📌 Simple Linear Regression — Theory + Intuition

1. What is Linear Regression?

- Linear Regression is a **supervised learning algorithm** used to **predict numerical values**.
- It finds a **relationship between two variables** — one **independent (input)** and one **dependent (output)**.
- Example: Predicting **salary (output)** based on **years of experience (input)**.

2. Types of Linear Regression

- **Simple Linear Regression**:
  - One input variable (X) and one output variable (Y).
  - We draw a **straight line** to best fit the data points.
- **Multiple Linear Regression**:
  - More than one input variable (e.g., experience, age, education).
- **Polynomial Regression** (extension):
  - Fits **curved lines** (non-linear relationships) using powers of X.

3. Simple Linear Regression

➢ ***Goal:***

- Find the **best straight line** (y = mx + c) that minimizes the difference between actual and predicted values.

➢ ***Components:***

- **X**: Independent variable (e.g., experience)
- **Y**: Dependent variable (e.g., salary)
- **m**: Slope — tells how much Y changes for each unit of X.
- **c**: Intercept — the value of Y when X = 0.

➢ ***Example:***

If the line is salary = 5000 × experience + 30000,
it means every extra year of experience adds Rs. 5000 to the salary.

4. Applying with Code

(Even though we're not using code here, here's what usually happens conceptually):

- We **feed the data** to the regression model.
- The model **calculates the best-fit line** using a method called **least squares** (minimizes the squared error between predicted and real values).

- Once trained, it can **predict Y** for any new X.

---

5. Visual & Intuitive Understanding

Intuition:

- Imagine plotting points on a graph — for example, years of experience vs salary.
- You draw a **line that comes as close as possible** to all the points.
- Some points are above the line, some below — but the line represents the **overall trend**.

The goal of the model:

- **Minimize the error** (difference between actual and predicted values).
- This error is called **residual**.
- Model tries to **adjust slope and intercept** so that the **total squared error** is as small as possible.

➢ Summary

Concept	Meaning
y = mx + c	Equation of a straight line
m (slope)	How much Y increases for 1 unit of X
c (intercept)	Value of Y when X is 0
Goal	Find best-fit line that minimizes error
Use Case	Predicting outcomes based on historical trends
