S-plus Examples

Finding Class Datasets

See what's in the class directory and in the WWW directory

```
tree1:~> ls /usr/class/stats202
DATA  WWW
tree1:~> ls /usr/class/stats202/DATA
housing.data  housing.info  income.data  income.info  spam.data
```

Copy the Boston housing data into your home directory:

```
tree1:~/stats202> cp /usr/class/stats202/DATA/housing.data housing.data
```

Read Data Into S-Plus

```
S-PLUS : Copyright (c) 1988, 2000 MathSoft, Inc.
S : Copyright Lucent Technologies, Inc.
Version 6.0 Release 1 for Sun SPARC, SunOS 5.6 : 2000
Working data will be in /afs/ir/users/i/b/ibelit/MySwork

> boston _ read.table("housing.data", header=T)
```

Getting Help

The best way to get help in S-Plus is to open help window at the beginning of you S-Plus session:

```
> help.start()
```

Working With Data

```
> #You can make comments with '#' in S-Plus
> dim(boston)
[1] 506  14
> names(boston)
 [1] "crim"    "zn"      "indus"   "chas"    "nox"     "rm"      "age"
 [8] "dis"     "rad"     "tax"     "ptratio" "b"       "lstat"   "medv"
> #the following command gives you a summary for each variable in the data:
> summary(boston)
       crim                zn              indus             chas
```

```
    Min.: 0.006320      Min.:  0.00      Min.: 0.460      Min.:0.00000
 1st Qu.: 0.082045   1st Qu.:  0.00   1st Qu.: 5.190   1st Qu.:0.00000
  Median: 0.256510    Median:  0.00    Median: 9.690    Median:0.00000
    Mean: 3.613524      Mean: 11.36      Mean:11.137      Mean:0.06917
 3rd Qu.: 3.677083   3rd Qu.: 12.50   3rd Qu.:18.100   3rd Qu.:0.00000
    Max.:88.976200      Max.:100.00      Max.:27.740      Max.:1.00000
> #Access parts of data
> boston[12:15, 1:3]
      crim   zn indus
12 0.11747 12.5  7.87
13 0.09378 12.5  7.87
14 0.62976  0.0  8.14
15 0.63796  0.0  8.14
> boston$crim
 [1]  0.00632  0.02731  0.02729  0.03237  0.06905  0.02985  0.08829  0.14455
 [9]  0.21124  0.17004  0.22489  0.11747  0.09378  0.62976  0.63796  0.62739
[17]  1.05393  0.78420  0.80271  0.72580  1.25179  0.85204  1.23247  0.98843
 ...
> sqrt(var(boston$rm))  #standard deviation of the number of rooms
[1] 0.7026171
```

Graphics

```
> #open a graphics window
> motif()

> #make a histogram for the variable "indus":
> hist(boston$indus)

> #get a density estimate and plot it:
> d _ density(boston$indus)
> plot(d$x, d$y, type="l")
> title("Proportion of Non-Retail Business Acres")

> #split the graphics window into 2 parts (1 by 2):
> par(mfrow=c(1,2))

> #place two histograms on same page:
> hist(boston$medv[boston$chas==1])
> title("tract bounds river")
> hist(boston$medv[boston$chas==0])
> title("otherwise")

> par(mfrow=c(1,1))  #back to one window
> #make a scatterplot of medv vs lstat:
```

```
> plot(boston$lstat, boston$medv)
> #add a smoothing spline to the existing graph:
> lines(smooth.spline(boston$lstat, boston$medv))
> title("medv vs lstat")

> #make pairwise scatterplots for selected variables:
> my.variables _ c(1,3,6,7,13,14)
> pairs(boston[, my.variables])

> #make 4 boxplots side by side
> #notice outliers in crime rates!
> boxplot(boston$indus, boston$lstat, boston$crim, boston$medv,
      names = c("indus","lstat", "crim", "medv"), main = "Boxplots")

> #save graph in a file
> printgraph(file="boxplot.ps")

> #the function brush allows interaction with the scatterplolt matrix:
> brush(as.matrix(boston))
> #try rotating points in 3D and highlighting them to see how they
> #are linked in the scatterplots.
```

Basic Data Manipulation

```
> #create your own data
> x_c(1,2,3,4)
> x
[1] 1 2 3 4
> x*3
[1]  3  6  9 12
> y _ matrix(c(1,2,3,4),2,2)
> y
     [,1] [,2]
[1,]    1    3
[2,]    2    4
> y[2,1]
[1] 2
> 5*y
     [,1] [,2]
[1,]    5   15
[2,]   10   20
> z _ matrix(c(4,5,6,7,8,9),3,2)
> z
     [,1] [,2]
[1,]    4    7
```

```
[2,]    5    8
[3,]    6    9
> t(z) #transpose
     [,1] [,2] [,3]
[1,]    4    5    6
[2,]    7    8    9
> z %*% y
     [,1] [,2]
[1,]   18   40
[2,]   21   47
[3,]   24   54
>
> #Note key difference:
> y*y #element by element multiplication
     [,1] [,2]
[1,]    1    9
[2,]    4   16
> y %*% y #matrix multiplication
     [,1] [,2]
[1,]    7   15
[2,]   10   22
```

Quitting S-Plus

Use function q() to quit S-Plus

```
> q()
```

Saving Your Work

Open a text editor window to save your commands and output. Cut and paste relevant code and output. If you are using emacs, just save the emacs file containing your S-Plus session.