Digital Commons @ George Fox University

Student Scholarship - College of Business

College of Business

Spring 2022

Fishing Without a Scale

Kaleb Miller

Luke Lemaitre

Follow this and additional works at: https://digitalcommons.georgefox.edu/gfsb_student

Part of the Aquaculture and Fisheries Commons, and the Leisure Studies Commons

Fishing Without a Scale

Kaleb Miller and Luke Lemaitre

George Fox University

Abstract

Our study looks at different species of fish and uses their cross-length and volume to find correlations in weight among the different fish. The different type of fish that we compare include perch, pike, parkki, whitefish, bream, roach, and smelt. The dependent variable that we are using is weight. Once we predict the weight of the individual fish, we will also compare that weight to other fish species. The result of the data are that there is linearity, where the adjusted R squared is between 0.94 and 0.97, depending on our model. Therefore, 94 to 97 percent of the variability in a fish's weight is explained by the regressors. For this first regressor, using volume, there was a correlation of 0.97 between the volume and the different types of fish. Then, for just the length (not length squared) and the different types of fish, there was a correlation of 0.94. In this study, we found that this method of estimating the weight of the fish before they are actually weighed on a scale was found to be most effective for box-shaped fish and not as effective for normal-shaped fish.

Keywords: Fishing without a scale, Perch, Pike, Parkki, Whitefish, Bream, Roach, Smelt, Height, Length, Width, Volume, Weight

JEL Classification: Q22, B23

Part 1: Introduction

Our research question is: Can we create a model to accurately predict the weight of different fish depending on their figure?

This is important because a lot of spring fish scales in the market today are either expensive, not durable (clamp breaks), or susceptible to rust when ocean fishing. Also, floor scales are very small. This means that if a long fish is placed on the scale, the weight estimate will not be accurate as part of the fish will be hanging off the scale. This is why it would be a lot more efficient to use a tape measure and measure a couple of dimensions of a fish to get an accurate fish weight. Moreover, many scales work by suspending the fish with a clamp and while it does not do "physical damage to a fish," it is definitely not the most humane way to weigh a fish (Hudson, 2022, para. 19). After reading this paper, you will know which fish dimensions give the most accurate weight estimate. Also, you will understand that since fish species have different shapes and sizes, some mathematical models may be better for certain species than others. For example, a Rainbow Trout is a longer fish, and a Bluegill is a tall but smaller fish. Lastly, you will understand how a fish's weight changes in different stages of its life. There are a couple of research papers on this topic already from Kaggle, such as "Fish Species Image Data" written by Sripaad Srinivasan and "Fish Market" written by Aung Pyae, and they use the same data set. However, while they all include some visualizations, none of them go as far as making inferences about the relationship between a fish's dimensions and its weight. They all focus more on the mathematical side, such as minimizing the mean squared error using a numerical method known as the gradient descent algorithm. They also compare different regression types on this data set, such as using linear regression, lasso regression, and ridge regression.

While those research papers contain useful and advanced content, this paper seeks to dive into a little more biology and applications. We found that the volume of a fish was the best multiple dimension predictor of weight. Also, from the "Fish Market" research paper, we saw that the distance between the head and tailfin herein, *dlength* of a fish was the best single dimension predictor for weight. This was a surprise to us because this measurement only uses one dimension. We originally thought that diagonal distance between the head and tail fin would be the best predictor of weight because this measurement used the height and length of a fish. A visual can be shown below.

Figure 1

In addition, keeping our research brief, we also included interaction variables with the species which would represent different rates of growth by fish species.

Part 2: Data Overview

The data set used in our research was found on Kaggle, containing 158 observations, seven different fish species, five single dimension variables, and a variable for weight. Here is our data dictionary.

Table 1

Variable Name	Description	Type
Weight	In grams	Continuous
Height	Distance between ventral fin to dorsal fin (centimeters)	Continuous
Dlength	Diagonal distance between bottom of fish's mouth to the top of the tailfin (centimeters)	Continuous
Clength	Distance across a fish, between its mouth and tailfin (centimeters)	Continuous
Thickness	The height of the fish if it was on its side (centimeters)	Continuous
Dwidth	Diagonal distance between the bottom left part of a fish's mouth to the top right part of a fish's mouth (centimeters)	Continuous

Volume	Height * Thickness * Clength (centimeters^3)	Continuous Interaction
Perch	Species name, Perch = 1 if true and Perch = 0 if false	Binary
Pike	Species name, Pike = 1 if true and Pike = 0 if false	Binary
Parkki	Species name, Parkki = 1 if true and Parkki = 0 if false	Binary
Smelt	Species name, Smelt = 1 if true and Smelt = 0 if false	Binary
Roach	Species name, Roach = 1 if true and Roach = 0 if false	Binary
Bream	Species name, Bream = 1 if true and Bream = 0 if false	Binary
Whitefish	Species name, Whitefish = 1 if true and Whitefish = 0 if false	Binary
PerchVol	Perch's Volume; if the species is not a Perch, PerchVol = 0	Continuous Interaction

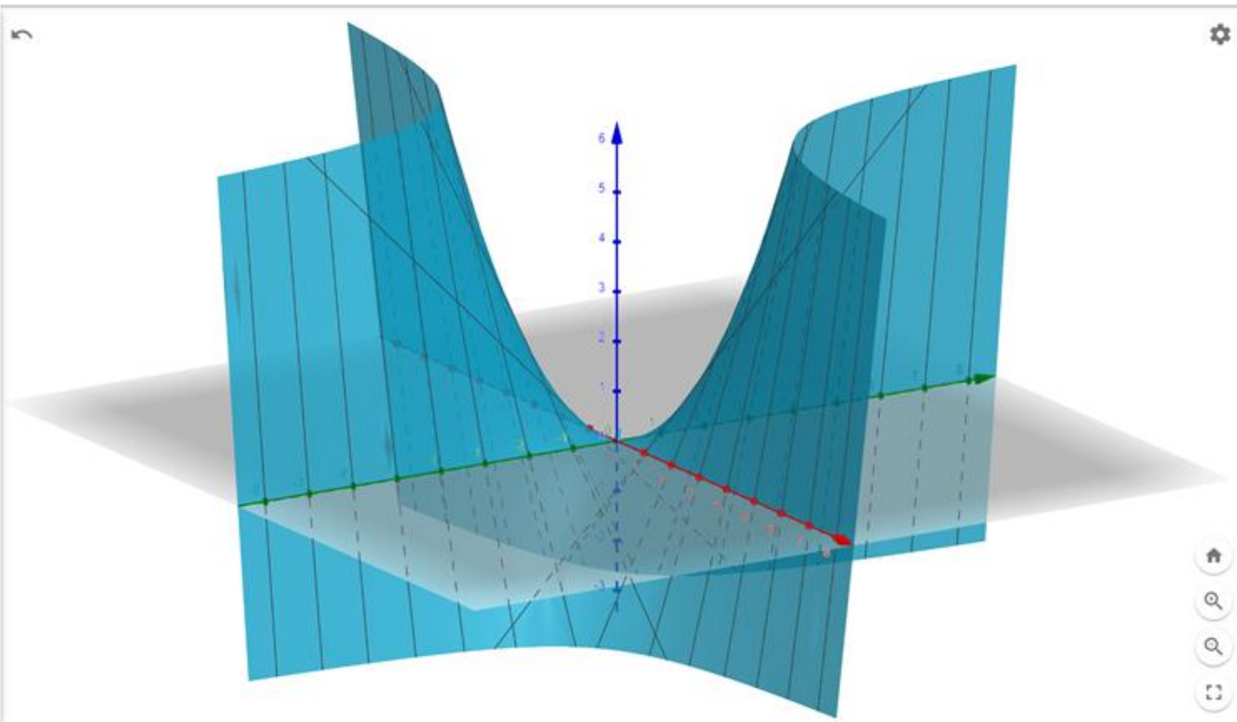
PikeVol	Pike's Volume; if the species is not a Pike, PikeVol= 0	Continuous Interaction
ParkkiVol	Parkki's Volume; if the species is not a Parkki, ParkkiVol = 0	Continuous Interaction
SmeltVol	Smelt's Volume; if the species is not a Smelt, SmeltVol = 0	Continuous Interaction
RoachVol	Roach's Volume; if the species is not a Roach, RoachVol = 0	Continuous Interaction
BreamVol	Bream's Volume; if the species is not a Bream, BreamVol = 0	Continuous Interaction
WhitefishVol	Whitefish's Volume; if the species is not a Whitefish, WhitefishVol= 0	Continuous Interaction

Table 2

Summary Statistics for the following variable:	Mean	Median	Minimum	Maximum	Standard Deviation	5% Percentile	95% Percentile
Weight	400.85	281.5	5.9	1650	357.7	9.8	1000.7
Height	26.247	25.2	7.5	59	9.9964	11.3	42
Dlength	28.416	27.3	8.4	63.4	10.716	11.8	45
Clength	31.227	29.4	8.8	68	11.61	13.1	48
Thickness	8.971	7.786	1.7284	18.957	4.2862	2.196	16.517
Dwidth	4.4175	4.2485	1.0476	8.142	1.6858	1.2772	7.3514
Volume	9996.6	6220.9	139.39	43378	9072.6	322.13	27849

Above are the summary statistics of the continuous variables. We can see that the standard deviation is the greatest for weight and volume. This makes sense because the biggest fish in the sample is a Pike and the smallest is a Parkki, and there is a significant difference in size between the two of these fish. In addition, since *Volume* is the product of *Height*, *Thickness*, and *Clength*, it is not surprising that the standard deviation is as high as 9072.6. So as values of x, y, and z get larger, the gradient vector increases in magnitude. Thus, implying a greater rate of change for volume. Therefore, as changes in x, y, and z occur for large values of x, y, and z, volume changes by a greater amount, causing more deviation. Although the function we defined above only exists in 4 dimensions, a 3D representation is shown below to convey the idea.

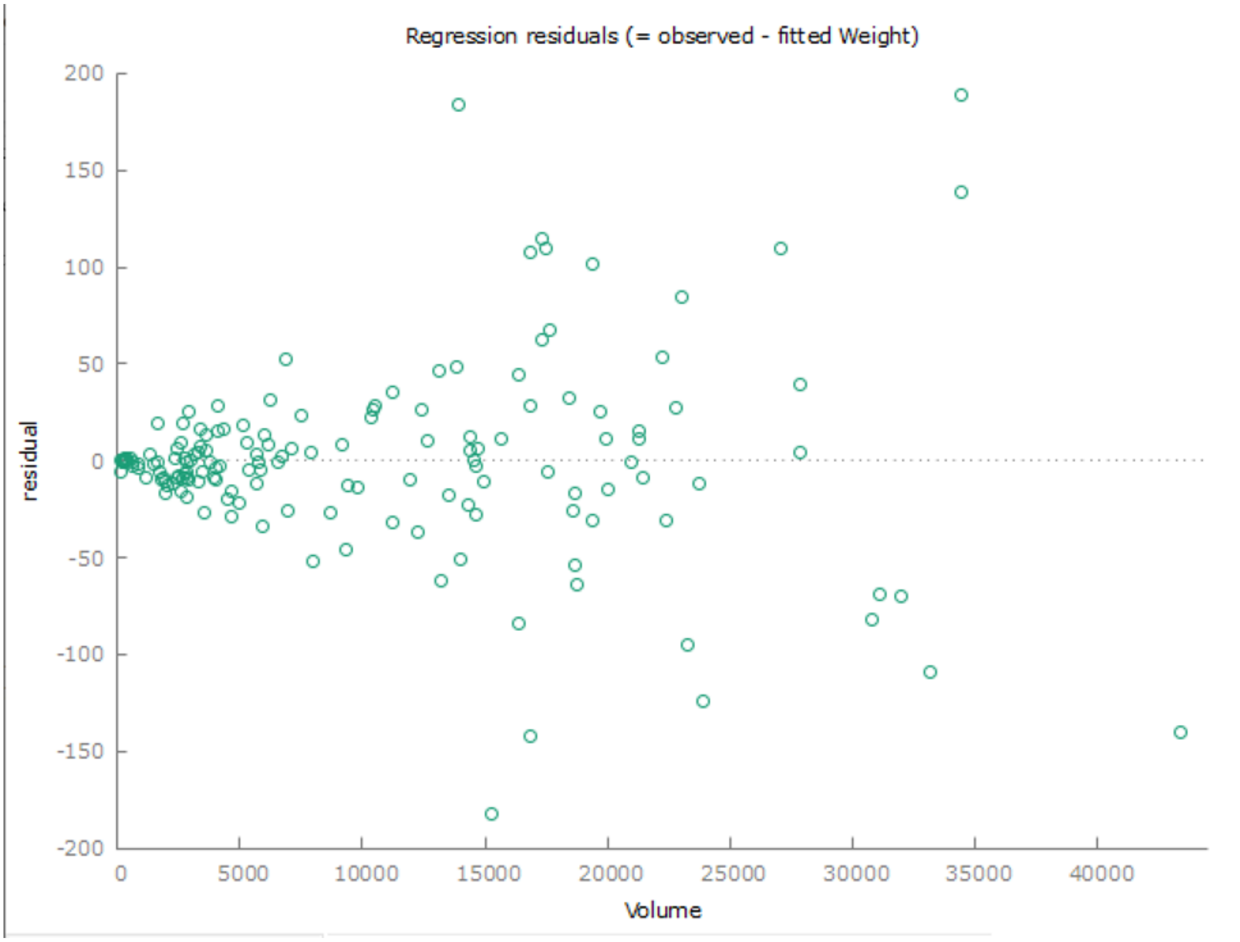
Figure 2



One of the flaws in this research project is that the data set seems to include very small fish in general. From the table above, the mean weight is 400.85 grams, which is 0.884 pounds. That being said, we had the idea of this project applying to all commercial fishing. However, some commercial fish include cod, tuna, and flounder. Cod weigh 11-26 pounds on average, tuna can weigh up to almost 1,500 pounds, and flounder weigh 1-20 pounds. Since the weight range of these fish goes beyond the weight range in the regression, the weight predictions for large fish may be inaccurate. In addition, we only have seven fish species included in the regression. This implies that if we are predicting the weight of one of the seven fish included the fitted weight will be very accurate. However, once we start forecasting for other fish species, the fitted weight may be a lot different than the actual weight. Later on, we will add some observations of fish outside of the weight and species range to see how well our model predicts the weight of these fish. Lastly, a flaw that we initially thought we had was that our standard errors would be too large because the data set only contained 158 observations. When the standard error becomes too large, it ultimately increases the length of our confidence interval for our coefficients. Thus, making it harder to reject the null hypothesis that $\beta_i = 0$ for some $i \in K$, where K is the set of all beta subscripts. However, in our analysis section, our standard errors were low, and almost every coefficient was statistically significant.

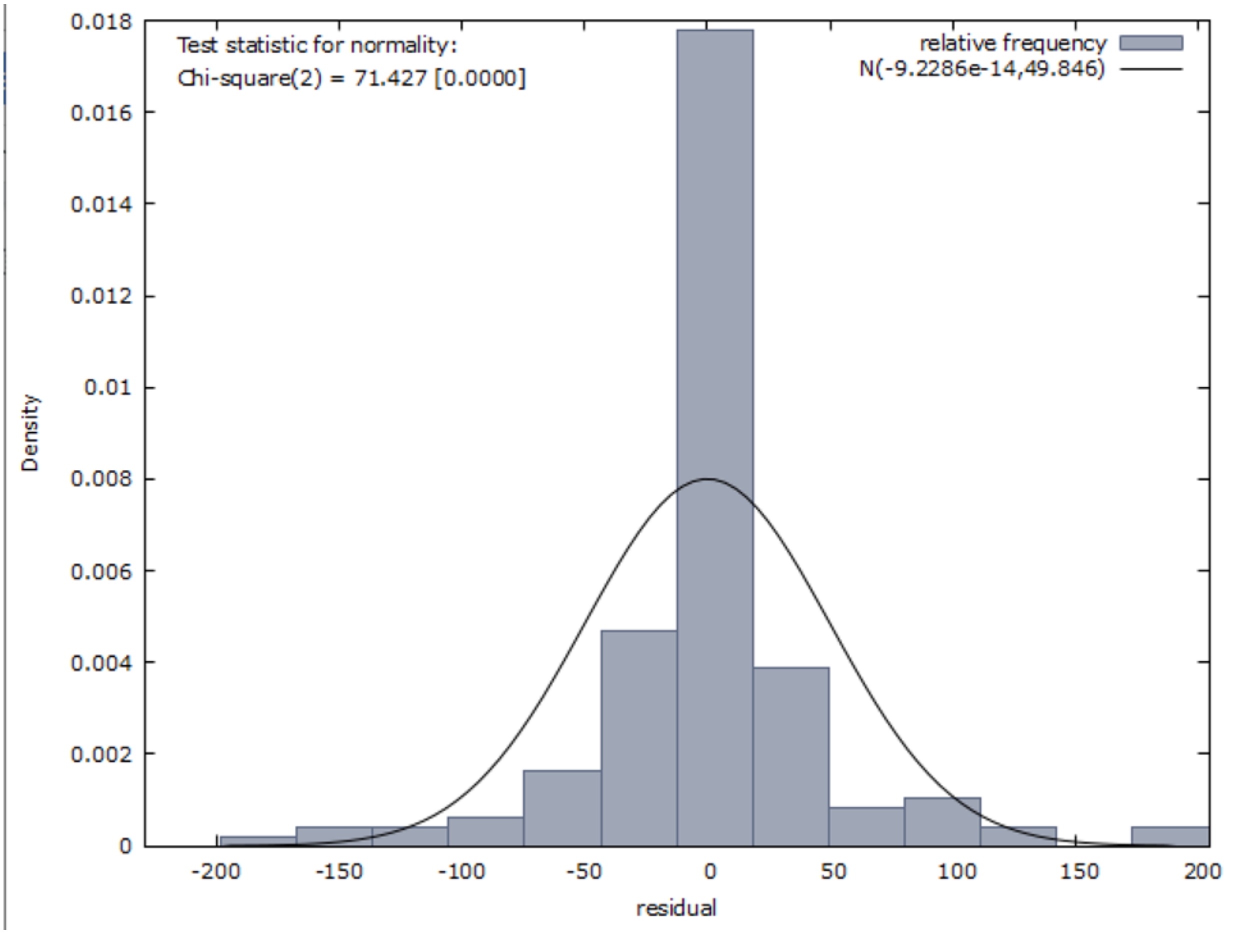
The only regression assumption our data meets is the linearity regression assumption. We have not met constant error variance as var $(u_i \mid X_i = x)$ is not constant. In other words, we do not have homoskedasticity. As far as the interpretation goes, this implies that as a fish grows more in a specific direction, the variation of the weight of the fish is not constant. As evidence, using the Breusch-Pagan test, we get a p-value less than 0.01 using the Chi-Square distribution. Therefore, we are over 99% confident that our error variance is not homoskedastic. We also made a scatterplot of the residuals and the volume, the scatterplot made a cone pattern, thus it is heteroskedastic. As a result, we will use heteroskedasticity-robust standard errors to evaluate the statistical significance of our coefficient estimates.

Figure 3



Our regression also did not meet the normal errors assumption. In fact, the test statistic for normality was 0.00000. This means that we can easily reject the null hypothesis that the errors are normally distributed.

Figure 4



Moreover, we can conclude that there is multicollinearity. If we run a regression and examine the correlation in the model, values are given for each regressor. If the value is greater than 10, there is a collinearity problem. Three out of the 14 regressors are below 10. It was expected that this condition was not going to be met because the data set only had dimension columns. Thus, when incorporating multiple regressors into the model, dimensions tend to move up and down together. Lastly, our regression does not meet the exogeneity condition (omitted variable bias). For example, *Clength* is not included in one of the regressions, but it is strongly correlated with *Volume*. Also, *Clength* is a determinant of weight. Therefore, this meets the two conditions for omitted variable bias.

Part 3: Methodology

We have created many scatterplots with our regressors and our outcome variable: weight. In ALMOST every case, if we were to fit a curve through the points, there would be a positive and increasing relationship between the regressors and weight. Therefore, we knew that a linear model would not be the best fit. Comparing the R squared of our single dimension regressors, we found that the R squared was the highest when we cubed the regressors. However, there was a very slight difference in the R squared between when we cubed the regressors and when we squared them. Also, when we cubed a variable, it made our coefficients for the single dimension regressors less statistically significant. For example, the coefficient for *Clength* had a p-value of 0.03 in the first regression equation, the coefficient for *Clength2* had a p-value of 0.0165, but the coefficient for *Clength3* had a p-value of 0.9727. Therefore, when we just used *Clength* and *Clength2*, the coefficients for each regressor were significant at the 1% level of significance. Below are our R squared values for each single dimension regressor.

1. Height2 (R squared = 0.847)
2. Dlength2 (R squared = 0.856)
3. Clength2 (R squared = 0.875)
4. Thickness2 (R squared = 0.536)
5. Dwidth2 (R squared = 0.827)

Since *Clength2* had the greatest R squared value, we decided to only use *Clength* and *Clength2* as our continuous variables in one of our regression equations. The next part of this equation included the binary variables of the different fish species. This gave us a different intercept for each fish species. When we created a new interaction variable, "volume," we found a linear relationship between weight and volume. This made sense because volume uses three dimensions and using 3 dimensions maximized our R squared. So, our population regression equation used volume as our continuous variable and our six different fish species to allow for different species intercepts. Then, the last part of our equation included an interaction variable, the product of the fish species and the volume. This gave each fish species a different slope in relation to its weight. This only made sense because different species grow differently.

Here is our population regression equation:

$$\text{Weight} = \beta_0 + \beta_1(\text{Volume}) + \beta_2(\text{Perch}) + \beta_3(\text{Pike}) + \beta_4(\text{Parkki}) + \beta_5(\text{Smelt}) + \beta_6(\text{Roach}) + \beta_7(\text{Bream}) + \beta_8(\text{PerchVol}) + \beta_9(\text{ParkkiVol}) + \beta_{10}(\text{PikeVol}) + \beta_{11}(\text{SmeltVol}) + \beta_{12}(\text{RoachVol}) + \beta_{13}(\text{BreamVol}) + u$$

Notice that we did not include *Whitefish* and *WhitefishVol*. This is because we wanted to avoid perfect multicollinearity and the dummy variable trap.

Part 4: Results and Interpretations

Table 3

Dependent Variable: Weight

Regressor	1	2	3	4
Clength2		0.475*** (0.063)		

Clength		4.262 (4.413)	37.928*** (1.762)	
Volume	0.042*** (0.001)			0.052*** (0.001)
Perch	-35.559 (25.24)	-12.457 (33.96)	31.215 (34.58)	33.169** (14.663)
Whitefish				
Pike	-118.088*** (29.338)	-473.922*** (45.001)	-358.490*** (47.004)	-12.576 (37.522)
Parkki	-97.515*** (25.191)	-6.599 (34.875)	60.964 (43.364)	28.202** (14.1)
Smelt	-80.458*** (26.785)	63.338 (46.012)	287.316*** (50.43)	29.346** (13.683)
Roach	-77.134*** (25.155)	-64.154* (34.117)	-20.782 (39.698)	34.812** (14.99)

Bream	-233.717*** (28.443)	-63.457* (34.203)	-66.309* (35.711)	71.311** (33.549)
PerchVol				-0.005*** (0.002)
WhitefishVol				
ParkkiVol				-0.015*** (0.001)
PikeVol				-0.01*** (0.003)
SmeltVol				-0.022*** (0.002)
RoachVol				-0.011*** (0.002)
BreamVol				-0.021*** (0.002)

Intercept	77.377*** (26.997)	-189.39** (84.752)	-770.552*** (67.47)	-28.398** (13.669)
F-statistic on all coefficients(p-value)	7.4e-129	6.0e-106	9.58e-80	1.5e-149
Adjusted R Squared	0.968569	0.965535	0.930350	0.980581

Column 1 Regression Equation:

Weight = 77.377 + .042(Volume) - 35.559(Perch) - 118.088(Pike) - 97.515(Parkki) -

(26.997) (.001) (25.24) (29.338) (25.191)

80.458(Smelt) - 77.134(Roach) - 233.717(Bream) + u

(26.785) (25.155) (28.443)

<u>Interpretation</u>: A one cm^3 increase in volume is correlated with a 0.042 gram increase in weight

Column 2 Regression Equation:

Weight = -189.39 + .475(Clength2) + 4.262(Clength) - 12.457(Perch) - 473.922(Pike) -

(84.752) (0.063) (4.413) (33.96) (45.001)

6.599(Parkki) + 63.338(Smelt) - 64.154(Roach) - 63.457(Bream) + u

(34.875) (46.012) (34.117) (34.203)

Interpretation: Given a mean Clength of 31.227 cm, a one-centimeter increase from 31.227 cm to 32.227 cm is correlated with a 34.403 gram increase in weight.

Column 3 Regression Equation

Weight = -770.552 + 37.928(Clength) + 31.215(Perch) - 358.490(Pike) + 60.964(Parkki)

(67.47) (1.762) (34.58) (47.004) (43.364)

+ 287.316(Smelt) - 70.512(Roach) - 66.309(Bream) + u

(50.43) (39.698) (37.511)

Interpretation: A one cm increase in Clength is correlated with a 37.928 gram increase in weight.

Column 4 Regression Equation

Weight = -23.398 + 0.052(Volume) + 33.169(Perch) - 12.576(Pike)

(13.669) (0.001) (14.663) (37.522)

\+ 28.202(Parkki) + 29.346(Smelt) + 34.812(Roach) + 71.311(Bream) - 0.005(PerchVol)

(14.01) (13.683) (14.99) (33.549) (0.002)

-0.01(PikeVol) - 0.015(ParkkiVol) - 0.022(SmeltVol) - 0.011(RoachVol) - 0.021(BreamVol) + u

(0.003) (0.001) (0.002) (0.002) (0.002)

<u>Interpretation</u>: A one cm^3 increase in volume is correlated with a 0.052 gram increase in weight

Sample Regression Functional Form:

Weight = + Volume + Perch + Pike + Parkki + Smelt + Roach + Bream + (PerchVol) + (ParkkiVol) + (PikeVol) + (SmeltVol) + (RoachVol) + (BreamVol)

Extra Regression Equation(Top 25% of Clength):

Weight = 1204.27 + 0.61(Clength2) - 31.251(Clength) -733.111(Pike) -

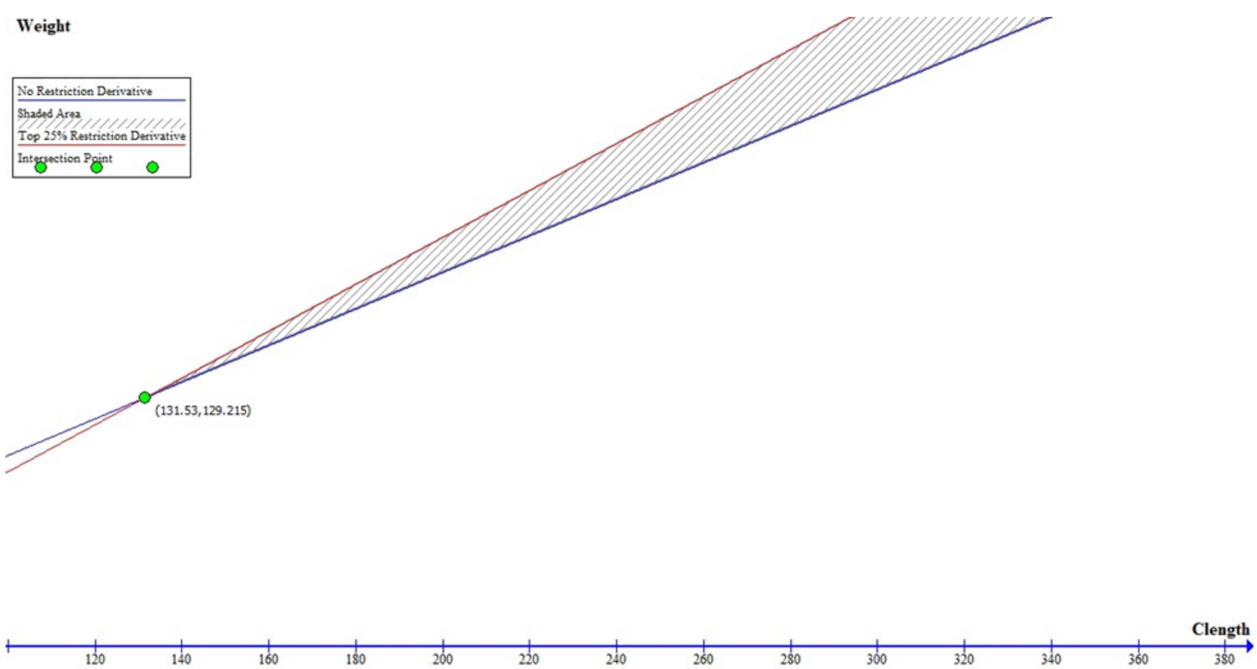
(878.961) (0.392) (34.735) (152.896)

569.165(Bream) + 0.024(PerchVol) + 0.014(PikeVol) + 0.016(BreamVol) + u

(201.404) (0.016) (0.013) (0.008)

<u>Note:</u> This regression equation has the same regressors as the ones in the column 2 regression equation. Also, by taking the derivative using the power rule, we can see that as Clength increases, weight increases at an increasing rate. Also, for any given fish over 131.53 cm or 51.78 inches in Clength, the equation restricting the top 25% of Clength in the sample reports a greater rate of change in weight than the equation using an unrestricted sample. Therefore, a change in Clength, *ceteris paribus*, increases a fish's weight more when the fish is bigger rather than smaller. Figure 5 displays this idea, using software called, "Graph".

Figure 5



From looking at the table, one asterisk means that our coefficient is significant to a 10% significance level, two asterisks mean that our coefficient is significant to a 5% significance level, and three asterisks mean that our coefficient is significant to a 1% significance level. For example, in column 4 of our regression equation, the coefficients on Volume, PerchVol, PikeVol, BreamVol, RoachVol, ParkkiVol, and SmeltVol are significant at a 1% significance level. Moreover, the coefficients on Bream, Roach, Parkki, Perch, and Smelt are significant at the 5% significance level. Lastly, the coefficient on Pike was statistically insignificant. As can be seen in the table, almost all of our coefficients were significant to some degree which is not what we were expecting. On top of that, in all of our regression equations, the p-value for the F-statistic was very close to 0, meaning that it is incredibly unlikely that all of our coefficients were 0. Before running the regression, we predicted that there would be more insignificant regressors because our sample size was so small. This was mentioned previously in one of our flaws. However, the low sample size did not seem to affect the significance of our regressors, especially the ones with a high standard deviation. Strangely, the regressors with the highest standard deviation like *Clength2* or *Volume* had the lowest standard errors which increased the statistical significance of those regression coefficients. In the column 1 regression equation, the only economically insignificant regression coefficient was the one for volume. A .042 gram increase in *Weight* given a one cm^3 increase in *Volume* is not an important result. To make this relationship economically significant, we can say a 200 cm^3 increase in *Volume* is correlated with an 8.4 gram increase in *Weight*. To avoid repetition, in the regression equations in columns 2 through 4, the only economically insignificant regression coefficient was the one for *Clength2*. It makes sense that the coefficients for *Volume* and *Clength2* are economically insignificant because both regressors have very large data points in the sample. Thus, a one-unit increase will not affect a fish's weight that much.

As mentioned previously, we were worried about the applications of our models in commercial fishing because most of the commercial fish species are a lot bigger than the fish used in this sample. This could potentially be a limitation to our research. We are now going to evaluate how well our regression equations predict the weight of two commercial fish.

Species: Wild Salmon, Weight = 1980 grams, Clength = 61.7 cm, Thickness= 11 cm, Height= 13.12 cm

Figure 6

Table 4

Column Equation	Continuous Variables	Actual Weight (g)	Predicted Weight (g)	Residual (g)
1	*Volume*	1980	367.9	1612.1
2	*Clength2, Clength*	1980	1517.8	462.2

3	*Clength*	1980	1298.7	681.3
4	*Volume*	1980	367.9	1612.1

Species: Black Crappie, Weight = 45.36 grams, Clength = 11.43 cm, Thickness= 1.524 cm, Height = 6.35 cm

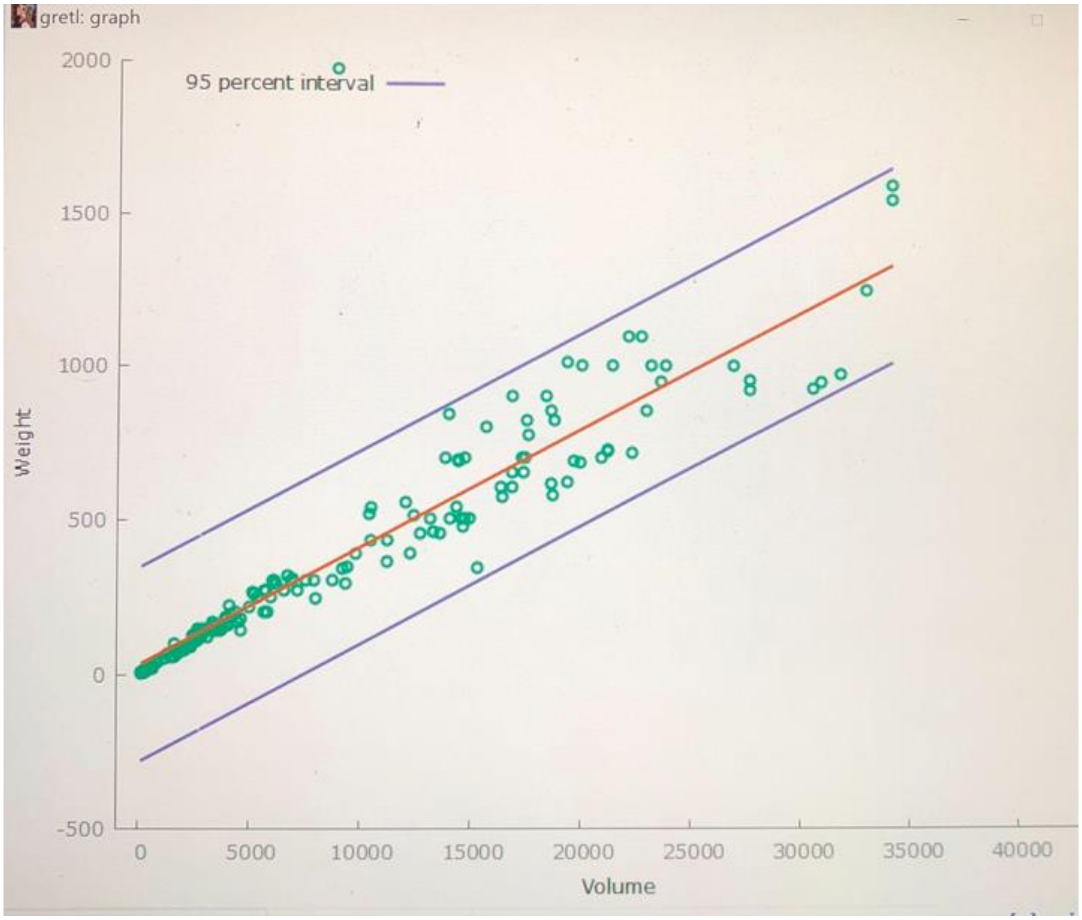
Figure 7

Table 7

Column Equation	Continuous Variables	Actual Weight (g)	Predicted Weight (g)	Residual (g)
1	*Volume*	45.36	23.13	22.23
2	*Clength2, Clength*	45.36	-56.52	101.88
3	*Clength*	45.36	-157.78	203.14
4	*Volume*	45.36	23.13	22.23

Note that since the only difference between the column 1 equation and the column 4 equation is the 6 individual species' volumes, the equations end up being equivalent because we are working with an entirely different species. Now, as mentioned before, some of the biggest issues with our data set are that the fish do not weigh much, and their bodies are very different from almost all commercial fish. Therefore, shown in Table 6 and Table 7 above, the predictions for weight were very far off on average. In Table 6, our models tried to predict the weight of a salmon. Notice that the regression equation with *Clength2* and *Clength* predicted the weight fairly well. The reasoning is that that is the one equation that depicts the quadratic relationship between *Clength* and *Weight*. As a fish grows more in *Clength*, the weight of the fish grows at an increasing rate. Therefore, we were able to get a decent approximation of the weight of the salmon in equation 2. We could have been even more accurate if our data set included some salmon because then we could create a new interaction variable called *Salmon_Clength*. This would take into account how quickly a salmon gains weight by changes in its *Clength*. Also, in Table 1, notice that the models that included volume were so far off in predicting the weight of the salmon. This is because those models do not represent the quadratic relationship between *Clength* and *Weight*. The salmon used for forecasting is about half the height, double the *Clength*, and weighs 1579.15 grams (3.481 lbs.) more than the mean fish in the sample. Note that the salmon is slightly longer in thickness than the mean fish in the sample. Now, the problem arising is that salmon and the average fish in the sample are not even a half a standard deviation apart in volume. However, they are about 4.41 standard deviations apart in weight. Figure 8 is the graph of the forecast with a 95% interval:

Figure 8



Moving on to Table 2, we used a Black Crappie. Based on the images, the difference in the bodies is clearly shown. The Black Crappie a shorter in *Clength*, taller in height, and weigh less than Salmon. Therefore, fish like the Black Crappie are very similar to the fish in the sample. Conversely from the Table 1 example, regression equations that include volume predict weight better. While the prediction is still inaccurate, the reason volume predicts weight better is because a Black Crappie's *Clength* is small. Because the *Clength* is small, the quadratic relationship does not impact the error of the prediction. Therefore, since Volume = Clength*Height*Thickness, not having a squared *Clength* will not cause issues in this case. However, since our prediction was still not accurate when using volume, it is important to analyze why. Black Crappies have a similar body to the fish in the sample. Although, the fish in the sample are a more extreme version. Figure 9 shows what a Bream looks like:

Figure 9

From this, looking at the regression line relating *Volume* to *Weight*, for a weight prediction to be accurate, the ordered pair of volume and weight would have to have a nearly linear relationship with the ordered pair, (9793.26 cm^3, 390 grams). That is approximately a fish having dimensions

Clength = 35 cm, Thickness = 9.485 cm, and Height = 29.5 cm. Therefore, unless the fish being forecasted has very similar dimensions (similar *Height* and *Clength*), a regression equation with volume will not be very accurate. The reason equations in column 2 and column 3 predict the weight of a Black Crappie the worst is because the slopes in those equations are so large, and the y-intercepts are so small. Therefore, for small values of *Clength*, there will be a negative output.

Part 5: Conclusion

Our research project tests many different measurement variables and their forms to see which ones have the highest adjusted R squared. In addition, we used seven different species binary variables: Pike, Parkki, Smelt, Roach, Bream, Perch, and Whitefish. This, along with creating interaction variables with each individual species and their volume and ultimately predict the fish's weight the best. These fish are all very small, so while our data are accurate for these types of small fish, it may not necessarily apply to all fish, especially larger fish. For scale, the biggest fish in our data set was 1200 grams or around two pounds. However, fish can get up to hundreds of pounds.

We used four different regression equations, all with the dependent variable of weight, just including different regressors in which we found that this method of estimating the weight of a fish was most effective with fish shaped closer to a box and not as effective for normal shaped fish. To avoid perfect multicollinearity for each regression we removed whitefish so not all the binary variables were included. First, we used the independent variables of volume and each of the different fish: *Perch, Bream, Parkki, Pike, Smelt* and *Roach*. Second, we used the independent variables of length and each of the different fish: *Perch, Bream, Parkki, Pike, Smelt* and *Roach*. Third, we looked at just the binary variables (whether it is that type of fish or not) for our six fish types: *Perch, Bream, Parkki, Pike, Smelt,* and *Roach* giving the average for each. Then, the fourth regression includes all of the binary variables for the different fish: *Perch, Bream, Parkki, Pike, Smelt,* and *Roach*, and then also the different volumes for those fish labeled as *PerchVol, BreamVol, ParkkiVol, SmeltVol,* and *RoachVol*.

Our study looks at different types of fish and uses their cross-length and volume to find correlations among these fish. The dependent variable that we are using is weight which is also a comparison that we are using among the different fish. The result of the data is that there is linearity, where the adjusted R squared, is 0.98. Therefore, coefficients and standard errors of the regression are reliable. For this first regressor, using volume, there was a correlation of 0.97 between the volume and the different types of fish. Then, for just the length, not length squared, and the different types of fish, there was a correlation of 0.94.

Works Cited

Hudson, S. (2022, March 27). *A guide to lip grippers*. Sport Fishing Mag. Retrieved April 5, 2022, from https://www.sportfishingmag.com/guide-to-lip-grippers/

Kaggle.com. 2022. *Kaggle: Your Machine Learning and Data Science Community*. [online] Available at: <https://www.kaggle.com/> [Accessed 25 April 2022].

Keretchashvili, G. (2021, December 20). *Fish Weight Prediction (Regression Analysis for beginners) — Part 1*. Medium. https://towardsdatascience.com/fish-weight-prediction-regression-analysis-for-beginners-part-1-8e43b0cb07e

Unsplash. (n.d.). Beautiful Free Images & Pictures | Unsplash. https://unsplash.com/