



Targets package

Pipeline tool for data science in R

Aud H. Halbritter

@audhalbritter

UNIVERSITY OF BERGEN





What is the targets package?

Pipeline tool for statistics and data science in R

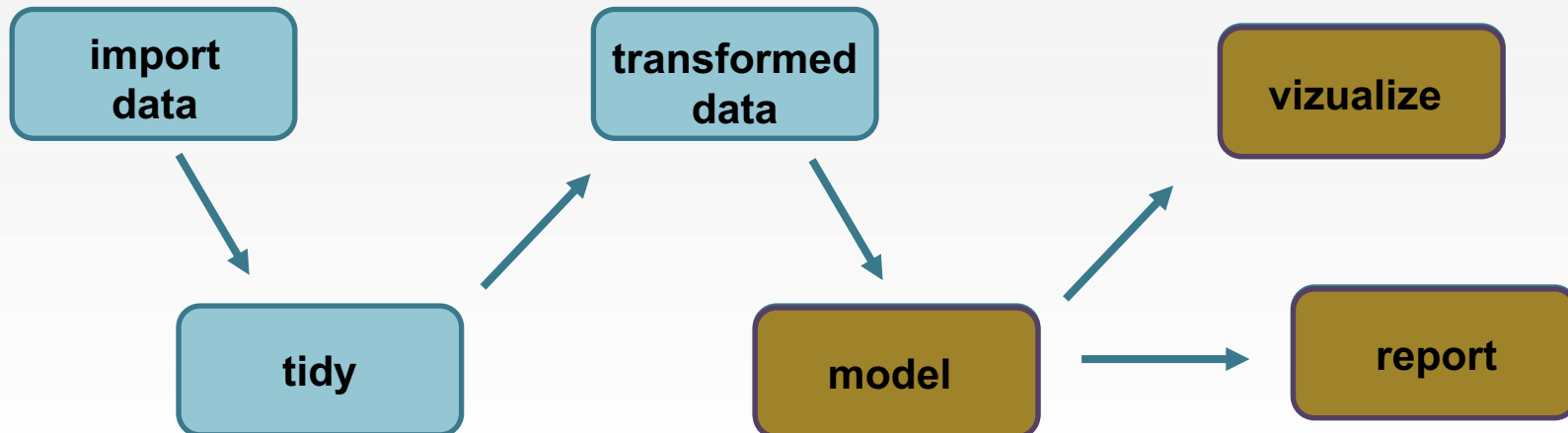
- Reproducible workflows avoiding repetition
- Skips costly running time for up to date tasks





When is targets useful?

- Code has long runtimes (slow or complex)
- Interconnected tasks with dependencies
- Different outputs (e.g. Rmarkdown)



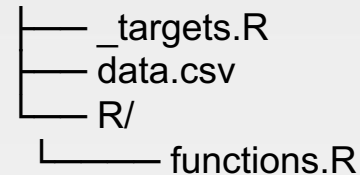
Source: <https://github.com/ropensci/targets>



How does it work?

File structure

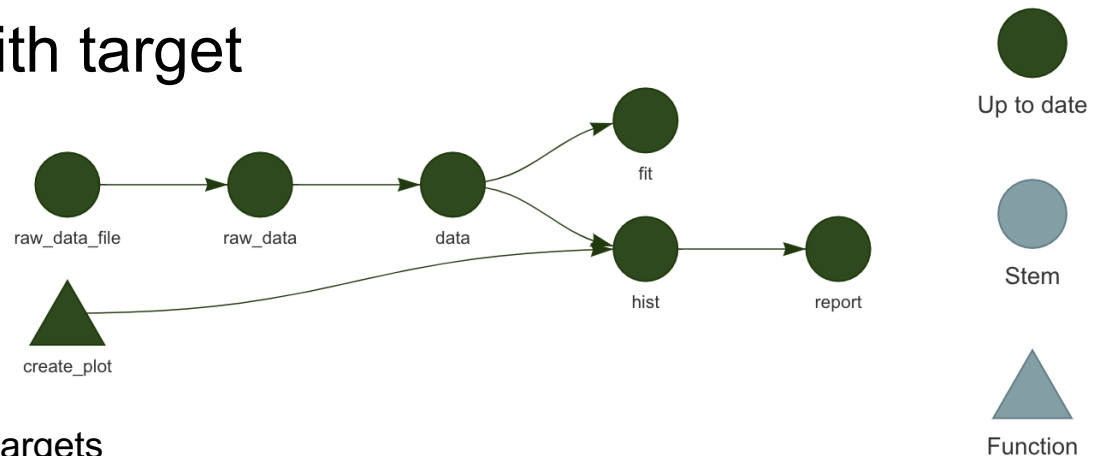
- Data
- R functions



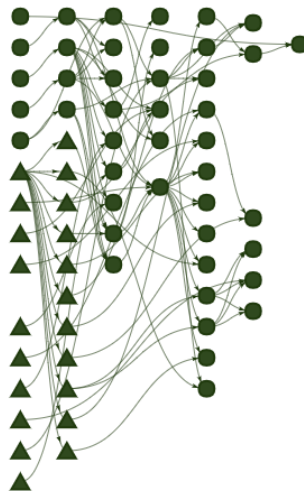
Target script file (`_target.R`)

Configures and defines the pipeline

- Loads packages and functions
- Makes pipeline with target objects



```
> tar_make()
● start target raw_data_file
● built target raw_data_file
● start target raw_data
● built target raw_data
● start target data
● built target data
● start target hist
● built target hist
● start target fit
● built target fit
● start target report
● built target report
● end pipeline
```





Further reading / watching

GitHub page:

<https://github.com/ropensci/targets>

User manual:

<https://books.ropensci.org/targets/>

Minimal example:

<https://github.com/wlandau/targets-minimal>

Larger example:

<https://github.com/richardjtelford/neotomaTargets>

