

Introducción a R

Josué Mendoza

IIMAS, UNAM



Familiarizarse con la sintáxis y las principales herramientas del lenguaje de programación R, lo cual facilitará la implementación de análisis de datos usando modelos psicométricos a través de este lenguaje.

Índice (Día 1)

Pre-requisitos

- Por qué usar R
- Instalar R
- Instalar RStudio
 - Introducción a RStudio (Cómo aprovecharlo).
- Instalar funciones adicionales

Uso básico de R

- Creación y manipulación de objetos
 - R como calculadora, asignación de variables
 - Vectores
 - Factores
- Estructuras de datos
 - Matrices
 - Listas
 - Data frames

Ideología GNU

Software libre significa que los usuarios de un programa tienen las cuatro libertades esenciales:

- La libertad de ejecutar el programa como lo desee, con cualquier propósito.
- La libertad de estudiar el funcionamiento del programa y adaptarlo a sus necesidades.
- La libertad de redistribuir copias para ayudar a los demás.
- La libertad de mejorar el programa y de publicar las mejoras, de modo que toda la comunidad se beneficie.



¿Qué es R?

Es un lenguaje de programación **libre**, principalmente orientado a análisis de datos.

La única limitación de R es **tu** habilidad para programar y **tu** creatividad.

¿Por qué usar R?

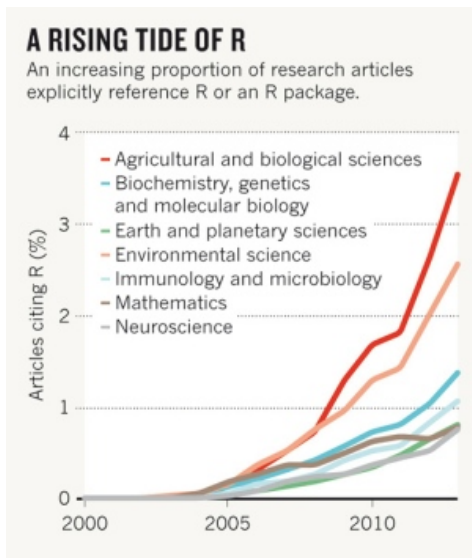


Figura 1:
Introducción a R

¿Por qué usar R?

Razones para amar R

- 1 Es software **libre**.
- 2 Es popular.
- 3 Es flexible (+ de 10,000 paquetes).
- 4 Soporte de la comunidad (la última actualización fue la semana pasada).

ugly graphics

low energy

community support

costs many \$\$\$

maps
spatial data

high energy

popularity

drop and flop

no ambitions

no limits

beautiful graphics

1.5GB storage required

shiny web applications

> 10.000 packages

SPSSCUCKS

point and click

no deeper understanding
no learning curve

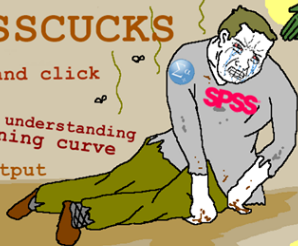
ugly output

machine learning

R-VOLUTIONARIES

doesn't require much disk space

open source



tidyverse

Rmarkdown

gdocs integration

data mining

object-oriented programming

LaTeX compatibility

unintuitive language

interactive visualization

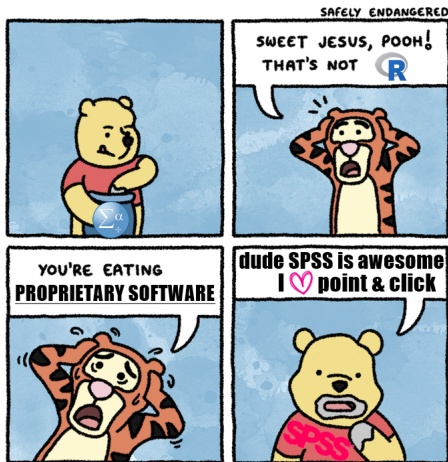
code completion

¡Empecemos!

¿Por qué usar R?

Todo tiene un costo.

Curva de aprendizaje más lenta.



1. Instalamos R

<https://cran.r-project.org>

2. Instalamos RStudio

<https://www.rstudio.com/products/rstudio/download/>

¿Qué es RStudio?

Un IDE.

Es decir, un entorno de desarrollo integrado o entorno de desarrollo interactivo, en inglés Integrated Development Environment (IDE), es una aplicación informática que proporciona servicios integrales para facilitarle al desarrollador o programador el desarrollo de software.

Normalmente, un IDE consiste de un editor de código fuente, herramientas de construcción automáticas y un depurador.

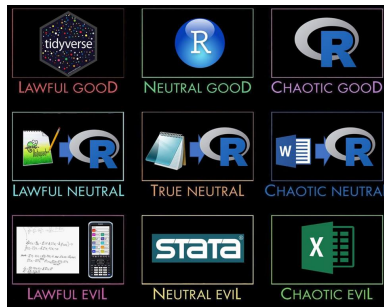
¿Por qué usarlo?

- Tiene soporte web (contenido interactivo).
- Puedes hacer proyectos.
- Te permite hacer control de versiones.
- Facilita creación de paquetes.
- Te permite leer diferentes tipos de archivos (e.g., LaTeX, Markdown, scripts de otros lenguajes).



RStudio®

What software will you be doing your statistical analysis in?



Instalar funciones adicionales

Además de las funciones que R trae por 'default', la comunidad ha desarrollado funciones específicas para cada análisis de datos.

El verdadero poder de R se debe a los miles de paquetes (+ 10,000) creados por la comunidad de estadísticos e investigadores.

Hasta este punto, hay un paquete para prácticamente todo lo que quieras hacer.

Ejemplos:

<https://github.com/leeper/meme>

<http://rmarkdown.rstudio.com/gallery.html>

Si quieren explorar...

<https://www.rdocumentation.org>

¡Manos a la obra!

<https://github.com/JosueMA/Modelizacion-Psicometrica-con-IRT>