Please cite as: Leventhal, B. (2018). Introduction to SPSS using Simulation via T tests, regression, and ANOVA. James Madison University.

PLEASE TYPE YOUR ANSWERS DIRECTLY INTO THIS WORD DOCUMENT TO SUBMIT.

Write-Up 1

An inferential statistics instructor is interested in testing the effects of three different instructional methods on statistics self-efficacy. The three instructional methods are:

1. online only course
2. lecture based course
3. hybrid course

The hybrid course consists of learning the material through an online mechanism and then meeting for weekly discussion and activities related to the material. The instructor collected the data with the help of a seasoned researcher. Graduate students at James Madison University registered in inferential statistics courses were asked if they would like to volunteer for the study. Students who volunteered were randomly assigned to one of three instructional methods.

The instructional methods were carefully developed to provide equivalent content with the only difference being the presentation of the material. The instructor was specifically interested in the students' statistical self-efficacy. To measure current statistics self-efficacy the researcher proposed using the CSSE. The CSSE was developed as a measure of students' *current* self-efficacy for performing basic statistics tasks (Finney & Schraw, 2003). Students responded to the 14 items with a Likert-type, 1-6 response scale where 1 represented "no confidence at all" to perform the task and 6 represented "complete confidence" to perform the task. Scores have the potential to range from 14 to 84 with higher scores indicating a higher degree of current statistics self-efficacy. The instructor assessed self-efficacy three times during the semester – the first day of class (PreS), the middle of the semester (MidS), and at the end of the semester (PostS).

Please use your assigned dataset. You will be using this dataset throughout the rest of the semester so please refer to your assigned dataset only.

Part 1: Organizing the data

Note: There will be nothing to submit for Part 1. This is used for formatting purposes. You will need to do this part in order to successfully complete Parts 2 and 3 as well as subsequent assignments.

1. Open the dataset using SPSS and click on Variable View.

2. You will note that none of the variables are formatted. Here we are going to change various properties of the variables in the dataset. Please make the following changes:

Variable Name	Label	Values	Measure
ID			
Age			Scale
Program		1 = "Online" 2 = "Lecture" 3 = "Hybrid"	Nominal
Gender		1 = "Male" 2 = "Female"	Nominal
PreS	Self-Efficacy Pre-Score		Scale
MidS	Self-Efficacy Mid-Score		Scale
PostS	Self-Efficacy Post-Score		Scale

Part 2: Describe the demographic characteristics of our sample.

1. What is the sample size?
2. We have information about the participant's *gender*. Provide the frequency and percentage for each *gender*.
3. Each participant provided their age.
   a. Compute, report, and interpret three measures of central tendency for the *age* variable.
   b. When reporting the mean, be sure to always report the standard deviation. What is the standard deviation of *age* for this group of participants? Interpret the standard deviation in everyday language.
   c. Provide a histogram and box-plot for the *age* variable. Does *age* appear to be normally distributed? How did you determine your answer? [Consider using the Rule of thumb for Skewness and Kurtosis.] Are there outliers? How did you determine your answer?
4. Why did we compute and report the frequency/percentage for gender, whereas we reported the mean/standard deviation for *age*?
5. Write up your findings as a brief "participants" section of an APA-style paper. (Note: If you are unfamiliar with the parts of an APA-style paper, consult the *Publication Manual of the American Psychological Association, 6th Edition.*) Include an APA-style table. Note: This will be used on future assignments.

Part 3: Report descriptive statistics for our variables.

1. Compute a new variable called *gain* that is the difference between Self-efficacy post-score and self-efficacy pre-score score (i.e. *gain = PostS-PreS*). **Note: There is nothing to submit for this question.**

2. Compute and interpret descriptive statistics (sample size, mean, standard deviation, skew, kurtosis, minimum and maximum score) for the *gain* variable by instructional program. Report the findings in an APA-style table. (Hint: you should have three means, three standard deviations, etc. -mean *gain* score for the online group, mean *gain* score for the lecture group, and mean *gain* score for the hybrid group).
3. Include a frequency histogram of the *gain* score for each group.
4. Include a side-by-side box-plot for *gain* score for each group.
5. Are there any outliers of *gain* score for any group that you examined? How did you determine this?

<u>SAVE this dataset overwriting the current dataset! This will be the file you work with for the remainder of the semester.</u>

<u>Write-Up 2</u>

The instructor has two initial researcher questions.

1) Does students' statistics self-efficacy change from the beginning to the end of the semester? In other words, ignoring all other variables (i.e. instructional method, gender, etc.), does statistics self-efficacy change from the statistics self-efficacy pre-score to the statistics self-efficacy post-score?
2) Is the change in statistics self-efficacy different for males and females? (Hint: you need to use the *gain* score variable created in the previous assignment, i.e., *gain*= postS-preS.)

Use SPSS for analysis to provide an APA-style write-up of the findings. Please include the following:

a. Start with a paragraph to introduce the study.
b. Provide a paragraph describing the participants (from Write-up 1).
c. In the third paragraph, present your research questions and alpha level.
d. Produce a table with descriptive statistics (Minimum, Maximum, Mean, SD, n) of scores on pre- and post-test for males and females. Based on these descriptive statistics, write a paragraph to describe the data to address your research questions.
e. Perform appropriate t-tests in SPSS regarding your research questions. Report your test result in two paragraphs, one for each research question (remember to test the assumption of homogeneity of variance when needed and report your result accordingly). State whether the assumptions for each test were met. If an assumption is violated for an independent samples t-test, interpret the correct results (for other t-tests, just state the violation and continue normally).
f. Write a conclusion paragraph summarizing the results.

<u>Write-Up 3</u>

The instructor has two more research questions:

1) Is there a significant correlation between self-efficacy pre-score and self-efficacy post-score, self-efficacy mid-score and self-efficacy post-score, and self-efficacy pre-score and self-efficacy mid-score?

2) What is the regression model to predict statistics post- statistics self-efficacy from pre-statistics self-efficacy?

Use SPSS for analysis to provide an APA-style write-up of the findings. Please include the following:

a. Start with a paragraph to introduce the study.
b. Provide a paragraph describing the participants (Write-up 1).
c. In the third paragraph, present your research questions and alpha level.
d. Investigate the correlation between self-efficacy pre-score and self-efficacy post-score, self-efficacy mid-score and self-efficacy post-score, and self-efficacy pre-score and self-efficacy mid-score. Provide a scatterplot for each pair and describe what information each scatterplot provides. Provide a correlation matrix in APA style with significant correlations flagged. Write a paragraph to summarize if the correlations are significant or not.
e. Use a simple linear regression model to predict post- statistics self-efficacy from pre-statistics self-efficacy. Report and interpret the estimates of the slope and intercept of the population regression line. Report the proportion of the observed variation in the post-scores that can be attributed to the pre- scores.

Write-Up 4

The instructor begins with one research question:

3) Is there significant mean difference of self-efficacy Post scores among the three instructional methods?

Use SPSS for analysis to provide an APA-style write-up of the findings. Please include the following:

1. Start with a paragraph to introduce the study.
2. Provide a paragraph describing the participants.
3. In the third paragraph, present your research hypotheses and alpha level.
4. Examine, compare, and present the descriptive statistics of students' post statistics self-efficacy by groups (using a table and a plot such as side-by-side boxplot). Make sure to include the group size, mean, and standard deviation. Speculate what the study might reveal based on the descriptive statistics and plot.
5. Perform a one-way between-subjects ANOVA. In the first paragraph of the analysis, specify the assumptions for the analysis and check whether the assumptions are held. Present a table for the test of normality and report the test for homogeneity of variance. Present the results of the omnibus test in the following paragraph. Include **an ANOVA summary table.** Report and explain the results. Remember to report the effect size and explain what it means.
6. If necessary, conduct post-hoc comparisons to find out how the teaching methods differ in student post statistics self-efficacy. If the homogeneity of variance assumption is satisfied, please use Scheffé adjustments. If not, please use Games-Howell adjustments.
7. At the end, provide a paragraph to summarize your findings and conclusion.

8. Please type out your answer on letter-size paper. Your paper should be in APA-style. The text of your paper should be double-spaced, numbered, and with appropriate font size.