

# Machine Learning Foundations

- by **Nagaraju Budigam**

# Learning Goals

- ❑ Decision Trees
- ❑ Naïve Bayes
- ❑ Gradient Descent
- ❑ Linear Regression
- ❑ Logistic Regression
- ❑ SVM
- ❑ Neural Networks
- ❑ Kernel Methods
- ❑ K-Means Clustering
- ❑ Hierarchical Clustering

# Recommend Apps

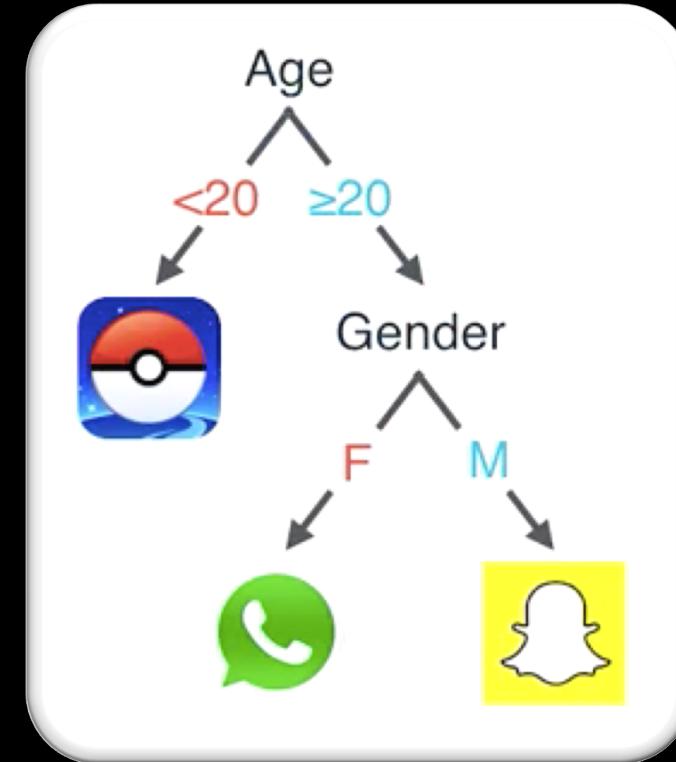
Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	

Quiz: Between Gender and Age, which one seems more decisive for predicting what app will the users download?

- Gender
- Age

# Decision Tree

Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	



# Detecting Spam e-Mails



# Detecting Spam e-Mails

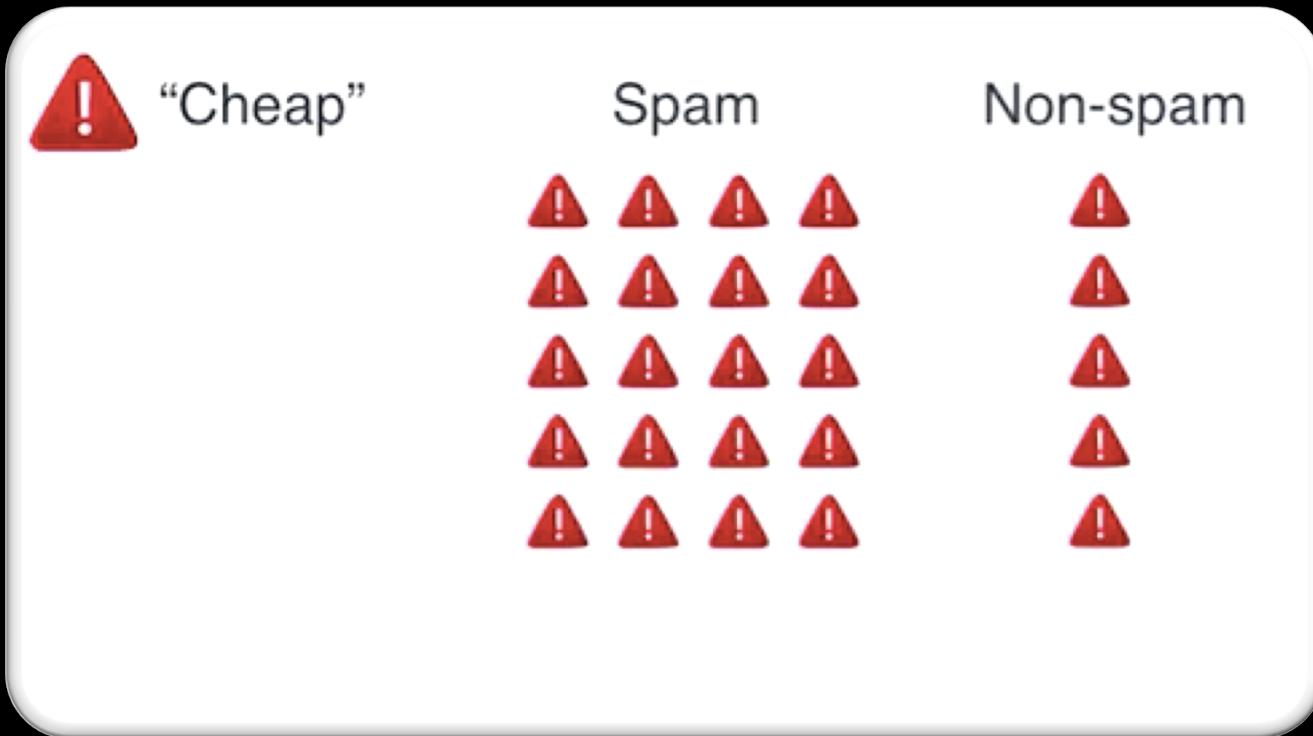
Spam



Non-spam



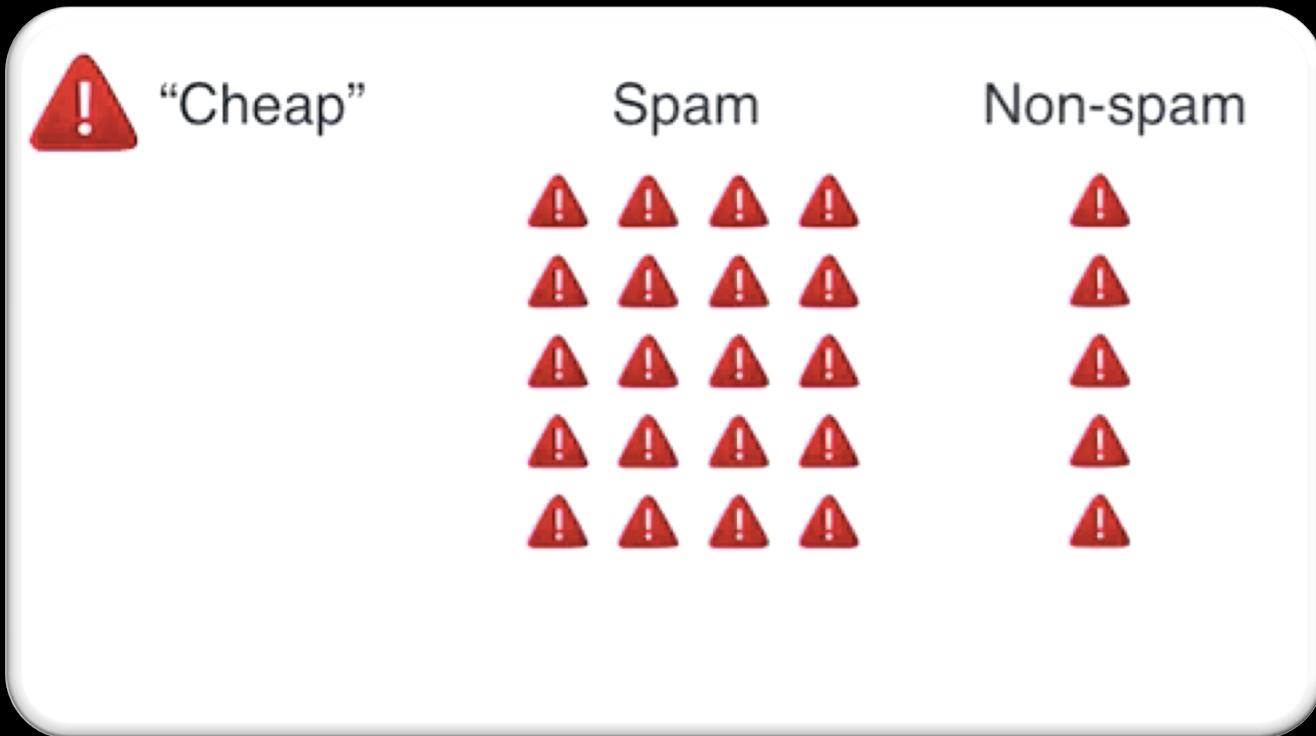
# Detecting Spam e-Mails



Quiz: If an email contains the word cheap what is the probability of it being it spam.

- 40%
- 60%
- 80%

# Detecting Spam e-Mails



Quiz: If an email contains the word cheap what is the probability of it being it spam.

- 40%
- 60%
- 80%

# Naïve Bayes Algorithm



“Cheap” → 80%



Spelling mistake → 70%



Missing title → 95%



etc...

**Quiz:** If an e-mail contains the word “cheap”, what is the probability of it being spam?

40%

60%

80%

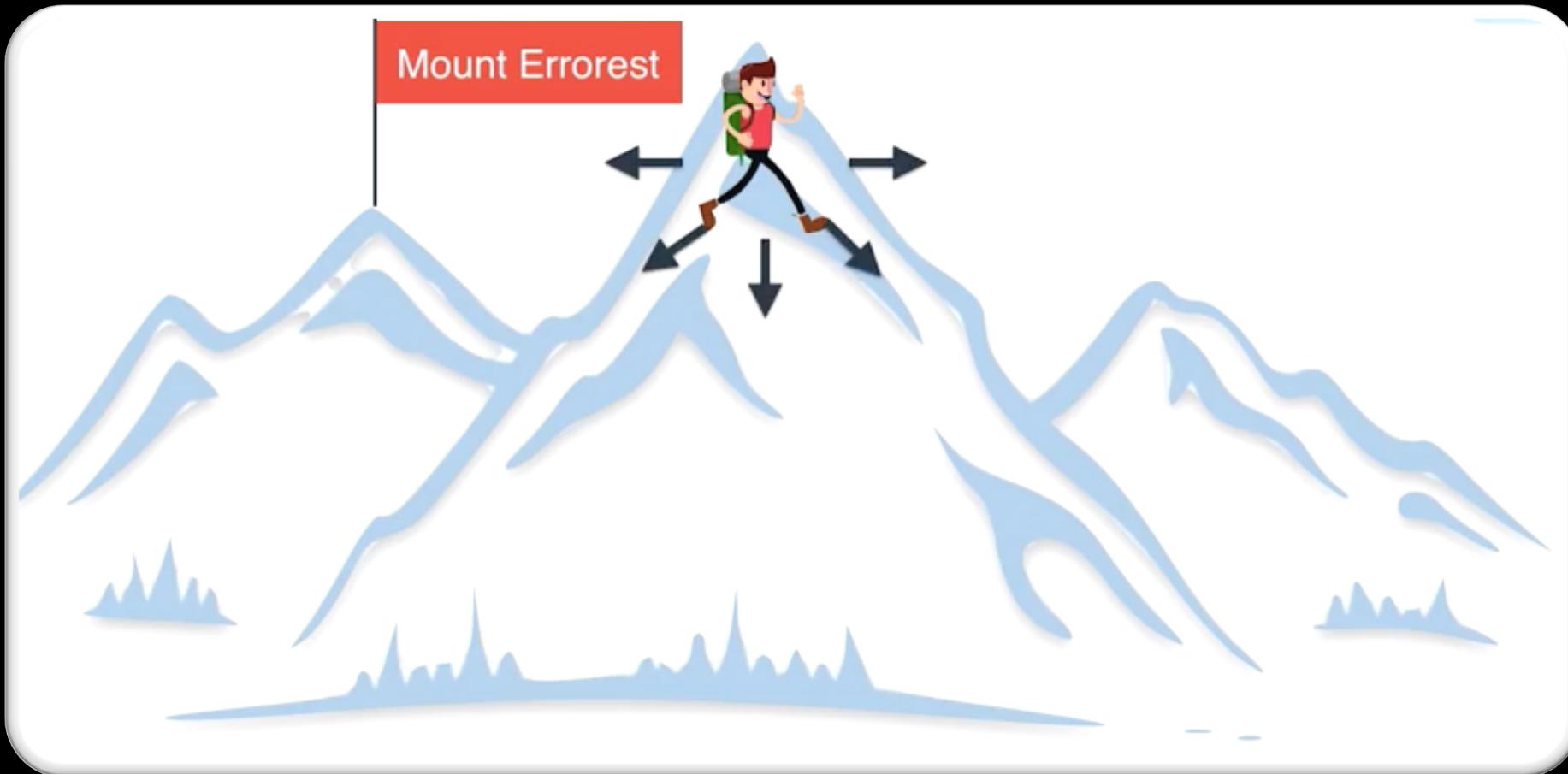
**Conclusion:**

If the e-mail contains the word “cheap”,  
The probability of it being spam is 80%

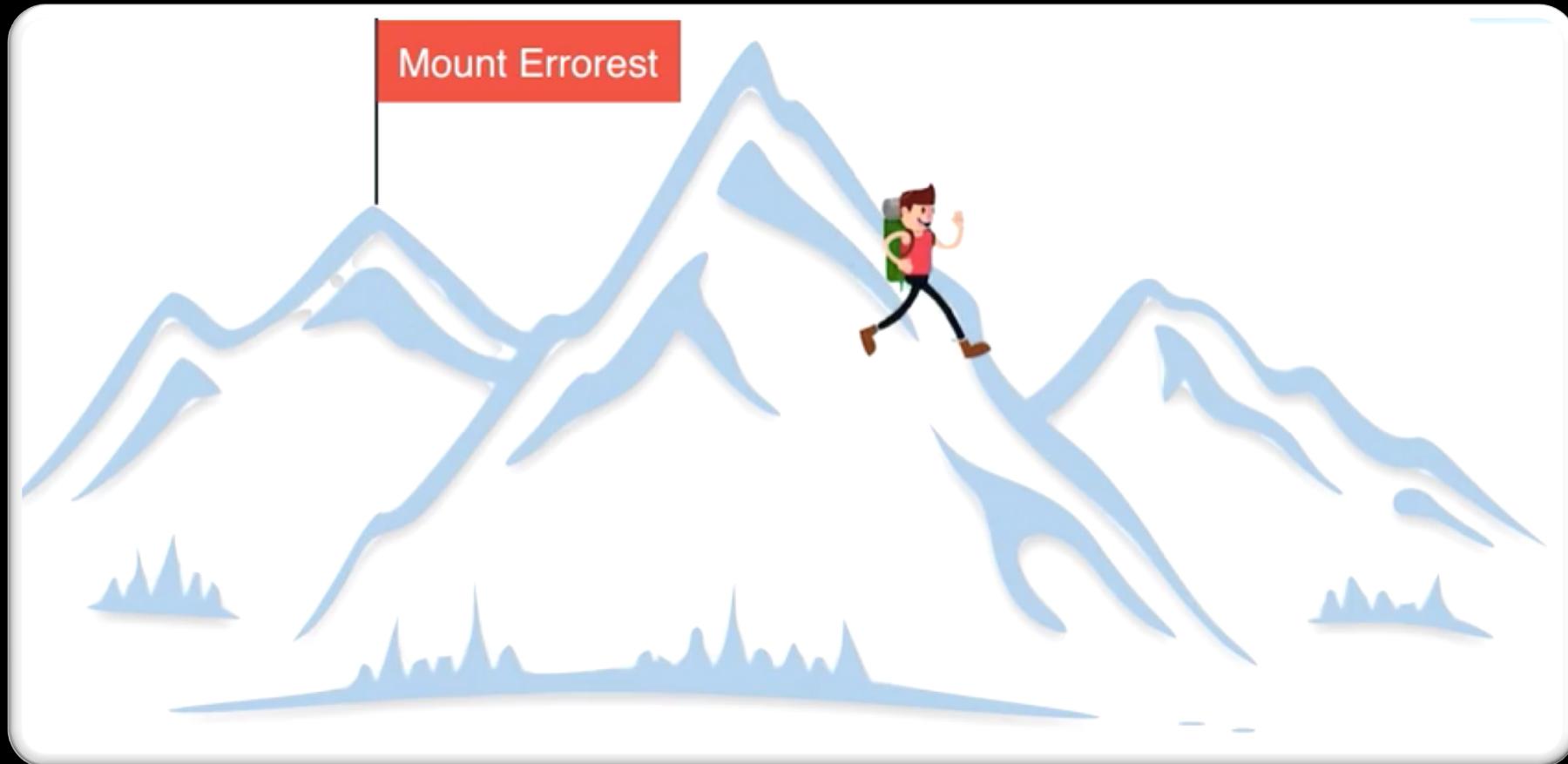
# Gradient Descent



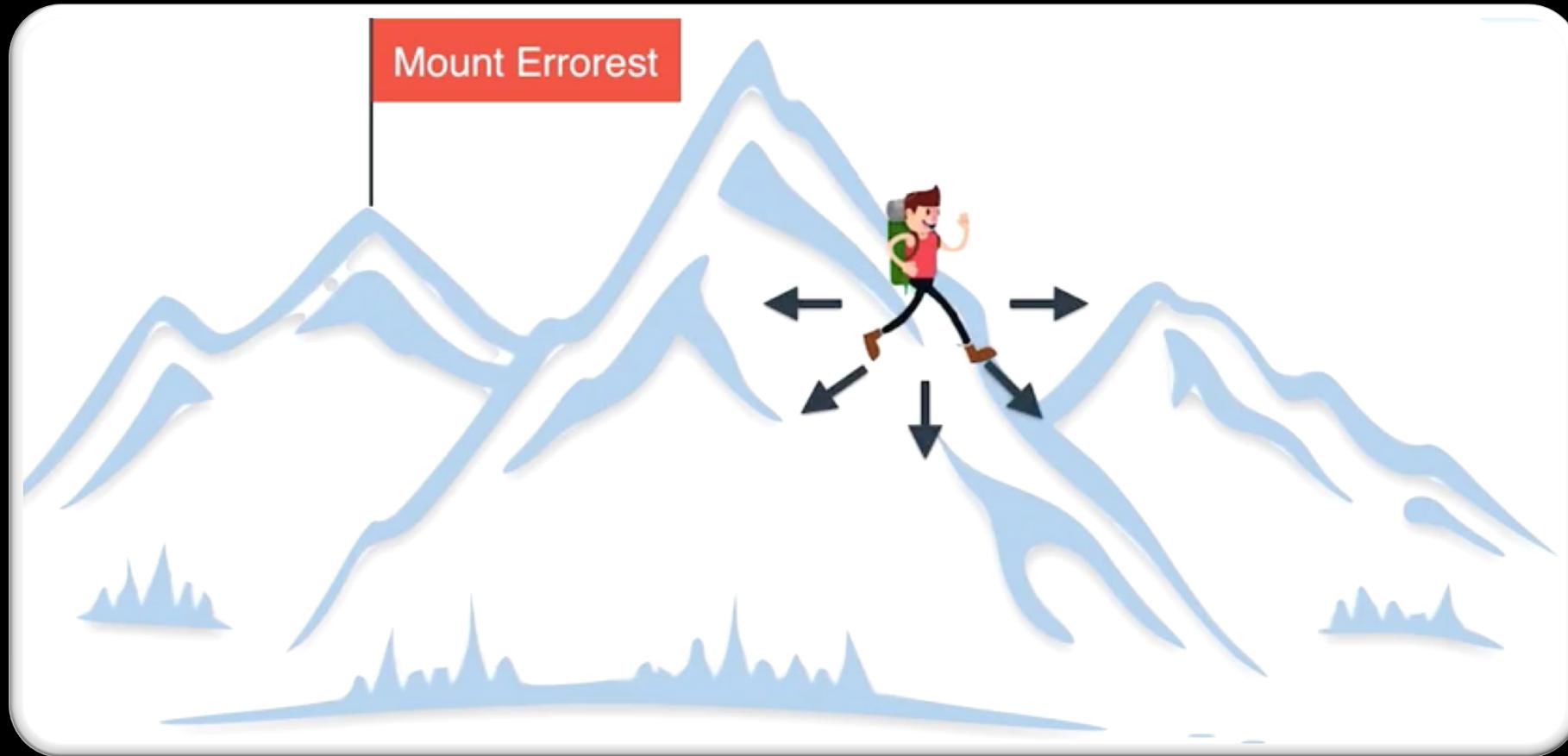
# Gradient Descent



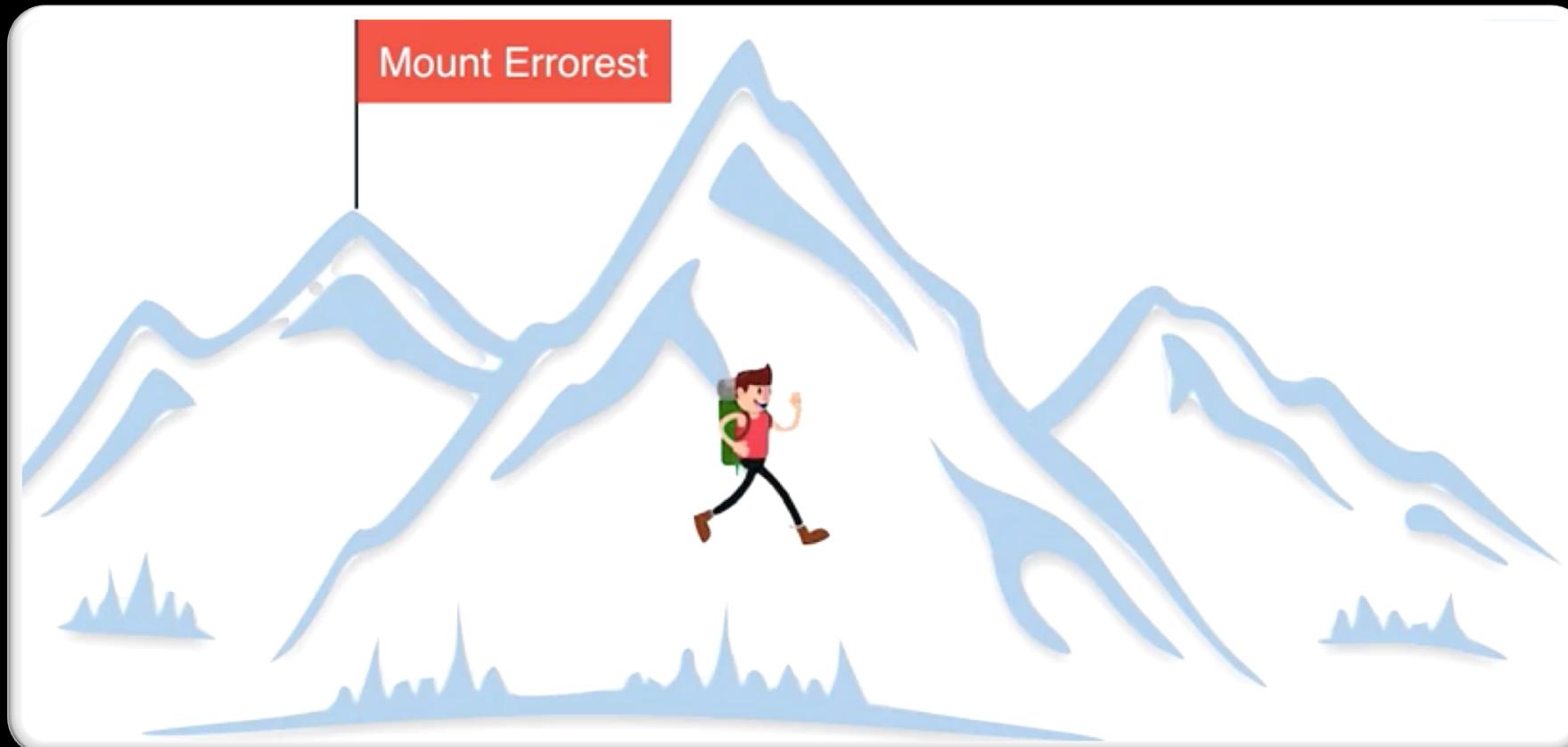
# Gradient Descent



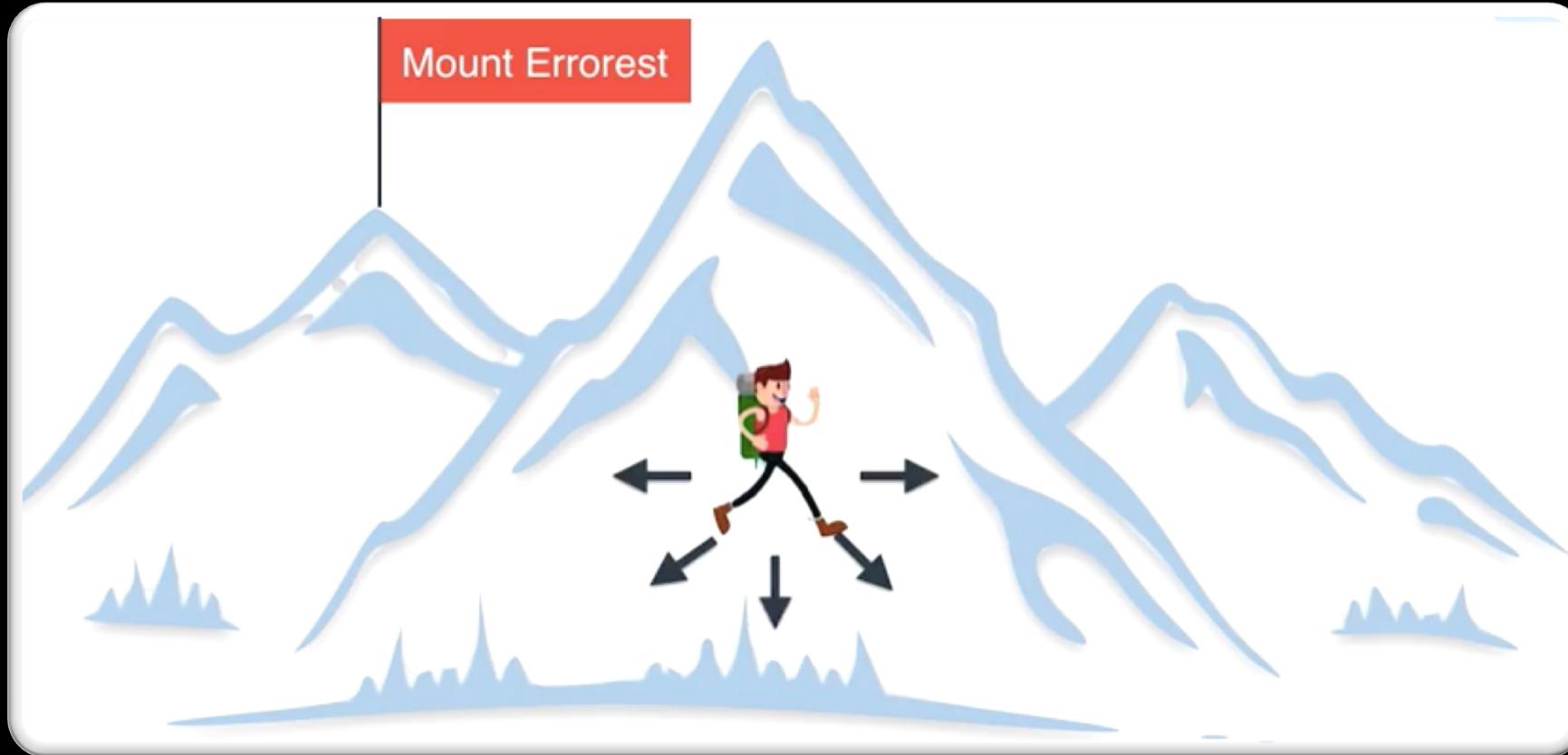
# Gradient Descent



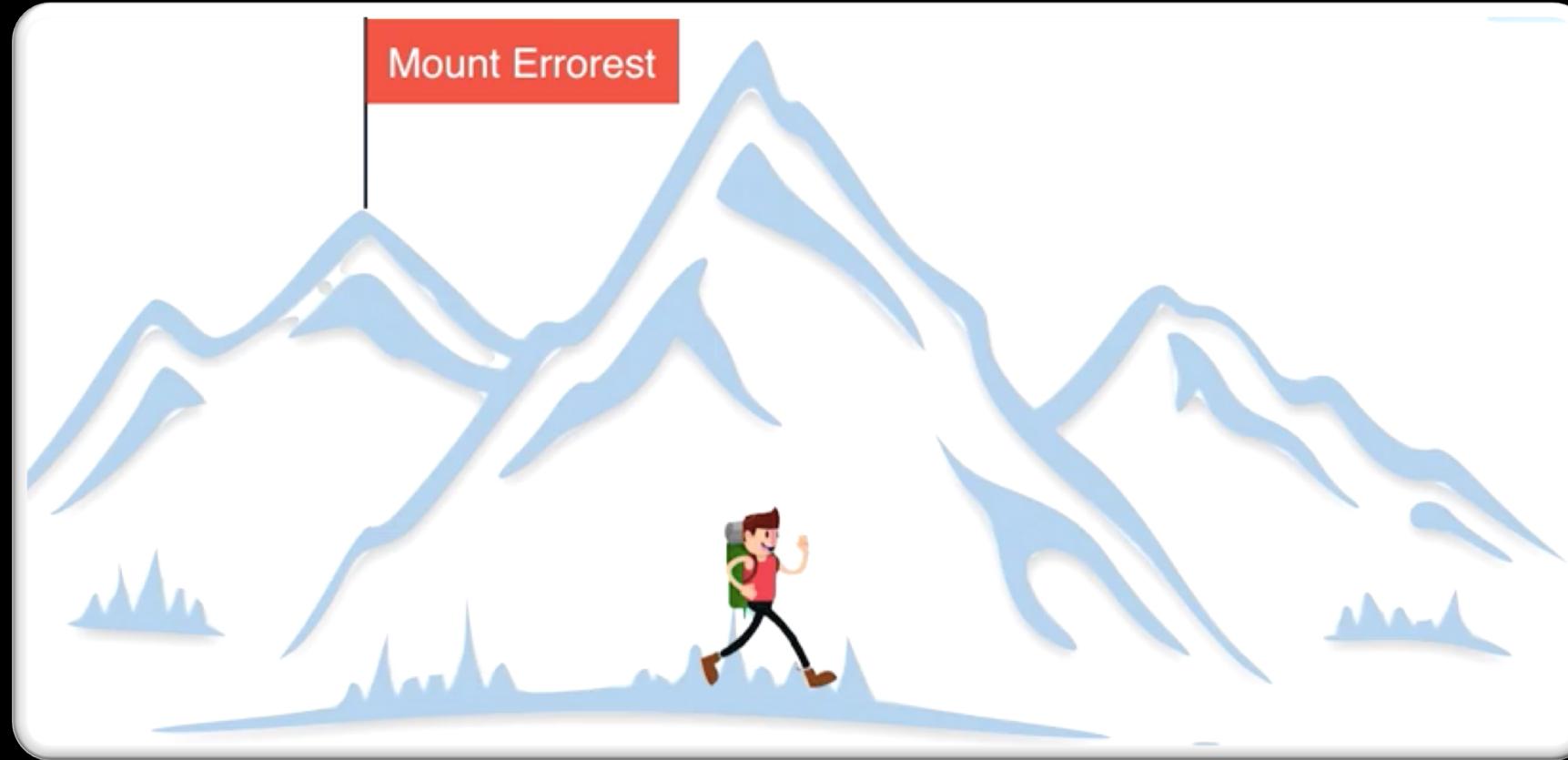
# Gradient Descent



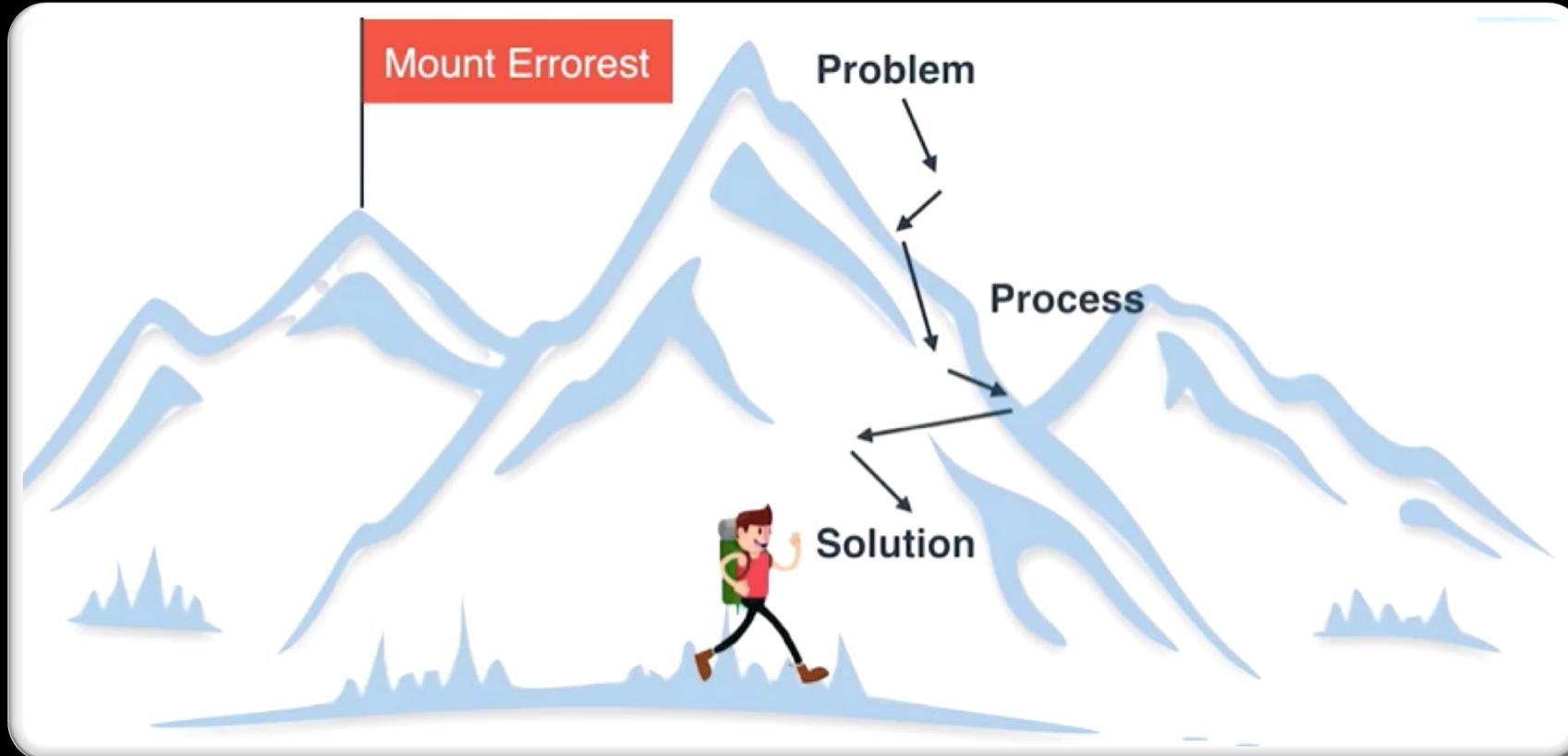
# Gradient Descent



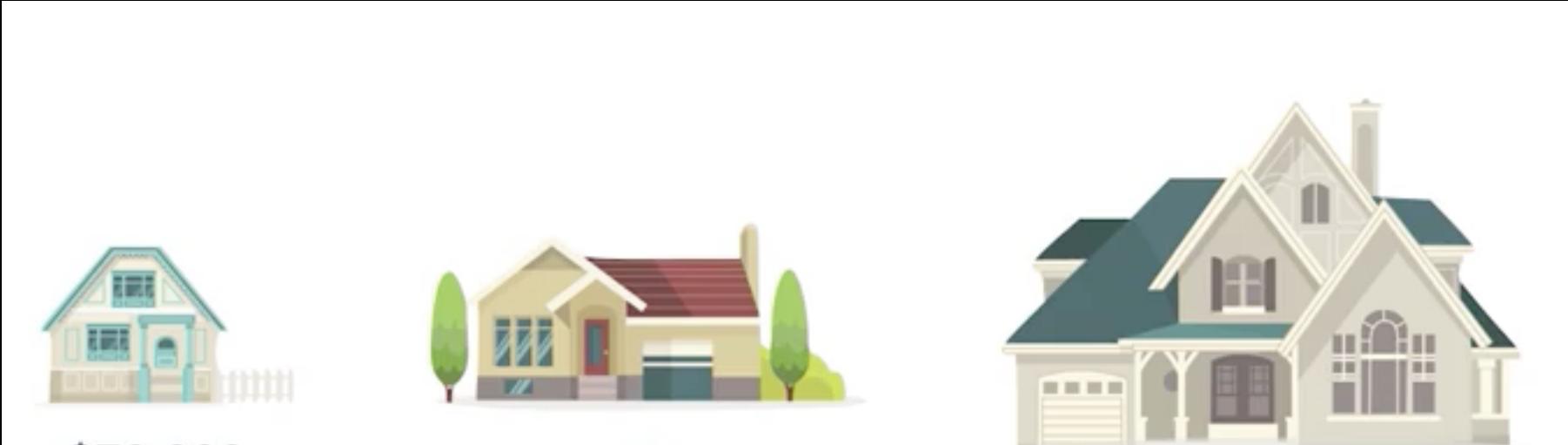
# Gradient Descent



# Gradient Descent



# Predict Price of a House

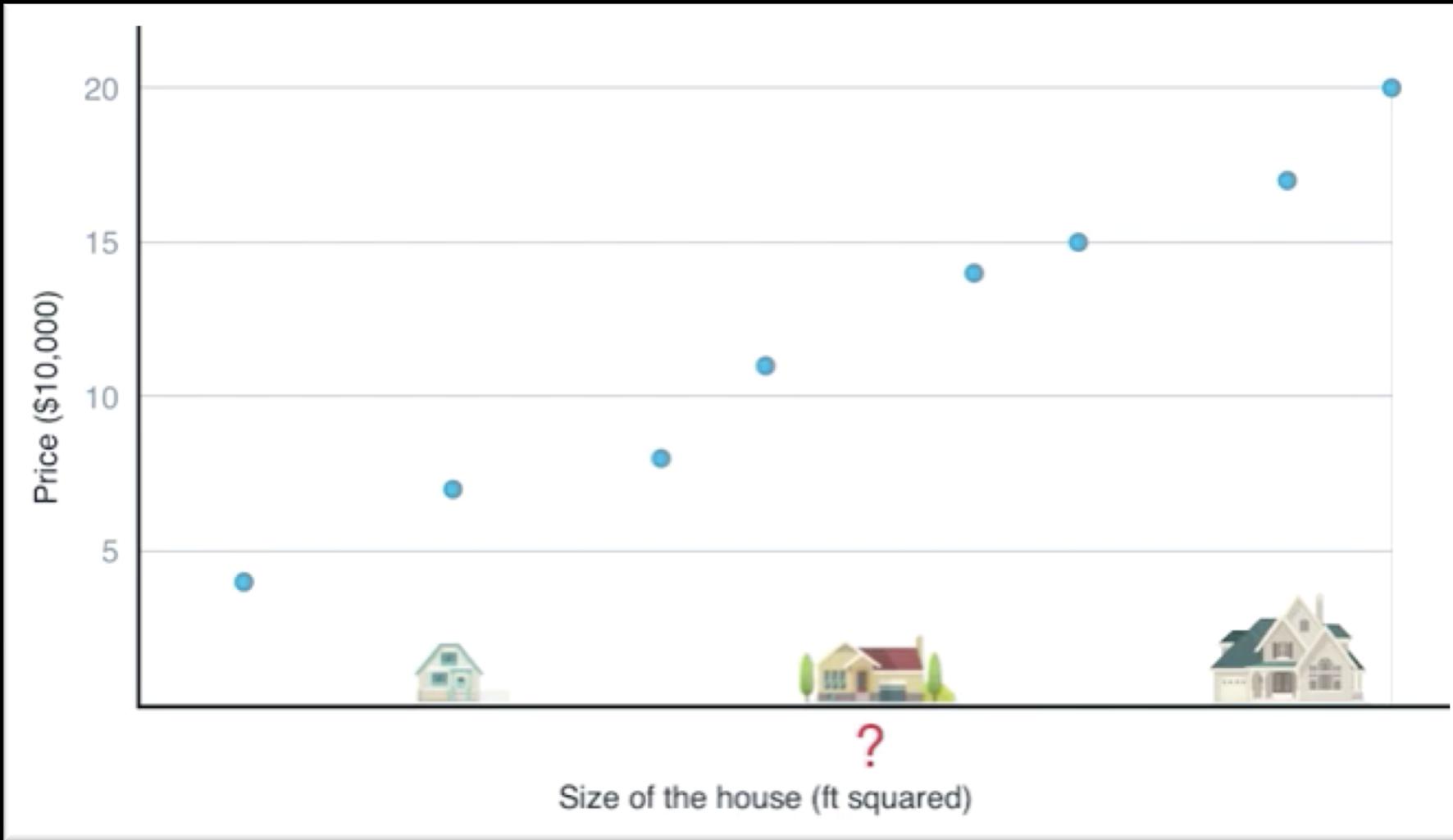


Rs. 7,00000

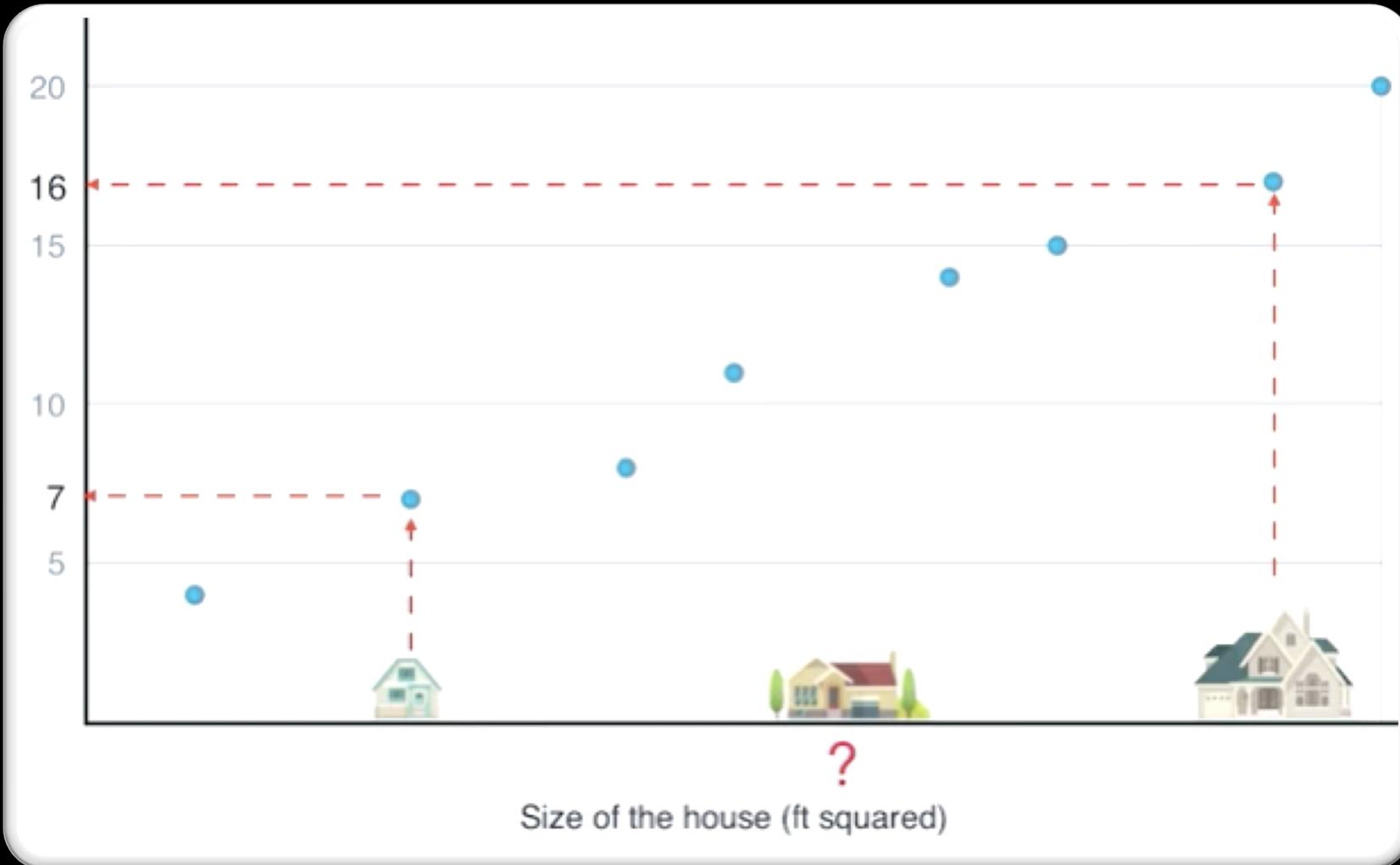
???

Rs. 16,00000

# Predict Price of a House



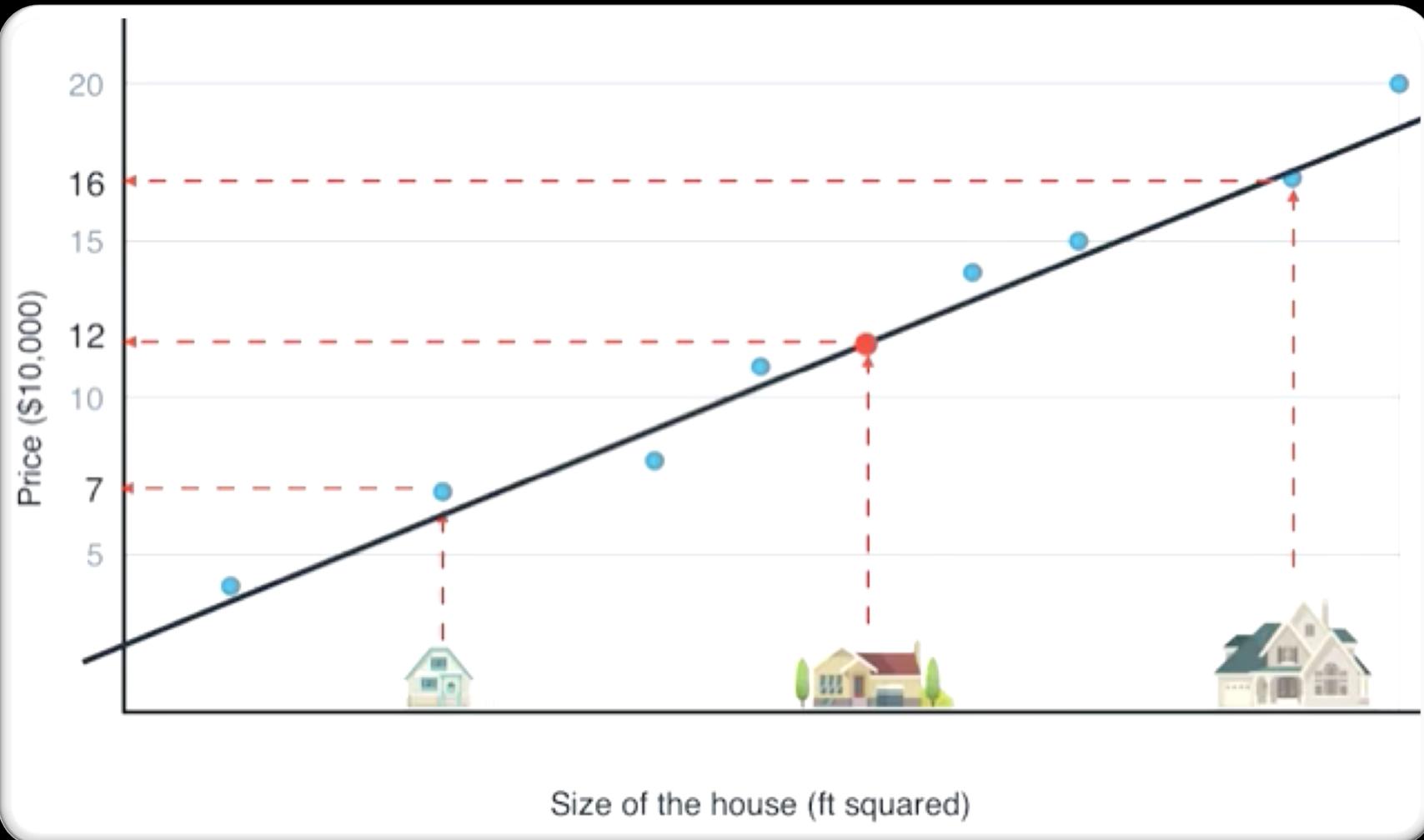
# Predict Price of a House



Quiz: what is the best estimate for the price of the house?

- Rs. 12,00000
- Rs. 20,00000
- Rs. 10,00000

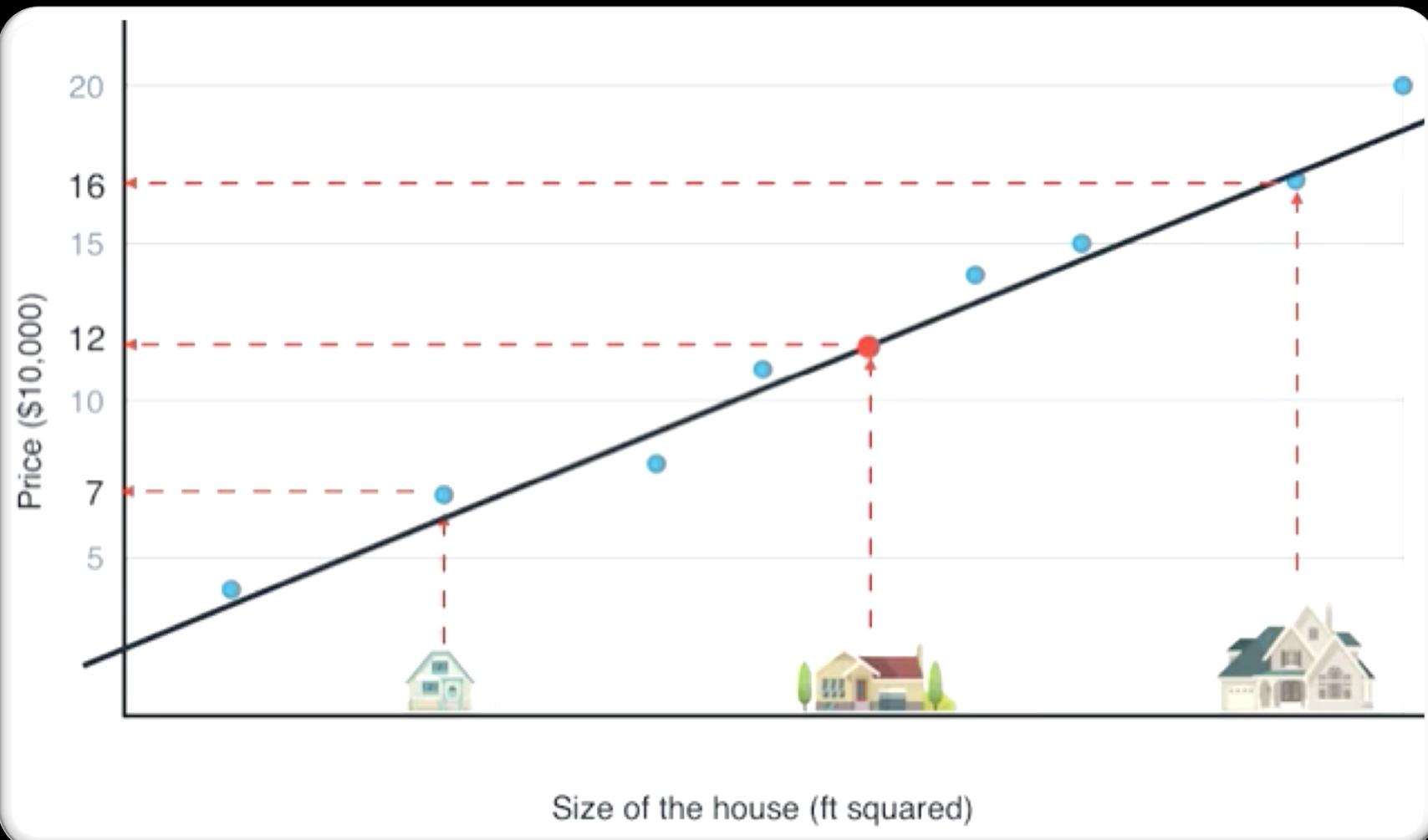
# Predict Price of a House



Quiz: what is the best estimate for the price of the house?

- Rs. 12,00000
- Rs. 20,00000
- Rs. 10,00000

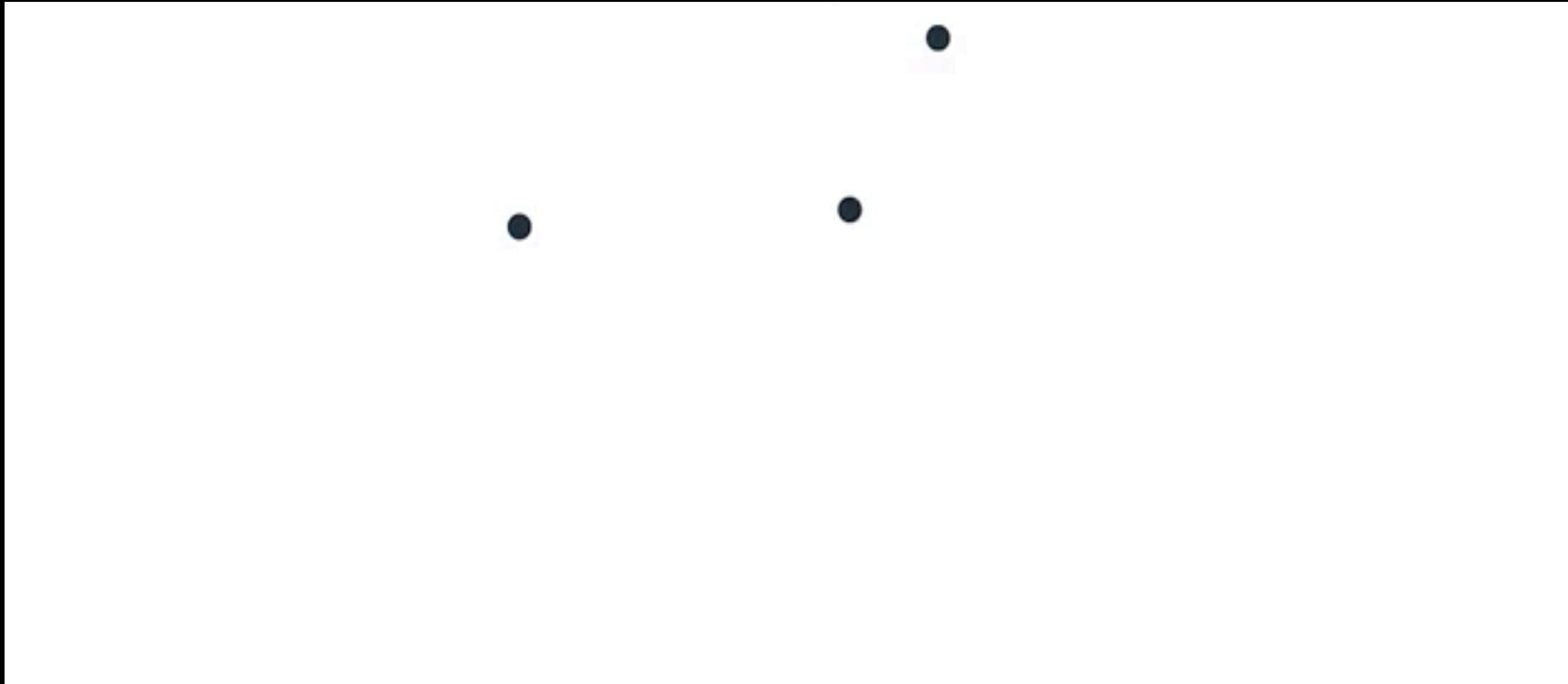
# Linear Regression



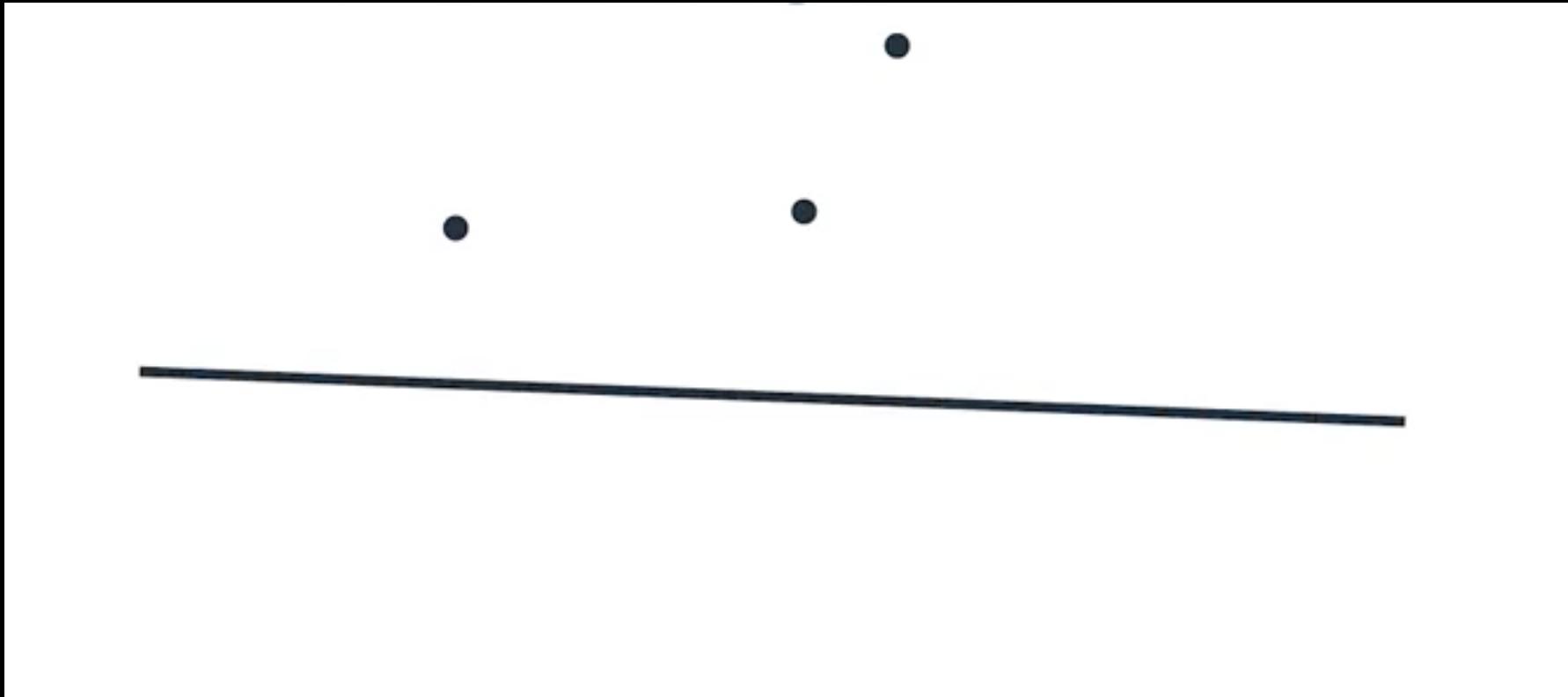
Quiz: what is the best estimate for the price of the house?

- Rs. 12,00000
- Rs. 20,00000
- Rs. 10,00000

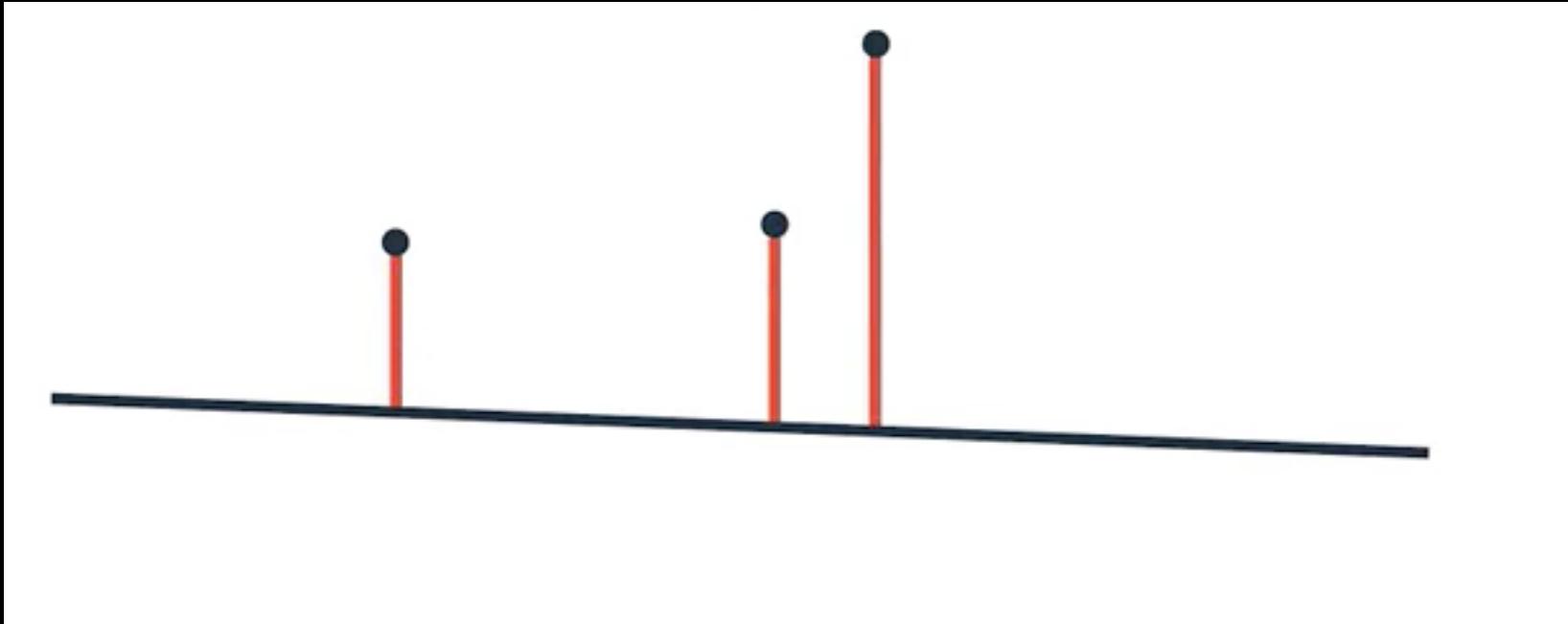
# Linear Regression



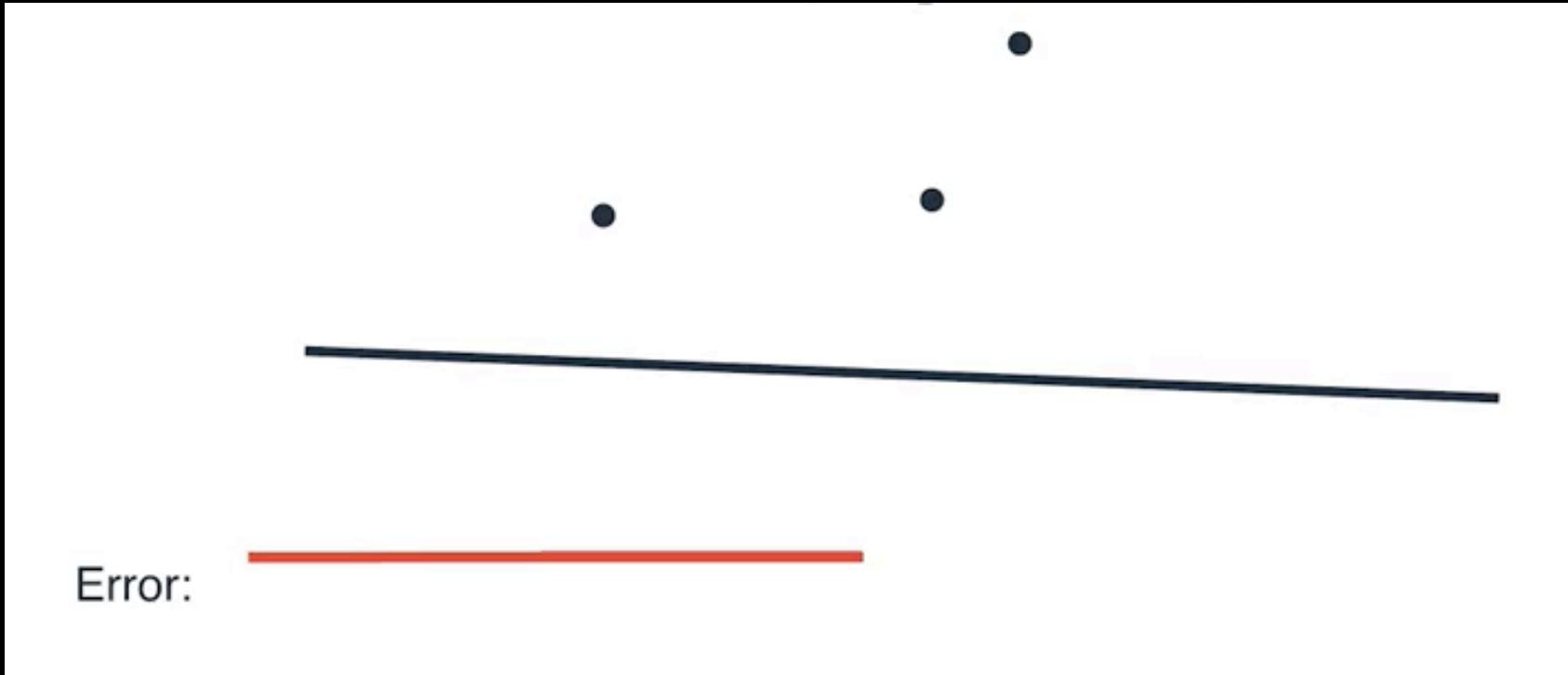
# Linear Regression



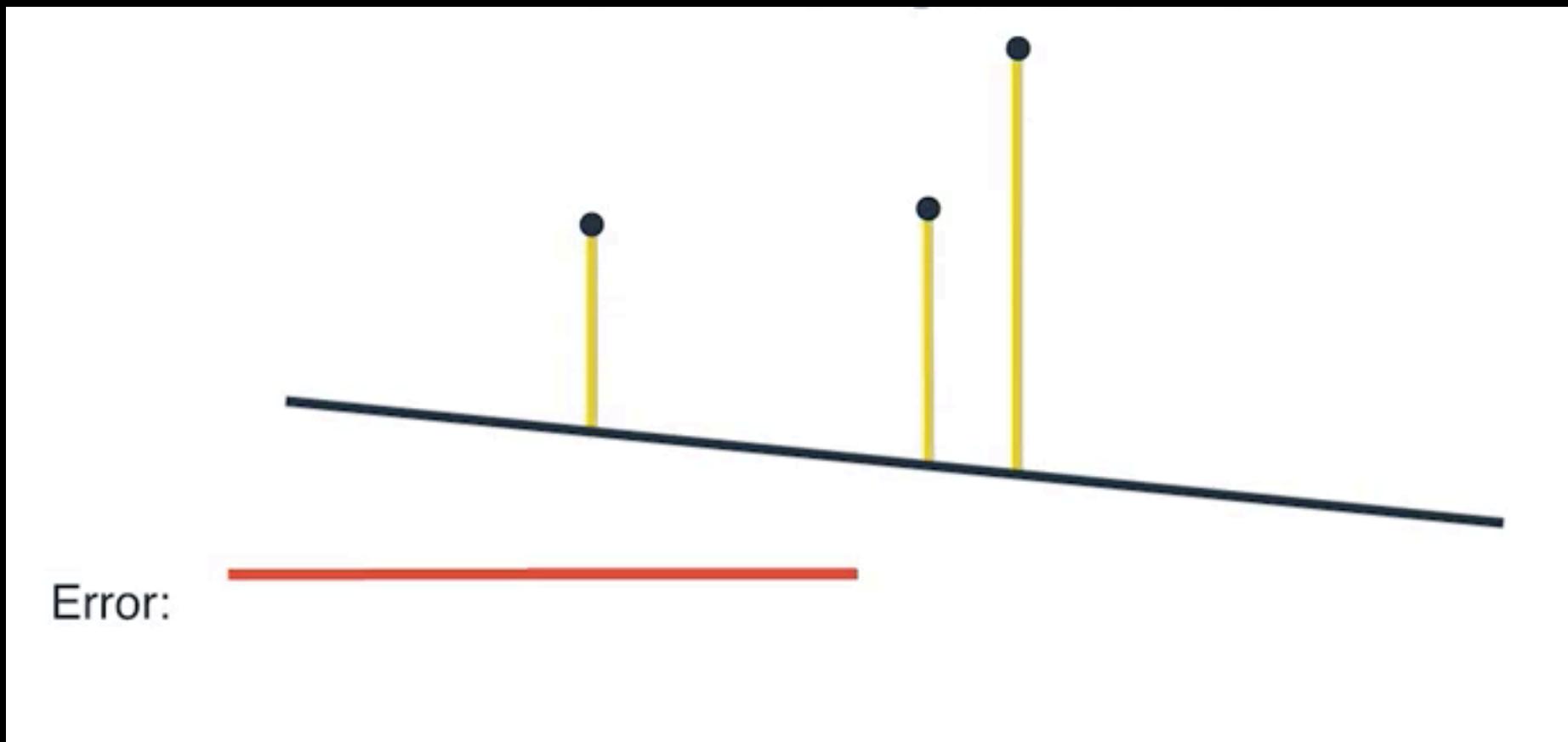
# Linear Regression



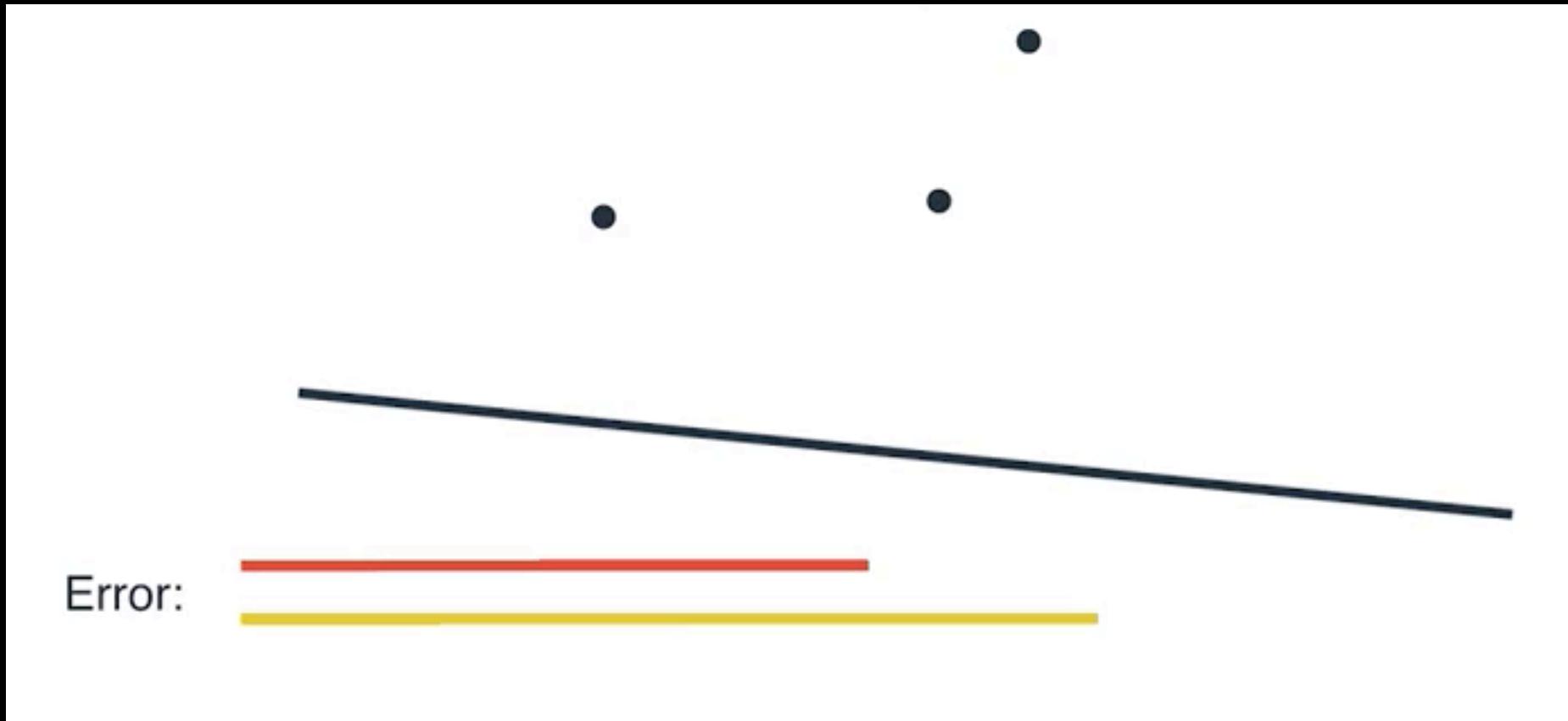
# Linear Regression



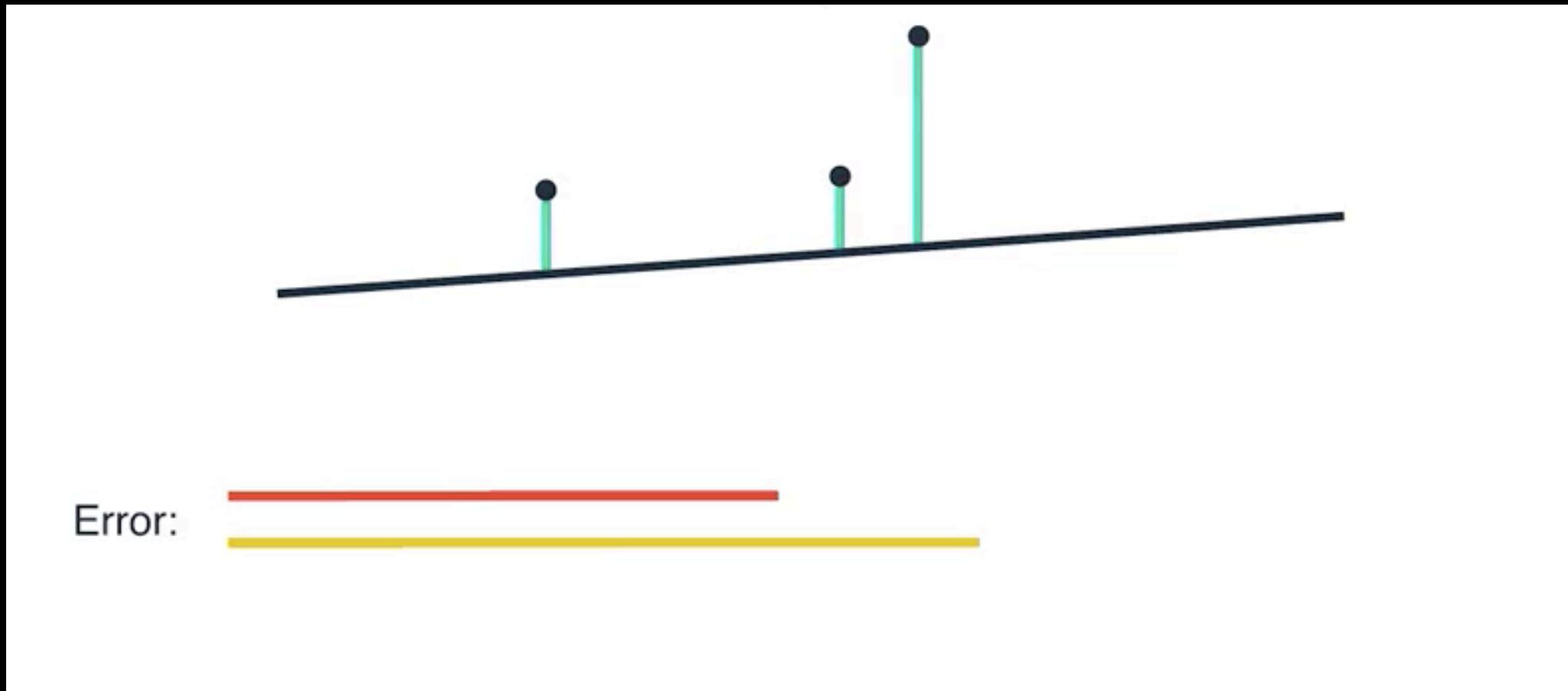
# Linear Regression



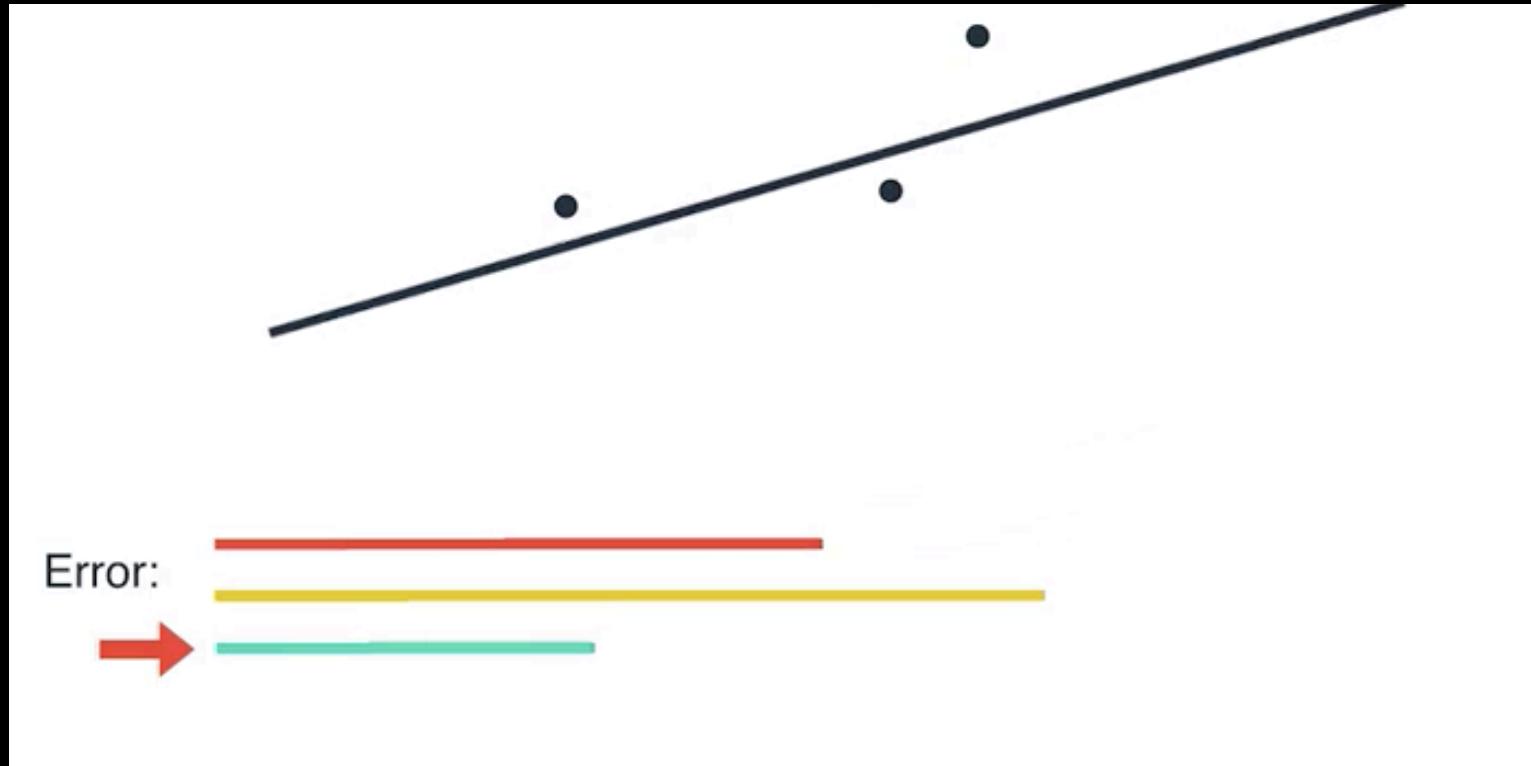
# Linear Regression



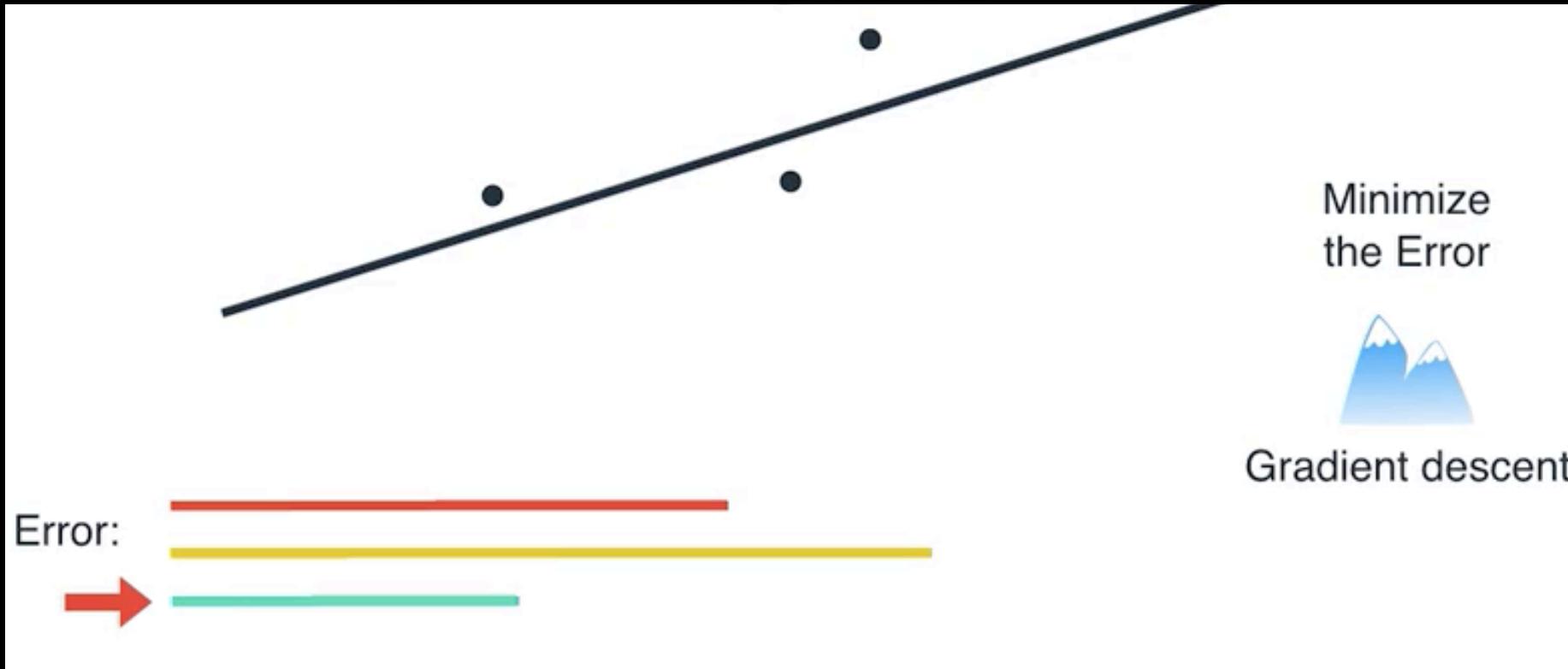
# Linear Regression



# Linear Regression



# Linear Regression



# Acceptance at a University



Test



Grades

# Acceptance at a University



Test



Grades

Student 1  
Test: 9/10   
Grades: 8/10

Student 2  
Test: 3/10   
Grades: 4/10

Student 3  
Test: 7/10   
Grades: 6/10

# Acceptance at a University



# Acceptance at a University



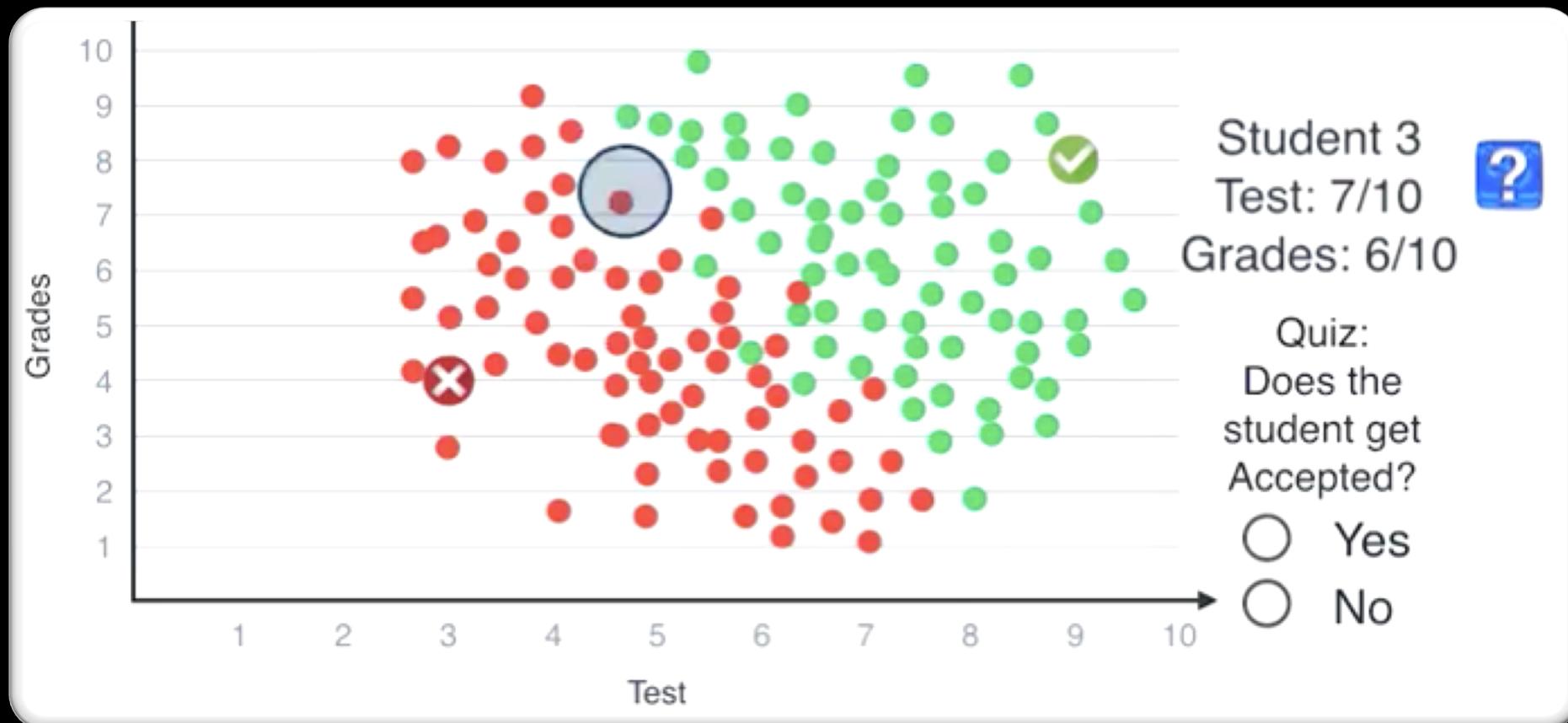
# Acceptance at a University



# Acceptance at a University



# Acceptance at a University



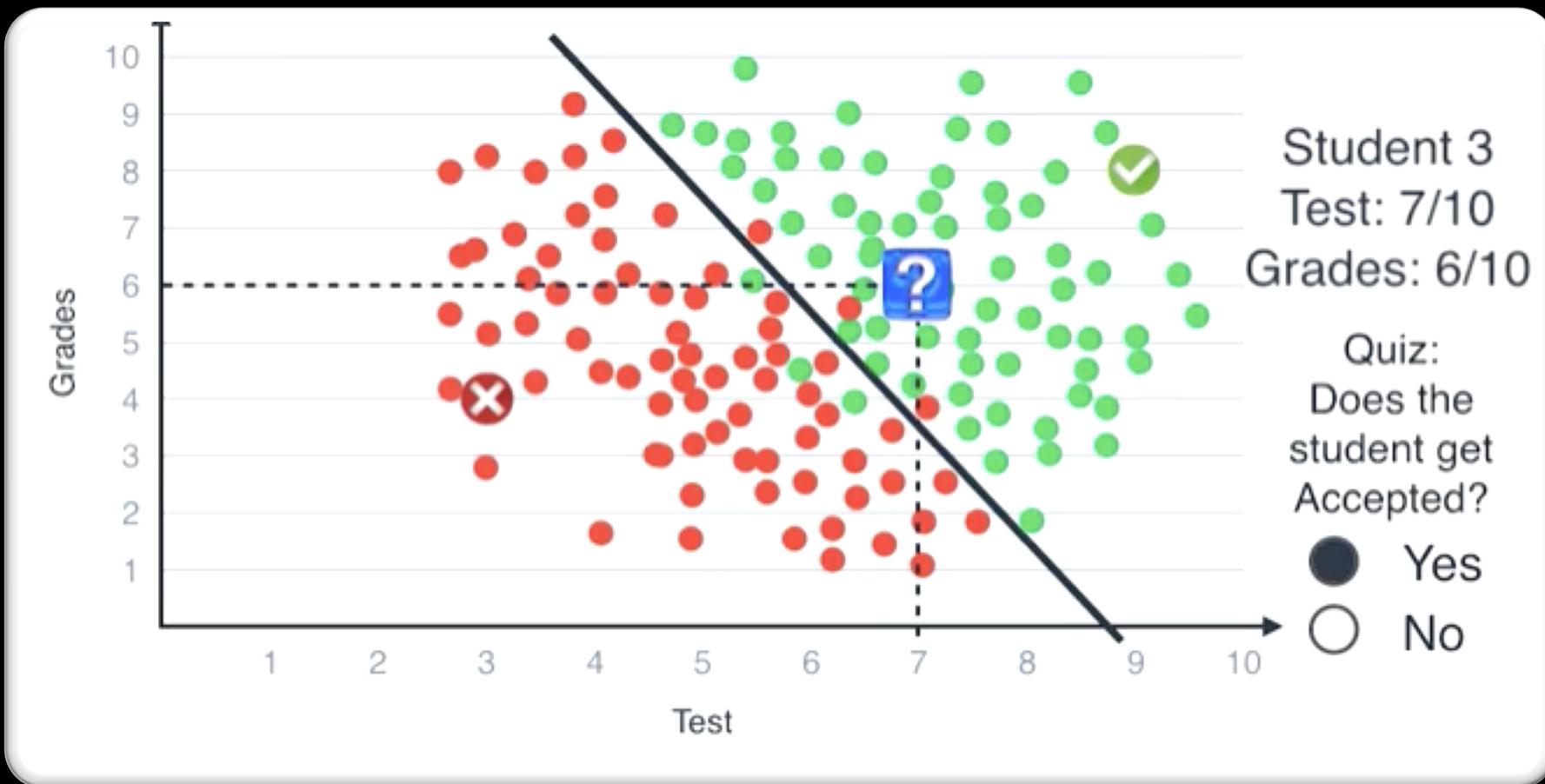
# Acceptance at a University



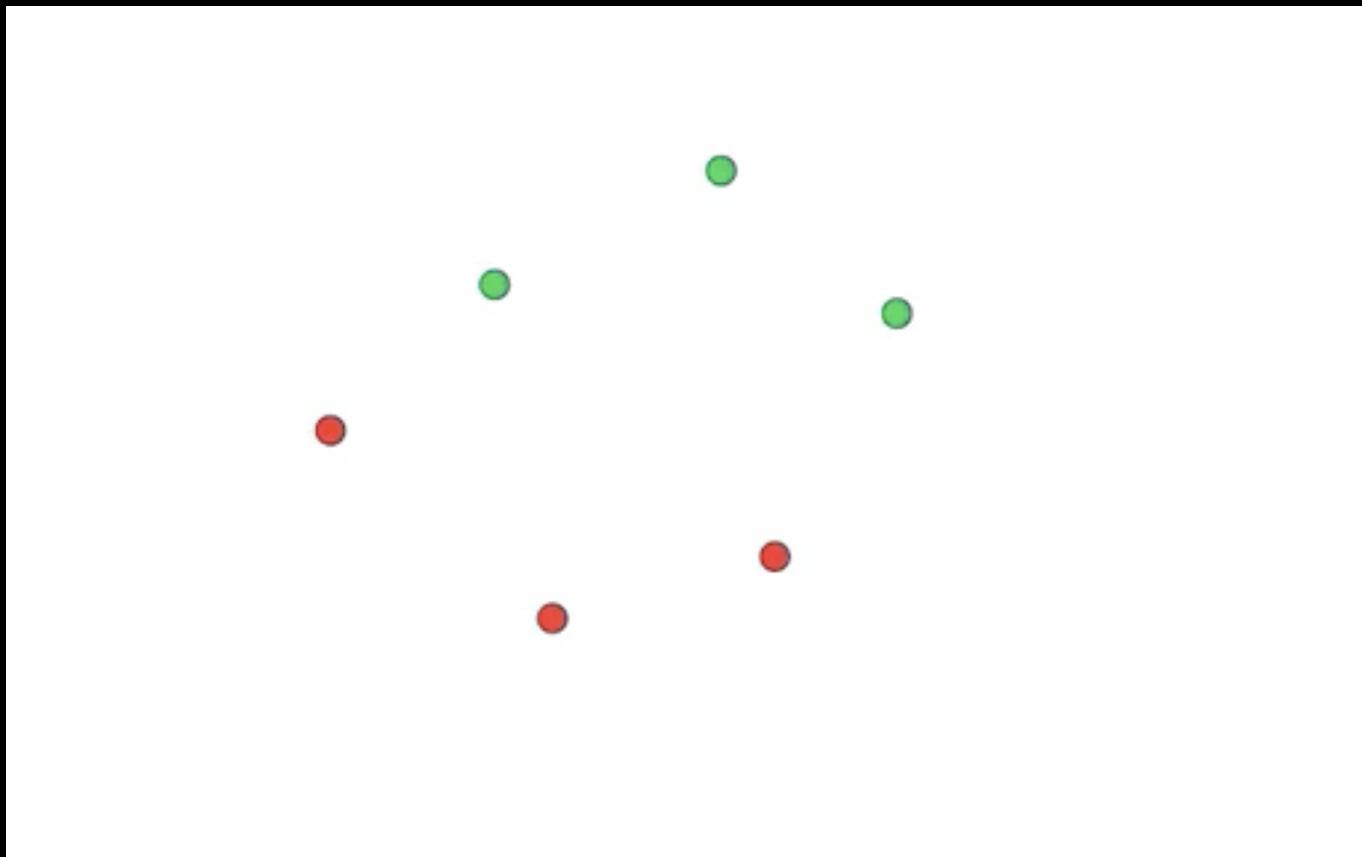
# Acceptance at a University



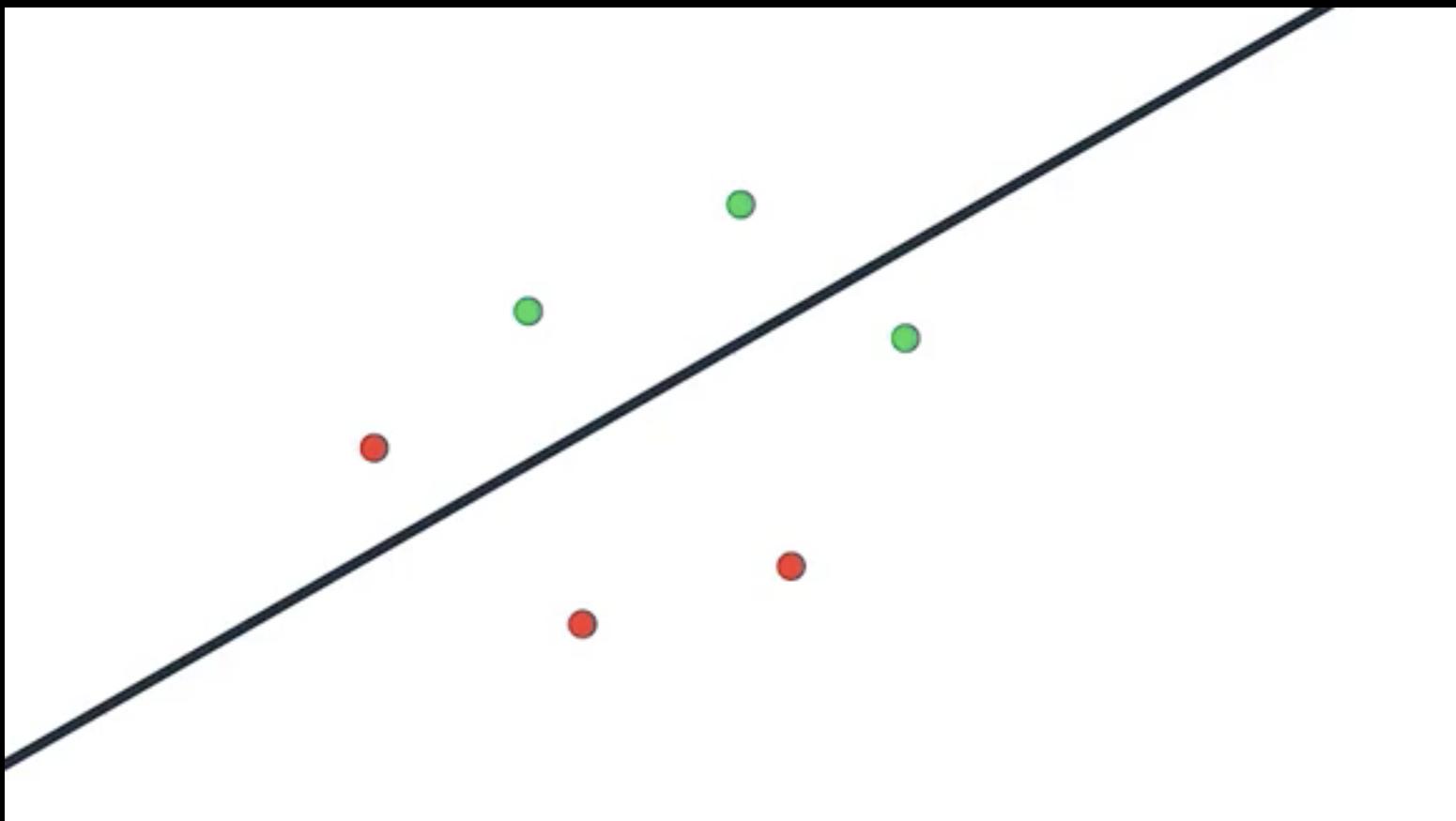
# Acceptance at a University



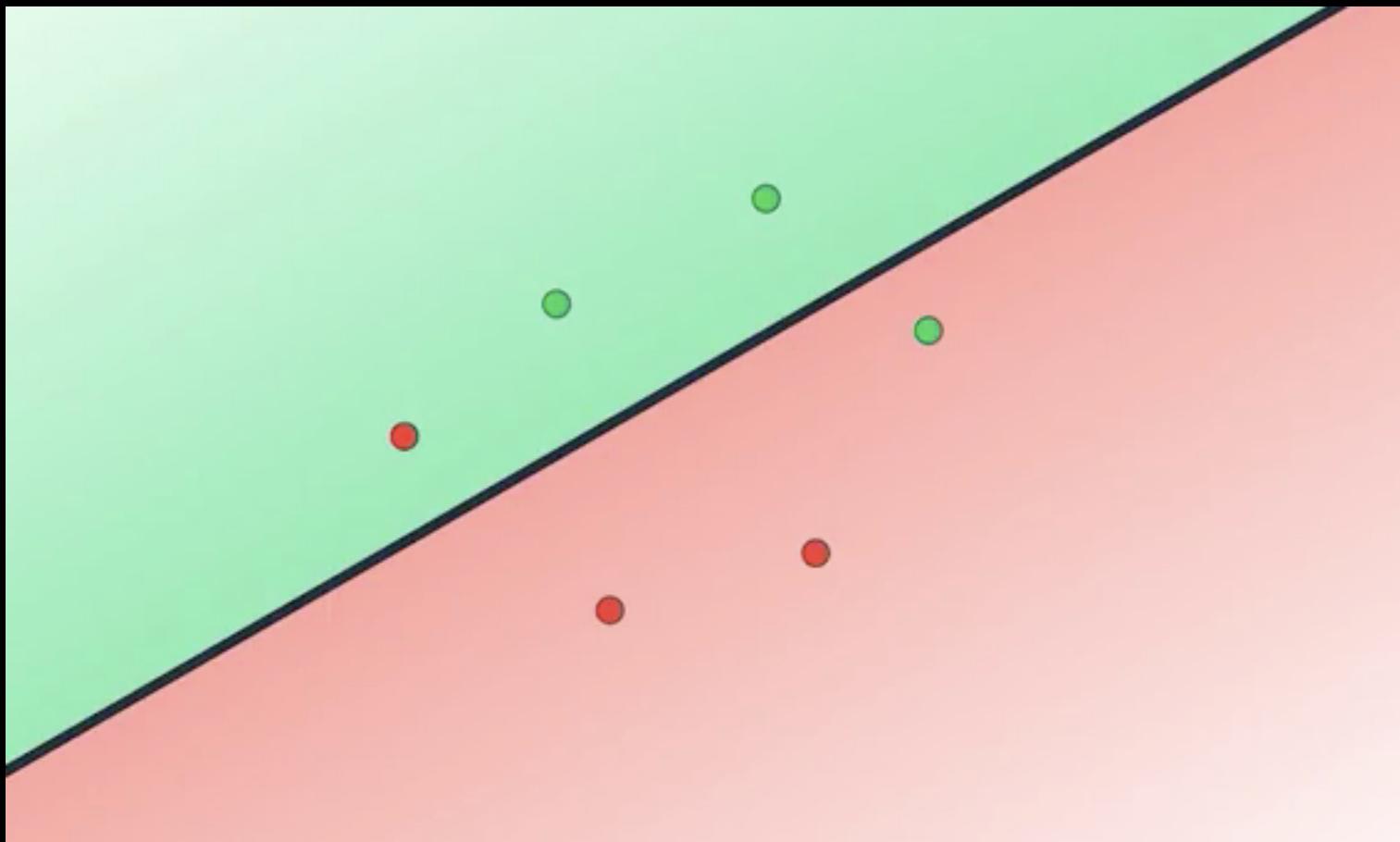
# Logistic Regression



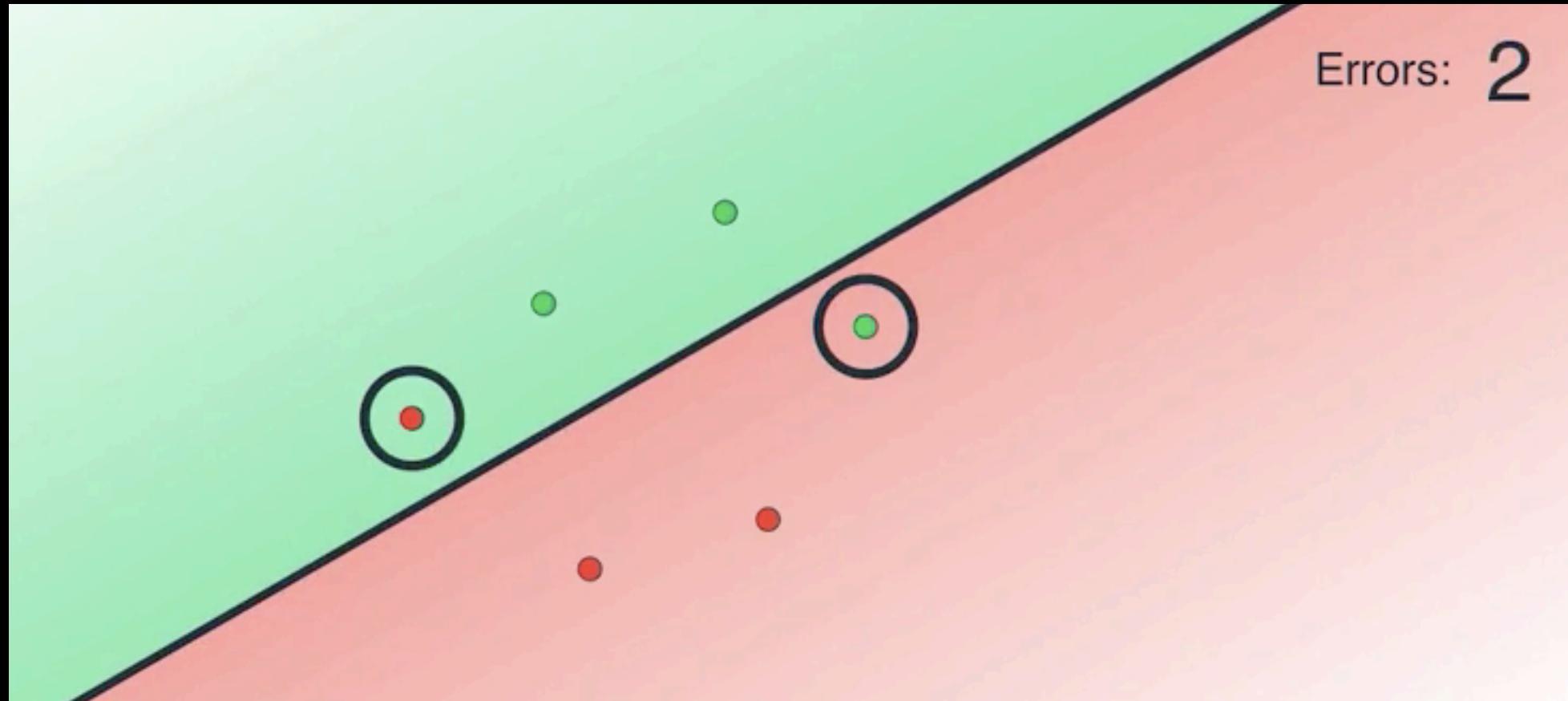
# Logistic Regression



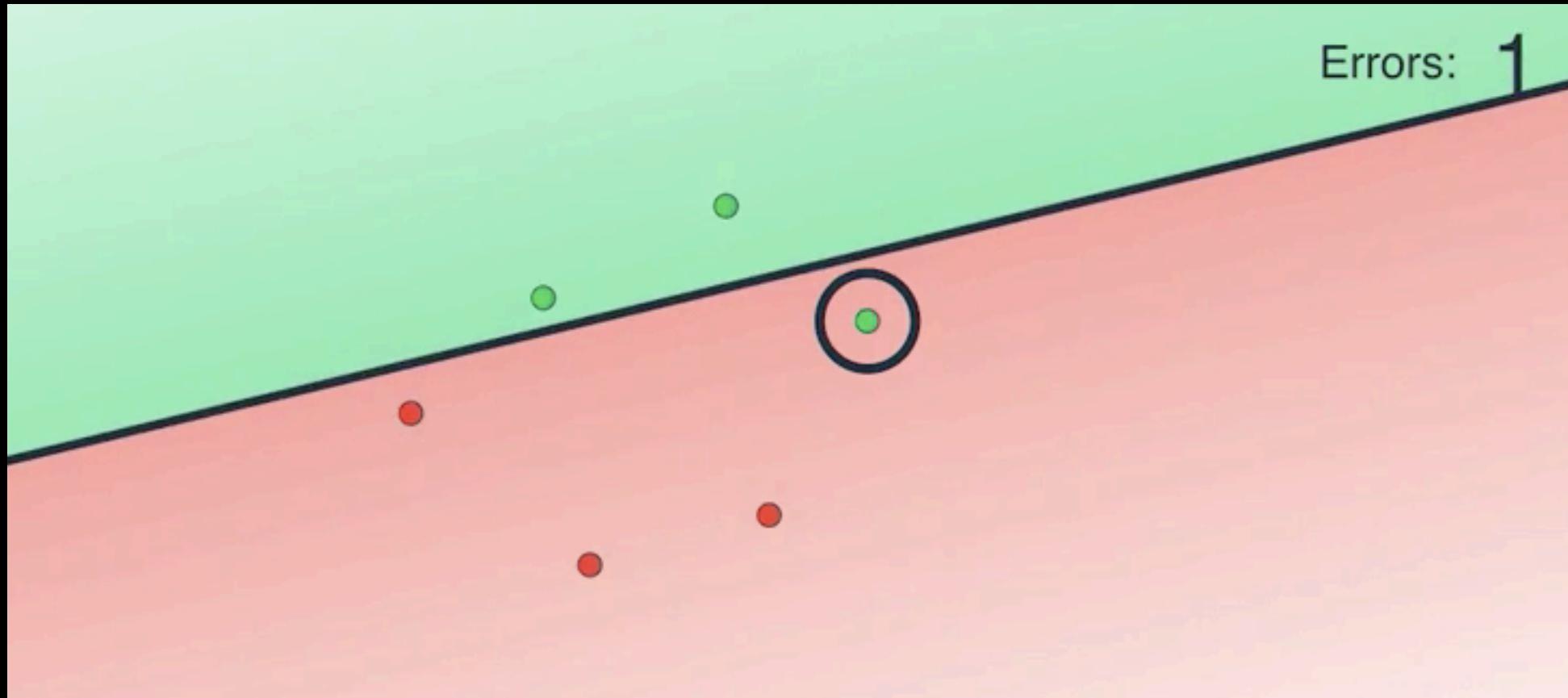
# Logistic Regression



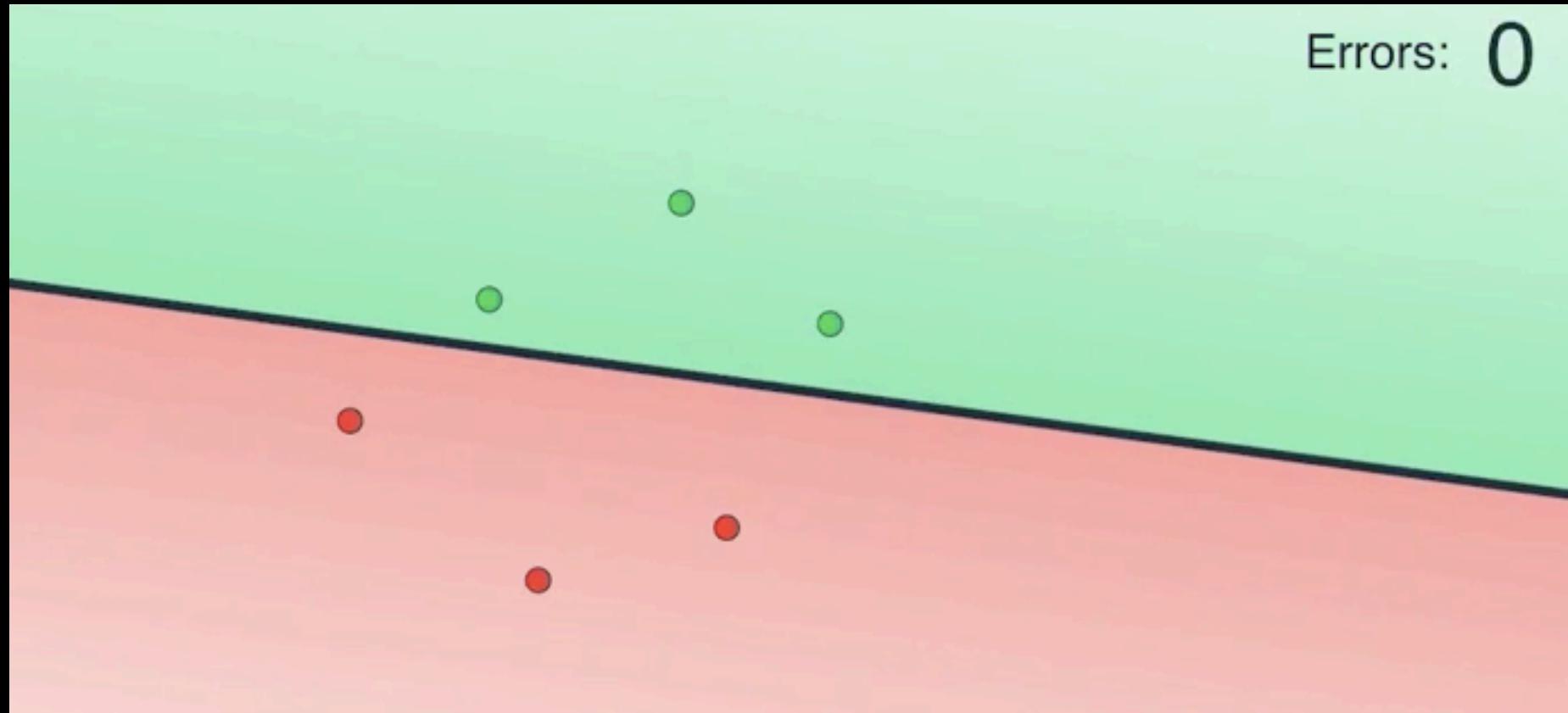
# Logistic Regression



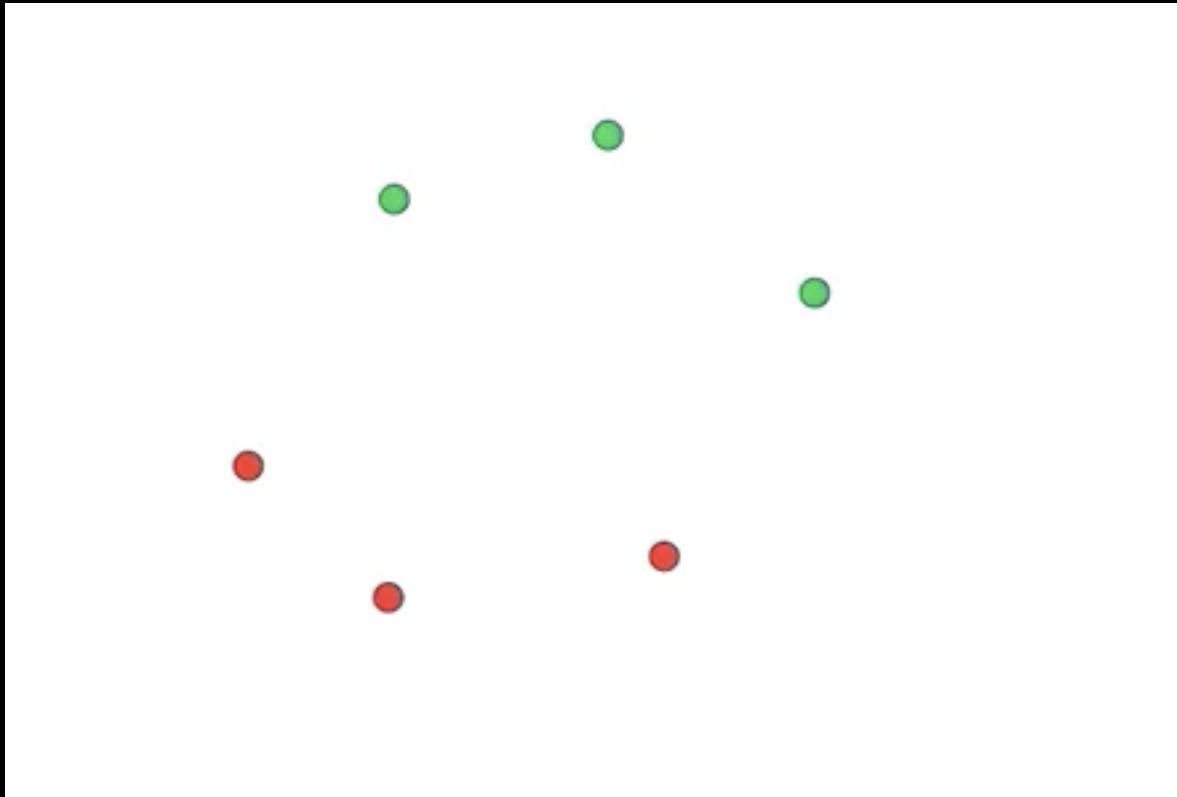
# Logistic Regression



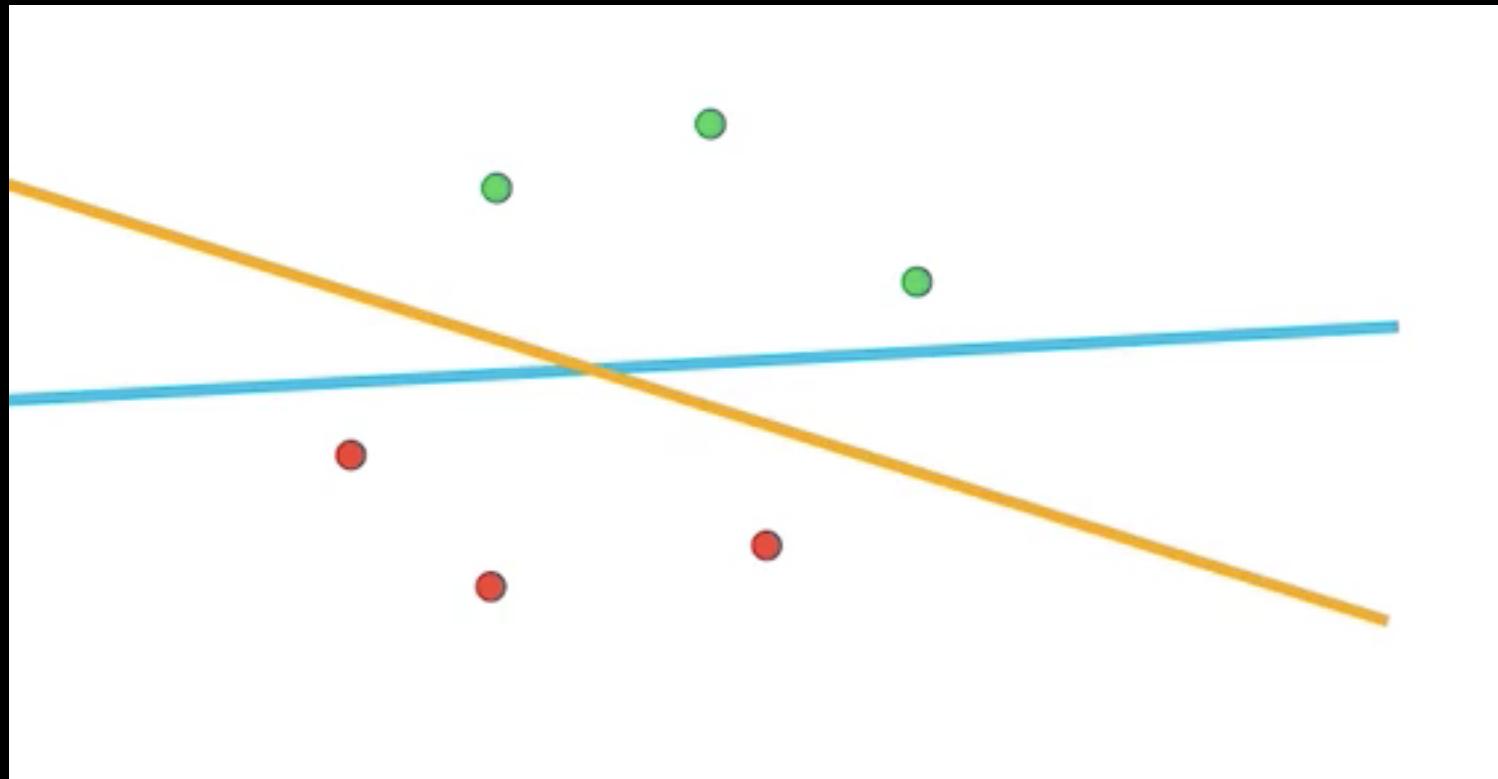
# Logistic Regression



# Cutting Data with Style

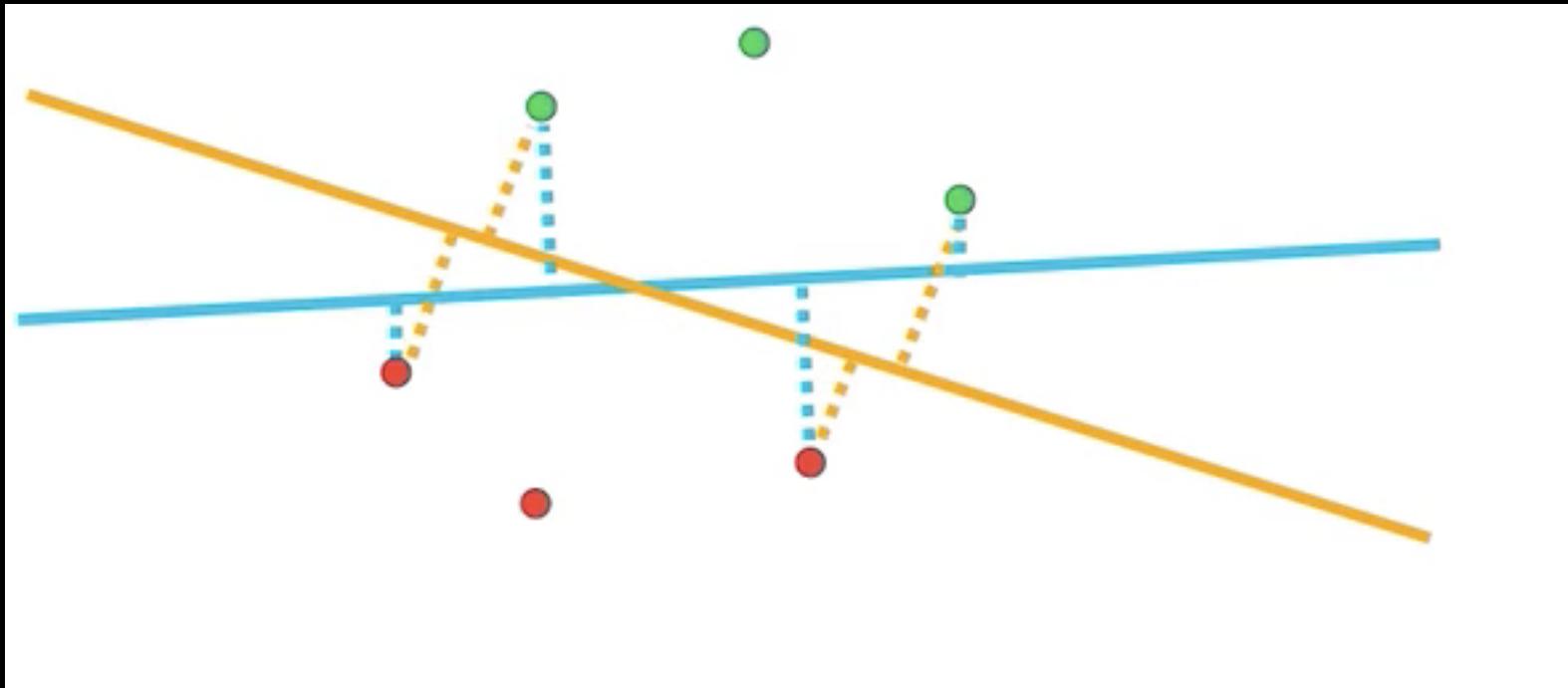


# Cutting Data with Style



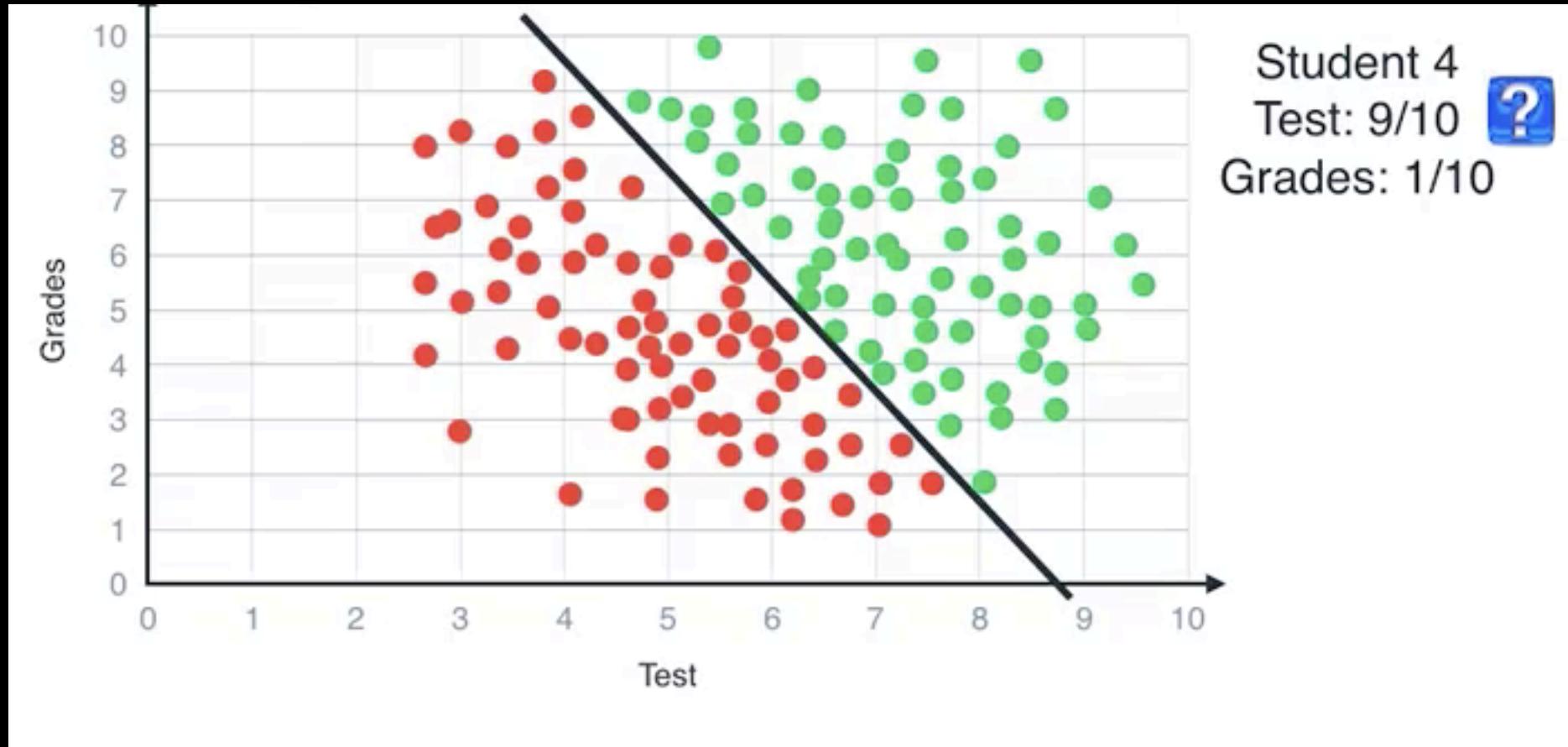
Quiz: Which Line is better?

# Support Vector Machine



Quiz: Which Line is better?

# Acceptance at University



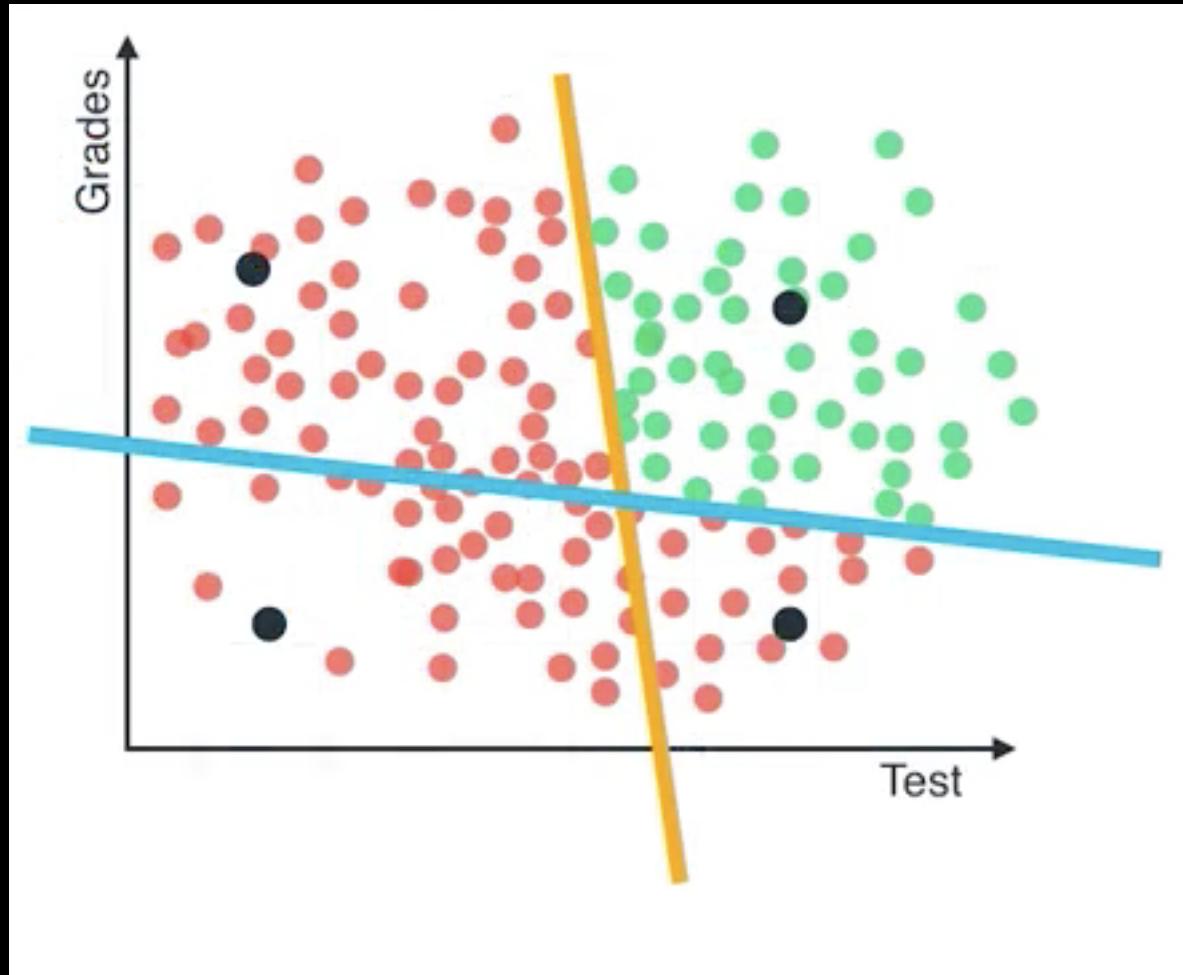
# Acceptance at University



# Acceptance at University



# Neural Network



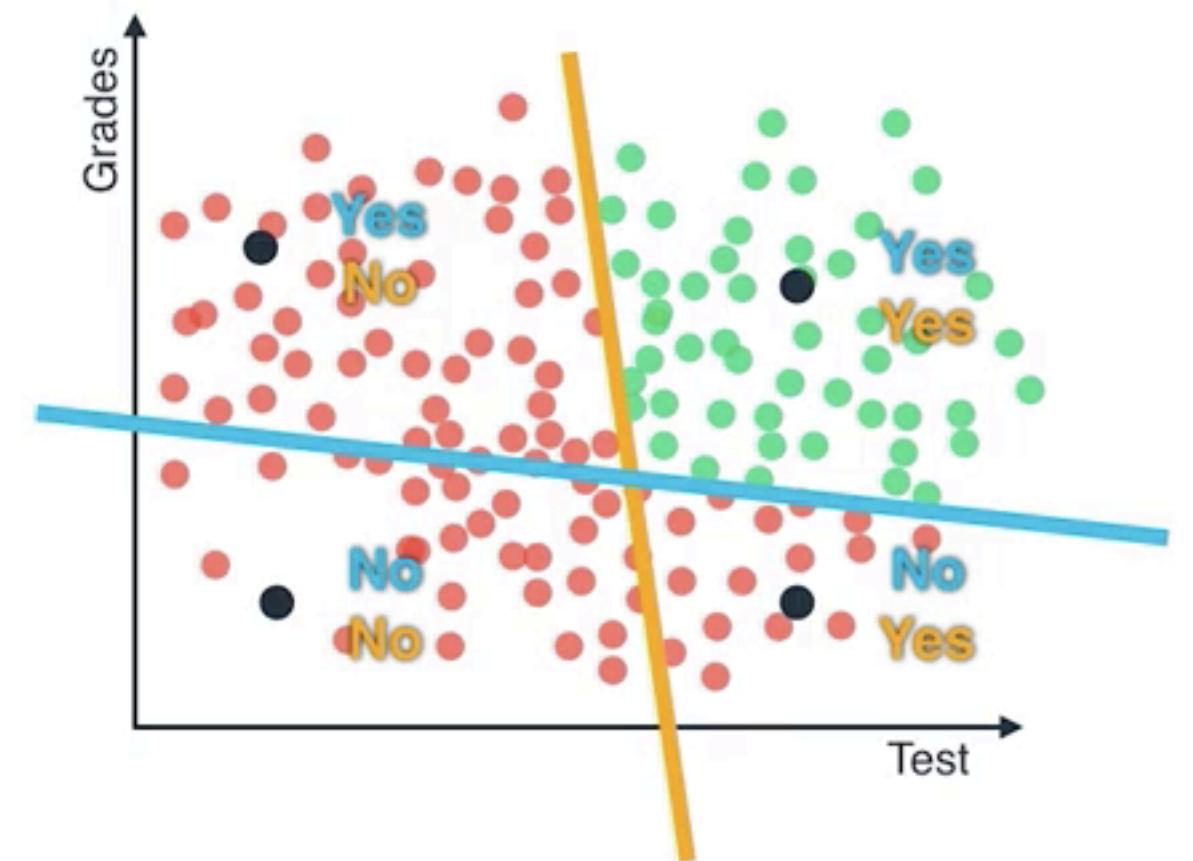
# Neural Network

**Question 1:**

Is the point over the **blue** line?

**Question 2:**

Is the point over the **orange** line?



# Neural Network

**Question 1:**

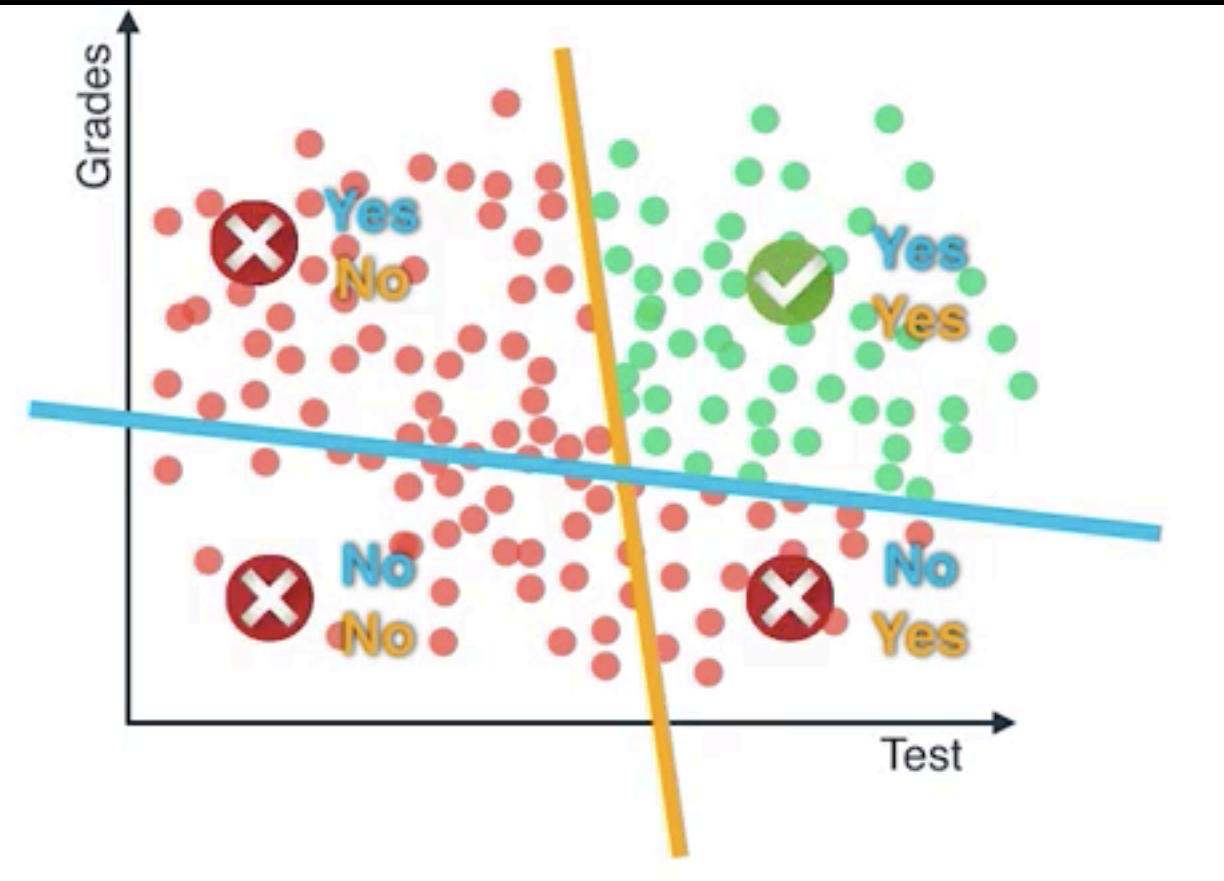
Is the point over the **blue** line?

**Question 2:**

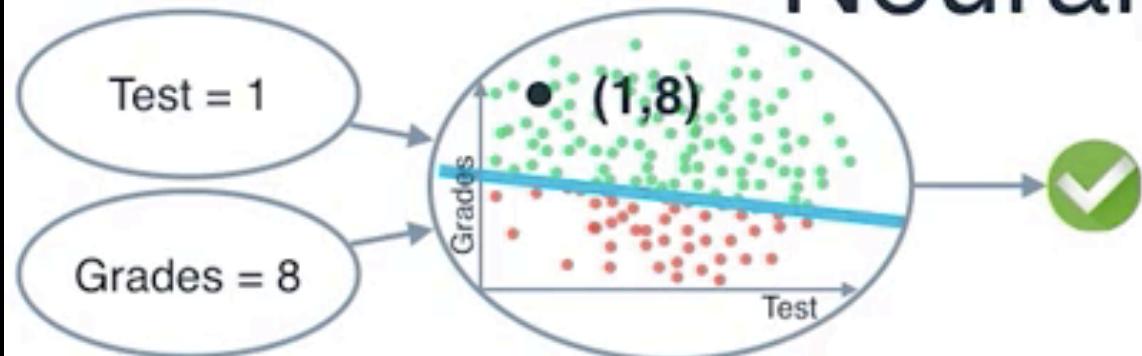
Is the point over the **orange** line?

**Question 3:**

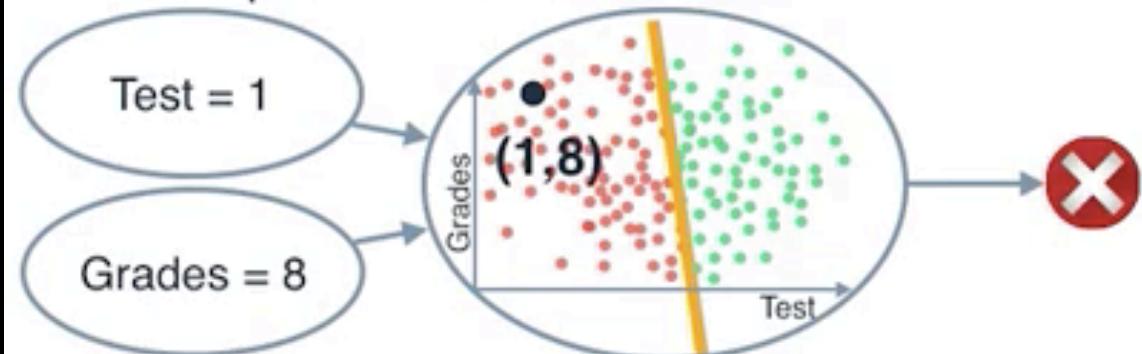
Are the answers to Questions  
1 and 2 both **yes**?



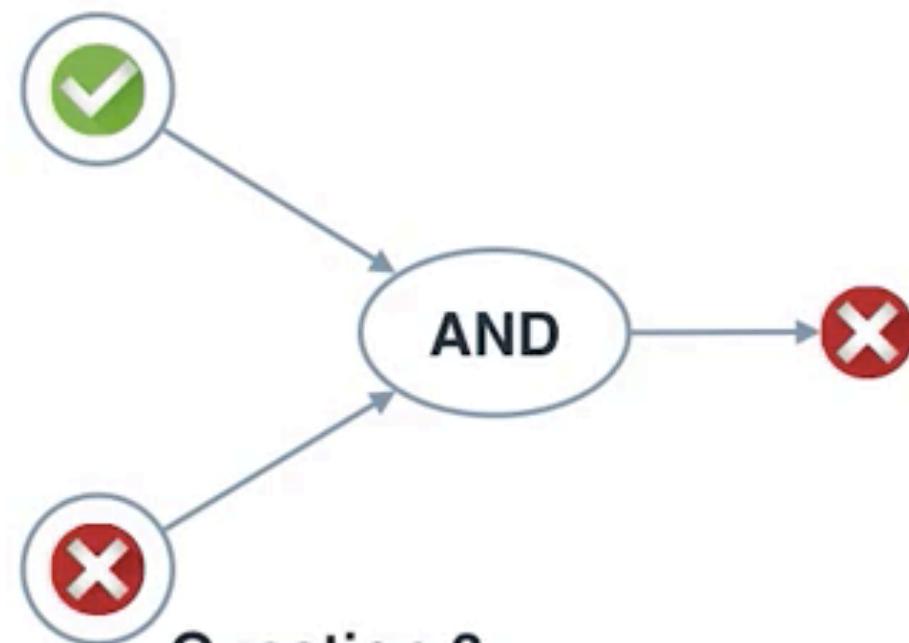
# Neural Network



**Question 1:**  
Is the point over the **blue** line?



**Question 2:**  
Is the point over the **orange** line?

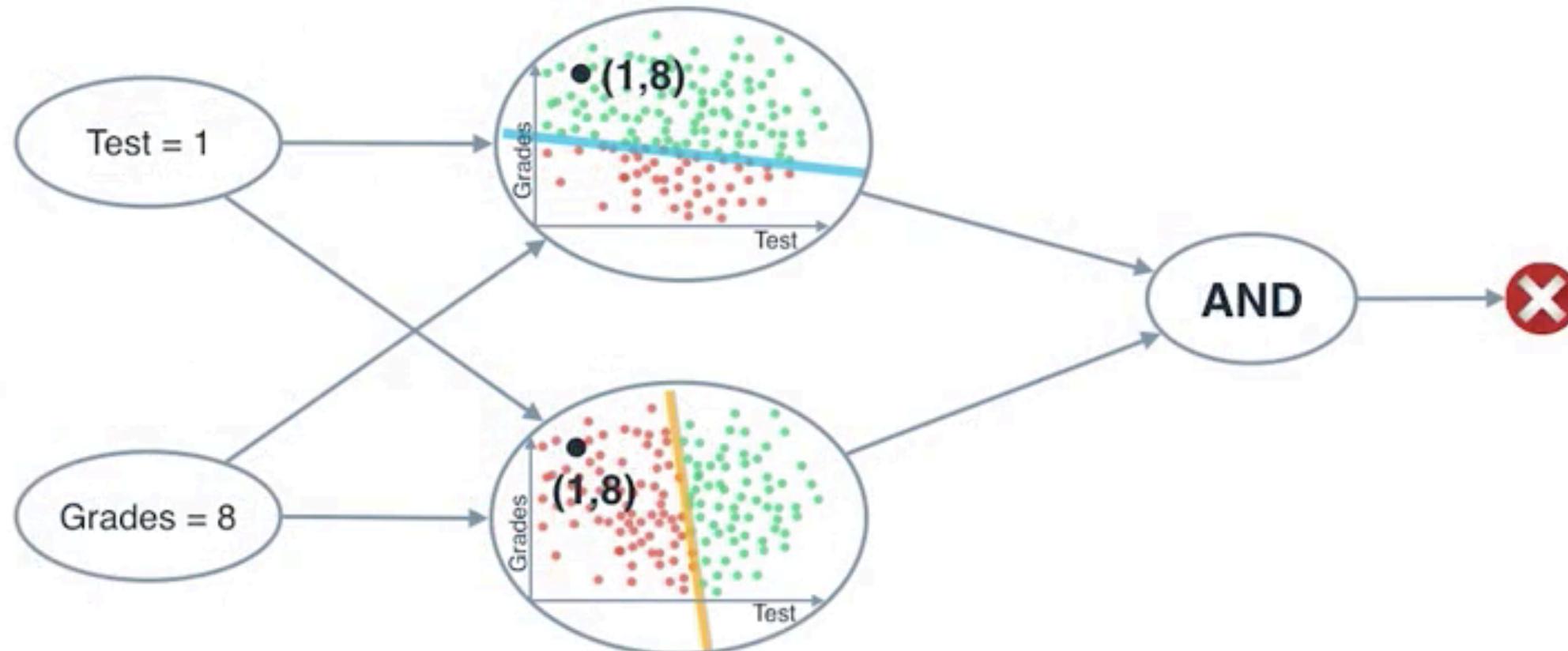


**Question 3:**  
Are the answers to Questions  
1 and 2 both **yes**?

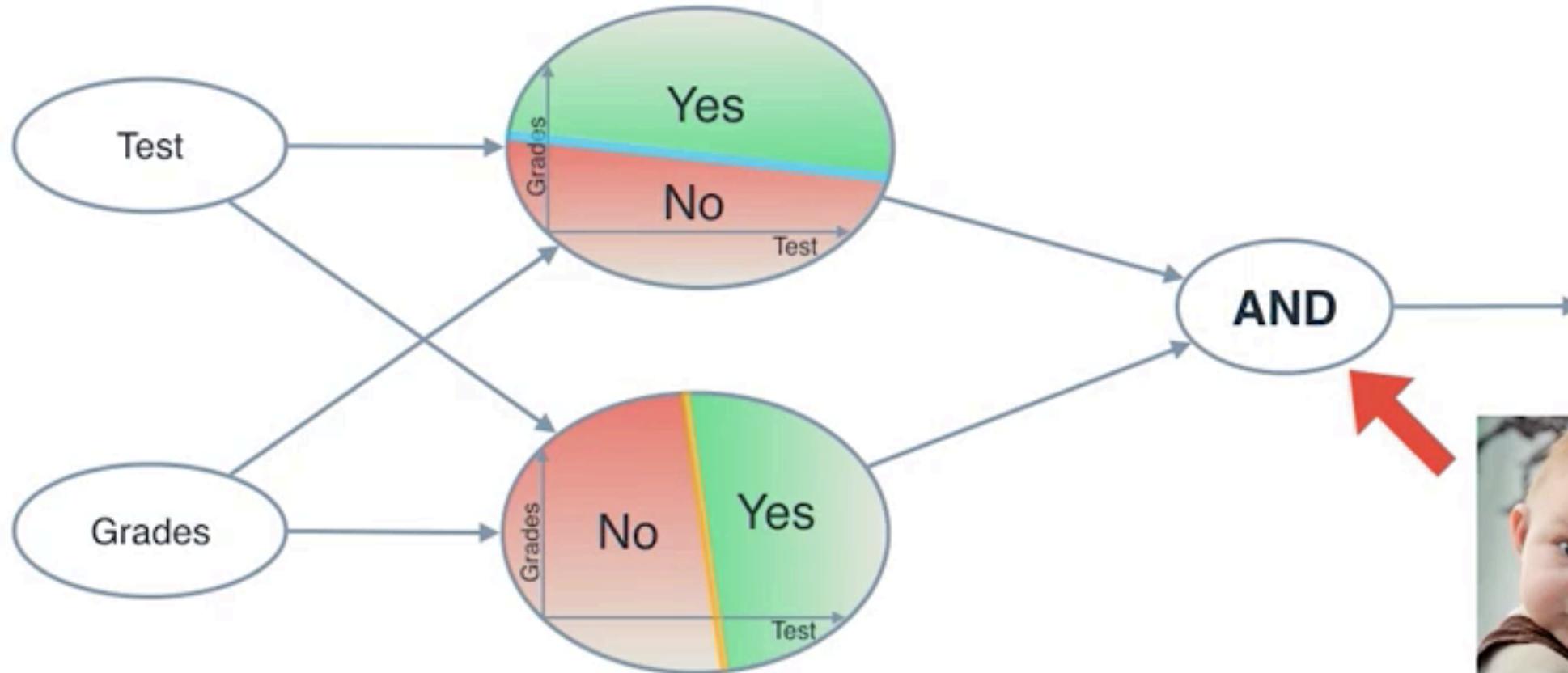
# Neural Network



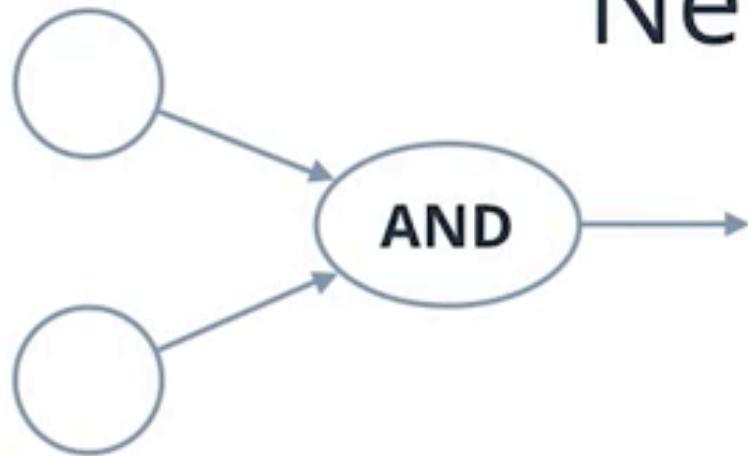
# Neural Network



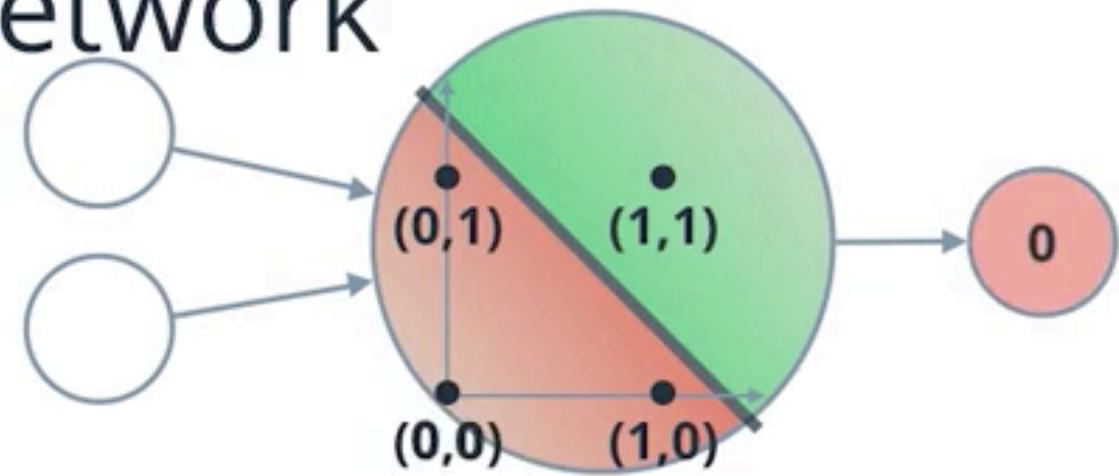
# Neural Network



# Neural Network

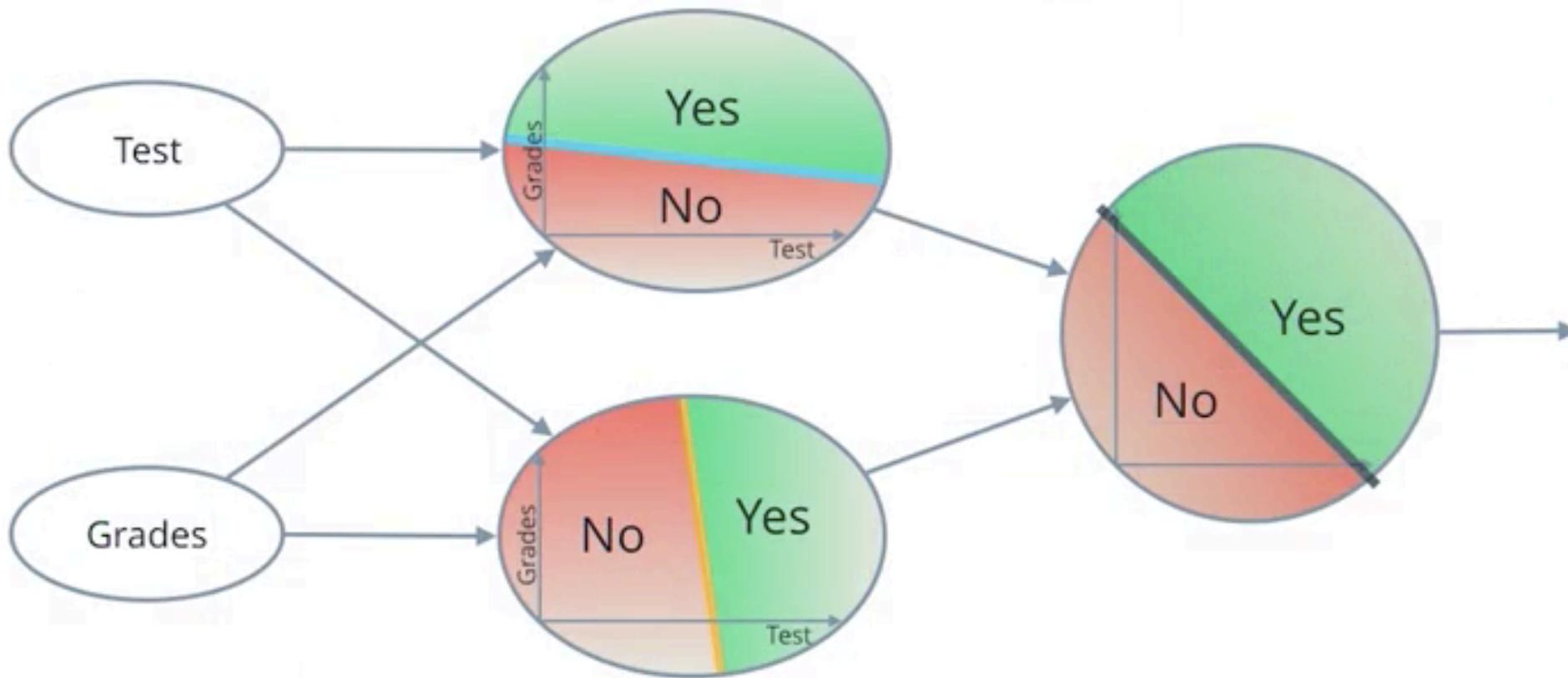


IN	IN	OUT
✓	✓	✓
✓	✗	✗
✗	✓	✗
✗	✗	✗

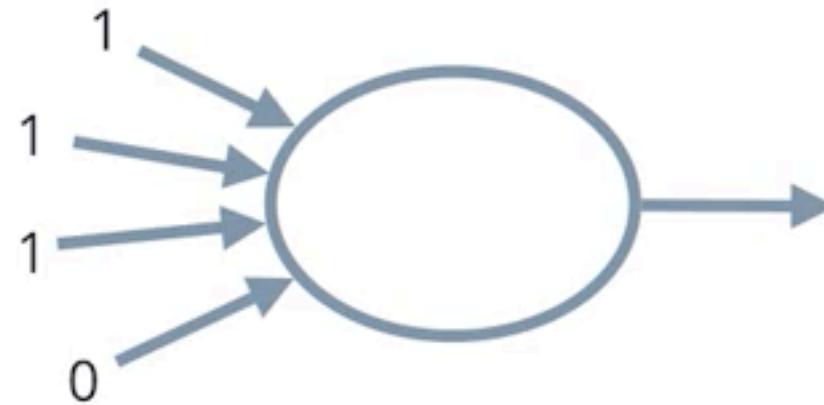


IN	IN	OUT
1	1	1
1	0	0
0	1	0
0	0	0

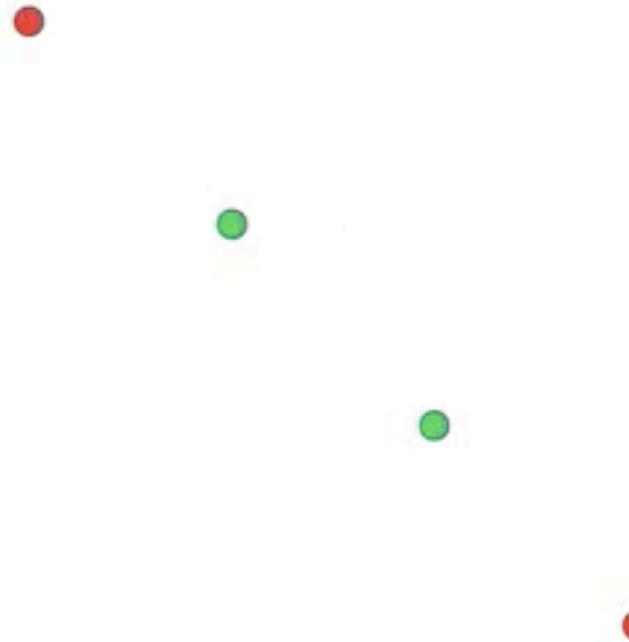
# Neural Network



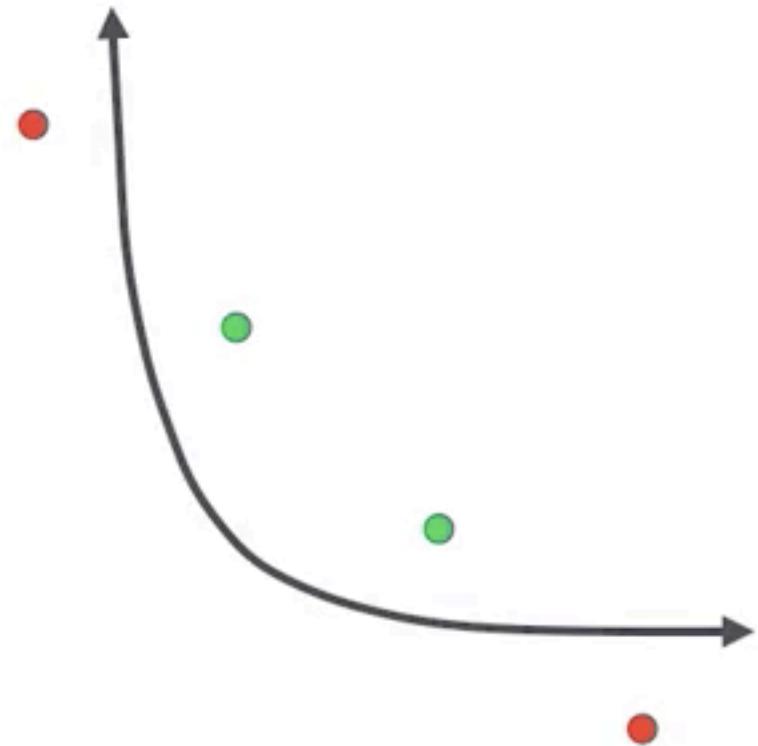
# Neural Network



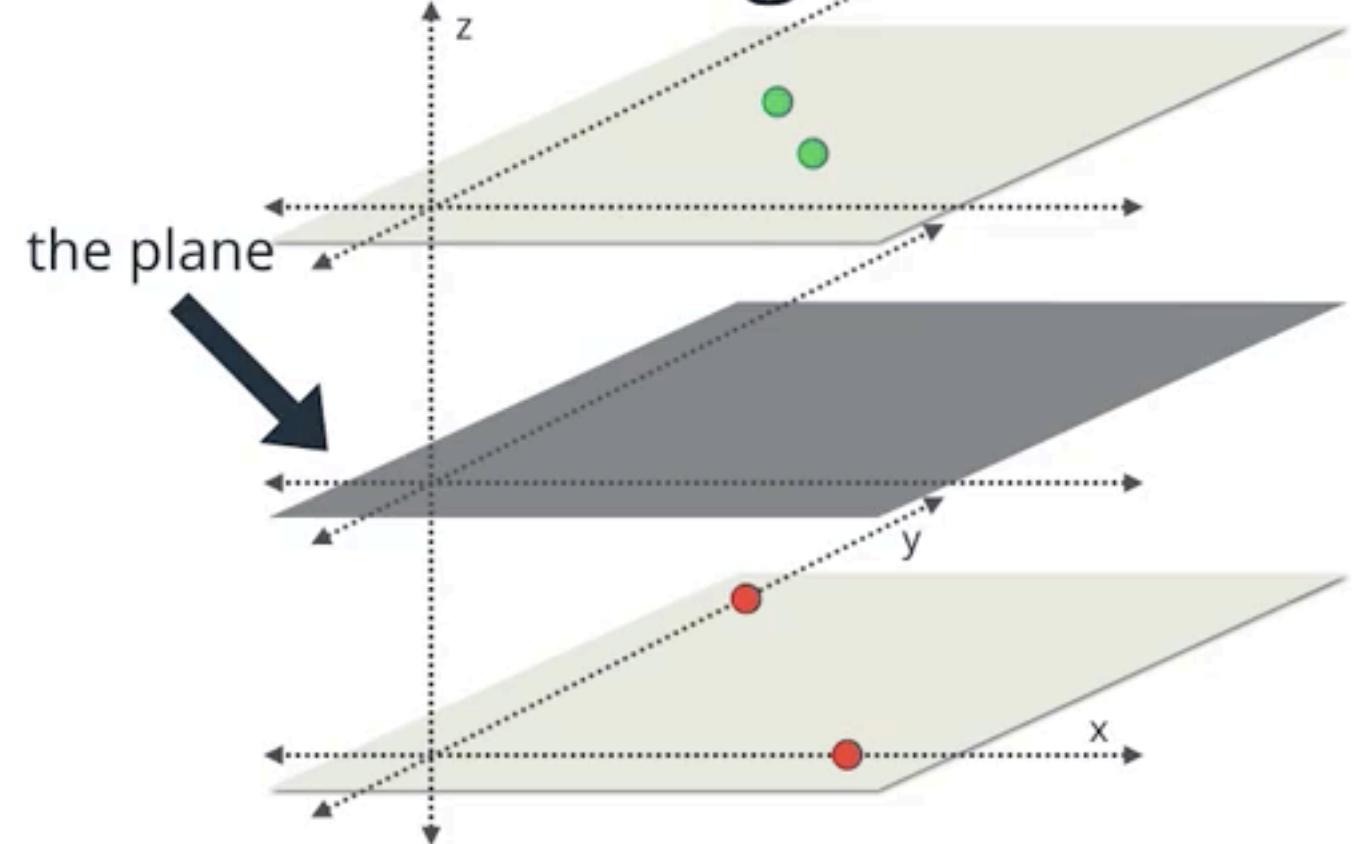
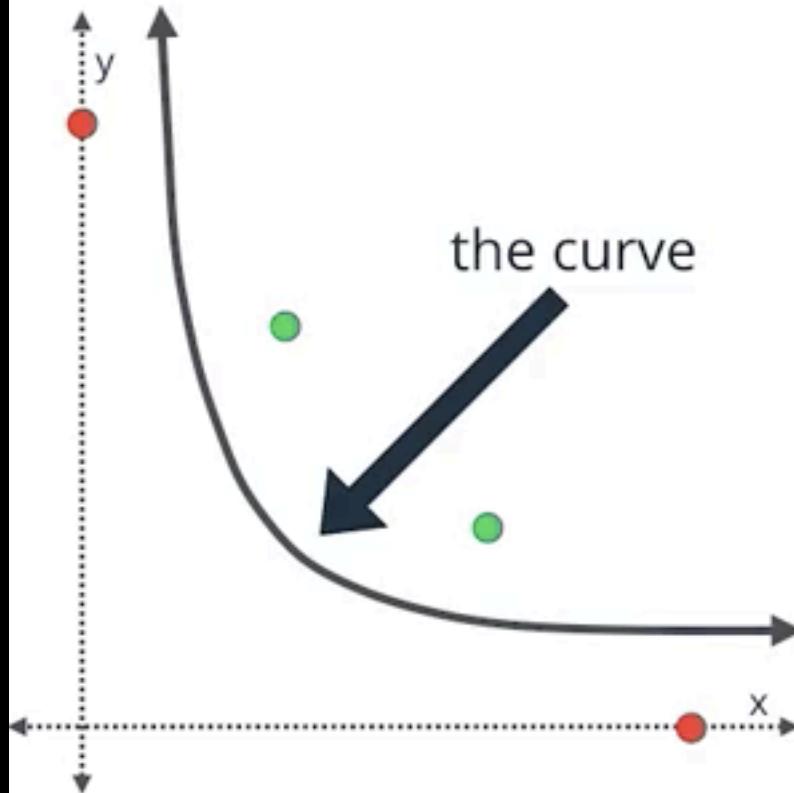
# When a line is not enough...

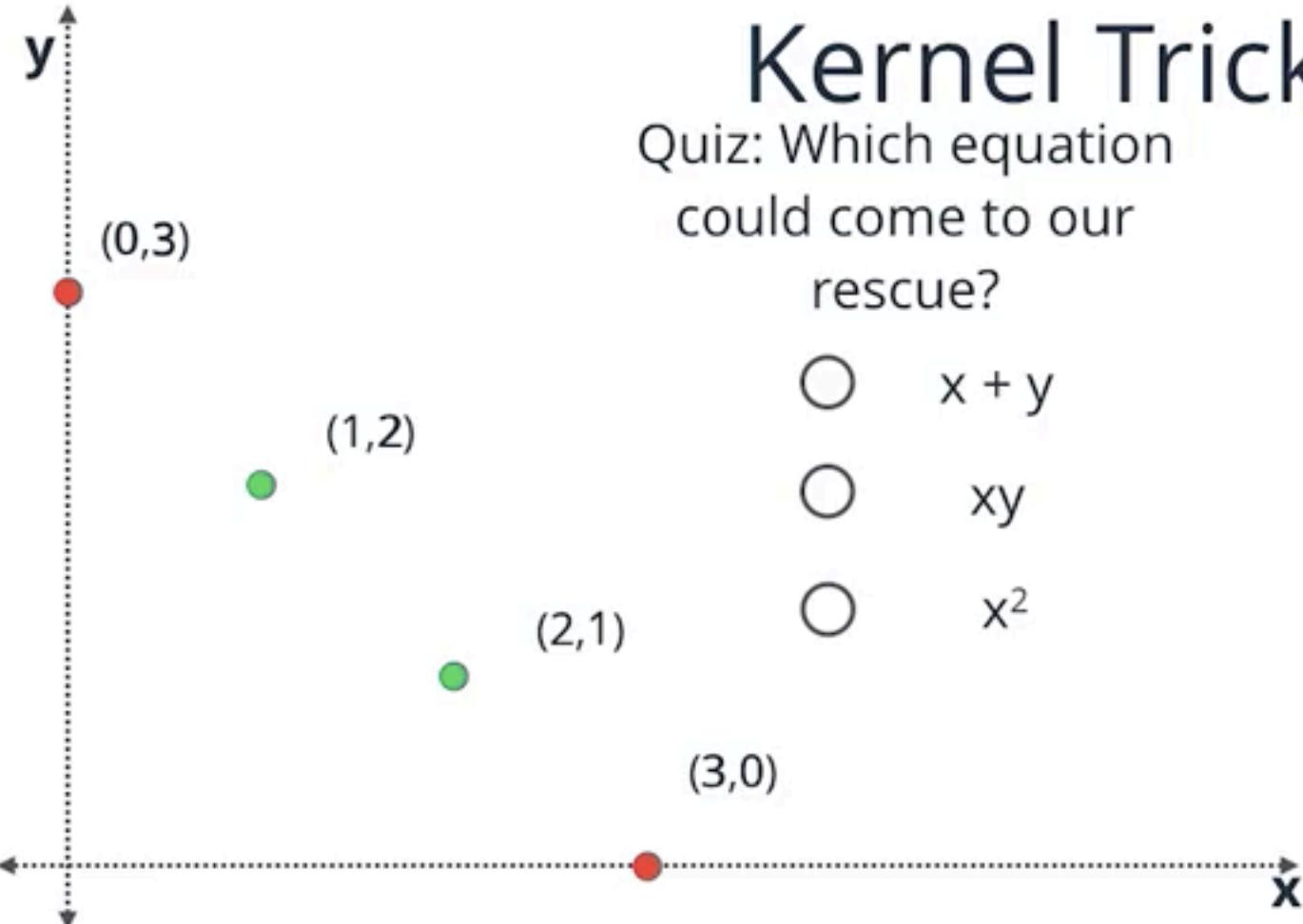


# When a line is not enough...



# When a line is not enough...





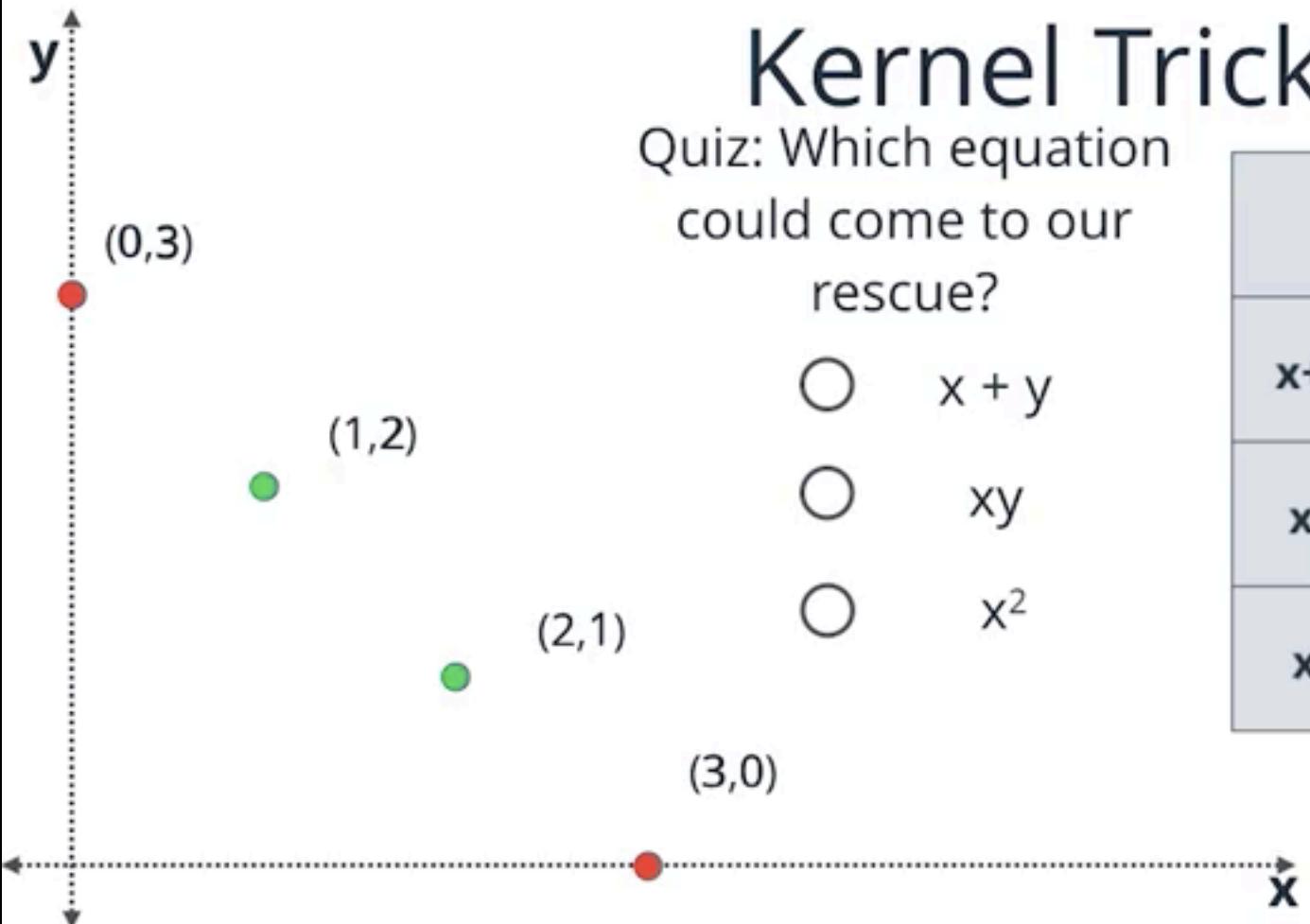
# Kernel Trick

Quiz: Which equation  
could come to our  
rescue?

$x + y$

$xy$

$x^2$

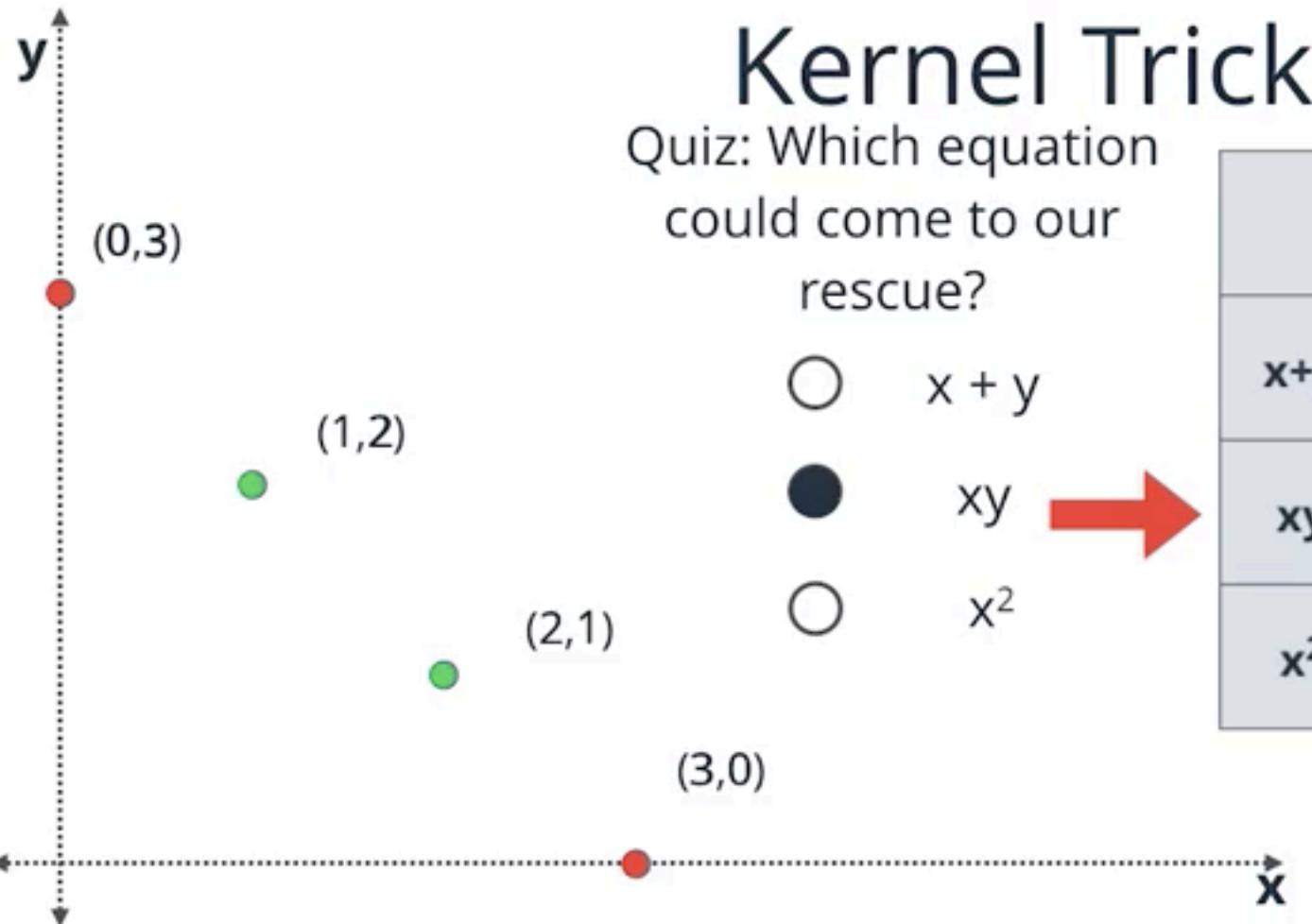


# Kernel Trick

Quiz: Which equation could come to our rescue?

- $x + y$
- $xy$
- $x^2$

	(0,3)	(1,2)	(2,1)	(3,0)
$x+y$	3	3	3	3
$xy$	0	2	2	0
$x^2$	0	1	4	9



# Kernel Trick

Quiz: Which equation could come to our rescue?

- $x + y$
- $xy$
- $x^2$

A 4x5 grid table comparing three functions across five data points. Red numbers indicate a value of 3, while green numbers indicate a value of 2.

	(0,3)	(1,2)	(2,1)	(3,0)
$x+y$	3	3	3	3
$xy$	0	2	2	0
$x^2$	0	1	4	9



02/02/18





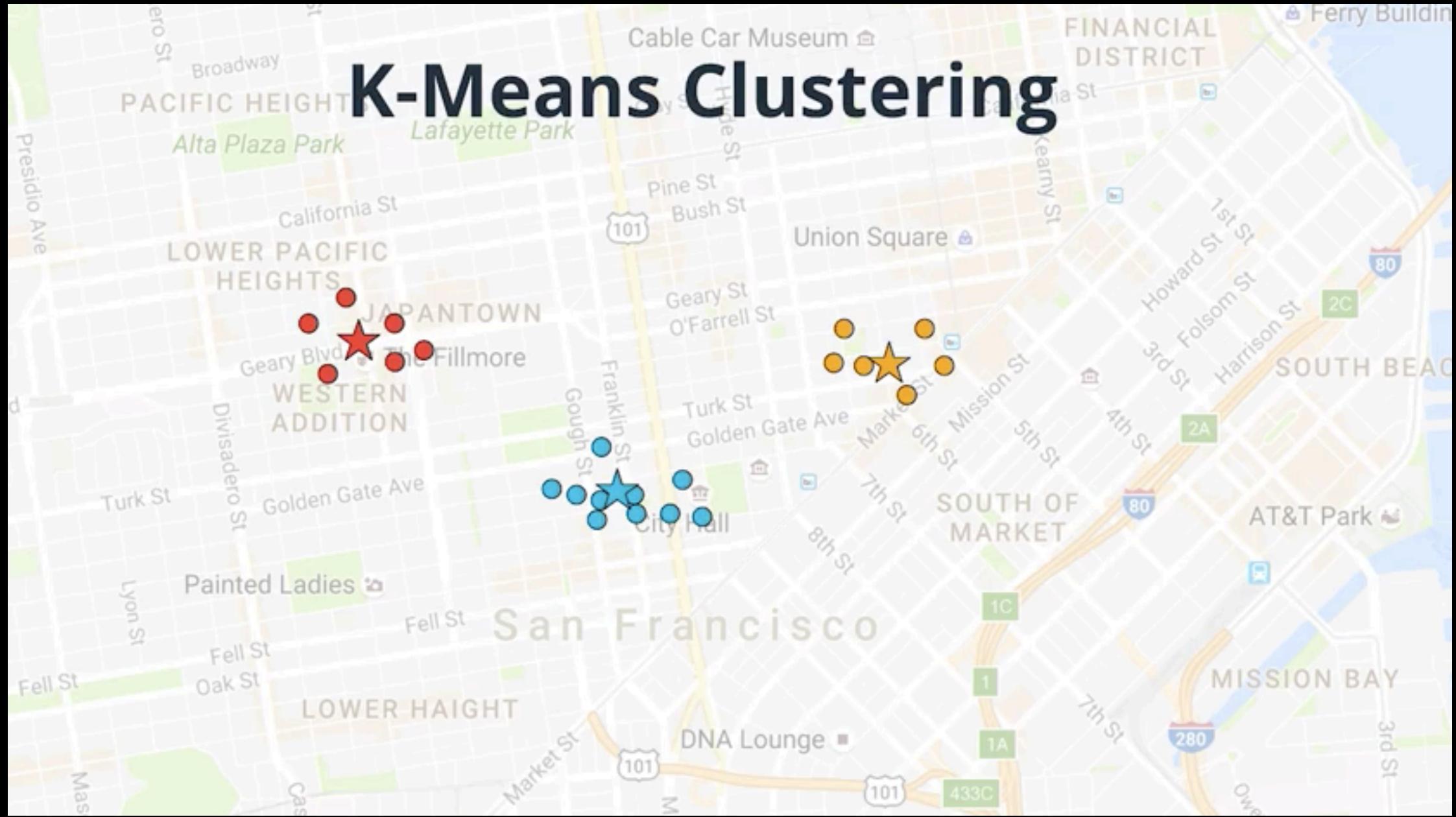
# Did anyone say pizza?



# Did anyone say pizza?



# Did anyone say pizza?



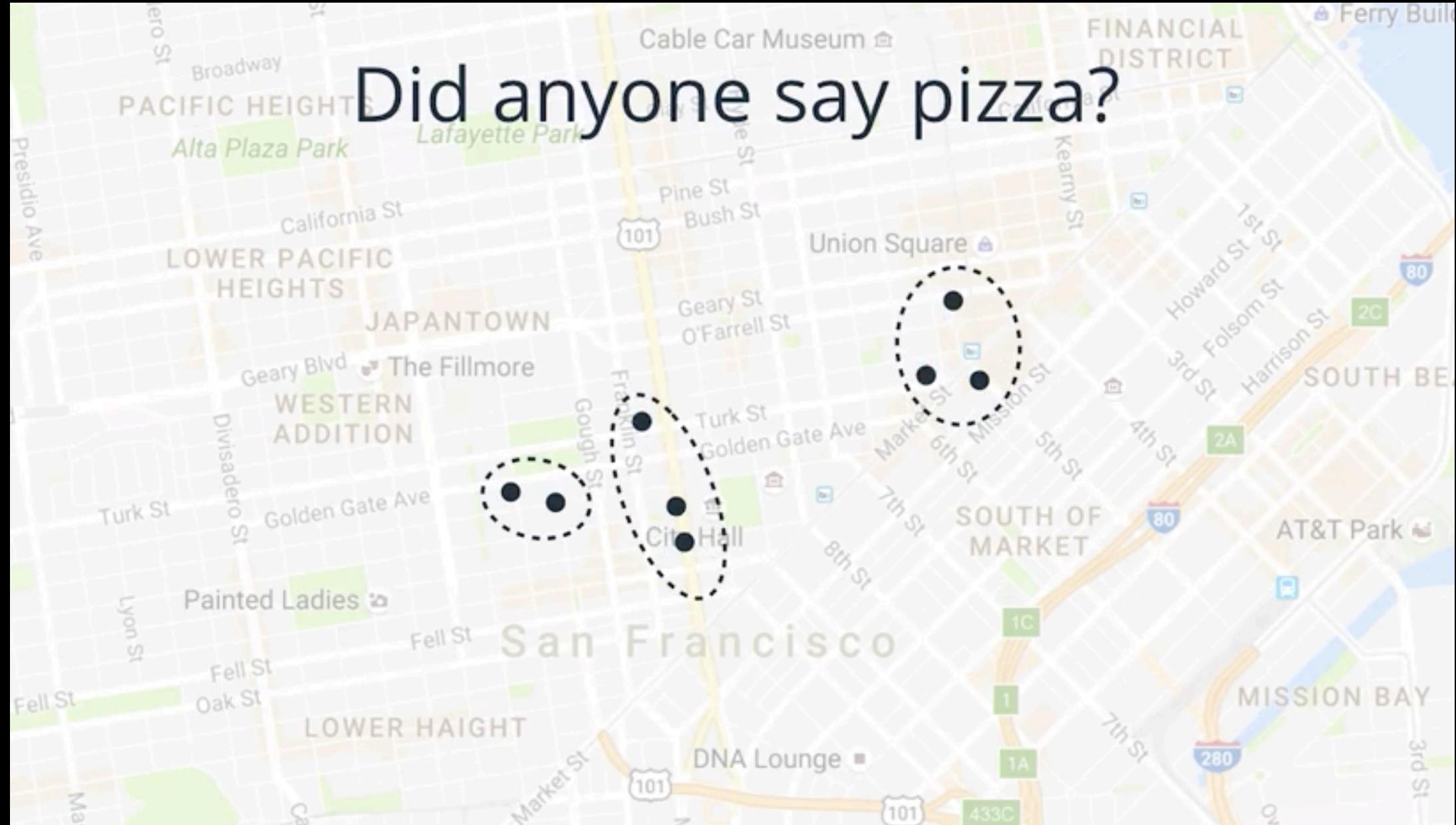


# Did anyone say pizza?



# Did anyone say pizza?





# Did anyone say pizza?



# Did anyone say pizza?



02/02/18

