MA 678 Final Project

Jiahao Liu
Master of Statistical Practice, Boston University

Content

Abstract

Spotify, one of the world's largest music providers, offers millions of tracks to people every day. Music is present in all parts of people's lives, so it got me thinking what types of music are most popular? What factors influence people's attitudes towards music? This report is a multilevel model analysis based on Spotify's data, and I use key, time signature, and audio mode as the group level. The result indicates that the popularity of a song depends on different factors.

Introduction

Music has long been an integral part of human life. And everyone has their own unique way of enjoying music and their unique type of music. Just as upbeat music can make people feel excited, soft music with a smooth melody can calm the mind. Sometimes, people also choose songs based on their favourite artists. From Whitney Houston's perfect soprano to the pop music performed by Rihanna and Taylor Swift today, or the Korean boy bands that are popular in Asia, music always brings people different surprises. Some psychotherapists choose music therapy to soothe their patients, as research findings have shown a strong connection between music and human mental health. For instance, the specific tones and rhythms of some music can help calm patients. To better understand this connection, we can look at people's love of music and other factors, including the pitch and rhythm of the music, to find the answer. The popularity of a song is a direct indication of how people respond to music, so I decided to use this as a dependent variable in a multilevel model. Also the model should contain random effects (song duration, tempo, loudness, etc), and fixed effects (key, audio mode, and time signature).

Method

The data is from Kaggle website (https://www.kaggle.com/datasets/yasserh/song-popularity-dataset). The origin dataset has 18836 rows and 15 columns, which is collected from spotify using their API. After cleaning and skimming the data, I found that it contained a lot of repetitive song

titles, so I filtered it to ensure the accuracy of the subsequent analysis. The final dataset has 13070 rows and 15 columns. The full description of the dataset shows below.

---

song_name | name of the song

song_popularity | The popularity of the song, measured from 0 to 100

song_duration_ms | the duration of the song in milliseconds

acousticness | measuring whether a song is acoustic or not

danceability | determine whether the song is suitable for dancing

energy | perceptual measure of activity and intensity, measured from 0.0 to 1.0

instrumentalness | whether a song contains no vocals, measured from 0.0 to 1.0

key | pitch of the song, eg. 0 = C, 1 = C♯/D ♭ , 2 = D

liveness | determine if the song is a live version

loudness | the average loudness value of a song decibels

audio_mode | modality of the song, 0 = minor, 1 = major

speechiness | detects the presence of spoken words in the song

tempo | speed of the song, measured from average beat duration

time_signature | counting the number of beats in a bar in a song

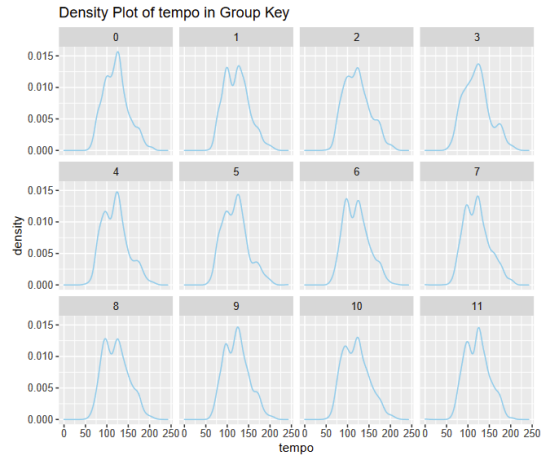
audio_valence | describes the musical positiveness conveyed by a song

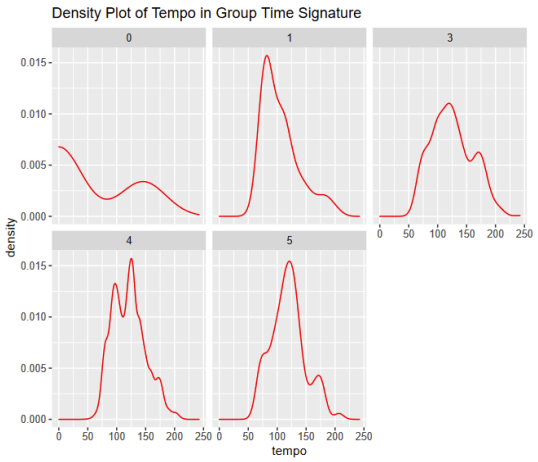
---

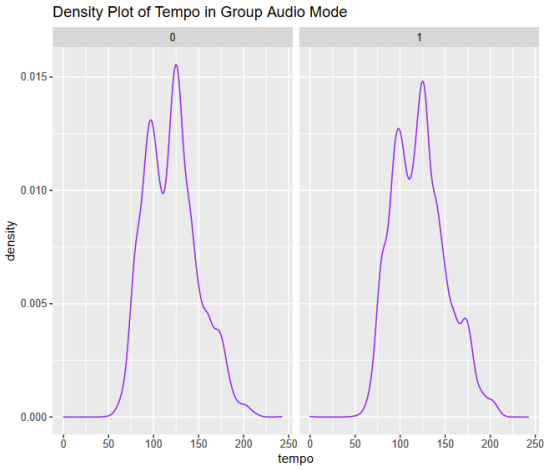
Then, I should check if the dataset contains missing data before choosing predictors. The result of the summary of the dataset shows below.

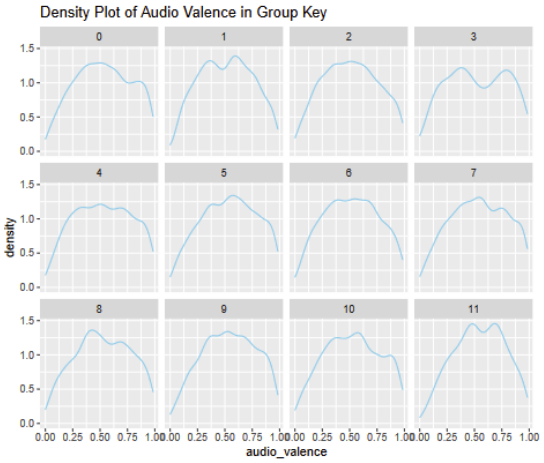
```
 song_name         song_popularity  song_duration_ms   acousticness       danceability
Length:13070       Min.   :  0.00   Min.   :  12000   Min.   :0.000001   Min.   :0.0000
Class :character   1st Qu.: 37.00   1st Qu.: 182967   1st Qu.:0.025000   1st Qu.:0.5242
Mode  :character   Median : 51.00   Median : 211486   Median :0.147000   Median :0.6370
                   Mean   : 48.49   Mean   : 218627   Mean   :0.277770   Mean   :0.6250
                   3rd Qu.: 63.00   3rd Qu.: 244506   3rd Qu.:0.479750   3rd Qu.:0.7410
                   Max.   :100.00   Max.   :1799346   Max.   :0.996000   Max.   :0.9870
    energy        instrumentalness        key           liveness         loudness
Min.   :0.00107   Min.   :0.000000   Min.   : 0.00   Min.   :0.0119   Min.   :-38.768
1st Qu.:0.49000   1st Qu.:0.000000   1st Qu.: 2.00   1st Qu.:0.0933   1st Qu.: -9.537
Median :0.66700   Median :0.000022   Median : 5.00   Median :0.1210   Median : -6.859
Mean   :0.63572   Mean   :0.096672   Mean   : 5.32   Mean   :0.1805   Mean   : -7.790
3rd Qu.:0.81500   3rd Qu.:0.005910   3rd Qu.: 8.00   3rd Qu.:0.2230   3rd Qu.: -5.041
Max.   :0.99900   Max.   :0.997000   Max.   :11.00   Max.   :0.9860   Max.   :  1.585
  audio_mode     speechiness         tempo         time_signature  audio_valence
Min.   :0.000   Min.   :0.0000   Min.   :  0.00   Min.   :0.000   Min.   :0.0000
1st Qu.:0.000   1st Qu.:0.0373   1st Qu.: 98.07   1st Qu.:4.000   1st Qu.:0.3340
Median :1.000   Median :0.0544   Median :120.03   Median :4.000   Median :0.5290
Mean   :0.633   Mean   :0.1008   Mean   :121.15   Mean   :3.952   Mean   :0.5286
3rd Qu.:1.000   3rd Qu.:0.1150   3rd Qu.:139.96   3rd Qu.:4.000   3rd Qu.:0.7300
Max.   :1.000   Max.   :0.9410   Max.   :242.32   Max.   :5.000   Max.   :0.9840
```

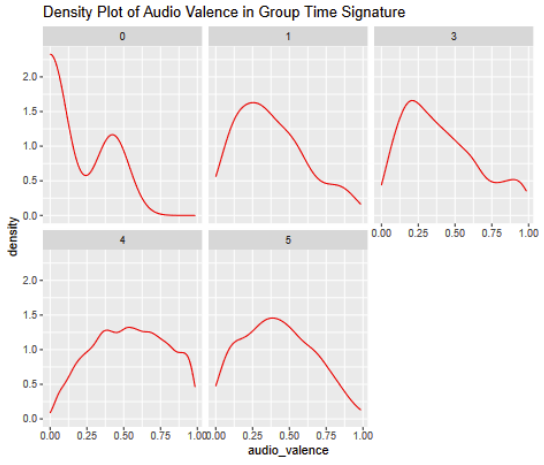
It shows that there is no missing data, and the current dataset has 12 variables and 13070 observations. I need to check the distribution of these variables based on 3 group levels.

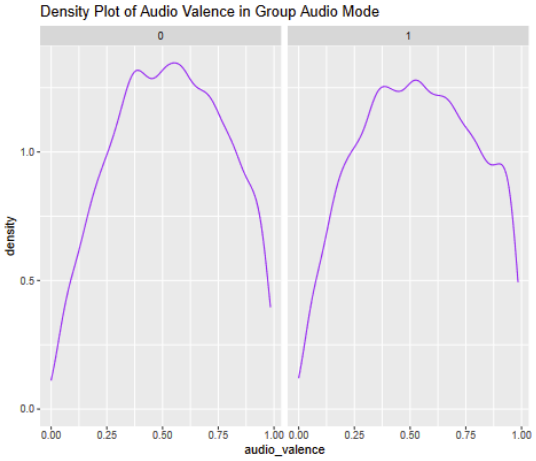












From these density plots, we can find that the relationship between group level and variables, for instance, each key group has a different tempo distribution, also the audio valence distribution in the time signature group exists variation. Other variables

also show such differences, so I put them in the appendix. Then I also do the following analysis to determine which factors should be contained in the model.

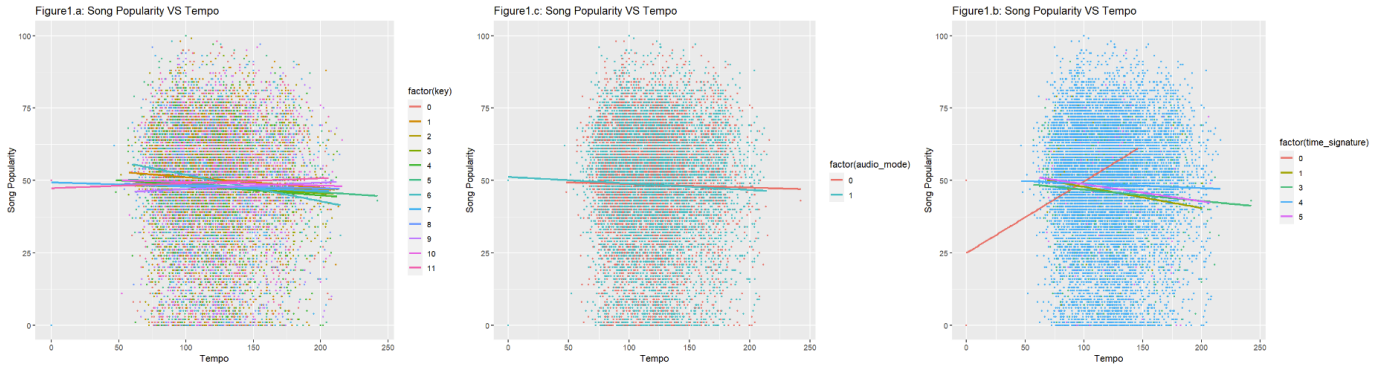


Figure1 indicates the relationship between song popularity and tempo, and figure1.a shows in group level key, figure2.b shows in group level audio mode, and figure1.c shows in group level time signature. We find that different tempo has different impact on song popularity, and the slope and intercept vary in different keys.

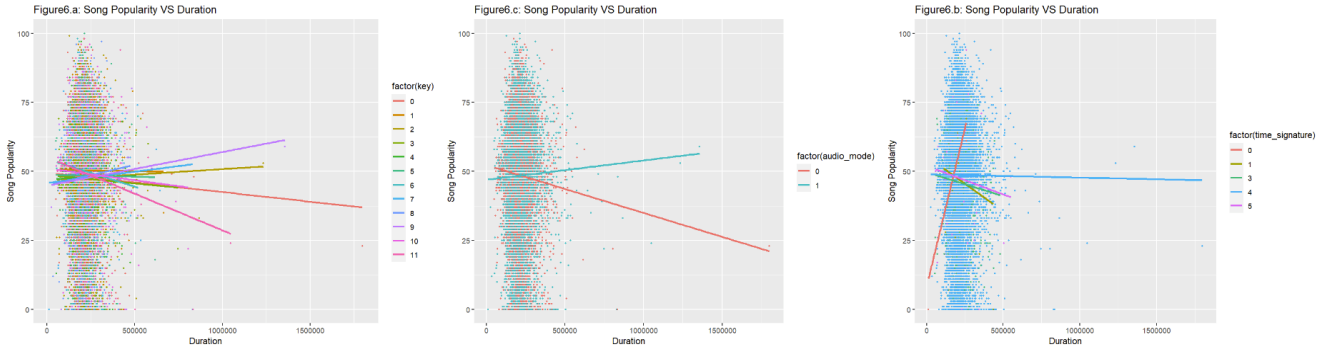
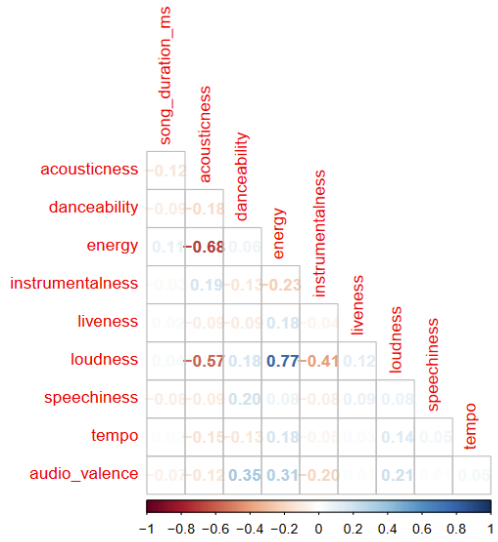


Figure6 indicates the relationship between song popularity and duration of the song, and a, b, c are in the group level key, audio mode, and time signature respectively. We find that the intercept and slope vary in all three plots, it is hard to define where to put this factor in the model, also duration is not a significant factor in our study of music, so I should drop it in the final model.
(I put other plots in the appdenix).

Then, I check the correlation between each pair of variables to get a more precise selection of variables .

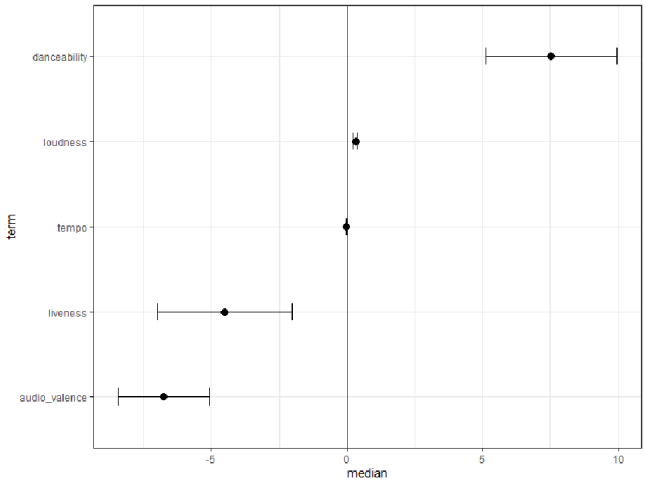


For the selection of variables, since loudness is the primary factor which indicates Psychological influences on the response to sound intensity, I include it in the model. Also the audio valence reflects the psychological impact of the song on people, so it is included. Sometimes, people prefer to experience the live version of the music, so liveness is also included. Morover, the table above demonstrates that energy and acousticness has high correlation (-0.68), also loudness and acoustiness has high correlation (-0.57), and loudness and energy has high correlation (0.77), so I will drop acousticness and energy. Additionally, the random effect of the key is important to variables: tempo and danceability. As a result, the model function shows below.

```
model <- lmer(song_popularity ~ tempo + danceability + audio_valence + livene
ss + loudness + (1 + tempo + danceability|key) + (1|time_signature) + (1|audi
o_mode), data = music_data)
```

I also check the ANOVA of this model to ensure all the variables are significant, and the summary of fixed effects shows below.

```
## Fixed effects:
##                  Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)    5.036e+01  1.592e+01  1.294e+04   3.164  0.00156 **
## tempo         -1.627e-02  6.278e-03  6.830e+02  -2.591  0.00977 **
## danceability   7.505e+00  1.299e+00  3.775e+01   5.778 1.17e-06 ***
## audio_valence -6.840e+00  7.664e-01  1.267e+04  -8.925  < 2e-16 ***
## liveness      -4.771e+00  1.220e+00  1.306e+04  -3.910 9.27e-05 ***
## loudness       3.043e-01  4.520e-02  1.286e+04   6.732 1.74e-11 ***
```



The summary of random effects shows below.

```
##    (Intercept)   tempo danceability
## 0      -0.4415  0.0014       0.1306
## 1       0.2288 -0.0007       0.8507
## 2      -0.8040  0.0025       0.2677
## 3       0.2578 -0.0008      -0.6062
## 4       0.1129 -0.0003      -0.4721
## 5       0.1656 -0.0005       0.0153
## 6       1.3714 -0.0042      -0.4203
## 7      -0.1107  0.0003      -0.5004
## 8      -0.3975  0.0012      -0.0023
## 9      -0.3423  0.0010      -0.3820
## 10      0.2703 -0.0008       0.2386
## 11     -0.3110  0.0010       0.8806
```

```
Random effects:
##  Groups         Name         Variance  Std.Dev.  Corr
##  key            (Intercept)  1.603e+00  1.265996
##                 tempo        1.507e-05  0.003882 -1.00
##                 danceability 1.869e+00  1.367054 -0.76  0.76
##  time_signature (Intercept)  2.530e+02 15.904768
##  audio_mode     (Intercept)  3.940e+02 19.850453
##  Residual                    3.988e+02 19.969965
## Number of obs: 13070, groups:  key, 12; time_signature, 5; audio_mode, 2
```

```
##   (Intercept)
## 0     -5.7595
## 1      1.4134
## 3      0.5875
## 4      2.2786
## 5      1.4799
```

```
##   (Intercept)
## 0     -0.2717
## 1      0.2717
```

Results

From the results in last part, the fixed effect model should be:

*song_popularity = 50.36 – 0.016*tempo + 7.505*danceability – 6.84*audio_valence – 4.771*liveness + 0.304*loudness*

For the model after adding random effect, I use key = 2 with 4 beats in a bar of a song in minor modality as an example:

*song_popularity = 51.5703 + 0.009*tempo + 7.7727*danceability – 6.84*audio_valence – 4.771*liveness + 0.304*loudness*

The interpretation of the coefficients in the models shows below, including both fixed effects and random effects.

Fixed Effects:

Intercept: the expected value of the popularity of a song is 50.36.

Tempo: the popularity of the song decreases as the speed of the music increases, for every unit increase in the average beat duration, song popularity decreases by -0.016.

Danceability: the songs suitable for dancing are more popular, and songs contribute to higher song popularity by 7.505.

Audio_valence: the song with higher valence sound (positive, happiness) decreases the song popularity by -6.84.

Liveness: the live version of the track reduces the popularity of the song by -4.771.

Loudness: the loudness of the music improves the popularity of the music by 0.304.

Random Effects:

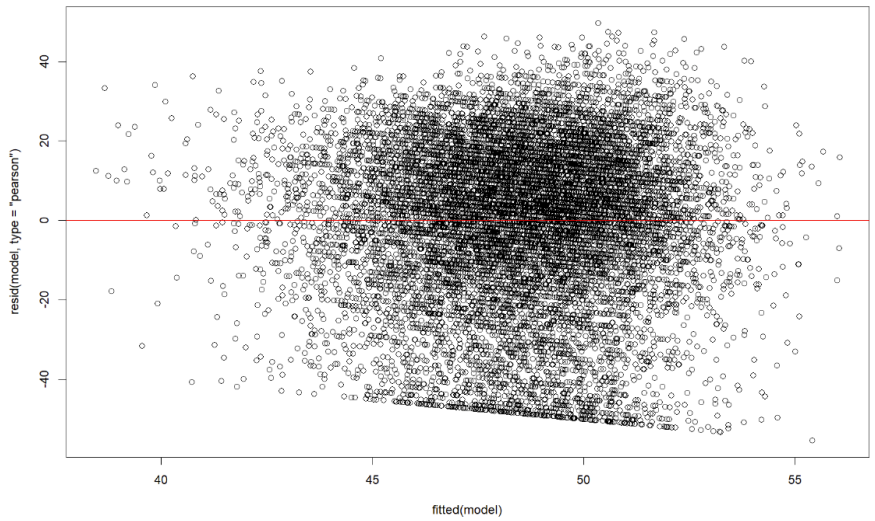
Sd(Intercept)-key: there exist differences between keys for the popularity of songs by 1.603.

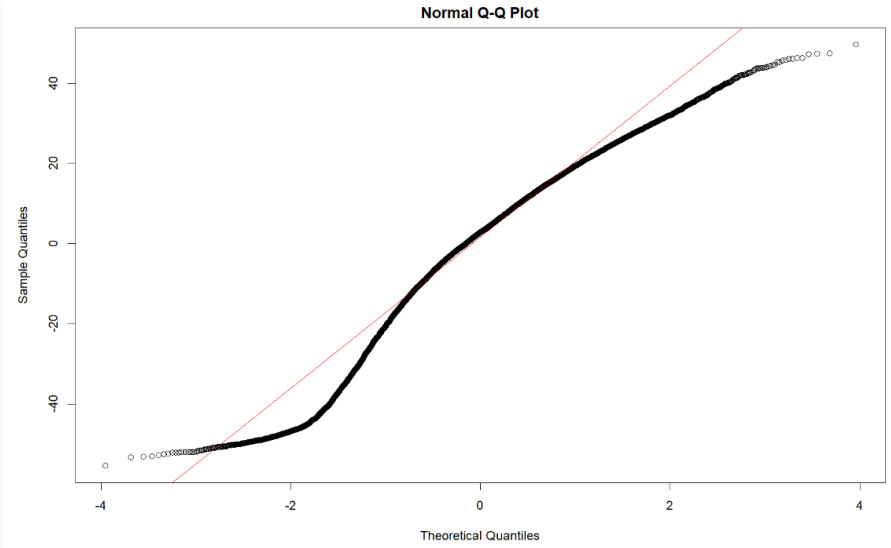
Sd(Tempo): there exist differences between tempo and song popularity in different key groups by 0.000015.

Sd(Danceability): there exist differences between danceability and song popularity in different key groups by 1.869.

Cov(Danceability, Intercept): the correlation between random slope and random intercept is -0.76, the high correlation means that those songs with high intercept values of popularity are more suitable for dancing.

I also drew a fitted value plot and QQ plot to check the model, the results show below.





The fitted value versus residuals plot looks random and the Q-Q plot also looks normally, so the model looks good.

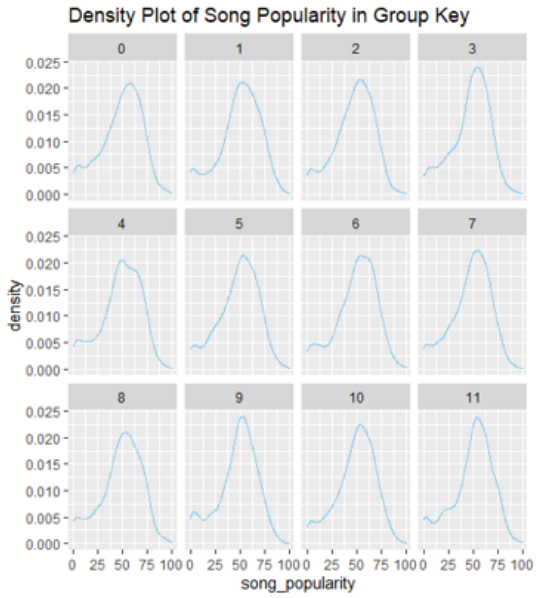
Discussion

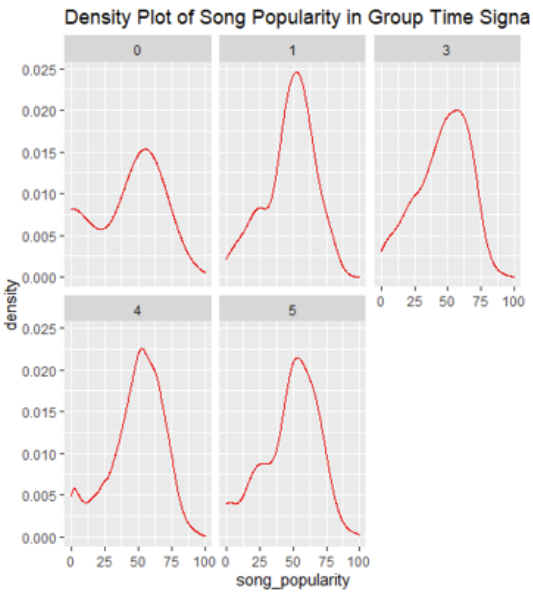
This project aims to know what factors can impact the popularity of a song. I use a multilevel model, and three group levels, including key, audio mode, and time

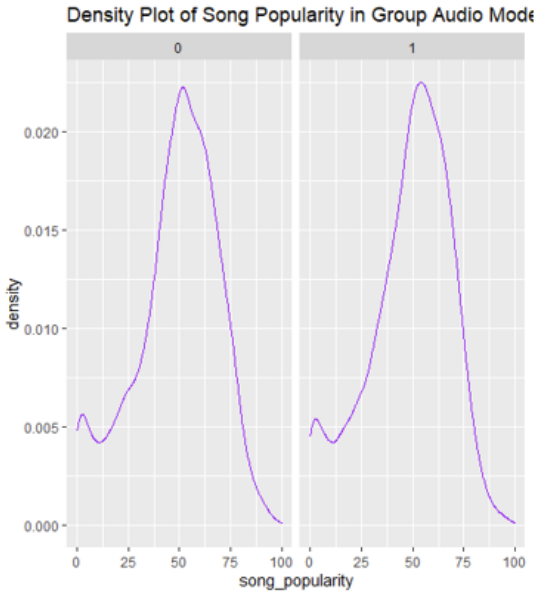
signature are the random effects in this model. From the results in the last part, we can find that the song suitable for dancing has higher popularity value, also the song decibels with high loudness value increase the popularity of the song. However, it is surprising that the valence with positivity decreases the popularity of the song. As a result, in terms of mental health, rhythmic and decibel songs have positive effects on mental health. Moreover, since the factor tempo has a negative impact on song popularity, slow-paced soprano songs are also better for mental health. Although the model check looks good, there are still some deficiencies in my model. I'm not sure about the year of release of the songs in the data, there is a link between the popularity of the songs and the year of release. It's possible that songs released earlier are more popular and that popular genres change over time.

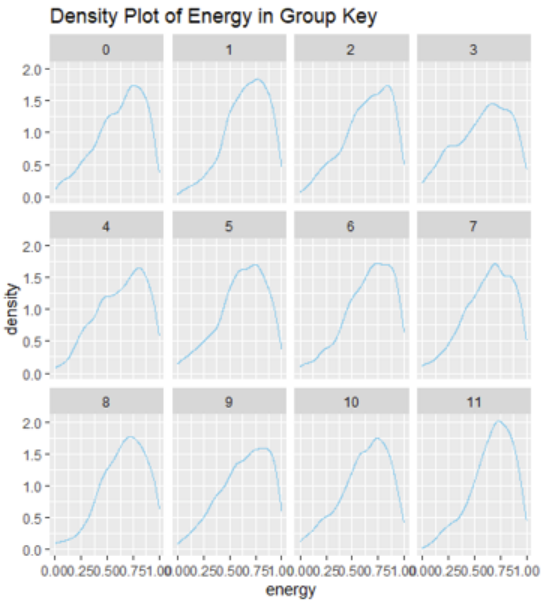
Appendix

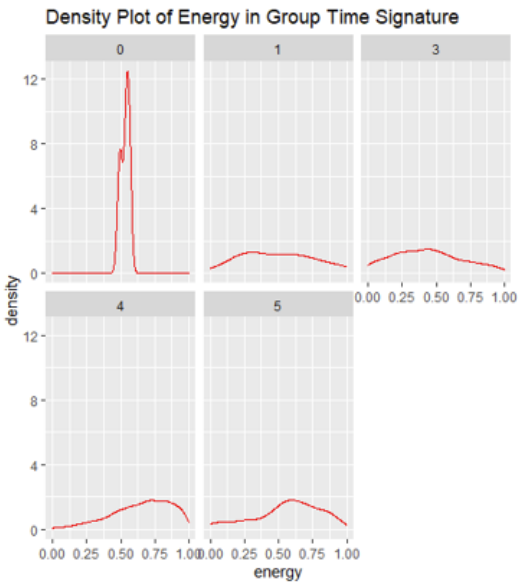
Density Plot

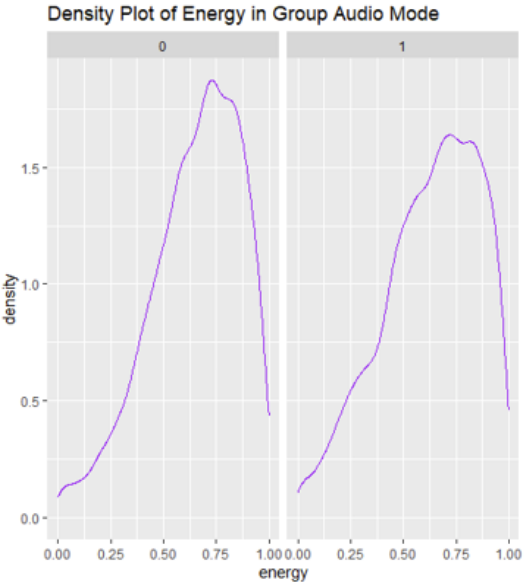


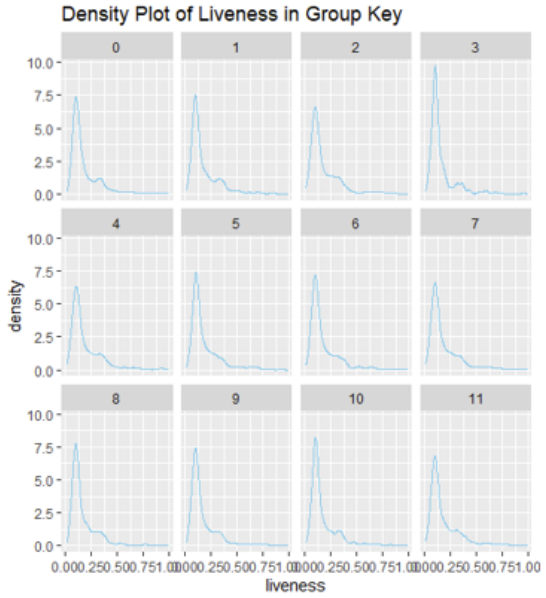


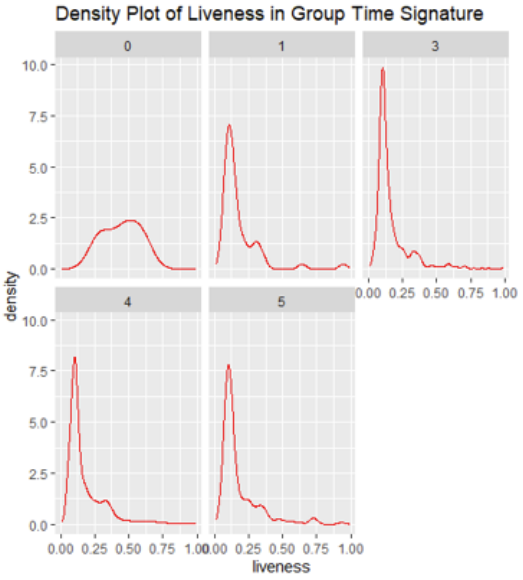


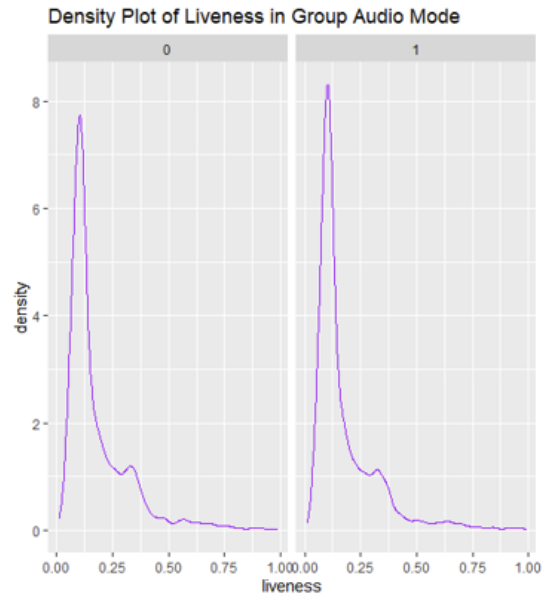


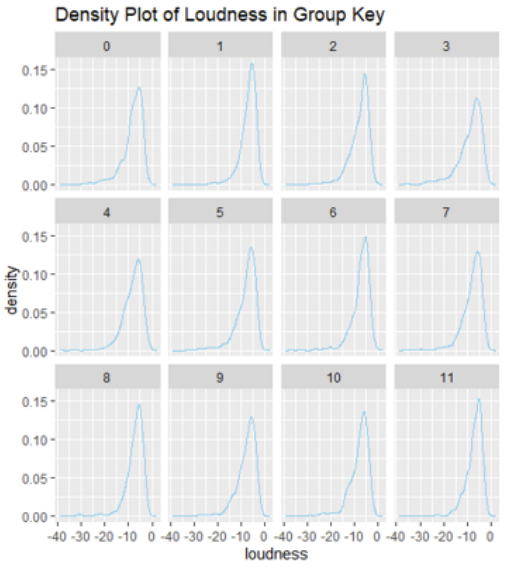


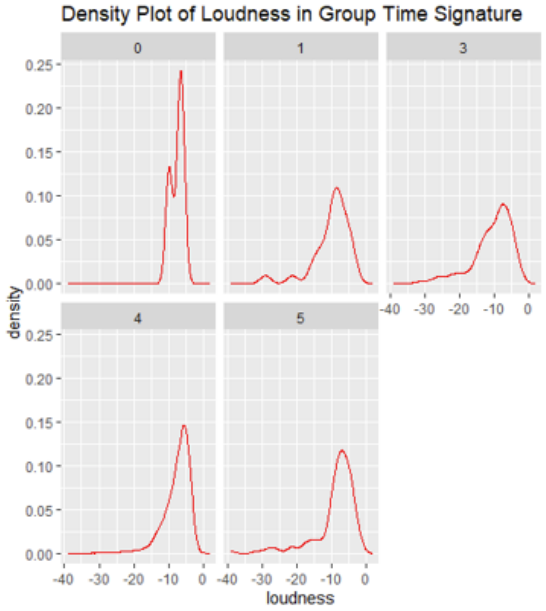


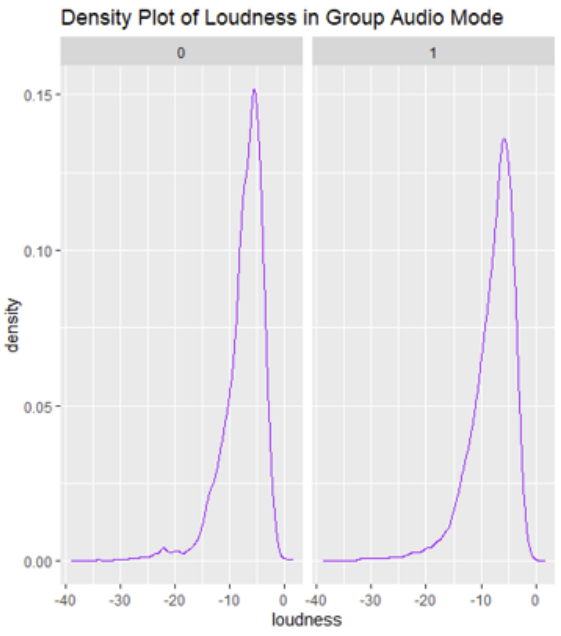

Density Plot of Liveness in Group Key
density
liveness

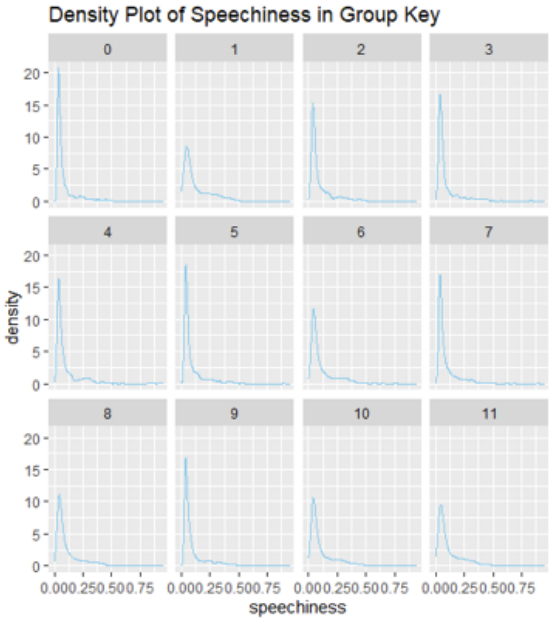

Density Plot of Liveness in Group Time Signature
density
liveness

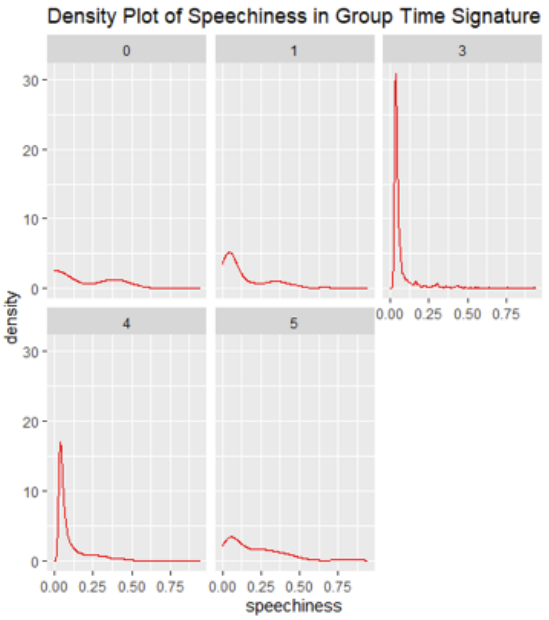

Density Plot of Liveness in Group Audio Mode
density
liveness

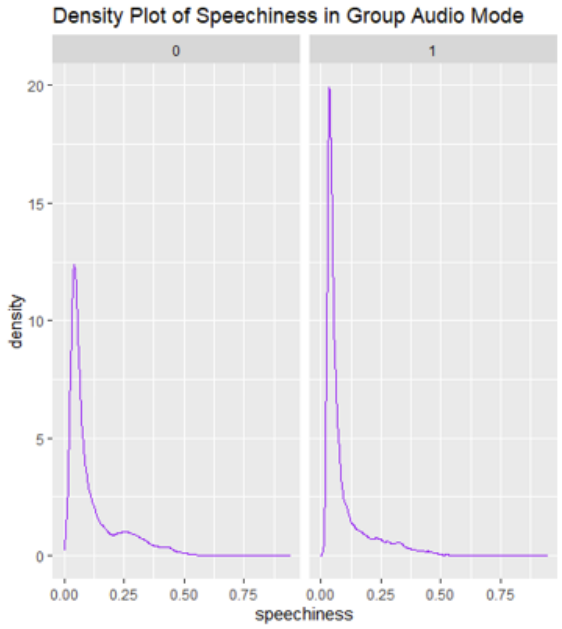

Density Plot of Loudness in Group Key
density
loudness

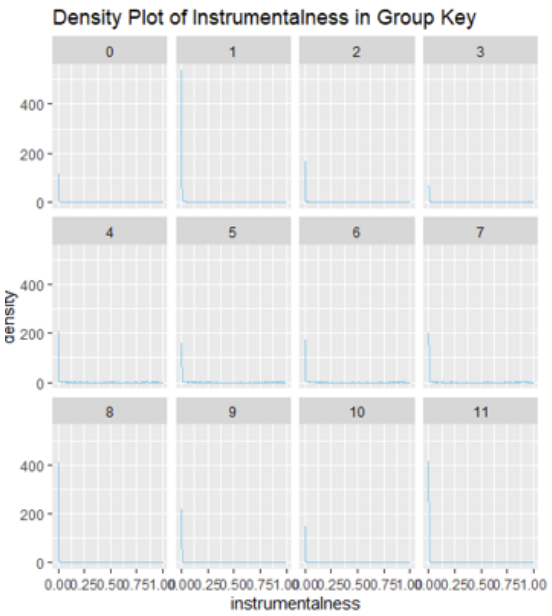

Density Plot of Loudness in Group Time Signature
density
loudness

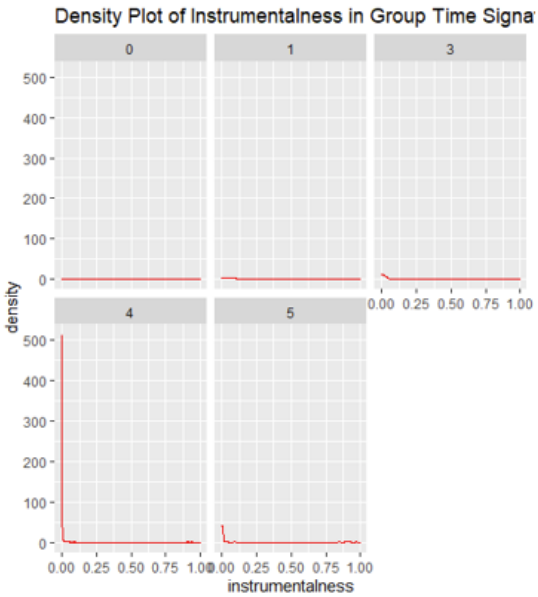

Density Plot of Loudness in Group Audio Mode
density
loudness

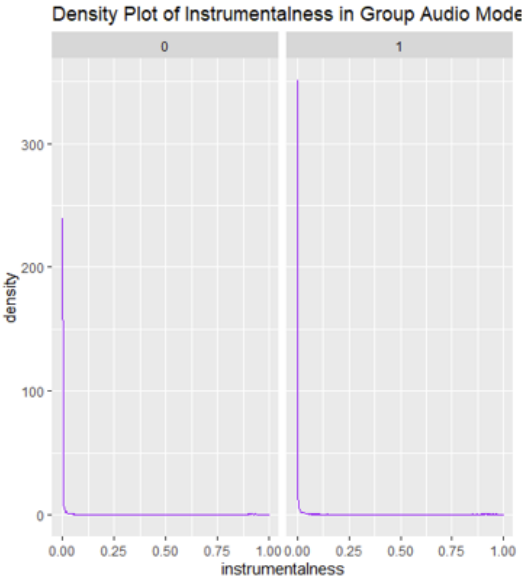

Density Plot of Speechiness in Group Key
density
speechiness

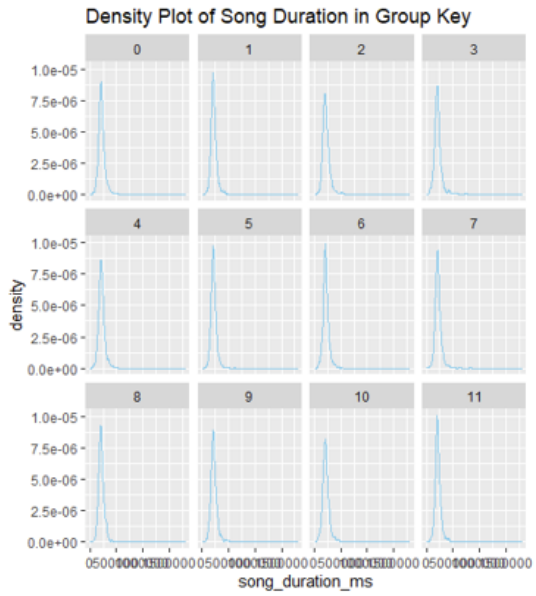

Density Plot of Speechiness in Group Time Signature
density
speechiness

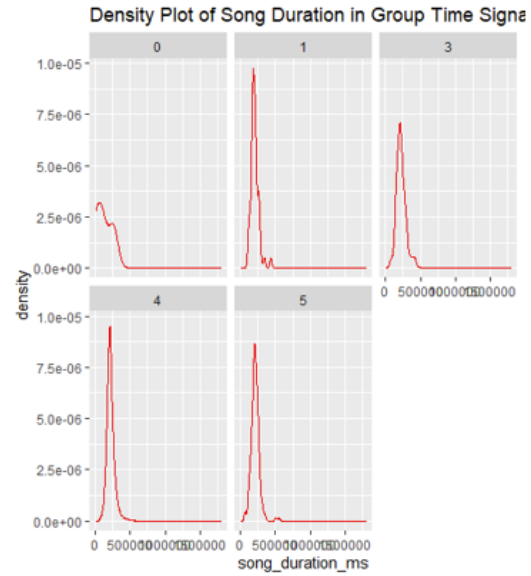

Density Plot of Speechiness in Group Audio Mode
density
speechiness

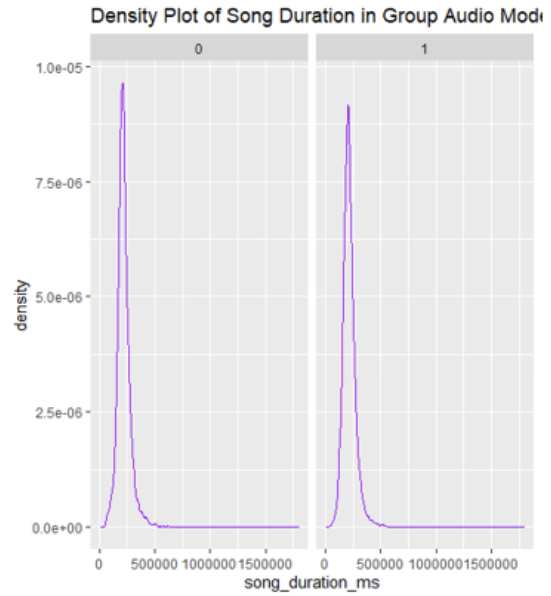

Density Plot of Instrumentalness in Group Key
density
instrumentalness

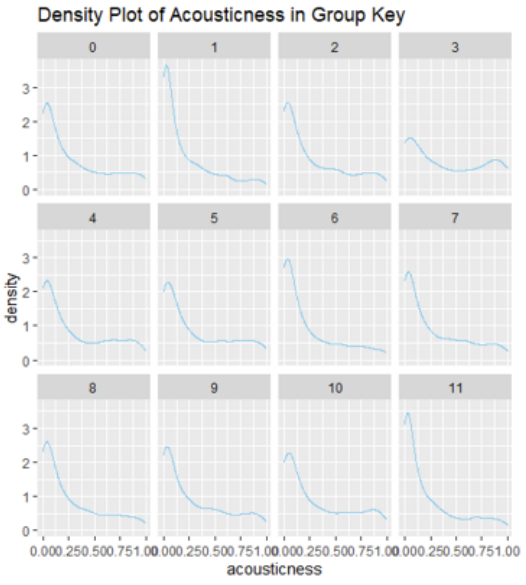

Density Plot of Instrumentalness in Group Time Signa
density
instrumentalness

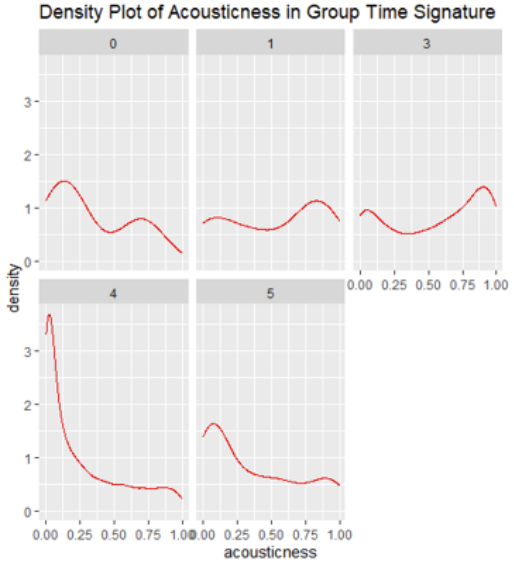

Density Plot of Instrumentalness in Group Audio Mode
density
instrumentalness

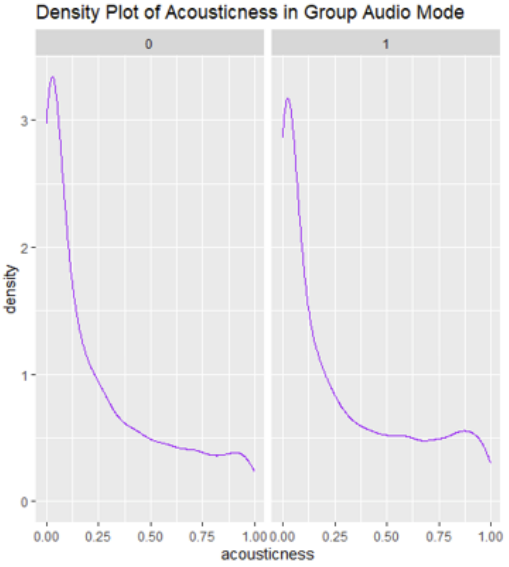

Density Plot of Song Duration in Group Key
0
1
2
3
4
5
6
7
8
9
10
11
density
song_duration_ms

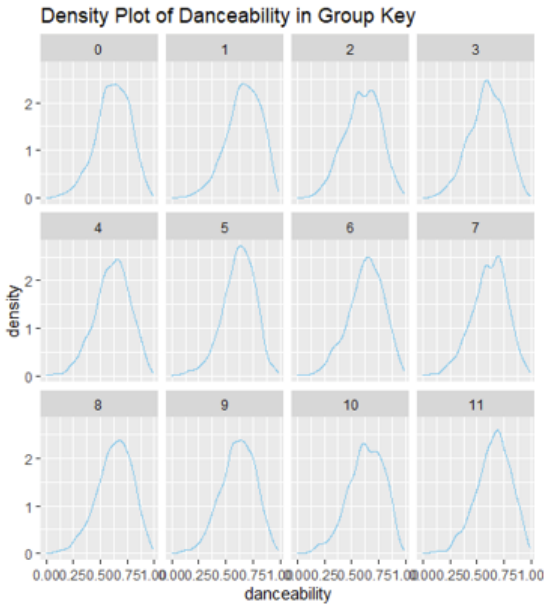

Density Plot of Song Duration in Group Time Signa
0
1
3
4
5
density
song_duration_ms

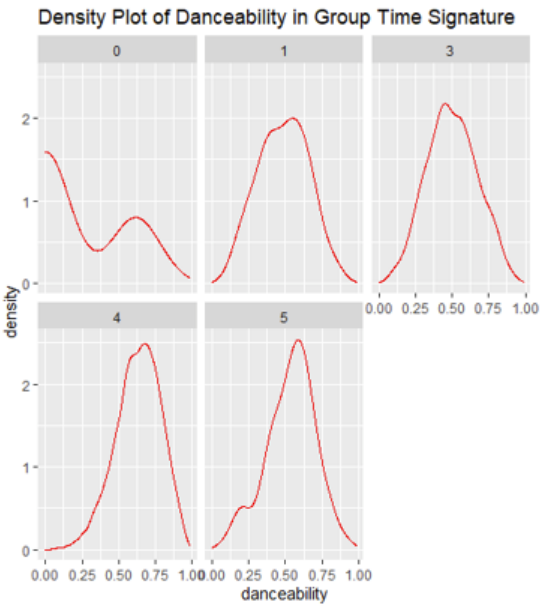

Density Plot of Song Duration in Group Audio Mod
0
1
density
song_duration_ms

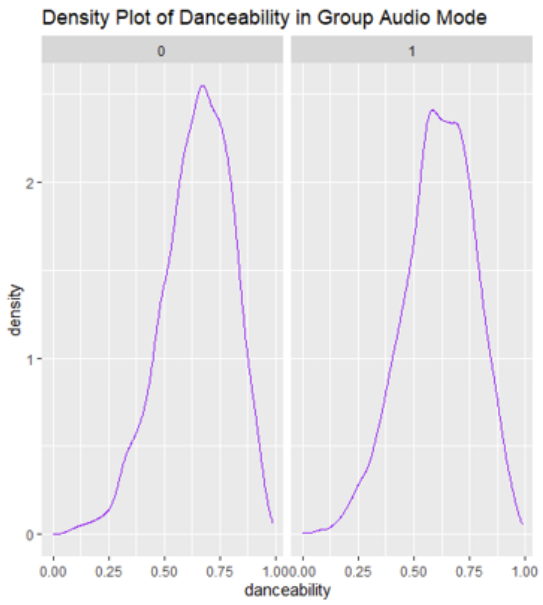

Density Plot of Acousticness in Group Key
0
1
2
3
4
5
6
7
8
9
10
11
density
acousticness


Density Plot of Acousticness in Group Time Signature
0
1
3
4
5
density
acousticness

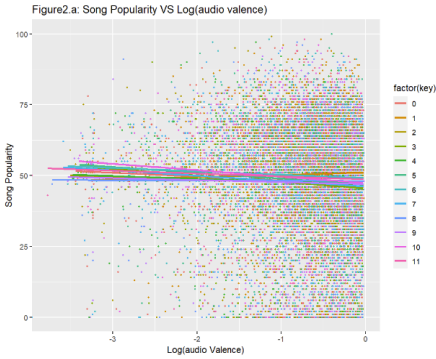

Density Plot of Acousticness in Group Audio Mode
0
1
density
acousticness

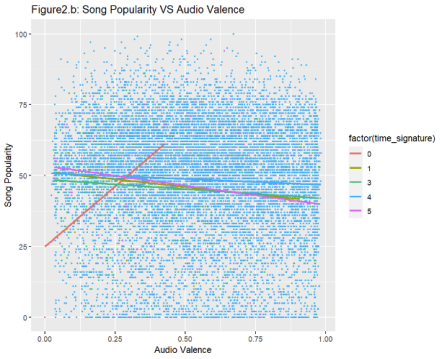






Variable Plot





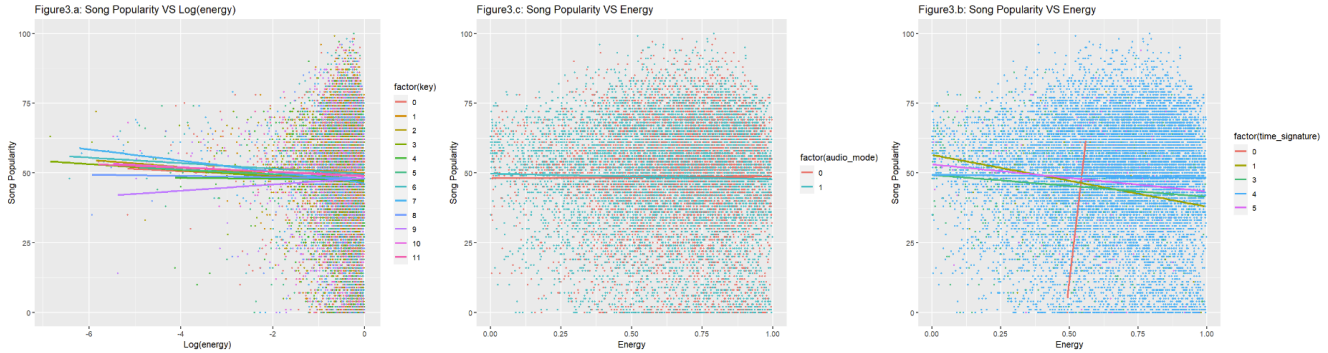


Figure3.a: Song Popularity VS Log(energy)
Song Popularity
Log(energy)
factor(key)
Figure3.c: Song Popularity VS Energy
Song Popularity
Energy
factor(audio_mode)
Figure3.b: Song Popularity VS Energy
Song Popularity
Energy
factor(time_signature)

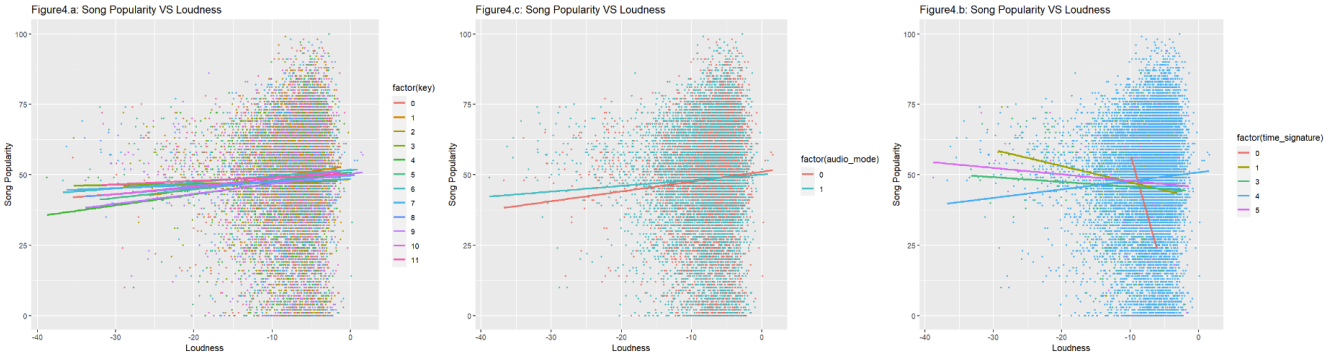


Figure4.a: Song Popularity VS Loudness
Song Popularity
Loudness
factor(key)
Figure4.c: Song Popularity VS Loudness
Song Popularity
Loudness
factor(audio_mode)
Figure4.b: Song Popularity VS Loudness
Song Popularity
Loudness
factor(time_signature)

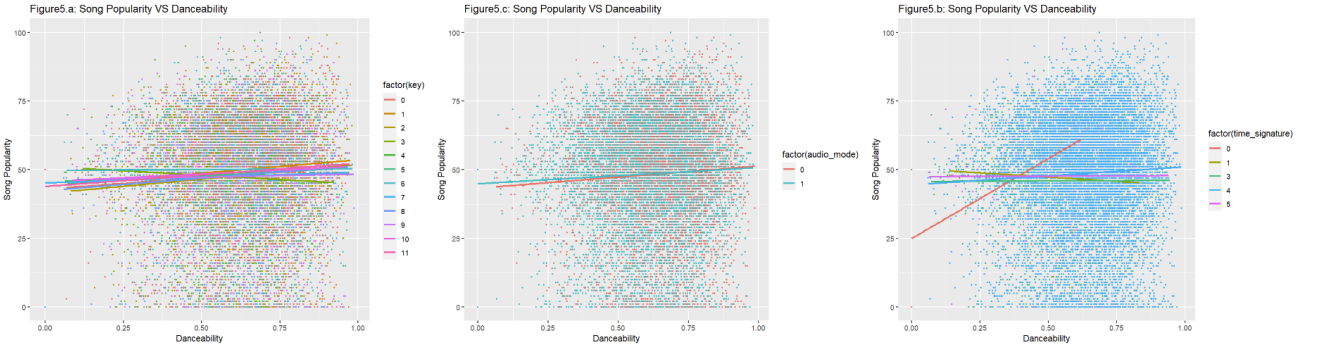


Figure5.a: Song Popularity VS Danceability
Song Popularity
Danceability
factor(key)
Figure5.c: Song Popularity VS Danceability
Song Popularity
Danceability
factor(audio_mode)
Figure5.b: Song Popularity VS Danceability
Song Popularity
Danceability
factor(time_signature)

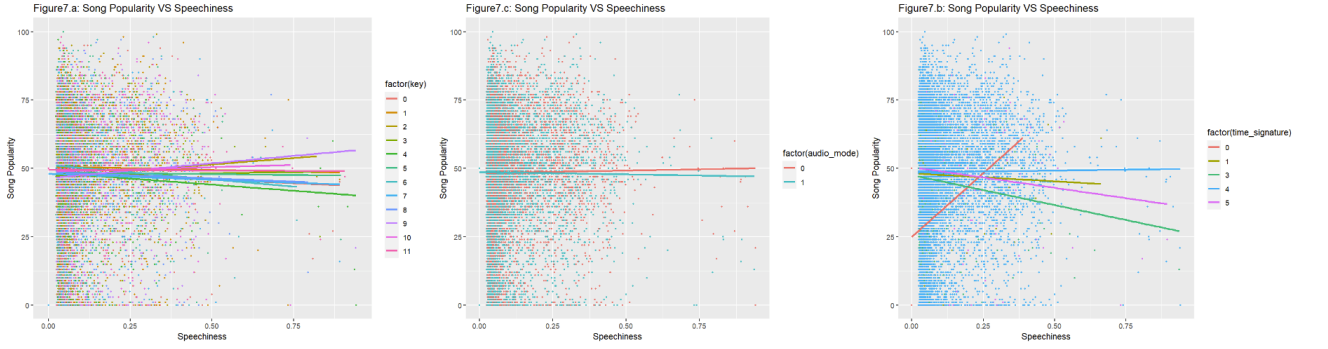
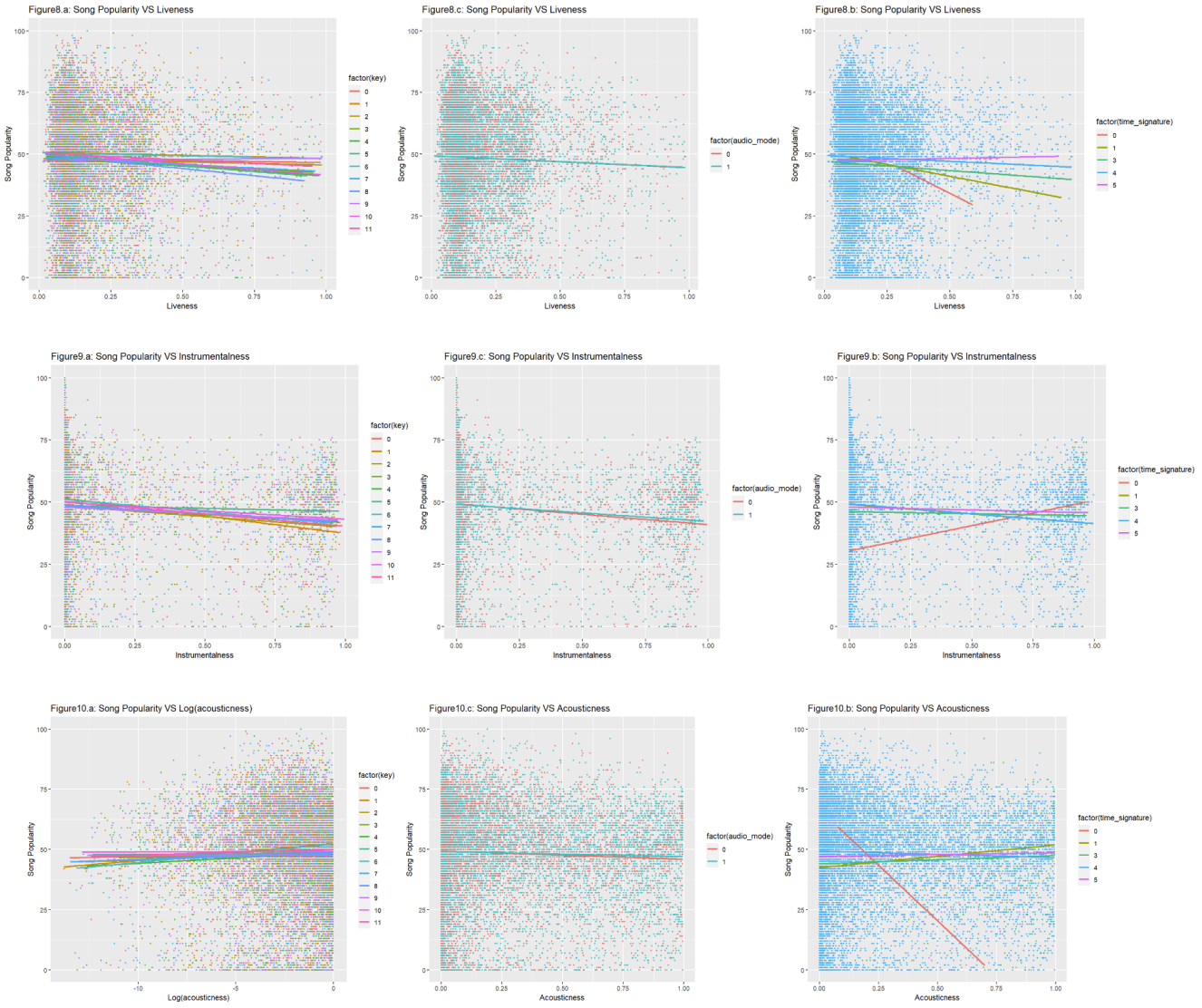


Figure7.a: Song Popularity VS Speechiness
Song Popularity
Speechiness
factor(key)
Figure7.c: Song Popularity VS Speechiness
Song Popularity
Speechiness
factor(audio_mode)
Figure7.b: Song Popularity VS Speechiness
Song Popularity
Speechiness
factor(time_signature)



Correlation Matrix