Hosting Data Packages via ‘drat’: A Case Study with Hurricane Exposure Data

Data reproducibility session, UseR! 2017

Brooke Anderson, Colorado State University
✉: brooke.anderson@colostate.edu
🐦: @gbwanderson
🌍: www.github.com/geanders

Dirk Eddelbuettel, Debian and R Projects, Ketchum Trading
✉: edd@debian.org
🐦: @eddelbuettel
🌍: www.github.com/eddelbuettel
Motivation

Goal: Create a publicly-accessible package with historical data (1988–present) on county-level exposure to Atlantic basin tropical storms in the United States (with exposure based on distance to storm track, rain, wind, floods, or tornadoes).
Motivation

"Data" package

hurricaneexposuredata
Host with drat.

"Code" package

hurricaneexposure
Available through CRAN.
Motivation

"Data" package

hurricaneexposedata
Host with drat.

"Code" package

hurricaneexposure
Available through CRAN.

Challenges

1. Host a package with drat.
2. Connect a CRAN package with a package hosted with drat.
Hosting a package with **drat**

**drat creates R repositories**
Provides functions to create and managed repositories for use by `install.packages()`, `update.packages()` etc

**Extensive resources**
- Drat Basics for Package Users
- Drat Basics for Package Authors
- Drat FAQs
Coding a CRAN package to use a drat package

"Data" package

'hurricaneexposedata'
Host with drat.

"Code" package

'hurricaneexposure'
Available through CRAN.

Challenges

1. Host a package with drat.
2. Connect a CRAN package with a package hosted with drat.
“Unfortunately it is out of our area of expertise to do deep level troubleshooting on an individual’s computer to find the cause of the issue. In this case I would recommend using a colleague’s computer and see if that is successful in creating a submission.”

– Example of recent non-R technical support
Technical support experience, R

1. Search GitHub
1. Technical support experience, R

2. Brooke Anderson
@gbwanderson

Can anyone point me to a CRAN package that imports a drat repo package? I'm trying to fix `build_win` build errors. @eddelbuettel #rstats
1. Search GitHub

2. Brooke Anderson
@gbwanderson

Can anyone point me to a CRAN package that imports a drat repo package? I'm trying to fix `build_win` build errors. @eddelbuettel #rstats

3. Hosting Data Packages via drat: A Case Study with Hurricane Exposure Data
by G. Brooke Anderson, Dirk Eddelbuettel
What do you need in a CRAN package to use drat?

- “Not much”, really.
- Tell R about the additional repository via
  - repos argument directly in `install.packages()` et al, or
  - setting another repos field in `options()`,
  - using drat helper functions `addRepo()` or `drat:::add()`
- Full details about repos in `help(download.packages)`
What is this thing about GitHub?

- Hosting on github particularly easy because
  - every github repo has an (optional) web presence
  - a github username plus ‘drat’ is a unique URL

- So if we assume drat as the (github) repository name
- Then we only need username: addRepo("geanders")
Coding an optional data package to use drat

Making use of R packaging infrastructure

- The DESCRIPTION file has field Additional_Repositories
- Optional packages can reside on Additional_Repositories
  - Mandatory packages (Imports:, Depends:, LinkingTo:) cannot
  - But Suggests: can point there – feature we use
- Good use case for drat repo as Additional_Repositories
- But also ensure you test for presence of optional package
How do we test for optional package?

```r
.pkgenv <- new.env(parent=emptyenv())

.onLoad <- function(libname, pkgname) {
  has_data <- requireNamespace("hurricaneexposuredata", quietly = TRUE)
  .pkgenv[['has_data']] <- has_data
}

.onAttach <- function(libname, pkgname) {
  if (!.pkgenv$has_data) {
    msg <- paste("To use this package, you must install the",
                 "hurricaneexposuredata package. To install that",
                 "package, run `install.packages('hurricaneexposuredata',",
                 "repos='https://geanders.github.io/drat/', type='source')`.",
                 "See the `hurricaneexposure` vignette for more details.")
    msg <- paste(strwrap(msg), collapse="\n")
    packageStartupMessage(msg)
  }
}

hasData <- function(has_data = .pkgenv$has_data) {
  if (!has_data) {
    msg <- paste("To use this function, you must have the",
                 "`hurricaneexposuredata` package installed. See the",
                 "`hurricaneexposure` package vignette for more details.")
    msg <- paste(strwrap(msg), collapse="\n")
    message(msg)
    return(invisible(NULL))
  }
}
```
Article in The R Journal

drat documentation

- Drat Basics for Package Users
- Drat Basics for Package Authors
- Drat FAQs

Acknowledgments