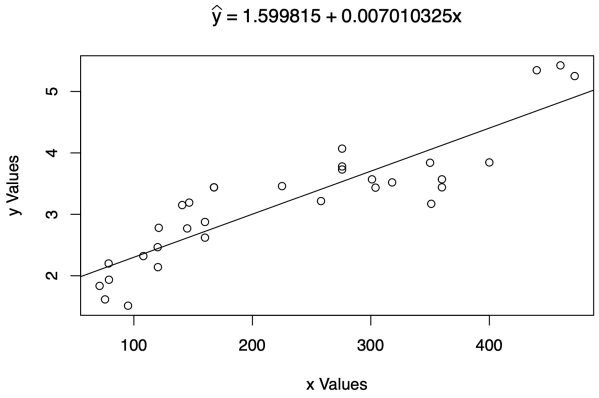
Chapter 5 Exam Prep Worksheet: Bivariate Data

Statypus Insight: From Producer to Consumer

On your exams, you will not be asked to write the R code to generate a scatterplot or calculate a line of best fit. Instead, you will be handed the raw output of the machinery and asked to evaluate it. This turns you into a **Critical Consumer**.

Below is an exact replica of question from a previous exam.

Raw Exam Output:
The output below shows the result of running a linear regression analysis of the variables `mtcars$disp`, which we call the `x Values`, and `mtcars$wt`, which we call the `y Values`.




```
## $Correlation
## [1] 0.8879799
```

Exam Question 1A: How strong is the relationship between `mtcars$disp` and `mtcars$wt`? Give a numerical strength based on the coefficient of determination, and interpret what that number represents in context.

Bill the Statypus says: Notice the Format Whiplash! On the test, the correlation coefficient (r) is handed to you. To find out exactly 'how strong' the relationship is, you need to square it to find r^2. Grab your calculator: $0.8879799 \times 0.8879799$.

Sally the Statypus says: Don't just write a number! A proper 'Statypus Standard' interpretation of r^2 requires you to explain the percentage of variance in the y-variable that is explained by the x-variable. Be sure to use the actual names of the variables, not just x and y!

Your Turn (Write your answer exactly as you would on the test):

Exam Question 1B: If the variable `mtcars$disp` is measured in cubic inches and `mtcars$wt` is measured in thousands of pounds, what would you predict that a vehicle's weight would be if its engine had a displacement of 105 cubic inches?

Sally the Statypus says: See that equation at the top of the output: $\hat{y} = \ldots$? That is your 'Prediction Machine.' Since x is the displacement, simply plug 105 in for x and do the arithmetic. You've got this!

Bill the Statypus says: Don't forget the units!

Exam Question 1C: What range of values of `mtcars$disp` would you feel comfortable making weight predictions? Explain.

Sally the Statypus says: Alert! Question 1C is a trap if you aren't paying attention. Look at the x-axis on the scatterplot description. Where does our actual data start and stop? If you try to use the 'Prediction Machine' outside of that window, what vocabulary word describes that sort of calculation? Is it ok to do that?

Your Turn (Explain your reasoning for 1C):