

Investigación reproducible

Guillermo Ayala Gallego

Investigación reproducible

Guillermo Ayala Gallego

2025-04-16

¿Qué es investigación reproducible?

- Estamos acostumbrados a que los investigadores en sus publicaciones comenten los resultados obtenidos en sus investigaciones.
 - Han obtenido unos datos, los han analizado y, finalmente, los han interpretado y discutido.
 - Hemos de creer que han producido los datos correctamente, que han aplicado un tratamiento adecuado (que han hecho lo que dicen que han hecho es un mínimo) y que, finalmente, la interpretación de los resultados de las técnicas utilizadas es correcta.
 - Si no tenemos los datos, el código y la interpretación: ¿cómo sabemos que el trabajo es correcto?
 - Es una cuestión de fe.
-
- ¿Dónde están los datos?
 - ¿Qué análisis de los datos han realizado? Deme el código.
 - La interpretación es lo único que tenemos: la publicación científica
 - Por investigación reproducible entendemos procedimientos que permitan **reproducir** la investigación completa.
 - Personalmente añadiría que los distintos elementos han de ser totalmente libres.

Documento dinámico

- Donald Knuth
- Yihui Xie
- Datos del documento YAML o en LaTeX
- Chunks de texto y de código intercalados.
- Los chunks de texto usualmente en Markdown o en LaTeX.
- Los chunks de código en diferentes lenguajes (aquí nos interesan **R** y **Bash**).

Herramientas básicas

- Markdown
 1. Markdown
 2. Pandoc Markdown
- Pandoc
 1. Es un programa que convierte textos escritos en distintos lenguajes de marcas.
 2. Pandoc
- knitr
 1. Desarrollado por Yihui Xie.
 2. Nos permite trabajar con Markdown y LaTeX en chunk de texto y con R como lenguaje de programación.
 3. knitr
- RMarkdown
 1. Combinamos R con Markdown.
 2. rmarkdown
- Quarto
 1. Es una evolución de RMarkdown.
 2. Permite trabajar con distintos lenguajes de programación.
 3. Aquí nos interesan:
 1. R

2. Bash

Entornos de desarrollo

- RStudio
 1. Es la opción más sencilla
 2. Muy bueno para RMarkdown y Quarto
 3. Es mi opción para docencia
- emacs
 1. Emacs Speaks Statistics
 2. Modo polymode para Markdown
 3. Modo quarto-mode para Quarto
 4. En Debian/Ubuntu tenemos el paquete quarto
 5. Es mi opción para investigación