Hierarchical Final HW

Ryan Huan

Question 1

a

```
> data=read.csv("ratings.csv")
> data = data[!is.na(data$Popular),]
> data = data[!is.na(data$Classical),]
> attach(data)
> lm1a=lm(Classical~as.factor(Instrument)+as.factor(Harmony)+as.factor(Voice))
> lm1a1=lm(Classical~as.factor(Instrument)+as.factor(Harmony))
> lm1a2=lm(Classical~as.factor(Instrument)+as.factor(Voice))
> lm1a3=lm(Classical~as.factor(Harmony)+as.factor(Voice))
> rbind(AIC(lm1a,lm1a1,lm1a2,lm1a3))

      df      AIC
lm1a   9 11230.45
lm1a1  7 11242.69
lm1a2  6 11275.96
lm1a3  7 11908.94

> rbind(BIC(lm1a,lm1a1,lm1a2,lm1a3))

      df      BIC
lm1a   9 11282.84
lm1a1  7 11283.43
lm1a2  6 11310.89
lm1a3  7 11949.69
```

There are some NA in the Classical rating column. For analysis purposes, those observations with NA in the Classical rating section are deleted. Since only 27 observations are deleted, the impact of this precedure on the data set is very minimal.
As we can see, The model with all three variables has the smallest value in both AIC and BIC compared to other three models with only two of the three variables. This indicates that the three main experimental factors should be kept in the model.

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

```
> lmer1b0=lmer(Classical~Instrument+Harmony+Voice+(1|Subject))
> rbind(BIC(lm1a),BIC(lmer1b0))

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

> rbind(AIC(lm1a),AIC(lmer1b0))

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

Method 2: LRT

```
> exactRLRT(lmer1b0)

        simulated finite sample distribution of RLRT.

        (p-value based on 10000 simulated values)

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

For Method 1, the model with random intercept has smaller values in both AIC and BIC comparison with the model without the random intercept, indicating that we should keep the random intercept. For Method 2, P-value is less than 0.05, so we reject $H_0 : \tau^2 = 0$ and keep the random intercept. Both methods indicate that random intercept should be included in the model.

iii.

```
> lmer1b1=lmer(Classical~Instrument+Voice+(1|Subject))
> lmer1b2=lmer(Classical~Instrument+Harmony+(1|Subject))
> lmer1b3=lmer(Classical~Harmony+Voice+(1|Subject))
> rbind(AIC(lmer1b0,lmer1b1,lmer1b2,lmer1b3))

        df      AIC
lmer1b0 10 10491.51
lmer1b1  7 10552.74
lmer1b2  8 10505.58
lmer1b3  8 11423.04

> rbind(BIC(lmer1b0,lmer1b1,lmer1b2,lmer1b3))

        df      BIC
lmer1b0 10 10549.73
lmer1b1  7 10593.49
lmer1b2  8 10552.15
lmer1b3  8 11469.60
```

Model 1b1,1b2 and 1b3 are three models each without one of the three main experimental factors. As we can see, these three models all have bigger AIC and BIC than the model with all three main experimental factors. This indicates that the full model is the best model and the three main experimental factors are all useful.

c

i

```
> lmer1c=lmer(Classical~Instrument+Harmony+Voice+(1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:V
> rbind(BIC(lmer1c),BIC(lm1a),BIC(lmer1b0))

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

> rbind(AIC(lmer1c),AIC(lm1a),AIC(lmer1b0))

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

As we can see, when we try to compare the three models with both AIC and BIC, the model in part c with all three new random effect terms is the best as it has the smallest value in both AIC and BIC.

ii

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

        df      AIC
lmer1c  12 10075.51
lmer1c1 10 10176.17
lmer1c2  9 10101.74
lmer1c3 10 10092.66

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

        df      BIC
lmer1c  12 10145.37
lmer1c1 10 10234.38
lmer1c2  9 10154.13
lmer1c3 10 10150.87

> summary(lmer1c1)

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

REML criterion at convergence: 10156.17

Random effects:
 Groups             Name        Variance Std.Dev.
 Subject:Harmony    (Intercept) 0.40247  0.6344
 Subject:Voice      (Intercept) 0.01863  0.1365
 Subject:Instrument (Intercept) 4.00784  2.0020
 Residual                       2.44861  1.5648
Number of obs: 2493, groups: Subject:Harmony, 280; Subject:Voice, 210; Subject:Instrument, 210
```

```
Fixed effects:
              Estimate Std. Error t value
(Intercept)    5.83842    0.17604   33.17
HarmonyI-V-IV -0.03000    0.13919   -0.22
HarmonyI-V-VI  0.77049    0.13917    5.54
HarmonyIV-I-V  0.05633    0.13912    0.40
Voicepar3rd   -0.40645    0.08023   -5.07
Voicepar5th   -0.37017    0.08017   -4.62

Correlation of Fixed Effects:
            (Intr) HI-V-I HI-V-V HIV-I- Vcpr3r
HrmnyI-V-IV -0.394
HrmnyI-V-VI -0.394  0.499
HrmnyIV-I-V -0.395  0.500  0.500
Voicepar3rd -0.228 -0.002  0.001  0.002
Voicepar5th -0.227 -0.001 -0.002 -0.001  0.500
```

The best model is the model with all three main experimental factors as it has the smallest value in both AIC and BIC compared to the model with only two of the three main experimental factors.

The standard deviation of subject:Harmony group is 0.63. So it accounts for "personal biases" varied by type of harmony. The range is 0.6344*2 = 1.3. This suggests the variability that could not be accounted for using a fixed effect covariate. The standard deviation of subject:Voice group is 0.1365. So it accounts for "personal biases" varied by type of voice. The range is 0.1365*2 = .27. This suggests the variability that could not be accounted for using a fixed effect covariate. The standard deviation of subject:Instrument group is 2. So it accounts for "personal biases" varied by type of instrument. The range is 2*2 = 4. This suggests the variability that could not be accounted for using a fixed effect covariate.

iii

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

$$\alpha_{0j[i]} = \beta_{01} + \eta_{1j}, \eta_{1j} \overset{i.i.d}{\sim} N(0, \tau_1^2)$$

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

$$\alpha_{0l[i]} = \beta_{03} + \eta_{3l}, \eta_{3l} \overset{i.i.d}{\sim} N(0, \tau_3^2)$$

Question 2

a

```
> data1 = data[!is.na(data$ConsNotes),]
> data2 = data1[!is.na(data1$PachListen),]
> data3 = data2[!is.na(data2$ClsListen),]
> data3$KnowAxis[is.na(data3$KnowAxis)]=0
> data3$CollegeMusic[is.na(data3$CollegeMusic)]=0
> data3$NoClass[is.na(data3$NoClass)]=0
> data3$APTheory[is.na(data3$APTheory)]=0
```

```
> data3$Composing[is.na(data3$Composing)]=0
> data3$GuitarPlay[is.na(data3$GuitarPlay)]=0
> data3$X1stInstr[is.na(data3$X1stInstr)]=0
> data3$X2ndInstr[is.na(data3$X2ndInstr)]=0
> summary(data3)

       X             Subject        Harmony      Instrument        Voice
 Min.   :   1   15     :  36   I-IV-V:519   guitar:695   contrary:690
 1st Qu.: 777   17     :  36   I-V-IV:518   piano :682   par3rd  :691
 Median :1404   18b    :  36   I-V-VI:517   string:696   par5th  :692
 Mean   :1348   19     :  36   IV-I-V:519
 3rd Qu.:1922   20     :  36
 Max.   :2520   22     :  36
                (Other):1857
  Selfdeclare        OMSI        X16.minus.17     ConsInstr
 Min.   :1.000   Min.   : 11.0   Min.   :-4.00   Min.   :0.000
 1st Qu.:2.000   1st Qu.: 67.0   1st Qu.: 0.00   1st Qu.:1.670
 Median :2.000   Median :146.0   Median : 1.00   Median :3.000
 Mean   :2.581   Mean   :243.6   Mean   : 1.72   Mean   :3.063
 3rd Qu.:3.000   3rd Qu.:345.0   3rd Qu.: 3.00   3rd Qu.:4.330
 Max.   :6.000   Max.   :970.0   Max.   : 9.00   Max.   :5.000

   ConsNotes     Instr.minus.Notes   PachListen      ClsListen
 Min.   :0.000   Min.   :-4.0000   Min.   :0.000   Min.   :0.000
 1st Qu.:1.000   1st Qu.:-0.6700   1st Qu.:5.000   1st Qu.:1.000
 Median :3.000   Median : 0.6700   Median :5.000   Median :3.000
 Mean   :2.524   Mean   : 0.5382   Mean   :4.566   Mean   :2.183
 3rd Qu.:5.000   3rd Qu.: 2.0000   3rd Qu.:5.000   3rd Qu.:3.000
 Max.   :5.000   Max.   : 4.3300   Max.   :5.000   Max.   :5.000

    KnowRob          KnowAxis        X1990s2000s    X1990s2000s.minus.1960s1970s
 Min.   :0.0000   Min.   :0.0000   Min.   :0.000   Min.   :-4.000
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:3.000   1st Qu.: 0.000
 Median :0.0000   Median :0.0000   Median :5.000   Median : 2.000
 Mean   :0.9138   Mean   :0.9725   Mean   :4.028   Mean   : 2.003
 3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:5.000   3rd Qu.: 3.000
 Max.   :5.0000   Max.   :5.0000   Max.   :5.000   Max.   : 5.000
 NA's   :136                       NA's   :72      NA's   :108
  CollegeMusic       NoClass          APTheory       Composing
 Min.   :0.0000   Min.   :0.0000   Min.   :0.000   Min.   :0.000
 1st Qu.:1.0000   1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.000
 Median :1.0000   Median :1.0000   Median :0.000   Median :0.000
 Mean   :0.7742   Mean   :0.8828   Mean   :0.219   Mean   :1.066
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.000   3rd Qu.:2.000
 Max.   :1.0000   Max.   :8.0000   Max.   :1.000   Max.   :5.000

   PianoPlay       GuitarPlay        X1stInstr       X2ndInstr         first12
 Min.   :0.000   Min.   :0.0000   Min.   :0.000   Min.   :0.0000   guitar:720
 1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.000   1st Qu.:0.0000   piano :603
 Median :0.000   Median :0.0000   Median :0.000   Median :0.0000   string:750
 Mean   :1.073   Mean   :0.7665   Mean   :1.183   Mean   :0.2219
 3rd Qu.:1.000   3rd Qu.:1.0000   3rd Qu.:2.000   3rd Qu.:0.0000
 Max.   :5.000   Max.   :5.0000   Max.   :5.000   Max.   :4.0000
```

```
   Classical         Popular
 Min.   : 0.000   Min.   : 0.000
 1st Qu.: 4.000   1st Qu.: 4.000
 Median : 6.000   Median : 6.000
 Mean   : 5.729   Mean   : 5.546
 3rd Qu.: 8.000   3rd Qu.: 7.000
 Max.   :19.000   Max.   :19.000
```

```
> attach(data3)
```

```
The following objects are masked from data:

    APTheory, Classical, ClsListen, CollegeMusic, Composing, ConsInstr,
    ConsNotes, first12, GuitarPlay, Harmony, Instr.minus.Notes,
    Instrument, KnowAxis, KnowRob, NoClass, OMSI, PachListen,
    PianoPlay, Popular, Selfdeclare, Subject, Voice, X, X16.minus.17,
    X1990s2000s, X1990s2000s.minus.1960s1970s, X1stInstr, X2ndInstr
```

```
> lmer2a=lmer(data3$Classical~data3$Instrument+data3$Harmony+data3$Voice+(1|Subject:Instrument)+(1|Subje
> anova(lmer2a)
```

```
Analysis of Variance Table
                 Df  Sum Sq Mean Sq F value
data3$Instrument  2 293.650 146.825  58.802
data3$Harmony     3 108.568  36.189  14.494
data3$Voice       2  65.092  32.546  13.034
```

X16.minus.17

```
> lmer2a.X16.minus.17=update(lmer2a,.~.+X16.minus.17)
> anova(lmer2a.X16.minus.17,lmer2a)
```

```
Data:
Models:
lmer2a: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer2a:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
lmer2a.X16.minus.17: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer2a.X16.minus.17:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer2a.X16.minus.17:     X16.minus.17
                    Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a              12 8398.6 8466.3 -4187.3   8374.6
lmer2a.X16.minus.17 13 8396.0 8469.3 -4185.0   8370.0 4.6524      1    0.03101

lmer2a
lmer2a.X16.minus.17 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ClsListen

```
> data3$clsListen=ifelse(ClsListen>2,1,0)
> lmer2a.ClsListen=update(lmer2a.X16.minus.17,
+ .~.+data3$clsListen)
> anova(lmer2a.ClsListen,lmer2a.X16.minus.17)
```

```
Data:
Models:
```

```
lmer2a.X16.minus.17: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer2a.X16.minus.17:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer2a.X16.minus.17:     X16.minus.17
lmer2a.ClsListen: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer2a.ClsListen:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer2a.ClsListen:     X16.minus.17 + data3$clsListen
                    Df    AIC    BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer2a.X16.minus.17 13 8396.0 8469.3  -4185   8370.0
lmer2a.ClsListen    14 8392.1 8471.0  -4182   8364.1 5.9315      1    0.01487 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

APTheory

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

```
Data:
Models:
lmer2a.ClsListen: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer2a.ClsListen:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer2a.ClsListen:     X16.minus.17 + data3$clsListen
lmer2a.APTheory: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer2a.APTheory:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer2a.APTheory:     X16.minus.17 + data3$clsListen + APTheory
                 Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer2a.ClsListen 14 8392.1 8471.0 -4182.0   8364.1
lmer2a.APTheory  15 8390.2 8474.7 -4180.1   8360.2 3.897      1    0.04837 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

There are NA in several columns. For analysis purpose, Those NAs need to be modified. For the variable ConsNotes, as there is no way i could know how much did those people who answered "NA" in this question concentrate on the notes while listening, I deleted all the observations that have NAs in the "ConsNotes" section. Same procedure is conducted on variables "PachListen" and "ClsListen".
For the variable "X1stInstr", Although there are several NAs in this column, I assume that those people that answered NA actually play no musical instrument. Therefore, they can be treated as "0" in this column. Same procedure is conducted on variables "KnowRob", "KnowAxis", "X1990s2000s", "CollegeMusic", "NoClass","ApTheory", "Composing", "GuitarPlay", and "X2ndInstr". The rest of the variables do not have NAs in their columns respectively.
The variable selection procedure is as follow: one variable is added to Model "lmer2a" one at a time. The model with one additonal variable is compared with model "lmer2a" using ANOVA. If anova suggests that this variable should be included in the model, the new model becomes the model we want to use to compare. Otherwise, The old model would be used to compare. After analyzing all the variables, three additonal variables should be added to the model. They are: "X16.minus.17", "ClsListen", and "APTheory". When I first included "X16.minus.17" in the model with the orginal 3 fixed effects and 3 random effects, this variable decreases the AIC and BIC. The p-value from the anova table is less than 0.05. After that, "ClsListen" is added to the model with "X16.minus.17" included. The variable is treated as a categorical variable with rating higher than 2 as a cateogory and the rest as another category. When comparing this model with the model with only "X16.minus.17" and original variables, this new model has a smaller AIC and BIC. Similiarly, "APThoery" is also added to the model.

b

```
> lmer2b0=update(lmer2a.APTheory,.~.+(1|Subject:APTheory))
> lmer2b1=update(lmer2a.APTheory,.~.+(1|Subject:X16.minus.17))
```

```
> lmer2b2=update(lmer2a.APTheory,.~.+(1|Subject:ClsListen))
> lmer2b3=update(lmer2a.APTheory,.~.+(1|Subject:APTheory)+(1|Subject:X16.minus.17))
> lmer2b4=update(lmer2a.APTheory,.~.+(1|Subject:APTheory)+(1|Subject:ClsListen))
> lmer2b5=update(lmer2a.APTheory,.~.+(1|Subject:X16.minus.17)+(1|Subject:ClsListen))
> lmer2b6=update(lmer2a.APTheory,.~.+(1|Subject:APTheory)+(1|Subject:X16.minus.17)+(1|Subject:ClsListen
> rbind(BIC(lmer2a.APTheory),BIC(lmer2b0),BIC(lmer2b1),BIC(lmer2b2),BIC(lmer2b3),BIC(lmer2b4),BIC(lmer2b

         [,1]
[1,] 8497.624
[2,] 8480.866
[3,] 8480.866
[4,] 8480.866
[5,] 8488.503
[6,] 8488.503
[7,] 8488.503
[8,] 8496.139

> rbind(AIC(lmer2a.APTheory),AIC(lmer2b0),AIC(lmer2b1),AIC(lmer2b2),AIC(lmer2b3),AIC(lmer2b4),AIC(lmer2b

         [,1]
[1,] 8413.072
[2,] 8390.678
[3,] 8390.678
[4,] 8390.678
[5,] 8392.678
[6,] 8392.678
[7,] 8392.678
[8,] 8394.678
```

Model lmer2b0,lmer2b1,and lmer2b2 have the same AIC and BIC. Their AIC and BIC values are the smallest among the comparsion group. Either of the models is a good fit for the data. I will choose model lmer2b2 with (1|Subject:ClsListen)) as the additional random effect.

c

```
> summary(lmer2b2)

Linear mixed model fit by REML ['lmerMod']
Formula: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +      (1 | Subject:Instrument

REML criterion at convergence: 8358.678

Random effects:
 Groups             Name        Variance Std.Dev.
 Subject:Harmony    (Intercept) 0.35448  0.5954
 Subject:Voice      (Intercept) 0.01348  0.1161
 Subject:Instrument (Intercept) 1.26339  1.1240
 Subject:ClsListen  (Intercept) 1.08533  1.0418
 Residual                       2.50926  1.5841
Number of obs: 2073, groups: Subject:Harmony, 232; Subject:Voice, 174; Subject:Instrument, 174; Subject:

Fixed effects:
                      Estimate Std. Error t value
(Intercept)            3.98719    0.32066  12.434
data3$Instrumentpiano  1.41111    0.22570   6.252
```

```
data3$Instrumentstring  3.17621    0.22535  14.095
data3$HarmonyI-V-IV    -0.03193    0.14803  -0.216
data3$HarmonyI-V-VI     0.82899    0.14806   5.599
data3$HarmonyIV-I-V     0.06838    0.14799   0.462
data3$Voicepar3rd      -0.41538    0.08795  -4.723
data3$Voicepar5th      -0.37496    0.08793  -4.264
X16.minus.17           -0.08230    0.05716  -1.440
data3$clsListen         0.51365    0.34595   1.485
APTheory                0.59105    0.41516   1.424

Correlation of Fixed Effects:
              (Intr) dt3$Instrmntp dt3$Instrmnts d3$HI-V-I d3$HI-V-V d3$HIV
dt3$Instrmntp -0.352
dt3$Instrmnts -0.351  0.499
dt3$HI-V-IV   -0.231  0.000         0.000
dt3$HI-V-VI   -0.231  0.000         0.000         0.500
dt3$HIV-I-V   -0.231  0.000         0.000         0.500     0.500
dt3$Vcpr3rd   -0.138  0.000         0.000         0.000     0.001     0.001
dt3$Vcpr5th   -0.137  0.000         0.000        -0.001    -0.001    -0.001
X16.mins.17   -0.341  0.000         0.000         0.000     0.000     0.000
dt3$clsLstn   -0.515  0.001         0.000         0.000     0.000     0.000
APTheory      -0.203  0.002         0.000         0.000     0.000     0.000
              dt3$V3 dt3$V5 X16..1 dt3$cL
dt3$Instrmntp
dt3$Instrmnts
dt3$HI-V-IV
dt3$HI-V-VI
dt3$HIV-I-V
dt3$Vcpr3rd
dt3$Vcpr5th    0.500
X16.mins.17    0.000  0.000
dt3$clsLstn    0.000  0.000  0.035
APTheory       0.000  0.000  0.052 -0.185
```

Keep other variable constant, when the instrument is piano, the classical rating increases by 1.41 compared to when the instrument is guitar.
Keep other variable constant, when the instrument is string instrument, the classical rating increases by 3.18 compared to when the instrument is guitar.
Keep other variable constant, when the Harmony is I-V-IV, the classical rating decreases by .03 compared to when the harmony is I-IV-V.
Keep other variable constant, when the Harmony is I-V-VI, the classical rating increases by .83 compared to when the harmony is I-IV-V.
Keep other variable constant, when the Harmony is IV-I-V, the classical rating increases by 0.07 compared to when the harmony is I-IV-V.
Keep other variable constant, when the Voice is par 3rd, the classical rating decreases by 0.42 compared to when the voice is contrary.
Keep other variable constant, when the Voice is par 5rd, the classical rating decreases by 0.37 compared to when the voice is contrary.
Keep other variable constant, one unit increase in auxiliary measure of listener's ability to distinguish classical vs popular music decreases the classical rating by 0.08.
Keep other variable constant, when people rated themselves higher than 2 in the listening frequency to classical music , the classical rating increases by 0.51 compared to people rated less than or equal to 2.
People that took Ap music theory class in high school rated higher in classical rating by 0.59 than people that did not take Ap music theory class in high school.

The standard deviation of subject:Harmony group is 0.6. So it accounts for "personal biases" varied by type of harmony. The range is 0.6*2 = 1.2. This suggests the variability that could not be accounted for using a fixed effect covariate.
The standard deviation of subject:Voice group is 0.01. So it accounts for "personal biases" varied by type of Voice. The range is 0.01*2 = 0.02. This suggests the variability that could not be accounted for using a fixed effect covariate.
The standard deviation of subject:Instrument group is 0.01. So it accounts for "personal biases" varied by type of Instrument. The range is 1.26*2 = 2.52. This suggests the variability that could not be accounted for using a fixed effect covariate.
The standard deviation of subject:ClsListen group is 1.086. So it accounts for "personal biases" varied by type of Clslisten. The range is 1.08*2 = 2.16. This suggests the variability that could not be accounted for using a fixed effect covariate.

Question 3

```
> data3$declare=ifelse(data3$Selfdeclare>2,1,0)
> lmer3=update(lmer2b2,.~.+data3$declare)
> lmer3a=update(lmer3,.~.+data3$declare*data3$Instrument)
> anova(lmer3,lmer3a)

Data:
Models:
lmer3: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3:     data3$declare
lmer3a: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3a:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3a:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3a:     data3$declare + data3$Instrument:data3$declare
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3  17 8371.1 8466.9 -4168.5   8337.1
lmer3a 19 8371.2 8478.3 -4166.6   8333.2 3.8713      2     0.1443

> lmer3b=update(lmer3,.~.+data3$declare*data3$Voice)
> anova(lmer3,lmer3b)

Data:
Models:
lmer3: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3:     data3$declare
lmer3b: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3b:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3b:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3b:     data3$declare + data3$Voice:data3$declare
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3  17 8371.1 8466.9 -4168.5   8337.1
lmer3b 19 8373.7 8480.8 -4167.8   8335.7 1.3929      2     0.4984

> lmer3c=update(lmer3,.~.+data3$declare*data3$Harmony)
> anova(lmer3,lmer3c)
```

```
Data:
Models:
lmer3: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3:     data3$declare
lmer3c: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3c:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3c:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3c:     data3$declare + data3$Harmony:data3$declare
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3  17 8371.1 8466.9 -4168.5   8337.1
lmer3c 20 8345.9 8458.7 -4153.0   8305.9 31.146      3  7.918e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> lmer3d=update(lmer3c,.~.+data3$declare*X16.minus.17)
> anova(lmer3c,lmer3d)

Data:
Models:
lmer3c: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3c:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3c:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3c:     data3$declare + data3$Harmony:data3$declare
lmer3d: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3d:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3d:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3d:     data3$declare + data3$Harmony:data3$declare + X16.minus.17:data3$declare
       Df    AIC    BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3c 20 8345.9 8458.7  -4153   8305.9
lmer3d 21 8341.9 8460.3  -4150   8299.9 6.0164      1    0.01417 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> lmer3e=update(lmer3d,.~.+data3$declare*data3$clsListen)
> anova(lmer3d,lmer3e)

Data:
Models:
lmer3d: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3d:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3d:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3d:     data3$declare + data3$Harmony:data3$declare + X16.minus.17:data3$declare
lmer3e: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3e:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3e:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3e:     data3$declare + data3$Harmony:data3$declare + X16.minus.17:data3$declare +
lmer3e:     data3$clsListen:data3$declare
       Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer3d 21 8341.9 8460.3 -4150.0   8299.9
lmer3e 22 8343.8 8467.9 -4149.9   8299.8  0.06      1     0.8065

> lmer3f=update(lmer3d,.~.+data3$declare*APTheory)
> anova(lmer3d,lmer3f)
```

```
Data:
Models:
lmer3d: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3d:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3d:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3d:     data3$declare + data3$Harmony:data3$declare + X16.minus.17:data3$declare
lmer3f: data3$Classical ~ data3$Instrument + data3$Harmony + data3$Voice +
lmer3f:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice) +
lmer3f:     X16.minus.17 + data3$clsListen + APTheory + (1 | Subject:ClsListen) +
lmer3f:     data3$declare + data3$Harmony:data3$declare + X16.minus.17:data3$declare +
lmer3f:     APTheory:data3$declare
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3d 21 8341.9 8460.3 -4150.0   8299.9
lmer3f 22 8340.4 8464.4 -4148.2   8296.4 3.5062      1    0.06114 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The two statistically significant (p value less than 0.05) interactions are between declare & harmony and declare & X16.minus.17. These two interactions should be included in the model.

Question 4

a

```
> lm4a0=lm(Popular~as.factor(Instrument)+as.factor(Harmony)+as.factor(Voice))
> anova(lm4a0)

Analysis of Variance Table

Response: Popular
                        Df  Sum Sq Mean Sq  F value Pr(>F)
as.factor(Instrument)    2  2369.8 1184.88 229.7816 <2e-16 ***
as.factor(Harmony)       3    26.2    8.73   1.6935 0.1663
as.factor(Voice)         2     8.8    4.38   0.8504 0.4274
Residuals             2065 10648.3    5.16
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As we can see from the p-values, instrument is the only significant factor for the popular rating. However, in order to fit the three random effects ((1|Subject:Instrument), (1|Subject:Harmony), and (1|Subject:Voice)), the three main effects will be all kept.

b

```
> attach(data3)

The following objects are masked from data3 (position 3):

    APTheory, Classical, ClsListen, CollegeMusic, Composing, ConsInstr,
    ConsNotes, first12, GuitarPlay, Harmony, Instr.minus.Notes,
    Instrument, KnowAxis, KnowRob, NoClass, OMSI, PachListen,
    PianoPlay, Popular, Selfdeclare, Subject, Voice, X, X16.minus.17,
    X1990s2000s, X1990s2000s.minus.1960s1970s, X1stInstr, X2ndInstr
The following objects are masked from data:
```

```
APTheory, Classical, ClsListen, CollegeMusic, Composing, ConsInstr,
ConsNotes, first12, GuitarPlay, Harmony, Instr.minus.Notes,
Instrument, KnowAxis, KnowRob, NoClass, OMSI, PachListen,
PianoPlay, Popular, Selfdeclare, Subject, Voice, X, X16.minus.17,
X1990s2000s, X1990s2000s.minus.1960s1970s, X1stInstr, X2ndInstr
```

```
> lmer4b=lmer(Popular~Instrument+Harmony+Voice+(1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Vo
```

GuitarPlay

```
> data3$guitarPlay=ifelse(GuitarPlay>2,1,0)
> lmer4b.GuitarPlay=update(lmer4b,.~.+data3$guitarPlay)
> anova(lmer4b.GuitarPlay,lmer4b)
```

```
Data:
Models:
lmer4b: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b:     (1 | Subject:Harmony) + (1 | Subject:Voice)
lmer4b.GuitarPlay: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b.GuitarPlay:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay
                  Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer4b            12 8477.3 8544.9 -4226.6   8453.3
lmer4b.GuitarPlay 13 8475.4 8548.7 -4224.7   8449.4 3.887      1    0.04866 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

X1990s2000s

```
> data3$x1990s2000s=ifelse(X1990s2000s>2,1,0)
> lmer4b.X1990s2000s=update(lmer4b.GuitarPlay,.~.+data3$x1990s2000s)
> anova(lmer4b.GuitarPlay,lmer4b.X1990s2000s)
```

```
Data:
Models:
lmer4b.GuitarPlay: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b.GuitarPlay:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay
lmer4b.X1990s2000s: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4b.X1990s2000s:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4b.X1990s2000s:     data3$x1990s2000s
                   Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4b.GuitarPlay  13 8475.4 8548.7 -4224.7   8449.4
lmer4b.X1990s2000s 14 8183.8 8262.2 -4077.9   8155.8 293.59      1  < 2.2e-16

lmer4b.GuitarPlay
lmer4b.X1990s2000s ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> lmer4b0=update(lmer4b.X1990s2000s,.~.+(1|Subject:GuitarPlay))
> lmer4b1=update(lmer4b.X1990s2000s,.~.+(1|Subject:X1990s2000s))
> lmer4b2=update(lmer4b.X1990s2000s,.~.+(1|Subject:GuitarPlay)+(1|Subject:X1990s2000s))
> rbind(BIC(lmer4b.X1990s2000s),BIC(lmer4b0),BIC(lmer4b1),BIC(lmer4b2))
```

```
         [,1]
[1,] 8279.457
[2,] 8265.172
[3,] 8265.172
[4,] 8272.773
```

```
> rbind(AIC(lmer4b.X1990s2000s),AIC(lmer4b0),AIC(lmer4b1),AIC(lmer4b2))

         [,1]
[1,] 8201.037
[2,] 8181.151
[3,] 8181.151
[4,] 8183.151

> summary(lmer4b1)

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

REML criterion at convergence: 8151.151

Random effects:
 Groups              Name        Variance Std.Dev.
 Subject:Harmony     (Intercept) 0.34685  0.5889
 Subject:Voice       (Intercept) 0.02481  0.1575
 Subject:Instrument  (Intercept) 1.07711  1.0378
 Subject:X1990s2000s (Intercept) 0.91315  0.9556
 Residual                        2.66631  1.6329
Number of obs: 2001, groups: Subject:Harmony, 224; Subject:Voice, 168; Subject:Instrument, 168; Subject:

Fixed effects:
                  Estimate Std. Error t value
(Intercept)        6.66750    0.45525  14.646
Instrumentpiano   -0.96622    0.21584  -4.476
Instrumentstring  -2.64370    0.21543 -12.272
HarmonyI-V-IV     -0.04412    0.15183  -0.291
HarmonyI-V-VI     -0.26761    0.15187  -1.762
HarmonyIV-I-V     -0.22942    0.15179  -1.511
Voicepar3rd        0.14687    0.09429   1.558
Voicepar5th        0.14766    0.09427   1.566
data3$guitarPlay   0.64386    0.49145   1.310
data3$x1990s2000s -0.04868    0.46404  -0.105

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r Vcpr5t
Instrumntpn -0.237
Instrmntstr -0.237  0.499
HrmnyI-V-IV -0.167  0.000     0.000
HrmnyI-V-VI -0.167  0.000     0.000     0.500
HrmnyIV-I-V -0.167  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
dat3$gtrPly  0.000  0.001     0.000     0.001  0.001  0.000  0.000  0.000
d3$19902000 -0.853  0.000     0.000     0.000  0.000  0.000  0.000  0.000
            dt3$gP
Instrumntpn
Instrmntstr
HrmnyI-V-IV
HrmnyI-V-VI
HrmnyIV-I-V
Voicepar3rd
```

```
Voicepar5th
dat3$gtrPly
d3$19902000 -0.154
```

The procedure is the same as in question 2. The additional two variables that should be added in the model are GuitarPlay and X1990s2000s.
After reexamining the random effects, either (1|Subject:GuitarPlay) or (1|Subject:X1990s2000s) should be included in the model. I choose (1|Subject:X1990s2000s) to be in the model. Keep other variable constant, when the instrument is piano, the popular rating decreases by 0.97 compared to when the instrument is guitar.
Keep other variable constant, when the instrument is string instrument, the popular rating decreases by 2.64 compared to when the instrument is guitar.
Keep other variable constant, when the Harmony is I-V-IV, the popular rating decreases by .04 compared to when the harmony is I-IV-V.
Keep other variable constant, when the Harmony is I-V-VI, the popular rating decreases by .27 compared to when the harmony is I-IV-V.
Keep other variable constant, when the Harmony is IV-I-V, the popular rating decreases by 0.23 compared to when the harmony is I-IV-V.
Keep other variable constant, when the Voice is par 3rd, the popular rating increases by 0.146 compared to when the voice is contrary.
Keep other variable constant, when the Voice is par 5rd, the popular rating increases by 0.147 compared to when the voice is contrary.
Keep other variable constant, when people rated themselves higher than 2 in the listening frequency to pop and rock music from the 90's and 2000's , the popular rating decreases by 0.05 compared to people rated less than or equal to 2.
Keep other variable constant, when people rated themselves higher than 2 in the proficiency of guitar play , the popular rating increases by 0.65 compared to people rated less than or equal to 2.

The standard deviation of subject:Harmony group is 0.35. So it accounts for "personal biases" varied by type of harmony. The range is 0.35*2 = 0.7. This suggests the variability that could not be accounted for using a fixed effect covariate.
The standard deviation of subject:Voice group is 0.025. So it accounts for "personal biases" varied by type of Voice. The range is 0.025*2 = 0.05. This suggests the variability that could not be accounted for using a fixed effect covariate.
The standard deviation of subject:Instrument group is 1.08. So it accounts for "personal biases" varied by type of Instrument. The range is 1.08*2 = 2.16. This suggests the variability that could not be accounted for using a fixed effect covariate.
The standard deviation of subject:X1990s2000s group is 0.91. So it accounts for "personal biases" varied by type of Instrument. The range is 0.91*2 = 1.82. This suggests the variability that could not be accounted for using a fixed effect covariate.

c

```
> lmer4c=update(lmer4b1,.~.+data3$declare)
> lmer4c1=update(lmer4c,.~.+data3$declare*Instrument)
> anova(lmer4c,lmer4c1)

Data:
Models:
lmer4c: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare
lmer4c1: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c1:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
```

```
lmer4c1:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c1:     Instrument:data3$declare
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4c  16 8166.4 8256.1 -4067.2   8134.4
lmer4c1 18 8168.7 8269.5 -4066.3   8132.7 1.7464      2     0.4176
```

```
> lmer4c2=update(lmer4c,.~.+data3$declare*Voice)
> anova(lmer4c,lmer4c2)

Data:
Models:
lmer4c: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare
lmer4c2: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c2:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c2:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c2:     Voice:data3$declare
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4c  16 8166.4 8256.1 -4067.2   8134.4
lmer4c2 18 8167.5 8268.4 -4065.8   8131.5 2.9109      2     0.2333
```

```
> lmer4c3=update(lmer4c,.~.+data3$declare*Harmony)
> anova(lmer4c,lmer4c3)

Data:
Models:
lmer4c: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare
lmer4c3: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c3:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c3:     Harmony:data3$declare
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4c  16 8166.4 8256.1 -4067.2   8134.4
lmer4c3 19 8157.3 8263.8 -4059.7   8119.3 15.103      3   0.001731 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> lmer4c4=update(lmer4c3,.~.+data3$declare*X16.minus.17)
> anova(lmer4c3,lmer4c4)

Data:
Models:
lmer4c3: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c3:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c3:     Harmony:data3$declare
lmer4c4: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c4:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c4:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c4:     X16.minus.17 + Harmony:data3$declare + data3$declare:X16.minus.17
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4c3 19 8157.3 8263.8 -4059.7   8119.3
lmer4c4 21 8160.0 8277.6 -4059.0   8118.0 1.3311      2      0.514
```

```
> lmer4c5=update(lmer4c3,.~.+data3$declare*data3$clsListen)
> anova(lmer4c5,lmer4c3)

Data:
Models:
lmer4c3: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c3:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c3:     Harmony:data3$declare
lmer4c5: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c5:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c5:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c5:     data3$clsListen + Harmony:data3$declare + data3$declare:data3$clsListen
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer4c3 19 8157.3 8263.8 -4059.7   8119.3
lmer4c5 21 8160.4 8278.0 -4059.2   8118.4 0.9836      2     0.6115

> lmer4c6=update(lmer4c3,.~.+data3$declare*APTheory)
> anova(lmer4c6,lmer4c3)

Data:
Models:
lmer4c3: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c3:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c3:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c3:     Harmony:data3$declare
lmer4c6: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer4c6:     (1 | Subject:Harmony) + (1 | Subject:Voice) + data3$guitarPlay +
lmer4c6:     data3$x1990s2000s + (1 | Subject:X1990s2000s) + data3$declare +
lmer4c6:     APTheory + Harmony:data3$declare + data3$declare:APTheory
        Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer4c3 19 8157.3 8263.8 -4059.7   8119.3
lmer4c6 21 8160.7 8278.4 -4059.4   8118.7 0.588      2     0.7453
```

The only statistically significant (p value less than 0.05) interaction is between declare & harmony. The interaction should be included in the model.

Question 5

Brief Write-up

Findings for Classical and Popular Ratings

Background

Dr.Jimenez is interested in what factors influence listeners' identification of music as "classical" or "popular". The data were collected by Dr. Jimenez and Vincent Rossi.

Data Modification

During the analysis process, NAs have been found in the dataset. For analysis purposes, Those NAs need to be modified. All the observations that had NAs in either "classical rating" or "popular rating" are deleted. For the variable ConsNotes, as there is no way I could know how much did those people who answered "NA" in this question concentrate on the notes while listening, I deleted all the observations that have NAs in the "ConsNotes" section. Same procedure is conducted on variables "PachListen" and "ClsListen". For the variable "X1stInstr", although there are several NAs in this column, I assume that those people that answered NA actually play no musical instrument. Therefore, they can be treated as "0" in this column. Same procedure is conducted on variables "KnowRob", "KnowAxis", "X1990s2000s", "CollegeMusic", "NoClass","ApTheory", "Composing", "GuitarPlay", and "X2ndInstr". The rest of the variables do not have NAs in their columns respectively.

Result

Classical Rating

The three main experimental factors "Instrument", "Voice" and "Harmony" all have statistically significant influence on classical rating. As mentioned in question 2, keep other variable constant, when the instrument is piano, the classical rating increases by 1.41 compared to when the instrument is guitar. Keep other variable constant, when the instrument is string instrument, the classical rating increases by 3.18 compared to when the instrument is guitar. Keep other variable constant, when the Harmony is I-V-IV, the classical rating decreases by .03 compared to when the harmony is I-IV-V. Keep other variable constant, when the Harmony is I-V-VI, the classical rating increases by .83 compared to when the harmony is I-IV-V. Keep other variable constant, when the Harmony is IV-I-V, the classical rating increases by 0.07 compared to when the harmony is I-IV-V. Keep other variable constant, when the Voice is par 3rd, the classical rating decreases by 0.42 compared to when the voice is contrary. Keep other variable constant, when the Voice is par 5rd, the classical rating decreases by 0.37 compared to when the voice is contrary.

Some other individual covariates are also useful for predicting classical rating score. One unit increase in auxiliary measure of listener's ability to distinguish classical vs. popular music decreases the classical rating by 0.08. When people rated themselves higher than 2 in the listening frequency to classical music, the classical rating increases by 0.51 compared to people rated less than or equal to 2. People that took AP music theory class in high school rated higher in classical rating by 0.59 than people that did not take AP music theory class in high school.

There are also several random effects. They account for "personal biases" in ratings. These effects cannot be explained by fixed effects. There are four random effects in the model. They accounts for "personal biases" varied by type of voice, instrument, harmony, and classical music listening frequency. This is not a standard repeated measures model as I included other variance component.

Popular Rating

Among the three main experimental factors, the only statistically significant factor is "instrument". But in order to fit the three random effects Subject & Instrument, Subject & Harmony, and Subject & Voice, the three main effects will be all kept. And the final model indicates that when the instrument is piano, the popular rating decreases by 0.97 compared to when the instrument is guitar. When the instrument is string instrument, the popular rating decreases by 2.64 compared to when the instrument is guitar. When the Harmony is I-V-IV, the popular rating decreases by .04 compared to when the harmony is I-IV-V. When the Harmony is I-V-VI, the popular rating decreases by .27 compared to when the harmony is I-IV-V. When the Harmony is IV-I-V, the popular rating decreases by 0.23 compared to when the harmony is I-IV-V. When the Voice is par 3rd, the popular rating increases by 0.146 compared to when the voice is contrary. When the Voice is par 5rd, the popular rating increases by 0.147 compared to when the voice is contrary.

Some other individual covariates are also useful for predicting popular rating score. When people rated themselves higher than 2 in the listening frequency to pop and rock music from the 90's and 2000's, the popular rating decreases by 0.05 compared to people rated less than or equal to 2. When people rated themselves higher than 2 in the proficiency of guitar play, the popular rating increases by 0.65 compared to people rated less than or equal to 2.

There are also several random effects. They account for "personal biases" in ratings. There are four random effects in the model. They accounts for "personal biases" varied by harmony, voice, instrument, and listening frequency to 1990s and 2000s popular music. This is not a standard repeated measures model as I included other variance component.

Interesting Findings

1. People tend to have high classical rating when the instrument is a string instrument and have high popular rating when the instrument is guitar.
2. When the voice is par 5rd, people tend to have low classical rating and high popular rating.
3. People that took AP music theory class in high school tend to rate classical music higher.
4. People tend to have a lower popular rating when they actually listen to more popular music from 1990s and 2000s.

4: 20
5: 20

40

very nice job