BANA 201A: Statistics for Data Science

Professor Rick So

Group Project Report

Team 12B

Becky Wang

Vy Nguyen

Dennis Wu

Hsiang-Han (Cyan) Huang

Kenjee Koh

Introduction

Research Question

In this project, our team wants to know if **listeners in the U.S. shift from upbeat, high-energy music to more mellow genres after the COVID-19 pandemic**.

Interest and Forecast benefits: 1. People faced much more levels of stress, anxiety and isolation during the pandemic, their music preferences may have shifted. 2. Analyzing these shifts can provide insights into the psychological and social impacts of crises on cultural consumption.

Difficulty: 1. Range the pre-covid and post-covid period is subjective. 2. Define the song's genre as we don't have enough reference value and volume of ranking songs.

Data Collection

Goal

To conduct our analyses, we needed to be able to reference datasets for each of the pre-COVID years (which we defined as the years 2018 and 2019) as well as the post-COVID years (which we defined as the years 2022 and 2023). These datasets needed to not only include the top streamed songs of each of these years, but also include metric information about each song's audio features (danceability, energy, key, valence, tempo, etc) that would help us define which songs are considered sad/happy.

Python, Spotify and Apple Music APIs

To collect accurate data that provided the information we needed, we leveraged Python (with the use of the Pandas and NumPy libraries) in conjunction with the Spotify and Apple Music APIs. These APIs gave us the ability to "communicate" with the Spotify and Apple Music streaming platforms' servers to retrieve the accurate, official data that we needed. With this, we produced 3 datasets of the top 100 songs of each year (in the format of CSVs for more efficient data analysis with tools such as Excel and SPSS), which included all the in-depth information needed for 2019, 2022, and 2023. Below contains a brief data dictionary covering the data provided by the APIs:

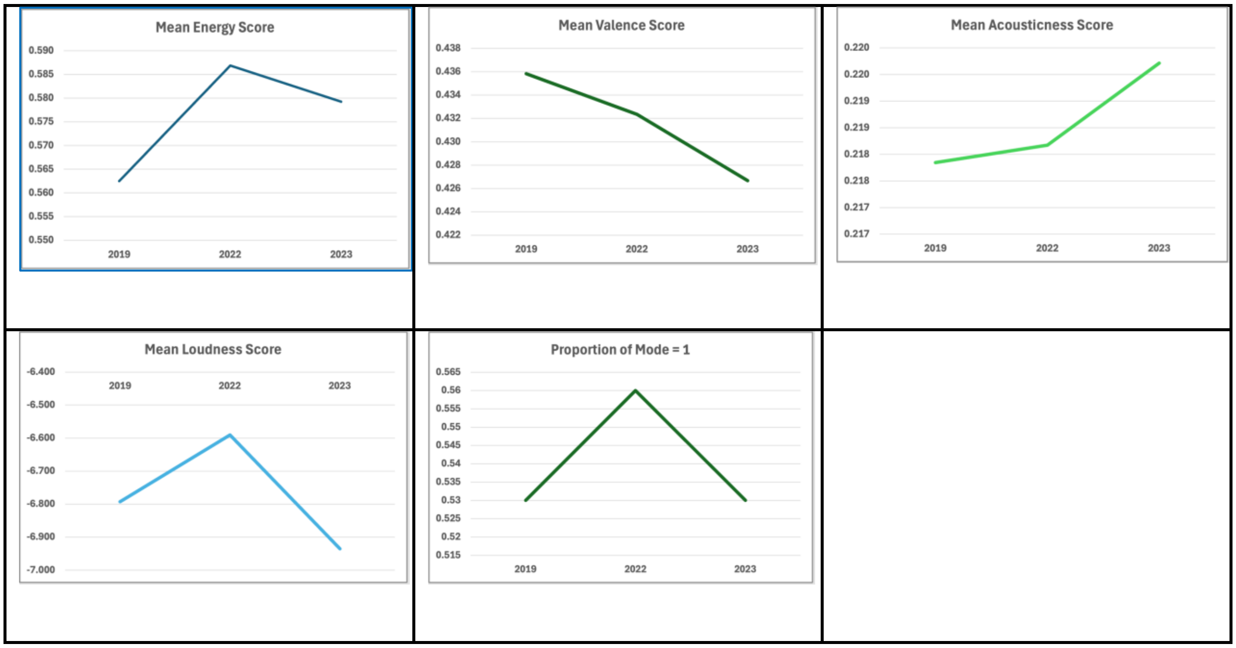
Data Variable	Description
Danceability	Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
Energy	Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has

	high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.
Key	The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key was detected, the value is -1.
Loudness	The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.
Mode	Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.
Speechiness	Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.
Acousticness	A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
Instrumentalness	Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.
Liveness	Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides a strong likelihood that the track is live.
Valence	A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).
Tempo	The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

Statistical Analysis

Overview of Sample Datasets

Before conducting statistical analysis, our team first looked at the sample mean scores for 5 key features that can define whether a song is upbeat or mellow. The line graphs below illustrate the general patterns of sample means of Energy, Valence, Acousticness, Loudness, and Mode throughout 3 years: 2019, 2022 and 2023.



Contrast to our expectations of a shift to more mellow genres post-pandemic, the higher means of Energy, Loudness, and Mode show that US listeners favored more upbeat and high-energy music in 2022 compared to 2019. However, this shift was temporary, as 2023 saw a return to more acoustic and less positive music, with drops in Energy, Valence, Loudness, and Mode, and a rise in Acousticness.

Method 1: 2 Population Means Hypothesis Testing

Despite slight changes in the sample means of the five features (1-2 decimal points), we conducted statistical tests to determine if these differences were significant. Our hypotheses were:

- **Null hypothesis:** Mean feature scores remained the same between 2 years.
- **Alternative hypothesis:** Mean feature scores were different between 2 years.

We performed t-tests for Energy, Valence, Acousticness, and Loudness, comparing means between 2019 and 2022, and 2022 and 2023. For Mode (a binary variable), we tested for two proportions using z-distribution. The results showed that all p-values were greater than any significant level, leading us to fail to reject the null hypothesis, indicating no significant differences between the years. From the hypothesis test, we conclude that US listeners did not shift from upbeat, high-energy music to more mellow genres after the Covid-19.

Interestingly, their music preferences had not significantly changed despite the emotional impact of the pandemic.

Year Comparison (p-value)	Energy	Valence	Acousticness	Loudness	Mode
2019 vs 2022	0.2001	0.9091	0.9916	0.5004	0.6701
2022 vs 2023	0.7093	0.8555	0.9588	0.2208	0.6701

Method 2: Classifying Songs as Upbeat or Sad Using 11 Metrics and a Simplified Scoring System

Method 2 builds on the foundation of Method 1 but applies a more comprehensive evaluation using 11 key metrics. The goal is to classify songs as either upbeat or sad by transforming their feature scores into a simple +1, 0, or -1 system. This method offers a streamlined classification, allowing us to focus on meaningful differences in a song's characteristics while minimizing the impact of minor fluctuations.

Metric	+1 (Upbeat)	0 (Neutral)	-1 (Sad)
Danceability	> 0.7	0.4 - 0.7	< 0.4
Energy	> 0.6	0.5 - 0.6	< 0.5
Key	Major (0, 2, 4, 5, 7, 9, 11)	Ambiguous/Neutral	Minor (1, 3, 6, 8, 10)
Loudness	> -5 dB	-5 dB to -10 dB	< -10 dB
Mode	Major mode (1)	Ambiguous/Neutral	Minor mode (0)
Speechiness	< 0.33	0.33 - 0.66	> 0.66
Acousticness	< 0.3	0.3 - 0.5	> 0.5
Instrumentalness	< 0.5	0.3 - 0.5	> 0.5
Liveness	< 0.6	0.6 - 0.8	> 0.8
Valence	> 0.6	0.35 - 0.6	< 0.35
Tempo	> 120 BPM	90 - 120 BPM	< 90 BPM

Using our classification system, we analyzed the top 100 streamed songs for 2019, 2022, and 2023. The results showed that in 2019, 92% of songs were classified as upbeat and 8% as sad. In 2022, the proportion of upbeat songs slightly increased to 94%, with 6% sad, while in 2023, the proportions returned to 92% upbeat and 8% sad. To assess whether there was a significant change in the proportion of upbeat songs across these years, we performed a z-test comparing the proportions between 2019 and 2022, and 2022 and 2023. The null hypothesis (H0) was that the proportion of upbeat songs remained the same between the two years, while the alternative hypothesis (Ha) suggested a difference. The p-value of 0.5794 was significantly greater than the 0.05 threshold, leading us to fail to reject the null hypothesis. Therefore, we fail to reject the null hypotheses across all features and cannot conclude that there is statistically significant difference in the proportion of upbeat songs between 2019, 2022, and 2023, consistent with the findings of Method 1.

Interpretation: Several factors may explain the stability in listeners' preferences for upbeat music. Certain genres and styles, such as pop, hip-hop, and electronic dance music, continue to dominate streaming charts. Additionally, during challenging periods like the Covid-19 pandemic, listeners may have turned to upbeat music as a form of mood regulation. Cultural and environmental factors, such as the music played in public spaces like coffee shops, gyms, and bars, where upbeat music is often streamed, may also have contributed to the consistently high proportion of upbeat songs.

Conclusion

Summary of Findings: Based on our analysis, we found no statistically significant change in music preferences between 2019, 2022, and 2023. Both of our methods led to the same conclusion: U.S. listeners did not shift from upbeat music to more mellow genres following the Covid-19 pandemic.

Limitations and Implications: While we used generally accepted criteria to classify songs as either upbeat or sad, setting precise thresholds for musical metrics is inherently subjective. Additionally, our analysis focused on the top 100 streamed songs, which may skew the results toward popular genres and artists, limiting the generalizability of our findings to niche or less popular genres. Furthermore, we did not analyze lyrical content, which plays a crucial role in conveying a song's emotional tone. Due to the lack of advanced Natural Language Processing (NLP) tools, we were unable to incorporate sentiment or emotion analysis of lyrics. Future research could integrate lyrical analysis using NLP techniques, providing a more comprehensive understanding of a song's emotional impact.