



# ESTADÍSTICA



CC-By 4.0 Ángel Vázquez Hernández 2025



Proyecto STEAM

**4 EDUCACIÓN DE CALIDAD**

(Diseño de [Inma P.nitas](#))

La Agenda 2030 establece la [Educación de Calidad](#) como uno de los Objetivos de Desarrollo Sostenible.

Bienvenide, bienvenida o bienvenido al Módulo IV del Ámbito Científico Tecnológico de ESPA.



Constantemente nos están bombardeando con estudios estadísticos sobre tal o cual cuestión de actualidad: desde la publicidad

hasta la política, pasando por noticias sobre medio ambiente, economía o salud, vivimos entre informes que pretenden ser una descripción de la realidad ¿Pero cómo se hacen esos informes? ¿Hasta que punto son fiables?

En esta situación de aprendizaje vamos a intentar responder a algunas de esas preguntas.

## Sumario

POBLACIÓN, MUESTRA Y TIPOS DE VARIABLES.....	2
Población y muestra.	
Representatividad de una muestra.....	2
Variables cualitativas, cuantitativas discretas y cuantitativas continuas.....	2
PARÁMETROS DE LOCALIZACIÓN O POSICIÓN.....	3
REPRESENTACIONES GRÁFICAS.....	4
Tablas de frecuencias.....	4
Gráficos estadísticos.....	5
PARÁMETROS DE DISPERSIÓN.....	6
REPRESENTACIONES GRÁFICAS DE UNA VARIABLE BIDIMENSIONAL.....	9



*Estatua de Florence Nightingale,  
Waterloo Place, Londres (Dominio  
público)*



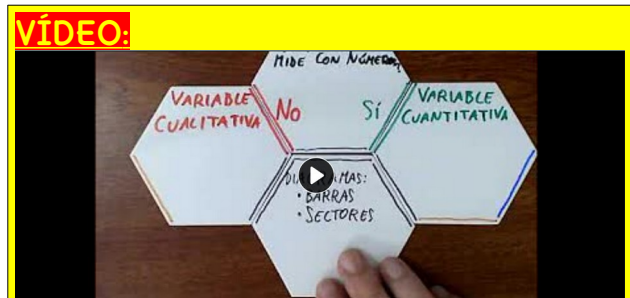
En 1854 [Florence Nightingale](#) llegó a un hospital sobrecargado y la mortalidad era altísima. Nightingale, como casi todo el personal sanitario de mediados del siglo XIX, nada sabía de microbios. Pero era muy buena en estadística, y sus informes sobre la influencia de las condiciones sanitarias en la mortalidad de los soldados heridos durante la Guerra de Crimea influyeron de forma decisiva en el diseño, construcción y gestión de hospitales.

## POBLACIÓN, MUESTRA Y TIPOS DE VARIABLES

### Población y muestra. Representatividad de una muestra

- **Población:** sistema del que se está midiendo una variable.
- **Muestra:** parte del sistema en el que realmente se está midiendo una variable. Se dice que es representativa si podemos suponer que los resultados obtenidos en la muestra pueden ser extrapolados a la población.

### Variables cualitativas, cuantitativas discretas y cuantitativas continuas



Variables cualitativas y cuantitativas

Otros vídeos auto-reproducción



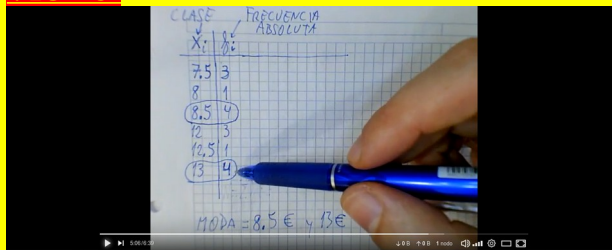
Las variables estadísticas pueden ser cualitativas o cuantitativas.

- **Variables cualitativas:** sin valor numérico. Ejemplos: color, ideología, gustos, sexo, género, etc.
- **Variables cuantitativas:** valores numéricos. Pueden ser discretas o continuas:
  - Discretas: números enteros. Ejemplo: número de hijos.
  - Continuas: números reales. Ejemplos: altura, masa, etc.

## PARÁMETROS DE LOCALIZACIÓN O POSICIÓN

- **Frecuencia absoluta ( $f_i$ ):** número de veces que aparece un valor en un conjunto de datos.
- **Número de datos ( $N$ ):**  $N = \sum f_i$
- **Moda:** valor de mayor frecuencia. Si hay varios de igual frecuencia se habla de distribución bimodal, trimodal, etc.

### VÍDEO:



La moda es el valor (o los valores) que más se repite.

- **Media:** media aritmética de los datos.

$$\bar{x} = \frac{\sum (x_i \cdot f_i)}{N}$$

**SIMULACIÓN DIGITAL:**

¿Cuál es la cantidad promedio de agua por recipiente?

Predecir la Media  
 Media  
 Gradualidades  
 Nivel de Agua del Contenedor

Número de Contenedores: 5

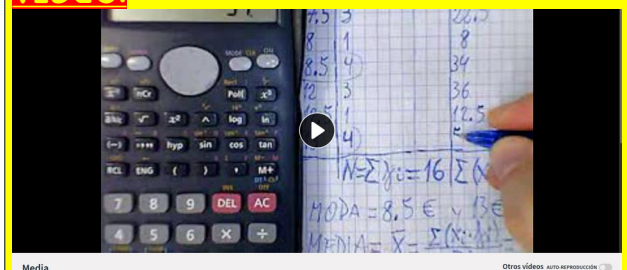
Media: Distribuye y Equilibra

*Media del agua contenida en un conjunto de recipientes*

(Imagen: [Distribuye y Equilibra, CC-BY PhET Interactive Simulations University of Colorado Boulder](https://phet.colorado.edu/))

<https://phet.colorado.edu/>

### VÍDEO:



En el cálculo de la media aritmética es necesario tener en cuenta la frecuencia.

- **Mediana:** valor intermedio de una serie de datos ordenados de menor a mayor. Si el número de datos es par la mediana es la media aritmética de los dos datos intermedios.

**SIMULACIÓN DIGITAL:**

Media y mediana son, junto a la moda, las formas más habituales de representar con pocos valores toda una serie de datos.

(Imagen: [Centro y Variabilidad, CC-By, PhET Interactive Simulations University of Colorado Boulder https://phet.colorado.edu](https://phet.colorado.edu))

## REPRESENTACIONES GRÁFICAS

### Tablas de frecuencias

- Frecuencia relativa ( $f_r$ ):  $f_r = \frac{f_i}{N}$
- Porcentaje:  $f_r \cdot 100$

**VÍDEO:**

7.5	3	22.5	0.1875
8	1	8	0.0625
8.5	4	34	0.25
10	3	24	0.25
12.5	1	12.5	0.25
13	4	52	0.25
16	165	1	

Frecuencia relativa y porcentaje son dos formas de expresar la misma información.


**VÍDEO:**

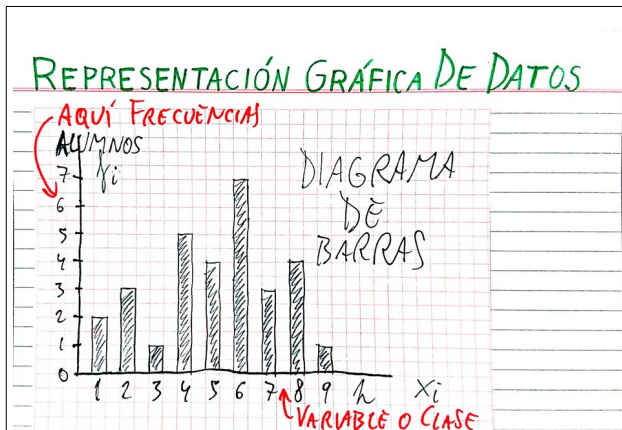
La mediana es el valor intermedio de una serie ordenada de datos.

## Gráficos estadísticos

Se utiliza una multitud de gráficos estadísticos, pero los más habituales suelen ser variantes de los diagramas de barras y de los diagramas de sectores:

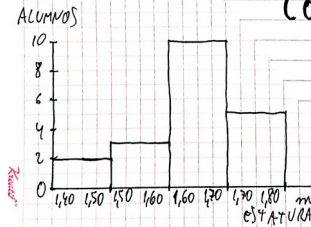


 En un diagrama de barras el tamaño de cada barra es directamente proporcional a la frecuencia de la clase que representa.



Ejemplo de diagrama de barras sencillo. En el eje vertical se ha representado la frecuencia absoluta, y en el eje horizontal la variable o clase.

LAS VARIABLES CONTINUAS, SI SE AGRUPAN EN INTERVALOS, SUELEN REPRESENTARSE MEDIANTE BARRAS CONTIGUAS.



Las variables cuantitativas continuas suelen ser agrupadas en intervalos:

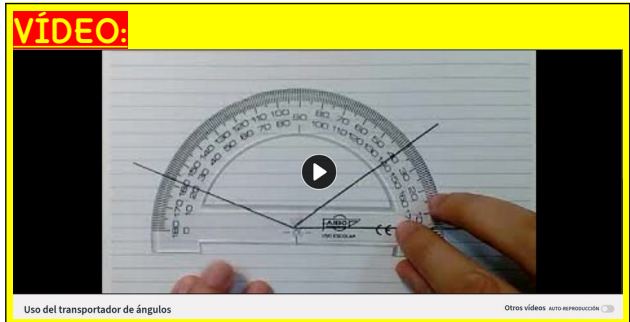
$[a,b]$  desde  $a$  hasta  $b$ , ambos inclusive

$[a,b)$  desde  $a$  hasta  $b$ , excepto  $b$

$(a,b]$  desde  $a$  hasta  $b$ , excepto  $a$

$(a,b)$  desde  $a$  hasta  $b$ , excepto  $a$  y  $b$

Marca de clase:  $x_i = \frac{a+b}{2}$



 En el SI los ángulos se miden en radianes, pero la unidad más habitual es el grado sexagesimal (símbolo  $^\circ$ ). Se pueden medir con un transportador.

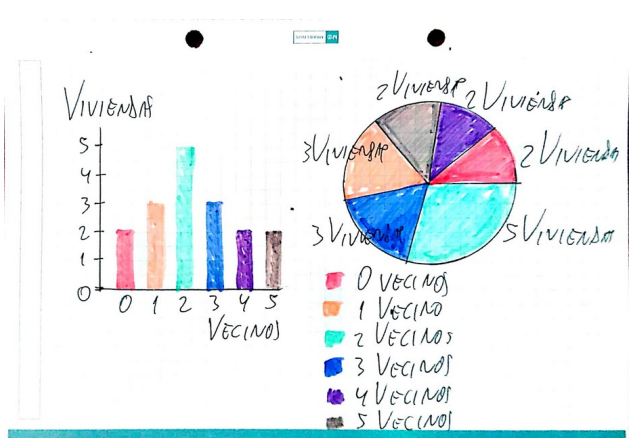
**VÍDEO:**



Diagrama de sectores

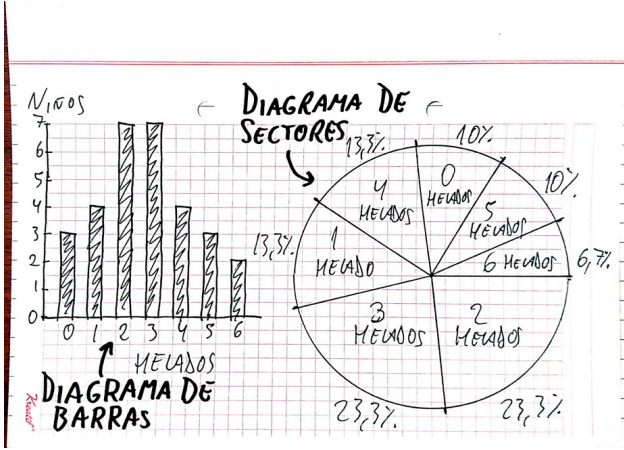


En un diagrama de sectores el tamaño de cada sector es directamente proporcional al tamaño de la clase que representa.




El uso de una leyenda de colores puede facilitar mucho la interpretación de los gráficos.

## PARÁMETROS DE DISPERSIÓN



Ejemplo de diagrama de barras y diagrama de sectores. En el diagrama de sectores, en este caso, se han representado las frecuencias absolutas y los porcentajes (es imprescindible indicar a que clase corresponde cada sector).

**VÍDEO:**



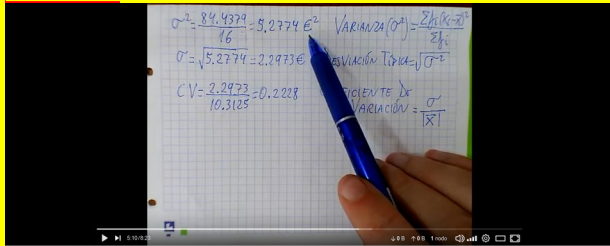
Rango y desviación media

$\bar{x} = 10,3125 \text{ €}$   
 Rango  $R = 13 - 7,5 = 5,5$   
 $DM = \frac{36}{16} = 2,25 \text{ €}$



- Rango o recorrido:** diferencia entre el valor máximo y el mínimo de la variable.
- Desviación media:** media aritmética de los valores absolutos de las desviaciones.

$$DM = \frac{\sum (f_i \cdot |x_i - \bar{x}|)}{N}$$

**VÍDEO:**

Varianza, desviación típica y coeficiente de variación. Otros vídeos auto-educación



- Varianza:  
media aritmética de los cuadrados de las desviaciones.

$$\sigma^2 = \frac{\sum (f_i \cdot (x_i - \bar{x})^2)}{N}$$

- Desviación típica: raíz cuadrada de la varianza.

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

- Coefficiente de variación: cociente entre la desviación típica y el valor absoluto de la media.

$$CV = \frac{\sigma}{|\bar{x}|}$$

1. En la baraja *Científicos y científicas* ¿Cuántas mujeres y cuántos hombres hay? Representa los datos en un diagrama de barras y uno de sectores.
2. En la baraja *Científicos y científicas* ¿Cuántas personas hay en cada área de conocimiento? Representa los datos en un diagrama de barras y uno de sectores.

3. En la baraja *Científicos y científicas* ¿Cómo se distribuyen por antigüedad? Representa los datos en un diagrama de barras y uno de sectores. Calcula moda, media y mediana.
4. Las edades de un grupo de víctimas de delitos sexuales se repartían en los siguientes intervalos: 18-20 años (1), 21-30 años (3), 31-40 años (11), 41-50 años (10), 51-64 años (4), 65-74 años (3). Representa los datos en un diagrama de barras y uno de sectores. Calcula moda, media y mediana.



5. Los precios de la carta de una pizzería, en euros, son:

7.5   7.5   7.5   8  
           8.5   8.5   8.5  
           8.5   12   12   12  
 12.5   13   13   13   13

Realiza un estudio estadístico completo.



6. Los precios, en euros, de las raciones de un bar son:

3   5   5.5   8  
    8   10   10

Realiza un estudio estadístico completo.

7. El número de piezas de fruta en unos tarros de conserva es:

6	6	7	8	6	6	6
7	7	7	8	8	7	6
8	7	7	6	6	6	7
8						

Realiza un estudio estadístico completo.

8. El número de hijos de una serie de familias es:

0	0	3	2	1	1	3
2	2	2	3	0	1	1
2	3	2	0	1	1	2
3						

Realiza un estudio estadístico completo.

9. El número de huevos que hay en un conjunto de nidos es:

1	1	3	2	3	2	2
1	2	3	2	2	1	3
2	1	2				

Realiza un estudio estadístico completo.

10. En una serie de cajas hay las siguientes gominolas:

6	6	6	7	8	7	7
6	6	8	6	6	8	7
7	7	8	8	8	6	6
6	9	9	9			

Realiza un estudio estadístico completo.

11. Los precios, en euros, de los "durum" de un restaurante turco son:

4	4	4.5	4.5	5	5	4.5
6.5	6.5	7	7	7	7	

Realiza un estudio estadístico completo.



12. Los precios, en euros, de la carta de una pizzería son:

8	9	8	8
	8.5	8.5	8.5
8	9	9	
12.5	13.5	12.5	12.5
	13	13	13
12.5			

Realiza un estudio estadístico completo.



13. Los precios de unos juguetes, en euros, son:

2	2	5	8
	7	5	6
6	4	4	2
	8	6	6
5	4	3	2
	6	7	8

Realiza un estudio estadístico completo.

14. Los precios, en euros, de los complementos de un restaurante de comida rápida son:

2.7	2.7	2.7	2.7	2.7	2.7	1.95
4.95	1.75	3.25	1.95	3.5	1.95	

Realiza un estudio estadístico completo.



15. Las calificaciones de un grupo de alumnos y alumnas son las siguientes:

2	4	4	8	7	6	4
5	5	4	6	9	1	0
7	4	4	6	5	5	3
4	1					

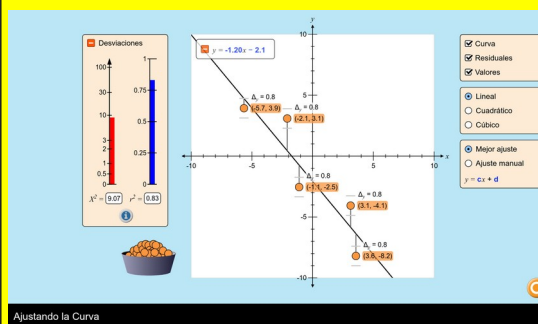
Realiza un estudio estadístico completo.

En las secciones de Ciencia Morada [iA comer!](#) y [De viaje](#) puedes encontrar más ejemplos de datos para practicar.



## REPRESENTACIONES GRÁFICAS DE UNA VARIABLE BIDIMENSIONAL

### SIMULACIÓN DIGITAL:



Ajustando la Curva

PhET

(Imagen: *Ajustando la Curva*. CC-By PhET Interactive Simulations University of Colorado Boulder <https://phet.colorado.edu>)

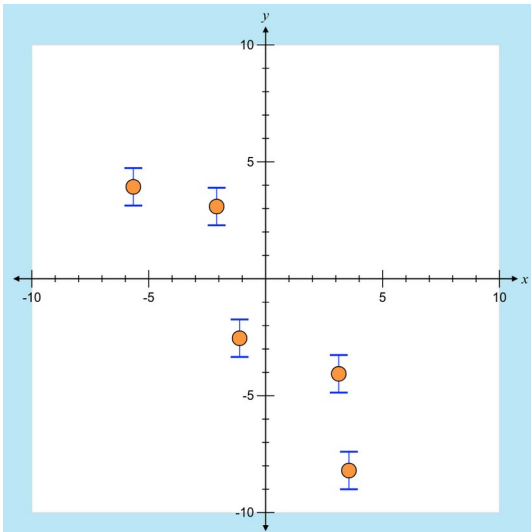


Si tenemos dos series de datos (x e y, por ejemplo) supuestamente relacionados entre ellos por una función

desconocida, podemos buscar esa función intentando que las desviaciones respecto a la función teórica sean lo más pequeñas posibles. Un método muy utilizado para lograrlo es el llamado **método de mínimos cuadrados**.

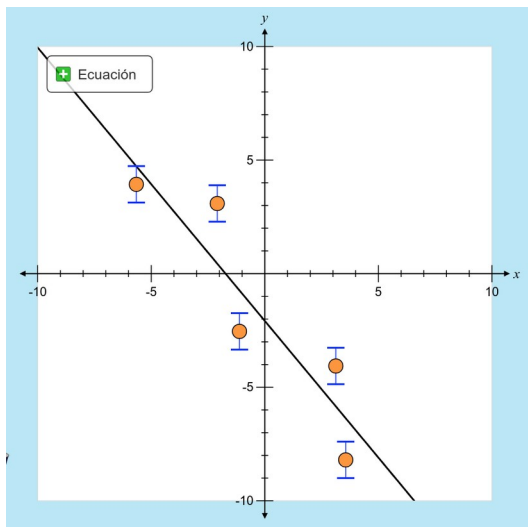
Pasos:

1. **Recopilar los datos y representarlos graficamente:** utilizar los pares de datos para crear una nube de N puntos en un sistema de dos ejes coordenados.



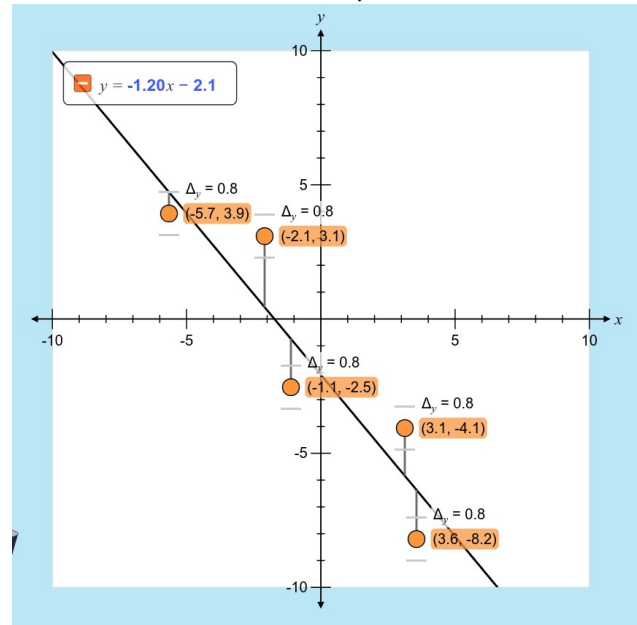
*Nube de puntos obtenida al emparejar dos series de valores.*

2. **Definir el tipo de modelo a utilizar:** representar los datos gráficamente en forma de una nube de puntos, y establecer el tipo de función que más se puede aproximar (lineal, cuadrática, etc).



*En este caso optamos por una función lineal*

3. **Ajustar la línea de regresión:** se busca la función lineal que menos se desvíe de los puntos.



Supongamos que la recta que estamos buscando tiene la forma  $y=mx+n$ . Para hallar los valores  $m$  y  $n$  debemos hacer lo siguiente:

- a) Calcular  $x^2$  y  $xy$  para cada punto.
- b) Calcular  $\Sigma x$ ,  $\Sigma y$ ,  $\Sigma x^2$  y  $\Sigma xy$ . Si estamos trabajando en una tabla será la suma de cada columna.

- c) Calcular la pendiente  $m$ :

$$m = \frac{N \Sigma(xy) - \Sigma x \Sigma y}{N \Sigma(x^2) - (\Sigma x)^2}$$

- d) Calcular la ordenada en el origen  $n$ :

$$n = \frac{\Sigma y - m \Sigma x}{N}$$

- e) Construir la función lineal.

X	Y	X <sup>2</sup>	XY
-5,7	3,9	32,49	-22,23
-2,1	3,1	4,41	-6,51
-1,1	-2,5	1,21	2,75
3,1	-4,1	9,61	-12,71
3,6	-8,2	12,96	-29,52
-2,2	-7,8	60,68	-68,22
N=5			

$$m = \frac{N \cdot \sum(XY) - \sum X \cdot \sum Y}{N \cdot \sum(X^2) - (\sum X)^2}$$

$$m = \frac{5 \cdot (-68,22) - (-2,2) \cdot (-7,8)}{5 \cdot 60,68 - (-2,2)^2} = -1,20$$

$$n = \frac{\sum Y - m \cdot \sum X}{N} = \frac{-7,8 - (-1,20) \cdot (-22)}{5} = -2,09$$

$$y = -1,20x - 2,09$$

*Cálculos de los que se deduce la ecuación de la recta que mejor define a la nube de puntos de este problema.*



**¡CUIDADO!** Algunas veces hay que eliminar del estudio uno o varios puntos. Esto ocurre cuando, debido a alguna circunstancia atípica

o a algún error de medida, uno de los puntos “se aleja” de la nube. No siempre es posible determinar cuando hay que hacerlo, pero a veces eliminar esos puntos anómalos dan lugar a un resultado más fiel a la realidad.

Una vez obtenida la función (una función lineal en este caso, aunque es posible ajustar la nube a otros tipos de funciones) podemos utilizar dicha función para realizar una estimación de valores que desconocemos.

Gracias por tu atención. Puedes dejar un comentario en mi [libro de visitas](#).



**GRACIAS**



CC-BY 4.0 Ángel Vázquez Hernández 2025

Usted es libre de:

- **Compartir** – copiar y redistribuir el material en cualquier medio o formato
- **Adaptar** – remezclar, transformar y crear a partir del material

para cualquier finalidad, incluso comercial.

El licenciador no puede revocar estas libertades mientras cumpla con los términos de la licencia.

Bajo las condiciones siguientes:

- **Reconocimiento** – Debe [reconocer adecuadamente](#) la autoría, proporcionar un enlace a la licencia e [indicar si se han realizado cambios](#). Puede hacerlo de cualquier manera razonable, pero no de una manera que sugiera que tiene el apoyo del licenciador o lo recibe por el uso que hace.
- **No hay restricciones adicionales** – No puede aplicar términos legales o [medidas tecnológicas](#) que legalmente restrinjan realizar aquello que la licencia permite.

Avisos:

- No tiene que cumplir con la licencia para aquellos elementos del material en el dominio público o cuando su utilización esté permitida por la aplicación de [una excepción o un límite](#). Los derechos de los usuarios bajo los límites o las excepciones, como el uso justo o el trato justo, no quedan afectados por las licencias CC. [Más información](#).
- No se dan garantías. La licencia puede no ofrecer todos los permisos necesarios para la utilización prevista. Por ejemplo, otros derechos como los de [publicidad, privacidad, o los derechos morales](#) pueden limitar el uso del material.