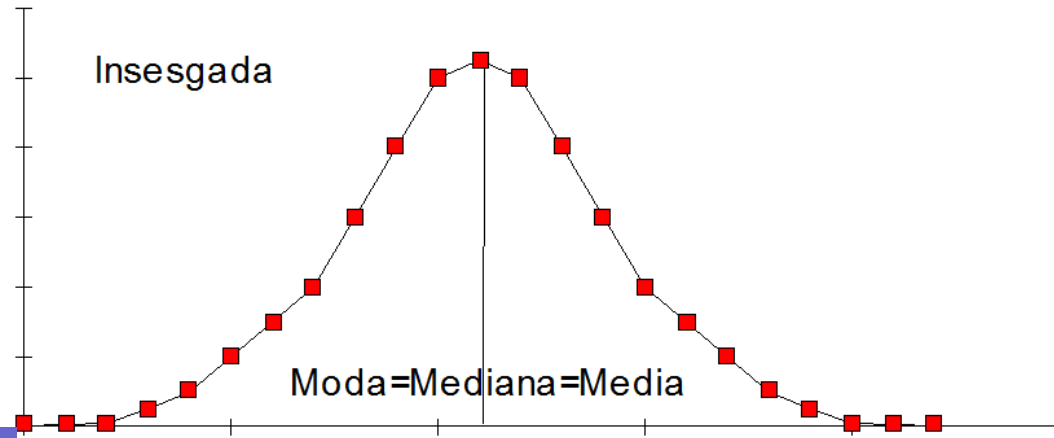
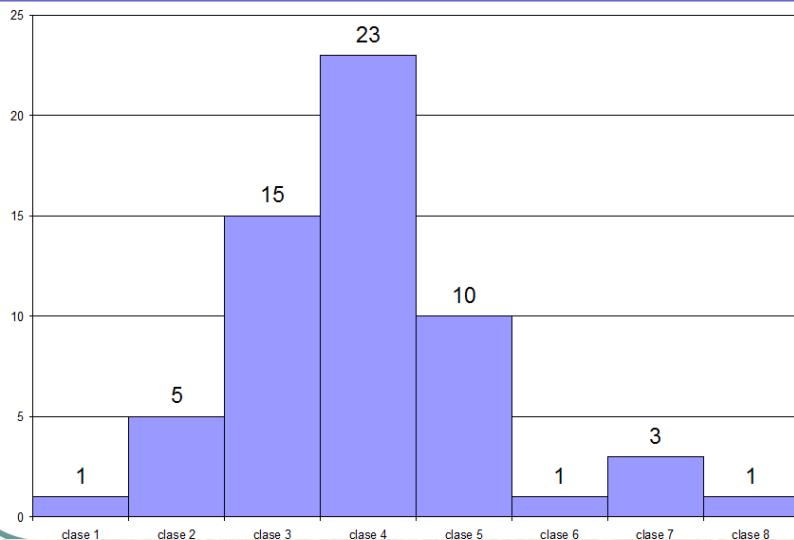


# Medidas tendencia central (MTC)

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$



## Histograma



$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

# Contenido

---

## # Introducción

## # Media

- datos no agrupados
- datos agrupados
- Otras medias: ponderal, geométrica, armónica y cuadrática

## # Mediana

- datos no agrupados
- datos agrupados

## # Moda

- datos no agrupados
- datos agrupados

## # Percentiles

---

# Posición ó localización

---

La característica más importante que describe o resume un grupo de datos **es su posición.**

Los conjuntos de datos muestran una tendencia definida a agruparse o resumirse en **torno a cierto punto**, por lo que para cualquier conjunto particular de datos, es posible seleccionar un **valor típico para describir, representar o resumir todo el conjunto de datos.**

De acuerdo a la organización de los datos hay dos formas de estimar este tipo de medidas ya sea **para datos no agrupados o agrupados**

---

# Propiedades deseables

---

- # Definida objetivamente a partir de los datos de la serie.
  - # Que dependa de todas las observaciones.
  - # De significado sencillo y fácil de entender.
  - # De cálculo rápido y fácil.
  - # Poco sensible a las fluctuaciones del muestreo (valor parecido al de la población)
  - # Adecuado a cálculos algebraicos posteriores.
-



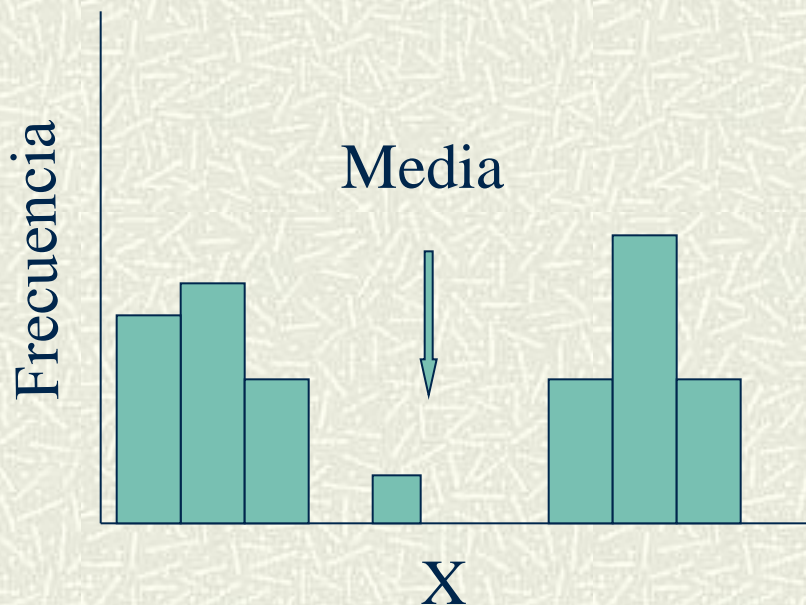
# Objetivo

---

- # El objetivo es resumir en un solo punto numérico la posición o localización de los datos que estamos analizando en la distribución.
  - # Se llaman MTC debido a que se ha procedido a escoger como valores típicos o medidas de posición a aquellos que están situados en la parte central de la distribución, por eso reciben el nombre de medidas de tendencia central.
-

# Media aritmética o promedio

- # Es la medida más común de localización y representa el **centro de un grupo de datos**.
- # El valor obtenido es más preciso que la precisión asociada con cada observación por lo cual siempre se representa con **un dígito más** que los utilizados en la medición



# Procedimiento de cálculo:

---

- # Paso 1. Se suman los datos individuales desde 1 hasta n o N
- # Paso 2. Se divide el tamaño de la muestra o la población

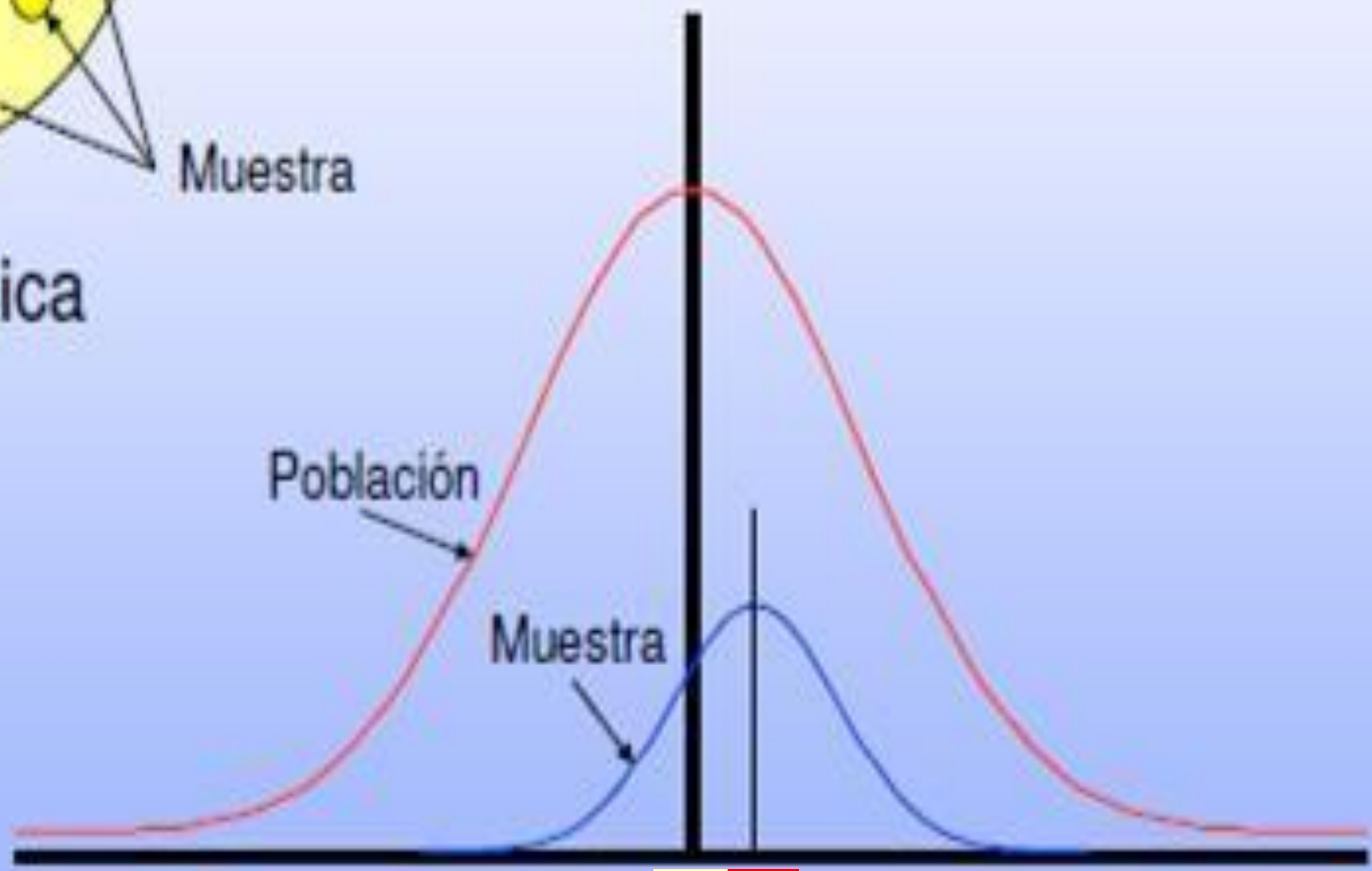
$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

$$\mu = \frac{\sum_{i=1}^N X_i}{N}$$

---



Relación física



Relación  $\mu$   $\bar{x}$  estadística



# Ejemplo

**Ingresos mensuales en dólares**

<b>1000</b>	<b>1110</b>	<b>1010</b>	<b>1070</b>	<b>1030</b>	<b>1000</b>
<b>1150</b>	<b>990</b>	<b>1090</b>	<b>1080</b>	<b>1150</b>	<b>1200</b>
<b>1050</b>	<b>1030</b>	<b>1120</b>	<b>1050</b>	<b>1030</b>	<b>1150</b>
<b>1230</b>	<b>1170</b>	<b>1180</b>	<b>1110</b>	<b>1160</b>	<b>1100</b>
<b>1100</b>	<b>1060</b>	<b>1130</b>	<b>1105</b>	<b>935</b>	<b>1210</b>

$$\bar{X} = \frac{\sum_{i=1}^{30} X_i}{30} = \frac{1000 + 1150 + 1050 + 1230 + 1100 + 1110 + 990 + 1030 + 1170 \cdots + 1210}{30}$$

$$\bar{X} = \frac{32800}{30} = 1.093,33$$

## Procedimiento de cálculo:

---

- # Paso 1. Definir las clases y determinar el punto medio de cada una
- # Paso 2. Determinar la frecuencia absoluta de cada clase

$$\bar{X} = \sum \frac{X_k * n_k}{n}$$

$X_k$  = punto medio de la clase

$n_k$  = frecuencia de la clase k

$n$  = tamaño de la muestra

---

# Datos agrupados

## Ejemplo

Limites reales de clase		MARCA DE CLASE	FRECUENCIA ABSOLUTA	FRECUENCIA ABSOLUTA ACUMULADA	FRECUENCIA RELATIVA	FRECUENCIA RELATIVA ACUMULADA
Li	Ls	$X_k$	$n_k$	$N_k$	$n_k/n$	$N_k/n$
930	980]	955	1	1	1/30	1/30
980	1030	1005	7	8	7/30	8/30
1030	1080	1055	5	13	5/30	13/30
1080	1130	1105	8	21	8/30	21/30
1130	1180	1155	6	27	6/30	27/30
1180	1230	1205	3	30	3/30	30/30=1
			30		30/30=1	

## Datos agrupados

Calcular la media aritmética

---

$$\bar{X} = \frac{(955*1) + (1005*7) + (1055*5) + (1105*8) + (1155*6) + (1205*3)}{30}$$

$$\bar{X} = 1088,33$$

La media aritmética es 1088,33

---



# Mediana ( $M_e$ )

---

La **mediana**  $M_e$  de un conjunto de datos  $x_1, x_2, \dots, x_n$ , es el valor  $x_i$  que se encuentra en el punto medio o centro, cuando se ordenan los valores de menor a mayor.

La interpretación geométrica de la mediana, es que es el valor que divide un histograma en dos partes iguales.

---

# Procedimiento de cálculo: Mediana

- # Paso 1.- Ordenar de menor a mayor los valores  $x_i$  del conjunto de datos individuales,  $i = 1, 2, \dots, n$
- # Paso 2.- Identificar si  $n$  es impar o par

$$x = \begin{cases} x_{([n+1])/2} & \text{impar} \\ \frac{x_{(n/2)} + x_{(n/2+1)}}{2} & \text{par} \end{cases}$$

# Ejemplo

## Mediana

Encontrar la mediana del siguiente conjunto de datos que corresponden al tiempo en segundos, requerido por una cajera para marcar la compra de artículos en un supermercado que utiliza verificadores automáticos

{ 10, 15, 62, 53, 11, 38, 75, 112, 40, 22, 57 }.

■ Paso 1: Ordenar el conjunto de datos:

{ 10, 11, 15, 22, 38, 40, 53, 57, 62, 75, 112 }.

■ Paso 2: Identificar  $n = 11$ , impar

## Mediana

# **Paso 3:** Como  $n = 11$  es impar, la mediana  $Me$  es:

{	10,	11,	15,	22,	38,	40,	53,	57,	62,	75,	112	}	Datos
	1	2	3	4	5	6	7	8	9	10	11		Indice

$$m = x_{\frac{n+1}{2}}$$

$$\frac{n+1}{2} = \frac{11+1}{2} = \frac{12}{2} = 6$$

$$x_{\frac{n+1}{2}} = x_6 = 40$$

$$m = x_6 = 40$$



# Procedimiento de localización y cálculo:

Datos agrupados

Mediana

- # En el caso de que los datos estén agrupados, se obtiene primero la **clase de la mediana**
- # La **clase de la mediana** se define como la primera clase que aparece en la tabla, para la cual la frecuencia acumulada,  $N_k$ , sea igual o mayor a la mitad de la suma de todas las frecuencias absolutas, esto es:

- ▣ **Paso 1.-** Obtener la Clase de la Mediana, es decir, el primer intervalo que cumpla la condición:

$$N_k \geq \frac{n}{2}$$

donde  $n$  - es el número total de datos del conjunto

$k$  - es el número del intervalo de clase que cumple la condición,  $j = 1, 2, \dots, k$

$N_k$ - es la frecuencia acumulada del intervalo de clase  $k$

# Paso 2.- Calcular la mediana con la siguiente ecuación:

$$\# \text{Me} = L_i + \frac{[(n/2) - N_{k-1}]}{n_k} * i$$

**Donde:**

$L_i$  = Límite inferior del intervalo que corresponde a la **clase mediana**.

$n$  = Total de datos

$N_{k-1}$  = Frecuencia acumulada de la clase anterior a la clase donde esta la mediana

$n_k$  = Frecuencia absoluta donde esta la mediana.

$i$  = Tamaño del intervalo de clase (amplitud o distancia del intervalo)

# Datos agrupados

## Ejemplo

Me



Limites reales de clase		MARCA DE CLASE	FRECUENCIA ABSOLUTA	FRECUENCIA ABSOLUTA ACUMULADA	FRECUENCIA RELATIVA	FRECUENCIA RELATIVA ACUMULADA
Li	Ls	$X_k$	$n_k$	$N_k$	$n_k/n$	$N_k/n$
930 -	980]	955	1	1	0,033	0,033
980 -	1030	1005	7	8	0,233	0,266
1030 -	1080	1055	5	13	0,166	0,432
1080 -	1130	1105	8	21	0,266	0,698
1130 -	1180	1155	6	27	0,200	0,898
1180 -	1230	1205	3	30	0,100	0,998
			30		0,998	



## #Paso 1.- Obtener la Clase de la Mediana

$$n = 30 \quad \text{y} \quad n/2 = 15,$$

$$k = 6 \text{ intervalos de clase}$$

En el **4 intervalo** se cumple que:

$$F_4 = \mathbf{21} \geq 15.$$

Por tanto el **cuarto intervalo** se identificará como la clase mediana, esto es:

**[1080, 1130]** es el intervalo de clase.

$$F_4 = \mathbf{21}$$

# Datos agrupados

## # Paso 2.- Calcular la mediana:

$L_i = 1080$  Limite inferior de la clase de la mediana

$n = 30$  Datos u observaciones

$n/2 = 15$  Datos u observaciones

$N_k = N_3 = 13$  Frec.acumul.absol anterior

$n_k = n_4 = 8$  Frec. Absoluta de la clase de la mediana

$i = L_{sup} - L_{inf} = 50$

La interpretación de este valor, es que la mitad de las familias entrevistadas ganan menos de 1.092,50 dólares.

No.	Limite real clase	$n_k$	$N_k$
3	(1030,1080]	5	13
4	(1080,1130]	8	21

$$Me = 1080 + \frac{[15 - 13]}{8} * 50$$

$$Me = 1080 + 12,50$$

$$Me = 1092,50$$

# Moda ( $M_o$ )

---

- # Es la observación que se presenta con mayor frecuencia en la muestra o población
  - # Si los datos tienen una sola moda son unimodales, si tienen 2 son bimodales y así sucesivamente
  - # No son comparables en su totalidad los resultados obtenidos entre datos individuales y agrupados
-

# Procedimiento de cálculo:

---

- # **Paso 1.** Ordenar los datos en forma ascendente
- # **Paso 2.** Identificar el o los datos con mayor frecuencia



# Ejemplo

## Moda

En el conjunto de datos dados para medir el ingreso promedio anual de las familias que habitan en un determinado sector, se tiene:

**{ 935, 990, 1000, 1000, 1010, 1030, 1030, 1030, 1050, 1050, 1060, 1070, 1080, 1090, 1100, 1100, 1105, 1110, 1110, 1120, 1130, 1150, 1150, 1150, 1160, 1170, 1180, 1200, 1210, 1230 }**

En este conjunto hay dos valores que aparecen más veces :

**1030 ( 3 veces)**

**1150 ( 3 veces)**

**por tanto la moda es = 1030 y 1150**

y en este caso se dice que el conjunto de datos es **bimodal**

## Procedimiento de cálculo: **Moda**

**Paso 1.** La Moda estará representada por la clase que posee la más alta frecuencia, denominándose **clase modal**.

**Paso 2.** El cálculo de la Moda se obtiene con la siguiente expresión:

$$M_o = L_i + \frac{[d_1]}{d_1 + d_2} * i$$

Donde:

$L_i$  = Límite inferior del intervalo de clase, **la clase modal**.

$d_1$  = diferencia de la frecuencia de la clase modal y la clase anterior a la clase modal

$d_2$  = diferencia de la frecuencia de la clase modal y la frecuencia posterior a la clase modal

$i$  = Tamaño del intervalo de clase.

### Ejemplo

Calcular la moda del ingreso promedio mensual de las familias que habitan en un determinado sector

Intervalo de **clase modal**: el 4, (1080 – 1130) ya que

$n_4 = 8$  (la mayor frecuencia absoluta de la tabla)

$L_i = 1080$  dólares

$d_1 = n_4 - n_3 = 8 - 5 = 3$

$d_2 = n_4 - n_5 = 8 - 6 = 2$

$i = 50$

No.	Limite clase	$n_k$	$N_k$
3	1030    1080	5	13
4	1080    1130	8	21
5	1130    1180	6	27

# Datos agrupados

## Moda

Sustituyendo los datos:

$$M_o = L_i + \left[ \frac{d_1}{d_1 + d_2} \right] * i = 1080 + \frac{3}{3+2} * 50 = 1110,0$$

### **Respuesta/**

Este valor de \$ 1.110,0 se interpreta como el ingreso familiar que con más frecuencia se da.



# Relación entre la Media, Mediana y Moda

## Situaciones de Simetría

---

Si las medidas de tendencia central se presentan en el siguiente orden de magnitud:

**Moda < Mediana < Media    o**

**Media < Mediana < Moda**

Se dice que el polígono de frecuencias (histograma) es asimétrico, lo que indica que los datos se encuentran distribuidos con algún grado de tendencia

---

# Asimétrica Positiva

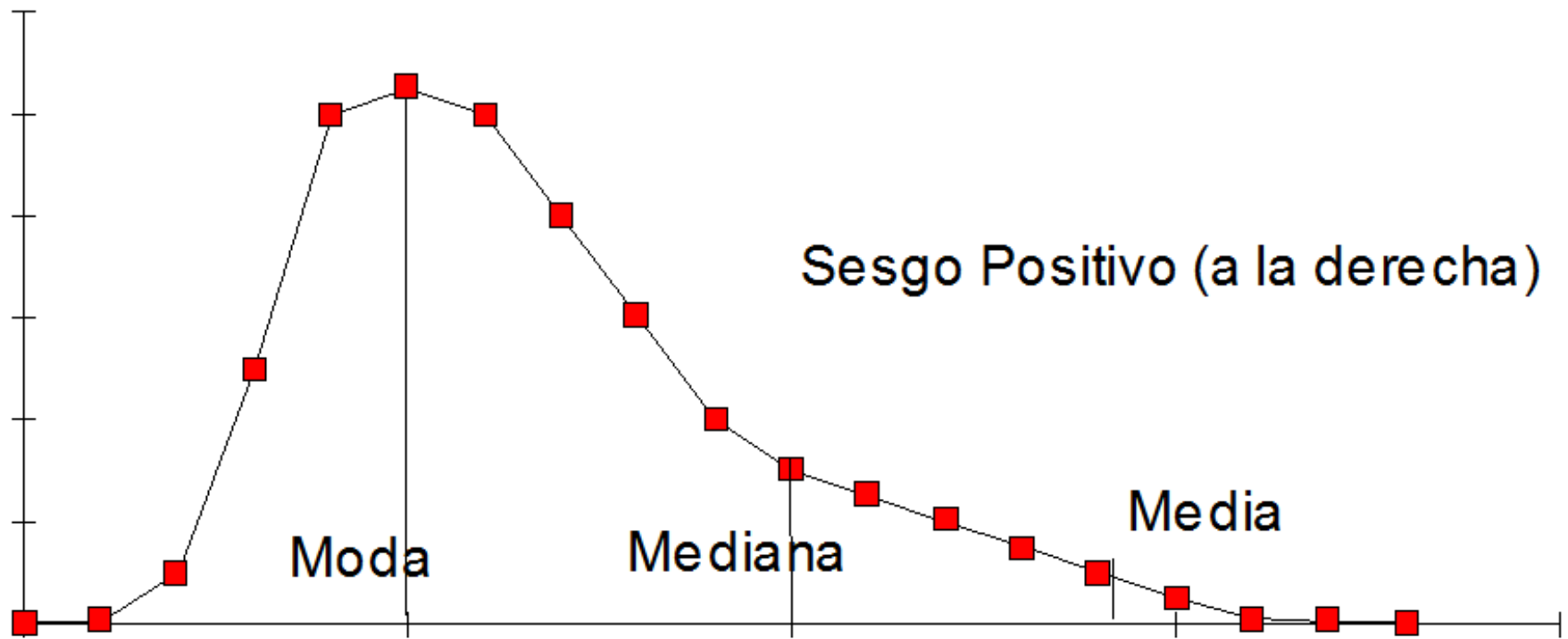
**Curva sesgada a la derecha o con sesgo positivo:**

**Moda < Mediana < Media**

**en este caso la mayoría de las observaciones se encuentran por debajo de la Media.**



# Moda < Mediana < Media

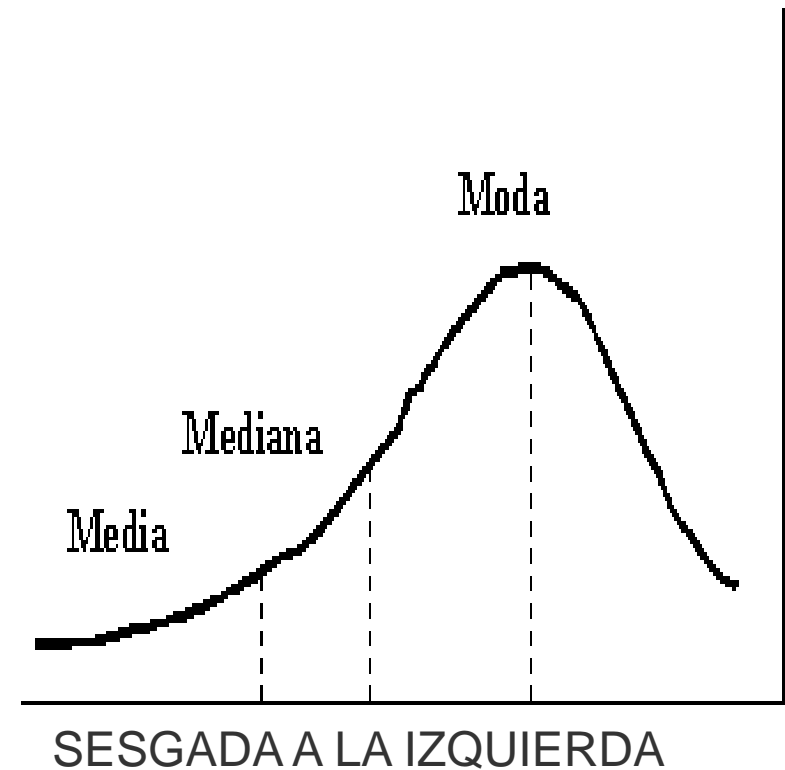


# Asimétrica Negativa

Curva sesgada a la izquierda o con sesgo negativo:

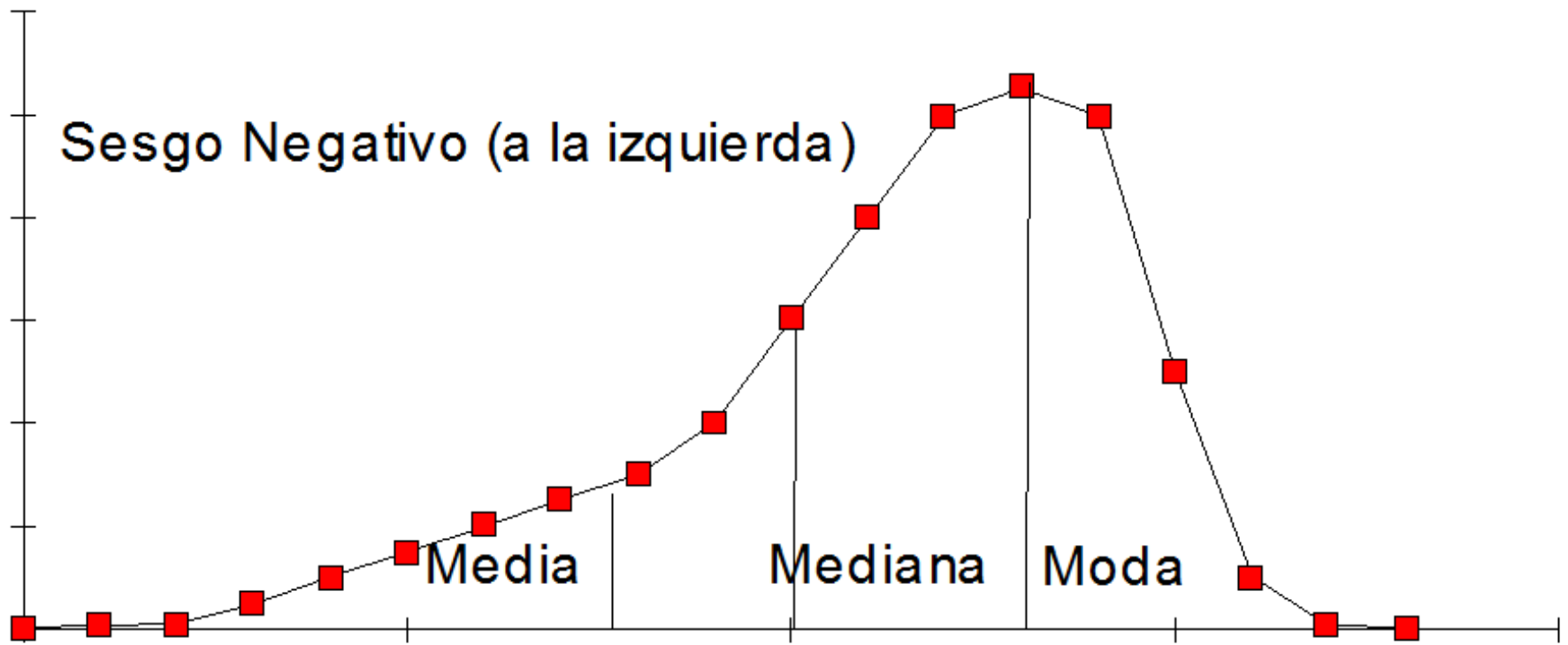
**Media < Mediana < Moda**

en este caso la mayoría de las observaciones se encuentran por arriba de la Media





# Media < Mediana < Moda



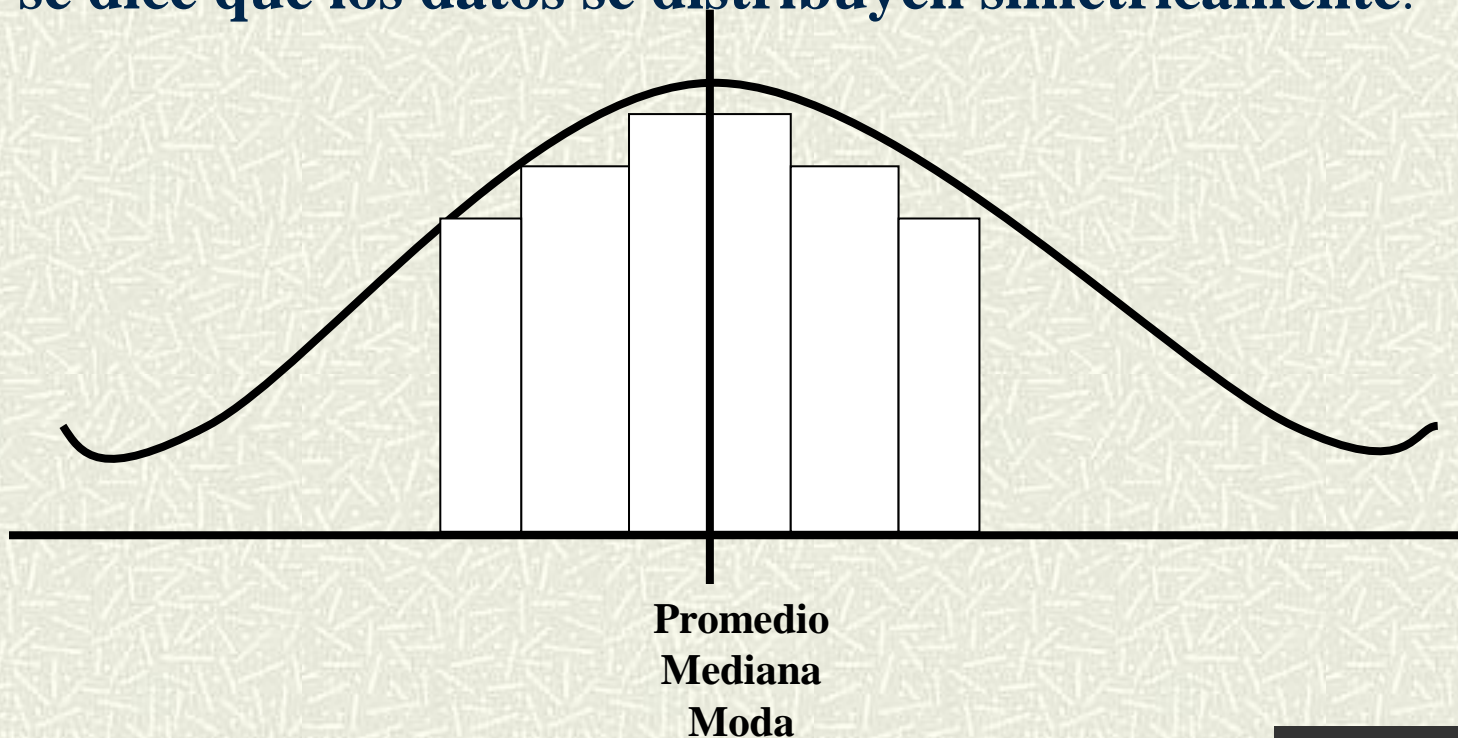
# Simétrica

## Curvas Simétricas o Normales

Si la Media, Moda y Mediana son iguales, esto es,

$$\text{Media} = \text{Mediana} = \text{Moda}$$

se dice que los datos se distribuyen simétricamente.



# Resumen

---

En el ejemplo de los ingresos por familia el resumen de los resultados en dólares es:

- **Media = 1.093,33**
- **Mediana = 1.100,0**
- **Moda = 1.030,0 y 1.150,0 (bimodal)**

Por tanto: **Moda < Media < Mediana** y el sesgo que presenta es a la izquierda

---

# Percentiles

---

- # Cuando un conjunto ordenado de datos se divide en 100 partes, los puntos de división reciben el nombre de percentiles
  - # Un valor tal que, cuando menos,  $p$  por ciento de los elementos son menores que o iguales a ese valor, y al menos  $(100-p)$  por ciento de los elementos tiene valor mayor que o igual a ese valor. El 50 percentil es la mediana
  - # Los cuartiles corresponden a los percentiles 25, 50 y 75. Se pueden usar para dividir al conjunto de datos en cuatro partes, donde cada una contiene aproximadamente el 25% de los datos
-



## Procedimiento de cálculo:

---

- # **Paso 1.** Ordenar los datos de manera ascendente
- # **Paso 2.** Calcular un índice  $i$ , en donde  $p$  es el percentil de interés y  $n$  es la cantidad de elementos

$$i = \left( \frac{p}{100} \right) n$$

- # **Paso 3.**
    - Si  $i$  no es entero, se redondea. El valor entero inmediatamente superior que  $i$  indica la posición del  $p$ -ésimo percentil
    - Si  $i$  es entero, el  $p$ -ésimo percentil es el promedio de los valores de los datos ubicados en los lugares  $i$  e  $i+1$
-

## Ejemplo

Determinar el 85° percentil de los dato adjuntos

Paso 1. 2210, 2255, 2350, 2380, 2390, 2420, 2440, 2450, 2550, 2630, 2825

Paso 2.

$$i = \left( \frac{85}{100} \right) \times 12 = 10.2$$

Paso 3. Como  $i$  no es entero, redondeamos. El lugar del 85° es el siguiente mayo que 10.2, o sea 11 es decir 2630

## Procedimiento de cálculo

- # **Paso 1.** Ordenar las clases y determinar las frecuencias porcentuales por clase y acumulativa
- # **Paso 2.** Utilizar la fórmula descrita

$$P_m = L_i + \frac{[(m/100) n - N_{k-1}]}{n_k} * i$$

# Donde:

$L_i$  : Límite inferior percentil de interés

$m$  : Percentil de interés

$N_{k-1}$  : Frec. Absoluta acumulada de la clase anterior a la clase percentil interés

$i$  : Amplitud de clase

$n_k$  : Frecuencia Absoluta de la clase del percentil de interés

$n$  : Tamaño de la muestra