

In the name of who surround me with blessing that I can not count

Al: From the pages of *science*, to the future of our **lives**.

by **Ahmadreza Anaami**

# Goals

In this lab you will:

- have a brief introduction on *AI* and *ML*
  - basic concept of machine learning
- main algorithm
  - supervised
    - regression
    - classification
  - unsupervised
- Learn to implement linear regression



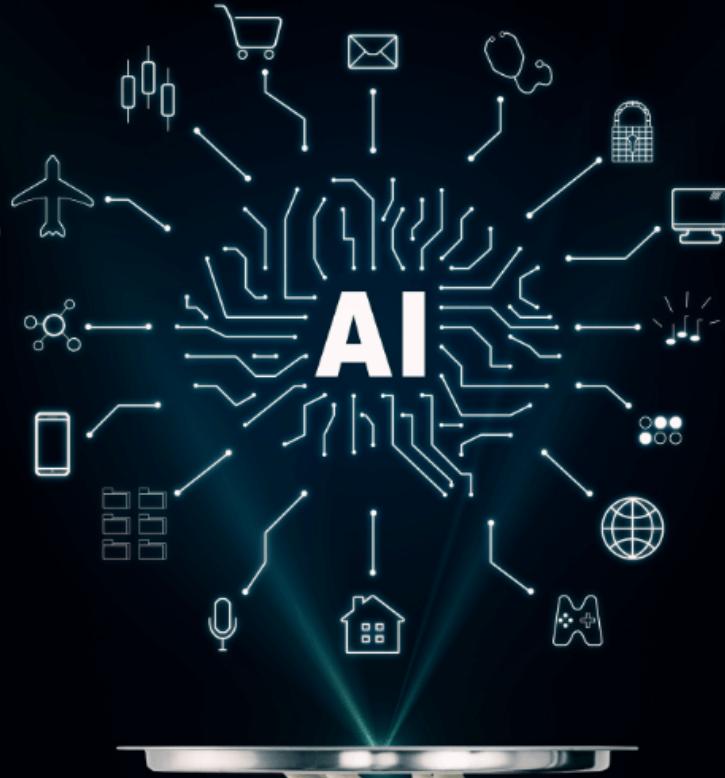
# What Is AI or Artificial Intelligence?

Artificial intelligence or AI is the branch of computer science that studies machine intelligence.

## EXAMPLES OF APPLICATIONS

- Search engines (Google)
- Content recommendations (Netflix, YouTube)
- Self-driving vehicles
- Automatic language translation
- Facial recognition
- Computer games
- Spam filters

An AI is a computer system that performs tasks that usually require human intelligence.

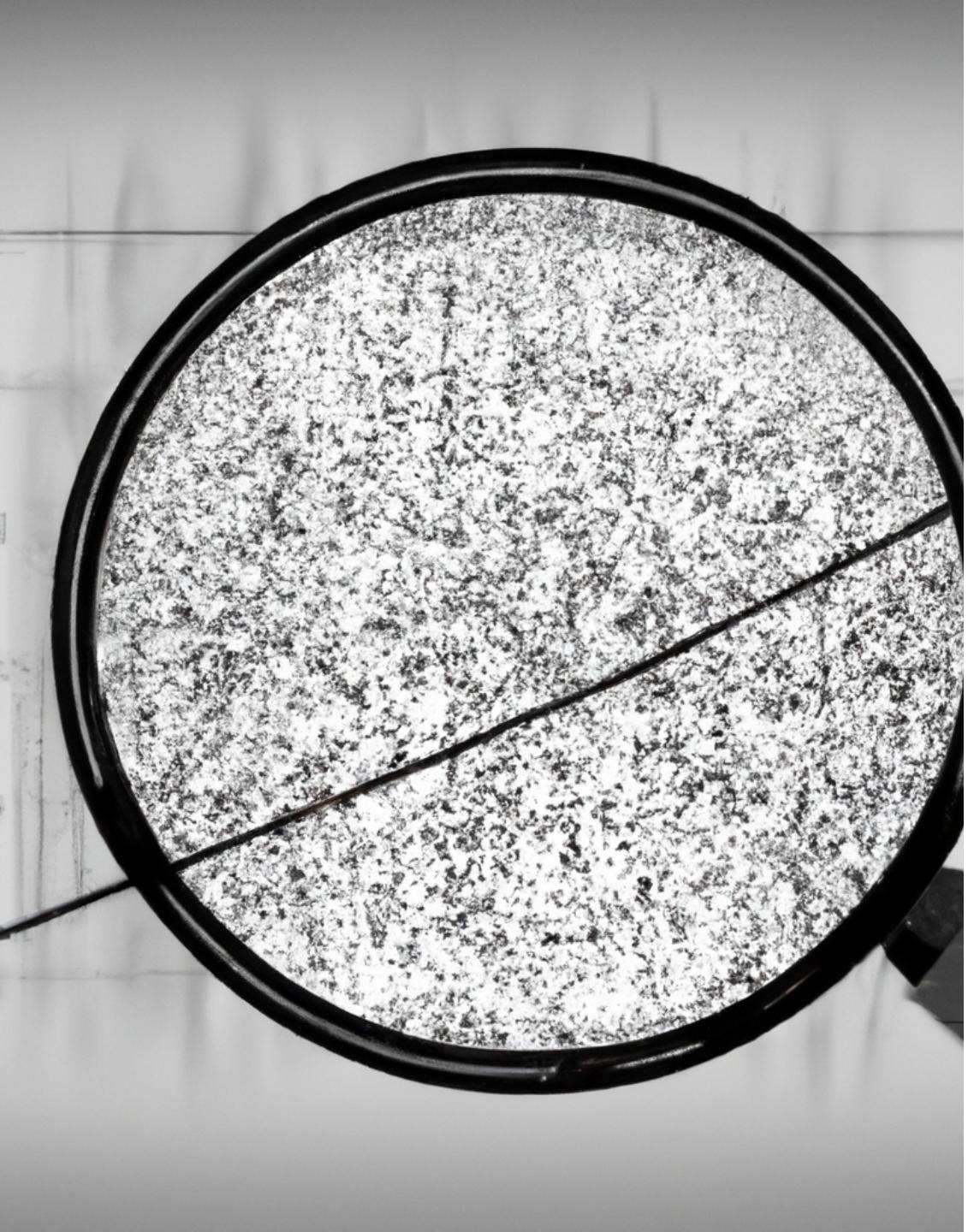


# machine learning

Subfield of AI in order to make intelligent machine

Field of study that gives the computers the ability to learn without explicitly programmed

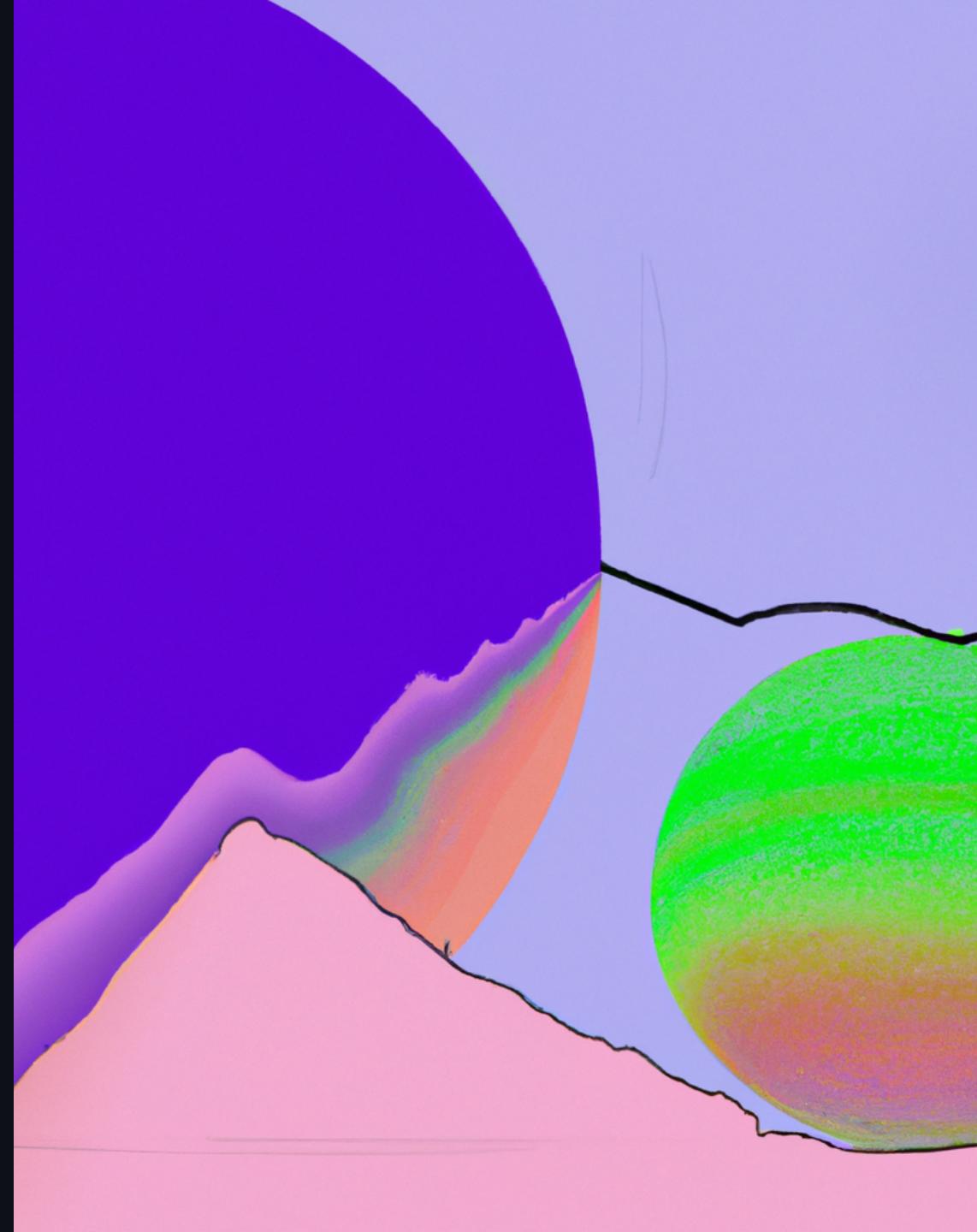
- Arthur samuel 1959



## Supervised learning

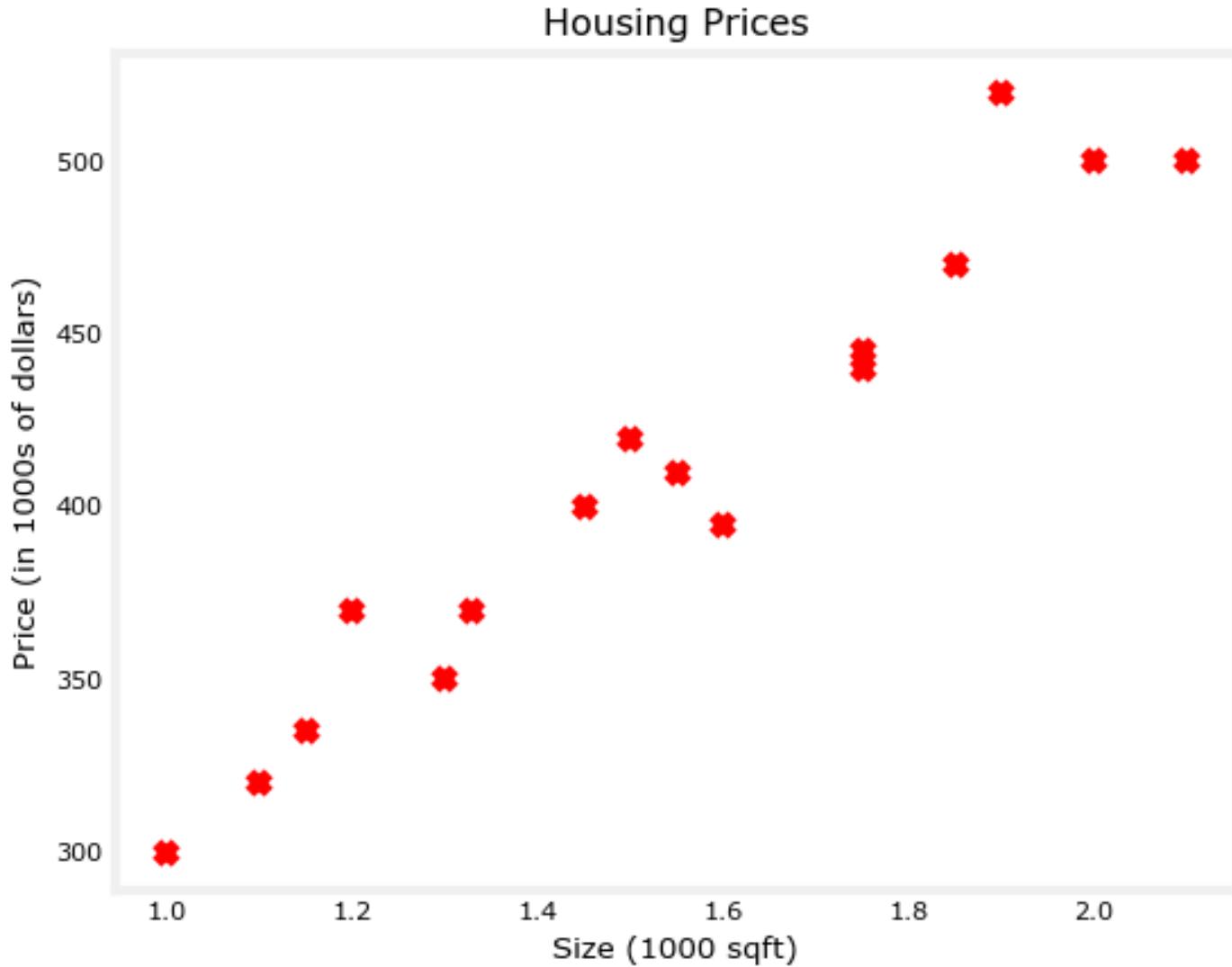
- learn from being given right answers
  - correct pairs of input ( $x$ ) and output ( $y$ )

<i>input</i>	<i>output</i>	<i>application</i>
email	spam?(0,1)	
audio	text transcript	
English	Spanish	
ad-User	click?(0,1)	
image-radar	position of other car	
image of phone	defect?(0,1)	



# Regression

predict a `number` from  
infinitely many possible  
outputs



# Classification

many inputs

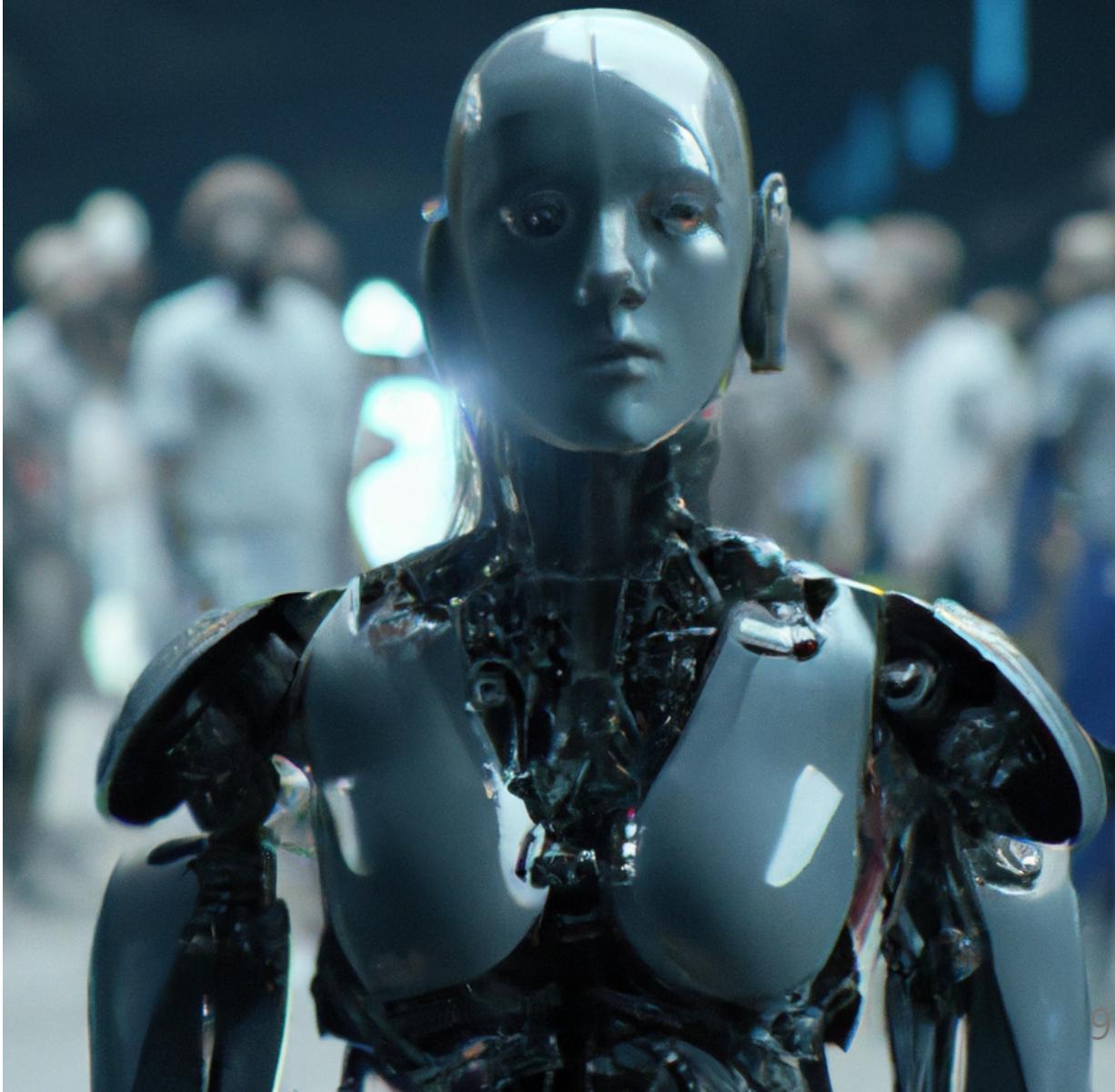
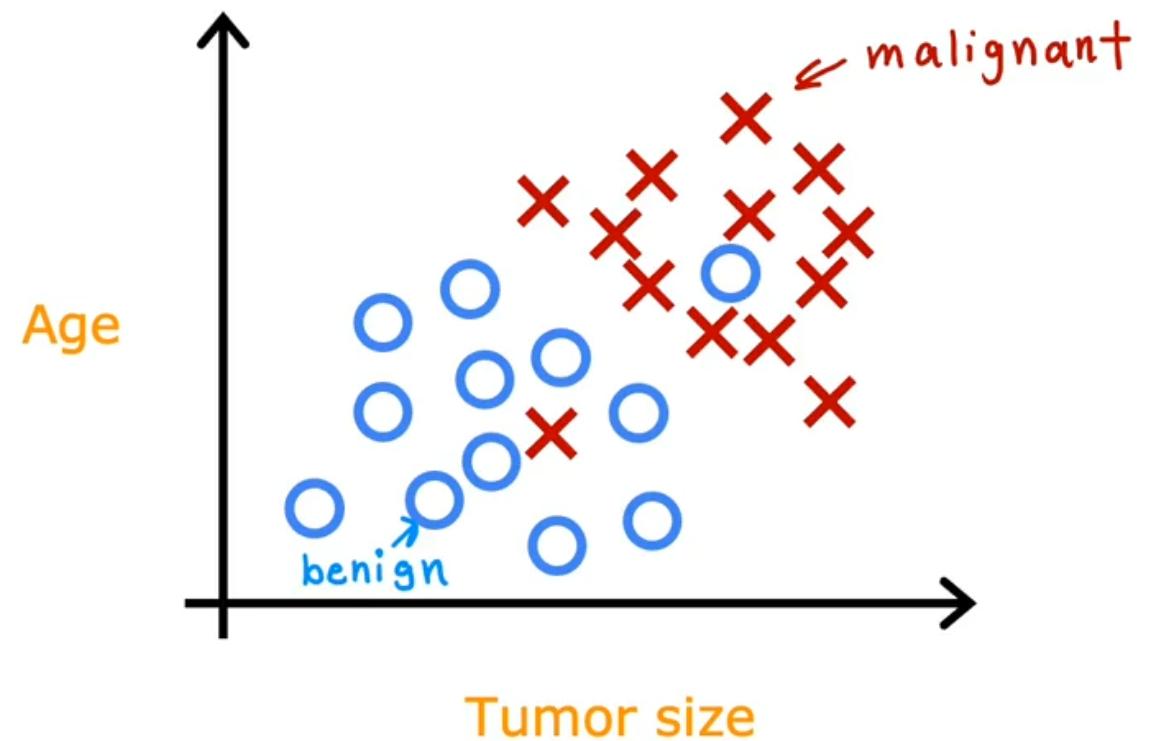
some outputs called

categories

<i>Size</i>	<i>diagnosis</i>
6	1
8	1
2	0
5	0
1	0



## Two or more inputs



# Recap



In every living man a child is hidden that wants to play

# Supervised learning

Learns from being given “right answers”

Regression

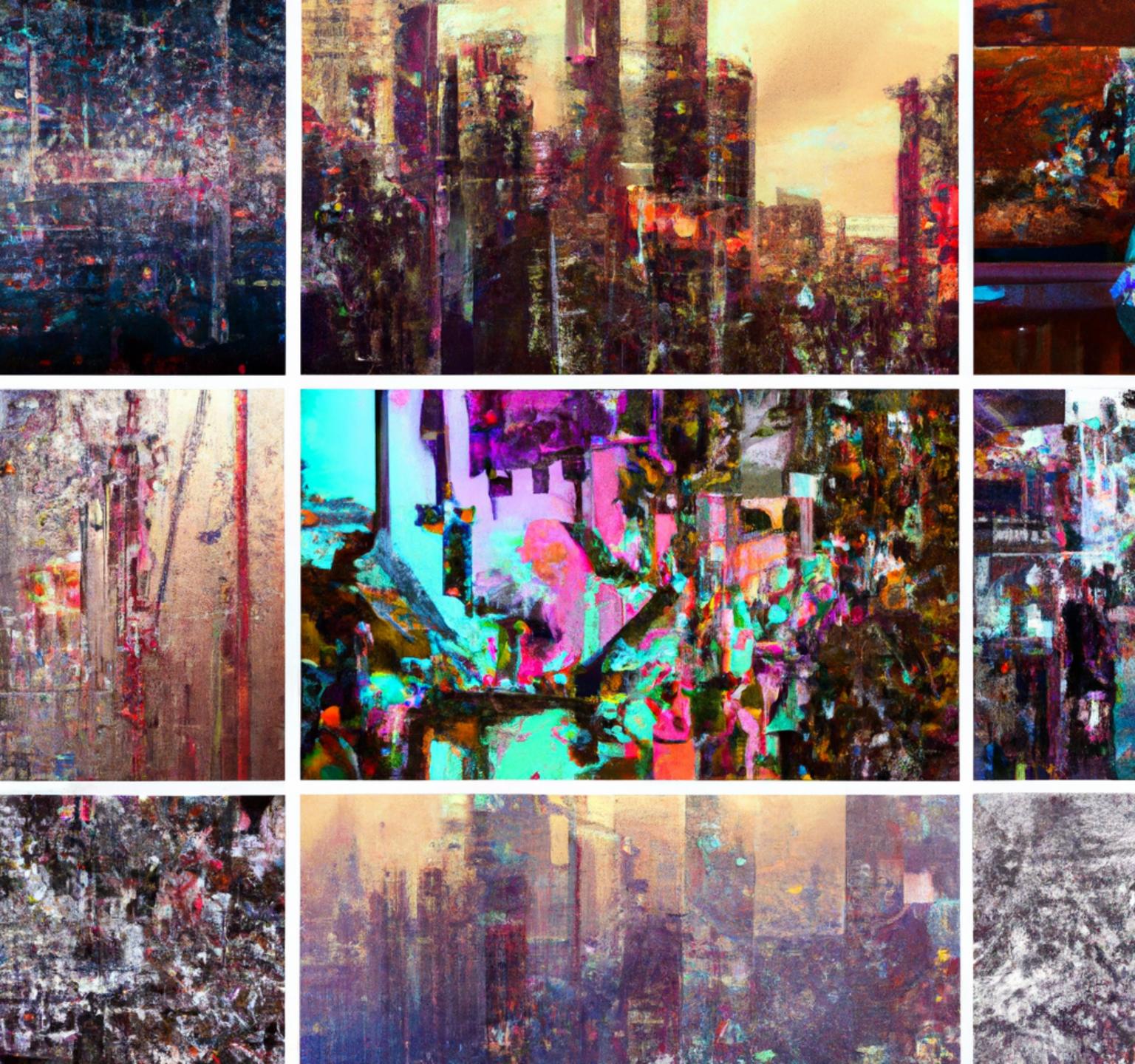
Predict a number

infinitely many possible outputs

Classification

predict categories

small number of possible outputs

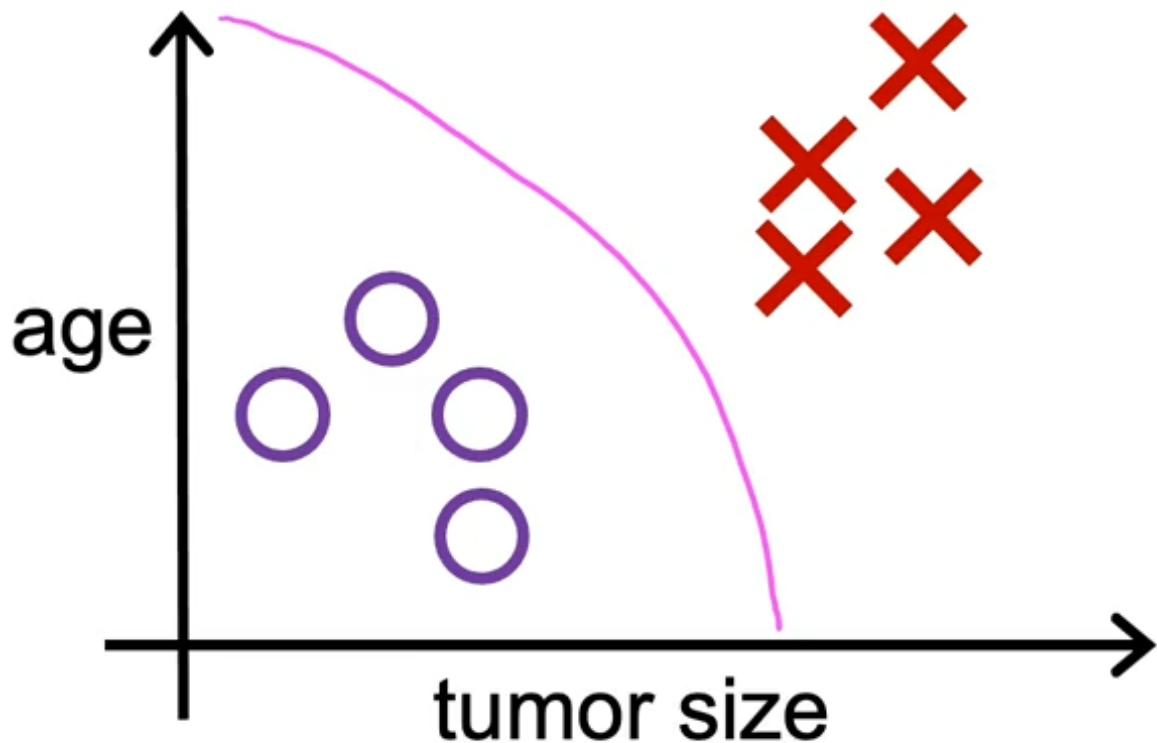


# Unsupervised learning

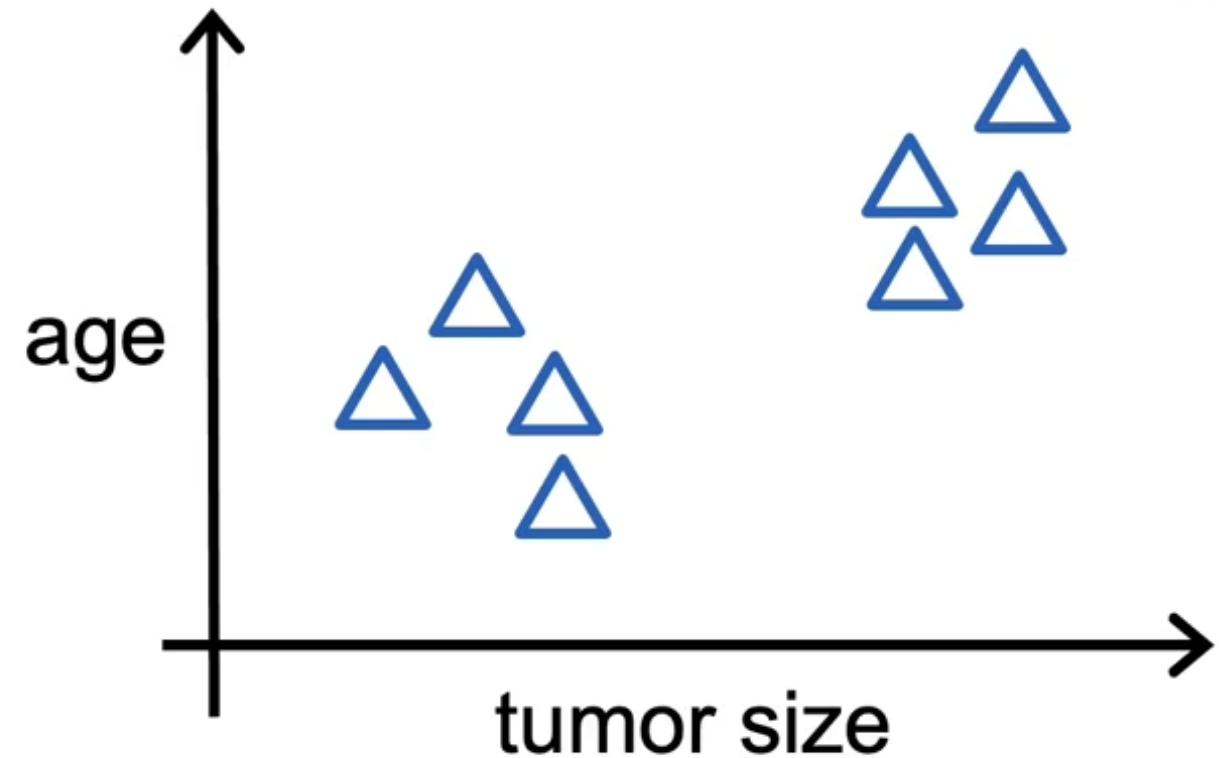
Find something interesting in (unlabeled) Data

- Clustering
- anomaly detection
- Dimensionality reduction

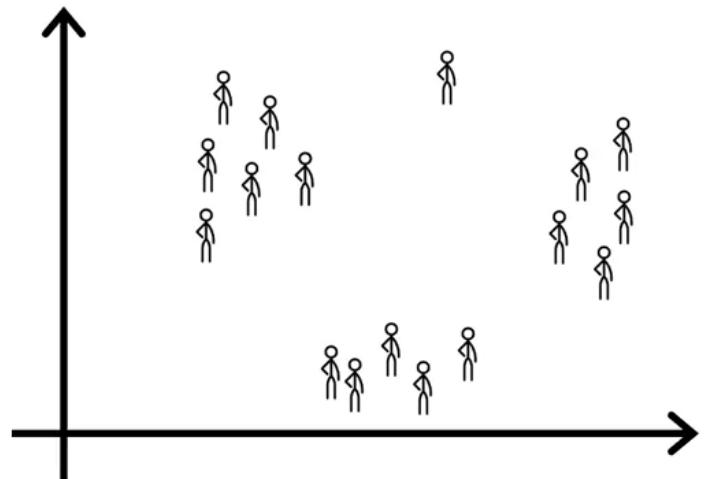
**Supervised learning**  
Learn from data **labeled**  
with the “**right answers**”



**Unsupervised learning**  
Find something interesting  
in **unlabeled** data.



## Clustering: Grouping customers



# Question

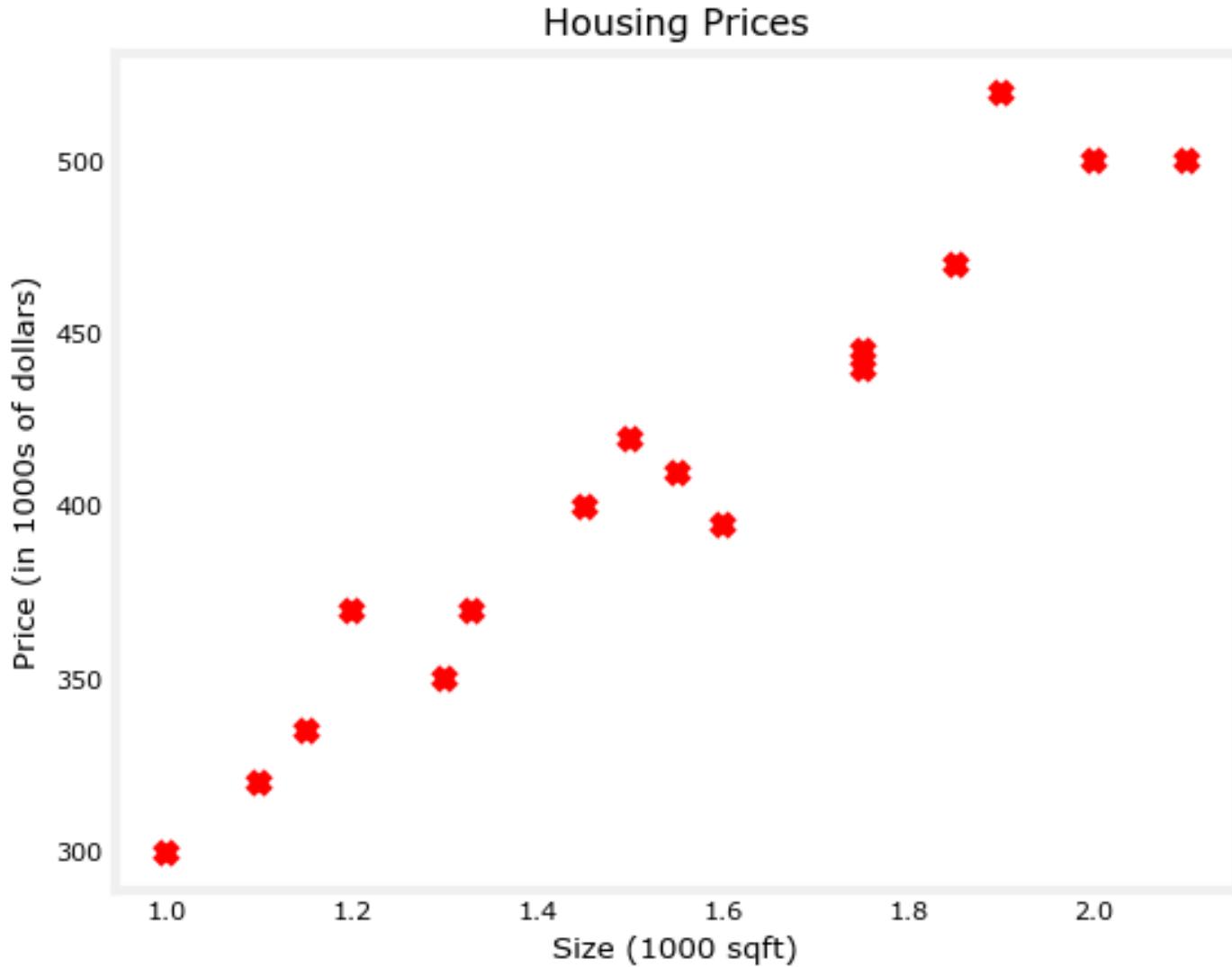
Of the following examples, which would you address using an **unsupervised** learning algorithm?

- Given email labeled as spam/not spam, learn a spam filter.
- Given a set of news articles found on the web, group them into sets of articles about the same story.
- Given a database of customer data, automatically discover market segments and group customers into different market segments.
- Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not

## linear regression

$$f_{w,b}(x) = wx + b$$

w , b > parameters , weight  
x > single feature



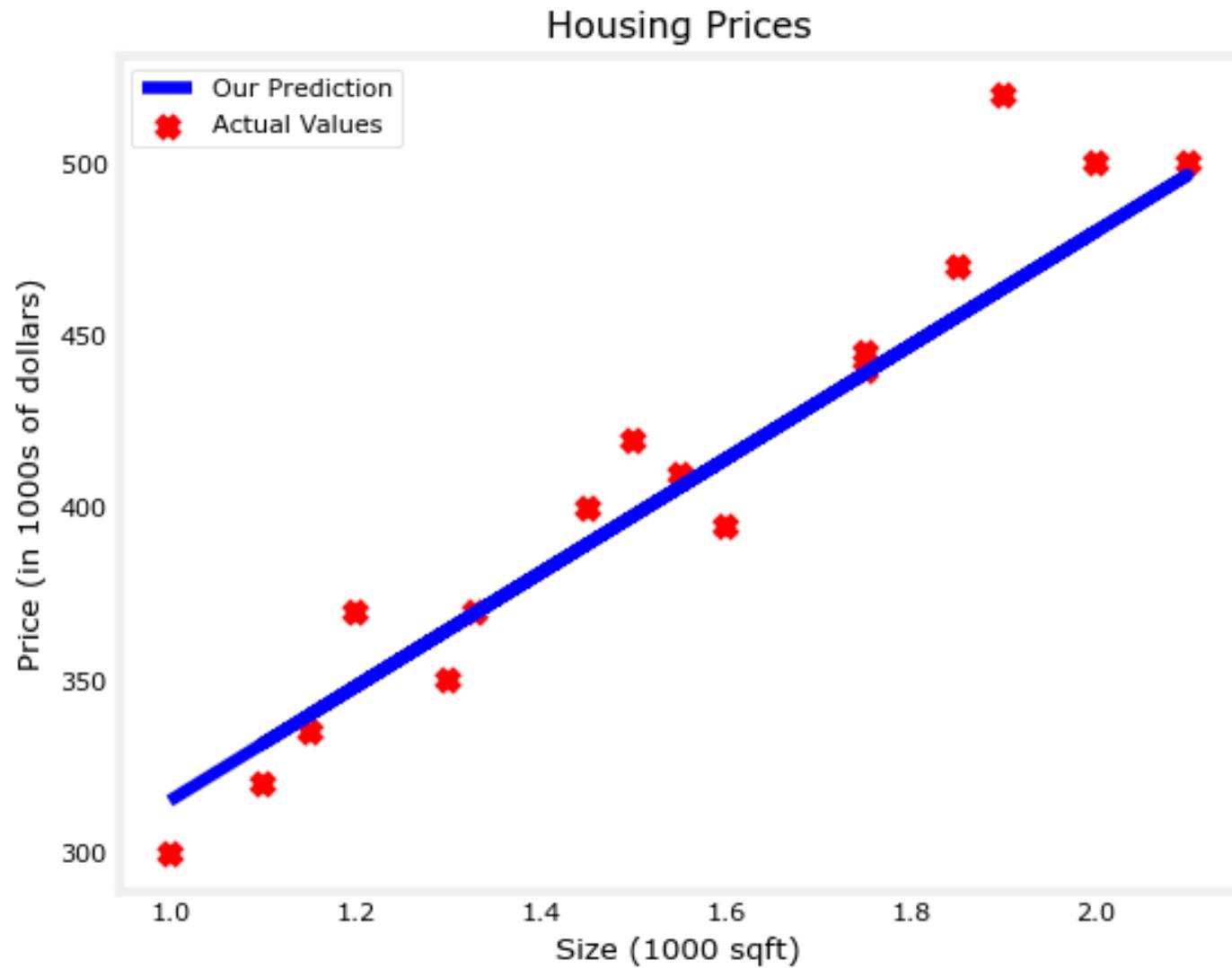
# Cost function

$$J(w, b) = \frac{1}{2m} \sum_{i=0}^{m-1} (f_{w,b}(x^{(i)}) - y^{(i)})^2$$

where

$$y\text{Hat} = f_{w,b}(x^{(i)}) = wx^{(i)} + b$$

m = number of example

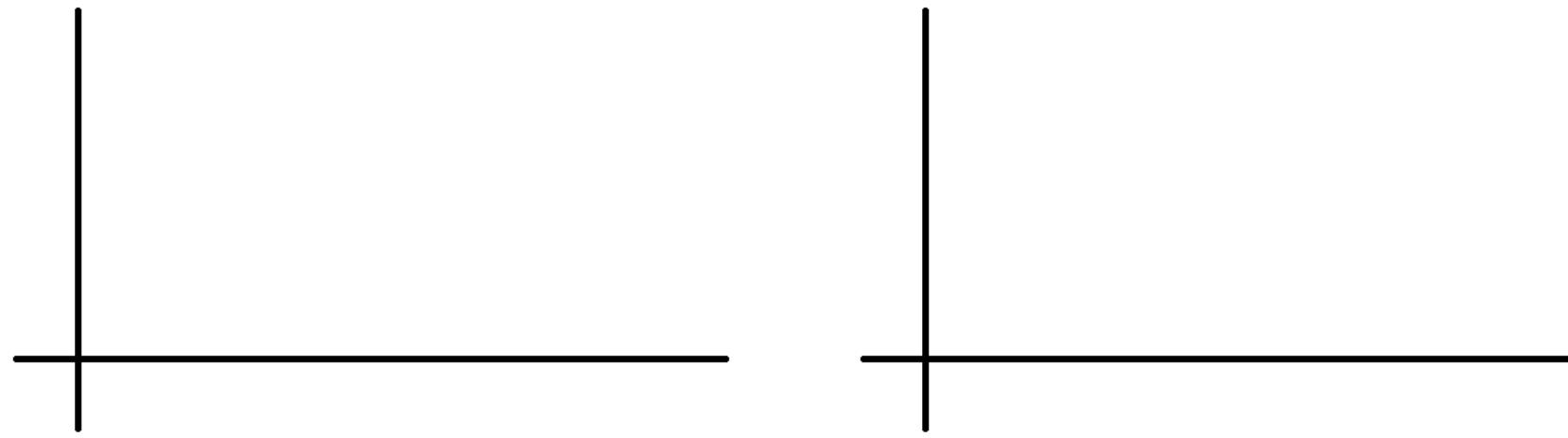


# Example

choose w to  
minimize J(w)

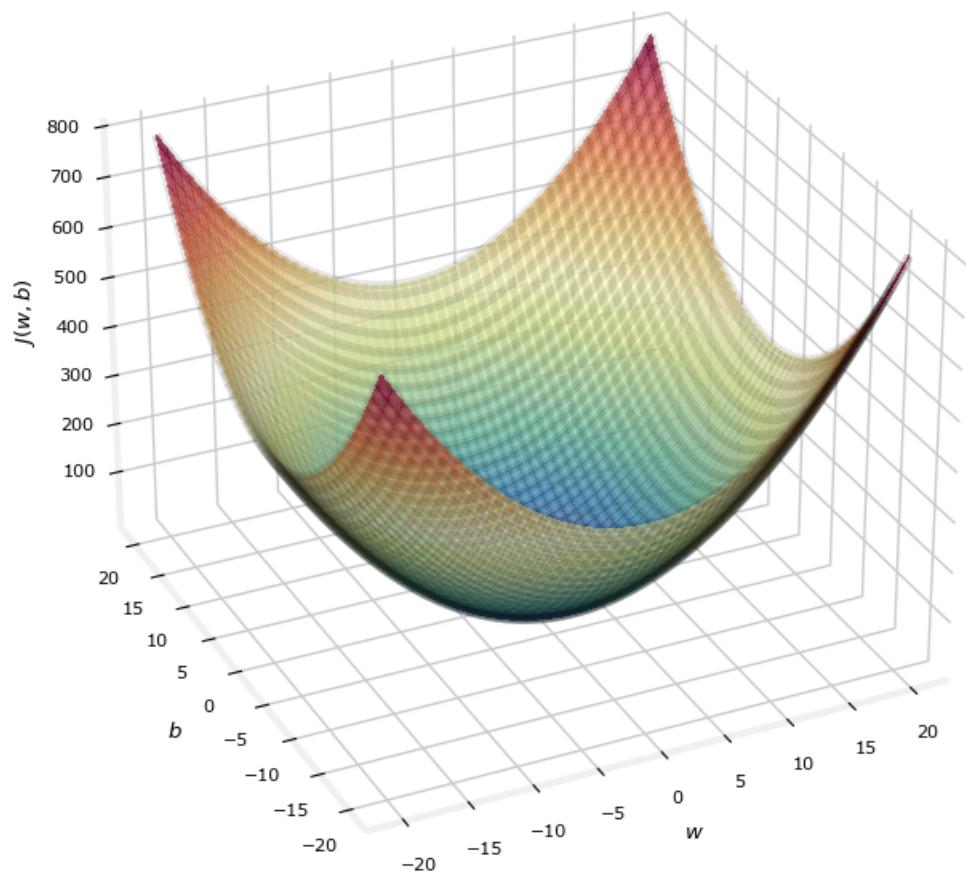
$$y\text{Hat} = f_w(x^{(i)}) = wx^{(i)}$$

$$J(w) = \frac{1}{2m} \sum_{i=0}^{m-1} (f_w(x^{(i)}) - y^{(i)})^2$$



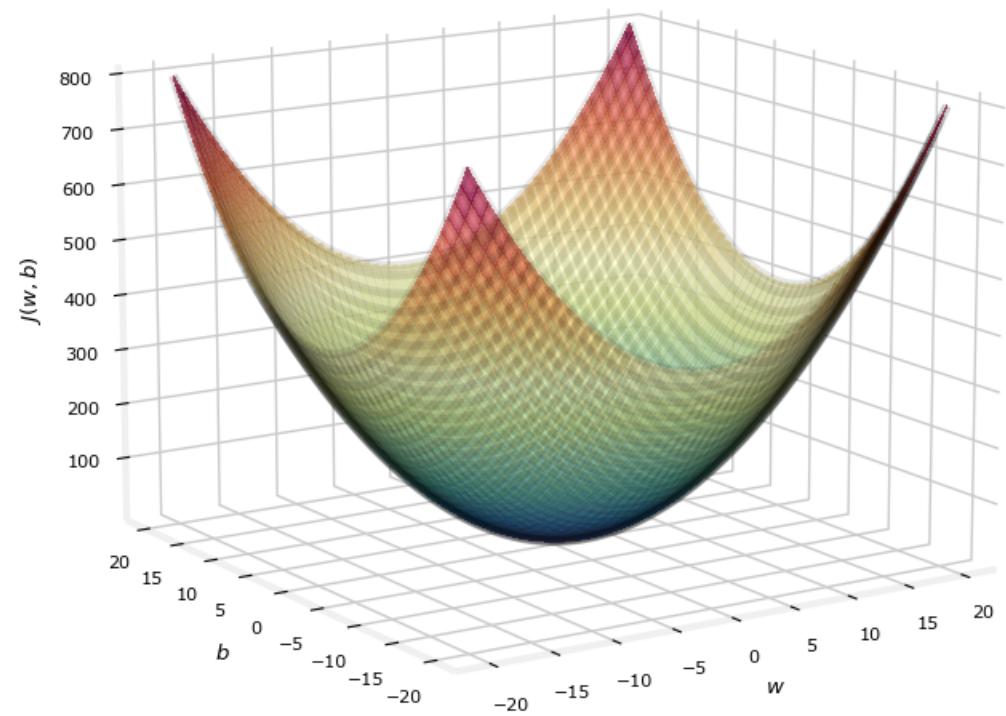
$J(w, b)$

[You can rotate this figure]

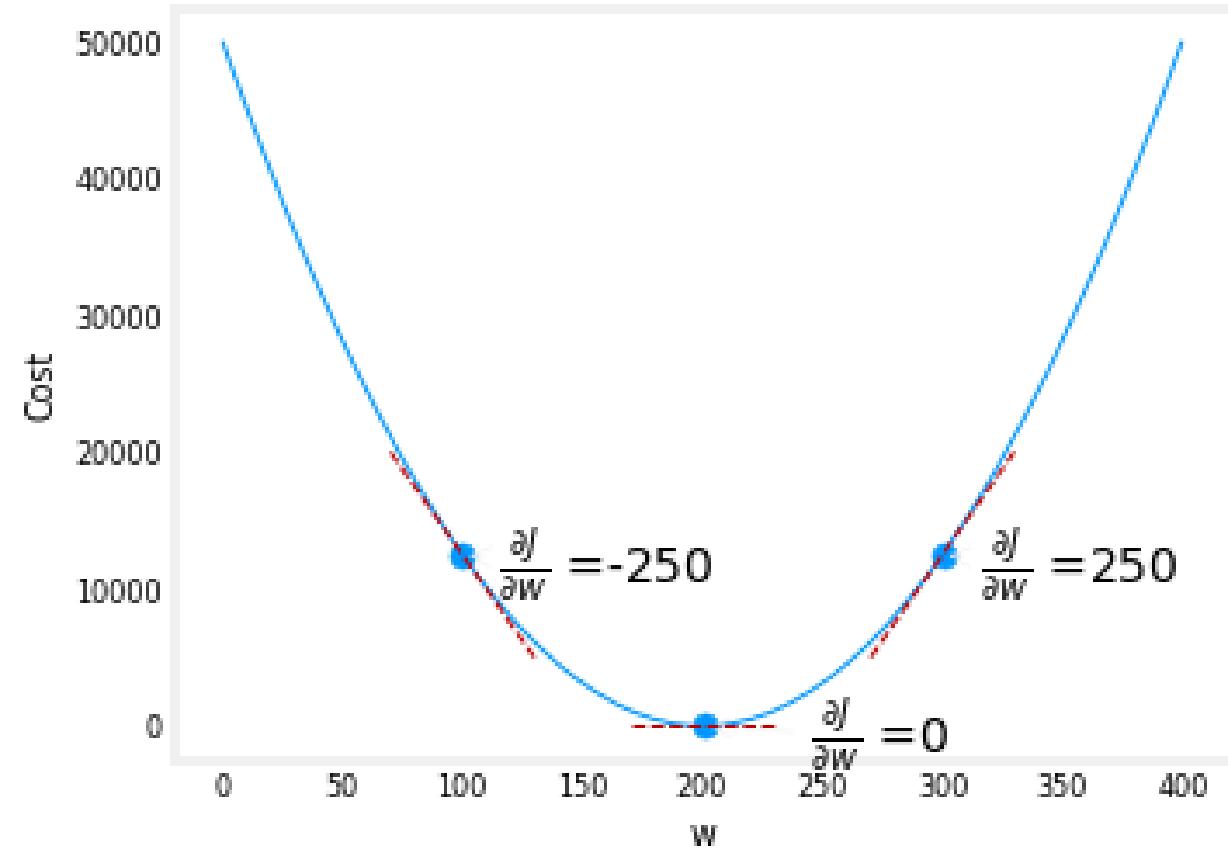


$J(w, b)$

[You can rotate this figure]



Cost vs w, with gradient; b set to 100



## Gradient descent

a more systematic way to minimize  $j(w,b)$

repeat until convergence: {

$$w = w - \alpha \frac{\partial}{\partial w} J(w, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

}

## Derivative of the Cost func

$$J(w, b) = \frac{1}{2m} \sum_{i=0}^{m-1} (f_{w,b}(x^{(i)}) - y^{(i)})^2$$

$$\frac{\partial}{\partial w} J(w, b) = \frac{1}{m} \sum_{i=0}^{m-1} (f_{w,b}(x^{(i)}) - y^{(i)}) x^{(i)}$$

$$\frac{\partial}{\partial b} J(w, b) = \frac{1}{m} \sum_{i=0}^{m-1} (f_{w,b}(x^{(i)}) - y^{(i)})$$

# optional

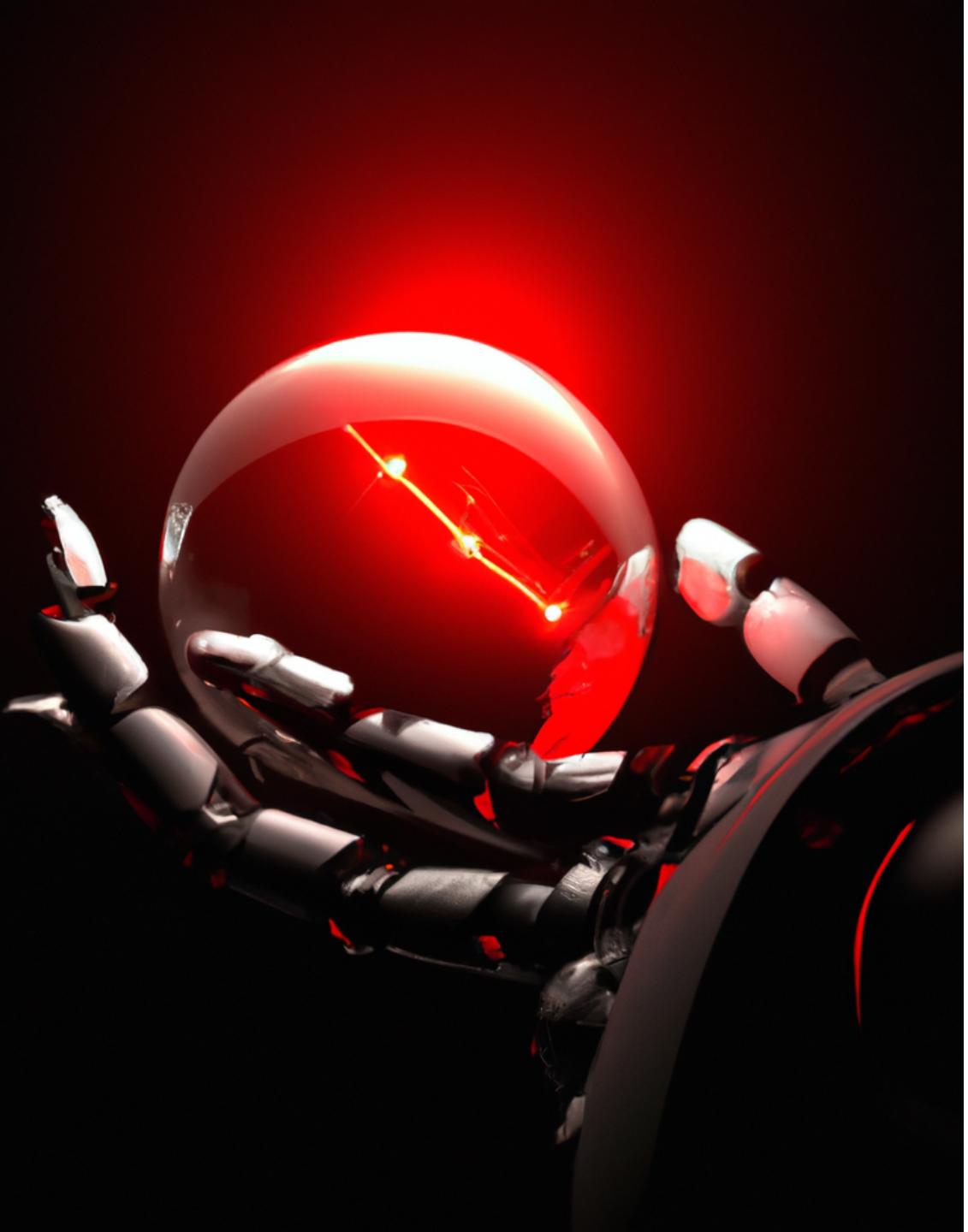
if **alpha** is too small :

if **alpha** is too big :

$$\alpha \frac{\partial}{\partial w} J(w, b)$$

calculate **Derivative** : ☺





## Lets get our hands dirty

- step 1
  - read our data
- step 2
  - write our functions
    - calculate  $\hat{y}$
    - calculate cost
    - calculate gradient
    - implement gradient descent
- step 3
  - congratulations