

CBS

Colegio Bautista Shalom



Estadística I

Cuarto PCOC

Segundo Bimestre

Contenidos

MEDIDAS DE TENDENCIA CENTRAL

- ✓ LA MEDIA ARITMÉTICA.
- ✓ MEDIA ARITMÉTICA PONDERADA.
 - PROPIEDADES DE LA MEDIA ARITMÉTICA.
- ✓ MEDIA ARITMÉTICA MUESTRAL.
- ✓ MEDIA ARITMÉTICA GEOMÉTRICA.
- ✓ LA MEDIANA.
- ✓ LA MODA.

NOTA: conforme avances en tu aprendizaje, encontrarás ejercicios a realizar. Sigue las instrucciones de tu catedrático(a).

MEDIDAS DE TENDENCIA CENTRAL

El primer conjunto de indicadores de la estadística corresponde a las *Medidas de Tendencia Central*.

Existen varios procedimientos para expresar matemáticamente las medidas de tendencia central, de los cuales, los más conocidos son: la media aritmética, la moda y la mediana.

Es importante recordar los siguientes términos:

Individuo: cualquier elemento que porte información sobre el fenómeno que se estudia. Así, si estudiamos la altura de nuestros compañeros dentro de la clase, cada alumno es un individuo; si estudiamos el precio de la vivienda, cada vivienda es un individuo.

Población: conjunto de todos los individuos (personas, objetos, animales, entre otros) que porten información sobre el fenómeno que se estudia. Ejemplificando: si estudiamos el precio de la vivienda en una ciudad, la población será el total de las viviendas de dicha ciudad.

Muestra: subconjunto que seleccionamos de la población. Es decir, al momento de estudiar el precio de la vivienda de una ciudad, lo normal será no recoger información sobre todas las viviendas de la ciudad (sería una labor muy compleja), sino que se suele seleccionar un subgrupo (muestra) que se entienda que es suficientemente representativo.

Son indicadores estadísticos que muestran hacia qué valor (o valores) se agrupan los datos. Aquellos valores que normalmente se encuentran ubicados en la parte central de un conjunto o agrupación de datos.

Estas, como medidas estadísticas pretenden "resumir" la información de la "muestra" para poder tener así un mejor conocimiento de determinada población. (Ellas permiten analizar los datos en torno a un valor central). Entre éstas están la media aritmética, la moda y la mediana.

Las medidas de tendencia central (media, mediana y moda) sirven como puntos de referencia para interpretar las calificaciones que se obtienen en una prueba. Al describir las características típicas de conjuntos de datos y, como hay varias formas de hacerlo, existen y se utilizan varios tipos de promedios. Se les llama medidas de tendencia central porque general mente la acumulación más alta de datos se encuentra en los valores intermedios.

¡Para analizar!

Imagina que tu compañero de adelante obtiene en el examen de bimestre 35 puntos. Este puntaje, por sí mismo tiene muy poco significado a menos que podamos conocer el total de puntos que obtiene alguno de tus compañeros al someterse a dicho examen, saber cuál es la calificación menor y mayor que se obtiene, y cuán variadas son esas calificaciones.

Es decir, para que una calificación tenga significado hay que contar con elementos de referencia generalmente relacionados con ciertos criterios estadísticos.

En este sentido las medidas de tendencia central sirven como puntos de referencia para interpretar las calificaciones que se obtienen en una prueba.

Volviendo a nuestro ejemplo, digamos que la calificación promedio en la prueba que hizo el alumno fue de 65 puntos. Con este dato podemos decir que la calificación del alumno se ubica notablemente sobre el promedio. Pero si la calificación promedio fue de 20 puntos, entonces la conclusión sería muy diferente, debido a que se ubicaría muy por debajo del promedio de la clase.

En resumen, el propósito de las medidas de tendencia central es:

- ✓ Mostrar en qué lugar se ubica la persona promedio o típica del grupo.
- ✓ Sirve como un método para comparar o interpretar cualquier puntaje en relación con el puntaje central o típico.
- ✓ Sirve como un método para comparar el puntaje obtenido por una misma persona en dos diferentes ocasiones.
- ✓ Sirve como un método para comparar los resultados medios obtenidos por dos o más grupos.

Las medidas de tendencia central más comunes son: Media Aritmética, La Mediana y La Moda.

¡Para recordar!

La distribución de frecuencia es la representación estructurada, en forma de tabla, de toda la información que se ha recogido sobre la variable que se estudia.

Variable (Valor)	Frecuencias Absolutas		Frecuencias Relativas	
	Simple	Acumulada	Simple	Acumulada
X	X	X	X	X
X ₁	n ₁	n ₁	f ₁ = n ₁ / n	f ₁
X ₂	n ₂	n ₁ + n ₂	f ₂ = n ₂ / n	f ₁ + f ₂
...	...	...	...	...
X _{n-1}	n _{n-1}	n ₁ + n ₂ + .. + n _{n-1}	f _{n-1} = n _{n-1} / n	f ₁ + f ₂ + .. + f _{n-1}
X _n	n _n	S n	f _n = n _n / n	S f

En donde:

Siendo **X** los distintos valores que puede tomar la variable.

Siendo **n** el número de veces que se repite cada valor.

Siendo **f** el porcentaje que la repetición de cada valor supone sobre el total.

La cantidad de tramos en los cuales, se agrupa la información es a criterio del analista.

La regla dice:

"Mientras más tramos se utilicen menos información se pierde, pero puede que menos representativa e informativa sea la tabla"

Las medidas de tendencia central nos ayudan en el análisis estadístico, facilitándonos la información sobre la serie de datos que estamos analizando. Estas nos permiten conocer determinadas y diversas características de esta serie de datos.

Las medidas de posición son de dos tipos:

- 1. Medidas de Posición Central:** son aquellas que informan los valores medios del conjunto de datos que se nos presenta para analizar.
- 2. Medidas de Posición No Centrales:** estas nos informan de cómo se distribuye el resto de los valores de la serie.

LA MEDIA ARITMÉTICA

Es el conjunto finito de números que es igual a la suma de todos los valores dividido entre el número de sumandos que intervienen. Comúnmente se le conoce con el nombre de *promedio*.

Si el conjunto en cuestión es una muestra aleatoria, tal como se designa a los individuos de una población estadística, se llamará media muestral y se convertirá en uno de los principales estadísticos muestrales.

Como medida de posición central la media aritmética nos sirve para realizar el cálculo cada valor por el número de veces que se repite. La suma de todos estos productos se divide por el total de datos de la muestra:

$$x_m = \frac{(x_1 \cdot n_1) + (x_2 \cdot n_2) + (x_3 \cdot n_3) + \dots + (x_{n-1} \cdot n_{n-1}) + (x_n \cdot n_n)}{n}$$

Por ejemplo:

Si quiero conocer la Media Aritmética o promedio que tengo en una determinada materia del colegio, debo sumar los números de cada una de las notas que obtuve en los exámenes y dividirlos por la cantidad de pruebas, es decir, si mis notas durante el año fueron de 4, 5, 7, 8 y 10, la media aritmética o promedio en cuestión será de 6,80.

Siempre que queramos obtener un promedio debemos disponer de dos cantidades de las cuales precisamente podemos lograr su punto medio. Siempre necesitaremos de otras cifras porque no se puede promediar una cifra consigo misma.

En el caso que sean varias las cifras, deberemos, como ya dijimos, sumarlas a todas y más luego dividir las por la cantidad de números que intervienen, es decir, si fueron cinco cifras dividir las por ese número.

Datos No Agrupados:

La media aritmética ($\bar{X}$), de una cantidad finita de números ($X_1, X_2, X_3, \dots, X_n$), es igual a la suma de todos ellos dividida entre el número de sumandos (n).

Simbólicamente se expresa así:

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

X= cualquier dato

Número total de datos

$$\bar{X} = \frac{(X_1 + X_2 + X_3 + \dots + X_n)}{n}$$

Por ejemplo: calcular la media aritmética de los números 10, 12, 36, 25 y 58.

$$\bar{x} = \frac{10 + 12 + 36 + 25 + 58}{5} = \frac{121}{5} = 24.2$$

Datos Agrupados:

La fórmula correspondiente para su cálculo es la siguiente:

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

Frecuencia por la marca de clase de cualquier renglón

Número total de datos

En donde: k = última clase.

Nota: la media muestral se denota $\bar{X}$, la media poblacional se conoce como μ .

$$\bar{X} = \frac{f_1 X_1 + f_2 X_2 + f_3 X_3 + \dots + f_n X_n}{f_1 + f_2 + f_3 + \dots + f_n}$$

¿Cuándo se debería utilizar este tipo de media? Para responder a esta interrogante se presentan una ilustración sobre la aplicación de esta medida de tendencia central.

Por ejemplo: Se desea estimar el rendimiento promedio de las llantas de cierta marca. Para ello se toma una muestra de cuatro automóviles a los que se les coloca esta marca de llanta.

Una vez las llantas se desgastan completamente se anota el número de kilómetros recorridos por cada auto, encontrándose los siguientes valores:

Número de Auto	Recorrido (kms)
1	56,000
2	42,000
3	23,000
4	73,000

Con base a la tabla anterior, se procede a calcular el promedio de la siguiente manera:

$$\bar{X} = \frac{(56,000 + 42,000 + 23,000 + 73,000)}{4} = 48,500 \text{ Kilómetros}$$

Por tanto, se puede concluir que el rendimiento promedio de las llantas de esta marca (vida útil) es de 48,500 kilómetros.

La media aritmética, o simplemente media, de un conjunto de N números $X_1, X_2, X_3, \dots, X_N$ se denota por $\bar{X}$; y se encuentra definida por:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N} = \frac{\sum_{j=1}^N X_j}{N} = \frac{\sum X}{N}$$

Por ejemplo: la media aritmética de los números 8, 3, 5, 12 y 10 es.

$$\bar{X} = \frac{8 + 3 + 5 + 12 + 10}{5} = \frac{38}{5} = 7.6$$

Si los números $X_1, X_2, \dots, X_K$ ocurren $f_1, f_2, \dots, f_K$ veces, respectivamente (o sea, con frecuencias $f_1, f_2, \dots, f_K$), la media aritmética es:

$$\bar{X} = \frac{f_1 X_1 + f_2 X_2 + \dots + f_K X_K}{f_1 + f_2 + \dots + f_K} = \frac{\sum_{j=1}^K f_j X_j}{\sum_{j=1}^K f_j} = \frac{\sum fX}{\sum f} = \frac{\sum fX}{N} \quad (2)$$

En donde: $N = \sum f$ es la frecuencia total (es decir, el número total de casos).

Por ejemplo: si: 5, 8, 6 y 2 ocurren con frecuencias 3, 2, 4, y 1, respectivamente, su media aritmética es

$$\bar{X} = \frac{(3)(5) + (2)(8) + (4)(6) + (1)(2)}{3 + 2 + 4 + 1} = \frac{15 + 16 + 24 + 2}{10} = 5.7$$

MEDIA ARITMÉTICA PONDERADA

Es una media o promedio de cantidades a las que se ha asignado una serie de coeficientes, llamados pesos, para tener en cuenta adecuadamente su importancia relativa.

$$\bar{X}_w$$

Datos No Agrupados:

La media ponderada de un grupo de datos $X_1, X_2, \dots, X_k$, con sus correspondientes pesos $w_1, w_2, \dots, w_k$, puede obtenerse a través de la siguiente fórmula:

$$\bar{X} = \frac{w_1 X_1 + w_2 X_2 + \dots + w_k X_k}{w_1 + w_2 + \dots + w_k} = \frac{\sum wX}{\sum w}$$

Por ejemplo: en el curso de estadística, la nota semestral se calcula como una media ponderada. Por cuanto que el promedio de laboratorios representa el 30% de la nota semestral. El promedio de ejercicios parciales representa el 30% y el examen semestral el restante 40%.

Si obtiene en este curso los siguientes promedios al final del semestre: laboratorios 90 pts. Parciales 75% pts. Y en el examen semestral 70 pts.; el promedio semestral se calcula de la siguiente forma:

$$\bar{X}_w = \frac{90(0.30) + 75(0.30) + 70(0.40)}{0.30 + 0.30 + 0.40} = \frac{27.0 + 22.5 + 28.0}{1} = 77.5$$

La nota semestral de 77.5.

¿Cuándo se debería utilizar este tipo de media? Se incluye una ilustración para responder a esta pregunta.

Cuando se trabaja con la media aritmética simple, se asume que a cada observación se le da la misma importancia.

Sin embargo, en ciertos casos, puede querer darse mayor peso o importancia a algunas de las observaciones y entonces se aplica la media ponderada.

Por ejemplo: en la clase de Estadística, para determinar la nota que un alumno obtendrá en el curso se asignan pesos de importancia, de la siguiente forma: Unidad I (20% del curso), Unidad II (35% del curso), Unidad III (20% del curso), Unidad IV (15% de la calificación), Unidad V (20% de la calificación).

Si las calificaciones de un alumno son 80 en la primera unidad, 50 en la segunda, 80 en la tercera unidad, 100 en la cuarta unidad y 80 en la última unidad, obtiene la siguiente tabla:

Unidad	Ponderación (w_i)	Datos (X_i)
I	20% = 0.2	80
II	25% = 0.35	50
III	20% = 0.2	80
IV	15% = 0.15	100
V	20% = 0.10	80

La media ponderada de las notas del alumno se determina de la siguiente manera:

$$\bar{X}_w = 80(0.2) + 50(0.35) + 80(0.20) + 100(0.15) + 80(0.10) / 1 = 72.50$$

El promedio ponderado entonces para este alumno es de una calificación de 72.50 puntos sobre 100, que era el total máximo.

PROPIEDADES DE LA MEDIA ARITMÉTICA

Propiedad Primera. La suma algebraica de las desviaciones de un conjunto de número respecto de su media aritmética es cero.

Ejemplo:

Las desviaciones de los números: 8, 3, 5, 12 y 10

Respecto de su media aritmética 7.6 son: $8 - 7.6$, $3 - 7.6$, $5 - 7.6$, $12 - 7.6$ y $10 - 7.6$

Es decir: 0.4, - 4.6, - 2.6, 4.4 y 2.4

Con suma algebraica: $0.4 - 4.6 - 2.6 + 4.4 + 2.4 = 0$

Propiedad Segunda. La suma de los cuadrados de las desviaciones de un conjunto de números X_j respecto de un cierto número a es mínima, si sólo si:

$$a = \bar{X}$$

Propiedad Tercera. Si f_1 números tiene media m_1 , f_2 números tiene media m_2 , ... f_k números tienen media m_k , entonces la media de todos los números es:

$$\bar{X} = \frac{f_1 m_1 + f_2 m_2 + \dots + f_k m_k}{f_1 + f_2 + \dots + f_k}$$

Es decir, una media aritmética ponderada de todas las medias.

Si A es una media aritmética supuesta o conjeturada (que puede ser cualquier número) y si $d_j = X_i - A$ son las desviaciones de X_i respecto de A , las ecuaciones (1) y (2) se convierten, respectivamente, en:

$$\bar{X} = A + \frac{\sum_{j=1}^N d_j}{N} = A + \frac{\sum d}{N}$$

$$\bar{X} = A + \frac{\sum_{j=1}^K f_j d_j}{\sum_{j=1}^K f_j} = A + \frac{\sum f d}{N}$$

donde $N = \sum_{j=1}^K f_j = \sum f$. Observe que las fórmulas (5) y (6) se resumen en la ecuación:

$$\bar{X} = A + \bar{d}$$

EXERCICIOS A continuación se te presentan problemas de hallar el valor de la media aritmética simple o ponderada.

PROBLEMA 01. Un grupo de 6 amigas tienen distintas edades. Son las siguientes: 2 de ellas tienen 28 años y otras 2 tienen 32 años, el resto tienen 29 y 30 años respectivamente. Calcula la media aritmética del grupo.

PROBLEMA 02. Calcula la media aritmética de la altura de un equipo de cinco jugadores de baloncesto que miden: 1.92, 1.95, 1.83, 1.76 y 1.69.

PROBLEMA 03. En clase de inglés 10 alumnos han sacado las siguientes notas: 7, 6.5, 4, 1, 9, 5, 8, 8.5, 2, 5.5. Siendo 10 la mayor nota y 0 la más baja. Calcula la media aritmética de las notas de la clase.

PROBLEMA 04. Calcula la media aritmética ponderada hallando la nota promedio de dos asignaturas que tienen valores diferentes. Química tiene un valor de 3 créditos y Física vale 2 créditos. Si Pedro ha sacado un 8 en la primera y un 7 en la segunda.

PROBLEMA 05. Para calcular la nota final del curso de literatura en donde cada apartado ha tenido distinta importancia. Los dos primeros trabajos tienen valor de 20% y 20% respectivamente, y el examen de 60%; las calificaciones respectivas son de 6.4, 9.2 y 8.1

PROBLEMA 06. En un partido de baloncesto, se tiene la siguiente anotación en los jugadores de un equipo. Calcular la media de anotación del equipo.

PROBLEMA 07. La altura en cm de los jugadores de un equipo de baloncesto está en la siguiente tabla. Calcular la media.

Intervalo	x_i	f_i	$x_i \cdot f_i$
[160, 170)	165	1	165
[170, 180)	175	2	350
[180, 190)	185	4	740
[190, 200)	195	3	585
[200, 210)	205	2	410
		12	2250

PROBLEMA 08. En un partido de baloncesto, se tiene la siguiente anotación en los jugadores de un equipo. 0, 1, 3, 4, 5, 6, 7, 8, 47 . Calcular la media de anotación del equipo.

PROBLEMA 09. José cosechó del árbol 4 peras, Catalina – 2 peras, y María – 6. Los niños juntaron sus frutas y se las repartieron en forma igualitaria. ¿Cuántas peras obtuvo cada uno? Calcular la media aritmética.

PROBLEMA 10. A los cursillos del inglés asistieron 15 personas el lunes, el martes – 10, el miércoles – 12, el jueves – 11, el viernes – 7, el sábado – 14, el domingo – 8. Calcular asistencia media de los cursillos por la semana. Calcular la media aritmética.

PROBLEMA 11. El piloto estuvo yendo dos horas a la velocidad de 120 km a la hora y horas a la velocidad de 90 km a la hora. Calcule la velocidad media del coche durante la carrera. Calcular la media aritmética de las velocidades del coche por cada hora del camino.

PROBLEMA 12. La media aritmética de 3 números es 6, y la media aritmética de otros 7 números es 3. ¿Cuánta es la media aritmética de estos 10 números?

PROBLEMA 13. Tomamos cuatro elementos evaluables y calculamos la nota final de una asignatura usando la media ponderada de las notas que han obtenido el alumno. Damos diferentes pesos según la importancia de la siguiente forma: un peso de 3 al examen inicial, de 1 al trabajo entregable, 2 al trabajo final y 4 al examen final. En la tabla siguiente se muestra las notas de un alumno y sus pesos.

Notas del alumno		
Elementos evaluados	Nota	Peso
Examen inicial	5,2	3
Trabajo entregable	8,2	1
Trabajo final	7,4	2
Examen final	5,7	4

PROBLEMA 14. Las calificaciones en una clase de lenguaje tienen cuatro componentes, con los pesos indicados: lectura (25 %), lectura comprensiva (45 %), ortografía (20 %) y elaboración de trabajo (10%). Si un alumno fue calificado respectivamente con un: 6, 5, 7 y 9. ¿Cuál será su nota media? Halla su nota media si todos los componentes tuvieran el mismo peso. ¿Cómo conseguiría sacar mayor nota?

PROBLEMA 15. En una empresa hay 5 trabajadores que ganan Q.200; 4 que ganan Q.250; 8 que ganan Q.175 y 3 que ganan Q.300. ¿Cuál es el promedio de salarios de la empresa?

SALARIO Y_i	NUMERO DE TRABAJADORES n_i	TOTAL $n_i Y_i$
200	5	1000
250	4	1000
175	8	1400
300	3	900
		$\Sigma n_i Y_i = 4300$

MEDIA ARITMÉTICA MUESTRAL

Representa el centro físico del conjunto de datos y se define como la suma de los valores observados, dividido por el total de observaciones.

Si:

$$x_1, x_2, \dots, x_n$$

son "n" observaciones numéricas, entonces la media aritmética de estas "n" observaciones, se define como:

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

Cuando se trata de datos agrupados (tabla de frecuencias) la media aritmética se puede aproximar mediante la expresión:

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

En donde m_i y f_i son respectivamente el punto medio y la frecuencia del i -ésimo intervalo.

Por ejemplo. Se toman 10 mediciones del diámetro interno de los tornillos para los pistones de un motor de un vehículo.

Los datos (en mm) son: 74.001, 74.003, 74.015, 74.000, 74.002, 74.005, 74.001, 74.001, 74.002, 74.004.

La media muestral del diámetro interno de los tornillos es:

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{74.001 + 74.003 + \dots + 74.004}{10} \quad \bar{x} = 74.0034.$$

Ahora, la media aritmética para la siguiente tabla de frecuencias se encuentra dada por:

Clases	Marca de clase m_i	Frecuencias f_i
33-34	33.5	6
34-35	34.5	13
35-36	35.5	22
36-37	36.5	12
37-38	37.5	6
38-39	38.5	5

$$\bar{x} = \frac{\sum_{i=1}^{10} m f_i}{64} = \frac{33.5 \times 6 + 34.5 \times 13 + 35.5 \times 22 + 36.5 \times 12 + 37.5 \times 6 + 38.5 \times 5}{64} \quad \bar{x} = 35.71875.$$

determina las soluciones de cada problema de media aritmética y muestral que tu catedrático(a) te indique.

MEDIA ARITMÉTICA GEOMÉTRICA

La media geométrica (MG) de un conjunto de n números positivos se define como la n ésima raíz del producto de n números.

Ventajas y desventajas:

- ✓ En su cálculo intervienen todos los valores de la distribución.
- ✓ Los valores extremos tienen menor influencia que en la media aritmética.
- ✓ Es única.
- ✓ Su cálculo es más complicado que el de la media aritmética.

La media geométrica de números positivos no pasa de la media aritmética de estos mismos números.

Si los números: $x_1, x_2, \dots, x_n$

no son todos iguales, la media geométrica de estos números es menor que su media aritmética.

Demostración. De la igualdad $g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$ se deduce que $1 = \sqrt[n]{\frac{x_1}{g} \cdot \frac{x_2}{g} \cdot \dots \cdot \frac{x_n}{g}}$,

o sea,

$$\frac{x_1}{g} \cdot \frac{x_2}{g} \cdot \dots \cdot \frac{x_n}{g} = 1$$

Debido a que el producto de estos " n números" positivos es igual a 1, resulta (por el teorema 1) que la suma de los mismos no es menor que " n ", es decir:

$$\frac{x_1}{g} + \frac{x_2}{g} + \dots + \frac{x_n}{g} \geq n$$

Multiplicando por " g " y dividiendo entre " n " ambos miembros de la última desigualdad, obtenemos:

$$a = \frac{x_1 + x_2 + \dots + x_n}{n} \geq g$$

Notemos que la igualdad tiene lugar sólo cuando:

$$\frac{x_1}{g} = \frac{x_2}{g} = \dots = \frac{x_n}{g} = 1,$$

Es decir:

$$x_1 = x_2 = \dots = x_n = g$$

Por el contrario, si los números $x_1, x_2, \dots, x_n$ no son todos iguales, se tiene

$$a > g$$

La media geométrica (MG), de un conjunto de “ n ” números positivos se define como la *raíz n -ésima* del producto de los “ n ” números.

Por tanto, la fórmula para la media geométrica es dada por:

$$MG = \sqrt[n]{(X_1)(X_2) \dots (X_n)}$$

Existen dos usos principales de la media geométrica:

1. Para promediar porcentajes, índices y cifras relativas y
2. Para determinar el incremento porcentual promedio en ventas, producción u otras actividades o series económicas de un periodo a otro.

Por ejemplo. Supóngase que las utilidades obtenidas por una compañía constructora en cuatro proyectos fueron de 3, 2, 4 y 6%, respectivamente. ¿Cuál es la media geométrica de las ganancias?

En este ejemplo X y así la media geométrica es determinada por:

$$MG = \sqrt[4]{(3)(2)(4)(6)} = 3.464101615$$

y así la media geométrica de las utilidades es el 3.46%.

Cálculo de la Media Aritmética para Datos Agrupados

La media aritmética de los valores anteriores es 3.75%. Aunque el valor 6% no es muy grande, hace que la media aritmética se incline hacia valores elevados. La media geométrica no se ve tan afectada por valores extremos.

Cuando los datos se presentan en una distribución de frecuencias, todos los valores que caen dentro de un intervalo de clase dado se consideran iguales a la marca de clase, o punto medio del intervalo. Las fórmulas (2) y (6) son válidas para tales datos agrupados y se interpretan X_j como la marca de clase, f_j como su correspondiente frecuencia de clase.

A como cualquier marca de clase conjeturada o supuesta y $d_j = X_j - A$ como las desviaciones de X respecto de A . Los cálculos con las fórmulas (2) y (6) se llaman métodos largos y métodos cortos, respectivamente.

Si todos los intervalos de clase son del mismo tamaño c , las desviaciones $d_j = X_j - A$ pueden expresarse como $c u_j$ en donde u_j serían números enteros positivos, negativos o cero, es decir, 0, ± 1 , ± 2 , $\pm 3, \dots$, y la fórmula (6) se convierte en:

$$\bar{X} = A + \left(\frac{\sum_{j=1}^K f_j u_j}{N} \right) = A + \left(\frac{\sum f u}{N} \right) c$$

que es equivalente a la ecuación:

$$\bar{X} = A + c \bar{u}$$

Esto se conoce como método de codificación para calcular la media. Es un método corto y debe usarse siempre para datos agrupados con intervalos de clase de tamaños iguales.

MEDIA GEOMÉTRICA UTILIZANDO MICROSOFT EXCEL:

Si el crecimiento de las ventas en un negocio fue en los tres último años de 26%, 32% y 28%, hallar la media anual del crecimiento.

Solucionemos...

Si tomamos en cuenta que el factor de crecimiento (FC) se calcula de la siguiente manera:

$$FC = 1 + \left(\frac{ti}{100} \right)$$

Entonces, por ejemplo para el primer año el factor de crecimiento es de 1.26.

$$FC = 1 + \left(\frac{26}{100} \right) = 1.26$$

Así que:

$$MG = \sqrt[3]{(1.26)(1.32)(1.28)}$$

$$MG = \sqrt[3]{2.128896} = 1.286$$

■ calcular la raíz n-ésima de un número, es equivalente a elevar el número a la potencia $1/n$. Ejemplo $2.128896^{1/3} = 1.286$.

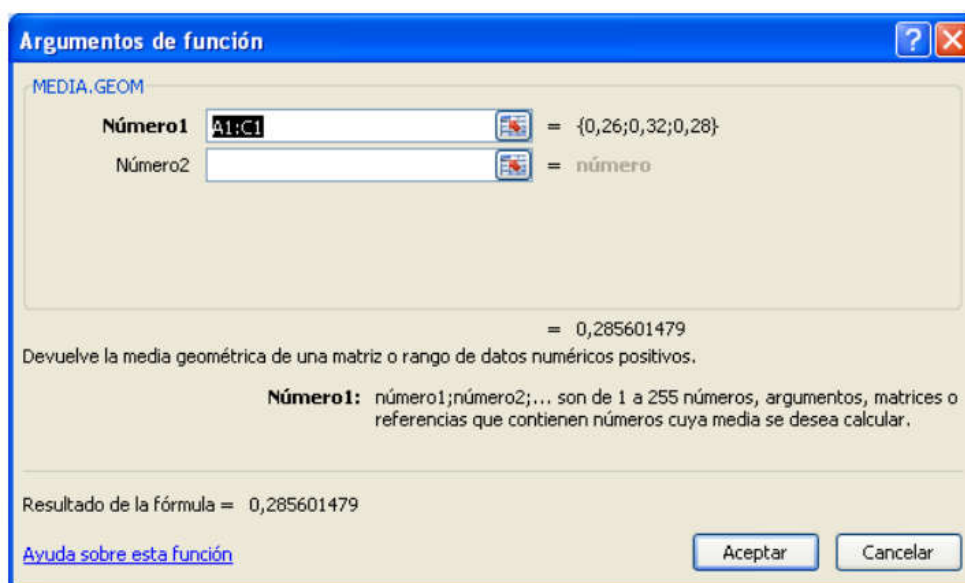
Se puede concluir entonces que la media anual de crecimiento fue del 28.6%.

Para calcular la media geométrica utilizando Excel, sigue estos pasos:

■ Copia la serie de datos a una hoja de Excel

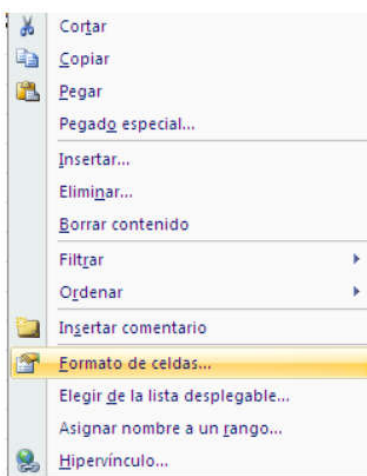
	A	B	C
1	26%	32%	28%

■ En una celda vacía, por ejemplo E1, inserta la función MEDIA.GEOM y especifica el rango de las celdas que contienen los datos en la caja de texto correspondiente a "Número1".



Pulsa el botón "Aceptar" y te aparecerá el resultado en forma de número fraccionario.

■ Para que aparezca en puntos porcentuales, sólo especifica un formato de celda "Porcentaje" con un decimal a la celda E1. Para ello, posiciona el cursor sobre dicha celda y pulsa el botón derecho del ratón para que aparezca el menú contextual. Después elige la opción "formato de celdas".



En el cuadro de diálogo emergente selecciona la categoría porcentaje y disminuye el número de posiciones decimales a 1.



Pulsa el botón Aceptar y obtendrás el resultado.

a continuación se te presentan problemas de hallar el valor de la media aritmética (según aplique cada uno de los casos).

Problema 01. La estadística de salarios de un grupo de trabajadores es la siguiente: el promedio gana 2,100.00 y sumados es igual a Q 8, 400.00. ¿Cuántos trabajadores son los analizados?

Mes	X
Ene	2
Feb	3
Mar	6
Abr	3
Mayo	9
Jun	4
	27

Problema 02. A continuación, se te presenta una serie de datos simple con la que debes de calcular el valor de la media.

Problema 03. A continuación se presenta el Impuesto sobre la Renta, pagado por un grupo de contribuyentes, en miles de quetzales. Calcular el valor de la media de los impuestos pagados.

ISR Pagado	No. Contribuyentes
35 - 40	4
41 - 46	5
47 - 52	8
53 - 58	14
59 - 64	8
65 - 70	4
71 - 76	3

Problema 04. A continuación se presenta una serie de agrupada en clases. Calcular el valor de la media.

Clases	f	x	fx
35 - 40	4	37.5	150
41 - 46	5	43.5	217.5
47 - 52	8	49.5	396
53 - 58	14	55.5	777
59 - 64	8	61.5	492
65 - 70	4	67.5	270
71 - 76	3	73.5	220.5
	N = 46		Σ 2523

Problema 5. José cosechó del árbol 4 peras, Catalina – 2 peras, y María – 6. Los niños juntaron sus frutas y se las repartieron en forma igualitaria. ¿Cuántas peras obtuvo cada uno? Calcula la media aritmética.

Problema 6. Al curso de refuerzo de estadística asistieron 15 alumnos del Colegio Shalom:

Días	Asistencia de Alumnos
Lunes	10
Martes	10
Miércoles	12
Jueves	11
Viernes	7
Sábado	14
Domingo	8

Problema 7. El piloto estuvo yendo dos horas a la velocidad de 120 km a la hora y horas a la velocidad de 90 km a la hora. Calcule la velocidad media del coche durante la carrera.

Problema 8. La media aritmética de 3 números es 6, y la media aritmética de otros 7 números es 3. ¿Cuál es la media aritmética de estos 10 números?

Problema 09. Se te pide calcular la Mg de 2, 4 y 8.

Problema 10. Los gastos de una empresa en los últimos 5 meses fueron los siguientes:

Meses	Gastos
Marzo	90,000
Abril	80,000
Mayo	60,000
Junio	50,000
Julio	45,000

La administración de la empresa solicita:

- a) La tasa promedio geométrica mensual de los gastos,
- b) Los gastos para el mes de agosto 2015,
- c) El promedio geométrico de gastos mensuales, al mes de julio.

Problema 11. En tarde calurosa del sábado, Cristian un empleado de un kiosco de bebidas sirvió en total 50 helados durante la mañana de ese día. Vendió 5 helados de Q 0.50, 15 de Q 0.75, otras 15 de Q 0.90, y otras 15 de Q 1.10.

Problema 12. Un ciclista recorrió la distancia entre la Ciudad de Guatemala y Zacapa (150 Km.) a una velocidad de 50 Km. por hora y empleó 3 horas. Regresó a una velocidad de 30 Km. por hora y tardó 5 horas. ¿Cuál es la velocidad promedio del recorrido?

Problema 13. Una empresa quiere entregar un pedido de 600 unidades. Asigna igual número de unidades a cada trabajador, si cuenta con tres trabajadores para elaborarlas. Calcula:

- El promedio de unidades por hora.
- La cantidad de horas que necesita cada uno de los trabajadores.
- Comprobación.

Problema 14. El examen final del curso se valora como 3 veces los exámenes parciales y un estudiante obtuvo en el examen final 80 puntos, 75 y 82 en los exámenes parciales. De forma ponderada calcula la nota final.

Problema 15. Calcular el salario promedio de la siguiente tabla de trabajadores, empleando la media aritmética.

Salario (X)	No. de Empleados (F)
Q 15,000	18
Q 20,000	35
Q 25,000	29

Problema 16. Calcular la media geométrica del número de hermanos que tienen Berta, Borja y Diana si tienen 2, 2 y 4 y respectivamente.

Problema 17. En una empresa quieren saber la **proporción media de mujeres** en los diferentes departamentos. Para ello, se recoge el porcentaje de mujeres en los cinco principales departamentos.

Como es la media de porcentajes, calculamos la **media geométrica** que es más representativa.

Porcentaje de mujeres por departamento	
Departamento	Porcentaje
Producción	32,6%
Compras	53,5%
Marketing	28,9%
Recursos Humanos	48,2%
Administración	67,4%

Problema 18. Una aldea sufre un proceso rápido de envejecimiento. El primer año aumentan los mayores de 65 años un 10%, el segundo año, un 20%, el tercer año un 30% y el cuarto año, un 40%. Encuentra la media geométrica si la población inicial es de 100 mayores de 65 años.

Problema 19. El director ejecutivo de la empresa White-Knuckle Airlines desea determinar la tasa de crecimiento promedio en los ingresos con base en las cifras dadas en la tabla. Si la tasa de crecimiento promedio es menor que el promedio industrial del 10%, se asumirá una nueva campaña publicitaria.

Ingresos para White-Knuckle Airlines		
Año	Ingreso	Porcentaje del año anterior
1992	US\$ 50,000	-- --
1993	55,000	55/50 = 1.10
1994	66,000	66/55 = 1.20
1995	60,000	60/66 = 0.91
1996	78,000	78/60 = 1.30

Problema 20. Una muestra de 50 negociantes de antigüedades en el sudeste de Estados Unidos reveló las siguientes ventas (en dólares) en el año pasado:

Ventas (miles de dólares)	frecuencia	Puntos medio X	f * X	FA
\$100 a \$120	5	110000	550000	5
\$120 a \$140	7	130000	910000	12
\$140 a \$160	9	150000	1350000	21
\$160 a \$180	16	170000	2720000	37
\$180 a \$200	10	190000	1900000	47
\$200 a \$220	3	210000	630000	50
TOTAL	50		8060000	

BERNARDINO determina las soluciones de cada problema de media aritmética (y sus tipos) que tu catedrático/a te indique.

LA MEDIANA

Sea $x_1, x_2, \dots, x_n$ una muestra aleatoria de n observaciones. Mediante la escritura $x_{(1)}$ se indica el elemento menor de la muestra; por $x_{(2)}$ el elemento que le sigue al menor y así sucesivamente hasta llegar a $x_{(n)}$ que representa al elemento mayor.

La Mediana, de un conjunto de observaciones es el valor para el cual, cuando todas las observaciones se ordenan de manera creciente, la mitad de éstas es menor que este valor y la otra mitad mayor.

Sea $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ una muestra aleatoria de n observaciones, la Mediana de estos datos se denota y se define de la siguiente manera:

$$\tilde{X} = \begin{cases} X_{(\frac{n+1}{2})} & \text{Si } n \text{ es un número impar} \\ X_{\frac{n}{2}} + X_{\frac{n}{2}+1} & \text{Si } n \text{ es un número par} \end{cases}$$

Por ejemplo:

Tienes las duraciones en horas de un cierto tipo de lámparas de alumbrado público:

612, 623, 666, 744, 883, 898, 964, 970, 983, 1003, 1016, 1022, 1029, 1058, 1085, 1088, 1122, 1135, 1197, 1201.
Como hay 20 datos y se encuentran ordenados, entonces la mediana es dada por:

$$\begin{aligned} \text{Mediana} &= \frac{X_{(\frac{20}{2})} + X_{(\frac{20}{2})+1}}{2} \\ &= \frac{X_{10} + X_{11}}{2} \\ &= \frac{1003 + 1016}{2} \\ &= 1009.5 \end{aligned}$$

Entendámonos: determina el valor de la mediana en los problemas que se te dan a continuación. Desarrolla cada uno en hojas aparte y sigue las instrucciones de tu catedrático(a).

Problema 01. Sea una distribución estadística que viene dada por la siguiente tabla.

x_i	61	64	67	70	73
f_i	5	18	42	27	8

Problema 02. Realiza el cálculo de la mediana de la siguiente serie de números: 5, 3, 6, 5, 4, 5, 2, 8, 6, 5, 4, 8, 3, 4, 5, 4, 8, 2, 5, 4.

Problema 03. Una distribución estadística viene dada por la siguiente tabla:

	[10, 15)	[15, 20)	[20, 25)	[25, 30)	[30, 35)
f_i	3	5	7	4	2

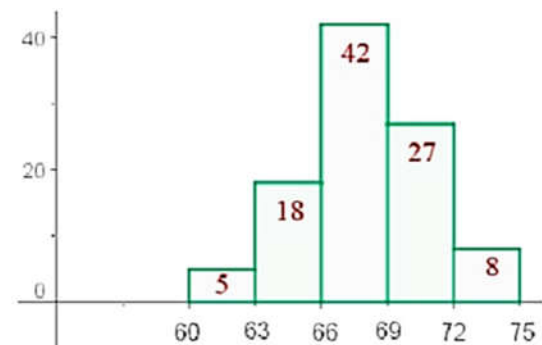
Problema 04. Conforme a los valores de la siguiente tabla de distribución de frecuencias. Las alturas de los jugadores de un equipo de baloncesto vienen dadas por la tabla.

Altura	[170, 175)	[175, 180)	[180, 185)	[185, 190)	[190, 195)	[195, 2.00)
Nº de jugadores	1	3	4	8	5	2

Problema 05. Dada la siguiente distribución estadística, determina la mediana.

	[0, 5)	[5, 10)	[10, 15)	[15, 20)	[20, 25)	[25, ∞)
f_i	3	5	7	8	2	6

Problema 06. A continuación, se te presenta una gráfica de tipo histograma de determinada distribución correspondiente al peso de 100 alumnos del Colegio Shalom.



Problema 07. Calcula la mediana de la distribución de frecuencias descrita en la siguiente tabla.

x_i	f_i	F_i	n_i	$x_i \cdot f_i$
1	4	4	0.08	4
2	4	8	0.08	8
3	8	16	0.16	24
4	7	23	0.14	28
5	5	28	0.1	25
6	10	38	0.2	60
7	7	45	0.14	49
8	5	50	0.1	40
	50			238

Problema 08. En una fundación para el cuidado de los niños, un pediatra obtuvo la siguiente tabla sobre los meses de edad de 50 bebés, en su consulta al momento de andar por primera vez.

Meses	Niños
9	1
10	4
11	9
12	16
13	11
14	8
15	1

Problema 09. Se ha aplicado un test de satisfacción en el trabajo a 88 empleados de una fábrica obteniéndose la tabla de datos adjunta.

Intervalos	f_i
[38-44)	7
[44-50)	8
[50-56)	15
[56-62)	25
[62-68)	18
[68-74)	9
[74-80)	6

Problema 10. Se escoge un salón de clases del Colegio Shalom, con un total de 25 estudiantes a quienes se les pide calificar un programa de determinado canal educativo de 1 a 5, según su preferencia.

(5 = Excelente 4 = Bueno 3 = Regular 2 = No muy bueno 1 = Fatal)

Los resultados son los siguientes:

1 3 3 4 1
 2 2 2 5 1
 4 5 1 5 3
 5 1 4 1 2
 2 1 2 3 5

Determina la mediana de los resultados anteriores.

LA MODA

La moda es la medida que se relaciona con la frecuencia con que se presenta el dato o los datos con mayor incidencia, con lo que se considera la posibilidad de que exista más de una moda para un conjunto de datos. La notación más frecuente es la siguiente: Mo y $\hat{x}$.

Esta medida se puede aparecer tanto para datos cualitativos como cuantitativos. Se dice que cuando un conjunto de datos tiene una moda la muestra es unimodal, cuando tiene dos modas bimodal, cuando la muestra contiene más de un dato repetido se dice que es multimodal y un último caso es cuando ningún dato tiene una frecuencia, en dicho caso se dice que la muestra es amodal.

Ejemplos:

1.- Determinar la moda del siguiente conjunto de datos:

a) 1, 2, 3, 3, 4, 5, 6, 7, 7, 3, 1, 9, 3

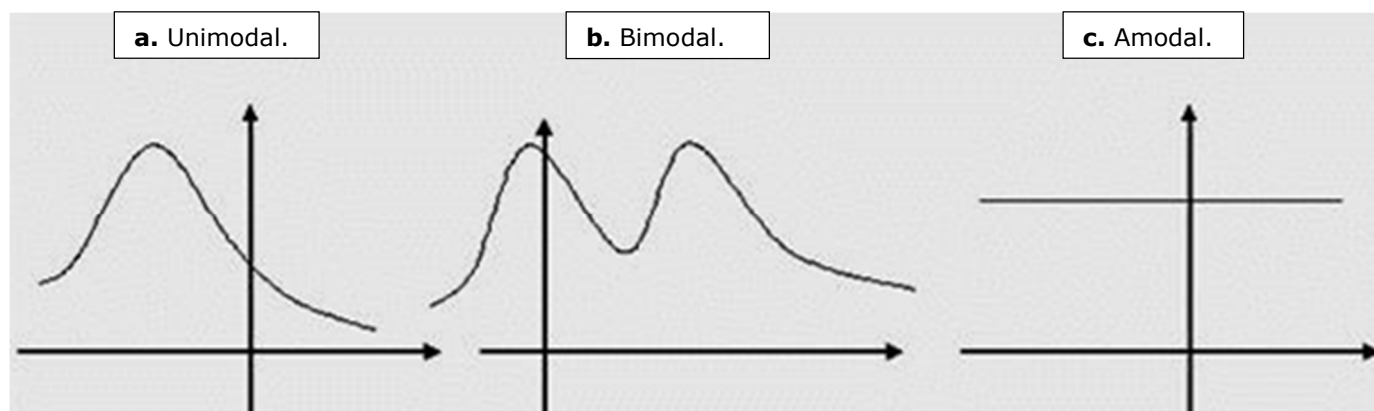
la moda de este conjunto de datos es igual a 3 y si considera unimodal

b) 1, 2, 3, 4, 4, 5, 2, 1, 3, 4, 2, -3, 4, 6, 3, 3

las modas de este conjunto de datos son 3 y 4 ya que ambas tienen la más alta frecuencia, por lo que la muestra es bimodal

c) 1, 2, 3, 4, 5, 6, 7, 8, 9

La muestra no contiene ningún dato repetido por lo que se considera que la muestra es amodal. Gráficamente eso se puede reflejar mediante el análisis de un histograma de frecuencias.



Para determinar la moda de datos agrupados en clases de igual tamaño su cálculo se puede realizar de la siguiente forma:

$$Mo = L_i + \frac{\Delta f_i}{\Delta f_i + \Delta f_s} A$$

En donde:

L_i = límite inferior o frontera inferior.

Δf_i = Exceso de la frecuencia modal sobre la clase modal inferior inmediata.

Δf_s = Exceso de la frecuencia modal sobre la clase modal superior inmediata.

A = Anchura o intervalo de la clase modal.

En ocasiones la expresión para el cálculo de la moda suele presentarse de la siguiente forma:

$$MO = L_i + \frac{f_m - f_{(m-1)}}{2f_m - f_{(m-1)} - f_{(m+1)}} A$$

En donde:

f_m = Frecuencia de clase modal

$f_{(m-1)}$ = Frecuencia de clase premodal

$f_{(m+1)}$ = Frecuencia de clase posmodal

Aunque la expresión se ve un poco diferente en realidad se trata de una misma ecuación, ya que el exceso de la clase modal inferior se puede determinar como:

$$\Delta f_i = f_m - f_{(m-1)}$$

y el exceso de la clase modal superior se determina como:

$$\Delta f_s = f_m - f_{(m+1)}$$

por lo que basta sustituir estos valores en una de ellas para encontrar la otra expresión.

Por ejemplo: determinar a partir de la tabla presentada, en el ejemplo de la media, cual es la moda:

Tabla de frecuencias reportadas por la clínica médica.

Clases: (Datos en años)	Punto medio de cada clase: x_i	Frecuencias de cada clase: f_i
$10 \leq x < 20$	15	8
$20 \leq x < 30$	25	20
$30 \leq x < 40$	35	14
$40 \leq x < 50$	45	8
$50 \leq x < 60$	55	2
$60 \leq x < 70$	65	2
$70 \leq x < 80$	75	1
		55 enfermos atendidos

Ahora bien, identificamos que:

$$Li = 20; \quad f_m = 20; \quad f_{(m-1)} = 8; \quad f_{(m+1)} = 14; \quad A = 10;$$

Al momento de sustituir, tendremos:

$$Mo = L_i + \frac{f_m - f_{(m-1)}}{2f_m - f_{(m-1)} - f_{(m+1)}} A = 20 + \frac{20 - 8}{2(20) - 8 - 14} = 20.666$$

Pese a que el valor de la moda no pueda constituir un dato real, para el ejercicio, se puede asumir que ese es el parámetro de mayor ocurrencia.

Actividad 1: a continuación se te presentan problemas con agrupación de números y/o tablas de frecuencias, en las que debes calcular la moda. Desarrolla cada uno en hojas aparte y sigue las instrucciones de tu catedrático(a).

Problema 01: los siguientes datos corresponden a las notas obtenidas por los alumnos del Estadística: 3, 5, 6, 5, 8, 9, 4, 10, 6, 2. Determina la moda.

Problema 02: sea una distribución estadística que viene dada por la siguiente tabla. Calcula la moda.

x_i	61	64	67	70	73
f_i	5	18	42	27	8

Problema 03: determina la moda de la siguiente serie de números: 5, 3, 6, 5, 4, 5, 2, 8, 6, 5, 4, 8, 3, 4, 5, 4, 8, 2, 5, 4.

Problema 04: se te presenta a continuación una distribución estadística, dada por la siguiente tabla. Encuentra la moda.

	[10, 15)	[15, 20)	[20, 25)	[25, 30)	[30, 35)
f_i	3	5	7	4	2

Problema 05: dada la siguiente tabla estadística, encuentra la moda.

	[0, 5)	[5, 10)	[10, 15)	[15, 20)	[20, 25)	[25, ∞)
f_i	3	5	7	8	2	6

Se te presenta a continuación una serie de ejercicios de repaso general de los temas vistos en el curso de estadística (durante el bimestre), determina la media (según sea su tipo), mediana y moda.

Ejercicio 01. Determina la media, la mediana y la moda de los siguientes números:

25 15 28 29 25 26 21 26

Ejercicio 02: determina la media, la mediana y la moda de los siguientes números.

15 16 19 15 14 16 20 15 17

Ejercicio 03: en un estudio que se realizó en un asilo de ancianos, se tomó las edades de los adultos mayores que pueden caminar sin dificultades. Buscar la media, la mediana y la moda de las siguientes edades, e indicar si es muestra o población.

69 73 65 70 71 74 65 69 60 62

Ejercicio 04. Hallar: Media, moda, mediana, de la siguiente distribución.

X_i	1	2	3	4	5	6
n_i	2	15	9	6	3	1

Ejercicio 05: de las 283 personas encuestadas el año pasado sobre si se encontraban afiliados a algún seguro médico, 86 contestaron afirmativamente.

Con los resultados afirmativos y clasificados según la edad obtenemos la siguiente tabla:

Edad	25-35	35-45	45-55	55-65	
Nº personas	45	23	15	3	86
Marca de clase	30	40	50	60	
$X_i n_i$	1350	920	750	180	3200
Ni	45	68	83	86	

Determina la media, mediana y moda.

INFORMACIÓN (INCLUÍDA EN ESTE DOCUMENTO EDUCATIVO) TOMADA DE:**Sitios web:**

1. <https://es.plusmaths.com/ejercicios/media-aritmetica>
2. <https://www.sangakoo.com/es/temas/media-aritmetica>
3. http://es.onlinemschool.com/math/library/arithmetic_mean/
4. <https://matesnoaburridas.wordpress.com/2017/05/30/media-ponderada/>
5. http://calculo.cc/temas/temas_estadistica/estadistica/teoria/centralizacion.html
6. https://unac.edu.pe/documentos/organizacion/vri/cdcitra/Informes_Finales_Investigacion/IF_JUNIO_2012/IF_CALDERON%20OTOYA_FCA/capitulo%203.pdf
7. <https://www.sangakoo.com/es/temas/media-geometrica>
8. <http://www.universoformulas.com/estadistica/descriptiva/media-geometrica/>
9. https://www.uaeh.edu.mx/docencia/VI_Lectura/licenciatura/documentos/LEC4.pdf
10. <http://www.stadcenterecuador.com/estadisticas/ejercicios/14-basicos/20-ejercicios-resueltos-mediana-moda-y-media-geometrica>