Exercise – Basic Summaries

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. Summarize age in the following ways:

   (a) Construct age-frequency (number-at-age) and percentage-at-age tables.

```
> ( agetbl <- xtabs(~age,data=lkt) )
age
 3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 21 23 24 25 27 32 34 43 50
 1  2  1 10  7  6  7  7 10  6  3  3  3  1  6  2  1  1  1  1  2  1  1  1  1  1
> ( ageptbl <- round(prop.table(agetbl)*100,1) )
age
   3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18   19
 1.2  2.3  1.2 11.6  8.1  7.0  8.1  8.1 11.6  7.0  3.5  3.5  3.5  1.2  7.0  2.3  1.2
  21   23   24   25   27   32   34   43   50
 1.2  1.2  1.2  2.3  1.2  1.2  1.2  1.2  1.2
```

   (b) Construct a bar chart of the percentage-at-age table.

```
> barplot(ageptbl,xlab="Age",ylab="Percentage of Fish",ylim=c(0,12))
```
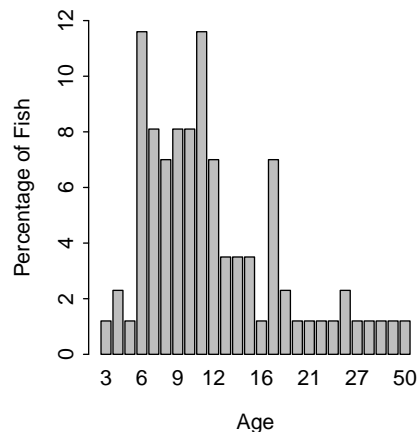

   (c) Construct an age-percentage table separated by sex (e.g., what percentage of males were age-17?).

```
> agetbl2 <- xtabs(~sex+age,data=lkt)
> round(prop.table(agetbl2,margin=1)*100,1)
```

```
       age
sex     3     4     5     6     7     8     9    10    11    12    13    14    15    16    17    18
  F   0.0   0.0   0.0   8.5   8.5   8.5  10.6  10.6   8.5  10.6   6.4   2.1   2.1   2.1   8.5   4.3
  M   2.6   5.1   2.6  15.4   7.7   5.1   5.1   5.1  15.4   2.6   0.0   5.1   5.1   0.0   5.1   0.0
       age
sex    19    21    23    24    25    27    32    34    43    50
  F   2.1   2.1   0.0   0.0   0.0   2.1   0.0   0.0   2.1   0.0
  M   0.0   0.0   2.6   2.6   5.1   0.0   2.6   2.6   0.0   2.6
```
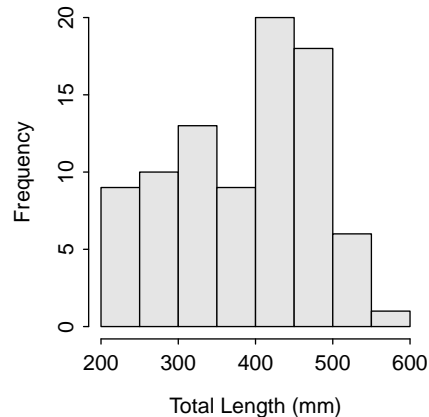
3. Summarize total length in the following ways:

   (a) Compute summary statistics of total length for all fish.

```
> Summarize(~tl,data=lkt,digits=1)
      n     mean       sd      min       Q1   median       Q3      max percZero
   86.0    387.0     92.5    206.0    325.0    415.0    461.0    567.0      0.0
```

   (b) Construct a histogram of total length using 50-mm length increments.

```
> hist(~tl,data=lkt,right=TRUE,breaks=seq(200,600,50),xlab="Total Length (mm)")
```
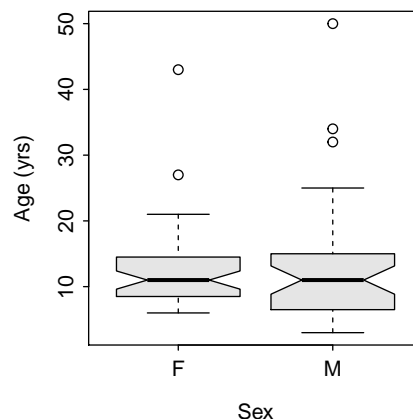

   (c) Compute summary statistics of total length separately for each sex.

```
> Summarize(tl~sex,data=lkt,digits=1)
  sex  n  mean     sd min  Q1 median  Q3 max percZero
1   F 47 396.0  80.6 225 331    415 460 544        0
2   M 39 376.3 105.2 206 284    413 464 567        0
```
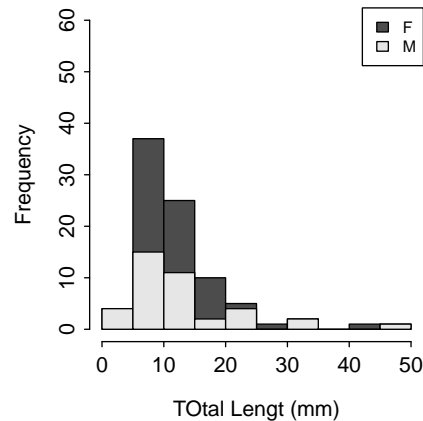
   (d) Construct a notched boxplot of total length separately for each sex.

```
> boxplot(age~sex,data=lkt,xlab="Sex",ylab="Age (yrs)",col="gray90",notch=TRUE)
```


   (e) Construct a stacked histogram of total length separately for each sex.

```
> histStack(age~sex,data=lkt,xlab="TOtal Lengt (mm)",breaks=seq(200,600,50),
          ylim=c(0,60),col="gray.colors",legend="topright")
```


4. Examine the following relationships (graphically and, if appropriate, numerically):

   (a) Between total length and otolith radius.

   ```
   > plot(otorad~tl,data=lkt,xlab="Total Length (mm)",ylab="Otolith Radius (mm)",pch=16)
   ```
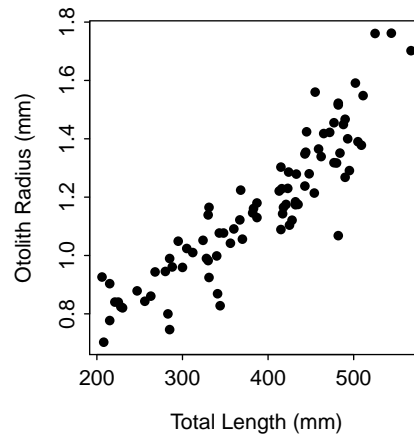

5. If you have time ...

   (a) Construct a plot of total length, with approximate 95% confidence intervals, by age.

   ```
   > smrzTLAge <- Summarize(tl~age,data=lkt,digits=1)
   Warning:  To continue, variable(s) on RHS of formula were converted to a factor.
   > smrzTLAge <- within(smrzTLAge, {
      LCI <- mean-1.96*sd/sqrt(n)
      UCI <- mean+1.96*sd/sqrt(n)
   })
   > with(smrzTLAge,plotCI(fact2num(age),mean,ui=UCI,li=LCI,pch=16,xlab="Age",ylab="Mean TL"))
   ```
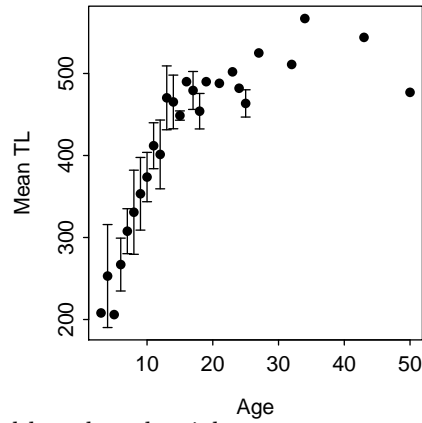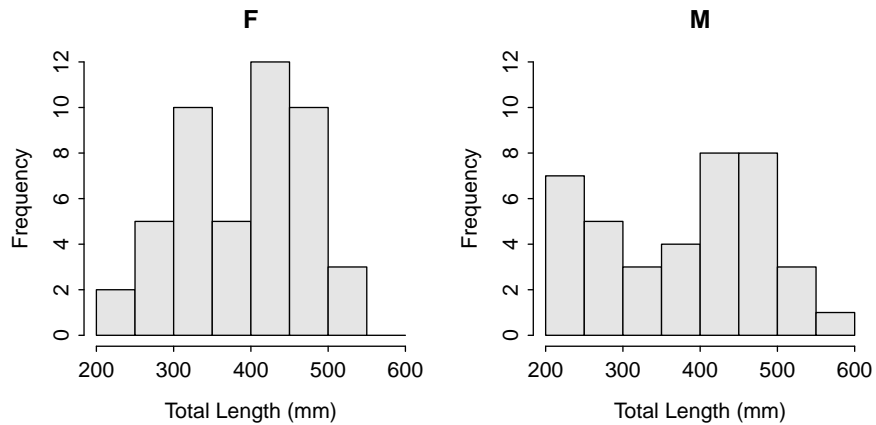
(b) Construct a plot between total length and weight.

```
> plot(w~tl,data=lkt,xlab="Total Length (mm)",ylab="Weight (g)",pch=16)
```
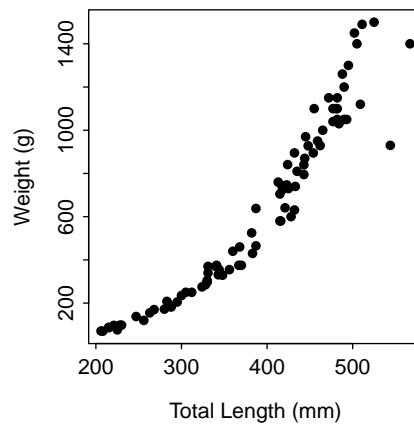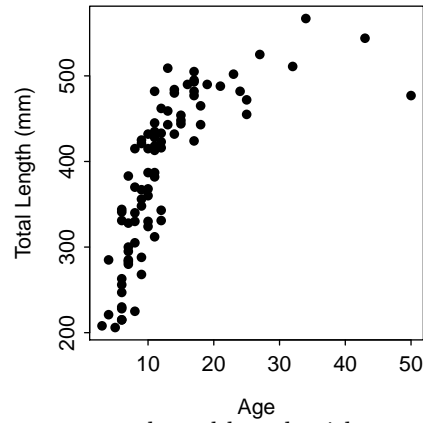

(c) Construct separate histograms of total length for males and females.

```
> hist(tl~sex,data=lkt,right=TRUE,breaks=seq(200,600,50),xlab="Total Length (mm)")
```


(d) Examine the relationship between age and total length.

```
> plot(tl~age,data=lkt,xlab="Age",ylab="Total Length (mm)",pch=16)
```

(e) Examine the relationship between age and total length with separate symbols for different sexes.

```
> pts <- c(16,3)
> plot(tl~age,data=lkt,xlab="Age",ylab="Total Length (mm)",pch=pts[sex])
> legend("bottomright",pch=pts,legend=levels(lkt$sex))
```