

Simple Linear Regression

Prediction

Prof. Maria Tackett

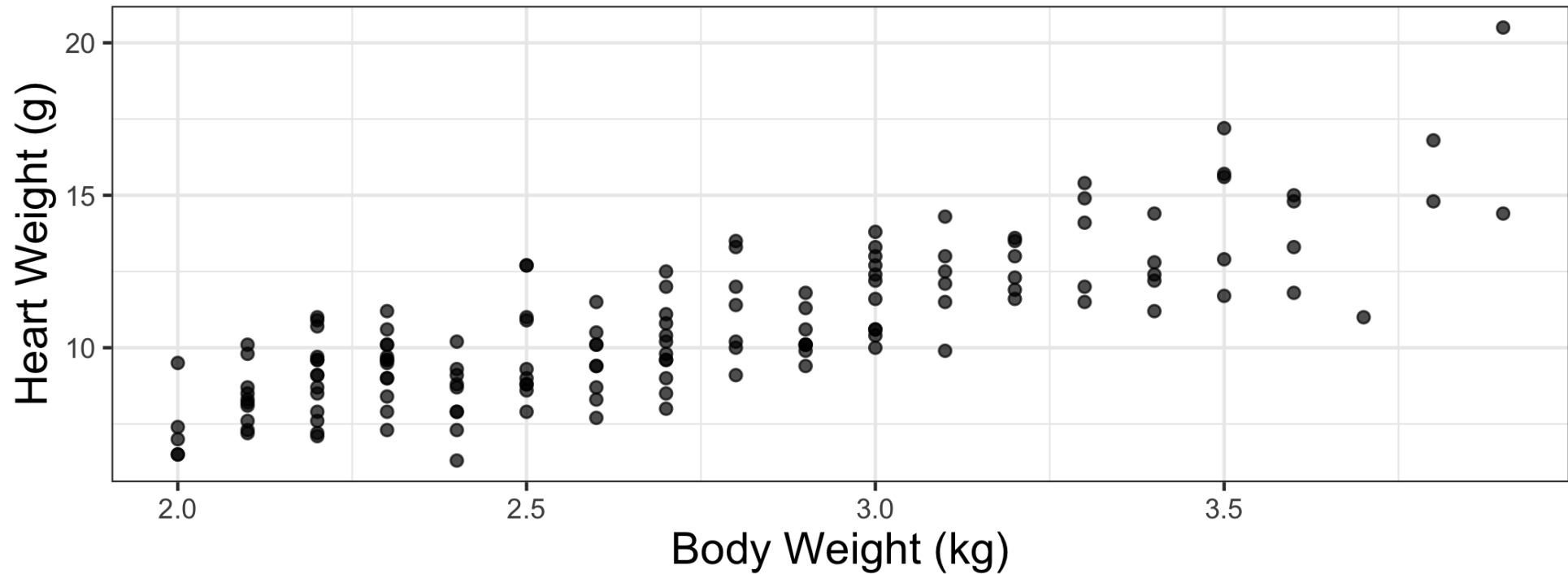
[Click here for PDF of slides](#)

Topics

- Predict the response given a value of the predictor variable
- Use intervals to quantify the uncertainty in the predicted values
- Define *extrapolation* and why we should avoid it

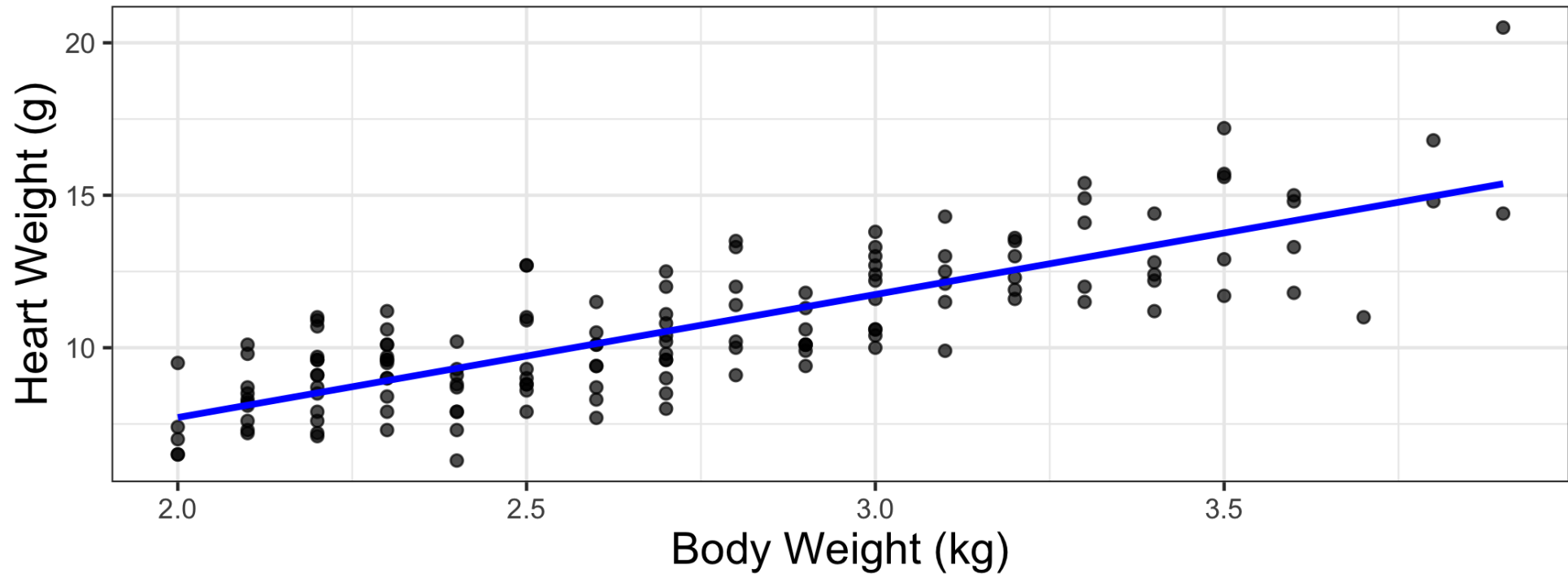
Cats data

The data set contains the heart weight (**Hwt**) and body weight (**Bwt**) for 144 domestic cats.



Cats data

We want to fit a model so we can use a cat's body weight to predict how much its heart weighs.



The model

$$\hat{H}_{wt} = -0.357 + 4.034 \times Bwt$$

term	estimate	std.error	statistic	p.value
(Intercept)	-0.357	0.692	-0.515	0.607
Bwt	4.034	0.250	16.119	0.000

Prediction

We can use the regression model to

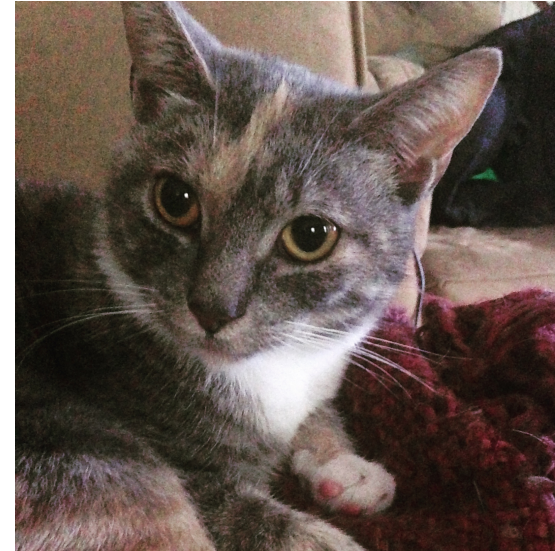
Estimate the mean response when the predictor variable is equal to a value x_0

Predict the response for an individual observation with a value of the predictor equal to x_0

Calculating a predicted value

My cat Mindy weighs about 3.18 kg (7 lbs).

Based on this model, about how much does her heart weigh?



$$\begin{aligned}\hat{Hwt} &= -0.357 + 4.034 \times \mathbf{3.18} \\ &= \mathbf{12.471\text{ g}}\end{aligned}$$

Uncertainty in predictions

Confidence interval for the mean response

$$\hat{y} \pm t_{n-2}^* \times \mathbf{SE}_{\hat{\mu}}$$

Prediction interval for an individual observation

$$\hat{y} \pm t_{n-2}^* \times \mathbf{SE}_{\hat{y}}$$

Standard errors

$$SE_{\hat{\mu}} = \hat{\sigma}_{\epsilon} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

$$SE_{\hat{y}} = \hat{\sigma}_{\epsilon} \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

Standard errors

$$SE_{\hat{\mu}} = \hat{\sigma}_{\epsilon} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

$$SE_{\hat{y}} = \hat{\sigma}_{\epsilon} \sqrt{\mathbf{1} + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

Confidence interval

The 95% **confidence interval** for the *mean* heart weight of cats that weigh 3.18 kg is

fit	lwr	upr
12.472	12.143	12.801

We are 95% confident that mean heart weight for the subset of cats that weigh 3.18 kg is between 12.143 g and 12.801 g.

Prediction interval

The 95% **prediction interval** for an *individual* cat (Mindy) that weighs 3.18 kg is

fit	lwr	upr
12.472	9.582	15.362

We can predict with 95% confidence that Mindy's heart weighs between 9.582 g and 15.362 g.

Comparing intervals

Caution! Extrapolation

We should not use the model to predict for values of X far outside the range of values used to fit the model.

This is called **extrapolation**.

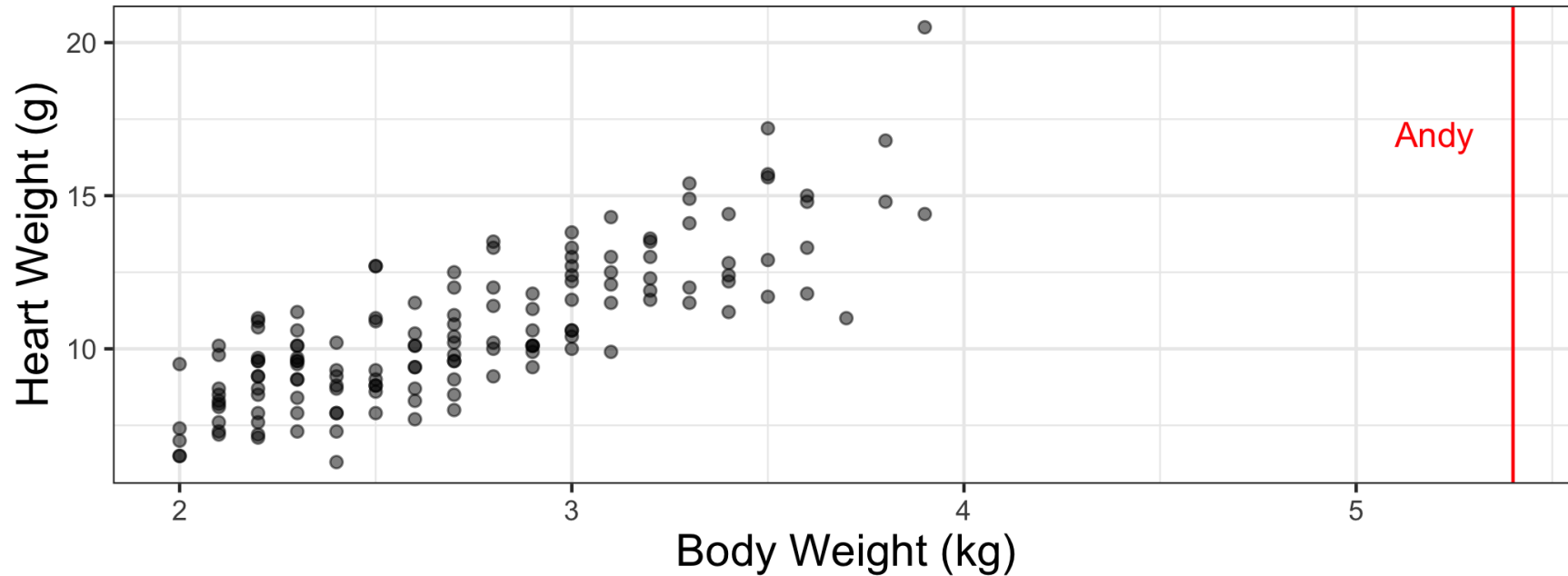
Predict Andy's heart weight?

My cat Andy weighs about 5.44 kg (12 lbs).

Should we use this regression model to predict how much his heart weighs?



Predict Andy's heart weight?



We should not use this model to predict Andy's heart weight, since that would be **extrapolation**.

Recap

- Predicted the response given a value of the predictor variable
- Used intervals to quantify the uncertainty in the predicted values
 - Confidence interval for the mean response
 - Prediction interval for individual response
- Defined **extrapolation** and why we should avoid it