After preparing answers for your assigned questions, you are to record a video (no more than six minutes long) that addresses each question. **You should start your video by giving your name. Then state the part and question number before giving your answer and explanation.** Make sure all assigned questions are addressed. **End your video by addressing anyone that you worked with on this exam**

Name

Part 1: Q 8
Part I (15 points):
Choose the most appropriate analysis from the following list of options: two-sample mean test, One-Way ANOVA, Two-Way ANOVA, Chi-squared independence, Chi-squared homogeneity, Simple Linear Regression, Multiple Linear Regression with quantitative predictors, or Multiple Linear Regression with a mix of quantitative and categorical predictors. In your recorded video, explain your choice of analysis in context and how the variables will be used in the model/analysis (explanatory/response, categorical/quantitative).
Then discuss the scope of inference for the study in context.

8) A researcher is interested in assessing the relationship between the modulus (elasticity in Newtons per square meter) and thickness (in millimeters) of the thorax of a moth, under the assumption that differences in thickness might be related to differences in the modulus. A single moth is obtained for measurements and locations for testing across the thorax area are randomly selected from all possible testing areas using a random number generator for the locations. A total of n=480 different locations are selected

```
SLR

Modulus is quantitative and thickness is quantitative
Interested in the relationship between the two

Thickness as explanatory and modulus as response?
Wording suggests that modulus is the explanatory?

One moth, so only applicable to that one moth.
Random locations, so causal inference for that one moth
```

Part 2: Q 16

16) State the method of analysis needed for these data, whether the fitted model is additive or has an interaction, and then report the estimated model (`lmI_P`) for `PE`, defining any indicator variables used. Report the simplified models for any two of the `AT_bin` levels and explain how you did the simplification. Provide a conclusion for the test of the term, in context, that explores whether the slopes should differ in the model. What does this test suggest about the inclusion of this term?

```
Multiple Linear Regression with a mix of quantitative and
categorical predictors

Fitted model has an interaction

"the relationship between energy output
and ambient pressure to vary by temperature" (Temp is
categorical)

PE is electrical power output (Quantitative) MW(Mega Watts
Ambient Pressure(AP) (Quantitative) millibars
Temp(AT) (Categorical) Degrees Celcius


Estimated Model is:
Estimated the PE = 561.424 - 0.081(AP) - 393.605(Imedium) -
723.560(Ihigh) + 0.368(Imedium)(AP) + 0.676(Ihigh)(AP)

(4.32, 14, 21, 34.3
Imedium = 1 when the 14-21
IHigh = 1 when temp 21-34.3

For low temperature:
Estimated the PE = 561.424 - 0.081(AP)

For Medium temperature:
Estimated the PE = 561.424 - 0.081(AP) - 393.605(Imedium) -
723.560(0) + 0.368(Imedium)(AP) + 0.676(0)(AP)
=167.819 + 0.287(AP)
```

Sloped should differ
Interaction test:
$F_{2,194} = 4.666$, p-value = 0.0105

We have strong evidence against the null hypothesis that all three temperature categories have the same slope (above) and can conclude at least one of the three temperature categories has a different slope in the Combined Cycle Power Plant in the study.
This term should be included.

<u>Who I worked With</u>