Lab 16 - Mixed ANOVA

Nick Sumpter (Edited by Eddie-Williams Owiredu & Guy Twa)

2023-10-16

Today's Lab

This lab will test your ability to perform a two-way mixed ANOVA with a between-subjects variable and a within-subjects (repeated-measures) variable. Mixed ANOVAs test 3 different things:

1. Is there a difference in the mean of an outcome variable between levels of an independent grouping variable?

2. Is there a difference in the mean of an outcome variable between levels of a repeated-measures grouping variable?

3. Is there an interaction between the two (i.e. does the effect of one variable depends on the level of the other)?

This is particularly common in designs assessing the effect of a treatment on two independent groups pre- and post-treatment application. The mixed ANOVA is the final ANOVA type learned in this course.

Loading Packages and Data

```
library(ez)
library(dae)
library(car)
library(pastecs)
library(tidyverse)

theme_set(theme_bw())

setwd("/Users/eddie-williamsowiredu/Desktop/grd770_23/Lab16")

load("Mixed_Factorial_ANOVA.RData")

source("functions.R")
```

Getting to Know Your Data

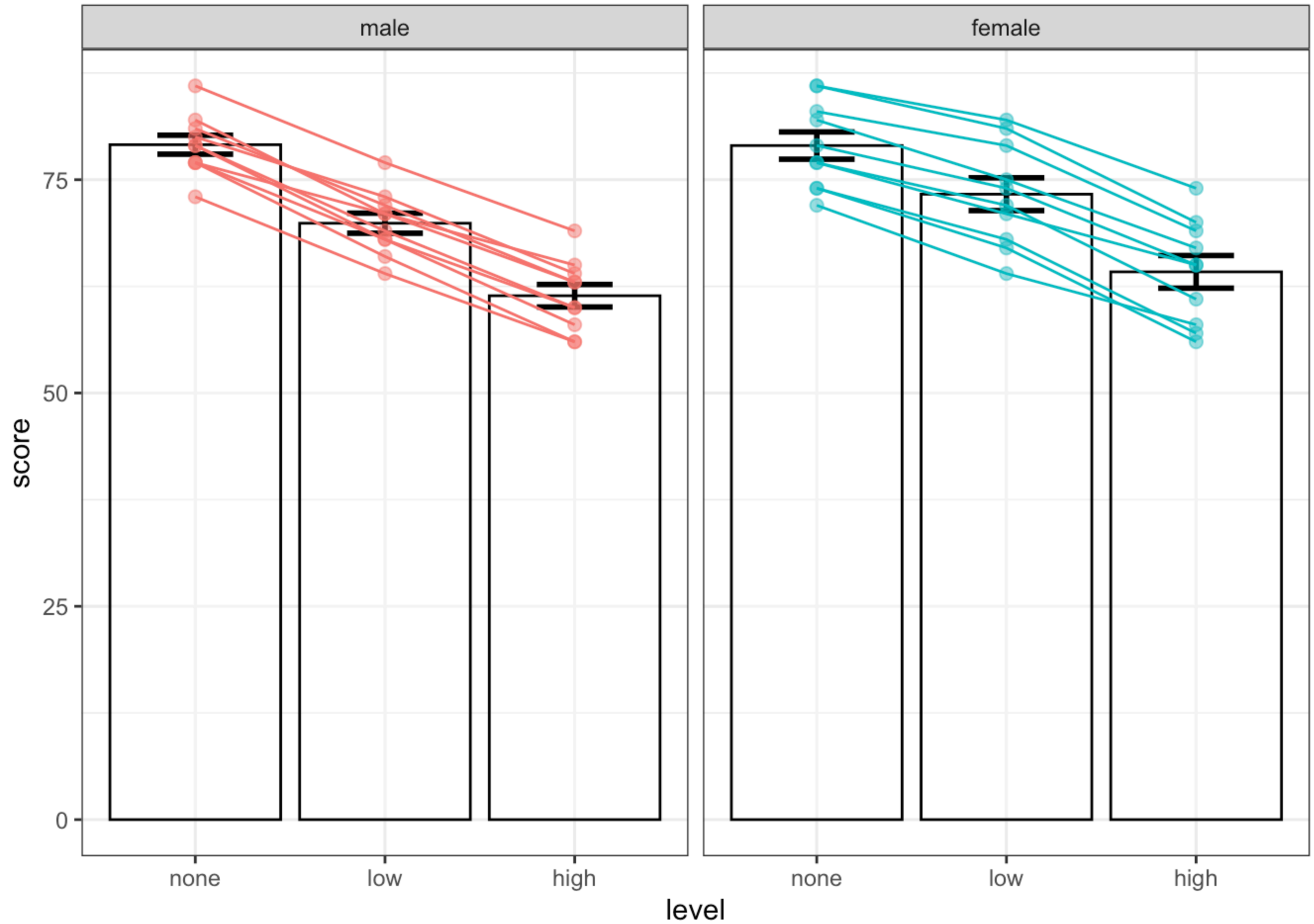
Today's dataset is called `noise` and is very similar to the dataset we used for the one-way repeated measures ANOVA. It has 20 observations and 5 variables:

- `ID` (factor, 20 levels): Subject ID
- `gender` (factor, 2 levels): our between subject's factor
- `none` (numeric): score on a perceptual task with no background noise
- `low` (numeric): score on a perceptual task with low amount of background noise
- `high` (numeric): score on a perceptual task with high amount of background noise

Notice we have dropped a level of the within-subjects group (medium noise level) and added a between subjects factor of gender.

Let's try to visualize the mean value of the outcome at each combination of our grouping variables. First we will need to transform our data into long format then order our noise levels.

```
noise_long <- noise %>%
  pivot_longer(cols = none:high, names_to = "level", values_to = "score") %>%
  mutate(level = factor(level, levels = c("none", "low", "high"), ordered = TRUE))

ggplot(data = noise_long, mapping = aes(x = level, y = score, color = gender)) +
  geom_bar(stat = "summary", fun = mean, fill = "white", color = 'black', alpha = 0.4) +
  stat_summary(fun.data = mean_se, geom = "errorbar", width = 0.4, color = 'black', size
= 1) +
  geom_point(size = 2, alpha = 0.5, show.legend = FALSE) +
  geom_line(mapping = aes(group = ID), show.legend = FALSE) +
  facet_wrap(~ gender)
```



Based on an eyeball test, we can see that there seems to be a main effect of noise level but not of gender. Based on the slope of each line, there may be a slightly stronger effect of noise level in males vs females.

Assumptions of Mixed ANOVA

Assumptions of a mixed ANOVA are a combination of assumptions from between- and within-subjects one-way ANOVAs:

1. Normality of model residuals

   After running the model, you should test normality of the model residuals across all group combinations (male-high, female-low, etc.).

2. Homogeneity of variance across levels of each between-subjects variable

   After running the model, you should test whether the model residuals show homogeneity of variance across all levels of your between-subjects grouping variable.

3. Sphericity across levels of each within-subjects variable

   Refer to the Mauchly's Test for Sphericity result in the `ezANOVA` function to determine whether this assumption is met.

   You will notice that this assumption is tested for both the main effect of your repeated-measures variable and the interaction effect. Basically, any term involving a repeated-measure variable will have sphericity tested. You should treat each term individually, determining whether you need to change the reported degrees of freedom for each one separately.

4. Independence

   When assessing independence in mixed ANOVA, ignore within-subjects factors. As long as the between-subjects groups have completely independent samples, independence is maintained. In our case, we can assume that this assumption is met.

Running a Mixed ANOVA

Using the long format of your data, use the `ezANOVA` function to run a two-way mixed ANOVA. The mixed ANOVA is essentially a combination of all three previous ANOVAs. Basically, name the independent variables in the `between` option and the repeated measures variables in the `within` option. Otherwise, all other inputs should remain the same.

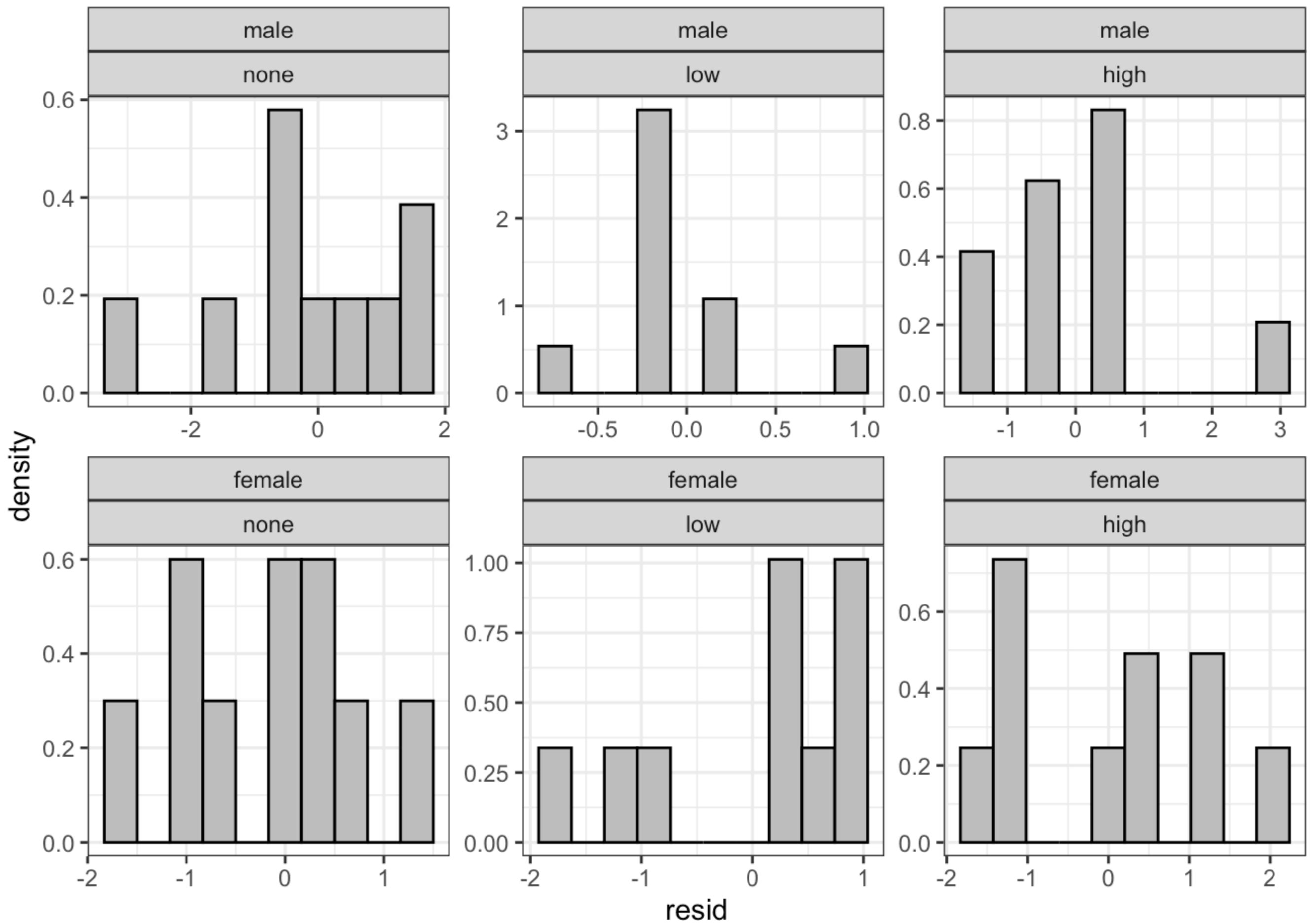
```
mod <- ezANOVA(data = noise_long,
               dv = score,
               within = level,
               between = gender,
               wid = ID,
               type = 3,
               return_aov = TRUE)

mod
```

```
## $ANOVA
##         Effect DFn DFd           F            p p<.05        ges
## 2       gender   1  18   0.9142756 3.516494e-01       0.04607764
## 3        level   2  36 757.0010593 3.867798e-30     * 0.67335776
## 4 gender:level   2  36  10.0201271 3.470909e-04     * 0.02656187
##
## $`Mauchly's Test for Sphericity`
##         Effect        W          p p<.05
## 3        level 0.689824 0.04258685     *
## 4 gender:level 0.689824 0.04258685     *
##
## $`Sphericity Corrections`
##         Effect       GGe        p[GG] p[GG]<.05      HFe        p[HF] p[HF]<.05
## 3        level 0.7632562 1.579728e-23         * 0.819612 4.205652e-25         *
## 4 gender:level 0.7632562 1.257776e-03         * 0.819612 9.246430e-04         *
##
## $aov
##
## Call:
## aov(formula = formula(aov_formula), data = data)
##
## Grand Mean: 71.15
##
## Stratum 1: ID
##
## Terms:
##                    gender Residuals
## Sum of Squares    62.0167 1220.9667
## Deg. of Freedom         1        18
##
## Residual standard error: 8.235987
## Estimated effects are balanced
##
## Stratum 2: ID:level
##
## Terms:
##                     level gender:level Residuals
## Sum of Squares  2646.7000      35.0333   62.9333
## Deg. of Freedom         2            2        36
##
## Residual standard error: 1.322176
## Estimated effects may be unbalanced
```

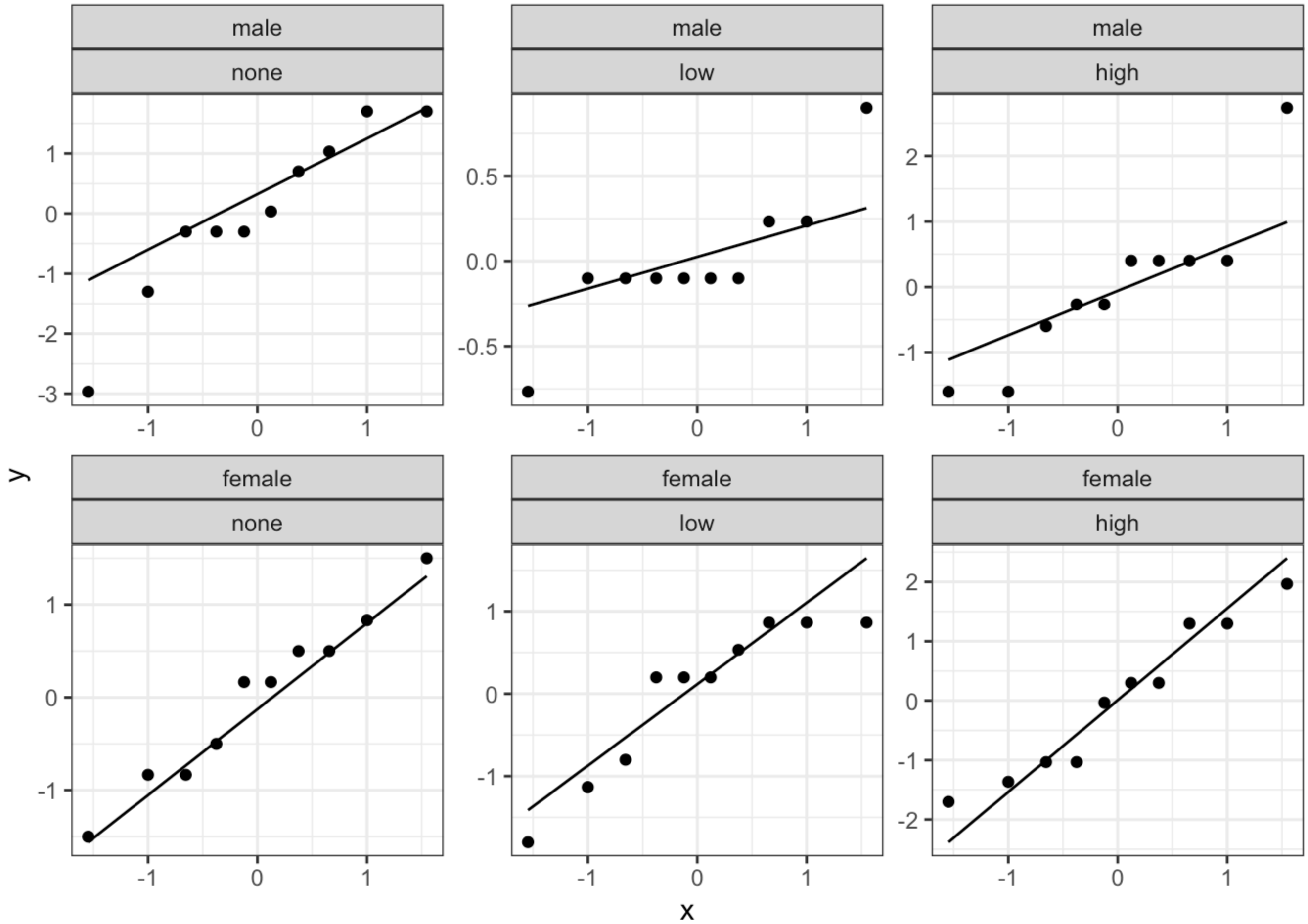
Testing Normality and Homogeneity of Variance

We need to extract the residuals of the model in a similar way to how we did for the repeated-measures ANOVA.

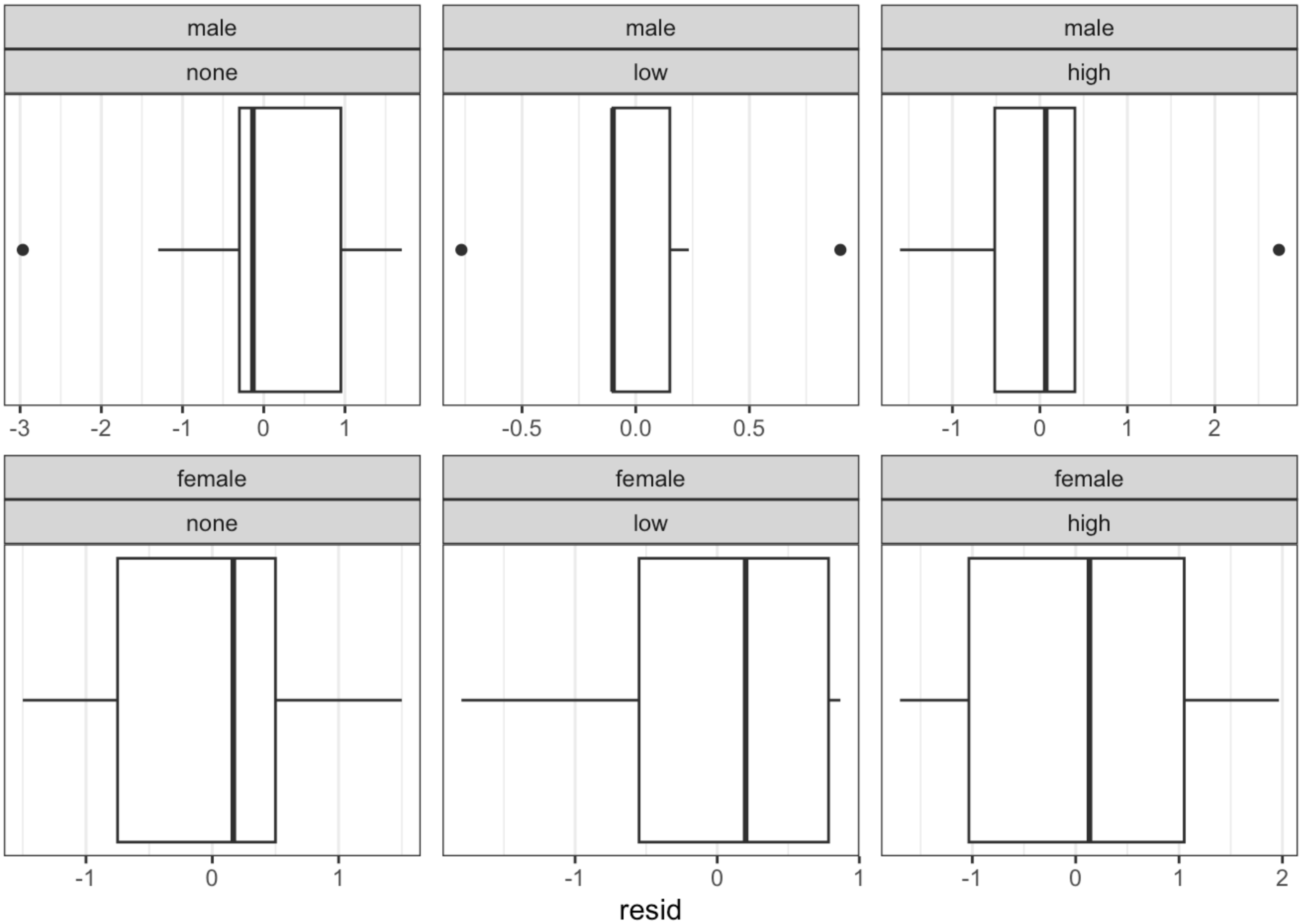
```
# Extracting residuals
residuals <- tibble(group1 = noise_long$gender,
                    group2 = noise_long$level,
                    resid = residuals.aovlist(mod$aov))

# Normality within groups
ggplot(data = residuals, mapping = aes(x = resid)) +
  geom_histogram(mapping = aes(y = ..density..), bins = 10, fill = 'gray', color = 'blac
k') +
  facet_wrap(~ group1 + group2, scales = "free")
```




```
ggplot(data = residuals, mapping = aes(sample = resid)) +
  geom_qq() +
  geom_qq_line() +
  facet_wrap(~ group1 + group2, scales = "free")
```




```
ggplot(data = residuals, mapping = aes(x = resid)) +
  geom_boxplot() +
  theme(axis.ticks.y = element_blank(),
        axis.text.y = element_blank(),
        panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank()) +
  facet_wrap(~ group1 + group2, scales = "free")
```




```
stat.desc.clean(dataset = residuals, variable = resid, group1, group2)
```

```
## # A tibble: 6 × 8
## # Groups:   group1, group2 [6]
##   group1 group2 skewness skew.2SE kurtosis kurt.2SE normtest.W normtest.p
##   <fct>  <ord>     <dbl>    <dbl>    <dbl>    <dbl>      <dbl>      <dbl>
## 1 male   none    -0.600   -0.437   -0.567   -0.212       0.922     0.378
## 2 male   low      0.410    0.298    0.327    0.123       0.826     0.0296
## 3 male   high     0.649    0.472   -0.0295  -0.0111      0.875     0.115
## 4 female none    -0.0597  -0.0435  -1.26    -0.470       0.971     0.896
## 5 female low     -0.700   -0.509   -1.10    -0.413       0.855     0.0665
## 6 female high     0.119    0.0864  -1.63    -0.609       0.934     0.492
```

```
# Homogeneity of variance
leveneTest(resid ~ group1, data = residuals)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  1  0.2411 0.6253
##       58
```

The male-low group seems to exhibit some deviations from normality, but not enough to be concerned. All other groups are close to normality and the homogeneity of variance assumption has been met between genders.

Reporting the Results of a Mixed ANOVA

You will need to report on the assumption of sphericity first for each necessary term, in this case the repeated measures main effect and interaction term. In our case, both the main effect and the interaction term had significant deviations from sphericity. They both had a GGe value of 0.76 and thus the HFe should be used as a correction factor for the degrees of freedom along with the adjusted p-value.

"Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of noise level, $W = 0.690$, $p = 0.043$, and the interaction effect of noise level and gender, $W = 0.690$, $p = 0.043$. The Greenhouse-Geisser epsilon value was above 0.75, therefore degrees of freedom were corrected using Huynh-Feldt sphericity corrections."

Next, you will need to report the different main and interaction effects separately:

"The main effect of noise level on score in a perceptual task was significant, $F(1.64,29.5) = 757.0$, $p = 4.2e-25$. However, the main effect of gender on score in a perceptual task was non-significant, $F(1,18) = 0.914$, $p = 0.352$. There was also a significant interaction of gender and noise level, $F(1.64,29.5) = 10.0$, $p = 9.2e-4$. This suggests that the effect of background noise level on score in a perceptual task depends upon the gender of the individual. Post-hoc tests revealed …"

Independent Practice

For these exercises, we are going to be using the `anxiety` dataset, which you should load from the `anxiety.RData` file on Canvas. We hypothesize that the score on an anxiety test will significantly differ based on the amount of exercise the individual performs (t1 = none, t2 = moderate, t3 = high, repeated in each individual). We also hypothesize that difference in anxiety test score between each exercise level will depend on the group that individual belongs to, indicating an interaction between the two variables. Investigate this using a two-way mixed ANOVA.

1. Get to know your data: describe your variables and their distribution
2. Run the mixed ANOVA
3. Assumptions
   a. Determine whether assumptions are met
   b. If assumptions are not met, describe what you will do to account for this. If possible, modify your model to meet the assumptions
4. Report your findings as you would describe them in the results section