

Usando y Enseñando R para Investigación Reproducible

Rayna M. Harris

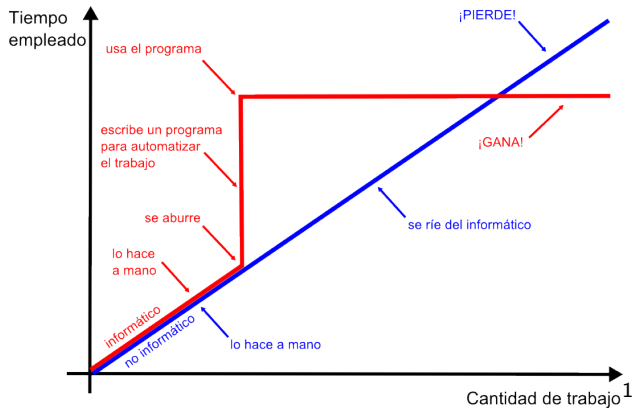
Twitter y Instagram y GitHub: @raynamharris

Página web: <https://raynamharris.github.io>

27 Marzo 2018

R-Ladies Buenos Aires

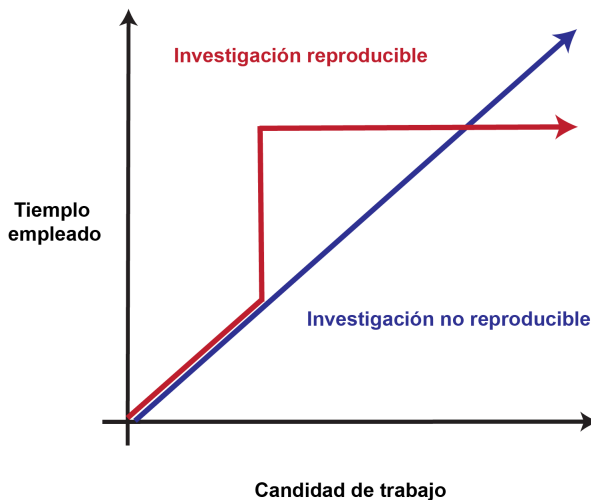
La programación es importante porque permite automatizar tareas.



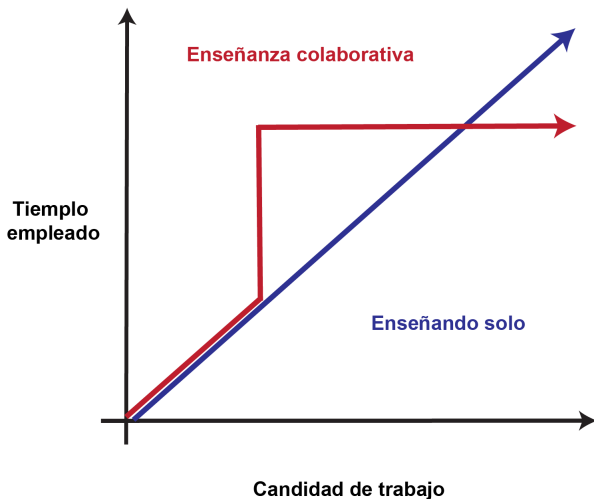
¹http:

//www.mclibre.org/consultar/python/otros/lenguajes-programacion.html

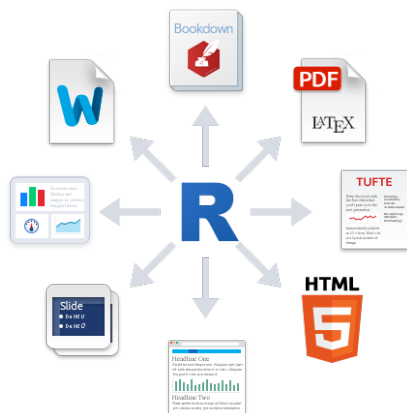
R permite estadísticas reproducibles y visualización de datos



La enseñanza colaborativa también ahorra tiempo con una mayor cantidad de trabajo



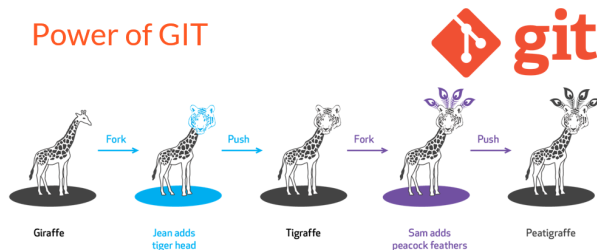
Consejo 1: Usa R Markdown para crear y compartir documentos



2

²https://rmarkdown.rstudio.com/authoring_quick_tour.html

Consejo 2: Usa el control de versiones para colaborar con otros y con vos en el futuro

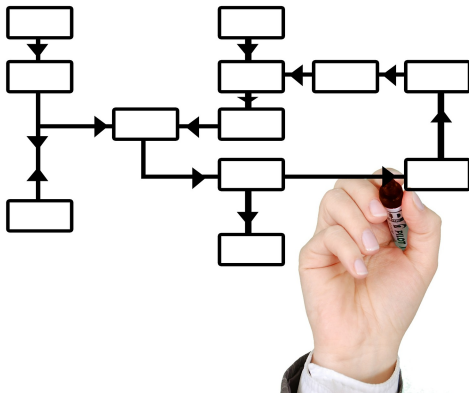


3

³<http://technetnepal.net/blogs/shirishamaharjan/archive/2017/05/07/expand-horizons-change-attitudes-git-and-github-workshop.aspx>

Consejo 3: Documenta tu flujo de trabajo

Porque probablemente sea único y complejo



4

⁴<https://www.blogdelfotografo.com/workflow-flujo-trabajo-foto/>

Ejemplo de archivo README.md

GitHub, Inc. [US] | <https://github.com/raynamharris/FMR1CA1rnaseq> 🔍 ☆

UNIXworkflow	organization	22 days ago
data	now with fig 1	4 months ago
docs	better intro	22 minutes ago
figures	better intro	22 minutes ago
results	with integration	3 months ago
scripts	rename adobe files	4 days ago
.gitignore	better intro	22 minutes ago
CITATION	zenodo	4 months ago
README.md	updated readme	just now

Repository organization and workflow

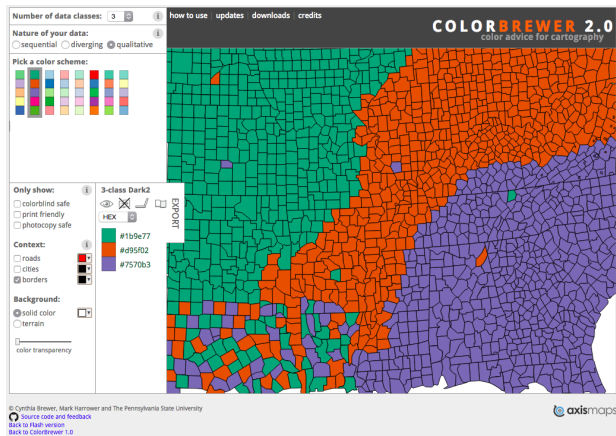
1. [UNIXworkflow](#) - runs on The Texas Advanced Computing Center's (TACC) cluster 'Stampede'
2. [data](#) - mostly raw data and meta data, some results included
3. [docs](#) - my talk for the R-Ladies Buenos Aires group
4. [scripts](#)
 - o [00_cpfiles.sh](#) - a bash script for copying files into and out of this repo
 - o [01_behavior.Rmd](#) - output of a knittable analysis script for wrangling the avoidance behavior data. The output is use for the [Figures](#), [Rmd](#) scripts that accompany each figure
 - **Figures 1, 2, 3, 4, 5, 6, 7** - Because I had so many behavior figures and accompanying statistics, I created a Rmd script containing the code to make the figure and calculate the statistics. The data were generated in [01_behavior.Rmd](#).
 - o [02_RNAseq.Rmd](#) - output of an knittable analysis of differential gene expression of RNA-seq data with DESeq2, GO-WMU, and WGCNA
 - o [03_wgcna.R](#) - R script for weighted gene co-expression network analysis

5

⁵<https://github.com/raynamharris/FMR1CA1rnaseq>

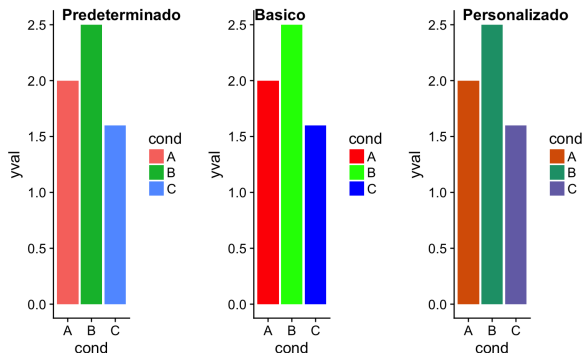
Consejo 4: Desarrolla tu propia paleta de colores

Colorbrewer⁶ te ayuda a elegir colores amigables para daltónicos



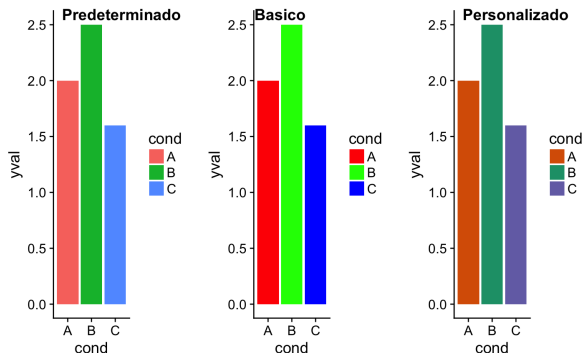
⁶<http://colorbrewer2.org/>

Ejemplos de paletas de colores en ggplot



- Basico: + `scale_fill_manual(values=c("red", "green", "blue"))`

Ejemplos de paletas de colores en ggplot



- ▶ Basico: + `scale_fill_manual(values=c("red", "green", "blue"))`
- ▶ Personalizado: + `scale_fill_manual(values=c("#d95f02", "#1b9e77", "#7570b3"))`

Tu podés convertir HEX a RGB para usar la misma paletas para las ilustraciones

#d95f02

rgb(217,95,2)



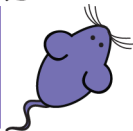
#1b9e77

rgb(27,158,119)



#7570b3

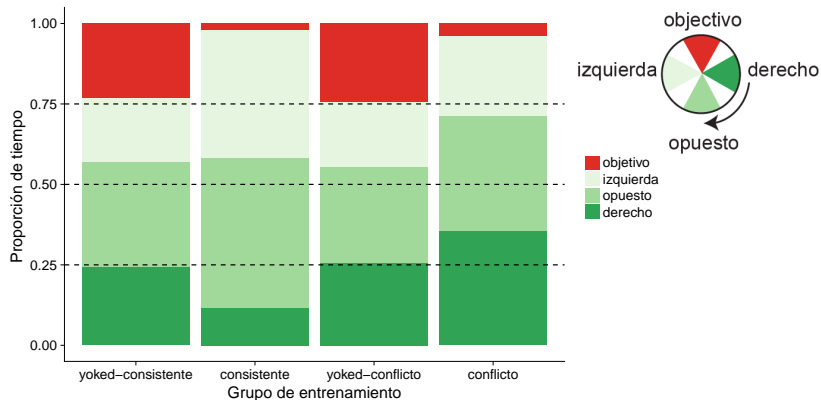
rgb(117,112,179)



7

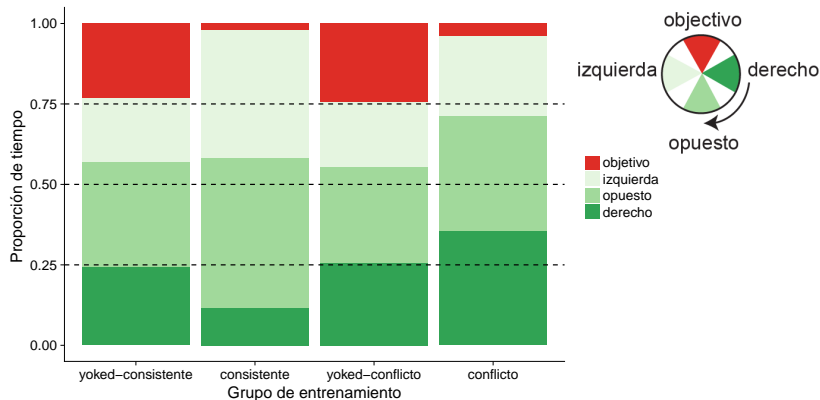
⁷<https://www.webpagefx.com/web-design/hex-to-rgb/>

Consejo 5: Usa leyendas gráficas



- Leyendas gráficas transmitir más información

Consejo 5: Usa leyendas gráficas



- ▶ Leyendas gráficas transmitir más información
- ▶ Use **cowplot**⁸ para agregar imágenes dentro de R

⁸**cowplot** <https://cran.r-project.org/web/packages/cowplot/index.html>









Punto medio resumen

- ▶ Consejo 1: Usa **R Markdown** para crear y compartir documentos
- ▶ Consejo 2: Usa el control de versiones para la colaboración
- ▶ Consejo 3: Documenta tu flujo de trabajo
- ▶ Consejo 4: Desarrolla tu propia paleta de colores
- ▶ Consejo 5: Usa leyendas gráficas









Desarrollo colaborativo de la lección



Our lessons in English

Lesson	Site	Repository	Reference	Instructor Guide	Maintainer(s)
The Unix Shell					Gabriel Devenyi , Ashwin Srinath , Colin Morris, Will Pitchers
Version Control with Git					Ivan Gonzalez , Daisie Huang , Nima Hejazi , Katherine Koziar

Our lessons in Spanish

Lesson	Site	Repository	Reference	Instructor Guide	Maintainer(s)
La Terminal de Unix					Heladia Saldago
Control de versiones con Git					Rayna M. Harris

9

⁹<https://software-carpentry.org/lessons/>

Deseo 1: Me ayudás a mejorar las nuevas lecciones en español de Software Carpentry

R for Reproducible Scientific Analysis

R para Análisis Científicos Reproducibles

Una introducción a R para no-programadores utilizando los datos de [Gapminder][gapminder]. Consulta la página <https://swcarpentry.github.io/r-novice-gapminder> para visualizar el material y [la documentación de la plantilla de la lección][lesson-example] para obtener las instrucciones sobre cómo formatear, compilar y enviar material, o ejecuta `make` en este directorio para obtener una lista de comandos útiles.

El objetivo de esta lección es enseñar a los programadores principiantes a escribir códigos modulares y las buenas prácticas en el uso de R para el análisis de datos. R y un conjunto de paquetes desarrollados por terceros, se usa comúnmente en diversas disciplinas científicas para el análisis estadístico. Encontramos que muchos científicos que asisten a los talleres de to Software Carpentry utilizan R y quieren aprender más. La importancia de estos materiales es proporcionar a los asistentes una base sólida en los fundamentos de R, y enseñar las mejores prácticas del cómputo científico: desglose del análisis en módulos, automatización tareas y encapsulamiento.

Ten en cuenta que este taller se enfoca en los fundamentos del lenguaje de programación R, y no en el análisis estadístico.

A lo largo de este taller se utilizan una variedad de paquetes, desarrollados por terceros, los cuales no necesariamente son los mejores, ni se ven todas las funcionalidades de los mismos, pero son paquetes que consideramos útiles y han sido elegidos principalmente por su facilidad de uso.

🌟 Prerequisites

Understand that computers store data and instructions (programs, scripts etc.) in files. Files are organised in directories (folders). Know how to access files not in the working directory by specifying the path.

Horario

	Setup	Download files required for the lesson
00:00	1. Introducción a R y RStudio	¿Cómo orientarse en RStudio? ¿Cómo interactuar con R? ¿Cómo administrar tu entorno? ¿Cómo instalar paquetes?
00:55	2. Gestión de proyectos con RStudio	¿Cómo pudo gestionar mis proyectos en R?
01:25	3. buscando ayuda	¿Cómo puedo obtener ayuda en R?

Como podés ayudarme

- ▶ Leer y comentar o editar en GitHub¹⁰

¹⁰<https://swcarpentry.github.io/r-novice-gapminder-es/>

Como podés ayudarme

- ▶ Leer y comentar o editar en GitHub¹⁰
- ▶ Participar en el **Bug BBQ**¹¹ el Abril 11 y 12

¹⁰<https://swcarpentry.github.io/r-novice-gapminder-es/>

¹¹<https://carpentries.github.io/2018-04-bug-bbq/>

Como podés ayudarme

- ▶ Leer y comentar o editar en GitHub¹⁰
- ▶ Participar en el **Bug BBQ**¹¹ el Abril 11 y 12
- ▶ Hacer videos de vos leyendo y codificando junto con la lección¹²

¹⁰<https://swcarpentry.github.io/r-novice-gapminder-es/>

¹¹<https://carpentries.github.io/2018-04-bug-bbq/>

¹²https://www.youtube.com/watch?v=bXxBeNkKmJE&list=PLNfIROxhN1JAtIbznqSoFODe0Z_WSVL4D

Deseo 2: Convertirse en una instructora certificada

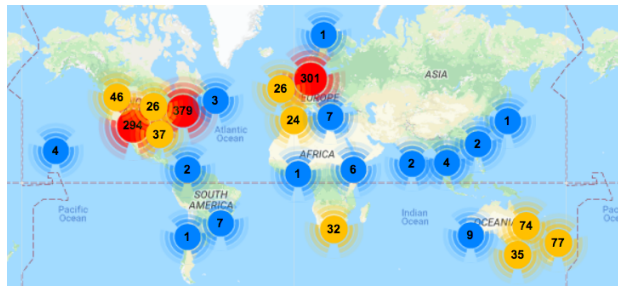
- ▶ Ahora, no hay instructoras en Argentina :(
- ▶ Aplicá aquí: <http://carpentries.github.io/instructor-training/>



13

Deseo 2: Convertirse en una instructora certificada

- ▶ Ahora, no hay instructoras en Argentina :(
- ▶ Aplicá aquí: <http://carpentries.github.io/instructor-training/>
- ▶ Usa el **Group Name** “R Ladies Buenos Aires”



13

¹³<https://software-carpentry.org/team/>

Deseo 3: ¡Asiste a nuestro primer taller de español!

Secure | <https://raynamharris.github.io/2018-04-07-BuenosAires/>

Home Code of Conduct License Improve this page Search...

R'lyeh Hacklab, Buenos Aires

7 Abril 2018
13:00 - 20:00

Instructors: Rayna Harris, Juli Arancio, Paola Prieto, Cynthia, Vladimir

Helpers: Juli Arancio, Paola Prieto, Cynthia, Vladimir

Registration	
SAT, APRIL 7, 2018, 1:00 PM - 6:00 PM PDT Sales end on April 26	
RSVP FREE	11 Remaining
QTY: 1	FREE
REGISTER	

14

¹⁴<https://raynamharris.github.io/2018-04-07-BuenosAires/>

Deseo 4: Organizar unos talleres en el futuro

Want to Know More?

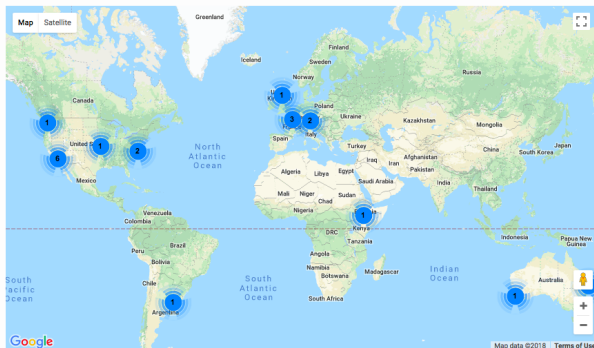
See our [FAQ](#).

Want to Run a Workshop?

See our [operations guide](#).

Want to Host a Workshop?

Please [fill in this form](#).



Future Workshops

15

¹⁵<https://software-carpentry.org/workshops/>

Deseo 5: Adoptar la práctica del desarrollo colaborativo de lecciones

10 Reglas Simples para el Desarrollo Colaborativo de las Lecciones

1. Aclarar tu audiencia



2. Hacer módulos de lecciones



3. Enseñe lo mejor prácticas para desarrollo de lecciones



4. Fomentar y empoderar contribuyentes



5. Construir comunidad alrededor a las lecciones



6. Publicar periódicamente y reconocer contribuciones



7. Evaluar lecciones en varias escalas



8. Reduzca, reutilizar, reciclan



9. Enlace a otros recursos



10. No podés complacer a todos



16

¹⁶Devenyi et al. 2018 PLOS Comp Bio
<http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005963>

Resumen

- ▶ Consejo 1: Usa *R Markdown* para la reproducibilidad
- ▶ Consejo 2: Usa el control de versiones para la colaboración
- ▶ Consejo 3: Documenta tu flujo de trabajo
- ▶ Consejo 4: Desarrolla tu propia paleta de colores
- ▶ Consejo 5: Usa leyendas gráficas
- ▶
- ▶ Deseo 1: Me ayudás a mejorar las lecciones en español
- ▶ Deseo 2: Convertirse en una instructora certificada
- ▶ Deseo 3: ¡Asiste a nuestro primer taller de español!
- ▶ Deseo 4: Organizar unos talleres en el futuro
- ▶ Deseo 5: Adoptar la práctica del desarrollo colaborativo de lecciones

Pensamiento concluyente

- ▶ Yo creo que todos aprenden más cuando la ciencia y la educación está abiertas y reproducibles

Pensamiento concluyente

- ▶ Yo creo que todos aprenden más cuando la ciencia y la educación está abiertas y reproducibles
- ▶ Yo creo que la mejor manera de aprender es enseñando

Pensamiento concluyente

- ▶ Yo creo que todos aprenden más cuando la ciencia y la educación está abiertas y reproducibles
- ▶ Yo creo que la mejor manera de aprender es enseñando
- ▶ Recuerda que nadie es re buena al principio, pero todas mejoramos con la práctica

Pensamiento concluyente

- ▶ Yo creo que todos aprenden más cuando la ciencia y la educación está abiertas y reproducibles
- ▶ Yo creo que la mejor manera de aprender es enseñando
- ▶ Recuerda que nadie es re buena al principio, pero todas mejoramos con la práctica
- ▶ Recuerda que vos podés hacer lo que quieras

¡Gracias por tu atención y ayuda! ¡Mantengámonos en contacto!

Email: rayna.harris@gmail.com

Twitter y GitHub y Instrgram y Speaker Deck: @raynamharris

Diapositivas acá¹⁷ y acá¹⁸

¹⁷<https://github.com/raynamharris/FMR1CA1rnaseq>

¹⁸<https://speakerdeck.com/raynamharris/>