PSYC*6380: Minor Assignment 4 (Non-Parametric Analysis)

Due: Monday, February 14th, 2022 @ 11:59pm

Part 1:

Background:

After hearing a few of Scott's questionable life stories, a researcher became interested in predicting whether any good may come out of dangerous and potentially life-threatening experiences. Specifically, the researcher wanted to assess whether having a near death experience may predict future psychological resilience.

To do this, the researcher collected data from a sample of 100 participants. Upon coming to the lab, each participant was asked to recall a time when they had experienced something they considered "dangerous". A research assistant later coded these responses into: 1) experiences that constituted a "near death experience"; or, 2) experiences that were not life-threatening. The researcher also administered a clinical measure of psychological resilience to each participant.

Your task is to load and prepare the data, and then use a non-parametric analytic technique to look at the relation between whether participants had a near death experience and their psychological resilience.

Filename: "*nonParametricAssignmentData.csv*"

Structure: Comma-separated values

Variables:

Participant ID Number (*ID*): 1-100

Previous Near Death Experience (*Experience*): 1 = "Yes"; 2 = "No"

Psychological Resilience (*Resilience*): 1-100 scale

Missing Data Code(s): N/A (no values are missing in this file)

Your Task:

Use the techniques we covered in today's class (and draw on material from our first class on using *R* to load and clean data) to test for parametric assumptions, and address any violated assumptions using an appropriate non-parametric analytic technique.

Once you've done this, please answer the following questions about the results. Unless otherwise indicated, you do not need to provide full-sentence answers; just the numbers the questions request are fine:

1. Load the data file, and then test for violations of the normality assumption using the *Shapiro-Wilk Test*. Please report both the "*w*" statistic and its associated *p*-value **(0.5 marks)**. To further explore any issues with non-normal data, please also report the skewness and kurtosis of participants' resilience scores. **(0.5 marks)**.

2. Test for violations of the homogeneity of variance assumption using the *Brown-Forscythe Test*. Please report both the test statistic and its associated *p*-value **(0.5 marks)**. To further explore any issues with heterogeneous variances, please also report the standard deviation for each of the two groups that are being compared. **(0.5 marks)**.

   Hint. The "*describeBy*" function may help you with the second part of this question.

3. Please create both a histogram and Q-Q plot for participants' resilience scores. Save both of these plots as either images or .pdf documents; and include them with your submission. For full marks, both graphs should be formatted to be consistent with APA style guidelines **(0.5 marks for histogram; 0.5 marks for QQ-plot)**.

4. Run a *Mann-Whitney Test* that compares the resilience scores that participants in the 'near-death experience' condition received vs. those that the 'no near-death experience' condition received. Please report: 1) The associated *u* statistic; 2) the predicted difference in location; and, 3) the 95% confidence intervals around that predicted difference in location **(0.5 marks).** Based on your interpretation of the confidence intervals, how would you best characterize the predicted difference between the two "*Experience*" groups? Please be as exacting as possible in your interpretation. **(0.5 marks).**

5. Please further explore the results of the Mann-Whitney test further by also calculating the Mann-Whitney *r'* effect size. Please report both the Mann-Whitney *z*-score and its associated *r'* value **(0.5 marks).** What does this *r'* value represent, conceptually? **(0.5 marks).**

Part 2:

Background:

To help the researcher better understand how normality and other parametric assumptions work (and to better interpret the results of your analysis), please answer the following conceptual questions in a short answer (i.e., 1-2 sentences) format:

1. In our class today, Scott said that the assumption of normality actually refers to the need for the sampling distribution of scores to be normally-distributed (not the scores themselves). What did he mean by this? What is a sampling distribution, and why do we need it to be normally-distributed? **(1 mark).**

2. In our class today, we went over several transformations that can help correct skewed data. Scott mentioned that *z*-scores are a transformation too. Does that mean *z*-scores will correct skewed data? Please explain why or why not. **(1 mark).**

3. Although there are several mathematically-viable options for handling violations to the normality assumption, they're often considered just a tad questionable. What theoretical questions might you ask yourself when you're considering whether it is appropriate to delete, transform, or replace outliers in your data set? **(1 mark).**

Please provide your full *R* script with your submission and leave comments in your script (i.e., using "#") explaining what each command you wrote does. **(1 mark for including a full script; 1 mark for including appropriate commenting).**

Good Luck!