

Cómo usar Rmarkdown

Julio César Iturra Sanhueza - jciturra@uc.cl

23 octubre, 2018

Resumen

El presente documento tiene por objetivo facilitarles la vida a través de la elaboración de documentos a través de Rmarkdown. La gracia de los documentos con extensión `.rmd` permite facilitar el uso de códigos para el cálculo matemático, estimación de modelos, y, sobre todo, el reporte de sus resultados en tablas de calidad. Dentro de las posibilidades, Rmarkdown permite elaborar documentos que combinan el código nativo de R, como también es posible emplear código `.tex`, o lo que coloquialmente se le llama “formato latex”.

YAML header

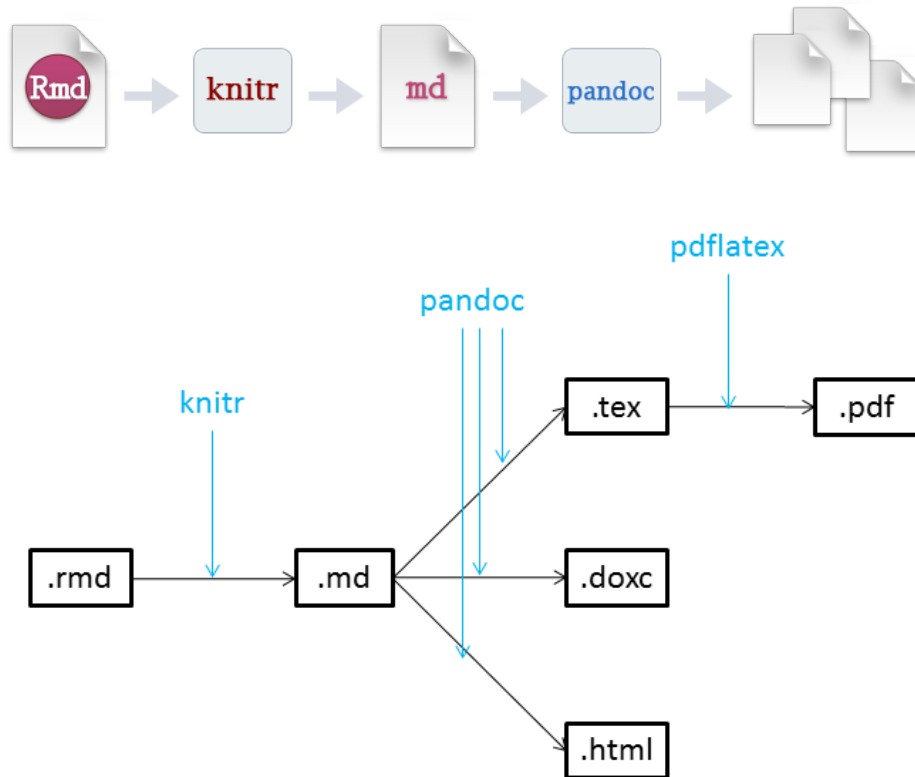
El YAML header corresponde al lugar donde ustedes pueden determinar las opciones de su documento. Generalmente comenzamos poniendo el nombre, autoría y fecha de su documento. Agregué otras opciones que son útiles para dar formato a su documento, tales como el tamaño de la fuentes (fontsize), el interlineado (linestretch), márgenes (geometry) y abstract.

```
---
title: "Cómo usar Rmarkdown"
author: "Julio Iturra - jciturra@uc.cl"
date: '`r format(Sys.time(), "%d %B, %Y")`'
fontsize: 11pt
linestretch: "1.0"
geometry: margin=0.78in
output: pdf_document
header-includes:
  - \usepackage[spanish,es-tabla]{babel} #Para que aparezan "Tabla" y "Figura" en español
  - \usepackage[utf8]{inputenc}
  - \usepackage{booktabs}
abstract: "El presente documento tiene por objetivo facilitarles la vida ..."
---
```

La opción `output` tiene tres opciones para convertir su documento en `.pdf`, `.html` y `.doc`. A mi me gusta realizarlo en pdf para los informes y html para mis reportes de datos y exploración, dado que los trabajos en html pueden ser abiertos en cualquier sistema operativo sin la necesidad de un software especializado como Adobe o Nitro reader (por dar algunos ejemplos).

Introducción

Este documento está enfocado al reporte en formato .pdf, dado que me parece más adecuado para el reporte de informes y trabajos académicos. A modo general, la figura a continuación muestra cómo funciona Rmarkdown cuando estamos empleando formato .pdf.



El primer cuadro es su documento en Rmarkdown, donde escriben su código en R y lo combinan con texto. [Knitr](#) es el paquete que convierte todo lo que ustedes escriben a formato Markdown (.md), lo cual posteriormente es transformado por [Pandoc](#) en cualquiera de los formatos que ustedes necesiten (.pdf, .html o .doc). El resultado de este proceso es su documento final.

Software necesario para compilar en .pdf

Cuando estamos empleando documentos en .pdf, Pandoc requiere de que ustedes tengan instalado en su computador alguno de los paquetes base para escribir documentos en \LaTeX . Para esto deben instalar alguno de los paquetes disponibles, en mi caso uso [Miktex](#), el cual tiene soporte para Windows, Mac y Linux. Hasta ahora no me ha dado problemas y lo recomiendo.

1. Ir a la web de [Miktex](#)
2. Descargar el instalador y ejecutarlo
3. Esperar que termine la descarga de los paquetes, lo cual puede demorar varios minutos.
4. Cuando esté todo listo, reinicien su computador.

Finalizado este paso, pueden comenzar a trabajar con Rmarkdown sin problemas.

Elaborar títulos y enumeración

#Título grande

##Título mediano

##Título pequeño

@ Para crear numeración con interrupciones (autonumerado)

1. Para crear numeración según sus necesidades

a. Para crear sub-numeración con letras, también se puede realizar con números.

* Una viñeta cuadrada

Ecuaciones y símbolos matemáticos

$e = mc^2$

$\alpha \chi^2 \beta$

Hay una hoja de consejos para usar Rmarkdwon que pueden encontrar [aquí](#). Aparecen muchos más detalles de lo que se señala en este documento.

1. Usar chunks

- el ctrl + alt + i

```
95 ~~~{r, eval=FALSE, include=TRUE, message=FALSE, warning=FALSE, results='asis'}
96 {r, eval=FALSE, include=TRUE, message=FALSE, warning=FALSE, results='asis'}
97
98 summary(mtcars)
99 x <- c(1,2,3,4)
100
101 #Esto es un Chunk
102 ~~~
```

- `eval=TRUE` Sirve para determinar si queremos que se vean nuestros resultados.
- `include=TRUE` Sirve para determinar si queremos que se incluya nuestro código.
- `message=FALSE` Sirve para determinar si queremos los mensajes emergentes.
- `warning=FALSE` Sirve para determinar si queremos las advertencias de R.
- `results='asis'` Permite que el código creado por las funciones de R sea empleado en la compilación.
- ¡run! (flecha verde o ctrl+shift+enter)

Ejemplo:

```
#{r, results='hold'}
revolucion <- "Marx"
revolucion

velociraptor <- "rawr"
velociraptor

## [1] "Marx"
## [1] "rawr"
```

2. Paquetes especializados para reporte de tablas

- knitr
- kableExtra
- xtable
- texreg
- stargazer

Instalemos los paquetes:

```
install.packages("knitr")
install.packages("kableExtra")
install.packages("xtable")
install.packages("texreg")
install.packages("stargazer")
```

Carguemos los paquetes:

```
library(knitr)
library(kableExtra)
library(xtable)
library(texreg)
library(stargazer)
library(dplyr) #principalmente para usar el operador %>%
```

Ejemplo con regresión usando texreg

El paquete `texreg` es muy útil cuando se trata de reportar modelos de regresión, tiene bastantes funcionalidades, pero las principales son el hecho de que permite reportar modelos anidados en una sola tabla.

```
#Un logit fome...
pl <- lm(voto ~ sexo + edad + educon + ecivil + ppolt+ socconf+confl, data=coes, link="logit")

texreg::texreg(pl, #Si son más modelos ponemos list(m1,m2,m3)
  digits = 3, #dígitos de la tabla
  float.pos="h!", #permite dejar la tabla fija en su lugar
  scalebox = 0.70, #indica que la tabla tiene una proporción del 75% c
  caption = "Modelo Logit", #título del modelo
  custom.coef.names=c("(Intercepto)", #Podemos asignar nombres a los coeficientes
    "Mujer", "Edad", "Educación",
    "Soltero/a", "Viudo/a", "Sep/Div/Anu",
    "Centro", "Derecha",
    "No sabe/No responde",
    "Confianza Social",
    "Per.Conflicto"),
  custom.model.names = "Modelo 1") #si son más modelos = c("Modelo1", "Modelo2")
```

	Modelo 1
(Intercepto)	0,056 (0,068)
Mujer	0,004 (0,020)
Edad	0,009*** (0,001)
Educación	0,034*** (0,005)
Soltero/a	-0,108*** (0,025)
Viudo/a	-0,007 (0,045)
Sep/Div/Anu	-0,060 (0,034)
Centro	-0,049 (0,028)
Derecha	0,095** (0,032)
No sabe/No responde	-0,075** (0,028)
Confianza Social	-0,008 (0,010)
Per.Conflicto	0,048*** (0,014)
R ²	0.140
Adj. R ²	0.135
Num. obs.	1947
RMSE	0.436

*** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$

Tabla 1: Modelo Logit

Con modelos anidados

```
#``{r , echo=TRUE, message=FALSE, results='asis'}
m1=glm(voto~sexo+edad+educon+ecivil+ppol+socconf,
      data=coes, family=binomial) #Educación continua
m2=glm(voto~sexo+edad+educ +ecivil+ppol+socconf,
      data=coes, family="binomial") #Educación categórica
m3=glm(voto~sexo+edad+educon+ecivil+ppol+socconf+ppol*socconf,
      data=coes, family="binomial") #Interacción Posición Política*Confianza Social
texreg::texreg(list(m1,m2,m3), digits = 3,float.pos="h!",scalebox=0.50,
  caption = "Modelos Logit",
  custom.model.names = c("Modelo 1", "Modelo 2", "Modelo 3"),
  custom.coef.names = c("(Intercepto)","Mujer","Edad",
    "Educación","Soltero/a","Viudo/a",
    "Sep/Div/Anu","Centro","Derecha",
    "No sabe/No responde","Confianza Social",
    "Primaria","Secundaria","Técnica superior",
    "Universitaria","Centro:Conf.Social",
    "Derecha:Conf.Social","NS/NR:Conf.Social"))
```

	Modelo 1	Modelo 2	Modelo 3
(Intercepto)	-1,911*** (0,329)	-1,344*** (0,346)	-1,683*** (0,341)
Mujer	-0,008 (0,108)	-0,012 (0,108)	-0,010 (0,108)
Edad	0,045*** (0,004)	0,044*** (0,004)	0,046*** (0,004)
Educación	0,185*** (0,029)		0,190*** (0,029)
Soltero/a	-0,484*** (0,127)	-0,531*** (0,129)	-0,468*** (0,128)
Viudo/a	0,025 (0,273)	-0,026 (0,273)	-0,004 (0,274)
Sep/Div/Anu	-0,363* (0,182)	-0,357* (0,182)	-0,372* (0,183)
Centro	-0,248 (0,149)	-0,254 (0,149)	-0,571** (0,202)
Derecha	0,567** (0,188)	0,575** (0,188)	-0,016 (0,248)
No sabe/No responde	-0,391** (0,148)	-0,401** (0,148)	-0,664*** (0,200)
Confianza Social	-0,024 (0,054)	-0,020 (0,054)	-0,305** (0,111)
Primaria		0,281 (0,257)	
Secundaria		0,344 (0,202)	
Técnica superior		0,673** (0,233)	
Universitaria		1,124*** (0,234)	
Centro:Conf.Social			0,338* (0,145)
Derecha:Conf.Social			0,677*** (0,201)
NS/NR:Conf.Social			0,277 (0,148)
AIC	2201.572	2213.736	2194.613
BIC	2262.887	2291.773	2272.649
Log Likelihood	-1089.786	-1092.868	-1083.306
Deviance	2179.572	2185.736	2166.613
Num. obs.	1947	1947	1947

***p < 0.001, **p < 0.01, *p < 0.05

Tabla 2: Modelos Logit

Kable y KableExtra

```
#```{r tabla, echo=TRUE, message=FALSE, results='asis'}
age = c(1,2,3,4,5)
ageF = factor(age, labels=c("35-44", "45-54", "55-64", "65-74", "75-84"))
deaths1 = c(32,104,206,186,102)
deaths2 = c(2,12,28,28,31)

#Person-Years
py1 = c(52407,43248,28612,12663,5317)
py2 = c(18790,10673,5710,2585,1462)

tabla <- as.data.frame(cbind(ageF,deaths1,py1,deaths2,py2)) #creamos el data.frame
tabla$ageF = factor(age, labels=c("35 - 44", "45 - 54", "55 - 64", "65 - 74", "75 - 84"))

#----nuestra tabla-----
kable(tabla, booktabs = TRUE, escape = FALSE, align = c("lcccc"),
      caption = "Muertes por enfermedad cardiaca después de 10 años",
      col.names = linebreak(c("Grupo de\n edad", "Muertes\n ", "Persona-años\n ",
                              "Muertes\n ", "Persona-años\n"), align = "c")) %>%
kable_styling(latex_options = c("hold_position")) %>%
add_header_above(c(" " = 1, "Fumadores" = 2, "No fuma" = 2))
```

Tabla 3: Muertes por enfermedad cardiaca después de 10 años

Grupo de edad	Fumadores		No fuma	
	Muertes	Persona-años	Muertes	Persona-años
35 - 44	32	52407	2	18790
45 - 54	104	43248	12	10673
55 - 64	206	28612	28	5710
65 - 74	186	12663	28	2585
75 - 84	102	5317	31	1462

Algunos tutoriales que me gustan:

- [Escribir una Tesis en Rmarkdown](#) (Van Espen, 2017)
- [Rmarkdown ultimate Guide](#) (Xie et al. 2018)
- [KableExtra y Latex](#) (Zhu, 2018)