

**Debes saber...**

La mediana (Me) de un conjunto de  $n$  datos  $X = \{x_1, x_2, \dots, x_n\}$  se puede calcular de dos formas, dependiendo de  $n$ . En ambos casos, los datos se ordenan de menor a mayor.

Si  $n$  es par, corresponde al promedio de los valores centrales.

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

Si  $n$  es impar, corresponde al dato central.

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

Los cuartiles  $Q_1$  y  $Q_3$  se pueden calcular como la mediana de la mitad inferior y de la mitad superior de los datos, respectivamente.

**Ayuda**

Para utilizar estos indicadores en la comparación de conjuntos de datos, es importante que estos sean del mismo tipo, se encuentren en las mismas unidades, y sus promedios sean iguales o similares.

## Comparación de conjuntos de datos

Un equipo de fútbol femenino necesita contratar a una delantera, para lo cual está observando a dos candidatas. En los últimos 10 partidos del campeonato, registraron las siguientes cantidades de goles:

Canales:	1	0	3	0	4	1	0	0	0	3
Carvajal:	1	1	2	0	1	1	2	1	1	2

La DT del equipo observa que ambas marcaron 12 goles en 10 partidos, con un promedio de 1,2 goles por partido. Para comparar mejor el rendimiento de ellas utiliza otros indicadores, como se muestra:

**Paso 1** → Calcula el rango de goles marcados por ambas jugadoras.

$$R_{\text{Canales}} = 4 - 0 = 4 \quad R_{\text{Carvajal}} = 2 - 0 = 2$$

El mayor rango que presenta Canales puede indicar que en algunos partidos anota muchos goles, pero en otros no anota, mientras que los de Carvajal están más repartidos.

**Paso 2** → Calcula la varianza y la desviación estándar.

Varianza	$\sigma^2_{\text{Canales}} = 2,16$	$\sigma^2_{\text{Carvajal}} = 0,36$
Desviación estándar	$\sigma_{\text{Canales}} \approx 1,47$	$\sigma_{\text{Carvajal}} = 0,6$

Estos indicadores confirman que los goles de Carvajal presentan menor dispersión, lo que se refleja en que cada partido hace una cantidad de goles más similar entre ellos que los de Canales.

**Paso 3** → Calcula los indicadores de posición: mediana y cuartiles.

	$Q_1 = 0$	$Me = 0,5$	$Q_3 = 3$						
Canales:	0	0	0	0	1	1	3	3	4
Carvajal:	0	1	1	1	1	1	2	2	2
	$Q_1 = 1$	$Me = 1$	$Q_3 = 2$						

Se puede confirmar que la dispersión es menor en el caso de Carvajal, observando que las diferencias entre la mediana y los cuartiles  $Q_1$  y  $Q_3$  es menor que en el caso de Canales.



El fútbol femenino ha tenido un gran impulso en nuestro país, a partir de la realización del mundial femenino sub-20 2008. Esto ha llevado a un mayor progreso que se ha reflejado, entre otros, por la Copa Libertadores Femenina, ganada por Colo Colo en 2012.



**¿A cuál de las jugadoras escogerá la DT?**

Discutan en grupos y fundamenten su opinión.

La jugadora escogida por la DT dependerá de lo que busque. Si consideramos los promedios de goles por partido ambos son iguales, pero el promedio de Carvajal resulta mucho más **representativo**, ya que presenta una cantidad de goles por partido más homogénea.

A la DT del equipo puede parecerle más confiable en este sentido, ya que quizás no marcará tantos goles en cada partido, pero sí es muy probable que cada partido marque al menos un gol. Si va a jugar pocos partidos en los que debe asegurar anotar una gran cantidad de goles puede ser conveniente que se escoja a Canales, que si bien no convierte en todos los partidos cuando lo hace convierte más de uno, en general.

### En resumen

Podemos comparar dos o más conjuntos de datos de acuerdo a sus medidas de tendencia central —como el promedio y la mediana— y de la dispersión que muestran. Así, podemos juzgar cuál de ellos posee un promedio más representativo, es decir, aquél conjunto cuyos valores son más cercanos al promedio.

### Razona y comenta

- ¿A qué jugadora habrías escogido tú? Fundamenta y discute con tus compañeros.
- Si en los dos semestres del año tuviste el mismo promedio pero distinta dispersión, ¿en cuál de ellos podría decirse que te fue mejor?

### Repaso

1. **Determina** la mediana y los cuartiles  $Q_1$  y  $Q_3$  de los siguientes conjuntos de datos.

- 4; 3; 5; 3; 5; 4; 4; 1
- 6; 2; 12; 11; 5; 14; 13; 14; 14; 2
- 0; 1; -8; 2; 6; -4; 7; 3; 10; -2
- 7; -6; -1; -3; 3; 5; -6; 2; 7; 6; -5; -2; -9

2. **Construye** en cada caso un conjunto de  $n$  números con las condiciones dadas.

- $n = 7$ , Med = 6
- $n = 12$ , Med = 15
- $n = 30$ ,  $Q_1 = 2$ ;  $Q_3 = 3$
- $n = 29$ , Med = 0,  $Q_1 = -5$ ;  $Q_3 = 1$
- $n = 31$ , Med = 0,  $Q_1 = 0$ ;  $Q_3 = 1$

### Práctica guiada

3. **Determina** en cada caso qué conjunto tiene un promedio más representativo. Para ello utiliza los distintos indicadores como se muestra en la página anterior.

Ejemplo

$$X = \{17; 11; 11; 13; 22; 13; 2; 22; 7; 2\}$$

$$Y = \{17; 10; 8; 16; 17; 14; 7; 13; 7; 11\}$$

Conjunto X:

$$\begin{aligned}\bar{X} &= 12 & R &= 20 \\ \sigma^2 &= 45,4 & \sigma &\approx 6,74 \\ Q_1 &= 7 & Q_2 &= 12 \\ Q_3 &= 17\end{aligned}$$

Conjunto Y:

$$\begin{aligned}\bar{Y} &= 12 & R &= 10 \\ \sigma^2 &= 14,2 & \sigma &\approx 3,77 \\ Q_1 &= 8 & Q_2 &= 12 \\ Q_3 &= 16\end{aligned}$$

El promedio es más representativo para el conjunto Y.

a)  $X = \{12; 2; 5; 7; 6; 10; 6; 10; 0; 2\}$

$$Y = \{10; 0; 5; 7; 9; 12; 3; 9; 2; 3\}$$

b)  $X = \{17; 3; 5; 13; 17; 15; 1; 10; 1; 8\}$

$$Y = \{13; 6; 8; 10; 14; 12; 4; 13; 5; 5\}$$

4. **Analiza** la siguiente situación.

Paulina trabaja en una ferretería, y ha recibido una muestra de diez clavos (medidos en pulgadas) y una de diez varas de madera, medidas en metros.

$$\text{Clavos: } 2; 2,5; 3,4; 2,6; 3,3; 3,5; 2,1; 2,3; 2,1$$

$$\text{Varas: } 3,3; 3; 3,5; 3,2; 3,5; 3,6; 2,7; 3,5; 3,5$$

Ya que las medidas son muy distintas entre sí, va a comparar la dispersión de ambas muestras mediante el coeficiente de variación que se define como:

$$CV = \frac{\sigma}{\bar{X}}$$

En general, este valor se se expresa en porcentaje. Así, obtiene que:

Clavos	Varas
$\bar{X} = 2,56$	$\bar{X} = 3,28$
$\sigma \approx 0,59$	$\sigma \approx 0,28$
$CV \approx \frac{0,59}{2,56} \approx 0,23 = 23\%$	$CV \approx \frac{0,28}{3,28} \approx 0,09 = 9\%$

Lo anterior permite concluir que la muestra de clavos es más heterogénea que la de varas. Por lo tanto, la distribuidora que le envió los clavos parece ser más cuidadosa en las medidas de sus productos.

**Juzga** qué conjunto es más homogéneo, utilizando su coeficiente de variación.

- a)  $X = \{203; 75; 5; 235; 193; 165; 47; 240; 37; 0\}$   
 $Y = \{3; 0; 1; 5; 5; 6; 1; 4; 3; 2\}$
- b)  $X = \{2; 0; 0; 2; 2; 0; 2; 0; 0\}$   
 $Y = \{47; 16; 2; 46; 44; 32; 4; 36; 1; 12\}$

### Aplica

Resuelve los siguientes problemas.

5. Bárbara desea comprobar la efectividad de un fertilizante para plantas. Para ello, cultivó dos maceteros con 20 plantas cada uno; la única diferencia entre ellos fue que a uno le agregó fertilizante. Luego de dos semanas, los tamaños en cm de las plantas eran los siguientes.

Sin fertilizante				
11	10	15	12	13
12	13	10	11	14
13	11	14	12	15
10	12	14	13	12

Con fertilizante				
15	12	15	14	14
13	14	11	11	15
13	12	13	13	15
11	13	16	14	12

- a) El fertilizante, ¿hace crecer más las plantas? Justifica tu respuesta.
- b) Si el fertilizante aumenta el promedio de los tamaños pero aumenta la dispersión, ¿podría decirse que es efectivo? Justifica tu respuesta y discute con tus compañeros.
- c) Si el fertilizante mantiene el promedio de los tamaños pero disminuye la dispersión, ¿podría decirse que es efectivo? Justifica tu respuesta y discute con tus compañeros.
6. Sergio es contador, y está analizando los sueldos de los trabajadores de una empresa, que son los siguientes:

Empresa A:

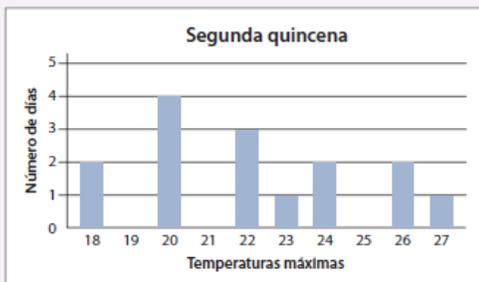
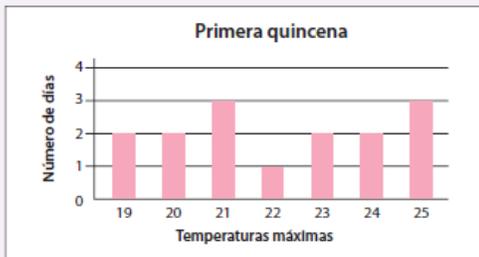
\$300 000	\$300 000
\$300 000	\$300 000
\$6 000 000	

Empresa B:

\$100 000	\$180 000
\$700 000	\$500 000
\$200 000	

- a) ¿Cuál de las empresas presenta una mayor dispersión en sus sueldos?
- b) ¿Qué indicador(es) serían más útiles para analizar la dispersión? Justifica.
- c) En general, ¿hay ocasiones en que los indicadores de dispersión pueden distorsionarse? Da ejemplos y explica qué harías tú en esos casos para solucionarlos.
7. En algunos países de Latinoamérica, las notas van de 1 a 10. Jorge tiene un amigo ecuatoriano, Eusebio, con el que compara sus notas:
- Jorge: 4,5 5 5,2 6,7 6,1 5,8
- Eusebio: 6,2 7,8 3,1 9,6 5,4 7,7
- a) ¿Es útil utilizar el rango para comparar las dispersiones de sus notas? Justifica.
- b) ¿Qué indicador(es) puede(n) resultar más convenientes en este caso? Justifica.

8. Los siguientes gráficos de barra muestran las temperaturas máximas alcanzadas en las dos quincenas de noviembre en Puerto Montt.



- a) Calcula los rangos de las dos variables estadísticas.  
 b) Calcula las desviaciones medias de las dos variables estadísticas. ¿Cuál de las dos variables tiene mayor dispersión?
9. Analiza la información de las tablas y luego responde.

Las tablas muestran las notas de dos cursos diferentes obtenidas en una misma prueba de matemática.

Curso A	
Notas	Frecuencia
7,0	3
6,7	5
6,3	4
6,0	8
4,0	8
3,4	2
3,0	4
Total	34

Curso B	
Notas	Frecuencia
6,0	2
5,5	5
5,3	9
5,2	5
5,1	3
5,0	10
Total	34

- a) ¿Cuál es la desviación estándar de los cursos?  
 b) ¿Qué curso tiene mejor rendimiento? Fundamenta tu respuesta.
10. **Conexiones:** David obtuvo 630 puntos en la PSU de matemáticas, y su hermano Andrés obtuvo 615 cuatro años después. ¿Es cierto que, necesariamente, a David le fue mejor que a Andrés? Investiga y justifica tu respuesta.
11. **Conexiones:** investiga respecto a los valores del coeficiente de variación, para los cuales se considera que un conjunto de datos es muy homogéneo, homogéneo, heterogéneo o muy heterogéneo.
12. **Desafío:** Mariela observa un conjunto X de datos, cuyo coeficiente de variación es, aproximadamente, 27,4%.
- 5; 4; 4; 3; 3; 2; 2; 3; 3; 4; 4; 5
- a) Si Mariela decide restar 1 a cada valor, ¿cuál es el valor del coeficiente de variación ahora?  
 b) Al restar 1 a cada valor, ¿cambia la dispersión de los datos? Utiliza resultados anteriores para justificar tu respuesta.  
 c) En general, ¿se podrá utilizar siempre el coeficiente de variación para comparar dos conjuntos de datos? ¿Qué alternativas crees que podrían utilizarse en los casos en que no? Discute con tus compañeros.

### Reflexiona

- ¿Por qué es importante no utilizar solo el promedio al comparar conjuntos de datos?
- ¿Has visto en los medios de comunicación un mal uso de un promedio? ¿Crees que pueden ser engañosos?