Psy/Educ 6600: Unit 4 Homework

ANOVA - Without Repeated Measures

Your Name

Spring 2018

Contents

PREPARATION

Load Packages

Make sure the packages are **installed** *(Package tab)*

```
library(tidyverse)    # Loads several very helpful 'tidy' packages
library(readxl)       # Read in Excel datasets
library(furniture)    # Nice tables (by our own Tyson Barrett)
library(afex)         # Analysis of Factorial Experiments
library(emmeans)      # Estimated marginal means (Least-squares means)
library(lsmeans)      # Least-Squares Means
library(multcomp)     # Simultaneous Inference in General Parametric Models
```

Other Datasets for Section B's

```
data_wait <- data.frame(child = c(10, 12, 15, 11,  5,  7,  2),
                        woman = c(17, 13, 16, 12,  7,  8,  3),
                        man   = c(20, 25, 14, 17, 12, 18,  7))

data_food <- data.frame(green = c(3, 7, 1, 0, 9, 2),
                        red   = c(3, 4, 5, 6, 4, 6),
                        blue  = c(2, 0, 4, 6, 4, 1))

data_undergrad <- data.frame(class      = c( 1,  1,  1,  1,
                                             2,  2,  2,  2,
                                             3,  3,  3,  3,
                                             4,  4,  4,  4),
                             humanities = c( 2,  4,  3,  7,
                                             3,  4,  6,  5,
                                             7,  8,  7,  7,
                                            10, 12,  9, 13),
                             science    = c( 5,  6,  9, 10,
                                            10, 12, 16, 14,
                                            14, 15, 13, 12,
                                            16, 18, 16, 19),
                             business   = c( 7,  8,  7, 12,
                                            20, 13, 16, 15,
                                            20, 25, 22, 21,
                                            30, 33, 34, 29))

data_memory <- data.frame(incidental_agree     = c(8, 7, 7, 9, 4),
                          incidental_disagree  = c(2, 3, 2, 4, 4),
                          intentional_agree    = c(6, 8, 9, 5, 8),
                          intentional_disagree = c(7, 9, 8, 5, 7))
```

Ihno's Dataset for Section C's

Import Data, Define Factors, and Compute New Variables

- Make sure the **dataset** is saved in the same *folder* as this file
- Make sure the that *folder* is the **working directory**

  NOTE: I added the second line to convert all the variables names to lower case. I still kept the `F` as a capital letter at the end of the five factor variables.

```
ihno_clean <- read_excel("Ihno_dataset.xls") %>%
  dplyr::rename_all(tolower) %>%
  dplyr::mutate(genderF = factor(gender,
                                 levels = c(1, 2),
                                 labels = c("Female",
                                            "Male"))) %>%
  dplyr::mutate(majorF = factor(major,
                                levels = c(1, 2, 3, 4,5),
                                labels = c("Psychology",
                                           "Premed",
                                           "Biology",
                                           "Sociology",
                                           "Economics"))) %>%
  dplyr::mutate(reasonF = factor(reason,
                                 levels = c(1, 2, 3),
                                 labels = c("Program requirement",
                                            "Personal interest",
                                            "Advisor recommendation"))) %>%
  dplyr::mutate(exp_condF = factor(exp_cond,
                                   levels = c(1, 2, 3, 4),
                                   labels = c("Easy",
                                              "Moderate",
                                              "Difficult",
                                              "Impossible"))) %>%
  dplyr::mutate(coffeeF = factor(coffee,
                                 levels = c(0, 1),
                                 labels = c("Not a regular coffee drinker",
                                            "Regularly drinks coffee")))  %>%
  dplyr::mutate(hr_base_bps = hr_base / 60)
```

TUTORIALS

1-Way ANOVA with `afex::aov_4()`

Test for Violations of *Homoegeity of Variance (HOV).*

> **NOTE:** We learned how to do this in chapter 7

```
# Levene's Test of HOV
data_name %>%
  car::leveneTest(continuous_var ~ group_var,
                  data = .,
                  center = "mean")
```

Fitting Models - formuals

The `aov_4()` function from the `afex` package fits ANOVA models (oneway, two-way, repeated measures, and mixed design). It needs at least two arguments:

1. formula: `continuous_var ~ group_var + (1|id_var)` *one observation per subject and **`id_var`** is **distinct** for each subject*
2. dataset: `data = .` *we use the period to signify that the datset is being piped from above*

Here is an outline of what your syntax should look like when you **fit and save a one-way ANOVA**. Of course you will replace the dataset name and the variable names, as well as the name you are saving it as.

> **NOTE:** The `aov_4()` function works on data in LONG format only. Each observation needs to be on its one line or row with seperate variables for the group membership (categorical factor or `fct`) and the continuous measurement (numberic or `dbl`).

```
# One-way ANOVA: fit and save
aov_name <- data_name %>%
  afex::aov_4(continuous_var ~ group_var + (1|id_var),
              data = .)
```

Basic Output - stored name of model

By running the name you saved you model under, you will get a brief set of output, including a measure of **Effect Size**.

> **NOTE:** The `ges` is the *generalized eta squared*. In a one-way ANOVA, the eta-squared effect size is the same value, ie. generalized (η^2_g) and partial η^2_p) are the same.

```
# Display basic ANOVA results (includes effect size)
aov_name
```

Fuller Output - add `$Anova` on model name

To fully fill out a standard ANOVA table and compute other effect sizes, you will need a more complete set of output, including the **Sum of Squares** components, you will need to add `$Anova` at the end of the model name before running it.

> **NOTE:** IGNORE the first line that starts with `(Intercept)`! Also, the 'mean sum of squares' are not included in this table, nor is the **Total** line at the bottom of the standard ANOVA table. You will need to manually compute these values and add them on the homework page. Remember that `Sum of Squares (SS)` and `degrees of freedom (df)` add up, but `Mean Sum of Squreas (MS)` do not add up. Also: `MS = SS/df` for each term.

```
# Display fuller ANOVA results (includes sum of squares)
aov_name$Anova
```

Post Hoc Comparisons with `emmeans::emmeans()`

All Pairwise Comparisons with `pairs()`

There are two steps to conduct all possible pairwise comparisons:

1. `emmeans(~ group_var)` - Calculate the *Estimated Marinal Means*
2. `pairs()` - Determine if each pair is *significantly different*

Within the `pairs()` function there are several options for controling for multiple comparisons, including:

- `adjust = "none"` - Fisher's LSD
- `adjust = "tukey"` - Tukey's HSD
- `adjust = "bon"` - Bonferroni

```
# Pairwise post hoc: Tukey's HSD adjustment for multiple comparisons
aov_name %>%
  emmeans::emmeans(~ group_var) %>%        # Calculate Estimated Marinal Means
  pairs(adjust = "tukey")                  # Is each pair signif different?
```

Contrast Statements with `emmenas::contrast()`

There are two steps to conduct a contrast comparison:

1. `emmeans(~ group_var)` - Calculate the *Estimated Marinal Means*
2. `contrast()` - Determine if each pair is *significantly different*

Inside the contrast statement, list the named sets of linear contrast weights. We will only be doing one-at-a-time, but we must still use a nested `list`.

> **NOTE:** You must provide one weight (c_i) for each of the k groups. If you wish to ignore a group, that group's weight is $c_i = 0$. The sum total of all the weights must be zero ($\sum c_i = 0$), so use positive and negative numbers.

```
# Contrast statement : Impossible vs. Rest
aov_name %>%
  emmeans::emmeans(~ group_var) %>%
  emmeans::contrast(list("your contrast name" = c(c_1, c_2, ... , c_k)))
```

Two-way ANOVA Models with `afex::aov_4()`

Fitting Models -include interaction in formula

The only difference between a one-way and two-way ANOVA's syntax is the inclusion of a second grouping variable in the formula.

> **NOTE:** The astric (`*`) is used to designate the interaction and main effects between two factors. `group_var1*group_var2` is short for `group_var1 + group_var2 + group_var1:group_var2`. The colon (`:`) designates an interaction.

```
# Two-way ANOVA: fit and save
aov_name <- data_name %>%
  afex::aov_4(continuous_var ~ group_var1*group_var2 + (1|id_var),
              data = .)
```

Cell Means: Displaying in a Grid and Plotting

For a two-way ANOVA, we often would like to see a grid of the means for all combinations of the two grouping factors. This is may be achieved by the following steps:

1. `group_by` - group observations by both of the grouping variables
2. `summarise` - compute the mean of each combination subgroup
3. `spread` -spread the means into a grid pattern

```
# Raw data: 2-way table of means (i.e. cell means)
data_long %>%
  dplyr::group_by(group_var1, group_var2) %>%
  dplyr::summarise(mean = mean(continuous_var)) %>%
  tidyr::spread(key = group_var1,
                value = mean)
```

Chapter 12. 1-way ANOVA Models

`data_wait` Reaction Time to Voice Calling for Help

12B-4 1-Way ANOVA - introduction

TEXTBOOK QUESTION: *A social psychologist wants to know how long people will wait before responding to cries for help from an unknown person and whether the gender or age of the person in need of help makes any difference. One at a time, subjects sit in a room waiting to be called for an experiment. After a few minutes they hear cries for help from the next room, which are actually on a tape recording. The cries are in either an adult male's, an adult female's, or a child's voice; seven subjects are randomly assigned to each condition. The dependent variable is the number of seconds from the time the cries begin until the subject gets up to investigate or help. (a) Calculate the F ratio. (b) Find the critical F ($\alpha = .05$). (c) What is your statistical conclusion? (d) Present the results of the ANOVA in a summary table. (e) Calculate η^2 using Formula 12.10.*

```r
# Display the raw dataset: wide format
data_wait
```

```
  child woman man
1    10    17  20
2    12    13  25
3    15    16  14
4    11    12  17
5     5     7  12
6     7     8  18
7     2     3   7
```

First, the data must be restructured from **wide** to **long** format, so that each observation is on its own line. All categorical variables must be declared as factors. We also must add a distinct indicator variable (id number variable).

```r
# convert the dataset: wide --> long
data_wait_long <- data_wait %>% 
  tidyr::gather(key = caller_type,                  # new var name = groups
                value = delay_time,                 # new var name = measurements
                child, woman, man) %>%              # all old variable names
  dplyr::mutate(id = row_number()) %>%              # create a sequential id variable
  dplyr::select(id, caller_type, delay_time) %>%    # reorder the variables
  dplyr::mutate_at(vars(id, caller_type), factor)   # declare factors

data_wait_long %>% head(n = 10)                     # display the top 10 rows only
```

```
   id caller_type delay_time
1   1       child         10
2   2       child         12
3   3       child         15
4   4       child         11
5   5       child          5
6   6       child          7
7   7       child          2
8   8       woman         17
9   9       woman         13
10 10       woman         16
```

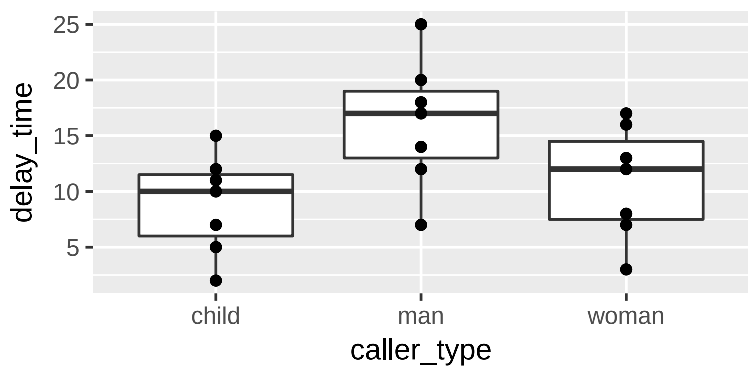
Second, check the summary statistics for each group.

```
# Raw data: summary table
data_wait_long %>%
  dplyr::group_by(caller_type) %>%          # divide into groups
  furniture::table1(delay_time,             # gives M(SD)
                    output = "markdown")     # add chunk option: results="asis"
```

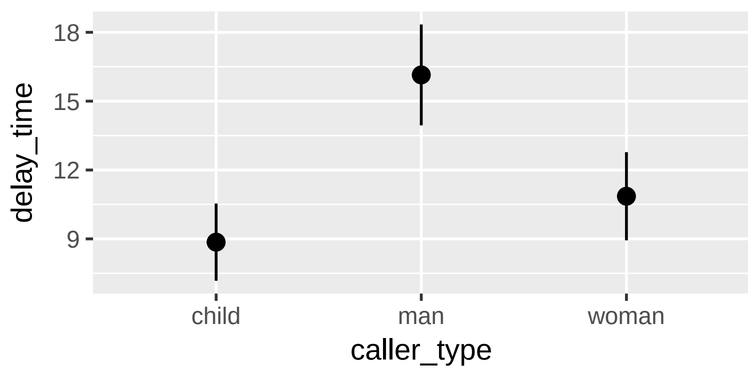
	child	man	woman
	n = 7	n = 7	n = 7
delay_time			
	8.9 (4.5)	16.1 (5.8)	10.9 (5.1)

Third, plot the data to eyeball the potential effect. Remember the center line in each box represents the median, not the mean.

```
# Raw data: boxplots
data_wait_long %>%
  ggplot(aes(x = caller_type,
             y = delay_time)) +
  geom_boxplot() +
  geom_point()
```




```
# Raw data: plot M(SE)
data_wait_long %>%
  ggplot(aes(x = caller_type,
             y = delay_time)) +
  stat_summary()
```



DIRECTIONS: Fit an one-way ANOVA model for the differences in mean `delay_time` for each of the three independent `caller_type` groups with the `afex::aov_4()` function and save the results under the name `aov_wait_time`.

```
# One-way ANOVA: fit and save
```

Remember, since you are saving your model to a name *(aov_wait_time)*, there will not be any output, except a message about setting contrasts to `contr.sum`.

DIRECTIONS: Request the omnibus F value by typing the name you saved your fitted model as above (`aov_wait_time`).

```
# Display basic ANOVA results (includes effect size)
```

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

`data_food` Icing Color Effect on Appetite

12B-5 Another One-Way ANOVA

TEXTBOOK QUESTION: *A psychologist is interested in the relationship between color of food and appetite. To explore this relationship, the researcher bakes small cookies with icing of one of three different colors (green, red, or blue). The researcher offers cookies to subjects while they are performing a boring task. Each subject is run individually under the same conditions, except for the color of the icing on the cookies that are available. Six subjects are randomly assigned to each color. The number of cookies consumed by each subject during the 30-minute session is shown in the following table. (a) Calculate the F ratio. (b) Find the critical F ($\alpha = .01$). (c) What is your statistical decision with respect to the null hypothesis? (d) Present your results in the form of a summary table.*

```r
# Display the raw dataset: wide format
data_food
```

```
  green red blue
1     3   3    2
2     7   4    0
3     1   5    4
4     0   6    6
5     9   4    4
6     2   6    1
```

First, the data must be restructured from **wide** to **long** format, so that each observation is on its own line. All categorical variables must be declared as factors. We also must add a distinct indicator variable.

```r
# convert the dataset: wide --> long
data_food_long <- data_food %>%
  tidyr::gather(key = icing_color,                # new var name = groups
                value = cookies_ate,               # new var name = measurements
                green, red, blue) %>%              # all old variable names
  dplyr::mutate(id = row_number()) %>%             # create a sequential id variable
  dplyr::select(id, icing_color, cookies_ate) %>%  # reorder the variables
  dplyr::mutate_at(vars(id, icing_color), factor)  # declare factors

data_food_long %>% head(n = 10)
```

```
   id icing_color cookies_ate
1   1       green           3
2   2       green           7
3   3       green           1
4   4       green           0
5   5       green           9
6   6       green           2
7   7         red           3
8   8         red           4
9   9         red           5
10 10         red           6
```

DIRECTIONS: Request the summary statistics for each group using the `table1()` function from the `furniture` package, after piping a `dplyr::group_by(group_var)` step.

```
# Raw data: summary table
```

DIRECTIONS: Plot the raw data for each group using the `stat_summary()` layer in `ggplot(aes(x = group_var, y = contin_var))`.

```
# Raw data: plot M(SE)'s
```

DIRECTIONS: Fit an one-way ANOVA model for the difference in mean `cookies_ate` for each of the three independent `icing_color` groups with the `afex::aov_4()` function and save the results under the name `aov_food_time`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

12B-6 The Effect of Larger Mean Values

TEXTBOOK QUESTION: *Suppose that the data in Exercise 5 had turned out differently. In particular, suppose that the number of cookies eaten by subjects in the green condition remains the same, but each subject in the red condition ate 10 more cookies than in the previous data set, and each subject in the blue condition ate 20 more. (a) Calculate the F ratio. Is the new F ratio significant at the .01 level? (b) Which part of the F ratio has changed from the previous exercise and which part has remained the same? (c) Put your results in a summary table to facilitate comparison with the results of Exercise 5. (d) Calculate estimated* ω^2 *with Formula 12.12 and adjusted* η^2 *with Formula 12.14. Are they the same? Explain.*

BEFORE you restructured from **wide** to **long** format, add 10 to the red counts and add 20 to the blue counts.

```
# Revised wide dataset
data_food_long2 <- data_food %>%
  dplyr::mutate(red   = 10 + red) %>%              # NEW VALUES = 10 + OLD !!!
  dplyr::mutate(blue  = 20 + blue) %>%             # NEW VALUES = 20 + OLD !!!
  tidyr::gather(key = icing_color,                 # new var name = groups
                value = cookies_ate,               # new var name = measurements
                green, red, blue) %>%              # all old variable names
  dplyr::mutate(id = row_number()) %>%             # create a sequential id variable
  dplyr::select(id, icing_color, cookies_ate) %>%  # reorder the variables
  dplyr::mutate_at(vars(id, icing_color), factor)  # declare factors

data_food_long2 %>% head(n = 10)
```

```
   id icing_color cookies_ate
1   1       green           3
2   2       green           7
3   3       green           1
4   4       green           0
5   5       green           9
6   6       green           2
7   7         red          13
8   8         red          14
9   9         red          15
10 10         red          16
```

DIRECTIONS: Request the summary statistics for each group using the `table1()` function from the `furniture` package, after piping a `dplyr::group_by(group_var)` step.

```
# Raw data: summary table
```

DIRECTIONS: Plot the raw data for each group using the `stat_summary()` layers in `ggplot(aes(x = group_var, y = contin_var))`.

```
# Raw data: plot M(SE)
```

DIRECTIONS: Fit an one-way ANOVA model for the difference in mean `cookies_ate` for each of the three independent `icing_color` groups with the `afex::aov_4()` function and save the results under the name `aov_food_time2`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

`ihno_clean` Ihno's Dataset

12C-1 Does Post-Quiz Heart Rate Differ by Difficulty Level?

TEXTBOOK QUESTION: *Perform a one-way ANOVA to test whether the different experimental conditions had a significant effect on postquiz heart rate. Request descriptive statistics and an HOV test. Calculate eta squared from your ANOVA output, and present your results in APA style.*

DIRECTIONS: Request the summary statistics for each group using the `table1()` function from the `furniture` package, after piping a `dplyr::group_by(group_var)` step.

```
# Raw Data: summary table
```

DIRECTIONS: Plot the raw data for each group using the `stat_summary()` layer in `ggplot(aes(x = group_var, y = contin_var))`.

```
# Raw data: plot M(SE)
```

DIRECTIONS: Use the `leveneTest()` function from the `car` package to test if the data give any evidence of a violation of *Homoegeity of Variance (HOV)*.

```
# Levene's Test of HOV
```

DIRECTIONS: Fit an one-way ANOVA model for the difference in mean `hr_post` for each of the three independent `exp_condF` groups *(make sure to use the factor version)* with the `afex::aov_4()` function and save the results under the name `aov_hr_post` for future use.

> **NOTE:** The identification variable is called `sub_num` in this dataset, not `id`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

12C-2a Do the Math and Stat Quiz Scores Differ by College Major?

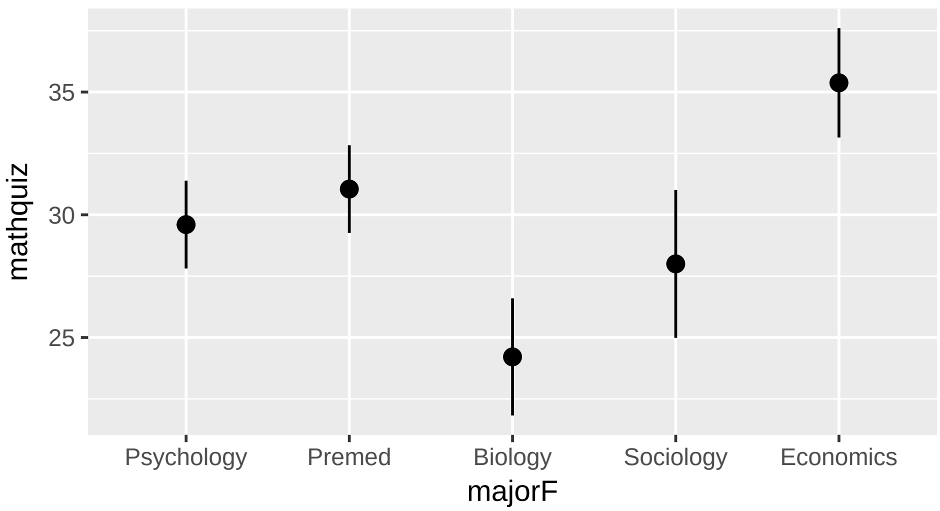
TEXTBOOK QUESTION: *Using college major as the independent variable, perform a one-way ANOVA to test for significant differences in both mathquiz and statquiz . Request descriptive statistics and an HOV test. Based on the HOV test, for which DV should you consider performing an alternative ANOVA test? For whichever DV yields a p value between .05 and .1, report its results as a trend. For whichever DV yields a p value less than .05, calculate the corresponding value of eta squared, and report the ANOVA results, along with the means for the groups, in APA style.*

```
# Raw Data: summary table
ihno_clean %>%
  dplyr::group_by(majorF) %>%
  furniture::table1(mathquiz, statquiz,      # gives M(SD)
                    output = "markdown")     # add chunk option: results="asis")
```

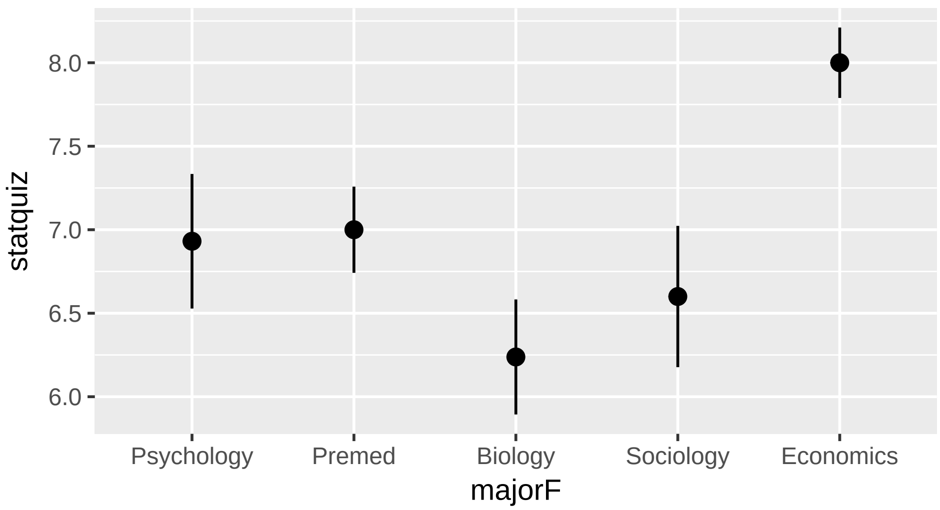
	Psychology	Premed	Biology	Sociology	Economics
	n = 29	n = 25	n = 21	n = 15	n = 10
mathquiz					
	29.6 (8.9)	31.0 (8.2)	24.2 (10.4)	28.0 (10.4)	35.4 (6.3)
statquiz					
	6.9 (2.2)	7.0 (1.3)	6.2 (1.6)	6.6 (1.6)	8.0 (0.7)

DIRECTIONS: Plot the raw data for each group using the `stat_summary()` layer in `ggplot(aes(x = group_var, y = contin_var))`. Do this TWICE, once with `y = mathquiz` and then again with `y = statquiz`.

```
# Raw data: plot M(SE) - Math Quiz
ihno_clean %>%
  ggplot(aes(x = majorF,
             y = mathquiz)) +
  stat_summary()
```




```
# Raw data: plot M(SE) - Stat Quiz
ihno_clean %>%
  ggplot(aes(x = majorF,
             y = statquiz)) +
  stat_summary()
```



Math Quiz - All Five Majors

DIRECTIONS: Use the `car::leveneTest()` to test for violations of HOV.

```r
# Levene's Test of HOV
```

DIRECTIONS: Fit an one-way ANOVA model using `afex::aov_4()`.

> **NOTE:** Because some of the students are missing the `mathquiz` variable, you will need to preceed the `aov_4()` step with `dplyr::filter(complete.cases(mathquiz, majorF))` in the pipeline.

```r
# One-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```r
# Display basic ANOVA results (includes effect size)
```

Stat Quiz - All Five Majors

DIRECTIONS: Use the `car::leveneTest()` to test for violations of HOV.

```
# Levene's Test of HOV
```

DIRECTIONS: Fit an one-way ANOVA model using `afex::aov_4()`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

12C-3 Remove Two Majors and Repeat

Math Quiz - Only Three Majors

TEXTBOOK QUESTION: *Repeat Exercise 2 after using Select Cases to eliminate all of the psychology and premed students.*

> **NOTE:** You will need to preceed Levene's Test with `dplyr::filter(majorF %in% c("Biology", "Sociology", "Economics"))` in the pipeline in order to subset the data.

DIRECTIONS: Use the `car::leveneTest()` to test for violations of *HOV*.

```
# Levene's Test of HOV
```

DIRECTIONS: Fit an one-way ANOVA model using `afex::aov_4()`.

> **NOTE:** Here you will need both the filter step for subsetting majors and the filter step to restrict to complete cases. The order of the two `dplyr::filter()` steps does not matter.

```
# One-way ANOVA: fit and display
```

Stat Quiz - Only Three Majors

DIRECTIONS: Use the `car::leveneTest()` to test for violations of HOV.

> **NOTE:** You will need to preceed Levene's Test and the ANOVA with `dplyr::filter(majorF %in% c("Biology", "Sociology", "Economics"))` in the pipeline in order to subset the data.

```
# Levene's Test of HOV
```

DIRECTIONS: Fit an one-way ANOVA model using `afex::aov_4()`.

```
# One-way ANOVA: fit and display
```

12C-5 Phobia Group vs. Difference (Pre-Post) Heart Rate

TEXTBOOK QUESTION: *Use Recode to create a grouping variable from phobia , such that Group 1 contains those with phobia ratings of 0, 1, or 2; Group 2 = 3 or 4; and Group 3 = 5 or more (you might call the new variable Phob_group). Then use Transform to create another new variable, hr_diff , that equals hr_pre minus hr_base . Perform a one-way ANOVA on hr_diff using Phob_group as the factor. Request descriptive statistics. Report the results in APA style, including the means of the three groups. Explain what this ANOVA demonstrates, in terms of the variables involved.*

```
data_new <- ihno_clean %>%
  dplyr::mutate(phob_group = case_when(phobia <3           ~ 1,
                                       phobia %in% c(3, 4)  ~ 2,
                                       phobia >= 5          ~ 3)) %>%
  dplyr::mutate(phob_group = factor(phob_group,
                                    levels = c(1, 2, 3),
                                    labels = c("Low", "Moderate", "High"))) %>%
  dplyr::mutate(hr_diff = hr_pre - hr_base)
```

DIRECTIONS: Request the summary statistics for each group using the `table1()` function from the `furniture` package, after piping a `dplyr::group_by(group_var)` step.

```
# Raw data: summary table
```

DIRECTIONS: Use the `car::leveneTest()` to test for violations of *HOV*.

```
# Levene's Test of HOV
```

DIRECTIONS: Fit and save a one-way ANOVA model using `afex::aov_4()`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

Chapter 13. Multiple Comparisons

`ihno_clean` Ihno's Dataset

13C-1a One-Way ANOVA: LSD and Tukey as Post Hoc tests

TEXTBOOK QUESTION: *(A) Redo the one-way ANOVA requested in exercise #1 in Section C of the previous chapter, selecting both LSD and Tukey as Post Hoc tests. For postquiz heart rate, which pairs of experimental conditions differ significantly from each other, according to each test? Can you justify using the results of the LSD test?*

DIRECTIONS: Using the ANOVA model saved as `aov_hr_post` previously, request all pair wise post hoc comparisons, by first piping `emmeans::emmeans(~ group_var)` followed by `pairs(adjust = "none")` to utilize Fisher's LSD correction for multiple comparisons.

```
# Pairwise post hoc: Fisher's LSD adjustment for multiple comparisons
```

DIRECTIONS: Repeat the above, but use `pairs(adjust = "tukey")` to utilize Tukey's HSD correction for multiple comparisons.

```
# Pairwise post hoc: Tukey's HSD adjustment for multiple comparisons
```

13C-1c Contrast: Impossible vs. Others

TEXTBOOK QUESTION: *(C) Perform a contrast to compare the "impossible" condition with the other three for postquiz heart rate. How does the significance of this contrast compare to the one-way ANOVA? Explain. Looking at the means for the four conditions, design a contrast that you think would capture a large proportion of between-group variance.*

DIRECTIONS: Using the sample recipe code chunk in the tutorial section, perform a contrast to compare the "impossible" condition with the other three for postquiz heart rate.

```
# Contrast statement : Impossible vs. Rest
```

13C-2a Post Hoc Pairwise: Tukey and Bonferroni

TEXTBOOK QUESTION: *Redo the one-way ANOVA requested in Section C, exercise 2 of the previous chapter just for the mathquiz variable, selecting both Tukey and Bonferroni as Post Hoc tests in each case. Why is it problematic to use HSD with major as the factor in this dataset? Given the results of the post hoc tests, does the Tukey or Bonferroni test seem to have greater power when testing all possible pairs of means?*

DIRECTIONS: Fit an one-way ANOVA model for the difference in mean `mathquiz` for each `majorF` and save the results under the name `aov_math_major`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request all pairwise post hoc comparisons TWICE, once via Tukey's HSD with the `adjust = "tukey"` option and a second time with `adjust = "bon"` within the `pairs()` function, applied after piping a `emmeans(~ group_var)` step to the ANOVA model.

```
# Pairwise post hoc: Tukey's HSD adjustment for multiple comparisons
```

```
# Pairwise post hoc: Bonferroni adjustment for multiple comparisons
```

13C-2b Contrast: (Biology and Sociology) vs. other three majors

TEXTBOOK QUESTION: *Redo the one-way ANOVA requested in Section C, exercise 2 of the previous chapter just for the statquiz variable and request a contrast that compares the average of the Biology and Sociology majors to the average of the other three majors. Would this contrast be significant if it had been planned? Would this contrast be significant according to Scheffe's test?*

DIRECTIONS: Fit an one-way ANOVA model using `afex::aov_4()` and add via pipes both `emmeans::emmeans(~ group_var)` and `contrast()` with appropriate weights.

```
# Contrast statement: Bio and Soc vs. rest
```

13C-4a One-Way ANOVA: prequiz anxiety by Phobia Group - LSD and Bonferroni

TEXTBOOK QUESTION: *Perform a one-way ANOVA on the prequiz anxiety measurement (anx_pre) using the grouping variable you created in Section C, exercise 5 of the previous chapter (based on phobia ratings). Select both LSD and Bonferroni as your post hoc tests. Which pairs differ significantly for each test?*

DIRECTIONS: Fit an one-way ANOVA model for the difference in mean `anx_pre` for each `phob_group` and save the results under the name `aov_anx_phob`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request all pairwise post hoc comparisons TWICE, once via Fisher's LSD with the `adjust = "none"` option and a second time with `adjust = "bon"` within the `pairs()` function, applied after piping a `emmeans(~ group_var)` step to the ANOVA model.

```
# Pairwise post hoc: Fisher's LSD adjustment for multiple comparisons
```

```
# Pairwise post hoc: Bonferroni adjustment for multiple comparisons
```

13C-4b Contrast: Students (low or moderate) phobia vs. high

TEXTBOOK QUESTION: *Perform a contrast that compares students who had reported low or moderate phobia with those reporting high phobia. Calculate the effect size for this contrast. Is it small, medium, or large?*

DIRECTIONS: Starting with the previously fitted `aov_anx_phob` ANOVA model, add via pipes both `emmeans::emmeans(~ group_var)` and `contrast()` with appropriate weights.

```
# Contrast statement: high vs. rest
```

Chapter 14. Two-Way ANOVA Models

`data_undergrad` Effect of on Class & Major on Employment Experctations

14B-7a 3x4 Two ANOVA

TEXTBOOK QUESTION: *A college is conducting a study of its students' expectations of employment upon graduation. Students are sampled by class and major area of study and are given a score from 0 to 35 according to their responses to a questionnaire concerning their job preparedness, goal orientation, and so forth. The data appear in the following table. (a) Perform a two-way ANOVA and create a summary table.*

```
# convert the dataset: wide --> long
data_undergrad_long <- data_undergrad %>%
  dplyr::mutate(class = factor(class,
                               levels = c(1, 2, 3, 4),
                               labels = c("Freshmen",
                                          "Sophomores",
                                          "Juniors",
                                          "Seniors"))) %>%
  tidyr::gather(key = major,
                value = expect_employ,
                humanities, science, business) %>%
  dplyr::mutate(id = row_number()) %>%
  dplyr::select(id, class, major, expect_employ)

data_undergrad_long %>% head(n = 12)
```

```
   id      class      major expect_employ
1   1   Freshmen humanities             2
2   2   Freshmen humanities             4
3   3   Freshmen humanities             3
4   4   Freshmen humanities             7
5   5 Sophomores humanities             3
6   6 Sophomores humanities             4
7   7 Sophomores humanities             6
8   8 Sophomores humanities             5
9   9    Juniors humanities             7
10 10    Juniors humanities             8
11 11    Juniors humanities             7
12 12    Juniors humanities             7
```

DIRECTIONS: Fit a two-way ANOVA model for the difference in mean `expect_employ` for each of the combinations between the four-level `class` factor and three-level `major` factor with the `afex::aov_4()` function and save the results under the name `aov_employ`.

```
# Two-way ANOVA: fit and save
```

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

14B-7b Plot Cell Means

TEXTBOOK QUESTION: *(B) Draw a graph of the cell means. Does the interaction obscure the interpretation of the main effects?*

Here is the 3x4 grid of cell means, giving the average for each of the 12 combinations of `class` and `major`.

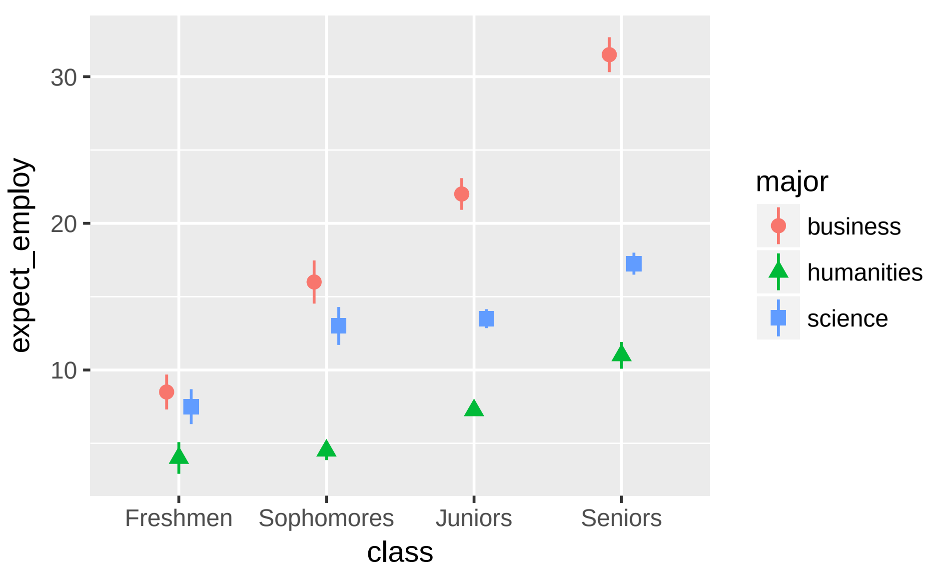
```
# Raw data: 2-way table of means (i.e. cell means)
data_undergrad_long %>%
  dplyr::group_by(class, major) %>%
  dplyr::summarise(mean = mean(expect_employ)) %>%
  tidyr::spread(key = class,
                value = mean)
```

```
# A tibble: 3 x 5
  major      Freshmen Sophomores Juniors Seniors
* <chr>         <dbl>      <dbl>   <dbl>   <dbl>
1 business       8.50       16.0    22.0    31.5
2 humanities     4.00        4.50    7.25   11.0
3 science        7.50       13.0    13.5    17.2
```

To incorporated a second grouping variable into the plot, we can use `shape` and/or `color`. I prefer to use both to ensure that the color distinction is not completely lost when photo coping or if a reader is color blind.

> **NOTE:** The inclusion of `> **NOTE:**` within the `stat_summary()` function causes the points to be slightly offset so that points are not drawn on top of each other.

```
# Raw data: plot M(SE)
data_undergrad_long %>%
  ggplot(aes(x = class,
             y = expect_employ,
             shape = major,
             color = major)) +
  stat_summary(position=position_dodge(width=0.25))
```



14B-7c Pairwise Post Hoc with Tukey's HSD

TEXTBOOK QUESTION: *(C) Use Tukey's HSD to determine which pairs of class years differ significantly.*

DIRECTIONS: Request the summary statistics for `expect_employ` within each `class` using the `table1()` function from the `furniture` package, after piping a `dplyr::group_by(group_var)` step.

DIRECTIONS: Plot the raw data for each `class` using the `stat_summary()` layer in `ggplot(aes(x = group_var, y = contin_var))`.

```
# Raw data: plot M(SE)
```

DIRECTIONS: Request all pairwise post hoc comparisons via Tukey's HSD with the `adjust = "tukey"` option in the `pairs()` function, applied after piping a `emmeans(~ group_var)` step to the ANOVA model.

```
# Pairwise post hoc: Tukey's HSD adjustment for multiple comparisons
```

14B-7d 2x2 Contrast Statements to Test Extremes

TEXTBOOK QUESTION: *For just the freshmen and seniors, calculate the three possible interaction contrasts. Which, if any, would be significant according to Scheffe's test?*

The following code chunk will display the means for all combinations of the two grouping variables so that you can figure out which order to set up the contrast weights (c_i's).

```
# Request all emmeans: see ORDER for contrast weights to be entered below
# aov_employ %>%
#   emmeans::emmeans(~ class*major)
```

```
# 2x2 Contrast statement (Freshmen vs. Seniors): Humanities vs. Science
# aov_employ %>%
#   emmeans::emmeans(~ class*major) %>%
#   emmeans::contrast(list("fr-sr X Hum-Sc" = c( 0, 0, 0,  0,
#                                                1, 0, 0, -1,
#                                               -1, 0, 0,  1)))
```

```
# 2x2 Contrast statement (Freshmen vs. Seniors): Humanities vs. Business
# aov_employ %>%
#   emmeans::emmeans(~ class*major) %>%
#   emmeans::contrast(list("fr-sr X Hum-bus" = c(1, 0, 0, -1,
#                                               -1, 0, 0,  1,
#                                                0, 0, 0,  0)))
```

```
# 2x2 Contrast statement (Freshmen vs. Seniors): Science vs. Business
# aov_employ %>%
#   emmeans::emmeans(~ class*major) %>%
#   emmeans::contrast(list("fr-sr X Hum-bus" = c(1, 0, 0, -1,
#                                                0, 0, 0,  0,
#                                               -1, 0, 0,  1)))
```

data_memory Effect of Warning and Attitude on Memory

14B-8a 2x2 Two-Way ANOVA

TEXTBOOK QUESTION: *The data from Exercise 12B8 for a four group experiment on attitudes and memory are reproduced below. Considering the relationships among the four experimental conditions, it should be obvious that it makes sense to analyze these data with a two-way ANOVA. (A) Perform a two-way ANOVA and create a summary table of your results. (Note : You can use the summary table from Exercise 12B8 as the basis for a new table.)*

```r
# convert the dataset: wide --> long
data_memory_long <- data_memory %>%
  tidyr::gather(key = warning_attitude,
                value = recall) %>%
  tidyr::separate(warning_attitude,
                  into = c("warning", "attitude"),
                  remove = FALSE) %>%
  dplyr::mutate(id = row_number()) %>%
  dplyr::mutate_at(vars(id, warning_attitude, warning, attitude), factor) %>%
  dplyr::select(id, warning_attitude, warning, attitude, recall)

data_memory_long
```

```
   id     warning_attitude     warning attitude recall
1   1     incidental_agree  incidental    agree      8
2   2     incidental_agree  incidental    agree      7
3   3     incidental_agree  incidental    agree      7
4   4     incidental_agree  incidental    agree      9
5   5     incidental_agree  incidental    agree      4
6   6  incidental_disagree  incidental disagree      2
7   7  incidental_disagree  incidental disagree      3
8   8  incidental_disagree  incidental disagree      2
9   9  incidental_disagree  incidental disagree      4
10 10  incidental_disagree  incidental disagree      4
11 11    intentional_agree intentional    agree      6
12 12    intentional_agree intentional    agree      8
13 13    intentional_agree intentional    agree      9
14 14    intentional_agree intentional    agree      5
15 15    intentional_agree intentional    agree      8
16 16 intentional_disagree intentional disagree      7
17 17 intentional_disagree intentional disagree      9
18 18 intentional_disagree intentional disagree      8
19 19 intentional_disagree intentional disagree      5
20 20 intentional_disagree intentional disagree      7
```

DIRECTIONS: Fit a two-way ANOVA model for the difference in mean `recall` for each of the combinations between the two-level `warning` factor and two-level `attitude` factor with the `afex::aov_4()` function and save the results under the name `aov_memory_2way`.

```
# Two-way ANOVA: fit and save
```

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

14B-8b One-Way ANOVA: one 4 level factor

TEXTBOOK QUESTION: *(B) Compare your summary table to the one you produced for Exercise 12B8.*

> **NOTE:** We did not do Exercise 12B8, but we can do it here.

DIRECTIONS: Fit an one-way ANOVA model for the difference in mean `recall` for each of the four independent `warning_attitude` groups with the `afex::aov_4()` function and save the results under the name `aov_memory_1way`.

```
# One-way ANOVA: fit and save
```

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

14B-8c Plot Means to Aid Interpretation

TEXTBOOK QUESTION: * (C) What conclusions can you draw from the two-way ANOVA?*

```
data_memory_long %>%
  dplyr::group_by(warning_attitude) %>%
  furniture::table1(recall,                    # gives M(SD)
                    output = "markdown")       # add chunk option: results="asis"
```

	incidental_agree	incidental_disagree	intentional_agree	intentional_disagree
	n = 5	n = 5	n = 5	n = 5
recall				
	7.0 (1.9)	3.0 (1.0)	7.2 (1.6)	7.2 (1.5)

`ihno_clean` Ihno's Dataset

14C-1a 5x2 ANOVA: Major and Gender on Math Quiz

TEXTBOOK QUESTION: *Using college major and gender as your independent variables, perform a two-way ANOVA on mathquiz . Request descriptive statistics and an HOV test. Calculate the ordinary eta squared for each factor, and report your results in APA style.*

DIRECTIONS: Fit a two-way ANOVA model for the difference in mean `mathquiz` for each of the combinations between the five-level `majorF` factor and two-level `genderF` factor with the `afex::aov_4()` function and save the results under the name `aov_math_2way`.

```
# Two-way ANOVA: fit and save
```

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

14C-1b Follow-up Comparisons: by major only

TEXTBOOK QUESTION: *Given the ANOVA results, perform an appropriate follow-up test. Explain your results in terms of the descriptive statistics.*

DIRECTIONS: Request the summary statistics for each group using the `table1()` function from the `furniture` package, after piping a `dplyr::group_by(group_var)` step.

```
# Raw data: summary table
```

DIRECTIONS: Fit an one-way ANOVA model using `afex::aov_4()`. Add on via pipes both `emmeans::emmeans(~ group_var)` and `pairs()`. Make sure to indicate `adjust = "tukey"` wintin the pairs command.

```
# One-way ANOVA: fit and pairwise with Tukey's HSD
```

14C-4a 2x3 ANOVA: Phobia Group and Gender on Math Quiz

TEXTBOOK QUESTION: *Using the phobia grouping variable you created for computer exercise 5 in Chapter 12 and gender as your IVs, perform a two-way ANOVA on mathquiz . Request the appropriate post hoc test and a plot of the cell means, and report the results in APA style.*

DIRECTIONS: Plot the raw data for each group using the `stat_summary()` layer in `ggplot(aes(x = group_var1, y = contin_var))`. Utilize the `shape` and `color` options for `group_var2`. Also consider dodging the position of the groups to avoid overplotting.

```r
# Raw data: plot M(SE)
```

Here is the 2x3 grid of cell means, giving the average for each of the 6 combinations of `genderF` and `phob_group`.

```r
# Raw data: 2-way table of means (i.e. cell means)
data_new %>%
  dplyr::group_by(phob_group, genderF) %>%
  dplyr::summarise(mean = mean(mathquiz, na.rm = TRUE)) %>%
  tidyr::spread(key = phob_group,
                value = mean)
```

```
# A tibble: 2 x 4
  genderF   Low Moderate  High
* <fct>   <dbl>    <dbl> <dbl>
1 Female   34.4     27.9  27.1
2 Male     31.1     23.6  26.4
```

DIRECTIONS: Fit a two-way ANOVA model for the difference in mean `mathquiz` for each of the combinations between the five-level `phob_group` factor and two-level `genderF` factor with the `afex::aov_4()` function and save the results under the name `aov_math_phob_gender`.

```
# Two-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

DIRECTIONS: Request all pairwise post hoc comparisons via Fisher's LSD with the `adjust = "none"` option in the `pairs()` function, applied after piping a `emmeans(~ group_var)` step to the ANOVA model above.

```
# Pairwise post hoc: Fisher's LSD adjustment for multiple comparisons
```

14C-4b Repeat without the Moderate Group

TEXTBOOK QUESTION: *Repeat part a (except for the post hoc test) after deleting the moderate phobia group from the analysis. What type of interaction do you see in the plot? Test the simple main effect of phobia for each gender. Do you need to follow up any of the simple main effects with pairwise comparisons? Explain.*

DIRECTIONS: Repeat the previous ANOVA model, but preceed it by a `dplyr::filter(phob_group != "Moderate")` step in the pipeline and save the results under the name `aov_math_phob2_gender`.

```
# Two-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

14C-5a 2x3 ANOVA: Coffee Drinking and Phobia Group on Post Quiz Heart Rate

TEXTBOOK QUESTION: *Using the phobia grouping variable you created for computer exercise #5 in Chapter 12 (do not drop any phobia groups for this exercise) and coffee (regular coffee drinker or not) as your IVs, perform a two-way ANOVA on the postquiz heart rate. Request an HOV test, observed power, and a plot of the cell means. (A) Does the HOV test give you cause for concern? Explain the ANOVA results in terms of the plot you created.*

Here is the 2x3 grid of cell means, giving the average for each of the 6 combinations of `coffeeF` and `phob_group`.

```
# Raw data: 2-way table of means (i.e. cell means)
data_new %>%
  dplyr::group_by(coffeeF, phob_group) %>%
  dplyr::summarise(mean = mean(hr_post)) %>%
  tidyr::spread(key = phob_group,
                value = mean)
```

```
# A tibble: 2 x 4
# Groups:   coffeeF [2]
  coffeeF                       Low Moderate  High
* <fct>                       <dbl>    <dbl> <dbl>
1 Not a regular coffee drinker  69.7     72.8  73.6
2 Regularly drinks coffee       71.8     75.8  75.5
```

DIRECTIONS: Plot the raw data for each group using the `stat_summary()` layer in `ggplot(aes(x = group_var1, y = contin_var))`. Utilize the `shape` and `color` options for `group_var2`. Also consider dodging the position of the groups to avoid overplotting.

```
# Raw data: plot M(SE)'s
```

DIRECTIONS: Use the `car::leveneTest()` to test for violations of *HOV*. Since this is a two-way ANOVA situation, be sure to include the correct formula: `contin_var = group_var1*group_var2`.

```
# Levene's Test of HOV
```

DIRECTIONS: Fit a two-way ANOVA model for the difference in mean `hr_post` for each of the combinations between the five-level `phob_group` factor and two-level `coffeeF` factor with the `afex::aov_4()` function and save the results under the name `aov_hrpost_phob_coffee`.

```
# Two-way ANOVA: fit and save
```

DIRECTIONS: Request the F value by typing the name you saved your fitted model as above.

```
# Display basic ANOVA results (includes effect size)
```

14C-5b Follow-up Comparisons

TEXTBOOK QUESTION: *Request an appropriate post hoc test to follow-up your ANOVA results, and report the results. Calculate the ordinary eta squared for each main effect; how large is each effect?* ~~*Does the observed power make sense in each case?*~~

DIRECTIONS: Request the more complete summary table by adding `$Anova` at the end of the name you saved your fitted model as above.

```
# Display fuller ANOVA results (includes sum of squares)
```

DIRECTIONS: Request all pairwise post hoc comparisons with the `pairs()` function, applied after piping a `emmeans(~ group_var)` step to the ANOVA model. Only do this for significant main effects with at least three factor levels.

Do NOT worry about observed power!