Exam 1

Last Name: _____

First Name: _____

I hereby state that I have not communicated with or gained information in any way from other students or any outside resource during this exam. I agree to abide by the rules stated below, and to abide by the Wake Forest Honor Code. All work is my own. I understand that any violation of this agreement will be reported to the Honor Council and will result, at minimum, in a 0 on this exam.

Signature : _____

- -

All work on this exam must be your own.

1. You have 50 minutes to complete the exam.

2. Show all your work on the open ended questions in order to get partial credit. No credit will be given for open ended questions where no work is shown, even if the answer is correct.

3. You are allowed a calculator, however you may not share a calculator with another student during the exam. The calculator must be only a calculator, and may not be connected to the internet.

4. You are allowed to ask clarification questions to me, but you may not ask anyone else.

5. You are **<u>not</u>** allowed a cell phone, even if you intend to use it as a calculator or for checking the time. You are **<u>not</u>** allowed a music device or headphones, notes, books, or other resources.

6. You may **<u>not</u>** communicate with anyone other than myself during the exam.

7. Write clearly and be clear. Make it easy to find your answers.

Good luck!

- -

The Data We have a client who is interested in examining a recent advertising campaign for DoorDash, a popular food service in which food from restaurants is picked up and delivered to your door. The client wants to examine the relationship between exposure to advertising and the use of the service at a local university. They have data on 150 randomly selected students from this university, and the data were collected during the month of September when the campaign was active.

We have information on X = the amount of exposure a student has to the ad campaign during the month (measured in minutes, between 0 to 6 minutes. Decimals are permitted for values less than a whole minute) and Y = the amount of money the student spent on DoorDash during that month (in US dollars).

1. Based on the information we have so far, write down a population model you would suggest the client consider to model the desired relationship. Use appropriate notation.

The client fits the model and obtains the following output:

Model 1

	Estimate	Std. Error
(Intercept)	4.50	0.12
AdExposure	1.26	0.08

2. Based on the output above, write down the equation of the fitted model.

3. Interpret the slope and intercept of the fitted model.

4. What is the predicted amount spent by a student who was exposed to 3 minutes of DoorDash ads? Show all your work.

Model 1

	Estimate	Std. Error
(Intercept)	4.50	0.12
AdExposure	1.26	0.08

The client is interested in whether there is a positive relationship between ad exposure and the amount students spend on DoorDash.

5. The client builds a 99% confidence interval for the slope, and interprets it as follows. "There is a 99% probability that the slope 1.26 falls between 1.18 and 1.34." (a) Is the client's confidence interval and interpretation correct? (b) If so, show work to back up the client's claim. If not, provide the correct confidence interval and interpretation. For reference, the critical value is $t^* = 2.61$.

Model 1

	Estimate	Std. Error
(Intercept)	4.50	0.12
AdExposure	1.26	0.08

6. Write down null and alternative hypotheses, in terms of one or more parameters in the model, which allow you to test whether there is a positive relationship between ad exposure and the amount students spend on DoorDash. Then calculate the test statistic for this hypothesis test. What distribution would you use to calculate a p-value with this test statistic? (Provide the name of the distribution, and the degrees of freedom if applicable. You do not need to actually calculate a p-value).

To help assess regression assumptions, you create the following residual plot:
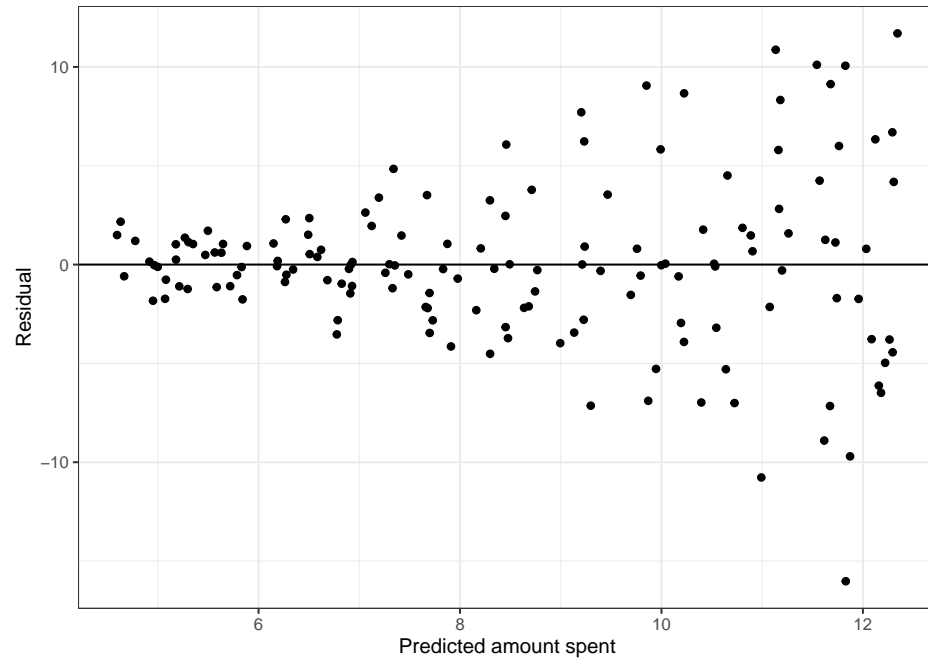
7. Which regression assumption is violated here? How does violation of that assumption affect our confidence interval?

8. Your client looks at the residual plot and writes the following summary. "The points are randomly scattered above and below the horizontal line, with no clear pattern to the residuals. Therefore, the independence and randomness assumptions are appropriate." Explain to your client why their summary incorrectly assesses the independence and randomness assumptions. Then correctly assess whether independence and randomness are reasonable for this data.

9. In order to check for outliers, what kind of plot would you need to make? Be clear about what information would be displayed on each axis, and which points in the plot you would call outliers. Give your answer in 1–3 sentences.

You are done!!! Whooo!!!