

Estadística

@ferstudiesss_

Estadística

Población

conjunto de todos los elementos a estudiar

Muestra

subconjunto representativo, significativo y aleatorio de la Población

Variable

característica a estudiar, propia de la Población

• **cualitativa:** no se pueden expresar en números

nominal: no tiene orden

ordinal: existe un orden

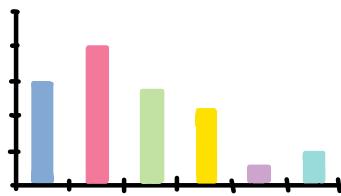
• **cuantitativa:** se expresa mediante un número

discreta: siempre valores exactos

continua: puede tomar valores decimales

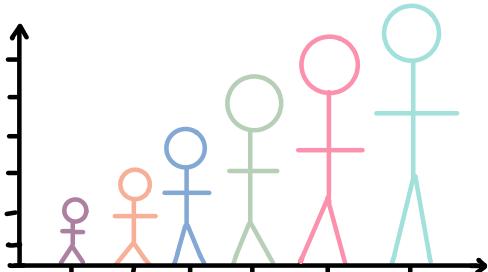
GRÁFICOS

gráfico de barras



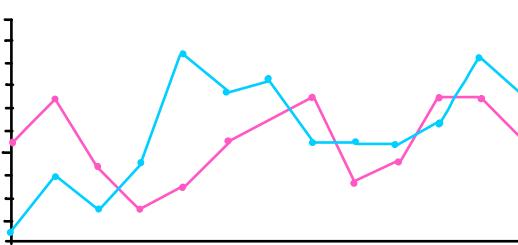
representa la Frecuencia mediante barras, de datos cualitativos o cuantitativos discretos

pictograma



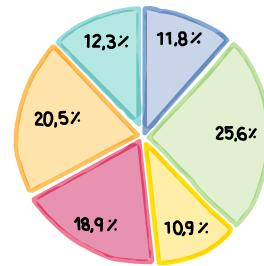
representan la Frecuencia mediante figuras en vez de barras

polígono de frecuencias



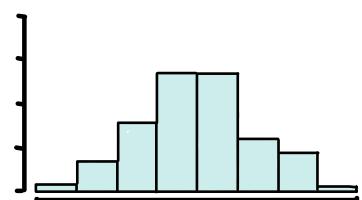
representa datos que han sido agrupados en intervalos, formando un polígono al unir los pares ordenados

gráfico circular



representa la PROPORCIÓN de los datos en porcentajes

histograma



representa datos que han sido agrupados en intervalos

TABLAS DE FRECUENCIAS

X variable o intervalo	X_i marca de clase	f_i Frecuencia absoluta	F_i Frecuencia absoluta acumulada	h_i Frecuencia relativa	H_i Frecuencia relativa acumulada
datos obtenidos en el estudio, pueden estar agrupados en intervalos o cada uno separado	en el caso de que existan intervalos, equivale al dato medio de cada uno, es decir la suma de los límites divididos en dos	es la cantidad de veces que se repite cada dato, la suma de todas es el total de datos que hay	es la suma de las frecuencias de los datos anteriores y del dato actual	es la frecuencia absoluta dividida en el total de datos, se puede expresar en porcentaje o en número decimal, su suma debe dar uno	es la suma de las frecuencias relativas anteriores y del dato actual

MEDIDAS DE TENDENCIA CENTRAL

Moda

es el dato que más se repite, es decir el que tiene mayor frecuencia

DATOS NO AGRUPADOS

Observar todos los datos obtenidos, aquel que más se repite corresponde a la moda

DATOS AGRUPADOS

$$Mo = Li + ai \cdot \frac{d_1}{d_1 + d_2}$$

el intervalo modal es el que tiene mayor frecuencia absoluta

Li = límite inferior del intervalo modal

ai = amplitud del intervalo, es la resta del límite superior e inferior

d_1 = la Frecuencia modal menos la Frecuencia anterior

d_2 = la Frecuencia modal menos la Frecuencia posterior

Mediana

es el valor central, es decir el que ocupa el lugar central de todos los datos cuando estos están ordenados

DATOS NO AGRUPADOS

ordenar los datos de menor a mayor, si el total de datos es impar, la mediana es el dato central, si el total es par, la mediana es la suma de los dos datos centrales dividido en dos

DATOS AGRUPADOS

$$Me = Li + ai \cdot \frac{\frac{N}{2} - F_{i-1}}{f_i}$$

$\frac{N}{2}$ = el total de datos dividido en dos, el resultado se debe ubicar en las F y ese será el intervalo a observar

Li = límite inferior del intervalo

ai = amplitud del intervalo, es la resta de los ambos límites

F_{i-1} = Frecuencia absoluta acumulada anterior

f_i = Frecuencia absoluta del intervalo

Media aritmética

es el promedio de cada dato respecto al total

DATOS NO AGRUPADOS

sumar todos los datos y dividirlo en el número total de datos

DATOS AGRUPADOS

$$\bar{x} = \frac{\sum X_i \cdot f_i}{N}$$

Σ = suma de

X_i = marca de clase

f_i = Frecuencia absoluta

N = total de datos

Rango

es una medida de dispersión de los datos

$$R = X_{\max} - X_{\min}$$

↓ ↓
dato dato

mayor menor

MEDIDAS DE POSICION

Las medidas de posición relativa dividen datos ordenados de menor a mayor en partes iguales

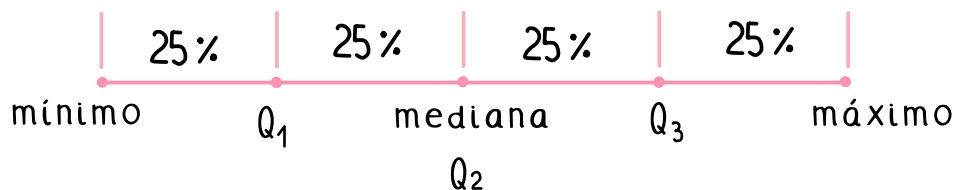
Cuartiles

divide los datos en cuatro partes iguales que equivalen al 25% cada una

Q_1 el primer cuartil corresponde al 25%

Q_2 el segundo cuartil corresponde al 50%

Q_3 el tercer cuartil corresponde al 75%



$$Q_K = \frac{K \cdot N}{4}$$

$K \rightarrow$ número del cuartil

$N \rightarrow$ total de datos

↓
el resultado es la
posición del cuartil K
↓
frecuencia acumulada

$$P_K = \frac{K \cdot N}{100}$$

$K \rightarrow$ número del Percentil

$N \rightarrow$ total de datos

↓
el resultado es la
posición del Percentil K
↓
frecuencia acumulada

Percentiles

divide los datos en cien partes iguales que equivalen al 1% cada una

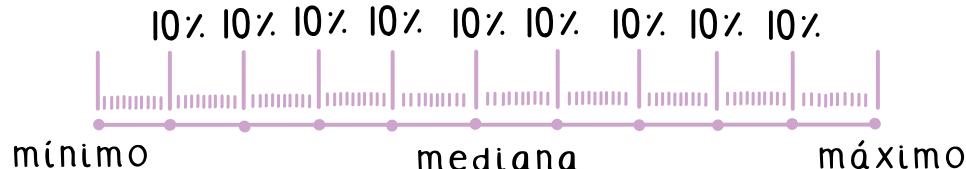
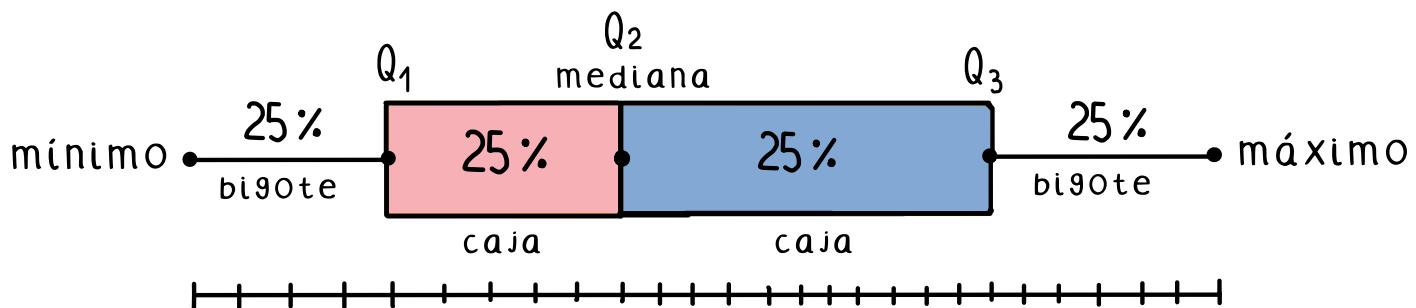


Diagrama de caja

representa la distribución Y dispersión de los datos ordenados de menor a mayor, escritos sin repetición



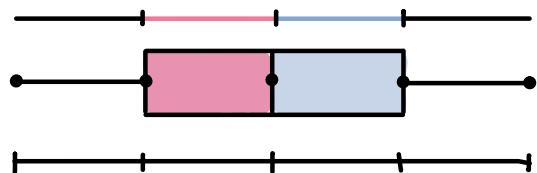
➡ RANGO = dato máximo – dato mínimo

mientras mayor sea el rango, más dispersos o separados están los datos

➡ RANGO INTERCUARTIL = cuartil 3 – cuartil 1

CASO 1

distribución simétrica

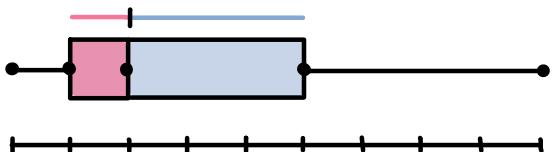


los bigotes son iguales Y la mediana se encuentra al centro de la caja

media = mediana = moda

CASO 2

asimétrica positiva



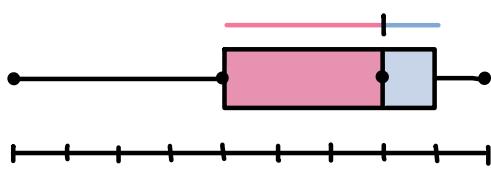
el Primer bigote es más pequeño que el segundo Y la mediana está más cerca del Primer cuartil

media > mediana > moda

los datos están más separados hacia la derecha de la caja

CASO 3

asimétrica negativa



el Primer bigote es más grande que el segundo Y la mediana está más cerca del tercer cuartil

media < mediana < moda

los datos están más separados hacia la izquierda de la caja