# Linear Regression Problem Statement

## Title: Predicting Sales Based on Advertising Budget

### Objective:

Develop a predictive model using linear regression to forecast sales based on the advertising budget.

### Context:

In the business domain, understanding the impact of advertising expenditure on sales is critical for optimizing marketing strategies and maximizing return on investment (ROI). By leveraging data on advertising budgets and corresponding sales figures, we can create a model to predict future sales and provide actionable insights for marketing and financial planning.

## Dataset Description:

### 1. Advertising\_Budget:

- Description: The amount of money spent on advertising, measured in thousands of dollars.  
- Type: Numeric  
- Range: 10 to 100 thousand dollars  
- Example Values: 15.25, 45.78, 89.34

### 2. Sales:

- Description: The sales revenue generated, measured in thousands of dollars.  
- Type: Numeric  
- Range: Approximately 20 to 320 thousand dollars  
- Example Values: 68.92, 160.25, 290.13

## Problem Statement:

Develop a linear regression model to predict sales revenue based on the advertising budget. The model should help businesses identify how much investment in advertising is generally required to achieve a certain level of sales, thereby aiding in the development of effective marketing strategies and budget allocations.

## Additional Details:

### 1. Data Collection:

The data was collected from a sample of businesses that reported their weekly advertising budget and corresponding sales figures.

### 2. Data Preprocessing:

- Ensure there are no missing or outlier values in the dataset.  
- Normalize the data if necessary to improve model performance.

### 3. Model Evaluation:

- Split the data into training and testing sets.  
- Use appropriate metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared to evaluate the model's performance.

### 4. Practical Application:

- Use the model to provide personalized budget recommendations to businesses.  
- Implement the model in financial planning tools to help businesses optimize their advertising expenditures.

## Conclusions:

1. Mean Absolute Error (MAE): The average absolute error between the predicted and actual sales.  
2. Mean Squared Error (MSE): The average squared error.  
3. Root Mean Squared Error (RMSE): The typical error magnitude.  
4. R-squared (R²): The proportion of the variance in the sales that is predictable from the advertising budget.

The linear regression model shows a strong correlation between the advertising budget and sales, making it a useful tool for predicting sales revenue based on advertising expenditure. This can help businesses optimize their marketing budgets and improve financial planning.