Project Outline

Description of Data:

This dataset was collected in 2020 to examine young adults' descriptions of lasting impacts of early experiences of (dis)trustful interactions with teachers. Eleven items were modified from the Ledbetter and Vik's (2012) Parental Privacy Invasions Scale (PPIS) and twelve items were modified from Levine and McCornack's (1991) Suspicion Scale to assess young adult and teacher trust interactions. The PPIS is on a five-point Likert scale, ranging from never to very often. The Suspicion Scale is on a seven-point Likert scale, ranging from strongly disagree to strongly agree.

106 participants were recruited primarily through an undergraduate student body located in Minnesota. The participation criteria required that participants be at least 18 years old and have interacted with a teacher in a school setting. Demographic information, such as age, gender identity, and ethnic-racial identity was also collected.

Preparatory work:

1. Convert the dataset from wide format into long using the gather() function within the `tidyr` package
2. Recode categorical variables into numerical data using recode() function within the `dplyr` package
3. Use clean\_names() function to tidy the variable names of the data within the `janitor` package

Meeting the Final Requirements:

To meet the requirements for the final paper, we must include two exploratory data visualizations and summary statistics of the data.

To meet the data visualization requirement, we will include the following visualizations:

* In the dataset, there is a subscale that measures student perceptions of their own trustworthiness, as well as privacy preferences when interacting with teachers. To examine the relationship between these variables, we will create a scatterplot with geom\_point() and also plot the line of best fit with geom\_smooth().
* For the second visualization, we will be examining categorical, rather than continuous, data. In our dataset, there is a subscale measuring student’s experiences interacting with teachers, especially in regards to the teachers trying to modify the student’s behavior. We want to plot this data separated by gender, which was measured in our dataset. To do this, we may use either geom\_col() or geom\_boxplot(), depending on what communicates the information more effectively.

To meet the requirements for including summary statistics, we will include the following summary statistics:

* Summary statistics of the subscales of the Parental Privacy Invasions Scale and the Suspicion scale.
  + Will also include summary statistics of the student’s experiences interacting with teachers subscale grouped by gender, to go alongside our second visualization.
* Summary statistics of the age variable in the dataset.
* Counts for categorical variables in the dataset (gender, ethnicity).
* To calculate summary statistics, we will use the summarize() function.
* To calculate counts, we will use the count() function.
* To create tables of our data, we will use the kable() function in the knitr package.

With these inclusions, we will meet the main requirements of the assignment, namely including two data visualizations and summary statistics.