



Visualización avanzada

4. Una gramática de los gráficos

Tipos de datos -> tipos de gráficos

Componentes de una gramática genérica para cualquier tipo de gráfico

ggplot2

- Componentes de la gramática

- Tipos frecuentes de gráficos

- Desarrollo completo de un ejemplo de “scatterplot”

ggplot: la gramática de los gráficos

Para conocer sus orígenes:

[https://byrneslab.net/classes/biol607/readings/wickham_layered-grammar.p
df](https://byrneslab.net/classes/biol607/readings/wickham_layered-grammar.pdf)

Tiene como objetivo estandarizar muchos de los detalles que hacen un lío desarrollar un gráfico (como dibujar leyendas) así como proporcionar un **modelo potente de gráficos complejos basado en la superposición de capas.** Dedicaremos una lección completa a este paquete esencial.

¿Por qué una “gramática de gráficos”?

Ya hemos visto la falta de consistencia de las funciones gráficas en diferentes paquetes.

Para solucionarlo Hadley Wickham, de Rstudio, desarrolló ggplot.

Usaremos la excelente “chuleta” (cheatsheet) de la propia Rstudio:

<https://raw.githubusercontent.com/rstudio/cheatsheets/main/data-visualization.pdf>

que también está traducida a español:

https://raw.githubusercontent.com/rstudio/cheatsheets/main/translations/spanish/data-visualization_es.pdf

Componentes de una gramática genérica

Mínimo requerido:

- Datos

- Geoms –componentes visuales para representar los puntos

- Propiedades estéticas –tamaño, color, posición

Adicional:

- Coordenadas

- Transformación estadística

- Facetas



Tipos de datos condicionan tipos de gráficos

Conceptos clave:

Tipos de variable (también “nivel de medida”...)

- Nivel de intervalo: los números mantienen significado cuando se suman y se restan.
 - De entre las de nivel de intervalo:
 - discretas (“números enteros”) -> `as.factor`
 - continuas (“números racionales”) -> `as.numeric`
 - porcentajes (normalmente de 0 a 100)
 - probabilidades (estrictamente entre 0 y 1)
 - etc.
 - Categóricas o nominales: dos números diferentes sólo indican permanencia a categorías distintas, no suman ni permiten otras operaciones. Por ejemplo: código postal, matrícula... -> `as.factor`
 - Ordinales: números mantienen significado de orden, pero no tienen propiedades de intervalo. Por ejemplo, algunas “rating scales”.



plotnine

<https://plotnine.readthedocs.io/en/stable/gallery.html>

<https://plotnine.readthedocs.io/en/stable/api.html>

¡Gracias!

Pedro Concejero

pedro.concejerocezoz@gmail.com