Free Assignment 7: Between-subject ANOVA

When completing this assignment, be sure to create a script file with all the commands you type into R and turn that in with the assignment.

In this assignment, we are going to cover interactions using the wages dataset we used in a previous dataset. As a reminder, here are the variables:

- ID: person ID
- WAGE: wage (dollars per hour)
- OCCUPATION: occupation(1=Management, 2=Sales, 3=Clerical , 4=Service, 5=Professional, 6=Other)
- SECTOR: sector of employment(0= other, 1=Manufacturing, 2=Construction)
- UNION: Union membership (1=yes, 0=no)
- EDUCATION: Years of education (12 = high school diploma, 16= completed college, etc.)
- EXPERIENCE: Years of work experience
- AGE: Age in years
- SEX: Sex (0 – male, 1 – female)
- MARR: Married (0 – no, 1 – yes)
- RACE: Race (0 – other, 1 – white, 2 – Hispanic)
- SOUTH:Southern region (1 – yes, 0 – no)

Step 1: The data are in the file "IA5WageData.csv". Load the data into R as the dataframe "wages". Also, load the `tidyverse` and `psych` packages into R.

In this study, we are going to investigate whether sex and marital status predict wages. If we remember Interactive Assignment 5, we examined the wage gap in earnings, that men make more than women. So we expect to find a main effect of sex in our data. However, we'll also investigate how marital status plays a role.

First, we have to alter our data. Since the factors in our data are given as numbers, we need to change them to factor variables. To do this for sex and race, we do the following code.

Step 2: Type the following code into R in order to change SEX and MARR from numeric variables to factor variables.

```
wages$SEX = as.factor(wages$SEX)
wages$MARR = as.factor(wages$MARR)
```

Step 3: Use the describeBy package to fill out the following table for wages to examine the main effects and the interactions for sex and marriage as factors predicting wage. You should have two tables for the main effects and one table for the interaction.

Step 4: Use ggplot to create a plot looking at the variables. Use SEX as your "x" variable and MARR as your "fill" variable. Feel free to change the labels in order to make the plot look better.

Based on these means and the plot, what main effects and/or interactions do you expect?

Step 5: Use the `aov()` command to examine whether there is an interaction between SEX and MARR. Are there any significant main effects? Interactions? Please type the F and P values in your annotations, using APA formatting.

Step 6: In your annotations, answer this question: Using your answers for questions 1 and 2, what do you think these data mean? What is your interpretation about the effects of sex and marriage on WAGE?

Step 7: For the rest of this lab, I want you to investigate the following question. A researcher has the idea that the wage gap between men and women will be higher in the South than the rest of the country. Can you test this hypothesis using ANOVA? Go through the steps we did for the rest of this lab and test this hypothesis.

First, you need to turn SOUTH into a factor variable, and then do the rest of the steps.

What results did you get? Do you think there is any evidence for the researcher's hypothesis?

Step 8: Knit your document and submit the knitted document.