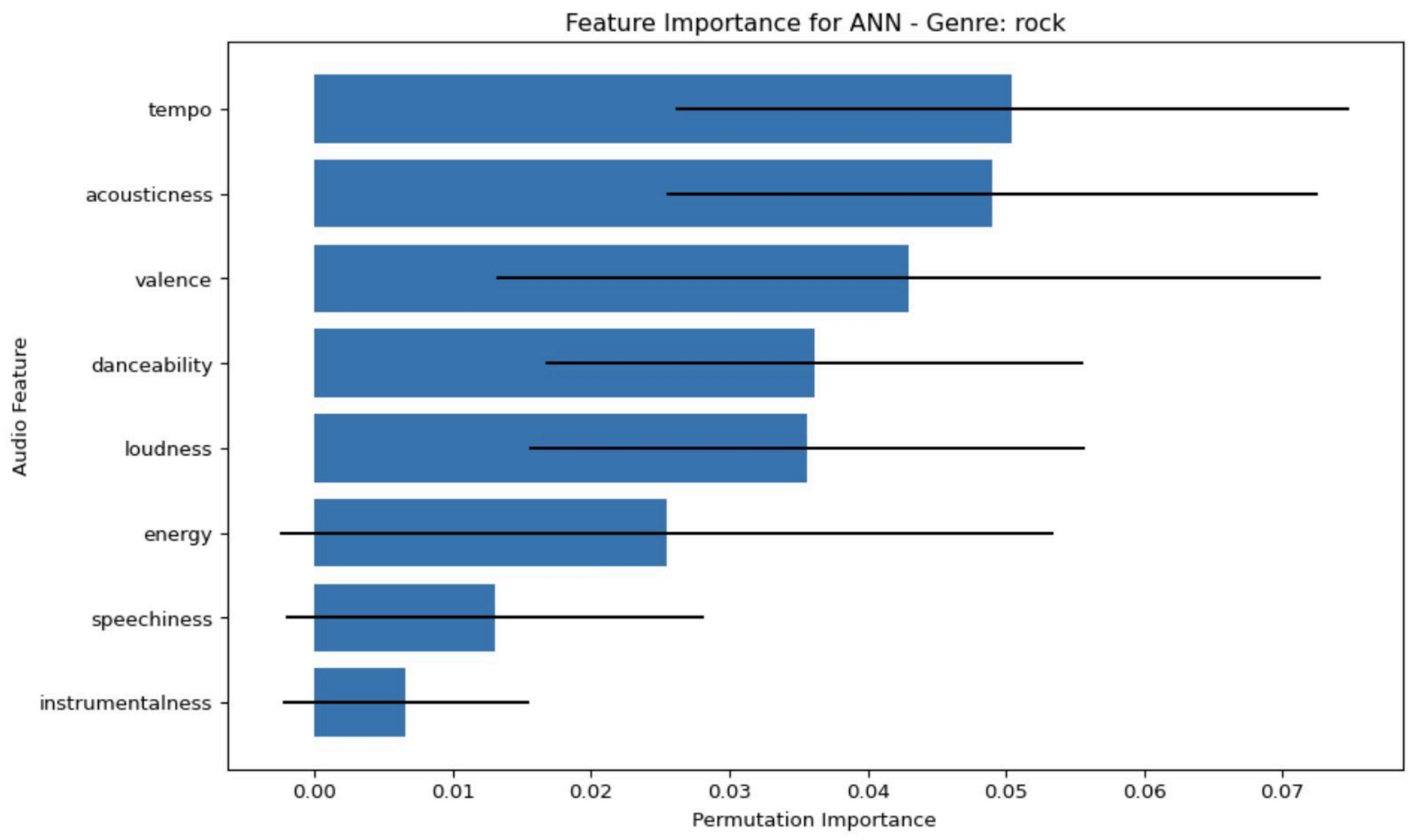
Feature Fingerprint

Permutation Importance
Retrain model after shuffling features to see which features impact model performance

- **Goal**: See which features impact popularity of song within genre
- **Method**:
  - use Sklearn's Permutation Importance
  - visualize feature importance for genres individually
- **Results**
  - genres with varying feature lists and scale of importance
  - features that model learns as most important in accurately predicting popularity

Feature Importance within Genre



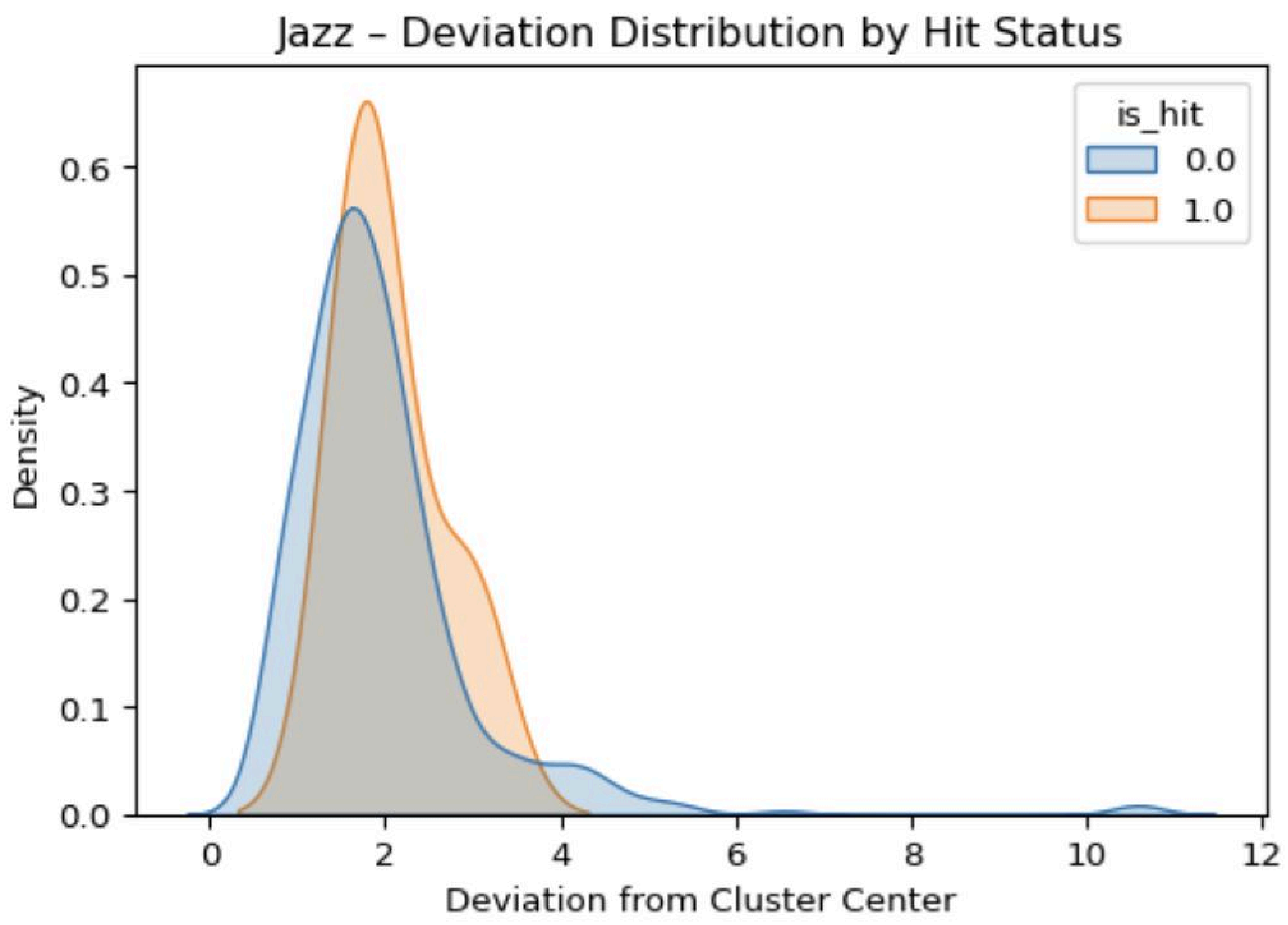


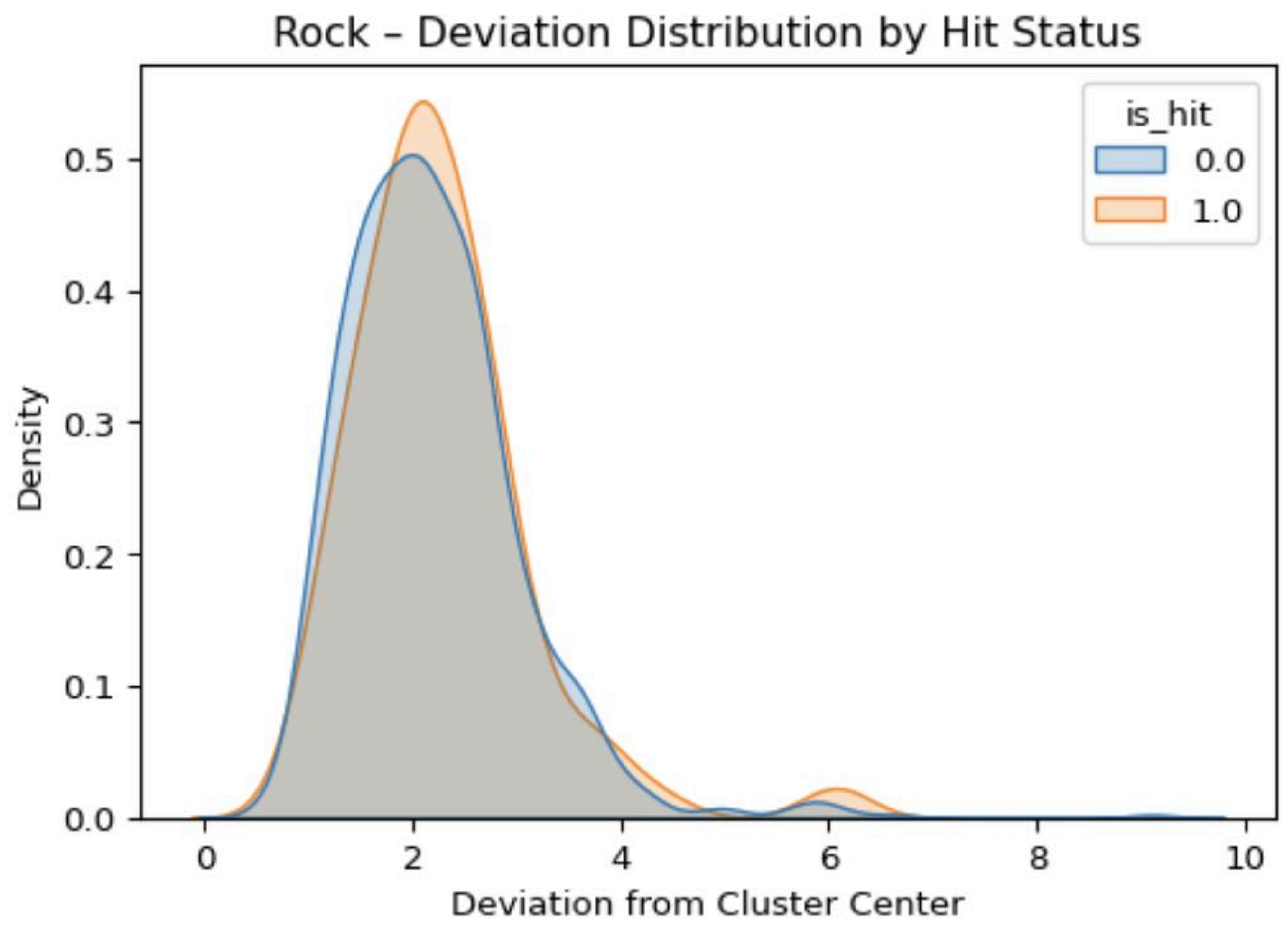
Deviations from Norm

Clustering
Create natural clusterings of songs within a genre

- **Goal**: See how deviation from genre clusters impacts popularity of the song
- **Method**:
  - use Sklearn's KMeans Clustering model
- **Results**
  - visualization of distribution of songs within certain deviations of the norm and its popularity score
  - most genres had similar distributions and distances from the mean for hit and non-hit songs

Deviation Distribution based on hit status





Monte Carlo Experiment

Class Imbalances impacted the f1 score evaluations of the model

- **Purpose**: Understand how adjusting the minority class proportion impact precision, recall, and f1 score
- **Design**:
  - dataset generation with 3 levels of 'is_hit' = 1 proportions
  - 20 iterations per level, 5000 samples each
- **Results**
  - Lack of visible results
  - Likely due to proportion of hit to non-hit songs
  - Trained on huge amount of hit songs, → high recall, low precision

	Model	Prevalence	Mean_Accuracy	SD_Accuracy	Mean_F1_Score	SD_F1_Score	Mean_Precision	SD_Precision	Mean_Recall	SD_Recall
0	ANN	0.01	0.98925	0.002673	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1	ANN	0.20	0.80200	0.009542	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	ANN	0.55	0.53505	0.014529	0.663646	0.025910	0.547713	0.015740	0.846307	0.075711
3	RandomForest	0.01	0.98925	0.002673	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	RandomForest	0.20	0.80200	0.009465	0.000476	0.002130	0.050000	0.223607	0.000239	0.001070
5	RandomForest	0.55	0.51810	0.014385	0.616899	0.013457	0.544628	0.017254	0.712371	0.026998

Question 1

Best features to predict popularity?

- For select genres:
  - **Jazz**: loudness, instrumentalness, acousticness
  - **Country**: energy
  - **Rock**: tempo, acousticness, valence
  - **Pop**: acousticness, speechiness, tempo

Question 2

Different measures of success per genre?

- Some genres depend on lots of features
  - rock, pop, soul
- Others only needed couple features
  - techno, heavy metal, country

Question 3

“Weird” but popular?

- Deviation from the norm did not impact song’s popularity
- Most songs tend to be certain deviation from cluster center

Learning Process

- **Importance of choosing the best model**
  - some questions require unique solutions/ models → not a one-size-fits-all situation
- **Evaluating perfomance on certain metrics**
  - accuracy is not the best indicator of model performance
  - F1 evaluation metrics help with cohesiveness
- **Documentation**
  - many layers involved with all 4 projects
  - for results and work to be of use → documentation and file organization is essential

Applying Findings in the Real World

Music Production

- aligns producers with popular listener preferences

Streaming Platforms

- recommend songs/artists that align with genre's key features

Music market

- allocate resources and time to songs that align with predicted popularity

Thank you!

Q&A