Home Work On Data Analysis 1

Md Shamsuzzaman

10.11.2017

1 Executive Summary

This study has been carried out to find out linear relationship among the weight, length and width of 12 Perch fishes which were caught from a lake in Finland. The statistical analysis has been performed through linear model for response variable weight and other predictor variables. Before the building the regression model, the variables have been analysed through the graphical analysis and correlation. The assumptions of linearity according to the linear regression analysis by checking linearity through scatter plot, normality through Q-Q plot, and constant variance through residual plot have been checked. In this analysis, the linear relationship has been performed firstly, between the length and width in relation to the weight. After that, a transformation has been performed for length and width respectively by using square, and these transformed length and width has been analysed against the weight. Finally, a new variable has been formulated as the product of variable length and width, and it has been analysed for in relation to weight. Along this linear regression test, a correlation test (Pearson correlation test) has been performed to double check the relationships revealed by the linear models. Adjusted R square value was one of the main important indicators to identify the relationship along with the scatter plot. According to the performed tests, it has been found that weight prediction in relation to new variable is the best model which shows stronger relationship (99 *percent*) than other models.

2 Introduction

A linear regression analysis has been applied on this data set to find out the best model. For getting the best model, the predictor variables have been transformed by using square.

3 Data Collection

This data set is consisted of weight, length and width of 12 Perch fishes which were caught in a lake of Finland.

4 Data Analysis and Summary

Before the building the regression model, we can analyze and understand the variables through the graphical analysis and correlation study below. In this exercise, we will try to build a simple regression model that we can use to predict weight of Perch by establishing a significant linear relationship with length and width of the Perch. Before going to the syntax, we have to check the assumptions of linearity according to the linear regression analysis by checking linearity through scatter plot, normality through Q-Q plot, and constant variance through residual plot.

4.1 Answer of question No. a)

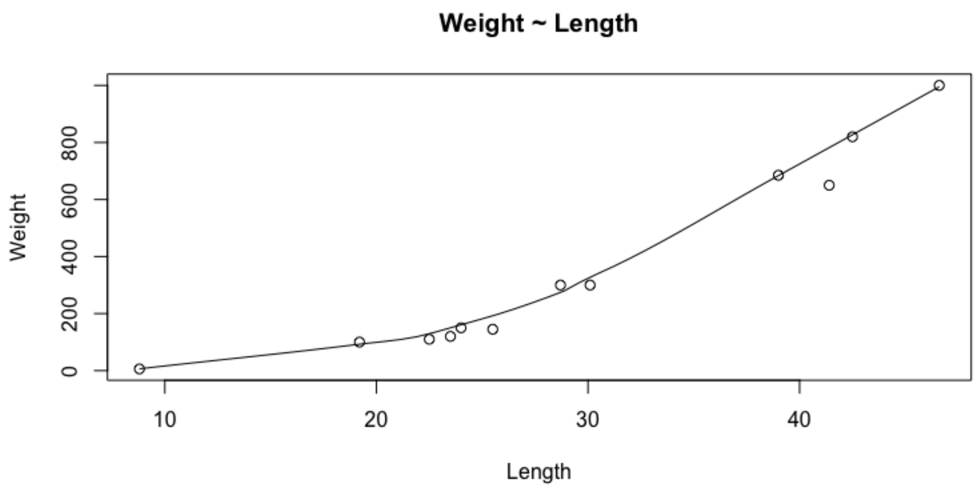

Fig 1. Scatter plot of Length and Weight

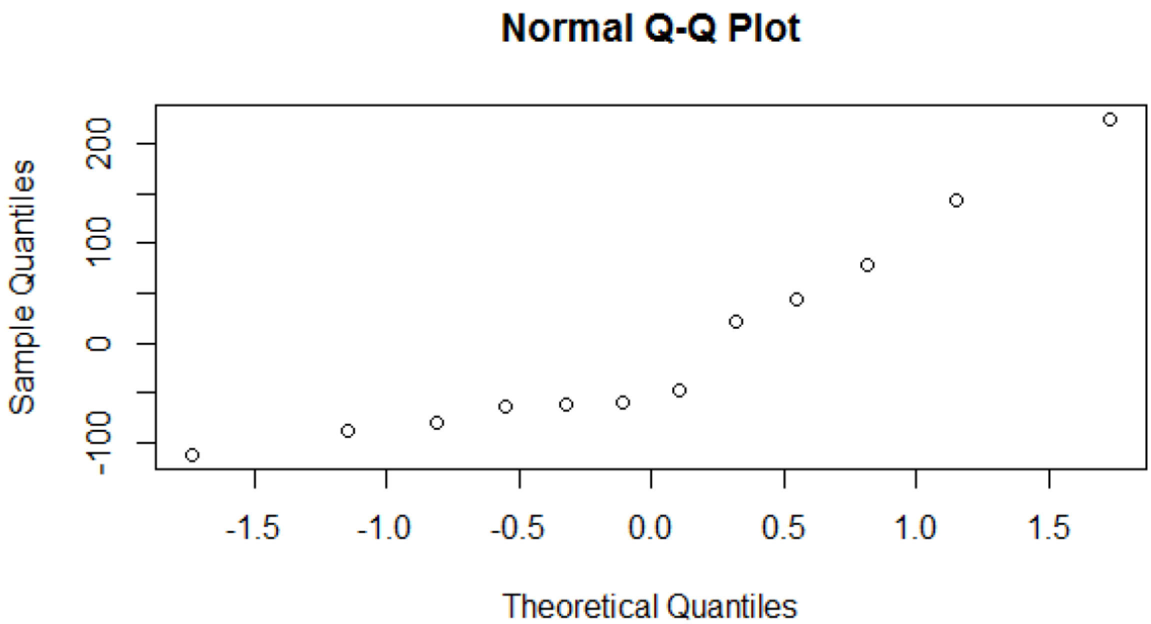

Fig 2. Q-Q plot of Length and Weight

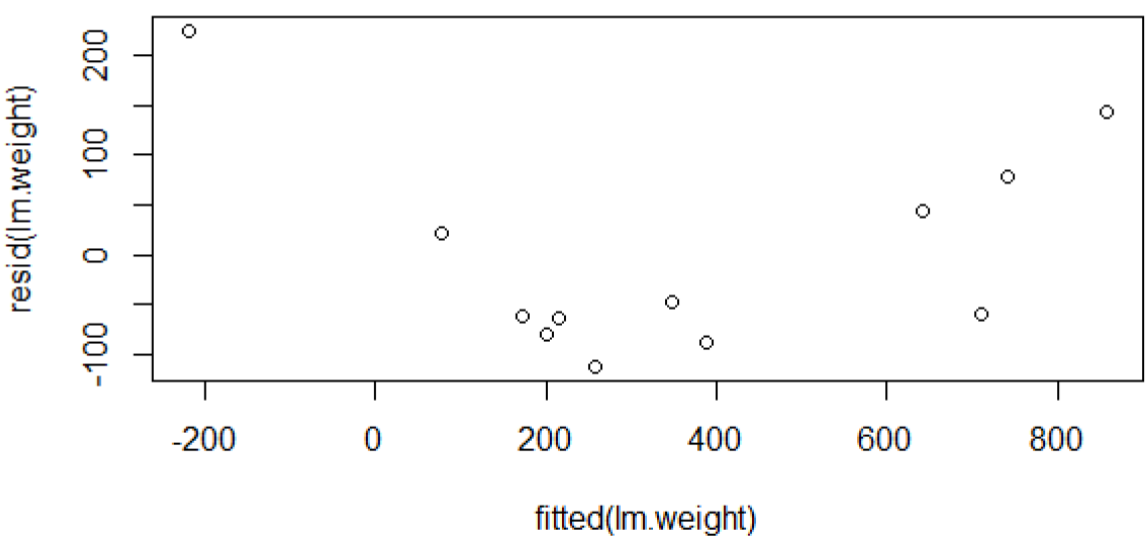

Fig 3. Residual plot of Length and Weight

From the figures (1- 3) above, scatter plot shows the linearity between length and weight while QQ plot shows normality. On the other hand, residual plot indicates the constant variance of them. We have seen the linear relationship between weight and length by computing the correlation. From the built linear model for Weight and Length, we also have established the relationship between the predictor and response in the form of a mathematical formula for Weight as a function for Length. For the above output, you can notice the ?Coefficients? part having two components: Intercept: -468.91, Length: 28.46. These are also called the beta coefficients. So far, we have built the linear model and we can predict the weight value if a corresponding value of length is known. Now, before using a regression model, we have to check the statistical significance of this model. Now, we can do it by printing the summary statistics for linearMod.

```
Call:
lm(formula = Weight ~ Length, data = perch)

Residuals:
    Min      1Q  Median      3Q     Max
-111.86  -68.11  -53.67   52.75  224.35

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -468.914     92.547  -5.067 0.000487 ***
Length        28.462      2.967   9.592 2.33e-06 ***
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

Residual standard error: 109.4 on 10 degrees of freedom
Multiple R-squared:  0.902,Adjusted R-squared:  0.8922
F-statistic:    92 on 1 and 10 DF,  p-value: 2.326e-06

Pearson?s product-moment correlation
data:  length and weight
t = 9.5919, df = 10, p-value = 2.326e-06
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
0.8260526 0.9861332
sample estimates:
cor
0.9497184
```

Here, we have the null hypothesis: The coefficients associated with the variables is equal to zero and the alternate hypothesis: The coefficients are not equal to zero. The p-value

(2.326e-06) which is less than pre-determined p-value 0.05. So, the null hypothesis is rejected and alternates are accepted. It means, there is the existence of relationship between weight and length of the Perch. Here we can see, the multiple R-squared: 0.902 and Adjusted R-squared: 0.8922. So, it means that this model can explain 90 percent variation in weight. It can be concluded that this is the good model to predict weight in respect to length. Here, also the Pearson correlation test shows a strong relationship between them.

4.2 Answer of question No. b)

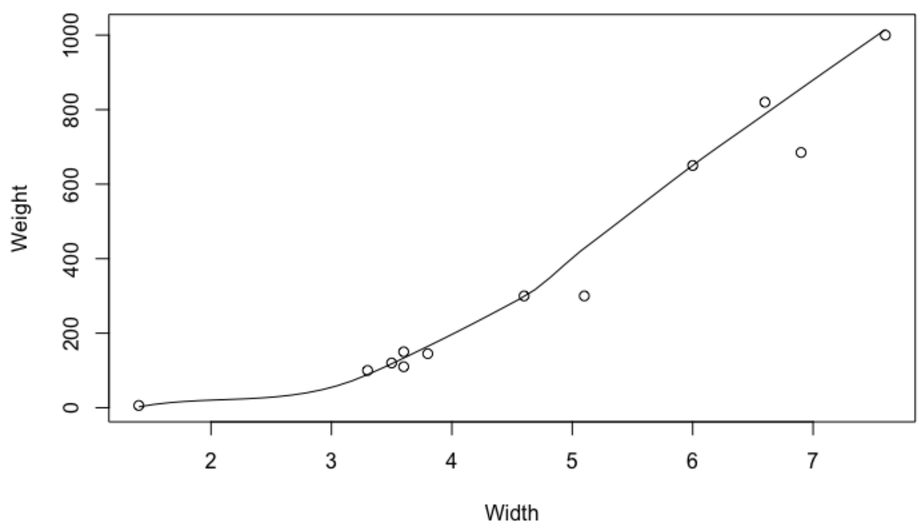

Fig 4.

Scatter plot of Width and Weight

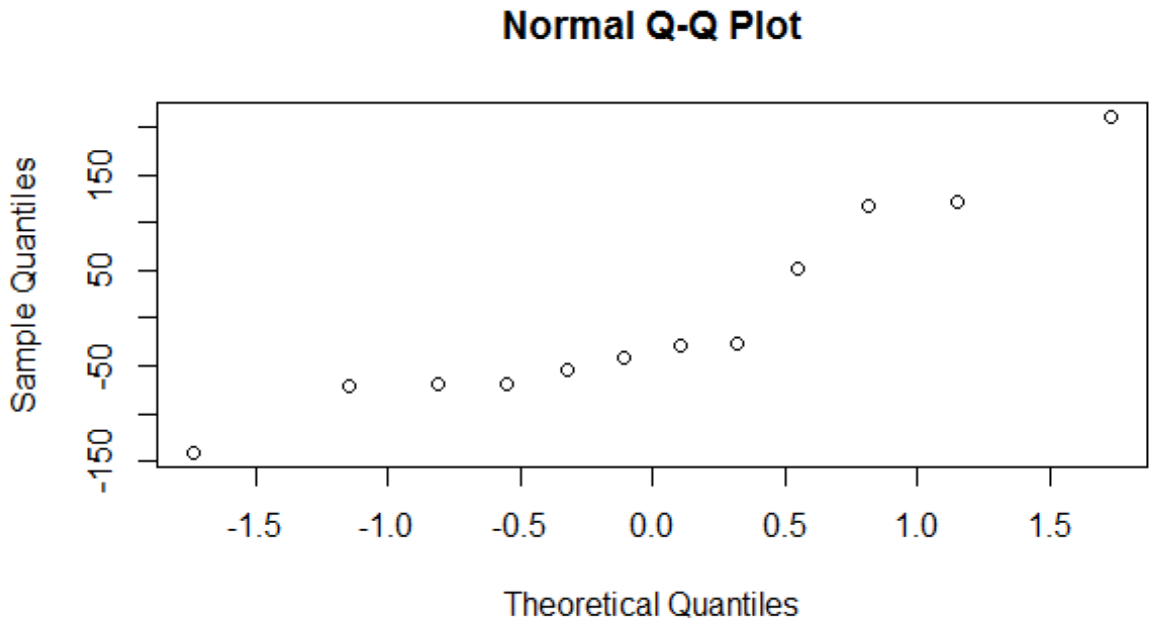

Fig 5. Q-Q plot of Width and Weight

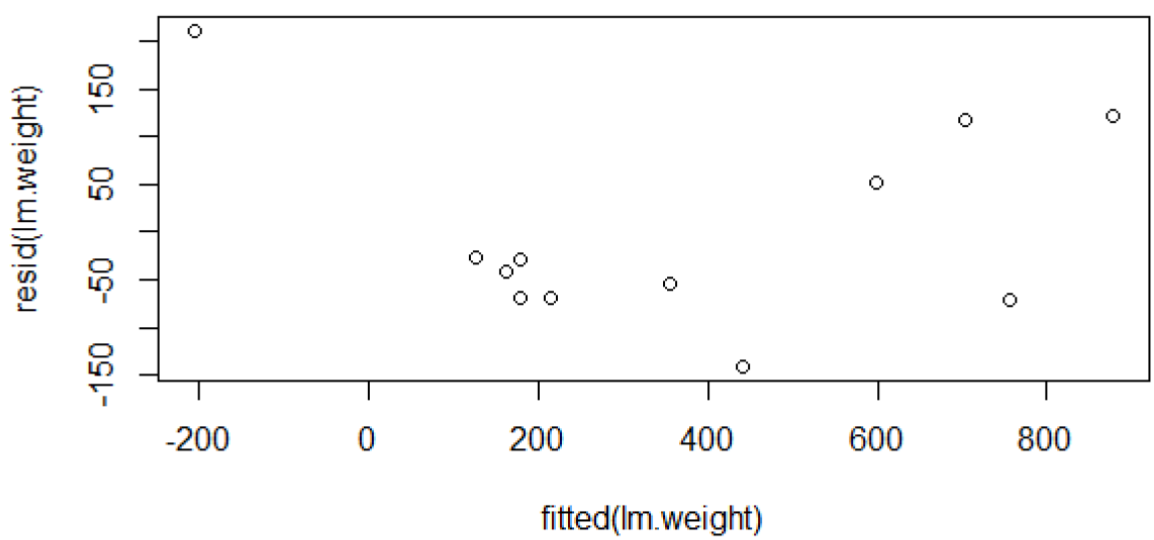

Fig 6. Residual plot of Width and Weight

From the figures (4 - 6) above, scatter plot shows the linearity between width and weight while QQ plot shows normality. On the other hand, residual plot indicates the constant variance of them.

```
> summary(linearMod2)

Call:
lm(formula = Weight ~ Width, data = perch)

Residuals:
    Min      1Q  Median      3Q     Max
-141.16  -69.17  -35.49   67.97  210.86

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -449.44      89.27  -5.034 0.000511 ***
Width         174.63      17.93   9.741 2.02e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 107.9 on 10 degrees of freedom
Multiple R-squared:  0.9047,Adjusted R-squared:  0.8951
F-statistic: 94.88 on 1 and 10 DF,  p-value: 2.021e-06

Pearson?s product-moment correlation
   data:  width and weight
   t = 9.7408, df = 10, p-value = 2.021e-06
   alternative hypothesis: true correlation is not equal to 0
   95 percent confidence interval:
   0.8306467 0.9865306
   sample estimates:
   cor
   0.9511338
```

Here, we have the null hypothesis: The coefficients associated with the variables is equal to zero and the alternate hypothesis: The coefficients are not equal to zero. The p-value (2.021e-06) which is less than pre-determined p-value 0.05. So, the null hypothesis is rejected and alternates are accepted. It means, there is the existence of relationship between width and weight of the Perch. Here we can see, the multiple R-squared: 0.9047 and Adjusted R-squared: 0.8951. So, it means that this model can explain about 90 percent variation in weight. It can be concluded that this linear relationship refers it a good model to predict weight in respect to width. Here, also the Pearson correlation test shows a strong relationship between them by indicating that it is more than of length and width.

4.3 Answer of question No. c)

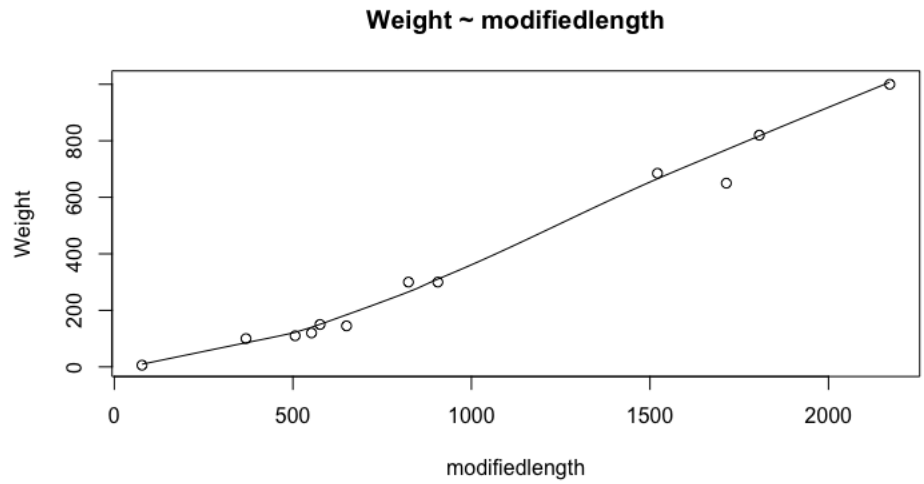

Fig 7. Scatter plot of modified length and weight

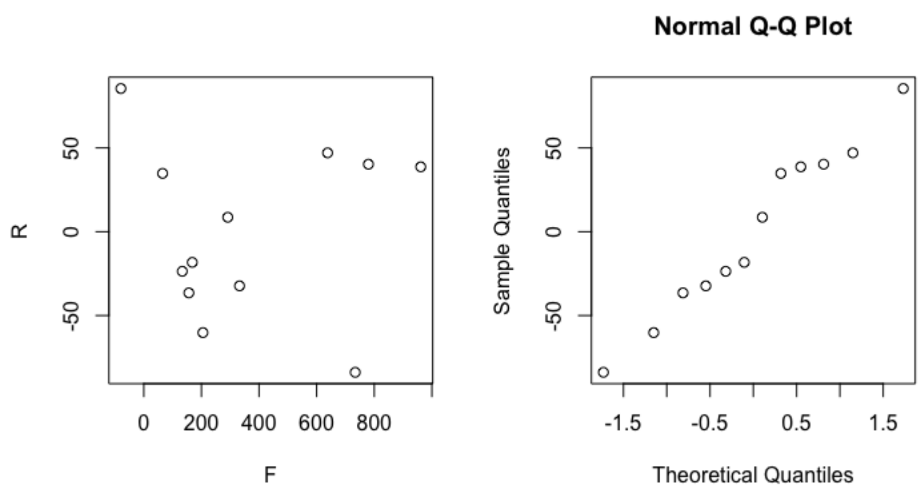

Fig 8. Residual plot (left) and Q-Q plot (right) of Weight vs Modifiedlength

Here, the length data has been transformed by using square. From the figures (7 - 8) above, scatter plot shows the linearity between transformed length and weight while QQ plot shows normality. On the other hand, residual plot indicates the constant variance of them.

```
> lm(weight~modifiedlength)
Call:
lm(formula = weight ~ modifiedlength)
Coefficients:
(Intercept)    modifiedlength
-117.991        0.497
> model.1=lm(weight~modifiedlength)
> summary(model.1)
Call:
lm(formula = weight ~ modifiedlength)
Residuals:
Min      1Q  Median      3Q     Max
-83.871 -33.353  -4.842  39.086  85.402
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -117.991     27.877  -4.233  0.00174 **
modifiedlength      0.497      0.024  20.707 1.53e-09 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
Residual standard error: 52.76 on 10 degrees of freedom
Multiple R-squared:  0.9772,Adjusted R-squared:  0.9749
F-statistic: 428.8 on 1 and 10 DF,  p-value: 1.528e-09

Pearson?s product-moment correlation
data:  modifiedlength and weight
t = 20.707, df = 10, p-value = 1.528e-09
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
0.9583087 0.9968841
sample estimates:
cor
0.9885388
```

Here, we have the null hypothesis: The coefficients associated with the variables is equal to zero and the alternate hypothesis: The coefficients are not equal to zero. The p-value (1.528e-09) which is less than pre-determined p-value 0.05. So, the null hypothesis is rejected and alternates are accepted. It means, there is the existence of relationship between transformed length and weight of the Perch. Here we can see, the multiple R-squared:

0.9772 and Adjusted R-squared: 0.9749. So, it means that this model can explain 95 percent variation in weight. It can be concluded that this linear relationship refers it a good model to predict weight in respect to transformed/modified length. Here, also the Pearson correlation test shows a strong relationship (97 percent) between them by indicating that it is more than of previous models.

4.4 Answer of question No. d)

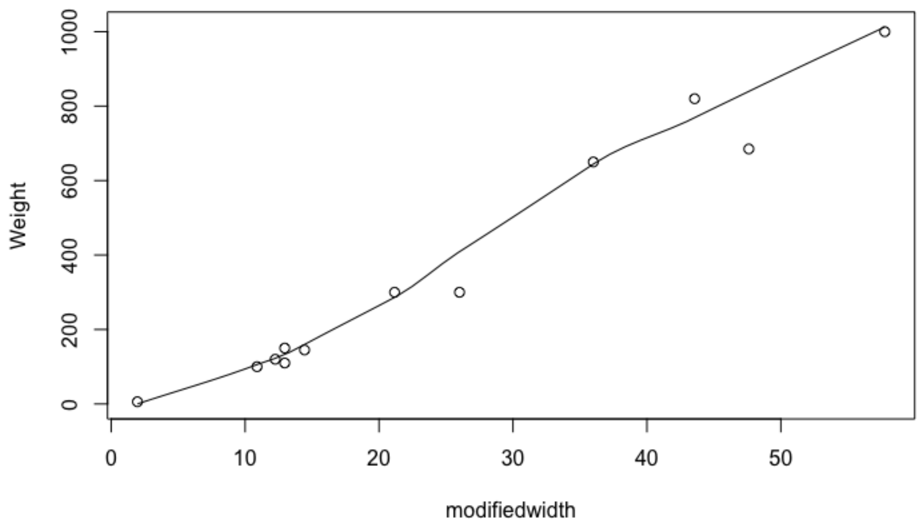

Fig 9: Scatter plot of Weight and Modifiedwidth

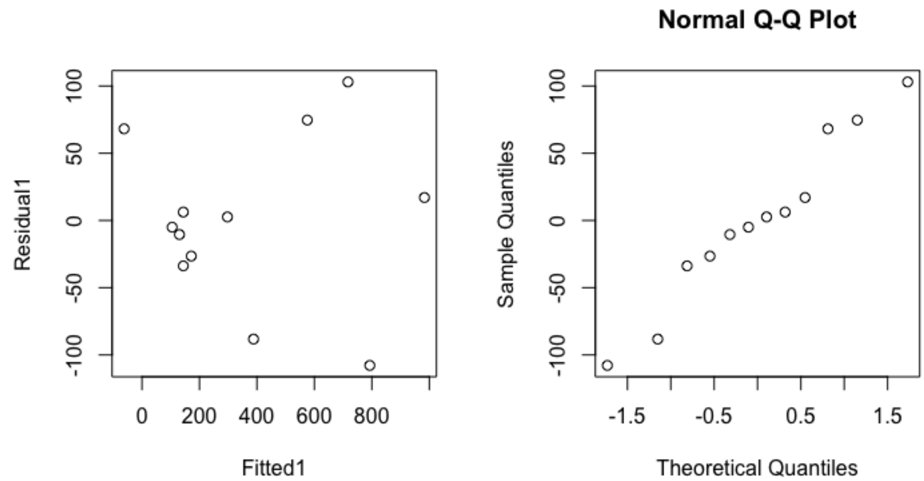

Fig 10: Residual plot (left) and Q-Q plot (right) of Weight vs Modifiedlength

Here, the width data has been transformed by using square. From the figures (9 - 10) above, scatter plot shows the linearity between transformed modified width while QQ plot shows normality. On the other hand, residual plot indicates the constant variance of them.

```
> lm(weight~modifiedwidth)
Call:
lm(formula = weight ~ modifiedwidth)
Coefficients:
(Intercept)     modifiedwidth
-98.99        18.73
> model.1=lm(weight~modifiedwidth)
> summary(model.1)
Call:
lm(formula = weight ~ modifiedwidth)
Residuals:
Min       1Q   Median       3Q      Max
-107.821  -28.314   -1.185   29.834  103.042
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  -98.989     33.671   -2.94   0.0148 *
modifiedwidth     18.732      1.126   16.64 1.28e-08 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
Residual standard error: 65.24 on 10 degrees of freedom
```

```
Multiple R-squared:  0.9651,Adjusted R-squared:  0.9617
F-statistic: 276.9 on 1 and 10 DF,  p-value: 1.284e-08

> cor.test(newwidth,weight)
   Pearson?s product-moment correlation
   data:  newwidth and weight
   t = 16.641, df = 10, p-value = 1.284e-08
   alternative hypothesis: true correlation is not equal to 0
   95 percent confidence interval:
   0.9365643 0.9952098
   sample estimates:
   cor
   0.9824196
```

Here, we have the null hypothesis: The coefficients associated with the variables is equal to zero and the alternate hypothesis: The coefficients are not equal to zero. The p-value (1.284e-09) which is less than pre-determined p-value 0.05. So, the null hypothesis is rejected and alternates are accepted. It means, there is the existence of relationship between transformed modified width and weight of the Perch. Here we can see, the multiple R-squared: 0.9651 and Adjusted R-squared: 0.9617. So,it means that this model can explain 96 percent variation in weight. It can be concluded that this linear relationship refers it a good model to predict weight in respect to transformed/modified width. Here, also the Pearson correlation test shows a strong relationship (98 percent) between them.

4.5 Answer of question No. e)

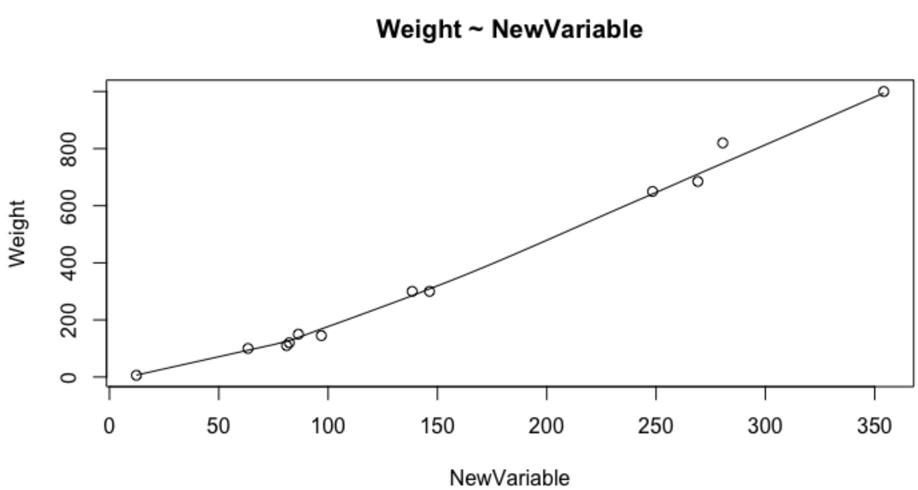

Fig 11.

Scatter plot of newvariable and width

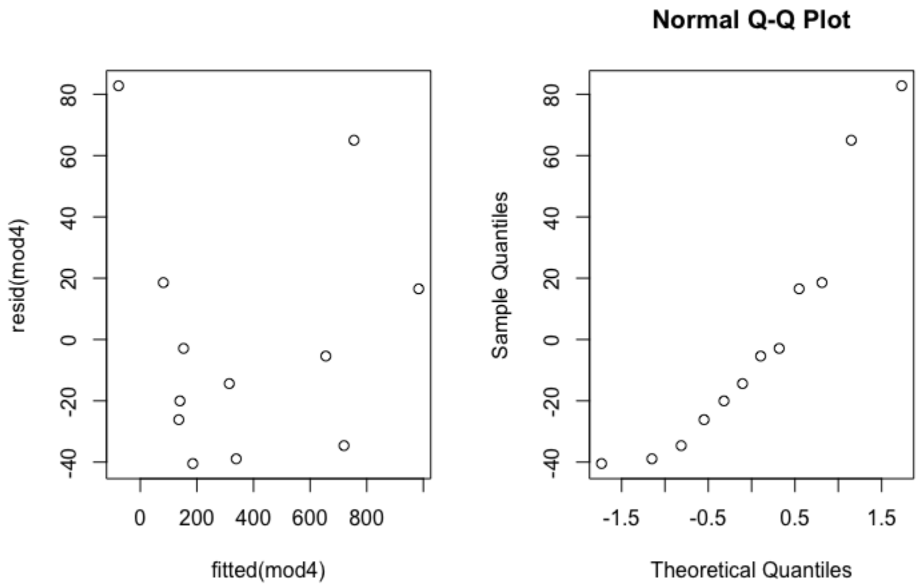

Fig 12.
Residual plot (left) and Q-Q plot (right) of newvariable and width

Here, a new variable has been created by multiplying length and width. From the figures (9 - 10) above, scatter plot shows the linearity between transformed modified width while QQ plot shows normality. On the other hand, residual plot indicates the constant variance of them.

```
> lm(weight~newvariable)
Call:
lm(formula = weight ~ newvariable)
Coefficients:
(Intercept)  newvariable
-115.102            3.102
> model.1=lm(weight~newvariable)
> summary(model.1)
Call:
lm(formula = weight ~newvariable)
Residuals:
Min     1Q Median     3Q    Max
-40.47 -28.27  -9.90  17.04  82.79
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)     -115.1021    21.8703  -5.263 0.000367 ***
newvariable    3.1019     0.1179  26.318 1.45e-10 ***
```

```
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
Residual standard error: 41.69 on 10 degrees of freedom
Multiple R-squared:  0.9858,Adjusted R-squared:  0.9843
F-statistic: 692.6 on 1 and 10 DF,  p-value: 1.446e-10

Pearson?s product-moment correlation
data:  newvariable and weight
t = 26.318, df = 10, p-value = 1.446e-10
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
0.9738714 0.9980615
sample estimates:
cor
0.9928582
```

Here, we have the null hypothesis: The coefficients associated with the variables is equal to zero and the alternate hypothesis: The coefficients are not equal to zero. The p-value (1.446e-10) which is less than pre-determined p-value 0.05. So, the null hypothesis is rejected and alternates are accepted. It means, there is the existence of relationship between new variable and weight of the Perch. Here we can see, the multiple R-squared: 0.9858 and Adjusted R-squared: 0.9843. it means that this model can explain 98 percent variation in weight. It can be concluded that this linear relationship refers it a good model to predict weight in respect to a new variable. Here, also the Pearson correlation test shows a strong relationship (99 percent) between them which more than other models.

4.6 Answer of question No. f)

The findings of the models are given below.

```
Length and weight
p-value = 2.326e-06
Adjusted R-squared:  0.8922.
Pearson correlation: .9497

width vs weight
p-value = 2.021e-06
Adjusted R-squared:  0.8951
Pearson correlation: .9511

length2 vs weight
p-value =1.528e-09
Adjusted R-squared:  0.9749
Pearson correlation: .9895
```

```
 width2 vs weight
p-value = 1.284e-09
Adjusted R-squared:  0.9617
Pearson correlation: .9824

 length*width
p-value =1.446e-10
Adjusted R-squared:  0.9843
Pearson correlation: .9928
```

From the Adjusted r and p values, it has been found that weight prediction in relation to new variable is the best model which shows stronger relationship (99 *percent*) than other models.

5 Conclusion

The analysis indicates that weight prediction in relation to new variable is the best model which shows stronger relationships (99 *percent*) than other models. Here through this regression analysis, I have learnt to conduct the test for response variable and predictor variables to justify the relationship among them.

6 Appendix

6.1 R Code

```
attach(perch)
View(perch)
str(perch)
head(perch)
tail(perch)
summary(perch)
plot(Length,Weight)
abline(lm(Weight~Length))
lm(Weight~Length)
model.1=lm(Weight~Length)
summary(model.1)
lm.Weight <-lm(Weight~Length)
fitted(lm.Weight)
resid(lm.Weight)
plot(fitted(lm.Weight),resid(lm.Weight))
qqnorm(resid(lm.Weight))
```

```
cor.test(Length,Weight)
plot(Width,Weight)
abline(lm(Weight~Width))
lm(Weight~Width)
model.2=lm(Weight~Width)
summary(model.2)
lm.Weight1 <-lm(Weight~Width)
fitted(lm.Weight1)
resid(lm.Weight1)
plot(fitted(lm.Weight1),resid(lm.Weight1))
qqnorm(resid(lm.Weight1))
cor.test(Width,Weight)
nLength <-Length^2
plot(nLength,Weight)
abline(lm(Weight~nLength))
lm(Weight~nLength)
model.3=lm(Weight~nLength)
summary(model.3)
lm.Weight2 <-lm(Weight~nLength)
fitted(lm.Weight2)
resid(lm.Weight2)
plot(fitted(lm.Weight2),resid(lm.Weight2))
qqnorm(resid(lm.Weight2))
cor.test(nLength,Weight)
nWidth <-Width^2
plot(nWidth,Weight)
abline(lm(Weight~nWidth))
lm(Weight~nWidth)
model.4=lm(Weight~nWidth)
summary(model.4)
lm.Weight3 <-lm(Weight~nWidth)
fitted(lm.Weight3)
a<-resid(lm.Weight3)
plot(fitted(lm.Weight3),resid(lm.Weight3))
qqnorm(resid(lm.Weight3))
cor.test(nWidth,Weight)
nLength_nWidth <-Length*Width
print(nLength_nWidth)
plot(nLength_nWidth,Weight)
abline(lm(Weight~nLength_nWidth))
lm(Weight~nLength_nWidth)
```

```
model.5=lm(Weight~nLength_nWidth)
summary(model.5)
lm.Weight4 <-lm(Weight~nLength_nWidth)
fitted(lm.Weight4)
resid(lm.Weight4)
plot(fitted(lm.Weight4),resid(lm.Weight4))
qqnorm(resid(lm.Weight4))
cor.test(nLength_nWidth,Weight)
```

6.2 Log File

```
> attach(perch)
> View(perch)
> str(perch)
Classes 'tbl_df', 'tbl' and 'data.frame': 12 obs. of  3 variables:
 $ Weight: num  5.9 300 100 300 110 685 120 650 150 820 ...
 $ Length: num  8.8 28.7 19.2 30.1 22.5 39 23.5 41.4 24 42.5 ...
 $ Width : num  1.4 5.1 3.3 4.6 3.6 6.9 3.5 6 3.6 6.6 ...
> head(perch)
  Weight Length Width
1    5.9    8.8   1.4
2  300.0   28.7   5.1
3  100.0   19.2   3.3
4  300.0   30.1   4.6
5  110.0   22.5   3.6
6  685.0   39.0   6.9
> tail(perch)
   Weight Length Width
7     120   23.5   3.5
8     650   41.4   6.0
9     150   24.0   3.6
10    820   42.5   6.6
11    145   25.5   3.8
12   1000   46.6   7.6
> summary(perch)
     Weight           Length          Width      
 Min.   :   5.9   Min.   : 8.80   Min.   :1.400  
 1st Qu.: 117.5   1st Qu.:23.25   1st Qu.:3.575  
 Median : 225.0   Median :27.10   Median :4.200  
 Mean   : 365.5   Mean   :29.32   Mean   :4.667  
 3rd Qu.: 658.8   3rd Qu.:39.60   3rd Qu.:6.150  
```

```
 Max.   :1000.0   Max.   :46.60   Max.   :7.600
> plot(Length,Weight)
> abline(lm(Weight~Length))
> lm(Weight~Length)

Call:
lm(formula = Weight ~ Length)

Coefficients:
(Intercept)       Length
    -468.91        28.46

> model.1=lm(Weight~Length)
> summary(model.1)

Call:
lm(formula = Weight ~ Length)

Residuals:
    Min      1Q  Median      3Q     Max
-111.86  -68.11  -53.67   52.75  224.35

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -468.914     92.547  -5.067 0.000487 ***
Length        28.462      2.967   9.592 2.33e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 109.4 on 10 degrees of freedom
Multiple R-squared:  0.902,Adjusted R-squared:  0.8922
F-statistic:    92 on 1 and 10 DF,  p-value: 2.326e-06

> lm.Weight <-lm(Weight~Length)
> fitted(lm.Weight)
         1          2          3          4          5          6          7          8
-218.45016  347.94021   77.55285  387.78676  171.47688  641.09703  199.93870  709.40541
         9         10         11         12
 214.16962  740.71342  256.86236  857.40691
> resid(lm.Weight)
         1          2          3          4          5          6          7          8
```

```
 224.35016  -47.94021   22.44715  -87.78676  -61.47688   43.90297  -79.93870  -59.40541
         9         10         11         12
 -64.16962   79.28658 -111.86236  142.59309
> plot(fitted(lm.Weight),resid(lm.Weight))
> qqnorm(resid(lm.Weight))
> cor.test(Length,Weight)

Pearson's product-moment correlation

data:  Length and Weight
t = 9.5919, df = 10, p-value = 2.326e-06
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8260526 0.9861332
sample estimates:
      cor
0.9497184

> plot(Width,Weight)
> abline(lm(Weight~Width))
> lm(Weight~Width)

Call:
lm(formula = Weight ~ Width)

Coefficients:
(Intercept)        Width
     -449.4        174.6

> model.2=lm(Weight~Width)
> summary(model.2)

Call:
lm(formula = Weight ~ Width)

Residuals:
    Min      1Q  Median      3Q     Max
-141.16  -69.17  -35.49   67.97  210.86

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept)  -449.44      89.27  -5.034 0.000511 ***
Width         174.63      17.93   9.741 2.02e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 107.9 on 10 degrees of freedom
Multiple R-squared:  0.9047,Adjusted R-squared:  0.8951
F-statistic: 94.88 on 1 and 10 DF,  p-value: 2.021e-06

> lm.Weight1 <-lm(Weight~Width)
> fitted(lm.Weight1)
        1         2         3         4         5         6         7         8
-204.9631  441.1642  126.8320  353.8497  179.2207  755.4964  161.7578  598.3303
        9        10        11        12
 179.2207  703.1077  214.1465  877.7367
> resid(lm.Weight1)
         1          2          3          4          5          6          7          8
 210.86308 -141.16423  -26.83203  -53.84973  -69.22073  -70.49644  -41.75783   51.66966
         9         10         11         12
 -29.22073  116.89226  -69.14653  122.26326
> plot(fitted(lm.Weight1),resid(lm.Weight1))
> qqnorm(resid(lm.Weight1))
> cor.test(Width,Weight)

Pearson's product-moment correlation

data:  Width and Weight
t = 9.7408, df = 10, p-value = 2.021e-06
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8306467 0.9865306
sample estimates:
      cor
0.9511338

> nLength <-Length^2
> plot(nLength,Weight)
> abline(lm(Weight~nLength))
> lm(Weight~nLength)

Call:
```

```
lm(formula = Weight ~ nLength)

Coefficients:
(Intercept)      nLength
   -117.991        0.497

> model.3=lm(Weight~nLength)
> summary(model.3)

Call:
lm(formula = Weight ~ nLength)

Residuals:
    Min      1Q  Median      3Q     Max
-83.871 -33.353  -4.842  39.086  85.402

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -117.991     27.877  -4.233  0.00174 **
nLength        0.497      0.024  20.707 1.53e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 52.76 on 10 degrees of freedom
Multiple R-squared:  0.9772,Adjusted R-squared:  0.9749
F-statistic: 428.8 on 1 and 10 DF,  p-value: 1.528e-09

> lm.Weight2 <-lm(Weight~nLength)
> fitted(lm.Weight2)
        1         2         3         4         5         6         7         8
-79.50240 291.39433  65.22809 332.30852 133.62219 637.96719 156.48484 733.87102
        9        10        11        12
168.28892 779.74044 205.19221 961.30464
> resid(lm.Weight2)
         1          2          3          4          5          6          7          8
 85.402398   8.605671  34.771910 -32.308524 -23.622192  47.032812 -36.484838 -83.871017
         9         10         11         12
-18.288921  40.259557 -60.192214  38.695358
> plot(fitted(lm.Weight2),resid(lm.Weight2))
> qqnorm(resid(lm.Weight2))
> cor.test(nLength,Weight)
```

```
Pearson's product-moment correlation

data:  nLength and Weight
t = 20.707, df = 10, p-value = 1.528e-09
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9583087 0.9968841
sample estimates:
      cor
0.9885388

> nWidth <-Width^2
> plot(nWidth,Weight)
> abline(lm(Weight~nWidth))
> lm(Weight~nWidth)

Call:
lm(formula = Weight ~ nWidth)

Coefficients:
(Intercept)       nWidth
     -98.99        18.73

> model.4=lm(Weight~nWidth)
> summary(model.4)

Call:
lm(formula = Weight ~ nWidth)

Residuals:
     Min       1Q   Median       3Q      Max
-107.821  -28.314   -1.185   29.834  103.042

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -98.989     33.671   -2.94   0.0148 *
nWidth        18.732      1.126   16.64 1.28e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 65.24 on 10 degrees of freedom
Multiple R-squared:  0.9651,Adjusted R-squared:  0.9617
F-statistic: 276.9 on 1 and 10 DF,  p-value: 1.284e-08

> lm.Weight3 <-lm(Weight~nWidth)
> fitted(lm.Weight3)
        1         2         3         4         5         6         7         8
-62.27487 388.21930 104.99802 297.37120 143.77236 792.82113 130.47295 575.34765
        9        10        11        12
143.77236 716.95829 171.49508 982.94652
> a<-resid(lm.Weight3)
> plot(fitted(lm.Weight3),resid(lm.Weight3))
> qqnorm(resid(lm.Weight3))
> cor.test(nWidth,Weight)

Pearson's product-moment correlation

data:  nWidth and Weight
t = 16.641, df = 10, p-value = 1.284e-08
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9365643 0.9952098
sample estimates:
      cor
0.9824196

> nLength_nWidth <-Length*Width
> print(nLength_nWidth)
 [1]  12.32 146.37  63.36 138.46  81.00 269.10  82.25 248.40  86.40 280.50  96.90 354.16
> plot(nLength_nWidth,Weight)
> abline(lm(Weight~nLength_nWidth))
> lm(Weight~nLength_nWidth)

Call:
lm(formula = Weight ~ nLength_nWidth)

Coefficients:
   (Intercept)  nLength_nWidth
      -115.102           3.102

> model.5=lm(Weight~nLength_nWidth)
```

```
> summary(model.5)

Call:
lm(formula = Weight ~ nLength_nWidth)

Residuals:
   Min     1Q Median     3Q    Max 
-40.47 -28.27  -9.90  17.04  82.79 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)    -115.1021    21.8703  -5.263 0.000367 ***
nLength_nWidth    3.1019     0.1179  26.318 1.45e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 41.69 on 10 degrees of freedom
Multiple R-squared:  0.9858,Adjusted R-squared:  0.9843 
F-statistic: 692.6 on 1 and 10 DF,  p-value: 1.446e-10

> lm.Weight4 <-lm(Weight~nLength_nWidth)
> fitted(lm.Weight4)
        1         2         3         4         5         6         7         8 
-76.88658 338.92385  81.43467 314.38777 136.15229 719.62070 140.02967 655.41126 
        9        10        11        12 
152.90258 754.98243 185.47258 983.46878 
> resid(lm.Weight4)
         1          2          3          4          5          6          7          8 
 82.786582 -38.923846  18.565327 -14.387774 -26.152286 -34.620705 -20.029668  -5.411261 
         9         10         11         12 
 -2.902575  65.017573 -40.472583  16.531216 
> plot(fitted(lm.Weight4),resid(lm.Weight4))
> qqnorm(resid(lm.Weight4))
> cor.test(nLength_nWidth,Weight)

Pearson's product-moment correlation

data:  nLength_nWidth and Weight
t = 26.318, df = 10, p-value = 1.446e-10
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
```

```
 0.9738714 0.9980615
sample estimates:
      cor
0.9928582
```