

Name: Ameya Satish Khond
SJSU ID: 019136845

```
In [15]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

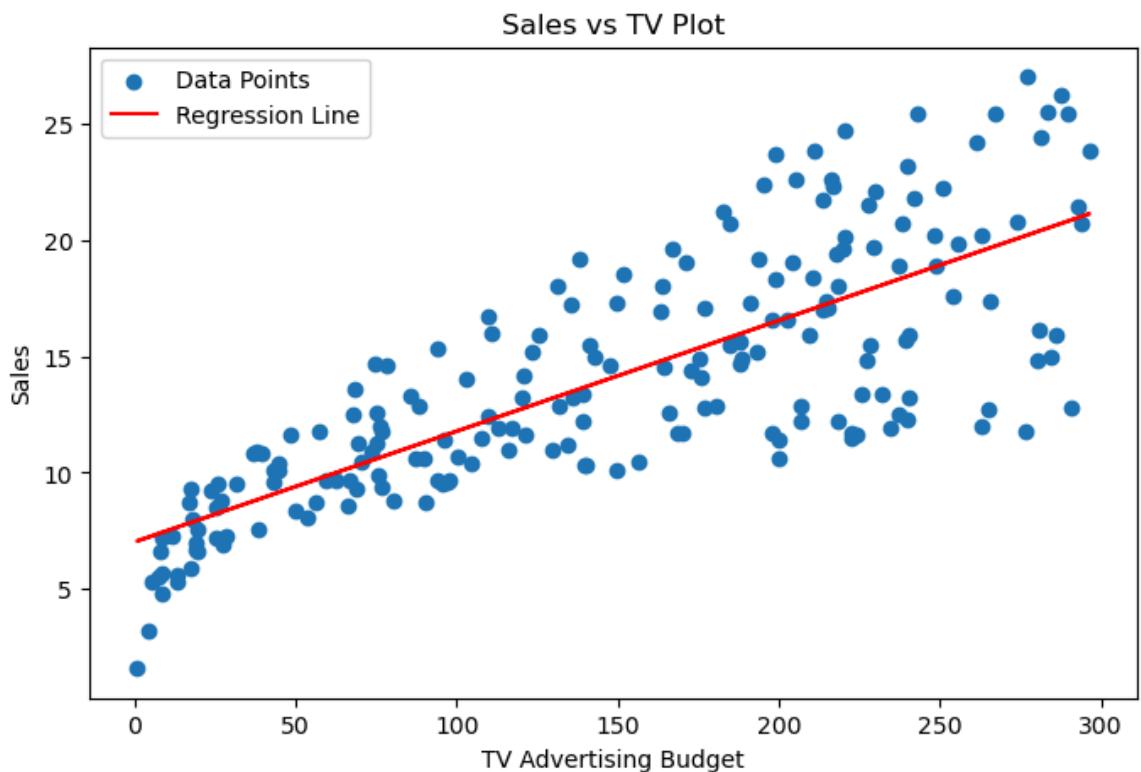
data = pd.read_csv("Advertising.csv")

x = data["TV"]
y = data["Sales"]

coeffs = np.polyfit(x, y, 1)
slope, intercept = coeffs
y_pred = slope * x + intercept
```

```
In [16]: # 1. Create a scatter plot of the data
# 2. Least square regression line

plt.figure(figsize=(8,5))
plt.scatter(x, y, label="Data Points")
plt.plot(x, y_pred, label="Regression Line", color = 'red')
plt.xlabel("TV Advertising Budget")
plt.ylabel("Sales")
plt.title("Sales vs TV Plot")
plt.legend()
plt.show()
```



In [17]: # 3. Compute R-squared and correlation coefficient

```
ss_res = np.sum((y - y_pred)**2)
ss_tot = np.sum((y - y.mean())**2)
r_squared = 1 - (ss_res / ss_tot)
correlation = np.corrcoef(x, y)[0, 1]

print("Slope:", slope)
print("\nIntercept:", intercept)
print("\nCoefficient of Determination (R²):", r_squared)
print("\nCorrelation Coefficient:", correlation)
```

Slope: 0.04753664043301971

Intercept: 7.032593549127703

Coefficient of Determination (R²): 0.611875050850071

Correlation Coefficient: 0.7822244248616066

In []: