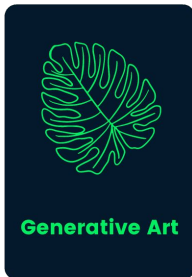
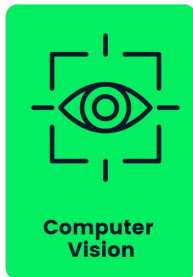
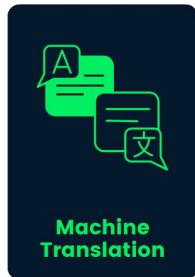
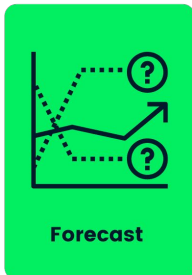
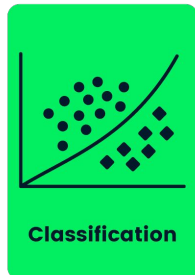
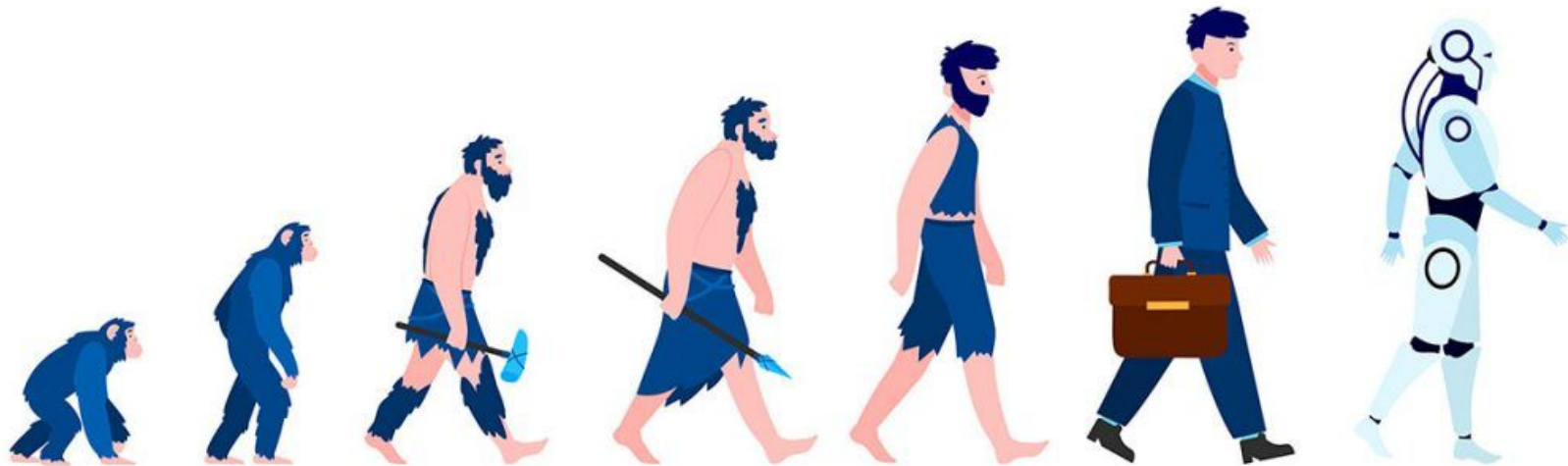


# Introducción al Machine Learning



# Historia del Machine Learning

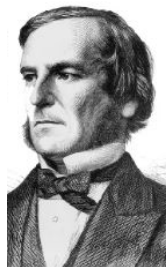


# Línea del tiempo del Machine Learning

1764

Se publica el  
teorema de Bayes

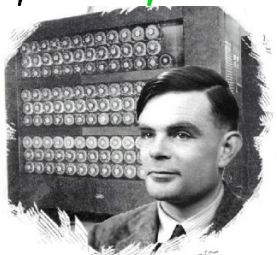
$$P[A_n/B] = \frac{P[B/A_n] \cdot P[A_n]}{\sum P[B/A_i] \cdot P[A_i]}$$



George Boole inventa la  
lógica **Booleana**

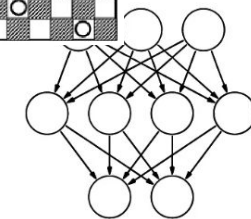
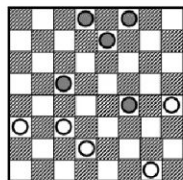
1847

Alan Turing propone  
una **máquina** que  
pueda **aprender**



1936

En **IBM** se crean los **primeros**  
**programas** de computadora /  
Se crea la **primer red**  
**neuronal** "Adaline"



1952 / 1959

Joffrey Hinton inventa el  
término "**Deep Learning**"



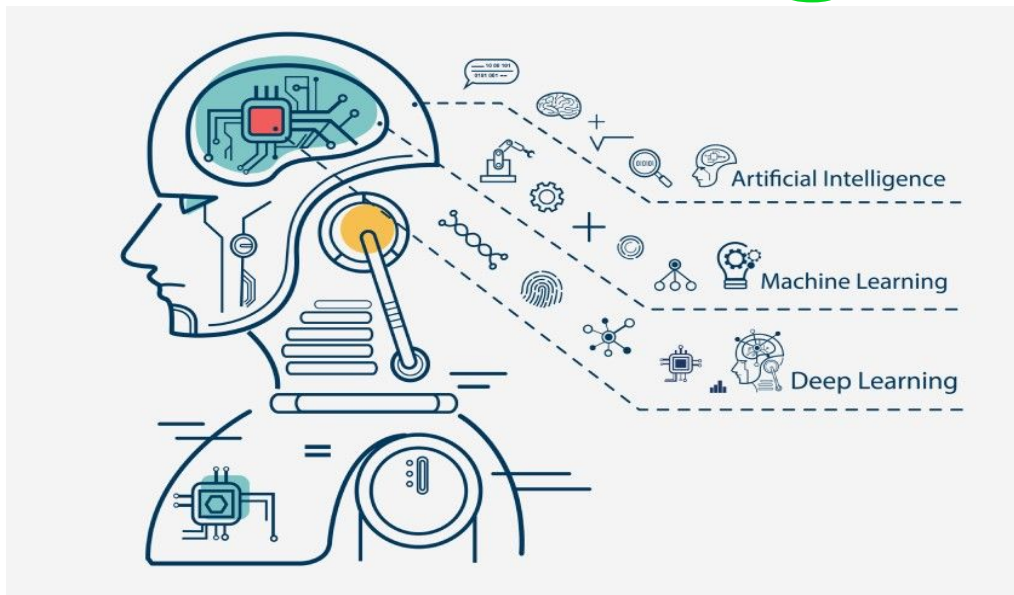
2006

2022

Sale al público  
**ChatGPT**



# Definición de Machine Learning



# ¿Cómo funciona el Aprendizaje Automático?

Es un **paradigma de programación** que “calcula” los parámetros de un modelo para obtener las salidas deseadas según los ejemplos de entrada.

## Programación Tradicional:



Ejemplos:

$D = A + B - C$   
 $y = x1 - x2(x3)$   
 $out = \max(0, inp)$

## Aprendizaje Automático:

Ejemplos:

E1:  $[0 \ 1 \ 1 \ 0] \rightarrow$  "Gato"

E2:  $[1 \ 0 \ 0 \ 1] \rightarrow$  "Perro"

Resultado: Modelo



# El Aprendizaje automático (ML) dentro de la IA

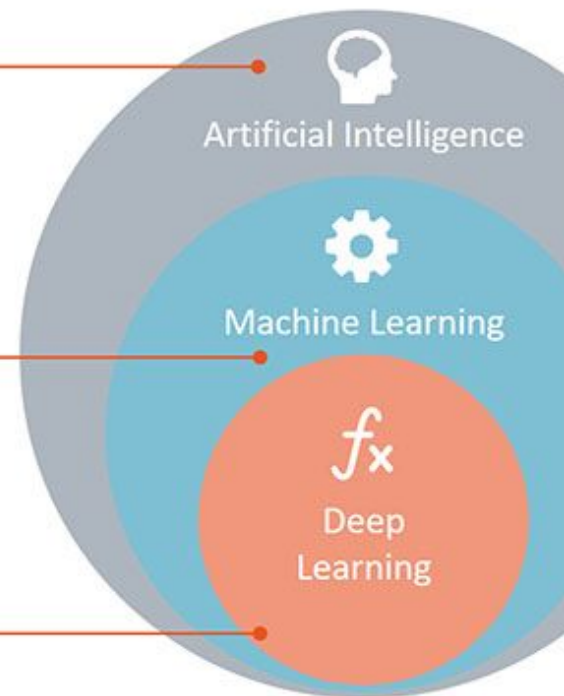
## Algoritmos que **imitan el razonamiento humano**

(Sistemas basados en reglas, Sistemas expertos, Lógica proposicional, Lógica difusa, Métodos de búsqueda)

## Algoritmos estadísticos que **aprenden de ejemplos**

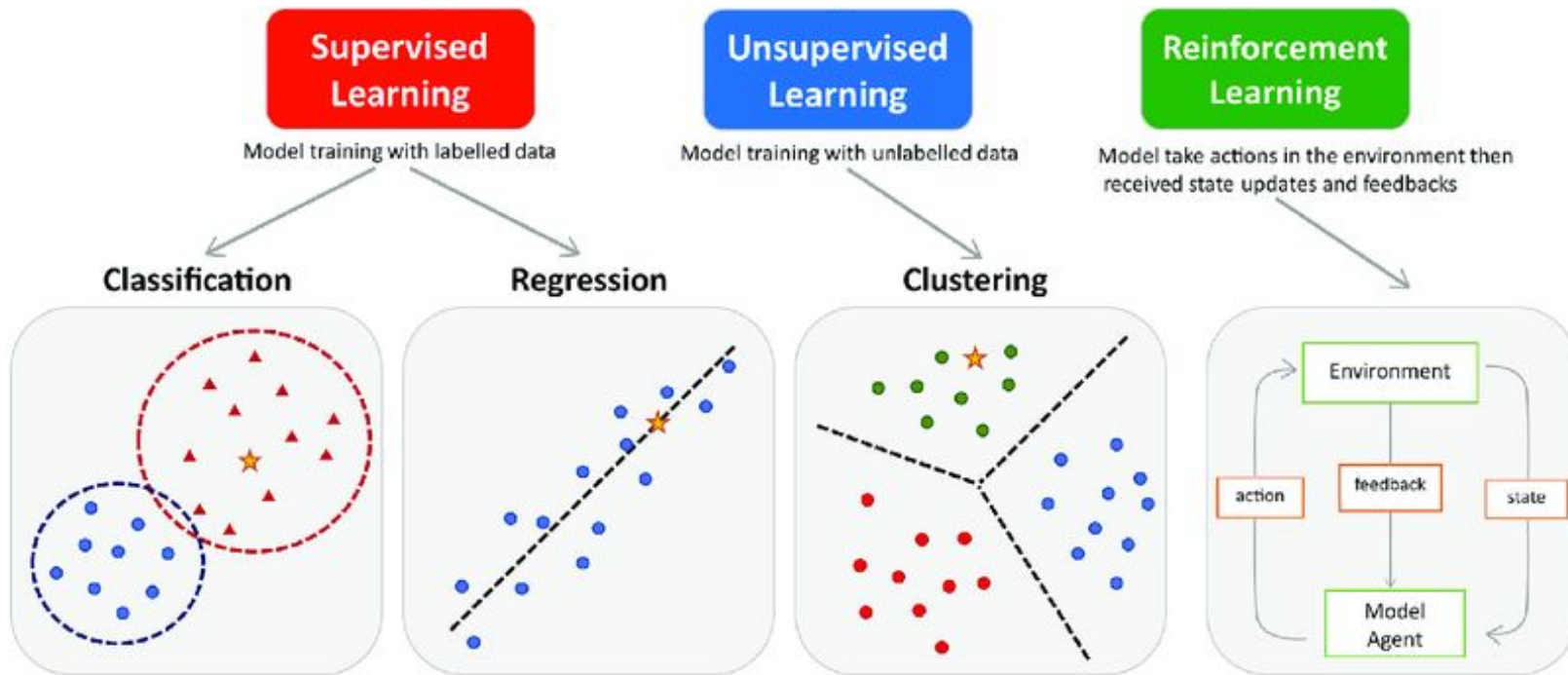
(Máquinas de soporte vectorial, Regresiones lineales, Regresión logística, Árboles de decisión)

**Redes neuronales multicapa que analizan profundamente** (DNN, CNN, RNN, Transformers)



# Tipos de Machine Learning

Los algoritmos de Machine Learning se dividen *principalmente en 3 tipos* según la *naturaleza del problema* que desean atacar:



# CLASIFICACIÓN





# Te lo explico con manzanas

A ver, si eres tan “humano” ¿Cómo sabes que esto es una **manzana**?



*Yo sé que esto es una manzana porque es / tiene:*

# Características

Son aquellos **atributos** que tiene un objeto que lo hace ser ese objeto. En ML una característica es cualquier cosa que se pueda **representar numéricamente**.



Las imágenes son un conjunto de **pixeles** que se forman de valores de **colores** del 0 a 255



Los textos son un conjunto de **palabras**, las cuáles pueden tener un **valor numérico asignado** cada una



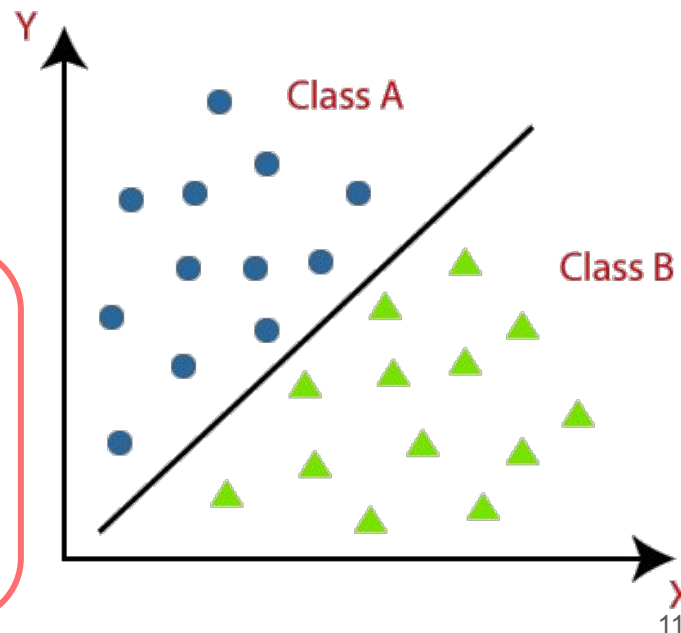
El audio es un grupo de **sonidos** que se representa por sus valores de **frecuencias en el tiempo**

# Características (Mas simples)

No todo tiene que ser pixeles, frecuencias o diccionarios de palabras, las características también pueden ser (*cualquier cosa que se pueda representar numéricamente*):

- El **precio promedio** de las casas en una zona por año
- La **cantidad de bacterias** en una muestra
- La **cantidad de alumnos** en un curso
- Un valor de **probabilidad** del 0 al 100%
- La **velocidad** de un vehículo

*¿Qué otra cosa puede ser una característica?*



# Peras y Manzanas

Imagina que quieres crear un modelo que pueda clasificar peras de manzanas.

¿Qué características podrías usar para diferenciarlas? Por ejemplo:

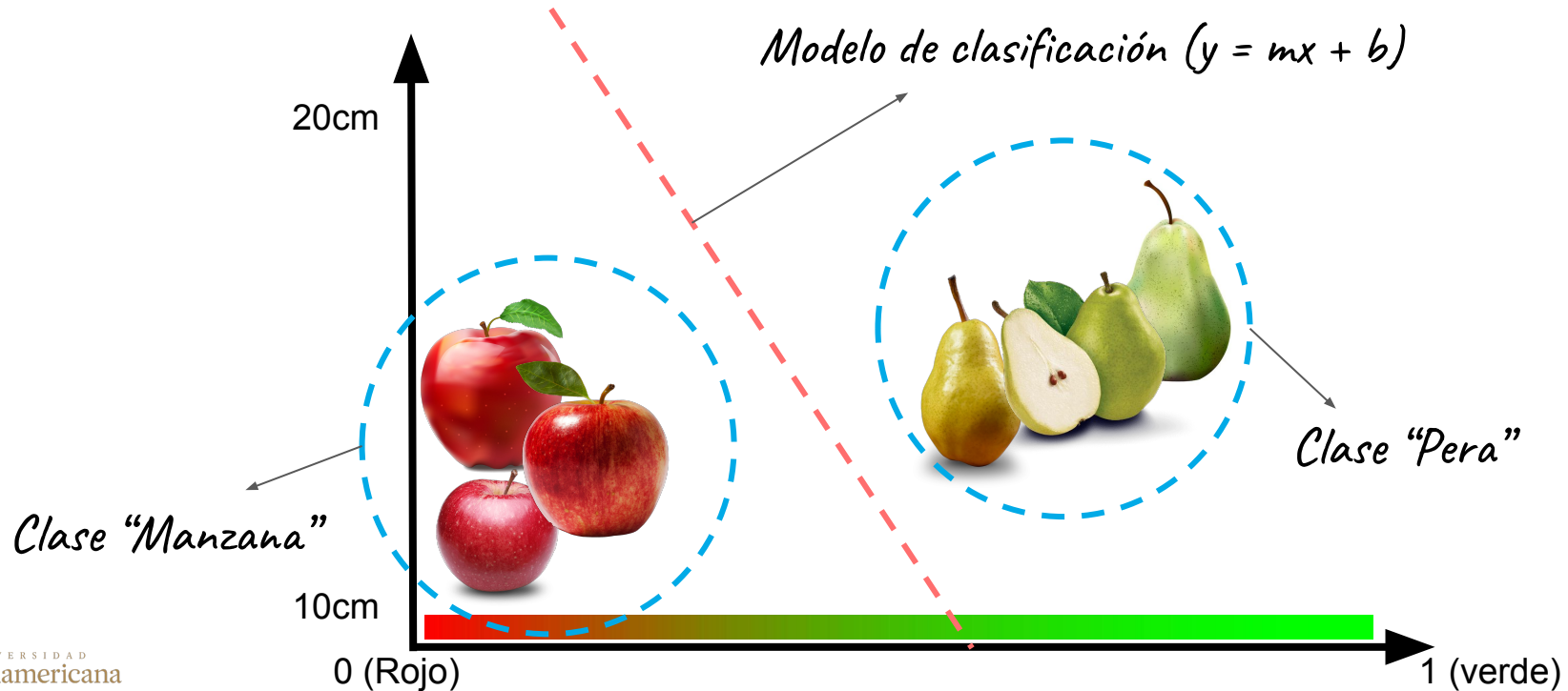
- **Color** [0 representa rojo y 1 representa verde]
- **Altura** [10cm - 20cm]

A cada uno de los **objetos**  
(en este caso peras o  
manzanas), los conocemos  
como **EJEMPLOS**



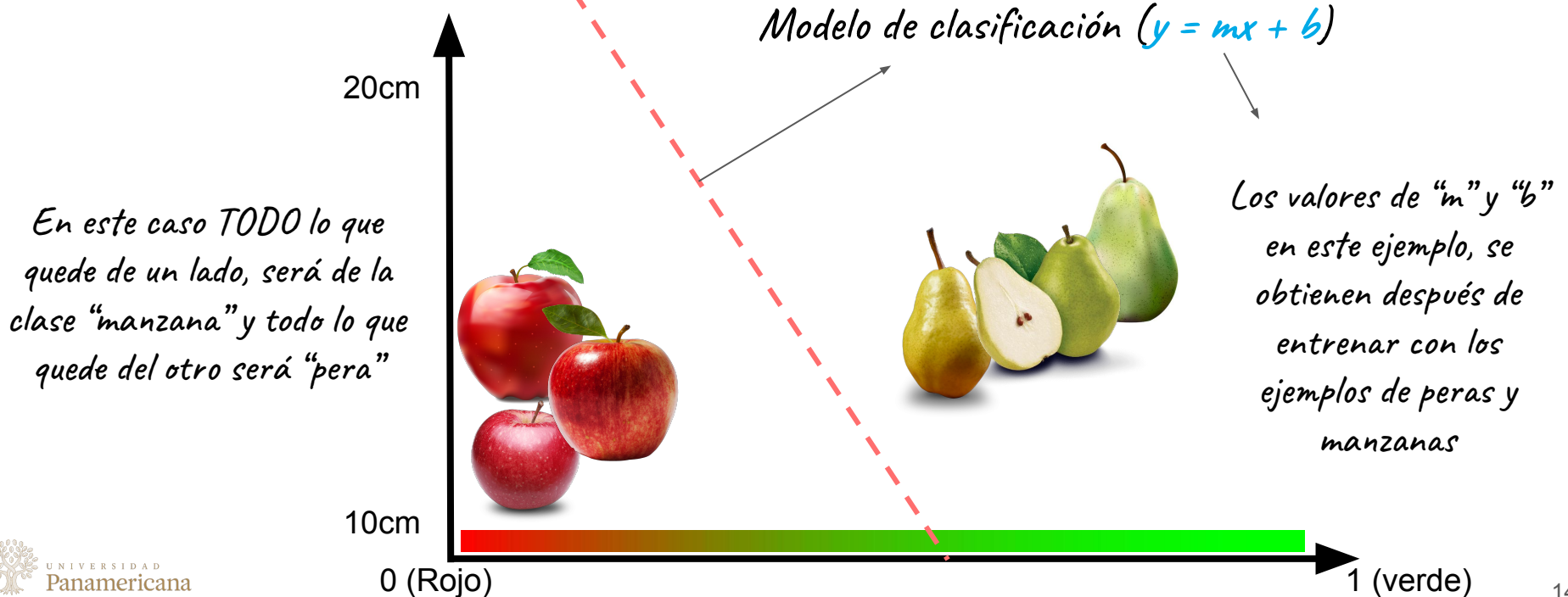
# Peras y Manzanas

El objetivo de un **algoritmo de clasificación** es crear un **modelo** que pueda separar cada **clase** de objeto:



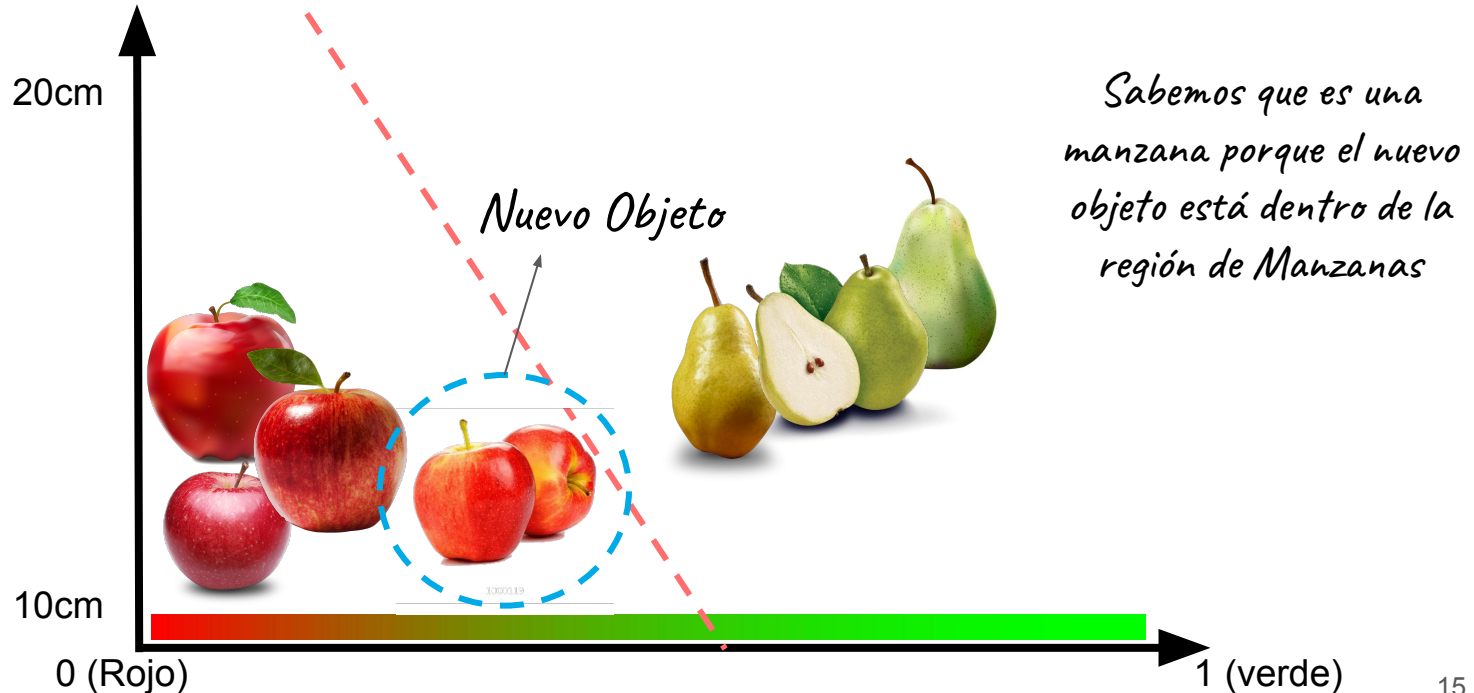
# Peras y Manzanas

Se llama **aprendizaje automático** (Machine Learning) porque a partir de los **ejemplos**, el algoritmo calcula los mejores **parámetros** para el **modelo**. A esto se le conoce como **entrenamiento**:



# Peras y Manzanas

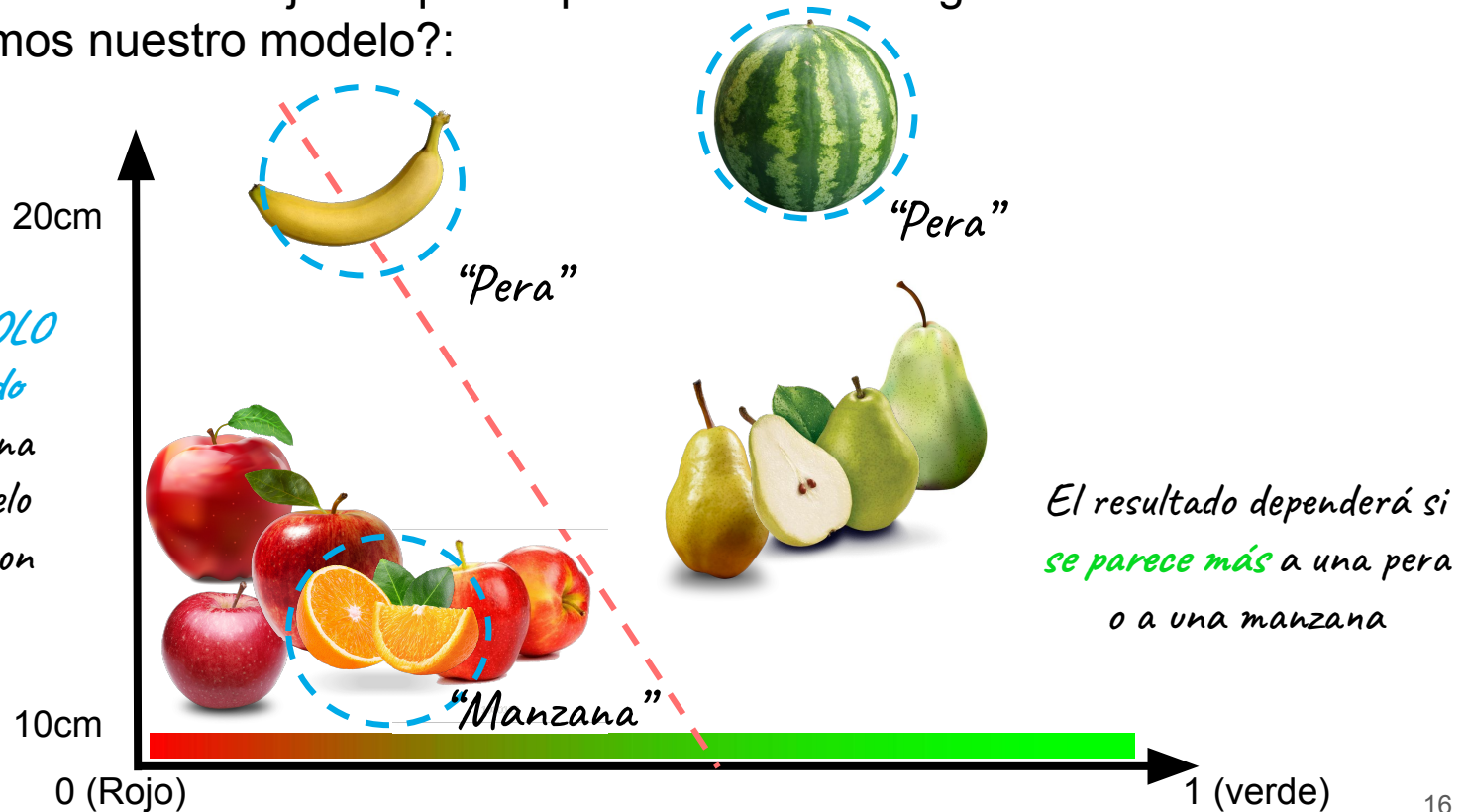
Como el espacio queda dividido en 2 partes, al agregar un nuevo objeto, podemos decir si es **más probable** que sea una manzana o una pera **según la región** en la que se encuentre por sus **características**:



# Peras y Manzanas

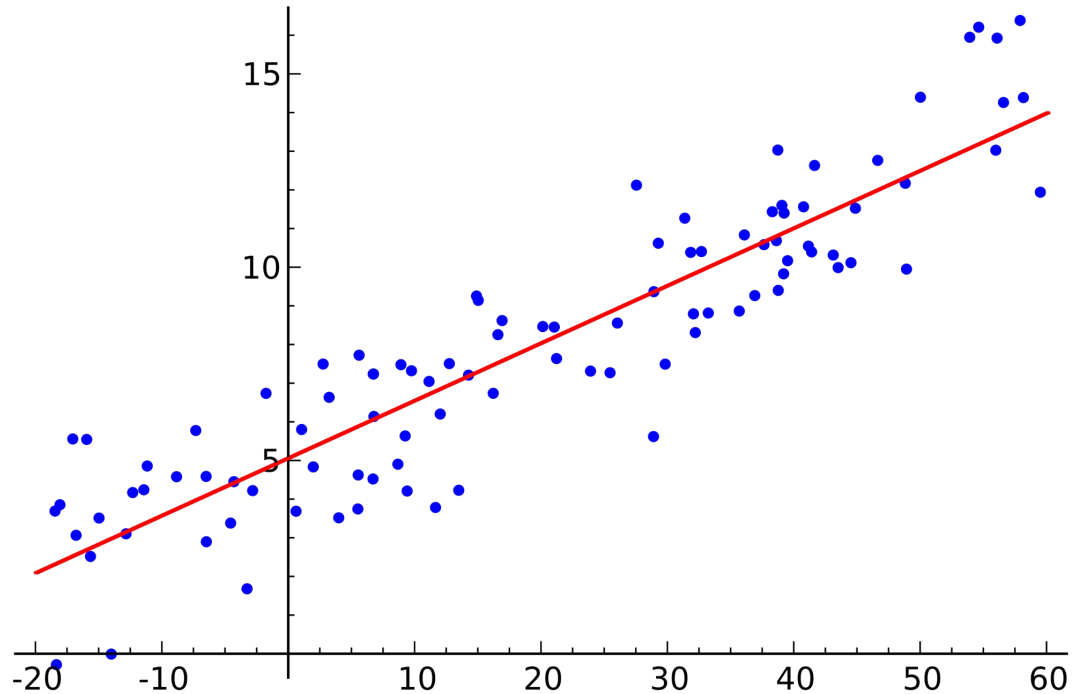
¿Qué pasa si introducimos objetos que no pertenezcan a ninguna de las clases con las que entrenamos nuestro modelo?:

Cada objeto nuevo *SOLO* podrá ser clasificado como Pera o Manzana porque nuestro modelo solo fue entrenado con esas 2 clases





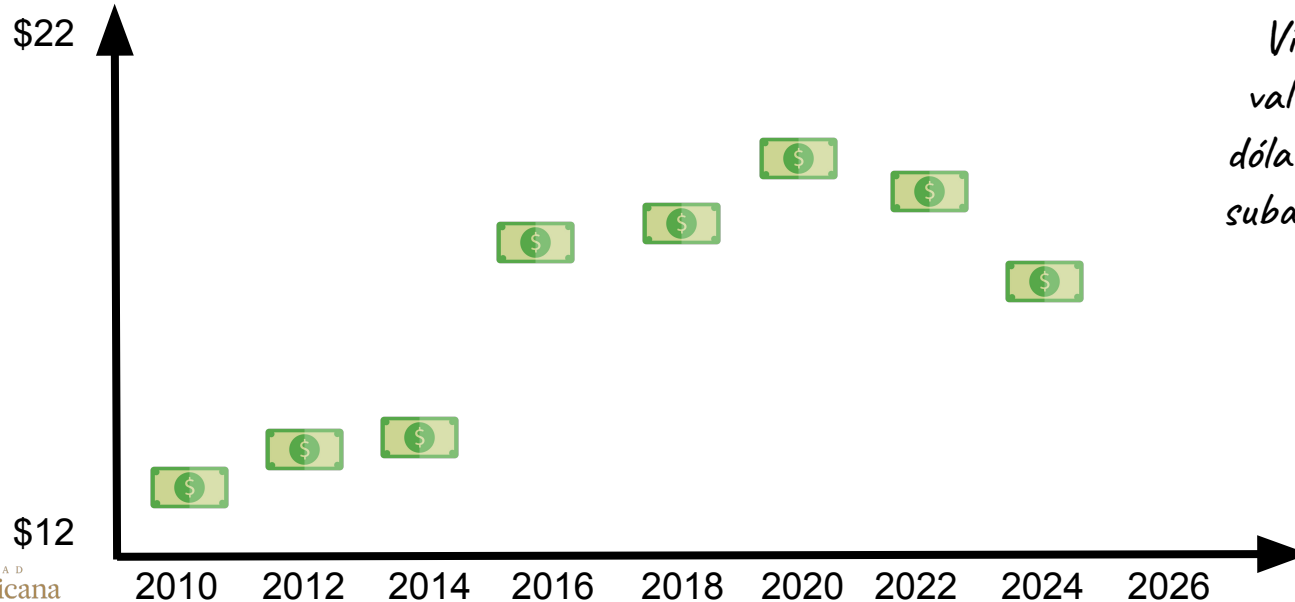
# REGRESIÓN



# Modelos de regresión

Un modelo de regresión es una **función** capaz de **relacionar ejemplos** del mismo tipo de objetos y **predecir su comportamiento** a través de una variable.

Por ejemplo, el valor promedio de un dólar por cada año:

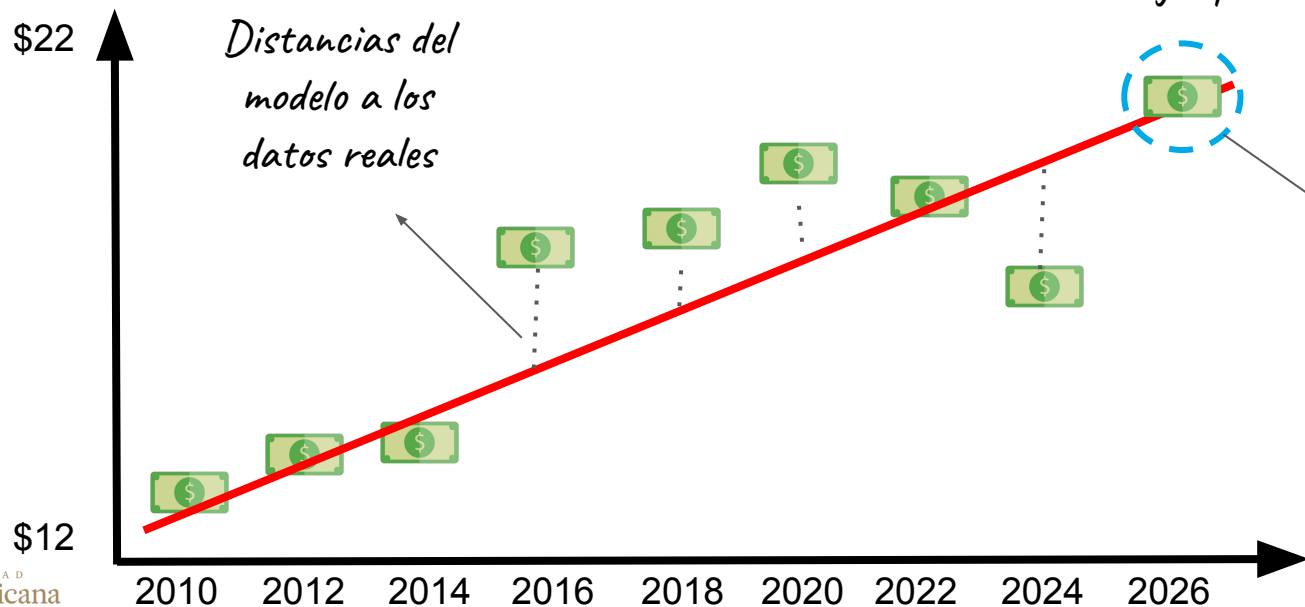


*Viendo el histórico de valores que ha tenido el dólar, ¿Crees que el precio suba o baje para el 2026?*

# Modelos de regresión

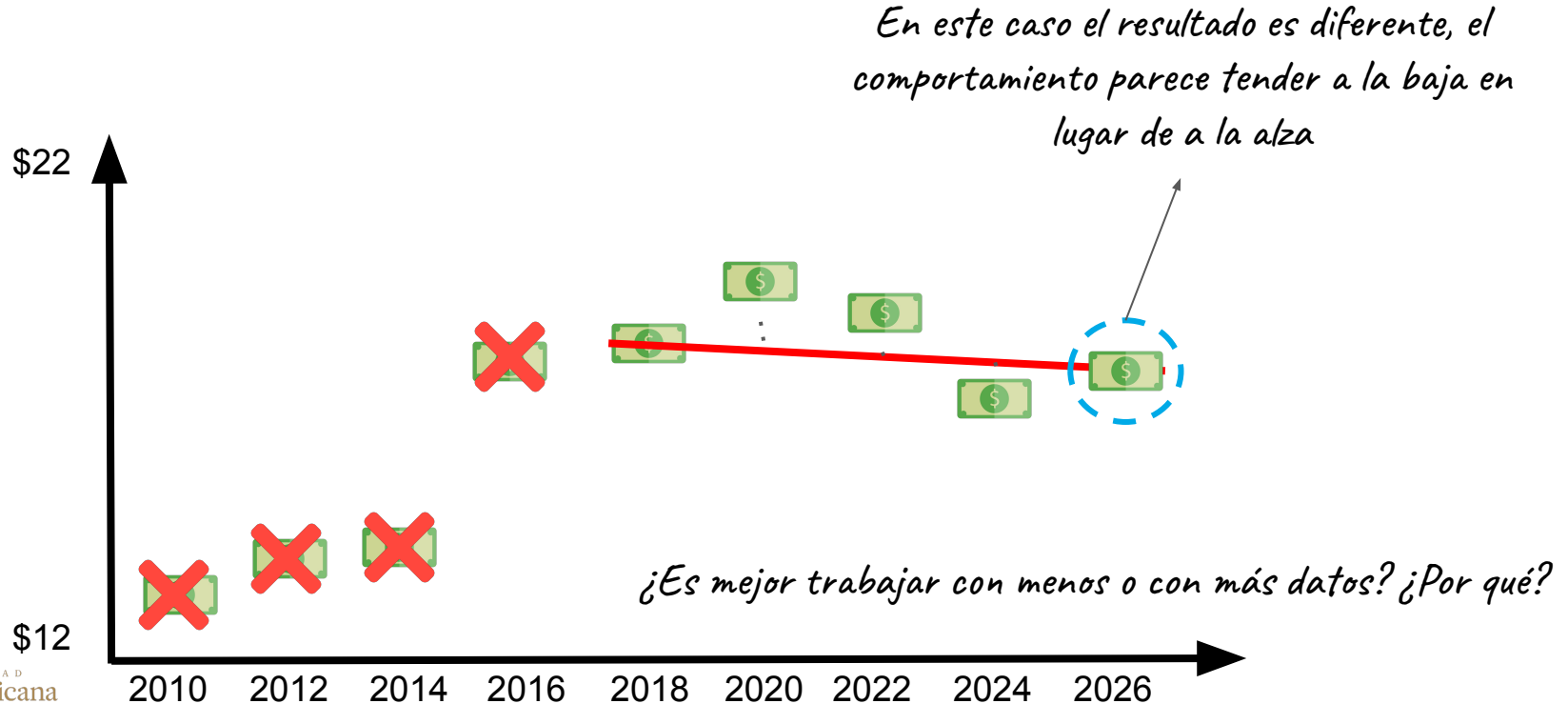
El modelo de regresión se **adapta a los datos**, y genera una **función** que describe el **comportamiento** o **tendencia** de los mismos:

*El modelo intentará tener la MENOR distancia entre la función y cada uno de los ejemplos con los que fue creado*

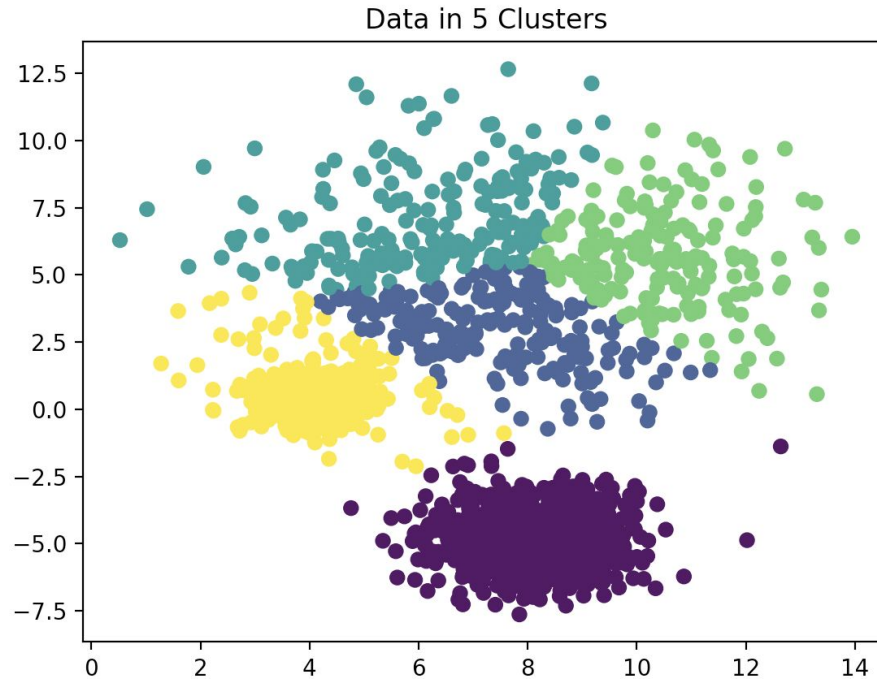


# Modelos de regresión

Pero... ¿Qué pasa si solo consideramos del año 2018 en adelante? ¿El resultado de la predicción es diferente?:



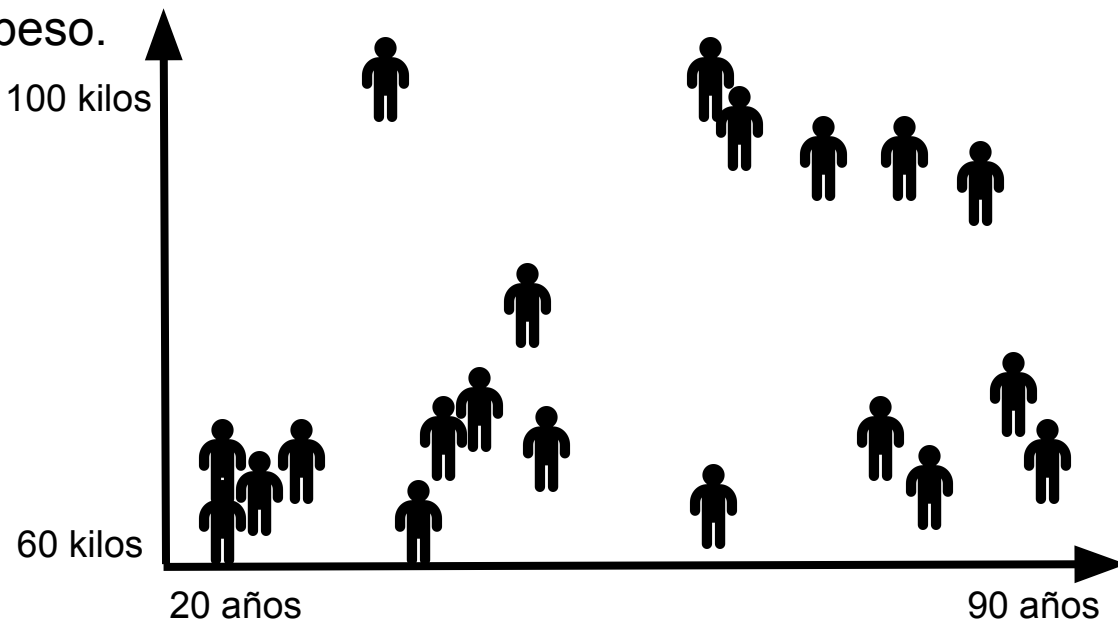
# CLUSTERIZACIÓN



# Clusterización

La clusterización es un método de **aprendizaje NO supervisado** que consiste en **crear grupos** de objetos que estén **más cerca entre ellos** según sus **características**.

Imagina que tienes un conjunto de personas de las cuáles lo único que conoces es su edad y peso.

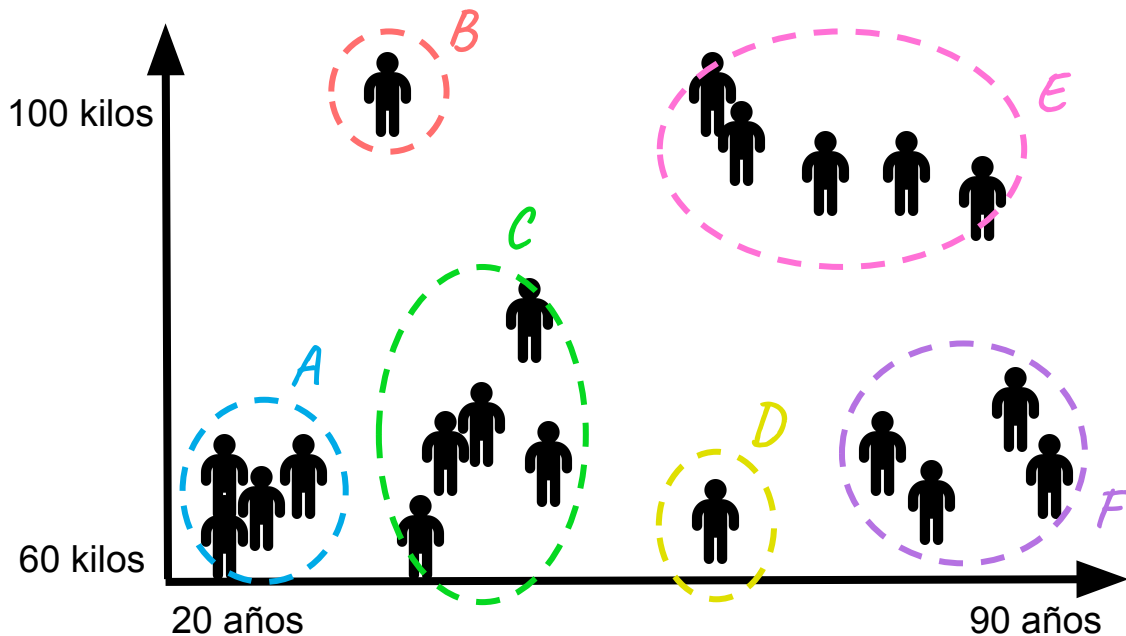


*Con estas personas,  
¿En cuántos grupos  
crees que podrías  
dividir las?*

# Clusterización

Podrías dividir este grupo de personas en 2, 3, 4, 5, 6, 7 grupos o incluso mas.

Vamos a dividirlos en 6 grupos diferentes:



*¿Crees que quedaría mejor la división si hay más/menos grupos?  
¿Por qué?*

# Clusterización

Cada grupo tiene **características similares** entre los participantes de ese grupo por lo que ahora, podríamos ponerles **nombre (etiqueta)** o usar los grupos para tomar **decisiones**, por ejemplo:

