# 📘 Simple Linear Regression – Complete Notes

**🔹 Introduction to Regression**

Regression is a supervised learning technique used to model the relationship between input features (independent variables) and a continuous output variable (dependent variable).

**What is Regression?**

- It’s a predictive modeling technique.

- Used when the target variable is continuous (e.g., house price, temperature, salary).

- Unlike classification, which predicts categories, regression predicts quantitative values.

**Types of Regression**

1. Simple Linear Regression: Only one independent variable (X) and one dependent variable (Y). Equation of a line is used to model the relationship.

2. Multiple Linear Regression: Involves two or more independent variables (X₁, X₂, ..., Xₙ). The model becomes:

y = b₀ + b₁x₁ + b₂x₂ + ... + bₙxₙ

**🔹 Simple Linear Regression (SLR)**

Simple Linear Regression finds the best-fitting straight line to describe the relationship between a single input feature (X) and a target variable (Y).

Equation: ŷ = mx + b

- ŷ is the predicted value

- m is the slope

- b is the intercept

Objective: Minimize the Mean Squared Error (MSE)

MSE = (1/n) × Σ(yᵢ - ŷᵢ)²

**🔹 Code Example (Implementation using Python)**

Step-by-step implementation using NumPy and Matplotlib:

import numpy as np

import matplotlib.pyplot as plt

X = np.array([1, 2, 3, 4, 5])

Y = np.array([2, 4, 5, 4, 5])

def predict(x, m, b):

return m \* x + b

mean\_x = np.mean(X)

mean\_y = np.mean(Y)

m = np.sum((X - mean\_x) \* (Y - mean\_y)) / np.sum((X - mean\_x) \*\* 2)

b = mean\_y - m \* mean\_x

Y\_pred = predict(X, m, b)

plt.scatter(X, Y)

plt.plot(X, Y\_pred, color='red')

plt.show()

**🔹 Intuition Behind Simple Linear Regression**

- Fit a line such that the total vertical distance (residuals) from each point to the line is minimized.

- The assumption is a linear relationship between X and Y.

**🔹 How to Find Slope (m) and Intercept (b)**

Slope (m) formula:

m = Σ(xᵢ - x̄)(yᵢ - ȳ) / Σ(xᵢ - x̄)²

Intercept (b) formula:

b = ȳ - m × x̄

These values are calculated using the Least Squares method, which minimizes the Mean Squared Error.

**🔍 Alternate Explanation**

- Center data using mean of x and y.

- Measure how y changes with x (slope).

- Adjust the vertical position (intercept).

- This is the basis of Ordinary Least Squares (OLS).

**✅ Summary**

- Simple Linear Regression predicts a continuous output using a single input.

- Model: ŷ = mx + b

- Objective: Minimize Mean Squared Error (MSE)

- Parameters are calculated using the Least Squares method.