## Activity 1.1 – Practice with Data Management in Prep

**Observational Units:** nursing faculty (one per rows).

**Goal of study:** Determine to what extent secondary traumatic stress levels (stress induced by exposure to student trauma) is associated with functional impairment.

**Variables.** We plan to employ the following variables to meet the goal.

* *Participant ID.* You will need to add a key to each row to identify the participant.
* *Numeric Scale.* The raw data involve two types of responses which need to be converted to a numeric scale. In both cases, we want to score to increase with frequency by 1 unit each step, with the least frequent label coded as 1 and the most frequent as 5.
  + ForQ41\_1 through Q41\_17, we will score *Never* as 1, up to *Very Often as 5.*
  + ForQ39\_1 through Q39\_7, we will score *None* as 1, up to *A Great Deal as 5.*
* *Secondary Traumatic Score* is the aggregated total numeric score of columns Q41\_1 through Q41\_17 for each participant.
* *Secondary Traumatic Stress Scale (STSS)* is ordinal variable determined using the Secondary Traumatic Score using the following table. This is the explanatory variable.

A white rectangular box with black text

Description automatically generated

* *Functional Impairment* is the aggregated total numeric score for columns Q39\_1 through Q39\_7 and represents the response variable.

**Today’s tasks**. In the first activity, you prototyped the data management process for the above analysis in Tableau Prep. In this follow-up activity, we will create a git/GitHub data repository to house the data, our solutions, analysis, etc. This will include.

1. A summary table and side-by-side box plot showing to what extent the mean functional impairment scores differ by STS category.

A screenshot of a computer

Description automatically generated

A graph with dots and lines

Description automatically generated

**Functional impairment scores are correlated with the STS category, as the STS category worsens the functional impairment also increases. There is a big range for people in the moderate STS category. There is a lot of overlap between functional impairment based on STS category however, so you can’t decide what category a person belongs in based off their functional impairment level.**

1. A WORD document containing the summaries from above, along with a short summary of your findings.
2. A GitHub repository containing
   1. A Tableau Prep file containing all your data management work.

**Screenshot (final Prep flow):**

* 1. The raw/original data and the final CSV in a data folder.

**A screenshot of a computer

Description automatically generatedScreenshot (contents of the data folder on GitHub):**

* 1. The WORD document containing our outcomes and summary.

**Upload to D2L + screenshot of document in repository.**

**A screenshot of a computer

Description automatically generated**

1. **(As time permits)** Start implementing the data management process in an R Colab notebook using the function from the tidyverse. Save your progress on this notebook to your GitHub repository.
2. **(As time permits)** Pull your repository onto your local machine and verify that you can run the R notebook in our r-notebook conda environment (see activity 1.2). Make some changes, commit these changes locally, then push them to GitHub.

**Final Deliverables.** Upload this document, including the screenshots and a link to your repository, to the D2L assignment folder.

**Appendix – The Corresponding Survey**



