

## 1. Empleos en lenguajes de programación

### Código:

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
import numpy as np

# Cargar los datos desde un archivo CSV
archivo_csv = 'datos.csv' # Asegúrate de tener el archivo con este nombre
df = pd.read_csv(archivo_csv, encoding='latin1') # Usamos 'latin1' para manejar caracteres especiales

# Mostrar los primeros registros para asegurarnos de que se cargó correctamente
print(df.head())

# Limpiar los nombres de las columnas (eliminar espacios extras)
df.columns = df.columns.str.strip()

# Convertir el porcentaje a formato numérico, quitando el signo de '%' y convirtiendo a float
df['Porcentaje_de_Empleos'] = df['Porcentaje_de_Empleos'].replace('%', '', regex=True).astype(float)

# Crear un índice ficticio como variable independiente
df = df.sort_values(by="Porcentaje_de_Empleos", ascending=False).reset_index(drop=True)
df["Ranking"] = np.arange(1, len(df) + 1)

# Separar las variables independiente (Ranking) y dependiente (Porcentaje_de_Empleos)
X = df[["Ranking"]].values # Variable independiente
y = df["Porcentaje_de_Empleos"].values # Variable dependiente

# Crear y ajustar el modelo de regresión lineal
model = LinearRegression()
model.fit(X, y)

# Generar predicciones
y_pred = model.predict(X)

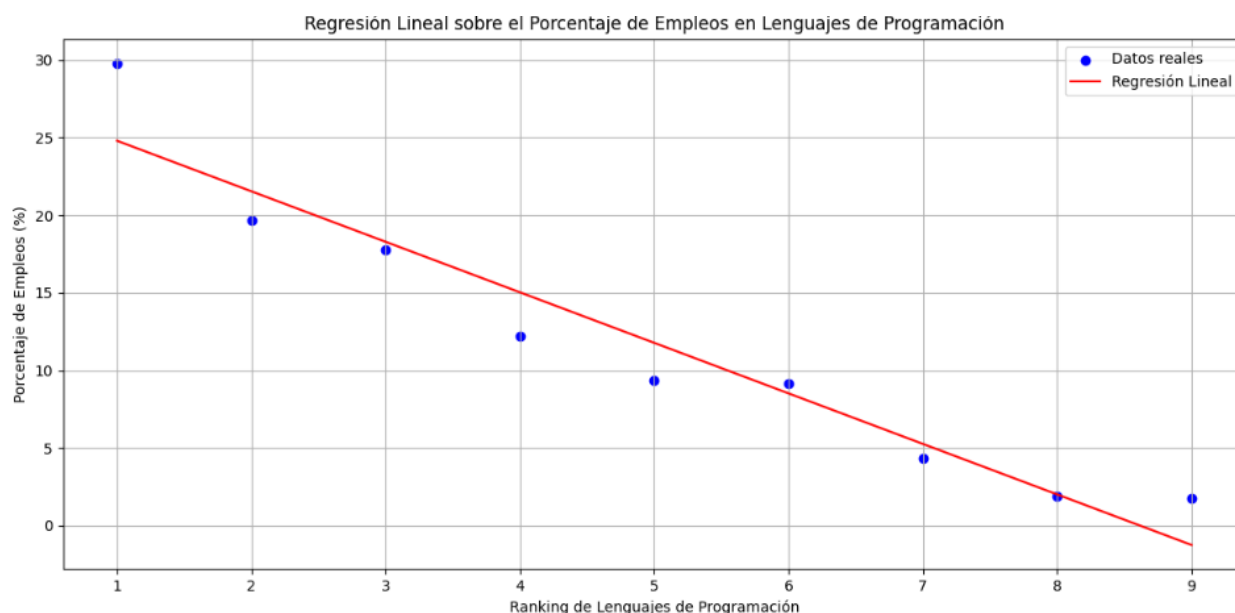
# Mostrar los coeficientes del modelo
print("Pendiente (coeficiente):", model.coef_[0])
print("Intercepto:", model.intercept_)

# Graficar los datos reales y la regresión lineal
plt.figure(figsize=(12, 6))
plt.scatter(df["Ranking"], df["Porcentaje_de_Empleos"], color="blue", label="Datos reales")
plt.plot(df["Ranking"], y_pred, color="red", label="Regresión Lineal")

# Etiquetas y título
plt.xlabel("Ranking de Lenguajes de Programación")
plt.ylabel("Porcentaje de Empleos (%)")
plt.title("Regresión Lineal sobre el Porcentaje de Empleos en Lenguajes de Programación")
plt.legend()
plt.grid(True)
plt.tight_layout()

# Mostrar el gráfico
plt.show()
```

## Grafica:



## Hipótesis:

El porcentaje de empleos asociados a lenguajes de programación disminuye a medida que el lenguaje ocupa una posición más alta en el ranking numérico (es decir, se vuelve menos popular). Los lenguajes más demandados en el mercado laboral tienden a tener un porcentaje significativo de empleos relacionados, mientras que los lenguajes menos populares presentan una menor representación en términos de empleabilidad. La relación entre el ranking y el porcentaje de empleos sigue una tendencia lineal decreciente, aunque es posible que existan ciertas excepciones debido a demandas específicas en nichos de mercado o tecnología emergente.