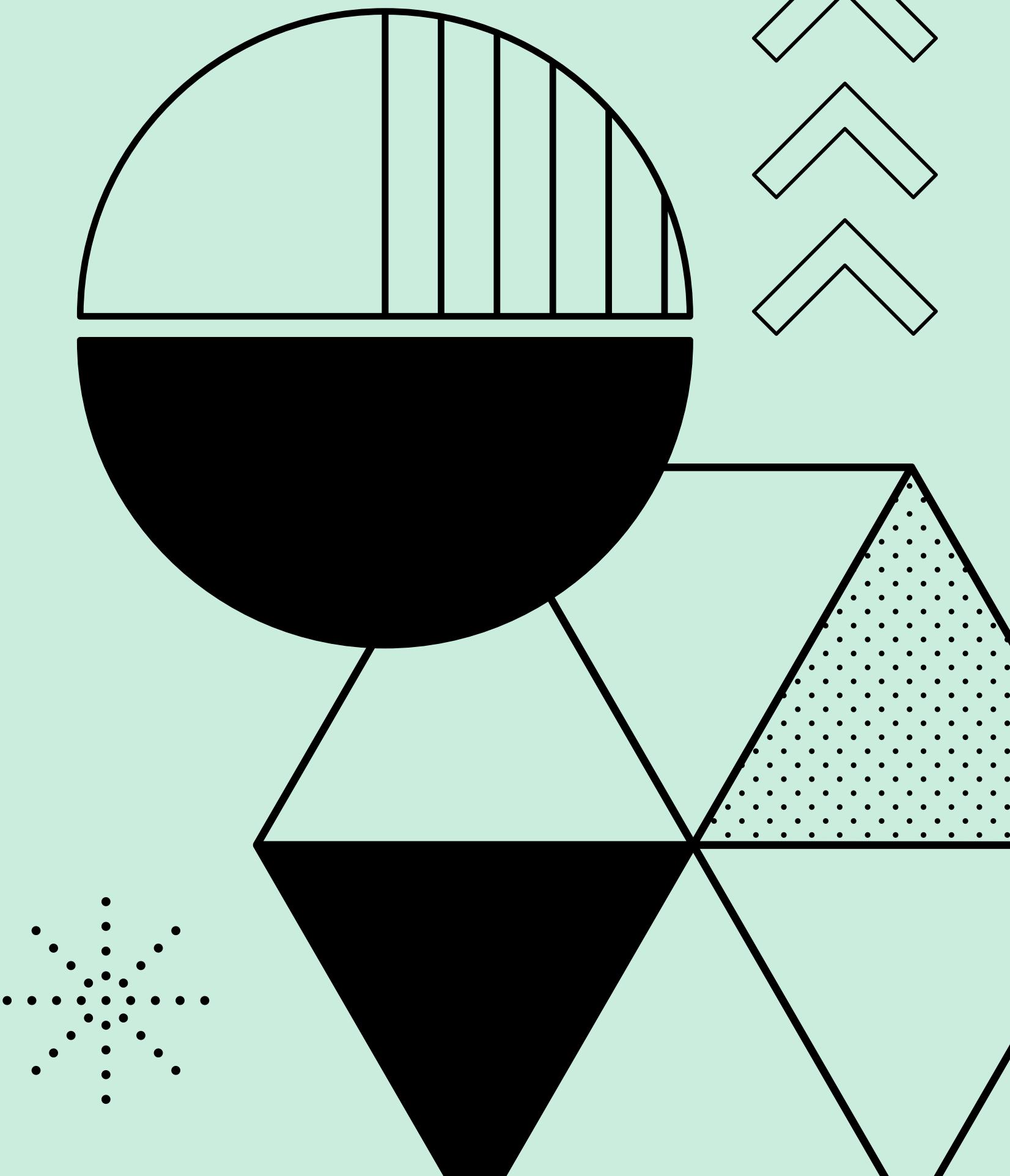
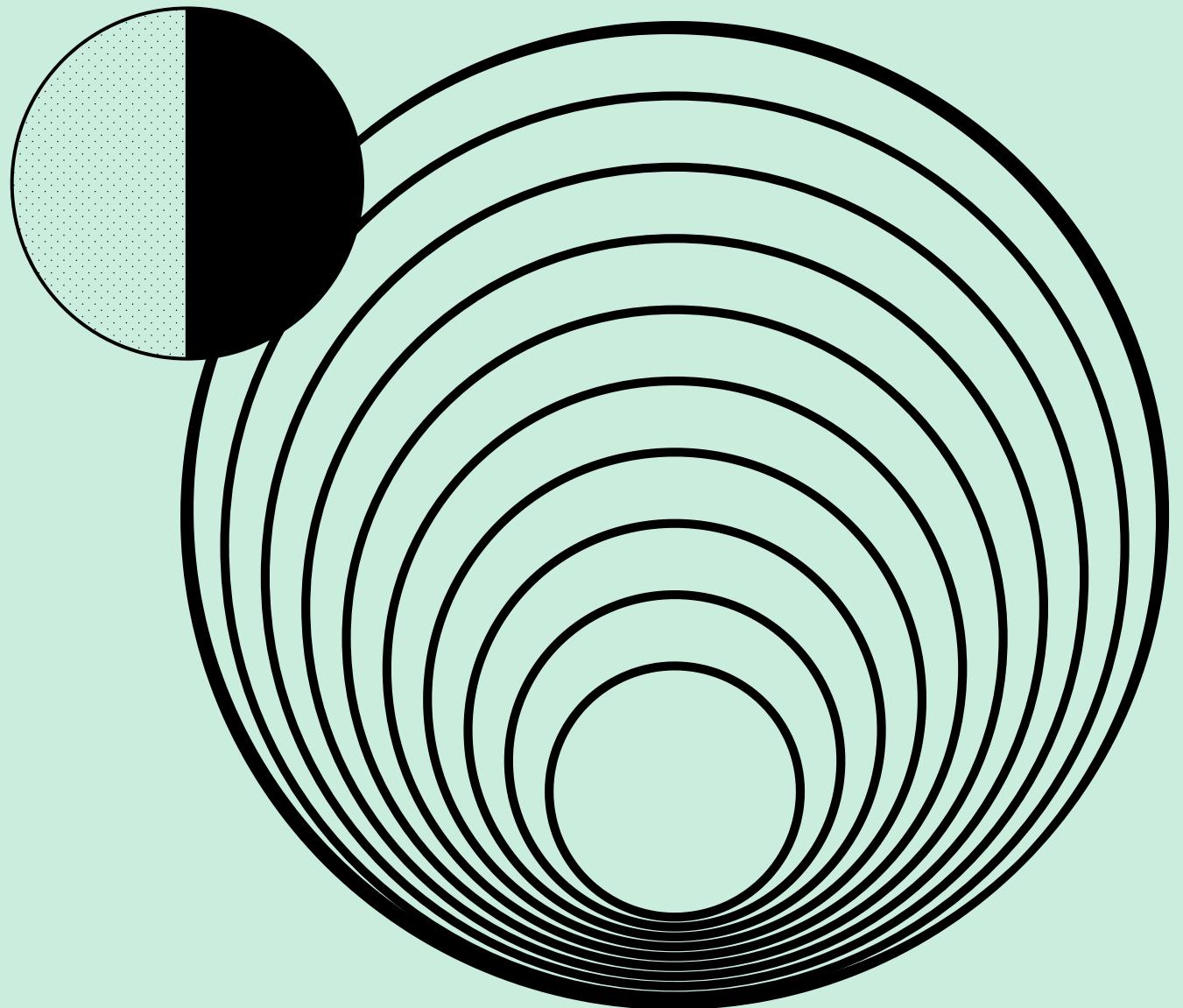


ESTADÍSTICA DESCRIPTIVA.

Por: Antonio Rodríguez Mejía.
Noviembre de 2025.



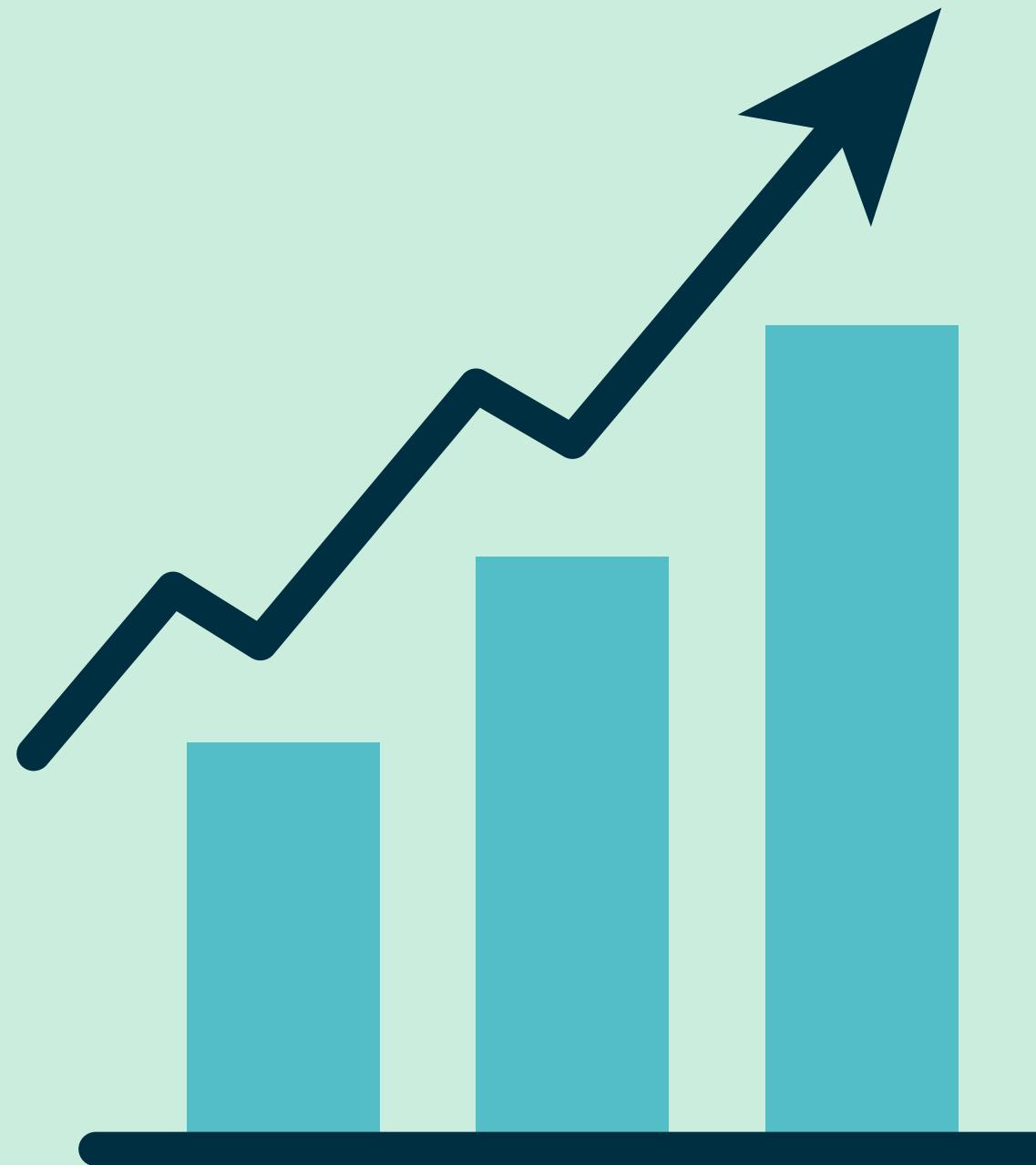


Contenidos.

- 1.** Introducción y tablas de frecuencia.
- 2.** Medidas de tendencia central, variabilidad y posición.
- 3.** Correlación, gráficos y visualización.
- 4.** Integración, práctica de proyecto y cierre.

Estadística descriptiva

Rama de la estadística que organiza, resume y presenta datos para describir sus características principales mediante medidas numéricas y gráficos. Se enfoca en el resumen de datos, la identificación de tendencias (tendencia central) y la cuantificación de la variabilidad (dispersión) de un conjunto de datos sin hacer inferencias sobre una población más grande.



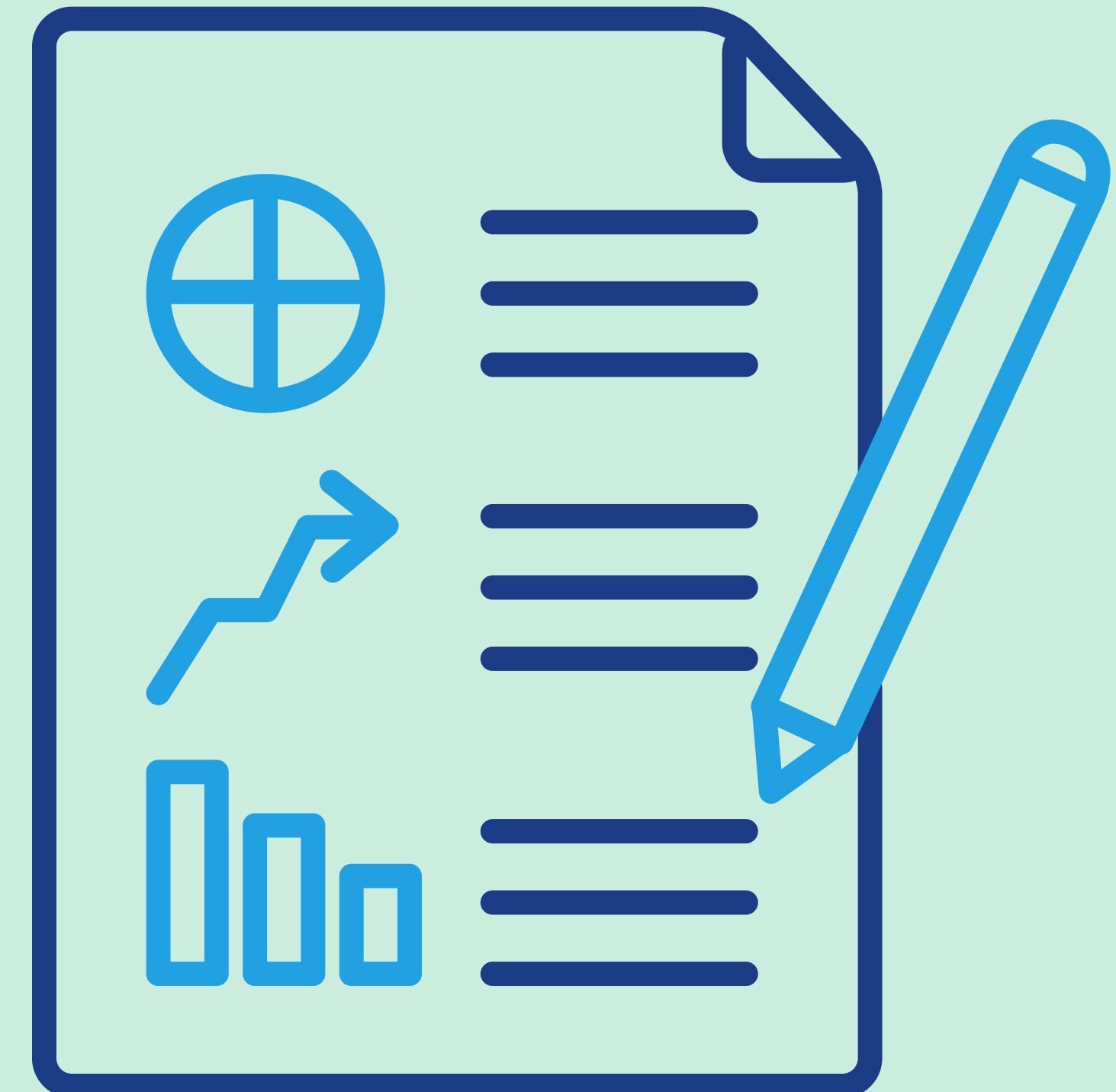
Tipos de datos

Cualitativos:

- Ordinales: Atributos que presentan orden.
- Nominales: Atributos que no presentan orden.

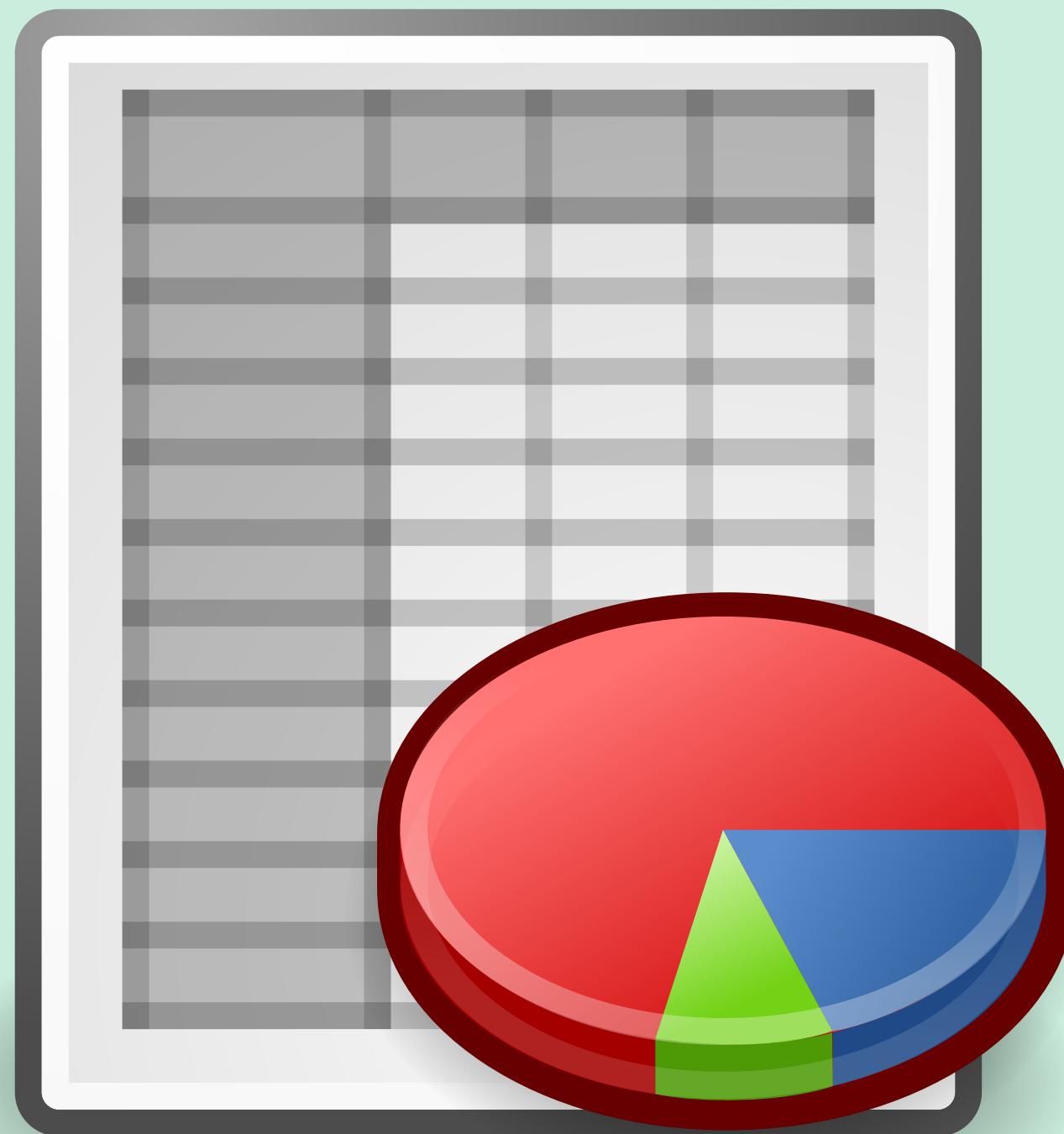
Cuantitativos:

- Discretos: Resultan de conteos, números enteros (-2, -1 ,0 ,1 ,2, ...).
- Continuos: Resultan de mediciones, por ejemplo, medidas físicas (peso, temperatura, altura).



Tablas de frecuencia

Es una herramienta estadística que organiza y resume la información sobre la distribución de datos. Consiste en una estructura de columnas que muestra las distintas categorías o valores de una variable, junto con el número de veces que cada categoría ocurre en un conjunto de datos. Esta tabla proporciona una visión general y ordenada de la frecuencia con la que aparecen los valores en una muestra o población, lo que permite analizar patrones, identificar tendencias y extraer conclusiones significativas.



Elementos



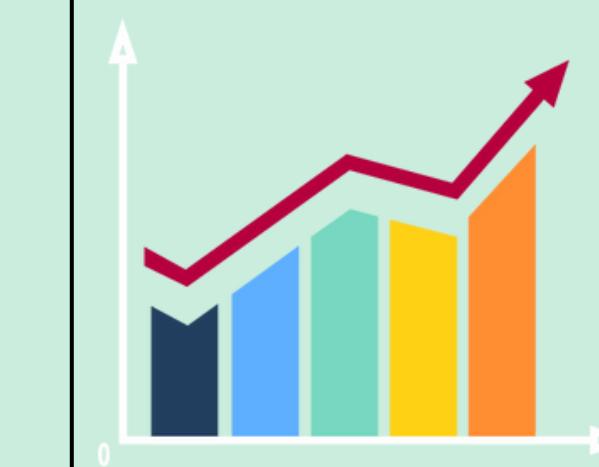
Valores o categorías: Son los diferentes valores o categorías que se están analizando en la variable.



Frecuencia absoluta: Es el número de veces que cada valor o categoría ocurre en el conjunto de datos.



Frecuencia relativa: Es la proporción o porcentaje de veces que cada valor o categoría ocurre en el conjunto de datos.



Frecuencia acumulada: Es la suma acumulada de las frecuencias absolutas o relativas a medida que se avanza a través de los valores o categorías en la tabla.

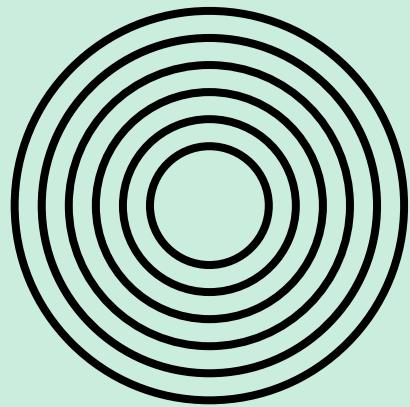
Valor	Frecuencia Absoluta	Frecuencia Absoluta Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada	Frecuencia Relativa (%)	Frecuencia Relativa Acumulada (%)
1	5	5	0.25	0.25	25%	25%
2	8	13	0.40	0.65	40%	65%
3	3	16	0.15	0.80	15%	80%
4	2	18	0.10	0.90	10%	90%
5	6	24	0.30	1.00	30%	100%

Tablas de contingencia



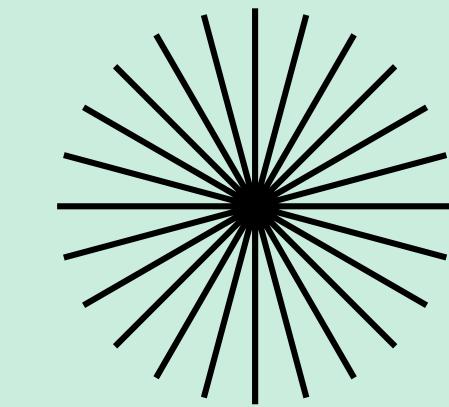
¿Qué es?

Es una herramienta utilizada en la rama de la estadística, la cual consiste en crear al menos dos filas y dos columnas para representar datos categóricos en términos de conteos de frecuencia.



Objetivo

Ordenar la información recolectada para un estudio cuando los datos se encuentran divididos de forma bidimensional, esto significa a que se relaciona con dos factores cualitativos.



Importancia

La tabla ofrece un método simple de agrupar variables, que minimiza el potencial de confusión o error al proporcionar resultados claros.

Preferencia	Hombres	Mujeres
Perros	36	22
Gatos	8	26
Sin preferencia	2	6

Medidas de tendencia central

Las medidas de tendencia central describen el centro de un conjunto de datos. Miden lo que es típico y pueden representarse mediante la media, la moda o la mediana, según la asimetría de los valores. La tendencia central ayuda a identificar cuándo un nuevo valor se aleja considerablemente de la distribución normal o a resumir rápidamente una gran cantidad de valores de un período o grupo específico.



Medidas de tendencia central



Media: es el número "promedio"; se encuentra al sumar todos los datos y dividir entre el número de datos.

Mediana: es el número de en medio; se encuentra al ordenar todos los puntos de datos y elegir el que está en medio (o si hay dos números en medio, tomar la media de esos dos números).

Moda: es el número más frecuente, es decir, el número que se repite el mayor número de veces.

Medidas de variabilidad

Los estadísticos de tendencia central o posición nos indican donde se sitúa un grupo de puntuaciones. Los de variabilidad o dispersión nos indican si esas puntuaciones o valores están próximas entre sí o si por el contrario están o muy dispersas.



Medidas de tendencia central



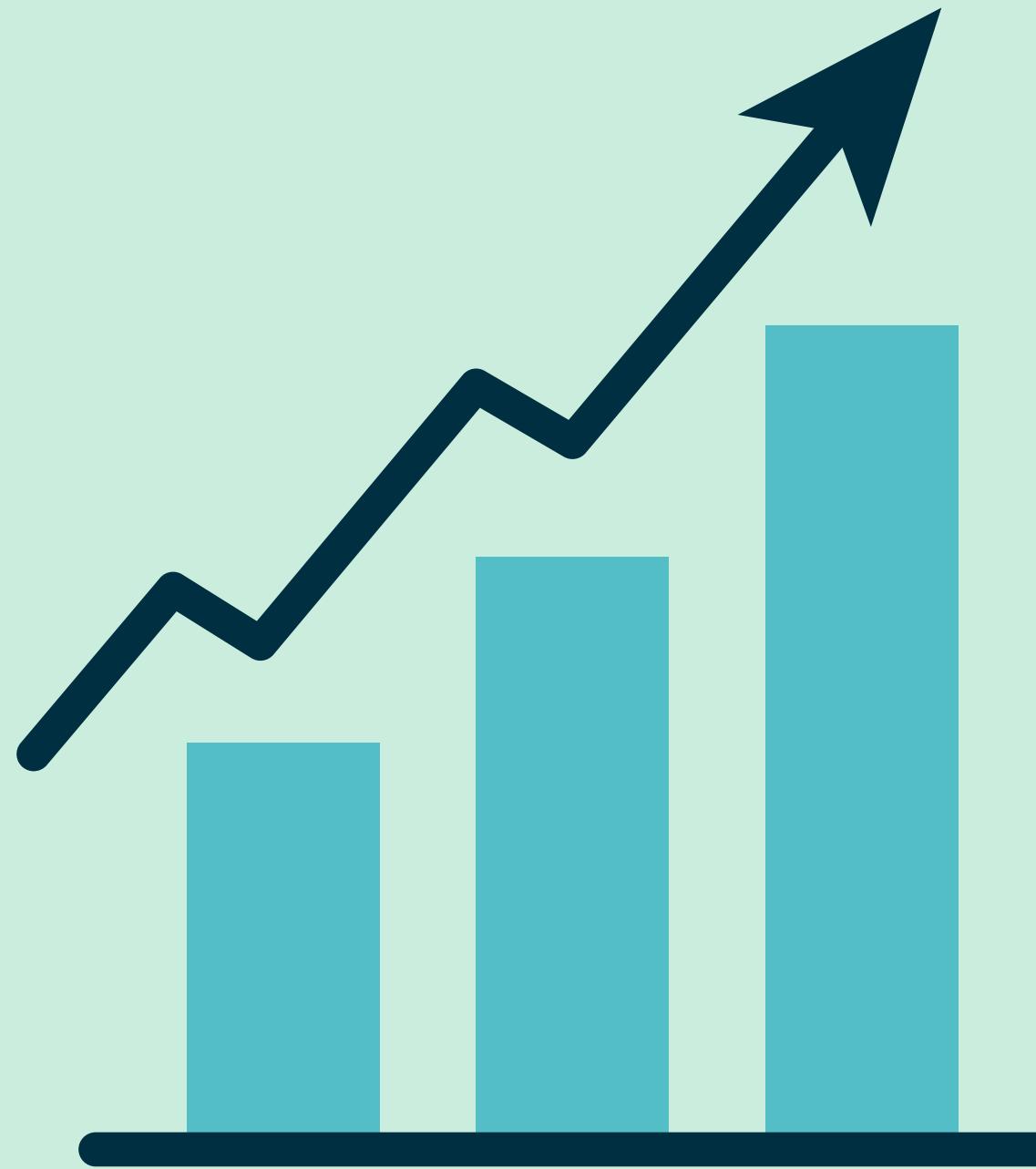
Rango: Es la diferencia entre el valor máximo y el valor mínimo del conjunto de datos.

La varianza: Es el promedio de las diferencias al cuadrado de cada dato respecto a la media. Indica qué tanto se alejan los datos en promedio de la media.

Desviación estándar: Es la raíz cuadrada de la varianza. Proporciona una medida de la distancia promedio de cada dato con respecto a la media y se expresa en las mismas unidades que los datos originales.

Cuartil

Separa los valores en cuatro grupos iguales y existen tres cuartiles. Es el más frecuente. El cuartil uno (Q1) son los datos menores y el tres (Q3) los mayores. Por otro lado, el cuartil dos (Q2) se corresponde con la mediana (M_e) que es un estadístico de posición que divide la distribución de los datos a la mitad. Los valores del cuantil serían 0.25 (Q1), 0.5(Q2) y 0.75 (Q3).



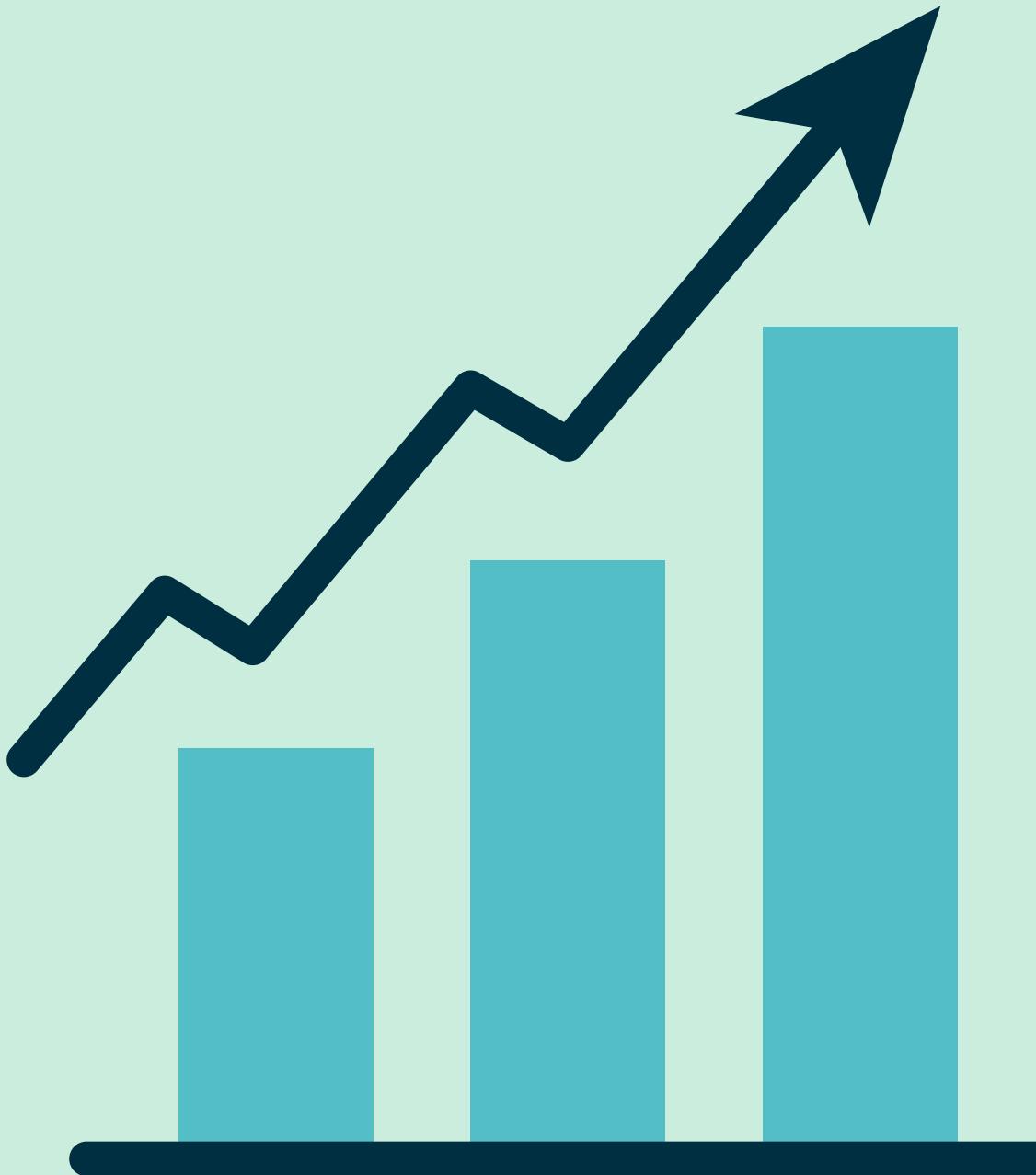
Decil

En este caso se dividen en diez partes y, por tanto, hay nueve deciles. Una vez más, este tampoco es demasiado frecuente. Sus valores serían de 0.1 a 0.9.

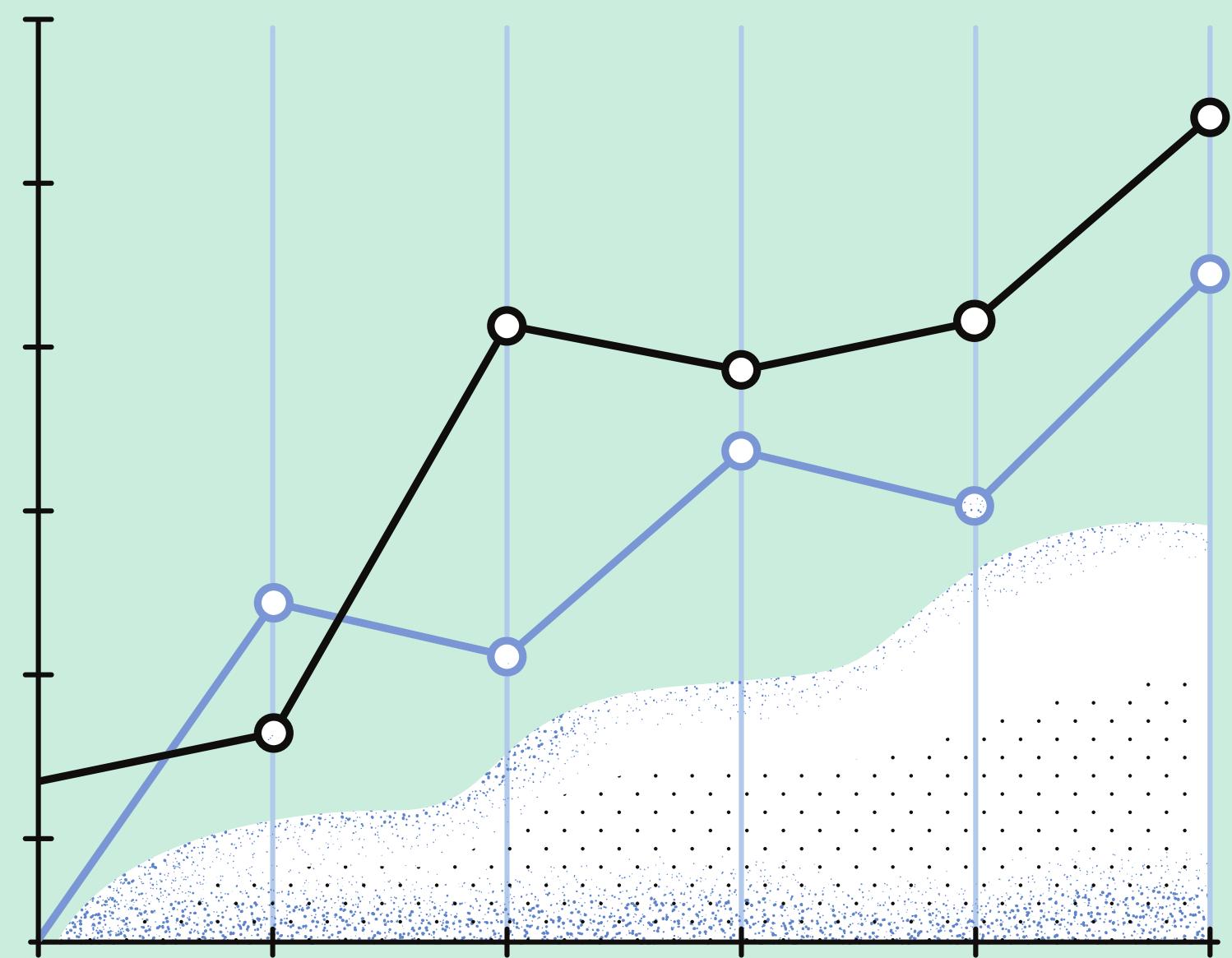


Percentil

Estamos ante una variante en que la distribución se divide en cien partes iguales. Puede ser de interés para muestras muy numerosas. Sus valores van de 0.01 a 0.99.



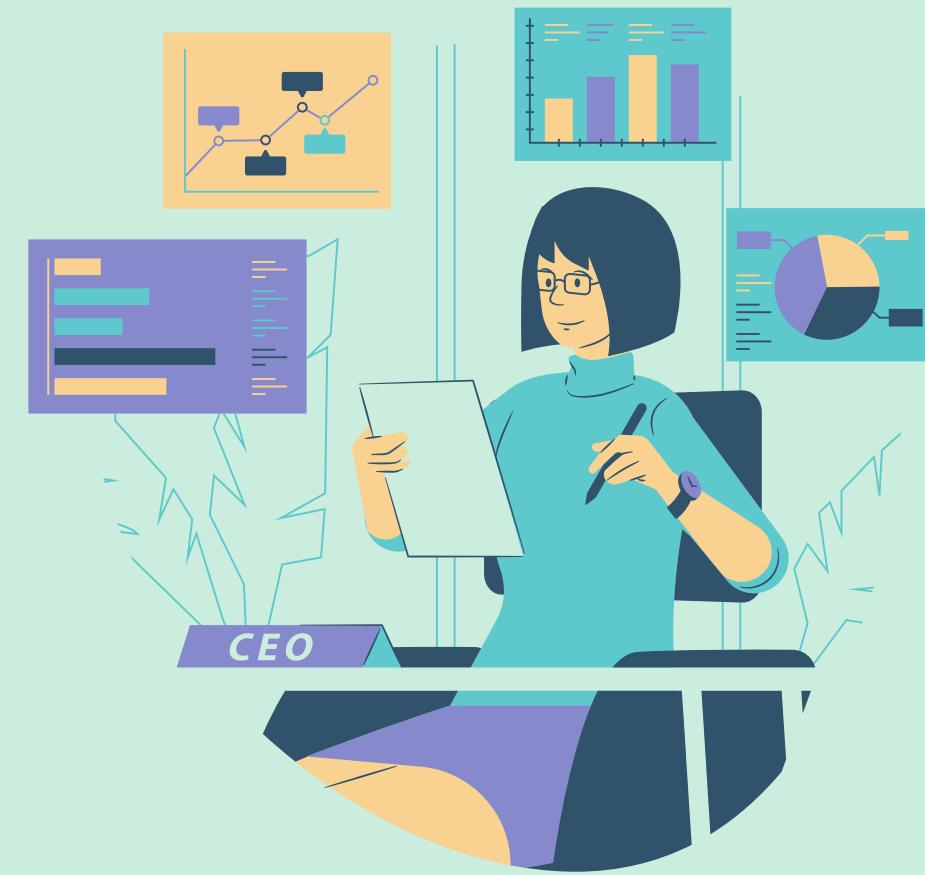
Coeficiente de variación: Es una medida estadística de dispersión relativa que compara la desviación estándar con la media de un conjunto de datos. Se calcula dividiendo la desviación estándar entre la media



Importancia de la Visualización en Estadística

La visualización de datos es una parte fundamental de la estadística descriptiva porque permite.

Una gráfica clara puede transmitir en segundos lo que un análisis numérico tardaría párrafos en explicar.



01

Comprender patrones y tendencias no visibles en tablas numéricas.

02

Identificar valores extremos, errores o datos atípicos.

03

Comparar grupos y distribuciones. Explorar relaciones entre variables.

04

Comunicar resultados de forma clara, intuitiva y profesional.

Tipos de Gráficos y Cuándo Usarlos

Tipo de variable	Objetivo	Gráfico recomendado
Cualitativa	Comparar categorías	Barras, pastel (menos recomendado), treemap
Cuantitativa	Analizar distribución	Histograma, boxplot, violín
Dos cuantitativas	Evaluar relación	Dispersión (scatterplot)
Cualitativa + Cuantitativa	Comparar grupos	Boxplot, violin plot, barras agrupadas

Conceptos de Distribución de Datos (visión gráfica)

Antes de ver correlación, es importante entender el comportamiento de una distribución cuantitativa:

- Simétrica
- Sesgada a la derecha (positiva)
- Sesgada a la izquierda (negativa)
- Unimodal / multimodal
- Con presencia de outliers

Estos rasgos son más fácilmente identificables usando histogramas y boxplots.

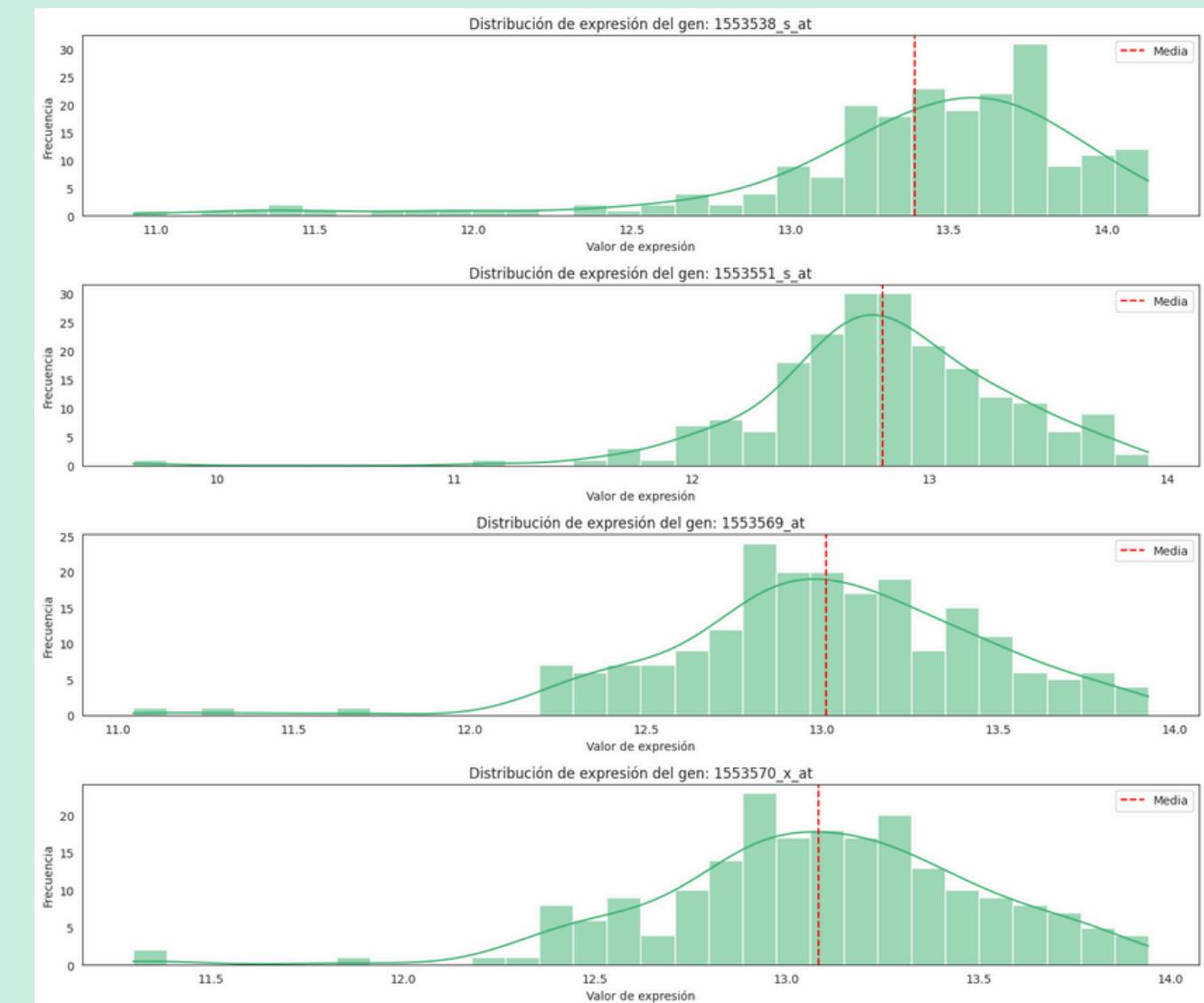


Histogramas

- Permiten visualizar la frecuencia de valores en intervalos.
- Son útiles para estudiar la forma, centro y variabilidad de una distribución.

Aspectos clave a interpretar:

- ¿Cuál es el intervalo más frecuente?
- ¿La distribución es simétrica o sesgada?
- ¿Hay valores atípicos?
- ¿La dispersión parece amplia o concentrada?



Boxplot (Diagrama de Caja y Bigotes)

- El boxplot resume una variable cuantitativa con:
- Mediana (línea central)
- Q1 y Q3 (caja)
- Rango intercuartílico (IQR)
- Líneas (“bigotes”) hasta $1.5 \times \text{IQR}$
- Puntos aislados = posibles outliers

Se utiliza para:

- Identificar valores atípicos
- Comparar grupos
- Evaluar simetría y variabilidad

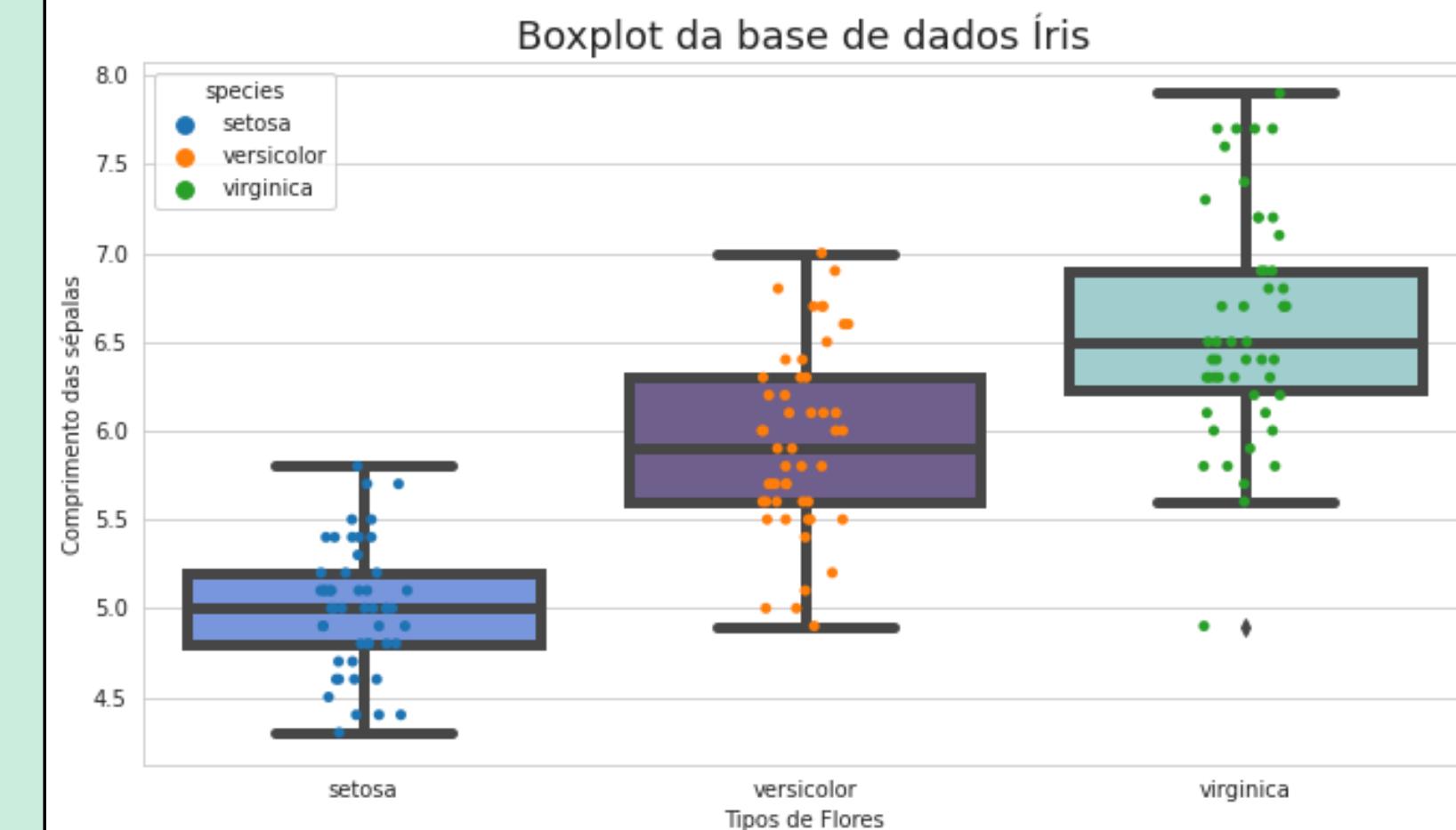
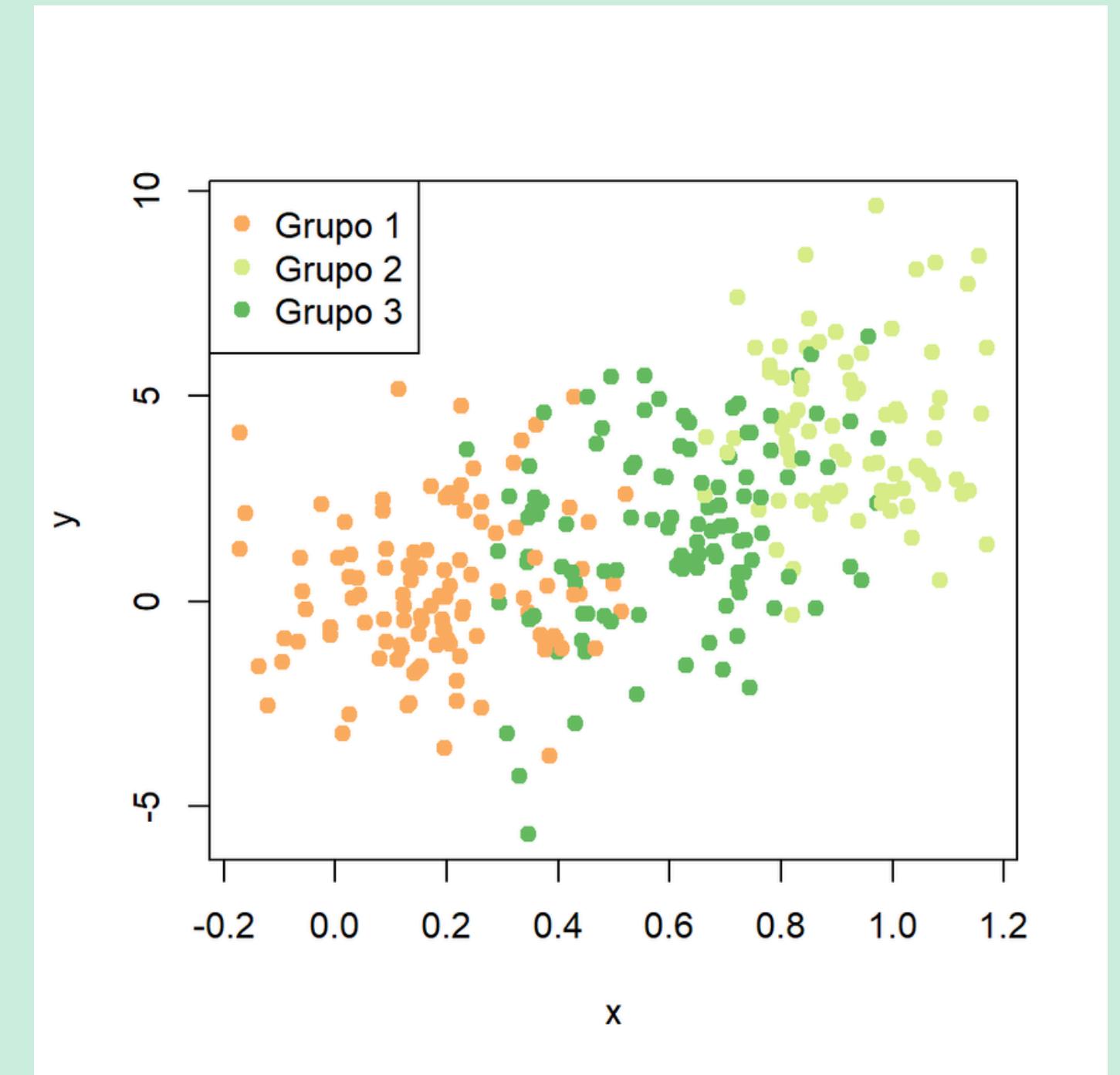


Diagrama de Dispersión (Scatterplot)

- Utilizado para examinar la relación entre dos variables cuantitativas.
- Cada punto representa un individuo u observación.
- Posible incluir línea de tendencia (recta de regresión) para evaluación visual.
- Se interpreta:
- Forma (lineal o no lineal)
- Dirección (positiva o negativa)
- Intensidad (qué tan agrupados están los puntos)
- Outliers o patrones inusuales





**¡¡¡MUCHAS
GRACIAS!!!**

