Something Fishy – Student Handout

## Description of Dataset

Scientists are interested in monitoring the health of trout perch in the Oil Sands Region of Canada. As part of a larger study, fish were collected from the Athabasca River and Peace River, and several characteristics were measured, including the *Weight* of each fish (in grams) and the *Length* of each fish (in millimeters). The dataset used in this analysis includes data from 2088 trout perch. The first six rows of this dataset are shown to the right.

## Simple Linear Regression

Using the fish data: make a scatterplot and fit a linear regression model using $Length$ as the explanatory model and $Weight$ as the response variable.

1. What regression model will you fit to explain Weight using Length as a predictor?
2. What is the regression equation?
3. What is the predicted (fitted) value for the fish with ID 4394? (See data excerpt on page 1.)
4. What is the residual for the fish with ID 4394?
5. Interpret the slope.
6. Interpret the $R^{2}$ value.

1. Examine the residual plots and comment on any potential problems.

## Background[[1]](#endnote-1)

1. Sketch a graph of each of the following:
	1. $y=x$
	2. $y=x^{2}$
	3. $y=x^{3}$



* 1. Considering only Quadrant I (positive x values and positive y values), discuss the similarities and differences among these graphs in a sentence or two.
1. Think about a rectangular prism.
	1. What is the formula for the volume of a rectangular prism?
	2. How many dimensions is the *length* of a rectangular prism?
	3. How many dimensions is the *volume* of a rectangular prism?
	4. Would the *weight* of a rectangular prism be more closely related to its length or its volume?
2. What shape does the scatterplot you made earlier have? What type of function does it look like?
3. If a linear regression using $Length$ as a predictor for $Weight$ does not seem to be appropriate, what variable might be an appropriate predictor for $Weight$? Why?

## New Model

Thinking about your answers in the *Background* section, propose and fit a new regression model that may be a better fit for the data.

1. Propose a new model to explain Weight using Length as a predictor. Why did you choose this?
2. What is the regression equation?
3. What is the predicted (fitted) value for the fish with ID 4394? (See data excerpt on page 1.)
4. What is the residual for the fish with ID 4394?
5. Interpret the slope.
6. Interpret the $R^{2}$ value.
7. Examine the residual plots and comment on any potential problems.
8. Overall, does this model appear to be a good fit for the data?

## Model Comparison

Compare the model you proposed to the original Simple Linear Regression Model.

1. Which model has a higher $R^{2}$ value? What does this mean?
2. For which model is the fitted slope easier to interpret?
3. Compare the residual plots for the two models and comment on the similarities and differences.

## Overall Comparison

Answer the following questions based on the class discussion of all of the models.

1. Out of all of the models, which model do you prefer?
2. Why?
1. Image By Dimitrios Vrettos - https://github.com/karbonk/matc3-algebra1/blob/master/lbr/chap19/fig001\_car.pgf, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=49383602 [↑](#endnote-ref-1)