

## S-plus tutorial

### Finding datasets:

See what's in my directory:

```
galton:> ls
WWW                pclab.ps          ratozone.dat
S-emacs.ps        getstarted.ps    ratozone         ratozone.doc
```

```
galton:> more ratozone.dat
```

```
41.0  10.1
38.4   6.1
24.9  20.4
25.9   7.3
21.9  14.3
18.3  15.5
13.1  -9.9
27.3   6.8
28.5  28.2
```

etc.

### Read the data into Splus:

```
galton:> Splus
S-PLUS : Copyright (c) 1988, 1996 MathSoft, Inc.
S : Copyright AT&T.
Version 3.4 Release 1 for Sun SPARC, SunOS 5.3 : 1996
Working data will be in /afs/ir/users/s/a/sandra/.Data
```

```
> rat_read.table("ratozone.dat")
```

```
> rat
```

```
      V1    V2
1  41.0  10.1
2  38.4   6.1
3  24.9  20.4
4  25.9   7.3
5  21.9  14.3
6  18.3  15.5
```

etc.

```
> ## make comments in your code with '#'
```

```
> #What's in my directory? All are saved from session to session.
> ls()
[1] ".Last.value" ".Random.seed" "rat"

> ##Housekeeping
> x_c(1,2,3)
> rm(x)
```

### Quitting S-plus:

```
> q()
```

### Getting Help:

Two ways to get help. Both give examples.

1. Open help window. Good idea to do this at start of session.

```
> help.start(gui="motif")
```

2. If you are in line edit mode, type help and topic in ().

```
> help(plot)
Plots - Generic function
DESCRIPTION:
    Creates a plot on the current graphics device.
    This function is generic (see Methods); method functions...
etc.
```

### Working with a dataset

```
> summary(rat)
      V1          V2
Min.   :-16.90   Min.   :-15.900
1st Qu.: 18.75   1st Qu.: -5.225
Median : 22.70   Median : 11.100
Mean   : 22.43   Mean    : 11.050
3rd Qu.: 26.95   3rd Qu.: 17.350
Max.   : 41.00   Max.    : 54.600
      NA's      : 1.000

> ##Reference case number 12
> rat[12,]
```

```
      V1 V2
12 21.8 14
```

```
> #Reassign case #12
> rat[12,]_c(21.9,15)
> rat #revised data
```

```
      V1    V2
1  41.0  10.1
```

```
...
```

```
11  17.4 -12.9
12  21.9  15.0 <--
13  15.4   6.6
```

```
...
```

```
> #name variables
> names(rat)_c("control","ozone")
> rat
```

```
      control ozone
1    41.0  10.1
2    38.4   6.1
3    24.9  20.4
```

```
...
```

Loading datasets with headers

```
tree0:> more ratozone2.dat
```

```
control ozone
  41.0  10.1
  38.4   6.1
  24.9  20.4
```

```
> rat<-read.data('ratozone2.dat', header = T)
```

```
> rat
```

```
      control ozone
1    41.0  10.1
2    38.4   6.1
3    24.9  20.4
```

```
...
```

```
> rat$control
```

```
[1] 41.0 38.4 24.9 25.9 21.9 18.3 13.1 27.3 28.5 -16.9 17.4 21.8
```

```

[13] 15.4 27.4 19.2 22.4 17.7 26.0 29.4 21.4 22.7 26.0 26.6
> rat$ozone
[1] 10.1 6.1 20.4 7.3 14.3 15.5 -9.9 6.8 28.2 17.9 -12.9 14.0
[13] 6.6 12.1 15.7 39.9 -15.9 54.6 -14.7 44.1 -9.0 -9.0 NA

> #The following command "attaches" the dataset, so you can work
> #with the individual variables in your commands. Without the
> #command below, Splus won't know that you are referencing individual
> #variables in the dataset.

> attach(rat)
> control
 1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
41 38.4 24.9 25.9 21.9 18.3 13.1 27.3 28.5 -16.9 17.4 21.8 15.4 27.4 19.2
 16 17 18 19 20 21 22 23
22.4 17.7 26 29.4 21.4 22.7 26 26.6
> summary(control)
Min. 1st Qu. Median Mean 3rd Qu. Max.
-16.9 18.75 22.7 22.43 26.95 41

> mean(control)
[1] 22.42609
> mean(ozone)
[1] NA
> mean(ozone, na.rm = T)
[1] 11.00909

> delta <- mean(control) - mean(ozone,na.rm = T)
> delta
[1] 11.417

> stdc <- sqrt(var(control))
> stdo <- sqrt(var(ozone,na.rm=T))
Error in var(ozone): missing values in x not allowed
Dumped
> ozone == "NA"
[1] F F F F F F F F F F F F F F F F F F T
> ozone <- ozone[ozone != "NA"] # gets rid of the 'NA' entries.
> ozone
[1] 10.1 6.1 20.4 7.3 14.3 15.5 -9.9 6.8 28.2 17.9 -12.9 14.0
[13] 6.6 12.1 15.7 39.9 -15.9 54.6 -14.7 44.1 -9.0 -9.0
> stdo <-sqrt(var(ozone))

> stdc
[1] 10.77675

```

```
> stdo
[1] 19.01711

> detach('rat')
> ## now you can't reference individual var's
> ## good to detach data if you plan to work with another
> ## dataset while you're in the same session.
```

## Graphics:

```
> # open a graphics window
> motif()
> boxplot(control,ozone,names = c("Controls","Ozone"),main='Weight Gains')
> # save graph in a file
> printgraph(file="boxplot.ps")

> # now the graph is saved in your account. Print as follows if
printer is Sweet1.
    elaine6:> lpr -Psweet1 boxplot.ps

> par(mfrow = c(2,1)) # Split the graphic window in two
> hist(control,xlab="Controls", xlim= c(-20,60))
> title('Weight Gains')
> hist(ozone,xlab="Ozone", xlim= c(-20,60))

> par(mfrow = c(1,1)) # Back to one window
> dev.off() # turn off the graphics device
```

## Basic Data Manipulation:

Create your own data/basic arithmetic

```
> 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

```

### Saving Your Work:

1. Open a text editor window to save your commands and output. Cut and paste relevant code and output.
2. If you know emacs, you can run Splus in emacs and save your session in emacs. This is called S-mode. Begin by typing Meta-x followed by S. On most machines, this is: Control-[ followed by S.