Software Engineering Standards in the R Community

by Oliver Keyes, Jennifer Bryan, David Robinson

Abstract R flourishes because of its wide range of user-submitted packages, providing generalised solutions to real-world problems. Crucial to these packages being useful, adopted and trusted by useRs is that they follow reasonable software engineering expectations, from unit tests to user-friendly documentation. We performed a quantitative and qualitative analysis of CRAN-hosted packages, examining their use of unit tests, internal consistency and documentation standards, along with many other variables. We report on this analysis and suggest best practises for writing new R packages, along with proposals to improve the standards of existing, widely-used packages.

Introduction

Best Practices in Software Engineering

There is no single list of "best practices" for writing high quality software, but there are some general traits that such software possesses. From the perspective of the user, software should be accurate, fast-running and easy to use. From the perspective of the developer, the internal code should be maintainable, portable and lend itself to being tested McConnell (2004). We can point to specific conventions or expectations that are built on these traits:

1. **Unit and integration tests**: code that tests whether code run by the user is fit for use, by performing operations and checking that the results are as expected. This touches on both "accuracy" (it prevents the wrong result being provided to the user) and "maintainability" (it gives the developer a way of conveniently checking that modifications have not broken functionality before releasing software). In R, unit tests can be created in an *ad-hoc* fashion or using a pre-existing testing framework, such as *RUnit* Burger et al. (2015) or *testthat* Wickham (2011).
2. **Documentation**, both for individual functions and

Practices in CRAN

Some figures and analyses, like Figure 1.

Bibliography

M. Burger, K. Juenemann, and T. Koenig. *RUnit: R Unit Test Framework*, 2015. URL `http://CRAN.R-project.org/package=RUnit`. R package version 0.4.28. [p1]

S. McConnell. *Code Complete*. Microsoft Press, 2nd edition, 2004. ISBN 0735619670. [p1]

H. Wickham. testthat: Get started with testing. *The R Journal*, 3:5–10, 2011. URL `http://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf`. [p1]

Oliver Keyes
Wikimedia Foundation

`ironholds@gmail.com`

Jennifer Bryan
University of British Columbia

`jenny@stat.ubc.ca`

David Robinson
Princeton University

`admiral.david@gmail.com`

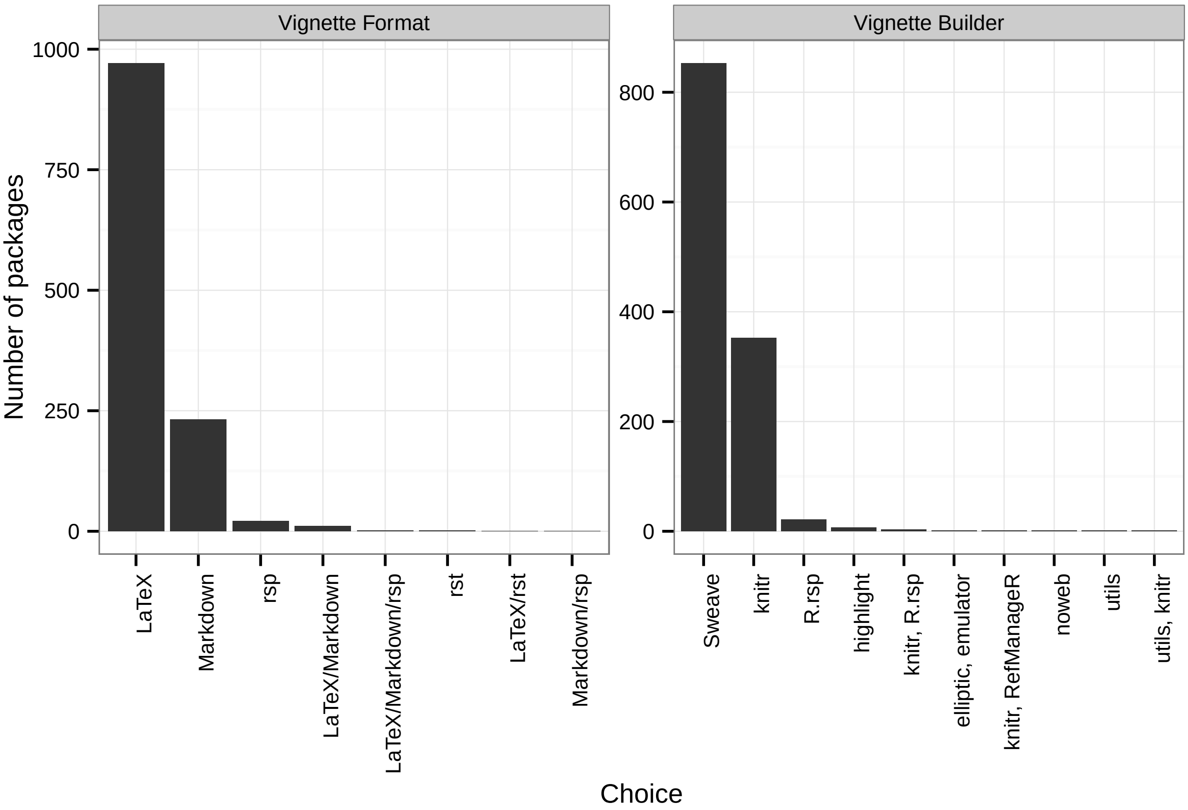

Figure 1: Distribution of the choice of vignette builder and format, among the 19.7% of CRAN packages that use vignettes.