Do Listeners Detect Classical or Popular Music?

Minami Makino
mlam@andrew.cmu.edu

2019 December 6

Abstract

We address the question of the extent of the effect of Instrument, Harmonic Motion, and Voice on the listeners' choices of whether the music is classical or popular. We examine data on the choices of listeners at the University of Pittsburgh collected by Jimenez and Rossi (2012). We proceed with our analysis by utilizing the multi-level model to analyze each subject separately in random effects. From our analysis, we found that the experimental factors, Voice Leading, Harmonic Motion, and Instrument, all affect Classical and Popular ratings in different ways. In our analysis, Voice Leading helps listeners detect Popular music while Harmonic Motion and Instrument help listeners detect Classical music. Some suggestions for future studies are to get a wider variety group of subjects for a more unbiased study, and have background information on subjects to understand why they made their choices.

1 Introduction

When listening to music, most listeners do not think deeply about the way music is made and what contributes to what they are listening to. Although we may be listening to our everyday playlist without much thought, much effort goes into the music we listen to, whether that is the instruments or voices. There are many ways to detect music, but in our case, we are interested in the way instrument, harmonic motion, and voice assist listeners in discerning the type of music played. In particular, we are interested in whether listeners can identify more with popular music or classical music when given these factors.

This study explores the depth of which experimental factor, or combination of the three factors, have the strongest influence on ratings of classical and popular music. We are given 27 variables from the original dataset that all contains the information needed to understand how they relate and affect the two ratings. In addition to to answering the main question above, we will address the following questions:

- Does instrument exert the strongest influence among the three design factors (Instrument, Harmonic Motion, Voice Leading), as the researchers suspect?
- Among the levels of Harmonic Motion does I-V-VI have a strong association with classical ratings? Does it seem to matter whether the respondent is familiar with with one or the other (or both) of the Pachelbel rants/comedy bits?
- Among the levels of Voice Leading, does contrary motion have a strong association with classical ratings?
- Are there differences in the way that musicians and non-musicians identify classical music?
- Are there differences in the things that drive classical and popular ratings?

The data was collected by Ivan Jimenez and Vincent Rossi in 2012 at the University of Pittsburgh.

2 Methods

The study was conducted by asking 70 undergraduates from the University of Pittsburgh to listen to 36 musical stimuli each and have them rate how classical and popular the music sounds, both on a scale from 1 to 10 (Variables described later in this section). The two rating scales are said to be independent, as a musical piece can be rated as both popular and classical, neither popular nor classical, or mostly popular and not classical, or vice versa. In our given data, we have 2520 observations as each listener had 36 musical stimuli, and the following are the variables in this study:

$$
\begin{aligned}
Classical &= \text{1 to 10: 1 = not at all, 10 = very classical sounding} \\
Popular &= \text{1 to 10: 1 = not at all, 10 = very popular sounding} \\
Subject &= \text{Unique subject ID} \\
Harmony &= \text{Harmonic motion: I-V-VI, I-VI-V, I-V-IV, IV-I-V} \\
Instrument &= \text{Instrument: String Quartet, Piano, Electric Guitar} \\
VoiceLeading &= \text{Voice leading: Contrary Motion, Parallel 3rds, Parallel 5ths} \\
Selfdeclare &= \text{Are you a musician?: 1-6 (1 = not at all)} \\
OMSI &= \text{Score on a test of musical knowledge} \\
X16.minus.17 &= \text{Auxiliary measure of distinguishing classical vs. popular music} \\
ConsInstr &= \text{Concentration on instrument while listening: 0-5 (0 = not at all)} \\
ConsNotes &= \text{Concentration on notes while listening: 0-5 (0 = not at all)} \\
Instr.minus.Notes &= \text{Difference between previous two variables} \\
PachListen &= \text{Familiarity with Pachelbel's Canon in D: 0-5 (0 = not at all)} \\
ClsListen &= \text{How much do you listen to classical music?: 0-5 (0 = not at all)} \\
KnowRob &= \text{Heard of Rob Paravonian's Pachelbel Rant: 0-5 (0 = not at all)} \\
KnowAxis &= \text{Heard of Axis of Evil's Comedy of 4 Pachelbel chords in popular music: 0-5} \\
X1990s2000s &= \text{Listen to pop and rock from 90's and 2000's?: 0-5 (0 = not at all)} \\
X90s00s.minus.60s70s &= \text{Prev minus variable referring to 60s and 70s} \\
CollegeMusic &= \text{Have you taken music classes in college? (0 = no, 1 = yes)} \\
NoClass &= \text{How many music classes have you taken?} \\
APTheory &= \text{Took AP Music Theory in High School? (0 = no, 1 = yes)} \\
Composing &= \text{Have you done any music composing?: 0-5 (0 = not at all)} \\
PianoPlay &= \text{Do you play piano?: 0-5 (0 = not at all)} \\
GuitarPlay &= \text{Do you play guitar?: 0-5 (0 = not at all)} \\
X1stInstr &= \text{Proficiency at first instrument?: 0-5 (0 = not at all)} \\
X2ndInstr &= \text{Proficiency at second instrument?: 0-5 (0 = not at all)} \\
first12 &= \text{Which instrument was presented to subject in first 12 stimuli?}
\end{aligned}
$$

For our analysis, we relied on multi-level models to best test which factors affect listeners' ratings of music. We use the multi-level models to remove the bias from our study. We want to be able to analyze the data by subject and not clump them together. With multi-level models, we are able to treat each subject separately by adding in random effects.

3 Results

To begin our analysis, we first clean our original data that was given to us. We cleaned the data by removing columns we do not need. In this case, we do not need X and first12, but we also removed X2ndInstr as it was a column full of NAs and felt that would affect our study, thus proceeded to remove it from the dataset. Because we removed X2ndInstr, we removed X1stInstr as we would end up with a small dataset, almost half of the original dataset. Going off of this theory, we also removed any row that contained any NAs to keep the dataset consistent. We also rounded the values in columns in ConsInstr, Classical, and Popular since they should be integers from the ratings. There were also values of 0 and 19 in Popular and Classical that are out of range of ratings, thus we removed those rows from the dataset. We also made CollegeMusic 0 if NoClass 0 since if they have taken music classes in college, then NoClass should not be 0.

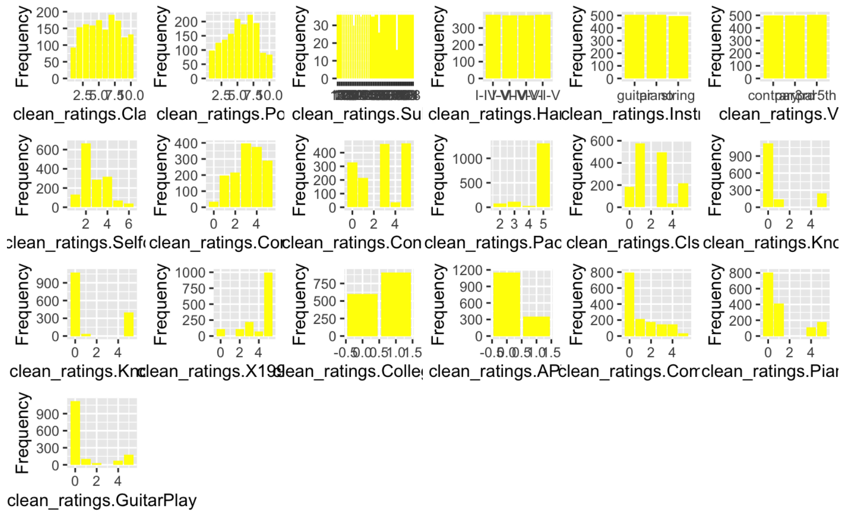

Figure 1: EDA of categorical variables

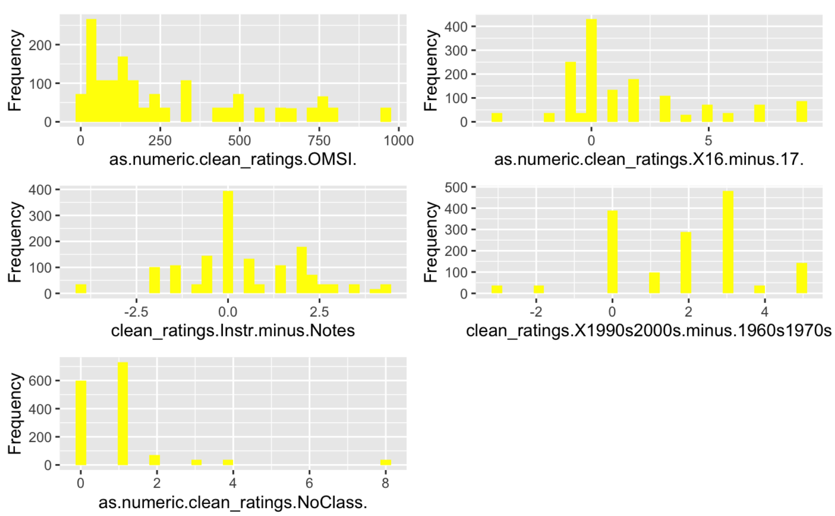

Figure 2: EDA of continuous variables

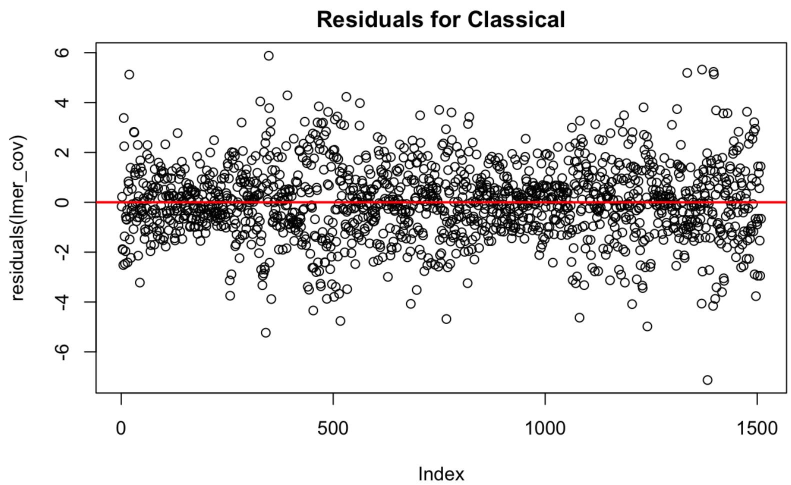

Figure 3: EDA of continuous variables

After in-depth analysis, the following is our final model for Classical ratings:

$$\begin{aligned} Classical = \beta_0 + \beta_1 Harmony + \beta_2 Instrument + \beta_3 Voice + \beta_4 PachListen + \\ \beta_5 ClsListen + \beta_6 APTheory + \beta_7 PianoPlay + \\ (1 + Harmony + Instrument|Subject) + \epsilon \end{aligned} \tag{1}$$

We do the same analysis for Popular ratings and we get:

$$\begin{aligned} Popular = \beta_0 + \beta_1 Harmony + \beta_2 Instrument + \beta_3 Voice + \beta_4 ConsInstr + \\ \beta_5 Instr.minus.Notes + \beta_6 ClsListen + \\ (1 + Harmony + Instrument|Subject) + \epsilon \end{aligned} \tag{2}$$

We perform exploratory data analysis on our cleaned up data and see if there are any further transformations needed for the variables. We plot the categorical variables as bar plots and nothing seemed too out of place. On the other hand, we plot the continuous variables as histograms to see if any variables are heavily skewed. We can transform the OMSI variable with square root to normalize it. We tried to transform the NoClass variable with both square root and log, but it did not make a big difference, thus we proceeded with the analysis without the transformation for that variable.

When we look at the conditional residuals in Figure 3 for the Classical model in equation 1, we see that there is approximately mean 0 with no grouping structure, homoskedastic, and seems relatively normal with no concerning outliers.

Now looking at the conditional residuals in Figure 4 for the Popular model in equation 2, we see the same thing as we did for the Classical model with mean 0 with no grouping structure, is homoskedastic, and has no extreme outliers.

Coefficient interpretations for equation 1 (Can all be referenced to Figure 15 below):

- A unit increase in the level piano of Instrument will result in the Classical rating to increase by 1.60573 points.
- A unit increase in the level string quartet of Instrument will result in the Classical rating to increase by 3.49656 points.
- A unit increase in level 3 of PachListen will result in the Classical rating to decrease by 0.74797 points.

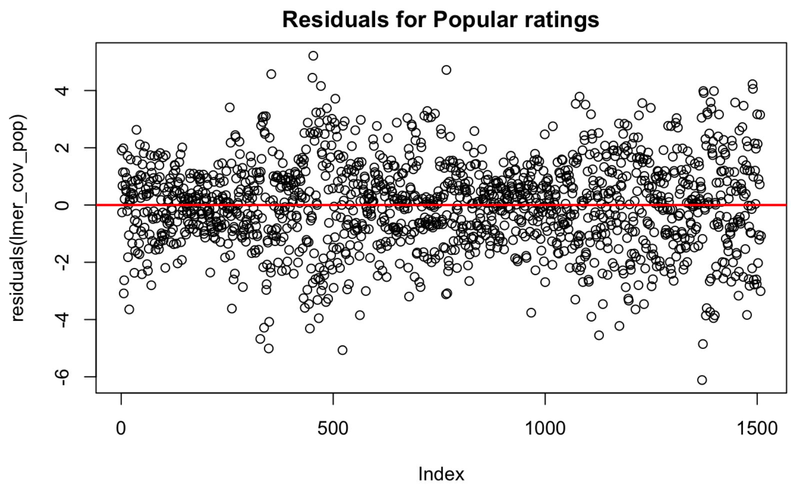

Figure 4: EDA of continuous variables

- A unit increase in level 4 of PachListen will result in the Classical rating to increase by 1.82366 points.
- A unit increase in level 5 of PachListen will result in the Classical rating to increase by 0.04754 points.
- A unit increase in level 1 of ClsListen will result in the Classical rating to decrease by 0.15776 points.
- A unit increase in level 3 of ClsListen will result in the Classical rating to increase by 0.81040 points.
- A unit increase in level 4 of ClsListen will result in the Classical rating to increase by 1.04752 points.
- A unit increase in level 5 of ClsListen will result in the Classical rating to increase by 0.43465 points.
- A unit increase in level 1 of APTheory will result in the Classical rating to increase by 0.67653 points.
- A unit increase in level 1 of PianoPlay will result in the Classical rating to increase by 0.07453 points.
- A unit increase in level 4 of PianoPlay will result in the Classical rating to increase by 1.43100 points.
- A unit increase in level 5 of PianoPlay will result in the Classical rating to increase by 0.35643 points.
- A unit increase in level Par3rd of Voice will result in the Classical rating to decrease by 0.38689 points.
- A unit increase in level Par5th of Voice will result in the Classical rating to decrease by 0.28674 points.
- A unit increase in level I-V-IV of Harmony will result in the Classical rating to decrease by 0.01938 points.
- A unit increase in level I-V-VI of Harmony will result in the Classical rating to increase by 0.86564 points.
- A unit increase in level IV-I-V of Harmony will result in the Classical rating to increase by 0.09424 points.

Coefficient interpretations for equation 2 (Can all be referenced to Figure 16 below):

- A unit increase in level 1 of ConsInstr will result in the Popular rating to increase by 1.06146 points.
- A unit increase in level 2 of ConsInstr will result in the Popular rating to increase by 1.28420 points.

- A unit increase in level 3 of ConsInstr will result in the Popular rating to increase by 0.68373 points.
- A unit increase in level 4 of ConsInstr will result in the Popular rating to increase by 0.93375 points.
- A unit increase in level 5 of ConsInstr will result in the Popular rating to increase by 1.80108 points.
- A unit increase in Instr.minus.Notes will result in the Popular rating to increase by 1.06146 points.
- A unit increase in level 1 of ClsListen will result in the Popular rating to increase by 1.16868 points.
- A unit increase in level 3 of ClsListen will result in the Popular rating to increase by 0.75006 points.
- A unit increase in level 4 of ClsListen will result in the Popular rating to increase by 0.63851 points.
- A unit increase in level 5 of ClsListen will result in the Popular rating to increase by 0.39988 points.
- A unit increase in level Par3rd of Voice will result in the Popular rating to increase by 0.16215 points.
- A unit increase in level Par5th of Voice will result in the Popular rating to increase by 0.23180 points.
- A unit increase in level I-V-IV of Harmony will result in the Popular rating to increase by 0.01855 points.
- A unit increase in level I-V-VI of Harmony will result in the Popular rating to decrease by 0.29207 points.
- A unit increase in level IV-I-V of Harmony will result in the Popular rating to decrease by 0.31316 points.
- A unit increase in level piano of Instrument will result in the Popular rating to decrease by 1.14552 points.
- A unit increase in level string quartet of Instrument will result in the Popular rating to decrease by 2.93279 points.

We will proceed to answer our research questions below.

3.1 Does Instrument exert the strongest influence?

We want to discover if the Instrument factor has the strongest influence on both Classical and Popular ratings. To analyze if Instrument exerts the strongest influence, we look at our summary from the following model with the three experimental factors on Classical ratings:

$$Classical = \beta_0 + \beta_1 Harmony + \beta_2 Instrument + \beta_3 Voice + (1|Subject) + \epsilon \tag{3}$$

We have included the output of the summary of equation 3: Looking at the output in Figure 5 of this model, we can observe that the coefficients for instrument is by far the highest out of all the factors, and with coefficients of 1.60993 and 3.465. The t-values also backs this theory of instrument having the strongest influence on classical ratings since they also have the highest values with 13.449 and 28.822.

To also see if all levels of instruments have the most influence on Classical ratings even in the final model, we used the summary of the following model to observe the effects:

$$\begin{aligned} Classical = \beta_0 + \beta_1 Instrument - 1 + \beta_2 Harmony + \beta_3 Voice + \beta_4 PachListen + \\ \beta_5 ClsListen + \beta_6 APTheory + \beta_7 PianoPlay + \\ (Harmony + Instrument|Subject) + \epsilon \end{aligned} \tag{4}$$

When we look at Figure 6, we can also see here that all levels of instrument have the highest coefficients and t-values over all other variables, with coefficients 3.238 for guitar, 4.84 for piano, 6.73 for string. The

```
Fixed effects:
                          Estimate Std. Error t value
(Intercept)                3.87831    0.22899  16.937
voice_factorpar3rd        -0.38880    0.12001  -3.240
voice_factorpar5th        -0.28853    0.11984  -2.408
instrument_factorpiano     1.60993    0.11971  13.449
instrument_factorstring    3.46500    0.12022  28.822
harmony_factorI-V-IV      -0.02406    0.13858  -0.174
harmony_factorI-V-VI       0.88080    0.13844   6.362
harmony_factorIV-I-V       0.09427    0.13821   0.682
```

Figure 5: Output of equation 3

```
Fixed effects:
                          Estimate Std. Error t value
instrument_factorguitar    3.23833    0.76139   4.253
instrument_factorpiano     4.84406    0.78481   6.172
instrument_factorstring    6.73489    0.78350   8.596
as.factor(PachListen)3    -0.74785    0.78560  -0.952
as.factor(PachListen)4     1.82342    1.22094   1.493
as.factor(PachListen)5     0.04777    0.64025   0.075
as.factor(ClsListen)1     -0.15792    0.46475  -0.340
as.factor(ClsListen)3      0.81053    0.48491   1.672
as.factor(ClsListen)4      1.04769    1.01355   1.034
as.factor(ClsListen)5      0.43447    0.58913   0.737
as.factor(APTheory)1       0.67663    0.35590   1.901
as.factor(PianoPlay)1      0.07448    0.37139   0.201
as.factor(PianoPlay)4      1.43074    0.54495   2.625
as.factor(PianoPlay)5      0.35635    0.46896   0.760
voice_factorpar3rd        -0.38689    0.09848  -3.929
voice_factorpar5th        -0.28674    0.09835  -2.915
harmony_factorI-V-IV      -0.01938    0.12543  -0.155
harmony_factorI-V-VI       0.86564    0.23123   3.744
harmony_factorIV-I-V       0.09424    0.11776   0.800
```

Figure 6: Output of instrument from equation 4

t-values are 4.253, 6.172, and 8.596, respectively. This proves that instrument has the strongest influence over Classical ratings.

We do the same analysis for Popular ratings and the following is our model for Popular:

$$Popular = \beta_0 + \beta_1 Harmony + \beta_2 Instrument + \beta_3 Voice + (1|Subject) + \epsilon \tag{5}$$

We can see from the output for the Popular model that the coefficients and t-values for instrument are the highest out of all three factors. The coefficients are -1.15647 for piano and -2.93727 for string quartet, which are significantly greater than zero with t-values -9.887 and -25.005, respectively.

$$\begin{aligned} Popular = \beta_0 + \beta_1 Instrument - 1 + \beta_2 Harmony + \beta_3 Voice + \beta_4 ConsInstr + \\ \beta_5 Instr.minus.Notes + \beta_6 ClsListen + \\ (Harmony + Instrument|Subject) + \epsilon \end{aligned} \tag{6}$$

We also see that from the final model of Popular, all levels of instrument, string quartet, piano, and electric guitar, are all significant with the highest coefficient estimates and t-values out of all other experimental factors. We can conclude that for both Classical and Popular ratings, Instrument is the most influential out of all experimental factor.

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	6.89567	0.22175	31.097
voice_factorpar3rd	0.16481	0.11727	1.405
voice_factorpar5th	0.23052	0.11710	1.969
instrument_factorpiano	-1.15647	0.11697	-9.887
instrument_factorstring	-2.93727	0.11747	-25.005
harmony_factorI-V-IV	0.01282	0.13540	0.095
harmony_factorI-V-VI	-0.30370	0.13527	-2.245
harmony_factorIV-I-V	-0.31193	0.13504	-2.310

Figure 7: Output of instrument from equation 5

Fixed effects:

	Estimate	Std. Error	t value
instrument_factorguitar	5.03829	0.93172	5.408
instrument_factorpiano	3.89276	0.94873	4.103
instrument_factorstring	2.10549	0.94772	2.222
as.factor(ConsInstr)1	1.06146	0.89711	1.183
as.factor(ConsInstr)2	1.28419	0.87762	1.463
as.factor(ConsInstr)3	0.68373	0.87317	0.783
as.factor(ConsInstr)4	0.93375	0.88715	1.053
as.factor(ConsInstr)5	1.80108	0.87644	2.055
Instr.minus.Notes	0.04763	0.07862	0.606
as.factor(ClsListen)1	1.16868	0.41079	2.845
as.factor(ClsListen)3	0.75006	0.44490	1.686
as.factor(ClsListen)4	0.63851	0.91252	0.700
as.factor(ClsListen)5	0.39988	0.51194	0.781
voice_factorpar3rd	0.16215	0.10079	1.609
voice_factorpar5th	0.23180	0.10064	2.303
harmony_factorI-V-IV	0.01855	0.12756	0.145
harmony_factorI-V-VI	-0.29207	0.17918	-1.630
harmony_factorIV-I-V	-0.31316	0.15097	-2.074

Figure 8: Output of instrument from equation 6

3.2 Does I-V-VI from Harmony have a strong association with Classical ratings? Does it matter if respondent knows Pachelbel?

To determine if I-V-VI from Harmony has a strong association with Classical ratings, we use the model:

$$\begin{aligned} Classical = \beta_0 + \beta_1 Harmony - 1 + \beta_2 Instrument + \beta_3 Voice + \beta_4 PachListen + \\ \beta_5 ClsListen + \beta_6 APTheory + \beta_7 PianoPlay + \\ (Harmony + Instrument | Subject) + \epsilon \end{aligned} \tag{7}$$

From our output from Figure 9, we can see that I-V-VI has a coefficient of 3.878 and a t-value with 4.948, which are higher than all the other coefficients and t-values of the other levels of harmony. The I-V-VI level coefficient estimate is significantly greater than zero and since the t-value is highest out of all levels, the I-V-VI level is the most significant. For Pachelbel, the higher the level, the more familiar the respondent is with Pachelbel rants and comedy bits. If we look at the different levels of PachListen, we can see that the coefficient is the highest for level 4 with a value of 1.04769. This tells us that it does matter if the listener is somewhat familiar with Pachelbel's rants and comedy bits to influence Classical ratings.

```
Fixed effects:
                         Estimate Std. Error t value
harmony_factorI-IV-V      3.01354    0.75944   3.968
harmony_factorI-V-IV      2.99431    0.76914   3.893
harmony_factorI-V-VI      3.87827    0.78375   4.948
harmony_factorIV-I-V      3.10913    0.75766   4.104
as.factor(PachListen)3   -0.74728    0.78578  -0.951
as.factor(PachListen)4    1.83856    1.22164   1.505
as.factor(PachListen)5    0.04562    0.64039   0.071
as.factor(ClsListen)1    -0.15539    0.46491  -0.334
as.factor(ClsListen)3     0.81458    0.48507   1.679
as.factor(ClsListen)4     1.05247    1.01381   1.038
as.factor(ClsListen)5     0.43422    0.58932   0.737
as.factor(APTheory)1      0.67863    0.35598   1.906
as.factor(PianoPlay)1     0.07078    0.37147   0.191
as.factor(PianoPlay)4     1.42895    0.54507   2.622
as.factor(PianoPlay)5     0.35500    0.46906   0.757
instrument_factorpiano    1.60676    0.23595   6.810
instrument_factorstring   3.49485    0.30383  11.503
```

Figure 9: Output of instrument

3.3 Does Contrary Motion from Voice have a strong association with Classical ratings?

Like how we analyzed which level of harmony has the strongest association with Classical ratings, we ran a similar model as equation 7 displayed before:

$$\begin{aligned} Classical = \beta_0 + \beta_1 Voice - 1 + \beta_2 Instrument + \beta_3 Harmony + \beta_4 PachListen + \\ \beta_5 ClsListen + \beta_6 APTheory + \beta_7 PianoPlay + \\ (Harmony + Instrument|Subject) + \epsilon \end{aligned} \tag{8}$$

We observe the effect of voice from Figure 10. When we look at all three levels of voice, we see a very distinct difference between contrary motion and the other two levels. The coefficient estimate of contrary motion is significantly greater than zero as it has the value of 3.23859 and the t-value is 4.254, which means it is the most significant level since it is the highest t-value out of all t-values for the other two levels of voice. We can conclude that contrary motion from voice has a strong association with Classical ratings.

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

We want to know if there are any differences in the way musicians and non-musicians identify classical music. People who self-identify as musicians may be influenced by effects that do not influence non-musicians. We dichotimize the Selfdeclare variable by setting anyone with a level 2 and lower as 0 and anyone with a level 3 and higher as 1 and set the variable as Musician to analyze this hypothesis. While checking which predictors best interact with Musician, we find that the best model is:

$$\begin{aligned} Classical = \beta_0 + \beta_1 Harmony + \beta_2 Instrument + \beta_3 Voice + \\ (1 + Harmony + Instrument|Subject) + \\ (\beta 4 PachListen + \beta 5 ClsListen + \\ \beta 6 APTheory + \beta 7 PianoPlay) * Musician + \epsilon \end{aligned} \tag{9}$$

We check our output in Figure 11, and see that the coefficient estimates and t-values with musicians are

Fixed effects:

	Estimate	Std. Error	t value
voice_factorcontrary	3.23859	0.76131	4.254
voice_factorpar3rd	2.85170	0.76129	3.746
voice_factorpar5th	2.95185	0.76122	3.878
as.factor(PachListen)3	-0.74810	0.78551	-0.952
as.factor(PachListen)4	1.82387	1.22083	1.494
as.factor(PachListen)5	0.04732	0.64017	0.074
as.factor(ClsListen)1	-0.15761	0.46470	-0.339
as.factor(ClsListen)3	0.81030	0.48486	1.671
as.factor(ClsListen)4	1.04744	1.01344	1.034
as.factor(ClsListen)5	0.43480	0.58907	0.738
as.factor(APTheory)1	0.67646	0.35586	1.901
as.factor(PianoPlay)1	0.07460	0.37134	0.201
as.factor(PianoPlay)4	1.43126	0.54488	2.627
as.factor(PianoPlay)5	0.35650	0.46890	0.760
harmony_factorI-V-IV	-0.01938	0.12544	-0.155
harmony_factorI-V-VI	0.86564	0.23124	3.744
harmony_factorIV-I-V	0.09424	0.11776	0.800
instrument_factorpiano	1.60573	0.23570	6.813
instrument_factorstring	3.49656	0.30384	11.508

Figure 10: Output of voice

generally the same as those that are not with musicians. For example, we see that at level 5 of PachListen, it is significant when it interacts with Musician with a t-value of -2.684, and when it does not interact with Musician, it has a t-value of 3.128. However, we have one variable where the variable interacted with Musician is significant and the one without is not significant which is PachListen at level 3. The t-value of PachListen at level 3 interacting with Musician is -2.794 and the t-value without the interaction is 0.776. But overall we can say that there is not much difference in the way musicians and non-musicians identify classical music when we dichotomize the Selfdeclare variable at level 2.

We also want to see if there is a difference between the way musicians and non-musicians identify classical music when we dichotomize at different levels of Selfdeclare. First we dichotomize at level 3 with the same model and we get the output in Figure 12. We can see there is not much difference in the values, but there is one thing that is noticeable and that is the fact that the interaction between PianoPlay and Musician is not included anymore. While we keep this idea in mind, we check the outputs of when we dichotomize Selfdeclare at 4 and 5. As we dichotomize at higher levels, we can see that more and more of our interactions with Musician does not show in our outputs as shown in Figure 13 and Figure 14. This implies that if a person self-claims that he or she is a musician, then it is not going to matter how much the musician listened to Pachelbel, listened to classical music, took AP Music Theory, and played piano, since they should be able to identify classical music without those skills.

3.5 Are there differences in what drives classical and popular ratings?

When comparing the covariates that drive the classical and popular ratings, there are noticeable differences between the two models. When looking at the classical model, we included Instrument, PachListen, ClsListen, APTheory, PianoPlay, Voice, and Harmony as fixed effects, and had Harmony and Instrument as random effects. However, for our popular model, we included ConsInstr, Instr.minus.Notes, ClsListen, Voice, Harmony, and Instrument as our fixed effects, and had Harmony and Instrument as our random effects. As the outputs show in Figure 15 and Figure 16, the coefficient estimates for Instrument, Voice, and Harmony have opposite signs for the different models. In the classical model, we see that Voice has negative coefficients, Instrument has positive coefficients, Harmony I-V-IV has a negative coefficient and the other two levels have positive coefficients. For the popular model, all of the signs are the opposite of what

```
Fixed effects:
                                Estimate Std. Error t value
(Intercept)                      0.77782    0.84760   0.918
harmony_factorI-V-IV            -0.01911    0.12479  -0.153
harmony_factorI-V-VI             0.86626    0.23131   3.745
harmony_factorIV-I-V             0.09462    0.11829   0.800
instrument_factorpiano           1.60729    0.23561   6.822
instrument_factorstring          3.49654    0.30378  11.510
voice_factorpar3rd              -0.38641    0.09851  -3.923
voice_factorpar5th              -0.28597    0.09838  -2.907
as.factor(PachListen)3           0.62540    0.80641   0.776
as.factor(PachListen)4           2.39810    1.03010   2.328
as.factor(PachListen)5           2.40303    0.76820   3.128
as.factor(ClsListen)1           -0.13630    0.40567  -0.336
as.factor(ClsListen)3            1.21502    0.42853   2.835
as.factor(ClsListen)4           -0.33476    0.92716  -0.361
as.factor(ClsListen)5           -0.72927    0.55618  -1.311
as.factor(APTheory)1             0.57371    0.42761   1.342
as.factor(PianoPlay)1            1.80357    0.44653   4.039
as.factor(PianoPlay)4            2.86089    0.52657   5.433
as.factor(PianoPlay)5           -0.60638    0.51971  -1.167
Musician                         5.69345    1.24529   4.572
as.factor(PachListen)3:Musician -4.00172    1.43229  -2.794
as.factor(PachListen)5:Musician -2.86889    1.06896  -2.684
as.factor(ClsListen)1:Musician  -2.25403    0.68361  -3.297
as.factor(ClsListen)3:Musician  -3.11904    0.69971  -4.458
as.factor(APTheory)1:Musician    0.90431    0.60425   1.497
as.factor(PianoPlay)1:Musician  -3.40941    0.69542  -4.903
as.factor(PianoPlay)4:Musician  -2.93308    0.91095  -3.220
```

Figure 11: Output of Musician with cutoff at 2

is listed for those variables in the classical model. In the case of Instrument in the classical model, since the coefficients are positive, if a string quartet, piano, or electric guitar is played in the song, then it will be heard as more classical. We can see with our coefficient estimates that the two models differ greatly.

4 Discussion

While interested in what influences a listener to decide whether a song is classical or popular, we particularly are interested in the way Instrument, Harmonic Motion, and Voice Leading affect a listener's perspective on what genre a song is. In our exploratory analysis, we checked bar plots and histograms for skewness and if there were any variables we should transform. We found that it would be best to transform OMSI with a square root transformation and did not transform NoClass since it did not improve the skewness at all. Overall, there were no alarming variables after the data cleanup in the beginning.

We wanted to understand which experimental factors affected Classical and Popular ratings and went into detail about the influence of the factors. As a summary of our research questions, we found that Instrument exerts the most influence on both Classical and Popular ratings out of all three factors since the coefficient estimates were significantly greater than zero and the t-values were the most significant. The level of I-V-VI in Harmonic Motion has the strongest association with Classical ratings out of all other levels in Harmonic Motion. It does matter if matter whether the respondent is familiar with one or more of Pachelbel's comedy bits and rants as the more familiar one is with it, the higher the Classical ratings. The contrary motion level has the strongest association with Classical ratings compared to the other levels. There is not much difference in the way musicians and non-musicians identify classical music when we set the cut off of Selfdeclare at

```
Fixed effects:
                                    Estimate Std. Error t value
(Intercept)                          3.15078    0.70032   4.499
harmony_factorI-V-IV                -0.01760    0.12853  -0.137
harmony_factorI-V-VI                 0.86752    0.23156   3.746
harmony_factorIV-I-V                 0.09544    0.11997   0.796
instrument_factorpiano               1.60973    0.23554   6.834
instrument_factorstring              3.49660    0.30388  11.507
voice_factorpar3rd                  -0.38601    0.09837  -3.924
voice_factorpar5th                  -0.28636    0.09825  -2.915
as.factor(PachListen)3              -1.32399    0.72656  -1.822
as.factor(PachListen)4               1.84942    1.11821   1.654
as.factor(PachListen)5               0.02887    0.56678   0.051
as.factor(ClsListen)1               -0.08174    0.44149  -0.185
as.factor(ClsListen)3                1.16221    0.46841   2.481
as.factor(ClsListen)4               -0.25768    1.09862  -0.235
as.factor(ClsListen)5               -0.54100    0.63003  -0.859
as.factor(APTheory)1                 0.80549    0.47818   1.684
as.factor(PianoPlay)1                0.47593    0.40810   1.166
as.factor(PianoPlay)4                1.77161    0.47818   3.705
as.factor(PianoPlay)5                1.44065    0.65744   2.191
Musician_3                           0.79949    0.74290   1.076
as.factor(PachListen)3:Musician_3    0.27710    1.28169   0.216
as.factor(ClsListen)1:Musician_3    -2.48033    0.95815  -2.589
as.factor(ClsListen)3:Musician_3    -3.44752    0.84276  -4.091
as.factor(APTheory)1:Musician_3      0.96694    0.80906   1.195
```

Figure 12: Output of Musician with cutoff at 3

level 2. However, as the cut off of the level is set higher, we see that the covariates do not affect the way musicians choose classical music. There are differences that drive the classical and popular ratings such as having different covariates and the signs of the coefficient estimates.

The strength of this study is the fact that there are many different variables to use to determine what best predicts the Classical and Popular ratings. Another strength I can point out is that in our models, we utilize multilevel models which includes random effects. This allows us to treat each subject separately and share information across these subjects to make better estimates. The weakness of this study is that our subjects are undergraduates in the University of Pittsburgh, which can heavily bias our study as they are young and from a certain area, resulting in a certain taste in music.

Some implications from this model are that Voice, Harmony, and Instrument have different effects and influences on the ratings of Classical and Popular music. We were able to see that Voice has a greater influence on listeners choosing Popular music and the other two factors, Harmony and Instrument, were more effective in helping listeners choose Classical music. This is logical as Classical music is heavily reliant on instruments and harmony to make the melody, while Popular music is based on the lyrics and voice to have listeners sing along.

Some unanswered questions are what kinds of subjects picked songs as classical over popular and vice versa, and the reason behind their choices. It would be insightful to see the types of people making certain choices and the reason why they chose it. For future research, it would be helpful to survey a more diverse group of people such as graduate students and faculty to get different age groups. It would also add value to the research by getting the subjects' backgrounds such as their major, area they are from, and what part of the song made them pick their choice of ratings as that would help us understand the mindsets behind these choices.

```
Fixed effects:
                              Estimate Std. Error t value
(Intercept)                    3.26906    0.65786   4.969
harmony_factorI-V-IV          -0.01991    0.12943  -0.154
harmony_factorI-V-VI           0.86653    0.23136   3.745
harmony_factorIV-I-V           0.09453    0.11729   0.806
instrument_factorpiano         1.60679    0.23597   6.809
instrument_factorstring        3.49644    0.30391  11.505
voice_factorpar3rd            -0.38603    0.09829  -3.928
voice_factorpar5th            -0.28613    0.09816  -2.915
as.factor(PachListen)3        -0.78414    0.67118  -1.168
as.factor(PachListen)4         2.17763    1.06237   2.050
as.factor(PachListen)5         0.01517    0.54849   0.028
as.factor(ClsListen)1         -0.11495    0.40133  -0.286
as.factor(ClsListen)3          0.99458    0.42033   2.366
as.factor(ClsListen)4          4.60521    1.20796   3.812
as.factor(ClsListen)5          0.25584    0.50955   0.502
as.factor(APTheory)1           0.38124    0.32120   1.187
as.factor(PianoPlay)1          0.07393    0.32020   0.231
as.factor(PianoPlay)4          1.52865    0.46617   3.279
as.factor(PianoPlay)5          1.53577    0.49338   3.113
Musician_4                    -3.67843    0.88936  -4.136
as.factor(APTheory)1:Musician_4  1.58136  1.06970   1.478
```

Figure 13: Output of Musician with cutoff at 4

5 Technical Appendix

```
library(tidyverse)
library(dplyr)
library(magrittr)
ratings <- read.csv("~/Downloads/ratings.csv")
#clean data

## remove columns with X and first12 , and X2ndInstr since it contains only NAs
## drop all rows that have NAs
## Round for ConsInstr, Classical, and Popular since they should be integers
## Make CollegeMusic 0 if NoClass is 0
clean_ratings <- ratings %>%
  select(-c("X", "X2ndInstr", "X1stInstr", "first12")) %>%
  drop_na() %>%
  mutate_at(vars("ConsInstr", "Classical", "Popular"), funs(round(., 0))) %>%
  filter(Classical != 0) %>%
  filter(Classical != 19) %>%
  filter(Popular != 0) %>%
  filter(Popular != 19) %>%
  mutate(CollegeMusic = ifelse(NoClass == 0, 0, CollegeMusic))


summary(clean_ratings)

#change experimental factors as factors
harmony_factor <- as.factor(clean_ratings$Harmony)
instrument_factor <- as.factor(clean_ratings$Instrument)
```

```
Fixed effects:
                          Estimate Std. Error t value
(Intercept)                3.08719    0.74214   4.160
harmony_factorI-V-IV      -0.01959    0.12446  -0.157
harmony_factorI-V-VI       0.86512    0.23104   3.744
harmony_factorIV-I-V       0.09389    0.11651   0.806
instrument_factorpiano     1.60531    0.23580   6.808
instrument_factorstring    3.49658    0.30391  11.505
voice_factorpar3rd        -0.38695    0.09852  -3.928
voice_factorpar5th        -0.28667    0.09839  -2.914
as.factor(PachListen)3    -0.60753    0.76393  -0.795
as.factor(PachListen)4     1.73652    1.18674   1.463
as.factor(PachListen)5     0.20196    0.62264   0.324
as.factor(ClsListen)1     -0.23708    0.45211  -0.524
as.factor(ClsListen)3      0.89724    0.47162   1.902
as.factor(ClsListen)4      1.21924    0.98558   1.237
as.factor(ClsListen)5      0.34912    0.57412   0.608
as.factor(APTheory)1       0.81277    0.35092   2.316
as.factor(PianoPlay)1      0.08211    0.36128   0.227
as.factor(PianoPlay)4      1.33855    0.52995   2.526
as.factor(PianoPlay)5      0.61477    0.48239   1.274
Musician_5                -1.58418    0.94455  -1.677
```

Figure 14: Output of Musician with cutoff at 5

```
voice_factor <- as.factor(clean_ratings$Voice)
subject_factor <- as.factor(clean_ratings$Subject)
#fit linear regression with interaction between all three
rating_reg1 <- lm(Classical ~ harmony_factor + instrument_factor +
                     voice_factor +
                     (harmony_factor*instrument_factor*voice_factor),
                  data = clean_ratings)
#linear regression as a final model
final_rating_reg1 <- lm(Classical ~ instrument_factor +
                     (harmony_factor*voice_factor),
                  data = clean_ratings)
summary(rating_reg1)
summary(final_rating_reg1)

##Method of finding what should be used as random effect
#harmony
lm_unpooled_contrast_from_mean_harmony <- lm(Classical ~ harmony_factor,
                                             data = clean_ratings)
anova(lm_unpooled_contrast_from_mean_harmony)
s_harmony <-
  sum(coef(summary(lm_unpooled_contrast_from_mean_harmony))[,4] < 0.05)
s_harmony/length(unique(clean_ratings$Harmony))
hist(coef(lm_unpooled_contrast_from_mean_harmony)[-1])

#instrument
lm_unpooled_contrast_from_mean_instrument <-
  lm(Classical ~ instrument_factor, data = clean_ratings)
anova(lm_unpooled_contrast_from_mean_instrument)
```

```
Fixed effects:
                          Estimate Std. Error t value
(Intercept)                3.23847    0.76135   4.254
instrument_factorpiano     1.60573    0.23570   6.812
instrument_factorstring    3.49656    0.30384  11.508
as.factor(PachListen)3    -0.74797    0.78555  -0.952
as.factor(PachListen)4     1.82366    1.22088   1.494
as.factor(PachListen)5     0.04754    0.64021   0.074
as.factor(ClsListen)1     -0.15776    0.46472  -0.339
as.factor(ClsListen)3      0.81040    0.48488   1.671
as.factor(ClsListen)4      1.04752    1.01349   1.034
as.factor(ClsListen)5      0.43465    0.58910   0.738
as.factor(APTheory)1       0.67653    0.35588   1.901
as.factor(PianoPlay)1      0.07453    0.37136   0.201
as.factor(PianoPlay)4      1.43100    0.54491   2.626
as.factor(PianoPlay)5      0.35643    0.46893   0.760
voice_factorpar3rd        -0.38689    0.09848  -3.929
voice_factorpar5th        -0.28674    0.09835  -2.915
harmony_factorI-V-IV      -0.01938    0.12543  -0.155
harmony_factorI-V-VI       0.86564    0.23123   3.744
harmony_factorIV-I-V       0.09424    0.11776   0.800
```

Figure 15: Output of Classical final model

```
s_instrument <-
  sum(coef(summary(lm_unpooled_contrast_from_mean_instrument))[,4] < 0.05)
s_instrument/length(unique(clean_ratings$Instrument))
hist(coef(lm_unpooled_contrast_from_mean_instrument)[-1])

#voice
lm_unpooled_contrast_from_mean_voice <- lm(Classical ~ voice_factor,
                                           data = clean_ratings)
anova(lm_unpooled_contrast_from_mean_voice)
s_voice <- sum(coef(summary(lm_unpooled_contrast_from_mean_voice))[,4] < 0.05)
s_voice/length(unique(clean_ratings$Voice))
hist(coef(lm_unpooled_contrast_from_mean_voice)[-1])

lm_unpooled_contrast_from_mean_subject <- lm(Classical ~ subject_factor,
                                           data = clean_ratings)
anova(lm_unpooled_contrast_from_mean_subject)
s_sub <- sum(coef(summary(lm_unpooled_contrast_from_mean_voice))[,4] < 0.05)
s_sub/length(unique(clean_ratings$Subject))
hist(coef(lm_unpooled_contrast_from_mean_subject)[-1])

#fitting lmer per subject

library(lme4)
lmer_all <- lmer(Classical ~ 1 + voice_factor + instrument_factor +
                    harmony_factor + (1 | subject_factor),
                  data = clean_ratings, REML = FALSE)
summary(lmer_all)

#which factors should be random effects
```

```
Fixed effects:
                            Estimate Std. Error t value
(Intercept)                  5.03829    0.93172   5.408
as.factor(ConsInstr)1        1.06146    0.89711   1.183
as.factor(ConsInstr)2        1.28420    0.87762   1.463
as.factor(ConsInstr)3        0.68373    0.87317   0.783
as.factor(ConsInstr)4        0.93375    0.88715   1.053
as.factor(ConsInstr)5        1.80108    0.87644   2.055
Instr.minus.Notes            0.04763    0.07862   0.606
as.factor(ClsListen)1        1.16868    0.41079   2.845
as.factor(ClsListen)3        0.75006    0.44490   1.686
as.factor(ClsListen)4        0.63851    0.91252   0.700
as.factor(ClsListen)5        0.39988    0.51194   0.781
voice_factorpar3rd           0.16215    0.10079   1.609
voice_factorpar5th           0.23180    0.10064   2.303
harmony_factorI-V-IV         0.01855    0.12756   0.145
harmony_factorI-V-VI        -0.29207    0.17918  -1.630
harmony_factorIV-I-V        -0.31316    0.15097  -2.074
instrument_factorpiano      -1.14552    0.22578  -5.074
instrument_factorstring     -2.93279    0.26737 -10.969
```

Figure 16: Output of Popular final model

```
lmer_best <- lmer(Classical ~ 1 + voice_factor + harmony_factor +
                    instrument_factor +
                    (harmony_factor + instrument_factor + 1 | subject_factor),
                  data = clean_ratings, REML = FALSE,
                  control = lmerControl(optimizer = 'bobyqa'))
summary(lmer_best)
anova(lmer_all, lmer_best)

#best model
lmer_cov <- lmer(Classical ~ 1 + instrument_factor + as.factor(PachListen) +
                   as.factor(ClsListen) + as.factor(APTheory) +
                   as.factor(PianoPlay) + voice_factor + harmony_factor +
                   (harmony_factor + instrument_factor + 1 | subject_factor),
                 data = clean_ratings, REML = FALSE,
                 control = lmerControl(optimizer = 'bobyqa'))
summary(lmer_cov)


#look at all levels of instrument
lmer_instr <- lmer(Classical ~ instrument_factor - 1 + as.factor(PachListen) +
                   as.factor(ClsListen) + as.factor(APTheory) +
                   as.factor(PianoPlay) + voice_factor + harmony_factor +
                   (harmony_factor + instrument_factor | subject_factor),
                 data = clean_ratings, REML = FALSE,
                 control = lmerControl(optimizer = 'bobyqa'))

summary(lmer_instr)

#look at all levels of voice
lmer_voice <- lmer(Classical ~ voice_factor - 1 + as.factor(PachListen) +
```

```
                 as.factor(ClsListen) + as.factor(APTheory) +
                 as.factor(PianoPlay) + harmony_factor + instrument_factor +
                 (harmony_factor + instrument_factor | subject_factor),
               data = clean_ratings, REML = FALSE,
               control = lmerControl(optimizer = 'bobyqa'))
summary(lmer_voice)

#look at all levels of harmony
lmer_harmony <- lmer(Classical ~ harmony_factor - 1 + as.factor(PachListen) +
                 as.factor(ClsListen) + as.factor(APTheory) +
                 as.factor(PianoPlay) + instrument_factor +
                 (harmony_factor + instrument_factor | subject_factor),
               data = clean_ratings, REML = FALSE,
               control = lmerControl(optimizer = 'bobyqa'))
summary(lmer_harmony)

#plot residuals for best model of classical
plot(residuals(lmer_cov), main = "Residuals for Classical")
abline(h=0, col = "red", lwd = 2)

#dichotomize musician with cutoff at 2
clean_ratings$Musician <- ifelse(clean_ratings$Selfdeclare <= 2, 0, 1)
#cutoff at 3
clean_ratings$Musician_3 <- ifelse(clean_ratings$Selfdeclare <= 3, 0, 1)
#cutoff at 4
clean_ratings$Musician_4 <- ifelse(clean_ratings$Selfdeclare <= 4, 0, 1)
#cutoff at 5
clean_ratings$Musician_5 <- ifelse(clean_ratings$Selfdeclare <= 5, 0, 1)
#check that cutoff at 2 is about half
length(which(clean_ratings$Musician == 1))/length(clean_ratings$Musician)

#model with musician as interaction
classical_musi <- lmer(Classical ~ 1 + harmony_factor + instrument_factor + voice_factor +
                       (harmony_factor +
                          instrument_factor + 1 | subject_factor) +
                       (as.factor(PachListen) +
                          as.factor(ClsListen) + as.factor(APTheory) +
                          as.factor(PianoPlay))*Musician,
                     data = clean_ratings, REML = FALSE,
                     control = lmerControl(optimizer = 'bobyqa'))
summary(classical_musi)

#model with cutoff 3
musi_3 <- lmer(Classical ~ 1 + harmony_factor + instrument_factor + voice_factor +
                       (harmony_factor +
                          instrument_factor + 1 | subject_factor) +
                       (as.factor(PachListen) +
                          as.factor(ClsListen) + as.factor(APTheory) +
                          as.factor(PianoPlay))*Musician_3,
                     data = clean_ratings, REML = FALSE,
                     control = lmerControl(optimizer = 'bobyqa'))
```

```
summary(musi_3)

#model with cutoff 4
musi_4 <- lmer(Classical ~ 1 + harmony_factor + instrument_factor + voice_factor +
                          (harmony_factor +
                             instrument_factor + 1 | subject_factor) +
                          (as.factor(PachListen) +
                             as.factor(ClsListen) + as.factor(APTheory) +
                             as.factor(PianoPlay))*Musician_4,
                        data = clean_ratings, REML = FALSE,
                        control = lmerControl(optimizer = 'bobyqa'))
summary(musi_4)

#model with cutoff 5
musi_5 <- lmer(Classical ~ 1 + harmony_factor + instrument_factor + voice_factor +
                          (harmony_factor +
                             instrument_factor + 1 | subject_factor) +
                          (as.factor(PachListen) +
                             as.factor(ClsListen) + as.factor(APTheory) +
                             as.factor(PianoPlay))*Musician_5,
                        data = clean_ratings, REML = FALSE,
                        control = lmerControl(optimizer = 'bobyqa'))
summary(musi_5)

#model of pop with each subject
lmer_pop <- lmer(Popular ~ 1 + voice_factor + instrument_factor +
                    harmony_factor + (1 | subject_factor),
                  data = clean_ratings, REML = FALSE)
summary(lmer_pop)

#which should be random
lmer_best_pop <- lmer(Popular ~ 1 + voice_factor + harmony_factor +
                         instrument_factor +
                         (harmony_factor +
                            instrument_factor + 1 | subject_factor),
                       data = clean_ratings, REML = FALSE,
                       control = lmerControl(optimizer = 'bobyqa'))
summary(lmer_best_pop)
anova(lmer_pop, lmer_best_pop)

#best model
lmer_cov_pop <- lmer(Popular ~ 1 + as.factor(ConsInstr) +
                        Instr.minus.Notes + as.factor(ClsListen) +
                        voice_factor + harmony_factor + instrument_factor +
                        (harmony_factor + instrument_factor + 1 | subject_factor),
                      data = clean_ratings, REML = FALSE,
                      control = lmerControl(optimizer = 'bobyqa'))
summary(lmer_cov_pop)

#see all instrument levels
lmer_cov_pop_instr <- lmer(Popular ~ instrument_factor - 1 + as.factor(ConsInstr) +
```

```
                           Instr.minus.Notes + as.factor(ClsListen) +
                           voice_factor + harmony_factor + instrument_factor +
                           (harmony_factor + instrument_factor | subject_factor),
                        data = clean_ratings, REML = FALSE,
                        control = lmerControl(optimizer = 'bobyqa'))
summary(lmer_cov_pop_instr)
#plot residuals of popular
plot(residuals(lmer_cov_pop), main = "Residuals for Popular ratings")
abline(h=0, col = "red", lwd = 2)

#popular with musician interaction
popular_musi <- lmer(Popular ~ 1 + harmony_factor + instrument_factor +
                           (harmony_factor +
                              instrument_factor + 1 | subject_factor) +
                           (as.factor(ConsInstr) +
                              Instr.minus.Notes + as.factor(ClsListen) +
                              voice_factor)*Musician, data = clean_ratings,
                        REML = FALSE, control = lmerControl(optimizer = 'bobyqa'))
summary(popular_musi)

#eda
library(gridExtra)
library(grid)
continuous_ratings <- data.frame(as.numeric(clean_ratings$OMSI), as.numeric(clean_ratings$X
continuous_eda <- list()
for (i in names(continuous_ratings)) {
  continuous_eda[[i]] = ggplot(continuous_ratings) + geom_histogram(aes_string(i), fill = "
}
do.call("grid.arrange", c(continuous_eda, ncol = 2))

hist(sqrt(as.numeric(clean_ratings$OMSI)))
hist(sqrt(as.numeric(clean_ratings$NoClass)))


cate_ratings <- data.frame(clean_ratings$Classical, clean_ratings$Popular, clean_ratings$Su
cate_eda <- list()
for (i in names(cate_ratings)) {
  cate_eda[[i]] = ggplot(cate_ratings) + geom_bar(aes_string(i), fill = "yellow") + labs(y
}
do.call("grid.arrange", c(cate_eda, nrow = 4, ncol = 6))
```

References

Jimenex, I. and Rossi, V. (2012). Ratings.csv. University of Pittsburgh. https://canvas.cmu.edu/courses/11853/files/folder/hw10

Jimenez, I. and Rossi, V. (2013). *The Influence of Timbra, Harmony, and Voice Leading on Listener's Distinction between Popular and Classical Music*. University of Pittsburgh. https://canvas.cmu.edu/courses/11853/files/folder/hw10