R Introduction/Review Assignment

The `RuffeBio.csv` file contains a variety of biological measurements from Ruffe (*Gymnocephalus cernuus*) captured in the St. Louis River Harbor. Use these data to answer the following questions.

```
> setwd("C:/aaaWork/Web/fishR/Courses/MNAFS2013/CourseMaterial")
> ruf <- read.csv("RuffeBio.csv",header=TRUE)
> str(ruf)
'data.frame': 40 obs. of  10 variables:
 $ fishID  : int  60 61 62 63 64 65 66 67 68 69 ...
 $ locShort: Factor w/ 1 level "St. Louis R. (2007)": 1 1 1 1 1 1 1 1 1 1 ...
 $ year    : int  2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 ...
 $ month   : int  9 9 9 9 9 9 9 9 9 9 ...
 $ day     : int  20 20 20 20 20 20 20 20 20 20 ...
 $ date    : Factor w/ 1 level "9/20/2007": 1 1 1 1 1 1 1 1 1 1 ...
 $ tl      : int  134 111 110 115 92 88 95 90 99 107 ...
 $ wt      : num  24.6 14.7 12.3 16 8.3 7.8 9.7 8.2 11.7 13 ...
 $ sex     : Factor w/ 3 levels "female","male",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ maturity: Factor w/ 3 levels "","immature",..: 3 3 2 3 3 3 3 3 3 3 ...
```

1. How many variables are in this data frame?
   There are 10 variables in this data frame.

2. Data was recorded for how many Ruffe?
   Data was recorded on 40 Ruffe in this data frame.

3. Show the vector of *tl* measures?

   ```
   > ruf$tl
    [1] 134 111 110 115  92  88  95  90  99 107  NA  99 102 105  90 102 114  NA  56  90
   [21] 101 109 110 111 101  95  84 105 120 104 102  99  84  87  81  81  65  42  NA 115
   ```

   All of the total length measurements are shown above.

4. What is the *tl* for the 17th Ruffe?

   ```
   > ruf$tl[17]
   [1] 114
   ```

   The *tl* for the 17th measured individual is 114 mm.

5. Show all information for the 20th Ruffe.

   ```
   > ruf[20,]
      fishID            locShort year month day      date tl  wt    sex maturity
   20     79 St. Louis R. (2007) 2007     9  20 9/20/2007 90 7.6 female   mature
   ```

   All of the information for the 20th Ruffe is shown above.

6. Create a new data frame of just female Ruffe. How many fish are in this data frame?

   ```
   > df1 <- Subset(ruf,sex=="female")
   ```

   There are 31 Ruffe in this data frame.

7. Create a new data frame excluding Ruffe of unknown sex. How many fish are in this data frame?

   ```
   > df2 <- Subset(ruf,sex!="unknown")
   ```

   There are 39 Ruffe in this data frame.

8. Create a new data frame of Ruffe greater than 110 mm. How many fish are in this data frame?

```
> df3 <- Subset(ruf,tl>110)
```

There are 7 Ruffe in this data frame.

9. Create a new data frame of Ruffe between 80 and 120 mm. How many fish are in this data frame?

```
> df4 <- Subset(ruf,tl>80 & tl<120)
```

There are 32 Ruffe in this data frame.

10. Create a new data frame of male Ruffe less than 100 mm. How many fish are in this data frame?

```
> df5 <- Subset(ruf,sex=="male" & tl<100)
```

There are 5 Ruffe in this data frame.

11. Create a new variable in the original data frame that is Fulton's condition factor (i.e., the weight of the fish divided by the cubed length of the fish multiplied by 100000).

```
> ruf$fult <- ruf$wt/(ruf$tl^3)*100000
```

12. Compute the mean Fulton's condition factor.

```
> Summarize(~fult,data=ruf,digits=3)
       n     mean       sd      min       Q1   median       Q3      max percZero
  37.000    1.061    0.105    0.837    1.010    1.070    1.120    1.310    0.000
```

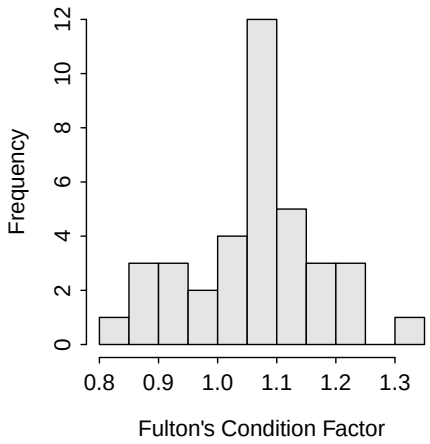
The mean Fulton's condition factor is 1.061 Ruffe in this data frame.

13. Compute the mean Fulton's condition factor separately for males and females.

```
> Summarize(fult~sex,data=ruf,digits=3)
      sex  n  mean    sd   min    Q1 median   Q3  max percZero
1  female 30 1.070 0.109 0.837 1.040  1.070 1.13 1.31        0
2    male  6 1.017 0.088 0.938 0.949  0.986 1.07 1.15        0
3 unknown  1 1.080    NA 1.080 1.080  1.080 1.08 1.08        0
```

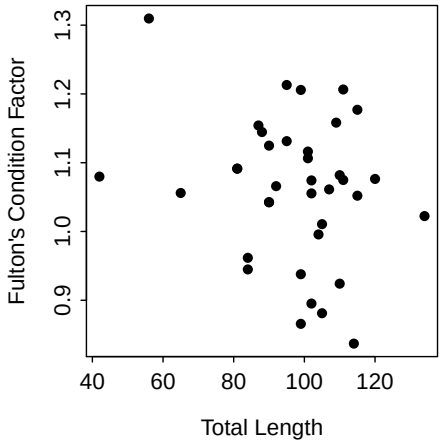
14. Create a histogram of Fulton's condition factor.

```
> hist(~fult,data=ruf,xlab="Fulton's Condition Factor",col="gray90")
```



15. Create a scatterplot of the relationship between Fulton's condition factor and total length.

```
> plot(fult~tl,data=ruf,xlab="Total Length",ylab="Fulton's Condition Factor",pch=16)
```



16. Fit a model that can be used to predict Fulton's condition factor from total length. Find the model coefficients and r^2 value and construct a plot that shows the fitted model.

```
> lm1 <- lm(fult~tl,data=ruf)
> summary(lm1)

Call:
lm(formula = fult ~ tl, data = ruf)

Residuals:
    Min      1Q  Median      3Q     Max
-0.2005 -0.0559  0.0134  0.0603  0.1899

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.199293   0.097239   12.33  2.7e-14
tl          -0.001420   0.000985   -1.44     0.16

Residual standard error: 0.104 on 35 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.056,Adjusted R-squared:  0.029
F-statistic: 2.08 on 1 and 35 DF,  p-value: 0.158
> plot(fult~tl,data=ruf,xlab="Total Length",ylab="Fulton's Condition Factor",pch=16)
> b <- coef(lm1)[1]
> m <- coef(lm1)[2]
> curve(m*x+b,from=40,to=135,add=TRUE,col="red",lwd=2)
```