Principles of Reproducible Research with R

Author: Andrea Foulkes, Gregory Matthews, Nicholas Reich

Biostatistics in Practice: Research Training in High-Performance Computing with R

This material is part of the **statsTeachR** *project*

Introduction and welcome...

Dyanamic R Reports

- Dynamic R documents allow a user to combine text, R code, and R output, including tables and figures, into one document.
- Why is this useful?
  - Writing code and producing reports are now one document rather than many.
  - When the analysis changes, the results in the report change automatically.
  - What else?
- There are several options for how to do this.
  - R Markdown
  - Sweave
  - knitr

Dyanamic R Reports: Summary

- R Markdown: creates HTML document
- Sweave: creates pdf document AND incorporates LaTeX
- knitr: $\approx$ Sweave + cacheSweave + pgfSweave + weaver + animation::saveLatex + R2HTML::RweaveHTML + highlight::HighlightWeaveLatex + 0.2 * brew + 0.1 * SweaveListingUtils + more

R Markdown

- R Markdown creates HTML files
- Reference: http://www.rstudio.com/ide/docs/authoring/using_markdown
- Markdown files (.Rmd) act just like text files, except they allow a user to embed R code in chunks
- The syntax for a chunk in R Markdown:

Regular text
```{r}
Code goes here
```

Sweave

- Sweave creates pdf files as output
- Reference: http://leisch.userweb.mwn.de/Sweave/
- Sweave Manual: http://www.stat.uni-muenchen.de/ leisch/Sweave/Sweave-manual.pdf
- Sweave not only integrates R code, but also LaTeX!
- The syntax for a chunk in Sweave:

Regular text with $LaTeX$ if you want it.
<<OPTIONS >>=
Code goes here
@

Sweave: Options

-
- Reference: http://leisch.userweb.mwn.de/Sweave/
- Sweave not only integrates R code, but also LaTeX!
- The syntax for a chunk in Sweave:

knitr

- Created by Ph.D. student Yihui Xie.
- Xie describes knitr $\approx$ Sweave + cacheSweave + pgfSweave + weaver + animation::saveLatex + R2HTML::RweaveHTML + highlight::HighlightWeaveLatex + 0.2 * brew + 0.1 * SweaveListingUtils + more
- knitr creates pdf files as output.
- It also allows the use of LaTeX, whereas R Markown does not.

Version control and reproducibility

GitHub increases productivity and transparency