

# New Tools for Reproducible Research with R

Yihui Xie<sup>1,\*</sup>, J.J. Allaire<sup>2</sup>

1. Department of Statistics, Iowa State University

2. RStudio

\*Contact author: [xie@yihui.name](mailto:xie@yihui.name)

**Keywords:** Reproducible, Sweave, knitr, RStudio, LyX

The practice of reproducible research (Schwab et al. [4]) seeks to make scholarship and data analysis more transparent, verifiable, recreateable, and reliable. *R* includes a rich set of tools for reproducible research based on Sweave (Leisch [1]), a system for dynamic document authoring.

In this talk we will discuss and demonstrate several new tools that have the potential to make reproducible research with *R* more flexible, powerful, and accessible to a broader audience. The first of these is the **knitr** package (Xie [6]), which follows in the steps of Sweave and adds several new capabilities including caching, tikz graphics, and code highlighting and reformatting. There are also features motivated by feedback from Sweave users such as conditional evaluation and code externalization. Graphics capabilities have also been significantly enhanced. This includes caching, displaying multiple plots per chunk, control over individual plot sizes, flexible arrangement (side by side or even as animations), and compatibility with non-standard plots from other packages like **rgl** and **rggobi**. Knitr supports traditional Sweave documents as well as provides several built-in output hooks which can be used to write results in  $\LaTeX$ , HTML and Markdown.

One of the difficulties in persuading users of traditional authoring environments to move to a reproducible workflow is the availability of tools that are strong for both document editing and computation. We'll demonstrate some new capabilities of the RStudio integrated development environment (RStudio [3]) designed to bridge this gap. RStudio combines TeX authoring features such as spell-checking, error navigation, and smart PDF previewing with Sweave and knitr savvy code editing tools. We will also demonstrate the support for Sweave and knitr that have been added to the LyX document editor (LyX [2]).

One of the cornerstones of reproducible research is authoring in plain-text formats. Recently several lightweight markup languages (Wikipedia [5]) such as Org-mode, Markdown, and AsciiDoc have emerged and have the potential to make authoring reproducible research more straightforward and accessible to a much broader audience. We'll discuss these new formats and the various ways that they can be used with *R*.

## References

- [1] Leisch, F. (2002). *Sweave: Dynamic Generation of Statistical Reports Using Literate Data Analysis*. Physica Verlag, Heidelberg. ISBN 3-7908-1517-9.
- [2] LyX (2012). LyX 1.6.1 - The Document Processor. <http://www.lyx.org>.
- [3] RStudio (2012). Rstudio: Integrated development environment for r. <http://www.rstudio.org>.
- [4] Schwab, M., M. Karrenbach, and J. Claerbout (2000). Making scientific computations reproducible. *Computing in Science Engineering* 2(6), 61–67.
- [5] Wikipedia (2012). Lightweight markup language. [http://en.wikipedia.org/wiki/Lightweight\\_markup\\_language](http://en.wikipedia.org/wiki/Lightweight_markup_language).
- [6] Xie, Y. (2012). knitr: A general-purpose package for dynamic report generation in r. <http://yihui.github.com/knitr/>.