

## 1.1 Pensamiento Estadístico:

### •Tipos de Variables

En estadística se encuentran dos tipos de variables las Cualitativas y Cuantitativas a su vez cada una compone dos sub tipos.

#### Cualitativas (Cualidades): Nominal y Ordinal

--Nominal; No cuentan con orden o distinción entre los datos recabados, no hay un mas o menos.  
--Ordinal; Con algún tipo de Orden se distinguen los datos de mayor o menor importancia, de un antes o un después.

#### Cuantitativas (Cuantificables): Discretas y Continuas

--Discretas; Son datos numéricos enteros o únicos, no hay un intermedio entre ello  
--Continuas; Datos provenientes de medidas tienden a ser infinitos respecto al nivel precisión del instrumento utilizado al obtener las muestras.



### •Tipos de Muestra

Las muestras son un segmento de la **Población** y estas pueden ser obtenidas de dos formas Probabilística y no Probabilística que a su vez están segmentadas en varios tipos

#### Probabilístico (todos los elementos son igual de probables):

--Aleatorio simple; Requiere un **marco muestral** (una lista completa de la población) donde al azar son selectos.  
--Sistemático; El primer individuo es selecto al azar y de el parten los siguientes por intervalos regulares.  
--Estratificado; Dividiendo a la población por características similares (edad,sexo,etnia,etc).  
--Conglomerado; Tomando grupos de la población ya agrupados por la estructura de la sociedad o el entorno.

#### No Probabilístico (los elementos no comparten igual probabilidad):

--Bola de Nieve; El individuo te lleva a otro (principalmente usado cuando el sujeto es difícil o de muy selecto acceso).  
--Intencional; Donde se busca una característica muy especifica en la muestra.  
--Conveniencia; Se eligen a los que tienen mas cerca.  
--Cuotas; Similar al estratificado, pero la elección de los individuos dentro del estrato no es azarosa, sino a juicio del investigador.



● **Medidas de tendencia central**

Describiendo el centro de los datos, surgen 3 métodos que dependiendo de la situación son en mayor medida eficaces en su uso, son las siguientes, Media, Mediana y Moda.

**1, 5, 2, 7, 2, 4, 3**

**Media**  $\frac{1+5+2+7+2+4+3}{7} = 3,43$

**Mediana** 1, 2, 2, **3**, 4, 5, 7

**Moda** 1, **2**, **2**, 3, 4, 5, 7

**Media;** Es el proceso de obtener un promedio, surgiendo de sumar todas las variables cuantitativas para luego dividirlo por el numero de variables

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

“La media (o promedio) de X es la suma de todos los valores  $X_i$  desde el primero hasta el n-ésimo, entre n.”

**Mediana;** Se busca lo que esta en medio de las variables agrupadas de manera Ordinal, en caso de que las variables sean pares en su totalidad las dos variables del centro se les busca su Media (promedio).

**MEDIANA (Me)**

Halla la mediana de los siguientes datos:  
7, 8, 9, 10, 11, 12, 6, 5

**1. Ordenar Datos**

De menor a mayor: ~~5~~, ~~6~~, ~~7~~, **8, 9**, 10, 11, 12

↑ ó ↓  
(Elige solo una forma de ordenar los datos).

De mayor a menor: 12, 11, 10, **9, 8**, ~~7~~, ~~6~~, ~~5~~

El Total de datos es par.  
Hay 8 datos.

**2. Tachar extremos hasta dejar dos datos en el centro.**

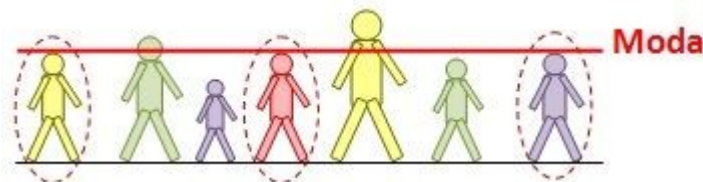
**3. Sumar valores centrales**

$8 + 9 = 17$   
ó  
 $9 + 8 = 17$

**4. Dividir ÷ 2**

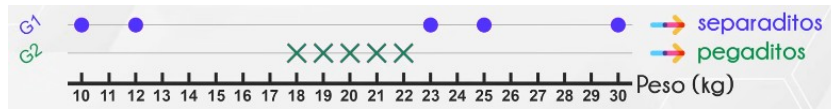
17	2
10	8,5
0	0

**Moda;** Es la o las variables de mayor aparición, la mas repetida, **la Moda es la medida "todoterreno"** porque es la que requiere menos condiciones para ser utilizada.



### ●Medidas de dispersión (variabilidad)

Nos permite observar que tan dispersos o separados están los datos de unos a otros, medidas estadísticas que muestran la variabilidad en la distribución de los datos.



**Rango (Amplitud);** es la medida mas simple, es la diferencia del dato mas grande entre el mas pequeño.



$$R = x_{max} - x_{min}$$

**Varianza;**Es expresada en unidades al cuadrado<sup>2</sup> por lo que es difícil de interpretar, pero es la medida que deriva a la Desviación estándar.

¿Quieres saber mas?

---**Poblacional ( $\sigma^2$ ):**"Para hallar la varianza, toma **cada dato** ( $x_i$ ), réstale el **promedio** ( $\mu$ ), eleva ese resultado al **cuadrado** para que no haya negativos, **suma** todos esos resultados y, finalmente, saca el **promedio** dividiendo entre el total de datos ( $N$ )."

Población

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

---**Muestra ( $s^2$ ):**Es casi igual a la anterior, pero usamos el promedio de la **muestra** ( $\bar{x}$ ) y dividimos por **n-1** para ser más conservadores y compensar el error de no tener todos los datos."

Muestra

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

**Desviación estándar ( $\sigma$ );** Es la raíz de la Varianza, lo que permite pasar de unidades al cuadrado a solo unidades.

---Poblacional

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

---Muestra

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

**Desviación Media(DM);**

Es el Promedio de las distancias.

$$DM = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

### Coeficiente de Variación (CV)

Esta medida de dispersión nos permite comparar grupos que tienen unidades de medida diferentes, presentando en porcentaje el cociente de la Desviación estándar entre su media.

Coeficiente de variación:

Población	Muestra
$CV = \frac{\sigma}{\mu}$	$CV = \frac{s}{\bar{x}}$

Coeficiente de variación en forma porcentual:

Población	Muestra
$CV = \frac{\sigma}{\mu} \cdot 100\%$	$CV = \frac{s}{\bar{x}} \cdot 100\%$