

The Prediction Perimeter

Bill the Statypus says: If you don't have the chapter 5 data loaded, please do it now.

1. The Meteorologist's Dilemma

Mathematics is happy to draw a line that goes on forever. However, biology is not. In this block, we learn why our "Prediction Machine" has a warranty—and why predictions far away from our observed data are inherently risky.

Statypus Insight:

Think about the weather. Meteorologists use current data to predict the future. A forecast for 1 hour from now is very safe because it is a tiny step away from known data. A forecast for 10 days from now is much shakier.

Sally the Statypus says: I agree with Bill here and really implore you to read Section 5.5 until you have read the definitions of **Interpolation** and **Extrapolation**.

2. Establishing the Advisable Range

To help us stay in the "Safety Zone," we calculate what we call the **Advisable Range**. This provides a guideline for when extrapolation is most likely to be reliable.

Coding Corner: Finding the Span

Use R to find the minimum and maximum Height of the fish in our data. Chapter 4 can help.

The Advisable Range

Record the observed limits of your x -variable (Height):

x_{min} : _____ x_{max} : _____

Calculating the 5% Buffer

We consider extrapolation "safe" if we stay within a 5% buffer of the range of the data.

1. Calculate the Range Length: $x_{max} - x_{min} =$ _____
2. Calculate 5% of that Range Length (*Buffer*): _____
3. Your **Advisable Range** is $[x_{min} - Buffer, x_{max} + Buffer]$.

Advisable Range: _____

3. The Software Safety Switch

Sally the Statypus says: The `LinearRegression()` function honors the **Advisable Range**. If you ask it for a prediction outside that guideline, it won't break—it just won't do the math for you.

Coding Corner: Testing the Perimeter

Try to use `LinearRegression` to predict the weight of a fish that is 0 cm tall.

Statypus Insight: The Refusal

Notice that `LinearRegression` just simply overlooks the value of `x0` and still runs the rest of the function. It warns you and leaves it up to you if you want to make the calculation.

4. Rogue Math (The Manual Override)

Bill the Statypus says: I designed this function to refuse to calculate that for you because a 0cm fish is absurd. But if you're determined to be a rogue scientist, you can always ignore the safety switch and use the output to do the math yourself.

Manual Extrapolation

Use your blueprints (a and b) from the previous block to calculate \hat{y} when $x = 0$:

$$\hat{y} = a + b(0) = \text{_____} \text{ grams.}$$

Reflection: Guideline vs. Law

1. If you found a new fish species and its height was just **1%** higher than your x_{max} , would you feel comfortable extrapolating? Why or why not?
2. **Bill's Meteorologist Check:** If a line predicts the temperature will rise 2 degrees every hour, and it's currently 80 degrees at noon, the math says it will be 128 degrees in 24 hours. Where is the error here?