

PEC 2

Estudio de técnicas de visualización de datos

VISUALIZACIÓN DE DATOS
Máster Universitario en Ciencia de Datos

Miguel Angel Sosvilla Luis

Publicación de las visualizaciones

- http://msosvi.github.io/uoc/m2859_pec2_msosvilla.html
- http://msosvi.github.io/uoc/m2859_pec2_msosvilla_presentacion.pdf



Técnica 1: Heatmap

Un heatmap o mapa de color es una representación gráfica de datos en la que los valores individuales contenidos en una matriz se representan con colores de tal manera que las variaciones de color permite hacer una lectura muy intuitiva de las variaciones de los datos.

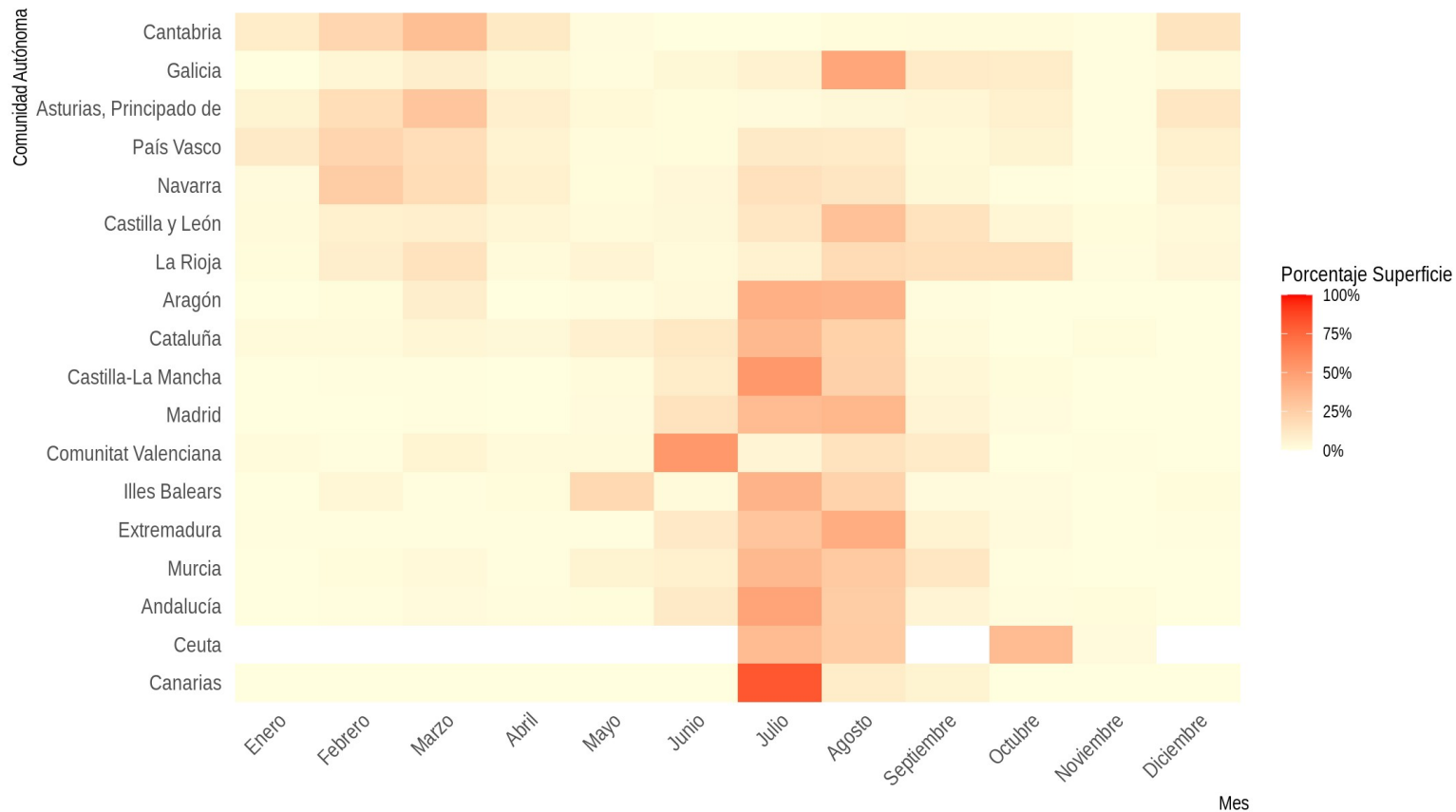
Aunque el nombre es reciente, esta técnica de visualización se usa desde el siglo XIX.

Es útil para mostrar una visión general de datos numéricos, revelando correlaciones entre variables o patrones temporales.

Uno de los usos más habituales es la representación gráfica de la matriz de correlación de un conjunto de variables numéricas.

¿En qué meses se quema más superficie?

Porcentaje de superficie quemada en cada mes sobre el total de superficie quemada



Fuente de los datos: Civio (<https://datos.civio.es/dataset/todos-los-incendios-forestales>)

Técnica 2: Parallel Coordinates Plot

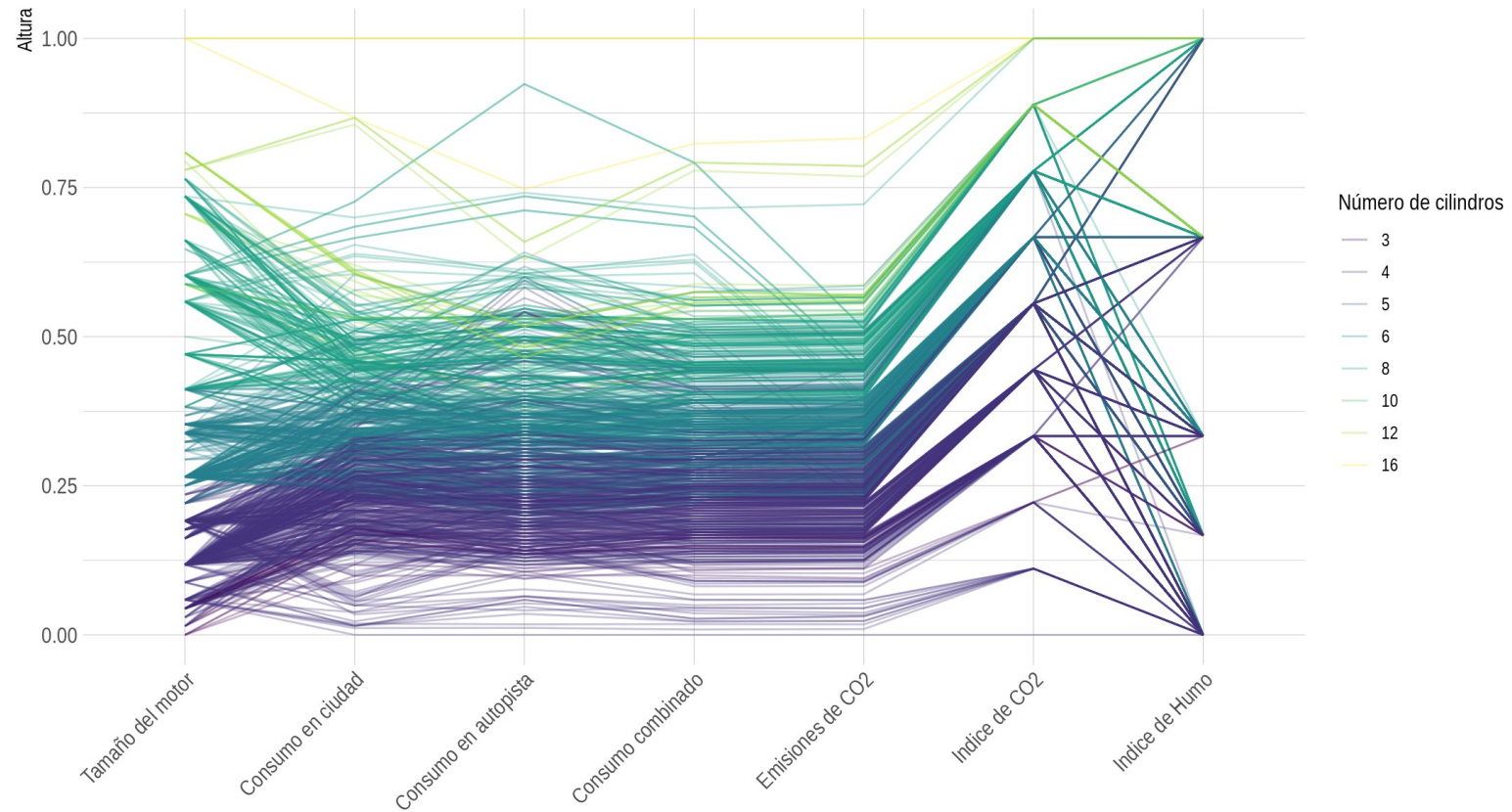
Un gráfico de coordenadas paralelas permite comparar variables numéricas continuas de un conjunto de datos, permitiendo localizar patrones en las variables o agrupaciones en las observaciones.

Una de sus principales ventajas es que permite la comparación de muchas variables en un solo gráfico, preservando cada una su escala y unidades (aunque se puede hacer una normalización previa de las variables).

Cada variable tiene su propio eje en el gráfico y todos los ejes se colocan de forma paralela. Los valores se representan como líneas conectadas o través de cada eje.

Relación del Consumo de Combustible y Emisiones de Co2

Modelos de coches en venta en Canadá en 2022



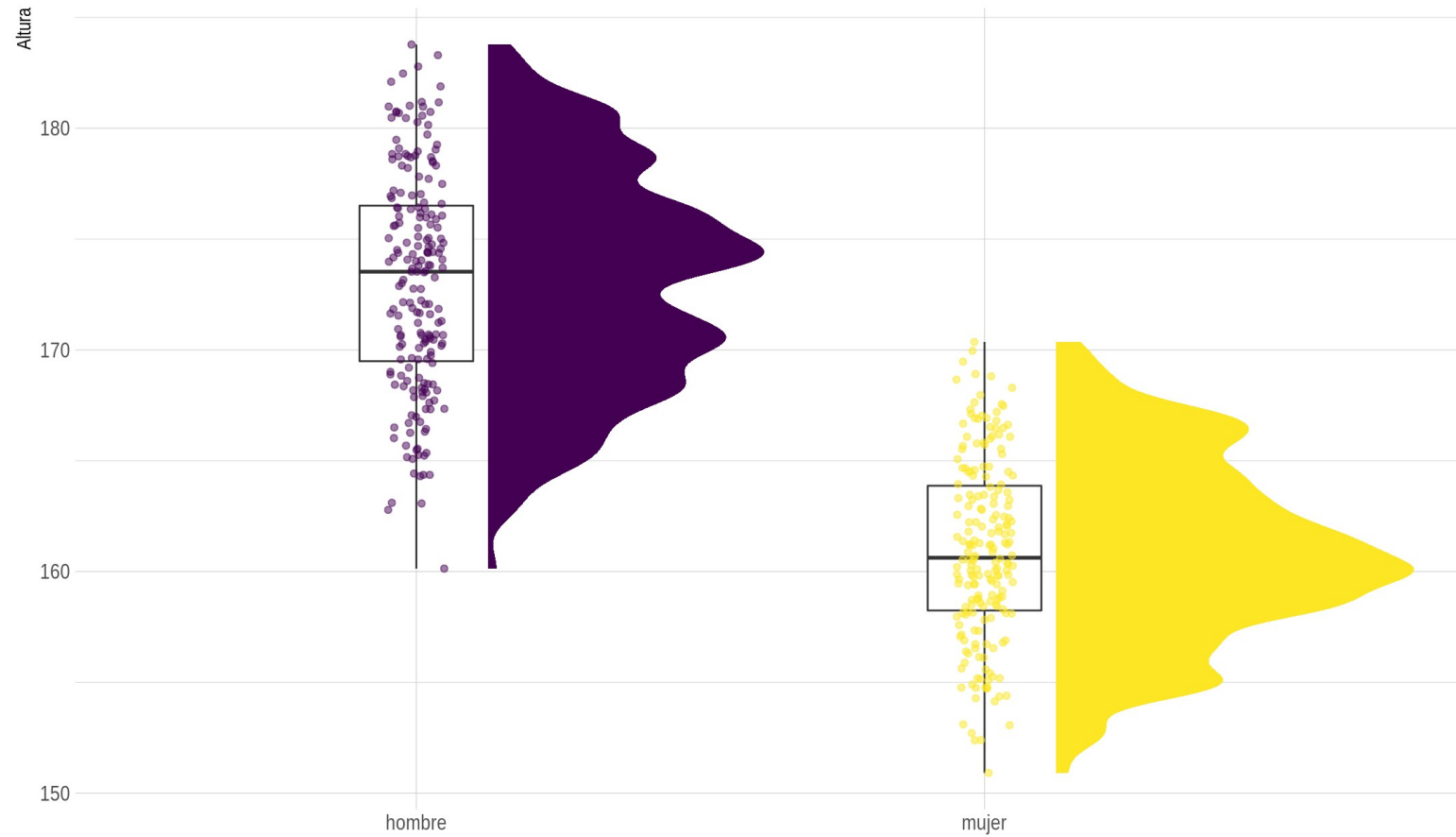
Fuente de los datos: Kaggle (<https://www.kaggle.com/datasets/rinichristy/2022-fuel-consumption-ratings>)

Técnica 3: Raincloud Plot

El “raincloud plot” es un gráfico que permite visualizar de forma rápida, intuitiva y al mismo tiempo los datos en bruto, la distribución y las estadísticas de resumen de una variable numérica.

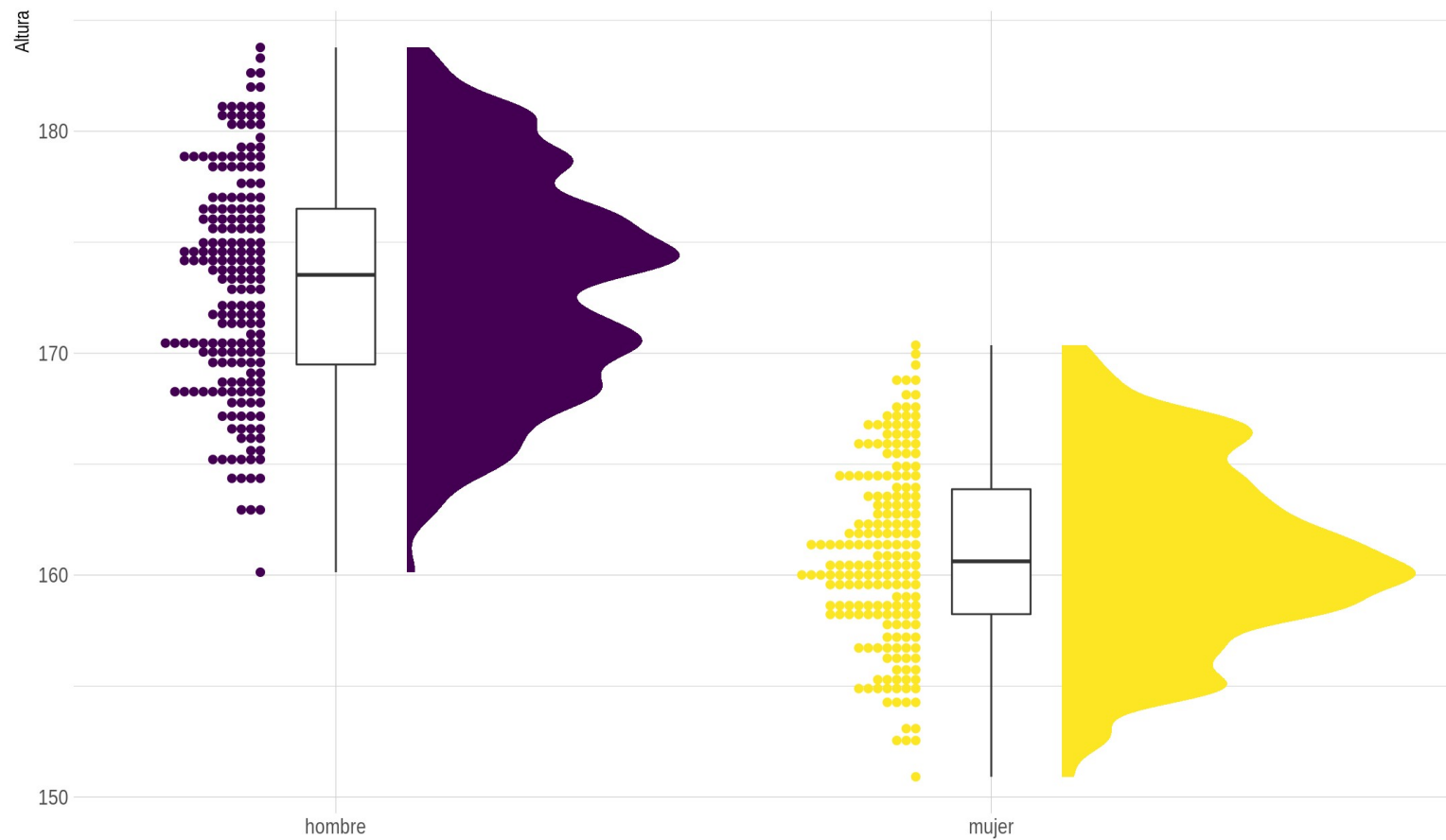
Es una combinación de tres tipos de gráficos: un gráfico de “medio” violín, un bloxpote y un jitter plot. El “raincloud plot” fue presentado en 2019 como un intento de solventar las carencias para informar sobre la distribución y el número de casos que tienen estos tipos de gráficos por separado.

Distribución de Alturas Medias de Hombres y Mujeres Por Países



Fuente de los datos: Kaggle (<https://www.kaggle.com/datasets/majyhain/height-of-male-and-female-by-country-2022>)

Distribución de Alturas Medias de Hombres y Mujeres Por Países



Fuente de los datos: Kaggle (<https://www.kaggle.com/majhain/height-of-male-and-female-by-country-2022>)