

Tools and best practices for R package development

Dario Righelli | Università degli Studi di Padova

Abstract

The R scripting language is widely popular in the statistical field for data analysis modeling, but often resulting in custom scripts or functions mostly interpretable only by its own developer.

Nowadays, reproducibility is becoming always more requested, not only by the scientific community, but also by scientific journals that require to share, together with the scientific insights, data and software necessary to reproduce the findings.

Despite it could be reasonably easy to implement a data analysis model in a custom script, it is not always straightforward to implement such models in a reproducible way.

One way to meet these requirements is to develop custom packages to facilitate not only the reproducibility of the analysis, but also to allow third-party users to reuse the methods in their analysis.

In this seminar, we will explore how to create a new package starting from a set of existing functions and tools to facilitate the documentation (roxygen), testing (test_that), code version control (git), and continuous integration (Github Actions).