# Lab 06: Regression with a Single Categorical Predictor (e.g., Independent Samples T-Test).

As always, indicate your answers using a different color font or shading to clearly separate your answers from the questions. When you are finished, save the file as "lab05\_FIRST\_LAST" using **your** first name and **your** last name, and then upload the file as a Word Document or .pdf on Canvas.

Start by opening the data\_SAT.csv datafile in Jamovi. These data show the results of hypothetical SAT preparation course for high school students. Students were randomly assigned to either take the prep course or do their own preparation and then their SAT test scores were assessed.

## Question 1.

1A. Create a **boxplot** showing SAT scores as a function of whether or not students took the course. Be sure to turn on the data points:

1B. Next, conduct an **independent samples t-test** to analyze your data using the t-test tab. Turn on the mean difference with 95% confidence interval, and the descriptive statistics. Paste the tables from the output below:

1C. Next turn on the assumption checks from the t-test tab (include the homogeneity of variance test, normality test, and the QQ Normal Plot). Insert those outputs below:

1D. Finally, provide a written interpretation of these results. Be sure to address the observed effect of the intervention, the statistical significance of your hypothesis test, and provide an interpretation of the confidence interval.

Question 2

Next, we want to analyze the same data, but explicitly using a regression model. (Note that the t-test tab is effectively running this regression, it just does so implicitly.) In the regression tab, make SAT scores the dependent variable and Group the independent variable. Recall that Jamovi requires that categorical predictors go in the “Factors” box whereas continuous predictors go in the “Covariates” box.

2A. **Regress SAT scores onto the Group variable**. Turn on the 95% confidence interval under the Model Coefficients tab, and turn on the Shapiro-Wilk Test of Normality and the QQ Normal plot of residuals under the Assumption Checks tab. Insert those outputs below:

2B. Note that the *r2* for the model is 0.196. **Explain what this *r2* value means**, be sure to address the magnitude of the *r2* and if you think the explained variance is small, moderate, or large.

2C. Note that *r2* is a “standardized” measure of effect size. That is, it gives an indication of the size of the effect, but it has been converted to proportion of variance explained. In contrast, the mean difference and it’s 95% confidence interval are “raw” measures of effect size. **Explain what this slope and 95% confidence interval mean**, be sure to address the magnitude of the slope and if you think the effect of the SAT prep course is small, moderate, or large

Question 3

3A. Provide a written reflection of the two different methods you have for conducting this analysis: using the t-test tab and using the regression tab. Both allow you to conduct an independent samples t-test, but one allows you to do it explicitly whereas the other does it implicitly in the background. Which of these two methods do you prefer and why?