36-763 Homework 5

Cong Shi

December 10, 2013

1. The Three Main Factors

(a)

```
> data=read.csv("ratings.csv",header=T)
> data[is.na(data)]=0
> attach(data)
> m1afull=lm(Classical~Instrument+Harmony+Voice)
> m1ah=lm(Classical~Instrument+Voice)
> m1ai=lm(Classical~Harmony+Voice)
> m1av=lm(Classical~Instrument+Harmony)
> anova(m1ah,m1afull)

Analysis of Variance Table

Model 1: Classical ~ Instrument + Voice
Model 2: Classical ~ Instrument + Harmony + Voice
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2515 14132
2   2512 13871  3    261.26 15.772 3.677e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m1ai,m1afull)

Analysis of Variance Table

Model 1: Classical ~ Harmony + Voice
Model 2: Classical ~ Instrument + Harmony + Voice
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2514 18142
2   2512 13871  2    4271.4 386.78 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m1av,m1afull)
```

```
Analysis of Variance Table

Model 1: Classical ~ Instrument + Harmony
Model 2: Classical ~ Instrument + Harmony + Voice
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2514 13955
2   2512 13871  2    84.813 7.6799 0.0004729 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We set all empty entries to 0 before starting the analysis. From the anova tests, it can be seen that all three main factors are significant.

Table 1: Model Comparison

Model	Adjusted R^2	AIC	BIC
Classical Instrument + Voice	0.2344	11508.36	11543.35
Classical Harmony + Voice	0.01677	12139.85	12180.67
Classical Instrument + Harmony	0.2437	11478.7	11519.52
Classical Instrument + Harmony + Voice	0.2477	11467.34	11519.82

Table 1 demonstrates that the full model is the best model compared to the reduced models under the adjusted R^2, AIC and BIC criteria. The summary of the best linear model is as follows.

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

Residuals:
    Min      1Q  Median      3Q     Max
-6.8627 -1.7222 -0.0164  1.7794 11.4705

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       4.31722    0.13240  32.608  < 2e-16 ***
Instrumentpiano   1.27381    0.11466  11.109  < 2e-16 ***
Instrumentstring  3.16881    0.11466  27.637  < 2e-16 ***
HarmonyI-V-IV    -0.05571    0.13240  -0.421 0.673931
HarmonyI-V-VI     0.73508    0.13240   5.552 3.12e-08 ***
HarmonyIV-I-V     0.04349    0.13240   0.328 0.742565
Voicepar3rd      -0.41393    0.11466  -3.610 0.000312 ***
Voicepar5th      -0.35845    0.11466  -3.126 0.001791 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.35 on 2512 degrees of freedom
Multiple R-squared:  0.2498,        Adjusted R-squared:  0.2477
F-statistic: 119.5 on 7 and 2512 DF,  p-value: < 2.2e-16
```

(b)

i.

The hierarchical linear model can be written as

$$\begin{aligned} y_i &= \alpha_{i[j]} + \alpha_1 + \alpha_2 + \alpha_3 + \epsilon_i, \epsilon_i \overset{i.i.d}{\sim} N(0, \sigma^2) \\ \alpha_i &= \beta + \eta_i, \end{aligned}$$

where y_i stands for the classical rating of each observation, $i = 1, 2, ..., 70$ are the 70 participants and $j = 1, 2, ..., 36$ are the 36 repeated measurements for each participant.α_1, α_2 and α_3 are parameters for Instrument, Harmony and Voice respectively. According to the design of the study, α_1 has 4 levels, while α_2 and α_3 both have 3 levels.

ii.

```
> library(lme4)
> library(arm)
> m1bfull=lmer(Classical~Instrument+Harmony+Voice+(1|Subject))
```

We compare the model with and without random intercept under AIC,BIC and LRT criterion.

Table 2: Model Comparison

Model		AIC	BIC
Classical	Instrument + Harmony + Voice	11467.34	11519.82
Classical	Instrument+Harmony+Voice+(1|Subject)	10755.46	10813.78

It can be seen that the model with repeated measures is preferable under both AIC and BIC.

```
> lrtest(m1afull,m1bfull)

Likelihood ratio test

Model 1: Classical ~ Instrument + Harmony + Voice
Model 2: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1   9 -5724.7
2  10 -5367.7  1 713.88  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The likelihood ratio test also favors the model with repeated measures with p-value<0.0001. Therefore, we conclude that the random effect is needed.

iii.

```
> m1bh=lmer(Classical~Instrument+Voice+(1|Subject))
> m1bi=lmer(Classical~Harmony+Voice+(1|Subject))
> m1bv=lmer(Classical~Instrument+Harmony+(1|Subject))
> anova(m1bh,m1bfull)

Data:
Models:
m1bh: Classical ~ Instrument + Voice + (1 | Subject)
m1bfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1bh     7 10795 10836 -5390.5    10781
m1bfull 10 10733 10792 -5356.6    10713 67.847      3  1.233e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m1bi,m1bfull)

Data:
Models:
m1bi: Classical ~ Harmony + Voice + (1 | Subject)
m1bfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1bi     8 11655 11701 -5819.4    11639
m1bfull 10 10733 10792 -5356.6    10713 925.61      2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m1bv,m1bfull)

Data:
Models:
m1bv: Classical ~ Instrument + Harmony + (1 | Subject)
m1bfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1bv     8 10751 10798 -5367.7    10735
m1bfull 10 10733 10792 -5356.6    10713 22.231      2  1.488e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the anova tests, it can be seen that all three main factors are still significant with random effect presenting.
Thus, a summary of the best model with repeated measures is as follows.

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

REML criterion at convergence: 10735.46

Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 1.732    1.316
 Residual             3.809    1.952
Number of obs: 2520, groups: Subject, 70

Fixed effects:
                 Estimate Std. Error t value
(Intercept)       4.31722    0.19194   22.49
Instrumentpiano   1.27381    0.09523   13.38
Instrumentstring  3.16881    0.09523   33.28
HarmonyI-V-IV    -0.05571    0.10996   -0.51
HarmonyI-V-VI     0.73508    0.10996    6.69
HarmonyIV-I-V     0.04349    0.10996    0.40
Voicepar3rd      -0.41393    0.09523   -4.35
Voicepar5th      -0.35845    0.09523   -3.76

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r
Instrumntpn -0.248
Instrmntstr -0.248  0.500
HrmnyI-V-IV -0.286  0.000     0.000
HrmnyI-V-VI -0.286  0.000     0.000     0.500
HrmnyIV-I-V -0.286  0.000     0.000     0.500  0.500
Voicepar3rd -0.248  0.000     0.000     0.000  0.000  0.000
Voicepar5th -0.248  0.000     0.000     0.000  0.000  0.000  0.500
```

(c)

i.

```
> m1cfull=lmer(Classical~Instrument+Harmony+Voice+(1|Subject:Instrument)+
+                (1|Subject:Harmony)+(1|Subject:Voice))
```

Table 3: Model Comparison

Model	AIC	BIC	DIC
Linear Model	11467.34	11519.82	-
Model with Single Ramdon Intercept	10755.46	10813.78	10690.9
Model with Combined Random Intercept	10307.76	10377.75	10247.9

Table 3 indicates the AIC, BIC and DIC of each model. It can be seen that the model with combined random effect term is preferable.

ii.

```
> m1ch=lmer(Classical~Instrument+Voice+(1|Subject:Instrument)+
+             (1|Subject:Harmony)+(1|Subject:Voice))
> m1ci=lmer(Classical~Harmony+Voice+(1|Subject:Instrument)+
+             (1|Subject:Harmony)+(1|Subject:Voice))
> m1cv=lmer(Classical~Instrument+Harmony+(1|Subject:Instrument)+
+             (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m1ch,m1cfull)

Data:
Models:
m1ch: Classical ~ Instrument + Voice + (1 | Subject:Instrument) + (1 |
m1ch:     Subject:Harmony) + (1 | Subject:Voice)
m1cfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m1cfull:     (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1ch     9 10322 10374 -5152.0    10304
m1cfull 12 10290 10360 -5132.9    10266 38.146      3  2.632e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m1ci,m1cfull)

Data:
Models:
m1ci: Classical ~ Harmony + Voice + (1 | Subject:Instrument) + (1 |
m1ci:     Subject:Harmony) + (1 | Subject:Voice)
m1cfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m1cfull:     (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1ci    10 10390 10448 -5184.8    10370
m1cfull 12 10290 10360 -5132.9    10266 103.69      2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m1cv,m1cfull)

Data:
Models:
m1cv: Classical ~ Instrument + Harmony + (1 | Subject:Instrument) +
m1cv:     (1 | Subject:Harmony) + (1 | Subject:Voice)
```

```
m1cfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m1cfull:     (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1cv    10 10314 10372 -5147.0    10294
m1cfull 12 10290 10360 -5132.9    10266 28.094      2  7.933e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the anova tests, it can be seen that all three main factors are still significant with combined random effect presenting.

iii.

The hierarchical linear model can be written as

$$\begin{aligned} y_i &= \alpha_{1i[j]} + \alpha_{2m[j]} + \alpha_{3n[j]} + \epsilon_i, \epsilon_i \overset{i.i.d}{\sim} N(0, \sigma^2) \\ \alpha_{1i} &= \beta_1 + \eta_{1i[j]}, \eta_{1i[j]} \overset{i.i.d}{\sim} N(0, \tau_1^2) \\ \alpha_{2m} &= \beta_2 + \eta_{2m[j]}, \eta_{2m[j]} \overset{i.i.d}{\sim} N(0, \tau_2^2) \\ \alpha_{3n} &= \beta_3 + \eta_{3n[j]}, \eta_{3n[j]} \overset{i.i.d}{\sim} N(0, \tau_3^2) \end{aligned}$$

where y_i stands for the classical rating of each observation, $i = 1, 2, ..., 70$ are the 70 participants and $j = 1, 2, ..., 36$ are the 36 repeated measurements for each participant.α_1, α_2 and α_3 are parameters for Instrument, Harmony and Voice respectively. According to the design of the study, β_1 has 4 levels, while β_2 and β_3 both have 3 levels.

2. Individual Covariates

(a)

The best model from section 1 is the model with combined random effect.

1. Selfdeclare

```
> m21=lmer(Classical~Instrument+Harmony+Voice+factor(Selfdeclare)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m1cfull,m21)
```

```
Data:
Models:
m1cfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m1cfull:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m21: Classical ~ Instrument + Harmony + Voice + factor(Selfdeclare) +
m21:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
```

```
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1cfull 12 10290 10360 -5132.9    10266
m21     17 10295 10394 -5130.5    10261 4.7902      5      0.442
```

The p-value is 0.44, so Selfdeclare is insignificant and we drop it.

2. OMSI

```
> m22=lmer(Classical~Instrument+Harmony+Voice+OMSI+(1|Subject:Instrument)+
+                (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m1cfull,m22)

Data:
Models:
m1cfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m1cfull:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m22: Classical ~ Instrument + Harmony + Voice + OMSI + (1 | Subject:Instrument) +
m22:     (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1cfull 12 10290 10360 -5132.9    10266
m22     13 10292 10368 -5132.9    10266 0.0088      1     0.9254
```

The p-value is 0.93, so OMSI is highly insignificant.

3. Instr.minus.Notes

```
> m25=lmer(Classical~Instrument+Harmony+Voice+Instr.minus.Notes+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m1cfull,m25)

Data:
Models:
m1cfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m1cfull:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m25: Classical ~ Instrument + Harmony + Voice + Instr.minus.Notes +
m25:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1cfull 12 10290 10360 -5132.9    10266
m25     13 10291 10366 -5132.4    10265 1.1403      1     0.2856
```

The p-value is 0.28, so Instr.minus.Notes is insignificant and we drop it.

4. X16.minus.17

```
> m22=lmer(Classical~Instrument+Harmony+Voice+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m1cfull,m22)

Data:
Models:
m1cfull: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m1cfull:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m22: Classical ~ Instrument + Harmony + Voice + factor(X16.minus.17) +
m22:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m1cfull 12 10290 10360 -5132.9    10266
m22     24 10282 10422 -5117.1    10234 31.706     12   0.001536 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.002, so X16.minus.17 is highly significant and we include it in our model.

5. ConsInstr

```
> m23=lmer(Classical~Instrument+Harmony+Voice+factor(ConsInstr)+
+            factor(X16.minus.17)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m22,m23)

Data:
Models:
m22: Classical ~ Instrument + Harmony + Voice + factor(X16.minus.17) +
m22:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m23: Classical ~ Instrument + Harmony + Voice + factor(ConsInstr) +
m23:     factor(X16.minus.17) + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
m23:     (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m22 24 10282 10422 -5117.1    10234
m23 37 10290 10506 -5108.2    10216 17.784     13     0.1659
```

The p-value is 0.17, so ConsInstr is insignificant and we drop it.

6. ConsNotes

```
> m24=lmer(Classical~Instrument+Harmony+Voice+factor(ConsNotes)+
+            factor(X16.minus.17)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m22,m24)
```

```
Data:
Models:
m22: Classical ~ Instrument + Harmony + Voice + factor(X16.minus.17) +
m22:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m24: Classical ~ Instrument + Harmony + Voice + factor(ConsNotes) +
m24:     factor(X16.minus.17) + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
m24:     (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m22 24 10282 10422 -5117.1    10234
m24 28 10279 10442 -5111.4    10223 11.444      4      0.022 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.02, so ConsNotes is significant and we include it in the model.

7. PachListen

```
> m26=lmer(Classical~Instrument+Harmony+Voice+factor(PachListen)+
+            factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+
+            (1|Subject:Voice))
> anova(m24,m26)
```

```
Data:
Models:
m24: Classical ~ Instrument + Harmony + Voice + factor(ConsNotes) +
m24:     factor(X16.minus.17) + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
m24:     (1 | Subject:Voice)
m26: Classical ~ Instrument + Harmony + Voice + factor(PachListen) +
m26:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m26:     (1 | Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m24 28 10279 10442 -5111.4    10223
m26 33 10271 10463 -5102.3    10205 18.086      5    0.00284 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.003, so PachListen is significant and we include it in the model.

8. ClsListen

```
> m27=lmer(Classical~Instrument+Harmony+Voice+factor(ClsListen)+
+            factor(PachListen)+factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m26,m27)
```

```
Data:
Models:
m26: Classical ~ Instrument + Harmony + Voice + factor(PachListen) +
m26:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m26:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m27: Classical ~ Instrument + Harmony + Voice + factor(ClsListen) +
m27:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m27:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m26 33 10271 10463 -5102.3    10205
m27 37 10274 10489 -5099.8    10200 4.9744      4     0.2899
```

The p-value is 0.29, so ClsListen is insignificant and we drop it.

8. KnowRob

```
> m28=lmer(Classical~Instrument+Harmony+Voice+factor(KnowRob)+
+            factor(PachListen)+factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+
+            (1|Subject:Voice))
> anova(m26,m28)

Data:
Models:
m26: Classical ~ Instrument + Harmony + Voice + factor(PachListen) +
m26:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m26:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m28: Classical ~ Instrument + Harmony + Voice + factor(KnowRob) +
m28:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m28:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m26 33 10271 10463 -5102.3    10205
m28 35 10273 10477 -5101.5    10203 1.5602      2     0.4584
```

The p-value is 0.46, so KnowRob is insignificant and we drop it.

9. KnowAxis

```
> m29=lmer(Classical~Instrument+Harmony+Voice+factor(KnowAxis)+
+            factor(PachListen)+factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+
+            (1|Subject:Voice))
> anova(m26,m29)

Data:
Models:
```

```
m26: Classical ~ Instrument + Harmony + Voice + factor(PachListen) +
m26:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m26:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m29: Classical ~ Instrument + Harmony + Voice + factor(KnowAxis) +
m29:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m29:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m26 33 10271 10463 -5102.3    10205
m29 35 10273 10478 -5101.7    10203 1.2383      2     0.5384
```

The p-value is 0.54, so KnowAxis is insignificant and we drop it.

10. X1990s2000s

```
> m210=lmer(Classical~Instrument+Harmony+Voice+factor(X1990s2000s)+
+             factor(PachListen)+factor(ConsNotes)+factor(X16.minus.17)+
+             (1|Subject:Instrument)+(1|Subject:Harmony)+
+             (1|Subject:Voice))
> anova(m26,m210)
```

```
Data:
Models:
m26: Classical ~ Instrument + Harmony + Voice + factor(PachListen) +
m26:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m26:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m210: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s) +
m210:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m210:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m26  33 10271 10463 -5102.3    10205
m210 37 10278 10493 -5101.8    10204 1.0959      4     0.8949
```

The p-value is 0.89, so X1990s2000s is insignificant and we drop it.

11. X1990s2000s.minus.1960s1970s

```
> m211=lmer(Classical~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+             factor(PachListen)+factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m26,m211)
```

```
Data:
Models:
m26: Classical ~ Instrument + Harmony + Voice + factor(PachListen) +
m26:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
```

```
m26:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m26  33 10271 10463 -5102.3    10205
m211 41 10265 10504 -5091.3    10183 22.066      8   0.004795 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.005, so X1990s2000s.minus.1960s1970s is significant and we include it in the model.

12. CollegeMusic

```
> m212=lmer(Classical~Instrument+Harmony+Voice+factor(CollegeMusic)+
+             factor(PachListen)+factor(X1990s2000s.minus.1960s1970s)+
+             factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m211,m212)

Data:
Models:
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m212: Classical ~ Instrument + Harmony + Voice + factor(CollegeMusic) +
m212:     factor(PachListen) + factor(X1990s2000s.minus.1960s1970s) +
m212:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m212:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m211 41 10265 10504 -5091.3    10183
m212 42 10266 10511 -5091.1    10182 0.3292      1     0.5661
```

The p-value is 0.57, so CollegeMusic is insignificant and we drop it.

13. NoClass

```
> m213=lmer(Classical~Instrument+Harmony+Voice+factor(NoClass)+
+             factor(PachListen)+factor(X1990s2000s.minus.1960s1970s)+
+             factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m211,m213)

Data:
Models:
```

```
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m213: Classical ~ Instrument + Harmony + Voice + factor(NoClass) +
m213:     factor(PachListen) + factor(X1990s2000s.minus.1960s1970s) +
m213:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m213:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m211 41 10265 10504 -5091.3    10183
m213 46 10270 10539 -5089.3    10178 4.0202      5     0.5465
```

The p-value is 0.55, so NoClass is insignificant and we drop it.

14. APTheory

```
> m214=lmer(Classical~Instrument+Harmony+Voice+factor(APTheory)+
+             factor(PachListen)+factor(X1990s2000s.minus.1960s1970s)+
+             factor(ConsNotes)+factor(X16.minus.17)+
+             (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m211,m214)

Data:
Models:
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m214: Classical ~ Instrument + Harmony + Voice + factor(APTheory) +
m214:     factor(PachListen) + factor(X1990s2000s.minus.1960s1970s) +
m214:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m214:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m211 41 10265 10504 -5091.3    10183
m214 42 10266 10511 -5091.0    10182 0.5554      1     0.4561
```

The p-value is 0.46, so APTheory is insignificant and we drop it.

15. Composing

```
> m215=lmer(Classical~Instrument+Harmony+Voice+factor(Composing)+
+             factor(PachListen)+factor(X1990s2000s.minus.1960s1970s)+
+             factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m211,m215)

Data:
Models:
```

```
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m215: Classical ~ Instrument + Harmony + Voice + factor(Composing) +
m215:     factor(PachListen) + factor(X1990s2000s.minus.1960s1970s) +
m215:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m215:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m211 41 10265 10504 -5091.3    10183
m215 46 10265 10533 -5086.6    10173 9.3985      5    0.09419 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.09, so Composing is insignificant and we drop it.

16. PianoPlay

```
> m216=lmer(Classical~Instrument+Harmony+Voice+factor(PianoPlay)+
+             factor(PachListen)+factor(X1990s2000s.minus.1960s1970s)+
+             factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m211,m216)

Data:
Models:
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m216: Classical ~ Instrument + Harmony + Voice + factor(PianoPlay) +
m216:     factor(PachListen) + factor(X1990s2000s.minus.1960s1970s) +
m216:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m216:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m211 41 10265 10504 -5091.3    10183
m216 45 10267 10529 -5088.5    10177 5.5915      4     0.2318
```

The p-value is 0.23, so PianoPlay is insignificant and we drop it.

17. GuitarPlay

```
> m217=lmer(Classical~Instrument+Harmony+Voice+factor(GuitarPlay)+
+             factor(PachListen)+factor(X1990s2000s.minus.1960s1970s)+
+             factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m211,m217)
```

```
Data:
Models:
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m217: Classical ~ Instrument + Harmony + Voice + factor(GuitarPlay) +
m217:     factor(PachListen) + factor(X1990s2000s.minus.1960s1970s) +
m217:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m217:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m211 41 10265 10504 -5091.3    10183
m217 45 10271 10533 -5090.3    10181 1.9451      4     0.7459
```

The p-value is 0.75, so GuitarPlay is insignificant and we drop it.

18. X1stInstr

```
> m218=lmer(Classical~Instrument+Harmony+Voice+factor(X1stInstr)+
+             factor(PachListen)+factor(X1990s2000s.minus.1960s1970s)+
+             factor(ConsNotes)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m211,m218)

Data:
Models:
m211: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m211:     factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) +
m211:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m218: Classical ~ Instrument + Harmony + Voice + factor(X1stInstr) +
m218:     factor(PachListen) + factor(X1990s2000s.minus.1960s1970s) +
m218:     factor(ConsNotes) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m218:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m211 41 10265 10504 -5091.3    10183
m218 46 10265 10533 -5086.5    10173 9.4661      5    0.09186 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.09, so X1stInstr is insignificant and we drop it.
Thus, the variables we include in the model are $16.minus.17$, $PachListen$ and $X1990s2000s.minus.X1960s197$

(b)

Acording to what we have found in 2(a), the current best model is

$$Classical \sim Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +$$

$$factor(PachListen) + factor(X16.minus.17) +$$
$$(1|Subject : Instrument) + (1|Subject : Harmony) + (1|Subject : Voice))$$

```
> m2b=lmer(Classical~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+             +factor(PachListen)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> m2bi=lmer(Classical~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+            factor(PachListen)+factor(X16.minus.17)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> m2bh=lmer(Classical~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+            factor(PachListen)+factor(X16.minus.17)+
+            (1|Subject:Instrument)+(1|Subject:Voice))
> m2bv=lmer(Classical~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+            factor(PachListen)+factor(X16.minus.17)+
+            (1|Subject:Harmony)+(1|Subject:Instrument))
> anova(m2bh,m2b);anova(m2bi,m2b);anova(m2bv,m2b)

Data:
Models:
m2bh: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m2bh:     factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m2bh:     (1 | Subject:Voice)
m2b: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m2b:     +factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m2b:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m2bh 36 10352 10562 -5139.9    10280
m2b  37 10272 10487 -5098.7    10198 82.266      1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Data:
Models:
m2bi: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m2bi:     factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Harmony) +
m2bi:     (1 | Subject:Voice)
m2b: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m2b:     +factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m2b:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m2bi 36 10822 11032 -5375.1    10750
m2b  37 10272 10487 -5098.7    10198 552.62      1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Data:
Models:
m2bv: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m2bv:     factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Harmony) +
m2bv:     (1 | Subject:Instrument)
m2b: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m2b:     +factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
m2b:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m2bv 36 10270 10480 -5098.8    10198
m2b  37 10272 10487 -5098.7    10198 0.0439      1      0.834
```

According to the anova tests, both the combined random effects of Instrument and Harmony are significant, while the combined random effect of Voice is insignificant.

```
> lrtest(m2bh,m2b);lrtest(m2bi,m2b);lrtest(m2bv,m2b)

Likelihood ratio test

Model 1: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
    factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
    (1 | Subject:Voice)
Model 2: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
    +factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
    (1 | Subject:Harmony) + (1 | Subject:Voice)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  36 -5142.1
2  37 -5099.4  1 85.515  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Likelihood ratio test

Model 1: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
    factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Harmony) +
    (1 | Subject:Voice)
Model 2: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
    +factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
    (1 | Subject:Harmony) + (1 | Subject:Voice)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  36 -5384.7
2  37 -5099.4  1 570.72  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Likelihood ratio test
```

```
Model 1: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
    factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Harmony) +
    (1 | Subject:Instrument)
Model 2: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
    +factor(PachListen) + factor(X16.minus.17) + (1 | Subject:Instrument) +
    (1 | Subject:Harmony) + (1 | Subject:Voice)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  36 -5099.5
2  37 -5099.4  1 0.1544     0.6943
```

The result of likelihood ratio test is consistent to that of anova test. Therefore, we should drop the term $(1|Subject:Voice)$ from the model.

(c)

The best model we have is

$$\begin{aligned} Classical \sim\ & Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) + \\ & factor(PachListen) + factor(ConsNotes) + factor(X16.minus.17) + \\ & (1|Subject:Instrument) + (1|Subject:Harmony) \end{aligned}$$

A summary of the mdel is as follows.

```
> m2full=lmer(Classical~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+             factor(PachListen)+factor(ConsNotes)+factor(X16.minus.17)+
+               (1|Subject:Instrument)+(1|Subject:Harmony))
> summary(m2full)$coe
```

```
                                      Estimate Std. Error     t value
(Intercept)                         1.56901093 1.53053284  1.02514033
Instrumentpiano                     1.27380952 0.24169308  5.27035994
Instrumentstring                    3.16880952 0.24169308 13.11088233
HarmonyI-V-IV                      -0.05571429 0.13997458 -0.39803146
HarmonyI-V-VI                       0.73507937 0.13997458  5.25152053
HarmonyIV-I-V                       0.04349206 0.13997458  0.31071402
Voicepar3rd                        -0.41392857 0.07868428 -5.26062625
Voicepar5th                        -0.35845238 0.07868428 -4.55557827
factor(X1990s2000s.minus.1960s1970s)-3 -0.14002835 1.32365518 -0.10578915
factor(X1990s2000s.minus.1960s1970s)-2 -2.44444444 1.24957718 -1.95621726
factor(X1990s2000s.minus.1960s1970s)0   1.17817356 0.97336334  1.21041497
factor(X1990s2000s.minus.1960s1970s)1   1.54499773 1.18650795  1.30213854
factor(X1990s2000s.minus.1960s1970s)2   1.34036213 0.97543346  1.37411950
factor(X1990s2000s.minus.1960s1970s)3   0.78161930 0.94364649  0.82829673
factor(X1990s2000s.minus.1960s1970s)4   1.15086550 1.21378433  0.94816308
factor(X1990s2000s.minus.1960s1970s)5   0.46242852 1.06323551  0.43492577
```

```
factor(PachListen)1                    0.69877160 1.11046837  0.62925845
factor(PachListen)2                    0.42392771 0.77474386  0.54718435
factor(PachListen)3                   -1.74798413 0.73726602 -2.37090017
factor(PachListen)4                   -0.40008279 1.10125774 -0.36329623
factor(PachListen)5                    0.36754293 0.58813484  0.62492970
factor(ConsNotes)1                     0.77968455 0.41770120  1.86660836
factor(ConsNotes)3                     0.20784590 0.31931576  0.65091026
factor(ConsNotes)4                    -2.65462878 0.93917011 -2.82656864
factor(ConsNotes)5                    -0.09523170 0.35367401 -0.26926405
factor(X16.minus.17)-2                 2.31990011 1.41532984  1.63912329
factor(X16.minus.17)-1                 1.17670486 0.97340925  1.20884907
factor(X16.minus.17)-0.5               3.51830483 1.35681780  2.59305622
factor(X16.minus.17)0                  2.33788418 0.97870593  2.38875039
factor(X16.minus.17)1                  2.40907509 1.02254282  2.35596499
factor(X16.minus.17)2                  1.92235698 1.01345008  1.89684428
factor(X16.minus.17)3                  0.73975032 1.03089635  0.71757972
factor(X16.minus.17)4                  0.76386380 1.10072988  0.69396117
factor(X16.minus.17)5                 -0.12247115 1.06189128 -0.11533304
factor(X16.minus.17)6                 -0.64789918 1.29697678 -0.49954571
factor(X16.minus.17)7                 -0.08330306 1.18749302 -0.07015036
factor(X16.minus.17)9                  0.89343333 1.07563370  0.83061114
```

The coefficient of a categorical variable means difference between adjacent groups (note that the categorical variables are ordinal); the coefficient of a numerical variable means one unit increase in the variable would cause the corresponding unit of change(depend on the sign of the coefficient) in the response variable.

3. Musicians vs. Non-musicians

```
> table(Selfdeclare)

Selfdeclare
  1   2   3   4   5   6
576 936 468 432  72  36
```

According to the distribution of Selfdeclare, we categorize "1" and "2" into non-musician and "3", "4" and "5" into musician. An indicator variable musician is created based on the categorization.

```
> anova(m3h,m3)

Data:
Models:
m3: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3:     PachListen + +ConsNotes + X16.minus.17 + musician + (1 |
m3:     Subject:Instrument) + (1 | Subject:Harmony)
```

```
m3h: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3h:     PachListen + +ConsNotes + X16.minus.17 + musician:Harmony +
m3h:     (1 | Subject:Instrument) + (1 | Subject:Harmony)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m3  16 10286 10379 -5127.0    10254
m3h 19 10266 10377 -5113.9    10228 26.079      3   9.18e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P-value<0.001, hence, musician:Harmony is significant.

```
> anova(m3i,m3)

Data:
Models:
m3: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3:     PachListen + +ConsNotes + X16.minus.17 + musician + (1 |
m3:     Subject:Instrument) + (1 | Subject:Harmony)
m3i: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3i:     PachListen + X16.minus.17 + musician:Instrument + (1 | Subject:Instrument) +
m3i:     (1 | Subject:Harmony)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m3  16 10286 10379 -5127.0    10254
m3i 17 10287 10386 -5126.7    10253 0.5771      1     0.4475
```

P-value=0.45, hence, musician:Instrument is insignificant.

```
> anova(m3v,m3)

Data:
Models:
m3: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3:     PachListen + +ConsNotes + X16.minus.17 + musician + (1 |
m3:     Subject:Instrument) + (1 | Subject:Harmony)
m3v: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3v:     PachListen + +ConsNotes + X16.minus.17 + musician:Voice +
m3v:     (1 | Subject:Instrument) + (1 | Subject:Harmony)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m3  16 10286 10379 -5127.0    10254
m3v 18 10290 10394 -5126.8    10254 0.4254      2     0.8084
```

P-value=0.81, hence, musician:Voice is insignificant.

```
> anova(m3x1990,m3)
```

```
Data:
Models:
m3x1990: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3x1990:     PachListen + ConsNotes + X16.minus.17 + musician:X1990s2000s.minus.1960s1970s
m3x1990:     (1 | Subject:Instrument) + (1 | Subject:Harmony)
m3: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3:     PachListen + +ConsNotes + X16.minus.17 + musician + (1 |
m3:     Subject:Instrument) + (1 | Subject:Harmony)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m3x1990 16 10286 10380 -5127.2    10254
m3      16 10286 10379 -5127.0    10254 0.4852      0  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P-value<0.001, hence, musician:X1990s2000s.minus.X1960s1970s is significant.

```
> anova(m3p,m3)
```

```
Data:
Models:
m3: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3:     PachListen + +ConsNotes + X16.minus.17 + musician + (1 |
m3:     Subject:Instrument) + (1 | Subject:Harmony)
m3p: Classical ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m3p:     PachListen + ConsNotes + X16.minus.17 + musician:PachListen +
m3p:     (1 | Subject:Instrument) + (1 | Subject:Harmony)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m3  16 10286 10379 -5127.0    10254
m3p 23 10288 10422 -5120.8    10242 12.297      7    0.09121 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P-value=0.09, hence, musician:PachListen is insignificant.

```
> anova(m3c,m3)
```

```
Data:
Models:
m3c: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3c:     PachListen + ConsNotes + X16.minus.17 + musician:ConsNotes +
m3c:     (1 | Subject:Instrument) + (1 | Subject:Harmony)
m3: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3:     PachListen + +ConsNotes + X16.minus.17 + musician + (1 |
m3:     Subject:Instrument) + (1 | Subject:Harmony)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m3c 16 10286 10380 -5127.1    10254
```

```
m3  16 10286 10379 -5127.0    10254 0.2292      0  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P-value<0.001, hence, musician:ConsNotes is significant.

```
> anova(m3x16,m3)

Data:
Models:
m3x16: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3x16:     PachListen + ConsNotes + X16.minus.17 + musician:X16.minus.17 +
m3x16:     (1 | Subject:Instrument) + (1 | Subject:Harmony)
m3: Classical ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m3:     PachListen + +ConsNotes + X16.minus.17 + musician + (1 |
m3:     Subject:Instrument) + (1 | Subject:Harmony)
      Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
m3x16 16 10283 10377 -5125.7    10251
m3    16 10286 10379 -5127.0    10254     0      0          1
```

P-value=1, hence, musician:X16.minus.17 is insignificant.

Thus, we include the random effects of $musician : Harmony$, $musician : X1990s2000s.minus.X1960s1970s$ and $musician : ConsNotes$.

4. Classical vs. Popular

(a)

i. Linear Model

```
> m4a1full=lm(Popular~Instrument+Harmony+Voice)
> m4a1h=lm(Popular~Instrument+Voice)
> m4a1i=lm(Popular~Harmony+Voice)
> m4a1v=lm(Popular~Instrument+Harmony)
> anova(m4a1h,m4a1full)

Analysis of Variance Table

Model 1: Popular ~ Instrument + Voice
Model 2: Popular ~ Instrument + Harmony + Voice
  Res.Df   RSS Df Sum of Sq      F  Pr(>F)
1   2515 13605
2   2512 13571  3    33.879 2.0903 0.09943 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> anova(m4a1i,m4a1full)

Analysis of Variance Table

Model 1: Popular ~ Harmony + Voice
Model 2: Popular ~ Instrument + Harmony + Voice
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2514 16351
2   2512 13571  2    2780.4 257.33 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m4a1v,m4a1full)

Analysis of Variance Table

Model 1: Popular ~ Instrument + Harmony
Model 2: Popular ~ Instrument + Harmony + Voice
  Res.Df   RSS Df Sum of Sq      F Pr(>F)
1   2514 13586
2   2512 13571  2    15.033 1.3913  0.249
```

From the anova tests, it can be seen that only Instrument(p-vaule<0.0001) is significant, Harmony(p-value=0.099) and Voice(p-value=0.25) are insignificant.

ii. Single Random Effect Model

```
> m4a2full=lmer(Popular~Instrument+Harmony+Voice+(1|Subject))
> m4a2h=lmer(Popular~Instrument+Voice+(1|Subject))
> m4a2i=lmer(Popular~Harmony+Voice+(1|Subject))
> m4a2v=lmer(Popular~Instrument+Harmony+(1|Subject))
> anova(m4a2h,m4a2full)

Data:
Models:
m4a2h: Popular ~ Instrument + Voice + (1 | Subject)
m4a2full: Popular ~ Instrument + Harmony + Voice + (1 | Subject)
         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4a2h     7 10704 10745 -5345.1    10690
m4a2full 10 10701 10759 -5340.6    10681 9.0108      3    0.02915 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(m4a2i,m4a2full)
```

```
Data:
Models:
m4a2i: Popular ~ Harmony + Voice + (1 | Subject)
m4a2full: Popular ~ Instrument + Harmony + Voice + (1 | Subject)
         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4a2i     8 11344 11391 -5664.2    11328
m4a2full 10 10701 10759 -5340.6    10681 647.31      2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> anova(m4a2v,m4a2full)
```

```
Data:
Models:
m4a2v: Popular ~ Instrument + Harmony + (1 | Subject)
m4a2full: Popular ~ Instrument + Harmony + Voice + (1 | Subject)
         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4a2v     8 10701 10748 -5342.6    10685
m4a2full 10 10701 10759 -5340.6    10681 4.0023      2     0.1352
```

From the anova tests, it can be seen that only Instrument(p-value<0.0001) and Harmony(p-value=0.03) are significant, Voice(p-value=0.14) is insignificant.

iii. Combined Random Effect Model

```
> m4a3full=lmer(Popular~Instrument+Harmony+Voice+(1|Subject:Instrument)+
+                 (1|Subject:Harmony)+(1|Subject:Voice))
> m4a3h=lmer(Popular~Instrument+Voice+(1|Subject:Instrument)+
+               (1|Subject:Harmony)+(1|Subject:Voice))
> m4a3i=lmer(Popular~Harmony+Voice+(1|Subject:Instrument)+
+               (1|Subject:Harmony)+(1|Subject:Voice))
> m4a3v=lmer(Popular~Instrument+Harmony+(1|Subject:Instrument)+
+               (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m4a3h,m4a3full)
```

```
Data:
Models:
m4a3h: Popular ~ Instrument + Voice + (1 | Subject:Instrument) + (1 |
m4a3h:     Subject:Harmony) + (1 | Subject:Voice)
m4a3full: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m4a3full:     (1 | Subject:Harmony) + (1 | Subject:Voice)
         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4a3h     9 10329 10382 -5155.5    10311
m4a3full 12 10330 10400 -5152.7    10306 5.5989      3     0.1328
```

```
> anova(m4a3i,m4a3full)
```

```
Data:
Models:
m4a3i: Popular ~ Harmony + Voice + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
m4a3i:     (1 | Subject:Voice)
m4a3full: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m4a3full:     (1 | Subject:Harmony) + (1 | Subject:Voice)
         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4a3i    10 10403 10461 -5191.5    10383
m4a3full 12 10330 10400 -5152.7    10306 77.543      2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> anova(m4a3v,m4a3full)
```

```
Data:
Models:
m4a3v: Popular ~ Instrument + Harmony + (1 | Subject:Instrument) + (1 |
m4a3v:     Subject:Harmony) + (1 | Subject:Voice)
m4a3full: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m4a3full:     (1 | Subject:Harmony) + (1 | Subject:Voice)
         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4a3v    10 10330 10389 -5155.2    10310
m4a3full 12 10330 10400 -5152.7    10306 4.8508      2    0.08844 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the anova tests, it can be seen that only Instrument(p-value<0.0001) is significant, Harmony(p-value=0.13) and Voice(p-value=0.088) are insignificant.

Table 4: Model Comparison

Model	AIC	BIC	DIC
Linear Model	11467.34	11434.7	-
Model with Single Ramdon Intercept	10717.47	10764.13	10668.8
Model with Combined Random Intercept	10334.42	10375.25	10311.5

Hence, the best model to estimate the Popular score is the model with combined random effect.

(b)

1. Selfdeclare

```
> m41=lmer(Popular~Instrument+Harmony+Voice+factor(Selfdeclare)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m4a3full,m41)
```

```
Data:
Models:
m4a3full: Popular ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
m4a3full:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m41: Popular ~ Instrument + Harmony + Voice + factor(Selfdeclare) +
m41:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
         Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4a3full 12 10330 10400 -5152.7    10306
m41      17 10322 10422 -5144.3    10288 16.988      5   0.004522 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.005, so Selfdeclare is significant and we keep it.

2. OMSI

```
> m42=lmer(Popular~Instrument+Harmony+Voice+OMSI+factor(Selfdeclare)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m41,m42)
```

```
Data:
Models:
m41: Popular ~ Instrument + Harmony + Voice + factor(Selfdeclare) +
m41:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m42: Popular ~ Instrument + Harmony + Voice + OMSI + factor(Selfdeclare) +
m42:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
m41 17 10322 10422 -5144.3    10288
m42 18 10319 10424 -5141.3    10283 5.925      1    0.01493 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.01, so OMSI is significant so we keep it.

3. Instr.minus.Notes

```
> m43=lmer(Popular~Instrument+Harmony+Voice+Instr.minus.Notes+
+            OMSI+factor(Selfdeclare)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m42,m43)
```

```
Data:
Models:
m42: Popular ~ Instrument + Harmony + Voice + OMSI + factor(Selfdeclare) +
m42:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
```

```
m43: Popular ~ Instrument + Harmony + Voice + Instr.minus.Notes +
m43:     OMSI + factor(Selfdeclare) + (1 | Subject:Instrument) + (1 |
m43:     Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
m42 18 10319 10424 -5141.3    10283
m43 19 10319 10430 -5140.5    10281  1.61      1     0.2045
```

The p-value is smaller than 0.20, so Instr.minus.Notes is insignificant and we drop it.

4. X16.minus.17

```
> m44=lmer(Popular~Instrument+Harmony+Voice+factor(X16.minus.17)+
+            OMSI+factor(Selfdeclare)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m42,m44)

Data:
Models:
m42: Popular ~ Instrument + Harmony + Voice + OMSI + factor(Selfdeclare) +
m42:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m44: Popular ~ Instrument + Harmony + Voice + factor(X16.minus.17) +
m44:     OMSI + factor(Selfdeclare) + (1 | Subject:Instrument) + (1 |
m44:     Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m42 18 10319 10424 -5141.3    10283
m44 30 10331 10506 -5135.3    10271 11.952     12     0.4496
```

The p-value is 0.45, so X16.minus.17 is insignificant and we drop it.

5. ConsInstr

```
> m45=lmer(Popular~Instrument+Harmony+Voice+ConsInstr+
+         OMSI+factor(Selfdeclare)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m42,m45)

Data:
Models:
m42: Popular ~ Instrument + Harmony + Voice + OMSI + factor(Selfdeclare) +
m42:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m45: Popular ~ Instrument + Harmony + Voice + ConsInstr + OMSI + factor(Selfdeclare) +
m45:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m42 18 10319 10424 -5141.3    10283
m45 19 10320 10431 -5141.0    10282 0.5583      1      0.455
```

The p-value is 0.46, so ConsInstr is insignificant and we drop it.

6. PachListen

```
> m46=lmer(Popular~Instrument+Harmony+Voice+factor(PachListen)+
+            OMSI+factor(Selfdeclare)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m42,m46)

Data:
Models:
m42: Popular ~ Instrument + Harmony + Voice + OMSI + factor(Selfdeclare) +
m42:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m46: Popular ~ Instrument + Harmony + Voice + factor(PachListen) +
m46:     OMSI + factor(Selfdeclare) + (1 | Subject:Instrument) + (1 |
m46:     Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m42 18 10319 10424 -5141.3    10283
m46 23 10321 10455 -5137.3    10275 7.8988      5     0.1619
```

The p-value is 0.16, so PachListen is insignificant and we drop it.

7. ClsListen

```
> m47=lmer(Popular~Instrument+Harmony+Voice+factor(ClsListen)+
+            OMSI+factor(Selfdeclare)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m42,m47)

Data:
Models:
m42: Popular ~ Instrument + Harmony + Voice + OMSI + factor(Selfdeclare) +
m42:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m47: Popular ~ Instrument + Harmony + Voice + factor(ClsListen) +
m47:     OMSI + factor(Selfdeclare) + (1 | Subject:Instrument) + (1 |
m47:     Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m42 18 10319 10424 -5141.3    10283
m47 22 10321 10449 -5138.5    10277 5.5147      4     0.2384
```

The p-value is 0.23, so ClsListen is insignificant and we drop it.

8. KnowRob

```
> m48=lmer(Popular~Instrument+Harmony+Voice+factor(KnowRob)+
+            OMSI+factor(Selfdeclare)+(1|Subject:Instrument)+
+            (1|Subject:Harmony)+(1|Subject:Voice))
> anova(m42,m48)
```

```
Data:
Models:
m42: Popular ~ Instrument + Harmony + Voice + OMSI + factor(Selfdeclare) +
m42:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m48: Popular ~ Instrument + Harmony + Voice + factor(KnowRob) + OMSI +
m48:     factor(Selfdeclare) + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
m48:     (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m42 18 10319 10424 -5141.3    10283
m48 20 10315 10432 -5137.4    10275 7.7595      2    0.02066 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.02, so KnowRob is significant and we keep it.

9. KnowAxis

```
> m49=lmer(Popular~Instrument+Harmony+Voice+factor(KnowAxis)+
+            OMSI+factor(Selfdeclare)+factor(KnowRob)+
+            (1|Subject:Instrument)+(1|Subject:Harmony)+
+            (1|Subject:Voice))
> anova(m48,m49)
```

```
Data:
Models:
m48: Popular ~ Instrument + Harmony + Voice + factor(KnowRob) + OMSI +
m48:     factor(Selfdeclare) + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
m48:     (1 | Subject:Voice)
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m48 20 10315 10432 -5137.4    10275
m49 22 10313 10441 -5134.3    10269 6.2197      2    0.04461 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.45, so KnowAxis is significant and we keep it.

10. X1990s2000s

```
> m410=lmer(Popular~Instrument+Harmony+Voice+factor(X1990s2000s)+
+             OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+             (1|Subject:Instrument)+(1|Subject:Harmony)+
+             (1|Subject:Voice))
> anova(m49,m410)
```

```
Data:
Models:
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m410: Popular ~ Instrument + Harmony + Voice + factor(X1990s2000s) +
m410:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m410:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m49  22 10313 10441 -5134.3    10269
m410 26 10318 10470 -5133.2    10266 2.2986      4      0.681
```

The p-value is 0.68, so X1990s2000s is insignificant and we drop it.

11. X1990s2000s.minus.1960s1970s

```
>   m411=lmer(Popular~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+               OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m49,m411)
```

```
Data:
Models:
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m411: Popular ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m411:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m411:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m49  22 10313 10441 -5134.3    10269
m411 30 10311 10486 -5125.6    10251 17.472      8    0.02556 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.026, so X1990s2000s.minus.1960s1970s is significant and we keep it.

12. CollegeMusic

```
>   m412=lmer(Popular~Instrument+Harmony+Voice+factor(CollegeMusic)+
+               OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
>   anova(m49,m412)
```

```
Data:
Models:
```

```
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m412: Popular ~ Instrument + Harmony + Voice + factor(CollegeMusic) +
m412:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m412:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m49  22 10313 10441 -5134.3    10269
m412 23 10315 10449 -5134.3    10269 0.0083      1     0.9273
```

The p-value is 0.93, so CollegeMusic is insignificant and we drop it.

11. X1990s2000s.minus.1960s1970s

```
>   m411=lmer(Popular~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+               OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m49,m411)

Data:
Models:
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m411: Popular ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m411:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m411:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m49  22 10313 10441 -5134.3    10269
m411 30 10311 10486 -5125.6    10251 17.472      8    0.02556 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.026, so X1990s2000s.minus.1960s1970s is significant and we keep it.

12. CollegeMusic

```
>   m412=lmer(Popular~Instrument+Harmony+Voice+factor(CollegeMusic)+
+               OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
>   anova(m49,m412)

Data:
Models:
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
```

```
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m412: Popular ~ Instrument + Harmony + Voice + factor(CollegeMusic) +
m412:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m412:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m49  22 10313 10441 -5134.3    10269
m412 23 10315 10449 -5134.3    10269 0.0083      1     0.9273
```

The p-value is 0.93, so CollegeMusic is insignificant and we drop it.

13. APTheory

```
>   m413=lmer(Popular~Instrument+Harmony+Voice+factor(APTheory)+
+               OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m41,m413)

Data:
Models:
m41: Popular ~ Instrument + Harmony + Voice + factor(Selfdeclare) +
m41:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
m413: Popular ~ Instrument + Harmony + Voice + factor(APTheory) + OMSI +
m413:     factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m413:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m41  17 10322 10422 -5144.3    10288
m413 23 10314 10449 -5134.3    10268 19.968      6   0.002806 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.80, so APTheory is insignificant and we drop it.

14. PianoPlay

```
>   m416=lmer(Popular~Instrument+Harmony+Voice+factor(PianoPlay)+
+               OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m49,m416)

Data:
Models:
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
```

```
m416: Popular ~ Instrument + Harmony + Voice + factor(PianoPlay) +
m416:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m416:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m49  22 10313 10441 -5134.3    10269
m416 26 10316 10468 -5132.2    10264 4.1872      4     0.3813
```

The p-value is 0.38, so PianoPlay is insignificant and we drop it.

15. GuitarPlay

```
>   m417=lmer(Popular~Instrument+Harmony+Voice+factor(GuitarPlay)+
+               OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m49,m417)

Data:
Models:
m49: Popular ~ Instrument + Harmony + Voice + factor(KnowAxis) + OMSI +
m49:     factor(Selfdeclare) + factor(KnowRob) + (1 | Subject:Instrument) +
m49:     (1 | Subject:Harmony) + (1 | Subject:Voice)
m417: Popular ~ Instrument + Harmony + Voice + factor(GuitarPlay) +
m417:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m417:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m49  22 10313 10441 -5134.3    10269
m417 26 10313 10464 -5130.4    10261 7.8771      4    0.09619 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.10, so GuitarPlay is insignificant and we drop it.

16. X1stInstr

```
>   m418=lmer(Popular~Instrument+Harmony+Voice+factor(X1stInstr)+
+               factor(X1990s2000s.minus.1960s1970s)+OMSI+
+               factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+               (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> anova(m411,m418)

Data:
Models:
m411: Popular ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m411:     OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) +
m411:     (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
```

```
m418: Popular ~ Instrument + Harmony + Voice + factor(X1stInstr) +
m418:     factor(X1990s2000s.minus.1960s1970s) + OMSI + factor(Selfdeclare) +
m418:     factor(KnowRob) + factor(KnowAxis) + (1 | Subject:Instrument) +
m418:     (1 | Subject:Harmony) + (1 | Subject:Voice)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m411 30 10311 10486 -5125.6    10251
m418 35 10311 10515 -5120.5    10241 10.118      5    0.07195 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is 0.07, so X1stInstr is insignificant and we drop it.
Thus, the variables we include in the model are $Selfdeclare$, $OMSI$, $KnowRob$, $KnowAxis$ and $X1990s2000s.minus.X1960s1970s$.

Acording to what we have found in 4(a), the current best model is

$$\begin{aligned} Popular \sim\ & Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) + \\ & OMSI + factor(Selfdeclare) + factor(KnowRob) + factor(KnowAxis) + \\ & (1|Subject:Instrument) + (1|Subject:Harmony) + (1|Subject:Voice)) \end{aligned}$$

A summaryof the best model is as follows.

```
> m4b=lmer(Popular~Instrument+Harmony+Voice+factor(X1990s2000s.minus.1960s1970s)+
+          OMSI+factor(Selfdeclare)+factor(KnowRob)+factor(KnowAxis)+
+         (1|Subject:Instrument)+(1|Subject:Harmony)+(1|Subject:Voice))
> summary(m4b)$coe
```

```
                                          Estimate  Std. Error       t value
(Intercept)                             4.399735578 1.069717155   4.112989642
Instrumentpiano                        -1.030952381 0.242269974  -4.255386521
Instrumentstring                       -2.557023810 0.242269974 -10.554439617
HarmonyI-V-IV                          -0.052063492 0.137976849  -0.377334984
HarmonyI-V-VI                          -0.293492063 0.137976849  -2.127110932
HarmonyIV-I-V                          -0.194603175 0.137976849  -1.410404544
Voicepar3rd                             0.159642857 0.085434166   1.868606719
Voicepar5th                             0.167738095 0.085434166   1.963360825
factor(X1990s2000s.minus.1960s1970s)-3 -0.755700858 1.466618917  -0.515267360
factor(X1990s2000s.minus.1960s1970s)-2  2.018151299 1.409708396   1.431609051
factor(X1990s2000s.minus.1960s1970s)0   1.648454090 1.074425702   1.534265316
factor(X1990s2000s.minus.1960s1970s)1   0.550319450 1.174964911   0.468370966
factor(X1990s2000s.minus.1960s1970s)2   0.894376389 1.119311609   0.799041466
factor(X1990s2000s.minus.1960s1970s)3   1.026862565 1.072263829   0.957658496
factor(X1990s2000s.minus.1960s1970s)4   0.811619460 1.261644616   0.643302757
factor(X1990s2000s.minus.1960s1970s)5   1.531763758 1.115691110   1.372928174
OMSI                                    0.001272357 0.001058718   1.201790907
```

```
factor(Selfdeclare)2                  0.875043662 0.337719305   2.591038326
factor(Selfdeclare)3                  0.903840526 0.364973789   2.476453249
factor(Selfdeclare)4                 -0.065846278 0.633893898  -0.103875867
factor(Selfdeclare)5                  1.014462548 0.860537195   1.178871238
factor(Selfdeclare)6                  0.120996398 1.390801855   0.086997582
factor(KnowRob)1                      0.624525619 0.529511896   1.179436427
factor(KnowRob)5                      0.788300438 0.454692356   1.733700662
factor(KnowAxis)1                     2.469314734 1.003900457   2.459720700
factor(KnowAxis)5                    -0.001157297 0.360795990  -0.003207621
```

The coefficient of a categorical variable means difference between adjacent groups (note that the categorical variables are ordinal); the coefficient of a numerical variable means one unit increase in the variable would cause the corresponding unit of change(depend on the sign of the coefficient) in the response variable.

(c)

Models are built accordingly, and we examine their anova test results.

```
> anova(m4ch,m4c)

Data:
Models:
m4c: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m4c:     OMSI + KnowRob + KnowAxis + musician + (1 | Subject:Instrument) +
m4c:     (1 | Subject:Harmony)
m4ch: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m4ch:     OMSI + KnowRob + KnowAxis + musician:Harmony + (1 | Subject:Instrument) +
m4ch:     (1 | Subject:Harmony)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4c  25 10313 10459 -5131.7    10263
m4ch 28 10307 10470 -5125.4    10251 12.444      3   0.006008 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P-value=0.006, hence, musician:Harmony is significant.

```
> anova(m4ci,m4c)

Data:
Models:
m4c: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m4c:     OMSI + KnowRob + KnowAxis + musician + (1 | Subject:Instrument) +
m4c:     (1 | Subject:Harmony)
m4ci: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
```

```
m4ci:     OMSI + KnowRob + KnowAxis + musician:Instrument + (1 | Subject:Instrument) +
m4ci:     (1 | Subject:Harmony)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4c  25 10313 10459 -5131.7    10263
m4ci 27 10316 10473 -5130.9    10262 1.6305      2     0.4425
```

P-value=0.44, hence, musician:Instrument is insignificant.

```
> anova(m4cv,m4c)
```

the exploration of interactions with musician was incomplete

```
Data:
Models:
m4c: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m4c:     OMSI + KnowRob + KnowAxis + musician + (1 | Subject:Instrument) +
m4c:     (1 | Subject:Harmony)
m4cv: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m4cv:     OMSI + KnowRob + KnowAxis + musician:Voice + (1 | Subject:Instrument) +
m4cv:     (1 | Subject:Harmony)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4c  25 10313 10459 -5131.7    10263
m4cv 27 10316 10474 -5131.2    10262 0.9401      2      0.625
```

P-value=0.625, hence, musician:Voice is insignificant.

```
> anova(m4crob,m4c)
```

```
Data:
Models:
m4c: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m4c:     OMSI + KnowRob + KnowAxis + musician + (1 | Subject:Instrument) +
m4c:     (1 | Subject:Harmony)
m4crob: Popular ~ Instrument + Harmony + Voice + factor(X1990s2000s.minus.1960s1970s) +
m4crob:     OMSI + KnowRob + KnowAxis + musician:KnowRob + (1 | Subject:Instrument) +
m4crob:     (1 | Subject:Harmony)
       Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4c    25 10313 10459 -5131.7    10263
m4crob 27 10316 10473 -5130.8    10262 1.7656      2     0.4136
```

P-value=0.41, hence, musician:KnowRob is insignificant.

```
> anova(m4co,m4c)
```

```
Data:
Models:
m4c: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
```

```
m4c:     OMSI + KnowRob + KnowAxis + musician + (1 | Subject:Instrument) +
m4c:     (1 | Subject:Harmony)
m4co: Popular ~ Instrument + Harmony + Voice + X1990s2000s.minus.1960s1970s +
m4co:     OMSI + KnowRob + KnowAxis + musician + musician:OMSI + (1 |
m4co:     Subject:Instrument) + (1 | Subject:Harmony)
     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
m4c  25 10313 10459 -5131.7    10263
m4co 26 10314 10466 -5131.2    10262 0.8783      1     0.3487
```

P-value=0.35, hence, musician:PachListen is insignificant.

Thus, we include the random effects of $musician : Harmony$.

Report

Introduction

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". 36 musical stimuli were presented 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.

Method

Based on the dataset given, we fit three types of models, namely, basic linear models, hierarchical models with single random effects and hierarchical models with combined random effects, to analyze the two response variables: Classical and Popular. We then choose the best model based under AIC, BIC, DIC and LRT criteria.

Results

The best model of Classical score is the hierarchical models with combined random effects. Otter than the three main effect Instrument, Harmony and Voice, the auxiliary measure of listeners' ability distinguish classical vs. popular music, how much the listeners concentrate on the notes while listening, how familiar of them with Pachelbel's Canon in D and the difference between prev variable and a similar variable referring to 60's and 70's pop and rock also affect the listeners' decisions. Among these fixed effects, the most significant one is Instrument. In addition, we also note that a combined random effect eg. $(1|Subject:Instrument)$ is also important because there might exist personal biases in rating.

Similar situations are seen in the analysis of Popular score. The factors that influence the score are score on a test of musical knowledge, a self-declare on whether the subject is a musician or not, whether the subject has heard of Rob Paravonian's Pachelbel Rant, whether the subject has heard of Axis of Evil's Comedy bit on the 4 Pachelbel chords in popular music and again, the difference between prev variable and a similar variable referring to 60's and 70's pop and rock. A new indicator defining whether a subject is a musician or not is highly significant in the best model.

?

Discussion

One thing we need to pay attention to is the missing values in the dataset. Although it is a relatively large dataset, some variables have most of their values to be NA. We treated the NA values as 0 in this study. Yet we still need to check the generalization ability of this dataset.

4: 18
5: 18 well organized but could add a bit more information