

Análisis y Regresión Lineal – 2025

Lara Cellini, Bernabe Moro,
Franco Dalla Gasperina, Joaquín Gabriel Sanchez

11 de octubre de 2025

1. Regresión Lineal Simple

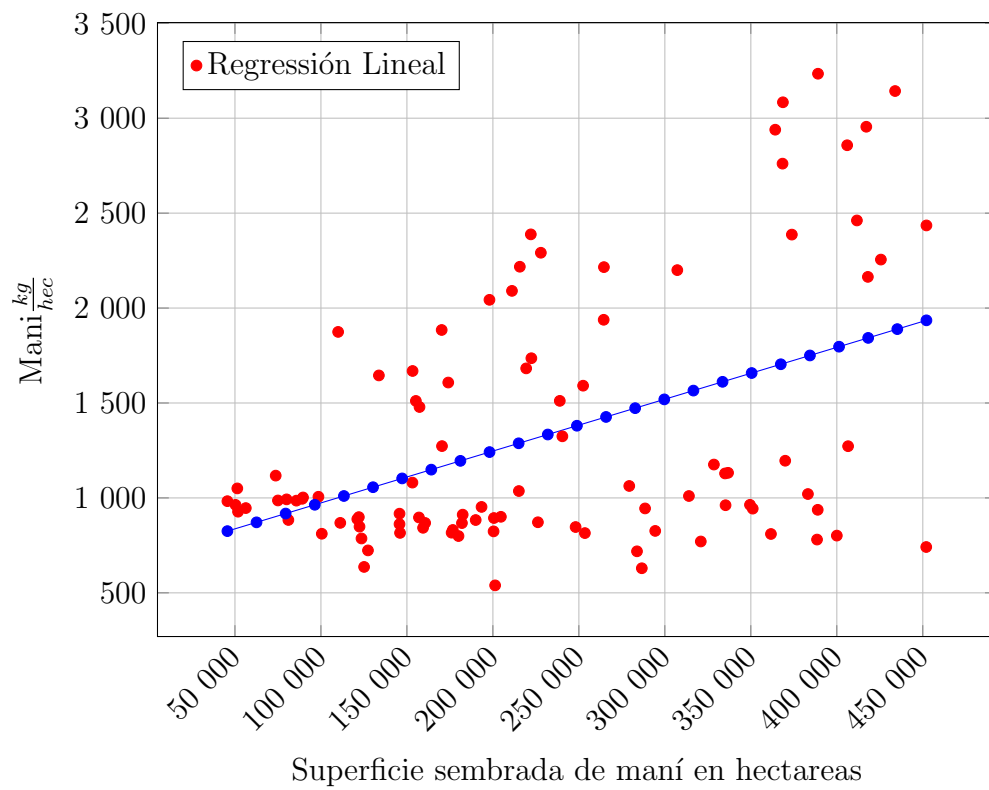
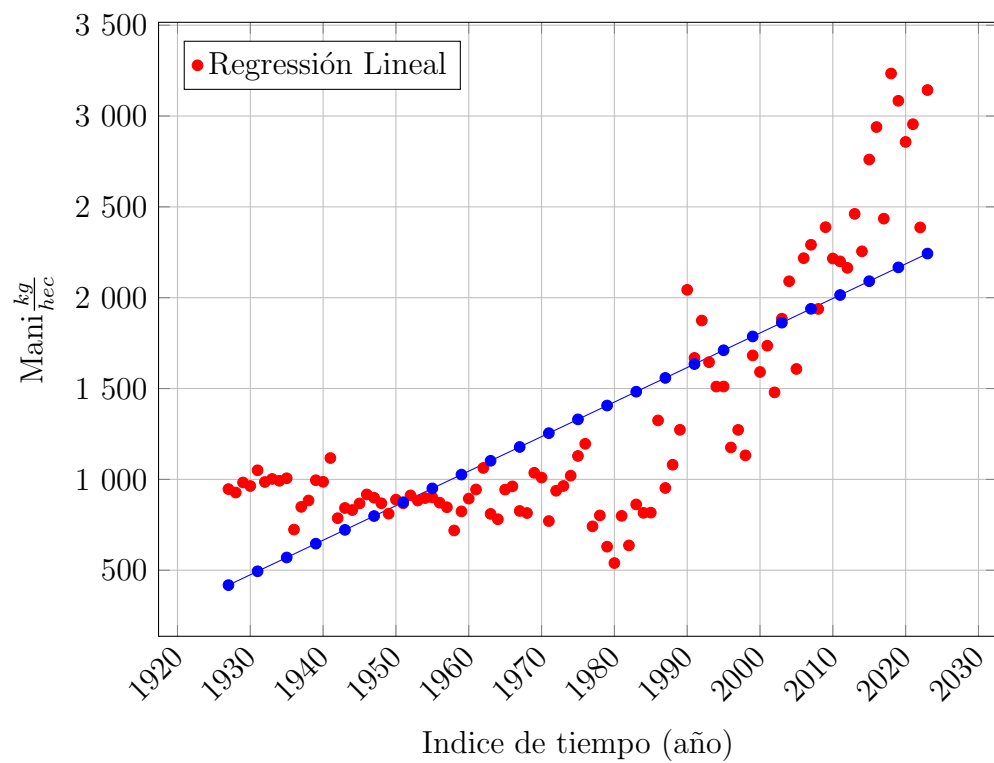
- a) Se define como la variable respuesta y , representante del rendimiento de kilos por hectárea de maní producido, el cual en nuestro conjunto de datos medidos posee un intervalo que va desde $539,43 \frac{kg}{hec}$ a $3234,21 \frac{kg}{hec}$. Además, se definen como variables independientes: la variable x_1 que representa el índice de tiempo, la cual va del año 1927 al año 2023; y la variable x_2 que representa la superficie sembrada en hectareas; y va de 45.606 a 452.118 hectáreas.

El razonamiento detrás de la elección de la variable independiente “Índice de tiempo” es el de visualizar la relación entre los avances tecnológicos de producción en el rendimiento de las cosechas. Por otro lado, el motivo por el cual se puede elegir la “Superficie sembrada en hectáreas” es que tiene una mayor correlación con la variable dependiente ya que esta es uno de los factores que la componen; además, es notable que la relación entre la superficie sembrada y el rendimiento de producción no forman una relación 1 a 1, esto a juzgar por como el valor máximo de hectáreas sembradas es diez veces mayor a el mínimo observado, mientras que el rendimiento máximo medido es 6 veces mayor que el mínimo.

Finalmente, la variable independiente fue seleccionada para este estudio debido a que nos interesa tener la capacidad de predecir el rendimiento de las cosechas anuales en base a las tecnologías utilizadas y/o la disponibilidad de tierras.

b)

Y	$\hat{y}_i = \hat{\beta}_1 x_i + \hat{\beta}_0$	$\hat{\sigma}^2$
x_1	$\hat{y}_i = 19,010073638x_i - 36214,433785565$	677384,754900883
	R^2	r
	0,618736208	0,786597869
	$IC(\beta_1)$	$IC(\beta_0)$
	[13,085213851; 24,934933425]	[-53073,09130; -19355,776266]
	$ICM(Y)$	$IP(y)$
	$-36214,433785565 + 19,010073638 - 1,9852$ $\sqrt{677384,754900883(\frac{1}{97} + \frac{(x^* - 1975)^2}{76048})};$ $-36214,433785565 + 19,010073638 + 1,9852$ $\sqrt{677384,754900883(\frac{1}{97} + \frac{(x^* - 1975)^2}{76048})}$	$\hat{Y}^* - 1,9852$ $\sqrt{677384,754900883(1 + \frac{1}{97} + \frac{(x^* - 1975)^2}{76048})};$ $\hat{Y}^* + 1,9852$ $\sqrt{677384,754900883(1 + \frac{1}{97} + \frac{(x^* - 1975)^2}{76048})}$
Y	$\hat{y}_i = \hat{\beta}_1 x_i + \hat{\beta}_0$	$\hat{\sigma}^2$
x_2	$\hat{y}_i = 0,002731115x_i + 700,57505221$	1405735,9731028215
	R^2	r
	0,208786109	0,456931186
	$IC(\beta_1)$	$IC(\beta_0)$
	[-21,109573105; 21,115035335]	[158,233779405; 1242,916325015]
	$ICM(Y)$	$IP(y)$
	$700,57505221 + 0,002731115 - 1,9852$ $\sqrt{1405735,9731028215(\frac{1}{97} + \frac{(x^* - 230633,494845361)^2}{1243287228824,25})};$ $700,57505221 + 0,002731115 + 1,9852$ $\sqrt{1405735,9731028215(\frac{1}{97} + \frac{(x^* - 230633,494845361)^2}{1243287228824,25})}$	$\hat{Y}^* - 1,9852$ $\sqrt{1405735,9731028215(1 + \frac{1}{97} + \frac{(x^* - 230633,494845361)^2}{1243287228824,25})};$ $\hat{Y}^* + 1,9852$ $\sqrt{1405735,9731028215(1 + \frac{1}{97} + \frac{(x^* - 230633,494845361)^2}{1243287228824,25})}$



- c) Como se puede observar en las tablas, el valor de R^2 en el modelo que utiliza como variable independiente la superficie sembrada es mucho menor que el valor de R^2 el modelo que toma en cuenta los años. Mientras más alto es el valor de R^2 , mejor será nuestro modelo de regresión lineal simple para explicar la variación del rendimiento. Esto significa que es factible negar el hecho de que la variación de la superficie sembrada con maní explique el aumento en el rendimiento de las cosechas.

Por otro lado, r indica la dirección de la relación entre las dos variables. Al ser r un valor dentro del intervalo de 0 a 1, se puede suponer que existe una correlación positiva fuerte entre el pasar de los años y el aumento del rendimiento de las cosechas. Al aumentar x (los años), aumenta también y (el rendimiento). Esto es visible en la gráfica que muestra el modelo con variable independiente x_1 .

En cambio, al considerar el r que corresponde al modelo que toma en cuenta las hectáreas sembradas como variable predictiva, se observa que la relación también sería positiva, mas no tan fuerte. Por lo tanto, es posible suponer que no está tan fuertemente relacionado el aumento en la cantidad de hectáreas sembradas con el aumento en el rendimiento de las cosechas, aunque tampoco es posible afirmar que la correlación sea nula.

$$\text{Con } x_1 : r = \sqrt{R^2} = \sqrt{0,618736208} = 0,786597869.$$

$$\text{Con } x_2 : r = \sqrt{R^2} = \sqrt{0,208786109} = 0,456931186.$$

2. Regresión Lineal Múltiple

- a)
- b)
- c)