

R como herramienta de escritura
académico-científica con visualización estática y
dinámica

Alexander Andrade

2020-02-29

Contents

1	Resumen	1
	Bibliografía básica	1
2	Requisitos para el taller	3
	2.1 Conocimiento previo (Recomendado)	3
	2.2 Software	3
3	Introducción	7
4	Visualización Estática y Dinámica	9
	4.1 Visualización Estática	9
	4.2 Visualización Dinámica	11
5	Demuestra lo aprendido	15
6	Acerca del autor	17

Chapter 1

Resumen

R es un lenguaje de programación estadística, que permite: importar, gestionar, analizar y presentar información (datos), con un alto nivel de reproducibilidad. Consolidándose así, como el software ideal en: estadística, machine learning, ciencia de datos, etc.

En el taller denominado *R como herramienta de escritura académico-científica con visualización estática y dinámica*, se estudian las ventajas que ofrece el paquete `rmarkdown`, el cual permite generar reportes estáticos y dinámicos, con el fin de divulgar tanto los resultados de un estudio, como su metodología y el código utilizado.

Particularmente, el taller se enfoca en la creación de libros académico-científicos (en formato HTML, EPUB y PDF) mediante el paquete `bookdown`, mismo que puede entenderse como una extensión de `rmarkdown`.

Al finalizar el taller, los asistentes serán capaces de:

- Entender de manera global, el funcionamiento de `rmarkdown`, su sintaxis y sus principales características.
- Construir libros digitales mediante el paquete `bookdown` (similares al libro que se está visualizando en este momento).
- Conocer la diferencia entre visualización estática y dinámica.
- Incorporar las bondades de `bookdown` en la divulgación científica.

Bibliografía básica

Libros:

- *Bookdown: Authoring Books and Technical Documents with R Markdown* (Xie Yihui, 2016)

- *R markdown: The definitive guide* (Xie et al., 2018)
- *Dynamic Documents with R and knitr* (Xie, 2015)

Paquetes:

- *Bookdown* (Xie, Yihui, 2019)
- *Rmarkdown* (Allaire et al., 2019)

Chapter 2

Requisitos para el taller

2.1 Conocimiento previo (Recomendado)

- Entendimiento básico de R y RStudio.
- Nivel básico de `rmarkdown`

¿Qué ocurre si nunca he trabajado con R, Rstudio o rmarkdown?

Previo a este taller, se impartirán charlas y talleres en los siguientes temas:

- R/RStudio
- Rmarkdown
- R shiny

Los cuales brindarán un importante background teórico-práctico sobre R.

2.2 Software

2.2.1 Recursos indispensables

1. Lenguaje de programación estadística **R**: <https://cran.r-project.org/>
2. Interfaz de desarrollo integrado **RStudio**: <https://rstudio.com/products/rstudio/download/>
3. **MiKTeX**: <https://miktex.org/download>

Nota: Se recomienda que la instalación de los recursos previamente mencionados, se realice en el orden en el que fueron enumerados.

2.2.2 Recursos recomendados

1. Lector de PDF
2. Lector de EPUB

Nota: El lector de archivos PDF y archivos EPUB queda a disposición de los asistentes. Una posible opción es FDReader (recomendado).

2.2.3 Paquetes

1. tidyverse

```
install.packages("tidyverse")
```

2. bookdown

```
install.packages("bookdown")
```

3. shiny

```
install.packages("shiny")
```

4. knitr

```
install.packages("knitr")
```

5. DT

```
install.packages("DT")
```

6. plotly

```
install.packages("plotly")
```

7. tinytex

```
install.packages("tinytex")  
tinytex::install_tinytex()
```

8. phantomjs


```
install.packages("webshot")  
webshot::install_phantomjs()
```


Chapter 3

Introducción

La reproducibilidad es sumamente importante en investigación científica y en el proceso de enseñanza-aprendizaje, principalmente cuando se está enseñando (o aprendiendo) un lenguaje de programación. En tal contexto, el paquete `bookdown` permite compartir el código y los resultados de un experimento en forma de libros virtuales, ayudando al investigador o docente, a enfocarse en el aspecto científico y facilitando el proceso de redacción.

Principales ventajas de `bookdown`

- Reproducibilidad de la investigación.
- Existe una gran cantidad de temas o templates que permiten crear libros virtuales estéticamente agradables.
- La optimización del paquete con la interfaz de RStudio logra un vínculo amigable para el usuario.
- La estructura de la herramienta resalta las características propias de la redacción científica.
- Los formatos de salida del libro automáticamente contemplan: HTML, PDF, EPUB.
- En formato HTML es posible presentar los resultados de forma dinámica, es decir, el lector podrá interactuar con la información que se comparte en el libro.
- Es posible crear quiz interactivos para el lector, de tal manera que su aprendizaje se vea reforzado mediante la práctica.

Chapter 4

Visualización Estática y Dinámica

Como se mencionó en el capítulo anterior, una de las ventajas de `bookdown` es la visualización dinámica de resultados. En este apartado se presentan varios ejemplos con la intención de resaltar las bondades de la visualización dinámica, sobre la visualización estática. Es importante tener en claro que la visualización dinámica de resultados es posible únicamente en la versión HTML del libro digital.

4.1 Visualización Estática

4.1.1 Ejemplo 1

A continuación se presenta una tabla con las diez primeras observaciones del dataset `Iris`. Este conjunto de datos está conformado por medidas de ciertas características físicas de pétalos y sépalos sobre 150 plantas, tomando en cuenta tres especies: Setosa, Versicolor y Virginica.

```
knitr::kable(head(iris, 10))
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa

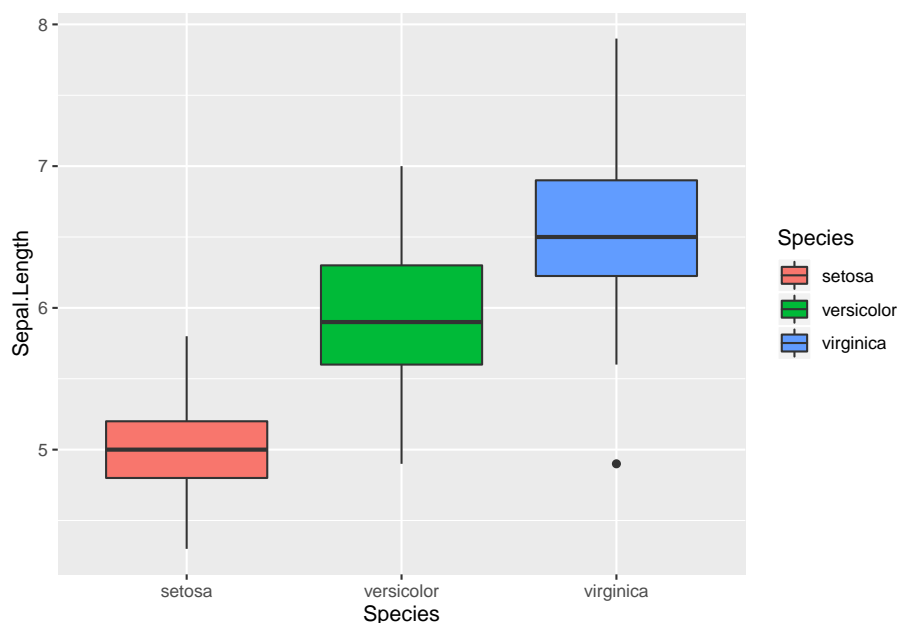
Aún cuando la tabla es estéticamente agradable, únicamente se observa una cierta cantidad de datos, los cuales no siguen ningún orden específico, y además, no es posible filtrar la información según ciertos parámetros de interés.

4.1.2 Ejemplo 2

La información del dataset `Iris` se resume en un diagrama de caja (boxplot), el cual compara la longitud del sépalo de las plantas, tomando en cuenta las tres especies bajo estudio.

```
library(tidyverse)

iris %>%
  ggplot(aes(Species, Sepal.Length, fill = Species)) +
  geom_boxplot()
```



El boxplot es una herramienta visual muy importante y útil a la hora de comparar una variable medida en muestras de diferentes poblaciones. Si bien el gráfico nos presenta información relevante sobre los individuos (plantas), esta información puede aprovecharse de mejor manera con técnicas de visualización dinámica.

4.2 Visualización Dinámica

4.2.1 Ejemplo 3

Nuevamente se presenta el dataset *Iris*, pero ahora mediante una tabla dinámica. En este caso el lector puede interactuar con la tabla, realizando las siguientes acciones:

- Cambiar el número de observaciones que se muestran (**Show entries**). Por defecto se pueden mostrar 10, 25, 50 o 100 observaciones. Este rango de valores se puede cambiar según las necesidades del investigador.
- Buscar (**Search**) un tipo de planta o un número en particular. Por ejemplo, el lector puede buscar la palabra “*virginica*”, de tal forma que se muestren únicamente las observaciones de esta especie.
- Ordenar el dataset según una de las variables, de forma ascendente o descendente. Para ello, el lector simplemente debe dar clic en el nombre de la variable que se quiera ordenar.

- Visualizar (*Previous*) diferentes porciones del dataset.

```
DT::datatable({iris})
```

Show entries

Search:

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa

Showing 1 to 10 of 150 entries

Previous 2 3 4 5 ... 15 Next

4.2.2 Ejemplo 4

En este ejemplo se presenta el mismo boxplot del Ejemplo 2, en formato dinámico. En este gráfico, el lector puede realizar las siguientes acciones:

- Al pasar el puntero del mouse sobre el boxplot, es posible obtener información específica sobre los valores de: cuartiles, máximo, mínimo, datos atípicos.
- En la parte superior del gráfico aparece una barra con las siguientes opciones: descargar el gráfico, realizar zoom, centrar la imagen, auto escalar, etc. El lector puede hacer uso de cualquiera de estas opciones sin preocuparse por afectar permanentemente el documento.
- Al dar clic sobre una de las tres especies que se muestran en la leyenda (*Species*), el lector podrá habilitar o deshabilitar cualquiera de los tres diagramas de caja que se muestran.

```
p <- iris %>%
  ggplot(aes(Species, Sepal.Length, fill = Species)) +
  geom_boxplot()

plotly::ggplotly(p)
```


4.2.3 Ejemplo 5

Finalmente, se presenta otra forma de visualización dinámica del boxplot que hemos analizado en los ejemplos 2 y 4, en dichos ejemplos se ha visualizado únicamente la variable longitud del sépalo (*Sepal.Length*). A continuación, el lector podrá escoger la variable del dataset *Iris* que prefiera analizar.

```
knitr::include_app("https://vaashub.shinyapps.io/iris_boxplot/",  
  height = "650px")
```

Selection

Variable

Sepal.Length ▼

Plot

Chapter 5

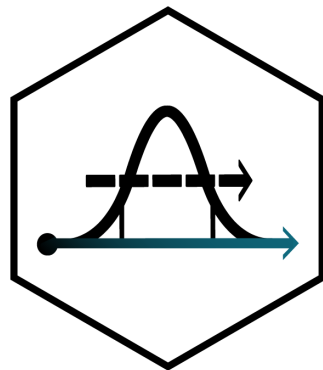
Demuestra lo aprendido

Quiz

[Start Over](#)

Chapter 6

Acerca del autor



ESTADÍSTICA
PARA NO ESTADÍSTICOS

ALEXANDER ANDRADE

Asesor Estadístico en Estadística para No Estadísticos

Estadística para No Estadísticos brinda asesoría y tutoría estadística en:

- Data Science para: Instituciones, Empresas y Profesionales
- Trabajos de Titulación (Pregrado y Posgrado)
- Artículos Científicos (Papers)
- Artículos de Revisión Bibliográfica mediante Minería de Texto

¿Tienes dudas o comentarios?

Contáctanos: www.estadisticaparanostadisticos.com

Bibliography

Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., and Iannone, R. (2019). *rmarkdown: Dynamic Documents for R*. R package version 2.0.

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y., Allaire, J. J., and Golemund, G. (2018). *R markdown: The definitive guide*. Chapman and Hall/CRC.

Xie Yihui (2016). *Bookdown: Authoring Books and Technical Documents with R Markdown*. Chapman and Hall/CRC.

Xie, Yihui (2019). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.16.