

DEV DAYS VILNIUS 2018

---

# TREND ANALYSIS AND MACHINE LEARNING IN PHP

DEV DAYS VILNIUS 2018

---

# TREND ANALYSIS AND MACHINE LEARNING IN PHP

ME?



Sam  
Knows

oh  
Communities



PHP-FIG

---

**MICHAEL CULLUM**

**@MICHAELCULLUMUK**

# STATISTICS

# ARTIFICIAL INTELLIGENCE





# MACHINE LEARNING

# MACHINE LEARNING



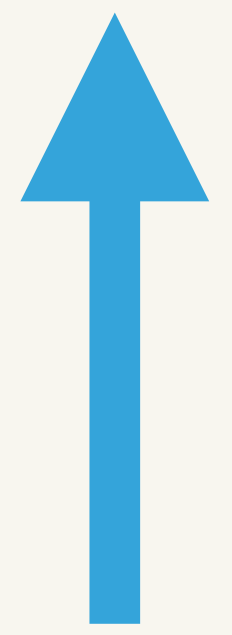
**Cause  
&  
Effect**



Process

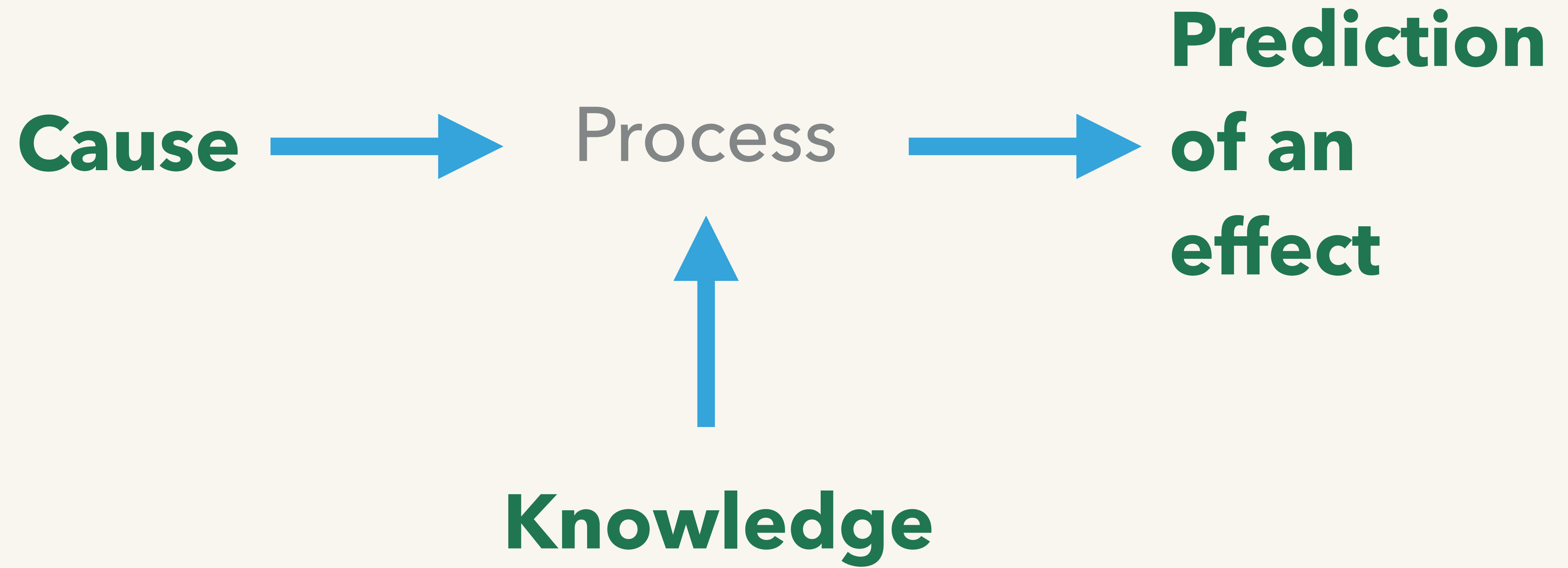


**Knowledge**

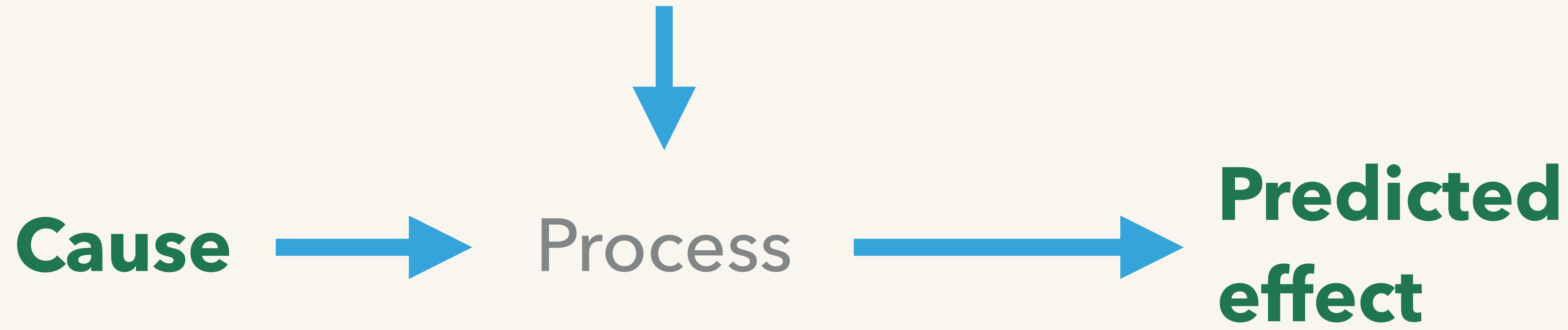
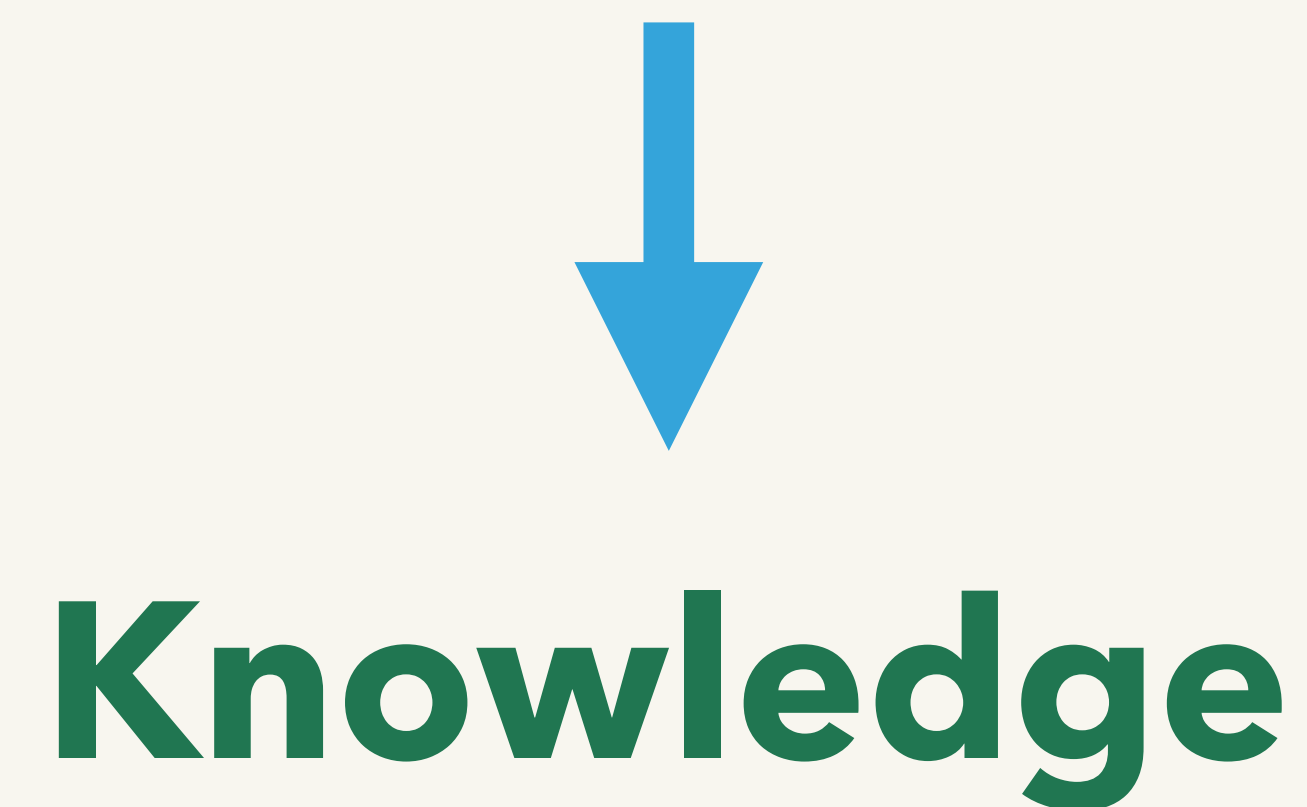


Context

**LEARNING**



USING



**Cause    Effect**

$$1 + 1 = 2$$

$$1 + 2 = 3$$

$$2 + 1 = 3$$

$$1 + 3 = 4$$

$$3 + 1 = 4$$

**EXAMPLE**

# Knowledge

$$1 + 1 = 2$$

$$1 + 2 = 3$$

$$2 + 1 = 3$$

$$1 + 3 = 4$$

$$3 + 1 = 4$$

# Knowledge

$$1 + 1 = 2$$

$$1 + 2 = 3$$

$$2 + 1 = 3$$

$$1 + 3 = 4$$

$$3 + 1 = 4$$

**Cause**

$$3 + 1 =$$

**Predicted Effect**

4

**EXAMPLE**

$$1x + 3 = 4$$

$$1x + 3 = 5$$

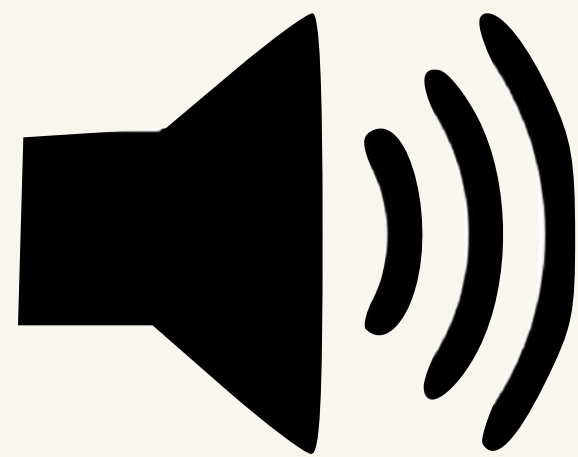
$$1x + 4 = 4$$

$$2x + 3 = 7$$

$$x = 10 \quad 1x + 4 = ?$$



"eggs"



**INFERENCE**

# MACHINE LEARNING AS A 4-STEP PROCESS

# 1. ACQUIRE DATA

# 2. TRAIN MODEL

# 3. ASK YOUR QUESTION

# 4. GET PREDICTED ANSWER

# 1. ACQUIRE DATA



**GOOD LUCK**

# 2. TRAIN MODEL

**SUPERVISED LEARNING**

**UNSUPERVISED LEARNING**

# SUPERVISED LEARNING

# KNOWN OUTCOMES

**QUANTATIVE – CLASSIFICATION**  
**QUALITATIVE – REGRESSION**

**QUANTATIVE – CLASSIFICATION**

**QUALITATIVE – REGRESSION**

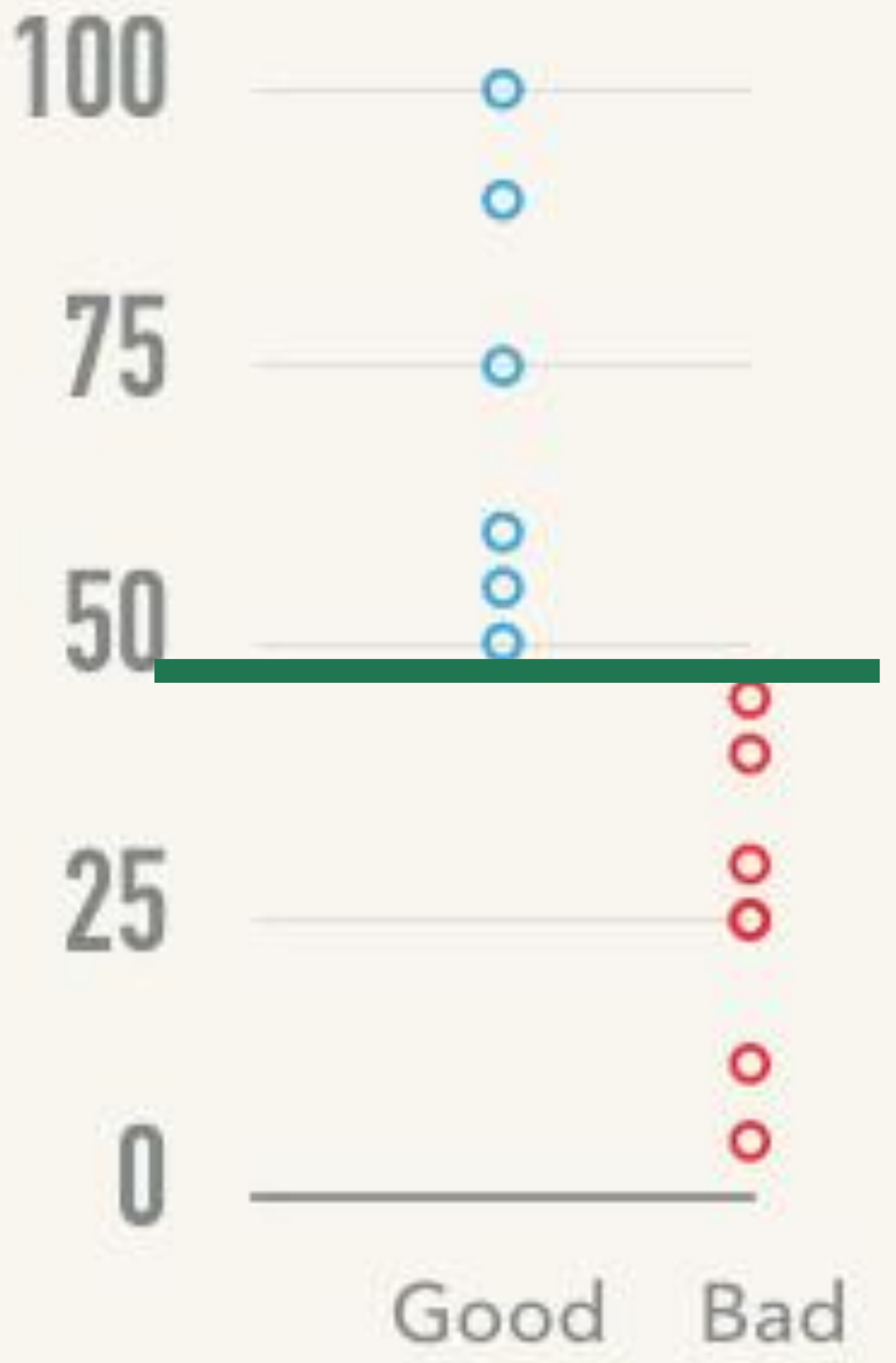


Rating	Conclusion
100	Good
25	Bad
50	Good
40	Bad

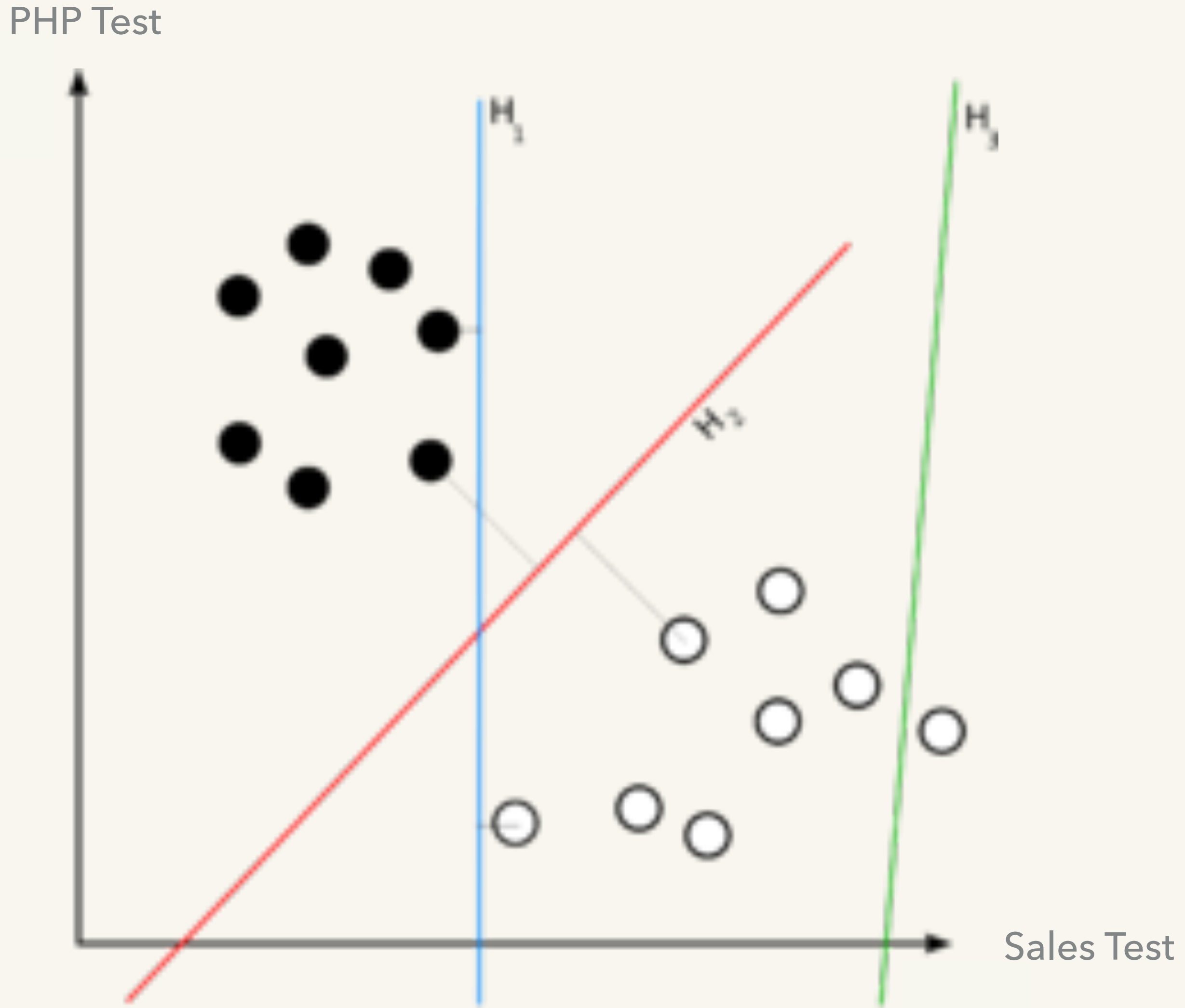


**CLASSIFICATION**

Rating	Conclusion
100	Good
25	Bad
50	Good
40	Bad



# CLASSIFICATION



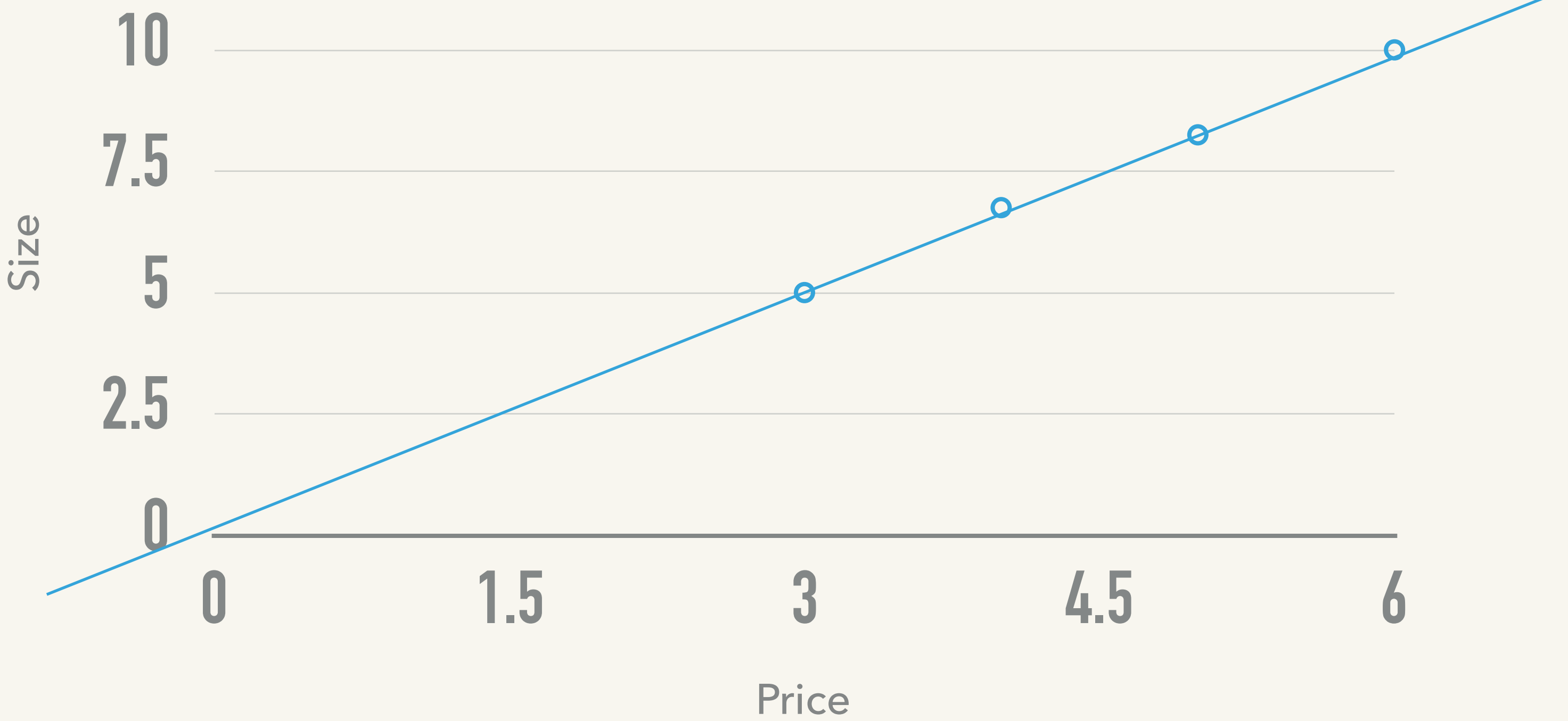
Item	Value
Black	PHP Devs
White	Sales

**LINEAR CLASSIFIER**

QUANTITATIVE – CLASSIFICATION

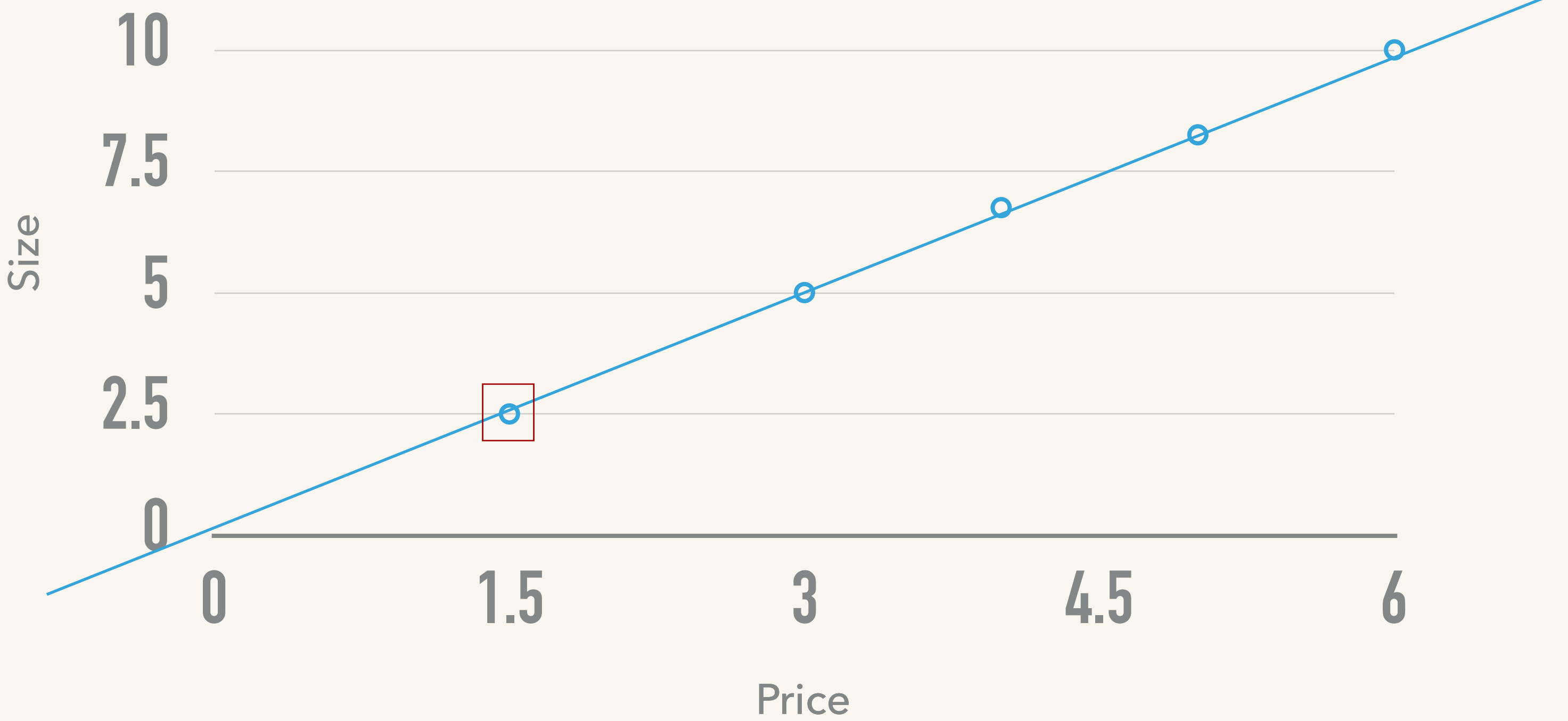
QUALITATIVE – REGRESSION

Price	Size
5.00	8.25
6.00	10
4.00	6.75
3.00	5



**REGRESSION**

Price	Size
5.00	8.25
6.00	10
4.00	6.75
3.00	5



**REGRESSION**

# UNSUPERVISED LEARNING



# DISCOVERY

# ASSOCIATION CLUSTERING

# ASSOCIATION CLUSTERING

Item 1	Item 2
T-shirt	Shorts
Shorts	T-shirt
Suit	Black Shoes
Socks	Underwear

Item 1	Item 2
Black shoes	Suit
Socks	Underwear
Underwear	Socks
T-shirt	Socks

**ASSOCIATION**

Item 1	Item 2
T-shirt	Shorts
Shorts	T-shirt
Suit	Black Shoes
Socks	Underwear

Item 1	Item 2
Black shoes	Suit
Socks	Underwear
Underwear	Socks
T-shirt	Socks

People who buy socks, also often buy underwear

People who who buy underwear always buy socks

# ASSOCIATION

Item 1	Item 2
T-shirt	Shorts
Shorts	T-shirt
<b>Suit</b>	<b>Black Shoes</b>
Socks	Underwear

Item 1	Item 2
<b>Black shoes</b>	<b>Suit</b>
Socks	Underwear
Underwear	Socks
T-shirt	Socks

People who buy socks, also often buy underwear

People who who buy underwear always buy socks

People who buy suits always buy black shoes

People who buy black shoes, always buy suits

**ASSOCIATION**

Item 1	Item 2
T-shirt	Shorts
Shorts	T-shirt
Suit	Black Shoes
Socks	Underwear

Item 1	Item 2
Black shoes	Suit
Socks	Underwear
Underwear	Socks
T-shirt	Socks

People who buy socks, also often buy underwear

People who who buy underwear always buy socks

People who buy suits always buy black shoes

People who buy black shoes, always buy suits

People who buy shorts always buy t-shirts

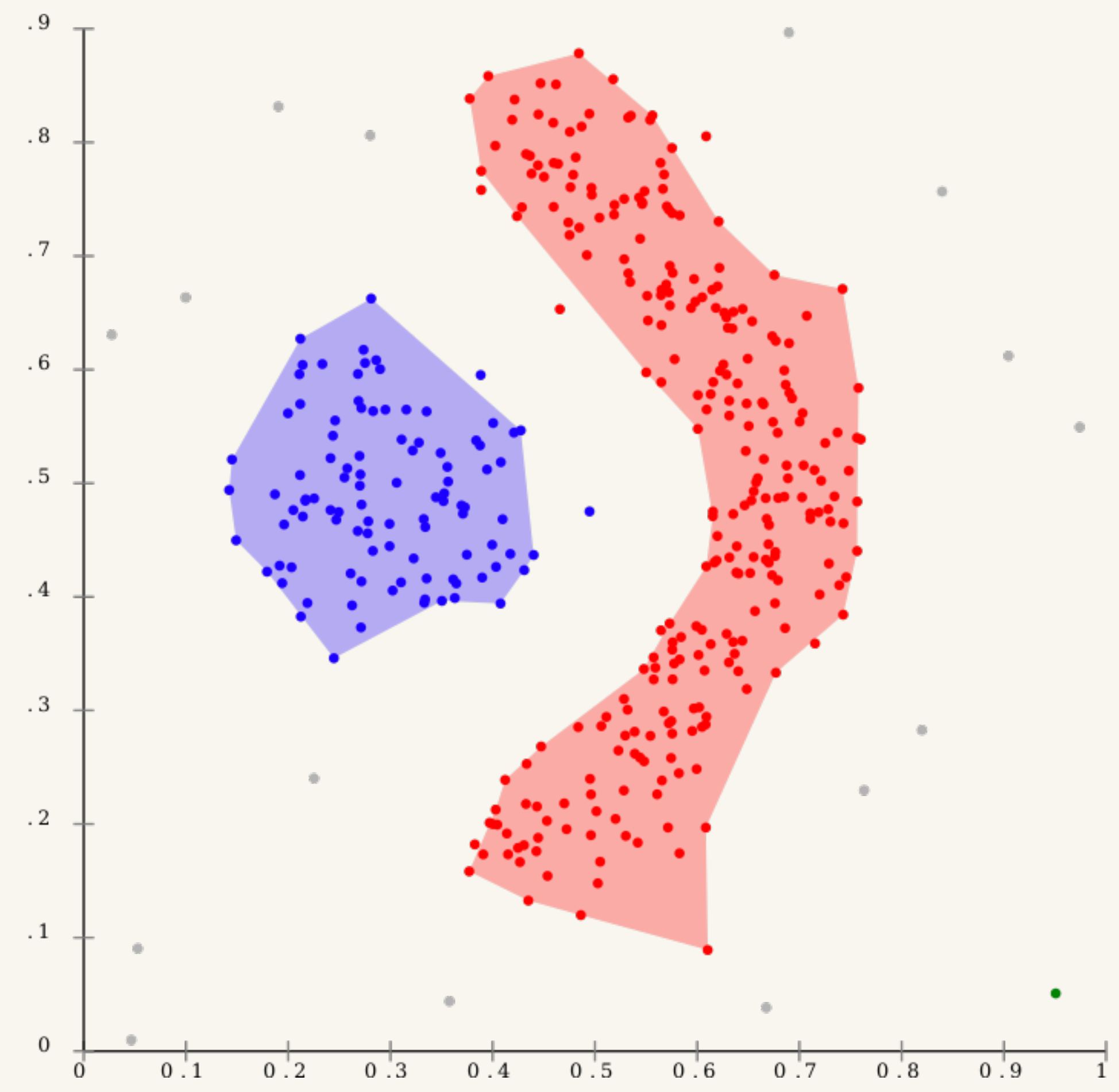
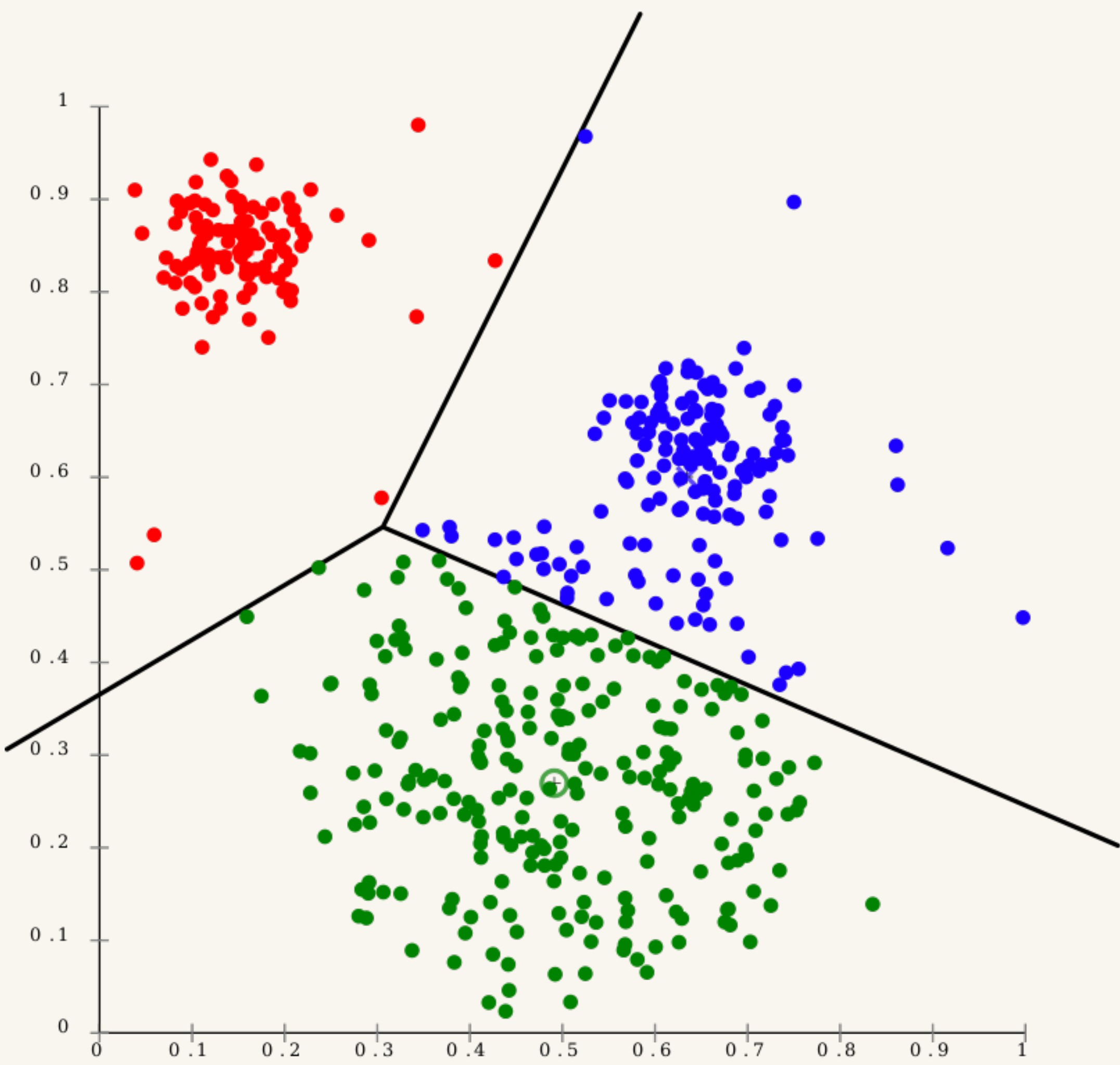
People who buy t-shirts, also often buy shorts

**ASSOCIATION**

# ASSOCIATION CLUSTERING



CLUSTER ANALYSIS IS THE TASK OF  
GROUPING A SET OF OBJECTS IN SUCH A  
WAY THAT OBJECTS IN THE SAME GROUP  
ARE MORE SIMILAR TO EACH OTHER  
THAN TO THOSE IN OTHER GROUPS.



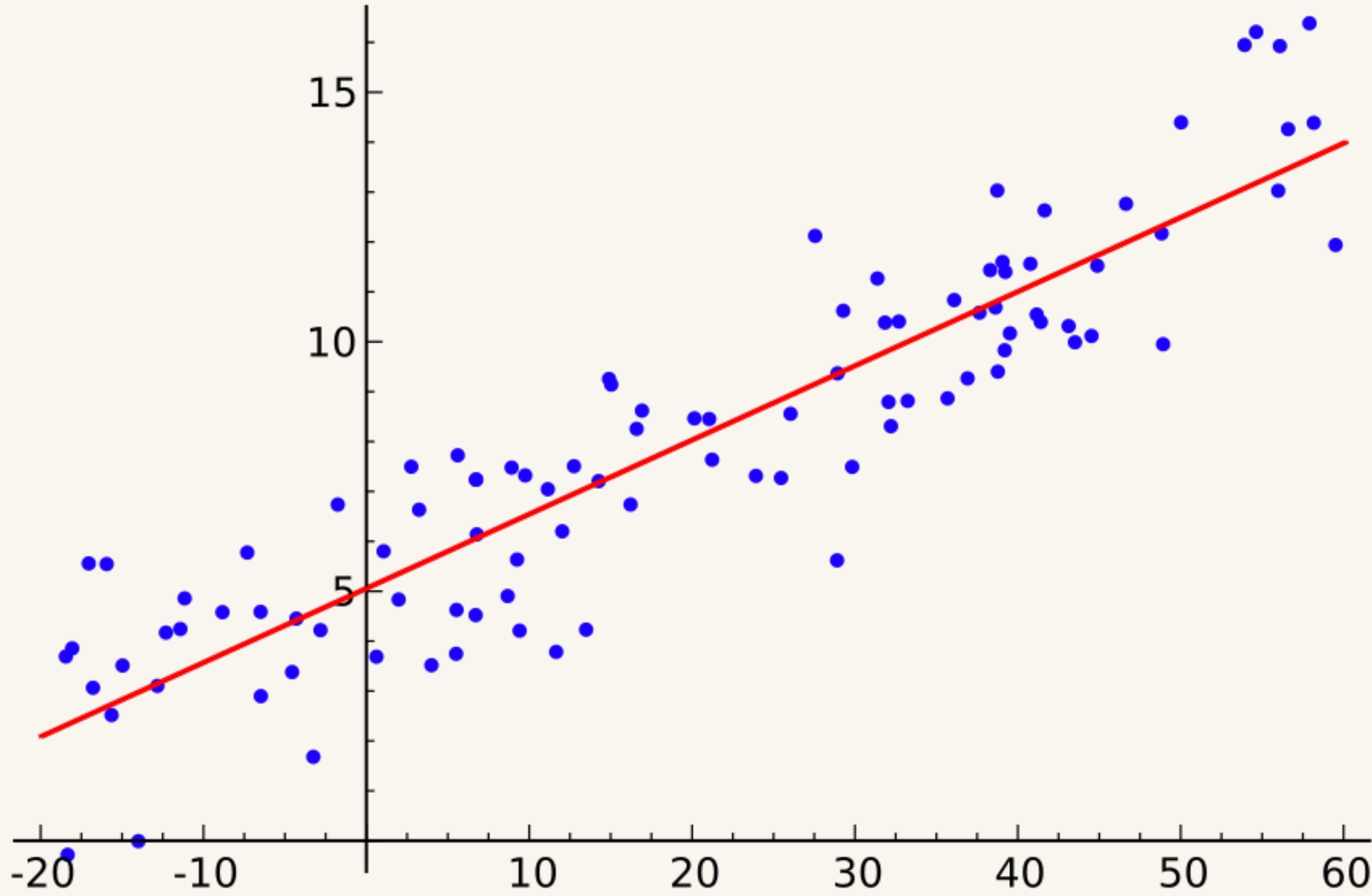
# CLUSTER ANALYSIS

**3. QUESTION**

**4. PREDICTED ANSWER**

# ALGORITHMS

# LEAST SQUARES



**LEAST SQUARES REGRESSION LINE**

# MATHS

$$A + BX = Y$$



- ▶ Each point has an x and y value
- ▶ We need an equation of a line
- ▶ We move the line an infinite number of times
- ▶ Each time, we draw a box between every point, and the line, with one corner on the line, and another on the point
- ▶ The correct line is the one where the sum of the area of all the squares is smallest

# DEMO

# PHP

# DEMO



# PHP-ML

[php-ai/php-ml](https://github.com/php-ai/php-ml)

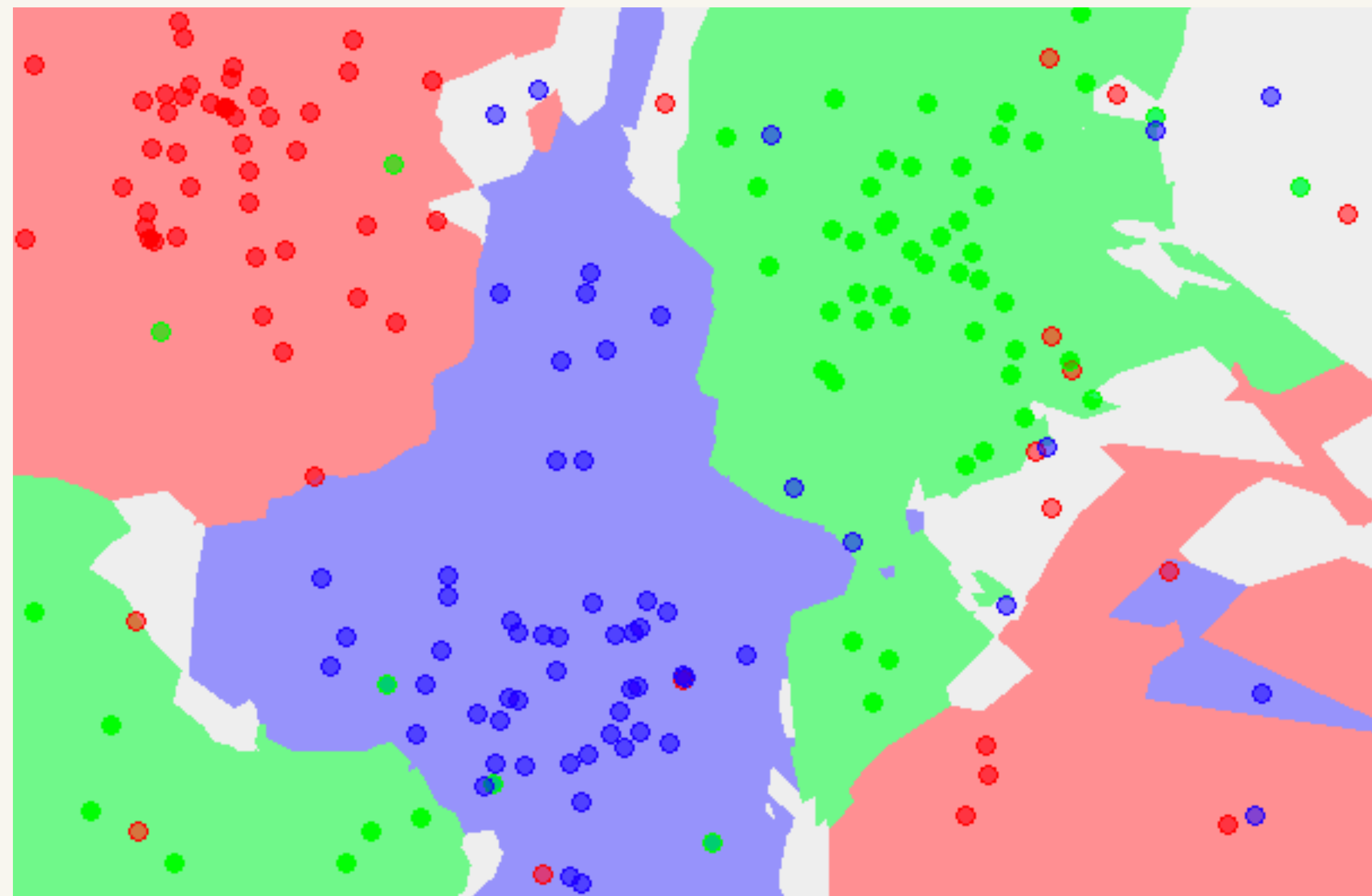
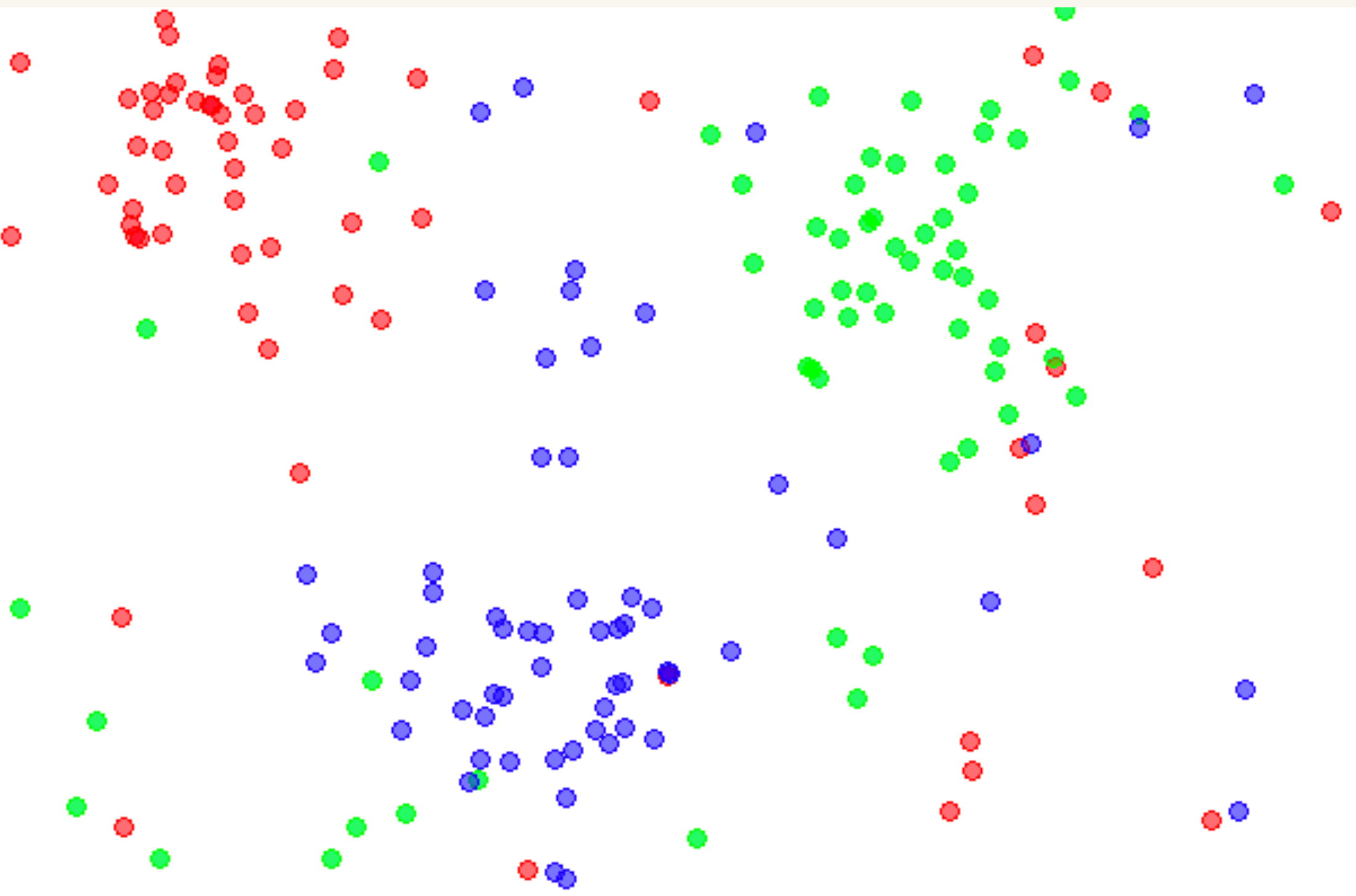
DEMOS

```
$samples = [[60], [61], [62], [63], [65]];
$targets = [3.1, 3.6, 3.8, 4, 4.1];

$regression = new LeastSquares();
$regression->train($samples, $targets);

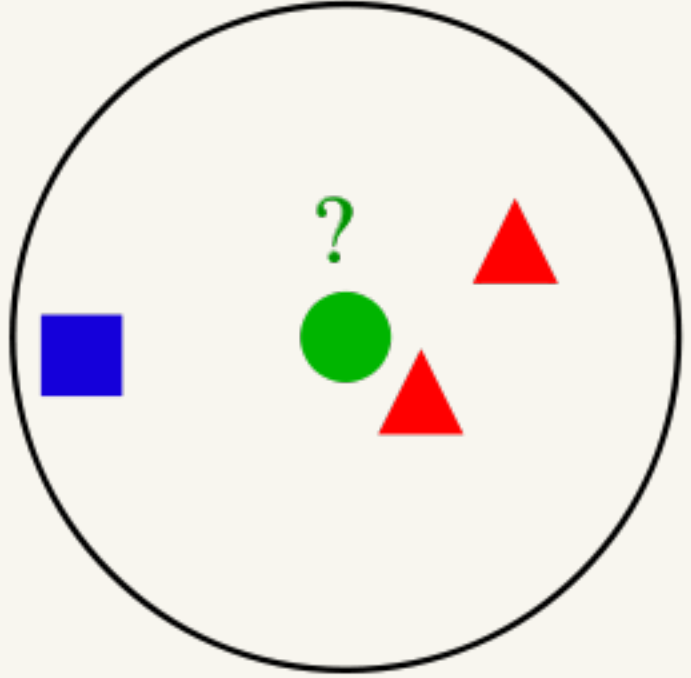
echo $regression->predict([64]);
```

# NEAREST NEIGHBOUR

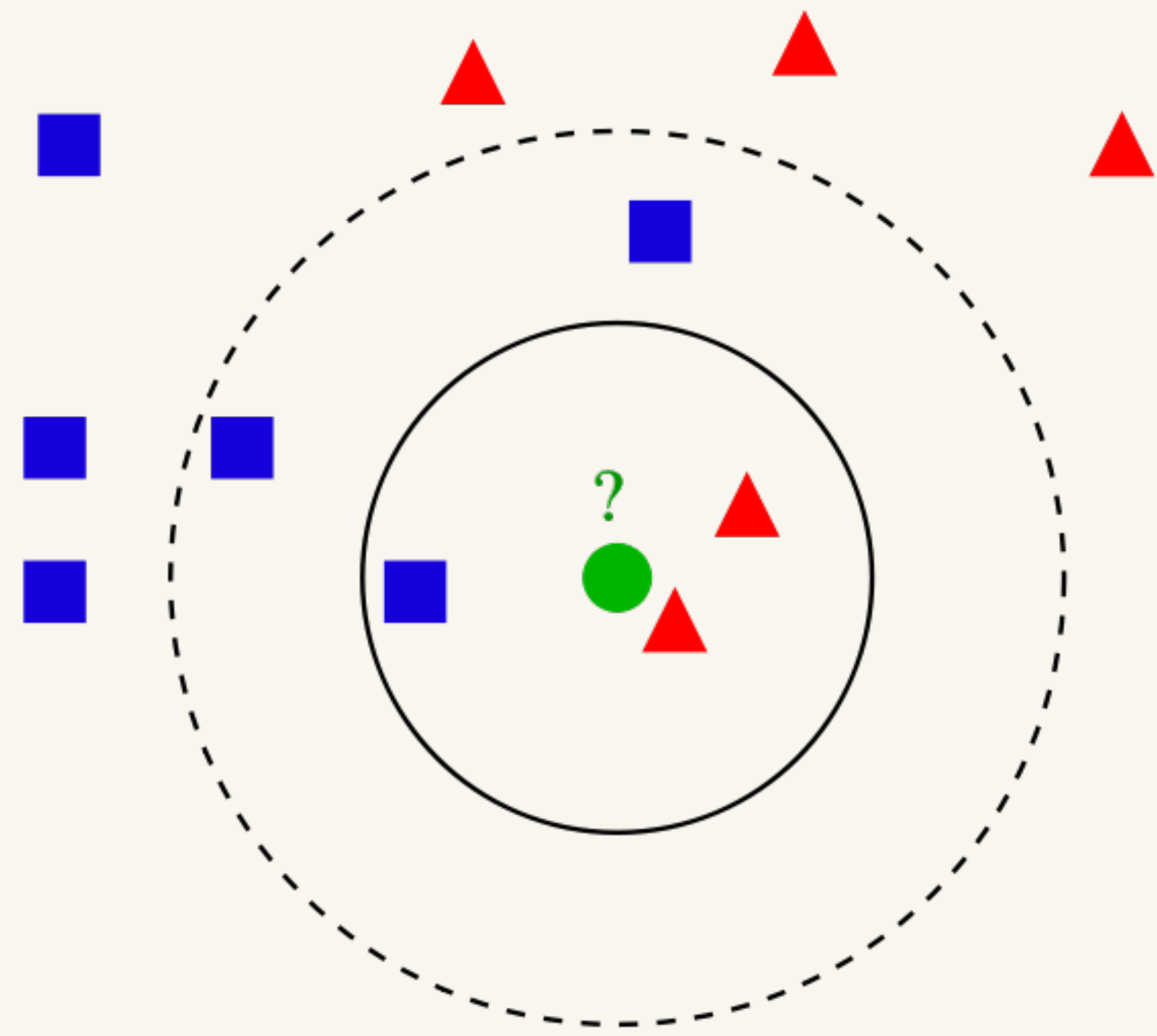


**NEAREST NEIGHBOUR**





# 3-NEAREST NEIGHBOUR

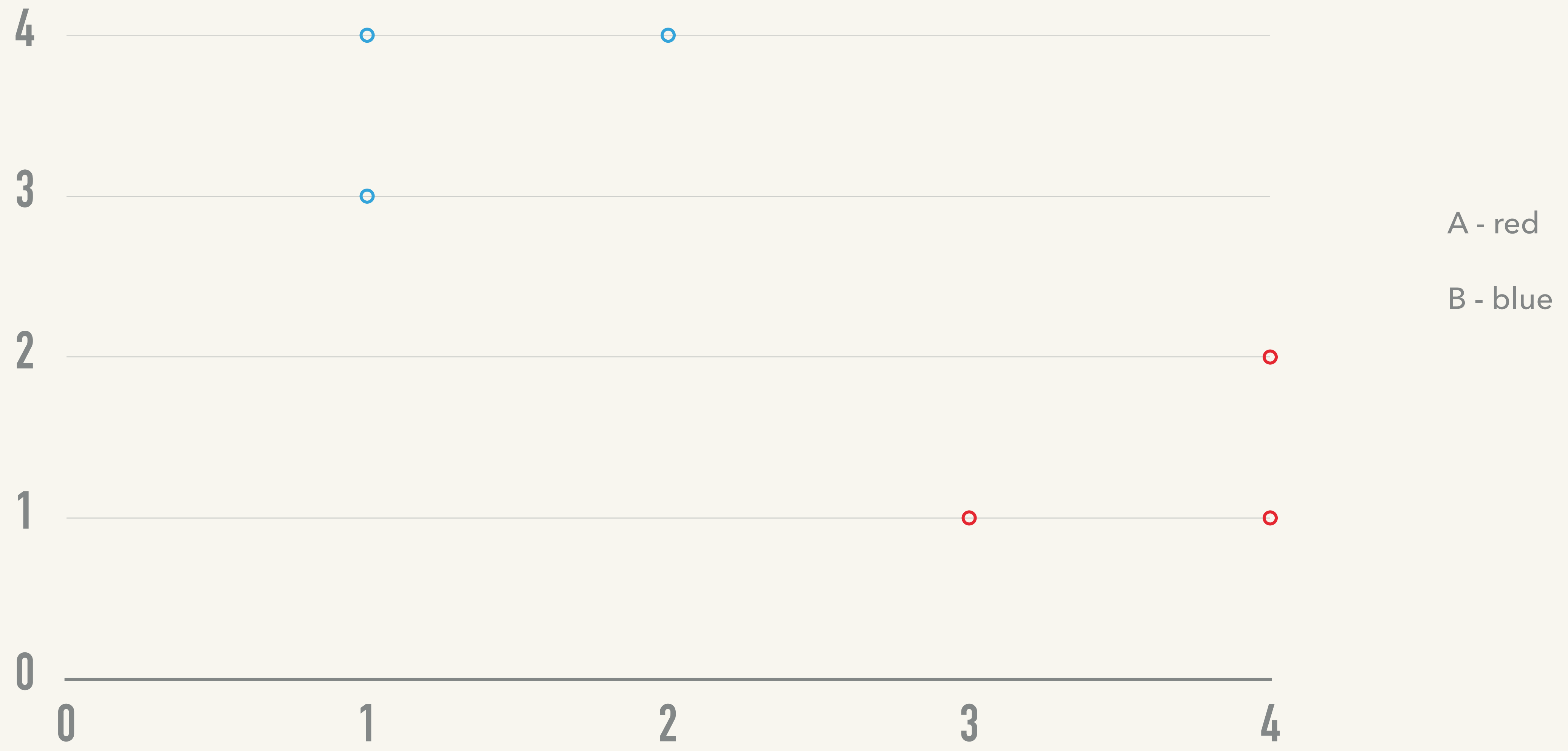


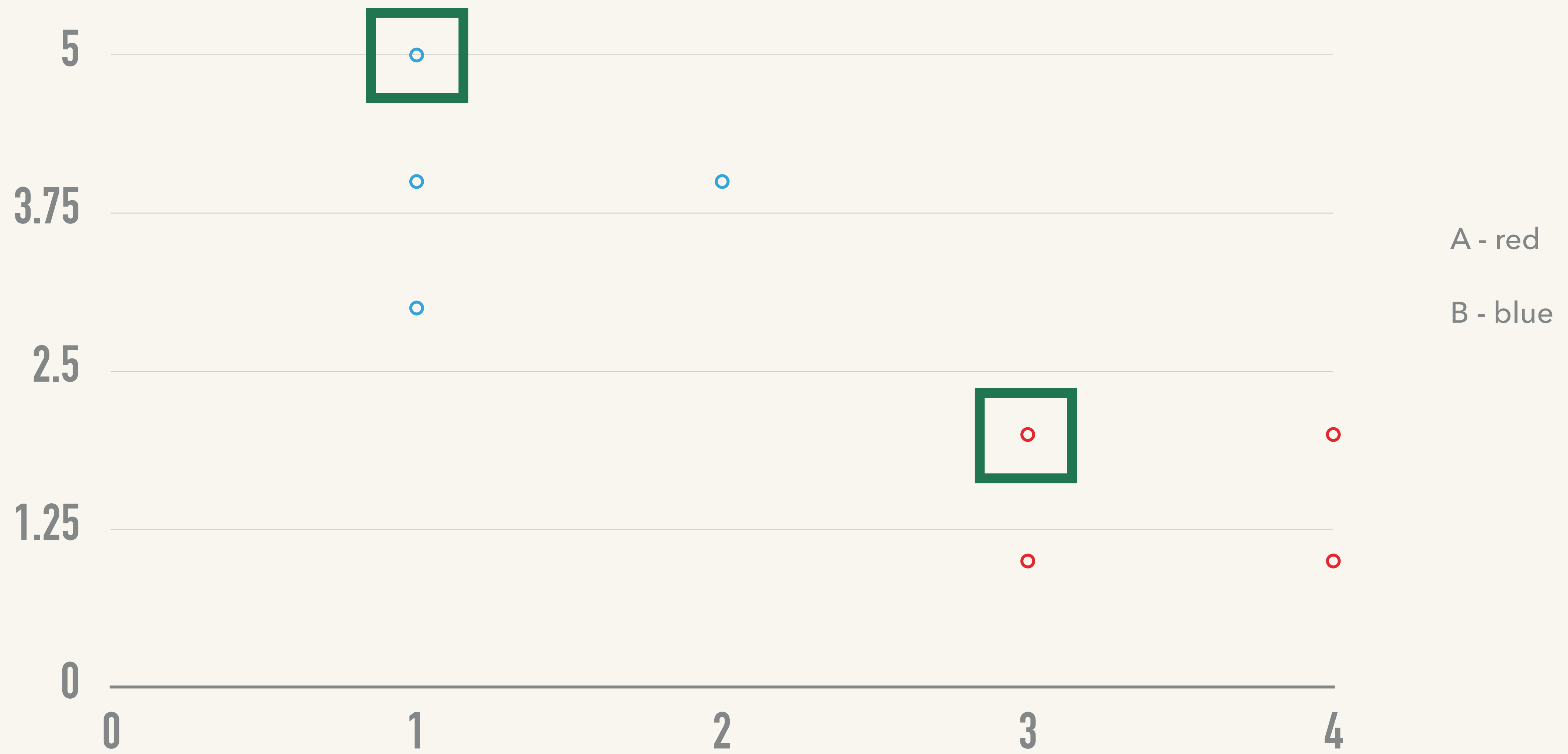
**5-NEAREST NEIGHBOUR**

# MATHS

# PHP

# DEMO





```
$samples = [[1, 3], [1, 4], [2, 4], [3, 1], [4, 1], [4, 2]];
$labels  = ['a', 'a', 'a', 'b', 'b', 'b'];

$classifier = new KNearestNeighbors();
$classifier->train($samples, $labels);

echo $classifier->predict([3, 2]);
```



# USES OF MACHINE LEARNING

# NUMERICAL ANALYSIS

# EXCEPTIONS

# E-COMMERCE

# FAULT DETECTION

# ROOT CAUSE ANALYSIS

# CLASSIFICATION

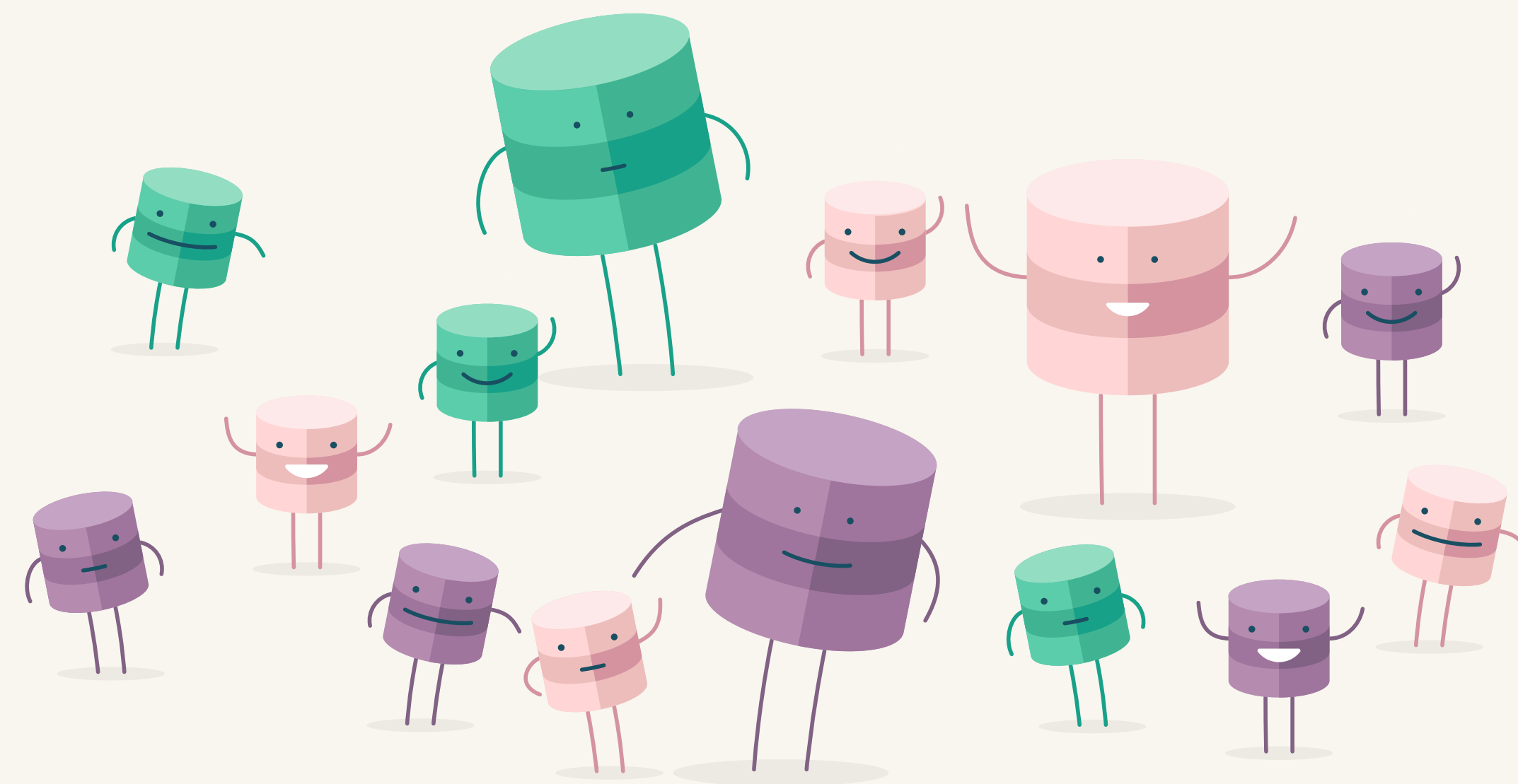
# NATURAL LANGUAGE PROCESSING



# ANALYSIS OF SUPPORT QUERIES

# ANALYSIS OF LARGE NUMBERS OF DOCUMENTS

# FUN



@MICHAELCULLUMUK

---

**THANKS - AČIŮ**

**ANY QUESTIONS?**

SLIDES

---

**BIT.LY/DEVDDAYS-ML**

DEV DAYS VILNIUS 2018

---

# TREND ANALYSIS AND MACHINE LEARNING IN PHP