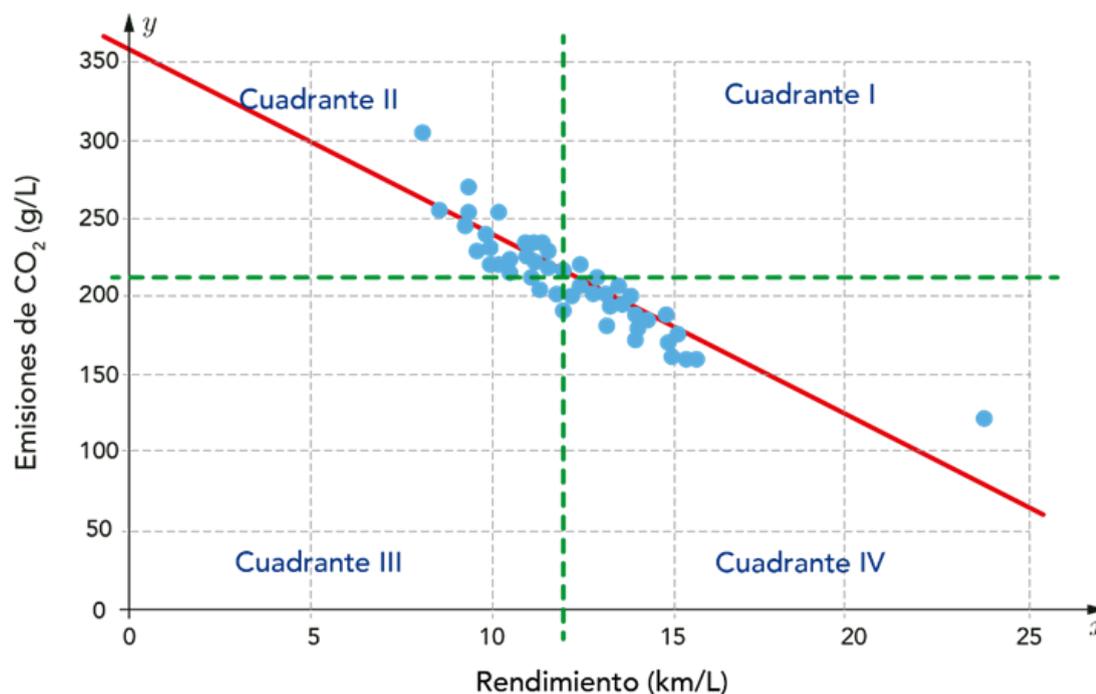


## Guía Práctica

### Calefaccionando el hogar

#### Actividad 1

En el siguiente gráfico se representa la relación entre los rendimientos de una muestra de vehículos y sus emisiones de CO<sub>2</sub>. Las líneas segmentadas representan la media de cada variable.



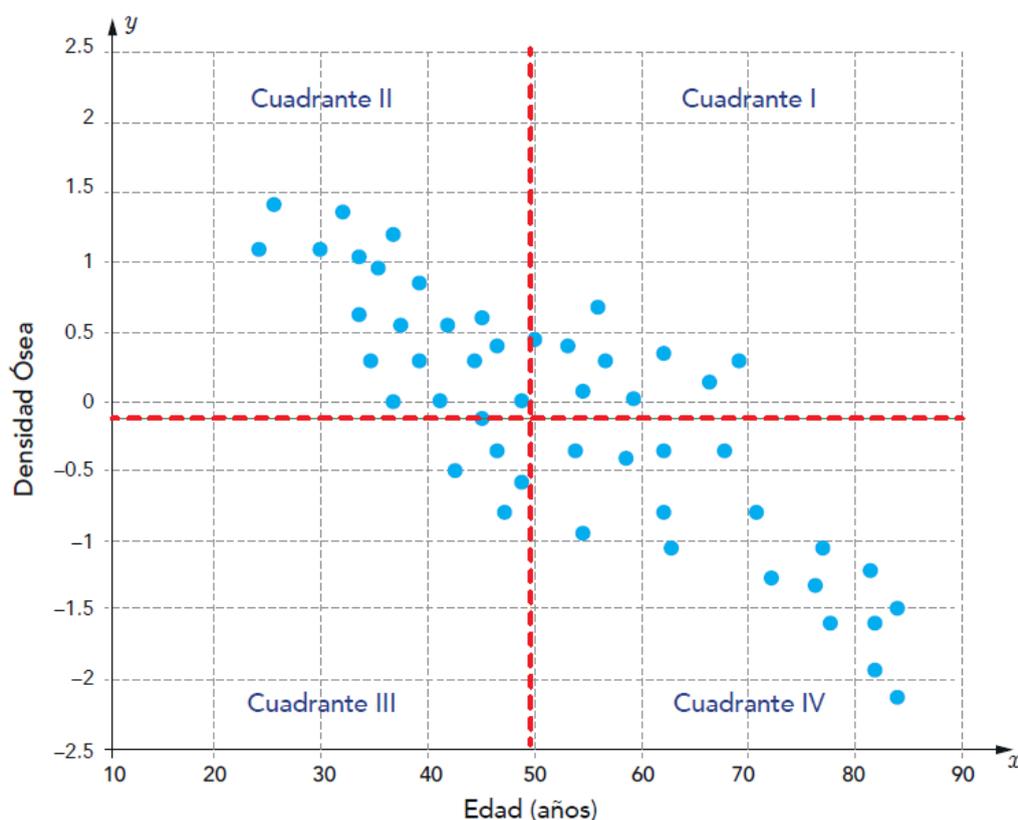
1. En relación con la información del gráfico, determina si las siguientes afirmaciones son verdaderas o falsas.

Afirmaciones	V	F
Los puntos en el Cuadrante IV son automóviles de alto rendimiento.		
Los puntos en el Cuadrante II son automóviles que emiten poco CO <sub>2</sub> .		
La diferencia entre los puntos del Cuadrante IV y la media de rendimiento es positiva.		
La diferencia entre los puntos del Cuadrante III y la media de emisión de CO <sub>2</sub> es positiva.		

2. ¿Qué signo tiene la suma de los productos entre estas dos variables? ¿Por qué?

## Actividad 2

En el siguiente gráfico se representa la relación entre la edad y la densidad ósea de una muestra de personas.



1. ¿Qué signo tendrá la suma de productos  $(x_i - \bar{x})(y_i - \bar{y})$ ? Justifica tu respuesta considerando los cuadrantes dibujados.
2. ¿Qué valor estimas debiera tener el coeficiente de correlación?
3. Si se añade a la muestra una persona de 20 años con densidad ósea menor que  $-1$ , ¿cómo cambiaría el coeficiente de correlación?

### Actividad 3

Considera las siguientes variables y los coeficientes de correlación entre ellas:

Variables	Coefficiente de correlación
A y B	$-0,9$
A y C	$0,9$
B y C	$x$

1. ¿Qué signo tendrá el coeficiente de correlación  $x$ ? ¿Podrías aproximar su magnitud? Justifica tu respuesta.

## Actividad 4

Descarga el archivo llamado [Base de datos - Actividad 4 Guía Práctica](#) y usa las herramientas de Excel o Google Sheets para responder las siguientes preguntas.

Talla de calzado	Estatura (cm)
36	160
36,5	165
37	166
37	166
37	168
37,5	170
38	175
38	173
38	180
39	180
40	185
40	182
41	183
42	184
43	183
43	185
43	185
44	190

1. Calcula la suma de productos  $(x_i - \bar{x})(y_i - \bar{y})$  para esta muestra. ¿Qué representa ese valor?
  
2. Calcula el coeficiente de correlación de la muestra. ¿Qué representa ese valor?

## Solucionario

- 
- Act. 1**
1.
  - Verdadero
  - Falso
  - Verdadero
  - Falso
- 
2. Los puntos en el cuadrante II tienen una diferencia positiva con la media de emisión de CO<sub>2</sub> y negativa con la de rendimiento; mientras que los puntos del cuadrante IV tienen una diferencia negativa con la media de emisión de CO<sub>2</sub> y positiva con la de rendimiento. De esta manera, se tiene que el producto  $(x_i - \bar{x}) \cdot (y_i - \bar{y})$  es negativo para los puntos ubicados en el cuadrante II y IV. Como esos puntos son los que más contribuyen en la suma, ésta será negativa.
- 
- Act. 2**
1. Como la mayoría de los puntos se encuentran en los cuadrantes II y IV, el producto entre las diferencias entre cada coordenada y la media será negativa. Con lo que la suma tendrá ese signo.
- 
2. Aproximadamente debe ser menor que  $-0,6$ .
- 
3. El dato por añadir constituye un dato atípico. Este se aleja de la tendencia lineal, por lo que al añadirlo aumentará la dispersión. Esto último implica que el coeficiente de correlación será más cercano a cero.
- 
- Act. 3**
1. Como A y C crecen conjuntamente, pero B decrece a medida que A crece, se espera que la asociación entre B y C sea negativa. Es decir, el signo del coeficiente de correlación será negativo.  
No es fácil estimar un valor para el coeficiente de correlación, sin embargo, dado que tanto como para A y B, como para A y C la relación es bastante fuerte, se espera que para B y C también haya una asociación fuerte, negativa, por lo que se espera que sea cercano a  $-1$ .
- 
- Act. 4**
1. La suma de productos  $(x_i - \bar{x})(y_i - \bar{y})$  es 358,2. Al ser positiva da cuenta de una relación positiva entre las variables.
- 
2. La correlación es 0,9. Al ser positiva da cuenta de una asociación positiva entre las variables. Al ser cercana a 1 da cuenta de una posible relación lineal fuerte entre ellas.
-