

Release

December 2016

Hadley Wickham

[@hadleywickham](https://twitter.com/hadleywickham)

Chief Scientist, RStudio

R CMD check

Automated checking

Runs automated checks for common problems in R packages.

While it has false positives, it is useful for local packages.

If you want to submit to CRAN, you **must pass** R CMD check cleanly.

<http://r-pkgs.had.co.nz/check.html>



This will be frustrating!

==> R CMD build rv2

- * checking for file 'rv2/DESCRIPTION' ... OK
- * preparing 'rv2':
- * checking DESCRIPTION meta-information ... OK
- * installing the package to build vignettes
- * creating vignettes ... OK
- * checking for LF line-endings in source and make files
- * checking for empty or unneeded directories
- * building 'rv2_0.1.tar.gz'

==> R CMD check rv2_0.1.tar.gz

- * using log directory '/Users/hadley/Documents/courses/13-devtools/rv2.Rcheck'
- * using R version 3.0.2 (2013-09-25)
- * using platform: x86_64-apple-darwin10.8.0 (64-bit)
- * using session charset: UTF-8
- * checking for file 'rv2/DESCRIPTION' ... OK
- * this is package 'rv2' version '0.1'
- * checking package namespace information ... OK
- * checking package dependencies ... OK
- * checking if this is a source package ... OK
- * checking if there is a namespace ... OK
- * checking for executable files ... OK
- * checking for hidden files and directories ... OK
- * checking for portable file names ... OK

- * checking for sufficient/correct file permissions ... OK
- * checking whether package 'rv2' can be installed ... OK* checking installed package size ... OK
- * checking package directory ... OK
- * checking DESCRIPTION meta-information ... OK
- * checking top-level files ... OK
- * checking for left-over files ... OK
- * checking index information ... OK
- * checking package subdirectories ... OK
- * checking R files for non-ASCII characters ... OK
- * checking R files for syntax errors ... OK
- * checking whether the package can be loaded ... OK
- * checking whether the package can be loaded with stated dependencies ... OK
- * checking whether the package can be unloaded cleanly ... OK
- * checking whether the namespace can be loaded with stated dependencies ... OK
- * checking whether the namespace can be unloaded cleanly ... OK
- * checking loading without being on the library search path ... OK
- * checking dependencies in R code ... OK
- * checking S3 generic/method consistency ... OK
- * checking replacement functions ... OK
- * checking foreign function calls ... OK
- * checking R code for possible problems ... OK
- * checking Rd files ... OK
- * checking Rd metadata ... OK
- * checking Rd cross-references ... OK
- * checking for missing documentation entries ... OK
- * checking for code/documentation mismatches ... OK

- * checking Rd \usage sections ... OK
- * checking Rd contents ... OK
- * checking for unstated dependencies in examples ... OK
- * checking installed files from 'inst/doc' ... OK
- * checking files in 'vignettes' ... OK
- * checking examples ... OK
- * checking for unstated dependencies in tests ... OK
- * checking tests ...
 - Running 'testthat.R' OK
- * checking for unstated dependencies in vignettes ...
OK
- * checking package vignettes in 'inst/doc' ... OK
- * checking running R code from vignettes ...
 - 'clt.Rmd' ... OK
- OK
- * checking re-building of vignette outputs ... OK
- * checking PDF version of manual ... OK

R CMD check succeeded

Types of problem

ERROR

Must fix!

WARNING

Fix if submitting to CRAN

NOTE

Fix if submitting to CRAN

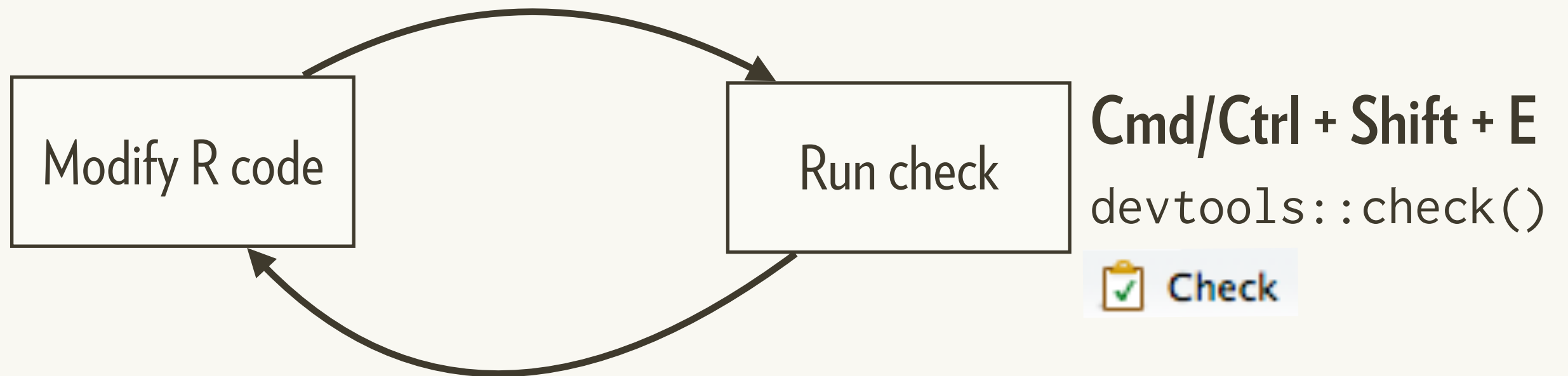
It is possible to submit with a NOTE, but it's best avoided

	Local	CRAN
ERROR	✓	✓
WARNING		✓
NOTE		✓

Change working directory/project to:

[check-me]

Check workflow



Repeat the check workflow until all problems are fixed.

There are **eight** problems for you to find and fix.
Don't look at the answers until you've made an attempt at finding them all.

~~See `'.../rv2.Rcheck/00check.log'` for details.~~

You can ignore

WARNING 'qpdf' is needed for checks on size
reduction of PDFs

No peeking until you've made an attempt!

1. @returns instead of @return in probability.r
2. Error in vignette
3. Description field needs to end in a period.
4. print generic is function(x, ...)
5. plot.rv needs to be explicit (see next slide)
6. Typo in moments documentation (SD not S)
7. # instead of #' in simulation.R
8. Missing "Suggests: testthat" in description

```
#' @export
plot.rv <- function(x, ...) {
  name <- deparse(substitute(x))
  ylim <- range(0, probs(x))

  graphics::plot(as.numeric(x), probs(x), type = "h",
    ylim = ylim, xlab = name,
    ylab = paste0("P(", name, ")"), ...)
  graphics::points(as.numeric(x), probs(x), pch = 20)
  graphics::abline(h = 0, col = "gray")
}
```

CRAN

Submitting to CRAN

```
# Always check with R-devel on windows  
devtools::build_win()
```

```
devtools::use_cran_comments()  
# Update cran-comments.md (next slide)
```

```
# Then submit to CRAN with:  
devtools::release()
```

```
# This asks you a number of questions to check that  
# you've done everything to make the submission as  
# painless as possible
```

```
# It then builds and uploads to CRAN.  
# You'll get a confirmation email a few minutes later
```

cran-comments.md

```
# Create with devtools::use_cran_comments()
```

```
## Test environments
```

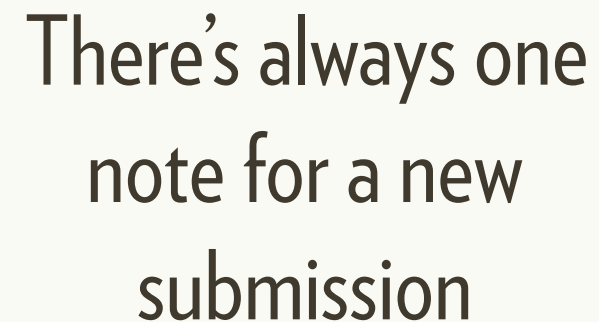
```
* local OS X install, R 3.2.2
```

```
* win-builder (devel and release)
```

```
## R CMD check results
```

```
0 errors | 0 warnings | 1 note
```

```
* This is a new release.
```

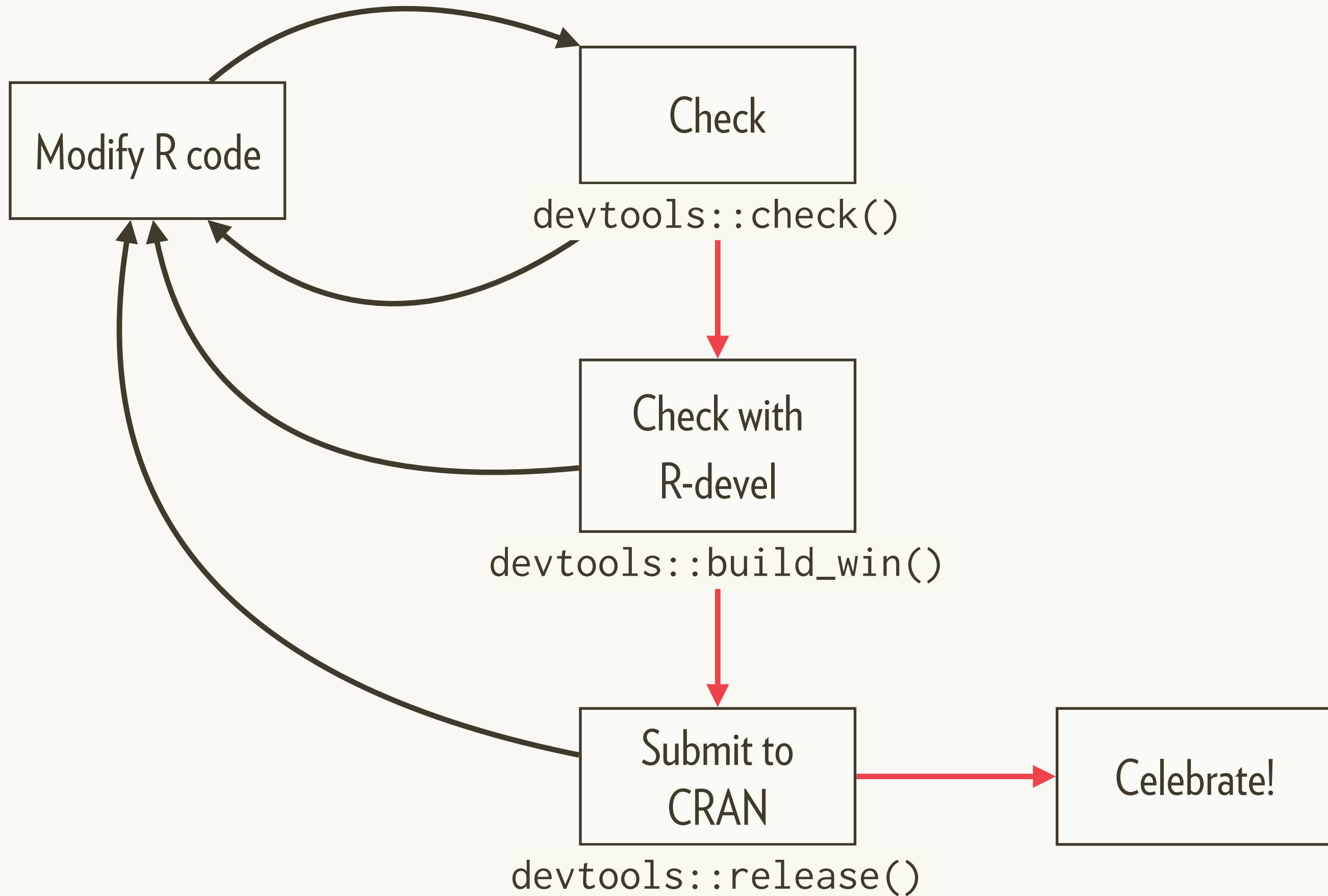


There's always one
note for a new
submission

```
## Reverse dependencies
```

```
This is a new release, so there are no reverse dependencies.
```

Release



If your submission fails

The CRAN maintainers are overwhelmed. ~80% of package submissions need human intervention because of a problem.

Don't worry if you receive a curt response. CRAN treats everyone the same way (even members of R core!) — they're not singling you out.

Ignore any ad-hominem attacks and respond to the content.

Never respond by email. Simply update your cran-comments.md & resubmit. It is not worth arguing.

For resubmission:

This is a resubmission. Compared to the last submission, I have:

- * First change.
- * Second change.
- * Third change

Test environments

- * local OS X install, R 3.2.2
- * win-builder (devel and release)

R CMD check results

...

Devtools guarantee

If a CRAN maintainer yells at you because of a bug in devtools, I'll send you an hand-written apology card. Just forward the email to me, along with your mailing address.

No one will use your package if they don't know it exists



`r-packages@stat.math.ethz.ch`

Git + github

A home on the web

GitHub, Inc. [US] <https://github.com/hadley/dplyr>

This repository Search or type a command Explore Gist Blog Help

hadley / dplyr

Unwatch 31 Star 109 Fork 19

Plyr specialised for data frames: faster & with remote datastores — Edit

935 commits 7 branches 0 releases 4 contributors

branch: master dplyr / Show language statistics

more on #152

romainfrancois authored 2 days ago latest commit a91dc1206d

R	Only create if _doesn't_ exist	3 days ago
data	Recompress nasa data	3 days ago
demo	Add trailing nl	13 days ago
inst	more on #152	2 days ago
man-roxygen	Shift manlp docs into individual srcs	2 months ago
man	Fix incorrect uses of ll with l	3 days ago
src	more on #152	2 days ago
tests	Move to cran version of testthat	3 days ago
vignettes	Vignettes can't use error = TRUE	4 days ago
.Rbuildignore	Ignore man-roxygen in built package	a month ago
.gitignore	Get vignettes workign	2 months ago

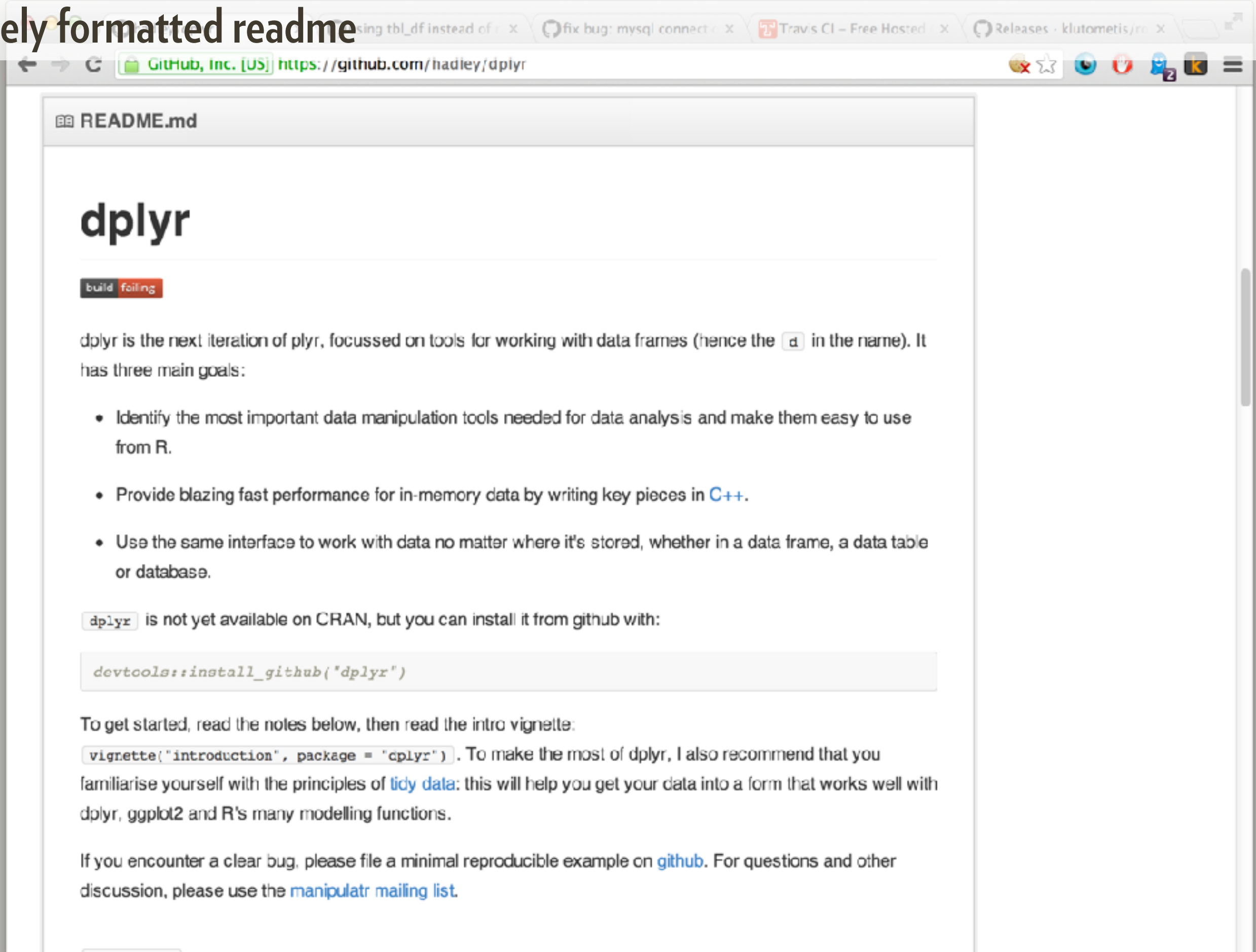
SSH clone URL
git@github.com:hadley:dplyr

You can clone with [HTTPS](#), [SSH](#), or [Subversion](#).

Clone in Desktop

Download ZIP

Nicely formatted readme



The screenshot shows a web browser window with the URL `https://github.com/hadley/dplyr`. The browser's address bar and tabs are visible at the top. The main content area displays the README for the `dplyr` package. The title `dplyr` is prominently displayed at the top left. Below the title, there is a status bar indicating `build` and `failing`. The text describes `dplyr` as the next iteration of `plyr`, focused on data frames. It lists three main goals: identifying important data manipulation tools, providing fast performance by writing key pieces in C++, and using a consistent interface for data manipulation regardless of storage. A code block shows how to install `dplyr` from GitHub using `devtools::install_github("dplyr")`. The text also provides instructions on how to get started, including reading the introduction vignette and familiarizing oneself with tidy data principles. Finally, it offers guidance on how to report bugs and where to ask questions.

README.md

dplyr

build failing

dplyr is the next iteration of plyr, focussed on tools for working with data frames (hence the `d` in the name). It has three main goals:

- Identify the most important data manipulation tools needed for data analysis and make them easy to use from R.
- Provide blazing fast performance for in-memory data by writing key pieces in C++.
- Use the same interface to work with data no matter where it's stored, whether in a data frame, a data table or database.

`dplyr` is not yet available on CRAN, but you can install it from github with:

```
devtools::install_github("dplyr")
```

To get started, read the notes below, then read the intro vignette:

`vignette("introduction", package = "dplyr")`. To make the most of dplyr, I also recommend that you familiarise yourself with the principles of [tidy data](#): this will help you get your data into a form that works well with dplyr, ggplot2 and R's many modelling functions.

If you encounter a clear bug, please file a minimal reproducible example on [github](#). For questions and other discussion, please use the [manipulatr mailing list](#).

See what's happening

using tbl_df instead of raw x

GitHub, Inc. [US] https://github.com/hadley/dplyr/commits/master


This repository Search or type a command Explore Gist Blog Help

hadley / dplyr


Unwatch 31 Star 109 Fork 19


branch: master dplyr / Commits


Dec 07, 2013


 **more on #152** a91dc1206d +
romainfranco's authored 2 days ago Browse code →


Dec 06, 2013


 **using tbl_df instead of raw data.frame for the test** 864c8b7b9f +
romainfranco's authored 2 days ago Browse code →


 **fix test** 58caddc3b3 +
romainfranco's authored 3 days ago Browse code →


 **using NamedListAccumulator** 0c7d40db1a +
romainfranco's authored 3 days ago Browse code →

 **handling date and time constants** 457dd7e8f6 +
romainfranco's authored 3 days ago Browse code →

 **handling special cases in summarise. closes #153** eb2c3566f4 +
romainfranco's authored 3 days ago Browse code →

 **summarise handles constants (#153)** ea873897a9 +
romainfranco's authored 3 days ago Browse code →

 **dealing with the case summarise(df, log(z)) from #153** 7d0fe5d829 +
romainfranco's authored 3 days ago Browse code →

 **dealing with the case summarise(df, diff(z)) from #153** d278dfcec8 +
romainfranco's authored 3 days ago Browse code →

See exactly what changed

using tbl_df instead of raw data.frame for the test

master

romainfrancois authored 2 days ago

1 parent 58eaddc commit 864c8b7b9fae5329d99d818bd78b2b1686bdb167

Showing 2 changed files with 3 additions and 3 deletions.

Show Diff Stats

2

inst/tests/test-summarise.r

View file @ 864c8b7

69	69	@@ -69,7 +69,7 @@ test_that("summarise handles constants (#153)", {
70	70	today <- Sys.Date()
71	71	now <- Sys.time()
72		- res <- summarise(df, int = 1L, num = 1.0,
	72	+ res <- summarise(tbl_df(df), int = 1L, num = 1.0,
73	73	str = "foo", bool = TRUE, date = today, time = now)
74	74	expect_equal(res\$int, 1L)
75	75	expect_equal(res\$num, 1.0)

4

src/dplyr.cpp

View file @ 864c8b7

1024	1024	@@ -1024,14 +1024,14 @@ SEXP summarise_not_grouped(DataFrame df, List args, Environment env){
1025	1025	Rcpp::Shelter<SEXP> __ ;
1026	1026	for(int i=0; i<nexpr; i++){
		SEXP name = names[i] ;
1027		-
1028		- boost::scoped_ptr<Result> res(get_handler(args[i], subsets, env)) ;
	1027	+ Result* res = get_handler(args[i], subsets, env) ;

Track issues

Commits · hadley/dplyr

using tbl_df instead of raw

Issues · hadley/dplyr

GitHub, Inc. [US]https://github.com/hadley/dplyr/issues?milestone=2&state=open

GitHub

This repository

Search or type a command

ExploreGistBlogHelp

hadley

Unwatch31Star109Fork19

Browse Issues

Milestones

New Issue

Everyone's Issues4

Assigned to you0

Created by you3

Mentioning you1

Milestone: v0.1

4 open issues · Due in 7 days

Labels

bug1

enhancement2

Manage Labels

New label

New label name

Clear milestone and label filters

4 Open30 ClosedSort: Newest

☐

Close

Label

Assignee

Milestone

☐

.data.frame versions not using the correct environment

#159

Opened by romainfrancois 2 days ago

2 comments

☐

Mutate fails when creating a constant

bug

#152

Opened by hadley 4 days ago

8 comments

☐

C++ code needs to regularly check for user interrupts

enhancement

#123

Opened by hadley 21 days ago

4 comments

☐

Plyr compatibility

enhancement

#29

Opened by hadley 4 months ago

2 comments

Keyboard shortcuts available

Talk things over

needs to regi x

using tbl_df instead of x x

fix bug: mysql connect x

Travis CI - Free Hosted x

Releases · klutometis/r x

GitHub, Inc. [US] https://github.com/hadley/dplyr/issues/123

hadley opened this issue 21 days ago

C++ code needs to regularly check for user interrupts

Edit

Open

4 comments

Labels

enhancement

romainfrancois is assigned

Milestone: v0.1

So if you accidentally miss-specify a call you can abort it.

2 participants

romainfrancois referenced this issue from a commit 14 days ago

romainfran... adding code to check for interrupts. #123

romainfrancois commented 11 days ago

I'm adding calls to `R_CheckUserInterrupt` here and there. But I'm wondering if this plays well with C++.

I think this uses long jumps and that possibly bypasses C++ destructors. I might ask this on R-devel.

romainfrancois commented 11 days ago

This comes from Simon's contribution to [a thread on R-devel](#) :

```
static void chkIntFn(void *dummy) {
  R_CheckUserInterrupt();
}

// this will call the above in a top-level context so it won't longjmp-out of your context
bool checkInterrupt() {
  return (R_ToplevelExec(chkIntFn, NULL) == FALSE);
}

// your code somewhere ...
```

Receive pull requests

GitHub, Inc. [US] <https://github.com/hadley/dplyr/pull/51>

This repository Search or type a command Explore Gist Blog Help

hadley / dplyr

Unwatch 31 Star 109 Fork 19

Merged hadley merged 1 commit into `hadley:master` from `szilard:master` about 1 month ago 2 #51

Conversation Commits 1 Files Changed 1

szilard opened this pull request 2 months ago
fix bug: mysql connection password was not passed
Edit

No one is assigned No milestone

it was working (only) if user had no password

szilard added a commit 2 months ago
szilard fix bug: mysql connection password was not passed 3ebb93f

hadley referenced this pull request from a commit 2 months ago
hadley Merge pull request #51 from szilard/master ebc7672

Merged hadley merged commit ebc7672 into `hadley:master` from `szilard:master` 2 months ago

Closed hadley closed the pull request 2 months ago

... and review online

The screenshot shows a web browser window displaying a GitHub pull request for the `dplyr` repository. The browser's address bar shows the URL `https://github.com/hadley/dplyr/pull/51/files`. The GitHub navigation bar at the top includes the repository name `hadley / dplyr`, a search bar, and links for `Explore`, `Gist`, `Blog`, and `Help`. The user `hadley` is logged in, as indicated by the profile icon and name in the top right. The pull request header shows it was merged by `hadley` into `hadley:master` from `szilard:master` about a month ago. It has 2 reviews, 109 stars, and 19 forks. Below the header, tabs for `Conversation`, `Commits` (1), and `Files Changed` (1) are visible. A summary states: "Showing 1 changed file with 1 addition and 1 deletion." A `Show Diff Stats` button is also present. The main area displays the diff for the file `R/src-mysql.r`. The diff shows a change on line 98, where a minus sign (`-`) is replaced by a plus sign (`+`) to add a `password` parameter to the `dbi_connect` function call. The code snippet is as follows:

```
@@ -95,7 +95,7 @@ src_mysql <- function(dbname, host = NULL, port = GL, user = "root",
95 95   }
96 96
97 97   con <- dbi_connect(MySQL(), dbname = dbname , host = host, port = port,
98 -   username = user, ...)
98 +   username = user, password = password, ...)
99 99   info <- db_info(con)
100 100
101 101   src_sql("mysql", con,
```

 A sidebar on the right contains icons for code navigation and search. At the bottom, a yellow tip box reads: "Tip: You can add notes to lines in a file. Hover to the left of a line to make a note." The footer of the page includes the copyright notice "© 2013 GitHub, Inc.", links for `Terms`, `Privacy`, `Security`, and `Contact`, the GitHub logo, and links for `Status`, `API`, `Training`, `Shop`, `Blog`, and `About`.

GitHub, Inc. [US] <https://github.com/hadley/dplyr/pull/51/files>

This repository Search or type a command Explore Gist Blog Help hadley

hadley / dplyr Unwatch 31 Star 109 Fork 19

Merged hadley merged 1 commit into hadley:master from szilard:master about 1 month ago 2 #51

Conversation Commits 1 Files Changed 1

Showing 1 changed file with 1 addition and 1 deletion. Show Diff Stats

R/src-mysql.r Open View file @ 3ebb93f

```
@@ -95,7 +95,7 @@ src_mysql <- function(dbname, host = NULL, port = GL, user = "root",
95 95   }
96 96
97 97   con <- dbi_connect(MySQL(), dbname = dbname , host = host, port = port,
98 -   username = user, ...)
98 +   username = user, password = password, ...)
99 99   info <- db_info(con)
100 100
101 101   src_sql("mysql", con,
```

Tip: You can add notes to lines in a file. Hover to the left of a line to make a note

© 2013 GitHub, Inc. Terms Privacy Security Contact Status API Training Shop Blog About

Find out as soon as you break something

The screenshot shows the Travis CI web interface for the repository **klutometis/roxygen**. The build status is **passing**. The left sidebar lists recent repositories, with **klutometis/roxygen** highlighted. The main content area shows the build details for the current build, including the commit hash **b510eaa (master)** and the message **First pass at functions in usage. #164**. A terminal window at the bottom displays the build script, which includes cloning the repository, checking out the specific commit, and running tests.

Travis CI for Private Repositories

Search all repositories

My Repositories Recent

- ropensci/RMendeley** 3
1 min 17 sec
3 minutes ago
- klutometis/roxygen** 97
3 min 55 sec
about 2 hours ago
- hadley/dplyr** 192
3 min 51 sec
2 days ago
- hadley/testthat** 12
1 min 33 sec
2 days ago
- hadley/assertthat** 10
1 min 30 sec
2 days ago

klutometis/roxygen

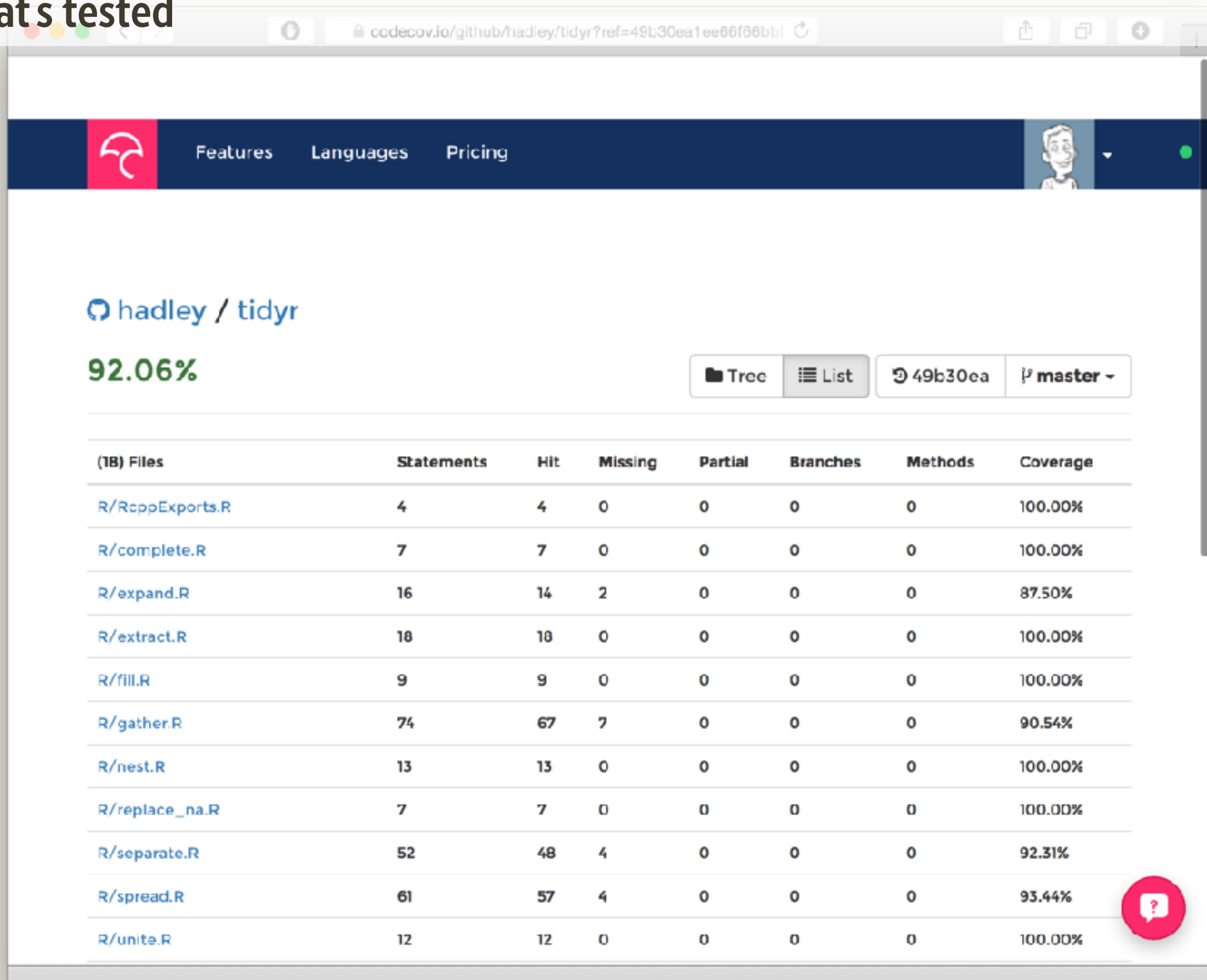
build passing

Current Build History Pull Requests Branch Summary

Build	97	Commit	b510eaa (master)
State	Passed	Compare	93fe512416d8...b510eaa00df
Finished	about 2 hours ago	Author	hadley
Duration	3 min 55 sec	Committer	hadley
Message	First pass at functions in usage. #164		

```
1 Using worker: worker-linux-5-2.bb.travis-ci.org:travis-linux-3
2
3 $ export global=[
4 $ export WARNINGS_ARE_ERRORS=1
5 $ export CC=gcc
6 $ git clone --depth=50 --branch=master
13 $ cd klutometis/roxygen
14 $ git checkout -qf b510eaa00df548d02305d5f233dab3d54420b9f
15 $ gcc --version
16 gcc (Ubuntu/Linaro 4.6.3-1ubuntu5) 4.6.3
17 Copyright (C) 2011 Free Software Foundation, Inc.
18 This is free software; see the source for copying conditions. There is NO
19 warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
20
21 $ curl -OL http://raw.githubusercontent.com/craigcitro/r-
26 $ chmod 755 ./travis-tool.sh
27 $ ./travis-tool.sh bootstrap
322 $ ./travis-tool.sh install_deps
574 $ ./travis-tool.sh run_tests
575 + CRAN=http://cran.rstudio.com
```

See what's tested



```
80
81 #' @export
82 gather_.data.frame <- function(data, key_col, value_col, gather_cols,
83                                na.rm = FALSE, convert = FALSE) {
84   ## Return if we're not doing any gathering
16 16 if (length(gather_cols) == 0) {
86   1 return(data)
87 16 }
88
89 15 gather_idx <- match(gather_cols, names(data))
90 15 if (anyNA(gather_idx)) {
91   missing_cols <- paste(gather_cols[is.na(gather_idx)], collapse = ", ")
92   stop("Unknown column names: ", missing_cols, call. = FALSE)
93 15 }
94 15 id_idx <- setdiff(seq_along(data), gather_idx)
95
96   ## Get the attributes if common, NULL if not.
97 15 args <- normalize_melt_arguments(data, gather_idx, factorsAsStrings = TRUE)
98 15 measure.attributes <- args$measure.attributes
99 15 factorsAsStrings <- args$factorsAsStrings
100 15 valueAsFactor <- "factor" %in% measure.attributes$class
101
102 15 df <- melt_dataframe(data,
103 15   id_idx = 1L,
104 15   gather_idx = 1L,
105 15   as.character(key_col),
106 15   as.character(value_col),
107 15   as.pairlist(measure.attributes),
108 15   as.logical(factorsAsStrings),
109 15   as.logical(valueAsFactor)
110 15 )
111
112 14 if (na.rm) {
113   1 missing <- is.na(df[[value_col]])
114   1 df <- df[!missing, ]
115 14 }
116
117 14 if (convert) {
118   df[[key_col]] <- type.convert(as.character(df[[key_col]]), as.is = TRUE)
119 14 }
120
```

Beautiful release notes

using tbl_df instead of r x

fix bug: mysql connect x

Travis CI - Free Hosted x

Releases · klutometis/roxygen

GitHub, Inc. [US]https://github.com/klutometis/roxygen/releases

This repository

Search or type a command

ExploreGistBlogHelp

hadley

klutometis / roxygen

Unwatch10

Star65

Fork32

Releases / Tags

Draft a new release

Latest release

v3.0.0

abe5d0d

roxygen2 3.0.0

hadley released this 3 days ago · 3 commits to master since this release

Edit

Roxygen2 now fully supports S4 and RC (reference classes) - you should no longer need to manually add `@alias` or `@usage` tags for S4 classes, methods and generics, or for RC classes.

- The default usage definitions are much better, generating the correct usage for data sets ([#122](#)), S3 methods (without additional `@method` tag), S4 generics, S4 methods, and for replacement ([#119](#)) and infix functions. Backslashes in function arguments in are correctly escaped. Usage statements also use a more sophisticated line wrapping algorithm so that they should cause fewer problems with the R CMD check line limit. ([#89](#), [#125](#)).
- S4 classes, S4 methods, and RC classes are given better default topics, and the file names corresponding to those topics are shorter.
- S4 methods will automatically inherit parameter documentation from their generic.
- `@slot name description` allows you to document the slots of a S4 class.

Devtools comes with helpers to get you started

```
use_git()  
use_github()  
use_travis()  
use_coverage()
```

Easy to install a package from github

```
library(devtools)  
install_github("hadley/devtools")
```

```
# Install specific tag  
install_github("hadley/devtools@v1.4.1")
```

```
# Github is a viable location for package  
# distribution, but it doesn't yet have the  
# reach or authority of CRAN. Currently  
# about ~4,000 packages on github.
```

More resources

<http://happygitwithr.com>

<http://r-pkgs.had.co.nz/git.html>

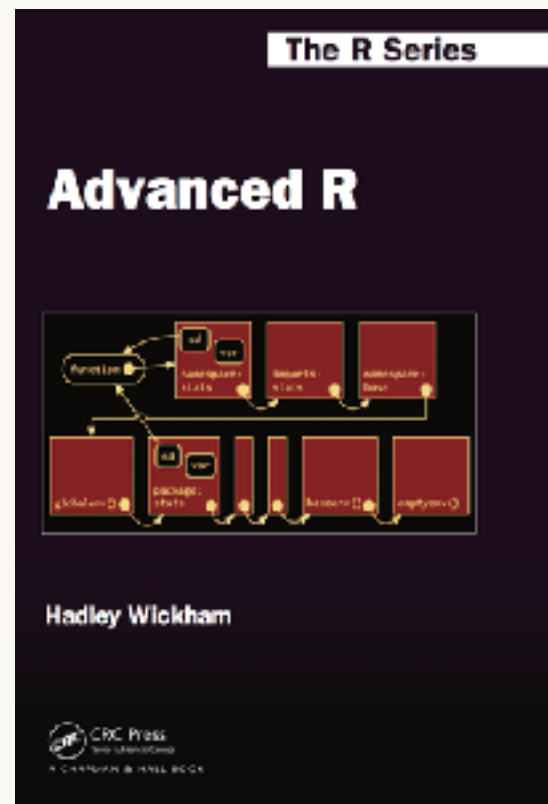
<http://help.github.com/>

<http://try.github.com/>

<http://gitref.org/>

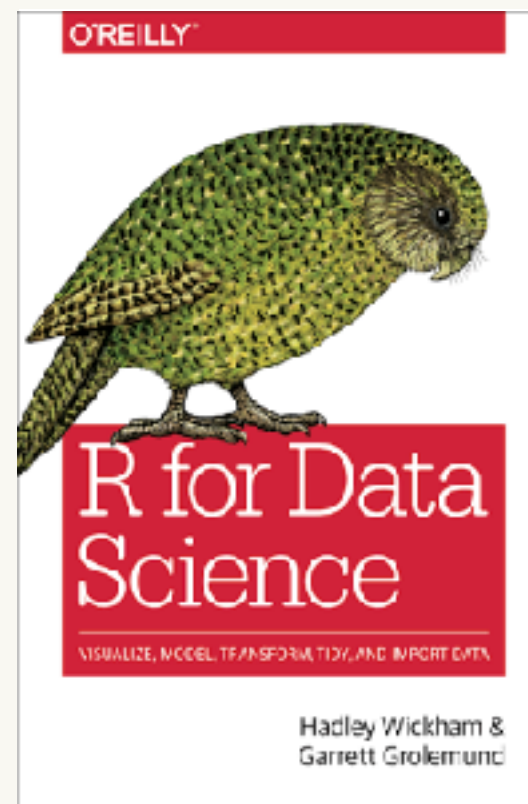
http://www.rstudio.com/ide/docs/version_control/overview

Learning more



<http://adv-r.had.co.nz/>

<http://amzn.com/1466586966>



<http://r4ds.had.co.nz>

Use AUTHD discount



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

Use AUTHD discount

Other resources



[r] score:5 is:question closed:no



<https://rweekly.org>

<https://stat.ethz.ch/mailman/listinfo/r-devel>

<https://groups.google.com/d/forum/rdevtools>

Friendly, low-traffic mailing list for package development questions

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/>