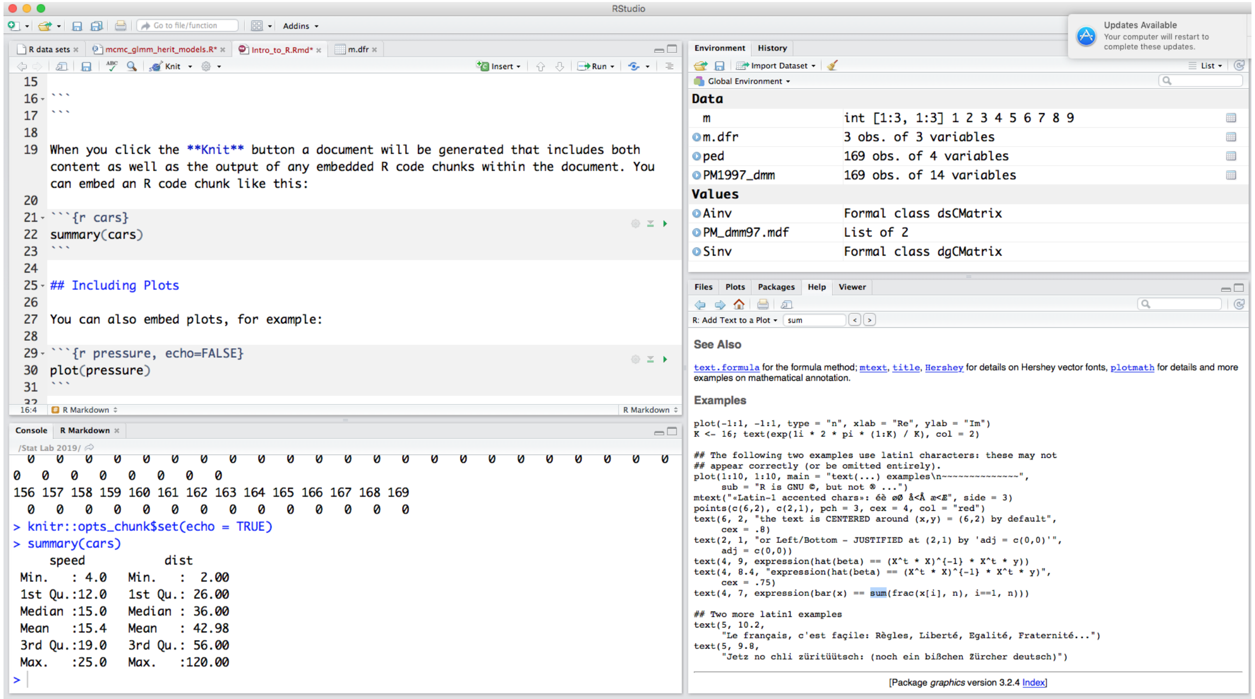
Introduction to R Studio

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com. I will be using R Markdown to develop R demos in class, and will share R Markdown code so that you can become familiar with it as well.

RStudio

There are a couple versions of R available, though I find students prefer RStudio because of its attractive IDE (Interactive Development Environment). After downloading and installing the appropriate version of R from its website (https://cran.r-project.org or https://www.rstudio.com), you can open R to begin; we will assume you have installed R Studio. When you first open R Studio, your screen will typically be divided into 4 frames.



```
```

When you click the **Knit** button a document will be generated that includes both
content as well as the output of any embedded R code chunks within the document. You
can embed an R code chunk like this:

```{r cars}
summary(cars)
```

Including Plots

You can also embed plots, for example:

```{r pressure, echo=FALSE}
plot(pressure)
```
```

```
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
0   0   0   0   0   0   0   0
156 157 158 159 160 161 162 163 164 165 166 167 168 169
  0   0   0   0   0   0   0   0   0   0   0   0   0   0
> knitr::opts_chunk$set(echo = TRUE)
> summary(cars)
     speed           dist       
 Min.   : 4.0   Min.   :  2.00  
 1st Qu.:12.0   1st Qu.: 26.00  
 Median :15.0   Median : 36.00  
 Mean   :15.4   Mean   : 42.98  
 3rd Qu.:19.0   3rd Qu.: 56.00  
 Max.   :25.0   Max.   :120.00  
> 
```

Data	
m	int [1:3, 1:3] 1 2 3 4 5 6 7 8 9
m.dfr	3 obs. of 3 variables
ped	169 obs. of 4 variables
PM1997_dmm	169 obs. of 14 variables
Values	
Ainv	Formal class dsCMatrix
PM_dmm97.mdf	List of 2
Sinv	Formal class dgCMatrix

See Also

text.formula for the formula method; mtext, title, Hershey for details on Hershey vector fonts, plotmath for details and more examples on mathematical annotation.

Examples

```
plot(-1:1, -1:1, type = "n", xlab = "Re", ylab = "Im")
K <- 16; text(exp(1i * 2 * pi * (1:K) / K), col = 2)

## The following two examples use latin1 characters: these may not
## appear correctly (or be omitted entirely).
plot(1:10, 1:10, main = "text(...) examples\n~~~~~~~~~~~~~~",
     sub = "R is GNU ©, but not ® ...")
mtext("«Latin-1 accented chars»: éè øØ å<Å æ<Æ", side = 3)
points(c(6,2), c(2,1), pch = 3, cex = 4, col = "red")
text(6, 2, "the text is CENTERED around (x,y) = (6,2) by default",
     cex = .8)
text(2, 1, "or Left/Bottom - JUSTIFIED at (2,1) by 'adj = c(0,0)'",
     adj = c(0,0))
text(4, 9, expression(hat(beta) == (X^t * X)^{-1} * X^t * y))
text(4, 8.4, "expression(hat(beta) == (X^t * X)^{-1} * X^t * y)",
     cex = .75)
text(4, 7, expression(bar(x) == sum(frac(x[i], n), i==1, n)))

## Two more latin1 examples
text(5, 10.2,
     "Le français, c'est façile: Règles, Liberté, Egalité, Fraternité...")
text(5, 9.8,
     "Jetz no chli züritüütsch: (noch ein bißchen Zürcher deutsch)")
```

[Package *graphics* version 3.2.4 Index]

The upper left-hand frame is used for code development, the lower left-hand frame is the console–as though R commands were being entered manually, the upper right-hand frame includes information about active R variables and data objects, while the lower right-hand screen includes plots, help and R packages.

A first R session

In this first R session, we will assume that we have saved tab-delimited data from the first lecture in the file Examp9p4.txt and placed it in our STAT 515 folder. To simplify reading the data set into R, select Session, then Set Working Directory->Choose Directory, browse to STAT 515 and select Open. R will now assume this is the location to read input files and place output files.
```

The commands below read the tab-delimited table, store it in a data frame, which is R's main data object, print the data frame (Example_1), and then print summary information about the variables in the data frame.

```
Example_1=read.table("Examp9p4.txt",header=T)
Example_1
```

```
##       Group Value
## 1       New    80
## 2  Standard    79
## 3       New    76
## 4  Standard    73
## 5       New    70
## 6  Standard    72
## 7       New    80
## 8  Standard    62
## 9       New    66
## 10 Standard    76
## 11      New    85
## 12 Standard    68
## 13      New    79
## 14 Standard    70
## 15      New    71
## 16 Standard    86
## 17 Standard    75
## 18      New    81
## 19 Standard    68
## 20      New    76
## 21 Standard    73
## 22 Standard    66
```

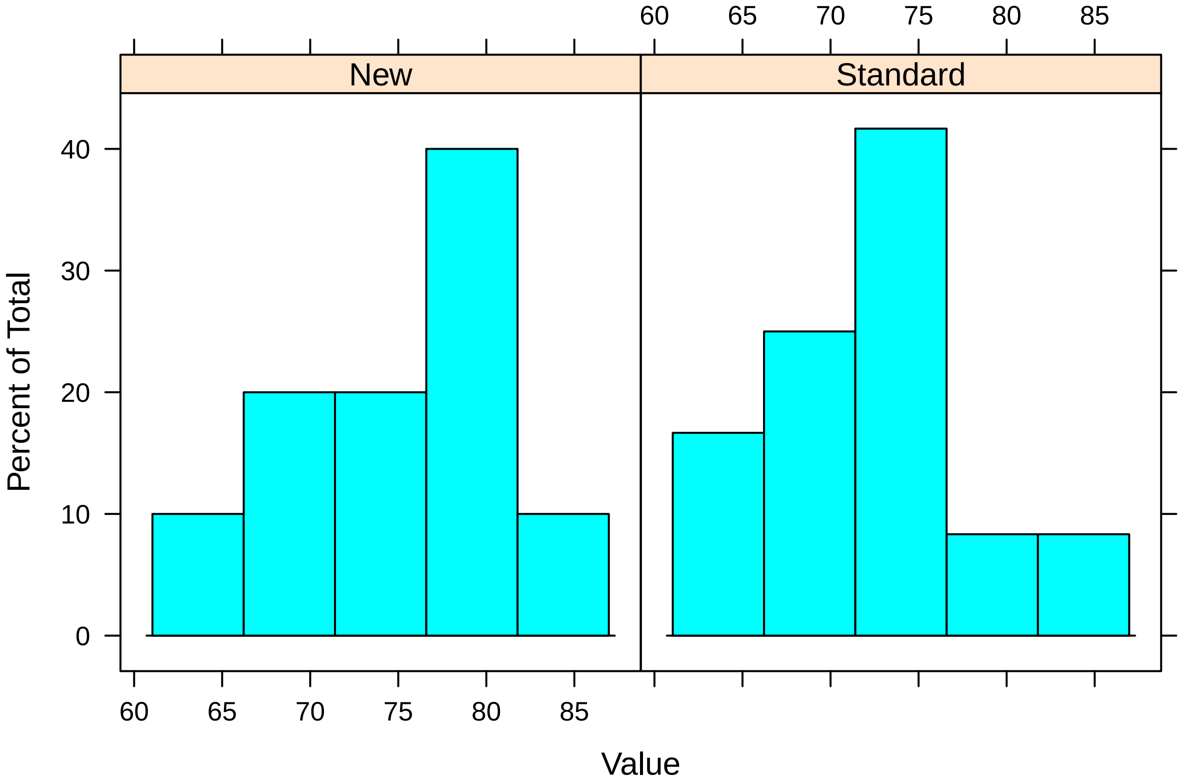
```
summary(Example_1)
```

```
##       Group        Value
##  New     :10   Min.   :62.00
##  Standard:12   1st Qu.:70.00
##                Median :74.00
##                Mean   :74.18
##                3rd Qu.:79.00
##                Max.   :86.00
```

Note the changes in the other frames; the environment frame now contains information about Example_1, and the console includes the commands and command output.

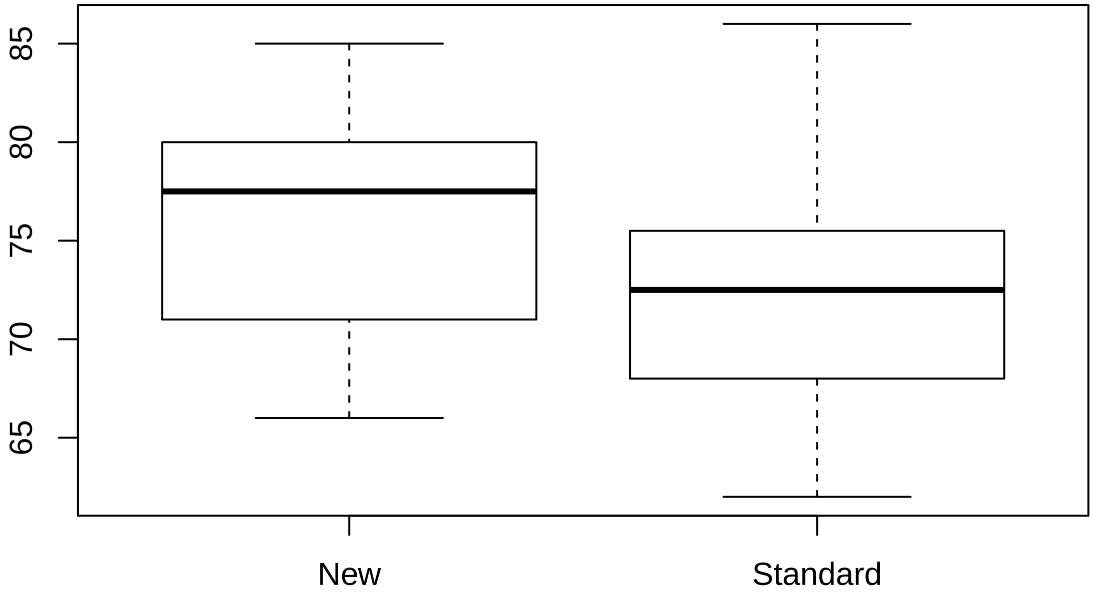
The next set of commands can print histograms side by side; they require loading an R package, lattice, to do so; lattice has a simpler syntax for side-by-side histograms that R's default hist() command lacks; look for the graph in the lower right-hand frame under the Plots tab. If you want to repeat a command (or re-execute it after a slight edit), go to the console and hit the "up" arrow on your keypad, and R will retrieve the command. Now hit the return key and the command will be executed again.

```
library(lattice)
histogram(~ Value | Group, data=Example_1)
```



Side-by-side boxplots do not require the lattice package (though since lattice has been loaded, its data sets and functions are available for the remainder of this R session). Note that the boxplot command supports an equation syntax that is not available in the default hist() command.

```
boxplot(Value~Group, data=Example_1)
```


New
Standard