

Reproducible Research in R and \LaTeX

February 21, 2014

Reproducible Research

...refers to the idea that the ultimate product of research is the paper with the full computational environment used to produce the results in the paper such as the data, code and these can be used to reproduce the results.

A paper by Robert Gentleman and Duncan Temple Lang

www.reproducibleresearch.org

a web site by Roger Koenker

Literate programming

Literate programming means that text, data, and computer code are interwoven in a single self-contained document.

What we are used to

A research document involving multiple files with figures and tables cut and paste from various places. For instance,

- a statistical software e.g stata do-file
- an excel spreadsheet with results
- an excel spreadsheet with data
- a directory with filenames like *old.doc* and *new.doc*
- a word document with tables and figures cut and paste from various places

*Changes to the stata do-file are not automatically propagated to the excel spread-sheet or to the Word document.

Benefits of Literate Programming

Integration of analysis and reporting

- No transcription errors
- Automatic updating of results. i.e dynamic reports
- Reproducible results

Benefits of \LaTeX

- Widespread use by the R community
- Produces aesthetically beautiful documents
- You don't have to worry about layout, it's all automatic.
- Free
- Scientific features like the references are available in \LaTeX .

Disadvantage:

Fairly steep learning curve

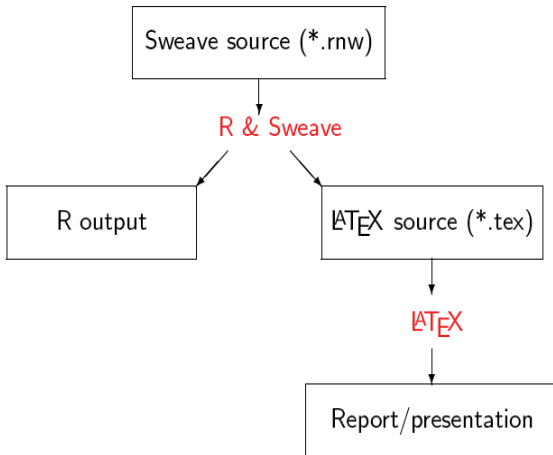
Intergrating R and \LaTeX with Sweave

Sweave

- An R package written by Friedrich Leisch
- Included with the base distribution of R
- Supports integration of R with \LaTeX

What does Sweave do?

Essentially requires a single source document *a.Rnwfile*.



What does it Work?

1. Write the LATEX file, but with extension .Rnw (or .Snw) instead of .tex:e.g **myfile.Rnw**.
2. The file will also contain R code segments, suitably separated from the LATEX segments.
3. Within R, execute Sweave("myfile.Rnw"), assuming myfile.Rnw is in the working directory of R.
4. This executes the code segments and will produce the file myfile.tex.
5. Run LATEX on ("myfile.tex") and obtain your report.

What you need

- A \LaTeX distribution (e.g., MiKTeX)
- An editor configured for use with R and \LaTeX (E.g., R-Studio, Tinn-R, LyX, Kile)
- Instructions that are comprehensive but limited in scope.

Resources

Introduction to LaTeX

A Sweave demo

Sweave and beyond

Beamer tutorial