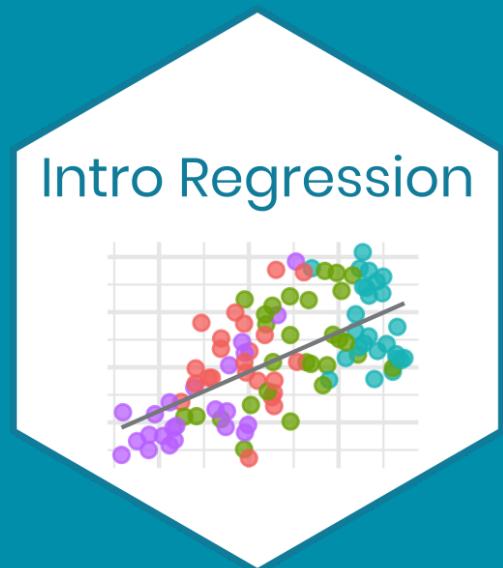


# Simple Linear Regression

## Prediction

Dr. Maria Tackett



# Topics

# Topics

- Predict the response given a value of the predictor variable

# Topics

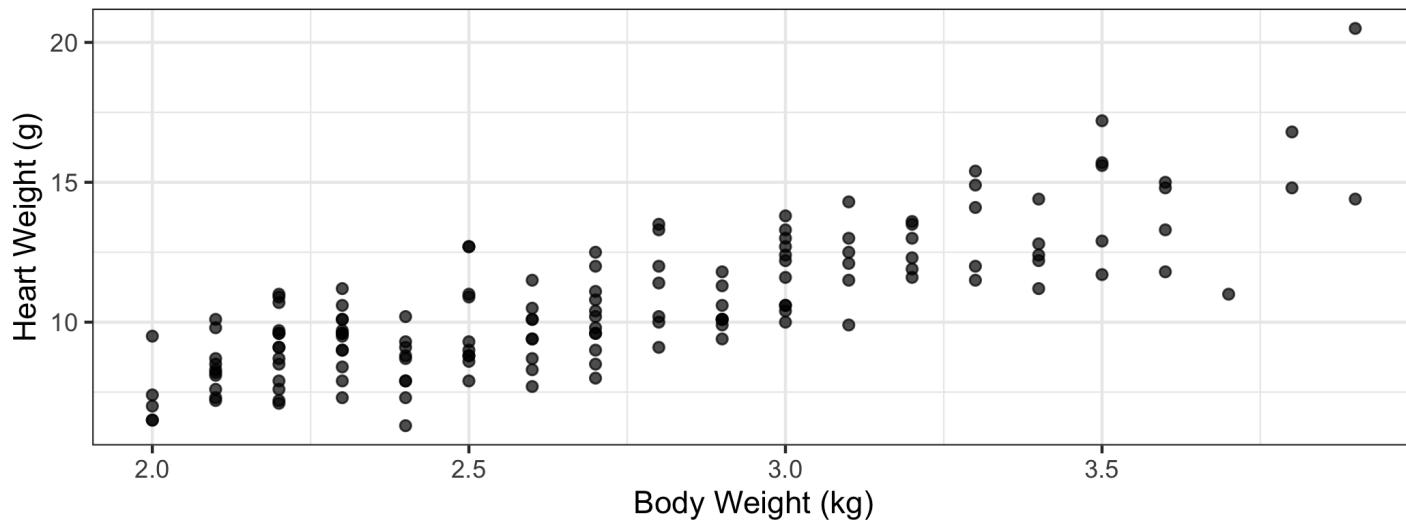
- Predict the response given a value of the predictor variable
- Use intervals to quantify the uncertainty in the predicted values

# Topics

- Predict the response given a value of the predictor variable
- Use intervals to quantify the uncertainty in the predicted values
- Discuss why extrapolation is unwise when using a model for prediction

# Cats data

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

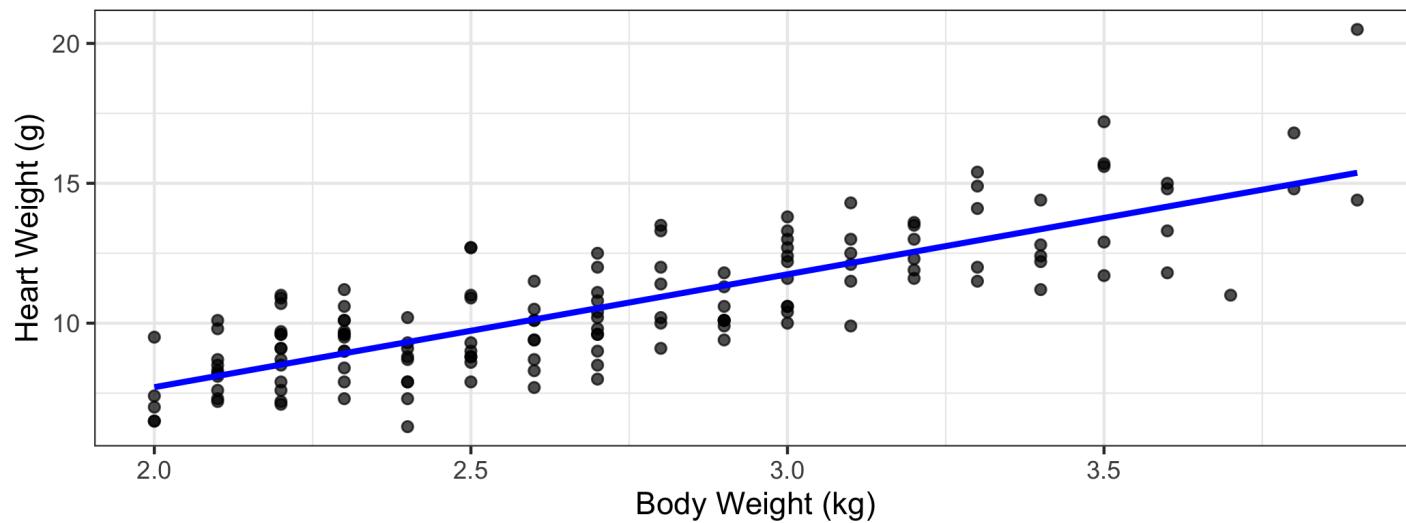


# Cats data

The weight of a cat's heart is used to determine the appropriate dosage for heart medicine.

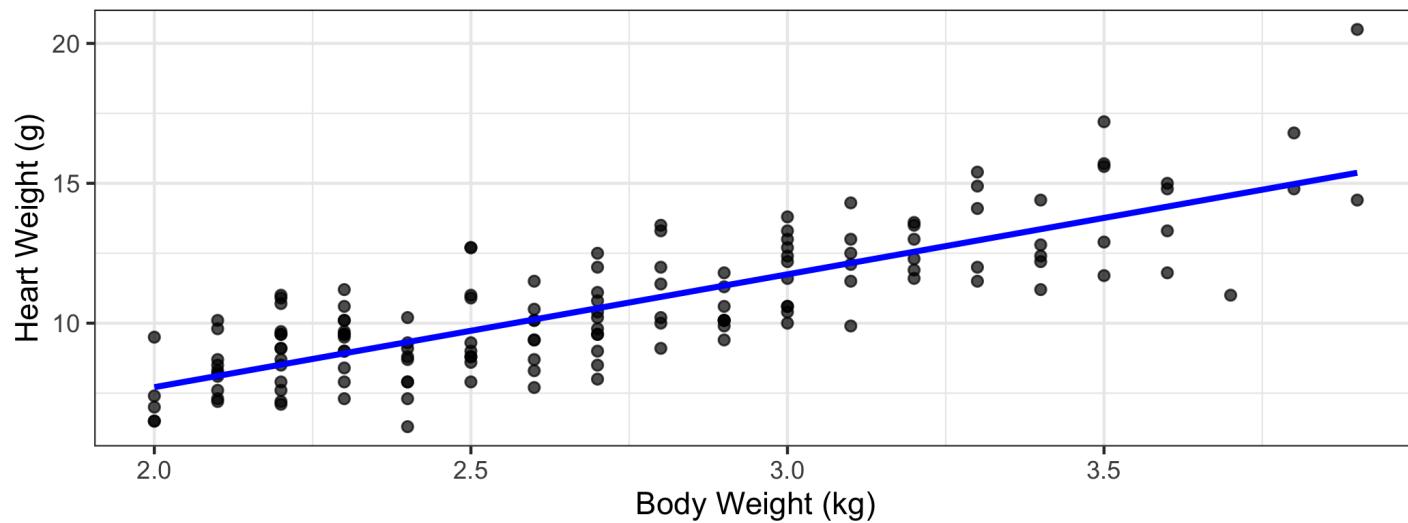
# Cats data

The weight of a cat's heart is used to determine the appropriate dosage for heart medicine.



# Cats data

The weight of a cat's heart is used to determine the appropriate dosage for heart medicine.



We want to fit a model that uses a cat's body weight to predict how much its heart weighs.

# The model

$$\hat{Hwt} = -0.357 + 4.034 \times Bwt$$

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 predict

# Prediction

We can use the regression model to predict

1. The mean response when the predictor variable is equal to a value  $x_0$

# Prediction

We can use the regression model to predict

1. The mean response when the predictor variable is equal to a value  $x_0$
2. The response for an individual observation with a value of the predictor equal to  $x_0$

# Prediction

We can use the regression model to predict

1. The mean response when the predictor variable is equal to a value  $x_0$
2. The response for an individual observation with a value of the predictor equal to  $x_0$

The predicted value will be the same in both cases, but the uncertainty around the prediction will differ.

# Calculating a predicted value

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

What is her predicted heart weight?



# Calculating a predicted value

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

What is her predicted heart weight?



$$\begin{aligned}\hat{H}_{wt} &= -0.357 + 4.034 \times \mathbf{3.18} \\ &= \mathbf{12.471} \text{ grams}\end{aligned}$$

# Uncertainty in predictions

Confidence interval for the mean response  $\mu_Y$

$$\hat{y} \pm t^* \times \mathbf{SE}_{\hat{\mu}}$$

Prediction interval for an individual observation

$$\hat{y} \pm t^* \times \mathbf{SE}_{\hat{y}}$$

## Standard errors

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

## Standard errors

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

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

# 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

# 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 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 that weighs 3.18 kg is

fit	lwr	upr
12.472	9.582	15.362

# Prediction interval

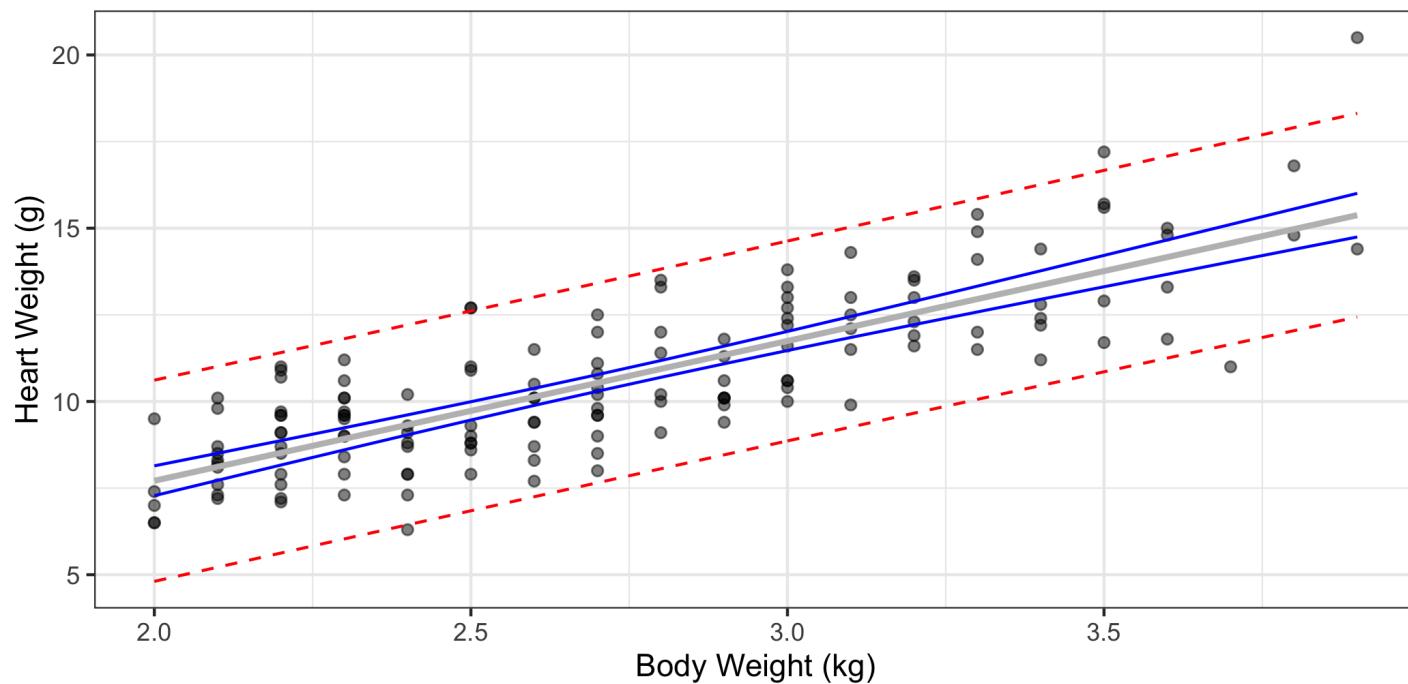
The 95% **prediction interval** for an *individual* cat that weighs 3.18 kg is

fit	lwr	upr
12.472	9.582	15.362

We are 95% confident the heart weight for an individual cat that weighs 3.18 kg is between 9.582 g and 15.362 g.

# Comparing intervals

Confidence interval for mean — Prediction interval for ind



# 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**.

# 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**.

If we extrapolate, the predictions are unreliable since we can't be sure the same linear relationship holds outside of the range of values in our data set.

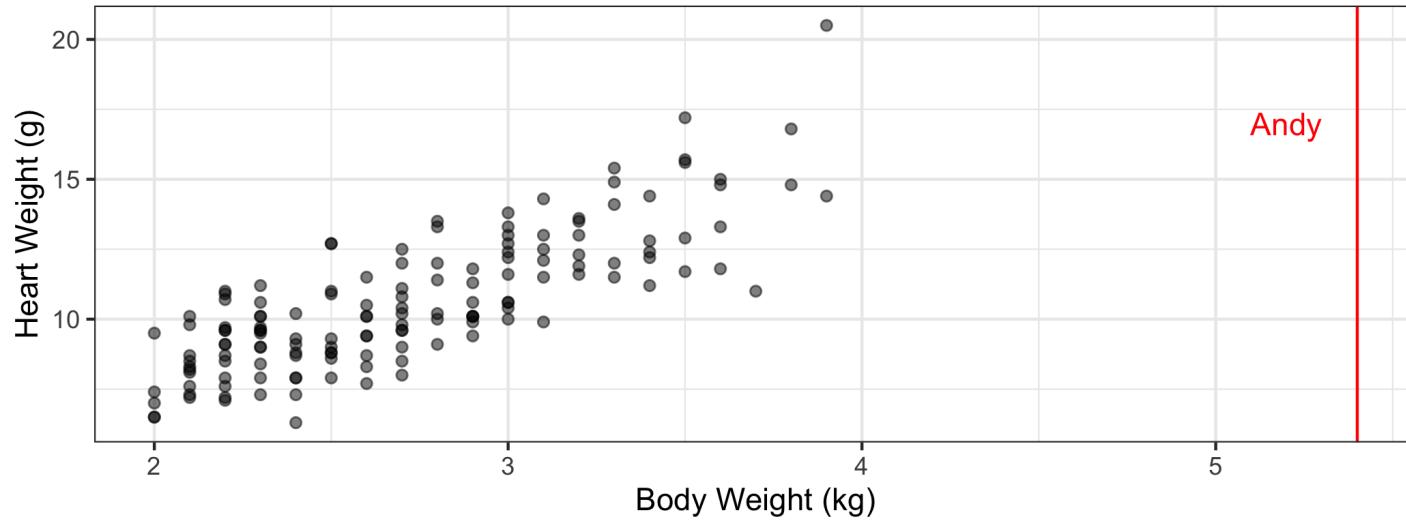
# Predict Andy's heart weight?

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

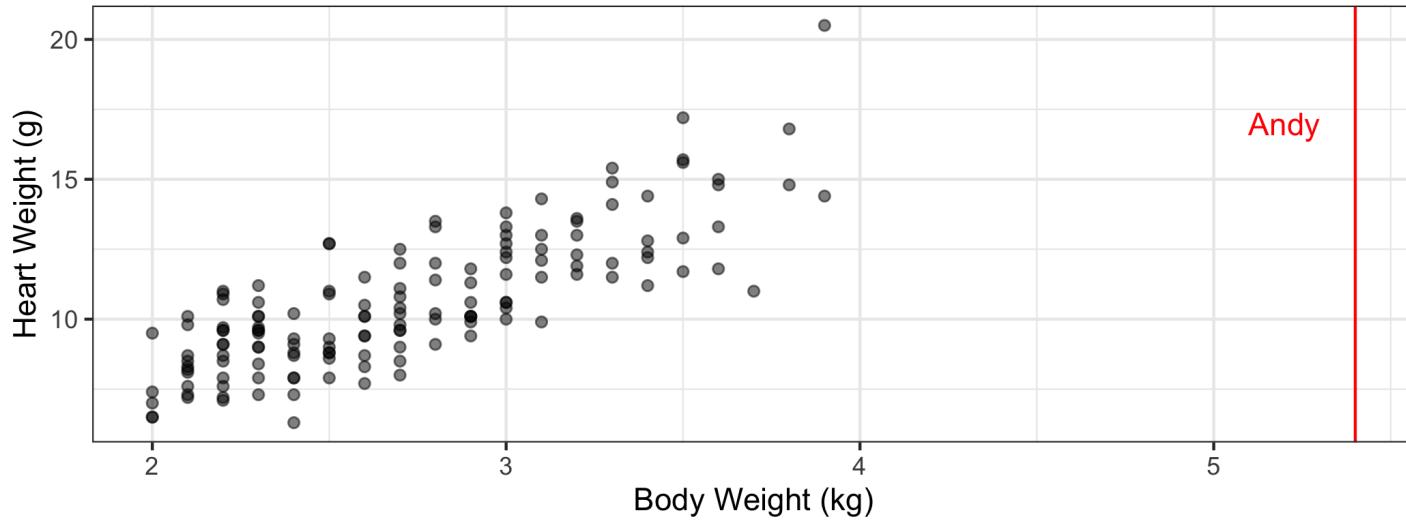
Should we use this regression model to predict his heart weight?



# Predict Andy's heart weight?



# Predict Andy's heart weight?



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

# Recap

# Recap

- Predicted the response given a value of the predictor variable

# 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

# 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
- Discussed why extrapolation is unwise when using a model for prediction

# Simple Linear Regression

## Prediction

Dr. Maria Tackett

