# Quiz 3

**Name:**

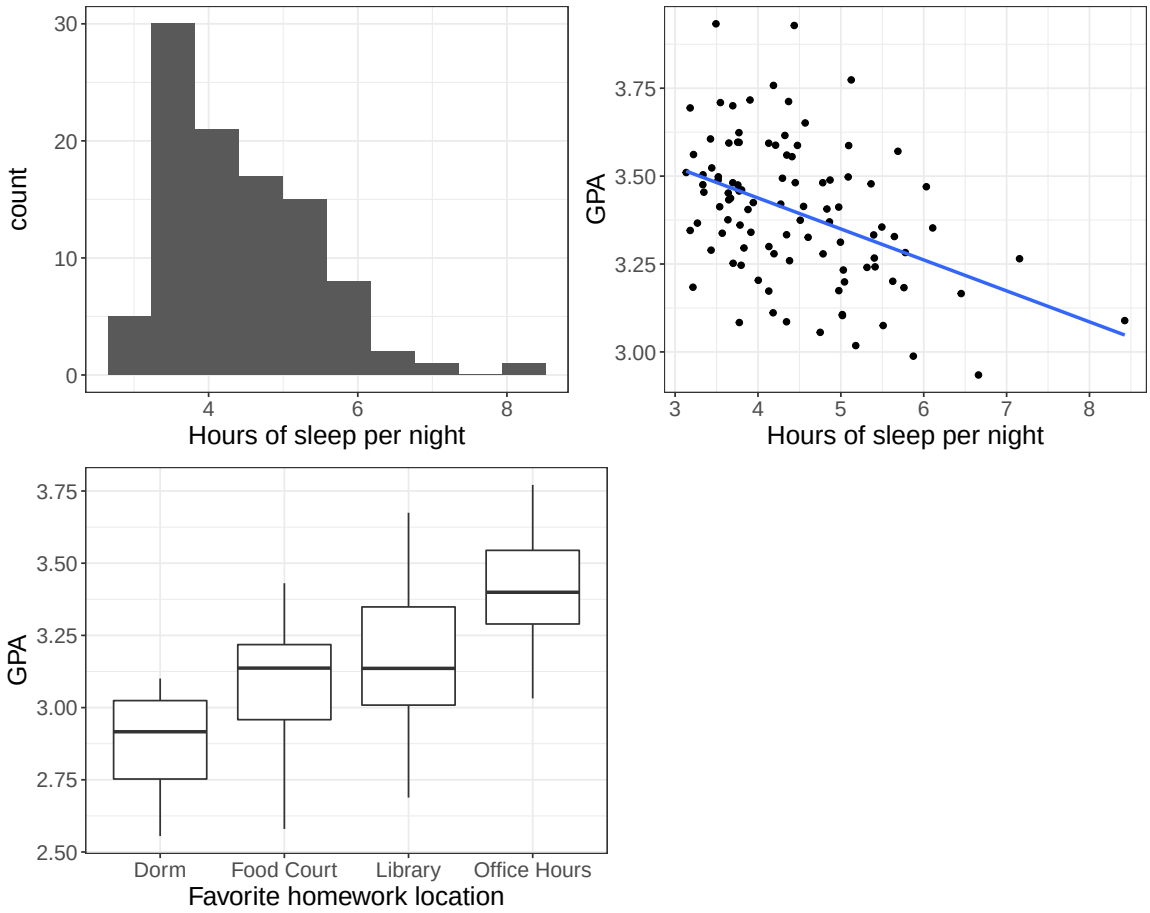
**Instructions:** You have 20 minutes to complete this quiz. You may use any notes or resources, including your laptop and the internet (R documentation, StackOverflow, Wikipedia, etc.). You may not communicate with anyone else during the quiz.

There is one question. Relax and do your best! Try to answer as much as possible.

The scenario is described on pages 2-3. The questions are on page 4.

## Scenario

Imagine you are a statistician studying factors associated with student success in college. One of your collaborators conducts an observational study in which they survey 100 Wake Forest students, and record their GPA, the average number of hours they sleep per night, and their favorite place to do homework. After analyzing the data, they make the following plots:



Your collaborator then writes the following report on their analysis:

"To test whether there is a relationship between amount of sleep and GPA, we carried out a hypothesis test for the hypotheses $H_0 : \beta_1 = 0$ vs. $H_A : \beta_1 \neq 0$, where $\beta_1$ is the true slope for the linear regression model between sleep and GPA in the population.

The p-value for the hypothesis test is 0.000013. This means that there is only a 0.000013 probability that there is no relationship between sleep and GPA, so we have strong evidence for a relationship. The estimated slope is $-0.088$ and the estimated intercept is 3.79. This means that an additional hour of sleep is associated with a GPA of 3.79, on average. Unfortunately, the normality assumption for this linear model is not appropriate because the distribution of hours of sleep is right-skewed, as we can see in the histogram.

Finally, from the boxplots we can see that there is a positive, linear relationship between students' favorite place to do homework, and their GPA.

Based on these plots and the fitted model, we recommend that students who want to improve their GPA set their alarms for one hour earlier, and avoid doing any homework in their dorm rooms."

**Question:** There are no errors in your collaborator's plots, and they correctly calculated the estimated slope (-0.088) and intercept (3.79) of the fitted regression line. The reported p-value (0.000013) is also correct. Unfortunately, they made several errors in their interpretation of the figures, estimated regression line, and/or p-value. List these errors below, and explain to your collaborator why each of their errors is wrong.