

Usando y Enseñando R para Investigación Reproducible

Rayna M. Harris

Twitter: @raynamharris

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

27 Marzo 2018

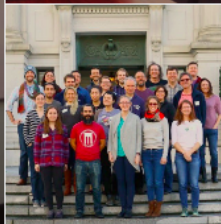
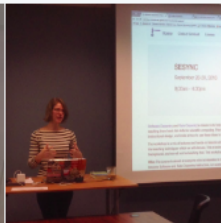
R-Ladies Buenos Aires

¿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! #actuallylivingscientist











¹<https://twitter.com/raynamharris>









Soy voluntaria de 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

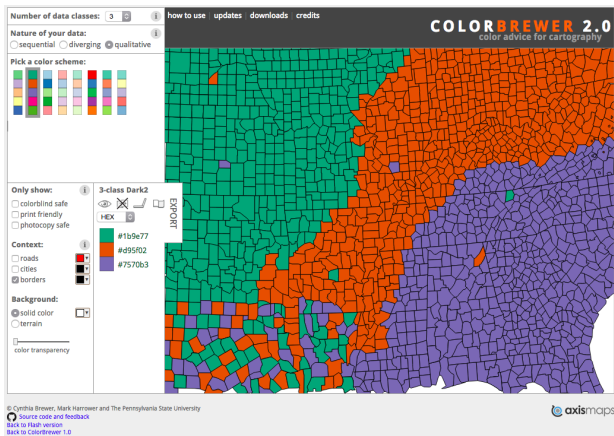
- ▶ Pero antes de comenzar
- ▶ Recuerda que vos podés hacer lo que quieras
- ▶ Recuerda que nadie es re buena al principio
- ▶ Yo creo que la mejor manera de aprender es a enseñar

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 a enseñar
- ▶ Yo creo que todos aprenden más cuando la ciencia y la educación están abiertas

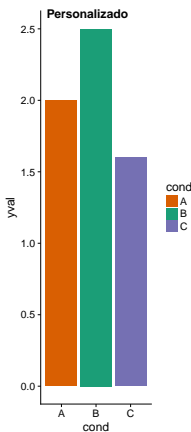
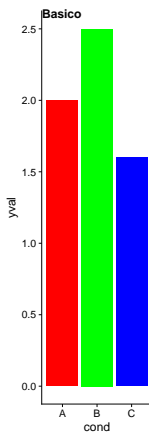
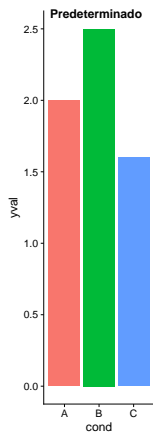
Deseo 1: Desarrolla tu propia paleta de colores

Colorbrewer³ te ayuda a elegir colores amigables para el daltónico



³<http://colorbrewer2.org/>

Ejemplos de paleta 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 necesita

#d95f02

rgb(217,95,2)

#1b9e77|

rgb(27,158,119)

#7570b3|

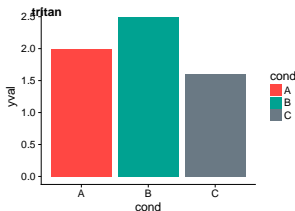
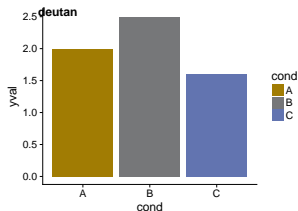
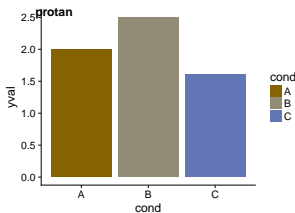
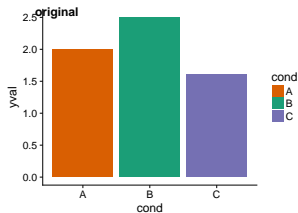
rgb(117,112,179)

4

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

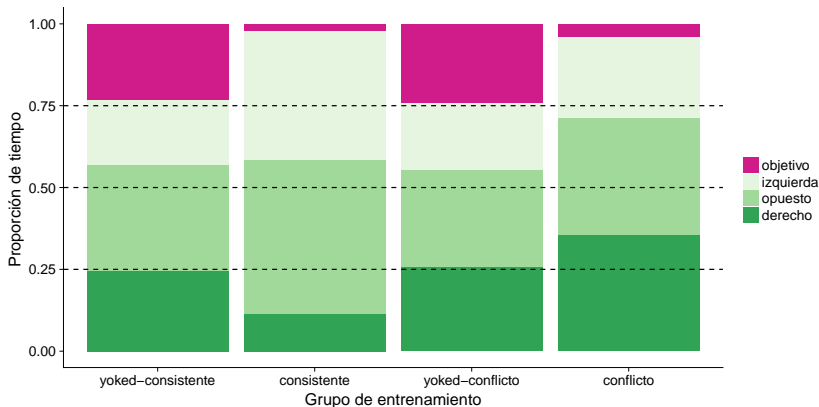
Confirme que tu palet es amigable con los daltónicos

colorblindr⁵ simula deficiencias de visión del color



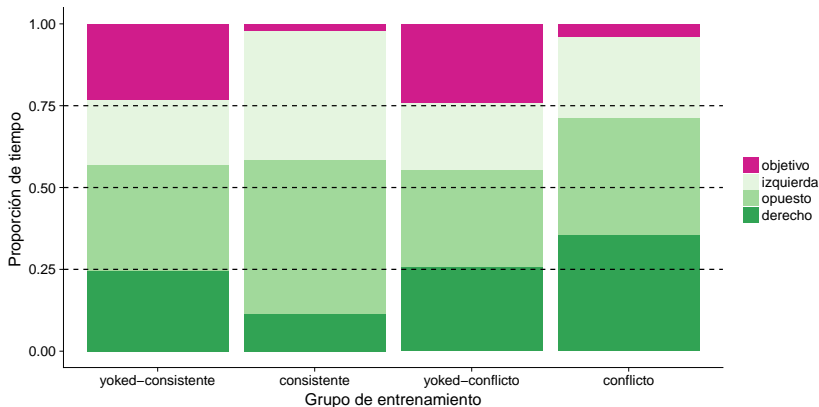
⁵<https://github.com/clauswilke/colorblindr>

Deseo 2: Usa leyendas graficas



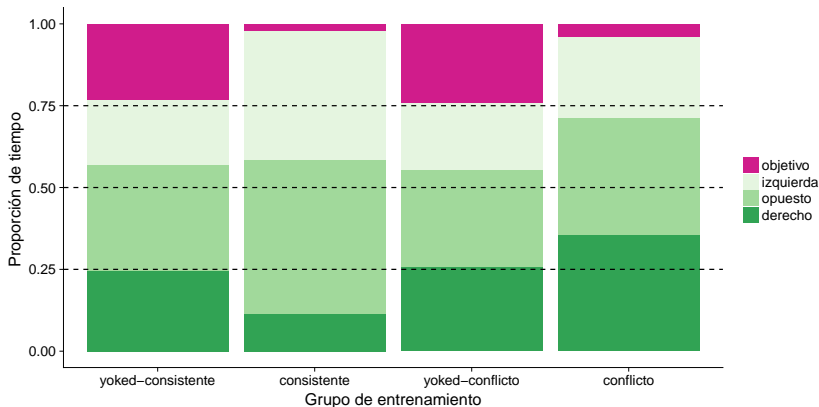
- ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?

Deseo 2: Usa leyendas graficas



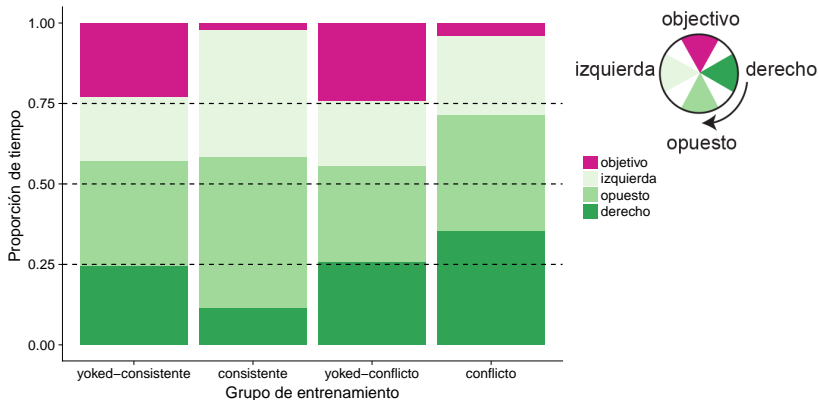
- ▶ ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?
- ▶ ¿Por qué el rosa “objetivo”?

Deseo 2: Usa leyendas graficas



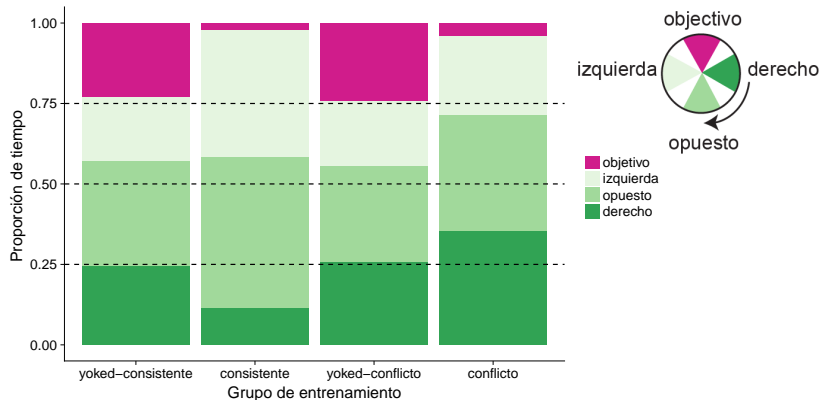
- ▶ ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?
- ▶ ¿Por qué el rosa “objetivo”?
- ▶ ¿Por qué hay líneas discontinuas?

Deseo 2: Usa leyendas graficas



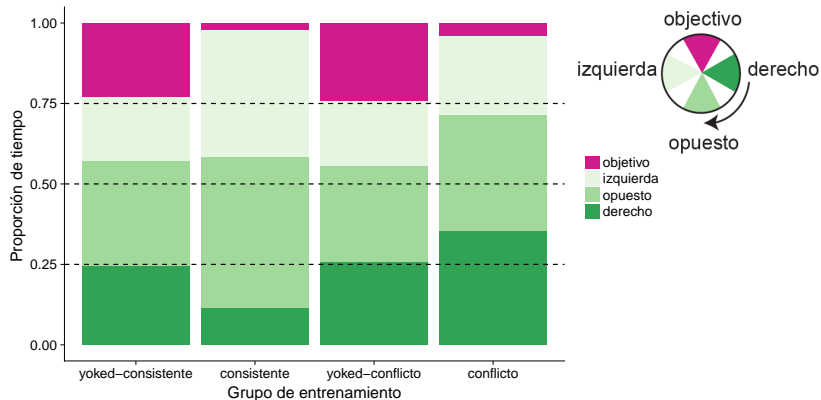
- “objetivo”, “izquierda”, etc. son cuadrantes de una arena

Deseo 2: Usa leyendas graficas



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

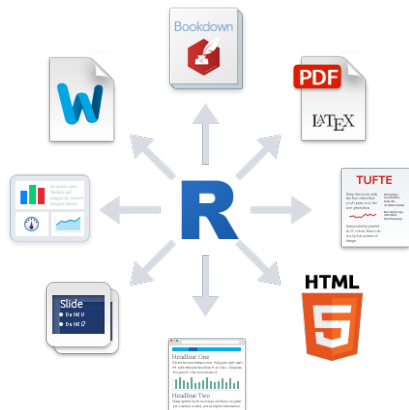
Deseo 2: Usa leyendas graficas



- ▶ “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>

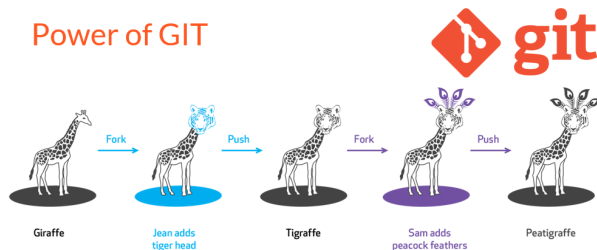
Deseo 3: Usa *R Markdown* para la reproducibilidad



7

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

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

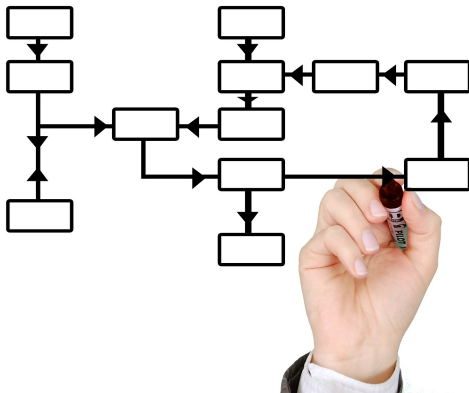


8

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

Deseo 5: Documenta su flujo de trabajo

Porque probablemente sea único y complejo



Por ejemplo, puede 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_NCBI GEO.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_NCBI GEO.md:** Prep for submitting files to NCBI's Gene Expression Omnibus database

10

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

Pruebe múltiples estrategias de organización y haga 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 graficas
- ▶ Deseo 3: Usa *R Markdown* para la reproducibilidad
- ▶ Deseo 4: Usa el control de versiones para la colaboración
- ▶ Deseo 5: Documenta su flujo de trabajo

Deseo 6: Me ayuda mejor las nuevas lecciones en español de Software Carpentry

R for Reproducible Scientific Analysis

Introduction to R for non-programmers using gapminder data.

The goal of this lesson is to teach novice programmers to write modular code and best practices for using R for data analysis. R is commonly used in many scientific disciplines for statistical analysis and its array of third-party packages. We find that many scientists who come to Software Carpentry workshops use R and want to learn more. The emphasis of these materials is to give attendees a strong foundation in the fundamentals of R, and to teach best practices for scientific computing: breaking down analyses into modular units, task automation, and encapsulation.

Note that this workshop will focus on teaching the fundamentals of the programming language R, and will not teach statistical analysis.

A variety of third party packages are used throughout this workshop. These are not necessarily the best, nor are they comprehensive, but they are packages we find useful, and have been chosen primarily for their usability.

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.

Schedule

	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 puedo gestionar mis proyectos en R?
01:25	3. buscando ayuda	¿Cómo puedo obtener ayuda en R?
01:45	4. Estructuras de datos	¿Cómo puedo leer datos en R? ¿Cuáles son los tipos de datos básicos en R? ¿Cómo represento la información categórica en R?

Como podes ayudarme

- ▶ Leer y comentar o editar en GitHub¹²

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

Como podes 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 podes ayudarme

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

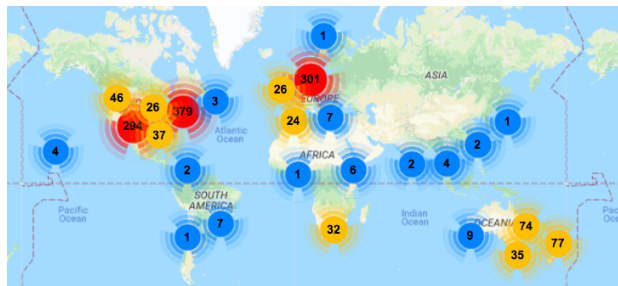
¹²<https://swcarpentry.github.io/r-novice-gapminder-es/>

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

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

Deseo 7: Convertirse en una instructor certificada

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



15

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

Deseo 8: Organizar y / o asistir a un taller

Want to Know More?

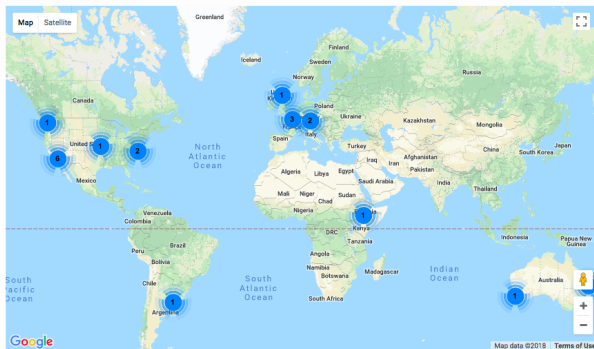
See our [FAQ](#).

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! ¡Mantengamonos en contacto!

Twitter: @raynamharris

Email: rayna.harris@gmail.com