

## MEDIDAS DE DISPERSIÓN EMPLEANDO EXCEL

Las medias de tendencia central o posición nos indican donde se sitúa un dato dentro de una distribución de datos. Las medidas de dispersión, variabilidad o variación nos indican si esos datos están próximos entre sí o si están dispersos, es decir, nos indican cuán esparcidos se encuentran los datos. Estas medidas de dispersión nos permiten apreciar la distancia que existe entre los datos a un cierto valor central e identificar la concentración de los mismos en un cierto sector de la distribución, es decir, permiten estimar cuán dispersas están dos o más distribuciones de datos.

Estas medidas permiten evaluar la confiabilidad del valor del dato central de un conjunto de datos, siendo la media aritmética el dato central más utilizado. Cuando existe una dispersión pequeña se dice que los datos están dispersos o acumulados cercanamente respecto a un valor central, en este caso el dato central es un valor muy representativo. En el caso que la dispersión sea grande el valor central no es muy confiable. Cuando una distribución de datos tiene poca dispersión toma el nombre de distribución homogénea y si su dispersión es alta se llama heterogénea.

### 1) DESVIACIÓN MEDIA O DESVIACIÓN PROMEDIO

La desviación media o desviación promedio es la media aritmética de los valores absolutos de las desviaciones respecto a la media aritmética.

#### 1.1) PROPIEDADES

Guarda las mismas dimensiones que las observaciones. La suma de valores absolutos es relativamente sencilla de calcular, pero esta simplicidad tiene un inconveniente: Desde el punto de vista geométrico, la distancia que induce la desviación media en el espacio de observaciones no es la *natural* (no permite definir ángulos entre dos conjuntos de observaciones). Esto hace que sea muy engorroso trabajar con ella a la hora de hacer inferencia a la población.

Cuando mayor sea el valor de la desviación media, mayor es la dispersión de los datos. Sin embargo, no proporciona una relación matemática precisa entre su magnitud y la posición de un dato dentro de una distribución.

La desviación media al tomar los valores absolutos mide una observación sin mostrar si la misma está por encima o por debajo de la media aritmética.

#### 1.2) MÉTODOS DE CÁLCULO

##### 1.2.1) Para Datos No Agrupados

Se emplea la fórmula:

$$DM = \frac{\sum |x - \bar{x}|}{n}$$

##### Ejemplo ilustrativo:

Calcular la desviación media de la distribución: 3, 8, 8, 8, 9, 9, 9, 18

##### Solución:

Se calcula la media aritmética.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{3 + 8 + 8 + 8 + 9 + 9 + 9 + 18}{8} = 9$$

Se calcula la desviación media.

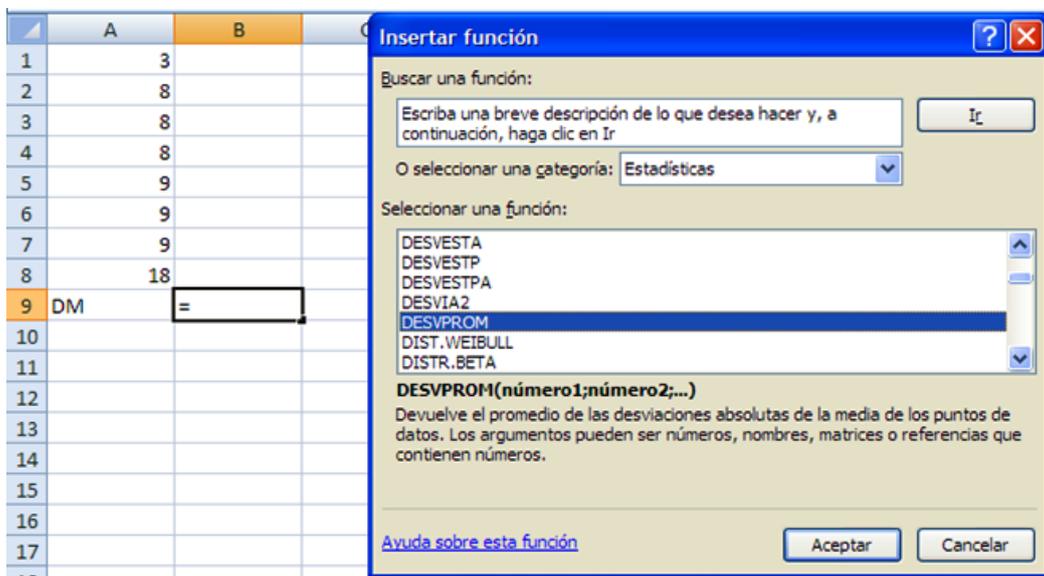
$$DM = \frac{\sum |x - \bar{x}|}{n}$$

$$DM = \frac{|3 - 9| + |8 - 9| + |8 - 9| + |8 - 9| + |9 - 9| + |9 - 9| + |9 - 9| + |18 - 9|}{8}$$

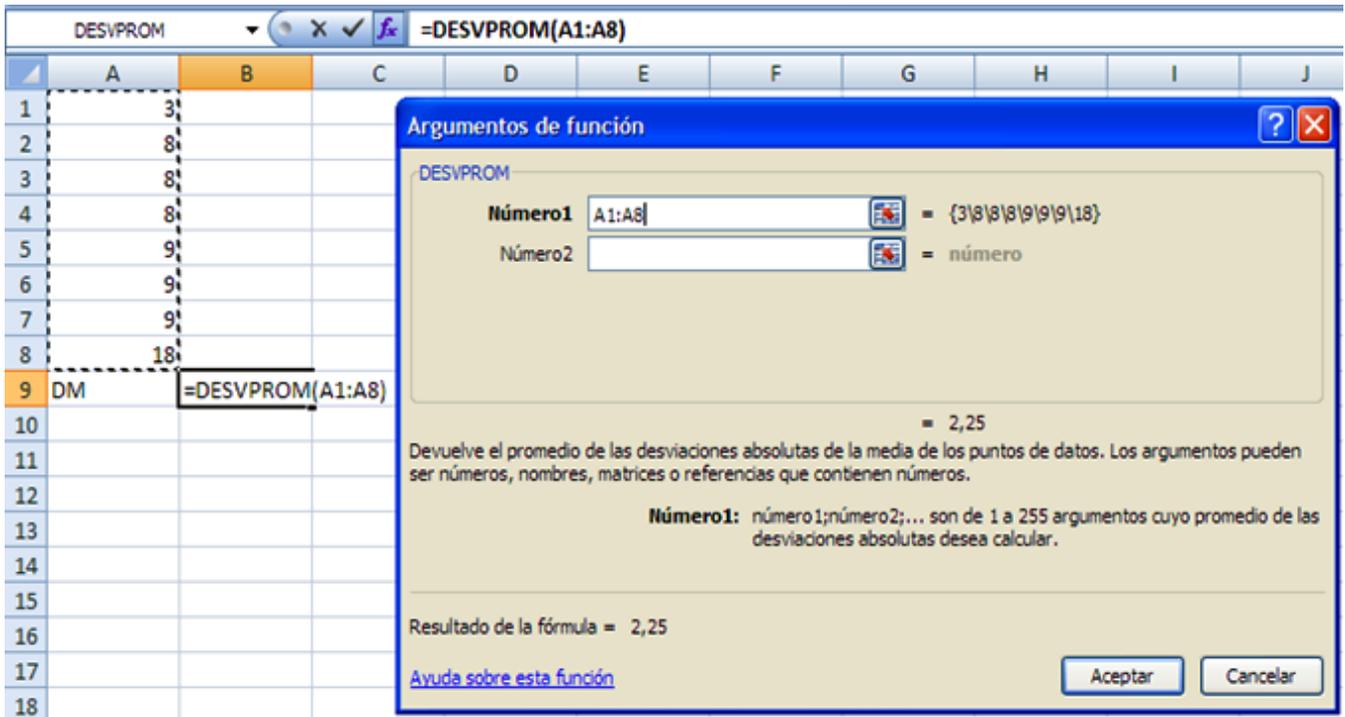
$$DM = \frac{6 + 1 + 1 + 1 + 0 + 0 + 0 + 9}{8} = \frac{18}{8} = 2,25$$

Empleando Excel se calcula de la siguiente manera:

a) Se inserta la función DESVPROM como se muestra en la siguiente figura:



b) Pulsar en Aceptar y seleccionar las celdas.



c) Pulsar en Aceptar.

	A	B	C	D
1	3			
2	8			
3	8			
4	8			
5	9			
6	9			
7	9			
8	18			
9	DM	2,25	=DEVPROM(A1:A8)	

### 1.2.2) Para Datos Agrupados en Tablas de Frecuencia

Se emplea la fórmula:

$$DM = \frac{\sum f|x - \bar{x}|}{n}$$

**Ejemplo ilustrativo:** Calcular la desviación media en base a la siguiente tabla sobre las calificaciones de un estudiante en 12 asignaturas evaluadas sobre 10.

Calificación	Cantidad de asignaturas
6	4
7	2
8	3

9	2
10	1
Total	12

**Solución:**

Se calcula la media aritmética.

$$\bar{x} = \frac{\sum fx}{n} = \frac{4 \cdot 6 + 2 \cdot 7 + 3 \cdot 8 + 2 \cdot 9 + 1 \cdot 10}{12} = \frac{24 + 14 + 24 + 18 + 10}{12} = \frac{90}{12} = 7,5$$

Se llena la siguiente tabla:

$x$	$f$	$ x - \bar{x} $	$f x - \bar{x} $
6	4	1,5	6
7	2	0,5	1
8	3	0,5	1,5
9	2	1,5	3
10	1	2,5	2,5
Total	12		14

Se emplea la ecuación de la desviación media.

$$DM = \frac{\sum f|x - \bar{x}|}{n} = \frac{14}{12} = 1,167$$

**1.2.3) Para Datos Agrupados en Intervalos**

Se emplea la fórmula:

$$DM = \frac{\sum f|xm - \bar{x}|}{n}$$

Donde xm es la marca de clase.

**Ejemplo ilustrativo:**

Calcular la desviación media de un curso de 40 estudiantes en la asignatura de Estadística en base a la siguiente tabla:

Calificación	Cantidad de estudiantes
2-4	6
4-6	8
6-8	16
8-10	10
Total	40

## Solución:

Para calcular la media aritmética se llena la siguiente tabla:

Intervalo	$f$	$xm$	$f \cdot xm$
2-4	6	3	18
4-6	8	5	40
6-8	16	7	112
8-10	10	9	90
Total	40		260

Calculando la media aritmética se obtiene:

$$\bar{x} = \frac{\sum f \cdot xm}{n} = \frac{260}{40} = 6,5$$

Para calcular la desviación media se llena la siguiente tabla:

Intervalo	$f$	$xm$	$ xm - \bar{x} $	$f xm - \bar{x} $
2-4	6	3	3,5	21
4-6	8	5	2,5	12
6-8	16	7	0,5	8
8-10	10	9	2,5	25
Total	40			66

$$DM = \frac{66}{40} = 1,65$$

### TAREA DE INTERAPRENDIZAJE

1) Conteste a las siguientes preguntas.

1.1) ¿Cuál es la diferencia entre medidas de tendencia central y medidas de dispersión?

1.2) ¿Qué permiten o que se logra con las medidas de dispersión?

1.3) ¿En qué caso una distribución de datos toma el nombre de homogénea?. Explique con un ejemplo.

1.4) ¿En qué caso una distribución de datos toma el nombre de heterogénea?. Explique con un ejemplo.

2) Realice un organizador gráfico sobre la desviación media.

3) Calcule la desviación media de las siguientes distribuciones empleando la ecuación y mediante Excel.

3.1) 6, 8, 7, 2, 4, 5, 8 y 9

MD= 1,875

3.2) 10, 4, 8, 9, 6, 10, 8 y 10

MD= 1,625

4) Crear y resolver un ejercicio similar al anterior.

5) Calcule la desviación media empleando los datos de la siguiente tabla:

$x$	$f$
20	2
19	2
18	3
17	4
16	6
15	3
14	3
Total	23

$$DM = 1,46$$

6) Cree y resuelva un ejercicio similar al ejemplo presentado en el cálculo de la desviación media para datos agrupados en tablas de frecuencias.

7) La siguiente tabla muestra la cantidad de dinero que se gasta por semana un grupo de 50 personas. Calcular la desviación media.

Intervalo	$f$
10-20	8
20-30	15
30-40	12
40-50	15
Total	50

$$DM = 9,456$$

8) Cree y resuelva un ejercicio con datos agrupados en intervalos y calcular

8.1) La media geométrica.

8.2) La media armónica.

8.3) La mediana.

8.4) El cuartil N°1 y el cuartil N°3 empleando las ecuaciones y mediante un histograma para la fra(%).

8.5) La Moda empleando la ecuación y mediante un histograma.

8.6) La desviación media.

## 2) VARIANZA Y DESVIACIÓN ESTÁNDAR

La varianza es la media aritmética de los cuadrados de las desviaciones respecto a la media aritmética, es decir, es el promedio de las desviaciones de la media elevadas al cuadrado. La desviación estándar o desviación típica es la raíz de la varianza.

La varianza y la desviación estándar proporcionan una medida sobre el punto hasta el cual se dispersan las observaciones alrededor de su media aritmética.

### 2.1) PROPIEDADES

- La varianza y desviación estándar (o cualquier otra medida de dispersión) indican el grado en que están dispersos los datos en una distribución. A mayor medida, mayor dispersión.

- La varianza es un número muy grande con respecto a las observaciones, por lo que con frecuencia se vuelve difícil para trabajar.

- Debido a que las desviaciones son elevadas al cuadrado y la varianza siempre se expresa en términos de los datos originales elevados al cuadrado, se obtiene unidades de medida de los datos que no tiene sentido o interpretación lógica. Por ejemplo, si se calcula la varianza de una distribución de datos medidos en metros, segundos, dólares, etc., se obtendrá una varianza mediada en metros cuadrados, segundos cuadrados, dólares cuadrados, respectivamente, unidades de medida que no tienen significado lógico respecto a los datos originales.

- Para solucionar las complicaciones que se tiene con la varianza, se halla la raíz cuadrada de la misma, es decir, se calcula la desviación estándar, la cual es un número pequeño expresado en unidades de los datos originales y que tiene un significado lógico respecto a los mismos.

A pesar de lo anterior, es difícil describir exactamente qué es lo que mide la desviación estándar. Sin embargo, hay un resultado útil, que lleva el nombre del matemático ruso Pafnuty Lvovich Chebyshev, y se aplica a todos los conjuntos de datos. Este teorema de Chebyshev establece que para todo conjunto de datos, por lo menos  $1 - 1/k^2$  de las observaciones están dentro de  $k$  desviaciones estándar de la media, en donde  $k$  es cualquier número mayor que 1. Este teorema se expresa de la siguiente manera:

$$1 - \frac{1}{k^2}$$

Así por ejemplo, si se forma una distribución de datos con  $k = 3$  desviaciones estándar por debajo de la media hasta 3 desviaciones estándar por encima de la media, entonces por lo menos

$$1 - \frac{1}{3^2} = 1 - \frac{1}{9} = \frac{9 - 1}{9} = \frac{8}{9} = 0,8889 = 88,89\%$$

*Interpretación:* El 88,89% de todas las observaciones estarán dentro  $\pm 3$  desviaciones de la media.

## 2.2) MÉTODOS DE CÁLCULO

### 2.2.1) Para Datos No Agrupados

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N}$$

Donde:

$x_i$  = *observaciones individuales de la población*

$\mu$  = *media aritmética poblacional*

$N$  = *número de observaciones de la población*

La desviación estándar poblacional se calcula con:

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

La varianza de la muestra se calcula con:

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

Donde:

$x_i$  = *observaciones individuales de la muestra*

$\bar{x}$  = *media aritmética de la muestra*

$n$  = *número de observaciones de la muestra*

La desviación estándar de una muestra se calculó con:

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

#### Notas:

1) Para el cálculo de la varianza de una muestra se divide por  $n-1$  en lugar de  $N$ , debido a que se tiene  $n-1$  grados de libertad en la muestra. Otra razón por la que se divide por  $n-1$  es debido a que una muestra generalmente está un poco menos dispersa que la población de la cual se tomó. Al dividir para  $n-1$  en lugar de  $N$  se cumple con la tendencia y sentido lógico de que la varianza y desviación estándar de la muestra deben tener un valor más pequeño que la varianza y desviación estándar de la población.

2) En la realidad, salvo indicación expresa, no se calcula la varianza y la desviación estándar de la población, ya que para ahorrar tiempo, esfuerzo, dinero, etc. es mejor trabajar con datos que representan a la muestra.

### Ejemplo ilustrativo N° 1

Considere que los siguientes datos corresponden al sueldo de una población: \$350, \$400, \$500, \$700 y \$1000

1) Calcular la desviación estándar.

2) ¿Cuál es el intervalo que está dentro de  $k = 2$  desviaciones estándar de la media?. ¿Qué porcentaje de las observaciones se encuentran dentro de ese intervalo?

#### Solución:

1) Para la calcular la desviación estándar se sigue los siguientes pasos:

a) Se calcula la media aritmética.

$$\mu = \frac{\sum x_i}{N} = \frac{350 + 400 + 500 + 700 + 1000}{5} = \frac{2950}{5} = \$ 590$$

$$\mu = \frac{2950}{5}$$

$$\mu = \$ 590$$

b) Se aplica la respectiva fórmula para calcular la varianza

$$\begin{aligned}\sigma^2 &= \frac{\sum (x_i - \mu)^2}{N} \\ \sigma^2 &= \frac{(350 - 590)^2 + (400 - 590)^2 + (500 - 590)^2 + (700 - 590)^2 + (1000 - 590)^2}{5} \\ \sigma^2 &= \frac{57600 + 36100 + 8100 + 12100 + 168100}{5}\end{aligned}$$

$$\sigma^2 = \frac{282000}{5}$$

$$\sigma^2 = \$^2 56400$$

c) Se calcula la desviación estándar.

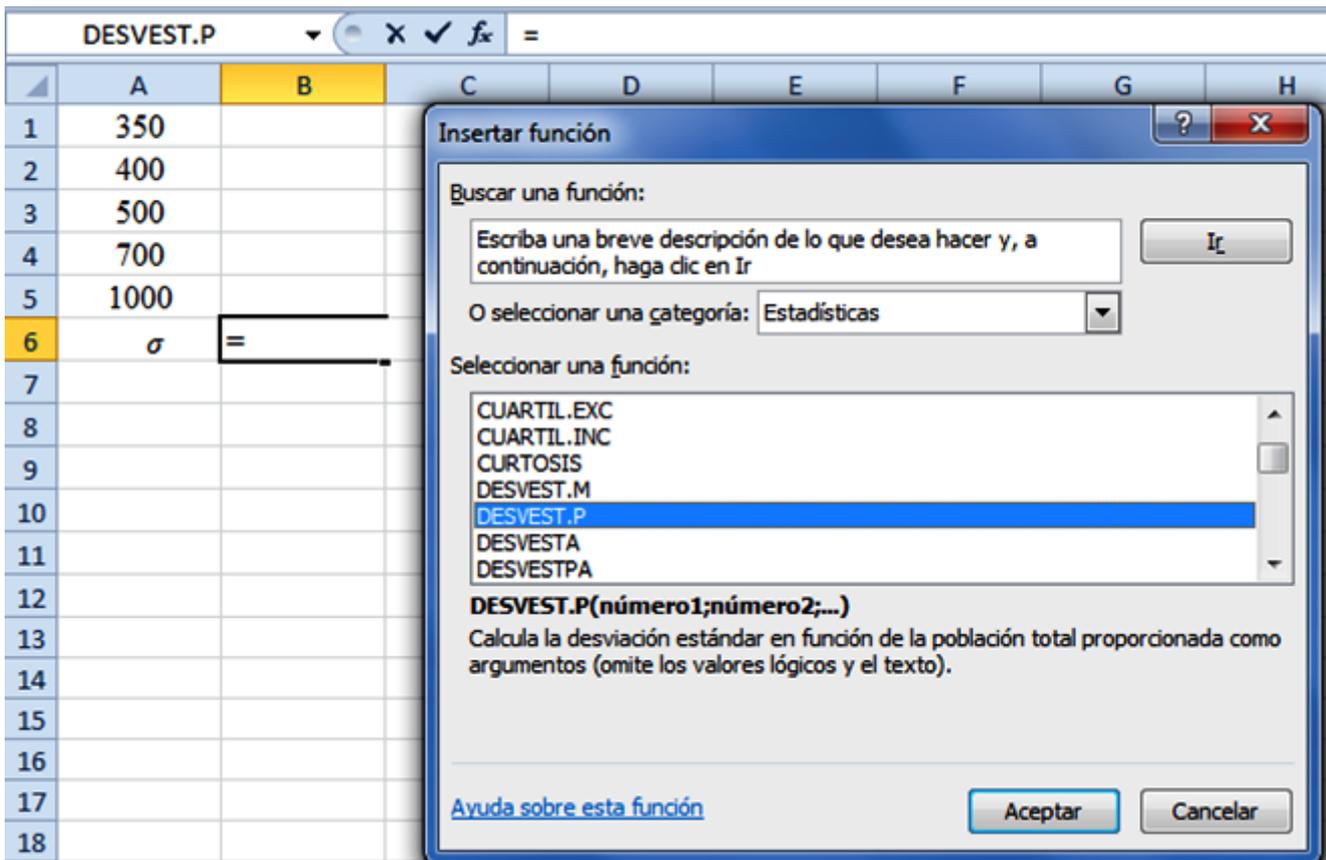
$$\sigma = \sqrt{\sigma^2}$$

$$\sigma = \sqrt{\$^2 56400}$$

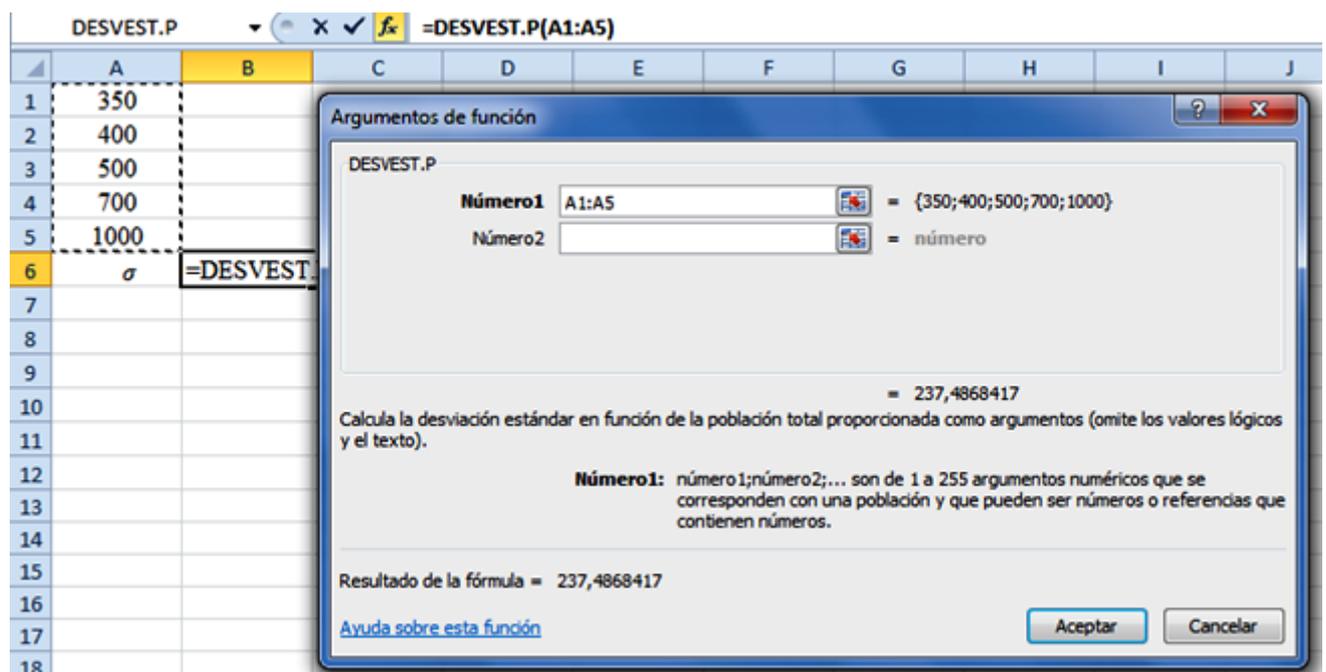
$$\sigma = \$237,4868$$

Empleando Excel se calcula de la siguiente manera:

a) Se inserta la función DESVEST.P. Clic en Aceptar



b) Pulsar en Aceptar y seleccionar las celdas.



c) Pulsar en Aceptar.

DESVEST.P				
	A	B	C	D
1	350			
2	400			
3	500			
4	700			
5	1000			
6	$\sigma$	237,48684	=DESVEST.P(A1:A5)	

2) Cálculo del intervalo de  $k = 2$  desviaciones estándar de la media.

Se transportan 2 desviaciones estándar ( $2 \times \$ 237,4868$ ) = \$ 474,97 por encima y por debajo de la media  $\mu = \$590$

Por lo tanto se tiene un intervalo desde  $\$ 590 - \$474,97 = \$ 115,03$  hasta  $\$ 590 + \$474,97 = \$ 1064,97$

Aplicando el Teorema de Chebyshev

$$1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 1 - \frac{1}{4} = \frac{4 - 1}{4} = \frac{3}{4} = 0,75 = 75\%$$

*Interpretación:* Se puede afirmar de que por lo menos el 75% los sueldos están entre \$ 115,03 y \$ 1064,97

### Ejemplo ilustrativo N° 2

Dos empresas, A y B, venden sobres de café instantáneo de 350 gramos. Se seleccionaron al azar en los mercados cinco sobres de cada una de las compañías y se pesaron cuidadosamente sus contenidos. Los resultados fueron los siguientes.

A	B
350,14	350,09
350,18	350,12
349,98	350,20
349,99	349,88
350,12	349,95

1) ¿Qué empresa proporciona más café en sus sobres?

2) ¿Qué empresa llena sus sobres de manera más consistente?

**Solución:**

a) Se calcula las medias aritméticas.

$$\bar{x}_A = \frac{\sum x_i}{n} = \frac{350,14 + 350,18 + 349,98 + 349,99 + 350,12}{5}$$

$$\bar{x}_A = \frac{1750,41}{5} = 350,082$$

$$\bar{x}_B = \frac{\sum x_i}{n} = \frac{350,09 + 350,12 + 350,20 + 349,88 + 349,95}{5} = \frac{1750,24}{5} = 350,048$$

$$\bar{x}_B = \frac{1750,24}{5} = 350,048$$

*Interpretación:* Como la media aritmética de la empresa A es mayor que la de la empresa B, por lo tanto la empresa A proporciona más café en sus sobres.

b) Se calcula las desviaciones estándar.

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

A	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	B	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
350,14	0,058	0,003364	350,09	0,042	0,001764
350,18	0,098	0,009604	350,12	0,072	0,005184
349,98	-0,102	0,010404	350,20	0,152	0,023104
349,99	-0,092	0,008464	349,88	-0,168	0,028224
350,12	0,038	0,001444	349,95	-0,098	0,009604
Total		0,03328			0,06788

$$s_A = \sqrt{\frac{0,03328}{5 - 1}} = \sqrt{\frac{0,03328}{4}} = 0,0912$$

$$s_B = \sqrt{\frac{0,06788}{5 - 1}} = \sqrt{\frac{0,06788}{4}} = 0,13$$

*Interpretación:* Como la desviación estándar de la empresa A es menor a la desviación estándar de la empresa B, por lo tanto la empresa A es más consistente al llenar los sobres de café.

Empleando Excel se calcula de la siguiente manera:

Se sigue los pasos para el cálculo de la desviación estándar de una población pero utilizando la función DESVEST.M como se muestra en las siguientes figuras:

DESVEST.M				
	A	B	C	D
1	A			
2	350,14			
3	350,18			
4	349,98			
5	349,99			
6	350,12			
7	$S_A$	0,091214	=DESVEST.M(A2:A6)	

DESVEST.M				
	A	B	C	D
1	B			
2	350,09			
3	350,12			
4	350,2			
5	349,88			
6	349,95			
7	$S_B$	0,130269	=DESVEST.M(A2:A6)	

### 2.2.2) Para Datos Agrupados en Tablas de Frecuencia

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum f(x_i - \mu)^2}{N}$$

Donde:

f = frecuencia absoluta.

La desviación estándar poblacional se calcula con:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum f(x_i - \mu)^2}{N}}$$

La varianza de la muestra se calcula con:

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

La desviación estándar de una muestra se calcula con:

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

### Ejemplo ilustrativo

Calcular la desviación estándar de los siguientes datos correspondientes a una muestra.

Calificaciones	f
4	3
5	6
6	4
7	13
8	7
10	6
Total	39

**Solución:**

a) Se llena la siguiente tabla:

Calificaciones	f	fx
4	3	12
5	6	30
6	4	24
7	13	91
8	7	56
10	6	60
Total	39	273

b) Se calcula la media aritmética.

$$\bar{x} = \frac{\sum f \cdot x_i}{n} = \frac{273}{39} = 7$$

c) Se llena la siguiente tabla:

Calificaciones	f	$f x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f(x_i - \bar{x})^2$
4	3	12	-3	9	27
5	6	30	-2	4	24
6	4	24	-1	1	4
7	13	91	0	0	0
8	7	56	1	1	7
10	6	60	3	9	54
Total	39	273			116

d) Se calcula la desviación estándar

$$s = \sqrt{\frac{116}{39 - 1}} = \sqrt{\frac{116}{38}} = \sqrt{3,0526} = 1,747$$

### 2.2.3) Para Datos Agrupados en Intervalos

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum f(xm_i - \mu)^2}{N}$$

Donde:

$f$  = frecuencia absoluta

$xm$  = marca de clase

La desviación estándar poblacional se calcula con:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum f(xm_i - \mu)^2}{N}}$$

La varianza de la muestra se calcula con:

$$s^2 = \frac{\sum f(xm_i - \bar{x})^2}{n - 1}$$

La desviación estándar de una muestra se calcula con:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum f(xm_i - \bar{x})^2}{n - 1}}$$

### Ejemplo ilustrativo

Calcular la desviación estándar de los siguientes datos correspondientes a una muestra.

Intervalo	f
60-65	5
65-70	20
70-75	40
80-85	27
85-90	8
Total	100

**Solución:**

a) Se llena la siguiente tabla:

Intervalo	$f$	$xm$	$f \cdot xm$
60-65	5	62,5	312,5
65-70	20	67,5	1350
70-75	40	72,5	2900
80-85	27	82,5	2227,5
85-90	8	87,5	700
Total	100		7490

b) Se calcula la media aritmética.

$$\bar{x} = \frac{\sum f \cdot xm}{n} = \frac{7490}{100} = 74,9$$

c) Se llena la siguiente tabla:

Intervalo	$f$	$xm$	$f \cdot xm$	$xm_i - \bar{x}$	$(xm_j - \bar{x})^2$	$f(xm_j - \bar{x})^2$
60-65	5	62,5	312,5	-12,4	153,76	768,8
65-70	20	67,5	1350	-7,4	54,76	1095,2
70-75	40	72,5	2900	-2,4	5,76	230,4
80-85	27	82,5	2227,5	7,6	57,76	1559,52
85-90	8	87,5	700	12,6	158,76	1270,08
Total	100		7490			4924

d) Se calcula la desviación estándar.

$$s = \sqrt{\frac{4924}{100 - 1}}$$

$$s = \sqrt{\frac{4924}{99}}$$

$$s = \sqrt{49,737}$$

$$s = 7,052$$

**TAREA DE INTERAPRENDIZAJE**

- 1) Elabore un organizador gráfico sobre la varianza y desviación estándar.
- 2) Consulte sobre la biografía de Pafnuty Lvovich Chebyshev y elabore un organizador gráfico sobre la misma.
- 3) Calcule la desviación estándar de las siguientes distribuciones correspondientes a una población empleando la ecuación y mediante Excel.

3.1) 10, 12, 14, 16, 18, 20 y 40

9,3

3.2) 30, 20, 50, 40, 60, 80 y 90

23,73

4) Cree y resuelva un ejercicio similar al anterior.

5) Cree y resuelva un ejercicio similar al ejemplo presentado para el cálculo de la desviación estándar de una muestra para datos sin agrupar.

6) Calcule la desviación estándar de una muestra dada la siguiente tabla:

$x_i$	$f$
14	6
15	5
18	4
20	12
30	7
40	6

8,903

7) El atraso diario al trabajo en la empresa D & M en el año pasado tuvo un promedio de 78 empleados con una desviación estándar de 13. Se recolectó una muestra de datos para el año en curso y se ubicaron en la siguiente tabla:

Número de empleados atrasados	Días en los que ese número estuvo atrasado
50 - 59	2
60 - 69	4
70 - 79	5
80 - 89	6
90 - 99	3
Total	20

7.1) Calcule la media aritmética.

76,5

7.2) Calcule la mediana empleando la fórmula y mediante un histograma para la fra (%).

78

7.3) Calcule la moda utilizando la fórmula y mediante un histograma para la frecuencia absoluta.

82,5

7.4) Calcule la desviación estándar.

12,4

8) Cree y resuelva un ejercicio similar al anterior.

### 3) OTRAS MEDIDAS DE DISPERSIÓN

#### 3.1) RANGO, AMPLITUD TOTAL O RECORRIDO

Dada una serie de valores  $x_1, x_2, \dots, x_n$ , su recorrido es la diferencia aritmética entre el máximo y el mínimo de estos valores.

$$Re = x_{\text{máx}} - x_{\text{mín}}$$

Es la medida de dispersión más sencilla y también, por tanto, la que proporciona menos información. Además, esta información puede ser errónea, pues el hecho de que no influyan más de dos valores del total de la serie puede provocar una deformación de la realidad.

#### Ejemplo ilustrativo:

Calcula el rango de las siguientes distribuciones:

- 1) 4, 6, 8, 10, 12, 14, 16
- 2) 5, 10, 13, 13, 14, 15, 17

#### Solución:

$$Re_1 = 16 - 4 = 12$$

$$Re_2 = 17 - 5 = 12$$

En Excel se inserta la función  $\text{MAX}(\text{Celdas}) - \text{MIN}(\text{Celdas})$  como muestra la siguiente figura:

	A	B	C	D	E	F	G	H	
1	4				5				
2	6				10				
3	8				13				
4	10				13				
5	12				14				
6	14				15				
7	16				17				
8									
9	Recorrido	12	=MAX(A1:A7)-MIN(A1:A7)		Recorrido	12	=MAX(E1:E7)-MIN(E1:E7)		

Ambas series tienen rango 12, pero están desigualmente distribuidas, pues mientras la primera se distribuye uniformemente a lo largo de todo el recorrido, la segunda tiene una mayor concentración en el centro.

La amplitud es una medida de dispersión cuya ventaja es la facilidad con que se calcula. Tiene en cambio las siguientes desventajas:

- En su cálculo sólo intervienen dos elementos del conjunto.
- Al aumentar el número de observaciones, puede esperarse que aumente la variabilidad. Puesto que la amplitud no tiene en cuenta el tamaño del conjunto, no es una medida adecuada para comparar la variabilidad de dos grupos de observaciones, a menos que éstos sean del mismo tamaño.

**Nota:** Cuando los datos están agrupados en intervalos se calcula la amplitud sacando la diferencia entre la marca de clase mayor y la marca de clase menor.

### 3.2) AMPLITUD INTERCUARTÍLICA

La amplitud intercuartílica es la distancia entre el tercer cuartil  $Q_3$  y el primer cuartil  $Q_1$ .

Amplitud intercuartílica = tercer cuartil - primer cuartil =  $Q_3 - Q_1$

### 3.3) RANGO SEMI-INTERCUARTIL O DESVIACIÓN CUARTÍLICA

La desviación cuartílica es la mitad de la distancia entre el tercer cuartil y el primero

$$DQ = \frac{Q_3 - Q_1}{2}$$

#### **Ejemplo ilustrativo:**

Si el tercer cuartil = 24 y el primer cuartil = 10. ¿Cuál es la desviación cuartílica?

#### **Solución:**

La amplitud intercuartílica es  $24 - 10 = 14$ ;

Por lo tanto, la desviación cuartílica es:

$$DQ = \frac{14}{2}$$

### 3.4) RANGO PERCENTIL O AMPLITUD CUARTÍLICA

Cada conjunto de datos tiene 99 percentiles, que dividen el conjunto en 100 partes iguales.

La amplitud cuartílica es la distancia entre dos percentiles establecidos.

El rango percentil o amplitud cuartílica 10 a 90 es la distancia entre el 10° y 90° percentiles, definida por

$$\text{Rango percentil } 10 - 90 = P_{90} - P_{10}$$

### **TAREA DE INTERAPRENDIZAJE**

- 1) Realice un organizador gráfico sobre las otras medidas de dispersión.
- 2) Plantee una distribución de datos sin agrupar y calcule el rango, la amplitud intercuartílica, el rango semi-intercuartil y el rango percentil de manera manual y empleando Excel.
- 3) Plantee una distribución de datos agrupados en tablas de frecuencia y calcule el rango, la amplitud intercuartílica, el rango semi-intercuartil y el rango percentil.
- 4) Plantee una distribución de datos agrupados en intervalos y calcule el rango, la amplitud intercuartílica, el rango semi-intercuartil y el rango percentil.

## 4) DISPERSIÓN RELATIVA O COEFICIENTE DE VARIACIÓN

Las medidas de dispersión anteriores son todas medidas de variación absolutas. Una medida de dispersión relativa de los datos, que toma en cuenta su magnitud, está dada por el coeficiente de variación.

El Coeficiente de variación (*CV*) es una medida de la dispersión relativa de un conjunto de datos, que se obtiene dividiendo la desviación estándar del conjunto entre su media aritmética y se expresa generalmente en términos porcentuales.

### 4.1) PROPIEDADES

- Puesto que tanto la desviación estándar como la media se miden en las unidades originales, el *CV* es una medida independiente de las unidades de medición.
- Debido a la propiedad anterior el *CV* es la cantidad más adecuada para comparar la variabilidad de dos conjuntos de datos.

### 4.2) MÉTODOS DE CÁLCULO

Para una población se emplea la siguiente fórmula:

$$CV = \frac{\sigma}{\mu} \cdot 100\%$$

Donde:

*CV* = Coeficiente de variación

$\sigma$  = desviación estándar de la población

$\mu$  = media aritmética de la población

Para una muestra se emplea la siguiente fórmula:

$$CV = \frac{s}{\bar{x}} \cdot 100\%$$

Donde:

*CV* = Coeficiente de variación

*s* = desviación estándar de la muestra

$\bar{x}$  = media aritmética de la muestra

### Ejemplo ilustrativo N° 1

Mathías, un estudiante universitario, tiene las siguientes calificaciones en las 10 asignaturas que recibe en su carrera: 8, 7, 10, 9, 8, 7, 8, 10, 9 y 10. Josué, un compañero de Mathías, tiene las siguientes calificaciones: 8, 9, 8, 7, 8, 9, 10, 7, 8 y 10. ¿Cuál estudiante tiene menor variabilidad en sus calificaciones?

**Solución:**

Como se está tomando en cuenta todas las asignaturas, se debe calcular el coeficiente de variación poblacional.

*Sin agrupar los datos empleando Excel se calcula el coeficiente de variación tal como se muestra en la siguiente figura:*

	A	B	C	D	E	F	G	H
1	<b>Mathias</b>				<b>Josué</b>			
2	8				7			
3	7				7			
4	10				8			
5	9				8			
6	8				8			
7	7				8			
8	8				9			
9	10				9			
10	9				10			
11	10				10			
12	$\mu$	8,6	=PROMEDIO(A2:A11)		$\mu$	8,4	=PROMEDIO(E2:E11)	
13	$\sigma$	1,1135529	=DESVEST.P(A2:A11)		$\sigma$	1,0198039	=DESVEST.P(E2:E11)	
14	CV	12,948289	=(B13/B12)*100		CV	12,140523	=(F13/F12)*100	

Agrupando los datos en tablas de frecuencias se calcula así:

a) Se agrupa las calificaciones y se realiza el cálculo la media aritmética

Para Mathías se obtiene:

Calificaciones ( $x_i$ )	$f$	$fx_i$
7	2	14
8	3	24
9	2	18
10	3	30
Total	10	86

$$\mu = \frac{\sum fx_i}{N} = \frac{86}{10} = 8,6$$

Para Josué se obtiene:

Calificaciones ( $x_i$ )	$f$	$fx_i$
7	2	14
8	4	32
9	2	18
10	2	20
Total	10	84

$$\mu = \frac{\sum fx_i}{N} = \frac{84}{10} = 8,4$$

b) Se calcula la desviación estándar

Para Mathías se obtiene:

Calificaciones ( $x_i$ )	$f$	$fx_i$	$(x_i - \mu)^2$	$f(x_i - \mu)^2$
7	2	14	2,56	5,12
8	3	24	0,36	1,08
9	2	18	0,16	0,32
10	3	30	1,96	5,88
Total	10	86		12,4

$$\sigma = \sqrt{\frac{\sum f(x_i - \mu)^2}{N}} = \sqrt{\frac{12,4}{10}} = 1,1136$$

Para Josué se obtiene:

Calificaciones ( $x_i$ )	$f$	$fx_i$	$(x_i - \mu)^2$	$f(x_i - \mu)^2$
7	2	14	1,96	3,92
8	4	32	0,16	0,64
9	2	18	0,36	0,72
10	2	20	2,56	5,12
Total	10	84		10,4

$$\sigma = \sqrt{\frac{\sum f(x_i - \mu)^2}{N}} = \sqrt{\frac{10,4}{10}} = 1,0198$$

c) Se calcula el coeficiente de variación

Para Mathías se obtiene:

$$CV = \frac{\sigma}{\mu} = \frac{1,1136}{8,6} = 0,129 = 12,9\%$$

Empleando Excel es como muestra la siguiente figura:

	A	B	C	D	E	F	G
1	$x_i$	$f$	$f(x_i - \mu)^2$				
2	7	2	5,12	=B2*(A2-SB\$7)^2			
3	8	3	1,08	=B3*(A3-SB\$7)^2			
4	9	2	0,32	=B4*(A4-SB\$7)^2			
5	10	3	5,88	=B5*(A5-SB\$7)^2			
6	Total	10	12,4	=SUMA(C2:C5)			
7	$\mu$	8,6		=SUMAPRODUCTO(A2:A5;B2:B5)/SUMA(B2:B5)			
8							
9	$\sigma$	1,1135529		=RCUAD(C6/SUMA(B2:B5))			
10							
11	CV	12,948289		=(B9/B7)*100			

Para Josué se obtiene:

$$CV = \frac{\sigma}{\mu} = \frac{1,0198}{8,4} = 0,121 = 12,1\%$$

Empleando Excel es como muestra la siguiente figura:

	A	B	C	D	E	F
1	$x_i$	f	$f(x_i - \mu)^2$			
2	7	2	3,92	=B2*(A2-\$B\$7)^2		
3	8	4	0,64	=B3*(A3-\$B\$7)^2		
4	9	2	0,72	=B4*(A4-\$B\$7)^2		
5	10	2	5,12	=B5*(A5-\$B\$7)^2		
6	Total	10	10,4	=SUMA(C2:C5)		
7	$\mu$	8,4	=SUMAPRODUCTO(A2:A5;B2:B5)/SUMA(B2:B5)			
8						
9	$\sigma$	1,0198039	=RCUAD(C6/SUMA(B2:B5))			
10						
11	CV	12,140523	=(B9/B7)*100			

*Interpretación:* Por lo tanto, Josué tiene menor variabilidad en sus calificaciones

### Ejemplo ilustrativo N° 2

Se saca una muestra de un curso de la Universidad UTN sobre las calificaciones en las asignaturas de Matemática y Estadística, resultados que se presentan en las siguientes tablas. ¿En qué asignatura existe mayor variabilidad?. Realice los cálculos empleando Excel

Matemática	
Intervalos	f
2 - 4	8
5 - 7	12
8 - 10	20
Total	40

Estadística	
Intervalos	f
2 - 4	8
5 - 7	14
8 - 10	18
Total	40

**Solución:**

Los cálculos para la asignatura de Matemática empleando Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I
1	Intervalos		f	xm			$f(xm_i - \bar{x})^2$		
2	2	4	10	3	=PROMEDIO(A2:B2)		152,1	=C2*(D2-SB\$7)^2	
3	5	7	8	6	=PROMEDIO(A3:B3)		6,48	=C3*(D3-SB\$7)^2	
4	8	10	22	9	=PROMEDIO(A4:B4)		97,02	=C4*(D4-SB\$7)^2	
5	Total		40	=SUMA(C2:C4)			255,6	=SUMA(G2:G4)	
6									
7	$\bar{x}$	6,9	=SUMAPRODUCTO(D2:D4;C2:C4)/C5						
8									
9	s	2,56	=RCUAD(G5/(C5-1))						
10									
11	CV	37,1	=(B9/B7)*100						

Los cálculos para la asignatura de Estadística empleando Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I
1	Intervalos		f	xm			$f(xm_i - \bar{x})^2$		
2	2	4	8	3	=PROMEDIO(A2:B2)		112,5	=C2*(D2-SB\$7)^2	
3	5	7	14	6	=PROMEDIO(A3:B3)		7,875	=C3*(D3-SB\$7)^2	
4	8	10	18	9	=PROMEDIO(A4:B4)		91,125	=C4*(D4-SB\$7)^2	
5	Total		40	=SUMA(C2:C4)			211,5	=SUMA(G2:G4)	
6									
7	$\bar{x}$	6,75	=SUMAPRODUCTO(D2:D4;C2:C4)/C5						
8									
9	s	2,33	=RCUAD(G5/(C5-1))						
10									
11	CV	34,5	=(B9/B7)*100						

*Interpretación:* Por lo tanto el curso presenta mayor variabilidad en la asignatura de Matemática.

**TAREA DE INTERAPRENDIZAJE**

1) Calcule el coeficiente de variación de las siguientes distribuciones de datos referentes a poblaciones. Realice los cálculos empleando la fórmula respectiva y utilizando Excel.

- 1.1) 6, 8, 10, 4, 7, 8, 9, 8, 4 y 6 27,105 %
- 1.1) 6, 6, 8, 8, 8, 8, 9, 9, 5 y 7 17,31 %

2) Calcule el coeficiente de variación de las siguientes distribuciones de datos referentes a muestras. Realice los cálculos empleando la fórmula respectiva y utilizando Excel.

2.1) 6, 8, 10, 9, 8, 7, 8, 10, 9 y 10

15,93 %

2.2) 6, 6, 8, 10, 8, 10, 9, 9, 5 y 7

22,45 %

3) Calcule el coeficiente de variación de manera manual y empleando Excel empleando los datos de la siguiente tabla correspondientes a una población.

$x_i$	$f$
7	4
8	8
9	12
10	6
Total	30

10,88%

4) Cree y resuelva un ejercicio similar al presentado en el ejemplo 1 con datos sin agrupar y agrupando en tablas de frecuencias. Resolver de manera manual y empleando Excel.

5) Calcule el coeficiente de variación de manera manual y empleando Excel utilizando los datos de la siguiente tabla correspondientes a una muestra.

Intervalos	$f$
2 - 4	10
5 - 7	8
8 - 10	22
Total	40

37,1 %

6) Cree y resuelva un ejercicio similar al presentado en el ejemplo 2. Resuelva de manera manual y empleando Excel.

## REFERENCIAS BIBLIOGRÁFICAS

SUÁREZ, Mario, (2011), *Medidas de dispersión*, [www.monografias.com/trabajos89/](http://www.monografias.com/trabajos89/)

SUÁREZ, Mario, (2012), *Interaprendizaje de Estadística Básica*, Universidad técnica de Norte TAPIA, Fausto Ibarra, Ecuador.