

# Usando y Enseñando R para Investigación Reproducible

Rayna M. Harris

Twitter: @raynamharris

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

27 Marzo 2018

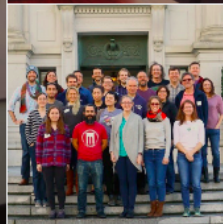
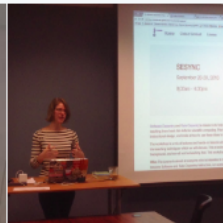
R-Ladies Buenos Aires

# ¿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











<sup>1</sup><https://twitter.com/raynamharris>









# Soy voluntaria de Software Carpentry



## Our lessons in English

| Lesson                   | Site  | Repository  | Reference   | Instructor Guide  | Maintainer(s)   |
|--------------------------|---|---|---|---|---|
| The Unix Shell           |  |  |  |  | <a href="#">Gabriel Devenyi</a> , <a href="#">Ashwin Srinath</a> , <a href="#">Colin Morris</a> , <a href="#">Will Pitchers</a> |
| Version Control with Git |  |  |  |  | <a href="#">Ivan Gonzalez</a> , <a href="#">Daisie Huang</a> , <a href="#">Nima Hejazi</a> , <a href="#">Katherine Koziar</a>   |

## Our lessons in Spanish

| Lesson                       | Site  | Repository  | Reference   | Instructor Guide  | Maintainer(s)                   |
|------------------------------|---|---|---|---|---------------------------------|
| La Terminal de Unix          |  |  |  |  | <a href="#">Heladia Saldago</a> |
| Control de versiones con Git |  |  |  |  | <a href="#">Rayna M. Harris</a> |

2

<sup>2</sup><https://software-carpentry.org/lessons/>

# Mis 10 mejores deseos (en lugar de mis 10 mejores consejos)



# Mis 10 mejores deseos (en lugar de mis 10 mejores consejos)



- ▶ Recuerda: Vos podés hacer lo que quieras

# Mis 10 mejores deseos (en lugar de mis 10 mejores consejos)

- ▶
- ▶ Recuerda: Vos podés hacer lo que quieras
- ▶ Recuerda: Nadie es re buena al principio

# Mis 10 mejores deseos (en lugar de mis 10 mejores consejos)

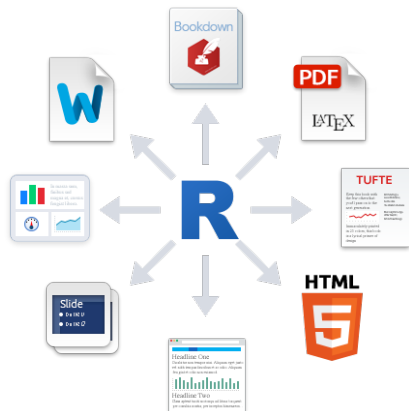
- ▶
- ▶ Recuerda: Vos podés hacer lo que quieras
- ▶ Recuerda: Nadie es re buena al principio
- ▶ Yo creo que la mejor manera de aprender es a enseñar

# Mis 10 mejores deseos (en lugar de mis 10 mejores consejos)

- ▶
- ▶ Recuerda: Vos podés hacer lo que quieras
- ▶ Recuerda: 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: Usa *R Markdown* para la reproducibilidad

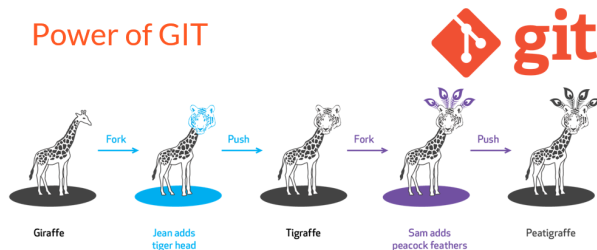


3

---

<sup>3</sup>[https://rmarkdown.rstudio.com/authoring\\_quick\\_tour.html](https://rmarkdown.rstudio.com/authoring_quick_tour.html)

## Deseo 2: Usa el control de versiones para la colaboración



4

<sup>4</sup><http://technetnepal.net/blogs/shirishamaharjan/archive/2017/05/07/expand-horizons-change-attitudes-git-and-github-workshop.aspx>



# 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_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

5

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

# Pruebe múltiples estrategias de organización y haga lo que funcione mejor para vos

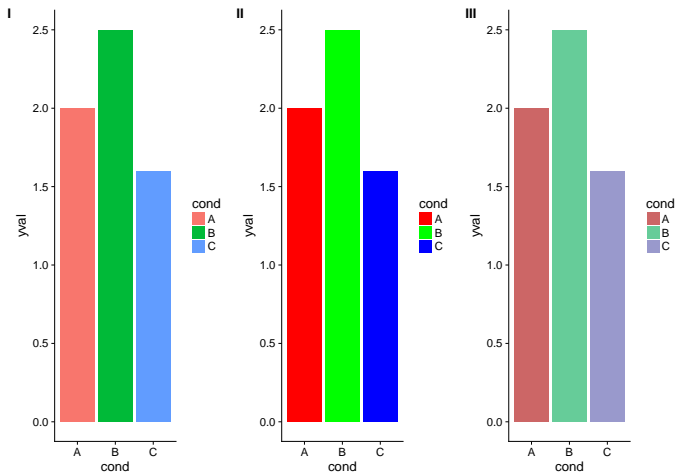
|                    |                   |              |
|--------------------|-------------------|--------------|
| 01_retention.md    | with integration  | 2 months ago |
| 02_RNAseq.Rmd      | with integration  | 2 months ago |
| 02_RNAseq.md       | with integration  | 2 months ago |
| 03_wgona.R         | wgona 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 |

6

---

<sup>6</sup><https://github.com/raynamharris/FMR1CA1rnaseq>

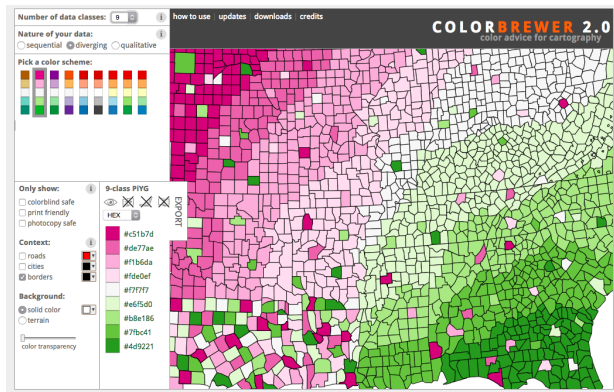
## Deseo 4: Desarrolla tu propia paleta de colores



- ▶ I: predeterminado
- ▶ II: `values=c("red", "green", "blue")`
- ▶ III: `values=c("#CC6666", "#66CC99", "#9999CC")`

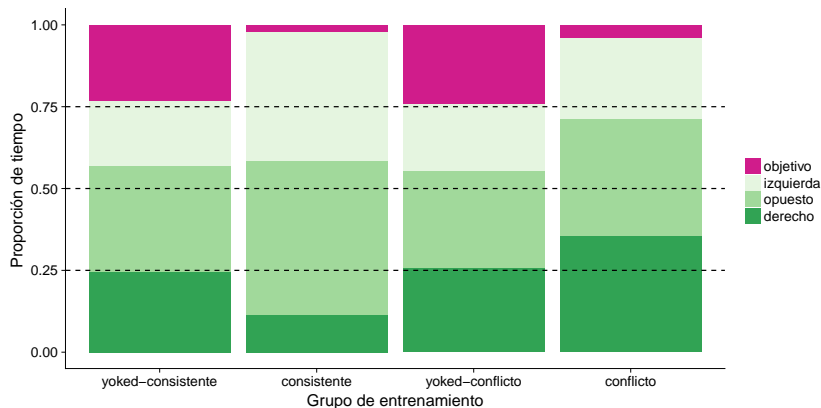
# Deseo 4: Desarrolla tu propia paleta de colores

Colorbrewer<sup>7</sup> te ayuda a elegir colores amigables para el daltónico



<sup>7</sup><http://colorbrewer2.org/>

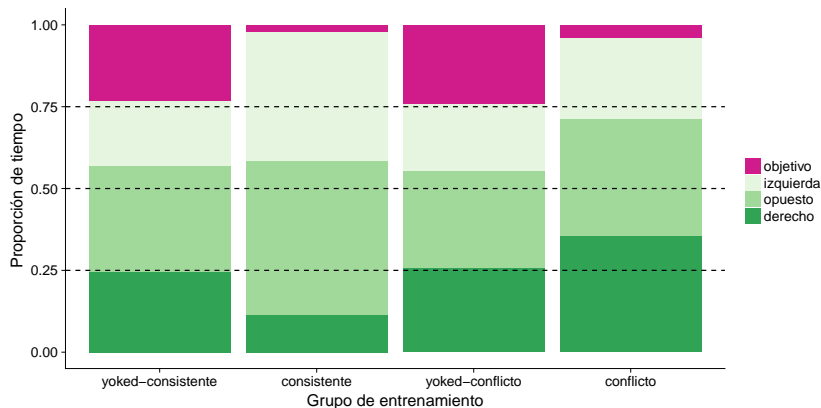
## Deseo 5: Usa leyendas graficas



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

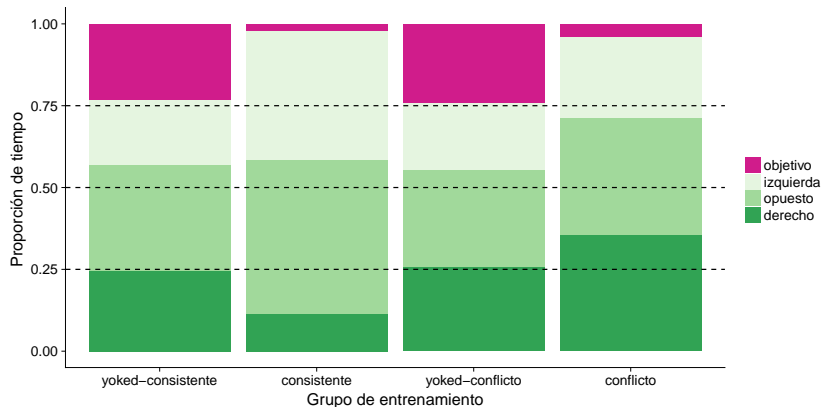


## Deseo 5: Usa leyendas graficas



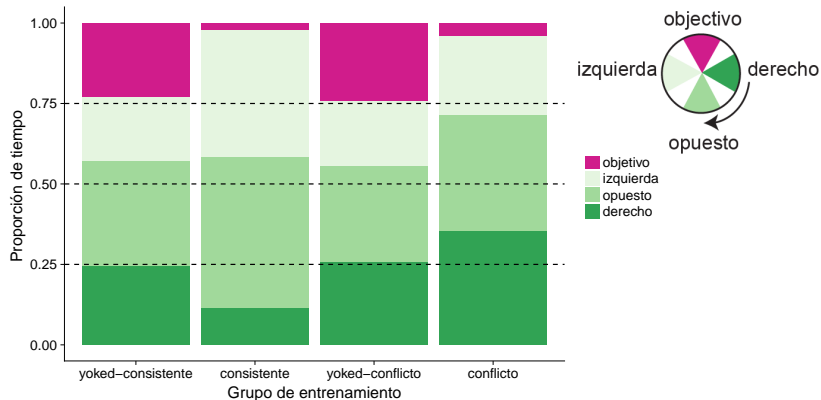
- ▶ ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?
- ▶ ¿Por qué el rosa “objetivo”?

# Deseo 5: Usa leyendas graficas



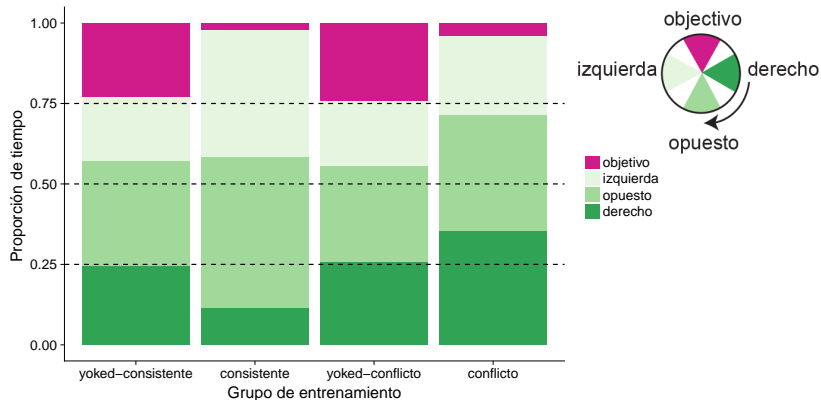
- ▶ ¿Qué significa “objetivo”, “izquierda”, “opuesto” y “derecho”?
- ▶ ¿Por qué el rosa “objetivo”?
- ▶ ¿Por qué hay líneas discontinuas?

## Deseo 5: Usa leyendas graficas



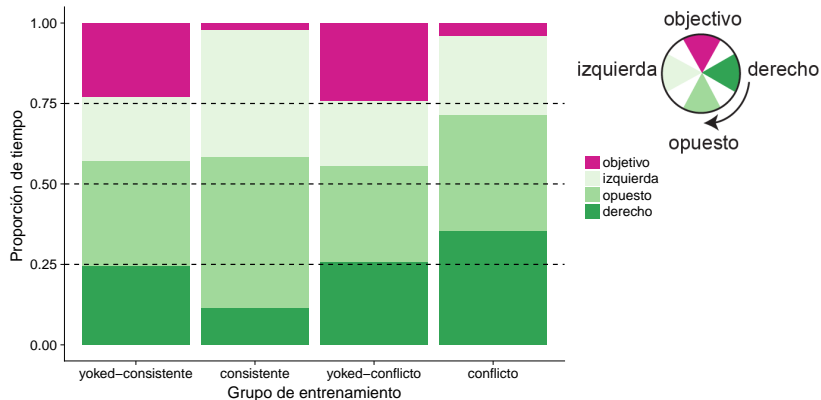
- “objetivo”, “izquierda”, etc. son cuadrantes de una arena

## Deseo 5: Usa leyendas graficas



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

## Deseo 5: Usa leyendas graficas



- ▶ “objetivo”, “izquierda”, etc. son cuadrantes de una arena
- ▶ Yo quería mostrar un uso desproporcionado del espacio
- ▶ **cowplot**<sup>8</sup> es genial!

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

# En resumen

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

# Deseo 6: Me ayuda mejor las lecciones 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?<br>¿Cómo interactuar con R?<br>¿Cómo administrar tu entorno?<br>¿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?<br>¿Cuáles son los tipos de datos básicos en R?<br>¿Cómo represento la información categórica en R? |

# Como puedes ayudar

- ▶ Leer y comentar o editar en GitHub



# Como puedes ayudar

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

# Como puedes ayudar

- ▶ Leer y comentar o editar en GitHub
- ▶ Participar en el **Bug BBQ** el Abril 11 y 12
- ▶ Organizar un taller o reunion para usarla

# Como puedes ayudar

- ▶ Leer y comentar o editar en GitHub
- ▶ Participar en el **Bug BBQ** el Abril 11 y 12
- ▶ Organizar un taller o reunion para usarla
- ▶ Haga videos de usted leyendo y codificando junto con la lección

## Deseo 7: Convertirse en una instructor certificada

- ▶ Aplicar aquí: <http://carpentries.github.io/instructor-training/>

**¡Gracias por su atención! ¡Mantengamonos en contacto!**

Twitter: @raynamharris Email: rayna.harris@gmail.com