# Multiple Linear Regression

Definition - Multiple Linear Regression is an extension of Simple Linear Regression that models the relationship **between multiple input features (independent variables) and a continuous output (dependent variable)**. It is used in machine learning to predict a target value by considering the combined influence of multiple factors, assuming a linear relationship between the inputs and the target.

## Equation

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$$

- Y is the predicted output (dependent variable)
- X_1, X_2, …, X_n are the input features (independent variables)
- B_0 is the intercept (the value of y when all x's are zero)
- B_1, B_2, …, B_N are the coefficients representing the rate of change of y with respect to each input feature x_n

Feature Scaling - If the algorithm uses OLS method for optimization, then feature scaling generally not required because OLS can handle different feature magnitudes.

OLS computes the coefficients directly through matrix operations and does not rely on iterative optimization like Gradient Descent. Therefore, the algorithm is less sensitive to the scale of features.

## How It Works

- Apply Feature Scaling if necessary.
- Using the raining dataset, the algorithm estimates the coefficients using methods like OLS or Gradient Descent. The algorithm finds the optimal line (or hyperplane in the case of multiple dimensions) that minimizes the sum of squared residuals (the differences between predicted and actual values).
- Once the model is trained, it can be used to predict the dependent variable y for new input data by plugging the input features x_1, x_2, …, x_n into the equation.

Example - Suppose you want to predict house prices (Y) based on square footage (X_1), number of bedrooms (X_2), and age of the house (X_3).

The equation of the model would look like:

Price = B_0 + B_1.(Square Footage) + B_2.(Bedrooms) + B_3(Age)

The algorithm of the model learns the coefficients B_1, B_2, B_3 from the training data and these coefficients represent the impact of each feature on the house price.

Then the model can be used to predict house prices for new or unseen input data. The model uses the equation to compute predictions.