



Using the rmd2pdf-guide: All the News that is Fit to Print

Paul Johnson, Director,
CRMDA <pauljohn@ku.edu>

Brent Kaplan, GRA,
CRMDA <bkaplan4@ku.edu>



Feb 22, 2018

See <http://crmda.ku.edu/guides> for updates.

1 Nuts and Bolts

1.1 Minor header edits only

The header has a place to insert a title, subtitle, email addresses. Do not change the `pdf_document` settings unless you want to get serious.

1.2 To compile this document

To compile, either start R and use the function `rmd2pdf` or use the shell script `rmd2pdf.sh` that is in the same directory as the `Rmd` file. The function `rmd2pdf` (same as the script) supplies settings. The user can supply many arguments to control the table of contents and so forth.

Compiling the document will also produce an R file that is a “tangled” version of the code chunks in the document.

1.3 Document customization: essentials

The document header depends on a \LaTeX template called “guide-boilerplate.tex”. That will be copied to the theme directory the first time a document is compiled in this directory. If you do want to edit the theme boilerplate, edit `theme/guide-boilerplate.tex`. Make sure the compiler uses that document. If you use our compiler script or the R function `rmd2pdf`, it is necessary to supply the template argument with the function call.

That is, it is not sufficient to insert a template name in the YAML header, and it is also not sufficient to use Rstudio’s “knit” menu. One must specify the template argument with `rmd2pdf`.

2 Much \LaTeX Syntax is allowed

This is explained in the `stationery` package vignette `Rmarkdown`.

Use `\[` and `\]` for display equations:

$$\Sigma_{gt} = \Lambda_{gt} \Psi_{gt} \Lambda'_{gt} + \Theta_{gt}$$

Because the backend is PDF and LaTeX will compile the document, numbered equations and cross references will succeed. Use `\begin{equation}` and `\end{equation}`:

$$\Sigma_{gt} = \Lambda_{gt} \Psi_{gt} \Lambda'_{gt} + \Theta_{gt} \tag{1}$$

That is equation (1). If we put in another equation, you will see it is automatically numbered:

$$y_i = \beta_0 + \beta_1 X_i \tag{2}$$

That's equation (2).

2.1 Additional packages

In addition, if you insert \LaTeX features that require packages that are not currently in the report-boilerplate.tex, then those packages can be inserted into the preamble by YAML header markup like so:

```
1 header-includes:
2 - \usepackage{xcolor}
3 - \usepackage{amsmath}
4 - \usepackage{amssymb}
5 - \usepackage{fancybox}
```

That example has the package `amssymb` because some non-standard math symbols were needed.

3 R code chunks

In guide documents, we do generally include input and output from code chunks. We have not yet mastered steps necessary to customize the output of code chunks when using Rmarkdown, so the output received here is beyond our control at the moment. This is a principal reason why working with Sweave is more desirable for our purposes.

First, suppose we create a data frame

```
1 set.seed(1234)
2 x <- rnorm(100, 0, 1)
3 y <- rnorm(100, 0, 1)
4 dat <- data.frame(x=x, y=y)
```

Of course, because this is guide, we are also likely to have some ugly raw R output.

```
1 summary(dat)
```

```

1      x              y
2  Min.    :-2.3457   Min.    :-2.85576
3  1st Qu.: -0.8953   1st Qu.: -0.55927
4  Median  :-0.3846   Median   : 0.03280
5  Mean    :-0.1568   Mean     : 0.04124
6  3rd Qu.: 0.4712   3rd Qu.: 0.62764
7  Max.    : 2.5490   Max.     : 3.04377

```

Readers should notice the document template numbers the lines in both input and output chunks. This can be changed by altering the template.

We can show ugly output if we really need to. But we'd rather include lovely graphs or nicely formatted \LaTeX tables.

When tables are concerned, the placement of tables is the major issue. Sometimes we like tables that are “numbered floating” objects. However, because this is a guide document, we might want the \LaTeX table output to show “right here” (prevent it from floating into a numbered table). The following code shows that the `print.xtable` method allows us to control that.

The following creates output that does not go into a floating object. Note the output will blurt into the document, immediately after the code chunk itself:

```

1 library(xtable)
2 xt <- xtable(head(dat, 10), caption="Ten Lines from One Data Frame",
3              label="tab:ex2")
4 print(xt, comment=FALSE, include.rownames=FALSE, floating=FALSE)

```

x	y
-1.21	0.41
0.28	-0.47
1.08	0.07
-2.35	-0.50
0.43	-0.83
0.51	0.17
-0.57	-0.90
-0.55	0.17
-0.56	0.35
-0.89	-0.05

Note that because the `print.xtable` arguments include “`floating=FALSE`”, the table caption is not printed with the output table. The concept of caption is meaningful only for floating tables, as we illustrate next.

A more professional result is obtained if we allow `xtable` to create a “floating” table object. The following code creates Table 1.

```

1 print(xt, comment=FALSE, include.rownames=FALSE, floating=TRUE,
2       caption.placement = "top")

```

For whatever reason, the custom is that captions (titles) of tables are placed **on top** while titles for figures are placed on the bottom.

Table 1: Ten Lines from One Data Frame

x	y
-1.21	0.41
0.28	-0.47
1.08	0.07
-2.35	-0.50
0.43	-0.83
0.51	0.17
-0.57	-0.90
-0.55	0.17
-0.56	0.35
-0.89	-0.05

As explained in the `stationery` vignette `code_chunks`, many options are available for code chunks.

Bibliographical citation test is appropriate.

This document is produced with R (R Core Team 2017). Here's a citation that excludes some author names (2013, 37). Note that to get the full parenthesized statement with names and dates, we insert hard brackets `[`, an `@` sign, the bibtex tag, and a closer `]`. If we don't want their names, we insert a `-` sign. It is also possible to refer to a group of projects (Hsiao 2014; Fitzmaurice, Laird, and Ware 2011; McCullagh 1983).

The style of the bibliography will not match what we generally want. It is close to what we want. The default is based on the Chicago author-year format.

4 Session Info

This is a required element in a guide

```
1 sessionInfo()
```

```
1 R version 3.4.3 (2017-11-30)
2 Platform: x86_64-pc-linux-gnu (64-bit)
3 Running under: Ubuntu 17.10
4
5 Matrix products: default
6 BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.7.1
7 LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.7.1
8
9 locale:
10  [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
11  [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
12  [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
13  [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
```

```

14 [9] LC_ADDRESS=C LC_TELEPHONE=C
15 [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
16
17 attached base packages:
18 [1] stats graphics grDevices utils datasets base
19
20 other attached packages:
21 [1] xtable_1.8-2 stationery_0.61
22
23 loaded via a namespace (and not attached):
24 [1] Rcpp_0.12.15 quadprog_1.5-5 rprojroot_1.3-2
25 [4] digest_0.6.15 plyr_1.8.4 backports_1.1.2
26 [7] magrittr_1.5 stats4_3.4.3 evaluate_0.10.1
27 [10] stringi_1.1.6 pbivnorm_0.6.0 openxlsx_4.0.17
28 [13] rmarkdown_1.8 tools_3.4.3 stringr_1.2.0
29 [16] foreign_0.8-69 kutils_1.34 yaml_2.1.16
30 [19] compiler_3.4.3 mnormt_1.5-5 htmltools_0.3.6
31 [22] knitr_1.19 lavaan_0.5-23.1097 methods_3.4.3

```

```

1 if (!is.null(warnings())){
2   print("Warnings:")
3   warnings()
4 }

```

References

- Diggle, Peter, Patrick Heagerty, Kung-Yee Liang, and Scott Zeger. 2013. *Analysis of Longitudinal Data*. 2 edition. Oxford: Oxford University Press.
- Fitzmaurice, Garrett M., Nan M. Laird, and James H. Ware. 2011. *Applied Longitudinal Analysis*. 2nd ed. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley.
- Hsiao, Cheng. 2014. *Analysis of Panel Data*. Third edition. Econometric Society Monographs 54. New York, NY: Cambridge University Press.
- McCullagh, P. 1983. *Generalized Linear Models*. Monographs on Statistics and Applied Probability 37. London: Chapman; Hall.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.