Final_vpisztor

Vincent Pisztora

December 18, 2015

1 a 9/9
b 5/9
c 6/9

2 a 5/9
b 9/9
c 9/9

3 9/9

4 a 9/9
b 6/9
c 7/9

5 10/10

Total 84/100

NUMBER 1

```
library(car)
library(stats)
library(lme4)
```

```
## Loading required package: Matrix
```

```
df = read.csv("ratings.csv")
df = df[,!(names(df) %in% c("first12", "X"))]
```

PART A

Next we transform all categorical variables into factors. The remaining ordinal equi-iterval variables will be treated as numeric.

```
df$Subject = as.factor(df$Subject)
df$CollegeMusic = as.factor(df$CollegeMusic)
df$APTheory = as.factor(df$APTheory)
```

To determine the which of the three main experimental factors are most influential on Classical ratings, we will regress all observed predictors (including Popular rating) on Classical. We will then examine the magnitude and significance of each of the main factors on rating.

First we assess the fit of our model. We find that the constant error variance and normality of errors assumptions hold well using the qq and residual plots. When reviewing the summary regression results we notice that many of the coefficients are not calculated.

We suspect missing data and find that the majority of the X2ndInstr and X1stInstr field are coded as NA. Interpreting this in context we will take the remedial measure of assuming NA to mean 0 or no exerience at all. Making this coding change we are now left with 979 incomplete observations. These we will remove from the sample instead of making further assumptions. Through this process we also eliminate all observations with non-zero X2ndInstr, thus we will drop this field entirely as well. Important to note that this may add a very specific selection bias to our results as it could be interpreted as removing most very musical subjects from the analysis.

We notice that if Subject is being controlled for, many fields contain only one observation per subject, making it impossible to calculate a coefficient. Thus for now we will remove Subject from the regression to assess Instrument, Harmony, and Voice overall.

Finally, we will also remove Instr.minus.Notes from all future analysis as it provides no additional information as a function of ConInstr and ConsNotes.

Our resulting model shows a significant effect of Instrument, Harmony, and Voice on Classical rating. There is nuance in the details though (all interpretations below should be read as "... when controlling for all other predictors":
Instrument:
- guitar: baseline

- piano: Samples featuring this instrument have significantly higher Classical ratings than the baseline.
- string: Samples featuring this instrument have significantly higher Classical ratings than the baseline.

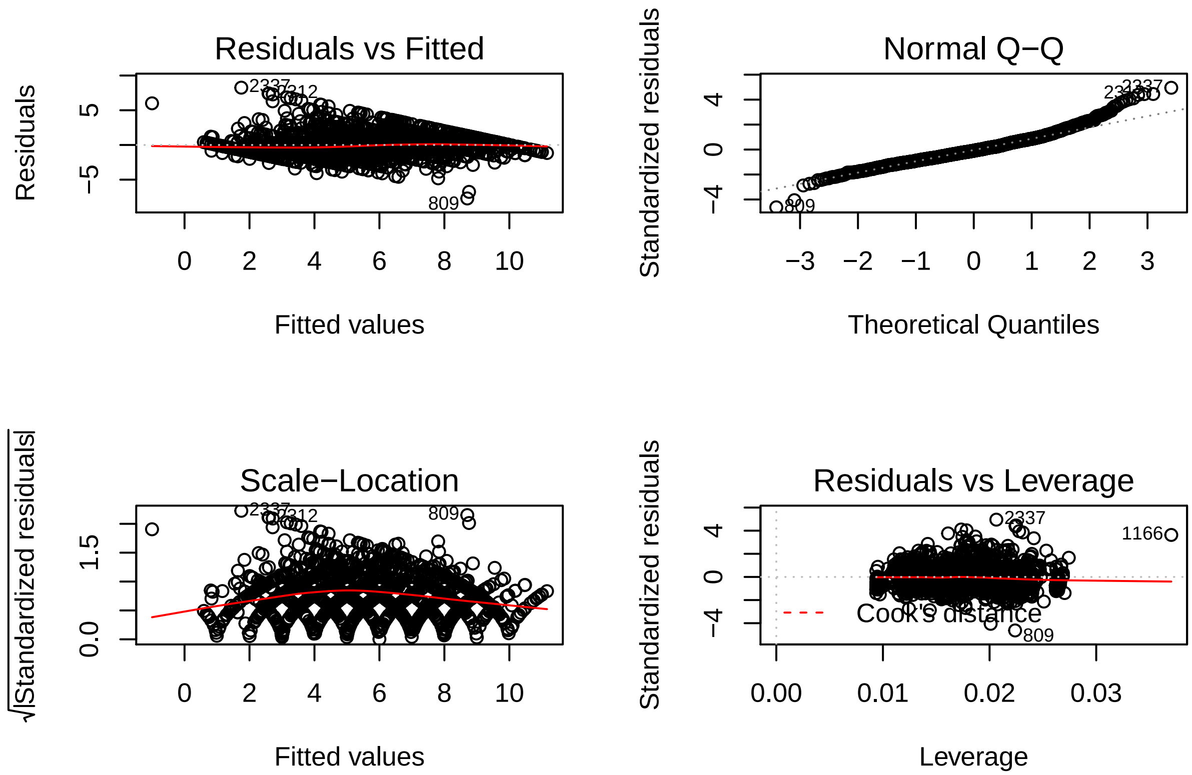
Voice:
- contrary: baseline
- par3rd: Samples featuring this type of voice have significantly lower Classical ratings than the baseline.
- par5th: Samples featuring this type of voice do not have significantly different Classical ratings than the baseline.

Harmony:
- I-IV-V: baseline
- I-V-IV: Samples of this pattern do not have significantly different Classical ratings than the baseline.
- I-V-VI: Samples of this pattern have on average significantly higher Classical ratings than the baseline.
- IV-I-V: Samples of this pattern do not have significantly different Classical ratings than the baseline.

We note that overall in terms of magnitudes it appears that the Instrument is most influential in Classical rating with strings specifically averaging 1.9 units higher in Classical rating than guitar samples. It is
9 important to understand the interpretation here as being "as compared to the baseline level". In otherwords we are saying that our data contains the largest rating differences between the various guitar and other instrument samples, as compared to differences between baseline and other across our other variables.

```r
#str(na.omit(df))
#summary(df$X2ndInstr)
df$X2ndInstr = ifelse(is.na(df$X2ndInstr), 0, df$X2ndInstr)
df$X1stInstr = ifelse(is.na(df$X1stInstr), 0, df$X1stInstr)
df = na.omit(df)
df = df[,!(names(df) %in% c("Instr.minus.Notes", "X2ndInstr"))]

f.1 = lm(data = df, formula = Classical ~ . - Subject)
par(mfrow = c(2,2))
plot(f.1)
```




```r
summary(f.1)
```

```
##
## Call:
## lm(formula = Classical ~ . - Subject, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.7054 -1.0828 -0.1568  0.9569  8.2523
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 6.2831091  0.4572948  13.740  < 2e-16 ***
## HarmonyI-V-IV               0.0152324  0.1213129   0.126  0.90009
## HarmonyI-V-VI               0.7095148  0.1214926   5.840 6.38e-09 ***
## HarmonyIV-I-V              -0.0812238  0.1213311  -0.669  0.50332
## Instrumentpiano             1.0214167  0.1075855   9.494  < 2e-16 ***
## Instrumentstring            1.9088032  0.1206453  15.822  < 2e-16 ***
## Voicepar3rd                -0.3011620  0.1051813  -2.863  0.00425 **
## Voicepar5th                -0.1710722  0.1052102  -1.626  0.10416
## Selfdeclare                -0.5552572  0.0934754  -5.940 3.53e-09 ***
## OMSI                        0.0012367  0.0004137   2.989  0.00284 **
## X16.minus.17               -0.0058035  0.0209305  -0.277  0.78161
## ConsInstr                   0.0398323  0.0423270   0.941  0.34682
## ConsNotes                  -0.0717101  0.0383734  -1.869  0.06185 .
```

```
## PachListen                     0.0181800  0.0625867   0.290  0.77149    
## ClsListen                      0.2440720  0.0429666   5.681 1.61e-08 ***
## KnowRob                        0.1383228  0.0332306   4.163 3.32e-05 ***
## KnowAxis                       0.0756495  0.0256049   2.954  0.00318 ** 
## X1990s2000s                    0.2551362  0.0499012   5.113 3.58e-07 ***
## X1990s2000s.minus.1960s1970s   0.0232614  0.0412564   0.564  0.57296    
## CollegeMusic1                  0.1229991  0.1484442   0.829  0.40747    
## NoClass                       -0.0595611  0.0496848  -1.199  0.23080    
## APTheory1                      0.7183365  0.1502955   4.779 1.93e-06 ***
## Composing                      0.3102425  0.0485365   6.392 2.18e-10 ***
## PianoPlay                      0.2724478  0.0341066   7.988 2.69e-15 ***
## GuitarPlay                    -0.2150196  0.0536166  -4.010 6.36e-05 ***
## X1stInstr                      0.0369862  0.0294350   1.257  0.20911    
## Popular                       -0.5543990  0.0197134 -28.123  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1.684 on 1514 degrees of freedom
## Multiple R-squared:  0.6263, Adjusted R-squared:  0.6199 
## F-statistic:  97.6 on 26 and 1514 DF,  p-value: < 2.2e-16
```

PART B

SUBPART I

see other submitted file

SUBPART II

To determine whether a random intercept is helpful, we fit a random-intercept model and then compare its AIC and BIC to the AIC and BIC of a fixed effects only model. In this case we observe that random intercept is not a better model as both BIC and AIC values are higher.

```r
fit.2 = lmer(data = df, Classical ~ . + (1 | Subject) -
    Selfdeclare -
    OMSI -
    X16.minus.17 -
    ConsInstr -
    ConsNotes -
    PachListen -
    ClsListen -
    KnowRob -
    KnowAxis -
    X1990s2000s -
    X1990s2000s.minus.1960s1970s -
    CollegeMusic -
    NoClass -
    APTheory -
    Composing -
    PianoPlay -
    GuitarPlay -
    X1stInstr)
summary(fit.2)
```

strange way to specify a model --
what's left?

```
## Linear mixed model fit by REML ['lmerMod']
```

```
## Formula:
## Classical ~ . + (1 | Subject) - Selfdeclare - OMSI - X16.minus.17 -
##     ConsInstr - ConsNotes - PachListen - ClsListen - KnowRob -
##     KnowAxis - X1990s2000s - X1990s2000s.minus.1960s1970s - CollegeMusic -
##     NoClass - APTheory - Composing - PianoPlay - GuitarPlay -
##     X1stInstr
##    Data: df
##
## REML criterion at convergence: 5604.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9265 -0.5424 -0.0028  0.5138  4.6123
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.5679   0.7536
##  Residual             2.1986   1.4828
## Number of obs: 1541, groups:  Subject, 43
##
## Fixed effects:
##                    Estimate Std. Error t value
## (Intercept)         8.19167    0.81098   10.10
## Subject17          -0.49193    1.12180   -0.44
## Subject19          -1.00647    1.12169   -0.90
## Subject20          -0.72710    1.12163   -0.65
## Subject22          -1.88870    1.12226   -1.68
## Subject23          -0.07029    1.12163   -0.06
## Subject26          -0.48527    1.12163   -0.43
## Subject29           0.60290    1.12165    0.54
## Subject30           0.76981    1.12229    0.69
## Subject31          -0.43335    1.12711   -0.38
## Subject32           0.42150    1.12171    0.38
## Subject37          -1.50014    1.12198   -1.34
## Subject38          -1.15034    1.12162   -1.03
## Subject40          -0.95754    1.12163   -0.85
## Subject42           1.03097    1.12173    0.92
## Subject44.1         0.38729    1.12165    0.35
## Subject44.2        -0.41531    1.12297   -0.37
## Subject45           0.34807    1.12163    0.31
## Subject46          -0.77126    1.12162   -0.69
## Subject47           3.33469    1.12297    2.97
## Subject48           0.73449    1.12349    0.65
## Subject49           0.11261    1.12194    0.10
## Subject52          -0.39048    1.12165   -0.35
## Subject53          -0.99971    1.12305   -0.89
## Subject55           0.51309    1.12163    0.46
## Subject56           2.56850    1.12189    2.29
## Subject57          -1.69444    1.12162   -1.51
## Subject59           2.45908    1.12170    2.19
## Subject60           1.94599    1.12175    1.73
## Subject61           0.78923    1.12162    0.70
## Subject63           0.31710    1.12182    0.28
## Subject64           1.69787    1.12207    1.51
```

subject should not be a fixed effect in this model

```
## Subject66         0.85947    1.12162    0.77
## Subject71        -1.22208    1.12198   -1.09
## Subject74         0.57039    1.12189    0.51
## Subject78         1.67633    1.12196    1.49
## Subject80         0.47717    1.12166    0.43
## Subject81        -1.42150    1.12171   -1.27
## Subject82        -1.89386    1.12171   -1.69
## Subject83         0.78899    1.12252    0.70
## Subject93         1.16343    1.12164    1.04
## Subject94        -1.71222    1.12227   -1.53
## Subject98         0.33647    1.12190    0.30
## HarmonyI-V-IV     0.01455    0.10681    0.14
## HarmonyI-V-VI     0.68258    0.10700    6.38
## HarmonyIV-I-V    -0.10425    0.10685   -0.98
## Instrumentpiano   0.90315    0.09547    9.46
## Instrumentstring  1.62954    0.11035   14.77
## Voicepar3rd      -0.28245    0.09263   -3.05
## Voicepar5th      -0.14902    0.09266   -1.61
## Popular          -0.64696    0.02000  -32.34
```

```
##
## Correlation matrix not shown by default, as p = 51 > 20.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

```
fit.3 = lm(data = df, Classical ~ . -
    Selfdeclare -
    OMSI -
    X16.minus.17 -
    ConsInstr -
    ConsNotes -
    PachListen -
    ClsListen -
    KnowRob -
    KnowAxis -
    X1990s2000s -
    X1990s2000s.minus.1960s1970s -
    CollegeMusic -
    NoClass -
    APTheory -
    Composing -
    PianoPlay -
    GuitarPlay -
    X1stInstr)
summary(fit.3)
```

```
##
## Call:
## lm(formula = Classical ~ . - Selfdeclare - OMSI - X16.minus.17 -
##     ConsInstr - ConsNotes - PachListen - ClsListen - KnowRob -
##     KnowAxis - X1990s2000s - X1990s2000s.minus.1960s1970s - CollegeMusic -
##     NoClass - APTheory - Composing - PianoPlay - GuitarPlay -
##     X1stInstr, data = df)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.3050 -0.8043 -0.0041  0.7619  6.8390
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       8.19167    0.29958  27.344  < 2e-16 ***
## Subject17        -0.49193    0.35007  -1.405 0.160155
## Subject19        -1.00647    0.34971  -2.878 0.004059 **
## Subject20        -0.72710    0.34952  -2.080 0.037672 *
## Subject22        -1.88870    0.35153  -5.373 8.99e-08 ***
## Subject23        -0.07029    0.34954  -0.201 0.840654
## Subject26        -0.48527    0.34954  -1.388 0.165253
## Subject29         0.60290    0.34959   1.725 0.084813 .
## Subject30         0.76981    0.35165   2.189 0.028744 *
## Subject31        -0.43335    0.36672  -1.182 0.237513
## Subject32         0.42150    0.34977   1.205 0.228368
## Subject37        -1.50014    0.35064  -4.278 2.00e-05 ***
## Subject38        -1.15034    0.34950  -3.291 0.001020 **
## Subject40        -0.95754    0.34952  -2.740 0.006225 **
## Subject42         1.03097    0.34984   2.947 0.003259 **
## Subject44.1       0.38729    0.34958   1.108 0.268092
## Subject44.2      -0.41531    0.35380  -1.174 0.240635
## Subject45         0.34807    0.34954   0.996 0.319514
## Subject46        -0.77126    0.34951  -2.207 0.027487 *
## Subject47         3.33469    0.35380   9.425  < 2e-16 ***
## Subject48         0.73449    0.35544   2.066 0.038962 *
## Subject49         0.11261    0.35051   0.321 0.748049
## Subject52        -0.39048    0.34958  -1.117 0.264173
## Subject53        -0.99971    0.35406  -2.824 0.004813 **
## Subject55         0.51309    0.34952   1.468 0.142314
## Subject56         2.56850    0.35035   7.331 3.72e-13 ***
## Subject57        -1.69444    0.34949  -4.848 1.38e-06 ***
## Subject59         2.45908    0.34975   7.031 3.11e-12 ***
## Subject60         1.94599    0.34992   5.561 3.17e-08 ***
## Subject61         0.78923    0.34950   2.258 0.024081 *
## Subject63         0.31710    0.35013   0.906 0.365262
## Subject64         1.69787    0.35093   4.838 1.45e-06 ***
## Subject66         0.85947    0.34950   2.459 0.014040 *
## Subject71        -1.22208    0.35064  -3.485 0.000506 ***
## Subject74         0.57039    0.35035   1.628 0.103726
## Subject78         1.67633    0.35060   4.781 1.91e-06 ***
## Subject80         0.47717    0.34963   1.365 0.172523
## Subject81        -1.42150    0.34977  -4.064 5.07e-05 ***
## Subject82        -1.89386    0.34979  -5.414 7.16e-08 ***
## Subject83         0.78899    0.35238   2.239 0.025302 *
## Subject93         1.16343    0.34955   3.328 0.000895 ***
## Subject94        -1.71222    0.35159  -4.870 1.24e-06 ***
## Subject98         0.33647    0.35039   0.960 0.337065
## HarmonyI-V-IV     0.01455    0.10681   0.136 0.891648
## HarmonyI-V-VI     0.68258    0.10700   6.379 2.37e-10 ***
## HarmonyIV-I-V    -0.10425    0.10685  -0.976 0.329374
## Instrumentpiano   0.90315    0.09547   9.460  < 2e-16 ***
```

```
## Instrumentstring  1.62954    0.11035  14.767  < 2e-16 ***
## Voicepar3rd      -0.28245    0.09263  -3.049 0.002334 **
## Voicepar5th      -0.14902    0.09266  -1.608 0.107990
## Popular          -0.64696    0.02000 -32.341  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.483 on 1490 degrees of freedom
## Multiple R-squared:  0.7149, Adjusted R-squared:  0.7054
## F-statistic: 74.74 on 50 and 1490 DF,  p-value: < 2.2e-16
```

```r
AIC(fit.2)
```

```
## [1] 5710.052
```

```r
BIC(fit.2)
```

```
## [1] 5993.082
```

```r
AIC(fit.3)
```

```
## [1] 5639.366
```

```r
BIC(fit.3)
```

```
## [1] 5917.055
```

SUBPART III

Our new model shows a significant effect of Instrument, Harmony, and Voice on Classical rating. Directionally they are identical to the previous (non-random intercept) model. There is nuance in the details though (all interpretations below should be read as "… when controlling for all other predictors":

Instrument:
- guitar: baseline
- piano: Samples featuring this instrument have significantly higher Classical ratings than the baseline.
- string: Samples featuring this instrument have significantly higher Classical ratings than the baseline.

it wold be good to commetn on the size of the effects here also

Voice:
- contrary: baseline
- par3rd: Samples featuring this type of voice have significantly lower Classical ratings than the baseline.
- par5th: Samples featuring this type of voice do not have significantly different Classical ratings than the baseline.

5

Harmony:
- I-IV-V: baseline
- I-V-IV: Samples of this pattern do not have significantly different Classical ratings than the baseline.
- I-V-VI: Samples of this pattern have on average significantly higher Classical ratings than the baseline.
- IV-I-V: Samples of this pattern do not have significantly different Classical ratings than the baseline.

PART C

SUBPART I

Fitting the new model with three separate interaction random effects on the intercept, we observe that this model has a lower AIC than the ones previously calculated. However, the BIC value falls between the previous two models. We will however interpret the higher lower AIC as indicative of the models superior performance. Interestingly it assigns a zero variance (i.e. no randomness) to the Voice:Subject interaction.

```r
fit.4 = lmer(data = df, Classical ~ . + (1 | Subject:Instrument) +
                                        (1 | Subject:Harmony) +
                                        (1 | Subject:Voice) -
    Selfdeclare -
    OMSI -
    X16.minus.17 -
    ConsInstr -
    ConsNotes -
    PachListen -
    ClsListen -
    KnowRob -
    KnowAxis -
    X1990s2000s -
    X1990s2000s.minus.1960s1970s -
    CollegeMusic -
    NoClass -
    APTheory -
    Composing -
    PianoPlay -
    GuitarPlay -
    X1stInstr)
summary(fit.4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Classical ~ . + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
##     (1 | Subject:Voice) - Selfdeclare - OMSI - X16.minus.17 -
##     ConsInstr - ConsNotes - PachListen - ClsListen - KnowRob -
##     KnowAxis - X1990s2000s - X1990s2000s.minus.1960s1970s - CollegeMusic -
##     NoClass - APTheory - Composing - PianoPlay - GuitarPlay -
##     X1stInstr
##    Data: df
##
## REML criterion at convergence: 5413
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4122 -0.5169 -0.0092  0.4820  5.7194
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2377   0.4876
##  Subject:Voice      (Intercept) 0.0000   0.0000
##  Subject:Instrument (Intercept) 0.5752   0.7584
##  Residual                       1.6436   1.2820
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Voice, 129; Subject:Instrument, 129
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)        7.80392    0.57981  13.460
## Subject17         -0.54833    0.77074  -0.711
## Subject19         -0.97201    0.77057  -1.261
```

```
## Subject20        -0.71457    0.77049  -0.927
## Subject22        -1.78216    0.77141  -2.310
## Subject23        -0.08596    0.77050  -0.112
## Subject26        -0.46960    0.77050  -0.609
## Subject29         0.57940    0.77052   0.752
## Subject30         0.87947    0.77146   1.140
## Subject31        -0.46214    0.77939  -0.593
## Subject32         0.38233    0.77060   0.496
## Subject37        -1.58004    0.77100  -2.049
## Subject38        -1.15660    0.77048  -1.501
## Subject40        -0.96850    0.77049  -1.257
## Subject42         0.98710    0.77063   1.281
## Subject44.1       0.40923    0.77052   0.531
## Subject44.2      -0.57041    0.77245  -0.738
## Subject45         0.36373    0.77050   0.472
## Subject46        -0.77909    0.77048  -1.011
## Subject47         3.17959    0.77245   4.116
## Subject48         0.83996    0.77292   1.087
## Subject49         0.03741    0.77094   0.049
## Subject52        -0.36855    0.77052  -0.478
## Subject53        -0.83991    0.77257  -1.087
## Subject55         0.52406    0.77049   0.680
## Subject56         2.49957    0.77087   3.243
## Subject57        -1.69444    0.77048  -2.199
## Subject59         2.42148    0.77059   3.142
## Subject60         1.89742    0.77067   2.462
## Subject61         0.79549    0.77048   1.032
## Subject63         0.37663    0.77077   0.489
## Subject64         1.78717    0.77113   2.318
## Subject66         0.85477    0.77048   1.109
## Subject71        -1.14218    0.77100  -1.481
## Subject74         0.63932    0.77087   0.829
## Subject78         1.59800    0.77098   2.073
## Subject80         0.50459    0.77054   0.655
## Subject81        -1.38233    0.77060  -1.794
## Subject82        -1.93460    0.77061  -2.510
## Subject83         0.66209    0.77180   0.858
## Subject93         1.18066    0.77050   1.532
## Subject94        -1.60412    0.77144  -2.079
## Subject98         0.26597    0.77088   0.345
## HarmonyI-V-IV     0.01086    0.13996   0.078
## HarmonyI-V-VI     0.69698    0.14010   4.975
## HarmonyIV-I-V    -0.08876    0.14000  -0.634
## Instrumentpiano   0.96811    0.18365   5.271
## Instrumentstring  1.80215    0.19186   9.393
## Voicepar3rd      -0.28983    0.08011  -3.618
## Voicepar5th      -0.16154    0.08016  -2.015
## Popular          -0.59056    0.02006 -29.442

##
## Correlation matrix not shown by default, as p = 51 > 20.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

again, subject should not be a fixed effect in these models!

```r
AIC(fit.4)
```

```
## [1] 5522.959
```

```r
BIC(fit.4)
```

```
## [1] 5816.669
```

SUBPART II

Our new model shows a significant effect of Instrument, Harmony, and Voice on Classical rating. Directionally they are almost to the previous models - the only difference being the significance of an additional Voice level. Detail below (all interpretations below should be read as "... when controlling for all other predictors":
Instrument:
- guitar: baseline
- piano: Samples featuring this instrument have significantly higher Classical ratings than the baseline.
- string: Samples featuring this instrument have significantly higher Classical ratings than the baseline.

Voice:
6
- contrary: baseline
- par3rd: Samples featuring this type of voice have significantly lower Classical ratings than the baseline.
- par5th: Samples featuring this type of voice have significantly lower Classical ratings than the baseline.

Harmony:
- I-IV-V: baseline
- I-V-IV: Samples of this pattern do not have significantly different Classical ratings than the baseline.
- I-V-VI: Samples of this pattern have on average significantly higher Classical ratings than the baseline.
- IV-I-V: Samples of this pattern do not have significantly different Classical ratings than the baseline.

Observing the random effect variance estimates, we also note that the "individual bias" is largest for subject with respect to the effect of instruments, followed by harmony. Interestingly as noted above, the model assigns a zero variance (i.e. no randomness) to the Voice:Subject interaction. All random effects are however much smaller (at least by 3x) than the residual variance - making them very small indeed.

SUBPART III

see other submitted file

NUMBER 2

PART A

We begin with model 4, as it was determined to be best above. Next we fit a full model using all individual variables. Next we observe which are significant and create a new reduced model containing only these variables.

```r
fit.5 = lmer(data = df, Classical ~ . + (1 | Subject:Instrument) +
                                        (1 | Subject:Harmony) +
                                        (1 | Subject:Voice) -
    Subject)
summary(fit.5)
```

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

```
##
## REML criterion at convergence: 5593.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.2606 -0.5184 -0.0080  0.4941  5.8438
##
## Random effects:
##  Groups             Name        Variance  Std.Dev.
##  Subject:Harmony    (Intercept) 2.737e-01 5.232e-01
##  Subject:Voice      (Intercept) 1.827e-14 1.352e-07
##  Subject:Instrument (Intercept) 1.023e+00 1.011e+00
##  Residual                       1.637e+00 1.280e+00
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Voice, 129; Subject:Instrument, 129
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                 6.4155773  1.0350673   6.198
## HarmonyI-V-IV               0.0106192  0.1457095   0.073
## HarmonyI-V-VI               0.7028564  0.1458422   4.819
## HarmonyIV-I-V              -0.0833837  0.1457424  -0.572
## Instrumentpiano             0.9942412  0.2334647   4.259
## Instrumentstring            1.8656881  0.2398514   7.779
## Voicepar3rd                -0.2935790  0.0799580  -3.672
## Voicepar5th                -0.1664003  0.0800031  -2.080
## Selfdeclare                -0.5643751  0.2243191  -2.516
## OMSI                        0.0012528  0.0009922   1.263
## X16.minus.17               -0.0053622  0.0496123  -0.108
## ConsInstr                   0.0406907  0.1016148   0.400
## ConsNotes                  -0.0663893  0.0917719  -0.723
## PachListen                  0.0144299  0.1495052   0.097
## ClsListen                   0.2453888  0.1031536   2.379
## KnowRob                     0.1337506  0.0790849   1.691
## KnowAxis                    0.0802155  0.0612369   1.310
## X1990s2000s                 0.2510584  0.1191225   2.108
## X1990s2000s.minus.1960s1970s  0.0277147  0.0986131   0.281
## CollegeMusic1               0.1055427  0.3553281   0.297
## NoClass                    -0.0507189  0.1186690  -0.427
## APTheory1                   0.6958708  0.3576051   1.946
## Composing                   0.3131754  0.1162797   2.693
## PianoPlay                   0.2735267  0.0815202   3.355
## GuitarPlay                 -0.2184469  0.1284032  -1.701
## X1stInstr                   0.0384870  0.0700448   0.549
## Popular                    -0.5696086  0.0198684 -28.669

##
## Correlation matrix not shown by default, as p = 27 > 20.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

```
fit.6 = lmer(data = df, Classical ~ . + (1 | Subject:Instrument) +
                                        (1 | Subject:Harmony) +
                                        (1 | Subject:Voice) -
```

```r
             OMSI -
             X16.minus.17 -
             ConsInstr -
             ConsNotes -
             PachListen -
             KnowRob -
             KnowAxis -
             X1990s2000s.minus.1960s1970s -
             CollegeMusic -
             NoClass -
             APTheory -
             GuitarPlay -
             X1stInstr -
             Subject)
summary(fit.6)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Classical ~ . + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
##     (1 | Subject:Voice) - OMSI - X16.minus.17 - ConsInstr - ConsNotes -
##     PachListen - KnowRob - KnowAxis - X1990s2000s.minus.1960s1970s -
##     CollegeMusic - NoClass - APTheory - GuitarPlay - X1stInstr -
##     Subject
##    Data: df
##
## REML criterion at convergence: 5569.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1965 -0.5262 -0.0113  0.4870  5.8636
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2753   0.5247
##  Subject:Voice      (Intercept) 0.0000   0.0000
##  Subject:Instrument (Intercept) 1.1106   1.0538
##  Residual                       1.6370   1.2795
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Voice, 129; Subject:Instrument, 129
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)        6.79410    0.49640  13.687
## HarmonyI-V-IV      0.01038    0.14596   0.071
## HarmonyI-V-VI      0.70296    0.14609   4.812
## HarmonyIV-I-V     -0.08320    0.14599  -0.570
## Instrumentpiano    0.99413    0.24206   4.107
## Instrumentstring   1.86780    0.24818   7.526
## Voicepar3rd       -0.29379    0.07995  -3.675
## Voicepar5th       -0.16665    0.08000  -2.083
## Selfdeclare       -0.43823    0.12852  -3.410
## ClsListen          0.22831    0.07460   3.061
## X1990s2000s        0.24508    0.07352   3.334
```

```
## Composing         0.29907    0.09242   3.236
## PianoPlay         0.30209    0.07206   4.192
## Popular          -0.56888    0.01980 -28.726
##
## Correlation of Fixed Effects:
##             (Intr) HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r Vcpr5t
## HrmnyI-V-IV -0.146
## HrmnyI-V-VI -0.156  0.499
## HrmnyIV-I-V -0.156  0.499  0.500
## Instrumntpn -0.267  0.000  0.004  0.003
## Instrmntstr -0.299 -0.001  0.008  0.008  0.506
## Voicepar3rd -0.068 -0.001 -0.002 -0.001 -0.004    -0.011
## Voicepar5th -0.065 -0.001 -0.004 -0.003 -0.005    -0.014     0.502
## Selfdeclare -0.475  0.000  0.000  0.000  0.000    -0.002     0.001  0.001
## ClsListen   -0.307  0.000  0.001  0.001  0.002     0.004    -0.001 -0.001
## X1990s2000s -0.651  0.000  0.000  0.000  0.000    -0.002     0.000  0.000
## Composing    0.181  0.000 -0.001 -0.001 -0.003    -0.007     0.001  0.001
## PianoPlay    0.163  0.000  0.000  0.000  0.000     0.002     0.000 -0.001
## Popular     -0.262 -0.004  0.035  0.034  0.094     0.241    -0.048 -0.058
##             Slfdcl ClsLst X19902 Cmpsng PinPly
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## Voicepar3rd
## Voicepar5th
## Selfdeclare
## ClsListen   -0.066
## X1990s2000s -0.037  0.178
## Composing   -0.555 -0.206  0.079
## PianoPlay   -0.434 -0.211  0.016  0.166
## Popular     -0.009  0.015 -0.008 -0.028  0.010
```

5

very difficult to tell what your result in part (a) is!

PART B

Once we fit the reduced model we observe as previously that the Voice:Subject interaction is estimated as zero. Thus it can be removed from the model. The updated model can be seen below.

```
fit.7 = lmer(data = df, Classical ~ . + (1 | Subject:Instrument) +
                                        (1 | Subject:Harmony) -
              OMSI -
              X16.minus.17 -
              ConsInstr -
              ConsNotes -
              PachListen -
              KnowRob -
              KnowAxis -
              X1990s2000s.minus.1960s1970s -
              CollegeMusic -
              NoClass -
              APTheory -
              GuitarPlay -
              X1stInstr -
```

```r
             Subject)
summary(fit.7)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Classical ~ . + (1 | Subject:Instrument) + (1 | Subject:Harmony) -
##     OMSI - X16.minus.17 - ConsInstr - ConsNotes - PachListen -
##     KnowRob - KnowAxis - X1990s2000s.minus.1960s1970s - CollegeMusic -
##     NoClass - APTheory - GuitarPlay - X1stInstr - Subject
##    Data: df
##
## REML criterion at convergence: 5569.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1965 -0.5262 -0.0113  0.4870  5.8636
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2753   0.5247
##  Subject:Instrument (Intercept) 1.1106   1.0538
##  Residual                       1.6370   1.2795
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Instrument, 129
##
## Fixed effects:
##                  Estimate Std. Error t value
## (Intercept)       6.79410    0.49640  13.687
## HarmonyI-V-IV     0.01038    0.14596   0.071
## HarmonyI-V-VI     0.70296    0.14609   4.812
## HarmonyIV-I-V    -0.08320    0.14599  -0.570
## Instrumentpiano   0.99413    0.24206   4.107
## Instrumentstring  1.86780    0.24818   7.526
## Voicepar3rd      -0.29379    0.07995  -3.675
## Voicepar5th      -0.16665    0.08000  -2.083
## Selfdeclare      -0.43823    0.12852  -3.410
## ClsListen         0.22831    0.07460   3.061
## X1990s2000s       0.24508    0.07352   3.334
## Composing         0.29907    0.09242   3.236
## PianoPlay         0.30209    0.07206   4.192
## Popular          -0.56888    0.01980 -28.726
##
## Correlation of Fixed Effects:
##             (Intr) HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r Vcpr5t
## HrmnyI-V-IV -0.146
## HrmnyI-V-VI -0.156  0.499
## HrmnyIV-I-V -0.156  0.499  0.500
## Instrumntpn -0.267  0.000  0.004  0.003
## Instrmntstr -0.299 -0.001  0.008  0.008  0.506
## Voicepar3rd -0.068 -0.001 -0.002 -0.001 -0.004    -0.011
## Voicepar5th -0.065 -0.001 -0.004 -0.003 -0.005    -0.014     0.502
## Selfdeclare -0.475  0.000  0.000  0.000  0.000    -0.002     0.001  0.001
## ClsListen   -0.307  0.000  0.001  0.001  0.002     0.004    -0.001 -0.001
```

```
## X1990s2000s -0.651  0.000  0.000  0.000  0.000    -0.002     0.000  0.000
## Composing    0.181  0.000 -0.001 -0.001 -0.003    -0.007     0.001  0.001
## PianoPlay    0.163  0.000  0.000  0.000  0.000     0.002     0.000 -0.001
## Popular     -0.262 -0.004  0.035  0.034  0.094     0.241    -0.048 -0.058
##             Slfdcl ClsLst X19902 Cmpsng PinPly
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## Voicepar3rd
## Voicepar5th
## Selfdeclare
## ClsListen   -0.066
## X1990s2000s -0.037  0.178
9 ## Composing   -0.555 -0.206  0.079
## PianoPlay   -0.434 -0.211  0.016  0.166
## Popular     -0.009  0.015 -0.008 -0.028  0.010
```

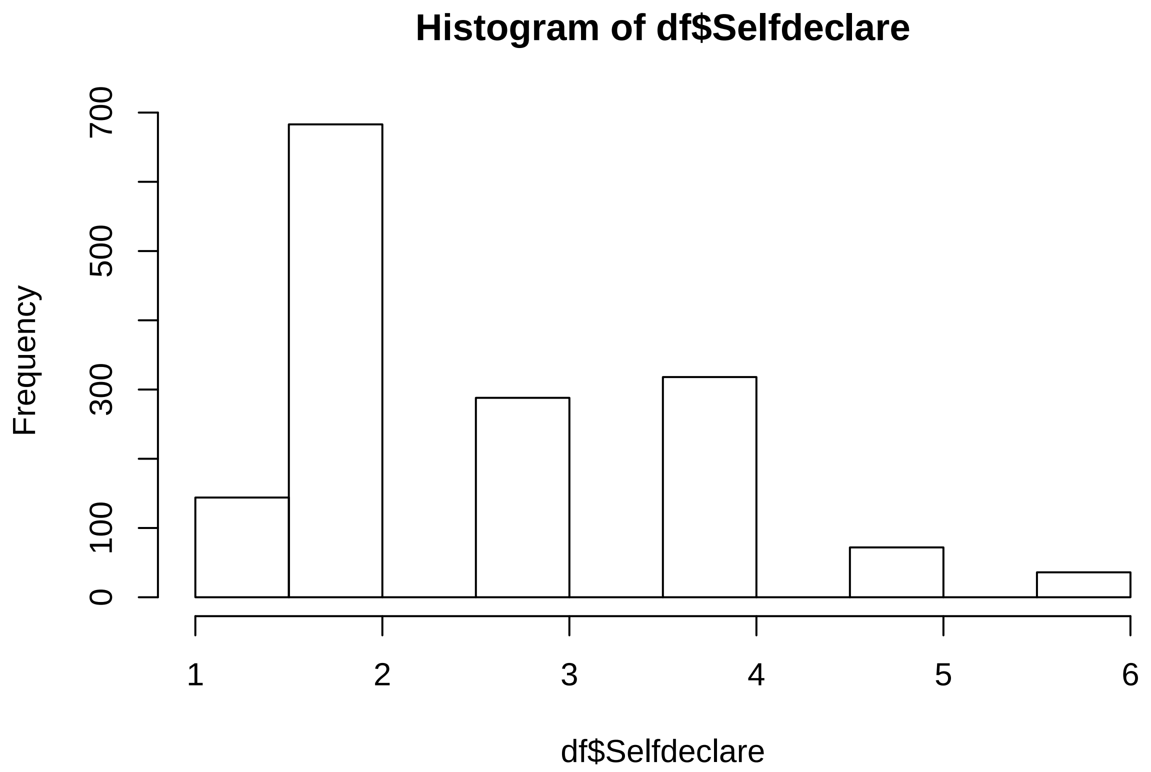
PART C

Variable interpretations (all interpretations are "... controlling for other predictors"):
- Harmony, Instrument, Voice: Interpretation the same as in part 1) c) II)
- Selfdeclare: With each additional "unit of certainty" of being a musician, an individual will on average rate a sample as being .43 units less Classical.
- ClsListen: With each additional unit of listing to classical music, an individual will on average rate a sample
9 as being .23 units more Classical.
- X1990s2000s: With each additional unit of listening to 90s and 2000s music, an individual will on average rate a sample as being .25 units more Classical.
- Composing: With each additional unit of music composition, an individual will on average rate a sample as being .3 units more Classical.
- PianoPlay: With each additional unit of piano play, an individual will on average rate a sample as being .3 units more Classical.
- Popular: The more an individual rates a particular sample as being Popular, the less they will rate the same sample as being Classical (at a rate of -.57 Classical units for each Popular unit). This is interesting as it calls into question the independence assumption made by the study creators.

NUMBER 3

First we determine the distribution of the Selfdeclare variable to determine the median to be 2. Next we code the Musician indicator.

```r
par(mfrow = c(1,1))
hist(df$Selfdeclare)
```




```
median(df$Selfdeclare)
```

```
## [1] 2
```

```
df2 = df
df2$Musician = ifelse(df2$Selfdeclare > 2, "Yes", "No")
```

Next modeling the interaction between being a Musician and other predictors, we find only a few interactions to be significant. Specifically (all interpretations are "… controlling for other predictors"): - Musician: On average musicians rate a sample as .16 units more Classical than non-musicians.
- Musician:HarmonyI-V-VI: On average musicians rate a HarmonyI-V-VI sample as .81 units more Classical
9 than non-musicians.
- Musician:PianoPlay: On average musicians report playing less piano than non-musicians (.43 units less).
- Musician:Popular: On average musicians rate a sample as .11 units more Popular than non-musicians.
Also note that we now remove the Selfdeclare variable as a predictor as it is related to the new Musician indicator.

```
fit.8 = lmer(data = df2, Classical ~ . + (1 | Subject:Instrument) +
                                         (1 | Subject:Harmony) -
              OMSI -
              X16.minus.17 -
              ConsInstr -
              ConsNotes -
              PachListen -
              KnowRob -
```

```r
              KnowAxis -
              X1990s2000s.minus.1960s1970s -
              CollegeMusic -
              NoClass -
              APTheory -
              GuitarPlay -
              X1stInstr -
              Selfdeclare -
              Subject +
              Musician*Harmony +
              Musician*Instrument +
              Musician*Voice +
              Musician*ClsListen +
              Musician*X1990s2000s +
              Musician*Composing +
              Musician*PianoPlay +
              Musician*Popular)
summary(fit.8)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Classical ~ . + (1 | Subject:Instrument) + (1 | Subject:Harmony) -
##     OMSI - X16.minus.17 - ConsInstr - ConsNotes - PachListen -
##     KnowRob - KnowAxis - X1990s2000s.minus.1960s1970s - CollegeMusic -
##     NoClass - APTheory - GuitarPlay - X1stInstr - Selfdeclare -
##     Subject + Musician * Harmony + Musician * Instrument + Musician *
##     Voice + Musician * ClsListen + Musician * X1990s2000s + Musician *
##     Composing + Musician * PianoPlay + Musician * Popular
##    Data: df2
##
## REML criterion at convergence: 5566.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3639 -0.5346 -0.0079  0.4770  5.7105
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2558   0.5058
##  Subject:Instrument (Intercept) 1.1707   1.0820
##  Residual                       1.6295   1.2765
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Instrument, 129
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                  5.703056   0.790726   7.212
## HarmonyI-V-IV                0.009662   0.194906   0.050
## HarmonyI-V-VI                0.349512   0.195055   1.792
## HarmonyIV-I-V               -0.129763   0.195005  -0.665
## Instrumentpiano              1.097144   0.339467   3.232
## Instrumentstring             1.955490   0.352016   5.555
## Voicepar3rd                 -0.266674   0.109090  -2.445
```

```
## Voicepar5th                  -0.137023   0.108914  -1.258
## ClsListen                     0.212927   0.103806   2.051
## X1990s2000s                   0.333715   0.149457   2.233
## Composing                     0.224956   0.215458   1.044
## PianoPlay                     0.540245   0.136681   3.953
## Popular                      -0.619131   0.028682 -21.586
## MusicianYes                   0.158627   0.987167   0.161
## HarmonyI-V-IV:MusicianYes    -0.001673   0.286103  -0.006
## HarmonyI-V-VI:MusicianYes     0.811291   0.286873   2.828
## HarmonyIV-I-V:MusicianYes     0.105080   0.286159   0.367
## Instrumentpiano:MusicianYes  -0.248112   0.496767  -0.499
## Instrumentstring:MusicianYes -0.246606   0.509001  -0.484
## Voicepar3rd:MusicianYes      -0.045985   0.160010  -0.287
## Voicepar5th:MusicianYes      -0.070106   0.160070  -0.438
## ClsListen:MusicianYes        -0.026980   0.157001  -0.172
## X1990s2000s:MusicianYes      -0.121244   0.173935  -0.697
## Composing:MusicianYes        -0.108017   0.237914  -0.454
## PianoPlay:MusicianYes        -0.431573   0.159876  -2.699
## Popular:MusicianYes           0.109955   0.039779   2.764

##
## Correlation matrix not shown by default, as p = 26 > 20.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

NUMBER 4

First we will fit the three base models (no random effect, random intercept, random interaction). Then selecting the best model we will add in all significant subject level factors. Finally we will assess the interaction effect of Musician on the final model.

We find that both the AIC and BIC are lower for the random interaction model. Fitting all subject level factors we find that the X16.minus.17, X1990s2000s, Composing, GuitarPlay, Classical are significantly other than zero.

```r
fit.p.1 = lm(data = df, Popular ~ . -
    Selfdeclare -
    OMSI -
    X16.minus.17 -
    ConsInstr -
    ConsNotes -
    PachListen -
    ClsListen -
    KnowRob -
    KnowAxis -
    X1990s2000s -
    X1990s2000s.minus.1960s1970s -
    CollegeMusic -
    NoClass -
    APTheory -
    Composing -
    PianoPlay -
    GuitarPlay -
    X1stInstr)
```

```r
fit.p.2 = lmer(data = df, Popular ~ . + (1 | Subject) -
    Selfdeclare -
    OMSI -
    X16.minus.17 -
    ConsInstr -
    ConsNotes -
    PachListen -
    ClsListen -
    KnowRob -
    KnowAxis -
    X1990s2000s -
    X1990s2000s.minus.1960s1970s -
    CollegeMusic -
    NoClass -
    APTheory -
    Composing -
    PianoPlay -
    GuitarPlay -
    X1stInstr)

fit.p.3 = lmer(data = df, Popular ~ . + (1 | Subject:Instrument) +
                                        (1 | Subject:Harmony) +
                                        (1 | Subject:Voice) -
    Selfdeclare -
    OMSI -
    X16.minus.17 -
    ConsInstr -
    ConsNotes -
    PachListen -
    ClsListen -
    KnowRob -
    KnowAxis -
    X1990s2000s -
    X1990s2000s.minus.1960s1970s -
    CollegeMusic -
    NoClass -
    APTheory -
    Composing -
    PianoPlay -
    GuitarPlay -
    X1stInstr)
AIC(fit.p.1)
```

```
## [1] 5616.724
```

```r
AIC(fit.p.2)
```

```
## [1] 5688.174
```

```r
AIC(fit.p.3)
```

```
## [1] 5593.375
```

```r
BIC(fit.p.1)
```

```
## [1] 5894.414
```

```r
BIC(fit.p.2)
```

```
## [1] 5971.204
```

```r
BIC(fit.p.3)
```

```
## [1] 5887.085
```

```r
fit.p.4 = lmer(data = df, Popular ~ . + (1 | Subject:Instrument) +
                                        (1 | Subject:Harmony) +
                                        (1 | Subject:Voice) -
    Subject)
summary(fit.p.4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Popular ~ . + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
##     (1 | Subject:Voice) - Subject
##    Data: df
##
## REML criterion at convergence: 5715.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.2071 -0.5166 -0.0011  0.4641  5.5242
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2705   0.5201
##  Subject:Voice      (Intercept) 0.0000   0.0000
##  Subject:Instrument (Intercept) 1.1708   1.0820
##  Residual                       1.7782   1.3335
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Voice, 129; Subject:Instrument, 129
##
## Fixed effects:
##                               Estimate Std. Error t value
## (Intercept)                   8.455395   1.084581   7.796
## HarmonyI-V-IV                 0.024194   0.147693   0.164
## HarmonyI-V-VI                 0.266878   0.148869   1.793
## HarmonyIV-I-V                -0.214779   0.147643  -1.455
## Instrumentpiano              -0.128420   0.250377  -0.513
## Instrumentstring             -0.807659   0.259481  -3.113
## Voicepar3rd                  -0.056826   0.083685  -0.679
## Voicepar5th                   0.049276   0.083484   0.590
## Selfdeclare                  -0.425046   0.237138  -1.792
## OMSI                          0.001525   0.001048   1.455
## X16.minus.17                  0.107875   0.052320   2.062
```

```
## ConsInstr                      0.057681   0.107313   0.537
## ConsNotes                      0.083887   0.096913   0.866
## PachListen                    -0.225492   0.157786  -1.429
## ClsListen                      0.137164   0.109084   1.257
## KnowRob                        0.159487   0.083496   1.910
## KnowAxis                       0.112811   0.064643   1.745
## X1990s2000s                    0.297738   0.125751   2.368
## X1990s2000s.minus.1960s1970s -0.064367   0.104135  -0.618
## CollegeMusic1                  0.200821   0.375241   0.535
## NoClass                       -0.055009   0.125326  -0.439
## APTheory1                      0.726781   0.377656   1.924
## Composing                      0.306399   0.122838   2.494
## PianoPlay                      0.077492   0.086381   0.897
## GuitarPlay                    -0.289667   0.135529  -2.137
## X1stInstr                     -0.090336   0.073943  -1.222
## Classical                     -0.617696   0.021543 -28.673
```

```
##
## Correlation matrix not shown by default, as p = 27 > 20.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

```r
fit.p.5 = lmer(data = df, Popular ~   (1 | Subject:Instrument) +
                                      (1 | Subject:Harmony) +
                                      (1 | Subject:Voice) +
                                      Harmony +
                                      Instrument +
                                      Voice +
                                      X16.minus.17 +
                                      X1990s2000s +
                                      Composing +
                                      GuitarPlay +
                                      Classical)
summary(fit.p.5)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Popular ~ (1 | Subject:Instrument) + (1 | Subject:Harmony) +
##     (1 | Subject:Voice) + Harmony + Instrument + Voice + X16.minus.17 +
##     X1990s2000s + Composing + GuitarPlay + Classical
##    Data: df
##
## REML criterion at convergence: 5698
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1797 -0.5102  0.0003  0.4585  5.5431
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2664   0.5162
##  Subject:Voice      (Intercept) 0.0000   0.0000
##  Subject:Instrument (Intercept) 1.3527   1.1630
##  Residual                       1.7787   1.3337
```

```
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Voice, 129; Subject:Instrument, 129
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)        8.48327    0.48499  17.492
## HarmonyI-V-IV      0.02383    0.14706   0.162
## HarmonyI-V-VI      0.25963    0.14822   1.752
## HarmonyIV-I-V     -0.21526    0.14701  -1.464
## Instrumentpiano   -0.14345    0.26670  -0.538
## Instrumentstring -0.83683    0.27512  -3.042
## Voicepar3rd       -0.05355    0.08369  -0.640
## Voicepar5th        0.05170    0.08349   0.619
## X16.minus.17      -0.01834    0.04383  -0.418
## X1990s2000s        0.11750    0.09303   1.263
## Composing          0.21442    0.11101   1.932
## GuitarPlay         0.03078    0.09737   0.316
## Classical         -0.60957    0.02136 -28.537
##
## Correlation of Fixed Effects:
##             (Intr) HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r Vcpr5t
## HrmnyI-V-IV -0.152
## HrmnyI-V-VI -0.131  0.496
## HrmnyIV-I-V -0.150  0.500  0.497
## Instrumntpn -0.250  0.000  0.016  0.001
## Instrmntstr -0.218  0.000  0.034  0.002  0.513
## Voicepar3rd -0.101 -0.001 -0.013  0.000 -0.013    -0.028
## Voicepar5th -0.097 -0.001 -0.012 -0.001 -0.010    -0.021     0.504
## X16.mins.17 -0.485  0.000 -0.006  0.000 -0.007    -0.015     0.006  0.004
## X1990s2000s -0.856  0.000  0.002  0.000  0.002     0.004    -0.002 -0.001
## Composing   -0.357 -0.001  0.005  0.000  0.004     0.011    -0.004 -0.003
## GuitarPlay   0.330  0.001  0.001  0.000  0.002     0.000     0.000  0.000
## Classical   -0.156  0.001 -0.122 -0.009 -0.132    -0.279     0.103  0.077
##             X16..1 X19902 Cmpsng GtrPly
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## Voicepar3rd
## Voicepar5th
## X16.mins.17
## X1990s2000s  0.467
## Composing    0.016  0.287
## GuitarPlay  -0.231 -0.374 -0.695
## Classical    0.054 -0.015 -0.040 -0.001
```

```
fit.p.6 = lmer(data = df, Popular ~   (1 | Subject:Instrument) +
                                      (1 | Subject:Harmony) +
                                      Harmony +
                                      Instrument +
                                      Voice +
                                      X16.minus.17 +
                                      X1990s2000s +
```

```r
                                    Composing +
                                    GuitarPlay +
                                    Classical)
summary(fit.p.6)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Popular ~ (1 | Subject:Instrument) + (1 | Subject:Harmony) +
##     Harmony + Instrument + Voice + X16.minus.17 + X1990s2000s +
##     Composing + GuitarPlay + Classical
##    Data: df
##
## REML criterion at convergence: 5698
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1797 -0.5102  0.0003  0.4585  5.5431
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2664   0.5162
##  Subject:Instrument (Intercept) 1.3527   1.1630
##  Residual                       1.7787   1.3337
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Instrument, 129
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)        8.48327    0.48499  17.492
## HarmonyI-V-IV      0.02383    0.14706   0.162
## HarmonyI-V-VI      0.25963    0.14822   1.752
## HarmonyIV-I-V     -0.21526    0.14701  -1.464
## Instrumentpiano   -0.14345    0.26670  -0.538
## Instrumentstring  -0.83683    0.27512  -3.042
## Voicepar3rd       -0.05355    0.08369  -0.640
## Voicepar5th        0.05170    0.08349   0.619
## X16.minus.17      -0.01834    0.04383  -0.418
## X1990s2000s        0.11750    0.09303   1.263
## Composing          0.21442    0.11101   1.932
## GuitarPlay         0.03078    0.09737   0.316
## Classical         -0.60957    0.02136 -28.537
##
## Correlation of Fixed Effects:
##             (Intr) HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r Vcpr5t
## HrmnyI-V-IV -0.152
## HrmnyI-V-VI -0.131  0.496
## HrmnyIV-I-V -0.150  0.500  0.497
## Instrumntpn -0.250  0.000  0.016  0.001
## Instrmntstr -0.218  0.000  0.034  0.002  0.513
## Voicepar3rd -0.101 -0.001 -0.013  0.000 -0.013    -0.028
## Voicepar5th -0.097 -0.001 -0.012 -0.001 -0.010    -0.021     0.504
## X16.mins.17 -0.485  0.000 -0.006  0.000 -0.007    -0.015     0.006  0.004
## X1990s2000s -0.856  0.000  0.002  0.000  0.002     0.004    -0.002 -0.001
## Composing   -0.357 -0.001  0.005  0.000  0.004     0.011    -0.004 -0.003
```

```
## GuitarPlay   0.330  0.001  0.001  0.000  0.002     0.000     0.000  0.000
## Classical   -0.156  0.001 -0.122 -0.009 -0.132    -0.279     0.103  0.077
##             X16..1 X19902 Cmpsng GtrPly
## HrmnyI-V-IV
## HrmnyI-V-VI
## HrmnyIV-I-V
## Instrumntpn
## Instrmntstr
## Voicepar3rd
## Voicepar5th
## X16.mins.17
## X1990s2000s  0.467
## Composing    0.016  0.287
## GuitarPlay  -0.231 -0.374 -0.695
## Classical    0.054 -0.015 -0.040 -0.001
```

PART A

From model fit.p.6, we observe a significant effect of only Instrument on Popular rating, with the effect of Harmony and Voice being not significantly other than zero. Detail below (all interpretations below should be read as "… when controlling for all other predictors":
Instrument:
- guitar: baseline
- piano: Samples featuring this instrument do not have Popular ratings significantly different than the baseline.
- string: Samples featuring this instrument have significantly lower Popular ratings than the baseline.
Voice:
- contrary: baseline
- par3rd: Samples featuring this type of voice do not have significantly different Popular ratings than the baseline.

9

- par5th: Samples featuring this type of voice do not have significantly different Popular ratings than the baseline.
Harmony:
- I-IV-V: baseline
- I-V-IV: Samples of this pattern do not have significantly different Popular ratings than the baseline.
- I-V-VI: Samples of this pattern do not have significantly different Popular ratings than the baseline.
- IV-I-V: Samples of this pattern do not have significantly different Popular ratings than the baseline.

PART B

Variable interpretations of model fit.p.6 (all interpretations are "… controlling for other predictors"):
- Harmony, Instrument, Voice: Interpretation the same as in part 4) b)
- X16.minus.17: For each additional unit of being able to discern popular from classical music, an individual will on average rate a sample as being .01 units less Popular. - X1990s2000s: With each additional unit of listening to 90s and 2000s music, an individual will on average rate a sample as being .12 units more Popular.
- Composing: With each additional unit of music composition, an individual will on average rate a sample as being .3 units more Classical.

6

- GuitarPlay: With each additional unit of guitar play, an individual will on average rate a sample as being .21 units more Popular.
- Classical: The more an individual rates a particular sample as being Classical, the less they will rate the same sample as being Popular (at a rate of -.61 Popular units for each Classical unit). This is interesting as it calls into question the independence assumption made by the study creators.

should not be in the model

PART C

Next modeling the interaction between being a Musician and other predictors, we find only one interaction to be significant. Specifically (all interpretations are "… controlling for other predictors"): - Musician:Classical: On average musicians rate a sample as .11 units more Classical than non-musicians.

these results are strange compared to others' analysis and my sense of the data

7

```r
fit.p.7 = lmer(data = df2, Popular ~ (1 | Subject:Instrument) +
                                     (1 | Subject:Harmony) +
                                     Harmony +
                                     Instrument +
                                     Voice +
                                     X16.minus.17 +
                                     X1990s2000s +
                                     Composing +
                                     GuitarPlay +
                                     Classical +
              Musician*Harmony +
              Musician*Instrument +
              Musician*Voice +
              Musician*X16.minus.17 +
              Musician*X1990s2000s +
              Musician*Composing +
              Musician*GuitarPlay +
              Musician*Classical)
summary(fit.p.7)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Popular ~ (1 | Subject:Instrument) + (1 | Subject:Harmony) +
##     Harmony + Instrument + Voice + X16.minus.17 + X1990s2000s +
##     Composing + GuitarPlay + Classical + Musician * Harmony +
##     Musician * Instrument + Musician * Voice + Musician * X16.minus.17 +
##     Musician * X1990s2000s + Musician * Composing + Musician *
##     GuitarPlay + Musician * Classical
##    Data: df2
##
## REML criterion at convergence: 5703
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1188 -0.5061 -0.0053  0.4662  5.6248
##
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.2862   0.5349
##  Subject:Instrument (Intercept) 1.3621   1.1671
##  Residual                       1.7733   1.3316
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Instrument, 129
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                  8.713529   0.754699  11.546
## HarmonyI-V-IV                0.006407   0.204980   0.031
## HarmonyI-V-VI                0.310380   0.205263   1.512
## HarmonyIV-I-V               -0.214188   0.204979  -1.045
## Instrumentpiano             -0.094928   0.367446  -0.258
## Instrumentstring            -0.786858   0.384503  -2.046
## Voicepar3rd                 -0.007279   0.114311  -0.064
## Voicepar5th                  0.021031   0.113749   0.185
```

```
## X16.minus.17                  -0.029094   0.055291  -0.526
## X1990s2000s                    0.110760   0.156258   0.709
## Composing                      0.149320   0.210542   0.709
## GuitarPlay                     0.646196   0.363348   1.778
## Classical                     -0.663116   0.031023 -21.375
## MusicianYes                   -0.707316   1.068655  -0.662
## HarmonyI-V-IV:MusicianYes      0.039354   0.300882   0.131
## HarmonyI-V-VI:MusicianYes     -0.173029   0.304593  -0.568
## HarmonyIV-I-V:MusicianYes     -0.009175   0.300779  -0.031
## Instrumentpiano:MusicianYes   -0.051691   0.536536  -0.096
## Instrumentstring:MusicianYes -0.006273   0.553916  -0.011
## Voicepar3rd:MusicianYes       -0.109375   0.167594  -0.653
## Voicepar5th:MusicianYes        0.068931   0.167211   0.412
## X16.minus.17:MusicianYes       0.040391   0.096561   0.418
## X1990s2000s:MusicianYes        0.037159   0.205929   0.180
## Composing:MusicianYes          0.132192   0.255347   0.518
## GuitarPlay:MusicianYes        -0.686605   0.381739  -1.799
## Classical:MusicianYes          0.105702   0.043375   2.437
```

```
##
## Correlation matrix not shown by default, as p = 26 > 20.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

problem 5: 10