Stat 414 - Day 27
Case Study Solutions

Part 1: Stage fright can be a serious problem for performers, and understanding the personality underpinnings of performance anxiety is an important step in determining how to minimize its impact. Sadler and Miller (2010) studied the emotional state of musicians before performances and factors which may affect their emotional state. Data were collected on 37 undergraduate music majors over the course of an academic year. Students completed diaries prior to performances, including the Positive Affect Negative Affect (PANAS) before each performance. The negative affect measure of this instrument is used as a measure of performance anxiety. Factors include type of performance (solo, large ensemble, small ensemble), audience (orchestral vs. keyboard/vocalist), age, gender, instrument, and years studying.

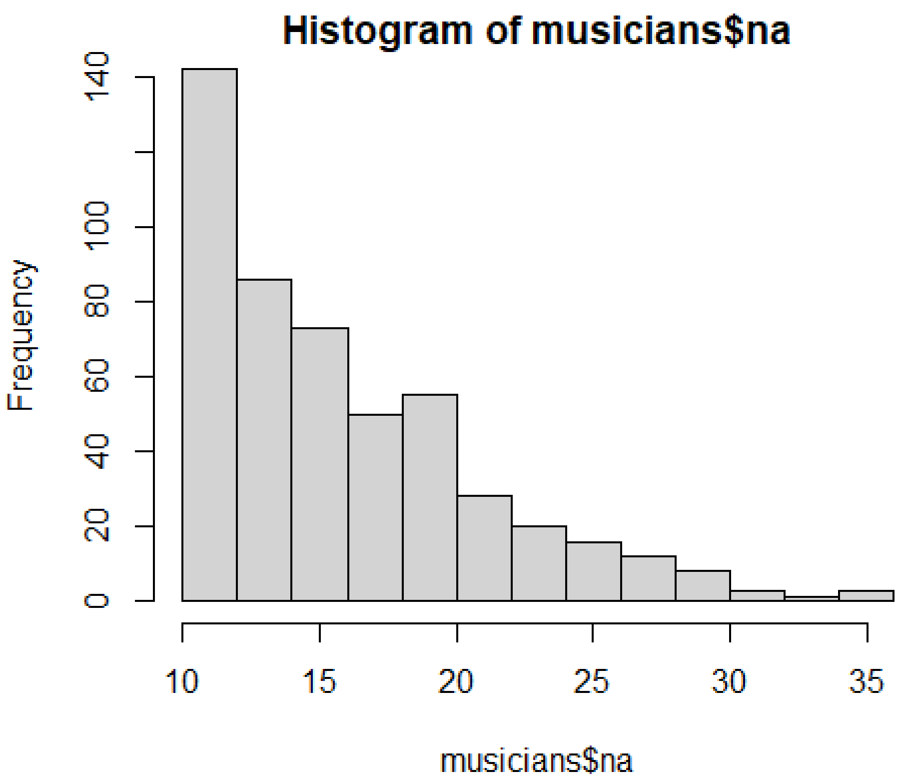
Load in the data and confirm the variables I mentioned.

```
musicians = read.delim("https://www.rossmanchance.com/stat414/data/musicians.txt" ,
"\t", header=TRUE)
head(musicians)
##   subjnum id diary previous perform_type1      memory          audience pa na
## 1       1  1     1        0          Solo Unspecified        Instructor 40 11
## 2       1  1     2        1 LargeEnsemble      Memory PublicPerformance 33 19
## 3       1  1     3        2 LargeEnsemble      Memory PublicPerformance 49 14
## 4       1  1     4        3          Solo      Memory PublicPerformance 41 19
## 5       1  1     5        4          Solo      Memory          Students 31 10
## 6       1  1     6        5          Solo      Memory          Students 33 13
##   age gender instrument1 years_study mpqab mpqsr mpqpem mpqnem mpqcon
## 1  18 Female       voice           3    16     7     52     16     30
## 2  18 Female       voice           3    16     7     52     16     30
## 3  18 Female       voice           3    16     7     52     16     30
## 4  18 Female       voice           3    16     7     52     16     30
## 5  18 Female       voice           3    16     7     52     16     30
## 6  18 Female       voice           3    16     7     52     16     30
```

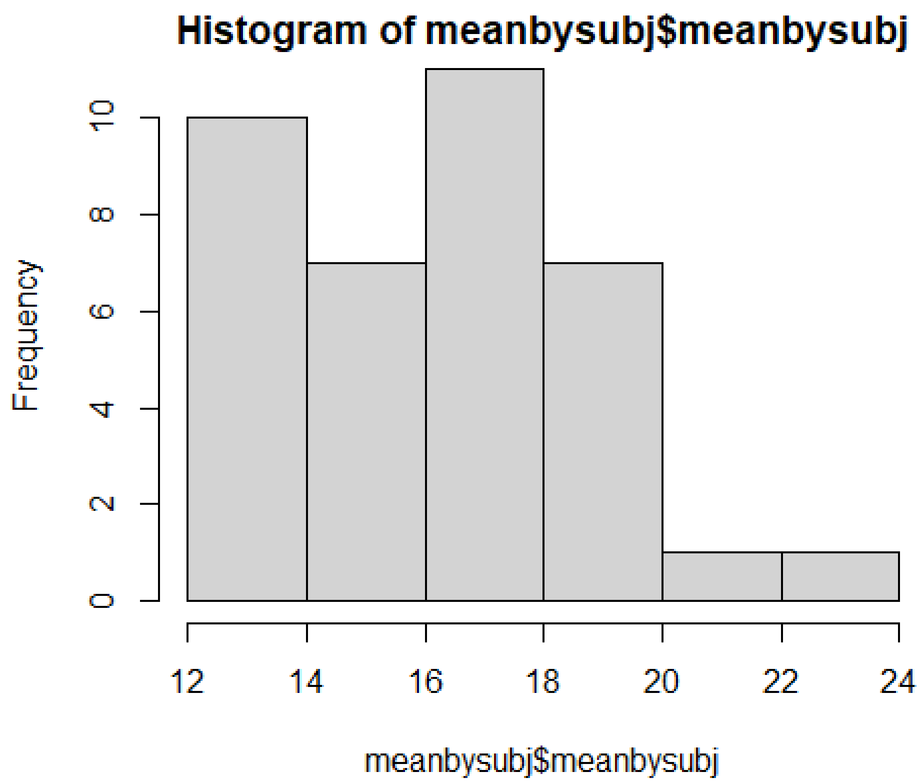
First do some data exploration.

1. Is there much variation in na across performances? Is there much variation in na across musicians?

```
hist(musicians$na)
```




```
meanbysubj <- musicians %>%
  group_by(subjnum) %>%
  summarise(meanbysubj = mean(na, na.rm = TRUE), instrument = head(instrument1, 1),
performtype = head(perform_type1, 1)) %>%
  ungroup()

hist(meanbysubj$meanbysubj)
```



The first histogram includes all the observations (across performances). The second histogram graphs the mean na for each musician (averaging over the number of performances for that musician). There appears to be quite a bit of variability across performances (ranging from 10 to 35 or so) but less in the averages by subject (which is not unexpected for averages to vary less than individual observations).

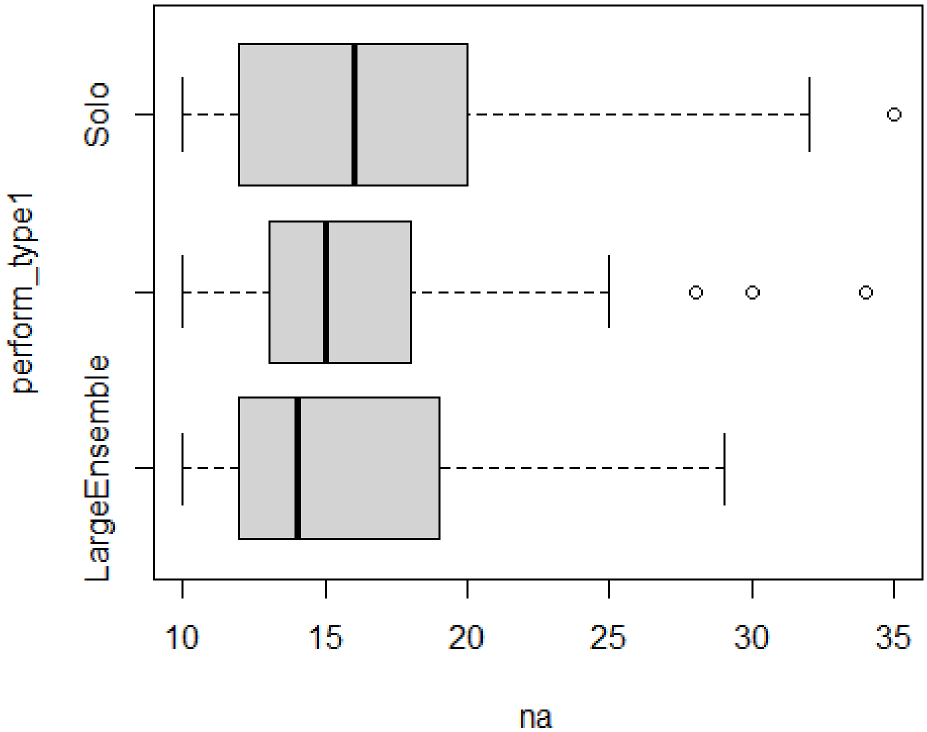
2. Confirm the types of variables/number of categories and start thinking about how you will incorporate these into a model.

```
table(musicians$perform_type1)
##
## LargeEnsemble SmallEnsemble          Solo
##           136            82           279
table(musicians$instrument1)
##
## keyboard(pianoororgan)   orchestralinstrument                  voice
##                     75                    235                    187
```

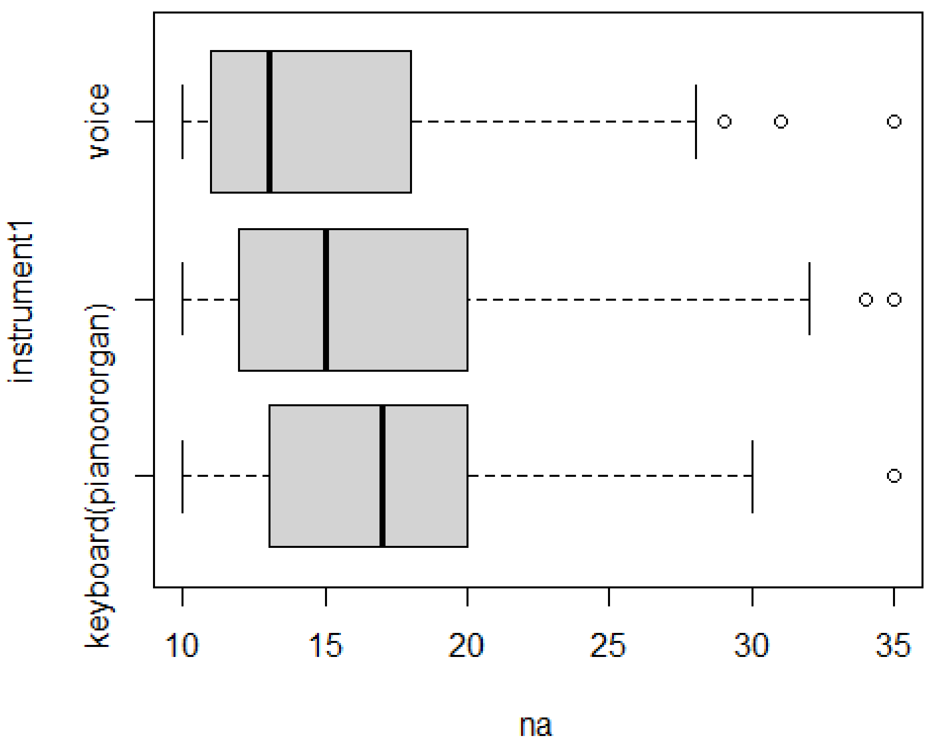
Three performance types and three instruments. So slope coefficients will represent differences in mean na between categories (vs. a reference group if using indicator coding)

3. Does negative affect seem to vary by performance type? What about instrument? Do you think either will be a useful variable to include?

```
boxplot(na ~ perform_type1, data = musicians, horizontal=TRUE)
```

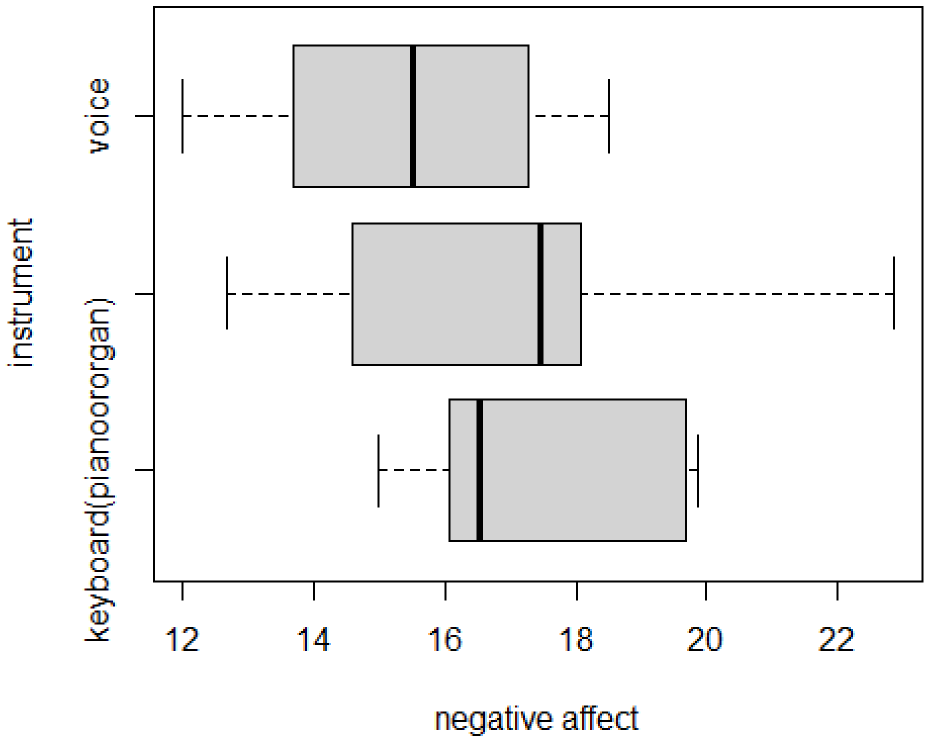



```
boxplot(na ~ instrument1, data = musicians, horizontal=TRUE)
```




```
#what is the following doing?
meanbysubj <- musicians %>%
  group_by(subjnum) %>%
  summarise(meanbysubj = mean(na, na.rm = TRUE), instrument = head(instrument1, 1))
```

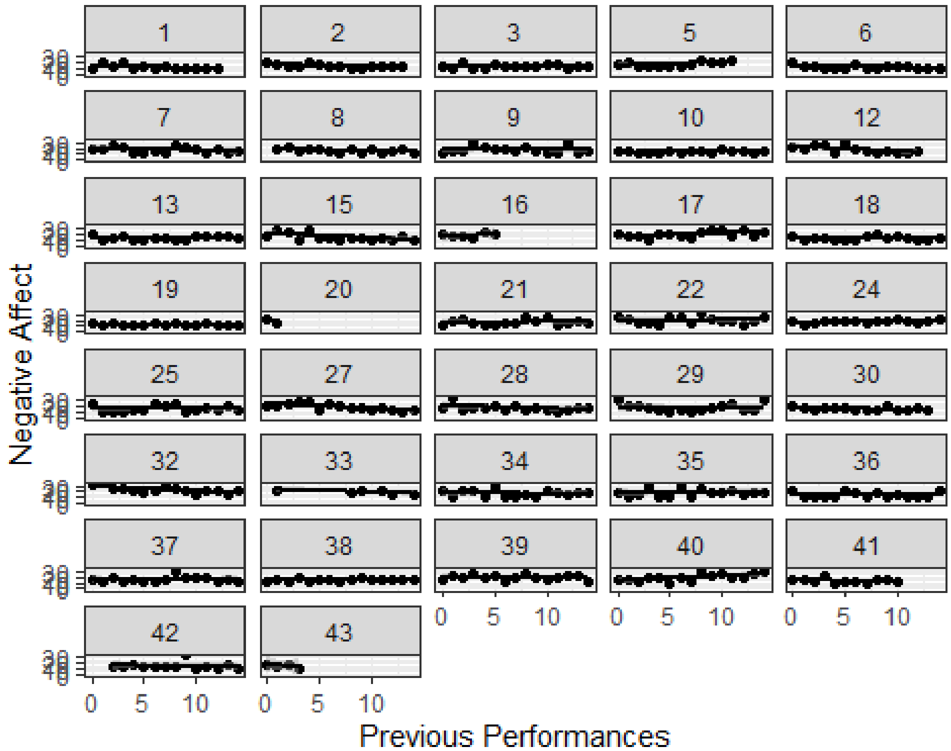
```
%>%
  ungroup()

boxplot(meanbysubj$meanbysubj ~ meanbysubj$instrument, horizontal = TRUE, ylab="ins
trument", xlab="negative affect")
```



There isn't that much variation in na by performance type (from the first graph, ignoring musicians differences). Similarly, we don't see much of a consistent difference in the first boxplot by instrument across all the performances. But when you look at the average anxiety at the musician level, there appears to be less anxiety on average for the vocalists. (The group_by is taking the first instrument listed for each subject since that is a subject-level variable and doesn't change across performances.)

4. Does the relationship between negative affect and number of previous performances appear to differ across musicians? What does that suggest including in the model?

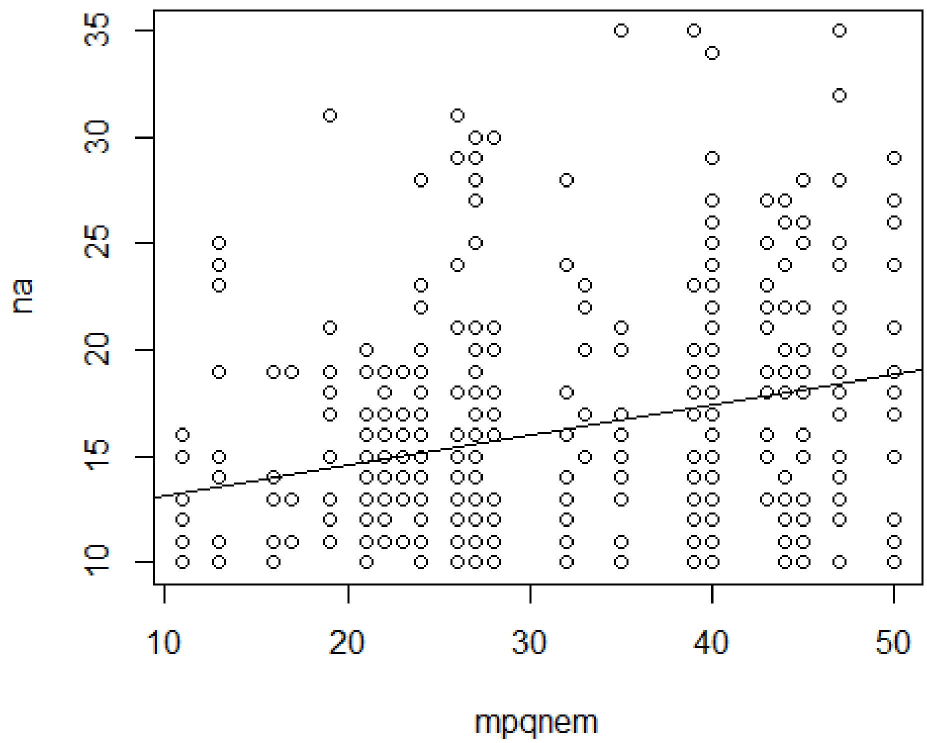
```
ggplot(musicians,aes(x=previous,y=na)) +
  geom_point() + geom_smooth(method="lm",color="black") +
  facet_wrap(~id,ncol=5) +
  labs(x="Previous Performances",y="Negative Affect") +
  theme_bw()
## `geom_smooth()` using formula 'y ~ x'
## Warning in qt((1 - level)/2, df): NaNs produced
## Warning in max(ids, na.rm = TRUE): no non-missing arguments to max; returning
## -Inf
```



The main thing I notice is different number of performances across the musicians, but most of these scatterplots look weakly positive? Seeing different slopes between these graphs would suggest including random slopes in the model (an interaction between previous performances and musician).

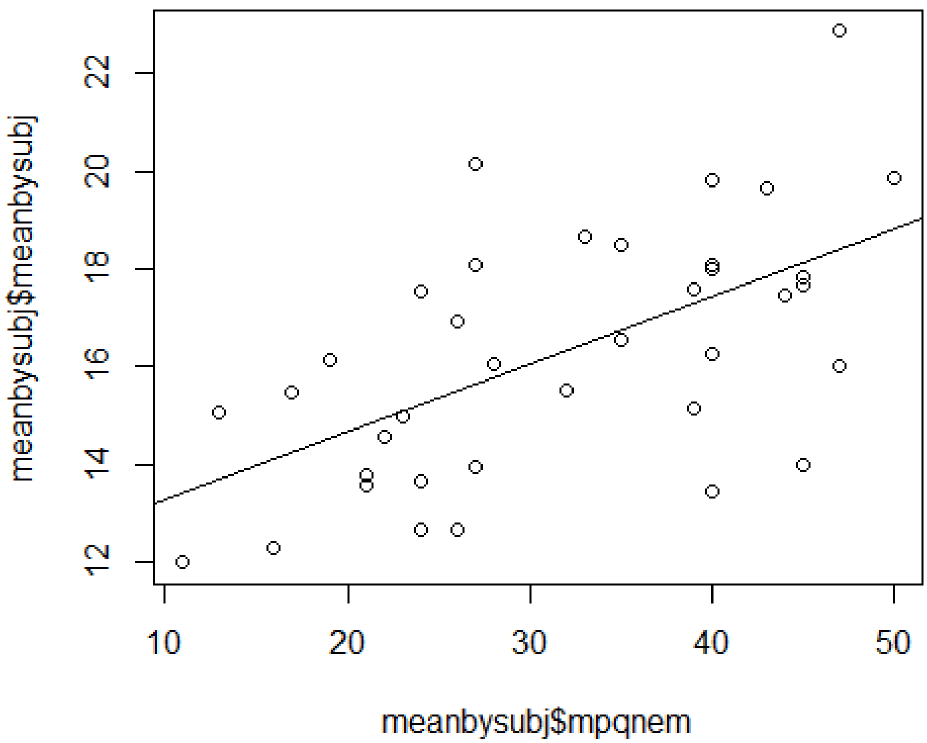
The following plots look at the musicians' negative emotionality composite scale from the MPQ instrument (NEM).

```
plot(na ~ mpqnem, data = musicians); abline(lm(na~mpqnem, data = musicians))
```



Note, mpgnem is a musician level variable (a person's overall disposition). It's not totally clear from the graph above, but I verified that this value stayed constant for each musician. In this case, it doesn't matter whether we use the first value or the average (they are identical). If we do see some changes in values over the course the study for any of the musicians, then you could treat as a Level 1 variable or you could take the average, realizing this will ignore any "measurement error" in that value.

```
meanbysubj <- musicians %>%
  group_by(subjnum) %>%
  summarise(meanbysubj = mean(na, na.rm = TRUE), instrument = head(instrument1, 1),
mpqnem = head(mpqnem, 1)) %>%  this should match taking the mean for mpqnem
  ungroup()

plot(meanbysubj$meanbysubj ~ meanbysubj$mpqnem)
abline(lm(meanbysubj$meanbysubj ~ meanbysubj$mpqnem))
```



(c) What interesting pattern do you notice? Why is it expected? What's the difference between the two plots? Is one better than the other?

The first graph looks at na vs. mpqnem for each performance. Because mpqnem is a subject-level variable that's the same for every musician, that's why we get the stacks. If we instead aggregate ot the musician level and look at mean negative anxiety (averaging across the performances) vs. mpqnem, we have have one observation per musician. For the second graph (aggregating), we see that the average na tends to be larger for individual with more negative emotionality. The second graph is probably better but in many cases will tell the same story about the relationship (see next paragraph). The only worry would be if the number of performances were very different across the performers (and could throw off some of the group comparisons). (We don't mind showing lots of dots for the same individual when exploring in the graphs, we worry about the "inflated sample size" more when carrying out tests of significance/inference.)

The two plots can differ when the number of performances differs among musicians. Suppose, for instance, one vocalist who is very anxious had twice as many performances as anybody else. In that case, plot (a) would show a higher level of anxiety for vocalists compared to other instruments than plot (b). On the other hand, if a keyboardist had only one performance in which she felt no anxiety, that one performance would count just as much in plot (b) as a keyboardist with 20 performances averaged together, so plot (b) would show a lower anxiety level for keyboardists.

Some of the questions we might want to explore: Which characteristics of individual performances are most associated with performance anxiety? Which characteristics of student musicians are most associated with performance anxiety? Are any of these associations statistically significant? Does the significance remain after controlling for other covariates? But of course, need to account for lack of independence in performances by the same musician.

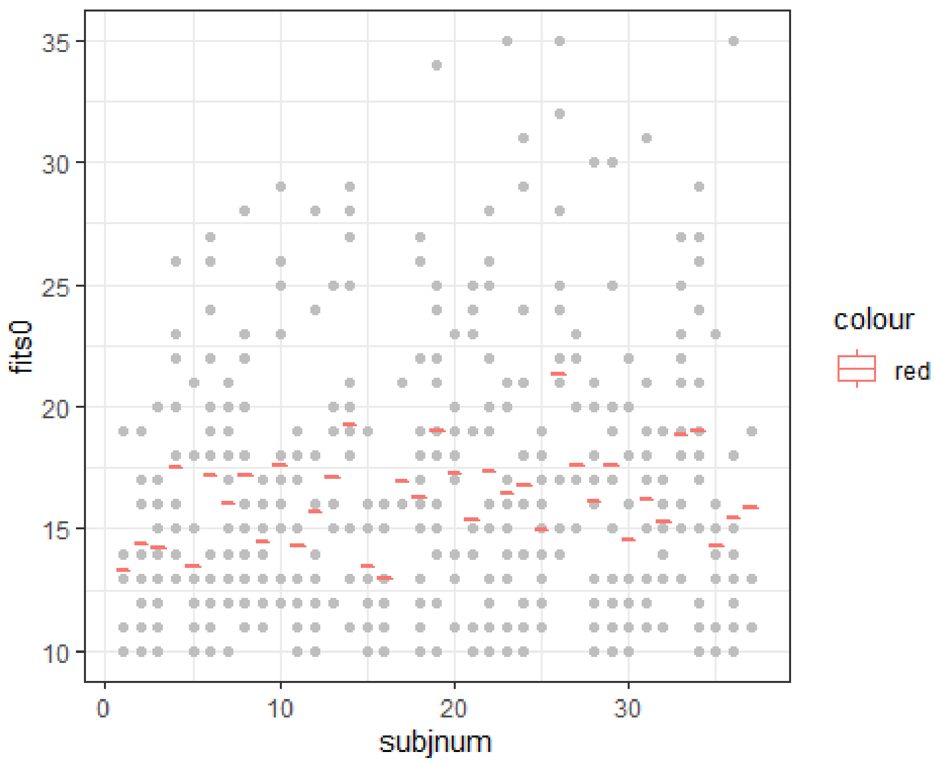
(d) Identify Level 1 and Level 2. Identify some variables at each level.

Level 1: performance (variable: who the performance was in front of)

Lever 2: musician (variables: type of instrument, baseline anxiety measures)

Start with a random intercepts (null) model to assess the variation in performance anxiety (“na”) among the musicians.

```
model0 = lmer(na ~ 1 + (1 | subjnum), data = musicians)
summary(model0)
## Linear mixed model fit by REML ['lmerMod']
## Formula: na ~ 1 + (1 | subjnum)
##    Data: musicians
##
## REML criterion at convergence: 3005.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.9041 -0.6894 -0.2076  0.5284  4.1286
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  subjnum  (Intercept)  4.95    2.225
##  Residual             22.46    4.739
## Number of obs: 497, groups:  subjnum, 37
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  16.2370     0.4279   37.94
fits0 = fitted.values(model0, level = 1)
ggplot(musicians,  aes(y = fits0, x = subjnum, group= factor(subjnum), col="red"))
+
  geom_point(aes(y=na, x= subjnum), col="grey") +
  geom_boxplot() +  #of the one value, to show a line
  theme_bw()
```



(e) What is the ICC?

```
4.95/(4.95 + 22.46)
## [1] 0.180591
```

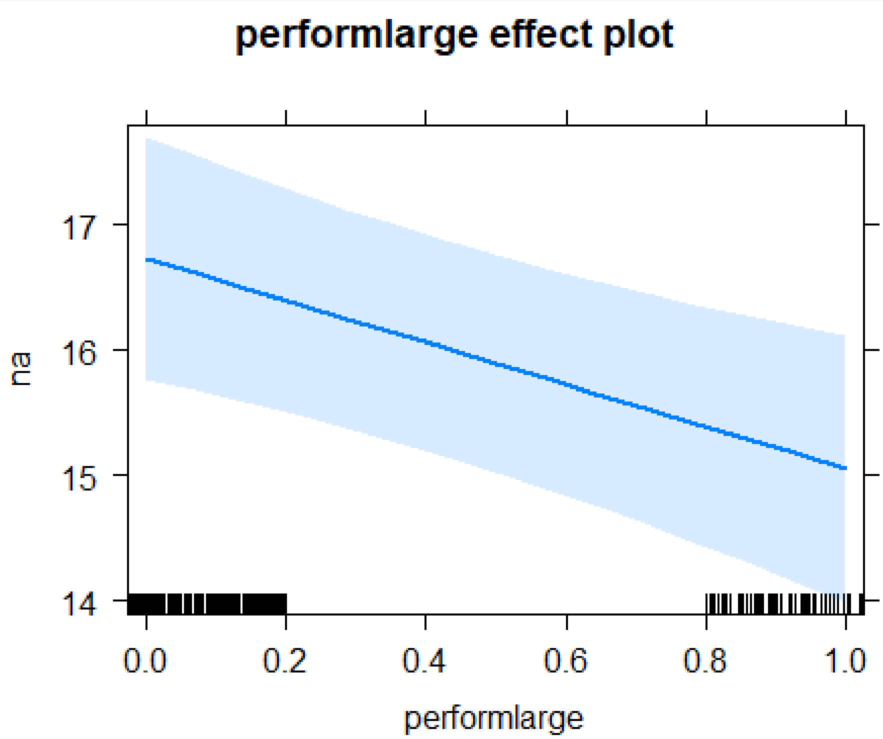
About 18% of the variation in na is among musicians, the rest is across performances within musicians.

So again, what we are doing with a multilevel model is not that different from including “subject” in the model to account for that source of variation (and source of dependence). The multilevel model has two main consequences: shrinkage, and the ability to ask different research questions, like about level 2 variables explaining some of the subject to subject differences.

Suppose we want to predict performance anxiety based on the type of performance (large ensemble or not). Fit a model that looks at the effects of type of performance (large vs. small/solo), allowing this effect to vary by musician.

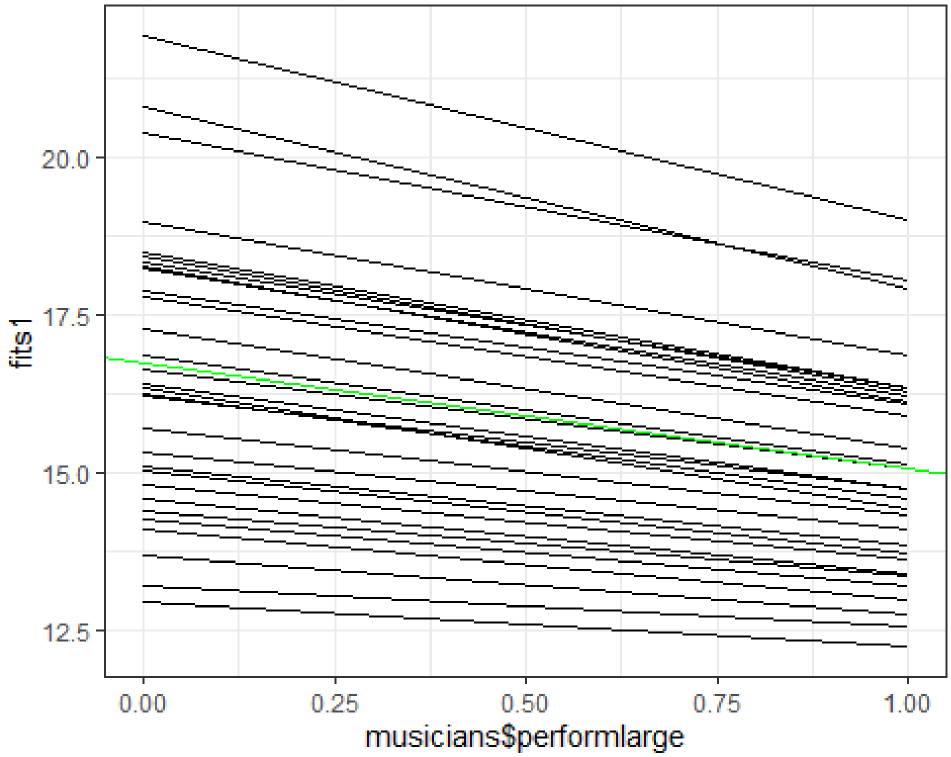
```
#First convert performance type to a binary variable, just to simplify things a bit
initially
musicians$performlarge = as.numeric(musicians$perform_type1 == "LargeEnsemble")
head(musicians$performlarge)
## [1] 0 1 1 0 0 0
model1 = lmer(na ~ 1 + performlarge + (1 + performlarge | subjnum), data = musician
s)
summary(model1, corr= F)
## Linear mixed model fit by REML ['lmerMod']
## Formula: na ~ 1 + performlarge + (1 + performlarge | subjnum)
```

```
##    Data: musicians
##
## REML criterion at convergence: 2994
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.9892 -0.6827 -0.1977  0.4839  4.1398
##
## Random effects:
##  Groups   Name         Variance Std.Dev. Corr
##  subjnum  (Intercept)   6.3330  2.5165
##           performlarge  0.7429  0.8619   -0.76
##  Residual              21.7712  4.6660
## Number of obs: 497, groups:  subjnum, 37
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   16.7297     0.4908   34.09
## performlarge  -1.6762     0.5425   -3.09
#install.packages("effects")
library(effects)
## Loading required package: carData
## lattice theme set by effectsTheme()
## See ?effectsTheme for details.
plot(allEffects(model1))
```




```
fits1 = fitted.values(model1, level =1)
qplot(musicians$performlarge, fits1, group = factor(musicians$subjnum), geom=c("lin
```

```
e")) +
  theme_bw() +
  geom_abline(intercept=fixef(model1)[1], slope = fixef(model1)[2], color="green")
```



*(f) Explain what it means in plain language for this model to have "random intercepts." (*Hint: What does the fixed slope for the 0/1 performance type variable represent?) What does it mean for this model to have "random slopes" for that variable?

The random intercepts pertain to the small and solo performances (performlarge = 0), so they tell us about the musician to musician variability in (average) anxiety for the small and solo performances. (Different musicians have more/less anxieity on average during small and solo performances.) The random slopes represent the change in anxiety between smaller performances and large performances. So the random slopes allow the "effect" of the performance type to vary by musician. (Some musicians may have similar anxiety for both types of performances ⇒ flatter slope, some may have less anxiety with larger performances compared to smaller performances ⇒ more "negative" slope. We probably don't expect any positive slopes, but could be!)

(g) What does "sigma-hat" represent here in this new model? What do the two level 2 variance components represent?

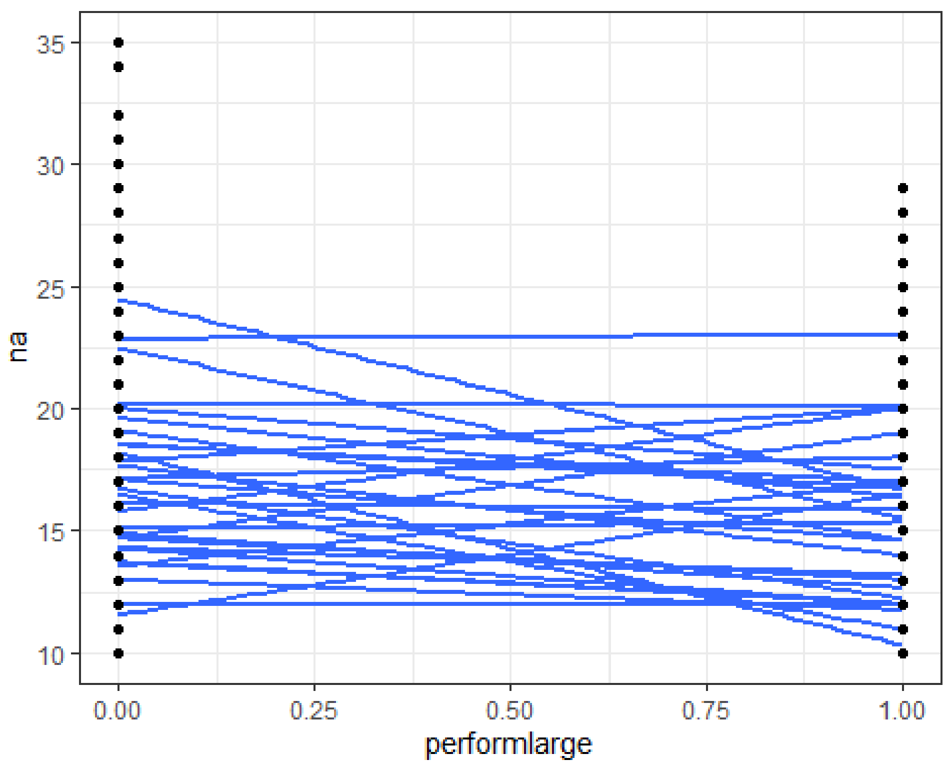
The Level 1 variance is the model-estimated variability within subjects' performances (after adjusting for type of performance), that is the "average" unexplained variation in the na measures of different performances by the same performer. The level 2 variances represent the variation in the intercepts (na for smaller performances) and slopes (change in na between small and large performances) and the covariance between the intercepts and slopes.

(See k) The negative correlation between the intercepts and slopes indicates that subjects with more anxiety with small performances tend to have larger decreases (more negative slopes) in their anxiety with larger performances (corresponding to the "fanning in" we see in the above graph). In other words, those with less anxiety in small performance tend to see less of a difference (slope closer to zero) between large and small performances.

(h) The above graph shows the fitted equations from the multilevel model for each performer. How do you think the graph will differ if we fit a separate line for each performer?

There would very likely be more variability among the lines vs. all going in the same general direction as the model (above graph) won't perfectly describe each individual. In other words, we see shrinkage in the above graph where the lines are encouraged to have the same general pattern and lines from the multilevel model will be a weighted combination of the performer's own (OLS) line and the overall green line averaged across all the performers. Below are the OLS fits.

```
ggplot(musicians) +
  aes(x = performlarge, y = na, group= subjnum) +
  stat_smooth(method = "lm",  se = FALSE) +
  geom_point() +
  theme_bw()
## `geom_smooth()` using formula 'y ~ x'
```



(i) Interpret your model output: Do the signs of the coefficients of the fixed effects make sense in context? What do you learn about the effect of large ensemble performances on anxiety? How much of the performance-to-performance variation is explained by the type of performance? How did the intercept variance change? Does this surprise you?

```
summary(model1)
## Linear mixed model fit by REML ['lmerMod']
## Formula: na ~ 1 + performlarge + (1 + performlarge | subjnum)
##    Data: musicians
##
## REML criterion at convergence: 2994
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.9892 -0.6827 -0.1977  0.4839  4.1398
##
## Random effects:
##  Groups   Name         Variance Std.Dev. Corr
##  subjnum  (Intercept)   6.3330  2.5165
##           performlarge  0.7429  0.8619   -0.76
##  Residual              21.7712  4.6660
## Number of obs: 497, groups:  subjnum, 37
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   16.7297     0.4908   34.09
## performlarge  -1.6762     0.5425   -3.09
##
## Correlation of Fixed Effects:
##             (Intr)
## performlarg -0.453
```

We see a negative coefficient of performlarge, showing a decrease of 1.676 in anxiety (on average across the performers) moving from small/solo performances to large performances (makes sense to me that anxiety would go down for larger performances compared to smaller). The effect appears to be statistically significant ($t = -3.09 < -2.00$). The residual variance (performance to performance within a musician) is now 21.77 from 22.46, a reduction of (22.46 - 21.77)/22.46 = .0307, or about 3%. In other words, the type of performance explains about 3% of the unexplained within-performer variability, which is not a lot.

The intercept variance is now 6.33 (and represents the variability in intercepts for the small performances), up from 4.95, so it has increased (but keep in mind these two “intercept variances” for these two models mean different things). This increase in the variation of the intercepts is a little surprising (but we’ve seen it before when we allow for random slopes) but reflects a hidden relationship between performance type and musician. Some musicians are more likely to have more large performances than other musicians, and it turns out the musicians with a higher proportion of large performances tended to be more anxious than musicians with a lower proportion of large performances. For example, musician 22 has a lot of large performances, but is actually a more anxious musician overall, so after accounting for type of performance, the “residual” for that musician (that musician’s effect) needs to be larger. This happens for several musicians and so the unexplained musician-to-musician variability for small performances increases.

Recall what we learned about how the (Level 2) variance changes with x in a random slopes model: $Var(y_{ij}) = \tau_0^2 + x_{ij}\tau_1^2 + 2x_{ij}\tau_{01}$. For this model, when x = 1 (larger performances) we have 6.333 + 1(.7429) + 21(-.76)(2.5165)(0.8619) = 3.77. When x = 0 (smaller performances), we have 6.333 (variability in na scores is larger for smaller performances than for larger performances). The point is, for large performances, there is less variation among musicians (3.77) than when we didn't account for performance size (4.95). So when we force the slopes (performance-size effect) to be the same for every musician (not random slopes, ignoring performance type) we were getting more of an estimated "average" variability in na across musicians.

(j) Which is larger, the variation in the intercepts or in the slopes? What does that tell you in context?

The variation in the intercepts is larger, so there is more person-to-person variation in negative anxiety (in the small performances) than in the effect of performance size on negative anxiety.

(k) Interpret the slope/intercept correlation in this context. Are the effects "fanning in" or "fanning out"? Why does this relationship make sense in context?

The variance is minimized at (.76)(2.5165)(.8619) ≈ 1.65, so beyond our 0/1 values for performance size. This means the lines are fanning in. Because the slopes are generally negative, the musicians who are more anxious with small performances tend to have a bigger decrease (smaller/more negative slope) in anxiety in moving to large performances. Those who are less anxious with small performances tend to decrease for large performances as well, but don't have as much "room" to drop. This is also consist with the larger variation in na values with small performances, but more consistency in the larger performances.

(l) Write out a (new) model (by level and then composite) that also uses the type of performance (large ensemble or not) with random intercepts and slopes that depend on type of instrument (orchestral or not).

Level 1: $na_{ij} = \beta_{0j} + \beta_{1j} largeperf_{ij} + \epsilon_{ij}$;

Level 2: $\beta_{0j} = \beta_{00} + \beta_{01} instrument_j + u_{0j}$;

$\beta_{1j} = \beta_{10} + \beta_{11} instrument_j + u_{1j}$.

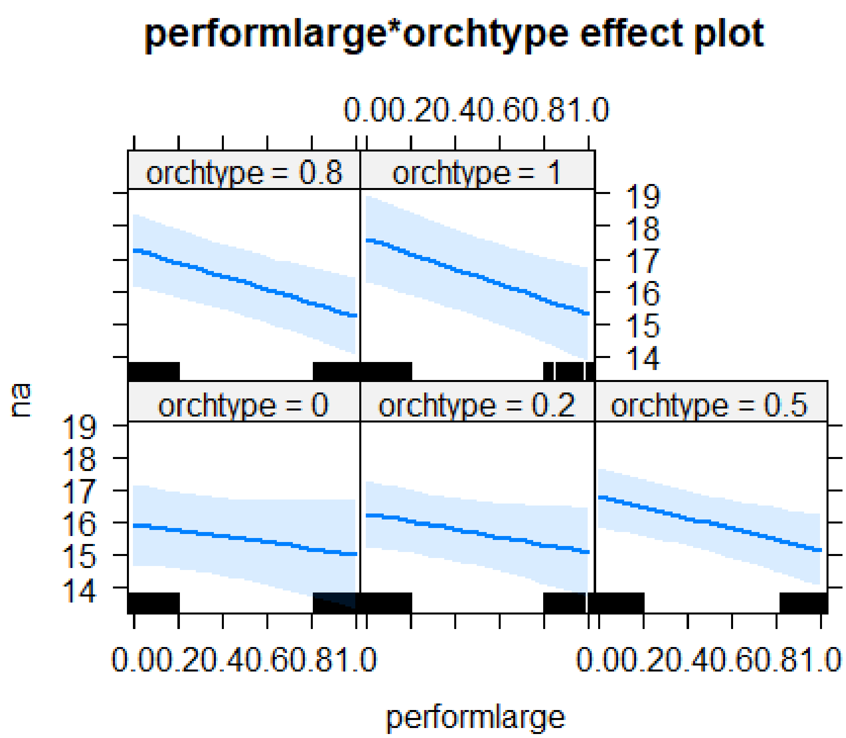
Composite: $na_{ij} = \beta_{00} + \beta_{01} instrument_j + u_{0j} + \beta_{10} largeperf_{ij}$

$+\beta_{11} largeperf_{ij} \times instrument_j + u_{1j} \times largeperf_{ij} + \epsilon_{ij}$;

Fit the model for (l): Make a binary variable for orchestra. Include orchestra, largeperformance, and their interaction as fixed effects, and then random intercepts and random slopes for performance type).

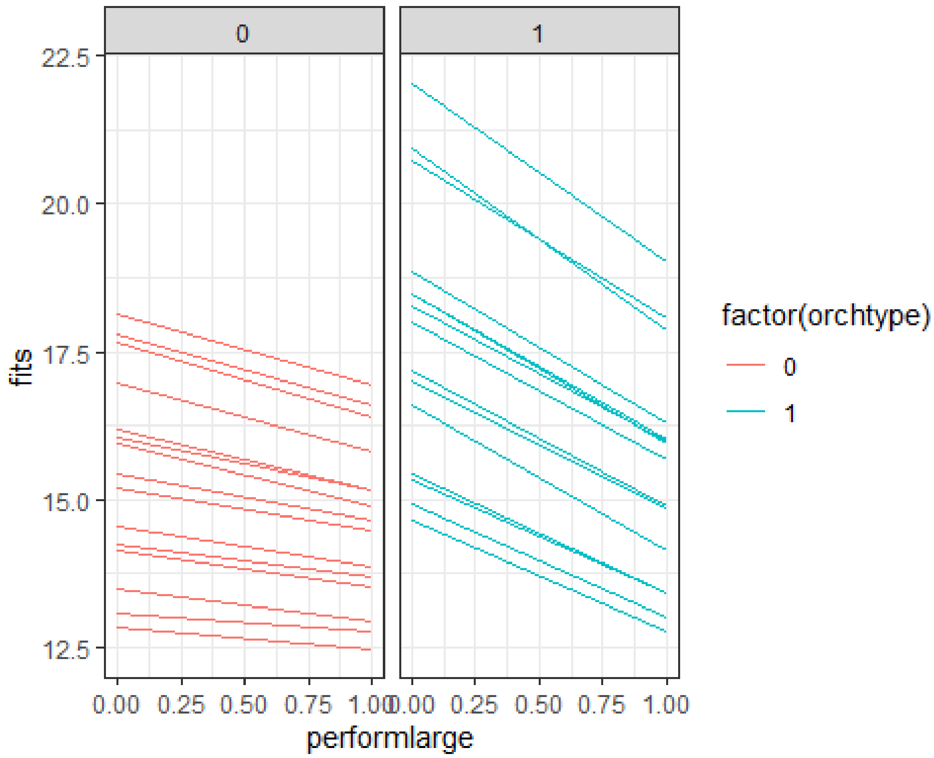
```
musicians$orchtype = ifelse(musicians$instrument1 == "orchestralinstrument", 1,0)
model2 =  lmer(na ~ performlarge*orchtype + (performlarge | subjnum), data = musici
ans)
summary(model2, corr=FALSE)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: na ~ performlarge * orchtype + (performlarge | subjnum)
##    Data: musicians
##
## REML criterion at convergence: 2987
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.9404 -0.6625 -0.1771  0.4796  4.1860
##
## Random effects:
##  Groups   Name         Variance Std.Dev. Corr
##  subjnum  (Intercept)   5.655   2.3781
##           performlarge  0.452   0.6723   -0.63
##  Residual              21.807   4.6698
## Number of obs: 497, groups:  subjnum, 37
##
## Fixed effects:
##                        Estimate Std. Error t value
## (Intercept)             15.9297     0.6415  24.833
## performlarge            -0.9106     0.8452  -1.077
## orchtype                 1.6926     0.9452   1.791
## performlarge:orchtype   -1.4239     1.0992  -1.295
plot(allEffects(model2))  #focus on 0/1 or keep as TRUE/FALSE variable
```

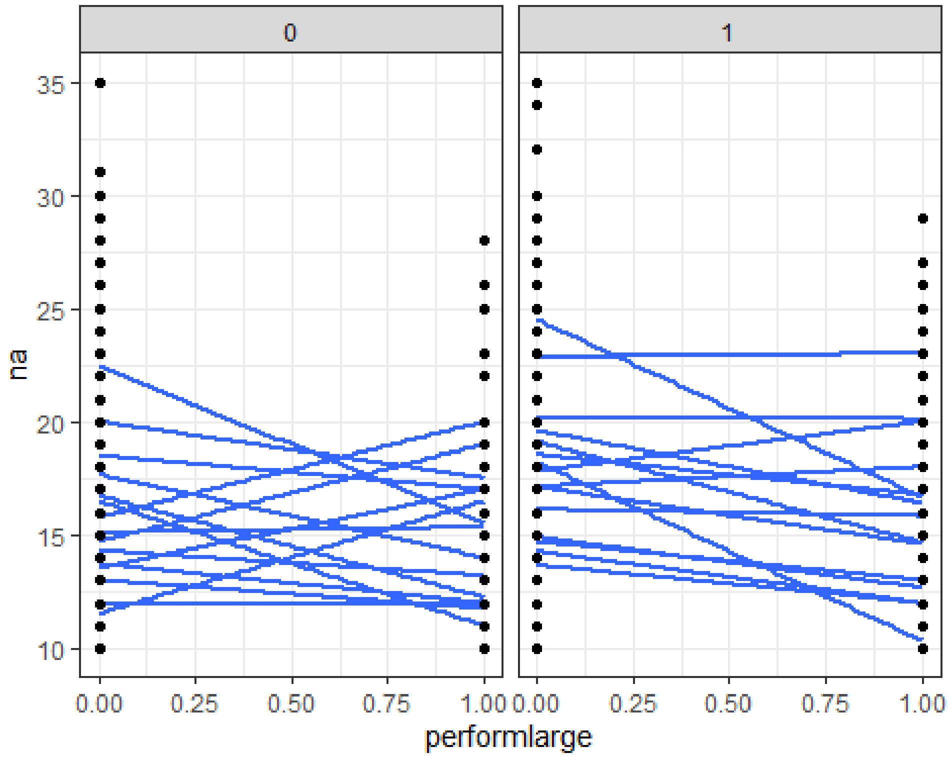



```
#musicians$orchtype = (musicians$instrument1 == "orchestralinstrument")
#model2 =  lmer(na ~ performlarge*orchtype + (performlarge | subjnum), data = music
ians)
```

```
#plot(allEffects(model2))

fits = fitted.values(model2, level =1)
ggplot(musicians, aes(y = fits, x= performlarge, group = factor(subjnum), col=facto
r(orchtype))) +
  facet_wrap(~orchtype) +
  geom_line()+
  theme_bw()
```




```
#"raw data"
ggplot(musicians) +
  aes(x = performlarge, y = na, group= subjnum) +
  stat_smooth(method = "lm",  se = FALSE) +
  facet_wrap(~orchtype) +
  geom_point() +
  theme_bw()
## `geom_smooth()` using formula 'y ~ x'
```



(m) Interpret the interaction between performance type and orchestra type in context.

If the musician has an orchestral instrument (rather than voice or keyboard), then they experience a (1.42) larger decrease (on average) in anxiety moving from small to large performance types than the voice and keyboardists. However, this may not be statistically significant with a t-value of -1.295 > -2.00.

(n) How much variability in the intercepts does including type of instrument explain? How much variability in the slopes?

Now these comparisons make a bit more sense because we are focusing on changes in Level 2 variance from adding a Level 2 variable.

Variability in intercepts is now 5.655, down from 6.33, a decrease of `r{(6.33 - 5.655)/6.33} = about 11%

Variability in slopes is now 0.452, down from 0.7429, a decrease of 'r{(.7429 - .452)/.7429} = about 39%

How did the estimate of within group variation change?

Our residual standard error is now 21.807, about the same as 21.77. We don't expect it to change when we add a Level 2 variable, and the change can be explained by small numerical adjustments as we simultaneously estimate all of these parameters.

(o) Summarize what you learn about the effect of type of instrument on the intercepts and the slopes.

So knowing the type of instrument tells us more about why some musicians see a sharper drop in negative anxiety moving from small to large performance types than others. It explains a bit (but not as much) about why some musicians have more anxiety on average for smaller performances than others.

(p) Maybe with the interaction between performance type and instrument type we no longer need the random slopes. Investigate this. Document how you did so (both the model equations and the R code).

We could remove the random slopes component.

```
model2b =   lmer(na ~ performlarge*orchtype + (1 | subjnum), data = musicians)
anova(model2, model2b)
## refitting model(s) with ML (instead of REML)
## Data: musicians
## Models:
## model2b: na ~ performlarge * orchtype + (1 | subjnum)
## model2: na ~ performlarge * orchtype + (performlarge | subjnum)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## model2b    6 3003.6 3028.9 -1495.8   2991.6
## model2     8 3007.2 3040.8 -1495.6   2991.2 0.4302  2     0.8065
```

We notice AIC and BIC values are better for model 2b, without the random slopes, and the more complicated model is not significantly better in terms of log likelihood (p-value = .8065). Using the same slope for performance type (difference in anxiety between large and small performances) on everyone, differing only by the type of instrument, seems to be as well-fitting of a model as also allowing a different slope for every musician.