# Overview:

**Linear Regression**

Types of Linear Regression

How Linear Regression Works

Evaluation Metrics

Assumptions of Linear Regression

# Linear Regression

Linear regression allows us to make predictions about the dependent variable based on the values of the independent variables

Predict continuous numerical value

The linear regression equation is typically represented

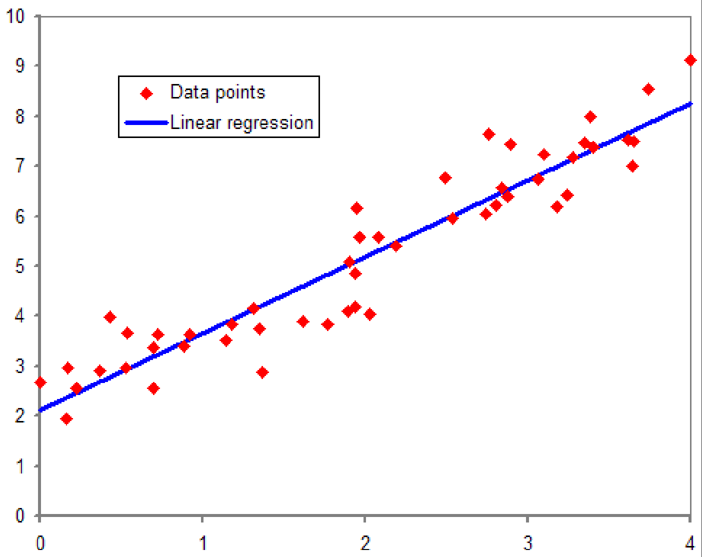
$$y=mx+b$$

y is the predicted value (dependent variable)

m is the slope of the line

x is the input features(independent variable)

b is the y-intercept



# Type of Linear Regression

## Simple Linear Regression:

Simple Linear Regression helps to find the linear relationship between One independent and one dependent feature.

y=mx+b

| YearsExperience | Salary |
|---|---|
| 1.1 | 35342 |
| 1.3 | 46200 |
| 1.5 | 38731 |
| 2 | 44526 |
| 2.2 | 39895 |
| 2.9 | 55645 |
| 3 | 60180 |
| 3.2 | 54445 |
| 3.2 | 64446 |
| 3.7 | 57175 |
| 3.9 | 63218 |
| 4 | 55794 |
| 4 | 57967 |
| 4.1 | 58080 |

## Multi Linear Regression:

In multi linear regression we have two or more than two input variable

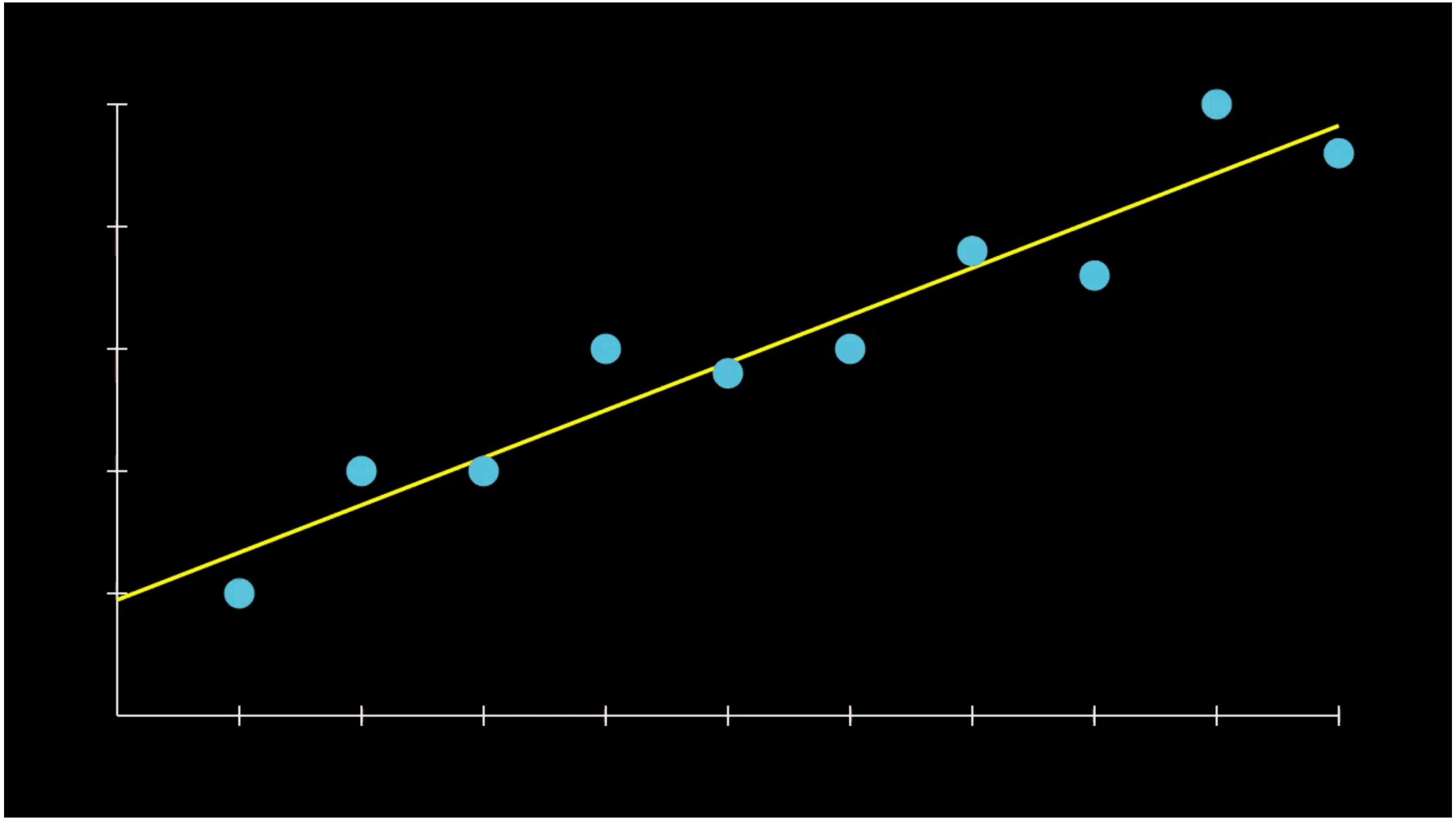
y=m1x1+m2x2+m3x3+⋯+mnxn+b

| y | $X_1$ | $X_2$ |
|---|---|---|
| 140 | 60 | 22 |
| 155 | 62 | 25 |
| 159 | 67 | 24 |
| 179 | 70 | 20 |
| 192 | 71 | 15 |
| 200 | 72 | 14 |
| 212 | 75 | 14 |
| 215 | 78 | 11 |

# How Linear Regression Works:

- Linear regression is a supervised machine learning method.
- That is used to finds a linear equation that best describes the correlation of the independent variables with the dependent variable.
- This is achieved by fitting a line to the data using least squares.
- The line tries to minimize the sum of the squares of the residuals.
- The residual is the distance between the line and the actual value of the independent variable.
- Finding the line of best fit is an iterative process
- we use mostly Mean Squared Error which is the average of the squared difference between actual value and predicted value

$$MSE = \frac{1}{N} \sum_{i=1}^{n} (y_i - (mx_i + b))^2$$



# Gradient Descent:

- Gradient Descent is an optimization algorithm used to minimize the cost function (like MSE) by updating model parameters iteratively.

## Gradient descent

**One feature**

repeat {

$$\underbrace{w} = w - \alpha \underbrace{\frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}\left(x^{(i)}\right) - y^{(i)}\right)x^{(i)}}_{\frac{\partial}{\partial w}J(w,b)}$$

$$b = b - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}\left(x^{(i)}\right) - y^{(i)}\right)$$

simultaneously update $w$, $b$

}

**$n$ features ($n \geq 2$)**

repeat {

$j=1$

$$\underbrace{w_1} = w_1 - \alpha \underbrace{\frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right) - y^{(i)}\right)x_1^{(i)}}_{\frac{\partial}{\partial w_1}J(\vec{w},b)}$$

$\vdots$

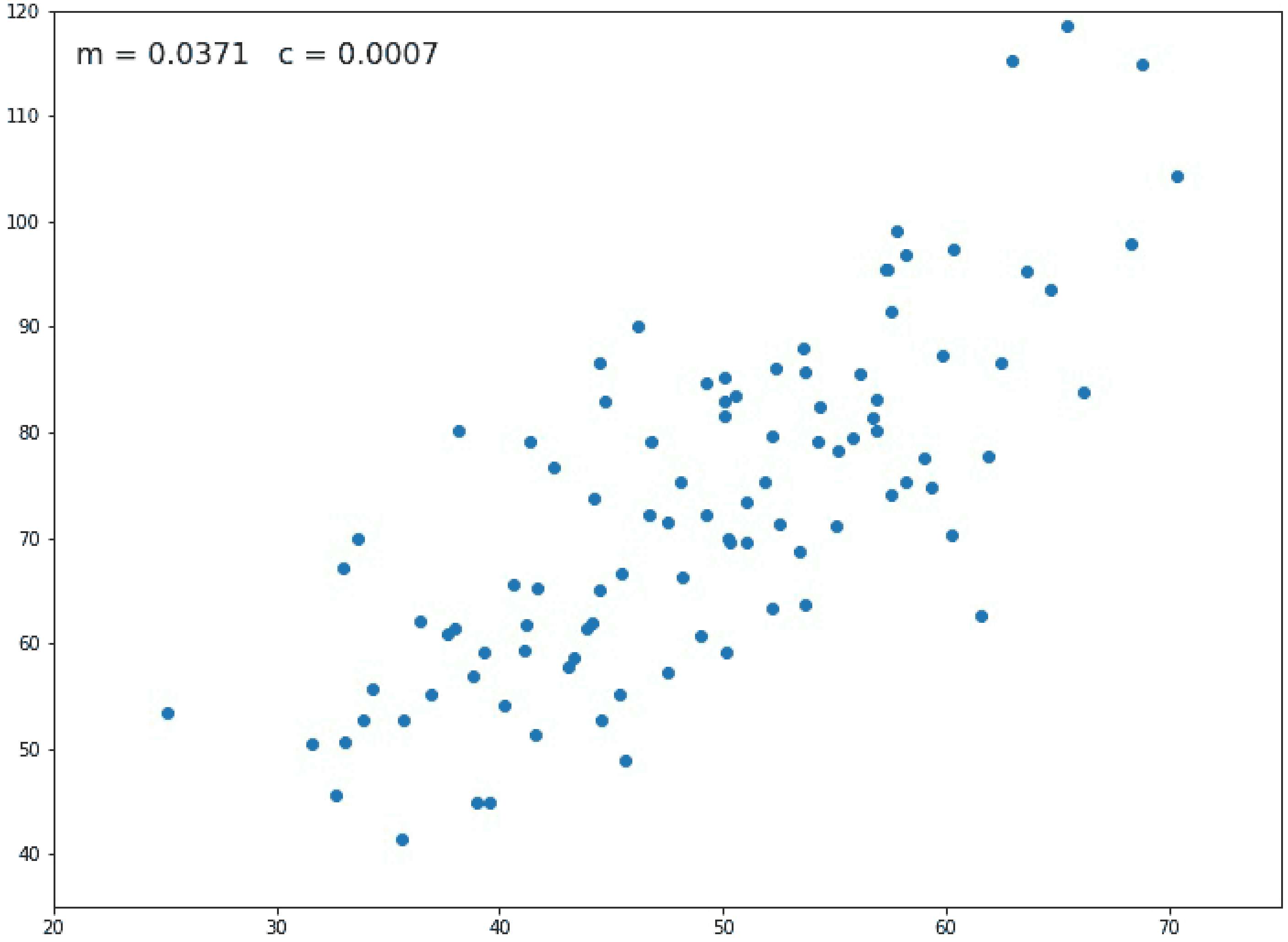
$j=n$

$$w_n = w_n - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right) - y^{(i)}\right)x_n^{(i)}$$

$$b = b - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right) - y^{(i)}\right)$$

simultaneously update $w_j$ (for $j = 1, \cdots, n$) and $b$

}

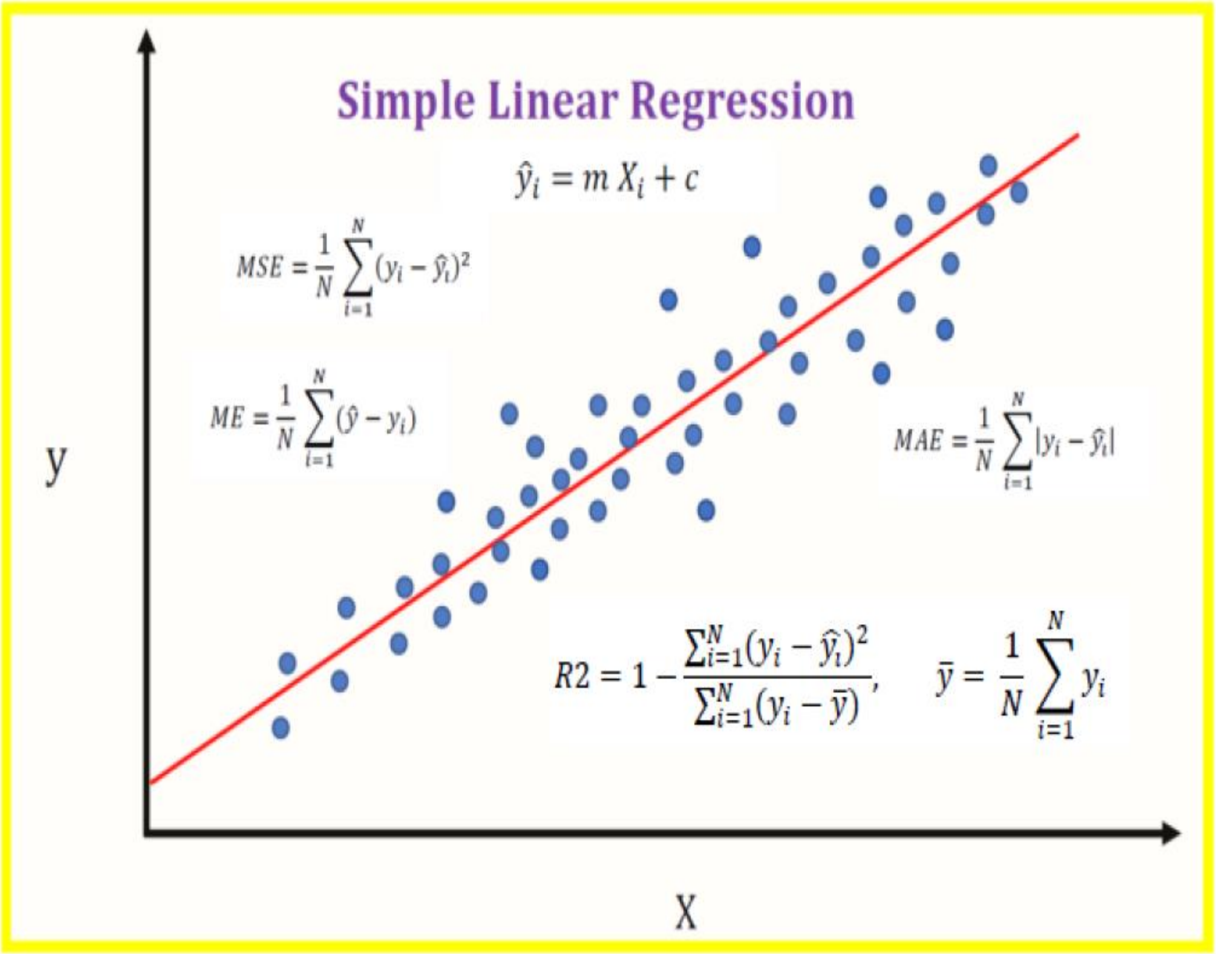

m = 0.0371   c = 0.0007
120
110
100
90
80
70
60
50
40
20
30
40
50
60
70

# Evaluation Metrics:

**MAE:** measures the average absolute differences between predicted and actual values.

**MSE:** measures the average of the squares of the errors, which gives more weight to larger errors.

**RMSE:** is the square root of MSE, providing error in the same units as the target variable.



# Assumptions of Linear Regression:

**Linearity**: The relationship between the independent variable and the dependent variable must be linear.

**Independence of error:** The observations should be independent of each other.

**Multicollinearity**: when two or more independent variable are not correlated to each other.

**Homoscedasticity**: The spread of errors should be roughly the same for all values of X.

**Normal Distribution of Residuals**: The prediction errors should follow a normal distribution