

This Python program performs a linear regression analysis on an advertising dataset to predict sales based on advertising expenditures in different media channels (TV, Radio, and Newspaper). The program follows these key steps:

**1. Data Loading & Preprocessing:**

- Reads the dataset from a CSV file.
- Displays the first few rows of the data.
- Checks for missing values.

**2. Feature Selection & Data Splitting:**

- Defines independent variables (TV, Radio, and Newspaper) as features (X).
- Sets Sales as the target variable (Y).
- Splits the dataset into training (80%) and testing (20%) subsets.

**3. Model Training & Prediction:**

- Uses the `LinearRegression` model from `sklearn` to train on the training data.
- Predicts sales for the test dataset.

**4. Model Evaluation:**

- Calculates key regression evaluation metrics:
  - **Mean Absolute Error (MAE):** Measures the average absolute difference between actual and predicted values.
  - **Mean Squared Error (MSE):** Computes the average squared differences.
  - **Root Mean Squared Error (RMSE):** Indicates how much error is expected in predictions.
  - **R<sup>2</sup> Score:** Determines how well the model explains the variability in sales.

**5. Visualization:**

- Creates a scatter plot comparing actual vs. predicted sales values to assess model performance visually.

This program provides insights into how advertising budgets across different channels influence sales and demonstrates basic machine learning concepts using Python.