Modelling fish

For this application exercise, we will work with data on fish. The dataset we will use, called `fish`, is on two common fish species in fish market sales.

```
library(tidyverse)
library(tidymodels)

fish <- read_csv("data/fish.csv")
```

The data dictionary is below:

variable	description
`species`	Species name of fish
`weight`	Weight, in grams
`length_vertical`	Vertical length, in cm
`length_diagonal`	Diagonal length, in cm
`length_cross`	Cross length, in cm
`height`	Height, in cm
`width`	Diagonal width, in cm

Visualizing the model

We're going to investigate the relationship between the weights and heights of fish.

- **Demo:** Create an appropriate plot to investigate this relationship. Add appropriate labels to the plot.

```
# add code here
```

- **Your turn (5 minutes):**

  - If you were to draw a a straight line to best represent the relationship between the heights and weights of fish, where would it go? Why?

    Add response here.

  - Now, let R draw the line for you. Refer to the documentation at https://ggplot2.tidyverse.org/reference/geom_smooth.html. Specifically, refer to the `method` section.

```
# add code here
```

  - What types of questions can this plot help answer?

Add response here.

- **Your turn (3 minutes):**

  - We can use this line to make predictions. Predict what you think the weight of a fish would be with a height of 10 cm, 15 cm, and 20 cm. Which prediction is considered extrapolation?

Add response here.

  - What is a residual?

Add response here.

Model fitting

- **Demo:** Fit a model to predict fish weights from their heights.

```
# add code here
```

- **Your turn (3 minutes):** Predict what the weight of a fish would be with a height of 10 cm, 15 cm, and 20 cm using this model.

```
# add code here
```

- **Demo:** Calculate predicted weights for all fish in the data and visualize the residuals under this model.

```
# add code here
```

Model summary

- **Demo:** Display the model summary including estimates for the slope and intercept along with measurements of uncertainty around them. Show how you can extract these values from the model output.

```
# add code here
```

- **Demo:** Write out your model using mathematical notation.

Add response here.

Correlation

We can also assess correlation between two quantitative variables.

- **Your turn (5 minutes):**
  - What is correlation? What are values correlation can take?

  Add response here.

  - Are you good at guessing correlation? Give it a try! https://www.rossmanchance.com/applets/2021/guesscorrelation/GuessCorrelation.html
- **Demo:** What is the correlation between heights and weights of fish?

```
# add code here
```

Adding a third variable

- **Demo:** Does the relationship between heights and weights of fish change if we take into consideration species? Plot two separate straight lines for the Bream and Roach species.

```
# add code here
```

Fitting other models

- **Demo:** We can fit more models than just a straight line. Change the following code below to read `method = "loess"`. What is different from the plot created before?

```
# add code here
```