# read data from a file called worms.csv to create a

# dataframe called worms

worms <- read.csv("c:\\temp\\worms.csv")

names(worms)

attach(worms)

worms

# all rows, just columns 1 to 3

worms[,1:3]

# all columns just rows 5 to 15

worms[5:15,]

# all columns but only selected rows (area > 3 and slope < 3)

worms[Area>3 & Slope <3,]

# sort the rows by increasing area

worms[order(Area),]

# only the columns with numeric data

worms[order(Area),c(2,3,5,7)]

# sort into decending order with just two columns

worms[rev(order(worms[,5])),c(5,7)]

# using tapply with a specified dataframe

with(worms,tapply(Worm.density,Vegetation,mean))

# using aggregate to summarise multiple variables by facor levels

aggregate(worms[,c(2,3,5,7)],list(Vegetation),mean)

aggregate(worms[,c(2,3,5,7)],list(Community=Vegetation),mean)

# multiple explanatory variables

aggregate(worms[,c(2,3,5,7)],

list(Moisture=Damp,Community=Vegetation),mean)

# aggregate and tapply compared

with(worms,tapply(Slope,list(Damp,Vegetation),mean))

# plotting your data

data <- read.csv("c:\\temp\\das.csv")

attach(data)

head(data)

# finding the identity of the outlier

which(y > 10)

y[50]

# plots with categorical explanatory variables

yields <- read.csv("c:\\temp\\fertyield.csv")

attach(yields)

head(yields)

table(treatment)

which(treatment == "nitogen")

# scatterplots

data <- read.csv("c:\\temp\\scatter.csv")

attach(data)

head(data)

plot(x,y,pch=21,bg="red")

# box and whisker plots

data <- read.csv("c:\\temp\\weather.data.csv")

attach(data)

head(data)

plot(factor(month),upper)

# data for coplot

data <- read.csv("c:\\temp\\coplot.csv")

attach(data)

head(data)

# scale the plotting area to accommodate two plots side by side

windows(7,4)

par(mfrow=c(1,2))

plot(x,y)

plot(z,y)

# using coplot

windows(7,7)

coplot(y~x|z,pch=16,panel=panel.smooth)

# factorial data

data <- read.csv("c:\\temp\\np.csv")

attach(data)

head(data)

windows(7,4)

par(mfrow=c(1,2))

plot(nitrogen,yield,main="N")

plot(phosphorus,yield,main="P")

tapply(yield,list(nitrogen,phosphorus),mean)

barplot(tapply(yield,list(nitrogen,phosphorus),mean),

beside=TRUE,xlab="phosphorus")

legend(locator(1),legend=c("no","yes"),title="nitrogen",

fill=c("black","lightgrey"))