
IS THERE A CORRELATION BETWEEN THE DESCRIPTIVE METRICS OF SONGS PROVIDED BY SPOTIFY AND SONG POPULARITY ON VARIOUS PLATFORMS?

PROJECT PROPOSAL

Adam Gauthier

100947233

Department of Electronics

Carleton University

Ottawa, Ontario

adamgauthier@cmail.carleton.ca

Alex Mougeot

101071728

Department of Electronics

Carleton University

Ottawa, Ontario

alexmougeot@cmail.carleton.ca

Andrew Dybka

101041087

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario

andrewdybka@cmail.carleton.ca

Jacob Laboissonniere

101031913

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario

jacoblaboissonniere@cmail.carleton.ca

Laura Colley

101054805

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario

lauracolley@cmail.carleton.ca

Nayeeb Mowla

101041742

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario

nayeebmowla@cmail.carleton.ca

February 12, 2021

ABSTRACT

The goal of this project is to investigate the correlation between metrics that describe songs and their popularity via several platforms. Song metrics will be obtained from Spotify's built-in features, and popularity evaluated on the platforms TikTok, WhoSampled, YouTube, Billboard Top 40, Twitter, and Google using various popularity metrics.

Keywords music · Spotify · song metrics · popularity

Detailed Problem Statement

The application Spotify contains a wide range of songs, intended to fit any user's music taste and preferred genre. To determine whether the future popularity of these songs can be predicted based on their inherent characteristics, correlations must be investigated between a known set of song metrics and song popularity at various times on several platforms. Analysis does not consider the factor of an artist's previous popularity but is restricted to the innate qualities of the song itself, such as acousticness, danceability, and key.

The specific question that will be investigated is: *what musical factor was maximized among the most popular songs measured via the platforms TikTok, WhoSampled, YouTube, Billboard Top 40, Twitter, and Google?*

As the question above will be evaluated using numerical metrics of song usages and popularity, ethical considerations are slim. Any data that could be connected to a specific user (i.e. Twitter, TikTok, or YouTube usernames) will be anonymized before data analysis occurs.

Motivation & Value

The results from this project can provide insight for artists and content creators into the type of songs they should create. Artists may use the results to better target their goals for their songs and content. TikTok creators, for example, can use the results to see which songs typically provide the best success for views and for starting trends. Musical artists can use the results to see the kind of songs that hit the charts in different periods and which types of songs are sampled more or are used in different content more frequently.

Data Set

Data for song metrics will be gathered from “Spotify Dataset 1921-2020, 160k+ Tracks” compiled by Yamac Eren Ay [1]. This dataset contains 16 features that Spotify uses for music categorization and recommendation, including duration_ms, key, mode, time_signature, and danceability, among others. Popularity data trends for certain songs will be collected using APIs and scraping from additional sources. Some examples of these possible sources are the TikTok API, the Google Trends API, and scraping of WhoSampled.com.

References

[1] Yamac Eren Ay. Spotify dataset 1921-2020, 160k+ tracks. May 2020.

Appendix

Project Repository: [View on Github](#)