Stat 021 Test #1 Corrections

Multiple Choice
Question 1:
  The correct answers should be choice A and C. Choice A is correct because the lower bound of the confidence interval is higher than the average prevalence of lung cancer across the entire state, which is 31.6. Our confidence interval is stating that we are 90% sure that the true value, aka the true average prevalence of lung cancer within this rural area, falls between 37.2 and 40.1. Therefore, it is likely that the prevalence of lung cancer in this area is higher than the state-wide average. Choice C is correct because the definition of a confidence interval is saying that if we were to conduct many, many samples we would find that 1-alpha % of the time, our true value for the parameter can be found within the lower and upper bound. Therefore, choice C is stating exactly this: if the study were to be repeated with new random samples, the resulting confidence intervals would contain the true prevalence of lung cancer in this area 90% of the time.

Question 5:
The correct answers should be choice C and D. Choice C is correct because making a confidence interval for the slope requires assumptions for inference. Choice D is also correct because we want to check for statistical significance.

> **Commented [1]:** Double check

Question 6:
In my answer, I did not mention that in step 2, I should find estimates for the intercept and the slope. Therefore I would run a SLR using the predictor and response and get estimates for the slope and intercept.

> **Commented [2]:** Do I need to add anything else?

Question 7:
**Noting that we met in office hours, and you said that you would give points for sufficient explanation/justification
The incorrect part of my explanation was where I assigned values to SS_{Model} and SSE. I incorrectly wrote that SS_{Model} is equal to 6%, when in reality $SS_{Model} / SS_{Total} = 6\%$.

Question 9 part c:
The correct answer I should have written was to run a confidence interval for the slope. If 100 falls within the bounds of our confidence interval, then we can say that our runner's rate of burning calories is not different from this average for all people in the age group. If 100 does not fall within the bounds of our confidence interval, then we can say that our runner's rate of burning calories is different from this average for all people in the age group.