



Relaciones causales

Estudiamos más a fondo las variables que se relacionan.

9.1 Relaciones de automóviles usados

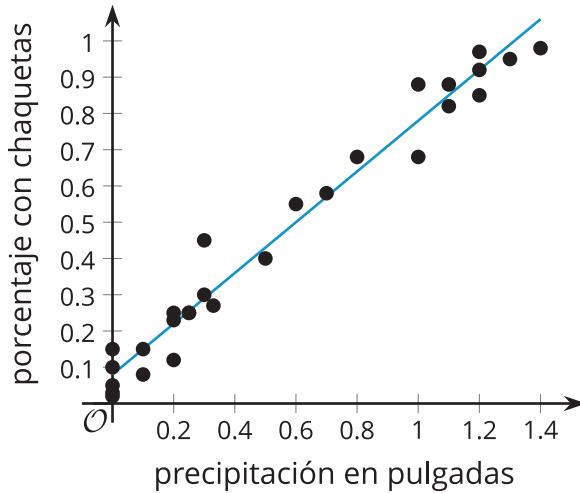
En cada caso, describe la intensidad de la relación entre las variables y el signo que esperas que tenga la relación. Prepárate para explicar tu razonamiento.

1. El precio de un automóvil usado y el precio de venta original de ese automóvil.
2. El precio de un automóvil usado y el número de portavasos que tiene ese automóvil.
3. El precio de un automóvil usado y el número de cambios de aceite que ha tenido ese automóvil.
4. El precio de un automóvil usado y el número de millas recorridas por ese automóvil.

9.2 ¿Causa o efecto?

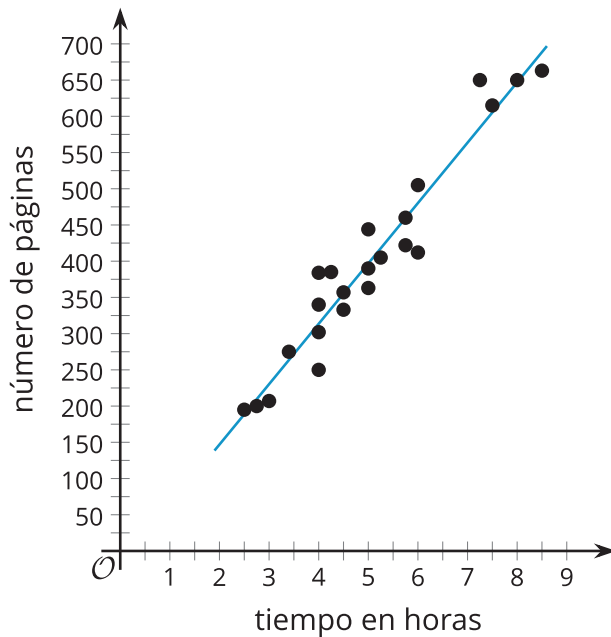
Cada uno de los diagramas de dispersión muestra una relación fuerte. Escribe una o dos oraciones que describan cómo crees que se relacionan las variables.

1.



Durante el mes de abril, Elena lleva un registro del número de pulgadas de lluvia que caen cada día y del porcentaje de personas que asisten a la escuela con chaquetas impermeables ese mismo día.

2.



El club de lectura de una escuela le ofrece a sus miembros una lista de 100 libros. Llevan un registro del número de páginas de cada libro que se lee y de la cantidad de tiempo que toma leer el libro.

A scatter plot with a linear regression line. The x-axis is labeled 'número de boletos no vendidos' and ranges from 0 to 200. The y-axis is labeled 'nivel de ruido en decibeles' and ranges from 0 to 100. There are 20 data points represented by black dots. A blue line represents the linear regression, showing a negative correlation. The data points are roughly as follows:

número de boletos no vendidos	nivel de ruido en decibeles
10	85
20	75
25	73
30	74
35	72
40	72
60	63
60	75
85	63
100	65
105	52
110	55
115	69
140	62
150	64
155	65
160	67
170	63
175	53
175	55
180	51
180	52

En un lugar hacen fiestas los festivales. El día de cada fiesta cuentan el número de boletos que no se vendieron y el nivel de ruido de la fiesta.

A scatter plot showing the relationship between height (altura en pies) on the x-axis and the number of rings (número de anillos) on the y-axis. The x-axis ranges from 0 to 70, and the y-axis ranges from 0 to 100. A blue regression line indicates a positive correlation. The data points are as follows:

altura en pies	número de anillos
38	60
39	65
40	60
41	65
42	65
43	55
44	55
45	65
46	65
47	75
48	80
49	65
50	65
50	70
51	70
52	80
53	85
54	95
55	85
56	85
57	90
58	100
59	85
60	95
65	85
65	100

Los pinos crecen en un bosque. Un arboricultor mide la altura, en pies, de los pinos de un bosque y cuenta el número de anillos que encuentra en las muestras de troncos de los árboles.

En cada caso, describe una pareja de variables que cumpla la condición. Explica tu razonamiento.

1. Las variables tienen una **relación causal**.
2. Las variables están fuertemente relacionadas, pero un tercer factor podría ser la causa de los cambios de las variables.
3. Las variables solo se relacionan débilmente.



¿Estás listo para más?

1. Busca en noticias o en publicidad 2 o 3 afirmaciones de causalidad o correlación. Para cada una de las afirmaciones, lee o mira los artículos o la publicidad y responde estas preguntas.
 - a. ¿Cuál es la afirmación?
 - b. ¿Qué evidencia se presenta para sustentar esa afirmación?
 - c. ¿Parece haber alguna evidencia de causalidad o de correlación? Explica cómo pensaste.
2. Escoge la afirmación que tenga menos evidencia o ninguna evidencia. Describe un experimento u otra forma en que podrías recolectar datos para mostrar si hay correlación o causalidad.



Resumen de la lección 9

Los seres humanos están programados para buscar conexiones y luego usarlas para aprender sobre su entorno. Una forma de identificar conexiones es buscar una pareja de variables que tengan una relación. Para descubrir cómo están relacionadas, necesitamos controlar una de las variables y ver si hay cambios en la otra variable. Por ejemplo, si observamos que las personas que tienden a comer muchas más calorías también tienen un riesgo más alto de tener un ataque cardíaco, podríamos preguntarnos si reducir nuestro consumo de calorías mejoraría nuestra salud.

Un error común que las personas tienden a cometer al usar la estadística es pensar que todas las relaciones entre variables son causales. Los diagramas de dispersión pueden mostrar únicamente una relación entre las dos variables. Para determinar si el cambio de una variable realmente hace que la otra variable cambie, es decir, si hay una **relación causal**, se debe entender mejor el contexto y descartar otras opciones.

Por ejemplo, supondríamos que hay una relación positiva y fuerte entre el número de tablas de *snowboard* alquiladas y las ventas de chocolate caliente durante los meses de septiembre a enero. Esto no significa que un aumento en las tablas de *snowboard* alquiladas cause que las personas compren más chocolate caliente. Tampoco significa que el aumento en las ventas de chocolate caliente cause que las personas alquilen más tablas de *snowboard*. Lo más probable es que haya una tercera variable, como un clima más frío, que haga que ambas variables aumenten al mismo tiempo.

Por otro lado, algunas veces sí hay una relación causal. Se podría establecer una relación positiva y fuerte entre las ventas de chocolate caliente y las ventas de malvaviscos, porque al comprar chocolate caliente es posible que se quiera agregar malvaviscos pequeños a la bebida. Por lo tanto, un aumento en las ventas de chocolate caliente realmente causa un aumento en las ventas de malvaviscos.

Encontrar relaciones con la ayuda del coeficiente de correlación es una muy buena forma de darse cuenta de que hay una conexión entre las variables. Para determinar si la relación es causal, el siguiente paso, por lo general, consiste en diseñar con cuidado un experimento que aísle y controle de manera precisa una de las variables para determinar cómo influye en la otra variable.