Taller de introducción a R

R-Ladies Bilbao

18.00-20.00h
Martes
11-18-25
febrero 2020
Wikitoki

b1lba0
data
lab

bilbaodatalab.wikitoki.org

Taller de introducción a R

Lugar: **Wikitoki**
http://wikitoki.org/

Organizado por: **Bilbao Data Lab**
http://bilbaodatalab.wikitoki.org/

Impartido por: **Rladies Bilbao**
https://www.meetup.com/es-ES/rladies-bilbao/

Necesidades: **R y Rstudio instalados**
https://cran.r-project.org/
https://rstudio.com/products/rstudio/download/

Material:
http://bilbaodatalab.wikitoki.org/actividad/taller-intro-r-2020/

Sesión 1: Introducción y operaciones básicas (18:00-20:00, 11 febrero 2020)

1. Intro a R y Rstudio
– Qué es R
– Qué es RStudio
2. Operaciones y conceptos básicos de R

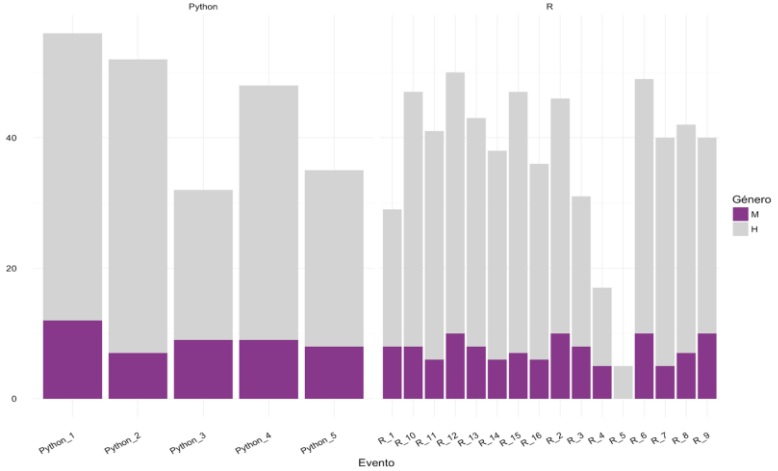
Sesión 2: Importar, gestionar y primeros gráficos (18:00-20:00, 18 febrero 2020)

3. Importar exportar datos
4. Procesar, gestionar datos
5. Gráficos simples

Posibilidad de trabajar con vuestros datos

Sesión 3: Más gráficos y algo de programación (18:00-20:00, 25 febrero 2020)

6. (más) Gráficos (ggplot2 etc.)
7. Programación (if else, for, while, apply..)
8. Dudas/Preguntas



Principales MeetUps de programación en Madrid

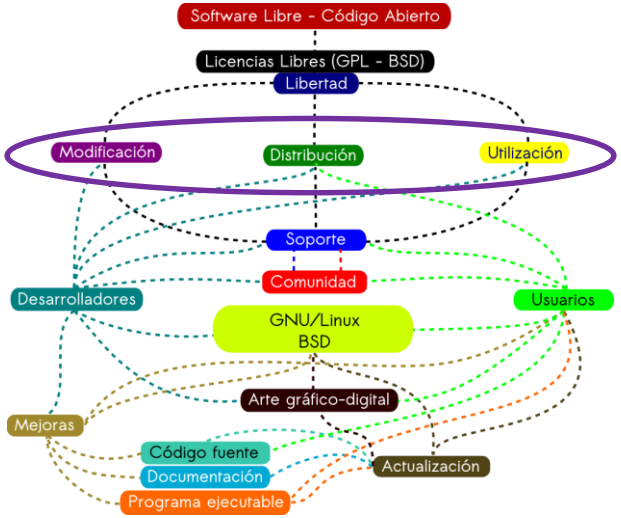
R Ladies

Organización mundial que **promueve la diversidad de género** en la comunidad de **R** mediante **encuentros y mentorías** en un entorno **amigable** y **seguro.**

https://rladies.org/

¿Qué es R?

Lenguaje de **programación** de **código abierto y gratuito** con énfasis en estadística y gráficos.

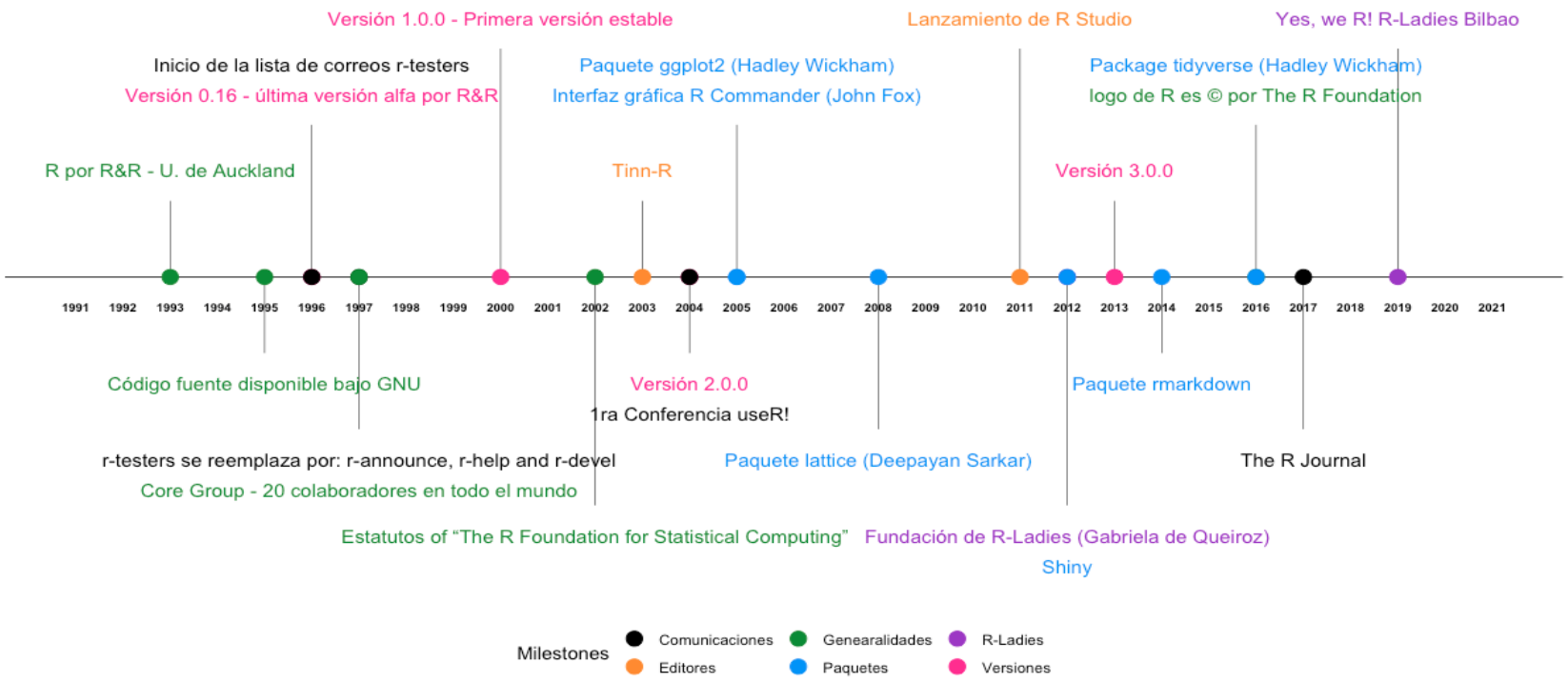


R compila y se ejecuta en en una amplia variedad de plataformas UNIX, Linux, FreeBSD, Windows y MacOs.

https://www.r-project.org/

https://es.wikipedia.org/wiki/Código_abierto

Casi 3 décadas con R...



Creada en R! Thanks to Ben Alex Keen

http://benalexkeen.com/creating-a-timeline-graphic-using-r-and-ggplot2/

¿Por qué utilizar R?

- R es un software libre, gratis para todo el mundo.
- Facilita desarollar trabajos reproducibles.
- R es una herramienta poderosa que puede manejar muchos tipos de datos.
- R va aumentando su popularidad cada vez más.
- Otros usuarios de R te pueden ayudar a resolver tus dudas:
  - StackOverflow
  - R-Bloggers
  - Twitter: #rstats
- R es flexible y puedes usar o añadir tus propios **paquetes**.

Fuente: https://www.r-bloggers.com/why-use-r-five-reasons/

Paquetes de R

- Las funciones básicas que por defecto facilita R se pueden **ampliar** mediante paquetes.

- Cada paquete es una **colección de funciones** diseñadas para atender una tarea específica (manejar bases de datos, análisis espacial, creación de mapas, visualización, métodos estadísticos específicos…).

- Los paquetes **se instalan una única vez** en el equipo, la primera vez que se necesitan, y luego **se cargan cada vez** que van a ser usados.

- Los paquetes están alojados en **REPOSITORIOS** desde los cuales podemos descargarlos. Los repositorios mas populares de paquetes R son:

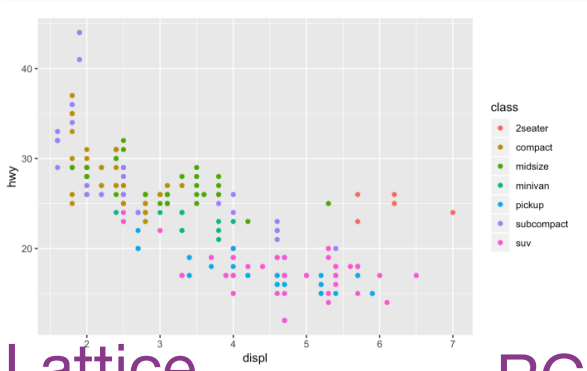
CRAN: el oficial**,** coordinado por la fundación R. Para que un paquete pueda ser publicado en el mismo, necesita **pasar por diferentes pruebas** para asegurar que el paquete cumple con las políticas de CRAN.

Bioconductor: repositorio específico para **bioinformática**. Como CRAN, tiene sus propias políticas de publicaciones y procesos de revisión

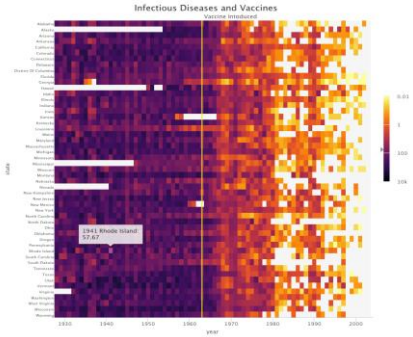
Github: Su popularidad procede del espacio ilimitado que proporciona para el alojamiento de proyectos open source, la integración con git (un software de control de versiones) y, la facilidad de compartir y colaborar con otras personas. No proporciona procesos de control.

Paquetes de visualización

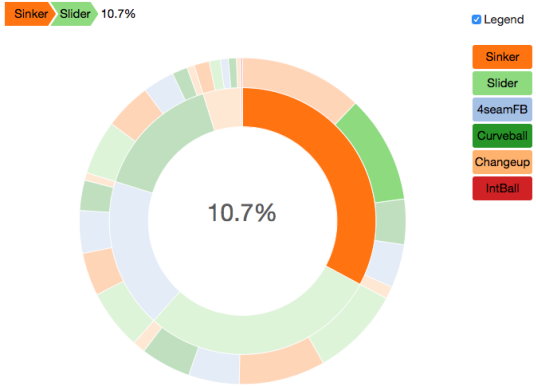
ggplot2
Hadley Wickham



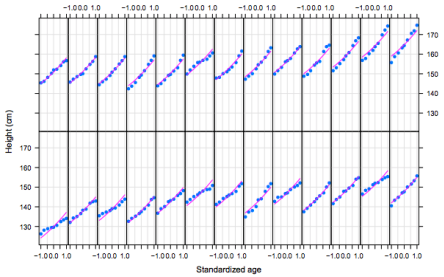
highcharter
Joshua Kunst



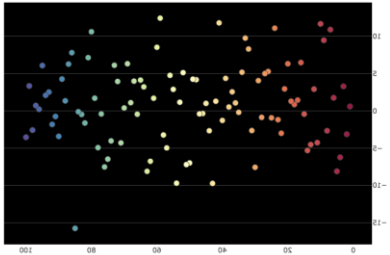
sunbustR
Kent Russell, Kerry Rodden, Mike Bostock, Kevin Warne



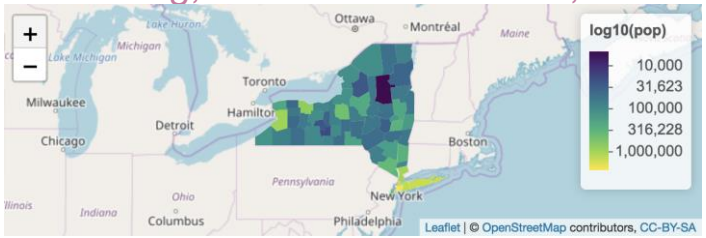
Lattice
Deepayan Sarkar



RColorBrewer
Erich Neuwirth



Leaflet
Joe Cheng, Bhaskar Karambelkar, Yihui Xie



Fuente: https://mode.com/blog/r-data-visualization-packages

Web Apps con Shiny

Fuente: https://omaymas.shinyapps.io/Exploring_Climate_Change_1900-2014/

Climate Change in Major Cities (1900-2014)

Type/Select one or more cities:

City

Madrid Bangkok

Type/Select one or more months:

Month

Jan Apr Jul Oct

Select type of plot:

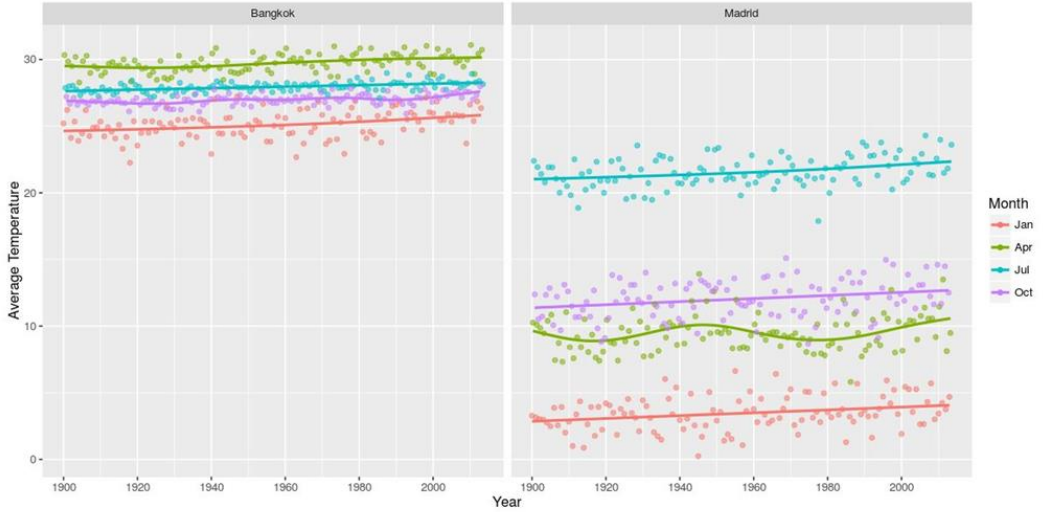
Plots

- [x] GAM Plot
- [x] Point Plot

Note that : "Early data was collected by technicians using mercury thermometers, where any variation in the visit time impacted measurements. In the 1940s, the construction of airports caused many weather stations to be moved. In the 1980s, there was a move to electronic thermometers that are said to have a cooling bias."

Source

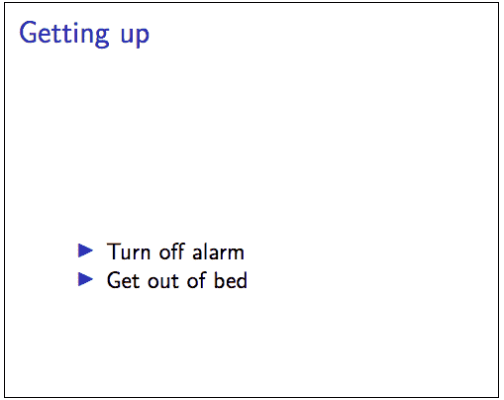
This Shiny App provides a fast and easy way to explore the "Earth Surface Temperature Data" published on Kaggle and compiled originally by "Berkeley Earth". Here you can plot the average temprature over time from 1900-2014 in 100 major cities. The model used is a Generalized Additive Model(GAM)



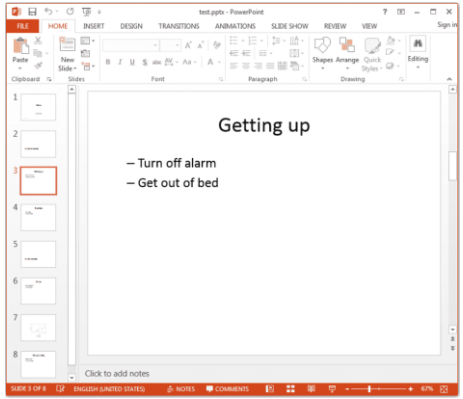
Más aplicaciones de este tipo en "Show Me Shiny": https://www.showmeshiny.com/

Presentaciones con R-Markdown

Beamer

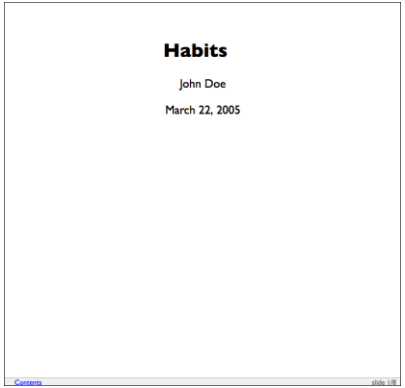


PowerPoint



ioslides

Slidy



Fuente: https://rmarkdown.rstudio.com/lesson-11.html

Artículos científicos con R-Markdown

Fuente: https://sites.google.com/site/rodriguezsanchezf/news/writingpapersinrmarkdown

A template for writing manuscripts in Rmarkdown

Jeffrey W. Hollister [1,2] Francisco Rodriguez-Sanchez [* 3]

[1] *US Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Atlantic Ecology Division, 27 Tarzwell Drive Narragansett, RI, 02882, USA*

[2] *Big Name University, Department of R, City, BN, 01020, USA*

[3] *Estacion Biologica de Donana (EBD-CSIC), E-41092 Sevilla, Spain*

[*] *corresponding author: f.rodriguez.sanc@gmail.com*

Write your abstract here.

Keywords: rmarkdown, reproducible science

INTRODUCTION

Write your introduction here. You can cite bibliography like this (Yan and Gerstein 2011, Sutherland et al. 2011), if you provide a `BibTeX` file with references. See http://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html for more information. Or you could also use knitcitations or RefManageR to fetch bibliographic metadata automatically from the web. For example, citing a paper can be as easy as providing its DOI (Clark and Gelfand 2006) or even just a few keywords (Ricklefs 2008). They will then automagically appear in the list of cited references.

You can even specifiy the desired output format for your bibliography by including a style file for a specific journal (e.g. "ecology.csl"). Many different bibliography styles (CSL files) can be obtained at http://citationstyles.org/ or https://github.com/citation-style-language/styles.

Autor del template de la imagen: Jeffrey W. Hollister
https://github.com/jhollist/manuscriptPackage

Más templates de este estilo en el paquete de R rmdTemplates (autor: Francisco Rodríguez-Sanchez). *Disponible en:*
https://github.com/Pakillo/rmdTemplates

Libros y páginas web con R-Markdown

Libros

Paquete: bookdown
Autor: Yihui Xie

Fuente: https://bookdown.org/yihui/bookdown/

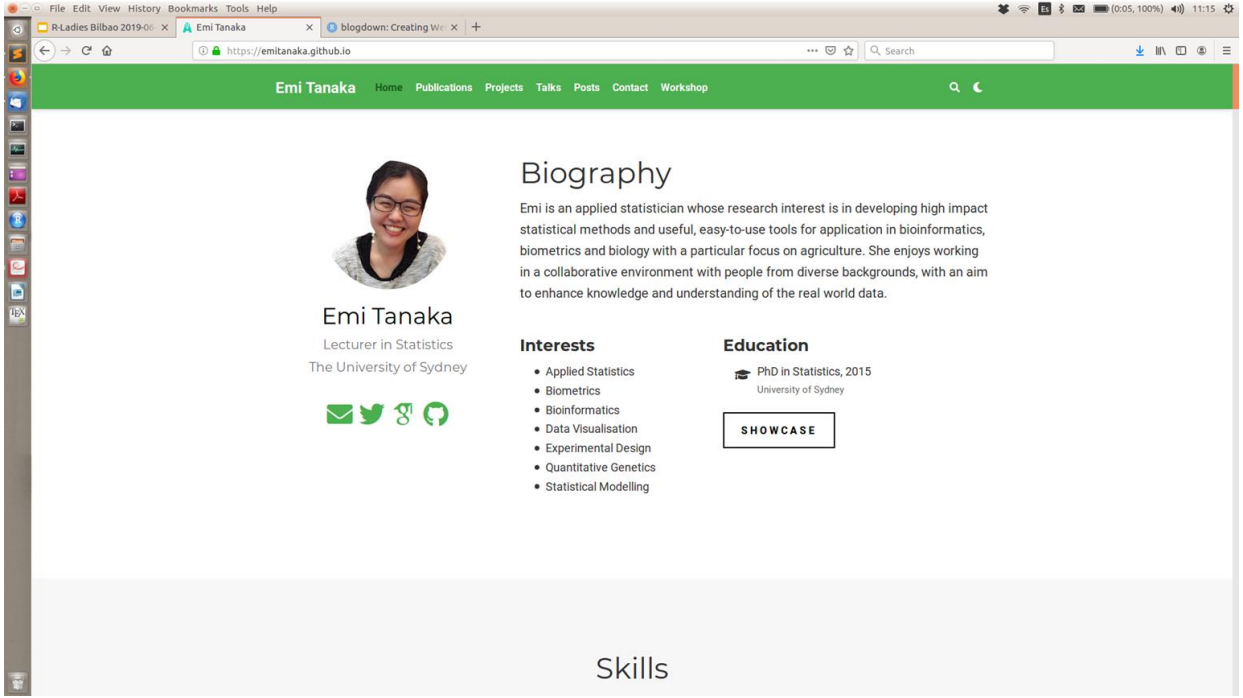
Páginas web

Paquete: blogdown
Autores: Yihui Xie, Amber Thomas, Alison Presmanes Hill

Fuente: https://bookdown.org/yihui/blogdown/

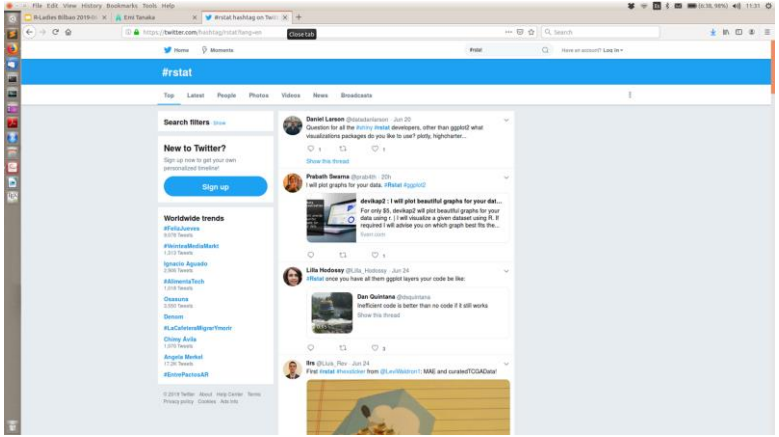
Páginas web con R-Markdown

https://emitanaka.github.io/



Twitter desde R

- **#rstats**



- Paquete para acceder a la API de Twitter:
  Paquete: twitteR
  Autor: Jeff Gentry

- Paquete para emojis:
  Paquete: emojifont
  Autor: Guangchuang Yu

- GIFs:

☆☆☆☆☆☆☆☆

Danielle Navarro
https://compcogscisydney.org/rstats/

Programación orientada a objetos

Todo en R, variables, datos, funciones, resultados, etc., se guarda en forma de objetos con el nombre que les asignemos.
Clases principals:

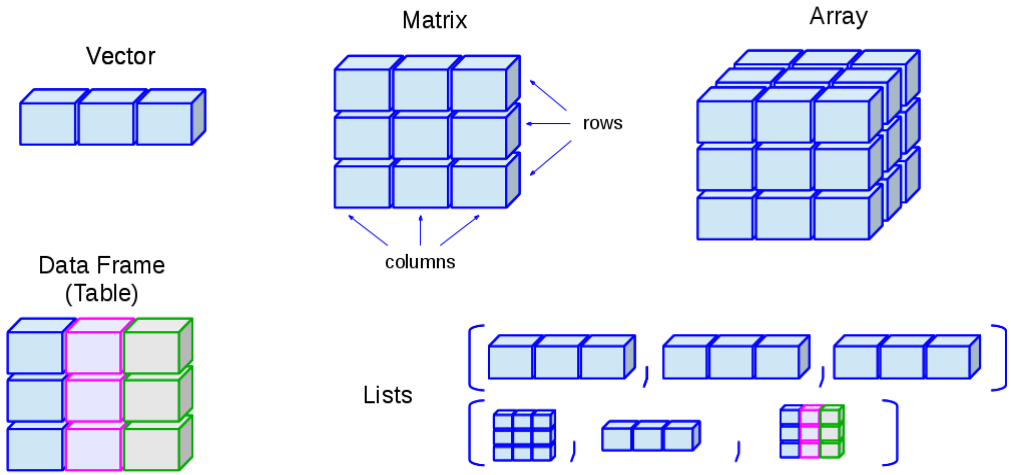
Vectores: una dimension.
Tipos: Numerico, Logico, texo, factor,…

Matrices: dos simensiones (datos del mismo tipo).

Arrays: más de dos dimensiones (con datos del mismo tipo).

Data Frames: Generalización las matrices, donde se pueden almacenar diferentes tipos de datos.

Listas: coleccion de objetos de diferente clase.



Funciones: objetos que permiten hacer operaciones sobre otros objetos (ejemplo: sum)

Expresiones, Salidas de funciones, …

http://venus.ifca.unican.es/Rintro/dataStruct.html

¿Qué es RStudio?

R Studio es un **IDE** (entorno de desarrollo integrado) para el **lenguaje de programación R**.

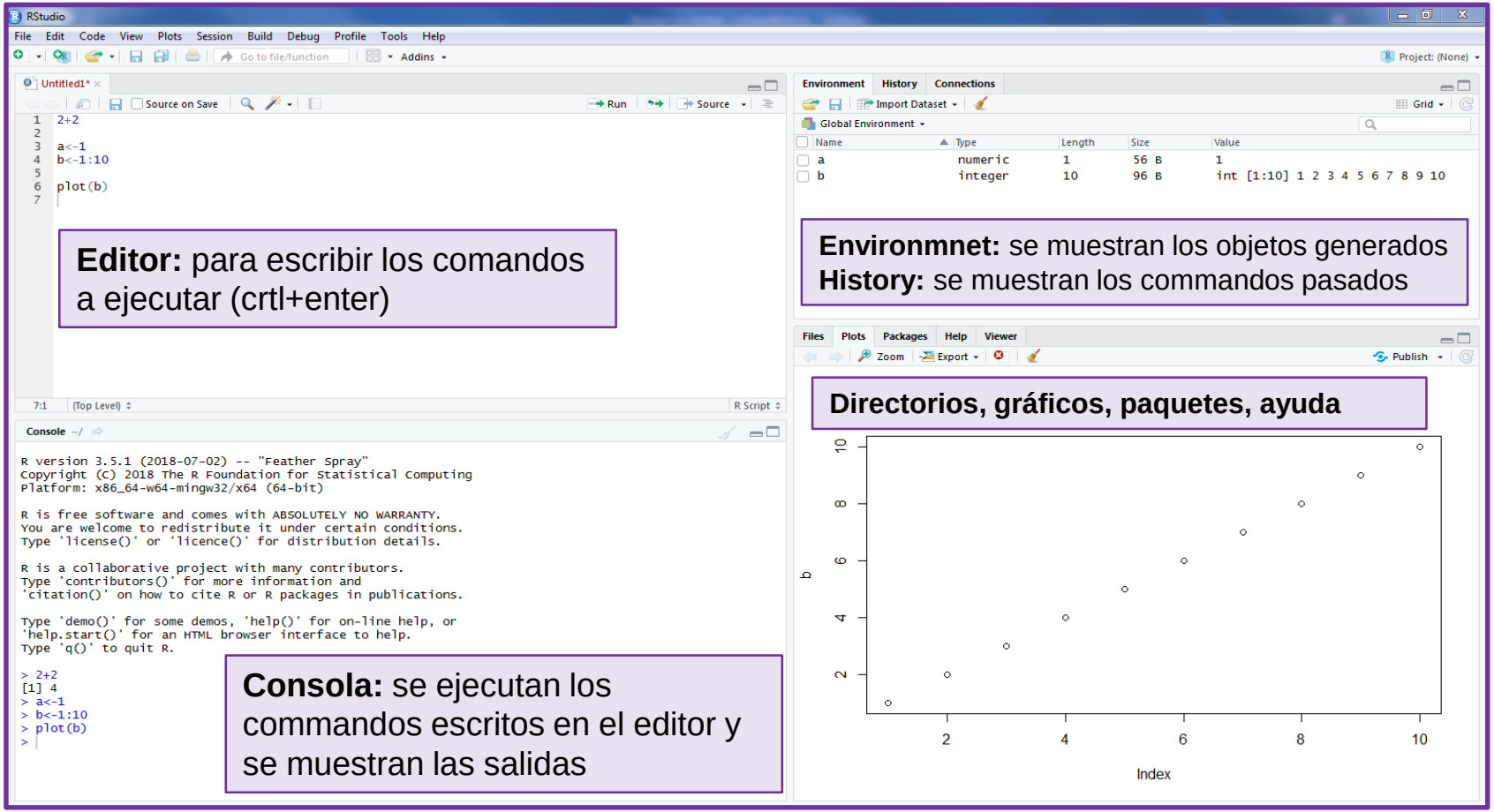
Es una **interfaz** que permite usar R de manera mas **cómoda** en algunos aspectos.

Está disponible para Windows, Mac y Linux

Existen otros programas o editores como TinnR, RKWard,… pero RStudio es el que **más utilizado**.

https://rstudio.com

RStudio: 4 paneles



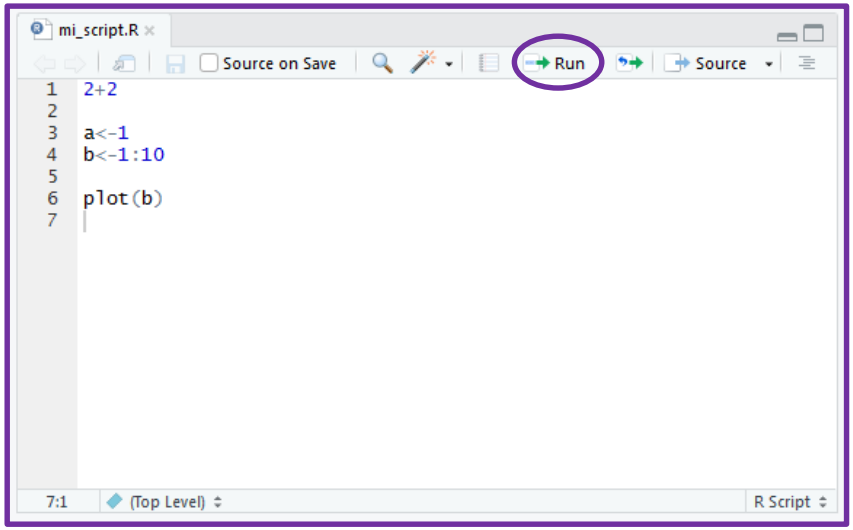
RStudio: Editor

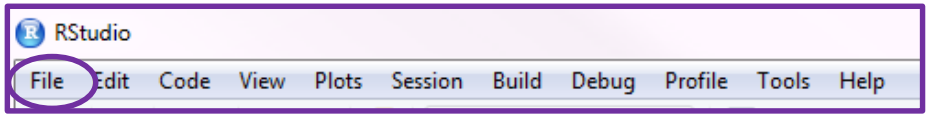
Al conjunto de comandos que escribimos en el editor se le llama **SCRIPT**.

Para ejecutar una línea clicamos en *Run* o pulsamos ctrl+enter.

Los scripts de R se guardan como archivos con extensión **.R**.

Para guardarlo usaremos la barra superior de Rstudio (File -> Save as…) o el icono





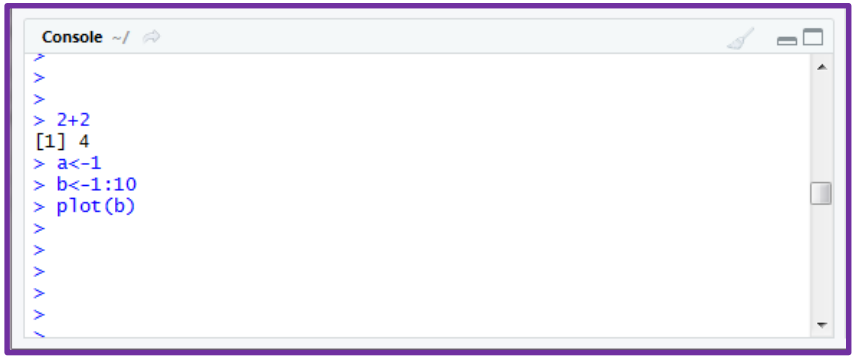
RStudio: Consola

La consola sirve para:

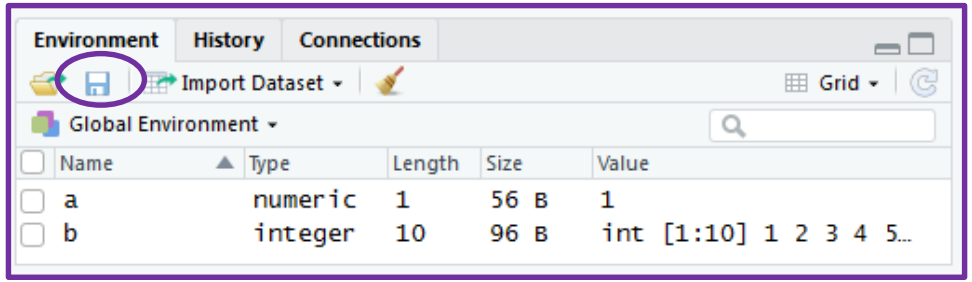
Ejecutar los comandos escritos en el editor y **visualizar las salidas**.

También podemos **escribir directamente en la consola**, esto no formara parte del script, ya que no está escrito en el editor.

La salida de la consola no se guarda.

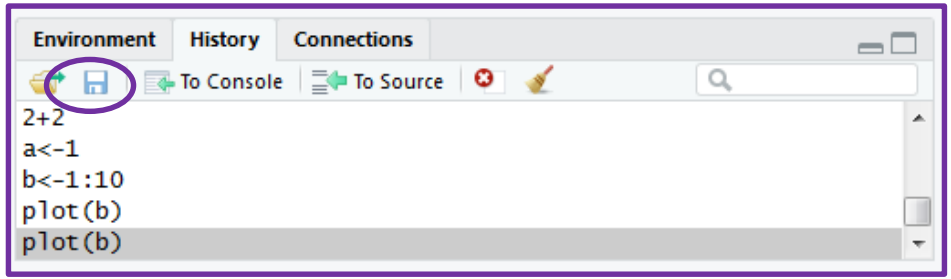


RStudio: Environment



Podemos guardar los objetos creados en un fichero con extensión **.RData** y cargarlos depués cuando queramos.

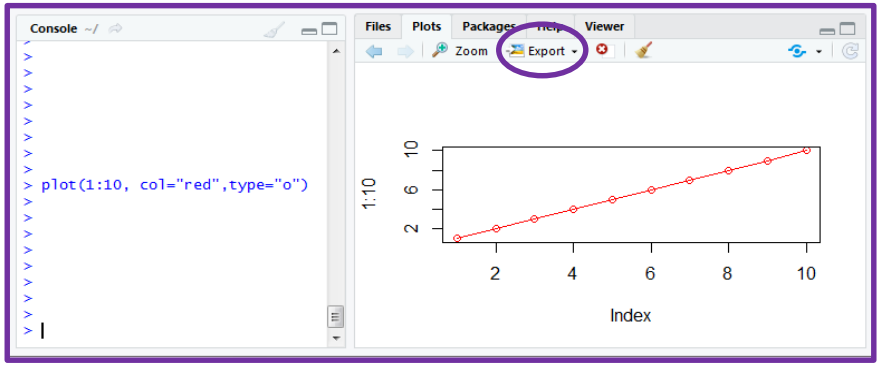
RStudio: History



En **.Rhistory** se guardan las instrucciones ejecutadas en R. RStudio por defecto guarda automáticamente el fichero .Rhistory en el directorio de trabajo actual, incluso cuando no guardamos .RData.

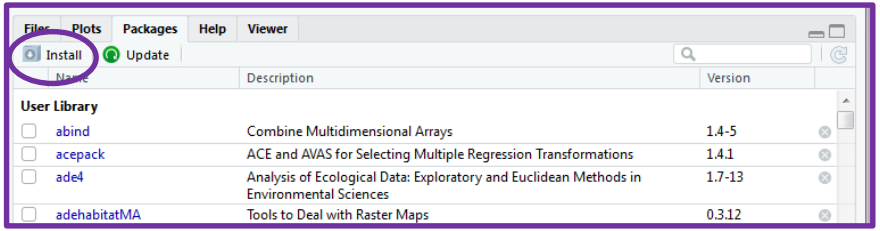
RStudio: Plots

Los gráficos se visualizan automáticamente en el panel “Plots” desde donde se puedes exportar/guardar.

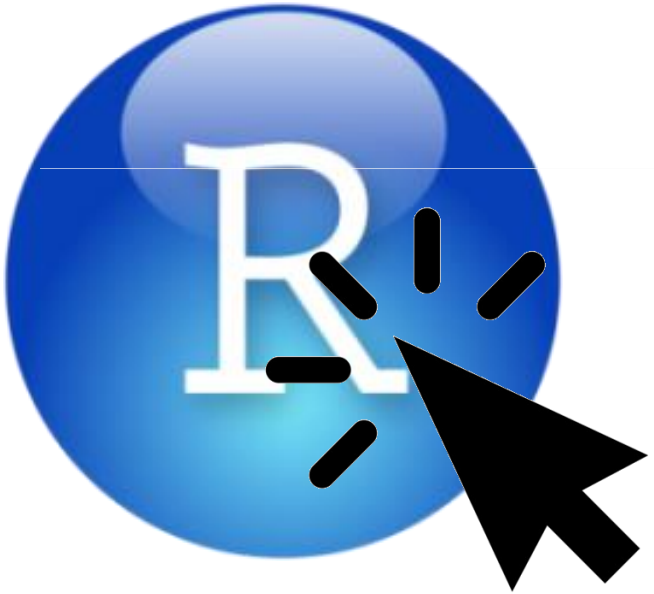


RStudio: Packages

En esta ventana vemos todo los paquetes instalados y el numero de versión. Podemos clicar “Install” para instalar un nuevo paquete metiendo su nombre.



RStudio: Help

En la ventana “Help” se muestra la ayuda solicitada.

3, 2, 1, …. Abramos R!