Exercise – Back-Calculation

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 **MN98WaeYep.csv** file into a data frame in R.

```
> setwd("C:/aaaWork/Web/fishR/courses/Vermont2014/CourseMaterial/Exercises")
> mn <- read.csv("Data/MN98WaeYep.csv")
> str(mn)
'data.frame': 1124 obs. of  17 variables:
 $ species: Factor w/ 2 levels "WAE","YEP": 1 1 1 1 1 1 1 2 2 2 ...
 $ lake   : Factor w/ 16 levels "Bean Lake","Bingham Lake",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ yearcap: int  1998 1998 1998 1998 1998 1998 1998 1998 1998 1998 ...
 $ fishID : int  75 46 63 90 85 62 2 76 67 33 ...
 $ agecap : int  2 2 2 2 2 2 6 1 1 1 ...
 $ lencap : int  416 425 426 439 444 446 590 168 176 178 ...
 $ radcap : num  4.98 5.74 5.3 4.95 6.48 ...
 $ rad1   : num  2.27 2.6 2.59 2.36 3.08 ...
 $ rad2   : num  4.27 4.86 4.51 4.04 5.48 ...
 $ rad3   : num  4.98 5.74 5.3 4.95 6.48 ...
 $ rad4   : num  NA NA NA NA NA ...
 $ rad5   : num  NA NA NA NA NA ...
 $ rad6   : num  NA NA NA NA NA ...
 $ rad7   : num  NA NA NA NA NA ...
 $ rad8   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad9   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad10  : num  NA NA NA NA NA NA NA NA NA NA ...
```

2. Isolate the Lake Shetek Walleye data.

```
> df <- Subset(mn,species=="WAE" & lake=="Lake Shetek")
> str(df)
'data.frame': 81 obs. of  17 variables:
 $ species: Factor w/ 1 level "WAE": 1 1 1 1 1 1 1 1 1 1 ...
 $ lake   : Factor w/ 1 level "Lake Shetek": 1 1 1 1 1 1 1 1 1 1 ...
 $ yearcap: int  1998 1998 1998 1998 1998 1998 1998 1998 1998 1998 ...
 $ fishID : int  155 153 171 85 172 175 154 48 87 174 ...
 $ agecap : int  1 1 1 1 1 1 1 1 1 1 ...
 $ lencap : int  185 189 190 191 197 198 205 207 208 209 ...
 $ radcap : num  2.17 1.9 2.02 1.97 2.13 ...
 $ rad1   : num  1.3 1.15 1.21 1.1 1.27 ...
 $ rad2   : num  2.17 1.9 2.02 1.97 2.13 ...
 $ rad3   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad4   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad5   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad6   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad7   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad8   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad9   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ rad10  : num  NA NA NA NA NA NA NA NA NA NA ...
```

3. Is "plus-growth" recorded for your chosen data? Explain.

```
> head(df,n=3)
```

```
  species        lake yearcap fishID agecap lencap radcap  rad1  rad2 rad3 rad4 rad5
1     WAE Lake Shetek    1998    155      1    185  2.165 1.304 2.165   NA   NA   NA
2     WAE Lake Shetek    1998    153      1    189  1.903 1.150 1.903   NA   NA   NA
3     WAE Lake Shetek    1998    171      1    190  2.024 1.209 2.024   NA   NA   NA
  rad6 rad7 rad8 rad9 rad10
1   NA   NA   NA   NA    NA
2   NA   NA   NA   NA    NA
3   NA   NA   NA   NA    NA
```

```
> tail(df,n=3)
```

```
   species        lake yearcap fishID agecap lencap radcap  rad1  rad2  rad3  rad4
79     WAE Lake Shetek    1998    179      5    543  7.245 2.627 4.600 5.878 6.765
80     WAE Lake Shetek    1998    140      6    596  7.231 3.275 4.841 5.743 6.282
81     WAE Lake Shetek    1998    141      6    630  6.964 2.214 3.762 5.081 5.813
    rad5  rad6  rad7 rad8 rad9 rad10
79 7.123 7.245    NA   NA   NA    NA
80 6.632 6.947 7.231   NA   NA    NA
81 6.304 6.753 6.964   NA   NA    NA
```

"Plus-growth" is recorded because one more "anu" measurement appears in the data file then the assessed age. For example, fish 155 was 1-year-old but two radial measurements were recorded and fish 1215 was 6-years-old but seven radial measurements were recorded.

4. Reshape the data frame from "wide" to "long" format so that it will be suitable for adding a back-calculated total length variable. Make sure to remove unneccesary "NA"s and the "plus-growth", if it was recorded.

```
> varying1 <- paste("rad",1:10,sep="")
> dfl <- reshape(df,direction="long",
                 idvar="fishID",        # what identifies unique fish
                 varying=varying1,      # declare the repeated measurements
                 v.names="rad",         # name for repeat meas in long format
                 timevar="age",         # name of var that identifies the repeat
                 times=1:10)            # values in timevar for repeat
> dfl <- Subset(dfl,!is.na(rad))        # remove all of the NAs
> dfl <- Subset(dfl,agecap-age>=0)      # remove the "plus" growth
> str(dfl)
```

```
'data.frame': 172 obs. of  9 variables:
 $ species: Factor w/ 1 level "WAE": 1 1 1 1 1 1 1 1 1 1 ...
 $ lake   : Factor w/ 1 level "Lake Shetek": 1 1 1 1 1 1 1 1 1 1 ...
 $ yearcap: int  1998 1998 1998 1998 1998 1998 1998 1998 1998 1998 ...
 $ fishID : int  155 153 171 85 172 175 154 48 87 174 ...
 $ agecap : int  1 1 1 1 1 1 1 1 1 1 ...
 $ lencap : int  185 189 190 191 197 198 205 207 208 209 ...
 $ radcap : num  2.17 1.9 2.02 1.97 2.13 ...
 $ age    : int  1 1 1 1 1 1 1 1 1 1 ...
 $ rad    : num  1.3 1.15 1.21 1.1 1.27 ...
```

5. Add a variable that is the Fraser-Lee back-calculated total length if the "correction factor" is 55 mm.

```
> k <- 55  # use Carlander intercept of k=55 mm
> dfl$bcFL <- (dfl$rad/dfl$radcap)*(dfl$lencap-k)+k
```

6. Compute the mean length-at-back-calculated-age.

```
> Summarize(bcFL~age,data=dfl)
```

```
Warning:  To continue, variable(s) on RHS of formula were converted to a factor.
```

```
  age  n  mean    sd min  Q1 median  Q3 max percZero
1   1 81 155.7 35.88 108 131    147 172 300        0
2   2 68 264.1 59.85 179 216    257 310 417        0
3   3 10 439.1 32.05 385 420    447 455 485        0
4   4  7 506.4 23.55 475 488    511 524 535        0
5   5  4 539.9 32.46 498 526    543 557 576        0
6   6  2 593.6 26.70 575 584    594 603 613        0
```