Exploring the Nature of Music by Genres : What Makes the Music Sound Classical or Popular?

Justin Kim

Carnegie Mellon University
justink2@andrew.cmu.edu

Abstract

The purpose of this study is to investigate how characteristics of music - instrument, harmony, voice - and audiences' personal preferences or experiences of music affect whether the music sounds classical or popular. We examine data from Jimenez et al.(2015), mainly by performing linear mixed effects models. We find statistically significant evidence that instrument plays the most significant role in how the music sounds classical. We also find that musicians tend to find the sound less classical than non-musicians do, depending on their musical "traits." Lastly, we find that attributes of music and audiences' musical preference and experiences play different roles in terms of whether the music sound classical or popular.

---

1. Introduction

Music has always been an indispensable part of our daily lives, as it plays a role not just as a form of entertainment, but also as a means of expressing emotions and conveying messages. Its "versatility" comes from the fact that it consists of various kinds of genres. From classical music to hip-hop, music encompasses audiences of all generations. By combining sounds of different instruments, harmonies, and voices, different kinds of music are generated.

However, such fact raises an intriguing question; what characteristics of music make people to differentiate music by genre? For instance, when given a piece of music, some people identify it as a country music, while others hear it as a popular music. Some claim that instruments played defines the genre of the song, while others claim that harmony is the significant factor. To explore such nature of music, we compare and examine audiences' responses to music of two genres that are often recognized to be very different, classical and popular music. Specifically, based on the ratings given by audiences on two different scales, we will address the following questions:

- What experimental factor or combinations of factors - instrument, harmonic motion, and voice leading - has the strongest influence on 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, vs. popular, ratings?

Regarding experimental factors, we investigate in depth on whether instrument has the strongest influence among all three factors and whether harmonic motion of I-V-vi and voice leading of contrary motion have a particularly strong association with classical ratings.

2. Methods

The data used is from Jimenez et al.(2015). The data was collected by conducting a designed experiment intended to measure the influence of instrument, harmonic motion, and voice leading on listeners' identification of music as "classical" or "popular." They presented 36 musical stimuli to 70 participants and asked the participants to rate the music on two different scales:

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

Participants were told that a piece could be rated as both classical and popular, neither classical nor popular, or mostly classical and not popular (or vice versa), so that the scales should have functioned more or less independently.

The 36 stimuli were chosen by completely crossing these factors:

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

Beside the variables discussed above, data used is represented in the variables illustrated in Figure 1.

For the given data, we perform several data cleaning methods. We remove variables "first12," "X," "X1stInstr" and "X2ndInstr," which are not used in the analysis or have a lot of missing values. As many other variables also have missing values, we first simply remove observations that have missing values in "Classical" or "Popular," as they are used as response variables for linear models in analysis and only few missing values exist. Then, we perform random forest imputation as it effectively deals with missing values in both continuous and categorical variables using simple codes.[1]

There are also "mistyped" values in the data. Variable "ConsInstr," which was measured in discrete numbers from 0 to 5, has values with decimal places. Thus, we round up the decimals and recalculate "Instr.minus.Notes." Furthermore, for the analysis, we dichotomize variable "Selfdeclare" such that it can be expressed as a binary variable.

For the analysis, based on the exploratory data analysis[2], we construct statistically significant linear mixed effects models. For each research question, we construct linear mixed effects model with different response and predictor variables that fully explains the research question. We select variables based on our intuition on the data and variable selection methods based on AIC, log-likelihood ratio test, and ANOVA test.[3]

[1]Appx : 6.9
[2]Appx : 6.1
[3]Appx : 6.2-6.8

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

Figure 1: Variables in the dataset

3. Results

3.1. *Influence of instrument, harmonic motion, and voice leading on classical ratings*

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	4.241e+00	2.513e-01	9.312e+01	16.878	< 2e-16	***
Instrumentpiano	1.370e+00	1.710e-01	7.012e+01	8.008	1.75e-11	***
Instrumentstring	3.127e+00	2.366e-01	6.998e+01	13.216	< 2e-16	***
HarmonyI-V-IV	1.710e-01	1.616e-01	6.614e+02	1.058	0.29034	
HarmonyI-V-VI	9.607e-01	2.144e-01	1.600e+02	4.481	1.41e-05	***
HarmonyIV-I-V	-9.656e-02	1.594e-01	1.537e+03	-0.606	0.54466	
Voicepar3rd	-2.700e-01	1.522e-01	2.142e+03	-1.775	0.07607	.
Voicepar5th	-2.367e-01	1.525e-01	2.141e+03	-1.552	0.12084	
KnowRob	-1.040e-01	1.084e-01	6.955e+01	-0.960	0.34062	
KnowAxis	1.082e-01	9.930e-02	6.939e+01	1.090	0.27944	
HarmonyI-V-IV:Voicepar3rd	-3.649e-01	2.156e-01	2.144e+03	-1.692	0.09079	.
HarmonyI-V-VI:Voicepar3rd	-6.805e-01	2.159e-01	2.143e+03	-3.152	0.00164	**
HarmonyIV-I-V:Voicepar3rd	4.862e-01	2.155e-01	2.142e+03	2.256	0.02419	*
HarmonyI-V-IV:Voicepar5th	-1.893e-01	2.161e-01	2.142e+03	-0.876	0.38108	
HarmonyI-V-VI:Voicepar5th	-4.258e-01	2.158e-01	2.143e+03	-1.973	0.04859	*
HarmonyIV-I-V:Voicepar5th	7.636e-02	2.154e-01	2.142e+03	0.354	0.72302	
HarmonyI-V-IV:KnowRob	2.385e-02	5.855e-02	1.346e+02	0.407	0.68445	
HarmonyI-V-VI:KnowRob	4.396e-01	9.924e-02	7.162e+01	4.430	3.32e-05	***
HarmonyIV-I-V:KnowRob	2.545e-03	5.689e-02	5.609e+02	0.045	0.96434	
HarmonyI-V-IV:KnowAxis	-4.137e-02	5.327e-02	1.313e+02	-0.777	0.43874	
HarmonyI-V-VI:KnowAxis	-1.692e-01	9.070e-02	7.084e+01	-1.865	0.06625	.
HarmonyIV-I-V:KnowAxis	-4.850e-02	5.189e-02	5.531e+02	-0.935	0.35038	

Figure 2: Fixed Effect of Linear Mixed Effects Model I

	Estimate	Std. Error	df	t value	Pr(>|t|)	
Instrumentguitar	4.241e+00	2.513e-01	9.312e+01	16.878	< 2e-16	***
Instrumentpiano	5.610e+00	2.436e-01	1.033e+02	23.031	< 2e-16	***
Instrumentstring	7.367e+00	2.329e-01	1.174e+02	31.630	< 2e-16	***

Figure 3: Fixed Effect of Linear Mixed Effects Model I (No-Intercept : Instrument)

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)	
HarmonyI-IV-V	4.241e+00	2.513e-01	9.312e+01	16.878	< 2e-16	***
HarmonyI-V-IV	4.412e+00	2.626e-01	1.019e+02	16.799	< 2e-16	***
HarmonyI-V-VI	5.201e+00	2.523e-01	1.193e+02	20.617	< 2e-16	***
HarmonyIV-I-V	4.144e+00	2.549e-01	9.320e+01	16.258	< 2e-16	***

Figure 4: Fixed Effect of Linear Mixed Effects Model I (No-Intercept : Harmony)

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)	
Voicecontrary	4.241e+00	2.513e-01	9.312e+01	16.878	< 2e-16	***
Voicepar3rd	3.970e+00	2.514e-01	9.329e+01	15.796	< 2e-16	***
Voicepar5th	4.004e+00	2.514e-01	9.337e+01	15.925	< 2e-16	***

Figure 5: Fixed Effect of Linear Mixed Effects Model I (No-Intercept : Voice)

We examined what experimental factors, or combination of factors, have the strongest influence on classical ratings by constructing linear mixed effects model.[4] Illustrated in Figure 2, 3, and 4, the results indicate that there are statistically significant association between all levels of Instrument, Harmony, Voice and classical ratings. We see from the coefficients of Instrument, Harmony, and Voice, shown in Figure 2, that Instrument has the largest average effect size on classical ratings. Among instruments, string has the largest positive effect on classical ratings in expectation.

From Figure 4, comparing coefficients of each level in Harmony, we see that I-V-vi has the strongest association with classical ratings; result in Figure 2 also indicates that there is a statistically significant difference in between the average effect size of I-V-vi and the I-IV-V, the baseline of Harmony. Also, looking at the interaction term of Harmony with KnowRob and KnowAxis, we find statistically significant evidence that association between I-V-Vi and classical ratings differs in positive direction by KnowRob. However, we do not find any statistical evidence that association between Harmony and classical ratings differs in terms of KnowAxis.

From Figure 5, we see that contrary motion has the largest coefficient size among other levels. However, result in Figure 2 shows that there is no statistically significant difference in between the average effect of each level of Voice on classical ratings, which suggests that there is no statistical evidence of contrary motion having the strongest association with classical ratings.

[4] Appx : 6.2-6.4

3.2. Differences in the way musicians and non-musicians identify classical music

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	5.14423	0.55800	89.64203	9.219	1.25e-14	***
Instrumentpiano	1.36945	0.17106	70.11290	8.006	1.76e-11	***
Instrumentstring	3.12675	0.23656	69.98006	13.218	< 2e-16	***
HarmonyI-V-IV	0.13653	0.17176	562.19638	0.795	0.427006	
HarmonyI-V-VI	0.64958	0.24045	131.13126	2.702	0.007815	**
HarmonyIV-I-V	-0.16052	0.16909	1050.39362	-0.949	0.342679	
Voicepar3rd	-0.26986	0.15212	2140.09799	-1.774	0.076218	.
Voicepar5th	-0.23621	0.15248	2139.81161	-1.549	0.121496	
Composing	0.14581	0.12503	74.22763	1.166	0.247253	
X16.minus.17	-0.08338	0.04452	70.65880	-1.873	0.065222	.
Selfdeclare	-1.60361	0.55410	81.24749	-2.894	0.004881	**
ClsListen	-0.07401	0.10839	70.53055	-0.683	0.496979	
APTheory	1.08798	0.34131	70.34513	3.188	0.002141	**
ConsInstr	0.03697	0.09826	70.67747	0.376	0.707840	
ConsNotes	-0.18570	0.08199	71.90360	-2.265	0.026527	*
X1990s2000s	-0.14077	0.08687	70.35511	-1.620	0.109626	
GuitarPlay	1.19915	0.45595	70.61772	2.630	0.010474	*
HarmonyI-V-IV:Voicepar3rd	-0.36554	0.21560	2142.08506	-1.695	0.090131	.
HarmonyI-V-VI:Voicepar3rd	-0.68024	0.21583	2141.09405	-3.152	0.001646	**
HarmonyIV-I-V:Voicepar3rd	0.48627	0.21550	2140.58009	2.256	0.024143	*
HarmonyI-V-IV:Voicepar5th	-0.18898	0.21602	2140.27775	-0.875	0.381769	
HarmonyI-V-VI:Voicepar5th	-0.42309	0.21576	2141.63048	-1.961	0.050016	.
HarmonyIV-I-V:Voicepar5th	0.07591	0.21538	2140.53107	0.352	0.724553	
HarmonyI-V-IV:Selfdeclare	0.04655	0.18709	148.24311	0.249	0.803855	
HarmonyI-V-VI:Selfdeclare	1.21982	0.32535	70.76967	3.749	0.000359	***
HarmonyIV-I-V:Selfdeclare	0.06749	0.18227	343.04926	0.370	0.711416	
Selfdeclare:ClsListen	0.58512	0.18203	70.32394	3.214	0.001974	**
Selfdeclare:GuitarPlay	-1.41997	0.46414	70.40031	-3.059	0.003137	**

Figure 6: Fixed Effect in Linear Mixed Effects Model II

We examined if there are any differences in the way that musicians and non-musicians identify classical music, by constructing another linear mixed effects model.[5] From Figure 6, we see that Selfdeclare has statistically significant negative association with classical ratings; a unit increase in Selfdeclare is associated with decrease in classical ratings by approximately 1.6 unit.

We also see that interaction terms of Selfdeclare with ClsListen and GuitarPlay have statistically significant association with classical ratings; former term has positive association, while latter term has negative association with classical ratings. Furthermore, we find that there is a statistically significant

[5] Appx : 6.5-6.6

difference in the average effect size on classical ratings between interaction term of Selfdeclare and Harmony I-V-vi, and its baseline, Selfdeclare and I-IV-V.

3.3. Difference in factors that drive classical, vs. popular, ratings

```
Fixed effects:
                      Estimate Std. Error        df t value Pr(>|t|)    
(Intercept)          5.911e+00  6.361e-01 7.784e+01   9.292 2.99e-14 ***
Instrumentpiano     -9.487e-01  1.622e-01 7.006e+01  -5.849 1.44e-07 ***
Instrumentstring    -2.604e+00  2.321e-01 7.003e+01 -11.222  < 2e-16 ***
HarmonyI-V-IV       -2.847e-02  9.864e-02 1.594e+02  -0.289 0.773249    
HarmonyI-V-VI       -2.733e-01  1.447e-01 7.247e+01  -1.889 0.062854 .  
HarmonyIV-I-V       -1.859e-01  1.073e-01 9.094e+01  -1.732 0.086622 .  
Voicepar3rd          1.654e-01  7.761e-02 2.144e+03   2.132 0.033159 *  
Voicepar5th          1.600e-01  7.756e-02 2.143e+03   2.063 0.039195 *  
Composing           -3.648e-01  1.720e-01 7.148e+01  -2.121 0.037382 *  
X16.minus.17         1.503e-01  4.122e-02 7.113e+01   3.646 0.000503 ***
Selfdeclare         -7.723e-01  3.521e-01 7.236e+01  -2.193 0.031497 *  
PachListen          -2.572e-01  1.093e-01 7.026e+01  -2.352 0.021484 *  
APTheory            -1.799e-02  3.104e-01 7.043e+01  -0.058 0.953932    
ConsInstr            1.324e-01  9.388e-02 7.060e+01   1.410 0.162811    
ConsNotes            9.064e-02  7.921e-02 7.235e+01   1.144 0.256284    
X1990s2000s          2.377e-01  7.904e-02 7.049e+01   3.007 0.003656 ** 
GuitarPlay          -4.421e-01  1.284e-01 7.278e+01  -3.443 0.000959 ***
KnowRob              1.954e-01  8.552e-02 7.036e+01   2.284 0.025365 *  
KnowAxis             7.721e-03  6.963e-02 7.023e+01   0.111 0.912019    
Composing:Selfdeclare 1.071e+00  2.310e-01 7.420e+01   4.637 1.48e-05 ***
```

Figure 7: Fixed Effect in Linear Mixed Effects Model III

We examined if there are any differences in the things that drive classical or popular ratings by constructing and comparing two similar linear mixed effects models that explain classical ratings or popular ratings well.[6] Comparing Figure 6 and 7, we found many differences in directions of average effects of same variables in both models. For instance, we see that guitar and all levels of voice leading have relatively the strongest average effect on popular ratings. Also, we find that GuitarPlay, which is statistically significant in both models, have different average effect sizes and directions.

Regarding variables that are not used in both models, we find that KnowRob and PachListen have statistically significant association with popular ratings. In terms of Selfdeclare and its interaction term, we observe that its interaction with ClsLiten and GuitarPlay have statistically significant association with classical ratings, while its interaction with Composing have significant association with popular ratings.

[6] Appx : 6.7-6.8

4. Discussion

We examined the given data to explore and examine the relationship between the characteristics of music and ratings that 60 participants had given on two different scales, Classical or Popular. We first found that three main experimental factors, instrument, harmony, and voice play significant roles when listeners rate the music. As discussed in the Result section, we found that instrument exerts the strongest influence, among the three experimental factors, on classical ratings. Such result is somewhat expected as many classical music consists of instrument sounds only.

We also observed that there is a statistically significant evidence that music played by string quartet or piano is likely to receive higher score as a classical sounding than music played by guitar. As many classical music indeed tends to have string or piano sounds, it is not surprising that many participants tend to think music played with such instruments as a classic music.

Among levels of harmonic motion, we found that I-V-vi itself not only has a significant association with classical ratings, but it also has significant association in terms of participants' familiarity with Rob Paravonian's Pachelbel Rant. Such result suggests that I-V-vi may have been used frequently in famous classical music or in Pachelbel Rant. Among levels of voice leading, we did not find any statistical evidence that contrary motion has notable association with classical ratings. Such result may be due to the fact that voice itself is not often used in classical music.

We investigated the differences in the way musicians and non-musicians identify classical music. We found that musicians tend to give lower scores on classical ratings. But we also see that musicians who frequently listens to classical music tend to give relatively higher scores, while those who play guitar tend to much lower scores. Such fact suggests that for musicians, those who may have more musical knowledge and experience than non-musicians do, their musical preference or experience affect their decisions on how the music sounds. Furthermore, we find that musicians tend to give higher scores on classical ratings when the music has harmonic motion of I-V-vi. This again strongly suggests that there is a strong association with I-V-vi and classical soundings.

Lastly, we examined the differences in factors that may affect listeners' tendency to give scores on classical or popular ratings. Among many differences, one noticeable result was the difference in the influence of instrument and voice on each rating. Unlike for classical ratings, music with guitar sounds or voice tends to receive higher scores on popular ratings. It was also expected as they are more frequently used in popular music than in classical music.

In terms of other factors, we observed that listeners who can play guitar tends to give higher scores on classical ratings and lower scores on popular ratings. It was surprising in that we intuitively thought that listeners' experiences with guitar would lead them to give higher score on popular ratings and low score on classical ratings. One possible reasoning is that since they know "well" about guitar, they eventually know more about popular music. Thus, they either have better abilities to distinguish classical sounding from popular music or tend to be more analytical with popular soundings, which eventually led them to carefully evaluate the music piece and give moderate rates.

Another noteworthy result was about factors that drive classical or popular ratings from musicians. We first see that musicians tend to give lower scores on both classical and popular ratings, which can also be explained by the reasoning mentioned right above. For popular ratings, we found that musicians who have experience in composing tend to give relatively higher scores. Compared to the result discussed in fourth paragraph, we see different aspects of musical experience have different influence on musicians.

For future research, we believe that data with more variables that measures individual musical experience or preference should be collected, as we have found some variability in the influence of musical factors over each subject. Data with more variables would give more guidance on how musical experience and preference of listeners shape how they perceive the sounding.

5. References

Jimenez, Ivan, Rossi, Vincent. (2015) *The Influence of Timbre, Harmony, and Voice-leading on Listeners' Distinction between Popular and Classical Music* Pittsburgh : University of Pittsburgh

6. Appendix

6.1. Exploratory Data Analysis

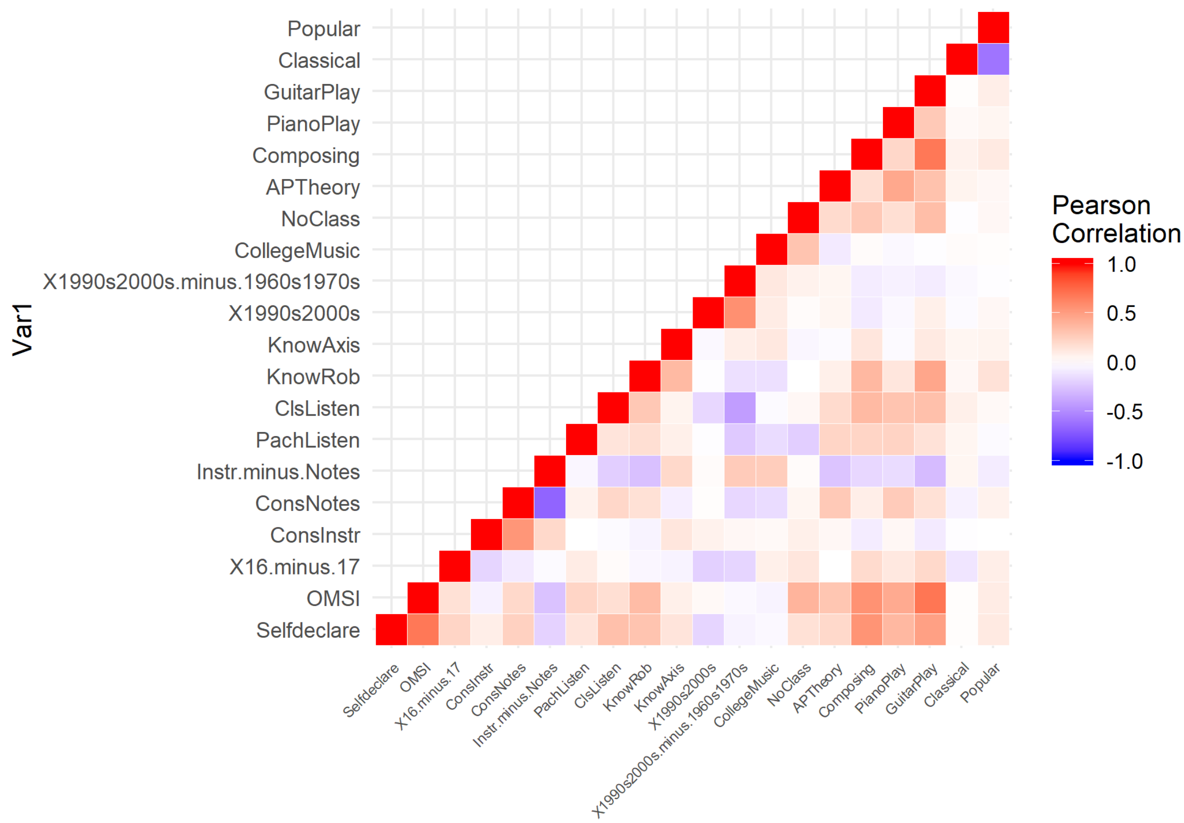

Figure 8: Correlation Heatmap of Numerical Variables

6.2. How Linear Mixed Effects Model I was Constructed

We examined three main experimental factors using both conventional linear models and analysis of variance test. Considering all possible combination of variables, using stepwise backward AIC variable selection method, the model with Instrument, Harmony, Voice, and interaction of Harmony and Voice appeared to explain the Classical well, as shown by the ANOVA test result.

```
Analysis of Variance Table

Model 1: Classical ~ Harmony + Instrument + Voice
Model 2: Classical ~ Instrument + Harmony * Voice
  Res.Df   RSS Df Sum of Sq      F  Pr(>F)
1   2485 13108
2   2479 13026  6     81.28 2.5781 0.01715 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 9: ANOVA Test Result 1

We also used ANOVA test to see whether the random intercept is needed in the model. The test result implies that there is a statistically significant evidence that the model with random intercept fits better to the data.

```
Models:
lm.1: Classical ~ Instrument + Harmony * Voice
lmer.1: Classical ~ Instrument + Harmony * Voice + (1 | Subject)
       Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lm.1   15 11227 11314 -5598.5    11197
lmer.1 16 10458 10551 -5213.1    10426 770.84      1  < 2.2e-16 ***
```

Figure 10: ANOVA Test Result 2

We tried to find the best combination of random effect terms by utilizing lmerTest package and performing backward elimination of random effects. The best model we found was the same model with random effects of (Instrument—Subject) and (Harmony—Subject), but not (1—Subject).

As our research question involved possible interaction between KnowRob,KnowAxis, and Harmony, we included interaction terms. The formula for final model is presented below. Summary of the full model is also shown below.

```
Classical ~ Instrument + Harmony * Voice + Harmony * KnowRob +
    Harmony * KnowAxis + (Instrument | Subject) + (Harmony |
    Subject)
```

Figure 11: Linear Mixed Effects Model I Formula(R)

```
     AIC      BIC   logLik deviance df.resid 
  9933.1  10160.1  -4927.5   9855.1     2454 

Scaled residuals: 
    Min      1Q  Median      3Q     Max 
-4.7386 -0.5764  0.0210  0.5606  6.1485 

Random effects:
 Groups    Name             Variance Std.Dev. Corr       
 Subject   (Intercept)      1.216131 1.10278             
           Instrumentpiano  1.632564 1.27772  -0.68      
           Instrumentstring 3.512075 1.87405  -1.00  0.66
 Subject.1 (Intercept)      1.700224 1.30393             
           HarmonyI-V-IV    0.033365 0.18266   0.58      
           HarmonyI-V-VI    1.122985 1.05971  -0.40  0.12
           HarmonyIV-I-V    0.005219 0.07224   0.51 -0.12  0.15
 Residual                   2.417291 1.55476             
Number of obs: 2493, groups:  Subject, 70

Fixed effects:
                            Estimate Std. Error         df t value Pr(>|t|)    
(Intercept)                4.241e+00  2.513e-01  9.312e+01  16.878  < 2e-16 ***
Instrumentpiano            1.370e+00  1.710e-01  7.012e+01   8.008 1.75e-11 ***
Instrumentstring           3.127e+00  2.366e-01  6.998e+01  13.216  < 2e-16 ***
HarmonyI-V-IV              1.710e-01  1.616e-01  6.614e+02   1.058  0.29034    
HarmonyI-V-VI              9.607e-01  2.144e-01  1.600e+02   4.481 1.41e-05 ***
HarmonyIV-I-V             -9.656e-02  1.594e-01  1.537e+03  -0.606  0.54466    
Voicepar3rd               -2.700e-01  1.522e-01  2.142e+03  -1.775  0.07607 .  
Voicepar5th               -2.367e-01  1.525e-01  2.141e+03  -1.552  0.12084    
KnowRob                   -1.040e-01  1.084e-01  6.955e+01  -0.960  0.34062    
KnowAxis                   1.082e-01  9.930e-02  6.939e+01   1.090  0.27944    
HarmonyI-V-IV:Voicepar3rd -3.649e-01  2.156e-01  2.144e+03  -1.692  0.09079 .  
HarmonyI-V-VI:Voicepar3rd -6.805e-01  2.159e-01  2.143e+03  -3.152  0.00164 ** 
HarmonyIV-I-V:Voicepar3rd  4.862e-01  2.155e-01  2.142e+03   2.256  0.02419 *  
HarmonyI-V-IV:Voicepar5th -1.893e-01  2.161e-01  2.142e+03  -0.876  0.38108    
HarmonyI-V-VI:Voicepar5th -4.258e-01  2.158e-01  2.143e+03  -1.973  0.04859 *  
HarmonyIV-I-V:Voicepar5th  7.636e-02  2.154e-01  2.142e+03   0.354  0.72302    
HarmonyI-V-IV:KnowRob      2.385e-02  5.855e-02  1.346e+02   0.407  0.68445    
HarmonyI-V-VI:KnowRob      4.396e-01  9.924e-02  7.162e+01   4.430 3.32e-05 ***
HarmonyIV-I-V:KnowRob      2.545e-03  5.689e-02  5.609e+02   0.045  0.96434    
HarmonyI-V-IV:KnowAxis    -4.137e-02  5.327e-02  1.313e+02  -0.777  0.43874    
HarmonyI-V-VI:KnowAxis    -1.692e-01  9.070e-02  7.084e+01  -1.865  0.06625 .  
HarmonyIV-I-V:KnowAxis    -4.850e-02  5.189e-02  5.531e+02  -0.935  0.35038    
```

Figure 12: Linear Mixed Effects Model I

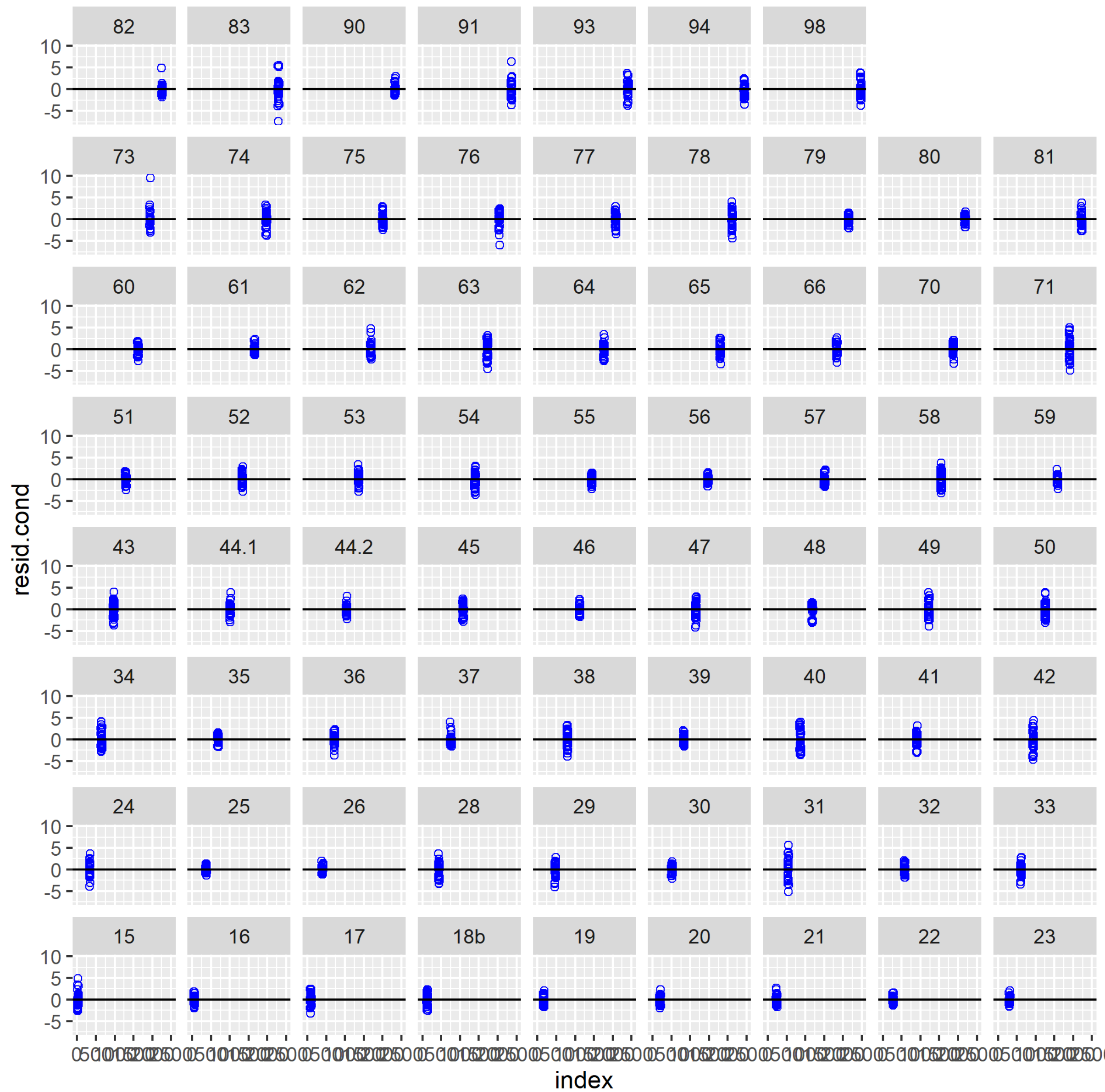

Figure 13: Conditional Residuals Plot - Model I

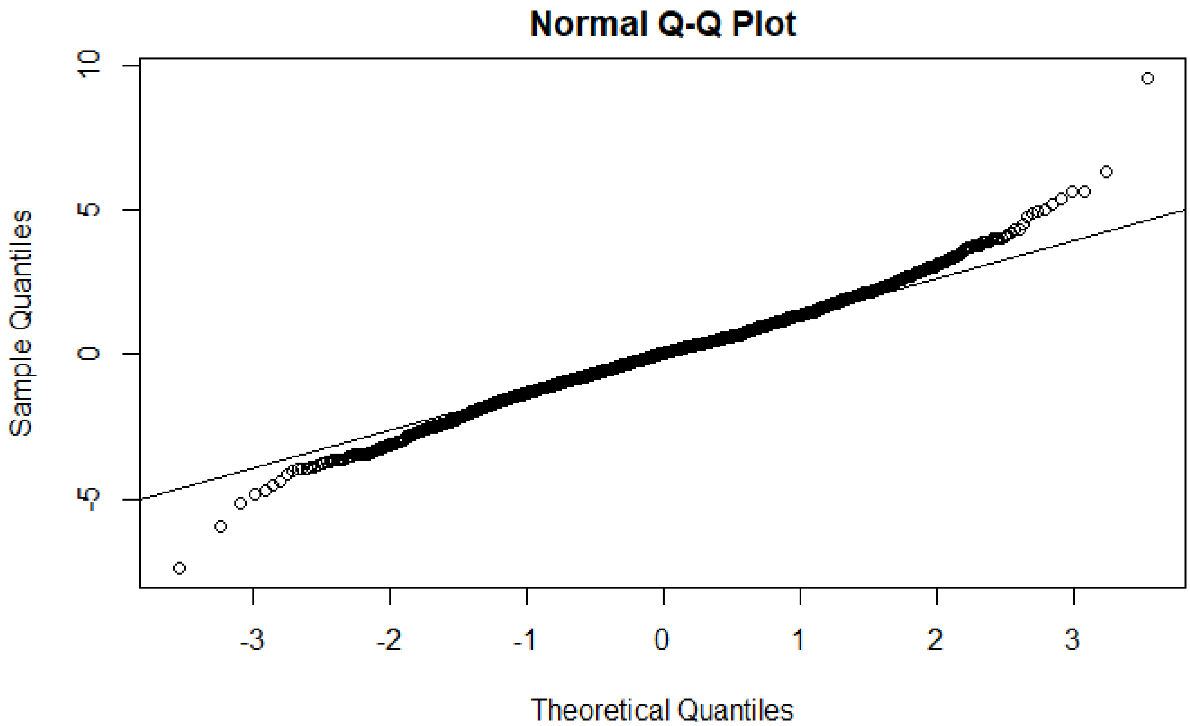

Figure 14: Normal Q-Q Plot of Conditional Residuals - Model I

We see that conditional residuals are scattered randomly about zero. Also, we do not see any serious normality issue in Normal Q-Q plot.

6.4. Focusing on Fixed Effects

Investigating the influence of all levels of Instrument, Harmony, and Voice, we focused only on interpreting fixed effects. For this study, variance components were used mainly to construct linear models that better explain classical or popular ratings, by taking variability in the influence of instruments and harmonic motion over each subject into account. However, as we were focused on finding "general" trend, we did not feel the necessity to look closely into random effects. Also, in terms of variances and standard deviations shown in the results, we did not find any abnormal trend that requires interpretation. It is very obvious that there is some variability among each subject, as each subject has different "tendency" to perceive the sound. Considering such fact and looking at the results, we did not find any particular factor that has unusual variability.

6.5. How Linear Mixed Effects Model II was Constructed

Basaed on the linear mixed effects model I, for fixed effects, we first selected variables that we thought could help predict classical ratings well. Running stepwise AIC variable selection in backward direction, we found that all the variables that we selected were suitable for model, except the KnowRob, NoClass, KnowAxis,Selfdeclare, Instr.minus.Notes, and PianoPlay. Since Selfdeclare is necessary for our research, we did not remove the variable. For rest of un-selected variables, we thought that remaining variables still gave us good interpretation in terms of music, so we excluded them from the model.

We then added random effects of three experimental factors to the model with fixed-effect-only model. Utilizing random effect variable selection in lmerTest, we found that random effects of Instrument and Harmony lead to an improvement; ANOVA test below also proves the point.

```
Models:
lm.2: Classical ~ Instrument + Harmony * Voice + Composing + X16.minus.17 +
lm.2:     Selfdeclare + ClsListen + APTheory + ConsInstr + ConsNotes +
lm.2:     X1990s2000s + GuitarPlay
lmer.22: Classical ~ Instrument + Harmony * Voice + Composing + X16.minus.17 +
lmer.22:     Selfdeclare + ClsListen + APTheory + ConsInstr + ConsNotes +
lmer.22:     X1990s2000s + GuitarPlay + (Harmony | Subject) + (Instrument |
lmer.22:     Subject)
        Df     AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lm.2    24 11103.9 11244 -5528.0  11055.9
lmer.22 40  9940.4 10173 -4930.2   9860.4 1195.5     16  < 2.2e-16 ***
```

Figure 15: ANOVA Test Result 3

For research purpose, we added interaction of Selfdeclare and all variables in the model. Then, we ran variable selection method utilizing lmerTest (can be used to test either fixed-effect or random-effect) on both random-effects and fixed-effects. For fixed effects that were not interaction terms and were not selected in variable-selection method, we did not remove them from the model. For interaction terms, we found that selected ones were Harmony:Selfdeclare, ClsListen:Selfdeclare, and GuitarPlay:Selfdeclare. We decided to keep them in the model, and remove interaction terms that were not selected.

```
Classical ~ Instrument + Harmony * Voice + Composing + X16.minus.17 +
    Selfdeclare + ClsListen + APTheory + ConsInstr + ConsNotes +
    X1990s2000s + GuitarPlay + Harmony:Selfdeclare + ClsListen:Selfdeclare +
    GuitarPlay:Selfdeclare + (Harmony | Subject) + (Instrument |
    Subject)
```

Figure 16: Linear Mixed Effects Model II Formula(R)

```
Random effects:
 Groups    Name              Variance Std.Dev. Corr
 Subject   (Intercept)       1.26398  1.1243
           HarmonyI-V-IV     0.04454  0.2110    0.57
           HarmonyI-V-VI     1.23395  1.1108   -0.55  0.18
           HarmonyIV-I-V     0.01573  0.1254    0.69  0.34  0.00
 Subject.1 (Intercept)       1.30664  1.1431
           Instrumentpiano   1.63337  1.2780   -0.61
           Instrumentstring  3.51142  1.8739   -1.00  0.66
 Residual                    2.41644  1.5545
Number of obs: 2493, groups:  Subject, 70

Fixed effects:
                             Estimate Std. Error         df t value Pr(>|t|)
(Intercept)                   5.14423    0.55800   89.64203   9.219 1.25e-14 ***
Instrumentpiano               1.36945    0.17106   70.11290   8.006 1.76e-11 ***
Instrumentstring              3.12675    0.23656   69.98006  13.218  < 2e-16 ***
HarmonyI-V-IV                 0.13653    0.17176  562.19638   0.795 0.427006
HarmonyI-V-VI                 0.64958    0.24045  131.13126   2.702 0.007815 **
HarmonyIV-I-V                -0.16052    0.16909 1050.39362  -0.949 0.342679
Voicepar3rd                  -0.26986    0.15212 2140.09799  -1.774 0.076218 .
Voicepar5th                  -0.23621    0.15248 2139.81161  -1.549 0.121496
Composing                     0.14581    0.12503   74.22763   1.166 0.247253
X16.minus.17                 -0.08338    0.04452   70.65880  -1.873 0.065222 .
Selfdeclare1                 -1.60361    0.55410   81.24749  -2.894 0.004881 **
ClsListen                    -0.07401    0.10839   70.53055  -0.683 0.496979
APTheory                      1.08798    0.34131   70.34513   3.188 0.002141 **
ConsInstr                     0.03697    0.09826   70.67747   0.376 0.707840
ConsNotes                    -0.18570    0.08199   71.90360  -2.265 0.026527 *
X1990s2000s                  -0.14077    0.08687   70.35511  -1.620 0.109626
GuitarPlay                    1.19915    0.45595   70.61772   2.630 0.010474 *
HarmonyI-V-IV:Voicepar3rd    -0.36554    0.21560 2142.08506  -1.695 0.090131 .
HarmonyI-V-VI:Voicepar3rd    -0.68024    0.21583 2141.09405  -3.152 0.001646 **
HarmonyIV-I-V:Voicepar3rd     0.48627    0.21550 2140.58009   2.256 0.024143 *
HarmonyI-V-IV:Voicepar5th    -0.18898    0.21602 2140.27775  -0.875 0.381769
HarmonyI-V-VI:Voicepar5th    -0.42309    0.21576 2141.63048  -1.961 0.050016 .
HarmonyIV-I-V:Voicepar5th     0.07591    0.21538 2140.53107   0.352 0.724553
HarmonyI-V-IV:Selfdeclare1    0.04655    0.18709  148.24311   0.249 0.803855
HarmonyI-V-VI:Selfdeclare1    1.21982    0.32535   70.76967   3.749 0.000359 ***
HarmonyIV-I-V:Selfdeclare1    0.06749    0.18227  343.04926   0.370 0.711416
Selfdeclare1:ClsListen        0.58512    0.18203   70.32394   3.214 0.001974 **
Selfdeclare1:GuitarPlay      -1.41997    0.46414   70.40031  -3.059 0.003137 **
---
```

Figure 17: Linear Mixed Effects Model II

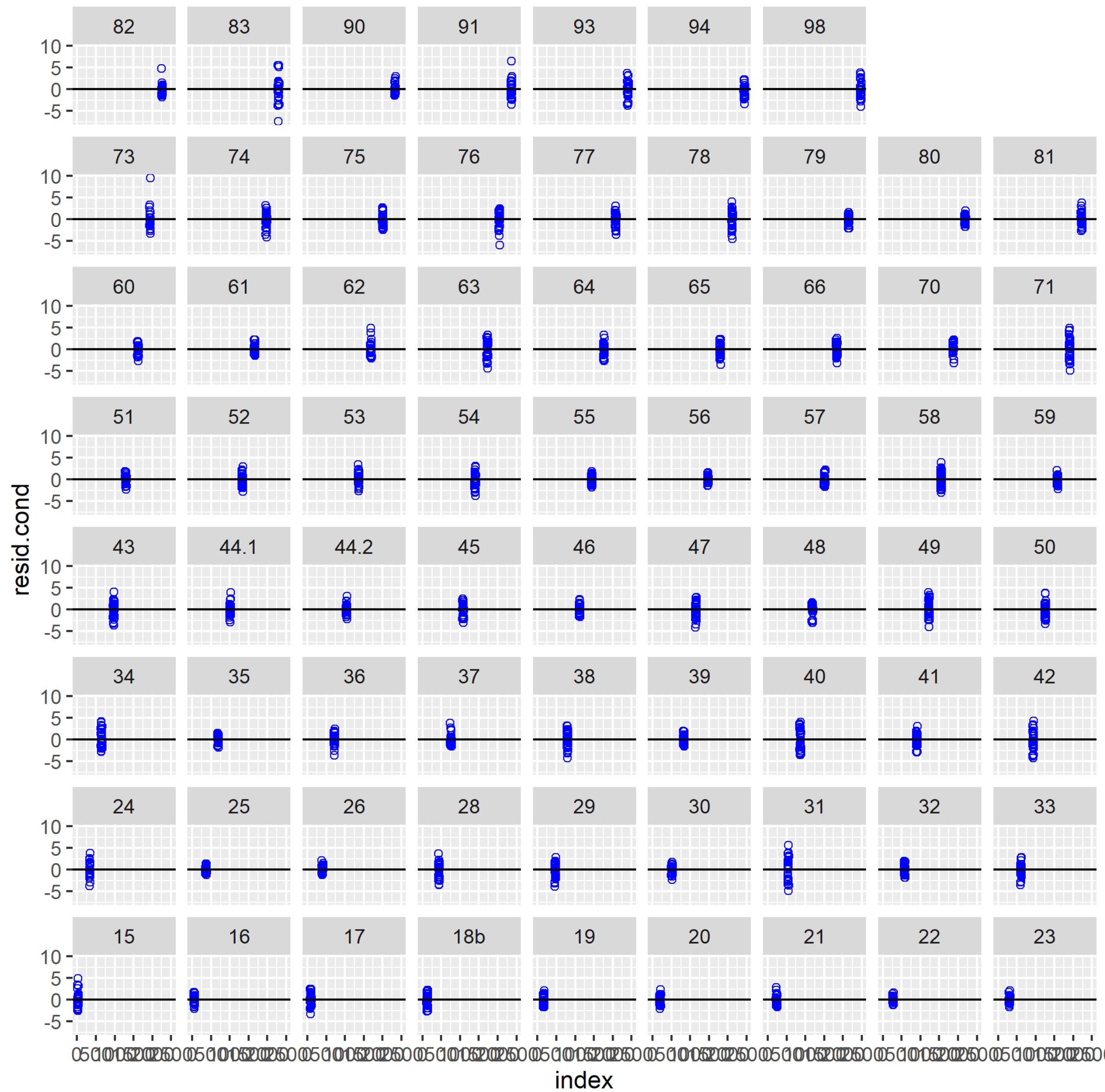

Figure 18: Conditional Residuals Plot - Model II

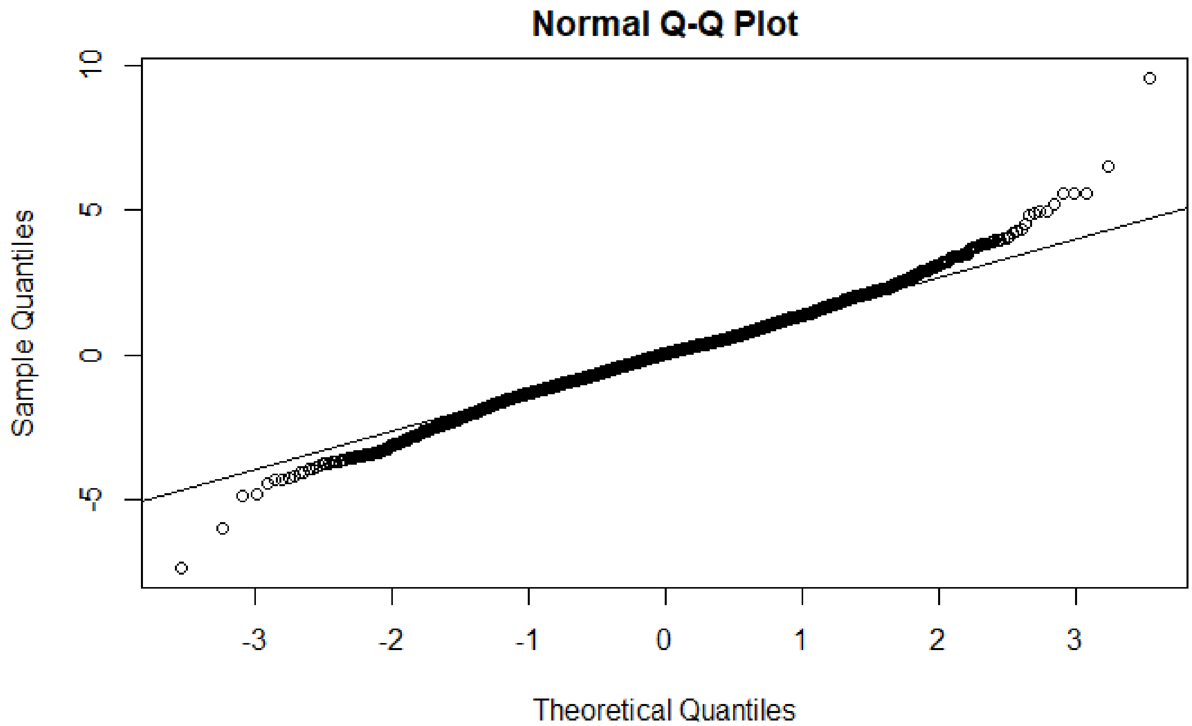

Figure 19: Normal Q-Q Plot of Conditional Residuals - Model II

We see that conditional residuals are scattered randomly about zero. Also, we do not see any serious normality issue in Normal Q-Q plot, except both tails have few outliers.

6.7. How Linear Mixed Effects Model III was Constructed

We examined three main experimental factors using both conventional linear models and analysis of variance test. Considering all possible combination of variables, the model with Instrument, Harmony, and Voice appeared to explain the Popular well. So we did not include any interaction terms.

Based on the findings, for fixed effects, we first selected variables that we thought could help predict Popular ratings well. Then, utilizing package olssr, which lets us run stepwise AIC variable selection in backward directions, we found that all the variables that we selected were suitable for model, except the NoClass, Instr.minus.Notes, OMSI, Voice, and PianoPlay. Since

Voice is necessary for our research, we did not remove the variable. For rest of un-selected variables, We thought that remaining variables still gave us good interpretation in terms of music, so we excluded them from the model.

We wanted to see if the interaction of Selfdeclare and other predictor variables play significant role in terms of popular ratings. Similar to the approach discussed in Appx 6.5, we found that interaction of Selfdeclare and Composing improves the model and has statistically significant association with Popular.

We focused on random effects of three main factors, Voice, Harmony, and Instrument. Thus we added random effect of those factors to the fixed-effect-only model and performed random effect term selection based on likelihood ratio test, utilizing package lmerTest. We found that random effect of Harmony and Instrument improved the model. ANOVA test result below shows that adding a random effect term of Instrument and Harmony indeed leads to an improvement in the model. We did not include random effect of other factors as we did not feel the necessity of them based on the goal of the study and the given data; also, automatic method suggests the same.

```
Models:
lm.333: Popular ~ Instrument + Harmony + Voice + Composing + X16.minus.17 +
lm.333:     Selfdeclare + PachListen + APTheory + ConsInstr + ConsNotes +
lm.333:     X1990s2000s + GuitarPlay + KnowRob + KnowAxis + Selfdeclare:Composing
lmer.33: Popular ~ Instrument + Harmony + Voice + Composing + X16.minus.17 +
lmer.33:     Selfdeclare + PachListen + APTheory + ConsInstr + ConsNotes +
lmer.33:     X1990s2000s + GuitarPlay + KnowRob + KnowAxis + Selfdeclare:Composing +
lmer.33:     (Harmony | Subject) + (Instrument | Subject)
        Df     AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lm.333  21 10871.9 10994 -5415.0  10829.9
lmer.33 37  9991.2 10206 -4958.6   9917.2 912.78     16  < 2.2e-16 ***
```

Figure 20: ANOVA Test Result 4

```
Popular ~ Instrument + Harmony + Voice + Composing + X16.minus.17 +
    Selfdeclare + PachListen + APTheory + ConsInstr + ConsNotes +
    X1990s2000s + GuitarPlay + KnowRob + KnowAxis + Selfdeclare:Composing +
    (Harmony | Subject) + (Instrument | Subject)
```

Figure 21: Linear Mixed Effects Model III Formula(R)

```
Random effects:
 Groups    Name              Variance Std.Dev. Corr
 Subject   (Intercept)       1.1439   1.0695
           HarmonyI-V-IV     0.1194   0.3455    0.85
           HarmonyI-V-VI     0.9025   0.9500   -0.24 -0.41
           HarmonyIV-I-V     0.2452   0.4952   -0.54 -0.76 -0.19
 Subject.1 (Intercept)       0.6833   0.8266
           Instrumentpiano   1.4131   1.1887   -0.85
           Instrumentstring  3.3505   1.8304   -0.98  0.72
 Residual                    2.4993   1.5809
Number of obs: 2493, groups:  Subject, 70

Fixed effects:
                        Estimate Std. Error         df t value Pr(>|t|)
(Intercept)            5.911e+00  6.361e-01  7.784e+01   9.292 2.99e-14 ***
Instrumentpiano       -9.487e-01  1.622e-01  7.006e+01  -5.849 1.44e-07 ***
Instrumentstring      -2.604e+00  2.321e-01  7.003e+01 -11.222  < 2e-16 ***
HarmonyI-V-IV         -2.847e-02  9.864e-02  1.594e+02  -0.289 0.773249
HarmonyI-V-VI         -2.733e-01  1.447e-01  7.247e+01  -1.889 0.062854 .
HarmonyIV-I-V         -1.859e-01  1.073e-01  9.094e+01  -1.732 0.086622 .
Voicepar3rd            1.654e-01  7.761e-02  2.144e+03   2.132 0.033159 *
Voicepar5th            1.600e-01  7.756e-02  2.143e+03   2.063 0.039195 *
Composing             -3.648e-01  1.720e-01  7.148e+01  -2.121 0.037382 *
X16.minus.17           1.503e-01  4.122e-02  7.113e+01   3.646 0.000503 ***
Selfdeclare1          -7.723e-01  3.521e-01  7.236e+01  -2.193 0.031497 *
PachListen            -2.572e-01  1.093e-01  7.026e+01  -2.352 0.021484 *
APTheory              -1.799e-02  3.104e-01  7.043e+01  -0.058 0.953932
ConsInstr              1.324e-01  9.388e-02  7.060e+01   1.410 0.162811
ConsNotes              9.064e-02  7.921e-02  7.235e+01   1.144 0.256284
X1990s2000s            2.377e-01  7.904e-02  7.049e+01   3.007 0.003656 **
GuitarPlay            -4.421e-01  1.284e-01  7.278e+01  -3.443 0.000959 ***
KnowRob                1.954e-01  8.552e-02  7.036e+01   2.284 0.025365 *
KnowAxis               7.721e-03  6.963e-02  7.023e+01   0.111 0.912019
Composing:Selfdeclare1 1.071e+00  2.310e-01  7.420e+01   4.637 1.48e-05 ***
```

Figure 22: Linear Mixed Effects Model III

6.8. Diagnostic Plot of Linear Mixed Effects Model III

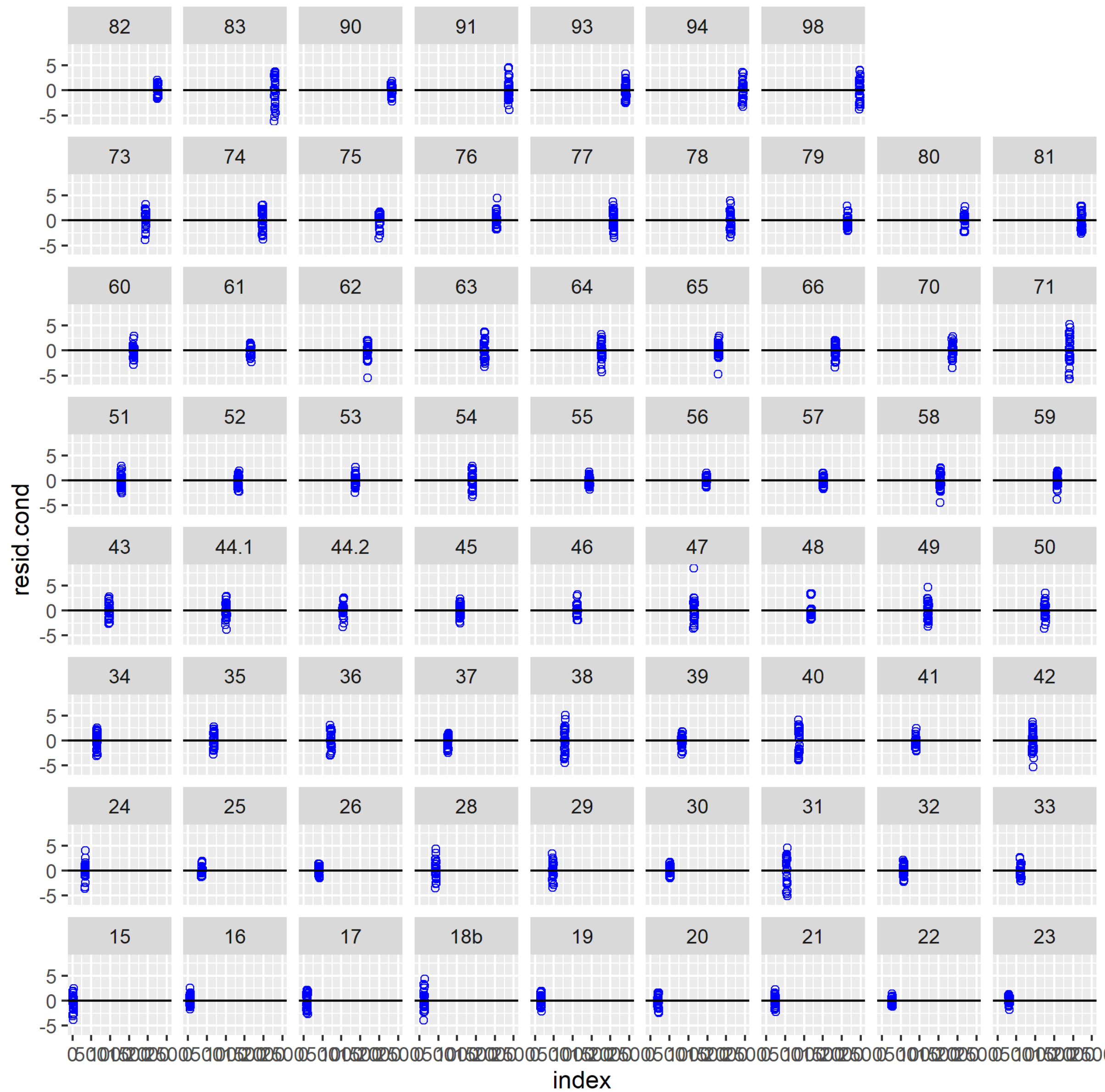

Figure 23: Conditional Residuals Plot - Model III

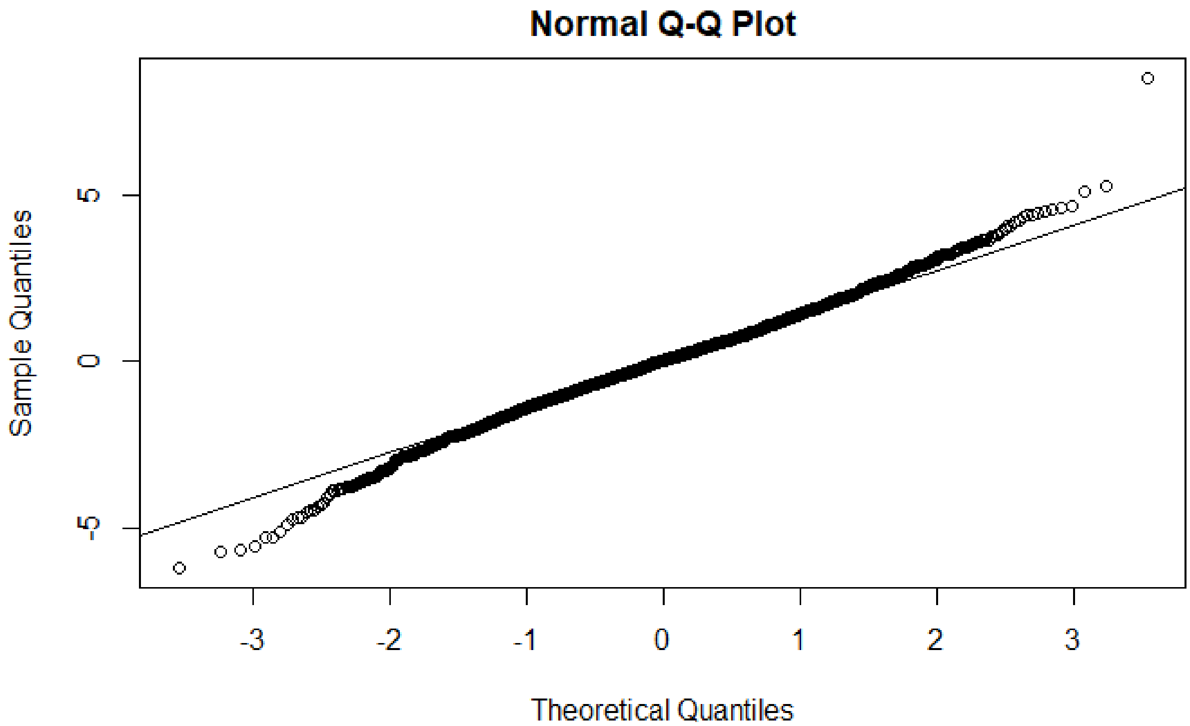

Figure 24: Normal Q-Q Plot of Conditional Residuals - Model III

We see that conditional residuals are scattered randomly about zero. Also, we do not see any serious normality issue in Normal Q-Q plot, except both tails have outliers and left tail is skewed in some extent.

6.9 Code Appendix

Justin Kim

11/26/2019

Data Cleaning/Missing Value Imputation

```
#remove columns not used in analysis
dat$X <- NULL
dat$first12 <- NULL

#These two columns contained so many NAs.Also, those variables did not seem
#necessary for modeling.
dat$X1stInstr <- NULL
dat$X2ndInstr <- NULL

#Rather than imputation, we simply remove missing response variables;
#relatively few NAs are in those variables
dat <- dat[is.na(dat$Classical) == FALSE,]
dat <- dat[is.na(dat$Popular) == FALSE,]
dat$Subject <- factor(dat$Subject)

#ConsInstr is not supposed to have decimal numbers.
#But since they do, we round up the decimals and calculate
#Instr.minus.Notes again
dat$ConsInstr <- round(dat$ConsInstr, 0)
for (i in 1:length(dat$ConsInstr)){
  if (is.na(dat$ConsNotes[i]) == FALSE){
    dat$Instr.minus.Notes[i] <- dat$ConsInstr[i] - dat$ConsNotes[i]
  }
  else{
    dat$Instr.minus.Notes[i] <- NA
  }
}

dat$ConsInstr <- factor(dat$ConsInstr)
dat$ConsNotes <- factor(dat$ConsNotes)
dat$KnowAxis <- factor(dat$KnowAxis)
dat$KnowRob <- factor(dat$KnowRob)
dat$X1990s2000s.minus.1960s1970s <- factor(dat$X1990s2000s.minus.1960s1970s)
dat$PachListen <- factor(dat$PachListen)
dat$ClsListen <- factor(dat$ClsListen)
dat$X1990s2000s <- factor(dat$X1990s2000s)
dat$CollegeMusic <- factor(dat$CollegeMusic)
dat$APTheory <- factor(dat$APTheory)
dat$Composing <- factor(dat$Composing)


set.seed(1234)
datt <- missForest(dat[,2:24], verbose = TRUE)
datt <- datt$ximp
datt <- cbind(dat$Subject, datt)
```

```r
colnames(datt)[1] <- "Subject"

rm(dat)
dat <- datt
rm(datt)
dat$Selfdeclare[dat$Selfdeclare %in% c(1,2)] <- 0
dat$Selfdeclare[dat$Selfdeclare %in% c(3,4,5,6)] <- 1

dat$Selfdeclare <- factor(dat$Selfdeclare)
```

Correlation Heatmap

```r
dat$PachListen <- unfactor(dat$PachListen)
dat$ClsListen <- unfactor(dat$ClsListen)
dat$X1990s2000s <- unfactor(dat$X1990s2000s)
dat$CollegeMusic <- unfactor(dat$CollegeMusic)
dat$APTheory <- unfactor(dat$APTheory)
dat$Composing <- unfactor(dat$Composing)
dat$ConsInstr <- unfactor(dat$ConsInstr)
dat$ConsNotes <- unfactor(dat$ConsNotes)
dat$KnowAxis <- unfactor(dat$KnowAxis)
dat$KnowRob <- unfactor(dat$KnowRob)
dat$X1990s2000s.minus.1960s1970s <- unfactor(dat$X1990s2000s.minus.1960s1970s)
dat$Selfdeclare <- unfactor(dat$Selfdeclare)

nums <- unlist(lapply(dat, is.numeric))
dat_n <- dat[,nums]
cormat <- round(cor(dat_n),2)
melted_cormat <- melt(cormat)

get_upper_tri <- function(cormat) {
  cormat[lower.tri(cormat)] <- NA
  return(cormat)
}

upper_tri <- get_upper_tri(cormat)
melted_cormat <- melt(upper_tri, na.rm = TRUE)

ggplot(data = melted_cormat, aes(Var2, Var1, fill = value)) +
  geom_tile(color = "white") +
  scale_fill_gradient2(
    low = "blue",
    high = "red",
    mid = "white",
    midpoint = 0,
    limit = c(-1, 1),
    space = "Lab",
    name = "Pearson\nCorrelation"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(
    angle = 45,
    vjust = 1,
```

```
    size = 6,
    hjust = 1
  )) +
  coord_fixed()

#ggsave("heatmap", plot=last_plot(), device = "png")

dat$Selfdeclare <- factor(dat$Selfdeclare)
```

Linear Mixed Effects Model I

```
lm.0 <- lm(Classical ~ (Harmony+Instrument+Voice)^2, data = dat)

olsrr::ols_step_backward_aic(lm.0)

lm.0 <- lm(Classical~Harmony+Instrument+Voice, data = dat)

lm.1 <- lm(Classical ~Instrument+ Harmony*Voice , data = dat)

anova(lm.0,lm.1)

summary(lm.1)
```

```
lmer.1 <- lmer(Classical~Instrument+Harmony*Voice+(1|Subject), data= dat, REML = FALSE)

anova(lmer.1,lm.1)
```

```
lmer.2 <- lmer(Classical~Instrument+Harmony*Voice+ (Instrument|Subject)+(Harmony|Subject), data= dat, RE
formula(lmer.2)
anova(lmer.1,lmer.2)
```

```
lmer.3 <- lmer(Classical~Instrument+Harmony*Voice+Harmony*KnowRob+Harmony*KnowAxis+(Instrument|Subject)+


# step_result <- step(lmer.3, reduce.fixed = TRUE, reduce.random = FALSE)
# lmer.3 <- get_model(step_result)
#
formula(lmer.3)
```

```
summary(lmer.3)

# lmer.3i <- lmer(Classical~Instrument-1+Harmony*Voice+Harmony*KnowRob+Harmony*KnowAxis+ (Instrument-1|
#
# lmer.3v <- lmer(Classical~0+Voice+Instrument+Harmony:Voice+Harmony+Harmony*KnowRob+Harmony*KnowAxis+(
#
# lmer.3h <- lmer(Classical~0+Harmony+Instrument+Harmony:Voice+Voice+Harmony*KnowRob+Harmony*KnowAxis+(
#
#
# summary(lmer.3i)
# summary(lmer.3v)
# summary(lmer.3h)
```

```
source("residual-functions.r")
resid.cond <- r.cond(lmer.3)
fit.cond <- yhat.cond(lmer.3)
```

```
attach(dat)
index <- 1:dim(dat)[1]


new.data <- data.frame(index,resid.cond,Subject)
names(new.data) <- c("index","resid.cond","Subject")
ggplot(new.data,aes(x=index,y=resid.cond)) +
  facet_wrap( ~ Subject, as.table=F) +
  geom_point(pch=1,color="Blue") +
    geom_hline(yintercept=0)

{qqnorm(resid.cond)
qqline(resid.cond)}


ggsave("rplot.png", plot=last_plot(), device = "png")
```

Linear Mixed Effects Model II

```
lm.2 <- lm(Classical~Instrument+Harmony*Voice+Composing+X16.minus.17+Selfdeclare +
             ClsListen + APTheory + ConsInstr + ConsNotes + Instr.minus.Notes + X1990s2000s +
             NoClass+ GuitarPlay + PianoPlay+KnowRob+KnowAxis, data= dat)

olsrr::ols_step_backward_aic(lm.2)

lm.2 <- lm(Classical~Instrument+Harmony*Voice+Composing+X16.minus.17+Selfdeclare
           + ClsListen + APTheory + ConsInstr + ConsNotes + X1990s2000s +GuitarPlay, data= dat)

lmer.22 <- lmer(Classical~Instrument+Harmony*Voice+Composing+X16.minus.17+
                  Selfdeclare + ClsListen + APTheory + ConsInstr + ConsNotes +
                  X1990s2000s +GuitarPlay+(Harmony|Subject)+(Instrument|Subject),
                data= dat, REML = FALSE, control = lmerControl(optimizer = 'bobyqa'))
```

```
lmer.4 <- lmer(Classical~Instrument+Harmony*Voice+Composing+X16.minus.17+
                 Selfdeclare+ClsListen + APTheory + ConsInstr + ConsNotes +
                 X1990s2000s +GuitarPlay+Instrument:Selfdeclare + Harmony:Selfdeclare+Voice:Selfdeclare
               +Composing:Selfdeclare+X16.minus.17:Selfdeclare+
                 ClsListen:Selfdeclare+APTheory:Selfdeclare+ConsInstr:Selfdeclare+ConsNotes:Selfdeclare
                 X1990s2000s:Selfdeclare+GuitarPlay:Selfdeclare+(Harmony|Subject)+(Instrument|Subject),
               data= dat, REML = FALSE, control = lmerControl(optimizer = 'bobyqa'))

# step_result <- step(lmer.3, reduce.fixed = TRUE, reduce.random = FALSE)
# lmer.3 <- get_model(step_result)
#
# formula(lmer.3)
lmer.4f <- lm(Classical~Instrument+Harmony*Voice+Composing+X16.minus.1
              +Selfdeclare+ClsListen + APTheory + ConsInstr + ConsNotes
              + X1990s2000s+GuitarPlay+Harmony:Selfdeclare+ClsListen:Selfdeclare+
                GuitarPlay:Selfdeclare, data= dat)


lmer.4 <- lmer(Classical~Instrument+Harmony*Voice+Composing+X16.minus.17+
                 Selfdeclare+ClsListen + APTheory + ConsInstr + ConsNotes + X1990s2000s+
```

```r
                 GuitarPlay+Harmony:Selfdeclare+ClsListen:Selfdeclare+GuitarPlay:Selfdeclare+
                 (Harmony|Subject)+(Instrument|Subject), data= dat, REML = FALSE, control = lmerControl

anova(lmer.4,lmer.4f)

formula(lmer.4)

summary(lmer.4)
```

```r
resid.cond <- r.cond(lmer.4)
fit.cond <- yhat.cond(lmer.4)

attach(dat)
index <- 1:dim(dat)[1]


new.data <- data.frame(index,resid.cond,Subject)
names(new.data) <- c("index","resid.cond","Subject")
ggplot(new.data,aes(x=index,y=resid.cond)) +
  facet_wrap( ~ Subject, as.table=F) +
  geom_point(pch=1,color="Blue") +
    geom_hline(yintercept=0)

ggsave("rplot2.png", plot=last_plot(), device = "png")

{qqnorm(resid.cond)
qqline(resid.cond)}


ggsave("rplot.png", plot=last_plot(), device = "png")

detach(dat)
```

Linear Mixed Effects Model III

```r
lm.00 <- lm(Popular ~ (Harmony+Instrument+Voice)^2, data = dat)

olsrr::ols_step_backward_aic(lm.00)

lm.11<- lm(Popular ~Instrument+ Harmony+Voice, data = dat)
```

```r
lmer.11 <- lmer(Popular~Instrument+Harmony+Voice+(1|Subject), data= dat, REML = FALSE)

anova(lmer.11,lm.11)
```

```r
lmer.22<- lmer(Popular~Instrument+Harmony+Voice+(Instrument|Subject)+
                 (Harmony|Subject), data= dat, REML = FALSE,
               control = lmerControl(optimizer = 'bobyqa'))

formula(lmer.22)
anova(lmer.11,lmer.22)

summary(lmer.22)
```

```r
lm.22 <- lm(Popular~Instrument+Harmony+Voice+Composing+X16.minus.17+
              Selfdeclare + PachListen + OMSI + APTheory + ConsInstr +
              ConsNotes + Instr.minus.Notes + X1990s2000s + NoClass+
              GuitarPlay + PianoPlay+KnowRob+KnowAxis, data= dat)

olsrr::ols_step_backward_aic(lm.22)

lm.22 <- lm(Popular~Instrument+Harmony+Voice+Composing+X16.minus.17+Selfdeclare
            + PachListen + APTheory + ConsInstr + ConsNotes + X1990s2000s
            +GuitarPlay+KnowRob+KnowAxis, data= dat)
summary(lm.22)
```

```r
lmer.33 <- lmer(Popular~Instrument+Harmony+Voice+Composing+X16.minus.17
                +Selfdeclare + PachListen + APTheory + ConsInstr +
                  ConsNotes + X1990s2000s+GuitarPlay+KnowRob+KnowAxis
                +Selfdeclare:Composing+(Harmony|Subject)+(Instrument|Subject),
                data= dat, REML = FALSE, control = lmerControl(optimizer = 'bobyqa'))

lm.333 <- lm(Popular~Instrument+Harmony+Voice+Composing+X16.minus.17+
               Selfdeclare + PachListen + APTheory + ConsInstr + ConsNotes
             + X1990s2000s+GuitarPlay+KnowRob+KnowAxis+Selfdeclare:Composing, data= dat)

anova(lmer.33, lm.333)

summary(lmer.33)

formula(lmer.33)
```

```r
resid.cond <- r.cond(lmer.33)
fit.cond <- yhat.cond(lmer.33)

attach(dat)
index <- 1:dim(dat)[1]


new.data <- data.frame(index,resid.cond,Subject)
names(new.data) <- c("index","resid.cond","Subject")
ggplot(new.data,aes(x=index,y=resid.cond)) +
  facet_wrap( ~ Subject, as.table=F) +
  geom_point(pch=1,color="Blue") +
    geom_hline(yintercept=0)

ggsave("rplot3.png", plot=last_plot(), device = "png")

{qqnorm(resid.cond)
qqline(resid.cond)}


detach(dat)
```