Behind Your Taste in Music:
Do You Hear Classical or Popular Music?

Melanie (Menglu) Huang*
mengluh@stat.cmu.edu

8 December 2019

Abstract

We address the question of what influences people's judgement on whether a piece of music is classical or popular. We examine data on 70 undergraduate students who listened to 36 musical stimuli and provided ratings for each musical stimulus, using linear mixed-effects models. We find that factors that influence classical and popular ratings are very different. Instrument exerts the most influence on classical ratings and harmonic motion is the most important factor in popular ratings. Our analysis is limited because of the sample size and certain biased exploratory variables, which could be improved by building a more comprehensive data set.

1 Introduction

Classical and popular music are two categories of music. One might think that classical and popular music are very different. However, categorization of music could be subjective, and it is often difficult to draw a distinct line between classical and popular music.

Despite the difficulty, it is generally very helpful for people who study music or are interested in learning music theory to understand what distinguishes one category of music from another. So, what makes a piece of music more or less classical? What makes it popular? This statistical analysis project seeks to find out important factors that drive people's decision in determining how classical or popular a piece of music sounds. In particular, we will address the following questions:

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

  - Does Instrument exert the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading)?

  - Among the levels of Harmonic Motion, does *I-V-VI* have a strong association (the strongest?) with classical ratings? Does it seem to matter whether the respondent is familiar with one or the other (or both) of the Pachelbel rants/Comedy bits?

  - Among the levels of Voice Leading, does *contrary motion* have a strong (the strongest?) association with classical ratings?

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

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

*Department of Statistics & Data Science, Carnegie Mellon University

2 Methods

The data for this study was collected by Ivan Jimenez, then a visiting scholar at the University of Pittsburgh, and a student Vincent Rossi in 2012. The two researchers presented 36 musical stimuli to 70 undergraduate students from the University of Pittsburgh, who were asked to rate the music on two scales:

Classical	=	How classical does the music sound (1 to 10, 1 = not at all, 10 = very classical)
Popular	=	How popular does the music sound (1 to 10, 1 = not at all, 10 = very popular)

The two scales are independent since a piece of music could be rated as both classical and popular, neither classical nor popular, or classical but not popular (or vice versa). After a closer examination at the two variables, we found some ratings were missing. We decided to remove the missing values because ratings were dependent variables in our model, and any data imputation would not be appropriate. We also found that some responses were not within the 1 to 10 range, which could be a data entry mistake. We changed ratings that were '19' to '10' and ratings that were '0' to '9' because we noticed how close the digit '9' and '0' were on the keyboard. In addition, we found some non-integer ratings. We believed they could be typos and we removed the decimals.

The data set contains other variables related to participants' music knowledge, and we found these variables have different degrees of missingness (see Figure 1 in page 4). Too much missing data can be a problem as we construct our model. Therefore, we considered either mode imputation or dropping variables from our analysis, depending on the level of missingness. As we can see from Figure 1, *X1stInstr* and *X2ndInstr* have more than 50% of missing values, therefore we decided to drop them from our analysis. For variables that have more than 5% but less than 50% of missing values, we removed the missing values.

We started from the variable that had the most missing values, specifically, *ConsNotes*. We checked the proportion of missingness after we removed the missing values for *ConsNotes*, and we repeated the procedure for the next variable. After removing missing values for *ConsNotes*, *KnowAxis*, *NoClass*, and *X1990s2000s*, the rest of variables in the data set has less than 5% missing values. For these variables, *APTheory*, *KnowRob*, and *PachListen*, we did mode imputation. We know any kind of data imputation could be dangerous, but since we are imputing a relatively small proportion of the whole dataset, and mode imputation replaces the missing values with the mode of the observed values for the variables, we believe our measure of data imputation effectively deals with missing values while still largely preserve the essentiality of the data set.

For our analysis, we constructed multilevel models, which allowed slopes and intercepts to vary from subjects to subjects to get the best fit because the 70 participants might have different perceptions of 'classical' and 'popular'. We performed stepwise procedures to automatically select fixed effects by backward elimination. Backward elimination starts with all of the predictor variables in the model, thus it could identify predictors that don't predict well individually but have considerable predictive power as a set. With these predictor variables, we constructed an LMER model by adding a random intercept for each subject. To examine the significance of the random intercept, we performed exact restricted likelihood ratio test (exactRLRT). We then forward fitted random effects to the LMER model using an automated fitting function, which selected random effects by way of log-likelihood ratio testing. Graphically, we looked at conditional residual plots and binned residual plot to check whether the assumptions of our models hold.

Our analysis is made possible by the R language and environment for statistical computing (R Core Team, 2017). The data are available in the file `ratings.csv` on canvas.

All the predictor variables in the data set are listed as follows:

Subject	=	Unique subject ID for each participant)
Harmony	=	4 levles of harmonic motion: $I-V-VI$, $I-VI-V$, $I-V-IV$, $IV-I-V$
Instrument	=	3 levels of instrument = String Quartet, Piano, Electric Guitar
Voice	=	3 levels of voice leading = Contrary Motion, Parallel 3rds Parallel 5ths
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 the previous 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 variables and 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 mbox(0−5, 0 = not at all)
X2ndInstr	=	Same, for second musical instrument

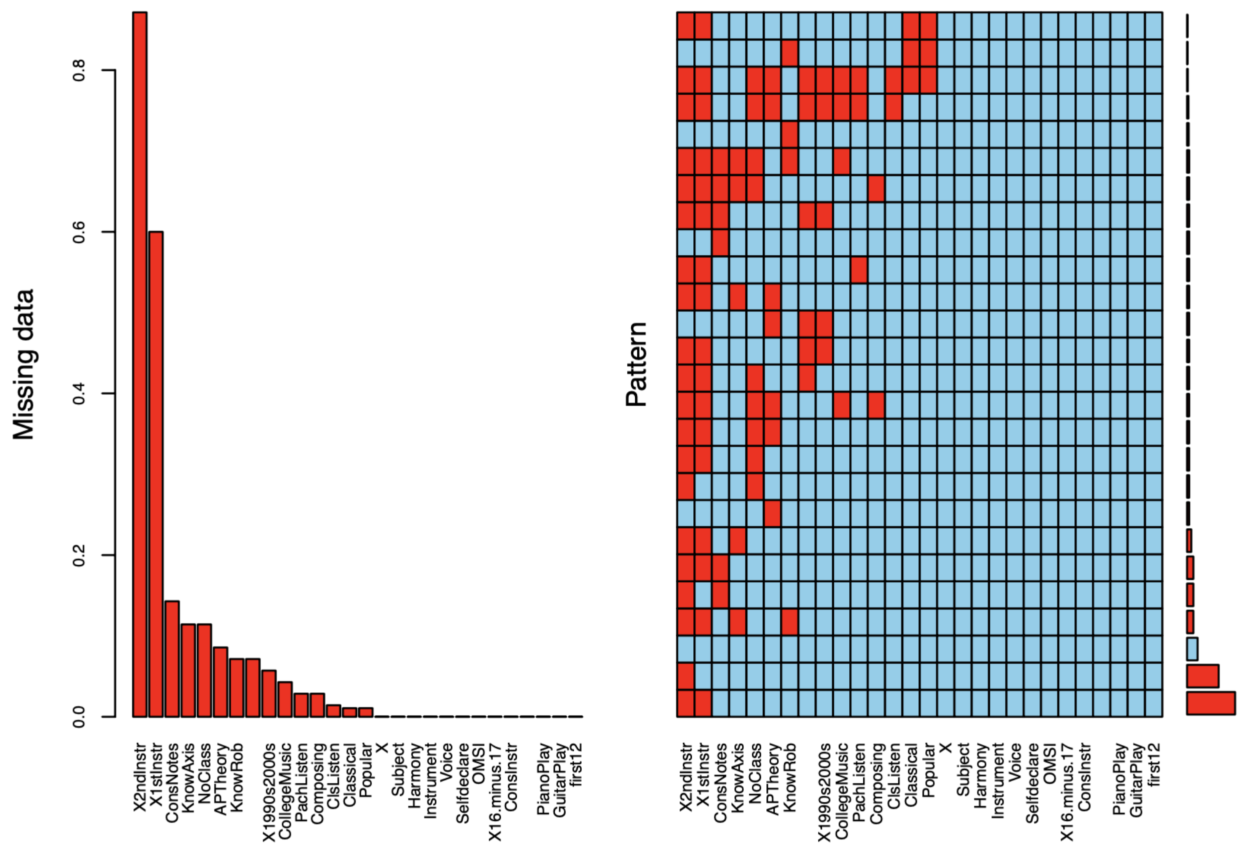


Figure 1: Missing Variable Plot

3 Results

3.1 What Affects Ratings?

In this section, we will present two models, one for classical ratings and one for popular ratings. Because we needed to account for subject-to-subject variability, the final models that we present will be multilevel mixed-effects models that incorporate random effects with fixed effects, constructed through the combination of three main experimental factors and other predictor variables.

3.1.1 EDA and Variable Transformations

Before we introduce our models, we want to discuss some variable transformations that we made based on our exploratory data analysis. Figure 2 on the next page displays a histogram matrix for all the numeric variables in the data set. We noticed that variable *OMSI* was rightly skewed and we decided to log-transform the variable in order to meet the normality assumption of frequency distribution in inferential statistics. As shown in Figure 3 on page 6, the frequency distribution of *OMSI* is more normally distributed after log transformation.

As part of our exploratory data analysis, we also took a look at the internal structure of the data set (See Appendix 1 for details). We found that some factor variables, specifically, *APTheory* and *CollegeMusic*,

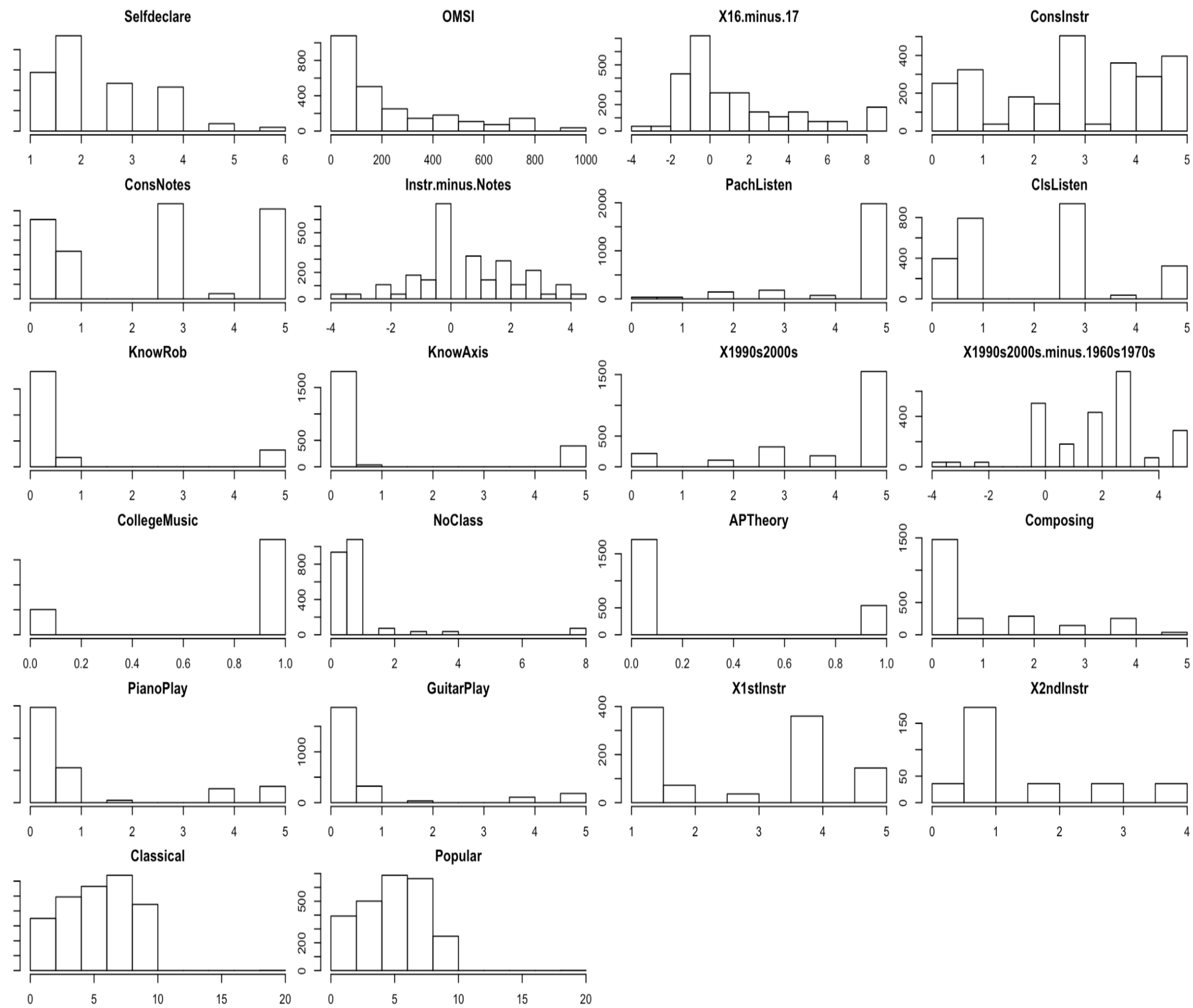


Figure 2: Histogram of Numeric Variables

were coded as integer. We factorized these two variables because treating them as continuous variables would assume the difference between adjacent levels of the variables is in some sense constant. We wanted to properly treat them as factor variables.

In addition, we explored the relationships between ratings and the three main design factors by constructing box-plots in Figure 4, from which we could see instrument not only had the most different mean ratings across its three levels, but for each level of instrument, its mean classical rating and mean popular rating were also very different. This finding lends supports to researchers' hypothesis that instrument exerts the strongest influence among the three design factors, but we need a complete model to make statistically valid conclusion.

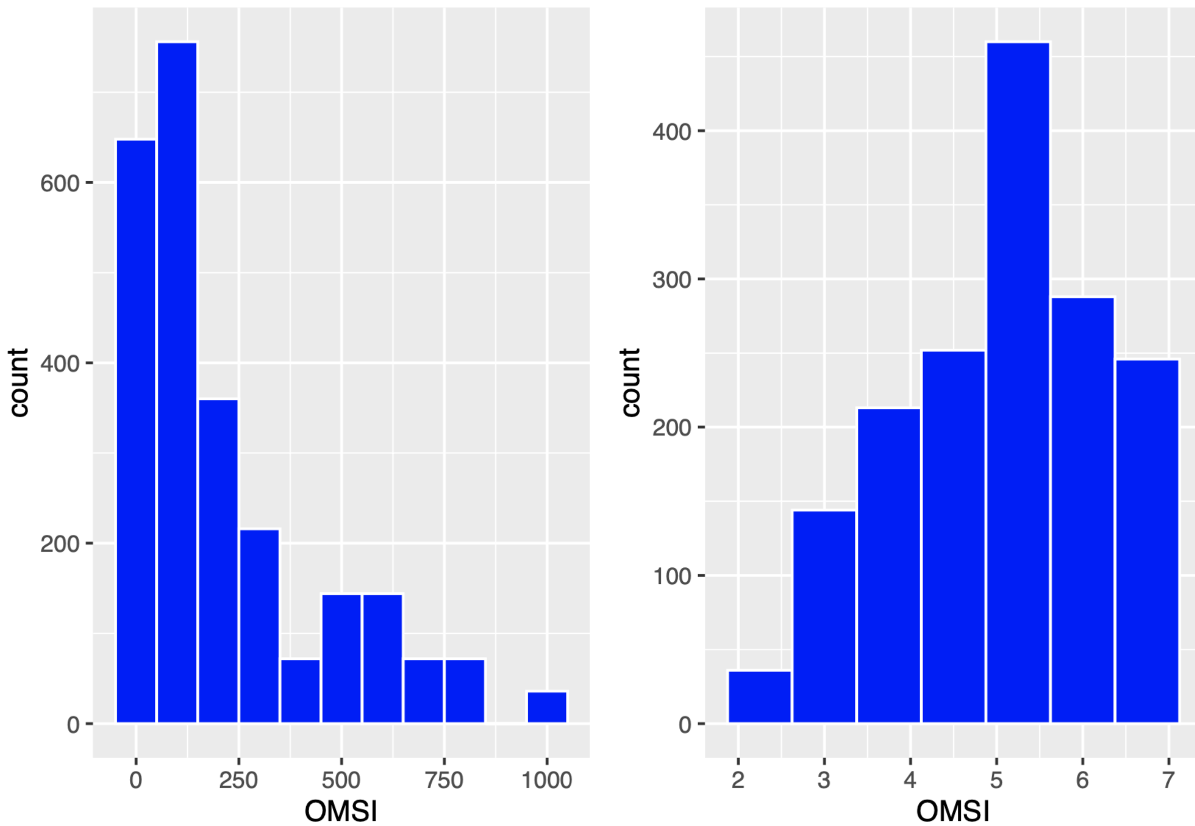


Figure 3: Histogram of OMSI (before and after log-transformation)

3.1.2 Constructing Multi-level Models

To understand the influence of the three main design factors on classical and popular ratings, we first fitted a conventional linear model on instrument, harmonic motion, voice leading, and all other predictor variables in the data set. We then utilized the stepAIC function in the R language. We used Bayesian information criterion (BIC) as our selection criterion by setting the number of degrees of freedom for the penalty to be $log(n)$ where n is the number of observations in the data set. We chose BIC because Akaike information criterion (AIC) tends to produce a more complex model and BIC will give a more interpretable model. During the variable selection process, one or more of the three main design factors was dropped, but we manually added it back to the model because we wanted to examine their influence on ratings. We also considered interacting the three main design factors, but stepAIC did not select any interaction term in the final models for either classical ratings or popular ratings (See Appendix 11 and 22 for more details).

After we found the fixed effects, we considered adding a random intercept to the model. Our reasoning is that people might have different understanding of what is classical and what is popular. The random intercept can account for personal biases in ratings by allowing intercept to vary from subject to subject. We compared the model with a random intercept and the one without random intercept using Analysis of Variance (ANOVA) test. The result showed that the model with a random intercept had lower AIC, BIC, and log likelihood. We also performed exact restricted likelihood ratio test (exactRLRT) to examine the significance of a random intercept, and we arrived at the same conclusion that adding a random intercept

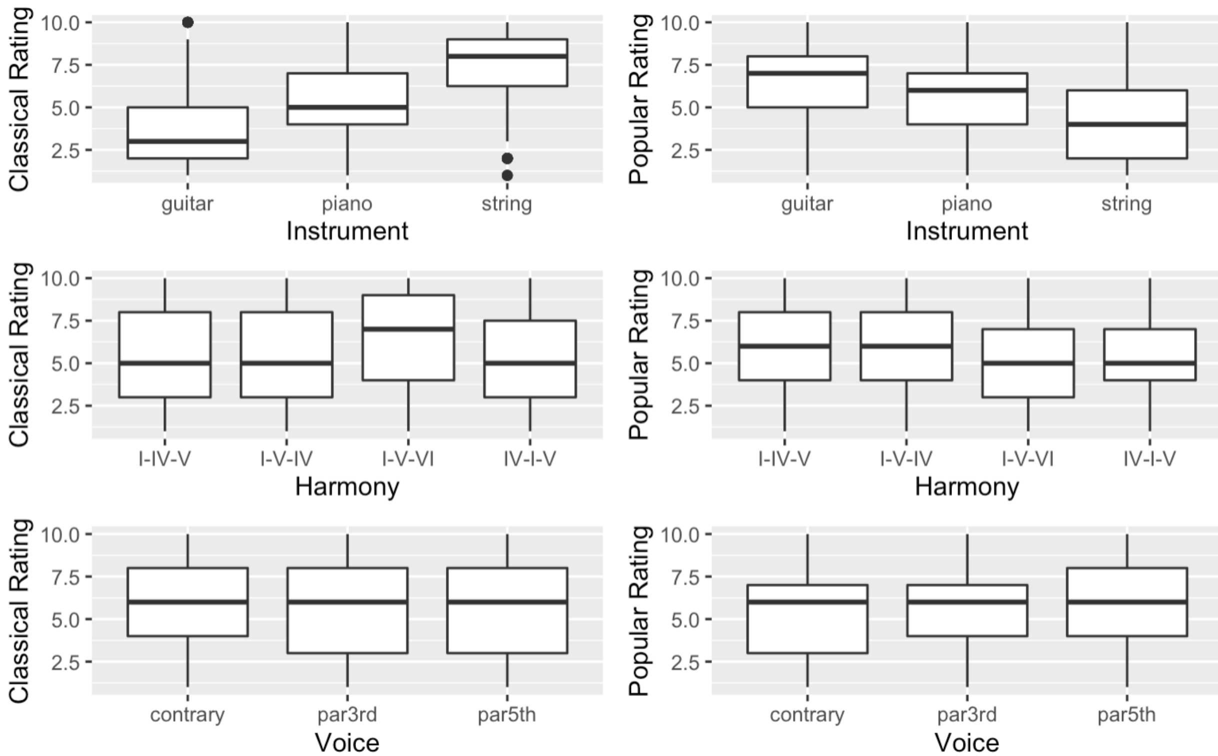


Figure 4: Box-plots for Ratings vs. Instrument, Harmony, Voice

improved the fit of our model (See Appendix 13 for more details). After accounting for the random effect, some variables became not statistically significant since the t-value of their estimates are either greater than -2 or less than 2 (The usual cutoff for significance test is ± 1.96, but since we have a lot of variables in the data set, it would be safe to increase the threshold). This is because adding variance components often increases the standard errors of fixed effect estimates. Therefore we dropped the insignificant fixed effects from the model.

Next we wanted to add more random effects to increase the fit of our model. From Figure 4, we can see that different instruments receive very different classical and popular ratings. The box-plot makes sense because guitar is the mainstream instrument in popular music, therefore music played with guitar will more likely to be associated with popular music. Similarly, music played by a string quartet could easily get higher classical ratings. However, participants could vary in the degree to which they are inclined to categorize music played by certain instruments. For example, a participant who is more knowledgeable in music theory might have a different understanding of what constitutes classical music and might not rate these music solely based on instruments. By the same token, people might have tendency to associate certain levels of harmonic motion with classical music or popular music, and this inclination varies from individual to individual. The same argument applies to different levels of voice. Therefore, we decided to add three random effects, (Harmony|Subject), (Instrument|Subject) and (Voice|Subject) to in order to better capture variability within subject.

We used LMERConvenienceFunctions in R and selected random effects by forward fitting. Our final models for classical ratings and popular ratings are presented as below:

$$Classical \sim \begin{cases} Harmony+Instrument+Voice+Selfdeclare+ClsListen+X1990s2000s+ \\ Composing+PianoPlay+(1|Subject)+(Harmony|Subject)+(Instrument|Subject) \end{cases} \tag{1}$$

$$Popular \sim \begin{cases} Harmony+Instrument+Voice+X16.minus.17+PachListen+(1|Subject)+ \\ (Harmony|Subject)+(Instrument|Subject) \end{cases} \tag{2}$$

We examined the validity of our model by checking conditional residuals and binnedplot, see Appendix for details.

3.1.3 What We Learn from the Models

To answer which design factors among *Instrument*, *Harmony*, *Voice* exerts the strongest influence on classical ratings, we examine the summary output of model (1). As as we can see from Table 1 on page 10, the coefficient estimates for *Instrument* and *Voice* are statistically significant. The coefficient estimates shown in the table are relative to the base level, which in this case of instrument, is guitar. From the values of coefficient estimates, we can tell that music played with string is expected to be approximately 3.5 points higher in classical ratings than music played with guitar, holding all other variables constant. Table 2 shows fixed effects and random effects for all levels of main design factors for classical ratings. String instrument appears to have the largest coefficient estimate (approximately 7.07), followed by piano with an estimate of 5.15. The coefficient for guitar is not particularly high, but it's among the higher ones. The standard deviation is small for each level of instrument. Therefore, we expect within two standard deviations of the distribution, or for approximately 95% of the participants, instrument has positive association with classical ratings.

Furthermore, standard deviation tells us the variation within subject for each level of random effects. For guitar, the variation within subject is approximately 0.83 point. For piano, the variation within subject is 0.96 point. For string, the variation within subject is 1.06 points. These show that there is a larger variation within group for string instrument than for guitar. In other words, classical ratings for music played with string instruments are expected to have larger variations than music played with guitar. This could be due to the varying degrees of music knowledge participants have and their perceptions of classical music associated with different instruments.

Back to Table 1, it seems like only one level of *Harmony*, specifically *I-V-VI*, is statistically significant. Therefore, among the four levels of *Harmony*, *I-V-VI* has the strongest association with classical ratings. Moreover, the coefficient estimates for *Harmony* is relatively small. For *Voice*, its coefficient estimates are significant. However, compared to contrary motion, music that has parallel 3rds as voice leading is expected to be approximately 0.36 point lower in classical ratings, holding all other variables constant. Hence, contrary motion has the strongest association with classical ratings. However, this is a relatively small effect when compared to the estimates for instrument. Therefore, based on our model for classical ratings, we believe instrument has stronger influence than harmonic motion and voice leading.

For popular ratings, instrument is not the most influential factor. As we can see from Table 3, the coefficient estimates for *Instrument* is statistically significant. For *Harmony*, only one level, *IV-I-V* is significant. For *Voice*, only *Voicepar5th* has a statistically significant coefficient estimate. Taking a closer look at *Instrument*, the coefficient estimates for piano and string are negative. This means that compared to the base case, or guitar, music played with piano is expected to receive a popular rating that's approximately 1.15 points lower, and music played with string instrument is expected to received a popular rating that's approximately 2.67 points lower, holding all other variables constant.

Table 4 shows fixed effects and random effects for all three levels of instruments. For guitar, the variation within subject is approximately 0.73 point. For piano, the variation within subject is 1.25 points. For string instrument, the variation within subject is 1.29 points. Again, there is a larger variation within subject for string instrument than for guitar. We would expect popular ratings for music played with string instruments to have larger variations than ratings for music played with guitar. This could be due to the same reason as commented above for classical ratings. Some participants might tend to strongly associate string instruments with classical music. As a result, they would rate these music as less popular.

By looking at the coefficient estimates for all levels of main design factors in Table 4, we can see that *HarmonyI-V-VI* is the most significant variable with a coefficient estimate of 9.11. The second most influential variable is *HarmonyI-V-IV*, which has a coefficient estimate of 8.45. Because the coefficient estimates for *Instrument* are not particularly large, our model suggests that *Instrument* does not exert the strongest influence on popular ratings.

Our final models for popular ratings and classical ratings do not include the variables *KnowRob* and *KnowAxis*. However, we wanted to know whether the participants are familiar with Rob Paravonian's Pachelbel Rant or Axis of Evil's Comedy bit matters in determining ratings. Therefore, we tried updating model (1) and model (2) by interacting *KnowRob* with *Harmony* and *KnowAxis* with *Harmony*. We found that adding the interaction terms did not improve the fit of our final models. (See Appendix 15 for details). Therefore, we conclude that participants' familiarity with Rob Paravonian's Pachelbel Rant or Axis of Evil's Comedy bit does not seem to affect ratings.

3.2 Musicians vs. non-Musicians

To investigate whether there are differences in the way that musicians and non-musicians identify classical music, we will present another model for classical ratings in this section.

3.2.1 Constructing Multi-level Model

We chose to dichotomize the variable *Selfdeclare* at responses less than or equal to 2 so that about half the participants are categorized as self-declared musicians and half not. In our new factor variable, which we call *Musician*, we assigned 0 to the non-musicians and 1 to musicians. We followed the same step in Section 3.1.2, but we added interaction terms between *Musician* and all the other predictor variables in the data set. Then we selected fixed effects by backward elimination and random effects by forward fitting, using automatic functions in R. Our model is presented as below:

$$Classical \sim \begin{cases} Musician+Harmony+Instrument+Voice+ConsNotes+PachListen+KnowRob+APTheory+ \\ Composing+PianoPlay+GuitarPlay+Musician : Harmony+Musician : Instrument+ \\ Musician : X16.munis.17+Musician : ConsNotes+Musician : PachListen+ \\ Musician : ClsListen+Musician : KnowAxis+Musician : X1990s2000s+ \\ Musician : NoClass+(1|Subject)+(Harmony|Subject)+(Instrument|Subject) \end{cases} \tag{3}$$

	Model (1) Classical Ratings
(Intercept)	3.54^{***}
	(0.59)
HarmonyI-V-IV	0.01
	(0.12)
HarmonyI-V-VI	0.87^{***}
	(0.22)
HarmonyIV-I-V	0.10
	(0.11)
Instrumentpiano	1.61^{***}
	(0.24)
Instrumentstring	3.53^{***}
	(0.29)
Voicepar3rd	-0.36^{***}
	(0.10)
Voicepar5th	-0.34^{***}
	(0.10)
Selfdeclare	-0.44^{**}
	(0.17)
ClsListen	0.21^{*}
	(0.10)
X1990s2000s	0.15
	(0.09)
Composing	0.22
	(0.12)
PianoPlay	0.31^{**}
	(0.10)
AIC	6601.85
BIC	6769.31
Log Likelihood	-3269.93
Num. obs.	1639
Num. groups: Subject	46
Var: Subject (Intercept)	0.00
Var: Subject.1 (Intercept)	0.94
Var: Subject.1 HarmonyI-V-IV	0.06
Var: Subject.1 HarmonyI-V-VI	1.62
Var: Subject.1 HarmonyIV-I-V	0.04
Cov: Subject.1 (Intercept) HarmonyI-V-IV	0.09
Cov: Subject.1 (Intercept) HarmonyI-V-VI	-0.42
Cov: Subject.1 (Intercept) HarmonyIV-I-V	-0.01
Cov: Subject.1 HarmonyI-V-IV HarmonyI-V-VI	0.21
Cov: Subject.1 HarmonyI-V-IV HarmonyIV-I-V	0.00
Cov: Subject.1 HarmonyI-V-VI HarmonyIV-I-V	0.10
Var: Subject.2 (Intercept)	0.69
Var: Subject.2 Instrumentpiano	2.17
Var: Subject.2 Instrumentstring	3.48
Cov: Subject.2 (Intercept) Instrumentpiano	-0.96
Cov: Subject.2 (Intercept) Instrumentstring	-1.52
Cov: Subject.2 Instrumentpiano Instrumentstring	1.76
Var: Residual	2.52

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1: Classical Ratings

	Fixed Effects		Random Effects
	All Levels of Harmony	Variance	Standard Deviation
HarmonyI-IV-V	3.54^{***}	0.94	0.97
	(0.59)		
HarmonyI-V-IV	3.54^{***}	1.16	1.08
	(0.59)		
HarmonyI-V-VI	4.41^{***}	1.71	1.31
	(0.60)		
HarmonyIV-I-V	3.64^{***}	0.94	0.97
	(0.59)		
	All Levels of Instrument	Variance	Standard Deviation
Instrumentguitar	3.54^{***}	0.69	0.83
	(0.59)		
Instrumentpiano	5.15^{***}	0.93	0.96
	(0.60)		
Instrumentstring	7.07^{***}	1.13	1.06
	(0.60)		
	All Levels of Voce Leading		
Voicecontrary	3.54^{***}		
	(0.59)		
Voicepar3rd	3.18^{***}		
	(0.59)		
Voicepar5th	3.19^{***}		
	(0.59)		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 2: All Levels of Main Design Factors for Classical Ratings (Model (1))

	Model (2) Popular Ratings
(Intercept)	8.42^{***}
	(0.82)
HarmonyI-V-IV	0.03
	(0.12)
HarmonyI-V-VI	−0.31
	(0.17)
HarmonyIV-I-V	-0.30^{*}
	(0.14)
Instrumentpiano	-1.15^{***}
	(0.22)
Instrumentstring	-2.67^{***}
	(0.30)
Voicepar3rd	0.12
	(0.10)
Voicepar5th	0.21^{*}
	(0.10)
X16.minus.17	0.14^{**}
	(0.04)
PachListen	-0.38^{*}
	(0.17)
AIC	6659.95
BIC	6811.20
Log Likelihood	-3301.98
Num. obs.	1639
Num. groups: Subject	46
Var: Subject (Intercept)	0.00
Var: Subject.1 (Intercept)	0.79
Var: Subject.1 HarmonyI-V-IV	0.09
Var: Subject.1 HarmonyI-V-VI	0.69
Var: Subject.1 HarmonyIV-I-V	0.29
Cov: Subject.1 (Intercept) HarmonyI-V-IV	0.07
Cov: Subject.1 (Intercept) HarmonyI-V-VI	0.03
Cov: Subject.1 (Intercept) HarmonyIV-I-V	-0.28
Cov: Subject.1 HarmonyI-V-IV HarmonyI-V-VI	0.01
Cov: Subject.1 HarmonyI-V-IV HarmonyIV-I-V	-0.14
Cov: Subject.1 HarmonyI-V-VI HarmonyIV-I-V	-0.16
Var: Subject.2 (Intercept)	0.54
Var: Subject.2 Instrumentpiano	1.81
Var: Subject.2 Instrumentstring	3.78
Cov: Subject.2 (Intercept) Instrumentpiano	-0.39
Cov: Subject.2 (Intercept) Instrumentstring	-1.33
Cov: Subject.2 Instrumentpiano Instrumentstring	1.84
Var: Residual	2.65

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3: Popular Ratings

	Fixed Effects		Random Effects
	All Levels of Harmony	Variance	Standard Deviation
HarmonyI-IV-V	8.42^{***}	0.79	0.89
	(0.82)		
HarmonyI-V-IV	8.45^{***}	1.01	1.00
	(0.82)		
HarmonyI-V-VI	9.11^{***}	1.54	1.24
	(0.83)		
HarmonyIV-I-V	8.13^{***}	0.52	0.72
	(0.82)		
	All Levels of Instrument	Variance	Standard Deviation
Instrumentguitar	8.42^{***}	0.54	0.73
	(0.82)		
Instrumentpiano	7.27^{***}	1.56	1.25
	(0.83)		
Instrumentstring	5.75^{***}	1.66	1.29
	(0.83)		
	All Levels of Voce Leading		
Voicecontrary	3.54^{***}		
	(0.59)		
Voicepar3rd	3.18^{***}		
	(0.59)		
Voicepar5th	3.19^{***}		
	(0.59)		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4: All Levels of Main Design Factors for Popular Ratings (Model (2))

3.2.2 What We Learn from our Model

In this section we will focus on discussing the interaction terms between *Musician* and other predictor variables. Figure 5 displays the fixed effects of Model (3). We would interpret random effects of the model the same way as we interpret before. For full summary of the model, please refer to Appendix 35.

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	1.226275	0.841907	1.457
Musician1	8.214199	1.523991	5.390
HarmonyI-V-IV	-0.008889	0.156950	-0.057
HarmonyI-V-VI	0.309182	0.268571	1.151
HarmonyIV-I-V	0.102222	0.152832	0.669
Instrumentpiano	1.880440	0.317007	5.932
Instrumentstring	3.956100	0.384018	10.302
Voicepar3rd	-0.359320	0.096057	-3.741
Voicepar5th	-0.342563	0.096060	-3.566
ConsNotes	-0.323952	0.071240	-4.547
PachListen	0.746766	0.151454	4.931
KnowRob	0.120258	0.068688	1.751
APTheory1	0.503903	0.257468	1.957
Composing	0.784157	0.103025	7.611
PianoPlay	0.529306	0.063341	8.356
GuitarPlay	0.243276	0.124741	1.950
Musician1:HarmonyI-V-IV	0.034386	0.233395	0.147
Musician1:HarmonyI-V-VI	1.229355	0.397931	3.089
Musician1:HarmonyIV-I-V	0.002972	0.227149	0.013
Musician1:Instrumentpiano	-0.618085	0.470303	-1.314
Musician1:Instrumentstring	-0.928322	0.568241	-1.634
Musician0:X16.minus.17	0.001766	0.050141	0.035
Musician1:X16.minus.17	-0.527580	0.075942	-6.947
Musician1:ConsNotes	1.015596	0.160933	6.311
Musician1:PachListen	-1.509119	0.257598	-5.858
Musician0:ClsListen	-0.084401	0.088439	-0.954
Musician1:ClsListen	-1.142037	0.179967	-6.346
Musician0:KnowAxis	-0.015310	0.068956	-0.222
Musician1:KnowAxis	0.368075	0.074069	4.969
Musician0:X1990s2000s	-0.185642	0.095008	-1.954
Musician1:X1990s2000s	-0.724623	0.156763	-4.622
Musician0:NoClass	0.150074	0.214673	0.699
Musician1:NoClass	-0.544017	0.104251	-5.218

Figure 5: Summary Output for Model (3) - Fixed effects only

As we can see from Figure 5, the interaction between *Musician* and *Harmony*, *X16.minus.17*, *ConsNotes*, *PachListen*, *ClsListen*, *KnowAxis*, *X1990s2000s*, *NoClass* estimates have t-values significantly different from zero. Take *Harmony* as an example, self-declared musicians are expected to rate music stimuli that are *HarmonyI-V-VI* more classical than music stimuli that are *HarmonyI-IV-V* by about 1.2 points. Since *I-V-VI* is the beginning progression for Pachelbel's *Canon in D*, a widely heard classical piece, it is reasonable that people tend to associate *I-V-VI* with classical music. In comparison, self-declared non-musicians might have heard Pachelbel's *Canon in D*, but they might not be able to discern the harmonic progression *I-V-VI*, not to mention associating it with classical music.

Furthermore, we found that if the participants were musicians, the better their abilities to distinguish classical and popular music, the lower classical ratings they will give to music stimuli. Similarly, the degree to which self-declared musicians concentrated on notes while listening, their familiarity with Pachelbel's *Canon in D* and Axis of Evil's Comedy bit, the frequency they listened to pop and rock from the 90's and 2000's, and numbers of music classes taken in college all have negative associations with classical ratings. To some extents, these variables demonstrate how knowledgeable participants are in music. Based on our model, it appears that the more self-declared musicians know about music, they tend to give lower classical ratings. We do not observe the same relationship for non-musicians, hence there is a difference in the way that musicians and non-musicians identify classical music.

3.3 Classical vs. Popular ratings

In this section, we will look into the differences in the variables that drive classical and popular ratings by referring back to model (1) and model (2):

$$Classical \sim \begin{cases} Harmony + Instrument + Voice + Selfdeclare + ClsListen + X1990s2000s + \\ Composing + PianoPlay + (1|Subject) + (Harmony|Subject) + (Instrument|Subject) \end{cases} \tag{1}$$

$$Popular \sim \begin{cases} Harmony + Instrument + Voice + X16.minus.17 + PachListen + (1|Subject) + \\ (Harmony|Subject) + (Instrument|Subject) \end{cases} \tag{2}$$

We will first compare the three main design factors. For classical ratings, *Instrument* appears to be the most influential variables. If the music stimuli are played with guitar, we expect they will be rated as less classical than music stimuli that are played with string instrument. For popular ratings, *HarmonyI-V-VI* exerts the most influence. For both classical and popular ratings, *Voice* seems to be the least influential among the three design factors as we can see from its coefficient estimates in Table 2 and Table 4. Moreover, both models let *Harmony* vary by subject and *Instrument* vary by subject, but did not have *Voice* varying by subject. This result suggests that there is not significant variation in participants' understanding of voice leading.

Next we will look at other fixed effects in the two models that might trigger differences in ratings. For classical ratings, *Selfdeclare*, *ClsListen*, and *PianoPlay* are statistically significant variables. *ClsListen* has a positive coefficient estimate, which means the more participants listen to classical music, the more classical they would rate music stimuli. Last but not least, for an additional level of skillfulness a participant is in playing piano, classical rating will increase by approximately 0.31 point, holding all other variables constant. This makes sense because many advanced piano music fall under the classical category. As one gets more skillful in piano, he or she might be biased toward classical music. Alternatively, a participant could be more comfortable with classical music because he or she understands it better, thereby giving higher classical ratings.

For popular ratings, *X16.minus.17* and *PachListen* are significant fixed effects. *X16.minus.17* is positively associated with popular ratings, which means for an additional score an individual has at distinguishing classical and popular music, popular ratings is expected to increase by 0.14, holding all other variables constant. Although this variable is significant, it is a relatively small value. This suggests that even if the participant really knows how to tell classical music apart from popular music, popular ratings will increase but only by a relatively small amount. This actually makes sense intuitively because for a person who knows how to distinguish classical and popular music, he or she should deem the two categories of music to be equal instead of being biased. The fact that *X16.minus.17* is positively associated with popular ratings might be due to the small sample size of our dataset or the relatively young age of the participants. *PachListen* is negatively associated with popular ratings. This suggests that the more familiar a participant is with Pachelbel's *Canon in D*, the less popular he or she will rate the music stimuli.

Based on our analysis, it appears that factors that drive classical and popular ratings are very different.

4 Discussion

From our previous analysis, we found that *Instrument, Harmony*, and *Voice* have different effects on classical and popular ratings. For classical ratings, *Instrument* has the strongest influence among the three main design factors. Among the three levels of *Instrument*, string instrument has the largest coefficient estimate, followed by piano and guitar. This result agrees with Figure 4 from Section 3.1.1, where string instrument has the highest mean classical ratings and guitar has the lowest mean classical ratings. For popular ratings, *Instrument* does not exert the most influence. Our model indicates that *HarmonyI-V-VI* is the most influential variable, followed by *HarmonyI-V-IV*.

Among the four levels of Harmonic Motion, *HarmonyI-V-VI* has the strongest association with classical ratings. This level of Harmonic Motion is also the most influential variable in the model for popular ratings. As for the variables *KnowRob* and *KnowAxis*, we tried fitting interaction between *Harmony* and these two variables. However, it did not improve the fit of our original models. Hence, we conclude that participants' familiarity with Rob Paravonian's Pachelbel Rant or Axis of Evil's Comedy bit does not seem to affect ratings. Among the three levels of Voice Leading, contrary motion has the strongest association with classical ratings.

We found that musicians and non-musicians perceive classical music differently. These differences are particularly obvious when we interact *Musicians* with *X16.minus.17*, *ConsNotes*, *PachListen*, *ClsListen*, *KnowAxis*, *X1990s2000s*, and *NoClass* one by one. We found that the more knowledgeable these self-declared musicians are, the more likely that they would give lower classical ratings to music stimuli.

Based on our models, classical and popular ratings seem to be driven by very different factors. In the final model that we constructed for classical ratings, besides the three main design factors, *Selfdeclare*, *ClsListen*, and *PianoPlay* are statistically significant. Whereas for popular ratings, variables *X16.minus.17* and *PachListen* are significant. *ClsListen* and *PachListen* are both about classical music. It is intuitive that when one of them is in the model for classical ratings, it is positively associated with the dependent variable. Whereas the other one is negatively associated with popular ratings.

All the variables that are significant in the models more or less relate to participants' music knowledge except for *Selfdeclare*. In some sense, *Selfdeclare* measures how much participants know about music. However, it entirely depends on participants' self-evaluation, which is very subjective. If we compare *Selfdeclare* to *OMSI*, we will see that some people who declare themselves as non-musicians (*Selfdeclare* ≤ 2) actually have scored higher in *OMSI* than people who declare themselves as musicians (*Selfdeclare* ≥ 3). In other words, there are inconsistencies in terms of who qualifies as 'musicians' and who doesn't, which could dampen the result of our analysis.

Moreover, our study is hampered by the many missing values in the data set. We have dropped two variables from our analysis because there are too many missing values. We have removed missing values in some variables, which further decreases the sample size of our data. Furthermore, we performed mode imputation on the data set. Although it was only for a small amount of data, mode imputation might cause bias because single imputation tends to underestimate error variation.

For future study, it would be better to have a complete data set. In addition, since certain variables in the current data set are either too subjective or useless, researchers should consider collecting more data and adding new explanatory variables. Based on this limited analysis, our results indicate that musicians and non-musicians view music differently. Perhaps a more appropriate variable such as occupation, which could be music-related or otherwise, might improve the biases in the study.

References

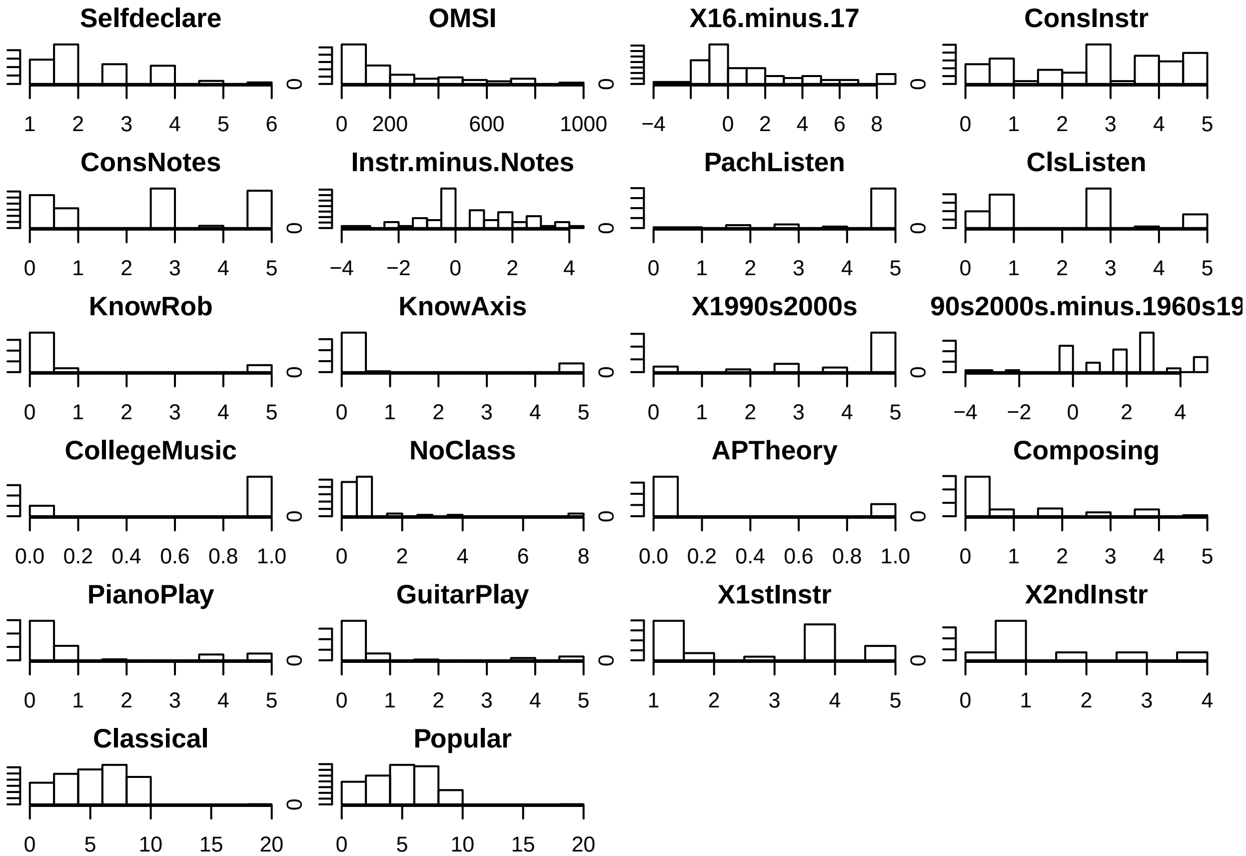
R Core Team (2017), *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

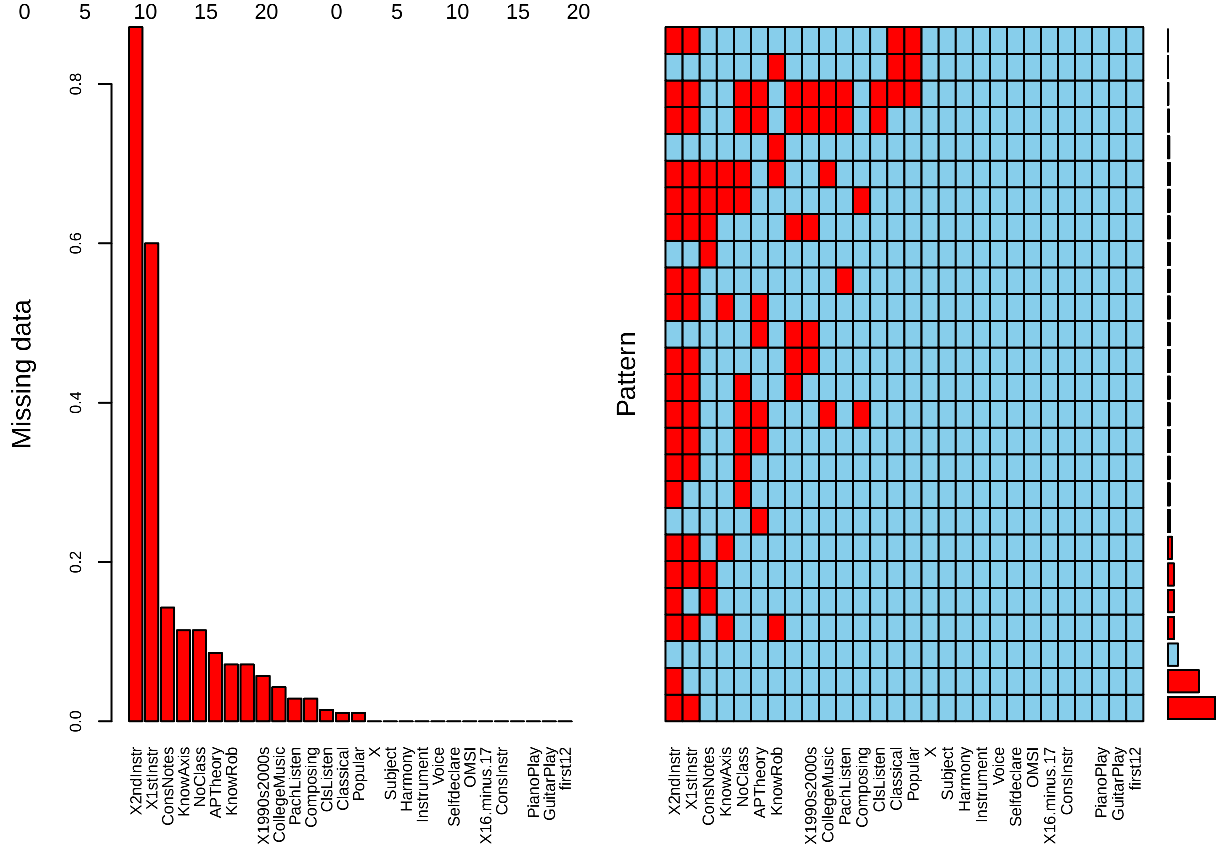
Technical Appendix

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

```
## 'data.frame':    2520 obs. of  28 variables:
##  $ X                          : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Subject                    : Factor w/ 70 levels "15","16","17",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ Harmony                    : Factor w/ 4 levels "I-IV-V","I-V-IV",..: 1 1 1 1 1 1 1 1 1 2 ...
##  $ Instrument                 : Factor w/ 3 levels "guitar","piano",..: 1 1 1 2 2 2 3 3 3 1 ...
##  $ Voice                      : Factor w/ 3 levels "contrary","par3rd",..: 1 2 3 1 2 3 1 2 3 1 ...
##  $ Selfdeclare                : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ OMSI                       : int  734 734 734 734 734 734 734 734 734 734 ...
##  $ X16.minus.17               : num  5 5 5 5 5 5 5 5 5 5 ...
##  $ ConsInstr                  : num  4.33 4.33 4.33 4.33 4.33 4.33 4.33 4.33 4.33 4.33 ...
##  $ ConsNotes                  : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ Instr.minus.Notes          : num  -0.67 -0.67 -0.67 -0.67 -0.67 -0.67 -0.67 -0.67 -0.67 -0.67 ..
##  $ PachListen                 : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ ClsListen                  : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ KnowRob                    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ KnowAxis                   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ X1990s2000s                : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ X1990s2000s.minus.1960s1970s: int  2 2 2 2 2 2 2 2 2 2 ...
##  $ CollegeMusic               : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ NoClass                    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ APTheory                   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ Composing                  : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ PianoPlay                  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ GuitarPlay                 : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ X1stInstr                  : int  4 4 4 4 4 4 4 4 4 4 ...
##  $ X2ndInstr                  : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ first12                    : Factor w/ 3 levels "guitar","piano",..: 3 3 3 3 3 3 3 3 3 3 ...
##  $ Classical                  : num  3 3 1 3 2 8 10 6 5 1 ...
##  $ Popular                    : num  9 7 8 7 8 3 1 4 5 8 ...
```

```
numeric_sub_data <- data.frame(data[,c(6:25,27,28)])

# make histogram for all the numeric variables that have selected
par(mfrow = c(6,4))
par(mar=c(2,1,2,1))
for (i in c(1:22)) {
  hist(numeric_sub_data[,i], main = names(numeric_sub_data)[i], xlab = "")
}

# missing variable plot
mice_plot <- aggr(data, c('navyblue', 'yellow'),
                  numbers=TRUE, sortVars=TRUE,
                  labels=names(data), cex.axis=.7,
                  gap=3, ylab=c('Missing data', 'Pattern'))
```


Selfdeclare
OMSI
X16.minus.17
ConsInstr
ConsNotes
Instr.minus.Notes
PachListen
ClsListen
KnowRob
KnowAxis
X1990s2000s
90s2000s.minus.1960s19
CollegeMusic
NoClass
APTheory
Composing
PianoPlay
GuitarPlay
X1stInstr
X2ndInstr
Classical
Popular


Missing data
Pattern
X2ndInstr
X1stInstr
ConsNotes
KnowAxis
NoClass
APTheory
KnowRob
X1990s2000s
CollegeMusic
PachListen
Composing
ClsListen
Classical
Popular
X
Subject
Harmony
Instrument
Voice
Selfdeclare
OMSI
X16.minus.17
ConsInstr
PianoPlay
GuitarPlay
first12

##

```
##  Variables sorted by number of missings:
##                         Variable      Count
##                        X2ndInstr 0.87142857
##                        X1stInstr 0.60000000
##                        ConsNotes 0.14285714
##                         KnowAxis 0.11428571
##                          NoClass 0.11428571
##                         APTheory 0.08571429
##                          KnowRob 0.07142857
##  X1990s2000s.minus.1960s1970s 0.07142857
##                      X1990s2000s 0.05714286
##                     CollegeMusic 0.04285714
##                       PachListen 0.02857143
##                        Composing 0.02857143
##                        ClsListen 0.01428571
##                        Classical 0.01071429
##                          Popular 0.01071429
##                                X 0.00000000
##                          Subject 0.00000000
##                          Harmony 0.00000000
##                       Instrument 0.00000000
##                            Voice 0.00000000
##                      Selfdeclare 0.00000000
##                             OMSI 0.00000000
##                     X16.minus.17 0.00000000
##                        ConsInstr 0.00000000
##                Instr.minus.Notes 0.00000000
##                        PianoPlay 0.00000000
##                       GuitarPlay 0.00000000
##                          first12 0.00000000
```

```
# proportion of missingness, used to check missing values
pMissing <- function(x){sum(is.na(x))/length(x)*100}

######## Data Cleaning #######
newdata <- data

# Get rid of NAs in Classical and Popular
newdata <- newdata[!is.na(newdata$Classical), ]
newdata <- newdata[!is.na(newdata$Popular), ]
# Get rid of decimal in Classical and Popular
levels(as.factor(data$Classical))
```

```
##  [1] "0"   "1"   "2"   "3"   "3.5" "4"   "4.2" "4.6" "5"   "6"   "7"   "8"
## [13] "9"   "9.5" "10"  "19"
```

```
levels(as.factor(data$Popular))
```

```
##  [1] "0"   "1"   "2"   "3"   "3.5" "4"   "4.2" "4.6" "5"   "6"   "6.8" "7"
## [13] "8"   "9"   "10"  "19"
```

```
newdata <- newdata[newdata$Classical == as.integer(newdata$Classical) &
                     newdata$Popular == as.integer(newdata$Popular), ]
# some values of Classical and Popular are not in the range (1,10)
newdata$Classical[which(newdata$Classical == 19)] <- 10
newdata$Classical[which(newdata$Classical == 0)] <- 9
```

```r
newdata$Popular[which(newdata$Popular == 19)] <- 10
newdata$Popular[which(newdata$Popular == 0)] <- 9

# drop these variables (missing >50%)
newdata$X1stInstr <- NULL
newdata$X2ndInstr <- NULL
# we do not consider these variables in the model
newdata$X <- NULL
newdata$first12 <- NULL

# ConsNotes, KnowAxis,NoClass, X1990s,  (5%<missing<50%)
newdata <- newdata[!is.na(newdata$ConsNotes), ]
apply(newdata, 2, pMissing)
```

```
##                    Subject                    Harmony
##                   0.000000                   0.000000
##                 Instrument                      Voice
##                   0.000000                   0.000000
##                Selfdeclare                       OMSI
##                   0.000000                   0.000000
##               X16.minus.17                  ConsInstr
##                   0.000000                   0.000000
##                  ConsNotes         Instr.minus.Notes
##                   0.000000                   0.000000
##                 PachListen                  ClsListen
##                   2.819549                   1.127820
##                    KnowRob                   KnowAxis
##                   6.390977                  10.150376
##                X1990s2000s X1990s2000s.minus.1960s1970s
##                   4.511278                   6.203008
##               CollegeMusic                    NoClass
##                   2.678571                   9.445489
##                   APTheory                  Composing
##                   9.445489                   1.550752
##                  PianoPlay                 GuitarPlay
##                   0.000000                   0.000000
##                  Classical                    Popular
##                   0.000000                   0.000000
```

```r
newdata <- newdata[!is.na(newdata$KnowAxis), ]
apply(newdata, 2, pMissing)
```

```
##                    Subject                    Harmony
##                   0.000000                   0.000000
##                 Instrument                      Voice
##                   0.000000                   0.000000
##                Selfdeclare                       OMSI
##                   0.000000                   0.000000
##               X16.minus.17                  ConsInstr
##                   0.000000                   0.000000
##                  ConsNotes         Instr.minus.Notes
##                   0.000000                   0.000000
##                 PachListen                  ClsListen
##                   3.138075                   1.255230
```

```
##                      KnowRob                     KnowAxis
##                     1.464435                     0.000000
##                  X1990s2000s X1990s2000s.minus.1960s1970s
##                     5.020921                     6.903766
##                 CollegeMusic                      NoClass
##                     2.981172                    10.512552
##                     APTheory                    Composing
##                     8.629707                     1.725941
##                    PianoPlay                   GuitarPlay
##                     0.000000                     0.000000
##                    Classical                      Popular
##                     0.000000                     0.000000
```

```r
newdata <- newdata[!is.na(newdata$NoClass), ]
apply(newdata, 2, pMissing)
```

```
##                      Subject                      Harmony
##                     0.000000                     0.000000
##                   Instrument                        Voice
##                     0.000000                     0.000000
##                  Selfdeclare                         OMSI
##                     0.000000                     0.000000
##                 X16.minus.17                    ConsInstr
##                     0.000000                     0.000000
##                    ConsNotes            Instr.minus.Notes
##                     0.000000                     0.000000
##                   PachListen                    ClsListen
##                     2.104033                     0.000000
##                      KnowRob                     KnowAxis
##                     1.636470                     0.000000
##                  X1990s2000s X1990s2000s.minus.1960s1970s
##                     4.208065                     4.208065
##                 CollegeMusic                      NoClass
##                     0.000000                     0.000000
##                     APTheory                    Composing
##                     4.208065                     0.000000
##                    PianoPlay                   GuitarPlay
##                     0.000000                     0.000000
##                    Classical                      Popular
##                     0.000000                     0.000000
```

```r
newdata <- newdata[!is.na(newdata$X1990s2000s), ]

# After removing the above four variables, the rest has <5% missing
apply(newdata, 2, pMissing)
```

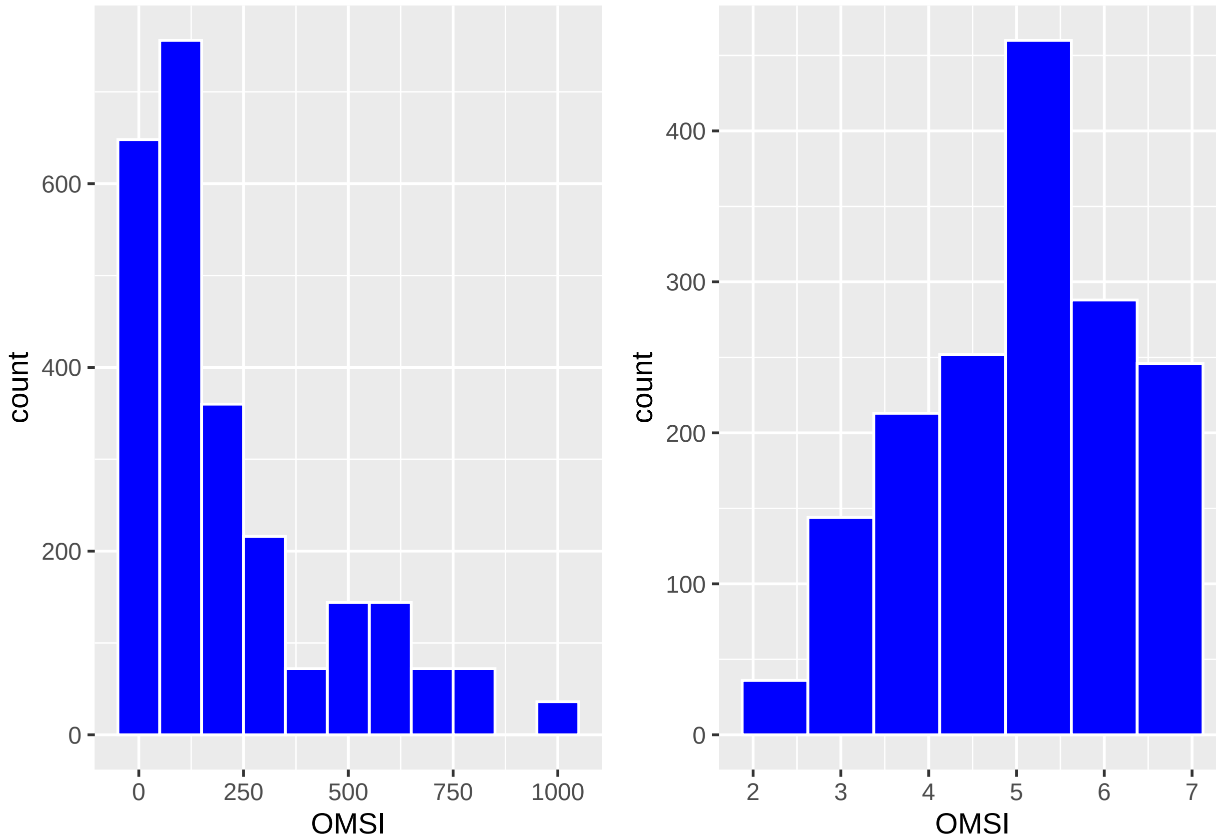
```
##                      Subject                      Harmony
##                     0.000000                     0.000000
##                   Instrument                        Voice
##                     0.000000                     0.000000
##                  Selfdeclare                         OMSI
##                     0.000000                     0.000000
##                 X16.minus.17                    ConsInstr
##                     0.000000                     0.000000
##                    ConsNotes            Instr.minus.Notes
```

```
##                     0.000000                     0.000000
##                   PachListen                    ClsListen
##                     2.196461                     0.000000
##                      KnowRob                     KnowAxis
##                     1.708359                     0.000000
##                  X1990s2000s X1990s2000s.minus.1960s1970s
##                     0.000000                     0.000000
##                 CollegeMusic                      NoClass
##                     0.000000                     0.000000
##                     APTheory                    Composing
##                     2.196461                     0.000000
##                    PianoPlay                   GuitarPlay
##                     0.000000                     0.000000
##                    Classical                      Popular
##                     0.000000                     0.000000
```

```
# Mode imputation on APTheory,KnowRob,PachListen
mode <- function(x) {
  ux <- unique(x)
  return(ux[which.max(tabulate(match(x, ux)))])
  }

newdata$APTheory[is.na(newdata$APTheory)] <- mode(newdata$APTheory[!is.na(newdata$APTheory)])
newdata$KnowRob[is.na(newdata$KnowRob)] <- mode(newdata$KnowRob[!is.na(newdata$KnowRob)])
newdata$PachListen[is.na(newdata$PachListen)] <- mode(newdata$PachListen[!is.na(newdata$PachListen)])

# no missing values!
apply(newdata, 2, pMissing)
```

```
##                      Subject                      Harmony
##                            0                            0
##                   Instrument                        Voice
##                            0                            0
##                  Selfdeclare                         OMSI
##                            0                            0
##                 X16.minus.17                    ConsInstr
##                            0                            0
##                    ConsNotes           Instr.minus.Notes
##                            0                            0
##                   PachListen                    ClsListen
##                            0                            0
##                      KnowRob                     KnowAxis
##                            0                            0
##                  X1990s2000s X1990s2000s.minus.1960s1970s
##                            0                            0
##                 CollegeMusic                      NoClass
##                            0                            0
##                     APTheory                    Composing
##                            0                            0
##                    PianoPlay                   GuitarPlay
##                            0                            0
##                    Classical                      Popular
##                            0                            0
```

```r
# log transform OMSI
newdata$OMSI <- log(newdata$OMSI)

p1 <- ggplot(data, aes(x=OMSI)) + geom_histogram(binwidth = 100, color="white", fill = "blue")
p2 <- ggplot(newdata, aes(x=OMSI)) + geom_histogram(binwidth = 0.75, color="white", fill="blue")
grid.arrange(p1, p2, nrow = 1)
```




```r
# factorize APTheory and CollegeMusic
newdata$APTheory <- as.factor(newdata$APTheory)
newdata$CollegeMusic <- as.factor(newdata$CollegeMusic)

# More EDA
flattenCorrMatrix <- function(cormat, pmat) {
  ut <- upper.tri(cormat)
  data.frame(
    row = rownames(cormat)[row(cormat)[ut]],
    column = rownames(cormat)[col(cormat)[ut]],
    cor = (cormat)[ut],
    p = pmat[ut]
  )
}

# correlation of mode imputed data
numeric_sub_data <- newdata[,5:24]
cor_matrix_imp <- rcorr(as.matrix(numeric_sub_data))
cor_imp <- flattenCorrMatrix(cor_matrix_imp$r, cor_matrix_imp$P)
correlated_imp <- cor_imp %>%
```

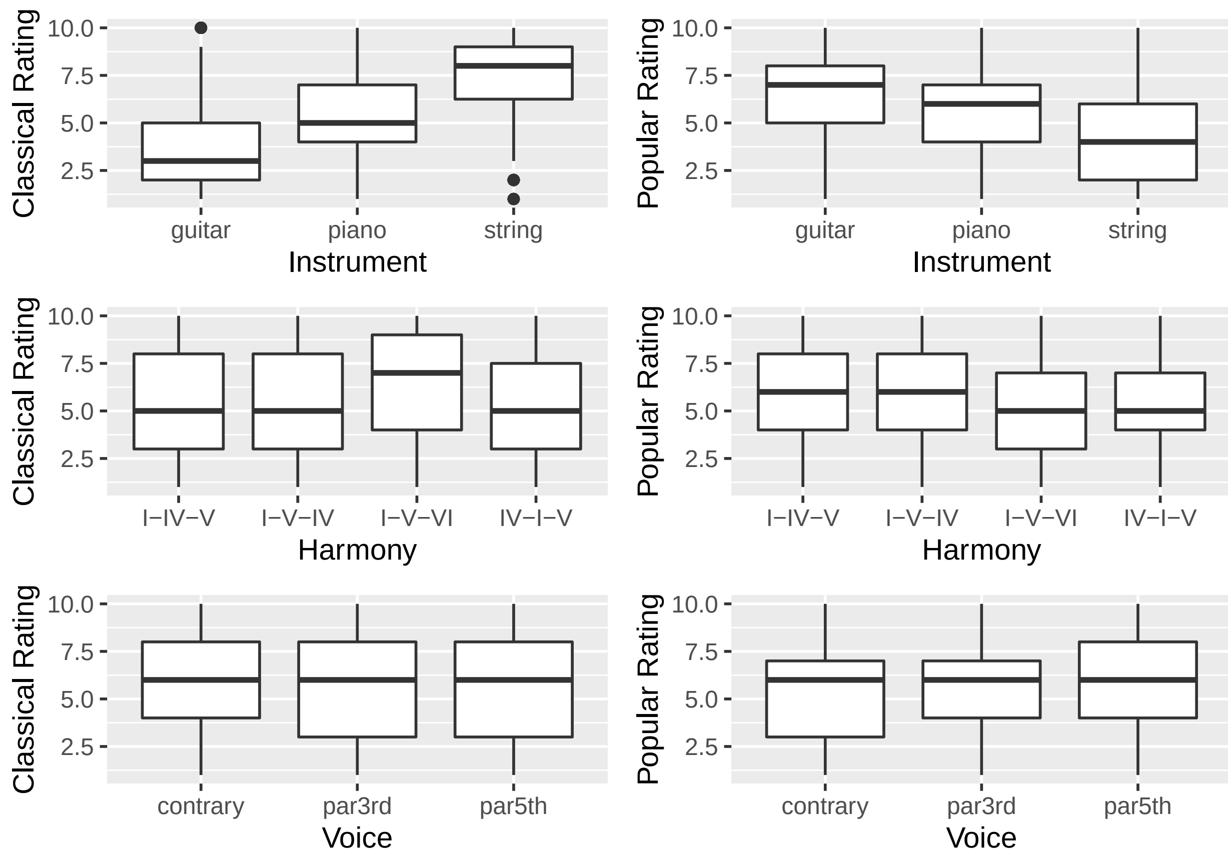
```r
  filter(p<0.05) %>%
  filter(cor < -0.6 | cor > 0.6)
correlated_imp
```

```
##           row                      column        cor p
## 1 Selfdeclare                        OMSI  0.6473301 0
## 2   ConsNotes            Instr.minus.Notes -0.7084273 0
## 3 X1990s2000s X1990s2000s.minus.1960s1970s  0.6355554 0
## 4   Composing                   GuitarPlay  0.6616028 0
## 5   Classical                      Popular -0.6023141 0
```

```r
not_correlated_imp <- cor_imp %>%
  filter(p<0.05) %>%
  filter(cor > -0.05 & cor < 0.05)
not_correlated_imp
```

```
##           row    column         cor          p
## 1 X1990s2000s   NoClass -0.04904349 0.04712400
## 2    KnowAxis  APTheory -0.04997806 0.04306690
## 3 X1990s2000s  APTheory  0.04917634 0.04652839
## 4  GuitarPlay Classical  0.04928329 0.04605350
## 5    KnowAxis   Popular  0.04876866 0.04837647
```

```r
# Boxplot for instrument
box1 <- ggplot(newdata,aes(x=factor(Instrument),y=Classical)) + geom_boxplot() + labs(x="Instrument",y=

box2 <- ggplot(newdata,aes(x=factor(Instrument),y=Popular)) + geom_boxplot() + labs(x="Instrument",y="P

# grid.arrange(box1,box2,ncol=2)
# Boxplot for Harmony
box3 <- ggplot(newdata,aes(x=factor(Harmony),y=Classical)) + geom_boxplot() + labs(x="Harmony",y="Class

box4 <- ggplot(newdata,aes(x=factor(Harmony),y=Popular)) + geom_boxplot() + labs(x="Harmony",y="Popular

# grid.arrange(box3,box4,ncol=2)
# Boxplot for Voice
box5 <- ggplot(newdata,aes(x=factor(Voice),y=Classical)) + geom_boxplot() + labs(x="Voice",y="Classical

box6 <- ggplot(newdata,aes(x=factor(Voice),y=Popular)) + geom_boxplot() +
labs(x="Voice",y="Popular Rating")
grid.arrange(box1,box2, box3,box4, box5,box6,ncol=2)
```




```
# Building Model (1)
# stepwise
fullmodel <- lm(Classical~Harmony + Instrument + Voice +
                  Selfdeclare + OMSI+X16.minus.17+ConsInstr+
                  ConsNotes+Instr.minus.Notes +PachListen+ClsListen+KnowRob+
                  KnowAxis+X1990s2000s+X1990s2000s.minus.1960s1970s+
                  CollegeMusic+NoClass+APTheory+Composing+PianoPlay+
                  GuitarPlay, data = newdata)

summary(fullmodel)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony + Instrument + Voice + Selfdeclare +
##     OMSI + X16.minus.17 + ConsInstr + ConsNotes + Instr.minus.Notes +
##     PachListen + ClsListen + KnowRob + KnowAxis + X1990s2000s +
##     X1990s2000s.minus.1960s1970s + CollegeMusic + NoClass + APTheory +
##     Composing + PianoPlay + GuitarPlay, data = newdata)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.4386 -1.4390 -0.0143  1.4809  6.2638
##
## Coefficients: (1 not defined because of singularities)
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   2.38047    0.52404   4.543 5.97e-06 ***
## HarmonyI-V-IV                 0.01121    0.14701   0.076 0.939238
```

```
## HarmonyI-V-VI                 0.86389    0.14729   5.865 5.42e-09 ***
## HarmonyIV-I-V                 0.10209    0.14692   0.695 0.487228
## Instrumentpiano               1.61526    0.12773  12.646  < 2e-16 ***
## Instrumentstring              3.52873    0.12695  27.796  < 2e-16 ***
## Voicepar3rd                  -0.36605    0.12747  -2.872 0.004137 **
## Voicepar5th                  -0.34592    0.12747  -2.714 0.006724 **
## Selfdeclare                  -0.46722    0.09498  -4.919 9.57e-07 ***
## OMSI                          0.05121    0.07674   0.667 0.504675
## X16.minus.17                 -0.04640    0.02540  -1.827 0.067895 .
## ConsInstr                     0.10971    0.05002   2.193 0.028426 *
## ConsNotes                    -0.19800    0.04366  -4.535 6.17e-06 ***
## Instr.minus.Notes                  NA         NA      NA       NA
## PachListen                    0.23628    0.08239   2.868 0.004187 **
## ClsListen                     0.24364    0.05164   4.718 2.59e-06 ***
## KnowRob                       0.10942    0.03980   2.749 0.006045 **
## KnowAxis                     -0.02356    0.03215  -0.733 0.463696
## X1990s2000s                   0.11787    0.05715   2.062 0.039322 *
## X1990s2000s.minus.1960s1970s  0.09627    0.05104   1.886 0.059446 .
## CollegeMusic1                -0.10908    0.16136  -0.676 0.499147
## NoClass                      -0.10949    0.05857  -1.869 0.061745 .
## APTheory1                     0.61591    0.16742   3.679 0.000242 ***
## Composing                     0.27357    0.05378   5.086 4.08e-07 ***
## PianoPlay                     0.30045    0.03856   7.792 1.17e-14 ***
## GuitarPlay                   -0.09484    0.06141  -1.544 0.122697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.106 on 1614 degrees of freedom
## Multiple R-squared:  0.4079, Adjusted R-squared:  0.3991
## F-statistic: 46.33 on 24 and 1614 DF,  p-value: < 2.2e-16
```

```r
step_full <- stepAIC(fullmodel, direction = "backward", k=log(1639),trace=0)
summary(step_full)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony + Instrument + Selfdeclare +
##     ConsNotes + PachListen + ClsListen + KnowRob + X1990s2000s +
##     APTheory + Composing + PianoPlay + GuitarPlay, data = newdata)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9245 -1.4305 -0.0058  1.4589  6.3338
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.98038    0.43084   4.597 4.63e-06 ***
## HarmonyI-V-IV      0.01133    0.14788   0.077  0.93896
## HarmonyI-V-VI      0.86548    0.14815   5.842 6.22e-09 ***
## HarmonyIV-I-V      0.10219    0.14778   0.691  0.48936
## Instrumentpiano    1.61633    0.12845  12.584  < 2e-16 ***
## Instrumentstring   3.52916    0.12770  27.637  < 2e-16 ***
## Selfdeclare       -0.43354    0.06677  -6.493 1.11e-10 ***
## ConsNotes         -0.12900    0.02901  -4.447 9.30e-06 ***
## PachListen         0.21036    0.07107   2.960  0.00312 **
```

```
## ClsListen         0.24295    0.03850   6.310 3.58e-10 ***
## KnowRob           0.13641    0.03393   4.020 6.09e-05 ***
## X1990s2000s       0.25981    0.03592   7.233 7.24e-13 ***
## APTheory1         0.67771    0.14659   4.623 4.08e-06 ***
## Composing         0.29096    0.05040   5.773 9.29e-09 ***
## PianoPlay         0.28515    0.03733   7.638 3.74e-14 ***
## GuitarPlay       -0.22772    0.05049  -4.510 6.94e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.119 on 1623 degrees of freedom
## Multiple R-squared:  0.3976, Adjusted R-squared:  0.392
## F-statistic:  71.4 on 15 and 1623 DF,  p-value: < 2.2e-16
```

```r
AIC(step_full)
```

```
## [1] 7130.068
```

```r
# checking interaction
interact_model <- lm(Classical~Harmony * Instrument * Voice +
                  Selfdeclare + OMSI+X16.minus.17+ConsInstr+
                  ConsNotes+Instr.minus.Notes +PachListen+ClsListen+KnowRob+
                  KnowAxis+X1990s2000s+X1990s2000s.minus.1960s1970s+
                  CollegeMusic+NoClass+APTheory+Composing+PianoPlay+
                  GuitarPlay, data = newdata)
step_interact <- stepAIC(interact_model, direction = "backward", k=log(1639),trace=0)
summary(step_interact)
```

```
##
## Call:
## lm(formula = Classical ~ Harmony + Instrument + Selfdeclare +
##     ConsNotes + PachListen + ClsListen + KnowRob + X1990s2000s +
##     APTheory + Composing + PianoPlay + GuitarPlay, data = newdata)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9245 -1.4305 -0.0058  1.4589  6.3338
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.98038    0.43084   4.597 4.63e-06 ***
## HarmonyI-V-IV      0.01133    0.14788   0.077  0.93896
## HarmonyI-V-VI      0.86548    0.14815   5.842 6.22e-09 ***
## HarmonyIV-I-V      0.10219    0.14778   0.691  0.48936
## Instrumentpiano    1.61633    0.12845  12.584  < 2e-16 ***
## Instrumentstring   3.52916    0.12770  27.637  < 2e-16 ***
## Selfdeclare       -0.43354    0.06677  -6.493 1.11e-10 ***
## ConsNotes         -0.12900    0.02901  -4.447 9.30e-06 ***
## PachListen         0.21036    0.07107   2.960  0.00312 **
## ClsListen          0.24295    0.03850   6.310 3.58e-10 ***
## KnowRob            0.13641    0.03393   4.020 6.09e-05 ***
## X1990s2000s        0.25981    0.03592   7.233 7.24e-13 ***
## APTheory1          0.67771    0.14659   4.623 4.08e-06 ***
## Composing          0.29096    0.05040   5.773 9.29e-09 ***
## PianoPlay          0.28515    0.03733   7.638 3.74e-14 ***
```

```
## GuitarPlay       -0.22772    0.05049  -4.510 6.94e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.119 on 1623 degrees of freedom
## Multiple R-squared:  0.3976, Adjusted R-squared:  0.392
## F-statistic:  71.4 on 15 and 1623 DF,  p-value: < 2.2e-16
```

```r
AIC(step_interact)
```

```
## [1] 7130.068
```

```r
# With the fixed effects, add a random intercept
lmer3 <- lme4::lmer(Classical~Harmony + Instrument + Voice +
                  Selfdeclare + ConsNotes+PachListen+ClsListen+
                  KnowRob + X1990s2000s+APTheory+Composing+PianoPlay+
                GuitarPlay+ (1|Subject), data = newdata,
              control = lmerControl(optimizer = "bobyqa"), REML = F)
summary(lmer3)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ConsNotes +
##     PachListen + ClsListen + KnowRob + X1990s2000s + APTheory +
##     Composing + PianoPlay + GuitarPlay + (1 | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6928.4   7036.4  -3444.2   6888.4     1619
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1734 -0.6251 -0.0107  0.6267  3.5347
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.7192   0.848
##  Residual             3.6944   1.922
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                    Estimate Std. Error t value
## (Intercept)        2.193749   1.068562   2.053
## HarmonyI-V-IV      0.009917   0.134165   0.074
## HarmonyI-V-VI      0.863938   0.134422   6.427
## HarmonyIV-I-V      0.102097   0.134081   0.761
## Instrumentpiano    1.613866   0.116661  13.834
## Instrumentstring   3.526691   0.115860  30.439
## Voicepar3rd       -0.367315   0.116333  -3.157
## Voicepar5th       -0.348192   0.116333  -2.993
## Selfdeclare       -0.432207   0.171093  -2.526
## ConsNotes         -0.127022   0.074298  -1.710
## PachListen         0.213195   0.181839   1.172
## ClsListen          0.245458   0.098531   2.491
## KnowRob            0.132449   0.086731   1.527
```

```
## X1990s2000s       0.261168   0.091994   2.839
## APTheory1         0.682681   0.372024   1.835
## Composing         0.293844   0.128583   2.285
## PianoPlay         0.283998   0.095246   2.982
## GuitarPlay       -0.234766   0.128513  -1.827
```

```r
#ConsNotes, Pachlisten, KnowRob, APTheory, GuitarPlay not significantly far from 0

anova(lmer3, step_full)
```

```
## Data: newdata
## Models:
## step_full: Classical ~ Harmony + Instrument + Selfdeclare + ConsNotes +
## step_full:     PachListen + ClsListen + KnowRob + X1990s2000s + APTheory +
## step_full:     Composing + PianoPlay + GuitarPlay
## lmer3: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ConsNotes +
## lmer3:     PachListen + ClsListen + KnowRob + X1990s2000s + APTheory +
## lmer3:     Composing + PianoPlay + GuitarPlay + (1 | Subject)
##           Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## step_full 17 7130.1 7221.9 -3548.0   7096.1
## lmer3     20 6928.4 7036.4 -3444.2   6888.4 207.67      3  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
exactRLRT(lmer3)
```

```
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
## data:
## RLRT = 216.78, p-value < 2.2e-16
```

```r
# p-value is significantly small, so it makes sense to include a random intercept

# take out not significant variables
lmer4.0 <- lme4::lmer(Classical~Harmony + Instrument + Voice +
                  Selfdeclare +ClsListen+
                   X1990s2000s+Composing+PianoPlay+
                (1|Subject), data = newdata,
              control = lmerControl(optimizer = "bobyqa"), REML = F)
summary(lmer4.0)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ClsListen +
##     X1990s2000s + Composing + PianoPlay + (1 | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6928.1   7009.1  -3449.0   6898.1     1624
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1877 -0.6145 -0.0137  0.6221  3.5663
```

```
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.9111   0.9545
##  Residual             3.6944   1.9221
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)       3.348499   0.604801   5.537
## HarmonyI-V-IV     0.009766   0.134165   0.073
## HarmonyI-V-VI     0.863901   0.134423   6.427
## HarmonyIV-I-V     0.102096   0.134082   0.761
## Instrumentpiano   1.612554   0.116662  13.822
## Instrumentstring  3.526720   0.115861  30.439
## Voicepar3rd      -0.367372   0.116333  -3.158
## Voicepar5th      -0.348143   0.116333  -2.993
## Selfdeclare      -0.499794   0.175441  -2.849
## ClsListen         0.250529   0.106435   2.354
## X1990s2000s       0.210995   0.097330   2.168
## Composing         0.235913   0.123021   1.918
## PianoPlay         0.298140   0.102966   2.896
```

```r
exactRLRT(lmer4.0)
```

```
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
## data:
## RLRT = 268.17, p-value < 2.2e-16
```

```r
# examine random effect
lmer4 <-ffRanefLMER.fnc(lmer4.0, ran.effects = c("(Harmony|Subject)",
                                       "(Instrument|Subject)", "(Voice|Subject)"),
                                        log.file=F)
```

```
## evaluating addition of (Harmony|Subject) to model
##  log-likelihood ratio test p-value = 5.028463e-08
##  adding (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 9.883115e-63
##  adding (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
##  log-likelihood ratio test p-value = 0.3519592
##  not adding (Voice|Subject) to model
```

```r
summary(lmer4)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ClsListen +
##     X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
```

```
##
##      AIC      BIC   logLik deviance df.resid
##   6601.9   6769.3  -3269.9   6539.9     1608
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6057 -0.5662  0.0087  0.5384  3.6428
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.00000  0.0000
##  Subject.1 (Intercept)      0.93569  0.9673
##            HarmonyI-V-IV    0.05701  0.2388    0.39
##            HarmonyI-V-VI    1.61792  1.2720   -0.34  0.71
##            HarmonyIV-I-V    0.03811  0.1952   -0.06  0.11  0.38
##  Subject.2 (Intercept)      0.69107  0.8313
##            Instrumentpiano  2.16630  1.4718   -0.79
##            Instrumentstring 3.48013  1.8655   -0.98  0.64
##  Residual                   2.51898  1.5871
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)        3.53795    0.59381   5.958
## HarmonyI-V-IV      0.00966    0.11625   0.083
## HarmonyI-V-VI      0.87017    0.21799   3.992
## HarmonyIV-I-V      0.10284    0.11440   0.899
## Instrumentpiano    1.60871    0.23765   6.769
## Instrumentstring   3.53280    0.29123  12.131
## Voicepar3rd       -0.36094    0.09608  -3.757
## Voicepar5th       -0.34480    0.09609  -3.588
## Selfdeclare       -0.44427    0.16922  -2.625
## ClsListen          0.21316    0.10267   2.076
## X1990s2000s        0.14740    0.09388   1.570
## Composing          0.21723    0.11868   1.830
## PianoPlay          0.30788    0.09931   3.100
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# Check the influence of KnowRob and KnowAxis, not significant
lmer4.test = update(lmer4, .~. +KnowRob:Harmony+KnowAxis:Harmony)
summary(lmer4.test)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ClsListen +
##     X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony |
##     Subject) + (Instrument | Subject) + Harmony:KnowRob + Harmony:KnowAxis
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6598.6   6809.3  -3260.3   6520.6     1600
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -4.6090 -0.5681  0.0168  0.5309  3.6258
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      2.296e-13 4.792e-07
##  Subject.1 (Intercept)      8.442e-01 9.188e-01
##            HarmonyI-V-IV    7.406e-02 2.721e-01  0.50
##            HarmonyI-V-VI    1.076e+00 1.037e+00 -0.29  0.67
##            HarmonyIV-I-V    2.747e-02 1.657e-01  0.19  0.13  0.18
##  Subject.2 (Intercept)      5.998e-01 7.745e-01
##            Instrumentpiano  2.167e+00 1.472e+00 -0.81
##            Instrumentstring 3.484e+00 1.867e+00 -0.97  0.64
##  Residual                   2.514e+00 1.585e+00
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                          Estimate Std. Error t value
## (Intercept)              3.528152   0.592061   5.959
## HarmonyI-V-IV            0.059663   0.139917   0.426
## HarmonyI-V-VI            0.726057   0.224715   3.231
## HarmonyIV-I-V            0.161137   0.134685   1.196
## Instrumentpiano          1.608721   0.237638   6.770
## Instrumentstring         3.533336   0.291334  12.128
## Voicepar3rd             -0.360823   0.095977  -3.759
## Voicepar5th             -0.345704   0.095983  -3.602
## Selfdeclare             -0.454407   0.170333  -2.668
## ClsListen                0.192197   0.102075   1.883
## X1990s2000s              0.138455   0.091728   1.509
## Composing                0.206141   0.116564   1.768
## PianoPlay                0.323932   0.097297   3.329
## HarmonyI-IV-V:KnowRob   -0.053327   0.101167  -0.527
## HarmonyI-V-IV:KnowRob   -0.035204   0.113543  -0.310
## HarmonyI-V-VI:KnowRob    0.394103   0.120715   3.265
## HarmonyIV-I-V:KnowRob   -0.004996   0.104514  -0.048
## HarmonyI-IV-V:KnowAxis   0.132181   0.080954   1.633
## HarmonyI-V-IV:KnowAxis   0.078305   0.091592   0.855
## HarmonyI-V-VI:KnowAxis  -0.067506   0.097796  -0.690
## HarmonyIV-I-V:KnowAxis   0.050363   0.083873   0.600
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
anova(lmer4.test, lmer4)
```

```
## Data: newdata
## Models:
## lmer4: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ClsListen +
## lmer4:     X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony |
## lmer4:     Subject) + (Instrument | Subject)
## lmer4.test: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ClsListen +
## lmer4.test:     X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony |
## lmer4.test:     Subject) + (Instrument | Subject) + Harmony:KnowRob + Harmony:KnowAxis
##            Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer4      31 6601.9 6769.3 -3269.9   6539.9
## lmer4.test 39 6598.6 6809.3 -3260.3   6520.6 19.206      8     0.0138 *
## ---
```

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

```r
# Checking all four levels of harmony
lmer4.1 <- lme4::lmer(Classical ~ Harmony-1 + Instrument + Voice + Selfdeclare + ClsListen +
    X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony-1 |
    Subject) + (Instrument | Subject), data = newdata,
    control = lmerControl(optimizer = "bobyqa"), REML = F)
summary(lmer4.1)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Harmony - 1 + Instrument + Voice + Selfdeclare +
##     ClsListen + X1990s2000s + Composing + PianoPlay + (1 | Subject) +
##     (Harmony - 1 | Subject) + (Instrument | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6601.9   6769.4  -3270.0   6539.9     1608
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6182 -0.5662  0.0109  0.5404  3.6603
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
##  Subject.1 HarmonyI-IV-V    0.9373   0.9681
##            HarmonyI-V-IV    1.1683   1.0809   0.98
##            HarmonyI-V-VI    1.7150   1.3096   0.41  0.58
##            HarmonyIV-I-V    0.9472   0.9732   0.98  0.97  0.47
##  Subject.2 (Intercept)      0.6870   0.8288
##            Instrumentpiano  2.1656   1.4716   -0.79
##            Instrumentstring 3.4799   1.8654   -0.98  0.64
##  Residual                   2.5204   1.5876
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                   Estimate Std. Error t value
## HarmonyI-IV-V      3.53564    0.59342   5.958
## HarmonyI-V-IV      3.54548    0.59785   5.930
## HarmonyI-V-VI      4.40582    0.60756   7.252
## HarmonyIV-I-V      3.63850    0.59363   6.129
## Instrumentpiano    1.60885    0.23763   6.771
## Instrumentstring   3.53280    0.29123  12.131
## Voicepar3rd       -0.36108    0.09611  -3.757
## Voicepar5th       -0.34481    0.09611  -3.587
## Selfdeclare       -0.44266    0.16911  -2.618
## ClsListen          0.21375    0.10260   2.083
## X1990s2000s        0.14682    0.09382   1.565
## Composing          0.21657    0.11859   1.826
## PianoPlay          0.30766    0.09924   3.100
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# Checking all four levels of instrument
lmer4.2 <- lme4::lmer(Classical ~ Instrument-1 + Voice + Harmony + Selfdeclare + ClsListen +
    X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony |
    Subject) + (Instrument-1 | Subject), data = newdata,
    control = lmerControl(optimizer = "bobyqa"), REML = F)
summary(lmer4.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Instrument - 1 + Voice + Harmony + Selfdeclare +
##     ClsListen + X1990s2000s + Composing + PianoPlay + (1 | Subject) +
##     (Harmony | Subject) + (Instrument - 1 | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6601.9   6769.3  -3269.9   6539.9     1608
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6057 -0.5662  0.0087  0.5384  3.6428
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      1.542e-16 1.242e-08
##  Subject.1 (Intercept)      9.357e-01 9.673e-01
##            HarmonyI-V-IV    5.701e-02 2.388e-01  0.39
##            HarmonyI-V-VI    1.618e+00 1.272e+00 -0.34  0.71
##            HarmonyIV-I-V    3.811e-02 1.952e-01 -0.06  0.11  0.38
##  Subject.2 Instrumentguitar 6.911e-01 8.313e-01
##            Instrumentpiano  9.336e-01 9.662e-01 -0.34
##            Instrumentstring 1.137e+00 1.066e+00 -0.93 -0.03
##  Residual                   2.519e+00 1.587e+00
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                  Estimate Std. Error t value
## Instrumentguitar  3.53795    0.59381   5.958
## Instrumentpiano   5.14666    0.59814   8.604
## Instrumentstring  7.07074    0.60194  11.747
## Voicepar3rd      -0.36094    0.09608  -3.757
## Voicepar5th      -0.34480    0.09609  -3.588
## HarmonyI-V-IV     0.00966    0.11625   0.083
## HarmonyI-V-VI     0.87017    0.21799   3.992
## HarmonyIV-I-V     0.10284    0.11440   0.899
## Selfdeclare      -0.44427    0.16922  -2.625
## ClsListen         0.21316    0.10267   2.076
## X1990s2000s       0.14740    0.09388   1.570
## Composing         0.21723    0.11868   1.830
## PianoPlay         0.30788    0.09931   3.100
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# Checking all four levels of voice
lmer4.3 <- lme4::lmer(Classical ~ Voice-1 + Instrument + Harmony + Selfdeclare + ClsListen +
```

```r
    X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony |
    Subject) + (Instrument | Subject), data = newdata,
    control = lmerControl(optimizer = "bobyqa"), REML = F)
summary(lmer4.3)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Voice - 1 + Instrument + Harmony + Selfdeclare +
##     ClsListen + X1990s2000s + Composing + PianoPlay + (1 | Subject) +
##     (Harmony | Subject) + (Instrument | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6601.9   6769.3  -3269.9   6539.9     1608
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6057 -0.5662  0.0087  0.5384  3.6428
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      5.371e-14 2.318e-07
##  Subject.1 (Intercept)      9.357e-01 9.673e-01
##            HarmonyI-V-IV    5.701e-02 2.388e-01  0.39
##            HarmonyI-V-VI    1.618e+00 1.272e+00 -0.34  0.71
##            HarmonyIV-I-V    3.811e-02 1.952e-01 -0.06  0.11  0.38
##  Subject.2 (Intercept)      6.911e-01 8.313e-01
##            Instrumentpiano  2.166e+00 1.472e+00 -0.79
##            Instrumentstring 3.480e+00 1.866e+00 -0.98  0.64
##  Residual                   2.519e+00 1.587e+00
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                  Estimate Std. Error t value
## Voicecontrary     3.53795    0.59381   5.958
## Voicepar3rd       3.17700    0.59377   5.351
## Voicepar5th       3.19315    0.59383   5.377
## Instrumentpiano   1.60871    0.23765   6.769
## Instrumentstring  3.53280    0.29123  12.131
## HarmonyI-V-IV     0.00966    0.11625   0.083
## HarmonyI-V-VI     0.87017    0.21799   3.992
## HarmonyIV-I-V     0.10284    0.11440   0.899
## Selfdeclare      -0.44427    0.16922  -2.625
## ClsListen         0.21316    0.10267   2.076
## X1990s2000s       0.14740    0.09388   1.570
## Composing         0.21723    0.11868   1.830
## PianoPlay         0.30788    0.09931   3.100
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
anova(lmer3, lmer4.0, lmer4)
```

```
## Data: newdata
## Models:
```

```
## lmer4.0: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ClsListen +
## lmer4.0:     X1990s2000s + Composing + PianoPlay + (1 | Subject)
## lmer3: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ConsNotes +
## lmer3:     PachListen + ClsListen + KnowRob + X1990s2000s + APTheory +
## lmer3:     Composing + PianoPlay + GuitarPlay + (1 | Subject)
## lmer4: Classical ~ Harmony + Instrument + Voice + Selfdeclare + ClsListen +
## lmer4:     X1990s2000s + Composing + PianoPlay + (1 | Subject) + (Harmony |
## lmer4:     Subject) + (Instrument | Subject)
##         Df    AIC    BIC  logLik deviance    Chisq Chi Df Pr(>Chisq)
## lmer4.0 15 6928.1 7009.1 -3449.0   6898.1
## lmer3   20 6928.4 7036.4 -3444.2   6888.4   9.6537      5    0.08566 .
## lmer4   31 6601.9 6769.3 -3269.9   6539.9 348.5464     11    < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# added Harmony|Subject and Instrument|Subject to lmer3

# check multicollinearity
vif(lmer4)
```

```
##                 GVIF Df GVIF^(1/(2*Df))
## Harmony     1.000028  3        1.000005
## Instrument  1.000009  2        1.000002
## Voice       1.000024  2        1.000006
## Selfdeclare 1.873539  1        1.368773
## ClsListen   1.264806  1        1.124636
## X1990s2000s 1.055315  1        1.027285
## Composing   1.561338  1        1.249535
## PianoPlay   1.388372  1        1.178292
```

```r
# plot residuals
# mlm facet plot
dev.new()

res.4 <- r.cond(lmer4)      ## standardized conditional residuals
robust.sd.4 <- diff(quantile(res.4,c(.025,.975)))/(2*1.96)
res.4 <- res.4/robust.sd.4
fit.4 <- yhat.cond(lmer4)

newdata.4 <- data.frame(subject=newdata$Subject,res_lmer4=res.4,fit_lmer4=fit.4)

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

mlm_facets(newdata.4,"subject",x="fit_lmer4",y="res_lmer4",params=resparams,
           lty=c(1,2,2),size=c(1,1,1))

res_lmer4=res.4
fit_lmer4=fit.4
dev.new()
par(mfrow=c(2,2))
# residuals look randomly scattered around zero. Assumptions met.
plot(fit_lmer4,res_lmer4)
```

```r
abline(h=0)
abline(h=2,lty=2)
abline(h=-2,lty=2)
qqnorm(res_lmer4)
qqline(res_lmer4)
library(arm)
# binnedplot shows that most residuals are within range
par(mfrow=c(2,2))
binnedplot(fit_lmer4, res_lmer4)
plot(fit_lmer4,res_lmer4)
```

```r
# Building Model (2)

# Examine fixed effect by stepwise
fullpop <- lm(Popular~Harmony + Instrument + Voice + Selfdeclare + OMSI + X16.minus.17+ConsInstr+ConsNo
Instr.minus.Notes+X1990s2000s+X1990s2000s.minus.1960s1970s+CollegeMusic+
NoClass+APTheory+Composing+PianoPlay+GuitarPlay, data = newdata)

steppop <- stepAIC(fullpop, direction = "backward", k=log(1639),trace=0)
summary(steppop)
```

```
##
## Call:
## lm(formula = Popular ~ Instrument + Selfdeclare + OMSI + X16.minus.17 +
##     ConsInstr + PachListen + ClsListen + KnowRob + CollegeMusic +
##     APTheory, data = newdata)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.727 -1.675  0.077  1.614  6.065
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       7.67702    0.43636  17.593  < 2e-16 ***
## Instrumentpiano  -1.13821    0.13215  -8.613  < 2e-16 ***
## Instrumentstring -2.66492    0.13139 -20.283  < 2e-16 ***
## Selfdeclare      -0.29806    0.07599  -3.922 9.14e-05 ***
## OMSI              0.26709    0.07060   3.783 0.000160 ***
## X16.minus.17      0.16732    0.02071   8.080 1.25e-15 ***
## ConsInstr         0.14838    0.04206   3.528 0.000430 ***
## PachListen       -0.52180    0.08218  -6.350 2.79e-10 ***
## ClsListen        -0.12999    0.03857  -3.370 0.000770 ***
## KnowRob           0.14999    0.03366   4.456 8.92e-06 ***
## CollegeMusic1     0.46676    0.13614   3.428 0.000622 ***
## APTheory1         0.43893    0.15082   2.910 0.003659 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.18 on 1627 degrees of freedom
## Multiple R-squared:  0.2552, Adjusted R-squared:  0.2501
## F-statistic: 50.67 on 11 and 1627 DF,  p-value: < 2.2e-16
```

```r
AIC(steppop)
```

```
## [1] 7219.545
```

```r
# Examine interaction
inter_pop <- lm(Popular~Harmony * Instrument * Voice + Selfdeclare + OMSI + X16.minus.17+ConsInstr+ConsI
Instr.minus.Notes+X1990s2000s+X1990s2000s.minus.1960s1970s+CollegeMusic+
NoClass+APTheory+Composing+PianoPlay+GuitarPlay, data = newdata)

step_inter_pop <- stepAIC(inter_pop, direction = "backward", k=log(1639), trace=0)
summary(step_inter_pop)
```

```
##
## Call:
## lm(formula = Popular ~ Instrument + Selfdeclare + OMSI + X16.minus.17 +
##     ConsInstr + PachListen + ClsListen + KnowRob + CollegeMusic +
##     APTheory, data = newdata)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.727 -1.675  0.077  1.614  6.065
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       7.67702    0.43636  17.593  < 2e-16 ***
## Instrumentpiano  -1.13821    0.13215  -8.613  < 2e-16 ***
## Instrumentstring -2.66492    0.13139 -20.283  < 2e-16 ***
## Selfdeclare      -0.29806    0.07599  -3.922 9.14e-05 ***
## OMSI              0.26709    0.07060   3.783 0.000160 ***
## X16.minus.17      0.16732    0.02071   8.080 1.25e-15 ***
## ConsInstr         0.14838    0.04206   3.528 0.000430 ***
## PachListen       -0.52180    0.08218  -6.350 2.79e-10 ***
## ClsListen        -0.12999    0.03857  -3.370 0.000770 ***
## KnowRob           0.14999    0.03366   4.456 8.92e-06 ***
## CollegeMusic1     0.46676    0.13614   3.428 0.000622 ***
## APTheory1         0.43893    0.15082   2.910 0.003659 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.18 on 1627 degrees of freedom
## Multiple R-squared:  0.2552, Adjusted R-squared:  0.2501
## F-statistic: 50.67 on 11 and 1627 DF,  p-value: < 2.2e-16
```

```r
AIC(step_inter_pop)
```

```
## [1] 7219.545
```

```r
# Don't need interaction since steppop and step_inter_pop are the same model


exp <- lme4::lmer(Popular ~ Harmony + Instrument + Voice + Selfdeclare + OMSI +
              X16.minus.17 + ConsInstr + PachListen + ClsListen + KnowRob +
              CollegeMusic + APTheory + (1|Subject),
            control = lmerControl(optimizer = "bobyqa"),
            data = newdata, REML = F)
# Selfdeclare, OMSI, ConsInstr, ClsListen, KnowRob, CollegeMusic, APTheory
# not significant
summary(exp)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Popular ~ Harmony + Instrument + Voice + Selfdeclare + OMSI +
##     X16.minus.17 + ConsInstr + PachListen + ClsListen + KnowRob +
##     CollegeMusic + APTheory + (1 | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6961.5   7064.2  -3461.8   6923.5     1620
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5690 -0.6433  0.0633  0.6480  3.0144
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.9294   0.964
##  Residual             3.7518   1.937
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)        7.67843    1.20525   6.371
## HarmonyI-V-IV      0.02838    0.13520   0.210
## HarmonyI-V-VI     -0.30234    0.13546  -2.232
## HarmonyIV-I-V     -0.29730    0.13512  -2.200
## Instrumentpiano   -1.14231    0.11757  -9.716
## Instrumentstring  -2.66660    0.11676 -22.839
## Voicepar3rd        0.12208    0.11723   1.041
## Voicepar5th        0.21629    0.11723   1.845
## Selfdeclare       -0.29709    0.21226  -1.400
## OMSI               0.26794    0.19738   1.357
## X16.minus.17       0.16651    0.05788   2.877
## ConsInstr          0.15191    0.11669   1.302
## PachListen        -0.51714    0.22926  -2.256
## ClsListen         -0.13122    0.10775  -1.218
## KnowRob            0.14529    0.09335   1.556
## CollegeMusic1      0.46085    0.37955   1.214
## APTheory1          0.43307    0.41676   1.039
```

```r
# exp has lower AIC
anova(exp, steppop)
```

```
## Data: newdata
## Models:
## steppop: Popular ~ Instrument + Selfdeclare + OMSI + X16.minus.17 + ConsInstr +
## steppop:     PachListen + ClsListen + KnowRob + CollegeMusic + APTheory
## exp: Popular ~ Harmony + Instrument + Voice + Selfdeclare + OMSI +
## exp:     X16.minus.17 + ConsInstr + PachListen + ClsListen + KnowRob +
## exp:     CollegeMusic + APTheory + (1 | Subject)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## steppop 13 7219.5 7289.8 -3596.8   7193.5
## exp     19 6961.5 7064.2 -3461.8   6923.5 270.02      6  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# add a random intercept
exactRLRT(exp)
```

```
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
## data:
## RLRT = 277.99, p-value < 2.2e-16
```

```r
# Take out the insignificant variables
exp.1 <- lme4::lmer(Popular ~ Harmony + Instrument + Voice +
             X16.minus.17 + PachListen + (1|Subject),
            control = lmerControl(optimizer = "bobyqa"),
            data = newdata, REML = F)

# examine other random effects
lmer_pop <- ffRanefLMER.fnc(exp.1,
                         ran.effects = c("(Harmony|Subject)",
                                      "(Instrument|Subject)",
                                      "(Voice|Subject)"),
                                       log.file=F)
```

```
## evaluating addition of (Harmony|Subject) to model
##  log-likelihood ratio test p-value = 0.0003374615
##  adding (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 7.167667e-61
##  adding (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
##  log-likelihood ratio test p-value = 0.8266013
##  not adding (Voice|Subject) to model
```

```r
summary(lmer_pop)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Popular ~ Harmony + Instrument + Voice + X16.minus.17 + PachListen +
##     (1 | Subject) + (Harmony | Subject) + (Instrument | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6660.0   6811.2  -3302.0   6604.0     1611
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8248 -0.5785  0.0525  0.6017  3.1605
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.00000  0.0000
##  Subject.1 (Intercept)      0.79058  0.8891
##            HarmonyI-V-IV    0.08761  0.2960    0.26
##            HarmonyI-V-VI    0.69339  0.8327    0.04  0.04
```

```
##            HarmonyIV-I-V    0.29276  0.5411   -0.59 -0.87 -0.36
##  Subject.2 (Intercept)      0.54009  0.7349
##            Instrumentpiano  1.80548  1.3437   -0.40
##            Instrumentstring 3.78490  1.9455   -0.93  0.70
##  Residual                   2.64922  1.6276
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                  Estimate Std. Error t value
## (Intercept)       8.42255    0.81993  10.272
## HarmonyI-V-IV     0.03083    0.12172   0.253
## HarmonyI-V-VI    -0.31019    0.16749  -1.852
## HarmonyIV-I-V    -0.29648    0.13878  -2.136
## Instrumentpiano  -1.14857    0.22159  -5.183
## Instrumentstring -2.67402    0.30317  -8.820
## Voicepar3rd       0.11617    0.09853   1.179
## Voicepar5th       0.21291    0.09854   2.161
## X16.minus.17      0.14167    0.04352   3.255
## PachListen       -0.38361    0.17006  -2.256
##
## Correlation of Fixed Effects:
##             (Intr) HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r Vcpr5t
## HrmnyI-V-IV -0.050
## HrmnyI-V-VI -0.042  0.326
## HrmnyIV-I-V -0.111  0.203  0.124
## Instrumntpn -0.073  0.001  0.000  0.000
## Instrmntstr -0.135  0.000  0.000  0.000  0.666
## Voicepar3rd -0.061  0.000  0.000  0.002  0.000     0.000
## Voicepar5th -0.060 -0.002 -0.001 -0.002  0.000     0.000     0.501
## X16.mins.17  0.081  0.000  0.000  0.000  0.000     0.000     0.001  0.000
## PachListen  -0.965  0.000 -0.001  0.001 -0.001     0.000     0.000  0.000
##             X16..1
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## Voicepar3rd
## Voicepar5th
## X16.mins.17
## PachListen  -0.165
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# Checking influence of KnowRob and KnowAxis
lmer_pop.test = update(lmer_pop, .~. +KnowRob:Harmony+KnowAxis:Harmony)
summary(lmer_pop.test)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Popular ~ Harmony + Instrument + Voice + X16.minus.17 + PachListen +
##     (1 | Subject) + (Harmony | Subject) + (Instrument | Subject) +
##     Harmony:KnowRob + Harmony:KnowAxis
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
```

```
##      AIC      BIC   logLik deviance df.resid
##   6662.4   6856.9  -3295.2   6590.4     1603
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8197 -0.5944  0.0455  0.6003  3.0238
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.00000  0.0000
##  Subject.1 (Intercept)      0.72275  0.8501
##            HarmonyI-V-IV    0.09156  0.3026    0.19
##            HarmonyI-V-VI    0.48038  0.6931    0.18  0.05
##            HarmonyIV-I-V    0.28548  0.5343   -0.56 -0.78 -0.57
##  Subject.2 (Intercept)      0.59092  0.7687
##            Instrumentpiano  1.80509  1.3435   -0.39
##            Instrumentstring 3.78808  1.9463   -0.93  0.70
##  Residual                   2.64552  1.6265
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                          Estimate Std. Error t value
## (Intercept)              8.329911   0.814500  10.227
## HarmonyI-V-IV           -0.005051   0.145000  -0.035
## HarmonyI-V-VI           -0.315302   0.181886  -1.734
## HarmonyIV-I-V           -0.244187   0.164264  -1.487
## Instrumentpiano         -1.148767   0.221542  -5.185
## Instrumentstring        -2.674503   0.303262  -8.819
## Voicepar3rd              0.116360   0.098464   1.182
## Voicepar5th              0.213970   0.098467   2.173
## X16.minus.17             0.148386   0.042879   3.461
## PachListen              -0.394116   0.169227  -2.329
## HarmonyI-IV-V:KnowRob    0.162030   0.096337   1.682
## HarmonyI-V-IV:KnowRob    0.161193   0.104046   1.549
## HarmonyI-V-VI:KnowRob   -0.104163   0.122021  -0.854
## HarmonyIV-I-V:KnowRob    0.058120   0.086377   0.673
## HarmonyI-IV-V:KnowAxis -0.007076   0.079606  -0.089
## HarmonyI-V-IV:KnowAxis  0.022880   0.086092   0.266
## HarmonyI-V-VI:KnowAxis  0.185525   0.101267   1.832
## HarmonyIV-I-V:KnowAxis  0.023768   0.071154   0.334
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
anova(lmer_pop.test, lmer_pop)
```

```
## Data: newdata
## Models:
## lmer_pop: Popular ~ Harmony + Instrument + Voice + X16.minus.17 + PachListen +
## lmer_pop:     (1 | Subject) + (Harmony | Subject) + (Instrument | Subject)
## lmer_pop.test: Popular ~ Harmony + Instrument + Voice + X16.minus.17 + PachListen +
## lmer_pop.test:     (1 | Subject) + (Harmony | Subject) + (Instrument | Subject) +
## lmer_pop.test:     Harmony:KnowRob + Harmony:KnowAxis
##               Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer_pop      28 6660.0 6811.2 -3302.0   6604.0
## lmer_pop.test 36 6662.4 6856.9 -3295.2   6590.4 13.543      8    0.09448 .
```

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

```r
# Compare lmer5 and exp
anova(lmer_pop, exp.1, exp)
```

```
## Data: newdata
## Models:
## exp.1: Popular ~ Harmony + Instrument + Voice + X16.minus.17 + PachListen +
## exp.1:     (1 | Subject)
## exp: Popular ~ Harmony + Instrument + Voice + Selfdeclare + OMSI +
## exp:     X16.minus.17 + ConsInstr + PachListen + ClsListen + KnowRob +
## exp:     CollegeMusic + APTheory + (1 | Subject)
## lmer_pop: Popular ~ Harmony + Instrument + Voice + X16.minus.17 + PachListen +
## lmer_pop:     (1 | Subject) + (Harmony | Subject) + (Instrument | Subject)
##          Df    AIC    BIC  logLik deviance    Chisq Chi Df Pr(>Chisq)
## exp.1    12 6956.0 7020.8 -3466.0   6932.0
## exp      19 6961.5 7064.2 -3461.8   6923.5   8.4763      7     0.2925
## lmer_pop 28 6660.0 6811.2 -3302.0   6604.0 319.5692      9     <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Check all levels of harmony
lmer_pop.1 <- lme4::lmer(Popular ~ Harmony-1 + Instrument + Voice + X16.minus.17 + PachListen +
    (1 | Subject) + (Harmony-1 | Subject) + (Instrument | Subject),
    control = lmerControl(optimizer = "bobyqa"), data = newdata, REML = F)
summary(lmer_pop.1)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Harmony - 1 + Instrument + Voice + X16.minus.17 + PachListen +
##     (1 | Subject) + (Harmony - 1 | Subject) + (Instrument | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6660.0   6811.2  -3302.0   6604.0     1611
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8248 -0.5785  0.0525  0.6017  3.1605
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
##  Subject.1 HarmonyI-IV-V    0.7906   0.8891
##            HarmonyI-V-IV    1.0138   1.0069   0.96
##            HarmonyI-V-VI    1.5372   1.2399   0.74  0.72
##            HarmonyIV-I-V    0.5163   0.7186   0.79  0.60  0.42
##  Subject.2 (Intercept)      0.5401   0.7349
##            Instrumentpiano  1.8055   1.3437   -0.40
##            Instrumentstring 3.7849   1.9455   -0.93  0.70
##  Residual                   2.6492   1.6276
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
```

```
##                   Estimate Std. Error t value
## HarmonyI-IV-V     8.42255    0.81993  10.272
## HarmonyI-V-IV     8.45338    0.82291  10.273
## HarmonyI-V-VI     8.11236    0.82994   9.775
## HarmonyIV-I-V     8.12607    0.81621   9.956
## Instrumentpiano  -1.14857    0.22159  -5.183
## Instrumentstring -2.67402    0.30317  -8.820
## Voicepar3rd       0.11617    0.09853   1.179
## Voicepar5th       0.21291    0.09854   2.161
## X16.minus.17      0.14167    0.04352   3.255
## PachListen       -0.38361    0.17006  -2.256
##
## Correlation of Fixed Effects:
##              HI-IV- HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r Vcpr5t
## HrmnyI-V-IV  0.989
## HrmnyI-V-VI  0.979  0.978
## HrmnyIV-I-V  0.986  0.980  0.969
## Instrumntpn -0.073 -0.072 -0.072 -0.073
## Instrmntstr -0.135 -0.135 -0.134 -0.136  0.666
## Voicepar3rd -0.061 -0.061 -0.060 -0.061  0.000     0.000
## Voicepar5th -0.060 -0.060 -0.059 -0.060  0.000     0.000     0.501
## X16.mins.17  0.081  0.081  0.080  0.082  0.000     0.000     0.001  0.000
## PachListen  -0.965 -0.962 -0.954 -0.969 -0.001     0.000     0.000  0.000
##              X16..1
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## Voicepar3rd
## Voicepar5th
## X16.mins.17
## PachListen  -0.165
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# Check all levels of instrument
lmer_pop.2 <- lme4::lmer(Popular ~ Instrument-1 + Harmony + Voice + X16.minus.17 + PachListen +
    (1 | Subject) + (Harmony | Subject) + (Instrument-1 | Subject),
    control = lmerControl(optimizer = "bobyqa"), data = newdata, REML = F)
summary(lmer_pop.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Instrument - 1 + Harmony + Voice + X16.minus.17 + PachListen +
##     (1 | Subject) + (Harmony | Subject) + (Instrument - 1 | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6660.0   6811.2  -3302.0   6604.0     1611
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8248 -0.5785  0.0525  0.6017  3.1605
```

```
##
## Random effects:
##  Groups    Name              Variance  Std.Dev.  Corr
##  Subject   (Intercept)       6.284e-13 7.927e-07
##  Subject.1 (Intercept)       7.906e-01 8.891e-01
##            HarmonyI-V-IV     8.761e-02 2.960e-01  0.26
##            HarmonyI-V-VI     6.934e-01 8.327e-01  0.04  0.04
##            HarmonyIV-I-V     2.928e-01 5.411e-01 -0.59 -0.87 -0.36
##  Subject.2 Instrumentguitar  5.401e-01 7.349e-01
##            Instrumentpiano   1.565e+00 1.251e+00  0.16
##            Instrumentstring  1.663e+00 1.289e+00 -0.83  0.41
##  Residual                    2.649e+00 1.628e+00
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                   Estimate Std. Error t value
## Instrumentguitar   8.42255    0.81993  10.272
## Instrumentpiano    7.27398    0.83364   8.726
## Instrumentstring   5.74853    0.83478   6.886
## HarmonyI-V-IV      0.03083    0.12172   0.253
## HarmonyI-V-VI     -0.31019    0.16749  -1.852
## HarmonyIV-I-V     -0.29648    0.13878  -2.136
## Voicepar3rd        0.11617    0.09853   1.179
## Voicepar5th        0.21291    0.09854   2.161
## X16.minus.17       0.14167    0.04352   3.255
## PachListen        -0.38361    0.17006  -2.256
##
## Correlation of Fixed Effects:
##              Instrmntg Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r Vcpr5t
## Instrumntpn   0.964
## Instrmntstr   0.933     0.963
## HrmnyI-V-IV -0.050    -0.049    -0.049
## HrmnyI-V-VI -0.042    -0.041    -0.041     0.326
## HrmnyIV-I-V -0.111    -0.109    -0.109     0.203  0.124
## Voicepar3rd -0.061    -0.060    -0.060     0.000  0.000  0.002
## Voicepar5th -0.060    -0.059    -0.059    -0.002 -0.001 -0.002  0.501
## X16.mins.17  0.081     0.080     0.080     0.000  0.000  0.000  0.001  0.000
## PachListen  -0.965    -0.949    -0.948     0.000 -0.001  0.001  0.000  0.000
##              X16..1
## Instrumntpn
## Instrmntstr
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Voicepar3rd
## Voicepar5th
## X16.mins.17
## PachListen  -0.165
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# Check all levels of voice
lmer_pop.3 <- lme4::lmer(Popular ~ Voice-1 + Harmony + Instrument + X16.minus.17 + PachListen +
    (1 | Subject) + (Harmony | Subject) + (Instrument | Subject),
```

```r
    control = lmerControl(optimizer = "bobyqa"), data = newdata, REML = F)
summary(lmer_pop.3)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Popular ~ Voice - 1 + Harmony + Instrument + X16.minus.17 + PachListen +
##     (1 | Subject) + (Harmony | Subject) + (Instrument | Subject)
##    Data: newdata
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6660.0   6811.2  -3302.0   6604.0     1611
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8248 -0.5785  0.0525  0.6017  3.1605
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      1.546e-09 3.932e-05
##  Subject.1 (Intercept)      7.906e-01 8.891e-01
##            HarmonyI-V-IV    8.761e-02 2.960e-01  0.26
##            HarmonyI-V-VI    6.934e-01 8.327e-01  0.04  0.04
##            HarmonyIV-I-V    2.928e-01 5.411e-01 -0.59 -0.87 -0.36
##  Subject.2 (Intercept)      5.401e-01 7.349e-01
##            Instrumentpiano  1.805e+00 1.344e+00 -0.40
##            Instrumentstring 3.785e+00 1.945e+00 -0.93  0.70
##  Residual                   2.649e+00 1.628e+00
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                  Estimate Std. Error t value
## Voicecontrary     8.42255    0.81993  10.272
## Voicepar3rd       8.53872    0.81986  10.415
## Voicepar5th       8.63546    0.81996  10.532
## HarmonyI-V-IV     0.03083    0.12172   0.253
## HarmonyI-V-VI    -0.31019    0.16749  -1.852
## HarmonyIV-I-V    -0.29648    0.13878  -2.136
## Instrumentpiano  -1.14857    0.22159  -5.183
## Instrumentstring -2.67402    0.30317  -8.820
## X16.minus.17      0.14167    0.04352   3.255
## PachListen       -0.38361    0.17006  -2.256
##
## Correlation of Fixed Effects:
##             Vccntr Vcpr3r Vcpr5t HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts
## Voicepar3rd  0.993
## Voicepar5th  0.993  0.993
## HrmnyI-V-IV -0.050 -0.050 -0.050
## HrmnyI-V-VI -0.042 -0.042 -0.042  0.326
## HrmnyIV-I-V -0.111 -0.111 -0.112  0.203  0.124
## Instrumntpn -0.073 -0.073 -0.073  0.001  0.000  0.000
## Instrmntstr -0.135 -0.135 -0.135  0.000  0.000  0.000  0.666
## X16.mins.17  0.081  0.082  0.082  0.000  0.000  0.000  0.000     0.000
## PachListen  -0.965 -0.965 -0.965  0.000 -0.001  0.001 -0.001     0.000
```

```
##             X16..1
## Voicepar3rd
## Voicepar5th
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## X16.mins.17
## PachListen  -0.165
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
vif(lmer_pop)
```

```
##                  GVIF Df GVIF^(1/(2*Df))
## Harmony      1.000019  3        1.000003
## Instrument   1.000004  2        1.000001
## Voice        1.000020  2        1.000005
## X16.minus.17 1.028079  1        1.013942
## PachListen   1.028079  1        1.013943
```

```
# plot residuals
# mlm facet plot
dev.new()
# residuals randomly scattered around zero
res.10 <- r.cond(lmer_pop)     ## standardized conditional residuals
robust.sd.10 <- diff(quantile(res.10,c(.025,.975)))/(2*1.96)
res.10 <- res.10/robust.sd.10
fit.10 <- yhat.cond(lmer_pop)

newdata.10 <- data.frame(subject=newdata$Subject,res_lmer_pop=res.10,fit_lmer_pop=fit.10)

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

mlm_facets(newdata.10,"subject",x="fit_lmer_pop",y="res_lmer_pop",params=resparams,
           lty=c(1,2,2),size=c(1,1,1))

res_lmer_pop=res.10
fit_lmer_pop=fit.10
dev.new()

# model assumption met
par(mfrow=c(2,2))
plot(fit_lmer_pop,res_lmer_pop)
binnedplot(fit_lmer_pop, res_lmer_pop)
```

```
# Building Model (3)
data4 <- newdata
######### Dichotomize Selfdeclare at 1,2 vs 3,4,5,6 ##########
data4$Musician <- rep(0, nrow(data4))
data4$Musician[which(data4$Selfdeclare >= 3)] <- 1
```

```r
data4$Musician <- as.factor(data4$Musician)

step_selfdeclare <- stepAIC(lm(Classical ~ Musician * (Harmony + Instrument + Voice +
                   OMSI+X16.minus.17+ConsInstr+
                  ConsNotes+Instr.minus.Notes +PachListen+ClsListen+KnowRob+
                  KnowAxis+X1990s2000s+X1990s2000s.minus.1960s1970s+
                  CollegeMusic+NoClass+APTheory+Composing+PianoPlay+
                  GuitarPlay),data4),
                  direction = "backward", k=log(1639), trace=0)
summary(step_selfdeclare)
```

```
##
## Call:
## lm(formula = Classical ~ Musician + Harmony + Instrument + X16.minus.17 +
##     ConsNotes + PachListen + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + NoClass + APTheory + Composing + PianoPlay +
##     GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
##     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
##     Musician:KnowRob + Musician:KnowAxis + Musician:X1990s2000s +
##     Musician:NoClass, data = data4)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.7525 -1.3208 -0.0296  1.1709  7.4836
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                1.549e+00  5.245e-01   2.954  0.00318 **
## Musician1                  7.790e+00  9.458e-01   8.237 3.63e-16 ***
## HarmonyI-V-IV             -8.889e-03  1.861e-01  -0.048  0.96191
## HarmonyI-V-VI              3.171e-01  1.868e-01   1.698  0.08977 .
## HarmonyIV-I-V              1.022e-01  1.861e-01   0.549  0.58293
## Instrumentpiano            1.874e+00  1.613e-01  11.615  < 2e-16 ***
## Instrumentstring           3.949e+00  1.616e-01  24.435  < 2e-16 ***
## X16.minus.17              -6.659e-02  3.090e-02  -2.155  0.03130 *
## ConsNotes                 -3.926e-01  4.520e-02  -8.686  < 2e-16 ***
## PachListen                 6.615e-01  9.421e-02   7.022 3.22e-12 ***
## ClsListen                 -4.785e-02  5.604e-02  -0.854  0.39325
## KnowRob                    2.935e-01  6.764e-02   4.339 1.52e-05 ***
## KnowAxis                  -1.052e-01  4.691e-02  -2.243  0.02506 *
## X1990s2000s               -1.782e-01  5.858e-02  -3.043  0.00238 **
## NoClass                    1.267e-01  1.336e-01   0.948  0.34305
## APTheory1                  7.824e-01  1.621e-01   4.827 1.52e-06 ***
## Composing                  7.013e-01  6.533e-02  10.734  < 2e-16 ***
## PianoPlay                  4.740e-01  3.923e-02  12.085  < 2e-16 ***
## GuitarPlay                 3.214e-01  8.090e-02   3.973 7.41e-05 ***
## Musician1:HarmonyI-V-IV    3.415e-02  2.769e-01   0.123  0.90186
## Musician1:HarmonyI-V-VI    1.200e+00  2.773e-01   4.328 1.60e-05 ***
## Musician1:HarmonyIV-I-V   -7.168e-05  2.767e-01   0.000  0.99979
## Musician1:Instrumentpiano -6.154e-01  2.407e-01  -2.556  0.01067 *
## Musician1:Instrumentstring -9.210e-01  2.388e-01  -3.856  0.00012 ***
## Musician1:X16.minus.17    -4.044e-01  5.735e-02  -7.051 2.64e-12 ***
## Musician1:ConsNotes        1.060e+00  1.005e-01  10.551  < 2e-16 ***
## Musician1:PachListen      -1.417e+00  1.608e-01  -8.812  < 2e-16 ***
```

```
## Musician1:ClsListen        -9.268e-01  1.130e-01  -8.199 4.90e-16 ***
## Musician1:KnowRob          -2.930e-01  9.013e-02  -3.251  0.00118 **
## Musician1:KnowAxis          4.135e-01  6.861e-02   6.026 2.08e-09 ***
## Musician1:X1990s2000s      -6.188e-01  1.091e-01  -5.674 1.65e-08 ***
## Musician1:NoClass          -7.524e-01  1.482e-01  -5.077 4.27e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.974 on 1607 degrees of freedom
## Multiple R-squared:  0.4821, Adjusted R-squared:  0.4721
## F-statistic: 48.25 on 31 and 1607 DF,  p-value: < 2.2e-16
```

```
# step BIC picks model that are more scientically makes sense

lmer5 <- lme4::lmer(Classical ~  Musician + Harmony + Instrument+Voice+ X16.minus.17 +
    ConsNotes + PachListen + ClsListen + KnowRob + KnowAxis +
    X1990s2000s + NoClass + APTheory + Composing + PianoPlay +
    GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
    Musician:ConsNotes + Musician:PachListen + Musician:ClsListen + Musician:KnowRob +
    Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass + (1|Subject),
    control = lmerControl(optimizer = "bobyqa"),
    REML = F, data = data4)
summary(lmer5)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Musician + Harmony + Instrument + Voice + X16.minus.17 +
##     ConsNotes + PachListen + ClsListen + KnowRob + KnowAxis +
##     X1990s2000s + NoClass + APTheory + Composing + PianoPlay +
##     GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
##     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
##     Musician:KnowRob + Musician:KnowAxis + Musician:X1990s2000s +
##     Musician:NoClass + (1 | Subject)
##    Data: data4
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6869.2   7063.6  -3398.6   6797.2     1603
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2290 -0.6203 -0.0168  0.5986  3.7445
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.2038   0.4514
##  Residual             3.5903   1.8948
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                 1.7865041  0.8568423   2.085
## Musician1                   7.7025411  1.5496174   4.971
## HarmonyI-V-IV              -0.0088889  0.1786450  -0.050
## HarmonyI-V-VI               0.3177827  0.1792737   1.773
## HarmonyIV-I-V               0.1022222  0.1786450   0.572
```

```
## Instrumentpiano             1.8725200  0.1548439  12.093
## Instrumentstring            3.9467103  0.1551170  25.443
## Voicepar3rd                -0.3646322  0.1146842  -3.179
## Voicepar5th                -0.3462141  0.1146834  -3.019
## X16.minus.17               -0.0666691  0.0517303  -1.289
## ConsNotes                  -0.3927171  0.0756375  -5.192
## PachListen                  0.6594674  0.1573953   4.190
## ClsListen                  -0.0456230  0.0936444  -0.487
## KnowRob                     0.2948678  0.1131395   2.606
## KnowAxis                   -0.1058664  0.0784392  -1.350
## X1990s2000s                -0.1759242  0.0980132  -1.795
## NoClass                     0.1277865  0.2221469   0.575
## APTheory1                   0.8044045  0.2690857   2.989
## Composing                   0.6967912  0.1089783   6.394
## PianoPlay                   0.4676432  0.0650286   7.191
## GuitarPlay                  0.3170440  0.1351648   2.346
## Musician1:HarmonyI-V-IV     0.0382888  0.2657588   0.144
## Musician1:HarmonyI-V-VI     1.2015382  0.2661842   4.514
## Musician1:HarmonyIV-I-V    -0.0002697  0.2655583  -0.001
## Musician1:Instrumentpiano  -0.5825705  0.2312179  -2.520
## Musician1:Instrumentstring -0.9189326  0.2292508  -4.008
## Musician1:X16.minus.17     -0.3999360  0.0956403  -4.182
## Musician1:ConsNotes         1.0541081  0.1673938   6.297
## Musician1:PachListen       -1.4057509  0.2687370  -5.231
## Musician1:ClsListen        -0.9161157  0.1882640  -4.866
## Musician1:KnowRob          -0.2936590  0.1500261  -1.957
## Musician1:KnowAxis          0.4088287  0.1141732   3.581
## Musician1:X1990s2000s      -0.6096047  0.1817540  -3.354
## Musician1:NoClass          -0.7539423  0.2467188  -3.056
```

```
# X16.minus.17,ClsListen, KnowAxis,X1990s2000s,NoClass,mus:KnowRob
# not significant
exactRLRT(lmer5)
```

```
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
## data:
## RLRT = 61.63, p-value < 2.2e-16
```

```
anova(lmer5, step_selfdeclare)
```

```
## Data: data4
## Models:
## step_selfdeclare: Classical ~ Musician + Harmony + Instrument + X16.minus.17 +
## step_selfdeclare:     ConsNotes + PachListen + ClsListen + KnowRob + KnowAxis +
## step_selfdeclare:     X1990s2000s + NoClass + APTheory + Composing + PianoPlay +
## step_selfdeclare:     GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
## step_selfdeclare:     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
## step_selfdeclare:     Musician:KnowRob + Musician:KnowAxis + Musician:X1990s2000s +
## step_selfdeclare:     Musician:NoClass
## lmer5: Classical ~ Musician + Harmony + Instrument + Voice + X16.minus.17 +
## lmer5:     ConsNotes + PachListen + ClsListen + KnowRob + KnowAxis +
## lmer5:     X1990s2000s + NoClass + APTheory + Composing + PianoPlay +
```

```
## lmer5:     GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
## lmer5:     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
## lmer5:     Musician:KnowRob + Musician:KnowAxis + Musician:X1990s2000s +
## lmer5:     Musician:NoClass + (1 | Subject)
##                  Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## step_selfdeclare 33 6914.4 7092.7 -3424.2   6848.4
## lmer5            36 6869.2 7063.6 -3398.6   6797.2 51.226      3  4.378e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# should include (1|Subject)

# Take out not significant variables,
lmer5.1 <- lme4::lmer(Classical ~ Musician + Harmony + Instrument+Voice+
    ConsNotes + PachListen + KnowRob + APTheory + Composing + PianoPlay +
    GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
    Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
    Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass + (1|Subject),
    control = lmerControl(optimizer = "bobyqa"),
    REML = F, data = data4)
# fit other random effects
lmer_6 <- ffRanefLMER.fnc(lmer5.1, ran.effects = c("(Harmony|Subject)","(Instrument|Subject)","(Voice|S
                                        log.file=F)
```

```
## evaluating addition of (Harmony|Subject) to model
##  log-likelihood ratio test p-value = 0.0001045479
##  adding (Harmony|Subject) to model
## evaluating addition of (Instrument|Subject) to model
##  log-likelihood ratio test p-value = 2.023426e-60
##  adding (Instrument|Subject) to model
## evaluating addition of (Voice|Subject) to model
##  log-likelihood ratio test p-value = 0.6202531
##  not adding (Voice|Subject) to model
```

```r
summary(lmer_6)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Classical ~ Musician + Harmony + Instrument + Voice + ConsNotes +
##     PachListen + KnowRob + APTheory + Composing + PianoPlay +
##     GuitarPlay + (1 | Subject) + (Harmony | Subject) + (Instrument |
##     Subject) + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
##     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
##     Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass
##    Data: data4
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6573.9   6849.4  -3236.0   6471.9     1588
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5203 -0.5690  0.0015  0.5527  3.6341
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
```

```
##  Subject   (Intercept)      0.00000  0.0000
##  Subject.1 (Intercept)      0.25354  0.5035
##            HarmonyI-V-IV    0.05634  0.2374   -0.86
##            HarmonyI-V-VI    1.23888  1.1130   -0.27  0.72
##            HarmonyIV-I-V    0.02445  0.1564    0.72 -0.27  0.48
##  Subject.2 (Intercept)      0.74429  0.8627
##            Instrumentpiano  2.09191  1.4463   -0.96
##            Instrumentstring 3.26438  1.8068   -0.82  0.62
##  Residual                   2.51771  1.5867
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                 1.226275   0.841907   1.457
## Musician1                   8.214199   1.523991   5.390
## HarmonyI-V-IV              -0.008889   0.156950  -0.057
## HarmonyI-V-VI               0.309182   0.268571   1.151
## HarmonyIV-I-V               0.102222   0.152832   0.669
## Instrumentpiano             1.880440   0.317007   5.932
## Instrumentstring            3.956100   0.384018  10.302
## Voicepar3rd                -0.359320   0.096057  -3.741
## Voicepar5th                -0.342563   0.096060  -3.566
## ConsNotes                  -0.323952   0.071240  -4.547
## PachListen                  0.746766   0.151454   4.931
## KnowRob                     0.120258   0.068688   1.751
## APTheory1                   0.503903   0.257468   1.957
## Composing                   0.784157   0.103025   7.611
## PianoPlay                   0.529306   0.063341   8.356
## GuitarPlay                  0.243276   0.124741   1.950
## Musician1:HarmonyI-V-IV     0.034386   0.233395   0.147
## Musician1:HarmonyI-V-VI     1.229355   0.397931   3.089
## Musician1:HarmonyIV-I-V     0.002972   0.227149   0.013
## Musician1:Instrumentpiano  -0.618085   0.470303  -1.314
## Musician1:Instrumentstring -0.928322   0.568241  -1.634
## Musician0:X16.minus.17      0.001766   0.050141   0.035
## Musician1:X16.minus.17     -0.527580   0.075942  -6.947
## Musician1:ConsNotes         1.015596   0.160933   6.311
## Musician1:PachListen       -1.509119   0.257598  -5.858
## Musician0:ClsListen        -0.084401   0.088439  -0.954
## Musician1:ClsListen        -1.142037   0.179967  -6.346
## Musician0:KnowAxis         -0.015310   0.068956  -0.222
## Musician1:KnowAxis          0.368075   0.074069   4.969
## Musician0:X1990s2000s      -0.185642   0.095008  -1.954
## Musician1:X1990s2000s      -0.724623   0.156763  -4.622
## Musician0:NoClass           0.150074   0.214673   0.699
## Musician1:NoClass          -0.544017   0.104251  -5.218
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
anova(lmer5, lmer5.1, lmer_6)
```

```
## Data: data4
## Models:
## lmer5.1: Classical ~ Musician + Harmony + Instrument + Voice + ConsNotes +
## lmer5.1:     PachListen + KnowRob + APTheory + Composing + PianoPlay +
```

```
## lmer5.1:     GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
## lmer5.1:     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
## lmer5.1:     Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass +
## lmer5.1:     (1 | Subject)
## lmer5: Classical ~ Musician + Harmony + Instrument + Voice + X16.minus.17 +
## lmer5:     ConsNotes + PachListen + ClsListen + KnowRob + KnowAxis +
## lmer5:     X1990s2000s + NoClass + APTheory + Composing + PianoPlay +
## lmer5:     GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
## lmer5:     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
## lmer5:     Musician:KnowRob + Musician:KnowAxis + Musician:X1990s2000s +
## lmer5:     Musician:NoClass + (1 | Subject)
## lmer_6: Classical ~ Musician + Harmony + Instrument + Voice + ConsNotes +
## lmer_6:     PachListen + KnowRob + APTheory + Composing + PianoPlay +
## lmer_6:     GuitarPlay + (1 | Subject) + (Harmony | Subject) + (Instrument |
## lmer_6:     Subject) + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
## lmer_6:     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
## lmer_6:     Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass
##         Df    AIC    BIC  logLik deviance    Chisq Chi Df Pr(>Chisq)
## lmer5.1 35 6870.9 7059.9 -3400.4   6800.9
## lmer5   36 6869.2 7063.6 -3398.6   6797.2   3.6807      1    0.05505 .
## lmer_6  51 6573.9 6849.4 -3236.0   6471.9 325.2670     15    < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# lmer_6 has lower AIC
# lmer_6 : Classical ~ Musician + Harmony + Instrument + Voice + ConsNotes +
#    PachListen + KnowRob + APTheory + Composing + PianoPlay +
#    GuitarPlay + (1 | Subject) + (Harmony | Subject) + (Instrument |
#    Subject) + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
#    Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
#    Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass


# Checking all levels of harmony
lmer.c.1 <- lme4::lmer(Classical ~ Harmony-1 + Instrument + Voice + ConsNotes +
    PachListen + KnowRob + APTheory + Composing + PianoPlay + Musician+
    GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
    Musician:ConsNotes + Musician:PachListen + Musician:ClsListen + (1|Subject)+
    (Instrument|Subject)+(Harmony-1|Subject)+
    Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass ,
    control = lmerControl(optimizer = "bobyqa"),
    REML = F, data = data4)
summary(lmer.c.1)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Harmony - 1 + Instrument + Voice + ConsNotes + PachListen +
##     KnowRob + APTheory + Composing + PianoPlay + Musician + GuitarPlay +
##     Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
##     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
##     (1 | Subject) + (Instrument | Subject) + (Harmony - 1 | Subject) +
##     Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass
##    Data: data4
## Control: lmerControl(optimizer = "bobyqa")
##
```

```
##      AIC      BIC   logLik deviance df.resid
##   6573.9   6849.4  -3236.0   6471.9     1588
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5203 -0.5690  0.0015  0.5527  3.6341
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
##  Subject.1 (Intercept)      0.7443   0.8627
##            Instrumentpiano  2.0919   1.4463   -0.96
##            Instrumentstring 3.2644   1.8068   -0.82  0.62
##  Subject.2 HarmonyI-IV-V    0.2535   0.5035
##            HarmonyI-V-IV    0.1034   0.3215   0.93
##            HarmonyI-V-VI    1.1885   1.0902   0.19  0.54
##            HarmonyIV-I-V    0.3906   0.6250   0.98  0.98  0.35
##  Residual                   2.5177   1.5867
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                            Estimate Std. Error t value
## HarmonyI-IV-V              1.226273   0.841907   1.457
## HarmonyI-V-IV              1.217384   0.838332   1.452
## HarmonyI-V-VI              1.535455   0.863920   1.777
## HarmonyIV-I-V              1.328495   0.845157   1.572
## Instrumentpiano            1.880440   0.317007   5.932
## Instrumentstring           3.956100   0.384017  10.302
## Voicepar3rd               -0.359320   0.096057  -3.741
## Voicepar5th               -0.342563   0.096060  -3.566
## ConsNotes                 -0.323952   0.071240  -4.547
## PachListen                 0.746766   0.151454   4.931
## KnowRob                    0.120258   0.068688   1.751
## APTheory1                  0.503904   0.257468   1.957
## Composing                  0.784157   0.103025   7.611
## PianoPlay                  0.529305   0.063341   8.356
## Musician1                  8.214199   1.523991   5.390
## GuitarPlay                 0.243276   0.124741   1.950
## HarmonyI-V-IV:Musician1    0.034386   0.233395   0.147
## HarmonyI-V-VI:Musician1    1.229355   0.397931   3.089
## HarmonyIV-I-V:Musician1    0.002972   0.227149   0.013
## Instrumentpiano:Musician1 -0.618085   0.470303  -1.314
## Instrumentstring:Musician1 -0.928322   0.568241  -1.634
## Musician0:X16.minus.17     0.001766   0.050141   0.035
## Musician1:X16.minus.17    -0.527580   0.075942  -6.947
## ConsNotes:Musician1        1.015596   0.160933   6.311
## PachListen:Musician1      -1.509120   0.257598  -5.858
## Musician0:ClsListen       -0.084401   0.088439  -0.954
## Musician1:ClsListen       -1.142036   0.179967  -6.346
## Musician0:KnowAxis        -0.015310   0.068956  -0.222
## Musician1:KnowAxis         0.368075   0.074069   4.969
## Musician0:X1990s2000s     -0.185642   0.095008  -1.954
## Musician1:X1990s2000s     -0.724623   0.156763  -4.622
## Musician0:NoClass          0.150073   0.214673   0.699
```

```
## Musician1:NoClass          -0.544017   0.104251  -5.218
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
# Checking all levels of instrument
lmer.c.2 <- lme4::lmer(Classical ~ Instrument-1 + Harmony+ Voice + ConsNotes +
    PachListen + KnowRob + APTheory + Composing + PianoPlay + Musician+
    GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
    Musician:ConsNotes + Musician:PachListen + Musician:ClsListen + (1|Subject)+
    (Instrument-1|Subject)+(Harmony|Subject)+
    Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass ,
    control = lmerControl(optimizer = "bobyqa"),
    REML = F, data = data4)
summary(lmer.c.2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Instrument - 1 + Harmony + Voice + ConsNotes + PachListen +
##     KnowRob + APTheory + Composing + PianoPlay + Musician + GuitarPlay +
##     Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
##     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
##     (1 | Subject) + (Instrument - 1 | Subject) + (Harmony | Subject) +
##     Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass
##    Data: data4
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6574.7   6850.2  -3236.4   6472.7     1588
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5058 -0.5675  0.0054  0.5456  3.6126
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.008914 0.09442
##  Subject.1 Instrumentguitar 0.880641 0.93842
##            Instrumentpiano  0.595923 0.77196  -0.43
##            Instrumentstring 1.565537 1.25121  -0.35  0.02
##  Subject.2 (Intercept)      0.114790 0.33881
##            HarmonyI-V-IV    0.056274 0.23722  -1.00
##            HarmonyI-V-VI    1.224189 1.10643  -0.57  0.57
##            HarmonyIV-I-V    0.032978 0.18160   0.54 -0.54  0.39
##  Residual                   2.513898 1.58553
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                            Estimate Std. Error t value
## Instrumentguitar           1.239181   0.843729   1.469
## Instrumentpiano            3.119664   0.836963   3.727
## Instrumentstring           5.195451   0.859887   6.042
## HarmonyI-V-IV             -0.008889   0.156833  -0.057
## HarmonyI-V-VI              0.309293   0.267411   1.157
## HarmonyIV-I-V              0.102222   0.153834   0.664
## Voicepar3rd               -0.359243   0.095985  -3.743
```

```
## Voicepar5th              -0.342937   0.095988  -3.573
## ConsNotes                -0.327172   0.071399  -4.582
## PachListen                0.751209   0.151783   4.949
## KnowRob                   0.135771   0.068748   1.975
## APTheory1                 0.552821   0.257764   2.145
## Composing                 0.774877   0.103204   7.508
## PianoPlay                 0.515994   0.063393   8.140
## Musician1                 7.978846   1.526671   5.226
## GuitarPlay                0.230529   0.124962   1.845
## HarmonyI-V-IV:Musician1   0.034004   0.233228   0.146
## HarmonyI-V-VI:Musician1   1.227992   0.396215   3.099
## HarmonyIV-I-V:Musician1   0.002107   0.228627   0.009
## Instrumentpiano:Musician1  -0.617530   0.470457  -1.313
## Instrumentstring:Musician1 -0.928492   0.568267  -1.634
## Musician0:X16.minus.17   -0.012501   0.050251  -0.249
## Musician1:X16.minus.17   -0.491742   0.076043  -6.467
## ConsNotes:Musician1       1.004368   0.161192   6.231
## PachListen:Musician1     -1.508573   0.258097  -5.845
## Musician0:ClsListen      -0.078507   0.088633  -0.886
## Musician1:ClsListen      -1.090727   0.180243  -6.051
## Musician0:KnowAxis       -0.022187   0.069100  -0.321
## Musician1:KnowAxis        0.336577   0.074105   4.542
## Musician0:X1990s2000s    -0.189865   0.095215  -1.994
## Musician1:X1990s2000s    -0.695685   0.156955  -4.432
## Musician0:NoClass         0.143913   0.215186   0.669
## Musician1:NoClass        -0.556724   0.104477  -5.329
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# Checking all levels of voice
lmer.c.3 <- lme4::lmer(Classical ~ Voice-1 + Harmony+ Instrument + ConsNotes +
    PachListen + KnowRob + APTheory + Composing + PianoPlay + Musician+
    GuitarPlay + Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
    Musician:ConsNotes + Musician:PachListen + Musician:ClsListen + (1|Subject)+
    (Instrument|Subject)+(Harmony|Subject)+
    Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass ,
    control = lmerControl(optimizer = "bobyqa"),
    REML = F, data = data4)
summary(lmer.c.3)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## Classical ~ Voice - 1 + Harmony + Instrument + ConsNotes + PachListen +
##     KnowRob + APTheory + Composing + PianoPlay + Musician + GuitarPlay +
##     Musician:Harmony + Musician:Instrument + Musician:X16.minus.17 +
##     Musician:ConsNotes + Musician:PachListen + Musician:ClsListen +
##     (1 | Subject) + (Instrument | Subject) + (Harmony | Subject) +
##     Musician:KnowAxis + Musician:X1990s2000s + Musician:NoClass
##    Data: data4
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   6573.9   6849.4  -3236.0   6471.9     1588
##
## Scaled residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -4.5203 -0.5690  0.0015  0.5527  3.6341
##
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr
##  Subject   (Intercept)      7.314e-14 2.704e-07
##  Subject.1 (Intercept)      7.443e-01 8.627e-01
##            Instrumentpiano  2.092e+00 1.446e+00 -0.96
##            Instrumentstring 3.264e+00 1.807e+00 -0.82  0.62
##  Subject.2 (Intercept)      2.535e-01 5.035e-01
##            HarmonyI-V-IV    5.634e-02 2.374e-01 -0.86
##            HarmonyI-V-VI    1.239e+00 1.113e+00 -0.27  0.72
##            HarmonyIV-I-V    2.445e-02 1.564e-01  0.72 -0.27  0.48
##  Residual                   2.518e+00 1.587e+00
## Number of obs: 1639, groups:  Subject, 46
##
## Fixed effects:
##                             Estimate Std. Error t value
## Voicecontrary               1.226274   0.841907   1.457
## Voicepar3rd                 0.866953   0.841909   1.030
## Voicepar5th                 0.883711   0.841903   1.050
## HarmonyI-V-IV              -0.008889   0.156950  -0.057
## HarmonyI-V-VI               0.309182   0.268571   1.151
## HarmonyIV-I-V               0.102222   0.152832   0.669
## Instrumentpiano             1.880440   0.317007   5.932
## Instrumentstring            3.956100   0.384017  10.302
## ConsNotes                  -0.323952   0.071240  -4.547
## PachListen                  0.746766   0.151454   4.931
## KnowRob                     0.120258   0.068688   1.751
## APTheory1                   0.503903   0.257468   1.957
## Composing                   0.784157   0.103025   7.611
## PianoPlay                   0.529305   0.063341   8.356
## Musician1                   8.214200   1.523991   5.390
## GuitarPlay                  0.243276   0.124741   1.950
## HarmonyI-V-IV:Musician1     0.034386   0.233395   0.147
## HarmonyI-V-VI:Musician1     1.229355   0.397931   3.089
## HarmonyIV-I-V:Musician1     0.002972   0.227149   0.013
## Instrumentpiano:Musician1  -0.618085   0.470303  -1.314
## Instrumentstring:Musician1 -0.928322   0.568241  -1.634
## Musician0:X16.minus.17      0.001766   0.050141   0.035
## Musician1:X16.minus.17     -0.527580   0.075942  -6.947
## ConsNotes:Musician1         1.015596   0.160933   6.311
## PachListen:Musician1       -1.509120   0.257598  -5.858
## Musician0:ClsListen        -0.084401   0.088439  -0.954
## Musician1:ClsListen        -1.142036   0.179967  -6.346
## Musician0:KnowAxis         -0.015310   0.068956  -0.222
## Musician1:KnowAxis          0.368075   0.074069   4.969
## Musician0:X1990s2000s      -0.185642   0.095008  -1.954
## Musician1:X1990s2000s      -0.724623   0.156763  -4.622
## Musician0:NoClass           0.150073   0.214673   0.699
## Musician1:NoClass          -0.544017   0.104251  -5.218
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
# check multicollinearity
vif(lmer_6)
```

```
##                              GVIF Df GVIF^(1/(2*Df))
## Musician                92.185592  1        9.601333
## Harmony                  6.127982  3        1.352756
## Instrument               3.366889  2        1.354588
## Voice                    1.000101  2        1.000025
## ConsNotes                3.138295  1        1.771523
## PachListen               2.196291  1        1.481989
## KnowRob                  2.363193  1        1.537268
## APTheory                 1.902246  1        1.379219
## Composing                3.890611  1        1.972463
## PianoPlay                1.865719  1        1.365913
## GuitarPlay               7.068154  1        2.658600
## Musician:Harmony         9.112344  3        1.445235
## Musician:Instrument      8.927307  2        1.728543
## Musician:X16.minus.17    8.941799  2        1.729244
## Musician:ConsNotes      16.423267  1        4.052563
## Musician:PachListen     62.374689  1        7.897765
## Musician:ClsListen      27.784681  2        2.295891
## Musician:KnowAxis        4.545156  2        1.460115
## Musician:X1990s2000s    62.836151  2        2.815480
## Musician:NoClass         5.701017  2        1.545212
```

```r
# plot residuals
# mlm facet plot
dev.new()

# random pattern. scatter around zero
res.6 <- r.cond(lmer_6)     ## standardized conditional residuals
robust.sd.6 <- diff(quantile(res.6,c(.025,.975)))/(2*1.96)
res.6 <- res.6/robust.sd.6
fit.6 <- yhat.cond(lmer_6)

newdata.6 <- data.frame(subject=newdata$Subject,res_lmer6=res.6,fit_lmer6=fit.6)

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

mlm_facets(newdata.6,"subject",x="fit_lmer6",y="res_lmer6",params=resparams,
           lty=c(1,2,2),size=c(1,1,1))

res_lmer6=res.6
fit_lmer6=fit.6
dev.new()
par(mfrow=c(2,2))

plot(fit_lmer6,res_lmer6)
abline(h=0)
abline(h=2,lty=2)
abline(h=-2,lty=2)
```

```r
qqnorm(res_lmer6)
qqline(res_lmer6)

# model assumption met
par(mfrow=c(2,2))
plot(fit_lmer6,res_lmer6)
binnedplot(fit_lmer6, res_lmer6)
```