Hierarchical Final Project

Ruicong(Ricky) She

December 10, 2013

Exercise 1

(a)

```
> music = read.csv("ratings.csv", header = T)
> attach(music)
> lm1=lm(Classical~Instrument+Harmony+Voice)
> lm2=lm(Classical~Instrument+Voice)
> lm3=lm(Classical~Instrument+Harmony)
> lm4=lm(Classical~Harmony+Voice)
> summary(lm1)

Call:
lm(formula = Classical ~ Instrument + Harmony + Voice)

Residuals:
    Min      1Q  Median      3Q     Max
-6.8718 -1.7137 -0.0297  1.7576 11.4766

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       4.34016    0.12987  33.420  < 2e-16 ***
Instrumentpiano   1.37359    0.11298  12.158  < 2e-16 ***
Instrumentstring  3.13312    0.11230  27.899  < 2e-16 ***
HarmonyI-V-IV    -0.03108    0.13008  -0.239 0.811168
HarmonyI-V-VI     0.76909    0.13008   5.913 3.83e-09 ***
HarmonyIV-I-V     0.05007    0.12997   0.385 0.700092
Voicepar3rd      -0.41247    0.11271  -3.660 0.000258 ***
Voicepar5th      -0.37058    0.11264  -3.290 0.001016 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.297 on 2485 degrees of freedom
  (27 observations deleted due to missingness)
Multiple R-squared:  0.255,        Adjusted R-squared:  0.2529
F-statistic: 121.5 on 7 and 2485 DF,  p-value: < 2.2e-16

> summary(lm2)
```

```
Call:
lm(formula = Classical ~ Instrument + Voice)

Residuals:
    Min      1Q  Median      3Q     Max
-6.3011 -1.5407 -0.1233  1.7433 11.3299

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)        4.5367     0.1038  43.724  < 2e-16 ***
Instrumentpiano    1.3730     0.1141  12.035  < 2e-16 ***
Instrumentstring   3.1334     0.1134  27.631  < 2e-16 ***
Voicepar3rd       -0.4134     0.1138  -3.633 0.000286 ***
Voicepar5th       -0.3690     0.1137  -3.244 0.001193 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.319 on 2488 degrees of freedom
  (27 observations deleted due to missingness)
Multiple R-squared:  0.2395,	Adjusted R-squared:  0.2383
F-statistic: 195.9 on 4 and 2488 DF,  p-value: < 2.2e-16


> summary(lm3)


Call:
lm(formula = Classical ~ Instrument + Harmony)

Residuals:
    Min      1Q  Median      3Q     Max
-6.9812 -1.8483 -0.0797  1.7874 11.7370

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       4.07966    0.11280  36.166  < 2e-16 ***
Instrumentpiano   1.37327    0.11330  12.120  < 2e-16 ***
Instrumentstring  3.13294    0.11262  27.818  < 2e-16 ***
HarmonyI-V-IV    -0.03234    0.13045  -0.248    0.804
HarmonyI-V-VI     0.76863    0.13045   5.892 4.33e-09 ***
HarmonyIV-I-V     0.05045    0.13034   0.387    0.699
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.303 on 2487 degrees of freedom
  (27 observations deleted due to missingness)
Multiple R-squared:  0.2502,	Adjusted R-squared:  0.2487
F-statistic: 165.9 on 5 and 2487 DF,  p-value: < 2.2e-16


> summary(lm4)


Call:
lm(formula = Classical ~ Harmony + Voice)
```

```
Residuals:
    Min      1Q  Median      3Q     Max 
-6.6170 -2.2463  0.1549  2.1549 13.1024 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)    5.84507    0.12896  45.326  < 2e-16 ***
HarmonyI-V-IV -0.02824    0.14910  -0.189  0.84979    
HarmonyI-V-VI  0.77194    0.14910   5.177 2.43e-07 ***
HarmonyIV-I-V  0.05249    0.14898   0.352  0.72461    
Voicepar3rd   -0.41065    0.12919  -3.179  0.00150 ** 
Voicepar5th   -0.37075    0.12911  -2.872  0.00412 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.632 on 2487 degrees of freedom
  (27 observations deleted due to missingness)
Multiple R-squared:  0.02044,	Adjusted R-squared:  0.01847 
F-statistic: 10.38 on 5 and 2487 DF,  p-value: 7.201e-10
```

```
> rbind(AIC(lm1,lm2,lm3,lm4))
```

```
    df      AIC
lm1  9 11230.45
lm2  6 11275.96
lm3  7 11242.69
lm4  7 11908.94
```

```
> rbind(BIC(lm1,lm2,lm3,lm4))
```

```
    df      BIC
lm1  9 11282.84
lm2  6 11310.89
lm3  7 11283.43
lm4  7 11949.69
```

From the results of AIC, BIC comparison, we found the full model with all the three main factors is the best (has smallest value).
From the 'summary' results of the models, we found Harmony I-V-IV and Harmony IV-I-V have very small coefficients, indicating they are not significantly different from Harmony I-IV-V.

(b) i

$$Classical_i = \alpha_{0j[i]} + \alpha_1 Instrument_i + \alpha_2 Harmony_i + \alpha_3 Voice_i + \epsilon_i, \epsilon_i \overset{i.i.d}{\sim} N(0, \sigma^2)$$

$$\alpha_{0j} = \beta_0 + \eta_j, \eta_j \overset{i.i.d}{\sim} N(0, \tau^2)$$

ii

Method 1: AIC, BIC Comparison

```
> lmer1b2 = lmer(Classical~ Instrument + Harmony + Voice + (1|Subject))
> rbind(AIC(lm1), AIC(lmer1b2))
```

```
         [,1]
[1,] 11230.45
[2,] 10491.51
```

```
> rbind(BIC(lm1), BIC(lmer1b2))
```

```
         [,1]
[1,] 11282.84
[2,] 10549.73
```

The model with random intercept has smaller values in both AIC and BIC compared with the model without it. It shows that we should keep the random intercept.

Method 2: LRT

```
> exactRLRT(lmer1b2)
```

```
        simulated finite sample distribution of RLRT.

        (p-value based on 10000 simulated values)

data:
RLRT = 763.3759, p-value < 2.2e-16
```

P-value is much less than 0.05, so we strongly reject $H_0 : \tau^2 = 0$ and keep the random intercept.

iii

```
> lmer1b31 = lmer(Classical~ Instrument + Voice + (1|Subject))
> lmer1b32 = lmer(Classical~ Instrument + Harmony + (1|Subject))
> lmer1b33 = lmer(Classical~ Harmony + Voice + (1|Subject))
> rbind(AIC(lmer1b2,lmer1b31,lmer1b32,lmer1b33))
```

```
         df      AIC
lmer1b2  10 10491.51
lmer1b31  7 10552.74
lmer1b32  8 10505.58
lmer1b33  8 11423.04
```

```
> rbind(BIC(lmer1b2,lmer1b31,lmer1b32,lmer1b33))
```

```
         df      BIC
lmer1b2  10 10549.73
lmer1b31  7 10593.49
lmer1b32  8 10552.15
lmer1b33  8 11469.60
```

As we can see from the results above, the model 1b2 has the smallest AIC and BIC. So the full model with all the three main factors (Instrument, Harmony and Voice) is the best.

(c) i

```
> lmer1c=lmer(Classical~Instrument+Harmony+Voice+(1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:V
> rbind(AIC(lmer1c),AIC(lmer1b2),AIC(lm1))
```

```
         [,1]
[1,] 10075.51
[2,] 10491.51
[3,] 11230.45
```

```
> rbind(BIC(lmer1c),BIC(lmer1b2),BIC(lm1))
```

```
         [,1]
[1,] 10145.37
[2,] 10549.73
[3,] 11282.84
```

As is shown in the results, the new model with all three new random effects has the smallest AIC and BIC. So it is the best.

ii

```
> lmer1c1=update(lmer1c,.~. -Instrument)
> lmer1c2=update(lmer1c,.~. -Harmony)
> lmer1c3=update(lmer1c,.~. -Voice1)
> rbind(AIC(lmer1c,lmer1c1,lmer1c2,lmer1c3))
```

```
        df      AIC
lmer1c  12 10075.51
lmer1c1 10 10176.17
lmer1c2  9 10101.74
lmer1c3 12 10075.51
```

```
> rbind(BIC(lmer1c,lmer1c1,lmer1c2,lmer1c3))
```

```
        df      BIC
lmer1c  12 10145.37
lmer1c1 10 10234.38
lmer1c2  9 10154.13
lmer1c3 12 10145.37
```

```
> summary(lmer1c)

Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +      (1 | Subject:Harmony

REML criterion at convergence: 10051.51

Random effects:
 Groups             Name        Variance Std.Dev.
 Subject:Harmony    (Intercept) 0.44307  0.6656
 Subject:Voice      (Intercept) 0.02809  0.1676
 Subject:Instrument (Intercept) 2.19850  1.4827
 Residual                       2.43753  1.5613
Number of obs: 2493, groups: Subject:Harmony, 280; Subject:Voice, 210; Subject:Instrument, 210

Fixed effects:
                 Estimate Std. Error t value
(Intercept)       4.34106    0.21435  20.252
Instrumentpiano   1.36384    0.26232   5.199
Instrumentstring  3.12836    0.26203  11.939
HarmonyI-V-IV    -0.03023    0.14317  -0.211
HarmonyI-V-VI     0.77063    0.14316   5.383
HarmonyIV-I-V     0.05618    0.14310   0.393
Voicepar3rd      -0.40699    0.08174  -4.979
Voicepar5th      -0.37084    0.08168  -4.540

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
Instrumntpn -0.611
Instrmntstr -0.611  0.500
HrmnyI-V-IV -0.333  0.000     0.000
HrmnyI-V-VI -0.333  0.000     0.000     0.499
HrmnyIV-I-V -0.333  0.000     0.000     0.500  0.500
Voicepar3rd -0.190 -0.001     0.000    -0.002  0.001  0.002
Voicepar5th -0.190 -0.001     0.000    -0.001 -0.002 -0.001  0.500
```

From the results of AIC and BIC comparison, it shows the best model is the model with all three main factors as it has the smallest value.
The variance of Subject/Voice combination is the smallest (variance=0.17); the variance of the Subject/Harmony combination is the second smallest (0.67); the variance of Subject/Instrument is the largest (1.48). They are all smaller than the variance of the residual (1.56).

iii

$$Classical_i = \alpha_{0j[i]} + \alpha_{0k[i]} + \alpha_{0l[i]} + \alpha_1 Instrument_i + \alpha_2 Harmony_i + \alpha_3 Voice_i + \epsilon_i, \epsilon_i \overset{i.i.d}{\sim} N(0, \sigma^2)$$

$$\alpha_{0j[i]} = \beta_{01} + \eta_j, \eta_j \overset{i.i.d}{\sim} N(0, \tau_1^2)$$

$$\alpha_{0k[i]} = \beta_{02} + \eta_k, \eta_k \overset{i.i.d}{\sim} N(0, \tau_2^2)$$

$$\alpha_{0l[i]} = \beta_{03} + \eta_l, \eta_l \overset{i.i.d}{\sim} N(0, \tau_3^2)$$

Exercise 2

a

Among all the many factors the data provided, there are several groups of them that within each group, they are very likely to be correlated. For instance, CollegeMusic, NoClass, APTheory seems to belong to a group. So waht I want to do is pick out one best factor (the one with biggest influence) from each group. Those particular groups are: Selfdeclare, (PachListen, ClsListen), (CollegeMusic, NoClass, APTheory as a group), (PianoPlay and GuitarPlay).
Actually another important thing I consider when choosing factors is that how many NA's does that column have. It is definitely not proper to choose a factor with many NA's.

Although we try to avoid it, there are still many NA's in the data, we need to do some cleaning work:

```
> music1=subset(music, Selfdeclare!="NA" & PachListen!="NA" & ClsListen!="NA" & ClsListen!="NA" & Colleg
```

Originally, there're 2520 obs, after the cleaning, 2088 are left. We think it's still adequate to work with.

First, we did some EDA for the factors we choose, following is one example. However, from the boxplots we got, there seems to be little difference between the many different levels.

```
> boxplot(Classical~Selfdeclare,music1)
```

So we relevel the factors in the same standard. The only exception is Selfdeclare, we recategorize the values into musician and non musician since it is more meaningful this way.

```
> music1$musician=ifelse(music1$Selfdeclare>1,"Y","N")
> music1$PachListen1=ifelse(music1$PachListen>2,"High","Low")
> music1$ClsListen1=ifelse(music1$ClsListen>2,"High","Low")
> music1$PianoPlay1=ifelse(music1$PianoPlay>2,"High","Low")
> music1$GuitarPlay1=ifelse(music1$GuitarPlay>2,"High","Low")
```

After dealing the the raw data, we can begin choosing our factors:

```
> lmer1c.new=lmer(Classical~Instrument+Harmony+Voice+(1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subje
> lmer2a.1=update(lmer1c.new,.~.+musician)
> anova(lmer2a.1,lmer1c.new)
```

```
Data: music1
Models:
lmer1c.new: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer1c.new:     (1 | Subject:Harmony) + (1 | Subject:Voice)
lmer2a.1: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician
           Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer1c.new 12 8375.1 8442.7 -4175.6   8351.1
lmer2a.1   13 8372.8 8446.1 -4173.4   8346.8 4.3026      1    0.03805 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Musician is significant, p-value=0.03805. So we add it into the model.

```
> lmer2a.2=update(lmer2a.1,.~.+PachListen1)
> anova(lmer2a.2,lmer2a.1)

Data: music1
Models:
lmer2a.1: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician
lmer2a.2: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.2:     PachListen1
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a.1 13 8372.8 8446.1 -4173.4   8346.8
lmer2a.2 14 8374.2 8453.1 -4173.1   8346.2 0.6156      1     0.4327
```

PachListen1 not significant, p-value=0.4327.

```
> lmer2a.3=update(lmer2a.1,.~.+ClsListen1)
> anova(lmer2a.3,lmer2a.1)

Data: music1
Models:
lmer2a.1: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician
lmer2a.3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.3:     ClsListen1
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a.1 13 8372.8 8446.1 -4173.4   8346.8
lmer2a.3 14 8366.8 8445.7 -4169.4   8338.8 8.0177      1   0.004632 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ClsListen1 significant, p-value=0.0046. So we add it into the model.

```
> lmer2a.4=update(lmer2a.3,.~.+CollegeMusic)
> anova(lmer2a.4,lmer2a.3)

Data: music1
Models:
lmer2a.3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.3:     ClsListen1
lmer2a.4: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.4:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.4:     ClsListen1 + CollegeMusic
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a.3 14 8366.8 8445.7 -4169.4   8338.8
lmer2a.4 15 8367.8 8452.4 -4168.9   8337.8 0.9618      1     0.3267
```

CollegeMusic not significant, p-value=0.3267.

```
> lmer2a.5=update(lmer2a.3,.~.+NoClass)
> anova(lmer2a.5,lmer2a.3)

Data: music1
Models:
lmer2a.3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.3:     ClsListen1
lmer2a.5: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.5:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.5:     ClsListen1 + NoClass
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a.3 14 8366.8 8445.7 -4169.4   8338.8
lmer2a.5 15 8368.6 8453.1 -4169.3   8338.6 0.2252      1     0.6351
```

NoClass not significant, p-value=0.6351.

```
> lmer2a.6=update(lmer2a.3,.~.+APTheory)
> anova(lmer2a.6,lmer2a.3)

Data: music1
Models:
lmer2a.3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.3:     ClsListen1
lmer2a.6: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.6:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.6:     ClsListen1 + APTheory
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a.3 14 8366.8 8445.7 -4169.4   8338.8
lmer2a.6 15 8365.4 8450.0 -4167.7   8335.4 3.3535      1    0.06706 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

APTheory marginally significant, p-value=0.06706. It is the best compared with NoClass and CollegeMusic. So we add it into the model.

```
> lmer2a.7=update(lmer2a.6,.~.+PianoPlay1)
> anova(lmer2a.7,lmer2a.6)

Data: music1
Models:
lmer2a.6: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.6:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.6:     ClsListen1 + APTheory
lmer2a.7: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.7:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.7:     ClsListen1 + APTheory + PianoPlay1
         Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer2a.6 15 8365.4 8450.0 -4167.7   8335.4
lmer2a.7 16 8367.2 8457.4 -4167.6   8335.2 0.224      1      0.636
```

PianoPlay1 is not significant, p-value=0.636.

```
> lmer2a.8=update(lmer2a.6,.~.+GuitarPlay1)
> anova(lmer2a.8,lmer2a.6)
```

```
Data: music1
Models:
lmer2a.6: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.6:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.6:     ClsListen1 + APTheory
lmer2a.8: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a.8:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a.8:     ClsListen1 + APTheory + GuitarPlay1
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a.6 15 8365.4 8450.0 -4167.7   8335.4
lmer2a.8 16 8366.1 8456.3 -4167.0   8334.1 1.3669      1     0.2423
```

GuitarPlay1 is not significant, p-value=0.2423.

```
> lmer2a=lmer2a.6
```

So the final variables (besides the three main factors) I would like to put into my model are musician, ClsListen1 and APTheory.

b

```
> lmer2b=update(lmer2a,.~.-(1 | Subject:Instrument) - (1 | Subject:Harmony) -(1 | Subject:Voice)+(1|Subj
> lm2b=lm(Classical~Instrument + Harmony + Voice + musician+factor(ClsListen)+ APTheory, music1)
> rbind(AIC(lmer2a),AIC(lmer2b),AIC(lm2b))
```

```
         [,1]
[1,] 8383.936
[2,] 8778.488
[3,] 9212.198
```

```
> rbind(BIC(lmer2a),BIC(lmer2b),BIC(lm2b))
```

```
         [,1]
[1,] 8468.487
[2,] 8851.765
[3,] 9296.749
```

From the comparison of the value of AIC and BIC, it is very clear that when the fixed effects are settled, the model with all the three random effects is still the best.

c

```
> summary(lmer2a)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +      (1 | Subject:Harmony
   Data: music1

REML criterion at convergence: 8353.936

Random effects:
 Groups             Name        Variance Std.Dev.
 Subject:Harmony    (Intercept) 0.50614  0.7114
 Subject:Voice      (Intercept) 0.03974  0.1993
 Subject:Instrument (Intercept) 1.90904  1.3817
 Residual                       2.43377  1.5601
Number of obs: 2073, groups: Subject:Harmony, 232; Subject:Voice, 174; Subject:Instrument, 174

Fixed effects:
                  Estimate Std. Error t value
(Intercept)       5.277776   0.444150  11.883
Instrumentpiano   1.483376   0.270164   5.491
Instrumentstring  3.328606   0.269867  12.334
HarmonyI-V-IV    -0.066033   0.163866  -0.403
HarmonyI-V-VI     0.761089   0.163899   4.644
HarmonyIV-I-V     0.005021   0.163828   0.031
Voicepar3rd      -0.409474   0.091783  -4.461
Voicepar5th      -0.339780   0.091760  -3.703
musicianY        -1.144951   0.348547  -3.285
ClsListen1Low    -0.652568   0.281386  -2.319
APTheory          0.523044   0.289208   1.809

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r Vcpr5t
Instrumntpn -0.304
Instrmntstr -0.304  0.499
HrmnyI-V-IV -0.184  0.000     0.000
HrmnyI-V-VI -0.184  0.000     0.000     0.500
HrmnyIV-I-V -0.184  0.000     0.000     0.500  0.500
Voicepar3rd -0.104  0.000     0.000     0.000  0.001  0.001
Voicepar5th -0.103  0.000     0.000    -0.001 -0.001 -0.001  0.500
musicianY   -0.771  0.000     0.000     0.000  0.000  0.000  0.000  0.000
ClsListn1Lw -0.648 -0.001     0.000     0.000  0.000  0.000  0.000  0.000
APTheory    -0.212  0.002     0.000     0.000  0.000  0.000  0.000  0.000
            muscnY ClsL1L
Instrumntpn
Instrmntstr
HrmnyI-V-IV
HrmnyI-V-VI
HrmnyIV-I-V
Voicepar3rd
Voicepar5th
musicianY
ClsListn1Lw  0.439
APTheory    -0.056  0.240
```

If the instrument is piano, the classical rating will increase by 1.48 compared to when the instrument is electric guitar, keeping all the other variables constant.

If the instrument is string quartet, the classical rating will increase by 3.33 compared to when the instrument is electric guitar, keeping all the other variables constant.
If the Harmony is I-V-VI, the classical rating will increase by 0.76 compared to when the harmony is I-IV-V, keeping all the other variables constant.
If the Harmony are IV-I-V and I-V-IV, the classical rating will increase by 0.01 and decrease by 0.07 respectively compared to when the harmony is I-IV-V, keeping all the other variables constant. They're very small differences.
If the Voice is par 3rd, the classical rating will decrease by 0.41 compared to when the voice is contrary, keeping all the other variables constant.
If the Voice is par 5th, the classical rating will decrease by 0.34 compared to when the voice is contrary, keeping all the other variables constant.
If a person is a self-declared musician, the classical rating will decrease by 1.14, keeping all the other variables constant.
If a person listens to classical music a lot, the classical rating will increase by 0.65, keeping all the other variables constant.
If a person has taken AP music theory class in high School, the classical rating will increase by 0.52, keeping all the other variables constant.

3

```
> lmer3.1=update(lmer2a,.~.+musician*Instrument)
> anova(lmer3.1,lmer2a)

Data: music1
Models:
lmer2a: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a:     ClsListen1 + APTheory
lmer3.1: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer3.1:     ClsListen1 + APTheory + Instrument:musician
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a  15 8365.4 8450.0 -4167.7   8335.4
lmer3.1 17 8367.3 8463.1 -4166.6   8333.3 2.1583      2     0.3399

> lmer3.2=update(lmer2a,.~.+musician*Harmony)
> anova(lmer3.2,lmer2a)

Data: music1
Models:
lmer2a: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a:     ClsListen1 + APTheory
lmer3.2: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer3.2:     ClsListen1 + APTheory + Harmony:musician
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a  15 8365.4 8450.0 -4167.7   8335.4
lmer3.2 18 8357.3 8458.8 -4160.7   8321.3 14.113      3   0.002755 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> lmer3.3=update(lmer2a,.~.+musician*Voice)
> anova(lmer3.3,lmer2a)

Data: music1
Models:
lmer2a: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a:     ClsListen1 + APTheory
lmer3.3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3.3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer3.3:     ClsListen1 + APTheory + Voice:musician
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a  15 8365.4 8450.0 -4167.7   8335.4
lmer3.3 17 8369.0 8464.9 -4167.5   8335.0 0.3872      2      0.824

> lmer3.4=update(lmer2a,.~.+musician*ClsListen)
> anova(lmer3.4,lmer2a)

Data: music1
Models:
lmer2a: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a:     ClsListen1 + APTheory
lmer3.4: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3.4:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer3.4:     ClsListen1 + APTheory + ClsListen + musician:ClsListen
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a  15 8365.4 8450.0 -4167.7   8335.4
lmer3.4 17 8367.8 8463.6 -4166.9   8333.8 1.6478      2     0.4387

> lmer3.5=update(lmer2a,.~.+musician*APTheory)
> anova(lmer3.5,lmer2a)

Data: music1
Models:
lmer2a: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer2a:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer2a:     ClsListen1 + APTheory
lmer3.5: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3.5:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer3.5:     ClsListen1 + APTheory + musician:APTheory
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a  15 8365.4 8450.0 -4167.7   8335.4
lmer3.5 16 8366.0 8456.2 -4167.0   8334.0 1.4543      1     0.2278
```

The only significant interaction is between dichotomized musician and Harmony, p-value=0.003.
The following output tells us the interaction between dichotomized musician and Harmony in detail:

```
> summary(lmer3.2)

Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +      (1 | Subject:Harmony
```

```
   Data: music1

REML criterion at convergence: 8340.745

Random effects:
 Groups             Name        Variance Std.Dev.
 Subject:Harmony    (Intercept) 0.45374  0.6736
 Subject:Voice      (Intercept) 0.03974  0.1994
 Subject:Instrument (Intercept) 1.93502  1.3911
 Residual                       2.43374  1.5600
Number of obs: 2073, groups: Subject:Harmony, 232; Subject:Voice, 174; Subject:Instrument, 174

Fixed effects:
                        Estimate Std. Error t value
(Intercept)              5.66513    0.48584  11.660
Instrumentpiano          1.48332    0.27182   5.457
Instrumentstring         3.32848    0.27152  12.259
HarmonyI-V-IV           -0.18182    0.36286  -0.501
HarmonyI-V-VI           -0.35354    0.36286  -0.974
HarmonyIV-I-V           -0.31313    0.36286  -0.863
Voicepar3rd             -0.40956    0.09178  -4.462
Voicepar5th             -0.33974    0.09176  -3.702
musicianY               -1.62287    0.42640  -3.806
ClsListen1Low           -0.65270    0.28065  -2.326
APTheory                 0.52303    0.28845   1.813
HarmonyI-V-IV:musicianY  0.14294    0.40323   0.354
HarmonyI-V-VI:musicianY  1.37619    0.40325   3.413
HarmonyIV-I-V:musicianY  0.39256    0.40320   0.974

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HrI-V-IV HrI-V-VI HrIV-I-V Vcpr3r Vcpr5t
Instrumntpn -0.279
Instrmntstr -0.280  0.499
HrmnyI-V-IV -0.373  0.000     0.000
HrmnyI-V-VI -0.373  0.000     0.000     0.500
HrmnyIV-I-V -0.373  0.000     0.000     0.500    0.500
Voicepar3rd -0.095  0.000     0.000     0.000    0.000    0.000
Voicepar5th -0.095  0.000     0.000     0.000    0.000    0.000    0.500
musicianY   -0.811  0.000     0.000     0.425    0.425    0.425    0.000  0.000
ClsListn1Lw -0.590 -0.001     0.000     0.000    0.000    0.000    0.000  0.000
APTheory    -0.193  0.002     0.000     0.000    0.000    0.000    0.000  0.000
HrmI-V-IV:Y  0.336  0.000     0.000    -0.900   -0.450   -0.450    0.000 -0.001
HrmI-V-VI:Y  0.336  0.000     0.000    -0.450   -0.900   -0.450    0.000 -0.001
HrmIV-I-V:Y  0.336  0.000     0.000    -0.450   -0.450   -0.900    0.001 -0.001
            muscnY ClsL1L APThry HI-V-IV: HI-V-VI:
Instrumntpn
Instrmntstr
HrmnyI-V-IV
HrmnyI-V-VI
HrmnyIV-I-V
Voicepar3rd
Voicepar5th
musicianY
ClsListn1Lw  0.358
```

```
APTheory    -0.045  0.240
HrmI-V-IV:Y -0.473  0.000  0.000
HrmI-V-VI:Y -0.473  0.000  0.000  0.500
HrmIV-I-V:Y -0.473  0.000  0.000  0.500    0.500
```

As we can see, musicians who hear harmony I-V-IV, I-V-VI and IV-I-V will rate the stimulus 0.14, 1.38 and 0.39 higher to be classical music compared to non-musicians who hear the same kind of Harmony. This result agrees with the researchers' second hypothesis.

4

a

```
> lm4a=lm(Popular~Instrument+Harmony+Voice)
> anova(lm4a)
```

```
Analysis of Variance Table

Response: Popular
             Df  Sum Sq Mean Sq  F value Pr(>F)
Instrument    2  2924.2 1462.08 287.0703 <2e-16 ***
Harmony       3    31.1   10.37   2.0367 0.1067
Voice         2    15.3    7.63   1.4984 0.2237
Residuals  2485 12656.4    5.09
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We first made a simple linear model for the three main variabls. As we can see from the p-values, compared with the other two main factors, Instrument is the only significant factor for the popular rating. However, we will still keep the other two main factors for the following research of adding random effects into the model.

b

First we should determine what variables we need. The method is alomst the same as what we did in 2(a).

```
Data: music1
Models:
lmerp: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp:     (1 | Subject:Harmony) + (1 | Subject:Voice)
lmerp.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmerp   12 8413.1 8480.7 -4194.5   8389.1
lmerp.1 13 8408.0 8481.2 -4191.0   8382.0 7.1034      1   0.007694 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Musician is significant, p-value=0.0077. So we add it into the model.

```
Data: music1
Models:
lmerp.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician
lmerp.2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.2:     PachListen1
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmerp.1 13 8408.0 8481.2 -4191.0   8382.0
lmerp.2 14 8407.2 8486.1 -4189.6   8379.2 2.8021      1    0.09414 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

PachListen1 is marginally significant, p-value=0.094. So we add it into the model.

```
Data: music1
Models:
lmerp.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician
lmerp.3: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.3:     PachListen1 + ClsListen1
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmerp.1 13 8408.0 8481.2 -4191.0   8382.0
lmerp.3 15 8408.7 8493.2 -4189.3   8378.7 3.3086      2     0.1912
```

ClsListen1 is not significant, p-value=0.19.

```
Data: music1
Models:
lmerp.2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.2:     PachListen1
lmerp.4: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.4:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.4:     PachListen1 + CollegeMusic
        Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmerp.2 14 8407.2 8486.1 -4189.6   8379.2
lmerp.4 15 8409.2 8493.7 -4189.6   8379.2 0.002      1     0.9646
```

CollegeMusic is not significant, p-value=0.97.

```
Data: music1
Models:
lmerp.2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.2:     PachListen1
lmerp.5: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.5:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.5:     PachListen1 + NoClass
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmerp.2 14 8407.2 8486.1 -4189.6   8379.2
lmerp.5 15 8408.9 8493.5 -4189.5   8378.9 0.2577      1     0.6117
```

NoClass is not significant, p-value=0.61.

```
Data: music1
Models:
lmerp.2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.2:     PachListen1
lmerp.6: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.6:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.6:     PachListen1 + APTheory
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmerp.2 14 8407.2 8486.1 -4189.6   8379.2
lmerp.6 15 8409.0 8493.5 -4189.5   8379.0 0.1931      1     0.6603
```

APTheory is not significant, p-value=0.66.

```
Data: music1
Models:
lmerp.2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.2:     PachListen1
lmerp.7: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.7:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.7:     PachListen1 + PianoPlay1
        Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmerp.2 14 8407.2 8486.1 -4189.6   8379.2
lmerp.7 15 8408.2 8492.7 -4189.1   8378.2 1.018      1      0.313
```

PianoPlay1 is not significant, p-value=0.31.

```
Data: music1
Models:
lmerp.2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.2:     PachListen1
lmerp.8: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmerp.8:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmerp.8:     PachListen1 + GuitarPlay1
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmerp.2 14 8407.2 8486.1 -4189.6   8379.2
lmerp.8 15 8404.9 8489.5 -4187.5   8374.9 4.2691      1    0.03881 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

GuitarPlay1 is significant, p-value=0.039. So we add it in the model.

The final model is:

```
> lmer4b=lmerp.8
> summary(lmer4b)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +      (1 | Subject:Harmony)
   Data: music1

REML criterion at convergence: 8392.577

Random effects:
 Groups             Name        Variance Std.Dev.
 Subject:Harmony    (Intercept) 0.46148  0.6793
 Subject:Voice      (Intercept) 0.03643  0.1909
 Subject:Instrument (Intercept) 1.75852  1.3261
 Residual                       2.52281  1.5883
Number of obs: 2073, groups: Subject:Harmony, 232; Subject:Voice, 174; Subject:Instrument, 174

Fixed effects:
                 Estimate Std. Error t value
(Intercept)       6.62423    0.47609  13.914
Instrumentpiano  -0.99957    0.26088  -3.832
Instrumentstring -2.78228    0.26057 -10.678
HarmonyI-V-IV    -0.01384    0.16017  -0.086
HarmonyI-V-VI    -0.29843    0.16021  -1.863
HarmonyIV-I-V    -0.20311    0.16013  -1.268
Voicepar3rd       0.19630    0.09256   2.121
Voicepar5th       0.18516    0.09254   2.001
musicianY         0.77028    0.30550   2.521
PachListen1Low    0.78204    0.42180   1.854
GuitarPlay1Low   -0.70613    0.34567  -2.043

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r Vcpr5t
Instrumntpn -0.273
Instrmntstr -0.274  0.499
HrmnyI-V-IV -0.168  0.000     0.000
HrmnyI-V-VI -0.168  0.000     0.000     0.500
HrmnyIV-I-V -0.168  0.000     0.000     0.500  0.500
Voicepar3rd -0.097  0.000     0.000     0.000  0.001  0.001
Voicepar5th -0.097  0.000     0.000    -0.001 -0.001 -0.001  0.500
musicianY   -0.642  0.001     0.000     0.000  0.000  0.000  0.000  0.000
PachLstn1Lw -0.093  0.000     0.000     0.000  0.000  0.000  0.000  0.000
GuitrPly1Lw -0.711 -0.001     0.000    -0.001 -0.001  0.000  0.000  0.000
            muscnY PchL1L
Instrumntpn
Instrmntstr
HrmnyI-V-IV
HrmnyI-V-VI
HrmnyIV-I-V
Voicepar3rd
Voicepar5th
musicianY
PachLstn1Lw  0.145
GuitrPly1Lw  0.177 -0.094
```

If the instrument is piano, the popular rating will decrease by 1.00 compared to when the instrument is electric guitar, keeping all the other variables constant.

If the instrument is string quartet, the popular rating will decrease by 2.78 compared to when the instrument is electric guitar, keeping all the other variables constant.
If the Harmony is I-V-VI, the popular rating will decrease by 0.30 compared to when the harmony is I-IV-V, keeping all the other variables constant.
If the Harmony is IV-I-V, the popular rating will decrease by 0.20 compared to when the harmony is I-IV-V, keeping all the other variables constant
If the Harmony is I-V-IV, there's little difference between the popular rating compared to when the Harmony is I-IV-V, keeping all the other variables constant
If the Voice is par 3rd, the popular rating will increase by 0.20 compared to when the voice is contrary, keeping all the other variables constant.
If the Voice is par 5th, the popular rating will increase by 0.19 compared to when the voice is contrary, keeping all the other variables constant.
If a person is a self-declared musician, the popular rating will increase by 0.77, keeping all the other variables constant.
If a person is very familiar with Pachelbel's Canon, the popular rating will decrease by 0.78, keeping all the other variables constant.
If a person can play guitar very well, the popular rating will increase by 0.71, keeping all the other variables constant.

c

```
> lmer4c.1=update(lmer4b,.~.+musician*Instrument)
> anova(lmer4c.1,lmer4b)

Data: music1
Models:
lmer4b: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4b:     PachListen1 + GuitarPlay1
lmer4c.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c.1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4c.1:     PachListen1 + GuitarPlay1 + Instrument:musician
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4b   15 8404.9 8489.5 -4187.5   8374.9
lmer4c.1 17 8408.6 8504.4 -4187.3   8374.6 0.3313      2     0.8473

> lmer4c.2=update(lmer4b,.~.+musician*Harmony)
> anova(lmer4c.2,lmer4b)

Data: music1
Models:
lmer4b: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4b:     PachListen1 + GuitarPlay1
lmer4c.2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c.2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4c.2:     PachListen1 + GuitarPlay1 + Harmony:musician
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4b   15 8404.9 8489.5 -4187.5   8374.9
lmer4c.2 18 8402.9 8504.4 -4183.5   8366.9 7.9541      3    0.04697 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> lmer4c.3=update(lmer4b,.~.+musician*Voice)
> anova(lmer4c.3,lmer4b)

Data: music1
Models:
lmer4b: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4b:     PachListen1 + GuitarPlay1
lmer4c.3: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c.3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4c.3:     PachListen1 + GuitarPlay1 + Voice:musician
         Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer4b   15 8404.9 8489.5 -4187.5   8374.9
lmer4c.3 17 8408.7 8504.6 -4187.4   8374.7 0.152      2     0.9268

> lmer4c.4=update(lmer4b,.~.+musician*PachListen1)
> anova(lmer4c.4,lmer4b)

Data: music1
Models:
lmer4b: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4b:     PachListen1 + GuitarPlay1
lmer4c.4: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c.4:     (1 | Subject:Harmony) + (1 | Subject:Voice) + musician +
lmer4c.4:     PachListen1 + GuitarPlay1 + musician:PachListen1
         Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer4b   15 8404.9 8489.5 -4187.5   8374.9
lmer4c.4 16 8405.8 8496.0 -4186.9   8373.8 1.112      1     0.2916
```

The only significant interaction is between dichotomized musician and Harmony, p-value=0.047.
The following output tells us the interaction between dichotomized musician and Harmony in detail:

```
> summary(lmer4c.2)

Linear mixed model fit by REML ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +      (1 | Subject:Harmony)
   Data: music1

REML criterion at convergence: 8385.449

Random effects:
 Groups             Name        Variance Std.Dev.
 Subject:Harmony    (Intercept) 0.43865  0.6623
 Subject:Voice      (Intercept) 0.03648  0.1910
 Subject:Instrument (Intercept) 1.77108  1.3308
 Residual                       2.52268  1.5883
Number of obs: 2073, groups: Subject:Harmony, 232; Subject:Voice, 174; Subject:Instrument, 174

Fixed effects:
                        Estimate Std. Error t value
```

```
(Intercept)               6.64694    0.51563  12.891
Instrumentpiano          -0.99949    0.26171  -3.819
Instrumentstring         -2.78221    0.26139 -10.644
HarmonyI-V-IV            -0.50505    0.36155  -1.397
HarmonyI-V-VI             0.12121    0.36155   0.335
HarmonyIV-I-V            -0.22222    0.36155  -0.615
Voicepar3rd               0.19630    0.09257   2.121
Voicepar5th               0.18512    0.09255   2.000
musicianY                 0.74223    0.39202   1.893
PachListen1Low            0.78207    0.42139   1.856
GuitarPlay1Low           -0.70621    0.34534  -2.045
HarmonyI-V-IV:musicianY  0.60662    0.40178   1.510
HarmonyI-V-VI:musicianY -0.51817    0.40180  -1.290
HarmonyIV-I-V:musicianY  0.02362    0.40176   0.059

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HrI-V-IV HrI-V-VI HrIV-I-V Vcpr3r Vcpr5t
Instrumntpn -0.253
Instrmntstr -0.253  0.499
HrmnyI-V-IV -0.351  0.000     0.000
HrmnyI-V-VI -0.351  0.000     0.000     0.500
HrmnyIV-I-V -0.351  0.000     0.000     0.500    0.500
Voicepar3rd -0.090  0.000     0.000     0.000    0.000    0.000
Voicepar5th -0.090  0.000     0.000     0.000    0.000    0.000    0.500
musicianY   -0.703  0.000     0.000     0.461    0.461    0.461    0.000  0.000
PachLstn1Lw -0.086  0.000     0.000     0.000    0.000    0.000    0.000  0.000
GuitrPly1Lw -0.656 -0.001     0.000     0.000    0.000    0.000    0.000  0.000
HrmI-V-IV:Y  0.316  0.000     0.000    -0.900   -0.450   -0.450    0.000 -0.001
HrmI-V-VI:Y  0.316  0.000     0.000    -0.450   -0.900   -0.450    0.000 -0.001
HrmIV-I-V:Y  0.316  0.000     0.000    -0.450   -0.450   -0.900    0.001 -0.001
            muscnY PchL1L GtrP1L HI-V-IV: HI-V-VI:
Instrumntpn
Instrmntstr
HrmnyI-V-IV
HrmnyI-V-VI
HrmnyIV-I-V
Voicepar3rd
Voicepar5th
musicianY
PachLstn1Lw  0.113
GuitrPly1Lw  0.138 -0.094
HrmI-V-IV:Y -0.512  0.000  0.000
HrmI-V-VI:Y -0.512  0.000  0.000  0.500
HrmIV-I-V:Y -0.512  0.000  0.000  0.500    0.500
```

For example: As we can see, musicians who hear harmony I-V-IV will rate the stimulus 0.61 higher to be popular music compared to non-musicians who hear the same kind of Harmony.

This is fine. however there is a simpler more comprehensive way to test interactions with "musician"

5

Summary of the research

Research Background

Dr. Jimenez and a student, Vincent Rossi from the University of Pittsburgh are interested in measuring 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, recruited from the population of undergraduates at the University of Pittsburgh, and asked the listeners to rate the music on 1-10 scale on how classical and popular the music is. So the two response variables for the research are called Classical and Popular.

Method

We use three types of models to fit the data, namely, basic linear models, hierarchical models with single random effects and hierarchical models with combined random effects. We then use AIC, BIC and LRT criteria to help choose the best model, and we look at the estimated coefficient and p-value to find out the influence of a specific variable.

Results

The influence of the three main experimental factors (Instrument, Harmony Voice) varies for classical rating and popular rating.
All the three main factors are significant for Classical rating. For Harmony, only I-V-IV is significantly different from I-IV-V, the other two are not significantly different from I-IV-V. For Voice, both parallel 3rds and parallel 5ths are significantly different from contrary motion, but there is no significant difference between parallel 3rds and parallel 5ths. For Instrument, a stimulus played by piano or string will both be rated higher than one played by guitar.
As for the popular rating, only Instrument has a significant influence. Specifically, a stimulus played by piano or string will both be rated lower than than one played by guitar.

There are also other variables included in the two models. They both have musician, which indicates whether one declare oneself as a musician or not. It seems that a self-declared musician will more likely to rate a stimulus higher on popular rating and lower on classical rating.
The unique covariats for classical rating shows that a person who listen to classical music a lot will rate 0.65 higher score for the classical rating than a person who does not listen to classical music a lot; a person who took AP music theory class will rate 0.52 higher for classical rating than a person who did not.
The unique covariats for popular rating shows if a person is very familiar with Pachelbel's Canon, the popular rating will decrease by 0.78; and if a person can play guitar very well, the popular rating will increase by 0.71.

Discussion

In this research, We need to especially pay attention to the missing values in the dataset. It seems that some variables have most of their values to be NA's, which will influence the results a lot.

How did you deal with the NA's?

4: 18
5: 18

36

nice job overall. should check text in #5 for readability (generally very good, just difficult in a couple of places)