PEC 2 · VISUALIZACIÓN DE DATOS · UOC

Tres técnicas, tres conjuntos de datos abiertos

Diagrama de dispersión · Slope chart · Ridgeline plot

Jorge Mateu Marín · Máster Universitario de Ciencia de Datos · UOC

PRESENTACIÓN

Jorge Mateu Marín · PEC 2

Visualización de Datos · Máster Universitario de Ciencia de Datos · UOC.

Investigar tres técnicas asignadas y buscar datos abiertos coherentes con cada una.

Demostrar el recorrido estructura del dato → pregunta → codificación visual.

Técnicas asignadas

I Diagrama de dispersión

II Slope chart

III Ridgeline plot

ENLACES PÚBLICOS

Tres visualizaciones sin registro

Técnica	Datos	URL directa
Diagrama de dispersión	PIB por habitante y esperanza de vida, 2022	pec2-visualizacion.pages.dev/dispersion
Slope chart	Cuota de renovables, UE, 2010 y 2023	pec2-visualizacion.pages.dev/pendiente
Ridgeline plot	Máximas diarias, Barcelona, 2024	pec2-visualizacion.pages.dev/crestas

Índice: pec2-visualizacion.pages.dev · Fuentes, preparación, alcance y tablas accesibles en las páginas.

Datos: Our World in Data (s.f.); Eurostat (2026); Open-Meteo (s.f.).

TÉCNICA 1 · GRUPO I

Diagrama de dispersión

Dos variables cuantitativas · un punto por observación · estudiar asociación y excepciones

Dos valores se convierten en una posición

Mecánica: x = primera medida; y = segunda. Color opcional para grupos; tamaño para una tercera medida.

Antecedentes: Herschel (1833) y Galton (1886), citados en Friendly y Denis (2005).

Ejemplos: superficie y precio por vivienda; dosis y respuesta por paciente.

Data Viz Project: Plot · Correlation / Distribution · Dot.

From Data to Viz: dos variables numéricas → Scatterplot.

Data Visualisation Catalogue: Patterns / Relationships.

Chartmaker: Scatter plot → Relationships / Multi-variate, con ficha propia.

Fuentes: Friendly y Denis (2005); Ferdio (s.f.); Holtz y Healy (2018); Ribecca (s.f.); Kirk (s.f.).

La legibilidad depende de la densidad

Estructura y escalas

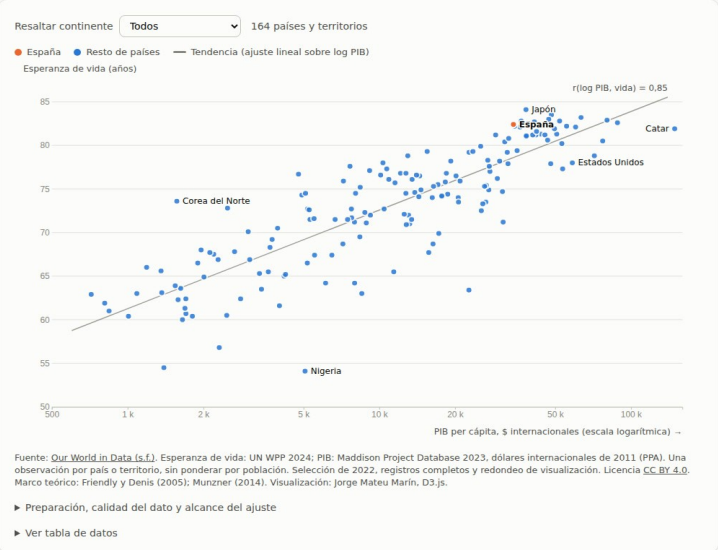
- Una fila por país o territorio; dos medidas cuantitativas.
- PIB por habitante y esperanza de vida: medidas de razón; PIB positivo.
- País y continente: dimensiones nominales. Cada punto pesa igual.
- Eje x logarítmico: distancias iguales = multiplicaciones iguales.

Restricciones y alternativas

- De decenas a pocos miles de puntos como guía; no hay un máximo universal.
- Solapamiento → transparencia, hexbin o densidad 2D.
- Categorías en un eje → strip plot o cajas; pocos puntos → prudencia.
- Subgrupos ocultos → resaltar o facetar. Asociación no implica causalidad.

Fuentes: Munzner (2014, caps. 2 y 5); Holtz y Healy (2018); Schwabish (2021, cap. 8).

Más PIB se asocia con mayor esperanza de vida



164 registros de 2022, OWID; selección, redondeo y una exclusión provisional documentada.

r = 0,85 con log10(PIB). EE. UU.: −3,3 años; Nigeria: −15,1 respecto a la recta.

Curva de Preston (1975): cada vez menos años por dólar. Son diferencias entre países; no efectos causales.

Audiencia: estudiantes y público general. España destacada; etiquetas selectivas.

Demostración: resaltar continente y consultar puntos.

Datos: Our World in Data (s.f.). Diseño: Munzner (2014); Coll Campmany (2023).

TÉCNICA 2 · GRUPO II

Slope chart

Dos estados comparables · una línea por categoría · nivel, cambio y ranking

La pendiente resume el cambio entre dos extremos

Mecánica: categoría → segmento; estado o fecha → lado; valor → altura en una escala común.

Antecedente: Tufte, 1983: ingresos fiscales de países en 1970 y 1979.

Ejemplos: intención de voto en dos encuestas; precios de productos antes y después.

Data Viz Project: Chart · Comparison / Trend over time · Line.

From Data to Viz y Catalogue: sin ficha propia en los catálogos consultados.

Chartmaker: Slope graph → Temporal / Trends; alias Slope chart. Ficha propia.

Interpretación: parentesco con un gráfico de líneas de dos extremos; ausencia de ficha no equivale a discrepancia.

Fuentes: Tufte (1983, pp. 158–159; 2011); Ferdio (s.f.); Kirk (s.f.); Schwabish (2021, cap. 5).

Misma medida, unidad y escala en ambos lados

Estructura y agregación

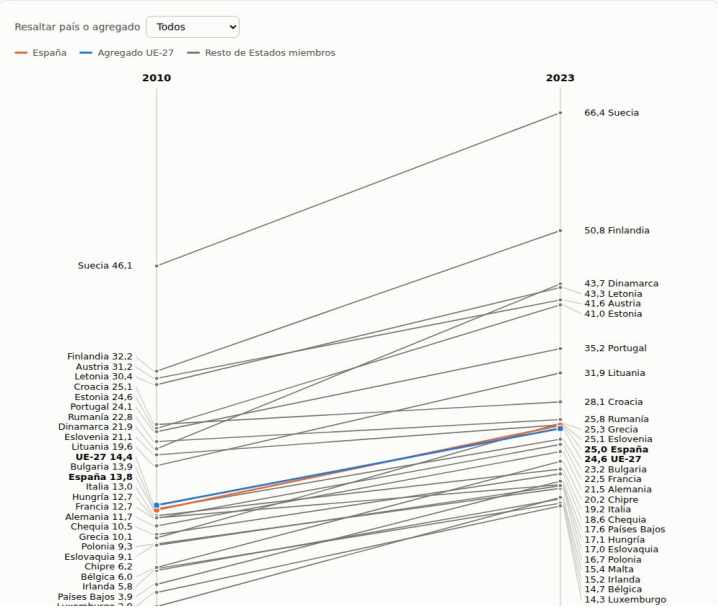
- Categoría × estado → medida; formato ancho o largo. Aquí: país, año y cuota.
- Años: escala temporal de intervalo. Cuota: % del consumo final bruto.
- 27 países y un agregado UE; el agregado no entra en el ranking nacional.
- Cambio expresado en puntos porcentuales.

Restricciones y alternativas

- Unas 5–30 líneas como guía; valores próximos pueden saturar antes.
- Etiquetas o cruces densos → selección, tabla o dumbbell; solo el puesto → bump chart.
- Trayectoria intermedia importante → serie anual completa.
- Extremos ausentes o unidades distintas impiden comparar la pendiente.

Fuentes: Munzner (2014, caps. 2 y 5); Tufte (2011); Gil Bellosta (s.f.).

Subir la cuota y mejorar el puesto son preguntas distintas



España: 13,8 → 25,0 %; puesto 13 en ambos años.

Dinamarca: +21,8 p.p.; puesto 9 → 3.

Croacia: aumenta, pero pasa del puesto 5 al 9.

UE: 14,4 → 24,6 %, agregado sobre el consumo total.

Objetivo UE 2030: 42,5 % (Directiva 2023/2413); la media queda lejos. La cuota no indica energía absoluta ni causalidad.

Audiencia: periodismo y responsables de energía.

Datos y reutilización: Eurostat (2026; s.f.). Interpretación y diseño: elaboración propia.

TÉCNICA 3 · GRUPO III

Ridgeline plot

Una distribución por grupo · eje horizontal común · comparar posición y dispersión

Una curva de densidad por grupo

Mecánica: valor → x; grupo → fila; densidad estimada → altura. Cada curva tiene área 1.

Antecedente visual: portada de Unknown Pleasures (1979), citada por Wilke (2017).

2017: Jenny Bryan propone «joyplot»; Claus Wilke crea ggjoy y lo sustituye por ggridges / ridgeline.

Ejemplos: edades por década; respuestas de probabilidad por expresión verbal.

From Data to Viz: variable numérica + grupos; función de distribución, ficha Ridgeline.

Data Viz Project y Catalogue: sin ficha propia en los catálogos consultados.

Chartmaker: Density plot → Categorical / Distributions; alias Ridgeline plot y joy plot.

El contraste separa estructura de datos, finalidad y cobertura de cada catálogo.

Fuentes: Wilke (2017); Holtz y Healy (2018); Kirk (s.f.); Schwabish (2021, cap. 6).

Comparar distribuciones exige controlar el suavizado

Estructura y estimación

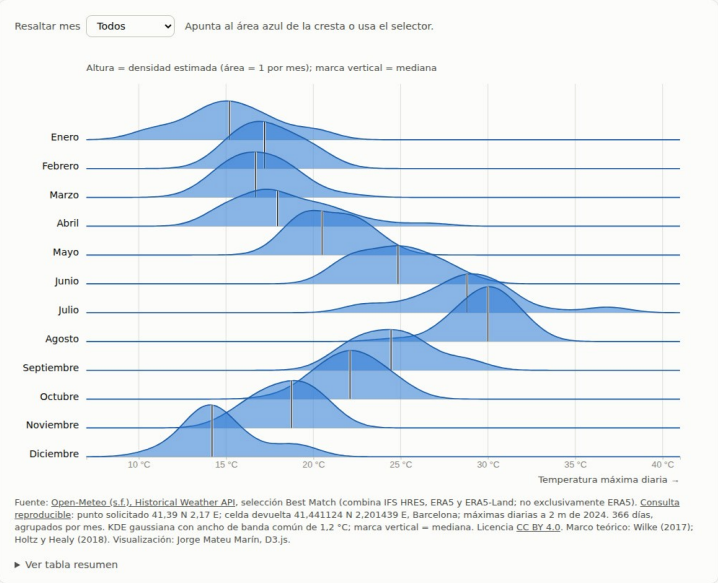
- Una fila por día: máxima en °C (intervalo) y mes (dimensión cíclica ordenada).
- 366 días → 12 grupos de 29–31 observaciones.
- KDE gaussiana: ancho de banda común 1,2 °C; misma escala de densidad.
- Mediana y tabla resumen complementan las formas.

Restricciones y alternativas

- From Data to Viz orienta hacia seis o más grupos; no es un requisito matemático.
- Decenas de datos por grupo dan una vista descriptiva, no formas estables.
- Hasta unas 30 crestas según altura y solapamiento; después, mapa de calor o paneles.
- Pocos datos → puntos; cifras exactas → tabla.

Fuentes: Munzner (2014); Holtz y Healy (2018); Wilke (s.f.); Gil Bellosta (s.f.).

Barcelona, 2024: cambia la posición y la dispersión



Open-Meteo Best Match: datos de modelo, no estación ni exclusivamente ERA5.

Medianas: enero 15,2 °C; agosto 30,0 °C.

Agosto: menor dispersión; 30 de 31 días ≥ 25 °C.

Julio: mayor dispersión; máxima de 37,1 °C.

Audiencia: docencia y divulgación meteorológica. Un año no demuestra una tendencia climática.

Demostración: áreas de las crestas, selector y tabla.

Datos: Open-Meteo (s.f.), Best Match, CC BY 4.0. Diseño: elaboración propia.

Tres preguntas y un mismo criterio

Dispersión: dos medidas → asociación → posición común.

Slope: dos estados → cambio → extremos conectados.

Ridgeline: datos por grupos → distribución → alineación horizontal.

La herramienta ejecuta una elección analítica fundamentada.

Referencias (selección; lista completa en Harvard en la web)

- Munzner, T. (2014) Visualization analysis and design. Boca Raton: A K Peters/CRC Press.
- Schwabish, J. (2021) Better data visualizations. Nueva York: Columbia University Press.
- Friendly, M. y Denis, D. (2005) «The early origins and development of the scatterplot», JHBS, 41(2), pp. 103–130.
- Tufte, E.R. (1983) The visual display of quantitative information. Cheshire: Graphics Press, pp. 158–159.
- Wilke, C.O. (2017) Goodbye joyplots. clauswilke.com (Consultado: 1 noviembre 2026).