

# R-{packaging}

BernR User Group Meeting at the SNSF

Rachel Heyard

19.12.2019

# Why?

anything that can be automated, should be automated - H. Wickham

- automate your work using functions ...
- have everything you need using one line of code ...

```
library("snf.preprocessing")
```

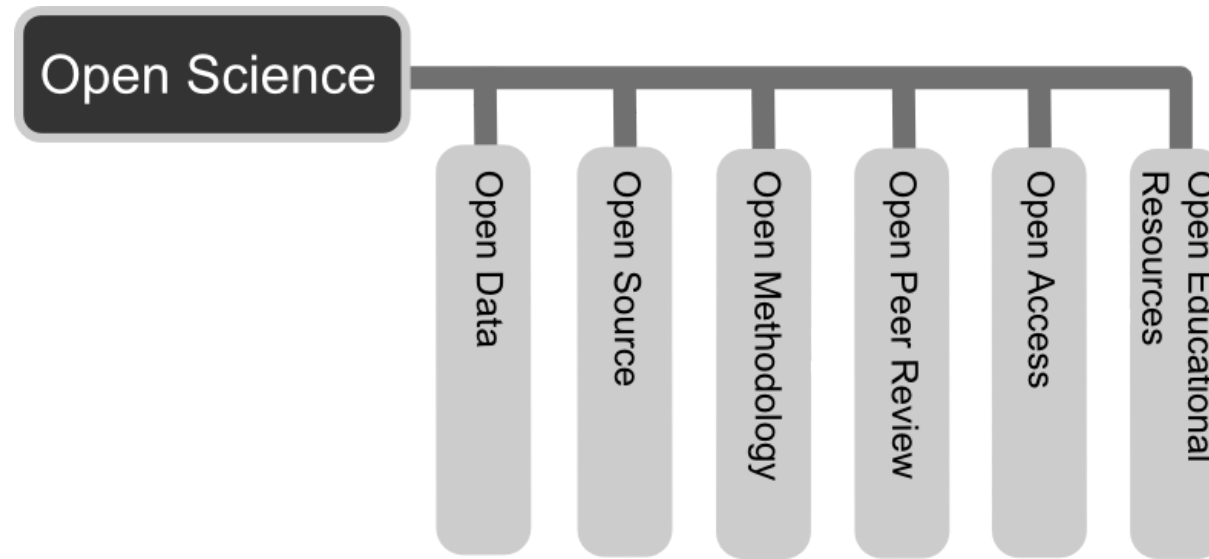
- easily share your code and functions with colleagues, or the world ...
- easily introduce corporate design ...

## Menu of the evening:

- two use cases: academic and enterprise
- demo

# Academic use case

- facilitates reproducible research
- in the open science framework



the open science framework (wiki)

- publish your functions in a package on CRAN or github

# Example {TBFmultinomial}


## Background

*during my time at UZH I wrote a methodological paper extending an approach to Bayesian variable selection to discrete competing risk models.*



RESEARCH PAPER | [Full Access](#)

## Dynamic clinical prediction models for discrete time-to-event data with competing risks—A case study on the OUTCOMEREA database

Rachel Heyard , Jean-François Timsit, Wafa Ibn Essaied, Leonhard Held, [on behalf of the COMBACTE-MAGNET consortium](#)

# Example {TBFmultinomial} (2)

**My thoughts..**

*tired of sourcing several R scripts with functions or - worse - copy-pasting code chunks*

***solution:*** *I build my own local package*

*I shared the .tar.gz with my collaborators, who could run my code easier*

*I uploaded and installed the package on the server where I did the computations*

**..getting more serious**

*I **should** make my code available before submitting the paper!*

*--> will facilitates peer review*

*why CRAN?*

# Example {TBFmultinomial} (3)

*submit my package on the CRAN website*

CRAN submission TBFmultinomial 0.1.0

Package Submission

An: Rachel Heyard

Dienstag, 31. Oktober 2017 15:53

[Details anzeigen](#)

Dear Rachel Heyard  
Someone has submitted the package TBFmultinomial to CRAN.  
You are receiving this email to confirm the submission as the maintainer of this package.  
To confirm the submission to CRAN, follow or copy & paste the following link into your browser:

[http://xmpalantir.wu.ac.at/cransubmit/conf\\_mail.php?code=e6801946f5f34d2f77a0ee38f4daf6c8](http://xmpalantir.wu.ac.at/cransubmit/conf_mail.php?code=e6801946f5f34d2f77a0ee38f4daf6c8)

If you did not submit the package or do not wish for it to be submitted to CRAN, simply ignore this email

Submission Information:

Submitter: Rachel Heyard <rachel.heyard@uzh.ch>

Package: TBFmultinomial

Version: 0.1.0

Title: TBF Methodology Extension for Multinomial Outcomes

Author(s): Rachel Heyard [aut, cre]

Maintainer: Rachel Heyard <rachel.heyard@uzh.ch>

Depends: VGAM, nnet, parallel, stats, stringr, plotrix, methods

Suggests: knitr, splines

Description: Extends the TBF methodology to multinomial regression models and discrete time-to-event models with competing risks.

License: GPL (>= 2)

Submitter's comment: Initial submission

# Example {TBFmultinomial} (4)

Re: CRAN submission TBFmultinomial 0.1.0

Donnerstag, 02. November 2017 13:32

An: Rachel Heyard, CRAN

[Details anzeigen](#)

Thanks, please explain the acronym TBF in your description.

Please add a reference for the method in the 'Description' field of your DESCRIPTION file in the form

authors (year) <doi:...>

authors (year) <arXiv:...>

authors (year, ISBN:...)

with no space after 'doi:', 'arXiv:' and angle brackets for auto-linking.

Please add more small executable examples in your Rd-files.

Something like

```
\examples{
  examples for users:
  executable in < 5 sec
  for checks
  \dontshow{
    examples for checks:
    executable in < 5 sec together with the examples above
    not shown to users
  }
  donttest{
    further examples for users (not used for checks)
  }
}
```

would be desirable.

Please fix and resubmit.

Best,

Re: CRAN submission TBFmultinomial 0.1.1

Donnerstag, 02. November 2017 16:29

An: Rachel Heyard, CRAN

[Details anzeigen](#)

Thanks, on CRAN now.

Best,

# Example `{TBFmultinomial}` (5)

available on [CRAN](#)

## 4 DEVELOPMENT OF CLINICAL PREDICTION MODELS FOR DISCRETE COMPETING RISKS DATA

We are considering discrete time-to-event models and need to account for the time-dependency of the cause-specific baseline hazard  $\theta_{0kr}$  in Equation 4. A method to determine which time-specification best fits the data is discussed in Section 4.1. Then we aim to select relevant predictors for our model using objective Bayesian variable selection as presented in Section 4.2. Choosing among baseline and time-dependent variables with different lags can result in a vast model space that needs to be considered and drastically slows down the computation of crucial quantities. Therefore we decide to use a sequential approach by first electing the best lags for the different time-dependent variables and then adding potential baseline predictors in the selection process. To further improve the model we proceed to cause-specific variable selection (CSVS), explained in Section 4.3, by dropping the outcome-specific coefficients that are not meaningful and correcting the relevant ones. In a last step, we repeat the variable selection and the CSVS steps on landmark-specific datasets that are conditioned on the fact that patients are still at risk at the landmark. Finally, we obtain dynamic prediction models for the timing of a VAP PA, see Section 4.4. The proposed methodology is implemented in R in the package `TBFmultinomial` available from R-CRAN. A vignette is provided with the package explaining its use. The `multinom()` function from the `nnet` package as well as the `vglm()` function from the `VGAM` package can be used to fit multinomial regression models with time-dependent intercepts resulting in discrete-time competing risks models.



# Enterprise use case

## at the Swiss National Science Foundation

to simplify the life of the analysts at the SNSF working with the *datahub* an **SNSF R environment** was set up, containing

- RStudio Connect,
- RStudio Server,
- and documented and unified data processing and plotting functions in an R-package

# Example {snf}-package collection

The {snf}-package collection includes two packages:

- **snf.preprocessing**: dedicated to import data from the SNSF *datahub* and process it to be used in monitoring reports *etc*
- **snf.plot**: contains plot-functions in the SNSF corporate design
- available on the SNSF-**miniCRAN**

# How?

- building a package with package documentation has become super easy using some helper packages:
  - devtools
  - roxygen2
  - pkgdown
  - fledge
  - ...
- package components:
  - scripts with code in R/
  - package metadata, the DESCRIPTION
  - documentation in man/
  - namespace, the NAMESPACE
  - vignette (vignette/), data (data/), ...

# Demo on RStudio Server

## Minimal steps

1. start a new *package-project*
2. add scripts with function(s) and document with roxygen2-code
3. `devtools::document()`
4. maybe: `devtools::check()`
5. build the (source) package
6. if needed: `pkgdown::build_site()`

# Convinced? Convaincu?

...time to ask more questions or start building yourself