Influence of Harmonic Motion, Instrument Type, and Voice Leading on Perception of Music as Classical or Popular

Sophia L. Hecht[1]
shecht@andrew.cmu.edu

16 December 2019

Abstract

The data in this paper is from a 2012 study performed by Ivan Jimenez, Sibelius Institute, and Vincent Rossi, University of Pittsburgh, to gather information on how harmonic motion, instrument type, and voice leading affect a listener's perception of music either as classical or popular. Data cleaning and variable transformations are addressed to make the data as usable as possible. It appears that of the 26 variables that harmonic motion, instrument type, and voice leading have the largest influence. A multilevel model was created and a discussion of classical versus popular ratings and a subject's self-identification as a musician takes place. Suggestions for future research are presented.

1 Introduction

The aim of this study is to provide statistical analysis on how to identify if a piece of music sounds classical or popular given certain variables. Given previous research, certain variables (e.g. contrary motion) are associated with classical music, while other variables are associated with popular music.

In this paper, the following questions will be addressed: (1) what experimental factors and/or combination of factors have the strongest influence on the perception of music? (2) what differences in perception of music are there for those who identify as musicians and those who do not? and (3) what differences are there in the variables that make music sound classical versus popular?

2 Methods

The data set used was collected in 2012 from 70 University of Pittsburgh undergraduate students by Ivan Jimenez, a composer and musicologist, and Vincent Rossi, a University of Pittsburgh student. Each of the 70 students was asked to listen to 36 pieces of music and rank them on two different scales: (1) how classical does the music sound? and (2) how popular does the music sound? The minimum score is one, meaning it does not sound classical or popular at all, and the

[1] Student, Masters of Statistical Practice, Department of Statistics & Data Science, Carnegie Mellon University

maximum score is ten, meaning it sounds very classical or very popular. The 36 pieces of music were selected based on their instrument, harmonic motion, and voice leading. A sufficient combination of all three factors was attempt.

The following is a summary of the variables used in this paper:

- Classical - x_1 - how classical does the piece of music presented sound? (1 = not classical at all, 10 = very classical), this rating was provided by the Subject

- Popular - x_2 - how classical does the piece of music presented sound? (1 = not popular at all, 10 = very popular), this rating was provided by the Subject

- Subject - x_3 - subject ID, one of the 70 undergraduate students who participated in the study

- Harmony - x_4 - harmonic motion (4 levels: I-V-vi, I-VI-V, I-V-IV, IV-I-V)

- Instrument - x_5 - instrument type (3 levels: string quartet, piano, electric guitar)

- Voice - x_6 - voice leading (3 levels: contrary motion, parallel 3rds, parallel 5ths)

- Selfdeclare - x_7 - are you a musician? (scale 1 to 6, 1 being "not at all")

- OMSI - x_8 - score on a musical knowledge exam

- X16.minus.17 - x_9 - additional measure of Subject's ability to distinguish classical versus popular music

- ConsInstr - x_{10} - how much did you concentrate on the instrument while listening? (scale 0 to 5, 0 being "not at all")

- ConsNotes - x_{11} - how much did you concentrate on the notes while listening? (scale 0 to 5, 0 being "not at all")

- Instr.minus.Notes - x_{12} - difference of ConsInstr and ConsNotes

- PachListen - x_{13} - how familiar are you with Pachelbel's Canon in D? (scale 0 to 5, 0 being "not at all")

- ClsListen - x_{14} - how much do you listen to classical music? (scale 0 to 5, 0 being "not at all")

- KnowRob - x_{15} - have you heard of Rob Paravonian's Pachelbel Rant? (scale 0 to 5, 0 being "not at all")
- KnowAxis - x_{16} - have you heard of Axis of Evil's comedy bit on the four Pachelbel chords in popular music? (scale 0 to 5, 0 being "not at all")
- X1990s2000s - x_{17} - how much do you listen to pop and rock from the 1990s and 2000s? (scale 0 to 5, 0 being "not at all")
- X1990s2000s.minus.1960s1970s - x_{18} - difference of X1990s2000s and a variable similar to X1990s2000s about 1960s and 1970s pop and rock
- CollegeMusic - x_{19} - have you taken music classes in college (binary, 0 is "no", 1 is "yes")

  NoClass - x_{20} - how many music classes have you taken?
- APTheory - x_{21} - did you take AP Music Theory in high school? (binary, 0 is "no", 1 is "yes")
- Composing - x_{22} - how much experience do you have composing music? (scale 0 to 5, 0 being "not at all")
- PianoPlay - x_{23} - do you play the piano? (scale 0 to 5, 0 being "not at all")
- GuitarPlay - x_{24} - do you play the guitar? (scale 0 to 5, 0 being "not at all")
- X1stInstr - x_{25} - how proficient are you at your first musical instrument? (scale 0 to 5, 0 being "not at all")
- X2ndInstr - x_{26} - how proficient are you at your second musical instrument? (scale 0 to 5, 0 being "not at all")

Figure 1 presents how many missing or NA values were in each column. Columns 24 (X1stInstr) and 25 (X2ndInstr) were identified as having extraordinary amounts of rows with NAs, 60% and 87% respectively. Rather than delete rows, we decided to delete these two columns.

Column	Number_of_NAs
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	360
11	0
12	72
13	36
14	180
15	288
16	144
17	180
18	108
19	288
20	216
21	72
22	0
23	0
24	1512
25	2196
26	0
27	27
28	27

Figure 1:[2]

The core statistical method used in this paper is multilevel models. Multilevel models can be notated formulaically as follows:

$$y_i = \alpha_{0j[i]} + \alpha_1 x_i + \varepsilon_i$$
$$\alpha_{0j} = \beta_0 + \beta_1 u_j + \eta_j$$

[2] Please view Appendix HW 10 #1 for code.

Key elements of multilevel models are fixed and random effects. Fixed effects are effects that attempt to be estimated in our model. They are “fixed but unknown”. Random effects are draws from a non-fixed distribution (Junker, 2019, pp. 7-8).

3 Results

4 Discussion

References

Junker, Brian, *Intro to Multi-level Models, II*. Presented during 36-617 Applied Linear Models, November 11, 2019, Pittsburgh, PA.

R Core Team (2017), *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Shether, S.J. (2009), *A Modern Approach to Regression with R*. New York: Springer Science+Business Media, LLC.

Appendix

hecht_sophia_hw_10

Sophia Hecht

12/14/2019

```
# libraries used
library(psych)
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(stats)
library(MASS)

# load in data
music_unedited <- read.csv("~/Desktop/Linear/ratings.csv")
```

1

Identify anything unusual in the data set.
Missingness

```
# cute little loop that identifies the number of NAs in each column
na_count <- c()
for (i in 1:ncol(music_unedited)) {
  na_count[i] <- sum(is.na(music_unedited[, i]))
}
print(data.frame(na_count))
```

```
##    na_count
## 1         0
## 2         0
## 3         0
## 4         0
## 5         0
## 6         0
## 7         0
## 8         0
## 9         0
## 10      360
## 11        0
## 12       72
## 13       36
## 14      180
## 15      288
## 16      144
## 17      180
## 18      108
## 19      288
## 20      216
```

```
## 21       72
## 22        0
## 23        0
## 24     1512
## 25     2196
## 26        0
## 27       27
## 28       27
```

Column 10 has 360 NAs. Column 12 has 72 NAs. Column 13 has 36 NAs. Column 14 has 180 NAs. Column 15 has 288 NAs. Column 16 has 144 NAs. Column 17 has 180 NAs. Column 18 has 108 NAs. Column 19 has 288 NAs. Column 20 has 216 NAs. Column 21 has 72 NAs. Column 24 has 1,512 NAs. Column 25 has 2,196 NAs. Column 27 has 27 NAs. Column 28 has 27 NAs. There are only 2,520 rows, so columns #24 and #25 have a concerning amount of NAs with 60% missing for column #24 and 87% missing for column #25. My solution is to delete these two columns because the missingness is overwhelming. I am concerned that even if I fill in the missing values, that the missingness is so severe that it will affect the model.

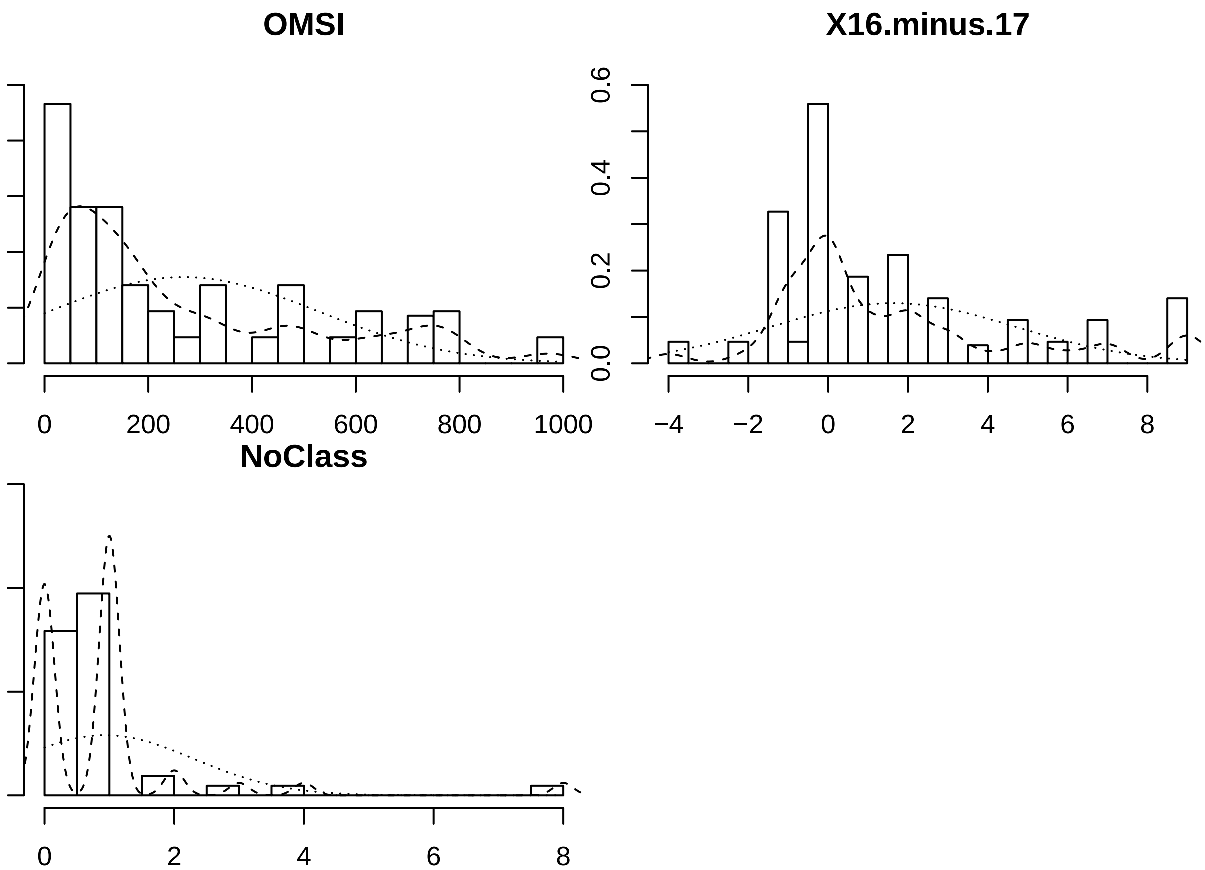
```
# delete columns #24 and #25
# I also deleted column #1 and #28 because Junker said so
music <- music_unedited
music <- music[, ! names(music_unedited) %in% c("X1stInstr", "X2ndInstr",
                                                 "first12", "X")]
```

For the other missing values, the highest one is still only 14% on the orginal data sets rows, therefore I have decided to delete all the leftover columns that have NAs.

```
# delete all the rows with NAs
music <- music[complete.cases(music), ]
```

This deleted 979 rows (~39% of the original data set). At this point, I have 1541 rows and 25 columns.
Transformations
To identify if there are any variables in desperate need of transformation, I am going to make a matrix of histograms for the numeric variables.
Numeric columns: OMSI, X16.minus.17, NoClass

```
# multi.hist() is from library "psych"
multi.hist(music[, c("OMSI", "X16.minus.17", "NoClass")])
```

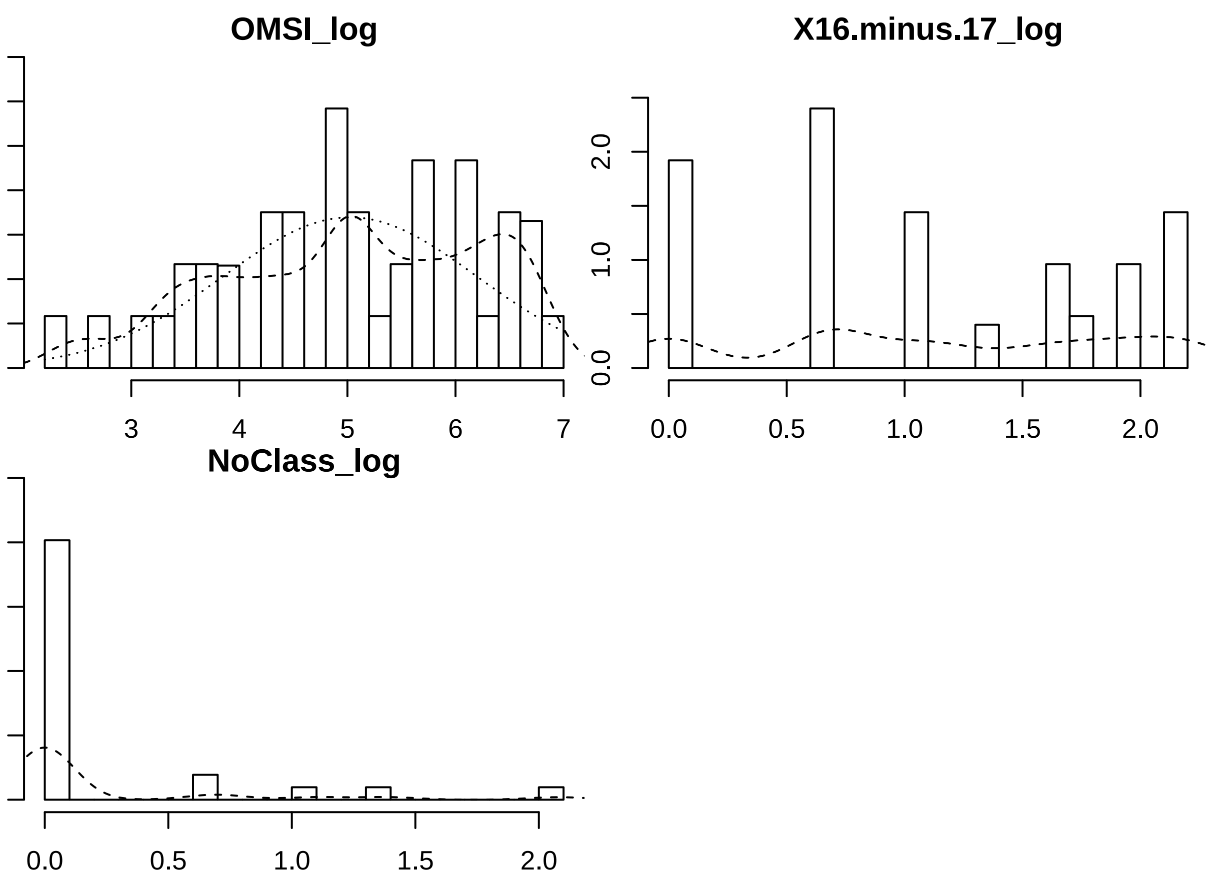


Looks like all three variables could go for some transformations! I am going to try logarthmic transformations.

```
# Logarithn Transformations for the 3 variables
# OMSI
music$OMSI_log <- NA
music$OMSI_log<- log(music$OMSI)
# X16.minus.17
music$X16.minus.17_log <- NA
music$X16.minus.17_log <- log(music$X16.minus.17)
```

```
## Warning in log(music$X16.minus.17): NaNs produced
```

```
# NoClass
music$NoClass_log <- NA
music$NoClass_log<- log(music$NoClass)
# try multi.hist() again, any improvement?
multi.hist(music[, c("OMSI_log", "X16.minus.17_log", "NoClass_log")])
```



It appears the only variable of the three numeric variables that responded well to the log transformations was OMSI. I am going to delete X16.minus.17_log and NoClass_log, because I am not going to use them in my analysis.

```
music <- music[, ! names(music) %in% c("X16.minus.17_log","NoClass_log")]
```

SUMMARY:

Got rid of NAs. Log transformed OMSI.
2 (a)
basic model :: Classical ~ Harmony + Instrument + Voice

```
# 5 ANOVA models with all the possible interactions
model_aov_1 <- aov(Classical ~ Harmony + Instrument + Voice, data = music)
model_aov_2 <- aov(Classical ~ Harmony*Instrument + Voice, data = music)
model_aov_3 <- aov(Classical ~ Harmony + Instrument*Voice, data = music)
model_aov_4 <- aov(Classical ~ Harmony*Voice + Instrument, data = music)
model_aov_5 <- aov(Classical ~ Harmony*Instrument*Voice, data = music)
summary(model_aov_1)
```

```
##               Df Sum Sq Mean Sq F value   Pr(>F)
## Harmony        3    200    66.5  12.871 2.63e-08 ***
## Instrument     2   3324  1661.8 321.504  < 2e-16 ***
## Voice          2     46    22.9   4.423   0.0122 *
## Residuals   1533   7924     5.2
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(model_aov_2)
```

```
##                       Df Sum Sq Mean Sq F value  Pr(>F)
## Harmony                3    200    66.5  12.854 2.7e-08 ***
## Instrument             2   3324  1661.8 321.072 < 2e-16 ***
## Voice                  2     46    22.9   4.417  0.0122 *
## Harmony:Instrument     6     20     3.4   0.657  0.6847
## Residuals           1527   7903     5.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(model_aov_3)
```

```
##                     Df Sum Sq Mean Sq F value   Pr(>F)
## Harmony              3    200    66.5  12.868 2.64e-08 ***
## Instrument           2   3324  1661.8 321.417  < 2e-16 ***
## Voice                2     46    22.9   4.422   0.0122 *
## Instrument:Voice     4     19     4.6   0.897   0.4652
## Residuals         1529   7905     5.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(model_aov_4)
```

```
##                  Df Sum Sq Mean Sq F value   Pr(>F)
## Harmony           3    200    66.5  12.953 2.34e-08 ***
## Voice             2     47    23.3   4.532   0.0109 *
## Instrument        2   3323  1661.3 323.469  < 2e-16 ***
## Harmony:Voice     6     81    13.5   2.625   0.0155 *
## Residuals      1527   7843     5.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(model_aov_5)
```

```
##                             Df Sum Sq Mean Sq F value   Pr(>F)
## Harmony                      3    200    66.5  12.940 2.39e-08 ***
## Instrument                   2   3324  1661.8 323.229  < 2e-16 ***
## Voice                        2     46    22.9   4.447   0.0119 *
## Harmony:Instrument           6     20     3.4   0.661   0.6811
## Harmony:Voice                6     81    13.5   2.619   0.0157 *
## Instrument:Voice             4     18     4.6   0.895   0.4658
## Harmony:Instrument:Voice    12     67     5.6   1.080   0.3730
## Residuals                 1505   7737     5.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# 5 simple linear models with all the possible interactions
model_simp_1 <- lm(Classical ~ Harmony + Instrument + Voice, data = music)
model_simp_2 <- lm(Classical ~ Harmony*Instrument + Voice, data = music)
model_simp_3 <- lm(Classical ~ Harmony + Instrument*Voice, data = music)
model_simp_4 <- lm(Classical ~ Harmony*Voice + Instrument, data = music)
model_simp_5 <- lm(Classical ~ Harmony*Instrument*Voice, data = music)
summary(model_simp_1)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony + Instrument + Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.0059 -1.5747 -0.0737  1.6963  6.4801
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.871116   0.163663  23.653  < 2e-16 ***
## HarmonyI-V-IV     0.001366   0.163755   0.008   0.9933
## HarmonyI-V-VI     0.845799   0.163863   5.162 2.77e-07 ***
## HarmonyIV-I-V     0.056710   0.163649   0.347   0.7290
## Instrumentpiano   1.654607   0.142025  11.650  < 2e-16 ***
## Instrumentstring  3.586789   0.141610  25.329  < 2e-16 ***
## Voicepar3rd      -0.407902   0.141885  -2.875   0.0041 **
## Voicepar5th      -0.297787   0.141886  -2.099   0.0360 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.273 on 1533 degrees of freedom
## Multiple R-squared:  0.3105, Adjusted R-squared:  0.3074
## F-statistic: 98.64 on 7 and 1533 DF,  p-value: < 2.2e-16
```

```r
summary(model_simp_2)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony * Instrument + Voice, data = music)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.127 -1.587 -0.093  1.609  6.413
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      4.06451    0.21643  18.780  < 2e-16 ***
## HarmonyI-V-IV                   -0.23256    0.28327  -0.821  0.41179
## HarmonyI-V-VI                    0.43019    0.28383   1.516  0.12980
## HarmonyIV-I-V                   -0.06977    0.28327  -0.246  0.80549
## Instrumentpiano                  1.32630    0.28383   4.673 3.23e-06 ***
## Instrumentstring                 3.33333    0.28327  11.767  < 2e-16 ***
## Voicepar3rd                     -0.40734    0.14198  -2.869  0.00417 **
```

```
## Voicepar5th                     -0.29781    0.14198  -2.098  0.03611 *
## HarmonyI-V-IV:Instrumentpiano    0.35375    0.40179   0.880  0.37876
## HarmonyI-V-VI:Instrumentpiano    0.65076    0.40218   1.618  0.10585
## HarmonyIV-I-V:Instrumentpiano    0.31110    0.40139   0.775  0.43843
## HarmonyI-V-IV:Instrumentstring   0.34884    0.40061   0.871  0.38402
## HarmonyI-V-VI:Instrumentstring   0.59694    0.40100   1.489  0.13679
## HarmonyIV-I-V:Instrumentstring   0.06977    0.40061   0.174  0.86177
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.275 on 1527 degrees of freedom
## Multiple R-squared:  0.3123, Adjusted R-squared:  0.3065
## F-statistic: 53.34 on 13 and 1527 DF,  p-value: < 2.2e-16
```

```r
summary(model_simp_3)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony + Instrument * Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7909 -1.5922 -0.0312  1.6189  6.5299
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     3.830638   0.200619  19.094  < 2e-16 ***
## HarmonyI-V-IV                   0.001269   0.163778   0.008   0.9938
## HarmonyI-V-VI                   0.845039   0.163886   5.156 2.85e-07 ***
## HarmonyIV-I-V                   0.056155   0.163672   0.343   0.7316
## Instrumentpiano                 1.657835   0.246267   6.732 2.36e-11 ***
## Instrumentstring                3.705374   0.245548  15.090  < 2e-16 ***
## Voicepar3rd                    -0.416719   0.245548  -1.697   0.0899 .
## Voicepar5th                    -0.166719   0.245548  -0.679   0.4973
## Instrumentpiano:Voicepar3rd    -0.041782   0.348022  -0.120   0.9045
## Instrumentstring:Voicepar3rd    0.067882   0.347003   0.196   0.8449
## Instrumentpiano:Voicepar5th     0.032336   0.348021   0.093   0.9260
## Instrumentstring:Voicepar5th -0.423397   0.347003  -1.220   0.2226
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.274 on 1529 degrees of freedom
## Multiple R-squared:  0.3121, Adjusted R-squared:  0.3072
## F-statistic: 63.08 on 11 and 1529 DF,  p-value: < 2.2e-16
```

```r
summary(model_simp_4)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony * Voice + Instrument, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -6.9281 -1.6273 -0.0412  1.7119  6.0526
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)                3.8032     0.2156  17.642  < 2e-16 ***
## HarmonyI-V-IV              0.2223     0.2827   0.786   0.4318
## HarmonyI-V-VI              1.2619     0.2833   4.454 9.03e-06 ***
## HarmonyIV-I-V             -0.3023     0.2822  -1.071   0.2842
## Voicepar3rd               -0.3101     0.2822  -1.099   0.2720
## Voicepar5th               -0.1917     0.2827  -0.678   0.4978
## Instrumentpiano            1.6554     0.1416  11.692  < 2e-16 ***
## Instrumentstring           3.5861     0.1412  25.404  < 2e-16 ***
## HarmonyI-V-IV:Voicepar3rd -0.4394     0.3995  -1.100   0.2715
## HarmonyI-V-VI:Voicepar3rd -0.7139     0.4002  -1.784   0.0747 .
## HarmonyIV-I-V:Voicepar3rd  0.7566     0.3995   1.894   0.0584 .
## HarmonyI-V-IV:Voicepar5th -0.2223     0.4002  -0.556   0.5786
## HarmonyI-V-VI:Voicepar5th -0.5314     0.4002  -1.328   0.1845
## HarmonyIV-I-V:Voicepar5th  0.3235     0.3995   0.810   0.4181
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.266 on 1527 degrees of freedom
## Multiple R-squared:  0.3176, Adjusted R-squared:  0.3118
## F-statistic: 54.66 on 13 and 1527 DF,  p-value: < 2.2e-16
```

```r
summary(model_simp_5)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony * Instrument * Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8488 -1.6512  0.0238  1.5581  6.7674
##
## Coefficients:
##                                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                                  3.93023    0.34578  11.366  < 2e-16
## HarmonyI-V-IV                               -0.23256    0.48900  -0.476  0.63444
## HarmonyI-V-VI                                1.23643    0.49190   2.514  0.01206
## HarmonyIV-I-V                               -0.48837    0.48900  -0.999  0.31809
## Instrumentpiano                              1.25581    0.48900   2.568  0.01032
## Instrumentstring                             3.60465    0.48900   7.371 2.77e-13
## Voicepar3rd                                  0.06977    0.48900   0.143  0.88657
## Voicepar5th                                 -0.37209    0.48900  -0.761  0.44682
## HarmonyI-V-IV:Instrumentpiano                0.71318    0.69361   1.028  0.30401
## HarmonyI-V-VI:Instrumentpiano                0.29181    0.69566   0.419  0.67493
## HarmonyIV-I-V:Instrumentpiano                0.60465    0.69155   0.874  0.38207
## HarmonyI-V-IV:Instrumentstring               0.65116    0.69155   0.942  0.34655
## HarmonyI-V-VI:Instrumentstring              -0.21318    0.69361  -0.307  0.75862
## HarmonyIV-I-V:Instrumentstring              -0.04651    0.69155  -0.067  0.94639
## HarmonyI-V-IV:Voicepar3rd                   -0.37209    0.69155  -0.538  0.59062
## HarmonyI-V-VI:Voicepar3rd                   -2.00388    0.69361  -2.889  0.00392
## HarmonyIV-I-V:Voicepar3rd                    0.41860    0.69155   0.605  0.54506
```

```
## HarmonyI-V-IV:Voicepar5th                     0.37209    0.69155   0.538  0.59062
## HarmonyI-V-VI:Voicepar5th                    -0.39922    0.69361  -0.576  0.56499
## HarmonyIV-I-V:Voicepar5th                     0.83721    0.69155   1.211  0.22623
## Instrumentpiano:Voicepar3rd                  -0.30233    0.69155  -0.437  0.66205
## Instrumentstring:Voicepar3rd                 -0.83721    0.69155  -1.211  0.22623
## Instrumentpiano:Voicepar5th                   0.51938    0.69361   0.749  0.45409
## Instrumentstring:Voicepar5th                  0.02326    0.69155   0.034  0.97318
## HarmonyI-V-IV:Instrumentpiano:Voicepar3rd  -0.41085    0.97946  -0.419  0.67493
## HarmonyI-V-VI:Instrumentpiano:Voicepar3rd   1.49834    0.98236   1.525  0.12741
## HarmonyIV-I-V:Instrumentpiano:Voicepar3rd  -0.03599    0.97946  -0.037  0.97069
## HarmonyI-V-IV:Instrumentstring:Voicepar3rd  0.20930    0.97800   0.214  0.83057
## HarmonyI-V-VI:Instrumentstring:Voicepar3rd  2.37597    0.97946   2.426  0.01539
## HarmonyIV-I-V:Instrumentstring:Voicepar3rd  1.04651    0.97800   1.070  0.28477
## HarmonyI-V-IV:Instrumentpiano:Voicepar5th  -0.66224    0.98236  -0.674  0.50033
## HarmonyI-V-VI:Instrumentpiano:Voicepar5th  -0.43909    0.98236  -0.447  0.65496
## HarmonyIV-I-V:Instrumentpiano:Voicepar5th  -0.84496    0.97946  -0.863  0.38845
## HarmonyI-V-IV:Instrumentstring:Voicepar5th -1.11628    0.97800  -1.141  0.25389
## HarmonyI-V-VI:Instrumentstring:Voicepar5th  0.03876    0.97946   0.040  0.96844
## HarmonyIV-I-V:Instrumentstring:Voicepar5th -0.69767    0.97800  -0.713  0.47573
##
## (Intercept)                                ***
## HarmonyI-V-IV
## HarmonyI-V-VI                              *
## HarmonyIV-I-V
## Instrumentpiano                            *
## Instrumentstring                           ***
## Voicepar3rd
## Voicepar5th
## HarmonyI-V-IV:Instrumentpiano
## HarmonyI-V-VI:Instrumentpiano
## HarmonyIV-I-V:Instrumentpiano
## HarmonyI-V-IV:Instrumentstring
## HarmonyI-V-VI:Instrumentstring
## HarmonyIV-I-V:Instrumentstring
## HarmonyI-V-IV:Voicepar3rd
## HarmonyI-V-VI:Voicepar3rd                  **
## HarmonyIV-I-V:Voicepar3rd
## HarmonyI-V-IV:Voicepar5th
## HarmonyI-V-VI:Voicepar5th
## HarmonyIV-I-V:Voicepar5th
## Instrumentpiano:Voicepar3rd
## Instrumentstring:Voicepar3rd
## Instrumentpiano:Voicepar5th
## Instrumentstring:Voicepar5th
## HarmonyI-V-IV:Instrumentpiano:Voicepar3rd
## HarmonyI-V-VI:Instrumentpiano:Voicepar3rd
## HarmonyIV-I-V:Instrumentpiano:Voicepar3rd
## HarmonyI-V-IV:Instrumentstring:Voicepar3rd
## HarmonyI-V-VI:Instrumentstring:Voicepar3rd *
## HarmonyIV-I-V:Instrumentstring:Voicepar3rd
## HarmonyI-V-IV:Instrumentpiano:Voicepar5th
## HarmonyI-V-VI:Instrumentpiano:Voicepar5th
## HarmonyIV-I-V:Instrumentpiano:Voicepar5th
## HarmonyI-V-IV:Instrumentstring:Voicepar5th
```

```
## HarmonyI-V-VI:Instrumentstring:Voicepar5th
## HarmonyIV-I-V:Instrumentstring:Voicepar5th
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.267 on 1505 degrees of freedom
## Multiple R-squared:  0.3267, Adjusted R-squared:  0.3111
## F-statistic: 20.87 on 35 and 1505 DF,  p-value: < 2.2e-16
```

Brief Summary:

From looking at the summaries, I believe that ANOVA is a better fit for making models than simple linear regression is. I would like to draw particular attention to ANOVA model 4 (model_aov_4) where all terms are significant at an alpha = 0.05 level. This included the interaction term of Harmony and Voice, suggesting that there is a significant difference in means of the variables. Because of model_aov_4, I have decided that I am going to include the interaction of harmony:voice in my future models.

2 (b) i.

I do not know why the photo will not stay with the numbering!! I am so sorry. Hopefully the labeling is good enough for you to tell where things are.

2 (b) i.

where ...

I = Instrument

H = Harmony

V = Voice

model :: Classical ~ Instrument + Harmony * Voice + (1 | Subject)

$$y_i = \alpha_{0j[i]} + \beta_2^{I} 1_{\{I_i=2\}} + \beta_3^{I} 1_{\{I_i=3\}} + \beta_2^{H} 1_{\{H_i=2\}} + \beta_3^{H} 1_{\{H_i=3\}} + \beta_4^{H} 1_{\{H_i=4\}} + \beta_2^{V} 1_{\{V_i=2\}} + \beta_3^{V} 1_{\{V_i=3\}} +$$
$$\beta_{22}^{HV} 1_{\{H_i=2, V_i=2\}} + \beta_{23}^{HV} 1_{\{H_i=2, V_i=3\}} + \beta_{24}^{HV} 1_{\{H_i=2, V_i=4\}} + \beta_{32}^{HV} 1_{\{H_i=3, V_i=2\}} + \beta_{33}^{HV} 1_{\{H_i=3, H_i=3\}} + \beta_{34}^{HV} 1_{\{H_i=3, V_i=4\}} +$$
$$\varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$\alpha_{0j} = \beta_0 + \eta_j, \quad \eta_j \overset{iid}{\sim} N(0, \tau^2)$$

2 (b) ii.

Test whether the random intercept is needed in the model. I am going to test whether the random intercept is needed by making a model with the random intercept and then comparing (with AIC) it to my favorite "simple" model which is model_aov_4 in this case. If the AIC is smaller for the model with the random intercept, then the random intercept is needed.

```r
# lmer() is from package "lme4"
# AIC() is from package "stats"
model_rand_int <- lmer(Classical ~ Instrument + Harmony*Voice + (1 | Subject),
                       data = music, REML = FALSE)
AIC(model_rand_int)
```

```
## [1] 6523.784
```

```r
AIC(model_aov_4)
```

```
## [1] 6910.615
```

Answer: Is the random effect needed?
The AICs show that the model with the random effect is favored. It has the lowest AIC ($6524 < 6911$).

2 (b) iii.

Create a model with Instrument, Harmony, and Voice on Classical using the repeated measures model with the random intercept for participants (aka what is above). I asked Laura Zhang what I was supposed to do here and she said that Junker said that I need to interpret the model, so here we go.

```r
summary(model_rand_int)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Instrument + Harmony * Voice + (1 | Subject)
##    Data: music
##
##      AIC      BIC   logLik deviance df.resid
##   6523.8   6609.2  -3245.9   6491.8     1525
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9954 -0.6377 -0.0015  0.6253  3.4122
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.420    1.191
##  Residual             3.668    1.915
## Number of obs: 1541, groups:  Subject, 43
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                   3.8078     0.2573  14.799
## Instrumentpiano               1.6456     0.1197  13.748
## Instrumentstring              3.5821     0.1193  30.029
## HarmonyI-V-IV                 0.2185     0.2389   0.915
## HarmonyI-V-VI                 1.2745     0.2394   5.323
## HarmonyIV-I-V                -0.3023     0.2385  -1.268
## Voicepar3rd                  -0.3101     0.2385  -1.300
## Voicepar5th                  -0.1955     0.2389  -0.818
## HarmonyI-V-IV:Voicepar3rd  -0.4356     0.3376  -1.290
## HarmonyI-V-VI:Voicepar3rd  -0.7302     0.3382  -2.159
## HarmonyIV-I-V:Voicepar3rd   0.7528     0.3376   2.230
## HarmonyI-V-IV:Voicepar5th  -0.2185     0.3382  -0.646
## HarmonyI-V-VI:Voicepar5th  -0.5402     0.3383  -1.597
## HarmonyIV-I-V:Voicepar5th   0.3273     0.3376   0.970

##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

The baseline is where Harmony is I-IV-V, Instrument is guitar, and Voice is contrary.
Given all other variables are held constant, if the Instrument is piano, then the rating of Classical is expected to increase by 1.6456, relative to the baseline.
Given all other variables are held constant, if the Instrument is a string, then the rating of Classical is expected to increase by 3.5821, relative to the baseline.
Given all other variables are held constant, if the Harmony is I-V-IV, then the rating of Classical is expected to increase by 0.2185, relative to the baseline.
Given all other variables are held constant, if the Harmony is I-V-VI, then the rating of Classical is expected to increase by 1.2745, relative to the baseline.
Given all other variables are held constant, if the Harmony is IV-I-V, then the rating of Classical is expected to decrease by 0.3023, relative to the baseline.
Given all other variables are held constant, if the Voice is par3rd, then the rating of Classical is expected to decrease by 0.3101, relative to the baseline.
Given all other variables are held constant, if the Voice is par5th, then the rating of Classical is expected to decrease by 0.1955, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-IV and Voice is par3rd, then the rating of Classical is expected to decrease by 0.4356, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-VI and Voice is par3rd, then the rating of Classical is expected to decrease by 0.7302, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to IV-I-V and Voice is par3rd, then the rating of Classical is expected to increase by 0.7528, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-IV and Voice is par5th, then the rating of Classical is expected to decrease by 0.2185, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-VI and Voice is par5th, then the rating of Classical is expected to decrease by 0.5402, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to IV-I-V and Voice is par5th, then the rating of Classical is expected to increase by 0.3273, relative to the baseline.
For the variance random effect of subject, the standard deviation is 1.191.

2 (c) i.

Edit model_rand_int by adding new random effects suggested and then decide which model is the best through an ANOVA test.

```
# anova() is from package "stats"
model_c_1 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Instrument | Subject), data = music, REML = FALSE)
model_c_2 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Harmony | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```
model_c_3 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Voice | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```
model_c_4 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Harmony + Instrument | Subject), data = music,
                  REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
model_c_5 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Voice + Instrument | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
model_c_6 <- lmer(Classical ~ Harmony + Instrument + Voice +
                    (Voice + Harmony + Instrument | Subject), data = music,
                  REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
model_c_7 <- lmer(Classical ~ Instrument + Harmony*Voice +
                    (Instrument + Harmony | Subject), data = music,
                  REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
model_c_8 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Voice + Harmony | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
model_c_9 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Voice + Instrument + Harmony | Subject), data = music,
                   REML = FALSE)
model_c_10 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Instrument + Voice | Subject), data = music, REML = FALSE)
model_c_11 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Harmony + Voice | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
model_c_12 <- lmer(Classical ~ Harmony*Voice + Instrument +
                    (Harmony + Instrument + Voice | Subject), data = music, REML = FALSE)
anova(model_rand_int, model_c_1, model_c_2, model_c_3, model_c_4, model_c_5,
      model_c_6, model_c_7, model_c_8, model_c_9, model_c_10, model_c_11, model_c_12)
```

```
## Data: music
## Models:
## model_rand_int: Classical ~ Instrument + Harmony * Voice + (1 | Subject)
## model_c_1: Classical ~ Harmony * Voice + Instrument + (Instrument | Subject)
## model_c_3: Classical ~ Harmony * Voice + Instrument + (Voice | Subject)
## model_c_2: Classical ~ Harmony * Voice + Instrument + (Harmony | Subject)
## model_c_5: Classical ~ Harmony * Voice + Instrument + (Voice + Instrument |
## model_c_5:     Subject)
## model_c_10: Classical ~ Harmony * Voice + Instrument + (Instrument + Voice |
## model_c_10:     Subject)
## model_c_4: Classical ~ Harmony * Voice + Instrument + (Harmony + Instrument |
## model_c_4:     Subject)
```

```
## model_c_7: Classical ~ Instrument + Harmony * Voice + (Instrument + Harmony |
## model_c_7:     Subject)
## model_c_8: Classical ~ Harmony * Voice + Instrument + (Voice + Harmony |
## model_c_8:     Subject)
## model_c_11: Classical ~ Harmony * Voice + Instrument + (Harmony + Voice |
## model_c_11:     Subject)
## model_c_6: Classical ~ Harmony + Instrument + Voice + (Voice + Harmony +
## model_c_6:     Instrument | Subject)
## model_c_9: Classical ~ Harmony * Voice + Instrument + (Voice + Instrument +
## model_c_9:     Harmony | Subject)
## model_c_12: Classical ~ Harmony * Voice + Instrument + (Harmony + Instrument +
## model_c_12:     Voice | Subject)
##                Df    AIC    BIC  logLik deviance    Chisq Chi Df Pr(>Chisq)
## model_rand_int 16 6523.8 6609.2 -3245.9   6491.8
## model_c_1      21 6271.9 6384.1 -3115.0   6229.9 261.8645      5  < 2.2e-16 ***
## model_c_3      21 6533.4 6645.6 -3245.7   6491.4   0.0000      0          1
## model_c_2      25 6487.0 6620.5 -3218.5   6437.0  54.4276      4  4.282e-11 ***
## model_c_5      30 6281.5 6441.7 -3110.8   6221.5 215.4998      5  < 2.2e-16 ***
## model_c_10     30 6281.7 6441.9 -3110.8   6221.7   0.0000      0          1
## model_c_4      36 6190.5 6382.7 -3059.2   6118.5 103.2022      6  < 2.2e-16 ***
## model_c_7      36 6190.1 6382.3 -3059.0   6118.1   0.3874      0  < 2.2e-16 ***
## model_c_8      36 6500.9 6693.1 -3214.4   6428.9   0.0000      0          1
## model_c_11     36 6499.7 6692.0 -3213.9   6427.7   1.1362      0  < 2.2e-16 ***
## model_c_6      45 6219.2 6459.6 -3064.6   6129.2 298.4670      9  < 2.2e-16 ***
## model_c_9      51 6198.5 6470.9 -3048.3   6096.5  32.7331      6  1.180e-05 ***
## model_c_12     51 6198.1 6470.5 -3048.1   6096.1   0.3687      0  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the ANOVA, the model with the lowest AIC and BIC is model_c_7 whos' formula is the following: Classical ~ Harmony*Voice + Instrument + (Instrument + Harmony | Subject).

2 (c) ii.

Interpret the model that you chose in 2 (c) ii (model_c_7).
Classical ~ Harmony*Voice + Instrument + (Instrument + Harmony | Subject)
Comment on the sizes of the variances of the random effects with respect to (a) each other and (b) the estimated residual variance.

```r
summary(model_c_7)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Instrument + Harmony * Voice + (Instrument + Harmony |
##     Subject)
##    Data: music
##
##      AIC      BIC   logLik deviance df.resid
##   6190.1   6382.3  -3059.0   6118.1     1505
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6550 -0.5648  0.0178  0.5238  3.4056
```

```
##
## Random effects:
##  Groups   Name              Variance Std.Dev. Corr
##  Subject  (Intercept)       1.68937  1.2998
##           Instrumentpiano   1.92062  1.3859   -0.26
##           Instrumentstring  3.68496  1.9196   -0.54  0.62
##           HarmonyI-V-IV     0.08731  0.2955    0.78 -0.75 -0.88
##           HarmonyI-V-VI     1.74687  1.3217    0.21 -0.40 -0.60  0.49
##           HarmonyIV-I-V     0.16079  0.4010    0.11 -0.22 -0.24  0.23  0.39
##  Residual                   2.44098  1.5624
## Number of obs: 1541, groups:  Subject, 43
##
## Fixed effects:
##                           Estimate Std. Error t value
## (Intercept)                 3.8034     0.2477  15.352
## Instrumentpiano             1.6532     0.2329   7.099
## Instrumentstring            3.5877     0.3085  11.630
## HarmonyI-V-IV               0.2127     0.2001   1.063
## HarmonyI-V-VI               1.2662     0.2808   4.510
## HarmonyIV-I-V              -0.3023     0.2039  -1.483
## Voicepar3rd                -0.3101     0.1945  -1.594
## Voicepar5th                -0.2038     0.1950  -1.045
## HarmonyI-V-IV:Voicepar3rd  -0.4298     0.2754  -1.560
## HarmonyI-V-VI:Voicepar3rd  -0.7074     0.2760  -2.563
## HarmonyIV-I-V:Voicepar3rd   0.7514     0.2754   2.728
## HarmonyI-V-IV:Voicepar5th  -0.2103     0.2760  -0.762
## HarmonyI-V-VI:Voicepar5th  -0.5236     0.2761  -1.896
## HarmonyIV-I-V:Voicepar5th   0.3356     0.2754   1.218

##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

The baseline is where Harmony is I-IV-V, Instrument is guitar, and Voice is contrary.
Given all other variables are held constant, if the Instrument is piano, then the rating of Classical is expected to increase by 1.6532, relative to the baseline.
Given all other variables are held constant, if the Instrument is a string, then the rating of Classical is expected to increase by 3.5877, relative to the baseline.
Given all other variables are held constant, if the Harmony is I-V-IV, then the rating of Classical is expected to increase by 0.2127, relative to the baseline.
Given all other variables are held constant, if the Harmony is I-V-VI, then the rating of Classical is expected to increase by 1.2662, relative to the baseline.
Given all other variables are held constant, if the Harmony is IV-I-V, then the rating of Classical is expected to decrease by 0.3023, relative to the baseline.
Given all other variables are held constant, if the Voice is par3rd, then the rating of Classical is expected to decrease by 0.3101, relative to the baseline.
Given all other variables are held constant, if the Voice is par5th, then the rating of Classical is expected to decrease by 0.2038, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-IV and Voice is par3rd, then the

rating of Classical is expected to decrease by 0.4298, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-VI and Voice is par3rd, then the rating of Classical is expected to decrease by 0.7074, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to IV-I-V and Voice is par3rd, then the rating of Classical is expected to increase by 0.7514, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-IV and Voice is par5th, then the rating of Classical is expected to decrease by 0.2103, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-VI and Voice is par5th, then the rating of Classical is expected to decrease by 0.5236, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to IV-I-V and Voice is par5th, then the rating of Classical is expected to increase by 0.3356, relative to the baseline.
The size of the variance of the random effects in respect to each other::
The highest variance is from String Instrument with a whooping 3.68496! wow The lowest is HarmonyI-V-IV at 0.08731.
The size of the variance of the random effects in respect to the estimated residual variance::
The estimated residual variance is 2.44098 (which seems high to me??) Most of the random effects variances are lower than that (actually all expect one, String Instruments).

2 (c) iii.

I do not know why the photo will not stay with the numbering!! I am so sorry. Hopefully the labeling is good enough for you to tell where things are.

2 (c) iii.

where...

I = Instrument
H = Harmony
V = Voice

model:: Classical ~ Instrument + Harmony * Voice + (Instrument + Harmony | Subject)

$$
\begin{aligned}
Y_i = \; & \alpha_{0j[i]} + \beta_2^I 1_{\{I_i=2\}} + \beta_3^I 1_{\{I_i=3\}} + \beta_2^H 1_{\{H_i=2\}} + \\
& \beta_3^H 1_{\{H_i=3\}} + \beta_4^H 1_{\{H_i=4\}} + \beta_2^V 1_{\{V_i=2\}} + \beta_3^V 1_{\{V_i=3\}} + \\
& \beta_{22}^{HV} 1_{\{H_i=2, V_i=2\}} + \beta_{23}^{HV} 1_{\{H_i=2, V_i=3\}} + \\
& \beta_{24}^{HV} 1_{\{H_i=2, V_i=4\}} + \beta_{32}^{HV} 1_{\{H_i=3, V_i=2\}} + \beta_{33}^{HV} 1_{\{H_i=3, V_i=3\}} + \\
& \beta_{34}^{HV} 1_{\{H_i=3, V_i=4\}} + \alpha_{2j[i]}^H 1_{\{H_i=2\}} + \alpha_{3j[i]}^H 1_{\{H_i=3\}} + \\
& \alpha_{4j[i]}^H 1_{\{H_i=4\}} + \alpha_{2j[i]}^I 1_{\{I_i=2\}} + \alpha_{3j[i]}^I 1_{\{I_i=3\}} + \\
& \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)
\end{aligned}
$$

$$\alpha_{0j} = \beta_0 + \eta_{0j}, \quad \eta_{0j} \overset{iid}{\sim} N(0, \tau^2)$$

$$\alpha_{2j}^I = \beta_2^I + \eta_{2Ij}, \quad \eta_{2jI} \overset{iid}{\sim} N(0, \tau^2)$$

$$\alpha_{3j}^I = \beta_3^I + \eta_{3Ij}, \quad \eta_{3Ij} \overset{iid}{\sim} N(0, \tau^2)$$

$$\alpha_{2j}^H = \beta_2^H + \eta_{2Hj}, \quad \eta_{2Hj} \overset{iid}{\sim} N(0, \tau^2)$$

$$\alpha_{3j}^H = \beta_3^H + \eta_{3Hj}, \quad \eta_{3Hj} \overset{iid}{\sim} N(0, \tau^2)$$

$$\alpha_{4j}^H = \beta_4^H + \eta_{4Hj}, \quad \eta_{4Hj} \overset{iid}{\sim} N(0, \tau^2)$$

3 (a)

Determine the individual covariates that should be added to model_c_4 as FIXED effects.

```r
# first step is to make sure that every column that needs to be a factor is a
# factor (as Junker said)
music$Harmony <- as.factor(music$Harmony)
music$Instrument <- as.factor(music$Instrument)
music$Voice <- as.factor(music$Voice)
music$Selfdeclare <- as.factor(music$Selfdeclare)
music$ConsInstr <- as.factor(music$ConsInstr)
music$ConsNotes <- as.factor(music$ConsNotes)
music$PachListen <- as.factor(music$PachListen)
music$ClsListen <- as.factor(music$ClsListen)
music$KnowRob <- as.factor(music$KnowRob)
music$KnowAxis <- as.factor(music$KnowAxis)
music$X1990s2000s <- as.factor(music$X1990s2000s)
music$CollegeMusic <- as.factor(music$CollegeMusic)
```

```r
music$APTheory <- as.factor(music$APTheory)
music$Composing <- as.factor(music$Composing)
music$PianoPlay <- as.factor(music$PianoPlay)
music$GuitarPlay <- as.factor(music$GuitarPlay)
```

Then do StepAIC, create a model, and make sure not to put in X or Popular or Classical and use the OMSI_log.

```r
# stepAIC() is from package "MASS"
model_simp_full <- lm(Classical ~  OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare
                      + ConsNotes + Instr.minus.Notes + PachListen + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory + Composing + PianoPlay + Subject +
                   GuitarPlay, data = music)
summary(model_simp_full)
```

```
##
## Call:
## lm(formula = Classical ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + Instr.minus.Notes + PachListen +
##     ClsListen + KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory + Composing + PianoPlay +
##     Subject + GuitarPlay, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8571 -1.8889 -0.1389  1.8611  6.6944
##
## Coefficients: (54 not defined because of singularities)
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -102.1697    28.1996  -3.623 0.000301 ***
## OMSI_log                       5.7115     1.9017   3.003 0.002714 **
## X16.minus.17                   0.7627     0.3305   2.308 0.021159 *
## ConsInstr0.67                 70.1057    16.4972   4.250 2.27e-05 ***
## ConsInstr1                    60.3760    15.0695   4.006 6.47e-05 ***
## ConsInstr1.67                 92.0110    21.8982   4.202 2.80e-05 ***
## ConsInstr2.33                 57.8953    14.0783   4.112 4.13e-05 ***
## ConsInstr2.67                 60.7954    15.6269   3.890 0.000104 ***
## ConsInstr3                   103.9358    26.1181   3.979 7.24e-05 ***
## ConsInstr3.33                123.2723    28.9466   4.259 2.18e-05 ***
## ConsInstr3.67                 87.4191    22.5175   3.882 0.000108 ***
## ConsInstr4                    77.7843    17.6535   4.406 1.13e-05 ***
## ConsInstr4.33                 96.9716    24.3275   3.986 7.04e-05 ***
## ConsInstr5                    96.6179    23.6356   4.088 4.59e-05 ***
## Selfdeclare2                   5.5914     2.3021   2.429 0.015267 *
## Selfdeclare3                  -2.0772     0.8376  -2.480 0.013253 *
## Selfdeclare4                 -21.5638     6.8779  -3.135 0.001751 **
## Selfdeclare5                 -15.4341     2.2553  -6.843 1.12e-11 ***
## Selfdeclare6                   4.5022     2.8151   1.599 0.109968
## ConsNotes1                   -11.5833     3.6622  -3.163 0.001593 **
## ConsNotes3                   -23.9557     6.3904  -3.749 0.000184 ***
## ConsNotes4                   -40.7967    10.3197  -3.953 8.07e-05 ***
## ConsNotes5                   -34.0591     8.5142  -4.000 6.64e-05 ***
```

```
## Instr.minus.Notes                NA         NA      NA       NA
## PachListen3                -38.2919     7.9675  -4.806 1.69e-06 ***
## PachListen4                      NA         NA      NA       NA
## PachListen5                -21.2283     5.8766  -3.612 0.000313 ***
## ClsListen1                   5.3836     1.9858   2.711 0.006782 **
## ClsListen3                  15.7480     3.9297   4.007 6.44e-05 ***
## ClsListen4                   7.0350     3.9332   1.789 0.073880 .
## ClsListen5                  28.8933     7.5882   3.808 0.000146 ***
## KnowRob1                     1.8651     1.0920   1.708 0.087852 .
## KnowRob5                    32.8840     8.5014   3.868 0.000114 ***
## KnowAxis1                    8.4372     2.3434   3.600 0.000328 ***
## KnowAxis5                  -40.6692    10.0218  -4.058 5.20e-05 ***
## X1990s2000s2                49.1798    10.7002   4.596 4.67e-06 ***
## X1990s2000s3                15.6187     3.9104   3.994 6.81e-05 ***
## X1990s2000s4                85.9962    20.3096   4.234 2.43e-05 ***
## X1990s2000s5                26.7339     5.5332   4.831 1.49e-06 ***
## X1990s2000s.minus.1960s1970s 3.0750     0.8790   3.498 0.000482 ***
## CollegeMusic1               -9.5230     2.5984  -3.665 0.000256 ***
## NoClass                      3.4189     0.9644   3.545 0.000404 ***
## APTheory1                   -5.0461     1.2675  -3.981 7.19e-05 ***
## Composing1                  -1.0043     1.6146  -0.622 0.534020
## Composing2                   8.5386     2.7884   3.062 0.002236 **
## Composing3                       NA         NA      NA       NA
## Composing4                       NA         NA      NA       NA
## Composing5                       NA         NA      NA       NA
## PianoPlay1                       NA         NA      NA       NA
## PianoPlay4                       NA         NA      NA       NA
## PianoPlay5                       NA         NA      NA       NA
## Subject17                        NA         NA      NA       NA
## Subject19                        NA         NA      NA       NA
## Subject20                        NA         NA      NA       NA
## Subject22                        NA         NA      NA       NA
## Subject23                        NA         NA      NA       NA
## Subject26                        NA         NA      NA       NA
## Subject29                        NA         NA      NA       NA
## Subject30                        NA         NA      NA       NA
## Subject31                        NA         NA      NA       NA
## Subject32                        NA         NA      NA       NA
## Subject37                        NA         NA      NA       NA
## Subject38                        NA         NA      NA       NA
## Subject40                        NA         NA      NA       NA
## Subject42                        NA         NA      NA       NA
## Subject44.1                      NA         NA      NA       NA
## Subject44.2                      NA         NA      NA       NA
## Subject45                        NA         NA      NA       NA
## Subject46                        NA         NA      NA       NA
## Subject47                        NA         NA      NA       NA
## Subject48                        NA         NA      NA       NA
## Subject49                        NA         NA      NA       NA
## Subject52                        NA         NA      NA       NA
## Subject53                        NA         NA      NA       NA
## Subject55                        NA         NA      NA       NA
## Subject56                        NA         NA      NA       NA
## Subject57                        NA         NA      NA       NA
```

```
## Subject59                      NA         NA      NA        NA
## Subject60                      NA         NA      NA        NA
## Subject61                      NA         NA      NA        NA
## Subject63                      NA         NA      NA        NA
## Subject64                      NA         NA      NA        NA
## Subject66                      NA         NA      NA        NA
## Subject71                      NA         NA      NA        NA
## Subject74                      NA         NA      NA        NA
## Subject78                      NA         NA      NA        NA
## Subject80                      NA         NA      NA        NA
## Subject81                      NA         NA      NA        NA
## Subject82                      NA         NA      NA        NA
## Subject83                      NA         NA      NA        NA
## Subject93                      NA         NA      NA        NA
## Subject94                      NA         NA      NA        NA
## Subject98                      NA         NA      NA        NA
## GuitarPlay1                    NA         NA      NA        NA
## GuitarPlay2                    NA         NA      NA        NA
## GuitarPlay4                    NA         NA      NA        NA
## GuitarPlay5                    NA         NA      NA        NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.47 on 1498 degrees of freedom
## Multiple R-squared:  0.2046, Adjusted R-squared:  0.1823
## F-statistic: 9.176 on 42 and 1498 DF,  p-value: < 2.2e-16
```

```
model_simp_full_2 <- lm(Classical ~  OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
                      + ConsNotes + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory, data = music)
summary(model_simp_full_2)
```

```
##
## Call:
## lm(formula = Classical ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + +ConsNotes + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + X1990s2000s.minus.1960s1970s + CollegeMusic +
##     NoClass + APTheory, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.8993 -1.8972 -0.0777  1.8833  7.0076
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   7.37447    5.83826   1.263 0.206739
## OMSI_log                     -0.13938    0.34410  -0.405 0.685498
## X16.minus.17                 -0.11632    0.11613  -1.002 0.316686
## ConsInstr0.67                -0.04947    3.65919  -0.014 0.989215
## ConsInstr1                   -2.80030    1.85650  -1.508 0.131668
## ConsInstr1.67                -0.12356    2.67326  -0.046 0.963139
## ConsInstr2.33                 0.37846    1.67657   0.226 0.821440
## ConsInstr2.67                 0.29000    1.85085   0.157 0.875515
```

```
## ConsInstr3                   -1.16543    3.60971  -0.323 0.746847
## ConsInstr3.33                 2.94888    4.14619   0.711 0.477054
## ConsInstr3.67                -0.63881    3.17943  -0.201 0.840789
## ConsInstr4                    2.98464    2.57941   1.157 0.247415
## ConsInstr4.33                -0.06545    3.19097  -0.021 0.983638
## ConsInstr5                    0.18642    2.76115   0.068 0.946179
## Selfdeclare2                 -2.59769    0.75137  -3.457 0.000561 ***
## Selfdeclare3                  0.31225    0.77622   0.402 0.687543
## Selfdeclare4                 -2.32070    1.17679  -1.972 0.048786 *
## Selfdeclare5                 -2.17218    0.95491  -2.275 0.023061 *
## Selfdeclare6                  1.11712    2.07790   0.538 0.590921
## ConsNotes1                    0.89964    0.83885   1.072 0.283681
## ConsNotes3                    0.06022    0.89385   0.067 0.946296
## ConsNotes4                   -1.45550    1.49539  -0.973 0.330548
## ConsNotes5                   -1.49087    1.20955  -1.233 0.217925
## ClsListen1                   -2.12643    1.07171  -1.984 0.047420 *
## ClsListen3                   -0.47784    1.40250  -0.341 0.733376
## ClsListen4                   -0.54242    1.00411  -0.540 0.589141
## ClsListen5                    1.09462    1.99913   0.548 0.584083
## KnowRob1                      0.02069    0.47559   0.044 0.965299
## KnowRob5                      0.78490    1.53444   0.512 0.609062
## KnowAxis1                     3.60199    1.10727   3.253 0.001167 **
## KnowAxis5                    -0.57756    1.44909  -0.399 0.690266
## X1990s2000s2                  3.43525    1.19198   2.882 0.004008 **
## X1990s2000s3                  0.98414    0.93737   1.050 0.293936
## X1990s2000s4                  3.84240    2.90892   1.321 0.186735
## X1990s2000s5                  4.37368    1.35525   3.227 0.001277 **
## X1990s2000s.minus.1960s1970s -0.23796    0.17948  -1.326 0.185102
## CollegeMusic1                -0.92691    0.33877  -2.736 0.006291 **
## NoClass                      -0.38480    0.22493  -1.711 0.087330 .
## APTheory1                     1.70794    0.60843   2.807 0.005063 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.523 on 1502 degrees of freedom
## Multiple R-squared:  0.1681, Adjusted R-squared:  0.147
## F-statistic: 7.986 on 38 and 1502 DF,  p-value: < 2.2e-16
```

```
stepAIC(model_simp_full_2, k = 2)
```

```
## Start:  AIC=2890.71
## Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
##     +ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory
##
##                                Df Sum of Sq    RSS    AIC
## - KnowRob                       2      2.04 9562.9 2887.0
## - OMSI_log                      1      1.04 9561.9 2888.9
## - X16.minus.17                  1      6.39 9567.3 2889.7
## - X1990s2000s.minus.1960s1970s  1     11.19 9572.1 2890.5
## <none>                                      9560.9 2890.7
## - NoClass                       1     18.63 9579.5 2891.7
## - CollegeMusic                  1     47.65 9608.5 2896.4
## - APTheory                      1     50.16 9611.0 2896.8
```

```
## - KnowAxis                   2     67.40 9628.3 2897.5
## - ConsNotes                  4    194.80 9755.7 2913.8
## - ClsListen                  4    284.05 9844.9 2927.8
## - Selfdeclare                5    306.84 9867.7 2929.4
## - ConsInstr                 11    391.01 9951.9 2930.5
## - X1990s2000s                4    356.96 9917.8 2939.2
##
## Step:  AIC=2887.04
## Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
##     ConsNotes + ClsListen + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory
##
##                               Df Sum of Sq    RSS    AIC
## - OMSI_log                     1      4.47 9567.4 2885.8
## <none>                                     9562.9 2887.0
## - NoClass                      1     25.48 9588.4 2889.1
## - X1990s2000s.minus.1960s1970s  1     27.92 9590.8 2889.5
## - X16.minus.17                 1     31.55 9594.5 2890.1
## - APTheory                     1     48.33 9611.2 2892.8
## - KnowAxis                     2     67.59 9630.5 2893.9
## - CollegeMusic                 1     59.64 9622.6 2894.6
## - ConsNotes                    4    217.93 9780.8 2913.8
## - ConsInstr                   11    417.79 9980.7 2930.9
## - X1990s2000s                  4    356.53 9919.4 2935.4
## - Selfdeclare                  5    387.36 9950.3 2938.2
## - ClsListen                    4    409.14 9972.1 2943.6
##
## Step:  AIC=2885.76
## Classical ~ X16.minus.17 + ConsInstr + Selfdeclare + ConsNotes +
##     ClsListen + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory
##
##                               Df Sum of Sq     RSS    AIC
## <none>                                      9567.4 2885.8
## - NoClass                      1     24.13  9591.5 2887.6
## - X16.minus.17                 1     27.13  9594.5 2888.1
## - X1990s2000s.minus.1960s1970s  1     37.40  9604.8 2889.8
## - KnowAxis                     2     65.08  9632.5 2892.2
## - CollegeMusic                 1     58.08  9625.5 2893.1
## - APTheory                     1     61.37  9628.8 2893.6
## - ConsNotes                    4    218.63  9786.0 2912.6
## - ConsInstr                   11    450.59 10018.0 2934.7
## - X1990s2000s                  4    366.28  9933.7 2935.7
## - Selfdeclare                  5    418.89  9986.3 2941.8
## - ClsListen                    4    408.56  9975.9 2942.2

##
## Call:
## lm(formula = Classical ~ X16.minus.17 + ConsInstr + Selfdeclare +
##     ConsNotes + ClsListen + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory, data = music)
##
## Coefficients:
##                  (Intercept)                  X16.minus.17
```

```
##                       8.03784                       -0.13766
##                 ConsInstr0.67                     ConsInstr1
##                      -1.14036                       -3.33680
##                 ConsInstr1.67                  ConsInstr2.33
##                      -1.35587                       -0.46166
##                 ConsInstr2.67                     ConsInstr3
##                      -0.35557                       -2.76034
##                 ConsInstr3.33                  ConsInstr3.67
##                       1.33044                       -2.06112
##                    ConsInstr4                  ConsInstr4.33
##                       2.16379                       -1.54388
##                    ConsInstr5                    Selfdeclare2
##                      -1.19268                       -2.56817
##                  Selfdeclare3                   Selfdeclare4
##                      -0.13478                       -1.99789
##                  Selfdeclare5                   Selfdeclare6
##                      -2.49257                        0.38325
##                    ConsNotes1                     ConsNotes3
##                       1.19706                        0.62266
##                    ConsNotes4                     ConsNotes5
##                      -0.64085                       -0.71317
##                    ClsListen1                     ClsListen3
##                      -1.91386                       -0.35117
##                    ClsListen4                     ClsListen5
##                      -0.33698                        0.82618
##                     KnowAxis1                      KnowAxis5
##                       3.17016                        0.05454
##                  X1990s2000s2                   X1990s2000s3
##                       3.21155                        0.69608
##                  X1990s2000s4                   X1990s2000s5
##                       2.32230                        3.63323
## X1990s2000s.minus.1960s1970s                  CollegeMusic1
##                      -0.19808                       -0.92449
##                       NoClass                      APTheory1
##                      -0.37742                        1.26435
```

```r
model_after_stepAIC <- lmer(Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
                              ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
                              X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory +
                              Harmony*Voice + Instrument + (Instrument + Harmony | Subject),
                            data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

Notes:

Notice in the summary of model_simp_full and model_simp_full it says "___ not defined because of singularities". This happens when variables are 100% collinear, therefore I progressively got rid of these variables. I ended up with model_simp_full_3. The step AIC of Model_simp_full_3 suggested variables.
3 (b)

```r
anova(model_after_stepAIC, model_c_7)
```

```
## Data: music
## Models:
## model_c_7: Classical ~ Instrument + Harmony * Voice + (Instrument + Harmony |
## model_c_7:     Subject)
## model_after_stepAIC: Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
## model_after_stepAIC:     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
## model_after_stepAIC:     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory +
## model_after_stepAIC:     Harmony * Voice + Instrument + (Instrument + Harmony | Subject)
##                     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_c_7           36 6190.1 6382.3 -3059.0   6118.1
## model_after_stepAIC 74 6174.8 6570.0 -3013.4   6026.8 91.286     38  2.809e-06
##
## model_c_7
## model_after_stepAIC ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Explanation:

I added the variables suggested by stepAIC() to the model I thought was best in Problem 2 (model_c_7. I then ran an ANOVA test to see which model performs better. Unfortunately model_c_7 has the smaller BIC, while model_after_stepAIC has the smaller AIC. model_c_7's BIC is 165.9 smaller than model_after_stepAIC's BIC, while model_after_stepAIC's AIC is 31.8 smaller than model_c_7's AIC. Therefore I have decided to continue on with model_c_7 because it's BIC is much lower than how much larger it's AIC is to model_after_stepAIC.
3 (c) Please view #2 (c) ii. It is exactly the same. Its weird how that worked out, but I am gonna trust my gut and AIC/BIC. ## 4
Make a new variable about whether or not a person self-identifies as a musician. 0 means you are a self-declared musician, 1 means you do not identify as a musician.

```r
music$Selfdeclare<- as.numeric(music$Selfdeclare)
music$Selfdeclare <- ifelse(music$Selfdeclare > 2, 1, 0)
music$Selfdeclare <- as.factor(music$Selfdeclare)
summary(music$Selfdeclare)
```

```
##   0   1
## 827 714
```

Now I am going to redo what I did in #3 and see if it makes a difference now that Selfdeclare is binary.

```r
model_selfdeclare_binary <- lm(Classical ~  OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
                                 ConsNotes + Instr.minus.Notes + PachListen + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory + Composing + PianoPlay + Subject +
                   GuitarPlay, data = music)
summary(model_selfdeclare_binary)
```

```
##
## Call:
```

```
## lm(formula = Classical ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + Instr.minus.Notes + PachListen +
##     ClsListen + KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory + Composing + PianoPlay +
##     Subject + GuitarPlay, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8571 -1.8889 -0.1389  1.8611  6.6944
##
## Coefficients: (50 not defined because of singularities)
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -3.25432    4.65298  -0.699 0.484407
## OMSI_log                      3.64583    0.54687   6.667 3.67e-11 ***
## X16.minus.17                  0.19962    0.07896   2.528 0.011573 *
## ConsInstr0.67                -8.87131    6.14288  -1.444 0.148903
## ConsInstr1                   -0.70769    5.58131  -0.127 0.899118
## ConsInstr1.67                 2.91678    7.47535   0.390 0.696454
## ConsInstr2.33                 1.21256    5.54293   0.219 0.826869
## ConsInstr2.67                 7.75527    4.87231   1.592 0.111662
## ConsInstr3                   -0.83865    7.02801  -0.119 0.905031
## ConsInstr3.33                 0.67118    7.90036   0.085 0.932308
## ConsInstr3.67                -1.87455    6.16492  -0.304 0.761119
## ConsInstr4                    4.02544    5.84092   0.689 0.490818
## ConsInstr4.33                 1.30566    6.85800   0.190 0.849033
## ConsInstr5                    3.21145    7.40759   0.434 0.664688
## Selfdeclare1                 -3.99124    1.21035  -3.298 0.000998 ***
## ConsNotes1                   -0.80888    0.91616  -0.883 0.377428
## ConsNotes3                    1.43757    1.41083   1.019 0.308388
## ConsNotes4                   -3.15224    2.29354  -1.374 0.169523
## ConsNotes5                   -0.46966    1.94452  -0.242 0.809177
## Instr.minus.Notes                  NA         NA      NA       NA
## PachListen3                  -6.03302    3.13626  -1.924 0.054589 .
## PachListen4                        NA         NA      NA       NA
## PachListen5                  -4.86227    2.03401  -2.390 0.016949 *
## ClsListen1                   -4.04798    0.97662  -4.145 3.59e-05 ***
## ClsListen3                   -2.21527    1.19973  -1.846 0.065020 .
## ClsListen4                  -20.85087    2.77688  -7.509 1.02e-13 ***
## ClsListen5                   -1.85222    1.21441  -1.525 0.127421
## KnowRob1                      0.21178    0.50937   0.416 0.677644
## KnowRob5                     -6.94931    1.64924  -4.214 2.66e-05 ***
## KnowAxis1                     4.34414    2.06083   2.108 0.035201 *
## KnowAxis5                     1.34683    2.09817   0.642 0.521031
## X1990s2000s2                 11.62089    3.07603   3.778 0.000164 ***
## X1990s2000s3                 -0.22931    1.77123  -0.129 0.897009
## X1990s2000s4                  3.97326    5.10434   0.778 0.436451
## X1990s2000s5                  2.44655    1.55198   1.576 0.115145
## X1990s2000s.minus.1960s1970s  0.25639    0.31141   0.823 0.410449
## CollegeMusic1                -4.34386    1.06504  -4.079 4.77e-05 ***
## NoClass                      -1.22676    0.92020  -1.333 0.182688
## APTheory1                    -3.24694    0.69062  -4.701 2.82e-06 ***
## Composing1                   -3.01167    0.85647  -3.516 0.000451 ***
## Composing2                    2.64419    1.03005   2.567 0.010353 *
## Composing3                   -0.98443    1.79612  -0.548 0.583716
```

```
## Composing4                  5.20148    1.05034   4.952 8.17e-07 ***
## Composing5                  8.34030    5.29057   1.576 0.115134
## PianoPlay1                  4.40673    0.64139   6.871 9.34e-12 ***
## PianoPlay4                       NA         NA      NA       NA
## PianoPlay5                       NA         NA      NA       NA
## Subject17                        NA         NA      NA       NA
## Subject19                        NA         NA      NA       NA
## Subject20                        NA         NA      NA       NA
## Subject22                        NA         NA      NA       NA
## Subject23                        NA         NA      NA       NA
## Subject26                        NA         NA      NA       NA
## Subject29                        NA         NA      NA       NA
## Subject30                        NA         NA      NA       NA
## Subject31                        NA         NA      NA       NA
## Subject32                        NA         NA      NA       NA
## Subject37                        NA         NA      NA       NA
## Subject38                        NA         NA      NA       NA
## Subject40                        NA         NA      NA       NA
## Subject42                        NA         NA      NA       NA
## Subject44.1                      NA         NA      NA       NA
## Subject44.2                      NA         NA      NA       NA
## Subject45                        NA         NA      NA       NA
## Subject46                        NA         NA      NA       NA
## Subject47                        NA         NA      NA       NA
## Subject48                        NA         NA      NA       NA
## Subject49                        NA         NA      NA       NA
## Subject52                        NA         NA      NA       NA
## Subject53                        NA         NA      NA       NA
## Subject55                        NA         NA      NA       NA
## Subject56                        NA         NA      NA       NA
## Subject57                        NA         NA      NA       NA
## Subject59                        NA         NA      NA       NA
## Subject60                        NA         NA      NA       NA
## Subject61                        NA         NA      NA       NA
## Subject63                        NA         NA      NA       NA
## Subject64                        NA         NA      NA       NA
## Subject66                        NA         NA      NA       NA
## Subject71                        NA         NA      NA       NA
## Subject74                        NA         NA      NA       NA
## Subject78                        NA         NA      NA       NA
## Subject80                        NA         NA      NA       NA
## Subject81                        NA         NA      NA       NA
## Subject82                        NA         NA      NA       NA
## Subject83                        NA         NA      NA       NA
## Subject93                        NA         NA      NA       NA
## Subject94                        NA         NA      NA       NA
## Subject98                        NA         NA      NA       NA
## GuitarPlay1                      NA         NA      NA       NA
## GuitarPlay2                      NA         NA      NA       NA
## GuitarPlay4                      NA         NA      NA       NA
## GuitarPlay5                      NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 2.47 on 1498 degrees of freedom
## Multiple R-squared:  0.2046, Adjusted R-squared:  0.1823
## F-statistic: 9.176 on 42 and 1498 DF,  p-value: < 2.2e-16
```

```
model_selfdeclare_binary_2 <- lm(Classical ~  OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare
                                  + ConsNotes + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory + Composing, data = music)
summary(model_selfdeclare_binary_2)
```

```
##
## Call:
## lm(formula = Classical ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + X1990s2000s.minus.1960s1970s + CollegeMusic +
##     NoClass + APTheory + Composing, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.1389 -1.8333 -0.0556  1.9011  7.3352
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   0.02068    3.21572   0.006 0.994870
## OMSI_log                      0.48996    0.24937   1.965 0.049621 *
## X16.minus.17                  0.20040    0.07106   2.820 0.004864 **
## ConsInstr0.67                10.46555    4.22217   2.479 0.013295 *
## ConsInstr1                    1.84043    2.77291   0.664 0.506973
## ConsInstr1.67                 4.47812    3.69550   1.212 0.225789
## ConsInstr2.33                 4.41007    2.94700   1.496 0.134744
## ConsInstr2.67                 4.84548    2.39717   2.021 0.043423 *
## ConsInstr3                    2.63884    3.40104   0.776 0.437934
## ConsInstr3.33                 2.61285    3.44596   0.758 0.448429
## ConsInstr3.67                 1.80764    3.04969   0.593 0.553452
## ConsInstr4                    6.96361    3.27557   2.126 0.033672 *
## ConsInstr4.33                 3.10255    3.23277   0.960 0.337351
## ConsInstr5                    3.50698    3.30510   1.061 0.288823
## Selfdeclare1                 -2.37284    0.75562  -3.140 0.001721 **
## ConsNotes1                    3.00997    0.73608   4.089 4.56e-05 ***
## ConsNotes3                    3.71626    0.60484   6.144 1.03e-09 ***
## ConsNotes4                   -0.40244    0.83635  -0.481 0.630451
## ConsNotes5                    2.63956    0.73540   3.589 0.000342 ***
## ClsListen1                   -0.87515    0.84379  -1.037 0.299825
## ClsListen3                    1.00411    1.10610   0.908 0.364134
## ClsListen4                   -6.24060    1.51378  -4.123 3.95e-05 ***
## ClsListen5                    1.39896    1.09006   1.283 0.199558
## KnowRob1                     -0.36469    0.44950  -0.811 0.417308
## KnowRob5                     -2.57240    0.54503  -4.720 2.58e-06 ***
## KnowAxis1                    -1.58911    1.47981  -1.074 0.283057
## KnowAxis5                     1.06231    0.67383   1.577 0.115116
## X1990s2000s2                 -0.21534    1.16545  -0.185 0.853431
## X1990s2000s3                 -4.53329    1.26463  -3.585 0.000348 ***
## X1990s2000s4                  3.98553    2.33341   1.708 0.087838 .
## X1990s2000s5                 -0.16385    1.08516  -0.151 0.880003
```

```
## X1990s2000s.minus.1960s1970s  0.18563    0.21011   0.884 0.377105
## CollegeMusic1                -2.44740    0.47113  -5.195 2.33e-07 ***
## NoClass                       1.14782    0.55529   2.067 0.038897 *
## APTheory1                    -0.38491    0.49099  -0.784 0.433193
## Composing1                   -1.52737    0.47546  -3.212 0.001344 **
## Composing2                    1.57326    0.69507   2.263 0.023749 *
## Composing3                   -4.08934    1.24190  -3.293 0.001015 **
## Composing4                    3.90822    0.80370   4.863 1.28e-06 ***
## Composing5                   -7.79138    3.41513  -2.281 0.022663 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.511 on 1501 degrees of freedom
## Multiple R-squared:  0.1766, Adjusted R-squared:  0.1552
## F-statistic: 8.257 on 39 and 1501 DF,  p-value: < 2.2e-16
```

```
stepAIC(model_selfdeclare_binary_2, k = 2)
```

```
## Start:  AIC=2876.76
## Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
##     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory +
##     Composing
##
##                                Df Sum of Sq     RSS    AIC
## - APTheory                      1      3.87  9466.3 2875.4
## - X1990s2000s.minus.1960s1970s  1      4.92  9467.3 2875.6
## <none>                                       9462.4 2876.8
## - KnowAxis                      2     28.82  9491.2 2877.4
## - OMSI_log                      1     24.34  9486.8 2878.7
## - NoClass                       1     26.94  9489.4 2879.1
## - X16.minus.17                  1     50.14  9512.6 2882.9
## - Selfdeclare                   1     62.17  9524.6 2884.8
## - CollegeMusic                  1    170.12  9632.5 2902.2
## - KnowRob                       2    193.83  9656.3 2904.0
## - X1990s2000s                   4    373.71  9836.1 2928.4
## - Composing                     5    403.18  9865.6 2931.1
## - ConsNotes                     4    410.83  9873.2 2934.2
## - ClsListen                     4    511.74  9974.2 2949.9
## - ConsInstr                    11    621.66 10084.1 2952.8
##
## Step:  AIC=2875.39
## Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
##     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + Composing
##
##                                Df Sum of Sq     RSS    AIC
## - X1990s2000s.minus.1960s1970s  1      2.28  9468.6 2873.8
## <none>                                       9466.3 2875.4
## - KnowAxis                      2     29.49  9495.8 2876.2
## - OMSI_log                      1     20.48  9486.8 2876.7
## - NoClass                       1     23.08  9489.4 2877.1
## - X16.minus.17                  1     51.28  9517.6 2881.7
## - Selfdeclare                   1     61.30  9527.6 2883.3
```

```
## - CollegeMusic                 1    181.29  9647.6 2902.6
## - KnowRob                      2    195.42  9661.7 2902.9
## - X1990s2000s                  4    370.87  9837.2 2926.6
## - Composing                    5    403.36  9869.7 2929.7
## - ConsNotes                    4    406.98  9873.3 2932.2
## - ClsListen                    4    512.30  9978.6 2948.6
## - ConsInstr                   11    617.89 10084.2 2950.8
##
## Step:  AIC=2873.76
## Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
##     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     CollegeMusic + NoClass + Composing
##
##               Df Sum of Sq     RSS    AIC
## <none>                      9468.6 2873.8
## - OMSI_log      1     21.60  9490.2 2875.3
## - NoClass       1     27.32  9495.9 2876.2
## - X16.minus.17  1     49.01  9517.6 2879.7
## - KnowAxis      2     87.41  9556.0 2883.9
## - Selfdeclare   1    100.53  9569.1 2888.0
## - CollegeMusic  1    223.84  9692.4 2907.8
## - KnowRob       2    254.07  9722.6 2910.6
## - Composing     5    411.92  9880.5 2929.4
## - ConsNotes     4    427.48  9896.1 2933.8
## - X1990s2000s   4    432.19  9900.8 2934.5
## - ClsListen     4    514.07  9982.6 2947.2
## - ConsInstr    11    615.62 10084.2 2948.8
```

```
##
## Call:
## lm(formula = Classical ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + CollegeMusic + NoClass + Composing, data = music)
##
## Coefficients:
##   (Intercept)       OMSI_log   X16.minus.17  ConsInstr0.67     ConsInstr1
##       2.76550        0.42648        0.19397        7.46173        0.05506
## ConsInstr1.67  ConsInstr2.33  ConsInstr2.67     ConsInstr3  ConsInstr3.33
##       2.25862        2.60895        3.35365        0.44555        0.10713
## ConsInstr3.67     ConsInstr4  ConsInstr4.33     ConsInstr5   Selfdeclare1
##      -0.01036        4.65677        1.12848        1.61187       -1.80642
##    ConsNotes1     ConsNotes3     ConsNotes4     ConsNotes5     ClsListen1
##       2.74875        3.47982       -0.51239        2.32312       -1.52112
##    ClsListen3     ClsListen4     ClsListen5       KnowRob1       KnowRob5
##       0.07366       -7.09357        0.45409       -0.12048       -2.83542
##     KnowAxis1      KnowAxis5   X1990s2000s2   X1990s2000s3   X1990s2000s4
##      -0.85083        1.52678       -0.02602       -4.05813        2.56614
##  X1990s2000s5  CollegeMusic1        NoClass     Composing1     Composing2
##       0.34994       -2.15516        0.77517       -1.47471        1.18957
##    Composing3     Composing4     Composing5
##      -3.46414        4.04893       -5.65137
```

```r
selfdeclare_after_stepAIC <- lmer(Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare + ConsNo
                                  + ClsListen + KnowRob + KnowAxis + X1990s2000s + CollegeMusic
                                  + NoClass + Composing + Instrument + Harmony*Voice +
                    (Instrument + Harmony | Subject), data = music,
                  REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
anova(model_c_7, selfdeclare_after_stepAIC)
```

```
## Data: music
## Models:
## model_c_7: Classical ~ Instrument + Harmony * Voice + (Instrument + Harmony |
## model_c_7:     Subject)
## selfdeclare_after_stepAIC: Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
## selfdeclare_after_stepAIC:     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
## selfdeclare_after_stepAIC:     CollegeMusic + NoClass + Composing + Instrument + Harmony *
## selfdeclare_after_stepAIC:     Voice + (Instrument + Harmony | Subject)
##                           Df    AIC    BIC  logLik deviance  Chisq Chi Df
## model_c_7                 36 6190.1 6382.3 -3059.0   6118.1
## selfdeclare_after_stepAIC 73 6158.3 6548.2 -3006.2   6012.3 105.73     37
##                           Pr(>Chisq)
## model_c_7
## selfdeclare_after_stepAIC   1.55e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

No difference from the results of #3. I am still going to favor with model_c_7.

Popular Analysis:

5 (a)

For (a) I am going to do the most basic of models:
Popular ~ Harmony + Instrument + Voice

```r
# 5 ANOVA models with all the possible interactions
pop_model_aov_1 <- aov(Popular ~ Harmony + Instrument + Voice, data = music)
pop_model_aov_2 <- aov(Popular ~ Harmony*Instrument + Voice, data = music)
pop_model_aov_3 <- aov(Popular ~ Harmony + Instrument*Voice, data = music)
pop_model_aov_4 <- aov(Popular ~ Harmony*Voice + Instrument, data = music)
pop_model_aov_5 <- aov(Popular ~ Harmony*Instrument*Voice, data = music)
summary(pop_model_aov_1)
```

```
##               Df Sum Sq Mean Sq F value Pr(>F)
## Harmony        3     27     9.0   1.736  0.158
## Instrument     2   2402  1200.8 232.912 <2e-16 ***
## Voice          2     16     8.0   1.550  0.213
## Residuals   1533   7903     5.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(pop_model_aov_2)
```

```
##                      Df Sum Sq Mean Sq F value Pr(>F)
## Harmony               3     27     9.0   1.733  0.158
## Instrument            2   2402  1200.8 232.409 <2e-16 ***
## Voice                 2     16     8.0   1.546  0.213
## Harmony:Instrument    6     14     2.3   0.447  0.847
## Residuals          1527   7890     5.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(pop_model_aov_3)
```

```
##                    Df Sum Sq Mean Sq F value Pr(>F)
## Harmony             3     27     9.0   1.736  0.158
## Instrument          2   2402  1200.8 232.855 <2e-16 ***
## Voice               2     16     8.0   1.549  0.213
## Instrument:Voice    4     19     4.7   0.905  0.460
## Residuals        1529   7885     5.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(pop_model_aov_4)
```

```
##                 Df Sum Sq Mean Sq F value Pr(>F)
## Harmony          3     27     9.0   1.740  0.157
## Voice            2     16     8.2   1.592  0.204
## Instrument       2   2401  1200.6 233.429 <2e-16 ***
## Harmony:Voice    6     50     8.3   1.609  0.141
## Residuals     1527   7854     5.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(pop_model_aov_5)
```

```
##                            Df Sum Sq Mean Sq F value Pr(>F)
## Harmony                     3     27     9.0   1.740  0.157
## Instrument                  2   2402  1200.8 233.340 <2e-16 ***
## Voice                       2     16     8.0   1.553  0.212
## Harmony:Instrument          6     14     2.3   0.449  0.846
## Harmony:Voice               6     50     8.3   1.605  0.142
## Instrument:Voice            4     19     4.7   0.908  0.458
## Harmony:Instrument:Voice   12     76     6.4   1.238  0.251
## Residuals                1505   7745     5.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# 5 simple linear models with all the possible interactions
pop_model_simp_1 <- lm(Popular ~ Harmony + Instrument + Voice, data = music)
pop_model_simp_2 <- lm(Popular ~ Harmony*Instrument + Voice, data = music)
```

```r
pop_model_simp_3 <- lm(Popular ~ Harmony + Instrument*Voice, data = music)
pop_model_simp_4 <- lm(Popular ~ Harmony*Voice + Instrument, data = music)
pop_model_simp_5 <- lm(Popular ~ Harmony*Instrument*Voice, data = music)
summary(pop_model_simp_1)
```

```
##
## Call:
## lm(formula = Popular ~ Harmony + Instrument + Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.0971 -1.7662  0.1574  1.4062 13.0888
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       6.84265    0.16345  41.863  < 2e-16 ***
## HarmonyI-V-IV     0.02258    0.16355   0.138    0.890
## HarmonyI-V-VI    -0.25261    0.16365  -1.544    0.123
## HarmonyIV-I-V    -0.24880    0.16344  -1.522    0.128
## Instrumentpiano  -1.14978    0.14184  -8.106 1.06e-15 ***
## Instrumentstring -3.02343    0.14143 -21.378  < 2e-16 ***
## Voicepar3rd       0.19579    0.14170   1.382    0.167
## Voicepar5th       0.23183    0.14170   1.636    0.102
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.271 on 1533 degrees of freedom
## Multiple R-squared:  0.2362, Adjusted R-squared:  0.2327
## F-statistic: 67.73 on 7 and 1533 DF,  p-value: < 2.2e-16
```

```r
summary(pop_model_simp_2)
```

```
##
## Call:
## lm(formula = Popular ~ Harmony * Instrument + Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.1124 -1.7236  0.1192  1.4356 13.0652
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.741289   0.216246  31.174  < 2e-16 ***
## HarmonyI-V-IV                  0.139535   0.283027   0.493    0.622
## HarmonyI-V-VI                 -0.017646   0.283580  -0.062    0.950
## HarmonyIV-I-V                 -0.193798   0.283027  -0.685    0.494
## Instrumentpiano               -1.133025   0.283580  -3.995 6.77e-05 ***
## Instrumentstring              -2.736434   0.283027  -9.668  < 2e-16 ***
## Voicepar3rd                    0.195747   0.141859   1.380    0.168
## Voicepar5th                    0.231549   0.141859   1.632    0.103
## HarmonyI-V-IV:Instrumentpiano  -0.008761   0.401438  -0.022    0.983
## HarmonyI-V-VI:Instrumentpiano  -0.221941   0.401828  -0.552    0.581
```

```
## HarmonyIV-I-V:Instrumentpiano   0.162269   0.401043   0.405    0.686
## HarmonyI-V-IV:Instrumentstring -0.341085   0.400260  -0.852    0.394
## HarmonyI-V-VI:Instrumentstring -0.482354   0.400651  -1.204    0.229
## HarmonyIV-I-V:Instrumentstring -0.325581   0.400260  -0.813    0.416
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.273 on 1527 degrees of freedom
## Multiple R-squared:  0.2376, Adjusted R-squared:  0.2311
## F-statistic:  36.6 on 13 and 1527 DF,  p-value: < 2.2e-16
```

```r
summary(pop_model_simp_3)
```

```
##
## Call:
## lm(formula = Popular ~ Harmony + Instrument * Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9435 -1.6910  0.1113  1.3769 13.0887
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.94348    0.20036  34.655  < 2e-16 ***
## HarmonyI-V-IV                  0.02253    0.16357   0.138    0.890
## HarmonyI-V-VI                 -0.25245    0.16368  -1.542    0.123
## HarmonyIV-I-V                 -0.24885    0.16346  -1.522    0.128
## Instrumentpiano               -1.27079    0.24595  -5.167 2.69e-07 ***
## Instrumentstring              -3.20460    0.24523 -13.068  < 2e-16 ***
## Voicepar3rd                    0.19366    0.24523   0.790    0.430
## Voicepar5th                   -0.06797    0.24523  -0.277    0.782
## Instrumentpiano:Voicepar3rd    0.02239    0.34757   0.064    0.949
## Instrumentstring:Voicepar3rd -0.01633    0.34656  -0.047    0.962
## Instrumentpiano:Voicepar5th    0.34004    0.34757   0.978    0.328
## Instrumentstring:Voicepar5th   0.55925    0.34656   1.614    0.107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.271 on 1529 degrees of freedom
## Multiple R-squared:  0.238,  Adjusted R-squared:  0.2325
## F-statistic: 43.42 on 11 and 1529 DF,  p-value: < 2.2e-16
```

```r
summary(pop_model_simp_4)
```

```
##
## Call:
## lm(formula = Popular ~ Harmony * Voice + Instrument, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.1275 -1.6902  0.0365  1.5054 12.9685
##
## Coefficients:
```

```
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                6.67027    0.21572  30.920  < 2e-16 ***
## HarmonyI-V-IV              0.04312    0.28294   0.152   0.8789
## HarmonyI-V-VI             -0.18354    0.28349  -0.647   0.5175
## HarmonyIV-I-V              0.34884    0.28238   1.235   0.2169
## Voicepar3rd                0.48062    0.28238   1.702   0.0890 .
## Voicepar5th                0.46500    0.28294   1.643   0.1005
## Instrumentpiano           -1.15036    0.14168  -8.120 9.54e-16 ***
## Instrumentstring          -3.02323    0.14126 -21.402  < 2e-16 ***
## HarmonyI-V-IV:Voicepar3rd -0.01212    0.39974  -0.030   0.9758
## HarmonyI-V-VI:Voicepar3rd -0.12114    0.40053  -0.302   0.7623
## HarmonyIV-I-V:Voicepar3rd -1.00508    0.39974  -2.514   0.0120 *
## HarmonyI-V-IV:Voicepar5th -0.05094    0.40052  -0.127   0.8988
## HarmonyI-V-VI:Voicepar5th -0.08766    0.40053  -0.219   0.8268
## HarmonyIV-I-V:Voicepar5th -0.79058    0.39974  -1.978   0.0481 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.268 on 1527 degrees of freedom
## Multiple R-squared:  0.241,  Adjusted R-squared:  0.2346
## F-statistic:  37.3 on 13 and 1527 DF,  p-value: < 2.2e-16
```

```
summary(pop_model_simp_5)
```

```
##
## Call:
## lm(formula = Popular ~ Harmony * Instrument * Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8140 -1.6429  0.0233  1.5476 12.8837
##
## Coefficients:
##                                             Estimate Std. Error t value
## (Intercept)                                 6.813953   0.345943  19.697
## HarmonyI-V-IV                               0.279070   0.489237   0.570
## HarmonyI-V-VI                              -0.456811   0.492141  -0.928
## HarmonyIV-I-V                               0.209302   0.489237   0.428
## Instrumentpiano                            -1.279070   0.489237  -2.614
## Instrumentstring                           -3.325581   0.489237  -6.797
## Voicepar3rd                                -0.116279   0.489237  -0.238
## Voicepar5th                                 0.325581   0.489237   0.665
## HarmonyI-V-IV:Instrumentpiano              -0.313953   0.693942  -0.452
## HarmonyI-V-VI:Instrumentpiano               0.374308   0.695992   0.538
## HarmonyIV-I-V:Instrumentpiano              -0.023256   0.691886  -0.034
## HarmonyI-V-IV:Instrumentstring             -0.395349   0.691886  -0.571
## HarmonyI-V-VI:Instrumentstring              0.445183   0.693942   0.642
## HarmonyIV-I-V:Instrumentstring              0.441860   0.691886   0.639
## HarmonyI-V-IV:Voicepar3rd                   0.186047   0.691886   0.269
## HarmonyI-V-VI:Voicepar3rd                   1.317276   0.693942   1.898
## HarmonyIV-I-V:Voicepar3rd                  -0.255814   0.691886  -0.370
## HarmonyI-V-IV:Voicepar5th                  -0.604651   0.691886  -0.874
## HarmonyI-V-VI:Voicepar5th                  -0.008306   0.693942  -0.012
## HarmonyIV-I-V:Voicepar5th                  -0.953488   0.691886  -1.378
```

```
## Instrumentpiano:Voicepar3rd                 0.441860   0.691886   0.639
## Instrumentstring:Voicepar3rd                1.348837   0.691886   1.950
## Instrumentpiano:Voicepar5th                -0.003322   0.693942  -0.005
## Instrumentstring:Voicepar5th                0.418605   0.691886   0.605
## HarmonyI-V-IV:Instrumentpiano:Voicepar3rd   0.104651   0.979929   0.107
## HarmonyI-V-VI:Instrumentpiano:Voicepar3rd  -1.642857   0.982833  -1.672
## HarmonyIV-I-V:Instrumentpiano:Voicepar3rd  -0.147841   0.979929  -0.151
## HarmonyI-V-IV:Instrumentstring:Voicepar3rd -0.697674   0.978474  -0.713
## HarmonyI-V-VI:Instrumentstring:Voicepar3rd -2.677741   0.979929  -2.733
## HarmonyIV-I-V:Instrumentstring:Voicepar3rd -2.093023   0.978474  -2.139
## HarmonyI-V-IV:Instrumentpiano:Voicepar5th   0.806202   0.982833   0.820
## HarmonyI-V-VI:Instrumentpiano:Voicepar5th  -0.138427   0.982833  -0.141
## HarmonyIV-I-V:Instrumentpiano:Voicepar5th   0.700997   0.979929   0.715
## HarmonyI-V-IV:Instrumentstring:Voicepar5th  0.860465   0.978474   0.879
## HarmonyI-V-VI:Instrumentstring:Voicepar5th -0.096346   0.979929  -0.098
## HarmonyIV-I-V:Instrumentstring:Voicepar5th -0.209302   0.978474  -0.214
##                                            Pr(>|t|)
## (Intercept)                                 < 2e-16 ***
## HarmonyI-V-IV                               0.56848
## HarmonyI-V-VI                               0.35345
## HarmonyIV-I-V                               0.66885
## Instrumentpiano                             0.00903 **
## Instrumentstring                           1.53e-11 ***
## Voicepar3rd                                 0.81217
## Voicepar5th                                 0.50584
## HarmonyI-V-IV:Instrumentpiano               0.65103
## HarmonyI-V-VI:Instrumentpiano               0.59079
## HarmonyIV-I-V:Instrumentpiano               0.97319
## HarmonyI-V-IV:Instrumentstring              0.56781
## HarmonyI-V-VI:Instrumentstring              0.52128
## HarmonyIV-I-V:Instrumentstring              0.52316
## HarmonyI-V-IV:Voicepar3rd                   0.78805
## HarmonyI-V-VI:Voicepar3rd                   0.05785 .
## HarmonyIV-I-V:Voicepar3rd                   0.71163
## HarmonyI-V-IV:Voicepar5th                   0.38230
## HarmonyI-V-VI:Voicepar5th                   0.99045
## HarmonyIV-I-V:Voicepar5th                   0.16838
## Instrumentpiano:Voicepar3rd                 0.52316
## Instrumentstring:Voicepar3rd                0.05142 .
## Instrumentpiano:Voicepar5th                 0.99618
## Instrumentstring:Voicepar5th                0.54526
## HarmonyI-V-IV:Instrumentpiano:Voicepar3rd   0.91497
## HarmonyI-V-VI:Instrumentpiano:Voicepar3rd   0.09482 .
## HarmonyIV-I-V:Instrumentpiano:Voicepar3rd   0.88010
## HarmonyI-V-IV:Instrumentstring:Voicepar3rd  0.47594
## HarmonyI-V-VI:Instrumentstring:Voicepar3rd  0.00636 **
## HarmonyIV-I-V:Instrumentstring:Voicepar3rd  0.03259 *
## HarmonyI-V-IV:Instrumentpiano:Voicepar5th   0.41218
## HarmonyI-V-VI:Instrumentpiano:Voicepar5th   0.88801
## HarmonyIV-I-V:Instrumentpiano:Voicepar5th   0.47450
## HarmonyI-V-IV:Instrumentstring:Voicepar5th  0.37933
## HarmonyI-V-VI:Instrumentstring:Voicepar5th  0.92169
## HarmonyIV-I-V:Instrumentstring:Voicepar5th  0.83065
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.268 on 1505 degrees of freedom
## Multiple R-squared:  0.2515, Adjusted R-squared:  0.2341
## F-statistic: 14.45 on 35 and 1505 DF,  p-value: < 2.2e-16
```

Popular appears to be more interesting with pop_model_simp_4 having significant interations between some Harmony and Voice and pop_model_simp_5 having significant interactions between some Instruments and Voice and some Harmony and Voice and all three Harmony, Instrument, and Voice.
Next I am going to do a random intercept model. I have chosen to go with the significant interaction in pop_model_simp_4 because I think the three way interaction in pop_model_simp_5 is too much.

```
pop_model_rand_int <- lmer(Popular ~ Instrument + Harmony*Voice + (1 | Subject),
                            data = music, REML = FALSE)
AIC(pop_model_simp_4)
```

```
## [1] 6912.784
```

```
AIC(pop_model_rand_int)
```

```
## [1] 6511.222
```

The random intercept significantly lowers the AIC. I will continue with pop_model_rand_int.

```
# anova() is from package "stats"
pop_model_c_1 <- lmer(Popular ~ Harmony*Voice + Instrument +
                    (Instrument | Subject), data = music, REML = FALSE)
pop_model_c_2 <- lmer(Popular ~ Harmony*Voice + Instrument +
                    (Harmony | Subject), data = music, REML = FALSE)
pop_model_c_3 <- lmer(Popular ~ Harmony*Voice + Instrument +
                    (Voice | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```
pop_model_c_4 <- lmer(Popular ~ Harmony*Voice + Instrument +
                    (Harmony + Instrument | Subject), data = music,
                  REML = FALSE)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0189444 (tol = 0.002, component 1)
```

```
pop_model_c_5 <- lmer(Popular ~ Harmony*Voice + Instrument +
                    (Voice + Instrument | Subject), data = music, REML = FALSE)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## unable to evaluate scaled gradient
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge: degenerate Hessian with 1 negative eigenvalues
```

```r
pop_model_c_6 <- lmer(Popular ~ Harmony + Instrument + Voice +
                     (Voice + Harmony + Instrument | Subject), data = music,
                   REML = FALSE)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0563408 (tol = 0.002, component 1)
```

```r
pop_model_c_7 <- lmer(Popular ~ Instrument + Harmony*Voice +
                     (Instrument + Harmony | Subject), data = music,
                   REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
pop_model_c_8 <- lmer(Popular ~ Harmony*Voice + Instrument +
                     (Voice + Harmony | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
pop_model_c_9 <- lmer(Popular ~ Harmony*Voice + Instrument +
                     (Voice + Instrument + Harmony | Subject), data = music,
                    REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
pop_model_c_10 <- lmer(Popular ~ Harmony*Voice + Instrument +
                     (Instrument + Voice | Subject), data = music, REML = FALSE)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## unable to evaluate scaled gradient

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge: degenerate Hessian with 1 negative eigenvalues
```

```r
pop_model_c_11 <- lmer(Popular ~ Harmony*Voice + Instrument +
                     (Harmony + Voice | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
pop_model_c_12 <- lmer(Popular ~ Harmony*Voice + Instrument +
                     (Harmony + Instrument + Voice | Subject), data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
anova(pop_model_rand_int, pop_model_c_1, pop_model_c_2, pop_model_c_3, pop_model_c_4,
      pop_model_c_5, pop_model_c_6, pop_model_c_7, pop_model_c_8, pop_model_c_9, pop_model_c_10,
      pop_model_c_11, pop_model_c_12)
```

```
## Data: music
## Models:
## pop_model_rand_int: Popular ~ Instrument + Harmony * Voice + (1 | Subject)
## pop_model_c_1: Popular ~ Harmony * Voice + Instrument + (Instrument | Subject)
## pop_model_c_3: Popular ~ Harmony * Voice + Instrument + (Voice | Subject)
## pop_model_c_2: Popular ~ Harmony * Voice + Instrument + (Harmony | Subject)
## pop_model_c_5: Popular ~ Harmony * Voice + Instrument + (Voice + Instrument |
## pop_model_c_5:     Subject)
## pop_model_c_10: Popular ~ Harmony * Voice + Instrument + (Instrument + Voice |
## pop_model_c_10:     Subject)
## pop_model_c_4: Popular ~ Harmony * Voice + Instrument + (Harmony + Instrument |
## pop_model_c_4:     Subject)
## pop_model_c_7: Popular ~ Instrument + Harmony * Voice + (Instrument + Harmony |
## pop_model_c_7:     Subject)
## pop_model_c_8: Popular ~ Harmony * Voice + Instrument + (Voice + Harmony | Subject)
## pop_model_c_11: Popular ~ Harmony * Voice + Instrument + (Harmony + Voice | Subject)
## pop_model_c_6: Popular ~ Harmony + Instrument + Voice + (Voice + Harmony + Instrument |
## pop_model_c_6:     Subject)
## pop_model_c_9: Popular ~ Harmony * Voice + Instrument + (Voice + Instrument +
## pop_model_c_9:     Harmony | Subject)
## pop_model_c_12: Popular ~ Harmony * Voice + Instrument + (Harmony + Instrument +
## pop_model_c_12:     Voice | Subject)
##                    Df    AIC    BIC  logLik deviance    Chisq Chi Df Pr(>Chisq)
## pop_model_rand_int 16 6511.2 6596.7 -3239.6   6479.2
## pop_model_c_1      21 6354.1 6466.2 -3156.0   6312.1 167.1711      5  < 2.2e-16
## pop_model_c_3      21 6520.3 6632.5 -3239.2   6478.3   0.0000      0   1.000000
## pop_model_c_2      25 6479.1 6612.6 -3214.6   6429.1  49.1989      4  5.307e-10
## pop_model_c_5      30 6368.5 6528.7 -3154.2   6308.5 120.6861      5  < 2.2e-16
## pop_model_c_10     30 6368.4 6528.6 -3154.2   6308.4   0.0306      0  < 2.2e-16
## pop_model_c_4      36 6300.0 6492.3 -3114.0   6228.0  80.4007      6  2.953e-15
## pop_model_c_7      36 6300.0 6492.3 -3114.0   6228.0   0.0003      0  < 2.2e-16
## pop_model_c_8      36 6496.7 6688.9 -3212.3   6424.7   0.0000      0   1.000000
## pop_model_c_11     36 6495.9 6688.2 -3212.0   6423.9   0.7569      0  < 2.2e-16
## pop_model_c_6      45 6325.7 6566.0 -3117.8   6235.7 188.2845      9  < 2.2e-16
## pop_model_c_9      51 6319.2 6591.6 -3108.6   6217.2  18.4199      6   0.005264
## pop_model_c_12     51 6318.9 6591.3 -3108.5   6216.9   0.2951      0  < 2.2e-16
##
## pop_model_rand_int
## pop_model_c_1      ***
## pop_model_c_3
## pop_model_c_2      ***
## pop_model_c_5      ***
## pop_model_c_10     ***
## pop_model_c_4      ***
## pop_model_c_7      ***
## pop_model_c_8
## pop_model_c_11     ***
## pop_model_c_6      ***
## pop_model_c_9      **
## pop_model_c_12     ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The ANOVA test tells me that pop_model_c_4 and pop_model_c_7 are the same AIC and BIC (wow!).

pop_model_c_4::
Popular ~ Harmony * Voice + Instrument + (Harmony + Instrument | Subject)
pop_model_c_7::
Popular ~ Instrument + Harmony * Voice + (Instrument + Harmony | Subject)

```r
summary(pop_model_c_7)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Popular ~ Instrument + Harmony * Voice + (Instrument + Harmony |
##     Subject)
##    Data: music
##
##      AIC      BIC   logLik deviance df.resid
##   6300.0   6492.3  -3114.0   6228.0     1505
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8069 -0.5800  0.0422  0.5779  4.9487
##
## Random effects:
##  Groups   Name             Variance Std.Dev. Corr
##  Subject  (Intercept)      1.2574   1.1213
##           Instrumentpiano  1.7848   1.3360   -0.07
##           Instrumentstring 2.6152   1.6172   -0.23  0.71
##           HarmonyI-V-IV    0.2091   0.4573    0.54 -0.21 -0.29
##           HarmonyI-V-VI    0.8234   0.9074   -0.01 -0.32 -0.23 -0.15
##           HarmonyIV-I-V    0.4766   0.6904   -0.37 -0.36 -0.45 -0.65 -0.03
##  Residual                  2.6577   1.6302
## Number of obs: 1541, groups:  Subject, 43
##
## Fixed effects:
##                           Estimate Std. Error t value
## (Intercept)                6.66942    0.23085  28.890
## Instrumentpiano           -1.14747    0.22787  -5.036
## Instrumentstring          -3.02358    0.26670 -11.337
## HarmonyI-V-IV              0.05470    0.21507   0.254
## HarmonyI-V-VI             -0.19017    0.24644  -0.772
## HarmonyIV-I-V              0.34884    0.22867   1.526
## Voicepar3rd                0.48062    0.20299   2.368
## Voicepar5th                0.47140    0.20343   2.317
## HarmonyI-V-IV:Voicepar3rd -0.02369    0.28739  -0.082
## HarmonyI-V-VI:Voicepar3rd -0.12254    0.28795  -0.426
## HarmonyIV-I-V:Voicepar3rd -1.00562    0.28739  -3.499
## HarmonyI-V-IV:Voicepar5th -0.05734    0.28794  -0.199
## HarmonyI-V-VI:Voicepar5th -0.08743    0.28810  -0.303
## HarmonyIV-I-V:Voicepar5th -0.79698    0.28738  -2.773

##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

They are pretty similar. I am just going to pick the one that was the most similar to Classical (pop_model_c_7).
Here is an interpretation of the model:: (tbh I copied it from 2 (c) ii.)
The baseline is where Harmony is I-IV-V, Instrument is guitar, and Voice is contrary.
Given all other variables are held constant, if the Instrument is piano, then the rating of Popular is expected to decrease by 1.14747, relative to the baseline.
Given all other variables are held constant, if the Instrument is a string, then the rating of Popular is expected to decrease by 3.02358, relative to the baseline.
Given all other variables are held constant, if the Harmony is I-V-IV, then the rating of Popular is expected to increase by 0.05470, relative to the baseline.
Given all other variables are held constant, if the Harmony is I-V-VI, then the rating of Popular is expected to decrease by 0.19017, relative to the baseline.
Given all other variables are held constant, if the Harmony is IV-I-V, then the rating of Popular is expected to increase by 0.34884, relative to the baseline.
Given all other variables are held constant, if the Voice is par3rd, then the rating of Popular is expected to increase by 0.48062, relative to the baseline.
Given all other variables are held constant, if the Voice is par5th, then the rating of Popular is expected to increase by 0.47140, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-IV and Voice is par3rd, then the rating of Popular is expected to decrease by 0.02369, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-VI and Voice is par3rd, then the rating of Popular is expected to decrease by 0.12254, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to IV-I-V and Voice is par3rd, then the rating of Popular is expected to decrease by 1.00562, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-IV and Voice is par5th, then the rating of Popular is expected to decrease by 0.05734, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to I-V-VI and Voice is par5th, then the rating of Popular is expected to decrease by 0.08743, relative to the baseline.
Given all other variables are held constant, when Harmony is fixed to IV-I-V and Voice is par5th, then the rating of Popular is expected to decrease by 0.79698, relative to the baseline.
The size of the variance of the random effects in respect to each other::
The highest variance is from String Instrument with a whooping 2.6152.The lowest is HarmonyI-V-IV at 0.2091.
The size of the variance of the random effects in respect to the estimated residual variance::
The estimated residual variance is 2.6577. Most of the random effects variances are lower than that (actually all expect one, String Instruments).

3 (b)

```
music_unedited <- read.csv("~/Desktop/Linear/ratings.csv")
music <- music_unedited
music <- music[, ! names(music_unedited) %in% c("X1stInstr", "X2ndInstr",
                                                "first12", "X")]
music <- music[complete.cases(music), ]
music$OMSI_log <- NA
music$OMSI_log<- log(music$OMSI)
music$Harmony <- as.factor(music$Harmony)
music$Instrument <- as.factor(music$Instrument)
music$Voice <- as.factor(music$Voice)
music$Selfdeclare <- as.factor(music$Selfdeclare)
music$ConsInstr <- as.factor(music$ConsInstr)
music$ConsNotes <- as.factor(music$ConsNotes)
```

```r
music$PachListen <- as.factor(music$PachListen)
music$ClsListen <- as.factor(music$ClsListen)
music$KnowRob <- as.factor(music$KnowRob)
music$KnowAxis <- as.factor(music$KnowAxis)
music$X1990s2000s <- as.factor(music$X1990s2000s)
music$CollegeMusic <- as.factor(music$CollegeMusic)
music$APTheory <- as.factor(music$APTheory)
music$Composing <- as.factor(music$Composing)
music$PianoPlay <- as.factor(music$PianoPlay)
music$GuitarPlay <- as.factor(music$GuitarPlay)
pop_model_simp_full <- lm(Popular ~  OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare
                      + ConsNotes + Instr.minus.Notes + PachListen + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory + Composing + PianoPlay + Subject +
                   GuitarPlay, data = music)
summary(pop_model_simp_full)
```

```
##
## Call:
## lm(formula = Popular ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + Instr.minus.Notes + PachListen +
##     ClsListen + KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory + Composing + PianoPlay +
##     Subject + GuitarPlay, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7222 -1.7500  0.1111  1.7778 10.7778
##
## Coefficients: (54 not defined because of singularities)
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   73.9657    26.2710   2.815 0.004934 **
## OMSI_log                      -2.4952     1.7716  -1.408 0.159207
## X16.minus.17                  -0.6236     0.3079  -2.025 0.043011 *
## ConsInstr0.67                -52.6029    15.3690  -3.423 0.000637 ***
## ConsInstr1                   -37.5360    14.0389  -2.674 0.007583 **
## ConsInstr1.67                -53.9935    20.4005  -2.647 0.008214 **
## ConsInstr2.33                -33.7489    13.1155  -2.573 0.010171 *
## ConsInstr2.67                -35.9749    14.5582  -2.471 0.013580 *
## ConsInstr3                   -66.9604    24.3318  -2.752 0.005995 **
## ConsInstr3.33                -71.8724    26.9669  -2.665 0.007777 **
## ConsInstr3.67                -56.4403    20.9775  -2.691 0.007213 **
## ConsInstr4                   -46.8558    16.4462  -2.849 0.004445 **
## ConsInstr4.33                -59.7228    22.6637  -2.635 0.008496 **
## ConsInstr5                   -58.0650    22.0192  -2.637 0.008450 **
## Selfdeclare2                  -5.2026     2.1447  -2.426 0.015391 *
## Selfdeclare3                  -3.4246     0.7803  -4.389 1.22e-05 ***
## Selfdeclare4                  12.7267     6.4075   1.986 0.047193 *
## Selfdeclare5                   5.8912     2.1011   2.804 0.005115 **
## Selfdeclare6                  -6.1371     2.6226  -2.340 0.019409 *
## ConsNotes1                     9.3538     3.4118   2.742 0.006186 **
## ConsNotes3                    15.4681     5.9533   2.598 0.009462 **
## ConsNotes4                    27.4095     9.6139   2.851 0.004418 **
```

```
## ConsNotes5                    22.4544     7.9319   2.831 0.004704 **
## Instr.minus.Notes                  NA         NA      NA       NA
## PachListen3                   17.2820     7.4226   2.328 0.020029 *
## PachListen4                        NA         NA      NA       NA
## PachListen5                    7.8740     5.4746   1.438 0.150567
## ClsListen1                    -3.5577     1.8500  -1.923 0.054651 .
## ClsListen3                    -9.3104     3.6610  -2.543 0.011085 *
## ClsListen4                    -6.1587     3.6642  -1.681 0.093015 .
## ClsListen5                   -20.9525     7.0693  -2.964 0.003086 **
## KnowRob1                      -0.7159     1.0173  -0.704 0.481706
## KnowRob5                     -22.5744     7.9199  -2.850 0.004427 **
## KnowAxis1                     -0.5104     2.1831  -0.234 0.815192
## KnowAxis5                     26.1287     9.3364   2.799 0.005199 **
## X1990s2000s2                 -22.9477     9.9684  -2.302 0.021470 *
## X1990s2000s3                  -8.0503     3.6430  -2.210 0.027269 *
## X1990s2000s4                 -51.8310    18.9206  -2.739 0.006228 **
## X1990s2000s5                 -15.6168     5.1548  -3.030 0.002491 **
## X1990s2000s.minus.1960s1970s  -1.8160     0.8189  -2.218 0.026719 *
## CollegeMusic1                  4.2012     2.4207   1.736 0.082850 .
## NoClass                       -2.2477     0.8984  -2.502 0.012463 *
## APTheory1                      0.8969     1.1808   0.760 0.447641
## Composing1                    -0.4979     1.5042  -0.331 0.740662
## Composing2                    -3.4045     2.5977  -1.311 0.190202
## Composing3                         NA         NA      NA       NA
## Composing4                         NA         NA      NA       NA
## Composing5                         NA         NA      NA       NA
## PianoPlay1                         NA         NA      NA       NA
## PianoPlay4                         NA         NA      NA       NA
## PianoPlay5                         NA         NA      NA       NA
## Subject17                          NA         NA      NA       NA
## Subject19                          NA         NA      NA       NA
## Subject20                          NA         NA      NA       NA
## Subject22                          NA         NA      NA       NA
## Subject23                          NA         NA      NA       NA
## Subject26                          NA         NA      NA       NA
## Subject29                          NA         NA      NA       NA
## Subject30                          NA         NA      NA       NA
## Subject31                          NA         NA      NA       NA
## Subject32                          NA         NA      NA       NA
## Subject37                          NA         NA      NA       NA
## Subject38                          NA         NA      NA       NA
## Subject40                          NA         NA      NA       NA
## Subject42                          NA         NA      NA       NA
## Subject44.1                        NA         NA      NA       NA
## Subject44.2                        NA         NA      NA       NA
## Subject45                          NA         NA      NA       NA
## Subject46                          NA         NA      NA       NA
## Subject47                          NA         NA      NA       NA
## Subject48                          NA         NA      NA       NA
## Subject49                          NA         NA      NA       NA
## Subject52                          NA         NA      NA       NA
## Subject53                          NA         NA      NA       NA
## Subject55                          NA         NA      NA       NA
## Subject56                          NA         NA      NA       NA
```

```
## Subject57                         NA         NA      NA       NA
## Subject59                         NA         NA      NA       NA
## Subject60                         NA         NA      NA       NA
## Subject61                         NA         NA      NA       NA
## Subject63                         NA         NA      NA       NA
## Subject64                         NA         NA      NA       NA
## Subject66                         NA         NA      NA       NA
## Subject71                         NA         NA      NA       NA
## Subject74                         NA         NA      NA       NA
## Subject78                         NA         NA      NA       NA
## Subject80                         NA         NA      NA       NA
## Subject81                         NA         NA      NA       NA
## Subject82                         NA         NA      NA       NA
## Subject83                         NA         NA      NA       NA
## Subject93                         NA         NA      NA       NA
## Subject94                         NA         NA      NA       NA
## Subject98                         NA         NA      NA       NA
## GuitarPlay1                       NA         NA      NA       NA
## GuitarPlay2                       NA         NA      NA       NA
## GuitarPlay4                       NA         NA      NA       NA
## GuitarPlay5                       NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.301 on 1498 degrees of freedom
## Multiple R-squared:  0.2334, Adjusted R-squared:  0.2119
## F-statistic: 10.86 on 42 and 1498 DF,  p-value: < 2.2e-16
```

```
pop_model_simp_full_2 <-lm(Popular ~  OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare
                      + ConsNotes + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory, data = music)
summary(pop_model_simp_full_2)
```

```
##
## Call:
## lm(formula = Popular ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + X1990s2000s.minus.1960s1970s + CollegeMusic +
##     NoClass + APTheory, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7222 -1.8705  0.0253  1.7394 10.7778
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   29.0539     5.3843   5.396 7.91e-08 ***
## OMSI_log                      -0.6318     0.3173  -1.991  0.04666 *
## X16.minus.17                  -0.3214     0.1071  -3.001  0.00274 **
## ConsInstr0.67                -21.2398     3.3747  -6.294 4.05e-10 ***
## ConsInstr1                   -10.6075     1.7121  -6.195 7.49e-10 ***
## ConsInstr1.67                -14.2183     2.4654  -5.767 9.77e-09 ***
## ConsInstr2.33                 -9.0745     1.5462  -5.869 5.39e-09 ***
```

```
## ConsInstr2.67                -11.1505     1.7069  -6.532 8.83e-11 ***
## ConsInstr3                   -22.5302     3.3290  -6.768 1.87e-11 ***
## ConsInstr3.33                -20.3459     3.8238  -5.321 1.19e-07 ***
## ConsInstr3.67                -19.7391     2.9322  -6.732 2.38e-11 ***
## ConsInstr4                   -14.5567     2.3788  -6.119 1.20e-09 ***
## ConsInstr4.33                -18.8875     2.9429  -6.418 1.85e-10 ***
## ConsInstr5                   -16.9859     2.5465  -6.670 3.58e-11 ***
## Selfdeclare2                  -2.0988     0.6929  -3.029  0.00250 **
## Selfdeclare3                  -4.9760     0.7159  -6.951 5.39e-12 ***
## Selfdeclare4                   6.4863     1.0853   5.977 2.84e-09 ***
## Selfdeclare5                  -1.5781     0.8807  -1.792  0.07334 .
## Selfdeclare6                  -2.7690     1.9163  -1.445  0.14868
## ConsNotes1                     4.7982     0.7736   6.202 7.18e-10 ***
## ConsNotes3                     5.8968     0.8243   7.153 1.32e-12 ***
## ConsNotes4                    11.4600     1.3791   8.310  < 2e-16 ***
## ConsNotes5                     9.2500     1.1155   8.292 2.44e-16 ***
## ClsListen1                    -0.1920     0.9884  -0.194  0.84598
## ClsListen3                    -2.1600     1.2934  -1.670  0.09513 .
## ClsListen4                     0.5654     0.9260   0.611  0.54156
## ClsListen5                    -9.9000     1.8437  -5.370 9.13e-08 ***
## KnowRob1                       0.7696     0.4386   1.755  0.07952 .
## KnowRob5                      -9.5563     1.4151  -6.753 2.06e-11 ***
## KnowAxis1                      1.3964     1.0212   1.367  0.17170
## KnowAxis5                      9.3807     1.3364   7.019 3.36e-12 ***
## X1990s2000s2                  -3.5400     1.0993  -3.220  0.00131 **
## X1990s2000s3                  -2.2643     0.8645  -2.619  0.00890 **
## X1990s2000s4                 -16.9516     2.6827  -6.319 3.47e-10 ***
## X1990s2000s5                  -6.2303     1.2499  -4.985 6.92e-07 ***
## X1990s2000s.minus.1960s1970s  -0.4625     0.1655  -2.794  0.00527 **
## CollegeMusic1                  1.0219     0.3124   3.271  0.00110 **
## NoClass                       -0.6507     0.2074  -3.137  0.00174 **
## APTheory1                     -3.0009     0.5611  -5.348 1.03e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.327 on 1502 degrees of freedom
## Multiple R-squared:  0.2142, Adjusted R-squared:  0.1943
## F-statistic: 10.77 on 38 and 1502 DF,  p-value: < 2.2e-16
```

```
stepAIC(pop_model_simp_full_2, k = 2)
```

```
## Start:  AIC=2641.23
## Popular ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
##     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory
##
##                                Df Sum of Sq    RSS    AIC
## <none>                                      8131.8 2641.2
## - OMSI_log                      1     21.46 8153.3 2643.3
## - X1990s2000s.minus.1960s1970s  1     42.27 8174.1 2647.2
## - X16.minus.17                  1     48.75 8180.6 2648.4
## - NoClass                       1     53.27 8185.1 2649.3
## - CollegeMusic                  1     57.92 8189.8 2650.2
## - APTheory                      1    154.85 8286.7 2668.3
```

```
## - KnowRob                    2    249.57 8381.4 2683.8
## - KnowAxis                   2    298.26 8430.1 2692.7
## - X1990s2000s                4    322.95 8454.8 2693.2
## - ConsInstr                 11    482.27 8614.1 2708.0
## - ConsNotes                  4    455.93 8587.8 2717.3
## - Selfdeclare                5    493.33 8625.2 2722.0
## - ClsListen                  4    619.39 8751.2 2746.3
```

```
##
## Call:
## lm(formula = Popular ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + X1990s2000s.minus.1960s1970s + CollegeMusic +
##     NoClass + APTheory, data = music)
##
## Coefficients:
##                 (Intercept)                      OMSI_log
##                     29.0539                       -0.6318
##                X16.minus.17                  ConsInstr0.67
##                     -0.3214                      -21.2398
##                  ConsInstr1                  ConsInstr1.67
##                    -10.6075                      -14.2183
##               ConsInstr2.33                  ConsInstr2.67
##                     -9.0745                      -11.1505
##                  ConsInstr3                  ConsInstr3.33
##                    -22.5302                      -20.3459
##               ConsInstr3.67                    ConsInstr4
##                    -19.7391                      -14.5567
##               ConsInstr4.33                    ConsInstr5
##                    -18.8875                      -16.9859
##                Selfdeclare2                  Selfdeclare3
##                     -2.0988                       -4.9760
##                Selfdeclare4                  Selfdeclare5
##                      6.4863                       -1.5781
##                Selfdeclare6                    ConsNotes1
##                     -2.7690                        4.7982
##                  ConsNotes3                    ConsNotes4
##                      5.8968                       11.4600
##                  ConsNotes5                    ClsListen1
##                      9.2500                       -0.1920
##                  ClsListen3                    ClsListen4
##                     -2.1600                        0.5654
##                  ClsListen5                      KnowRob1
##                     -9.9000                        0.7696
##                    KnowRob5                     KnowAxis1
##                     -9.5563                        1.3964
##                   KnowAxis5                  X1990s2000s2
##                      9.3807                       -3.5400
##                X1990s2000s3                  X1990s2000s4
##                     -2.2643                      -16.9516
##                X1990s2000s5  X1990s2000s.minus.1960s1970s
##                     -6.2303                       -0.4625
##               CollegeMusic1                       NoClass
##                      1.0219                       -0.6507
```

```
##                     APTheory1
##                       -3.0009
```

```
pop_model_after_stepAIC <- lmer(Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
    ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
    X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory +
                              Harmony*Voice + Instrument + (Instrument + Harmony | Subject),
                            data = music, REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```
anova(pop_model_after_stepAIC, pop_model_c_7)
```

```
## Data: music
## Models:
## pop_model_c_7: Popular ~ Instrument + Harmony * Voice + (Instrument + Harmony |
## pop_model_c_7:     Subject)
## pop_model_after_stepAIC: Classical ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
## pop_model_after_stepAIC:     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
## pop_model_after_stepAIC:     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory +
## pop_model_after_stepAIC:     Harmony * Voice + Instrument + (Instrument + Harmony | Subject)
##                         Df    AIC    BIC  logLik deviance  Chisq Chi Df
## pop_model_c_7           36 6300.0 6492.3 -3114.0   6228.0
## pop_model_after_stepAIC 74 6174.8 6570.0 -3013.4   6026.8 201.24     38
##                         Pr(>Chisq)
## pop_model_c_7
## pop_model_after_stepAIC  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Yet again, just like when I did it for classical, I will continue on with op_model_c_7. Please view #5(a) above for the interpretations for pop_model_c_7.
5 (c)
Make a new variable about whether or not a person self-identifies as a musician. 0 means you are a self-declared musician, 1 means you do not identify as a musician.

```
music$Selfdeclare<- as.numeric(music$Selfdeclare)
music$Selfdeclare <- ifelse(music$Selfdeclare > 2, 1, 0)
music$Selfdeclare <- as.factor(music$Selfdeclare)
summary(music$Selfdeclare)
```

```
##   0   1
## 827 714
```

Now I am going to redo what I did in #5 (b) and see if it makes a difference now that Selfdeclare is binary.

```
pop_model_selfdeclare_binary <- lm(Popular ~  OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare
                                   + ConsNotes + Instr.minus.Notes + PachListen + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory + Composing + PianoPlay + Subject +
                   GuitarPlay, data = music)
summary(pop_model_selfdeclare_binary)
```

```
##
## Call:
## lm(formula = Popular ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + Instr.minus.Notes + PachListen +
##     ClsListen + KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory + Composing + PianoPlay +
##     Subject + GuitarPlay, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7222 -1.7500  0.1111  1.7778 10.7778
##
## Coefficients: (50 not defined because of singularities)
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -6.67372    4.33476  -1.540  0.12387
## OMSI_log                     -0.11663    0.50947  -0.229  0.81896
## X16.minus.17                 -0.05914    0.07356  -0.804  0.42153
## ConsInstr0.67                 9.66908    5.72276   1.690  0.09132 .
## ConsInstr1                   11.11323    5.19960   2.137  0.03273 *
## ConsInstr1.67                16.28245    6.96410   2.338  0.01952 *
## ConsInstr2.33                11.17642    5.16384   2.164  0.03059 *
## ConsInstr2.67                 7.36211    4.53908   1.622  0.10503
## ConsInstr3                   15.71498    6.54736   2.400  0.01651 *
## ConsInstr3.33                23.30045    7.36004   3.166  0.00158 **
## ConsInstr3.67                14.10281    5.74329   2.456  0.01418 *
## ConsInstr4                   11.42724    5.44146   2.100  0.03589 *
## ConsInstr4.33                15.86825    6.38897   2.484  0.01311 *
## ConsInstr5                   15.53038    6.90097   2.250  0.02456 *
## Selfdeclare1                 -2.36934    1.12758  -2.101  0.03578 *
## ConsNotes1                    0.65432    0.85350   0.767  0.44342
## ConsNotes3                   -3.93350    1.31434  -2.993  0.00281 **
## ConsNotes4                   -2.26099    2.13668  -1.058  0.29014
## ConsNotes5                   -3.23623    1.81153  -1.786  0.07423 .
## Instr.minus.Notes                  NA         NA      NA       NA
## PachListen3                  -7.46040    2.92177  -2.553  0.01077 *
## PachListen4                        NA         NA      NA       NA
## PachListen5                  -5.80540    1.89490  -3.064  0.00223 **
## ClsListen1                    3.89495    0.90982   4.281 1.98e-05 ***
## ClsListen3                    4.81762    1.11768   4.310 1.74e-05 ***
## ClsListen4                   11.20073    2.58697   4.330 1.59e-05 ***
## ClsListen5                    3.53695    1.13136   3.126  0.00180 **
## KnowRob1                      0.12106    0.47453   0.255  0.79867
## KnowRob5                      7.71671    1.53645   5.022 5.71e-07 ***
## KnowAxis1                     1.73343    1.91989   0.903  0.36673
## KnowAxis5                    -6.40372    1.95467  -3.276  0.00108 **
## X1990s2000s2                  6.14764    2.86565   2.145  0.03209 *
## X1990s2000s3                  3.42259    1.65009   2.074  0.03823 *
## X1990s2000s4                 12.30371    4.75524   2.587  0.00976 **
## X1990s2000s5                  2.46409    1.44584   1.704  0.08854 .
## X1990s2000s.minus.1960s1970s  0.53285    0.29011   1.837  0.06645 .
## CollegeMusic1                -0.58053    0.99220  -0.585  0.55857
## NoClass                       1.59964    0.85727   1.866  0.06224 .
## APTheory1                    -0.62335    0.64339  -0.969  0.33277
## Composing1                    0.13228    0.79790   0.166  0.86835
```

```
## Composing2                     1.80738    0.95960   1.883  0.05983 .
## Composing3                    -0.07222    1.67328  -0.043  0.96558
## Composing4                    -2.88556    0.97851  -2.949  0.00324 **
## Composing5                    -6.57964    4.92874  -1.335  0.18209
## PianoPlay1                    -3.09099    0.59752  -5.173 2.61e-07 ***
## PianoPlay4                          NA         NA      NA       NA
## PianoPlay5                          NA         NA      NA       NA
## Subject17                           NA         NA      NA       NA
## Subject19                           NA         NA      NA       NA
## Subject20                           NA         NA      NA       NA
## Subject22                           NA         NA      NA       NA
## Subject23                           NA         NA      NA       NA
## Subject26                           NA         NA      NA       NA
## Subject29                           NA         NA      NA       NA
## Subject30                           NA         NA      NA       NA
## Subject31                           NA         NA      NA       NA
## Subject32                           NA         NA      NA       NA
## Subject37                           NA         NA      NA       NA
## Subject38                           NA         NA      NA       NA
## Subject40                           NA         NA      NA       NA
## Subject42                           NA         NA      NA       NA
## Subject44.1                         NA         NA      NA       NA
## Subject44.2                         NA         NA      NA       NA
## Subject45                           NA         NA      NA       NA
## Subject46                           NA         NA      NA       NA
## Subject47                           NA         NA      NA       NA
## Subject48                           NA         NA      NA       NA
## Subject49                           NA         NA      NA       NA
## Subject52                           NA         NA      NA       NA
## Subject53                           NA         NA      NA       NA
## Subject55                           NA         NA      NA       NA
## Subject56                           NA         NA      NA       NA
## Subject57                           NA         NA      NA       NA
## Subject59                           NA         NA      NA       NA
## Subject60                           NA         NA      NA       NA
## Subject61                           NA         NA      NA       NA
## Subject63                           NA         NA      NA       NA
## Subject64                           NA         NA      NA       NA
## Subject66                           NA         NA      NA       NA
## Subject71                           NA         NA      NA       NA
## Subject74                           NA         NA      NA       NA
## Subject78                           NA         NA      NA       NA
## Subject80                           NA         NA      NA       NA
## Subject81                           NA         NA      NA       NA
## Subject82                           NA         NA      NA       NA
## Subject83                           NA         NA      NA       NA
## Subject93                           NA         NA      NA       NA
## Subject94                           NA         NA      NA       NA
## Subject98                           NA         NA      NA       NA
## GuitarPlay1                         NA         NA      NA       NA
## GuitarPlay2                         NA         NA      NA       NA
## GuitarPlay4                         NA         NA      NA       NA
## GuitarPlay5                         NA         NA      NA       NA
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.301 on 1498 degrees of freedom
## Multiple R-squared:  0.2334, Adjusted R-squared:  0.2119
## F-statistic: 10.86 on 42 and 1498 DF,  p-value: < 2.2e-16
```

```
pop_model_selfdeclare_binary_2 <- lm(Popular ~  OMSI_log + X16.minus.17 + ConsInstr +
                                        Selfdeclare + ConsNotes + ClsListen +
                        KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
                   CollegeMusic + NoClass + APTheory + Composing, data = music)
summary(pop_model_selfdeclare_binary_2)
```

```
##
## Call:
## lm(formula = Popular ~ OMSI_log + X16.minus.17 + ConsInstr +
##     Selfdeclare + ConsNotes + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + X1990s2000s.minus.1960s1970s + CollegeMusic +
##     NoClass + APTheory + Composing, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7222 -1.7276  0.1065  1.7473 10.7778
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   4.94188    2.98381   1.656 0.097885 .
## OMSI_log                      0.41017    0.23138   1.773 0.076485 .
## X16.minus.17                  0.09037    0.06594   1.370 0.170737
## ConsInstr0.67               -15.53143    3.91768  -3.964 7.70e-05 ***
## ConsInstr1                  -11.01081    2.57294  -4.279 1.99e-05 ***
## ConsInstr1.67               -11.33197    3.42900  -3.305 0.000973 ***
## ConsInstr2.33               -10.36470    2.73447  -3.790 0.000156 ***
## ConsInstr2.67                -9.48274    2.22430  -4.263 2.14e-05 ***
## ConsInstr3                   -9.93156    3.15577  -3.147 0.001681 **
## ConsInstr3.33                -3.37478    3.19745  -1.055 0.291386
## ConsInstr3.67                -9.01910    2.82976  -3.187 0.001466 **
## ConsInstr4                  -10.59799    3.03935  -3.487 0.000503 ***
## ConsInstr4.33                -8.92405    2.99964  -2.975 0.002976 **
## ConsInstr5                  -10.93124    3.06675  -3.564 0.000376 ***
## Selfdeclare1                  1.12898    0.70113   1.610 0.107556
## ConsNotes1                   -0.88104    0.68299  -1.290 0.197260
## ConsNotes3                   -2.03064    0.56123  -3.618 0.000306 ***
## ConsNotes4                    3.27496    0.77603   4.220 2.59e-05 ***
## ConsNotes5                   -0.51382    0.68237  -0.753 0.451568
## ClsListen1                    3.49322    0.78294   4.462 8.74e-06 ***
## ClsListen3                    2.53971    1.02633   2.475 0.013450 *
## ClsListen4                    0.61612    1.40461   0.439 0.660986
## ClsListen5                    0.87964    1.01145   0.870 0.384615
## KnowRob1                     -0.65246    0.41709  -1.564 0.117956
## KnowRob5                      2.17714    0.50572   4.305 1.78e-05 ***
## KnowAxis1                     8.17878    1.37309   5.956 3.20e-09 ***
## KnowAxis5                     0.10866    0.62524   0.174 0.862051
## X1990s2000s2                  6.40235    1.08140   5.920 3.97e-09 ***
## X1990s2000s3                  8.97145    1.17343   7.645 3.69e-14 ***
```

```
## X1990s2000s4                   -4.34554    2.16514  -2.007 0.044924 *
## X1990s2000s5                    6.61333    1.00690   6.568 7.01e-11 ***
## X1990s2000s.minus.1960s1970s   -0.46538    0.19496  -2.387 0.017105 *
## CollegeMusic1                   2.58788    0.43716   5.920 3.99e-09 ***
## NoClass                        -2.67333    0.51524  -5.188 2.41e-07 ***
## APTheory1                      -1.04295    0.45558  -2.289 0.022202 *
## Composing1                      2.62604    0.44117   5.952 3.28e-09 ***
## Composing2                     -0.88874    0.64494  -1.378 0.168401
## Composing3                      7.17648    1.15234   6.228 6.13e-10 ***
## Composing4                     -0.61779    0.74574  -0.828 0.407564
## Composing5                     16.93106    3.16885   5.343 1.06e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.33 on 1501 degrees of freedom
## Multiple R-squared:  0.2127, Adjusted R-squared:  0.1922
## F-statistic:  10.4 on 39 and 1501 DF,  p-value: < 2.2e-16
```

```
stepAIC(pop_model_selfdeclare_binary_2, k = 2)
```

```
## Start:  AIC=2646.07
## Popular ~ OMSI_log + X16.minus.17 + ConsInstr + Selfdeclare +
##     ConsNotes + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory +
##     Composing
##
##                                Df Sum of Sq    RSS    AIC
## - X16.minus.17                  1     10.19 8157.1 2646.0
## <none>                                      8146.9 2646.1
## - Selfdeclare                   1     14.07 8160.9 2646.7
## - OMSI_log                      1     17.06 8163.9 2647.3
## - APTheory                      1     28.44 8175.3 2649.4
## - X1990s2000s.minus.1960s1970s  1     30.93 8177.8 2649.9
## - KnowRob                       2    100.71 8247.6 2661.0
## - NoClass                       1    146.11 8293.0 2671.5
## - KnowAxis                      2    199.44 8346.3 2679.3
## - CollegeMusic                  1    190.20 8337.1 2679.6
## - X1990s2000s                   4    370.62 8517.5 2706.6
## - ConsNotes                     4    379.07 8525.9 2708.2
## - Composing                     5    400.08 8546.9 2709.9
## - ClsListen                     4    417.29 8564.2 2715.1
## - ConsInstr                    11    657.11 8804.0 2743.6
##
## Step:  AIC=2646
## Popular ~ OMSI_log + ConsInstr + Selfdeclare + ConsNotes + ClsListen +
##     KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory + Composing
##
##                                Df Sum of Sq    RSS    AIC
## <none>                                      8157.1 2646.0
## - OMSI_log                      1     12.76 8169.8 2646.4
## - APTheory                      1     29.80 8186.9 2649.6
## - Selfdeclare                   1     36.25 8193.3 2650.8
## - X1990s2000s.minus.1960s1970s  1     38.01 8195.1 2651.2
```

```
## - KnowRob                     2     94.60 8251.7 2659.8
## - NoClass                     1    143.39 8300.4 2670.9
## - KnowAxis                    2    189.28 8346.3 2677.3
## - CollegeMusic                1    196.80 8353.9 2680.7
## - X1990s2000s                 4    364.11 8521.2 2705.3
## - ClsListen                   4    444.87 8601.9 2719.8
## - Composing                   5    463.93 8621.0 2721.2
## - ConsNotes                   4    467.47 8624.5 2723.9
## - ConsInstr                  11    696.49 8853.5 2750.3

##
## Call:
## lm(formula = Popular ~ OMSI_log + ConsInstr + Selfdeclare + ConsNotes +
##     ClsListen + KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
##     CollegeMusic + NoClass + APTheory + Composing, data = music)
##
## Coefficients:
##                  (Intercept)                      OMSI_log
##                       5.8899                        0.3479
##                ConsInstr0.67                    ConsInstr1
##                     -16.0675                      -10.8069
##                ConsInstr1.67                 ConsInstr2.33
##                     -11.2105                      -10.5537
##                ConsInstr2.67                    ConsInstr3
##                      -9.8423                      -10.3708
##                ConsInstr3.33                 ConsInstr3.67
##                      -3.5118                       -9.4726
##                   ConsInstr4                 ConsInstr4.33
##                     -10.4474                       -9.3629
##                   ConsInstr5                  Selfdeclare1
##                     -11.3174                        1.5899
##                   ConsNotes1                    ConsNotes3
##                      -1.3096                       -2.2871
##                   ConsNotes4                    ConsNotes5
##                       3.1489                       -0.7205
##                   ClsListen1                    ClsListen3
##                       3.3691                        2.2171
##                   ClsListen4                    ClsListen5
##                       1.0086                        0.6455
##                     KnowRob1                      KnowRob5
##                      -0.6234                        2.0952
##                    KnowAxis1                     KnowAxis5
##                       7.8211                        0.1960
##                 X1990s2000s2                  X1990s2000s3
##                       6.4210                        8.9649
##                 X1990s2000s4                  X1990s2000s5
##                      -4.7282                        6.6077
## X1990s2000s.minus.1960s1970s                 CollegeMusic1
##                      -0.5090                        2.6268
##                      NoClass                     APTheory1
##                      -2.6464                       -1.0667
##                   Composing1                    Composing2
##                       2.7640                       -1.0658
##                   Composing3                    Composing4
```

```
##                        7.3088                         -0.8180
##                    Composing5
##                       16.9897
```

```r
pop_selfdeclare_after_stepAIC <- lmer(Popular ~ OMSI_log + ConsInstr + Selfdeclare + ConsNotes +
                                    ClsListen + KnowRob + KnowAxis + X1990s2000s +
                                    X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass +
                                    APTheory + Composing + Instrument + Harmony*Voice +
                                    (Instrument + Harmony | Subject), data = music,
                                  REML = FALSE)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
anova(pop_model_c_7, pop_selfdeclare_after_stepAIC)
```

```
## Data: music
## Models:
## pop_model_c_7: Popular ~ Instrument + Harmony * Voice + (Instrument + Harmony |
## pop_model_c_7:     Subject)
## pop_selfdeclare_after_stepAIC: Popular ~ OMSI_log + ConsInstr + Selfdeclare + ConsNotes + ClsListen +
## pop_selfdeclare_after_stepAIC:     KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s +
## pop_selfdeclare_after_stepAIC:     CollegeMusic + NoClass + APTheory + Composing + Instrument +
## pop_selfdeclare_after_stepAIC:     Harmony * Voice + (Instrument + Harmony | Subject)
##                               Df    AIC    BIC logLik deviance  Chisq Chi Df
## pop_model_c_7                 36 6300.0 6492.3  -3114   6228.0
## pop_selfdeclare_after_stepAIC 74 6275.9 6671.1  -3064   6127.9 100.08     38
##                               Pr(>Chisq)
## pop_model_c_7
## pop_selfdeclare_after_stepAIC   1.72e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Go with model pop_model_c_7 because the BIC is 178.8 units smaller and only 24.1 units larger AIC.