MUSIC OR LYRICS

a duet about song genre classification

Tanya Shapiro | March 2nd, 2022

The Prelude

WHY THIS PROJECT?

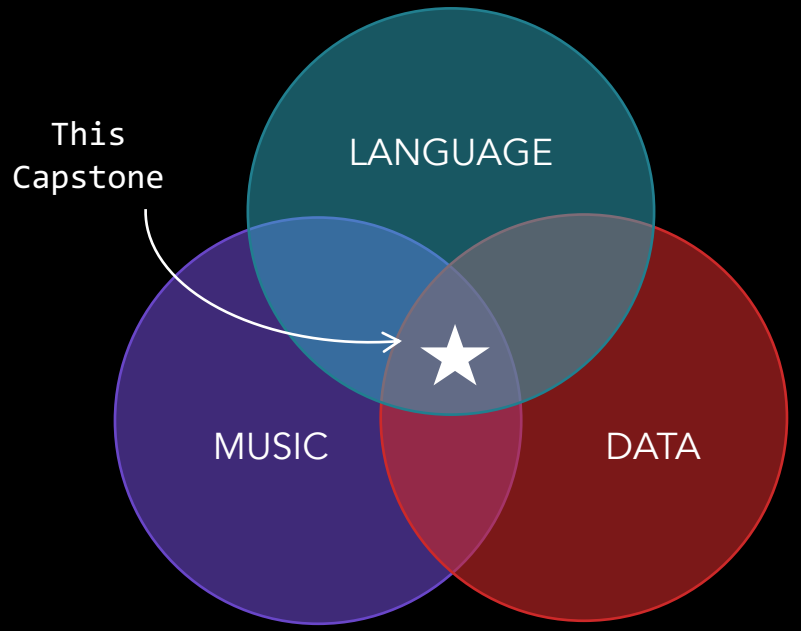
Project is the perfect intersection of my 3 interests:

+ **MUSIC**. Playing guitar since I was 8 years old.
+ LANGUAGE. Shakespeare nerd and lover of NLP.
+ DATA. This one doesn't need explaining…

OTHER PERSONAL MOTIVATORS

- Opportunity to test out data collection with APIs
- Practice more text analysis and NLP techniques
- Build an application to explore song features

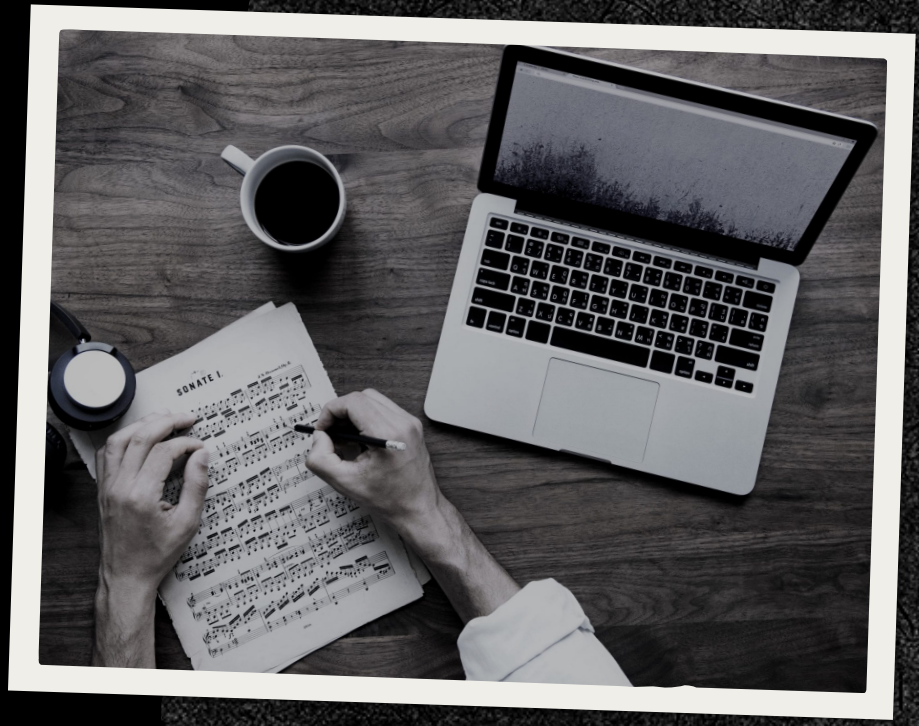
MY 3 PASSIONS



ME & MY LES PAUL

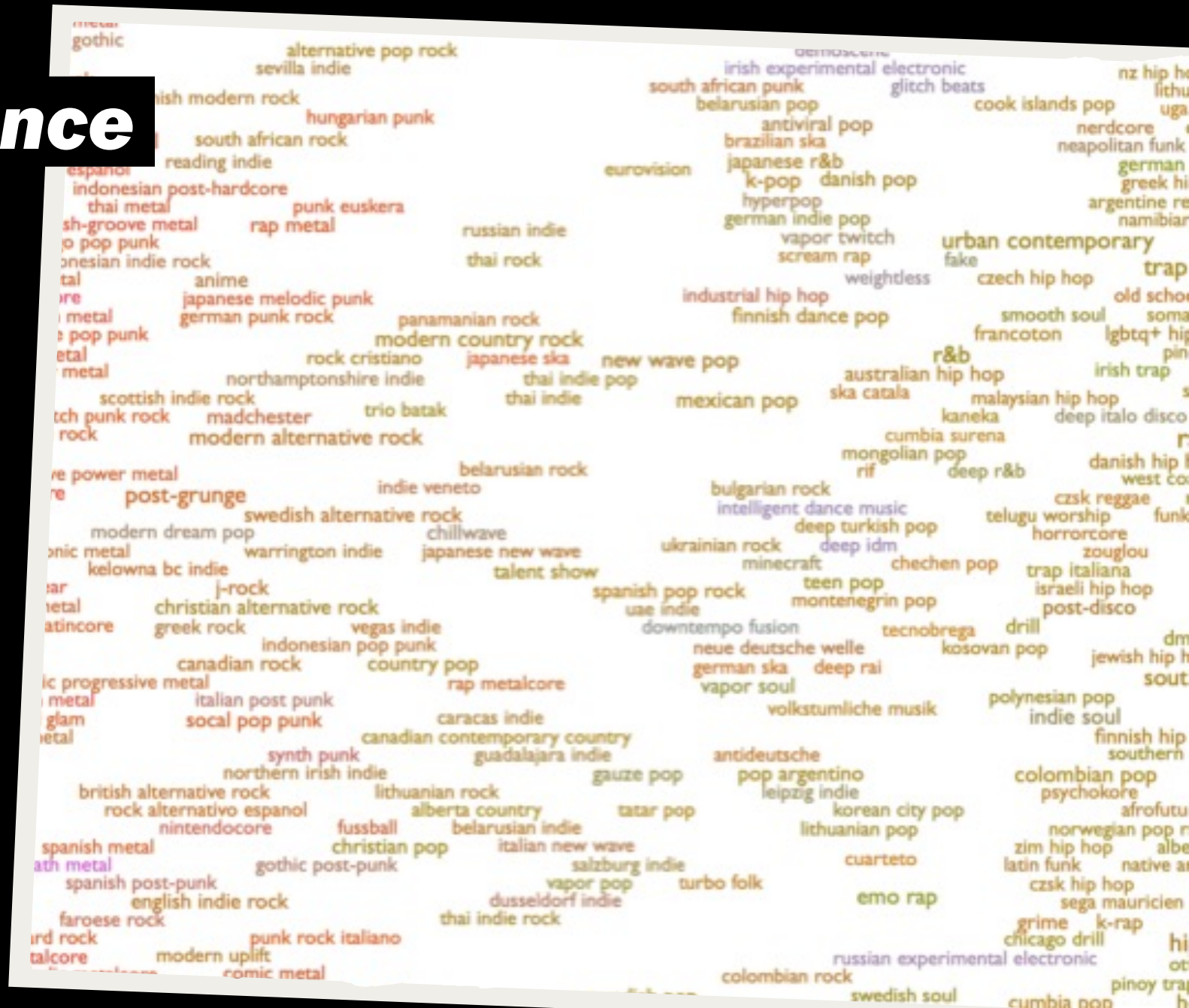
The Problem

- What part of a song is better at determining a song's music genre – the music (i.e. audio features) or the lyrics?
- Create a model that can accurately classify songs with at least 60% accuracy



Every Noise At Once Project

- Started in 2013 by Glennon McDonald
- "…an ongoing attempt at an algorithmically-generated, readability-adjusted scatter-plot of the musical genre-space, based on data tracked and analyzed for 5,782 genre-shaped distinctions by Spotify as of 2022-03-01"



Notes About The Data

Genre classification based on

Every Noise At Once Project

Based on 4 “The Sound of” playlists: Country, Dance Pop, Hip Hop, and Rock.

Collected audio feature data using **Spotipy**, Python wrapper for Spotify’s API.

Deduped songs overlapping genres.

Song lyrics collected using **Lyric Genius API**. Search based on Artist Name and Song Title.

Additional data cleaning.

Total Songs: 3,425

DATA COLLECTION

PRE-PROCEESSING

MODELING

RESULT ANALYSIS

Spotify API
Audio Data

Genius API
Lyric Data

Music Data
Total Songs – 3,425

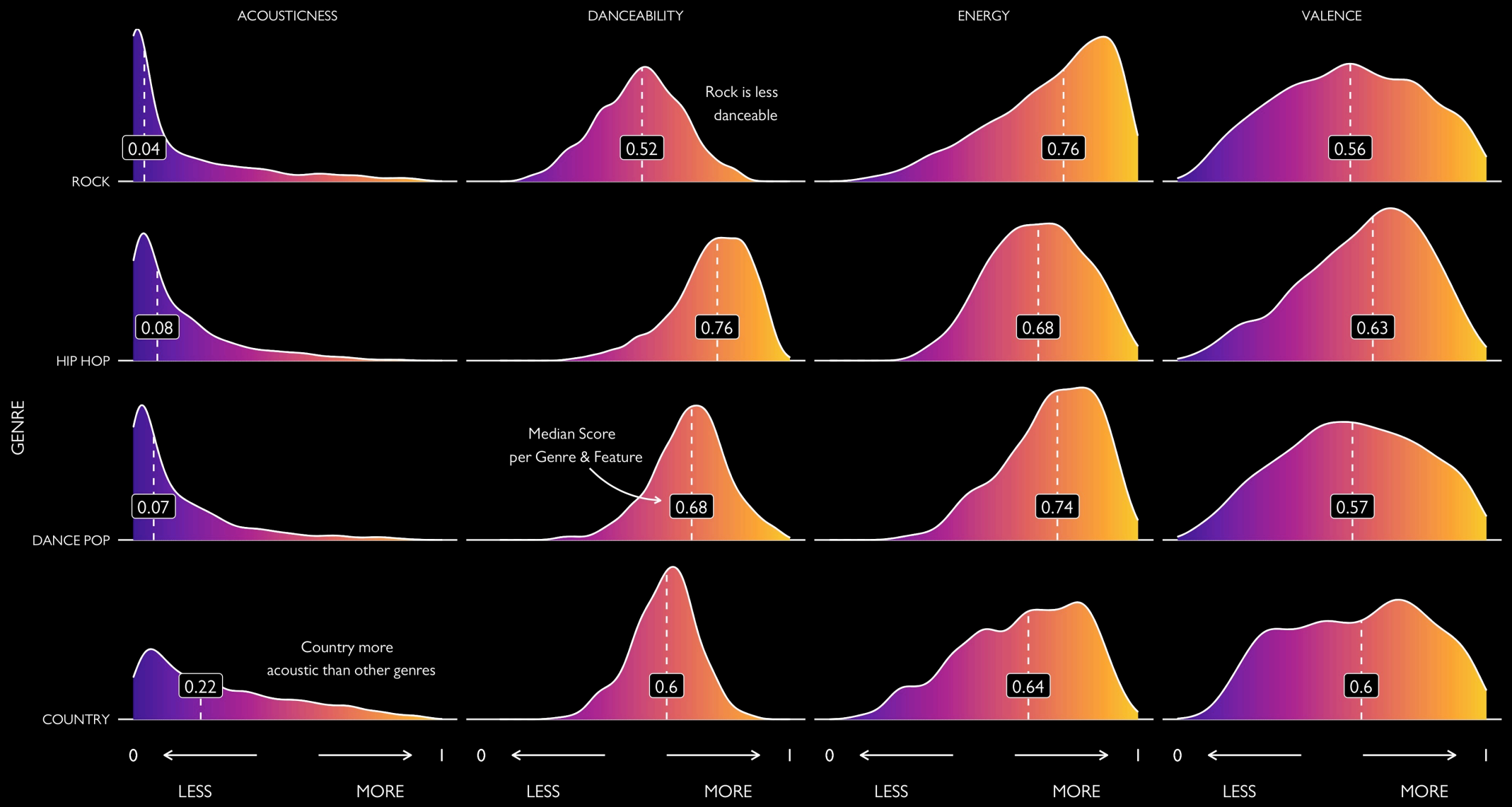
Audio Feature Models

Lyric NLP Models

Determine Best Model

MUSIC GENRE AUDIO PROFILES

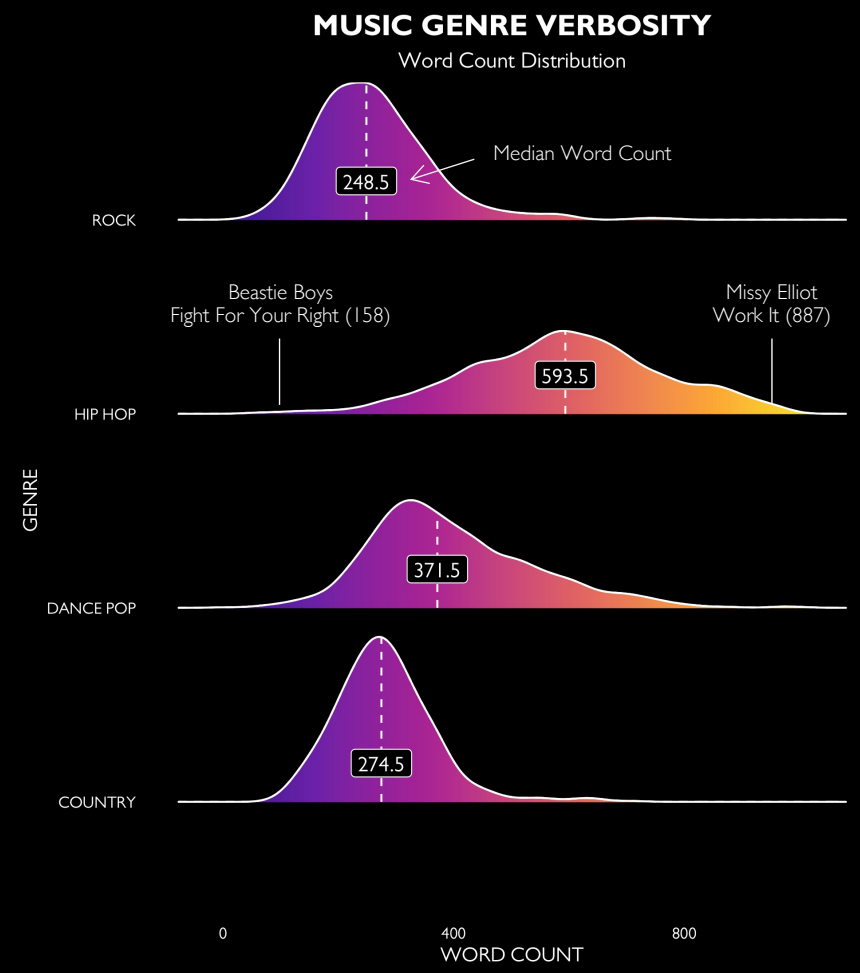
Audio features created by Spotify, scaled from 0 to 1

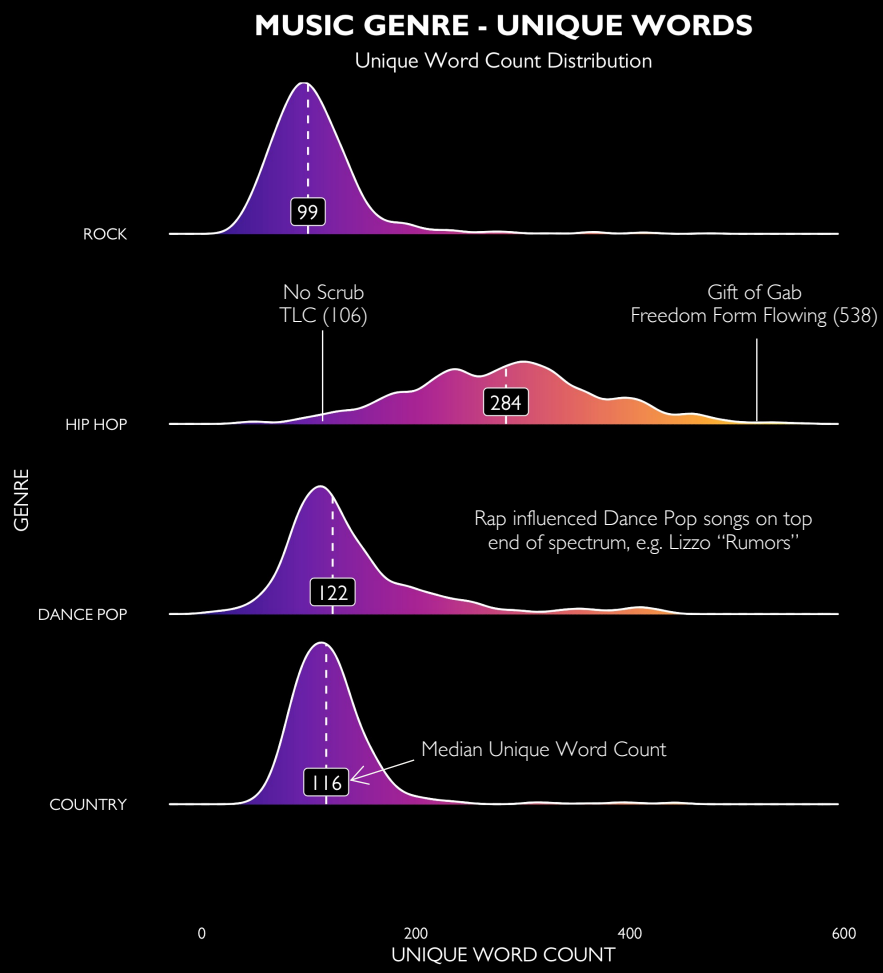


Data from Spotify API | Chart by @tanya_shapiro

Lyric Analysis

Hip Hop More Verbose Compared To Other Genres





Shiny Application

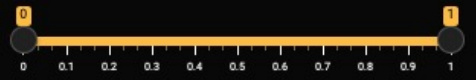
FILTERS

Change the inputs below to change results displayed in the main panel.

Music Genre

Country | Dance Pop | Hip Hop | Rock

Valence



Danceability

Energy

Acousticness

Instrumentalness

Speechiness

Top Grams | Word Cloud | Cross Analysis | Audio Profile

Search

	ARTIST	TRACK	GENRE	VALENCE	DANCEABILITY	ACOUSTICNESS	ENERGY
	The Doobie Brothers	What a Fool Believes	Rock	0.99	0.76	0.28	0.38
	Madonna	Material Girl	Dance Pop	0.98	0.74	0.33	0.88
	Austin Mahone	Mmm Yeah (feat. Pitbull)	Dance Pop	0.98	0.71	0.00	0.92
	Keyshia Cole	Last Night	Dance Pop	0.97	0.92	0.17	0.86
	John Mellencamp	Hurts So Good	Rock	0.97	0.79	0.04	0.74
	The Rolling Stones	Start Me Up - Remastered 2009	Rock	0.97	0.63	0.04	0.93
	Chuck Berry	Johnny B. Goode	Rock	0.97	0.53	0.74	0.80
	Mary J. Blige	Family Affair	Dance Pop	0.97	0.91	0.13	0.55
	Iggy Pop	Real Wild Child (Wild One)	Rock	0.97	0.59	0.02	0.85
	Ginuwine	Pony	Hip Hop	0.97	0.75	0.00	0.61

1–10 of 1552 rows

Previous 1 2 3 4 5 ... 156 Next

Audio Feature Model Results

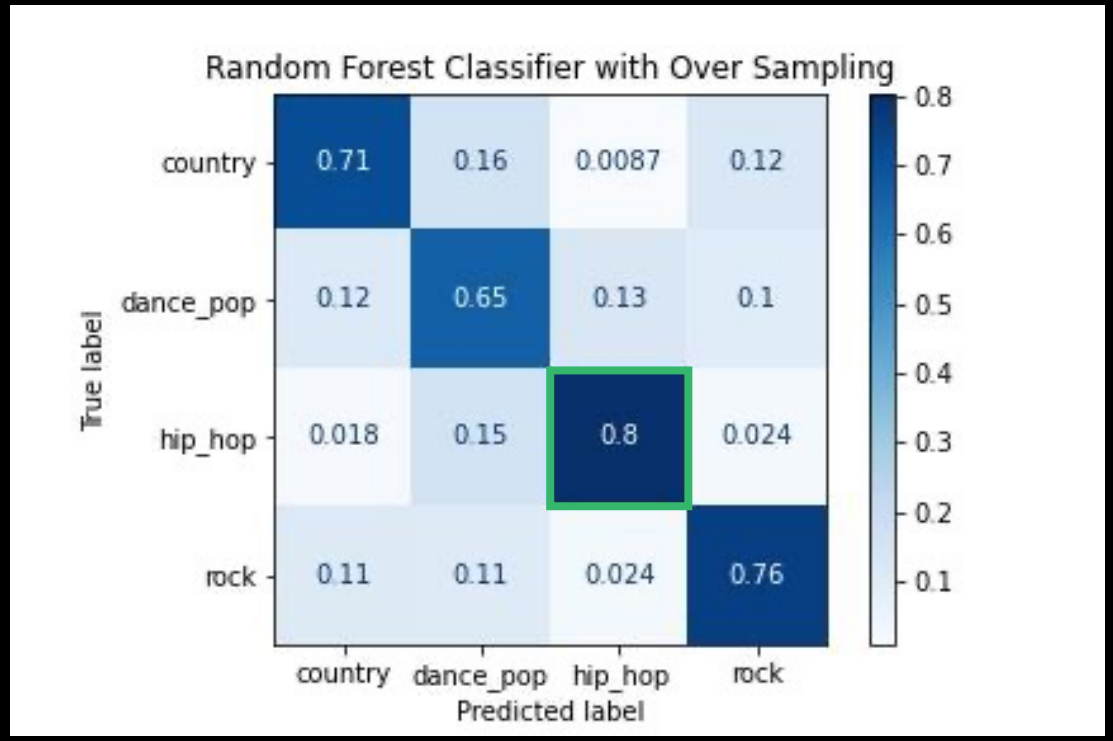
Model	Over-Sampling	Train Accuracy	Overall Accuracy	Accuracy Per Class			
				Country	Dance Pop	Hip Hop	Rock
Random Forest Classifier	**Yes**	**100.0%**	**73.1%**	**71.3%**	**65.5%**	**80.4%**	**75.6%**
Random Forest Classifier		100.0%	72.3%	62.6%	66.0%	81.0%	76.6%
Decision Tree Classifier		72.2%	66.9%	62.6%	57.4%	78.0%	69.3%
Decision Tree Classifier	Yes	73.1%	64.8%	73.0%	52.8%	73.8%	64.4%
KNN		50.5%	34.9%	19.1%	51.8%	31.5%	30.2%
KNN	Yes	54.1%	31.4%	34.8%	36.5%	33.9%	22.4%

Best Audio Models

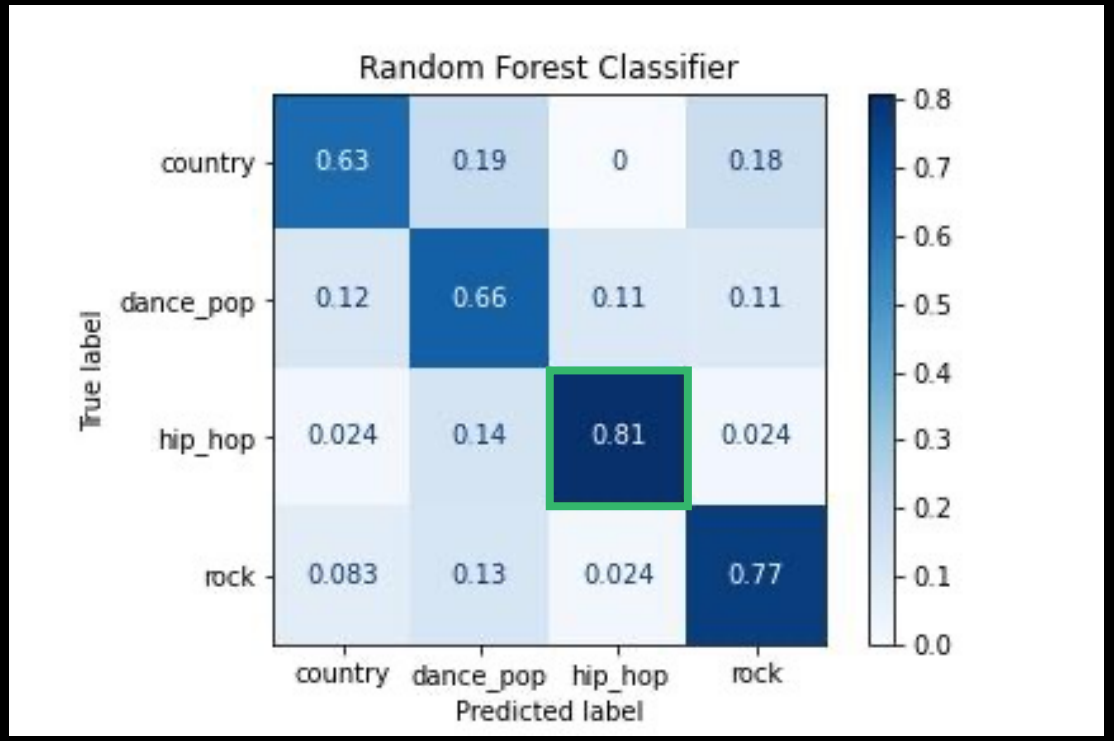
Random Forest Best Model for Audio Feature Data. Oversampling improved accuracy by 1%.

Hip Hop has best accuracy across all genres for both models (80-80%).

Random Forest (Oversampling) – R2 73%



Random Forest – R2 72%



Lyric Model Results

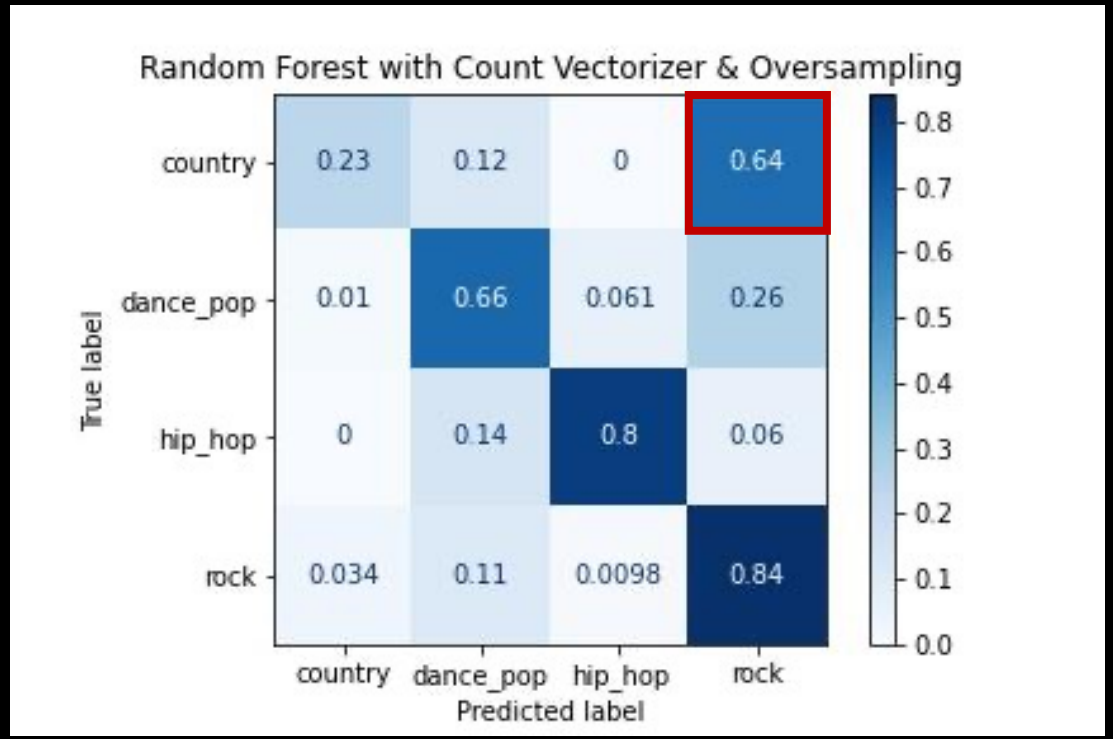
Model	Over-Sampling	Train Accuracy	Overall Accuracy	Accuracy Per Class			
				Country	Dance Pop	Hip Hop	Rock
Random Forest with CV	**Yes**	**99.8%**	**67.9%**	**23.5%**	**66.5%**	**79.8%**	**84.4%**
SVM with TFID	Yes	91.8%	66.9%	62.6%	70.6%	71.4%	62.0%
SVM with TFID		89.9%	66.7%	52.2%	70.1%	70.2%	68.8%
Random Forest with TFID	Yes	99.8%	65.1%	27.8%	61.4%	80.4%	77.1%
Random Forest with CV		99.8%	65.0%	8.7%	65.5%	79.8%	83.9%
Random Forest with TFID		99.8%	64.5%	13.0%	62.4%	79.8%	82.9%
SVM with CV	Yes	99.8%	62.8%	53.9%	57.9%	70.2%	66.3%
SVM with CV		99.8%	62.8%	53.9%	57.9%	70.2%	66.3%

Best Lyric Models

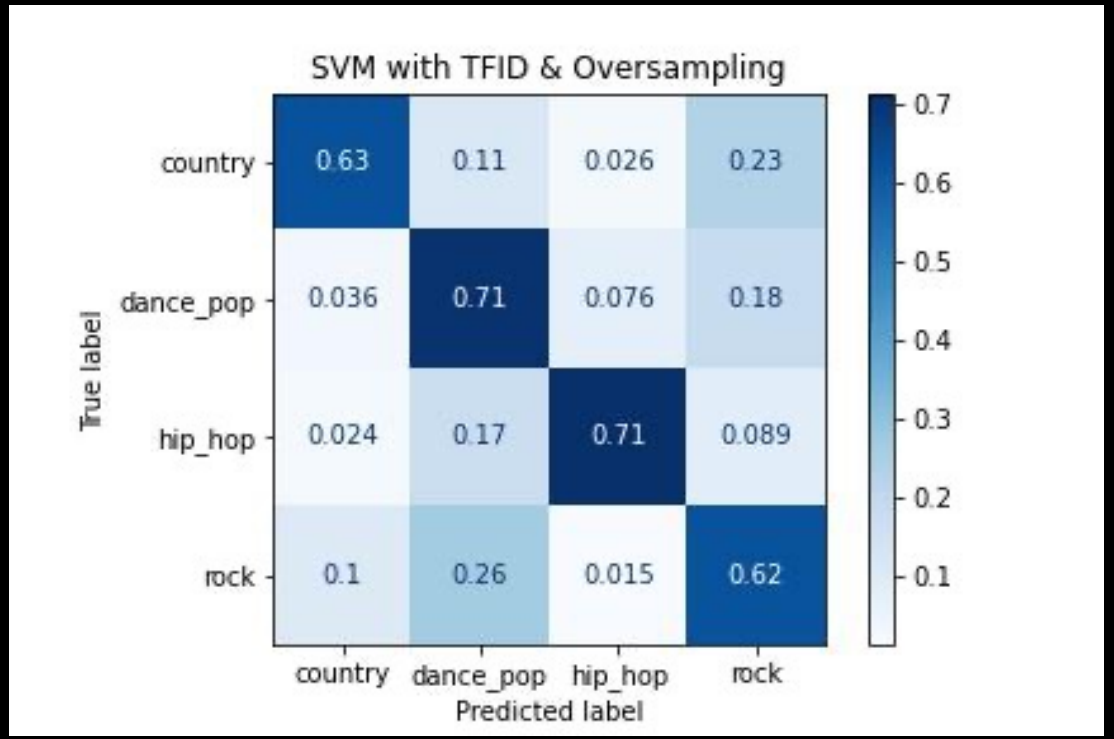
RFC – skewed range of accuracy results, Hip-Hop & Rock good, Country most often misclassified as Rock

SVM accuracy more evenly distributed across classes

RFC with CV (Oversampling) – R2 68%



SVM with TFID (Oversampling) – R2 67%



Outro - Key Findings

- **Close Call**. Best Audio Features model beat out best Lyric model, only by 5% points
- **Best Overall**. Although Random Forest (Audio) had the highest accuracy (73%), SVM Audio should not be overlooked – more balanced across genres with accuracy.
- **Oversampling**. This technique helped improve most model accuracy by ~1-2 % points
- **Accuracy by Genre.**
  + Hip Hop songs were classified correctly more often compared to songs in other genres (true for both audio and lyrics)
  + Country commonly misclassified as rock – small sample size issue or close genre overlap?

Encore – Ideas for Future Research

- **Model Tuning** – work with finding best parameters settings for different models
- **Songs Overlap Genres** – this problem might be better suited for multilabel classification
- **Get More Data** – sample sizes & class imbalances can impact model performance.
- **Genre Evolution** – future research should take into consideration time (Rock in the 60s vs. 80s) and sub-genre classifications (Rock might be too broad – Grunge vs. Indie?)
- **Exploring Rhyme Patterns** – is there a relationship between rhyme patterns and genre?
- **Combining Audio & Lyrics** – what would the modeling results look like if we tested both?
- **Additional Applications** – option to explore creating Recommendation Systems based on song features