

Literate programming with R and L^AT_EX

A brief overview

Glen Sargeant
U.S. Geological Survey

Sponsored by the
TWS Biometrics Working Group

October 2, 2010

Objectives

- ▶ Promote appreciation for literate programming with R/L^AT_EX
- ▶ Provide a *minimally sufficient* overview of major concepts
- ▶ Provide examples that can be modified to meet many needs

Motivation

*Let us change our traditional attitude to the construction of programs: Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on **explaining to human beings** what we want a computer to do.*

–Donald Knuth

Philosophy

A traditional computer program consists of a text file containing program code. Scattered in amongst the program code are comments which describe the various parts of the code.

*In literate programming the emphasis is reversed. Instead of writing code containing documentation, **the literate programmer writes documentation containing code.***

–Ross Williams

Benefits of literate programming

Integration of analysis and reporting

- ▶ Consistent formatting
- ▶ **No** transcription errors
- ▶ Automatic updating of documents
- ▶ Reproducible results

Benefits of L^AT_EX

- ▶ Widespread use by the R community
- ▶ Relative ease-of-use
- ▶ Beautifully typeset products
- ▶ Flexibility and extensibility
- ▶ Platform independence
- ▶ Transparency (text files)

Drawbacks of L^AT_EX

Programming with L^AT_EX is simultaneously easy and very, very difficult... like herding cats:

Drawbacks of L^AT_EX

Programming with L^AT_EX is simultaneously easy and very, very difficult... like herding cats:

- ▶ Making cats go where *they* wish is easy.

Drawbacks of L^AT_EX

Programming with L^AT_EX is simultaneously easy and very, very difficult... like herding cats:

- ▶ Making cats go where *they* wish is easy.
- ▶ Making cats go where *you* wish... not so much!

Other challenges

- ▶ Maintaining software
- ▶ Learning another programming language

Outline

Three brief examples:

- ▶ Writing a report
- ▶ Integrating with R
- ▶ Preparing a slide show

\LaTeX as a whole is overwhelming: the parts you *need* to master are not!

Writing a report

Parts of a \LaTeX program

A \LaTeX **source file** is a text file that produces printed text when it is processed. Such files consist of 2 parts: a **preamble** and a **body**.

Preamble

```
\documentclass{article}  
Global commands  
...
```

Body

```
\begin{document}  
Document text  
Local commands  
Graphics  
...  
\end{document}
```

Strategic tip!

To achieve a consistent "look" that can be modified easily, rely as much as possible on global controls exerted through the preamble and as little as possible on local formatting commands issued in the body.

What's in a preamble?

L^AT_EX features a package system similar to the one used by R. Packages extend capabilities of L^AT_EX and facilitate such key tasks as inclusion of graphics and the use of color. Packages used in the body must be loaded in the preamble:

```
\documentclass{article}  
\usepackage{graphicx}  
...
```

In addition, most aspects of output can be controlled in the preamble by setting layout and style **parameters** (for present purposes, we will be content with defaults).

Components of a report?

- ▶ Title page
- ▶ Abstract
- ▶ Sections
- ▶ Illustrations
- ▶ Cross-references
 - ▶ Lit cited
 - ▶ Index
 - ▶ List of figures
 - ▶ List of tables

\LaTeX will automatically format all of these components to specifications in the preamble, leaving you free* to concentrate on content!

A sample title page

```
\documentclass{article}
\begin{document}
\title{A sample title page}
\author{Glen A. Sargeant}
...
\maketitle
...
\end{document}
```



A complete report

Commands used to generate this report are shown in typewriter font. Dots (...) are placeholders for text.

A L^AT_EX article

Glen Sargant

July 19, 2010

Abstract

This skeleton of a document was produced to demonstrate document generation with L^AT_EX. Subsequent examples will demonstrate preparation of presentations and integration with R statistical software.

```
\begin{document}
```

```
\title{A \LaTeX \space{pt} article}
\author{Glen Sargant}
\maketitle
\begin{abstract}
...
\end{abstract}
```

Contents

Introduction	2
Methods	2
Data collection	2
Analyses	2
Results	2
Discussion	2
References	2

```
\tableofcontents
```

Introduction

```
\section{Introduction}
\addcontentline{toc}{section}{Introduction}
...
```

Methods

```
\section{Methods}
\addcontentline{toc}{section}{Methods}
...
```

Data collection

```
\subsection{Data collection}
\addcontentline{toc}{subsection}{Data collection}
...
```

Analyses

```
\subsection{Analyses}
\addcontentline{toc}{subsection}{Analyses}
...
```

Results

```
\section{Results}
\addcontentline{toc}{section}{Results}
...
```

Discussion

```
\section{Discussion}
\addcontentline{toc}{section}{Results}
...
```

References

[1] Leslie Lamport, *L^AT_EX: A Document Preparation System*. Addison Wesley, Massachusetts, 2nd Edition, 1994.

```
\addcontentline{toc}{section}{References}
\begin{thebibliography}{0}

\bibitem{Lamport94}
Leslie Lamport,
\emph{\LaTeX: A Document Preparation System}.
Addison Wesley, Massachusetts,
2nd Edition,
1994.

\end{thebibliography}

\end{document}
```

Paragraphs of text

Paragraphs of text can simply be typed. Line breaks and vertical alignment are handled automatically. Paragraphs may be separated by blank lines, but using hard returns (`\`) seems to improve vertical spacing.

What if your needs are more elaborate?

Commands and environments

Layout and style can be modified temporarily in the body. Most local formatting is accomplished with a **command** or an **environment**.

Commands apply to arguments or to subsequent text.

Environments treat enclosed text differently than surrounding text.

Sample commands

```
\fbox{Jack} (Jack-in-a-box)\|[1cm]
```

Jack (Jack-in-a-box)

```
\textbf{\colorbox{orange}{\textcolor{blue}  
{Bold-faced blue text on an orange background}}}\|[1cm]
```

Bold-faced blue text on an orange background

```
\emph{\fcolorbox{red}{yellow}  
{Italic text in a yellow box with a red border}}\|
```

Italic text in a yellow box with a red border

Command syntax

`\name[optional]{required}`

Some commands do not require any arguments and others require multiple arguments, each of which must be enclosed in curly braces. Some commands accept optional arguments, which must be enclosed in square braces.

A sample environment

```
\begin{itemize}
\item First item
\item Second item
\item Third item
\end{itemize}
```

- ▶ First item
- ▶ Second item
- ▶ Third item

Environments always are initiated with the command `\begin{name}` and terminated by `\end{name}`.

Do you really need a typesetting program?

Mastering \LaTeX requires substantial effort. Why bother?

Do you really need a typesetting program?

Mastering \LaTeX requires substantial effort. Why bother?

Sweave!

Sweave

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

Integrating R with L^AT_EX

Sweave source files

Once you have "mastered" \LaTeX , integrating R requires minimal additional effort.

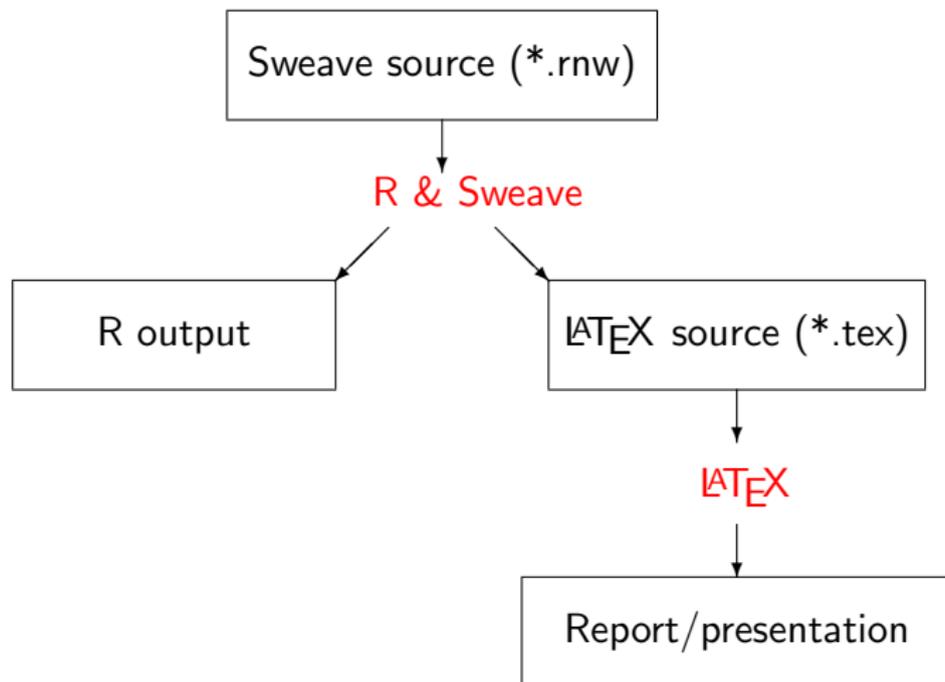
Sweave source files are just like \LaTeX source files *except* they may include executable R code and associated formatting commands.

```
\documentclass{article}  
\usepackage{C:/R/R-2.11.0/share/texmf/Sweave}  
\begin{document}
```

R code is identified thusly:

```
<<Sweave options>>=  
R commands  
@
```

What does Sweave do?



What does Sweave do?

Code is executed; code and results are incorporated in documentation.

Sweave input

```
<<fig=T,include=T>>=  
x <- rnorm(10)  
y <- 10*x + rnorm(10)  
plot(x,y)
```

Sweave output

```
\begin{Schunk}  
\begin{Sinput}  
> x <- rnorm(10)  
> y <- 10 * x + rnorm(10)  
> plot(x, y)  
\end{Sinput}  
\end{Schunk}  
\includegraphics{LaTeX-001}
```

Preparing a presentation

Presentations with L^AT_EX

The L^AT_EX "beamer" class formats documents for on-screen display with a... well, a "beamer."

```
\documentclass{beamer}
```

Terminology and syntax

What we would call a "slide" is known as a **frame**. Creating a frame is very easy:

```
\begin{document}  
\begin{frame}  
\frametitle{Title text}  
Content  
...  
...  
\end{frame}  
\end{document}
```

Dressing up reports and presentations

With practice, only creativity limits effects that can be achieved: good judgement should impose additional limits!

- ▶ Mathematics
- ▶ R code
- ▶ Illustrations
- ▶ Tables
- ▶ Layout
- ▶ Color
- ▶ Fonts
- ▶ Backgrounds

Sample results follow. See the presentation source code or references listed for additional information.

Mathematics

L^AT_EX excels for typesetting mathematics and PDF is a device-independent format. Font substitution and goofy spacing are never problems!

Mathematics can be displayed in line with text (e.g., x_i^j) or between lines of text, e.g.:

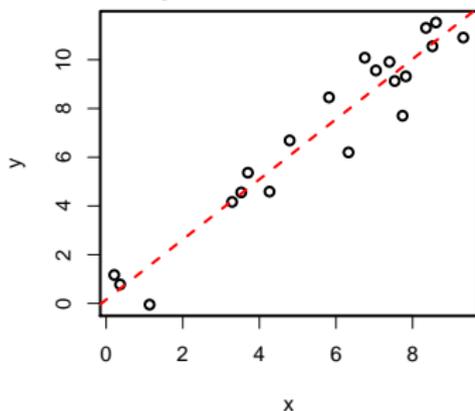
$$\mathbf{y} = \mathbf{x}\beta + \epsilon$$

See the [Short math guide for L^AT_EX](#), published by the American Mathematical Society.

R code and output

Sweave works with the Beamer document class, too!

```
> par(cex = 1.4, lwd = 3)
> x <- runif(20, 0, 10)
> y <- 1.2 * x + rnorm(20, 0, 1)
> plot(x, y)
> abline(lm(y ~ x), lty = 2, col = "red")
```



R code and output

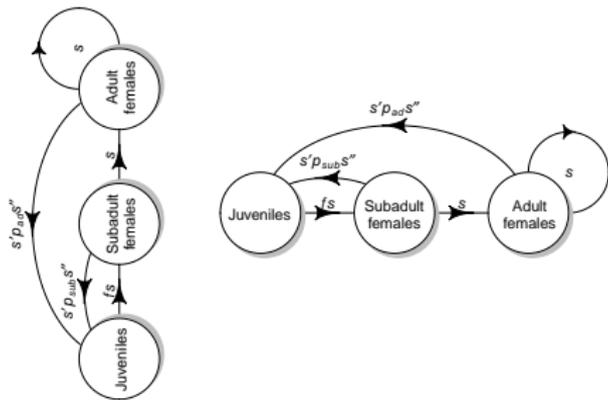
- ▶ Code can be hidden or displayed, as desired.
- ▶ No more shuttling between software packages for statistics, graphics, and presentations!
- ▶ No tedious updating of results and graphics!
- ▶ No transcription errors!

Visit the [Sweave Homepage](#) and read the (short) user manual for additional information.

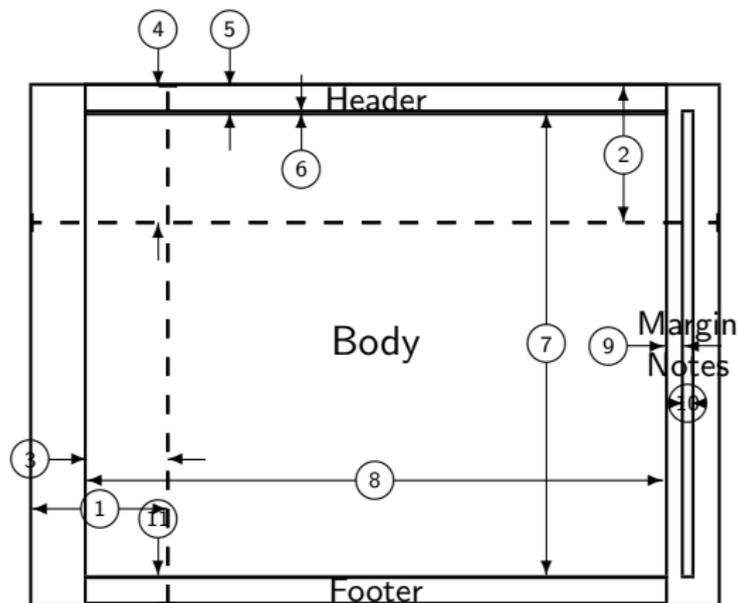


Illustrations

The `graphicx` package supports insertion of graphics. **Lengths** and optional arguments can be used to fine-tune automatic placement of text. The `textpos` package can be used to place illustrations exactly where you want them on slides.



Modifying lengths (`\usepackage{layout}`)



- | | | | |
|----|-------------------------------------|----|---|
| 1 | <code>one inch + \hoffset</code> | 2 | <code>one inch + \voffset</code> |
| 3 | <code>\oddsidemargin = -43pt</code> | 4 | <code>\topmargin = -72pt</code> |
| 5 | <code>\headheight = 14pt</code> | 6 | <code>\headsep = 0pt</code> |
| 7 | <code>\textheight = 244pt</code> | 8 | <code>\textwidth = 307pt</code> |
| 9 | <code>\marginparsep = 10pt</code> | 10 | <code>\marginparwidth = 4pt</code> |
| 11 | <code>\footskip = 14pt</code> | | <code>\marginparpush = 5pt</code> (not shown) |

Tables

Simple tables are easy to prepare but fine control of details can be challenging. See <http://en.wikibooks.org/wiki/LaTeX/Tables>. Use tables (*not columns*) if you want items in columns to be vertically aligned.

Column 1	Column 2	Column 3
Item 1-1	Item 1-2	Item 1-3
Item 2-1	Item 2-2	Item 2-3 (Example of text wrapping in a table)
⋮	⋮	⋮
Item 2-1	Item 2-2	Item 2-3

Slide backgrounds and illustrations

- ▶ Themes
- ▶ Colorthemes
- ▶ Font themes
- ▶ Graphics

Presentations with Beamer

Use "themes" to format presentations

Glen Sargeant

July 19, 2010

Beamer **themes** make it easy to achieve a consistent, visually pleasing style.

Specify a theme in the preamble

```
\usetheme{copenhagen}
```

Color themes

Beamer **color themes** make it easy to use *complementary* colors in a consistent way.

Specify a color theme in the preamble

```
\usetheme{copenhagen}  
\usecolortheme{crane}
```

Font themes

Beamer **font themes** make it easy to use *complementary* fonts in a consistent way.

Specify a font theme in the preamble

```
\usetheme{copenhagen}  
\usecolortheme{crane}  
\usefonttheme{serif}
```

Photographic backgrounds

```
{  
  \usebackgroundtemplate{[width = \paperwidth]{elkphoto.jpg}}  
  \begin{frame}  
  ...  
  \end{frame}  
}
```



Questions and closing remarks

?

Getting started

What do you need?

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

Good sources of instruction

L^AT_EX

- ▶ A wikibook; distinguished by excellent coverage of essentials
- ▶ A discussion forum; useful for resolving specific problems

Beamer

- ▶ An excellent Beamer tutorial
- ▶ A short but useful article

See also the Beamer documentation and example source files provided with your workshop materials!

The use of trade, product, or firm names does not imply endorsement by the U.S. Government.