Linear Regression

BCSE0105: MACHINE LEARNING

BCSE0105: MACHINE LEARNING

Objective: *To introduce students to the basic concepts and techniques of Machine Learning. To develop skills of using recent machine learning software for solving practical problems. To gain experience of doing independent study and research.*

Credits: 03

L–T–P-J: 3–0–0-0

Module No.	Content	Teaching Hours
I	**Introduction:** Machine Learning basics, Hypothesis space and inductive bias, training and test set, and cross-validation. **Introduction to Statistical Learning:** Bayesian Method. **Machine Learning:** Supervised (Regression, Classification) vs. Unsupervised (Clustering) Learning. **Data Preprocessing:** Imputation, Outlier management, One hot encoding, Dimensionality Reduction- feature extraction, Principal Component Analysis (PCA), Singular Value Decomposition **Supervised Learning:** Regression- Linear regression, Polynomial regression, Classification- Logistic regression, k-nearest neighbor classifier,	20
II	**Supervised Learning:** Decision tree classifier, Naïve Bayes classifier Support Vector Machine (SVM)Classifier, **Unsupervised Learning:** k-means clustering, Hierarchical clustering **Underfitting vs Overfitting:** Regularization and Bias/Variance. **Ensemble methods:** Bagging, Boosting, Improving classification with Ada-Boost algorithm.	20

Text Book:

- Tom M. Mitchell, Machine Learning. Tata McGraw-Hill Education, 2013.
- Alpaydin, E. . Introduction to machine learning. MIT press, 2009.

Reference Books:

- Harrington, P. , " Machine learning in action", Shelter Island, NY: Manning Publications Co, 2012.
- Bishop, C. M. . Pattern recognition and machine learning (information science and statistics) springer-verlag new york. Inc. Secaucus, NJ, USA. 2006

Find the equation of line ?

- Two Points are given (3, 5) and (9,10)
- Find equation of line
- What will be slope (m) and y intercept (c)?

- Y= m.X + c
- Y= 0.83X+2.5

Quiz

X	Y
2	4
3	9
5	25
9	81
7	49
11	121
10.5	WHAT?

Quiz

X	Y
2	4
3	9
5	25
9	81
7	49
11	121
10.5	110.25

How Did You Find That?

- You find the relation between X and Y

$$Y = X.X = X^2$$

$$Y = f(X)$$

Which one is dependent variable ?

ANSWER = Y

SO What is X?

Independent variable

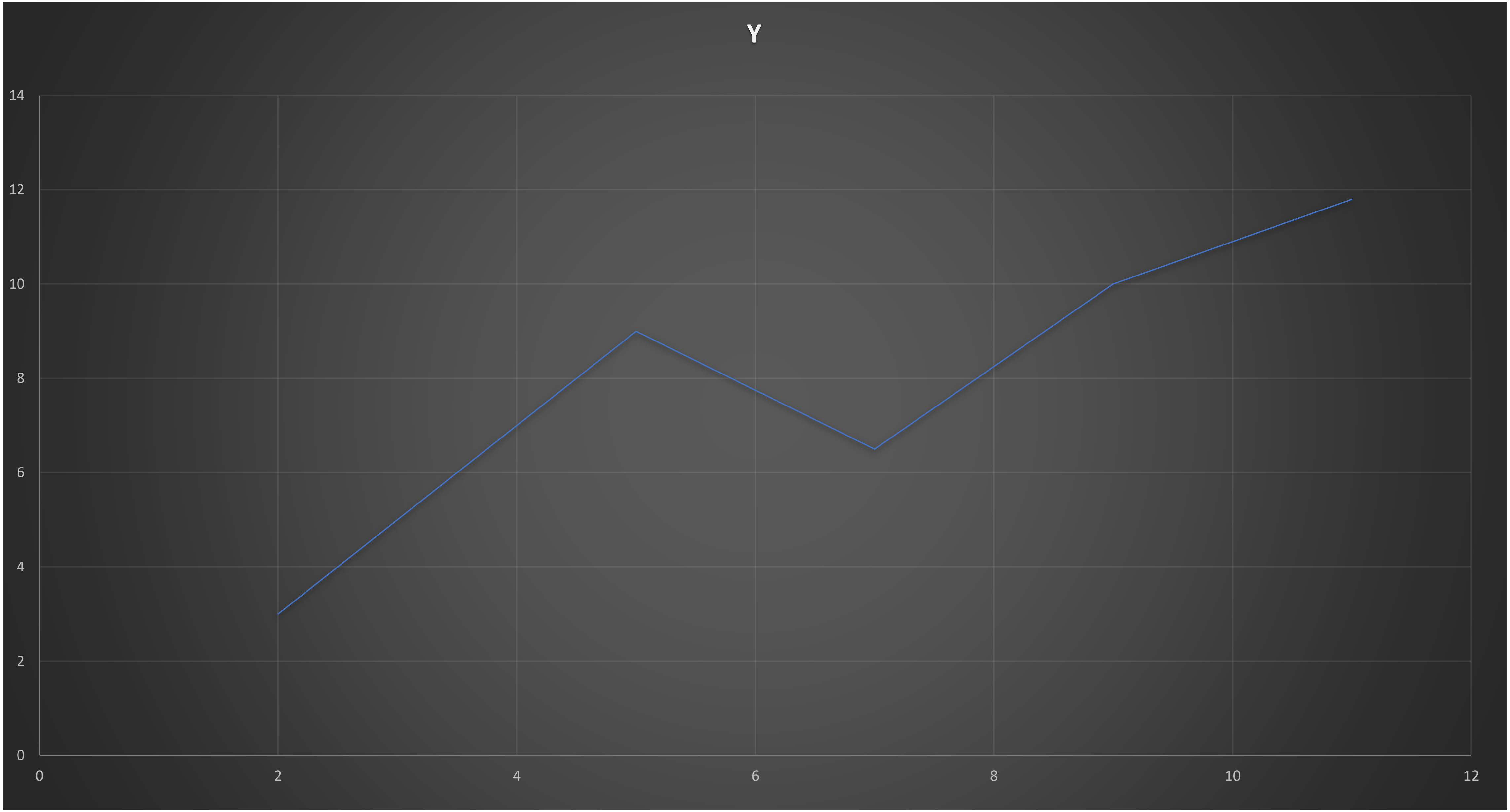
New Quiz

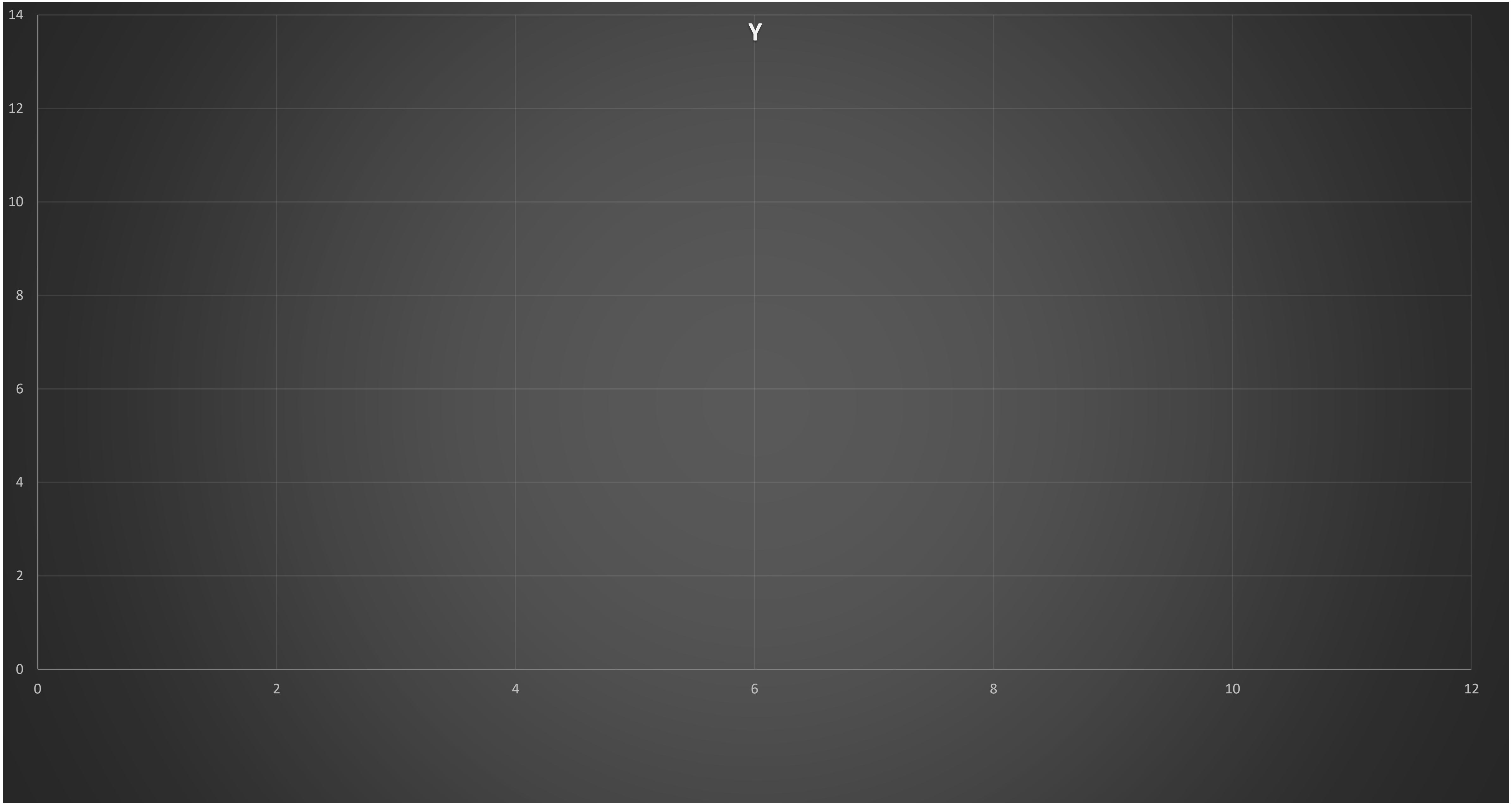
X	Y
2	3
3	5
5	9
9	10
7	6.5
11	11.8
10.5	WHAT?

New Quiz

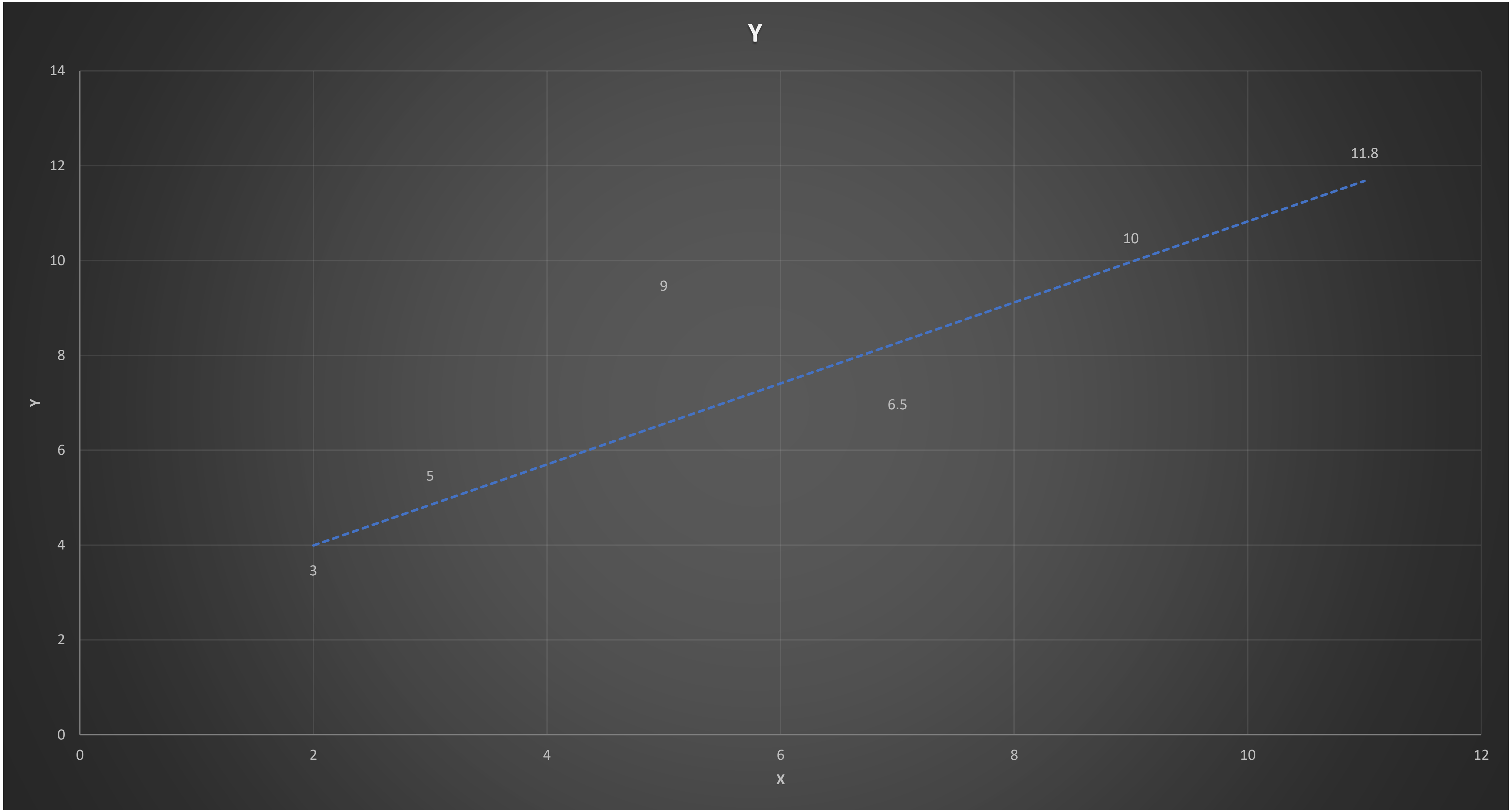
X	Y
2	3
3	5
5	9
9	10
7	6.5
11	11.8
10.5	Is it difficult to find out the relation ?

Graph is solution




Y
0
2
4
6
8
10
12
14
0
2
4
6
8
10
12

Approximation

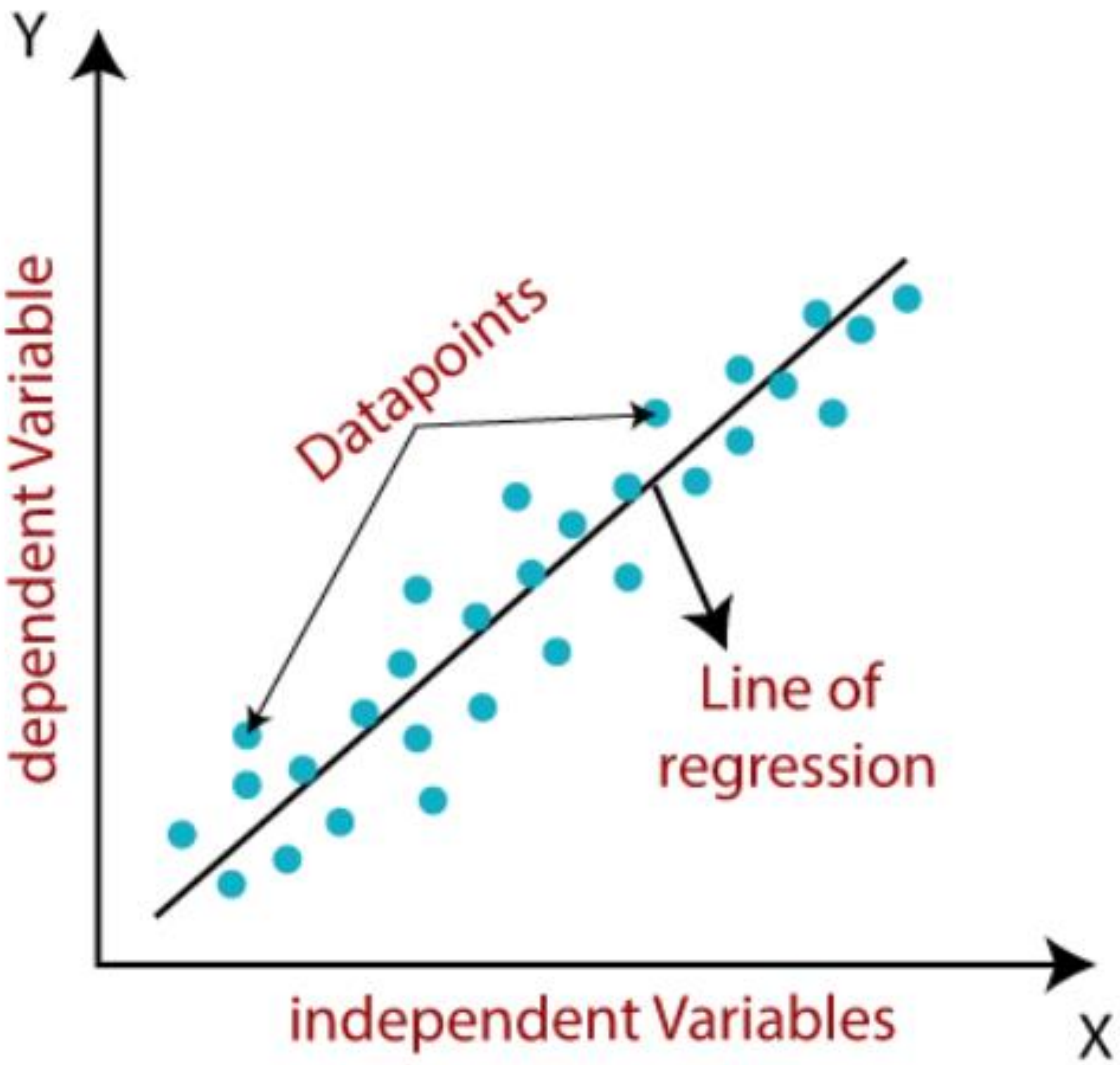


Linear Regression

- Finding the best fit line between dependent variable and Independent variable is called Linear Regression.

- Linear regression analysis is used **to predict the value of a variable based on the value of another variable**.

- Linear Regression is an **ML algorithm used for supervised learning**. Linear regression performs the task to predict a dependent variable(target) based on the given independent variable(s).


Y
dependent Variable
Datapoints
Line of regression
independent Variables
X

What is Regression ?

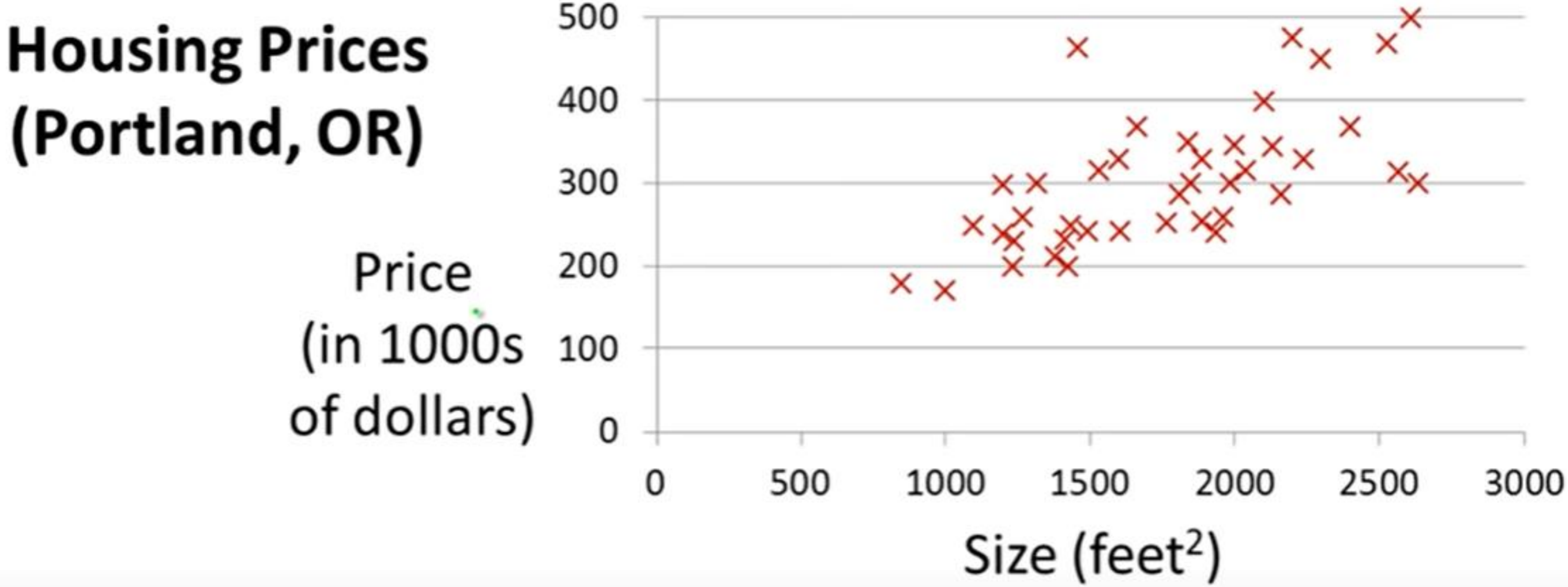
"Regression analysis is a form of predictive modelling technique which investigates the relationship between a dependent and independent variable"

Uses of Regression

Three major uses for regression analysis are

- Determining the strength of predictors
- Forecasting an effect, and
- Trend forecasting

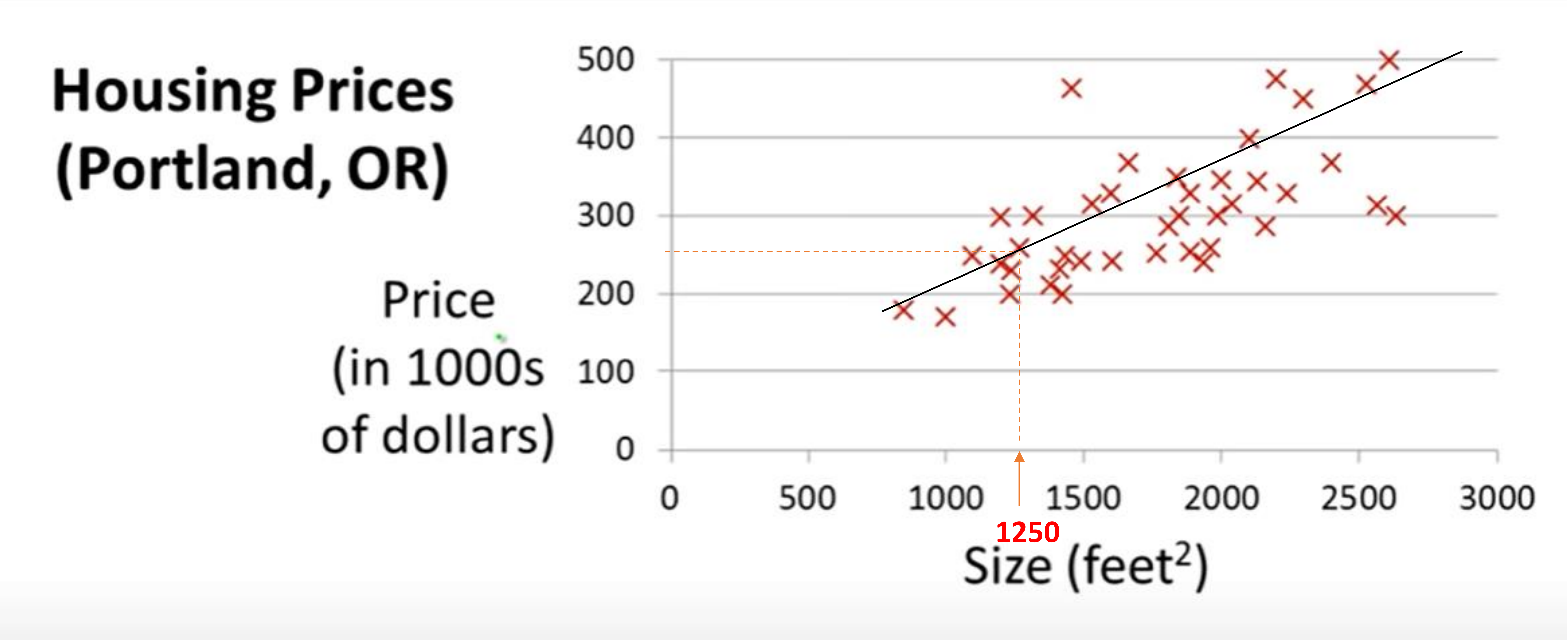
Example



Housing Prices
(Portland, OR)
Price
(in 1000s
of dollars)
500
400
300
200
100
0
0
500
1000
1500
2000
2500
3000
Size (feet^2)

To know the cost of House with 1250 sq. feet

Fit the straight line to the data



Points to remember

- Supervised Learning
  - Given the right answer for each example in the data.

- Regression
  - Predict real-valued output

- Classification
  - Discrete valued output

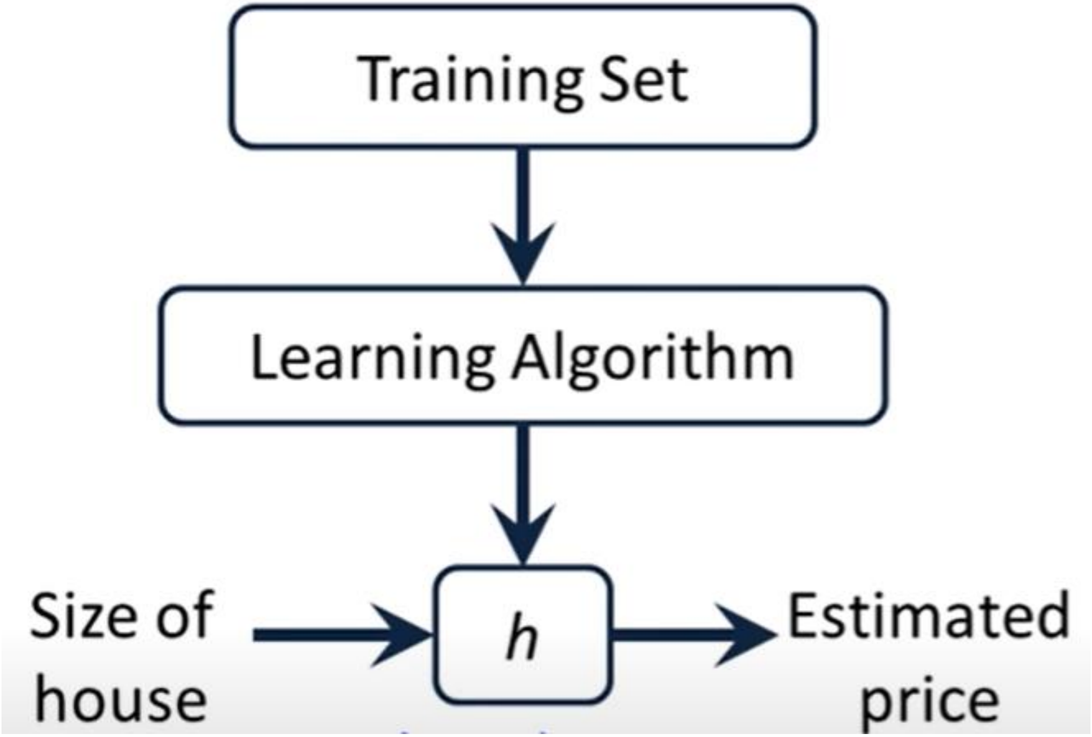
Notations

Let

- m→ number of training examples
- X’s → “input” variables/features
- Y’s → “output” variables/features
- (x,y)→ one training example
- (x(i),y(i))→i^{th} training example

Size in $feet^2$ (x)	Price ($) in 1000's (y)
2104	460
1416	232
1534	315
852	178
…	…

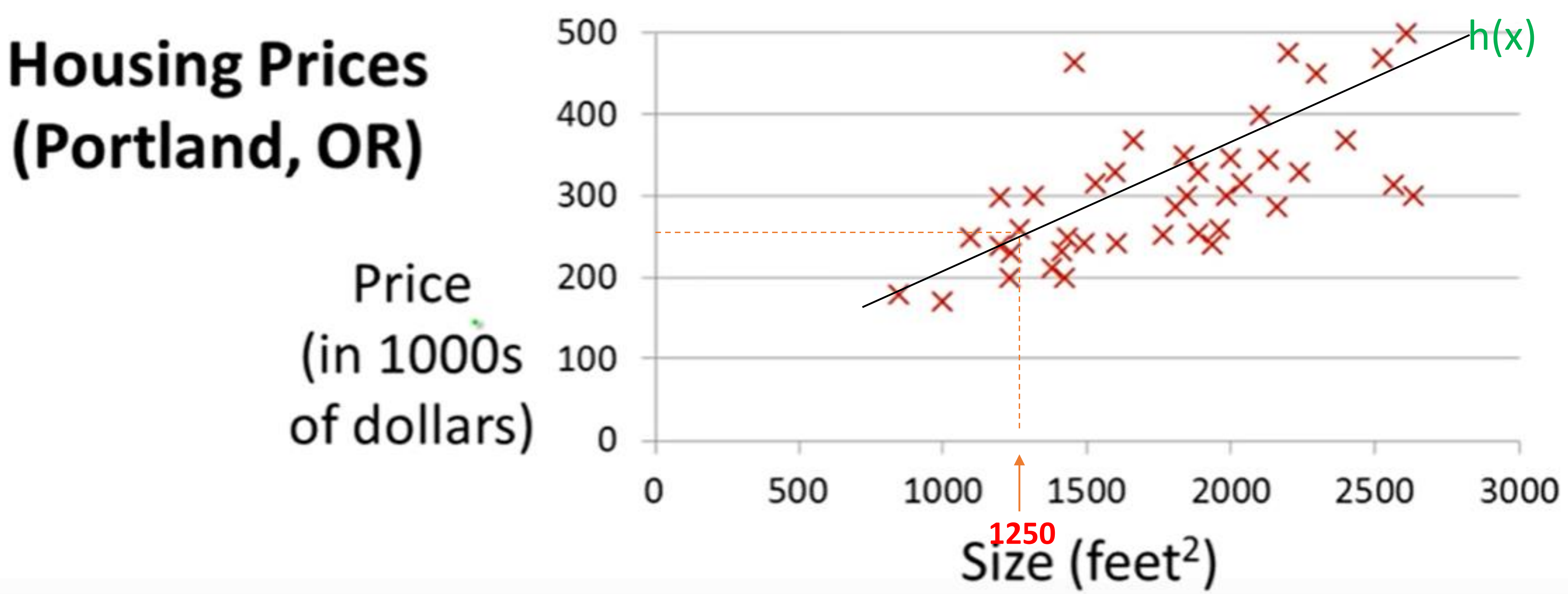
How Supervised Learning Algorithm works?



x Hypothesis Estimated value of y

h maps from x’s to y’s

How we can represent h?
h(x)=a+bx



This model is a linear regression with one variable

The error (residuals)

ERROR =ACTUAL DATA – PREDICTED DATA

e=Y(actual) −Y(predicted)

"Question"

"How can we find the best fit line?"

"Answer"

If Y(actual) =Y(predicted) then e=0

Or

Minimize the error

How to find best fit line

- **Derivation of linear regression equation :**
- given a set of n points (X_i, Y_i) on a scatterplot,
- find the best-fit line, Y_i' =*a*+ bX_i
- such that the sum of squared errors in Y, $\sum(Y_i - Y_i')^2$ is minimized.
- We can minimize error by using any optimization algorithm (eg. Gradient descent algorithm)

For Linear Regression

Hypothesis:

$$h_\theta(x) = \theta_0 + \theta_1 x$$

Parameters:

$$\theta_0, \theta_1$$

Cost Function:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$

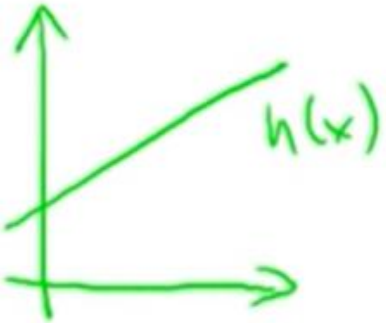
Understanding Cost Function

Hypothesis:

$$h_\theta(x) = \theta_0 + \theta_1 x$$

Parameters:

θ_0, θ_1



Cost Function:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

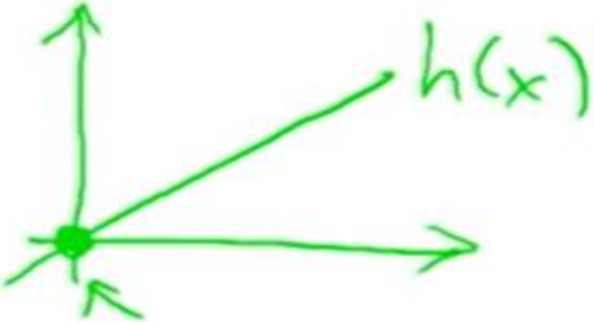
Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$

Simplified

$$h_\theta(x) = \theta_1 x$$

$\theta_0 = 0$

θ_1

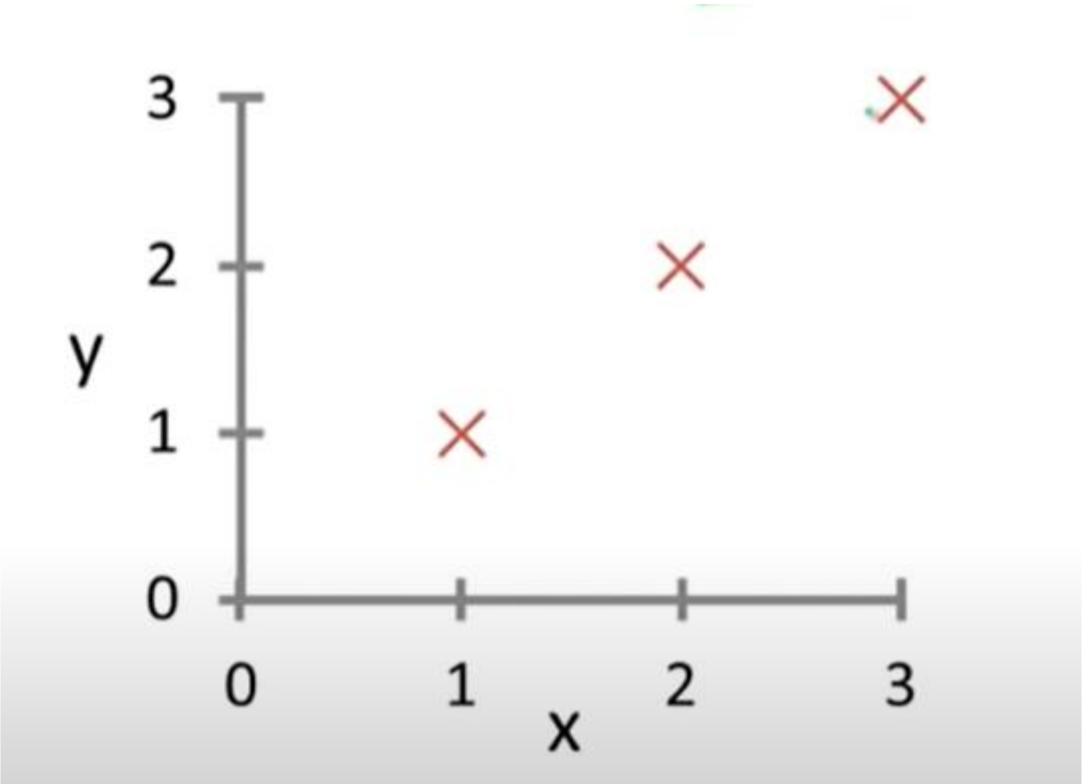


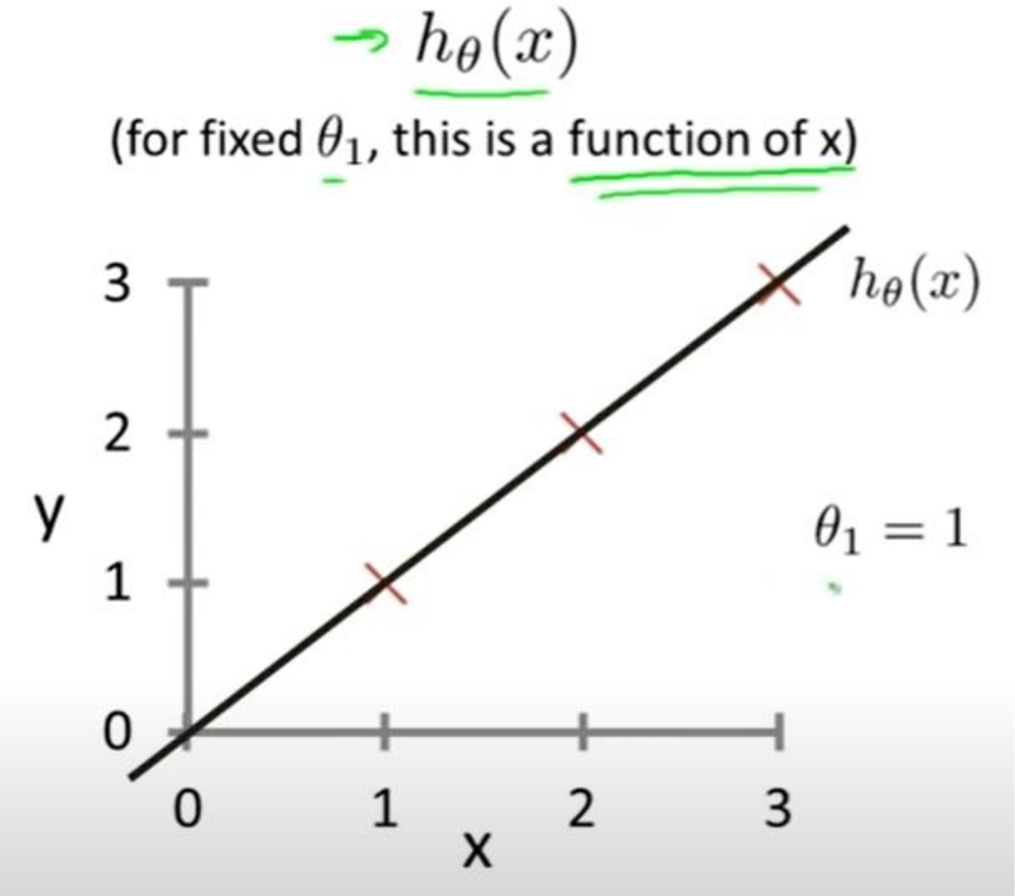
$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

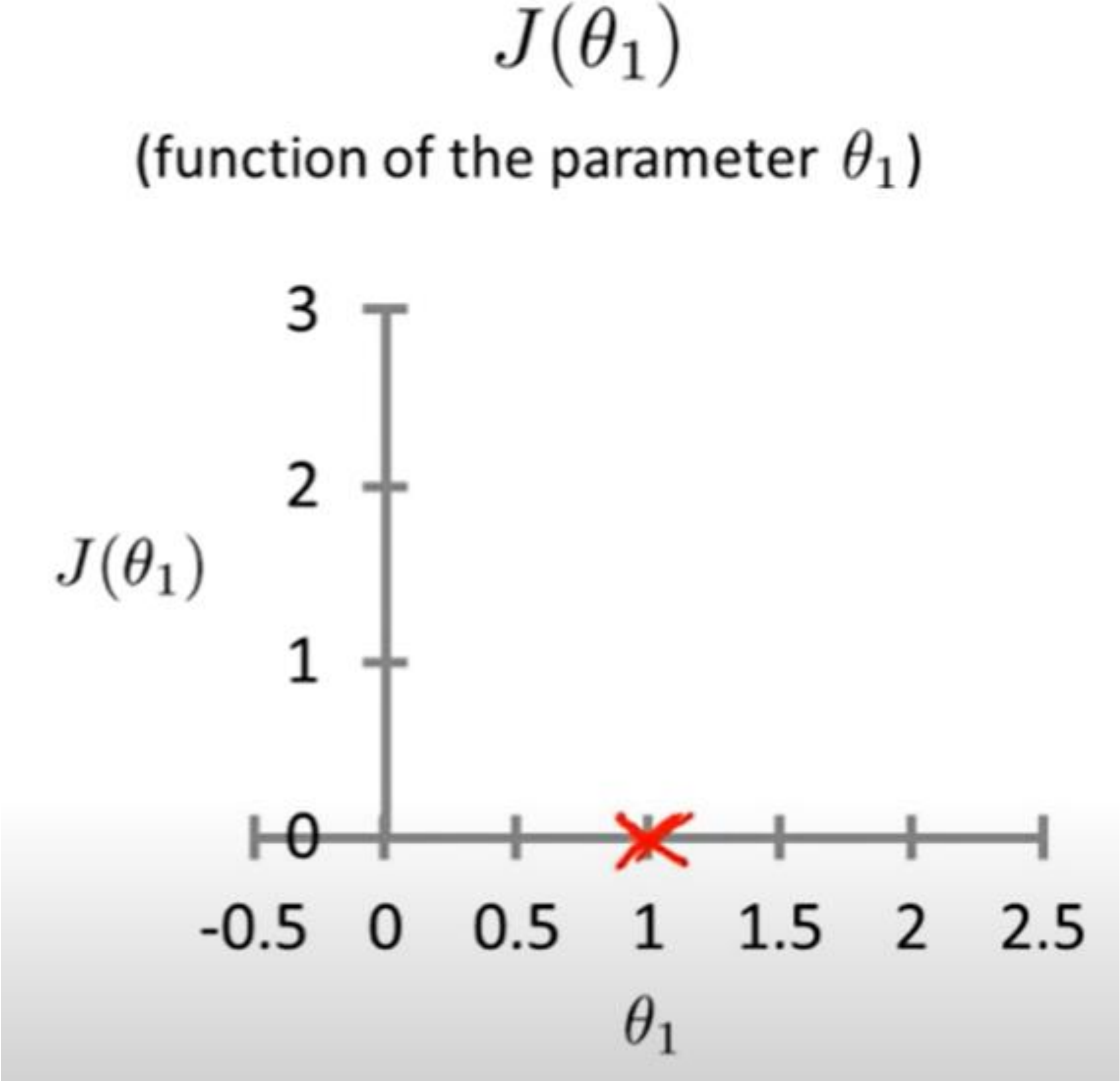
$\underset{\theta_1}{\text{minimize}}\ J(\theta_1)$

$\theta_1, x^{(i)}$

Let 3 data points (1,1), (2,2)and (3,3) are given as training set





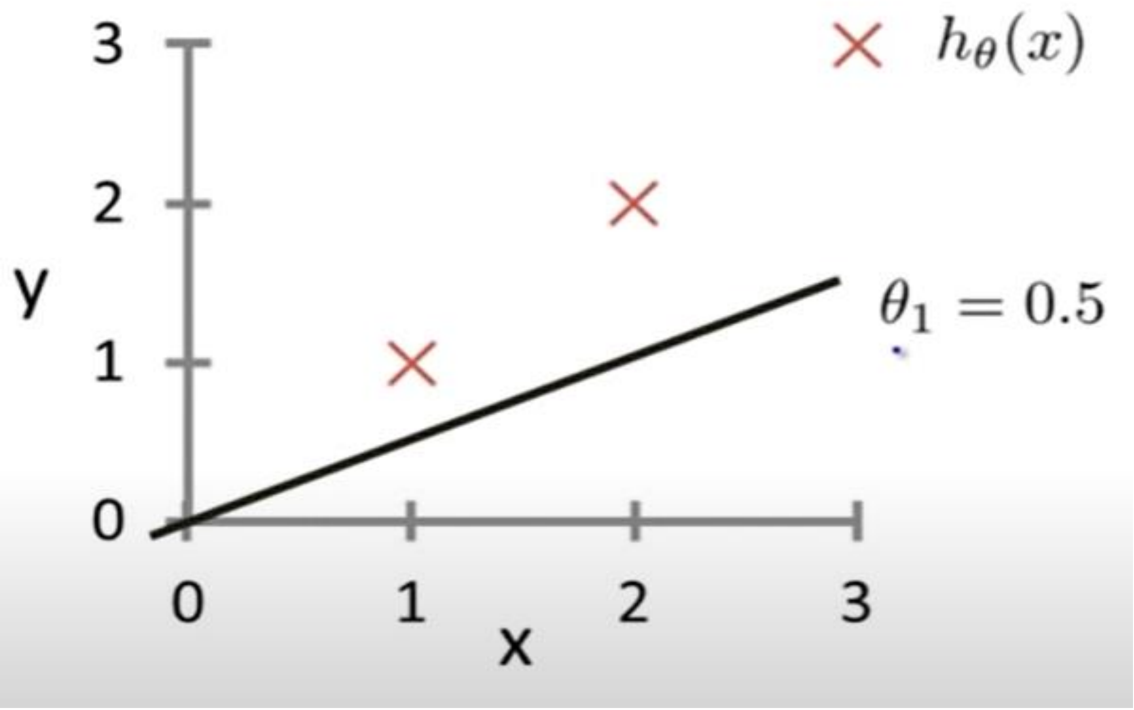


m=3, For $\theta_1 = 1$, find predicted y's and then J(1)=?

J(1)=0

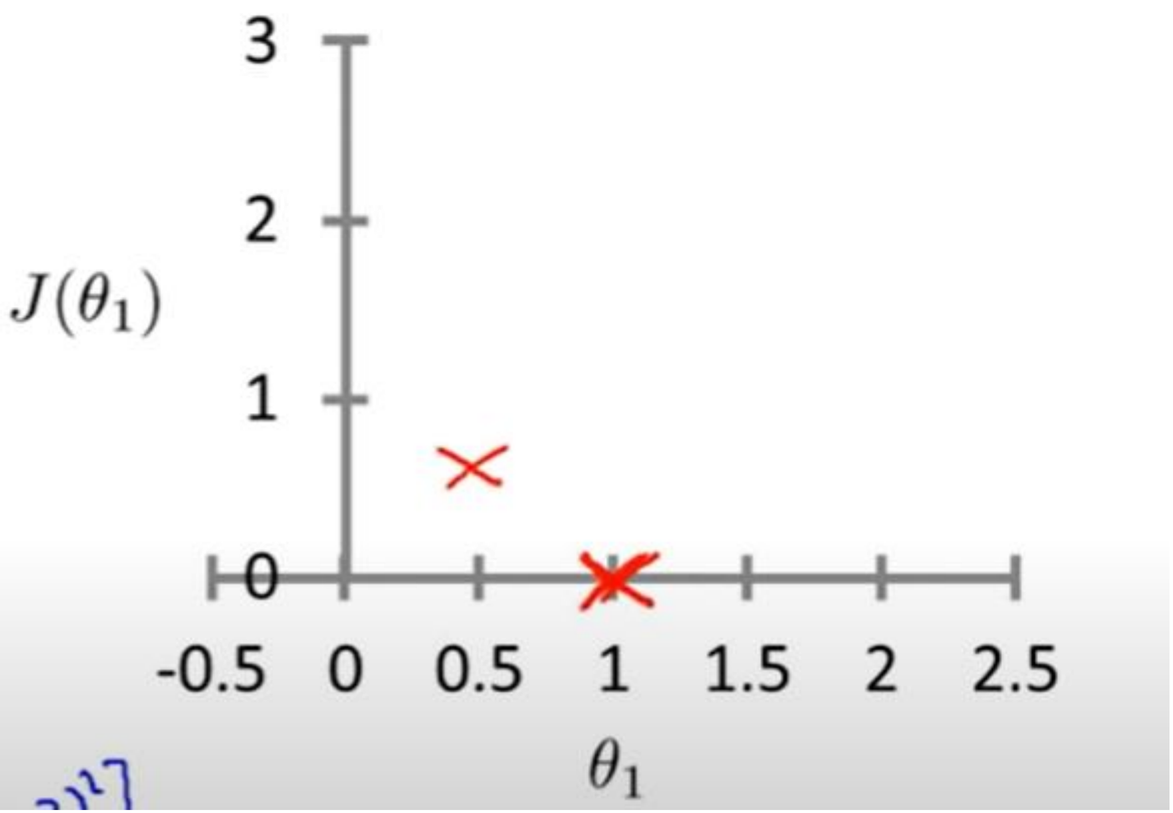
$h_\theta(x)$

(for fixed θ_1, this is a function of x)



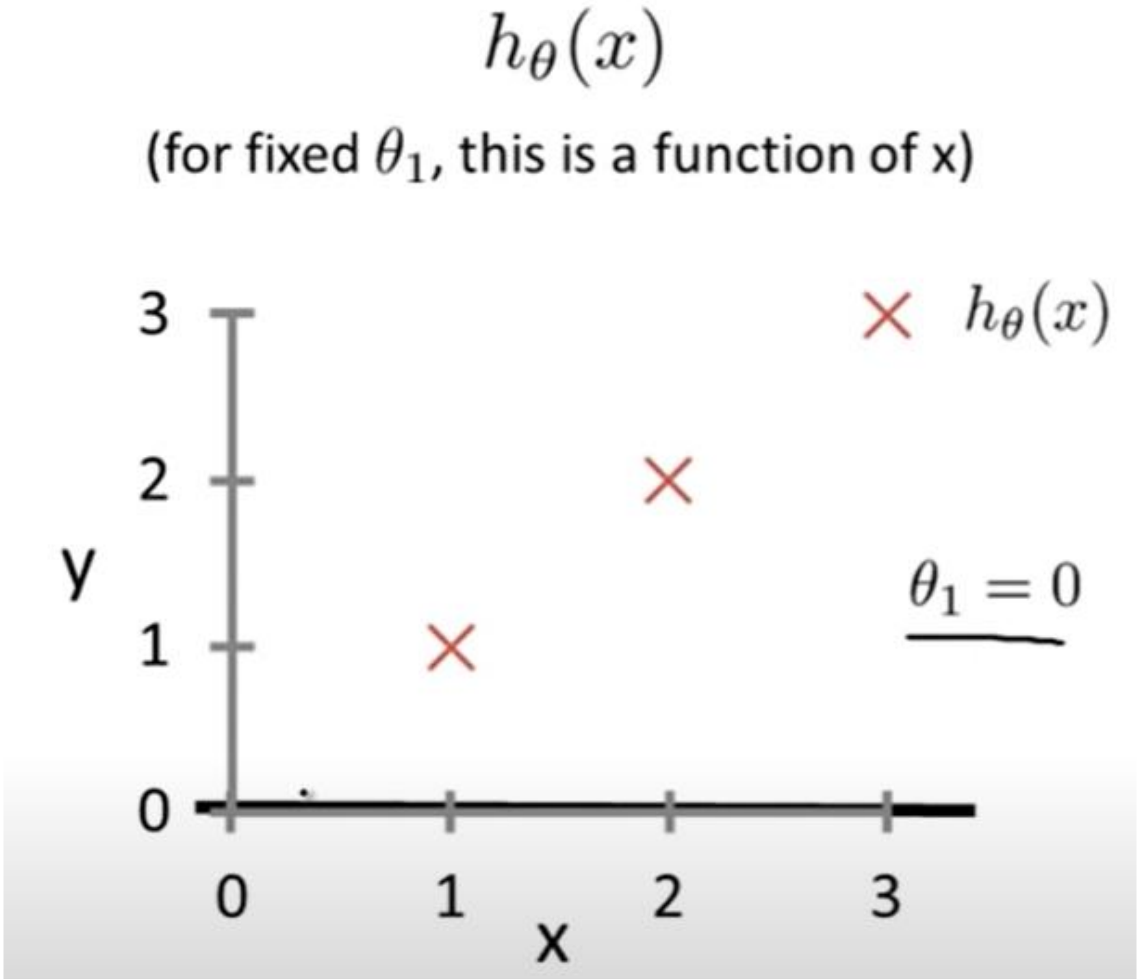
$J(\theta_1)$

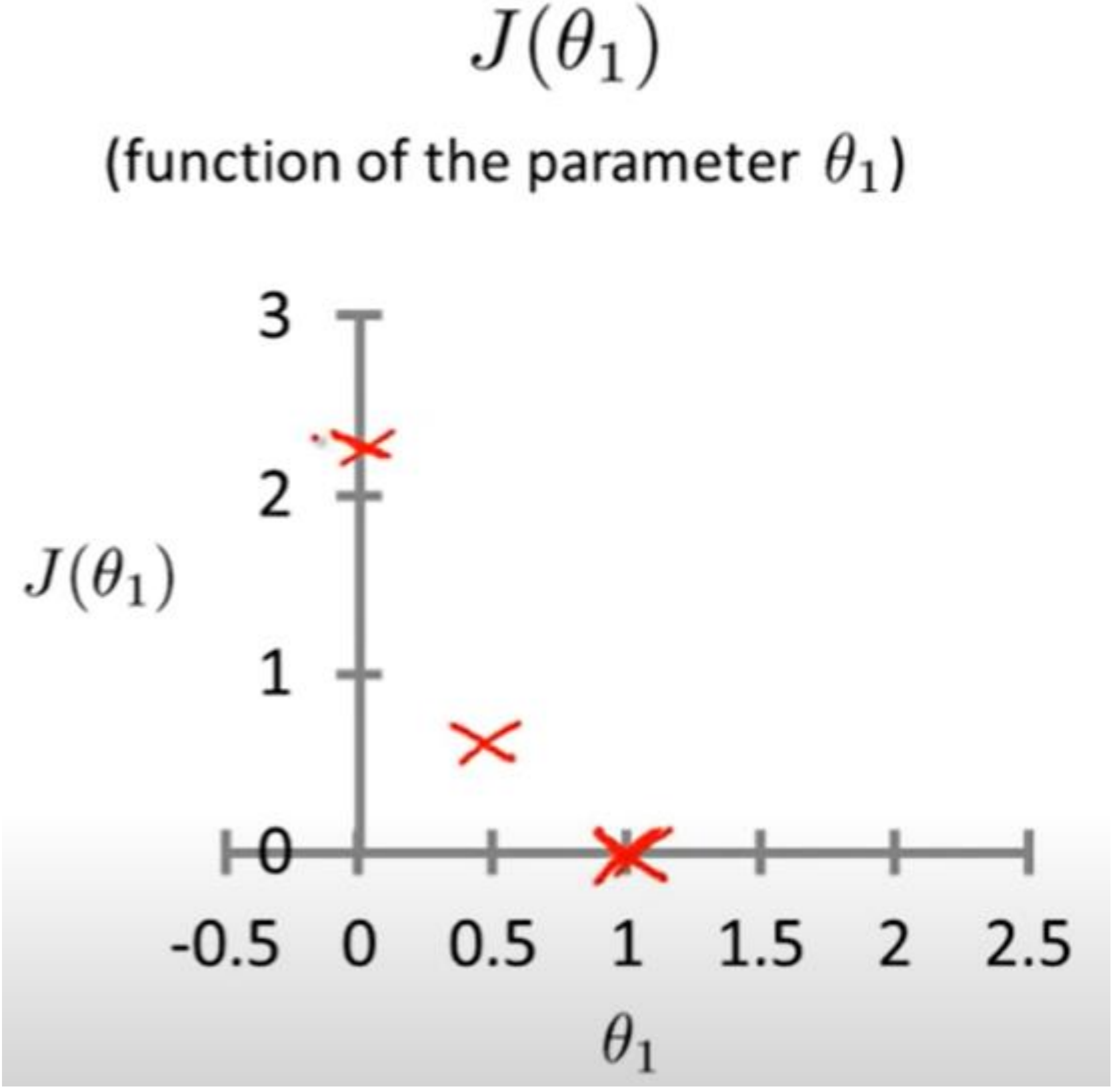
(function of the parameter θ_1)



m=3, For $\theta_1 = 0.5$, find predicted y's and then J(0.5)=?

J(0.5)=0.68

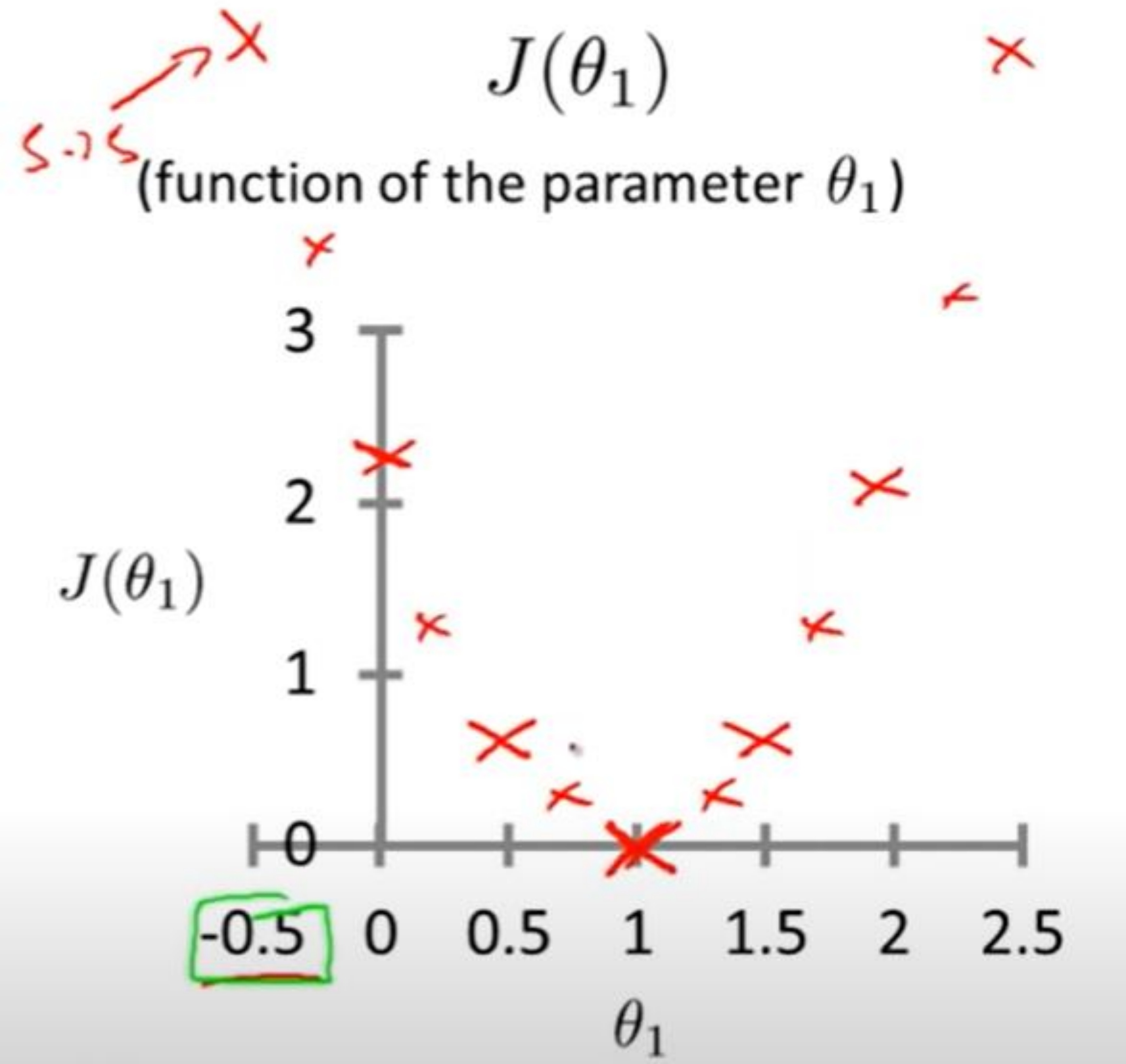




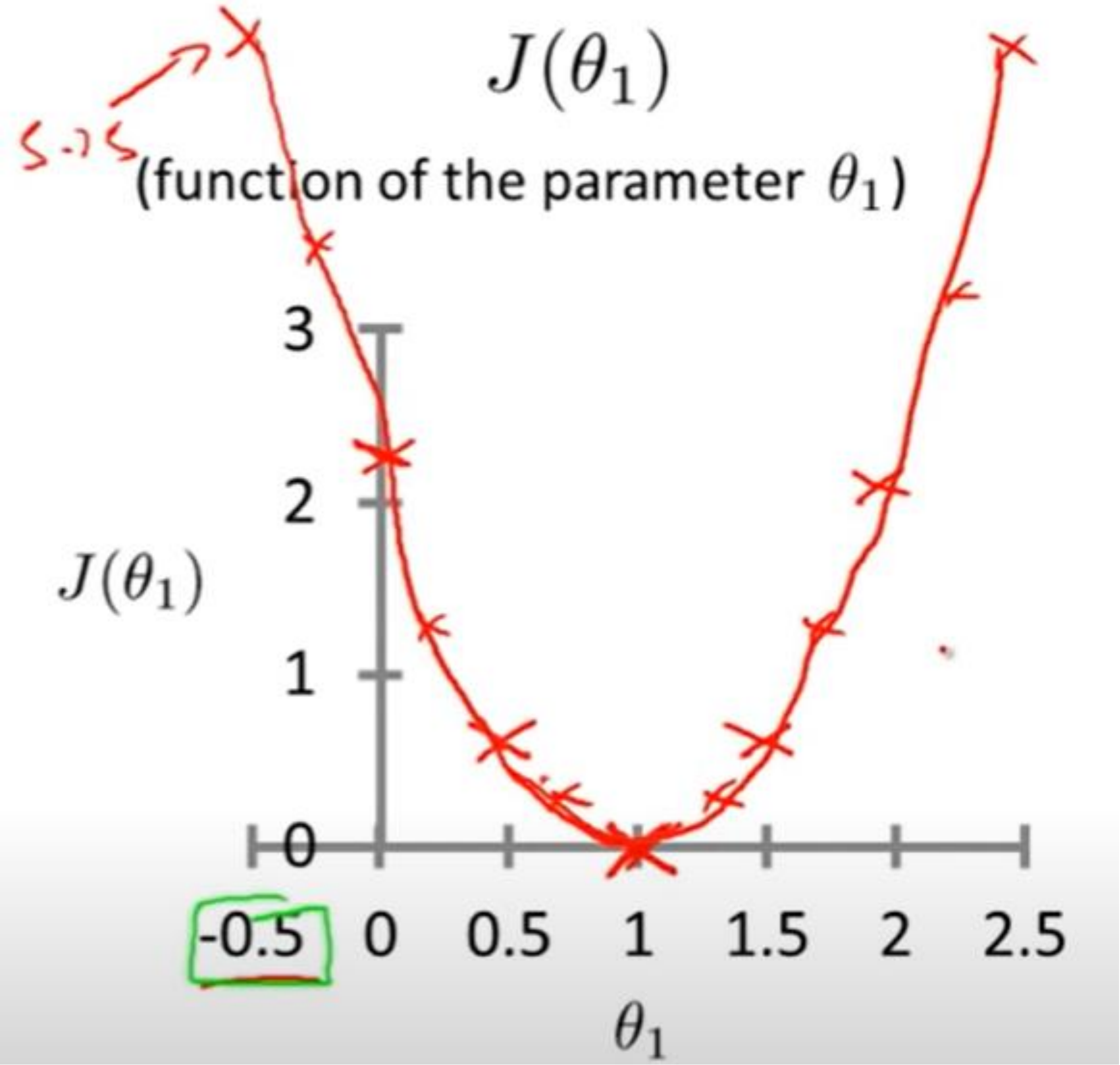
m=3, For $\theta_1 = 0$, find the predicted y's and then J(0)=?

J(0)=2.3

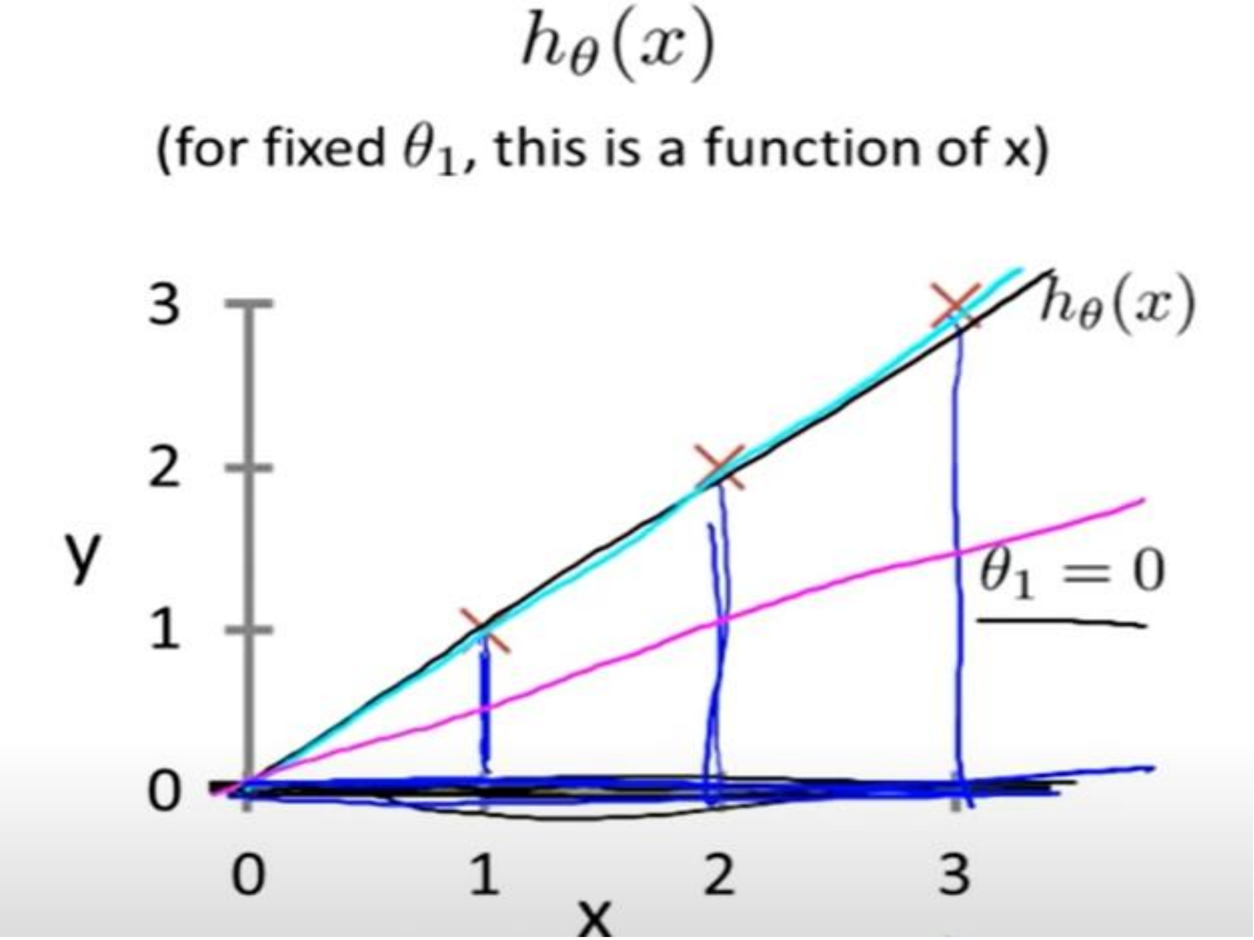
Keep on doing this for other values of $\boldsymbol{\theta_1}$

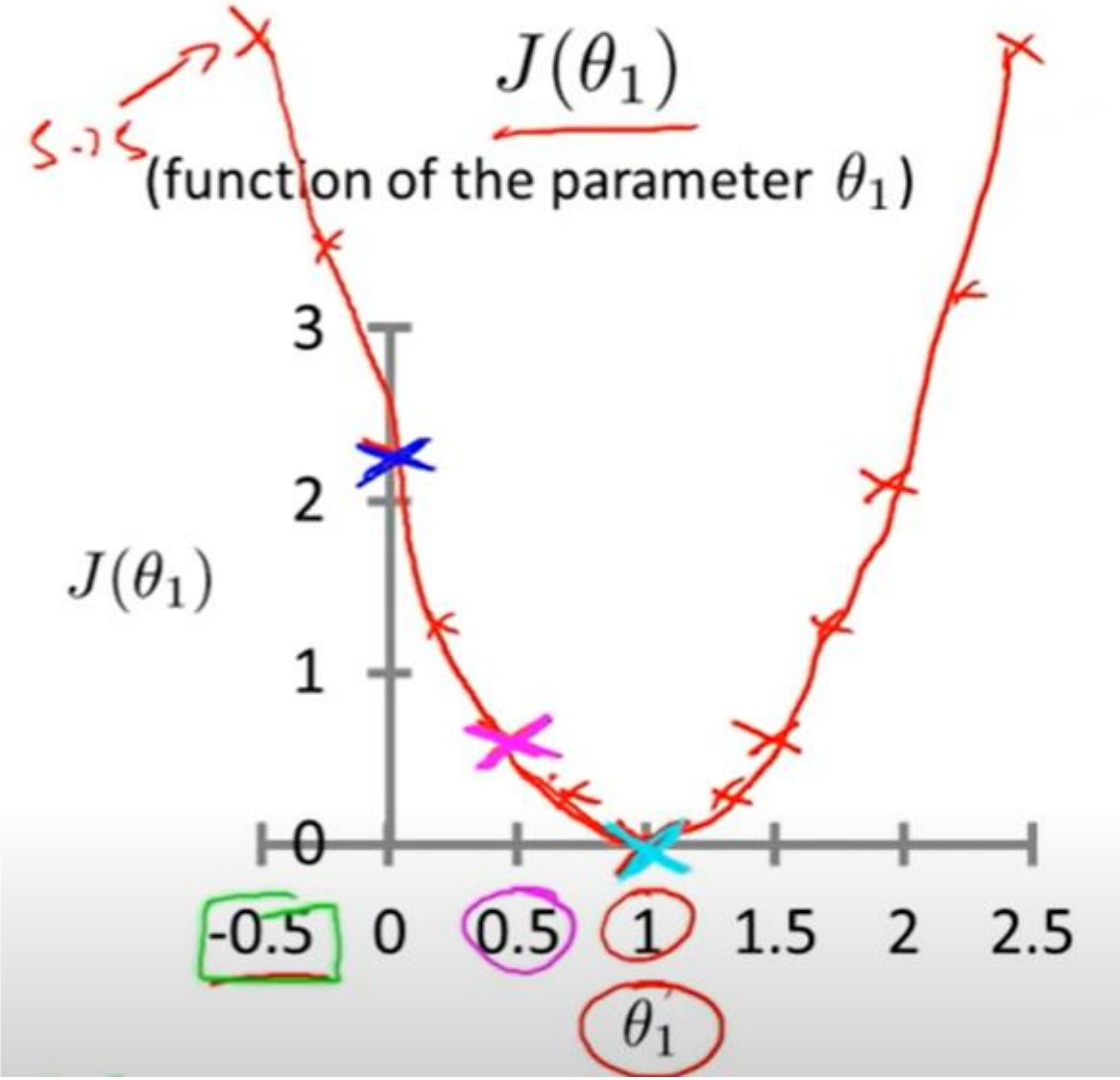


We get bowl shaped curve



For each value of θ_1 , we have different hypothesis (line or $h_\theta(x)$) and cost function ($J(\theta_1)$)





Remember our objective?

- To minimize the cost function ($\boldsymbol{J(\theta_1)}$)
- At $\theta_1 = 1,$
- We get the minimum value of $\boldsymbol{J(\theta_1)}$
- Therefore, the **best fit line** is the line corresponding to $\theta_1 = 1$

- But this is a manual method , we need some **algorithm** to minimize the cost function
- Gradient descent algorithm

Gradient Descent

- Gradient descent is used to minimize the MSE by calculating the gradient of the cost function.

- A regression model uses gradient descent to update the coefficients of the line by reducing the cost function.

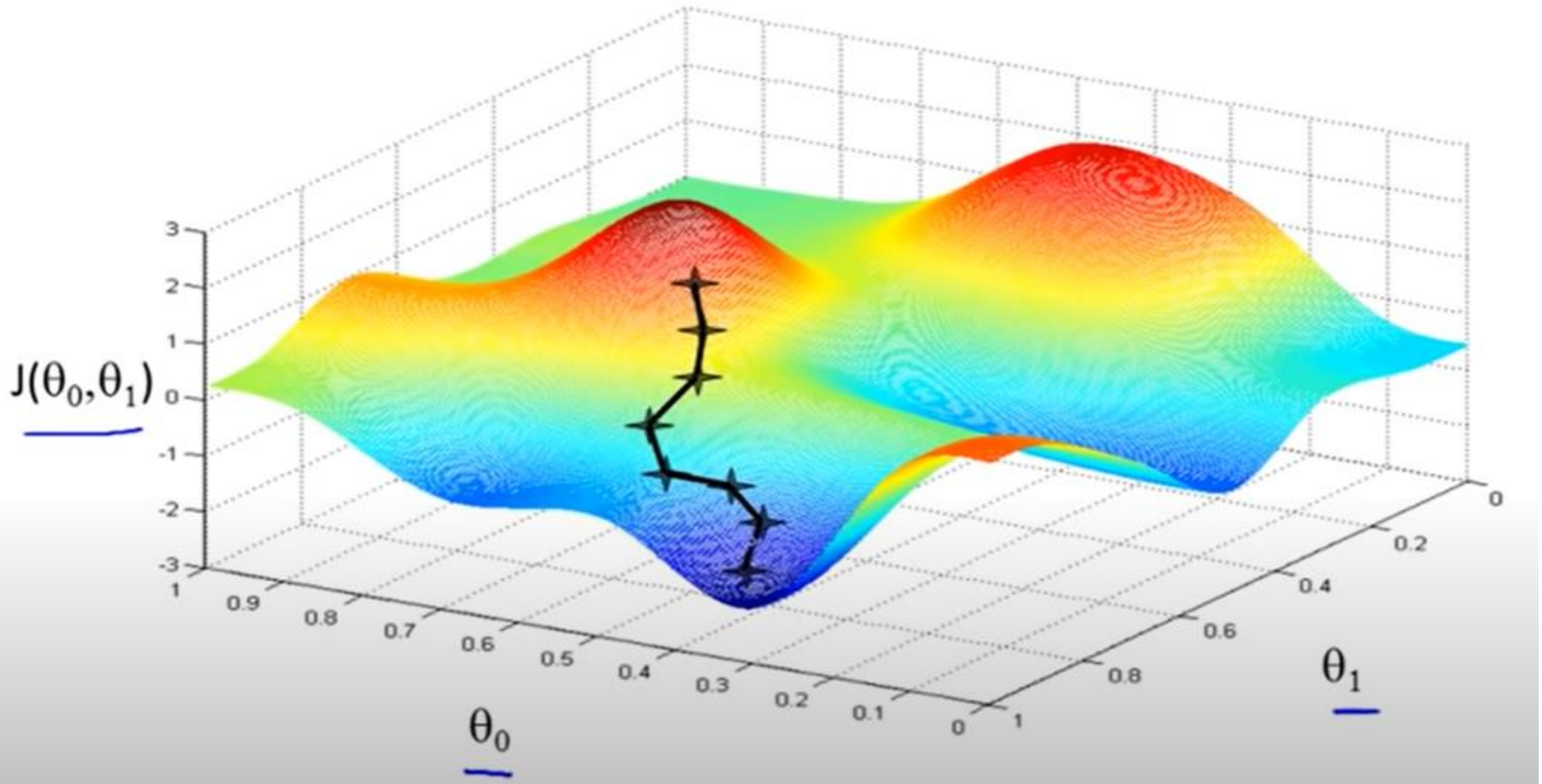
- It is done by a random selection of values of coefficient and then iteratively update the values to reach the minimum cost function.

Have some function $J(\theta_0, \theta_1)$

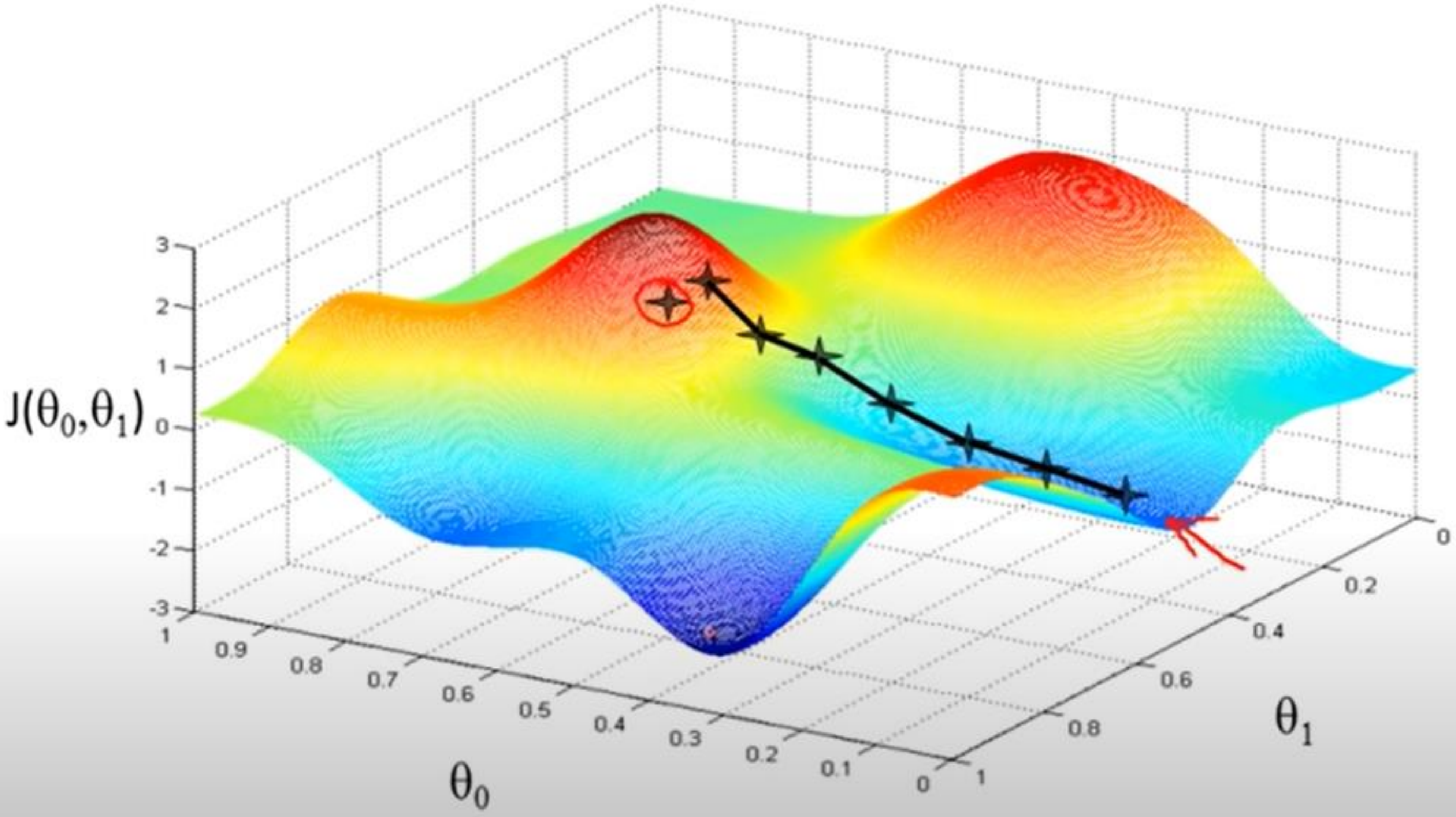
Want $\min_{\theta_0, \theta_1} J(\theta_0, \theta_1)$

Outline:

- Start with some θ_0, θ_1
- Keep changing θ_0, θ_1 to reduce $J(\theta_0, \theta_1)$ until we hopefully end up at a minimum



$J(\theta_0,\theta_1)$
3
2
1
0
-1
-2
-3
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
θ_0
0
0.2
0.4
0.6
0.8
1
θ_1



3
2
1
0
-1
-2
-3
$J(\theta_0,\theta_1)$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
θ_0
0
0.2
0.4
0.6
0.8
1
θ_1

Gradient Descent Algorithm

$$
\begin{aligned}
&\text{repeat until convergence } \{ \\
&\quad \theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \qquad (\text{for } j = 0 \text{ and } j = 1) \\
&\}
\end{aligned}
$$

Convergence means no further update.

α is learning rate (positive number)

Gradient Descent Algorithm

repeat until convergence {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \qquad (\text{for } j = 0 \text{ and } j = 1)$$

}

Correct: Simultaneous update

$$\text{temp0} := \theta_0 - \alpha \tfrac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$
$$\text{temp1} := \theta_1 - \alpha \tfrac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$$
$$\theta_0 := \text{temp0}$$
$$\theta_1 := \text{temp1}$$

Incorrect:

$$\text{temp0} := \theta_0 - \alpha \tfrac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$
$$\theta_0 := \text{temp0}$$
$$\text{temp1} := \theta_1 - \alpha \tfrac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$$
$$\theta_1 := \text{temp1}$$

Gradient Descent for Linear Regression

Gradient descent algorithm

repeat until convergence {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

$$(\text{for } j = 1 \text{ and } j = 0)$$

}

Linear Regression Model

$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Objective is to minimize $J(\theta_0, \theta_1)$

- Calculate the derivative term
- $\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_j} \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$
- $\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_j} \frac{1}{2m} \sum_{i=1}^{m} (\theta_0 + \theta_1 x^{(i)} - y^{(i)})^2$
- Put j=0, for θ_0, we get,
- $\frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})$

- Put j=1, for θ_1, we get,
- $\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}).\, x^{(i)}$

Gradient Descent for Linear Regression

repeat until convergence {

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)$$

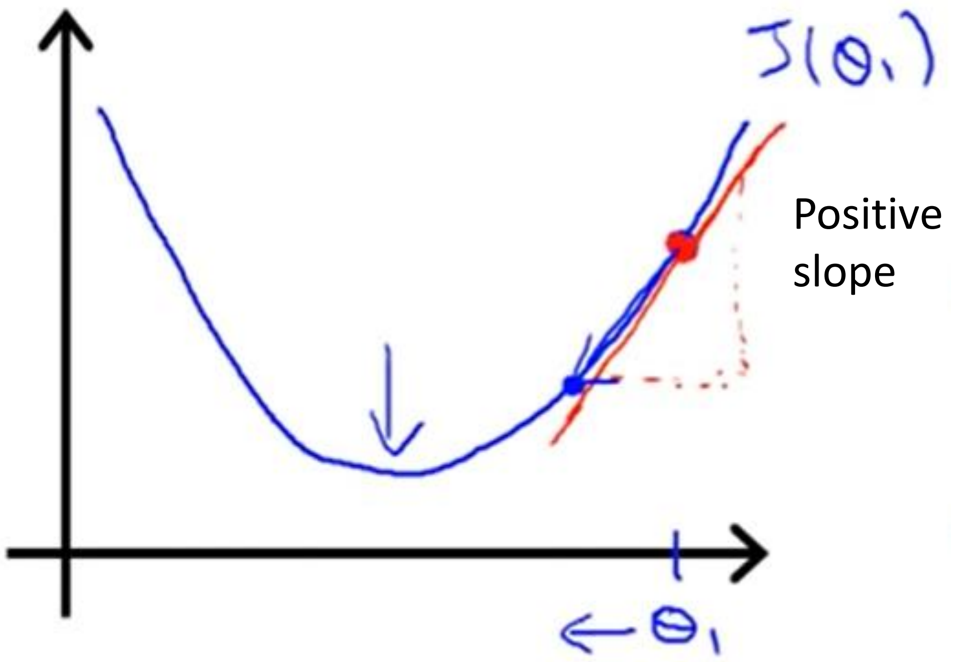
$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right) \cdot x^{(i)}$$

}

Update θ_0 *and* θ_1 simultaneously

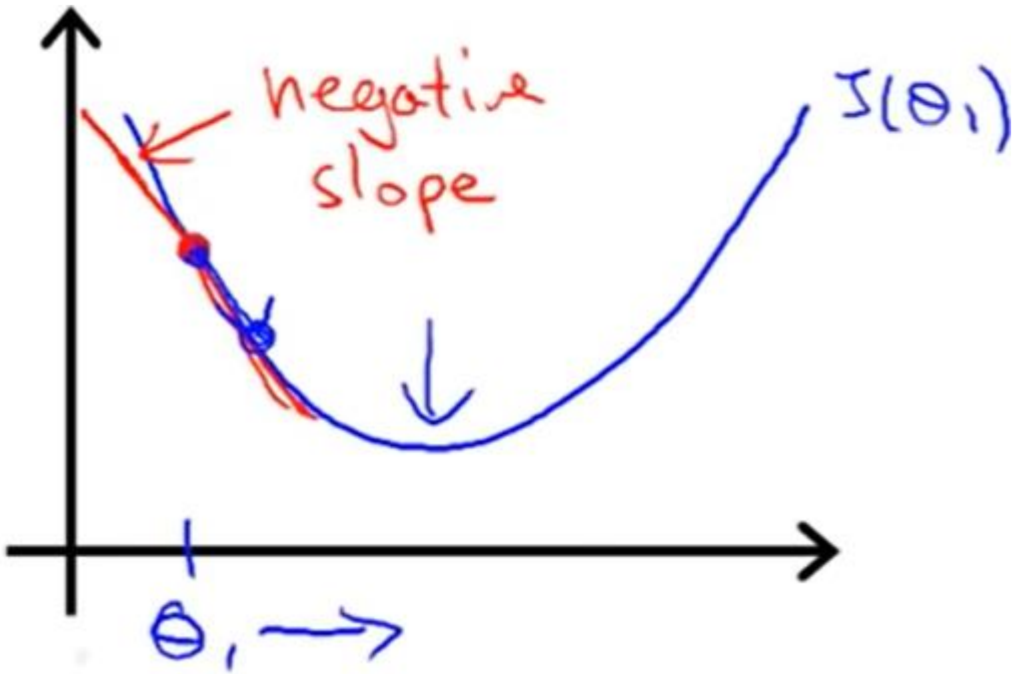
Significance of derivative term and learning rate

- Let $\theta \in R$
- Update in θ
- $\theta_1 = \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_1)$
- Where α is the learning rate
- And $\frac{\partial}{\partial \theta_1} J(\theta_1)$ is derivative term (slope)
- **If** $\boldsymbol{\frac{\partial}{\partial \theta_1} J(\theta_1) \geq 0}$
- **Then,** $\theta_1 = \theta_1 - \alpha(positive\ number)$
- Means, decrement in θ_1



Significance of derivative term and learning rate

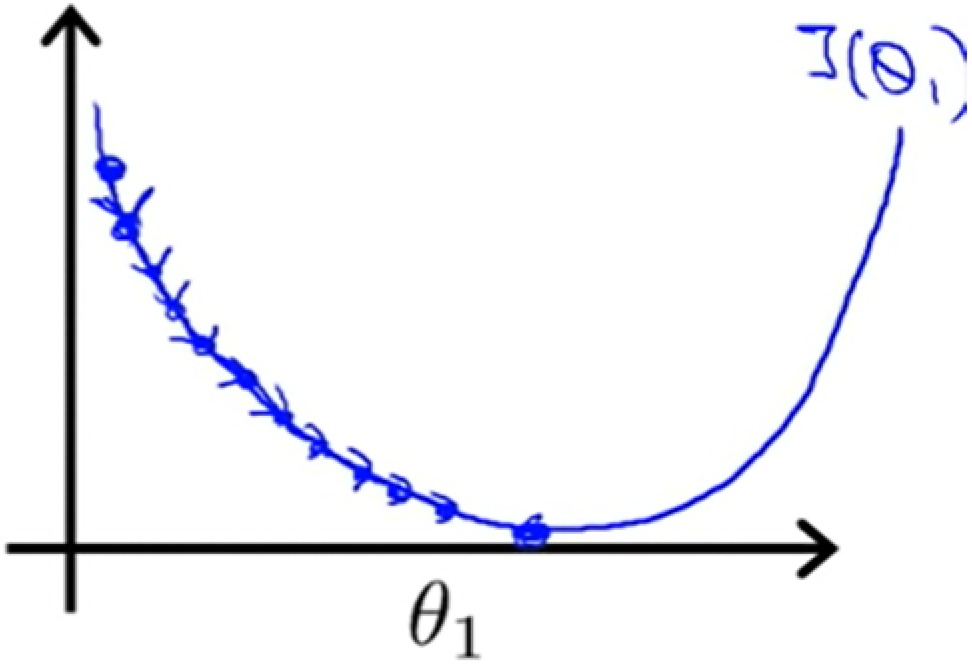
- **If** $\frac{\partial}{\partial \theta_1} J(\theta_1) \leq 0$
- **Then,** $\theta_1 = \theta_1 - \alpha(negative\ number)$
- Means, increment in θ_1



Small Learning rate

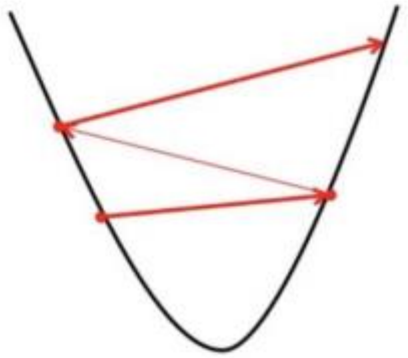
$$\theta_1 := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_1)$$

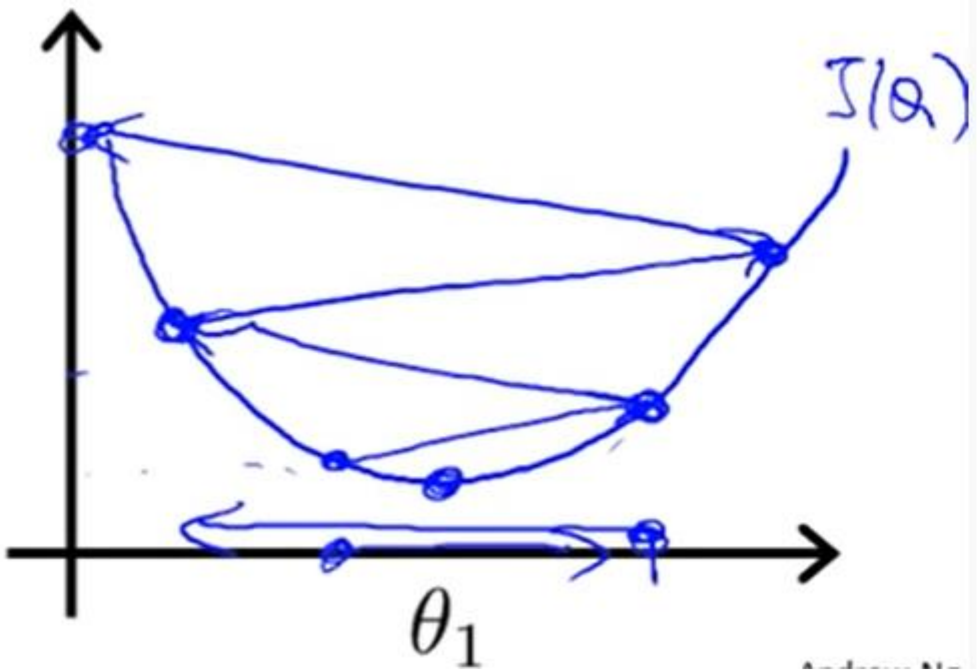
If α is too small, gradient descent can be slow.



Large Learning Rate

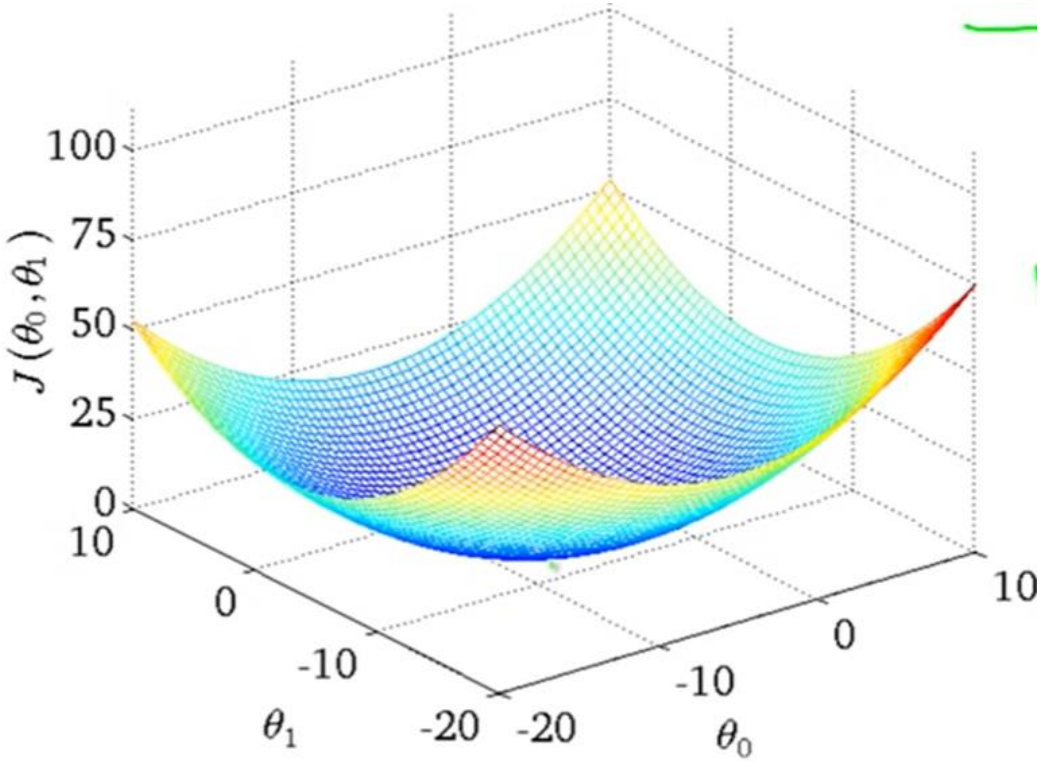
If α is too large, gradient descent can overshoot the minimum. It may fail to converge, or even diverge.





Convex function

- Cost function of linear regression is always a bowl shaped which is convex function



Applications and Advantage

- Medical researchers often use linear regression to understand the relationship between drug dosage and blood pressure of patients.
- Forecasting
- Prediction

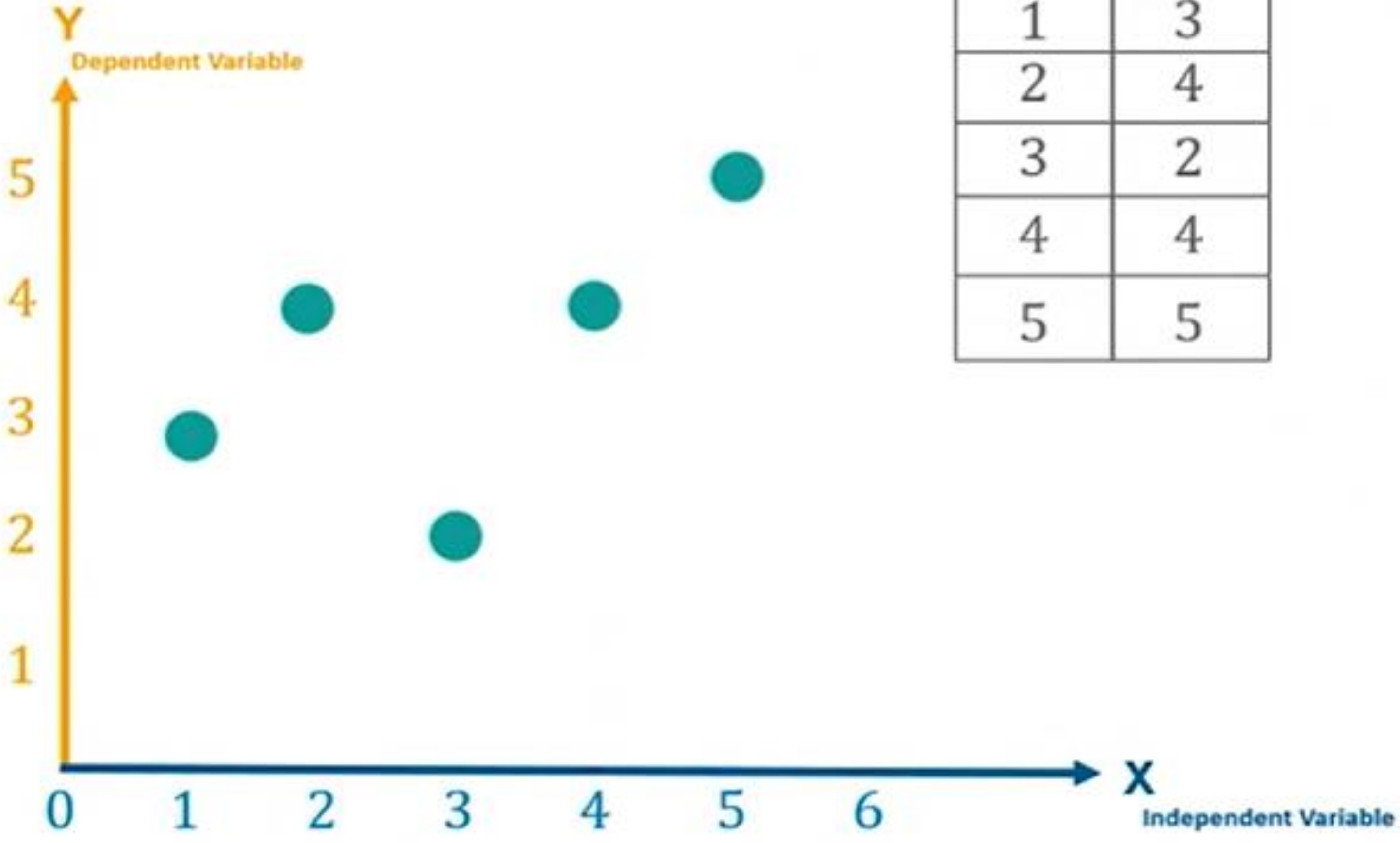
- The advantage of linear regression models is linearity: It makes the estimation procedure simple.

Limitation

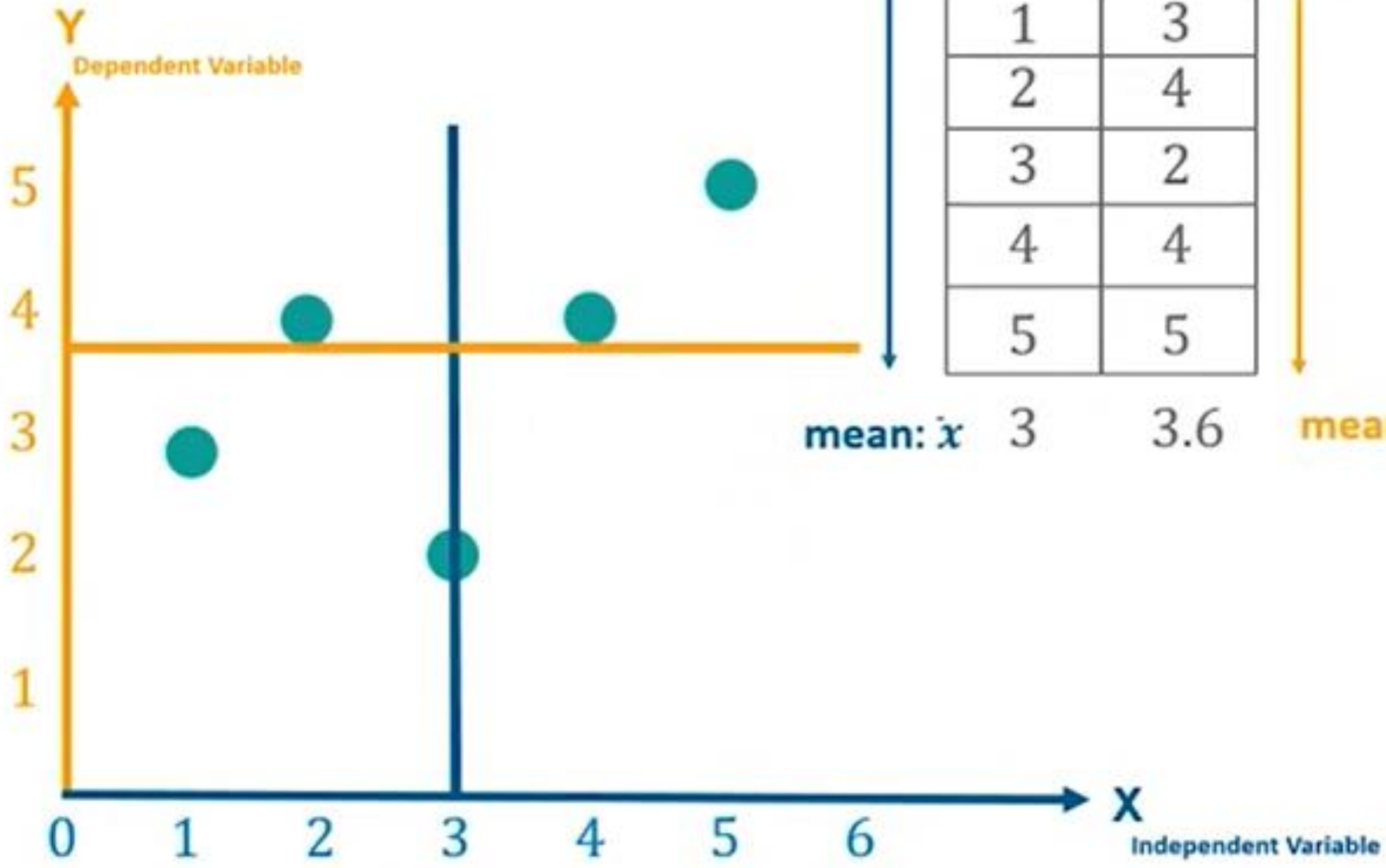
- Main limitation of Linear Regression is **the assumption of linearity between the dependent variable and the independent variables**. In the real world, the data is rarely linearly separable.

Example

Understanding Linear Regression Algorithm



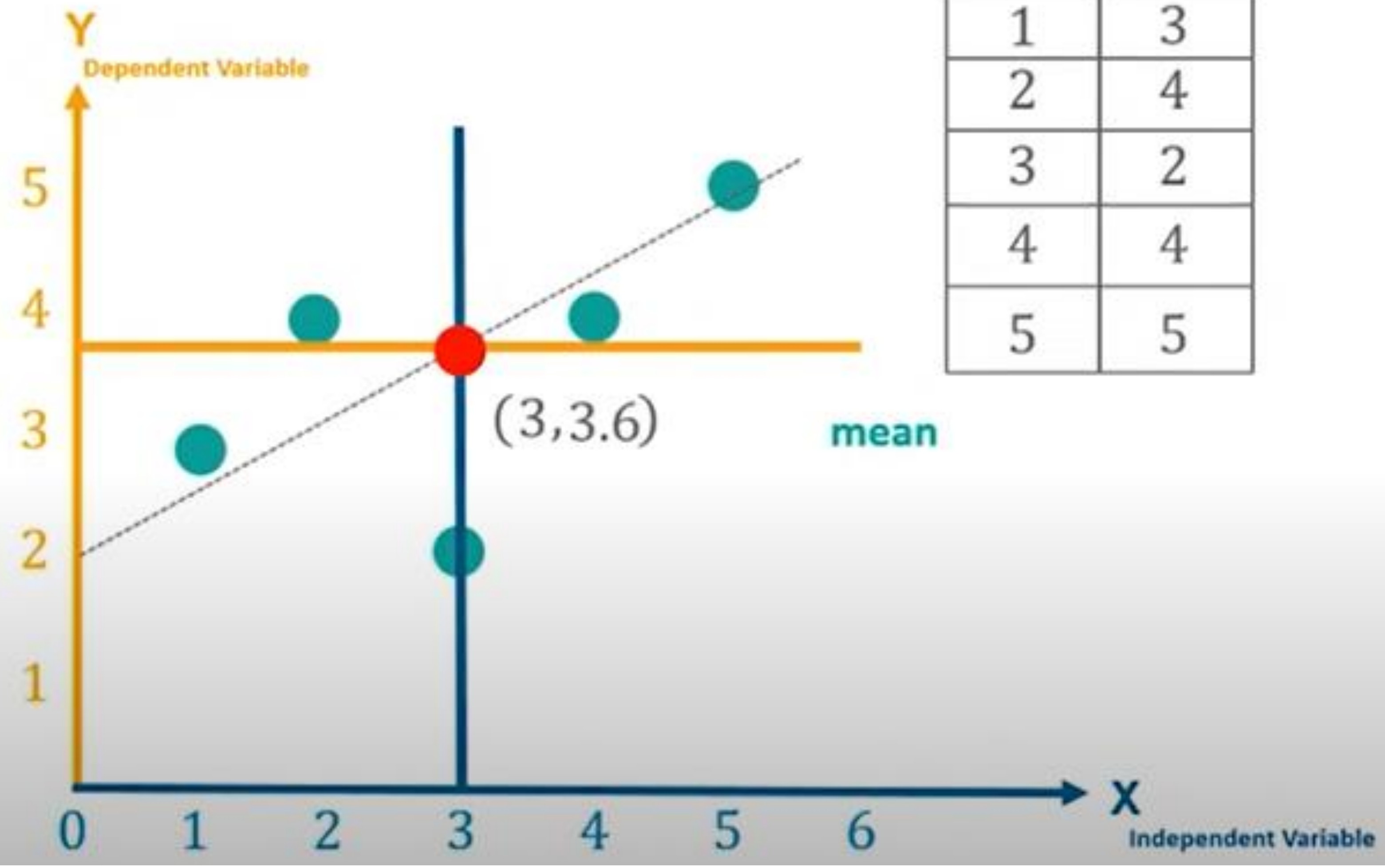
x	y
1	3
2	4
3	2
4	4
5	5



x	y
1	3
2	4
3	2
4	4
5	5

mean: $\dot{x}$ 3 3.6 **mean : $\dot{y}$**

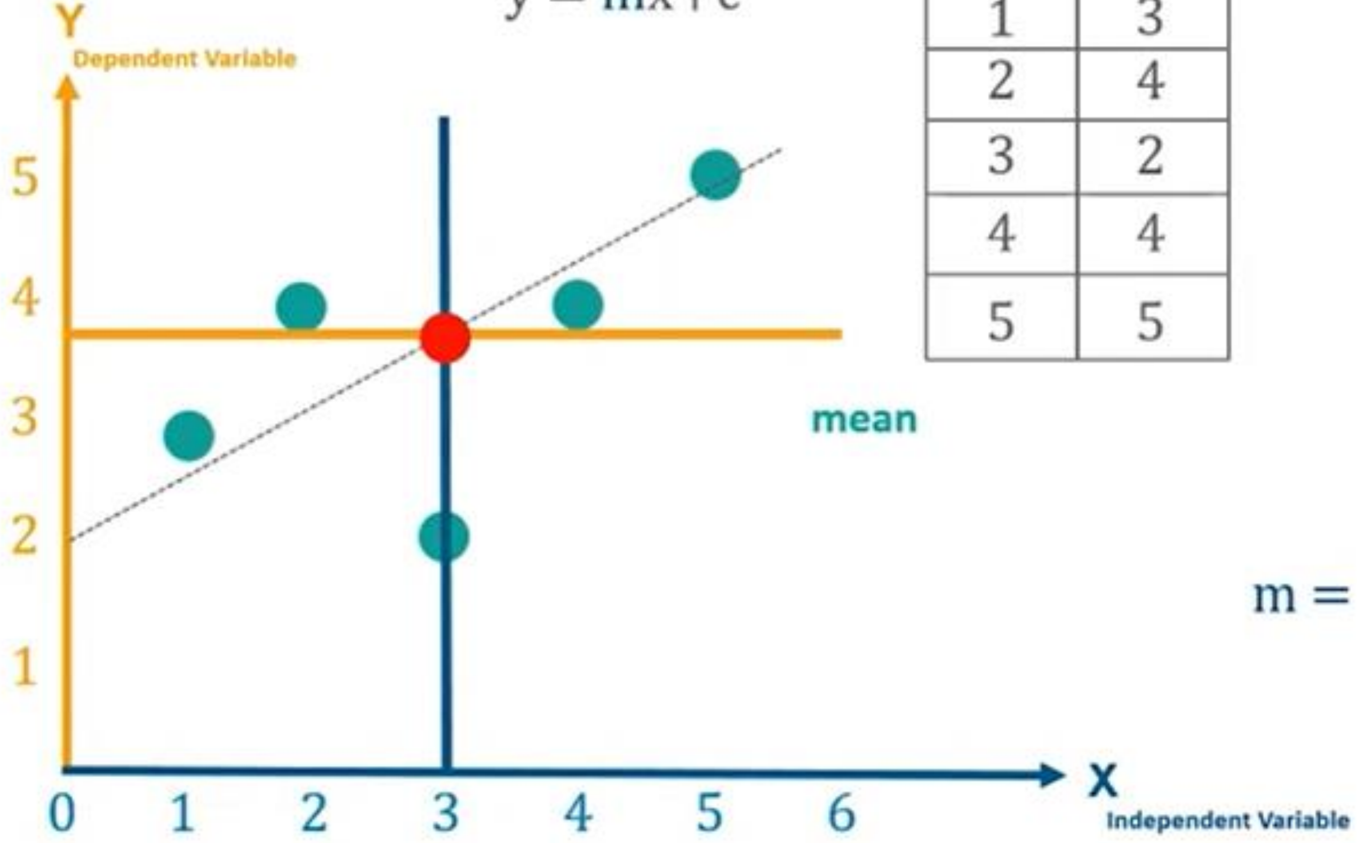
Understanding Linear Regression Algorithm



x	y
1	3
2	4
3	2
4	4
5	5

Understanding Linear Regression Algorithm

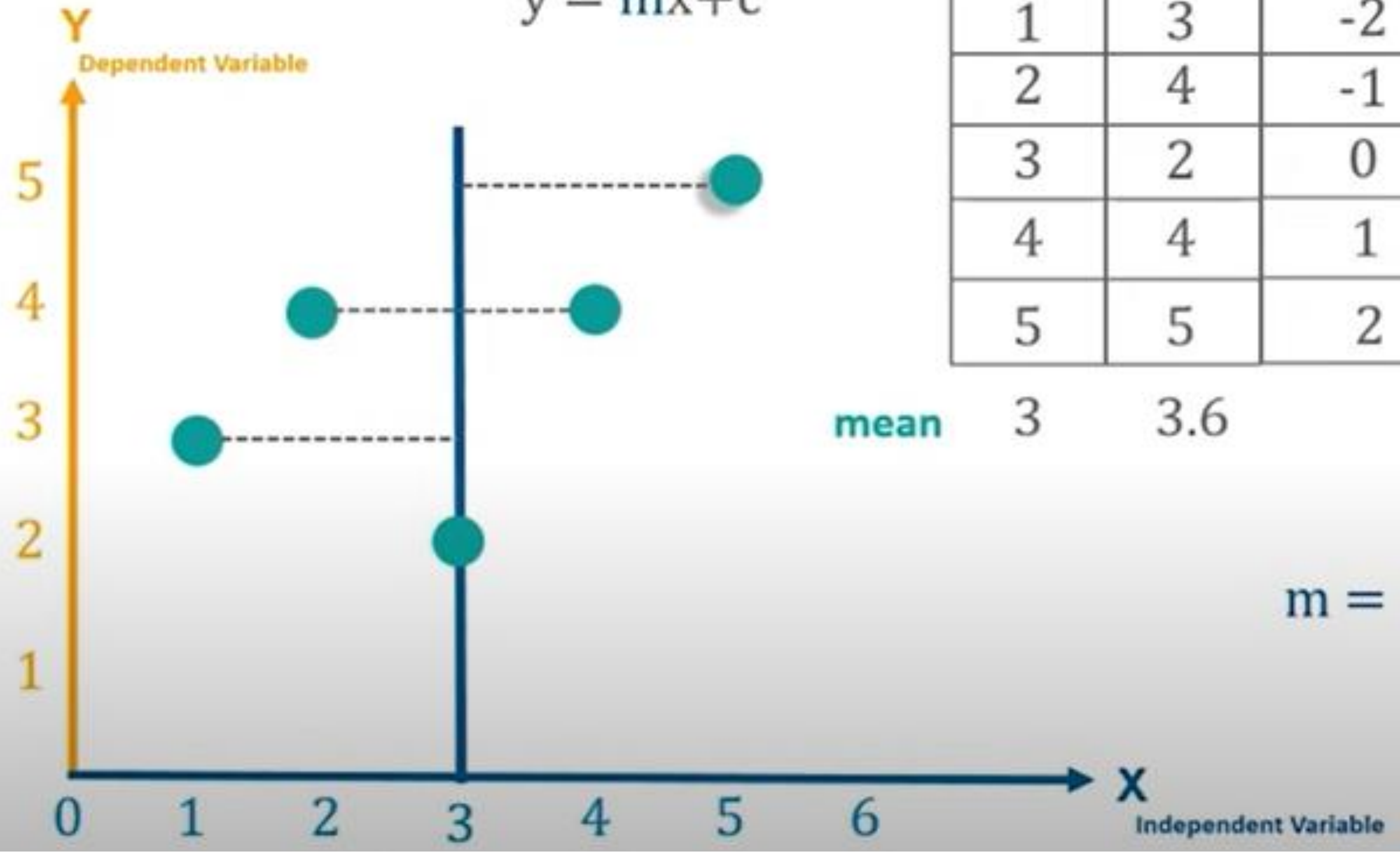
$$y = mx + c$$



x	y
1	3
2	4
3	2
4	4
5	5

$$m = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

Understanding Linear Regression Algorithm

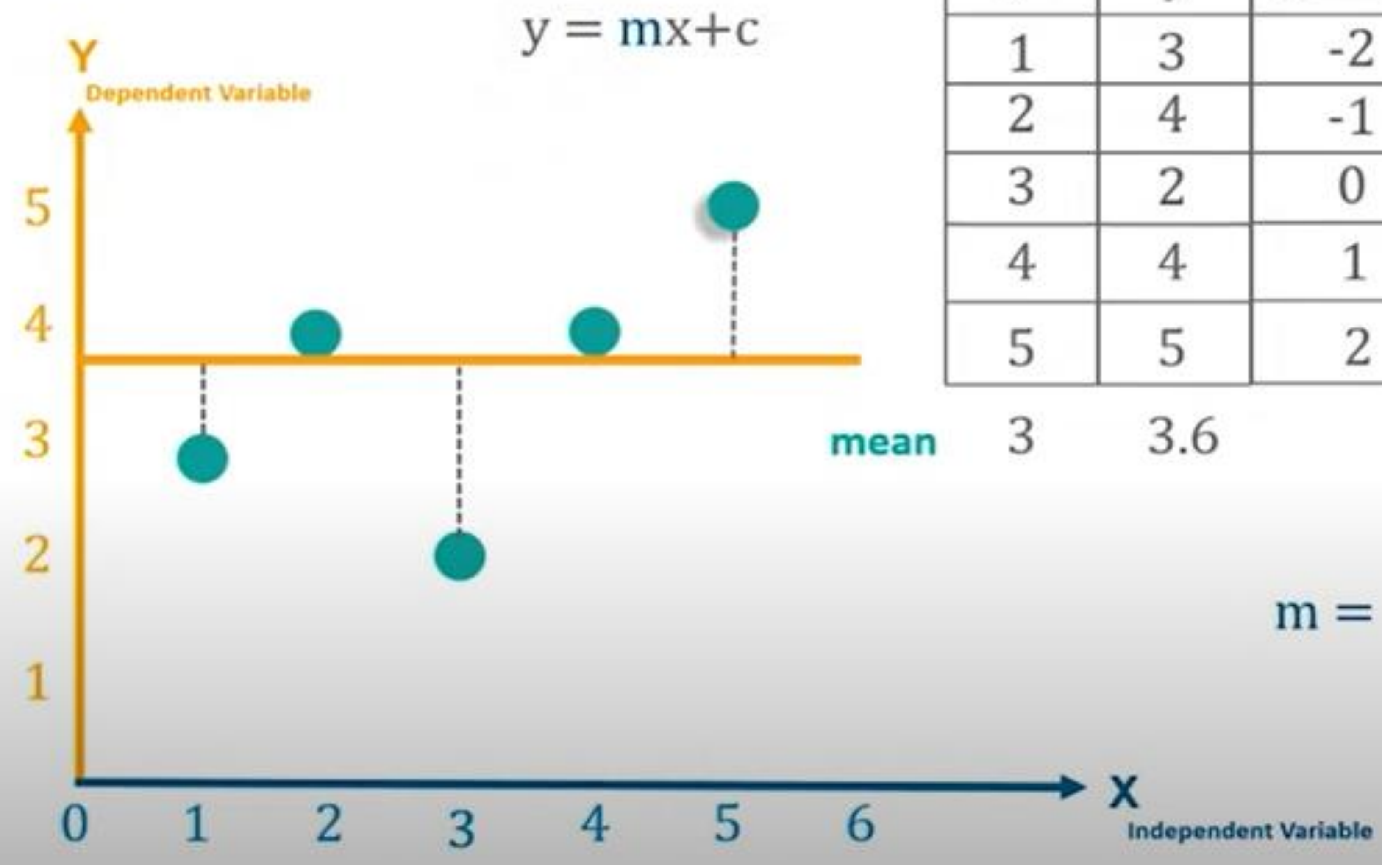


$$y = mx + c$$

x	y	$x - \bar{x}$
1	3	-2
2	4	-1
3	2	0
4	4	1
5	5	2
mean 3	3.6	

$$m = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

Understanding Linear Regression Algorithm

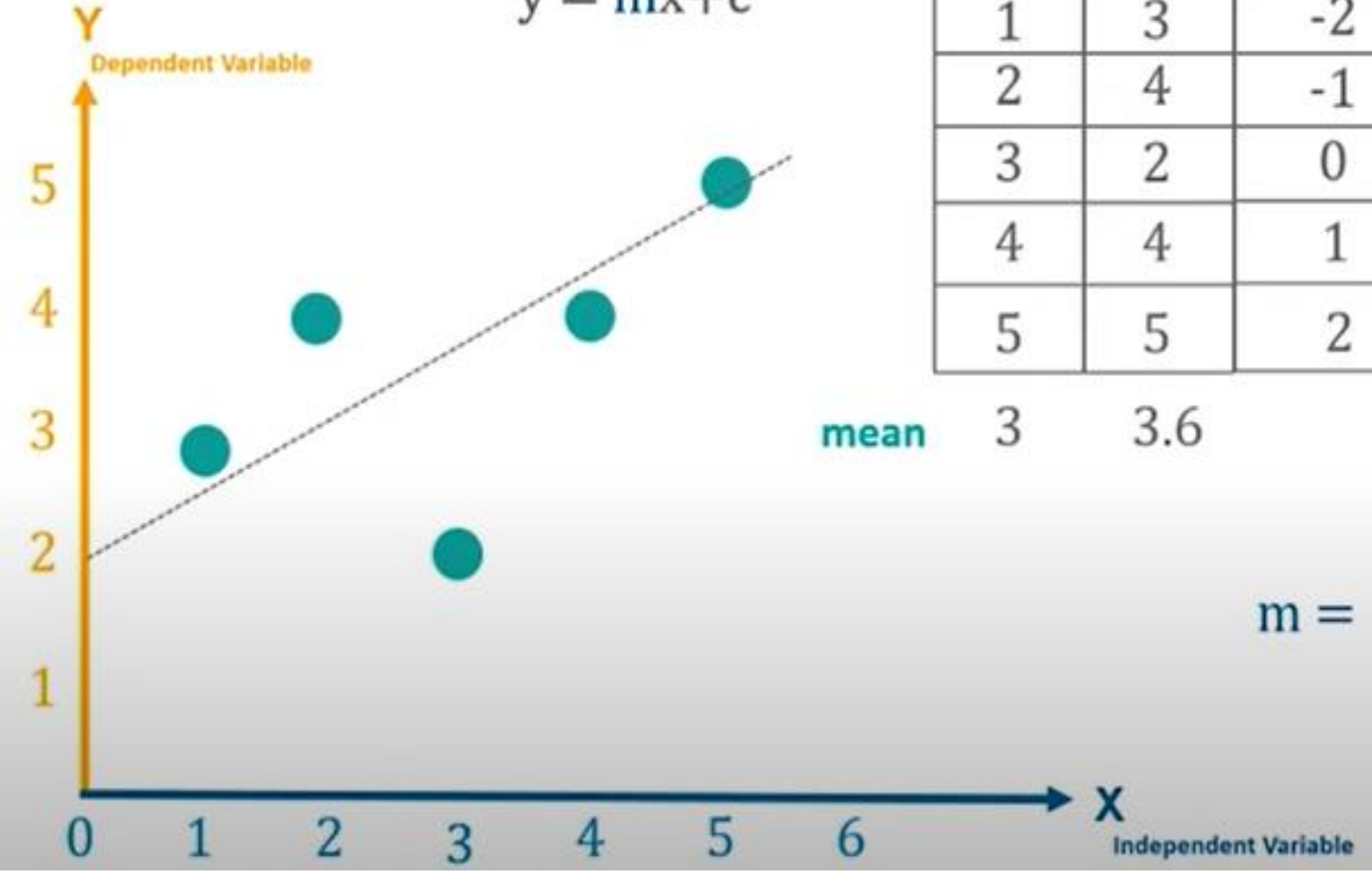


x	y	$x-\bar{x}$	$y-\bar{y}$
1	3	-2	-0.6
2	4	-1	0.4
3	2	0	-1.6
4	4	1	0.4
5	5	2	1.4
mean 3	3.6		

$$m = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

Understanding Linear Regression Algorithm

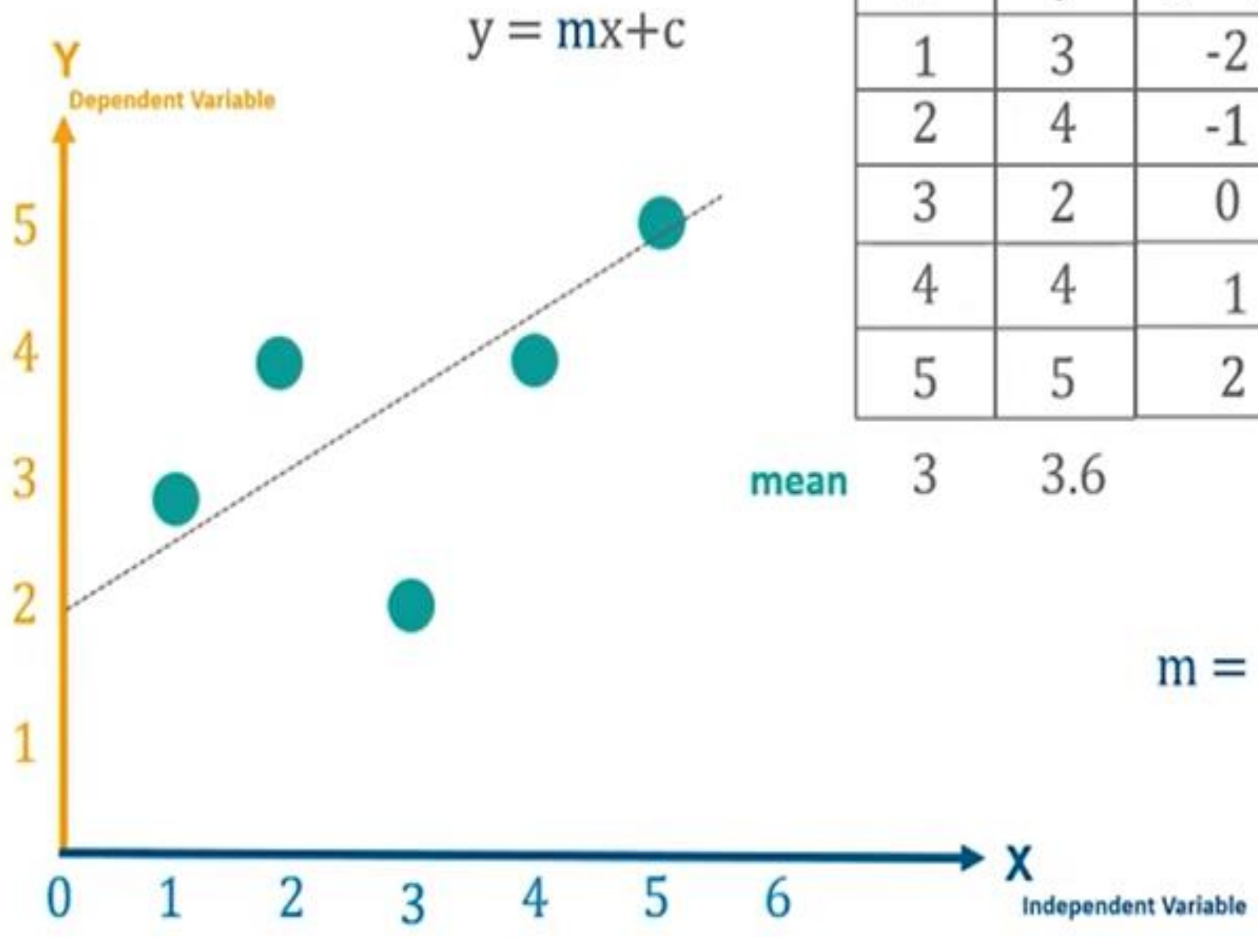
$$y = mx + c$$



	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$
	1	3	-2	-0.6	4
	2	4	-1	0.4	1
	3	2	0	-1.6	0
	4	4	1	0.4	1
	5	5	2	1.4	4
mean	3	3.6			

$$m = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

Understanding Linear Regression Algorithm

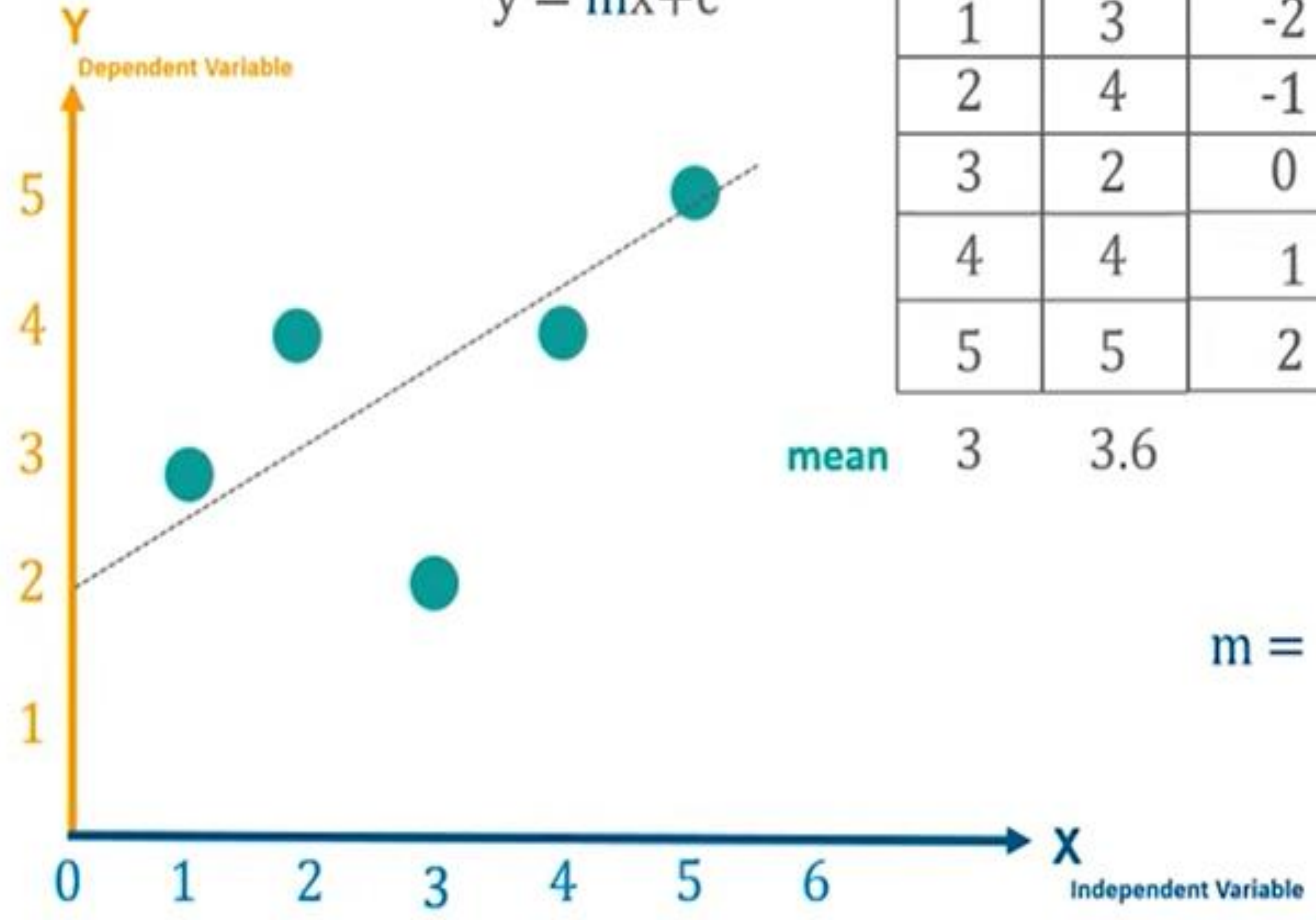


	x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(x - \bar{x})(y - \bar{y})$
	1	3	-2	-0.6	4	1.2
	2	4	-1	0.4	1	-0.4
	3	2	0	-1.6	0	0
	4	4	1	0.4	1	0.4
	5	5	2	1.4	4	2.8
mean	3	3.6				

$$m = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

Understanding Linear Regression Algorithm

$$y = mx + c$$

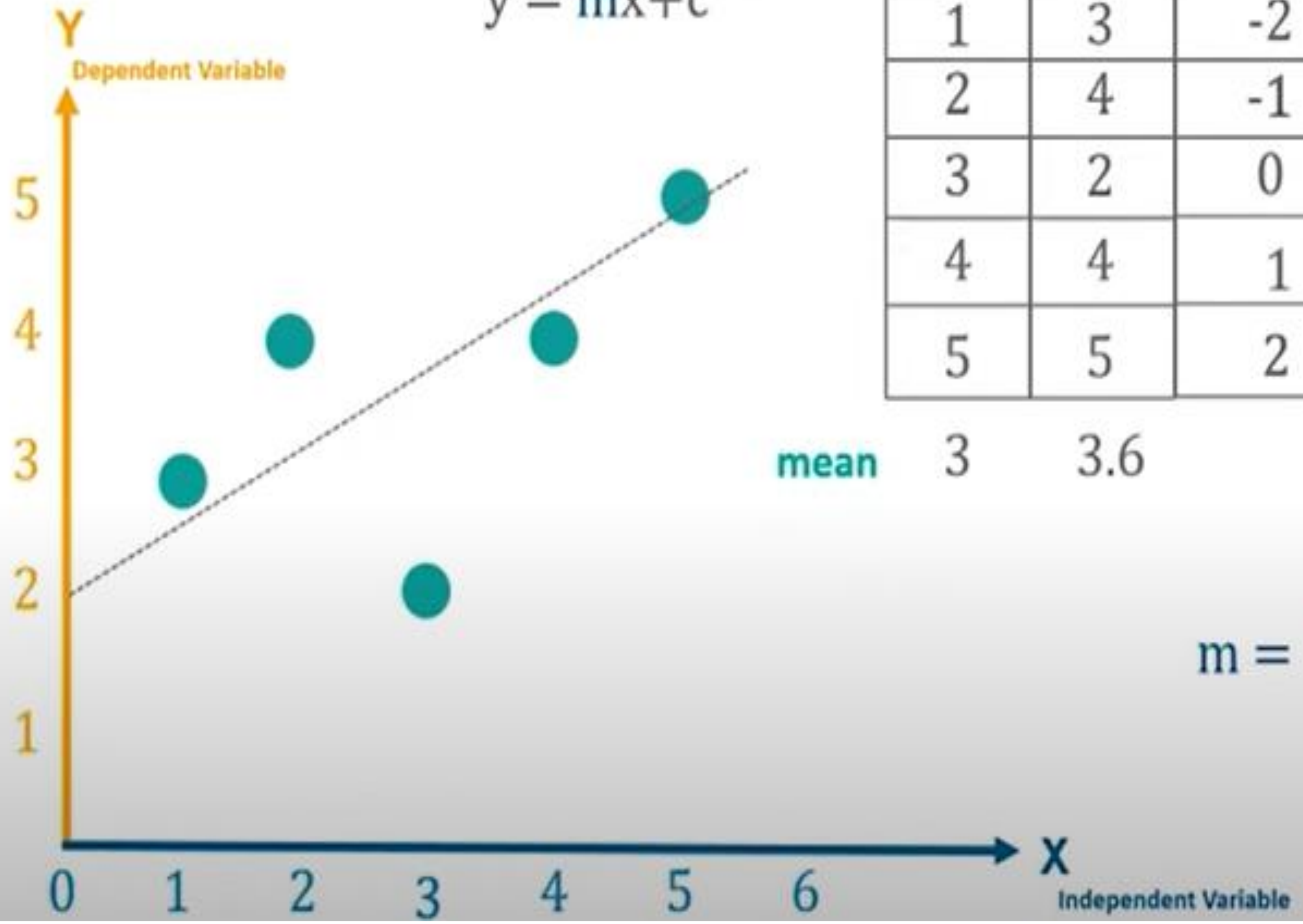


	x	y	$x-\dot{x}$	$y-\dot{y}$	$(x-\dot{x})^2$	$(x-\dot{x})(y-\dot{y})$
	1	3	-2	-0.6	4	1.2
	2	4	-1	0.4	1	-0.4
	3	2	0	-1.6	0	0
	4	4	1	0.4	1	0.4
	5	5	2	1.4	4	2.8
mean	3	3.6			$\Sigma = 10$	$\Sigma = 4$

$$m = \frac{\Sigma\,(x-\dot{x})(y-\dot{y})}{\Sigma\,(x-\dot{x})^2}$$

Understanding Linear Regression Algorithm

$$y = mx + c$$



x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$
1	3	-2	-0.6	4	1.2
2	4	-1	0.4	1	-0.4
3	2	0	-1.6	0	0
4	4	1	0.4	1	0.4
5	5	2	1.4	4	2.8
mean 3	3.6			$\Sigma = 10$	$\Sigma = 4$

$$m = \frac{\Sigma\,(x-\bar{x})(y-\bar{y})}{\Sigma\,(x-\bar{x})^2} = \frac{4}{10}$$

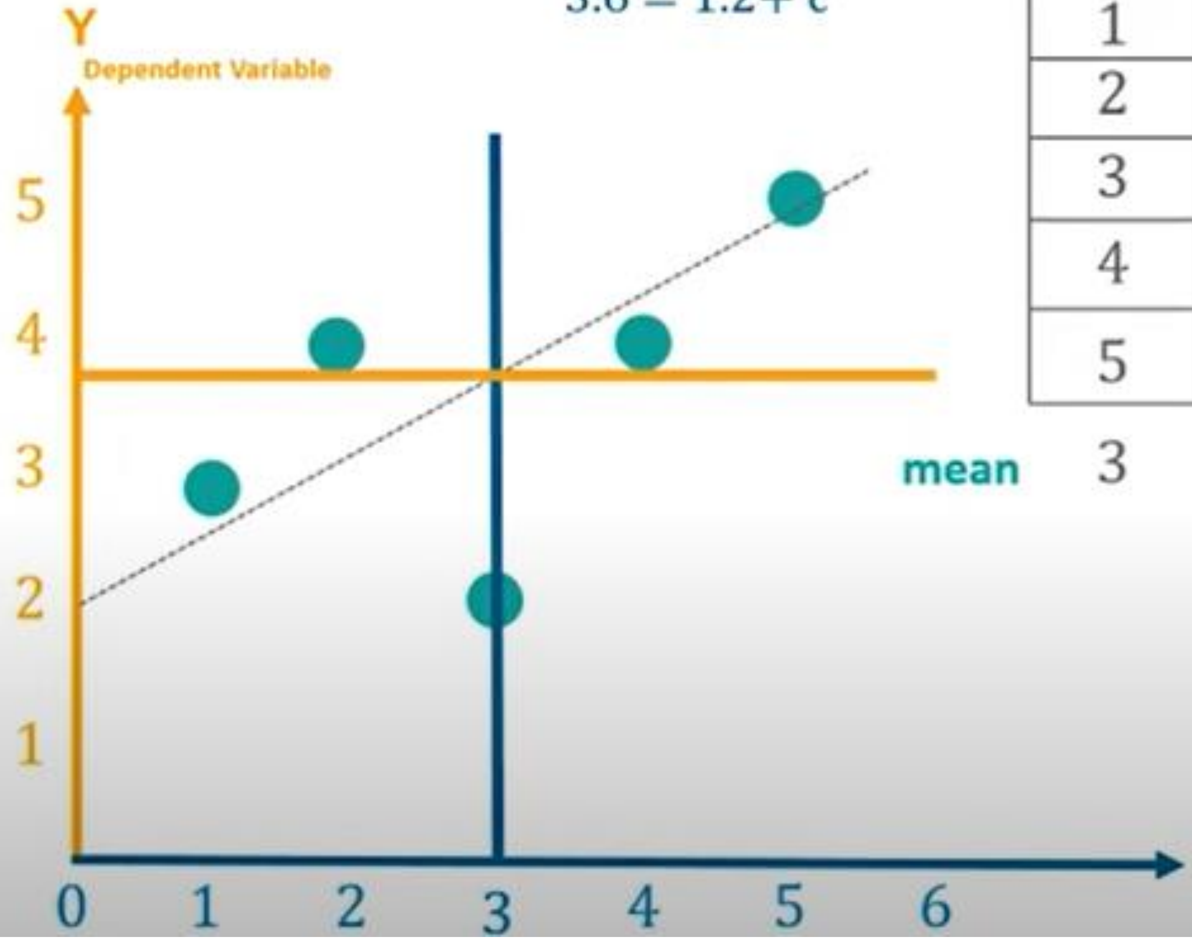
$$y = mx + c$$

$$3.6 = 0.4 \times 3 + c$$

Understanding Linear Regression Algorithm

$$y = mx + c$$

$$3.6 = 1.2 + c$$



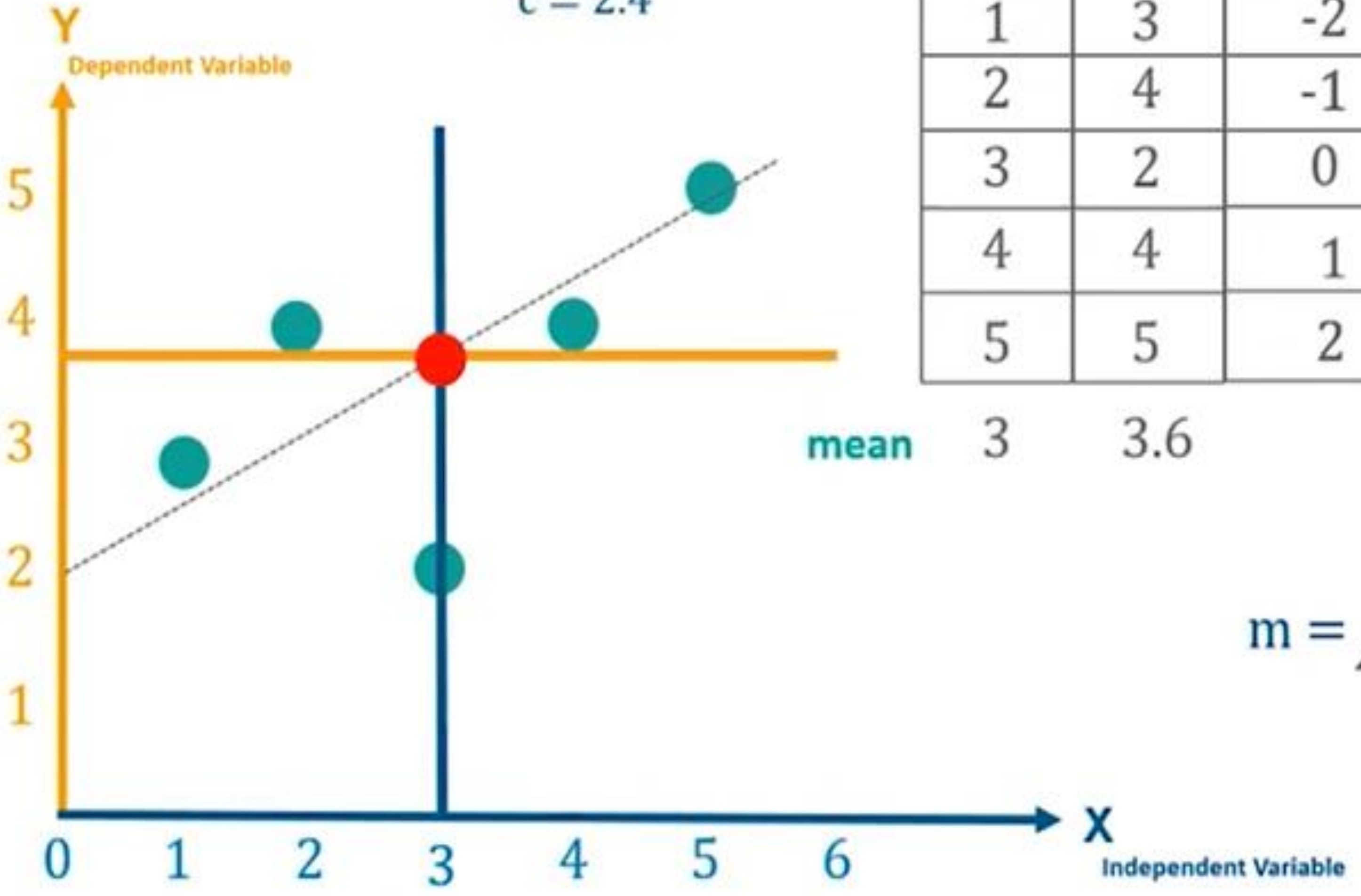
	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$
	1	3	-2	-0.6	4	1.2
	2	4	-1	0.4	1	-0.4
	3	2	0	-1.6	0	0
	4	4	1	0.4	1	0.4
	5	5	2	1.4	4	2.8
mean	3	3.6			$\Sigma = 10$	$\Sigma = 4$

$$m = \sum \frac{(x-\bar{x})(y-\bar{y})}{(x-\bar{x})^2} = \frac{4}{10}$$

Understanding Linear Regression Algorithm

$y = mx + c$

$c = 2.4$



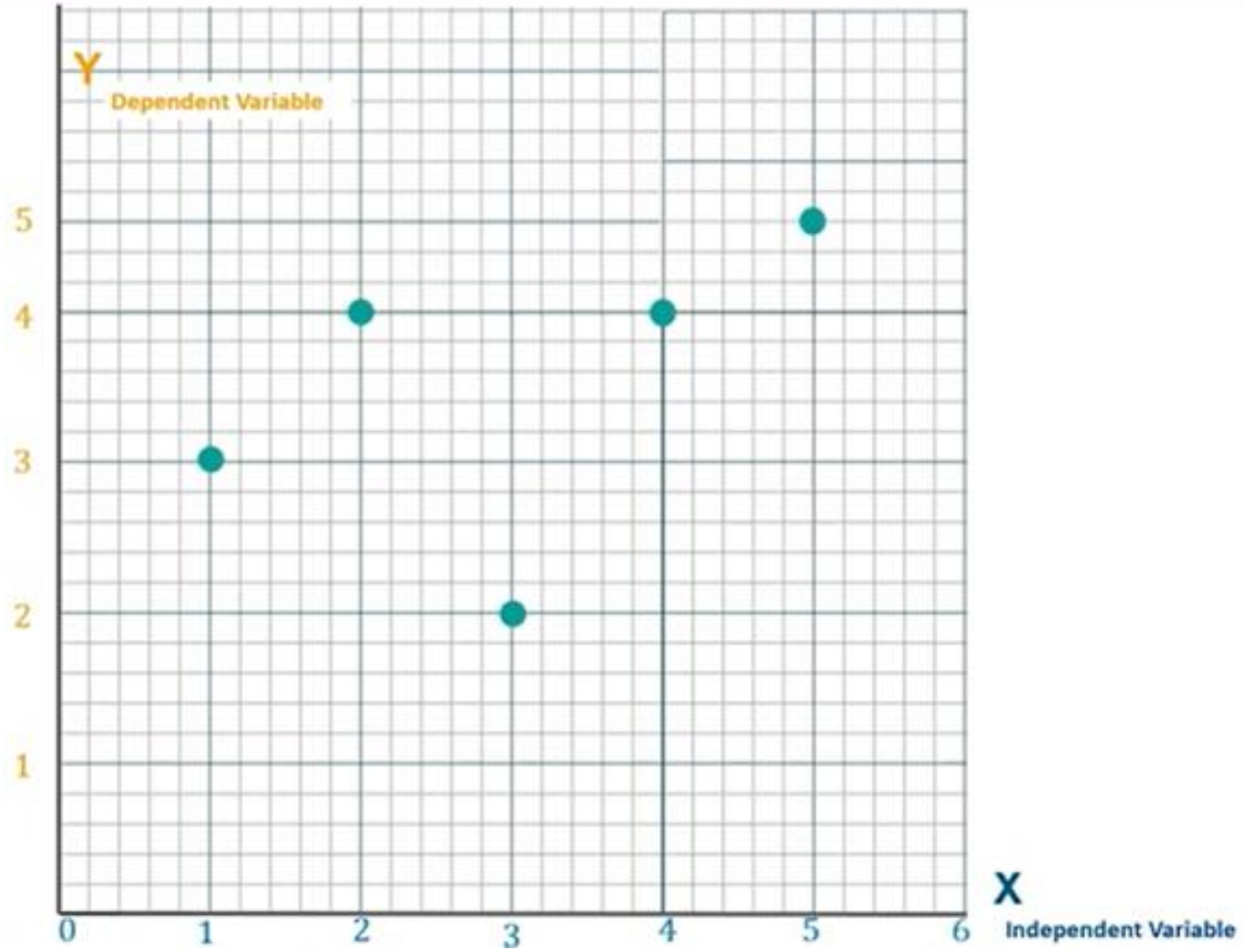
	x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(x - \bar{x})(y - \bar{y})$
	1	3	-2	-0.6	4	1.2
	2	4	-1	0.4	1	-0.4
	3	2	0	-1.6	0	0
	4	4	1	0.4	1	0.4
	5	5	2	1.4	4	2.8
mean	3	3.6			$\Sigma = 10$	$\Sigma = 4$

$$m = \sum \frac{(x - \bar{x})(y - \bar{y})}{(x - \bar{x})^2} = \frac{4}{10}$$

Thus the regression line is:

$$m = 0.4$$
$$c = 2.4$$
$$y = 0.4x + 2.4$$

Mean Square Error

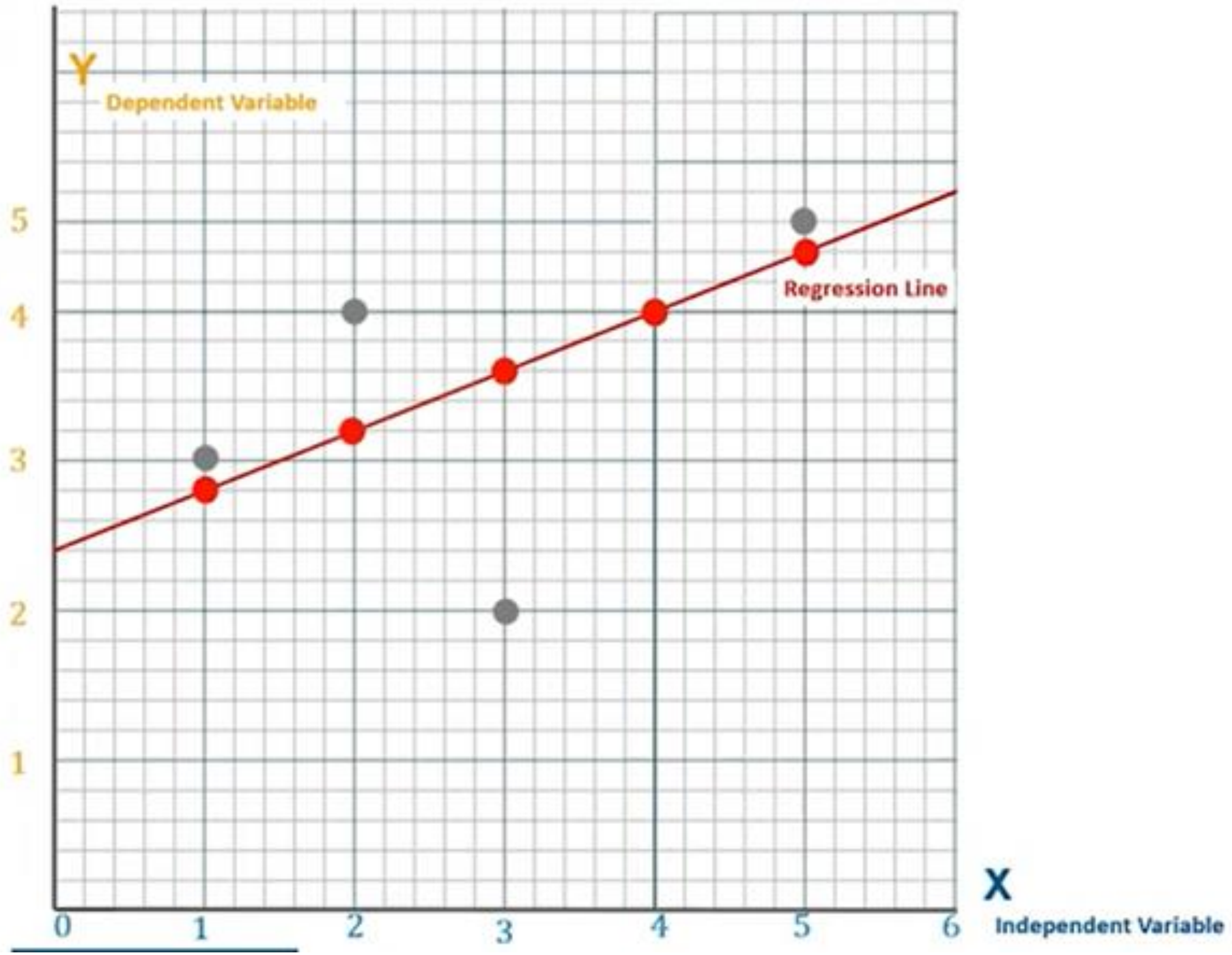


$m = 0.4$
$c = 2.4$
$y = 0.4x + 2.4$

For given $m = 0.4$ & $c = 2.4$, lets predict values for y for $x = \{1,2,3,4,5\}$

$y = 0.4 \times 1 + 2.4 = 2.8$
$y = 0.4 \times 2 + 2.4 = 3.2$
$y = 0.4 \times 3 + 2.4 = 3.6$
$y = 0.4 \times 4 + 2.4 = 4.0$
$y = 0.4 \times 5 + 2.4 = 4.4$

Mean Square Error

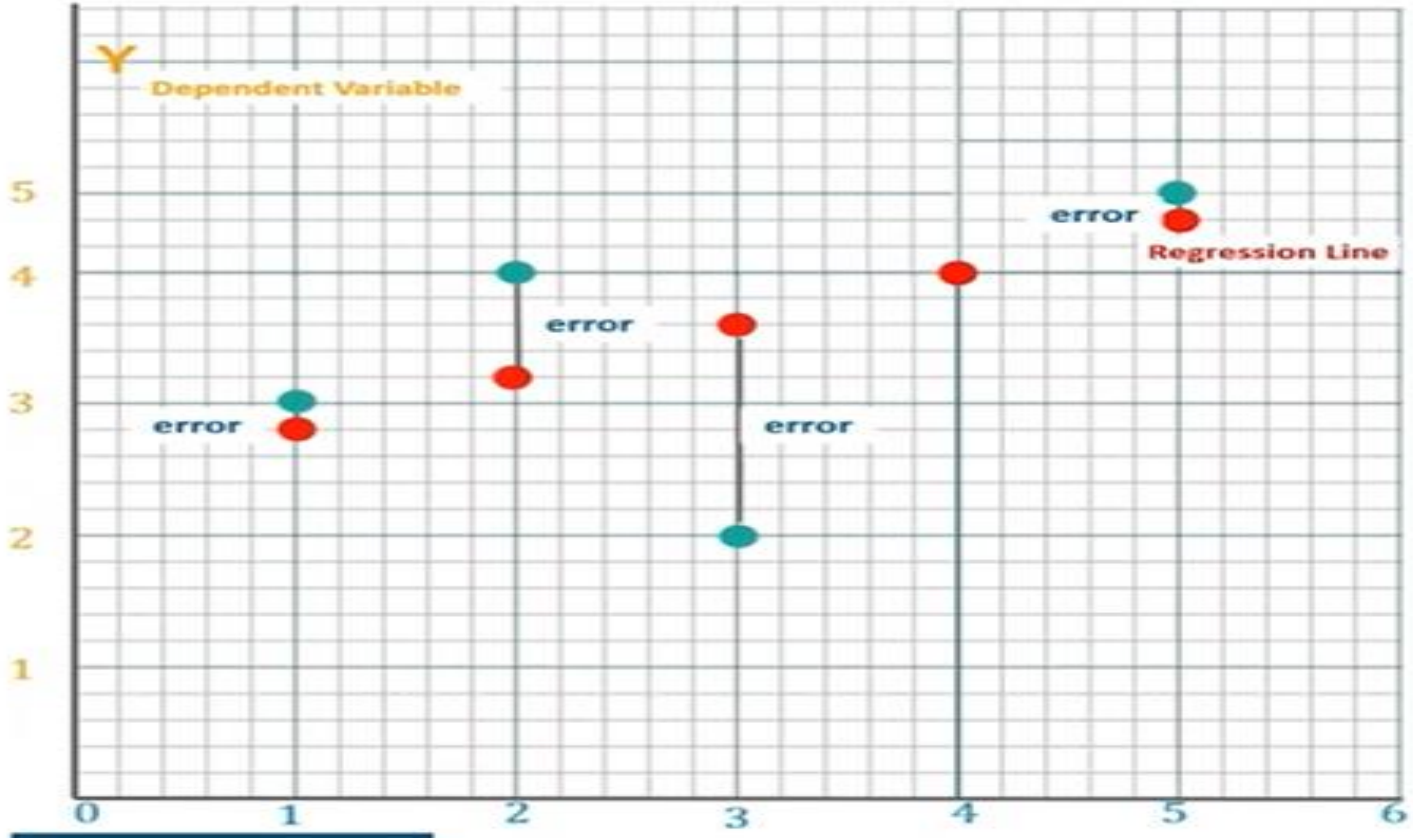


$m = 0.4$
$c = 2.4$
$y = 0.4x + 2.4$

For given m = 0.4 & c = 2.4, lets predict values for y for x = {1,2,3,4,5}

$y = 0.4 \times 1 + 2.4 = 2.8$
$y = 0.4 \times 2 + 2.4 = 3.2$
$y = 0.4 \times 3 + 2.4 = 3.6$
$y = 0.4 \times 4 + 2.4 = 4.0$
$y = 0.4 \times 5 + 2.4 = 4.4$

Now job is to find the distance between actual & predicted value



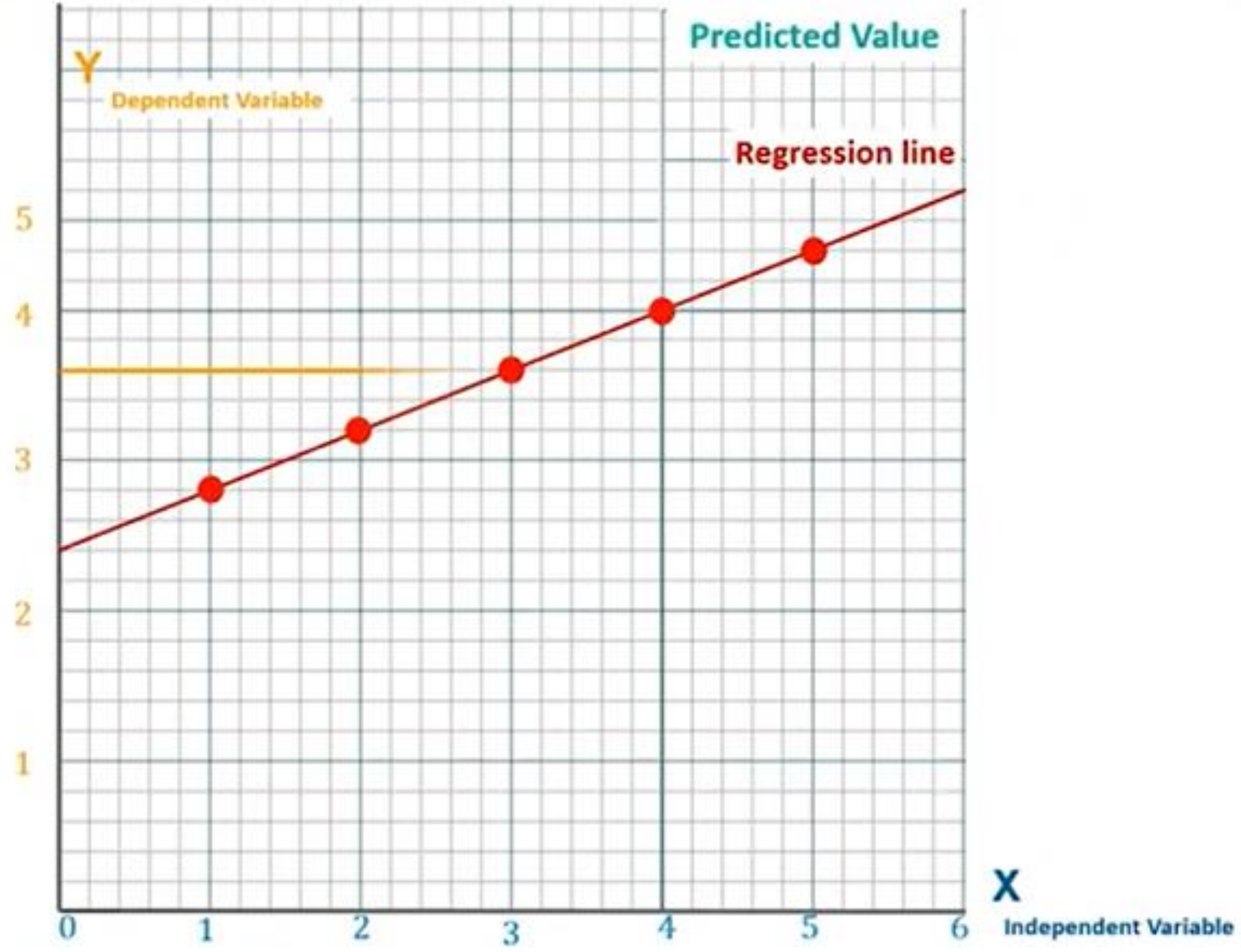
Now the job is to reduce the error between actual and predicted value.

The line with the least error is the line of linear regression (minimization using gradient descent algorithm)

To check the goodness of fit

R^2 Method

Calculation of R^2



x	y_p
1	2.8
2	3.2
3	3.6
4	4.0
5	4.4

Calculation of R^2

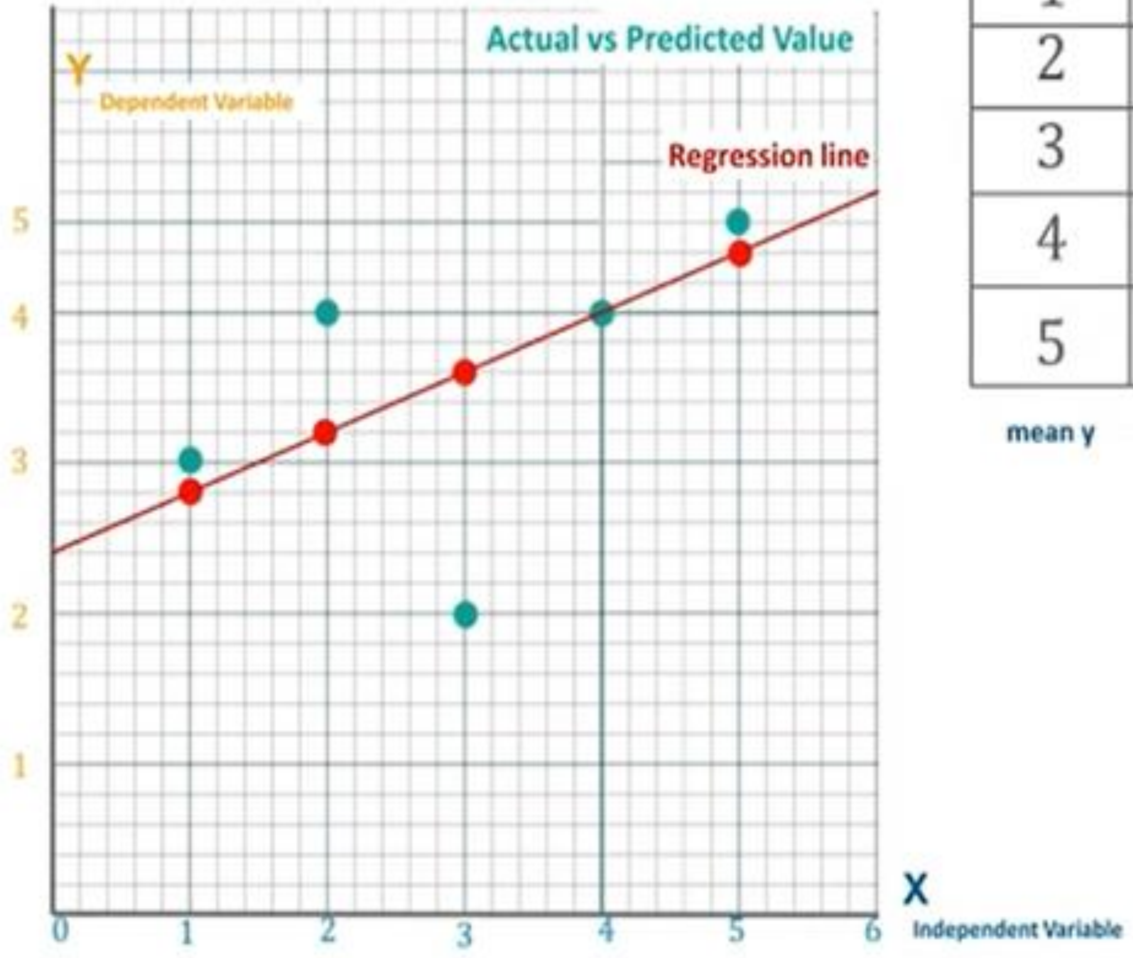
Distance actual - mean

vs

Distance predicted - mean

This is nothing but $R^2 = \dfrac{\sum (y_p - \bar{y})^2}{\sum (y - \bar{y})^2}$

Calculation of R^2

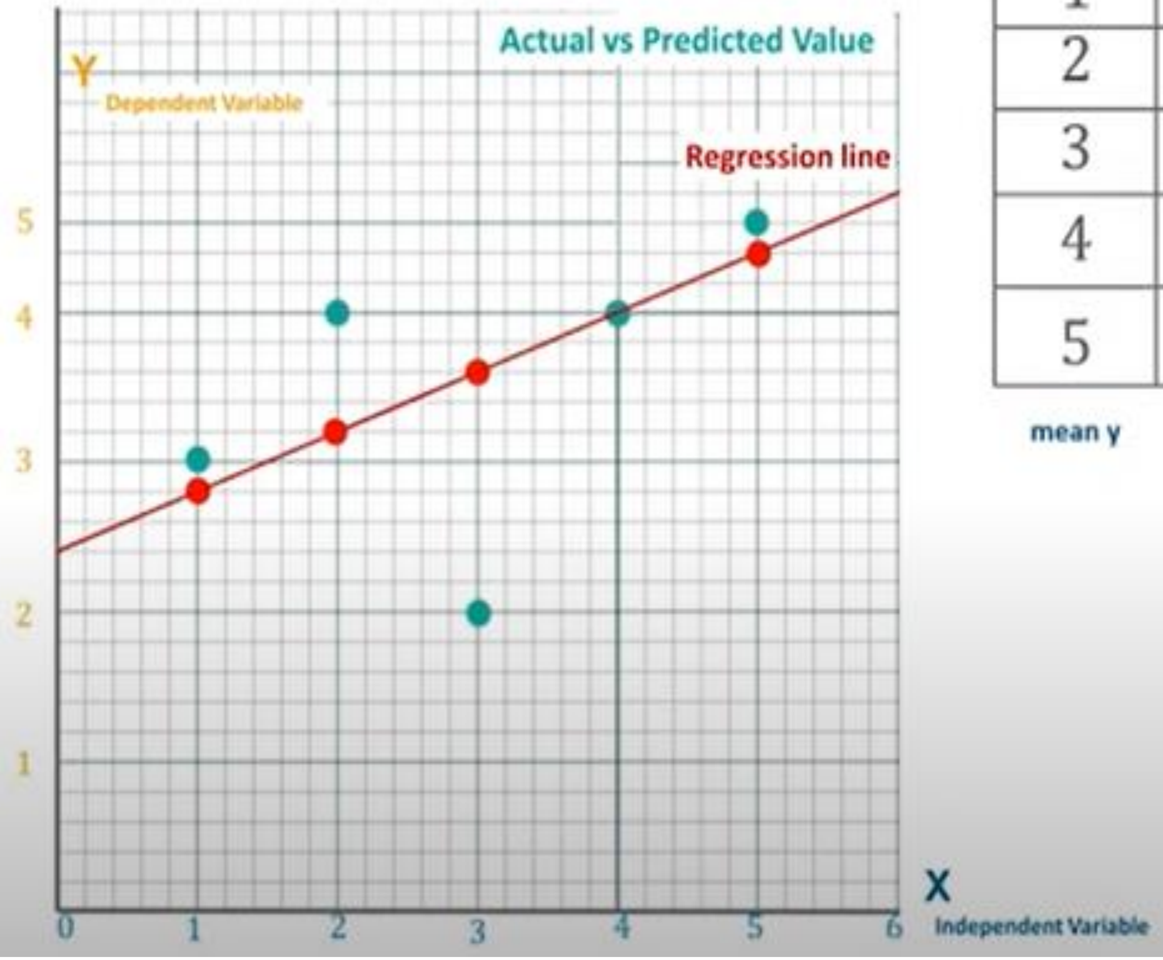


x	y	$y - \bar{y}$
1	3	– 0.6
2	4	0.4
3	2	–1.6
4	4	0.4
5	5	1.4

mean y 3.6

$$R^2 = \frac{\sum (y_p - \bar{y})^2}{\sum (y - \bar{y})^2}$$

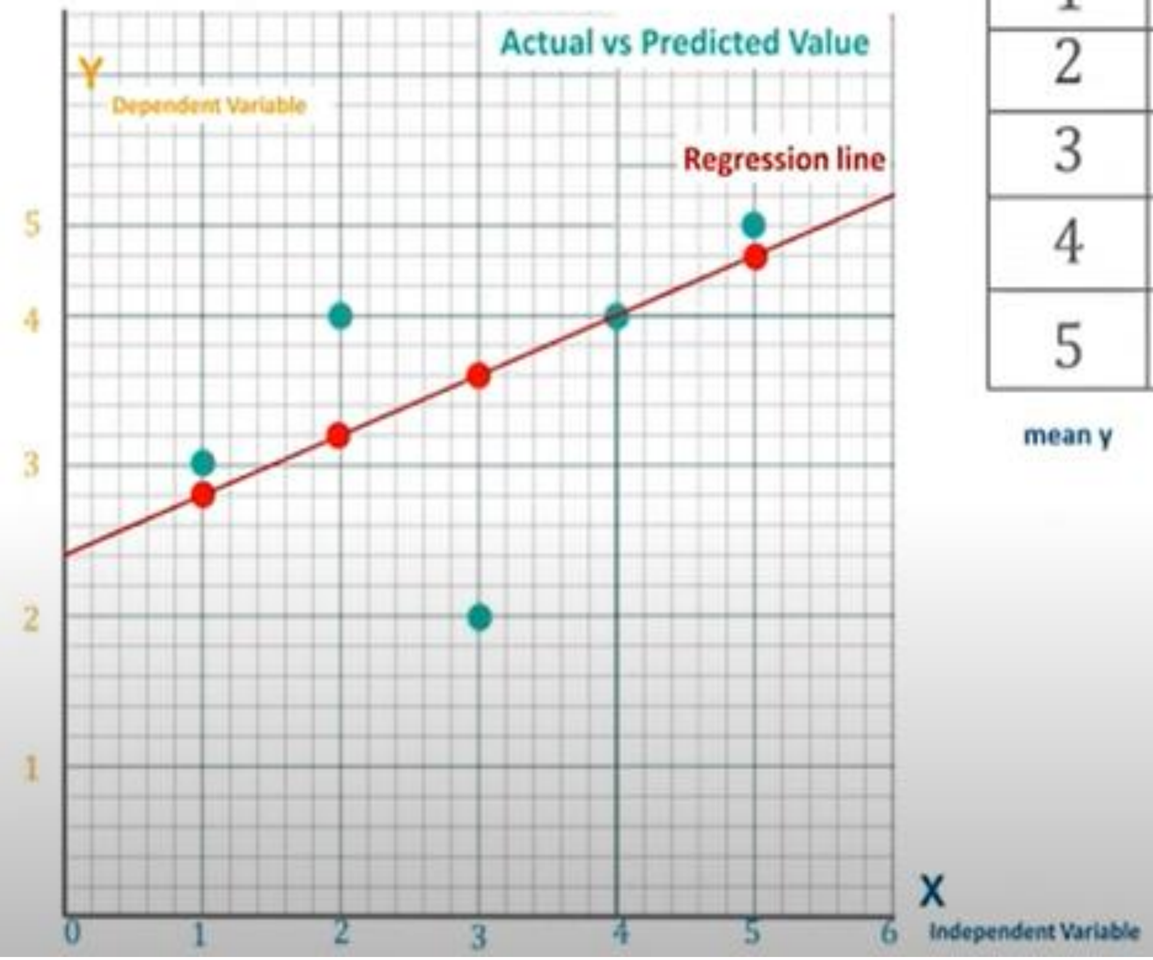
Calculation of R^2



x	y	$y - \bar{y}$	$(y - \bar{y})^2$	y_p	$(y_p - \bar{y})$
1	3	– 0.6	0.36		
2	4	0.4	0.16		
3	2	–1.6	2.56		
4	4	0.4	0.16		
5	5	1.4	1.96		

mean y 3.6

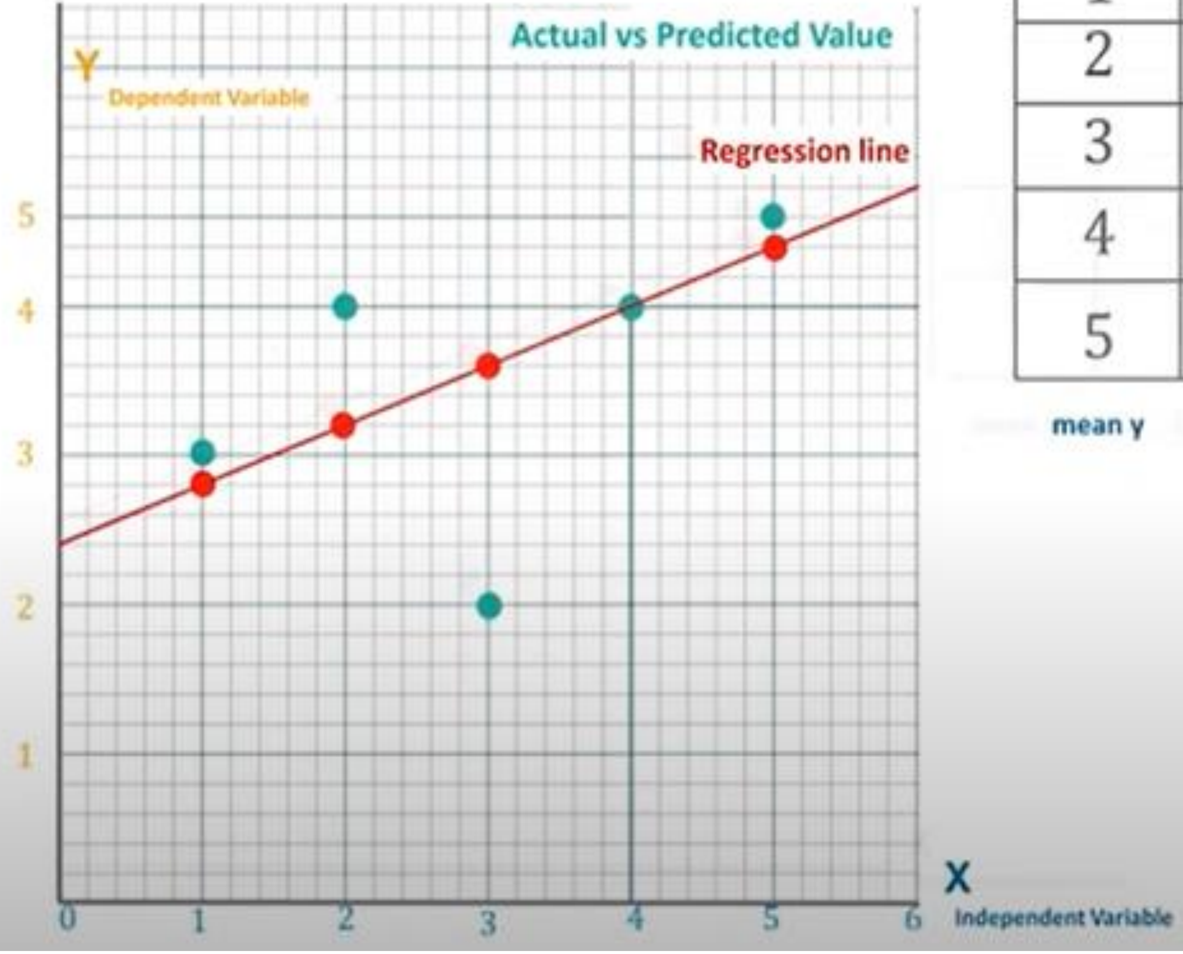
Calculation of R^2



x	y	$y - \bar{y}$	$(y - \bar{y})^2$	y_p	$(y_p - \bar{y})$
1	3	– 0.6	0.36	2.8	
2	4	0.4	0.16	3.2	
3	2	–1.6	2.56	3.6	
4	4	0.4	0.16	4.0	
5	5	1.4	1.96	4.4	

mean y 3.6

Calculation of R^2

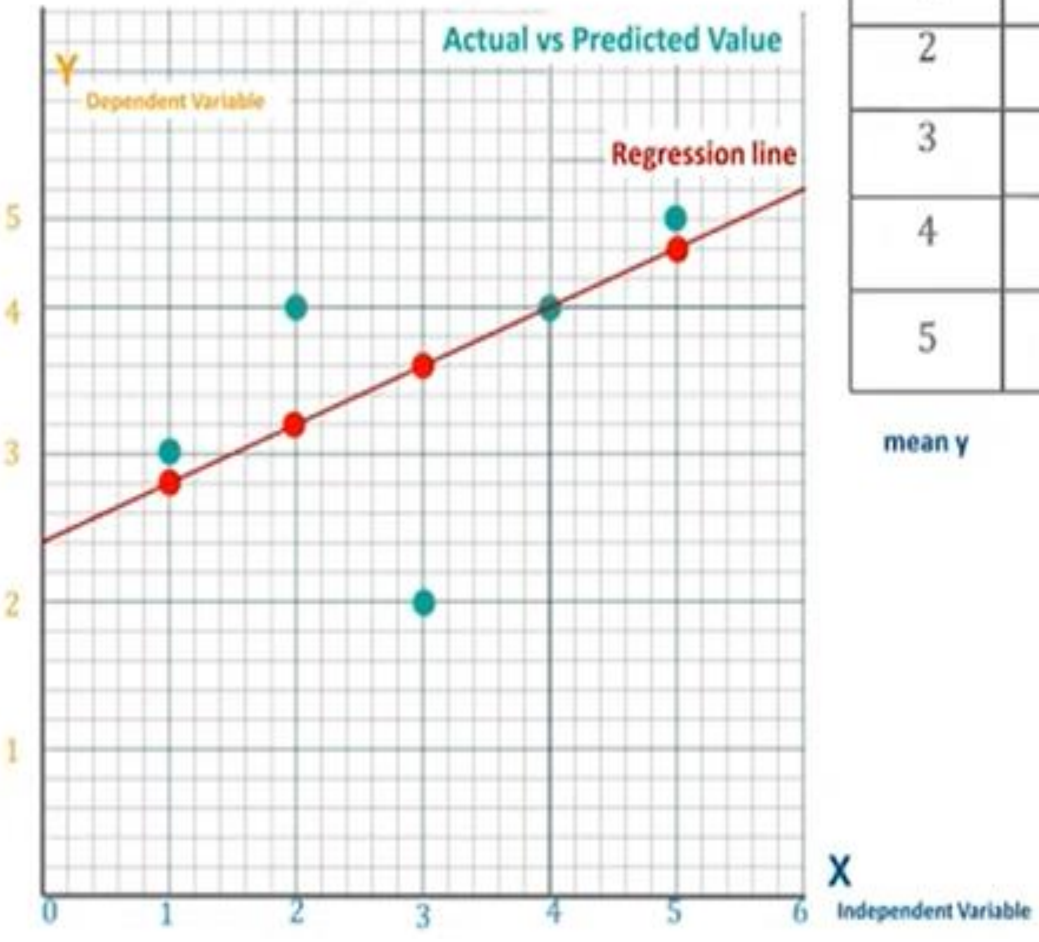


x	y	$y - \bar{y}$	$(y - \bar{y})^2$	y_p	$(y_p - \bar{y})$
1	3	– 0.6	0.36	2.8	-0.8
2	4	0.4	0.16	3.2	-0.4
3	2	−1.6	2.56	3.6	0
4	4	0.4	0.16	4.0	0.4
5	5	1.4	1.96	4.4	0.8

mean y 3.6

$$R^2 = \frac{\sum (y_p - \bar{y})^2}{\sum (y - \bar{y})^2}$$

Calculation of R^2

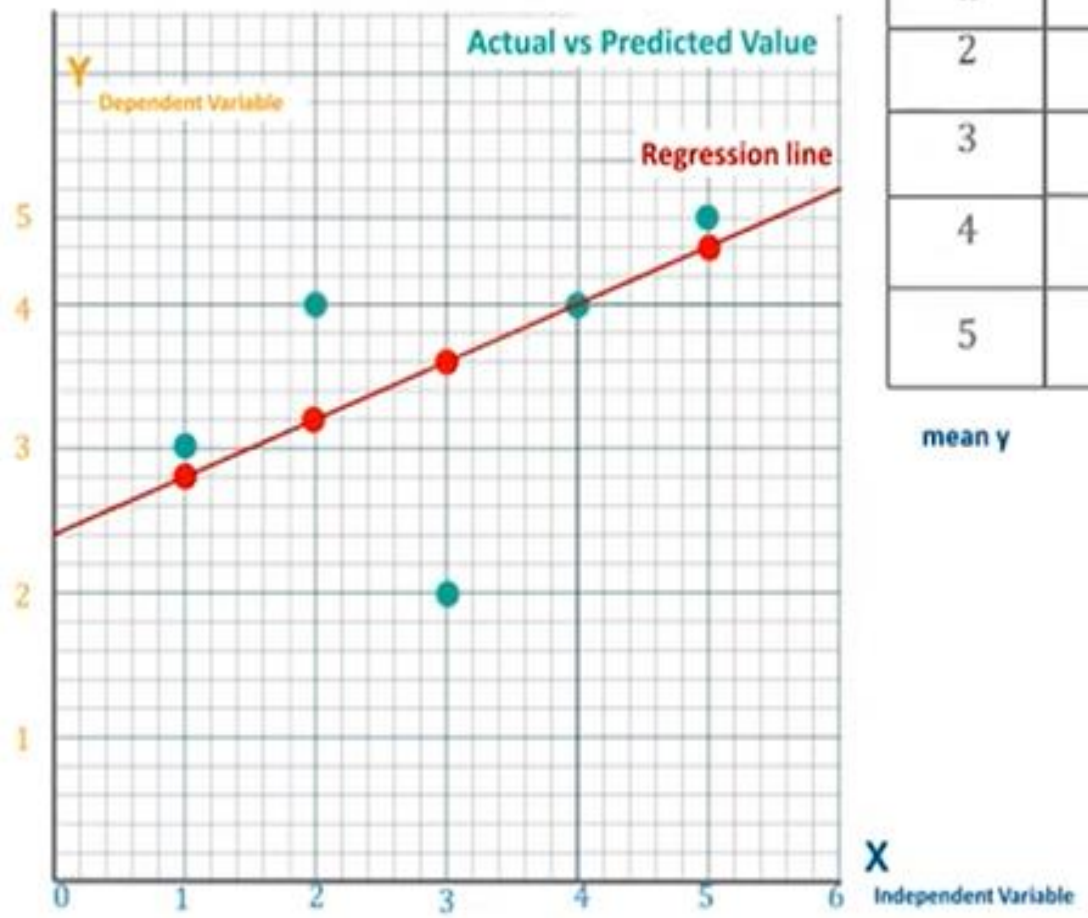


x	y	$y - y$	$(y - y)^2$	y_p	$(y_p - y)$	$(y_p - y)^2$
1	3	− 0.6	0.36	2.8	-0.8	0.64
2	4	0.4	0.16	3.2	-0.4	0.16
3	2	−1.6	2.56	3.6	0	0
4	4	0.4	0.16	4.0	0.4	0.16
5	5	1.4	1.96	4.4	0.8	0.64

mean y 3.6

$$\frac{\sum (y_p - \dot{y})^2}{\sum (y - \dot{y})^2}$$

Calculation of R^2

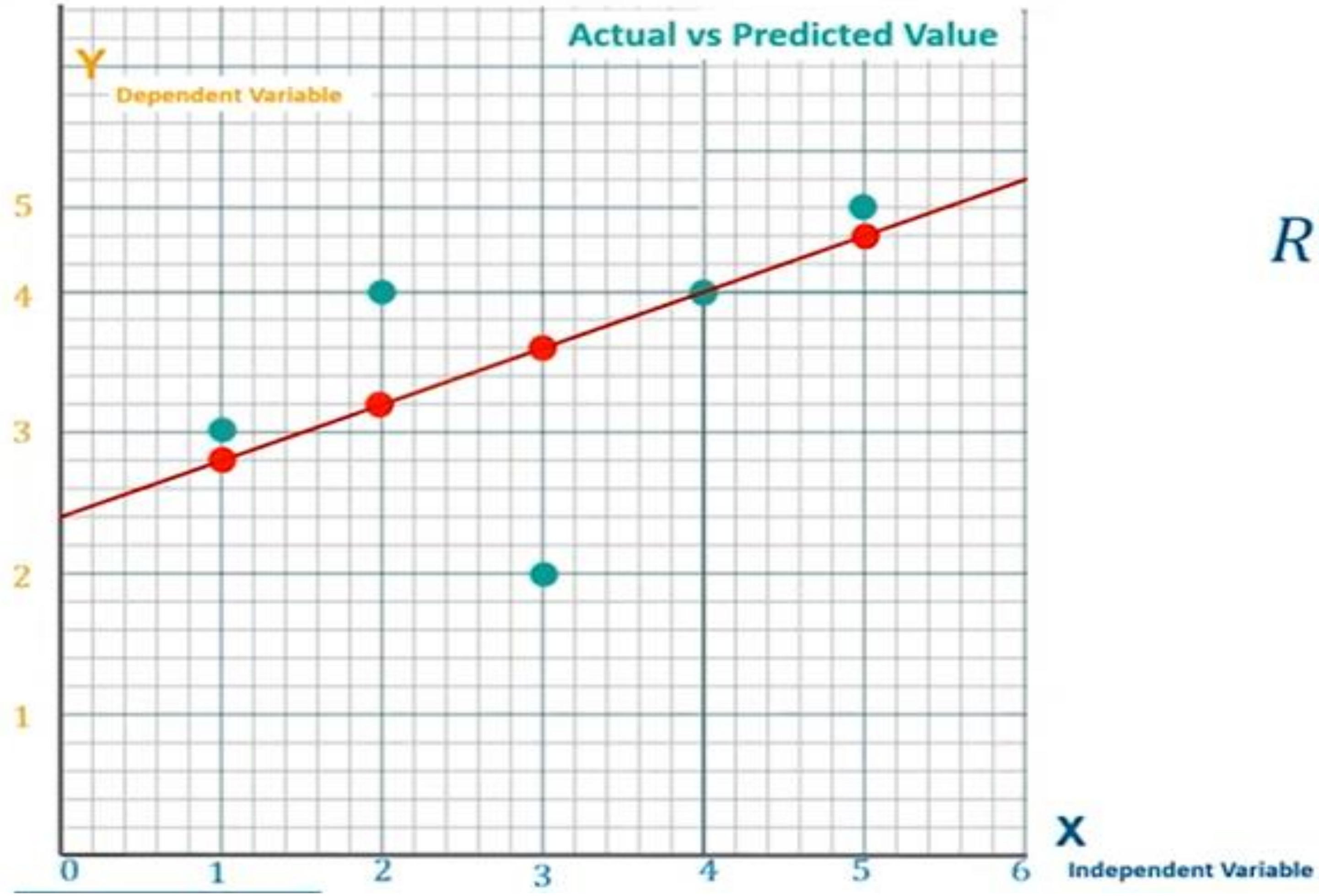


x	y	$y - y$	$(y - y)^2$	y_p	$(y_p - y)$	$(y_p - y)^2$
1	3	- 0.6	0.36	2.8	-0.8	0.64
2	4	0.4	0.16	3.2	-0.4	0.16
3	2	-1.6	2.56	3.6	0	0
4	4	0.4	0.16	4.0	0.4	0.16
5	5	1.4	1.96	4.4	0.8	0.64
mean y	3.6		Σ 5.2			Σ 1.6

$$R^2 = \frac{1.6}{5.2} = \frac{\Sigma\left((y_p - \bar{y}\right)^2}{\Sigma\,(y - \dot{y})^2}$$

0.3 is not a good fit .higher the value more fit the line

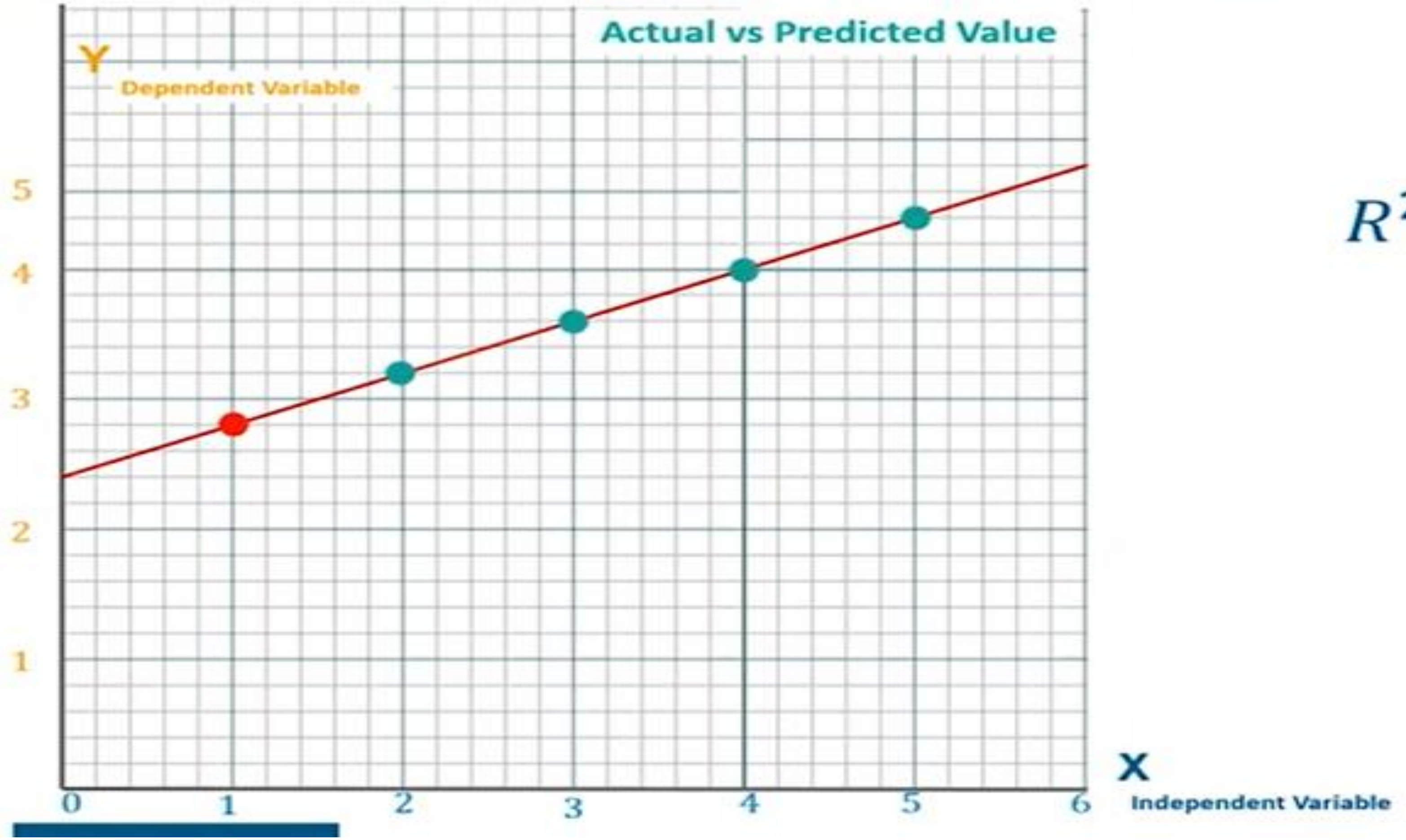
Calculation of R^2



$$R^2 \approx 0.3$$

Actual values lies on the regression line if R^2=1

Calculation of R^2



$R^2 \approx 1$

Try the same question with this method

$$a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$