

Due on Canvas on or before Aug 30

Submit a single document with the entire response in it

1. Install or access your software of choice (I'll assume R). Maybe also an IDE.
2. Read the GSS data into R
3. Run my cleaning script
4. Plan and conduct two different t-tests. For each one:
 - a. Identify a variable of interest for comparison and a grouping variable (use two completely different variables in the second test, no repetition). You can use any variables at all – ones I have already cleaned with my code or others – except JobSat and Female.
 - i. If it's not already cleaned, be sure to clean the variable appropriately.
 - ii. You may need to convert your grouping variable to a dichotomous form. Any reasonable criterion for doing so is fine.
 - b. Before doing the test, make a prediction. Write a paragraph that explains what result you think you will find and why you expect it. Think of this as a one-paragraph version of the theory section of a research article
 - c. Formally state the null hypothesis and the alternative hypothesis
 - d. Conduct the test. Include your R code and output for the t-test, confidence intervals, and effect size
 - e. Interpret the results, explaining what you found, and how you know. Refer both to p-values and to confidence intervals, plus the effect size
 - f. Explain whether or not your prediction was supported, and what you conclude from that support. This is the one-paragraph version of the discussion section in a research article
5. Use R's help or a web search to teach yourself how to get a correlation matrix from R, and how to assign p-values to those correlations. (Important note here: the value in this exercise is the search and the figuring out. Asking a classmate how to do it defeats the purpose. We want you to get to the point where you can think things like *"Huh . . . I'm going to need to use the Huber-White sandwich here to adjust for heteroskedasticity in responses. Let's see how that works in R."* Because no matter what the issue is, R can do it; you just need to find the way.)
6. Create a "Table 1" (Most empirical research articles have a table, often the first table, that reports the mean, SD, and correlations of variables in the study. Look through recent issues of AMJ or JAP if you are not yet familiar with this idiom.)
 - a. Choose four relevant variables from the data
 - b. Get the correlations among them, and associated p values, from R
 - c. Make a table in your word processor from these results, modeling the format you see in top journals
 - d. In one paragraph, interpret the results. What can you learn from this table? (In this instance, it's sufficient to rely solely on p-values -- though if you can add and use confidence intervals as well, that would be useful for the future)
 - e. Append your R code and the original R output