# Project on Linear Regression

## Overview

Linear regression is a statistical method used to model the relationship between two variables by fitting a linear equation to observed data. It is a type of supervised learning algorithm primarily used for regression tasks, where the goal is to predict the value of a continuous dependent variable based on one or more independent variables.  
  
The most basic form, simple linear regression, involves a single independent variable and is expressed by the equation:  
y = mx + c,  
where:  
- y = dependent variable (Salary)  
- m = slope (coefficient)  
- x = independent variable (Years of Experience)  
- c = constant (intercept)

## Salary Analysis Using Linear Regression

### Problem Statement

The objective of this project is to predict employee salaries based on their years of experience. This analysis supports more informed decision-making regarding salary planning, negotiations, and budgeting within organizations.

### Data Cleaning

The dataset had no missing or null values, ensuring clean and reliable data for analysis.

### Model Building

Using Python’s scikit-learn library, a simple linear regression model was built with:  
- Independent variable: YearsExperience  
- Dependent variable: Salary

### Model Evaluation

The model was trained and tested using an 80:20 data split. Below are the key performance metrics:  
- Root Mean Squared Error (RMSE): 59.1157  
- Slope (Coefficient): 9460.66  
- Constant (Intercept): 25521.23  
- R² Score: 0.9219  
  
These results indicate the model has a strong predictive capability with a high degree of accuracy and minimal error.