First Bytes 2014 - R Technology Workshop

R is the most popular free software environment for statistical computing and graphics. ggplot2 is a data visualization package for R that that can be used to produce publication-quality graphics. Both technologies and more (RStudio, KnitR, Slidify, and Shiny) will be explored in this workshop.

This is how my RStudio is configured:

```
sessionInfo()
```

```
## R version 3.0.2 (2013-09-25)
## Platform: x86_64-apple-darwin10.8.0 (64-bit)
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## loaded via a namespace (and not attached):
## [1] digest_0.6.4     evaluate_0.5.5   formatR_0.10     htmltools_0.2.4
## [5] knitr_1.6        rmarkdown_0.2.49 stringr_0.6.2    tools_3.0.2
## [9] yaml_2.1.13
```

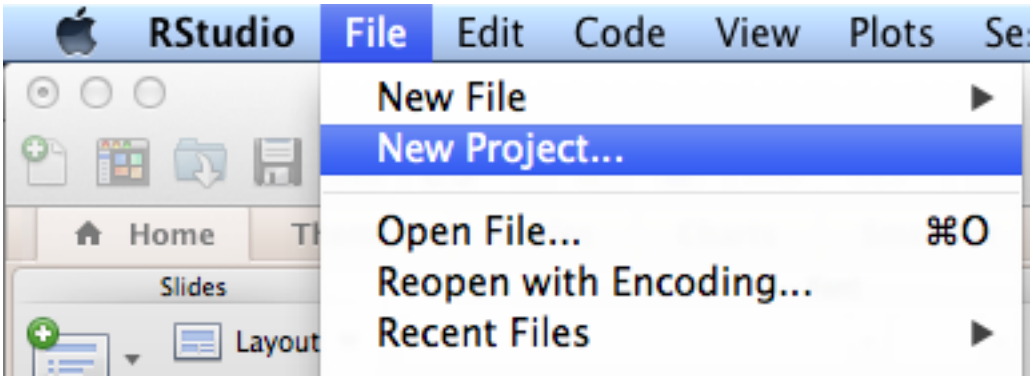
You also need to install LaTeX if you want to generate PDF files from KnitR.

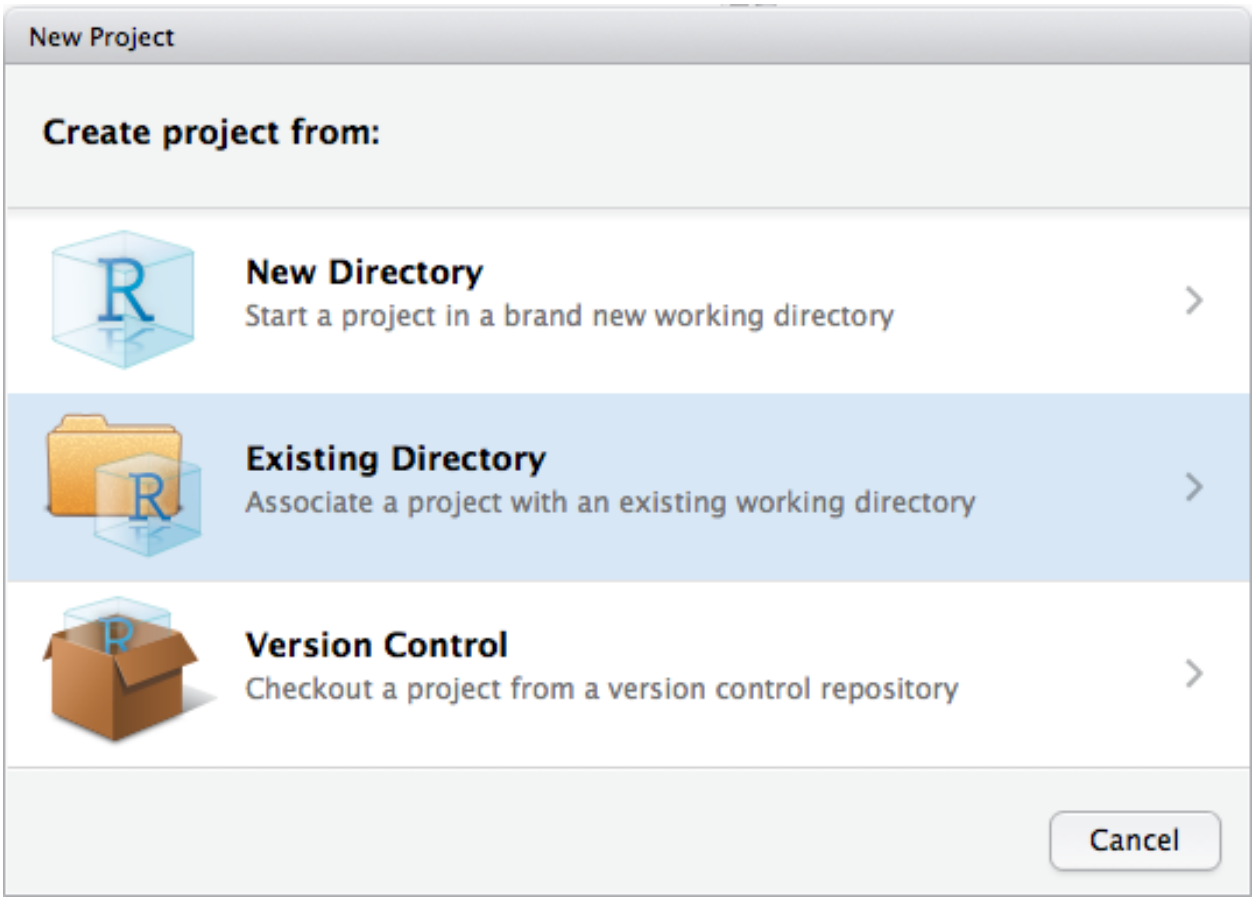
http://latex-project.org/ftp.html

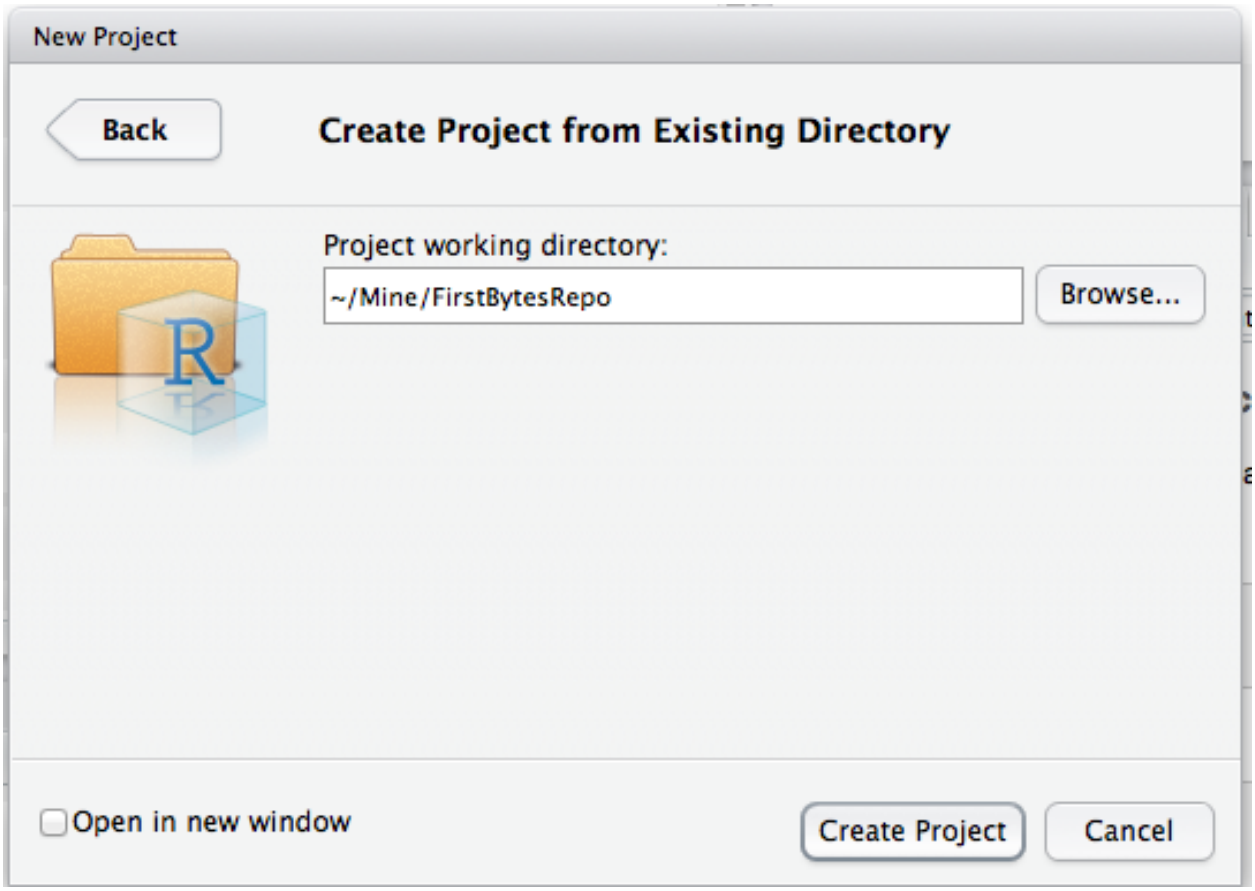
Getting Started - Clone the Workshop Repository:

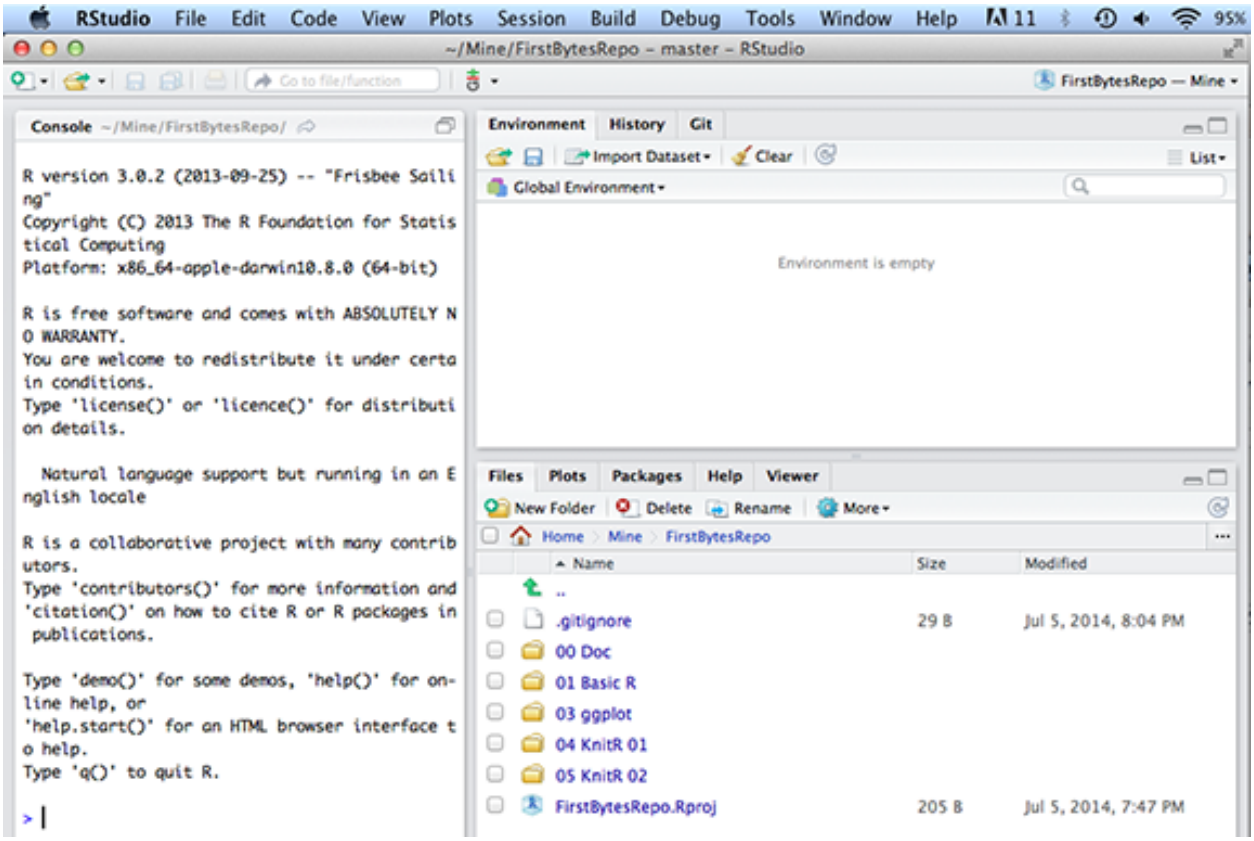
Phils-MacBook-Pro:Mine pcannata$ pwd
/Users/pcannata/Mine
Phils-MacBook-Pro:Mine pcannata$ git clone https://github.com/pcannata/FirstBytesRepo.git
Cloning into 'FirstBytesRepo'...
remote: Counting objects: 19, done.
remote: Compressing objects: 100% (12/12), done.
remote: Total 19 (delta 3), reused 19 (delta 3)
Unpacking objects: 100% (19/19), done.
Checking connectivity... done
Phils-MacBook-Pro:Mine pcannata$ ls -a FirstBytesRepo/
. .Rprofile.R 00 Doc 03 ggplot 05 KnitR 02
.. .git 01 Basic R 04 KnitR 01 FirstBytesRepo.Rproj

Getting Started - Create a New RStudio Project for the Workshop Repository code:






New Project
Back
Create Project from Existing Directory
Project working directory:
~/Mine/FirstBytesRepo
Browse...
Open in new window
Create Project
Cancel



Recommended Books:

http://www.amazon.com/Everyone-Advanced-Analytics-Graphics-Addison-Wesley/dp/0321888030/ref=sr_1_1?ie=UTF8&qid=1404611818&sr=8-1&keywords=R+for+everyone

http://www.amazon.com/Reproducible-Research-RStudio-Chapman-Series/dp/1466572841/ref=sr_1_1?ie=UTF8&qid=1404612129&sr=8-1&keywords=reproducible+research+with+r

Basic R Language Constructs

This is a tiny introduction to R.

See also http://cran.r-project.org/doc/manuals/r-devel/R-lang.html, http://www.r-tutor.com/r-introduction, and http://www.cookbook-r.com/

```
source("../01 Basic R/Basic.R", echo = TRUE)
```

```
##
## > "Variables"
## [1] "Variables"
##
## > v <- 211
##
## > v
## [1] 211
##
## > "Vectors"
## [1] "Vectors"
##
## > v1 <- c(1, 2, 3, 4, 5)
##
## > v1
## [1] 1 2 3 4 5
##
## > v2 <- 1:11
##
## > v2
##  [1]  1  2  3  4  5  6  7  8  9 10 11
##
## > v3 <- -5:5
##
## > v3
##  [1] -5 -4 -3 -2 -1  0  1  2  3  4  5
##
## > "Vector Operations"
## [1] "Vector Operations"
##
## > v1 + 2
## [1] 3 4 5 6 7
##
## > sqrt(v2)
##  [1] 1.000 1.414 1.732 2.000 2.236 2.449 2.646 2.828 3.000 3.162 3.317
##
## > v2 + v3
##  [1] -4 -2  0  2  4  6  8 10 12 14 16
##
## > length(4:22)
## [1] 19
##
## > mean(4:22)
## [1] 13
##
## > "Functions"
## [1] "Functions"
##
## > apropos("mean")
```

```
##  [1] ".colMeans"     ".rowMeans"     "colMeans"      "kmeans"
##  [5] "mean"          "mean.Date"     "mean.default"  "mean.difftime"
##  [9] "mean.POSIXct"  "mean.POSIXlt"  "rowMeans"      "weighted.mean"
```

R Dataframes

See also http://www.r-tutor.com/r-introduction/data-frame

```
source("../02 R Dataframes/Dataframes.R", echo = TRUE)
```

```
##
## > library("ggplot2", lib.loc = "/Library/Frameworks/R.framework/Versions/3.0/Resources/library")
##
## > head(diamonds)
##   carat       cut color clarity depth table price    x    y    z
## 1  0.23     Ideal     E     SI2  61.5    55   326 3.95 3.98 2.43
## 2  0.21   Premium     E     SI1  59.8    61   326 3.89 3.84 2.31
## 3  0.23      Good     E     VS1  56.9    65   327 4.05 4.07 2.31
## 4  0.29   Premium     I     VS2  62.4    58   334 4.20 4.23 2.63
## 5  0.31      Good     J     SI2  63.3    58   335 4.34 4.35 2.75
## 6  0.24 Very Good     J    VVS2  62.8    57   336 3.94 3.96 2.48
```

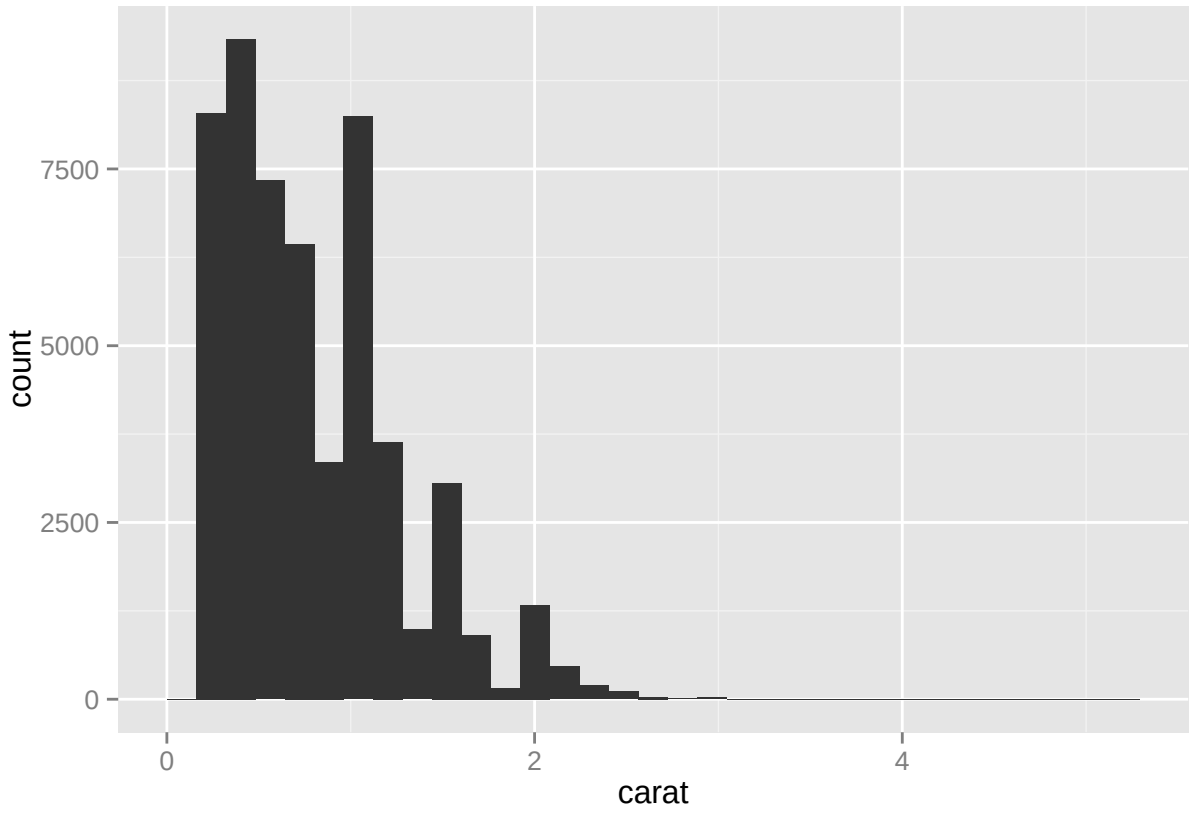
R ggplot2

See also http://ggplot2.org/

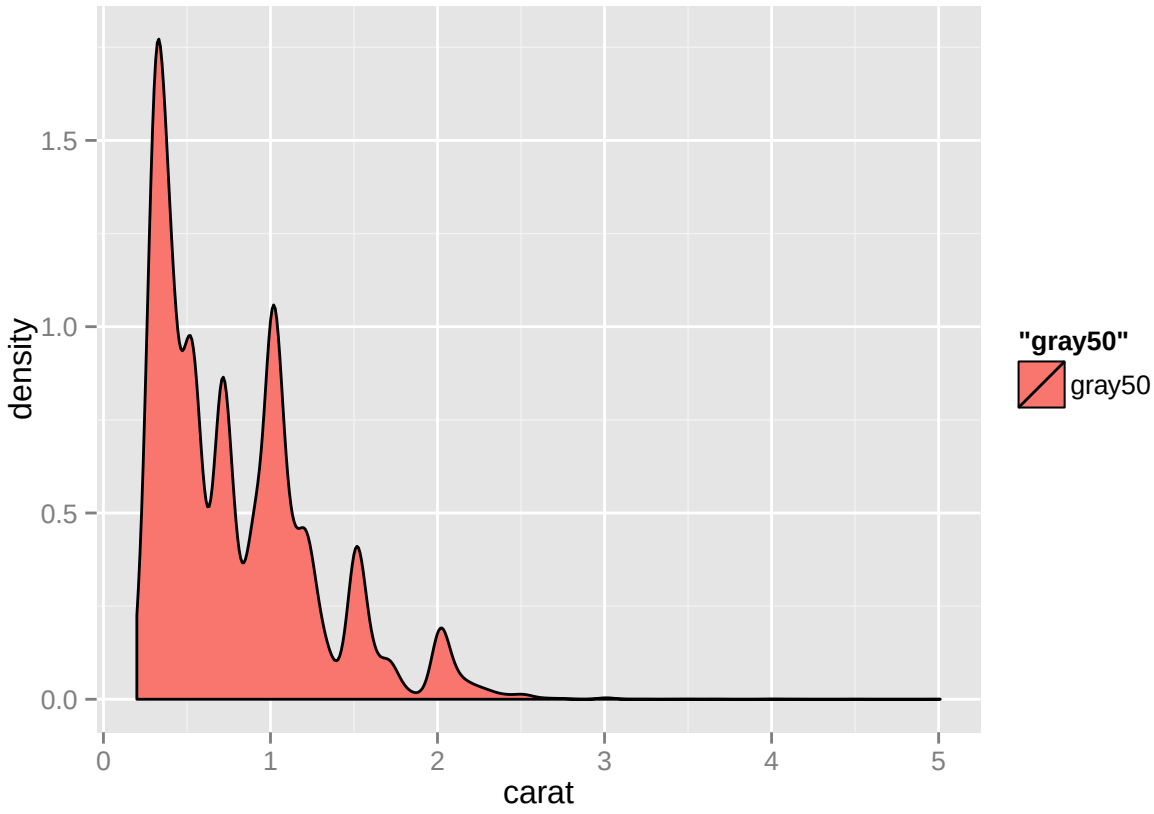
```
source("../03 ggplot/01 Histograms.R", echo = TRUE)
```

```
##
## > library("ggplot2", lib.loc = "/Library/Frameworks/R.framework/Versions/3.0/Resources/library")
##
## > ggplot(data = diamonds) + geom_histogram(aes(x = carat))

## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```




```
##
## > ggplot(data = diamonds) + geom_density(aes(x = carat,
## +     fill = "gray50"))
```


density
1.5
1.0
0.5
0.0
0
1
2
3
4
5
carat
"gray50"
gray50