

Usando y Enseñando R para Investigación Reproducible

Rayna M. Harris

Twitter: @raynamharris

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

27 Marzo 2018

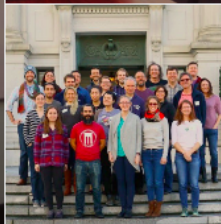
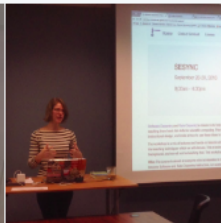
R-Ladies Buenos Aires

¿Quién soy?



Rayna Harris @raynamharris · 4 Feb 2017

I wear a lab coat to research microbes, molecules, & brains. I also teach & work w/ data scientists around the globe! #actuellivingscientist











¹<https://twitter.com/raynamharris>









Soy voluntaria en Software Carpentry



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

2

²**cowplot** <https://cran.r-project.org/web/packages/cowplot/index.html>

Mis mejores 8 deseos con consejos

- ▶ Pero antes de comenzar

Mis mejores 8 deseos con consejos

- ▶ Pero antes de comenzar
- ▶ Recuerda que vos podés hacer lo que quieras

Mis mejores 8 deseos con consejos

- ▶ Pero antes de comenzar
- ▶ Recuerda que vos podés hacer lo que quieras
- ▶ Recuerda que nadie es re buena al principio

Mis mejores 8 deseos con consejos

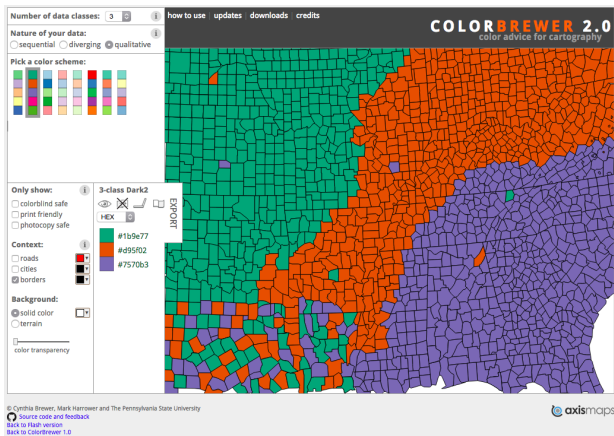
- ▶ Pero antes de comenzar
- ▶ Recuerda que vos podés hacer lo que quieras
- ▶ Recuerda que nadie es re buena al principio
- ▶ Yo creo que la mejor manera de aprender es enseñando

Mis mejores 8 deseos con consejos

- ▶ Pero antes de comenzar
- ▶ Recuerda que vos podés hacer lo que quieras
- ▶ Recuerda que nadie es re buena al principio
- ▶ Yo creo que la mejor manera de aprender es enseñando
- ▶ Yo creo que todos aprenden más cuando la ciencia y la educación están abiertas

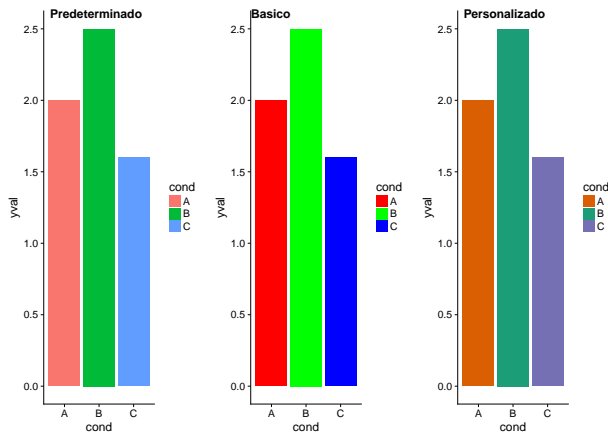
Deseo 1: Desarrolla tu propia paleta de colores

Colorbrewer³ te ayuda a elegir colores amigables para daltónicos



³<http://colorbrewer2.org/>

Ejemplos de paletas de colores



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

Convertir HEX a RGB si fuera necesario

#d95f02

rgb(217,95,2)

#1b9e77|

rgb(27,158,119)

#7570b3|

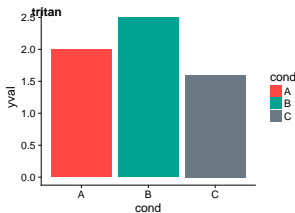
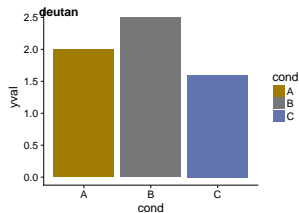
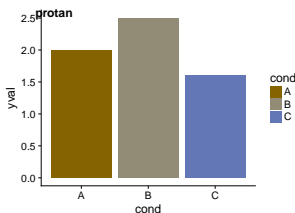
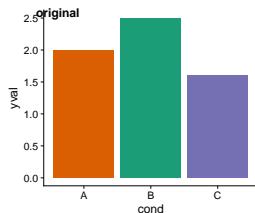
rgb(117,112,179)

4

⁴<https://www.webpagefx.com/web-design/hex-to-rgb/>

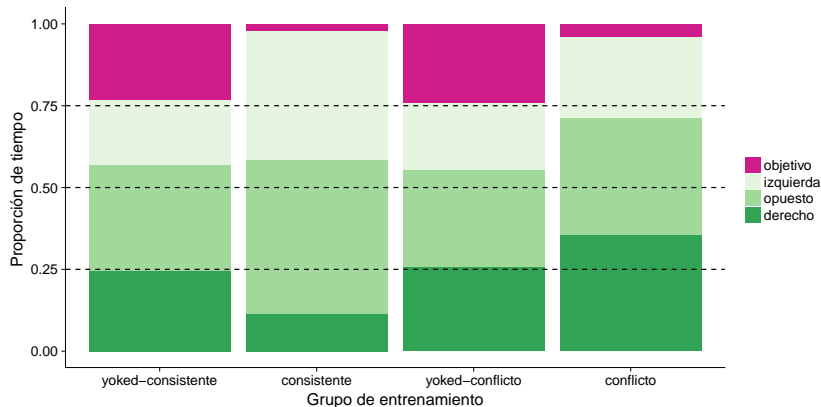
Confirmá que tu paleta es amigable con los daltónicos

colorblindr⁵ simula deficiencias de visión del color



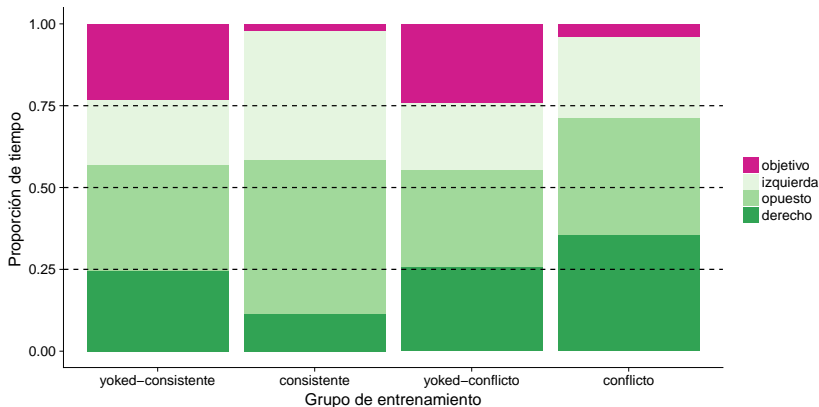
⁵<https://github.com/clauswilke/colorblindr>

Deseo 2: Usa leyendas gráficas



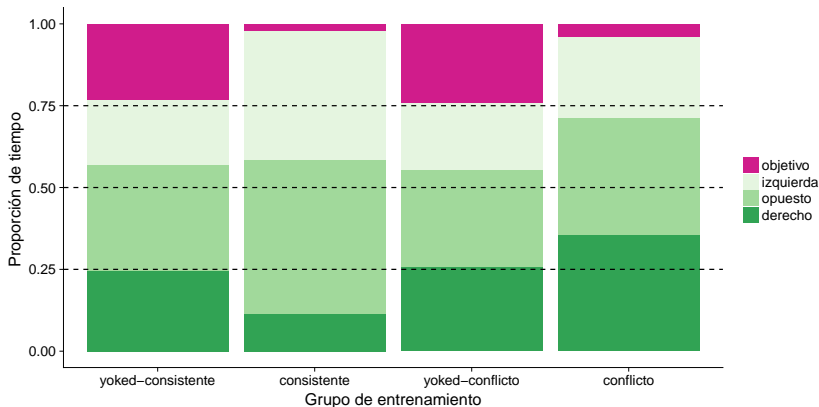
- ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?

Deseo 2: Usa leyendas gráficas



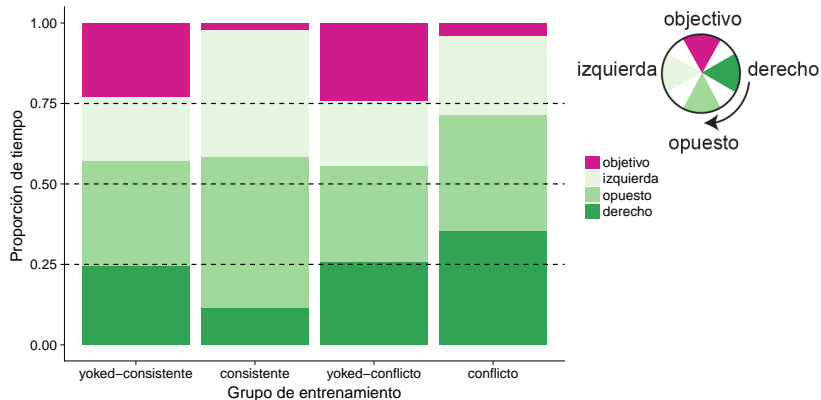
- ▶ ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?
- ▶ ¿Por qué el rosa “objetivo”?

Deseo 2: Usa leyendas gráficas



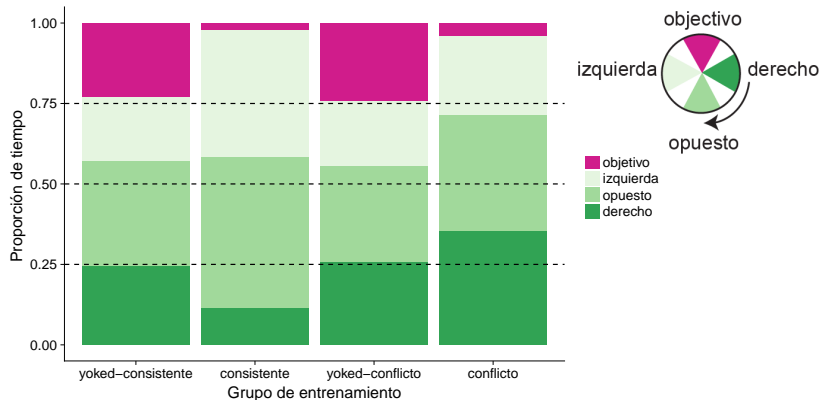
- ▶ ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?
- ▶ ¿Por qué el rosa “objetivo”?
- ▶ ¿Por qué hay líneas discontinuas?

Deseo 2: Usa leyendas gráficas



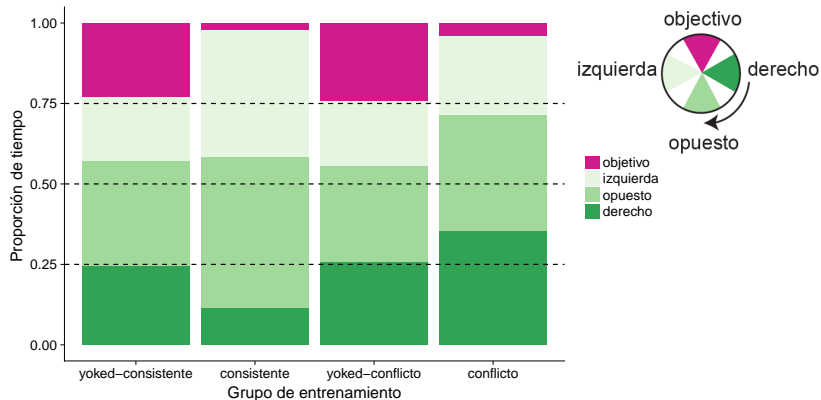
- “objetivo”, “izquierda”, etc. son cuadrantes de una arena

Deseo 2: Usa leyendas gráficas



- ▶ “objetivo”, “izquierda”, etc. son cuadrantes de una arena
- ▶ Yo quería mostrar un uso desproporcionado del espacio

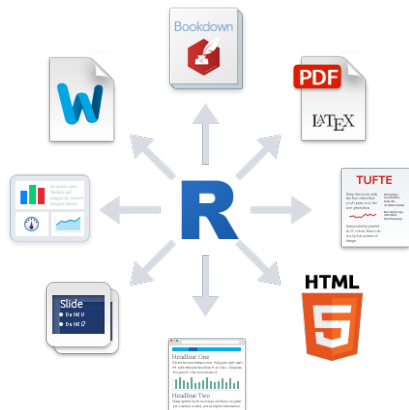
Deseo 2: Usa leyendas gráficas



- ▶ “objetivo”, “izquierda”, etc. son cuadrantes de una arena
- ▶ Yo quería mostrar un uso desproporcionado del espacio
- ▶ Use **cowplot**⁶ para agregar imágenes dentro de R

⁶**cowplot** <https://cran.r-project.org/web/packages/cowplot/index.html>

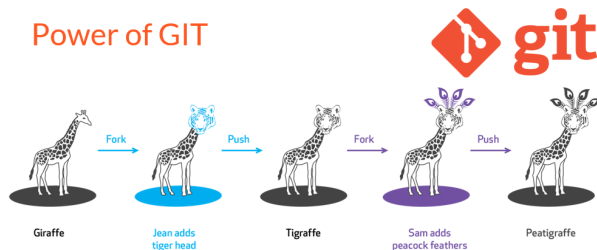
Deseo 3: Usa *R Markdown* para la reproducibilidad



7

⁷https://rmarkdown.rstudio.com/authoring_quick_tour.html

Deseo 4: Usa el control de versiones para colaborar con otros y con vos en el futuro

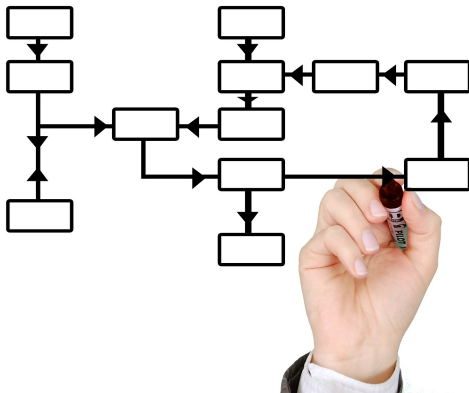


8

⁸<http://technetnepal.net/blogs/shirishamaharjan/archive/2017/05/07/expand-horizons-change-attitudes-git-and-github-workshop.aspx>

Deseo 5: Documenta tu flujo de trabajo

Porque probablemente sea único y complejo



Por ejemplo, podés enumerar los comandos por orden de operación

Branch: master FMR1CA1rnaseq / UNIXworkflow /

Create new file Upload files Find file History

raynamharris organization Latest commit 47e1c77 a minute ago

**

00_rawdata.md	behavior and multiqc	4 months ago
01_fastqc.md	multiqc	4 months ago
02_kallisto.md	multiqc ncbi	4 months ago
03_NCBIGEO.md	updated	3 months ago
README.md	organization	a minute ago

README.md

RNAseq workflow on TACC

The workflow

- **00_rawdata:** Download the data to scratch on Stampede with `00_gsaf_download.sh`.
- **01_fastqc:** Evaluate the quality of the reads using the program FastQC.
- **02_kallisto:** Quantify transcript-level expression using Kallisto
- **03_NCBIGEO.md:** Prep for submitting files to NCBI's Gene Expression Omnibus database

10

¹⁰<https://github.com/raynamharris/FMR1CA1rnaseq>

Probá múltiples estrategias de organización y hace lo que funcione mejor para vos

01_retention.md	with integration	2 months ago
02_RNAseq.Rmd	with integration	2 months ago
02_RNAseq.md	with integration	2 months ago
03_wgcn.R	wgcn and daytime	4 months ago
04_ephys.Rmd	ephys	3 months ago
04_ephys.md	ephys	3 months ago
05_Ceolin.Rmd	volcanos	3 months ago
05_Ceolin.md	three way venn	3 months ago
07_colorblindr.Rmd	touchups	3 months ago
07_colorblindr.md	touchups	3 months ago
08_results.Rmd	touchups	3 months ago
08_results.md	touchups	3 months ago
09_integration.Rmd	non DEGs	2 months ago
09_integration.md	retention	2 months ago
Fig1.Rmd	fig 2	3 months ago
Fig1.md	fig 2	3 months ago
Fig2.Rmd	ephys	3 months ago
Fig2.md	fig 4 5	3 months ago
Fig3.Rmd	fig 4 5	3 months ago
Fig3.md	fig 4 5	3 months ago
Fig4.Rmd	touchups	3 months ago

11

¹¹<https://github.com/raynamharris/FMR1CA1rnaseq>

Punto medio resumen

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

Deseo 6: 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 GitHub¹²

¹²<https://swcarpentry.github.io/r-novice-gapminder-es/>

Como podés ayudarme

- ▶ Leer y comentar o editar en GitHub¹²
- ▶ Participar en el **Bug BBQ**¹³ el Abril 11 y 12

¹²<https://swcarpentry.github.io/r-novice-gapminder-es/>

¹³<https://carpentries.github.io/2018-04-bug-bbq/>

Como podés ayudarme

- ▶ Leer y comentar o editar en GitHub¹²
- ▶ Participar en el **Bug BBQ**¹³ el Abril 11 y 12
- ▶ Hacer videos de vos leyendo y codificando junto con la lección¹⁴

¹²<https://swcarpentry.github.io/r-novice-gapminder-es/>

¹³<https://carpentries.github.io/2018-04-bug-bbq/>

¹⁴<https://www.youtube.com/watch?v=rQkfLaTdAvw>

Deseo 8: Organizar y/o asistir a un taller

Want to Know More?

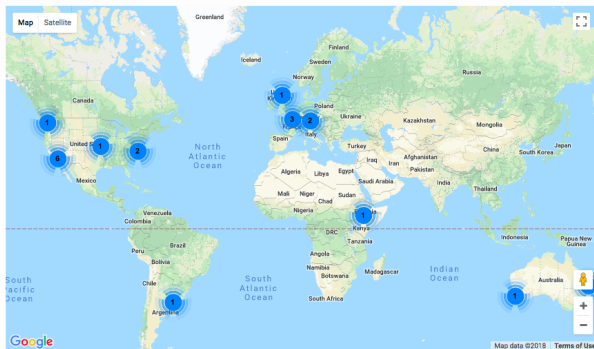
See our [FAQ](#).

Want to Run a Workshop?

See our [operations guide](#).

Want to Host a Workshop?

Please [fill in this form](#).



Future Workshops

16

¹⁶<https://software-carpentry.org/workshops/>

¡Gracias por tu atención! ¡Mantengámonos en contacto!

Twitter: @raynamharris

Email: rayna.harris@gmail.com