Guiding questions:

1. Click on “show estimates from linear regression”
2. In your own words, explain what the following terms represent in terms of the graph:
   1. Intercept (where the best fit line intercepts the y-axis):   
      **Expected points earned if the student does not study at all**
   2. Effect estimate (change in y per unit change in x):  
      **Points earned per hour studied**
   3. Standard error (accuracy of effect estimate):   
      **Error for the estimated effect estimate**
   4. P-value (significance):   
      **How significant the estimated effect estimate is**
3. Keeping everything else the same, play around with increasing and decreasing the number of observations. What is the general trend that for the p-value when the number of observations increases? **Decreases**
4. Keeping everything else the same, play around with the effect estimates (2nd slider bar) and error amounts (3rd slider bar). Note what happens to the significance of the association as you change these terms. **Higher effect estimate = more significant. Greater error amounts = less significant**
5. Set the sliders as follows:
   1. 1st slider: 50
   2. 2nd slider: 1
   3. 3rd slider: 10
6. Check the box to add “teacher” as a confounding variable.
7. Color by teacher to see how the data is confounded.
8. Check and uncheck the confounding variable box a couple of times, observing what happens to the p-value when a confounding variable is added. **Usually, the p-value is less significant if there’s confounding**.
9. Turn off color. Without color, do you feel as though you are missing important information about this graph? **Yes**
10. Increase the average point difference between Mr. A and Mr. B’s classes. How does this affect your ability to detect a significant effect of studying hours on test scores? **Increasing the difference decreases the ability to detect a significant effect.**
11. Uncheck the box that says “add teacher as a confounding variable”
12. Using the effect estimate and intercept from linear regression, write out y = mx + b. Double check this with the “show equation used to generate data”. In this scenario, what is x? What is y?

**X = Hours studied**

**Y = Test Scores**

1. Recheck the box that says “add teacher as a confounding variable”, and repeat question 12. What is x2?

**X1 = Hours studied**

**X2 = Whether or not they had teacher A (0 or 1)**