**Module 11 Lab Activity: Core Concepts in Analysis of Variance**

**PSY 652 Research Methods**

**Nov 13, 2019**

**Description of the dataset:**

The StressReduction.csv dataset includes the stress levels of 120 subjects who were randomly assigned to either no treatment or one of three stress reduction treatments (referred to as Programs). The variables are:

* Program = a text variable representing which of the four stress reduction treatments a subject was assigned to (None, Healthy Habits, Get Movin, and Recharge).
* ProgramCode = assigned a numeric value for program (0 = None, 1 = Healthy Habits, 2 = Get Movin, 3 = Recharge).
* StressLevel = reported stress level for each participant on a scale of 1-10, in which higher values represent higher stress levels.

1. Download the “StressReduction.csv” dataset from the Module 10 Lab dropbox folder and save it into a project folder.
2. Create a new R notebook from your project file and name it “StressReduction\_notebook”
3. Create a new R chunk with a first level header: “Load Libraries”
   1. Load the psych & tidyverse packages
4. Create a new R chunk with a first level header: “Import Data”
   1. Read in the “StressReduction.csv” dataset and assign it to an object named “stress”
5. Create a new R chunk with a first level header: “Factor the grouping variable”
   1. Copy and paste the following code into a new R chunk and click run:

stress <- stress %>% mutate(ProgramCode.f = factor(ProgramCode, levels = c(0,1,2,3),labels = c("None", "Healthy Habits", "Get Movin", "Recharge")))

This code uses the mutate function to create a factor version of the ProgramCode variable, which tells R to read this variable as a factor (i.e., categorical) variable. Confirm that this worked by opening the dataframe in your global environment.

1. Create a new R chunk with a first level header: “Get descriptives”
   1. Use any method to calculate descriptive statistics for the StressLevel variable (the mean you calculate in this step will represent the Grand Mean of StressLevel)
   2. Use the aggregate function to calculate the mean StressLevel for participants in each of the four treatment groups (i.e., the group means)
   3. In the white space, write 1-2 sentences interpreting how the group means differ from other groups and from the grand mean.
2. Create a new R chunk with a first level header: “Create boxplots of StressLevel across treatment groups”
   1. Use the ggplot function to create boxplots for StressLevel across the four treatment groups. All four boxplots should be *in the same plot*.

Hint: ggplot(*dataframe*, aes(y = *outcome\_variable*, group = *predictor\_variable*, color = *predictor\_variable*)) + geom\_boxplot()

* 1. In the white space, write 1-2 sentences explaining what this plot shows.

1. Create a new R chunk with a first level header: “Conduct an ANOVA”
   1. Use the lm function to build a model in which StressLevel is regressed on ProgramCode.f.
   2. Use the anova function to display anova output for this model
   3. Calculate the η2 for the model (hint: see the lecture slides for this formula)
   4. In the white space, answer the following questions:
      1. Interpret the F statistic. Is there a significant difference between group means for program level?
      2. Identify the dferr and dfhyp in this model
      3. Interpret the model η2
2. Create a new R chunk with a first level header: “Conduct planned contrasts”
   1. In a new R chunk with a second level header: “Create contrasts.” In this code chunk, create three contrasts with pairwise comparisons between StressLevel in the no treatment versus three treatment groups: None vs. Healthy Habits, None vs. Get Movin, and None vs. Recharge. Save each of these three contrasts to a new object.
   2. In a new R chunk with a second level header: “Bind contrasts to predictor variable”
      1. Use the contrasts and cbind functions to attach the three contrasts you just created to the predictor variable (i.e., ProgramCode.f). View this updated variable to make sure that your weights were assigned correctly (Hint: scroll to the bottom of this output to view weights).

Hint:

# attach the contrasts to your variable:

contrasts(*dataframe*$*variablename*) <- cbind(*contrast1*,*contrast2*, *contrast3*)

# view the variable to ensure that the contrasts were properly attached:

*dataframe$variablename*

1. Create a new R chunk with a second level header: “Run ANOVA with contrasts”
   1. Build an anova model in which StressLevel is regressed on ProgramCode.f (to which the contrasts have been attached) and get model output using the summary.lm function

Hint: summary.lm(*modelname*))

1. In the white space, answer the following questions:
   1. What does the estimate for the intercept represent (hint, you calculated this before!)
   2. Interpret each of the three contrasts. Is the mean StressLevel for any of the treatment groups significantly different from the “no treatment” group?
   3. What is the difference between a planned contrast and a post-hoc analysis? When would you use one method over the other?