Sick Beats

Effects of COVID-19 on the Attributes of Top Songs in the US

Mohammed Akel, Jenna Chu, Rocket Drew, Darren Fang
Brown University, Department of Computer Science

Introduction

Over the past two years, COVID-19 has affected every facet of our lives. Even the media we consume has adapted as more people worked from home and gained leisure time. In this vein, we investigated the effect of COVID-19 on music listening habits by examining the differences in attributes of top songs between periods of high and low cases.

Data

Since our project revolves around the impact of COVID-19 on attributes of music, all the data we collected falls within the timeframe of January 28, 2020 to February 28, 2022. We identify 5 days (within the past 2 years) in which the number of new COVID-19 cases was a local minimum, and 5 days in which the number was a local maximum. For each of these days, we found the top 100 ranked songs, and obtained the attributes of each song from Spotify, along with lyrics.

Our project uses five data sources:

- Our World in Data's Daily new confirmed COVID-19 cases per million people in the United States.
- Billboard's Top 100 songs chart, for the 10 days that are a local minimum or maximum
- Song data from Spotify API
- Lyrics data, drawn from Genius lyrics for each of the songs we analyze
- Sentiment analysis from Flair's NLP library

Methodology

Since our project draws on multiple datasets, we first joined the datasets together for our specific needs. We split our COVID-19 data into categorical data—5 days to be categorized as peaks and 5 days as troughs. For these two categories, we obtained the Billboard Top 100 songs in each day, as well as the Spotify song attributes for each song.

Our first analysis focuses on the differences between song attributes during COVID-19 peaks and troughs. Specifically, our statistical testing looks at loudness, energy, and valence, which refers to the positiveness conveyed by a track.

Our second analysis examines the sentiment of song lyrics. For peaks and troughs, we are interested in whether the sentiment of the lyrics differ. As follow-up investigations, we also assessed the word count of the lyrics in both phases.

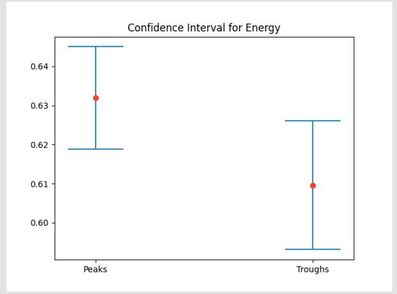
Hypotheses

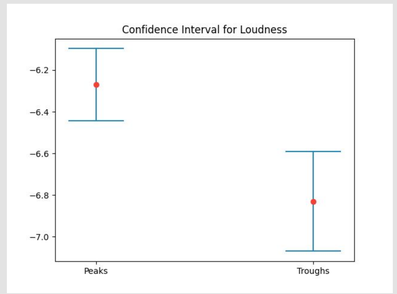
We expect that popular songs during COVID-19 peaks will have higher loudness, valence, and energy attributes than popular songs during COVID-19 troughs.

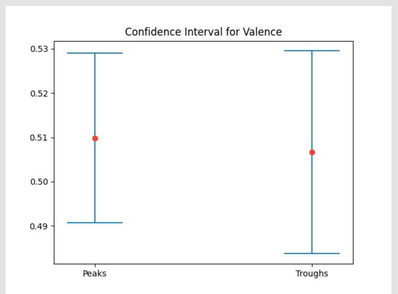
We expect that popular songs during COVID-19 peaks will have more negative lyric sentiment than popular songs during COVID-19 troughs.

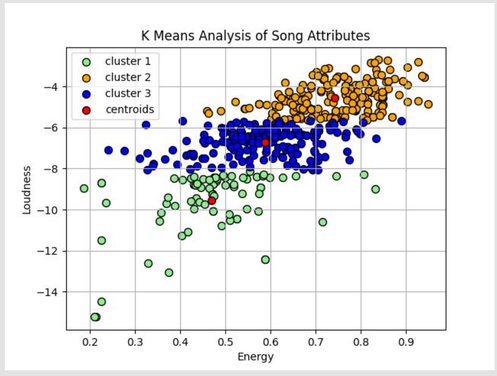
Song Attribute Analysis

To test our hypotheses, we conducted a number of statistical tests. We selected three attributes to represent our dataset—energy, loudness, and valence. Energy represents a perceptual measure of intensity and activity, loudness represents the volume of a song in decibels, and valence represents the musical positiveness conveyed by a track. For each of the three characteristics, we ran a two-tailed independent t-test not assuming equal variance. The confidence intervals of these attributes are displayed in the three figures below, and the results of our hypothesis tests are shown in the subsequent table. Overall, we found that energy and loudness are statistically significantly different between peaks and troughs, while valence does not yield statistical significance. We can conclude that the loudness and energy of songs are higher during peaks of new COVID-19 cases in the United States.









Summary of Hypothesis Testing Values		
Attribute	**P-Value**	**T-Stat**
Energy	0.03622	2.09800
Loudness	0.00020	3.72869
Valence	0.82926	0.21572

For the machine learning portion, we ran a k-means analysis on two song attributes to visually identify correlation. We chose to run our k-means analysis on loudness and energy since these were the two attributes with statistically significant differences. The results can be show in the k-means figure on the left. Our elbow plot suggested that the data can be best represented with three clusters. There is a positive correlation between loudness and energy of a song. This aligns with our hypotheses tests, which concluded that top songs during a peak of new COVID-19 cases have lower energy and lower loudness values.

Lyric Sentiment Analysis

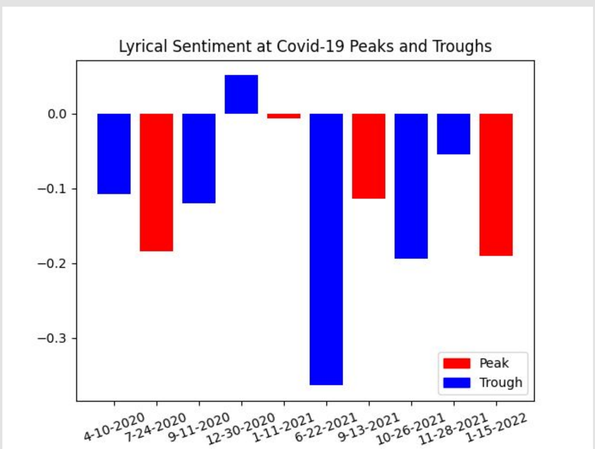
We investigated the impact of COVID-19 on the lyrics of songs that were popular during peaks and troughs.

After scraping lyrics from Genius.com, we used Flair's NLP library to estimate the average sentiment of the songs in each period as an index value in [-1, 1].

We found that, despite visibly high variance in the sentiment across periods, there was **no statistically significant difference** between peaks and troughs.

Consider, for example, the 9-11-2020 and 9-13-2021 samples. Almost exactly one year apart (and therefore loosely controlling for season), the former corresponded to a trough and the latter to a peak, but the difference in sentiment was negligible.

We also tested for differences in the count of total words and unique words, but **we did not find a statistically significant difference** in these metrics.



Conclusions

We found evidence of correlation between the number of new COVID-19 cases and the aura of songs people listen to. In particular, we looked at the Spotify generated attributes of energy, loudness and valance. We found that popular songs during COVID-19 peaks tend to be louder and more energetic than popular songs during COVID-19 troughs. In addition, we observed that there is no difference in valence for popular songs during COVID-19 peaks and troughs.

We also conducted lyrical sentiment analysis on the top songs that were popular during COVID-19 peaks and troughs. We found no statistically significant difference.

Limitations of Analysis

Our data suffered from some confounding variables. The time of year has a large impact on both new COVID-19 cases and types of songs being released. For example, two peaks occurred during the winter holiday season when holiday songs dominate the Billboard charts. Though COVID-19 remained prevalent since January 2020, our data may still have been impacted by seasonal listening habits.

Billboard Top 100 reliably aggregates data from multiple sources to identify the most popular songs for a given week. We are concerned, however, that Billboard over-represents new releases in their lists, thereby limiting the Top 100's ability to accurately capture national listening habits. If new releases are concentrated around a certain time of year, this distribution may present an additional confounding variable.

Our identification strategy for coding "high-Covid" and "low-Covid" periods also raises questions. We chose to examine peaks and troughs, but focusing on these local extrema potentially obscured the absolute number of confirmed cases at each period. For example, there were more cases during the "trough" at October 2021 than during the "peak" at July 2020. We also relied on new daily cases, rather than considering total active cases on a given day.

Acknowledgements

We would like to acknowledge Professor Lorenzo De Stefani, our project mentor Yutong Liang, as well as the CS1951A TAs for guidance throughout this project.