The Sound of Music: What Influences Perception of Classical and Popular Music?

Mitchell Pudil*
mpudil@andrew.cmu.edu

30 November 2019

Abstract

We analyze collected data from a designed experiment intended to measure the influence of instrument, harmonic motion, and voice leading on 70 listeners' identification of music as "classical" or "popular". We use linear mixed models, ANOVA tests, and stepwise regression to answer further questions such as how I-V-VI harmonic motion affects classical ratings, the effect that contrary motion has on classical ratings, if there are differences in how musicians vs. non-musicians identify classical music, and if there are differences in the things that drive classical vs. popular music. We determine that alone, instruments have the strongest effect among the three design factors. We also find that the effect of the I-V-VI harmony on the average ranking of classical music is highly dependent on whether the listener is familiar with the Pachelbel rant, is a self-declared musician, or the voice leading used. We conclude with a discussion on the strengths, weaknesses, and implications of the results and methodology.

1 Introduction

Instrumentation, vocals, and harmony all important when it comes to creating and listening to music. Composer and musicologist Ivan Jimenez, alongside his student Vincent Rossi, were interested in understanding how important each of these measures are in determining the perceived genre of music for individuals with various levels of proficiency in music. They wanted to know whether instrumentation affect perception more than harmonic motion or voice leading, and whether I-V-VI has the strongest relationship to classical music among all other types of harmonic motion. They were also interested in whether familiarity with Pachelbel's work influences the relationship between the I-V-VI harmony and the classical ratings. Further, they wanted to know whether contrary motion has the strongest association with classical ratings, if there are differences in the way musicians vs. non-musicians identify classical music, and if there are differences in the things that drive classical vs. popular ratings.

This paper will attempt to answer these problems by focusing on the main questions of interest:

- What experimental factor, or combination of factors, has the strongest influence on ratings?
- Are there differences in the way that musicians and non-musicians identify classical music?

*Master's Student of Statistical Practice at Carnegie Mellon University

- Are there differences in the things that drive classical, vs. popular, ratings?

Previous research has determined that there are distinct differences in the types of people who listen to popular versus classical music. Prieto-Rodríguez and Fernández-Blanco (2000) used a bivariate probit model to determine that people vary in their interest of these types of music by socioeconomic standing, among many other factors. Castell (1982) performed a study that looked at children's sensitivity to stylistic differences in classical and popular music and found that popular pieces were more recognizable among 9-11 year olds. However, the children had trouble explaining exactly what distinguished popular music from classical music. This paper then also builds upon this research by using mixed methods to determine how instrumentation, vocals, and harmony (and the combination of them) affect the rating of whether a piece sounds classical or popular.

This paper is structured as follows: section 2 will describe the data and the variables used in the study and models. Section 3 will focus on the statistical analysis, results, and interpretations of the questions. Section 4 will serve to review the findings, address the strengths and weaknesses of the methodology, and suggest implications of the results and opportunities for further research.

2 Methods

Composer and musicologist Ivan Jimenez, alongside student Vincent Rossi, collected data in a designed experiment in 2012 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 listeners, all of whom were recruited from the population of undergraduates at the University of Pittsburgh. They then asked the listeners to rate how classical and popular the music sounded, on a scale from 1-10 (10 being highest).

Listeners 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 instrument (string quartet, piano, and electric guitar), harmonic motion (I-V-vi, I-VI-V, I-V-IV, and IV-I-V), and voice leading (contrary motion, parallel 3rds, and parallel 5ths).

The researchers also included background information on each of the subjects. These variables include:

- 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= notat 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 Canonin 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 fromthe 90's and 2000's? (0-5, 0=not at all)
- X1990s2000s.minus.1960s1970s: Difference between prev variable and asimilar variable referring to 60'sand 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 inHigh 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 firstmusical instrument (0-5, 0=not at all)
- X2ndInstr: Same, for second musical instrument

Additionally, we constructed the following variables to more fully be able to perform the analyses:

- binaryrob: Did the individual answer 5 for KnowRob? (0 = no, 1 = yes)
- musician: Did the individual answer 3, 4, or 5 for Selfdeclare? (0 = no, 1 = yes) Note: 2 is the median of Selfdeclare.

We used a variety of statistical methods in order to answer the questions stated previously. Generally speaking, we first used regression imputation techniques to fill in some missing data[1]. We also used stepwise regression to find the models that would be the best fit for the lowest cost[2]. We also took into consideration explainability of the model, being sure that each of the coefficients were interpretatiale. We additionally performed variable selection and variable transformations to determine the factors (and combinations of factors through introduction of interaction terms) that would be the best predictors. We then used binned plots, QQ plots, and scatterplots of residuals versus fitted values for marginal, conditional, and random effects residuals to ensure the residuals of the models were normally distributed and that the model was specified correctly. For each of the questions, we relied on these techniques and their output. The specific process by which we use the aforementioned methods will be explained further in the results section, with related code available throughout the appendix.

Through performing the methods explained above, we hope to gain a better understanding about how instruments, harmonic motion, voice leading are related to the perception of a song being classified by an individual as either classical or popular.

[1] In some cases, however, we chose to disregard the variable completely (see results section

[2] This was done specifically through comparing BIC

3 Results

Before directly answering the questions stated in the above sections, we begin by describing the variables and the relationships among the variables through a brief exploratory data analysis (EDA), along with explaining the transformations of select variables.

3.1 General Exploratory Data Analysis and Variable Transformations

To begin the EDA, we first determined how many missing values we had for each variable. The figure below demonstrates that there were several missing values for many columns.

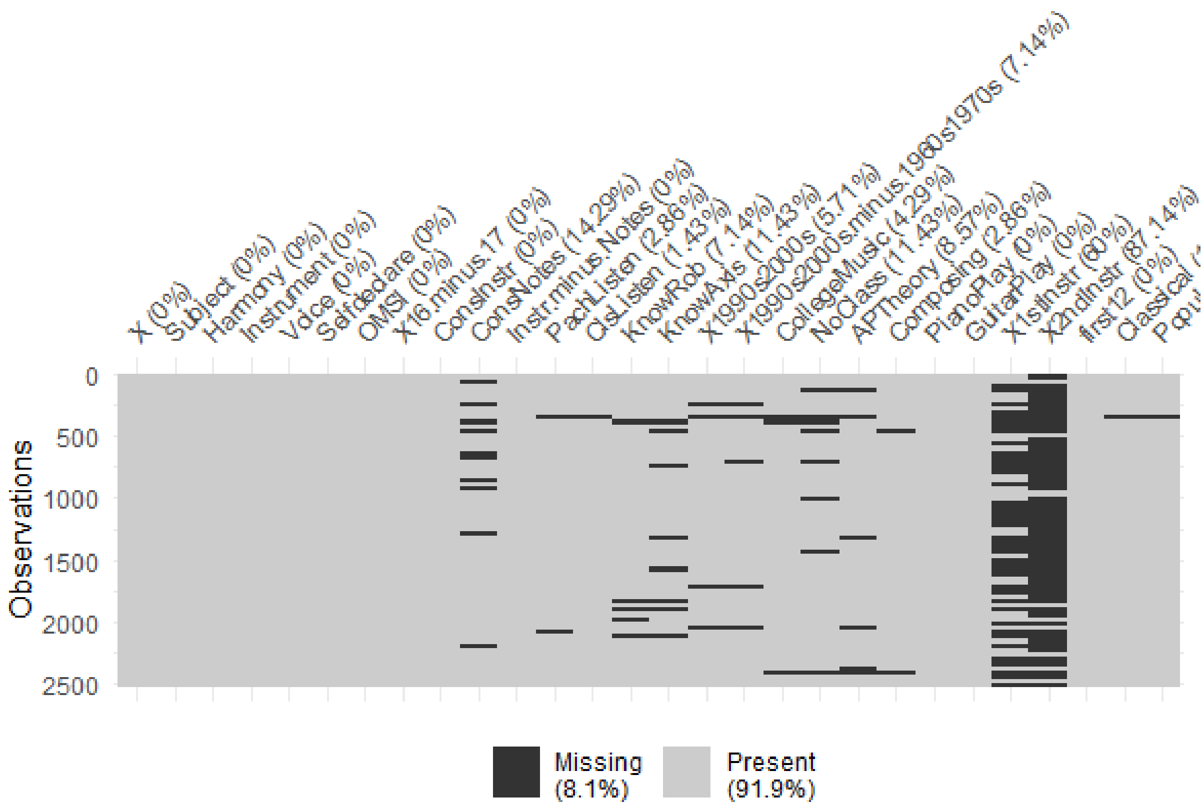


Figure 1: Missing Variables Present in Original Data

Fortunately, many of these variables did not end up being crucial to answer the proposed questions. However, a couple of variables that were important and had missing observations were KnowRob and PachListen. Fortunately, there appeared to be a statistically significant relationship between these two variables, and when one of the observations was missing, the other was not. In order to not lose so many observations, we chose to fill in the missing data through imputation, using a simple linear regression to approximate the value of the missing observation (see Appendix page 1). For other variables that had many missing observations (more than about 10%), we chose to delete the column entirely. This allowed us more observations to use, and did not affect the relationships among the more important variables (such as instrument, harmony, and voice).

As the next part of the EDA, we looked at each of the numeric variables[3] to see if there were any patterns, outliers, or variables that should be transformed. This plot is shown below as Figure 2.

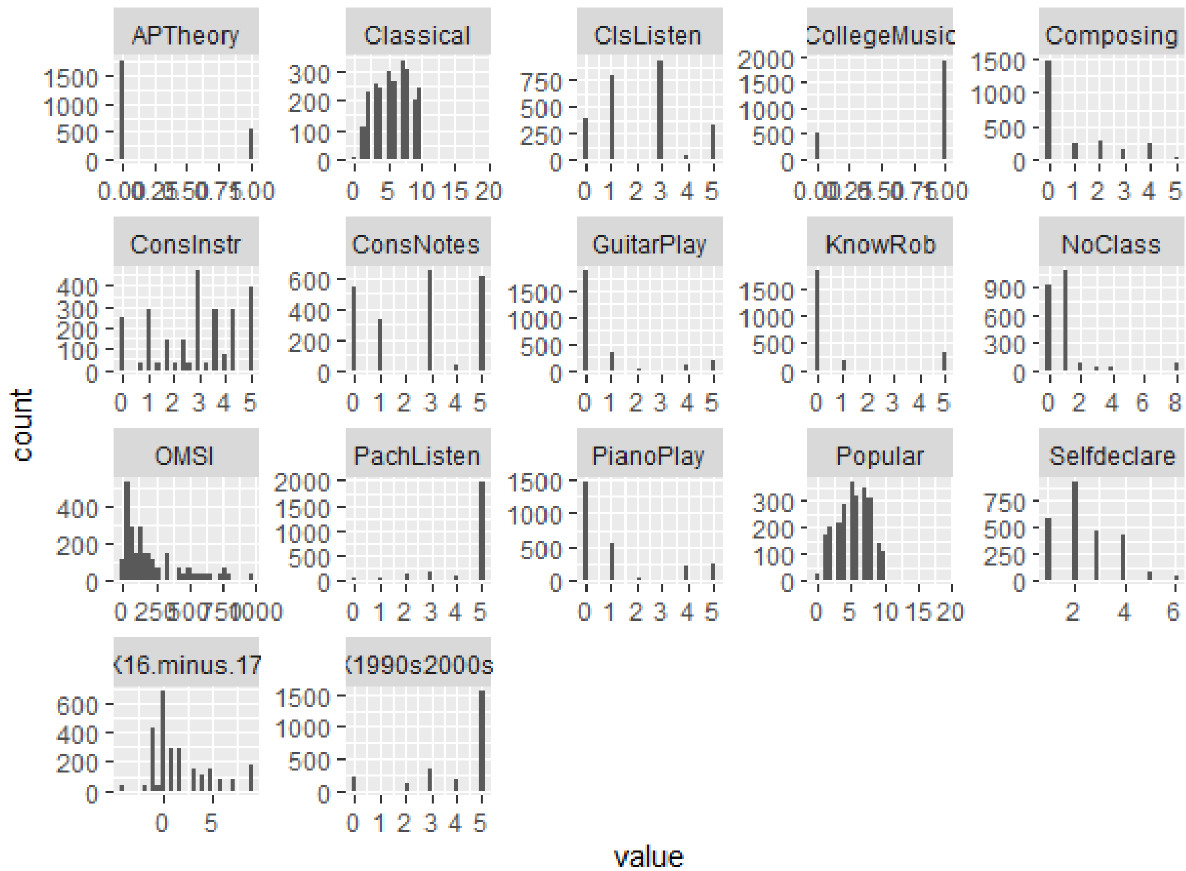


Figure 2: Plots of Numeric Variables

First, it was surprising to see that the maximum values for classical and popular ratings exceeded 10 since the maximum rating was to be 10[4]. For this reason, any observation with either a popular score or classical score greater than 10 was thrown out. Next, it appeared that the OMSI score should be logged, since it was skewed right. A power transform test affirmed that this was the case, so the OMSI score was logged for all models. All other strictly numeric variables seemed approximately normally distributed. For ease of interpretation later on, and because of the small number of observations in between the minimum and the maximum score[5], we dichotomized the KnowRob variable into a binary variable with a value of 0 (corresponding to a score less than or equal to 1) and 1 (corresponding to a score equal to 5).

[3] including some factor variables that only had values 0 and 1

[4] Fortunately, this only occurred twice in the study

[5] In fact, nobody put 2, 3, or 4 for KnowRob

Next, we were interested in understanding the relationships between these variables, so we constructed a correlation plot. This plot is shown below in Figure 3.

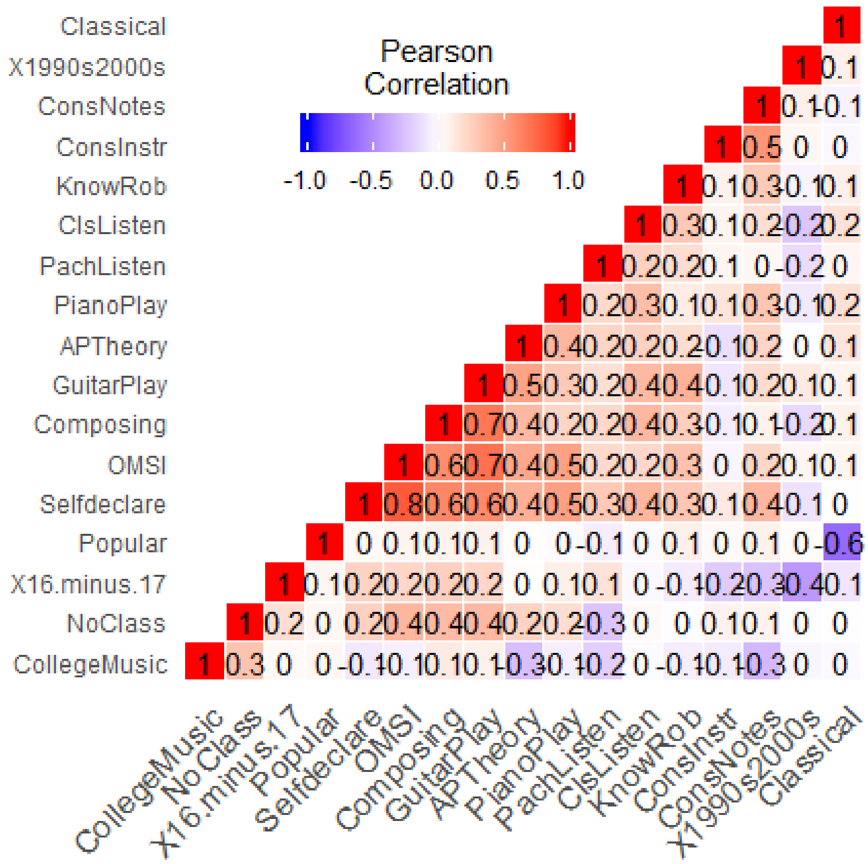


Figure 3: Correlation Plot of Numeric Variables

From the correlation plot, we found that there was a high correlation between many of the variables that would suggest a highly-musical person, including: Selfdeclare, OMSI score, GuitarPlay, PianoPlay, APTheory, and PachListen. As expected, there is a negative relationship between Classical and Popular. Overall, the correlations are not too surprising, suggesting that the results of the study have at least some measure of validity.

3.2 Factors with the Strongest Influence on Ratings

The first main question regards the relationship between predictor variables and their effect on classical ratings. All of the subquestions were answered using one model to predict classical ratings. To make the model, we first[6] started out by creating a linear regression model that examined only the influence of the three main experimental factors (with interaction terms) on classical ratings (see Appendix page 9). Using stepwise regression, we determined that it was crucial to include the interaction between voice and harmony, which we found to be statistically significant (see Appendix

[6]After performing the EDA described in the previous section

pages 9-10). We then found it appropriate to include a random intercept by comparing the linear model to one with a random intercept by using an ANOVA test. We went back and forth between optimizing random effects and fixed effects using stepwise regression and ANOVA testing, to create an optimal, yet explainable, model (see Appendix pages 11-15). We made sure that we did not use an intercept so we could better see what the individual effects are of each factor variable on classical ratings. Finally, we made sure that the chosen model had residuals that were normally distributed and appropriate for the model. Specifically, we relied on binned plots to help us determine whether the model overfit or underfit for various levels of predictions (see Appendix page 17). We also looked at marginal residuals and their relationship to the predicted values to make sure there was approximately equal spread. We used QQ plots of the conditional residuals and of the random effects residuals to further ensure normality of the residuals (see Appendix pages 18-20). The final model for determining classical ratings is shown below:

The first subquestion asks what the most important design factor was. In order to determine this, we first need to be sure to correctly interpret the regression coefficients as shown on Table I, since we are releveling the Voice leading factor. Specifically, the coefficients of the levels of the Voice leading factor have no baseline, which makes the interpretations differ from those of harmony or instrumentation. The coefficients for voice are the individual effects of the particular type of voice leading on classical ratings. On the other hand, the coefficients for instrument are the added effect on classical music for the level of instrumentation. For example, it was found that using the piano increased the average classical rating by 1.89 points (SE 0.28) compared to the electric guitar, and that using a string instrument increased the average score by 3.77 points (SE 0.38). To compare voice level to instrument, it is better to look at the spread of coefficients by using 0 as the coefficient for guitar. Doing this suggests that instrumentation has a much larger effect than voice on classical ratings since the coefficients are much different from one level to the next with instrumentation, but rather close together with voice, suggesting that regardless of the level of voice, the classical ratings increase by about 4 points. In fact, we can also see that instrumentation leads to the largest effect overall even when considering interactions between design factors by using the same logical comparisons.

The next subquestion regarded the effect of the I-V-VI harmonic motion on classical ratings and whether the relationship is affected by familiarity to Rob Paravonian's Pachelbel rant [7]. To begin answering this question, we graphed the relationship between harmony and classical ratings, separated by whether or not the subject was very familiar with the Pachelbel rant. This is shown as Figure 4 below.

[7] See https://www.youtube.com/watch?v=JdxkVQy7QLM

	Effect on Classical Ratings
Voicecontrary	4.26^{***}
	(0.37)
Voicepar3rd	3.98^{***}
	(0.37)
Voicepar5th	4.05^{***}
	(0.37)
HarmonyI-V-IV	0.16
	(0.21)
HarmonyI-V-VI	$0.49^{\prime}$
	(0.27)
HarmonyIV-I-V	-0.31
	(0.21)
binaryrob	0.21
	(0.48)
Instrumentpiano	1.89^{***}
	(0.28)
Instrumentstring	3.77^{***}
	(0.38)
musician	-0.26
	(0.46)
PianoPlay	0.28^{**}
	(0.10)
X16.minus.17	-0.07
	(0.05)
ConsNotes	-0.12
	(0.08)
HarmonyI-V-IV:binaryrob	0.06
	(0.34)
HarmonyI-V-VI:binaryrob	1.50^{**}
	(0.51)
HarmonyIV-I-V:binaryrob	0.11
	(0.33)
Voicepar3rd:HarmonyI-V-IV	$-0.49^{\prime}$
	(0.26)
Voicepar5th:HarmonyI-V-IV	-0.23
	(0.26)
Voicepar3rd:HarmonyI-V-VI	-0.70^{**}
	(0.26)
Voicepar5th:HarmonyI-V-VI	-0.56^{*}
	(0.26)
Voicepar3rd:HarmonyIV-I-V	0.67^{**}
	(0.26)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\prime}p < 0.1$

	Effect on Classical Ratings
Voicepar5th:HarmonyIV-I-V	0.15
	(0.26)
HarmonyI-V-IV:musician	0.03
	(0.24)
HarmonyI-V-VI:musician	1.08^{**}
	(0.35)
HarmonyIV-I-V:musician	0.07
	(0.23)
Instrumentpiano:musician	-0.55
	(0.42)
Instrumentstring:musician	-0.78
	(0.57)
AIC	7158.91
BIC	7438.70
Log Likelihood	-3528.45
Num. obs.	1783
Num. groups: Subject	50
Var: Subject (Intercept)	0.00
Var: Subject.1 HarmonyI-IV-V	0.70
Var: Subject.1 HarmonyI-V-IV	1.10
Var: Subject.1 HarmonyI-V-VI	1.23
Var: Subject.1 HarmonyIV-I-V	0.82
Cov: Subject.1 HarmonyI-IV-V HarmonyI-V-IV	0.84
Cov: Subject.1 HarmonyI-IV-V HarmonyI-V-VI	0.51
Cov: Subject.1 HarmonyI-IV-V HarmonyIV-I-V	0.73
Cov: Subject.1 HarmonyI-V-IV HarmonyI-V-VI	0.83
Cov: Subject.1 HarmonyI-V-IV HarmonyIV-I-V	0.85
Cov: Subject.1 HarmonyI-V-VI HarmonyIV-I-V	0.63
Var: Subject.2 Voicecontrary	0.33
Var: Subject.2 Voicepar3rd	0.37
Var: Subject.2 Voicepar5th	0.22
Cov: Subject.2 Voicecontrary Voicepar3rd	0.28
Cov: Subject.2 Voicecontrary Voicepar5th	0.24
Cov: Subject.2 Voicepar3rd Voicepar5th	0.28
Var: Subject.3 Instrumentguitar	0.87
Var: Subject.3 Instrumentpiano	1.28
Var: Subject.3 Instrumentstring	0.98
Cov: Subject.3 Instrumentguitar Instrumentpiano	0.20
Cov: Subject.3 Instrumentguitar Instrumentstring	-0.90
Cov: Subject.3 Instrumentpiano Instrumentstring	-0.01
Var: Residual	2.42

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 1: Classical Ratings Model

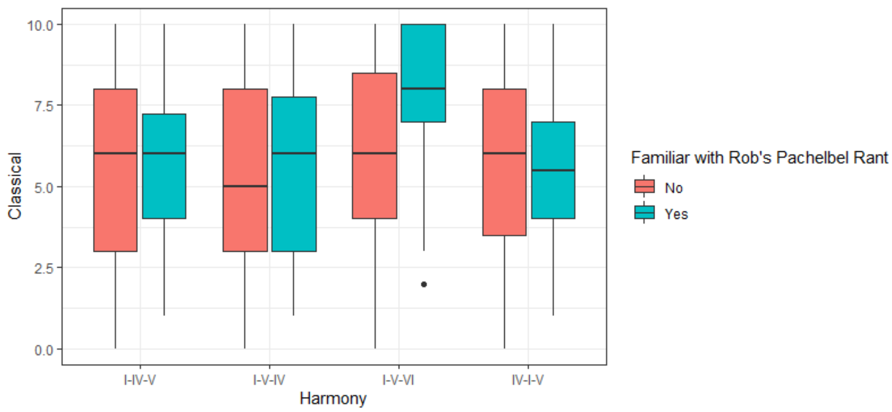


Figure 4: The Relationship Between Harmonic Motion and Classical Ratings Varies by Familiarity with the Pachelbel Rant

According to the graph, it appears that on average, harmony isn't a great predictor of classical rating, but that musicians tend to rate classical music higher than non-musicians if the there is a I-V-VI harmonic motion. Using the model shown in Table I, it was determined that alone, none of the levels of harmonic motion were statistically significant at the 5% level. However, when combined with voice, familiarity with Pachelbel, or whether the subject is a musician, the I-V-VI harmonic motion is statistically significant, as predicted by Figure 4. Specifically, individuals who are very familiar with the Pachelbel rant (score = 5) rated classical music about 1.5 points higher (SE 0.51).

While instrumentation was the single most influential factor among the three design factors, it is interesting to note that there was a large effect of several interaction terms between harmony and voice on classical ratings.

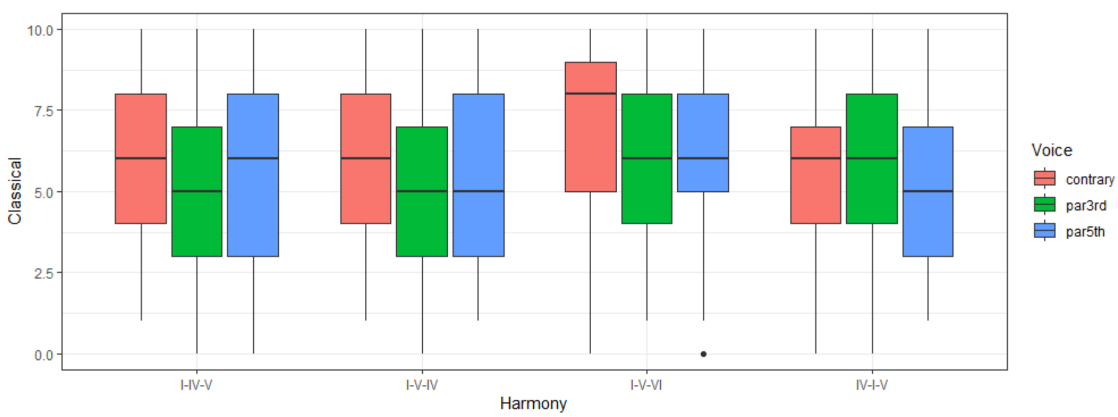


Figure 5: The Relationship Between Harmony and Classical Ratings Varies by Voice

Figure 5 above shows that there is likely some interaction effect between harmony and voice, especially when the harmony is I-V-VI and voice is contrary motion. Using Table I, we found that the I-V-VI harmonic motion has the strongest relationship with classical music among all other harmonic motions, but only when interaction terms are considered. That is, no level of harmonic motion is statistically significant by itself, but when interacted with voice, or even whether the individual is a musician or is familiar with the Pachelbel rant, the I-V-VI harmony has more of an effect on classical ratings than any other harmonic motion. Specifically, when combined with par 3rd voice leading, using I-V-VI as the harmony decreases average classical ratings by 0.7 points (SE 0.26) relative to using contrary motion and I-V-VI harmony. Additionally, the average musician in the study increased the classical rating by 1.08 points (SE 0.35) when the harmony was I-V-VI, compared to when the harmony was I-IV-V. Finally, when the individual was familiar with the Pachelbel rant and the harmony was I-V-VI, the individual increased the classical rating by 1.5 points (SE 0.51).

We also found that the variance of the harmony I-V-VI random effect for the classical model was 1.23. This suggests that people differ widely in how they score classical music when they hear the I-V-VI harmony. Specifically, about 95% of the baseline subjects[8] will rate the classical music somewhere between -2 and 3 points lower/higher if they hear the I-V-VI harmony, compared to IV-I-V. However, 95% of musicians that have heard Rob Pavalonian's rant tend to rate the I-V-VI harmony somewhere around 0.5 and 5.5 points higher than when they hear the IV-I-V harmony.

We were also curious in determining if contrary motion has a strong association with classical ratings. Using the same model from Table I, it is first evident that contrary motion has the largest coefficient, suggesting that on average, when contrary motion was used, it led to the highest classical rating. However, no level of voice leading had a statistically different effect than the other[9]. However, we can still get a sense of the effect that contrary motion has on classical ratings by again looking at interaction terms. Figure 6, essentially a flipped version of Figure 5, shows that when using I-V-VI, that contrary motion is rated higher on average than any other voice-harmony combination, compared to I-IV-V.

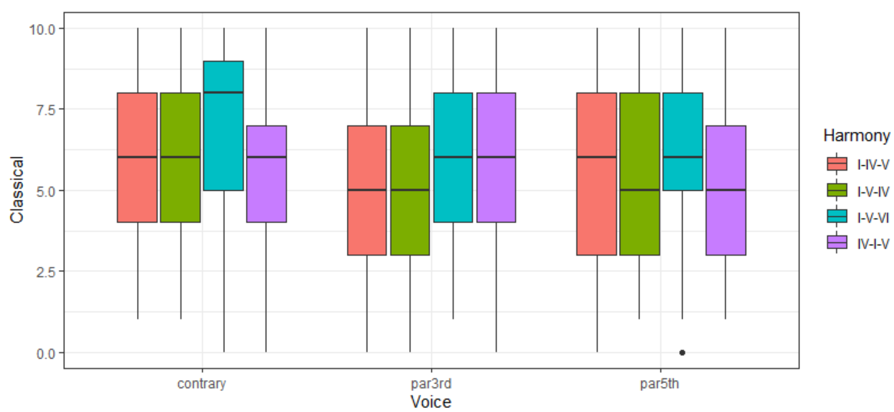


Figure 6: The Relationship Between Voice and Classical Ratings Varies by Harmony

The classical ratings model also shows that the interaction between contrary motion and I-V-VI is significant. According to the classical ratings model, when using the I-V-VI progression, using contrary motion increases the average classical ratings by 0.7 points (SE 0.25), holding all else constant, compared to using par3rd. On the other hand, IV-I-V paired with contrary motion decreases the average classical ratings by 0.67 (SE 0.25), holding all else constant, compared to

[8] i.e. have never heard Rob Pavalonian's rant and not a musician

[9] This can easily be eyeballed by comparing coefficients with standard error, but a t-test will also show that the estimates are not statistically different with p-value ¡ 0.01

par 3rd. The sizes on the coefficients for the variables and interaction terms with contrary motion is larger on average then the sizes of the statistically significant coefficients for the variables and interaction terms of the other voice leading levels.

Using the random effects shown in Table I, we find that individuals varyAlthough perhaps not by as much as when they hear the I-V-VI harmony in how they rate classical music when contrary motion is used. The variance of the contrary motion random effect is 0.33, suggesting that 95% of the baseline subjects will rate contrary motion somewhere between 3.6 and and 4.8 points, plus whatever the average effects are of the other design factorsThe number of points varies based on whether there is the I-V-VI harmony. Thus, the model as a whole leads us to conclude that contrary motion does have a strong relationship to classical ratings.

3.3 Differences in Classical Ratings Between Musicians and Non-musicians

Our second question asks if there are any differences in the way that musicians and non-musicians identify classical music. In order to answer this, we first examined the relationship between the three design factors and classical score, by whether or not the individual declared themselves a musician[10]. Figure 7 below shows this graph for the classical ratings.

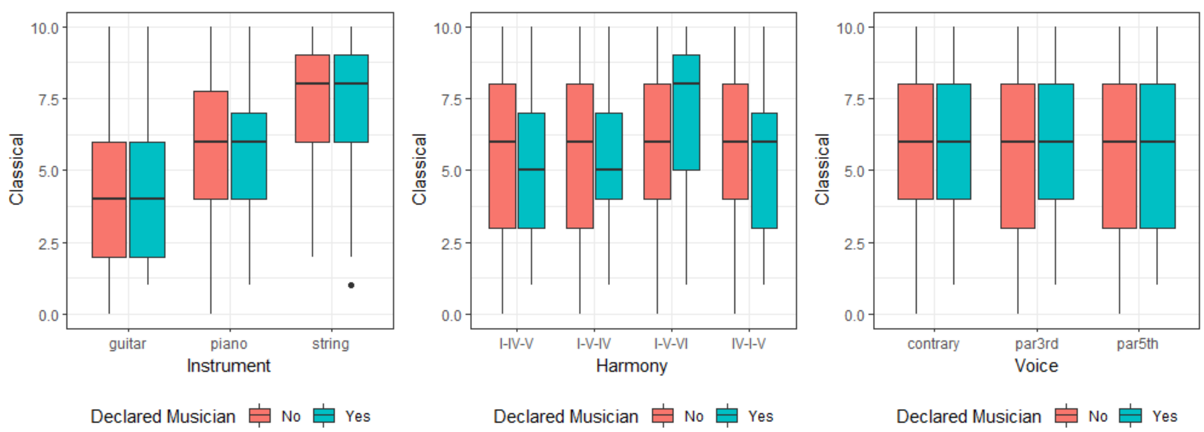


Figure 7: Boxplots of Classical Ratings by Design Factor and Musician

From Figure 7, it is difficult to tell whether being a musician affects the relationship of any of the levels of harmonic motion and classical ratings except for perhaps I-V-VI. It appears in that case that musicians tend to increase their classical ratings more than non-musicians.

Looking back at the classical ratings model shown in Table I, we in fact find that the only statistically significant variable with musician was the interaction term between musician and the I-V-VI harmonic motion. Specifically, we find that if the harmonic motion is I-V-VI, then the average effect of being a musician on classical rating increases by 1.08 points (SE 0.35) compared to non-musicians. We do not have enough evidence, however, to determine whether there are other ways in which musicians rate classical music differently than non-musicians.

[10]Here, we used the constructed "musician" variable by dichotomizing at the median, as explained above.

3.4 Differences in Predictors Between Classical and Popular Music

Our final question asks if there is any difference in the things that drive classical and popular music. In order to determine this, we first compared the ratings of each level of instrument, voice, and harmonic motion to the respective ratings. Figure 8 shows these boxplots for both ratings.

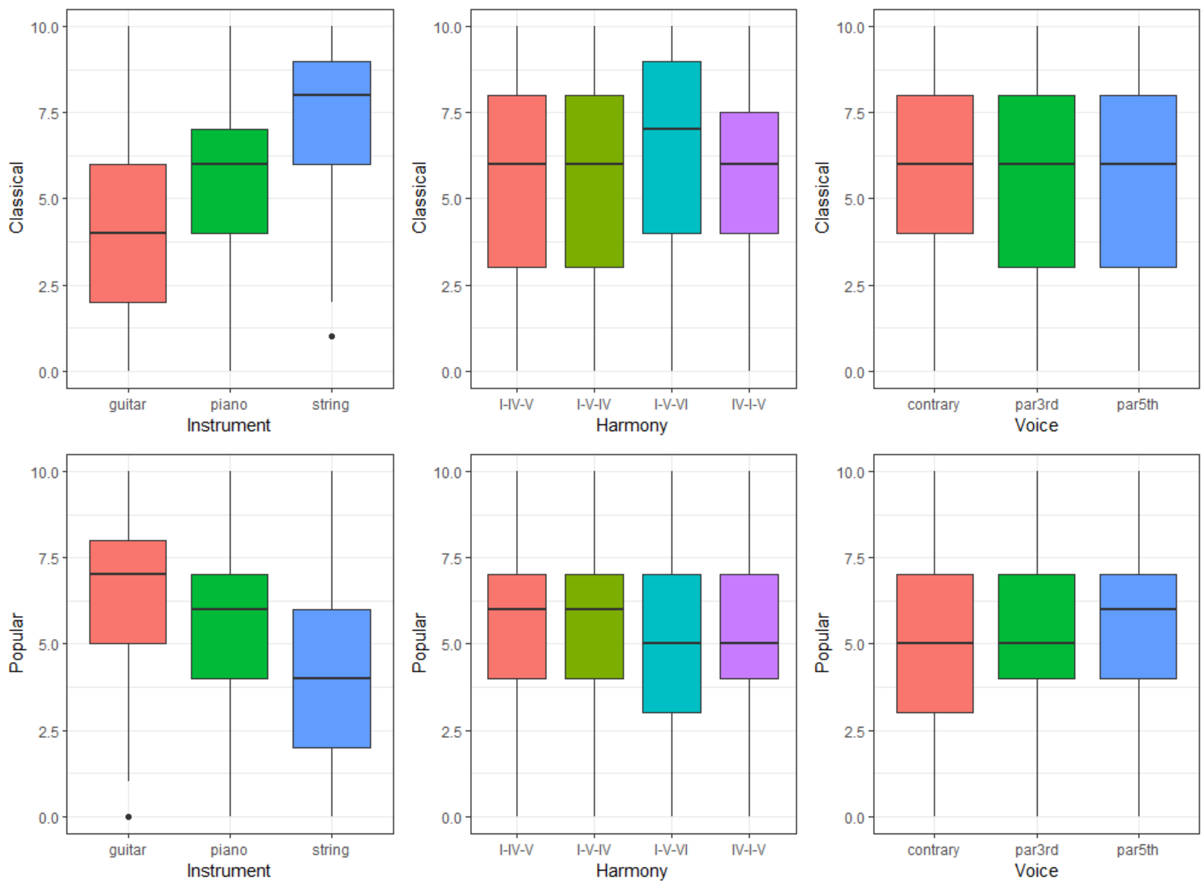


Figure 8: Boxplots of Classical and Popular Ratings by Design Factor

From Figure 8, it appears that instrument is the single most influential driver in determining classical ratings as well as popular ratings. Specifically, it seems that for each instrument type that makes classical ratings increase, it causes the corresponding popular ratings decrease. It is difficult to tell by just looking at the plots if there is any statistically significant difference with harmony or voice alone (and in fact we showed that alone[11], there were no significant levels of the harmony or voice factors). For this reason, we created another model to determine the extent to which various factors impacted the ratings of popular music, using similar techniques as we did to create the classical model[12] (see appendix pages 20-28). The model and related estimates

The most apparent difference here is that the only factors that were necessary to model the popular ratings were the three design factors and whether or not the individual was a musician

[11] i.e. other than interaction terms

[12] (i.e. stepwise regression, analysis of variance, similar transformations of variables, and residual analysis

	Effect on Popular Ratings
Voicecontrary	6.78^{***}
	(0.25)
Voicepar3rd	6.94^{***}
	(0.23)
Voicepar5th	6.99^{***}
	(0.25)
HarmonyI-V-IV	-0.13
	(0.15)
HarmonyI-V-VI	0.08
	(0.15)
HarmonyIV-I-V	-0.25
	(0.15)
musician	0.00
	(0.33)
Instrumentpiano	-1.09^{***}
	(0.21)
Instrumentstring	-2.77^{***}
	(0.28)
HarmonyI-V-IV:musician	0.19
	(0.23)
HarmonyI-V-VI:musician	-0.75^{**}
	(0.23)
HarmonyIV-I-V:musician	0.05
	(0.23)
AIC	7296.30
BIC	7438.94
Log Likelihood	-3622.15
Num. obs.	1783
Num. groups: Subject	50
Var: Subject (Intercept)	0.00
Var: Subject.1 Voicecontrary	1.10
Var: Subject.1 Voicepar3rd	0.81
Var: Subject.1 Voicepar5th	1.10
Cov: Subject.1 Voicecontrary Voicepar3rd	0.94
Cov: Subject.1 Voicecontrary Voicepar5th	1.09
Cov: Subject.1 Voicepar3rd Voicepar5th	0.94
Var: Subject.2 Instrumentguitar	0.20
Var: Subject.2 Instrumentpiano	1.36
Var: Subject.2 Instrumentstring	2.01
Cov: Subject.2 Instrumentguitar Instrumentpiano	-0.07
Cov: Subject.2 Instrumentguitar Instrumentstring	-0.55
Cov: Subject.2 Instrumentpiano Instrumentstring	0.98
Var: Residual	2.89

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

Table 2: Popular Ratings Model

(and various interaction terms). All other factors afterward were either unnecessary because they didn't add much more prediction to the ratings or not statistically significant.

Additionally, and as expected, the remaining variables have a much different effect on popular ratings than on classical ratings. However, there are similarities between which variables are statistically significant and which are not, as well as the absolute size of the coefficients. For example, we find that the single most important design factor category in both the popular and classical models is instrumentation. In the classical model, we determined that using a string quartet or piano increases ratings. However, using a piano decreases the popular ratings on average by about 1.09 points (SE 0.21) and using a string quartet decreases the popular ratings by 2.77 points (SE 0.28).

Furthermore, we found that the interaction between musician and the I-V-VI progression is statistically significant for popular ratings. Specifically, musicians tend to decrease their average vote by 0.75 points (SE 0.23) when the harmony is I-V-VI, compared to I-IV-V . Again, musician and harmony alone have no statistically significant effects. However, we do find that when using 5ths, the ranking for popular music is about 0.2 points higher on average[13].

Regarding random effects, we found that in the classical ratings model, it was crucial to add harmony, voice, and instrument random effects. This suggests that people vary in the degree to which they are inclined to call music with specific instruments, vocals, or harmony "classical." However, we found that it was not appropriate to include harmony random effects in the popular ratings model, suggesting that the degree to which individuals differ in how they rate classical music based on harmony was not an important effect to predict popular ratings. However, by comparing Table 1 with Table 2, we find that individuals tend to vary their effect of hearing particular voice leading styles much more when it comes to rating classical music (average S.D. among levels is 0.9) than with rating popular music (average S.D. among levels is 0.3). Further, we find that the degree to which individuals vary their classical ratings when they hear electric guitar is much higher than when they are rating popular music and hear the electric guitar. Specifically, the variance in the random effect for classical music when the guitar is played is 0.87, whereas for popular music, it is 0.2. This all suggests that there are many ways in which individuals rate popular music differently than classical music, and in the things that drive classical versus popular music.

4 Discussion

In analyzing the data for this experiment, we were primarily interested in discovering how voice, harmony, and instrument affected the ratings of classical and popular music. We also had several secondary questions, including the effect that contrary motion has on classical rating, or whether people who classify themselves as musicians are influenced by things that do not influence non-musicians.

We will begin by discussing the most important factors that affect classical rankings. In order to determine the solution to this question, we used techniques such as stepwise regression, variable transformation, and ANOVA comparisons to determine the model that would be the most interpretable, informative, and accurate. From that model, we found that the single most important, statistically significant design factor that contribute to a piece being classified as "Classical" was

[13]This value was found to be significant at the 0.05 level because of the small standard errors of all levels of the factor

the instrument played. Furthermore, we determined that classical rankings vary significantly based on the combination of voice leading and harmonic motion.

We further found the interaction between being a self-declared musician and the I-V-VI harmony to be statistically and practically significant, suggesting that musicians rated songs with that harmony even higher than non-musicians. We also found that the effect of I-V-VI varies by whether the individual was familiar with the Pachelbel . Specifically, if the individual was familiar and I-V-VI was used, then the subject rated the classical score about 1.5 points higher (SE 0.5).

The last part of the question regarding the influential factors on classical ratings asks about the effect of contrary motion. We determined that by itself, there is no evidence to suggest that contrary motion affects classical rating. However, when we interacted the variable with harmonic motion, we found that when the harmony is I-V-VI, that individuals rate songs with contrary motion about 0.7 points higher on average (SE 0.25) than when the harmony is I-IV-V.

The second major question we were interested in answering was whether musicians and non-musicians vary in the way they identify classical music. We found limited evidence to suggest many differences in the way that musicians and non-musicians identify classical music. As explained in the paragraph above, we were able to determine that musicians increased their classical ratings by about a point on average if I-V-VI was used. Other than that, we were unable to reject the hypothesis that musicians are more inclined to rate classical music higher, even considering other interaction terms.

Finally, we were interested in determining which factors affect the "Popular" score, and how the relevant predictors differ from classical ratings. Again, we used stepwise regression, transformation of variables, and analysis of variance to determine that only instrument, and the interaction between Harmony I-V-VI and musician were statistically significant variables[14], and that the best predictive model only used those variables (and some interactions) to determine popular ratings.

Generally speaking, this research paved the road for determining how people understand classical and popular music. Professors interested in helping non-musicians recognize classical music could, for example, introduce the students to the Pachelbel's rant or I-V-VI / harmonic motion progressions. Students could benefit from these studies by realizing how their personal biases and experience in music helps them identify particular songs. Singers and songwriters looking to make their music sound more classical or popular could also determine if and how they should change the progressions they use.

There are many strengths and weaknesses to this study. The major strength of the study is that it was performed in a controlled environment, and that many of the measurements were interpretable. Further, it examined dozens of combinations of harmony, voice, and instrumentation. Additionally, the researchers were sure to perform their experiment on subjects with various levels of musical backgrounds rather than just music students. This all allowed us to understand, to an extent, how individuals with various backgrounds determine how classical- or popular-sounding certain pieces are.

On the other hand, there are many weaknesses to the study. First, we only have at our disposal one example of each combination of stimuli. If we had more examples of the combinations of the stimuli, we would have been able to determine more precisely the effect that the design factors as well as music background has on classical and popular ratings. Secondly, the sample size is comprised of people from the same university, which questions external validity. It is likely that these students were not independently chosen. Further, they may all be about the same age and

[14] Recall that the significance level for Voice is potentially misleading since the coefficients have no baseline

have other similarities to each other. It would have been better to use people from many different locations for the study.

Moreover, it would have been more useful to have more background on the individuals that was not directly related to their musical expertise, for example their age, sex, and ethnicity (if considered ethical). While most of the questions were useful, the questions regarding how proficient the individual was in their first and second instruments should have been complimented with questions about how many instruments the individual could play. More than half of the subjects left that question blank. While it is possible that this was due to the subject not knowing any instruments, it could have also been the case that they left the questions blank for other reasons like they did for many of the other questions. Including a question about how many instruments the student could play would allow us to determine which students left the original instrument questions blank because they did not know how to play an instrument and which students left the question blank for other reasons. Finally, it would also be useful to include some measure of "truth" regarding the classical vs. popular ratings, either by using songs that are strictly classical or popular, or else ask professionals how classical or popular the songs sound and average their answers. This would allow us to determine how well different subjects are really able to classify popular versus classical music.

Further research should be aimed at first improving the quality of the test using the suggestions stated in the previous paragraphs. Subsequently, there are many other questions that are unanswered that would be interesting. For example, it might be interesting to determine the extent to which age affects people's ability to distinguish between popular and classical music. In fact, research by Juan Prieto-Rodríguez and Víctor Fernández-Blanco (2000) suggests that there are differences between the consumers of popular and classical music. It could also be interesting to find how preferences in music genre affects people's ranking or ability to classify certain songs. Finally, it may also be of interest to determine how ranking or classification of other genres such as rock or jazz is affected by musical ability.

References

Castell, K. C. (1982). Children's sensitivity to stylistic differences in "classical" and "popular" music. Psychology of Music, Spec Iss, 22-25.

Jimenez, Ivan Rossi, Vincent (2012). "The Influence of Timbre, Harmony, and Voice Leading on Listeners' Distinction between Popular and Classical Music." University of Pittsburgh, Pittsburgh, PA.

Prieto-Rodríguez, Juan and Víctor Fernández-Blanco. 2000. "Are Popular and Classical Music Listeners the Same People?" Journal of Cultural Economics 24(2):147–64.

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/.

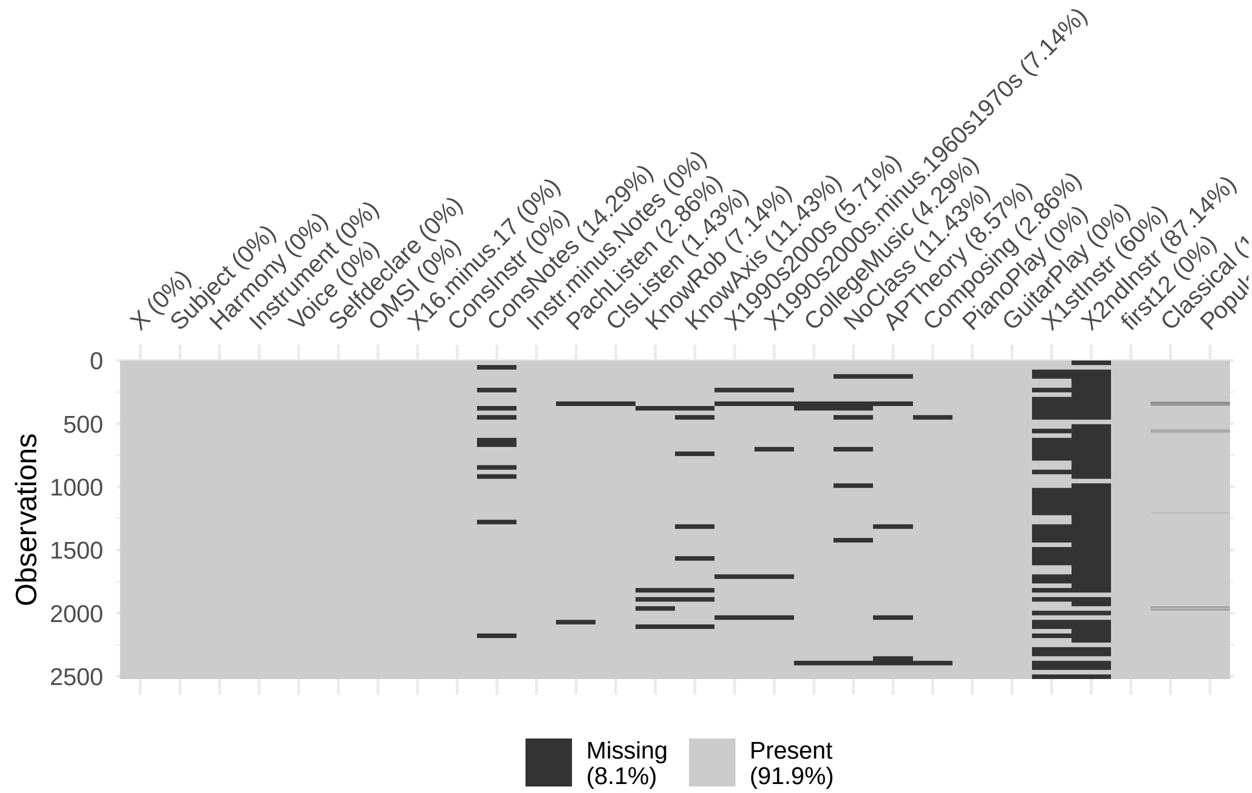
Appendix

Mitchell Pudil

11/30/2019

EDA and Variable Transformations

Let's begin by looking at the missing data



We notice there are several columns that have missing data.

A couple of the variables we will be using throughout the study are KnowRob and PachListen. However, both of these variables have many missing data points. However, nobody has both missing. We can exploit these observations and fill in the missing data by the average of the levels in the other.

First, determine the relationship between the two variables

```
summary(lm(ratings$KnowRob ~ ratings$PachListen))
```

```
##
## Call:
## lm(formula = ratings$KnowRob ~ ratings$PachListen)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9308 -0.9308 -0.9308  0.0692  4.0692
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         -0.50943    0.15245  -3.342 0.000846 ***
## ratings$PachListen  0.28805    0.03275   8.795  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.71 on 2266 degrees of freedom
##   (252 observations deleted due to missingness)
## Multiple R-squared:  0.03301,    Adjusted R-squared:  0.03258
## F-statistic: 77.35 on 1 and 2266 DF,  p-value: < 2.2e-16
```

```
summary(lm(ratings$PachListen ~ ratings$KnowRob))
```

```
##
## Call:
## lm(formula = ratings$PachListen ~ ratings$KnowRob)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.4329 -0.0058  0.5671  0.5671  0.5671
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.43286    0.02490 178.057   <2e-16 ***
## ratings$KnowRob  0.11460    0.01303   8.795   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.079 on 2266 degrees of freedom
##   (252 observations deleted due to missingness)
## Multiple R-squared:  0.03301,    Adjusted R-squared:  0.03258
## F-statistic: 77.35 on 1 and 2266 DF,  p-value: < 2.2e-16
```

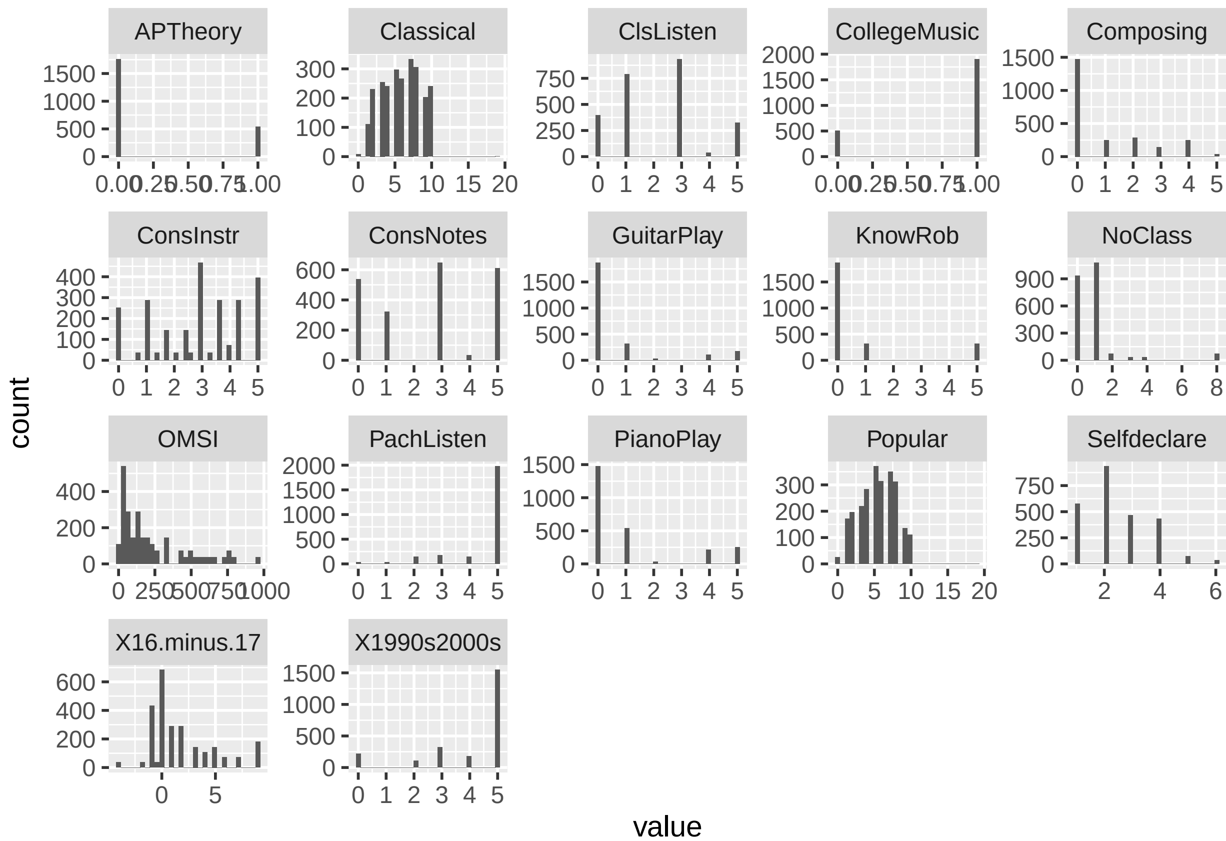
Then fill in missing data

```
for(i in 1:nrow(ratings)){
  if(is.na(ratings$PachListen[i])) {
    ratings$PachListen[i] <- round(4.43 + 0.114*ratings$KnowRob[i])
  }

  if(is.na(ratings$KnowRob[i])) {
    ratings$KnowRob[i] <- max(round(-0.5 + 0.28*ratings$PachListen[i]), 0)
  }
}
```

Plot Numeric Variables

```r
ratings[,-c(1:5, 11, 15, 17, 24, 25, 26)] %>%
  keep(is.numeric) %>%
  gather() %>%
  ggplot(aes(value)) +
    facet_wrap(~ key, scales = "free") +
    geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

## Warning: Removed 1278 rows containing non-finite values (stat_bin).
```

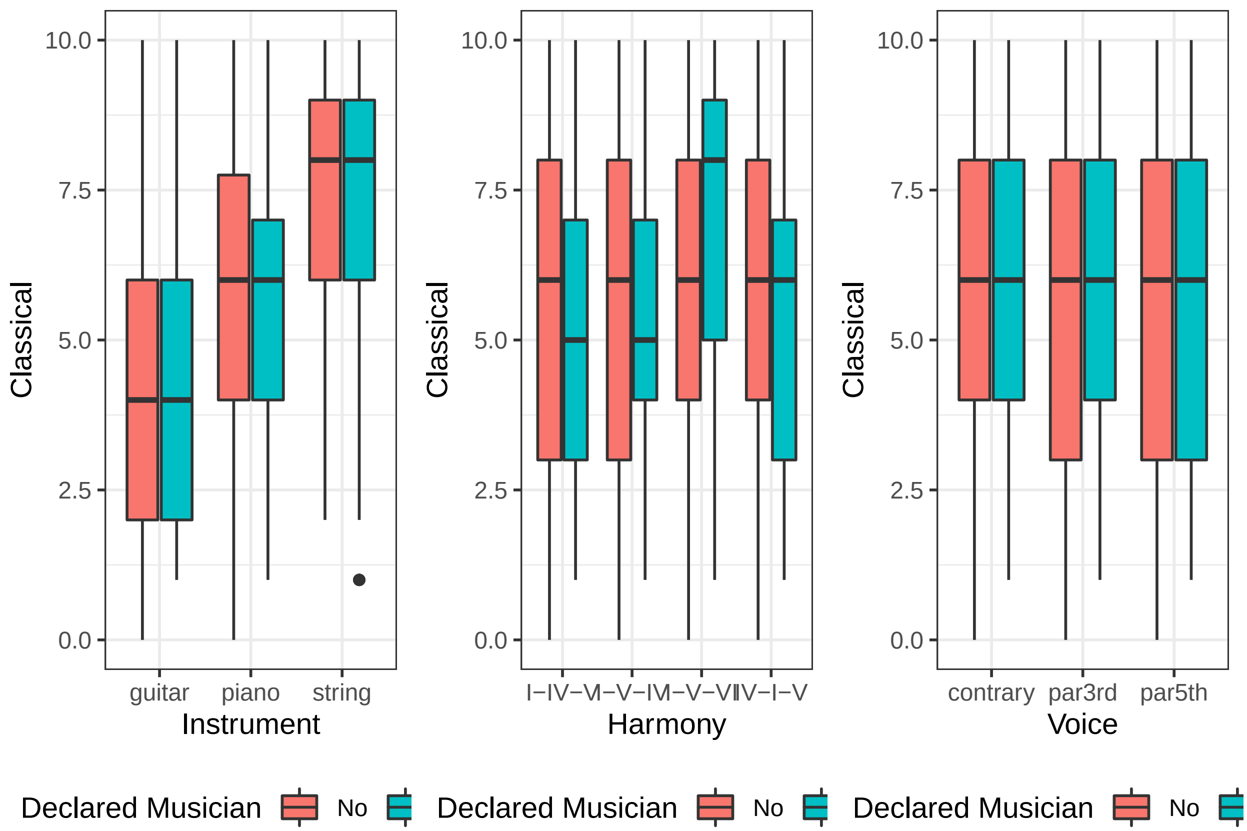


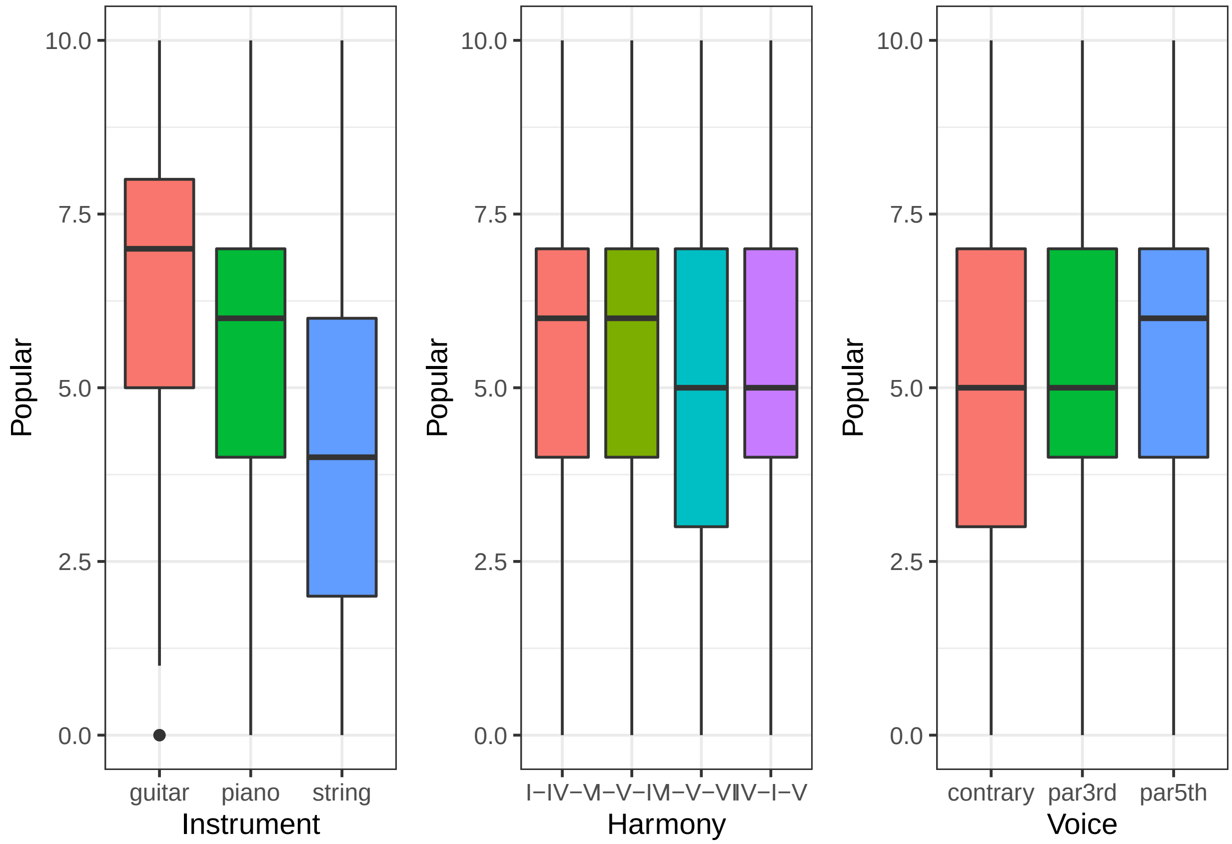
OMSI Score and classical/popular outliers

```r
powerTransform(OMSI ~ 1, data=ratings) # Should use log in model
```

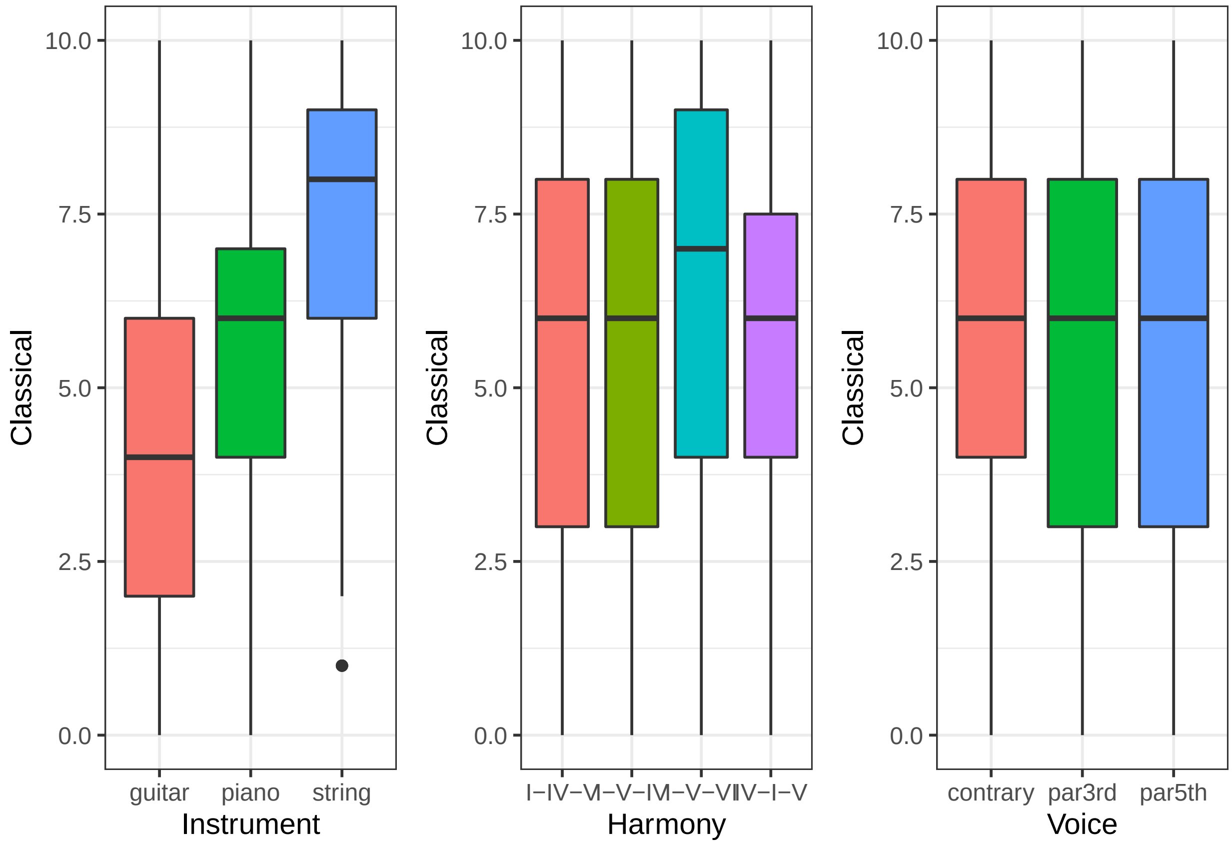
```
## Estimated transformation parameter
##          Y1
## 0.08565721
```

```r
ratings$binaryrob <- ifelse(ratings$KnowRob==5, 1, 0)
ratings <- subset(ratings, Popular <= 10 & Classical <= 10)
```

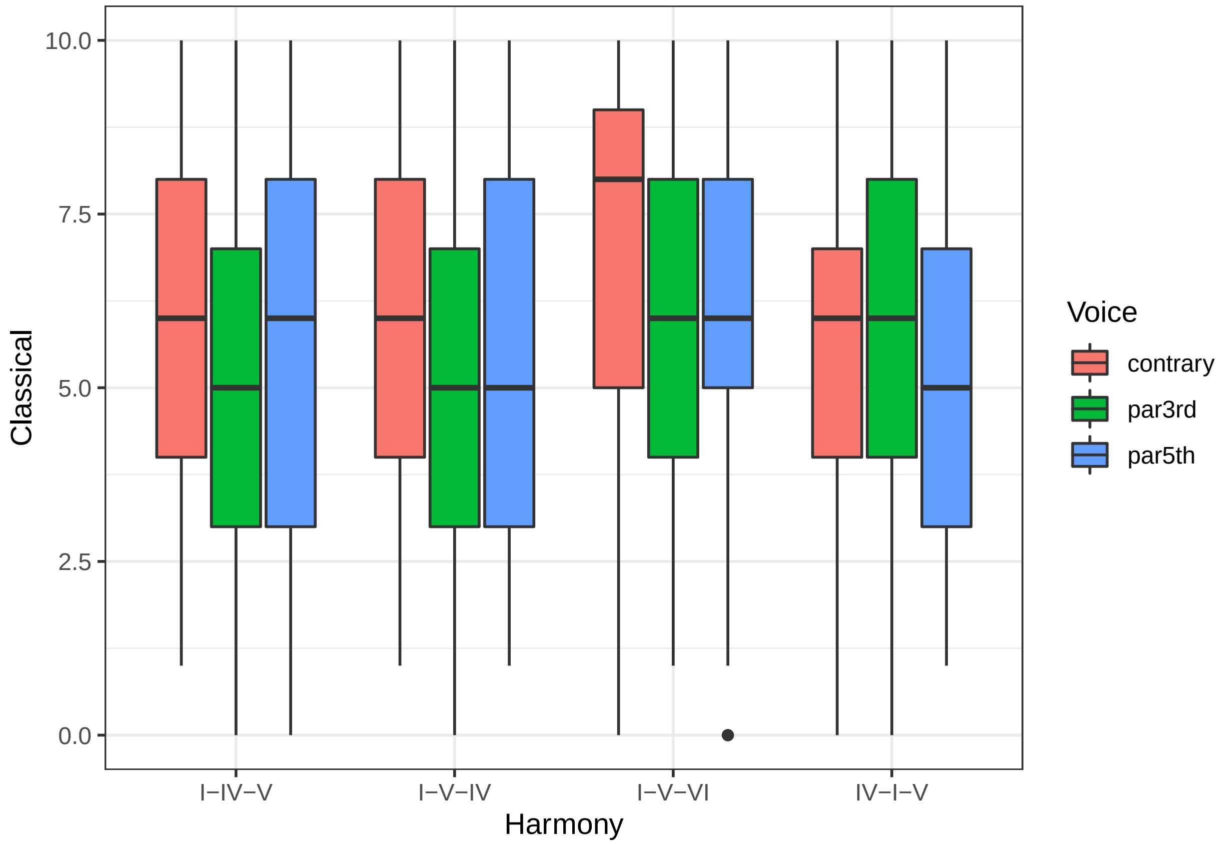

Classical
10.0
7.5
5.0
2.5
0.0
guitar
piano
string
Instrument
I−IV−V
I−V−IV
I−V−VI
IV−I−V
Harmony
contrary
par3rd
par5th
Voice
Declared Musician
No



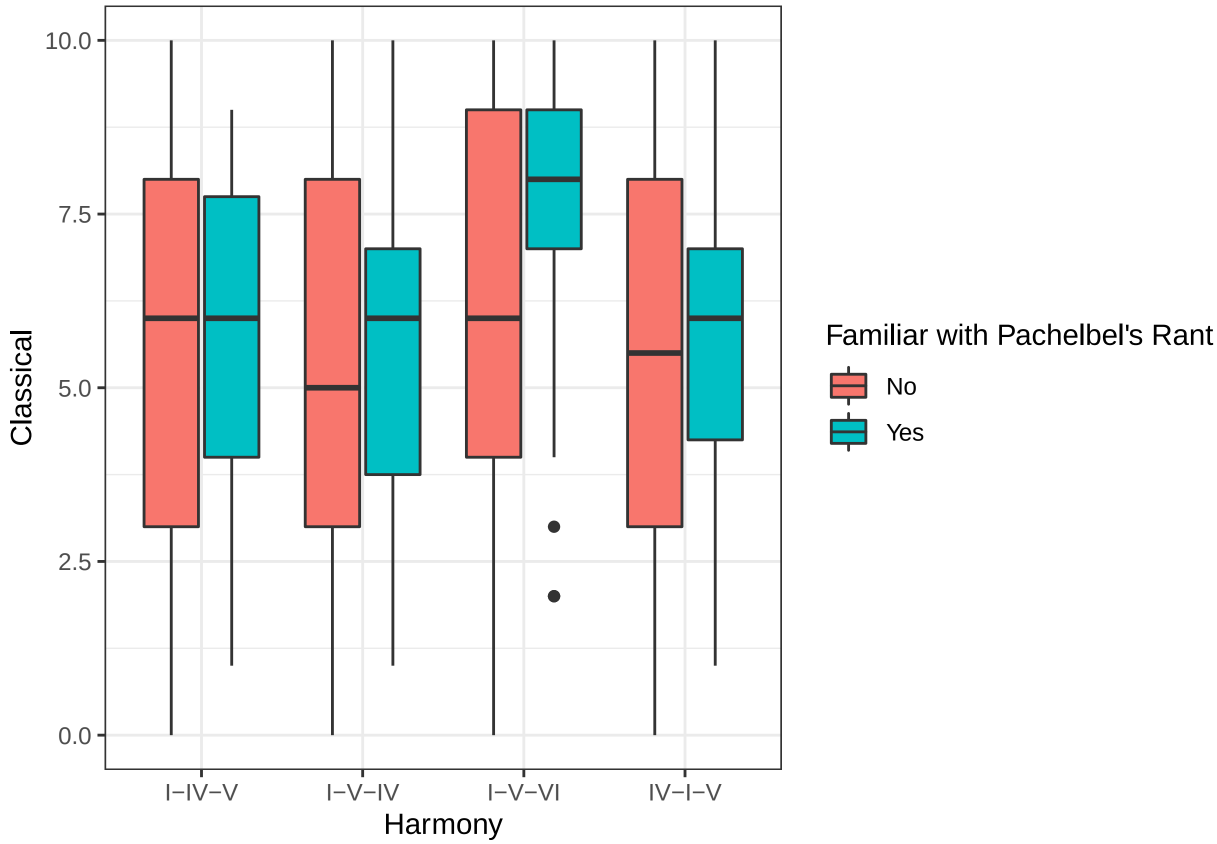
10.0
7.5
5.0
2.5
0.0
Popular
guitar
piano
string
Instrument
I−IV−V
I−V−IV
I−V−VI
IV−I−V
Harmony
contrary
par3rd
par5th
Voice



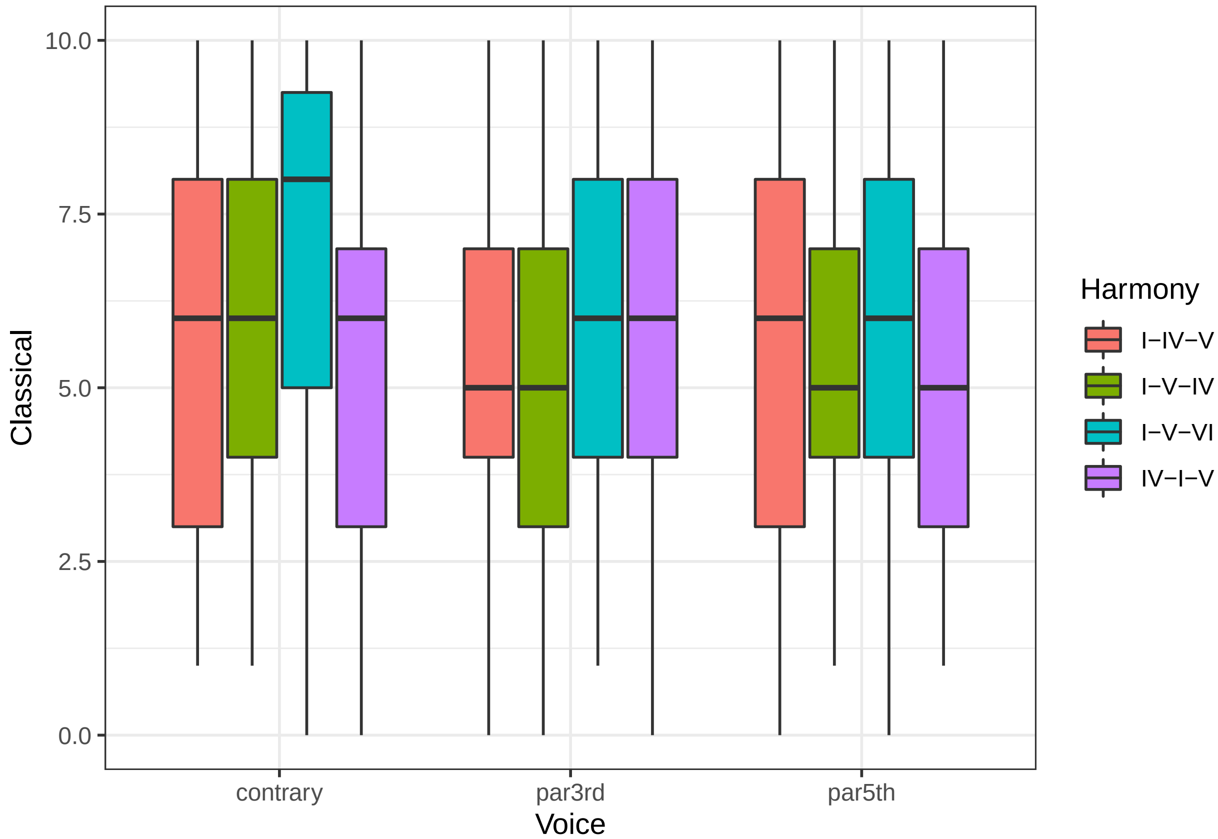
More boxplots to back up classical model



10.0
7.5
5.0
2.5
0.0
Classical
I−IV−V
I−V−IV
I−V−VI
IV−I−V
Harmony
Voice
contrary
par3rd
par5th




```r
ggplot(ratings2, aes(x=Voice, y=Classical, fill=Harmony)) + geom_boxplot() +
  theme_bw() + labs(fill= "Harmony")
```



Classical Model

Note: The following is a simplified version of the steps that were taken to produce the final model.

First, determine best linear model.

```
lm1 <- lm(Classical ~ Instrument*Voice*Harmony, data=ratings2)

step.model <- stepAIC(lm1, direction = "both",
                      trace = FALSE)
step.model[["call"]][["formula"]]
```

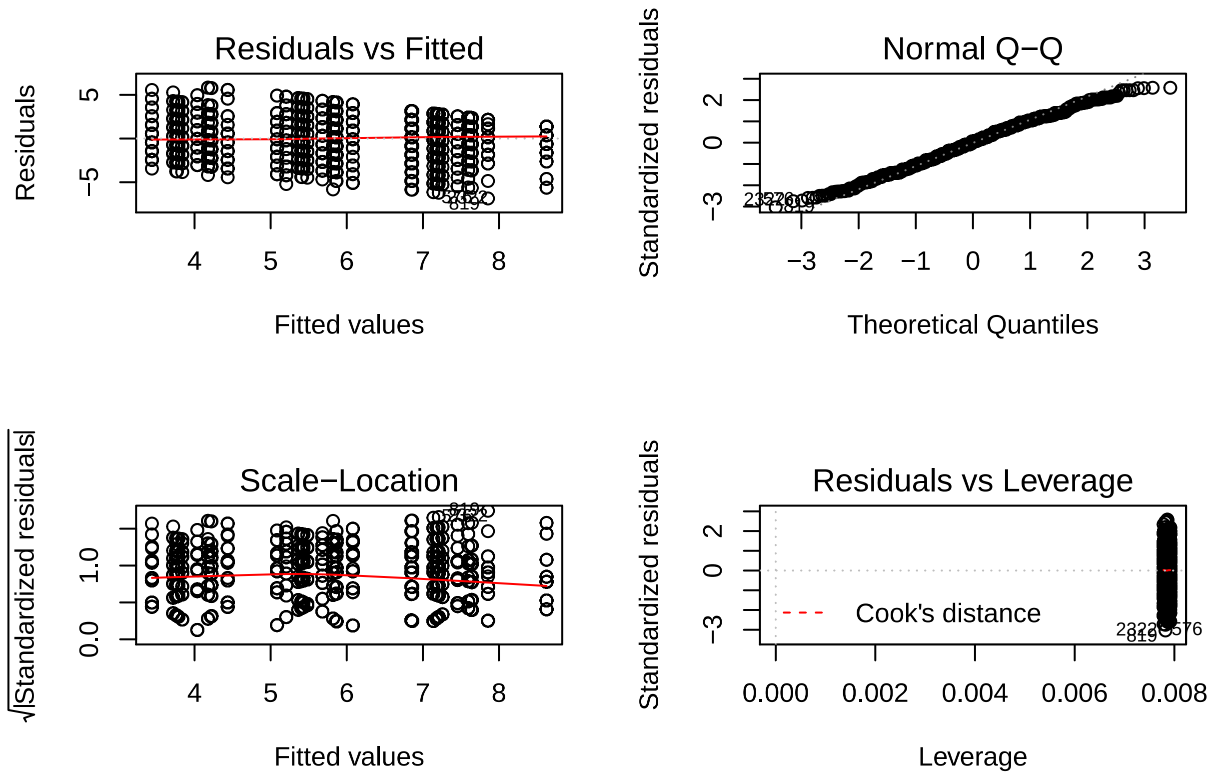
```
## Classical ~ Instrument + Voice + Harmony + Voice:Harmony
```

```
AIC(step.model)
```

```
## [1] 7996.031
```

Plot residuals

```
par(mfrow=c(2,2))
plot(step.model)
```



Determine if random intercept is important

```r
lmer.intercept.only <- lmer(Classical ~ Harmony*Voice + Instrument +
                              (1|Subject), data=ratings2, REML=FALSE,
                            control = lmerControl(optimizer = "bobyqa"))

anova(lmer.intercept.only, step.model)
```

```
## Data: ratings2
## Models:
## step.model: Classical ~ Instrument + Voice + Harmony + Voice:Harmony
## lmer.intercept.only: Classical ~ Harmony * Voice + Instrument + (1 | Subject)
##                     Df    AIC    BIC  logLik deviance  Chisq Chi Df
## step.model          15 7996.0 8078.3 -3983.0   7966.0
## lmer.intercept.only 16 7535.4 7623.2 -3751.7   7503.4 462.64      1
##                     Pr(>Chisq)
## step.model
## lmer.intercept.only  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pr(>Chisq) << 0.05, and AIC much smaller with intercept model, so we will update model.

Now compare with more random effects (note: no other covariates yet).

```
lmer.voice <- lmer(Classical ~Instrument + Harmony*Voice + (1 | Subject) +
                     (0 + Voice|Subject), data=ratings2, REML=FALSE,
                   control = lmerControl(optimizer= "bobyqa"))
```

```
## Warning: Model failed to converge with 1 negative eigenvalue: -1.2e+00
```

```
lmer.instrument <- lmer(Classical ~Instrument + Harmony*Voice + (1 | Subject) +
                          (0 + Instrument|Subject),
                        data=ratings2, REML=FALSE, control = lmerControl(optimizer= "bobyqa"))
```

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

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

```
lmer.voice.instrument <- lmer(Classical ~ Instrument + Harmony*Voice + (1 | Subject) +
                                (0 + Voice | Subject) +
                                (0 + Instrument | Subject), data=ratings2, REML=FALSE,
                              control = lmerControl(optimizer = "bobyqa"))
```

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

```
lmer.voice.harmony <- lmer(Classical ~ Instrument + Harmony*Voice + (1 | Subject) +
                             (0 + Harmony | Subject) +
                             (0 + Voice | Subject), data=ratings2, REML=FALSE,control = lmerControl(opti



lmer.voice.instrument.harmony <- lmer(Classical ~ Instrument + Harmony*Voice + (1 | Subject) +
                                        (0 + Voice | Subject) +
                                      (0 + Instrument | Subject) + (0 + Harmony | Subject), data=rating
                                      control = lmerControl(optimizer = "bobyqa"))
```

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

```
## Warning: Model failed to converge with 4 negative eigenvalues: -3.3e-02
## -1.6e-01 -7.7e+00 -6.6e+01
```

```
anova(lmer.voice, lmer.instrument, lmer.voice.instrument, lmer.voice.harmony, lmer.voice.instrument.harm
```

```
## Data: ratings2
## Models:
## lmer.voice: Classical ~ Instrument + Harmony * Voice + (1 | Subject) + (0 +
```

```
## lmer.voice:     Voice | Subject)
## lmer.instrument: Classical ~ Instrument + Harmony * Voice + (1 | Subject) + (0 +
## lmer.instrument:     Instrument | Subject)
## lmer.voice.instrument: Classical ~ Instrument + Harmony * Voice + (1 | Subject) + (0 +
## lmer.voice.instrument:     Voice | Subject) + (0 + Instrument | Subject)
## lmer.voice.harmony: Classical ~ Instrument + Harmony * Voice + (1 | Subject) + (0 +
## lmer.voice.harmony:     Harmony | Subject) + (0 + Voice | Subject)
## lmer.voice.instrument.harmony: Classical ~ Instrument + Harmony * Voice + (1 | Subject) + (0 +
## lmer.voice.instrument.harmony:     Voice | Subject) + (0 + Instrument | Subject) + (0 + Harmony |
## lmer.voice.instrument.harmony:     Subject)
##                               Df    AIC    BIC  logLik deviance    Chisq
## lmer.voice                    22 7547.2 7667.9 -3751.6   7503.2
## lmer.instrument               22 7239.2 7359.9 -3597.6   7195.2 308.0413
## lmer.voice.instrument         28 7250.1 7403.7 -3597.1   7194.1   1.0547
## lmer.voice.harmony            32 7512.9 7688.5 -3724.5   7448.9   0.0000
## lmer.voice.instrument.harmony 38 7168.7 7377.1 -3546.3   7092.7 356.2545
##                               Chi Df Pr(>Chisq)
## lmer.voice
## lmer.instrument                    0     <2e-16 ***
## lmer.voice.instrument              6     0.9835
## lmer.voice.harmony                 4     1.0000
## lmer.voice.instrument.harmony      6     <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The lmer with the instrument random effect works best (note that some models were not shown because of non-convergence).

Add covariates

```
class(ratings2$Instrument) <- class(ratings2$Harmony) <- class(ratings2$Voice) <- "factor"
ratings2$musician <- ifelse(ratings2$Selfdeclare <= 2, 0, 1)
#ratings2$binaryrob <- ifelse(ratings2$KnowRob == 5, 1, 0)

lmer.harmony.full <- lmer(Classical ~ Harmony*Voice + Instrument + musician*Harmony +
                            musician*Voice + musician*Instrument + Harmony*KnowRob*PachListen + log(OMSI
                       PianoPlay + GuitarPlay + X16.minus.17 + ConsInstr + ConsNotes + ClsListen + X1990
                         CollegeMusic + NoClass + APTheory + Composing
                         + (1 | Subject) + (0 + Harmony | Subject),
                     data=ratings2, REML=FALSE, control = lmerControl(optimizer = "bobyqa"))
```

Stepwise regression to determine best model with covariates

```
# summary(final_fm)
```

Determine if we should change random effects

```
lmer.fe.int <- lmer(Classical ~ Harmony*KnowRob + Voice + Instrument + musician +
                            PianoPlay + X16.minus.17 + ConsNotes + (0 + Harmony | Subject) +
                            Harmony:Voice + Harmony:musician + Instrument:musician +
                         (1 | Subject), data=ratings2, REML=FALSE, control = lmerControl(optimizer = "bo
```

```r
lmer.fe.voice <- lmer(Classical ~ Harmony*KnowRob + Voice + Instrument + musician +
                        PianoPlay + X16.minus.17 + ConsNotes +
                        Harmony:Voice + Harmony:musician + Instrument:musician +
                      (1 | Subject) + (0 + Voice | Subject), data=ratings2, REML=FALSE,
                      control = lmerControl(optimizer = "bobyqa"))


lmer.fe.harm <- lmer(Classical ~ Harmony*KnowRob + Voice + Instrument + musician +
                        PianoPlay + X16.minus.17 + ConsNotes +
                        Harmony:Voice + Harmony:musician + Instrument:musician +
                      (1 | Subject) + (0 + Harmony | Subject), data=ratings2, REML=FALSE,
                      control = lmerControl(optimizer = "bobyqa"))
```

```
## Warning: Model failed to converge with 1 negative eigenvalue: -1.6e+02
```

```r
lmer.fe.voice.harm.inst <- lmer(Classical ~ Harmony*KnowRob + Voice + Instrument + musician +
                        PianoPlay + X16.minus.17 + ConsNotes +
                        Harmony:Voice + Harmony:musician + Instrument:musician +
                      (1 | Subject) + (0 + Harmony | Subject) + (0 + Voice | Subject) +
                        (0 + Instrument | Subject), data=ratings2, REML=FALSE,
                      control = lmerControl(optimizer = "bobyqa"))
```

```r
anova(lmer.fe.int, lmer.fe.voice, lmer.fe.harm, lmer.fe.voice.harm.inst)
```

```
## Data: ratings2
## Models:
## lmer.fe.voice: Classical ~ Harmony * KnowRob + Voice + Instrument + musician +
## lmer.fe.voice:     PianoPlay + X16.minus.17 + ConsNotes + Harmony:Voice + Harmony:musician +
## lmer.fe.voice:     Instrument:musician + (1 | Subject) + (0 + Voice | Subject)
## lmer.fe.int: Classical ~ Harmony * KnowRob + Voice + Instrument + musician +
## lmer.fe.int:     PianoPlay + X16.minus.17 + ConsNotes + (0 + Harmony | Subject) +
## lmer.fe.int:     Harmony:Voice + Harmony:musician + Instrument:musician +
## lmer.fe.int:     (1 | Subject)
## lmer.fe.harm: Classical ~ Harmony * KnowRob + Voice + Instrument + musician +
## lmer.fe.harm:     PianoPlay + X16.minus.17 + ConsNotes + Harmony:Voice + Harmony:musician +
## lmer.fe.harm:     Instrument:musician + (1 | Subject) + (0 + Harmony | Subject)
## lmer.fe.voice.harm.inst: Classical ~ Harmony * KnowRob + Voice + Instrument + musician +
## lmer.fe.voice.harm.inst:     PianoPlay + X16.minus.17 + ConsNotes + Harmony:Voice + Harmony:musician
## lmer.fe.voice.harm.inst:     Instrument:musician + (1 | Subject) + (0 + Harmony | Subject) +
## lmer.fe.voice.harm.inst:     (0 + Voice | Subject) + (0 + Instrument | Subject)
##                         Df    AIC    BIC  logLik deviance   Chisq Chi Df
## lmer.fe.voice           35 7494.1 7686.1 -3712.1   7424.1
## lmer.fe.int             39 7480.5 7694.5 -3701.3   7402.5  21.594      4
## lmer.fe.harm            39 7480.5 7694.5 -3701.3   7402.5   0.000      0
## lmer.fe.voice.harm.inst 51 7158.5 7438.3 -3528.3   7056.5 345.985     12
##                         Pr(>Chisq)
## lmer.fe.voice
## lmer.fe.int              0.0002414 ***
## lmer.fe.harm             1.0000000
## lmer.fe.voice.harm.inst  < 2.2e-16 ***
## ---
```

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

It now appears that we should use voice, harmony, and instrument random effects.

```
lmer.fe.voice.harm.inst
```

```
## Linear mixed model fit by maximum likelihood  ['lmerModLmerTest']
## Formula: Classical ~ Harmony * KnowRob + Voice + Instrument + musician +
##     PianoPlay + X16.minus.17 + ConsNotes + Harmony:Voice + Harmony:musician +
##     Instrument:musician + (1 | Subject) + (0 + Harmony | Subject) +
##     (0 + Voice | Subject) + (0 + Instrument | Subject)
##    Data: ratings2
##       AIC       BIC    logLik  deviance  df.resid
##  7158.531  7438.320 -3528.265  7056.531      1732
## Random effects:
##  Groups    Name             Std.Dev.  Corr
##  Subject   (Intercept)      1.991e-05
##  Subject.1 HarmonyI-IV-V    8.449e-01
##            HarmonyI-V-IV    1.055e+00 0.95
##            HarmonyI-V-VI    1.109e+00 0.54  0.72
##            HarmonyIV-I-V    9.126e-01 0.96  0.89  0.63
##  Subject.2 Voicecontrary    5.597e-01
##            Voicepar3rd      5.946e-01 0.78
##            Voicepar5th      4.535e-01 0.87  0.99
##  Subject.3 Instrumentguitar 9.266e-01
##            Instrumentpiano  1.118e+00  0.17
##            Instrumentstring 9.902e-01 -0.99 -0.02
##  Residual                   1.555e+00
## Number of obs: 1783, groups:  Subject, 50
## Fixed Effects:
##              (Intercept)              HarmonyI-V-IV
##                 4.251337                   0.163046
##            HarmonyI-V-VI              HarmonyIV-I-V
##                 0.456219                  -0.304244
##                  KnowRob                Voicepar3rd
##                 0.062385                  -0.276201
##              Voicepar5th            Instrumentpiano
##                -0.208396                   1.886213
##         Instrumentstring                   musician
##                 3.765677                  -0.281390
##                PianoPlay               X16.minus.17
##                 0.281173                  -0.067127
##                ConsNotes      HarmonyI-V-IV:KnowRob
##                -0.125301                   0.002246
##    HarmonyI-V-VI:KnowRob      HarmonyIV-I-V:KnowRob
##                 0.296622                   0.012698
## HarmonyI-V-IV:Voicepar3rd  HarmonyI-V-VI:Voicepar3rd
##                -0.489231                  -0.702249
## HarmonyIV-I-V:Voicepar3rd  HarmonyI-V-IV:Voicepar5th
##                 0.670296                  -0.229224
## HarmonyI-V-VI:Voicepar5th  HarmonyIV-I-V:Voicepar5th
##                -0.558935                   0.153321
##   HarmonyI-V-IV:musician     HarmonyI-V-VI:musician
##                 0.032438                   1.048662
```

```
##    HarmonyIV-I-V:musician   Instrumentpiano:musician
##                 0.073178                  -0.552774
## Instrumentstring:musician
##                -0.776579
## convergence code 0; 1 optimizer warnings; 0 lme4 warnings
```

Relevel Voice so we can compare contrary

```r
#ratings3 <- within(ratings2, Voice <- relevel(Voice, ref = 2))

final.classical.releveled <- lmer(Classical ~ Voice + Harmony*binaryrob+ Instrument + musician +
                            PianoPlay + X16.minus.17 + ConsNotes +
                            Harmony:Voice + Harmony:musician + Instrument:musician -1 +
                         (1 | Subject) + (0 + Harmony | Subject) + (0 + Voice | Subject) +
                           (0 + Instrument | Subject), data=ratings2, REML=FALSE,
                         control = lmerControl(optimizer = "bobyqa"))

summary(final.classical.releveled)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula:
## Classical ~ Voice + Harmony * binaryrob + Instrument + musician +
##     PianoPlay + X16.minus.17 + ConsNotes + Harmony:Voice + Harmony:musician +
##     Instrument:musician - 1 + (1 | Subject) + (0 + Harmony |
##     Subject) + (0 + Voice | Subject) + (0 + Instrument | Subject)
##    Data: ratings2
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7158.9   7438.7  -3528.5   7056.9     1732
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6199 -0.5791  0.0274  0.5502  3.4722
##
## Random effects:
##  Groups    Name             Variance Std.Dev. Corr
##  Subject   (Intercept)      0.0000   0.0000
##  Subject.1 HarmonyI-IV-V    0.7012   0.8374
##            HarmonyI-V-IV    1.0994   1.0485   0.95
##            HarmonyI-V-VI    1.2252   1.1069   0.55  0.72
##            HarmonyIV-I-V    0.8187   0.9048   0.96  0.89  0.63
##  Subject.2 Voicecontrary    0.3270   0.5718
##            Voicepar3rd      0.3710   0.6091   0.79
##            Voicepar5th      0.2198   0.4688   0.88  0.99
##  Subject.3 Instrumentguitar 0.8665   0.9309
##            Instrumentpiano  1.2835   1.1329    0.19
##            Instrumentstring 0.9769   0.9884   -0.98 -0.01
##  Residual                   2.4174   1.5548
## Number of obs: 1783, groups:  Subject, 50
##
## Fixed effects:
```

```
##                              Estimate Std. Error         df t value
## Voicecontrary                 4.25630    0.37298   75.78880  11.412
## Voicepar3rd                   3.97996    0.37411   73.28222  10.638
## Voicepar5th                   4.04774    0.37009   73.74761  10.937
## HarmonyI-V-IV                 0.15959    0.21461  274.60966   0.744
## HarmonyI-V-VI                 0.48746    0.27394   99.71232   1.779
## HarmonyIV-I-V                -0.30602    0.20957  328.02620  -1.460
## binaryrob                     0.21105    0.48389   50.95506   0.436
## Instrumentpiano               1.88616    0.27702   49.57512   6.809
## Instrumentstring              3.76562    0.38061   49.99122   9.894
## musician                     -0.26233    0.46140   54.25792  -0.569
## PianoPlay                     0.28063    0.09539   50.39704   2.942
## X16.minus.17                 -0.06652    0.05152   50.37173  -1.291
## ConsNotes                    -0.12151    0.08407   50.36919  -1.445
## HarmonyI-V-IV:binaryrob       0.06262    0.34466   88.01617   0.182
## HarmonyI-V-VI:binaryrob       1.49544    0.50601   51.24660   2.955
## HarmonyIV-I-V:binaryrob       0.10677    0.32757   97.13036   0.326
## Voicepar3rd:HarmonyI-V-IV    -0.48903    0.25507 1474.27460  -1.917
## Voicepar5th:HarmonyI-V-IV    -0.22936    0.25526 1473.61199  -0.899
## Voicepar3rd:HarmonyI-V-VI    -0.70261    0.25516 1475.16277  -2.754
## Voicepar5th:HarmonyI-V-VI    -0.55938    0.25539 1474.78989  -2.190
## Voicepar3rd:HarmonyIV-I-V     0.67045    0.25507 1474.64461   2.628
## Voicepar5th:HarmonyIV-I-V     0.15331    0.25508 1474.55014   0.601
## HarmonyI-V-IV:musician        0.02522    0.23910   85.63812   0.105
## HarmonyI-V-VI:musician        1.08171    0.35287   50.79003   3.065
## HarmonyIV-I-V:musician        0.06815    0.22853   96.26374   0.298
## Instrumentpiano:musician     -0.55238    0.41888   50.10386  -1.319
## Instrumentstring:musician    -0.77650    0.57379   49.99350  -1.353
##                            Pr(>|t|)
## Voicecontrary               < 2e-16 ***
## Voicepar3rd                 < 2e-16 ***
## Voicepar5th                 < 2e-16 ***
## HarmonyI-V-IV               0.45773
## HarmonyI-V-VI               0.07821 .
## HarmonyIV-I-V               0.14518
## binaryrob                   0.66456
## Instrumentpiano            1.24e-08 ***
## Instrumentstring           2.30e-13 ***
## musician                    0.57201
## PianoPlay                   0.00492 **
## X16.minus.17                0.20254
## ConsNotes                   0.15455
## HarmonyI-V-IV:binaryrob     0.85625
## HarmonyI-V-VI:binaryrob     0.00471 **
## HarmonyIV-I-V:binaryrob     0.74517
## Voicepar3rd:HarmonyI-V-IV   0.05540 .
## Voicepar5th:HarmonyI-V-IV   0.36904
## Voicepar3rd:HarmonyI-V-VI   0.00597 **
## Voicepar5th:HarmonyI-V-VI   0.02866 *
## Voicepar3rd:HarmonyIV-I-V   0.00867 **
## Voicepar5th:HarmonyIV-I-V   0.54793
## HarmonyI-V-IV:musician      0.91626
## HarmonyI-V-VI:musician      0.00348 **
## HarmonyIV-I-V:musician      0.76619
```

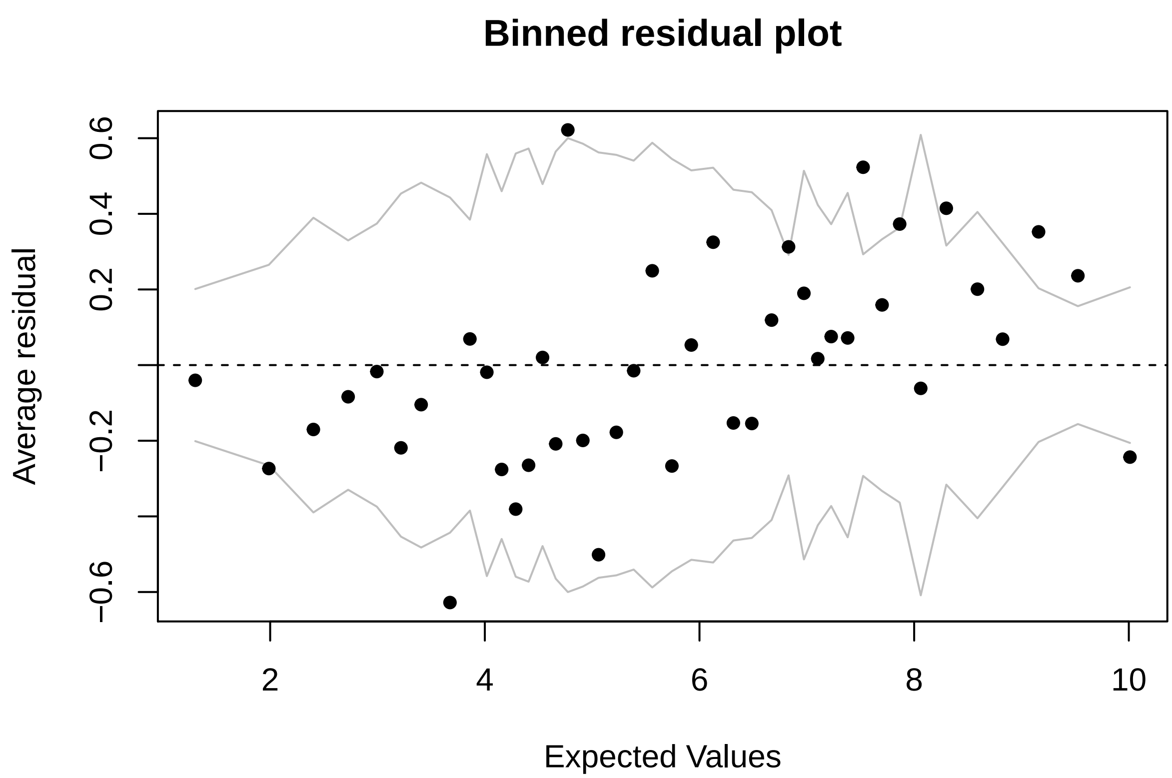
```
## Instrumentpiano:musician   0.19326
## Instrumentstring:musician  0.18205
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

See paper for regression output.

Check errors for Classical Model

We'll start by looking at the binned residuals

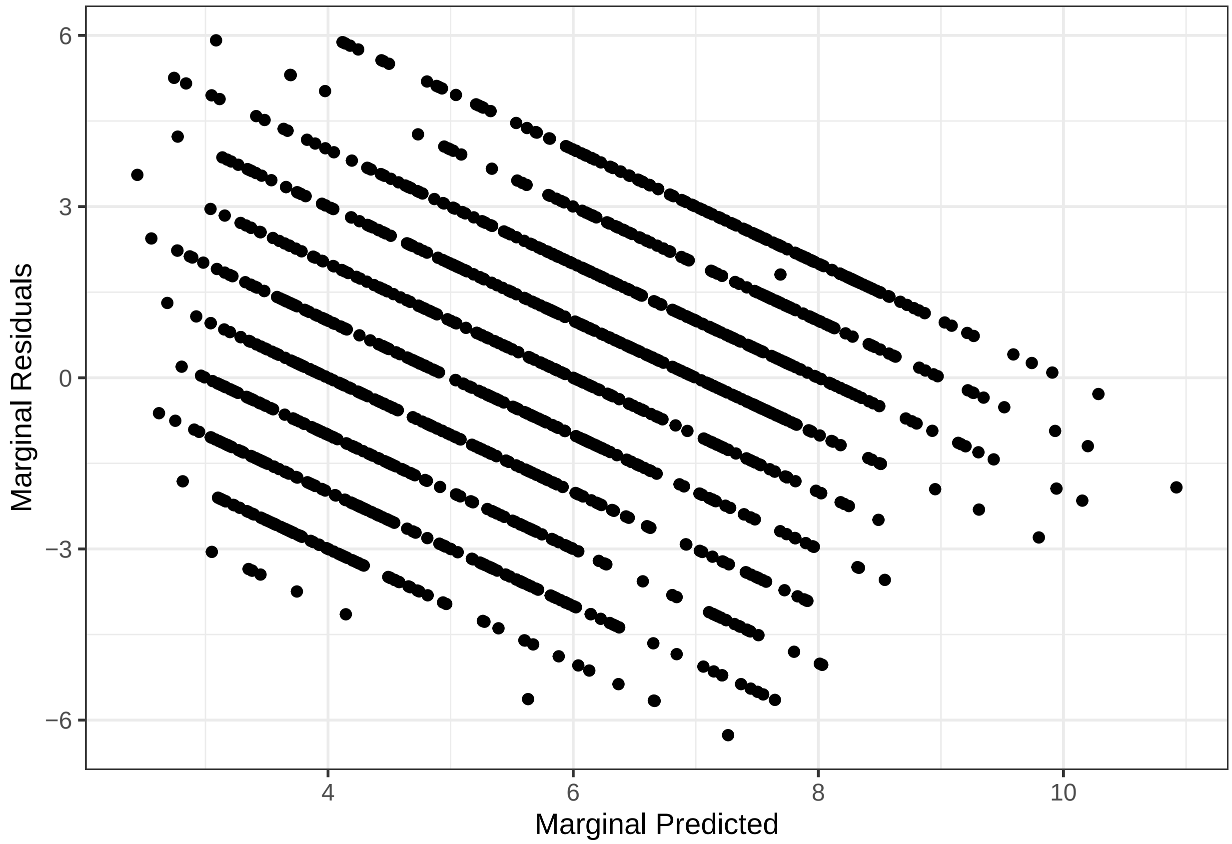
```
modelc <- final.classical.releveled
binnedplot(fitted(modelc),resid(modelc))
```



It appears that the majority of the residuals are within the bin.

Next, we look at marginal fitted values vs. residuals

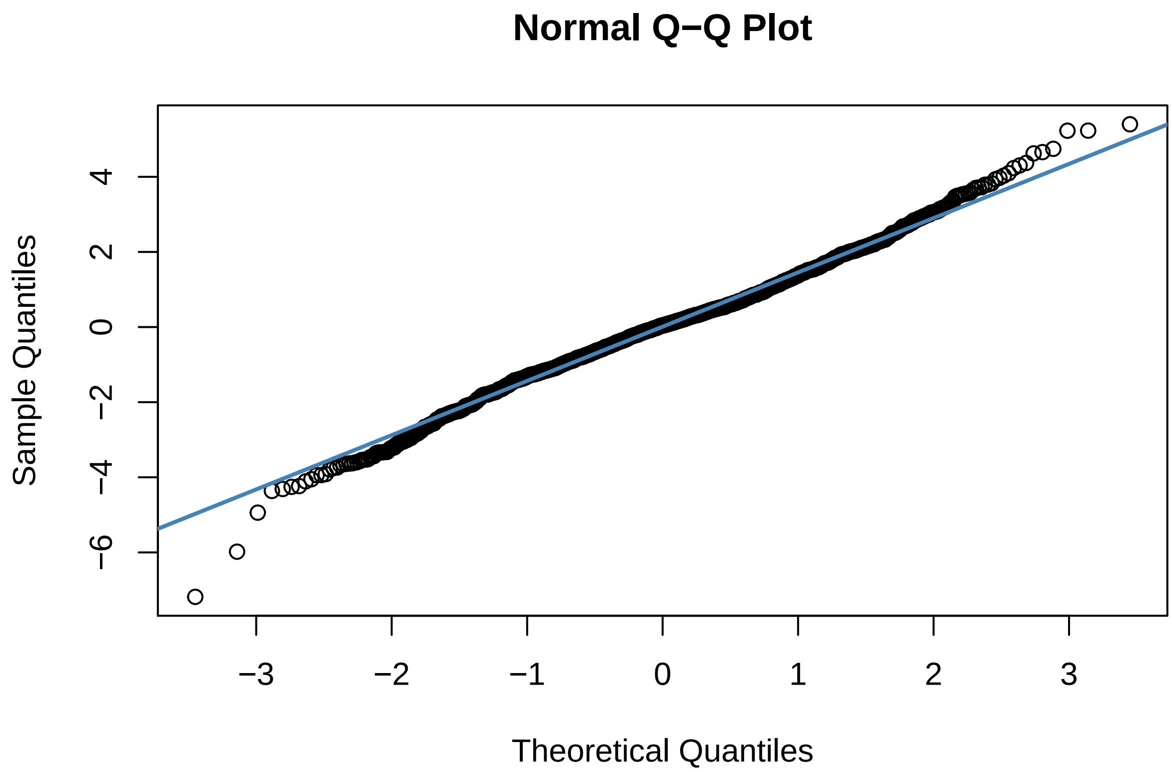
```
ggplot(mapping=aes(yhat.marg(modelc), r.marg(modelc))) + geom_point() +
  labs(x="Marginal Predicted", y="Marginal Residuals") + theme_bw()
```



In the marginal models residuals plot above, we don't care about trends, but are more focused on the spread of the points. Therefore, the marginal residuals plot above looks good.

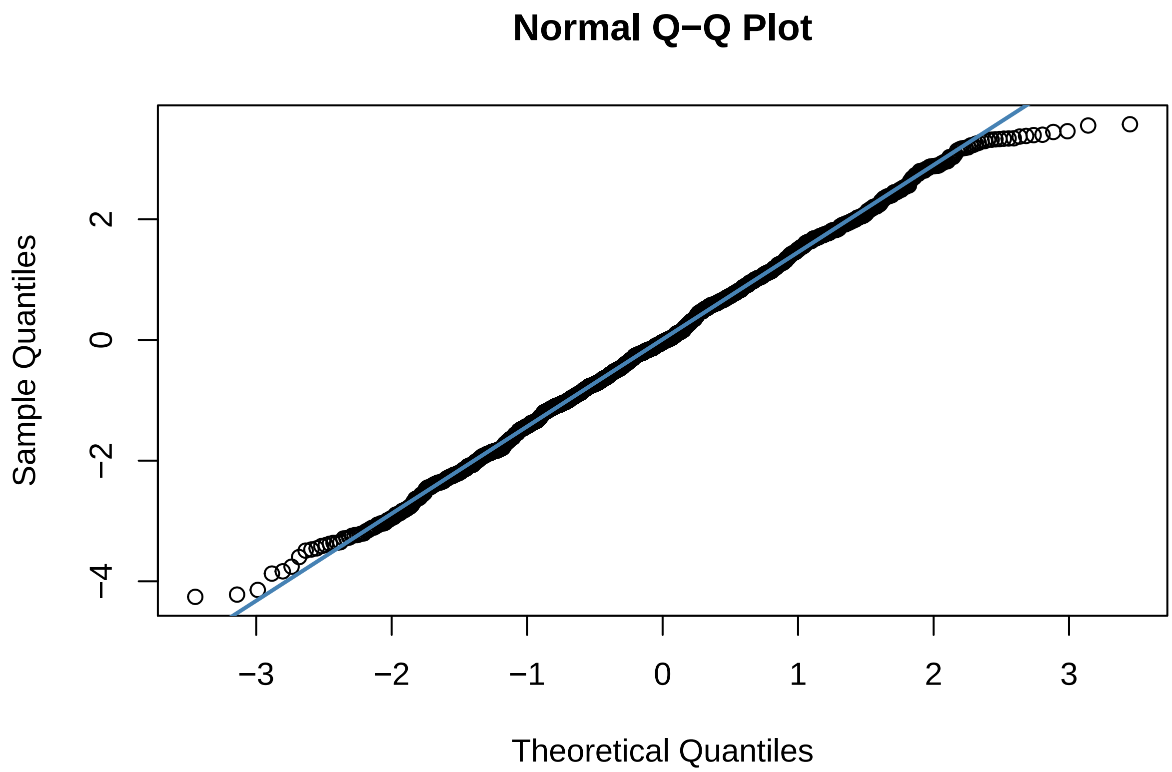
Now we will look at the QQ plot for conditional residuals.

```
qqnorm(r.cond(modelc))
qqline(r.reff(modelc), col = "steelblue", lwd = 2)
```



And the QQ plot for the random effects

```
qqnorm(r.reff(modelc))
qqline(r.reff(modelc), col = "steelblue", lwd = 2)
```



Both QQ plots look linear, suggesting normality of residuals.

We now move to determining the popular ratings model.

Popular Ratings Model

We'll start by running a stepwise regression to determine the optimal fixed effects.

```
lm.pop <- lm(Popular ~ Harmony*Voice + Instrument + musician*Harmony + musician*Voice + musician*Instru
               Harmony*KnowRob*PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 + ConsIns
               ConsNotes + ClsListen + X1990s2000s + CollegeMusic + NoClass + APTheory + Composing, data

step.model <- stepAIC(lm.pop, direction = "both",
                      trace = FALSE)
# summary(step.model)
```

```
lmer.pop.int <- lmer(Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
    PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
    ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
    APTheory + Composing + Harmony:musician + Instrument:musician +
    KnowRob:PachListen + (1 | Subject), data=ratings2,
                    control= lmerControl(optimizer = "bobyqa"), REML = FALSE)

lmer.pop.harm <- lmer(Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
    PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
```

```r
    ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
    APTheory + Composing + Harmony:musician + Instrument:musician +
    KnowRob:PachListen + (1 | Subject) + (0 + Harmony | Subject), data=ratings2,
                     control= lmerControl(optimizer = "bobyqa"), REML = FALSE)
```

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

## Warning: Model failed to converge with 1 negative eigenvalue: -2.9e+00
```

```r
lmer.pop.voice <- lmer(Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
    PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
    ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
    APTheory + Composing + Harmony:musician + Instrument:musician +
    KnowRob:PachListen +  (1 | Subject) + (0 + Voice | Subject), data=ratings2,
                     control= lmerControl(optimizer = "bobyqa"), REML = FALSE)
```

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

```r
lmer.pop.harm.inst <- lmer(Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
    PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
    ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
    APTheory + Composing + Harmony:musician + Instrument:musician +
    KnowRob:PachListen + (1 | Subject) + (0 + Harmony | Subject) +
                       (0 + Instrument | Subject), data=ratings2,
                     control= lmerControl(optimizer = "bobyqa"), REML = FALSE)
```

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

## Warning: Model failed to converge with 2 negative eigenvalues: -1.8e-05
## -8.4e+00
```

```r
lmer.pop.harm.voice <- lmer(Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
    PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
    ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
    APTheory + Composing + Harmony:musician + Instrument:musician +
    KnowRob:PachListen +
                       (0 + Voice | Subject), data=ratings2,
                     control= lmerControl(optimizer = "bobyqa"), REML = FALSE)
```

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

```r
lmer.pop.voice.inst <- lmer(Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
    PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
    ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
    APTheory + Composing + Harmony:musician + Instrument:musician +
    KnowRob:PachListen +
                       (0 + Instrument | Subject), data=ratings2,
                     control= lmerControl(optimizer = "bobyqa"), REML = FALSE)
```

There were a few models that failed to converge, but we can still compare the models that did converge.

```r
anova(lmer.pop.int, lmer.pop.harm, lmer.pop.voice, lmer.pop.harm.inst, lmer.pop.harm.voice, lmer.pop.vo
```

```
## Data: ratings2
## Models:
## lmer.pop.int: Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
## lmer.pop.int:     PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
## lmer.pop.int:     ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
## lmer.pop.int:     APTheory + Composing + Harmony:musician + Instrument:musician +
## lmer.pop.int:     KnowRob:PachListen + (1 | Subject)
## lmer.pop.harm.voice: Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
## lmer.pop.harm.voice:     PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
## lmer.pop.harm.voice:     ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
## lmer.pop.harm.voice:     APTheory + Composing + Harmony:musician + Instrument:musician +
## lmer.pop.harm.voice:     KnowRob:PachListen + (0 + Voice | Subject)
## lmer.pop.voice.inst: Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
## lmer.pop.voice.inst:     PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
## lmer.pop.voice.inst:     ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
## lmer.pop.voice.inst:     APTheory + Composing + Harmony:musician + Instrument:musician +
## lmer.pop.voice.inst:     KnowRob:PachListen + (0 + Instrument | Subject)
## lmer.pop.voice: Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
## lmer.pop.voice:     PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
## lmer.pop.voice:     ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
## lmer.pop.voice:     APTheory + Composing + Harmony:musician + Instrument:musician +
## lmer.pop.voice:     KnowRob:PachListen + (1 | Subject) + (0 + Voice | Subject)
## lmer.pop.harm: Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
## lmer.pop.harm:     PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
## lmer.pop.harm:     ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
## lmer.pop.harm:     APTheory + Composing + Harmony:musician + Instrument:musician +
## lmer.pop.harm:     KnowRob:PachListen + (1 | Subject) + (0 + Harmony | Subject)
## lmer.pop.harm.inst: Popular ~ Harmony + Voice + Instrument + musician + KnowRob +
## lmer.pop.harm.inst:     PachListen + log(OMSI) + PianoPlay + GuitarPlay + X16.minus.17 +
## lmer.pop.harm.inst:     ConsInstr + ConsNotes + X1990s2000s + CollegeMusic + NoClass +
## lmer.pop.harm.inst:     APTheory + Composing + Harmony:musician + Instrument:musician +
## lmer.pop.harm.inst:     KnowRob:PachListen + (1 | Subject) + (0 + Harmony | Subject) +
## lmer.pop.harm.inst:     (0 + Instrument | Subject)
##                     Df    AIC    BIC  logLik deviance   Chisq Chi Df
## lmer.pop.int        30 7525.8 7690.3 -3732.9   7465.8
## lmer.pop.harm.voice 35 7532.6 7724.6 -3731.3   7462.6   3.166      5
## lmer.pop.voice.inst 35 7306.7 7498.7 -3618.4   7236.7 225.883      0
## lmer.pop.voice      36 7534.6 7732.1 -3731.3   7462.6   0.000      1
## lmer.pop.harm       40 7510.9 7730.3 -3715.4   7430.9  31.736      4
## lmer.pop.harm.inst  46 7262.3 7514.6 -3585.1   7170.3 260.592      6
##                     Pr(>Chisq)
## lmer.pop.int
## lmer.pop.harm.voice     0.6744
## lmer.pop.voice.inst  < 2.2e-16 ***
## lmer.pop.voice          1.0000
## lmer.pop.harm        2.166e-06 ***
## lmer.pop.harm.inst   < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The smallest BIC overall appears to be the one where we used voice and instrument random effects. Let's do one more check with fixed effects to see if there are any more variables we should drop.

```
step_fm <- step(lmer.pop.voice.inst)
final_fm <- get_model(step_fm)
# summary(final_fm)
final_fm
```

```
## Linear mixed model fit by maximum likelihood  ['lmerModLmerTest']
## Formula: Popular ~ Harmony + Instrument + musician + (0 + Instrument |
##     Subject) + Harmony:musician
##    Data: ratings2
##       AIC       BIC    logLik  deviance  df.resid
##  7285.833  7379.096 -3625.917  7251.833      1766
## Random effects:
##  Groups   Name             Std.Dev. Corr
##  Subject  Instrumentguitar 1.102
##           Instrumentpiano  1.506    0.54
##           Instrumentstring 1.710    0.22 0.73
##  Residual                  1.705
## Number of obs: 1783, groups:  Subject, 50
## Fixed Effects:
##           (Intercept)           HarmonyI-V-IV           HarmonyI-V-VI
##               6.84581                -0.13095                 0.08099
##          HarmonyIV-I-V         Instrumentpiano        Instrumentstring
##              -0.25397                -1.09293                -2.76959
##              musician  HarmonyI-V-IV:musician  HarmonyI-V-VI:musician
##               0.13023                 0.18816                -0.75599
## HarmonyIV-I-V:musician
##               0.04709
```

The stepwise regression model suggests to only use the variables Harmony, Voice, musician, Instrument, and the interaction Harmony:musician.

Since adding back in Voice and Harmony random effects would cause negative eigenvalues, the final model then is:

```
lmer.pop.final <- lmer(Popular ~  Voice + Harmony*musician + Instrument -1 +
                       (1 | Subject) + (0 + Voice | Subject) + (0 + Instrument | Subject), data=ratings2
                     control= lmerControl(optimizer = "bobyqa"), REML = FALSE)

summary(lmer.pop.final)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: Popular ~ Voice + Harmony * musician + Instrument - 1 + (1 |
##     Subject) + (0 + Voice | Subject) + (0 + Instrument | Subject)
##    Data: ratings2
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   7296.3   7438.9  -3622.2   7244.3     1757
##
## Scaled residuals:
```

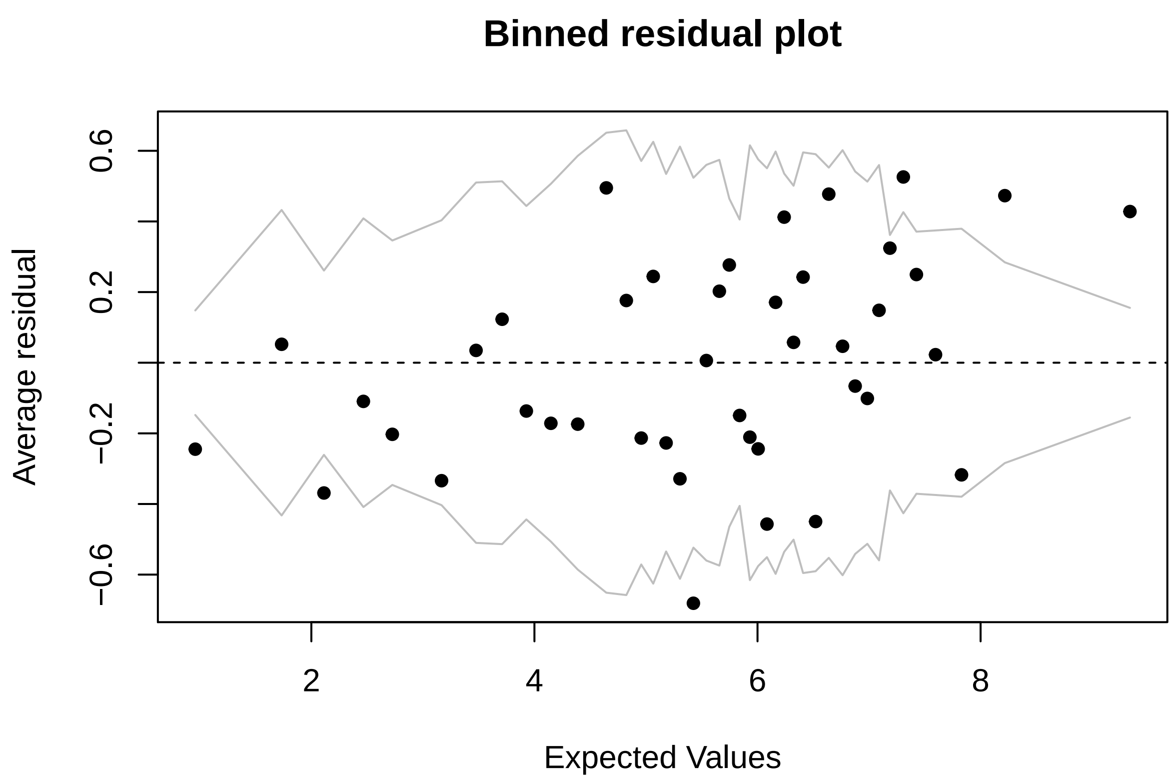
```
##     Min      1Q  Median      3Q     Max 
## -3.5708 -0.6051  0.0168  0.6162  3.2564 
## 
## Random effects:
##  Groups    Name             Variance  Std.Dev.  Corr       
##  Subject   (Intercept)      1.863e-11 4.317e-06            
##  Subject.1 Voicecontrary    1.105e+00 1.051e+00            
##            Voicepar3rd      8.106e-01 9.003e-01 1.00       
##            Voicepar5th      1.099e+00 1.049e+00 0.99  1.00 
##  Subject.2 Instrumentguitar 1.956e-01 4.423e-01            
##            Instrumentpiano  1.359e+00 1.166e+00 -0.13      
##            Instrumentstring 2.008e+00 1.417e+00 -0.87  0.59
##  Residual                   2.889e+00 1.700e+00            
## Number of obs: 1783, groups:  Subject, 50
## 
## Fixed effects:
##                           Estimate Std. Error         df t value Pr(>|t|)
## Voicecontrary            6.776e+00  2.456e-01  8.454e+01  27.591  < 2e-16
## Voicepar3rd              6.941e+00  2.331e-01  7.804e+01  29.773  < 2e-16
## Voicepar5th              6.990e+00  2.453e-01  8.194e+01  28.494  < 2e-16
## HarmonyI-V-IV           -1.310e-01  1.514e-01  1.584e+03  -0.865  0.38727
## HarmonyI-V-VI            8.033e-02  1.516e-01  1.584e+03   0.530  0.59623
## HarmonyIV-I-V           -2.540e-01  1.514e-01  1.584e+03  -1.677  0.09370
## musician                 1.601e-03  3.326e-01  7.419e+01   0.005  0.99617
## Instrumentpiano         -1.093e+00  2.088e-01  4.983e+01  -5.236 3.29e-06
## Instrumentstring        -2.770e+00  2.750e-01  4.996e+01 -10.070 1.28e-13
## HarmonyI-V-IV:musician  1.869e-01  2.296e-01  1.585e+03   0.814  0.41579
## HarmonyI-V-VI:musician -7.550e-01  2.296e-01  1.585e+03  -3.289  0.00103
## HarmonyIV-I-V:musician  4.636e-02  2.294e-01  1.585e+03   0.202  0.83989
##                           
## Voicecontrary          ***
## Voicepar3rd            ***
## Voicepar5th            ***
## HarmonyI-V-IV             
## HarmonyI-V-VI             
## HarmonyIV-I-V          .  
## musician                  
## Instrumentpiano        ***
## Instrumentstring       ***
## HarmonyI-V-IV:musician    
## HarmonyI-V-VI:musician ** 
## HarmonyIV-I-V:musician    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Correlation of Fixed Effects:
##             Vccntr Vcpr3r Vcpr5t HrI-V-IV HrI-V-VI HrIV-I-V musicn
## Voicepar3rd  0.912                                                
## Voicepar5th  0.916  0.912                                         
## HrmnyI-V-IV -0.308 -0.325 -0.309                                  
## HrmnyI-V-VI -0.307 -0.324 -0.308  0.499                           
## HrmnyIV-I-V -0.308 -0.325 -0.309  0.500    0.499                  
## musician    -0.595 -0.627 -0.596  0.228    0.227    0.228         
## Instrumntpn -0.196 -0.206 -0.197  0.000   -0.001    0.000    0.000
```

```
## Instrmntstr -0.292 -0.307 -0.292  0.000    0.000    0.000    0.000
## HrmnI-V-IV:  0.203  0.214  0.202 -0.659   -0.329   -0.330   -0.344
## HrmnI-V-VI:  0.202  0.214  0.202 -0.330   -0.660   -0.330   -0.344
## HrmnIV-I-V:  0.203  0.215  0.202 -0.330   -0.330   -0.660   -0.344
##             Instrmntp Instrmnts HI-V-IV: HI-V-VI:
## Voicepar3rd
## Voicepar5th
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## musician
## Instrumntpn
## Instrmntstr  0.707
## HrmnI-V-IV:  0.001     0.000
## HrmnI-V-VI:  0.001     0.000     0.499
## HrmnIV-I-V:  0.000     0.000     0.499    0.499
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

Check errors for Popular Model

We'll start by looking at the binned residuals

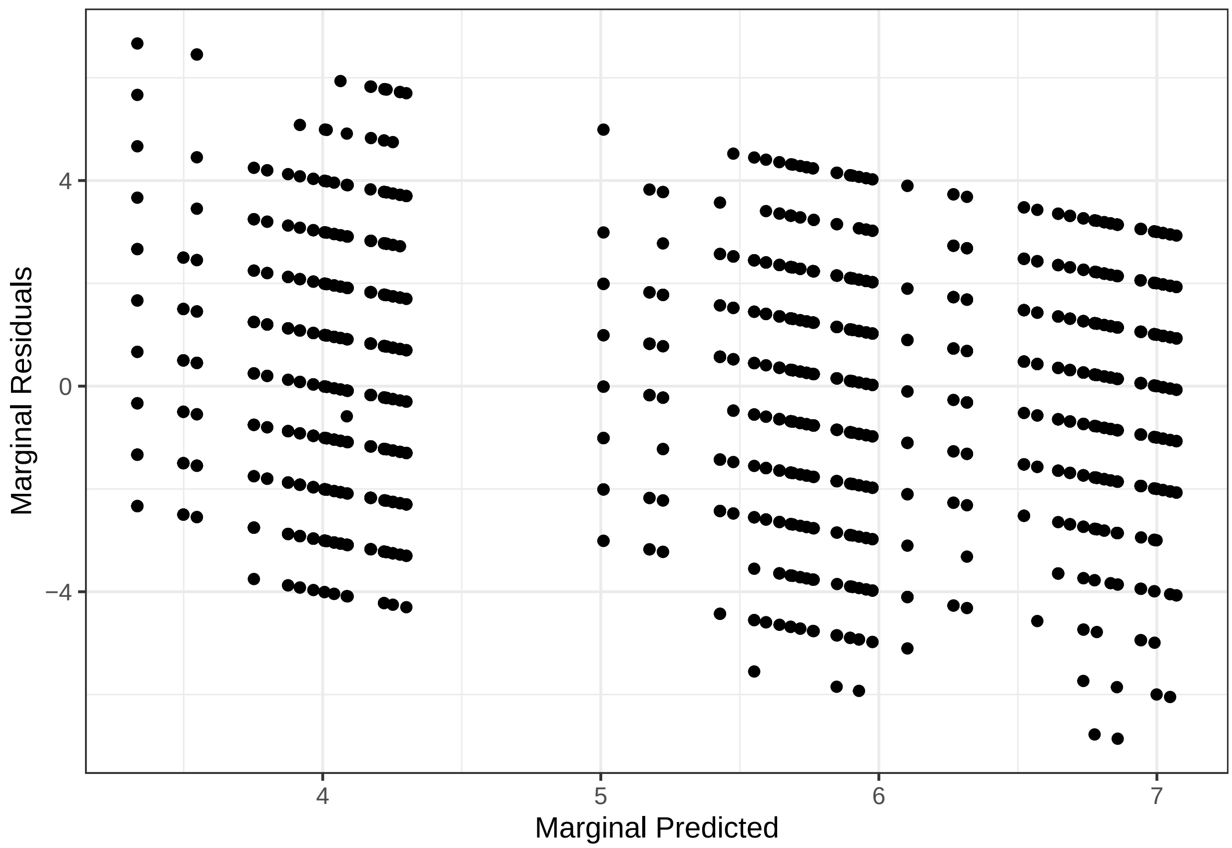
```
modelp <- lmer.pop.final
binnedplot(fitted(modelp),resid(modelp))
```



It appears that the majority of the residuals are within the bin.

Next, we look at marginal fitted values vs. residuals

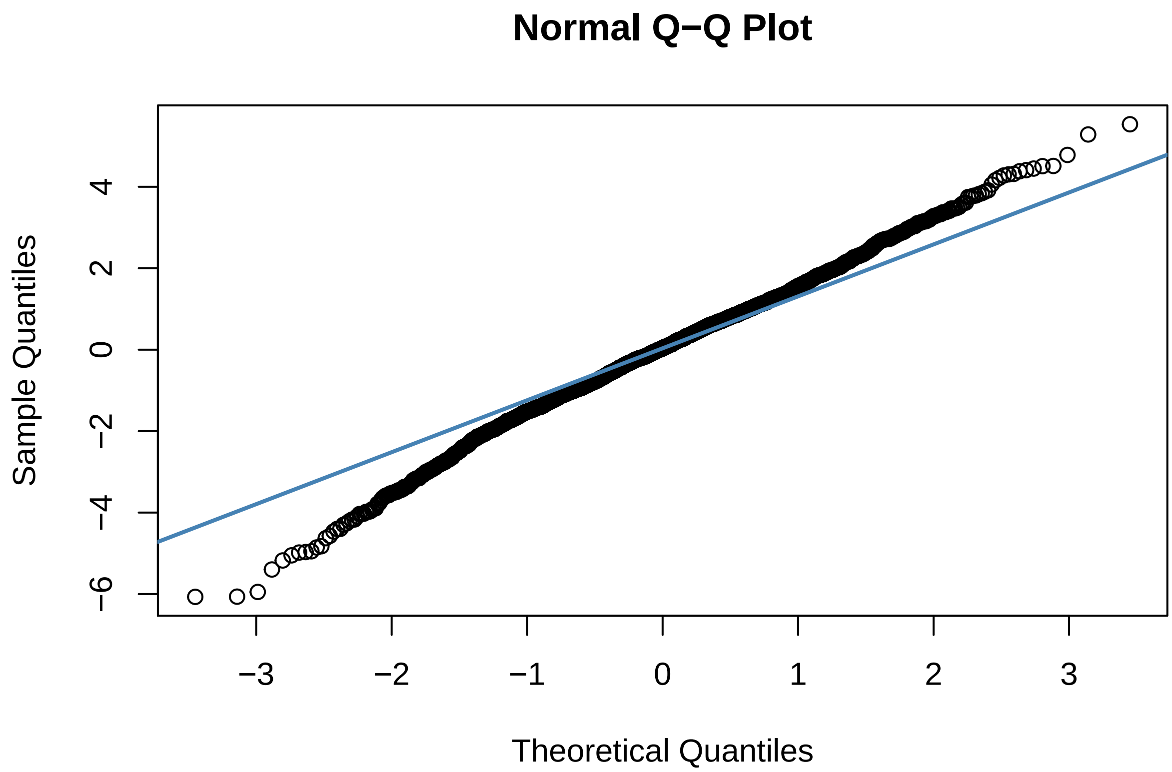
```
ggplot(mapping=aes(yhat.marg(modelp), r.marg(modelp))) + geom_point() +
  labs(x="Marginal Predicted", y="Marginal Residuals") + theme_bw()
```



In the marginal models residuals plot above, we don't care about trends, but are more focused on the spread of the points. Therefore, the marginal residuals plot above looks good.

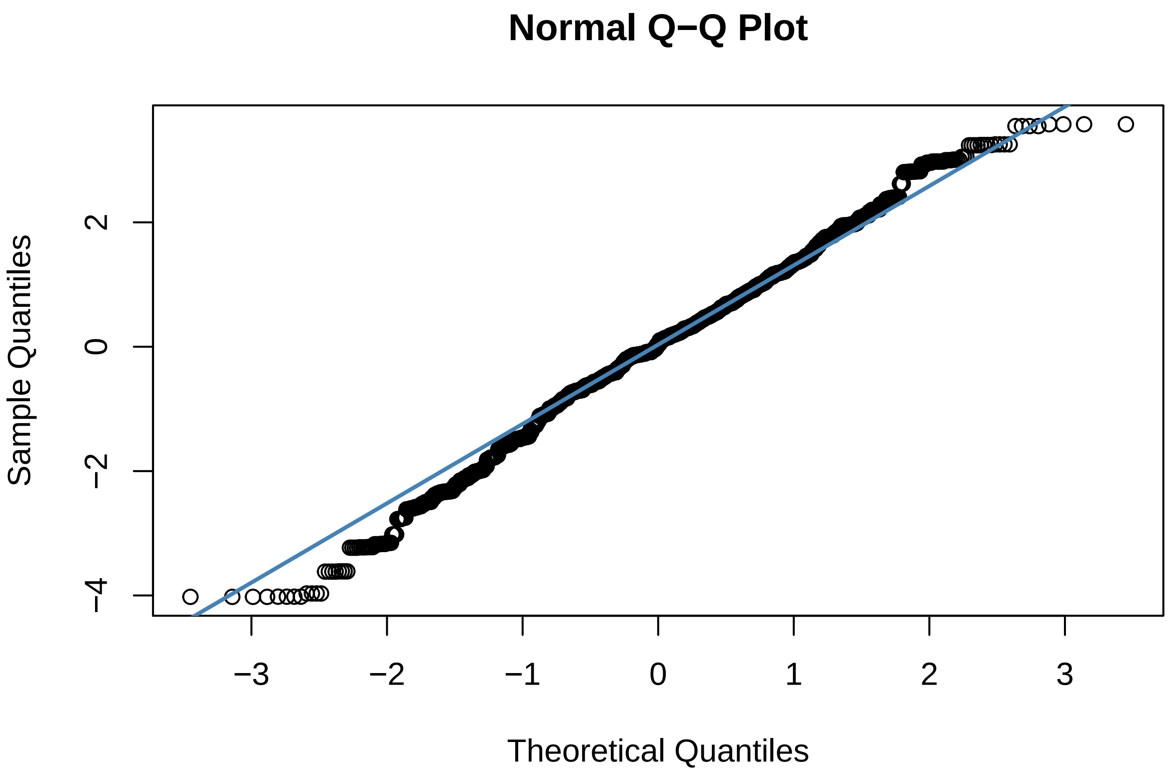
Now we will look at the QQ plot for conditional residuals.

```
qqnorm(r.cond(modelp))
qqline(r.reff(modelp), col = "steelblue", lwd = 2)
```



And the QQ plot for the random effects

```
qqnorm(r.reff(modelp))
qqline(r.reff(modelp), col = "steelblue", lwd = 2)
```



Both QQ plots look linear, suggesting normality of residuals.