**308B DA1: Assumptions and Diagnostics of the GLM**

Download the dataset “*308B DA1 Data.csv”* and complete the following tasks. You are always welcome to do extra cleaning before proceeding (e.g., rename or reorder variables), but that is not required for the assignment.

**Prompt:** Stress-related growth (sometimes called post-traumatic growth) refers to positive growth after a stressful or traumatic event. In the present study, athletes who had suffered serious injuries during sport participation were surveyed regarding their experiences while injured. The researchers aimed to determine if the severity of injury (*moderate = 0, severe =1*) predicts whether someone would experience growth (*0-100, 100 meaning more growth*) after their injury.

Question 1: Check for outliers and handle them accordingly. Describe your process in checking and handling outliers. What decision did you make and why?

Question 2: One assumption of the GLM is normality. In your own words, what is the impact to our data when the distribution is not normal?

Question 3: Check the assumptions of Normality (univariate) that are relevant to this dataset. List the assumptions you checked here and the degree to which they are violated.

Question 4: Another assumption of the GLM is homoscedasticity. In your own words…

* 1. Define homoscedasticity in the context of the GLM.
  2. How do we check for homoscedasticity when we are conducting a t-test and what are your options to consider when the assumption of homoscedasticity is violated?
  3. What is the impact to our data when the data is heteroscedastic?

Question 5: Check the assumption of homoscedasticity in this dataset. Please report what you find in APA format.

Question 6: Now that you’ve checked the assumptions, conduct the appropriate analysis to address the researcher’s question. Report your findings in APA format.

Question 7: In your own words, define collinearity and the issue the violation of the assumption would pose for your data.

Question 8: What is the relationship between models and residuals?

Question 9: How would you explain to a fellow student that the goal of checking assumptions is not about achieving perfection but about understanding how deviations from the ideal might influence your conclusions? Provide examples to illustrate your point.

Question 10: Identify the best visualization and best quantitative indicator of each of the following assumptions for an independent t-test.

1. Univariate Normality
2. Homoscedasticity