

# Statistical Computing

Michael Mayer

March 2023



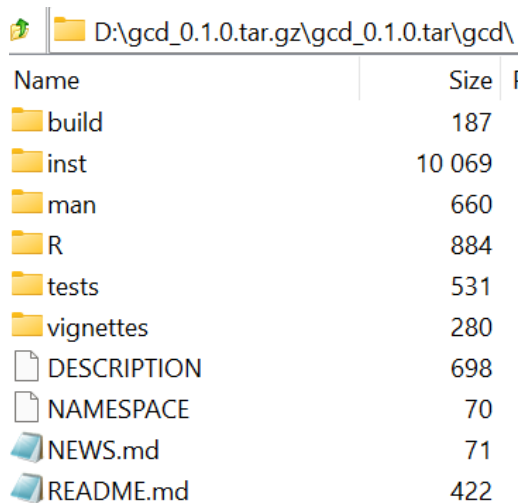
# R Packages: Do it yourself

- ▶ About 20'000 contributed packages on CRAN
- ▶ Many others not on CRAN
- ▶ Making the **second** package is easy

## Resources

- ▶ Writing extensions:  
<https://cran.r-project.org/doc/manuals/r-release/R-exts.html>
- ▶ R Packages (Hadley Wickham and Jennifer Bryan): <https://r-pkgs.org>

## Content of Simple Package “gcd”



A screenshot of a Windows File Explorer window showing the contents of the directory 'D:\gcd\_0.1.0.tar.gz\gcd\_0.1.0.tar\gcd\'. The window displays a list of files and folders with their names, sizes, and icons. The files and folders are: 'build' (187 bytes), 'inst' (10 069 bytes), 'man' (660 bytes), 'R' (884 bytes), 'tests' (531 bytes), 'vignettes' (280 bytes), 'DESCRIPTION' (698 bytes), 'NAMESPACE' (70 bytes), 'NEWS.md' (71 bytes), and 'README.md' (422 bytes).

Name	Size	Icon
build	187	Folder
inst	10 069	Folder
man	660	Folder
R	884	Folder
tests	531	Folder
vignettes	280	Folder
DESCRIPTION	698	Text file
NAMESPACE	70	Text file
NEWS.md	71	Markdown file
README.md	422	Markdown file

- ▶ R → R code
- ▶ man → Help files
- ▶ build, inst, vignettes  
→ Vignette
- ▶ tests → Unit tests
- ▶ Other files?

## How to create all this??

1. Put functions into one or more R scripts in folder “R”
2. Document functions with Roxygen

### The rest requires three things

- ▶ Package “usethis”: Cares about content of package
- ▶ Package “devtools”: Turns content into package
- ▶ Software bundle RTools: <https://cran.r-project.org/bin/windows/Rtools>

### Example

<https://github.com/mayer79/gcd>

# How to Contribute to other Projects?

## Github project

- ▶ You see issue
- ▶ or missing functionality
- ▶ or simply a typo

## Example

[https://github.com/  
duckdblabs/duckplyr](https://github.com/duckdblabs/duckplyr)

## Example Workflow

1. Fork project on Github
2. Clone your fork locally
3. Create “feature” branch
4. Commit changes to this branch and push
5. Open pull request (PR) to original project
6. If not okay, add more commits locally and push
7. When merged, delete feature branch

## More Complex Packages?

- ▶ What if your package uses other packages like "ggplot2"?  
→ Dependency hell, so minimize their use
- ▶ What if you use `print()`, `summary()` etc.?

### Example

Pandemic apero: <https://github.com/mayer79/apero>