# Exercise – Simple Linear Regression

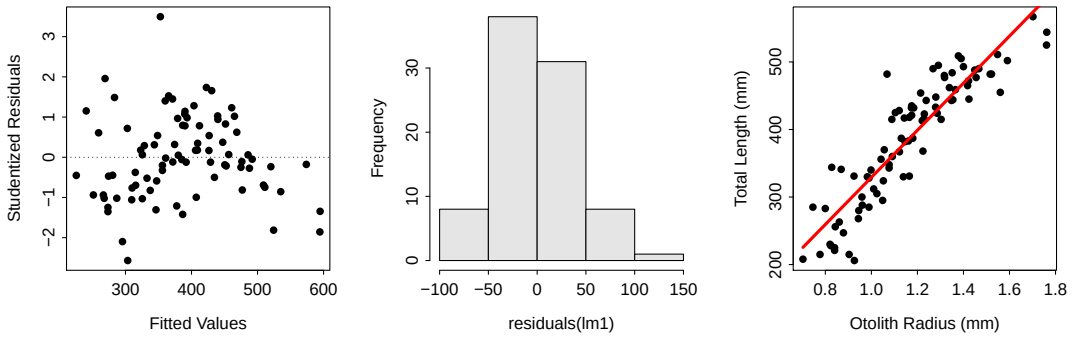
Answer the following questions with R code by creating (*and editing if you make a mistake*) an R script and iteratively running the code in RStudio.

1. Load the data in the **LakeTroutALTER.csv** file into a data frame in R.

```
> setwd("C:/aaaWork/Web/fishR/courses/Vermont2014/CourseMaterial/Exercises")
> lkt <- read.csv("Data/LakeTroutALTER.csv")
> str(lkt)
'data.frame': 86 obs. of  8 variables:
 $ id    : int  18 512 307 52 84 37 80 36 17 59 ...
 $ tl    : int  225 247 256 268 285 288 295 324 328 330 ...
 $ fl    : int  202 226 235 241 262 265 270 295 297 299 ...
 $ sl    : int  185 212 209 228 240 244 243 273 278 280 ...
 $ w     : int  76 138 120 170 185 182 205 275 285 297 ...
 $ otorad: num  0.84 0.879 0.843 0.944 0.99 ...
 $ age   : int  8 6 6 9 7 9 7 10 7 10 ...
 $ sex   : Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 1 ...
```

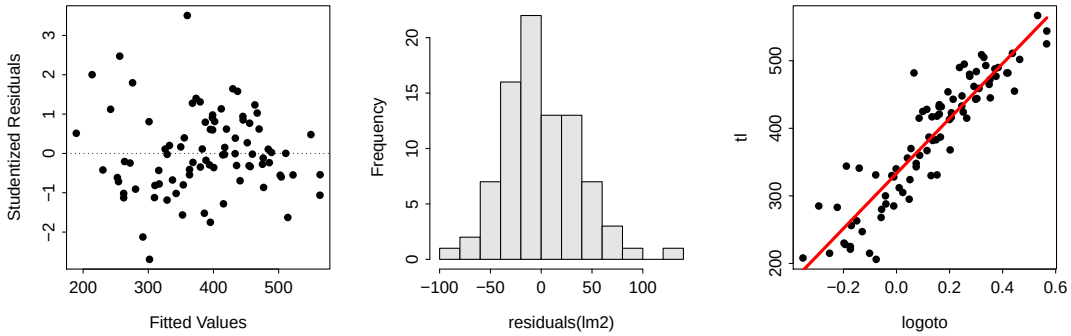
2. Fit the linear model betwween total length (dependent variable) and otolith radius. Assess the appropriateness of this model.

```
> lm1 <- lm(tl~otorad,data=lkt)
> residPlot(lm1)                  # Left
> hist(~residuals(lm1))           # Middle
> fitPlot(lm1,ylab="Total Length (mm)",xlab="Otolith Radius (mm)",pch=16)  # Right
```




3. Fit the linear model betwween total length (dependent variable) and the natural log of otolith radius. Assess the appropriateness of this model.

```
> lkt$logoto <- log(lkt$otorad)
> lm2 <- lm(tl~logoto,data=lkt)
> residPlot(lm2)
> hist(~residuals(lm2))
> fitPlot(lm2)
```




4. Answer the following questions with the total length and natural log of otolith radius model.

   (a) Is the relationship between total length ans otolith radius significant? How much variability in total length is explained by otolith radius?

```
> summary(lm2)

Call:
lm(formula = tl ~ logoto, data = lkt)

Residuals:
   Min     1Q Median     3Q    Max
-95.87 -22.42  -2.86  23.02 122.15

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   333.07       4.78    69.7   <2e-16
logoto        407.03      19.46    20.9   <2e-16

Residual standard error: 37.3 on 84 degrees of freedom
Multiple R-squared:  0.839,Adjusted R-squared:  0.837
F-statistic:  438 on 1 and 84 DF,  p-value: <2e-16
```

   The relationship between total length and log otolith radious is statistically significant ($p < 0.00005$) with 83.9% of the variability in total length explained by log otolith radius.

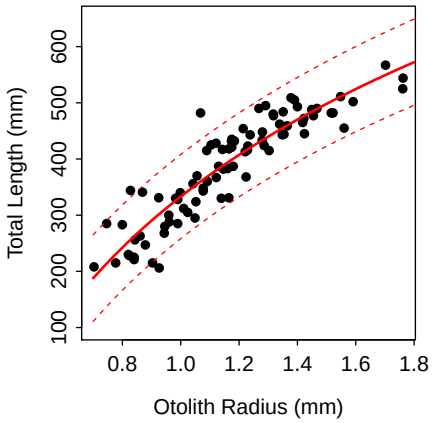
   (b) Predict the total length if the scale radius is 1.2 mm.

```
> ( p1.2 <- predict(lm2,data.frame(logoto=log(1.2)),interval="predict") )
    fit   lwr upr
1 407.3 332.6 482
```

   The predicted total length if the otolith radius is 1.2 mm is between 333 and 482, with a best guess of 407.

   (c) Construct a plot that illustrates the model with prediction intervals on the original scale.

```
> xs <- seq(0.7,1.8,0.01)
> pTL <- predict(lm2,data.frame(logoto=log(xs)),interval="prediction")
> plot(tl~otorad,data=lkt,pch=16,ylim=c(100,650),
      ylab="Total Length (mm)",xlab="Otolith Radius (mm)")
> lines(pTL[,"fit"]~xs,col="red",lwd=2)
> lines(pTL[,"lwr"]~xs,col="red",lty=2)
> lines(pTL[,"upr"]~xs,col="red",lty=2)
```




5. *If time permits ...* fit the length-weight regression for the Lake Trout data.

```
> lkt <- within(lkt,{
   logW <- log(w)
   logTL <- log(tl)
 })
> lm3 <- lm(logW~logTL,data=lkt)
> summary(lm3)


Call:
lm(formula = logW ~ logTL, data = lkt)

Residuals:
    Min      1Q  Median      3Q     Max
-0.5330 -0.0754  0.0238  0.0684  0.2118

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -12.9744     0.2943   -44.1   <2e-16
logTL         3.2295     0.0496    65.1   <2e-16

Residual standard error: 0.12 on 84 degrees of freedom
Multiple R-squared:  0.981,Adjusted R-squared:  0.98
F-statistic: 4.24e+03 on 1 and 84 DF,  p-value: <2e-16
> fitPlot(lm3)
```