```
#=====================================================================
#  file:          c:\Projects\CBP\Rcourse\snook.r
#  function:      length weight regression for snook
#
#  programmer:    Elgin S. Perry, Ph. D.
#
#  date:          5/11/2014
#
#  address:       2000 Kings Landing Rd.
#                 Huntingtown,  Md.  20639
#
#  voice phone:   (410)535-2949
#  email:         EPERRY@chesapeake.net
#=====================================================================

#install.packages()
#library(lattice) #Used for contour plots [contourplot()]
#library(nlme)    #used for gam Mixed model [gamm()]
#library(MASS)    #used for glm Mixed model [glmmPQL()]
library(mgcv)    #Wood's gam package
#library(chron)   #date functions
#library(doBy)    # Allows "BY processing similar to SAS
#library(FitAR)   #AR package from McLeod and Zhang
#library(Hmisc)   #stat function by Frank Harrell
#library(cluster) #cluster analysis routines
options(stringsAsFactors = FALSE)

source("C:/Projects/Rtp/dfsum.r")
source("C:/Projects/Rtp/RTF.r")

#  be sure to change \ to /
ProjRoot <- 'c:/Projects/CBP/Rcourse/'
setwd(ProjRoot);
RTFout <- paste(ProjRoot,"RTFexample.rtf",sep='')


datafile <- paste(ProjRoot,"snook.tdf",sep='');
a <- count.fields(datafile, sep = "\t", quote = "\"'", skip = 1,
             blank.lines.skip = TRUE, comment.char = "#")
range(a)
rbind(1:length(a),a)

snook <- read.table(datafile, header=TRUE, sep="\t", na.strings="NA", dec=".",
strip.white=TRUE,stringsAsFactors = FALSE)
dfsum(snook)
snook[snook$length==40 & snook$water.body=='Atlantic'&snook$season=="May-Oct",'wgt.mean'] <-
NA
#[1] "length"     "water.body" "season"     "wgt.mean"   "wgt.min"    "wgt.max"

snook$wb.col <- ifelse(snook$water.body=='Gulf','green','blue')
plot(snook$length,snook$wgt.mean)

snook$log.wgt <- log(snook$wgt.mean)

plot(snook$length,snook$log.wgt)


lm1 <- lm(log.wgt~length,data=snook)
plot(snook$length,snook$log.wgt)

psnook <- data.frame(length = 20:44)
lines(psnook$length,predict(lm1,newdata=psnook),lwd=2)

gam1 <- gam(wgt.mean ~ s(length),family=gaussian(),data=snook)
```

```
plot(snook$length,snook$wgt.mean)
lines(psnook$length,predict(gam1,newdata=psnook),lwd=2)


lm2 <- lm(log.wgt~length+as.factor(water.body),data=snook)
summary(lm2)

psnook.A <- data.frame(length = 20:44,water.body='Atlantic')
psnook.G <- data.frame(length = 20:44,water.body='Gulf')
plot(snook$length,snook$log.wgt,col=snook$wb.col)
lines(psnook.A$length,predict(lm2,newdata=psnook.A),lwd=2,col='blue')
lines(psnook.A$length,predict(lm2,newdata=psnook.G),lwd=2,col='green')

gam2 <- gam(wgt.mean ~ s(length)+water.body,family=gaussian(),data=snook)
summary(gam2)
plot(snook$length,snook$wgt.mean,col=snook$wb.col)
lines(psnook.A$length,predict(gam2,newdata=psnook.A),lwd=2,col='blue')
lines(psnook.G$length,predict(gam2,newdata=psnook.G),lwd=2,col='green')

snook$wbf <- as.factor(snook$water.body)
gam3 <- gam(wgt.mean ~ wbf+s(length)+ s(length, by=wbf),family=gaussian(),data=snook)
summary(gam3)
psnook.A <- data.frame(length = 20:44,wbf='Atlantic')
psnook.G <- data.frame(length = 20:44,wbf='Gulf')
plot(snook$length,snook$wgt.mean,col=snook$wb.col)
lines(psnook.A$length,predict(gam3,newdata=psnook.A),lwd=2,col='blue')
lines(psnook.G$length,predict(gam3,newdata=psnook.G),lwd=2,col='green')
```