

Nombre de la práctica	Ejercicio De Regresión Lineal			No.	4
Asignatura:	Simulación	Carrera:	Ingeniería sistemas	en	Duración de la práctica (Hrs)

Nombre del alumno: Jesús Navarrete Martínez

Grupo: 3501

II. Lugar de realización de la práctica (laboratorio, taller, aula u otro): Aula

III. Material empleado:

- Equipo de computo

IV. Desarrollo de la práctica:

Regresión Lineal: Costo de un incidente de seguridad.

En este ejercicio se explican los fundamentos básicos de la regresión lineal aplicada a un caso de uso sencillo relacionado con la ciberseguridad.

Enunciado del ejercicio

El ejercicio consiste en predecir el costo de un incidente de seguridad en base al número de equipos que se han visto afectados. El conjunto de datos es generado de manera aleatoria

1.- Generación del DataSet

```
import numpy as np

X = 2 * np.random.rand(100,1)
y = 4 + 3 * X + np.random.randn(100,1)

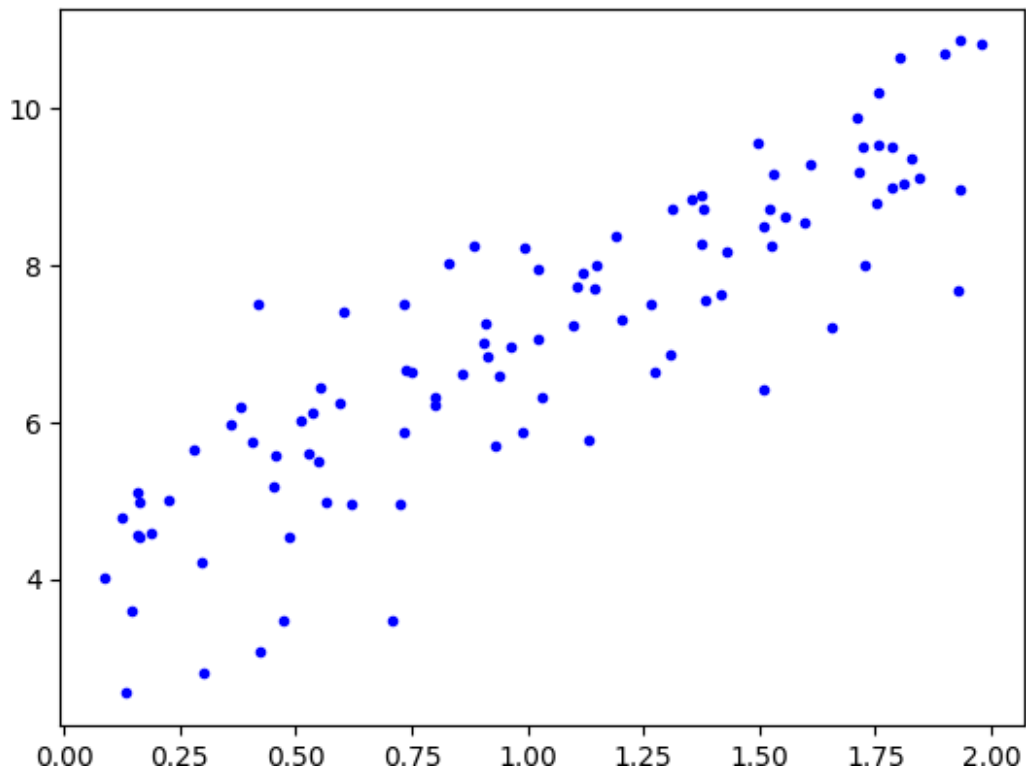
print("La longitud del DataSet es: ", len(X))

La longitud del DataSet es: 100
```

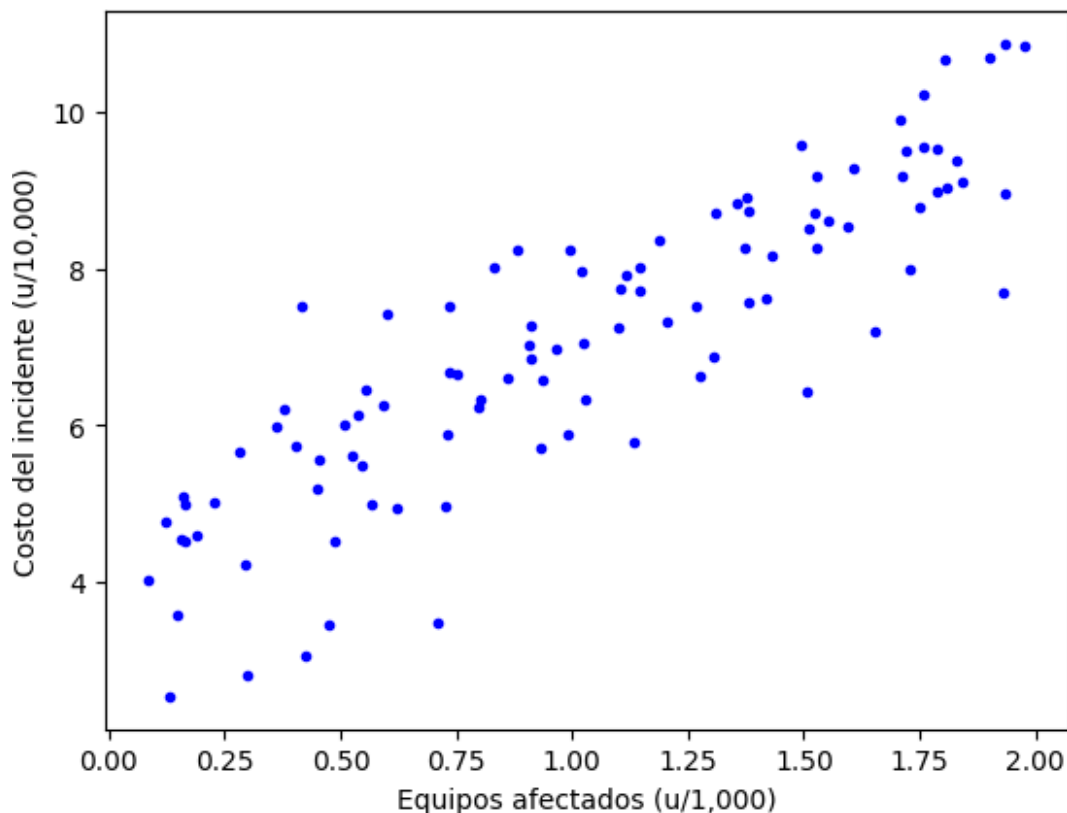
2.- Visualización del DataSet

```
import matplotlib.pyplot as plt
%matplotlib inline

plt.plot(X, y, "b.")
plt.show()
```



```
plt.plot(X,y,"b.")  
plt.xlabel("Equipos afectados (u/1,000)")  
plt.ylabel("Costo del incidente (u/10,000)")  
plt.show()
```



3.- Modificación del DataSet

```
import pandas as pd

data = {'No_Equipos_Afectados': X.flatten(), 'Costo': y.flatten()}
df= pd.DataFrame(data)
df.head(10)
```

	No_Equipos_Afectados	Costo
0	0.733469	7.514180
1	1.188510	8.371345
2	0.156090	4.553944
3	1.526334	8.257509
4	0.086367	4.021777
5	0.158250	5.092942
6	0.749744	6.640276
7	0.725684	4.962020
8	0.593276	6.246454
9	0.163875	4.991766

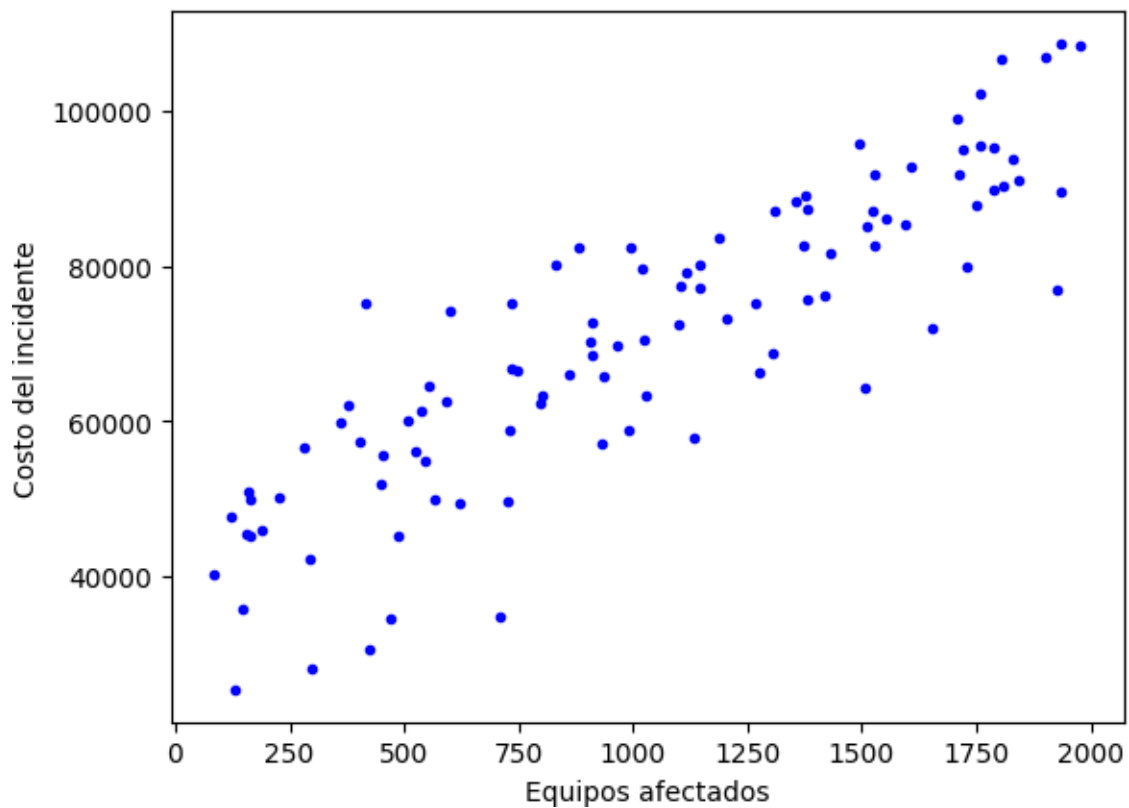
```
# Escalado del numero de equipos afectados
df['No_Equipos_Afectados'] =df['No_Equipos_Afectados'] * 1000
df['No_Equipos_Afectados']= df['No_Equipos_Afectados'].astype('int')
```



```
# Escalado del Costo
df['Costo'] = df['Costo']*10000
df['Costo'] = df['Costo'].astype('int')
df.head(10)
```

	No_Equipos_Afectados	Costo
0	733	75141
1	1188	83713
2	156	45539
3	1526	82575
4	86	40217
5	158	50929
6	749	66402
7	725	49620
8	593	62464
9	163	49917

```
plt.plot(df['No_Equipos_Afectados'],df['Costo'], "b.")
plt.xlabel("Equipos afectados")
plt.ylabel("Costo del incidente")
plt.show()
```



4.- Construcción del modelo

```
from sklearn.linear_model import LinearRegression

# Construcción del modelo y ajuste de la función de hipótesis
lin_reg = LinearRegression()
lin_reg.fit(df['No_Equipos_Afectados'].values.reshape(-1,1),
df['Costo'].values)

LinearRegression()

# Parámetro Tetha 0
lin_reg.intercept_

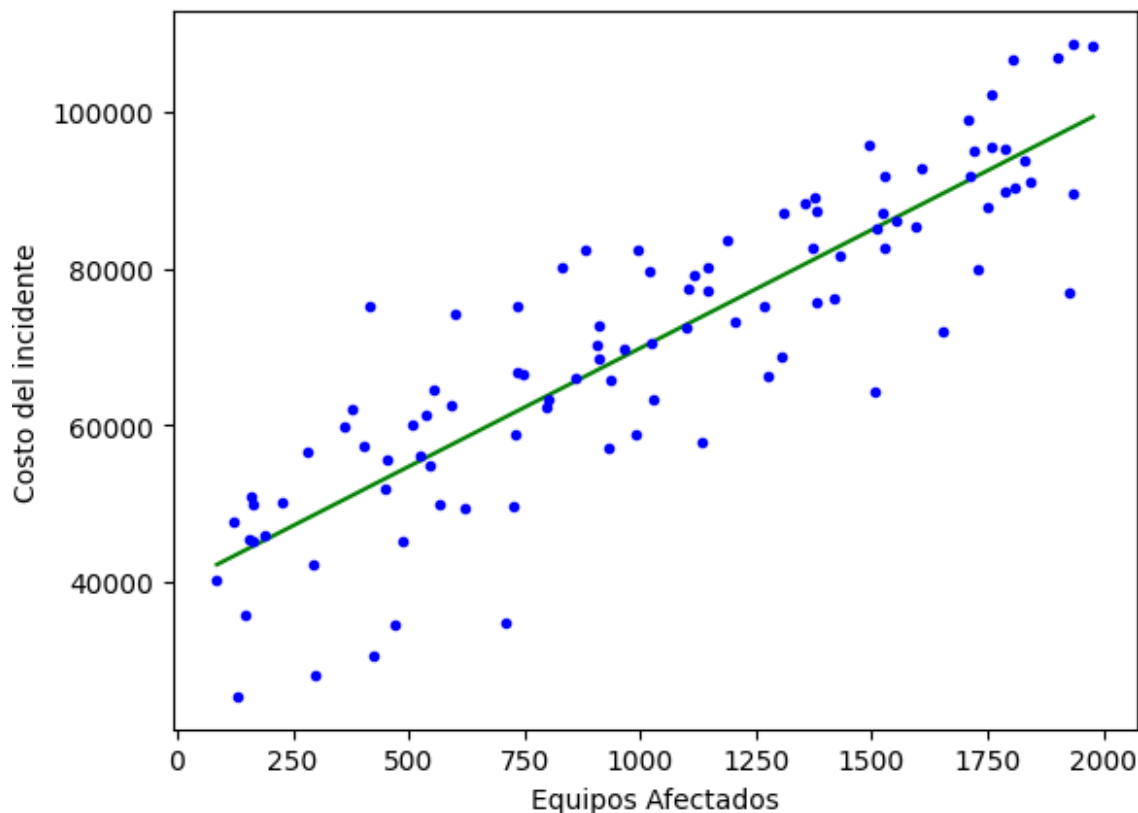
np.float64(39605.94898338528)

# Parámetro Tetha 1
lin_reg.coef_

array([30.22301543])

# Predicción para el valor mínimo y máximo para el conjunto de datos
de entrenamiento
X_min_max = np.array([[df["No_Equipos_Afectados"].min()],
[ df["No_Equipos_Afectados"].max() ]])
y_train_pred = lin_reg.predict(X_min_max)

# Representacion grafica de la funcion de hipotesis generada.
plt.plot(X_min_max, y_train_pred, "g-")
plt.plot(df['No_Equipos_Afectados'], df['Costo'], "b.")
plt.xlabel("Equipos Afectados")
plt.ylabel("Costo del incidente")
plt.show()
```



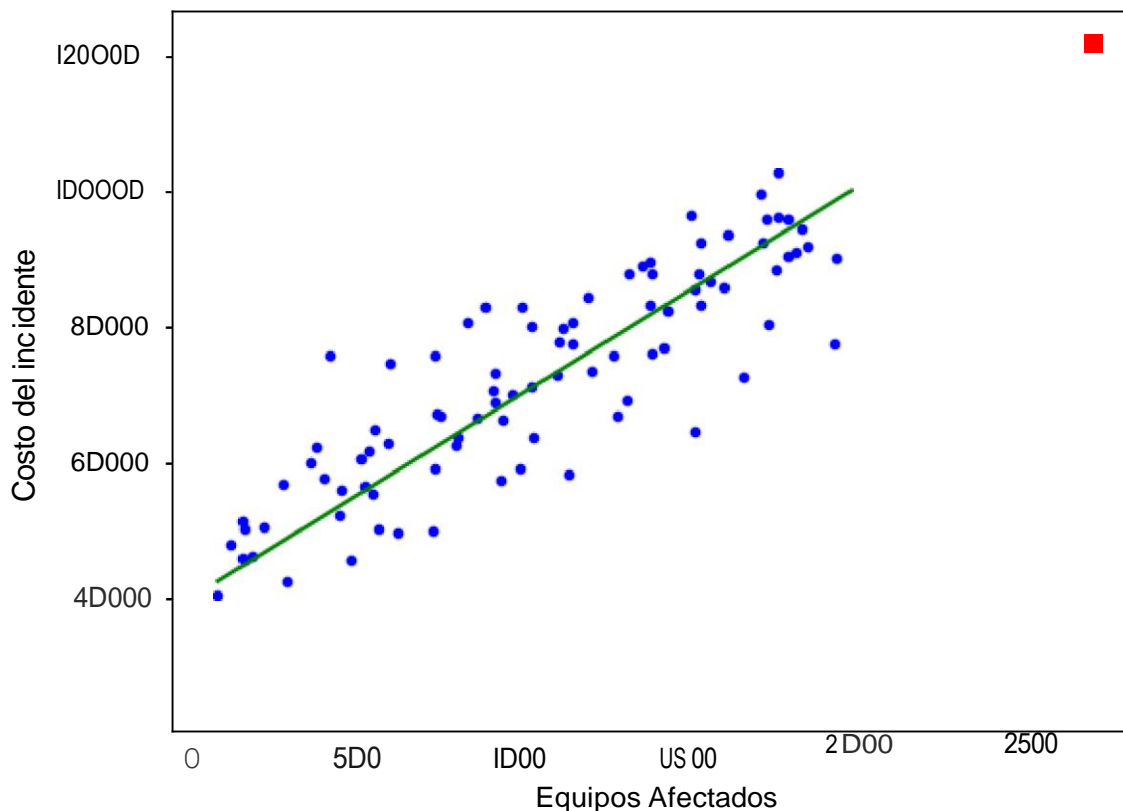
5.- Predicción de nuevos ejemplos

```
x_new = np.array([[2700]]) #Numero de equipos afectados

# Predicción del costo que tendría el incidente
Costo= lin_reg.predict(x_new)
print("El costo del incidente seria: $ ", int(Costo[0]))

El costo del incidente seria: $ 121208

plt.plot(df['No_Equipos_Afectados'], df['Costo'], "b.")
plt.plot(X_min_max, y_train_pred, "g-")
plt.plot(x_new, Costo, "rs")
plt.xlabel("Equipos Afectados")
plt.ylabel("Costo del incidente")
plt.show()
```



V. Conclusiones:

En esta práctica, se aplicaron los conceptos fundamentales de la regresión lineal para predecir el costo de un incidente de seguridad en función del número de equipos afectados. A través del análisis de los datos generados de manera aleatoria, se logró demostrar cómo la regresión lineal permite identificar relaciones lineales entre variables, facilitando la estimación de costos futuros basados en incidentes similares. Este tipo de modelos predictivos son esenciales en el ámbito de la ciberseguridad, ya que pueden ayudar a las organizaciones a anticipar los impactos financieros de futuros incidentes y a optimizar sus estrategias de mitigación. La precisión del modelo puede mejorarse utilizando conjuntos de datos reales y un análisis más profundo de otros factores que podrían influir en los costos, como la severidad del incidente y el tiempo de respuesta.