R packages

September 2016

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

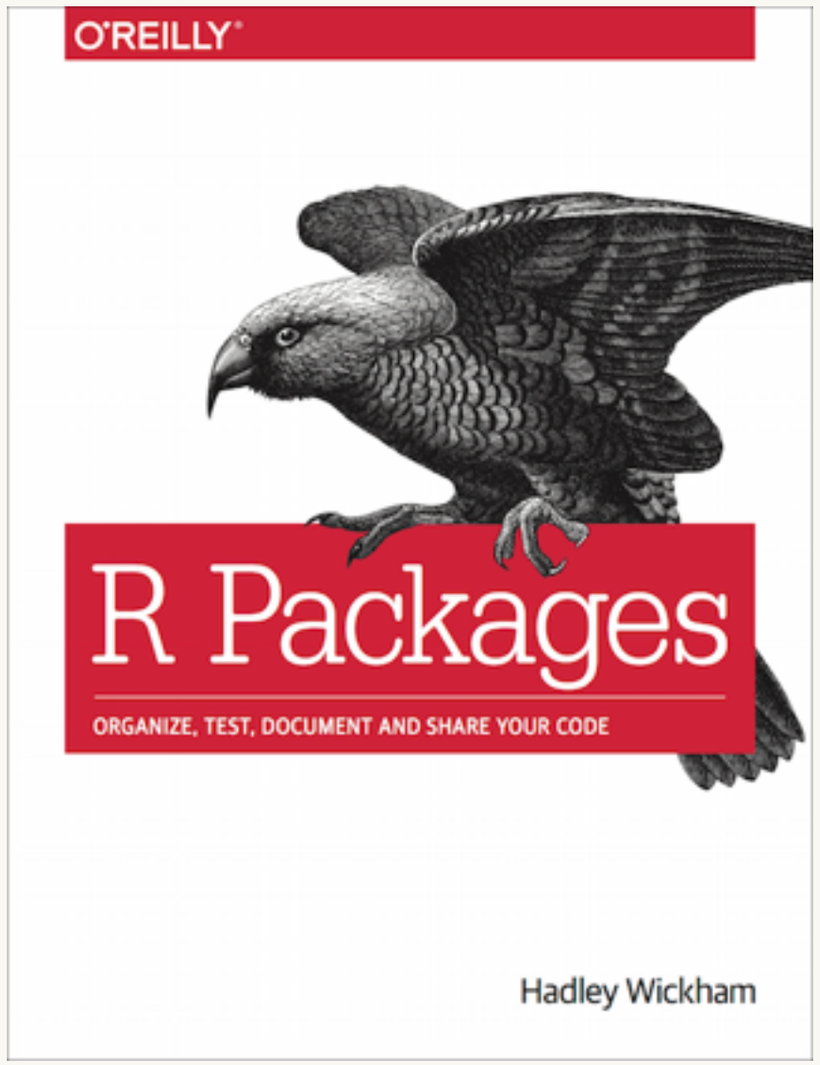
A package is a set of conventions that (with the right tools) makes your life easier

“Seriously, it doesn’t have to be about sharing your code (although that is an added benefit!). It is about saving yourself time.”

— *Hilary Parker*

One way to learn what we don't cover today

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



Or buy from O'Reilly with discount **AUTHD**

Another way is to look at what others do:

https://github.com/ropensci
https://github.com/ropensci/onboarding/issues

https://github.com/lme4/lme4/
https://github.com/ramnathv/slidify/
https://github.com/rstudio/httpuv/
https://github.com/wch/harbor/
https://github.com/yihui/knitr/
https://github.com/hadley/purrr

“Each [package] is perfect the way it is; and it can use a little improvement.”
—*Shunryu Suzuki*

Warm ups

Your turn

What are the six things most commonly found in a package? (Hint: the most common is R code).

What do they do?

How do you list them within R?

```r
# DESCRIPTION
packageDescription("ggplot2")

# R/
ls("package:ggplot2")

# man/
help(package = ggplot2)
```

```r
# data/
data(package = "ggplot2")

# vignettes/
vignette(package = "dplyr")
browseVignettes("dplyr")

# tests/
# NAMESPACE
```

Your turn

How can you access the source code for an R package?

Find the file that contains purrr::map().

Find the file that contains the documentation for tidyr::extract().

CRAN – Package tidyr

tidyr: Easily tidy data with spread and gather functions

tidyr is an evolution of reshape2. It's design specifically for data tidying (not general reshaping or aggregating) and works well with dplyr data pipelines.

Version:	0.1
Depends:	R (≥ 3.1.0)
Imports:	reshape2, dplyr (≥ 0.2)
Suggests:	knitr, testthat
Published:	2014-07-21
Author:	'Hadley Wickham' [aut, cre]
Maintainer:	'Hadley Wickham' <h.wickham at gmail.com>
License:	MIT + file LICENSE
URL:	https://github.com/hadley/tidyr
NeedsCompilation:	no
Materials:	README
CRAN checks:	tidyr results

Downloads:

Reference manual:	tidyr.pdf
Vignettes:	Tidy data
Package source:	tidyr_0.1.tar.gz
Windows binaries:	r-devel: tidyr_0.1.zip, r-release: tidyr_0.1.zip, r-oldrel: not available
OS X Snow Leopard binaries:	r-release: tidyr_0.1.tgz, r-oldrel: not available
OS X Mavericks binaries:	r-release: tidyr_0.1.tgz

Get the source code for a package

CRAN – Package tidyr

tidyr: Easily tidy data with spread and gather functions

tidyr is an evolution of reshape2. It's design specifically for data tidying (not general reshaping or aggregating) and works well with dplyr data pipelines.

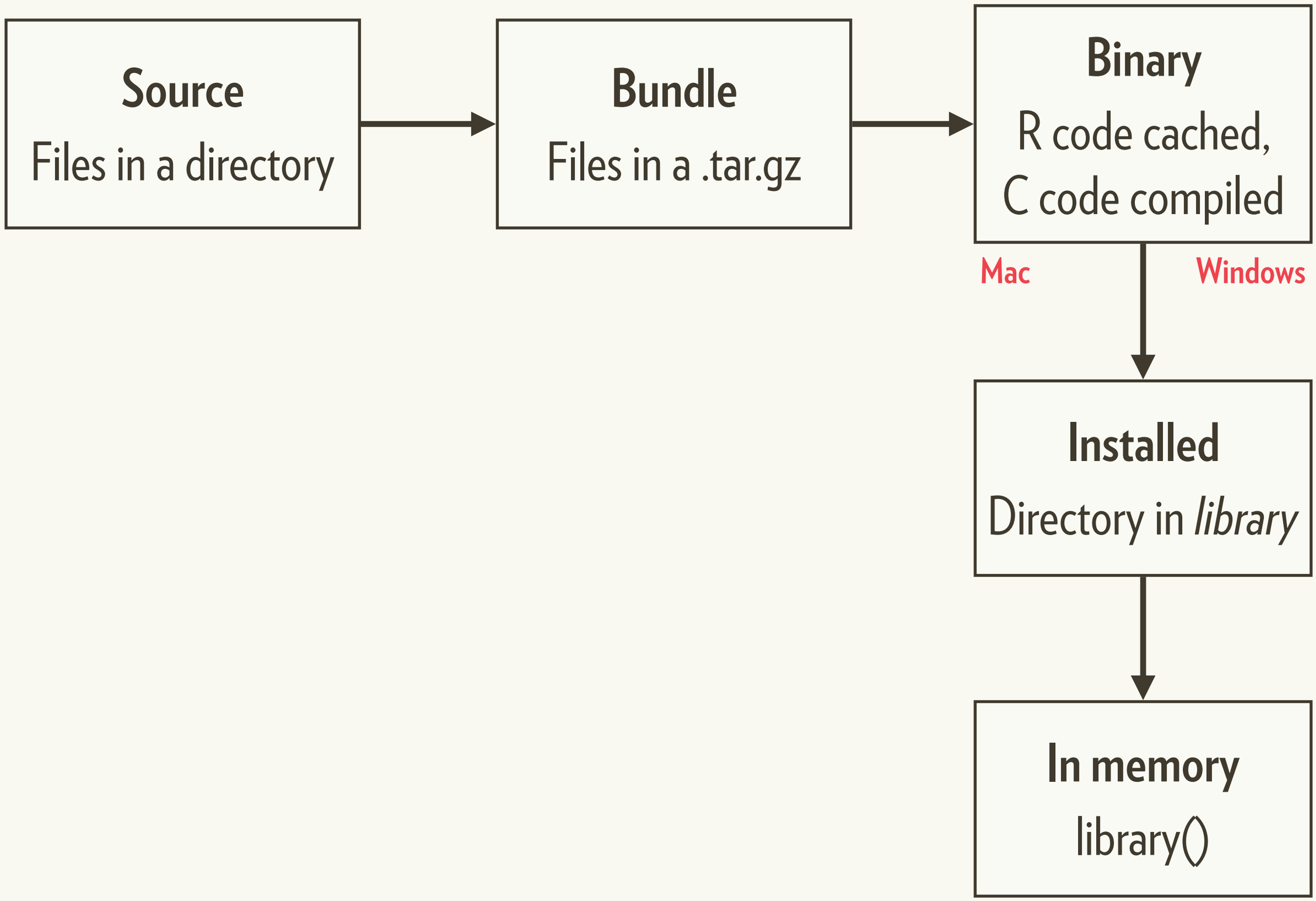
Version:	0.1
Depends:	R (≥ 3.1.0)
Imports:	reshape2, dplyr (≥ 0.2)
Suggests:	knitr, testthat
Published:	2014-07-21
Author:	'Hadley Wickham' [aut, cre]
Maintainer:	'Hadley Wickham' <h.wickham at gmail.com>
License:	MIT + file LICENSE
URL:	https://github.com/hadley/tidyr
NeedsCompilation:	no
Materials:	README
CRAN checks:	tidyr results

Downloads:

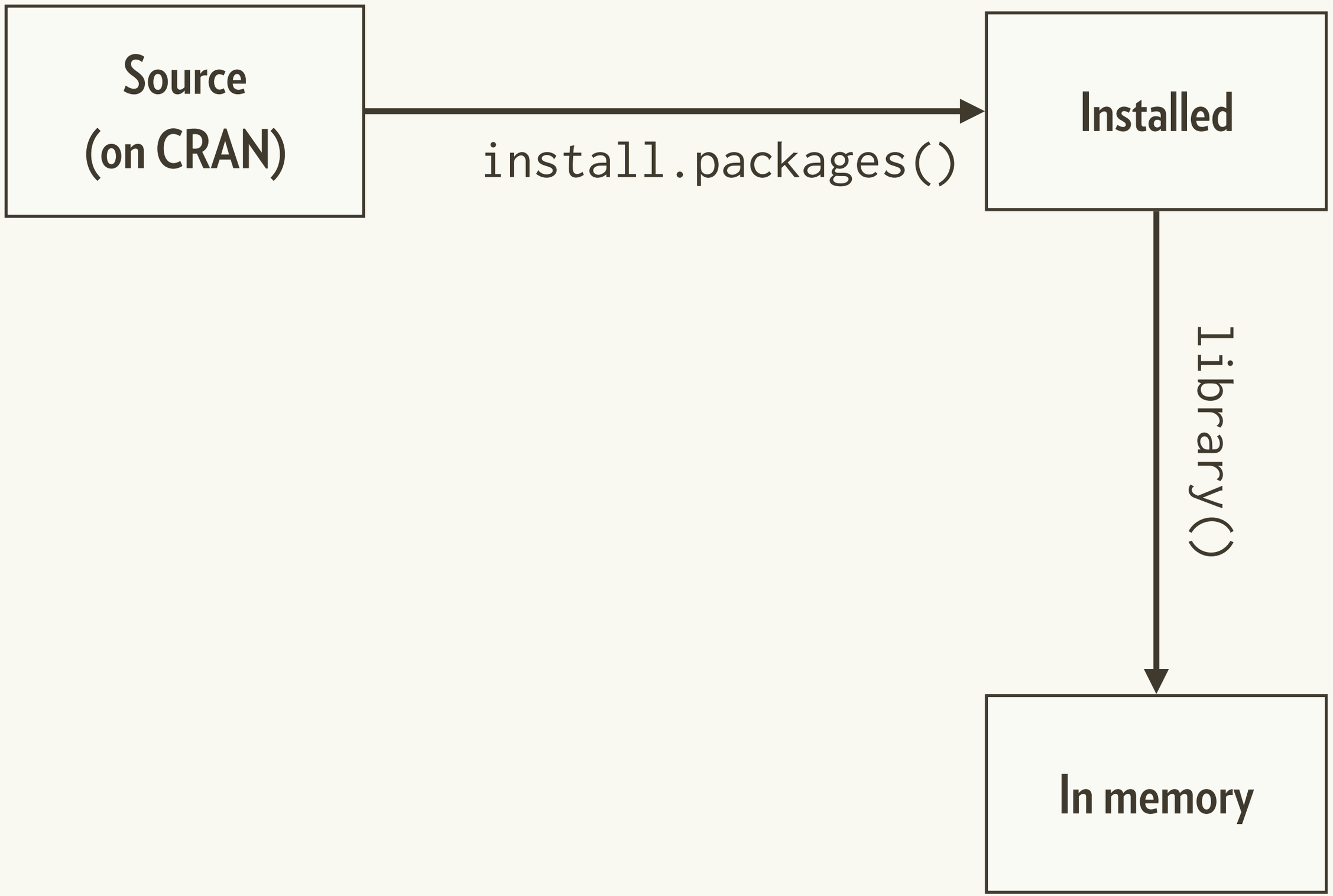
Reference manual:	tidyr.pdf
Vignettes:	Tidy data
Package source:	tidyr_0.1.tar.gz
Windows binaries:	r-devel: tidyr_0.1.zip, r-release: tidyr_0.1.zip, r-oldrel: not available
OS X Snow Leopard binaries:	r-release: tidyr_0.1.tgz, r-oldrel: not available
OS X Mavericks binaries:	r-release: tidyr_0.1.tgz

Don't look at these

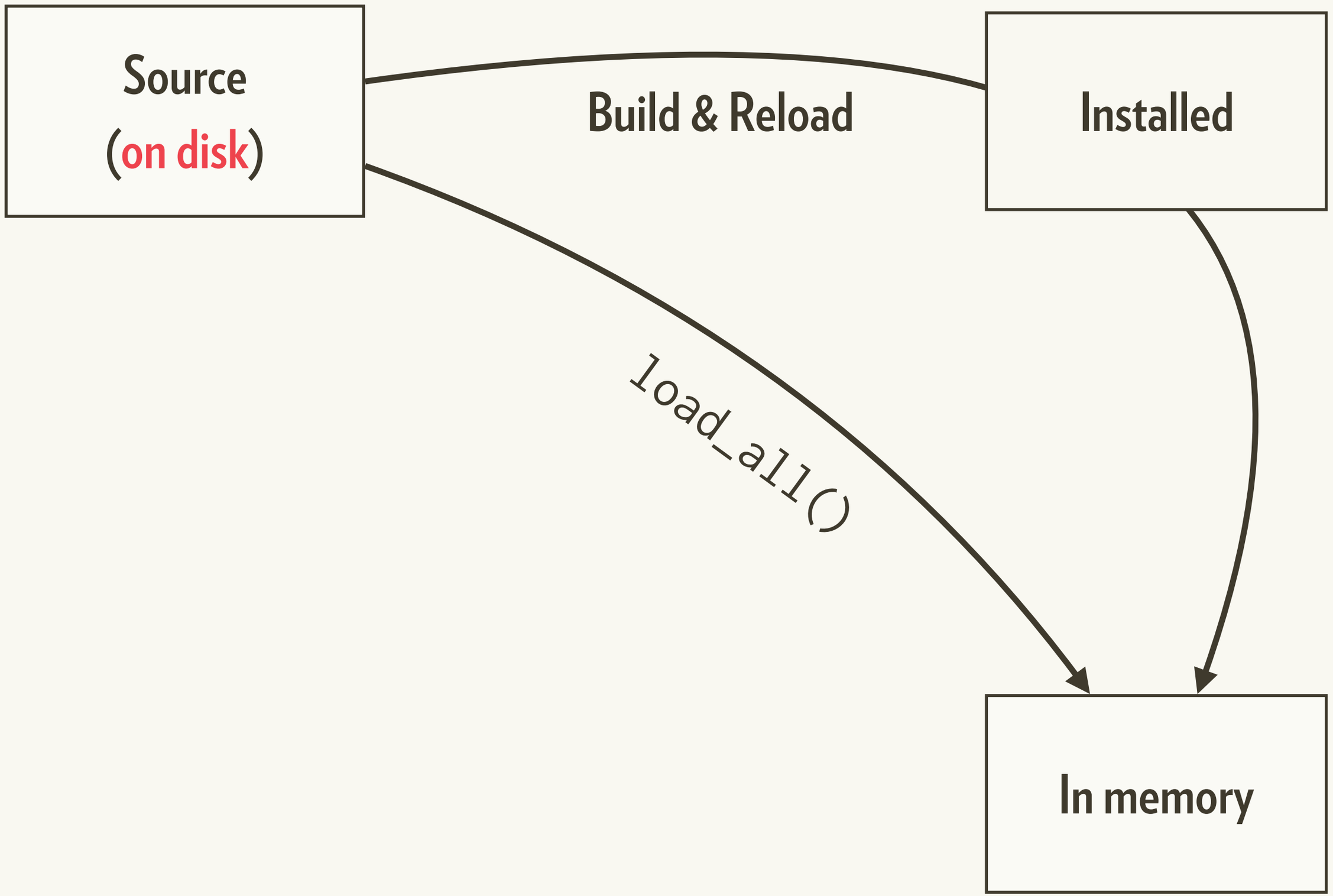
Types of package



These are what you've used in the past



This is what you'll learn today



Discrete random variables

Goal: model random variables in R

x	-1	0	1	2	3
P(x)	0.2	0.1	0.3	0.1	0.3

1.9

Mean

```
library(rv2)
dice <- rv(1:6)
dice
plot(dice)

P(dice > 3)
E(dice)
VAR(dice)

rsim(dice, 100)
plot(dice + dice)
P(dice > dice)
```