**🔢 What is Linear Regression?**

**Linear Regression** is a **supervised machine learning algorithm** used to **predict a continuous dependent variable (Y)** based on **one or more independent variables (X)**.

When there's only **one independent variable**, it’s called **Simple Linear Regression** (what you're doing).

**🎯 Goal of Linear Regression**

To **find the best-fitting straight line** (called the regression line) through the data points such that the difference between actual and predicted values is minimized.

**📈 Linear Regression Equation**

Y=mX+bY = mX + bY=mX+b

Where:

* YYY: Predicted value (e.g., per capita income)
* XXX: Input feature (e.g., year)
* mmm: Slope of the line (how much Y changes when X increases)
* bbb: Intercept (value of Y when X = 0)

**🧠 How the Model Learns**

The algorithm tries to **learn the values of m and b** that make the predicted Y values as **close as possible to the actual Y values**.

This is done using a method called **Ordinary Least Squares (OLS)**.

**⚙️ Behind the Scenes: Ordinary Least Squares (OLS)**

OLS minimizes the **Sum of Squared Errors (SSE)**:

SSE=∑i=1n(yi−y^i)2\text{SSE} = \sum\_{i=1}^{n} (y\_i - \hat{y}\_i)^2SSE=i=1∑n​(yi​−y^​i​)2

Where:

* yiy\_iyi​: Actual value
* y^i\hat{y}\_iy^​i​: Predicted value by the model
* The square is used to penalize larger errors more

The smaller the SSE, the better the line fits the data.