

# An Analysis of Unit Testing on CRAN

Rory Nolan

06/05/2017

Let us take a look at how many packages on CRAN employ unit tests by checking whether their source code contains a non-empty `tests/` directory (using <https://rdrr.io>). We examine first all packages and then the subset that were authored or updated in 2016 or 2017. We will also compare the popularity of the testing frameworks `testthat`, `RUnit`, `svUnit` `svUnit` and `unitizer` [code adapted from `mango`].

```
required_pkgs <- c("tidyverse", "stringr", "lubridate", "rvest", "RCurl")
invisible(lapply(required_pkgs, library, character.only = TRUE))
check_for_tests <- function(pkg_name) {
  message(pkg_name)
  rdrr_address <- paste0("https://rdrr.io/cran/", pkg_name)
  if (url.exists(rdrr_address)) {
    if (inherits(try(read_html(rdrr_address), silent = TRUE), "try-error")) {
      NA # Allow for error reading rdrr.io page
    } else {
      rdrr_page_text <- rdrr_address %>% read_html %>% html_text
      trimmed_lines <- str_split(rdrr_page_text, "\n") %>% unlist %>% str_trim()
      trimmed_lines %>% str_detect("^tests/.+") %>% any
    }
  } else {
    NA # allow for not found on rdrr.io
  }
}

download.file("http://cran.R-project.org/web/packages/packages.rds",
              "packages.rds", mode = "wb")
cran_packages_info <- readRDS("packages.rds")
cran_packages <- cran_packages_info[, "Package"]
post2015 <- ymd(cran_packages_info[, "Published"]) >= ymd("2016-01-01")
new_packages <- cran_packages[post2015]

tested_check <- cran_packages %>%
map_lgl(check_for_tests) %>%
set_names(cran_packages) %>%
na.omit # allow for not found on rdrr
n_packages <- length(tested_check)
n_tested_packages <- sum(tested_check)
print(c(n_packages, n_tested_packages))

#> [1] 10084 2566

new_packages <- intersect(new_packages,
                          names(tested_check)) # allow for not found on rdrr
n_new_packages <- length(new_packages)
n_tested_new_packages <- sum(tested_check[new_packages])
print(c(n_new_packages, n_tested_new_packages))

#> [1] 5715 1984
```

```

unit_test_packages <- c("testthat", "RUnit", "svUnit", "unitizer")
reverse_deps <- tools::package_dependencies(packages = unit_test_packages,
                                           cran_packages_info, recursive=FALSE, reverse=TRUE,
                                           which = c("Depends", "Imports", "LinkingTo", "Suggests"))
framework_use <- map_int(reverse_deps, length)
print(framework_use)

```

```

#> testthat    RUnit    svUnit unitizer
#>      1979      133      11      0

```

So we see that (at the time of writing) of the 10084 packages on CRAN, 25% (2566) are unit tested. Also, of the 5715 packages authored or updated since 1st January 2016, 35% (1984) are unit tested. We also see that testthat is the most popular testing framework, preferred in 93% of cases.