

# Predicting Hip-Hop Popularity

*Nicholas Indorf*



# Client & Issue

## *KC Makes Music*

- ❖ 2018-present, 30 songs
- ❖ ~24.5k monthly listeners
- ❖ ~10 songs finished, unreleased

## *Popularity*

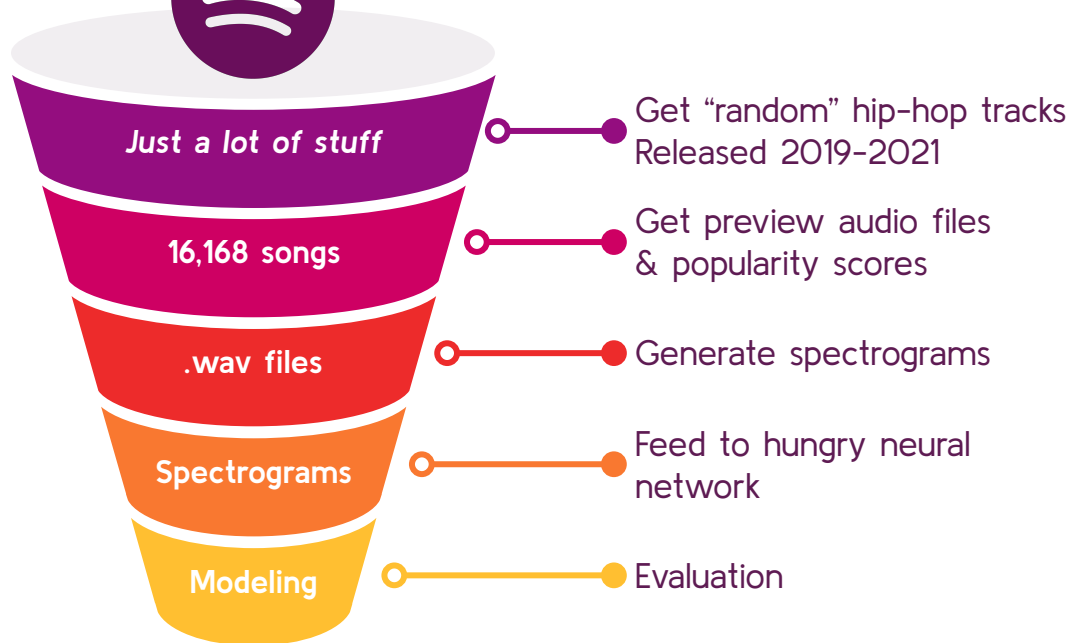
*based solely on audio*



# Data Gathering & Preparation

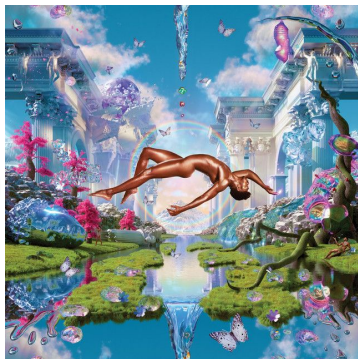


*Spotify Web API*

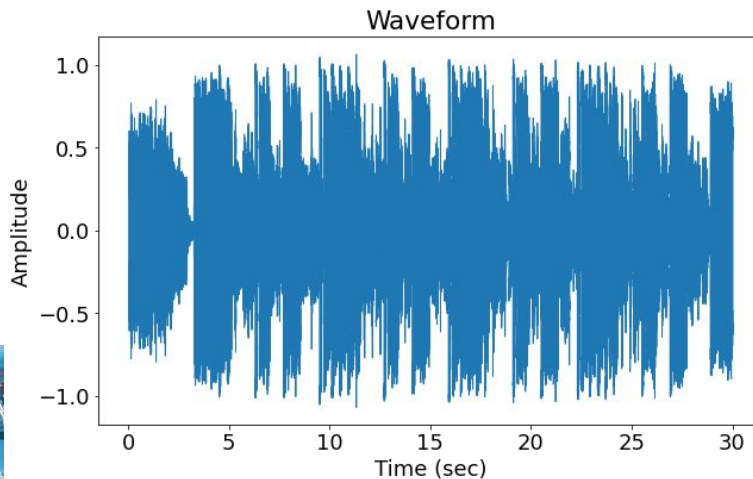


**Key Data:**  
*Preview audio & Popularity scores*

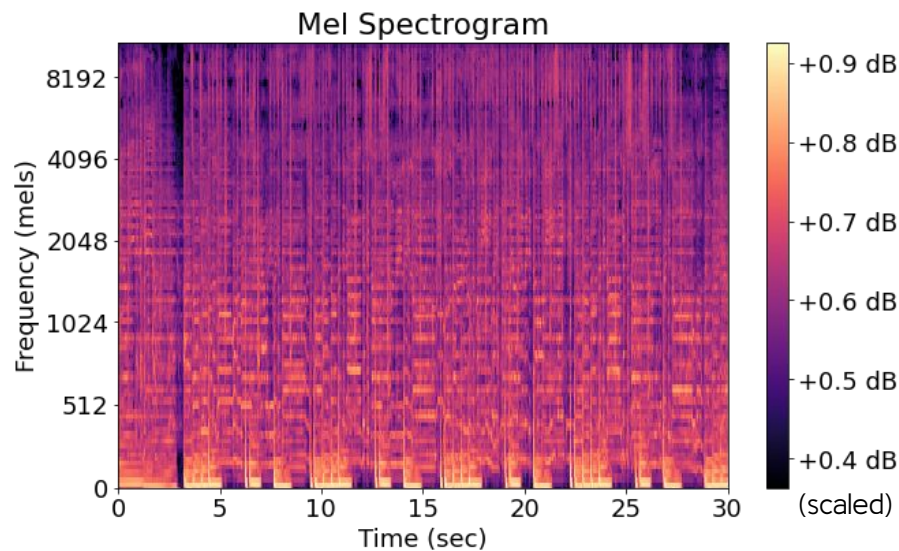
# Preview Audio



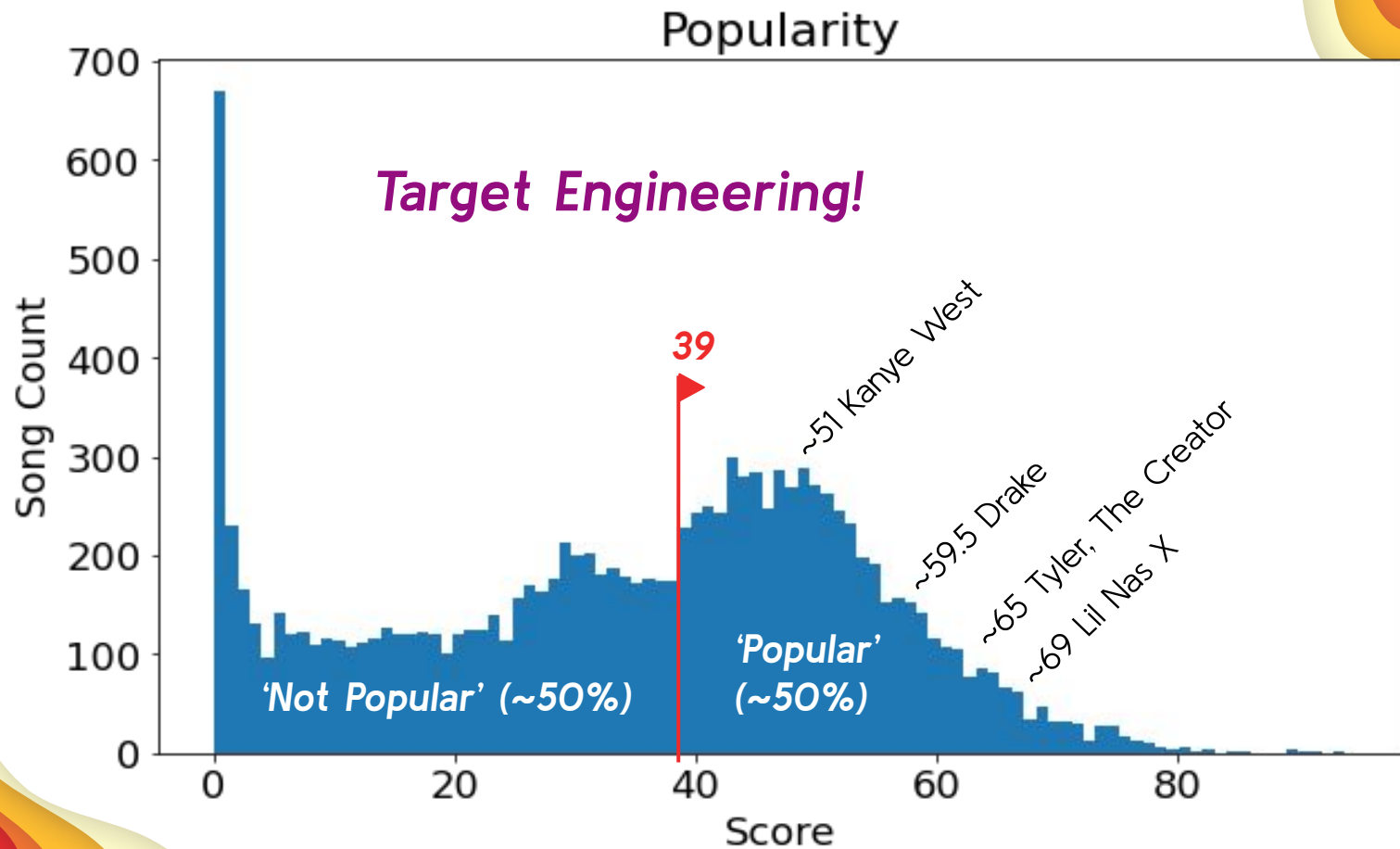
INDUSTRY BABY  
(feat. Jack Harlow)  
- Lil Nas X



~30 Seconds

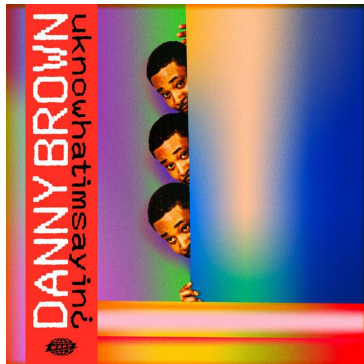


# Popularity





## Not Popular



Uknowwhatimsayin?  
- Danny Brown

**Pop = 35**



Look Over Your  
Shoulder (feat.  
Kendrick Lamar)  
- Busta Rhymes

**Pop = 33**

## Popular



INDUSTRY BABY  
(feat. Jack Harlow)  
- Lil Nas X

**Pop = 90**



WUSYANAME (feat.  
Youngboy Never  
Broke Again & Ty  
Dolla \$ign)  
- Tyler, The Creator

**Pop = 79**

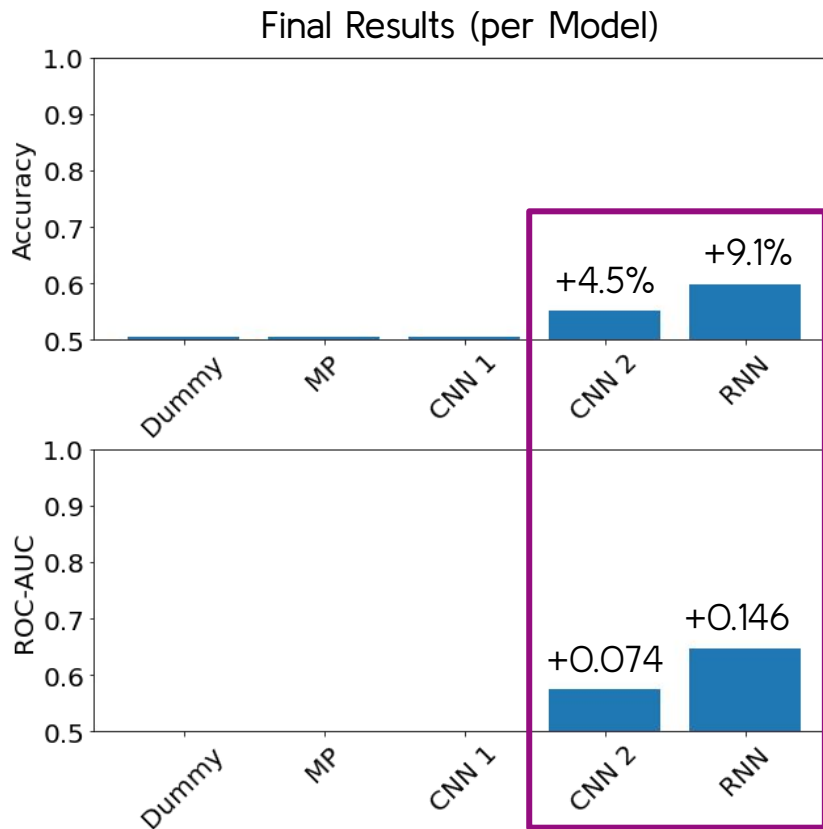
# Modeling & Evaluation

❖ Model: Neural Network

❖ Metric:

➤ Accuracy

➤ ROC-AUC



Promising!

Increasing model complexity →

# Discussion & Recommendation

- ❖ Only audio → **very limiting**
- ❖ Popularity → **confusing**
- ❖ Exposure is huge → TikTok, etc.

"I'm not surprised"

**Client**

***Sieve-type model &  
Continued development***



# Next Steps

- ❖ Incorporate non-audio features
  - E.g. exposure, release timing
- ❖ Expand dataset
- ❖ Feature importance analysis
  - What did it pick up on?

# Thank You



[linkedin.com/nicholas-indorf-data-scientist](https://www.linkedin.com/nicholas-indorf-data-scientist)



[nicholasindorf@gmail.com](mailto:nicholasindorf@gmail.com)



[github.com/Nindorph](https://github.com/Nindorph)

# Apdx. - Context

## Audio

### Music Genre Classification using Transfer Learning(Pytorch)



Aryan Khatana



Jun 26, 2020 · 6 min read



#### TensorFlow Speech Recognition Challenge

Can you build an algorithm that understands simple speech commands?



Google Brain · 1,313 teams · 4 years ago

#### Cornell Birdcall Identification

Build tools for bird population monitoring



Cornell Lab of Ornithology · 1,390 teams · a year ago

## Fine Detail

### Pump it Up: Data Mining the Water Table

HOSTED BY DRIVENDATA

ARTICLE



<https://doi.org/10.1038/s41467-020-15432-4>

OPEN

### Automatic diagnosis of the 12-lead ECG using a deep neural network

Antônio H. Ribeiro<sup>1,2✉</sup>, Manoel Horta Ribeiro<sup>1</sup>, Gabriela M. M. Paixão<sup>1,3</sup>, Derick M. Oliveira<sup>1</sup>, Paulo R. Gomes<sup>1,3</sup>, Jéssica A. Canazart<sup>1,3</sup>, Milton P. S. Ferreira<sup>1</sup>, Carl R. Andersson<sup>2</sup>, Peter W. Macfarlane<sup>4</sup>, Wagner Meira Jr.<sup>1</sup>, Thomas B. Schön<sup>2✉</sup> & Antonio Luiz P. Ribeiro<sup>1,3✉</sup>

## Similar



### Predicting Popularity on Spotify — When Data Needs Culture More than Culture Needs Data

A short, step-by-step stroll through an Introductory Machine Learning project using Spotify data.



Philip Pekar Jun 25 · 14 min read



Audio + Finely detailed insight

**Very Difficult!**

(Links to these projects on my GitHub)