**Simple Linear Regression Theory**

**Regression:** used to find the relationship between variables.

**Linear Relation:** A linearrelation means that the change in one variable is directly proportional (or almost proportional) to the change in another variable.

**Linear Regression:** Used to find the relationship between 2 variables.

**Equation:** Y=β0 +β1X+ε

**Y:** Dependent variable

**X:** Independent variable

**β1:** Slope of the line (how much y changes for unit changes in x)

**β0:** Intercept (value of y when x is zero)

**ε:** Error Term

**Formula for β1:** sum of product of deviations/ sum of square of deviations for x

**Formula for β2:** Mean of Y - (β1 \* mean of β1)

### **Assumptions of Linear Regression**

* Linear relationship between X and Y
* Residuals (errors) are normally distributed
* Homoscedasticity (equal variance of residuals)
* Independence of observations
* No/mild multicollinearity (Means predictors should be either not correlated or only weakly correlated with each other, not relevant for simple regression, but mention briefly)

### **How the Model Works**

* Goal: Find best fit line that minimizes error.
* **Ordinary Least Squares (OLS):** Ordinary Least Squares (OLS) is a method that finds the best-fit line by minimizing the total squared differences between actual and predicted values.

### **In Machine Learning**

* It’s the **foundation for regression problems**.
* Provides interpretability (easy to explain how X affects Y).
* Forms the base for more advanced algorithms like **Multiple Linear Regression, Logistic Regression, and Neural Networks**.