

```

# THIS IS A REALLY GREAT R SCRIPT THAT I CAN SAVE FOR FUTURE USE
# ANYTHING TO THE RIGHT OF THE POUND SIGN IS IGNORED BY R
# 17 APRIL 2008 COLVIN.MIKE@GMAIL.COM

# R CAN BE USED LIKE A CALCULATOR
# JUST PRESS CTL-R TO SEND A LINE TO THE R CONSOLE
2+2

# R IS OBJECT ORIENTED
# THIS CREATES AN OBJECT X
X <- 2+2

# LETS SEE WHAT X IS
X

# R ALSO HAS FUNCTIONS LIKE MEAN
# AND SOME ASSOCIATED HELP FOR THE FUNCTIONS
?mean          # Works! mean is an R command
help(mean)     # Works! mean is an R command
help(regression) # Fails! regression is not an R command
help.search("regression") # Well that works alright
RSiteSearch("regression") # We can also look at the R archive on the web

# ENOUGH OF THIS CRAP LETS DO SOMETHING IMPORTANT
# LIKE GETTING DATA INTO R
# IMPORTING *.CSV, *.TXT FILES
# NOTE- THE / IN THE FILE NAME ARE BACKWARDS (THIS IS DUE TO MULTIPLATFORM
# COMPATABILITY)
dataset<-read.csv("C:/Documents and Settings/Mike/My Documents/rClass/studydata.csv")
head(dataset) # Lets look at the first few records of the dataset

# THAT IS AN AWFUL AMOUNT TO TYPE
# SETTING A WORKING DIRECTORY WILL ALLEVIATE TYPING LONG FILENAMES
setwd("C:/Documents and Settings/Mike/My Documents/rClass")

# NOW WE CAN READ IN THE DATASET USING THE WORKING DIRECTORY
dataset<-read.csv("./studydata.csv")

# WE CAN ALSO READ FROM ACCESS OR EXCEL
require(RODBC)
channel<- odbcConnectAccess("./mydbase.mdb")
channel2<- odbcConnectExcel("./studyData.xls", readOnly=F)# you need to specify readOnly=f to save
again to the excel file

datasetAcces<-sqlFetch(channel, "Studydata")
datasetExcel<-sqlFetch(channel2, "studydata")

# WE CAN ALSO USE HEAD TO LOOK AT THE FIRST FEW ROWS OF THE DATASET
head(dataset)
# WE CAN ALSO LOOK AT THE FIRST 20 ROWS OF OUR DATASET
dataset[c(1:20),]

# WHAT IS THE DIMENSIONS OF THE DATASETS?
dim(dataset)
dim(datasetAcces)
dim(datasetExcel)

# SUMMARIZE THE DATASET

```

```

summary(dataset)
summary(datasetAcces)
summary(datasetExcel)

# LETS DO A QUICK QA TO MAKE SURE THE 3 DATASETS ARE THE SAME
sum(dataset$abundance)
sum(datasetAcces$abundance)
sum(datasetExcel$abundance) # ALL THE SUMS SHOULD BE THE SAME

length(dataset$site)
length(datasetAcces$site)
length(datasetExcel$site) # ALL HAVE THE SAME NUMBER OF OBSERVATIONS

# LETS LOOK AT THE WHOLE DATASET
plot(dataset)
plot(datasetAcces)
plot(datasetExcel)

# LETS SAVE THAT PLOT AS A JPEG TO PUT ON YOUR FRIDGE
# YOU CAN FIND THE PLOT IN YOUR WORKING DIRECTORY
savePlot("./myCoolPlot.jpg", type="jpg")
# OR SAVE IT TO YOUR DESKTOP
savePlot("C:/DOCUMENTS AND SETTINGS/Mike/Desktop/myCoolPlot.jpg", type="jpg")
# YOU CAN ALSO DO PDF, EPS, METAFILES, JUST TYPE ?savePlot for more info
savePlot("./myCoolPlot.pdf", type="pdf")
savePlot("./myCoolPlot.eps", type="eps")

# PLOTTING DATA
plot(abundance~area,dataset)

# I THINK THE POINTS NEED TO BE BIGGER AND FILLED
# CEX IS A MAGNIFICATION FACTOR AND PCH 1-25 ARE ALL SORTS OF SYBOLS
plot(abundance~area,dataset,cex=2,pch=19)
# CRAP MY ADVISOR WANTS BIGGER FONT FOR THE LABELS
# CEX.AXIS IS A MAGNIFICATION FOR THE AXIS FONT
# CEX.LAB IS A MAGFICIATION FOR THE LABEL FONT
plot(abundance~area,dataset,cex=1.5,pch=19,cex.axis=1.5,cex.lab=1.5)

# THE AXIS LABELS REALLY NEED TO BE CAPITALIZED AND ADD A MAIN TITLE
# XLAB IS A CUSTOM X LABEL
# YLAB IS A CUSTOM Y LABEL
plot(abundance~area,dataset,cex=1.5,pch=19,cex.axis=1.5,cex.lab=1.5,
      xlab="Area",ylab="Abundance", main="My really great plot")

# PERSONALLY, I DISLIKE HAVING THE NUMBERS ON THE YAXIS ORIENTED PARRELLELL
# TO THE AXIS
# LAS=1 FIXES THAT
plot(abundance~area,dataset,cex=1.5,pch=19,cex.axis=1.5,cex.lab=1.5,
      xlab="Area",ylab="Abundance", main="My really great plot",las=1)

# MAYBE THERE IS A RELATIONSHIP BETWEEN AREA AND ABUNDANCE
lm(abundance~area,dataset) # GIVES US THE PARAMETER ESTIMATES
summary(lm(abundance~area,dataset))# DOESN'T LOOK LIKE IT!

# SINCE R IS OBJECT ORIENTED WE CAN NAME THAT MODEL FIT SOMETHING
# FOR LATER USE
fit.area<- lm(abundance~area,dataset)
summary(fit.area) # STILL WORKS

```

```

# LETS TRY ANOTHER MODEL
fit.area.canopy<- lm(abundance~area + canopy,dataset)
summary(fit.area.canopy) # BETTER!

# WHAT ABOUT AN INTERACTION?
fit.area.canopy.interaction<- lm(abundance~area*canopy,dataset)
summary(fit.area.canopy.interaction) # WORKS!

# THIS IS ALL WELL AND GOOD BUT LETS CHECK SOME ASSUMPTIONS
plot(fit.area)
plot(fit.area.canopy)
plot(fit.area.canopy.interaction)

# IF WE NEEDED TO WE COULD LOG TRANSFORM ABUNDANCE
fit.area.canopy.log<- lm(log(abundance)~area + canopy,dataset)
summary(fit.area.canopy.log)

# WELL AIC IS POPULAR NOW WHAT ABOUT THOSE?
AIC(fit.area)
AIC(fit.area.canopy)
AIC(fit.area.canopy.interaction)

# CAN WE DO THAT MODEL SELECTION STUFF IN R?
# C CAN COMBINE A SERIES OF OBJECTS INTO A VECTOR
# DATA.FRAME CAN MAKE A DATAFRAME FOR R TO USE AND EXPORT
# "" ARE USED TO DENOTE TEXT
modelSelection<- data.frame(
  model=c("fit.area","fit.area.canopy","fit.area.canopy.interaction"),
  AIC=c(AIC(fit.area),AIC(fit.area.canopy),AIC(fit.area.canopy.interaction)))

# LETS LOOK AT OUR TABLE
modelSelection

# LETS WORK A LITTLE INFORMATION THEORETIC MAGIC
# LETS CALCULATE THE DELTA AIC
modelSelection$deltaAic<- modelSelection$AIC-min(modelSelection$AIC)
modelSelection$lik<-exp(-0.5*modelSelection$deltaAic)
modelSelection$weight<-modelSelection$lik/sum(modelSelection$lik)

# A MODEL SELECTION
modelSelection

# LETS EXPORT THIS SO WE CAN PUT IT IN OUR THESIS
write.csv(modelSelection,file="./modelSelection.csv")

# OR IT MIGHT MAKE SENSE TO KEEP IT IN THE SAME EXCEL FOLDER AS THE DATA
# FIX THIS
sqlSave(channel2, modelSelection)

# OR EVEN IN OUR DATABASE
sqlSave(channel, modelSelection)

# I WONDER IF THERE IS SOMETHING GOING ON THAT WE AREN'T SEEING HERE?
# TYPE="N" PRODUCES NO PLOT SO WE CAN ADD POINTS OURSELVES
plot(abundance~area,dataset,cex=1.5,pch=19,cex.axis=1.5,cex.lab=1.5,
      xlab="Area",ylab="Abundance", main="My really great plot",las=1,type="n")

```

```

# WE CAN ADD POINTS EASILY WITH DIFFERENT COLORS AND SHAPES (pch)
points(abundance~area,dataset,subset=lake=="lakeA",col="red",pch=15)
points(abundance~area,dataset,subset=lake=="lakeB",col="blue",pch=16)
points(abundance~area,dataset,subset=lake=="lakeC",col="green",pch=17)

# OR MAYBE AS A BOXPLOT
boxplot(abundance~lake, dataset,las=1)
boxplot(abundance~lake, dataset,las=1,col="gray")

# WHAT ABOUT SUBSETTING A DATASET?
newdataset<- subset(dataset, lake=="lakeA")
newdataset

# SOME MORE COMPLEX SUBSETTING
newdataset<- subset(dataset, (lake %in% c("lakeA","lakeB") & canopy <0.5))
newdataset
newdataset<- subset(dataset, lake != "lakeA")
newdataset

# PLOTTING A SHAPEFILE USING THE MAPTOOLS PACKAGE
install.packages("maptools")

require(maptools)
clearLake<- readShapePoly("c:/gis/clearlakenad83.shp",proj4string=CRS("+proj=utm +zone=15
+datum=NAD83"))
plot(clearLake,axes=T, las=1)
par(mar=c(5,5,1,1))
plot(clearLake,axes=T, las=1)

# LETS GET SOME POINTS TO PUT ON THE MAP
trawling<-read.csv("http://www.public.iastate.edu/~mcolvin/trawl.csv")
points( startUTMy~startUTMx, trawling)

# IT IS A LAKE, LETS MAKE IT BLUE
plot(clearLake,axes=T, las=1, col="blue")
points( startUTMy~startUTMx, trawling,pch=16)
savePlot("./trawlSites",type="pdf")

# OTHER USEFUL PACKAGES
install.packages("spatstat")
install.packages("vegan")
install.packages("RODBC") # ALLOWS R TO TALK TO EXCEL AND ACCESS
install.packages("lme4")# LINEAR AND NON LINEAR MIXED EFFECTS MODELS
install.packages("R2WinBUGS")# USE WINBUGS FROM R
install.packages("demogR") # Analysis of age-structured demographic models

# AN EXTENSION RMARK IS AVIALABLE HERE TO RUN MARK FROM R
# http://www.phidot.org/software/mark/rmark/
# WISP: Wildlife Simulation Package
# http://www.ruwpa.st-and.ac.uk/estimating.abundance/WiSP/index.html

```