Welcome to the Tidy Tools Workshop!

1. Get connected to the wifi

   rstudio_conf / tidyverse

2. Get the course materials

   usethis::use_course("http://rstd.io/tidy-tools")

3. Follow the instructions in **setup.R**

4. Stuck? Please ask for help!

Doesn't work? Install latest usethis from CRAN

Tidy tools

January 2018

Hadley Wickham
@hadleywickham
Chief Scientist, **RStudio**

Preliminaries

HELLO

my name is

Hadley

Lionel Henry

RStudio

Kara Woo

Datacamp

Your turn

This means that you have to work!

This course is very hands on, and while we're here to help you, the best resource is often the person sitting next to you.

Introduce yourself to your neighbours. Who are you and what are you using R for?

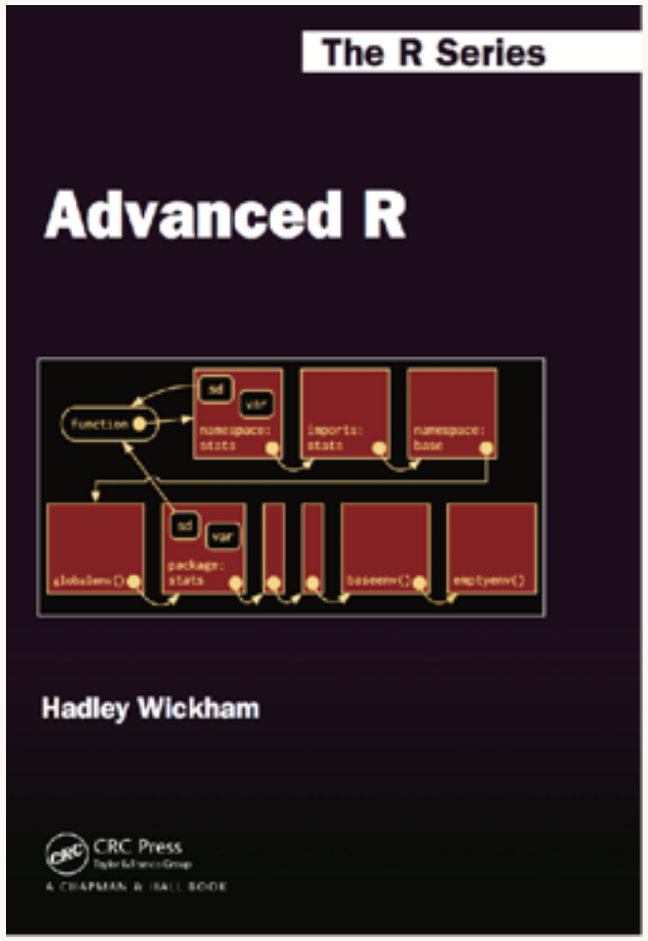
Writing functions

Individual functions to solve individual problems

Designing APIs

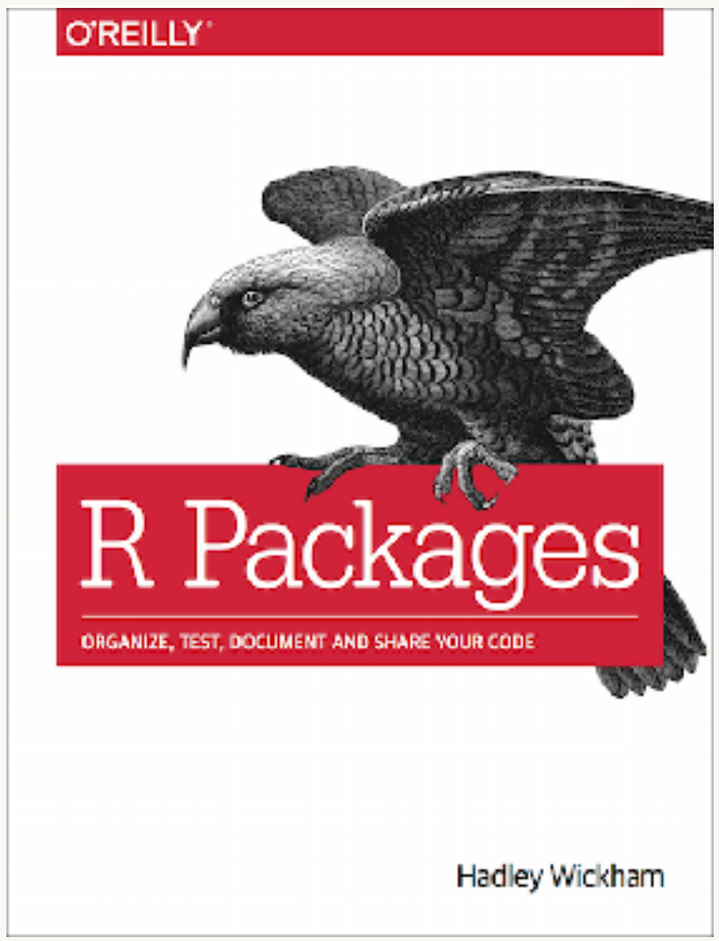
Family of functions that work together to solve family of problems

Much of the course is drawn from existing books



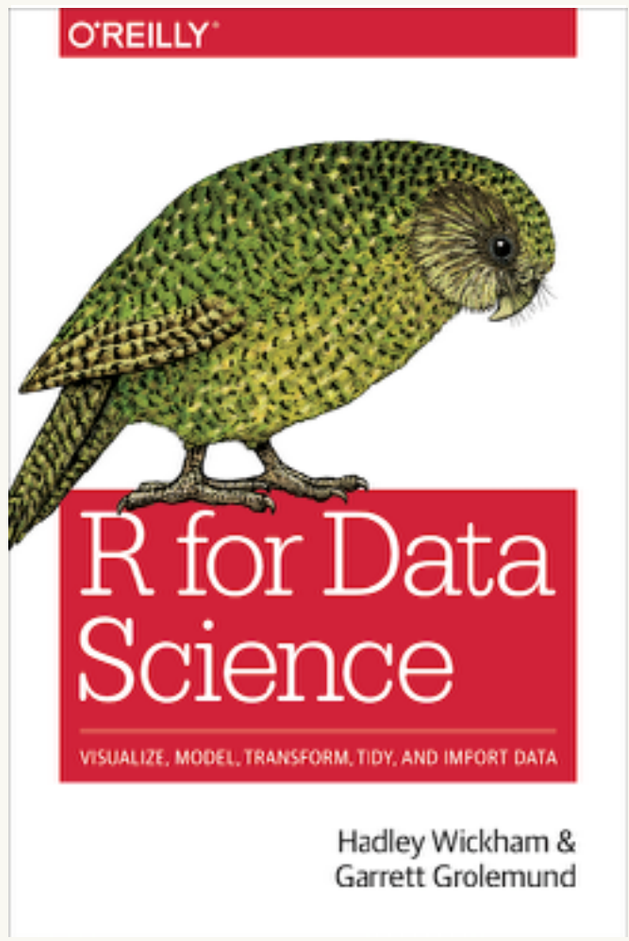
http://adv-r.hadley.nz/

Working on 2nd ed



http://r-pkgs.had.co.nz

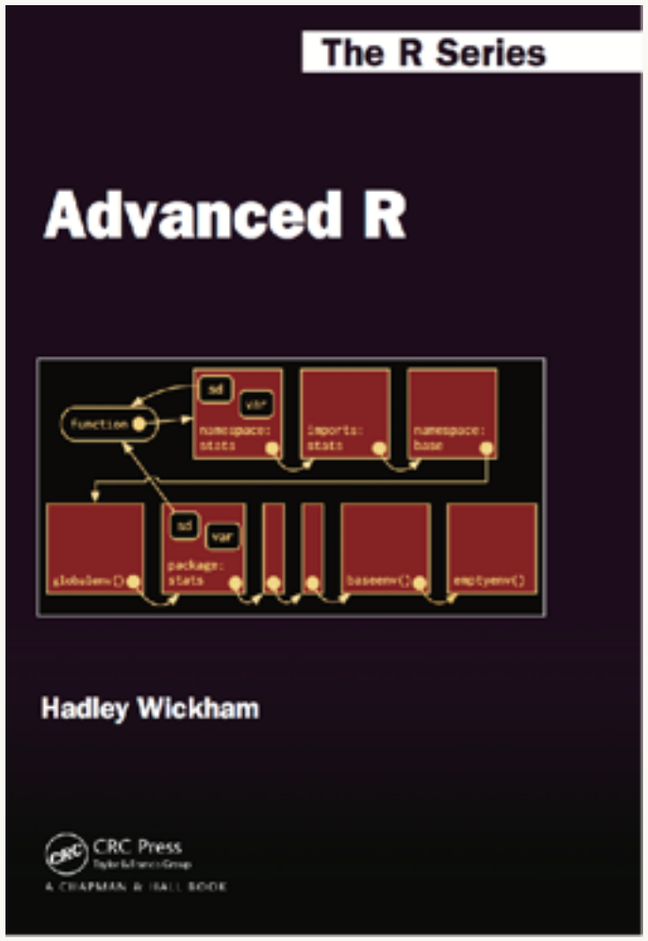
https://amzn.com/1491910399



http://r4ds.had.co.nz

https://amzn.com/1491910399

But the primary book does not yet exist



How R works

Tidy tools

Hadley Wickham

How to solve real problems with R

Warmups

Don’t expect to know all the answers!

Your turn

What are the four common types of atomic vectors?

(Bonus points for the two uncommon types)

What are the three properties of a vector?

```
typeof(TRUE)
typeof(1L)
typeof(1.5)
typeof("a")

# We'll talk about this (S3) later:
typeof(factor(1:10))
typeof(Sys.Date())
```

Every vector has three properties:

```
x <- 1:5

# 1. Type:
typeof(x)

# 2. Length
length(x)

# 3. Attributes
attributes(x)
# (we'll come back to those later)
```

Missing values

What does `NA == NA` return? Why?

What should you use instead?

There isn’t a single unknown value

```
age_john <- NA
age_mary <- NA
age_john == age_mary

is.na(x)
```

Your turn

What are the six types of thing that you can put inside []?

blank	include all
integer	**+ve**: include **0**: drop all **-ve**: exclude
logical	keep TRUEs
character	lookup by name

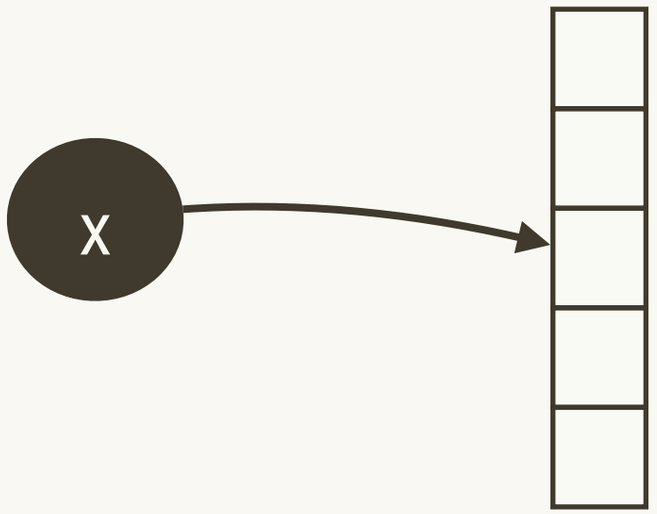
Use character subsetting for simple look ups

```
x <- c("m", "f", "u", "f", "f", "m", "m")
lookup <- c(m = "Male", f = "Female", u = NA)
lookup[x]

unname(lookup[x])
```

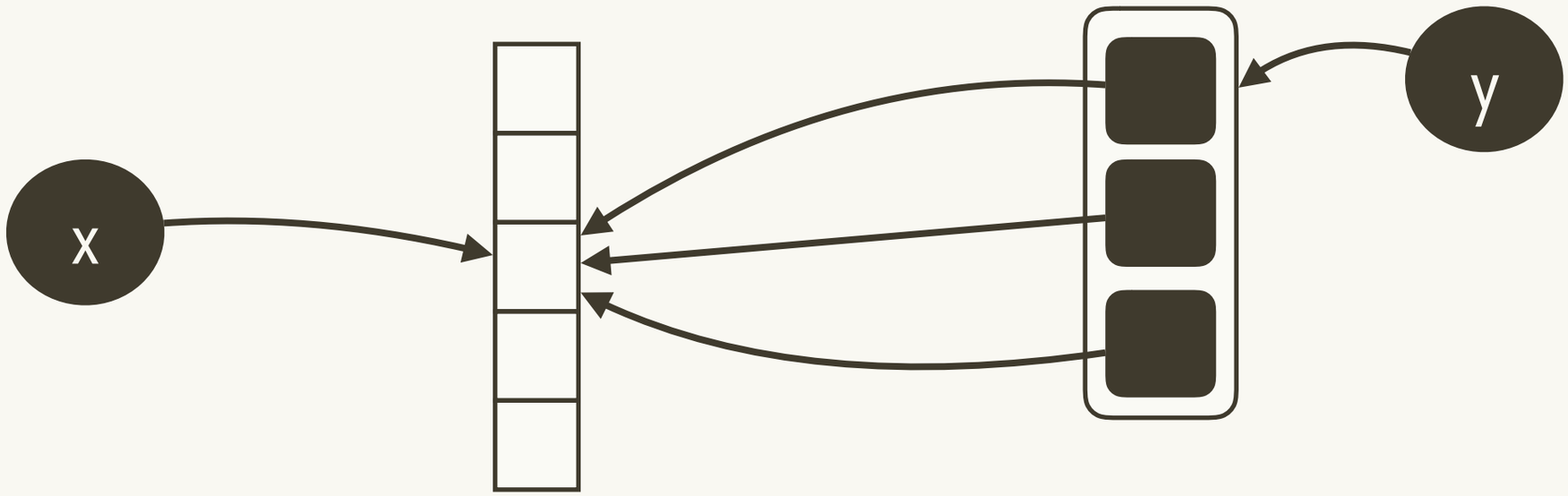
Your turn

```
x <- rnorm(1e6)
pryr::object_size(x)
#> 8 MB

# How big will y be?
y <- list(x, x, x)
pryr::object_size(y)

# What if we change a value?
y[[1]][[1]] <- NA
pryr::object_size(y)
```

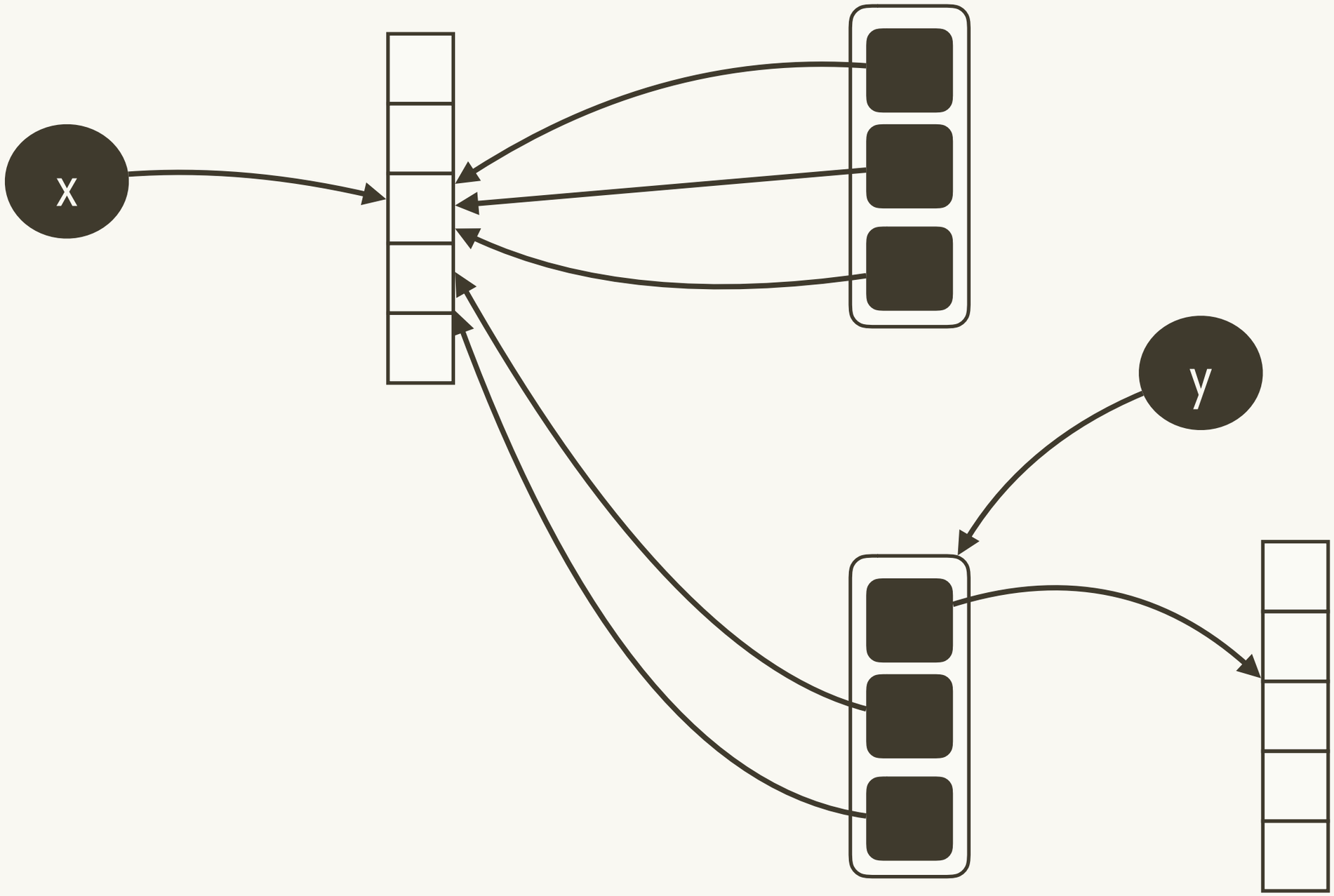
A name points to an object



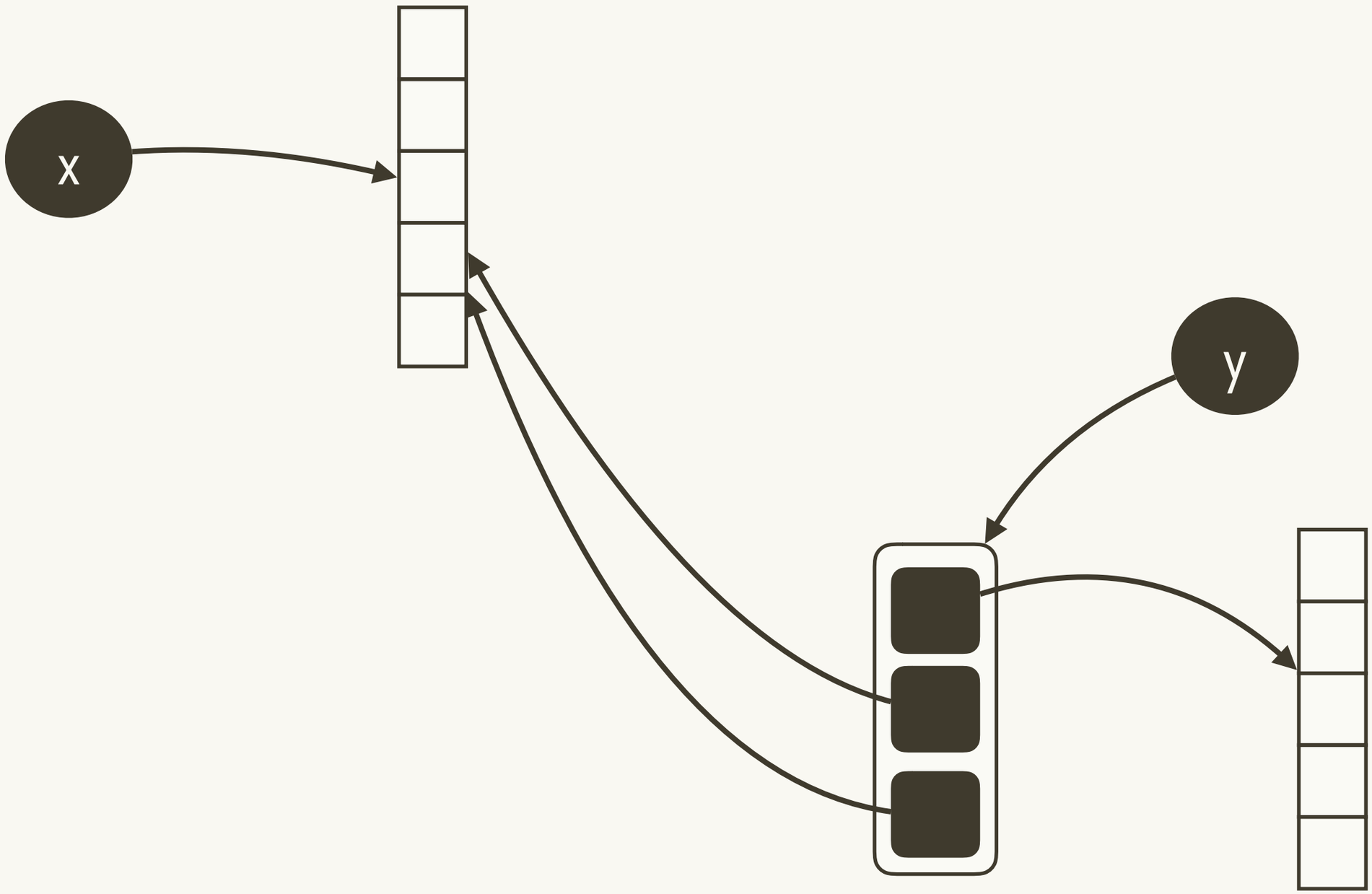
There can be multiple references to an object

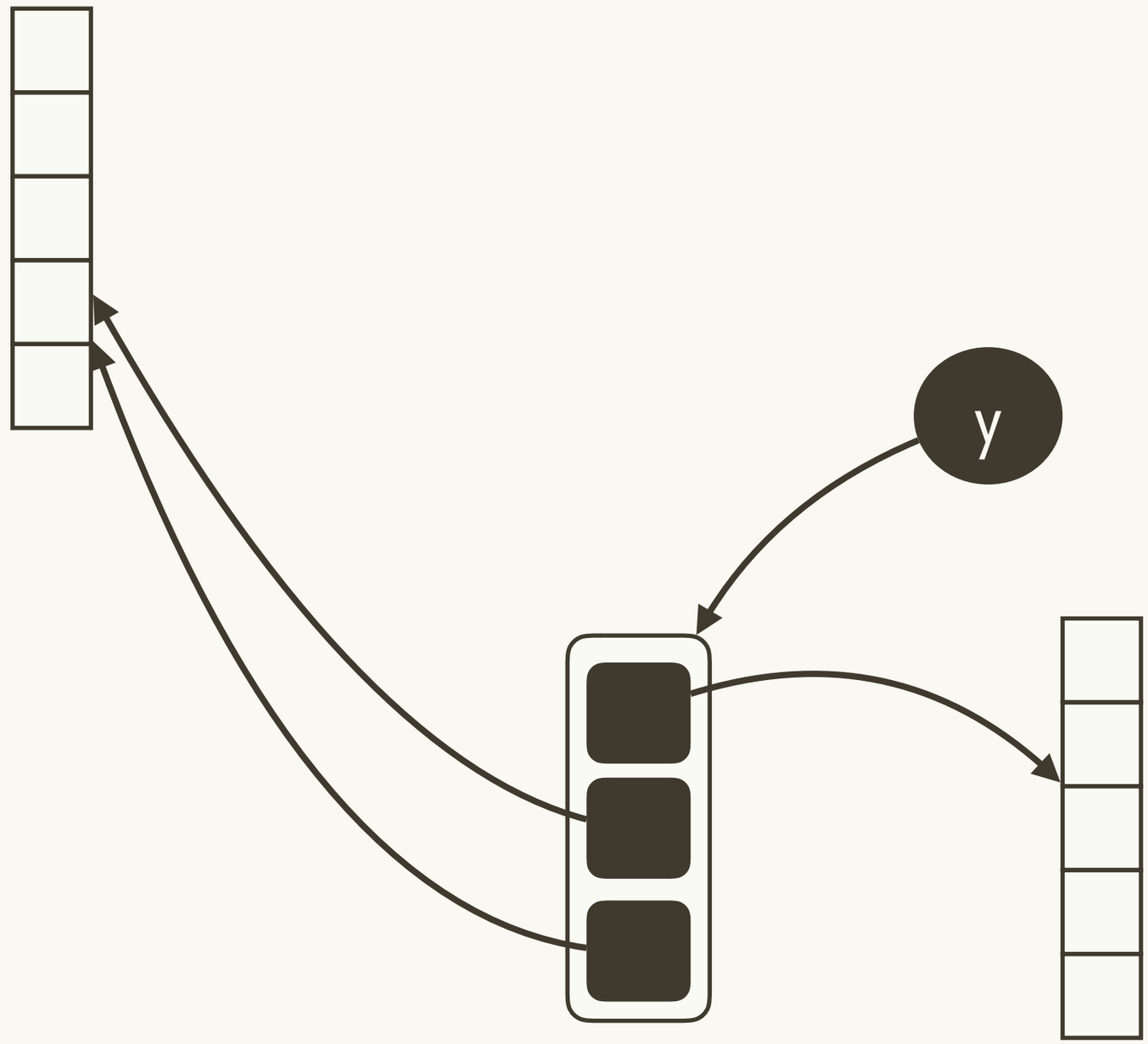


Modifying an object creates a copy

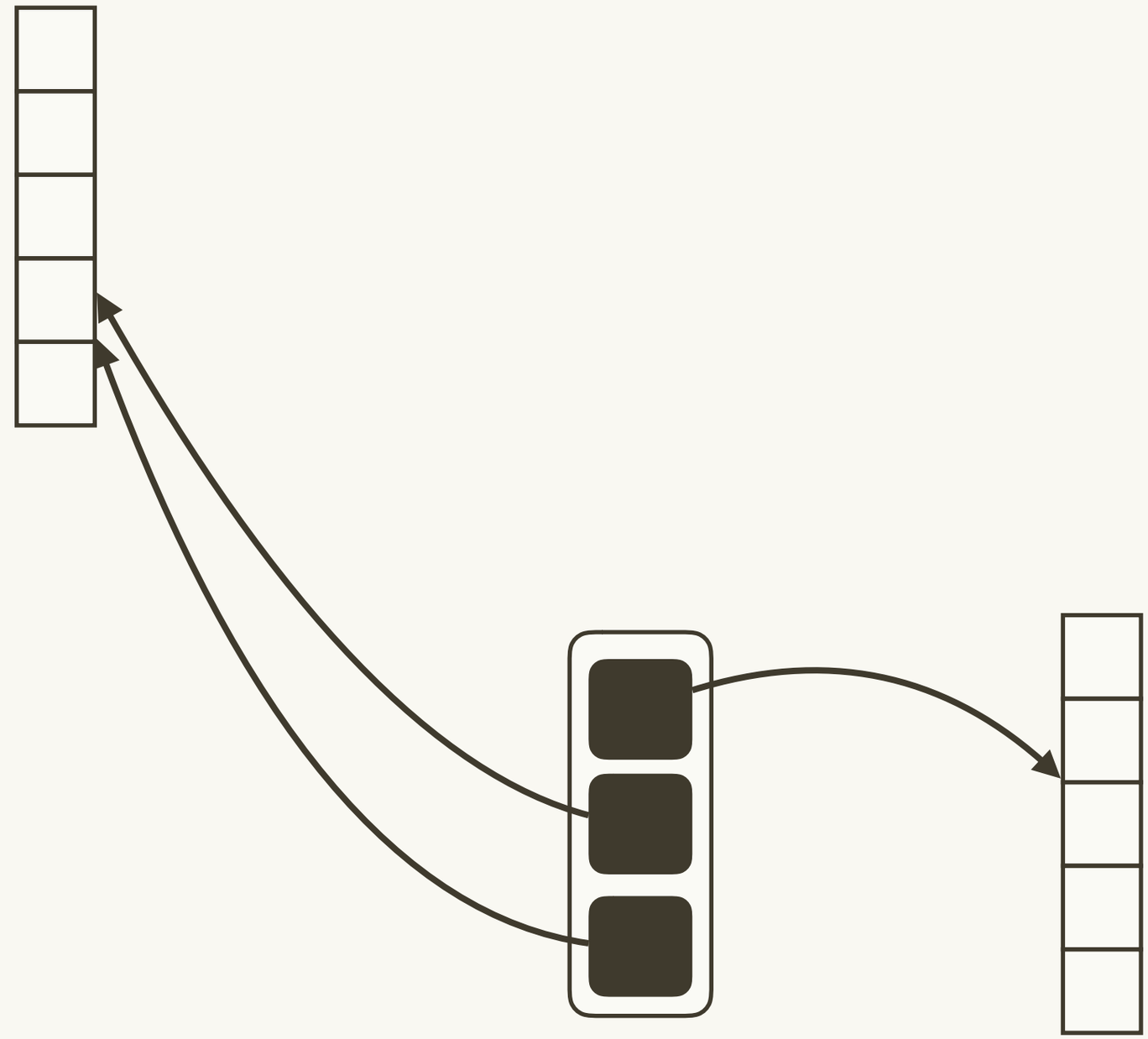


What happens after rm(x) and rm(y)?




y

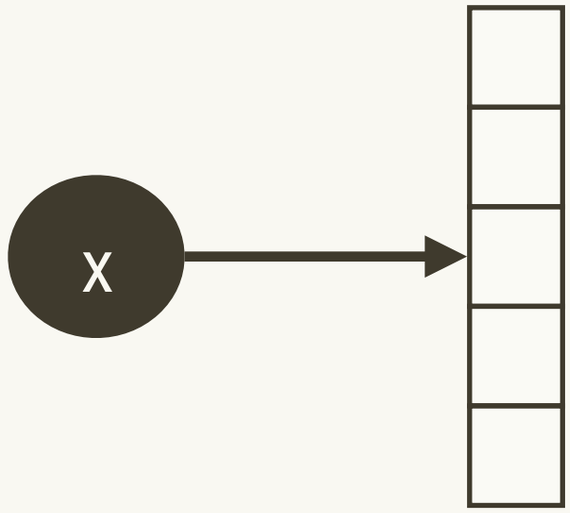
First x and y are removed



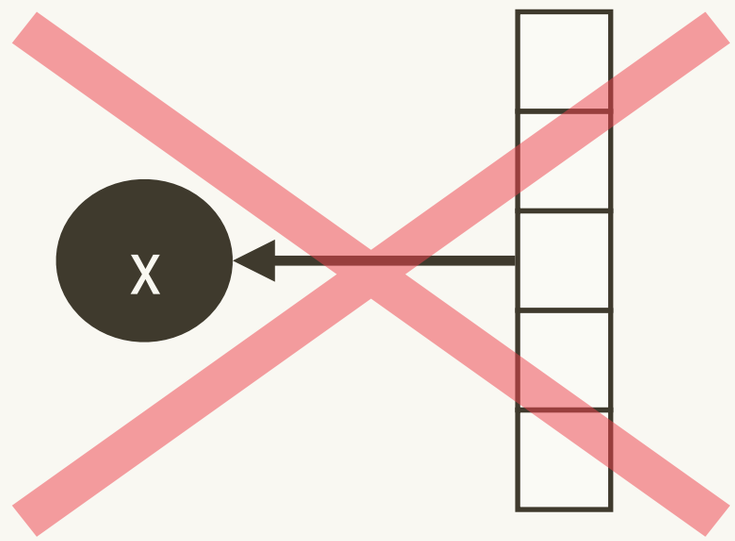
Then the objects they point to

Until nothing is left

The direction of the arrow is important



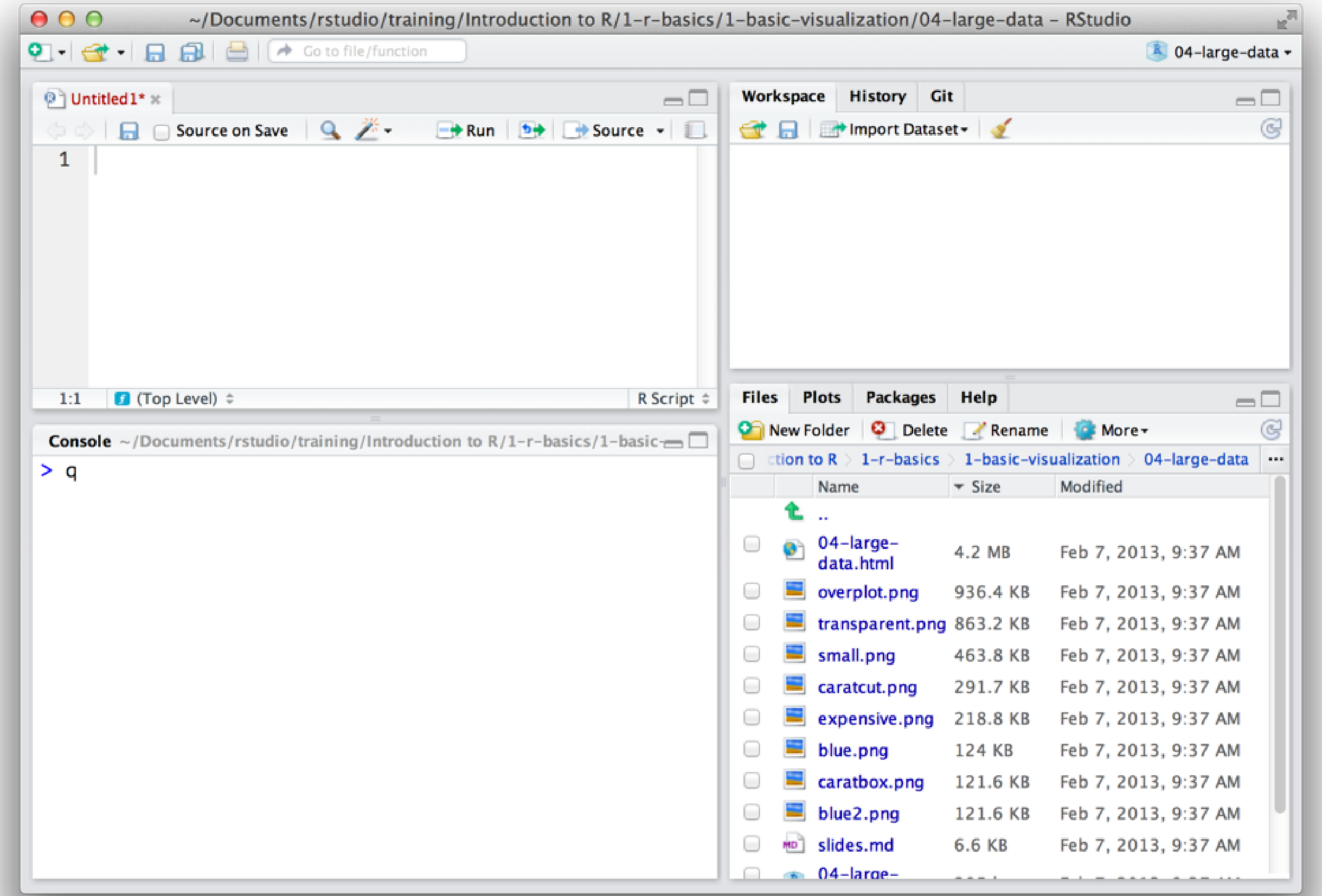
A name “has” an object

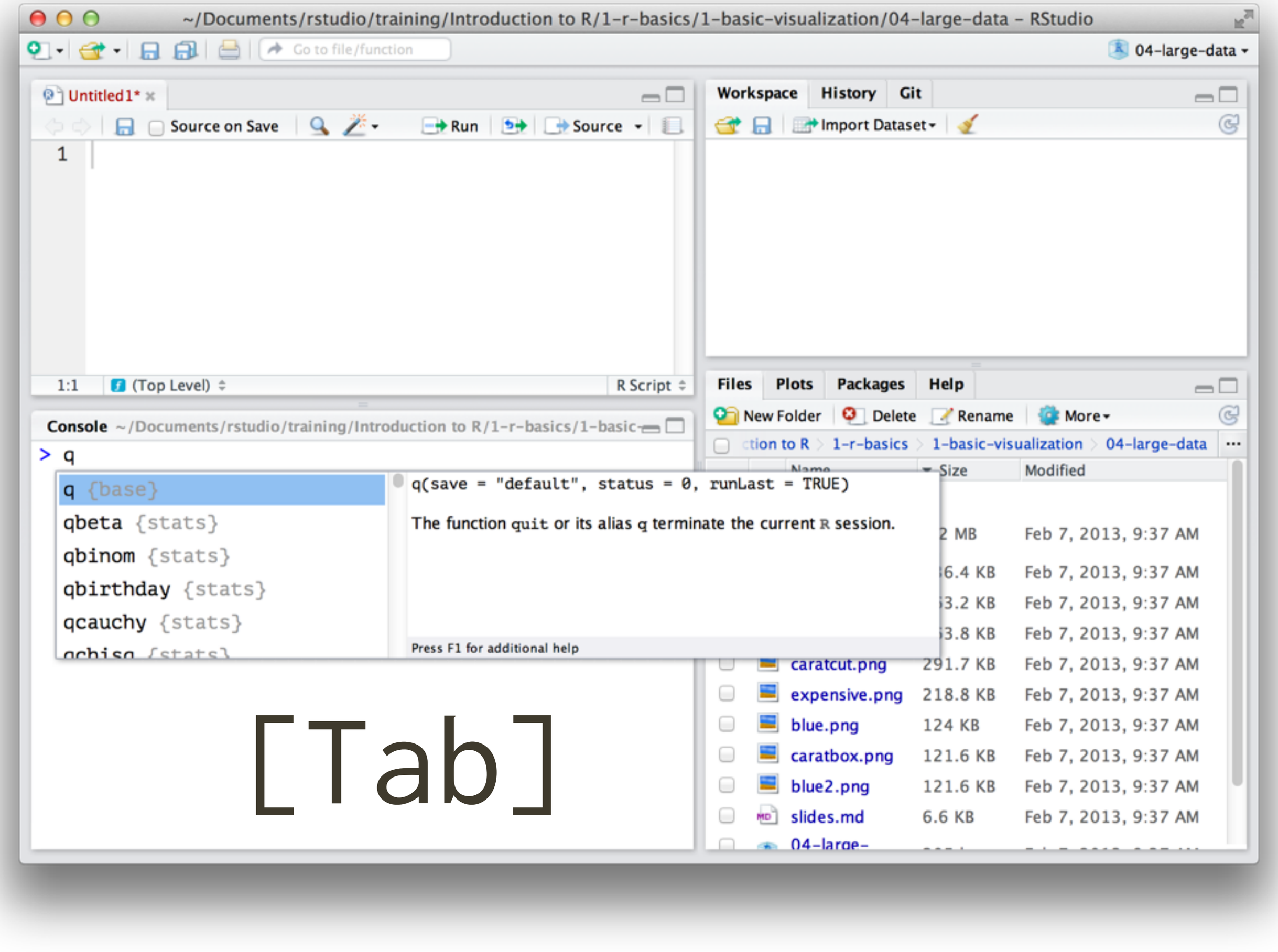


An object doesn’t have a name

RStudio

You don't have to use RStudio,
but if you do, try to master it!


~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-visualization/04-large-data - RStudio
Go to file/function
04-large-data
Untitled1*
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
Console ~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-
> q
Workspace
History
Git
Import Dataset
Files
Plots
Packages
Help
New Folder
Delete
Rename
More
tion to R
1-r-basics
1-basic-visualization
04-large-data
Name
Size
Modified
..
04-large-data.html 4.2 MB Feb 7, 2013, 9:37 AM
overplot.png 936.4 KB Feb 7, 2013, 9:37 AM
transparent.png 863.2 KB Feb 7, 2013, 9:37 AM
small.png 463.8 KB Feb 7, 2013, 9:37 AM
caratcut.png 291.7 KB Feb 7, 2013, 9:37 AM
expensive.png 218.8 KB Feb 7, 2013, 9:37 AM
blue.png 124 KB Feb 7, 2013, 9:37 AM
caratbox.png 121.6 KB Feb 7, 2013, 9:37 AM
blue2.png 121.6 KB Feb 7, 2013, 9:37 AM
slides.md 6.6 KB Feb 7, 2013, 9:37 AM
04-large-


~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-visualization/04-large-data - RStudio
Go to file/function
04-large-data
Untitled1*
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
Workspace
History
Git
Import Dataset
Files
Plots
Packages
Help
New Folder
Delete
Rename
More
1-r-basics
1-basic-visualization
04-large-data
Name
Size
Modified
Console ~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-
> q
q {base}
qbeta {stats}
qbinom {stats}
qbirthday {stats}
qcauchy {stats}
q(save = "default", status = 0, runLast = TRUE)
The function quit or its alias q terminate the current R session.
Press F1 for additional help
2 MB
Feb 7, 2013, 9:37 AM
6.4 KB
Feb 7, 2013, 9:37 AM
3.2 KB
Feb 7, 2013, 9:37 AM
3.8 KB
Feb 7, 2013, 9:37 AM
caratcut.png
291.7 KB
Feb 7, 2013, 9:37 AM
expensive.png
218.8 KB
Feb 7, 2013, 9:37 AM
blue.png
124 KB
Feb 7, 2013, 9:37 AM
caratbox.png
121.6 KB
Feb 7, 2013, 9:37 AM
blue2.png
121.6 KB
Feb 7, 2013, 9:37 AM
slides.md
6.6 KB
Feb 7, 2013, 9:37 AM
[Tab]

~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-visualization/04-large-data - RStudio

Go to file/function

04-large-data

Untitled1*

Source on Save | Run | Source

```
qplot(table - depth, data = diamonds,
qplot(day, data = email)
qplot(day, mails, data = daily, geom = "line", colo
qplot(day, mails, data = daily, geom = "smooth", co
qplot(day, variants, data = daily, geom = "line", c
qplot(wday, hour, data = wh, size = freq)
qplot(mpg, wt, data = mtcars)
qplot(mpg, wt, data = mtcars, colour = cyl)
```

```
> q
```

[Cmd/Ctrl + ↑]

Workspace | History | Git

Import Dataset

Files | Plots | Packages | Help

New Folder | Delete | Rename | More

ction to R > 1-r-basics > 1-basic-visualization > 04-large-data

Name	Size	Modified
..		
04-large-data.html	4.2 MB	Feb 7, 2013, 9:37 AM
overplot.png	936.4 KB	Feb 7, 2013, 9:37 AM
transparent.png	863.2 KB	Feb 7, 2013, 9:37 AM
small.png	463.8 KB	Feb 7, 2013, 9:37 AM
caratcut.png	291.7 KB	Feb 7, 2013, 9:37 AM
	B	Feb 7, 2013, 9:37 AM
		Feb 7, 2013, 9:37 AM
	B	Feb 7, 2013, 9:37 AM
	B	Feb 7, 2013, 9:37 AM
		Feb 7, 2013, 9:37 AM
04-large-		

~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-visualization/04-large-data - RStudio

Go to file/function

04-large-data

Untitled1*

Source on Save

Run

Source

```
1  library(ggplot2)
```

1:17 (Top Level) R Script

Console ~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-

```
>
```

Workspace History Git

Import Dataset

Files Plots Packages Help

New Folder Delete Rename More

tion to R 1-r-basics 1-basic-visualization 04-large-data

	Name	Size	Modified
	..		
	04-large-data.html	4.2 MB	Feb 7, 2013, 9:37 AM
	overplot.png	936.4 KB	Feb 7, 2013, 9:37 AM
	transparent.png	863.2 KB	Feb 7, 2013, 9:37 AM
	small.png	463.8 KB	Feb 7, 2013, 9:37 AM
	caratcut.png	291.7 KB	Feb 7, 2013, 9:37 AM
	expensive.png	218.8 KB	Feb 7, 2013, 9:37 AM
	blue.png	124 KB	Feb 7, 2013, 9:37 AM
	caratbox.png	121.6 KB	Feb 7, 2013, 9:37 AM
	blue2.png	121.6 KB	Feb 7, 2013, 9:37 AM
	slides.md	6.6 KB	Feb 7, 2013, 9:37 AM
	04-large-		

~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-visualization/04-large-data - RStudio

Go to file/function

04-large-data

Untitled1*

Source on Save | Run | Source

```
1  library(ggplot2)
```

[Cmd/ctrl + enter]

1:17 (Top Level) R Script

Workspace | History | Git

Import Dataset

Files | Plots | Packages | Help

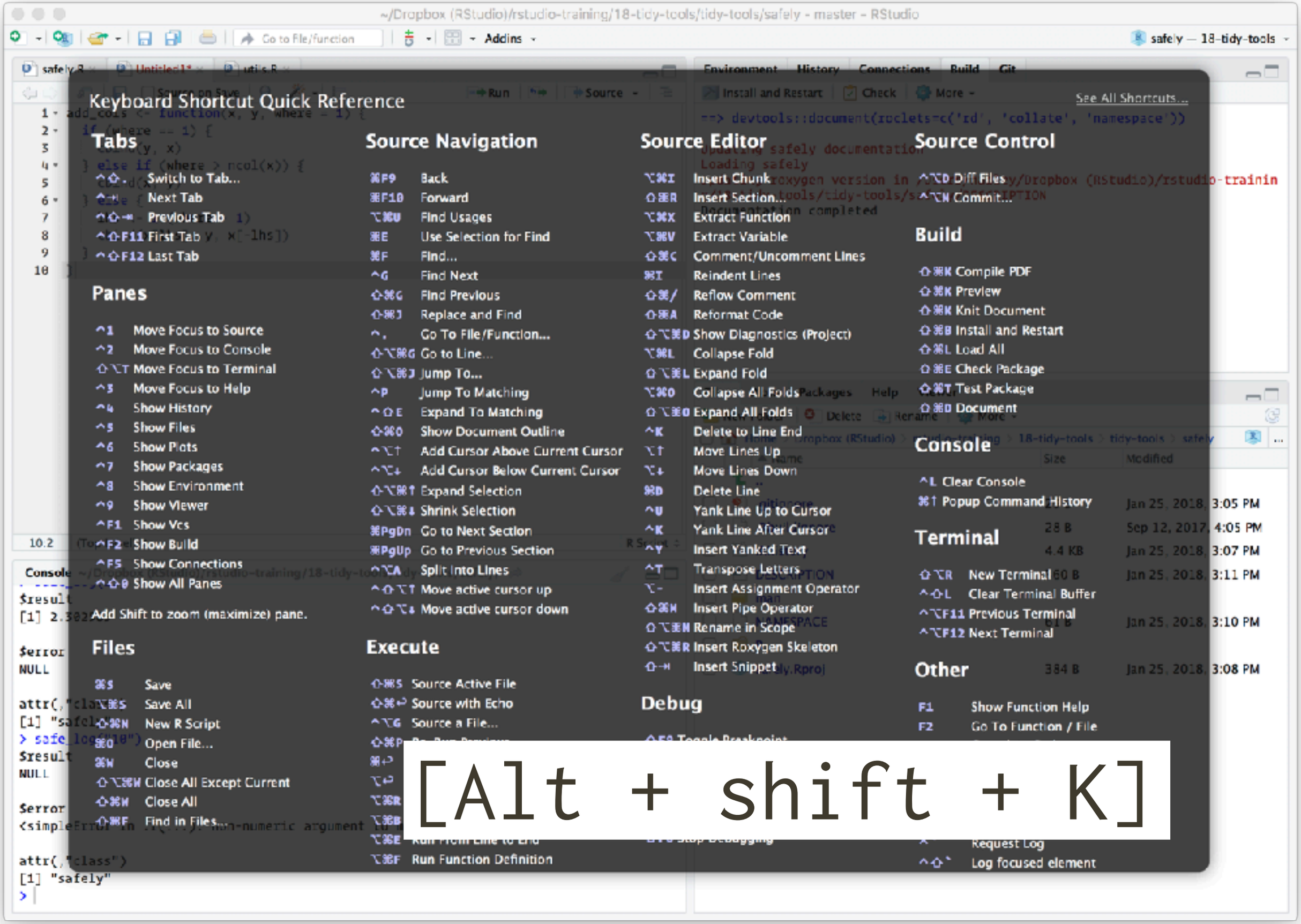
New Folder | Delete | Rename | More

tion to R › 1-r-basics › 1-basic-visualization › 04-large-data

Console ~/Documents/rstudio/training/Introduction to R/1-r-basics/1-basic-

```
> library(ggplot2)
>
```

	Name	Size	Modified
	..		
	04-large-data.html	4.2 MB	Feb 7, 2013, 9:37 AM
	overplot.png	936.4 KB	Feb 7, 2013, 9:37 AM
	transparent.png	863.2 KB	Feb 7, 2013, 9:37 AM
	small.png	463.8 KB	Feb 7, 2013, 9:37 AM
	caratcut.png	291.7 KB	Feb 7, 2013, 9:37 AM
	expensive.png	218.8 KB	Feb 7, 2013, 9:37 AM
	blue.png	124 KB	Feb 7, 2013, 9:37 AM
	caratbox.png	121.6 KB	Feb 7, 2013, 9:37 AM
	blue2.png	121.6 KB	Feb 7, 2013, 9:37 AM
	slides.md	6.6 KB	Feb 7, 2013, 9:37 AM
	04-large-		


~/Dropbox (RStudio)/rstudio-training/18-tidy-tools/tidy-tools/safely - master - RStudio
Keyboard Shortcut Quick Reference
Tabs
Switch to Tab...
Next Tab
Previous Tab
First Tab
Last Tab
Panes
Move Focus to Source
Move Focus to Console
Move Focus to Terminal
Move Focus to Help
Show History
Show Files
Show Plots
Show Packages
Show Environment
Show Viewer
Show Vcs
Show Build
Show Connections
Show All Panes
Add Shift to zoom (maximize) pane.
Files
Save
Save All
New R Script
Open File...
Close
Close All Except Current
Close All
Find in Files...
Source Navigation
Back
Forward
Find Usages
Use Selection for Find
Find...
Find Next
Find Previous
Replace and Find
Go To File/Function...
Go to Line...
Jump To...
Jump To Matching
Expand To Matching
Show Document Outline
Add Cursor Above Current Cursor
Add Cursor Below Current Cursor
Expand Selection
Shrink Selection
Go to Next Section
Go to Previous Section
Split Into Lines
Move active cursor up
Move active cursor down
Execute
Source Active File
Source with Echo
Source a File...
Run Function Definition
Source Editor
Insert Chunk
Insert Section...
Extract Function
Extract Variable
Comment/Uncomment Lines
Reindent Lines
Reflow Comment
Reformat Code
Show Diagnostics (Project)
Collapse Fold
Expand Fold
Collapse All Folds
Expand All Folds
Delete to Line End
Move Lines Up
Move Lines Down
Delete Line
Yank Line Up to Cursor
Yank Line After Cursor
Insert Yanked Text
Transpose Letters
Insert Assignment Operator
Insert Pipe Operator
Rename in Scope
Insert Roxygen Skeleton
Insert Snippet
Debug
Source Control
Diff Files
Commit...
Build
Compile PDF
Preview
Knit Document
Install and Restart
Load All
Check Package
Test Package
Document
Console
Clear Console
Popup Command History
Terminal
New Terminal
Clear Terminal Buffer
Previous Terminal
Next Terminal
Other
Show Function Help
Go To Function / File
Request Log
Log focused element
[Alt + shift + K]

Your turn

What's the shortcut for <- (assignment)?

What about %>% (pipe)?

How can you quickly comment a block of lines?

What are two ways to create multiple cursors? (Hint: google)

This work is licensed under the
Creative Commons Attribution-Noncommercial 3.0
United States License.

To view a copy of this license, visit
http://creativecommons.org/licenses/by-nc/3.0/us/