Linear Regression

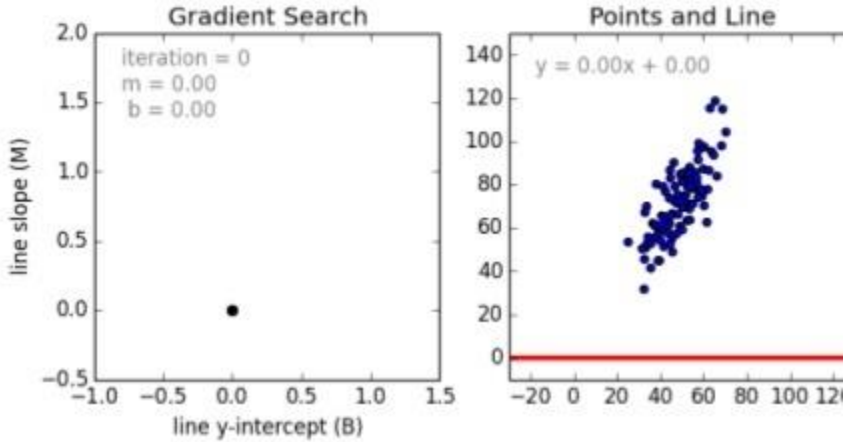
https://www.edureka.co/machine-learning-certification-training

edureka!

Linear Regression Algorithm

Agenda for Today's Session

- What is Regression?
- Regression Use-case
- Types of Regression – Linear vs Logistic Regression
- What is Linear Regression?
- Finding best fit regression line using Least Square Method
- Checking goodness of fit using R squared Method
- Implementation of Linear Regression using Python
  - Linear Regression Algorithm using Python from scratch
  - Linear Regression Algorithm using Python (scikit lib)



What is Regression?

"Regression analysis is a form of predictive modelli
technique which investigates the relationship betwe
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

Linear vs Logistic Regression

Basis	Linear Regression	Logistic Regression
Core Concept	The data is modelled using a straight line	The probability of some obtained event is represented as a linear function of a combinatio predictor variables.
Used with	Continuous Variable	Categorical Variable
Output/Prediction	Value of the variable	Probability of occurrenc event
Accuracy and Goodness of fit	measured by loss, R squared, Adjusted R squared etc.	Accuracy, Precision, Rec F1 score, ROC curve, Confusion Matrix, etc

What is Linear Regression?

"Linear Regression is a method to predict dependent variable (
based on values of independent variables (X). It can be used for
cases where we want to predict some continuous quantity."

Linear Regression

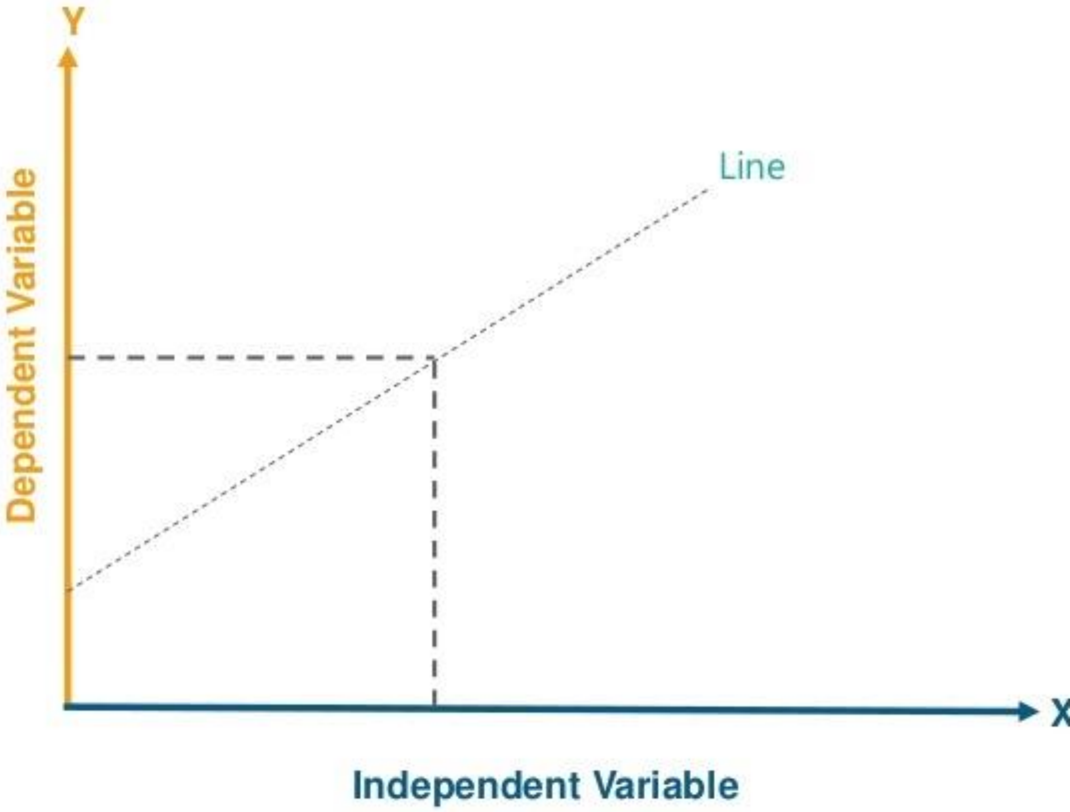
Selection Criteria

- Classification and Regression Capabilities
- Data Quality
- Computational Complexity
- Comprehensible and Transparent

Where is Linear Regression used?

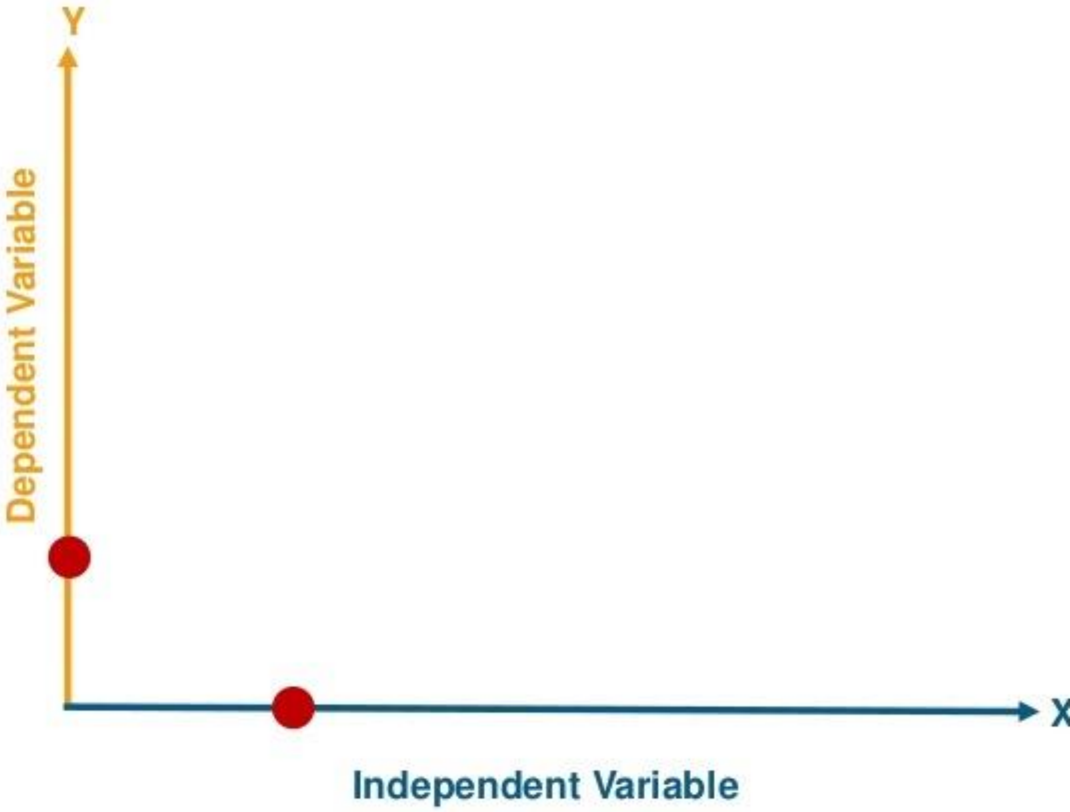
- Evaluating Trends and Sales Estimates
- Analyzing the Impact of Price Changes
- Assessment of risk in financial services and insurance domain

Understanding Linear Regression Algorithm



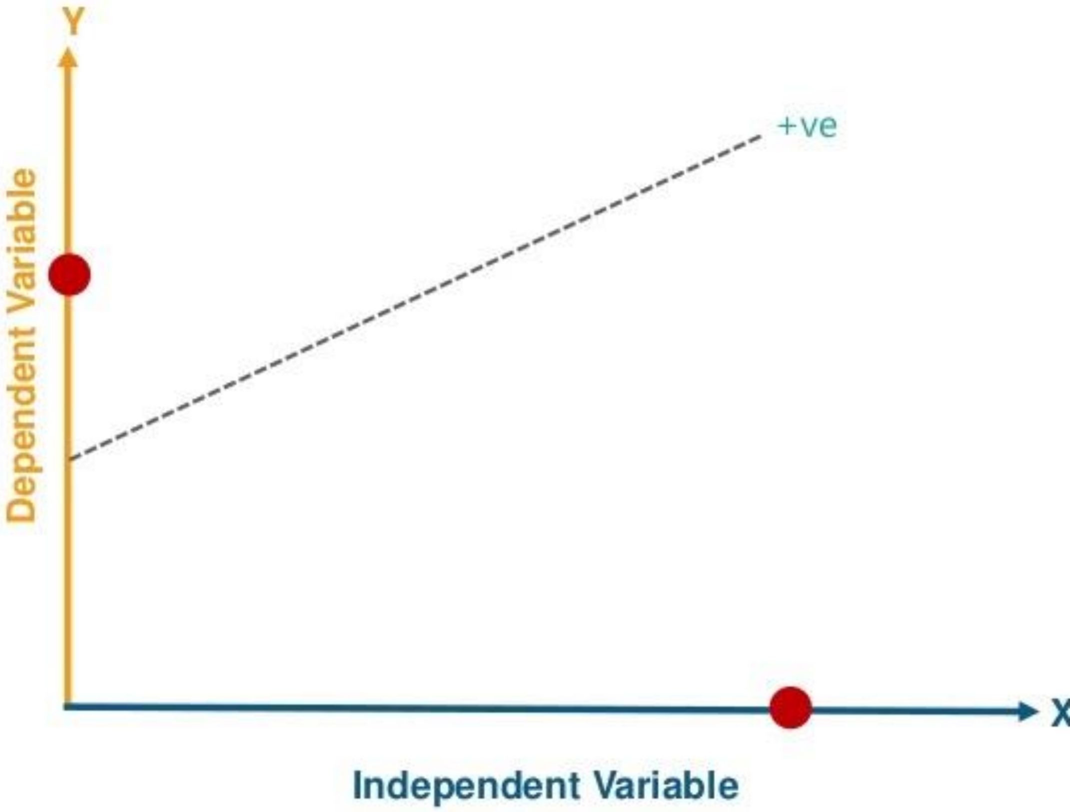
edureka!

Understanding Linear Regression Algorithm



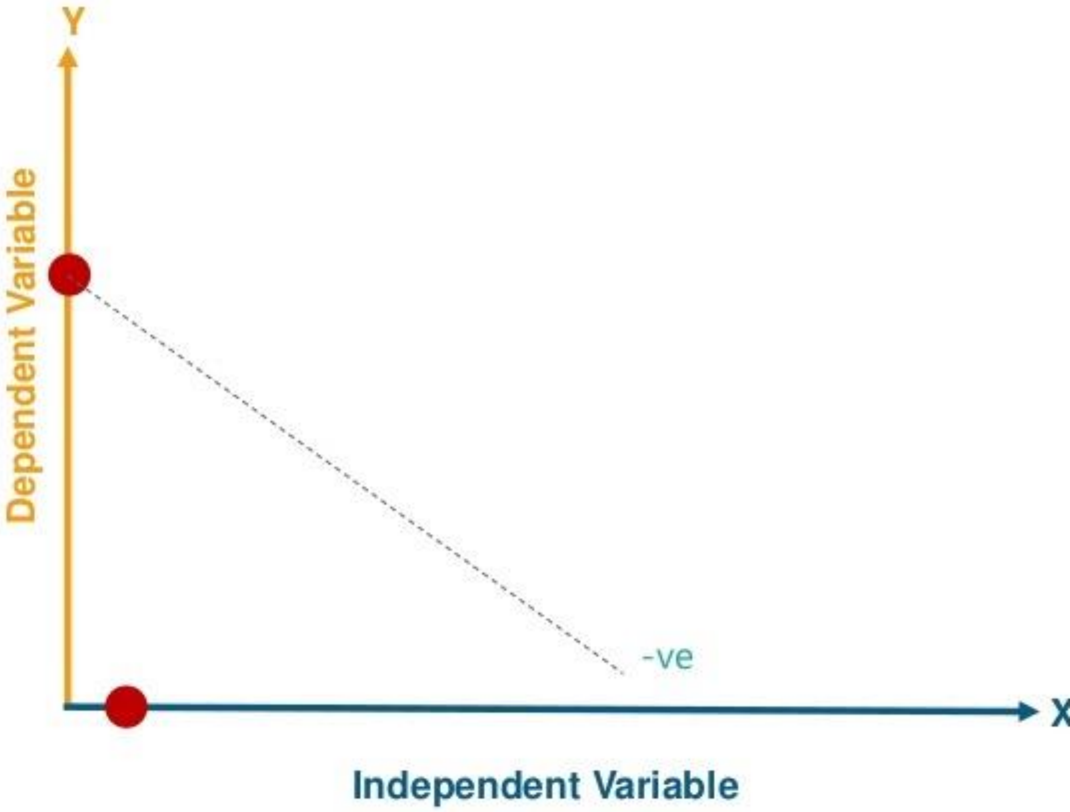
edureka!

Understanding Linear Regression Algorithm



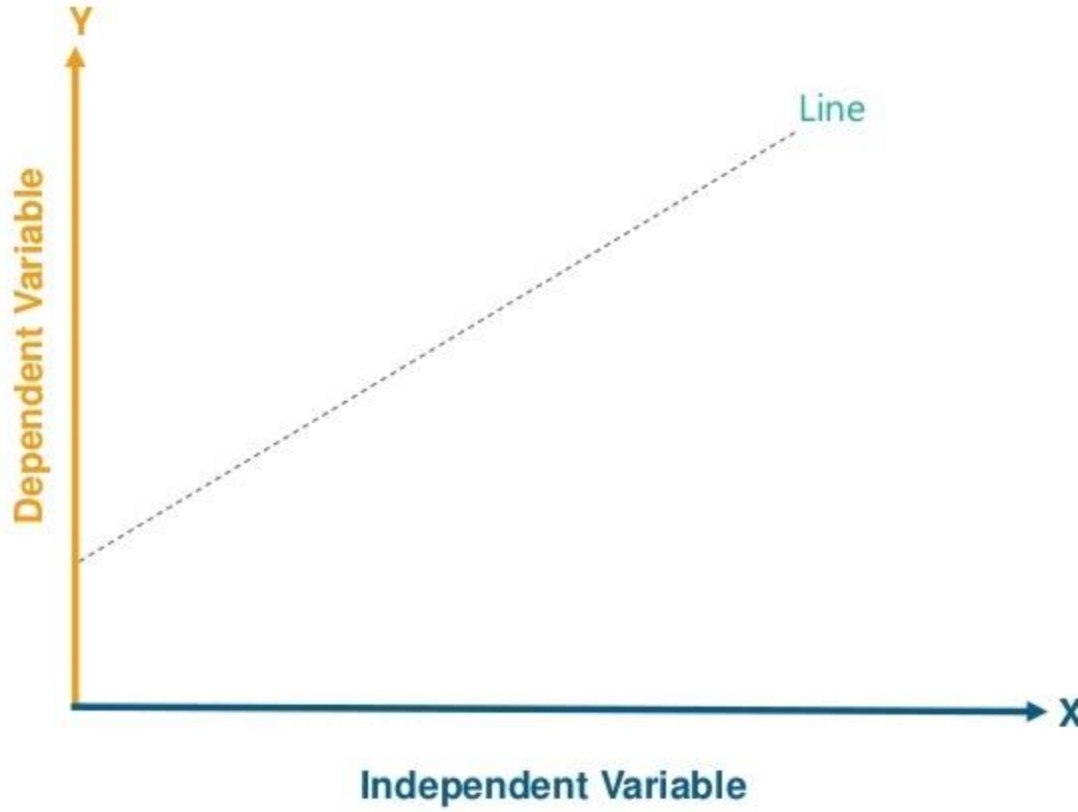
edureka!

Understanding Linear Regression Algorithm

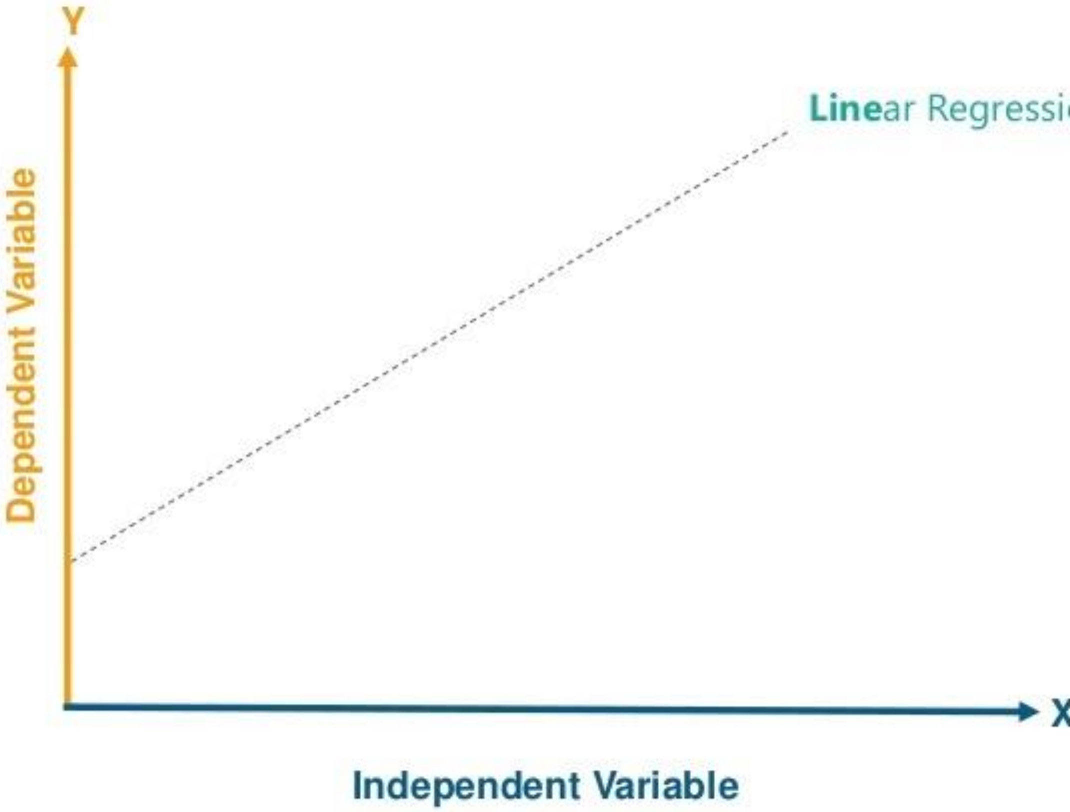


edureka!

Understanding Linear Regression Algorithm

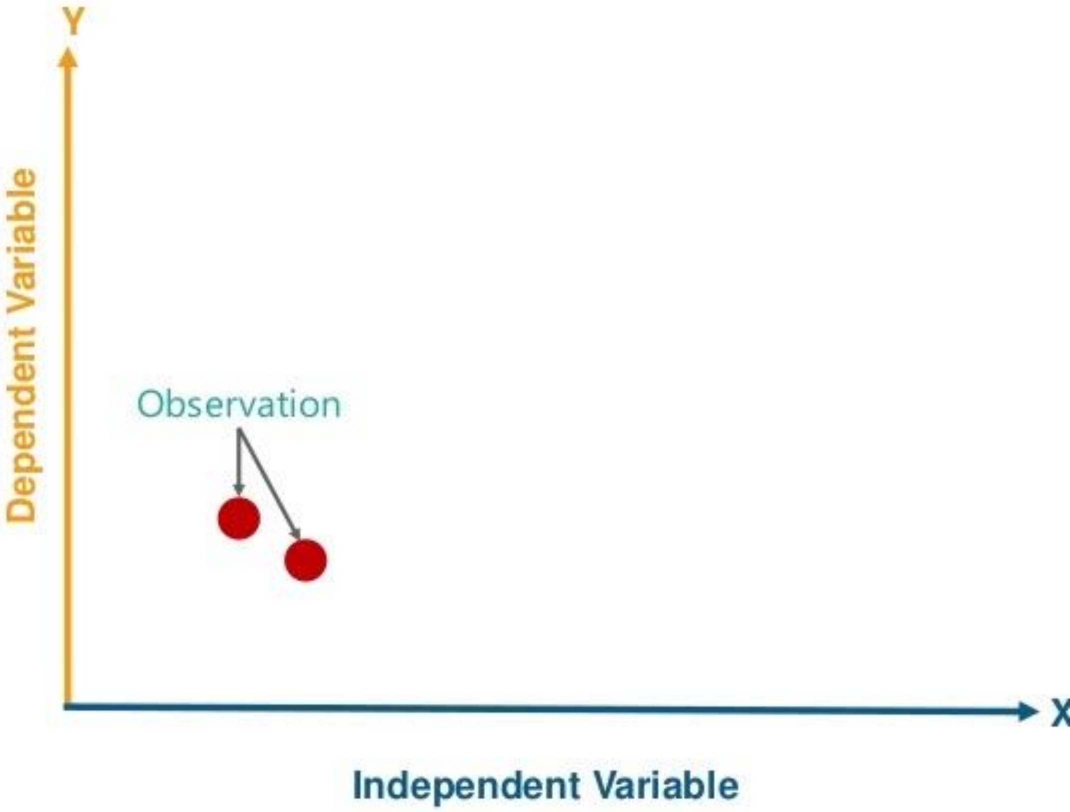


Understanding Linear Regression Algorithm



edureka!

Understanding Linear Regression Algorithm

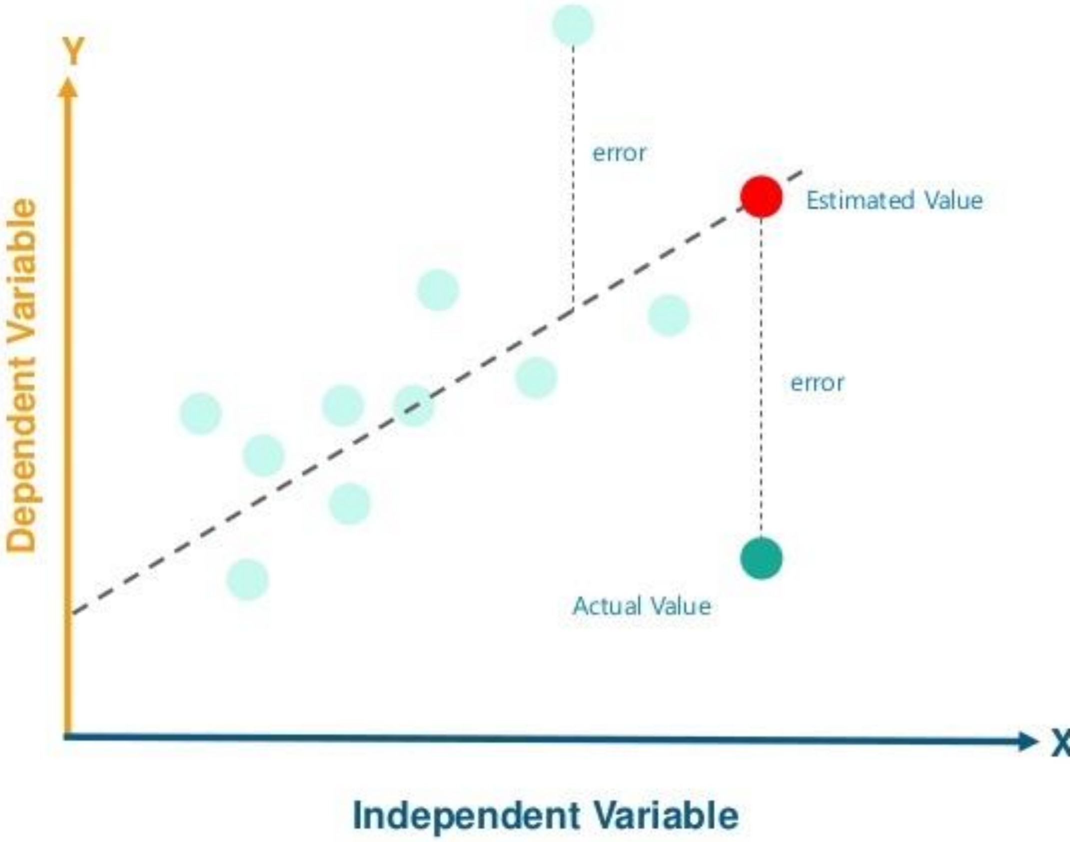


edureka!

Understanding Linear Regression Algorithm

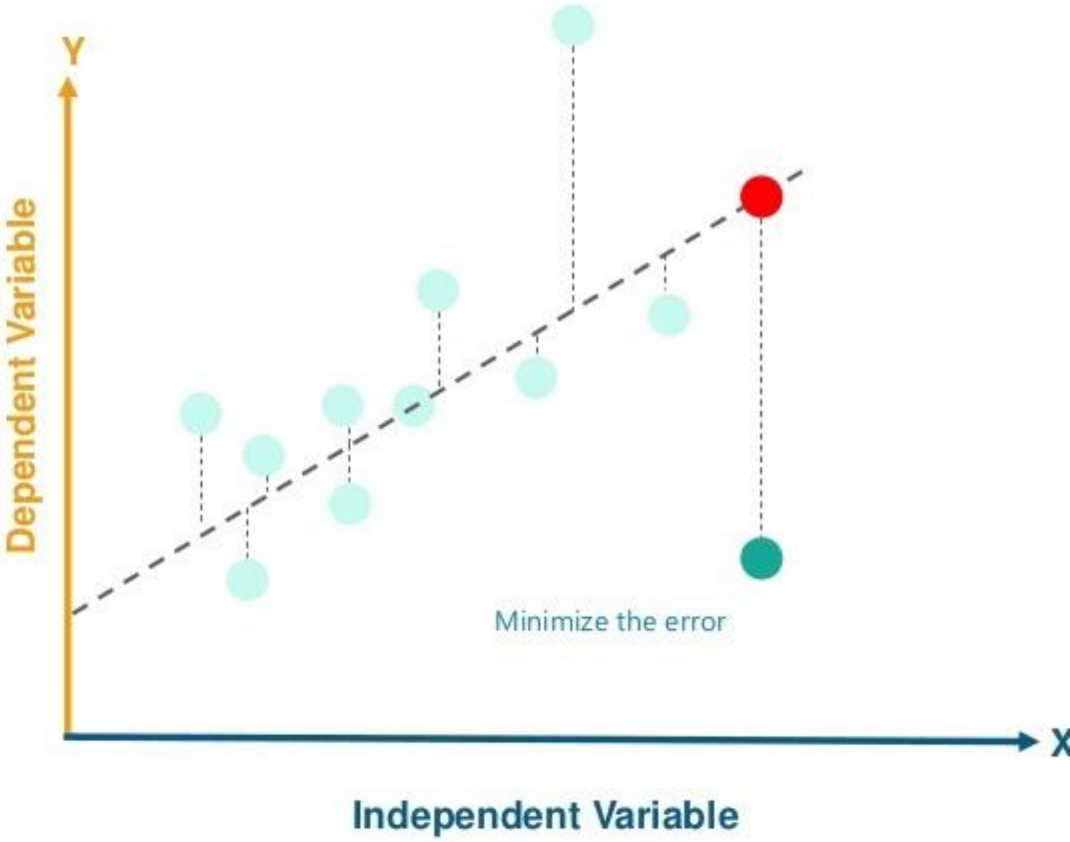
edureka!

Understanding Linear Regression Algorithm



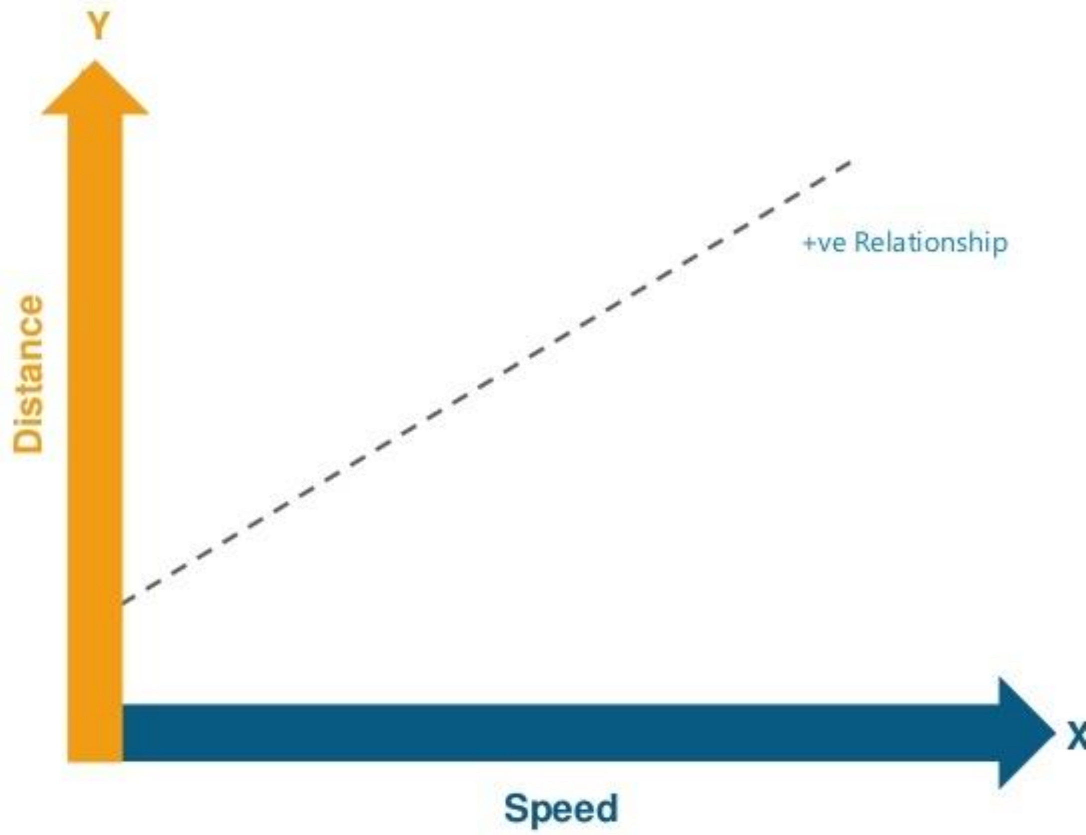
edureka!

Understanding Linear Regression Algorithm



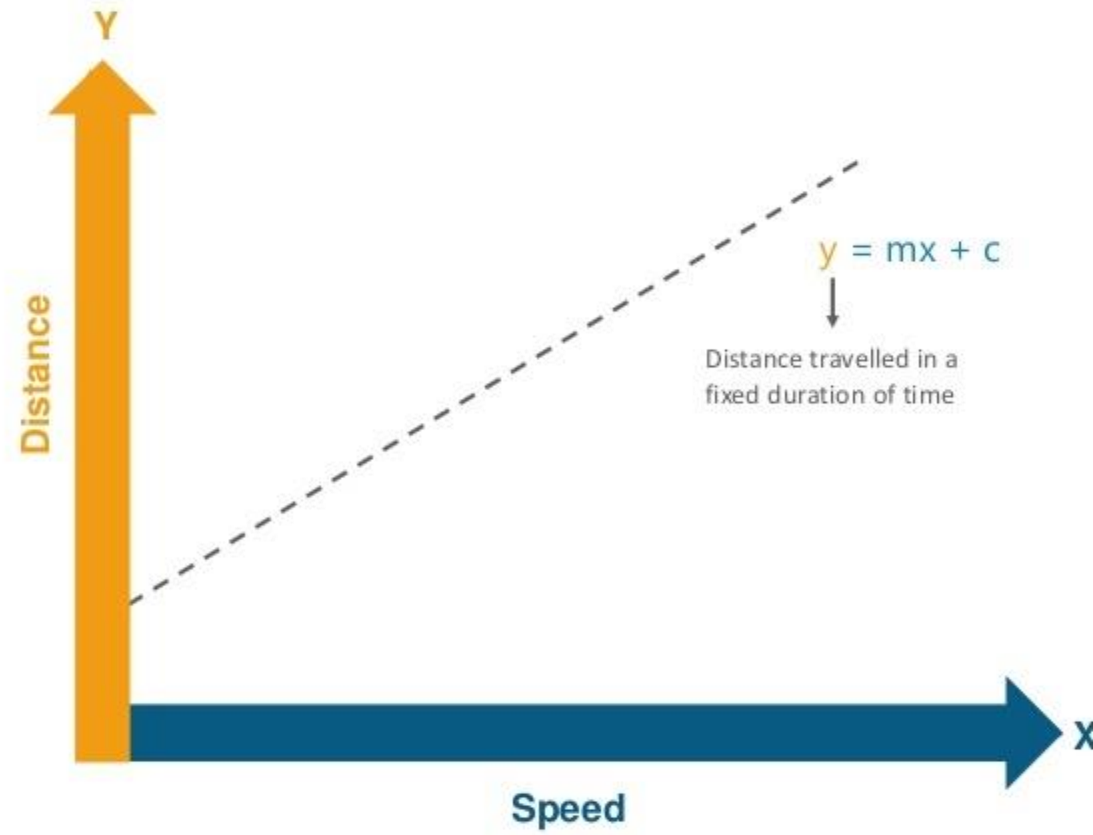
edureka!

Understanding Linear Regression Algorithm



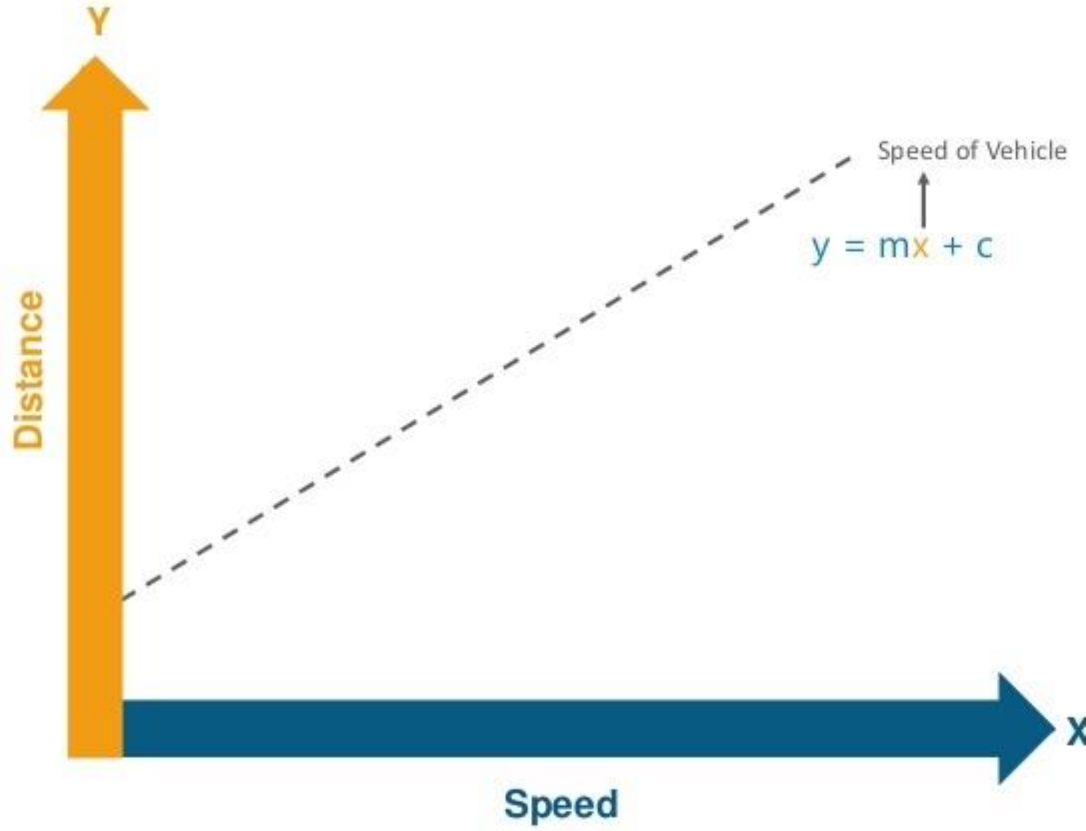
edureka!

Understanding Linear Regression Algorithm



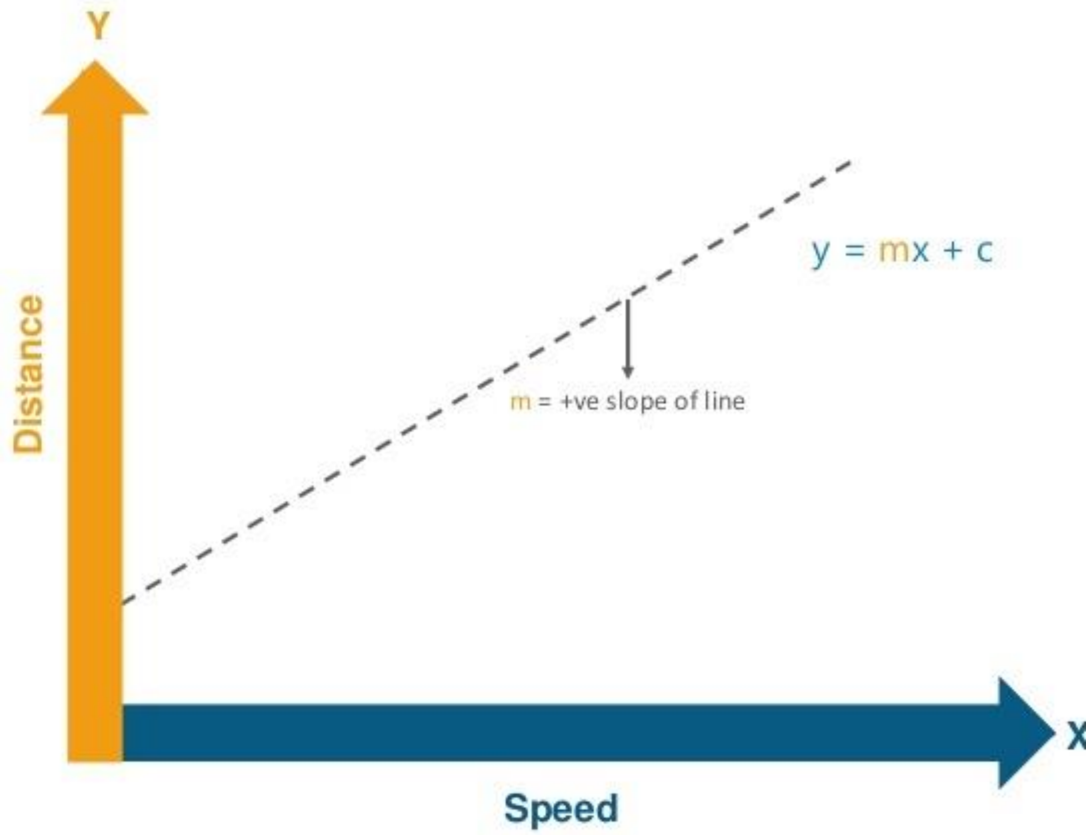
edureka!

Understanding Linear Regression Algorithm



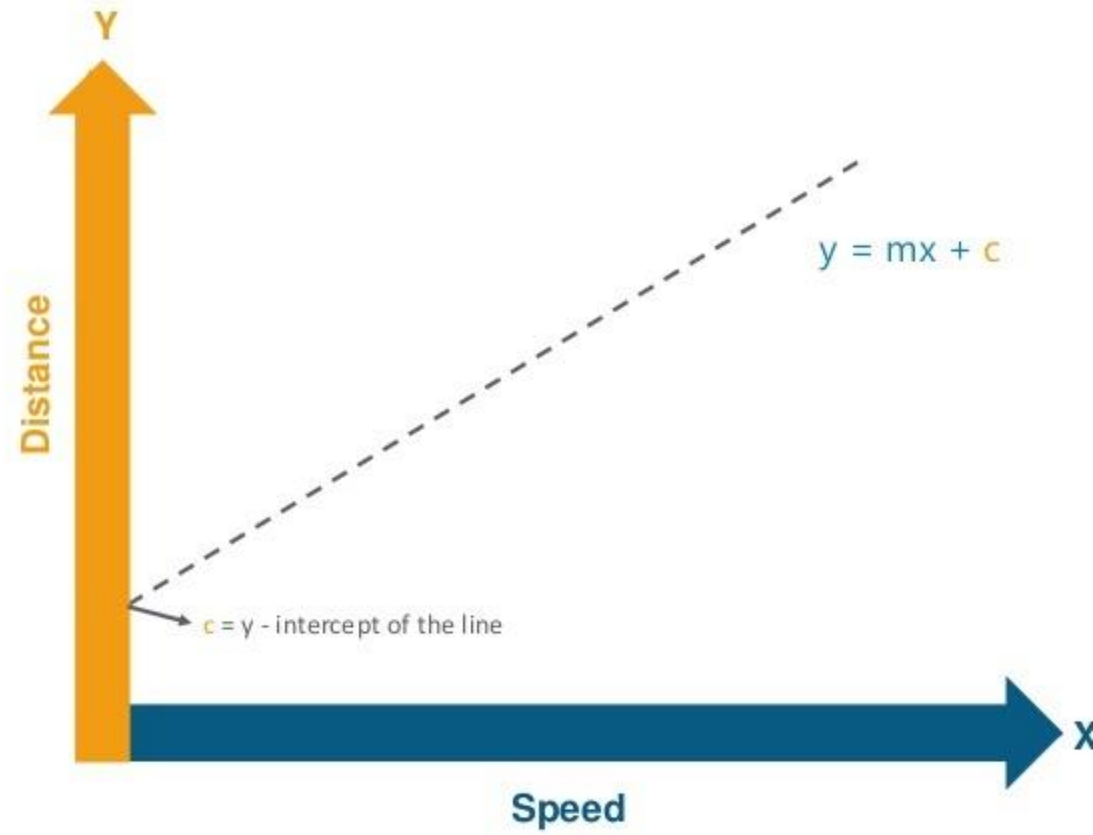
edureka!

Understanding Linear Regression Algorithm



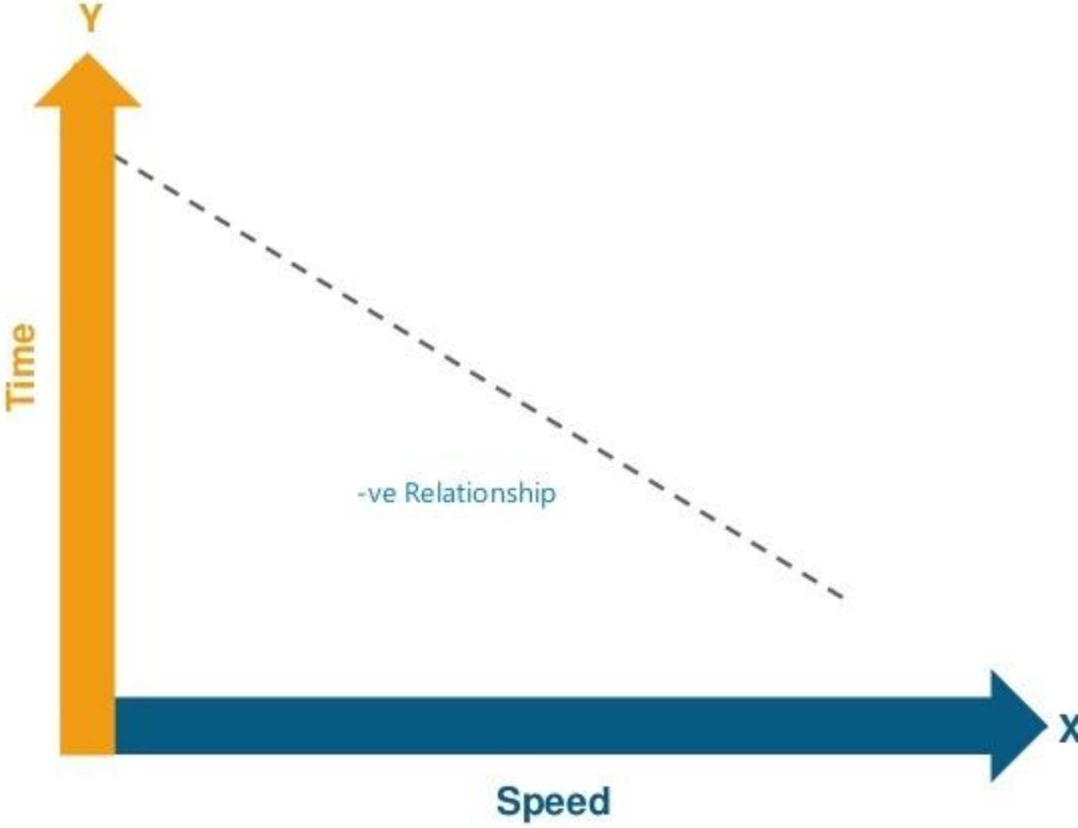
edureka!

Understanding Linear Regression Algorithm



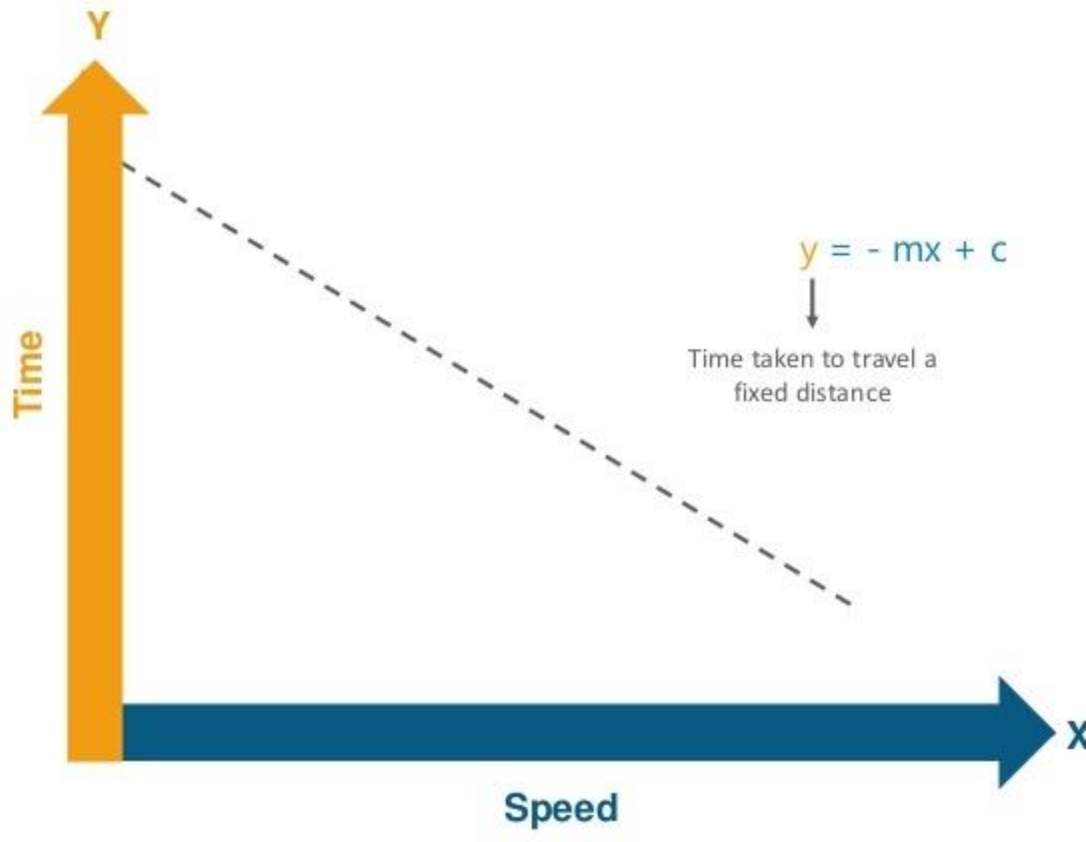
edureka!

Understanding Linear Regression Algorithm



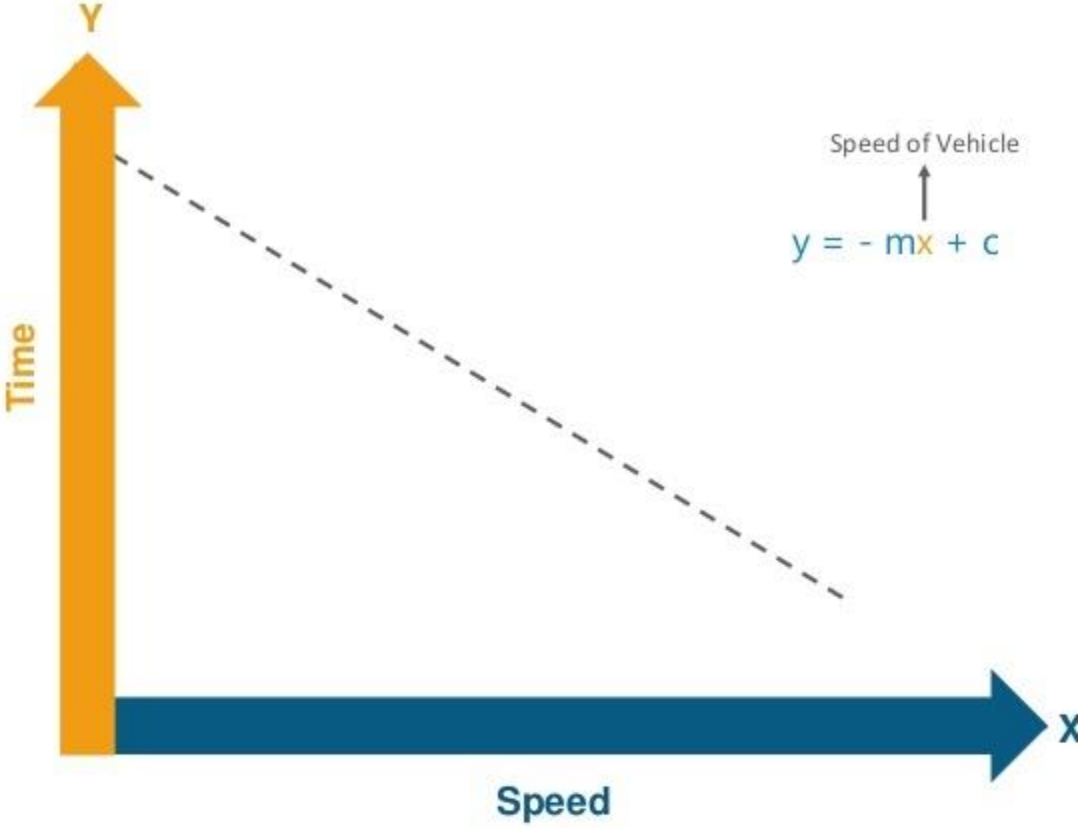
edureka!

Understanding Linear Regression Algorithm



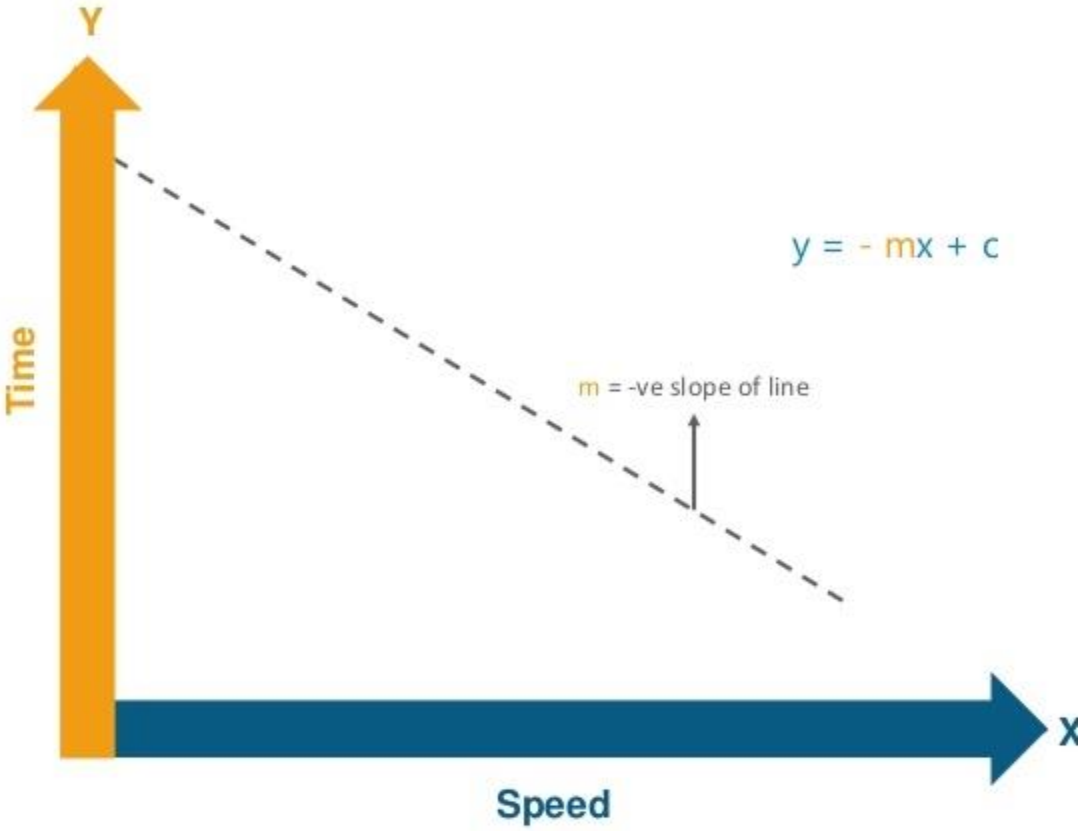
edureka!

Understanding Linear Regression Algorithm



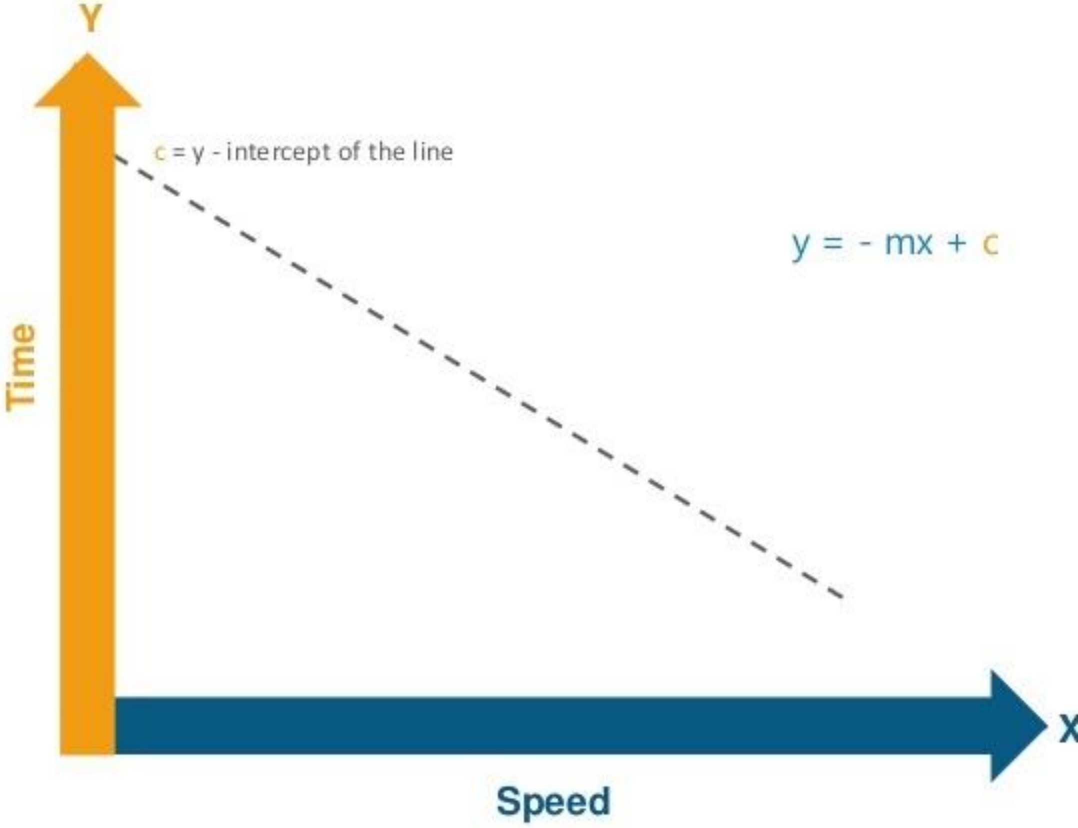
edureka!

Understanding Linear Regression Algorithm



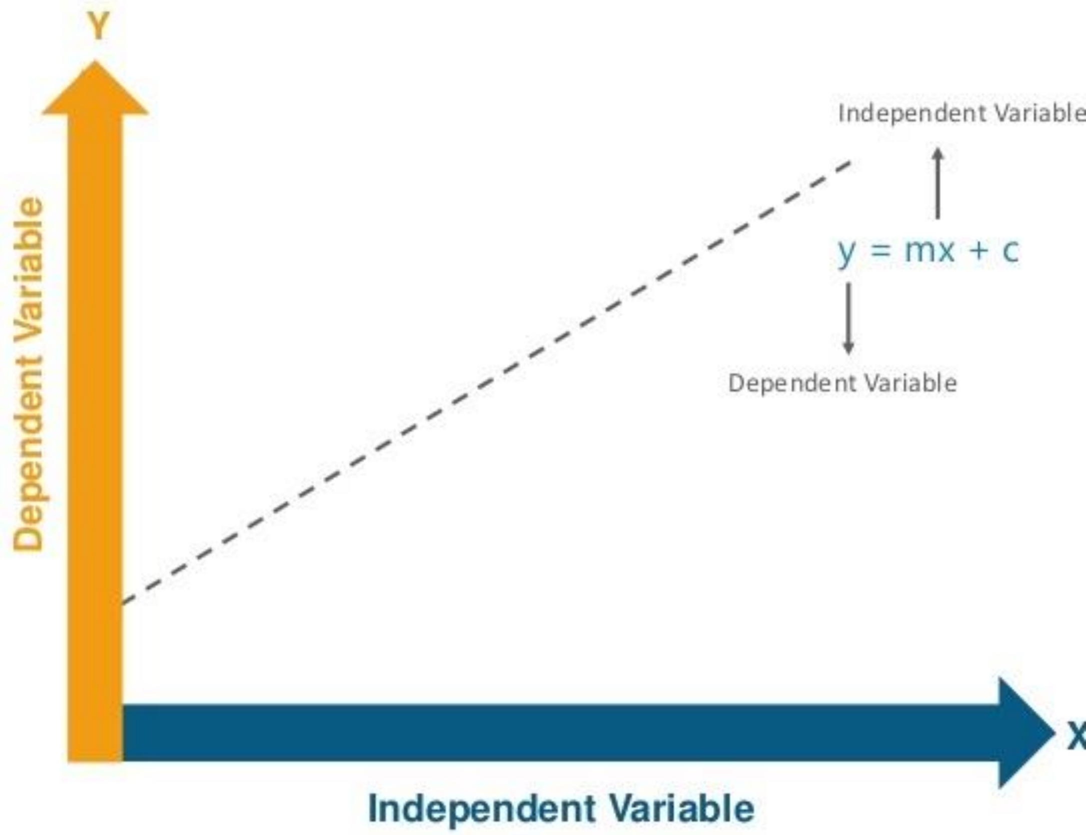
edureka!

Understanding Linear Regression Algorithm

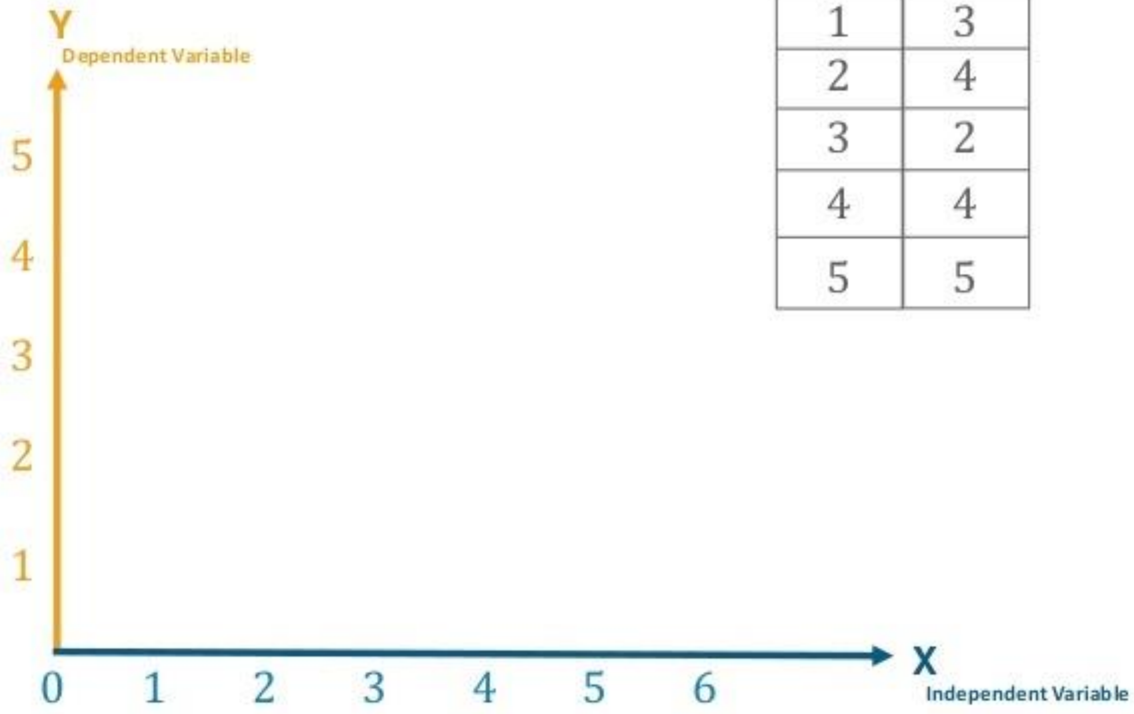


edureka!

Understanding Linear Regression Algorithm

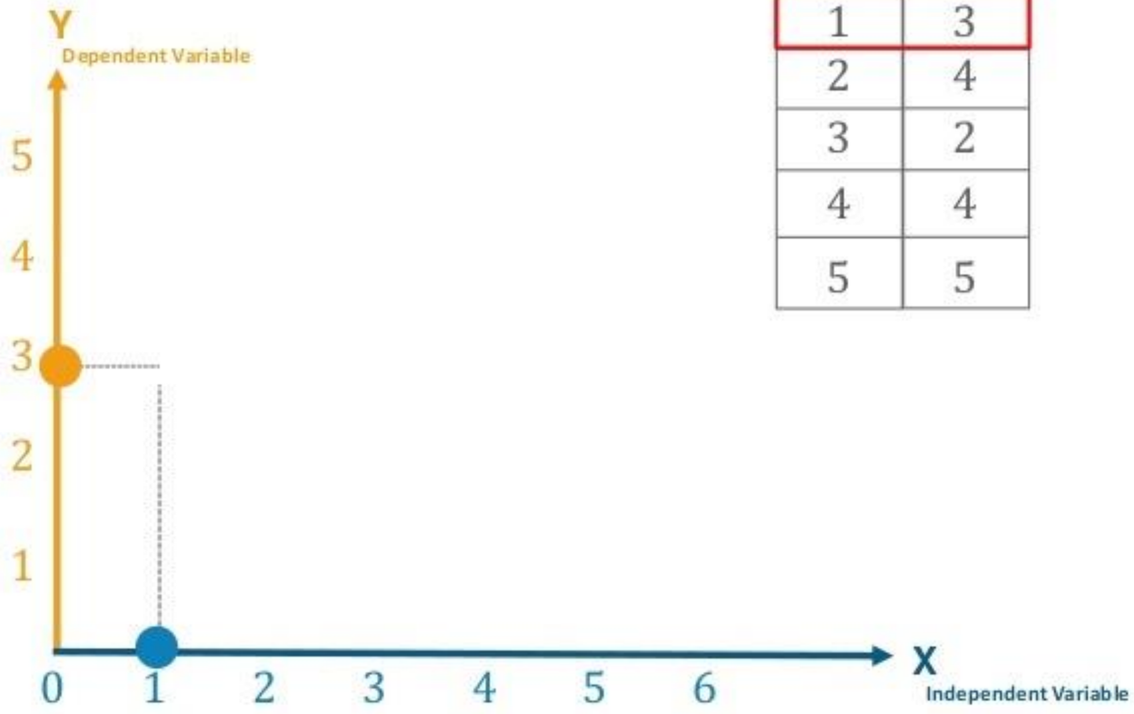


Understanding Linear Regression Algorithm



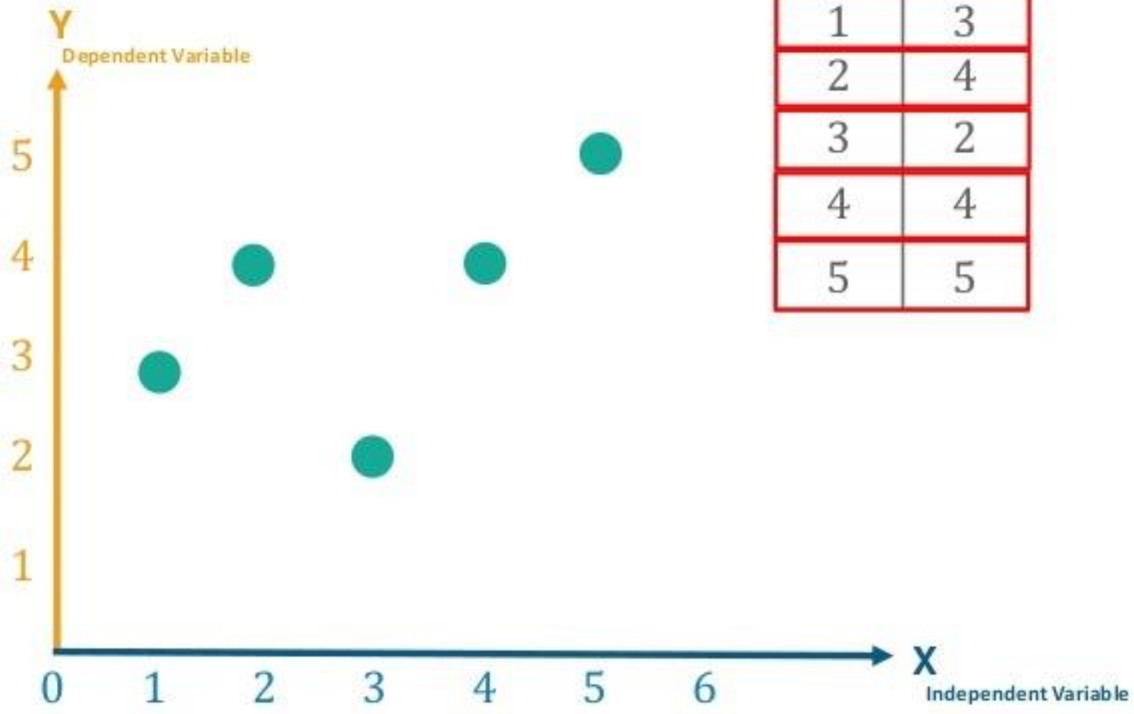
x	y
1	3
2	4
3	2
4	4
5	5

Understanding Linear Regression Algorithm



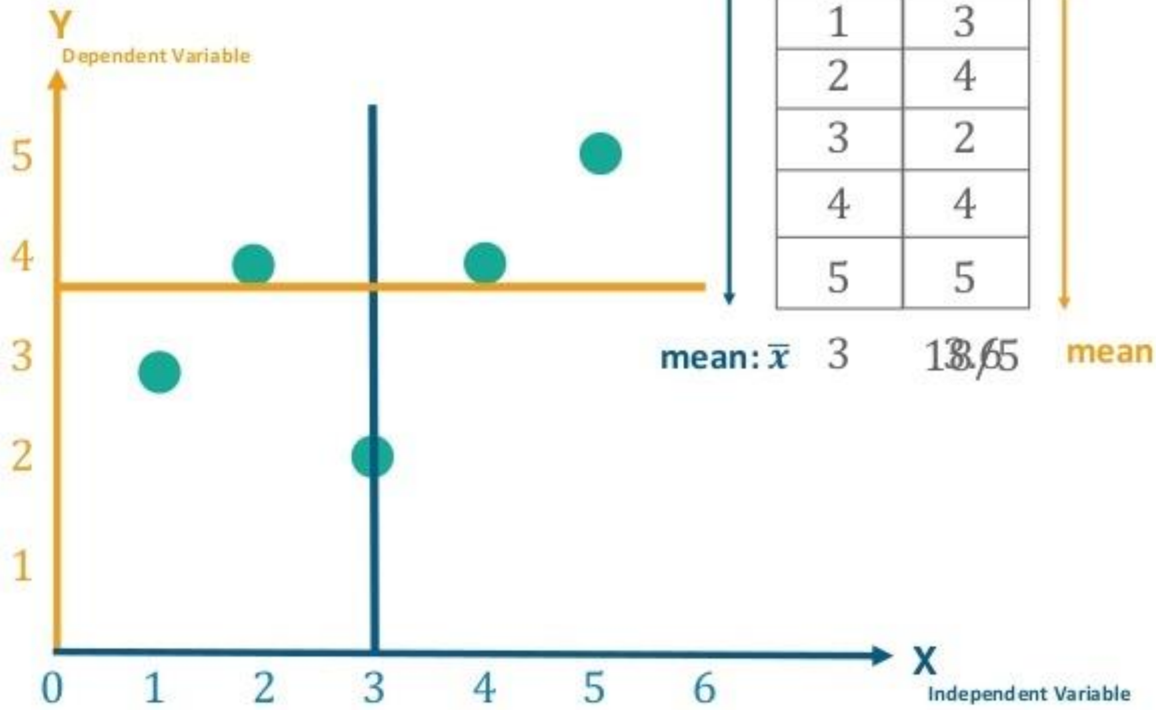
x	y
1	3
2	4
3	2
4	4
5	5

Understanding Linear Regression Algorithm



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

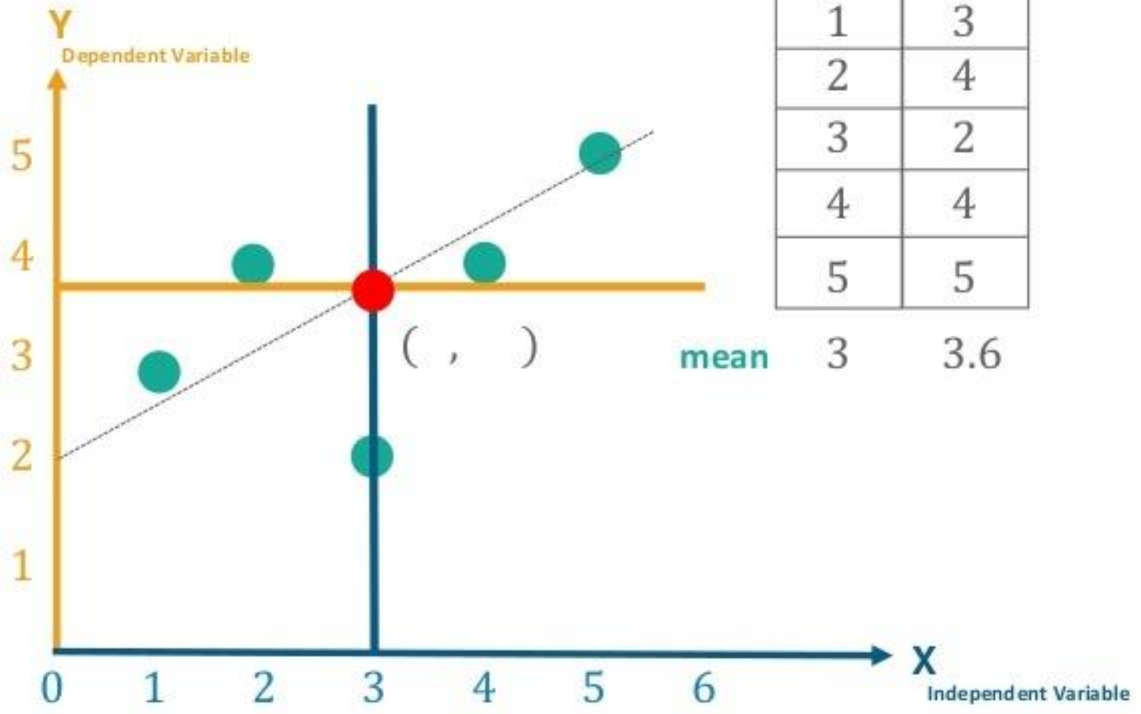
Understanding Linear Regression Algorithm



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

mean: $\overline{x}$ 3   18/5 3.6 mean : $\overline{y}$

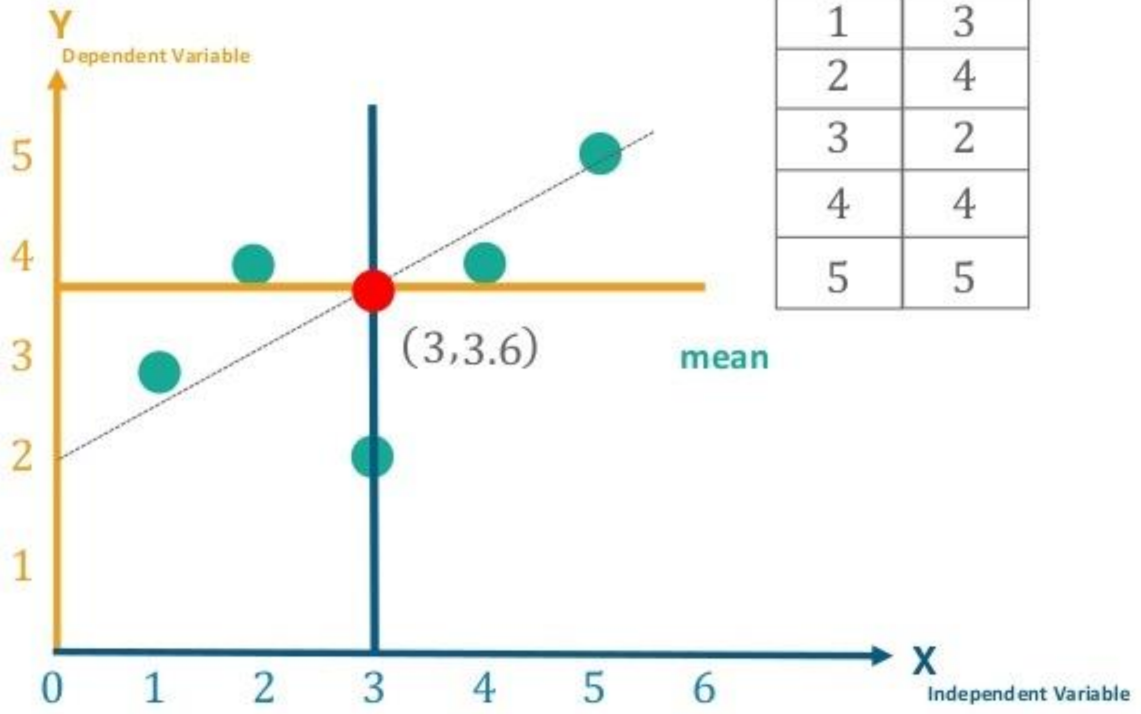
Understanding Linear Regression Algorithm



x	y
1	3
2	4
3	2
4	4
5	5
3	3.6

mean 3 3.6

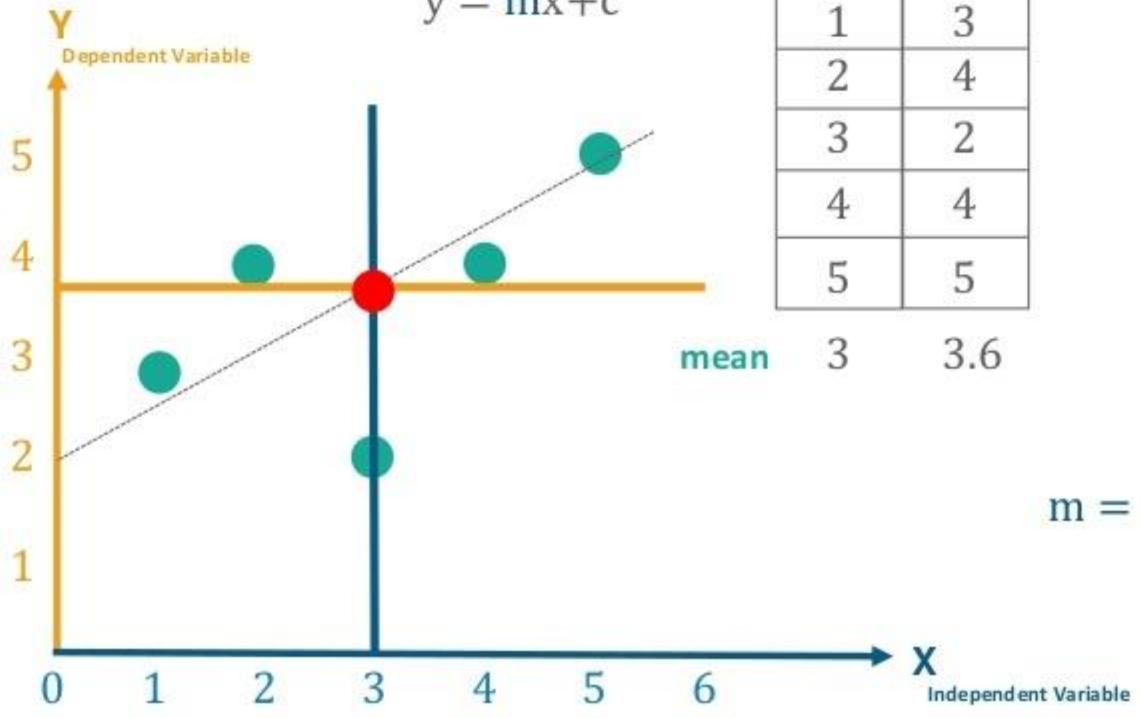
Understanding Linear Regression Algorithm



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

Understanding Linear Regression Algorithm

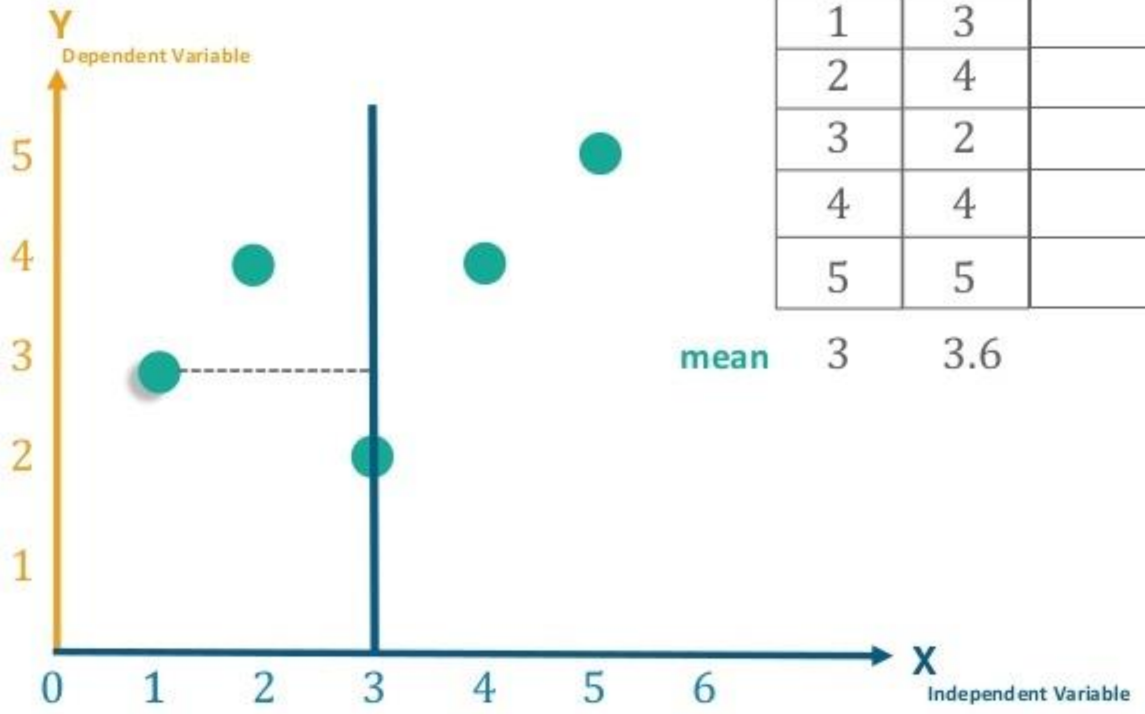
$y = mx+c$



x	y
1	3
2	4
3	2
4	4
5	5
mean 3	3.6

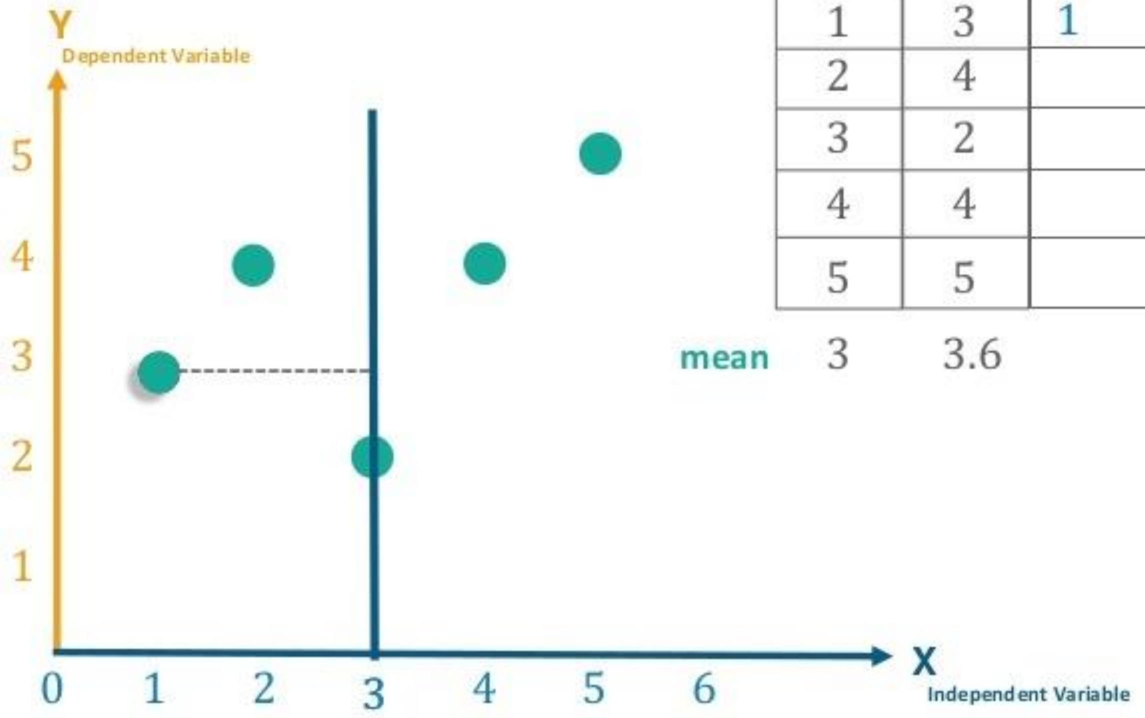
$$m = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

Understanding Linear Regression Algorithm



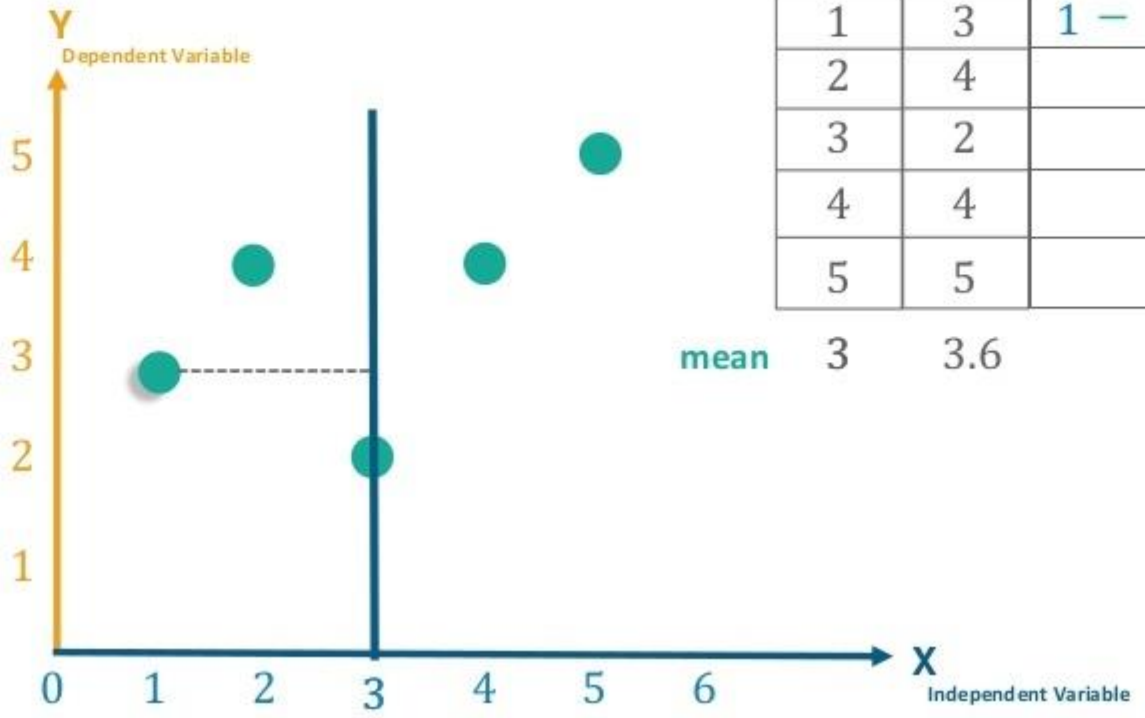
x	y	$x - \bar{x}$
1	3	
2	4	
3	2	
4	4	
5	5	
mean 3	3.6	

Understanding Linear Regression Algorithm



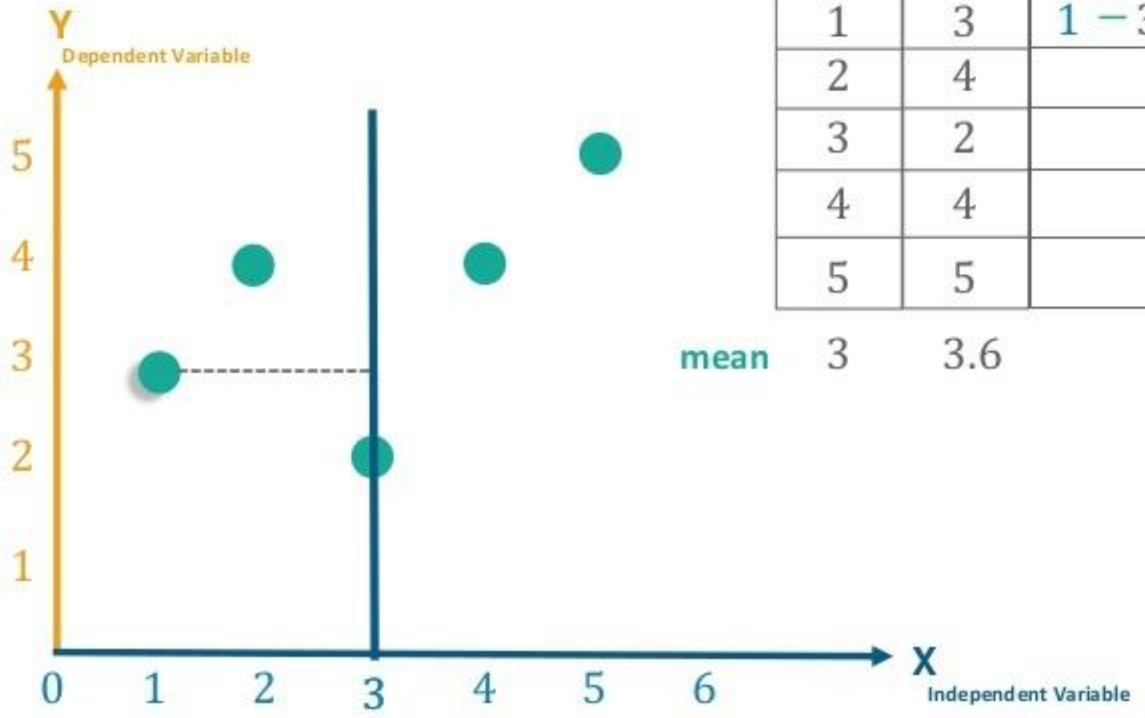
x	y	$x - \bar{x}$
1	3	1
2	4	
3	2	
4	4	
5	5	
mean 3	3.6	

Understanding Linear Regression Algorithm



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

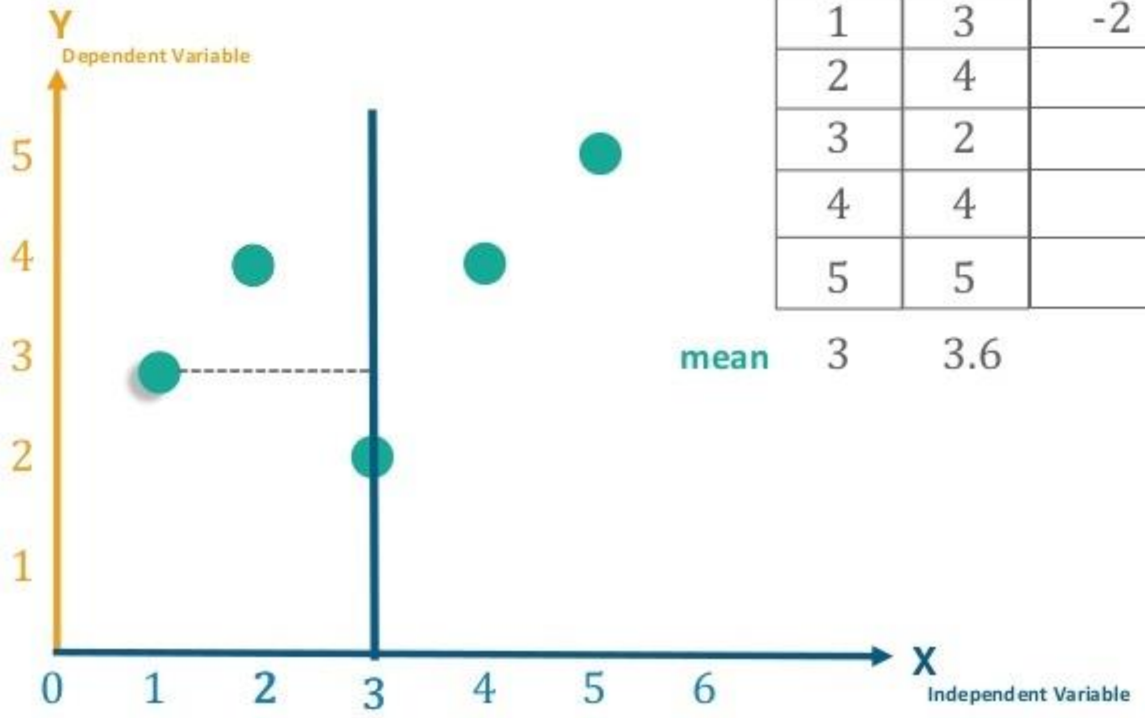
Understanding Linear Regression Algorithm



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

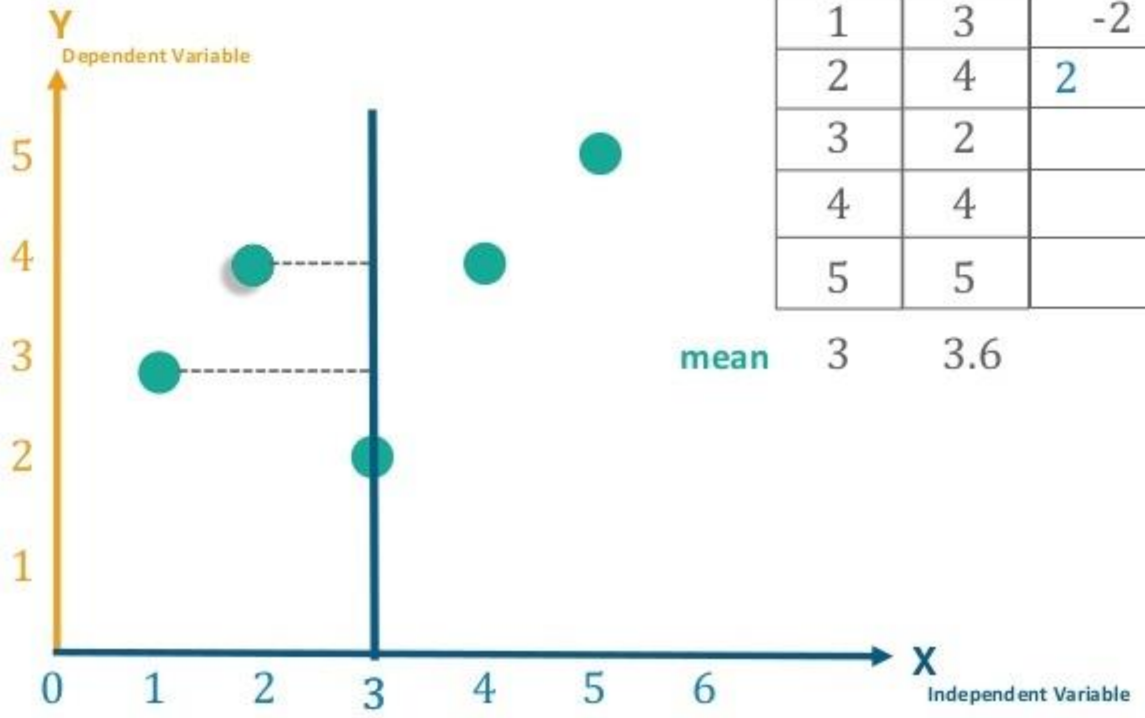
mean	3	3.6

Understanding Linear Regression Algorithm



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

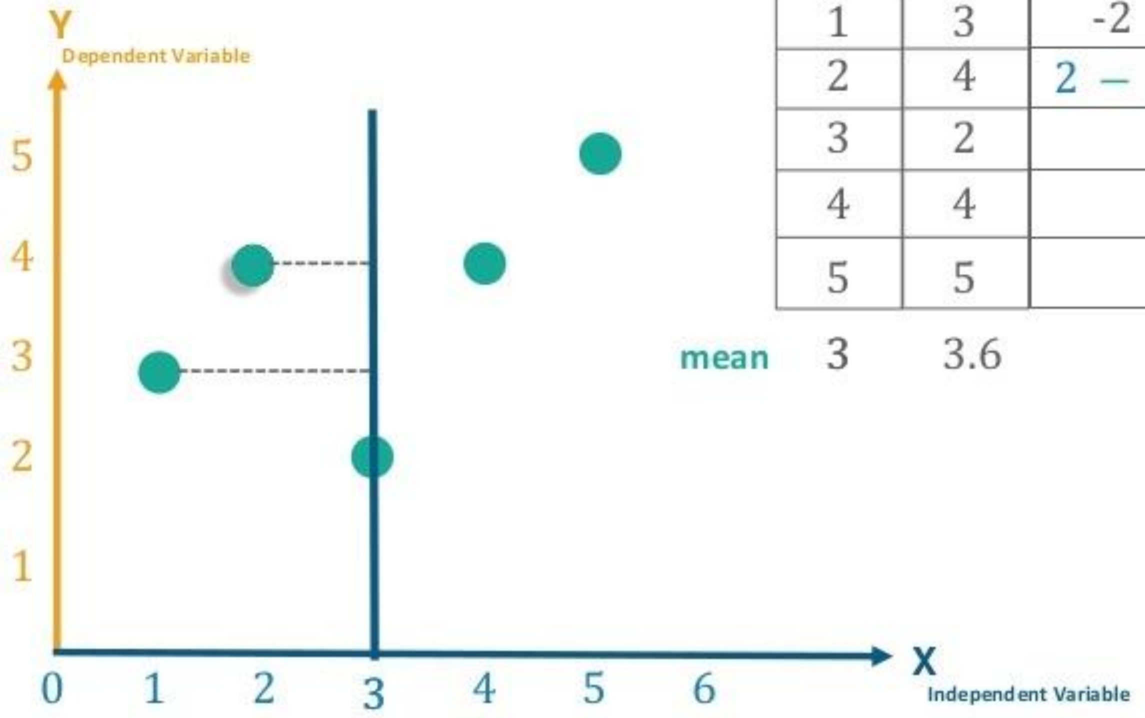
Understanding Linear Regression Algorithm



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

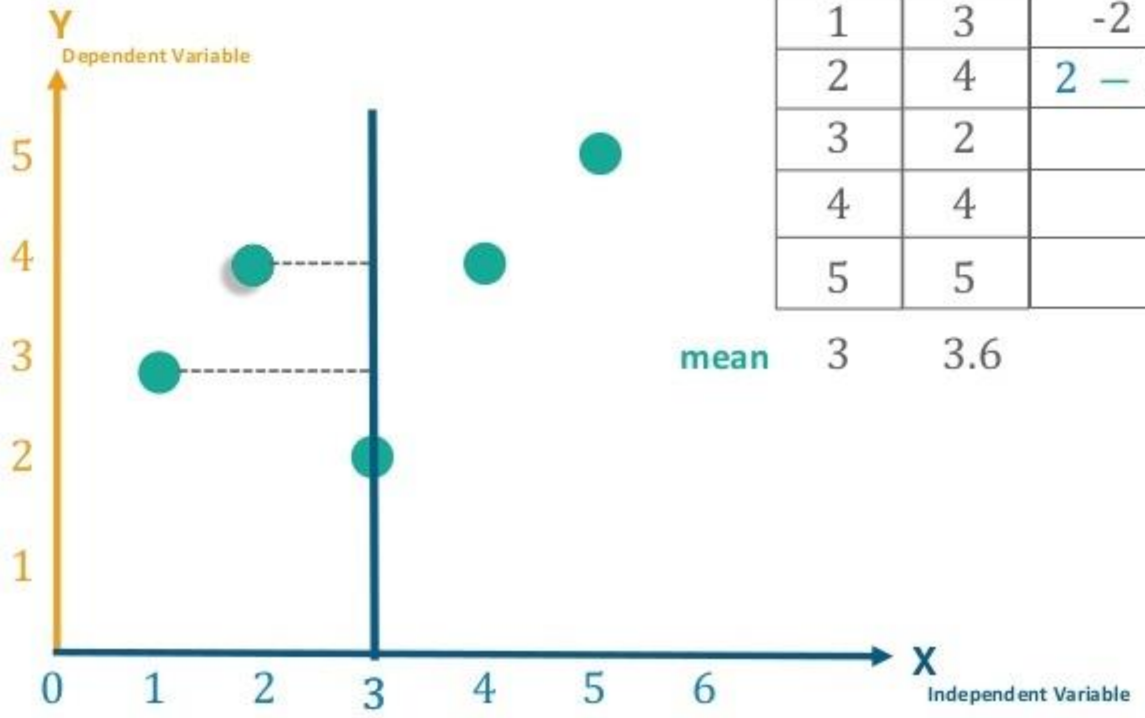
mean 3 3.6

Understanding Linear Regression Algorithm



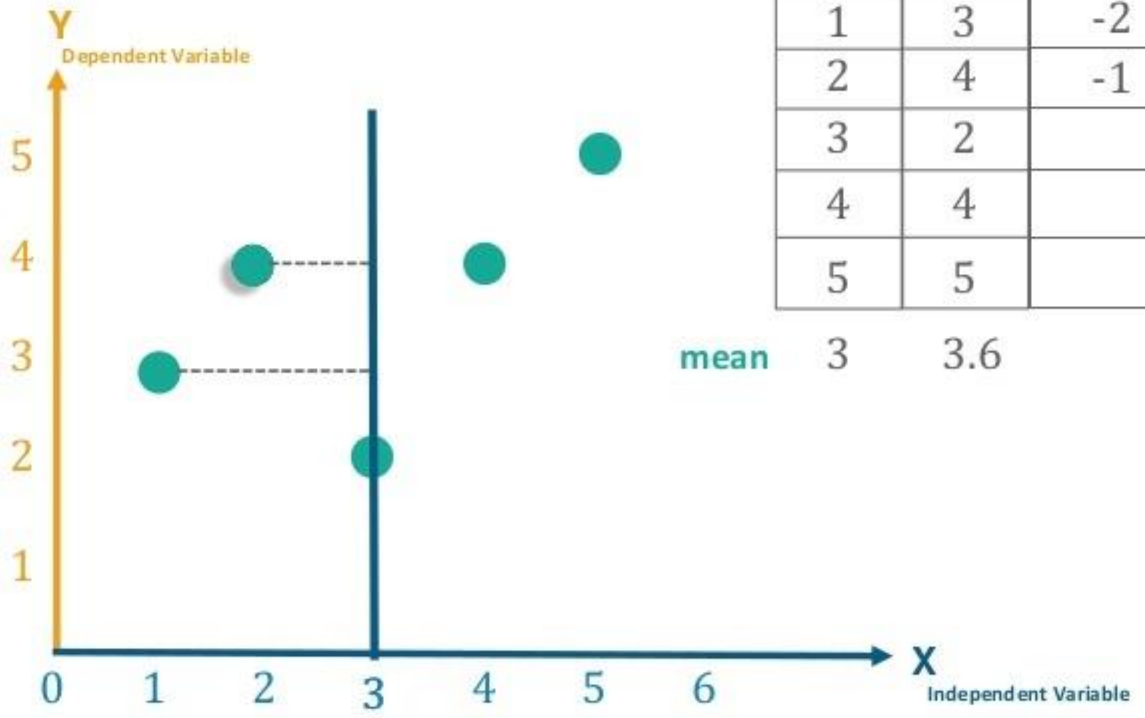
x	y	$x - \bar{x}$
1	3	-2
2	4	2 −
3	2	
4	4	
5	5	
mean 3	3.6	

Understanding Linear Regression Algorithm



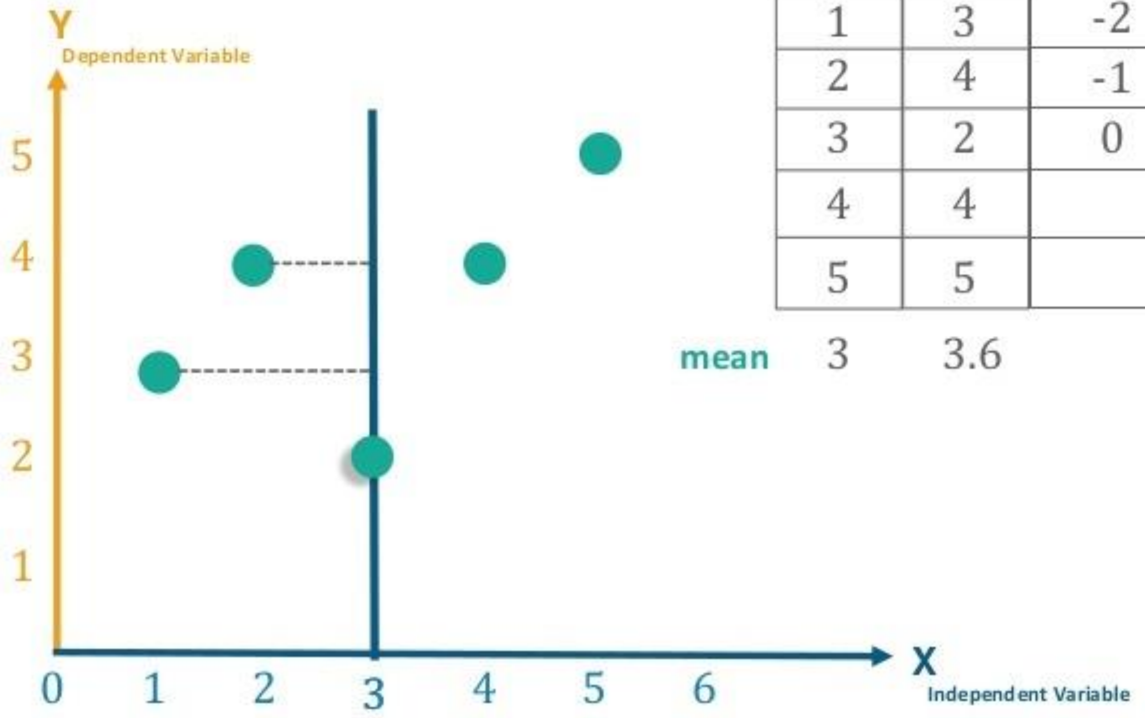
x	y	$x - \bar{x}$
1	3	-2
2	4	$2 - 3$
3	2	
4	4	
5	5	
mean 3	3.6	

Understanding Linear Regression Algorithm



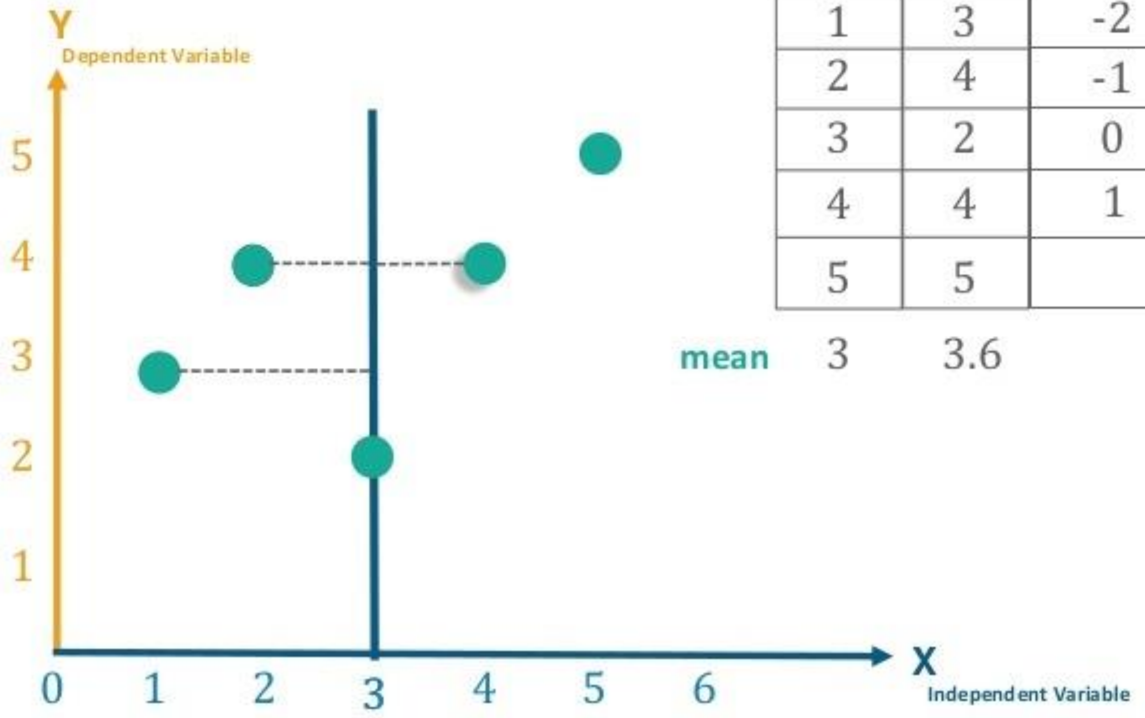
x	y	$x - \bar{x}$
1	3	-2
2	4	-1
3	2	
4	4	
5	5	
mean 3	3.6	

Understanding Linear Regression Algorithm



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

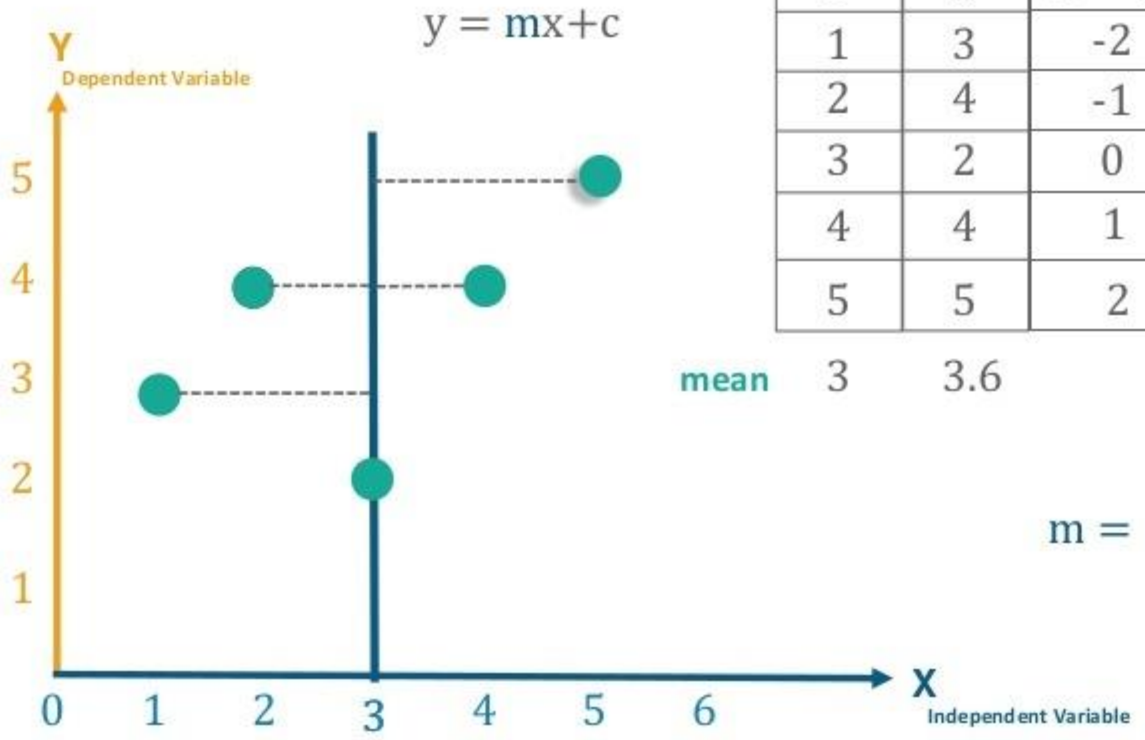
Understanding Linear Regression Algorithm



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

mean 3 3.6

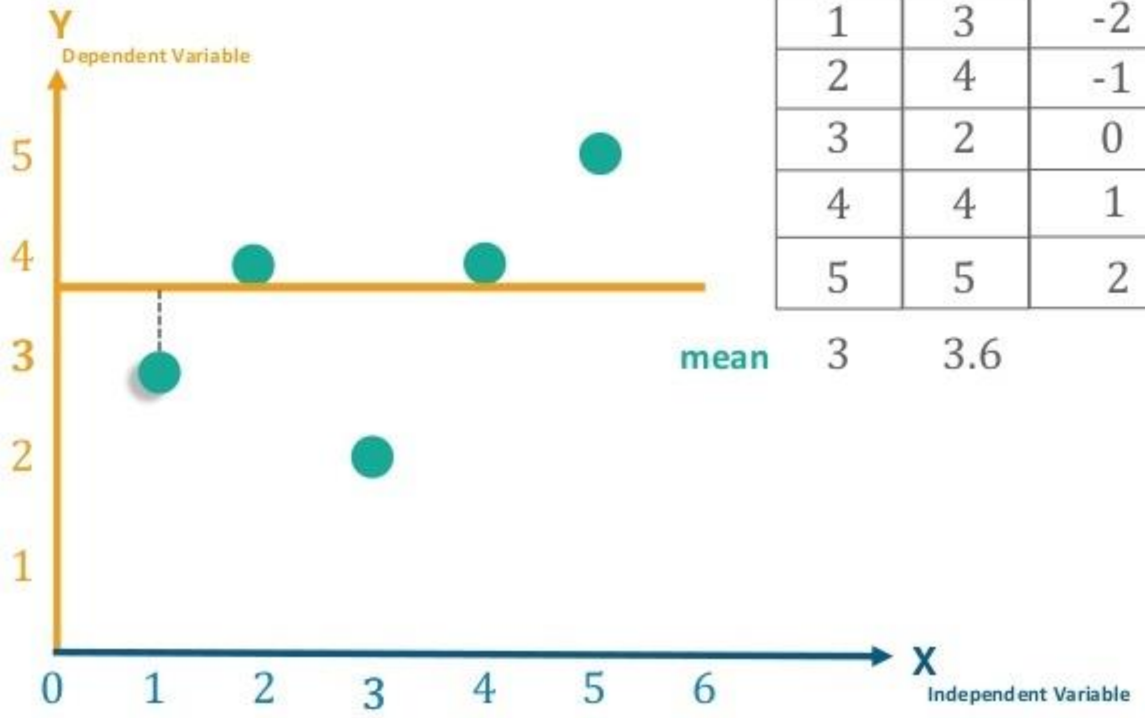
Understanding Linear Regression Algorithm



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

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

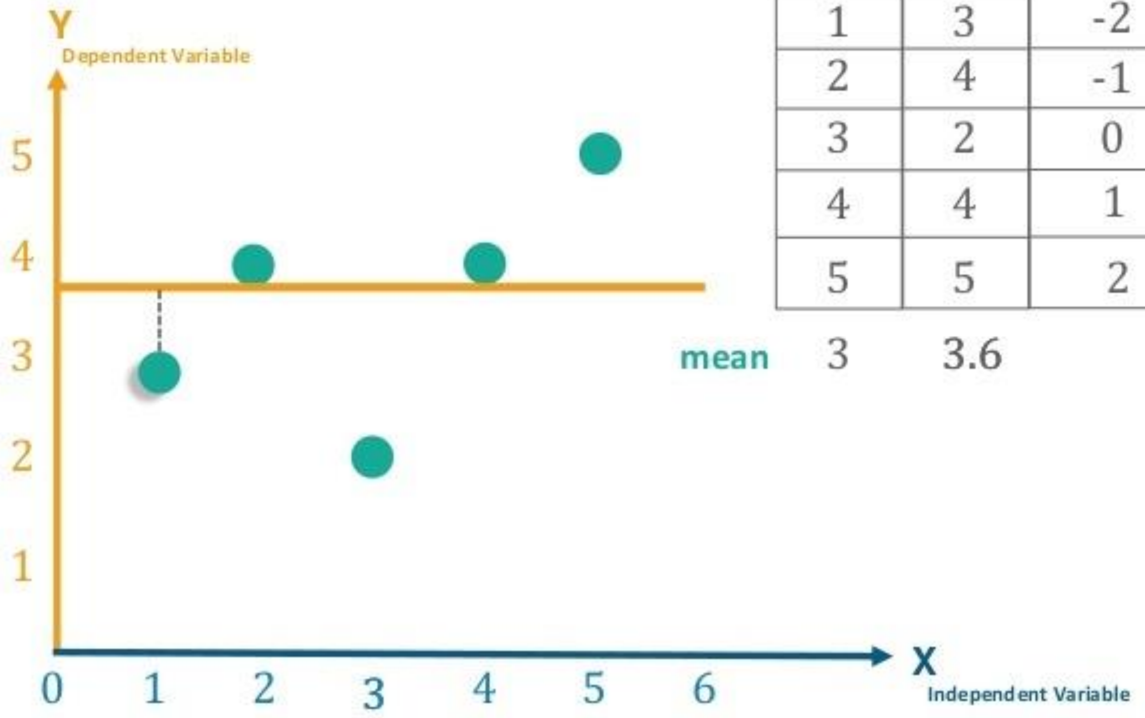
Understanding Linear Regression Algorithm



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

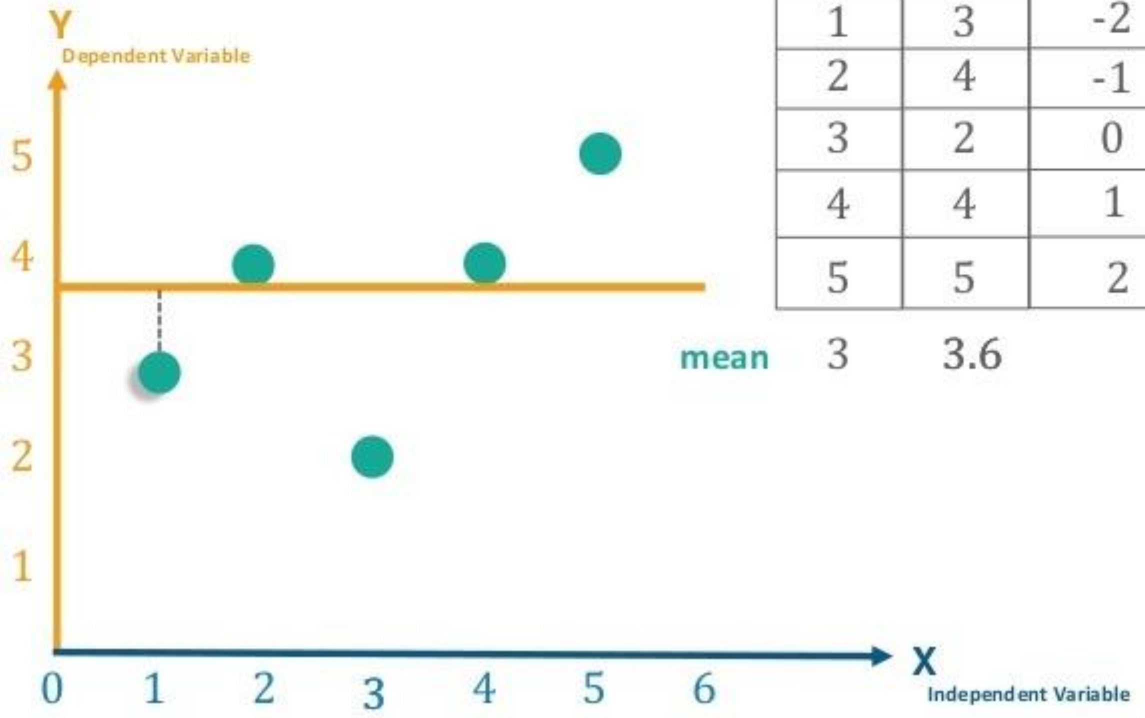
mean 3 3.6

Understanding Linear Regression Algorithm



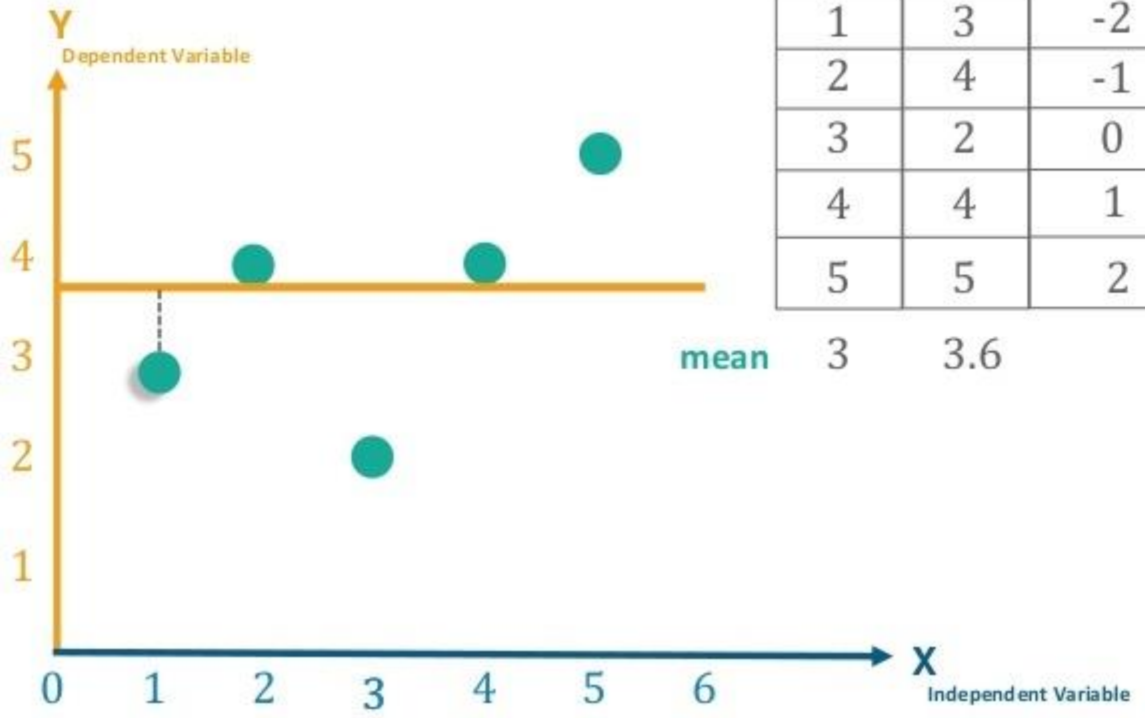
x	y	$x - \bar{x}$	$y - \bar{y}$
1	3	-2	3
2	4	-1	
3	2	0	
4	4	1	
5	5	2	
mean 3	3.6		

Understanding Linear Regression Algorithm



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

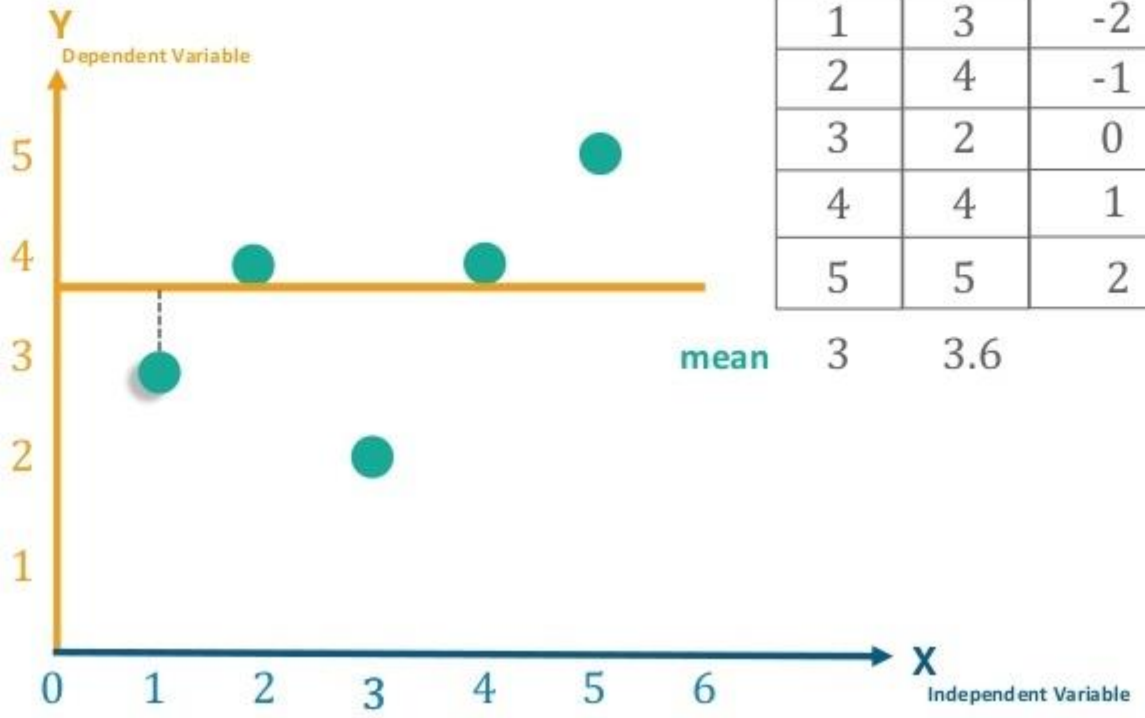
Understanding Linear Regression Algorithm



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

mean 3 3.6

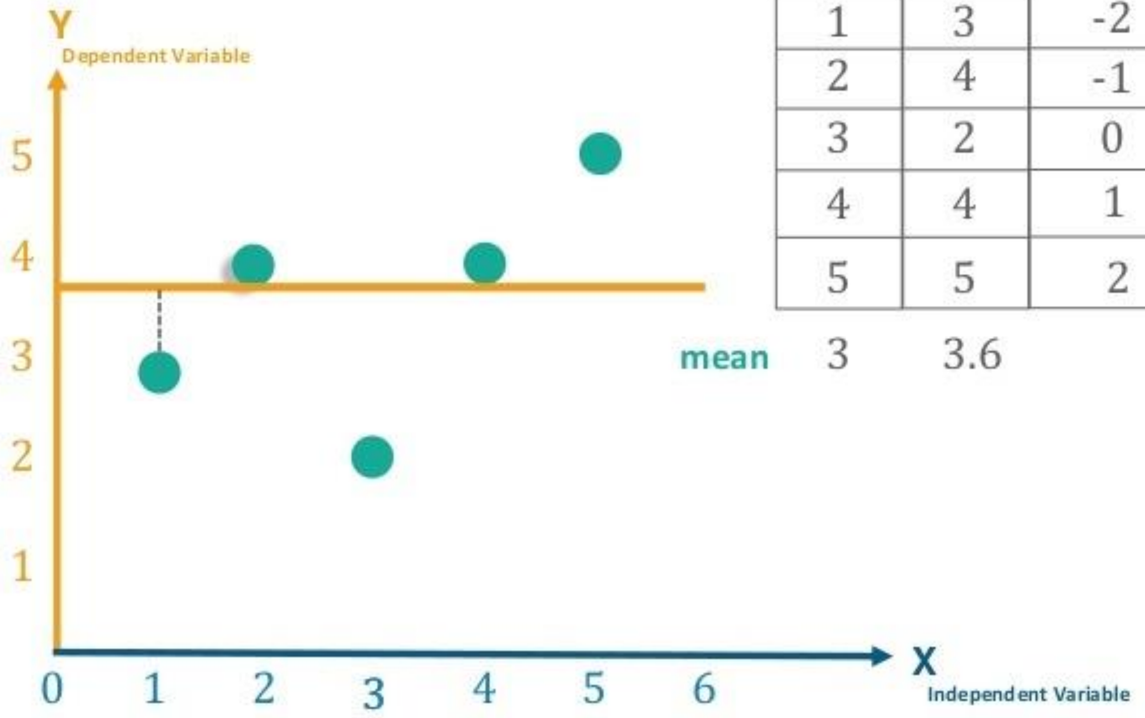
Understanding Linear Regression Algorithm



x	y	$x - \bar{x}$	$y - \bar{y}$
1	3	-2	-0.6
2	4	-1	
3	2	0	
4	4	1	
5	5	2	

mean 3 3.6

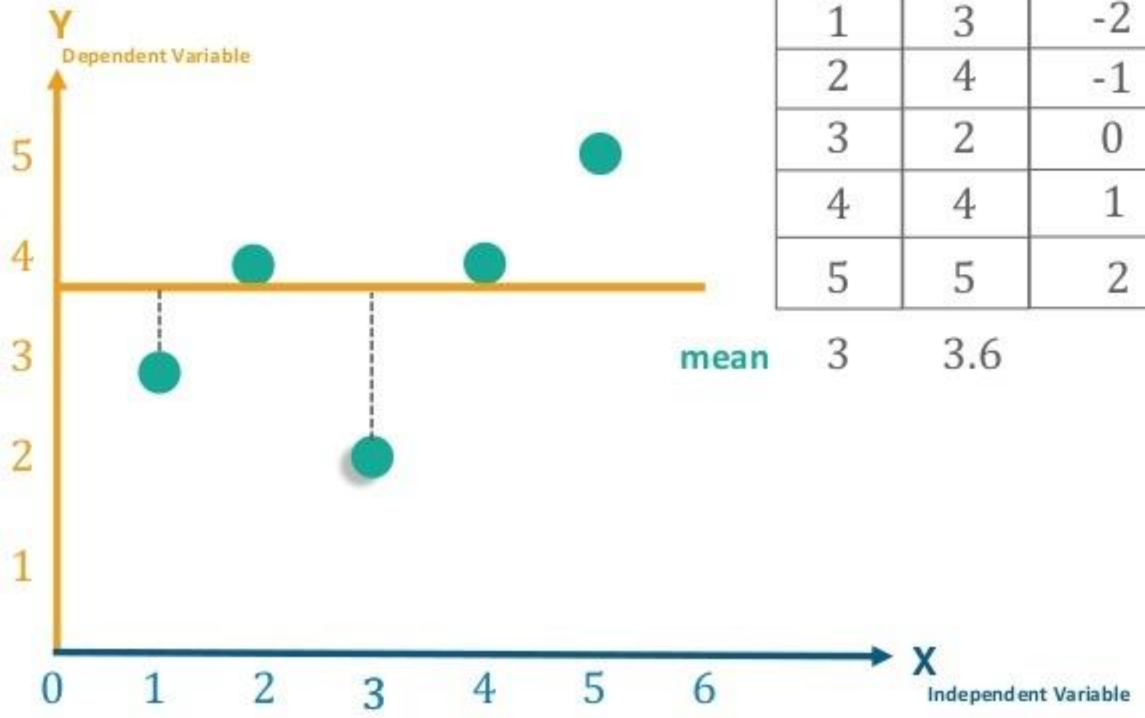
Understanding Linear Regression Algorithm



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

mean 3 3.6

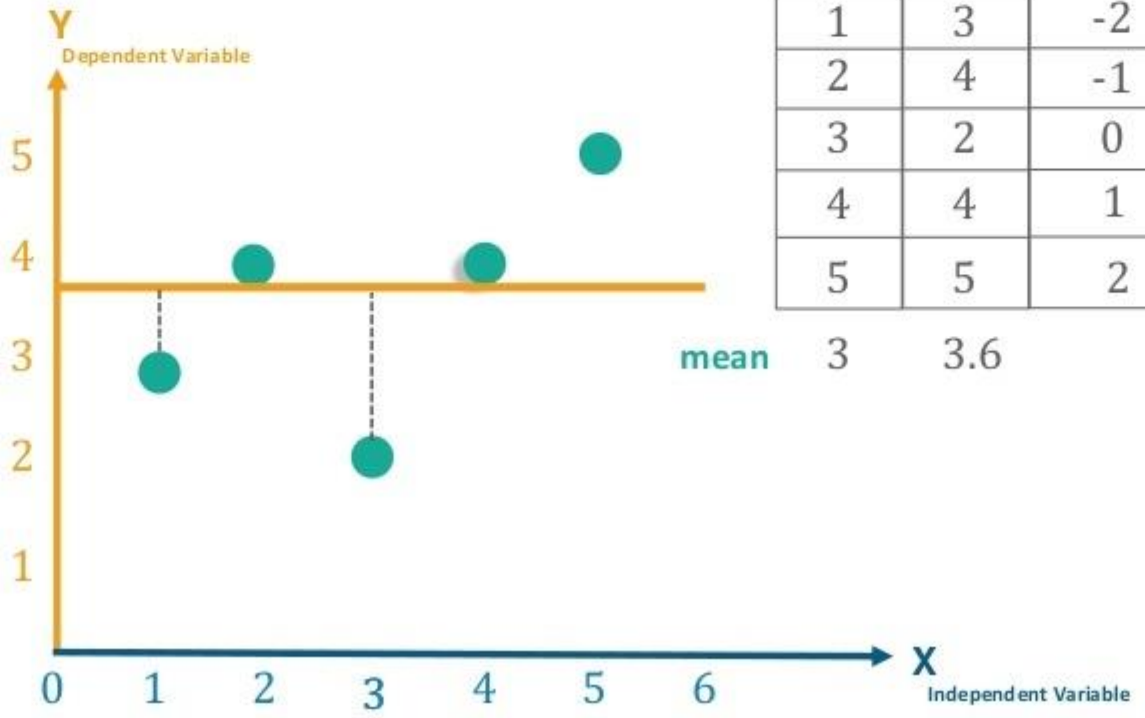
Understanding Linear Regression Algorithm



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	
5	5	2	

mean 3 3.6

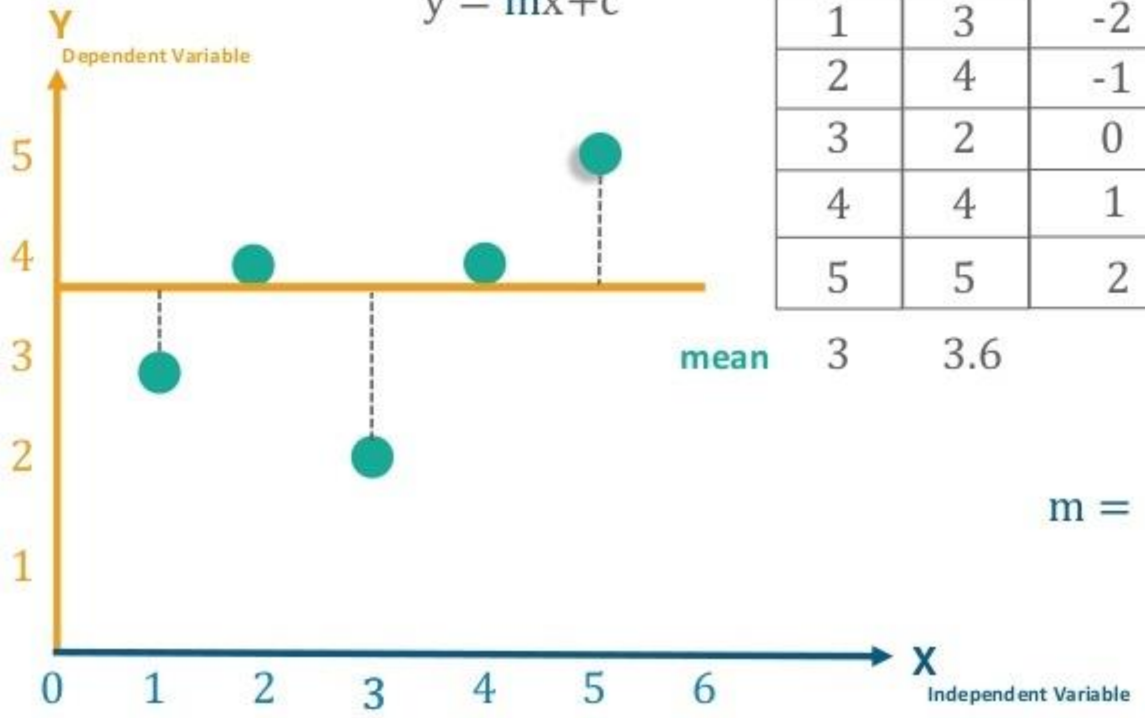
Understanding Linear Regression Algorithm



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	

mean 3 3.6

Understanding Linear Regression Algorithm

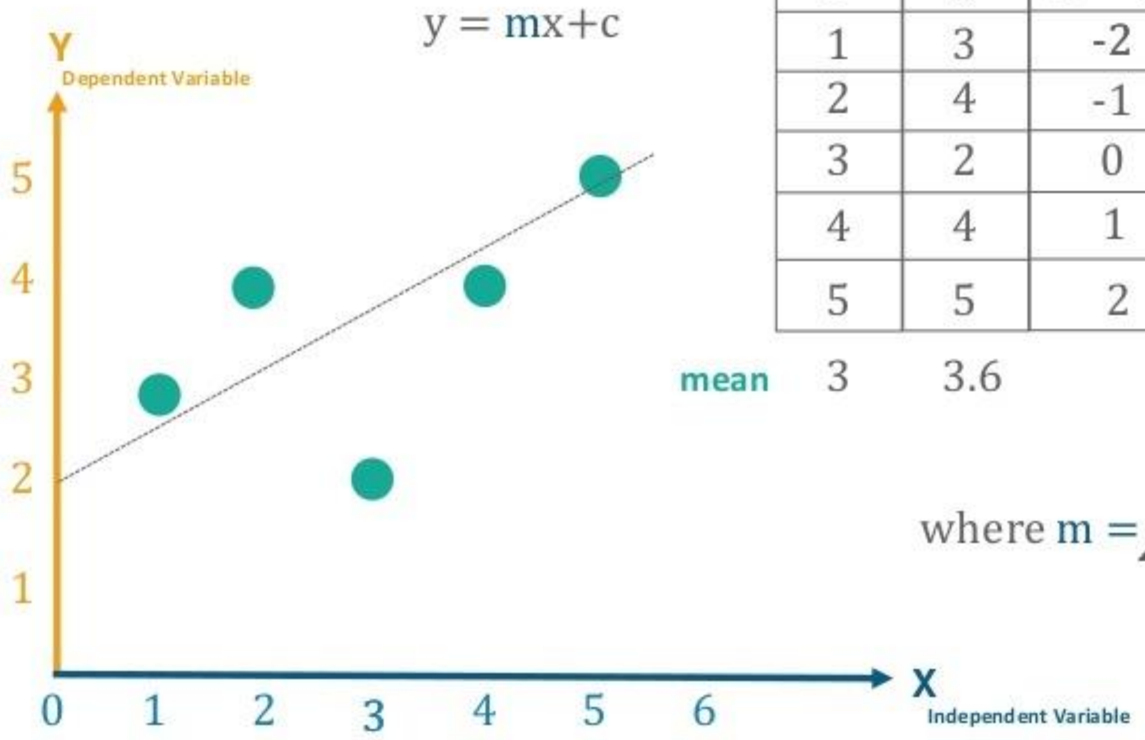


$$y = mx + c$$

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		

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

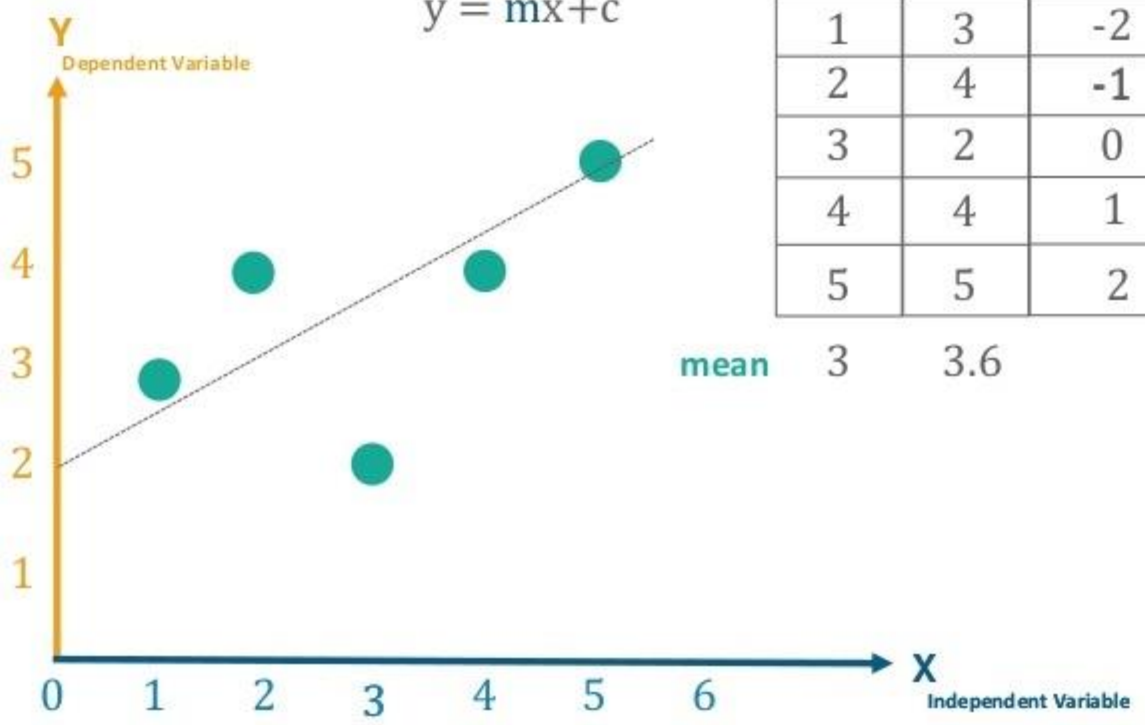
Understanding Linear Regression Algorithm



x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$
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			

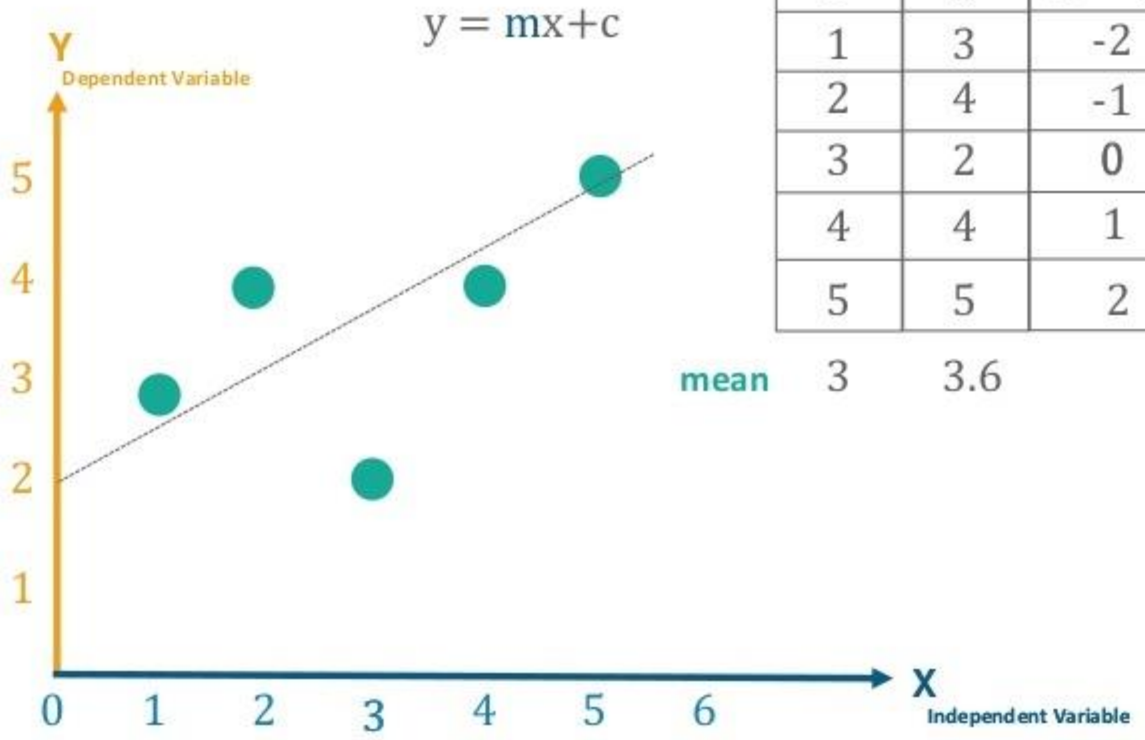
$$\text{where } m = \sum \frac{(x-\bar{x})(y-\bar{y})}{(x-\bar{x})^2}$$

Understanding Linear Regression Algorithm



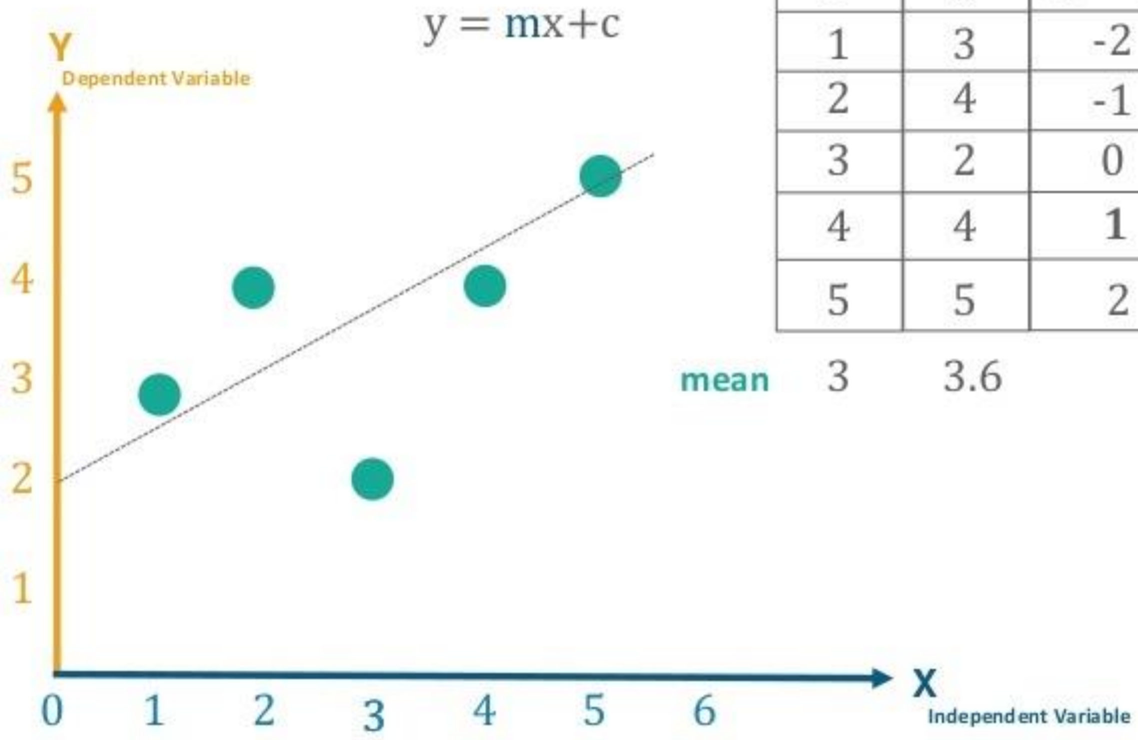
x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$
1	3	-2	-0.6	4
2	4	-1	0.4	
3	2	0	-1.6	
4	4	1	0.4	
5	5	2	1.4	
mean 3	3.6			

Understanding Linear Regression Algorithm



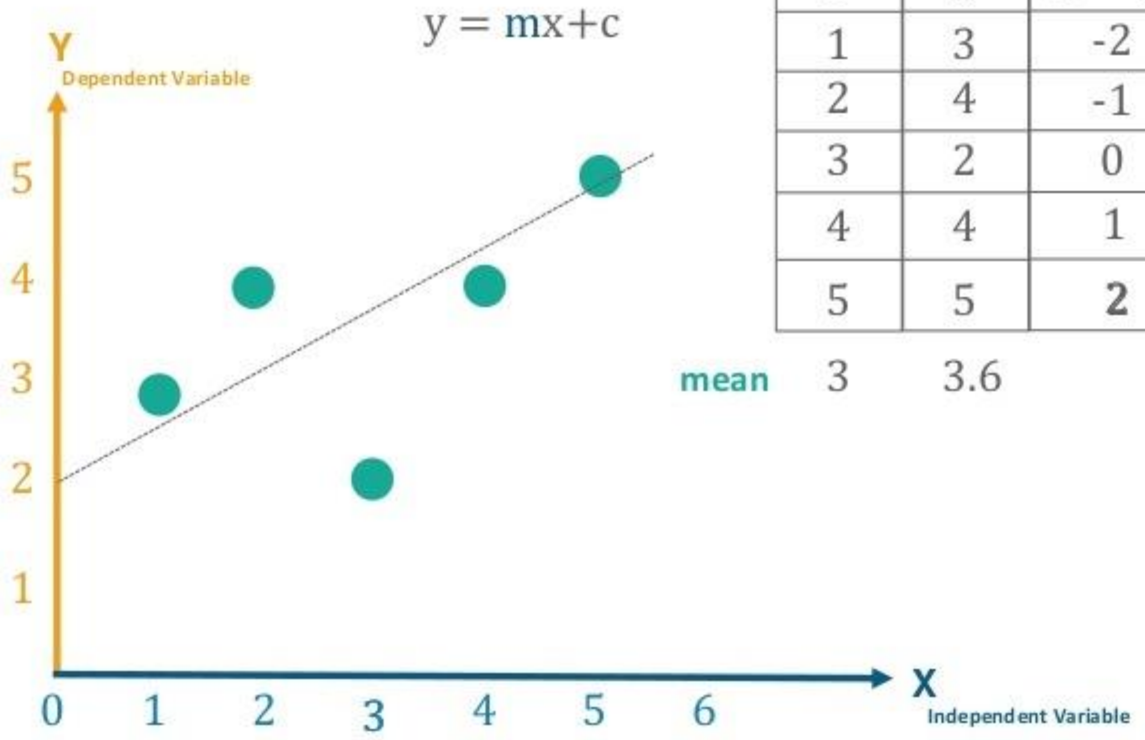
x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$
1	3	-2	-0.6	4
2	4	-1	0.4	1
3	2	0	-1.6	
4	4	1	0.4	
5	5	2	1.4	
mean	3	3.6		

Understanding Linear Regression Algorithm



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	
5	5	2	1.4	
mean 3	3.6			

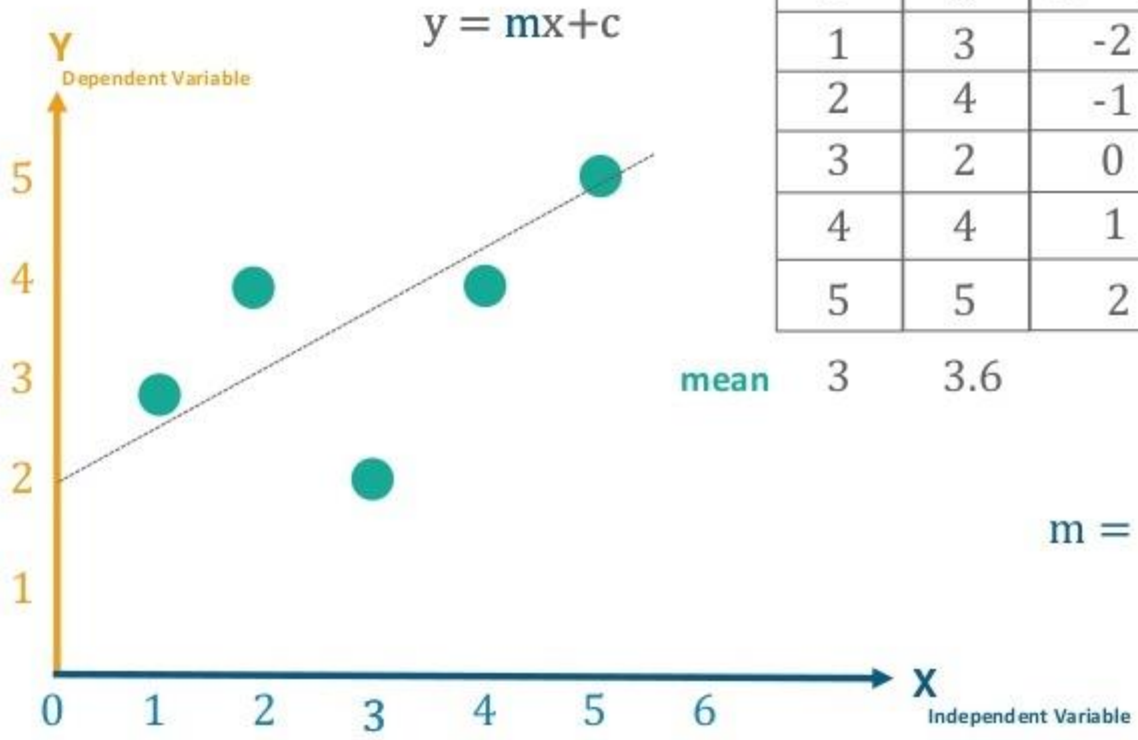
Understanding Linear Regression Algorithm



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	

mean 3 3.6

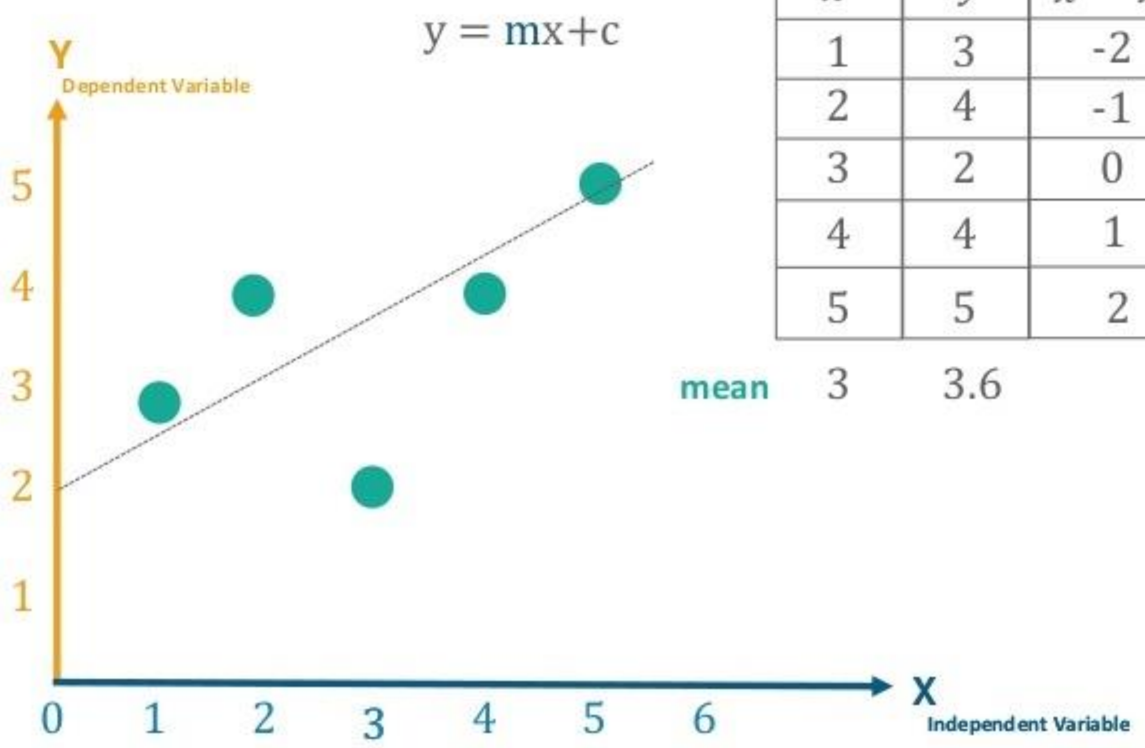
Understanding Linear Regression Algorithm



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			

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

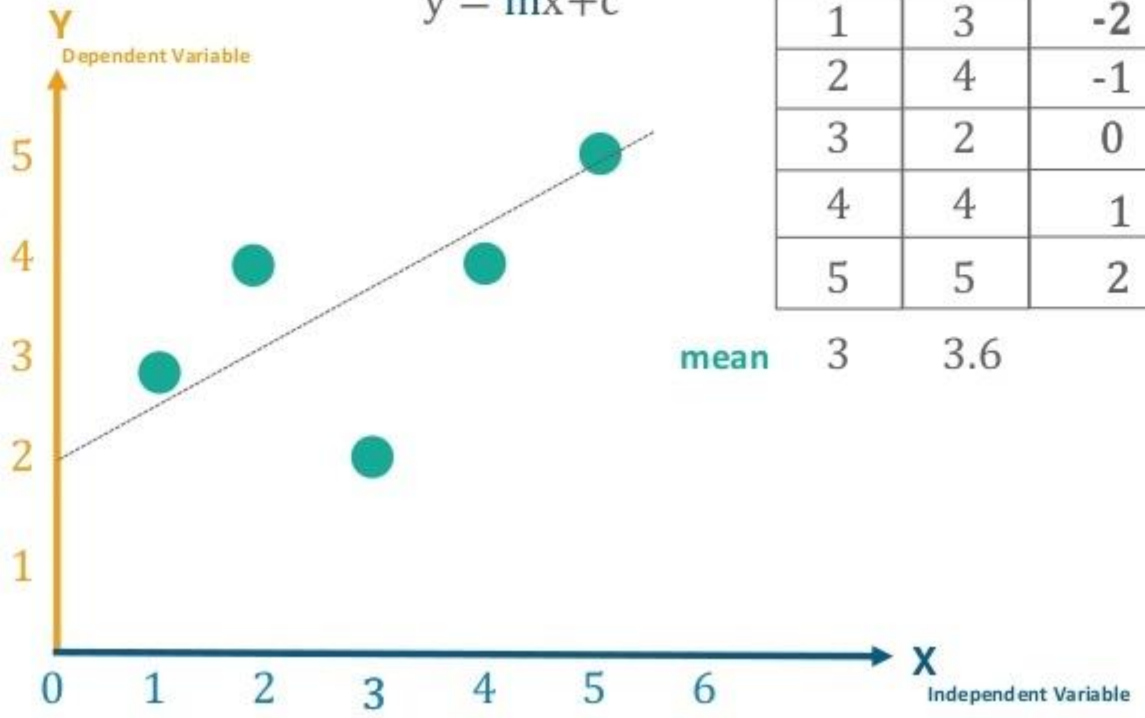
Understanding Linear Regression Algorithm



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			

Understanding Linear Regression Algorithm

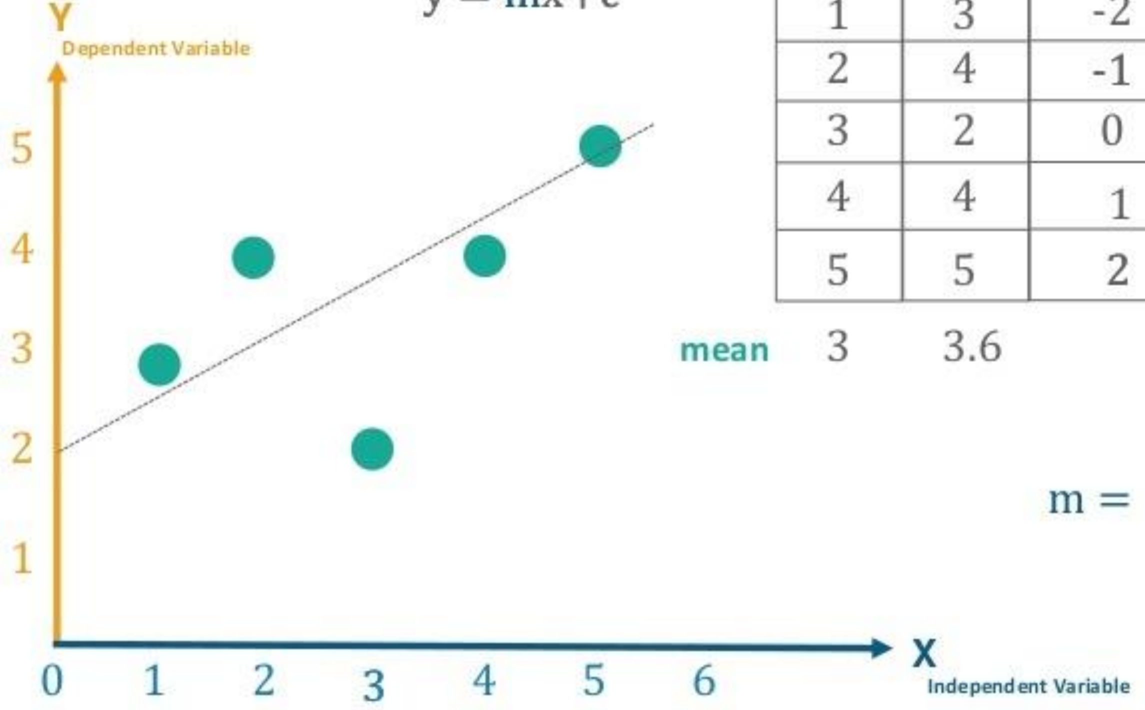
$y = mx+c$



x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$
1	3	-2	-0.6	4	
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				

Understanding Linear Regression Algorithm

$y = mx + c$

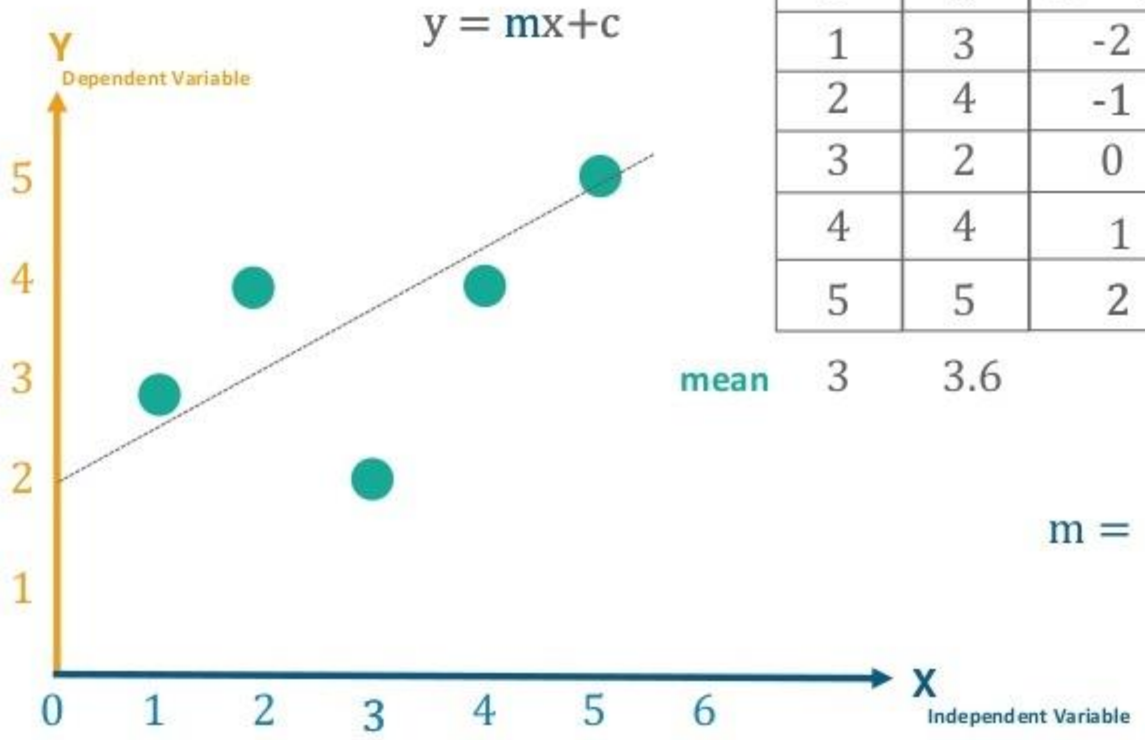


x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(x - \bar{x})(y - \bar{y}$
1	3	-2	-0.6	4	1.2
2	4	-1	0.4	1	-0.4
3	2	0	-1.6	0	0
4	4	1	0.4	1	0.4
5	5	2	1.4	4	2.8

mean 3 3.6

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

Understanding Linear Regression Algorithm

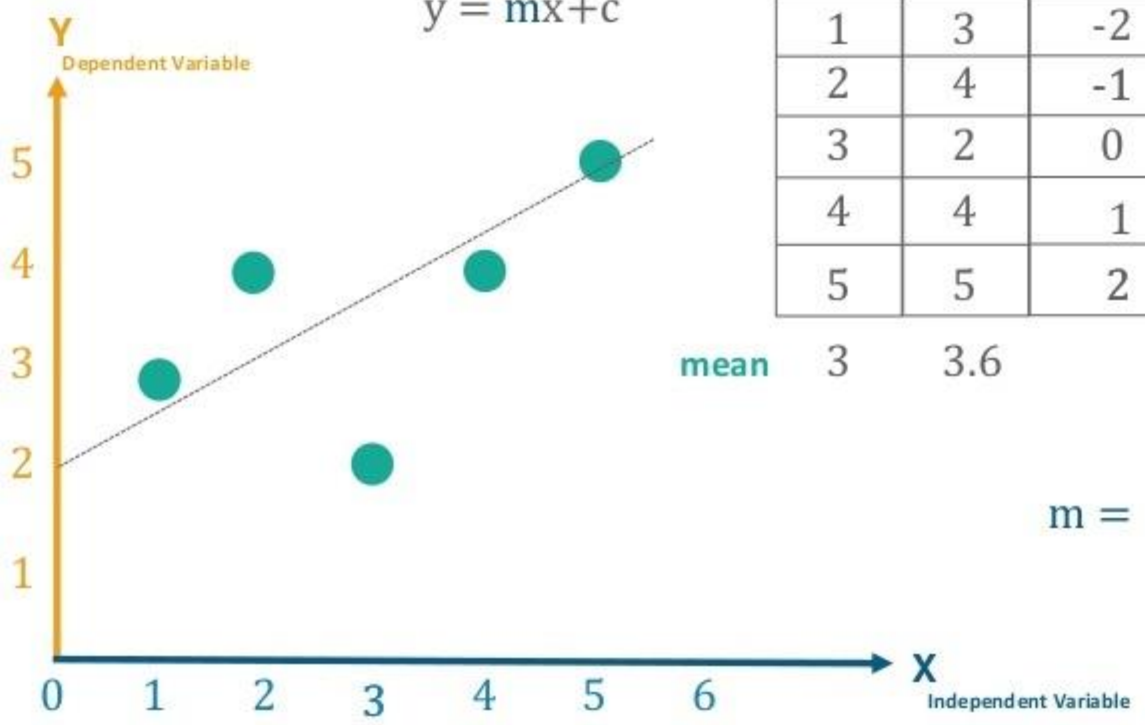


	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y}$
	1	3	-2	-0.6	4	1.2
	2	4	-1	0.4	1	-0.4
	3	2	0	-1.6	0	0
	4	4	1	0.4	1	0.4
	5	5	2	1.4	4	2.8
mean	3	3.6			$\Sigma = 10$	$\Sigma = 4$

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

Understanding Linear Regression Algorithm

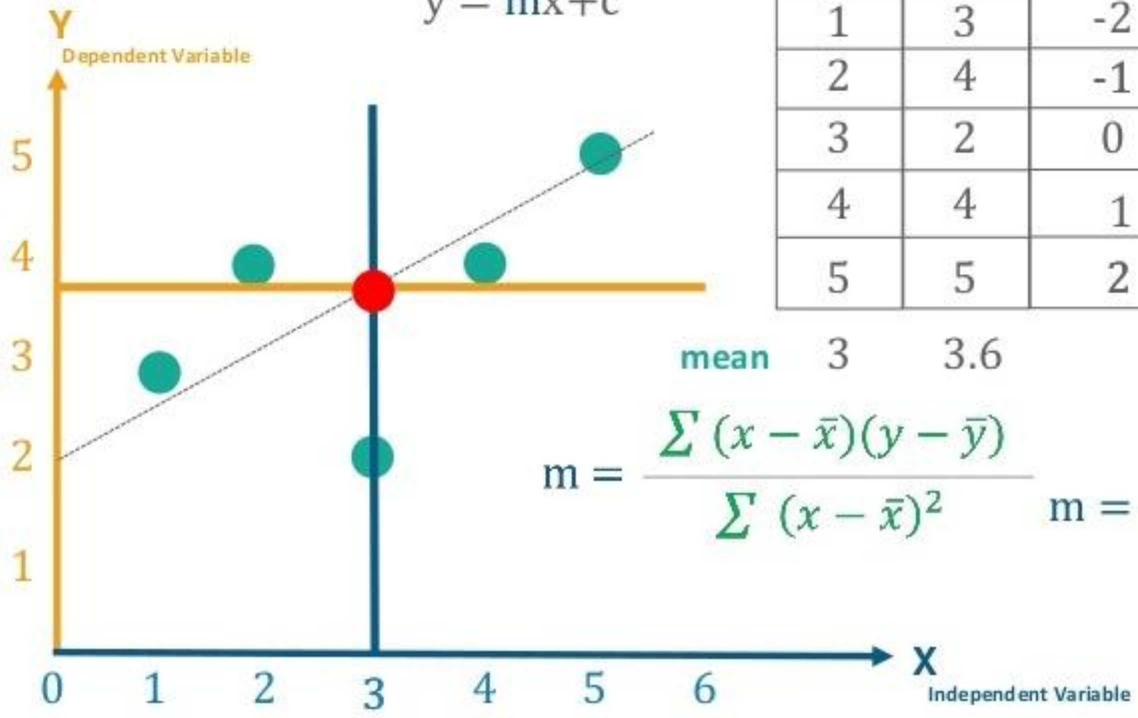
$y = mx + c$



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			$\Sigma = 10$	$\Sigma = 4$

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

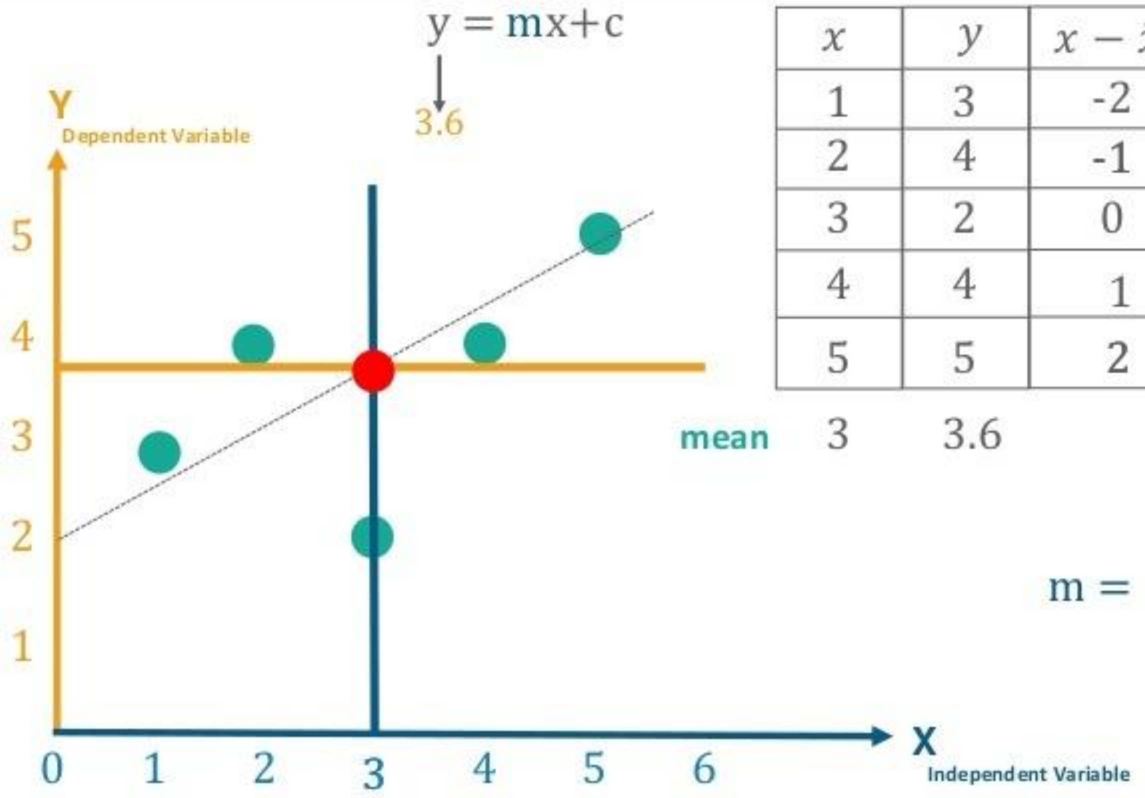
Understanding Linear Regression Algorithm



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			$\Sigma = 10$	$\Sigma = 4$

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

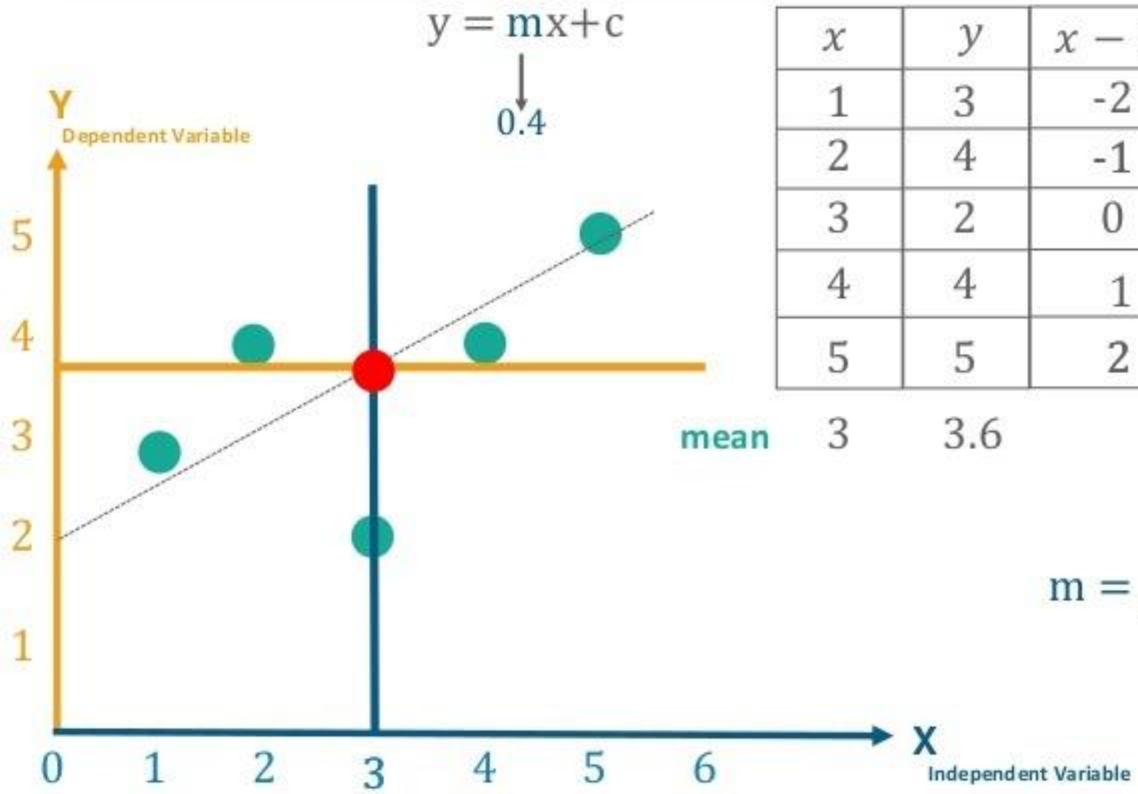
Understanding Linear Regression Algorithm



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			$\Sigma = 10$	$\Sigma = 4$

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

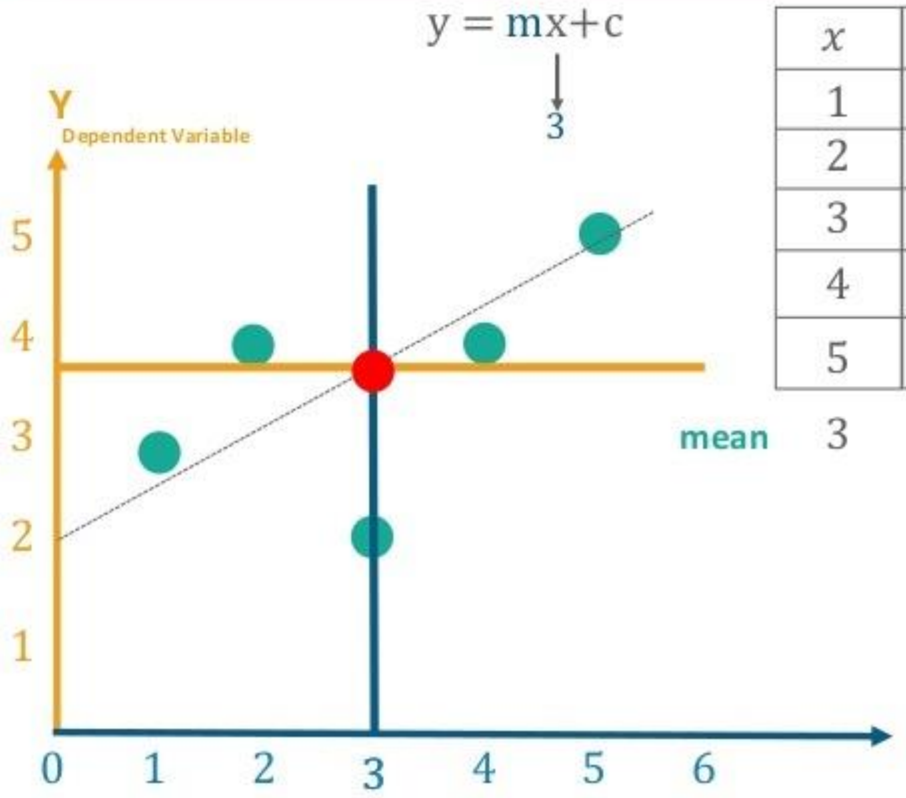
Understanding Linear Regression Algorithm



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			$\Sigma = 10$	$\Sigma = 4$

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

Understanding Linear Regression Algorithm



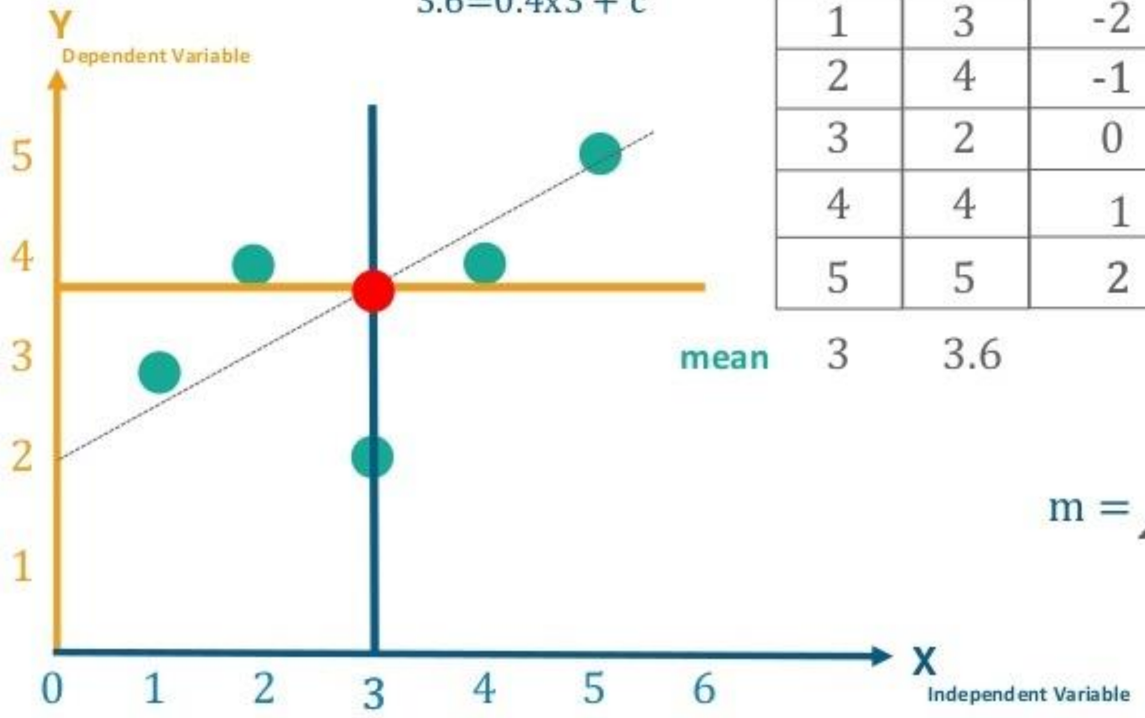
x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y}$
1	3	-2	-0.6	4	1.2
2	4	-1	0.4	1	-0.4
3	2	0	-1.6	0	0
4	4	1	0.4	1	0.4
5	5	2	1.4	4	2.8
mean 3	3.6			$\Sigma = 10$	$\Sigma = 4$

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

Understanding Linear Regression Algorithm

$y = mx + c$

$3.6 = 0.4x3 + c$

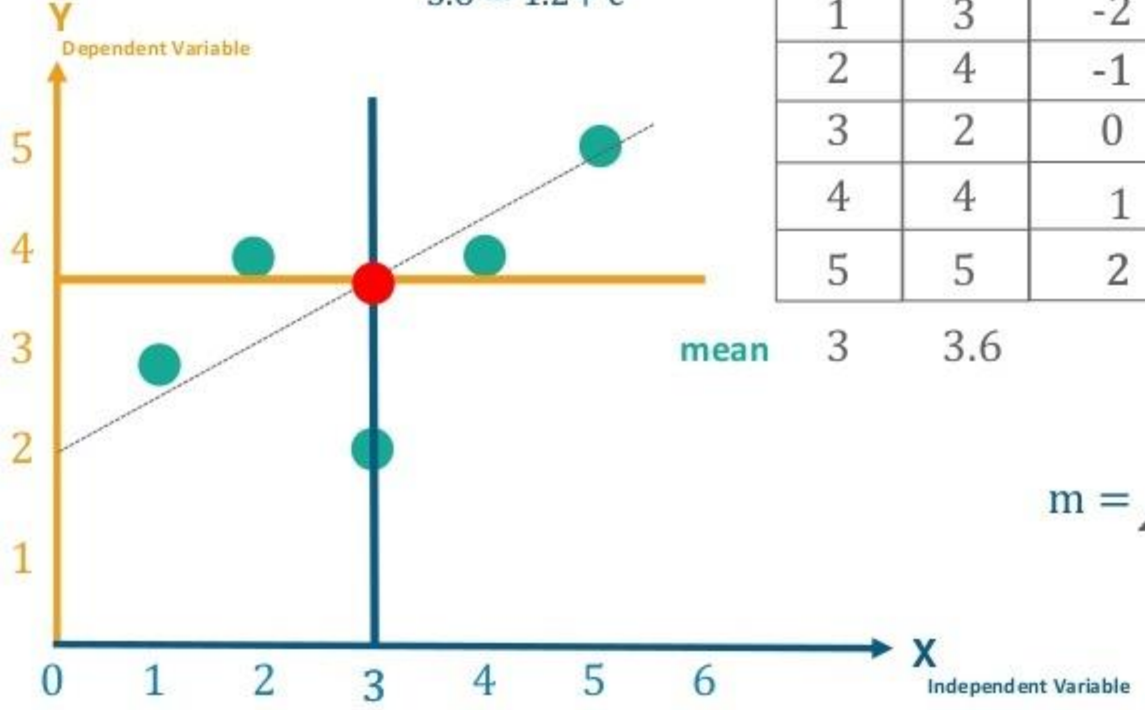


x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$
1	3	-2	-0.6	4	1.2
2	4	-1	0.4	1	-0.4
3	2	0	-1.6	0	0
4	4	1	0.4	1	0.4
5	5	2	1.4	4	2.8
mean 3	3.6			$\Sigma = 10$	$\Sigma = 4$

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

Understanding Linear Regression Algorithm

$$y = mx + c$$
$$3.6 = 1.2 + c$$



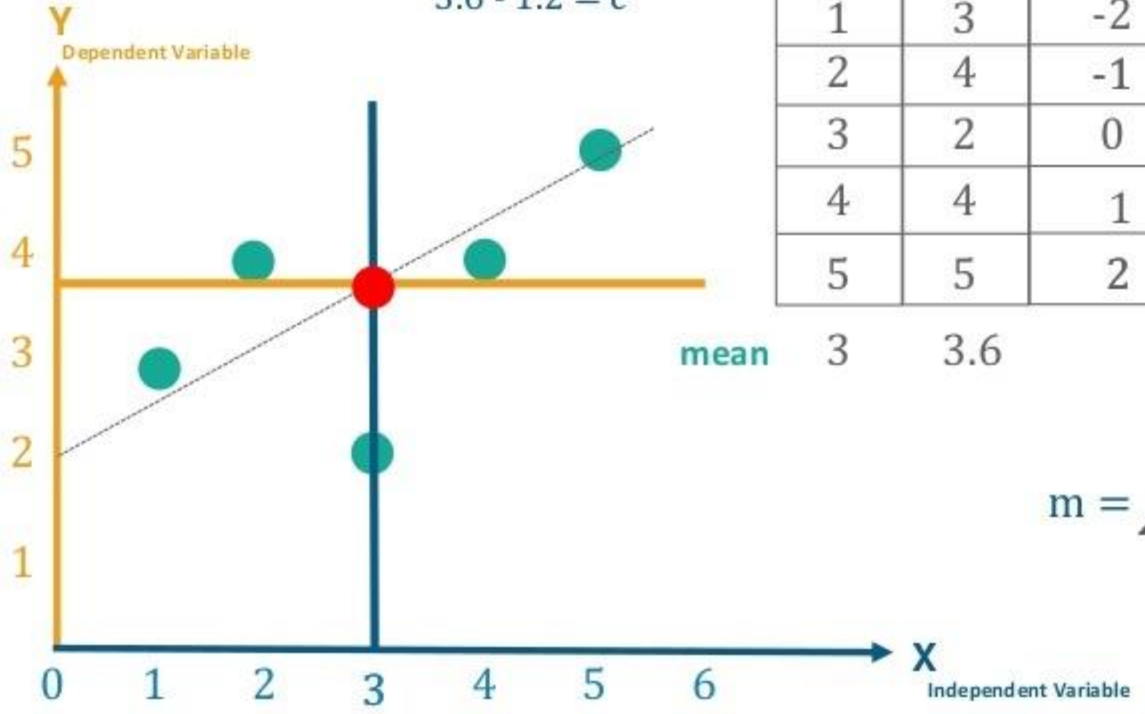
x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$
1	3	-2	-0.6	4	1.2
2	4	-1	0.4	1	-0.4
3	2	0	-1.6	0	0
4	4	1	0.4	1	0.4
5	5	2	1.4	4	2.8
mean 3	3.6			$\Sigma = 10$	$\Sigma = 4$

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

Understanding Linear Regression Algorithm

$$y = mx + c$$

$$3.6 - 1.2 = c$$



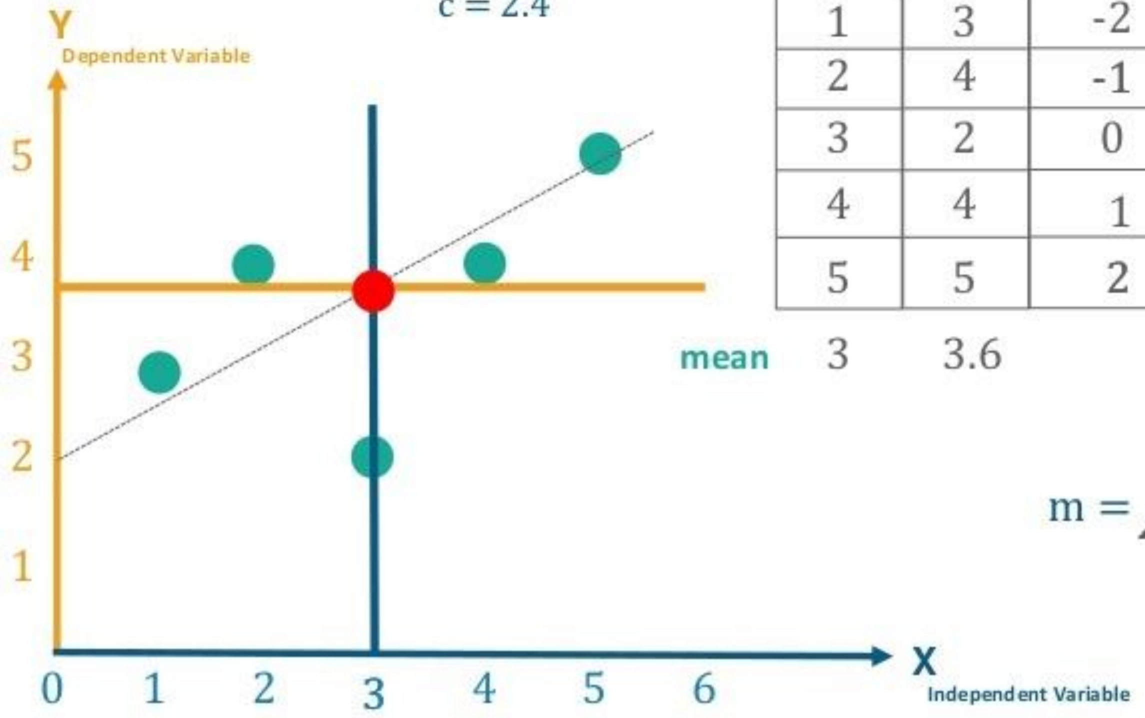
	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y}$
	1	3	-2	-0.6	4	1.2
	2	4	-1	0.4	1	-0.4
	3	2	0	-1.6	0	0
	4	4	1	0.4	1	0.4
	5	5	2	1.4	4	2.8
mean	3	3.6			$\Sigma = 10$	$\Sigma = 4$

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

Understanding Linear Regression Algorithm

$y = mx + c$

$c = 2.4$



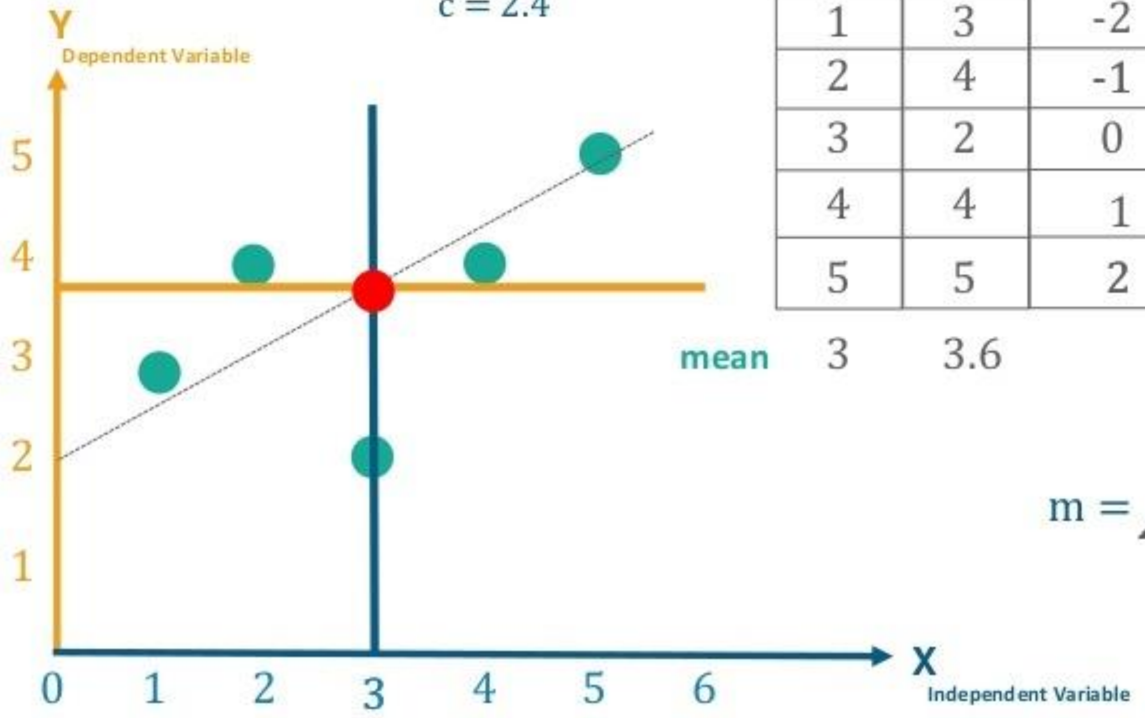
x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$
1	3	-2	-0.6	4	1.2
2	4	-1	0.4	1	-0.4
3	2	0	-1.6	0	0
4	4	1	0.4	1	0.4
5	5	2	1.4	4	2.8
mean 3	3.6			$\Sigma = 10$	$\Sigma = 4$

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

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			$\Sigma = 10$	$\Sigma = 4$

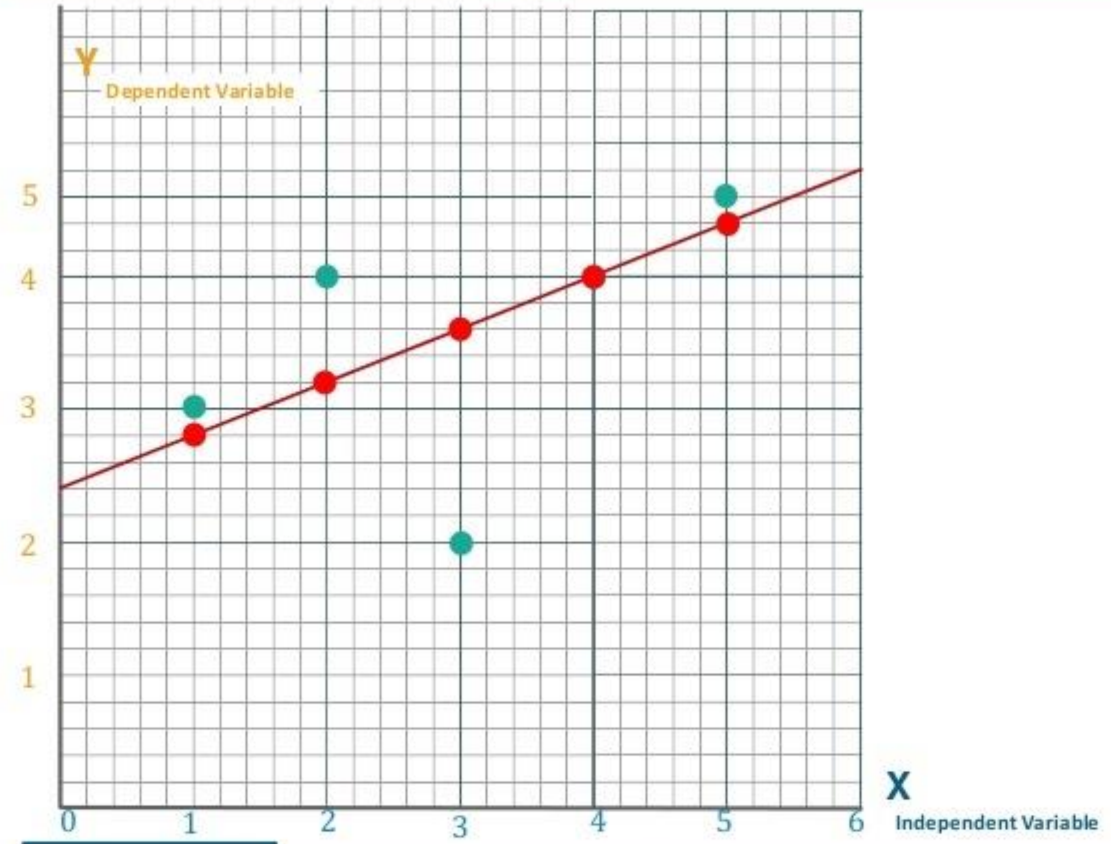
$$m = \sum \frac{(x-\bar{x})(y-\bar{y})}{(x-\bar{x})^2} = \frac{4}{10}$$

$m = 0.4$

$c = 2.4$

$y = 0.4x + 2.4$

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 $x = \{1,2,3,4,5\}$

$y = 0.4 \times 1 + 2.4 = 2.8$

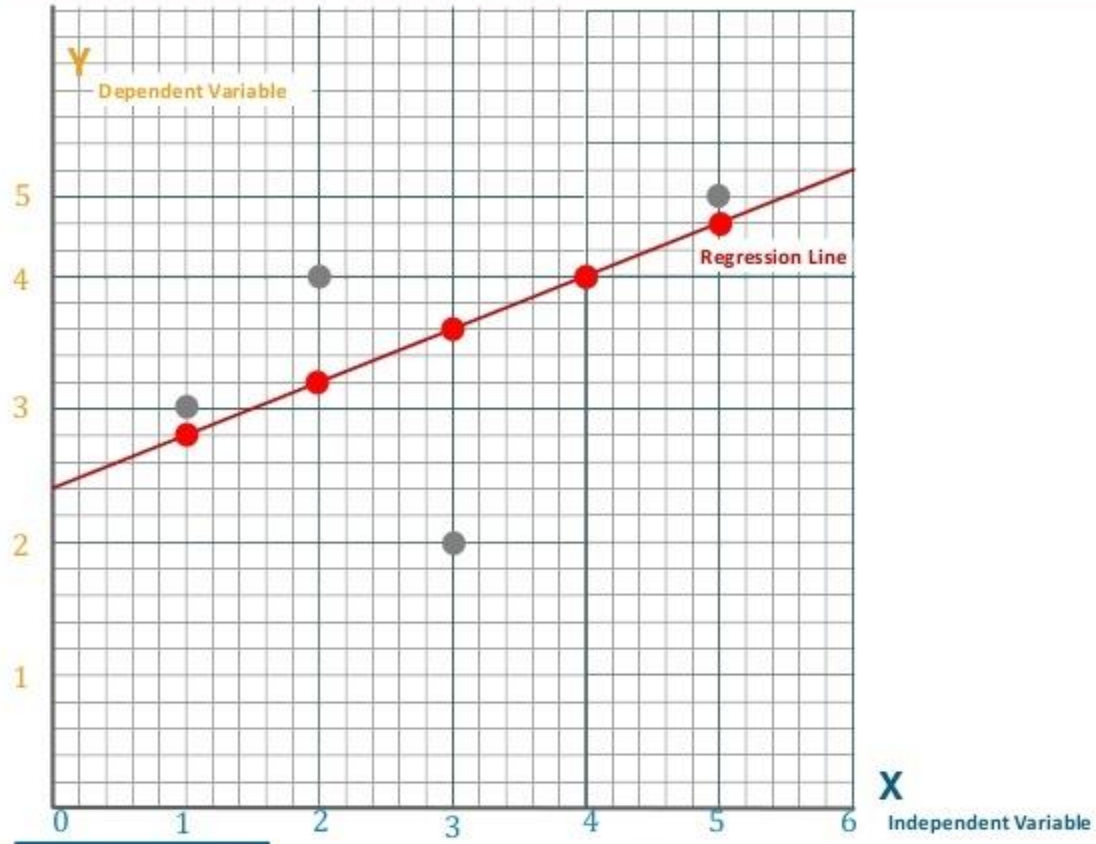
$y = 0.4 \times 2 + 2.4 = 3.2$

$y = 0.4 \times 3 + 2.4 = 3.6$

$y = 0.4 \times 4 + 2.4 = 4.0$

$y = 0.4 \times 5 + 2.4 = 4.4$

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 $x = \{1,2,3,4,5\}$

$y = 0.4 \times 1 + 2.4 = 2.8$

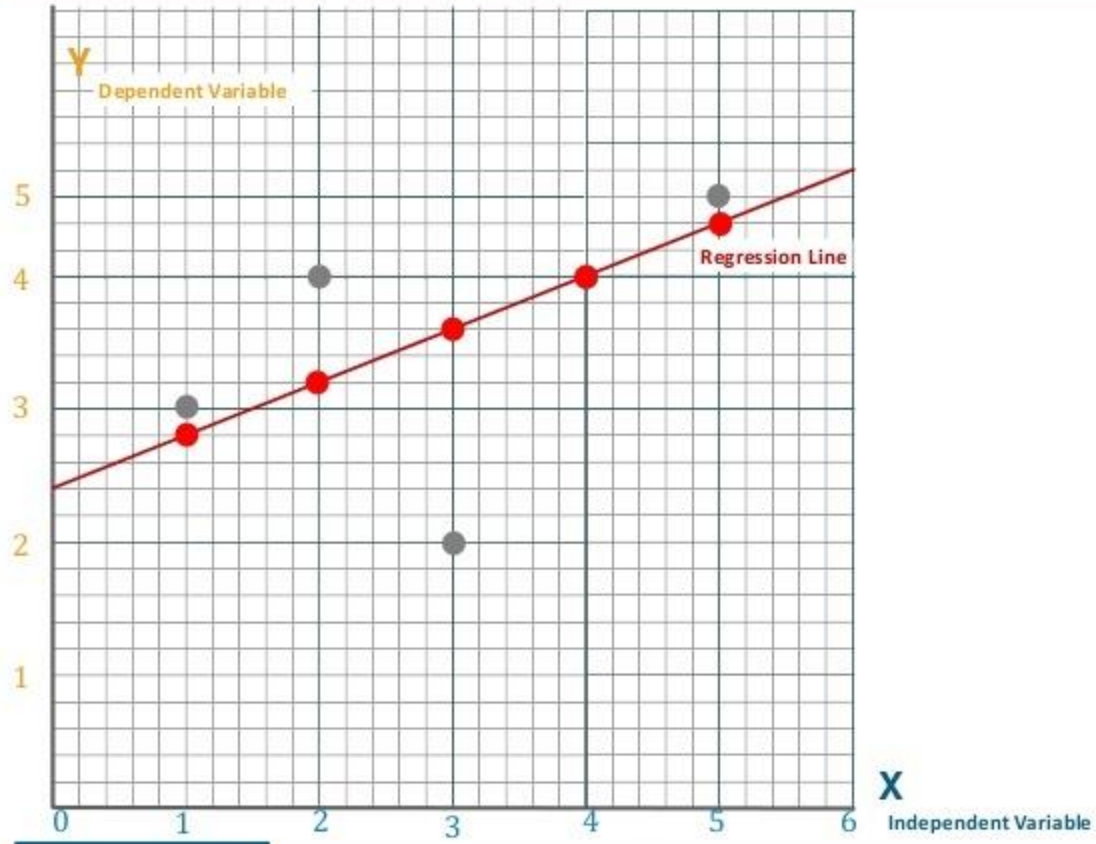
$y = 0.4 \times 2 + 2.4 = 3.2$

$y = 0.4 \times 3 + 2.4 = 3.6$

$y = 0.4 \times 4 + 2.4 = 4.0$

$y = 0.4 \times 5 + 2.4 = 4.4$

Mean Square Error

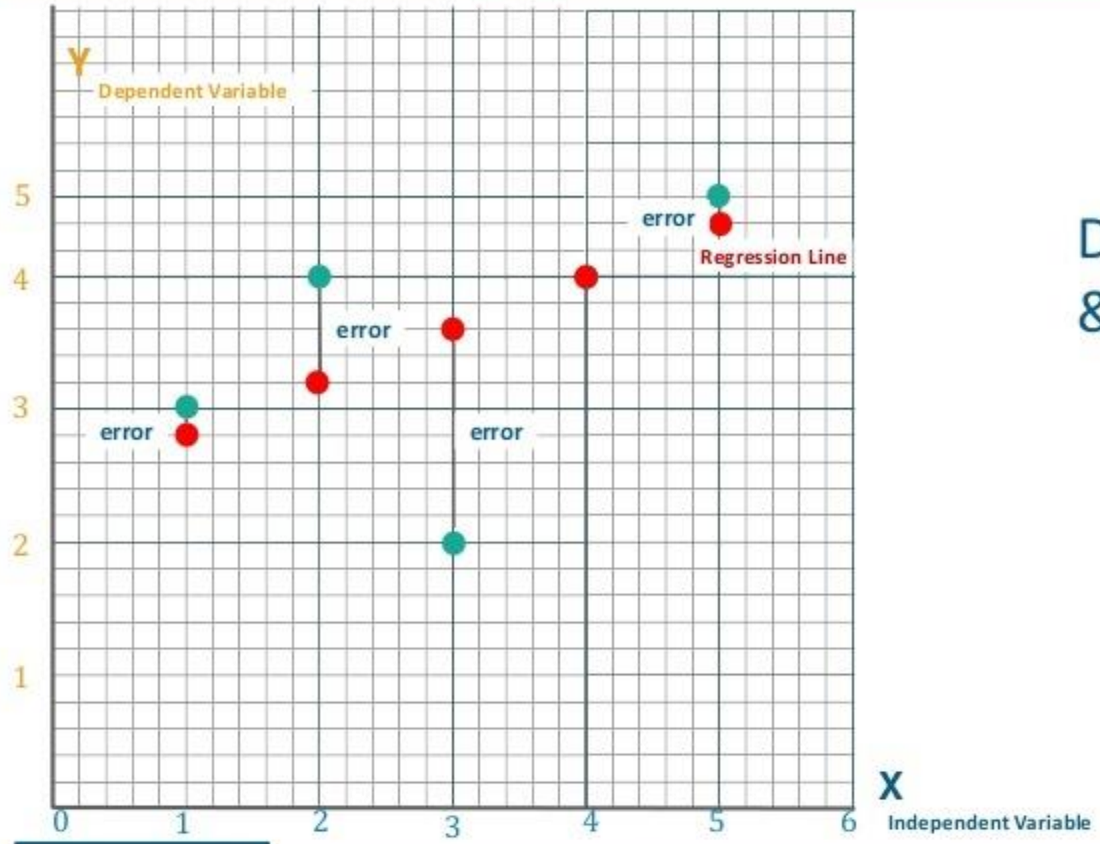


$m = 0.4$
$c = 2.4$
$y = 0.4x + 2.4$

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

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

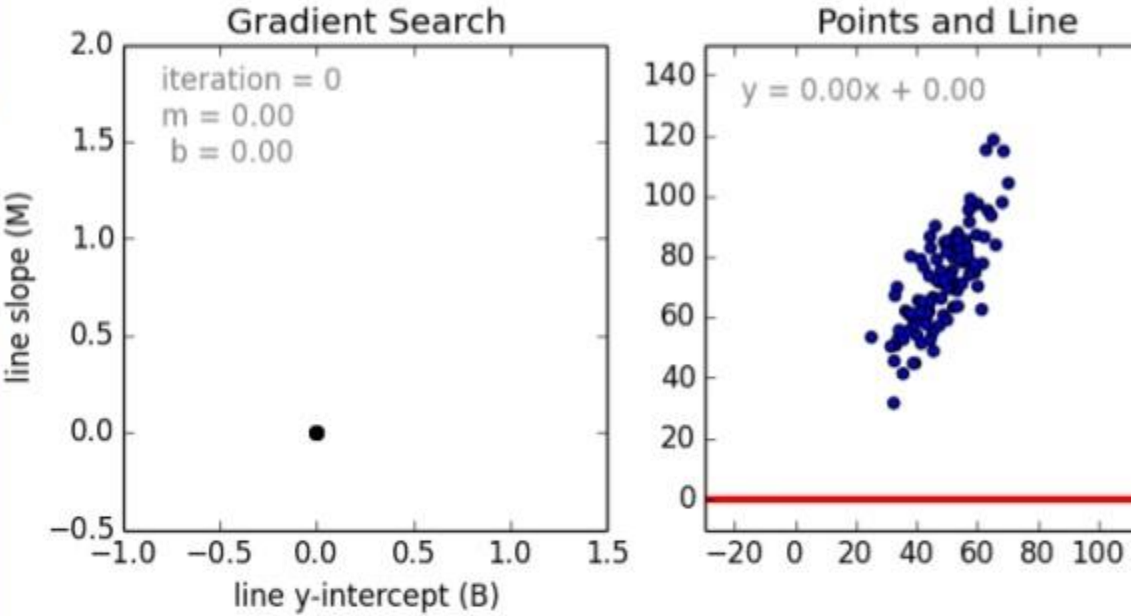
Mean Square Error



Distance between actual & predicted value

edureka!

Finding the best fit line



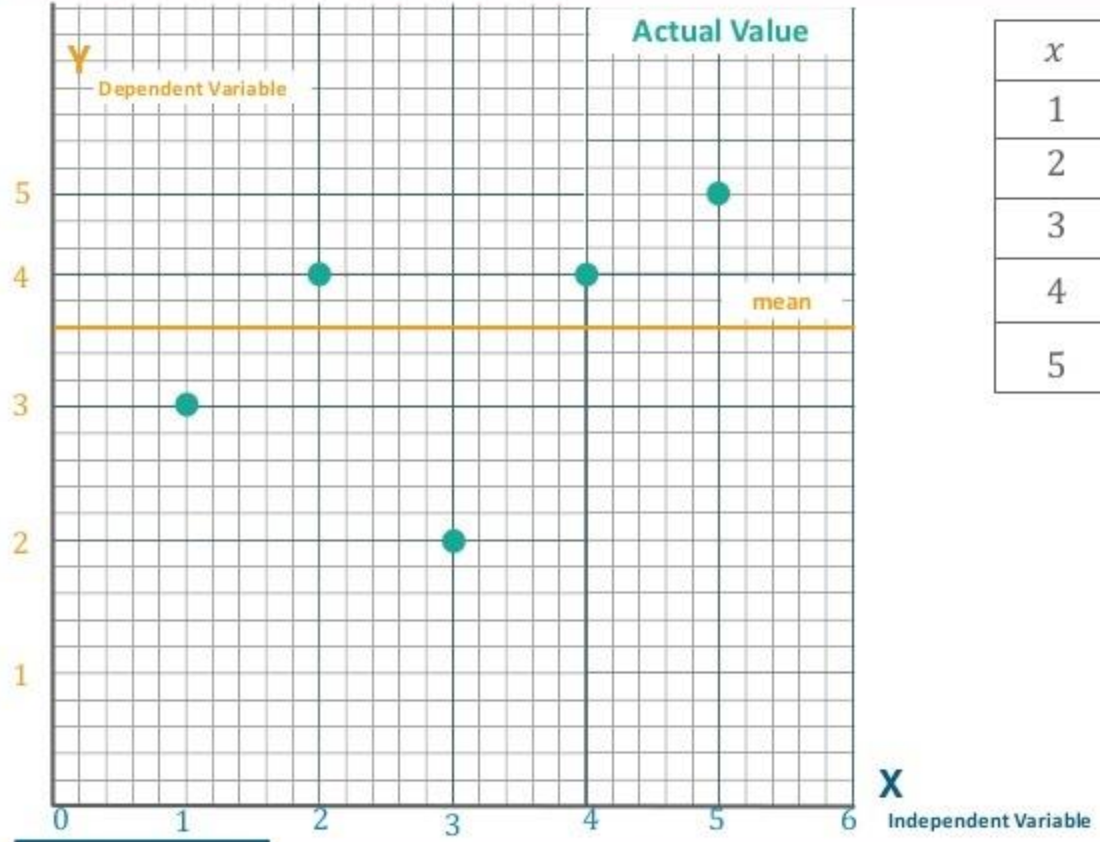
Let’s check the Goodness of fit

edureka!

What is R-Square?

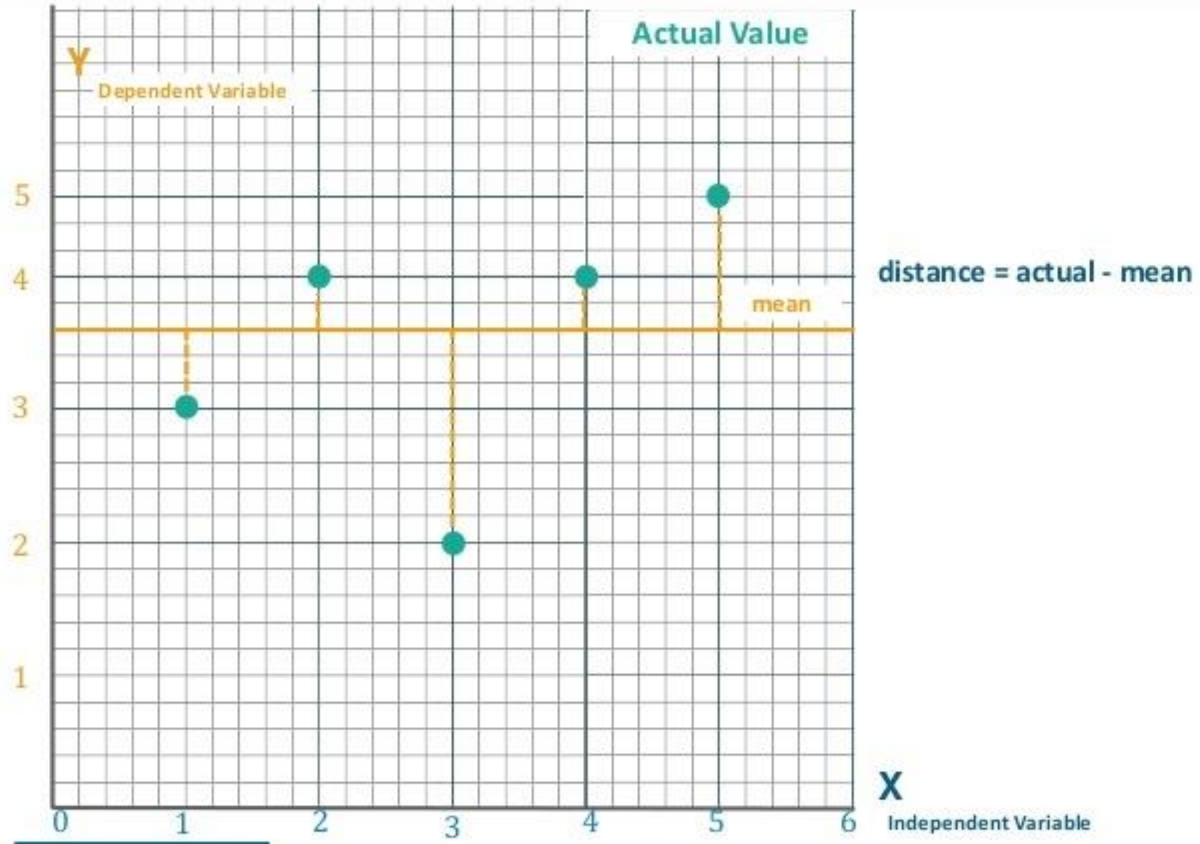
- **R-squared** value is a statistical measure of how close the data are to the fitted regression line
- It is also known as **coefficient of determination**, or the **coefficient of multiple determination**

Calculation of R^2

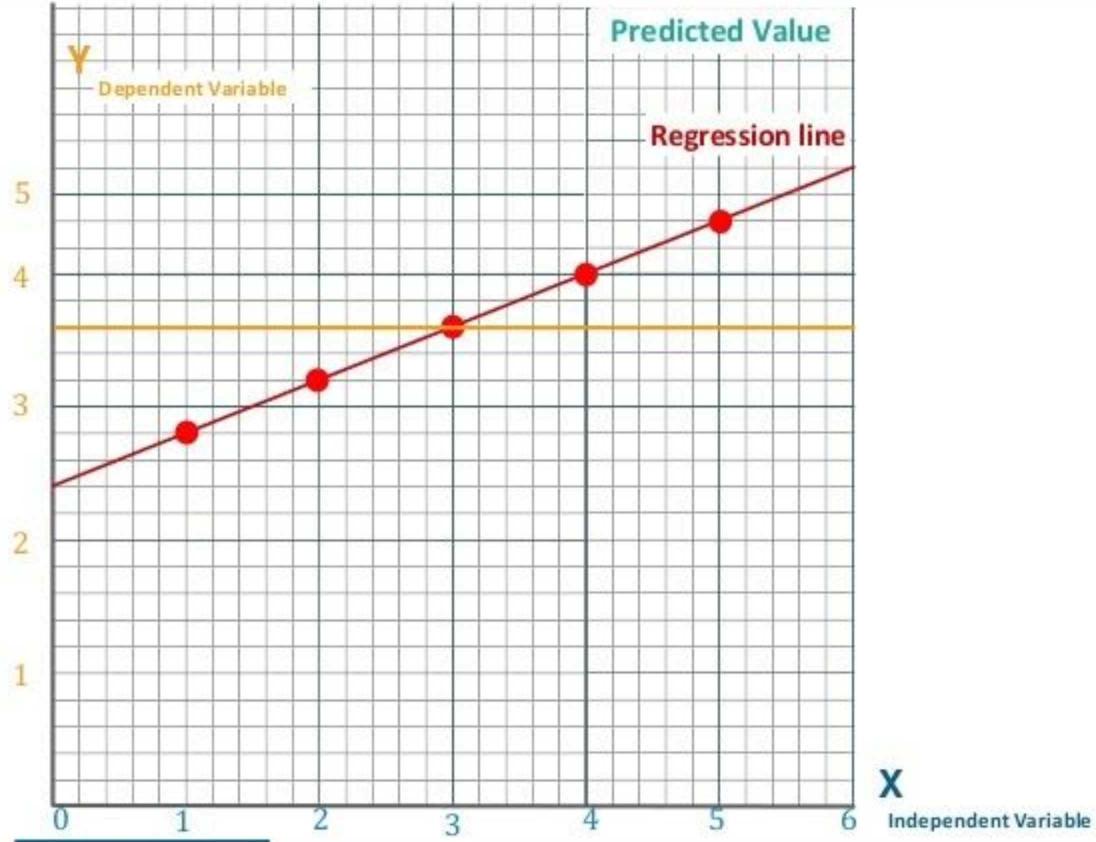


x	y
1	3
2	4
3	2
4	4
5	5

Calculation of R^2

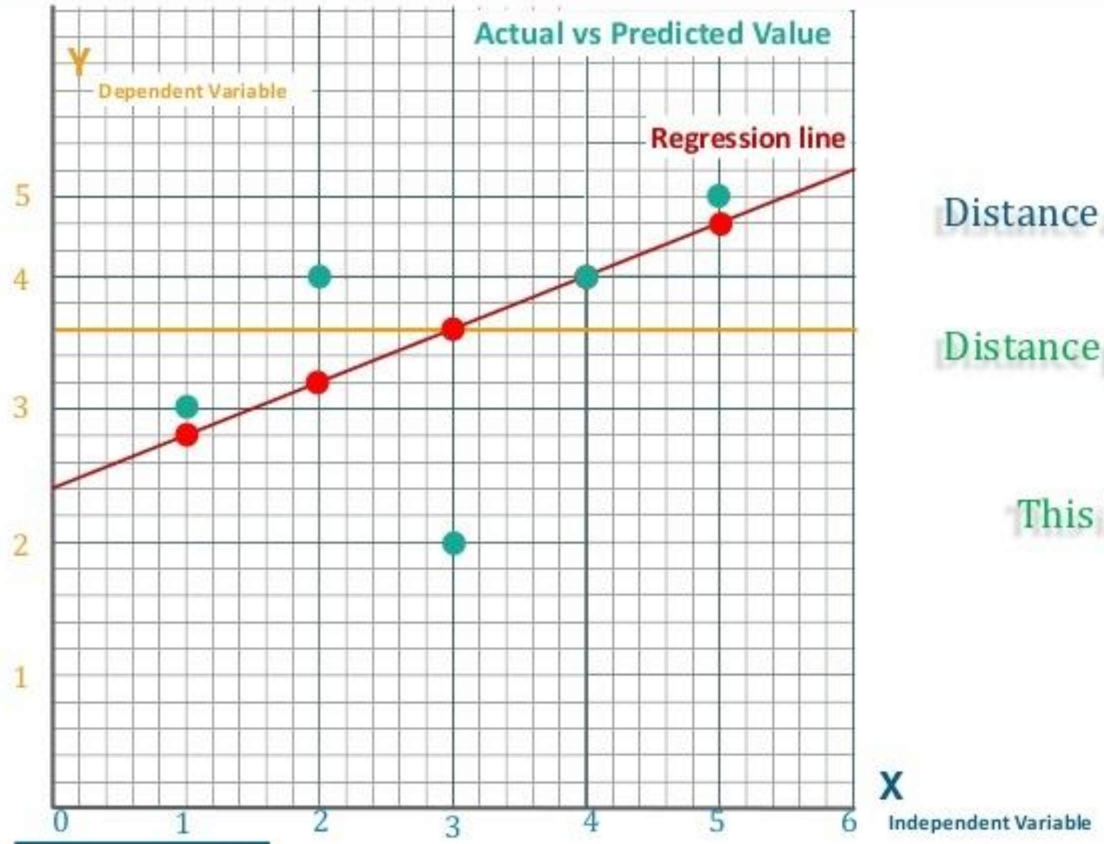


Calculation of R^2



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

Calculation of R^2



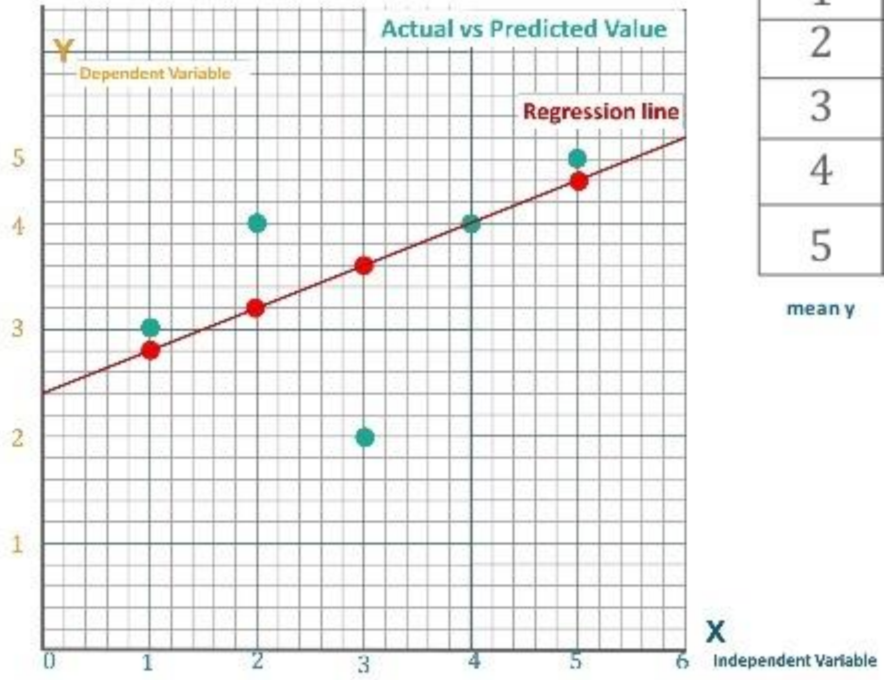
Distance actual - mean

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

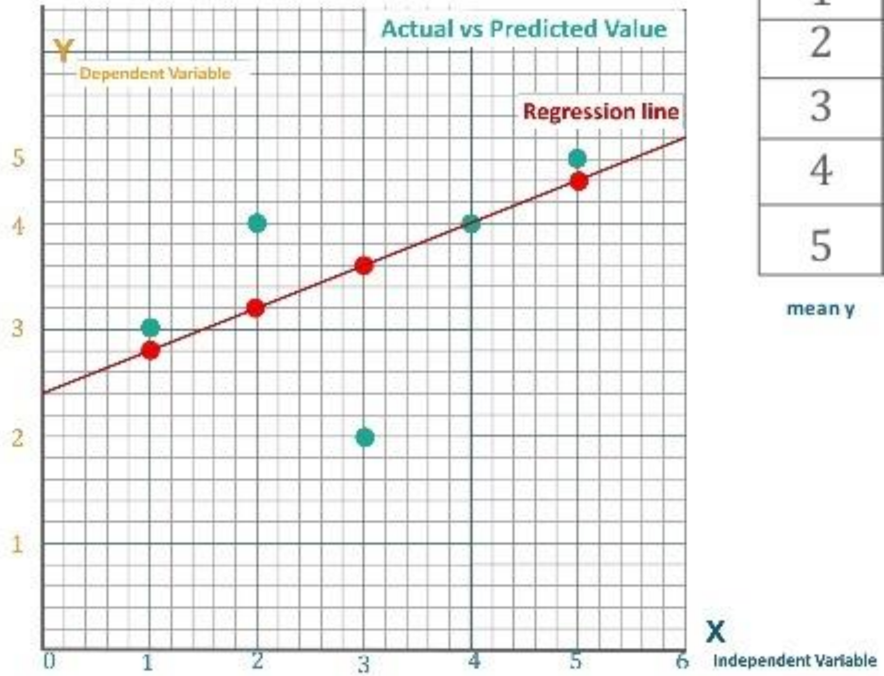


x	y	$y - \bar{y}$
1	3	– 0.6
2	4	0.4
3	2	–1.6
4	4	0.4
5	5	1.4

mean y 3.6

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

Calculation of R^2

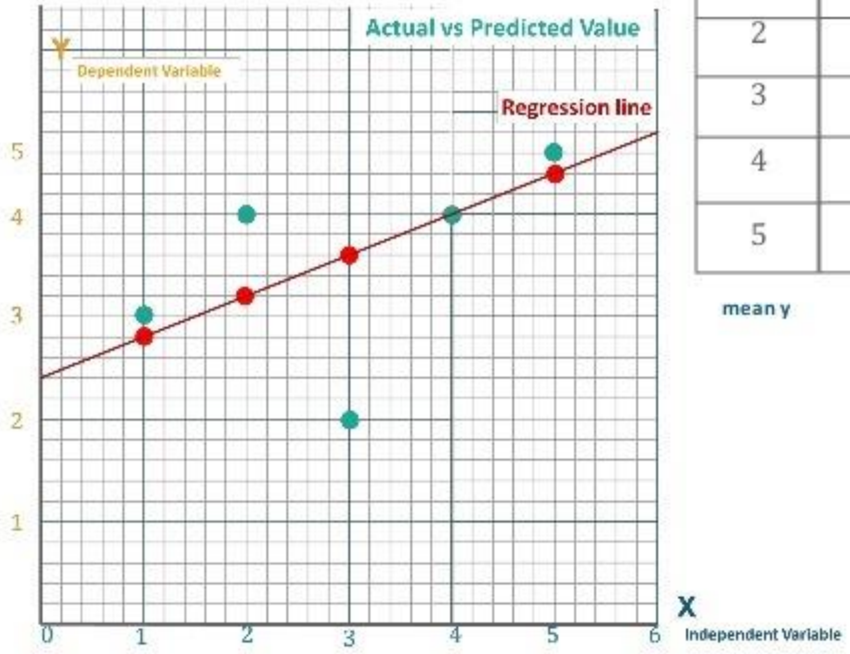


x	y	$y-\bar{y}$	$(y-\bar{y})^2$	y_p	$(y_p - \bar{y})$
1	3	– 0.6	3.6	2.8	-0.8
2	4	0.4	1.6	3.2	-0.4
3	2	−1.6	2.56	3.6	0
4	4	0.4	1.6	4.0	0.4
5	5	1.4	1.96	4.4	0.8

mean y 3.6

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

Calculation of R^2

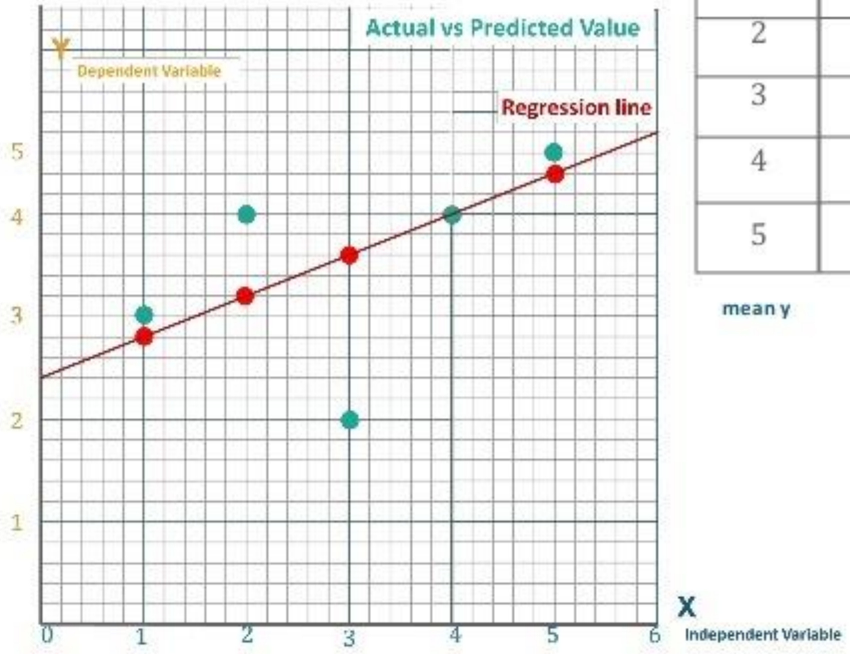


x	y	$y-\bar{y}$	$(y-\bar{y})^2$	y_p	$(y_p - \bar{y})$	$(y_p - \bar{y}$
1	3	- 0.6	3.6	2.8	-0.8	6.4
2	4	0.4	1.6	3.2	-0.4	1.6
3	2	−1.6	2.56	3.6	0	0
4	4	0.4	1.6	4.0	0.4	1.6
5	5	1.4	1.96	4.4	0.8	6.4

mean y 3.6

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

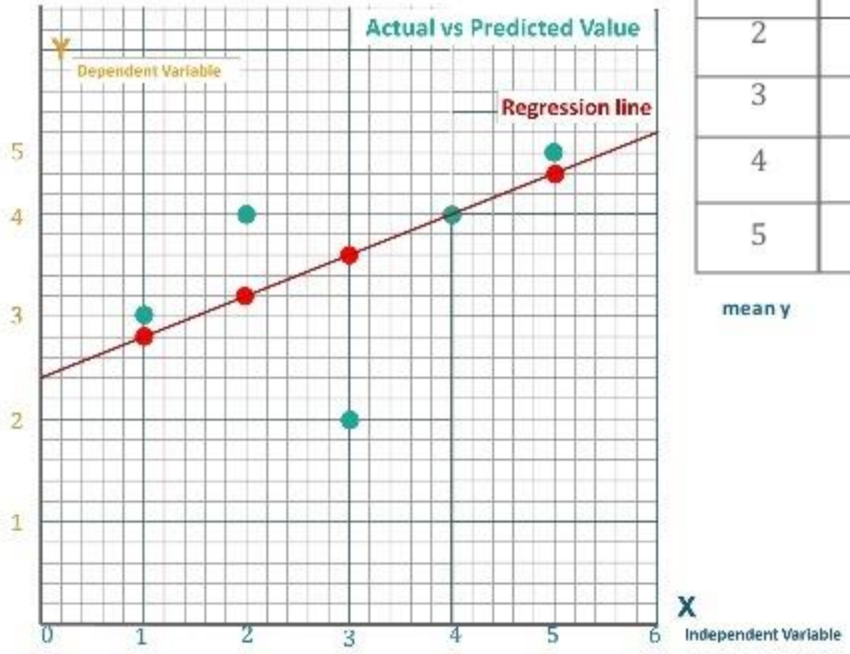
Calculation of R^2



x	y	$y-\bar{y}$	$(y-\bar{y})^2$	y_p	$(y_p - \bar{y})$	$(y_p - \bar{y}$
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

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

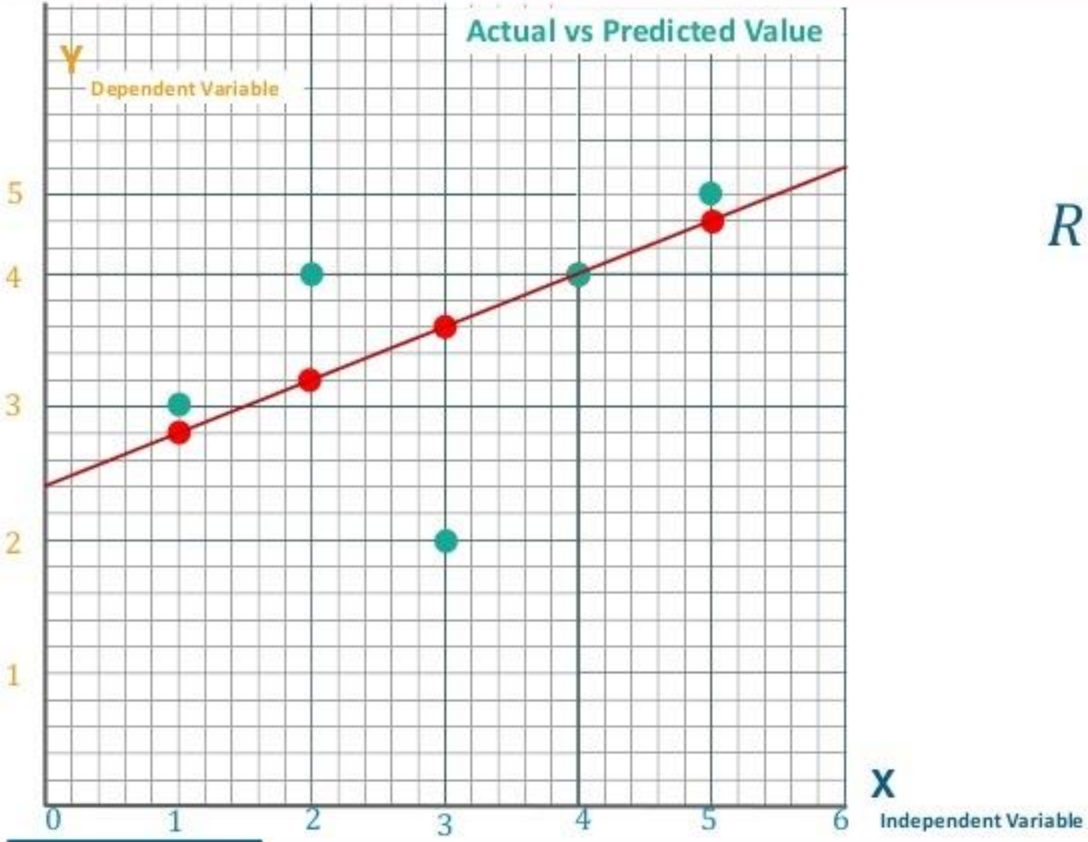
Calculation of R^2



x	y	$y-\bar{y}$	$(y-\bar{y})^2$	y_p	$(y_p - \bar{y})$	$((y_p -$
1	3	- 0.6	3.6	2.8	-0.8	6.4
2	4	0.4	1.6	3.2	-0.4	1.6
3	2	−1.6	2.56	3.6	0	0
4	4	0.4	1.6	4.0	0.4	1.6
5	5	1.4	1.96	4.4	0.8	6.4
mean y	3.6		11.32			16

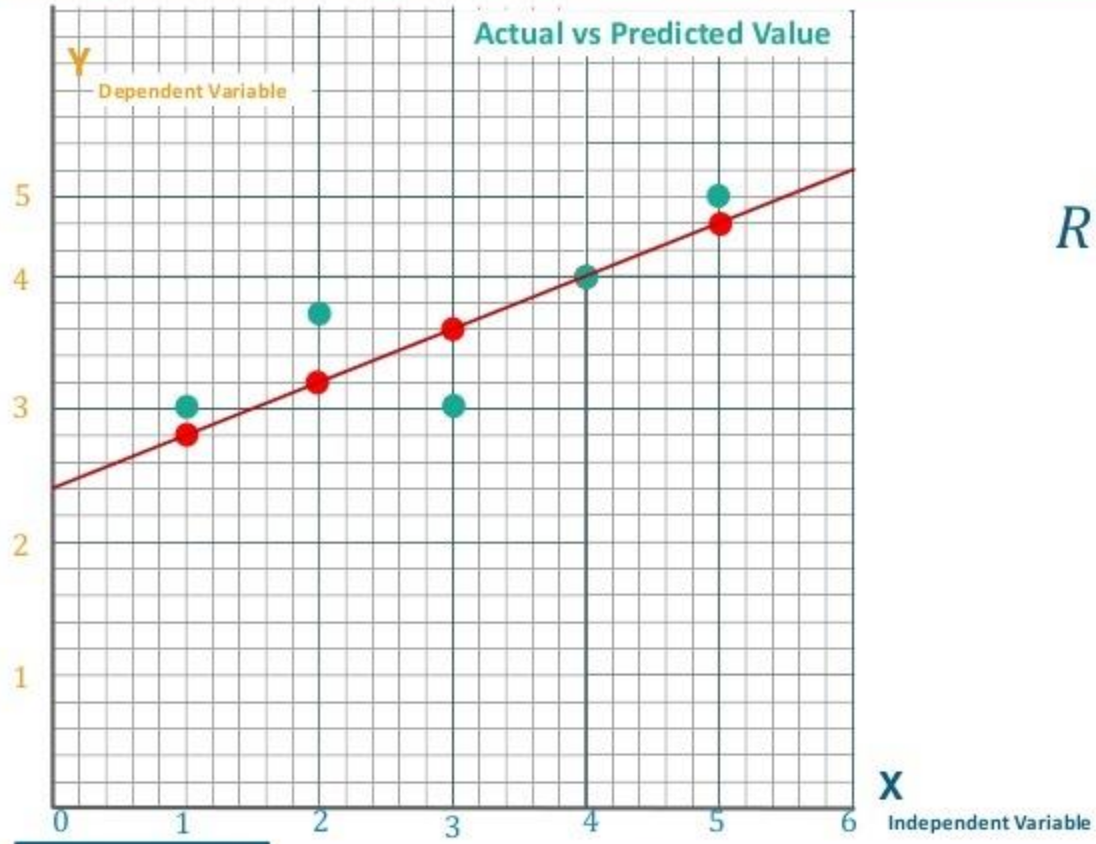
$$R^2 \approx 0.3$$

Calculation of R^2



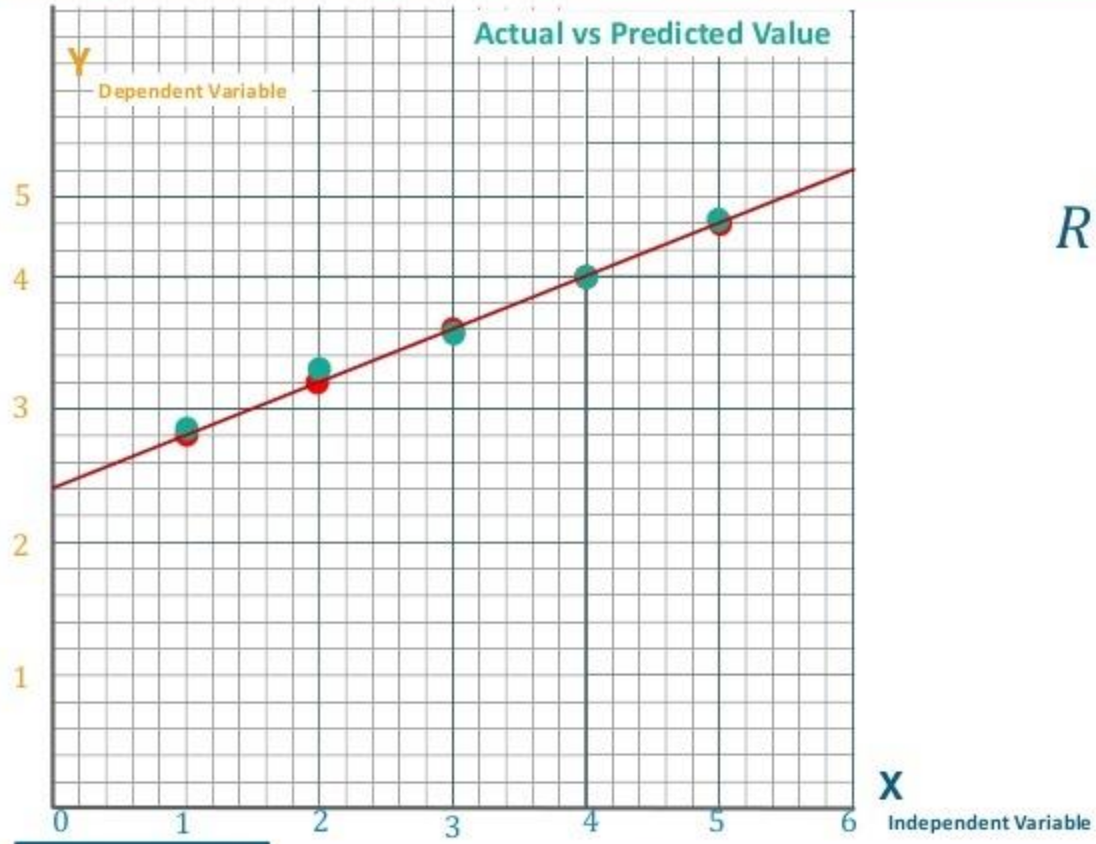
$$R^2 \approx 0.3$$

Calculation of R^2



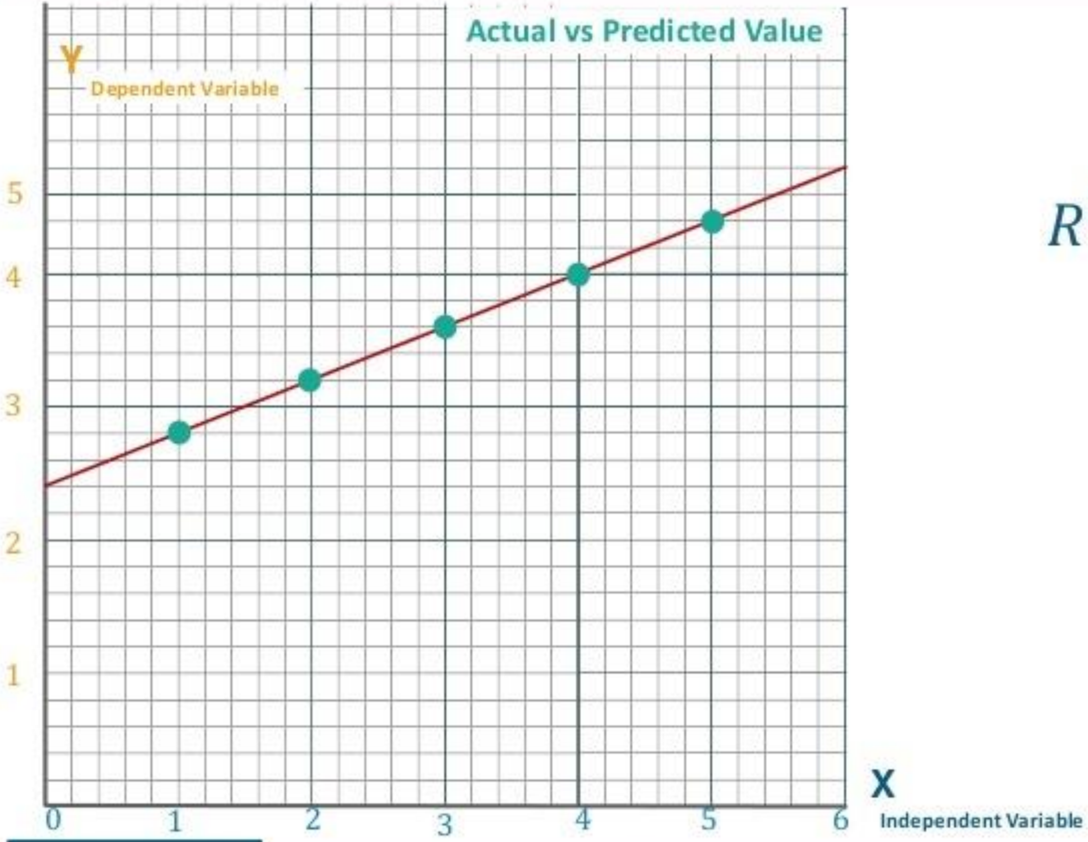
$$R^2 \approx 0.7$$

Calculation of R^2



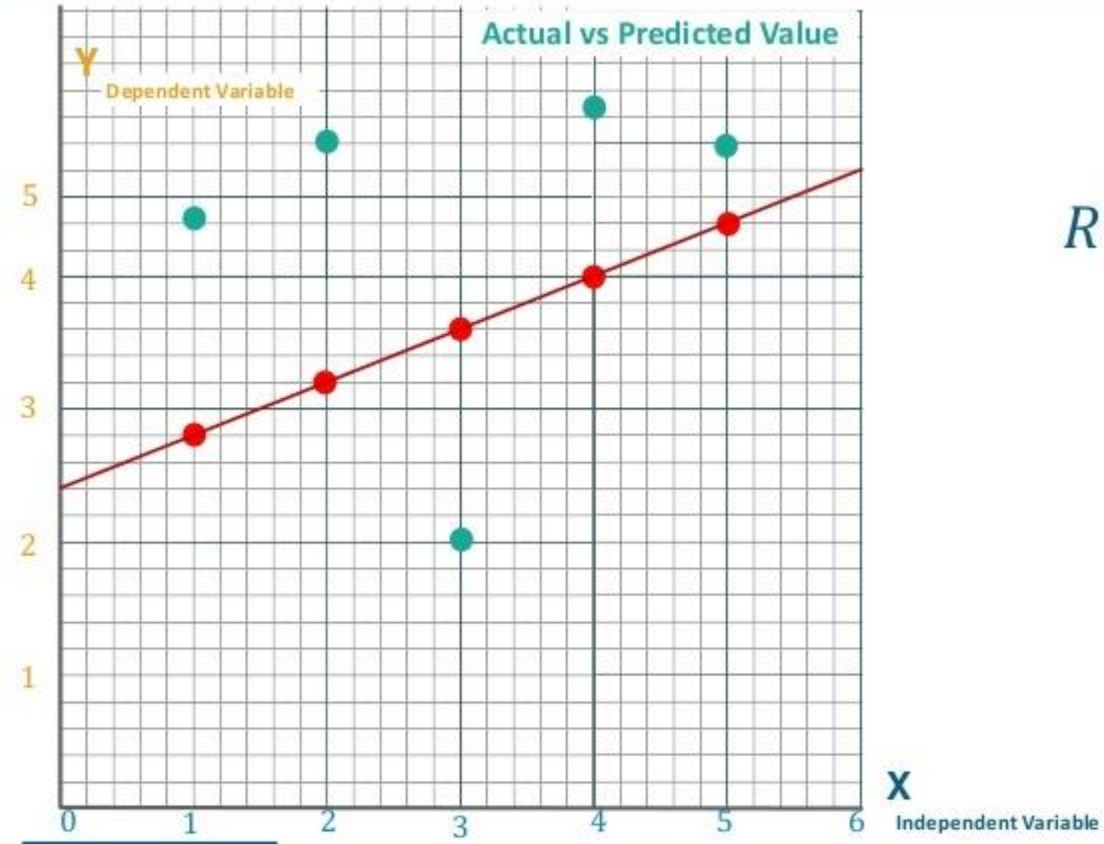
$$R^2 \approx 0.9$$

Calculation of R^2



$$R^2 \approx 1$$

Calculation of R^2



$$R^2 \approx 0.02$$

edureka!

Are Low R-squared values always bad?

edureka!

Are High R-squared values always good?

Let’s learn to code

600,000+

SATISFIED LEARNERS

Thank you! ☺

For more information please visit our website
www.edureka.co