

perch

Chris Parrish

January 8, 2016

perch

references:

- Cannon, et al., Stat2, chapter 03, examples 3.10, 3.12-3.13, 3.16

Import the data.

```
data <- read.csv("Perch.csv", header=TRUE)
head(data)
```

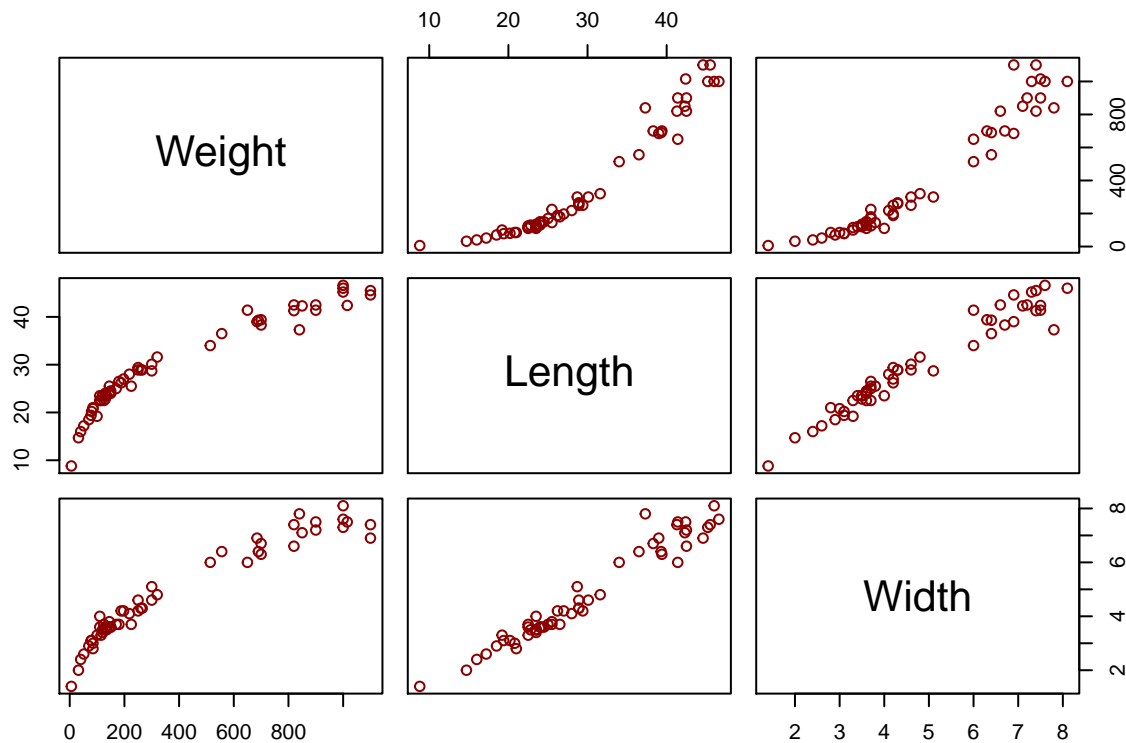
```
##   Obs Weight Length Width
## 1 104   5.9   8.8   1.4
## 2 105  32.0  14.7   2.0
## 3 106  40.0  16.0   2.4
## 4 107  51.5  17.2   2.6
## 5 108  70.0  18.5   2.9
## 6 109 100.0  19.2   3.3
```

```
dim(data)
```

```
## [1] 56 4
```

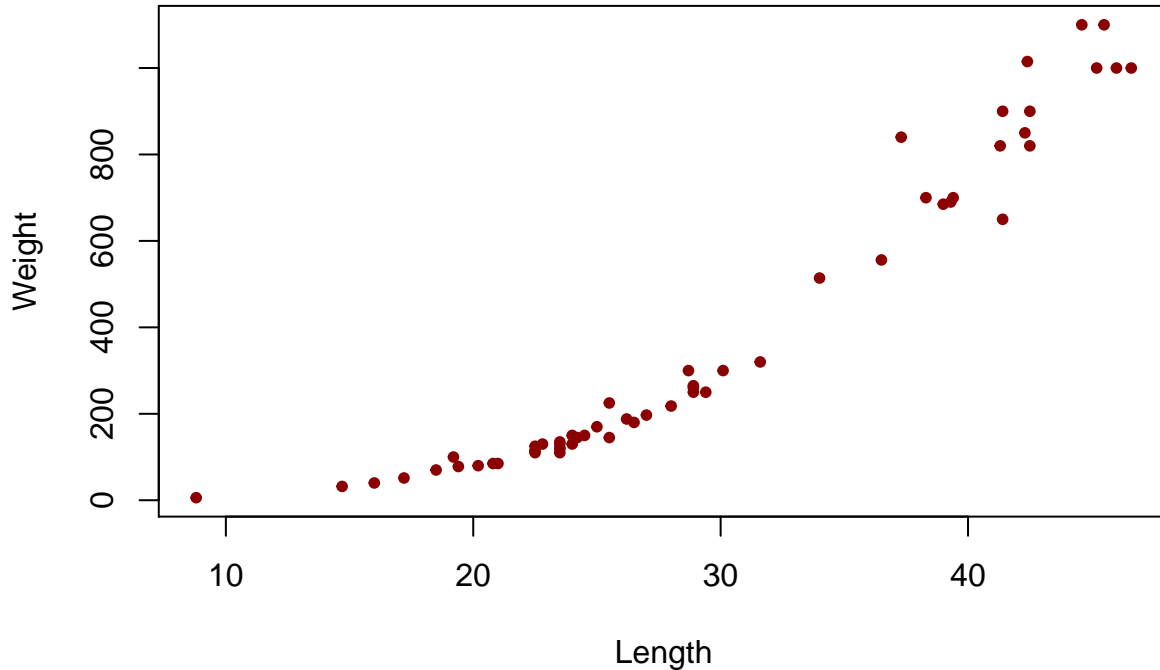
Scatterplot matrix.

```
pairs(~ Weight + Length + Width, data=data, col="darkred")
```



Separate linear models for length and width

```
# perch
plot(Weight ~ Length, data=data,
     pch=20, col="darkred")
```



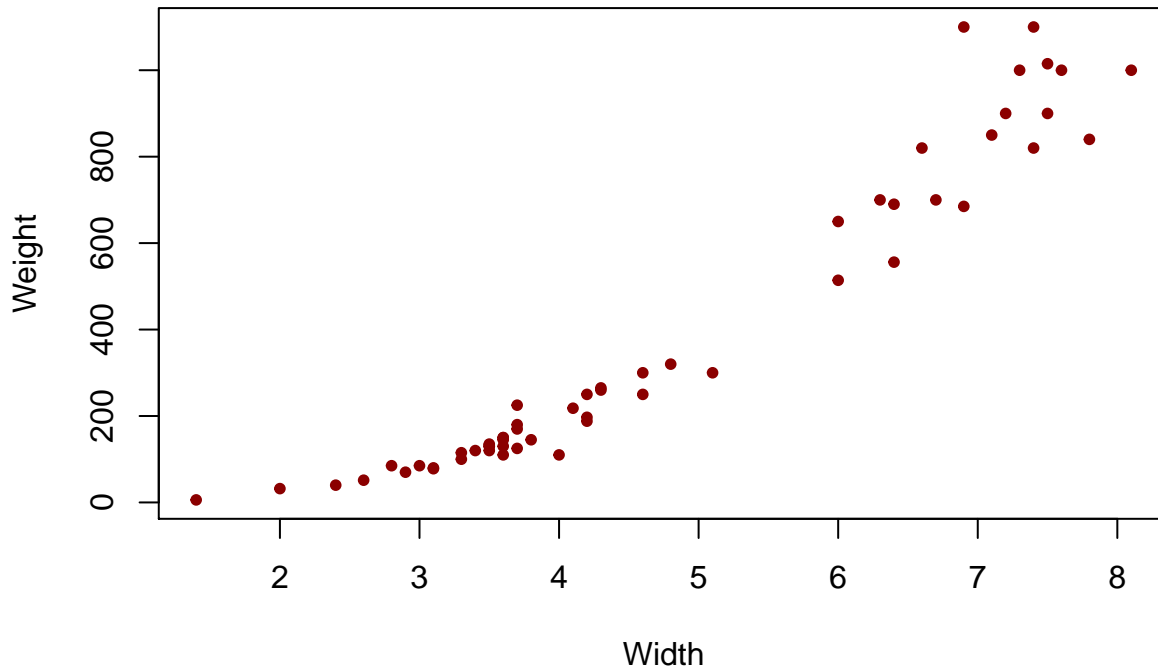
```
perch.length.lm <- lm(Weight ~ Length, data=data)
coef(perch.length.lm)
```

```
## (Intercept)      Length
## -652.78714    35.00089
```

```
options(show.signif.stars=FALSE)
summary(perch.length.lm)
```

```
##
## Call:
## lm(formula = Weight ~ Length, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -146.25  -57.86  -23.99   45.00  350.68
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -652.787    43.407  -15.04  <2e-16
## Length       35.001     1.398   25.03  <2e-16
##
## Residual standard error: 98.82 on 54 degrees of freedom
## Multiple R-squared:  0.9207, Adjusted R-squared:  0.9192
## F-statistic: 626.5 on 1 and 54 DF, p-value: < 2.2e-16
```

```
# width
plot(Weight ~ Width, data=data,
     pch=20, col="darkred")
```



```
perch.width.lm <- lm(Weight ~ Width, data=data)
coef(perch.width.lm)
```

```
## (Intercept)      Width
## -509.2893      188.1146
```

```
summary(perch.width.lm)
```

```
##
## Call:
## lm(formula = Weight ~ Width, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -150.09  -56.51  -15.59   31.39  311.30
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -509.289    35.594  -14.31  <2e-16
## Width       188.115     7.038   26.73  <2e-16
##
## Residual standard error: 93 on 54 degrees of freedom
## Multiple R-squared:  0.9297, Adjusted R-squared:  0.9284
## F-statistic: 714.5 on 1 and 54 DF,  p-value: < 2.2e-16
```

Multiple regression without interaction.

```
perch.lm1<- lm(Weight ~ Length + Width, data=data)
summary(perch.lm1)
```

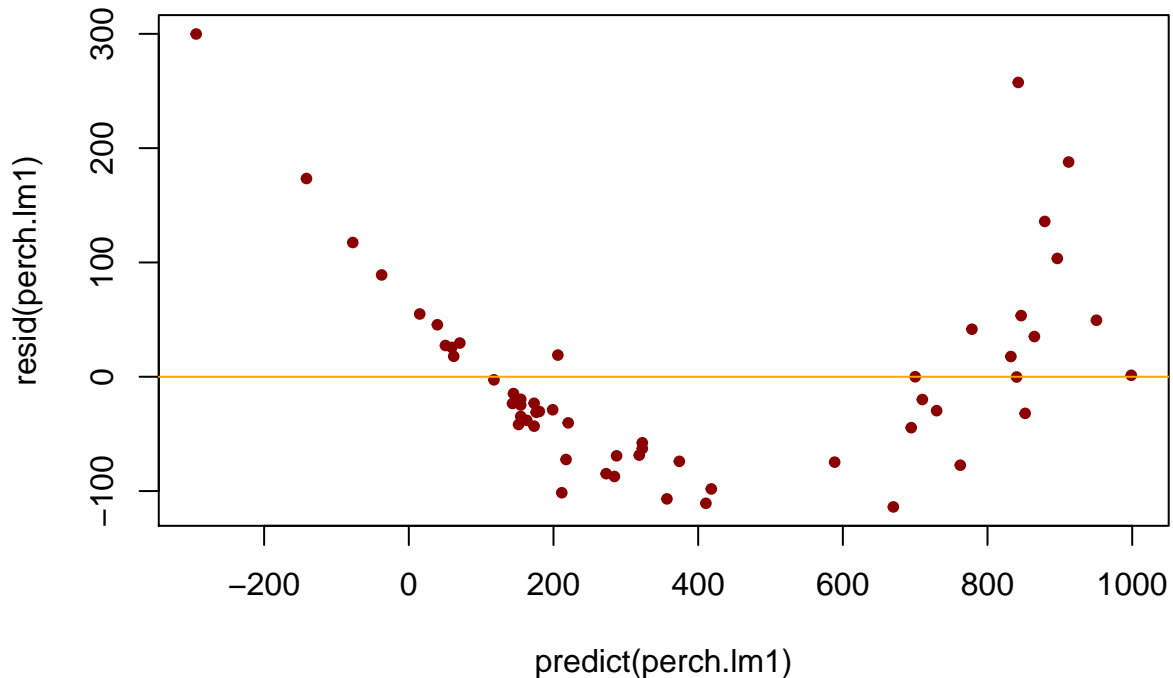
```
##
## Call:
## lm(formula = Weight ~ Length + Width, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -113.86  -59.02  -23.29   30.93  299.85
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -578.758     43.667  -13.254 < 2e-16
## Length       14.307       5.659   2.528 0.014475
## Width        113.500      30.265   3.750 0.000439
##
## Residual standard error: 88.68 on 53 degrees of freedom
## Multiple R-squared:  0.9373, Adjusted R-squared:  0.9349
## F-statistic: 396.1 on 2 and 53 DF,  p-value: < 2.2e-16
```

```
anova(perch.lm1)
```

```
## Analysis of Variance Table
##
## Response: Weight
##          Df Sum Sq Mean Sq F value    Pr(>F)
## Length    1 6118739 6118739  778.126 < 2.2e-16
## Width     1  110593  110593   14.064 0.0004388
## Residuals 53  416762    7863
```

Residuals without interaction.

```
plot(predict(perch.lm1), resid(perch.lm1),
      pch=20, col="darkred")
abline(h=0, col="orange")
```



Multiple regression with interaction.

```
perch.lm2 <- lm(Weight ~ Length * Width, data=data)
summary(perch.lm2)
```

```
##
## Call:
## lm(formula = Weight ~ Length * Width, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -140.106  -12.226    1.230    8.489  181.408
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  113.9349   58.7844   1.938  0.058
## Length       -3.4827    3.1521  -1.105  0.274
## Width        -94.6309   22.2954  -4.244 9.06e-05
## Length:Width   5.2412    0.4131  12.687 < 2e-16
##
## Residual standard error: 44.24 on 52 degrees of freedom
## Multiple R-squared:  0.9847, Adjusted R-squared:  0.9838
## F-statistic: 1115 on 3 and 52 DF, p-value: < 2.2e-16
```

```
anova(perch.lm2)
```

```
## Analysis of Variance Table
##
## Response: Weight
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Length     1 6118739 6118739 3126.571 < 2.2e-16
```

```
## Width      1 110593 110593 56.511 7.416e-10
## Length:Width 1 314997 314997 160.958 < 2.2e-16
## Residuals  52 101765 1957
```

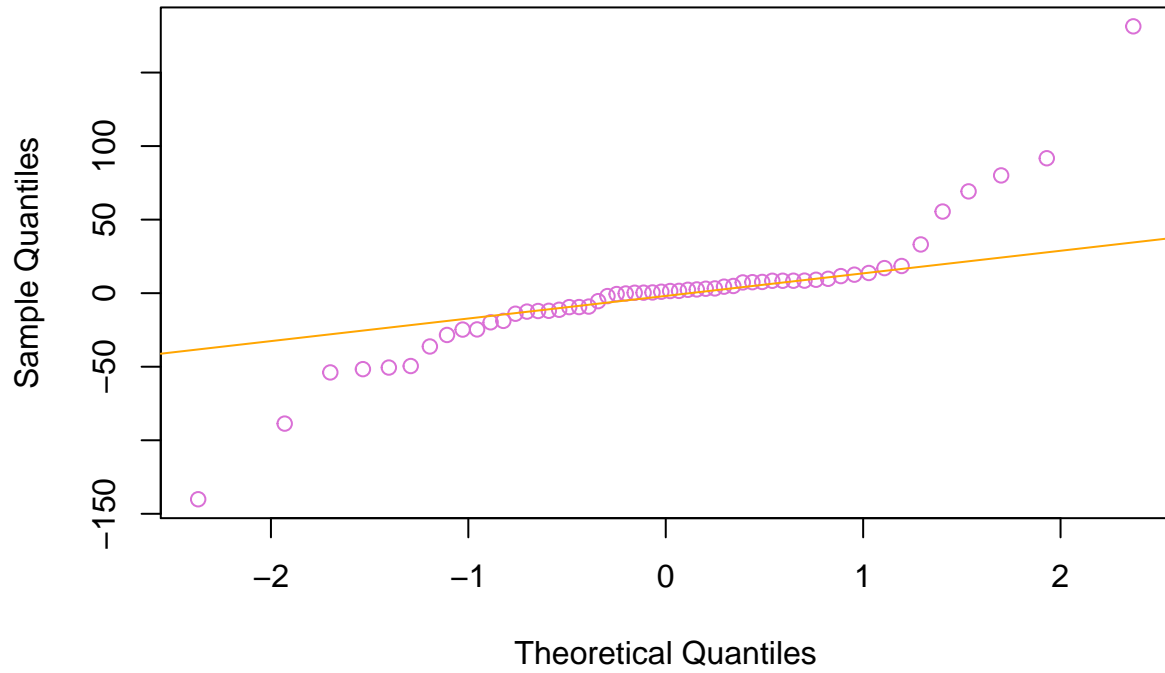
Residuals with interaction.

```
hist(resid(perch.lm2),
      col="wheat")
```

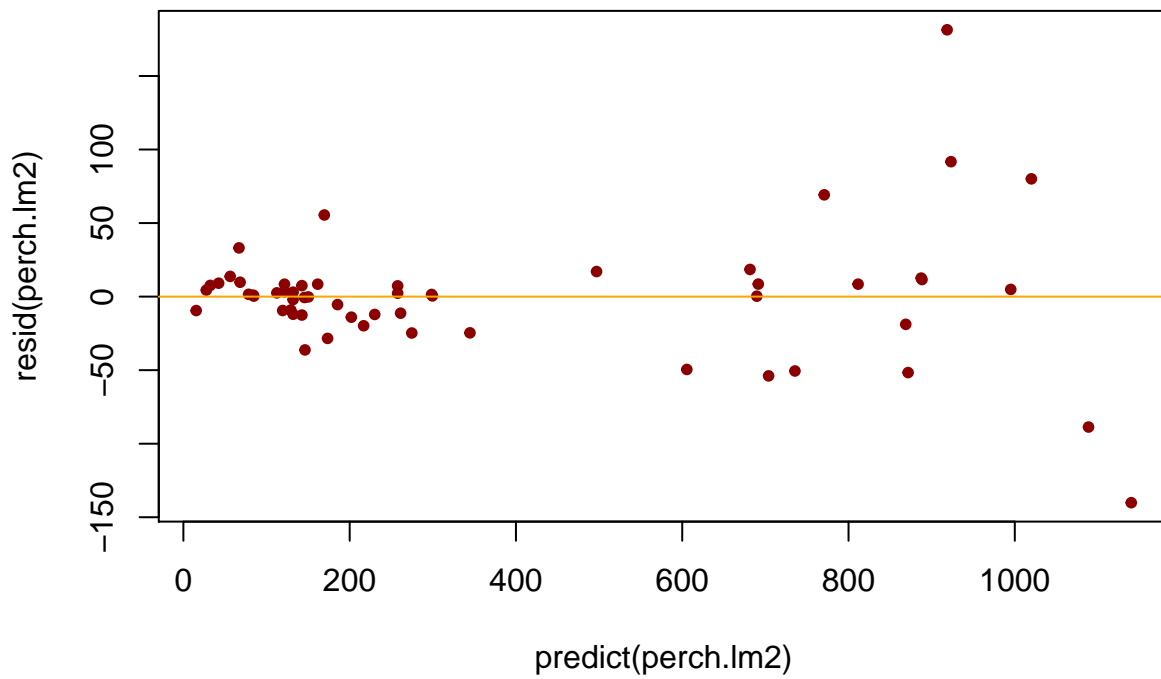


```
qqnorm(resid(perch.lm2),
        col="orchid")
qqline(resid(perch.lm2),
        col="orange")
```

Normal Q-Q Plot



```
plot(predict(perch.lm2), resid(perch.lm2),  
      pch=20, col="darkred")  
abline(h=0, col="orange")
```



Complete second-order model.

```
perch.lm3 <- lm(Weight ~ Length + Width + I(Length^2) + I(Width^2) + Length:Width, data=data)
summary(perch.lm3)
```

```
##
## Call:
## lm(formula = Weight ~ Length + Width + I(Length^2) + I(Width^2) +
##     Length:Width, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -117.175  -11.904    2.822   11.556  157.596
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  156.3486    61.4152   2.546  0.0140
## Length       -25.0007    14.2729  -1.752  0.0860
## Width         20.9772    82.5877   0.254  0.8005
## I(Length^2)   1.5719     0.7244   2.170  0.0348
## I(Width^2)    34.4058    18.7455   1.835  0.0724
## Length:Width  -9.7763     7.1455  -1.368  0.1774
##
## Residual standard error: 43.13 on 50 degrees of freedom
## Multiple R-squared:  0.986, Adjusted R-squared:  0.9846
## F-statistic: 704.6 on 5 and 50 DF,  p-value: < 2.2e-16
```

```
anova(perch.lm3)
```

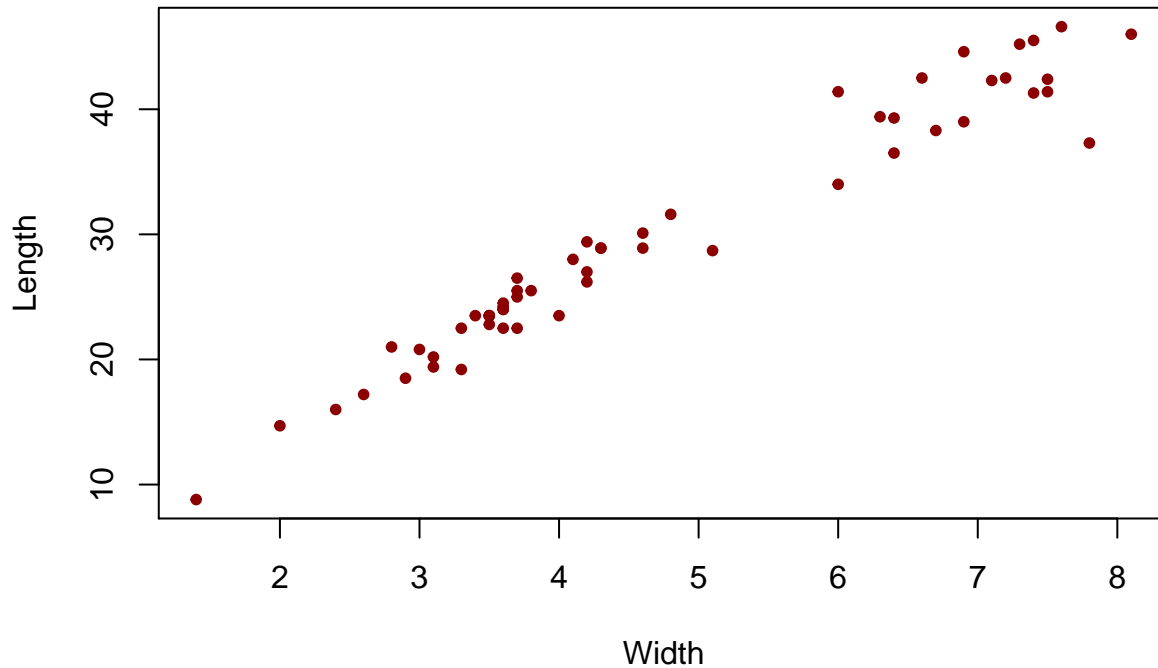
```
## Analysis of Variance Table
##
## Response: Weight
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## Length     1 6118739 6118739 3289.6413 < 2.2e-16
## Width      1  110593  110593   59.4585 4.667e-10
## I(Length^2) 1  314899  314899  169.3002 < 2.2e-16
## I(Width^2)  1    5381    5381    2.8932 0.09517
## Length:Width 1    3482    3482    1.8719 0.17737
## Residuals   50   93000    1860
```

Correlated predictors.

```
with(data,
      cor(cbind(Weight, Length, Width)))
```

```
##           Weight  Length  Width
## Weight 1.0000000 0.9595061 0.9642244
## Length 0.9595061 1.0000000 0.9751074
## Width  0.9642244 0.9751074 1.0000000
```

```
plot(Length ~ Width, data=data,
     pch=20, col="darkred")
```

```
perch.lm4 <- lm(Weight ~ Length + Width, data=data)
summary(perch.lm4)
```

```
##
## Call:
## lm(formula = Weight ~ Length + Width, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -113.86  -59.02  -23.29   30.93  299.85
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -578.758     43.667  -13.254 < 2e-16
## Length         14.307       5.659   2.528 0.014475
## Width         113.500      30.265   3.750 0.000439
##
## Residual standard error: 88.68 on 53 degrees of freedom
## Multiple R-squared:  0.9373, Adjusted R-squared:  0.9349
## F-statistic: 396.1 on 2 and 53 DF,  p-value: < 2.2e-16
```

```
perch.lm5 <- lm(Weight ~ Length + Width + Length:Width, data=data)
summary(perch.lm5)
```

```
##
## Call:
## lm(formula = Weight ~ Length + Width + Length:Width, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -140.106  -12.226    1.230    8.489  181.408
```

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 113.9349    58.7844   1.938   0.058
## Length      -3.4827     3.1521  -1.105   0.274
## Width       -94.6309    22.2954  -4.244 9.06e-05
## Length:Width  5.2412     0.4131  12.687 < 2e-16
##
## Residual standard error: 44.24 on 52 degrees of freedom
## Multiple R-squared:  0.9847, Adjusted R-squared:  0.9838
## F-statistic: 1115 on 3 and 52 DF,  p-value: < 2.2e-16
```

Collinearity.

```
cor.matrix <- with(data,
                    cor(cbind(Length, Width, Length^2, Width^2, Length * Width)))
var.names <- c("Length", "Width", "Length^2", "Width^2", "Length*Width")
rownames(cor.matrix) <- colnames(cor.matrix) <- var.names
cor.matrix
```

```
##           Length      Width Length^2  Width^2 Length*Width
## Length      1.0000000 0.9751074 0.9888600 0.9523922  0.9785746
## Width       0.9751074 1.0000000 0.9680791 0.9896026  0.9878073
## Length^2    0.9888600 0.9680791 1.0000000 0.9644068  0.9905755
## Width^2     0.9523922 0.9896026 0.9644068 1.0000000  0.9914553
## Length*Width 0.9785746 0.9878073 0.9905755 0.9914553  1.0000000
```

Nested (or incremental) F-tests for testing subsets of predictors.

```
# perch.lm5 <- lm(Weight ~ Length + Width + Length:Width, data=data)
# perch.lm3 <- lm(Weight ~ Length + Width + I(Length^2) + I(Width^2) + Length:Width, data=data)
anova(perch.lm5, perch.lm3)
```

```
## Analysis of Variance Table
##
## Model 1: Weight ~ Length + Width + Length:Width
## Model 2: Weight ~ Length + Width + I(Length^2) + I(Width^2) + Length:Width
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1      52 101765
## 2      50  93000  2   8764.6 2.3561 0.1052
```