

Guía práctica para interpretar gráficos estadísticos

Elaborado en septiembre de 2025

Índice

1. Gráfico de dispersión (Scatter)
2. Gráfico de barras
3. Diagrama de torta (Circular)
4. Matriz de correlación
5. Histograma
6. Diagrama de nulos
7. Diagrama de cajas (Boxplot)

Gráfico de dispersión (Scatter)

Permite analizar la relación entre dos variables numéricas representando cada par de valores como un punto en el plano.

Sirve para observar agrupaciones, alineamientos o patrones, lo que sugiere la existencia de relación (correlación positiva, negativa o nula).

Si los puntos forman una franja o línea, hay relación; si están dispersos sin patrón claro, las variables son independientes.

Facilita la detección de valores atípicos y tendencias generales, siendo fundamental en el análisis exploratorio de datos.

Gráfico de barras

Facilita comparar cantidades entre distintas categorías usando barras horizontales o verticales.

La altura o longitud de cada barra corresponde al valor de la categoría.

Permite rápidamente identificar la mayor, menor o igualdad entre categorías.

Es idóneo para comparar resultados de distintos periodos o grupos y observar cambios o tendencias.

Diagrama de torta (Circular)

Representa la proporción de cada categoría respecto del total mediante sectores de un círculo.

Cada sector muestra el porcentaje que representa la categoría sobre el total.

Es especialmente útil para ver cómo se reparte un conjunto de datos entre sus partes.

Cuando hay muchas categorías, pierde claridad; recomendable para menos de seis.

Matriz de correlación

Relaciona en forma matricial varias variables cuantitativas, mostrando su grado de correlación.

Cada celda indica la fuerza y sentido de la relación entre las variables.

Se representa con colores, símbolos o números: cerca de +1 (fuerte positiva), -1 (fuerte negativa), o 0 (nula).

Fundamental para analizar relaciones múltiples y detectar variables que evolucionan juntas o en direcciones opuestas.

Histograma

Muestra la distribución de los valores de una variable numérica agrupándolos en intervalos.

La altura de cada barra refleja la frecuencia de datos dentro de un rango.

Permite apreciar la forma de la distribución: simetría, sesgo, concentración, dispersión o presencia de datos extremos.

Clave para conocer la variabilidad y forma de los datos antes de un análisis más profundo.

Diagrama de nulos

Indica la presencia y distribución de valores faltantes en un conjunto de datos.

Se representa mediante una matriz donde se destacan los lugares donde faltan datos, usando colores o símbolos.

Esencial para evaluar la calidad y completitud antes de cualquier análisis posterior.

Ayuda a decidir estrategias de limpieza o imputación según la cantidad y distribución de los nulos.

Diagrama de cajas (Boxplot)

Visualiza de manera rápida la dispersión, tendencia central y presencia de valores atípicos.

La caja representa el rango intercuartílico (Q1 a Q3, el 50% central de los datos).

La línea central es la mediana, los bigotes muestran la extensión normal y los puntos aislados los atípicos.

Si la caja o los bigotes son desiguales, existe sesgo: la distribución está cargada hacia uno de los extremos.

Es fundamental para detectar simetría, dispersión y valores extremos, comparando varios grupos fácilmente.