Generalidades del curso

Maicel Monzón

Objetivo del curso

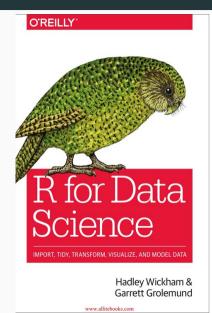
Que los sean capaces de usar R para desarrollar procesamientos estadísticos consistentes en:

- 1. Leer datos desde ficheros con formato comunes (xls,csv,sav,dat, etc.).
- 2. Ordenar datos.
- 3. Realizar transformaciones.
- 4. Realizar análisis y presentar los resultados de forma tabular y gráfica.

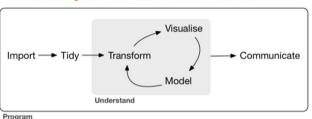
Plan temático del curso

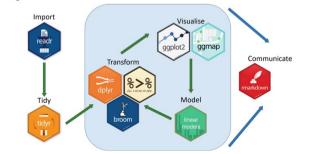
- · Módulo 1: Introducción a R y RStudio
- · Módulo 2: Importación de datos
- · Módulo 3: Ordenar datos
- · Módulo 4: Manipulación de datos
- · Módulo 5: Representación tabular
- · Módulo 6: Representación gráfica
- · Proyecto Final: Informes con R Markdown

Plan temático del curso desde el enfoque de ciencias de datos.



Enfoque de ciencia de datos





Plan temático del curso desde el enfoque de la investigación científica tradicional.

Enfoque del proceso de investigación científica

Metodología de la Investigación

ELEMENTOS BÁSICOS PARA LA INVESTIGACIÓN CLÍNICA

delimitación del problema.

delimitación del problema.

delimitación del problema.

delimitación del problema.

planificación.

planificación.

formudación de los objetivos.
selección de los métodos y técnicas a emplear.
determinación de la forma y procedimientos para la elaboración y análisis de los resultados.
procesamiento y análisis de los resultados.
confección del informe final.

publicación de los resultados en la práctica social.



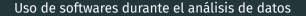
ROSA JIMÉNEZ PANEQUE

Tipos de estudios



Muchos de los procedimientos de este curso se pueden implemetar en analisis de datos de forticación de alimentos a gran escala.





"Se suele decir que los científicos de datos pasamos el 80 por ciento del tiempo **limpiando, preparando y re-organizando datos** y tan solo un 20 en el **análisis** propiamente dicho"

Hadley Wickham

"De 80% limpieza a 20% análisis, a la magia de los datos ordenados"

```
library(gtsummary)
trial %>%
  tbl_summary(by = trt, include = c(age, grade))  %>%
  add_p()
```

Characteristic	Drug A N = 98^{1}	Drug B N = 102^{1}	p-value ²
Age	46 (37, 60)	48 (39, 56)	0.7
Unknown	7	4	
Grade			0.9
1	35 (36%)	33 (32%)	
II	32 (33%)	36 (35%)	
III	31 (32%)	33 (32%)	
¹ Median (Q1, Q3); n	(%)		

²Wilcovon rank sum test: Pearson's Chi-squared test

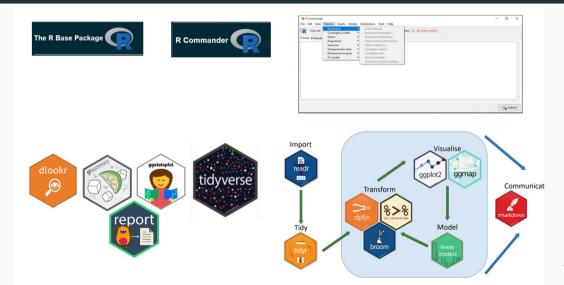
Tidyverse vs R base

Tidyverse ofrece una sintaxis más **intuitiva** y **legible**, es más **accesible para principiantes** y está optimizado para **análisis complejos**, mientras que R Base es más flexible y rápido en operaciones simples, pero con una **curva de aprendizaje más pronunciada**.

Curva de aprendizaje

```
Habilidad (Y)
Experto
                                 * (R Base)
                   * (Tidyverse)
Intermedio -
Básico - * (SPSS)
                                   Tiempo/Esfuerzo (X)
```

Tidyverse; como pequeñas elementos se puden convertir en sitemas poderosos y al alncance de todos.



¿Te sientes abrumado por R?

· No estás solo. Todos empezamos igual.

Cuando comencé con R hace 10 años...

- · No entendía nada.
- · Copiaba código de internet.
- · Lo adaptaba a mis necesidades.

El secreto está en intentarlo

Pasos simples para empezar:

- 1. Copia código.
- 2. Adáptalo.
- 3. Aprende mientras avanzas.

No necesitas ser un experto desde el principio.

¡Solo empieza!

¡Bienvenidos al curso!

¿Qué ofrece este curso?

- Estructura clara: 6 módulos divididos en sesiones matutinas (4 horas) y vespertinas (3 horas).
- Pausas activas: Aplicamos principios de neurodidáctica con pausas de 10-15 minutos entre bloques para evitar la fatiga mental.
- Storytelling: Historias como esta te inspirarán y te mostrarán que está bien cometer errores o depender de recursos externos al principio.
- Gamificación: Ejercicios prácticos individuales y en grupo para reforzar los conceptos aprendidos.

La metáfora de Nutrilandia (Gamificación)

Imagina un "Edificio de la Fortificación" donde cada elemento arquitectónico representa un micronutriente clave . Los estudiantes "construyen" el edificio completando módulos y desafíos, mientras aprenden sobre análisis de datos en R y la importancia de cada nutriente.

La metáfora de Nutrilandia (Gamificación)

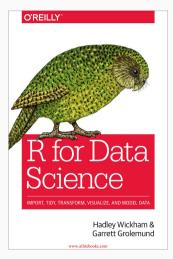
Imagina un "Edificio de la Fortificación" donde cada elemento arquitectónico representa un micronutriente clave. Los estudiantes "construyen" el edificio completando módulos y desafíos, mientras aprenden sobre análisis de datos en R y la importancia de cada nutriente.



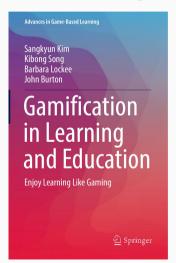
Cimientos: Hierro (datos limpios). Columnas: Zinc (manipulación de datos).

Pisos: Vitamina A (visualización). Ventanas: Ácido fólico (limpieza de datos).

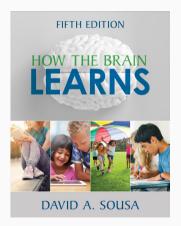
R for Data Science - Hadley Wickham (https://r4ds.had.co.nz/)



Gamification in Learning and Education - Sangkyun Kim

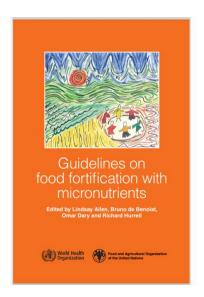


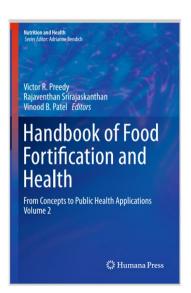
How the Brain Learns - David A. Sousa











Participación de

















Participación especial

