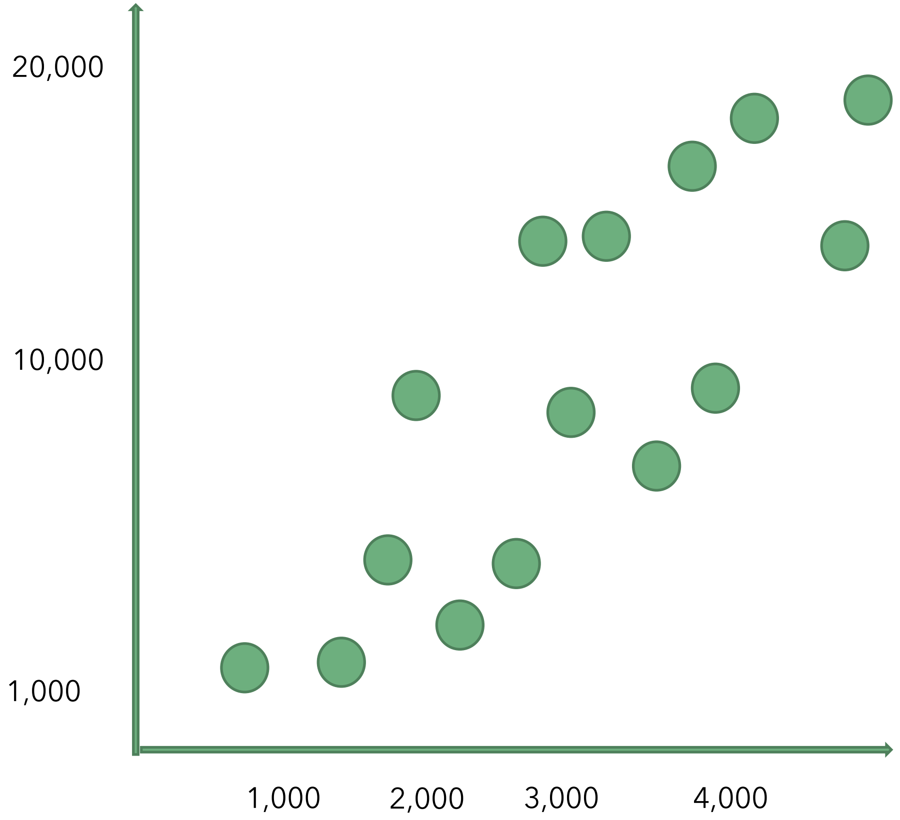
Linear Regression

Rafael Dávila Bugarín

Motivación



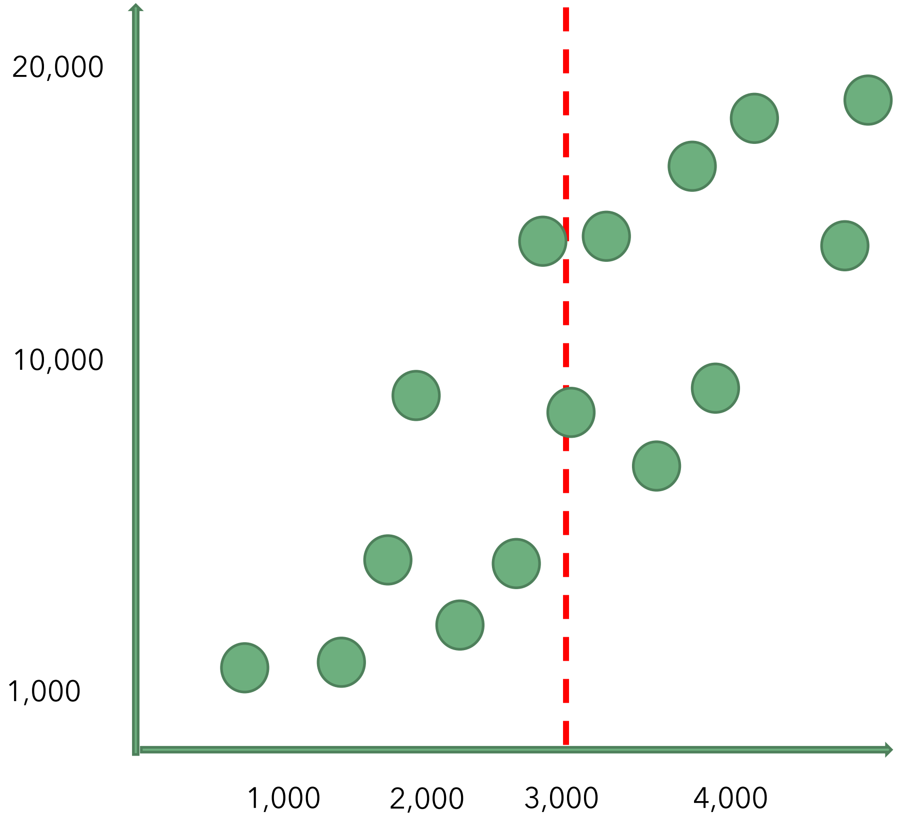
Emisiones CO_2
(Millones de toneladas por año)

#Automóviles
(Millones)

Los puntos en el plano son países, ¿Existe alguna relación entre CO_2 y el número de autos? ¿La relación es fuerte o débil? ¿positiva o negativa?

Mativación

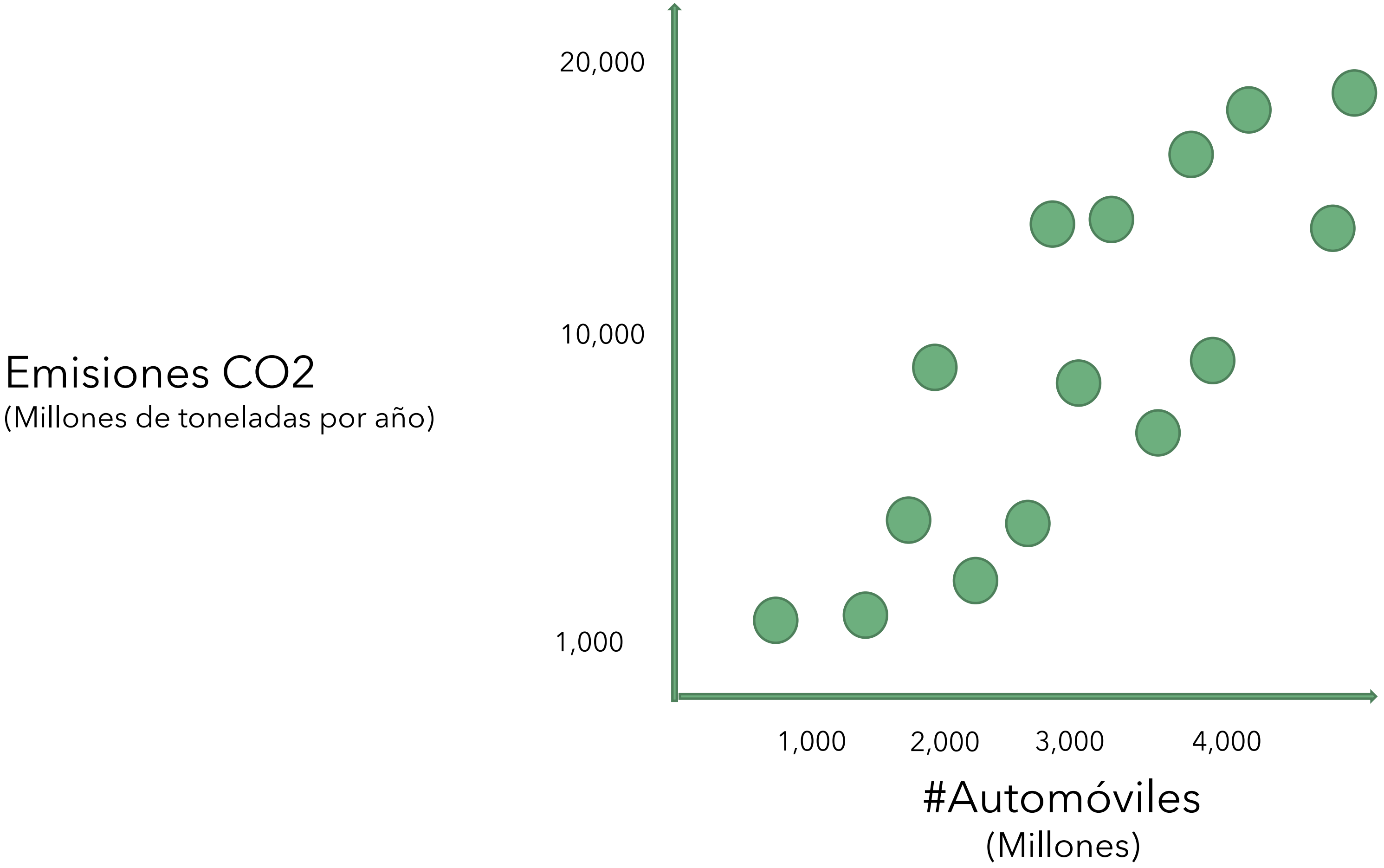
Emisiones CO2
(Millones de toneladas por año)



#Automóviles
(Millones)

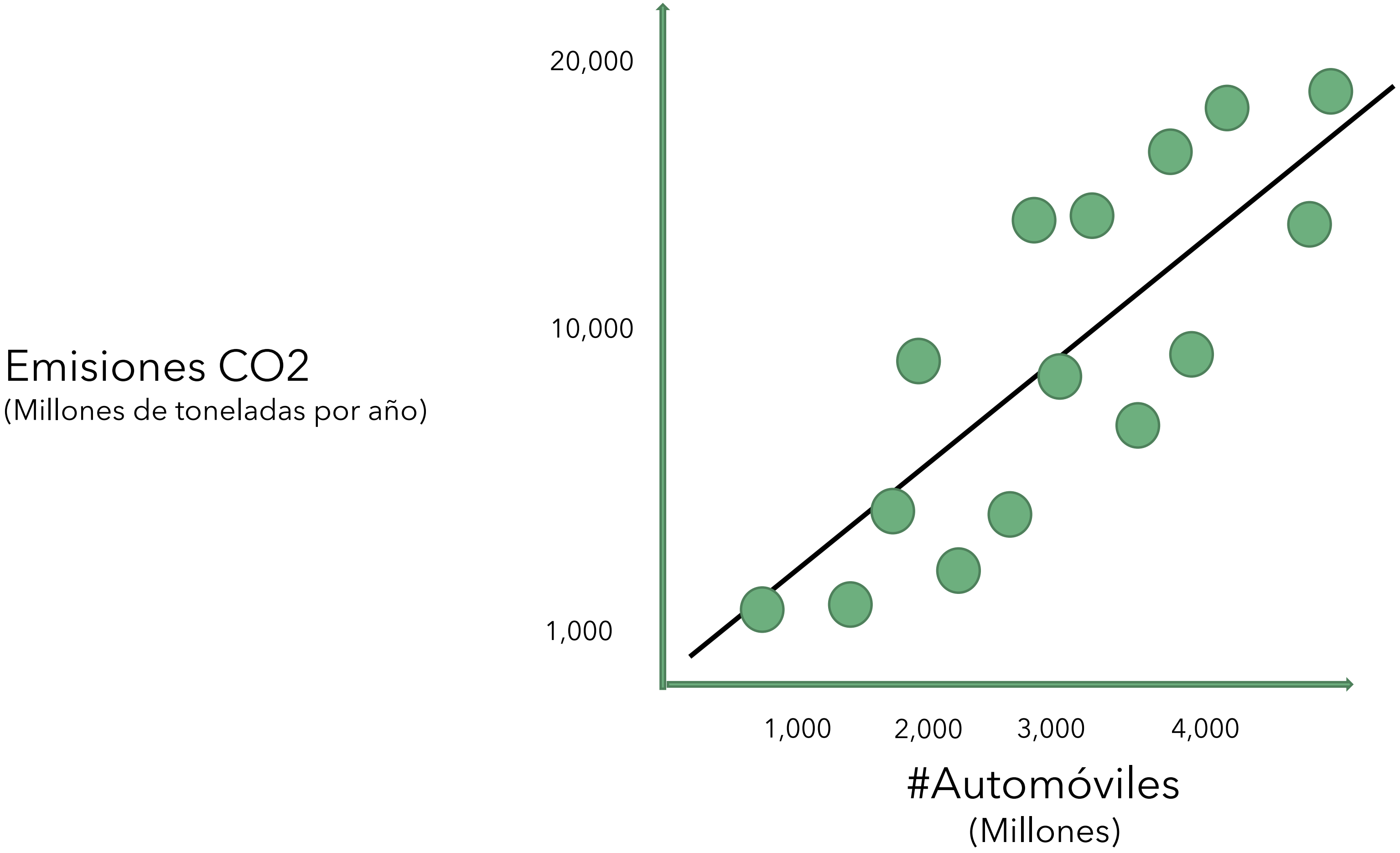
Si quisiera saber un país con 3mil coches cuántas toneladas de CO_2 tiene ¿qué harían?

Mativación



Pareciera que existe una tendencia y esta nos podría ayudar.

Mativación

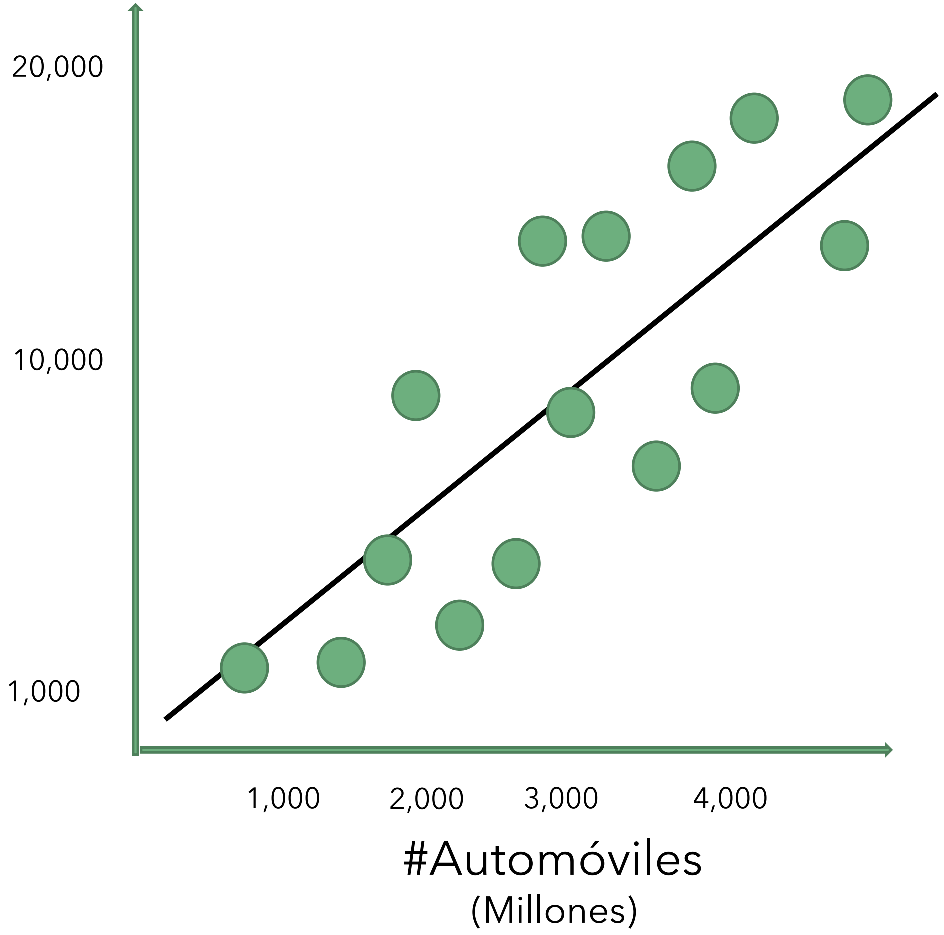


Recordando la ecuación de la recta:

Y=AX+B

Mativación

Emisiones CO2
(Millones de toneladas por año)



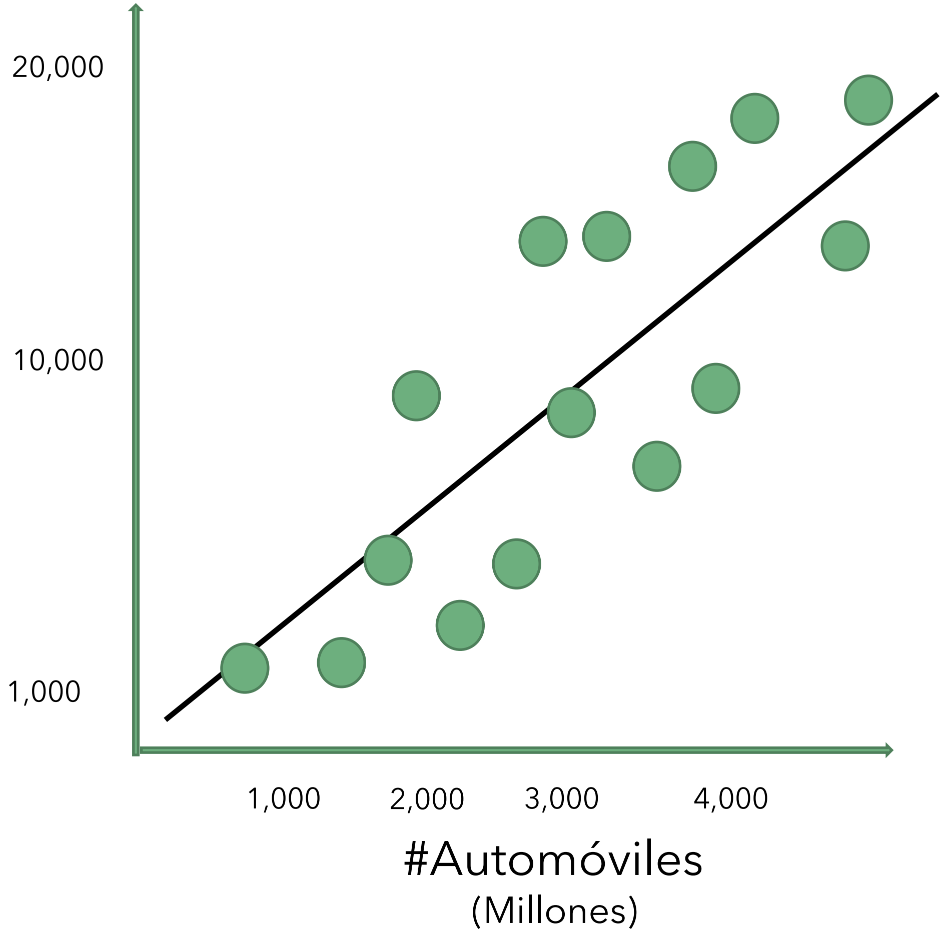
Recordando la ecuación de la recta:

Y=AX+B

CO2=A*(3,000 automóviles)+B

Mativación

Emisiones CO2
(Millones de toneladas por año)



Recordando la ecuación de la recta:

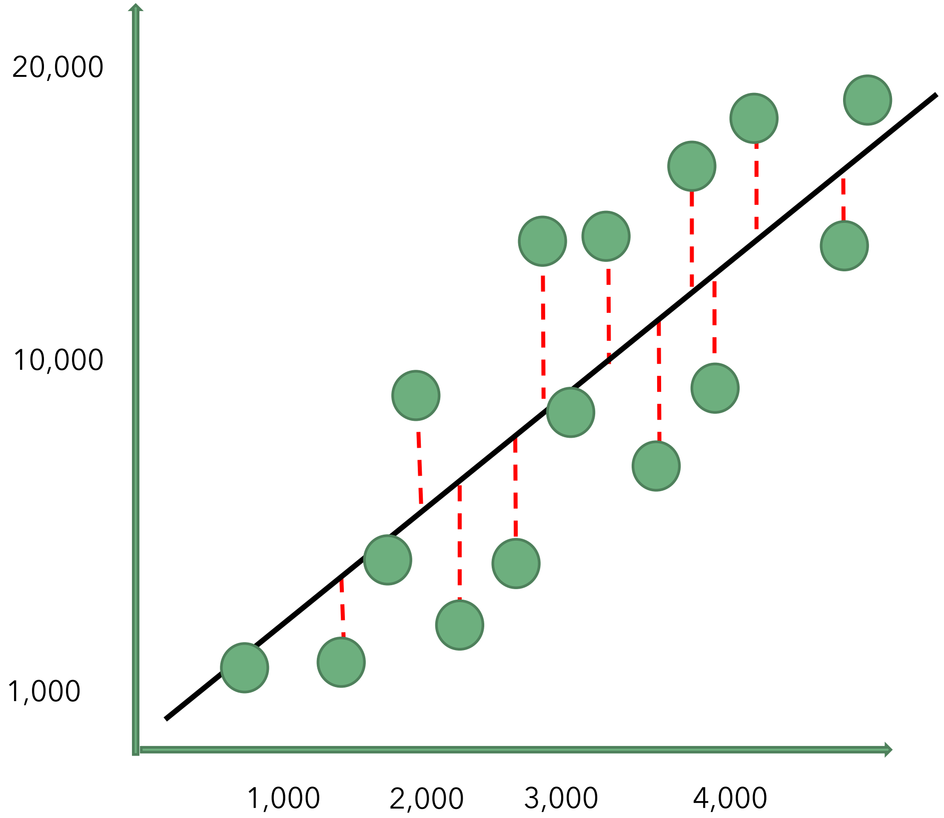
Y=AX+B

CO2=A*(3,000 automóviles)+B

¿Cómo encontramos A y B para que sea la mejor predicción?

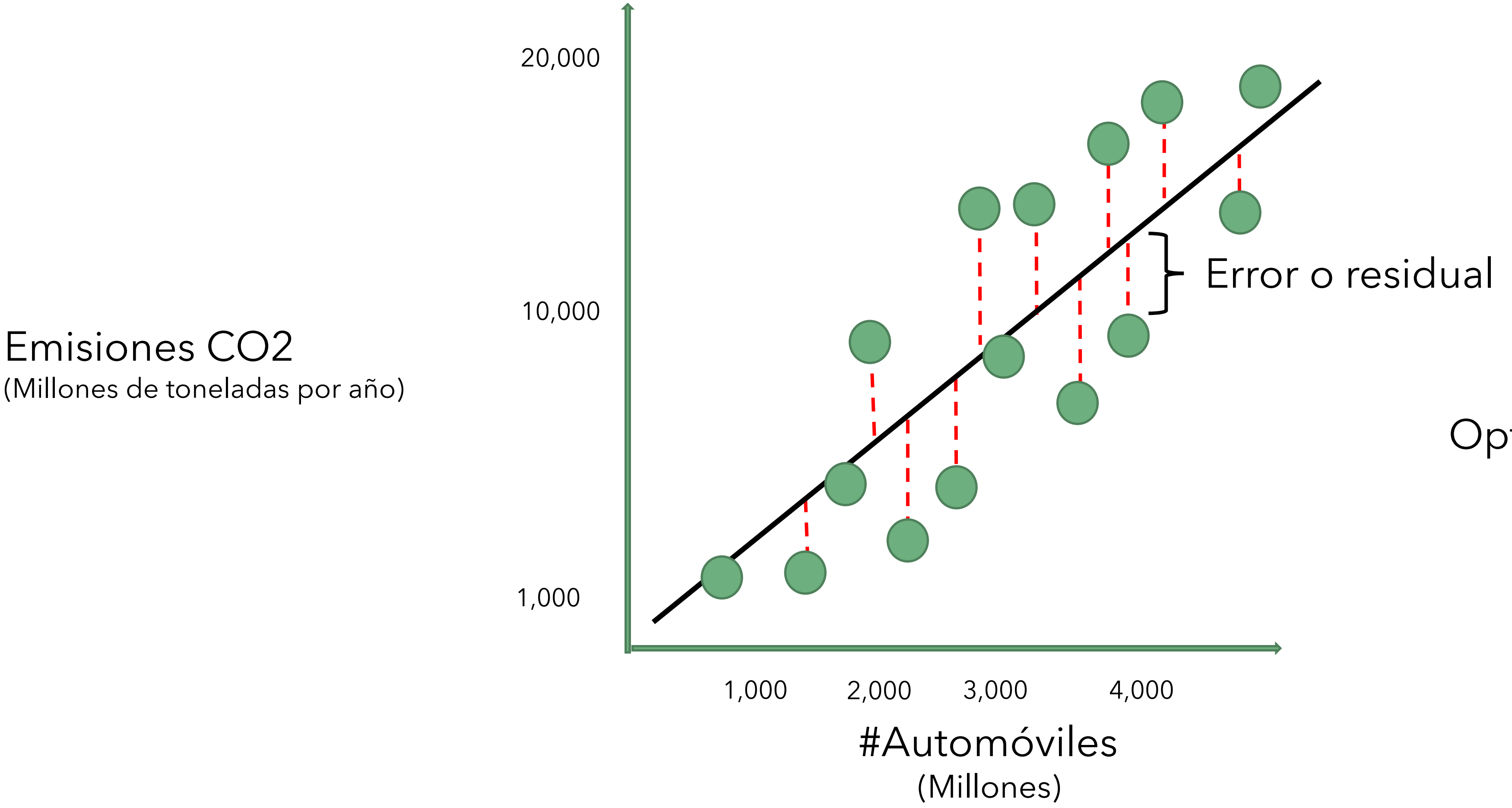
Mativación

Emisiones CO2
(Millones de toneladas por año)



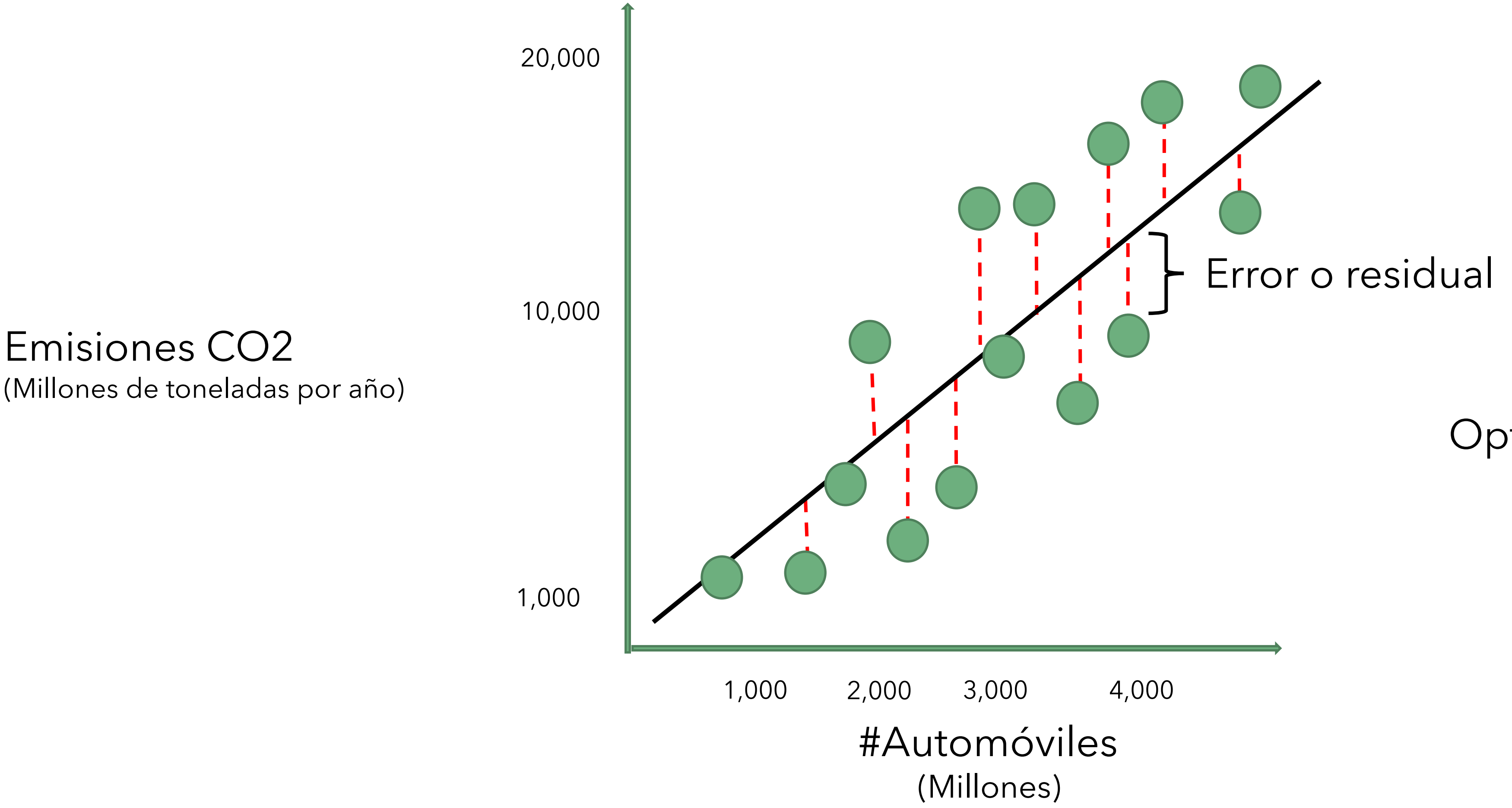
#Automóviles
(Millones)

Mativación



Optimizar

Mativación



Optimizar, disminuir los errores.

Mativación

$$\mathrm{RSS} = e_1^2 + e_2^2 + \cdots + e_n^2$$

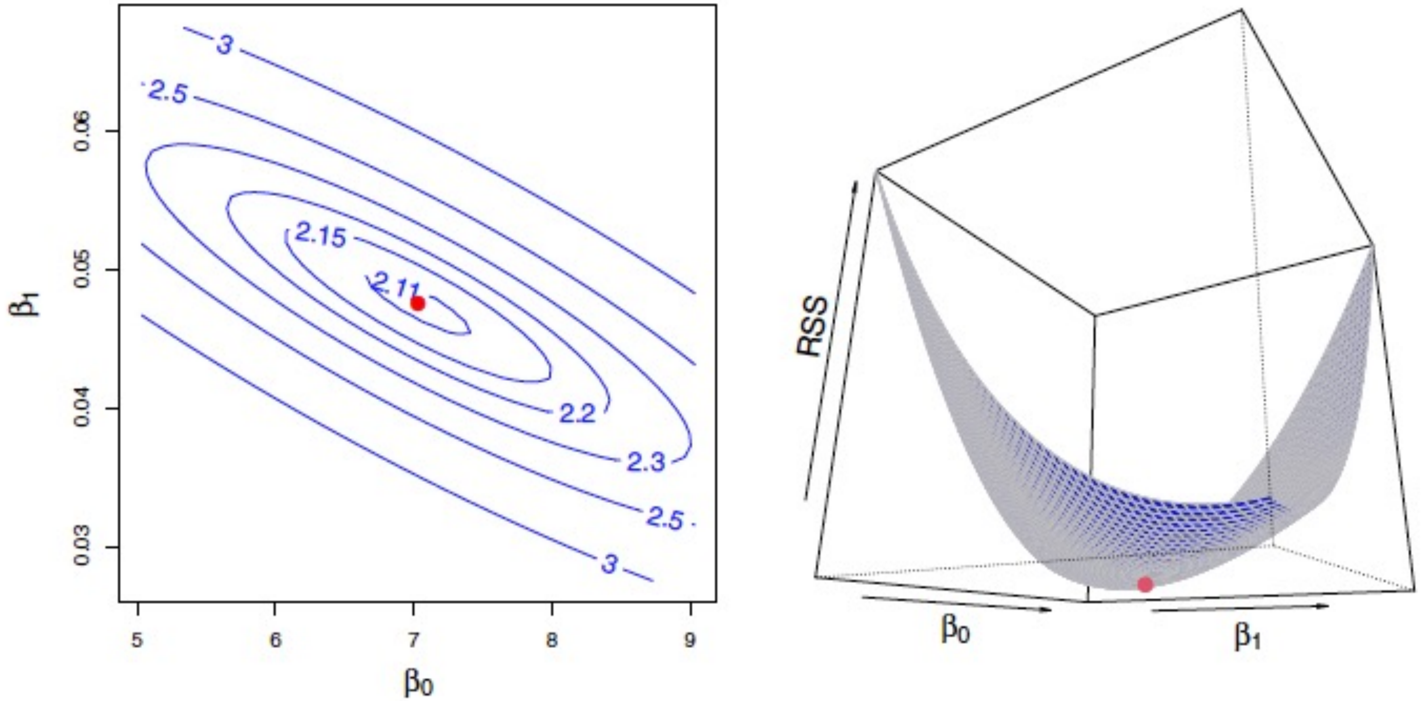
Mativación

$$\mathrm{RSS} = e_1^2 + e_2^2 + \cdots + e_n^2$$

$$\mathrm{RSS} = (y_1 - \hat{\beta}_0 - \hat{\beta}_1 x_1)^2 + (y_2 - \hat{\beta}_0 - \hat{\beta}_1 x_2)^2 + \cdots + (y_n - \hat{\beta}_0 - \hat{\beta}_1 x_n)^2$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

¿Qué es lo que estamos haciendo?



Source: Introduction to Statistical Learning. James et al.

Matricialmente

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{kn} \end{bmatrix}_{n\times k} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_n \end{bmatrix}_{k\times 1} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \vdots \\ \epsilon_n \end{bmatrix}_{n\times 1}$$

Matricialmente

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{kn} \end{bmatrix}_{n\times k} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_n \end{bmatrix}_{k\times 1} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \vdots \\ \epsilon_n \end{bmatrix}_{n\times 1}$$

$$y = X\beta + \epsilon$$

$$e = y - X\hat{\beta}$$

Matricialmente

$$y = X\beta + \epsilon$$

$$e = y - X\hat{\beta}$$

$$\begin{bmatrix} e_1 & e_2 & \dots & \dots & e_n \end{bmatrix}_{1\times n} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} e_1 \times e_1 + e_2 \times e_2 + \dots + e_n \times e_n \end{bmatrix}_{1\times 1}$$

Matricialmente

Sustituyendo

$$\begin{aligned} e'e &= (y - X\hat{\beta})'(y - X\hat{\beta}) \\ &= y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\ &= y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta} \end{aligned}$$

Maximizando

$$\frac{\partial e'e}{\partial \hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

Matricialmente

Sustituyendo

$$\begin{aligned} e'e &= (y - X\hat{\beta})'(y - X\hat{\beta}) \\ &= y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\ &= y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta} \end{aligned}$$

Maximizando

$$\frac{\partial e'e}{\partial \hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

Despejando

$$\hat{\beta} = (X'X)^{-1}X'y$$


Vamos a Python!