Startup 50

Startup 50: A Look Inside Overview of the Simple Linear Regression Analysis Model: The link between Marketing Spend (an independent variable) and Profit (a dependent variable) was assessed using a simple linear regression model. Eighty percent of the dataset was used to train the model, while the remaining twenty percent was set aside for testing.

Model parameters: a) Intercept (𝛽0) = 48923.31, meaning that 48923.31 units of profit would be expected if marketing spend was zero. This is the starting profit in the absence of any marketing expenditures.

b) R-Squared = -0.11, indicating that the model is ineffective in explaining the variation in profit. The model performs worse than a straightforward mean-based prediction for the dependent variable, according to the negative value. This could suggest that marketing expenditures alone is not a significant predictor of profit in this dataset.

Multiple Linear Regression Analysis

Model Overview: Multiple Linear Regression is used to predict Profit based on R&D Spend, Administration, Marketing Spend, and State. The dataset was processed, with categorical variables (State) encoded into numerical form to ensure compatibility with the regression model.

Model Parameters

a) Intercept (𝛽0) = 54080.72, indicating that if all independent variables (R&D Spend, Administration, Marketing Spend, and the encoded State variables) are zero, the model predicts a baseline profit of approximately 54080.72 units.

b) R-Squared value = 0.9001, which means that the independent variables collectively explain 90.01% of the variability in Profit. This indicates a strong relationship between the predictors and the target variable.

c) Mean Squared Error = 80929465.4910, which represents the average squared difference between the actual and predicted profits. While this number provides insight into the model's error magnitude, its interpretation depends on the scale of the target variable.