# Bops by Country - Explained

*Gahyun (Grace) Kim*

*12/07/2019*

## Introduction

In an era where the world feels smaller than ever before, Spotify allows artists to connect with their global fans in an even easier manner. Spotify is one of the most popular global streaming platforms that provides music, videos, and podcasts. Initially launched in 2008 for service in five European countries, it now has 248 million active users and is available in 79 countries [1]. Spotify has top song charts for each country their users are in, and displays the top 5 cities where an artist's listeners are located.

Some bands have noticed and utilized their popularity overseas. Gorillaz are a virtual band, active since 1998, whose music encompasses hip hop, rock, and electronic genres. Despite the band being from London, most of Gorillaz's regular monthly listeners on Spotify live in Mexico City (424,983 listeners on December 6th, 2019). Considering how the second city with the most listeners has less than half of that number (174,633 listeners on December 6th, 2019), Gorillaz's popularity in Mexico is notable. In response to the incredible feedback from Mexican listeners, they finished up their last tour in Mexico City instead of their hometown as most bands do [2].

For fledgling artists seeking opportunities, targeting markets overseas may seem appealing. The differences in what global listeners prefer pose the following questions: why are some artists more popular in certain countries? Do popular songs in different countries share certain traits?

This report attempts to answer those questions by comparing the differences between the top song charts for the United States and Taiwan. Taiwan was chosen because it was one of the first few markets that Spotify expanded to in Asia. Furthermore, it was assumed that more non-American songs would be on their charts; English is not commonly used in Taiwan, compared to Spotify's other early Asian markets such as Singapore or the Philippines. Based on the data collected, a multilevel regression model was fit for both USA and Taiwan to predict what kind of song characteristics would lead to more streams in each country.

## What are the most streamed songs and artists in USA and Taiwan?

The charts collected for this report were the USA and Taiwan Weekly Top 200 Spotify charts for five different weeks: the first week in January of 2017, the first week in July of 2017, the first week in January of 2018, the first week in July of 2018, and the first week in January of 2019.

On first glance, the Spotify charts for USA and Taiwan looked different.

The top songs on USA's charts have stream counts that are seven digits, with the first song racking up more than 9 million streams. The top artists are mostly rap and hip-hop artists, such as Migos and Drake.

Table 1: Top 10 Songs in USA and Their Characteristics (1st Week of Jan. 2017)

| Rank | Track Name | Artist | Streams | Danceability | Speechiness | Acousticness | Valence |
|---|---|---|---|---|---|---|---|
| 1 | Bad and Boujee (feat. Lil Uzi Vert) | Migos | 9212016 | 0.927 | 0.2440 | 0.0610 | 0.1750 |
| 2 | Fake Love | Drake | 8402709 | 0.928 | 0.2870 | 0.1050 | 0.6130 |
| 3 | Starboy | The Weeknd | 7634586 | 0.681 | 0.2820 | 0.1650 | 0.5350 |
| 4 | Closer | The Chainsmokers | 6384465 | 0.748 | 0.0338 | 0.4140 | 0.6610 |
| 5 | Black Beatles | Rae Sremmurd | 5745384 | 0.794 | 0.0649 | 0.1420 | 0.3550 |
| 6 | Caroline | Aminé | 4992851 | 0.945 | 0.4920 | 0.3030 | 0.7430 |
| 7 | I Don't Wanna Live Forever (Fifty Shades Darker) - From "Fifty Shades Darker (Original Motion Picture Soundtrack)" | ZAYN | 4874178 | 0.735 | 0.0585 | 0.0631 | 0.0862 |
| 8 | Bounce Back | Big Sean | 4818172 | 0.780 | 0.1410 | 0.1040 | 0.2730 |
| 9 | Bad Things (with Camila Cabello) | Machine Gun Kelly | 4805096 | 0.697 | 0.1460 | 0.2140 | 0.3050 |
| 10 | I Feel It Coming | The Weeknd | 4586635 | 0.773 | 0.1180 | 0.4260 | 0.5850 |

Meanwhile, the stream counts in Taiwan's charts are mostly five digits and the top few songs barely surpass 100 thousand streams. Some of the top artists are the same as USA's, such as the Chainsmokers and the

Weeknd, but rap artists are nowhere to be seen in the top 10. Instead, there are local artists such as NICKTHEREAL.

Table 2: Top 10 Songs in Taiwan and Their Characteristics (1st Week of Jan. 2017)

| Rank | Track Name | Artist | Streams | Danceability | Speechiness | Acousticness | Valence |
|---|---|---|---|---|---|---|---|
| 1 | Closer | The Chainsmokers | 151205 | 0.748 | 0.0338 | 0.4140 | 0.6610 |
| 2 | I Don't Wanna Live Forever (Fifty Shades Darker) - From "Fifty Shades Darker (Original Motion Picture Soundtrack)" | ZAYN | 118950 | 0.735 | 0.0585 | 0.0631 | 0.0862 |
| 3 | Starboy | The Weeknd | 118763 | 0.681 | 0.2820 | 0.1650 | 0.5350 |
| 4 |  | NICKTHEREAL | 110299 | 0.579 | 0.0446 | 0.0309 | 0.3730 |
| 5 | Let Me Love You | DJ Snake | 107646 | 0.476 | 0.0576 | 0.0784 | 0.1420 |
| 6 | Don't Wanna Know | Maroon 5 | 102903 | 0.783 | 0.0800 | 0.3380 | 0.4470 |
| 7 | Say You Won't Let Go | James Arthur | 102048 | 0.358 | 0.0590 | 0.6950 | 0.4940 |
| 8 | All We Know | The Chainsmokers | 98847 | 0.662 | 0.0307 | 0.0970 | 0.2960 |
| 9 | Alone | Alan Walker | 96448 | 0.676 | 0.0458 | 0.1860 | 0.1570 |
| 10 | Rockabye (feat. Sean Paul & Anne-Marie) | Clean Bandit | 93461 | 0.720 | 0.0523 | 0.4060 | 0.7420 |

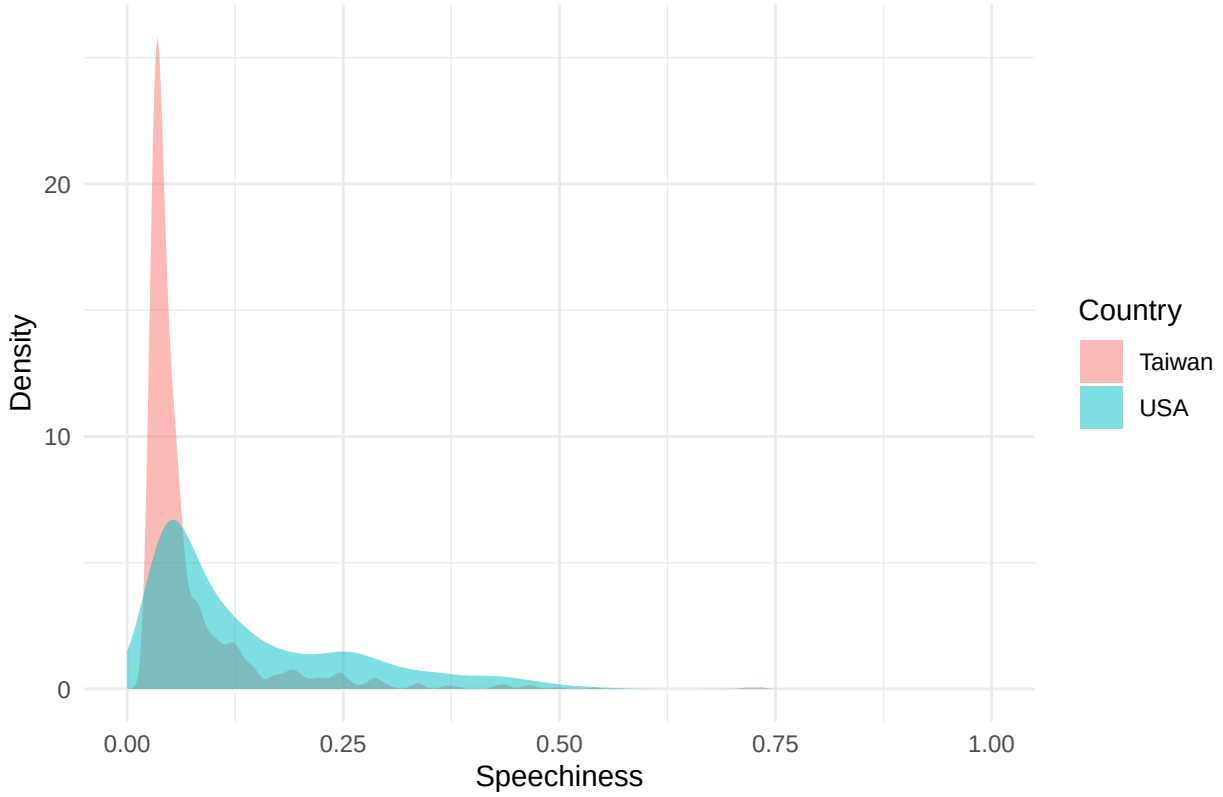
Taiwan's total stream counts are much lower than those of USA's because of two main reasons: Spotify is still relatively new in the Asian market and offers less Asian songs than local music streaming apps. Spotify originates from Sweden and expanded its service to USA in 2011 [3]. It did not begin service in Asia, including Taiwan, until 2013 [4]. Furthermore, Spotify offers fewer Asian songs compared to local music apps such as KKBOX (founded in 2004 in Taiwan) [5].

## How are the top songs in USA and Taiwan different or similar?

Spotify has numeric variables that represent different characteristics for every song they offer for streaming. The characteristics examined here are speechiness, acousticness, and danceability. (Refer to Appendix A for other track characteristics). Figures 1 to 3 are density plots, which reflect the distribution of data across a numeric variable. The data shown in the following plots is for all five of the charts collected.

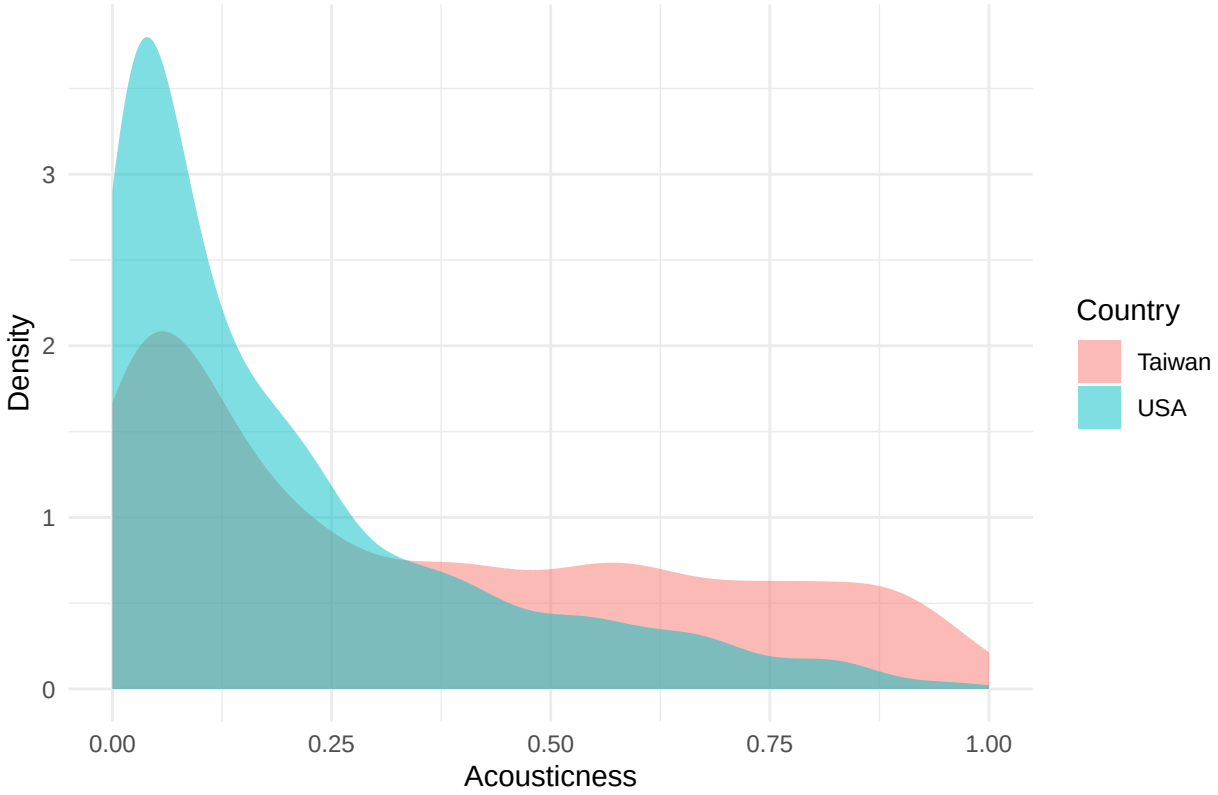
Speechiness measures the amount of spoken word in a song. The more exclusive speech-like material that a track has (i.e. audio book, poetry), the closer the metric is to 1.0. Rap music usually has speechiness values between 0.33 and 0.66 [6]. Figure 1 represents the distribution of all the songs in the charts collected for each country at all five time periods, across speechiness. As we noticed how there were fewer rap artists in Taiwan's top 10 songs compared to USA's, there are many more songs that have lower speechiness values in Taiwan's charts. Meanwhile, there is a relatively more even distribution of songs across the speechiness spectrum in USA's charts.

Figure 1: Distribution of Speechiness for USA and Taiwan's Top Songs



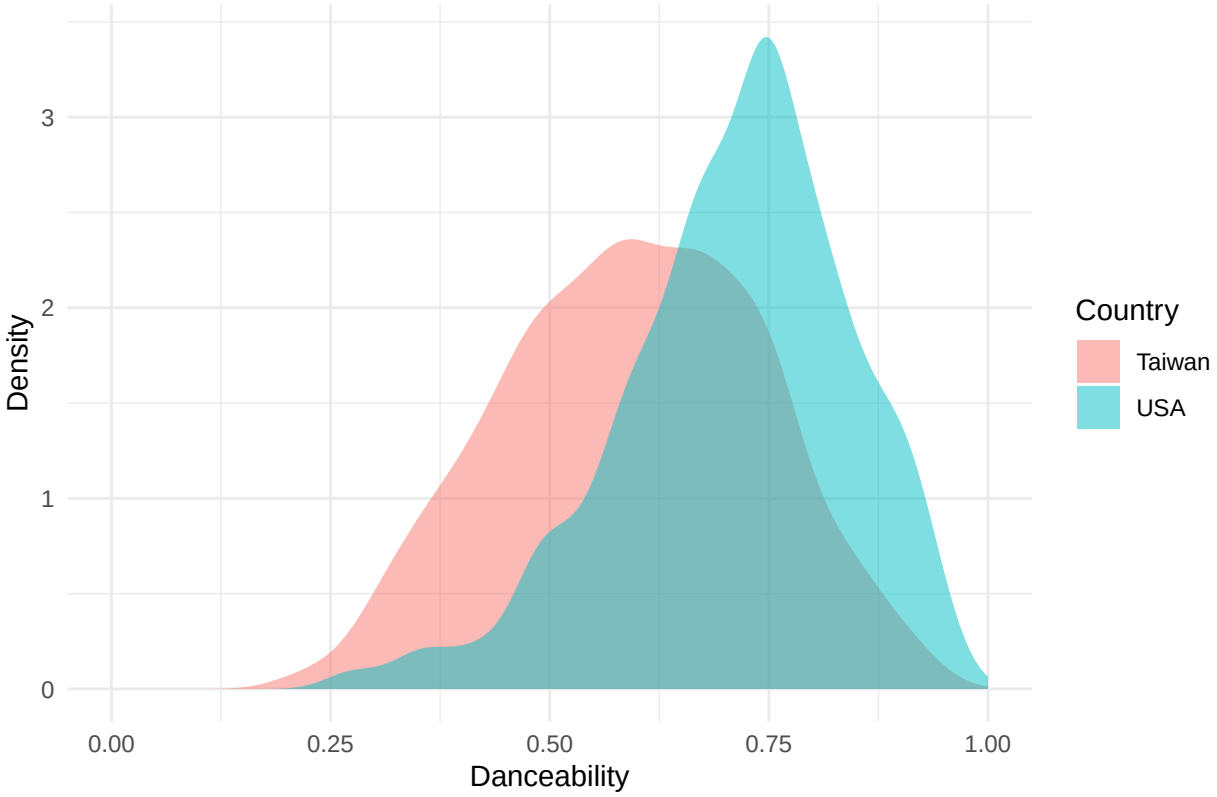

Acousticness is a measure that reflects how acoustic a song is. If Spotify has a higher confidence that the track is acoustic, then the track's acousticness is closer to 1.0 [6]. Most of the songs in both countries' top songs have lower acousticness, but Taiwan's top songs are distributed more evenly across the scale compared to USA's top songs; in Figure 2, the peak for Taiwan's songs is not as pronounced as that of USA's. Moreover, the median acousticness value for USA's top songs was 0.114 while it was 0.242 for Taiwan's top songs. The visualization of the data distribution and the median values imply that Taiwan has more top songs that are acoustic than USA does, for the charts collected.

Figure 2: Distribution of Acousticness for USA and Taiwan's Top Songs




Danceability indicates how appropriate a song is for dancing. The metric is based on a combination of tempo, rhythmic stability, beat strength, and overall regularity. The better a song is for dancing, the closer it is to 1.0 [6]. In Figure 3, USA's data has a higher and narrower peak around 0.75 while Taiwan's data has a rounder peak at a lower number. The median danceability value for USA's songs was 0.732, and the value for Taiwan was 0.599. Based on this density plot and the median values, we can observe that America's charts generally have more songs that are easier to dance to compared to Taiwan's charts.

Figure 3: Distribution of Danceability for USA and Taiwan's Top Songs




Although both countries had differences in the distribution of song characteristics for their top songs, the two countries' charts had one similarity; when all five charts were pooled together for each country, the number of artists were only a small fraction of the number of songs. For USA, there were only 269 artists among the 1000 songs compiled from five different weeks. For Taiwan, there were 323 artists in their 1000 songs. This introduces the possibility that the song's artist may affect the song being on the charts.

## Can we predict what kind of song would chart higher in USA and Taiwan?

Based on the song characteristics that seemed to differ noticeably for each country (speechiness, acousticness, danceability), a multilevel model was fitted to predict what kind of song would chart higher in USA and Taiwan.

Assuming that

- $i = 1, ..., 1000$ songs (from each country's set of charts)
- $j = 1, ...,$ n artists (269 for USA, 323 for Taiwan)
- $y_i =$ log of number of streams on Spotify for $i$th song
- $x =$ individual level predictors, such as speechiness, acousticness, and danceability

we can formulate a multilevel model for each country's data with intercepts varying by artist and song-level predictors:

$$y_i = \alpha_{j[i]} + \beta_{j[i]} x_i + e_{ij}$$

Based on the coefficients of the predictors in the models, it was observed that speechiness had a greater effect on the number of streams for USA's model (coefficient = 0.20430) than in Taiwan's model (coefficient = 0.1649). However, the coefficients of acousticness and danceability were negative for Taiwan's model (-0.3721 and -0.2254, respectively); this indicates that on average, an increase in acousticness and danceability leads to less streams on Taiwan's chart.

The following plots are residual plots and QQ plots for USA's multilevel model (Figures 4, 5) and Taiwan's model (Figures 6, 7). Ideally, the residual plots should exhibit an even spread around the centered line. However, there is a small collection of data points on the far right side of the x-axis for both plots (Figures 4, 6). For the QQ plots, the points should follow a linear pattern along the black line, but both plots had minor tail issues.

**Figure 4: Residual Plot for USA's Model**

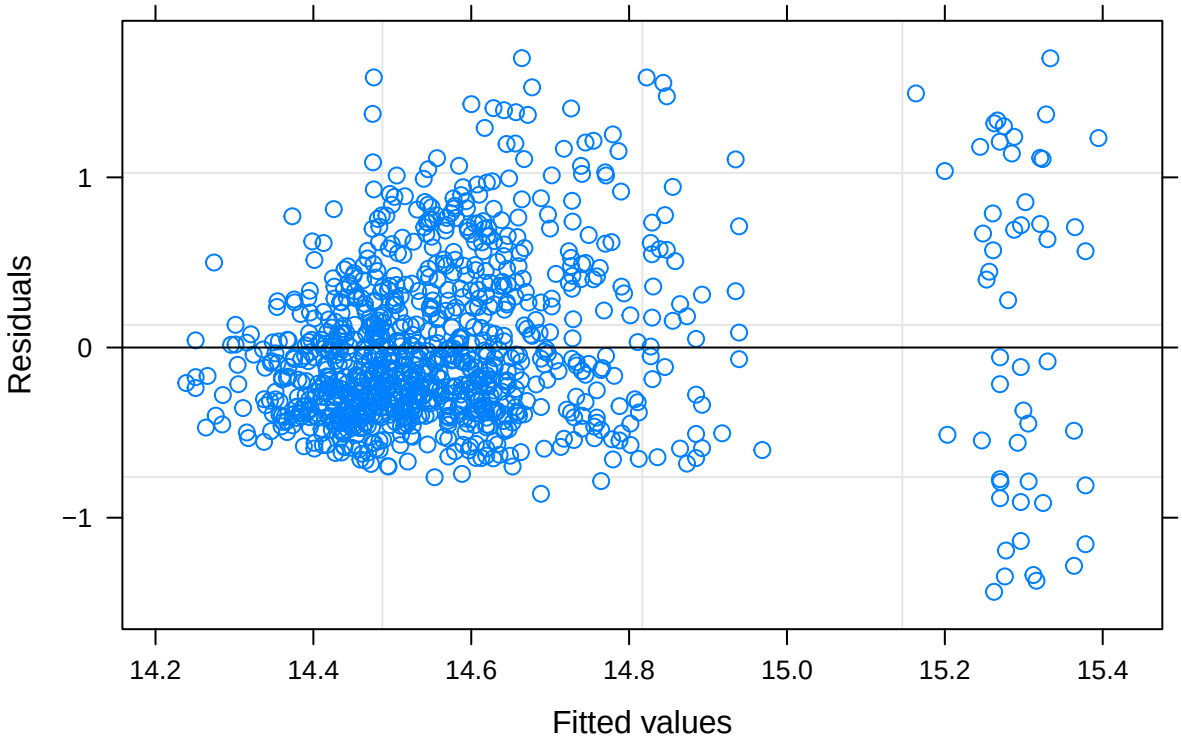


**Figure 5: Residual Plot for Taiwan's Model**

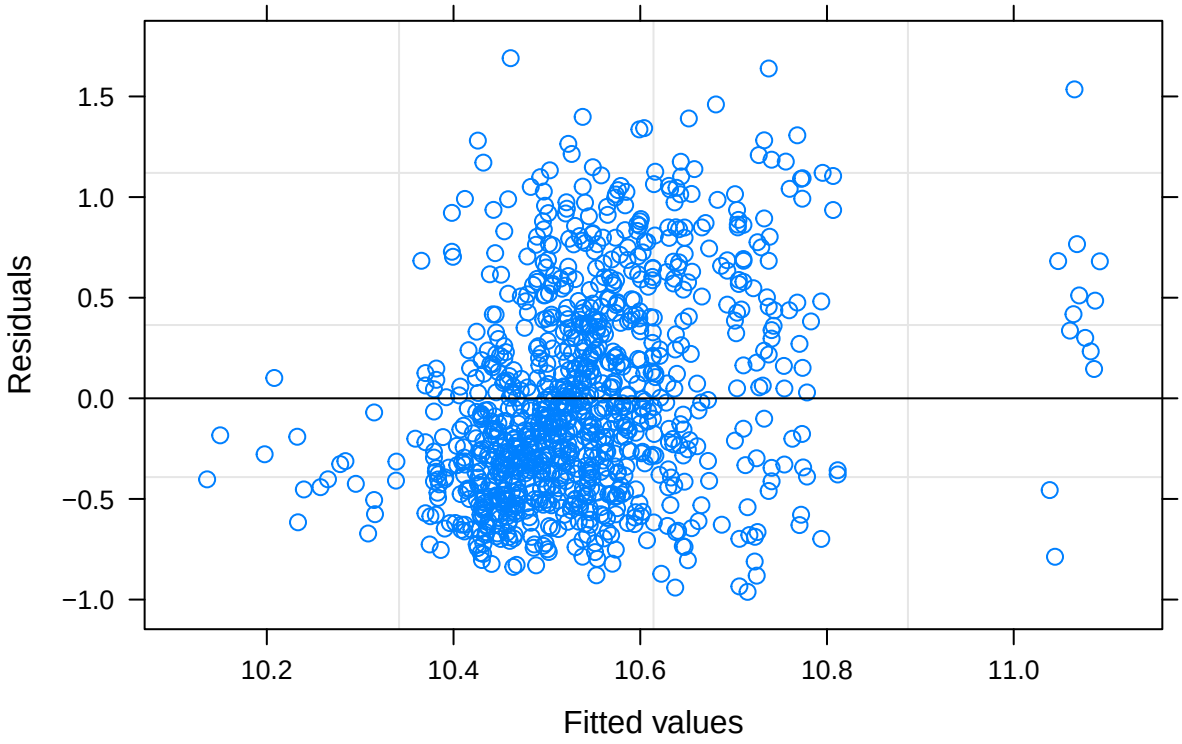



**Figure 6: QQ Plot for USA's Model**

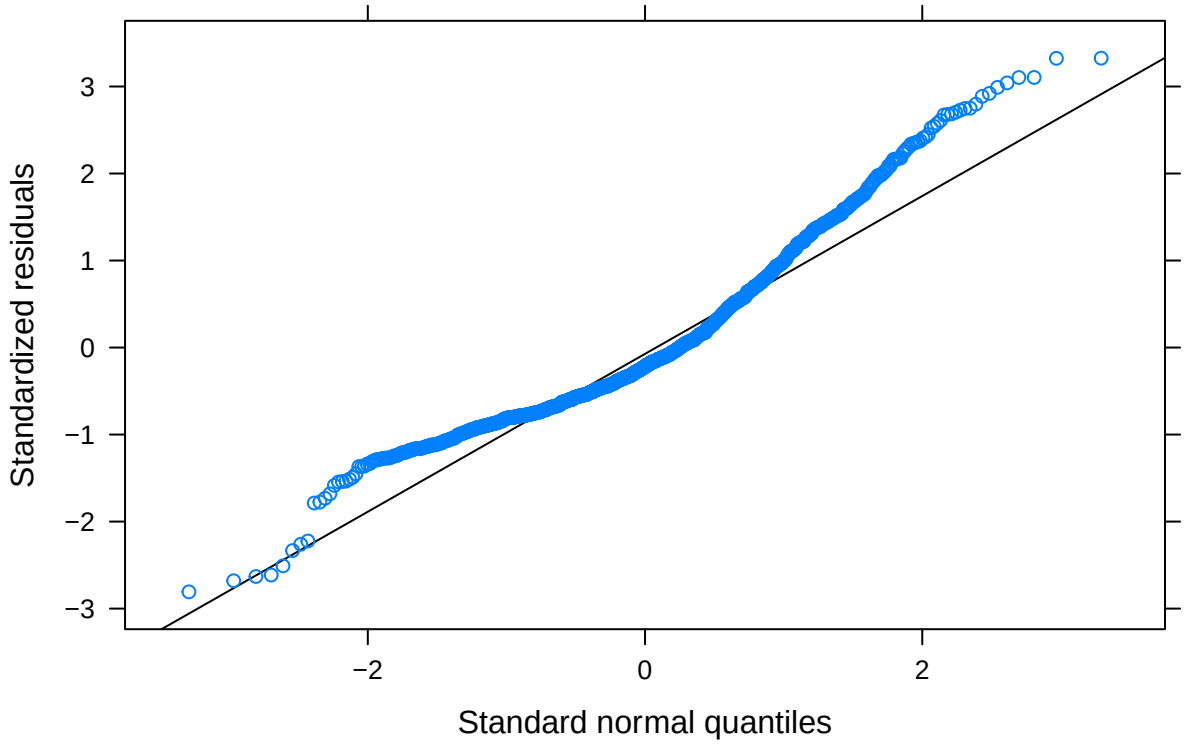
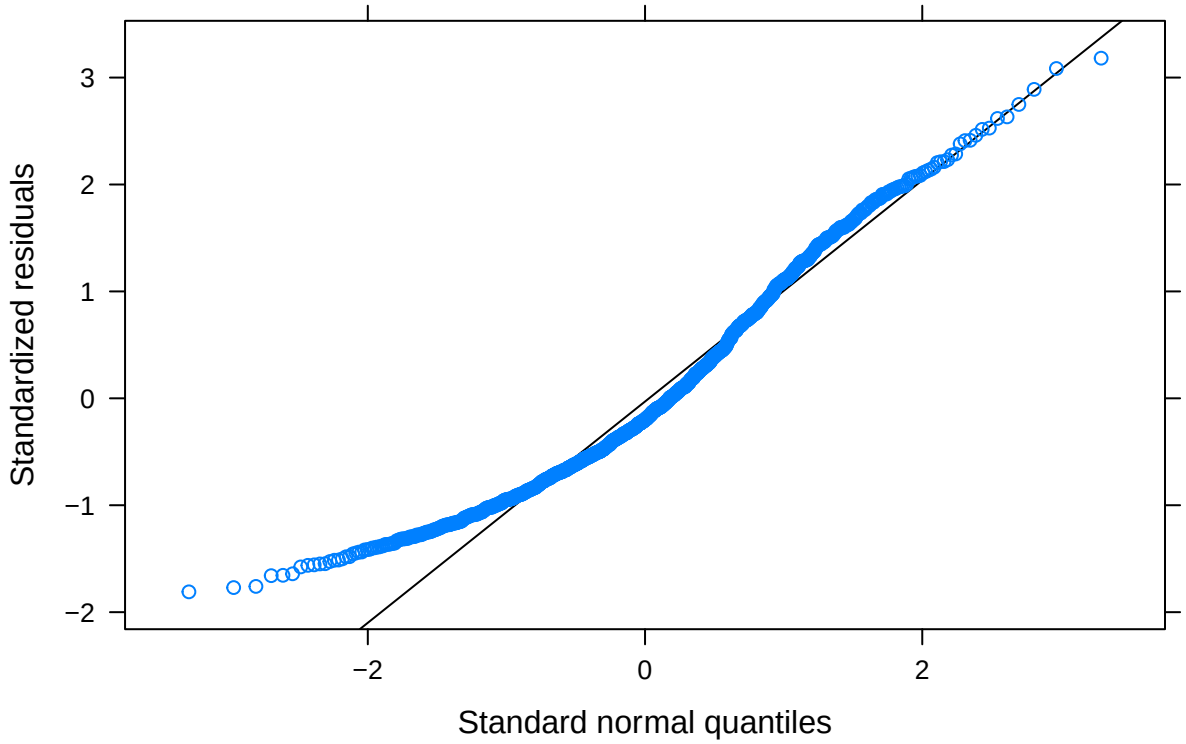


**Figure 7: QQ Plot for Taiwan's Model**




Refer to Appendix B for more information on the models.

### Limitations and Possible Areas of Improvement

One element that could improve the quality of the model is incorporating the release date of a song. For the weekly charts, the number of streams for each song are measured only for that specific week. Therefore, songs released closer to a given week have a higher chance of charting than other songs, especially if they are released by a popular artist. For example, Drake's album 'Scorpion' was released on June 29, 2018. Due to the album's release date in conjunction with Drake's popularity, all 25 songs from the album were within the top 35 songs for the first weekly chart of July 2018. If song release dates were added to the data and model, it would assist in making better predictions for how high songs can chart.

## Appendix A: Figures for Other Track Characteristics

Figure 9: Distribution of Valence for USA and Taiwan's Top Songs

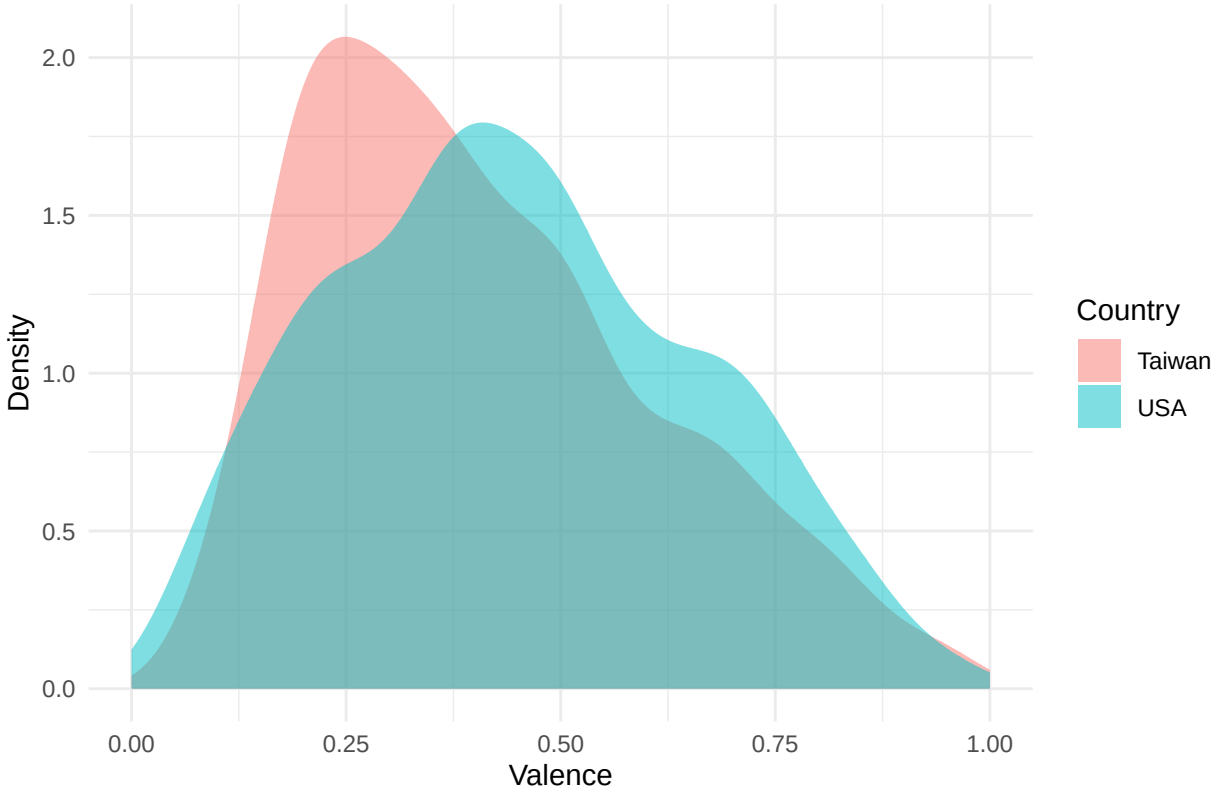


Figure 10: Distribution of Tempo for USA and Taiwan's Top Songs

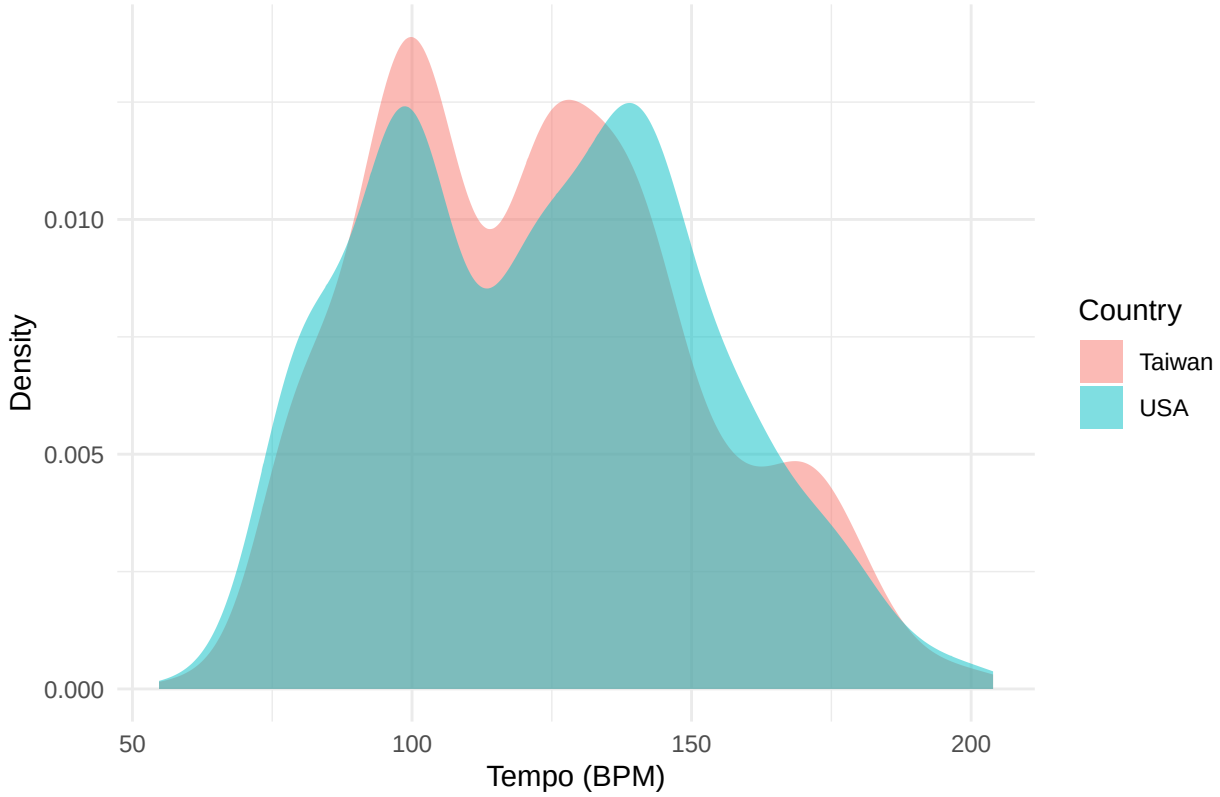



Figure 11: Distribution of Song Duration for USA and Taiwan's Top Songs

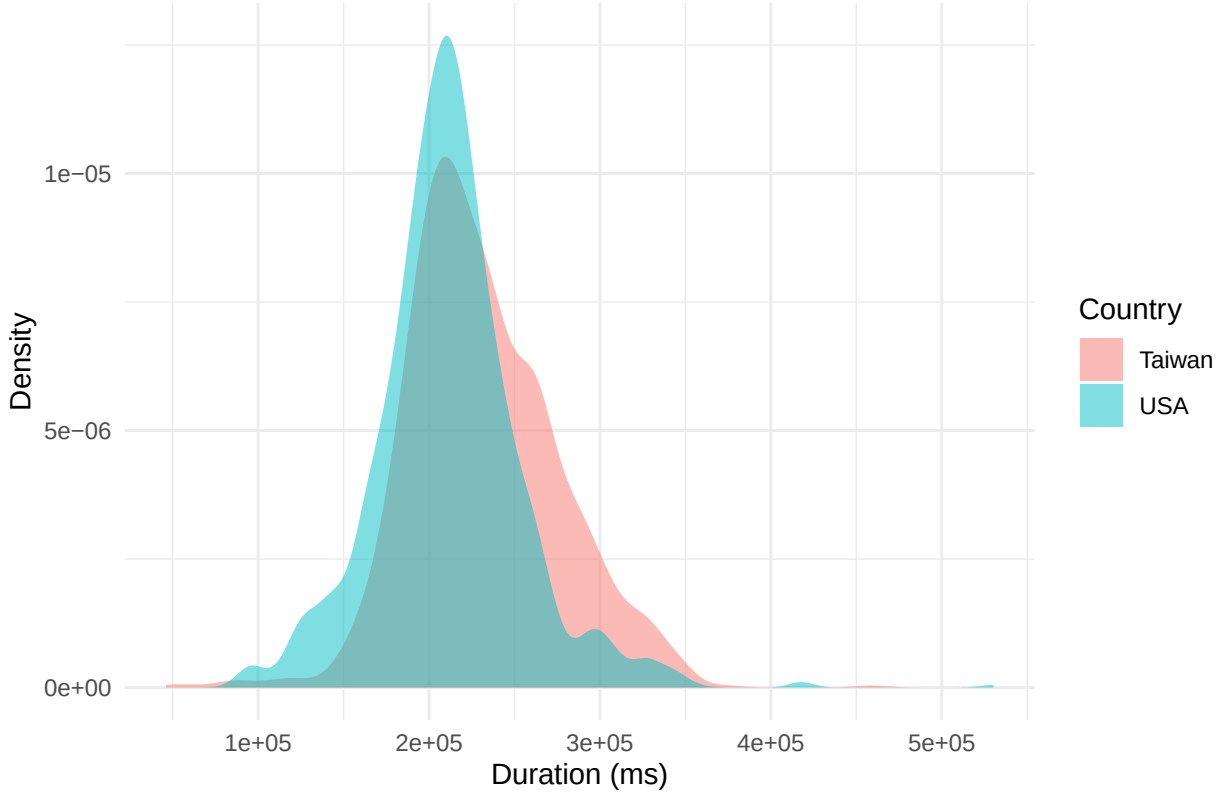
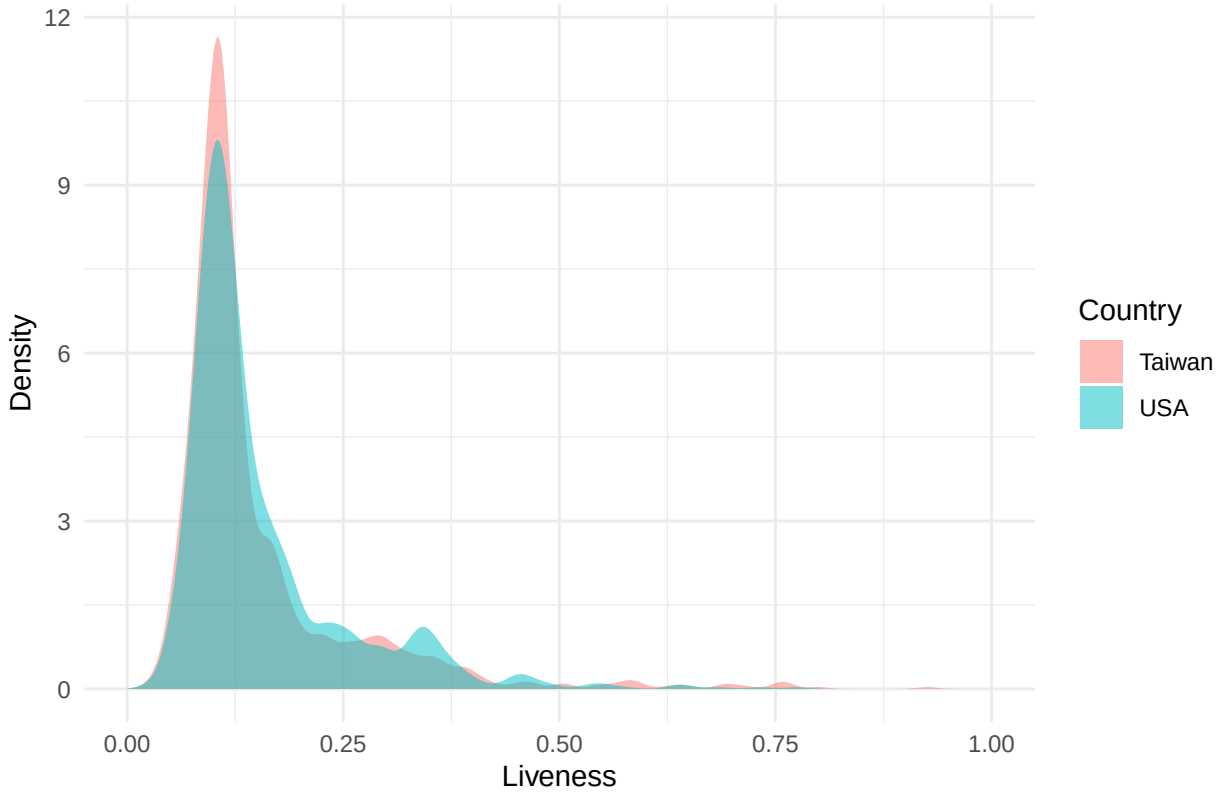


Figure 12: Distribution of Liveness for USA and Taiwan's Top Songs




## Appendix B: R Outputs for Models

```
summary(fit_usa)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log(Streams) ~ speechiness + acousticness * danceability + (1 |
##     Artist)
##    Data: usa
##
## REML criterion at convergence: 1588
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8075 -0.6845 -0.2210  0.5391  3.3255
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artist   (Intercept) 0.03141  0.1772
##  Residual             0.26147  0.5113
## Number of obs: 1000, groups:  Artist, 269
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                 14.20858    0.13243 107.294
## speechiness                  0.20430    0.15500   1.318
## acousticness                 0.02374    0.40846   0.058
## danceability                 0.37016    0.18410   2.011
```

```
## acousticness:danceability  0.08197    0.59562   0.138
##
## Correlation of Fixed Effects:
##             (Intr) spchns acstcn dncblt
## speechiness -0.051
## acousticnss -0.654  0.011
## danceabilty -0.965 -0.111  0.650
## acstcnss:dn  0.622 -0.008 -0.977 -0.645
```

```
summary(fit_taiwan)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log(Streams) ~ speechiness + acousticness * danceability + (1 |
##     Artist)
##    Data: taiwan
##
## REML criterion at convergence: 1666
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8099 -0.7277 -0.2003  0.6682  3.1814
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artist   (Intercept) 0.03261  0.1806
##  Residual             0.28226  0.5313
## Number of obs: 1000, groups:  Artist, 323
##
## Fixed effects:
##                            Estimate Std. Error t value
## (Intercept)                 10.6451     0.1347  79.003
## speechiness                  0.1649     0.2563   0.644
## acousticness                -0.3721     0.2560  -1.453
## danceability                -0.2254     0.2078  -1.085
## acousticness:danceability    0.6190     0.4490   1.379
##
## Correlation of Fixed Effects:
##             (Intr) spchns acstcn dncblt
## speechiness -0.113
## acousticnss -0.770  0.058
## danceabilty -0.960 -0.050  0.770
## acstcnss:dn  0.691 -0.031 -0.962 -0.739
```

## References


Cited sources:

1. https://newsroom.spotify.com/company-info/
2. https://newsroom.spotify.com/2018-11-19/mexico-city-is-now-the-worlds-music-streaming-mecca/
3. http://edition.cnn.com/2011/TECH/web/07/13/spotify.us/
4. https://thenextweb.com/media/2013/09/24/spotify-quietly-launches-in-taiwan-its-fourth-country-in-asia/
5. https://www.fastcompany.com/1576788/streaming-music-service-kkbox-chinese-spotify
6. https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

Some methods and code were referenced from the following sources:

1. https://www.rcharlie.com/spotifyr/
2. https://msmith7161.github.io/what-is-speechiness/