MEDIDAS DE TENDENCIA CENTRAL

- Estas medidas se denominan de **«tendencia central»** porque fijan su atención en el centro de la distribución o punto central sobre el que gravitan el conjunto de valores de la distribución.

MEDIDAS DE TENDENCIA CENTRAL

- **LA MEDIA ARITMETICA o MEDIA o PROMEDIO**
- **LA MEDIANA**
- **LA MODA**
- **Rango medio**

MEDIA o PROMEDIO

Son medidas que buscan posiciones (valores) con respecto a que los datos muestran tendencia a agruparse

- Media ('mean') Es la media aritmética (promedio) de los valores de una variable. Suma de los valores dividido por el tamaño muestral
  - Media de {2, 2, 3, 7} es (2+2+3+7)/4 = 3,5
  - Conveniente cuando los datos se concentran simétricamente con respecto a ese valor. Muy sensible a valores extremos
  - Centro de gravedad de los datos

Datos sin agrupar

$$= \frac{\sum_{i=1}^{n} x_i}{n} ¿$$

Datos agrupados

$$= \frac{\sum_{i=1}^{n} f_i x_i}{n} ¿$$

Observaciones sobre la media

- La media (o promedio), en todos los casos, es un número comprendido entre el mínimo y el máximo de los valores observados.
- El promedio no tiene por qué coincidir con alguno de los valores observados en la población.
- Si la distribución de la variable no es muy dispersa (porque se concentra en unos pocos valores) entonces el promedio es un buen indicador de la "posición" de la distribución.
- Como medida de tendencia central, tiene el defecto de estar <u>muy influido</u> por los valores extremos de la distribución. Ya que todas las observaciones intervienen en el cálculo de la media, la aparición de una observación extrema, hará que la media se desplace en esa dirección.

- No es recomendable usar la media como medida central en las distribuciones muy asimétricas;

Observaciones 2

En general, la media aritmética obtenida a partir de las marcas de clase xc, diferirá de la media obtenida con los valores reales, *xi*.

Es decir, habrá una pérdida de precisión que será tanto mayor cuanto mayor sea la diferencia entre los valores reales y las marcas de clase, o sea, cuanto mayores sean las amplitudes de los intervalos de clase *ai*.

La media calculada sobre datos agrupados en intervalos dependerá siempre de la división en intervalos de clase.

La **media aritmética** de un conjunto de datos es el cociente entre la suma de todos los datos y el número de estos.

Ejemplo: las notas de Juan el año pasado fueron:

5, 6, 4, 7, 8, 4, 6

Hay 7 datos

que suman 40

La nota media de Juan es:

$$\text{Nota media} = \frac{5+6+4+7+8+4+6}{7} = \frac{40}{7} = 5{,}7$$

Cálculo de la media aritmética cuando los datos se repiten.

1º. Se multiplican los datos por sus frecuencias absolutas respectivas, y se suman.
2º. El resultado se divide por el total de datos.

Ejemplo. Las notas de un grupo de alumnos fueron:

Notas	Frecuencia absoluta	Notas x F. absoluta
3	**5**	**15**
5	**8**	**40**
6	**10**	**60**
7	**2**	**14**
Total	**25**	**129**

Datos por frecuencias

$$\text{Media} = \frac{129}{25} = 5{,}1$$

Total de datos

Ilustración 3. En tarde calurosa del sábado, Cristian un empleado de un kiosco de bebidas sirvió en total 50 bebidas durante la mañana de ese día. Vendió 5 bebidas de \$0.50, 15 de \$0.75, otras 15 de \$0.90, y otras 15 de \$1.10. A continuación se muestra la media ponderada del precio de las bebidas vendidas por Cristian para ese día:

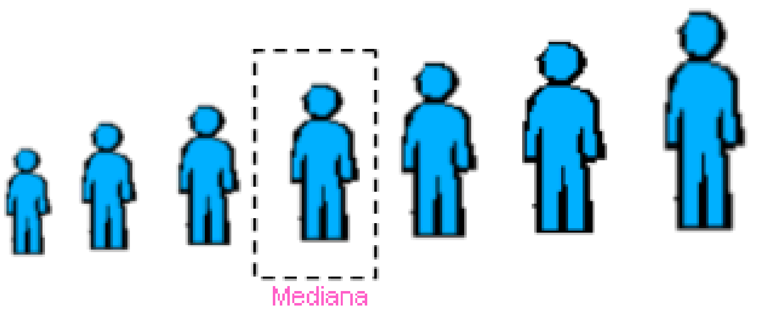
$$X_w = 5(\$0.50) + 15(\$0.75) + 15(\$0.90) + 15(\$1.15) / 50 = \$0.89$$

La media ponderada para el precio de las bebidas despachadas por Cristian en su kiosco, con base a los datos del día sábado, fue de \$0.89 por bebida.

Mediana

La mediana, a diferencia de la media no busca el valor central del recorrido de la variable según la cantidad de observaciones, sino que busca determinar el valor que tiene aquella observación que divide la cantidad de observaciones en dos mitades iguales. Por lo tanto es necesario atender a la ordenación de los datos, y debido a ello, este cálculo depende de la posición relativa de los valores obtenidos. Es necesario, antes que nada, ordenar los datos de menor a mayor (o viceversa).

$$med = x_{\left(\frac{N+1}{2}\right)} \qquad med = \frac{N}{2}$$



MEDIANA

- Mediana ('median') Es un valor que divide a las observaciones en dos grupos con el mismo número de individuos (percentil 50). Si el número de datos es par, se elige la media de los dos datos centrales

  - Mediana de 1, 2, 4, **5**, 6, 6, 8 es 5

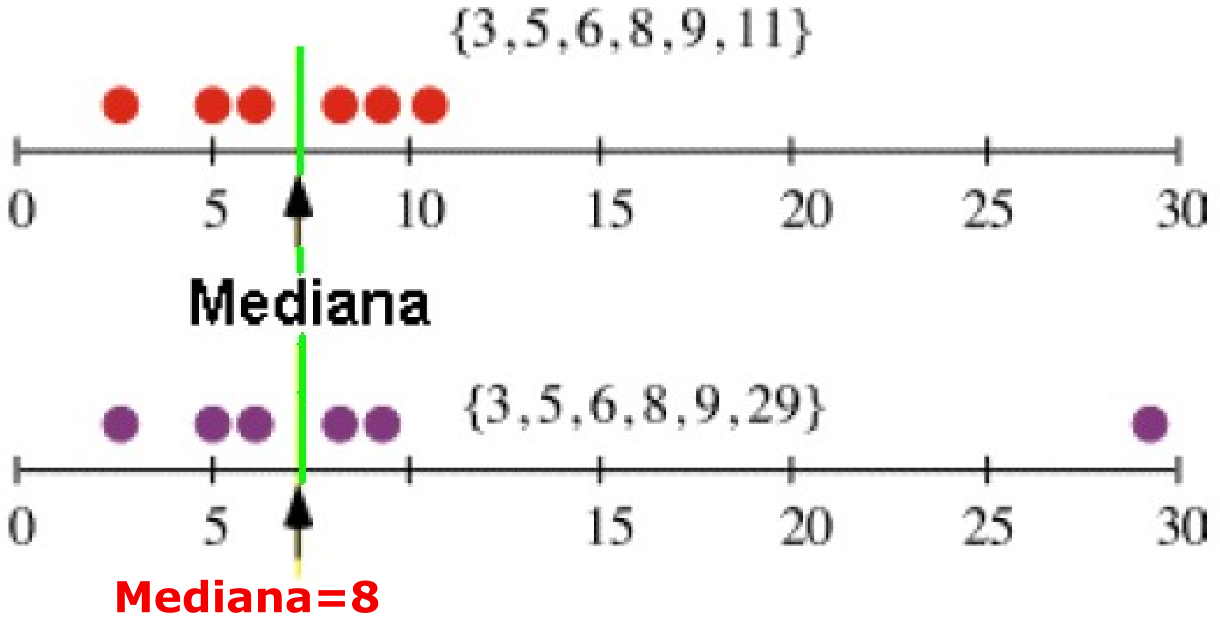
  - Mediana de 1, 2, 4, **5**, **6**, 6, 8, 9 es (5+6)/2 = 5,5

  - Es conveniente cuando los datos son asimétricos.

    No es sensible a valores extremos.

    - Mediana de 1, 2, 4, **5**, 6, 6, 80 es 5

MEDIANA



No es afectada por valores extremos

MEDIANA

Fórmula de la mediana para datos simples

$$Me = \begin{cases} X_{\frac{n+1}{2}} & si\ n\ es\ impar \\ \left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right) * \dfrac{1}{2} & si\ n\ es\ par \end{cases}$$

MEDIANA

Fórmula de la mediana para datos agrupados:

$$Me = L_i + \left[\frac{\frac{n}{2} - F_{i-1}}{f_i} \right] C_i$$

$[L_{i-1}, L_i)$ $i=1$ $i=5$

Intervalo	i
[60 – 70)	1
[70 – 80)	2
[80 – 90)	3
[90 – 100)	4
[100 – 110)	5
[110 – 120]	6

$i=1,\dots,6$

Se considera el menor intervalo i, tal que Fi >n/2

Donde:

n : número de datos de la muestra.

L_i : límite inferior del intervalo que contiene a la mediana.

F_{i-1}: Frecuencia absoluta acumulada inmediata inferior a la frecuencia acumulada F_i.

f_i : frecuencia absoluta del intervalo de la clase mediana (es aquel que contiene a F_i).

C_i : amplitud del intervalo de la clase mediana.

MEDIANA

Pasos para calcular la mediana (datos agrupados):

1ro. Se determinan las frecuencias acumulada absolutas (F_i).

2do. Se halla n/2. (suma de frecuencia dividida entre 2).

3ro. Ubicar n/2 en la columna de las F_i.

4to. Si n/2 no coincide con algún F_i, ubicar un F_i como la inmediata superior a n/2, para determinar el dato X_i o el intervalo Mediano correspondiente.

F_i = Es la Frecuencia absoluta acumulada inmediata superior a n/2.

F_{i-1} = es la Frecuencia absoluta acumulada inmediata anterior a F_i.

Ejemplo

	f_i
[60 – 70)	3
[70 – 80)	6
[80 – 90)	7
[90 – 100)	9
[100 – 110)	2
[110 – 120]	3
	30

MEDIANA

Punto medio del intervalo

	x_i	f_i	h_i	F_i	H_i
[60 – 70)	65	3	0.10	3	0.10
[70 – 80)	75	6	0.20	9	0.30
[80 – 90)	85	7	0.23	16	0.53
[90 – 100)	95	9	0.30	25	0.83
[100 – 110)	105	2	0.07	27	0.90
[110 – 120]	115	3	0.10	30	1.00
	Total =	30	1.00		

Intervalo mediano

MEDIANA

$$Me = L_i + \left[\frac{\frac{n}{2} - F_{i-1}}{f_i} \right] C_i$$

$$Me = 80 + \left[\frac{\frac{30}{2} - 9}{7} \right] 10 = 80 + \frac{15 - 9}{7} 10 = 80 + \left[\frac{6}{7} \right] 10 = 80 + 8{,}57 = 88{,}57$$

Ejemplo: Calcular la mediana para estos datos agrupados

	f_i
[60, 63)	5
[63, 66)	18
[66, 69)	42
[69, 72)	27
[72, 75)	8

Ejemplo:

Calcular la **mediana** de una distribución estadística que viene dada por la siguiente tabla:

	f_i	F_i
[60, 63)	5	5
[63, 66)	18	23
[66, 69)	42	65
[69, 72)	27	92
[72, 75)	8	100
	100	

100/2 = 50

Clase de la mediana: [66, 69)

$$Me = 66 + \frac{50 - 23}{42} \cdot 3 = 67.93$$

MODA

Es la medida de tendencia central, que se define como el valor que se presenta con mayor frecuencia en una serie o distribución de datos.

PARA DATOS SIN TABULAR:

Ejemplo 1: Encontrar la Moda en la serie: 6, 8, 6, 9, 10, 3, 6, 3

La moda es 6, porque se repite con mayor frecuencia, es una serie unimodal.

Ejemplo 2: Hallar la Moda en la serie: 4, 7, 8, 9, 10, 3, 8, 4

Presenta dos modas, que son 4 y 8, porque se repite con mayor frecuencia, es una serie bimodal.

b) DATOS TABULADOS

$$M_o = L_i + \left[\frac{d_1}{d_1 + d_2} \right] C_i$$

La clase modal es el intervalo con mayor frecuencia

L_i =Límite inferior de la clase Modal

d_1 = Diferencia de la frecuencia simple de la clase Modal y la frecuencia de la clase inmediata inferior.

d_2 =Diferencia de la frecuencia simple de la clase Modal y la frecuencia de la clase inmediata superior.

C =Tamaño o amplitud del intervalo.

MODA

Punto medio del intervalo

	x_i	f_i	h_i	F_i	H_i
[60 – 70)	65	3	0.10	3	0.10
[70 – 80)	75	6	0.20	9	0.30
[80 – 90)	85	7	0.23	16	0.53
[90 – 100)	95	9	0.30	25	0.83
[100 – 110)	105	2	0.07	27	0.90
[110 – 120]	115	3	0.10	30	1.00
	total	30	1.00		

Intervalo modal

MODA

$$M_o = L_i + \left[\frac{d_1}{d_1 + d_2} \right] C$$

$$L_i = 90$$

$$d_1 = 9 - 7 = 2$$

$$d_2 = 9 - 2 = 7$$

$$c = 10$$

Luego:

$$M_o = 90 + \left[\frac{2}{2+7} \right] 10 = 90 + \frac{20}{9} = 90 + 2{,}22 = 92{,}22$$

Ejemplo

Calcular la **moda** de una distribución estadística que viene dada por la siguiente tabla:

	f_i
[60, 63)	5
[63, 66)	18
[66, 69)	42
[69, 72)	27
[72, 75)	8
	100

- $$Mo = 66 + \frac{(42-18)}{(42-18)+(42-27)} \cdot 3 = 67.846$$

- $$Mo = 66 + \frac{27}{18+27} \cdot 3 = 67.8$$

Rango Medio

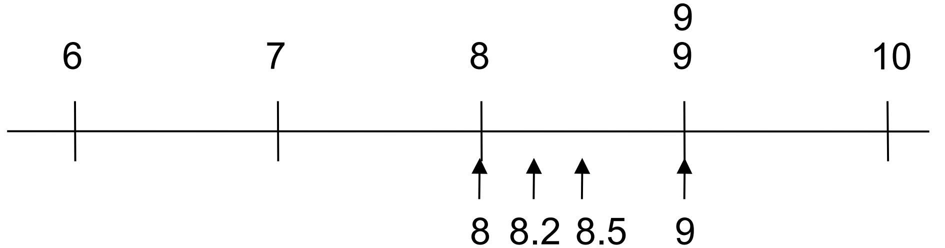
Número que está exactamente a la mitad del camino entre un dato con menor valor *Mín* y un dato con mayor valor *Máx*. Se encuentra promediando los valores mínimo y máximo.

$$\text{Rango Medio} = \frac{\text{valor mínimo} + \text{valor máximo}}{2}$$

$$\text{Rango Medio} = \frac{\textit{Mín} + \textit{Máx}}{2}$$

Nota…

Las cuatro medidas de tendencia central representan cuatro métodos distintos para describir el centro. Estos cuatro valores pueden ser iguales, aunque es más probable que sean diferentes. Para los datos muestrales 6, 7, 8, 9, 9, 10, la media es 8.2, la mediana es 8.5, la moda es 9 y el rango medio es 8.



Ejercicios

1. Considere la muestra 2, 4, 7, 8, 9. Encuentre:
    - La media
    - La mediana
    - La moda
    - El rango medio

2. A 15 estudiantes universitarios, elegidos aleatoriamente, se les solicitó mencionar el número de horas que durmieron la noche anterior. Los datos resultantes fueron, 5, 6, 6, 8, 7, 7, 9, 5, 4, 8, 11, 6, 7, 8, 7. Encontrar:
    - La media
    - La mediana
    - La moda
    - El rango medio

Ejercicios

A los reclutas de una academia de policía se les solicitó presentar un examen que mide la capacidad que tienen para hacer ejercicio. Esta capacidad (medida en minutos) se obtuvo para cada uno de los 20 reclutas:

25	27	30	33	30	32	30	34	30	27
26	25	29	31	31	32	34	32	33	30

a. Encuentre la media, la mediana, la moda y el rango medio.
b. Elabore una gráfica de barras para estos datos y localice la media, la mediana, la moda y el rango medio sobre la gráfica.
c. Describa la relación que hay entre los cuatro promedios (semejanza) y qué propiedades muestran los datos por las que dichos promedios son semejantes

Calcule la mediana y moda de los siguientes datos.

Ejemplo: La siguiente tabla muestra la inversión anual de 44 empresas.

Intervalo	Marca de clase m_i	Frecuencias f_i	h_i	Frecuencias acumuladas F_i	H_i
[4, 10[	7	1	0,025	1	0.025
[10, 16[	13	3	0,075	4	0.100
[16, 22[	19	6	0,150	10	0.250
[22, 28[	25	12	0,300	22	0.550
[28, 34[	31	11	0,275	33	0.825
[34, 40[	37	5	0,125	38	0.950
[40, 46]	43	2	0,050	40	1.000
		40	1,000		