

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

#Cargar datos
data = pd.read_csv('dataset_RegresionLineal.csv')
data.head()
```

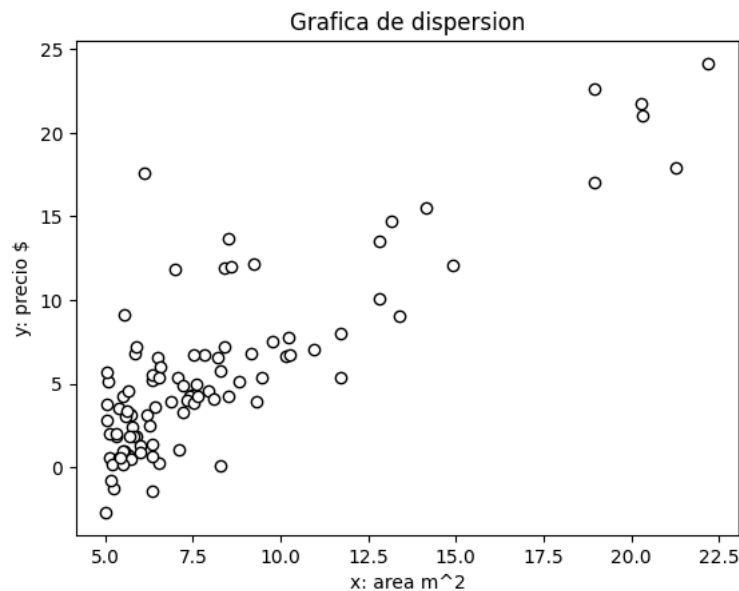
	x	y
0	6.1101	17.5920
1	5.5277	9.1302
2	8.5186	13.6620
3	7.0032	11.8540
4	5.8598	6.8233

Next steps:

[Generate code with data](#)[New interactive sheet](#)

```
x = np.array(data['x'])
y = np.array(data['y'])
m = np.size(x)

#Graficar datos
def graficarDatos():
    plt.plot(x, y, 'o', color='w', mec='black')
    plt.title('Grafica de dispersion')
    plt.xlabel('x: area m^2')
    plt.ylabel('y: precio $')
graficarDatos()
plt.show()
```



```
#ESTIMACION PARAMETRICA
x_media = np.mean(x)
y_media = np.mean(y)

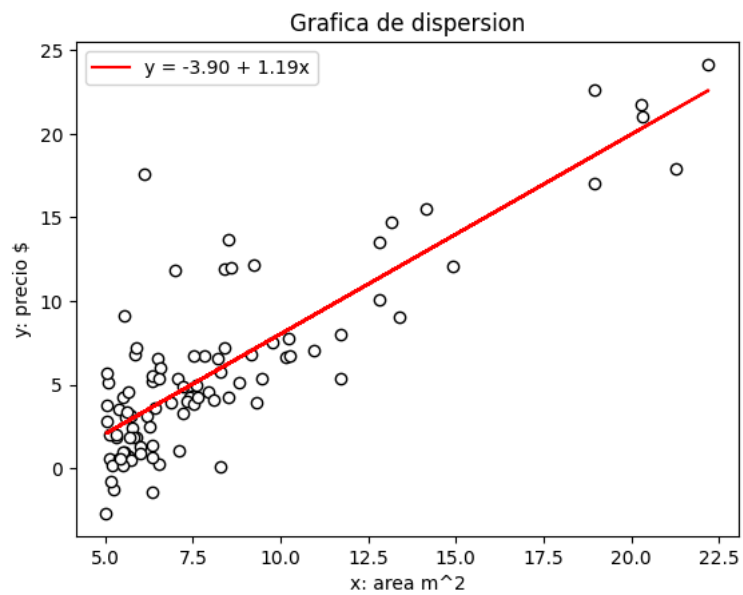
# Calcular  $\beta_1$  (a1) usando:  $\beta_1 = \text{cov}(x,y) / \text{var}(x)$ 
covarianza_xy = np.mean((x - x_media) * (y - y_media))
varianza_x = np.mean((x - x_media) ** 2)
a1 = covarianza_xy / varianza_x

# Calcular  $\beta_0$  (a0) usando:  $\beta_0 = \bar{y} - \beta_1 \cdot \bar{x}$ 
a0 = y_media - a1 * x_media

# Calcular hipótesis final
h = a0 + a1 * x
```

```
# Calcular error (costo)
J = (1 / (2 * m)) * np.sum(np.square(h - y))
```

```
#Graficar datos finales o modelo final
graficarDatos()
plt.plot(x, h, color='r', label=f'y = {a0:.2f} + {a1:.2f}x')
plt.legend()
plt.show()
```



```
# Resultados
print('Error final: ', J)
print('a0: ', a0)
```

```
Error final: 4.476971375975178
a0: -3.8957808783118537
```