

R Package tutorial

Dr Patricio Estévez Soto

JDI Open, Jan 26, 2021

What today is about

- to introduce participants to the development of programming packages in the context of academic research and reproducible science, and
- to produce a (basic) working R package

Packages and modular programming

What are packages?

- TL:DR: packages allow the distribution of (mainly) functions so that you don't have to do everything from scratch.
- "**Modular programming** is a [software design](#) technique that emphasizes separating the functionality of a [program](#) into independent, interchangeable **modules**, such that each contains everything necessary to execute only one aspect of the desired functionality." ([Wikipedia](#), 2021)
- Core language (R, Python, Go, Julia, JavaScript, etc.)
- Package/module (ggplot2, tensorflow, strings, Flux, MathJS)

Packages and science

- Most contemporary academic research uses computers
- Packages are a core element of a researcher's toolbox
- What packages do you use? (Language agnostic)

From user to developer



This Photo by Unknown Author is licensed under [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/)

1. Develop research methods and tools

texreg: Conversion of R Regression Output to LaTeX or HTML Tables

Converts coefficients, standard errors, significance stars, and goodness-of-fit statistics of statistical models into LaTeX tables or HTML tables/MS Word documents or to nicely formatted screen output for the R console for easy model comparison. A list of several models can be combined in a single table. The output is highly customizable. New model types can be easily implemented. (If the Zelig package, which this package enhances, cannot be found on CRAN, you can find it at <<https://github.com/IQSS/Zelig>>.)

Version: 1.37.5
Depends: R (≥ 3.5)
Imports: methods, stats, [httr](#)
Suggests: [broom](#) (≥ 0.4.2), [coda](#) (≥ 0.19.2), [ggplot2](#) (≥ 3.1.0), [huxtable](#) (≥ 4.2.0), [knitr](#) (≥ 1.22), [rmarkdown](#) (≥ 1.12.3), [sandwich](#) (≥ 2.3-1), [testthat](#) (≥ 2.0.0), [lmtest](#) (≥ 0.9-34)
Enhances: [AER](#), [alpaca](#), [bife](#), [biglm](#), [brms](#) (≥ 2.8.8), [btergm](#), [dynlm](#), [feisr](#) (≥ 1.0.1), [forecast](#), [gamlss](#), [gamlss.inf](#), [glmmTMB](#), [lfe](#), [lme4](#), [miceadds](#), [mlogit](#), [mnlogit](#), [MuMIn](#), [nlme](#), [nnet](#), [ordinal](#), [plm](#), [speedglm](#), [survival](#), [VGAM](#), [Zelig](#)
Published: 2020-06-18
Author: Philip Leifeld [aut, cre], Claudia Zucca [ctb]
Maintainer: Philip Leifeld <philip.leifeld@essex.ac.uk>
BugReports: <https://github.com/leifeld/texreg/issues>
License: [GPL-3](#)
URL: <http://github.com/leifeld/texreg/>
NeedsCompilation: no
SystemRequirements: pandoc (≥ 1.12.3) suggested for using wordreg function; LaTeX packages tikz, booktabs, dcolumn, rotating, thumbpdf, longtable, paralist for the vignette
Citation: [texreg citation info](#)
In views: [ReproducibleResearch](#)
CRAN checks: [texreg results](#)

pscl: Political Science Computational Laboratory

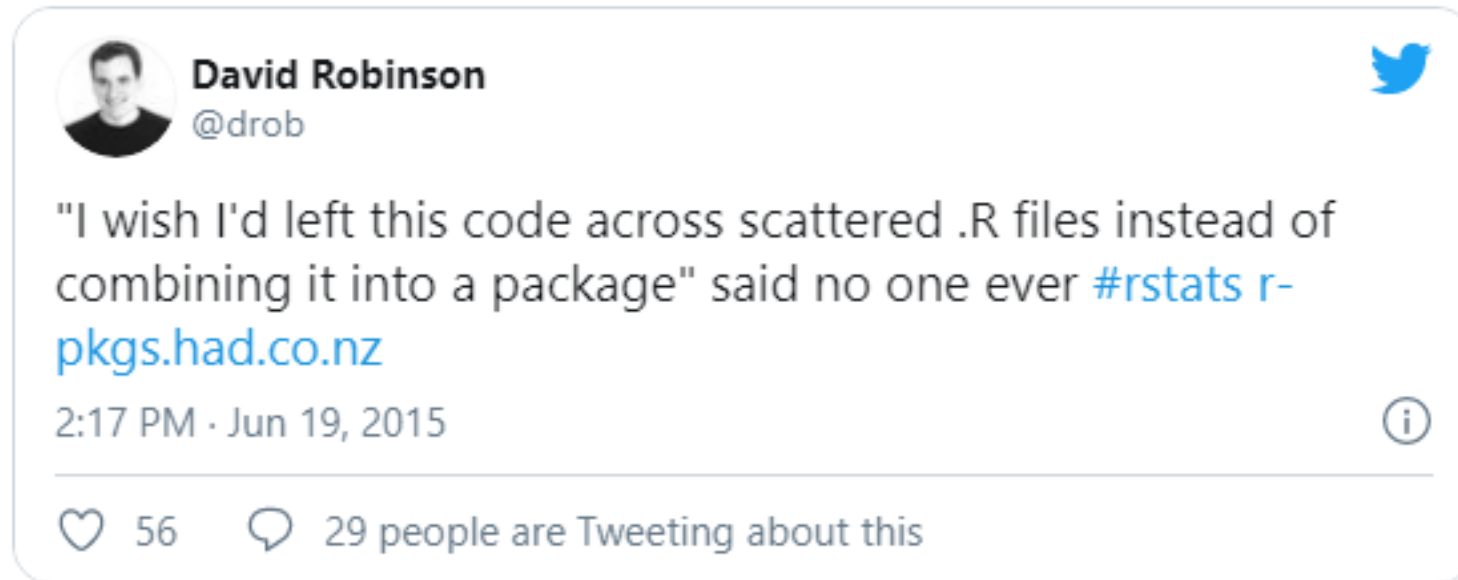
Bayesian analysis of item-response theory (IRT) models, roll call analysis; computing highest density regions; maximum likelihood estimation of zero-inflated and hurdle models for count data; goodness-of-fit measures for GLMs; data sets used in writing and teaching at the Political Science Computational Laboratory; seats-votes curves.

Version: 1.5.5
Imports: [MASS](#), datasets, grDevices, graphics, stats, utils
Suggests: [lattice](#), [MCMCpack](#), [car](#), [lmtest](#), [sandwich](#), [zoo](#), [coda](#), [vcd](#), [mvtnorm](#), [mgcv](#)
Published: 2020-03-07
Author: Simon Jackman, with contributions from Alex Tahk, Achim Zeileis, Christina Maimone, Jim Fearon and Zoe Meers
Maintainer: Simon Jackman <simon.jackman@sydney.edu.au>
License: [GPL-2](#)
Copyright: see file [COPYRIGHTS](#)
URL: <http://github.com/atahk/pscl>
NeedsCompilation: yes
Citation: [pscl citation info](#)
Materials: [NEWS](#)
In views: [Bayesian](#), [Econometrics](#), [Environmetrics](#), [Psychometrics](#), [SocialSciences](#)
CRAN checks: [pscl results](#)

2. To share reproducible research ([Marwick et al. 2018](#))

- Research compendia
 - A convenient method to wrap all your data and materials, including reproducibility scripts
- Papers with open data/materials get more
 - Citations
 - Publications
 - And more
- Licensing
 - For articles, documentation, examples: CC
 - For code: MIT, GPL, GNU (<https://choosealicense.com/>)
- Version control
 - Git + Github
- Persistence
 - DOI Repositories
 - Zenodo
 - OSF

3. To make your life easier



My personal R packages

- [prestevez/victim](#)

prestevez/victim: Victimisation Analysis and Reporting

A convenience package to facilitate analysis and reporting of victimisation data analysis. Though developed for this purpose, functions can be applied to different fields where the analytical needs are similar (e.g., studying prevalence, count data distributions).

Getting started

[README.md](#)

Browse package contents

- [Vignettes](#)
- [Man pages](#)
- [API and functions](#)
- [Files](#)

Search within the prestevez/victim package

Package details

Maintainer	
License	MIT
Version	0.0.0.9000
Package repository	View on GitHub
Installation	Install the latest version of this package by entering the following in R: <pre>install.packages("remotes") remotes::install_github("prestevez/victim")</pre>

prestevez/victim documentation built on Jan. 30, 2020, 7:14 p.m.

- [prestevez/crimeineq](#)

prestevez/crimeineq:

More about what it does (maybe more than one line) Use four spaces when indenting paragraphs within the Description.

Getting started

Browse package contents

- [Vignettes](#)
- [Man pages](#)
- [API and functions](#)
- [Files](#)

Search within the prestevez/crimeineq package

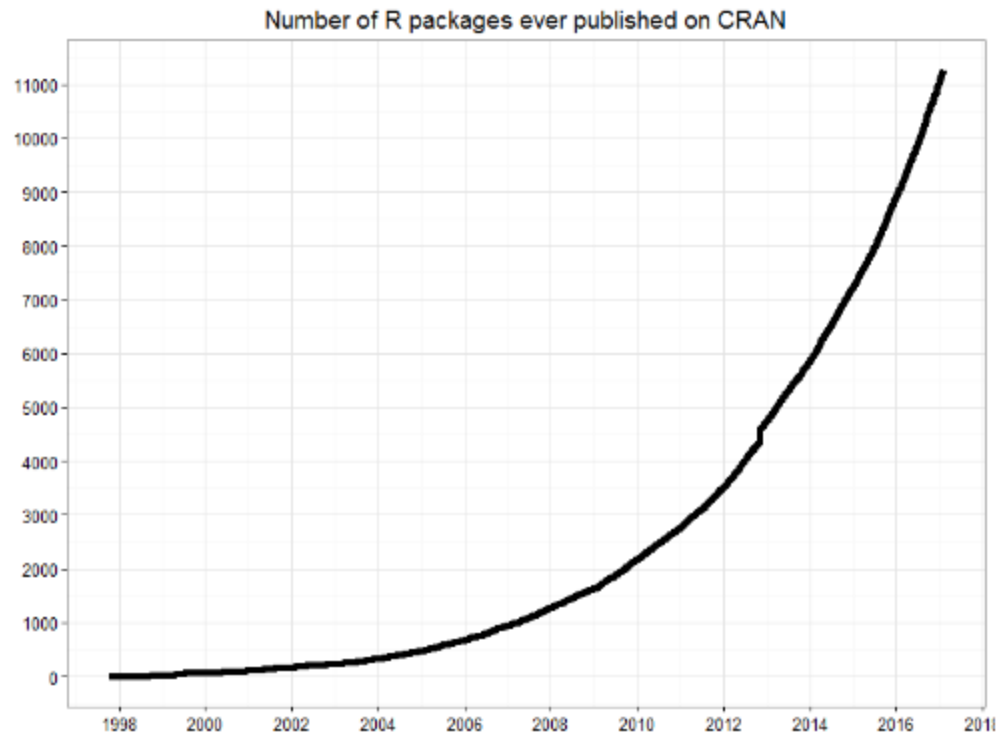
Package details

Author	Who wrote it
Maintainer	The package maintainer <yourself@somewhere.net>
License	What license is it under?
Version	0.1.0
Package repository	View on GitHub
Installation	Install the latest version of this package by entering the following in R: <pre>install.packages("remotes") remotes::install_github("prestevez/crimeineq")</pre>

prestevez/crimeineq documentation built on May 29, 2019, 7:39 a.m.

R packages

- Over 10,000 on CRAN



Source: <https://blog.revolutionanalytics.com/2017/01/cran-10000.html>

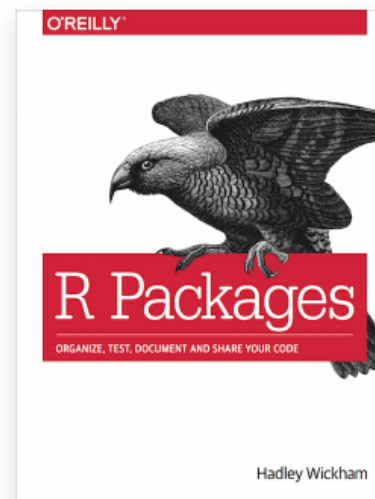


The book of packages

Welcome!

Welcome to R packages by [Hadley Wickham](#) and [Jenny Bryan](#). Packages are the fundamental units of reproducible R code. They include reusable R functions, the documentation that describes how to use them, and sample data. In this book you'll learn how to turn your code into packages that others can easily download and use. Writing a package can seem overwhelming at first. So start with the basics and improve it over time. It doesn't matter if your first version isn't perfect as long as the next version is better.

This is the work-in-progress 2nd edition of the book.



[Preface »](#)

Tutorial

- Go to: <https://www.prestevez.com/post/r-package-tutorial/>
- Open a new R studio instance.