**SPOTIFY TOP 200**

**Team Members:**

Kelvyn Guzman Francisco Zuluaga John Eagan  
Rob Matelyan Ryan Sabon

**Introduction**

The year 2020 was unprecedented for many different industries, including music. On the surface, it is expected that the number of streams, energy of the music, and danceability of the music were highly impacted by the pandemic. According to Spotify, energy is measured from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. 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. Historical data from Spotify has indicated that world events can affect listening trends; however, there are events that have had no impact on music trends.

It is evident that a singular event has the possibility of changing music trends. Therefore, understanding what variables can change trends is crucial. Here an analysis was performed in order to determine the impact of COVID-19 on music trends, such as the number of streams, energy of the music, and danceability. The team compared 2019 and 2020. The team also tracked the impact from the pandemic on the popularity of an individual song to see how a major event impacted the number of streams.

**Methods**

*Data Collection*

Data used for this analysis comes from Kaggle and Spotify through downloading of CSV files (<https://spotifycharts.com/regional/us/weekly/latest>) and Spotify API calls, respectively.

*Exploratory Analysis*

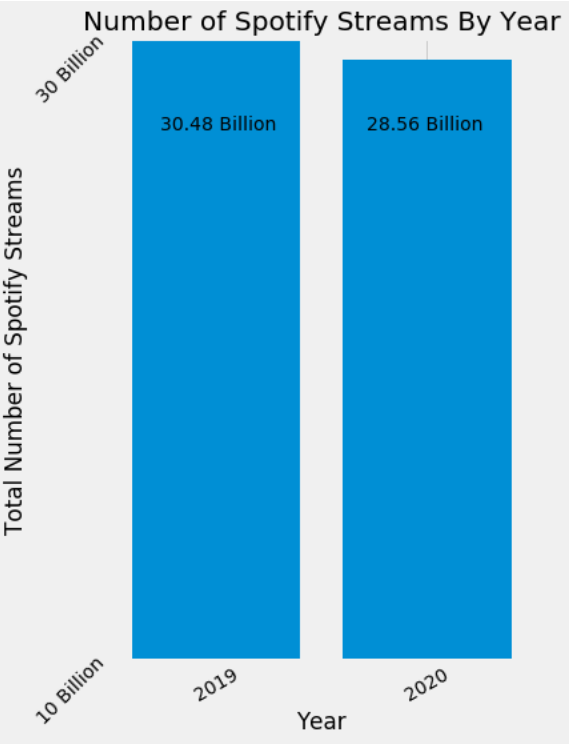
Exploratory analysis was performed in order to identify the quality of data and to determine the proper terms for the graphs and visualizations in order to model a relationship between all music features, danceability and energy,monthly streams and streams by year, and tracking the song “Blinding Lights” by The Weekend. Exploratory analysis consisted of a) transforming the raw data; and b) studying the data.

*Statistical Modeling*

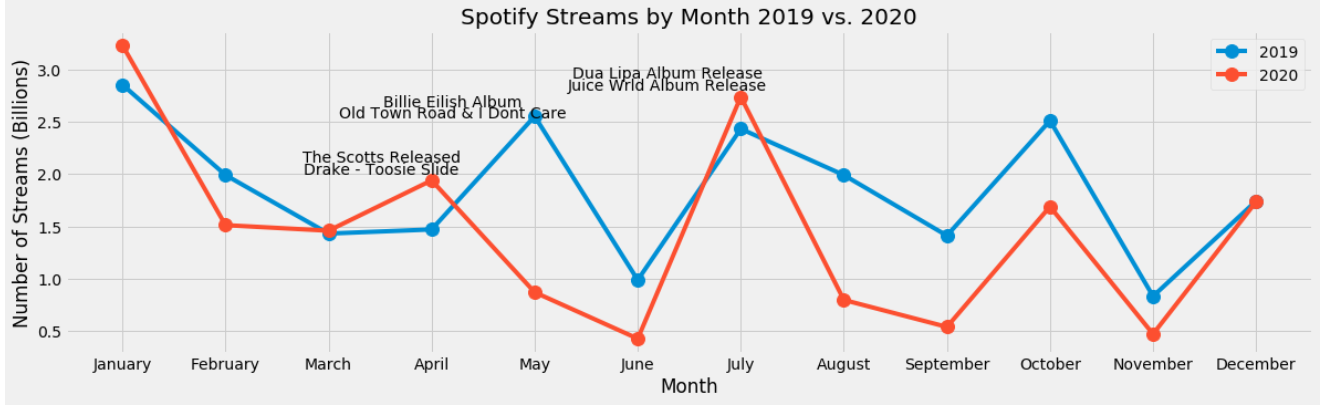
In order to identify the relationships between all music features, a multivariate correlation model was performed. The choice of the model was informed by the results of the exploratory analysis. In order to identify the relationship between danceability and energy, a univariate correlation model was performed based on the results of the exploratory analysis.

**Analysis**

**Streams in the year 2019 and 2020**

****

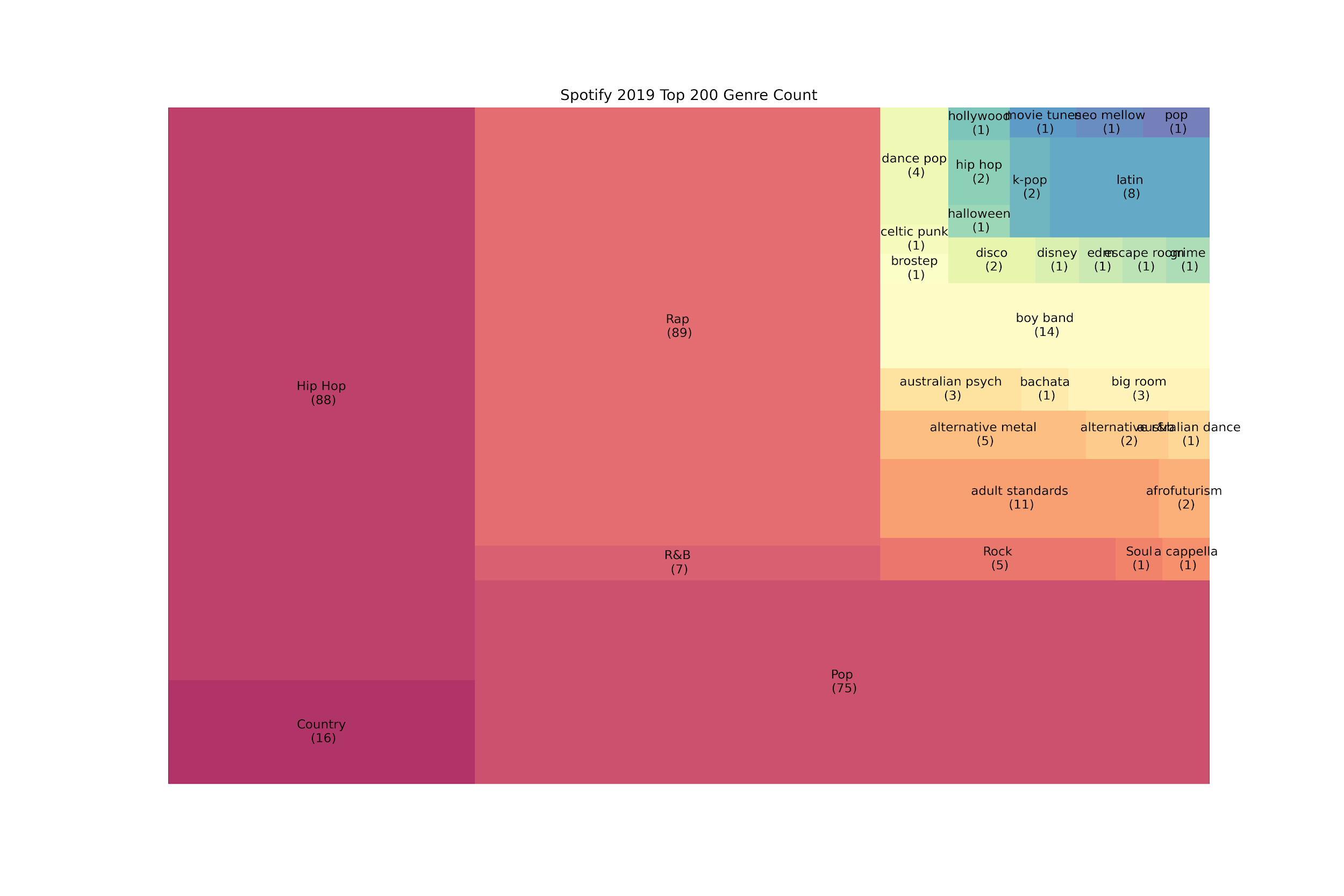
This analysis concludes that there were no major changes to listening trends between 2019 and 2020. The number of streams from 2019 to 2020 declined by 6.5%. This could be due to lockdowns where historically music has been listened to commuting to work or working out at gyms.



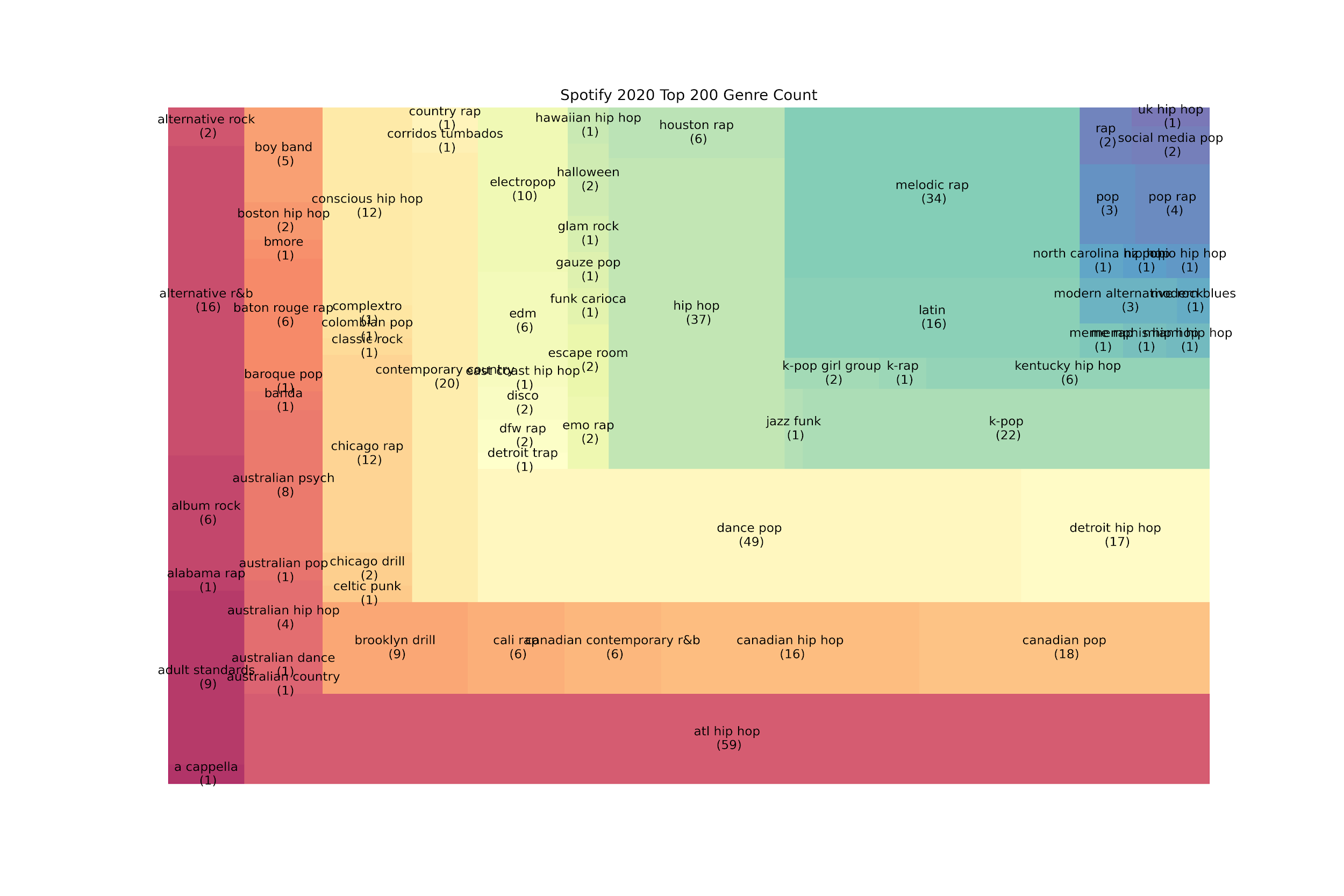
There was a deviation between the months of April and May that was possibly the result of aggressive national lockdowns, however the team was unable to find a direct correlation between lockdowns and the number of streams. Based on the data it could be argued that the streaming behavior was very similar year over year as the months that had spikes and declines in streaming volume were very similar YoY

**Unique Genres and Tracks**

Although 2019 had more total streams 2020 had more artists, tracks, and genres appear on the top 200. Using the Spotify API calls we were able to see what genre every artist and track is classified as. These plots show the difference between years.



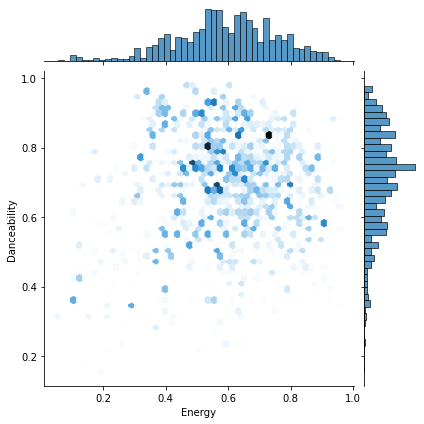
In 2019 there were 33 types of genres and 2020 more than doubled reaching 70. These treemaps keep track of every time a new song reaches the Top 200 and the total number of tracks associated with the genre.

****

This data shows that people were exploring new bands and new kinds of music. If we were able to look beyond the Top 200 I believe it would reflect this with a greater value.

**Audio Analysis in the United States**

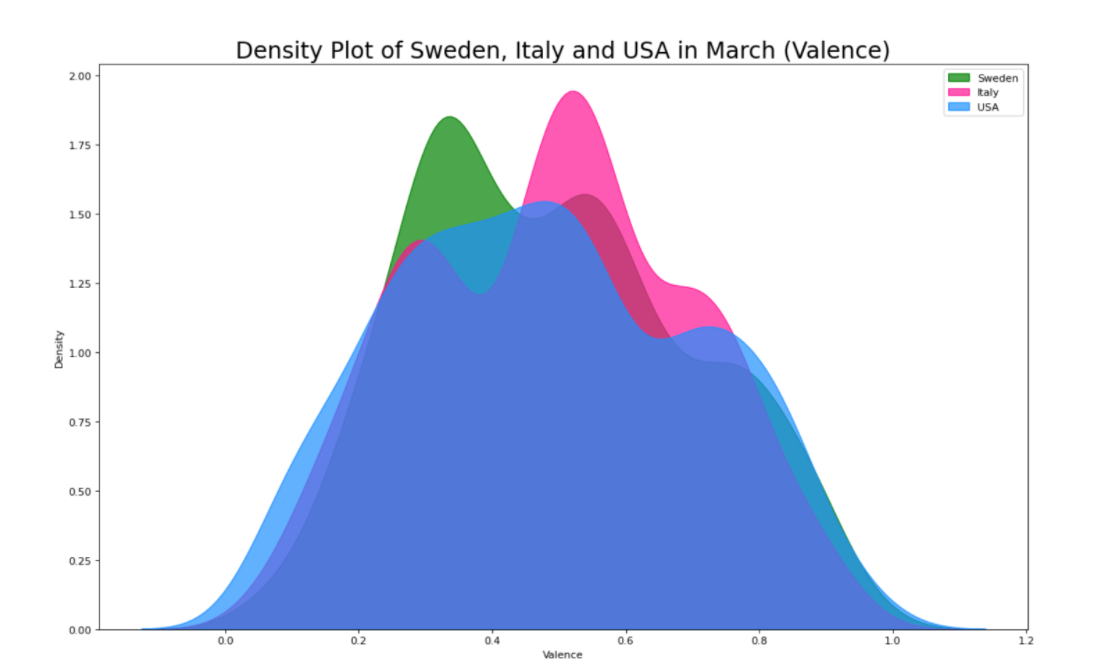
This analysis also suggests that there is a statistically significant association between danceability and energy. The majority of songs have energy between 0.6-0.8 on the index while most songs have a danceability of 0.6. Those figures remained relatively unchanged between 2019 and 2020.



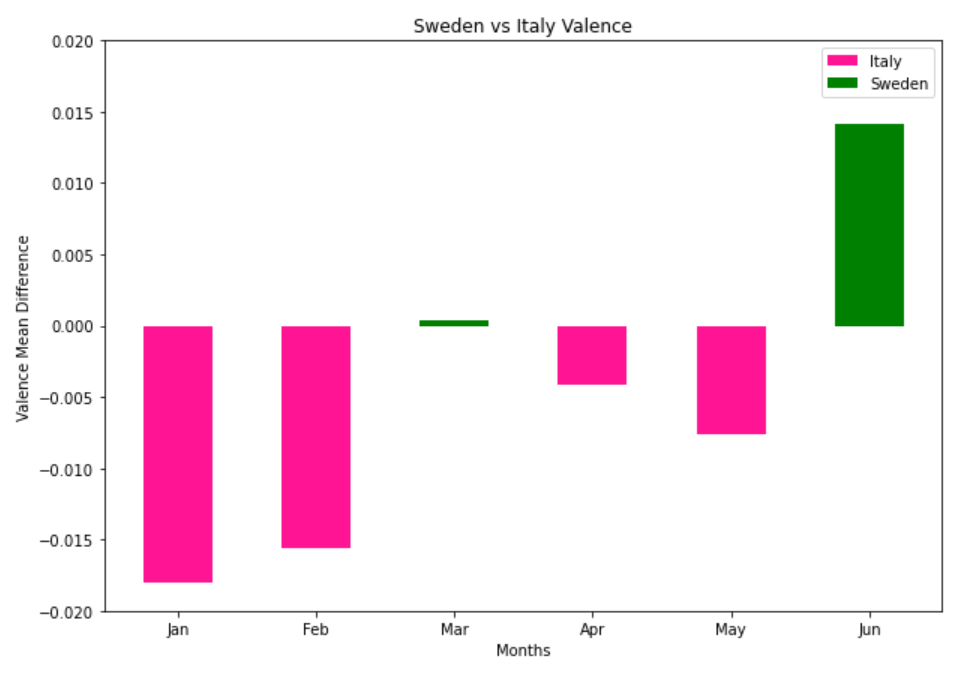
**Audio Analysis in Sweden, Italy and the United States**

****

These radar plots show the distribution of Spotify audio features in which you can notice that Valence, Energy and Danceability were very similar for all countries, except for Italy that has a higher energy value.

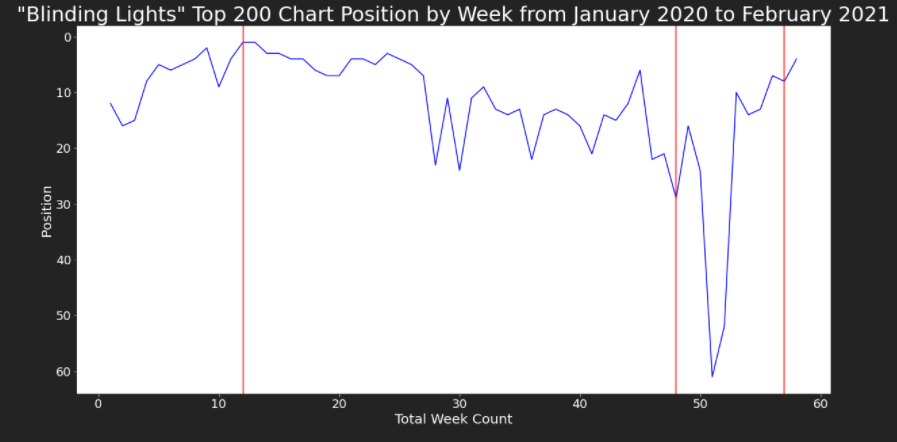


This density plot you can notice a more detailed separation in Valence values which measure music positiveness. The month of March 2020 shows that Italy was listening to more positive music than Sweden and the United States as highlighted by the pink peak.

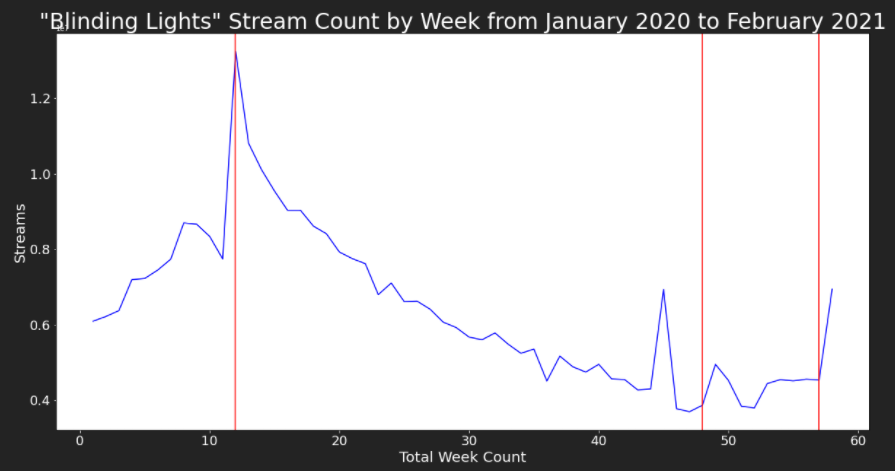


This bar plot was created using the average of Valence values for Italy and Sweden for the first six months of 2020. By using the difference of both values we were able to highlight which country listened to the most positive music per month. As the graph suggests Italy was leading in the months of January and February while Sweden with a small difference took over the month of march. The reason why these countries were chosen is due to how different they decided to deal with the pandemic; Italy decided to go into a lockdown while Sweden did not.

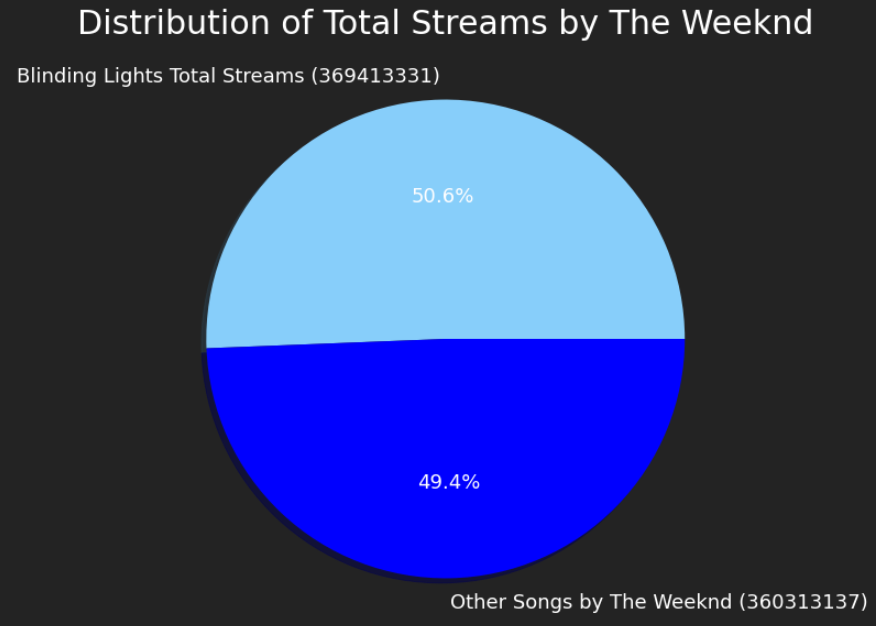
**A Closer Look Into The Top 200 Charts**

****

This plot tracks the Top 200 Chart position for the song “Blinding Lights” by The Weeknd. From this data it was determined that the pandemic had little to no effect on the popularity of the song and instead the popularity or ranking was influenced by external factors including its social media relevance and the artist’s public appearances.



As with the above plot, this line graph tracks The Weeknd’s hit song over time, but with a focus on total stream count by week. The weekly stream count reaches its highest position at the official release date of his album, After Hours. The stream count trends downward and seems to spike surrounding his live performances later in the year and in early 2021.



“Blinding Lights” accounted for 50.6% of the total streams for all of the Weeknd’s music that appeared on the top 200 chart in the above mentioned time period.

**References**

1. Spotify Weekly Top 200 Charts:   
   <https://spotifycharts.com/regional/us/weekly/latest>
2. Spotify Developer API:

<https://developer.spotify.com/documentation/web-api/>

**Glossary**

**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.

**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.

**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).

**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 strong likelihood that the track is live.