Back-Calculation Assignment

The file `MNBCData98.csv` contains radial measurements from scales for a variety of species captured in 1998 from a variety of lakes in southern Minnesota. Use these data to answer the questions below.

```
> setwd("C:/aaaWork/Web/fishR/Courses/MNAFS2013/CourseMaterial")
> mn <- read.csv("MNBCData98.csv",header=TRUE)
> str(mn)
'data.frame': 2065 obs. of  20 variables:
 $ species: Factor w/ 9 levels "BLC","BLG","LMB",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ gear   : Factor w/ 1 level "All": 1 1 1 1 1 1 1 1 1 1 ...
 $ lake   : Factor w/ 17 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 ...
 $ fish   : int  38 41 37 39 42 40 81 35 36 78 ...
 $ agecap : int  3 3 4 4 4 4 4 4 4 5 ...
 $ lencap : int  290 297 316 318 318 319 323 328 329 337 ...
 $ anu1   : num  1.093 0.982 1.141 1.119 1.059 ...
 $ anu2   : num  3.57 3.39 3.37 3.84 3.48 ...
 $ anu3   : num  5.08 5.16 4.6 5.34 4.8 ...
 $ anu4   : num  5.66 5.81 5.33 5.9 5.63 ...
 $ anu5   : num  NA NA 5.54 6.08 5.83 ...
 $ anu6   : num  NA NA NA NA NA ...
 $ anu7   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ anu8   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ anu9   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ anu10  : num  NA NA NA NA NA NA NA NA NA NA ...
 $ anu11  : logi  NA NA NA NA NA NA ...
 $ anu12  : logi  NA NA NA NA NA NA ...
 $ radcap : num  5.66 5.81 5.54 6.08 5.83 ...
```

1. Choose a lake and species to examine. You may find it useful to use the following code, which assumes that you named the data frame as `mn` in R and you are interested in Lake Shetek Walleye.

   ```
   > table(mn$lake,mn$species)
   > df <- Subset(mn,species=="WAE" & lake=="Lake Shetek")
   ```

   I chose to examine Lake Shetek Walleye as illustrated with the example `Subset()` function.

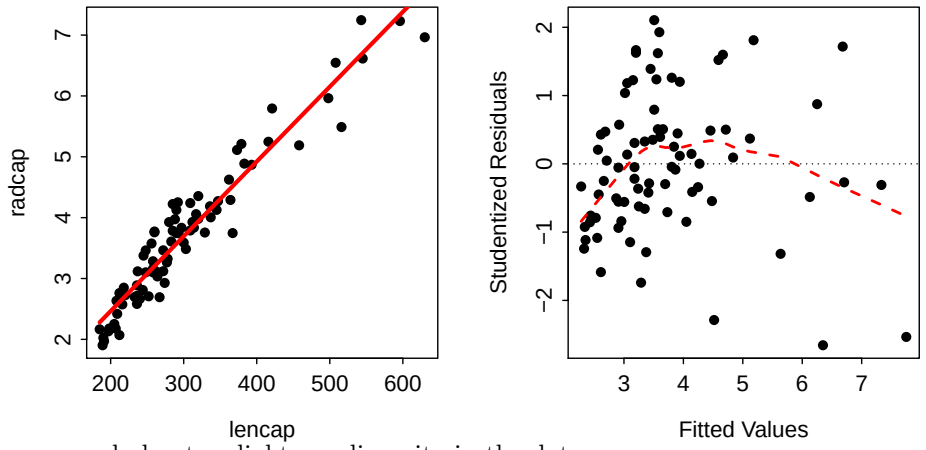
2. Is "plus-growth" recorded for your chosen data? Explain.

   ```
   > head(df,n=3)
        species gear        lake yearcap fish agecap lencap  anu1  anu2 anu3 anu4 anu5
   1135     WAE  All Lake Shetek    1998  155      1    185 1.304 2.165   NA   NA   NA
   1136     WAE  All Lake Shetek    1998  153      1    189 1.150 1.903   NA   NA   NA
   1137     WAE  All Lake Shetek    1998  171      1    190 1.209 2.024   NA   NA   NA
        anu6 anu7 anu8 anu9 anu10 anu11 anu12 radcap
   1135   NA   NA   NA   NA    NA    NA    NA  2.165
   1136   NA   NA   NA   NA    NA    NA    NA  1.903
   1137   NA   NA   NA   NA    NA    NA    NA  2.024
   > tail(df,n=3)
        species gear        lake yearcap fish agecap lencap  anu1  anu2  anu3  anu4
   1213     WAE  All Lake Shetek    1998  179      5    543 2.627 4.600 5.878 6.765
   1214     WAE  All Lake Shetek    1998  140      6    596 3.275 4.841 5.743 6.282
   1215     WAE  All Lake Shetek    1998  141      6    630 2.214 3.762 5.081 5.813
         anu5  anu6  anu7 anu8 anu9 anu10 anu11 anu12 radcap
   1213 7.123 7.245    NA   NA   NA    NA    NA    NA  7.245
   1214 6.632 6.947 7.231   NA   NA    NA    NA    NA  7.231
   1215 6.304 6.753 6.964   NA   NA    NA    NA    NA  6.964
   ```

"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.

3. Fit the regression required for use with the scale-proportional-hypothesis model. Save the intercept and slope for use below. Comment on the fit of this model.

```
> lm.SL <- lm(radcap~lencap,data=df)
> coef(lm.SL)
(Intercept)      lencap
   0.005703    0.012293
> ( a <- coef(lm.SL)[1] )
(Intercept)
   0.005703
> ( b <- coef(lm.SL)[2] )
 lencap
0.01229
> fitPlot(lm.SL,main="")
> residPlot(lm.SL,loess=TRUE)
```



I am a bit concerned about a slight non-linearity in the data.

4. Use the scale-proportional-hypothesis method to back-calculate length-at-age for your data.

```
> df2 <- gReshape(df,in.pre="anu",last.plus="agecap")
> df2 <- within(df2, lenSPH <- (anu/radcap)*(lencap+(a/b))-(a/b))
> view(df2)
    species gear        lake yearcap fish agecap lencap radcap age   anu lenSPH
8       WAE  All Lake Shetek    1998   48      1    207  2.177   1 1.201 114.03
25      WAE  All Lake Shetek    1998   83      2    248  3.102   1 1.032  82.21
116     WAE  All Lake Shetek    1998   26      2    267  2.693   2 2.092 207.30
128     WAE  All Lake Shetek    1998   99      2    290  4.127   2 3.430 240.95
146     WAE  All Lake Shetek    1998  106      2    362  4.625   2 3.893 304.61
148     WAE  All Lake Shetek    1998   93      2    367  3.747   2 3.239 317.17
```

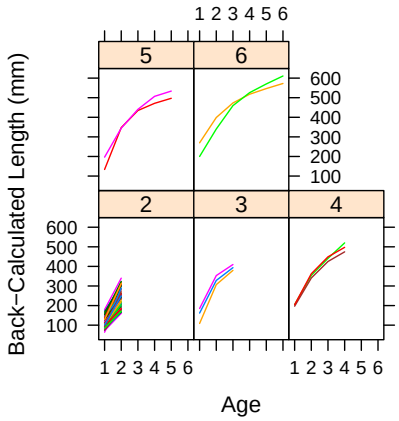
5. Compute the mean "back-calculated" length-at-age.

```
> Summarize(lenSPH~age,data=df2,digits=2)
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 81 123.6 37.28  64  96.9    115 144 270        0
2   2 68 252.6 59.64 161 204.0    249 302 399        0
```

```
3   3 10 430.5 29.12 380 413.0    438 447 473        0
4   4  7 501.9 21.97 471 486.0    507 519 526        0
5   5  4 536.9 30.60 497 525.0    540 552 570        0
6   6  2 591.7 27.05 573 582.0    592 601 611        0
```

6. Compute the mean "back-calculated" length-at-age for different ages-at-capture (and, perhaps, construct a plot). Any interesting observations from this summary?

```
> Summarize(lenSPH~age*agecap,data=df2,digits=2)
Warning:  To continue, variable(s) on RHS of formula were converted to a factor.
   age agecap  n  mean    sd min    Q1 median  Q3 max percZero
1    1      1 13 122.0 14.55 107 113.0    114 126 153        0
2    1      2 58 113.2 28.32  64  90.3    109 136 179        0
3    2      2 58 236.2 47.23 161 195.0    238 263 340        0
4    1      3  3 151.9 39.14 109 135.0    162 173 185        0
5    2      3  3 329.4 23.14 307 318.0    328 341 353        0
6    3      3  3 394.3 14.94 380 387.0    394 402 410        0
7    1      4  3 202.1  4.54 197 201.0    205 205 205        0
8    2      4  3 352.6 10.92 341 348.0    354 358 362        0
9    3      4  3 438.3 12.04 425 433.0    441 445 449        0
10   4      4  3 497.2 22.87 474 486.0    497 509 520        0
11   1      5  2 165.1 44.51 134 149.0    165 181 197        0
12   2      5  2 345.9  1.95 345 345.0    346 347 347        0
13   3      5  2 437.5  4.10 435 436.0    438 439 440        0
14   4      5  2 489.1 25.19 471 480.0    489 498 507        0
15   5      5  2 515.4 26.03 497 506.0    515 525 534        0
16   1      6  2 234.8 49.29 200 217.0    235 252 270        0
17   2      6  2 369.5 41.54 340 355.0    369 384 399        0
18   3      6  2 466.4  9.73 459 463.0    466 470 473        0
19   4      6  2 521.8  5.68 518 520.0    522 524 526        0
20   5      6  2 558.4 16.74 547 552.0    558 564 570        0
21   6      6  2 591.7 27.05 573 582.0    592 601 611        0
> library(lattice)
> xyplot(lenSPH~age|factor(agecap),groups=factor(fish),data=Subset(df2,agecap>1),
        type="l",ylab="Back-Calculated Length (mm)",xlab="Age")
```



Depends on chosen data.

7. (*Time Permitting*) Repeat the previous questions but using the body-proportional-hypothesis. How different are the summary results?

```
> lm.LS <- lm(lencap~radcap,data=df)
> ( c <- coef(lm.LS)[1] )
```

```
(Intercept)
      23.47
> ( d <- coef(lm.LS)[2] )
radcap
 74.94
> df2 <- within(df2, lenBPH <- lencap*(c+d*anu)/(c+d*radcap))
> Summarize(lenBPH~age,data=df2,digits=2)
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 81 137.5 36.77  82.7 112    128 156 283        0
2   2 68 257.6 59.83 169.0 209    252 306 407        0
3   3 10 434.4 30.56 382.0 416    441 451 478        0
4   4  7 504.0 22.75 473.0 487    508 521 530        0
5   5  4 538.3 31.54 497.0 525    541 555 573        0
6   6  2 592.6 26.94 574.0 583    593 602 612        0
> Summarize(lenBPH~age*agecap,data=df2,digits=2)
Warning:  To continue, variable(s) on RHS of formula were converted to a factor.
   age agecap  n  mean    sd   min  Q1 median  Q3 max percZero
1    1      1 13 132.0 12.93 119.0 124    127 135 160        0
2    1      2 58 127.6 26.95  82.7 106    125 150 192        0
3    2      2 58 240.7 46.22 169.0 200    241 267 343        0
4    1      3  3 167.5 37.88 126.0 151    175 188 201        0
5    2      3  3 335.1 23.09 313.0 323    333 346 359        0
6    3      3  3 396.4 15.33 382.0 388    395 404 412        0
7    1      4  3 218.1  5.21 212.0 216    220 221 222        0
8    2      4  3 361.0 11.19 349.0 356    363 367 371        0
9    3      4  3 442.3 12.17 429.0 437    446 449 452        0
10   4      4  3 498.3 22.84 475.0 487    498 510 521        0
11   1      5  2 181.1 42.54 151.0 166    181 196 211        0
12   2      5  2 353.9  1.31 353.0 353    354 354 355        0
13   3      5  2 441.4  4.75 438.0 440    441 443 445        0
14   4      5  2 490.8 25.06 473.0 482    491 500 508        0
15   5      5  2 515.8 25.94 497.0 507    516 525 534        0
16   1      6  2 251.1 45.74 219.0 235    251 267 283        0
17   2      6  2 380.0 38.46 353.0 366    380 394 407        0
18   3      6  2 472.7  8.13 467.0 470    473 476 478        0
19   4      6  2 525.7  6.57 521.0 523    526 528 530        0
20   5      6  2 560.8 17.11 549.0 555    561 567 573        0
21   6      6  2 592.6 26.94 574.0 583    593 602 612        0
```

Depends on data choice.

8. (*Time Permitting*) Repeat the previous questions but use a different species and lake combination.