How different factors influence classical and popular ratings

Yuanyuan Wang*
yuanyua4@andrew.cmu.edu

16 December 2019

Abstract

This study aims to learn how different factors, especially instruments, harmonic motion, and voice leading, influence listeners' identification of music as "classical" or "popular". We examined the data from Ivan Jimenez et al. (2013)[1], using different statistical analysis approaches and mixed-effects models. We find that the Instrument exerts the strongest influence among the three design factors. We also find that musicians differ from non-musician on Classical rating; Variables X16.minus.17 and PianoPlay influence Classical rating but do not affect Popular rating at all. The data is biased towards listeners who have limited musical knowledge. More survey needs to be done to collect data from listeners who have sophisticated musical knowledge.

1 Introduction

People can not live without music. Classical music and popular music are two major genres of music that have been in existence for a long time and loved by different groups of people.Classical music and popular music are different by definition. According to the Cambridge Dictionary, classical music is defined as a form of music developed from a European tradition mainly in the 18th and 19th centuries; While popular music is defined as the kind of music with words and strong rhythm that many young people enjoy listening to and dancing to. Researchers once stated that "popular music is music created by, and especially for, the enjoyment and enrichment of everyday people in their everyday lives" (Bowman, 2004, p. 36).

It is well-known that classical music differs from popular music in many areas, including chord progressions, melody, instruments, and voice leading, etc. And the appreciation of classical music requires skills and educations.

So how do listeners identify whether a piece of given music is classical or popular music? What factors have the strongest influence? Are there any differences in the way that musicians and non-musicians identify classical music?

All those questions are interesting and important. A better understanding of those questions could help us understand the listener's preconceptions about classical music and popular music and what specific knowledge needs to be promoted to help people better appreciate Classical music.

*Department of Statistics & Data Science, Carnegie Mellon University

[1]Ivan Jimenez and Vincent Rossi. *The Influence of Timbre, Harmony, and Voice Leading on Listeners' Distinction between Popular and Classical Music* Pittsburgh, PA: University of Pittsburgh.

However, very few researches have done in this area. This study aims to learn how different factors, especially instruments, harmonic motion, and voice leading, influence listeners' identification of music as "classical" or "popular".

In this paper, we will address the following questions:

- What experimental factor, or combination of factors, has the strongest influence on Classical ratings?

  Whether Instrument exerts the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading), as the researchers suspect?

  Among the different levels of Harmonic Motion, whether the Harmony motion I-V-VI has the strongest association with classical ratings?

  One particular harmonic progression, I-V-vi, might be frequently rated as classical, because it is the beginning progression for Pachelbel's Canon in D, which many people have heard. On the other hand, it is also a very common chord progression in popular music of the past 20 years. Whether the respondent is familiar with the Pachelbel rants/comedy bits matters or not for Classical ratings?

  Among the levels of Voice Leading, whether contrary motion has the strongest association with classical ratings?

- Are there any differences in the way that musicians and non-musicians identify classical music?

- Are there any differences in the factors that drive classical ratings and that drive popular ratings?

2 Methods

The data set for this study comes from Ivan Jimenez et al. (2013). It contains surveyed data from 70 listeners, who were recruited from the population of undergraduates at the University of Pittsburgh, on the 36 musical stimuli in the year of 2012. Listeners were asked to indicate the extent to which a series of three-chord successions were popular or classical music sounding. The 36 stimuli were chosen by completely crossing these factors:

- Instrument: String Quartet, Piano, Electric Guitar
- Harmonic Motion: I-V-vi, I-VI-V, I-V-IV, IV-I-V
- Voice Leading: Contrary Motion, Parallel 3rds, Parallel 5ths

The reader should refer to Ivan Jimenez et al. (2013) for the design of the experiment, eligibility, inclusion/exclusion criteria, and so forth, for the data set.

In all, a total of 27 variables and 2520 observations are represented in the data available to us. The definitions and the measurement are given in Figure 1.

A brief description of all variables in the data set follows:

Classical	How classical does the stimulus sound?
Popular	How popular does the stimulus sound?
Subject	Unique subject ID
Harmony	Harmonic Motion (4 levels)
Instrument	Instrument (3 levels)
Voice	Voice Leading (3 levels)
Selfdeclare	Are you a musician? (1-6, 1=not at all)
OMSI	Score on a test of musical knowledge
X16.minus.17	Auxiliary measure of listener's ability to distinguish classical vs popular music
ConsInstr	How much did you concentrate on the instrument while listening (0-5, 0=not at all)
ConsNotes	How much did you concentrate on the notes while listening? (0-5, 0=not at all)
Instr.minus.Notes	Difference between prev. two variables
PachListen	How familiar are you with Pachelbel's Canon in D (0-5, 0=not at all)
ClsListen	How much do you listen to classical music? (0-5, 0=not at all)
KnowRob	Have you heard Rob Paravonian's Pachelbel Rant (0-5, 0=not at all)
KnowAxis	Have you heard Axis of Evil's Comedy bit on the 4 Pachelbel chords in popular music? (0-5, 0=not at all)
X1990s2000s	How much do you listen to pop and rock from the 90's and 2000's? (0-5, 0=not at all)
X1990s2000s.minus.1960s1970s	Difference between prev variable and a similar variable referring to 60's and 70's pop and rock.
CollegeMusic	Have you taken music classes in college (0=no, 1=yes)
NoClass	How many music classes have you taken?
APTheory	Did you take AP Music Theory class in High School (0=no, 1=yes)
Composing	Have you done any music composing (0-5, 0=not at all)
PianoPlay	Do you play piano (0-5, 0=not at all)
GuitarPlay	Do you play guitar (0-5, 0=not at all)
X1stInstr	How proficient are you at your first musical instrument (0-5, 0=not at all)
X2ndInstr	Same, for second musical instrument
first12	In the experiment, which instrument was presented to the subject in the first 12 stimuli? (IGNORE FOR THIS ASSIGNMENT.)

Figure 1: Variable definitions for the surveyed data from Ivan Jimenez et al. (2013).

To deal with the missing value of the data, we removed six columns: X1stInstr, X2ndInstr, ConsNotes Instr.minus.Notes, NoClass, X1, and first12. After we delete these six columns, we discarded all the observations with NA values. We then replaced the unnormal value of 19 of column Popular with 10. After this, we have 1865 observations with 52 subjects left. Please refer to the Technical Appendix[2] page 1 to page 2 for details of data cleaning.

For our analysis, we used histogram plots to examine the distribution of the variables and performed the logarithm transformation to deal with the skewness of the data and dichotomized several variables. We then relied upon box plots to evaluate the relationships among variables.

We also relied on conventional linear models and mixed-effects models for modeling. We used step AIC, and step BIC for variable selection for the fixed effects part. We also relied on the ANOVA test and the log-likelihood ratio test for model comparisons. We performed the random effect test using function exactRLRT and later used the ffRanefLMER.fnc method to forward select the random effects. After random effects were fixed, we re-evaluated the fixed effect part by looking at the t-value of the coefficients. We used diagnostic plots, binned plots and the standardized conditional residuals plots for model validation.

To compare the difference between musicians and non-musicians, We dichotomized "Selfdeclare" at different values and created the new dummy variable Musicians.

To examined the influence of different values of categorical variables, we reorganized the model by putting that dummy variable first and using "-1" to get rid of the baseline level of the fixed effects and random effects.

We did all the analyses using the statistical computing tool R language and environment (R Core Team, 2017).

3 Results

3.1 Evaluation of the influences of experimental factors on Classical ratings

3.1.1 Model creation and model selection

After cleaning the data using the approaches we mentioned in the Method section, we transformed OMSI to log(OMSI) to deal with the skewness and dichotomized variables PachListen, KnowRob, KnowAxis, PianoPlay, and GuitarPlay. Please refer to Technical Appendix page 2 to page 3 for more technical details. Then we explored the relationship between variable Classical and other variables usingbox plots. Please refer to Technical Appendix page 4 to page 7 for details and conclusions of the EDA. Here we only highlighted that we find that variables APTheory, ClsListen, CollegeMusic, GuitarPlay, and KnowAxis do not seem to matter for Classical ratings, as shown in Figure 2.

[2] *Technical Appendix* is attached after the Reference section of this paper.

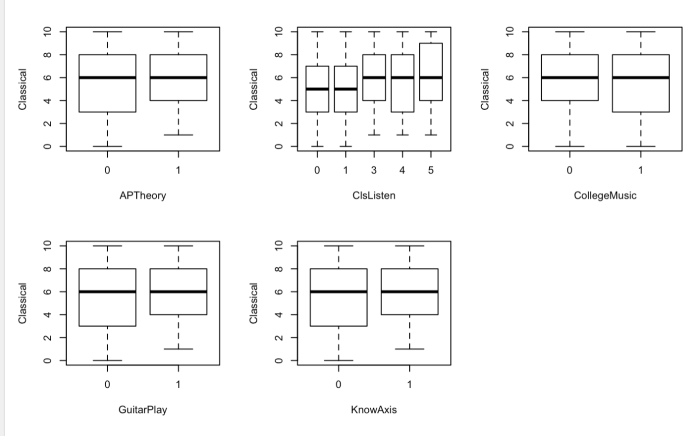


Figure 2: Box plots of variables that are not related to Classical ratings

After this, we examined the influence of the three main experimental factors(Instrument, Harmony, Voice) on Classical ratings. We first produced a base model by considering all the interactions among the three main experimental factors. Using stepAIC and ANOVA test for model comparison, the conventional linear model selected is lm2 and the variables selectedare Instrument, Harmony, Voice and the interactions between Harmony and Voice. Please see Technical Appendix page 9 to page 13 for more technical details about model lm2.

To examine whether variables KnowRob(Have you heard Rob Paravonian's Pachelbel Rant) and KnowAxis(Have you heard Axis of Evil's Comedy bit on the 4 Pachelbel chords in popular music) influence Classical ratings or not, we added the two variables and their interactions with Harmony to the fixed effects part. Using stepBIC and ANOVA test, the fixed effects part selected are Instrument, Harmony, Voice, the interactions between Harmony and Voice, and KnowAxis.

To account for "personal biases" for ratings, we decided to add the random intercept for each participant(1—Subject). Similarly, to account for the variations of "personal biases" with the types of instrument, type of harmony, and type of voice leading, we added all the three possible random slopes (Instrument+Harmony+Voice—Subject). Using function ffRanefLMER.fnc for forward selection of the random effects and re-examining the fixed effects part after the random effects are added, the final mixed effect model we get has the formula: lmer(Classical Instrument + Harmony + Voice +Harmony:Voice+ (1 — Subject) + (Instrument — Subject) + (Harmony — Subject) -1). So KnowRob and KnowAxis are not kept in the lmer model, which indicates that it doesn't matter whether the listener is familiar with either one of the Pachelbel rants or comedy bit. Please refer to Technical Appendix page 22 to page 27 for more technical details.

To determine which individual covariates should be added to the model as fixed effects, we first referred back to the findings of the EDA. Recalled that variables APTheory, ClsListen, CollegeMusic, and GuitarPlay don't seem to matter for Classical ratings.So we exclude these variables and added all the other variables as possible explanatory variables to build a full model. After using stepBIC for variable selection for the fixed effects part, using function ffRanefLMER.fnc for forward selection for the random effects part, and re-examining the t-value of the coefficients of the fixed effects part after random effects are added, the final mixed model we get is model lmer6.2. The following is the abbreviated form of model lmer6.2:

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Classical ~ Instrument + Harmony + Voice + X16.minus.17 + PianoPlay +
    (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) -      1
   Data: data
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   logLik deviance df.resid
  7478.9   7633.8  -3711.5   7422.9     1837

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.6926 -0.5842  0.0166  0.5449  3.6822

Random effects:
 Groups    Name             Variance  Std.Dev.  Corr
 Subject   (Intercept)      5.983e-08 0.0002446
 Subject.1 (Intercept)      7.395e-01 0.8599395
           Instrumentpiano  1.667e+00 1.2910689 -0.54
           Instrumentstring 3.549e+00 1.8838126 -1.00  0.60
 Subject.2 (Intercept)      1.053e+00 1.0263711
           HarmonyI-V-IV    5.753e-02 0.2398589  0.87
           HarmonyI-V-VI    1.869e+00 1.3670010 -0.36  0.05
           HarmonyIV-I-V    6.998e-02 0.2645403  0.11 -0.12  0.16
 Residual                   2.467e+00 1.5706033
Number of obs: 1865, groups:  Subject, 52

Fixed effects:
                  Estimate Std. Error t value
Instrumentguitar  3.97494    0.23499  16.916
Instrumentpiano   5.50315    0.25382  21.681
Instrumentstring  7.43247    0.24761  30.017
HarmonyI-V-IV     0.02992    0.10809   0.277
HarmonyI-V-VI     0.89608    0.21571   4.154
HarmonyIV-I-V     0.09070    0.10914   0.831
Voicepar3rd      -0.38972    0.08910  -4.374
Voicepar5th      -0.31014    0.08911  -3.480
X16.minus.17     -0.11195    0.04886  -2.291
PianoPlay         0.92342    0.36601   2.523
```

Figure 3: Summary output of Model lmer6.2

$$Classical_i = \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i + \alpha_3 Voice_i + \alpha_4 X16.minus.17 + \alpha_5 PianoPlay + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2) \quad (1)$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

The validation of lmer6.2 has been confirmed by examining the diagnostic plots of the standardized conditional residuals. The Figure 3 shows the summary output for model lmer6.2. Please see Technical Appendix page 27 to page 35 for more technical details.

3.1.2 Influences of experimental factors on Classical rating

To answer the question of what experimental factor, or combination of factors, has the strongest influence on Classical ratings, we can look at the Figure 3.We find that the Instrument did exert the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading),

```
Random effects:
 Groups    Name             Variance Std.Dev. Corr
 Subject   Instrumentguitar 0.73950  0.8599
           Instrumentpiano  1.21627  1.1028    0.15
           Instrumentstring 1.05715  1.0282   -0.99 -0.02
 Subject.1 (Intercept)      1.05344  1.0264
           HarmonyI-V-IV    0.05753  0.2399    0.87
           HarmonyI-V-VI    1.86869  1.3670   -0.36  0.05
           HarmonyIV-I-V    0.06998  0.2645    0.11 -0.12  0.16
 Residual                   2.46679  1.5706
Number of obs: 1865, groups:  Subject, 52

Fixed effects:
                 Estimate Std. Error t value
Instrumentguitar  3.97494    0.23499  16.916
Instrumentpiano   5.50315    0.25382  21.681
Instrumentstring  7.43247    0.24761  30.017
```

Figure 4: Coefficients of Instrument of Model lmer6.2

as the researchers suspected. According to the Figure 3, variable Instrument has the largest t-value and the largest coefficientsvalue among the three design factors (Instrument, Harmonic Motion, Voice Leading). This means the variable Instrument is most significant and has the strongest influence.

We can further interpret how different types of Instrument influence Classical rating by putting Instrument first and using "-1" to get rid of the baseline level of fixed effects and random effects. Looking at the Figure 4, among the three types of instrument, instrument string has the largest influence on Classical ratings, followed by instrument piano and instrument guitar. The coefficient of the fixed effect of Instrumentstring is around 7.4, which means on the group level, if the music is played by string, on average the rating for Classical music increases 7.4 points. But there is random variability across different listeners in the degree to which they are inclined to call music played by string "Classical". The standard deviation of the random effect of instrumentstring is around 1.0. So for almost 97.5% of the data, the overall coefficient of Instrumentstring is at least 5.4. This means on the individual level if the music is played by string, most of the time on average the rating for Classical music increases at least 5.4 points. In conclusion, the instrument string has the largest influence on Classical ratings among the three types of instruments.

To answer the question that among the different levels of Harmonic Motion whether I-V-VI have the strongest association with classical ratings, we can look at the Figure 5(Produced by putting variable Harmony of model lmer6.2 first and using "-1" to get rid of the baseline level of fixed effects and random effects). Looking at the fixed effects part, HarmonyI-V-VI has the largest t-value and the largest coefficient value among the four different types of Harmony. Looking at the part of the random effect, HarmonyI-V-VI has a similar standard deviation as the other three different types of Harmony. So overall, Harmonic Motion I-V-VI has the strongest association with classical ratings. The coefficient of the fixed effect of HarmonyI-V-VI is around 4.87, which means on the group level, if the music has the I-V-VI chord progression, on average the rating for Classical music increases 4.87 points. But there is random variability across different listeners in the degree to which they are inclined to call music with the I-V-VI chord progression "Classical". The standard deviation of the random effect of HarmonyI-V-VI is around 1.38. So for almost 97.5% of the data, the overall coefficient of Instrumentstring is at least 2.11. This means on the individual level if the music has the I-V-VI chord progression, most of the time on average the rating for Classical music increases

```
Random effects:
 Groups    Name             Variance Std.Dev. Corr
 Subject   (Intercept)      0.7395   0.8599
           Instrumentpiano  1.6669   1.2911   -0.54
           Instrumentstring 3.5488   1.8838   -1.00  0.60
 Subject.1 HarmonyI-IV-V    1.0534   1.0264
           HarmonyI-V-IV    1.5369   1.2397   1.00
           HarmonyI-V-VI    1.9108   1.3823   0.39  0.45
           HarmonyIV-I-V    1.1855   1.0888   0.97  0.96  0.42
 Residual                   2.4668   1.5706
Number of obs: 1865, groups:  Subject, 52

Fixed effects:
                Estimate Std. Error t value
HarmonyI-IV-V    3.97494    0.23499  16.916
HarmonyI-V-IV    4.00486    0.25405  15.764
HarmonyI-V-VI    4.87101    0.26791  18.182
HarmonyIV-I-V    4.06564    0.24033  16.917
```

Figure 5: Coefficients of Harmony of Model lmer6.2

at least 2.11 points. In summary, among the different levels of Harmonic Motion, Harmony I-V-VI has the strongest association with classical ratings.

To answer the question that whether the respondent is familiar with one or the other (or both) of the Pachelbel rants/comedy bits matters for Classical rating, we can first look at the Figure 6. As shown on the left-side box plot of the Figure 6, for listeners who have heard Rob Paravonian's Pachelbel Rant before, the mean of its classical rating is a little higher. But its IQR rage is fully covered by the IQR rage of the listeners who have never heard Rob Paravonian's Pachelbel Rant before, which indicates the influence of KnowRob on Classical rating is not much. As shown on the right-side box plot of the Figure 6, there is no difference in the mean of the classical rating between people who have heard Axis of Evil's Comedy bit and who have not. This means KnowAxis seems not to affect classical rating at all.

What's more, recalled that we have built different models by adding KnowRob and KnowAxis and their interactions with Harmony.The final mixed effect model we get has the formula: lmer(Classical Instrument + Harmony + Voice +Harmony:Voice+ (1 — Subject) + (Instrument — Subject) + (Harmony — Subject) -1). So KnowRob and KnowAxis are not kept in the lmer model, which indicates that it doesn't matter whether the listener is familiar with either one of the Pachelbel rants or comedy bit. Please refer to Technical Appendix page 22 to page 27 for more technical details. In conclusion, it doesn't matter for Classical rating whether the listener is familiar with either one of the Pachelbel rants or comedy bit.

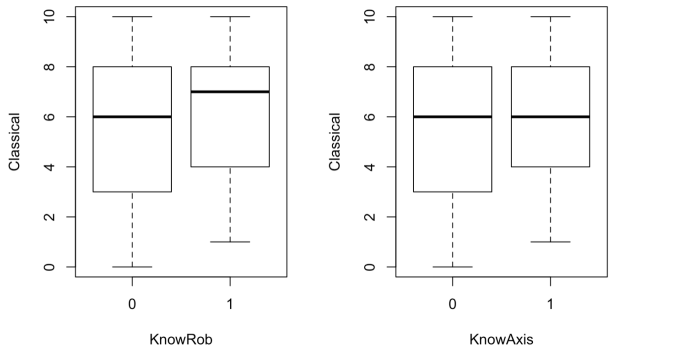


Figure 6: Boxplots of variable Classical and variable KnowRob and variable KnowAxis

To answer the question that among the levels of Voice Leading whether contrary motion has the strongest association with classical rating, we can look at the Figure 7(Produced by putting variable Voice of model lmer6.2 first and using "-1" to get rid of the baseline level of fixed effects). Variable Voice only has fixed effect. Voicecontrary has the largest t-value and the largest coefficient value among the three different types of Voice leading, which followed by Voicepar5th and Voicepar3rd. The coefficient of the fixed effect of Voicecontrary is around 3.97, which means on average, if the music uses a contrary motion for voice leading, the rating for Classical music increases 3.97 points. In summary, among the different levels of Voice Leading, Voice contrary has the strongest association with the classical rating.

```
Fixed effects:
                Estimate Std. Error t value
Voicecontrary    3.97494    0.23499  16.916
Voicepar3rd      3.58522    0.23495  15.260
Voicepar5th      3.66480    0.23499  15.596
```

Figure 7: Coefficients of Voice of Model lmer6.2

In conclusion, based on model lmer6.2, we find that the Instrument did exert the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading). Among the three types of instrument, instrument string has the largest influence on Classical ratings, followed by instrument piano and instrument guitar. Among the different levels of Harmonic Motion, Harmony I-V-VI has the strongest association with classical ratings. It doesn't matter whether the listener is familiar with either one of the Pachelbel rants/comedy bit for Classical rating. Among the different levels of Voice Leading, Voice contrary has the strongest association with the classical rating.

3.2 Differences between musicians and non-musicians on Classical rating

3.2.1 Model creation and model selection

Our researches hypotheses that people who are musicians may be influenced by things that do not influence non-musicians. To evaluate whether this hypothesis is true or not, we decided to use variable Selfdeclare to categorize participants into two groups: self-declared musicians and self-declared non-musicians.

We first dichotomized Selfdeclare at 2 and created a new dummy variable "Musicians" by defining "Non-Musicans" as Selfdeclare less and equal than 2 and "Musicians" as Selfdeclare bigger than 2. So we had 30 Non-Musicians and 22 Musicians. We first updated the fixed effects part of the model of lmer6.2 by adding the new variable Musicians and add its interaction with all the other explanatory variables.

After using stepAIC for variable selection for the fixed effects part, using function ffRanefLMER.fnc for forward selection for the random effects part, and re-examining the t-value of the coefficients of the fixed effects part after random effects are added, the final mixed model we get is model lmer8.2. The validation of model lmer8.2 has been confirmed by examining the diagnostic plots of the standardized conditional residuals. Please refer to Technical Appendix page 35 to page 40 for more technical details.

The following is the abbreviated form of model lmer8.2:

$$\begin{aligned} Classical_i &= \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i \\ &+\alpha_3 Voice_i + \alpha_4 X16.minus.17 + \alpha_5 PianoPlay + \alpha_6 Musicians{*}Harmony + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2) \end{aligned} \tag{2}$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

```
Random effects:
 Groups    Name             Variance  Std.Dev.  Corr
 Subject   (Intercept)      7.550e-14 2.748e-07
 Subject.1 (Intercept)      8.285e-01 9.102e-01
           Instrumentpiano  1.667e+00 1.291e+00 -0.51
           Instrumentstring 3.549e+00 1.884e+00 -0.99  0.60
 Subject.2 (Intercept)      1.177e+00 1.085e+00
           HarmonyI-V-IV    3.681e-02 1.919e-01  1.00
           HarmonyI-V-VI    1.423e+00 1.193e+00 -0.38 -0.38
           HarmonyIV-I-V    7.480e-02 2.735e-01  0.28  0.28 -0.22
 Residual                   2.469e+00 1.571e+00
Number of obs: 1865, groups:  Subject, 52

Fixed effects:
                                   Estimate Std. Error t value
Instrumentguitar                    3.46763    0.39864   8.699
Instrumentpiano                     4.99606    0.40992  12.188
Instrumentstring                    6.92523    0.40197  17.228
HarmonyI-V-IV                       0.10283    0.21357   0.482
HarmonyI-V-VI                       1.94897    0.39025   4.994
HarmonyIV-I-V                       0.40618    0.21984   1.848
Voicepar3rd                        -0.38959    0.08914  -4.370
Voicepar5th                        -0.31024    0.08915  -3.480
PianoPlay                           0.79593    0.39712   2.004
HarmonyI-IV-V:MusicansNon-Musicans  0.46053    0.40488   1.137
HarmonyI-V-IV:MusicansNon-Musicans  0.36340    0.45873   0.792
HarmonyI-V-VI:MusicansNon-Musicans -0.94297    0.45830  -2.058
HarmonyIV-I-V:MusicansNon-Musicans  0.04011    0.43434   0.092
```

Figure 8: Summary output of Model lmer8.2

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

We then dichotomized Selfdeclare at 3. So we had 39 Non-Musicians and 13 Musicians. Using similar approaches as mentioned above, the final mixed model we get is model lmer10.2, with the formula as lmer(Classical Instrument + Harmony + Voice + PianoPlay + X16.minus.17+Harmony:Musicans - 1+ (1 — Subject) + (Instrument — Subject) +(Harmony — Subject)). So Model lmer10.2 is the actuall the same model as the model lmer8.2, which means the results are not sensitive to where we dichotomize. Please to Technical Appendix page 40 to page 44 for more technical details.

We did not try dichotomizing "Selfdeclare" at 4 because this would give us 49 Non-Musicans and 3 Musicans, which would be problematic for analysis. So in conclusion, it seems that the results are not sensitive to where we dichotomize. We get the same model in the above two cases, which is the model lmer8.2. And the Figure 8 shows the summary output for the model lmer8.2.

3.2.2 Differences between musicians and non-musicians on Classical rating

We can evaluate the differences in the way that musicians and non-musicians identify classical music by looking at the Figure 8.We find that the interaction between HarmonyI-V-VI and MusicansNon-Musicans is significant and its coefficient is around -0.94. This means that if the subject (listener) is a self-declared non-Musician, the size of the effect of HarmonyI-V-VI on average Classical rating will additionally lower 0.94, compared to the subject(listener) who is a self-declared Musician.

In conclusion, the results are not sensitive to where we dichotomize Selfdeclare and we get

the same model as lmer8.2, with formula as lmer(Classical Instrument + Harmony + Voice + PianoPlay + X16.minus.17+Harmony:Musicans - 1+ (1 — Subject) + (Instrument — Subject) +(Harmony — Subject)). Based on model lmer8.2, if the subject (listener) is a self-declared non-Musician, the size of the effect of HarmonyI-V-VI on average Classical rating will additionally lower 0.94, compared to the subject(listener) who is a self-declared Musician.

3.3 Differences between factors that drive classical ratings and that drive popular ratings

3.3.1 Best Model for Popular rating

Before comparing the different factors that drive classical ratings and that drive popular ratings, we need building models for Popular ratings.

We first examined the influence of the three main experimental factors(Instrument, Harmony, Voice) on Popular ratings. We first produced a base model by considering all the interactions among the three main experimental factors. Using stepAIC and ANOVA test for model comparison, the variables selected for the fixed effects part are Instrument, Harmony and Voice. Please see Technical Appendix page 46 to page 48 for more technical details.

To determine which individual covariates should be added to the model as fixed effects, we added all other variables as possible explanatory variables to build a full model. UsingstepAIC, the fixed effects part selected are Instrument, Harmony, Voice, X16.minus.17, ConsInstr, PachListen,ClsListen, KnowRob, KnowAxis, X1990s2000s, X1990s2000s.minus.1960s1970s, Composing and PianoPlay.

To account for "personal biases" for ratings, we decided to add the random intercept for each participant(1—Subject). Similarly, to account for the variations of "personal biases" with the types of instrument, type of harmony, and type of voice leading, we added all the three possible random slopes (Instrument+Harmony+Voice—Subject). Using function ffRanefLMER.fnc for forward selection, the random effects selected are: (1 —Subject) + (Instrument — Subject) + (Harmony — Subject).

We added the selected random effects above and re-examined the fixed effects part. After dropping the variables of which the t-value of the coefficients are less than 2, we get the final mixed-effect model with the formula lmer(Popular Instrument + Harmony + Voice + X16.minus.17 + (1 — Subject) + (Instrument — Subject) +(Harmony — Subject) - 1, data = data). We then notice the coefficient of X16.minus.17 has a t-value less than 2; So we drop the variable X16.minus.17 too. The final model is lmer.4.3. The following is the abbreviated form of model lmer.4.3:

$$Popular_i = \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i + \alpha_3 Voice_i + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2) \quad (3)$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

```
Formula: Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |      Subject) + (Harmony | Subject) - 1
   Data: data
Control: lmerControl(optimizer = "bobyqa")

     AIC      BIC   loglik deviance df.resid
  7495.4   7639.2  -3721.7   7443.4     1839

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9873 -0.5734  0.0317  0.5610  3.2941

Random effects:
 Groups    Name             Variance Std.Dev. Corr
 Subject   (Intercept)      0.0000   0.0000
 Subject.1 (Intercept)      0.2965   0.5445
           Instrumentpiano  1.4316   1.1965   -0.65
           Instrumentstring 2.7342   1.6536   -1.00  0.66
 Subject.2 (Intercept)      1.5111   1.2293
           HarmonyI-V-IV    0.1136   0.3370    0.43
           HarmonyI-V-VI    1.0938   1.0459   -0.30 -0.38
           HarmonyIV-I-V    0.3314   0.5757   -0.52 -0.81 -0.17
 Residual                   2.5012   1.5815
Number of obs: 1865, groups:  Subject, 52

Fixed effects:
                 Estimate Std. Error t value
Instrumentguitar  6.74915    0.21325  31.649
Instrumentpiano   5.61683    0.23805  23.595
Instrumentstring  3.75867    0.25183  14.925
HarmonyI-V-IV    -0.01981    0.11362  -0.174
HarmonyI-V-VI    -0.34191    0.17826  -1.918
HarmonyIV-I-V    -0.25826    0.13071  -1.976
Voicepar3rd       0.20964    0.08973   2.336
Voicepar5th       0.24616    0.08973   2.743
```

Figure 9: Summary output of Model lmer.4.3

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

The validation of lmer.4.3 has been confirmed by examining the diagnostic plots of the standardized conditional residuals. The Figure 9 shows the summary output for model lmer.4.3. Please see Technical Appendix page 50 to page 56 for more technical details.

3.3.2 Differences between factors that drive classical ratings and that drive popular ratings

Recall on Section, the best model we get for Classical rating is model lmer6.2 and its abbreviated form is:

$$\begin{aligned} Classical_i = \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i \\ + \alpha_3 Voice_i + \alpha_4 X16.minus.17 + \alpha_5 PianoPlay + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2) \end{aligned}$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

Recalled the best model we get for Popular rating is model lmer.4.3 and its abbreviated form is:

$$\begin{aligned} Popular_i = \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i \\ + \alpha_3 Voice_i + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2) \end{aligned}$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

By comparing the above two formulas, we notice that variables X16.minus.17 (the auxiliary measure of listener's ability to distinguish classical vs popular music) and PianoPlay(listener plays Piano or not) influence Classical rating but do not affect Popular rating at all.

What's more, by comparing the Figure 9 with the Figure 3, we also notice the following differences:

- Comparing to that Instrument string has the largest influence on Classical ratings, the Instrument guitar has the largest influence on Popular ratings.

  Looking at the Figure 10(Produced by putting variable Instrument of model lmer.4.3 first and using "-1" to get rid of the baseline level of fixed effects and random effects). Looking at the fixed effects part, Instrument guitar has the largest t-value and the largest coefficient value among the three different types of Instrument. Looking at the random effects part, Instrument guitar has the smallest standard deviation among the three different types of Instrument. So overall, Instrument guitar has the strongest association with Popular ratings. The coefficient of the fixed effect of Instrument guitar is around 6.75, which means on the group level, if the music is played by guitar, on average the rating for Popular music increases 6.75 points. But there is random variability across different listeners in the degree to which they are inclined to call music played by guitar "Popular". The standard deviation of the random effect of Instrument guitar is around 0.54. So for almost 97.5% of the data, the overall coefficient of Instrumentstring is at least 5.67. This means on the individual level if the music is played by guitar, most of the time on average the rating for Popular music increases at least 5.67 points.

```
Random effects:
 Groups    Name             Variance Std.Dev. Corr
 Subject   Instrumentguitar 0.2965   0.5445
           Instrumentpiano  0.8765   0.9362   -0.25
           Instrumentstring 1.2302   1.1091   -1.00  0.27
 Subject.1 (Intercept)      1.5111   1.2293
           HarmonyI-V-IV    0.1136   0.3370    0.43
           HarmonyI-V-VI    1.0938   1.0459   -0.30 -0.38
           HarmonyIV-I-V    0.3314   0.5757   -0.52 -0.81 -0.17
 Residual                   2.5012   1.5815
Number of obs: 1865, groups:  Subject, 52

Fixed effects:
                Estimate Std. Error t value
Instrumentguitar  6.74915    0.21325  31.649
Instrumentpiano   5.61683    0.23805  23.595
Instrumentstring  3.75867    0.25183  14.925
HarmonyI-V-IV    -0.01981    0.11362  -0.174
HarmonyI-V-VI    -0.34191    0.17826  -1.918
HarmonyIV-I-V    -0.25826    0.13071  -1.976
Voicepar3rd       0.20964    0.08973   2.336
Voicepar5th       0.24616    0.08973   2.743
```

Figure 10: Coefficients of Instrument of model lmer.4.3

- Comparing to that Harmony I-V-VI has the strongest association with classical ratings among the different levels of Harmonic Motion, the different levels of Harmonic Motion seems to have very similar influences on Popular ratings.

  Turning to the Figure 11(Produced by putting variable Harmony of model lmer.4.3 first and using "-1" to get rid of the baseline level of fixed effects and random effects), the value of the coefficient of the fixed effects of the four different types of Harmony are very similar, varying from 6.40 to 6.75. Looking at the part of the random effects, they have similar standard deviations too, varying from 1.1 to 1.4. So overall, the different levels of Harmonic Motion seem to have very similar influences on Popular ratings.

```
Random effects:
 Groups    Name             Variance Std.Dev. Corr
 Subject   (Intercept)      0.2965   0.5445
           Instrumentpiano  1.4316   1.1965   -0.65
           Instrumentstring 2.7342   1.6536   -1.00  0.66
 Subject.1 HarmonyI-IV-V    1.5111   1.2293
           HarmonyI-V-IV    1.9806   1.4073   0.98
           HarmonyI-V-VI    1.8327   1.3538   0.68  0.61
           HarmonyIV-I-V    1.1100   1.0536   0.88  0.79  0.46
 Residual                   2.5012   1.5815
Number of obs: 1865, groups:  Subject, 52

Fixed effects:
                 Estimate Std. Error t value
HarmonyI-IV-V     6.74915    0.21325  31.649
HarmonyI-V-IV     6.72934    0.23350  28.819
HarmonyI-V-VI     6.40724    0.22742  28.173
HarmonyIV-I-V     6.49089    0.19433  33.402
Instrumentpiano  -1.13232    0.18874  -5.999
Instrumentstring -2.99048    0.24618 -12.147
Voicepar3rd       0.20964    0.08973   2.336
Voicepar5th       0.24616    0.08973   2.743
```

Figure 11: Coefficients of Harmony of model lmer.4.3

- Comparing to that Voice contrary has the strongest association with classical rating among the different levels of Voice Leading, the Voice par5th has the strongest association with Popular ratings.

  Looking at the Figure 12(Produced by putting variable Voice of model lmer.4.3 first and using "-1" to get rid of the baseline level of fixed effects). Variable Voice only has fixed effect. Voicepar5th has the largest t-value and the largest coefficient value among the three different types of Voice leading, which followed by Voicepar3rd and Voicecontrary. The coefficient of the fixed effect of Voicepar5th is around 7.0, which means on average, if the music uses par5th for voice leading, the rating for Popular music increases 7.0 points.

```
Fixed effects:
                 Estimate Std. Error t value
Voicecontrary     6.74915    0.21325  31.649
Voicepar3rd       6.95879    0.21322  32.636
Voicepar5th       6.99531    0.21326  32.801
HarmonyI-V-IV    -0.01981    0.11362  -0.174
HarmonyI-V-VI    -0.34191    0.17826  -1.918
HarmonyIV-I-V    -0.25826    0.13071  -1.976
Instrumentpiano  -1.13232    0.18874  -5.999
Instrumentstring -2.99048    0.24618 -12.147
```

Figure 12: Coefficients of Voice of model lmer.4.3

4 Discussion

In our analysis, the final mixed model we get for Classical rating is model lmer6.2 and its abbreviated form is:

$$Classical_i = \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i + \alpha_3 Voice_i + \alpha_4 X16.minus.17 + \alpha_5 PianoPlay + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

Based on model lmer6.2, we find that Instrument did exert the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading). Among the three types of instrument, instrument string has the largest influence on Classical ratings, followed by instrument piano and instrument guitar. Among the different levels of Harmonic Motion, Harmony I-V-VI has the strongest association with classical ratings. It doesn't matter whether the listener is familiar with either one of the Pachelbel rants/comedy bit for Classical rating. Among the different levels of Voice Leading, Voice contrary has the strongest association with the classical rating.

To evaluate the differences in the way that musicians and non-musicians identify classical music, the best-mixed model we get is lmer8.2 and its abbreviated form is:

$$Classical_i = \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i + \alpha_3 Voice_i + \alpha_4 X16.minus.17 + \alpha_5 PianoPlay + \alpha_6 Musicians * Harmony + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

Based on the model lmer8.2, we find that if the subject (listener) is a self-declared non-Musician, the size of the effect of HarmonyI-V-VI on average Classical rating will additionally lower 0.94, compared to the subject(listener) who is a self-declared Musician. And the results are not sensitive to where we dichotomize Selfdeclare.

The best model we get for Popular rating is model lmer.4.3 and its abbreviated form is:

$$Popular_i = \alpha_{0j[i]} + \alpha_{1j[i]} Instrument_i + \alpha_{2j[i]} Harmony_i + \alpha_3 Voice_i + \epsilon_i, \qquad \epsilon_i \sim^{iid} \mathcal{N}(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_{00} + \eta_{0j}, \qquad \eta_{0j} \sim^{iid} \mathcal{N}(0, \tau_0^2)$$

$$\alpha_{1j[i]} = \beta_{10} + \eta_{1j}, \qquad \eta_{1j} \sim^{iid} \mathcal{N}(0, \tau_1^2)$$

$$\alpha_{2j[i]} = \beta_{20} + \eta_{2j}, \qquad \eta_{2j} \sim^{iid} \mathcal{N}(0, \tau_2^2)$$

By comparing model lmer.4.3 with model lmer6.2, we find that variables X16.minus.17 (the auxiliary measure of listener's ability to distinguish classical vs popular music) and PianoPlay(listener plays Piano or not) influence Classical rating but do not affect Popular rating at all. What's more, we also notice that: comparing to that Instrument string has the largest influence on Classical ratings, the Instrument guitar has the largest influence on Popular ratings; comparing to that Harmony I-V-VI has the strongest association with classical ratings among the different levels of Harmonic Motion, the different levels of Harmonic Motion seems to have very similar influences on Popular ratings; comparing to that Voice contrary has the strongest association with classical rating among the different levels of Voice Leading, the Voice par5th has the strongest association with Popular ratings.

Our study was limited by the use of Ivan Jimenez et al. (2013) data. There are a lot of missing values for this original data set. There are multiple ways of dealing with missingness. Subjective decisions of different approaches of dealing with missingness could create different data set for analysis and thus lead to different variables selected and completely different models.

Table 1: the frequency table of OMSI Score

	OMSI <500	OMSI >= 500
Subjects	43	9
Percentage(%)	82.7%	17.3%

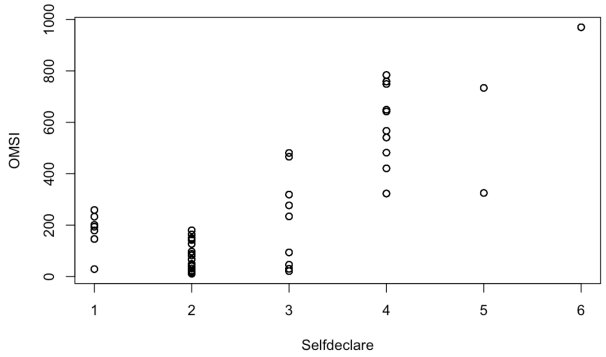


Figure 13: Scatter plot of OMSI score and Selfdeclare

Also, the data is biased toward listeners who have limited musical knowledge. As shown in the Table 1, about 82.7% of the surveyed listeners has an OMSI score that is less than 500. According to the website of the MARCS Institute, the Ollen Musical Sophistication Index (OMSI) respondents with a score greater than 500 should be classified as "more musically sophisticated" and those with a score less than 500 as "less musically sophisticated." Thus, the is very biased towards listeners who have limited musical knowledge. So the model we generalized based on this biased data set may not be useful for making a prediction for people who have sophisticated musical knowledge. What's more, because of this biased data, the conclusions we draw for the differences between musicians and non-musicians on Classical ratings may not capture the truth at all. As shown in the Figure 13, the score of the Variable Selfdeclare may not indicate the actual musical knowledge of the listener at all. To address this shortcoming, more survey needs to be done to collect data from listeners who have sophisticated musical knowledge and we should also consider using more subjective measurement such as OMSI to define Musicians and Non-Musicians.

In summary, we find that the Instrument exerts the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading). If the subject (listener) is a self-declared non-Musician, the size of the effect of HarmonyI-V-VI on average Classical rating will additionally lower 0.94. Variables X16.minus.17 and PianoPlay influence Classical rating but do not affect Popular rating at all. The data is biased towards listeners who have limited musical knowledge. More survey needs to be done to collect data from listeners who have sophisticated musical knowledge and we should also consider using more subjective measurement such as OMSI to define Musicians and Non-Musicians.

References

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/.

Ivan Jimenez et al. (2013), *The Influence of Timbre, Harmony, and Voice Leading on Listeners' Distinction between Popular and Classical Music.* Pittsburgh, PA: The University of Pittsburgh.

Dictionary of Cambridge *Definitions of Classical Music and Popular Music* dictionary.cambridge.org URL https://dictionary.cambridge.org/us/dictionary/english/tradition

Bowman, W. D. (2004), *Popular music: The very idea of listening to it. In C. X. Rodriguez (Ed.), Bridging the gap: Popular music and music education.* Reston,VA: The National Association for Music Education.

The MARCS Institute *What is the Ollen Musical Sophistication Index (OMSI)?* marcs-survey.uws.edu.au, The MARCS Institute. URL http://marcs-survey.uws.edu.au/OMSI/omsi.php.

Technical Appendix

Yuanyuan Wang

Dec/12/2019

1.1 Data Cleaning Approach

1) Missing values

We have a total of 2520 observations. Since column X1stInstr has 1512 missing values and column X2ndInstr has 2196 missing values, we decide to remove these two columns.

Also, column ConsNotes has 360 missing values and is not an important measurement, we decide to delete this column. Column Instr.minus.Notes is calculated based on column ConsNotes. So we delete the Column Instr.minus.Notes too.

NoClass has 288 missing values and is not an important measurement, we decide to delete this column.

2) Unneeded columns

Since we do not need X1 and first12 for the analysis, we also delete these two columns too.

After we delete these columns, we use na.omit to discard all the observations with NA values.

3) Unusual values

We notice column Popular has a value of 19, which is unnormal. We suspect that this is a typo and should be value of 10.So we replace the value of 19 with 10.

Then, we have 1865 observations with 52 subjects left.

```
data <- read.csv("ratings.csv")

# delete columns  "X1","X1stInstr","X2ndInstr","first12","ConsNotes",
# "Instr.minus.Notes","NoClass"

data <- data[,-c(1,10,11,19,24,25,26)]


# omit NA values
data <- na.omit(data)

# Replace observation with Popular of 19 with 10 10

data[data$Popular == 19,]$Popular <- 10

# examine the cleaned data using skim
skim(data)
```

```
## Skim summary statistics
##  n obs: 1865
##  n variables: 21
##
## -- Variable type:factor ---------------------------------------------------------------------------
##    variable missing complete    n n_unique
##     Harmony       0     1865 1865        4
```

```
##  Instrument       0     1865 1865        3
##     Subject       0     1865 1865       52
##       Voice       0     1865 1865        3
##                               top_counts ordered
##  I-I: 467, IV-: 467, I-V: 466, I-V: 465   FALSE
##     str: 624, gui: 623, pia: 618, NA: 0   FALSE
##          15: 36, 16: 36, 17: 36, 19: 36   FALSE
##     par: 622, par: 622, con: 621, NA: 0   FALSE
##
## -- Variable type:integer ------------------------------------------------------------------------
##                       variable missing complete    n   mean     sd p0 p25
##                       APTheory       0     1865 1865   0.23   0.42  0   0
##                      ClsListen       0     1865 1865   2.17   1.6   0   1
##                    CollegeMusic       0     1865 1865   0.81   0.39  0   1
##                      Composing       0     1865 1865   0.98   1.42  0   0
##                     GuitarPlay       0     1865 1865   0.76   1.6   0   0
##                       KnowAxis       0     1865 1865   1.08   2.04  0   0
##                        KnowRob       0     1865 1865   0.74   1.68  0   0
##                           OMSI       0     1865 1865 252.31 244.78 11  67
##                     PachListen       0     1865 1865   4.65   0.85  2   5
##                      PianoPlay       0     1865 1865   1.14   1.76  0   0
##                    Selfdeclare       0     1865 1865   2.61   1.16  1   2
##                    X1990s2000s       0     1865 1865   4.17   1.38  0   3
##  X1990s2000s.minus.1960s1970s       0     1865 1865   1.93   1.72 -3   0
##  p50 p75 p100     hist
##    0   0    1
##    3   3    5
##    1   1    1
##    0   2    5
##    0   0    5
##    0   0    5
##    0   0    5
##  147 325  970
##    5   5    5
##    0   1    5
##    2   3    6
##    5   5    5
##    2   3    5
##
## -- Variable type:numeric ------------------------------------------------------------------------
##      variable missing complete    n mean   sd p0  p25 p50  p75 p100
##     Classical       0     1865 1865 5.65 2.69  0 3      6 8      10
##     ConsInstr       0     1865 1865 3.05 1.46  0 2.33   3 4.33    5
##       Popular       0     1865 1865 5.37 2.55  0 3      6 7      10
##  X16.minus.17       0     1865 1865 1.66 2.96 -4 0      1 3       9
##      hist
##
##
##
##
```

1.2 Examine distribution and transformation

We first examine all the distributions of the numeric scoers. After examining the distributions, we find that:

1)It seems that OMSI are right-skewed and we could use log transformation to bring the distribution back to center.

2) PachListen either has a high value of 5 or very low values. It can be dichotomized to two groups, with 0 indicates orginal value less than 4, 1 indicates orginal value equal or more than 4;

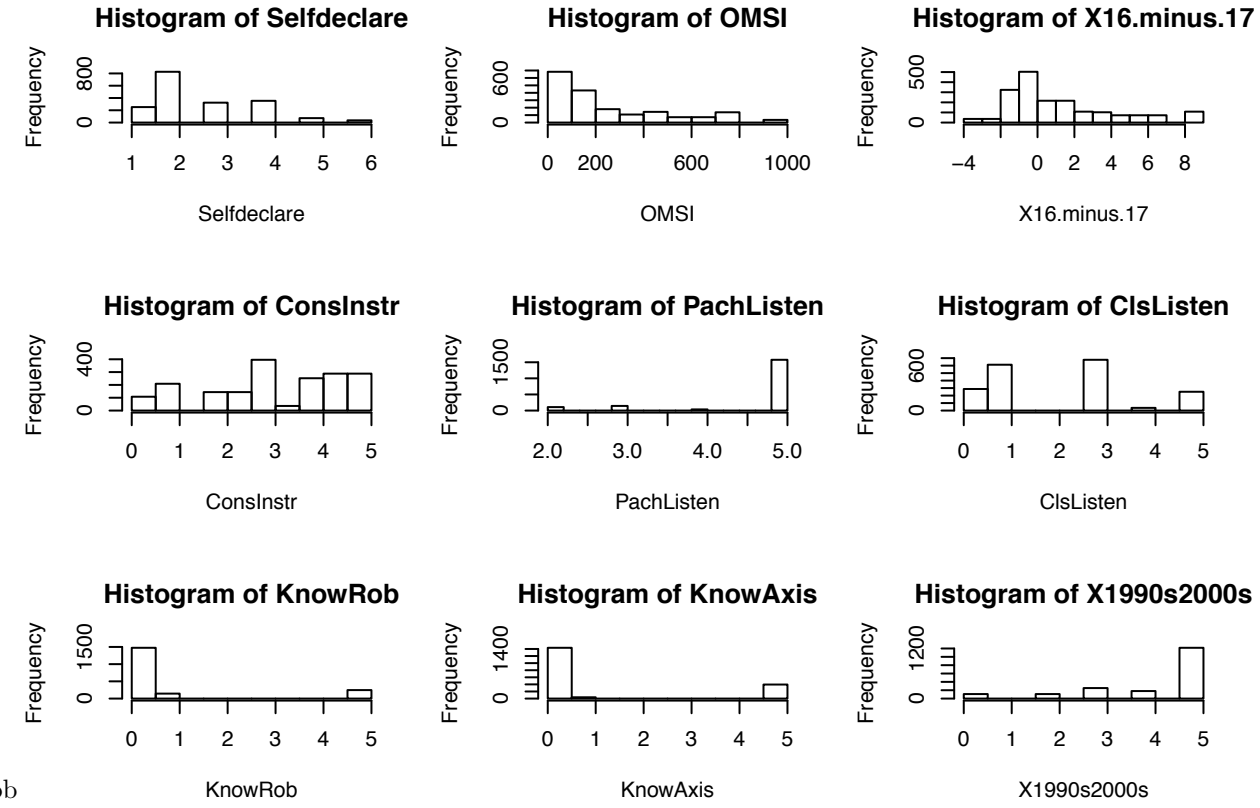
3) KnowRob either has a very high value or a very low value.It can be dichotomized to two groups, with 0 indicates orginal value less than 2, 1 indicates orginal value equal or more than 4;

4) KnowAxis either has a very high value or a very low value.It can be dichotomized to two groups, with 0 indicates orginal value less than 2, 1 indicates orginal value equal or more than 4;

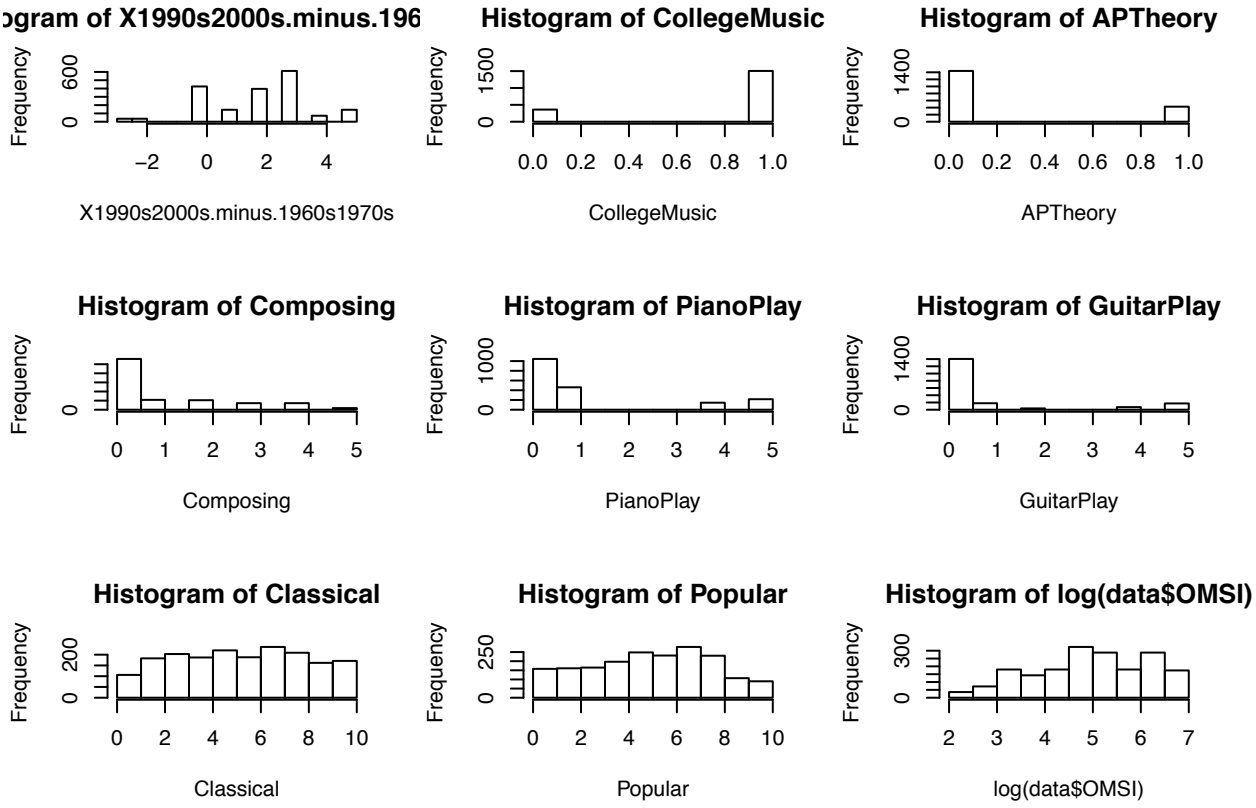
5) PianoPlay either has a very high value or a very low value.It can be dichotomized to two groups, with 0 indicates orginal value no more than 2, 1 indicates orginal value more than 3;

6) GuitarPlay either has a very high value or a very low value.It can be dichotomized to two groups, with 0 indicates orginal value no more than 2, 1 indicates orginal value more than 3;

So we changed the original datasets by transformation and dichotomizing these columns



1.2-1.bb



1.2-2.bb

1.3 Examine relationships

1) Examine the relationship with column Classical

We know that columns: 1) Harmony,Instrument,Voice, Subject are categorical variables;

2) APTheory,CollegeMusic,KnowAxis,KnowRob,PachListen,GuitarPlay and PianoPlay are actually categorical variables which are coded as numeric (0 or 1);

We treate all other columns as numeric; we could use scatter plots to examine the relationships between numeric variables. But since these are all discrete values, scatter plots won't give you any useful information.So we decide go with boxplot.

From the boxplots below, we find that:

Classical seems to be positively related with Harmony(I-V-VI) and Harmony(IV-I-V).

Classical seems to be postively related with Instrument(string).

Classical seems to be negatively related with Voice(par3rd).

There is large variance on the mean of the score for Classical between different subjects.

APTheory doesn't seem to matter for Classical.

ClsListen doesn't seem to matter for Classical.

CollegeMusic doesn't seem to matter for Classical.

GuitarPlay doesn't seem to matter for Classical.

KnowAxis doesn't seem to matter for Classical.

Classical seems to be positively related with KnowRob.

Classical seems to be positively related with PachListen.

Classical seems to be positively related with PianoPlay.

Classical seems to be postively related with X1990s2000s.

Classical seems to be different for different values of Composing.

Classical seems to be different for different values of ConsInstr.

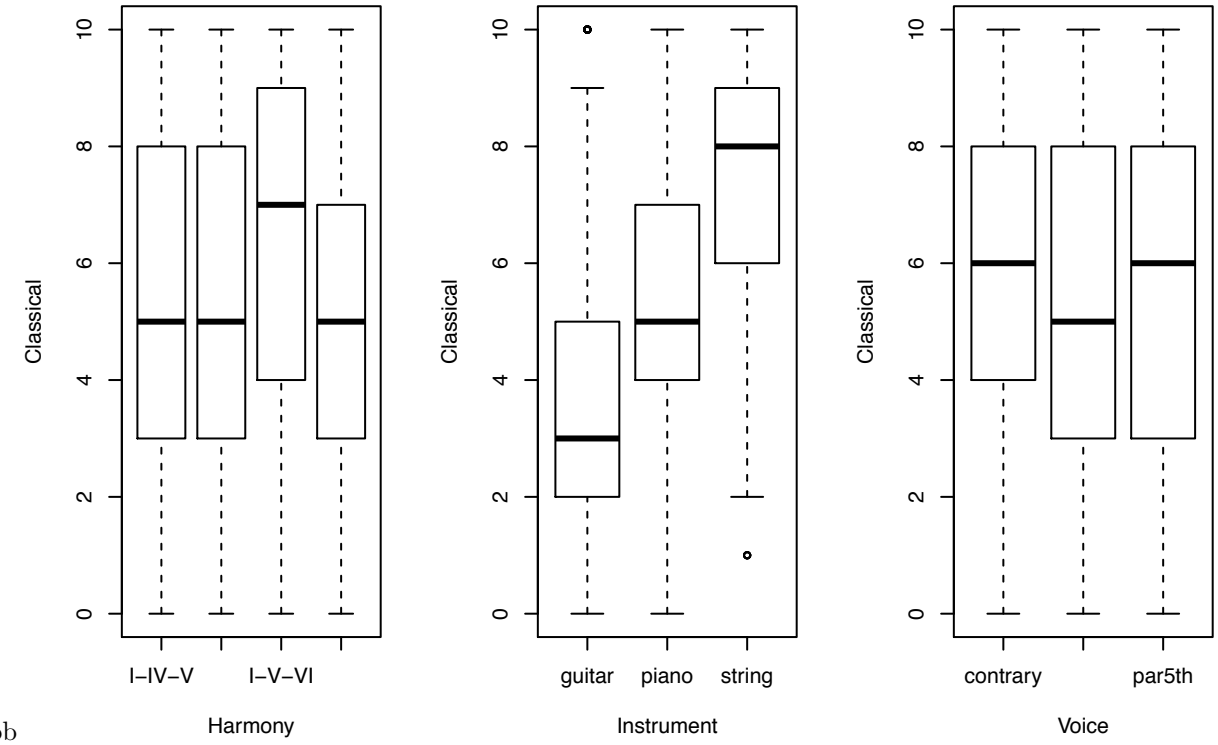
Classical seems to be different for different values of Selfdeclare.

Classical seems to be different for different values of LogOMSI.

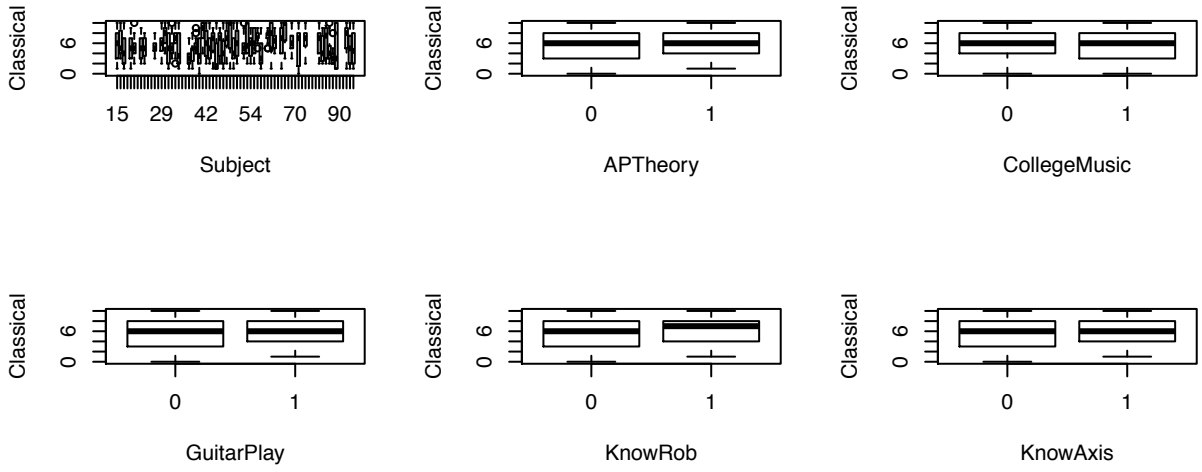
Classical seems to be different for different values of X16.minus.17.

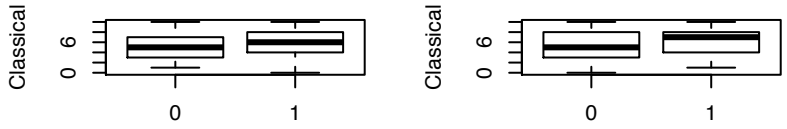
Classical seems to be different for different values of X1990s2000s.minus.1960s1970s.

```
## The following objects are masked from data (pos = 3):
##
##     APTheory, Classical, ClsListen, CollegeMusic, Composing,
##     ConsInstr, GuitarPlay, Harmony, Instrument, KnowAxis, KnowRob,
##     PachListen, PianoPlay, Popular, Selfdeclare, Subject, Voice,
##     X16.minus.17, X1990s2000s, X1990s2000s.minus.1960s1970s
```

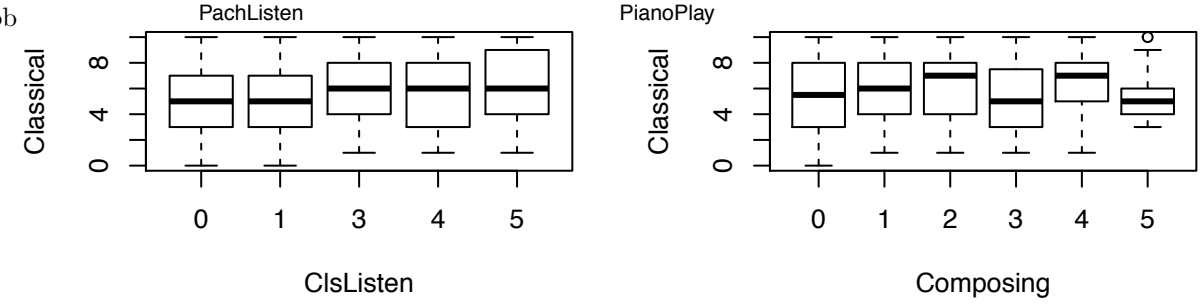


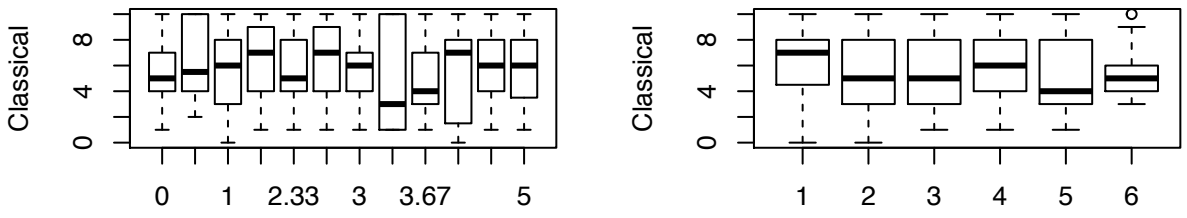
1.2.2-1.bb



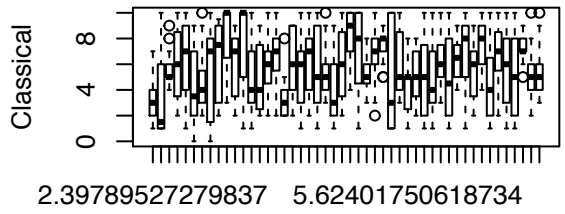


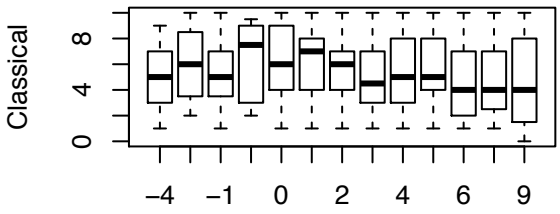
1.2.2-2.bb

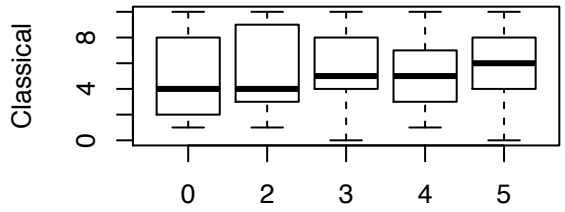


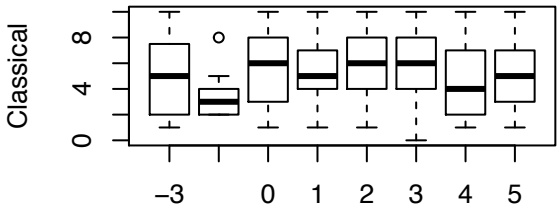


1.2.2-3.bb









1.2.2-4.bb

2) Examine the relationship with column Popular

From the boxplots below, we find that:

Popular seems to be negatively related with Harmony(I-V-VI) and Harmony(IV-I-V).

Popular seems to be negatively related with Instrument(string).

Popular seems to be negatively related with Voice(contrary).

There is large variance on the mean of the score for Popular between different subjects.

Popular seems to be positively related with APTheory.

Popular seems to be positively related with CollegeMusic.

Popular seems to be positively related with GuitarPlay.

Popular seems to be positively related with knowAxis. Popular seems to be positively related with KnowRob.

Popular seems to be positively related with PianoPlay.

Popular seems to be negatively related with PachListen.

Popular seems to be different for different values of ClsListen.

Popular seems to be different for different values of Composing.

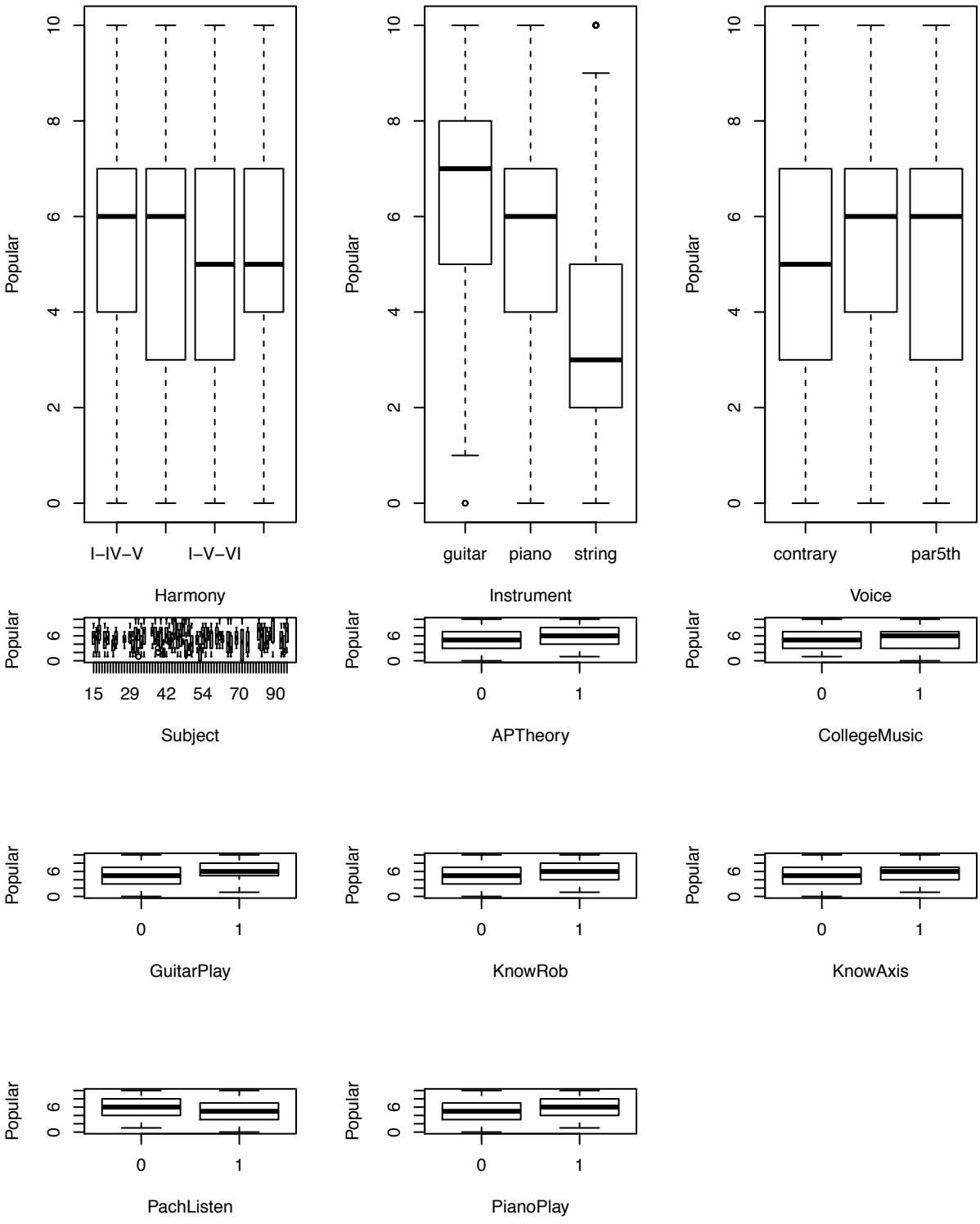
Popular seems to be different for different values of ConsInstr.

Popular seems to be different for different values of Selfdeclare.

Popular seems to be different for different values of LogOMSI.

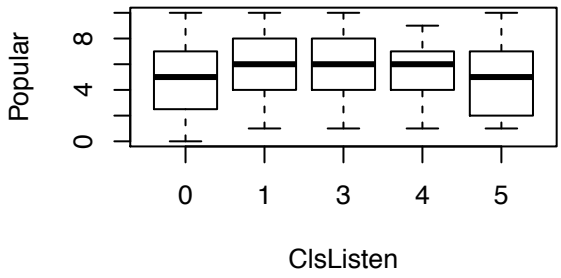
Popular seems to be different for different values of X16.minus.17.

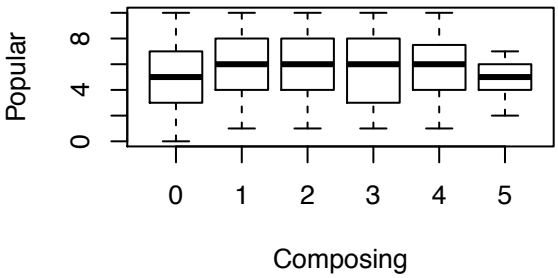
Popular seems to be different for different values of X1990s2000s.minus.1960s1970s.

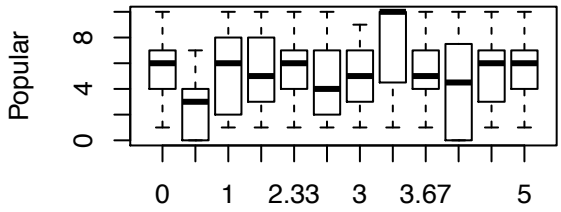

Popular
0
2
4
6
8
10
I–IV–V
I–V–VI
Harmony
guitar
piano
string
Instrument
contrary
par5th
Voice
15
29
42
54
70
90
Subject
APTheory
CollegeMusic
GuitarPlay
KnowRob
KnowAxis
PachListen
PianoPlay

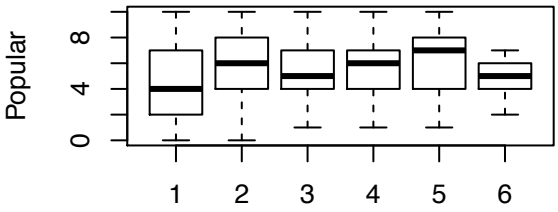
1.3-1.bb

1.3-2.bb

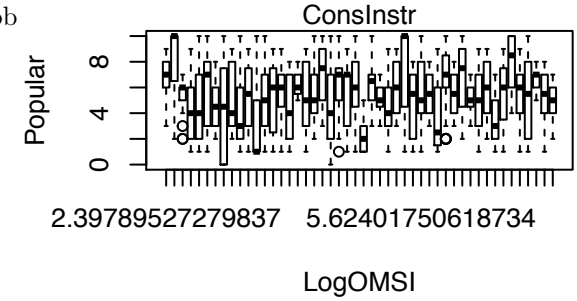


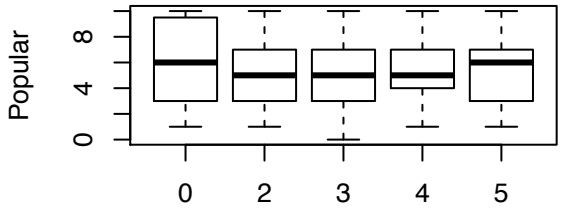


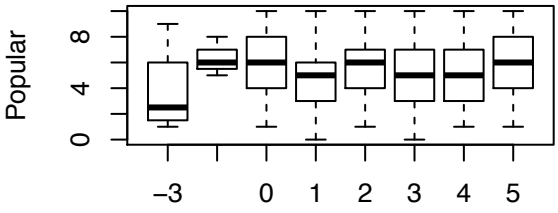




1.3-3.bb







1.3-4.bb

2 Model for Classical— Examing the three main effects:

2.(a) Conventional linear models for examining Instrument, Harmony & Voice model selection:

We first produce lm1 by considering interactions among Instrument, Harmony & Voice, using lm1 <- lm(Classical~ Instrument*Harmony*Voice, data = data). We notice from the summary output of lm1 that only variables Instrument, Harmony, and the interactions between Harmony and voice are significant.

So we then produce a simple model lm2 only using Instrument, Harmony, Voice and the interactions between Harmony and Voice (lm2 <- lm(Classical~ Instrument+Harmony*Voice, data = data)). And we notice that only only variables Instrument and Harmony are significant.

Using anova test, the simple model lm2 is prefered over lm1.

We then produce a simple model lm3 only using Instrument, Harmony and Voice.We notice that all the variables are significant.

Using anova test, model lm2 is actually prefered over lm3.

Using stepAIC, it selected the same model, which is lm2. So the final model is lm2 with (formula = Classical ~ Instrument + Harmony + Voice + Harmony:Voice - 1, data = data).

We can examine whether lm2 is valid or not using the normal Q-Q plot from the diagnostic plots and binnedplot. We notice the normal Q-Q plot almost follows a straight line. Looking at the Binned residual plot, we find almost all the points lie in the 95% confidence interval for the residulas, except the points at the two ends. And the points randomly scatter around the line at zero.

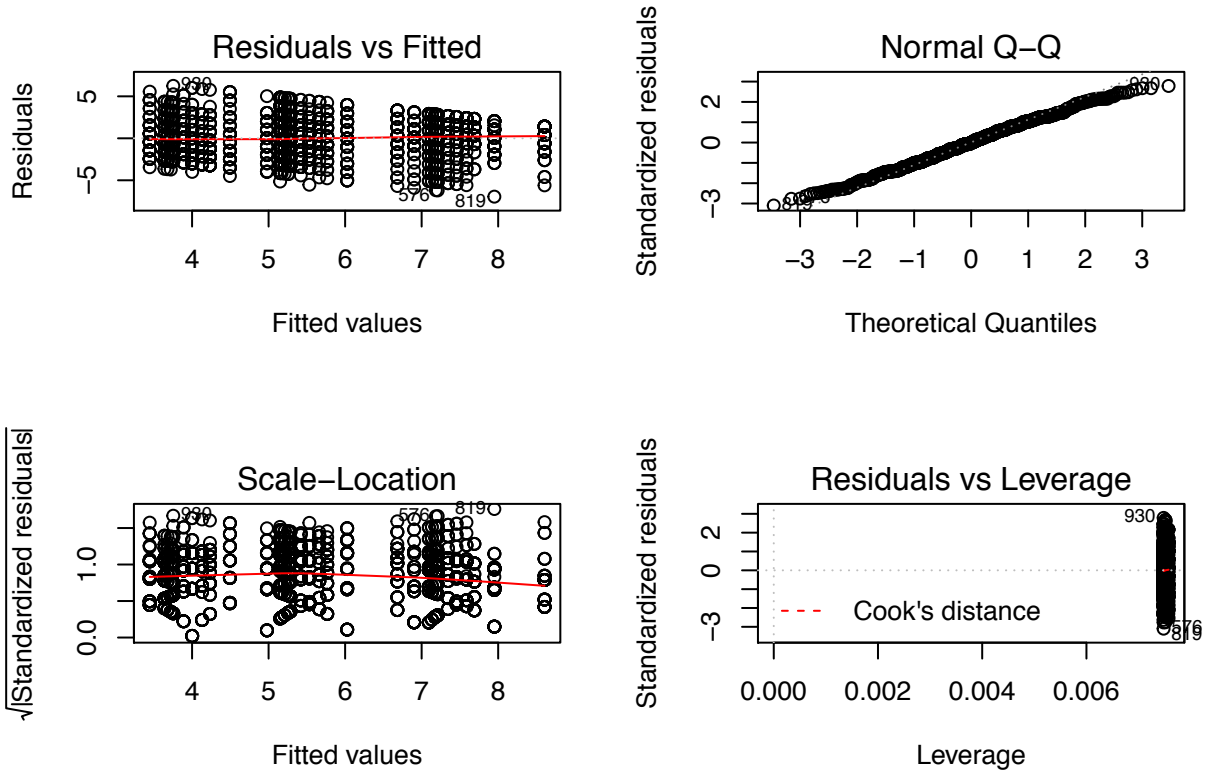
In conclusion, model lm2 seems valid.

```
##
## Call:
## lm(formula = Classical ~ Instrument * Harmony * Voice - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8365 -1.6538  0.0385  1.6275  6.5000
##
## Coefficients:
##                                              Estimate Std. Error t value
## Instrumentguitar                              4.13462    0.31233  13.238
## Instrumentpiano                               5.23077    0.31233  16.748
## Instrumentstring                              7.28846    0.31233  23.336
## HarmonyI-V-IV                                -0.32692    0.44170  -0.740
## HarmonyI-V-VI                                 0.96342    0.44386   2.171
## HarmonyIV-I-V                                -0.46154    0.44170  -1.045
## Voicepar3rd                                  -0.19231    0.44170  -0.435
## Voicepar5th                                  -0.55769    0.44170  -1.263
## Instrumentpiano:HarmonyI-V-IV                 0.84125    0.62618   1.343
## Instrumentstring:HarmonyI-V-IV                0.86538    0.62465   1.385
## Instrumentpiano:HarmonyI-V-VI                 0.57051    0.62771   0.909
## Instrumentstring:HarmonyI-V-VI                0.32504    0.62618   0.519
## Instrumentpiano:HarmonyIV-I-V                 0.63462    0.62465   1.016
## Instrumentstring:HarmonyIV-I-V                0.25000    0.62465   0.400
## Instrumentpiano:Voicepar3rd                   0.01923    0.62465   0.031
## Instrumentstring:Voicepar3rd                 -0.19231    0.62465  -0.308
## Instrumentpiano:Voicepar5th                   0.58183    0.62618   0.929
## Instrumentstring:Voicepar5th                  0.46154    0.62465   0.739
## HarmonyI-V-IV:Voicepar3rd                     0.03846    0.62465   0.062
## HarmonyI-V-VI:Voicepar3rd                    -1.40573    0.62618  -2.245
## HarmonyIV-I-V:Voicepar3rd                     0.48077    0.62465   0.770
## HarmonyI-V-IV:Voicepar5th                     0.55769    0.62465   0.893
## HarmonyI-V-VI:Voicepar5th                     0.11350    0.62618   0.181
## HarmonyIV-I-V:Voicepar5th                     0.94231    0.62465   1.509
## Instrumentpiano:HarmonyI-V-IV:Voicepar3rd  -0.95664    0.88448  -1.082
## Instrumentstring:HarmonyI-V-IV:Voicepar3rd -0.46154    0.88339  -0.522
```

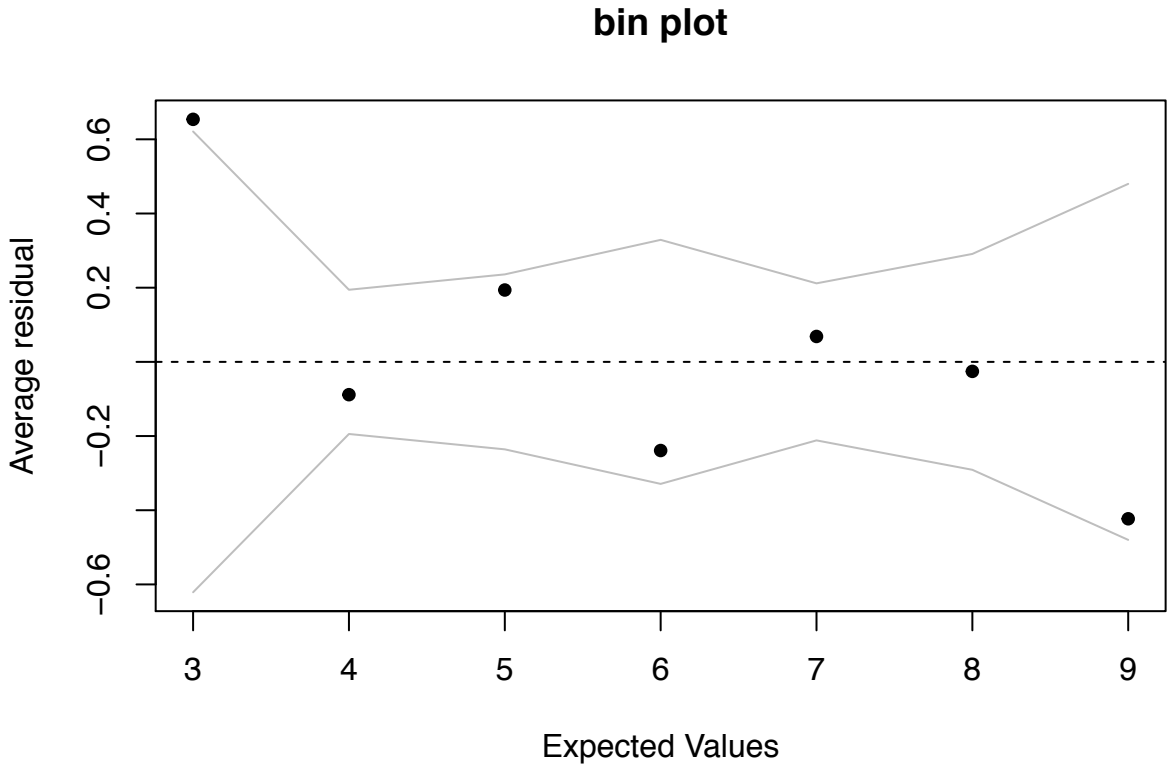
```
## Instrumentpiano:HarmonyI-V-VI:Voicepar3rd   0.75528    0.88664   0.852
## Instrumentstring:HarmonyI-V-VI:Voicepar3rd  1.46342    0.88448   1.655
## Instrumentpiano:HarmonyIV-I-V:Voicepar3rd  -0.33899    0.88448  -0.383
## Instrumentstring:HarmonyIV-I-V:Voicepar3rd  0.48077    0.88339   0.544
## Instrumentpiano:HarmonyI-V-IV:Voicepar5th  -0.83673    0.88664  -0.944
## Instrumentstring:HarmonyI-V-IV:Voicepar5th -1.40385    0.88339  -1.589
## Instrumentpiano:HarmonyI-V-VI:Voicepar5th  -0.92157    0.88664  -1.039
## Instrumentstring:HarmonyI-V-VI:Voicepar5th -0.75773    0.88448  -0.857
## Instrumentpiano:HarmonyIV-I-V:Voicepar5th  -1.00490    0.88448  -1.136
## Instrumentstring:HarmonyIV-I-V:Voicepar5th -1.09615    0.88339  -1.241
##                                           Pr(>|t|)
## Instrumentguitar                            <2e-16 ***
## Instrumentpiano                             <2e-16 ***
## Instrumentstring                            <2e-16 ***
## HarmonyI-V-IV                               0.4593
## HarmonyI-V-VI                               0.0301 *
## HarmonyIV-I-V                               0.2962
## Voicepar3rd                                 0.6633
## Voicepar5th                                 0.2069
## Instrumentpiano:HarmonyI-V-IV               0.1793
## Instrumentstring:HarmonyI-V-IV              0.1661
## Instrumentpiano:HarmonyI-V-VI               0.3635
## Instrumentstring:HarmonyI-V-VI              0.6038
## Instrumentpiano:HarmonyIV-I-V               0.3098
## Instrumentstring:HarmonyIV-I-V              0.6890
## Instrumentpiano:Voicepar3rd                 0.9754
## Instrumentstring:Voicepar3rd                0.7582
## Instrumentpiano:Voicepar5th                 0.3529
## Instrumentstring:Voicepar5th                0.4601
## HarmonyI-V-IV:Voicepar3rd                   0.9509
## HarmonyI-V-VI:Voicepar3rd                   0.0249 *
## HarmonyIV-I-V:Voicepar3rd                   0.4416
## HarmonyI-V-IV:Voicepar5th                   0.3721
## HarmonyI-V-VI:Voicepar5th                   0.8562
## HarmonyIV-I-V:Voicepar5th                   0.1316
## Instrumentpiano:HarmonyI-V-IV:Voicepar3rd   0.2796
## Instrumentstring:HarmonyI-V-IV:Voicepar3rd  0.6014
## Instrumentpiano:HarmonyI-V-VI:Voicepar3rd   0.3944
## Instrumentstring:HarmonyI-V-VI:Voicepar3rd  0.0982 .
## Instrumentpiano:HarmonyIV-I-V:Voicepar3rd   0.7016
## Instrumentstring:HarmonyIV-I-V:Voicepar3rd  0.5863
## Instrumentpiano:HarmonyI-V-IV:Voicepar5th   0.3454
## Instrumentstring:HarmonyI-V-IV:Voicepar5th  0.1122
## Instrumentpiano:HarmonyI-V-VI:Voicepar5th   0.2988
## Instrumentstring:HarmonyI-V-VI:Voicepar5th  0.3917
## Instrumentpiano:HarmonyIV-I-V:Voicepar5th   0.2560
## Instrumentstring:HarmonyIV-I-V:Voicepar5th  0.2148
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.252 on 1829 degrees of freedom
## Multiple R-squared:  0.873,  Adjusted R-squared:  0.8705
## F-statistic: 349.3 on 36 and 1829 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony * Voice - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9502 -1.6621 -0.0265  1.7139  6.2459
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar             3.8887     0.1947  19.969  < 2e-16 ***
## Instrumentpiano              5.4207     0.1948  27.828  < 2e-16 ***
## Instrumentstring             7.3444     0.1947  37.717  < 2e-16 ***
## HarmonyI-V-IV                0.2414     0.2553   0.946   0.3445
## HarmonyI-V-VI                1.2617     0.2557   4.934  8.8e-07 ***
## HarmonyIV-I-V               -0.1667     0.2549  -0.654   0.5133
## Voicepar3rd                 -0.2500     0.2549  -0.981   0.3269
## Voicepar5th                 -0.2102     0.2553  -0.823   0.4105
## HarmonyI-V-IV:Voicepar3rd   -0.4337     0.3608  -1.202   0.2295
## HarmonyI-V-VI:Voicepar3rd   -0.6671     0.3614  -1.846   0.0651 .
## HarmonyIV-I-V:Voicepar3rd    0.5291     0.3608   1.466   0.1427
## HarmonyI-V-IV:Voicepar5th   -0.1898     0.3614  -0.525   0.5995
## HarmonyI-V-VI:Voicepar5th   -0.4458     0.3614  -1.234   0.2175
## HarmonyIV-I-V:Voicepar5th    0.2422     0.3608   0.671   0.5021
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.251 on 1851 degrees of freedom
## Multiple R-squared:  0.8716, Adjusted R-squared:  0.8706
## F-statistic: 897.5 on 14 and 1851 DF,  p-value: < 2.2e-16

## Analysis of Variance Table
##
## Model 1: Classical ~ Instrument + Harmony * Voice - 1
## Model 2: Classical ~ Instrument * Harmony * Voice - 1
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1   1851 9382.3
## 2   1829 9277.6 22    104.75 0.9386 0.5427

## Analysis of Variance Table
##
## Model 1: Classical ~ Instrument + Harmony + Voice - 1
## Model 2: Classical ~ Instrument + Harmony * Voice - 1
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1   1857 9448.0
## 2   1851 9382.3  6    65.626 2.1578 0.04439 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice + Harmony:Voice -
##     1, data = data)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -6.9502 -1.6621 -0.0265  1.7139  6.2459
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar             3.8887     0.1947  19.969  < 2e-16 ***
## Instrumentpiano              5.4207     0.1948  27.828  < 2e-16 ***
## Instrumentstring             7.3444     0.1947  37.717  < 2e-16 ***
## HarmonyI-V-IV                0.2414     0.2553   0.946   0.3445
## HarmonyI-V-VI                1.2617     0.2557   4.934  8.8e-07 ***
## HarmonyIV-I-V               -0.1667     0.2549  -0.654   0.5133
## Voicepar3rd                 -0.2500     0.2549  -0.981   0.3269
## Voicepar5th                 -0.2102     0.2553  -0.823   0.4105
## HarmonyI-V-IV:Voicepar3rd   -0.4337     0.3608  -1.202   0.2295
## HarmonyI-V-VI:Voicepar3rd   -0.6671     0.3614  -1.846   0.0651 .
## HarmonyIV-I-V:Voicepar3rd    0.5291     0.3608   1.466   0.1427
## HarmonyI-V-IV:Voicepar5th   -0.1898     0.3614  -0.525   0.5995
## HarmonyI-V-VI:Voicepar5th   -0.4458     0.3614  -1.234   0.2175
## HarmonyIV-I-V:Voicepar5th    0.2422     0.3608   0.671   0.5021
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.251 on 1851 degrees of freedom
## Multiple R-squared:  0.8716, Adjusted R-squared:  0.8706
## F-statistic: 897.5 on 14 and 1851 DF,  p-value: < 2.2e-16
```



2.1-1.bb



2.1-2.bb

2 (b) random intercept model for examining Instrument, Harmony & Voice

(i) test whether random intercept is needed or not

We first produce the random intercept model lmer1 based on the best model we find in part 2(a). lmer1 <- lmer(Classical ~ Instrument+Harmony*Voice-1 + (1|Subject),data = data, REML = F, control = lmerControl(optimizer = 'bobyqa')).

the AIC for lmer1 is 7883.7. The BIC for lmer1 is 7972.2.

The AIC for lm2 is 8335.677. The BIC for lm2 is 8418.642.

So based on AIC and BIC, the random intercept model lmer1 is prepfered.

We can also do a random effect test using exactRLRT(lmer1). The test results are significant. So this double confirms that the random intercept model lmer1 is prepfered.

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Instrument + Harmony * Voice - 1 + (1 | Subject)
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7883.7   7972.2  -3925.8   7851.7     1849
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9682 -0.6324 -0.0032  0.6393  3.4125
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.368    1.170
```

```
##  Residual             3.661    1.913
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                            Estimate Std. Error t value
## Instrumentguitar            3.8926     0.2317  16.798
## Instrumentpiano             5.4162     0.2318  23.368
## Instrumentstring            7.3450     0.2317  31.697
## HarmonyI-V-IV               0.2380     0.2170   1.097
## HarmonyI-V-VI               1.2715     0.2174   5.849
## HarmonyIV-I-V              -0.1667     0.2167  -0.769
## Voicepar3rd                -0.2500     0.2167  -1.154
## Voicepar5th                -0.2136     0.2170  -0.984
## HarmonyI-V-IV:Voicepar3rd  -0.4303     0.3066  -1.403
## HarmonyI-V-VI:Voicepar3rd  -0.6803     0.3071  -2.215
## HarmonyIV-I-V:Voicepar3rd   0.5256     0.3066   1.714
## HarmonyI-V-IV:Voicepar5th  -0.1864     0.3071  -0.607
## HarmonyI-V-VI:Voicepar5th  -0.4521     0.3072  -1.472
## HarmonyIV-I-V:Voicepar5th   0.2457     0.3066   0.801
##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
## [1] 8335.677
## [1] 8418.642
## Using restricted likelihood evaluated at ML estimators.
## Refit with method="REML" for exact results.
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
## data:
## RLRT = 452.21, p-value < 2.2e-16
```

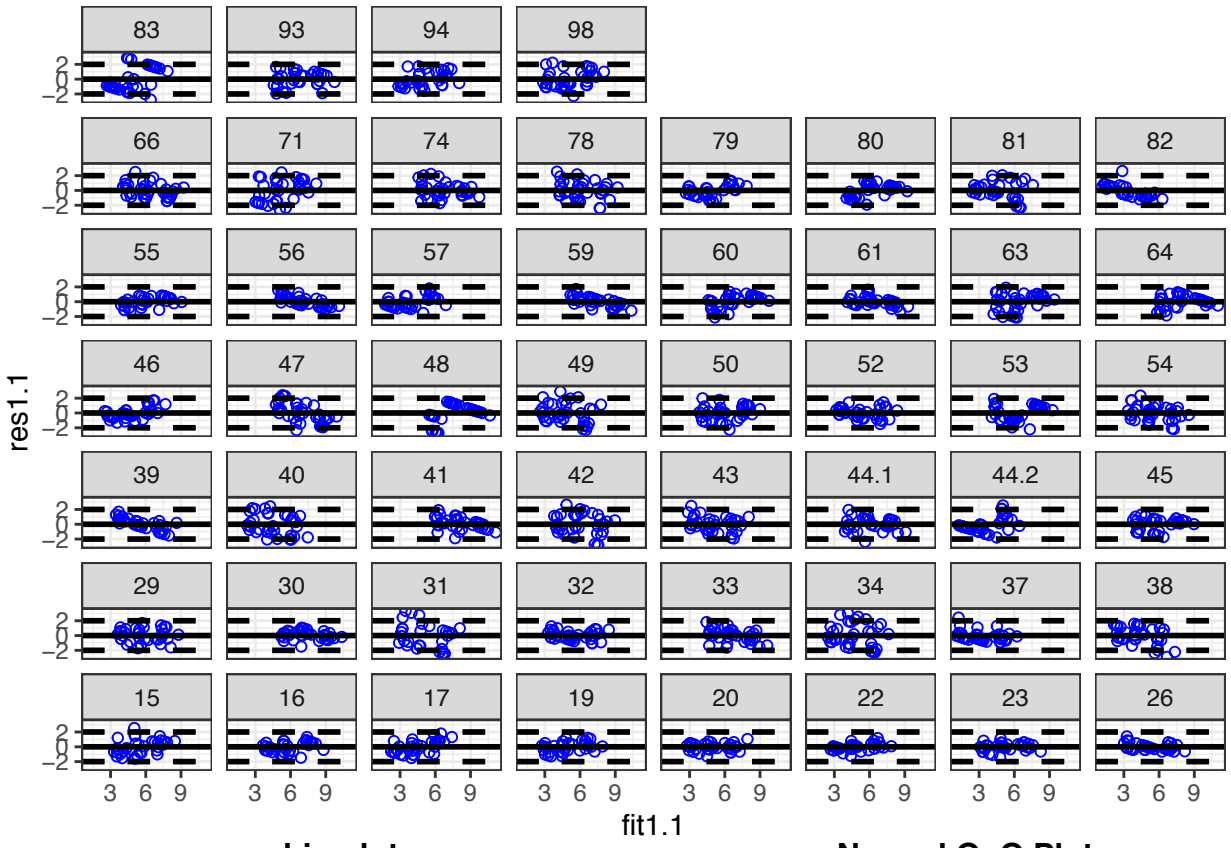
(ii) Validate model lmer1

Looking at the standardlized conditional residuals plots of model lmer1, all the residuals are centered at zero and there is no clear pattern identified. Looking at the Binned residual plot, we find almost all the points lie in the 95% confidence interval for the residulas, except the points at the two ends. And the points randomly scatter around the line at zero.We also notice the normal Q-Q plot almost follows a straight line.

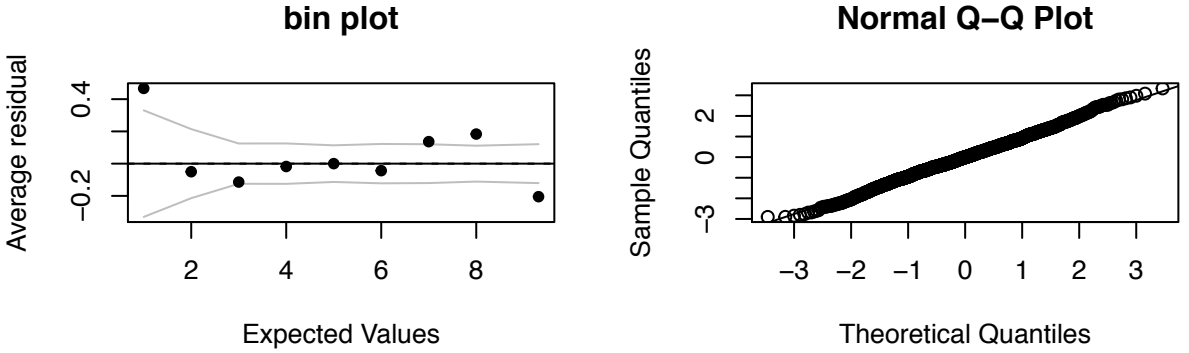
In conclusion, model lmer1 seems valid.

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Instrument + Harmony * Voice - 1 + (1 | Subject)
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7883.7   7972.2  -3925.8   7851.7     1849
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9682 -0.6324 -0.0032  0.6393  3.4125
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.368    1.170
##  Residual             3.661    1.913
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                            Estimate Std. Error t value
## Instrumentguitar            3.8926     0.2317  16.798
## Instrumentpiano             5.4162     0.2318  23.368
## Instrumentstring            7.3450     0.2317  31.697
## HarmonyI-V-IV               0.2380     0.2170   1.097
## HarmonyI-V-VI               1.2715     0.2174   5.849
## HarmonyIV-I-V              -0.1667     0.2167  -0.769
## Voicepar3rd                -0.2500     0.2167  -1.154
## Voicepar5th                -0.2136     0.2170  -0.984
## HarmonyI-V-IV:Voicepar3rd  -0.4303     0.3066  -1.403
## HarmonyI-V-VI:Voicepar3rd  -0.6803     0.3071  -2.215
## HarmonyIV-I-V:Voicepar3rd   0.5256     0.3066   1.714
## HarmonyI-V-IV:Voicepar5th  -0.1864     0.3071  -0.607
## HarmonyI-V-VI:Voicepar5th  -0.4521     0.3072  -1.472
## HarmonyIV-I-V:Voicepar5th   0.2457     0.3066   0.801
##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```



2.3-1.bb



2.3-2.bb

2.c lmer model ___ adding random effects

(i) adding random slope terms

Using anova test and log likelihood ratio test, the final model selected is: lmer3 <- lmer(Classical ~ Instrument+Harmony*Voice-1 + (1+Instrument+Harmony|Subject),data = data, REML = F, control = lmerControl(optimizer = 'bobyqa')).

(ii) Validate model lmer3

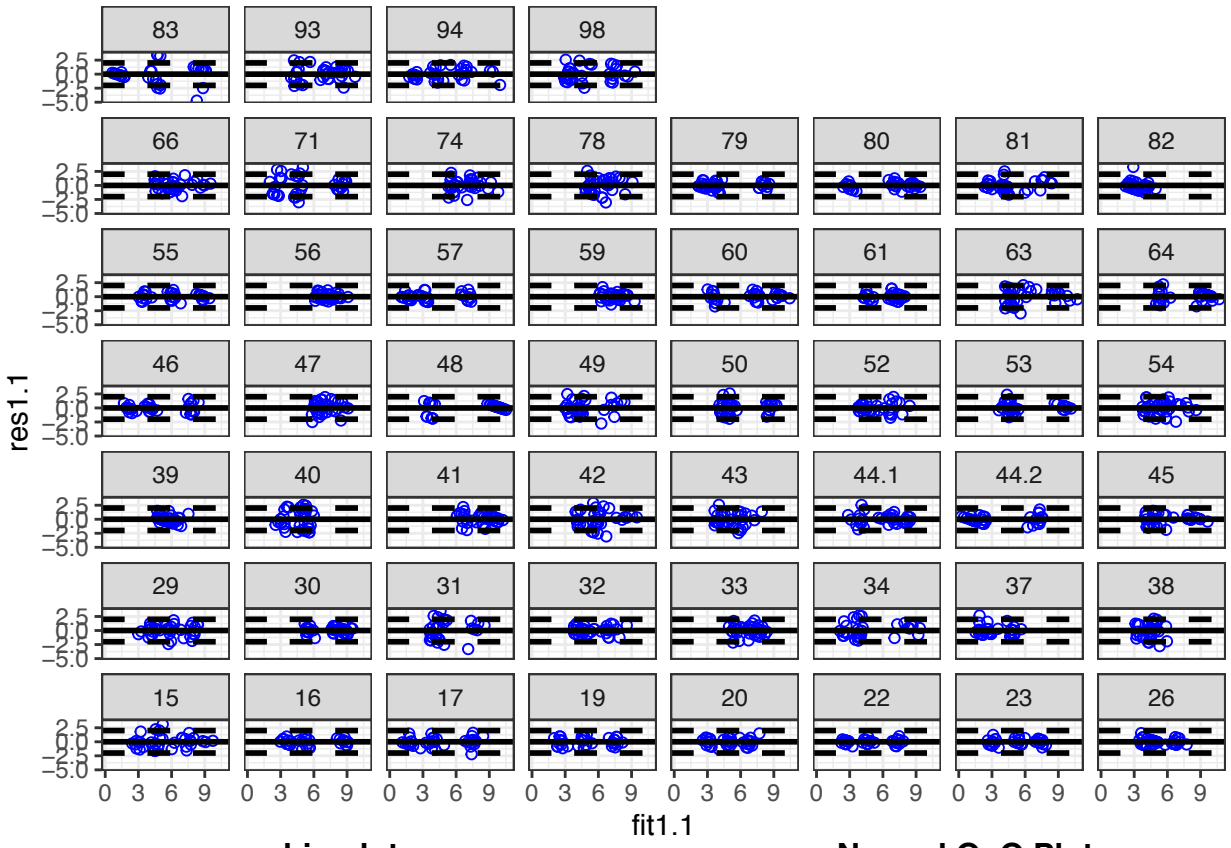
Looking at the standardlized conditional residuals plots of model lmer3, all the residuals are centered at zero and there is no clear pattern identified. Looking at the Binned residual plot, we find almost all the points lie in the 95% confidence interval for the residulas, except the points at the two ends. And the points randomly scatter around the line at zero.We also notice the normal Q-Q plot almost follows a straight line except a long tail on the right.

In conclusion, model lmer3 seems valid.

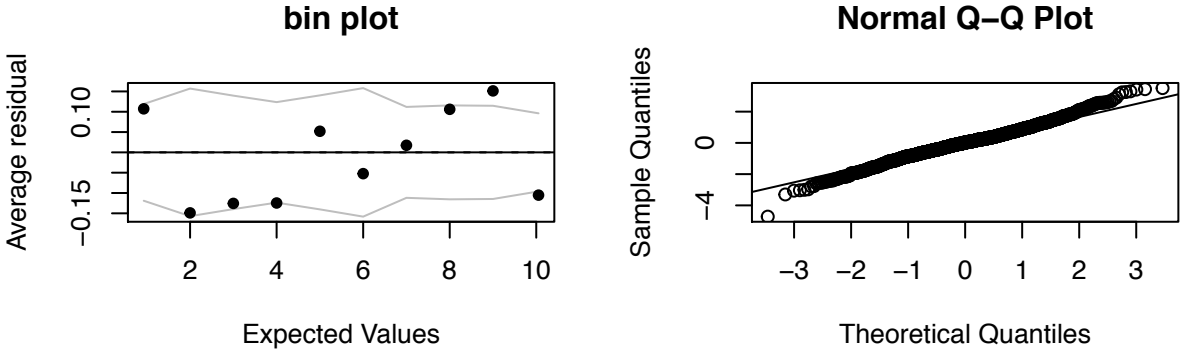
```
## Data: data
## Models:
## lmer1: Classical ~ Instrument + Harmony * Voice - 1 + (1 | Subject)
## lmer2: Classical ~ Instrument + Harmony * Voice - 1 + (1 + Instrument |
## lmer2:     Subject)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1 16 7883.7 7972.2 -3925.8   7851.7
## lmer2 21 7599.1 7715.3 -3778.6   7557.1 294.55      5  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## boundary (singular) fit: see ?isSingular

## Data: data
## Models:
## lmer2: Classical ~ Instrument + Harmony * Voice - 1 + (1 + Instrument |
## lmer2:     Subject)
## lmer3: Classical ~ Instrument + Harmony * Voice - 1 + (1 + Instrument +
## lmer3:     Harmony | Subject)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer2 21 7599.1 7715.3 -3778.6   7557.1
## lmer3 36 7459.3 7658.4 -3693.6   7387.3 169.88     15  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Warning in commonArgs(par, fn, control, environment()): maxfun < 10 *
## length(par)^2 is not recommended.

## boundary (singular) fit: see ?isSingular

## Data: data
## Models:
## lmer3: Classical ~ Instrument + Harmony * Voice - 1 + (1 + Instrument +
## lmer3:     Harmony | Subject)
## lmer4: Classical ~ Instrument + Harmony * Voice - 1 + (1 + Instrument +
## lmer4:     Harmony + Voice | Subject)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer3 36 7459.3 7658.4 -3693.6   7387.3
## lmer4 51 7468.2 7750.3 -3683.1   7366.2 21.065     15     0.1348

## boundary (singular) fit: see ?isSingular

## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 1.205228e-60
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 2.405792e-27
##  adding (Harmony|Subject) to model
## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.5729909
##  not adding (Voice|Subject) to model

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
```

```
## Classical ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) + Harmony:Voice - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7470.2   7647.2  -3703.1   7406.2     1833
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7999 -0.5771  0.0189  0.5509  3.6125
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      2.377e-09 4.875e-05
##  Subject.1 (Intercept)      9.232e-01 9.608e-01
##            Instrumentpiano  1.661e+00 1.289e+00 -0.51
##            Instrumentstring 3.551e+00 1.884e+00 -0.99  0.59
##  Subject.2 (Intercept)      1.300e+00 1.140e+00
##            HarmonyI-V-IV    1.979e-02 1.407e-01  1.00
##            HarmonyI-V-VI    1.793e+00 1.339e+00 -0.37 -0.37
##            HarmonyIV-I-V    9.921e-02 3.150e-01  0.05  0.05  0.07
##  Residual                   2.432e+00 1.560e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                            Estimate Std. Error t value
## Instrumentguitar             3.8892     0.2469  15.753
## Instrumentpiano              5.4184     0.2620  20.684
## Instrumentstring             7.3462     0.2449  30.002
## HarmonyI-V-IV                0.2321     0.1780   1.304
## HarmonyI-V-VI                1.2647     0.2567   4.927
## HarmonyIV-I-V               -0.1667     0.1819  -0.916
## Voicepar3rd                 -0.2500     0.1766  -1.416
## Voicepar5th                 -0.2185     0.1769  -1.235
## HarmonyI-V-IV:Voicepar3rd   -0.4244     0.2499  -1.698
## HarmonyI-V-VI:Voicepar3rd   -0.6620     0.2504  -2.644
## HarmonyIV-I-V:Voicepar3rd    0.5237     0.2500   2.095
## HarmonyI-V-IV:Voicepar5th   -0.1815     0.2504  -0.725
## HarmonyI-V-VI:Voicepar5th   -0.4404     0.2505  -1.758
## HarmonyIV-I-V:Voicepar5th    0.2506     0.2499   1.002
##
## 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
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Instrument + Harmony * Voice - 1 + (1 + Instrument +
##     Harmony | Subject)
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
```

```
##
##      AIC      BIC   logLik deviance df.resid
##   7459.3   7658.4  -3693.6   7387.3     1829
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6767 -0.5736  0.0201  0.5441  3.4730
##
## Random effects:
##  Groups   Name              Variance Std.Dev. Corr
##  Subject  (Intercept)       1.71019  1.3077
##           Instrumentpiano   1.66241  1.2893   -0.24
##           Instrumentstring  3.55547  1.8856   -0.52  0.59
##           HarmonyI-V-IV     0.09253  0.3042    0.83 -0.74 -0.73
##           HarmonyI-V-VI     1.97016  1.4036    0.06 -0.34 -0.58  0.26
##           HarmonyIV-I-V     0.11339  0.3367    0.18 -0.37 -0.25  0.33  0.26
##  Residual                   2.42113  1.5560
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                            Estimate Std. Error t value
## Instrumentguitar            3.8888     0.2258  17.219
## Instrumentpiano             5.4191     0.2602  20.827
## Instrumentstring            7.3459     0.2655  27.668
## HarmonyI-V-IV               0.2334     0.1815   1.286
## HarmonyI-V-VI               1.2650     0.2630   4.810
## HarmonyIV-I-V              -0.1667     0.1823  -0.914
## Voicepar3rd                -0.2500     0.1762  -1.419
## Voicepar5th                -0.2195     0.1765  -1.244
## HarmonyI-V-IV:Voicepar3rd  -0.4257     0.2494  -1.707
## HarmonyI-V-VI:Voicepar3rd  -0.6612     0.2498  -2.647
## HarmonyIV-I-V:Voicepar3rd   0.5237     0.2494   2.100
## HarmonyI-V-IV:Voicepar5th  -0.1805     0.2498  -0.723
## HarmonyI-V-VI:Voicepar5th  -0.4397     0.2499  -1.760
## HarmonyIV-I-V:Voicepar5th   0.2516     0.2494   1.009
##
## 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
```



2.4-1.bb



2.4-2.bb

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Instrument + Harmony * Voice - 1 + (1 + Instrument +
##     Harmony | Subject)
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7459.3   7658.4  -3693.6   7387.3     1829
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6767 -0.5736  0.0201  0.5441  3.4730
##
## Random effects:
##  Groups   Name             Variance Std.Dev. Corr
##  Subject  (Intercept)      1.71019  1.3077
```

```
##           Instrumentpiano  1.66241  1.2893   -0.24
##           Instrumentstring 3.55547  1.8856   -0.52  0.59
##           HarmonyI-V-IV    0.09253  0.3042    0.83 -0.74 -0.73
##           HarmonyI-V-VI    1.97016  1.4036    0.06 -0.34 -0.58  0.26
##           HarmonyIV-I-V    0.11339  0.3367    0.18 -0.37 -0.25  0.33  0.26
##  Residual                  2.42113  1.5560
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                            Estimate Std. Error t value
## Instrumentguitar            3.8888     0.2258  17.219
## Instrumentpiano             5.4191     0.2602  20.827
## Instrumentstring            7.3459     0.2655  27.668
## HarmonyI-V-IV               0.2334     0.1815   1.286
## HarmonyI-V-VI               1.2650     0.2630   4.810
## HarmonyIV-I-V              -0.1667     0.1823  -0.914
## Voicepar3rd                -0.2500     0.1762  -1.419
## Voicepar5th                -0.2195     0.1765  -1.244
## HarmonyI-V-IV:Voicepar3rd  -0.4257     0.2494  -1.707
## HarmonyI-V-VI:Voicepar3rd  -0.6612     0.2498  -2.647
## HarmonyIV-I-V:Voicepar3rd   0.5237     0.2494   2.100
## HarmonyI-V-IV:Voicepar5th  -0.1805     0.2498  -0.723
## HarmonyI-V-VI:Voicepar5th  -0.4397     0.2499  -1.760
## HarmonyIV-I-V:Voicepar5th   0.2516     0.2494   1.009
##
## 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
```

2.d lmer model ___ examing the Pachelbel

(i) examing Pachelbel for fixed effects part;

We know variables KnowRob(Have you heard Rob Paravonian's Pachelbel Rant) and KnowAxis(Have you heard Axis of Evil's Comedy bit on the 4 Pachelbel chords in popular music) are related to Pachelbel. To examine how these two variables affect people identity Harmony and Classical Music, we add these two variables and its interactions with Harmony to the fixed effects part of model lmer3.

Using stepBIC and anova, the best model we found for the fixed effects part are:lm.p2 <- lm(Classical ~ Instrument + Harmony + Voice + Harmony:Voice + KnowAxis-1)

```
## recall lm2 <-
# lm(formula = Classical ~ Instrument + Harmony * Voice - 1, data = data)


## produce a new model lm.p adding pachelbel related variables and
## its interactions with Harmony

lm.p <- lm(Classical~ Instrument+Harmony+Voice+KnowAxis+
KnowRob+Harmony:Voice+KnowAxis:Harmony+KnowRob:Harmony-1, data = data)

summary(lm.p)
```

```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice + KnowAxis +
##     KnowRob + Harmony:Voice + KnowAxis:Harmony + KnowRob:Harmony -
##     1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.5748 -1.5750 -0.0294  1.6069  6.3380
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar             3.7428     0.2012  18.603  < 2e-16 ***
## Instrumentpiano              5.2794     0.2012  26.238  < 2e-16 ***
## Instrumentstring             7.1988     0.2012  35.783  < 2e-16 ***
## HarmonyI-V-IV                0.2955     0.2655   1.113  0.26582
## HarmonyI-V-VI                1.1943     0.2660   4.490 7.56e-06 ***
## HarmonyIV-I-V               -0.1128     0.2652  -0.425  0.67065
## Voicepar3rd                 -0.2500     0.2521  -0.992  0.32141
## Voicepar5th                 -0.2122     0.2525  -0.840  0.40083
## KnowAxis                     0.7866     0.2694   2.920  0.00354 **
## KnowRob                     -0.1643     0.3245  -0.506  0.61270
## HarmonyI-V-IV:Voicepar3rd   -0.4328     0.3568  -1.213  0.22518
## HarmonyI-V-VI:Voicepar3rd   -0.6657     0.3573  -1.863  0.06261 .
## HarmonyIV-I-V:Voicepar3rd    0.5287     0.3568   1.482  0.13855
## HarmonyI-V-IV:Voicepar5th   -0.1878     0.3573  -0.526  0.59918
## HarmonyI-V-VI:Voicepar5th   -0.4532     0.3573  -1.268  0.20488
## HarmonyIV-I-V:Voicepar5th    0.2442     0.3568   0.685  0.49373
## HarmonyI-V-IV:KnowAxis      -0.3316     0.3814  -0.869  0.38474
## HarmonyI-V-VI:KnowAxis      -0.9395     0.3815  -2.463  0.01388 *
## HarmonyIV-I-V:KnowAxis      -0.3755     0.3810  -0.986  0.32443
## HarmonyI-V-IV:KnowRob        0.1127     0.4610   0.244  0.80694
## HarmonyI-V-VI:KnowRob        2.0472     0.4610   4.440 9.51e-06 ***
## HarmonyIV-I-V:KnowRob        0.1899     0.4589   0.414  0.67911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.226 on 1843 degrees of freedom
## Multiple R-squared:  0.875,  Adjusted R-squared:  0.8735
## F-statistic: 586.4 on 22 and 1843 DF,  p-value: < 2.2e-16
```

```
## Using step BIC selecting variables

summary(MASS::stepAIC(lm.p, k = log(nrow(data)),direction = "backward",
                      trace = F))
```

```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + KnowAxis - 1,
##     data = data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.476 -1.629 -0.086  1.807  6.338
##
```

```
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar  3.62926    0.13042  27.829  < 2e-16 ***
## Instrumentpiano   5.16031    0.13068  39.489  < 2e-16 ***
## Instrumentstring  7.08596    0.13039  54.345  < 2e-16 ***
## HarmonyI-V-IV     0.03253    0.14745   0.221    0.825
## HarmonyI-V-VI     0.88879    0.14753   6.024 2.04e-09 ***
## HarmonyIV-I-V     0.08994    0.14737   0.610    0.542
## KnowAxis          0.50080    0.12751   3.928 8.90e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.252 on 1858 degrees of freedom
## Multiple R-squared:  0.871,  Adjusted R-squared:  0.8706
## F-statistic:  1793 on 7 and 1858 DF,  p-value: < 2.2e-16
```

```
## BIC choose model
# lm(formula = Classical ~ Instrument + Harmony + KnowAxis - 1,
#     data = data)

## Add back the main effect Voice;
lm.p3 <- lm(Classical ~ Instrument + Harmony + Voice + KnowAxis-1)

## Add back the interaction of Harmony:Voice;

lm.p2 <- lm(Classical ~ Instrument + Harmony + Voice + Harmony:Voice + KnowAxis-1)


## compare model lm2 vs lm.p2

anova(lm2,lm.p2)
```

```
## Analysis of Variance Table
##
## Model 1: Classical ~ Instrument + Harmony * Voice - 1
## Model 2: Classical ~ Instrument + Harmony + Voice + Harmony:Voice + KnowAxis -
##     1
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1   1851 9382.3
## 2   1850 9304.4  1    77.994 15.508 8.522e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## signficant; the complex model lm.p2 is prefered;

## compare model lm.p3 vs lm.p2
anova(lm.p3,lm.p2)
```

```
## Analysis of Variance Table
##
## Model 1: Classical ~ Instrument + Harmony + Voice + KnowAxis - 1
## Model 2: Classical ~ Instrument + Harmony + Voice + Harmony:Voice + KnowAxis -
##     1
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1   1856 9369.8
```

```
## 2   1850 9304.4  6    65.442 2.1687 0.04335 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## signicant, lm.p2 is prefered;


## final fixed effect  model is :

# lm.p2 <- lm(Classical ~ Instrument + Harmony + Voice + Harmony:Voice +
# KnowAxis-1)

summary(lm.p2)
```

```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice + Harmony:Voice +
##     KnowAxis - 1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.3448 -1.5747 -0.0435  1.6280  6.3514
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar             3.7832     0.1958  19.320  < 2e-16 ***
## Instrumentpiano              5.3143     0.1959  27.127  < 2e-16 ***
## Instrumentstring             7.2390     0.1958  36.971  < 2e-16 ***
## HarmonyI-V-IV                0.2407     0.2543   0.947   0.3440
## HarmonyI-V-VI                1.2604     0.2548   4.947 8.20e-07 ***
## HarmonyIV-I-V               -0.1667     0.2539  -0.656   0.5117
## Voicepar3rd                 -0.2500     0.2539  -0.985   0.3250
## Voicepar5th                 -0.2109     0.2543  -0.829   0.4071
## KnowAxis                     0.5001     0.1270   3.938 8.52e-05 ***
## HarmonyI-V-IV:Voicepar3rd   -0.4330     0.3594  -1.205   0.2284
## HarmonyI-V-VI:Voicepar3rd   -0.6664     0.3600  -1.851   0.0643 .
## HarmonyIV-I-V:Voicepar3rd    0.5284     0.3594   1.470   0.1417
## HarmonyI-V-IV:Voicepar5th   -0.1891     0.3600  -0.525   0.5994
## HarmonyI-V-VI:Voicepar5th   -0.4437     0.3600  -1.233   0.2179
## HarmonyIV-I-V:Voicepar5th    0.2429     0.3594   0.676   0.4992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.243 on 1850 degrees of freedom
## Multiple R-squared:  0.8727, Adjusted R-squared:  0.8716
## F-statistic: 845.2 on 15 and 1850 DF,  p-value: < 2.2e-16
```

(ii) random effects part;

After we added the random effects part, we notice that the t-value of the co-efficeints of KnowAxis is less than 2; So we have to drop the variable KnowAxis; So the final model we get is lmer(Classical ~ Instrument + Harmony + Voice +Harmony:Voice+ (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) -1), which is the same model as lmer3.

So using the mixed effect model, KnowRob and KnowAxis are not in our final lmer model. So we can

conclude that it seems does not matter whether the respondent is familiar with any one of the Pachelbel rants/comedy bit.

```
## produce a lmer model by adding random intercept to the model lm.p2

lmerp <-  lmer(Classical ~ Instrument + Harmony + Voice + Harmony:Voice +
                KnowAxis
              - 1 + (1 |Subject),data = data,
    REML = F, control = lmerControl(optimizer = 'bobyqa'))


lmerp <- ffRanefLMER.fnc(lmerp,ran.effects = c("(Voice|Subject)",
                                               "(Instrument|Subject)",
                                              "(Harmony|Subject)"),log.file=F)
```

```
## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.9999959
##  not adding (Voice|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 1.696393e-60
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model
##  log-likelihood ratio test p-value = 1.526808e-27
##  adding (Harmony|Subject) to model
```

```
summary(lmerp)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + KnowAxis + (1 | Subject) +
##     (Instrument | Subject) + (Harmony | Subject) + Harmony:Voice -      1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7470.5   7653.0  -3702.2   7404.5     1832
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7456 -0.5793  0.0227  0.5499  3.5514
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.004324 0.06576
##  Subject.1 (Intercept)      1.089745 1.04391
##            Instrumentpiano  1.662768 1.28948  -0.47
##            Instrumentstring 3.551790 1.88462  -0.89  0.59
##  Subject.2 (Intercept)      1.013459 1.00671
##            HarmonyI-V-IV    0.053220 0.23069   0.75
##            HarmonyI-V-VI    1.873342 1.36870  -0.41  0.04
##            HarmonyIV-I-V    0.074909 0.27369   0.19 -0.21  0.16
##  Residual                   2.427344 1.55799
## Number of obs: 1865, groups:  Subject, 52
```

```
##
## Fixed effects:
##                            Estimate Std. Error t value
## Instrumentguitar            3.7983     0.2551  14.890
## Instrumentpiano             5.3273     0.2696  19.762
## Instrumentstring            7.2553     0.2571  28.224
## HarmonyI-V-IV               0.2318     0.1796   1.291
## HarmonyI-V-VI               1.2644     0.2596   4.871
## HarmonyIV-I-V              -0.1667     0.1804  -0.924
## Voicepar3rd                -0.2500     0.1764  -1.417
## Voicepar5th                -0.2183     0.1767  -1.235
## KnowAxis                    0.4302     0.3774   1.140
## HarmonyI-V-IV:Voicepar3rd  -0.4241     0.2497  -1.698
## HarmonyI-V-VI:Voicepar3rd  -0.6618     0.2501  -2.646
## HarmonyIV-I-V:Voicepar3rd   0.5238     0.2497   2.098
## HarmonyI-V-IV:Voicepar5th  -0.1817     0.2501  -0.726
## HarmonyI-V-VI:Voicepar5th  -0.4403     0.2502  -1.760
## HarmonyIV-I-V:Voicepar5th   0.2504     0.2497   1.003
##
## Correlation matrix not shown by default, as p = 15 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```
# formula
 # Classical ~ Instrument + Harmony + Voice + + Harmony:Voice+KnowAxis +
# (1 | Subject) +  (Instrument | Subject) + (Harmony | Subject)  -1

# we notice that the t-value of the co-efficeints of KnowAxis is less than 2;
# So we have to drop the variable Know Axis;

# So the final model we get is Classical ~ Instrument + Harmony + Voice +
# Harmony:Voice+
# (1 | Subject) +  (Instrument | Subject) + (Harmony | Subject)  -1

## which is the same model as lmer3.
```

Model for Classsical ___Adding other covariates besides the three main effects

3 (a) Adding individual covariates

From Part 2, we know the fixed effects part we have selected are Classical ~ Instrument + Harmony + Voice + Harmony:Voice - 1

Now, we can add other individual covariates.Based on the analysis result of 1.3,

we know APTheory, ClsListen,CollegeMusic, GuitarPlay doesn't seem to matter for Classical.So we exclude these three variables and add all the other variables as possible explanatory variables to build a full model: lm.full.

Using stepBIC, the final model selected are lm(formula = Classical ~ Instrument + Harmony + Selfdeclare + LogOMSI + X16.minus.17 + Composing + PianoPlay - 1, data = data)

Since the main effects should always be included, so we have to add the Voice back to the model.

So the final model for the fixed effects part are: lm.final <- lm(formula = Classical ~ Instrument + Harmony +Voice+ Selfdeclare + LogOMSI+ X16.minus.17 +Composing + PianoPlay - 1, data = data)

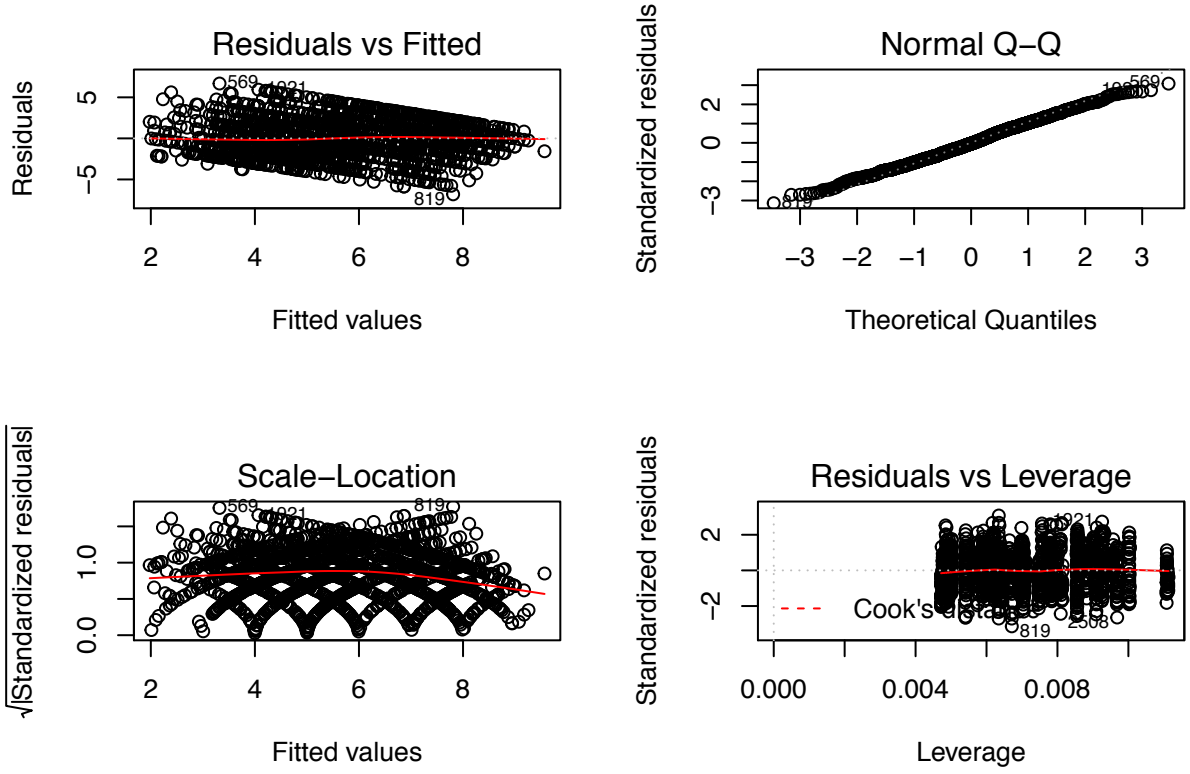
We then examine whether model lm.final is valid or not using the normal Q-Q plot from the diagnostic plots and binnedplot. We notice the normal Q-Q plot almost follows a straight line. Looking at the Binned residual plot, we find almost all the points lie in the 95% confidence interval for the residulas. And the points randomly scatter around the line at zero.

In conclusion, model lm.final seems valid.

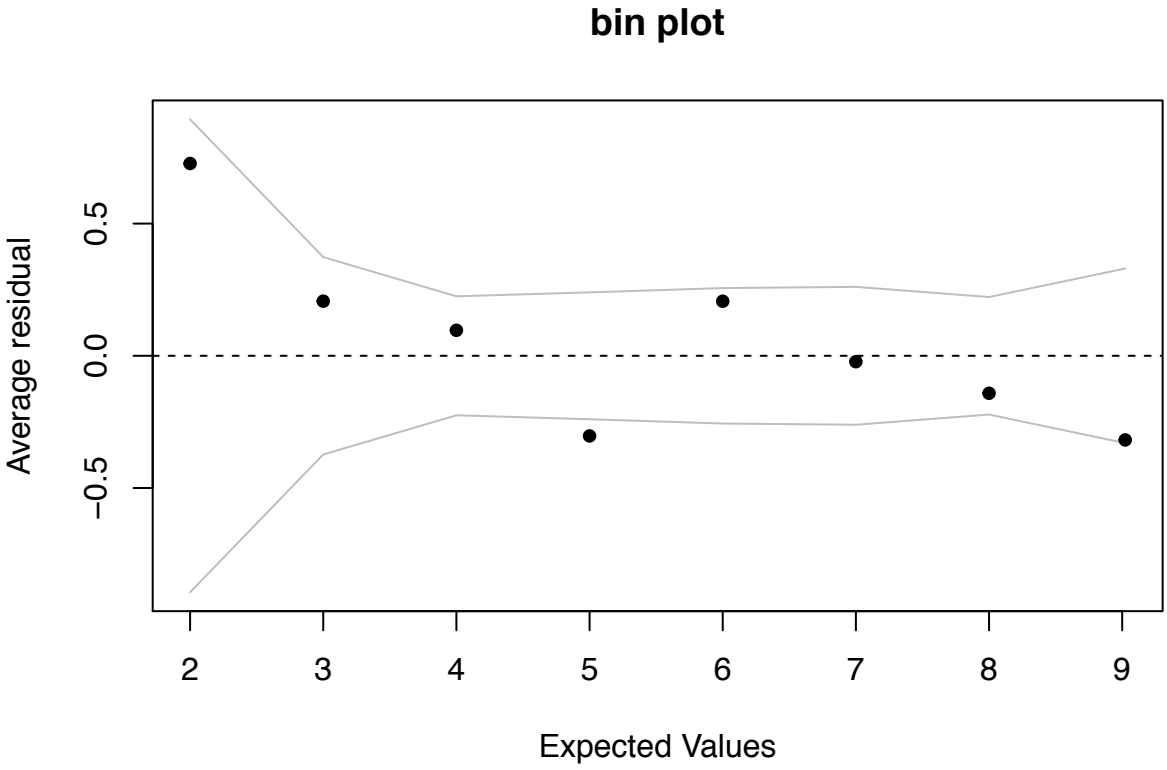
```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Selfdeclare +
##     LogOMSI + X16.minus.17 + Composing + PianoPlay - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8838 -1.6040 -0.0607  1.5927  6.5203
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar  3.36854    0.27032  12.461  < 2e-16 ***
## Instrumentpiano   4.89664    0.26961  18.162  < 2e-16 ***
## Instrumentstring  6.82462    0.27021  25.257  < 2e-16 ***
## HarmonyI-V-IV     0.03201    0.14290   0.224    0.823
## HarmonyI-V-VI     0.88911    0.14297   6.219 6.18e-10 ***
## HarmonyIV-I-V     0.08994    0.14282   0.630    0.529
## Selfdeclare      -0.38185    0.06290  -6.071 1.54e-09 ***
## LogOMSI           0.23405    0.05580   4.194 2.87e-05 ***
## X16.minus.17     -0.12593    0.01777  -7.086 1.95e-12 ***
## Composing         0.25162    0.04745   5.302 1.28e-07 ***
## PianoPlay         0.85062    0.13802   6.163 8.73e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.182 on 1854 degrees of freedom
## Multiple R-squared:  0.8792, Adjusted R-squared:  0.8784
## F-statistic:  1226 on 11 and 1854 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice + Selfdeclare +
##     LogOMSI + X16.minus.17 + Composing + PianoPlay - 1, data = data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.810 -1.568 -0.094  1.604  6.680
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar  3.60295    0.27899  12.914  < 2e-16 ***
## Instrumentpiano   5.13067    0.27827  18.438  < 2e-16 ***
## Instrumentstring  7.05865    0.27885  25.313  < 2e-16 ***
## HarmonyI-V-IV     0.03251    0.14254   0.228  0.81960
## HarmonyI-V-VI     0.88993    0.14262   6.240 5.41e-10 ***
## HarmonyIV-I-V     0.08975    0.14247   0.630  0.52877
## Voicepar3rd      -0.39242    0.12350  -3.178  0.00151 **
## Voicepar5th      -0.30793    0.12350  -2.493  0.01274 *
```

```
## Selfdeclare      -0.38194    0.06274  -6.087 1.39e-09 ***
## LogOMSI           0.23390    0.05566   4.202 2.77e-05 ***
## X16.minus.17     -0.12596    0.01773  -7.105 1.71e-12 ***
## Composing         0.25181    0.04734   5.319 1.17e-07 ***
## PianoPlay         0.85065    0.13768   6.179 7.93e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.177 on 1852 degrees of freedom
## Multiple R-squared:  0.8799, Adjusted R-squared:  0.879
## F-statistic:  1044 on 13 and 1852 DF,  p-value: < 2.2e-16
```



3.1-1.bb



3.1-2.bb

3 (b) check random effects:

Based on the analysis of 3.(a), we know the fixed effects part is:

(formula = Classical ~ Instrument + Harmony +Voice+ Selfdeclare + LogOMSI+ X16.minus.17 +Composing + PianoPlay - 1, data = data)

so we first produce a model as lmer5 <- lmer(Classical ~ Instrument + Harmony +Voice+ Selfdeclare + LogOMSI+ X16.minus.17 +Composing + PianoPlay - 1 + (1 |Subject),data = data, REML = F, control = lmerControl(optimizer = 'bobyqa')).

And then we use ffRanefLMER.fnc method to forward select all the possible random slopes. The final model selected is lmer6: Formula= Classical ~ Instrument + Harmony + Voice + Selfdeclare + LogOMSI + X16.minus.17 + Composing + PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) - 1

By examing the t-value of the coefficients of fixed effects of lmer6, we find that LogOMSI and Composing has the t-value much less than 2. So we can delete this variable LogOMSI and Composing. Thus, we have the lmer model:

lmer6.2 <- lmer(Classical ~ Instrument + Harmony + Voice + Selfdeclare + X16.minus.17 +PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) - 1,data = data,REML = F, control = lmerControl(optimizer = 'bobyqa')). We then notice the Selfdeclare has the t-value much less than 2. So we drop Selfdeclare too.

We have the final model lmer6.2 <- lmer(Classical ~ Instrument + Harmony + Voice + X16.minus.17 +PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) - 1,data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))

We then validate lmer6.2 by examining the conditional residual plots. Looking at the standardlized conditional residuals plots of model lmer6.2, all the residuals are centered at zero and there is no clear pattern identified. Looking at the Binned residual plot, we find almost all the points lie in the 95% confidence

interval for the residulas, except the points at the two ends. And the points randomly scatter around the line at zero.We also notice the normal Q-Q plot almost follows a straight line except a long tail on the right.

In conclusion, model lmer6.2 seems valid.

```
## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.9985862
##  not adding (Voice|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 1.803643e-59
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 3.354355e-27
##  adding (Harmony|Subject) to model

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + Selfdeclare + LogOMSI +
##     X16.minus.17 + Composing + PianoPlay + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7477.8   7649.3  -3707.9   7415.8     1834
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7022 -0.5920  0.0085  0.5495  3.7039
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      2.780e-13 5.273e-07
##  Subject.1 (Intercept)      5.737e-01 7.574e-01
##            Instrumentpiano  1.666e+00 1.291e+00 -0.57
##            Instrumentstring 3.549e+00 1.884e+00 -1.00  0.60
##  Subject.2 (Intercept)      8.700e-01 9.327e-01
##            HarmonyI-V-IV    6.213e-02 2.493e-01  0.93
##            HarmonyI-V-VI    1.874e+00 1.369e+00 -0.31  0.07
##            HarmonyIV-I-V    8.466e-02 2.910e-01  0.04  0.10  0.17
##  Residual                   2.466e+00 1.570e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  3.53819    0.66282   5.338
## Instrumentpiano   5.06643    0.67092   7.552
## Instrumentstring  6.99568    0.67281  10.398
## HarmonyI-V-IV     0.03017    0.10849   0.278
## HarmonyI-V-VI     0.89623    0.21594   4.150
## HarmonyIV-I-V     0.09074    0.11041   0.822
```

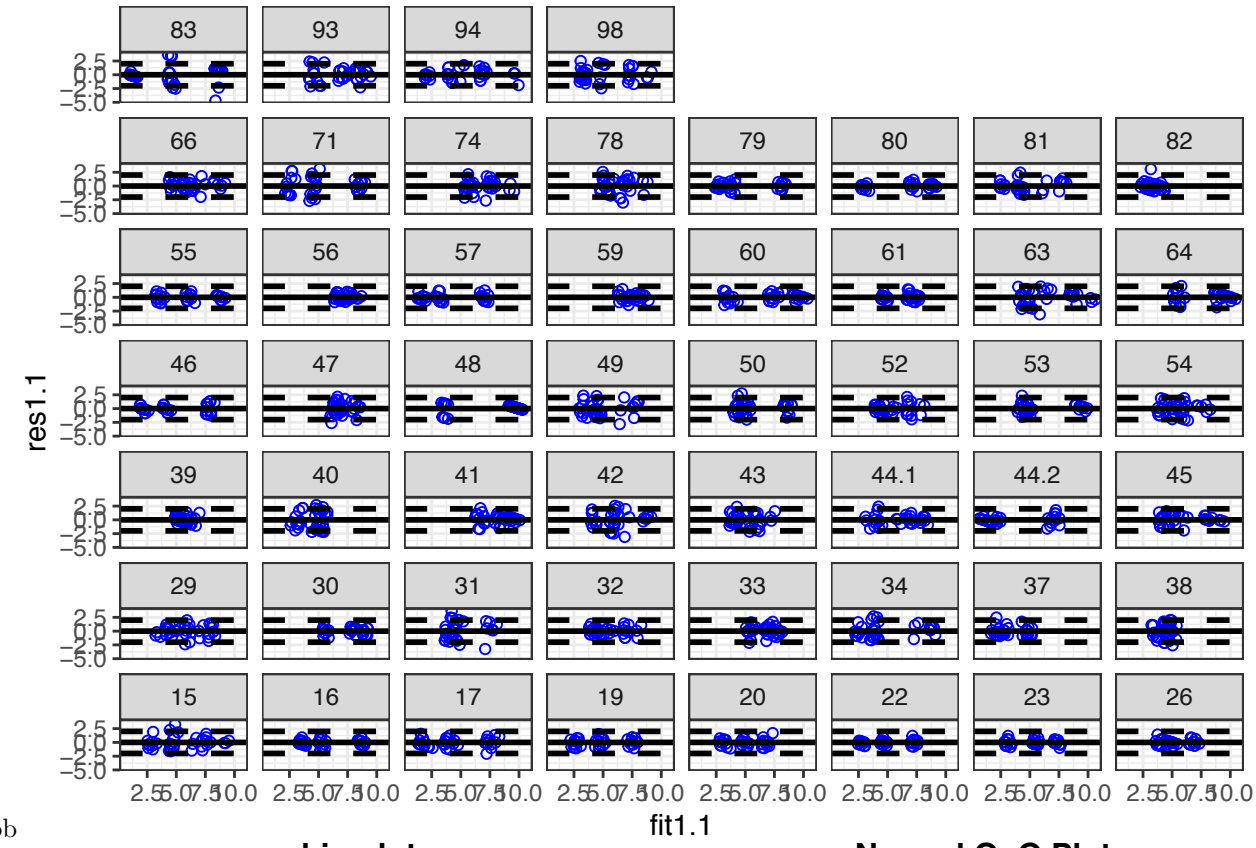
```
## Voicepar3rd      -0.38985    0.08909  -4.376
## Voicepar5th      -0.31024    0.08910  -3.482
## Selfdeclare      -0.36336    0.16640  -2.184
## LogOMSI           0.23294    0.14741   1.580
## X16.minus.17     -0.12384    0.04701  -2.634
## Composing         0.24205    0.12552   1.928
## PianoPlay         0.97496    0.36470   2.673

##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

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

## boundary (singular) fit: see ?isSingular

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + Selfdeclare + X16.minus.17 +
##     PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony |
##     Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7480.8   7641.2  -3711.4   7422.8     1836
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6912 -0.5836  0.0174  0.5495  3.6798
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      1.503e-08 0.0001226
##  Subject.1 (Intercept)      7.367e-01 0.8583173
##            Instrumentpiano  1.667e+00 1.2910624 -0.54
##            Instrumentstring 3.549e+00 1.8838161 -1.00  0.60
##  Subject.2 (Intercept)      1.025e+00 1.0125443
##            HarmonyI-V-IV    5.999e-02 0.2449234  0.87
##            HarmonyI-V-VI    1.870e+00 1.3673733 -0.34  0.06
##            HarmonyIV-I-V    7.152e-02 0.2674375  0.16 -0.09  0.16
##  Residual                   2.466e+00 1.5704744
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  4.09202    0.38548  10.615
## Instrumentpiano   5.62022    0.39734  14.145
## Instrumentstring  7.54954    0.39344  19.188
## HarmonyI-V-IV     0.02991    0.10830   0.276
## HarmonyI-V-VI     0.89611    0.21575   4.153
## HarmonyIV-I-V     0.09072    0.10926   0.830
## Voicepar3rd      -0.38972    0.08910  -4.374
## Voicepar5th      -0.31015    0.08910  -3.481
```

```
## Selfdeclare      -0.04934    0.13154  -0.375
## X16.minus.17     -0.10865    0.04998  -2.174
## PianoPlay         0.95716    0.37958   2.522
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.870
## Instrmntstr  0.749     0.832
## HrmnyI-V-IV -0.026    -0.025    -0.026
## HrmnyI-V-VI -0.172    -0.167    -0.169     0.241
## HrmnyIV-I-V -0.106    -0.103    -0.104     0.437  0.272
## Voicepar3rd -0.116    -0.113    -0.114    -0.001 -0.001  0.001
## Voicepar5th -0.116    -0.112    -0.113    -0.001 -0.002 -0.001  0.500
## Selfdeclare -0.795    -0.772    -0.779    -0.001  0.000  0.000  0.000
## X16.mins.17 -0.013    -0.013    -0.013    -0.001  0.000  0.000  0.000
## PianoPlay    0.062     0.060     0.061     0.001  0.000  0.000  0.000
##             Vcpr5t Slfdcl X16..1
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th
## Selfdeclare  0.000
## X16.mins.17  0.000 -0.215
## PianoPlay    0.000 -0.268 -0.057
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

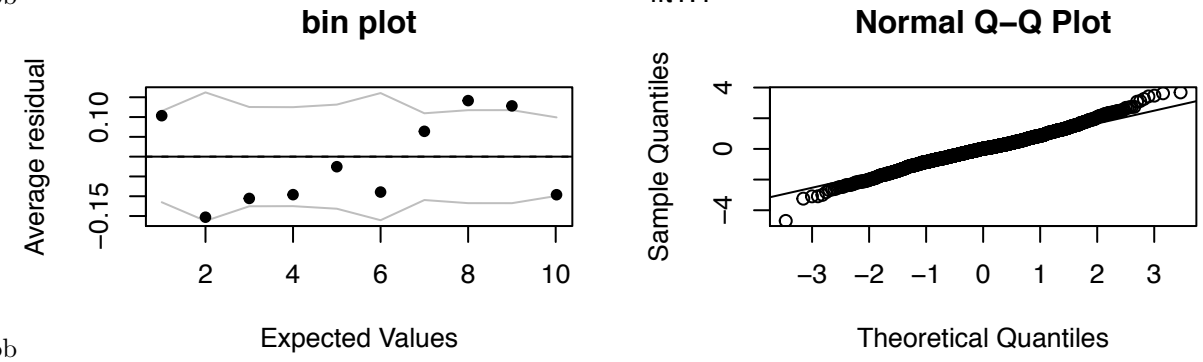
```
## boundary (singular) fit: see ?isSingular
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + X16.minus.17 + PianoPlay +
##     (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) -      1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7478.9   7633.8  -3711.5   7422.9     1837
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6926 -0.5842  0.0166  0.5449  3.6822
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      5.983e-08 0.0002446
##  Subject.1 (Intercept)      7.395e-01 0.8599395
##            Instrumentpiano  1.667e+00 1.2910689 -0.54
##            Instrumentstring 3.549e+00 1.8838126 -1.00  0.60
##  Subject.2 (Intercept)      1.053e+00 1.0263711
##            HarmonyI-V-IV    5.753e-02 0.2398589  0.87
##            HarmonyI-V-VI    1.869e+00 1.3670010 -0.36  0.05
```

```
##            HarmonyIV-I-V    6.998e-02 0.2645403  0.11 -0.12  0.16
##  Residual                   2.467e+00 1.5706033
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  3.97494    0.23499  16.916
## Instrumentpiano   5.50315    0.25382  21.681
## Instrumentstring  7.43247    0.24761  30.017
## HarmonyI-V-IV     0.02992    0.10809   0.277
## HarmonyI-V-VI     0.89608    0.21571   4.154
## HarmonyIV-I-V     0.09070    0.10914   0.831
## Voicepar3rd      -0.38972    0.08910  -4.374
## Voicepar5th      -0.31014    0.08911  -3.480
## X16.minus.17     -0.11195    0.04886  -2.291
## PianoPlay         0.92342    0.36601   2.523
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.667
## Instrmntstr  0.347     0.583
## HrmnyI-V-IV -0.047    -0.043    -0.044
## HrmnyI-V-VI -0.296    -0.274    -0.281     0.241
## HrmnyIV-I-V -0.183    -0.169    -0.173     0.435  0.271
## Voicepar3rd -0.190    -0.176    -0.180    -0.001 -0.001  0.001
## Voicepar5th -0.190    -0.175    -0.180    -0.001 -0.002 -0.001  0.500
## X16.mins.17 -0.309    -0.286    -0.294    -0.001  0.000  0.000  0.000
## PianoPlay   -0.258    -0.238    -0.244     0.001  0.000  0.000  0.000
##             Vcpr5t X16..1
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th
## X16.mins.17  0.000
## PianoPlay    0.000 -0.122
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```



3.2-1.bb



3.2-2.bb

4. Model for Classical – Musicans vs Non-Musicans

4.1 dichotomize "Selfdeclare" at 2

We firt produced a table summarising the 52 subjects(listeners) belong to each Selfdeclare score.

Selfdeclare: 1 2 3 4 5 6 number of subjects: 7 23 9 10 2 1

We can first dichotomize "Selfdeclare" at 2. Have a "Selfdeclare" score less and equal than 2 are categorized as "Non-Musicans". Have a "Selfdeclare" score bigger than 2 are categorized as "Musicans". We created a new column called Musicans, with value of "Non-Musicans" and "Musicans". So we have 30 Non-Musicans and 22 Musicans.

Recall that from part 3, we have the final model lmer6.2 as

lmer6.2 <- lmer(Classical ~ Instrument + Harmony + Voice +
X16.minus.17 + PianoPlay + (1 | Subject) + (Instrument |

Subject) + (Harmony | Subject) - 1,data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))

We can first examine the fixed effect part. Add the variable Musicans and add interaction between Musicans and all the other variables. So we have a full model:

lm.full <- lm(Classical ~ Instrument + Harmony + Voice + Musicans +
X16.minus.17 + PianoPlay+ Musicans:Instrument + Musicans:Harmony + Musicans:Voice + Musicans:X16.minus.17+Musicans:PianoPlay -1, data = data)

Using step AIC, the model selected is:

lm(formula = Classical ~ Instrument + Harmony + Voice + Musicans + X16.minus.17 + Composing + PianoPlay + Instrument:Musicans + Harmony:Musicans + Musicans:X16.minus.17 + Musicans:Composing - 1, data = data)

So the interactions of Musicans and Harmony, the interactions of Musicans and instrument, the interactions of Musicans and X16.minus.17, and the interacations of Musicans and Composing is kept.

We can add the random intercept and produce a model as

lmer7 <- lmer(Classical ~ Instrument + Harmony + Voice + Musicans + X16.minus.17 + Composing + PianoPlay + Instrument:Musicans + Harmony:Musicans + Musicans:X16.minus.17 + Musicans:Composing - 1 + (1 |Subject),data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

And then we use ffRanefLMER.fnc method to forward select all the possible random slopes.

The model selected is lmer8: Formula: Classical ~ Instrument + Harmony + Voice + Musicans + X16.minus.17 +
Composing + PianoPlay + (1 | Subject) + (Instrument | Subject) +
(Harmony | Subject) + Instrument:Musicans + Harmony:Musicans +
Musicans:X16.minus.17 + Musicans:Composing - 1

Examining the t-value of the fix effects of lmer8, we notice that the co-efficients of Musicans, Composing, Instrument:Musicans, Musicans:X16.minus.17, and Musicans:Composing all have t-value much less than 2. So we can delete these variables from the model.

Thus, the final model we get is lmer8.2:

lmer8.2 <- lmer(Classical ~ Instrument + Harmony + Voice + X16.minus.17 + PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony |
Subject) + Harmony:Musicans - 1,data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

Based on the summmary of the lmer8.2, the interaction between HarmonyI-V-VI and MusicansNon-Musicans is significant, with the co-efficient equals -1.03627. This means that if the subject (listener) is non-Musicans and the harmony is I-V-VI, an average the Classical rating will additionally decrease -1.03627.

The validation of model lmer8.2 has been confirmed by examining standardized conditional resisuals plots and binned plots.

```
##
##  1  2  3  4  5  6
##  7 23  9 10  2  1
```

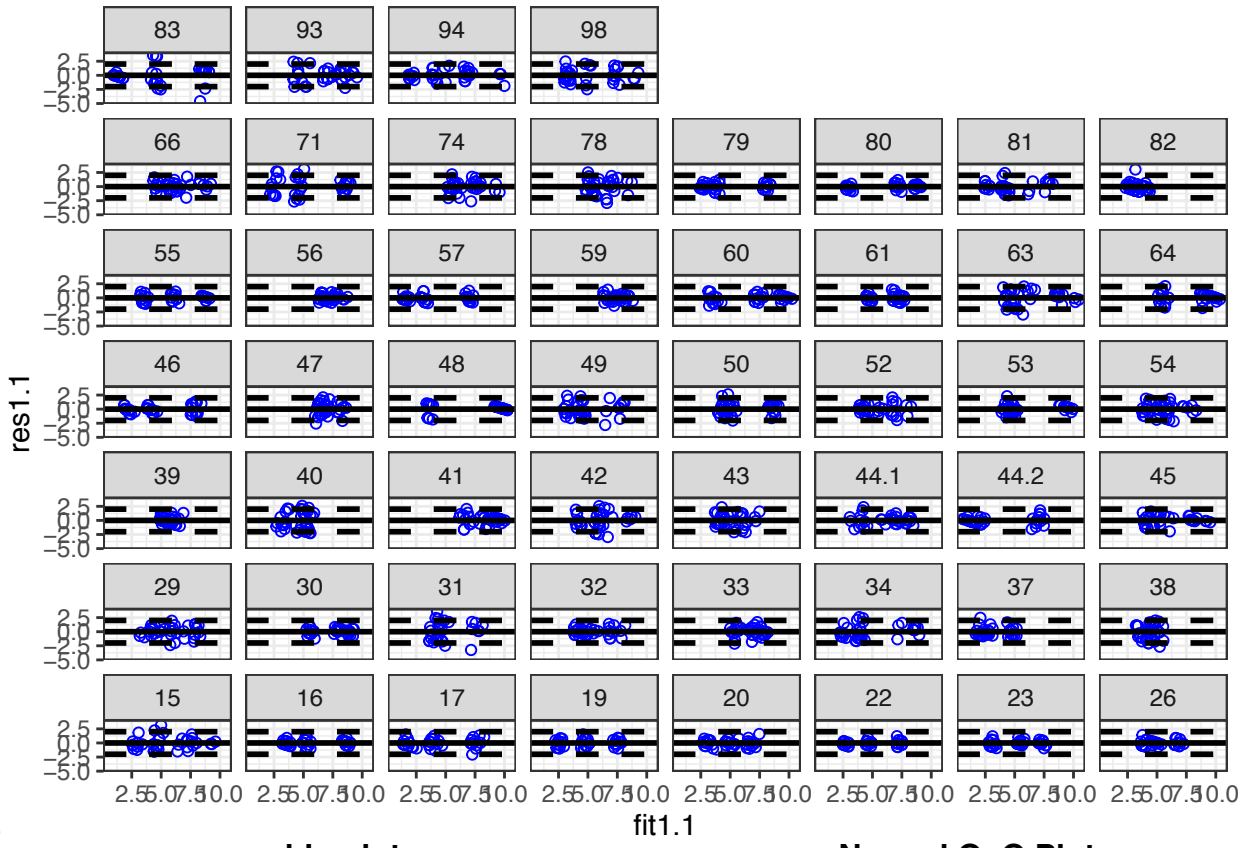
```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice + Musicans +
##     X16.minus.17 + PianoPlay + Instrument:Musicans + Harmony:Musicans +
##     Musicans:X16.minus.17 - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.3330 -1.4940 -0.0781  1.5562  6.6702
```

```
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar                      4.50989    0.22123  20.386  < 2e-16
## Instrumentpiano                       5.75635    0.22181  25.951  < 2e-16
## Instrumentstring                      7.46822    0.22123  33.758  < 2e-16
## HarmonyI-V-IV                         0.06047    0.22034   0.274 0.783790
## HarmonyI-V-VI                         1.52943    0.22035   6.941 5.37e-12
## HarmonyIV-I-V                         0.17217    0.22006   0.782 0.434112
## Voicepar3rd                          -0.39094    0.12390  -3.155 0.001629
## Voicepar5th                          -0.30757    0.12390  -2.482 0.013137
## MusicansNon-Musicans                 -0.65469    0.26431  -2.477 0.013341
## X16.minus.17                         -0.22618    0.02846  -7.947 3.30e-15
## PianoPlay                             0.71005    0.13094   5.423 6.65e-08
## Instrumentpiano:MusicansNon-Musicans  0.48889    0.25119   1.946 0.051774
## Instrumentstring:MusicansNon-Musicans 0.86451    0.25035   3.453 0.000566
## HarmonyI-V-IV:MusicansNon-Musicans   -0.04936    0.28963  -0.170 0.864705
## HarmonyI-V-VI:MusicansNon-Musicans   -1.10804    0.28974  -3.824 0.000136
## HarmonyIV-I-V:MusicansNon-Musicans   -0.14254    0.28942  -0.492 0.622429
## MusicansNon-Musicans:X16.minus.17     0.15281    0.03694   4.136 3.69e-05
##
## Instrumentguitar                     ***
## Instrumentpiano                      ***
## Instrumentstring                     ***
## HarmonyI-V-IV
## HarmonyI-V-VI                        ***
## HarmonyIV-I-V
## Voicepar3rd                          **
## Voicepar5th                          *
## MusicansNon-Musicans                 *
## X16.minus.17                         ***
## PianoPlay                            ***
## Instrumentpiano:MusicansNon-Musicans .
## Instrumentstring:MusicansNon-Musicans ***
## HarmonyI-V-IV:MusicansNon-Musicans
## HarmonyI-V-VI:MusicansNon-Musicans   ***
## HarmonyIV-I-V:MusicansNon-Musicans
## MusicansNon-Musicans:X16.minus.17    ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.184 on 1848 degrees of freedom
## Multiple R-squared:  0.8794, Adjusted R-squared:  0.8782
## F-statistic: 792.4 on 17 and 1848 DF,  p-value: < 2.2e-16

## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.9998618
##  not adding (Voice|Subject) to model
## evaluating addition of (Instrument|Subject) to model

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?
```

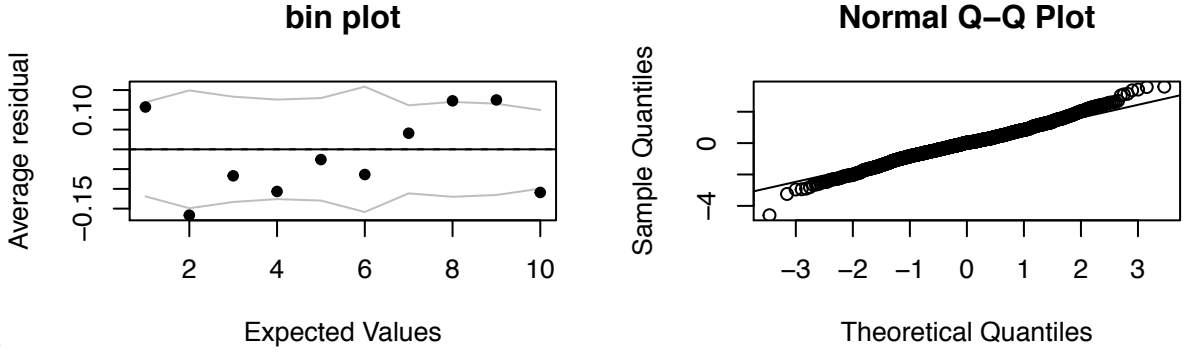
```
##  log-likelihood ratio test p-value = 2.156036e-58
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 1.241853e-21
##  adding (Harmony|Subject) to model

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + Musicans + X16.minus.17 +
##     PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony |
##     Subject) + Instrument:Musicans + Harmony:Musicans + Musicans:X16.minus.17 -
##     1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7481.9   7675.5  -3705.9   7411.9     1830
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6481 -0.5799  0.0165  0.5453  3.6801
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      4.037e-09 6.354e-05
##  Subject.1 (Intercept)      6.643e-01 8.151e-01
##            Instrumentpiano  1.610e+00 1.269e+00 -0.52
##            Instrumentstring 3.367e+00 1.835e+00 -1.00  0.58
##  Subject.2 (Intercept)      1.026e+00 1.013e+00
##            HarmonyI-V-IV    5.225e-02 2.286e-01  0.86
##            HarmonyI-V-VI    1.559e+00 1.249e+00 -0.36  0.01
##            HarmonyIV-I-V    6.278e-02 2.506e-01  0.14 -0.22  0.05
##  Residual                   2.467e+00 1.571e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                      Estimate Std. Error t value
## Instrumentguitar                      4.26373    0.39745  10.728
## Instrumentpiano                       5.50814    0.42762  12.881
## Instrumentstring                      7.22207    0.41884  17.243
## HarmonyI-V-IV                         0.05535    0.16583   0.334
## HarmonyI-V-VI                         1.54493    0.30988   4.986
## HarmonyIV-I-V                         0.17404    0.16705   1.042
## Voicepar3rd                          -0.38954    0.08911  -4.371
## Voicepar5th                          -0.31004    0.08912  -3.479
## MusicansNon-Musicans                 -0.40207    0.47053  -0.855
## X16.minus.17                         -0.15375    0.07979  -1.927
## PianoPlay                             0.86639    0.36705   2.360
## Instrumentpiano:MusicansNon-Musicans  0.49195    0.39952   1.231
## Instrumentstring:MusicansNon-Musicans 0.86552    0.54560   1.586
## HarmonyI-V-IV:MusicansNon-Musicans   -0.04424    0.21798  -0.203
## HarmonyI-V-VI:MusicansNon-Musicans   -1.12489    0.40781  -2.758
```

```
## HarmonyIV-I-V:MusicansNon-Musicans    -0.14441    0.21971  -0.657
## MusicansNon-Musicans:X16.minus.17      0.04789    0.10360   0.462

##
## Correlation matrix not shown by default, as p = 17 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

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

## boundary (singular) fit: see ?isSingular

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + PianoPlay + (1 | Subject) +
##     (Instrument | Subject) + (Harmony | Subject) + Harmony:Musicans -
##     1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7481.9   7653.3  -3709.9   7419.9     1834
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6572 -0.5730  0.0197  0.5456  3.6444
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      4.686e-08 0.0002165
##  Subject.1 (Intercept)      8.410e-01 0.9170649
##            Instrumentpiano  1.668e+00 1.2916508 -0.50
##            Instrumentstring 3.550e+00 1.8840501 -0.99  0.60
##  Subject.2 (Intercept)      1.185e+00 1.0887098
##            HarmonyI-V-IV    5.701e-02 0.2387610  0.79
##            HarmonyI-V-VI    1.561e+00 1.2493991 -0.41  0.01
##            HarmonyIV-I-V    6.647e-02 0.2578256  0.30 -0.15  0.06
##  Residual                   2.466e+00 1.5705084
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                    Estimate Std. Error t value
## Instrumentguitar                     3.6007     0.3163  11.383
## Instrumentpiano                      5.1291     0.3310  15.496
## Instrumentstring                     7.0584     0.3199  22.064
## HarmonyI-V-IV                        0.0549     0.1664   0.330
## HarmonyI-V-VI                        1.5450     0.3100   4.984
## HarmonyIV-I-V                        0.1740     0.1675   1.039
## Voicepar3rd                         -0.3895     0.0891  -4.371
## Voicepar5th                         -0.3100     0.0891  -3.479
## PianoPlay                            0.8488     0.3844   2.208
## HarmonyI-IV-V:MusicansNon-Musicans   0.3504     0.3487   1.005
## HarmonyI-V-IV:MusicansNon-Musicans   0.3066     0.3981   0.770
## HarmonyI-V-VI:MusicansNon-Musicans  -0.7745     0.3957  -1.957
## HarmonyIV-I-V:MusicansNon-Musicans   0.2060     0.3741   0.551
```

```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

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



4.1-1.bb



4.1-2.bb

4.2 dichotomize "Selfdeclare" at 3

recall the frequency table is shown as below:

Selfdeclare: 1 2 3 4 5 6 number of subjects: 7 23 9 10 2 1

We can dichotomize "Selfdeclare" at 3. Have a "Selfdeclare" score less and equal than 3 are categorized as "Non-Musicans". Have a "Selfdeclare" score bigger than 3 are categorized as "Musicans".We created a new column called Musicans, with value of "Non-Musicans" and "Musicans". So we have 39 Non-Musicans and 13 Musicans.

Using this new defination, the final model we get is lmer10.2:

lmer10.2 <- lmer(Classical ~ Instrument + Harmony + Voice + PianoPlay + X16.minus.17+ (1 | Subject) + (Instrument | Subject) +
(Harmony | Subject) + Harmony:Musicans - 1,data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

The validation of model lmer10.2 has been confirmed by examining standardized conditional resisuals plots and binned plots.

Model lmer10.2 is the same as the model we get if we dichotomize "Selfdeclare" at 2.Based on the summmary of the lmer10.2, the interaction between HarmonyI-V-VI and MusicansNon-Musicans is significant, with the co-efficient equals -1.04169, which is almost no difference from the value of -1.03627 of model lmer8.2;

This means that if the subject (listener) is non-Musicans and the harmony is I-V-VI, an average the Classical rating will additionally decrease 1.04169.

So far, it seems that the results are not sensitive to where we dichotomize. We get the same model in the above two cases. And the interaction between HarmonyI-V-VI and MusicansNon-Musicans stays significant and its co-efficient almost stays the same. Generally, if the subject (listener) is non-Musicans and the harmony is I-V-VI, an average the Classical rating will additionally decrease.

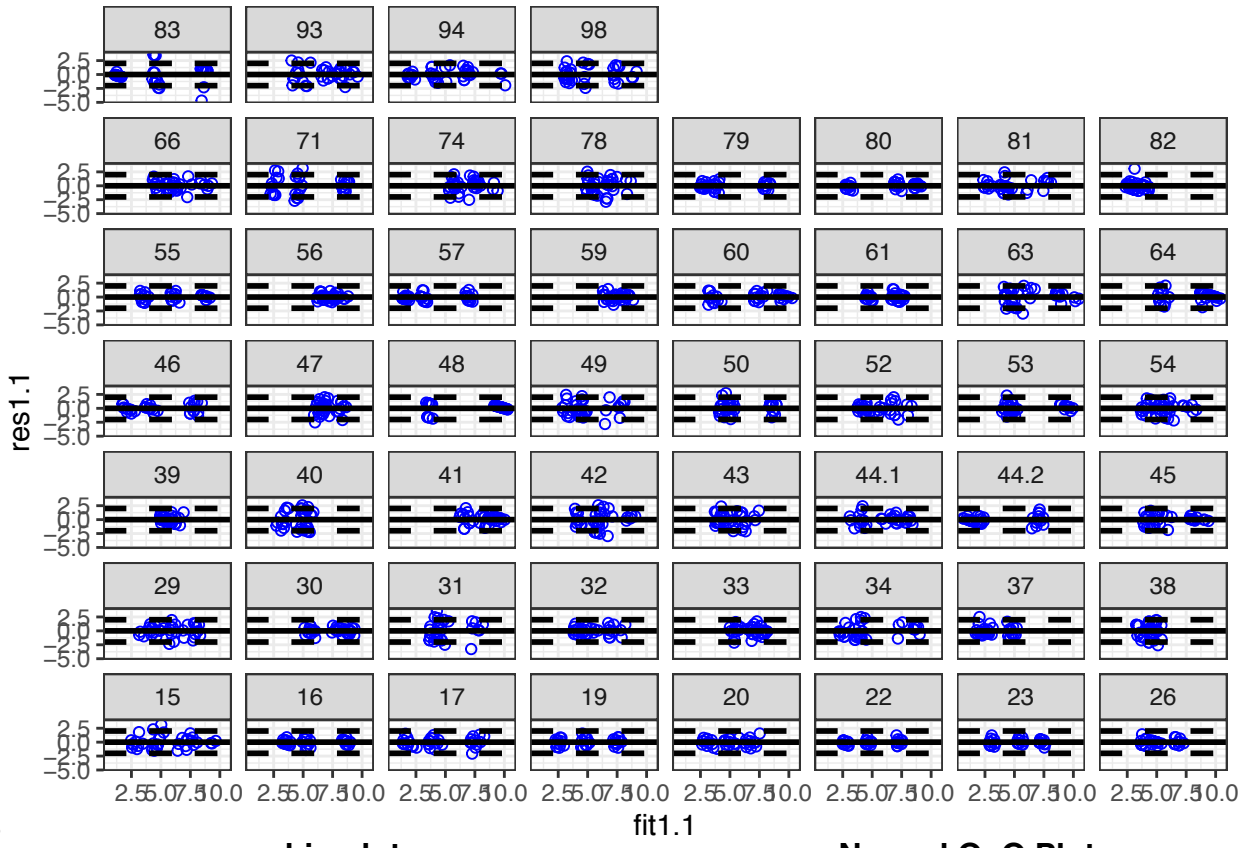
```
##
##  1  2  3  4  5  6
##  7 23  9 10  2  1
```

```
##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice + Musicans +
##     X16.minus.17 + PianoPlay + Instrument:Musicans + Harmony:Musicans +
##     Musicans:X16.minus.17 - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.3026 -1.5244 -0.0637  1.5689  6.6475
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar                      4.69231    0.30204  15.535  < 2e-16
## Instrumentpiano                       5.88637    0.30194  19.495  < 2e-16
## Instrumentstring                      7.28846    0.30204  24.131  < 2e-16
## HarmonyI-V-IV                         0.10898    0.28748   0.379 0.704661
## HarmonyI-V-VI                         1.92566    0.28749   6.698 2.79e-11
## HarmonyIV-I-V                         0.40446    0.28686   1.410 0.158716
## Voicepar3rd                          -0.39102    0.12393  -3.155 0.001630
## Voicepar5th                          -0.30858    0.12393  -2.490 0.012862
## MusicansNon-Musicans                 -0.72690    0.32481  -2.238 0.025344
## X16.minus.17                         -0.33831    0.06320  -5.353 9.71e-08
## PianoPlay                             0.66645    0.13620   4.893 1.08e-06
## Instrumentpiano:MusicansNon-Musicans  0.44373    0.28784   1.542 0.123343
## Instrumentstring:MusicansNon-Musicans 1.14741    0.28567   4.017 6.14e-05
## HarmonyI-V-IV:MusicansNon-Musicans   -0.10329    0.33142  -0.312 0.755347
## HarmonyI-V-VI:MusicansNon-Musicans   -1.37861    0.33148  -4.159 3.34e-05
## HarmonyIV-I-V:MusicansNon-Musicans   -0.41871    0.33088  -1.265 0.205875
## MusicansNon-Musicans:X16.minus.17     0.23085    0.06605   3.495 0.000485
##
## Instrumentguitar                     ***
```

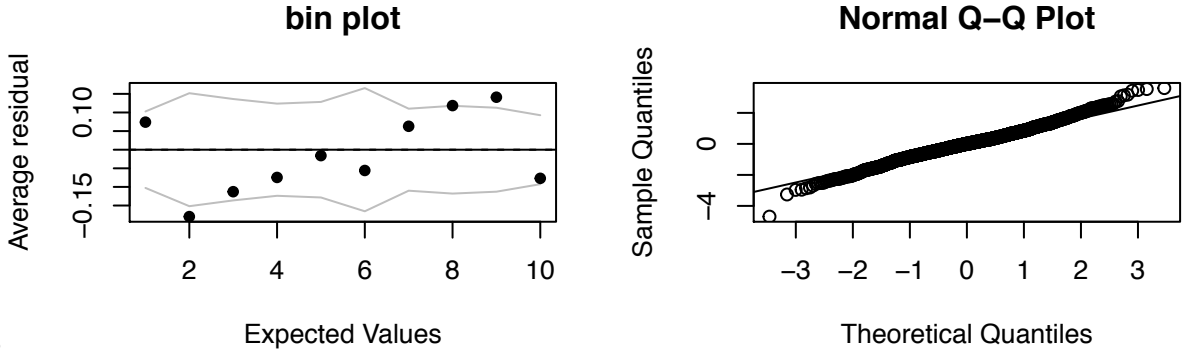
```
## Instrumentpiano                        ***
## Instrumentstring                       ***
## HarmonyI-V-IV
## HarmonyI-V-VI                          ***
## HarmonyIV-I-V
## Voicepar3rd                            **
## Voicepar5th                            *
## MusicansNon-Musicans                   *
## X16.minus.17                           ***
## PianoPlay                              ***
## Instrumentpiano:MusicansNon-Musicans
## Instrumentstring:MusicansNon-Musicans ***
## HarmonyI-V-IV:MusicansNon-Musicans
## HarmonyI-V-VI:MusicansNon-Musicans    ***
## HarmonyIV-I-V:MusicansNon-Musicans
## MusicansNon-Musicans:X16.minus.17     ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.185 on 1848 degrees of freedom
## Multiple R-squared:  0.8793, Adjusted R-squared:  0.8782
## F-statistic: 791.9 on 17 and 1848 DF,  p-value: < 2.2e-16

## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.9995236
##  not adding (Voice|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 2.52896e-57
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 1.017258e-21
##  adding (Harmony|Subject) to model

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + Musicans + X16.minus.17 +
##     PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony |
##     Subject) + Instrument:Musicans + Harmony:Musicans + Musicans:X16.minus.17 -
##     1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7478.7   7672.3  -3704.4   7408.7     1830
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7074 -0.5727  0.0064  0.5534  3.6441
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
```

```
##  Subject   (Intercept)      4.946e-09 7.032e-05
##  Subject.1 (Intercept)      6.641e-01 8.149e-01
##            Instrumentpiano  1.631e+00 1.277e+00 -0.54
##            Instrumentstring 3.302e+00 1.817e+00 -1.00  0.58
##  Subject.2 (Intercept)      1.008e+00 1.004e+00
##            HarmonyI-V-IV    5.934e-02 2.436e-01  0.85
##            HarmonyI-V-VI    1.504e+00 1.226e+00 -0.32 -0.02
##            HarmonyIV-I-V    4.286e-02 2.070e-01  0.29 -0.21 -0.19
##  Residual                   2.466e+00 1.570e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                       Estimate Std. Error t value
## Instrumentguitar                       4.32194    0.59715   7.238
## Instrumentpiano                        5.51843    0.63025   8.756
## Instrumentstring                       6.91809    0.61901  11.176
## HarmonyI-V-IV                          0.10275    0.21747   0.472
## HarmonyI-V-VI                          1.94858    0.39812   4.894
## HarmonyIV-I-V                          0.40622    0.21404   1.898
## Voicepar3rd                           -0.38956    0.08909  -4.373
## Voicepar5th                           -0.31018    0.08909  -3.482
## MusicansNon-Musicans                  -0.38117    0.63150  -0.604
## X16.minus.17                          -0.20333    0.17711  -1.148
## PianoPlay                              0.83464    0.38390   2.174
## Instrumentpiano:MusicansNon-Musicans   0.44251    0.45883   0.964
## Instrumentstring:MusicansNon-Musicans  1.14862    0.61712   1.861
## HarmonyI-V-IV:MusicansNon-Musicans    -0.09705    0.25073  -0.387
## HarmonyI-V-VI:MusicansNon-Musicans    -1.40315    0.45949  -3.054
## HarmonyIV-I-V:MusicansNon-Musicans    -0.42046    0.24691  -1.703
## MusicansNon-Musicans:X16.minus.17      0.09744    0.18523   0.526
##
## Correlation matrix not shown by default, as p = 17 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
## convergence code: 0
## boundary (singular) fit: see ?isSingular
## boundary (singular) fit: see ?isSingular
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + PianoPlay + (1 | Subject) +
##     (Instrument | Subject) + (Harmony | Subject) + Harmony:Musicans -
##     1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7479.6   7651.0  -3708.8   7417.6     1834
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7449 -0.5720  0.0078  0.5567  3.6404
##
```

```
## Random effects:
##  Groups    Name              Variance  Std.Dev.  Corr
##  Subject   (Intercept)       7.550e-14 2.748e-07
##  Subject.1 (Intercept)       8.285e-01 9.102e-01
##            Instrumentpiano   1.667e+00 1.291e+00 -0.51
##            Instrumentstring  3.549e+00 1.884e+00 -0.99  0.60
##  Subject.2 (Intercept)       1.177e+00 1.085e+00
##            HarmonyI-V-IV     3.681e-02 1.919e-01  1.00
##            HarmonyI-V-VI     1.423e+00 1.193e+00 -0.38 -0.38
##            HarmonyIV-I-V     7.480e-02 2.735e-01  0.28  0.28 -0.22
##  Residual                    2.469e+00 1.571e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                      Estimate Std. Error t value
## Instrumentguitar                      3.46763    0.39864   8.699
## Instrumentpiano                       4.99606    0.40992  12.188
## Instrumentstring                      6.92523    0.40197  17.228
## HarmonyI-V-IV                         0.10283    0.21357   0.482
## HarmonyI-V-VI                         1.94897    0.39025   4.994
## HarmonyIV-I-V                         0.40618    0.21984   1.848
## Voicepar3rd                          -0.38959    0.08914  -4.370
## Voicepar5th                          -0.31024    0.08915  -3.480
## PianoPlay                             0.79593    0.39712   2.004
## HarmonyI-IV-V:MusicansNon-Musicans    0.46053    0.40488   1.137
## HarmonyI-V-IV:MusicansNon-Musicans    0.36340    0.45873   0.792
## HarmonyI-V-VI:MusicansNon-Musicans   -0.94297    0.45830  -2.058
## HarmonyIV-I-V:MusicansNon-Musicans    0.04011    0.43434   0.092
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
##
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```



4.2-1.bb



4.2-2.bb

4.3 dichotomize "Selfdeclare" at 4

recall the frequency table is shown as below:

Selfdeclare: 1 2 3 4 5 6 number of subjects: 7 23 9 10 2 1

If we dichotomize "Selfdeclare" at 4, we will have 49 Non-Musicans and 3 Musicans. This will become problematic, because we will have too few data for Musicans to draw any reliable conclusions. So we should not try dichotomizing "Selfdeclare" at 4.

In conclusion, it seems that the results are not sensitive to where we dichotomize. We get the same model in the above two cases. And the interaction between HarmonyI-V-VI and MusicansNon-Musicans stays significant and its co-efficient stays the same. Generally, if the subject (listener) is non-Musicans and the harmony is I-V-VI, an average the Classical rating will additionally decrease.

5. Model for "Popular" rating

(a) Main effects: the influence of Instrument, Harmony & Voice on Popular ratings

Using stepAIC, we get the fixed effects as: lm(formula = Popular ~ Instrument + Harmony + Voice - 1, data = data).

Using ffRanefLMER.fnc to select the random effects, we get the final model lmer.2, (Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
Subject) + (Harmony | Subject) - 1).

The validation of lmer.2 is being confirmed by examing standardized conditional resisuals plots and binned plots.

Based on the summary of lmer.2, we find that: Instrument has the largest influence on Popular rating.And the instrument guitar has the largest positive influence on Popular rating out of the three instruments. If the Instrument is guitar, the Popular rating on average increases 6.74915. If the Instrument is piano, the Popular rating on average increases 5.61683.If the Instrument is string, the Popular rating on average increases 3.75867.

Also, HarmonyI-V-VI seems to have negative influence on Popular rating too. If the Harmony is I-V-VI, the Popular rating on average decreases 0.34981. If the Harmony is IV-I-V, the Popular rating on average decreases 0.25480.

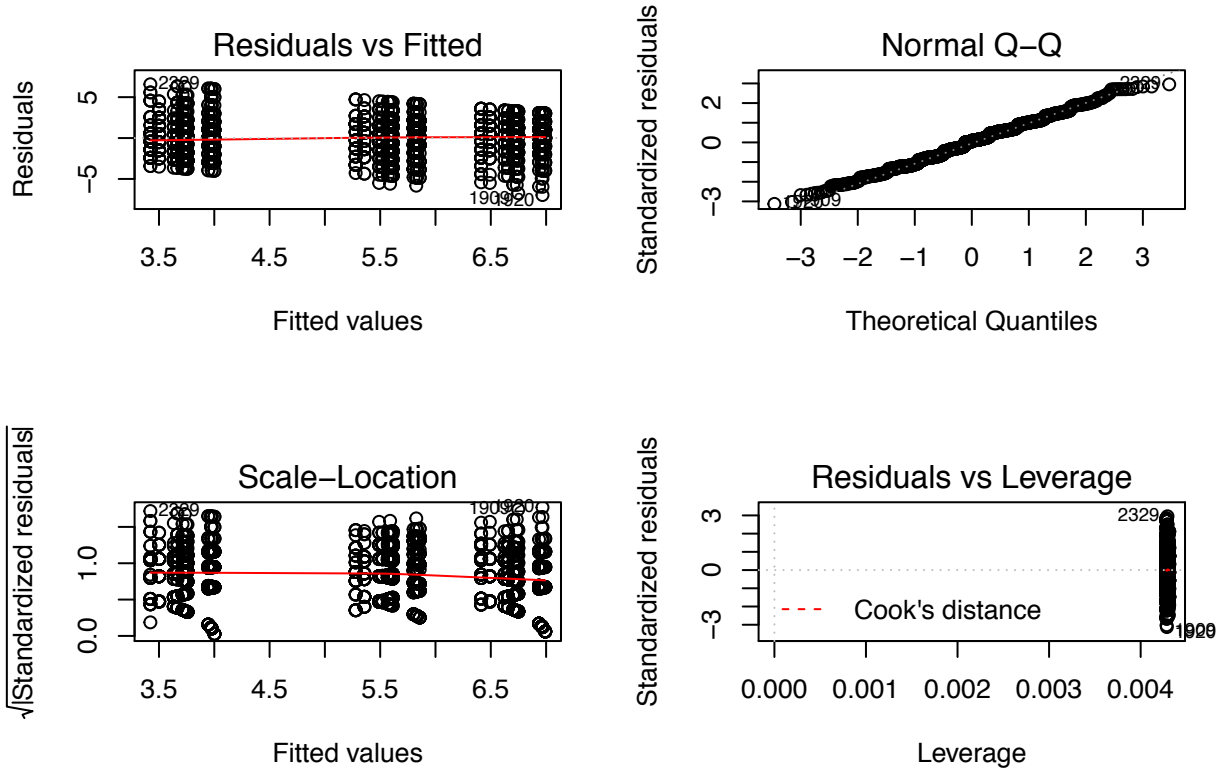
Voicepar3rd and Voicepar5th seems to have positive influence on Popular rating. If the Voice is par3rd, the Popular rating on average increases 0.22932. If the Voice is par5th, the Popular rating on average increases 0.25868.

Besides the fixed effects,there are random variability across different listeners in the degree to which they are inclined to call music played by certain instrument "Popular". The variance of the Instrumentguitar|Subject is 0.3190;the variance of the Instrumentpiano|Subject is 1.4681; the variance of the Instrumentstring|Subject is 3.1962.

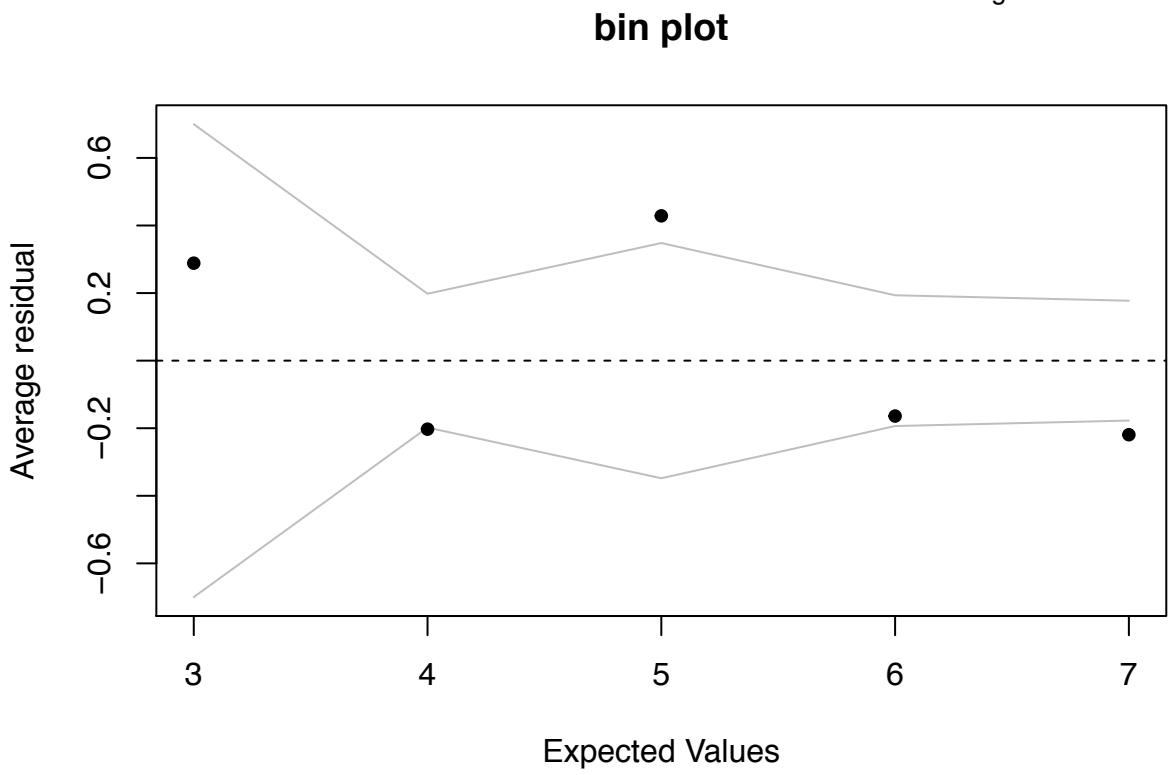
There are also random variability across different listeners in the degree to which they are inclined to call music with a certain harmony structure. The variance of HarmonyI-IV-V|Subject is 1.4406; The variance of HarmonyI-V-IV|Subject is 0.1215 ;the variance of HarmonyI-V-VI|Subject is 1.0585;the variance of HarmonyIV-I-V |Subject is 0.3054.

```
##
## Call:
## lm(formula = Popular ~ Instrument + Harmony + Voice - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9674 -1.6657  0.1669  1.4310  6.5777
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar  6.74837    0.14614  46.179   <2e-16 ***
## Instrumentpiano   5.61408    0.14627  38.382   <2e-16 ***
## Instrumentstring  3.75813    0.14606  25.730   <2e-16 ***
## HarmonyI-V-IV    -0.02443    0.14620  -0.167   0.8673
## HarmonyI-V-VI    -0.33585    0.14628  -2.296   0.0218 *
## HarmonyIV-I-V    -0.25703    0.14613  -1.759   0.0788 .
## Voicepar3rd       0.21199    0.12667   1.674   0.0944 .
## Voicepar5th       0.24345    0.12667   1.922   0.0548 .
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.233 on 1857 degrees of freedom
## Multiple R-squared:  0.8596, Adjusted R-squared:  0.859
## F-statistic:  1421 on 8 and 1857 DF,  p-value: < 2.2e-16
```

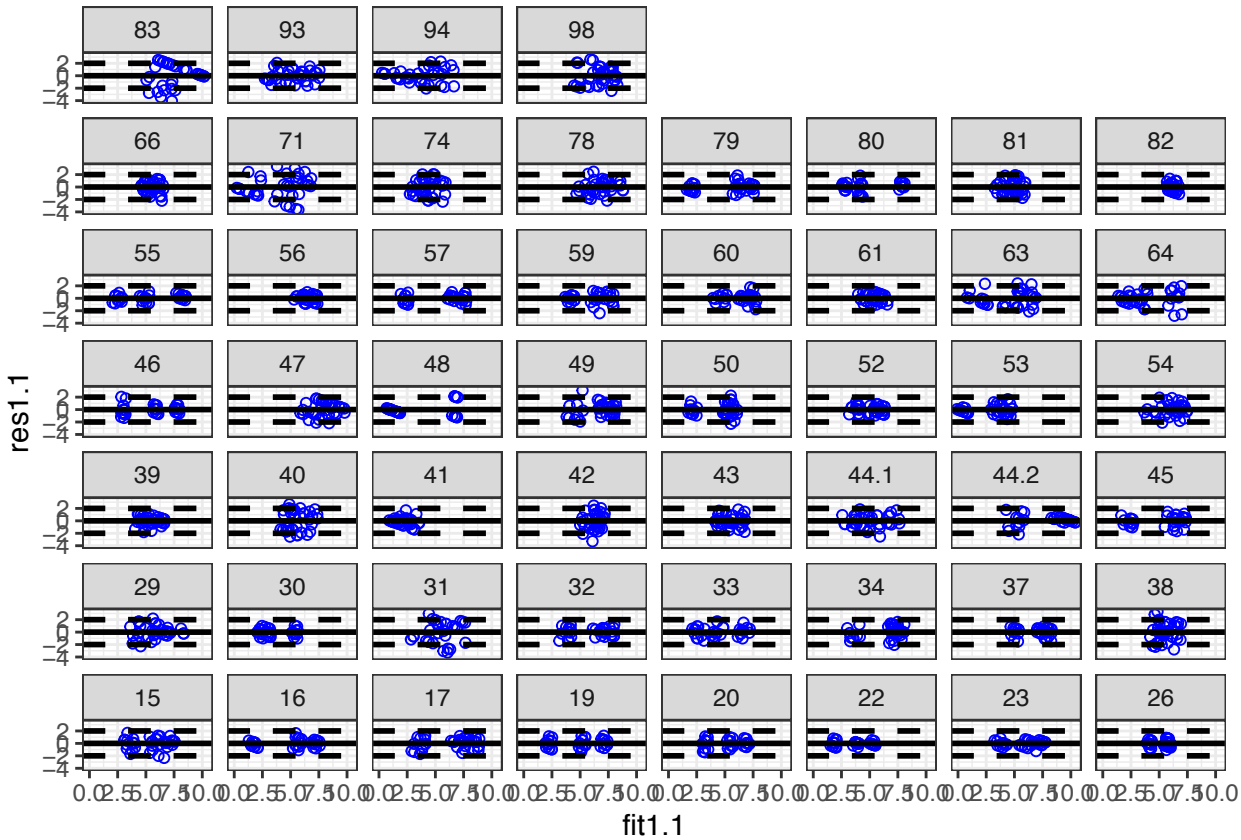


5.1-1.bb



5.1-2.bb

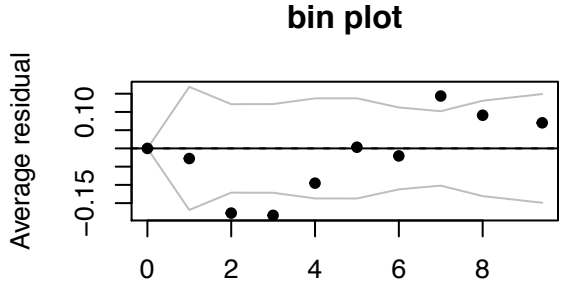
```
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 1.342741e-42
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 1.409666e-20
##  adding (Harmony|Subject) to model
## evaluating addition of (Voice|Subject) to model

## Warning in optwrap(optimizer, devfun, getStart(start, rho$lower, rho$pp), :
## convergence code 1 from bobyqa: bobyqa -- maximum number of function
## evaluations exceeded

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.4353067
##  not adding (Voice|Subject) to model
```

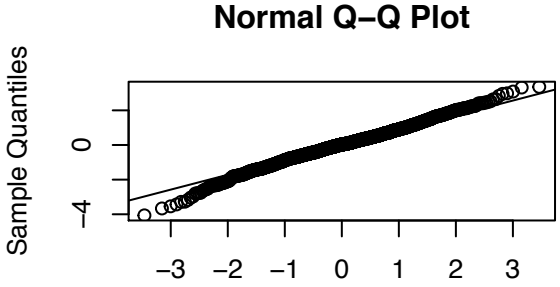


5.1-3.bb

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7495.4   7639.2  -3721.7   7443.4     1839
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name              Variance Std.Dev. Corr
##  Subject   (Intercept)       0.0000   0.0000
##  Subject.1 (Intercept)       0.2965   0.5445
##            Instrumentpiano   1.4316   1.1965   -0.65
##            Instrumentstring  2.7342   1.6536   -1.00  0.66
##  Subject.2 (Intercept)       1.5111   1.2293
##            HarmonyI-V-IV     0.1136   0.3370    0.43
##            HarmonyI-V-VI     1.0938   1.0459   -0.30 -0.38
##            HarmonyIV-I-V     0.3314   0.5757   -0.52 -0.81 -0.17
##  Residual                    2.5012   1.5815
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                   Estimate Std. Error t value
## Instrumentguitar  6.74915    0.21325  31.649
## Instrumentpiano   5.61683    0.23805  23.595
## Instrumentstring  3.75867    0.25183  14.925
## HarmonyI-V-IV    -0.01981    0.11362  -0.174
## HarmonyI-V-VI    -0.34191    0.17826  -1.918
## HarmonyIV-I-V    -0.25826    0.13071  -1.976
## Voicepar3rd       0.20964    0.08973   2.336
## Voicepar5th       0.24616    0.08973   2.743
##
## Correlation of Fixed Effects:
##              Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn   0.655
## Instrmntstr   0.450     0.687
## HrmnyI-V-IV -0.080    -0.071    -0.067
## HrmnyI-V-VI -0.336    -0.301    -0.285     0.136
## HrmnyIV-I-V -0.445    -0.399    -0.377     0.158  0.145
## Voicepar3rd -0.211    -0.188    -0.178    -0.001 -0.001  0.001
## Voicepar5th -0.210    -0.188    -0.178    -0.001 -0.002 -0.001  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```





5.1-4.bb

(b) Final Model for "Popular"___ adding other covariates

fixed effects part

From lmer.2, we know the fixed effects part we have selected are (Popular ~ Instrument + Harmony + Voice - 1).

Now, we can add other individual covariates as possible explanatory variables to build a full model: lm.full.

Using stepAIC, the final model selected are lm.final <- lm(formula = Popular ~ Instrument + Harmony + Voice + X16.minus.17 +ConsInstr+ PachListen + ClsListen + KnowRob + KnowAxis + X1990s2000s + X1990s2000s.minus.1960s1970s + Composing + PianoPlay - 1, data = data)

random effects part

We use ffRanefLMER.fnc to select random effects. The final model selected is lmer.4, (Popular ~ Instrument + Harmony + Voice + X16.minus.17 + ConsInstr +
PachListen + ClsListen + KnowRob + KnowAxis + X1990s2000s +
X1990s2000s.minus.1960s1970s + Composing + PianoPlay + (1 |
Subject) + (Instrument | Subject) + (Harmony | Subject) - 1 Data: data Control: lmerControl(optimizer = "bobyqa")

re-examine the fixed effects part

We notice that the co-efficients of following variables of lmer.4 have t-value less than 2: ConsInstr,PachListen,ClsListen, KnowRob, KnowAxis, X1990s2000s,X1990s2000s.minus.1960s1970s, Composing, PianoPlay.

So we delete them, and then produce a model lmer.4.2

lmer.4.2 <- lmer(Popular ~ Instrument + Harmony + Voice + X16.minus.17 + (1 | Subject) + (Instrument | Subject) +
(Harmony | Subject) - 1, data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

We notice that the co-efficients of X16.minus.17 has a t-value less than 2; So we drop the variable X16.minus.17 too. The final model is lmer.4.3 <-
lmer(Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument | Subject) +
(Harmony | Subject) - 1, data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

validate model lmer.4.3

The validation of model lmer.4.3 has been confirmed by examining standardized conditional resisuals plots and binned plots.

final model interpretation

Based on the summary output for lmer.4.3, we find that:

Instrument has the largest influence on Popular rating.And the instrument guitar has the largest positive influence on Popular rating out of the three instruments. If the Instrument is guitar, the Popular rating on average increases 5.29759. If the Instrument is piano, the Popular rating on average increases 4.22841.If the Instrument is string, the Popular rating on average increases 2.45492.

Also, HarmonyI-V-VI seems to have negative influence on Popular rating too. If the Harmony is I-V-VI, the Popular rating on average decreases 0.35005.

Voicepar3rd and Voicepar5th seems to have positive influence on Popular rating. If the Voice is par3rd, the Popular rating on average increases 0.22884. If the Voice is par5th, the Popular rating on average increases 0.25879.

Overall, Selfdeclare seems to have positive influence on Popular rating. If the Selfdeclare score is 2, the Popular rating on average increases 1.37668.If the Selfdeclare score is 5, the Popular rating on average increases 2.4390.

ClsListen1 seems to have positive influence on Popular rating. If the ClsListen score is 1, the Popular rating on average increases 0.87189.

X1990s2000s score seems to have negative influence on Popular rating. If the X1990s2000s score is 2, the Popular rating on average decreases 2.49621.If the X1990s2000s score is 3, the Popular rating on average decreases 3.20521.If the X1990s2000s score is 4, the Popular rating on average decreases 2.52953.If the X1990s2000s score is 5, the Popular rating on average decreases 2.80632.

X1990s2000s.minus.1960s1970s score seems to have positive influence on Popular rating. If X1990s2000s.minus.1960s1970s is -2, the Popular rating on average increases 3.84433 unit.If X1990s2000s.minus.1960s1970s is 0, the Popular rating on average increases 2.48689 unit. If X1990s2000s.minus.1960s1970s is 2, the Popular rating on average increases 2.46738 unit. If X1990s2000s.minus.1960s1970s is 3, the Popular rating on average increases 2.73177 unit. If X1990s2000s.minus.1960s1970s is 5, the Popular rating on average increases 3.40562 unit.

GuitarPlay4 score seems to have the postive influence on Popular rating. If GuitarPlay is rated as 4, the Popular rating on average increases 2.48366.

Besides the fixed effects,there are random variability across different listeners in the degree to which they are inclined to call music played by certain instrument "Popular". The variance of the Instrumentguitar|Subject is 0.3506;the variance of the Instrumentpiano|Subject is 1.4696;the variance of the Instrumentstring|Subject is 3.1972.

There are also random variability across different listeners in the degree to which they are inclined to call music with a certain harmony structure. The variance of HarmonyI-IV-V|Subject is 0.6871; The variance of HarmonyI-V-IV|Subject is 0.1474 ;the variance of HarmonyI-V-VI|Subject is 1.0685;the variance of HarmonyIV-I-V |Subject is 0.3225.

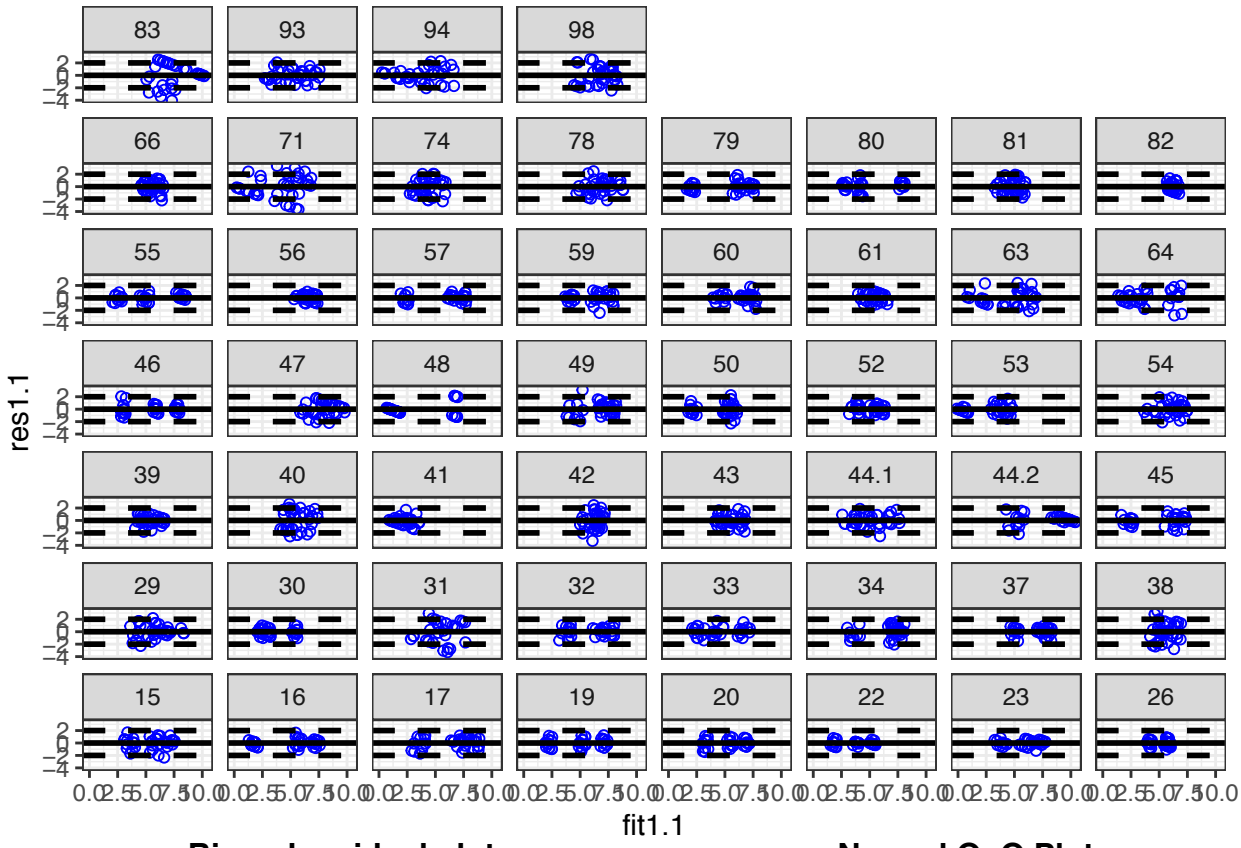
```
##
## Call:
## lm(formula = Popular ~ Instrument + Harmony + Voice + X16.minus.17 +
##     ConsInstr + PachListen + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + X1990s2000s.minus.1960s1970s + Composing +
##     PianoPlay - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.0984 -1.6060  0.0215  1.5755  6.3729
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar              6.58604    0.33631  19.583  < 2e-16 ***
## Instrumentpiano               5.45910    0.33628  16.234  < 2e-16 ***
## Instrumentstring              3.59643    0.33631  10.694  < 2e-16 ***
## HarmonyI-V-IV                -0.02294    0.14344  -0.160 0.872965
## HarmonyI-V-VI                -0.33520    0.14352  -2.336 0.019620 *
## HarmonyIV-I-V                -0.25703    0.14336  -1.793 0.073164 .
## Voicepar3rd                   0.21262    0.12427   1.711 0.087270 .
## Voicepar5th                   0.24408    0.12427   1.964 0.049674 *
## X16.minus.17                  0.08252    0.02030   4.066 4.99e-05 ***
## ConsInstr                     0.05560    0.03678   1.512 0.130743
## PachListen                   -0.79481    0.16275  -4.884 1.13e-06 ***
## ClsListen                    -0.11290    0.04050  -2.788 0.005364 **
```

```
## KnowRob                       0.52912    0.17186   3.079 0.002109 **
## KnowAxis                      0.34993    0.13960   2.507 0.012274 *
## X1990s2000s                   0.15199    0.04971   3.058 0.002264 **
## X1990s2000s.minus.1960s1970s -0.11167    0.04199  -2.660 0.007890 **
## Composing                     0.15516    0.04035   3.845 0.000125 ***
## PianoPlay                     0.36617    0.13982   2.619 0.008895 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.191 on 1847 degrees of freedom
## Multiple R-squared:  0.8656, Adjusted R-squared:  0.8643
## F-statistic: 660.7 on 18 and 1847 DF,  p-value: < 2.2e-16

## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.8623882
##  not adding (Voice|Subject) to model
## evaluating addition of (Instrument|Subject) to model

## 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

##  log-likelihood ratio test p-value = 9.191355e-43
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 2.187726e-21
##  adding (Harmony|Subject) to model

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + X16.minus.17 + ConsInstr +
##     PachListen + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + Composing + PianoPlay + (1 |
##     Subject) + (Instrument | Subject) + (Harmony | Subject) -      1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7503.1   7702.2  -3715.6   7431.1     1829
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9796 -0.5810  0.0273  0.5611  3.3068
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      4.141e-10 2.035e-05
##  Subject.1 (Intercept)      3.076e-01 5.546e-01
##            Instrumentpiano  1.430e+00 1.196e+00 -0.61
```

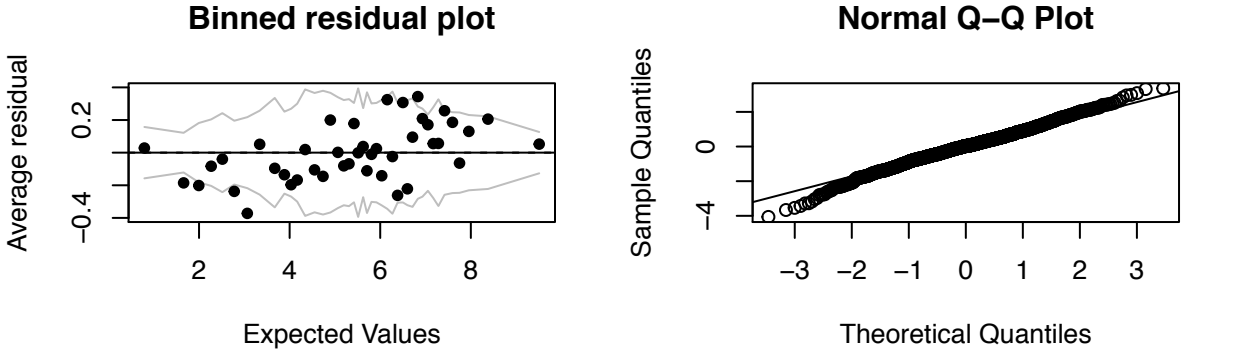
```
##            Instrumentstring 2.734e+00 1.653e+00 -1.00  0.66
##  Subject.2 (Intercept)      1.397e+00 1.182e+00
##            HarmonyI-V-IV    1.112e-01 3.335e-01  0.57
##            HarmonyI-V-VI    1.093e+00 1.045e+00 -0.38 -0.38
##            HarmonyIV-I-V    3.330e-01 5.771e-01 -0.61 -0.80 -0.17
##  Residual                   2.502e+00 1.582e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                Estimate Std. Error t value
## Instrumentguitar               6.438696   0.829827   7.759
## Instrumentpiano                5.306330   0.837004   6.340
## Instrumentstring               3.448310   0.840247   4.104
## HarmonyI-V-IV                 -0.020295   0.113430  -0.179
## HarmonyI-V-VI                 -0.341661   0.178213  -1.917
## HarmonyIV-I-V                 -0.257863   0.130837  -1.971
## Voicepar3rd                    0.210175   0.089732   2.342
## Voicepar5th                    0.246069   0.089732   2.742
## X16.minus.17                   0.114301   0.053475   2.137
## ConsInstr                      0.140265   0.096909   1.447
## PachListen                    -0.795515   0.429565  -1.852
## ClsListen                     -0.127351   0.106865  -1.192
## KnowRob                       -0.083714   0.449091  -0.186
## KnowAxis                       0.539041   0.368251   1.464
## X1990s2000s                    0.161779   0.130668   1.238
## X1990s2000s.minus.1960s1970s -0.130605   0.110497  -1.182
## Composing                      0.134894   0.106522   1.266
## PianoPlay                     -0.004595   0.369096  -0.012
##
## Correlation matrix not shown by default, as p = 18 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
## convergence code: 0
## boundary (singular) fit: see ?isSingular
## boundary (singular) fit: see ?isSingular
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + (1 |
##     Subject) + (Instrument | Subject) + (Harmony | Subject) -      1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7494.6   7643.9  -3720.3   7440.6     1838
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9937 -0.5791  0.0321  0.5574  3.3153
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      3.601e-09 6.001e-05
##  Subject.1 (Intercept)      3.685e-01 6.071e-01
```

```
##            Instrumentpiano  1.431e+00 1.196e+00 -0.62
##            Instrumentstring 2.734e+00 1.654e+00 -1.00  0.66
##  Subject.2 (Intercept)      1.531e+00 1.237e+00
##            HarmonyI-V-IV    1.146e-01 3.386e-01  0.41
##            HarmonyI-V-VI    1.096e+00 1.047e+00 -0.30 -0.37
##            HarmonyIV-I-V    3.369e-01 5.804e-01 -0.60 -0.77 -0.16
##  Residual                   2.501e+00 1.581e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                   Estimate Std. Error t value
## Instrumentguitar  6.60020    0.23252  28.385
## Instrumentpiano   5.46781    0.25342  21.576
## Instrumentstring  3.60975    0.26092  13.835
## HarmonyI-V-IV    -0.01990    0.11370  -0.175
## HarmonyI-V-VI    -0.34187    0.17836  -1.917
## HarmonyIV-I-V    -0.25805    0.13110  -1.968
## Voicepar3rd       0.20983    0.08972   2.339
## Voicepar5th       0.24613    0.08972   2.743
## X16.minus.17      0.08947    0.04961   1.803
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.702
## Instrmntstr  0.507     0.716
## HrmnyI-V-IV -0.078    -0.071    -0.070
## HrmnyI-V-VI -0.307    -0.282    -0.274     0.138
## HrmnyIV-I-V -0.447    -0.410    -0.398     0.163  0.147
## Voicepar3rd -0.193    -0.177    -0.172    -0.001 -0.001  0.001
## Voicepar5th -0.193    -0.177    -0.171    -0.001 -0.002 -0.001  0.500
## X16.mins.17 -0.355    -0.326    -0.316     0.000  0.000  0.000  0.001
##             Vcpr5t
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th
## X16.mins.17  0.000
## convergence code: 0
## boundary (singular) fit: see ?isSingular

## boundary (singular) fit: see ?isSingular

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7495.4   7639.2  -3721.7   7443.4     1839
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
##  Subject.1 (Intercept)      0.2965   0.5445
##            Instrumentpiano  1.4316   1.1965   -0.65
##            Instrumentstring 2.7342   1.6536   -1.00  0.66
##  Subject.2 (Intercept)      1.5111   1.2293
##            HarmonyI-V-IV    0.1136   0.3370    0.43
##            HarmonyI-V-VI    1.0938   1.0459   -0.30 -0.38
##            HarmonyIV-I-V    0.3314   0.5757   -0.52 -0.81 -0.17
##  Residual                   2.5012   1.5815
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  6.74915    0.21325  31.649
## Instrumentpiano   5.61683    0.23805  23.595
## Instrumentstring  3.75867    0.25183  14.925
## HarmonyI-V-IV    -0.01981    0.11362  -0.174
## HarmonyI-V-VI    -0.34191    0.17826  -1.918
## HarmonyIV-I-V    -0.25826    0.13071  -1.976
## Voicepar3rd       0.20964    0.08973   2.336
## Voicepar5th       0.24616    0.08973   2.743
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.655
## Instrmntstr  0.450     0.687
## HrmnyI-V-IV -0.080    -0.071    -0.067
## HrmnyI-V-VI -0.336    -0.301    -0.285     0.136
## HrmnyIV-I-V -0.445    -0.399    -0.377     0.158  0.145
## Voicepar3rd -0.211    -0.188    -0.178    -0.001 -0.001  0.001
## Voicepar5th -0.210    -0.188    -0.178    -0.001 -0.002 -0.001  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```



5.2-1.bb



5.2-2.bb

(c) Musicans vs Non-musicans

dichotomize "Selfdeclare" at 2

We can first dichotomize "Selfdeclare" at 2. Have a "Selfdeclare" score less and equal than 2 are categorized as "Non-Musicans". Have a "Selfdeclare" score bigger than 2 are categorized as "Musicans". We created a new column called Musicans, with value of "Non-Musicans" and "Musicans".

Recall that from part (b), we have the final model as lmer.4.3 <- lmer(Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument | Subject) +
(Harmony | Subject) - 1, data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

We can first examine the fixed effect part. Add the variable Musicans and add interaction between Musicans and the three main effects. So we have a full model: lm.full <- lm(Popular ~ Instrument + Harmony + Voice + Musicans + Instrument: Musicans+Harmony: Musicans + Voice : Musicans- 1, data = data)

Using step AIC, the model selected is: lm(formula = Popular ~ Instrument + Harmony + Voice + Musicans + Instrument:Musicans + Harmony:Musicans - 1, data = data)

we add random intercept and use ffRanefLMER.fnc to select random slope. The model selected is lmer.6:

Formula: Popular ~ Instrument + Harmony + Voice + Musicans + (1 | Subject) +
(Instrument | Subject) + (Harmony | Subject) + Instrument:Musicans +
Harmony:Musicans - 1

By examining the t-value of the fix effects of lmer.6, we notice that the co-efficients of Musicans has t-value much less than 2. So we can delete the variables Musicans; Thus, we get model lmer.7:

lmer.7 <- lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) + Instrument:Musicans+ Harmony:Musicans - 1,data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

Based on the summmary of the lmer.7, the interaction between HarmonyI-V-VI and MusicansNon-Musicans is significant, with the co-efficient equals 0.75703. This means that if the subject (listener) is non-Musicans and the harmony is I-V-VI, an average the Popular rating will additionally increase 0.75703. The interaction between Instrumentstring and MusicansNon-Musicans is signicant, with the co-efficient equals -1.00490. This means that if the subject (listener) is non-Musicans and the Instrument is string, an average the Popular rating will additionally decrease 1.00490.

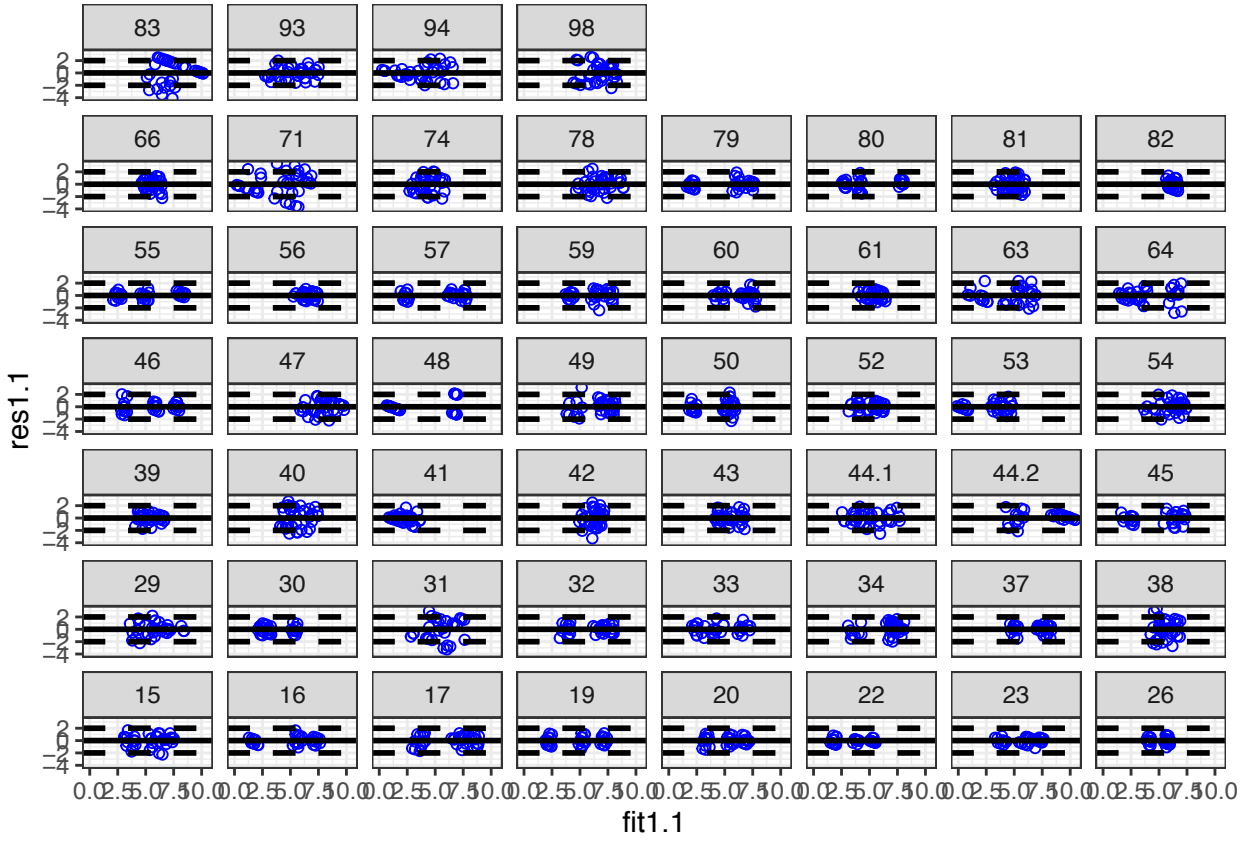
The validation of model lmer.7 has been confirmed by examining standardized conditional resisuals plots and binned plots.

```
##
## Call:
## lm(formula = Popular ~ Instrument + Harmony + Voice + Musicans +
##     Instrument:Musicans + Harmony:Musicans - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9771 -1.5863  0.0729  1.4740  6.4548
##
## Coefficients:
##                                       Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar                       6.76521    0.20648  32.765  < 2e-16
## Instrumentpiano                        5.91462    0.20729  28.533  < 2e-16
## Instrumentstring                       4.33718    0.20648  21.006  < 2e-16
## HarmonyI-V-IV                         -0.05146    0.22377  -0.230 0.818156
## HarmonyI-V-VI                         -0.76591    0.22377  -3.423 0.000633
## HarmonyIV-I-V                         -0.31996    0.22349  -1.432 0.152404
## Voicepar3rd                            0.21144    0.12582   1.680 0.093039
## Voicepar5th                            0.24365    0.12582   1.936 0.052966
## MusicansNon-Musicans                  -0.02808    0.25442  -0.110 0.912125
## Instrumentpiano:MusicansNon-Musicans  -0.49028    0.25509  -1.922 0.054763
## Instrumentstring:MusicansNon-Musicans -0.97535    0.25424  -3.836 0.000129
## HarmonyI-V-IV:MusicansNon-Musicans     0.04775    0.29413   0.162 0.871049
## HarmonyI-V-VI:MusicansNon-Musicans     0.74448    0.29425   2.530 0.011485
## HarmonyIV-I-V:MusicansNon-Musicans     0.10885    0.29392   0.370 0.711172
##
## Instrumentguitar                      ***
## Instrumentpiano                       ***
## Instrumentstring                      ***
## HarmonyI-V-IV
## HarmonyI-V-VI                         ***
## HarmonyIV-I-V
## Voicepar3rd                           .
## Voicepar5th                           .
```

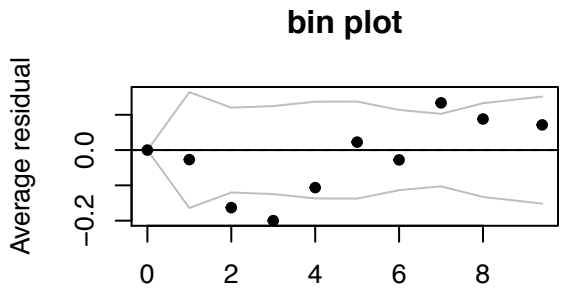
```
## MusicansNon-Musicans
## Instrumentpiano:MusicansNon-Musicans  .
## Instrumentstring:MusicansNon-Musicans ***
## HarmonyI-V-IV:MusicansNon-Musicans
## HarmonyI-V-VI:MusicansNon-Musicans    *
## HarmonyIV-I-V:MusicansNon-Musicans
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.218 on 1851 degrees of freedom
## Multiple R-squared:  0.8619, Adjusted R-squared:  0.8609
## F-statistic: 825.1 on 14 and 1851 DF,  p-value: < 2.2e-16

## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.9683414
##  not adding (Voice|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 1.606879e-39
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 1.43261e-18
##  adding (Harmony|Subject) to model

## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + Musicans + (1 | Subject) +
##     (Instrument | Subject) + (Harmony | Subject) + Instrument:Musicans +
##     Harmony:Musicans - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7498.8   7675.8  -3717.4   7434.8     1833
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0404 -0.5743  0.0195  0.5773  3.2956
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      1.426e-08 0.0001194
##  Subject.1 (Intercept)      2.735e-01 0.5229803
##            Instrumentpiano  1.374e+00 1.1720574 -0.63
##            Instrumentstring 2.503e+00 1.5821047 -1.00  0.65
##  Subject.2 (Intercept)      1.482e+00 1.2174460
##            HarmonyI-V-IV    1.177e-01 0.3430452  0.42
##            HarmonyI-V-VI    9.612e-01 0.9803950 -0.27 -0.41
##            HarmonyIV-I-V    3.332e-01 0.5772689 -0.51 -0.77 -0.21
##  Residual                   2.500e+00 1.5811799
## Number of obs: 1865, groups:  Subject, 52
##
```

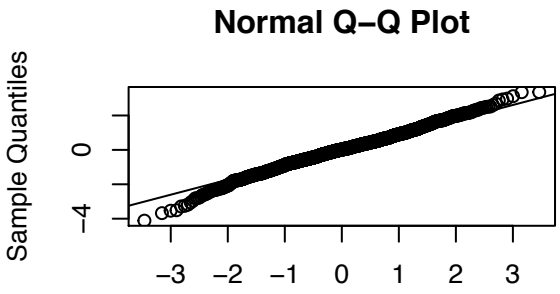
```
## Fixed effects:
##                                      Estimate Std. Error t value
## Instrumentguitar                      6.76656    0.31855  21.242
## Instrumentpiano                       5.91789    0.35893  16.487
## Instrumentstring                      4.33853    0.37421  11.594
## HarmonyI-V-IV                        -0.04197    0.17553  -0.239
## HarmonyI-V-VI                        -0.77876    0.26300  -2.961
## HarmonyIV-I-V                        -0.32279    0.20133  -1.603
## Voicepar3rd                           0.20952    0.08971   2.336
## Voicepar5th                           0.24617    0.08971   2.744
## MusicansNon-Musicans                 -0.02979    0.41379  -0.072
## Instrumentpiano:MusicansNon-Musicans -0.49197    0.37611  -1.308
## Instrumentstring:MusicansNon-Musicans -0.97511    0.47965  -2.033
## HarmonyI-V-IV:MusicansNon-Musicans    0.03826    0.23077   0.166
## HarmonyI-V-VI:MusicansNon-Musicans    0.75703    0.34607   2.187
## HarmonyIV-I-V:MusicansNon-Musicans    0.11168    0.26488   0.422
##
## 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
## boundary (singular) fit: see ?isSingular
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) + Instrument:Musicans + Harmony:Musicans -
##     1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7498.8   7675.8  -3717.4   7434.8     1833
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0404 -0.5743  0.0195  0.5773  3.2956
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
##  Subject.1 (Intercept)      0.2735   0.5230
##            Instrumentpiano  1.3737   1.1721   -0.63
##            Instrumentstring 2.5031   1.5821   -1.00  0.65
##  Subject.2 (Intercept)      1.4822   1.2174
##            HarmonyI-V-IV    0.1177   0.3430    0.42
##            HarmonyI-V-VI    0.9612   0.9804   -0.27 -0.41
##            HarmonyIV-I-V    0.3332   0.5773   -0.51 -0.77 -0.21
##  Residual                   2.5001   1.5812
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
```

```
##                                          Estimate Std. Error t value
## Instrumentguitar                          6.76656    0.31855  21.242
## Instrumentpiano                           5.91789    0.35893  16.487
## Instrumentstring                          4.33853    0.37421  11.594
## HarmonyI-V-IV                            -0.04197    0.17553  -0.239
## HarmonyI-V-VI                            -0.77876    0.26300  -2.961
## HarmonyIV-I-V                            -0.32279    0.20133  -1.603
## Voicepar3rd                               0.20952    0.08971   2.336
## Voicepar5th                               0.24617    0.08971   2.744
## Instrumentguitar:MusicansNon-Musicans    -0.02979    0.41379  -0.072
## Instrumentpiano:MusicansNon-Musicans     -0.52176    0.46737  -1.116
## Instrumentstring:MusicansNon-Musicans    -1.00490    0.48792  -2.060
## HarmonyI-V-IV:MusicansNon-Musicans        0.03826    0.23077   0.166
## HarmonyI-V-VI:MusicansNon-Musicans        0.75703    0.34607   2.187
## HarmonyIV-I-V:MusicansNon-Musicans        0.11168    0.26488   0.422

##
## 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
```



5.3-1.bb





5.3-2.bb

dichotomize "Selfdeclare" at 3

recall the frequency table is shown as below:

Selfdeclare: 1 2 3 4 5 6 number of subjects: 7 23 9 10 2 1

We can dichotomize "Selfdeclare" at 3. Have a "Selfdeclare" score less and equal than 3 are categorized as "Non-Musicans". Have a "Selfdeclare" score bigger than 3 are categorized as "Musicans".We created a new column called Musicans, with value of "Non-Musicans" and "Musicans".

Using this new defination, the final model we get is lmer.12 <- lmer(Popular ~ Instrument + Harmony + Voice + (1 | Subject) +
(Instrument | Subject) + (Harmony | Subject) - 1,data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

In the final model of lmer.12, variable Musicans and any interactions of Musicans with other variables are not kept. So this means if we dichotomize Selfdeclare" at 3, musicans and nonmusicans seems do not affect Popular ratings at all.

The validation of model lmer.12 has being confirmed by examing standardized conditional resisuals plots.

In conclusion, the results are sensitive to where we we dichotomize. If we we dichotomize "Selfdeclare" at 2,the interaction between HarmonyI-V-VI and MusicansNon-Musicans is significant, with the co-efficient equals 0.75703. This means that if the subject (listener) is non-Musicans and the harmony is I-V-VI, an average the Popular rating will additionally increase 0.75703. The interaction between Instrumentstring and MusicansNon-Musicans is signicant, with the co-efficient equals -1.00490. This means that if the subject (listener) is non-Musicans and the Instrument is string, an average the Popular rating will additionally decrease 1.00490.

However, if we dichotomize "Selfdeclare" at 3, musicans and nonmusicans seems do not affect Popular ratings at all.

```
data$Musicans[as.numeric(data$Selfdeclare) <= 3] <- "Non-Musicans"
data$Musicans[as.numeric(data$Selfdeclare) > 3] <- "Musicans"

lm.full <- lm(Popular ~ Instrument + Harmony + Voice + Musicans +
                Instrument: Musicans+Harmony: Musicans + Voice : Musicans -  1,
data = data)

lm.reduced <- MASS::stepAIC(lm.full, k = 2,direction = "backward"
                            ,trace = F)

summary(lm.reduced)
```

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

```
## lm(formula = Popular ~ Instrument + Harmony + Voice + Musicans +
##     Instrument:Musicans + Harmony:Musicans - 1, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9078 -1.6498  0.1619  1.4667  6.5615
##
## Coefficients:
##                                       Estimate Std. Error t value Pr(>|t|)
## Instrumentguitar                       6.87738    0.26234  26.216  < 2e-16
## Instrumentpiano                        5.96151    0.26420  22.564  < 2e-16
## Instrumentstring                       4.47994    0.26234  17.077  < 2e-16
## HarmonyI-V-IV                          0.03445    0.29229   0.118 0.906196
## HarmonyI-V-VI                         -1.11367    0.29230  -3.810 0.000143
## HarmonyIV-I-V                         -0.39684    0.29166  -1.361 0.173799
## Voicepar3rd                            0.21132    0.12600   1.677 0.093691
## Voicepar5th                            0.24435    0.12600   1.939 0.052624
## MusicansNon-Musicans                  -0.17118    0.29097  -0.588 0.556396
## Instrumentpiano:MusicansNon-Musicans  -0.29060    0.29262  -0.993 0.320786
## Instrumentstring:MusicansNon-Musicans -0.79087    0.29045  -2.723 0.006531
## HarmonyI-V-IV:MusicansNon-Musicans    -0.07718    0.33697  -0.229 0.818854
## HarmonyI-V-VI:MusicansNon-Musicans     1.03425    0.33703   3.069 0.002181
## HarmonyIV-I-V:MusicansNon-Musicans     0.18601    0.33642   0.553 0.580389
##
## Instrumentguitar                      ***
## Instrumentpiano                       ***
## Instrumentstring                      ***
## HarmonyI-V-IV
## HarmonyI-V-VI                         ***
## HarmonyIV-I-V
## Voicepar3rd                           .
## Voicepar5th                           .
## MusicansNon-Musicans
## Instrumentpiano:MusicansNon-Musicans
## Instrumentstring:MusicansNon-Musicans **
## HarmonyI-V-IV:MusicansNon-Musicans
## HarmonyI-V-VI:MusicansNon-Musicans    **
## HarmonyIV-I-V:MusicansNon-Musicans
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.221 on 1851 degrees of freedom
## Multiple R-squared:  0.8615, Adjusted R-squared:  0.8605
## F-statistic: 822.4 on 14 and 1851 DF,  p-value: < 2.2e-16
```

```
# Model selected:
# lm(formula = Popular ~ Instrument + Harmony + Voice + Musicans +
#     Instrument:Musicans + Harmony:Musicans - 1, data = data)


## add ramdom effect- random intercept

lmer.9 <-  lmer(Popular ~ Instrument + Harmony + Voice + Musicans +
    Instrument:Musicans + Harmony:Musicans - 1 + (1 |Subject),data = data,
```

```
    REML = F, control = lmerControl(optimizer = 'bobyqa'))


## use ffRanefLMER.fnc to select random slope

lmer.10 <- ffRanefLMER.fnc(lmer.9,ran.effects = c("(Voice|Subject)",
                                                  "(Instrument|Subject)",
                                                 "(Harmony|Subject)"),log.file=F)
```

```
## evaluating addition of (Voice|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 0.9385187
##  not adding (Voice|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 1.201156e-41
##  adding (Instrument|Subject) to model
## evaluating addition of (Harmony|Subject) to model

## boundary (singular) fit: see ?isSingular

##  log-likelihood ratio test p-value = 2.630712e-17
##  adding (Harmony|Subject) to model
```

```
summary(lmer.10)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + Musicans + (1 | Subject) +
##     (Instrument | Subject) + (Harmony | Subject) + Instrument:Musicans +
##     Harmony:Musicans - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7497.7   7674.7  -3716.8   7433.7     1833
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9682 -0.5757  0.0233  0.5696  3.2846
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      7.279e-09 8.532e-05
##  Subject.1 (Intercept)      2.796e-01 5.288e-01
##            Instrumentpiano  1.416e+00 1.190e+00 -0.65
##            Instrumentstring 2.618e+00 1.618e+00 -1.00  0.66
##  Subject.2 (Intercept)      1.475e+00 1.215e+00
##            HarmonyI-V-IV    1.137e-01 3.372e-01  0.41
##            HarmonyI-V-VI    8.946e-01 9.458e-01 -0.26 -0.36
##            HarmonyIV-I-V    3.292e-01 5.738e-01 -0.51 -0.78 -0.25
##  Residual                   2.500e+00 1.581e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
```

```
##                                       Estimate Std. Error t value
## Instrumentguitar                       6.87917    0.41213  16.692
## Instrumentpiano                        5.96763    0.46544  12.821
## Instrumentstring                       4.48173    0.48951   9.156
## HarmonyI-V-IV                          0.04881    0.22823   0.214
## HarmonyI-V-VI                         -1.13089    0.33496  -3.376
## HarmonyIV-I-V                         -0.40122    0.26162  -1.534
## Voicepar3rd                            0.20946    0.08971   2.335
## Voicepar5th                            0.24628    0.08971   2.745
## MusicansNon-Musicans                  -0.17300    0.47210  -0.366
## Instrumentpiano:MusicansNon-Musicans  -0.29492    0.43478  -0.678
## Instrumentstring:MusicansNon-Musicans -0.79086    0.55789  -1.418
## HarmonyI-V-IV:MusicansNon-Musicans    -0.09154    0.26315  -0.348
## HarmonyI-V-VI:MusicansNon-Musicans     1.05147    0.38653   2.720
## HarmonyIV-I-V:MusicansNon-Musicans     0.19039    0.30188   0.631
```

```
##
## 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 model selected:

# Formula: Popular ~ Instrument + Harmony + Voice + Musicans + (1 | Subject) +
#     (Instrument | Subject) + (Harmony | Subject) + Instrument:Musicans +
#     Harmony:Musicans - 1


# Examining the t-value of the fix effects of lmer10, we notice that the
# co-efficients of Musicans and Instrument:Musicans has t-value much less
# than 2. So we can delete these variables;

# Thus,  we get model lmer.11:

lmer.11 <-  lmer(Popular ~ Instrument + Harmony + Voice  + (1 | Subject) +
    (Instrument | Subject) + (Harmony | Subject) +
    Harmony:Musicans - 1,data = data,
    REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

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

```
summary(lmer.11)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) + Harmony:Musicans - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7495.7   7661.7  -3717.9   7435.7     1835
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9654 -0.5769  0.0252  0.5730  3.2863
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      3.432e-14 1.853e-07
##  Subject.1 (Intercept)      2.932e-01 5.414e-01
##            Instrumentpiano  1.432e+00 1.197e+00 -0.65
##            Instrumentstring 2.735e+00 1.654e+00 -1.00  0.66
##  Subject.2 (Intercept)      1.475e+00 1.215e+00
##            HarmonyI-V-IV    1.138e-01 3.373e-01  0.41
##            HarmonyI-V-VI    8.944e-01 9.457e-01 -0.26 -0.36
##            HarmonyIV-I-V    3.293e-01 5.738e-01 -0.51 -0.78 -0.25
##  Residual                   2.500e+00 1.581e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                    Estimate Std. Error t value
## Instrumentguitar                    7.08099    0.38482  18.401
## Instrumentpiano                     5.94847    0.39918  14.902
## Instrumentstring                    4.09042    0.40772  10.032
## HarmonyI-V-IV                       0.04899    0.22824   0.215
## HarmonyI-V-VI                      -1.13070    0.33494  -3.376
## HarmonyIV-I-V                      -0.40122    0.26163  -1.534
## Voicepar3rd                         0.20947    0.08971   2.335
## Voicepar5th                         0.24629    0.08971   2.745
## HarmonyI-IV-V:MusicansNon-Musicans -0.44215    0.42861  -1.032
## HarmonyI-V-IV:MusicansNon-Musicans -0.53387    0.47973  -1.113
## HarmonyI-V-VI:MusicansNon-Musicans  0.60911    0.46224   1.318
## HarmonyIV-I-V:MusicansNon-Musicans -0.25175    0.38087  -0.661
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HrI-V-IV HrI-V-VI HrIV-I-V
## Instrumntpn  0.885
## Instrmntstr  0.809     0.884
## HrmnyI-V-IV -0.098    -0.094    -0.092
## HrmnyI-V-VI -0.348    -0.335    -0.328     0.168
## HrmnyIV-I-V -0.484    -0.467    -0.457     0.167    0.127
## Voicepar3rd -0.117    -0.112    -0.110    -0.002    0.000    0.002
## Voicepar5th -0.116    -0.111    -0.109    -0.002   -0.003   -0.001
## HI-IV-V:MN- -0.835    -0.805    -0.789     0.088    0.312    0.434
## HI-V-IV:MN- -0.700    -0.675    -0.661    -0.397    0.199    0.309
## HI-V-VI:MN- -0.523    -0.504    -0.493    -0.040   -0.435    0.310
## HIV-I-V:MN- -0.608    -0.586    -0.574    -0.016    0.264   -0.198
##             Vcpr3r Vcpr5t HI-IV- HI-V-IV: HI-V-VI:
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th  0.500
## HI-IV-V:MN-  0.000 -0.001
```

```
## HI-V-IV:MN-  0.001  0.000  0.838
## HI-V-VI:MN-  0.000  0.001  0.626  0.585
## HIV-I-V:MN- -0.001  0.000  0.728  0.660    0.420
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# we notice that the Harmony:Musicans has t-value much less than 2. So we
# can drop the interaction of Harmony:Musicans

# The final model we get is lmer.12


lmer.12 <-  lmer(Popular ~ Instrument + Harmony + Voice  + (1 | Subject) +
    (Instrument | Subject) + (Harmony | Subject)  - 1,data = data,
    REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

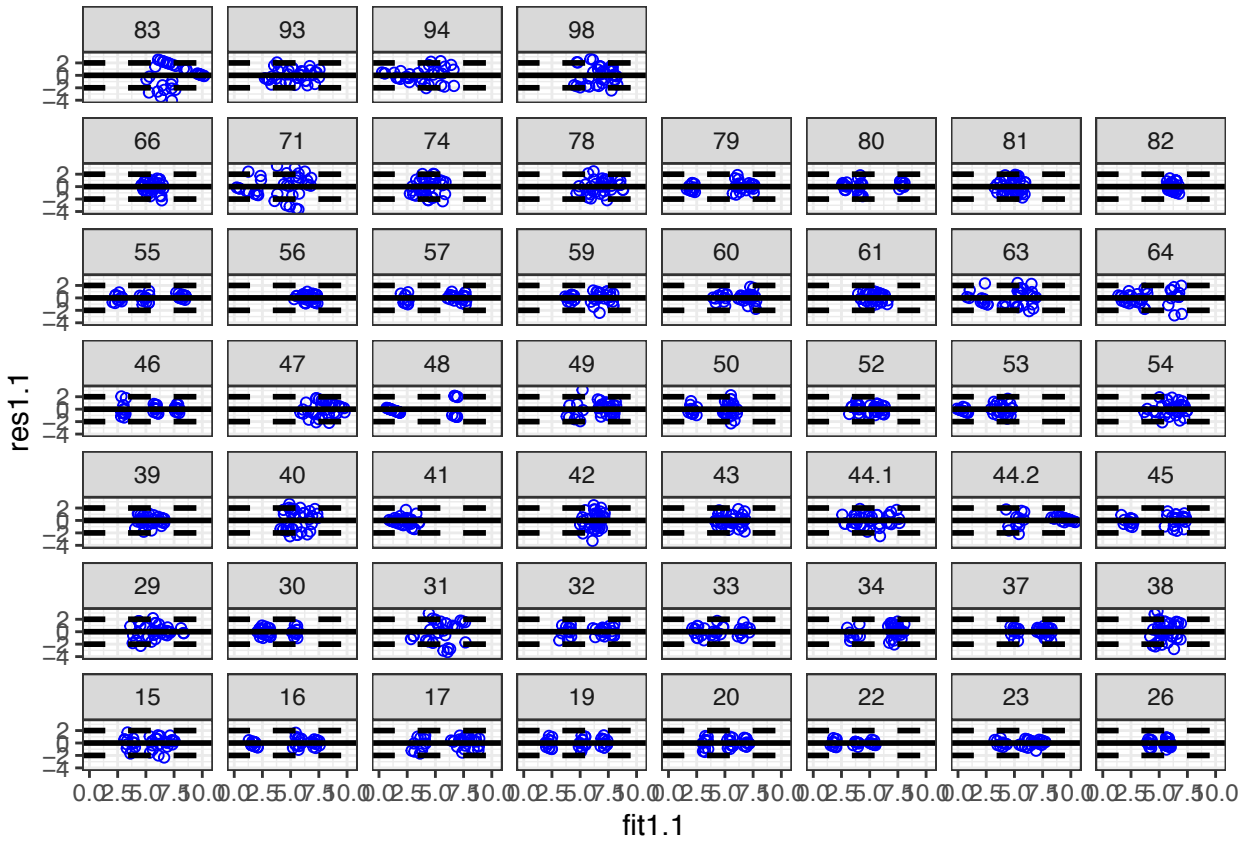
```
## boundary (singular) fit: see ?isSingular
```

```
summary(lmer.12)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7495.4   7639.2  -3721.7   7443.4     1839
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
##  Subject.1 (Intercept)      0.2965   0.5445
##            Instrumentpiano  1.4316   1.1965   -0.65
##            Instrumentstring 2.7342   1.6536   -1.00  0.66
##  Subject.2 (Intercept)      1.5111   1.2293
##            HarmonyI-V-IV    0.1136   0.3370    0.43
##            HarmonyI-V-VI    1.0938   1.0459   -0.30 -0.38
##            HarmonyIV-I-V    0.3314   0.5757   -0.52 -0.81 -0.17
##  Residual                   2.5012   1.5815
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  6.74915    0.21325  31.649
## Instrumentpiano   5.61683    0.23805  23.595
## Instrumentstring  3.75867    0.25183  14.925
## HarmonyI-V-IV    -0.01981    0.11362  -0.174
## HarmonyI-V-VI    -0.34191    0.17826  -1.918
## HarmonyIV-I-V    -0.25826    0.13071  -1.976
```

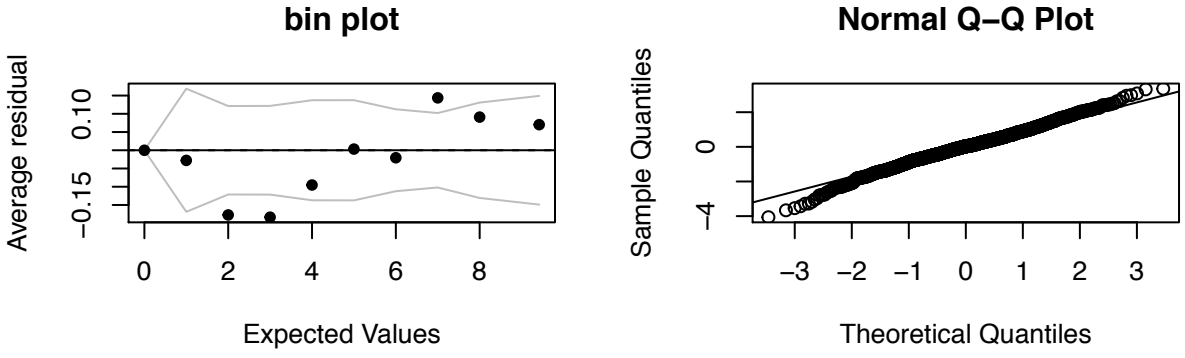
```
## Voicepar3rd       0.20964    0.08973   2.336
## Voicepar5th       0.24616    0.08973   2.743
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.655
## Instrmntstr  0.450     0.687
## HrmnyI-V-IV -0.080    -0.071    -0.067
## HrmnyI-V-VI -0.336    -0.301    -0.285     0.136
## HrmnyIV-I-V -0.445    -0.399    -0.377     0.158  0.145
## Voicepar3rd -0.211    -0.188    -0.178    -0.001 -0.001  0.001
## Voicepar5th -0.210    -0.188    -0.178    -0.001 -0.002 -0.001  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# In the final model of lmer.12, variable Musicans and any interactions of
# Musicans with other variables are not kept. So this means if we dichotomize
# "Selfdeclare" at 3, musicans and nonmusicans seems do not affect Popular
# ratings at all.


# This result is very differerent from if we dichotomize
# "Selfdeclare" at 2; if we dichotomize at 2, we have
# Instrumentstring:MusicansNon-Musicans and HarmonyI-V-VI:MusicansNon-Musicans
# are significant;



## validate lmer.12 using standardized conditional resisuals plots

res <- r.cond(lmer.12)     ## standardized conditional residuals
robust.sd <- diff(quantile(res,c(.025,.975)))/(2*1.96)
res <- res/robust.sd
fit <- yhat.cond(lmer.12)


newdata <- data.frame(Subject=data$Subject,res1.1=res,fit1.1=fit)

resparams <- data.frame(Subject=unique(data$Subject),
                        int1=0,slo1=0,
                        int2=2,slo2=0,
                        int3=-2,slo3=0)

mlm_facets(newdata,"Subject",x="fit1.1",y="res1.1",params=resparams,
           lty=c(1,2,2),size=c(1,1,1))
```



5.4-1.bb

```
res1.1 <- res
fit1.1 <- fit

par(mfrow=c(2,2))

binnedplot(round(fit1.1),round(res1.1),main="bin plot")
abline(h=0)
abline(h=2,lty=2)
abline(h=-2,lty=2)
qqnorm(res1.1)
qqline(res1.1)
```



5.4-2.bb

6 Summary

6.1 Examine the effects:

Classical Ratings:

Recall that from part 3, we have the final model lmer6.2 as

lmer6.2 <- lmer(Classical ~ Instrument + Harmony + Voice +
X16.minus.17 +PianoPlay + (1 | Subject) + (Instrument |
Subject) + (Harmony | Subject) - 1,data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))

1) Examine Instrument

Random effects: Groups Name Variance Std.Dev. Corr
Subject Instrumentguitar 0.73950 0.8599
Instrumentpiano 1.21627 1.1028 0.15
Instrumentstring 1.05715 1.0282 -0.99 -0.02
Subject.1 (Intercept) 1.05344 1.0264
HarmonyI-V-IV 0.05753 0.2399 0.87
HarmonyI-V-VI 1.86869 1.3670 -0.36 0.05
HarmonyIV-I-V 0.06998 0.2645 0.11 -0.12 0.16 Residual 2.46679 1.5706
Number of obs: 1865, groups: Subject, 52

Fixed effects: Estimate Std. Error t value Instrumentguitar 3.97494 0.23499 16.916 Instrumentpiano 5.50315 0.25382 21.681 Instrumentstring 7.43247 0.24761 30.017 HarmonyI-V-IV 0.02992 0.10809 0.277 HarmonyI-V-VI 0.89608 0.21571 4.154 HarmonyIV-I-V 0.09070 0.10914 0.831 Voicepar3rd -0.38972 0.08910 -4.374 Voicepar5th -0.31014 0.08911 -3.480 X16.minus.17 -0.11195 0.04886 -2.291 PianoPlay 0.92342 0.36601 2.523

2) Examine Harmony:

Random effects: Groups Name Variance Std.Dev. Corr
Subject.1 HarmonyI-IV-V 1.0534 1.0264
HarmonyI-V-IV 1.5369 1.2397 1.00
HarmonyI-V-VI 1.9108 1.3823 0.39 0.45
HarmonyIV-I-V 1.1855 1.0888 0.97 0.96 0.42 Residual 2.4668 1.5706
Number of obs: 1865, groups: Subject, 52

Fixed effects: Estimate Std. Error t value HarmonyI-IV-V 3.97494 0.23499 16.916 HarmonyI-V-IV 4.00486 0.25405 15.764 HarmonyI-V-VI 4.87101 0.26791 18.182 HarmonyIV-I-V 4.06564 0.24033 16.917

3) Examine Voice:

Fixed effects: Estimate Std. Error t value Voicecontrary 3.97494 0.23499 16.916 Voicepar3rd 3.58522 0.23495 15.260 Voicepar5th 3.66480 0.23499 15.596

```
# Recall that from part 3, we have the final model lmer6.2 as


# lmer6.2 <- lmer(Classical ~ Instrument + Harmony + Voice  +
#     X16.minus.17 +PianoPlay + (1 | Subject) + (Instrument |
#     Subject) + (Harmony | Subject) - 1,data = data,REML = F, control =
#       lmerControl(optimizer = 'bobyqa'))


summary(lmer6.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
```

```
## Classical ~ Instrument + Harmony + Voice + X16.minus.17 + PianoPlay +
##     (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) -      1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7478.9   7633.8  -3711.5   7422.9     1837
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6926 -0.5842  0.0166  0.5449  3.6822
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      5.983e-08 0.0002446
##  Subject.1 (Intercept)      7.395e-01 0.8599395
##            Instrumentpiano  1.667e+00 1.2910689 -0.54
##            Instrumentstring 3.549e+00 1.8838126 -1.00  0.60
##  Subject.2 (Intercept)      1.053e+00 1.0263711
##            HarmonyI-V-IV    5.753e-02 0.2398589  0.87
##            HarmonyI-V-VI    1.869e+00 1.3670010 -0.36  0.05
##            HarmonyIV-I-V    6.998e-02 0.2645403  0.11 -0.12  0.16
##  Residual                   2.467e+00 1.5706033
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                   Estimate Std. Error t value
## Instrumentguitar  3.97494    0.23499  16.916
## Instrumentpiano   5.50315    0.25382  21.681
## Instrumentstring  7.43247    0.24761  30.017
## HarmonyI-V-IV     0.02992    0.10809   0.277
## HarmonyI-V-VI     0.89608    0.21571   4.154
## HarmonyIV-I-V     0.09070    0.10914   0.831
## Voicepar3rd      -0.38972    0.08910  -4.374
## Voicepar5th      -0.31014    0.08911  -3.480
## X16.minus.17     -0.11195    0.04886  -2.291
## PianoPlay         0.92342    0.36601   2.523
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.667
## Instrmntstr  0.347     0.583
## HrmnyI-V-IV -0.047    -0.043    -0.044
## HrmnyI-V-VI -0.296    -0.274    -0.281     0.241
## HrmnyIV-I-V -0.183    -0.169    -0.173     0.435  0.271
## Voicepar3rd -0.190    -0.176    -0.180    -0.001 -0.001  0.001
## Voicepar5th -0.190    -0.175    -0.180    -0.001 -0.002 -0.001  0.500
## X16.mins.17 -0.309    -0.286    -0.294    -0.001  0.000  0.000  0.000
## PianoPlay   -0.258    -0.238    -0.244     0.001  0.000  0.000  0.000
##             Vcpr5t X16..1
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
```

```
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th
## X16.mins.17  0.000
## PianoPlay    0.000 -0.122
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# examine the impact of Instrument; put Instrument first,
# using -1 to get rid of the baseline;

lmer.I <- lmer(Classical ~ Instrument + Harmony + Voice +
X16.minus.17  +PianoPlay  + (Instrument-1 |
Subject) + (Harmony | Subject) - 1,data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

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

```
summary(lmer.I)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + X16.minus.17 + PianoPlay +
##     (Instrument - 1 | Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7476.9   7626.3  -3711.5   7422.9     1838
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6926 -0.5842  0.0166  0.5449  3.6822
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   Instrumentguitar 0.73950  0.8599
##            Instrumentpiano  1.21627  1.1028    0.15
##            Instrumentstring 1.05715  1.0282   -0.99 -0.02
##  Subject.1 (Intercept)      1.05344  1.0264
##            HarmonyI-V-IV    0.05753  0.2399    0.87
##            HarmonyI-V-VI    1.86869  1.3670   -0.36  0.05
##            HarmonyIV-I-V    0.06998  0.2645    0.11 -0.12  0.16
##  Residual                   2.46679  1.5706
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  3.97494    0.23499  16.916
## Instrumentpiano   5.50315    0.25382  21.681
## Instrumentstring  7.43247    0.24761  30.017
## HarmonyI-V-IV     0.02992    0.10809   0.277
## HarmonyI-V-VI     0.89608    0.21571   4.154
## HarmonyIV-I-V     0.09070    0.10914   0.831
## Voicepar3rd      -0.38972    0.08910  -4.374
## Voicepar5th      -0.31014    0.08911  -3.480
```

```
## X16.minus.17     -0.11195    0.04886  -2.291
## PianoPlay         0.92342    0.36601   2.523
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.667
## Instrmntstr  0.347     0.583
## HrmnyI-V-IV -0.047    -0.043    -0.044
## HrmnyI-V-VI -0.296    -0.274    -0.281     0.241
## HrmnyIV-I-V -0.183    -0.169    -0.173     0.435  0.271
## Voicepar3rd -0.190    -0.176    -0.180    -0.001 -0.001  0.001
## Voicepar5th -0.190    -0.175    -0.180    -0.001 -0.002 -0.001  0.500
## X16.mins.17 -0.309    -0.286    -0.294    -0.001  0.000  0.000  0.000
## PianoPlay   -0.258    -0.238    -0.244     0.001  0.000  0.000  0.000
##             Vcpr5t X16..1
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th
## X16.mins.17  0.000
## PianoPlay    0.000 -0.122
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# examine the impact of Harmony; put Harmony first,
# using -1 to get rid of the baseline;

lmer.H <- lmer(Classical ~ Harmony +Instrument + Voice +
    X16.minus.17 +PianoPlay + (Instrument|
Subject) + (Harmony-1 | Subject) - 1,
    data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

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

```
summary(lmer.H)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Harmony + Instrument + Voice + X16.minus.17 + PianoPlay +
##     (Instrument | Subject) + (Harmony - 1 | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7476.9   7626.3  -3711.5   7422.9     1838
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6926 -0.5842  0.0166  0.5449  3.6822
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
```

```
##  Subject   (Intercept)      0.7395   0.8599
##            Instrumentpiano  1.6669   1.2911   -0.54
##            Instrumentstring 3.5488   1.8838   -1.00  0.60
##  Subject.1 HarmonyI-IV-V    1.0534   1.0264
##            HarmonyI-V-IV    1.5369   1.2397   1.00
##            HarmonyI-V-VI    1.9108   1.3823   0.39  0.45
##            HarmonyIV-I-V    1.1855   1.0888   0.97  0.96  0.42
##  Residual                   2.4668   1.5706
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## HarmonyI-IV-V     3.97494    0.23499  16.916
## HarmonyI-V-IV     4.00486    0.25405  15.764
## HarmonyI-V-VI     4.87101    0.26791  18.182
## HarmonyIV-I-V     4.06564    0.24033  16.917
## Instrumentpiano   1.52821    0.20009   7.638
## Instrumentstring  3.45753    0.27597  12.529
## Voicepar3rd      -0.38972    0.08910  -4.374
## Voicepar5th      -0.31014    0.08911  -3.480
## X16.minus.17     -0.11195    0.04886  -2.291
## PianoPlay         0.92342    0.36601   2.523
##
## Correlation of Fixed Effects:
##             HI-IV- HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r Vcpr5t
## HrmnyI-V-IV  0.905
## HrmnyI-V-VI  0.639  0.656
## HrmnyIV-I-V  0.895  0.892  0.651
## Instrumntpn -0.328 -0.303 -0.287 -0.321
## Instrmntstr -0.540 -0.500 -0.474 -0.528  0.577
## Voicepar3rd -0.190 -0.176 -0.167 -0.185  0.000     0.000
## Voicepar5th -0.190 -0.176 -0.168 -0.186  0.000     0.000     0.500
## X16.mins.17 -0.309 -0.287 -0.271 -0.303  0.000     0.000     0.000  0.000
## PianoPlay   -0.258 -0.238 -0.226 -0.252  0.000     0.000     0.000  0.000
##             X16..1
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## Voicepar3rd
## Voicepar5th
## X16.mins.17
## PianoPlay   -0.122
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# examine the impact of Voice; put Voice first,
# using -1 to get rid of the baseline;

lmer.V <- lmer(Classical ~ Voice+Harmony +Instrument +
    X16.minus.17 +PianoPlay +(1|Subject) + (Instrument|
Subject) + (Harmony | Subject) - 1,
    data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

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

```
summary(lmer.V)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Voice + Harmony + Instrument + X16.minus.17 + PianoPlay +
##     (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) -      1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7478.9   7633.8  -3711.5   7422.9     1837
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6926 -0.5842  0.0166  0.5449  3.6822
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      4.019e-09 6.339e-05
##  Subject.1 (Intercept)      7.395e-01 8.599e-01
##            Instrumentpiano  1.667e+00 1.291e+00 -0.54
##            Instrumentstring 3.549e+00 1.884e+00 -1.00  0.60
##  Subject.2 (Intercept)      1.053e+00 1.026e+00
##            HarmonyI-V-IV    5.753e-02 2.399e-01  0.87
##            HarmonyI-V-VI    1.869e+00 1.367e+00 -0.36  0.05
##            HarmonyIV-I-V    6.998e-02 2.645e-01  0.11 -0.12  0.16
##  Residual                   2.467e+00 1.571e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Voicecontrary     3.97494    0.23499  16.916
## Voicepar3rd       3.58522    0.23495  15.260
## Voicepar5th       3.66480    0.23499  15.596
## HarmonyI-V-IV     0.02992    0.10809   0.277
## HarmonyI-V-VI     0.89608    0.21571   4.154
## HarmonyIV-I-V     0.09070    0.10914   0.831
## Instrumentpiano   1.52821    0.20009   7.638
## Instrumentstring  3.45753    0.27597  12.529
## X16.minus.17     -0.11195    0.04886  -2.291
## PianoPlay         0.92342    0.36601   2.523
##
## Correlation of Fixed Effects:
##             Vccntr Vcpr3r Vcpr5t HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts
## Voicepar3rd  0.928
## Voicepar5th  0.928  0.928
## HrmnyI-V-IV -0.047 -0.047 -0.047
## HrmnyI-V-VI -0.296 -0.296 -0.296  0.241
## HrmnyIV-I-V -0.183 -0.182 -0.183  0.435  0.271
## Instrumntpn -0.328 -0.328 -0.328  0.001  0.000  0.000
## Instrmntstr -0.540 -0.540 -0.540  0.000  0.000  0.000  0.577
## X16.mins.17 -0.309 -0.309 -0.309 -0.001  0.000  0.000  0.000     0.000
## PianoPlay   -0.258 -0.258 -0.258  0.001  0.000  0.000  0.000     0.000
```

```
##             X16..1
## Voicepar3rd
## Voicepar5th
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## X16.mins.17
## PianoPlay   -0.122
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

Popular Ratings:

Recall that from part 4, we have the final model lmer.4.3 <- lmer(Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument | Subject) +
(Harmony | Subject) - 1, data = data, REML = F, control = lmerControl(optimizer = 'bobyqa'))

1) Examine Instrument

Random effects: Groups Name Variance Std.Dev. Corr
Subject Instrumentguitar 0.2965 0.5445
Instrumentpiano 0.8765 0.9362 -0.25
Instrumentstring 1.2302 1.1091 -1.00 0.27
Subject.1 (Intercept) 1.5111 1.2293
HarmonyI-V-IV 0.1136 0.3370 0.43
HarmonyI-V-VI 1.0938 1.0459 -0.30 -0.38
HarmonyIV-I-V 0.3314 0.5757 -0.52 -0.81 -0.17 Residual 2.5012 1.5815
Number of obs: 1865, groups: Subject, 52

Fixed effects: Estimate Std. Error t value Instrumentguitar 6.74915 0.21325 31.649 Instrumentpiano 5.61683 0.23805 23.595 Instrumentstring 3.75867 0.25183 14.925 HarmonyI-V-IV -0.01981 0.11362 -0.174 HarmonyI-V-VI -0.34191 0.17826 -1.918 HarmonyIV-I-V -0.25826 0.13071 -1.976 Voicepar3rd 0.20964 0.08973 2.336 Voicepar5th 0.24616 0.08973 2.743

2) Examine Harmony

Random effects: Groups Name Variance Std.Dev. Corr
Subject.1 HarmonyI-IV-V 1.5111 1.2293
HarmonyI-V-IV 1.9806 1.4073 0.98
HarmonyI-V-VI 1.8327 1.3538 0.68 0.61
HarmonyIV-I-V 1.1100 1.0536 0.88 0.79 0.46 Residual 2.5012 1.5815
Number of obs: 1865, groups: Subject, 52

Fixed effects: Estimate Std. Error t value HarmonyI-IV-V 6.74915 0.21325 31.649 HarmonyI-V-IV 6.72934 0.23350 28.819 HarmonyI-V-VI 6.40724 0.22742 28.173 HarmonyIV-I-V 6.49089 0.19433 33.402

3) Examine Voice:

Fixed effects: Estimate Std. Error t value Voicecontrary 6.74915 0.21325 31.649 Voicepar3rd 6.95879 0.21322 32.636 Voicepar5th 6.99531 0.21326 32.801

```
# Recall that from part 5, we have the final model
# lmer.4.3 <-  lmer(Popular ~ Instrument + Harmony + Voice
# + (1 | Subject) + (Instrument | Subject) +
#     (Harmony | Subject) - 1, data = data,
#     REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

```r
summary(lmer.4.3)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7495.4   7639.2  -3721.7   7443.4     1839
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name              Variance Std.Dev. Corr
##  Subject   (Intercept)       0.0000   0.0000
##  Subject.1 (Intercept)       0.2965   0.5445
##            Instrumentpiano   1.4316   1.1965   -0.65
##            Instrumentstring  2.7342   1.6536   -1.00  0.66
##  Subject.2 (Intercept)       1.5111   1.2293
##            HarmonyI-V-IV     0.1136   0.3370    0.43
##            HarmonyI-V-VI     1.0938   1.0459   -0.30 -0.38
##            HarmonyIV-I-V     0.3314   0.5757   -0.52 -0.81 -0.17
##  Residual                    2.5012   1.5815
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  6.74915    0.21325  31.649
## Instrumentpiano   5.61683    0.23805  23.595
## Instrumentstring  3.75867    0.25183  14.925
## HarmonyI-V-IV    -0.01981    0.11362  -0.174
## HarmonyI-V-VI    -0.34191    0.17826  -1.918
## HarmonyIV-I-V    -0.25826    0.13071  -1.976
## Voicepar3rd       0.20964    0.08973   2.336
## Voicepar5th       0.24616    0.08973   2.743
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.655
## Instrmntstr  0.450     0.687
## HrmnyI-V-IV -0.080    -0.071    -0.067
## HrmnyI-V-VI -0.336    -0.301    -0.285     0.136
## HrmnyIV-I-V -0.445    -0.399    -0.377     0.158  0.145
## Voicepar3rd -0.211    -0.188    -0.178    -0.001 -0.001  0.001
## Voicepar5th -0.210    -0.188    -0.178    -0.001 -0.002 -0.001  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# examine the impact of Instrument; put Instrument first,
# using -1 to get rid of the baseline;

lmer.I <- lmer(Popular ~ Instrument + Harmony + Voice + (Instrument-1 |
Subject) + (Harmony | Subject) - 1,data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

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

```r
summary(lmer.I)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (Instrument - 1 | Subject) +
##     (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7493.4   7631.7  -3721.7   7443.4     1840
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   Instrumentguitar 0.2965   0.5445
##            Instrumentpiano  0.8765   0.9362   -0.25
##            Instrumentstring 1.2302   1.1091   -1.00  0.27
##  Subject.1 (Intercept)      1.5111   1.2293
##            HarmonyI-V-IV    0.1136   0.3370    0.43
##            HarmonyI-V-VI    1.0938   1.0459   -0.30 -0.38
##            HarmonyIV-I-V    0.3314   0.5757   -0.52 -0.81 -0.17
##  Residual                   2.5012   1.5815
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  6.74915    0.21325  31.649
## Instrumentpiano   5.61683    0.23805  23.595
## Instrumentstring  3.75867    0.25183  14.925
## HarmonyI-V-IV    -0.01981    0.11362  -0.174
## HarmonyI-V-VI    -0.34191    0.17826  -1.918
## HarmonyIV-I-V    -0.25826    0.13071  -1.976
## Voicepar3rd       0.20964    0.08973   2.336
## Voicepar5th       0.24616    0.08973   2.743
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.655
## Instrmntstr  0.450     0.687
## HrmnyI-V-IV -0.080    -0.071    -0.067
## HrmnyI-V-VI -0.336    -0.301    -0.285     0.136
## HrmnyIV-I-V -0.445    -0.399    -0.377     0.158  0.145
```

```
## Voicepar3rd -0.211    -0.188    -0.178    -0.001 -0.001  0.001
## Voicepar5th -0.210    -0.188    -0.178    -0.001 -0.002 -0.001  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# examine the impact of Harmony; put Harmony first,
# using -1 to get rid of the baseline;

lmer.H <- lmer(Popular ~ Harmony +Instrument + Voice + (Instrument |
Subject) + (Harmony-1 | Subject) - 1,data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

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

```
summary(lmer.H)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Harmony + Instrument + Voice + (Instrument | Subject) +
##     (Harmony - 1 | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7493.4   7631.7  -3721.7   7443.4     1840
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name              Variance Std.Dev. Corr
##  Subject   (Intercept)       0.2965   0.5445
##            Instrumentpiano   1.4316   1.1965   -0.65
##            Instrumentstring  2.7342   1.6536   -1.00  0.66
##  Subject.1 HarmonyI-IV-V     1.5111   1.2293
##            HarmonyI-V-IV     1.9806   1.4073   0.98
##            HarmonyI-V-VI     1.8327   1.3538   0.68  0.61
##            HarmonyIV-I-V     1.1100   1.0536   0.88  0.79  0.46
##  Residual                    2.5012   1.5815
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                   Estimate Std. Error t value
## HarmonyI-IV-V      6.74915    0.21325  31.649
## HarmonyI-V-IV      6.72934    0.23350  28.819
## HarmonyI-V-VI      6.40724    0.22742  28.173
## HarmonyIV-I-V      6.49089    0.19433  33.402
## Instrumentpiano   -1.13232    0.18874  -5.999
## Instrumentstring  -2.99048    0.24618 -12.147
## Voicepar3rd        0.20964    0.08973   2.336
## Voicepar5th        0.24616    0.08973   2.743
##
## Correlation of Fixed Effects:
##             HI-IV- HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r
## HrmnyI-V-IV  0.874
```

```
## HrmnyI-V-VI  0.674  0.632
## HrmnyIV-I-V  0.798  0.738  0.536
## Instrumntpn -0.303 -0.277 -0.284 -0.333
## Instrmntstr -0.406 -0.371 -0.381 -0.446  0.630
## Voicepar3rd -0.211 -0.193 -0.198 -0.231  0.000     0.000
## Voicepar5th -0.210 -0.192 -0.199 -0.231  0.000     0.000     0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# examine the impact of Voice; put Voice first,
# using -1 to get rid of the baseline;

lmer.V <- lmer(Popular ~ Voice+Harmony +Instrument +(1|Subject) + (Instrument|
Subject) + (Harmony | Subject) - 1,
    data = data,REML = F, control = lmerControl(optimizer = 'bobyqa'))
```

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

```
summary(lmer.V)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Voice + Harmony + Instrument + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7495.4   7639.2  -3721.7   7443.4     1839
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      1.263e-12 1.124e-06
##  Subject.1 (Intercept)      2.965e-01 5.445e-01
##            Instrumentpiano  1.432e+00 1.196e+00 -0.65
##            Instrumentstring 2.734e+00 1.654e+00 -1.00  0.66
##  Subject.2 (Intercept)      1.511e+00 1.229e+00
##            HarmonyI-V-IV    1.136e-01 3.370e-01  0.43
##            HarmonyI-V-VI    1.094e+00 1.046e+00 -0.30 -0.38
##            HarmonyIV-I-V    3.314e-01 5.757e-01 -0.52 -0.81 -0.17
##  Residual                   2.501e+00 1.582e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Voicecontrary     6.74915    0.21325  31.649
## Voicepar3rd       6.95879    0.21322  32.636
## Voicepar5th       6.99531    0.21326  32.801
## HarmonyI-V-IV    -0.01981    0.11362  -0.174
## HarmonyI-V-VI    -0.34191    0.17826  -1.918
## HarmonyIV-I-V    -0.25826    0.13071  -1.976
```

```
## Instrumentpiano  -1.13232    0.18874  -5.999
## Instrumentstring -2.99048    0.24618 -12.147
##
## Correlation of Fixed Effects:
##             Vccntr Vcpr3r Vcpr5t HI-V-I HI-V-V HIV-I- Instrmntp
## Voicepar3rd  0.911
## Voicepar5th  0.911  0.912
## HrmnyI-V-IV -0.080 -0.080 -0.080
## HrmnyI-V-VI -0.336 -0.336 -0.337  0.136
## HrmnyIV-I-V -0.445 -0.445 -0.445  0.158  0.145
## Instrumntpn -0.303 -0.303 -0.303  0.001  0.000  0.000
## Instrmntstr -0.406 -0.406 -0.406  0.000  0.000  0.000  0.630
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

6.2 Examine musicans vs non-musicans for classical ratings

In conclusion, it seems that the results are not sensitive to where we dichotomize (we dichotomize "Selfdeclare" at 2 or 3). We get the same model in the above two cases. And the interaction between HarmonyI-V-VI and MusicansNon-Musicans stays significant and its co-efficient stays the same. Generally, if the subject (listener) is non-Musicans and the harmony is I-V-VI, an average the Classical rating will additionally decrease 0.94297.

```
#Redall from part 4, when we get if we dichotomize "Selfdeclare"
# at 2,the final model we get is lmer8.2:

# Thus, the final model we get is lmer8.2:


# lmer8.2 <-  lmer(Classical ~ Instrument + Harmony + Voice +
# PianoPlay + (1 | Subject) + (Instrument | Subject) + (Harmony |
# Subject) + Harmony:Musicans  - 1,data = data, REML = F,
# control = lmerControl(optimizer = 'bobyqa'))


summary(lmer8.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + PianoPlay + (1 | Subject) +
##     (Instrument | Subject) + (Harmony | Subject) + Harmony:Musicans -
##     1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7481.9   7653.3  -3709.9   7419.9     1834
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6572 -0.5730  0.0197  0.5456  3.6444
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      4.686e-08 0.0002165
```

```
##  Subject.1 (Intercept)      8.410e-01 0.9170649
##            Instrumentpiano  1.668e+00 1.2916508 -0.50
##            Instrumentstring 3.550e+00 1.8840501 -0.99  0.60
##  Subject.2 (Intercept)      1.185e+00 1.0887098
##            HarmonyI-V-IV    5.701e-02 0.2387610  0.79
##            HarmonyI-V-VI    1.561e+00 1.2493991 -0.41  0.01
##            HarmonyIV-I-V    6.647e-02 0.2578256  0.30 -0.15  0.06
##  Residual                   2.466e+00 1.5705084
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                    Estimate Std. Error t value
## Instrumentguitar                     3.6007     0.3163  11.383
## Instrumentpiano                      5.1291     0.3310  15.496
## Instrumentstring                     7.0584     0.3199  22.064
## HarmonyI-V-IV                        0.0549     0.1664   0.330
## HarmonyI-V-VI                        1.5450     0.3100   4.984
## HarmonyIV-I-V                        0.1740     0.1675   1.039
## Voicepar3rd                         -0.3895     0.0891  -4.371
## Voicepar5th                         -0.3100     0.0891  -3.479
## PianoPlay                            0.8488     0.3844   2.208
## HarmonyI-IV-V:MusicansNon-Musicans   0.3504     0.3487   1.005
## HarmonyI-V-IV:MusicansNon-Musicans   0.3066     0.3981   0.770
## HarmonyI-V-VI:MusicansNon-Musicans  -0.7745     0.3957  -1.957
## HarmonyIV-I-V:MusicansNon-Musicans   0.2060     0.3741   0.551
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
##
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# HarmonyI-V-VI:MusicansNon-Musicans -0.94297    0.45830  -2.058


# If we get if we dichotomize "Selfdeclare"
# at 3, the final model we get is lmer10.2:

# lmer10.2 <-  lmer(Classical ~ Instrument + Harmony + Voice + PianoPlay +
#                     (1 | Subject) + (Instrument | Subject) +
#     (Harmony | Subject) +  Harmony:Musicans  - 1,data = data,
#     REML = F, control = lmerControl(optimizer = 'bobyqa'))

summary(lmer10.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + PianoPlay + (1 | Subject) +
##     (Instrument | Subject) + (Harmony | Subject) + Harmony:Musicans -
##     1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
```

```
##   7479.6   7651.0  -3708.8   7417.6     1834
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7449 -0.5720  0.0078  0.5567  3.6404
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      7.550e-14 2.748e-07
##  Subject.1 (Intercept)      8.285e-01 9.102e-01
##            Instrumentpiano  1.667e+00 1.291e+00 -0.51
##            Instrumentstring 3.549e+00 1.884e+00 -0.99  0.60
##  Subject.2 (Intercept)      1.177e+00 1.085e+00
##            HarmonyI-V-IV    3.681e-02 1.919e-01  1.00
##            HarmonyI-V-VI    1.423e+00 1.193e+00 -0.38 -0.38
##            HarmonyIV-I-V    7.480e-02 2.735e-01  0.28  0.28 -0.22
##  Residual                   2.469e+00 1.571e+00
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                                    Estimate Std. Error t value
## Instrumentguitar                    3.46763    0.39864   8.699
## Instrumentpiano                     4.99606    0.40992  12.188
## Instrumentstring                    6.92523    0.40197  17.228
## HarmonyI-V-IV                       0.10283    0.21357   0.482
## HarmonyI-V-VI                       1.94897    0.39025   4.994
## HarmonyIV-I-V                       0.40618    0.21984   1.848
## Voicepar3rd                        -0.38959    0.08914  -4.370
## Voicepar5th                        -0.31024    0.08915  -3.480
## PianoPlay                           0.79593    0.39712   2.004
## HarmonyI-IV-V:MusicansNon-Musicans  0.46053    0.40488   1.137
## HarmonyI-V-IV:MusicansNon-Musicans  0.36340    0.45873   0.792
## HarmonyI-V-VI:MusicansNon-Musicans -0.94297    0.45830  -2.058
## HarmonyIV-I-V:MusicansNon-Musicans  0.04011    0.43434   0.092
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
##
## convergence code: 0
## boundary (singular) fit: see ?isSingular
#
# HarmonyI-V-VI:MusicansNon-Musicans -0.94297    0.45830  -2.058
```

6.3 differences in the things that drive classical, vs. popular, ratings

X16.minus.17 and PianoPlay are significant variable in the model for Classical rating;

For classical rating:

Instrumentstring 7.43247 0.24761 30.017 X16.minus.17 -0.11195 0.04886 -2.291 PianoPlay 0.92342 0.36601 2.523

Fixed effects: HarmonyI-V-VI 4.87101 0.26791 18.182 ++ biggest

Voicecontrary 3.97494 0.23499 16.916 ++ biggest

For popular rating:

```
              Estimate Std. Error t value
```

Instrumentguitar 6.74915 0.21325 31.649

Fixed effects: HarmonyI-V-VI 6.40724 0.22742 28.173 –smallest

Fixed effects: Estimate Std. Error t value Voicecontrary 6.74915 0.21325 31.649 – smallest

```
# Recall that from part 3, we have the final model lmer6.2 as


# lmer6.2 <- lmer(Classical ~ Instrument + Harmony + Voice  +
#     X16.minus.17 +PianoPlay + (1 | Subject) + (Instrument |
#     Subject) + (Harmony | Subject) - 1,data = data,REML = F, control =
#       lmerControl(optimizer = 'bobyqa'))


summary(lmer6.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument + Harmony + Voice + X16.minus.17 + PianoPlay +
##     (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) -      1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7478.9   7633.8  -3711.5   7422.9     1837
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6926 -0.5842  0.0166  0.5449  3.6822
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      5.983e-08 0.0002446
##  Subject.1 (Intercept)      7.395e-01 0.8599395
##            Instrumentpiano  1.667e+00 1.2910689 -0.54
##            Instrumentstring 3.549e+00 1.8838126 -1.00  0.60
##  Subject.2 (Intercept)      1.053e+00 1.0263711
##            HarmonyI-V-IV    5.753e-02 0.2398589  0.87
##            HarmonyI-V-VI    1.869e+00 1.3670010 -0.36  0.05
##            HarmonyIV-I-V    6.998e-02 0.2645403  0.11 -0.12  0.16
##  Residual                   2.467e+00 1.5706033
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                   Estimate Std. Error t value
## Instrumentguitar  3.97494    0.23499  16.916
## Instrumentpiano   5.50315    0.25382  21.681
## Instrumentstring  7.43247    0.24761  30.017
## HarmonyI-V-IV     0.02992    0.10809   0.277
## HarmonyI-V-VI     0.89608    0.21571   4.154
```

```
## HarmonyIV-I-V     0.09070    0.10914   0.831
## Voicepar3rd      -0.38972    0.08910  -4.374
## Voicepar5th      -0.31014    0.08911  -3.480
## X16.minus.17     -0.11195    0.04886  -2.291
## PianoPlay         0.92342    0.36601   2.523
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.667
## Instrmntstr  0.347     0.583
## HrmnyI-V-IV -0.047    -0.043    -0.044
## HrmnyI-V-VI -0.296    -0.274    -0.281     0.241
## HrmnyIV-I-V -0.183    -0.169    -0.173     0.435  0.271
## Voicepar3rd -0.190    -0.176    -0.180    -0.001 -0.001  0.001
## Voicepar5th -0.190    -0.175    -0.180    -0.001 -0.002 -0.001  0.500
## X16.mins.17 -0.309    -0.286    -0.294    -0.001  0.000  0.000  0.000
## PianoPlay   -0.258    -0.238    -0.244     0.001  0.000  0.000  0.000
##             Vcpr5t X16..1
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th
## X16.mins.17  0.000
## PianoPlay    0.000 -0.122
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# Recall that from part 5, we have the final model
# lmer.4.3 <-  lmer(Popular ~ Instrument + Harmony + Voice
# + (1 | Subject) + (Instrument | Subject) +
#     (Harmony | Subject) - 1, data = data,
#     REML = F, control = lmerControl(optimizer = 'bobyqa'))

summary(lmer.4.3)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument + Harmony + Voice + (1 | Subject) + (Instrument |
##     Subject) + (Harmony | Subject) - 1
##    Data: data
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7495.4   7639.2  -3721.7   7443.4     1839
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9873 -0.5734  0.0317  0.5610  3.2941
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
```

```
##  Subject.1 (Intercept)      0.2965   0.5445
##            Instrumentpiano  1.4316   1.1965   -0.65
##            Instrumentstring 2.7342   1.6536   -1.00  0.66
##  Subject.2 (Intercept)      1.5111   1.2293
##            HarmonyI-V-IV    0.1136   0.3370    0.43
##            HarmonyI-V-VI    1.0938   1.0459   -0.30 -0.38
##            HarmonyIV-I-V    0.3314   0.5757   -0.52 -0.81 -0.17
##  Residual                   2.5012   1.5815
## Number of obs: 1865, groups:  Subject, 52
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  6.74915    0.21325  31.649
## Instrumentpiano   5.61683    0.23805  23.595
## Instrumentstring  3.75867    0.25183  14.925
## HarmonyI-V-IV    -0.01981    0.11362  -0.174
## HarmonyI-V-VI    -0.34191    0.17826  -1.918
## HarmonyIV-I-V    -0.25826    0.13071  -1.976
## Voicepar3rd       0.20964    0.08973   2.336
## Voicepar5th       0.24616    0.08973   2.743
##
## Correlation of Fixed Effects:
##             Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
## Instrumntpn  0.655
## Instrmntstr  0.450     0.687
## HrmnyI-V-IV -0.080    -0.071    -0.067
## HrmnyI-V-VI -0.336    -0.301    -0.285     0.136
## HrmnyIV-I-V -0.445    -0.399    -0.377     0.158  0.145
## Voicepar3rd -0.211    -0.188    -0.178    -0.001 -0.001  0.001
## Voicepar5th -0.210    -0.188    -0.178    -0.001 -0.002 -0.001  0.500
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```