Simple Linear Regression

Problem Statement – Using the advertising dataset analyse the relationship between 'TV advertising' and 'sales' using a simple linear regression model. Work on linear model using two different libraries: **statsmodels** and **SKLearn.**

Explanation – Out of all the independent/predictor variables present in the dataset 'TV' was very strongly correlated to 'Sales'. So built simple linear regression model with 'TV' as the predictor variable.

Building Model steps -

1. Study and clean data
2. Visualizing the Data
    a. Plot scatter plots and heatmaps between predictor and target variables to understand the relationship between them
3. Hypothesis testing in linear regression
    a. To determine the significance of beta coefficients.
4. Building a linear model
    a. Used statsmodels to build LR model
    b. OLS (Ordinary Least Squares) method in statsmodels to fit a line.
    c. Summary statistics
        i. F-statistic, R-squared, coefficients and their p-values.
5. Residual Analysis
    a. Histogram or Q-Q plot of the error terms to check normality.
    b. Plot of the error terms with X or y to check independence.
6. Predictions

Made predictions on the test set using the 'predict()' function.