

End-term project: Advanced Statistical Computing 2020

Yizhen Dai s2395479

This document provides an introduction to R Markdown, argues for its benefits, and presents a sample manuscript template intended for an academic audience. I include basic syntax to R Markdown and a minimal working example of how the analysis itself can be conducted within R with the `knitr` package.

Keywords: pandoc, r markdown, knitr

Introduction

This project aims at solving a modeling problem faced by an insurance company - ANV. Two of ANV business lines, [Professional liability insurance](#) (PLI) and [Workers' compensation](#) (WC), were affected by a huge claim from one client during the last year. Therefore, ANV comes to a reinsurance company for an insurance policy. For some threshold $t = 100, 110, \dots, 200$: - If $PLI + WC \leq t$, ANV pays the claim themselves. - If not, the reinsurance company pays the claim.

The price that reinsurance company asks depends on the threshold t :

$$P(t) = 40000 * e^{-t/7}$$

Of course, the policy is only reasonable if the expected over-threshold claim exceeds the price: $V(t) = E[(PLI + WC)1(PLI + WC > t)] > P(t)$. We will use statistical modeling to approximate $V(t)$

Methodolgy

Simulation study

Results

The lion's share of a R Markdown document will be raw text, though the front matter may be the most important part of the document. R Markdown uses [YAML](#) for its metadata and the fields differ from [what an author would use for a Beamer presentation](#). I provide a sample YAML metadata largely taken from this exact document and explain it below.

Markdown Syntax

```
# Introduction

**Lorem ipsum** dolor *sit amet*.

- Single asterisks italicize text *like this*.
- Double asterisks embolden text **like this**.
```

Start a new paragraph with a blank line separating paragraphs.

- This will start an unordered list environment, and this will be the first item.
- This will be a second item.
- A third item.
 - Four spaces and a dash create a sublist and this item in it.
- The fourth item.

1. This starts a numerical list.
2. This is no. 2 in the numerical list.

```
# This Starts A New Section
## This is a Subsection
### This is a Subsubsection
#### This starts a Paragraph Block.
```

> This will create a block quote, if you want one.

Want a table? This will create one.

Table Header	Second Header
Table Cell	Cell 2
Cell 3	Cell 4

Note that the separators **do not** have to be aligned.

Want an image? This will do it.

![caption for my image](path/to/image.jpg)

‘fig_caption: yes’ will provide a caption. Put that in the YAML metadata.

Almost forgot about creating a footnote.^[^1] This will do it again.^[^2]

^[^1]: The first footnote
^[^2]: The second footnote

Want to cite something?

- Find your biblatexkey in your bib file.
- Put an @ before it, like @smith1984, or whatever it is.
- @smith1984 creates an in-text citation (e.g. Smith (1984) says...)
- [@smith1984] creates a parenthetical citation (Smith, 1984)

That’ll also automatically create a reference list at the end of the document.

[In-text link to Google](http://google.com) as well.

That’s honestly it. Markdown takes the chore of markup from your manuscript (hence: “Mark-down”).

On that note, you could easily pass most LaTeX code through Markdown if you’re writing a LaTeX document. However, you don’t need to do this (unless you’re using the math environment) and probably shouldn’t anyway if you intend to share your document in HTML as well.

Using R Markdown with Knitr

`eval=FALSE` option just displays the R code (and does not run it), `tidy=TRUE` wraps long code so it does not run off the page. `include=FALSE` hide code and output from document. `echo=FALSE` hides only the code, but not the output.

```
library(stevemisc)
data(uniondensity)

M1 <- lm(union ~ left + size + concen, data = uniondensity)

library(arm)
coefplot(M1)
```

The implications for workflow are fairly substantial. Authors can rather quickly display the code they used to run the analyses in the document itself (likely in the appendix). As such, there’s little guesswork for reviewers and editors in understanding what the author did in the analyses reported in the manuscript.

Figure

`eval=FALSE` changes to `eval=TRUE`

```
coefplot(M1)
```

Table

add `results="asis"` to the brackets to start the R code chunk. The ensuing output will look like this (though the table may come on the next page).

```
stargazer(M1, title = "A Handsome Table", header = FALSE)
```

Footnote

Adding `echo="FALSE"` inside the brackets to start the R chunk will omit the presentation of the R commands. It will just present the table. This provides substantial opportunity for authors in doing their analyses. Now, the analysis and presentation in the form of a polished manuscript can be effectively simultaneous.¹

¹I’m not sure if I’m ready to commit to this myself since my workflow is still largely derived from [Rob J. Hyndman’s example](#). However, *knitr* has endless potential, especially when analyses can be stored in cache, saved as chunks, or loaded in the preamble of a document to reference later in the manuscript.

For my template, I'm pretty sure this is mandatory.²

How to cite

Perhaps the greatest intrigue of R Markdown comes with the [knitr package](#) provided by [Xie \(2013\)](#).

²The main reason I still use `pdflatex` (and most readers probably do as well) is because of LaTeX fonts. [Unlike others](#), I find standard LaTeX fonts to be appealing.

References

Xie, Yihui. 2013. *Dynamic Documents with R and knitr*. Boca Raton, FL: CRC Press.