ST 552 Final Report

Chloe Drummond

2022-11-25

Abstract

The purpose of this analysis is to predict a song's popularity based on specific attributes. This data has been sourced from Kaggle.com. The link to the specific data set is in the Appendix. The data includes data pulled from spotify's API and has been compiled into one large dataset with 19 variables. This analysis focuses on multiple linear regression for the model building. After the analysis, the final model includes 17 total predictors with 12 main effect terms and 5 interaction terms. The prediction RSE is around 0.58 and has an R-squared value of around 0.12.

Introduction

The purpose of this analysis is to see what song attributes the biggest effect on the popularity have of said song. For context I will explain each of the variables for a better understanding. For a further breakdown of the variables in this analysis and others that could be added on to an analysis in the future please refer to the Spotify API for developers. It is also referenced in the works cited. The variables in the dataset from kaggle included popularity which ranges from 0-100 and in an integer value and it is our target variable in this analysis. Next is acousticness, spotify API says acousticness is a numeric variable that is a "confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic." Artists shows the artists that produced the song, the name of which is found in the name variable. Danceability is next, and it is a numeric variable as well described as "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." Duration_ms is the length of the track in miliseconds. Energy is numeric and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. Death metal, for example, has a lot of energy, whereas a Bach prelude has a low energy level. Dynamic range, perceived loudness, timbre, onset rate, and general entropy are all perceptual features that contribute to this attribute. Energy ranges from 0 to 1.

Explicit, key, and mode are the categorical varibles that are in the initial dataset. Please see the plots in the Appendix or the 12 Keys of music or the pitch class sources to better understand the keys variable. ID is a unique code assigned for each track. Spotify describes Instrumentalness as a variable that "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 is numeric and detects the presence of a live audience in the track. Higher liveness means a stronger likelihood that the track was live, and it ranges also from 0-1. Loudness is exactly how it sounds, it measures the loudness of a track in decibels, and its values are around -60 to 0 db. Mode indicates whether the song is in a major scale or minor scale. Major is 1, minor is 0. Release_date is the date the song was released. It has varying format as well. Speechiness is next and spotify tells us that "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.". Tempo is the tracks overall bpm over the entire track, so it is esstentially the average speed of the track in beats per minute. Valence is spotify's way to describe the "musical positiveness conveyed by a track". The higher the valence the more positive the track and valence ranges from 0-1.

The purpose of my analysis was to see if I could build a multiple linear regression model to predict popularity of a song based on its other features. This could be useful especially for record producers that want to use the model to see if they can find the next big song or star based on the attributes and features a song has. I first had to clean the data as there were some issues with the dataset right off the bat. Then after data cleaning I moved to the model building steps where I split the data into a training and test set and then I plotted the interactions between the numerical and categorical variables. I also found the model variance inflation factors and removed energy from the analysis. I also transformed both the response variable, popularity and I transformed Instrumentalness. I tested multiple different transformations on multiple variables to help remedy some the assumption issues my model has and found that while not perfect, a square-root transformation on both popularity and instrumentalness works the best to improve model fit. Then after the model building, I moved on to variable selection. I ran forward, backwards, and stepwise

selections with both AIC and BIC criterion. I found that the best model with an AIC of -14700.81 and an R-squared value of 0.1305 to be the best and final model as it was the model with the best fit statistics and lowest AIC or BIC of all the models. Then with the final model I put the test data set through it and got back an r-squared of 0.1217, adjusted r-squared of 0.1149 and an rse of 0.5815.

Methods

Cleaning of the data

First I read in the following libraries, `stringr, rsample, corrplot, ggplot2, car,glmet, moments, bestNormalize`. Then I read in the dataset into the dataframe spotifydataraw. Then to make the artists variable more readable I removed the square brackets. Then I checked to see if any observations in the dataset were duplicated. By looking at table 1 I could see that there are no duplicated observations. I also checked a summary of the data and found an issue right away with Tempo. First, I removed variables that were not of interest to the analysis. For right now the only variable I will be removing is ID. This subset is extracted into the dataframe `spotifydata`. The summary of tempo shows that there are some observations with a tempo of 0 which is impossible for a song to be 0 bpm, so I removed all observations that had a tempo equal to 0 and updated `spotifydata`. Then I wanted to see if specific tracks were duplicated, meaning observations matched both the track name and the artist(s). I found that 13301 observations in the data set were the same song. I kept the first encountered observation of those repeated in the unique Spotify data set. Meaning `uniquespotifydata` is the set of all tracks just once.

Then I wanted to see of release date was any different information than year provided. Using the `substr` function I extracted the year from release date and then compared it to year to see if there were any observations that differed. Looking at Table 2 in the appendix it shows that there are no unique observations that have differing release_date years and the year it made the chart. So, knowing that I removed the variable release_date in `uniquespotifydata` Looking at Year and taking into account the target audience my model aims to assist, I knew that the more recent years would be the areas of interest as most record producers are going to be trying to find newer artists or newer songs that follow the trend from the current trend. With this I updated `uniquespotifydata` to include a new variable `yearbins` whichs splits the data into two groups, those tracks released before 2010 and those released in 2010 or after. Then I removed those that were released before 2010 and removed `yearbins` as it was no longer of use. Now I check the response variable popularity. Looking at histogram 2 and table 4 in the appendix its clear that a

large amount of the observations had a popularity of 0. According to the information from the kaggle dataset page, these songs all made the top 100 atleast one year. I believe that a population of zero wouldn't have been very illustrative of the analysis I was focusing on so I removed them to focus on songs with a popularity of atleast 1. This is `spotifyminpop`. Looking at the histogram 3 and table 5 in the appendix it's clear that popularity is much more normally distributed.

Then I need to do something artists to make it meaningful. Instead of trying to make it a factor with artist names as the levels, which I knew would be massive and very difficult to interpret accurately, I wanted to make use of the fact that some songs have more than one artist producing them. Does collaboration in turn behind the scenes make a song more popular? I checked for a "`,`" in the artists column which was the indicator that there was more than one artist and then coded a 1 if there was more than one artist and 0 if there was just one. This was assigned as morethan1artist. I bounded it to spotifyminpopart. Then to make duration_ms easier to interpret I made it into how long the song lasted in minutes by dividing by 60000. This gives us durationmin which I bound to spotifyminpopart to make spotifyminpopartdur and then removed duration_ms. Now I need to fix the structure of the categorical predictors. I conveniently place the factor variables together in the dataframe to make it easier to keep track of. Now I want to see the overview of my categorical variables, key, morethan1artist, explicit, and mode are distributed against popularity. Please see boxplots 1-4 in the appendix. After evaluating the boxplots, I subsetted the data to only include songs that had an instrumental score of 0.5 or less so I could focus on songs that a record producer would be analyzing. This also exposed many songs that were not actually songs. Which leads me to my next step which is the Removal of non-applicable songs. Unfortunately, there wasn't a simple way for me to filter out all these songs that were like brown noise and rain sounds. Most of them have normal numbers for their variables but usually they had at least one extreme case in one variable. But of course, that also wasn't always the case in the reverse, so I did searches for keywords like sleep, nature, noise, rain, sound, lullaby, and relax, and then combed through the searches by hand to remove the observations that did not belong in the analysis.

After this I moved to outlier removal. I wanted to evaluate an initial model to see if there are any outliers that can be removed. Since the model is still relatively large, I felt ok about removing what the outlierTest function told me along with points that were flagged in the residuals vs leverage plot. I did 5 rounds of outlier removal and ended with a final model shown by `modeloutlier5` in the Appendix under Outlier Removal Final. Now that the outliers have been

removed, I wanted to remove some of the overly influental points that were affecting the model. The code, plots and output for this can be found in the Appendix under Influential Point Removal. This point removal is done using the function `influencePlot` and by evaluating the plots of the models as well. I chose to remove the point if the studentized residual was over 0.5 and if when it was removed the model fit increased. This led to five rounds of removals of influential points to improve the fit noise of the data. Each round has been separated in the Appendix so the plots can be examined by each model easier. The final data frame of the removed data is shown under Infuential Point Removal Final titled `spotifyremoved`. This was the final step of my data cleaning process. To prep for the model training steps that follow I split the data into a training and test set. I chose to split 65/35 which is a bit against convention. After observing the data for the cleaning, I could tell that I was going to want a larger test set than 30% but 40% seemed to close to a 50/50 match and I knew I wanted to give the model as much data as I could to ensure I got decently predictive model. With this predicament I chose to go down the middle with a 35/65 split. This split gives `spotifytraining` and `spotify_test` and can be found in the Appendix under Data Split into Training and Test Data.

Model Building

The first thing I did in the model building process was to check the correlation of my numerical variables, `popularity`, `durationmin`, `tempo`, `loudness`, `acousticness`, `danceability`, `energy`, `instrumentalness`, `valence`, `liveness`, and `speechiness`. This is seen in the Appendix under Correlation Plot. Looking at the plot we can see that energy is highly correlated with both loudness and acousticness. It has a positive correlation with loudness and a negative correlation with acousticness. Acousticness also appears to be correlated with loudness. Knowing that we have some correlation within my predictors I need to check the variance inflation factors to see which variable can be accounted for by the other predictors. This is my next step in my model building. In the Appendix under the Variance Inflation, I started with an initial model `modelvif1` using the training data. Then I checked the model's variance inflation factors using the `vif` function. It showed that energy was accounted in the model by two other variables, so I chose to remove energy first. The other option that had a high VIF was loudness, but it was not as high as energy, so I still went with energy first. I removed energy from the model in `modelvif2`. By comparing the output with energy and without energy it seems like the removal of energy hurts the model fit. But when we look again at the VIF of the new model it seems like the multicollinearity has

been settled by removing energy. Therefore, I will go to the next step of the model building without including energy in the analysis.

The next step in the model building was to check to see if any predictors had interactions with any of the categorical predictors in my model. I did this by examining the interaction plots of all predictors against popularity and plotted the best fit lines of each level in the interaction variable. These plots can be found in the Appendix in the Interaction Plots Section. I also included commentary on each of the plots on if it showed a presence of an an interaction or not. The interactions that I could exclude from the full interaction model based on the information from the interaction plots are tempo:mode, loudness:mode, acousticness:explicit, acousticness:morethan1artist, danceability:explicit, danceability:morethan1artist and valence:explicit. This leads us to the next step of the model building process, interaction term removal. The full interaction model can be found in the Appendix under Interaction Terms in the first section Full Interaction Start. There also is the summary output and the four model plots looking at the Multiple R-squared and Adjusted R-squared of this model compared to those in the model before the interactions we see the interactions have greatly helped the model fit improve however we can see that some of the interactions are not significant. This leads to the sequential interaction term removals that follow Full Interaction Start in the Appendix. We end on `modelinteraction10` as the next interactions that are not significant, durationmin:explicit, tempo:morethan1artist, and instrumentalness:morethan1artist actually hurt the model when removed instead of improving the fit. When I tried to remove each of the interactions individually the R-squared values, both the Multiple and Adjusted, decreased and the RSE also increased instead of decreasing. So, while the model is displaying the interaction as insignificant, I chose to keep them as I wanted to also have a maximal model before I went to variable selection.

Looking at the model plots, it does seem like the fit has improved as the normality plot has become more linear than it was, and the residuals/leverage plot line is now more straight and same things with the Residuals vs fitted plot line. So, I moved to the next step in the model building, transformations. Transformations in this model are very limited as most of the plots I felt were acceptable. The issue in the plots appears in most clearly in the normal Q-Q plot. The line mostly fits the line except for the lower tail around -2 standardized residuals and then even less of a deviation around 3 standardized residuals. The other issue is shown in the Residuals vs fitted plot. This leads me to my first transformation. The plot of the fitted values looks curved at the lower fitted values, so I applied a square-root transformation to popularity and ran the model. Shown in the Appendix

in the Transformations section in the Popularity section. Looking at the model fit statistics, both R-squared values increased and the RSE also decreased. Looking at the model plots as well and looking at the range of my remaining variables I chose to also take the square root of the variable instrumentalness. It ranges from 0 to 0.5 with many more 0s than other values. I chose to take a square root as I wanted to rein in the upper values of the stop transformations here. Some of my variables like loudness include a majority of negative values but that implies a quiet song so I didnt want to change the sign. Also, most of the distributions against popularity were clusters more than spread points so it was hard to see any pattern that could transformable. Maybe with further experience I could explore this further. With this I move to the next step of model building, Variable selection. All variable selection models can be found in the Appendix labeled with the type of selection and also then by criterion. From the fit statistics found from each selection process, it seems like the AIC criterion provides a better fit than the BIC criterion offers. There were two models with the same lowest AIC score, the stepwise and backwards AIC selection. They both gave an AIC of -14700.81, however the stepwise AIC model gave just slightly high R-squared values and a smaller RSE, so I chose the stepwise AIC model as my final model for this analysis. After that I moved to testing the model. I tested my model by using the same variables as the model gave out but instead used the test data on the model. It gave me an R-squared of 0.1217, an adjusted R-squared of 0.1149, and an RSE of 0.5815. This result was as expected and a deeper explanation of the model can be found in the Conclusion and Discussion section of the paper.

Results

The final model includes 12 main effects and 5 interaction terms for a total of 17 predictors. Both speechiness and loudness main effects are marked as not significant however their interaction terms are significant, so they are included in the final model. The variables and coefficients for each variable included in the final model can be found in the Appendix under Stepwise in AIC. Most of the coefficients in the model don't exceed 1 except for the intercept which is only at 7.75 which suggests that little changes in the input numbers shouldn't change the popularity as drastically.

The final model tested on the test set gave a result that was to be expected. The residual standard error increased since my model might have been slightly overfitted to the training data. This would of course cause greater residuals and thus an increase in the residual standard error. From the training to test data the multiple R-squared dropped from 0.13 to 0.122, and the adjusted r-squared dropped also from 0.1265 to 0.1149. This again makes sense since the model was fitted around the training data and when introduced to the test data it is not able to account for the

extremities in the test set that may have been pruned out of the training data during model building or variable selection.

Also, the model fit in general is quite low for both the training and testing models. Looking at the plots it's clear that some regression assumptions are still slightly violated. The Normal Q-Q plots shows that the standardized residuals stray from the linear line at the ends of the residuals. For standardized residuals below -2 the model is wanting the values to be lower than what they are. For standardized residuals above 2 the model does the opposite and wants the points to be higher than they actually are.

The Residuals vs Fitted plot also is not randomly spread. It is fairly even around 0 except for fitted values between 6.5 and 7.0. The line on the graph no longer is straight and instead a negative slope appears. Since popularity (my target variable) is an integer variable instead of a continuous variable some of this issue might extend from that as well. We can see some evidence of this possible being an issue in the scale-location plot as well as an unusual pattern emerges. The residual vs leverage plot also shows an issue with the data as the points are mainly clustered around 0 leverage but there are a handful of points not within the cluster that are relatively randomly spread but are slightly sloped down.

Conclusion and Discussion

In conclusion, the multiple linear regression model I've developed might not be the best model to predict popularity. While the model is okay, it's still not fantastic. The assumptions that were violated as shown in the plots mentioned earlier in the results need to be fixed to be able to provide a model that could be suggestible. With more time and experience in transformations I hopefully could fix the assumption of linearity a little better than I did here. Also, with further knowledge and experience in model building I would like to see if I could make a more robust model that could be resistant to outliers more than this model is.

With more time it would be interesting to see how this model could be shifted to focus on popularity of tracks in a person's specific playlist. That way it could be implemented as a "this is what you like the most:danceability" on Spotify to allow users to be more acquainted with the complexities of music. With more knowledge in R as well it would be interesting to see how other types of models that are not regression model the data and see if any other model provides a better fit. Since the regression model run here did not provide a great model fit, I would recommend

refitting with new transformations or try a different type of model like a neural network or decision tree before using the model in any real application.

Works Cited

12 Keys of Music. (2022). Retrieved December 7, 2022, from Zebrakeys.com website: https://www.zebrakeys.com/lessons/beginner/musictheory/?id=12

Correlation and Regression with R. (2013). Retrieved December 7, 2022, from Bu.edu website: https://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/R/R5_Correlation-Regression/R5_Correlation-Regression_print.html

DeRuiter, S. (2019a, December 2). Chapter 10 Binary regression: Data with more than one trial per row | STAT 245 Course Notes. Retrieved December 7, 2022, from Github.io website: https://stacyderuiter.github.io/s245-notes-bookdown/binary-regression-data-with-more-than-one-trial-per-row.html#checking-linearity

DeRuiter, S. (2019b, December 2). Chapter 11 Collinearity and Multicollinearity | STAT 245 Course Notes. Retrieved December 5, 2022, from Github.io website: https://stacyderuiter.github.io/s245-notes-bookdown/collinearity-and-multicollinearity.html

factor. (2013, September 22). Which variance inflation factor should I be using: GVIF or $\text{GVIF}^{1/(2\cdot\text{df})}$? Retrieved December 5, 2022, from Cross Validated website: https://stats.stackexchange.com/questions/70679/which-variance-inflation-factor-should-i-be-using-textgvif-or-textgvif#:~:text=For%20the%20two%20continuous%20variables%2C%20GVIF(1,to%20the%20level%20of%20collinearity.

Francis Adrian Viernes. (2021, March 2). Linear Regression Models and Influential Points - Towards Data Science. Retrieved December 7, 2022, from Medium website: https://towardsdatascience.com/linear-regression-models-and-influential-points-4ee844adac6d

In. (2012, July 19). In what order should you do linear regression diagnostics? Retrieved December 7, 2022, from Cross Validated website: https://stats.stackexchange.com/questions/32600/in-what-order-should-you-do-linear-regression-diagnostics

Negi, E. (2020). Spotify-Data 1921-2020. Retrieved December 7, 2022, from Kaggle.com website: https://www.kaggle.com/datasets/ektanegi/spotifydata-19212020

Qualtrics. (2017, July 19). Interpreting Residual Plots to Improve Your Regression. Retrieved December 7, 2022, from Qualtrics.com website: https://www.qualtrics.com/support/stats-iq/analyses/regression-guides/interpreting-residual-plots-improve-regression/

Web API Reference | Spotify for Developers. (2022). Retrieved December 7, 2022, from Spotify.com website: https://developer.spotify.com/documentation/web-api/reference/#/operations/get-several-audio-features

Wikipedia Contributors. (2022, November 29). Pitch class. Retrieved December 7, 2022, from Wikipedia website: https://en.wikipedia.org/wiki/Pitch_class

<div align="center">**Appendix**</div>

Libraries and data read in and duplicated observation check

Table 1 Duplicated Observations

```
table(duplicated(spotifydataraw))

##
##  FALSE
## 169909

summary(spotifydataraw)

##   acousticness       artists          danceability      duration_ms
##  Min.   :0.0000   Length:169909      Min.   :0.0000    Min.   :   5108
##  1st Qu.:0.0945   Class :character   1st Qu.:0.4170    1st Qu.: 171040
##  Median :0.4920   Mode  :character   Median :0.5480    Median : 208600
##  Mean   :0.4932                      Mean   :0.5381    Mean   : 231406
##  3rd Qu.:0.8880                      3rd Qu.:0.6670    3rd Qu.: 262960
##  Max.   :0.9960                      Max.   :0.9880    Max.   :5403500
##      energy          explicit            id            instrumentalness
##  Min.   :0.0000   Min.   :0.00000   Length:169909      Min.   :0.000000
##  1st Qu.:0.2630   1st Qu.:0.00000   Class :character   1st Qu.:0.000000
##  Median :0.4810   Median :0.00000   Mode  :character   Median :0.000204
##  Mean   :0.4886   Mean   :0.08486                      Mean   :0.161937
##  3rd Qu.:0.7100   3rd Qu.:0.00000                      3rd Qu.:0.086800
##  Max.   :1.0000   Max.   :1.00000                      Max.   :1.000000
##      key             liveness         loudness           mode
##  Min.   : 0.000   Min.   :0.0000   Min.   :-60.000   Min.   :0.0000
##  1st Qu.: 2.000   1st Qu.:0.0984   1st Qu.:-14.470   1st Qu.:0.0000
##  Median : 5.000   Median :0.1350   Median :-10.474   Median :1.0000
##  Mean   : 5.201   Mean   :0.2067   Mean   :-11.370   Mean   :0.7086
##  3rd Qu.: 8.000   3rd Qu.:0.2630   3rd Qu.: -7.118   3rd Qu.:1.0000
##  Max.   :11.000   Max.   :1.0000   Max.   :  3.855   Max.   :1.0000
##      name            popularity      release_date        speechiness
##  Length:169909      Min.   :  0.00   Length:169909      Min.   :0.00000
##  Class :character   1st Qu.: 12.00   Class :character   1st Qu.:0.03490
##  Mode  :character   Median : 33.00   Mode  :character   Median :0.04500
##                     Mean   : 31.56                      Mean   :0.09406
##                     3rd Qu.: 48.00                      3rd Qu.:0.07540
##                     Max.   :100.00                      Max.   :0.96900
##      tempo           valence            year
##  Min.   :  0.00   Min.   :0.0000   Min.   :1921
##  1st Qu.: 93.52   1st Qu.:0.3220   1st Qu.:1957
##  Median :114.78   Median :0.5440   Median :1978
##  Mean   :116.95   Mean   :0.5321   Mean   :1977
##  3rd Qu.:135.71   3rd Qu.:0.7490   3rd Qu.:1999
##  Max.   :244.09   Max.   :1.0000   Max.   :2020
```

Removal of ID.

Removal of tempo = 0

Separation of unique and duplicated songs
```
uniquespotifydata <- subset(spotifydata, duplicated(spotifydata[,
    2:3], data = spotifytraining) == FALSE)
```

Release_date year check

Table 2
```
##
##    TRUE
## 153199
```

Year visualization

Histogram 1
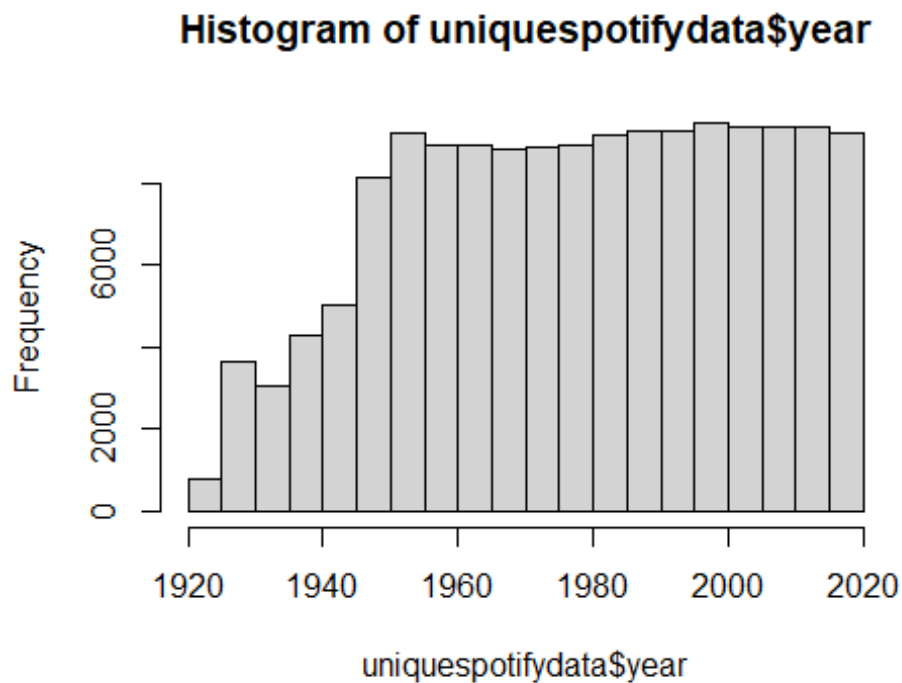
Year grouping

Table 3
```
##      before 2010 Recent 10 years
##           130887           22312
```

Popularity visualization

Histogram 2 and Histogram 3

popularity with 0

Frequency
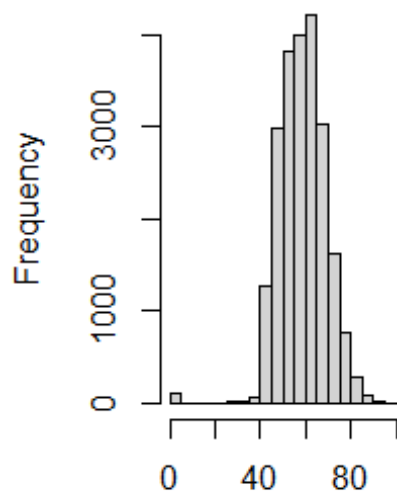
uniquespotifydata$popularity

popularity w/o 0

Frequency
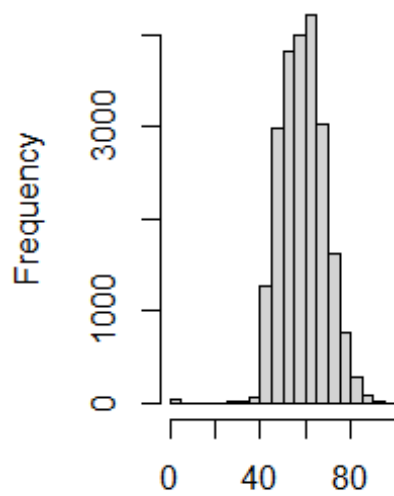
spotifyminpop$popularity
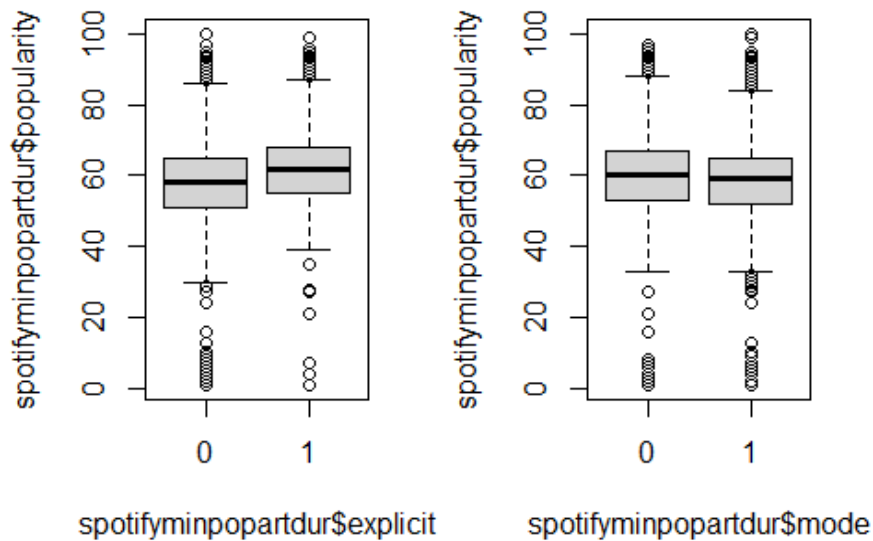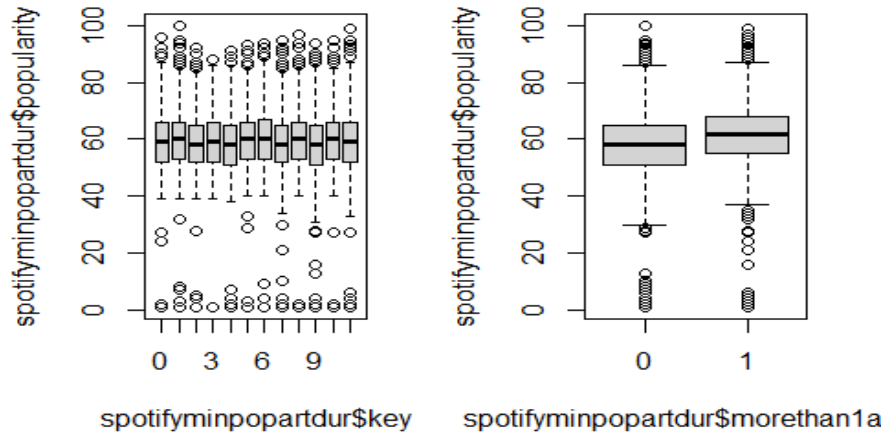
Instrumentalness

Removal of tracks that had a higher instrumentalness than 0.5

Used searches on sleep, nature, noise, rain, sound, lullaby, and relax

Initial Outlier Removal
```
## [1] 0.114301

## [1] 0.1153266

## [1] 9.430515
```
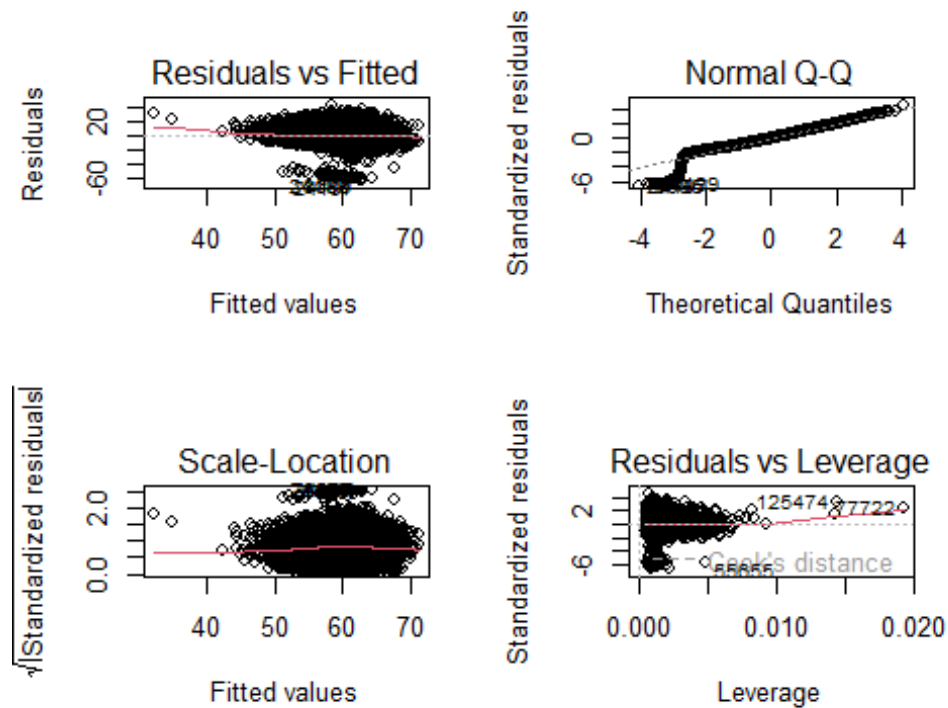


```
par(mfrow = c(1, 1))

##          24197            8085           24199           70600          126140              47
990
## 1.128438e-06 1.339584e-06 1.806562e-06 1.910192e-06 2.089137e-06 2.332777e
-06
##          77988           77987          169848           70610
## 2.512901e-06 2.622203e-06 2.704307e-06 2.820966e-06
```

Outlier Removal round 2
```
## [1] 0.1195466

## [1] 0.1185252

## [1] 9.325692
```

```
##        70636         161470         40206          16197         144251          78
013
## 1.984695e-06 2.847464e-06 3.914908e-06 4.284666e-06 5.345559e-06 6.828081e
-06
##         32235         161468         135501          70588
## 8.819674e-06 1.613125e-05 1.801590e-05 1.933536e-05
```

Outlier Removal round 3

```
## [1] 0.1223871

## [1] 0.1213683

## [1] 9.215796
```
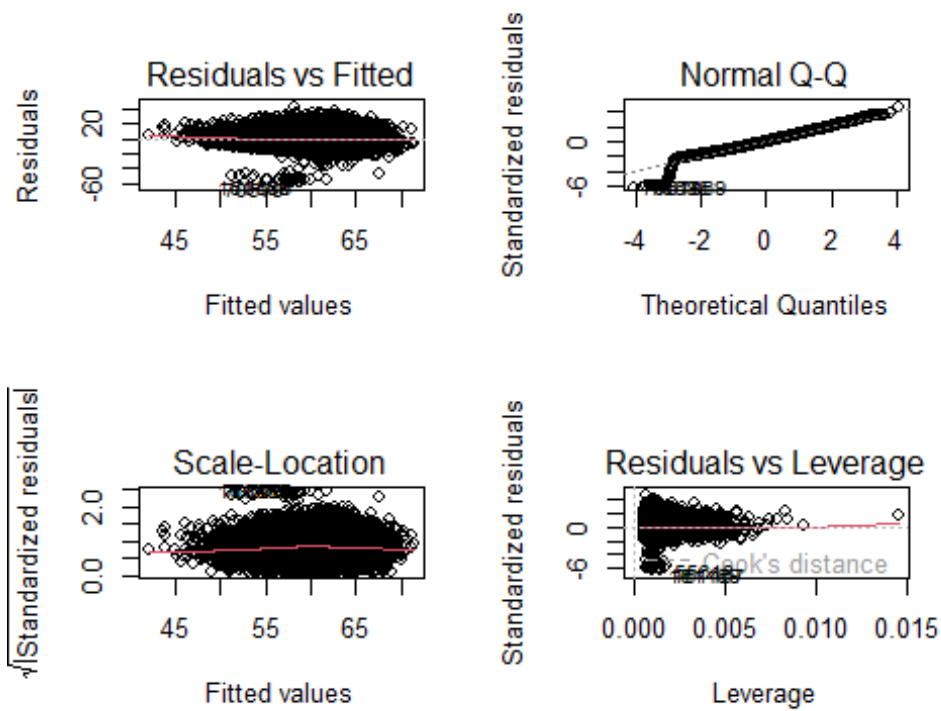
```
##         70578          70585         161469           8096          77989              8
104
## 1.350465e-05 1.387218e-05 1.648769e-05 1.778990e-05 2.359625e-05 2.496321e
-05
##         48010           8091          32177          78020
## 2.704001e-05 3.153942e-05 3.195312e-05 3.620858e-05
```

Outlier Removal round 4

```
summary(modeloutlier3)$r.sq
```

```
## [1] 0.1236292
```

```
summary(modeloutlier3)$adj.r.sq
```

```
## [1] 0.1226114
```

```
## [1] 9.126884
```

```
par(mfrow = c(2, 2))
```
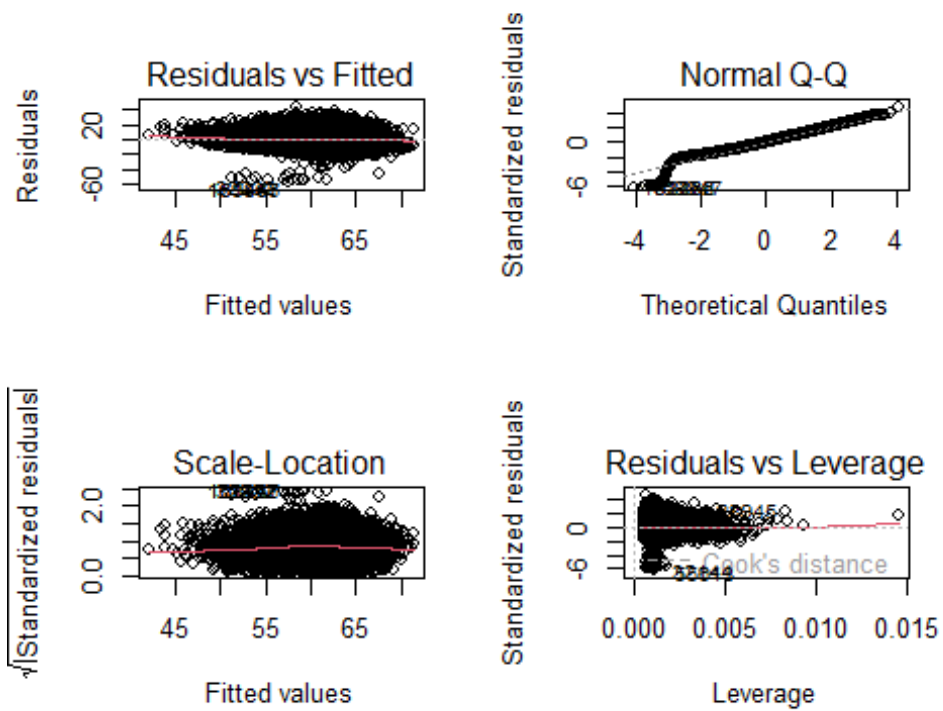
```
##         161466          169847          32232          70635          40203          144
253
## 2.440932e-05 2.476212e-05 3.268789e-05 3.915530e-05 5.381281e-05 5.446628e
-05
##          16213          24194          77985          48008
## 5.882778e-05 6.523048e-05 1.267362e-04 1.588924e-04
```

Outlier Removal round 5

```
## [1] 0.1251324

## [1] 0.1241156

## [1] 9.042857
```
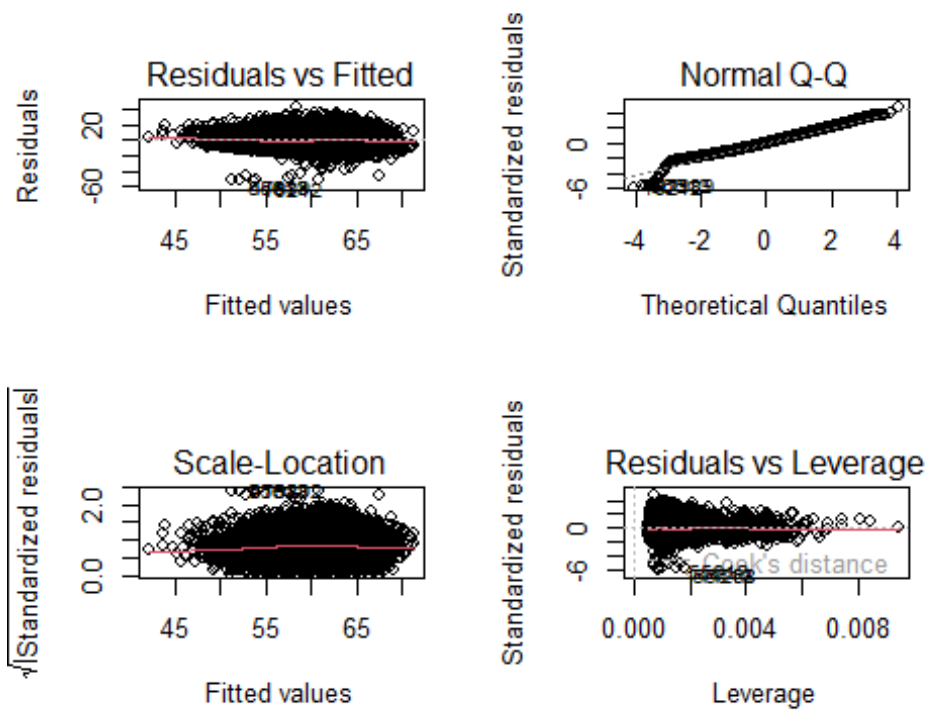
```
##         16212         55613         47989        144268        135500         24
192
## 0.0001179194 0.0003131498 0.0003559071 0.0005229612 0.0006544939 0.0009551
794
##        152879         55612         24168         40231
## 0.0052045154 0.0083312500 0.0138735362 0.0462259422
```

Outlier Removal Final

```
## [1] 0.1253734
```

```
## [1] 0.1243565
```

```
## [1] 8.991996
```
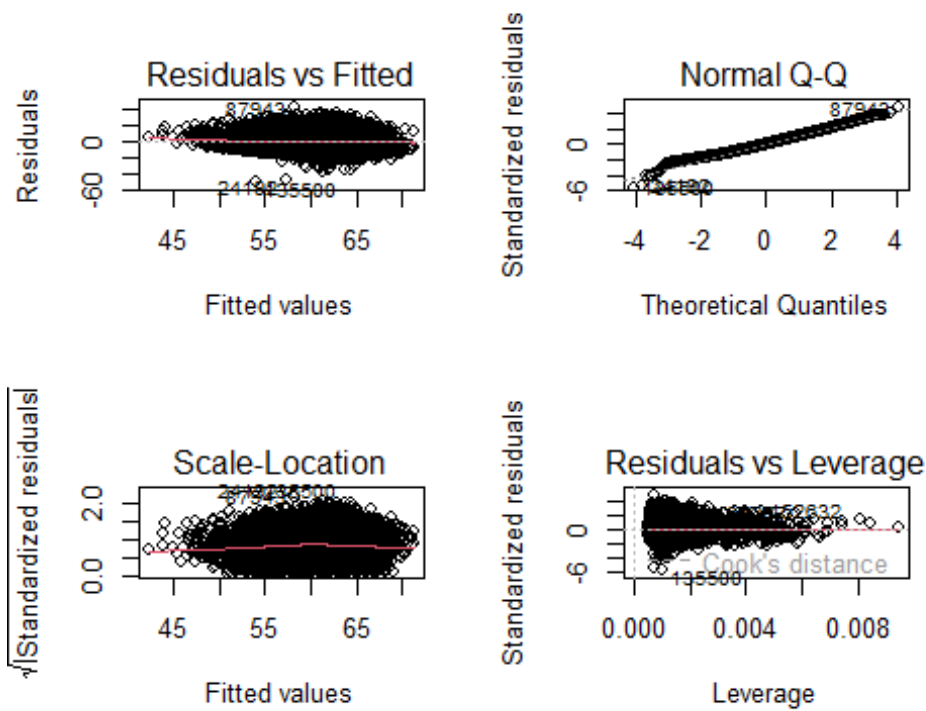
```
## [1] TRUE
```

Influential Point Removal 1
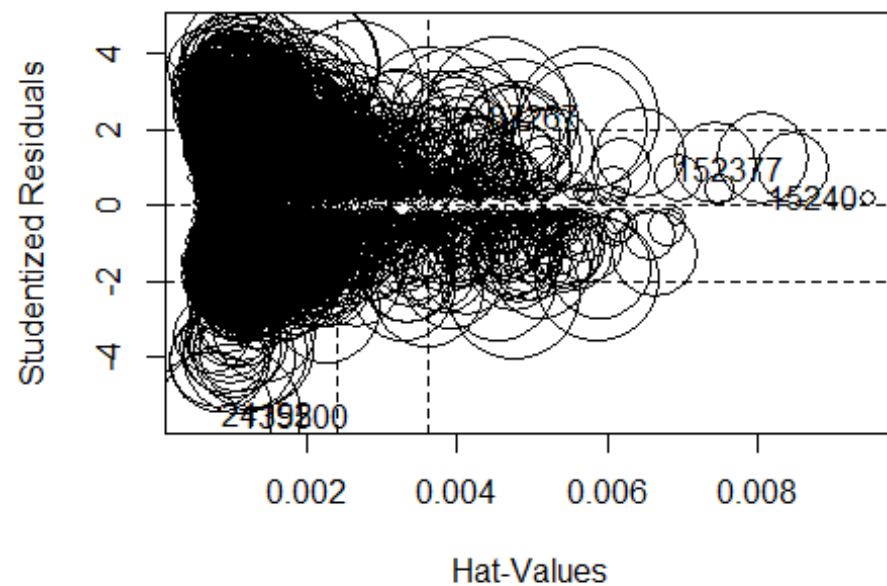
```
##            StudRes           Hat         CookD
## 15240    0.1898682 0.0094654836 1.378027e-05
## 24192   -5.5021829 0.0007538695 9.123012e-04
## 87267    2.2894175 0.0057513898 1.212547e-03
## 135500  -5.5686106 0.0009945061 1.232998e-03
## 152377   0.9435809 0.0084701948 3.042344e-04
```

```
modelinflupointremoval2 <- lm(popularity ~ durationmin + tempo +
    loudness + acousticness + danceability + energy + instrumentalness +
    valence + liveness + speechiness + key + mode + explicit +
    morethan1artist, data = spotifyinfluentialpointremoval1)
summary(modelinflupointremoval2)$r.sq
```

```
## [1] 0.1259627
```

```
## [1] 0.1249462
```

```
## [1] 8.982939
```
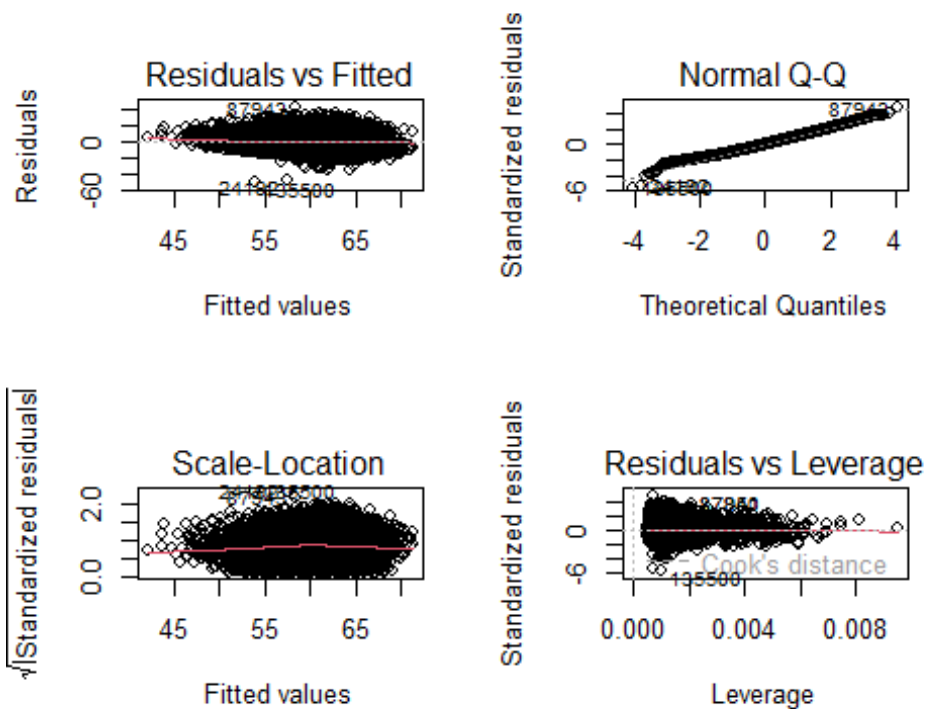
Influential Point Removal 2
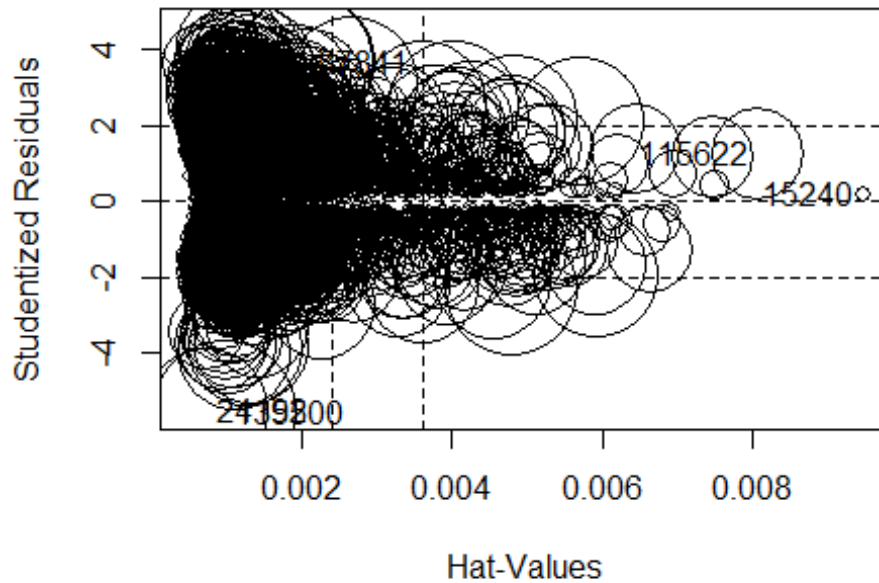

```
##            StudRes           Hat        CookD
## 15240    0.1897949 0.0094726588 1.378018e-05
## 24192   -5.5076308 0.0007539403 9.141917e-04
## 87841    3.6285769 0.0020930579 1.103997e-03
## 115622   1.2476261 0.0080833201 5.073781e-04
## 135500  -5.5730726 0.0009947518 1.235277e-03

summary(modelinflupointremoval3)$r.sq

## [1] 0.1265028

summary(modelinflupointremoval3)$adj.r.sq

## [1] 0.1254866

sqrt(deviance(modelinflupointremoval3)/df.residual(modelinflupointremoval3))

## [1] 8.967162
```
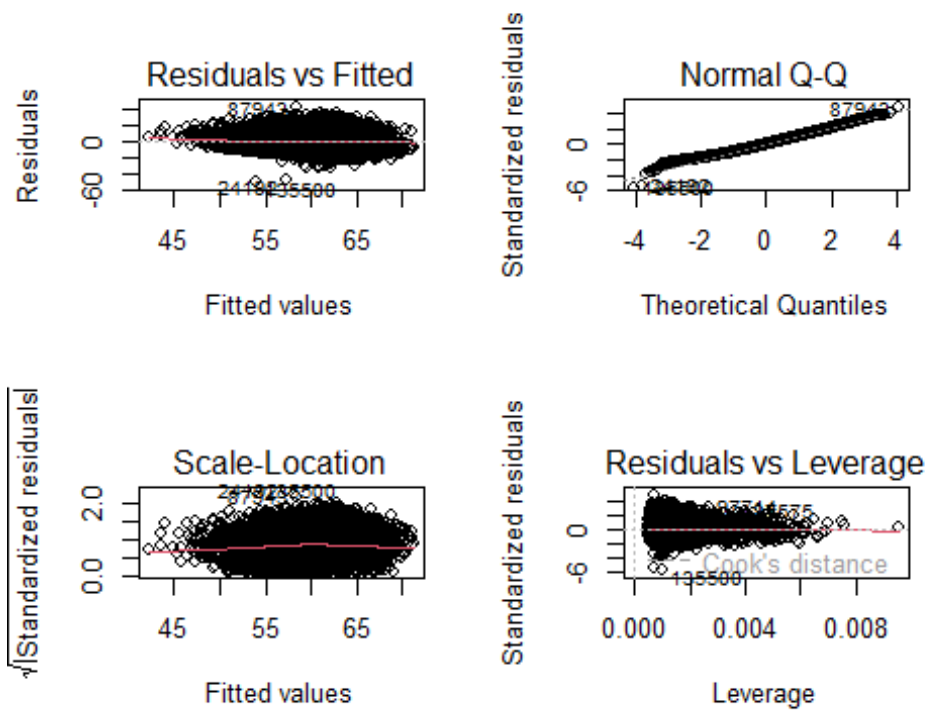
Influential Point Removal 3
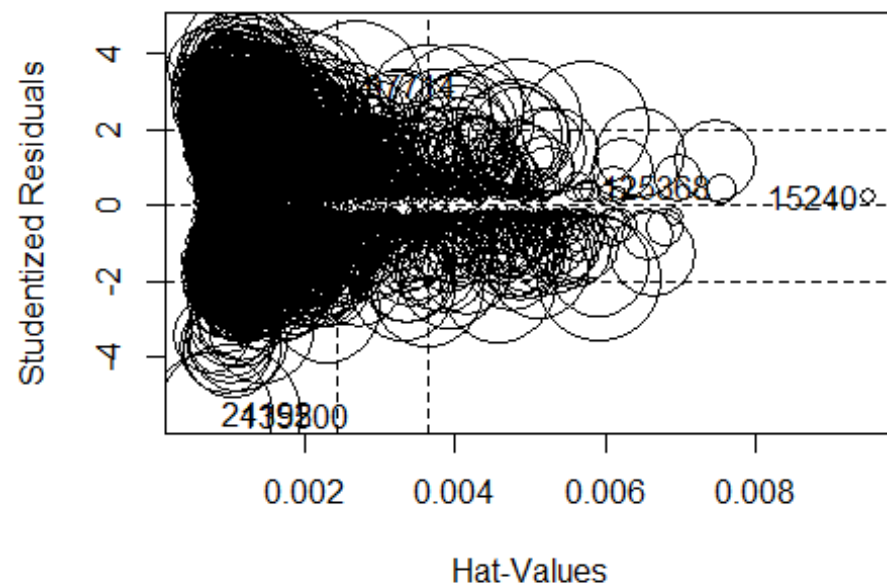
```
##           StudRes          Hat        CookD
## 15240   0.1993545 0.0095169634 1.527508e-05
## 24192  -5.5151783 0.0007541717 9.169764e-04
## 97714   3.1053075 0.0026839996 1.037617e-03
## 125368  0.4277046 0.0075489366 5.565980e-05
## 135500 -5.5823101 0.0009952142 1.239945e-03

summary(modelinflupointremoval4)$r.sq

## [1] 0.127072

summary(modelinflupointremoval4)$adj.r.sq

## [1] 0.1260562

sqrt(deviance(modelinflupointremoval4)/df.residual(modelinflupointremoval4))

## [1] 8.95771
```
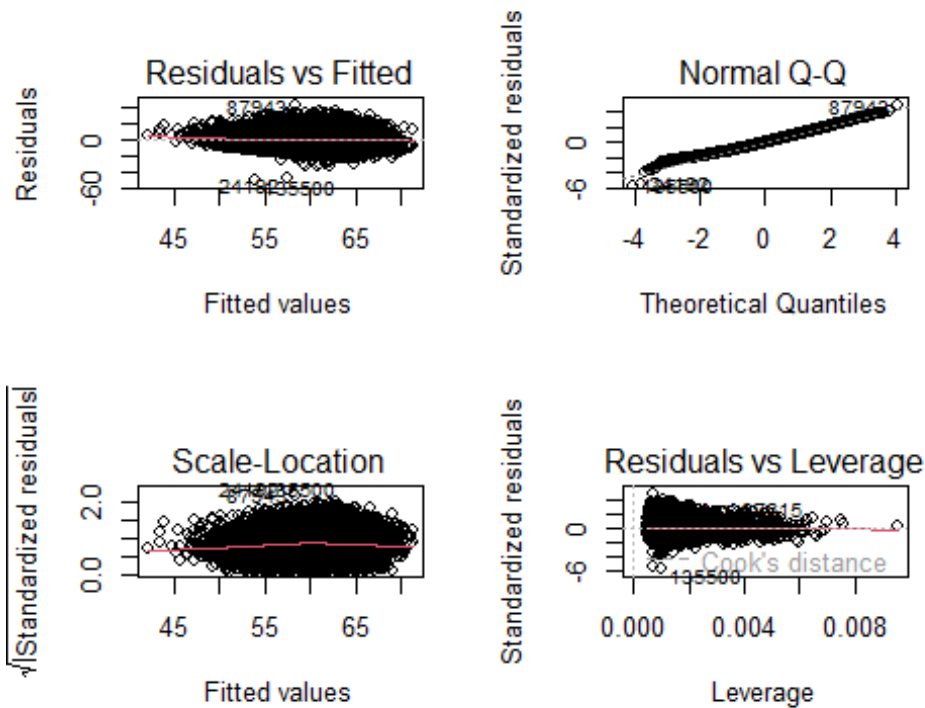
Influential Point Removal 4
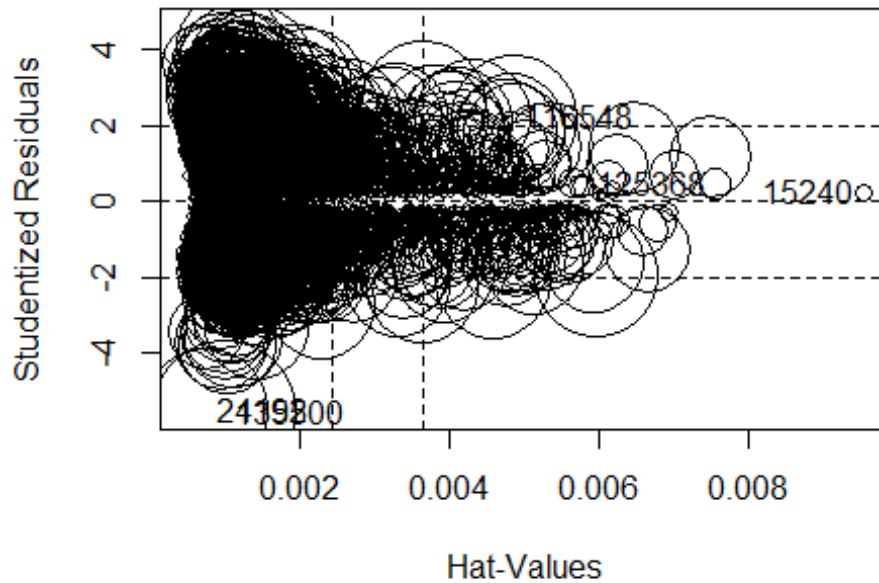

```
##              StudRes            Hat          CookD
## 15240     0.2162256 0.0095300255 1.799481e-05
## 24192    -5.5202979 0.0007543488 9.188927e-04
## 116548    2.2121394 0.0048617835 9.561262e-04
## 125368    0.4397087 0.0075619813 5.893040e-05
## 135500   -5.5869896 0.0009955997 1.242503e-03

summary(modelinflupointremoval5)$r.sq

## [1] 0.1275171

summary(modelinflupointremoval5)$adj.r.sq

## [1] 0.1265016

sqrt(deviance(modelinflupointremoval5)/df.residual(modelinflupointremoval5))

## [1] 8.94954
```
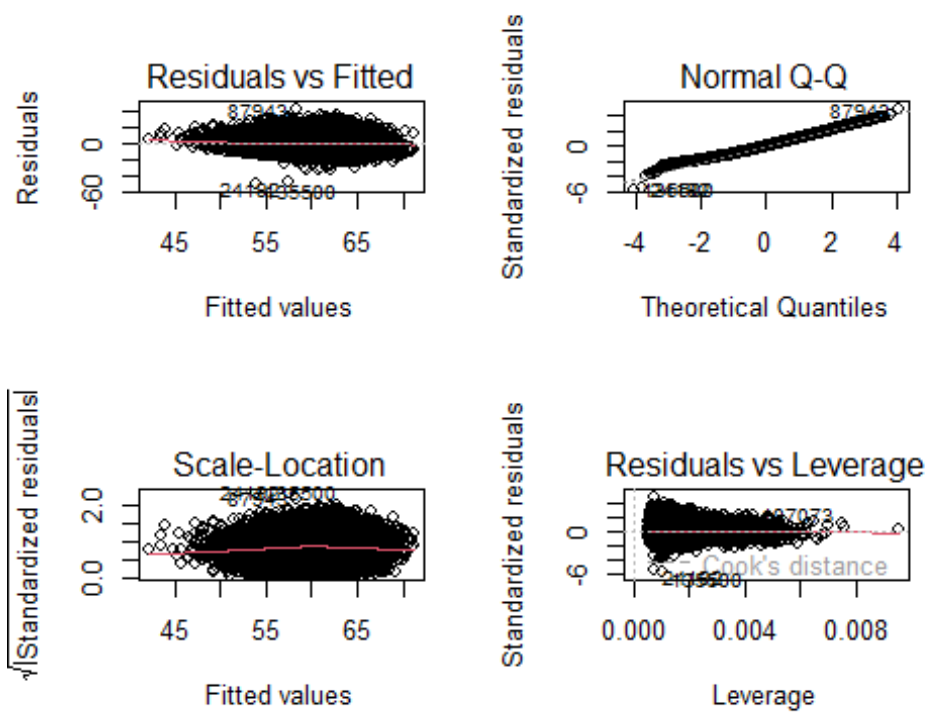
Influential Point Removal 5
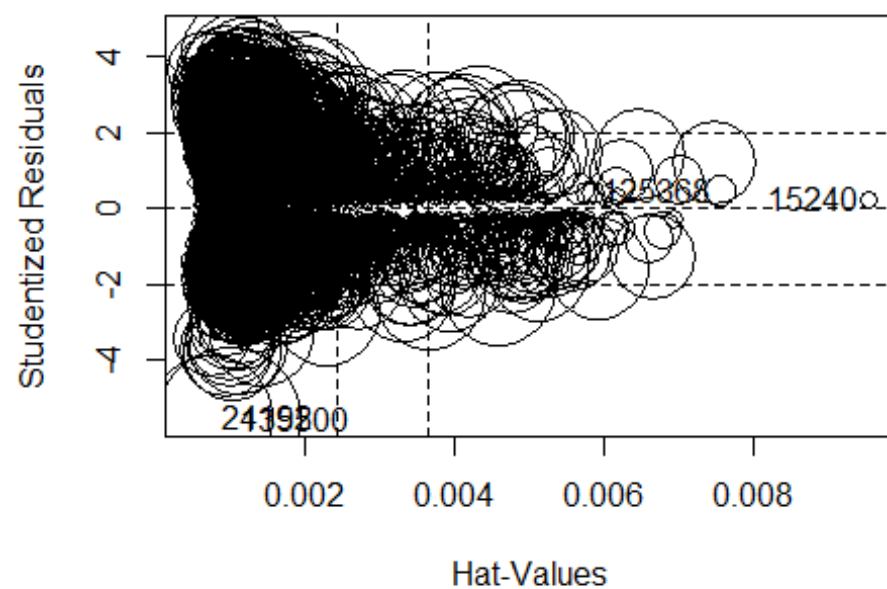
```
##           StudRes            Hat        CookD
## 15240    0.2211684 0.0095345045 1.883584e-05
## 24192   -5.5256043 0.0007543907 9.207083e-04
## 125368   0.4460282 0.0075637785 6.065097e-05
## 135500  -5.5896112 0.0009961606 1.244368e-03
```

```
summary(modelinflupointremoval6)$r.sq
```

```
## [1] 0.1280203
```

```
summary(modelinflupointremoval6)$adj.r.sq
```

```
## [1] 0.127005
```

```
sqrt(deviance(modelinflupointremoval6)/df.residual(modelinflupointremoval6))
```

```
## [1] 8.938277
```
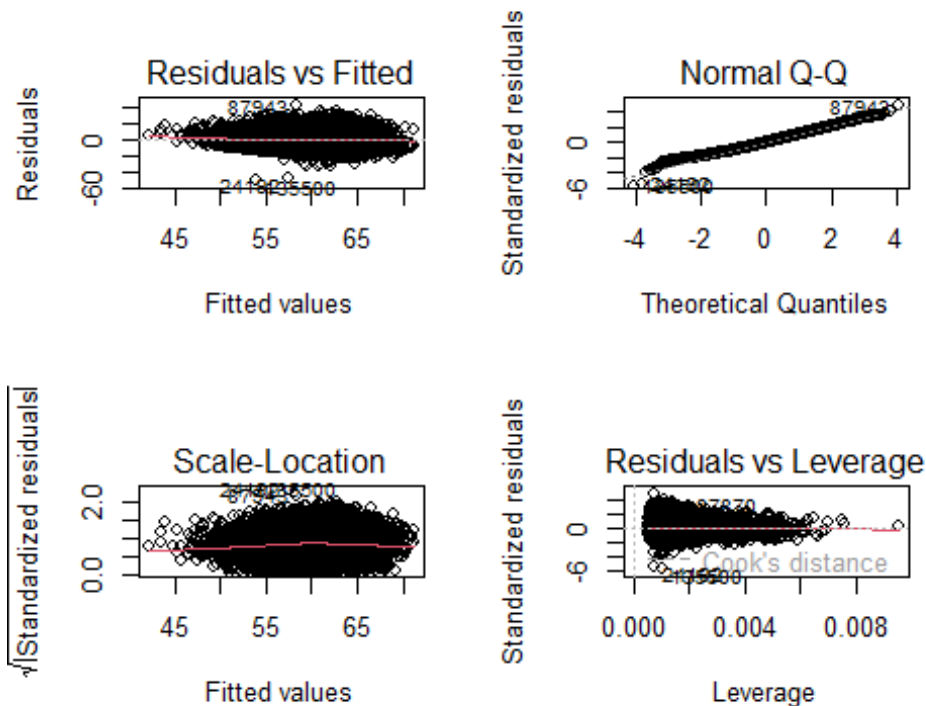

Influential Point Removal Final
```
influencePlot(modelinflupointremoval6)
```
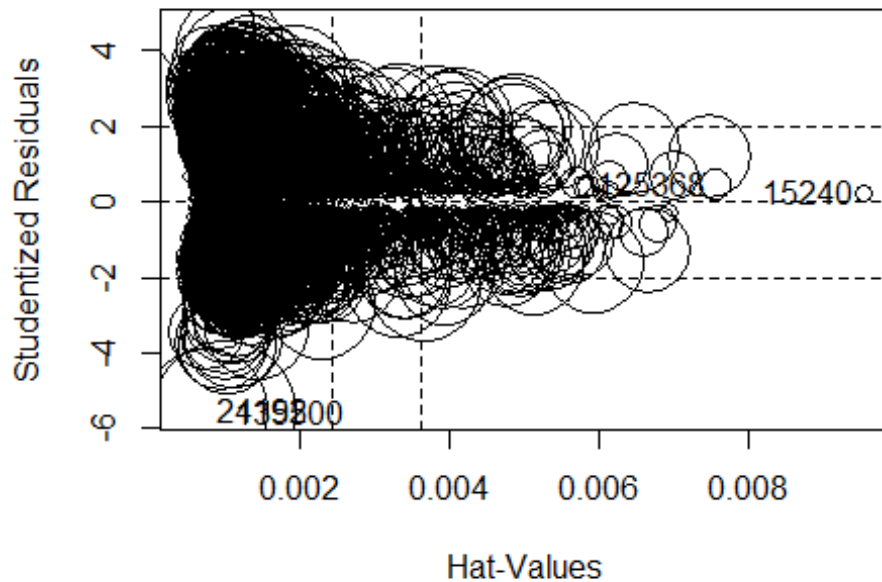
```
##             StudRes          Hat          CookD
## 15240     0.2187591 0.0095462284 1.845057e-05
## 24192    -5.5306708 0.0007545375 9.225742e-04
## 125368    0.4446798 0.0075640670 6.028712e-05
## 135500   -5.5953793 0.0009963496 1.247170e-03
```

```
spotifyremoved <- subset(spotifyinfluentialpointremoval5, select = c(popularity,
    durationmin, tempo, loudness, acousticness, danceability,
    energy, instrumentalness, valence, liveness, speechiness,
    key, mode, explicit, morethan1artist, rownames, name, artists))
```

Data Split into Training and Test Data

Correlation Plot
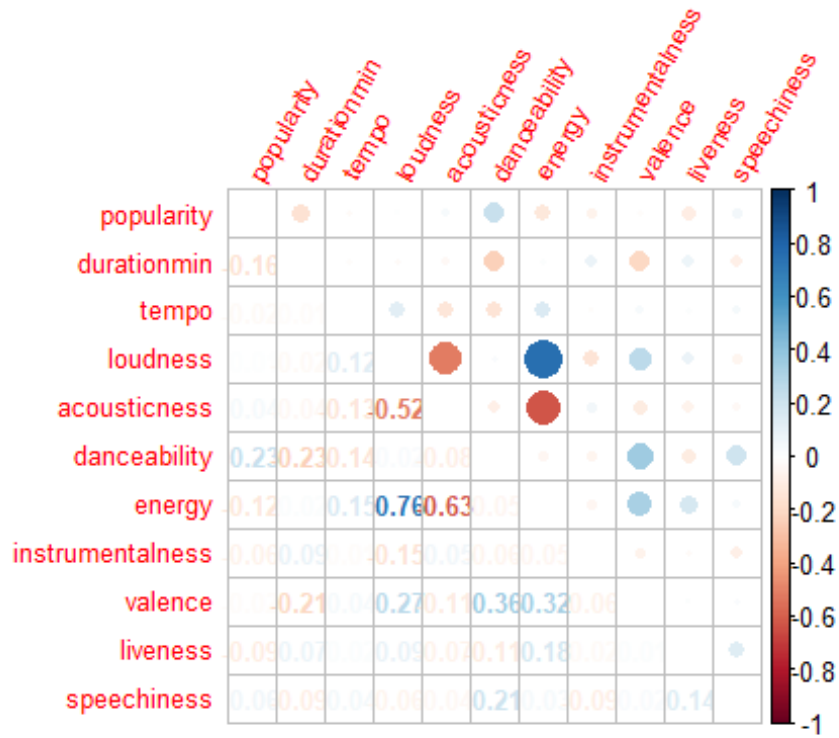
Variance Inflation

```
vif(modelvif1)

##                       GVIF Df GVIF^(1/(2*Df))
## durationmin      1.104773  1        1.051082
## tempo            1.063576  1        1.031298
## acousticness     1.808755  1        1.344900
## danceability     1.556882  1        1.247751
## energy           3.415914  1        1.848219
## loudness         2.544722  1        1.595219
## instrumentalness 1.055445  1        1.027348
## valence          1.491261  1        1.221172
## liveness         1.081486  1        1.039945
## speechiness      1.400466  1        1.183413
## key              1.125561 11        1.005391
## mode             1.097268  1        1.047506
## explicit         1.494141  1        1.222351
## morethan1artist  1.103608  1        1.050527

summary(modelvif1)$r.sq

## [1] 0.1290078
```

```
## removal of energy

##                   GVIF Df GVIF^(1/(2*Df))
## durationmin      1.100572  1        1.049081
## tempo            1.060095  1        1.029609
## acousticness     1.434333  1        1.197636
## danceability     1.479729  1        1.216441
## loudness         1.530232  1        1.237025
## instrumentalness 1.041172  1        1.020378
## valence          1.365750  1        1.168653
## liveness         1.055882  1        1.027561
## speechiness      1.368939  1        1.170017
## key              1.124264 11        1.005338
## mode             1.096222  1        1.047006
## explicit         1.497896  1        1.223886
## morethan1artist  1.101773  1        1.049654

## [1] 0.1138145
```

Interaction Plots

Durationmin

key

There is some interaction between durationmin and keys that needs to be accounted for
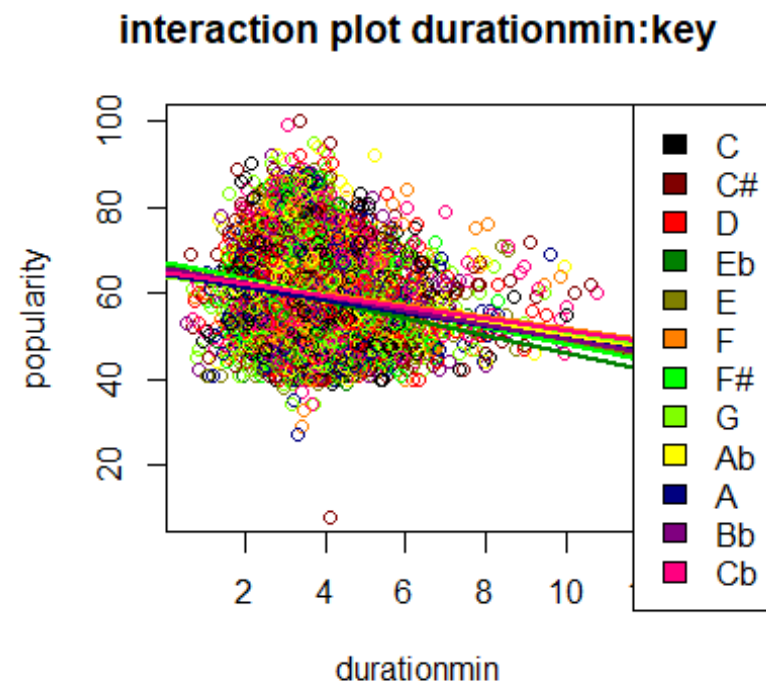
mode

There is an interaction between mode and durationmin

interaction plot durationmin:mode
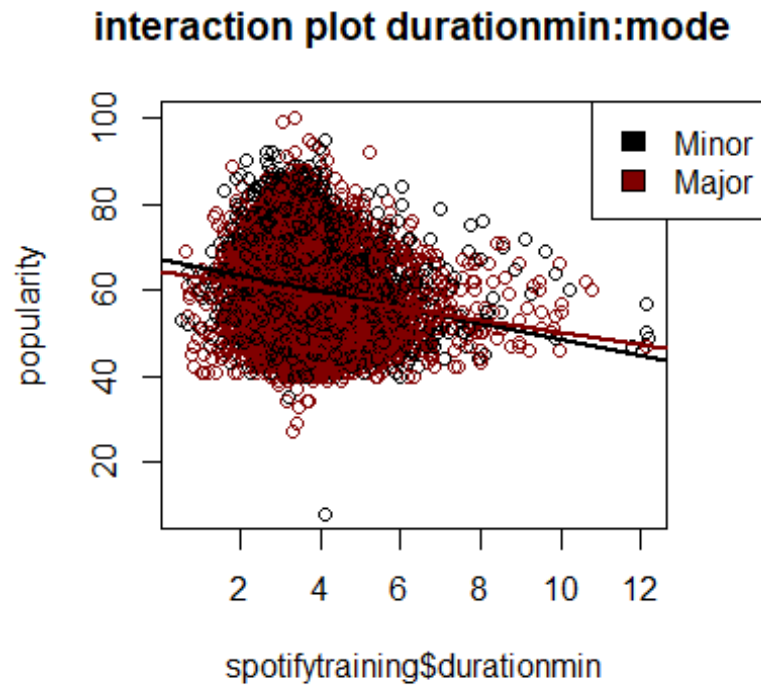
explicit

There is an interaction between durationmin and explicit

interaction plot durationmin:explicit
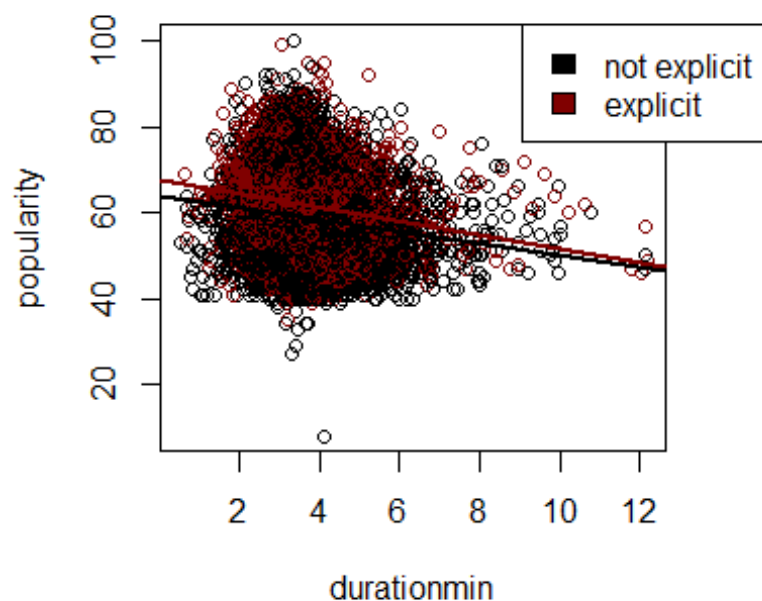
morethan1artist

There is an interaction between duration and more than 1 artist

nteraction plot durationmin:morethan1artist
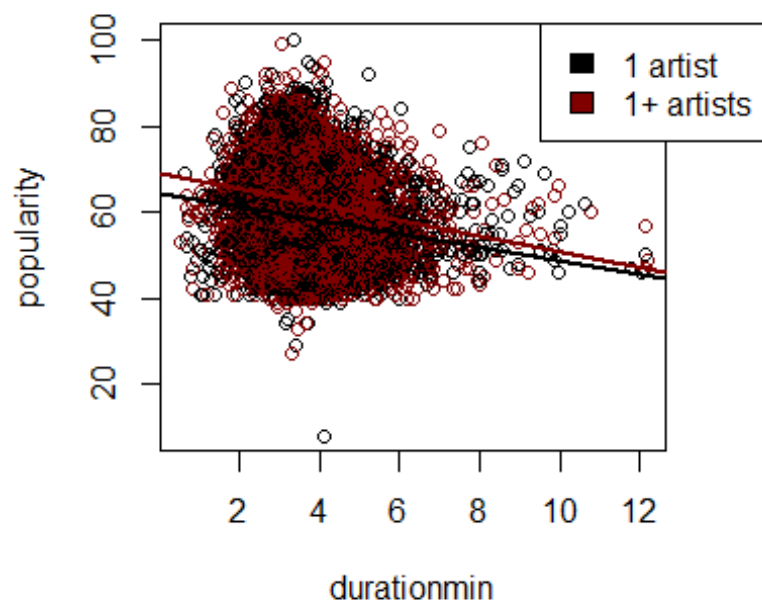
almost parallel lines but will still include in full model
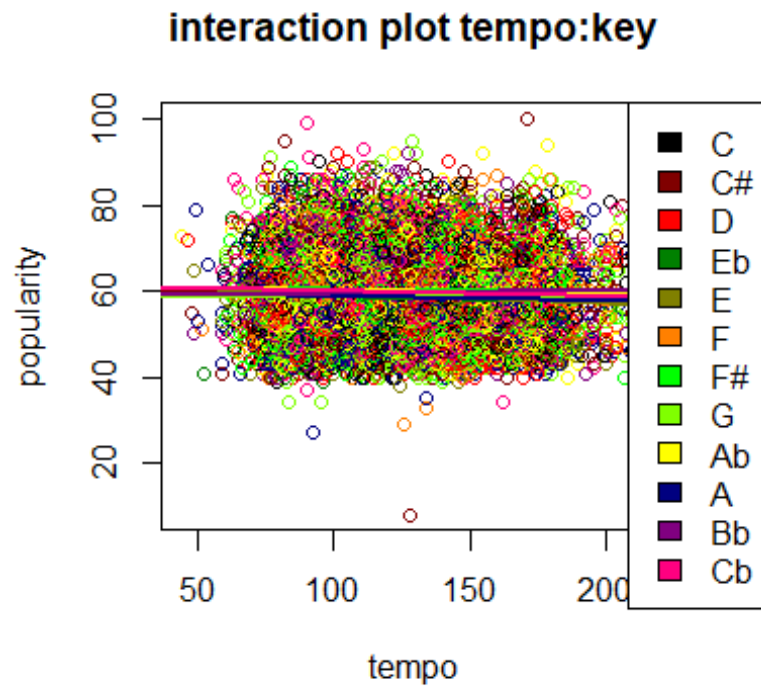
no interaction

interaction plot tempo:mode
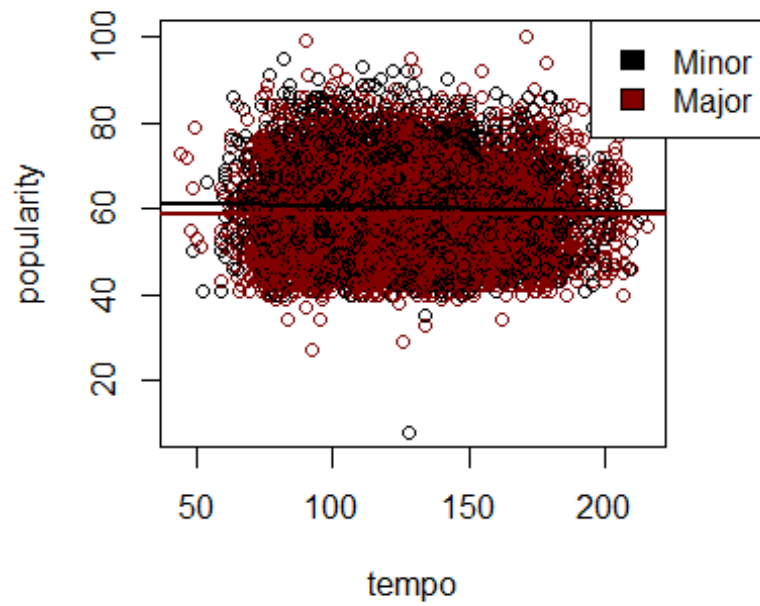
explicit

There is an interaction between tempo and explicit

interaction plot tempo:explicit

small interaction between tempo and morethan1artist

interaction plot tempo:morethan1artist
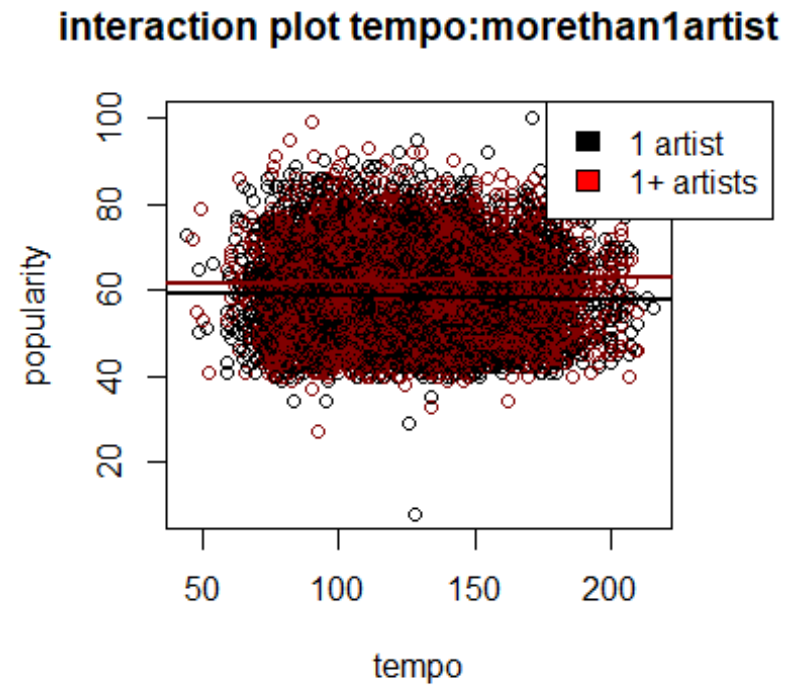
There is an interaction with Key and loudness

interaction plot loudness:key
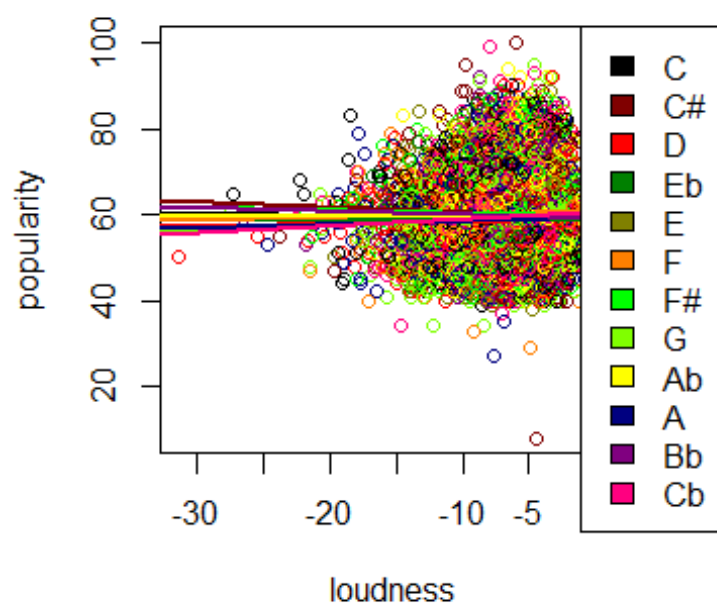
mode

No interaction between mode and loudness

interaction plot loudness:mode
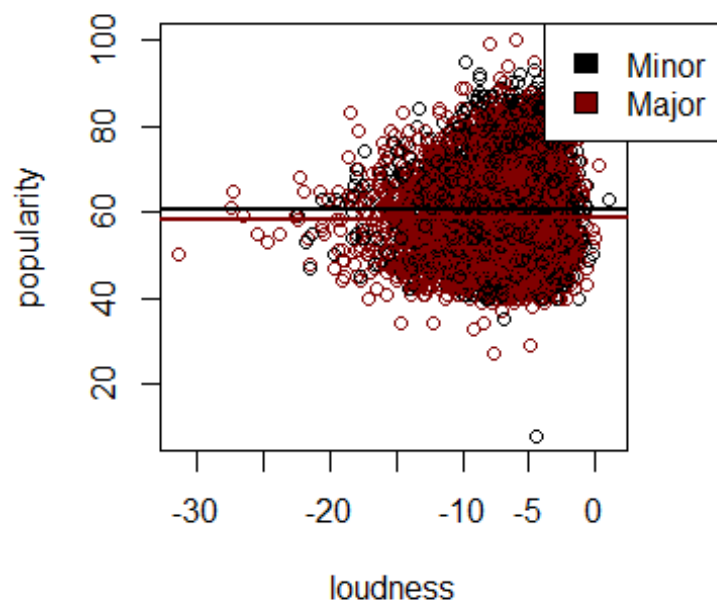
explicit

small interaction between explicit and loudness

interaction plot loudness:explicit
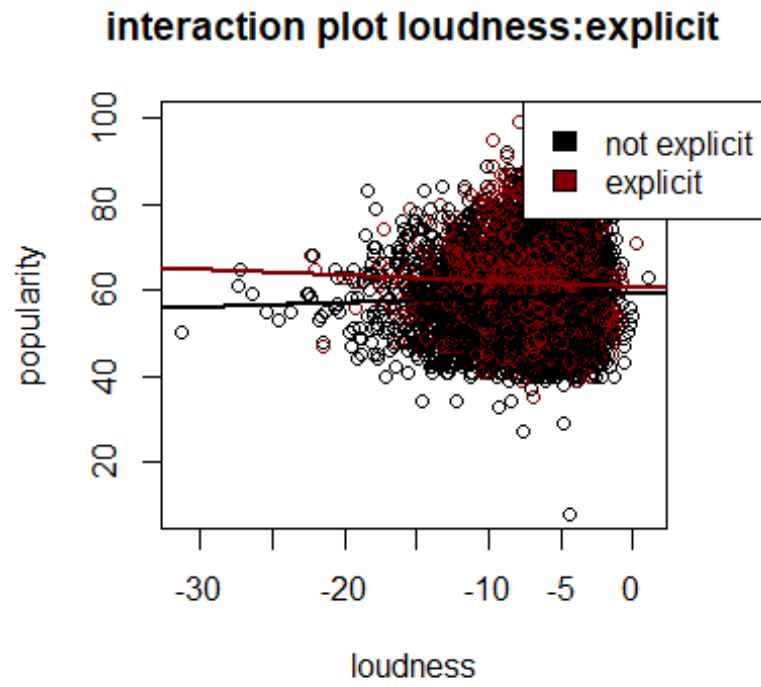
morethan1artist

small interaction between morethan1artist and loudness

interaction plot loudness:morethan1artist
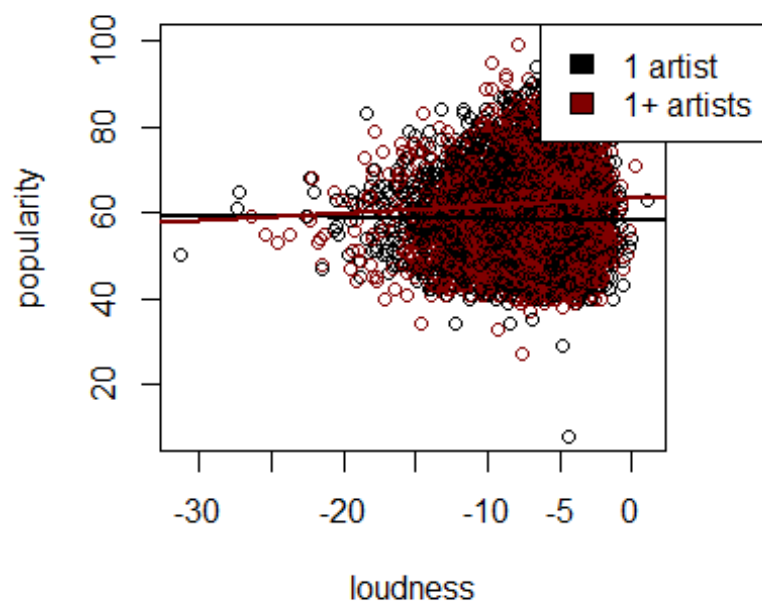
Acousticness

key

small interaction between acousticness and key

interaction plot acousticness:key
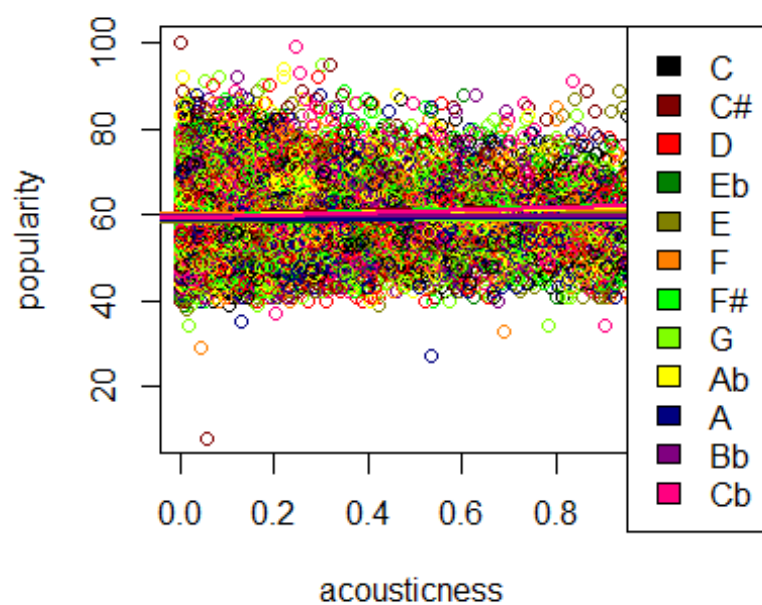
mode

small interaction between mode and acousticness

interaction plot acousticness:mode
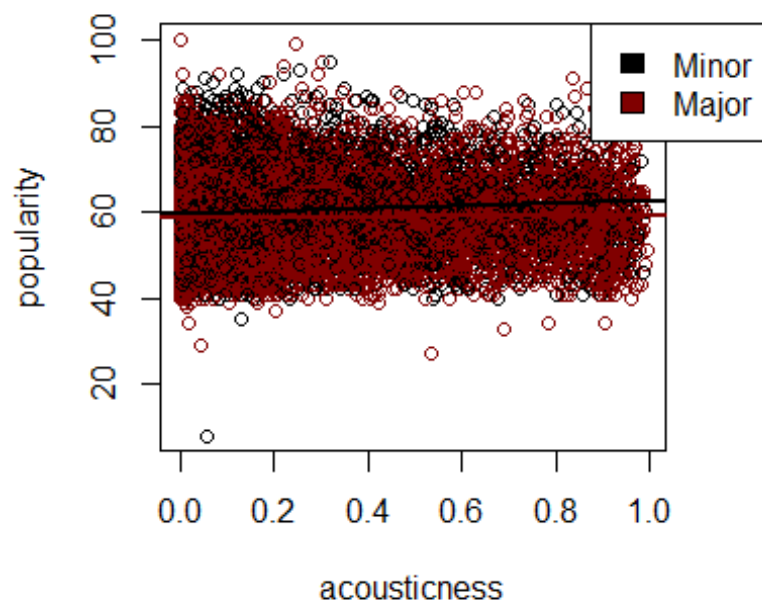
parallel lines so no interaction between acousticness and explicit
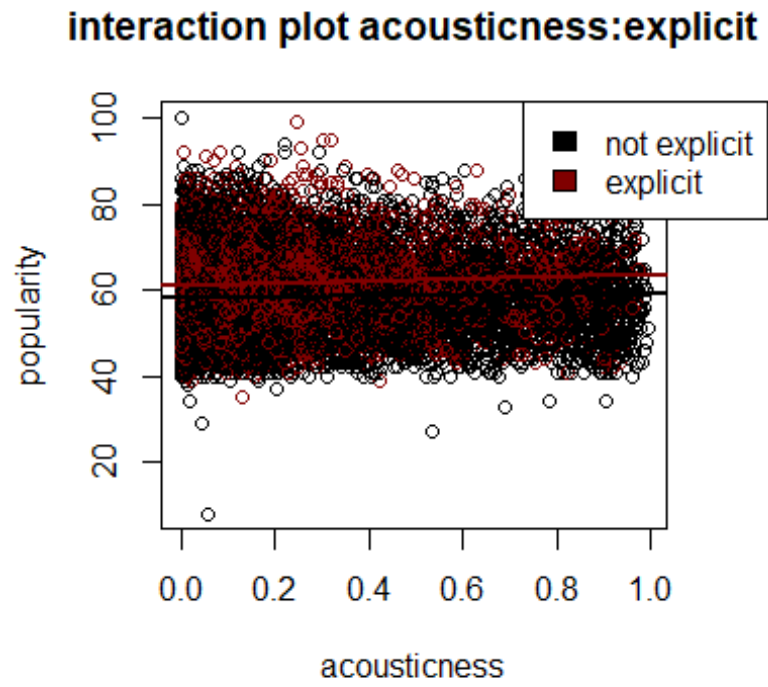
interaction plot acousticness:explicit

almost no interaction or difference will not include in full model

nteraction plot acousticness:morethan1artist
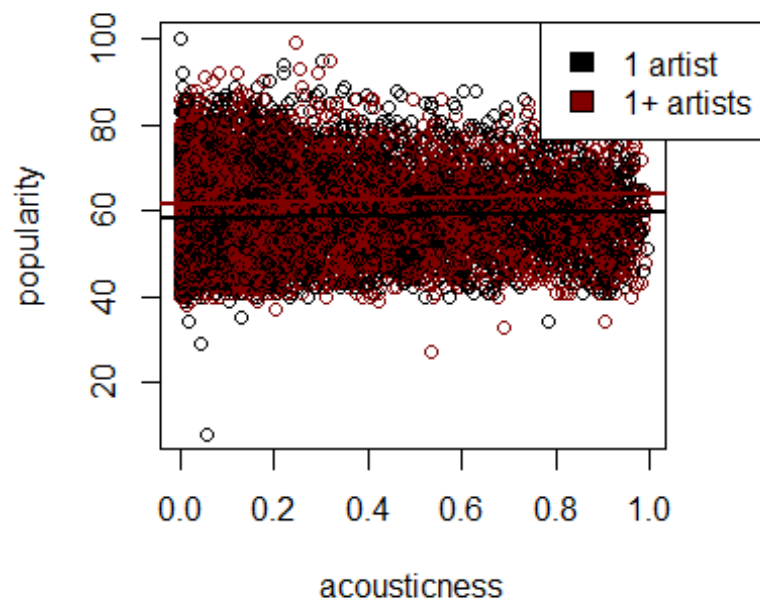
Danceability

key

There is an interaction between danceability and key

interaction plot danceability:key
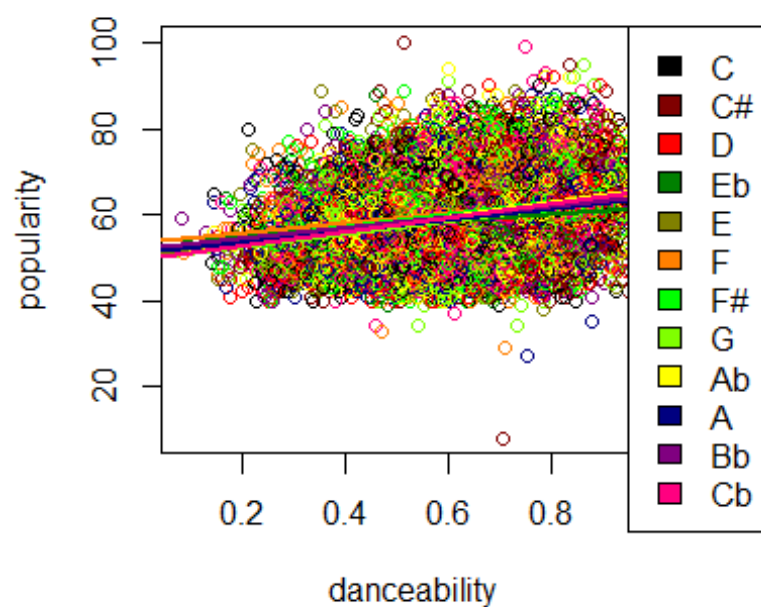
mode

Interaction is present

interaction plot danceability:mode
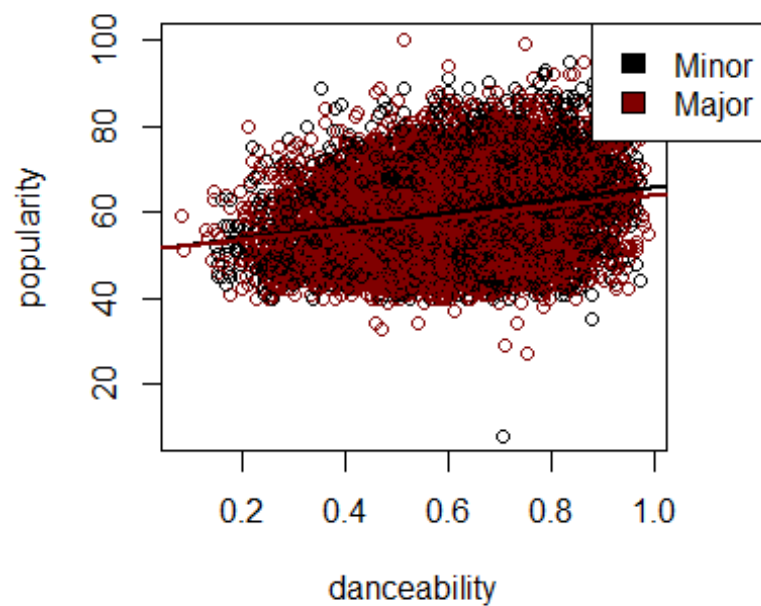
parallel lines so no interactions

interaction plot danceability:explicit

small interaction to include

nteraction plot danceability:morethan1artist
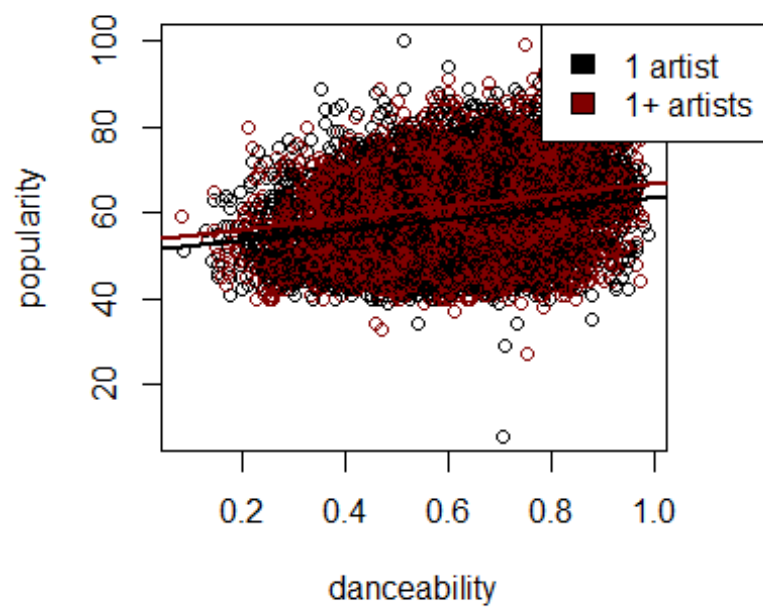
instrumentalness

key

interaction present between instrumentalness and key

interaction plot instrumentalness:key
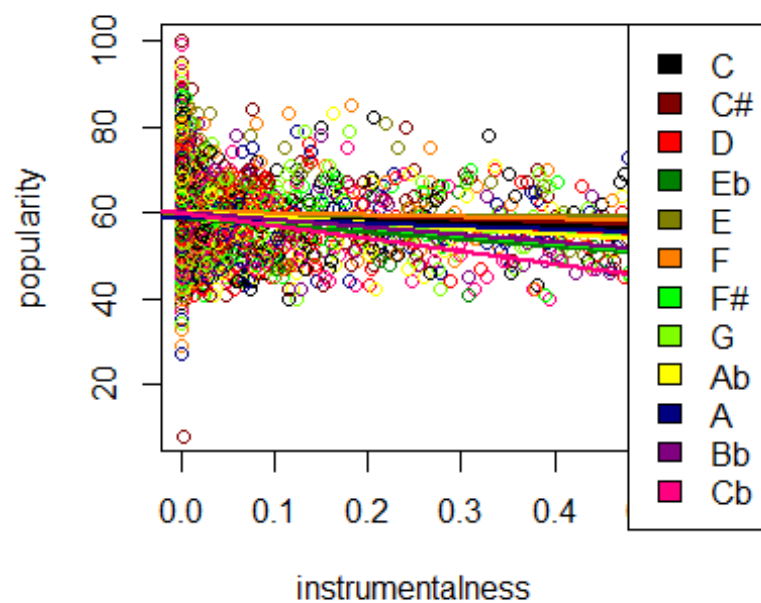
mode

small interaction present
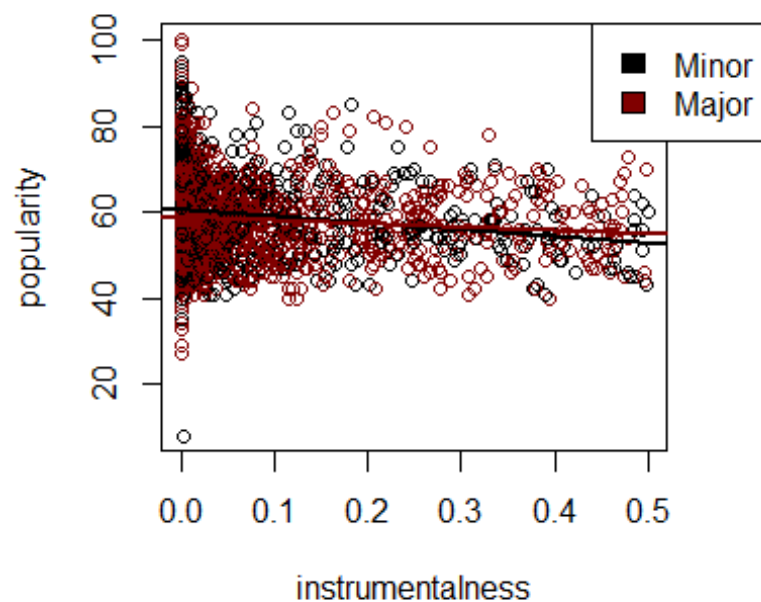
interaction plot instrumentalness:mode

interaction present

interaction plot instrumentalness:explicit
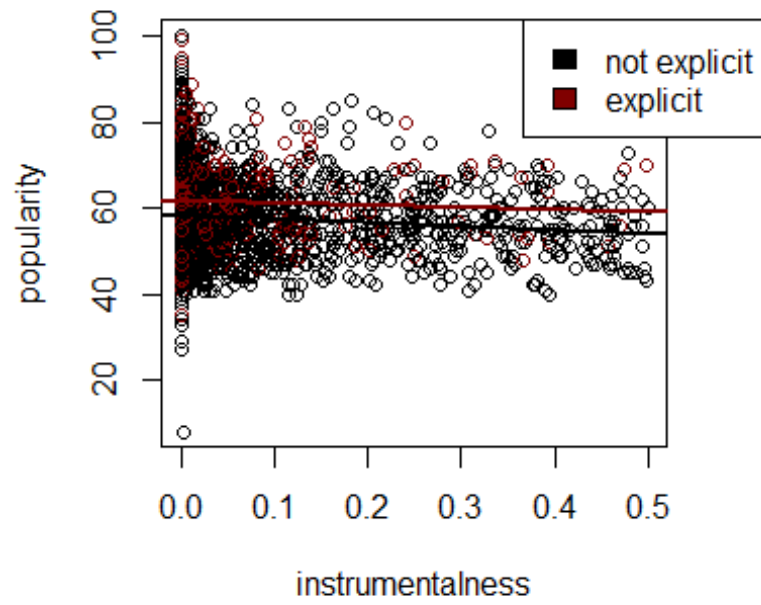
interaction is present

:raction plot instrumentalness:morethan1artist
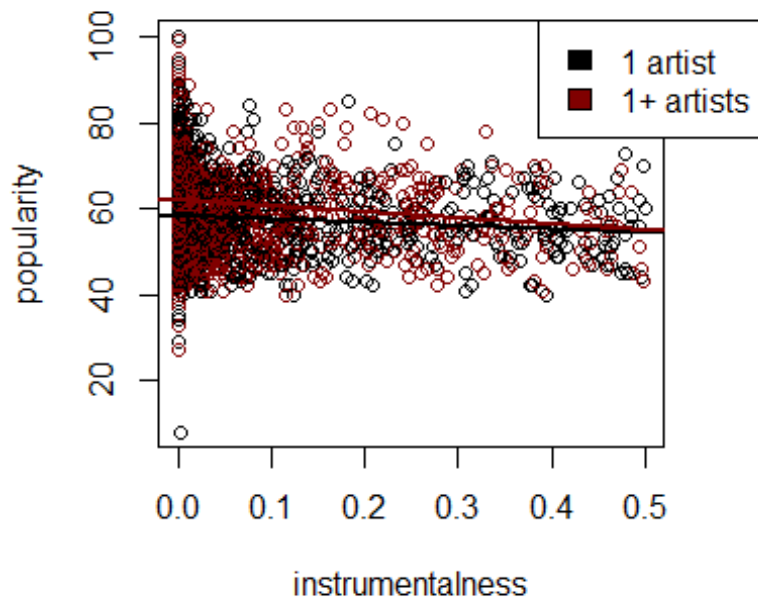
valence

key

some interactions present

interaction plot valence:key
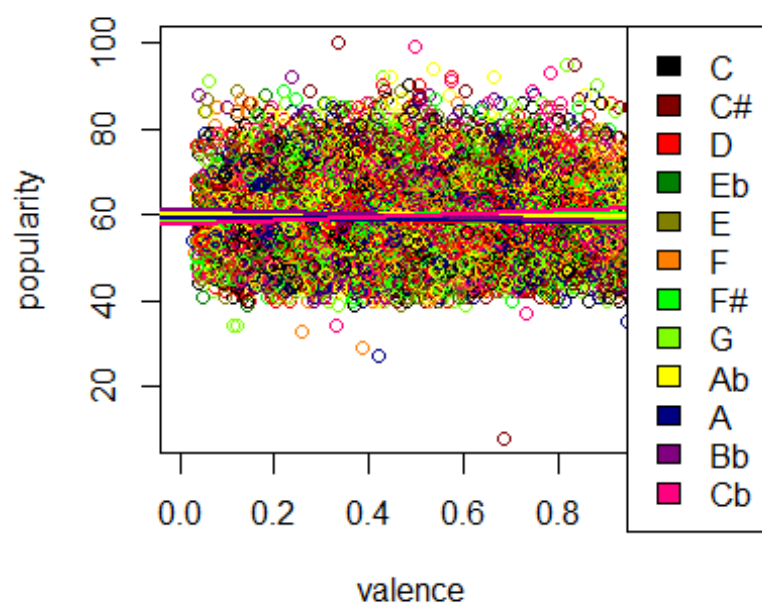
mode

some interaction present

interaction plot valence:mode
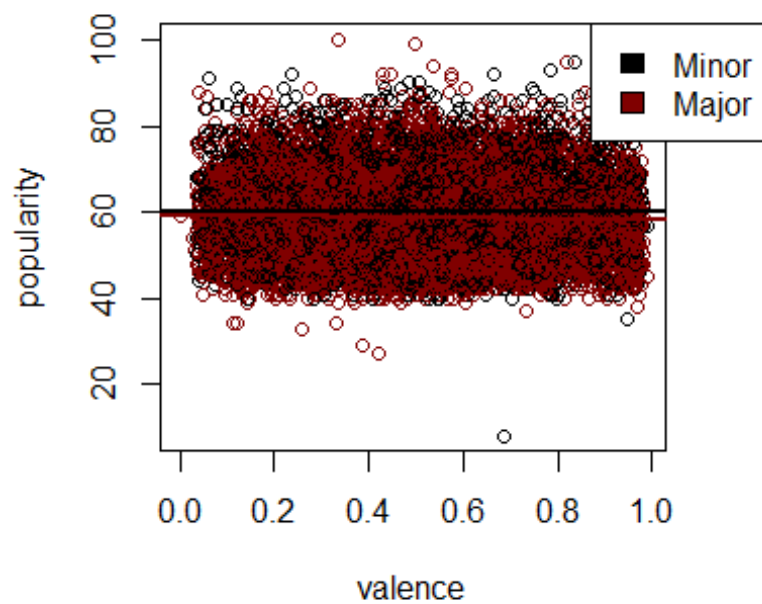
no interaction present
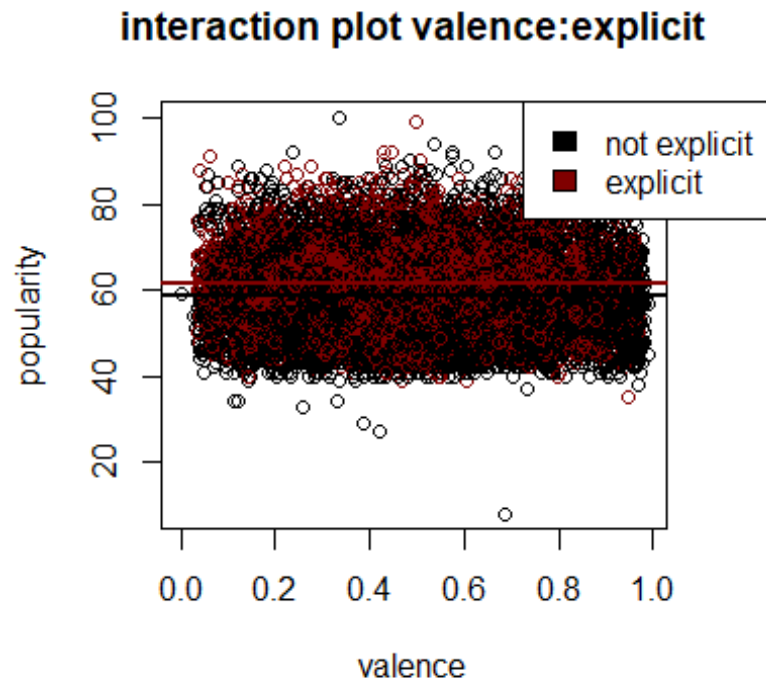
interaction plot valence:explicit

very small interaction

interaction plot valence:morethan1artist
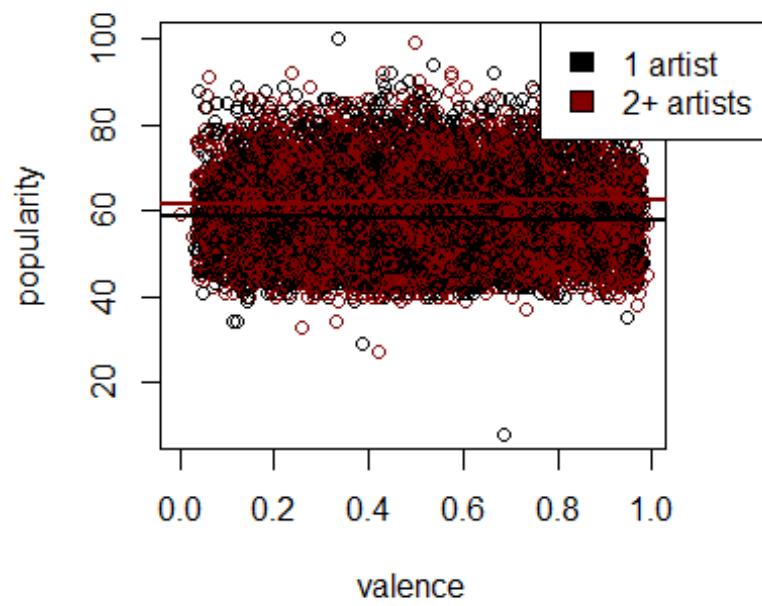
liveness

key

small interaction present

interaction plot liveness:key
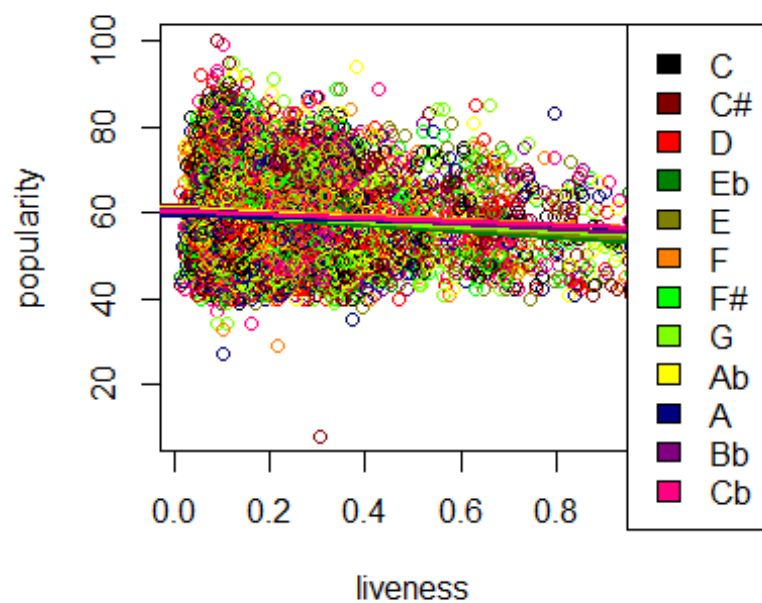
mode

small interaction present

interaction plot liveness:mode

small interaction present

interaction plot liveness:explicit

small interaction present

interaction plot liveness:morethan1artist
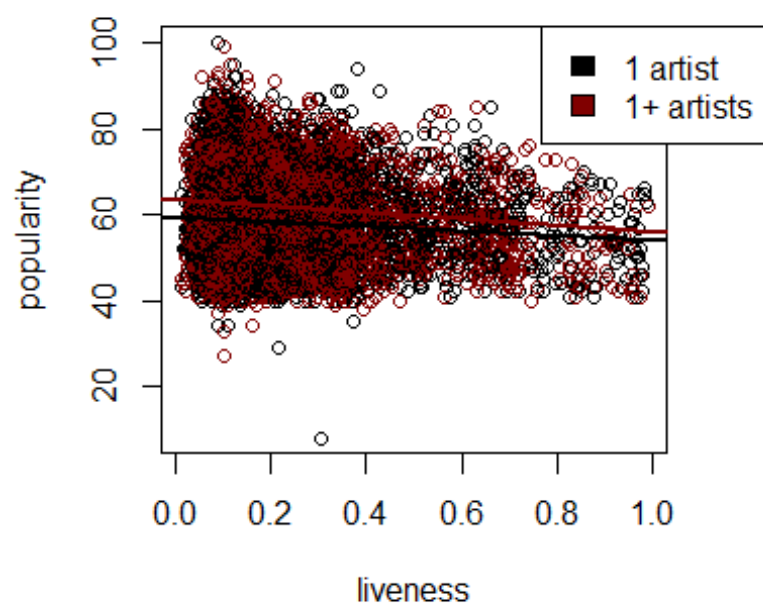
speechiness

key

Definite interaction present

interaction plot speechiness:key
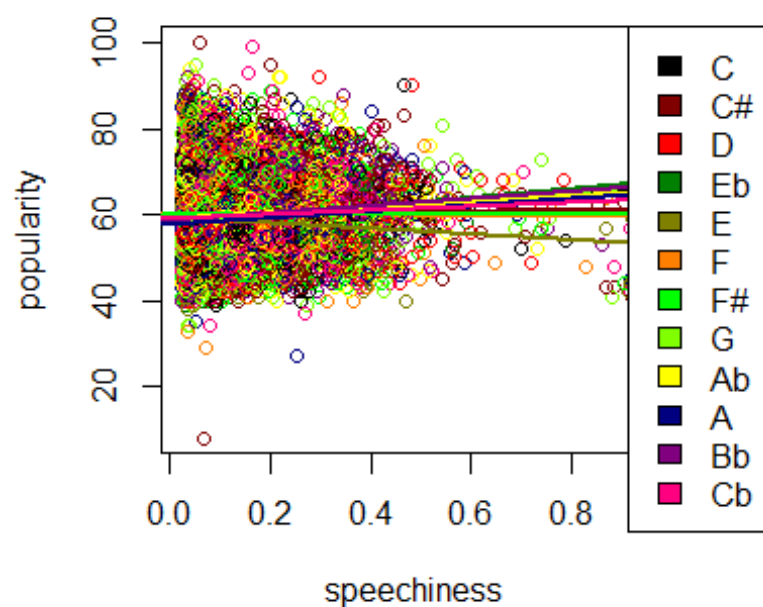
mode

interaction is present

interaction plot speechiness:mode
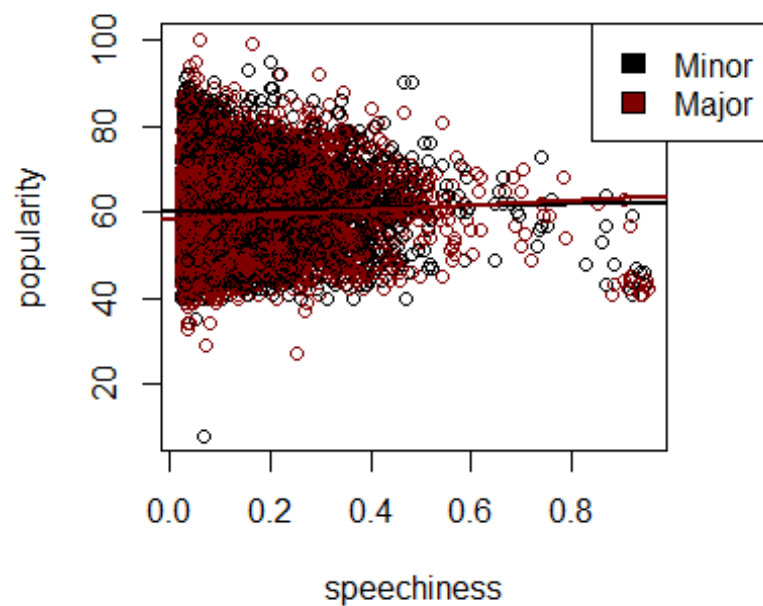
interaction is present

interaction plot speechiness:explicit
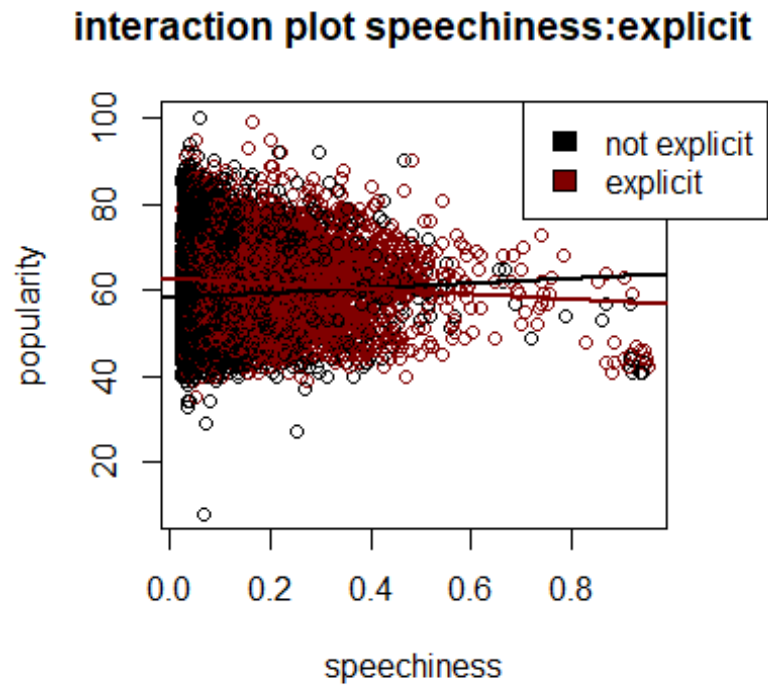
interaction is present

Interaction plot speechiness:morethan1artist
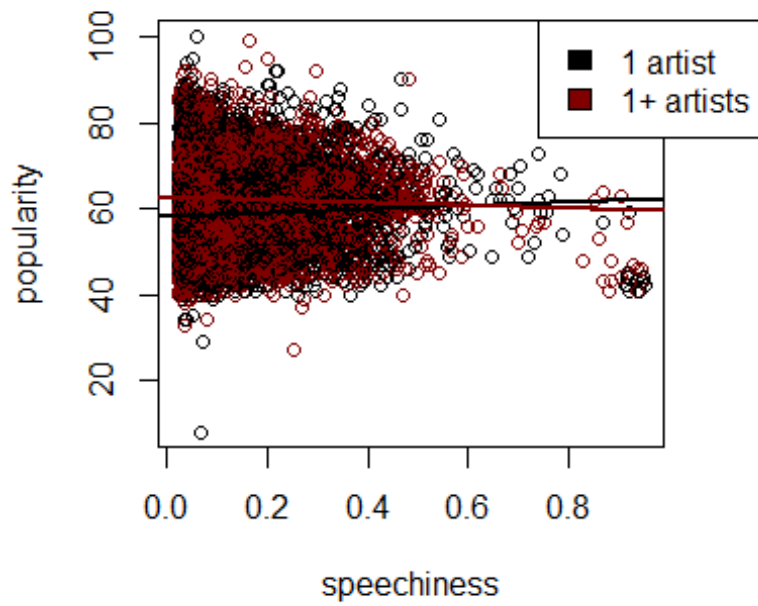
Interaction Terms

Full Interaction Start

```
## [1] 0.1299882

## [1] 0.1206791

## [1] 8.943591
```
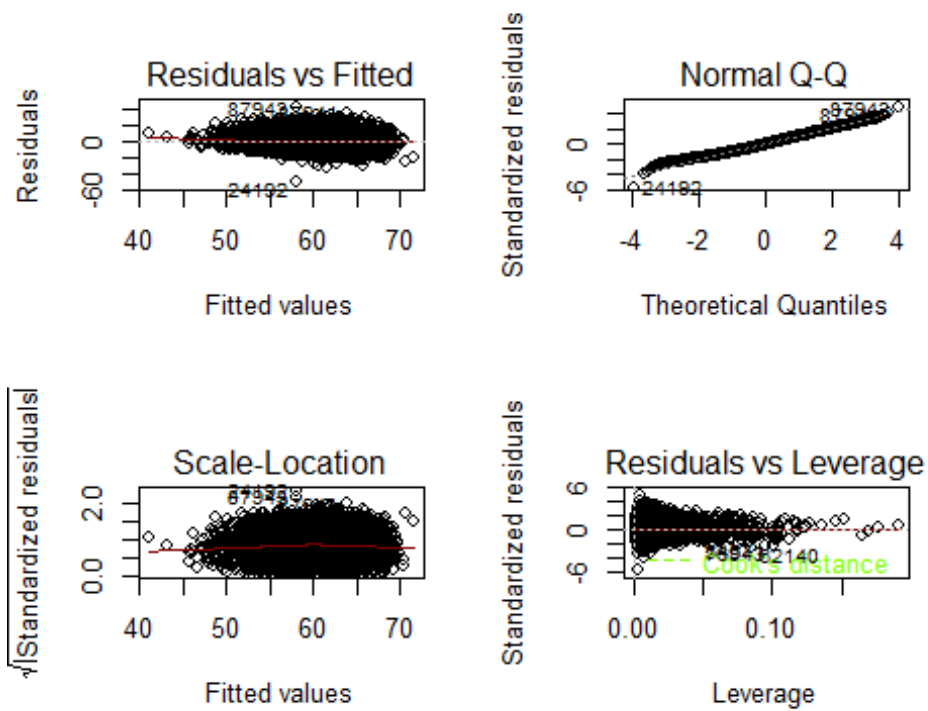
Interaction Term Removal 1

removal of speechiness:mode

```
## [1] 0.1299819
```

```
## [1] 0.1207389
```

```
## [1] 8.943287
```

Interaction Term Removal 2

removal of danceability:mode

```
## [1] 0.1299567
## [1] 0.1207797
## [1] 8.94308
```
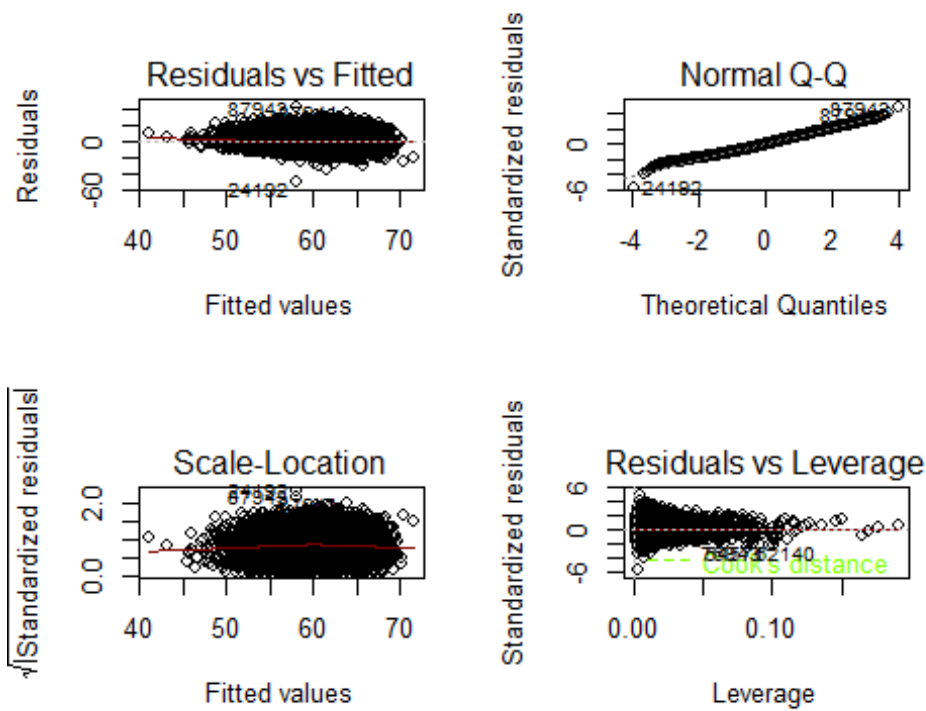
Interaction Term Removal 3

removal of instrumental:explicit

```
## [1] 0.1299565
## [1] 0.1208457
## [1] 8.942744
```
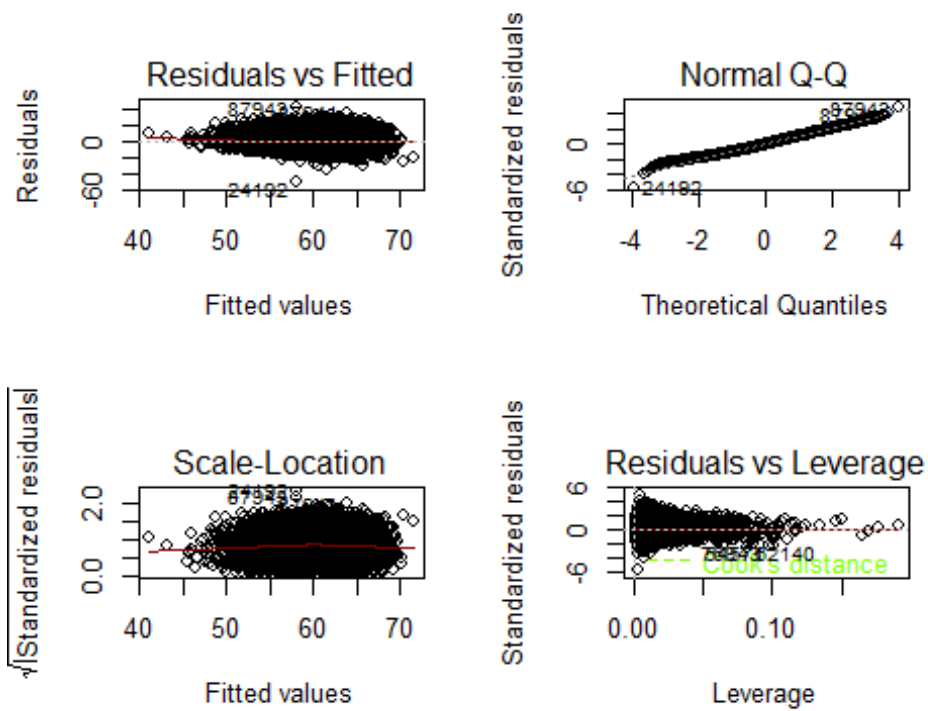
Interaction Term Removal 4

removal of liveness:mode

```
## [1] 0.1299451
```

```
## [1] 0.1209005
```

```
## [1] 8.942466
```
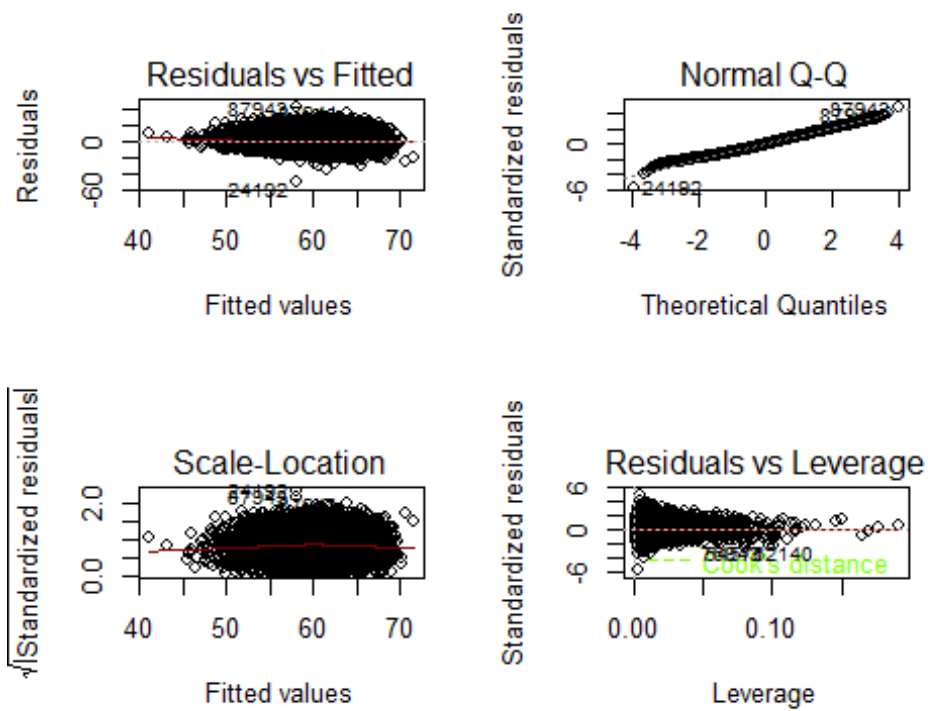
Interaction Term Removal 5

removal of liveness:explicit

```
## [1] 0.1299415
```

```
## [1] 0.120963
```

```
## [1] 8.942148
```

Interaction Term Removal 6

Removal of loudness:morethan1artist

```
## [1] 0.129909
```

```
## [1] 0.1209964
```

```
## [1] 8.941978
```

Interaction Term Removal 7

Removal of instrumentalness:mode

```
## [1] 0.129871
```

```
## [1] 0.1210242
```

```
## [1] 8.941836
```
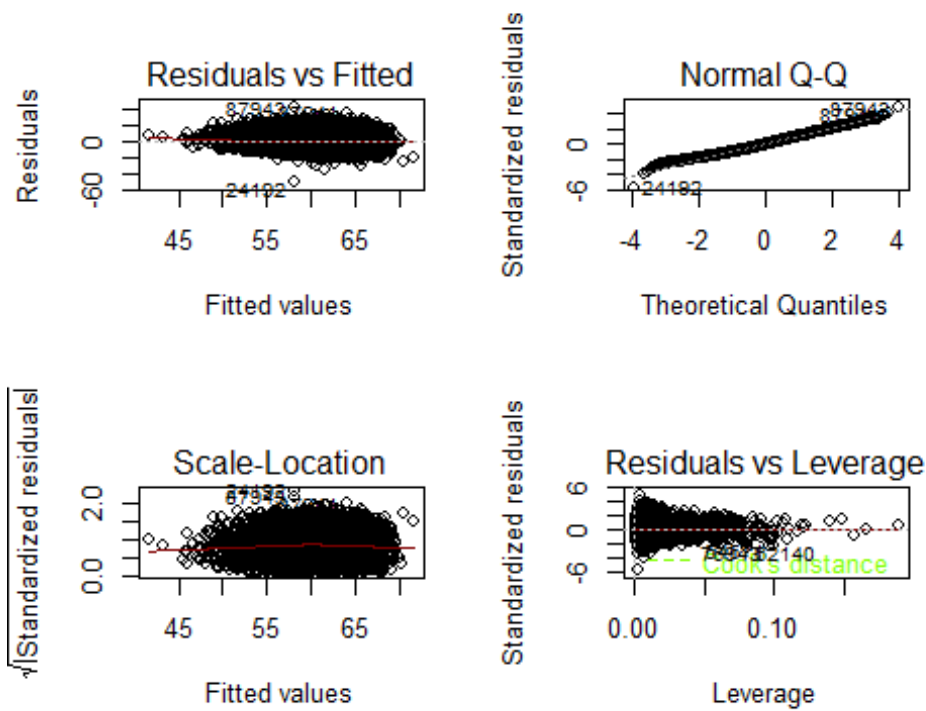
Interaction Term Removal 8

Removal of spotifytraining$duration:morethan1artist R.sq:0.1275 and adj.R.sq:0.1187 and RSE:8.98

```
## [1] 0.1298477
```

```
## [1] 0.1210669
```

```
## [1] 8.941619
```
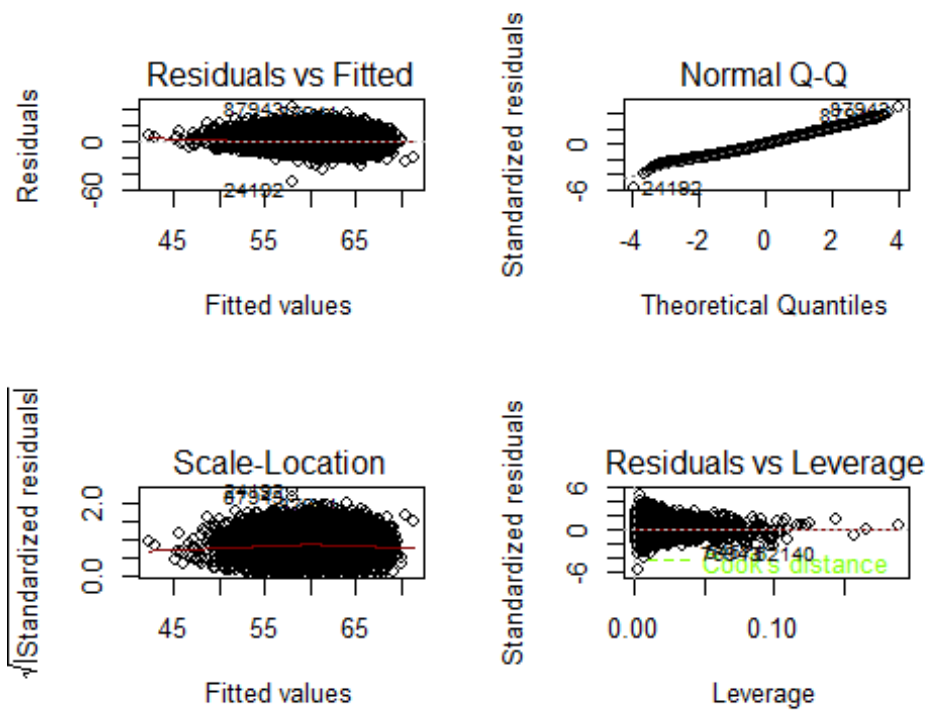
Interaction Term Removal 9

Removal of liveness:key R.sq:0.1272 and adj.R.sq:0.1191 and RSE:8.978

```
summary(modelinteraction9)$r.sq   ## 0.1311 and 0.123 and RSE:8.862

## [1] 0.1295853

plot(modelinteraction9)
```
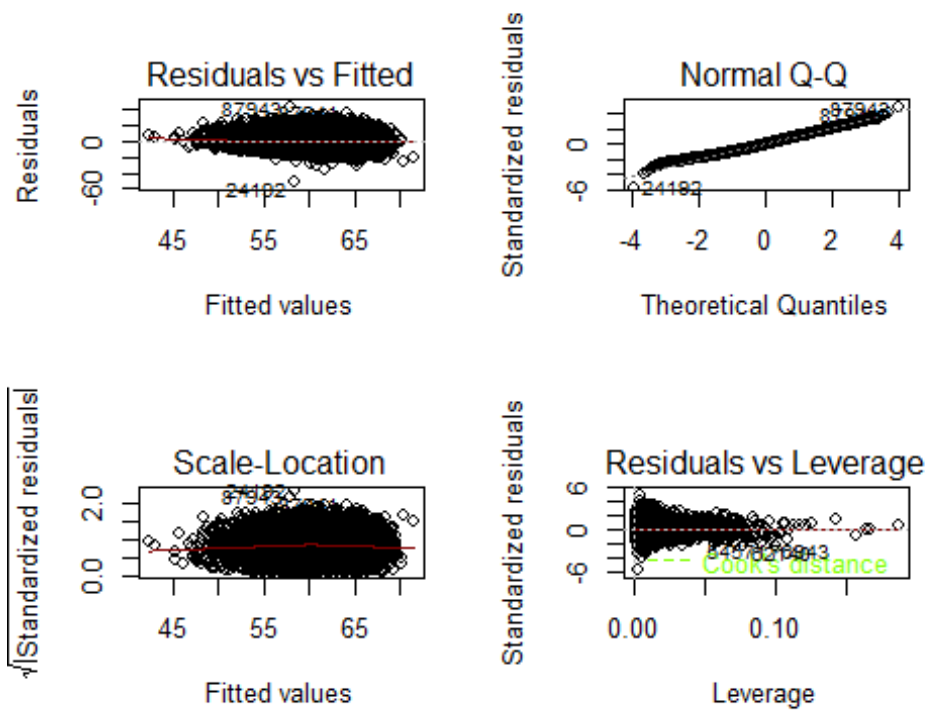
Interaction Term Removal 10

Removal of instrumentalness:morethan1artist1 R.sq:0.1271 and adj.R.sq:0.1191 and RSE:8.978

```
## [1] 0.1294975
```
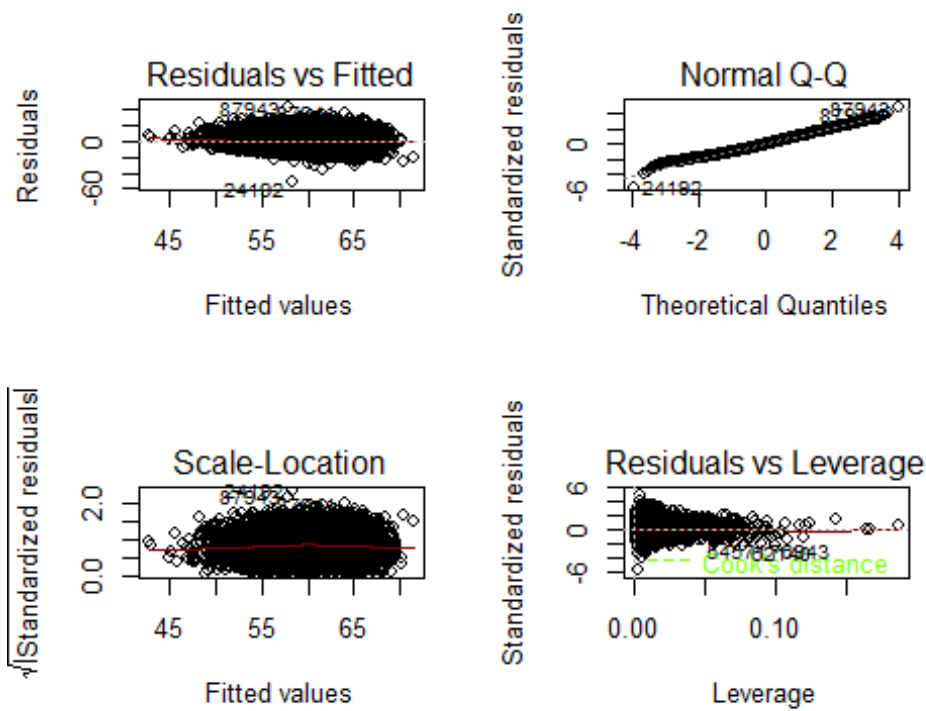
Transformations ##### Popularity R.sq:0.1264 and adj.R.sq:0.1184 and RSE:0.5814

```
## [1] 0.1288691
```
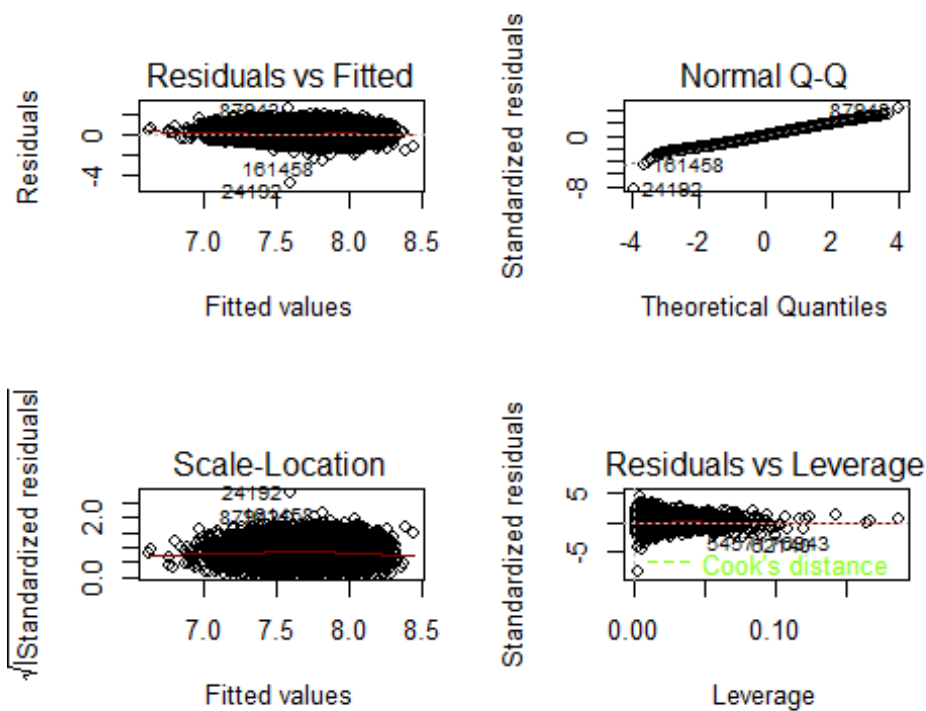
Instrumentalness

R-squared: 0.1268, Adjusted R-squared: 0.1188, RSE: 0.5813

```
## [1] 0.1292437
```
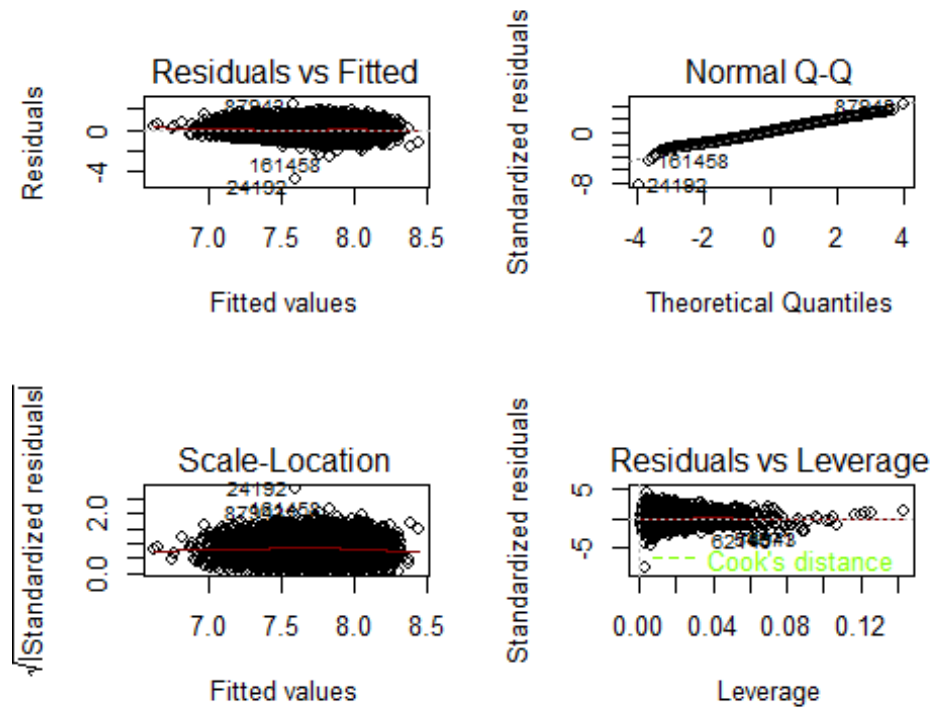
Variable Selection

Forward

AIC
```
## Start:  AIC=-12929.48
## sqrt(popularity) ~ 1
##
##                              Df Sum of Sq    RSS    AIC
## + danceability               1    267.312 4848.2 -13647
## + morethan1artist            1    148.007 4967.5 -13321
## + spotifytraining$duration   1    124.293 4991.2 -13257
## + explicit                   1    122.051 4993.4 -13251
## + liveness                   1     44.852 5070.6 -13046
## + sqrt(instrumentalness)     1     28.796 5086.7 -13003
## + mode                       1     25.791 5089.7 -12995
## + key                       11     24.933 5090.5 -12973
## + speechiness                1     17.124 5098.4 -12972
## + acousticness               1      6.600 5108.9 -12945
## + tempo                      1      2.029 5113.4 -12933
## + valence                    1      2.024 5113.5 -12933
```

```
## <none>                                       5115.5 -12930
## + loudness                        1      0.217 5115.3 -12928
##
## Step:  AIC=-13647.41
## sqrt(popularity) ~ danceability
##
##                             Df Sum of Sq    RSS    AIC
## + morethan1artist            1     85.991 4762.2 -13886
## + valence                    1     60.918 4787.2 -13815
## + spotifytraining$duration   1     57.359 4790.8 -13805
## + explicit                   1     40.909 4807.3 -13759
## + liveness                   1     24.900 4823.3 -13714
## + sqrt(instrumentalness)     1     16.859 4831.3 -13692
## + acousticness               1     15.150 4833.0 -13687
## + mode                       1     14.658 4833.5 -13686
## + key                       11     14.380 4833.8 -13665
## + tempo                      1      0.805 4847.4 -13648
## <none>                                     4848.2 -13647
## + speechiness                1      0.548 4847.6 -13647
## + loudness                   1      0.005 4848.2 -13645
##
## Step:  AIC=-13885.47
## sqrt(popularity) ~ danceability + morethan1artist
##
##                             Df Sum of Sq    RSS    AIC
## + spotifytraining$duration   1     68.756 4693.4 -14079
## + valence                    1     49.925 4712.2 -14025
## + liveness                   1     30.385 4731.8 -13969
## + explicit                   1     22.769 4739.4 -13948
## + acousticness               1     20.729 4741.4 -13942
## + sqrt(instrumentalness)     1     11.846 4750.3 -13917
## + mode                       1     10.381 4751.8 -13913
## + key                       11     12.368 4749.8 -13898
## <none>                                     4762.2 -13886
## + tempo                      1      0.628 4761.5 -13885
## + speechiness                1      0.554 4761.6 -13885
## + loudness                   1      0.165 4762.0 -13884
##
## Step:  AIC=-14078.55
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
## on
##
##                             Df Sum of Sq    RSS    AIC
## + valence                    1     67.893 4625.5 -14272
## + liveness                   1     26.484 4666.9 -14152
## + explicit                   1     20.981 4672.4 -14137
## + acousticness               1     17.077 4676.3 -14125
## + mode                       1     11.266 4682.2 -14109
## + sqrt(instrumentalness)     1      7.059 4686.4 -14097
## + key                       11     12.133 4681.3 -14091
```

```
## + speechiness                    1       1.379 4692.0 -14080
## <none>                                         4693.4 -14079
## + loudness                       1       0.340 4693.1 -14078
## + tempo                          1       0.146 4693.3 -14077
##
## Step:   AIC=-14272.01
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence
##
##                              Df Sum of Sq    RSS    AIC
## + liveness                    1   21.3046 4604.2 -14332
## + acousticness                1   11.2566 4614.3 -14303
## + mode                        1   10.0608 4615.5 -14299
## + sqrt(instrumentalness)      1    9.8758 4615.6 -14299
## + explicit                    1    8.4357 4617.1 -14294
## + valence:morethan1artist     1    7.0808 4618.4 -14291
## + key                        11   12.5727 4613.0 -14286
## + loudness                    1    3.2423 4622.3 -14279
## + speechiness                 1    2.6087 4622.9 -14278
## + tempo                       1    1.2772 4624.2 -14274
## <none>                                    4625.5 -14272
##
## Step:   AIC=-14331.94
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness
##
##                              Df Sum of Sq    RSS    AIC
## + explicit                    1   11.6617 4592.6 -14364
## + sqrt(instrumentalness)      1   10.9151 4593.3 -14362
## + mode                        1   10.8448 4593.4 -14362
## + acousticness                1    9.3261 4594.9 -14357
## + valence:morethan1artist     1    7.2341 4597.0 -14351
## + key                        11   12.4254 4591.8 -14346
## + loudness                    1    4.6393 4599.6 -14344
## + tempo                       1    1.2768 4602.9 -14334
## + speechiness                 1    0.7585 4603.5 -14332
## <none>                                    4604.2 -14332
## + liveness:morethan1artist    1    0.6714 4603.5 -14332
##
## Step:   AIC=-14363.95
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit
##
##                                  Df Sum of Sq    RSS    AIC
## + acousticness                    1   11.0053 4581.6 -14394
## + mode                            1    8.8039 4583.8 -14388
## + sqrt(instrumentalness)          1    8.7748 4583.8 -14388
```

```
## + valence:morethan1artist          1    7.3485 4585.2 -14383
## + speechiness                       1    6.4504 4586.1 -14381
## + loudness                          1    4.7405 4587.8 -14376
## + key                              11   11.4950 4581.1 -14376
## + tempo                             1    0.8527 4591.7 -14364
## <none>                                          4592.6 -14364
## + liveness:morethan1artist          1    0.6505 4591.9 -14364
## + spotifytraining$duration:explicit 1    0.4704 4592.1 -14363
##
## Step:  AIC=-14394.14
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness
##
##                                     Df Sum of Sq    RSS    AIC
## + loudness                          1   20.5718 4561.0 -14452
## + mode                              1    9.8304 4571.7 -14421
## + sqrt(instrumentalness)            1    9.6189 4571.9 -14420
## + speechiness                       1    6.9589 4574.6 -14412
## + valence:morethan1artist           1    6.5076 4575.0 -14411
## + key                              11   11.8057 4569.7 -14407
## + tempo                             1    1.9565 4579.6 -14398
## + liveness:morethan1artist          1    0.7424 4580.8 -14394
## + spotifytraining$duration:explicit 1    0.7395 4580.8 -14394
## <none>                                          4581.6 -14394
##
## Step:  AIC=-14452.5
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness
##
##                                     Df Sum of Sq    RSS    AIC
## + mode                              1    9.4708 4551.5 -14478
## + valence:morethan1artist           1    6.3590 4554.6 -14469
## + loudness:explicit                 1    6.0527 4554.9 -14468
## + sqrt(instrumentalness)            1    5.4850 4555.5 -14467
## + speechiness                       1    5.3434 4555.6 -14466
## + key                              11   11.4991 4549.5 -14464
## + tempo                             1    1.4752 4559.5 -14455
## + spotifytraining$duration:explicit 1    1.1732 4559.8 -14454
## + liveness:morethan1artist          1    0.6834 4560.3 -14452
## <none>                                          4561.0 -14452
##
## Step:  AIC=-14478.39
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode
##
##                                     Df Sum of Sq    RSS    AIC
```

```
## + speechiness                            1     6.1661 4545.3 -14495
## + valence:morethan1artist                1     6.1426 4545.4 -14494
## + sqrt(instrumentalness)                 1     6.0349 4545.5 -14494
## + loudness:explicit                      1     5.6630 4545.8 -14493
## + spotifytraining$duration:mode          1     3.2914 4548.2 -14486
## + valence:mode                           1     2.9113 4548.6 -14485
## + key                                   11     9.4540 4542.1 -14484
## + tempo                                  1     1.5852 4549.9 -14481
## + spotifytraining$duration:explicit  1     1.3137 4550.2 -14480
## <none>                                              4551.5 -14478
## + acousticness:mode                      1     0.6755 4550.8 -14478
## + liveness:morethan1artist               1     0.6512 4550.9 -14478
##
## Step:   AIC=-14494.57
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness
##
##                                        Df Sum of Sq    RSS     AIC
## + loudness:explicit                      1     7.1436 4538.2 -14514
## + sqrt(instrumentalness)                 1     6.9218 4538.4 -14513
## + valence:morethan1artist                1     6.6012 4538.7 -14512
## + speechiness:explicit                   1     5.3120 4540.0 -14508
## + spotifytraining$duration:mode          1     3.1912 4542.2 -14502
## + key                                   11     9.7917 4535.6 -14502
## + valence:mode                           1     2.9386 4542.4 -14501
## + speechiness:morethan1artist            1     2.7917 4542.6 -14501
## + tempo                                  1     2.0328 4543.3 -14499
## + spotifytraining$duration:explicit  1     1.1640 4544.2 -14496
## + acousticness:mode                      1     0.8364 4544.5 -14495
## + liveness:morethan1artist               1     0.7564 4544.6 -14495
## <none>                                              4545.3 -14495
##
## Step:   AIC=-14513.67
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + explicit:loudness
##
##                                        Df Sum of Sq    RSS     AIC
## + valence:morethan1artist                1     7.0411 4531.2 -14532
## + speechiness:explicit                   1     6.2580 4531.9 -14530
## + sqrt(instrumentalness)                 1     5.9941 4532.2 -14529
## + valence:mode                           1     3.3911 4534.8 -14522
## + spotifytraining$duration:mode          1     3.2694 4534.9 -14521
## + key                                   11     9.6034 4528.6 -14520
## + speechiness:morethan1artist            1     2.5284 4535.7 -14519
## + tempo                                  1     2.2265 4536.0 -14518
## + spotifytraining$duration:explicit  1     1.1152 4537.1 -14515
```

```
## <none>                                         4538.2 -14514
## + liveness:morethan1artist           1     0.6164 4537.6 -14514
## + acousticness:mode                  1     0.5611 4537.6 -14513
##
## Step:  AIC=-14532.5
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + explicit:loudness + morethan1artist:valence
##
##                                     Df Sum of Sq    RSS    AIC
## + speechiness:explicit               1     6.2595 4524.9 -14549
## + sqrt(instrumentalness)             1     6.1101 4525.0 -14549
## + spotifytraining$duration:mode      1     3.1361 4528.0 -14540
## + speechiness:morethan1artist        1     2.9078 4528.3 -14539
## + key                               11     9.4953 4521.7 -14539
## + valence:mode                       1     2.6489 4528.5 -14538
## + tempo                              1     2.3859 4528.8 -14538
## + spotifytraining$duration:explicit  1     0.7697 4530.4 -14533
## + liveness:morethan1artist           1     0.7138 4530.4 -14533
## <none>                                         4531.2 -14532
## + acousticness:mode                  1     0.5854 4530.6 -14532
##
## Step:  AIC=-14549.04
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + explicit:loudness + morethan1artist:valence +
##      explicit:speechiness
##
##                                     Df Sum of Sq    RSS    AIC
## + sqrt(instrumentalness)             1     6.0819 4518.8 -14565
## + spotifytraining$duration:mode      1     3.1828 4521.7 -14556
## + valence:mode                       1     2.8105 4522.1 -14555
## + key                               11     9.3785 4515.5 -14555
## + speechiness:morethan1artist        1     2.4496 4522.5 -14554
## + tempo                              1     2.1964 4522.7 -14554
## + spotifytraining$duration:explicit  1     0.9895 4523.9 -14550
## <none>                                         4524.9 -14549
## + liveness:morethan1artist           1     0.6206 4524.3 -14549
## + acousticness:mode                  1     0.5237 4524.4 -14549
##
## Step:  AIC=-14565.08
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + sqrt(instrumentalness) + explicit:loudness +
##      morethan1artist:valence + explicit:speechiness
##
##                                     Df Sum of Sq    RSS    AIC
```

```
## + spotifytraining$duration:mode      1     3.0039 4515.8 -14572
## + valence:mode                        1     2.8081 4516.0 -14571
## + key                                 11    9.3090 4509.5 -14571
## + speechiness:morethan1artist         1     2.3550 4516.5 -14570
## + tempo                               1     2.2228 4516.6 -14570
## + spotifytraining$duration:explicit   1     1.2437 4517.6 -14567
## + acousticness:mode                   1     0.6872 4518.1 -14565
## <none>                                              4518.8 -14565
## + liveness:morethan1artist            1     0.6060 4518.2 -14565
##
## Step:   AIC=-14572
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + sqrt(instrumentalness) + explicit:loudness +
##      morethan1artist:valence + explicit:speechiness + spotifytraining$durat
ion:mode
##
##                                       Df Sum of Sq    RSS    AIC
## + key                                 11    9.0360 4506.8 -14577
## + speechiness:morethan1artist         1     2.2875 4513.5 -14577
## + tempo                               1     2.2356 4513.6 -14577
## + valence:mode                        1     1.9087 4513.9 -14576
## <none>                                              4515.8 -14572
## + liveness:morethan1artist            1     0.5912 4515.2 -14572
## + acousticness:mode                   1     0.5738 4515.2 -14572
## + spotifytraining$duration:explicit   1     0.5555 4515.3 -14572
##
## Step:   AIC=-14576.87
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + sqrt(instrumentalness) + key + explicit:loudness
+
##      morethan1artist:valence + explicit:speechiness + spotifytraining$durat
ion:mode
##
##                                       Df Sum of Sq    RSS    AIC
## + speechiness:key                     11    9.9524 4496.8 -14584
## + tempo                               1     2.2827 4504.5 -14582
## + speechiness:morethan1artist         1     2.2115 4504.6 -14582
## + valence:mode                        1     1.7621 4505.0 -14580
## + acousticness:mode                   1     0.7434 4506.0 -14577
## <none>                                              4506.8 -14577
## + spotifytraining$duration:explicit   1     0.6135 4506.2 -14577
## + liveness:morethan1artist            1     0.5915 4506.2 -14577
## + valence:key                         11    7.2019 4499.6 -14576
## + sqrt(instrumentalness):key          11    7.0712 4499.7 -14576
## + loudness:key                        11    3.9861 4502.8 -14567
## + acousticness:key                    11    3.8093 4503.0 -14566
```

```
## + danceability:key                         11    3.0898 4503.7 -14564
## + spotifytraining$duration:key             11    2.8760 4503.9 -14563
##
## Step:  AIC=-14584.53
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + sqrt(instrumentalness) + key + explicit:loudness
+
##      morethan1artist:valence + explicit:speechiness + spotifytraining$durat
ion:mode +
##      speechiness:key
##
##                                   Df Sum of Sq    RSS    AIC
## + speechiness:morethan1artist      1    2.6739 4494.2 -14590
## + tempo                            1    2.0917 4494.7 -14589
## + valence:mode                     1    1.7720 4495.1 -14588
## + spotifytraining$duration:explicit  1    0.7409 4496.1 -14585
## + liveness:morethan1artist         1    0.7131 4496.1 -14585
## + acousticness:mode                1    0.6882 4496.1 -14585
## <none>                                        4496.8 -14584
## + valence:key                      11    7.3212 4489.5 -14584
## + sqrt(instrumentalness):key       11    6.6008 4490.2 -14582
## + loudness:key                     11    3.8957 4492.9 -14574
## + acousticness:key                 11    3.6200 4493.2 -14573
## + danceability:key                 11    3.4257 4493.4 -14573
## + spotifytraining$duration:key     11    3.2094 4493.6 -14572
##
## Step:  AIC=-14590.5
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness + loudness +
##      mode + speechiness + sqrt(instrumentalness) + key + explicit:loudness
+
##      morethan1artist:valence + explicit:speechiness + spotifytraining$durat
ion:mode +
##      speechiness:key + morethan1artist:speechiness
##
##                                   Df Sum of Sq    RSS    AIC
## + tempo                            1    2.0662 4492.1 -14595
## + valence:mode                     1    1.7480 4492.4 -14594
## + valence:key                      11    7.3783 4486.8 -14590
## + spotifytraining$duration:explicit  1    0.6782 4493.5 -14590
## <none>                                        4494.2 -14590
## + acousticness:mode                1    0.6651 4493.5 -14590
## + liveness:morethan1artist         1    0.4597 4493.7 -14590
## + sqrt(instrumentalness):key       11    6.5564 4487.6 -14588
## + loudness:key                     11    3.7934 4490.4 -14580
## + acousticness:key                 11    3.5226 4490.6 -14579
## + danceability:key                 11    3.4962 4490.7 -14579
```

```
## + spotifytraining$duration:key        11    3.2271 4490.9 -14578
##
## Step:  AIC=-14594.67
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + sqrt(instrumentalness) + key + tempo +
##     explicit:loudness + morethan1artist:valence + explicit:speechiness +
##     spotifytraining$duration:mode + speechiness:key + morethan1artist:spee
chiness
##
##                                    Df Sum of Sq    RSS     AIC
## + tempo:explicit                    1    3.0079 4489.1 -14602
## + valence:mode                      1    1.7968 4490.3 -14598
## + tempo:morethan1artist             1    1.6747 4490.4 -14598
## + valence:key                      11    7.4174 4484.7 -14595
## + acousticness:mode                 1    0.6781 4491.4 -14595
## <none>                                          4492.1 -14595
## + spotifytraining$duration:explicit  1   0.6262 4491.5 -14594
## + liveness:morethan1artist          1    0.4366 4491.6 -14594
## + sqrt(instrumentalness):key       11    6.6216 4485.5 -14592
## + loudness:key                     11    3.9173 4488.2 -14584
## + danceability:key                 11    3.5285 4488.6 -14583
## + acousticness:key                 11    3.5217 4488.6 -14583
## + spotifytraining$duration:key     11    3.1810 4488.9 -14582
## + tempo:key                        11    2.6611 4489.4 -14581
##
## Step:  AIC=-14601.66
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + sqrt(instrumentalness) + key + tempo +
##     explicit:loudness + morethan1artist:valence + explicit:speechiness +
##     spotifytraining$duration:mode + speechiness:key + morethan1artist:spee
chiness +
##     explicit:tempo
##
##                                    Df Sum of Sq    RSS     AIC
## + valence:mode                      1    1.7339 4487.3 -14605
## + tempo:morethan1artist             1    0.8801 4488.2 -14602
## + valence:key                      11    7.4955 4481.6 -14602
## + acousticness:mode                 1    0.6973 4488.4 -14602
## <none>                                          4489.1 -14602
## + spotifytraining$duration:explicit  1   0.5852 4488.5 -14601
## + liveness:morethan1artist          1    0.4340 4488.6 -14601
## + sqrt(instrumentalness):key       11    6.5738 4482.5 -14599
## + loudness:key                     11    3.8741 4485.2 -14591
## + acousticness:key                 11    3.5455 4485.5 -14590
## + danceability:key                 11    3.5348 4485.5 -14590
## + spotifytraining$duration:key     11    3.1864 4485.9 -14589
```

```
## + tempo:key                                 11    2.7563 4486.3 -14588
##
## Step:  AIC=-14604.84
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + sqrt(instrumentalness) + key + tempo +
##     explicit:loudness + morethan1artist:valence + explicit:speechiness +
##     spotifytraining$duration:mode + speechiness:key + morethan1artist:spee
chiness +
##     explicit:tempo + valence:mode
##
##                                         Df Sum of Sq    RSS     AIC
## + acousticness:mode                      1    0.9508 4486.4 -14606
## + tempo:morethan1artist                  1    0.8920 4486.5 -14606
## <none>                                             4487.3 -14605
## + spotifytraining$duration:explicit      1    0.5569 4486.8 -14604
## + liveness:morethan1artist               1    0.4391 4486.9 -14604
## + valence:key                           11    7.0147 4480.3 -14604
## + sqrt(instrumentalness):key            11    6.4652 4480.9 -14602
## + loudness:key                          11    3.6885 4483.7 -14594
## + acousticness:key                      11    3.6193 4483.7 -14594
## + danceability:key                      11    3.4985 4483.8 -14593
## + spotifytraining$duration:key          11    3.1034 4484.2 -14592
## + tempo:key                             11    2.7687 4484.6 -14591
##
## Step:  AIC=-14605.68
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + sqrt(instrumentalness) + key + tempo +
##     explicit:loudness + morethan1artist:valence + explicit:speechiness +
##     spotifytraining$duration:mode + speechiness:key + morethan1artist:spee
chiness +
##     explicit:tempo + valence:mode + acousticness:mode
##
##                                         Df Sum of Sq    RSS     AIC
## + tempo:morethan1artist                  1    0.8904 4485.5 -14606
## <none>                                             4486.4 -14606
## + spotifytraining$duration:explicit      1    0.5682 4485.8 -14605
## + liveness:morethan1artist               1    0.4197 4486.0 -14605
## + valence:key                           11    7.0440 4479.3 -14605
## + sqrt(instrumentalness):key            11    6.5766 4479.8 -14603
## + loudness:key                          11    3.8730 4482.5 -14595
## + danceability:key                      11    3.4553 4482.9 -14594
## + acousticness:key                      11    3.4471 4482.9 -14594
## + spotifytraining$duration:key          11    3.0597 4483.3 -14593
## + tempo:key                             11    2.7426 4483.7 -14592
##
## Step:  AIC=-14606.35
```

```
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + sqrt(instrumentalness) + key + tempo +
##     explicit:loudness + morethan1artist:valence + explicit:speechiness +
##     spotifytraining$duration:mode + speechiness:key + morethan1artist:spee
chiness +
##     explicit:tempo + valence:mode + acousticness:mode + morethan1artist:te
mpo
##
##                                       Df Sum of Sq    RSS    AIC
## <none>                                            4485.5 -14606
## + spotifytraining$duration:explicit   1    0.5769 4484.9 -14606
## + liveness:morethan1artist            1    0.4217 4485.1 -14606
## + valence:key                        11    7.0621 4478.4 -14606
## + sqrt(instrumentalness):key         11    6.5109 4479.0 -14604
## + loudness:key                       11    3.8507 4481.7 -14596
## + danceability:key                   11    3.5185 4482.0 -14595
## + acousticness:key                   11    3.4927 4482.0 -14595
## + spotifytraining$duration:key       11    3.0707 4482.4 -14594
## + tempo:key                          11    2.7252 4482.8 -14592

anova(forwardAIC)

## Analysis of Variance Table
##
## Response: sqrt(popularity)
##                             Df Sum Sq Mean Sq  F value     Pr(>F)
## danceability                 1  267.3 267.312 796.7810 < 2.2e-16 ***
## morethan1artist              1   86.0  85.991 256.3133 < 2.2e-16 ***
## spotifytraining$duration     1   68.8  68.756 204.9423 < 2.2e-16 ***
## valence                      1   67.9  67.893 202.3696 < 2.2e-16 ***
## liveness                     1   21.3  21.305  63.5031 1.730e-15 ***
## explicit                     1   11.7  11.662  34.7603 3.819e-09 ***
## acousticness                 1   11.0  11.005  32.8035 1.042e-08 ***
## loudness                     1   20.6  20.572  61.3188 5.219e-15 ***
## mode                         1    9.5   9.471  28.2298 1.095e-07 ***
## speechiness                  1    6.2   6.166  18.3794 1.823e-05 ***
## sqrt(instrumentalness)       1    6.9   6.922  20.6318 5.615e-06 ***
## key                         11    9.7   0.882   2.6279  0.002365 **
## tempo                        1    2.1   2.122   6.3245  0.011920 *
## explicit:loudness            1    6.2   6.230  18.5699 1.650e-05 ***
## morethan1artist:valence      1    7.2   7.211  21.4953 3.581e-06 ***
## explicit:speechiness         1    5.9   5.924  17.6563 2.664e-05 ***
## spotifytraining$duration:mode 1   2.7   2.742   8.1735  0.004257 **
## speechiness:key             11    9.8   0.887   2.6451  0.002211 **
## morethan1artist:speechiness  1    2.6   2.648   7.8941  0.004967 **
## explicit:tempo               1    3.0   3.008   8.9657  0.002756 **
## valence:mode                 1    1.7   1.734   5.1682  0.023020 *
## acousticness:mode            1    1.0   0.951   2.8339  0.092315 .
```

```
## morethan1artist:tempo              1    0.9   0.890   2.6542  0.103303
## Residuals                       13370 4485.5   0.335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

forwardAIC$coefficients

##                  (Intercept)                         danceability
##                 7.767951e+00                         7.896162e-01
##               morethan1artist1         spotifytraining$duration
##                 1.923002e-02                        -9.364216e-02
##                      valence                             liveness
##                -2.905865e-01                        -2.757123e-01
##                    explicit1                         acousticness
##                -1.155144e-01                         2.658353e-01
##                     loudness                                mode1
##                 1.973806e-02                        -7.977838e-02
##                  speechiness              sqrt(instrumentalness)
##                 2.287489e-01                        -2.132366e-01
##                         key1                                 key2
##                 3.455087e-02                        -2.873166e-02
##                         key3                                 key4
##                -4.858325e-02                         4.896025e-02
##                         key5                                 key6
##                 1.015961e-01                         7.248942e-02
##                         key7                                 key8
##                -3.574053e-02                         2.526435e-02
##                         key9                                key10
##                -2.714305e-02                        -2.405045e-02
##                        key11                                tempo
##                 3.027273e-02                        -3.565109e-05
##            explicit1:loudness         morethan1artist1:valence
##                -2.190629e-02                         2.313030e-01
##         explicit1:speechiness spotifytraining$duration:mode1
##                -4.446127e-01                         2.665946e-02
##              speechiness:key1                      speechiness:key2
##                -2.047157e-01                         1.337637e-02
##              speechiness:key3                      speechiness:key4
##                 1.826353e-01                        -7.246027e-01
##              speechiness:key5                      speechiness:key6
##                -4.313545e-01                        -3.530245e-01
##              speechiness:key7                      speechiness:key8
##                 1.543365e-01                        -4.839646e-02
##              speechiness:key9                     speechiness:key10
##                 3.179464e-03                         1.503345e-01
##             speechiness:key11     morethan1artist1:speechiness
##                -1.838087e-01                        -2.740968e-01
##                explicit1:tempo                         valence:mode1
##                 9.558513e-04                        -1.140180e-01
```

```
##              acousticness:mode1          morethan1artist1:tempo
##                  -7.119364e-02                    6.555610e-04

extractAIC(forwardAIC)

## [1]      44.00 -14606.35

## [1] 0.1231505
```

BIC

```
## Start:  AIC=-12921.97
## sqrt(popularity) ~ 1
##
##                            Df Sum of Sq    RSS    AIC
## + danceability              1   267.312 4848.2 -13632
## + morethan1artist           1   148.007 4967.5 -13306
## + spotifytraining$duration  1   124.293 4991.2 -13242
## + explicit                  1   122.051 4993.4 -13236
## + liveness                  1    44.852 5070.6 -13031
## + sqrt(instrumentalness)    1    28.796 5086.7 -12988
## + mode                      1    25.791 5089.7 -12980
## + speechiness               1    17.124 5098.4 -12957
## + acousticness              1     6.600 5108.9 -12930
## <none>                                   5115.5 -12922
## + tempo                     1     2.029 5113.4 -12918
## + valence                   1     2.024 5113.5 -12918
## + loudness                  1     0.217 5115.3 -12913
## + key                      11    24.933 5090.5 -12883
##
## Step:  AIC=-13632.4
## sqrt(popularity) ~ danceability
##
##                            Df Sum of Sq    RSS    AIC
## + morethan1artist           1    85.991 4762.2 -13863
## + valence                   1    60.918 4787.2 -13792
## + spotifytraining$duration  1    57.359 4790.8 -13782
## + explicit                  1    40.909 4807.3 -13737
## + liveness                  1    24.900 4823.3 -13692
## + sqrt(instrumentalness)    1    16.859 4831.3 -13670
## + acousticness              1    15.150 4833.0 -13665
## + mode                      1    14.658 4833.5 -13664
## <none>                                   4848.2 -13632
## + tempo                     1     0.805 4847.4 -13625
## + speechiness               1     0.548 4847.6 -13624
## + loudness                  1     0.005 4848.2 -13623
## + key                      11    14.380 4833.8 -13568
##
## Step:  AIC=-13862.96
## sqrt(popularity) ~ danceability + morethan1artist
##
##                            Df Sum of Sq    RSS    AIC
```

```
## + spotifytraining$duration  1      68.756 4693.4 -14048
## + valence                    1      49.925 4712.2 -13995
## + liveness                   1      30.385 4731.8 -13939
## + explicit                   1      22.769 4739.4 -13918
## + acousticness               1      20.729 4741.4 -13912
## + sqrt(instrumentalness)     1      11.846 4750.3 -13887
## + mode                       1      10.381 4751.8 -13883
## <none>                                      4762.2 -13863
## + tempo                      1       0.628 4761.5 -13855
## + speechiness                1       0.554 4761.6 -13855
## + loudness                   1       0.165 4762.0 -13854
## + key                        11     12.368 4749.8 -13793
##
## Step:  AIC=-14048.53
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on
##
##                             Df Sum of Sq    RSS    AIC
## + valence                    1     67.893 4625.5 -14234
## + liveness                   1     26.484 4666.9 -14115
## + explicit                   1     20.981 4672.4 -14099
## + acousticness               1     17.077 4676.3 -14088
## + mode                       1     11.266 4682.2 -14071
## + sqrt(instrumentalness)     1      7.059 4686.4 -14059
## <none>                                     4693.4 -14048
## + speechiness                1      1.379 4692.0 -14043
## + loudness                   1      0.340 4693.1 -14040
## + tempo                      1      0.146 4693.3 -14039
## + key                        11     12.133 4681.3 -13979
##
## Step:  AIC=-14234.49
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence
##
##                             Df Sum of Sq    RSS    AIC
## + liveness                   1    21.3046 4604.2 -14287
## + acousticness               1    11.2566 4614.3 -14258
## + mode                       1    10.0608 4615.5 -14254
## + sqrt(instrumentalness)     1     9.8758 4615.6 -14254
## + explicit                   1     8.4357 4617.1 -14250
## + valence:morethan1artist    1     7.0808 4618.4 -14246
## <none>                                     4625.5 -14234
## + loudness                   1     3.2423 4622.3 -14234
## + speechiness                1     2.6087 4622.9 -14233
## + tempo                      1     1.2772 4624.2 -14229
## + key                        11    12.5727 4613.0 -14166
##
## Step:  AIC=-14286.91
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
```

```
on +
##      valence + liveness
##
##                                 Df Sum of Sq    RSS     AIC
## + explicit                       1    11.6617 4592.6 -14311
## + sqrt(instrumentalness)         1    10.9151 4593.3 -14309
## + mode                           1    10.8448 4593.4 -14309
## + acousticness                   1     9.3261 4594.9 -14305
## + valence:morethan1artist        1     7.2341 4597.0 -14298
## + loudness                       1     4.6393 4599.6 -14291
## <none>                                        4604.2 -14287
## + tempo                          1     1.2768 4602.9 -14281
## + speechiness                    1     0.7585 4603.5 -14280
## + liveness:morethan1artist       1     0.6714 4603.5 -14279
## + key                           11    12.4254 4591.8 -14219
##
## Step:  AIC=-14311.43
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit
##
##                                       Df Sum of Sq    RSS     AIC
## + acousticness                        1    11.0053 4581.6 -14334
## + mode                                1     8.8039 4583.8 -14328
## + sqrt(instrumentalness)              1     8.7748 4583.8 -14328
## + valence:morethan1artist             1     7.3485 4585.2 -14323
## + speechiness                         1     6.4504 4586.1 -14321
## + loudness                            1     4.7405 4587.8 -14316
## <none>                                              4592.6 -14311
## + tempo                               1     0.8527 4591.7 -14304
## + liveness:morethan1artist            1     0.6505 4591.9 -14304
## + spotifytraining$duration:explicit   1     0.4704 4592.1 -14303
## + key                                11    11.4950 4581.1 -14240
##
## Step:  AIC=-14334.1
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##      valence + liveness + explicit + acousticness
##
##                                       Df Sum of Sq    RSS     AIC
## + loudness                            1    20.5718 4561.0 -14385
## + mode                                1     9.8304 4571.7 -14353
## + sqrt(instrumentalness)              1     9.6189 4571.9 -14353
## + speechiness                         1     6.9589 4574.6 -14345
## + valence:morethan1artist             1     6.5076 4575.0 -14344
## <none>                                              4581.6 -14334
## + tempo                               1     1.9565 4579.6 -14330
## + liveness:morethan1artist            1     0.7424 4580.8 -14327
## + spotifytraining$duration:explicit   1     0.7395 4580.8 -14327
## + key                                11    11.8057 4569.7 -14264
```

```
## 
## Step:   AIC=-14384.97
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness
## 
##                                  Df Sum of Sq    RSS    AIC
## + mode                            1    9.4708 4551.5 -14403
## + valence:morethan1artist         1    6.3590 4554.6 -14394
## + loudness:explicit               1    6.0527 4554.9 -14393
## + sqrt(instrumentalness)          1    5.4850 4555.5 -14392
## + speechiness                     1    5.3434 4555.6 -14391
## <none>                                         4561.0 -14385
## + tempo                           1    1.4752 4559.5 -14380
## + spotifytraining$duration:explicit  1  1.1732 4559.8 -14379
## + liveness:morethan1artist        1    0.6834 4560.3 -14378
## + key                            11   11.4991 4549.5 -14314
## 
## Step:   AIC=-14403.35
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode
## 
##                                  Df Sum of Sq    RSS    AIC
## + speechiness                     1    6.1661 4545.3 -14412
## + valence:morethan1artist         1    6.1426 4545.4 -14412
## + sqrt(instrumentalness)          1    6.0349 4545.5 -14412
## + loudness:explicit               1    5.6630 4545.8 -14410
## + spotifytraining$duration:mode   1    3.2914 4548.2 -14404
## <none>                                         4551.5 -14403
## + valence:mode                    1    2.9113 4548.6 -14402
## + tempo                           1    1.5852 4549.9 -14398
## + spotifytraining$duration:explicit  1  1.3137 4550.2 -14398
## + acousticness:mode               1    0.6755 4550.8 -14396
## + liveness:morethan1artist        1    0.6512 4550.9 -14396
## + key                            11    9.4540 4542.1 -14327
## 
## Step:   AIC=-14412.03
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness
## 
##                                  Df Sum of Sq    RSS    AIC
## + loudness:explicit               1    7.1436 4538.2 -14424
## + sqrt(instrumentalness)          1    6.9218 4538.4 -14423
## + valence:morethan1artist         1    6.6012 4538.7 -14422
## + speechiness:explicit            1    5.3120 4540.0 -14418
## <none>                                         4545.3 -14412
```

```
## + spotifytraining$duration:mode        1     3.1912 4542.2 -14412
## + valence:mode                          1     2.9386 4542.4 -14411
## + speechiness:morethan1artist           1     2.7917 4542.6 -14411
## + tempo                                 1     2.0328 4543.3 -14408
## + spotifytraining$duration:explicit  1     1.1640 4544.2 -14406
## + acousticness:mode                     1     0.8364 4544.5 -14405
## + liveness:morethan1artist              1     0.7564 4544.6 -14405
## + key                                  11     9.7917 4535.6 -14336
##
## Step:  AIC=-14423.62
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + explicit:loudness
##
##                                       Df Sum of Sq    RSS    AIC
## + valence:morethan1artist              1     7.0411 4531.2 -14435
## + speechiness:explicit                 1     6.2580 4531.9 -14433
## + sqrt(instrumentalness)               1     5.9941 4532.2 -14432
## + valence:mode                         1     3.3911 4534.8 -14424
## + spotifytraining$duration:mode        1     3.2694 4534.9 -14424
## <none>                                              4538.2 -14424
## + speechiness:morethan1artist          1     2.5284 4535.7 -14422
## + tempo                                1     2.2265 4536.0 -14421
## + spotifytraining$duration:explicit  1     1.1152 4537.1 -14417
## + liveness:morethan1artist             1     0.6164 4537.6 -14416
## + acousticness:mode                    1     0.5611 4537.6 -14416
## + key                                 11     9.6034 4528.6 -14348
##
## Step:  AIC=-14434.94
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + explicit:loudness + morethan1artist:valence
##
##                                       Df Sum of Sq    RSS    AIC
## + speechiness:explicit                 1     6.2595 4524.9 -14444
## + sqrt(instrumentalness)               1     6.1101 4525.0 -14444
## <none>                                              4531.2 -14435
## + spotifytraining$duration:mode        1     3.1361 4528.0 -14435
## + speechiness:morethan1artist          1     2.9078 4528.3 -14434
## + valence:mode                         1     2.6489 4528.5 -14433
## + tempo                                1     2.3859 4528.8 -14432
## + spotifytraining$duration:explicit  1     0.7697 4530.4 -14428
## + liveness:morethan1artist             1     0.7138 4530.4 -14428
## + acousticness:mode                    1     0.5854 4530.6 -14427
## + key                                 11     9.4953 4521.7 -14358
##
## Step:  AIC=-14443.98
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
```

```
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + explicit:loudness + morethan1artist:valence +
##     explicit:speechiness
##
##                                   Df Sum of Sq    RSS    AIC
## + sqrt(instrumentalness)           1    6.0819 4518.8 -14452
## <none>                                         4524.9 -14444
## + spotifytraining$duration:mode    1    3.1828 4521.7 -14444
## + valence:mode                     1    2.8105 4522.1 -14443
## + speechiness:morethan1artist      1    2.4496 4522.5 -14442
## + tempo                            1    2.1964 4522.7 -14441
## + spotifytraining$duration:explicit 1   0.9895 4523.9 -14437
## + liveness:morethan1artist         1    0.6206 4524.3 -14436
## + acousticness:mode                1    0.5237 4524.4 -14436
## + key                             11    9.3785 4515.5 -14367
##
## Step:  AIC=-14452.52
## sqrt(popularity) ~ danceability + morethan1artist + spotifytraining$durati
on +
##     valence + liveness + explicit + acousticness + loudness +
##     mode + speechiness + sqrt(instrumentalness) + explicit:loudness +
##     morethan1artist:valence + explicit:speechiness
##
##                                   Df Sum of Sq    RSS    AIC
## <none>                                         4518.8 -14452
## + spotifytraining$duration:mode    1    3.0039 4515.8 -14452
## + valence:mode                     1    2.8081 4516.0 -14451
## + speechiness:morethan1artist      1    2.3550 4516.5 -14450
## + tempo                            1    2.2228 4516.6 -14450
## + spotifytraining$duration:explicit 1   1.2437 4517.6 -14447
## + acousticness:mode                1    0.6872 4518.1 -14445
## + liveness:morethan1artist         1    0.6060 4518.2 -14445
## + key                             11    9.3090 4509.5 -14376

anova(forwardBIC)

## Analysis of Variance Table
##
## Response: sqrt(popularity)
##                           Df Sum Sq Mean Sq F value    Pr(>F)
## danceability               1  267.3 267.312 792.622 < 2.2e-16 ***
## morethan1artist            1   86.0  85.991 254.975 < 2.2e-16 ***
## spotifytraining$duration   1   68.8  68.756 203.873 < 2.2e-16 ***
## valence                    1   67.9  67.893 201.313 < 2.2e-16 ***
## liveness                   1   21.3  21.305  63.172 2.045e-15 ***
## explicit                   1   11.7  11.662  34.579 4.191e-09 ***
## acousticness               1   11.0  11.005  32.632 1.137e-08 ***
## loudness                   1   20.6  20.572  60.999 6.134e-15 ***
## mode                       1    9.5   9.471  28.082 1.181e-07 ***
```

```
## speechiness                    1     6.2    6.166   18.283 1.917e-05 ***
## sqrt(instrumentalness)          1     6.9    6.922   20.524 5.939e-06 ***
## explicit:loudness               1     6.2    6.216   18.431 1.774e-05 ***
## morethan1artist:valence         1     7.2    7.157   21.222 4.128e-06 ***
## explicit:speechiness            1     6.2    6.231   18.477 1.732e-05 ***
## Residuals                   13399 4518.8    0.337
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(forwardBIC)$r.sq
```

```
## [1] 0.1166379
```

```
summary(forwardBIC)$adj.r.sq
```

```
## [1] 0.1157149
```

```
extractAIC(forwardBIC)
```

```
## [1]     15.00 -14565.08
```

Backwards

AIC
```
## Start:  AIC=-14541.88
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:key + spotifytraining$duration:mode +
##     spotifytraining$duration:explicit + tempo:key + tempo:morethan1artist
+
##     tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##     acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                     Df Sum of Sq    RSS    AIC
## - tempo:key                         11    2.4055 4456.7 -14557
## - spotifytraining$duration:key      11    3.1540 4457.5 -14554
## - danceability:key                  11    3.7012 4458.0 -14553
## - loudness:key                      11    3.8329 4458.2 -14552
## - acousticness:key                  11    4.8029 4459.1 -14549
## - valence:key                       11    5.4069 4459.7 -14548
## - sqrt(instrumentalness):key        11    6.1832 4460.5 -14545
## - spotifytraining$duration:explicit  1    0.2979 4454.6 -14543
## - liveness:morethan1artist           1    0.4406 4454.8 -14543
## <none>                                             4454.3 -14542
## - acousticness:mode                  1    0.7883 4455.1 -14542
## - tempo:morethan1artist              1    0.8647 4455.2 -14541
## - valence:mode                       1    1.2314 4455.6 -14540
```

```
## - spotifytraining$duration:mode      1     1.3392 4455.7 -14540
## - speechiness:key                    11     8.7289 4463.1 -14538
## - tempo:explicit                      1     2.1176 4456.5 -14538
## - speechiness:morethan1artist         1     2.2172 4456.6 -14537
## - valence:morethan1artist             1     5.4381 4459.8 -14528
## - speechiness:explicit                1     5.9393 4460.3 -14526
## - loudness:explicit                   1     8.4834 4462.8 -14518
##
## Step:  AIC=-14556.64
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:key + spotifytraining$duration:mode +
##     spotifytraining$duration:explicit + tempo:morethan1artist +
##     tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##     acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:key       11     3.1683 4459.9 -14569
## - danceability:key                   11     3.5847 4460.3 -14568
## - loudness:key                       11     3.8937 4460.6 -14567
## - acousticness:key                   11     5.1785 4461.9 -14563
## - valence:key                        11     5.2527 4462.0 -14563
## - sqrt(instrumentalness):key         11     6.1855 4462.9 -14560
## - spotifytraining$duration:explicit   1     0.3258 4457.1 -14558
## - liveness:morethan1artist            1     0.4292 4457.2 -14557
## <none>                                             4456.7 -14557
## - acousticness:mode                   1     0.7956 4457.5 -14556
## - tempo:morethan1artist               1     0.9100 4457.6 -14556
## - valence:mode                        1     1.2421 4458.0 -14555
## - spotifytraining$duration:mode       1     1.3293 4458.1 -14555
## - tempo:explicit                      1     2.0498 4458.8 -14552
## - speechiness:morethan1artist         1     2.2414 4459.0 -14552
## - speechiness:key                    11     9.1536 4465.9 -14551
## - valence:morethan1artist             1     5.5502 4462.3 -14542
## - speechiness:explicit                1     5.9554 4462.7 -14541
## - loudness:explicit                   1     8.3950 4465.1 -14533
##
## Step:  AIC=-14569.11
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##     acousticness:key + acousticness:mode + danceability:key +
##     sqrt(instrumentalness):key + valence:key + valence:morethan1artist +
```

```
##      valence:mode + liveness:morethan1artist + speechiness:key +
##      speechiness:morethan1artist + speechiness:explicit
##
##                                   Df Sum of Sq    RSS    AIC
## - danceability:key                11    3.3818 4463.3 -14581
## - loudness:key                     11    3.9499 4463.9 -14579
## - acousticness:key                 11    5.0660 4465.0 -14576
## - valence:key                      11    5.3805 4465.3 -14575
## - sqrt(instrumentalness):key       11    6.1214 4466.0 -14573
## - liveness:morethan1artist          1    0.3998 4460.3 -14570
## - spotifytraining$duration:explicit 1    0.5200 4460.4 -14570
## <none>                                          4459.9 -14569
## - acousticness:mode                 1    0.8420 4460.7 -14569
## - tempo:morethan1artist             1    0.9229 4460.8 -14568
## - spotifytraining$duration:mode     1    1.2937 4461.2 -14567
## - valence:mode                      1    1.3102 4461.2 -14567
## - tempo:explicit                    1    2.0620 4462.0 -14565
## - speechiness:morethan1artist       1    2.2447 4462.2 -14564
## - speechiness:key                  11    9.0811 4469.0 -14564
## - valence:morethan1artist           1    5.5611 4465.5 -14554
## - speechiness:explicit              1    6.0068 4465.9 -14553
## - loudness:explicit                 1    8.5059 4468.4 -14546
##
## Step:  AIC=-14580.94
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##      acousticness + danceability + sqrt(instrumentalness) + valence +
##      liveness + speechiness + key + morethan1artist + mode + explicit +
##      spotifytraining$duration:mode + spotifytraining$duration:explicit +
##      tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##      acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##      valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##      speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                   Df Sum of Sq    RSS    AIC
## - loudness:key                     11     3.768 4467.1 -14592
## - acousticness:key                 11     5.142 4468.4 -14588
## - valence:key                      11     5.474 4468.8 -14586
## - sqrt(instrumentalness):key       11     6.048 4469.3 -14585
## - liveness:morethan1artist          1     0.382 4463.7 -14582
## - spotifytraining$duration:explicit 1     0.489 4463.8 -14582
## <none>                                          4463.3 -14581
## - acousticness:mode                 1     0.833 4464.1 -14580
## - tempo:morethan1artist             1     0.863 4464.2 -14580
## - spotifytraining$duration:mode     1     1.233 4464.5 -14579
## - valence:mode                      1     1.346 4464.6 -14579
## - tempo:explicit                    1     2.169 4465.5 -14576
## - speechiness:morethan1artist       1     2.187 4465.5 -14576
## - speechiness:key                  11     9.140 4472.4 -14576
```

```
## - valence:morethan1artist                        1       5.670 4469.0 -14566
## - speechiness:explicit                            1       5.838 4469.1 -14565
## - loudness:explicit                               1       8.361 4471.6 -14558
## - danceability                                    1     132.741 4596.0 -14190
##
## Step:  AIC=-14591.62
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - acousticness:key                     11       4.288 4471.3 -14601
## - sqrt(instrumentalness):key           11       6.276 4473.3 -14595
## - liveness:morethan1artist              1       0.390 4467.4 -14592
## - spotifytraining$duration:explicit     1       0.507 4467.6 -14592
## <none>                                                4467.1 -14592
## - valence:key                          11       7.353 4474.4 -14592
## - acousticness:mode                     1       0.820 4467.9 -14591
## - tempo:morethan1artist                 1       0.896 4468.0 -14591
## - spotifytraining$duration:mode         1       1.286 4468.3 -14590
## - valence:mode                          1       1.425 4468.5 -14589
## - tempo:explicit                        1       2.197 4469.3 -14587
## - speechiness:morethan1artist           1       2.324 4469.4 -14587
## - speechiness:key                      11       9.793 4476.8 -14584
## - valence:morethan1artist               1       5.807 4472.9 -14576
## - speechiness:explicit                  1       6.022 4473.1 -14576
## - loudness:explicit                     1       8.640 4475.7 -14568
## - danceability                          1     132.558 4599.6 -14201
##
## Step:  AIC=-14600.75
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + sqrt(instrumentalness):key + valence:key +
##     valence:morethan1artist + valence:mode + liveness:morethan1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - sqrt(instrumentalness):key           11       6.202 4477.5 -14604
## - valence:key                          11       6.631 4478.0 -14603
## - liveness:morethan1artist              1       0.395 4471.7 -14602
## - spotifytraining$duration:explicit     1       0.514 4471.9 -14601
```

```
## <none>                                         4471.3 -14601
## - tempo:morethan1artist          1    0.851 4472.2 -14600
## - acousticness:mode              1    1.068 4472.4 -14600
## - spotifytraining$duration:mode  1    1.356 4472.7 -14599
## - valence:mode                   1    1.417 4472.8 -14598
## - tempo:explicit                 1    2.180 4473.5 -14596
## - speechiness:morethan1artist    1    2.410 4473.8 -14596
## - speechiness:key               11    9.915 4481.3 -14593
## - valence:morethan1artist        1    5.999 4477.3 -14585
## - speechiness:explicit           1    6.028 4477.4 -14585
## - loudness:explicit              1    8.782 4480.1 -14576
## - danceability                   1  132.170 4603.5 -14212
##
## Step:  AIC=-14604.16
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:key + valence:morethan1artist +
##     valence:mode + liveness:morethan1artist + speechiness:key +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                   Df Sum of Sq    RSS    AIC
## - valence:key                     11    6.973 4484.5 -14605
## - liveness:morethan1artist         1    0.360 4477.9 -14605
## - spotifytraining$duration:explicit 1   0.520 4478.1 -14605
## <none>                                         4477.5 -14604
## - tempo:morethan1artist           1    0.919 4478.5 -14603
## - acousticness:mode               1    0.970 4478.5 -14603
## - spotifytraining$duration:mode   1    1.394 4478.9 -14602
## - valence:mode                    1    1.463 4479.0 -14602
## - tempo:explicit                  1    2.195 4479.7 -14600
## - speechiness:morethan1artist     1    2.455 4480.0 -14599
## - speechiness:key                11   10.394 4487.9 -14595
## - speechiness:explicit            1    6.007 4483.6 -14588
## - valence:morethan1artist         1    6.074 4483.6 -14588
## - sqrt(instrumentalness)          1    6.098 4483.6 -14588
## - loudness:explicit               1    8.878 4486.4 -14580
## - danceability                    1  132.696 4610.2 -14214
##
## Step:  AIC=-14605.28
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     liveness:morethan1artist + speechiness:key + speechiness:morethan1arti
st +
```

```
##     speechiness:explicit
##
##                                     Df Sum of Sq    RSS    AIC
## - liveness:morethan1artist           1      0.406 4484.9 -14606
## - spotifytraining$duration:explicit  1      0.561 4485.1 -14606
## <none>                                            4484.5 -14605
## - tempo:morethan1artist              1      0.901 4485.4 -14605
## - acousticness:mode                  1      0.941 4485.5 -14604
## - spotifytraining$duration:mode      1      1.389 4485.9 -14603
## - valence:mode                       1      1.974 4486.5 -14601
## - tempo:explicit                     1      2.133 4486.7 -14601
## - speechiness:morethan1artist        1      2.385 4486.9 -14600
## - speechiness:key                   11     10.288 4494.8 -14596
## - speechiness:explicit               1      5.939 4490.5 -14590
## - sqrt(instrumentalness)             1      6.107 4490.6 -14589
## - valence:morethan1artist            1      6.602 4491.1 -14588
## - loudness:explicit                  1      8.573 4493.1 -14582
## - danceability                       1    131.653 4616.2 -14219
##
## Step:  AIC=-14606.07
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                     Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:explicit  1      0.577 4485.5 -14606
## <none>                                            4484.9 -14606
## - tempo:morethan1artist              1      0.899 4485.8 -14605
## - acousticness:mode                  1      0.961 4485.9 -14605
## - spotifytraining$duration:mode      1      1.388 4486.3 -14604
## - valence:mode                       1      1.971 4486.9 -14602
## - tempo:explicit                     1      2.135 4487.1 -14602
## - speechiness:morethan1artist        1      2.621 4487.5 -14600
## - speechiness:key                   11     10.215 4495.1 -14598
## - speechiness:explicit               1      6.000 4490.9 -14590
## - sqrt(instrumentalness)             1      6.121 4491.0 -14590
## - valence:morethan1artist            1      6.544 4491.5 -14588
## - loudness:explicit                  1      8.680 4493.6 -14582
## - liveness                           1     21.272 4506.2 -14545
## - danceability                       1    131.762 4616.7 -14220
##
## Step:  AIC=-14606.35
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:morethan1artist + tempo:explicit
```

```
+
##     loudness:explicit + acousticness:mode + valence:morethan1artist +
##     valence:mode + speechiness:key + speechiness:morethan1artist +
##     speechiness:explicit
##
##                              Df Sum of Sq     RSS    AIC
## <none>                                    4485.5 -14606
## - tempo:morethan1artist        1     0.890 4486.4 -14606
## - acousticness:mode            1     0.949 4486.5 -14606
## - spotifytraining$duration:mode 1    1.898 4487.4 -14603
## - valence:mode                 1     2.000 4487.5 -14602
## - tempo:explicit               1     2.172 4487.7 -14602
## - speechiness:morethan1artist  1     2.677 4488.2 -14600
## - speechiness:key             11    10.102 4495.6 -14598
## - speechiness:explicit         1     5.844 4491.3 -14591
## - sqrt(instrumentalness)       1     5.937 4491.4 -14591
## - valence:morethan1artist      1     6.836 4492.3 -14588
## - loudness:explicit            1     8.744 4494.2 -14582
## - liveness                     1    21.402 4506.9 -14544
## - danceability                 1   132.827 4618.3 -14217

anova(backAIC)

## Analysis of Variance Table
##
## Response: sqrt(popularity)
##                              Df Sum Sq Mean Sq  F value    Pr(>F)
## spotifytraining$duration      1  124.3 124.293 370.4807 < 2.2e-16 ***
## tempo                         1    2.5   2.501   7.4557  0.006332 **
## loudness                      1    0.2   0.170   0.5053  0.477174
## acousticness                  1    6.5   6.471  19.2886 1.132e-05 ***
## danceability                  1  208.4 208.377 621.1113 < 2.2e-16 ***
## sqrt(instrumentalness)        1   11.1  11.139  33.2022 8.489e-09 ***
## valence                       1   92.8  92.770 276.5203 < 2.2e-16 ***
## liveness                      1   17.7  17.739  52.8739 3.754e-13 ***
## speechiness                   1    0.2   0.235   0.7018  0.402192
## key                          11   14.1   1.282   3.8214 1.630e-05 ***
## morethan1artist               1   84.8  84.766 252.6633 < 2.2e-16 ***
## mode                          1   10.7  10.701  31.8975 1.658e-08 ***
## explicit                      1   15.6  15.610  46.5279 9.421e-12 ***
## spotifytraining$duration:mode 1    2.7   2.750   8.1966  0.004203 **
## tempo:morethan1artist         1    1.3   1.348   4.0188  0.045015 *
## tempo:explicit                1    1.3   1.269   3.7837  0.051775 .
## loudness:explicit             1    7.3   7.306  21.7783 3.090e-06 ***
## acousticness:mode             1    0.8   0.817   2.4347  0.118703
## valence:morethan1artist       1    6.9   6.931  20.6597 5.534e-06 ***
## valence:mode                  1    1.9   1.915   5.7085  0.016897 *
## speechiness:key              11    9.8   0.889   2.6490  0.002177 **
## speechiness:morethan1artist   1    3.1   3.143   9.3679  0.002213 **
## speechiness:explicit          1    5.8   5.844  17.4201 3.015e-05 ***
```

```
## Residuals                        13370 4485.5    0.335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(backAIC)$r.sq

## [1] 0.1231505

summary(backAIC)$adj.r.sq

## [1] 0.1203305

## [1]     44.00 -14606.35
```

BIC
```
## Start:  AIC=-13618.88
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##      acousticness + danceability + sqrt(instrumentalness) + valence +
##      liveness + speechiness + key + morethan1artist + mode + explicit +
##      spotifytraining$duration:key + spotifytraining$duration:mode +
##      spotifytraining$duration:explicit + tempo:key + tempo:morethan1artist
+
##      tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##      acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##      valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##      speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                     Df Sum of Sq    RSS    AIC
## - tempo:key                         11    2.4055 4456.7 -13716
## - spotifytraining$duration:key      11    3.1540 4457.5 -13714
## - danceability:key                  11    3.7012 4458.0 -13712
## - loudness:key                      11    3.8329 4458.2 -13712
## - acousticness:key                  11    4.8029 4459.1 -13709
## - valence:key                       11    5.4069 4459.7 -13707
## - sqrt(instrumentalness):key        11    6.1832 4460.5 -13705
## - speechiness:key                   11    8.7289 4463.1 -13697
## - spotifytraining$duration:explicit  1    0.2979 4454.6 -13628
## - liveness:morethan1artist           1    0.4406 4454.8 -13627
## - acousticness:mode                  1    0.7883 4455.1 -13626
## - tempo:morethan1artist              1    0.8647 4455.2 -13626
## - valence:mode                       1    1.2314 4455.6 -13625
## - spotifytraining$duration:mode      1    1.3392 4455.7 -13624
## - tempo:explicit                     1    2.1176 4456.5 -13622
## - speechiness:morethan1artist        1    2.2172 4456.6 -13622
## <none>                                          4454.3 -13619
## - valence:morethan1artist            1    5.4381 4459.8 -13612
## - speechiness:explicit               1    5.9393 4460.3 -13610
## - loudness:explicit                  1    8.4834 4462.8 -13603
##
## Step:  AIC=-13716.19
```

```
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:key + spotifytraining$duration:mode +
##     spotifytraining$duration:explicit + tempo:morethan1artist +
##     tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##     acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS     AIC
## - spotifytraining$duration:key       11    3.1683 4459.9 -13811
## - danceability:key                   11    3.5847 4460.3 -13810
## - loudness:key                       11    3.8937 4460.6 -13809
## - acousticness:key                   11    5.1785 4461.9 -13805
## - valence:key                        11    5.2527 4462.0 -13805
## - sqrt(instrumentalness):key         11    6.1855 4462.9 -13802
## - speechiness:key                    11    9.1536 4465.9 -13793
## - spotifytraining$duration:explicit   1    0.3258 4457.1 -13725
## - liveness:morethan1artist            1    0.4292 4457.2 -13724
## - acousticness:mode                   1    0.7956 4457.5 -13723
## - tempo:morethan1artist               1    0.9100 4457.6 -13723
## - valence:mode                        1    1.2421 4458.0 -13722
## - spotifytraining$duration:mode       1    1.3293 4458.1 -13722
## - tempo:explicit                      1    2.0498 4458.8 -13720
## - speechiness:morethan1artist         1    2.2414 4459.0 -13719
## <none>                                            4456.7 -13716
## - valence:morethan1artist             1    5.5502 4462.3 -13709
## - speechiness:explicit                1    5.9554 4462.7 -13708
## - loudness:explicit                   1    8.3950 4465.1 -13700
##
## Step:  AIC=-13811.2
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##     acousticness:key + acousticness:mode + danceability:key +
##     sqrt(instrumentalness):key + valence:key + valence:morethan1artist +
##     valence:mode + liveness:morethan1artist + speechiness:key +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS     AIC
## - danceability:key                   11    3.3818 4463.3 -13906
## - loudness:key                       11    3.9499 4463.9 -13904
## - acousticness:key                   11    5.0660 4465.0 -13900
## - valence:key                        11    5.3805 4465.3 -13900
## - sqrt(instrumentalness):key         11    6.1214 4466.0 -13897
```

```
## - speechiness:key                        11     9.0811 4469.0 -13888
## - liveness:morethan1artist                1     0.3998 4460.3 -13820
## - spotifytraining$duration:explicit       1     0.5200 4460.4 -13819
## - acousticness:mode                       1     0.8420 4460.7 -13818
## - tempo:morethan1artist                   1     0.9229 4460.8 -13818
## - spotifytraining$duration:mode           1     1.2937 4461.2 -13817
## - valence:mode                            1     1.3102 4461.2 -13817
## - tempo:explicit                          1     2.0620 4462.0 -13814
## - speechiness:morethan1artist             1     2.2447 4462.2 -13814
## <none>                                                4459.9 -13811
## - valence:morethan1artist                 1     5.5611 4465.5 -13804
## - speechiness:explicit                    1     6.0068 4465.9 -13803
## - loudness:explicit                       1     8.5059 4468.4 -13795
##
## Step:  AIC=-13905.58
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##     acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                          Df Sum of Sq    RSS    AIC
## - loudness:key                           11     3.768 4467.1 -13999
## - acousticness:key                       11     5.142 4468.4 -13995
## - valence:key                            11     5.474 4468.8 -13994
## - sqrt(instrumentalness):key             11     6.048 4469.3 -13992
## - speechiness:key                        11     9.140 4472.4 -13983
## - liveness:morethan1artist                1     0.382 4463.7 -13914
## - spotifytraining$duration:explicit       1     0.489 4463.8 -13914
## - acousticness:mode                       1     0.833 4464.1 -13913
## - tempo:morethan1artist                   1     0.863 4464.2 -13912
## - spotifytraining$duration:mode           1     1.233 4464.5 -13911
## - valence:mode                            1     1.346 4464.6 -13911
## - tempo:explicit                          1     2.169 4465.5 -13909
## - speechiness:morethan1artist             1     2.187 4465.5 -13908
## <none>                                                4463.3 -13906
## - valence:morethan1artist                 1     5.670 4469.0 -13898
## - speechiness:explicit                    1     5.838 4469.1 -13898
## - loudness:explicit                       1     8.361 4471.6 -13890
## - danceability                            1   132.741 4596.0 -13522
##
## Step:  AIC=-13998.8
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
```

```
##        spotifytraining$duration:mode + spotifytraining$duration:explicit +
##        tempo:morethan1artist + tempo:explicit + loudness:explicit +
##        acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##        valence:key + valence:morethan1artist + valence:mode + liveness:moreth
## an1artist +
##        speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - acousticness:key                   11     4.288 4471.3 -14090
## - sqrt(instrumentalness):key         11     6.276 4473.3 -14084
## - valence:key                        11     7.353 4474.4 -14081
## - speechiness:key                    11     9.793 4476.8 -14074
## - liveness:morethan1artist            1     0.390 4467.4 -14007
## - spotifytraining$duration:explicit   1     0.507 4467.6 -14007
## - acousticness:mode                   1     0.820 4467.9 -14006
## - tempo:morethan1artist               1     0.896 4468.0 -14006
## - spotifytraining$duration:mode       1     1.286 4468.3 -14004
## - valence:mode                        1     1.425 4468.5 -14004
## - tempo:explicit                      1     2.197 4469.3 -14002
## - speechiness:morethan1artist         1     2.324 4469.4 -14001
## <none>                                              4467.1 -13999
## - valence:morethan1artist             1     5.807 4472.9 -13991
## - speechiness:explicit                1     6.022 4473.1 -13990
## - loudness:explicit                   1     8.640 4475.7 -13982
## - danceability                        1   132.558 4599.6 -13616
##
## Step:  AIC=-14090.48
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##        acousticness + danceability + sqrt(instrumentalness) + valence +
##        liveness + speechiness + key + morethan1artist + mode + explicit +
##        spotifytraining$duration:mode + spotifytraining$duration:explicit +
##        tempo:morethan1artist + tempo:explicit + loudness:explicit +
##        acousticness:mode + sqrt(instrumentalness):key + valence:key +
##        valence:morethan1artist + valence:mode + liveness:morethan1artist +
##        speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - sqrt(instrumentalness):key         11     6.202 4477.5 -14176
## - valence:key                        11     6.631 4478.0 -14175
## - speechiness:key                    11     9.915 4481.3 -14165
## - liveness:morethan1artist            1     0.395 4471.7 -14099
## - spotifytraining$duration:explicit   1     0.514 4471.9 -14098
## - tempo:morethan1artist               1     0.851 4472.2 -14097
## - acousticness:mode                   1     1.068 4472.4 -14097
## - spotifytraining$duration:mode       1     1.356 4472.7 -14096
## - valence:mode                        1     1.417 4472.8 -14096
## - tempo:explicit                      1     2.180 4473.5 -14093
## - speechiness:morethan1artist         1     2.410 4473.8 -14093
## <none>                                              4471.3 -14090
## - valence:morethan1artist             1     5.999 4477.3 -14082
```

```
## - speechiness:explicit                          1       6.028 4477.4 -14082
## - loudness:explicit                              1       8.782 4480.1 -14074
## - danceability                                   1     132.170 4603.5 -13709
##
## Step:  AIC=-14176.43
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:key + valence:morethan1artist +
##     valence:mode + liveness:morethan1artist + speechiness:key +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - valence:key                          11     6.973 4484.5 -14260
## - speechiness:key                      11    10.394 4487.9 -14250
## - liveness:morethan1artist              1     0.360 4477.9 -14185
## - spotifytraining$duration:explicit     1     0.520 4478.1 -14184
## - tempo:morethan1artist                 1     0.919 4478.5 -14183
## - acousticness:mode                     1     0.970 4478.5 -14183
## - spotifytraining$duration:mode         1     1.394 4478.9 -14182
## - valence:mode                          1     1.463 4479.0 -14182
## - tempo:explicit                        1     2.195 4479.7 -14179
## - speechiness:morethan1artist           1     2.455 4480.0 -14179
## <none>                                              4477.5 -14176
## - speechiness:explicit                  1     6.007 4483.6 -14168
## - valence:morethan1artist               1     6.074 4483.6 -14168
## - sqrt(instrumentalness)                1     6.098 4483.6 -14168
## - loudness:explicit                     1     8.878 4486.4 -14159
## - danceability                          1   132.696 4610.2 -13794
##
## Step:  AIC=-14260.1
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     liveness:morethan1artist + speechiness:key + speechiness:morethan1arti
st +
##     speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - speechiness:key                      11    10.288 4494.8 -14334
## - liveness:morethan1artist              1     0.406 4484.9 -14268
## - spotifytraining$duration:explicit     1     0.561 4485.1 -14268
## - tempo:morethan1artist                 1     0.901 4485.4 -14267
## - acousticness:mode                     1     0.941 4485.5 -14267
## - spotifytraining$duration:mode         1     1.389 4485.9 -14265
```

```
## - valence:mode                            1      1.974 4486.5 -14264
## - tempo:explicit                          1      2.133 4486.7 -14263
## - speechiness:morethan1artist             1      2.385 4486.9 -14262
## <none>                                                  4484.5 -14260
## - speechiness:explicit                     1      5.939 4490.5 -14252
## - sqrt(instrumentalness)                   1      6.107 4490.6 -14251
## - valence:morethan1artist                  1      6.602 4491.1 -14250
## - loudness:explicit                        1      8.573 4493.1 -14244
## - danceability                             1    131.653 4616.2 -13882
##
## Step:   AIC=-14333.9
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##      acousticness + danceability + sqrt(instrumentalness) + valence +
##      liveness + speechiness + key + morethan1artist + mode + explicit +
##      spotifytraining$duration:mode + spotifytraining$duration:explicit +
##      tempo:morethan1artist + tempo:explicit + loudness:explicit +
##      acousticness:mode + valence:morethan1artist + valence:mode +
##      liveness:morethan1artist + speechiness:morethan1artist +
##      speechiness:explicit
##
##                                         Df Sum of Sq    RSS    AIC
## - key                                   11      9.069 4503.9 -14411
## - liveness:morethan1artist               1      0.333 4495.1 -14342
## - spotifytraining$duration:explicit      1      0.450 4495.3 -14342
## - tempo:morethan1artist                  1      0.826 4495.6 -14341
## - acousticness:mode                      1      1.009 4495.8 -14340
## - spotifytraining$duration:mode          1      1.271 4496.1 -14340
## - speechiness:morethan1artist            1      1.978 4496.8 -14338
## - valence:mode                           1      1.979 4496.8 -14338
## - tempo:explicit                         1      2.269 4497.1 -14337
## <none>                                                 4494.8 -14334
## - speechiness:explicit                   1      6.072 4500.9 -14325
## - sqrt(instrumentalness)                 1      6.207 4501.0 -14325
## - valence:morethan1artist                1      6.383 4501.2 -14324
## - loudness:explicit                      1      8.241 4503.0 -14319
## - danceability                           1    133.286 4628.1 -13951
##
## Step:   AIC=-14411.41
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##      acousticness + danceability + sqrt(instrumentalness) + valence +
##      liveness + speechiness + morethan1artist + mode + explicit +
##      spotifytraining$duration:mode + spotifytraining$duration:explicit +
##      tempo:morethan1artist + tempo:explicit + loudness:explicit +
##      acousticness:mode + valence:morethan1artist + valence:mode +
##      liveness:morethan1artist + speechiness:morethan1artist +
##      speechiness:explicit
##
##                                         Df Sum of Sq    RSS    AIC
## - liveness:morethan1artist               1      0.334 4504.2 -14420
## - spotifytraining$duration:explicit      1      0.400 4504.3 -14420
```

```
## - acousticness:mode                    1      0.816 4504.7 -14418
## - tempo:morethan1artist                 1      0.826 4504.7 -14418
## - spotifytraining$duration:mode         1      1.458 4505.3 -14417
## - speechiness:morethan1artist           1      2.057 4505.9 -14415
## - valence:mode                          1      2.098 4506.0 -14415
## - tempo:explicit                        1      2.299 4506.2 -14414
## <none>                                              4503.9 -14411
## - speechiness:explicit                  1      6.194 4510.1 -14402
## - sqrt(instrumentalness)                1      6.249 4510.1 -14402
## - valence:morethan1artist               1      6.474 4510.4 -14402
## - loudness:explicit                     1      8.509 4512.4 -14396
## - danceability                          1    135.231 4639.1 -14024
##
## Step:  AIC=-14419.92
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:explicit  1      0.413 4504.6 -14428
## - tempo:morethan1artist              1      0.826 4505.0 -14427
## - acousticness:mode                  1      0.832 4505.0 -14427
## - spotifytraining$duration:mode      1      1.458 4505.7 -14425
## - valence:mode                       1      2.095 4506.3 -14423
## - speechiness:morethan1artist        1      2.258 4506.5 -14423
## - tempo:explicit                     1      2.299 4506.5 -14423
## <none>                                            4504.2 -14420
## - speechiness:explicit               1      6.245 4510.5 -14411
## - sqrt(instrumentalness)             1      6.261 4510.5 -14411
## - valence:morethan1artist            1      6.421 4510.6 -14410
## - loudness:explicit                  1      8.609 4512.8 -14404
## - liveness                           1     22.012 4526.2 -14364
## - danceability                       1    135.309 4639.5 -14032
##
## Step:  AIC=-14428.2
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:morethan1artist + tempo:explicit
+
##     loudness:explicit + acousticness:mode + valence:morethan1artist +
##     valence:mode + speechiness:morethan1artist + speechiness:explicit
##
##                              Df Sum of Sq    RSS    AIC
## - tempo:morethan1artist       1      0.820 4505.4 -14435
## - acousticness:mode           1      0.823 4505.4 -14435
```

```
## - spotifytraining$duration:mode   1      1.907 4506.5 -14432
## - valence:mode                     1      2.120 4506.7 -14431
## - speechiness:morethan1artist      1      2.307 4506.9 -14431
## - tempo:explicit                   1      2.329 4507.0 -14431
## <none>                                          4504.6 -14428
## - sqrt(instrumentalness)           1      6.104 4510.7 -14420
## - speechiness:explicit             1      6.109 4510.7 -14420
## - valence:morethan1artist          1      6.668 4511.3 -14418
## - loudness:explicit                1      8.671 4513.3 -14412
## - liveness                         1     22.126 4526.7 -14372
## - danceability                     1    136.211 4640.8 -14038
##
## Step:  AIC=-14435.26
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                   Df Sum of Sq    RSS    AIC
## - acousticness:mode                1      0.826 4506.3 -14442
## - spotifytraining$duration:mode    1      1.900 4507.3 -14439
## - valence:mode                     1      2.110 4507.6 -14438
## - speechiness:morethan1artist      1      2.163 4507.6 -14438
## - tempo:explicit                   1      3.121 4508.6 -14436
## <none>                                          4505.4 -14435
## - sqrt(instrumentalness)           1      6.056 4511.5 -14427
## - speechiness:explicit             1      6.215 4511.7 -14426
## - valence:morethan1artist          1      6.720 4512.2 -14425
## - loudness:explicit                1      8.657 4514.1 -14419
## - liveness                         1     22.166 4527.6 -14379
## - danceability                     1    136.935 4642.4 -14043
##
## Step:  AIC=-14442.31
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     valence:morethan1artist + valence:mode + speechiness:morethan1artist +
##     speechiness:explicit
##
##                                   Df Sum of Sq    RSS    AIC
## - valence:mode                     1      1.871 4508.1 -14446
## - spotifytraining$duration:mode    1      2.059 4508.3 -14446
## - speechiness:morethan1artist      1      2.190 4508.5 -14445
## - tempo:explicit                   1      3.101 4509.4 -14443
## <none>                                          4506.3 -14442
## - sqrt(instrumentalness)           1      5.872 4512.1 -14434
## - speechiness:explicit             1      6.282 4512.6 -14433
```

```
## - valence:morethan1artist            1      6.728 4513.0 -14432
## - loudness:explicit                   1      9.017 4515.3 -14425
## - liveness                            1     22.364 4528.6 -14385
## - acousticness                        1     29.587 4535.9 -14364
## - danceability                        1    137.054 4643.3 -14050
##
## Step:  AIC=-14446.24
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     valence:morethan1artist + speechiness:morethan1artist + speechiness:ex
plicit
##
##                                     Df Sum of Sq    RSS    AIC
## - speechiness:morethan1artist        1      2.210 4510.3 -14449
## - spotifytraining$duration:mode      1      2.939 4511.1 -14447
## - tempo:explicit                     1      3.169 4511.3 -14446
## <none>                                            4508.1 -14446
## - sqrt(instrumentalness)             1      5.848 4514.0 -14438
## - speechiness:explicit               1      6.162 4514.3 -14437
## - valence:morethan1artist            1      7.365 4515.5 -14434
## - loudness:explicit                  1      8.684 4516.8 -14430
## - liveness                           1     22.456 4530.6 -14389
## - acousticness                       1     29.603 4537.7 -14368
## - danceability                       1    137.395 4645.5 -14053
##
## Step:  AIC=-14449.17
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     valence:morethan1artist + speechiness:explicit
##
##                                     Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:mode      1      3.006 4513.4 -14450
## <none>                                            4510.3 -14449
## - tempo:explicit                     1      3.229 4513.6 -14449
## - sqrt(instrumentalness)             1      5.937 4516.3 -14441
## - speechiness:explicit               1      6.616 4517.0 -14439
## - valence:morethan1artist            1      7.035 4517.4 -14438
## - loudness:explicit                  1      8.991 4519.3 -14432
## - liveness                           1     21.947 4532.3 -14394
## - acousticness                       1     30.158 4540.5 -14369
## - danceability                       1    138.727 4649.1 -14052
##
## Step:  AIC=-14449.74
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
```

```
##     tempo:explicit + loudness:explicit + valence:morethan1artist +
##     speechiness:explicit
##
##                              Df Sum of Sq    RSS    AIC
## <none>                                    4513.4 -14450
## - tempo:explicit            1     3.241 4516.6 -14450
## - sqrt(instrumentalness)    1     6.116 4519.5 -14441
## - speechiness:explicit      1     6.570 4519.9 -14440
## - valence:morethan1artist   1     7.167 4520.5 -14438
## - loudness:explicit         1     8.896 4522.3 -14433
## - mode                      1     9.376 4522.7 -14431
## - liveness                  1    21.527 4534.9 -14395
## - acousticness              1    30.003 4543.4 -14370
## - spotifytraining$duration  1    64.607 4578.0 -14269
## - danceability              1   139.006 4652.4 -14052

anova(backBIC)

## Analysis of Variance Table
##
## Response: sqrt(popularity)
##                              Df Sum Sq Mean Sq  F value    Pr(>F)
## spotifytraining$duration      1  124.3 124.293 368.9380 < 2.2e-16 ***
## tempo                         1    2.5   2.501   7.4247  0.006442 **
## loudness                      1    0.2   0.170   0.5032  0.478093
## acousticness                  1    6.5   6.471  19.2083 1.181e-05 ***
## danceability                  1  208.4 208.377 618.5250 < 2.2e-16 ***
## sqrt(instrumentalness)        1   11.1  11.139  33.0639 9.112e-09 ***
## valence                       1   92.8  92.770 275.3688 < 2.2e-16 ***
## liveness                      1   17.7  17.739  52.6538 4.197e-13 ***
## speechiness                   1    0.2   0.235   0.6989  0.403173
## morethan1artist               1   86.1  86.092 255.5466 < 2.2e-16 ***
## mode                          1   13.1  13.120  38.9451 4.490e-10 ***
## explicit                      1   16.2  16.219  48.1422 4.146e-12 ***
## tempo:explicit                1    1.9   1.863   5.5295  0.018713 *
## loudness:explicit             1    7.4   7.378  21.8993 2.902e-06 ***
## valence:morethan1artist       1    7.2   7.185  21.3287 3.905e-06 ***
## speechiness:explicit          1    6.6   6.570  19.5030 1.012e-05 ***
## Residuals                 13397 4513.4   0.337
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(backBIC)$r.sq

## [1] 0.1177059

summary(backBIC)$adj.r.sq

## [1] 0.1166522

## [1]    17.00 -14577.31
```

Stepwise

AIC

```
## Start:  AIC=-14541.88
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:key + spotifytraining$duration:mode +
##     spotifytraining$duration:explicit + tempo:key + tempo:morethan1artist
+
##     tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##     acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                       Df Sum of Sq    RSS    AIC
## - tempo:key                           11    2.4055 4456.7 -14557
## - spotifytraining$duration:key        11    3.1540 4457.5 -14554
## - danceability:key                    11    3.7012 4458.0 -14553
## - loudness:key                        11    3.8329 4458.2 -14552
## - acousticness:key                    11    4.8029 4459.1 -14549
## - valence:key                         11    5.4069 4459.7 -14548
## - sqrt(instrumentalness):key          11    6.1832 4460.5 -14545
## - spotifytraining$duration:explicit    1    0.2979 4454.6 -14543
## - liveness:morethan1artist             1    0.4406 4454.8 -14543
## <none>                                              4454.3 -14542
## - acousticness:mode                    1    0.7883 4455.1 -14542
## - tempo:morethan1artist                1    0.8647 4455.2 -14541
## - valence:mode                         1    1.2314 4455.6 -14540
## - spotifytraining$duration:mode        1    1.3392 4455.7 -14540
## - speechiness:key                     11    8.7289 4463.1 -14538
## - tempo:explicit                       1    2.1176 4456.5 -14538
## - speechiness:morethan1artist          1    2.2172 4456.6 -14537
## - valence:morethan1artist              1    5.4381 4459.8 -14528
## - speechiness:explicit                 1    5.9393 4460.3 -14526
## - loudness:explicit                    1    8.4834 4462.8 -14518
##
## Step:  AIC=-14556.64
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:key + spotifytraining$duration:mode +
##     spotifytraining$duration:explicit + tempo:morethan1artist +
##     tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##     acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
```

```
##                                      Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:key       11    3.1683 4459.9 -14569
## - danceability:key                   11    3.5847 4460.3 -14568
## - loudness:key                       11    3.8937 4460.6 -14567
## - acousticness:key                   11    5.1785 4461.9 -14563
## - valence:key                        11    5.2527 4462.0 -14563
## - sqrt(instrumentalness):key         11    6.1855 4462.9 -14560
## - spotifytraining$duration:explicit   1    0.3258 4457.1 -14558
## - liveness:morethan1artist            1    0.4292 4457.2 -14557
## <none>                                             4456.7 -14557
## - acousticness:mode                   1    0.7956 4457.5 -14556
## - tempo:morethan1artist               1    0.9100 4457.6 -14556
## - valence:mode                        1    1.2421 4458.0 -14555
## - spotifytraining$duration:mode       1    1.3293 4458.1 -14555
## - tempo:explicit                      1    2.0498 4458.8 -14552
## - speechiness:morethan1artist         1    2.2414 4459.0 -14552
## - speechiness:key                    11    9.1536 4465.9 -14551
## - valence:morethan1artist             1    5.5502 4462.3 -14542
## + tempo:key                          11    2.4055 4454.3 -14542
## - speechiness:explicit                1    5.9554 4462.7 -14541
## - loudness:explicit                   1    8.3950 4465.1 -14533
## 
## Step:  AIC=-14569.11
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##     acousticness:key + acousticness:mode + danceability:key +
##     sqrt(instrumentalness):key + valence:key + valence:morethan1artist +
##     valence:mode + liveness:morethan1artist + speechiness:key +
##     speechiness:morethan1artist + speechiness:explicit
## 
##                                      Df Sum of Sq    RSS    AIC
## - danceability:key                   11    3.3818 4463.3 -14581
## - loudness:key                       11    3.9499 4463.9 -14579
## - acousticness:key                   11    5.0660 4465.0 -14576
## - valence:key                        11    5.3805 4465.3 -14575
## - sqrt(instrumentalness):key         11    6.1214 4466.0 -14573
## - liveness:morethan1artist            1    0.3998 4460.3 -14570
## - spotifytraining$duration:explicit   1    0.5200 4460.4 -14570
## <none>                                             4459.9 -14569
## - acousticness:mode                   1    0.8420 4460.7 -14569
## - tempo:morethan1artist               1    0.9229 4460.8 -14568
## - spotifytraining$duration:mode       1    1.2937 4461.2 -14567
## - valence:mode                        1    1.3102 4461.2 -14567
## - tempo:explicit                      1    2.0620 4462.0 -14565
## - speechiness:morethan1artist         1    2.2447 4462.2 -14564
## - speechiness:key                    11    9.0811 4469.0 -14564
```

```
## + spotifytraining$duration:key         11    3.1683 4456.7 -14557
## - valence:morethan1artist               1    5.5611 4465.5 -14554
## + tempo:key                             11    2.4198 4457.5 -14554
## - speechiness:explicit                  1    6.0068 4465.9 -14553
## - loudness:explicit                     1    8.5059 4468.4 -14546
##
## Step:  AIC=-14580.94
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##     acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - loudness:key                       11     3.768 4467.1 -14592
## - acousticness:key                   11     5.142 4468.4 -14588
## - valence:key                        11     5.474 4468.8 -14586
## - sqrt(instrumentalness):key         11     6.048 4469.3 -14585
## - liveness:morethan1artist            1     0.382 4463.7 -14582
## - spotifytraining$duration:explicit   1     0.489 4463.8 -14582
## <none>                                            4463.3 -14581
## - acousticness:mode                   1     0.833 4464.1 -14580
## - tempo:morethan1artist               1     0.863 4464.2 -14580
## - spotifytraining$duration:mode       1     1.233 4464.5 -14579
## - valence:mode                        1     1.346 4464.6 -14579
## - tempo:explicit                      1     2.169 4465.5 -14576
## - speechiness:morethan1artist         1     2.187 4465.5 -14576
## - speechiness:key                    11     9.140 4472.4 -14576
## + danceability:key                   11     3.382 4459.9 -14569
## + spotifytraining$duration:key        11     2.965 4460.3 -14568
## + tempo:key                          11     2.369 4460.9 -14566
## - valence:morethan1artist             1     5.670 4469.0 -14566
## - speechiness:explicit                1     5.838 4469.1 -14565
## - loudness:explicit                   1     8.361 4471.6 -14558
## - danceability                        1   132.741 4596.0 -14190
##
## Step:  AIC=-14591.62
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
```

```
##      speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - acousticness:key                     11     4.288 4471.3 -14601
## - sqrt(instrumentalness):key           11     6.276 4473.3 -14595
## - liveness:morethan1artist              1     0.390 4467.4 -14592
## - spotifytraining$duration:explicit     1     0.507 4467.6 -14592
## <none>                                             4467.1 -14592
## - valence:key                          11     7.353 4474.4 -14592
## - acousticness:mode                     1     0.820 4467.9 -14591
## - tempo:morethan1artist                 1     0.896 4468.0 -14591
## - spotifytraining$duration:mode         1     1.286 4468.3 -14590
## - valence:mode                          1     1.425 4468.5 -14589
## - tempo:explicit                        1     2.197 4469.3 -14587
## - speechiness:morethan1artist           1     2.324 4469.4 -14587
## - speechiness:key                      11     9.793 4476.8 -14584
## + loudness:key                         11     3.768 4463.3 -14581
## + danceability:key                     11     3.200 4463.9 -14579
## + spotifytraining$duration:key         11     3.133 4463.9 -14579
## + tempo:key                            11     2.354 4464.7 -14577
## - valence:morethan1artist               1     5.807 4472.9 -14576
## - speechiness:explicit                  1     6.022 4473.1 -14576
## - loudness:explicit                     1     8.640 4475.7 -14568
## - danceability                          1   132.558 4599.6 -14201
##
## Step:  AIC=-14600.75
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + sqrt(instrumentalness):key + valence:key +
##     valence:morethan1artist + valence:mode + liveness:morethan1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - sqrt(instrumentalness):key           11     6.202 4477.5 -14604
## - valence:key                          11     6.631 4478.0 -14603
## - liveness:morethan1artist              1     0.395 4471.7 -14602
## - spotifytraining$duration:explicit     1     0.514 4471.9 -14601
## <none>                                             4471.3 -14601
## - tempo:morethan1artist                 1     0.851 4472.2 -14600
## - acousticness:mode                     1     1.068 4472.4 -14600
## - spotifytraining$duration:mode         1     1.356 4472.7 -14599
## - valence:mode                          1     1.417 4472.8 -14598
## - tempo:explicit                        1     2.180 4473.5 -14596
## - speechiness:morethan1artist           1     2.410 4473.8 -14596
## - speechiness:key                      11     9.915 4481.3 -14593
## + acousticness:key                     11     4.288 4467.1 -14592
## + danceability:key                     11     3.274 4468.1 -14589
```

```
## + spotifytraining$duration:key       11      3.053 4468.3 -14588
## + loudness:key                        11      2.913 4468.4 -14588
## + tempo:key                           11      2.784 4468.6 -14587
## - valence:morethan1artist              1      5.999 4477.3 -14585
## - speechiness:explicit                 1      6.028 4477.4 -14585
## - loudness:explicit                    1      8.782 4480.1 -14576
## - danceability                         1    132.170 4603.5 -14212
##
## Step:  AIC=-14604.16
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:key + valence:morethan1artist +
##     valence:mode + liveness:morethan1artist + speechiness:key +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                       Df Sum of Sq    RSS    AIC
## - valence:key                         11      6.973 4484.5 -14605
## - liveness:morethan1artist             1      0.360 4477.9 -14605
## - spotifytraining$duration:explicit    1      0.520 4478.1 -14605
## <none>                                              4477.5 -14604
## - tempo:morethan1artist                1      0.919 4478.5 -14603
## - acousticness:mode                    1      0.970 4478.5 -14603
## - spotifytraining$duration:mode        1      1.394 4478.9 -14602
## - valence:mode                         1      1.463 4479.0 -14602
## + sqrt(instrumentalness):key          11      6.202 4471.3 -14601
## - tempo:explicit                       1      2.195 4479.7 -14600
## - speechiness:morethan1artist          1      2.455 4480.0 -14599
## - speechiness:key                     11     10.394 4487.9 -14595
## + acousticness:key                    11      4.214 4473.3 -14595
## + danceability:key                    11      3.225 4474.3 -14592
## + loudness:key                        11      2.998 4474.5 -14591
## + spotifytraining$duration:key        11      2.949 4474.6 -14591
## + tempo:key                           11      2.822 4474.7 -14591
## - speechiness:explicit                 1      6.007 4483.6 -14588
## - valence:morethan1artist              1      6.074 4483.6 -14588
## - sqrt(instrumentalness)               1      6.098 4483.6 -14588
## - loudness:explicit                    1      8.878 4486.4 -14580
## - danceability                         1    132.696 4610.2 -14214
##
## Step:  AIC=-14605.28
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     liveness:morethan1artist + speechiness:key + speechiness:morethan1arti
```

```
st +
##     speechiness:explicit
##
##                                     Df Sum of Sq    RSS    AIC
## - liveness:morethan1artist          1      0.406 4484.9 -14606
## - spotifytraining$duration:explicit 1      0.561 4485.1 -14606
## <none>                                          4484.5 -14605
## - tempo:morethan1artist             1      0.901 4485.4 -14605
## - acousticness:mode                 1      0.941 4485.5 -14604
## + valence:key                      11      6.973 4477.5 -14604
## - spotifytraining$duration:mode     1      1.389 4485.9 -14603
## + sqrt(instrumentalness):key       11      6.544 4478.0 -14603
## - valence:mode                      1      1.974 4486.5 -14601
## - tempo:explicit                    1      2.133 4486.7 -14601
## - speechiness:morethan1artist       1      2.385 4486.9 -14600
## - speechiness:key                  11     10.288 4494.8 -14596
## + loudness:key                     11      3.829 4480.7 -14595
## + danceability:key                 11      3.532 4481.0 -14594
## + acousticness:key                 11      3.484 4481.0 -14594
## + spotifytraining$duration:key     11      2.830 4481.7 -14592
## + tempo:key                        11      2.707 4481.8 -14591
## - speechiness:explicit              1      5.939 4490.5 -14590
## - sqrt(instrumentalness)            1      6.107 4490.6 -14589
## - valence:morethan1artist           1      6.602 4491.1 -14588
## - loudness:explicit                 1      8.573 4493.1 -14582
## - danceability                      1    131.653 4616.2 -14219
##
## Step:  AIC=-14606.07
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                     Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:explicit 1      0.577 4485.5 -14606
## <none>                                          4484.9 -14606
## - tempo:morethan1artist             1      0.899 4485.8 -14605
## + liveness:morethan1artist          1      0.406 4484.5 -14605
## - acousticness:mode                 1      0.961 4485.9 -14605
## + valence:key                      11      7.020 4477.9 -14605
## - spotifytraining$duration:mode     1      1.388 4486.3 -14604
## + sqrt(instrumentalness):key       11      6.511 4478.4 -14604
## - valence:mode                      1      1.971 4486.9 -14602
## - tempo:explicit                    1      2.135 4487.1 -14602
## - speechiness:morethan1artist       1      2.621 4487.5 -14600
## - speechiness:key                  11     10.215 4495.1 -14598
## + loudness:key                     11      3.846 4481.1 -14596
```

```
## + danceability:key                      11       3.539 4481.4 -14595
## + acousticness:key                      11       3.489 4481.4 -14594
## + spotifytraining$duration:key          11       2.833 4482.1 -14592
## + tempo:key                             11       2.690 4482.2 -14592
## - speechiness:explicit                   1       6.000 4490.9 -14590
## - sqrt(instrumentalness)                 1       6.121 4491.0 -14590
## - valence:morethan1artist                1       6.544 4491.5 -14588
## - loudness:explicit                      1       8.680 4493.6 -14582
## - liveness                               1      21.272 4506.2 -14545
## - danceability                           1     131.762 4616.7 -14220
##
## Step:   AIC=-14606.35
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##      acousticness + danceability + sqrt(instrumentalness) + valence +
##      liveness + speechiness + key + morethan1artist + mode + explicit +
##      spotifytraining$duration:mode + tempo:morethan1artist + tempo:explicit
+
##      loudness:explicit + acousticness:mode + valence:morethan1artist +
##      valence:mode + speechiness:key + speechiness:morethan1artist +
##      speechiness:explicit
##
##                                         Df Sum of Sq    RSS     AIC
## <none>                                                 4485.5 -14606
## + spotifytraining$duration:explicit  1     0.577 4484.9 -14606
## - tempo:morethan1artist                  1     0.890 4486.4 -14606
## + liveness:morethan1artist               1     0.422 4485.1 -14606
## - acousticness:mode                      1     0.949 4486.5 -14606
## + valence:key                           11     7.062 4478.4 -14606
## + sqrt(instrumentalness):key            11     6.511 4479.0 -14604
## - spotifytraining$duration:mode          1     1.898 4487.4 -14603
## - valence:mode                           1     2.000 4487.5 -14602
## - tempo:explicit                         1     2.172 4487.7 -14602
## - speechiness:morethan1artist            1     2.677 4488.2 -14600
## - speechiness:key                       11    10.102 4495.6 -14598
## + loudness:key                          11     3.851 4481.7 -14596
## + danceability:key                      11     3.518 4482.0 -14595
## + acousticness:key                      11     3.493 4482.0 -14595
## + spotifytraining$duration:key          11     3.071 4482.4 -14594
## + tempo:key                             11     2.725 4482.8 -14592
## - speechiness:explicit                   1     5.844 4491.3 -14591
## - sqrt(instrumentalness)                 1     5.937 4491.4 -14591
## - valence:morethan1artist                1     6.836 4492.3 -14588
## - loudness:explicit                      1     8.744 4494.2 -14582
## - liveness                               1    21.402 4506.9 -14544
## - danceability                           1   132.827 4618.3 -14217

anova(stepwiseAIC)

## Analysis of Variance Table
##
```

```
## Response: sqrt(popularity)
##                                    Df  Sum Sq Mean Sq  F value    Pr(>F)
## spotifytraining$duration            1   124.3 124.293 370.4807 < 2.2e-16 ***
## tempo                               1     2.5   2.501   7.4557  0.006332 **
## loudness                            1     0.2   0.170   0.5053  0.477174
## acousticness                        1     6.5   6.471  19.2886 1.132e-05 ***
## danceability                        1   208.4 208.377 621.1113 < 2.2e-16 ***
## sqrt(instrumentalness)              1    11.1  11.139  33.2022 8.489e-09 ***
## valence                             1    92.8  92.770 276.5203 < 2.2e-16 ***
## liveness                            1    17.7  17.739  52.8739 3.754e-13 ***
## speechiness                         1     0.2   0.235   0.7018  0.402192
## key                                11    14.1   1.282   3.8214 1.630e-05 ***
## morethan1artist                     1    84.8  84.766 252.6633 < 2.2e-16 ***
## mode                                1    10.7  10.701  31.8975 1.658e-08 ***
## explicit                            1    15.6  15.610  46.5279 9.421e-12 ***
## spotifytraining$duration:mode       1     2.7   2.750   8.1966  0.004203 **
## tempo:morethan1artist               1     1.3   1.348   4.0188  0.045015 *
## tempo:explicit                      1     1.3   1.269   3.7837  0.051775 .
## loudness:explicit                   1     7.3   7.306  21.7783 3.090e-06 ***
## acousticness:mode                   1     0.8   0.817   2.4347  0.118703
## valence:morethan1artist             1     6.9   6.931  20.6597 5.534e-06 ***
## valence:mode                        1     1.9   1.915   5.7085  0.016897 *
## speechiness:key                    11     9.8   0.889   2.6490  0.002177 **
## speechiness:morethan1artist         1     3.1   3.143   9.3679  0.002213 **
## speechiness:explicit                1     5.8   5.844  17.4201 3.015e-05 ***
## Residuals                       13370  4485.5   0.335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(stepwiseAIC)$coefficients[, "Estimate"]

##               (Intercept)      spotifytraining$duration
##              7.767951e+00                 -9.364216e-02
##                     tempo                      loudness
##             -3.565109e-05                  1.973806e-02
##              acousticness                  danceability
##              2.658353e-01                  7.896162e-01
##    sqrt(instrumentalness)                       valence
##             -2.132366e-01                 -2.905865e-01
##                  liveness                   speechiness
##             -2.757123e-01                  2.287489e-01
##                      key1                          key2
##              3.455087e-02                 -2.873166e-02
##                      key3                          key4
##             -4.858325e-02                  4.896025e-02
##                      key5                          key6
##              1.015961e-01                  7.248942e-02
##                      key7                          key8
##             -3.574053e-02                  2.526435e-02
##                      key9                         key10
```

```
##                  -2.714305e-02                       -2.405045e-02
##                          key11                    morethan1artist1
##                   3.027273e-02                        1.923002e-02
##                          mode1                            explicit1
##                  -7.977838e-02                       -1.155144e-01
## spotifytraining$duration:mode1               tempo:morethan1artist1
##                   2.665946e-02                        6.555610e-04
##                 tempo:explicit1                    loudness:explicit1
##                   9.558513e-04                       -2.190629e-02
##              acousticness:mode1            valence:morethan1artist1
##                  -7.119364e-02                        2.313030e-01
##                   valence:mode1                    speechiness:key1
##                  -1.140180e-01                       -2.047157e-01
##               speechiness:key2                    speechiness:key3
##                   1.337637e-02                        1.826353e-01
##               speechiness:key4                    speechiness:key5
##                  -7.246027e-01                       -4.313545e-01
##               speechiness:key6                    speechiness:key7
##                  -3.530245e-01                        1.543365e-01
##               speechiness:key8                    speechiness:key9
##                  -4.839646e-02                        3.179464e-03
##              speechiness:key10                   speechiness:key11
##                   1.503345e-01                       -1.838087e-01
##    speechiness:morethan1artist1               speechiness:explicit1
##                  -2.740968e-01                       -4.446127e-01
```

```
summary(stepwiseAIC)$r.sq
```

```
## [1] 0.1231505
```

```
summary(stepwiseAIC)$adj.r.sq
```

```
## [1] 0.1203305
```
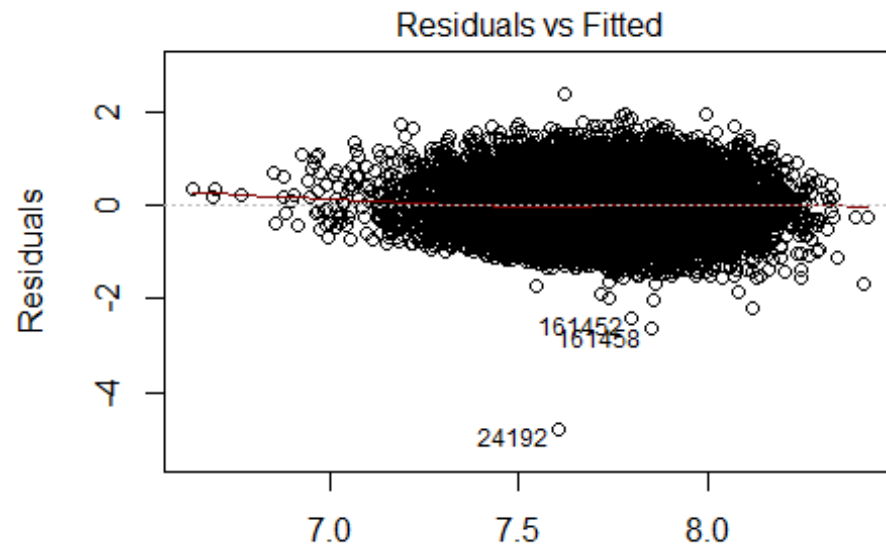
```
## [1]     44.00 -14606.35
```

Residuals vs Fitted

Residuals

161452
161458

24192

Fitted values
m(sqrt(popularity) ~ spotifytraining$duration + tempo + loudness + aco

Normal Q-Q

Standardized residuals

161452
161458

24192

Theoretical Quantiles
m(sqrt(popularity) ~ spotifytraining$duration + tempo + loudness + aco

Scale-Location

24192 ○

161458 ○
161452 ○

√|Standardized residuals|

Fitted values
m(sqrt(popularity) ~ spotifytraining$duration + tempo + loudness + aco

Residuals vs Leverage

Standardized residuals

62 142 376943 ○

---- Cook's distance
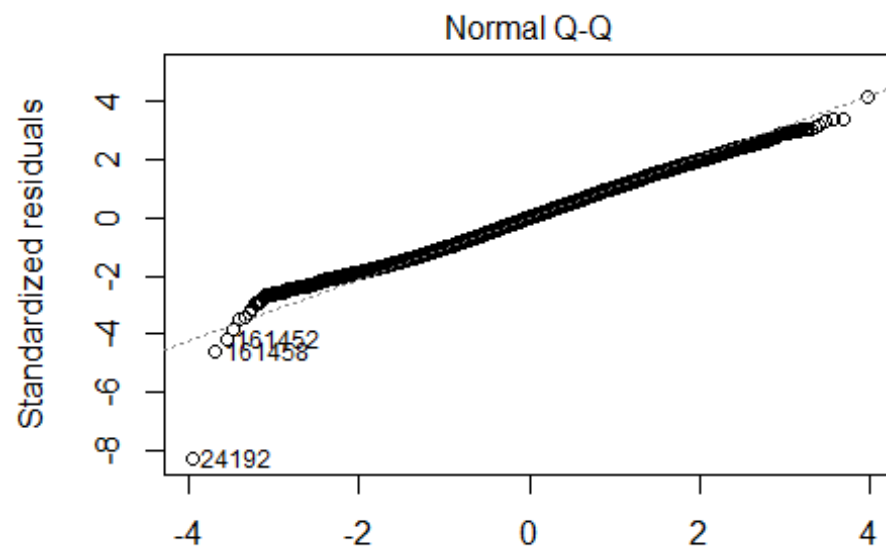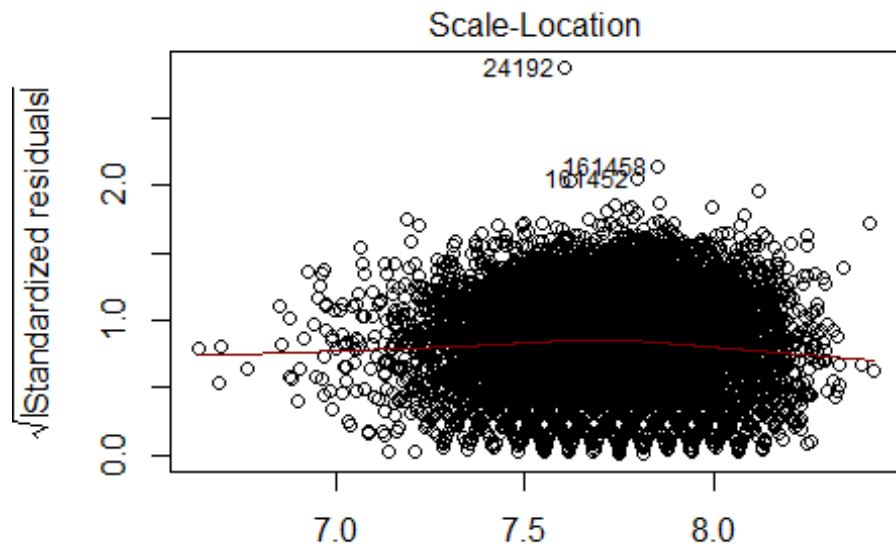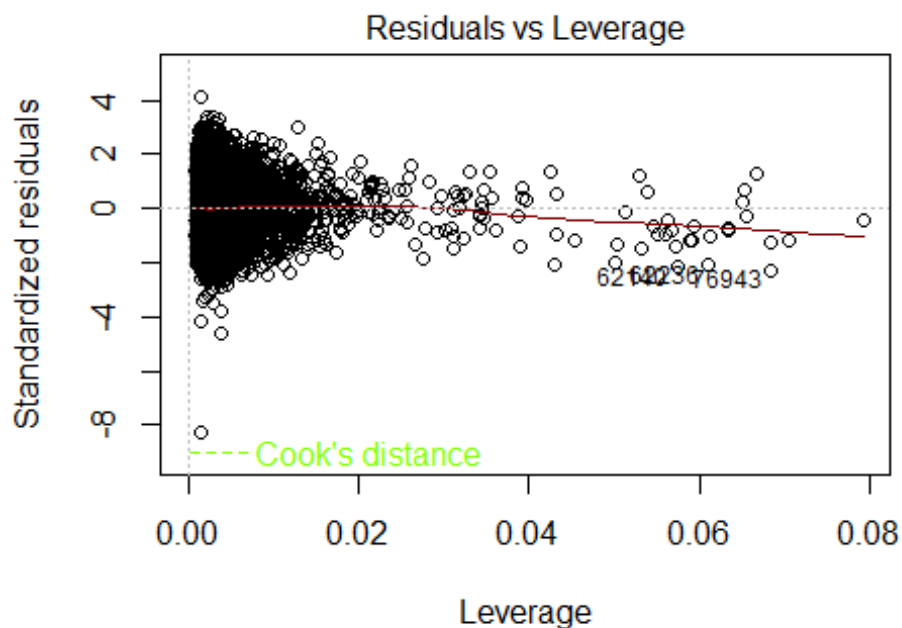
Leverage
m(sqrt(popularity) ~ spotifytraining$duration + tempo + loudness + aco

BIC
```
## Start:  AIC=-13618.88
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##      acousticness + danceability + sqrt(instrumentalness) + valence +
```

```
##       liveness + speechiness + key + morethan1artist + mode + explicit +
##       spotifytraining$duration:key + spotifytraining$duration:mode +
##       spotifytraining$duration:explicit + tempo:key + tempo:morethan1artist
+
##       tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##       acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##       valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##       speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - tempo:key                          11    2.4055 4456.7 -13716
## - spotifytraining$duration:key       11    3.1540 4457.5 -13714
## - danceability:key                   11    3.7012 4458.0 -13712
## - loudness:key                       11    3.8329 4458.2 -13712
## - acousticness:key                   11    4.8029 4459.1 -13709
## - valence:key                        11    5.4069 4459.7 -13707
## - sqrt(instrumentalness):key         11    6.1832 4460.5 -13705
## - speechiness:key                    11    8.7289 4463.1 -13697
## - spotifytraining$duration:explicit   1    0.2979 4454.6 -13628
## - liveness:morethan1artist            1    0.4406 4454.8 -13627
## - acousticness:mode                   1    0.7883 4455.1 -13626
## - tempo:morethan1artist               1    0.8647 4455.2 -13626
## - valence:mode                        1    1.2314 4455.6 -13625
## - spotifytraining$duration:mode       1    1.3392 4455.7 -13624
## - tempo:explicit                      1    2.1176 4456.5 -13622
## - speechiness:morethan1artist         1    2.2172 4456.6 -13622
## <none>                                              4454.3 -13619
## - valence:morethan1artist             1    5.4381 4459.8 -13612
## - speechiness:explicit                1    5.9393 4460.3 -13610
## - loudness:explicit                   1    8.4834 4462.8 -13603
##
## Step:  AIC=-13716.19
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##       acousticness + danceability + sqrt(instrumentalness) + valence +
##       liveness + speechiness + key + morethan1artist + mode + explicit +
##       spotifytraining$duration:key + spotifytraining$duration:mode +
##       spotifytraining$duration:explicit + tempo:morethan1artist +
##       tempo:explicit + loudness:key + loudness:explicit + acousticness:key +
##       acousticness:mode + danceability:key + sqrt(instrumentalness):key +
##       valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##       speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:key       11    3.1683 4459.9 -13811
## - danceability:key                   11    3.5847 4460.3 -13810
## - loudness:key                       11    3.8937 4460.6 -13809
## - acousticness:key                   11    5.1785 4461.9 -13805
## - valence:key                        11    5.2527 4462.0 -13805
```

```
## - sqrt(instrumentalness):key            11     6.1855 4462.9 -13802
## - speechiness:key                        11     9.1536 4465.9 -13793
## - spotifytraining$duration:explicit       1     0.3258 4457.1 -13725
## - liveness:morethan1artist                1     0.4292 4457.2 -13724
## - acousticness:mode                       1     0.7956 4457.5 -13723
## - tempo:morethan1artist                   1     0.9100 4457.6 -13723
## - valence:mode                            1     1.2421 4458.0 -13722
## - spotifytraining$duration:mode           1     1.3293 4458.1 -13722
## - tempo:explicit                          1     2.0498 4458.8 -13720
## - speechiness:morethan1artist             1     2.2414 4459.0 -13719
## <none>                                               4456.7 -13716
## - valence:morethan1artist                 1     5.5502 4462.3 -13709
## - speechiness:explicit                    1     5.9554 4462.7 -13708
## - loudness:explicit                       1     8.3950 4465.1 -13700
## + tempo:key                              11     2.4055 4454.3 -13619
##
## Step:  AIC=-13811.2
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##     acousticness:key + acousticness:mode + danceability:key +
##     sqrt(instrumentalness):key + valence:key + valence:morethan1artist +
##     valence:mode + liveness:morethan1artist + speechiness:key +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                          Df Sum of Sq    RSS    AIC
## - danceability:key                       11     3.3818 4463.3 -13906
## - loudness:key                           11     3.9499 4463.9 -13904
## - acousticness:key                       11     5.0660 4465.0 -13900
## - valence:key                            11     5.3805 4465.3 -13900
## - sqrt(instrumentalness):key             11     6.1214 4466.0 -13897
## - speechiness:key                        11     9.0811 4469.0 -13888
## - liveness:morethan1artist                1     0.3998 4460.3 -13820
## - spotifytraining$duration:explicit       1     0.5200 4460.4 -13819
## - acousticness:mode                       1     0.8420 4460.7 -13818
## - tempo:morethan1artist                   1     0.9229 4460.8 -13818
## - spotifytraining$duration:mode           1     1.2937 4461.2 -13817
## - valence:mode                            1     1.3102 4461.2 -13817
## - tempo:explicit                          1     2.0620 4462.0 -13814
## - speechiness:morethan1artist             1     2.2447 4462.2 -13814
## <none>                                               4459.9 -13811
## - valence:morethan1artist                 1     5.5611 4465.5 -13804
## - speechiness:explicit                    1     6.0068 4465.9 -13803
## - loudness:explicit                       1     8.5059 4468.4 -13795
## + spotifytraining$duration:key           11     3.1683 4456.7 -13716
## + tempo:key                              11     2.4198 4457.5 -13714
##
```

```
## Step:  AIC=-13905.58
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:key + loudness:expli
cit +
##     acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - loudness:key                       11     3.768 4467.1 -13999
## - acousticness:key                   11     5.142 4468.4 -13995
## - valence:key                        11     5.474 4468.8 -13994
## - sqrt(instrumentalness):key         11     6.048 4469.3 -13992
## - speechiness:key                    11     9.140 4472.4 -13983
## - liveness:morethan1artist            1     0.382 4463.7 -13914
## - spotifytraining$duration:explicit   1     0.489 4463.8 -13914
## - acousticness:mode                   1     0.833 4464.1 -13913
## - tempo:morethan1artist               1     0.863 4464.2 -13912
## - spotifytraining$duration:mode       1     1.233 4464.5 -13911
## - valence:mode                        1     1.346 4464.6 -13911
## - tempo:explicit                      1     2.169 4465.5 -13909
## - speechiness:morethan1artist         1     2.187 4465.5 -13908
## <none>                                            4463.3 -13906
## - valence:morethan1artist             1     5.670 4469.0 -13898
## - speechiness:explicit                1     5.838 4469.1 -13898
## - loudness:explicit                   1     8.361 4471.6 -13890
## + danceability:key                   11     3.382 4459.9 -13811
## + spotifytraining$duration:key       11     2.965 4460.3 -13810
## + tempo:key                          11     2.369 4460.9 -13808
## - danceability                        1   132.741 4596.0 -13522
##
## Step:  AIC=-13998.8
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:key + acousticness:mode + sqrt(instrumentalness):key +
##     valence:key + valence:morethan1artist + valence:mode + liveness:moreth
an1artist +
##     speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                Df Sum of Sq    RSS    AIC
## - acousticness:key             11     4.288 4471.3 -14090
## - sqrt(instrumentalness):key   11     6.276 4473.3 -14084
## - valence:key                  11     7.353 4474.4 -14081
```

```
## - speechiness:key                         11          9.793 4476.8 -14074
## - liveness:morethan1artist                 1          0.390 4467.4 -14007
## - spotifytraining$duration:explicit        1          0.507 4467.6 -14007
## - acousticness:mode                        1          0.820 4467.9 -14006
## - tempo:morethan1artist                    1          0.896 4468.0 -14006
## - spotifytraining$duration:mode            1          1.286 4468.3 -14004
## - valence:mode                             1          1.425 4468.5 -14004
## - tempo:explicit                           1          2.197 4469.3 -14002
## - speechiness:morethan1artist              1          2.324 4469.4 -14001
## <none>                                                       4467.1 -13999
## - valence:morethan1artist                  1          5.807 4472.9 -13991
## - speechiness:explicit                     1          6.022 4473.1 -13990
## - loudness:explicit                        1          8.640 4475.7 -13982
## + loudness:key                            11          3.768 4463.3 -13906
## + danceability:key                        11          3.200 4463.9 -13904
## + spotifytraining$duration:key            11          3.133 4463.9 -13904
## + tempo:key                               11          2.354 4464.7 -13901
## - danceability                             1        132.558 4599.6 -13616
##
## Step:   AIC=-14090.48
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##       acousticness + danceability + sqrt(instrumentalness) + valence +
##       liveness + speechiness + key + morethan1artist + mode + explicit +
##       spotifytraining$duration:mode + spotifytraining$duration:explicit +
##       tempo:morethan1artist + tempo:explicit + loudness:explicit +
##       acousticness:mode + sqrt(instrumentalness):key + valence:key +
##       valence:morethan1artist + valence:mode + liveness:morethan1artist +
##       speechiness:key + speechiness:morethan1artist + speechiness:explicit
##
##                                            Df Sum of Sq    RSS    AIC
## - sqrt(instrumentalness):key               11          6.202 4477.5 -14176
## - valence:key                             11          6.631 4478.0 -14175
## - speechiness:key                         11          9.915 4481.3 -14165
## - liveness:morethan1artist                 1          0.395 4471.7 -14099
## - spotifytraining$duration:explicit        1          0.514 4471.9 -14098
## - tempo:morethan1artist                    1          0.851 4472.2 -14097
## - acousticness:mode                        1          1.068 4472.4 -14097
## - spotifytraining$duration:mode            1          1.356 4472.7 -14096
## - valence:mode                             1          1.417 4472.8 -14096
## - tempo:explicit                           1          2.180 4473.5 -14093
## - speechiness:morethan1artist              1          2.410 4473.8 -14093
## <none>                                                       4471.3 -14090
## - valence:morethan1artist                  1          5.999 4477.3 -14082
## - speechiness:explicit                     1          6.028 4477.4 -14082
## - loudness:explicit                        1          8.782 4480.1 -14074
## + acousticness:key                        11          4.288 4467.1 -13999
## + danceability:key                        11          3.274 4468.1 -13996
## + spotifytraining$duration:key            11          3.053 4468.3 -13995
## + loudness:key                            11          2.913 4468.4 -13995
## + tempo:key                               11          2.784 4468.6 -13994
```

```
## - danceability                                 1    132.170 4603.5 -13709
##
## Step:   AIC=-14176.43
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##       acousticness + danceability + sqrt(instrumentalness) + valence +
##       liveness + speechiness + key + morethan1artist + mode + explicit +
##       spotifytraining$duration:mode + spotifytraining$duration:explicit +
##       tempo:morethan1artist + tempo:explicit + loudness:explicit +
##       acousticness:mode + valence:key + valence:morethan1artist +
##       valence:mode + liveness:morethan1artist + speechiness:key +
##       speechiness:morethan1artist + speechiness:explicit
##
##                                          Df Sum of Sq    RSS    AIC
## - valence:key                            11     6.973 4484.5 -14260
## - speechiness:key                        11    10.394 4487.9 -14250
## - liveness:morethan1artist                1     0.360 4477.9 -14185
## - spotifytraining$duration:explicit       1     0.520 4478.1 -14184
## - tempo:morethan1artist                   1     0.919 4478.5 -14183
## - acousticness:mode                       1     0.970 4478.5 -14183
## - spotifytraining$duration:mode           1     1.394 4478.9 -14182
## - valence:mode                            1     1.463 4479.0 -14182
## - tempo:explicit                          1     2.195 4479.7 -14179
## - speechiness:morethan1artist             1     2.455 4480.0 -14179
## <none>                                                 4477.5 -14176
## - speechiness:explicit                    1     6.007 4483.6 -14168
## - valence:morethan1artist                 1     6.074 4483.6 -14168
## - sqrt(instrumentalness)                  1     6.098 4483.6 -14168
## - loudness:explicit                       1     8.878 4486.4 -14159
## + sqrt(instrumentalness):key             11     6.202 4471.3 -14090
## + acousticness:key                       11     4.214 4473.3 -14084
## + danceability:key                       11     3.225 4474.3 -14082
## + loudness:key                           11     2.998 4474.5 -14081
## + spotifytraining$duration:key           11     2.949 4474.6 -14081
## + tempo:key                              11     2.822 4474.7 -14080
## - danceability                            1   132.696 4610.2 -13794
##
## Step:   AIC=-14260.1
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##       acousticness + danceability + sqrt(instrumentalness) + valence +
##       liveness + speechiness + key + morethan1artist + mode + explicit +
##       spotifytraining$duration:mode + spotifytraining$duration:explicit +
##       tempo:morethan1artist + tempo:explicit + loudness:explicit +
##       acousticness:mode + valence:morethan1artist + valence:mode +
##       liveness:morethan1artist + speechiness:key + speechiness:morethan1arti
## st +
##       speechiness:explicit
##
##                                          Df Sum of Sq    RSS    AIC
## - speechiness:key                        11    10.288 4494.8 -14334
## - liveness:morethan1artist                1     0.406 4484.9 -14268
```

```
## - spotifytraining$duration:explicit  1      0.561 4485.1 -14268
## - tempo:morethan1artist              1      0.901 4485.4 -14267
## - acousticness:mode                  1      0.941 4485.5 -14267
## - spotifytraining$duration:mode      1      1.389 4485.9 -14265
## - valence:mode                       1      1.974 4486.5 -14264
## - tempo:explicit                     1      2.133 4486.7 -14263
## - speechiness:morethan1artist        1      2.385 4486.9 -14262
## <none>                                            4484.5 -14260
## - speechiness:explicit               1      5.939 4490.5 -14252
## - sqrt(instrumentalness)             1      6.107 4490.6 -14251
## - valence:morethan1artist            1      6.602 4491.1 -14250
## - loudness:explicit                  1      8.573 4493.1 -14244
## + valence:key                       11      6.973 4477.5 -14176
## + sqrt(instrumentalness):key        11      6.544 4478.0 -14175
## + loudness:key                      11      3.829 4480.7 -14167
## + danceability:key                  11      3.532 4481.0 -14166
## + acousticness:key                  11      3.484 4481.0 -14166
## + spotifytraining$duration:key      11      2.830 4481.7 -14164
## + tempo:key                         11      2.707 4481.8 -14164
## - danceability                       1    131.653 4616.2 -13882
##
## Step:  AIC=-14333.9
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + key + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     liveness:morethan1artist + speechiness:morethan1artist +
##     speechiness:explicit
##
##                                    Df Sum of Sq    RSS    AIC
## - key                              11     9.069 4503.9 -14411
## - liveness:morethan1artist          1     0.333 4495.1 -14342
## - spotifytraining$duration:explicit  1     0.450 4495.3 -14342
## - tempo:morethan1artist             1     0.826 4495.6 -14341
## - acousticness:mode                 1     1.009 4495.8 -14340
## - spotifytraining$duration:mode     1     1.271 4496.1 -14340
## - speechiness:morethan1artist       1     1.978 4496.8 -14338
## - valence:mode                      1     1.979 4496.8 -14338
## - tempo:explicit                    1     2.269 4497.1 -14337
## <none>                                           4494.8 -14334
## - speechiness:explicit              1     6.072 4500.9 -14325
## - sqrt(instrumentalness)            1     6.207 4501.0 -14325
## - valence:morethan1artist           1     6.383 4501.2 -14324
## - loudness:explicit                 1     8.241 4503.0 -14319
## + speechiness:key                  11    10.288 4484.5 -14260
## + sqrt(instrumentalness):key       11     6.998 4487.8 -14250
## + valence:key                      11     6.867 4487.9 -14250
## + loudness:key                     11     3.978 4490.8 -14241
```

```
## + acousticness:key                          11     3.664 4491.1 -14240
## + danceability:key                           11     3.176 4491.6 -14239
## + tempo:key                                  11     2.945 4491.9 -14238
## + spotifytraining$duration:key               11     2.529 4492.3 -14237
## - danceability                                1   133.286 4628.1 -13951
##
## Step:  AIC=-14411.41
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     liveness:morethan1artist + speechiness:morethan1artist +
##     speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - liveness:morethan1artist              1     0.334 4504.2 -14420
## - spotifytraining$duration:explicit     1     0.400 4504.3 -14420
## - acousticness:mode                     1     0.816 4504.7 -14418
## - tempo:morethan1artist                 1     0.826 4504.7 -14418
## - spotifytraining$duration:mode         1     1.458 4505.3 -14417
## - speechiness:morethan1artist           1     2.057 4505.9 -14415
## - valence:mode                          1     2.098 4506.0 -14415
## - tempo:explicit                        1     2.299 4506.2 -14414
## <none>                                              4503.9 -14411
## - speechiness:explicit                  1     6.194 4510.1 -14402
## - sqrt(instrumentalness)                1     6.249 4510.1 -14402
## - valence:morethan1artist               1     6.474 4510.4 -14402
## - loudness:explicit                     1     8.509 4512.4 -14396
## + key                                  11     9.069 4494.8 -14334
## - danceability                          1   135.231 4639.1 -14024
##
## Step:  AIC=-14419.92
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + spotifytraining$duration:explicit +
##     tempo:morethan1artist + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                        Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:explicit     1     0.413 4504.6 -14428
## - tempo:morethan1artist                 1     0.826 4505.0 -14427
## - acousticness:mode                     1     0.832 4505.0 -14427
## - spotifytraining$duration:mode         1     1.458 4505.7 -14425
## - valence:mode                          1     2.095 4506.3 -14423
## - speechiness:morethan1artist           1     2.258 4506.5 -14423
## - tempo:explicit                        1     2.299 4506.5 -14423
```

```
## <none>                                                   4504.2 -14420
## + liveness:morethan1artist             1       0.334 4503.9 -14411
## - speechiness:explicit                 1       6.245 4510.5 -14411
## - sqrt(instrumentalness)               1       6.261 4510.5 -14411
## - valence:morethan1artist              1       6.421 4510.6 -14410
## - loudness:explicit                    1       8.609 4512.8 -14404
## - liveness                             1      22.012 4526.2 -14364
## + key                                 11       9.070 4495.1 -14342
## - danceability                         1     135.309 4639.5 -14032
##
## Step:  AIC=-14428.2
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:morethan1artist + tempo:explicit
+
##     loudness:explicit + acousticness:mode + valence:morethan1artist +
##     valence:mode + speechiness:morethan1artist + speechiness:explicit
##
##                                       Df Sum of Sq    RSS    AIC
## - tempo:morethan1artist                1       0.820 4505.4 -14435
## - acousticness:mode                    1       0.823 4505.4 -14435
## - spotifytraining$duration:mode        1       1.907 4506.5 -14432
## - valence:mode                         1       2.120 4506.7 -14431
## - speechiness:morethan1artist          1       2.307 4506.9 -14431
## - tempo:explicit                       1       2.329 4507.0 -14431
## <none>                                                 4504.6 -14428
## + spotifytraining$duration:explicit    1       0.413 4504.2 -14420
## + liveness:morethan1artist             1       0.347 4504.3 -14420
## - sqrt(instrumentalness)               1       6.104 4510.7 -14420
## - speechiness:explicit                 1       6.109 4510.7 -14420
## - valence:morethan1artist              1       6.668 4511.3 -14418
## - loudness:explicit                    1       8.671 4513.3 -14412
## - liveness                             1      22.126 4526.7 -14372
## + key                                 11       9.019 4495.6 -14350
## - danceability                         1     136.211 4640.8 -14038
##
## Step:  AIC=-14435.26
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     acousticness:mode + valence:morethan1artist + valence:mode +
##     speechiness:morethan1artist + speechiness:explicit
##
##                                       Df Sum of Sq    RSS    AIC
## - acousticness:mode                    1       0.826 4506.3 -14442
## - spotifytraining$duration:mode        1       1.900 4507.3 -14439
## - valence:mode                         1       2.110 4507.6 -14438
## - speechiness:morethan1artist          1       2.163 4507.6 -14438
```

```
## - tempo:explicit                         1      3.121 4508.6 -14436
## <none>                                                  4505.4 -14435
## + tempo:morethan1artist                   1      0.820 4504.6 -14428
## + spotifytraining$duration:explicit  1      0.407 4505.0 -14427
## + liveness:morethan1artist                1      0.346 4505.1 -14427
## - sqrt(instrumentalness)                   1      6.056 4511.5 -14427
## - speechiness:explicit                     1      6.215 4511.7 -14426
## - valence:morethan1artist                  1      6.720 4512.2 -14425
## - loudness:explicit                        1      8.657 4514.1 -14419
## - liveness                                 1     22.166 4527.6 -14379
## + key                                     11      9.019 4496.4 -14358
## - danceability                             1    136.935 4642.4 -14043
##
## Step:  AIC=-14442.31
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     valence:morethan1artist + valence:mode + speechiness:morethan1artist +
##     speechiness:explicit
##
##                                       Df Sum of Sq    RSS    AIC
## - valence:mode                             1      1.871 4508.1 -14446
## - spotifytraining$duration:mode       1      2.059 4508.3 -14446
## - speechiness:morethan1artist              1      2.190 4508.5 -14445
## - tempo:explicit                           1      3.101 4509.4 -14443
## <none>                                                  4506.3 -14442
## + acousticness:mode                        1      0.826 4505.4 -14435
## + tempo:morethan1artist                    1      0.823 4505.4 -14435
## - sqrt(instrumentalness)                   1      5.872 4512.1 -14434
## + spotifytraining$duration:explicit  1      0.397 4505.9 -14434
## + liveness:morethan1artist                 1      0.362 4505.9 -14434
## - speechiness:explicit                     1      6.282 4512.6 -14433
## - valence:morethan1artist                  1      6.728 4513.0 -14432
## - loudness:explicit                        1      9.017 4515.3 -14425
## - liveness                                 1     22.364 4528.6 -14385
## + key                                     11      8.826 4497.4 -14364
## - acousticness                             1     29.587 4535.9 -14364
## - danceability                             1    137.054 4643.3 -14050
##
## Step:  AIC=-14446.24
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     valence:morethan1artist + speechiness:morethan1artist + speechiness:ex
plicit
##
##                                       Df Sum of Sq    RSS    AIC
## - speechiness:morethan1artist              1      2.210 4510.3 -14449
```

```
## - spotifytraining$duration:mode       1      2.939 4511.1 -14447
## - tempo:explicit                       1      3.169 4511.3 -14446
## <none>                                              4508.1 -14446
## + valence:mode                         1      1.871 4506.3 -14442
## + tempo:morethan1artist                1      0.813 4507.3 -14439
## + acousticness:mode                    1      0.587 4507.6 -14438
## - sqrt(instrumentalness)               1      5.848 4514.0 -14438
## + spotifytraining$duration:explicit  1      0.422 4507.7 -14438
## + liveness:morethan1artist             1      0.356 4507.8 -14438
## - speechiness:explicit                 1      6.162 4514.3 -14437
## - valence:morethan1artist              1      7.365 4515.5 -14434
## - loudness:explicit                    1      8.684 4516.8 -14430
## - liveness                             1     22.456 4530.6 -14389
## + key                                 11      8.969 4499.2 -14368
## - acousticness                         1     29.603 4537.7 -14368
## - danceability                         1    137.395 4645.5 -14053
##
## Step:  AIC=-14449.17
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     spotifytraining$duration:mode + tempo:explicit + loudness:explicit +
##     valence:morethan1artist + speechiness:explicit
##
##                                      Df Sum of Sq    RSS    AIC
## - spotifytraining$duration:mode       1      3.006 4513.4 -14450
## <none>                                              4510.3 -14449
## - tempo:explicit                       1      3.229 4513.6 -14449
## + speechiness:morethan1artist          1      2.210 4508.1 -14446
## + valence:mode                         1      1.891 4508.5 -14445
## + tempo:morethan1artist                1      0.668 4509.7 -14442
## + acousticness:mode                    1      0.608 4509.7 -14442
## + liveness:morethan1artist             1      0.562 4509.8 -14441
## + spotifytraining$duration:explicit  1      0.471 4509.9 -14441
## - sqrt(instrumentalness)               1      5.937 4516.3 -14441
## - speechiness:explicit                 1      6.616 4517.0 -14439
## - valence:morethan1artist              1      7.035 4517.4 -14438
## - loudness:explicit                    1      8.991 4519.3 -14432
## - liveness                             1     21.947 4532.3 -14394
## + key                                 11      9.047 4501.3 -14372
## - acousticness                         1     30.158 4540.5 -14369
## - danceability                         1    138.727 4649.1 -14052
##
## Step:  AIC=-14449.74
## sqrt(popularity) ~ spotifytraining$duration + tempo + loudness +
##     acousticness + danceability + sqrt(instrumentalness) + valence +
##     liveness + speechiness + morethan1artist + mode + explicit +
##     tempo:explicit + loudness:explicit + valence:morethan1artist +
##     speechiness:explicit
##
```

```
##                                      Df Sum of Sq    RSS    AIC
## <none>                                            4513.4 -14450
## - tempo:explicit                      1     3.241 4516.6 -14450
## + spotifytraining$duration:mode       1     3.006 4510.3 -14449
## + valence:mode                        1     2.787 4510.6 -14448
## + speechiness:morethan1artist         1     2.276 4511.1 -14447
## + spotifytraining$duration:explicit   1     1.119 4512.2 -14444
## + acousticness:mode                   1     0.725 4512.6 -14442
## + tempo:morethan1artist               1     0.656 4512.7 -14442
## + liveness:morethan1artist            1     0.576 4512.8 -14442
## - sqrt(instrumentalness)              1     6.116 4519.5 -14441
## - speechiness:explicit                1     6.570 4519.9 -14440
## - valence:morethan1artist             1     7.167 4520.5 -14438
## - loudness:explicit                   1     8.896 4522.3 -14433
## - mode                                1     9.376 4522.7 -14431
## - liveness                            1    21.527 4534.9 -14395
## + key                                11     9.321 4504.0 -14373
## - acousticness                        1    30.003 4543.4 -14370
## - spotifytraining$duration            1    64.607 4578.0 -14269
## - danceability                        1   139.006 4652.4 -14052

anova(stepwiseBIC)

## Analysis of Variance Table
##
## Response: sqrt(popularity)
##                             Df Sum Sq Mean Sq  F value     Pr(>F)
## spotifytraining$duration     1  124.3 124.293 368.9380  < 2.2e-16 ***
## tempo                        1    2.5   2.501   7.4247   0.006442 **
## loudness                     1    0.2   0.170   0.5032   0.478093
## acousticness                 1    6.5   6.471  19.2083  1.181e-05 ***
## danceability                 1  208.4 208.377 618.5250  < 2.2e-16 ***
## sqrt(instrumentalness)       1   11.1  11.139  33.0639  9.112e-09 ***
## valence                      1   92.8  92.770 275.3688  < 2.2e-16 ***
## liveness                     1   17.7  17.739  52.6538  4.197e-13 ***
## speechiness                  1    0.2   0.235   0.6989   0.403173
## morethan1artist              1   86.1  86.092 255.5466  < 2.2e-16 ***
## mode                         1   13.1  13.120  38.9451  4.490e-10 ***
## explicit                     1   16.2  16.219  48.1422  4.146e-12 ***
## tempo:explicit               1    1.9   1.863   5.5295   0.018713 *
## loudness:explicit            1    7.4   7.378  21.8993  2.902e-06 ***
## valence:morethan1artist      1    7.2   7.185  21.3287  3.905e-06 ***
## speechiness:explicit         1    6.6   6.570  19.5030  1.012e-05 ***
## Residuals                13397 4513.4   0.337
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(stepwiseBIC)$r.sq

## [1] 0.1177059
```

```
summary(stepwiseBIC)$adj.r.sq
```

```
## [1] 0.1166522
```

```
## [1]      17.00 -14577.31
```

Testing model against testing data

```
modeltest <- lm(sqrt(spotify_test$popularity) ~ spotify_test$durationmin +
    spotify_test$tempo + spotify_test$loudness + spotify_test$acousticness +
    spotify_test$danceability + sqrt(spotify_test$instrumentalness) +
    spotify_test$valence + spotify_test$liveness + spotify_test$speechiness +
    spotify_test$key + spotify_test$morethan1artist + spotify_test$mode +
    spotify_test$explicit + spotify_test$durationmin:spotify_test$mode +
    spotify_test$tempo:spotify_test$explicit + spotify_test$loudness:spotify_
test$explicit:spotify_test$acousticness:spotify_test$mode +
    spotify_test$danceability:spotify_test$key + spotify_test$valence:spotify
_test$morethan1artist +
    spotify_test$valence:spotify_test$mode + spotify_test$speechiness:spotify
_test$key +
    spotify_test$speechiness:spotify_test$morethan1artist + spotify_test$spee
chiness:spotify_test$explicit,
    data = spotify_test)
```

```
summary(modeltest)
```

```
##
## Call:
## lm(formula = sqrt(spotify_test$popularity) ~ spotify_test$durationmin +
##      spotify_test$tempo + spotify_test$loudness + spotify_test$acousticness
+
##      spotify_test$danceability + sqrt(spotify_test$instrumentalness) +
##      spotify_test$valence + spotify_test$liveness + spotify_test$speechines
s +
##      spotify_test$key + spotify_test$morethan1artist + spotify_test$mode +
##      spotify_test$explicit + spotify_test$durationmin:spotify_test$mode +
##      spotify_test$tempo:spotify_test$explicit + spotify_test$loudness:spoti
fy_test$explicit:spotify_test$acousticness:spotify_test$mode +
##      spotify_test$danceability:spotify_test$key + spotify_test$valence:spot
ify_test$morethan1artist +
##      spotify_test$valence:spotify_test$mode + spotify_test$speechiness:spot
ify_test$key +
##      spotify_test$speechiness:spotify_test$morethan1artist + spotify_test$s
peechiness:spotify_test$explicit,
##      data = spotify_test)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.4258 -0.4166 -0.0037   0.4097   1.8980
##
## Coefficients:
```

```
##
Estimate
## (Intercept)
7.7813693
## spotify_test$durationmin
-0.1212871
## spotify_test$tempo
-0.0002203
## spotify_test$loudness
0.0121858
## spotify_test$acousticness
0.2825983
## spotify_test$danceability
0.9130866
## sqrt(spotify_test$instrumentalness)
-0.2737133
## spotify_test$valence
-0.2637835
## spotify_test$liveness
-0.2648211
## spotify_test$speechiness
0.1567550
## spotify_test$key1
0.1202388
## spotify_test$key2
0.1500884
## spotify_test$key3
0.3384312
## spotify_test$key4
-0.0351353
## spotify_test$key5
0.1867960
## spotify_test$key6
0.0056441
## spotify_test$key7
0.0616763
## spotify_test$key8
0.2709322
## spotify_test$key9
0.1891855
## spotify_test$key10
0.2398202
## spotify_test$key11
0.0423149
## spotify_test$morethan1artist1
0.0092579
## spotify_test$mode1
-0.0745777
## spotify_test$explicit1
-0.1212540
```

```
## spotify_test$durationmin:spotify_test$mode1
0.0368522
## spotify_test$tempo:spotify_test$explicit1
0.0018373
## spotify_test$danceability:spotify_test$key1
-0.0922836
## spotify_test$danceability:spotify_test$key2
-0.2750620
## spotify_test$danceability:spotify_test$key3
-0.4647007
## spotify_test$danceability:spotify_test$key4
0.0655542
## spotify_test$danceability:spotify_test$key5
-0.3328225
## spotify_test$danceability:spotify_test$key6
0.1464514
## spotify_test$danceability:spotify_test$key7
-0.1514633
## spotify_test$danceability:spotify_test$key8
-0.3074966
## spotify_test$danceability:spotify_test$key9
-0.2819471
## spotify_test$danceability:spotify_test$key10
-0.3226530
## spotify_test$danceability:spotify_test$key11
-0.0617045
## spotify_test$valence:spotify_test$morethan1artist1
0.2847656
## spotify_test$valence:spotify_test$mode1
-0.2369940
## spotify_test$speechiness:spotify_test$key1
-0.5169197
## spotify_test$speechiness:spotify_test$key2
0.0201300
## spotify_test$speechiness:spotify_test$key3
-0.1635318
## spotify_test$speechiness:spotify_test$key4
-0.4120405
## spotify_test$speechiness:spotify_test$key5
-0.0376515
## spotify_test$speechiness:spotify_test$key6
-0.4738737
## spotify_test$speechiness:spotify_test$key7
0.2513837
## spotify_test$speechiness:spotify_test$key8
-0.5362090
## spotify_test$speechiness:spotify_test$key9
-0.3997410
## spotify_test$speechiness:spotify_test$key10
0.0034172
```

```
## spotify_test$speechiness:spotify_test$key11
-0.2102842
## spotify_test$speechiness:spotify_test$morethan1artist1
-0.1239017
## spotify_test$speechiness:spotify_test$explicit1
-0.1006687
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit0  0.0062842
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit0  0.0147982
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit1  0.0064417
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit1 -0.0062374
##
Std. Error
## (Intercept)
0.1189665
## spotify_test$durationmin
0.0133500
## spotify_test$tempo
0.0002944
## spotify_test$loudness
0.0043365
## spotify_test$acousticness
0.0691730
## spotify_test$danceability
0.1451894
## sqrt(spotify_test$instrumentalness)
0.0703539
## spotify_test$valence
0.0602401
## spotify_test$liveness
0.0468826
## spotify_test$speechiness
0.2377916
## spotify_test$key1
0.1212097
## spotify_test$key2
0.1212547
## spotify_test$key3
0.1869578
## spotify_test$key4
0.1366155
## spotify_test$key5
0.1324131
## spotify_test$key6
0.1367182
## spotify_test$key7
0.1268232
```

```
## spotify_test$key8
0.1351785
## spotify_test$key9
0.1321885
## spotify_test$key10
0.1364681
## spotify_test$key11
0.1353616
## spotify_test$morethan1artist1
0.0407168
## spotify_test$mode1
0.0767094
## spotify_test$explicit1
0.0671699
## spotify_test$durationmin:spotify_test$mode1
0.0161870
## spotify_test$tempo:spotify_test$explicit1
0.0005087
## spotify_test$danceability:spotify_test$key1
0.1906401
## spotify_test$danceability:spotify_test$key2
0.1988774
## spotify_test$danceability:spotify_test$key3
0.3117948
## spotify_test$danceability:spotify_test$key4
0.2235378
## spotify_test$danceability:spotify_test$key5
0.2125736
## spotify_test$danceability:spotify_test$key6
0.2225692
## spotify_test$danceability:spotify_test$key7
0.2041906
## spotify_test$danceability:spotify_test$key8
0.2160013
## spotify_test$danceability:spotify_test$key9
0.2136963
## spotify_test$danceability:spotify_test$key10
0.2200342
## spotify_test$danceability:spotify_test$key11
0.2149935
## spotify_test$valence:spotify_test$morethan1artist1
0.0704908
## spotify_test$valence:spotify_test$mode1
0.0668087
## spotify_test$speechiness:spotify_test$key1
0.2656813
## spotify_test$speechiness:spotify_test$key2
0.3043858
## spotify_test$speechiness:spotify_test$key3
0.4089697
```

```
## spotify_test$speechiness:spotify_test$key4
0.3252498
## spotify_test$speechiness:spotify_test$key5
0.3192380
## spotify_test$speechiness:spotify_test$key6
0.3191833
## spotify_test$speechiness:spotify_test$key7
0.3174110
## spotify_test$speechiness:spotify_test$key8
0.3063258
## spotify_test$speechiness:spotify_test$key9
0.3000781
## spotify_test$speechiness:spotify_test$key10
0.3253213
## spotify_test$speechiness:spotify_test$key11
0.2995398
## spotify_test$speechiness:spotify_test$morethan1artist1
0.1363705
## spotify_test$speechiness:spotify_test$explicit1
0.1471995
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit0  0.0092902
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit0  0.0080133
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit1  0.0121320
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit1  0.0108032
##
t value
## (Intercept)
65.408
## spotify_test$durationmin
-9.085
## spotify_test$tempo
-0.748
## spotify_test$loudness
2.810
## spotify_test$acousticness
4.085
## spotify_test$danceability
6.289
## sqrt(spotify_test$instrumentalness)
-3.891
## spotify_test$valence
-4.379
## spotify_test$liveness
-5.649
## spotify_test$speechiness
0.659
```

```
## spotify_test$key1
0.992
## spotify_test$key2
1.238
## spotify_test$key3
1.810
## spotify_test$key4
-0.257
## spotify_test$key5
1.411
## spotify_test$key6
0.041
## spotify_test$key7
0.486
## spotify_test$key8
2.004
## spotify_test$key9
1.431
## spotify_test$key10
1.757
## spotify_test$key11
0.313
## spotify_test$morethan1artist1
0.227
## spotify_test$mode1
-0.972
## spotify_test$explicit1
-1.805
## spotify_test$durationmin:spotify_test$mode1
2.277
## spotify_test$tempo:spotify_test$explicit1
3.611
## spotify_test$danceability:spotify_test$key1
-0.484
## spotify_test$danceability:spotify_test$key2
-1.383
## spotify_test$danceability:spotify_test$key3
-1.490
## spotify_test$danceability:spotify_test$key4
0.293
## spotify_test$danceability:spotify_test$key5
-1.566
## spotify_test$danceability:spotify_test$key6
0.658
## spotify_test$danceability:spotify_test$key7
-0.742
## spotify_test$danceability:spotify_test$key8
-1.424
## spotify_test$danceability:spotify_test$key9
-1.319
```

```
## spotify_test$danceability:spotify_test$key10
-1.466
## spotify_test$danceability:spotify_test$key11
-0.287
## spotify_test$valence:spotify_test$morethan1artist1
4.040
## spotify_test$valence:spotify_test$mode1
-3.547
## spotify_test$speechiness:spotify_test$key1
-1.946
## spotify_test$speechiness:spotify_test$key2
0.066
## spotify_test$speechiness:spotify_test$key3
-0.400
## spotify_test$speechiness:spotify_test$key4
-1.267
## spotify_test$speechiness:spotify_test$key5
-0.118
## spotify_test$speechiness:spotify_test$key6
-1.485
## spotify_test$speechiness:spotify_test$key7
0.792
## spotify_test$speechiness:spotify_test$key8
-1.750
## spotify_test$speechiness:spotify_test$key9
-1.332
## spotify_test$speechiness:spotify_test$key10
0.011
## spotify_test$speechiness:spotify_test$key11
-0.702
## spotify_test$speechiness:spotify_test$morethan1artist1
-0.909
## spotify_test$speechiness:spotify_test$explicit1
-0.684
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit0    0.676
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit0    1.847
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit1    0.531
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit1   -0.577
##
Pr(>|t|)
## (Intercept)
< 2e-16
## spotify_test$durationmin
< 2e-16
## spotify_test$tempo
0.454297
```

```
## spotify_test$loudness
0.004967
## spotify_test$acousticness
4.45e-05
## spotify_test$danceability
3.38e-10
## sqrt(spotify_test$instrumentalness)
0.000101
## spotify_test$valence
1.21e-05
## spotify_test$liveness
1.68e-08
## spotify_test$speechiness
0.509781
## spotify_test$key1
0.321236
## spotify_test$key2
0.215833
## spotify_test$key3
0.070306
## spotify_test$key4
0.797044
## spotify_test$key5
0.158375
## spotify_test$key6
0.967071
## spotify_test$key7
0.626757
## spotify_test$key8
0.045080
## spotify_test$key9
0.152422
## spotify_test$key10
0.078903
## spotify_test$key11
0.754589
## spotify_test$morethan1artist1
0.820140
## spotify_test$mode1
0.330978
## spotify_test$explicit1
0.071088
## spotify_test$durationmin:spotify_test$mode1
0.022837
## spotify_test$tempo:spotify_test$explicit1
0.000307
## spotify_test$danceability:spotify_test$key1
0.628349
## spotify_test$danceability:spotify_test$key2
0.166686
```

```
## spotify_test$danceability:spotify_test$key3
0.136162
## spotify_test$danceability:spotify_test$key4
0.769334
## spotify_test$danceability:spotify_test$key5
0.117467
## spotify_test$danceability:spotify_test$key6
0.510557
## spotify_test$danceability:spotify_test$key7
0.458249
## spotify_test$danceability:spotify_test$key8
0.154610
## spotify_test$danceability:spotify_test$key9
0.187084
## spotify_test$danceability:spotify_test$key10
0.142590
## spotify_test$danceability:spotify_test$key11
0.774116
## spotify_test$valence:spotify_test$morethan1artist1
5.41e-05
## spotify_test$valence:spotify_test$mode1
0.000392
## spotify_test$speechiness:spotify_test$key1
0.051737
## spotify_test$speechiness:spotify_test$key2
0.947274
## spotify_test$speechiness:spotify_test$key3
0.689269
## spotify_test$speechiness:spotify_test$key4
0.205253
## spotify_test$speechiness:spotify_test$key5
0.906117
## spotify_test$speechiness:spotify_test$key6
0.137682
## spotify_test$speechiness:spotify_test$key7
0.428398
## spotify_test$speechiness:spotify_test$key8
0.080083
## spotify_test$speechiness:spotify_test$key9
0.182862
## spotify_test$speechiness:spotify_test$key10
0.991619
## spotify_test$speechiness:spotify_test$key11
0.482687
## spotify_test$speechiness:spotify_test$morethan1artist1
0.363609
## spotify_test$speechiness:spotify_test$explicit1
0.494065
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit0 0.498793
```

```
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit0 0.064831
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit1 0.595457
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit1 0.563709
##
## (Intercept)
***
## spotify_test$durationmin
***
## spotify_test$tempo
## spotify_test$loudness
**
## spotify_test$acousticness
***
## spotify_test$danceability
***
## sqrt(spotify_test$instrumentalness)
***
## spotify_test$valence
***
## spotify_test$liveness
***
## spotify_test$speechiness
## spotify_test$key1
## spotify_test$key2
## spotify_test$key3
.
## spotify_test$key4
## spotify_test$key5
## spotify_test$key6
## spotify_test$key7
## spotify_test$key8
*
## spotify_test$key9
## spotify_test$key10
.
## spotify_test$key11
## spotify_test$morethan1artist1
## spotify_test$mode1
## spotify_test$explicit1
.
## spotify_test$durationmin:spotify_test$mode1
*
## spotify_test$tempo:spotify_test$explicit1
***
## spotify_test$danceability:spotify_test$key1
## spotify_test$danceability:spotify_test$key2
## spotify_test$danceability:spotify_test$key3
```

```
## spotify_test$danceability:spotify_test$key4
## spotify_test$danceability:spotify_test$key5
## spotify_test$danceability:spotify_test$key6
## spotify_test$danceability:spotify_test$key7
## spotify_test$danceability:spotify_test$key8
## spotify_test$danceability:spotify_test$key9
## spotify_test$danceability:spotify_test$key10
## spotify_test$danceability:spotify_test$key11
## spotify_test$valence:spotify_test$morethan1artist1
***
## spotify_test$valence:spotify_test$mode1
***
## spotify_test$speechiness:spotify_test$key1
.
## spotify_test$speechiness:spotify_test$key2
## spotify_test$speechiness:spotify_test$key3
## spotify_test$speechiness:spotify_test$key4
## spotify_test$speechiness:spotify_test$key5
## spotify_test$speechiness:spotify_test$key6
## spotify_test$speechiness:spotify_test$key7
## spotify_test$speechiness:spotify_test$key8
.
## spotify_test$speechiness:spotify_test$key9
## spotify_test$speechiness:spotify_test$key10
## spotify_test$speechiness:spotify_test$key11
## spotify_test$speechiness:spotify_test$morethan1artist1
## spotify_test$speechiness:spotify_test$explicit1
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit0
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit0 .
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode0:spotify
_test$explicit1
## spotify_test$loudness:spotify_test$acousticness:spotify_test$mode1:spotify
_test$explicit1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5856 on 7169 degrees of freedom
## Multiple R-squared:  0.1264, Adjusted R-squared:  0.1197
## F-statistic: 18.86 on 55 and 7169 DF,  p-value: < 2.2e-16

par(mfrow = c(2, 2))
plot(modeltest)
```

Residuals vs Fitted	**Normal Q-Q**
Scale-Location	**Residuals vs Leverage**