**Missing Values Lab**

For the following lab worksheet please open the Lab 1 – Missing Values.sav file on Canvas.

As a researcher, you are interested in several factors that impact high school GPA. You pulled data from a historical survey file and noticed that there are several missing values. Being the good statistician that you are, you decided to impute the missing data for a complete data set. Before imputing the data, please analyze the pattern of missingness.

1. What % of variables have missing data? How % of cases are missing data? What % of all values are missing incomplete data?
2. What variable has the highest percentage of missing data (include count and percent missing)? What variable has the lowest percentage (include count and percent missing)?
3. What missing value pattern is most occurring? Describe what that pattern is.
4. Based on the missing values chart, are there any issues with monotonicity? How do you know?
5. Were there any significant patterns of missingness between variables? If so, which variables? Please report the t-test and p-value associated with the pattern
6. Can we assume the data is missing at random? Why or why not?
7. As a researcher, you want to run an inferential technique on the dataset. Based on this information and your analysis of the pattern of missingness, would you run an Expectation Maximization or Multiple Imputation method to estimate the values of the missing data? Please explain why.
8. Run the best missing values technique on the dataset and run an independent sample t-test using sex as the independent variable and number\_grade as the dependent variable. Please attach your output and write up your findings in APA format. Please be sure to interpret the right t-value.