

Generación de informes reproducibles con RMarkdown

Día del estadístico peruano

Mg. Jesús Eduardo Gamboa Unsihuay

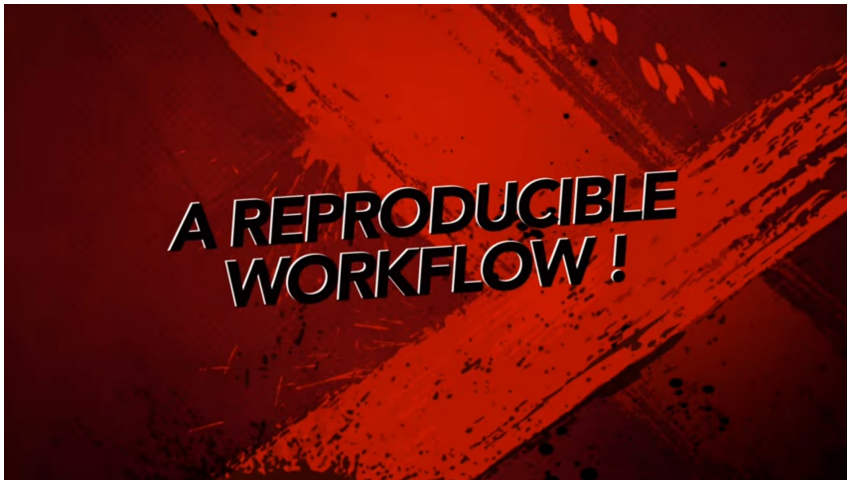
05/12/2020

- 1 **Introducción**
- 2 **Proyectos en RStudio**
- 3 **Trabajando con RMarkdown**

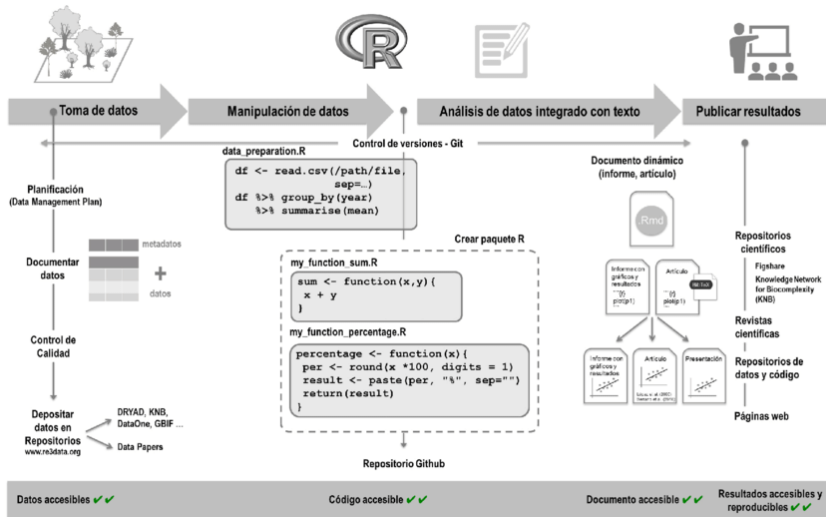
Section 1

Introducción

Introducción



<https://www.youtube.com/watch?v=s3JldKoA0zw>



Section 2

Proyectos en RStudio

Proyectos en RStudio

Reflexionemos acerca de nuestro modo de trabajo en RStudio:

- 1 ¿Tus datos son 'vecinos' de tus scripts?
- 2 ¿Cómo lees conjuntos de datos?
- 3 ¿Cargas el espacio de trabajo al iniciar sesión? ¿lo guardas al finalizar?

Los datos como 'vecinos' de los scripts

- Evitar casos como este:

```
read.csv("C:/Documentos/Curso Regresion/Trabajo 1/datos.csv")  
read.delim("clipboard")
```

- Datos vecinos de los scripts → facilitará la lectura.
- No será necesario indicar la ruta:

```
read.csv("datos.csv")
```


Lectura de datos

Ejemplo:

```
datos <- read.csv2("datos.csv", dec = ".")
```

datos

##		X	Y
##	1	2	12.06
##	2	6	32.83
##	3	1	8.28
##	4	6	32.39
##	5	4	21.59
##	6	3	17.61
##	7	10	53.10
##	8	7	35.74

Manipulación de datos

- Use R para limpiar, transformar y manipular conjuntos de datos
- Funciones de manipulación en el paquete **dplyr**: filter, select, mutate, etc.

```
library(dplyr)
datos2 <- datos %>%
  mutate(Z = Y - X)
```

No cargar ni guardar espacio de trabajo

- Guardar códigos en vez de almacenar el espacio de trabajo
- De ser necesario almacenar un objeto, hacerlo en formato RDS en vez de Rda o RData
- Ejemplo:

```
modelo <- lm(Y ~ X, data = datos)
saveRDS(modelo, "modeloRL.RDS")
```

Exportando objetos

```
write.csv(datos2,"datos2.csv")  
library(ggplot2)  
imagen = ggplot(datos, aes(X,Y)) +  
  geom_point() + theme_minimal()  
ggsave("imagen.png",imagen)
```

Algunos puntos adicionales

- No colocar **install.packages("...")** en código reproducible.
- Los proyectos son el punto de partida para la creación de paquetes

Section 3

Trabajando con RMarkdown

Trabajando con RMarkdown



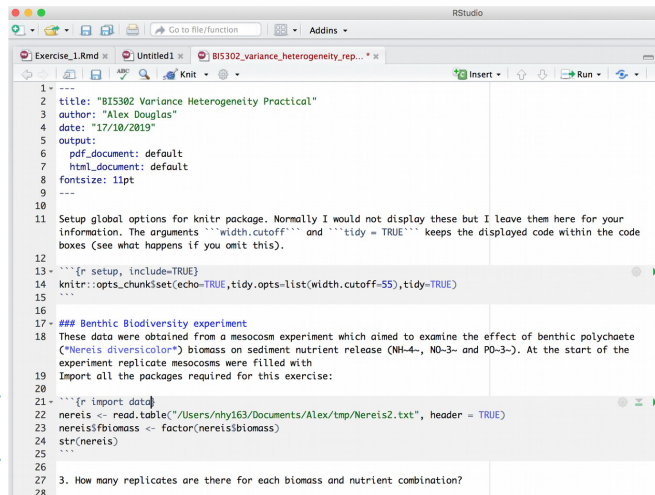
Texto + código + resultados

<https://rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>

YAML
header

formatted
text

code
chunk



```
1 ---
2 title: "BI5302 Variance Heterogeneity Practical"
3 author: "Alex Douglas"
4 date: "17/10/2019"
5 output:
6   pdf_document: default
7   html_document: default
8 fontsize: 11pt
9 ---
10
11 Setup global options for knitr package. Normally I would not display these but I leave them here for your
12 information. The arguments ``width.cutoff`` and ``tidy = TRUE`` keeps the displayed code within the code
13 boxes (see what happens if you omit this).
14
15 ```{r setup, include=TRUE}
16 knitr::opts_chunk$set(echo=TRUE, tidy.opts=list(width.cutoff=55), tidy=TRUE)
17 ```
18
19 ## Benthic Biodiversity experiment
20 These data were obtained from a mesocosm experiment which aimed to examine the effect of benthic polychaete
21 (*Nereis diversicolor*) biomass on sediment nutrient release (NH4~, NO3~ and PO4~). At the start of the
22 experiment replicate mesocosms were filled with
23 Import all the packages required for this exercise:
24
25 ```{r import data}
26 nereis <- read.table("/Users/nhy163/Documents/Alex/tmp/Nereis2.txt", header = TRUE)
27 nereis$fbiomass <- factor(nereis$biomass)
28 str(nereis)
29 ```
30
31 3. How many replicates are there for each biomass and nutrient combination?
```


- Documentos → HTML, PDF, Word
- Presentaciones → HTML, PDF, Power Point
- Dashboard estático (flexdashboard)
- Dashborad interactivo (shiny)
- Tutorial (learnr)
- Blog (distill)
- Libro (bookdown)
- Publicacion en GitHub
- Publicación en RPubS

Ejemplo Flexdashboard + RPubs



<https://rpubs.com/jesuseduardog/covid19peru>

Esta presentación está alojada en

<https://github.com/jeguns/reproducibilidad>

Contacto: jgamboa@lamolina.edu.pe