

Rosalyn Bathrick
I worked alone

- **Q1 (1 pt.):** Why would we want a model selection criterion to penalize the number of parameters in a model?

A: Because a model with many parameters can lose some of its power to identify the predictor acting upon the response variable.

- **Q2 (3 pts.):** In 2 - 3 short paragraphs, describe the meaning of the slope parameter β_1 in the context of the relationship between the predictor variable, x , and the response variable y .

The slope parameter represents the rate of change, or the relationship, between the predictor variable x and the response variable y . For every 1-unit change in x , we can expect a β_1 change in y (on average). When graphed, it also represents the slope of the line with x values on the x -axis and y -values on the y .

An example in the penguin dataset (made up). There is a relationship between flipper length and bill depth – they can be used to predict each other. For every 1mm increase in flipper length, we can expect a 0.5mm increase in bill length. So, for two penguins that have flippers that are 4mm different in length, we can also expect their bills to be 2mm different in length. In this example, $\beta_1=0.5$.

- **Q3 (1 pt.):** Based on the model table, what is the *base case* water treatment?

waterLow

- **Q4 (2 pts.):** What is the average plant mass, in grams, for the **low** water treatment?

2.4g (base case intercept)

- **Q5 (2 pts.):** What is the average plant mass, in grams, for the **medium** water treatment?

3.7g (base case intercept + waterMed)

- **Q6 (1 pt.):** Which of the following questions cannot be addressed with the model coefficient table? Select the correct answer or answers:

(B) Is water availability a significant predictor for plant biomass accumulation