



Creating R Reports with Sweave Pro's and Con's vs. LaTeX

Natalie Hawkins
SCHARP/VIDD/FHCRC

October 14, 2011

Outline

1. Sweave

- ▶ What is it?
- ▶ Why use it? / Why I use it.
- ▶ Examples of Sweave Final Products
- ▶ How to use it?
- ▶ Example .Rnw Files
- ▶ Sweave Editors
- ▶ Sweave Packages
- ▶ Other “Weaves”

2. LaTeX

- ▶ What is it?
- ▶ How to use it?

3. Sweave vs. LaTeX?

4. Help – Resources

What is Sweave? (practical)

Sweave is a function in the statistical programming language R that enables integration of R code into LaTeX or LyX documents.

LyX largely functions as a front-end to the LaTeX typesetting system; it has a GUI with menus.

A Sweave file:

1. has a .Rnw or .rnw extension
2. is a LaTeX file with R code embedded
3. can be compiled to produce a .pdf which contains the LaTeX text, the results of the R code (graphics or text), and optionally the R code itself.

What is Sweave? (theory)

Sweave is a *literate programming* language.

Literate programming is an approach to programming introduced by Donald Knuth as an alternative to the structured programming paradigm of the 1970s.

Donald Knuth, 73, Professor Stanford University, Math and Computer Science

- 1) *Introduced Literate Programming*
- 2) *Created TeX Typesetting System*

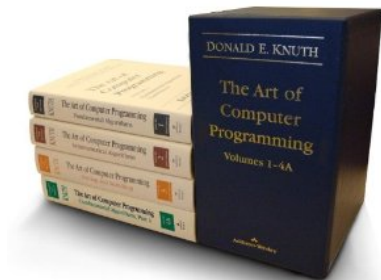


The *literate programming* paradigm moves away from writing programs in the manner and order imposed by the computer, and instead enables programmers to develop in the order of their thoughts.

Literate programs are written in an ordinary human language, much like the text of an essay, in which macros which hide abstractions and traditional source code are included. (source: Wikipedia)

The Knuth Books - The Art of Computer Programming

3)DK also wrote these books (\$198.46, Amazon.com)



“Computer science was a totally new field“, Knuth recalls, “with no real identity. And the standard of available publications was not that high. A lot of the papers coming out were quite simply wrong. So one of my motivations was to put straight a story that had been very badly told“.
(source: Wikipedia)

Why use Sweave?

The purpose is to create *dynamic reports*, which can be updated automatically if data or analysis change.

Another purpose is to enable *reproducible research*. The code and data are linked to the report (.pdf).

To create a *vignette for an R Pkg*; required by Bioconductor.

To create *slides using Beamer*.

Beamer is a LaTeX class for creating slides for presentations. It supports both pdfLaTeX and LaTeX + dvips. The name is taken from the German word Beamer, a pseudo-anglicism for video projector.

Why I use Sweave

1. My Rpkg, flowPlots, User Documentation
Sweave vignette is required by Bioconductor.
2. “Connecting to the Data Warehouse from R” doc
Show an example R session in document.
3. SBRI Project –
Avoid "reporting" in email.
where a file of plots is attached and "reporting" comments are made in the email text.
4. CHAVI Project –
Programmatically table survival analysis model results.
My first attempt in Powerpoint was not fast and not fun, so I didn't want to repeat
5. This Presentation – I saw Youyi's CAVD slides, and I thought
Beamer slides look nice.

Examples: Sweave Final Products

1. These Slides
2. Vignette for R pkg, flowPlots:
<http://www.bioconductor.org/packages/2.8/bioc/html/flowPlots.html>
3. Connecting to the Data Warehouse from R
<https://twiki.vidi.org/bin/view/PopSci/SCHARP/Pg2R>
4. SBRI Project: Analysis Report
(show Example)
5. CHAVI Project: Univariate and Multivariate Analysis Report
(show Example)

How to use Sweave? - 1) Write the .Rnw file

1. Start with a Sweave template, base.Rnw
2. Rename it for your report, say, report.Rnw
3. Add the LaTeX sections and text that you need
4. Add the R code
 - ▶ R code starts after: `<<>=`, and ends just before: `@`
options can be specified inside `<<>`
 - ▶ options:
 - ▶ 1) `echo = TRUE|FALSE`, show/hide the code itself
 - ▶ 2) `results = verbatim|tex|hide`, write the results as verbatim or as tex, or hide them
 - ▶ 3) `fig = TRUE|FALSE`, specify if a figure will be generated
 - ▶ `source("R code")` or write the code directly

How to Use Sweave? - 2) Compile the .Rnw file

Use the Sweave function in the R utils pkg

```
R> help(Sweave)
```

Sweave{utils} R Doc Automatic Generation of Reports

Description

Sweave provides a flexible framework for mixing text and R/S code for automatic report generation. The basic idea is to replace the code with its output, such that the final document only contains the text and the output of the statistical analysis: however, the source code can also be included.

Two functions:

1. Sweave() compiles the R code and produces a .tex document
2. Stangle() uses driver Rtangle() which extracts R code chunks

How to use Sweave? - 2) Compile the .Rnw file

1. *From a Terminal Window*

- ▶ Compile the Sweave File: (this creates report.tex)

```
msDos>R CMD Sweave report.Rnw
```

- ▶ Generate a .pdf file from the .tex file: (this creates report.pdf)

```
msDos>pdflatex report.tex
```

*Note: You ***might*** need to put Sweave.sty and Upquote.sty in your working directory.*

2. *From within R*

- ▶ Compile the Sweave File: (this creates report.tex)

```
R>Sweave("report.Rnw")
```

- ▶ Generate a .pdf file from the .tex file: (this creates report.pdf)

```
R>library(tools)
```

```
R>texi2dvi("report.tex", pdf=TRUE)
```

An Example .Rnw File (also Show Examples)

File Structure:

Some LaTeX (eg. begin document, title, author)

```
<< setX, echo = FALSE, results = hide, fig = FALSE >>=  
## This is a chunk of R code  
x = 1:10  
print(mean(x))
```

@

More LaTeX (eg, close with: end document)

Examples:

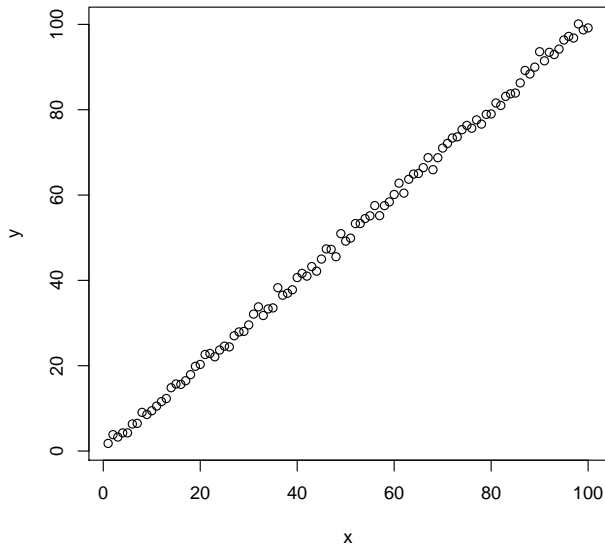
Example .Rnw files can be found on the SCHARP Twiki Stat Presentations page:

https:

[//twiki.vidi.org/bin/view/PopSci/SCHARP/StatPresentations](https://twiki.vidi.org/bin/view/PopSci/SCHARP/StatPresentations)

Plot generated from embedded R Code

This plot was generated when the Sweave slides were compiled



Sweave Editors

1. WordPad, Notepad
2. R Gui
3. RStudio, Vim, Emacs + ESS
4. Eclipse StatET plugin

Eclipse is a multi-language software development environment comprising an integrated development environment (IDE) and an extensible plug-in system. Eclipse is written mostly in Java.

Note: I use WordPad for Sweave. RStudio sounds interesting. I have used Eclipse for Java programming. I use R Gui on Windows for my R programming.

Sweave Packages

1. *weaver*, *pgfSweave*, *cacheSweave* – useful when you have time-consuming code chunks; it allows them to be cached if they don't change, so the report can be re-created faster.
2. *ascii* – Export R objects to several markup languages
3. *R2HTML* – can output HTML, rather than pdf
4. *exam* – automatic generation of exams
5. *svSweave* – clean-up R noweb files produced by LyX
6. *SweaveListingUtils* – avoids inconsistencies between vignette and documented source code

Other “Weaves”

1. *odfWeave()* – weaves “Word” and R; the input and output = a MicrosoftWord-like document.

Details:

- ▶ The markup language used is the Open Document Format (ODF), which is an open, non-proprietary format that encompasses text documents, presentations and spreadsheets.
- ▶ The file extension for the ODF file is .odt

Advantages:

- ▶ sharing, Word is more common than LaTeX
- ▶ the end-user can easily edit the document

Disadvantages:

- ▶ hasn't yet reached Version 1.0; it's at the 5th release, v 0.7.17.
- ▶ may have issues with PDFs, may require fiddling with PNGs, based on review of v 0.7.10
- ▶ may have issues with presentations and spreadsheets.

2. *PWeave* – weaves LaTeX and Python

Break: Any Sweave Questions?

Ok, moving on to "Creating Reports with LaTeX".

This approach was taken for the RV144 Analysis.

What is LaTeX?

LaTeX is a document markup language and document preparation system for the TeX typesetting program.

LaTeX is not a stand-alone typesetting program in itself, but document preparation software that runs on top of Donald E. Knuth's TeX typesetting system.

How to use LaTeX? - 1) Write the .tex file

1. Start with a LaTeX template, `base.tex`
2. Rename it for your report, say, `report.tex`
3. Add the LaTeX sections and text that you need
4. Add LaTeX code to input figures and tables
(possibly R output, maybe LaTeX-wrapped R output)

How to use LaTeX? - 2) Compile the .tex file

From a Terminal Window: (this creates report.pdf)

```
msDos> pdflatex report.tex
```

Sweave or LaTeX Reports: *Which Approach Should I Take?*

Note –

1. Sweave is “LaTeX with embedded R code”,
so Sweave “=” LaTeX,
if you don't embed R code to be dynamically run.
2. For example,
I have embedded R code to create Slide 14 in these slides;
without that slide (and code), these slides are “Sweave”; they are
also “LaTeX”.

Let's list the concerns we might have.

Sweave vs. LaTeX, version 1.0

Concern	Which One?
Want “Reproducible Research”	Sweave
Don’t Want to “Install” or “Use” Sweave	LaTeX
Want a Dynamic Report	Either
Have Multiple Contributors	Either
Have Time-Consuming Code Chunks	Either
Want To Use R Results in Multiple Reports	Either
Don’t Want To Include All Analysis in the Report	Either
Want OpenOffice Format as Output	“Sweave helps”

Note: I call this “Version 1.0” because we could talk a lot of details. Maybe we add to the list. Maybe we further explain “Which One?”.

Help - Resources

1. *Google* xxx, where xxx = what you need
2. *Wikibooks*
 - ▶ Sweave - R Programming/Publication Quality Output:
http://en.wikibooks.org/wiki/R_Programming/Publication_Quality_Ouput
 - ▶ LaTeX -
<http://en.wikibooks.org/wiki/LaTeX>
3. *Sweave Cran Task View* – Reproducible Research
<http://cran.r-project.org/web/views/ReproducibleResearch.html>
4. Quick and Dirty Guide to LaTeX:
<http://happymutant.com/latex/index.html>
5. *Beamer*:
[http://en.wikipedia.org/wiki/Beamer_\(LaTeX\)](http://en.wikipedia.org/wiki/Beamer_(LaTeX))
6. *SCHARP Twiki*: these slides plus example .Rnw and .tex files
<https://twiki.vidi.org/bin/view/PopSci/SCHARP/StatPresentations>

The End

