1 a 9/9
b 9/9
c 9/9

2 a 7/9
b 9/9
c 9/9

3 6/9

4 a 9/9
b 9/9
c 9/9

5 10/10

Total 95/100

Homework 5

Nikhil Kumar
36-763

18 December 2015

All work was done in R and the code can be seen at the end of this project in the R appendix.

1. (a) We are going to consider the linear model of `Classical` regressed on `Harmony`, `Instrument` and `Voice`. Note how the response variable could only take integer values, thus not making it a truly continuous variable. However, for the sake of this question, a typical linear model with normally distributed errors was made. In future analysis a poisson regression should also be considered. The following is the summary of this model.

	Estimate	Std. Error	t value	$Pr(>\|t\|)$
(Intercept)	4.3404	0.1299	33.41	0.0000
HarmonyI-V-IV	-0.0309	0.1301	-0.24	0.8122
HarmonyI-V-VI	0.7689	0.1301	5.91	0.0000
HarmonyIV-I-V	0.0510	0.1300	0.39	0.6947
Instrumentpiano	1.3736	0.1130	12.15	0.0000
Instrumentstring	3.1310	0.1123	27.87	0.0000
Voicepar3rd	-0.4128	0.1127	-3.66	0.0003
Voicepar5th	-0.3717	0.1127	-3.30	0.0010

Model 1(a)

The harmony variable appears to have insignificant factors. To examine that in detail, let us consider the reduced model of `Classical` regressed on `Harmony` and `Instrumnet`. Thus to check the adjusted signficicance of the `Harmony` variable, we can do analysis of variance test. The following is the summary of that test.

9

	Res.Df	RSS	Df	Sum of Sq	F	$Pr(>F)$
1	2488	13388.21				
2	2485	13114.89	3	273.32	17.26	0.0000

As seen by the low p-value, the effect of `Harmony` is significant in predicting `Classical`. Thus our final model for this part will be the linear model made in the previous step (model 1(a)). Based on model 1(a), we can see how the expected classical score changes for each factor level of each of the predictors. For example, as a subject's `Harmony` level changes from I-IV-V (baseline level) to I-V-IV, the `Classical` score is expected to decrease by 0.0309. Note how the expected change in `Classical` is positive for each of the `Instrument` levels and negative for each of the `Voice` levels.

(b) First, let $y =$ classical, $x_1 =$ harmony, $x_2 =$ voice, $x_3 =$ instrument
The syntax for the repeated measures model is: `lmer(y` $\sim$ x_1 `+` x_2 `+` x_3 `+ (1|Subject)`
Thus this model written as a hierarchical model is:
$y_i = \alpha_{j[i]} + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \epsilon_i$
$\epsilon \sim N(0, \sigma^2)$
$\alpha_{j[i]} = \beta_0 + \eta_j$
$\eta_j \sim N(0, \tau^2)$

The following summarizes the attributes for the measures model.

	Estimate	Std. Error	t value
(Intercept)	4.34	0.19	22.96
HarmonyI-V-IV	-0.03	0.11	-0.30
HarmonyI-V-VI	0.77	0.11	7.19
HarmonyIV-I-V	0.05	0.11	0.47
Instrumentpiano	1.38	0.09	14.77
Instrumentstring	3.13	0.09	33.80
Voicepar3rd	-0.42	0.09	-4.47
Voicepar5th	-0.38	0.09	-4.04

Fixed Effects for model 1(b)

	(Intercept)	HarmonyI-V-IV	HarmonyI-V-VI	HarmonyIV-I-V
(Intercept)	0.04	-0.01	-0.01	-0.01
HarmonyI-V-IV	-0.01	0.01	0.01	0.01
HarmonyI-V-VI	-0.01	0.01	0.01	0.01
HarmonyIV-I-V	-0.01	0.01	0.01	0.01
Instrumentpiano	-0.00	0.00	0.00	-0.00
Instrumentstring	-0.00	-0.00	-0.00	-0.00
Voicepar3rd	-0.00	-0.00	0.00	0.00
Voicepar5th	-0.00	-0.00	-0.00	-0.00

Correlation of Fixed Effects for Model 1(b)

	Instrumentpiano	Instrumentstring	Voicepar3rd	Voicepar5th
(Intercept)	-0.00	-0.00	-0.00	-0.00
HarmonyI-V-IV	0.00	-0.00	-0.00	-0.00
HarmonyI-V-VI	0.00	-0.00	0.00	-0.00
HarmonyIV-I-V	-0.00	-0.00	0.00	-0.00
Instrumentpiano	0.01	0.00	-0.00	-0.00
Instrumentstring	0.00	0.01	-0.00	-0.00
Voicepar3rd	-0.00	-0.00	0.01	0.00
Voicepar5th	-0.00	-0.00	0.00	0.01

Correlation of Fixed Effects for Model 1(b)

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	1.704	1.305
Residual		3.583	1.893

Random Effects for Model 1(b)

To test the performance of this model, relative to the one created in the previous part, we are going to consider the AIC, BIC, and in-sample MSE for these models. Note that for the MSE, conditional residuals were used. The following table summarizes these measurements.

nice to consider this.

9

	MSE	AIC	BIC
linear	5.26	11231.86	11284.25
repeated measures	3.48	10492.76	10550.97

Comparing Models 1(a), 1(b)

The repeated measures model is better for all three measurements. Thus, we have evidence to conclude that the random effect adds value to the model. We also observe the fact that fixed effect estimates are very similar to model 1(a) and the correlation between each level of each predictor are very low.

(c) The syntax (using the variables defined in part (c)) for this model is: `lmer(y` $\sim x_1$ `+` x_2 `+` x_3 `+ (1|Subject:`x_1`) + (1|Subject:`x_2`) + (1|Subject:`x_3`)` This model written as a hierarchical model is:
$y_i = \alpha_{j[i]} + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$
$\alpha_{0i} = \beta_{00} + \eta_{0i}, \eta_{0i} \sim N(0, \tau_0^2)$
$\alpha_{1i} = \beta_{10} + \eta_{1i}, \eta_{1i} \sim N(0, \tau_1^2)$
$\alpha_{2i} = \beta_{20} + \eta_{2i}, \eta_{2i} \sim N(0, \tau_2^2)$
This new model (model 1(c)) has the following performance measurements.

	MSE	AIC	BIC
model 1(a)	5.26	11231.86	11284.25
model 1 (b)	3.48	10492.76	10550.97
model 1(c)	2.11	10077.65	10147.50

Comparing Models 1(a), 1(b), 1(c)

All of these measurements are better for model 1(c) where there are random effects for each fixed effect. The following summarizes the fixed effects for this mode.

	Estimate	Std. Error	t value
(Intercept)	4.34	0.21	20.25
HarmonyI-V-IV	-0.03	0.14	-0.21
HarmonyI-V-VI	0.77	0.14	5.38
HarmonyIV-I-V	0.06	0.14	0.40
Instrumentpiano	1.36	0.26	5.20
Instrumentstring	3.13	0.26	11.93
Voicepar3rd	-0.41	0.08	-4.98
Voicepar5th	-0.37	0.08	-4.55

Fixed Effects for Model 1(c)

Again, notice how the fixed effects do not differ from the fixed effects from models 1(a) and 1(b). The following shows each random effect's variance.

9

Group	Variance
Subject:Harmony	0.44300
Subject:Instrument	2.19904
Subject:Voice	0.02802
Residual	2.43997

Variance for Model 1(c)'s Random Effects

The `Subject:Instrument` group has a variance that is similar to the residual's variance. The `Subject:Harmony` and `Subject:Voice` variance's are very similar relative the residual's variance.

7

9

9

2. (a) The covariates `Voice`, `OMSI`, `Composing`, `ClsListen` and `Instr.minus.Notes` were selected as possible variables to add to the model. They were chosen for intuitive reasons. Furthermore, the following table summarizes the performance of this new model compared to model 1(c).

	MSE	AIC	BIC
Model 1(c)	2.11	10077.65	10147.50
Model 2(a)	2.05	9673.74	9766.25

Comparing Models 1(c) and 2(a)

did you also check these covariates individually?

did you look at any other covariates?

Since all three measurements confirm that the new model is better, we have evidence to conclude that these new covariates add value to the model.

(b) We will consider two new sets of random effects. The first set of random effects to include in the model is of the form `(1|Subject:Harmony) + (1|Subject:Composing) + (1|Subject:Instrument) + (1|Subject:Voice)`. We will call this model 2(b)i. The second set of random effects to include in the model is of the form `(1|Subject:Harmony) + (1|Subject:Composing) + (1|Subject:Instrument) + (1|Subject:Voice) + (1|Selfdeclare:Voice)`. We will call this model 2(b)ii. Using our usual criteria of evaluating and comparing AIC/BIC/MSE values, we will see if these new models with new random effects perform better than model 2(a).

although this wasn't discussed explicitly in class, random effects are usually only fitted to account for design variables (inst, voice, harmony, and subject, in this case), rather than response variables (composing, self-declare, e.g.).

	MSE	AIC	BIC
Model 2(a)	2.05	9673.74	9766.25
Model 2(b)i	2.07	9641.73	9740.02
Model 2(b)ii	2.07	9643.73	9747.81

Model 2(b)i and 2(b)ii have much better AIC/BIC values than model 2(a). We would then remove model 2(a) from consideration. However, between model 2(b)i and model 2(b)ii, their AIC and MSE values are too close to each other to make a decision based on this criteria. However, there is a big enough difference in BIC between 2(b)i and model 2(b)ii. Thus, we would select model2(b)i as our final model for this part.

(c) For this question model 2(b)i was our final model (which for future parts will be called m2). The following table summarizes the fixed effects for this final model.

a little confusing for skim-readers (renaming the model)....

	Estimate	Std. Error	t value
(Intercept)	3.93172	0.37175	10.57612
HarmonyI-V-IV	-0.01258	0.14348	-0.08768
HarmonyI-V-VI	0.79749	0.14351	5.55714
HarmonyIV-I-V	0.05731	0.14345	0.39955
Instrumentpiano	1.40589	0.20766	6.77008
Instrumentstring	3.18769	0.20738	15.37091
Voicepar3rd	-0.37067	0.08110	-4.57031
Voicepar5th	-0.34571	0.08108	-4.26361
OMSI	-0.00040	0.00090	-0.44712
Composing	0.10915	0.14736	0.74068
ClsListen	0.14893	0.11531	1.29156
Instr.minus.Notes	0.05553	0.10252	0.54164

Final Model 2

As can be seen by the summary, the `Instrument` levels have the highest coefficients and thus the greatest expected effect on the response variable. This is especially true for when a subject changes their instrument from guitar (baseline) to string, the expected value of the response variable will increase by 3.18769, holding all other effects constant. All other coefficients corresponding to categorical variable have this type of an interpretation. On the other end of the spectrum, `OMSI` seems to have the least expected effect on the response variable. As the `OMSI` test score increases by one point, the expected value of the response variable increases only by 0.00040, holding all other effects constant. All other coefficients corresponding to numeric variables have this type of an interpretation.

3. In order to dichotomize `Selfdeclare` variable, the median was found to be 2. Since 2 represents the 50th percentile, all values that are less than 2 were coded as 0 and all values greater than or equal to 2 were coded as 1. With this new variable, `Selfdeclare.binary`, the following two models were considered. The first one was of the form `Classical ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes + Selfdeclare.binary + Selfdeclare.binary:Composing + (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Selfdeclare.binary:Voice)`, and the second one was of the form `lmer(Classical ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes + Selfdeclare.binary + Selfdeclare.binary*Composing + Selfdeclare.binary*ClsListen + (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Selfdeclare.binary:Voice)`. The first model will be called model 3i and the second model will be called model 3ii. Using our usual criteria of AIC/BIC/MSE, we will see which model is the best one. Furthermore, we will compare these two models with the final model from 2.

why these two models in particular? What about other covariates, and/or interactions of selfdeclare.binary with other covariates?

ok given above, but again it's a bit unusual to make a random effect out of a response variable.

6

	MSE	AIC	BIC
model 2	2.07	9641.73	9740.02
model 3i	2.08	9675.13	9779.20
model 3ii	2.08	9675.76	9785.62

According to this criteria, model 2 still performed better than either model 3i or model 3ii. This gives us some evidence the controlling for the dichotomized `Selfdeclare` does not add much value to the model. Furthermore, recall how model 2(b)ii also controlled for the non-dichotomized `Selfdeclare` as a random intercept, but that model was still not as good as model 2(b)i (which did not use the `Selfdeclare` variable in any way). Thus based on these models, and model selection criteria, we have some evidence to conclude that `Selfdeclare` (coded as either form) adds no predictive power when modeling how classical the stimulus sounds.

4. (a) Using a similar model to the one from 1(c) (where `Popular` is regressed on `Subject`, `Harmony`, and `Voice` in addition to each predictor's interaction with `Subject` as random effects), creates the following set of fixed effects.

9

	Estimate	Std. Error	t value
(Intercept)	6.58	0.21	31.77
HarmonyI-V-IV	-0.03	0.14	-0.18
HarmonyI-V-VI	-0.27	0.14	-1.93
HarmonyIV-I-V	-0.19	0.14	-1.32
Instrumentpiano	-0.95	0.25	-3.77
Instrumentstring	-2.61	0.25	-10.37
Voicepar3rd	0.16	0.08	1.97
Voicepar5th	0.16	0.08	1.95

Model 4(a)

In this model, the levels of `Voice` all have positive effects on how popular the stimulus sounds. The levels for the other two predictors are all negative. This is opposite to what was observed when `Classical` was the response variable.

(b) The same set of covariates that were used for question 2 were also considered for this question. In doing so, two models were created. The first one was of the form `lmer(Popular ~ Harmony +`

`Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes + (1|Subject:Harmony) + (1|Subject:Composing) + (1|Subject:Instrument) + (1|Subject:Voice))`. This model
9 will be called model 4(b)i. The second model is of the form `lmer(Popular ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes + (1|Subject:Harmony) + (1|Subject:Composing) + (1|Subject:Instrument) + (1|Subject:Voice) + (1|Selfdeclare:Voice))`. This model will be called model 4(b)ii. Using our usual model selection criteria, we have the following measurements.

	MSE	AIC	BIC
model 4(b)i	2.15	9669.00	9767.29
model 4(b)ii	2.15	9671.00	9775.07

Similar to the results from question 2, these two models have similar MSE and AIC values. However, the BIC does favor model 4(b)i. Thus based on this criteria, we will select model 4(b)i. Doing this results in the following fixed effects.

	Estimate	Std. Error	t value
(Intercept)	6.53264	0.35605	18.34742
HarmonyI-V-IV	-0.02875	0.13906	-0.20676
HarmonyI-V-VI	-0.30318	0.13908	-2.17984
HarmonyIV-I-V	-0.21228	0.13902	-1.52692
Instrumentpiano	-0.98791	0.19710	-5.01213
Instrumentstring	-2.69424	0.19681	-13.68961
Voicepar3rd	0.17796	0.08245	2.15839
Voicepar5th	0.16911	0.08243	2.05155
OMSI	0.00038	0.00086	0.44042
Composing	0.14508	0.14101	1.02887
ClsListen	-0.01298	0.11034	-0.11762
Instr.minus.Notes	-0.10870	0.09811	-1.10802

Model 4(b)

With this final model, the `Instrument` levels have the highest coefficients and thus the greatest expected effect on the response variable. This is especially true when a subject changes their instrument from guitar (baseline) to string, the expected value of the response variable will decrease by about 2.694. On the other end of the spectrum, `OMSI` seems to have the least expected effect on the response variable. As the `OMSI` score increases by one point, the expected value of the response variable increases only by 0.00038.

(c) Using the same dichotomized `Selfdeclare.binary` variable, two models will be considered. The first one will be of the form, `Popular ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes + Selfdeclare.binary + Selfdeclare.binary:Composing + (1|Subject:Harmony) + (1Subject:Instrument) + (1|Selfdeclare.binary:Voice)`, and the second one was of the form `lmer(Popular ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes + Selfdeclare.binary + Selfdeclare.binary*Composing + Selfdeclare.binary*ClsListen + (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Selfdeclare.binary:Voice)`. The first model will be called model 4(c)i and the second model will be called model 4(c)ii.

	MSE	AIC	BIC
model4(b)	2.15	9669.00	9767.29
model 4(c)i	2.16	9694.85	9798.93
model 4(c)ii	2.16	9698.63	9808.48

9

Based on this criteria, model 4(b) is better (the final model from the previous part). Thus, using the dichotomized `Selfdeclare.binary` model did not improve our modeling of the response variable. The fixed effects and interpretations are the same to what they were in the previous part.

5. See write-up on next page.

In this analysis, we were interested in creating two different models. The first model was to identify what variables, within the data set, affect how classical the musical stimulus sounds. The second model was to identify what variables, within the data set, affect how popular the musical stimulus sounds. For both research questions, traditional linear models and hierarchical models were considered.

It's confusing to flip back and forth between the two outcomes...

For the first research question of identifying what factors affect how popular a song sounds, three major predictors were considered. These were the harmony motions (I-IV-V, I-V-IV, I-V-VI, IV-I-V), voice leadings (contrary, par3rd, par5th) and the instrument (guitar, piano, string). The first model was a linear model in where the response variable of how classical a song was linearly regressed on these three main predictors. For this linear model, all three predictors were observed to be significant. The second model that was considered was a repeated measures model in where the previous model was fit with a random intercept for each participant. The third model considered was one where there was a random intercept for each person/harmony combination, person/voice combination, and person/instrument combinations. The reasoning behind this model was to account for all possible forms of bias that people may have when rating music. In order to compare between these three models, the in-sample mean squared error, AIC, and BIC measurement were all used. In doing so, the third model where there was a random intercept for each person/harmony, person/voice and person/instrument combination performed the best. This tells us that including other variance components adds values to the model.

Now we are not sure which outcome you are modeling...

Moving past the main predictors of harmony, voice and instrument, other predictors were also considered at this point. Specifically, the OMSI test of musical knowledge score for each subject, the self-reported composing experience of each subject, the self-reported score of how much classical music each subject listens to, and the difference between how much a subject concentrates on an instrument while listening to music and how much a subject concentrates on the notes while listening to music were considered. At first, 10 each of these covariates were added as fixed effects to the model from the previous part. Then these variable's interaction with each subject were considered as random intercepts in the model. Using the MSE/AIC/BIC criteria mentioned in the previous paragraph, the model where all these new covariates were added as fixed effects, but not as random intercepts was seen to be the best. Note that this represents the final model 2 from the earlier questions. The last covariate that was considered was the measurement of if they self declare themselves as musicians. This variable was first dichotomized by making the cut-off value equal to the median of this variable. This variable was then added as a fixed effect and as a fixed effect with its interaction with the voice variable. In addition the combination between the self-declare variable and the voice variable was added as a random intercept to the model. Unfortunately, adding this new variable did not show any improvement to the model and thus it was not considered. All of this analysis has shown that a hierarchical model performs very well when trying to predict how classical a song sounds. The final model that was selected form this analysis was the one displayed in question 2(c). With this final model, the instrument levels have the highest coefficients and thus the greatest expected effect on the response variable. This is especially true for when a subject changes their instrument from guitar (baseline) to string, the expected value of the response variable will increase by 3.18769, holding all other effects constant. On the other end of the spectrum, the OMSI score seems to have the least expected effect on the response variable. As this test score increases by one point, the expected value of the response variable increases only by 0.00040, holding all other effects constant.

The second research question was about what factors, within the data set, affect how popular a song sounds. The procedure for answering this question was very similar to the previous question. The three major covariates of harmony motions, voice leadings and the instrument were first considered. Linear and hierarchical models were made similar to those made in the first part of the research question. The model accounting for al possible forms of bias (question 1(c)) performed the best, again telling us that other variance components adds value to the model.

At this point, additional covariates were added to refine the model. The same set of covariates were added to the model. Again models where the covariates were added as fixed effects in addition to models were the covariate's interaction with each subject were added as random intercepts were considered. Similar to the previous question, not surprisingly, the former model performed the best (displayed as model 4(b) in the earlier questions). Lastly the dichotomized self-declare variable was considered. Again, similar to the previous question, this variable, both when added as a fixed effect and as a random intercept did not add any value to the model, thus telling us that the variable is insignificant in predicting the response variable. This analysis continued to show us that a hierarchical model performs very well when trying to predict how popular a song sounds. The final model that selected from this analysis was the one displayed in question 4(b). With this final model, the instrument levels have the highest coefficients and thus the greatest expected effect on the response variable. This is especially true when a subject changes their instrument from guitar (baseline) to string, the expected value of the response variable will decrease by about 2.694. On the other end of the spectrum, OMSI score seems to have the least expected effect on the response variable. As this score increases by one point, the expected value of the response variable increases only by 0.00038.

R Appendix

```
library(lme4)
library(arm)
library(xtable)

the.data = read.table("ratings.txt", sep = ",", header = TRUE)

### QUESTION 1 ###
## Part (a) ##
getMSE = function(object){
  mse = mean(residuals(object)^2)
  return(mse)
}

getModelSummary = function(list.of.objects){
  l = length(list.of.objects)
  model.summary = matrix(rep(NA, l*3), nrow = l, ncol = 3)
  for(i in 1:l){
    model.summary[i, ] = c(getMSE(list.of.objects[[i]]),
                           AIC(list.of.objects[[i]]),
                           BIC(list.of.objects[[i]]))
  }
  colnames(model.summary) = c("MSE", "AIC", "BIC")
  return(model.summary)
}

the.data$Classical = as.integer(the.data$Classical)

m1a = lm(Classical ~ Harmony + Instrument + Voice, data = the.data)
summary(m1a)
xtable(m1a)
m1a.noHarmony = lm(Classical ~ Instrument + Voice, data = the.data)
anova(m1a.noHarmony, m1a)

getModelSummary(list(m1a, m1a.noHarmony))
# m1a looks better

## Part (b) ##
m1b = lmer(Classical ~ Harmony + Instrument + Voice + (1|Subject), data = the.data)
xtable(getModelSummary(list(m1a, m1b)))

names(summary(m1b))
xtable(summary(m1b)$coefficients)

xtable(as.matrix(summary(m1b)$varcor))
m = as.matrix(summary(m1b)$varcor)

xtable(as.matrix(summary(m1b)$vcov)[,1:4])
xtable(as.matrix(summary(m1b)$vcov)[,5:8])

## Part (c) ##
m1c = lmer(Classical ~ Harmony + Instrument + Voice + (1|Subject:Harmony)
           + (1|Subject:Instrument) + (1|Subject:Voice), data = the.data)
```

```
summary.1c = getModelSummary(list(m1a, m1b, m1c))
row.names(summary.1c) = c("model 1(a)", "model 1 (b)", "model 1(c)")
xtable(summary.1c)
summary(m1c)
xtable(as.matrix(summary(m1c)$coefficients))

### QUESTION 2 ###
the.data$CollegeMusic = as.factor(the.data$CollegeMusic)
the.data$APTheory = as.factor(the.data$APTheory)

## Part (a) ##
m2a = lmer(Classical ~ Harmony + Instrument + Voice + OMSI + Composing +
             ClsListen + Instr.minus.Notes + (1|Subject:Harmony) +
             (1|Subject:Instrument) + (1|Subject:Voice), data = the.data)
summary.2a = getModelSummary(list(m1c, m2a))
row.names(summary.2a) = c("Model 1(c)", "Model 2(a)")
xtable(summary.2a)

## Part (b) ##
m2bi = lmer(Classical ~ Harmony + Instrument + Voice + OMSI + Composing +
              ClsListen + Instr.minus.Notes +
              (1|Subject:Harmony) +
              (1|Subject:Composing) +
              (1|Subject:Instrument) +
              (1|Subject:Voice),
            data = the.data)

m2bii = lmer(Classical ~ Harmony + Instrument + Voice + OMSI + Composing +
              ClsListen + Instr.minus.Notes +
              (1|Subject:Harmony) +
              (1|Subject:Composing) +
              (1|Subject:Instrument) +
              (1|Subject:Voice) +
              (1|Selfdeclare:Voice),
            data = the.data)
summary.2b = getModelSummary(list(m2a, m2bi, m2bii))
xtable(summary.2b)

m2b = m2bi

xtable(summary(m2b)$coefficients, digits = 5)

### QUESTION 3 ###
summary(the.data$Selfdeclare)
the.data$Selfdeclare.binary = ifelse(the.data$Selfdeclare <= 2, 0, 1)
table(the.data$Selfdeclare.binary)

m3i = lmer(Classical ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes +
            Selfdeclare.binary +  Selfdeclare.binary:Composing +
            (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Selfdeclare.binary:Voice), data = the.data

m3ii = lmer(Classical ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes
             Selfdeclare.binary +  Selfdeclare.binary*Composing + Selfdeclare.binary*ClsListen +
             (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Selfdeclare.binary:Voice), data = the.da
```

```
summary.3 = getModelSummary(list(m2b, m3i, m3ii))
row.names(summary.3) = c("m2", "m3i", "m3ii")
summary.3
xtable(summary.3)

### QUESTION 4 ###
## Part (a) ##
m4a = lmer(Popular ~ Harmony + Instrument + Voice + (1|Subject:Harmony) + (1|Subject:Instrument) +
             (1|Subject:Voice), data = the.data)
xtable(summary(m4a)$coefficient)

## Part (b) ##
m4bi = lmer(Popular ~ Harmony + Instrument + Voice + OMSI + Composing +
              ClsListen + Instr.minus.Notes +
              (1|Subject:Harmony) +
              (1|Subject:Composing) +
              (1|Subject:Instrument) +
              (1|Subject:Voice),
            data = the.data)

m4bii = lmer(Popular ~ Harmony + Instrument + Voice + OMSI + Composing +
               ClsListen + Instr.minus.Notes +
               (1|Subject:Harmony) +
               (1|Subject:Composing) +
               (1|Subject:Instrument) +
               (1|Subject:Voice) +
               (1|Selfdeclare:Voice),
             data = the.data)

summary.4b = getModelSummary(list(m4bi, m4bii))
row.names(summary.4b) = c("model 4(b)i", "model 4(b)ii")
xtable(summary.4b)

m4b = m4bi
xtable(summary(m4b)$coefficients, digits = 5)

## Part (c) ##
m4ci = lmer(Popular ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes +
             Selfdeclare.binary +  Selfdeclare.binary:Composing +
             (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Selfdeclare.binary:Voice), data = the.da

m4cii = lmer(Popular ~ Harmony + Instrument + Voice + OMSI + Composing + ClsListen + Instr.minus.Notes
              Selfdeclare.binary +  Selfdeclare.binary*Composing + Selfdeclare.binary*ClsListen +
              (1|Subject:Harmony) + (1|Subject:Instrument) + (1|Selfdeclare.binary:Voice), data = the.d

summary.4c = getModelSummary(list(m4b, m4ci, m4cii))
row.names(summary.4c) = c("model4(b)", "model 4(c)i", "model 4(c)ii")
xtable(summary.4c)
```