Problem Statement

As an owner of a startup, you wish to forecast the sales of your product to plan how much money should be spent on advertisements. This is because the sale of a product is usually proportional to the money spent on advertisements.

Predict the impact of TV advertising on your product sales by performing simple linear regression analysis.

Activity 1: Analyzing the Dataset

Create a Pandas DataFrame for **Advertising-Sales** dataset using the below link. This dataset contains information about the money spent on the TV, radio and newspaper advertisement (in thousand dollars) and their generated sales (in thousand units). The dataset consists of examples that are divided by 1000.

Dataset Link: https://raw.githubusercontent.com/jiss-sngce/CO_3/main/advertising.csv

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
df=pd.read_csv('https://raw.githubusercontent.com/mariya-sherin/DSML/main/advertising%20-%20advertising.csv')
print(df)
x=df['TV']
Y=df['Sales']
```

Output

```
      TV  Radio  Newspaper  Sales
0    230.1   37.8       69.2   22.1
1     44.5   39.3       45.1   10.4
2     17.2   45.9       69.3   12.0
3    151.5   41.3       58.5   16.5
4    180.8   10.8       58.4   17.9
..     ...    ...        ...    ...
195   38.2    3.7       13.8    7.6
196   94.2    4.9        8.1   14.0
197  177.0    9.3        6.4   14.8
```

198 283.6   42.0   66.2   25.5
199 232.1   8.6   8.7   18.4
[200 rows x 4 columns]

Activity 2: Train-Test Split

Activity 3: Model Training

Train the simple regression model using **training data** to obtain the best fit line

y=mx+c

```
x_train,x_test,y_train,y_test=train_test_split(x,Y,test_size=0.2)
x_train=x_train.values.reshape(-1,1)
x_test=x_test.values.reshape(-1,1)
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(x_train.reshape(-1, 1), y_train)
```

Activity 4: Model Prediction

For the TV advertising of $50,000, what is prediction for Sales? In order to predict this value, perform the following task:

- Based on the regression line, create a function sales_predicted() which takes a budget to be used for TV advertising as an input and returns the corresponding units of Sales.
- Call the function sales_predicted() and pass the amount spent on TV advertising.

Note: To predict the sales for TV advertising of $50,000, pass 50 as parameter to sales_predicted() function as the original data of this dataset consists of examples that are divided by 1000. Also, the value obtained after calling sales_predicted(50) must be multiplied by 1000 to obtain the predicted units of sales.

```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(x_train.reshape(-1, 1), y_train)
def sales_predicted(budget):
  return model.predict([[budget]])
print(sales_predicted(50))
slope = model.coef_[0]
intercept = model.intercept_
print("Slope:", slope)
print("Intercept:", intercept)
```

Output

```
[9.70450587]
Slope: 0.056202952396790455
Intercept: 6.894358251320877
```