

Modelos de aprendizaje no supervisado & ética

Taller comercio y desarrollo regional

Módulo Aprendizaje de máquinas

Curso - Taller

Hoy

1. Modelos de aprendizaje no supervisado

¿Estructuras detrás de los datos? ¿Para qué nos sirve reducir dimensionalidad y agrupar filas?

2. IA puede

El ML es poderoso y todo, ¿pero en qué situaciones de la vida es útil?

3. Ética de la inteligencia artificial

El machine learning no es perfecto. ¿Cómo podemos usarlo cuando otras personas pueden salir afectadas?

4. Taller diferenciado

Tiempo para informe y para identificar aplicaciones concretas.



Photo by [The Coach Space](#) from [Pexels](#)

Modelos de aprendizaje no supervisado

Clustering y reducción de
dimensionalidad



Photo by [Mehrad Vosoughi](#) from [Pexels](#)

Recordemos que contamos con:

- Datos que tienen características de cada observación PERO NO CONTAMOS CON ETIQUETAS, así que no queremos predecir algo en particular.

Identificador	Dato útil 1	Dato útil 2
Alfredo	12	7432
Juliana	16	9375

Recordemos que contamos con:

- Análisis de reducción de dimensionalidad
- Análisis de agrupamiento o *clustering*

Identificador	Dato útil 1	Dato útil 2
Alfredo	12	7432
Juliana	16	9375

PCA

Análisis de componentes principales



Photo by [Mehrad Vosoughi](#) from [Pexels](#)

El reto

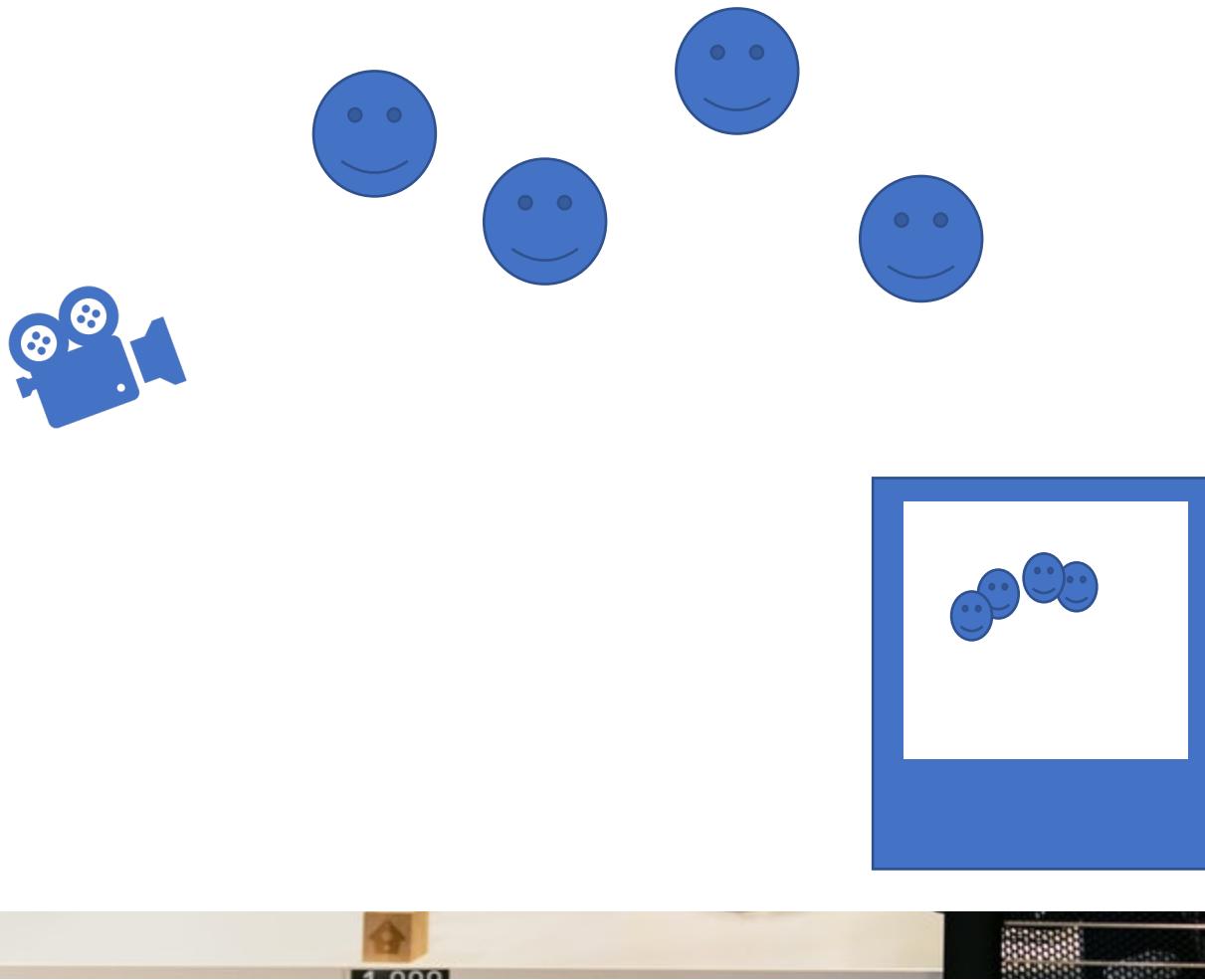
- Tenemos una base de datos con muchas columnas, y queremos representarla en menos dimensiones, sin perder mucha información

X1	X2	X3	X4
...
...

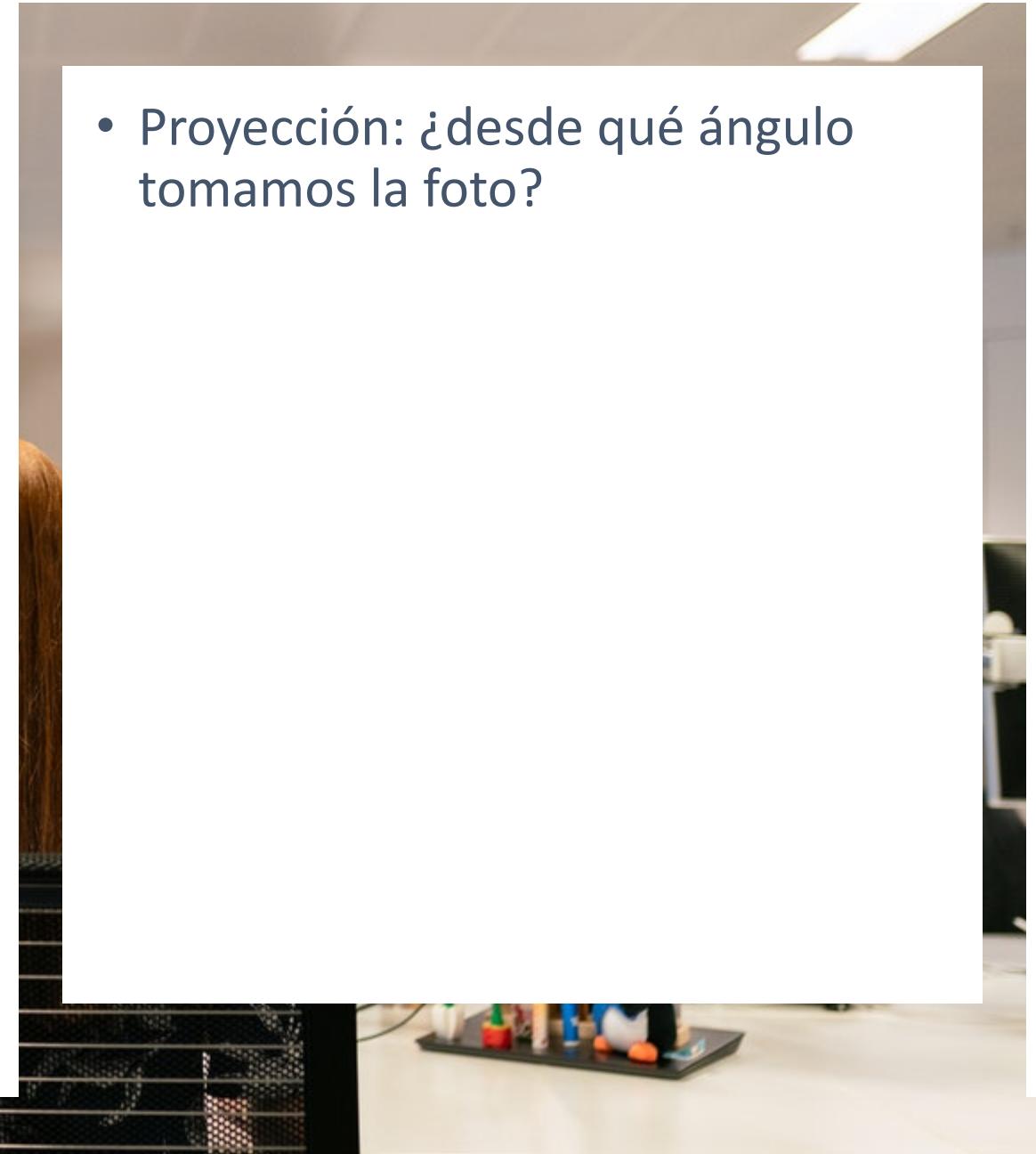


Z1	Z2
...	...
...	...

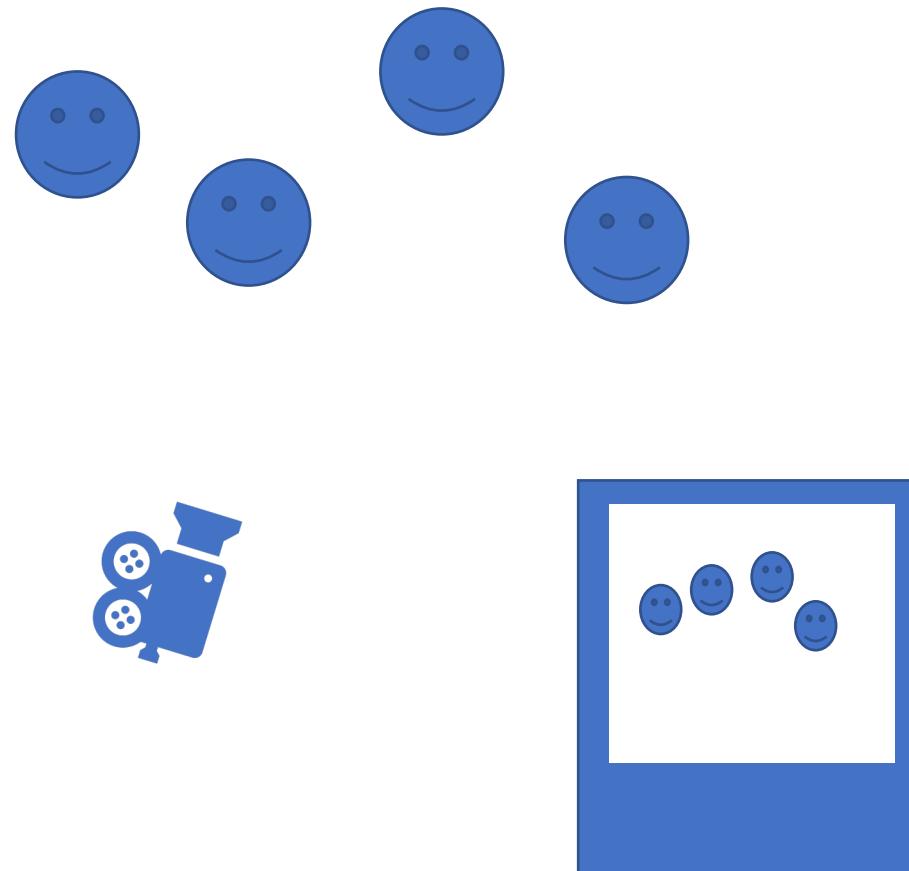
Análisis de Componentes Principales (PCA)



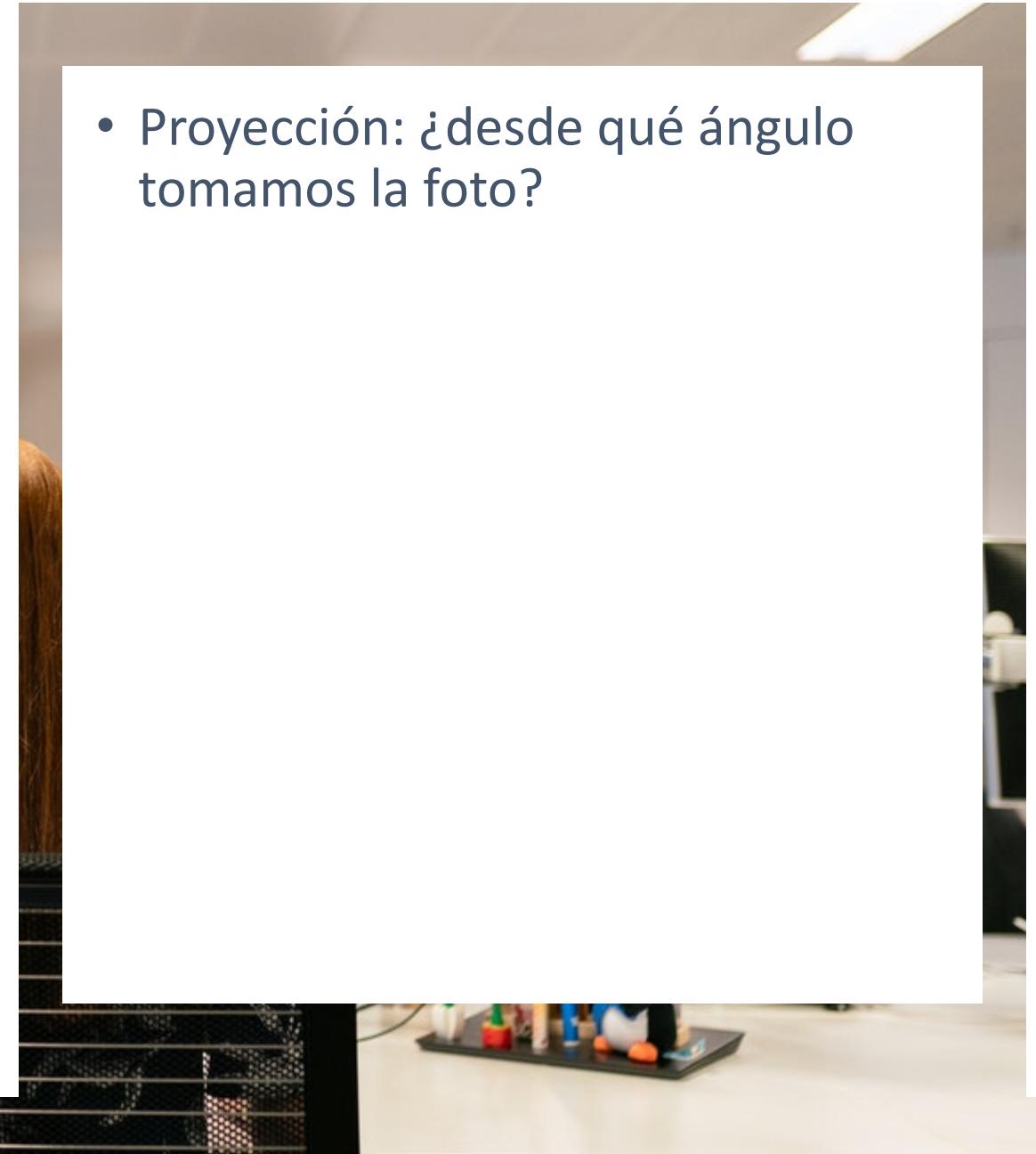
- Proyección: ¿desde qué ángulo tomamos la foto?



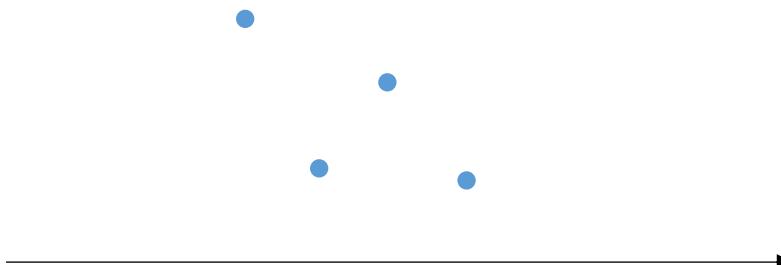
Análisis de Componentes Principales (PCA)



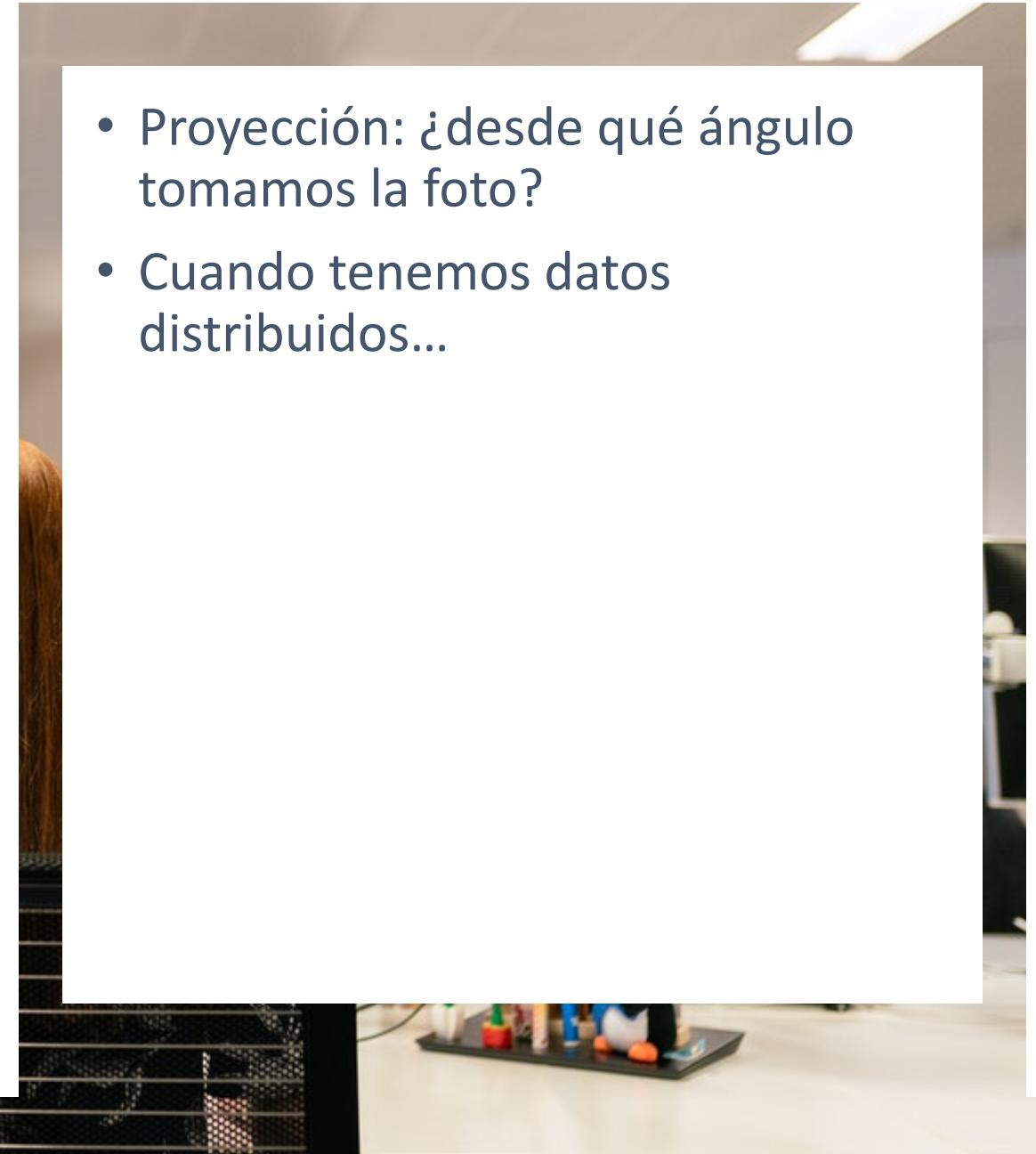
- Proyección: ¿desde qué ángulo tomamos la foto?



Análisis de Componentes Principales (PCA)



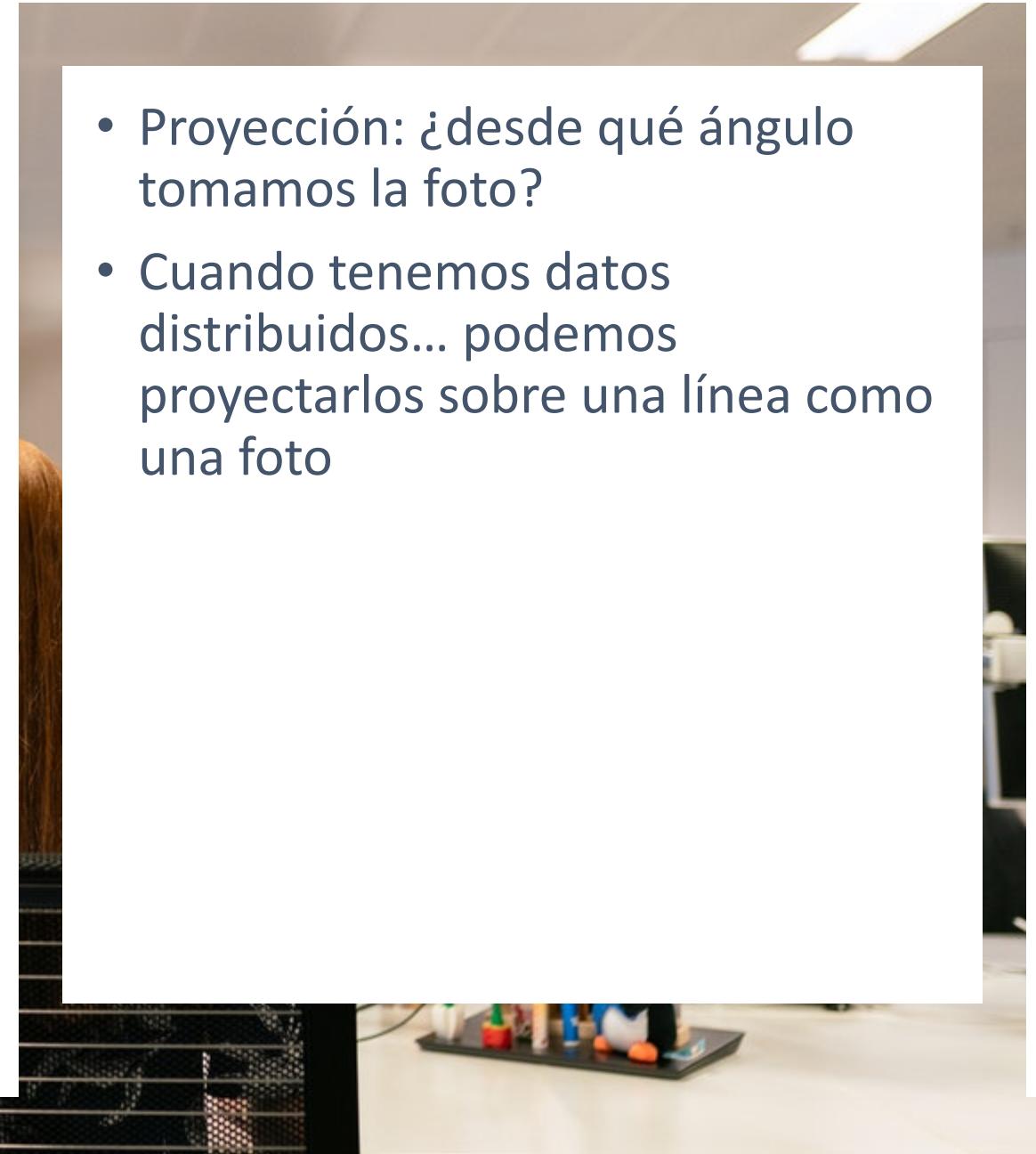
- Proyección: ¿desde qué ángulo tomamos la foto?
- Cuando tenemos datos distribuidos...



Análisis de Componentes Principales (PCA)



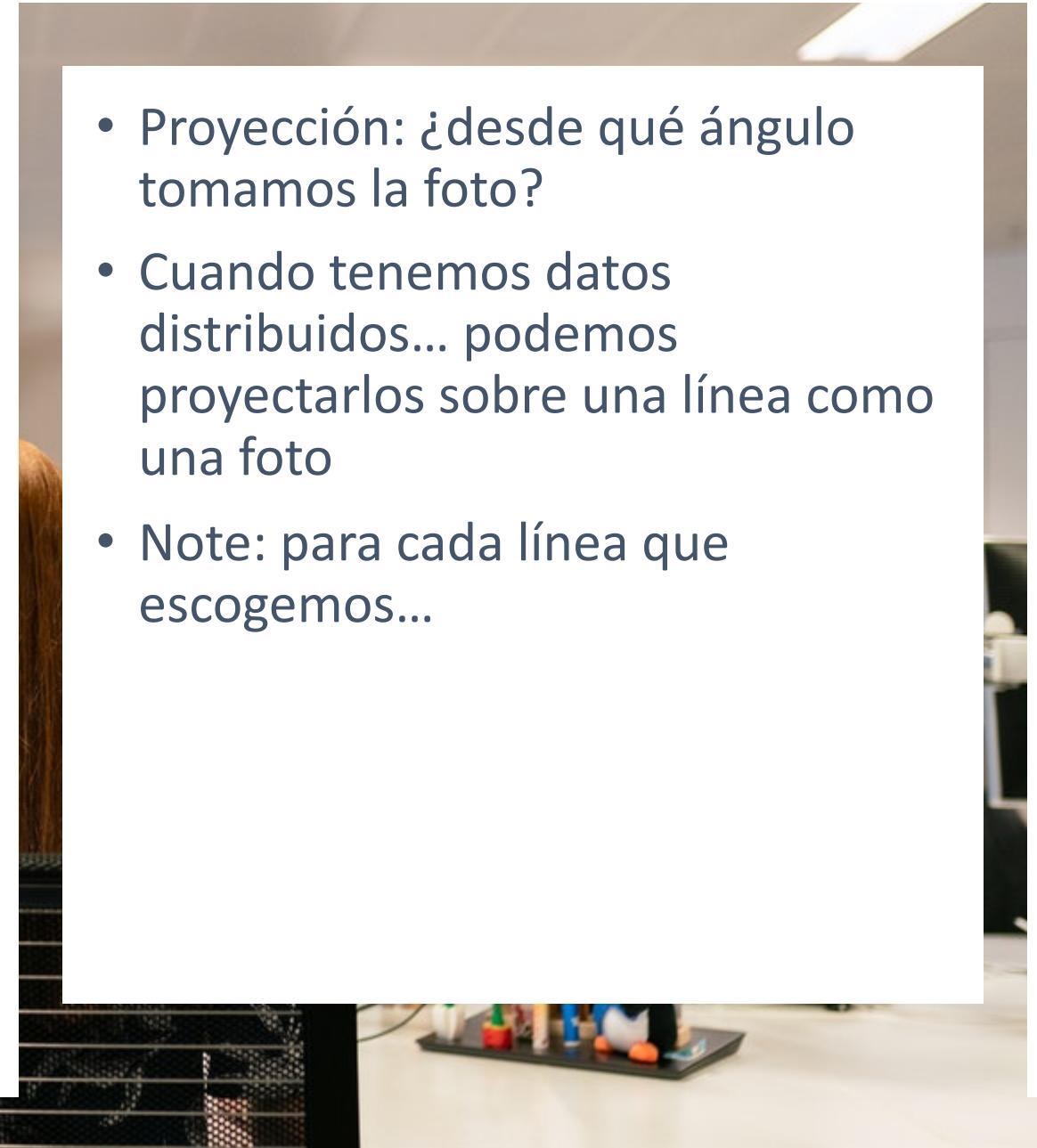
- Proyección: ¿desde qué ángulo tomamos la foto?
- Cuando tenemos datos distribuidos... podemos proyectarlos sobre una línea como una foto



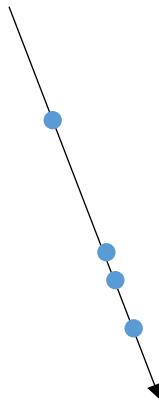
Análisis de Componentes Principales (PCA)

-
-
-
-

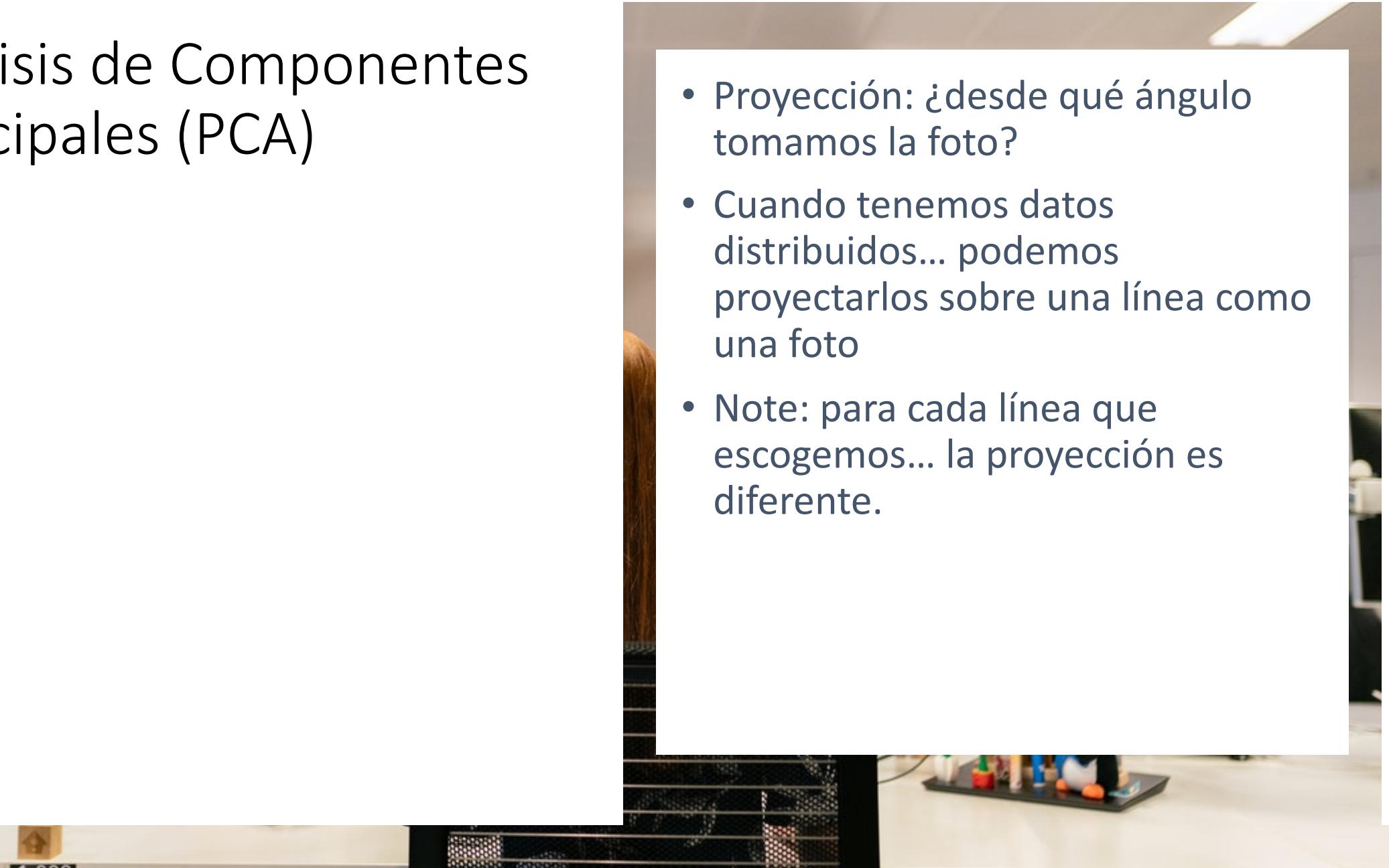
- Proyección: ¿desde qué ángulo tomamos la foto?
- Cuando tenemos datos distribuidos... podemos proyectarlos sobre una línea como una foto
- Note: para cada línea que escogemos...



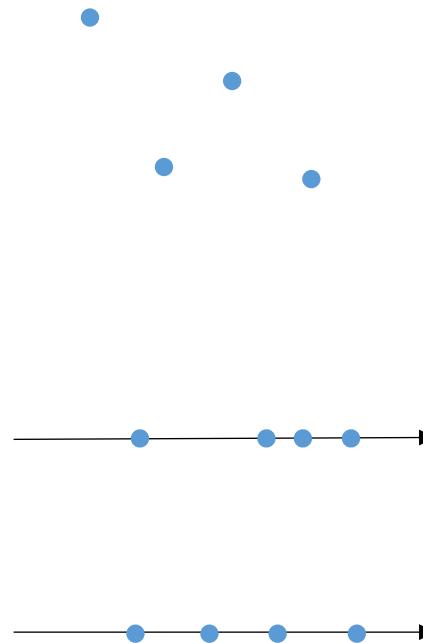
Análisis de Componentes Principales (PCA)



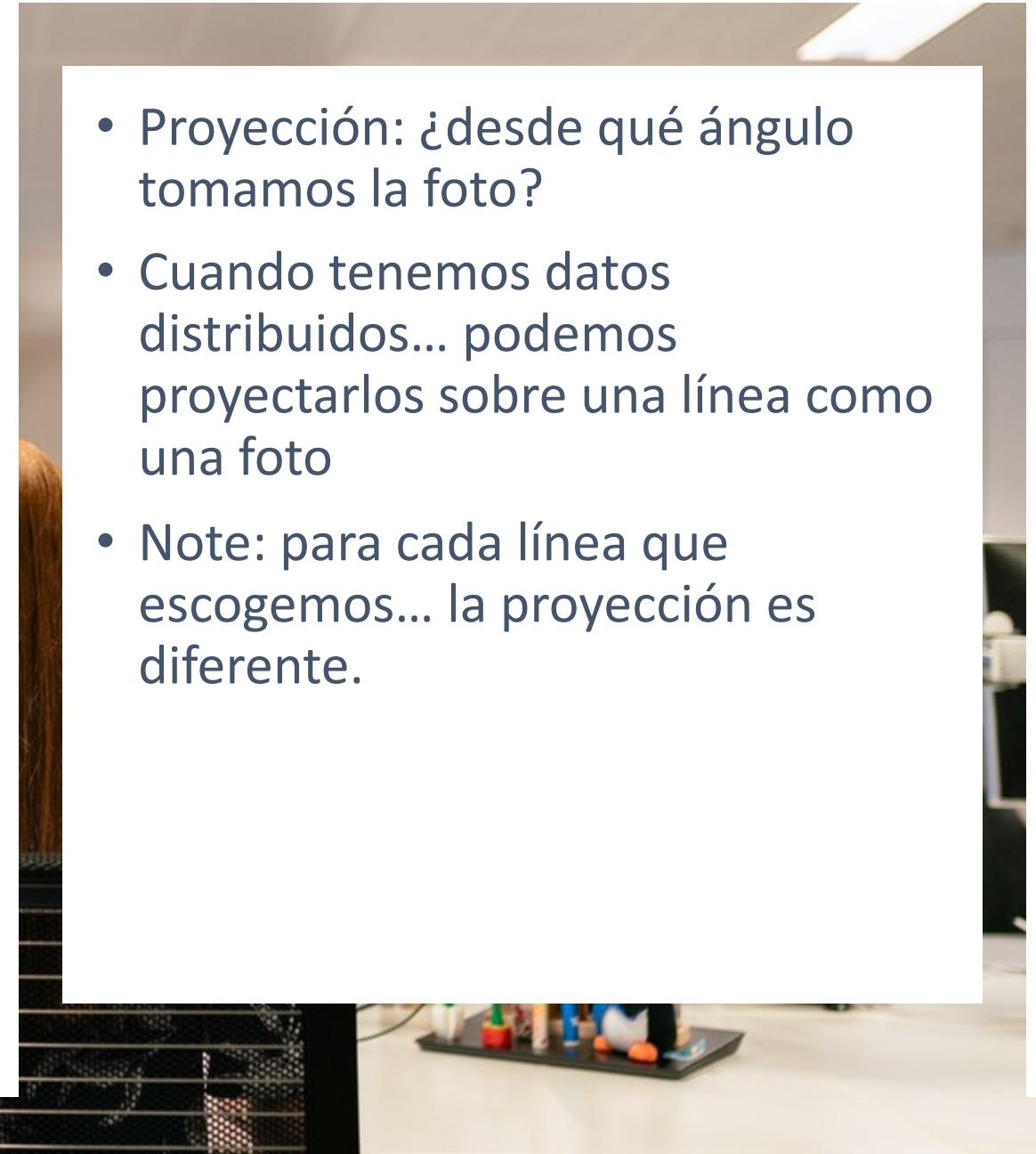
- Proyección: ¿desde qué ángulo tomamos la foto?
- Cuando tenemos datos distribuidos... podemos proyectarlos sobre una línea como una foto
- Note: para cada línea que escogemos... la proyección es diferente.



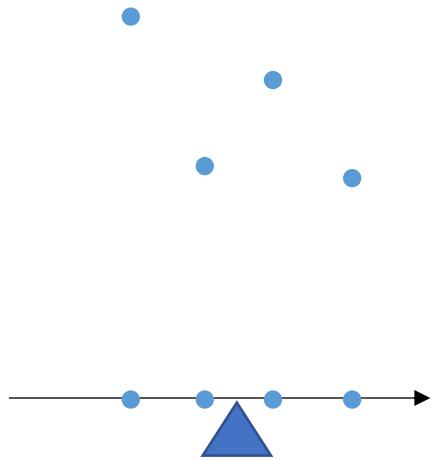
Análisis de Componentes Principales (PCA)



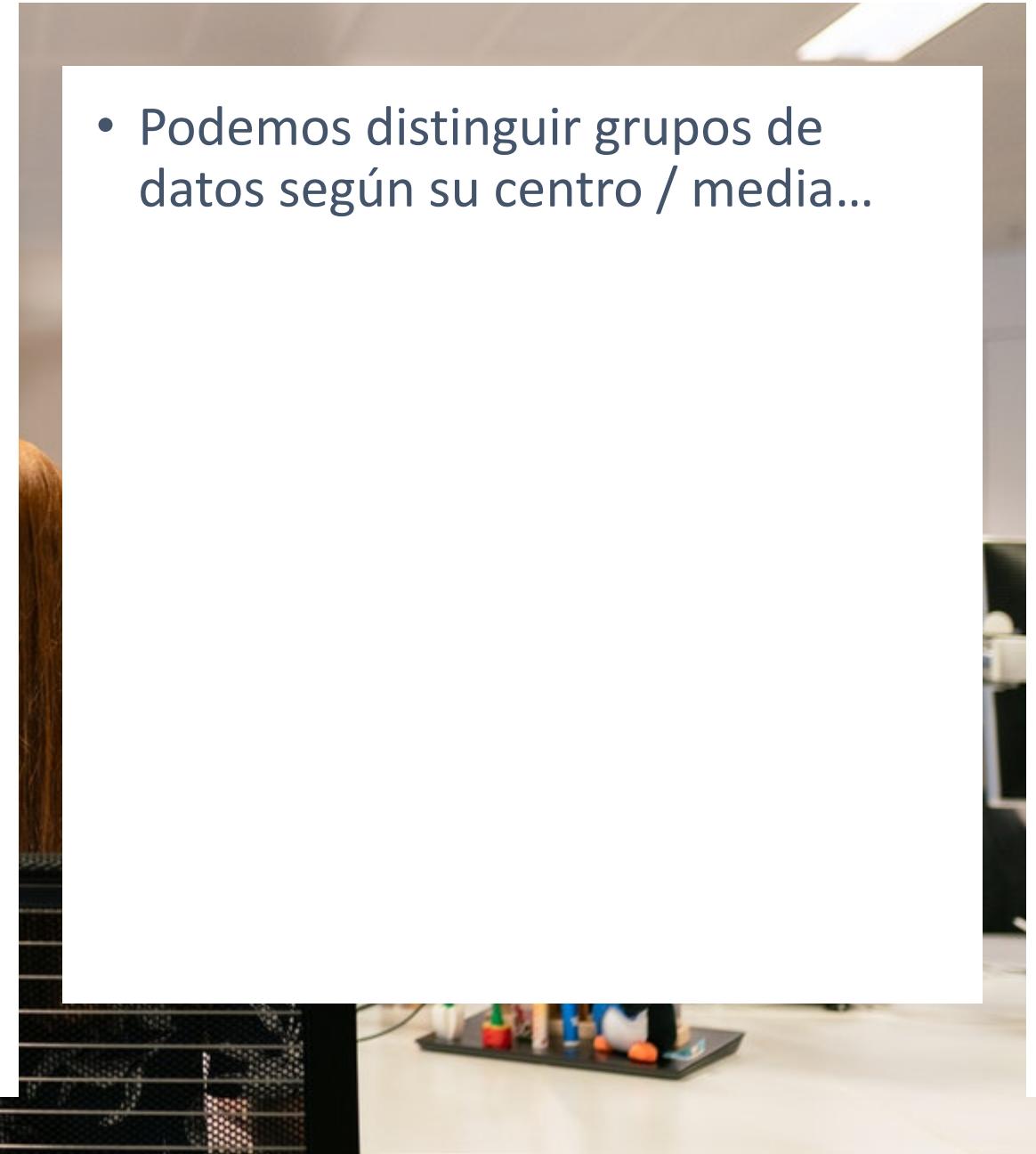
- Proyección: ¿desde qué ángulo tomamos la foto?
- Cuando tenemos datos distribuidos... podemos proyectarlos sobre una línea como una foto
- Note: para cada línea que escogemos... la proyección es diferente.



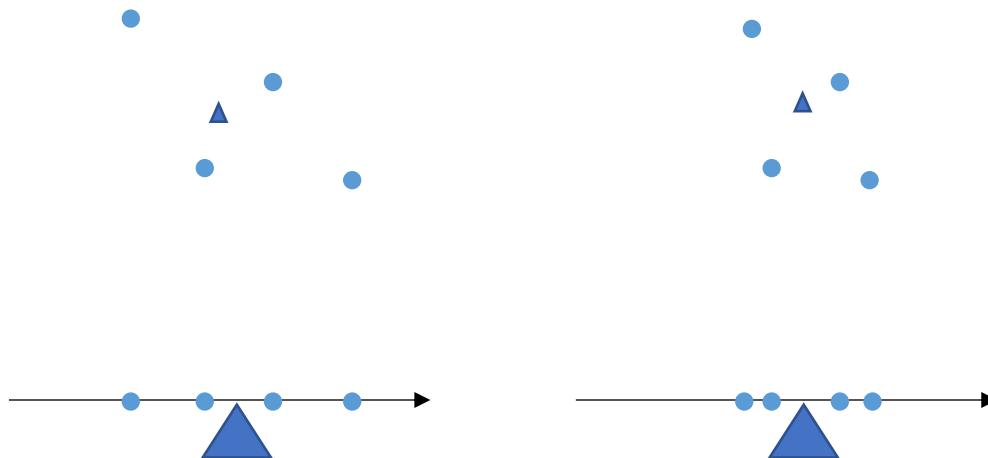
Análisis de Componentes Principales (PCA)



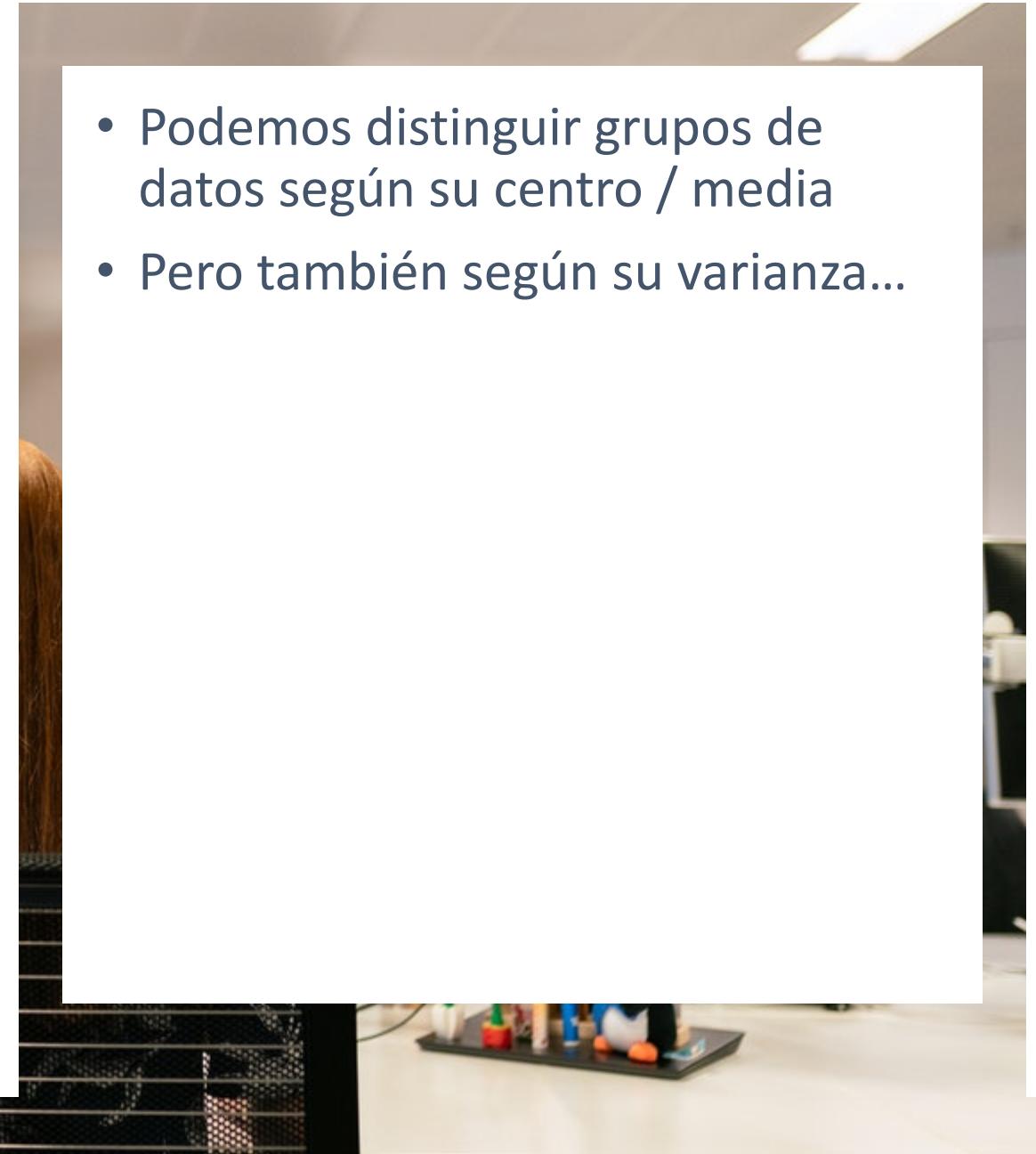
- Podemos distinguir grupos de datos según su centro / media...



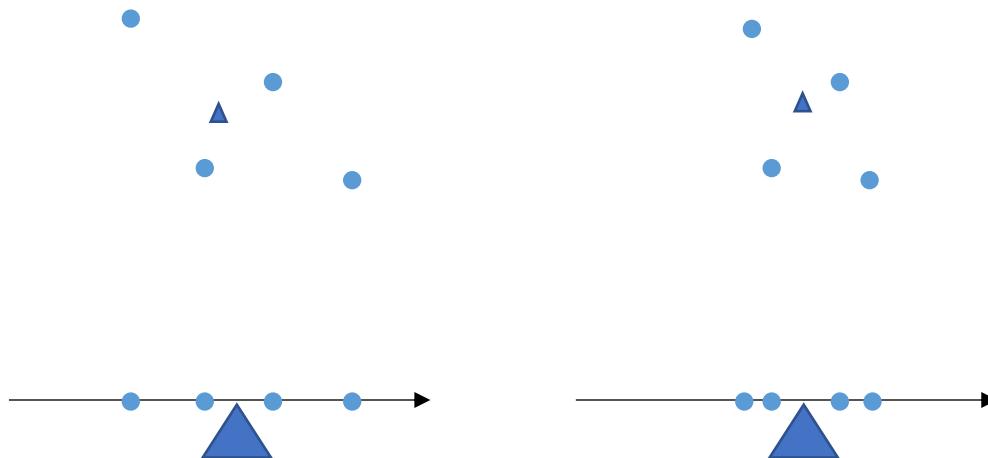
Análisis de Componentes Principales (PCA)



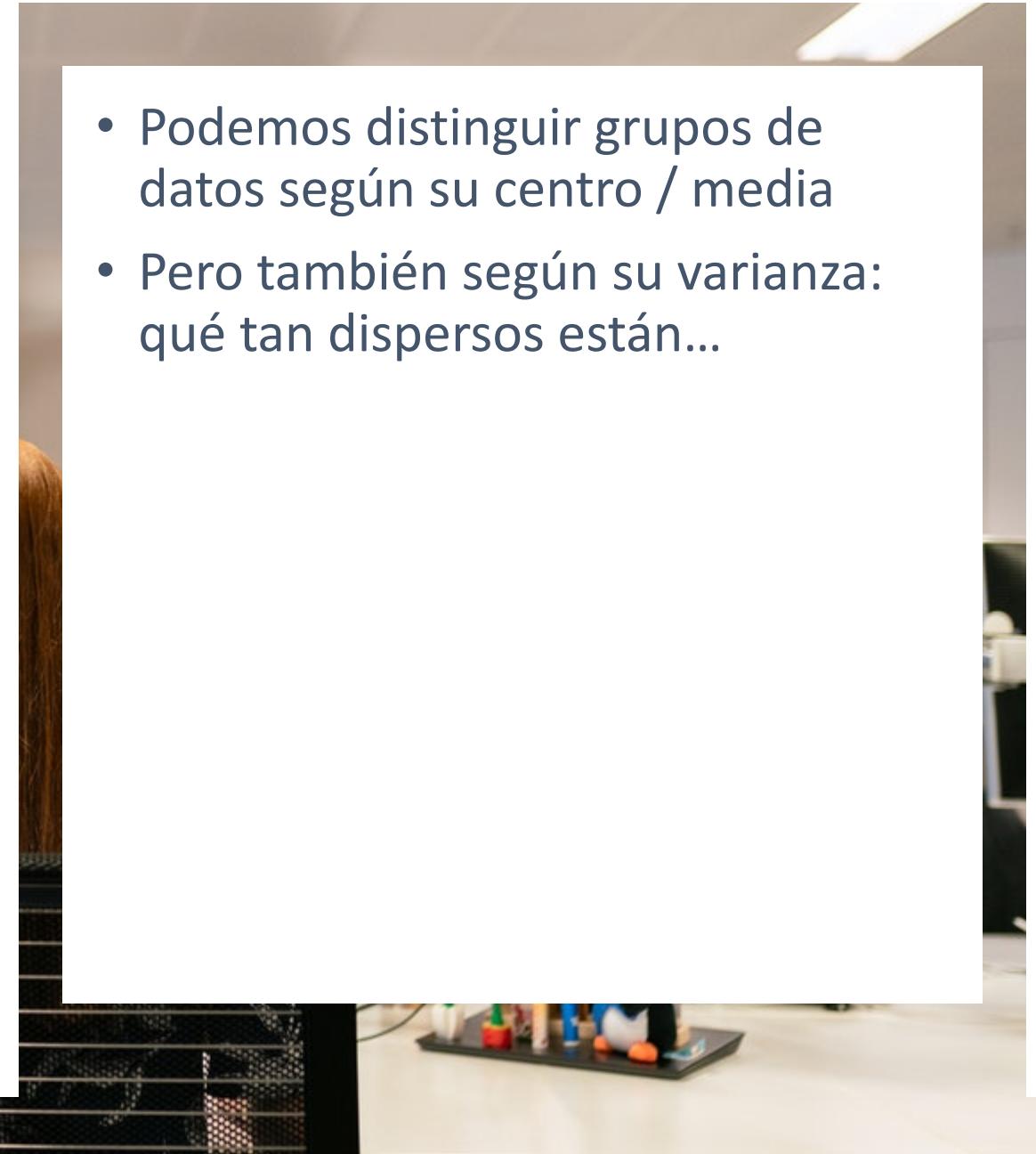
- Podemos distinguir grupos de datos según su centro / media
- Pero también según su varianza...



Análisis de Componentes Principales (PCA)



- Podemos distinguir grupos de datos según su centro / media
- Pero también según su varianza: qué tan dispersos están...

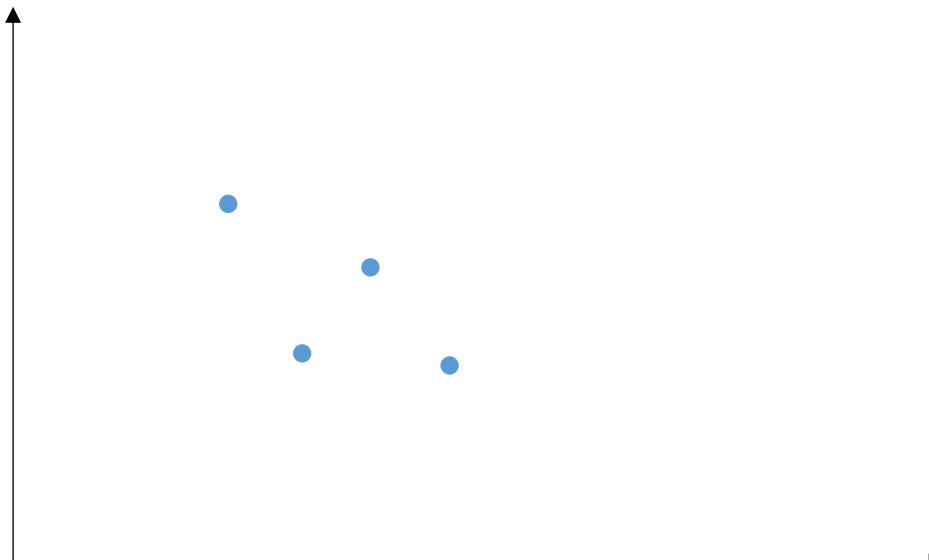


Análisis de Componentes Principales (PCA)



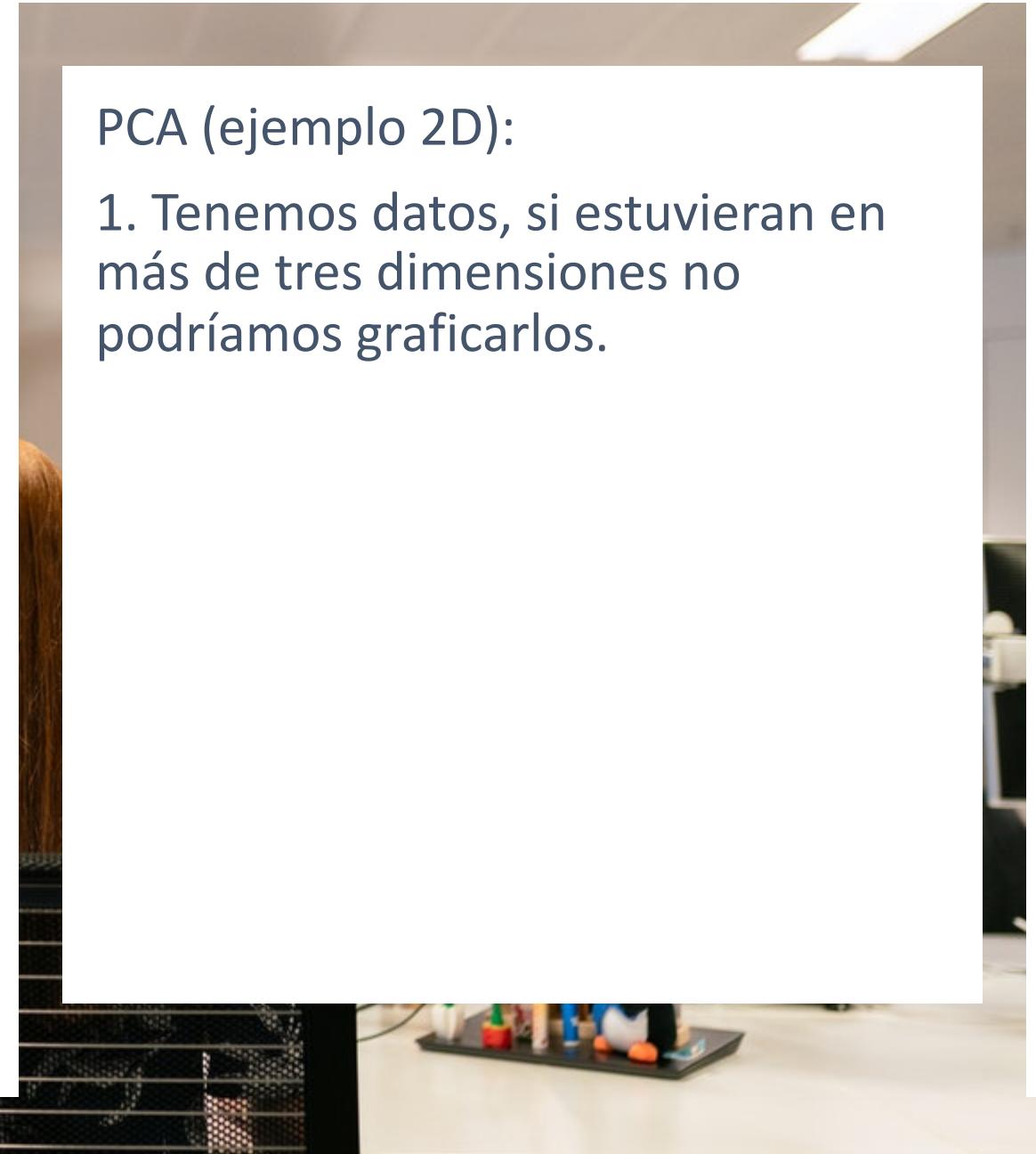
- Podemos distinguir grupos de datos según su centro / media
- Pero también según su varianza: qué tan dispersos están
- O la covarianza: muy relacionada con la correlación

Análisis de Componentes Principales (PCA)

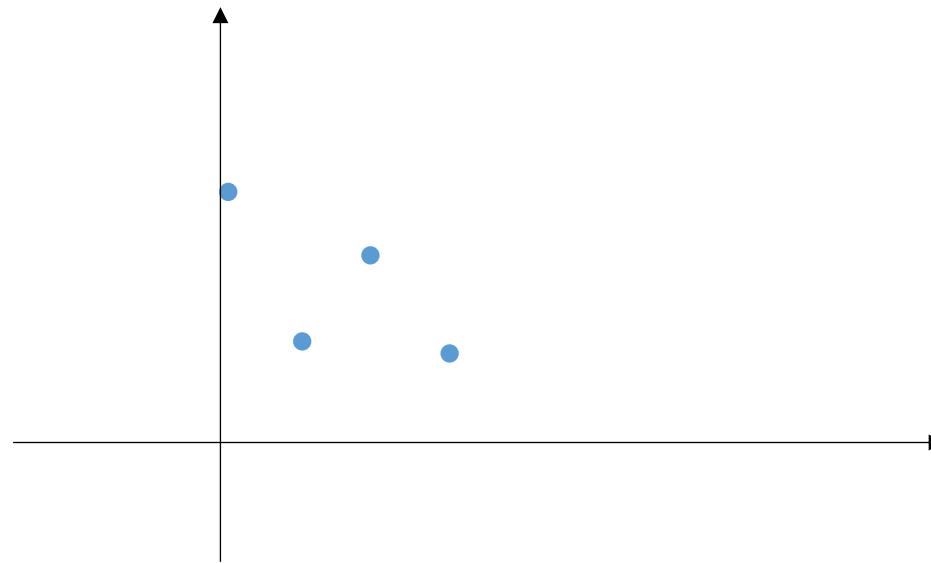


PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.



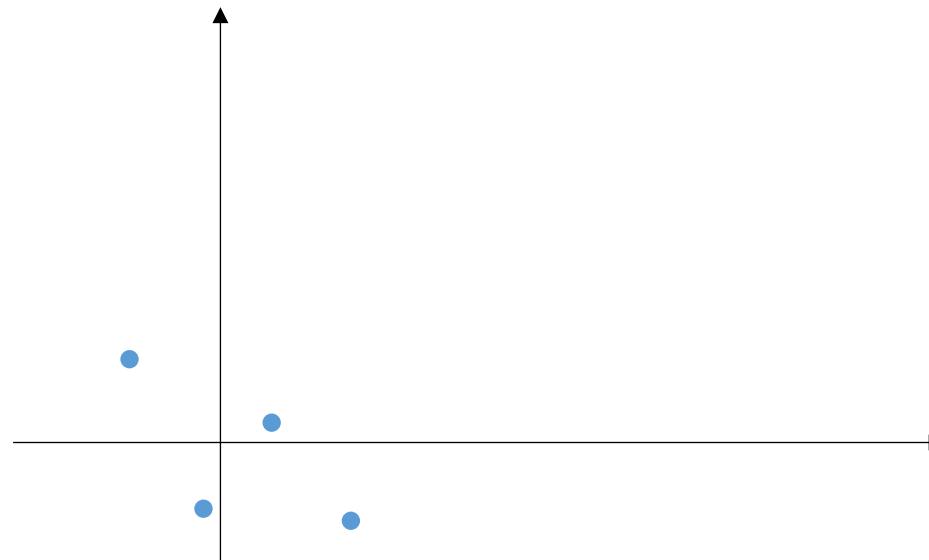
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.

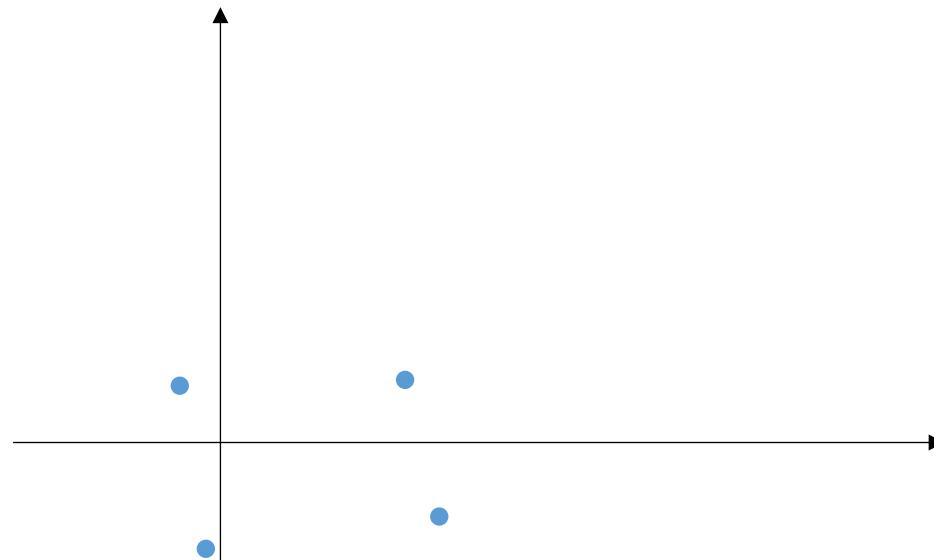
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.

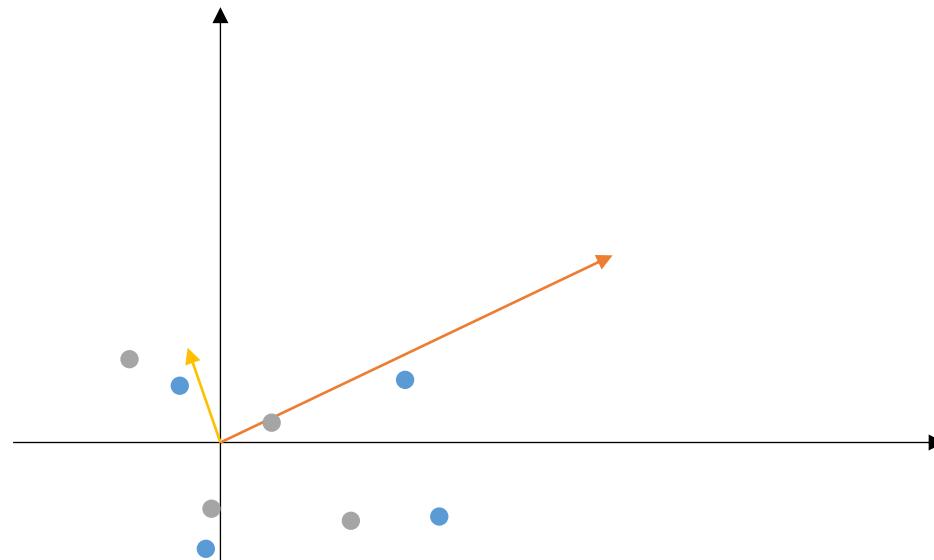
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano.

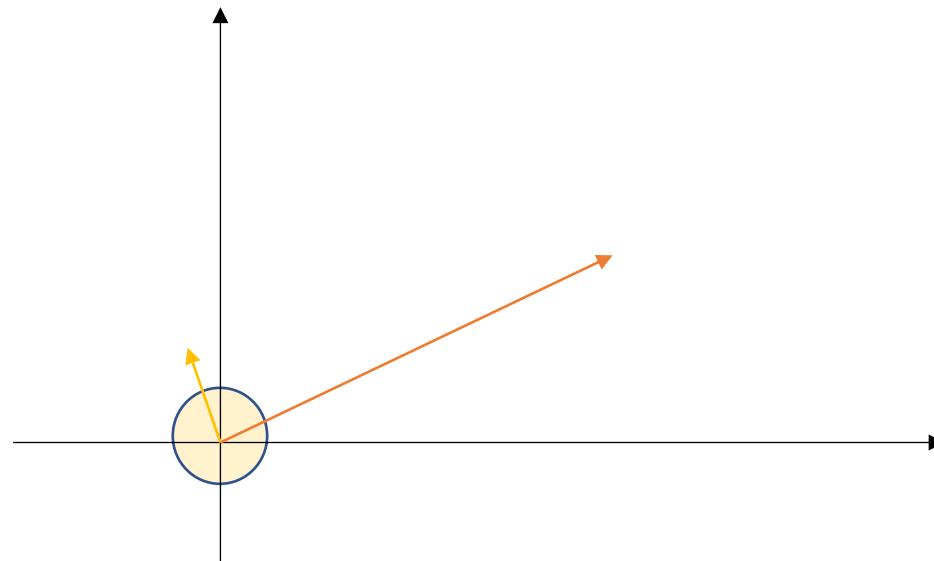
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: “autovectores”
 2. Magnitud: “autovalores”

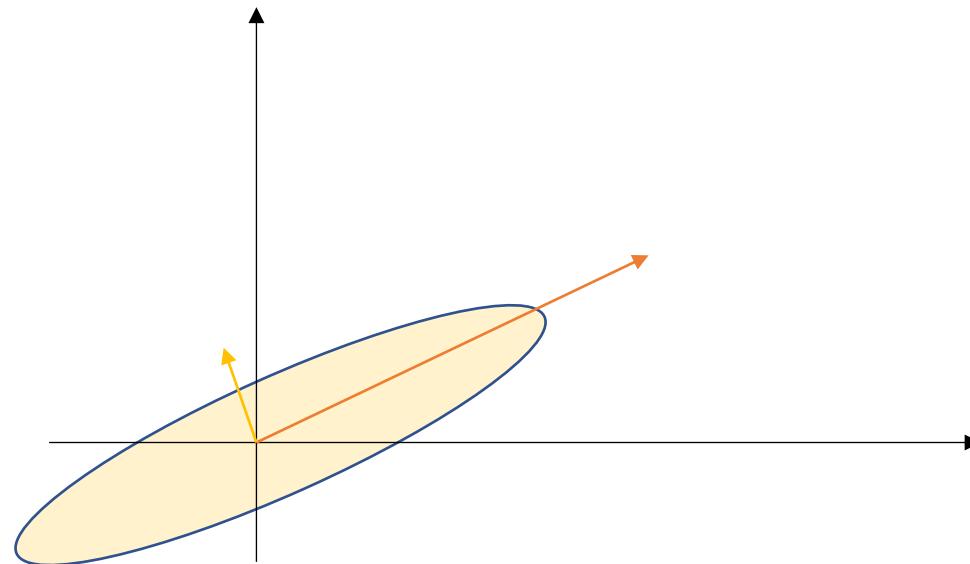
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: “autovectores”
 2. Magnitud: “autovalores”

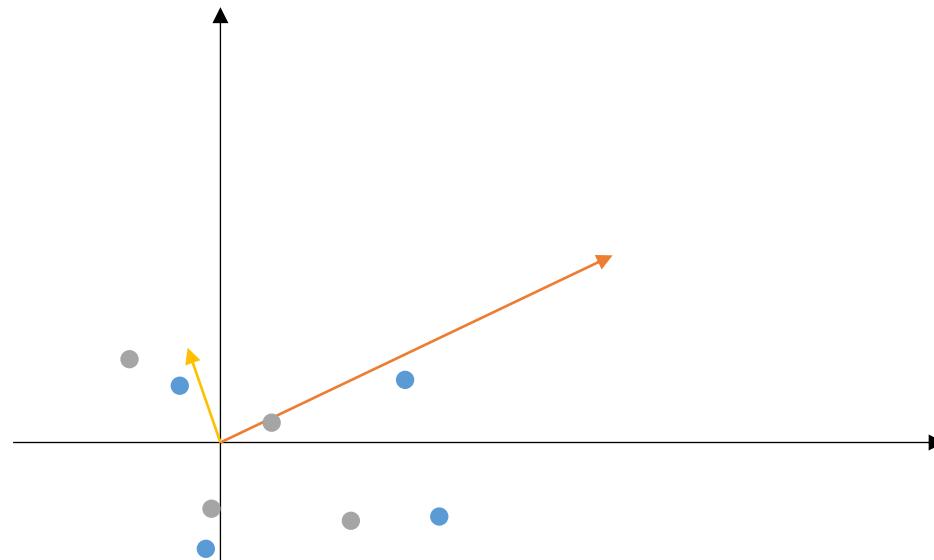
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: “autovectores”
 2. Magnitud: “autovalores”

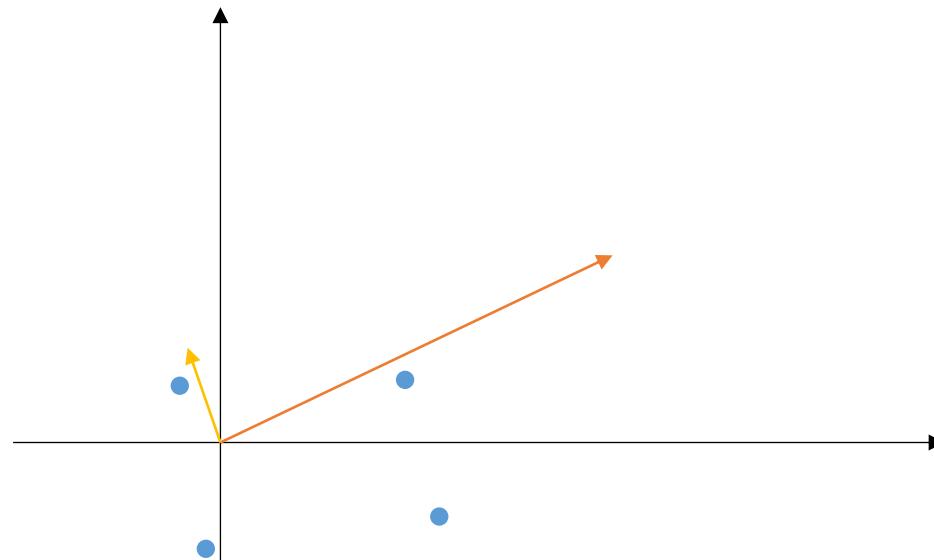
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: “autovectores”
 2. Magnitud: “autovalores”

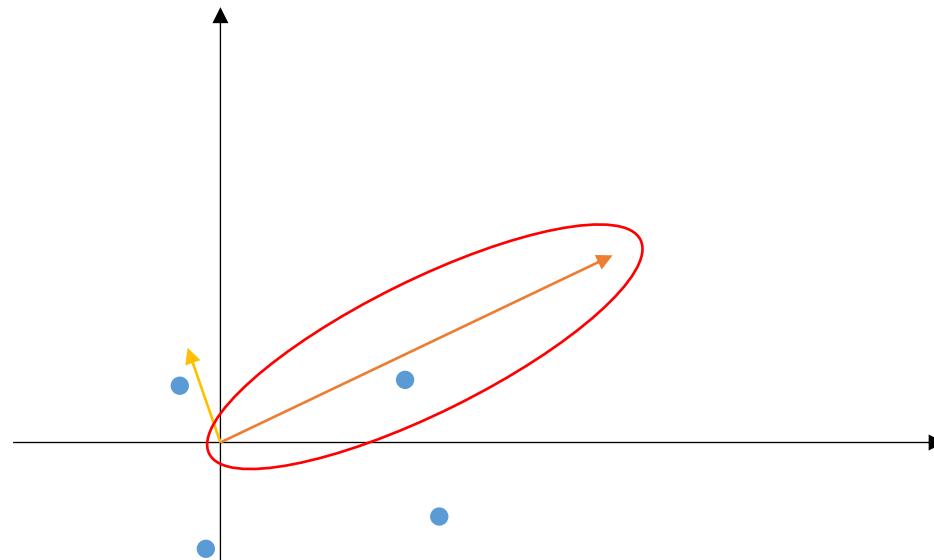
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: “autovectores”
 2. Magnitud: “autovalores”
4. Escogemos las líneas con mayor varianza

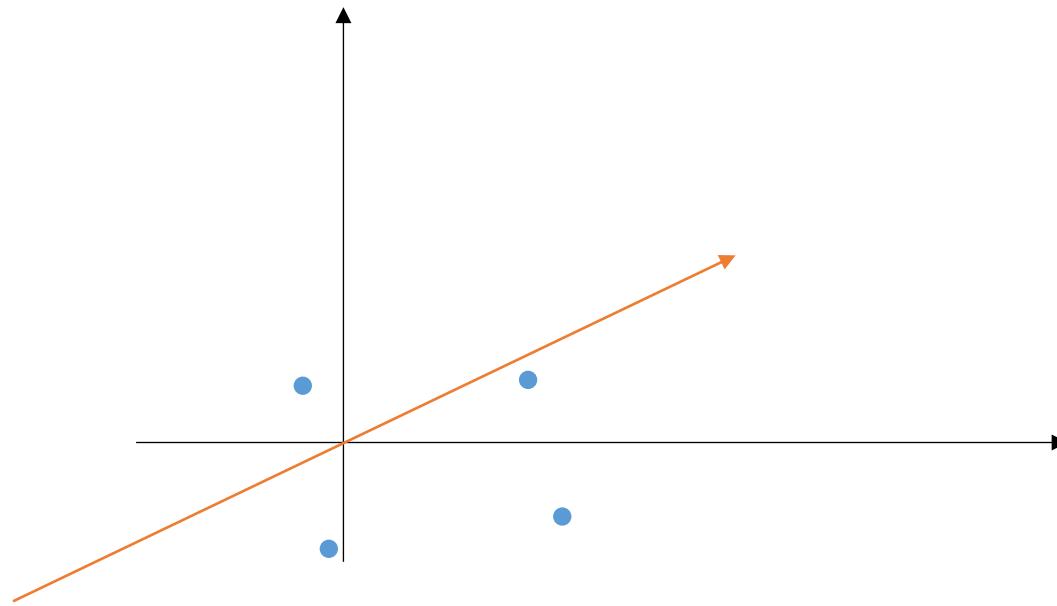
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: “autovectores”
 2. Magnitud: “autovalores”
4. Escogemos las líneas con mayor varianza

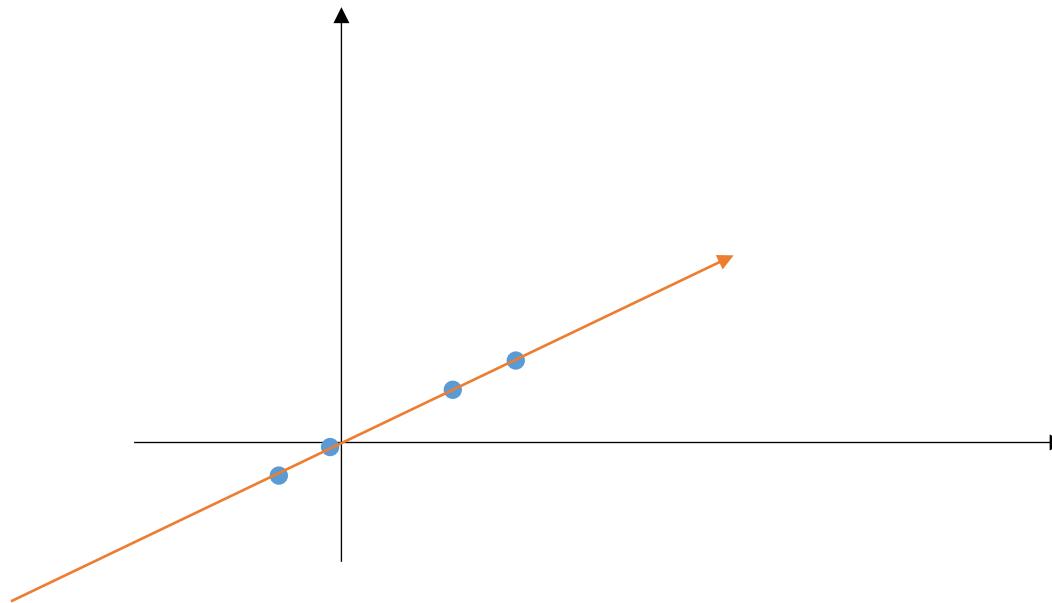
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: "autovectores"
 2. Magnitud: "autovalores"
4. Escogemos las líneas con mayor varianza y proyectamos...

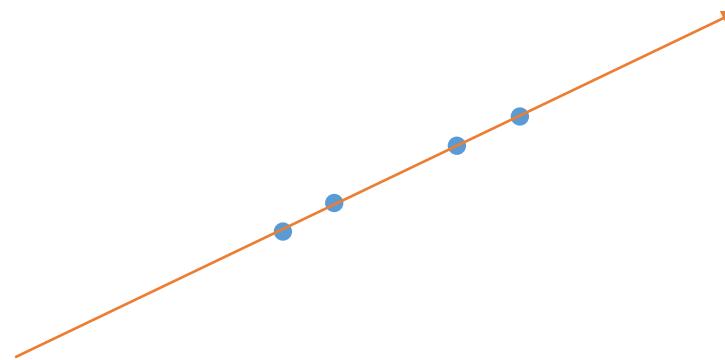
Análisis de Componentes Principales (PCA)



PCA (ejemplo 2D):

1. Tenemos datos, si estuvieran en más de tres dimensiones no podríamos graficarlos.
2. Los centramos.
3. Considerando las varianzas y covarianzas... estiramos y giramos el plano. ¿Para donde?
 1. Dirección: “autovectores”
 2. Magnitud: “autovalores”
4. Escogemos las líneas con mayor varianza y proyectamos...

Análisis de Componentes Principales (PCA)



Ahora cada dato anterior tiene una coordenada en una **nueva variable** (en lugar de 2). **Note:** no sólo eliminamos una columna, creamos una nueva a partir de las anteriores.

Tomamos una foto y conservamos una cantidad razonable de información.

¡Esto funciona con más de 2D!

Clustering

K-medias

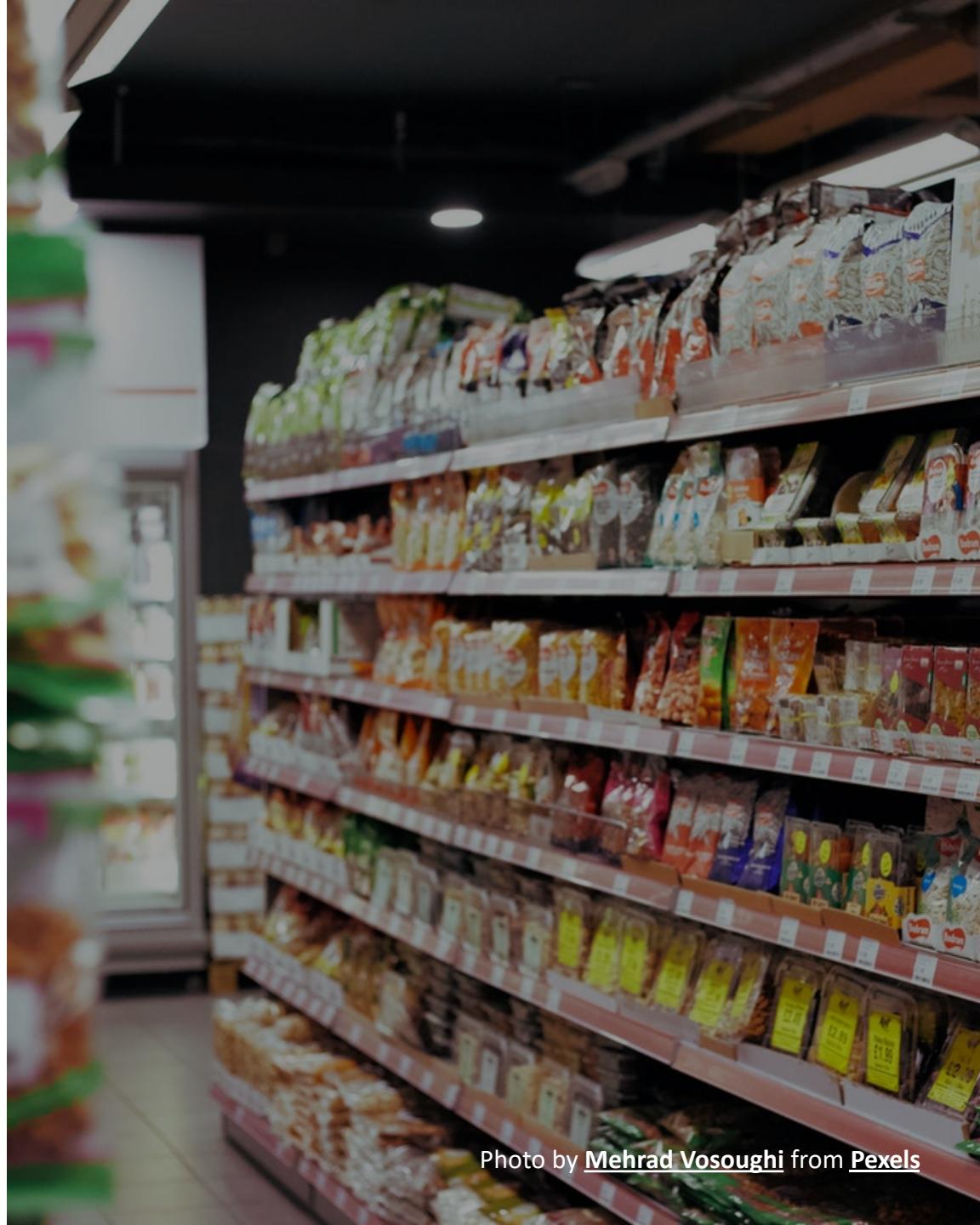


Photo by [Mehrad Vosoughi](#) from [Pexels](#)

El reto

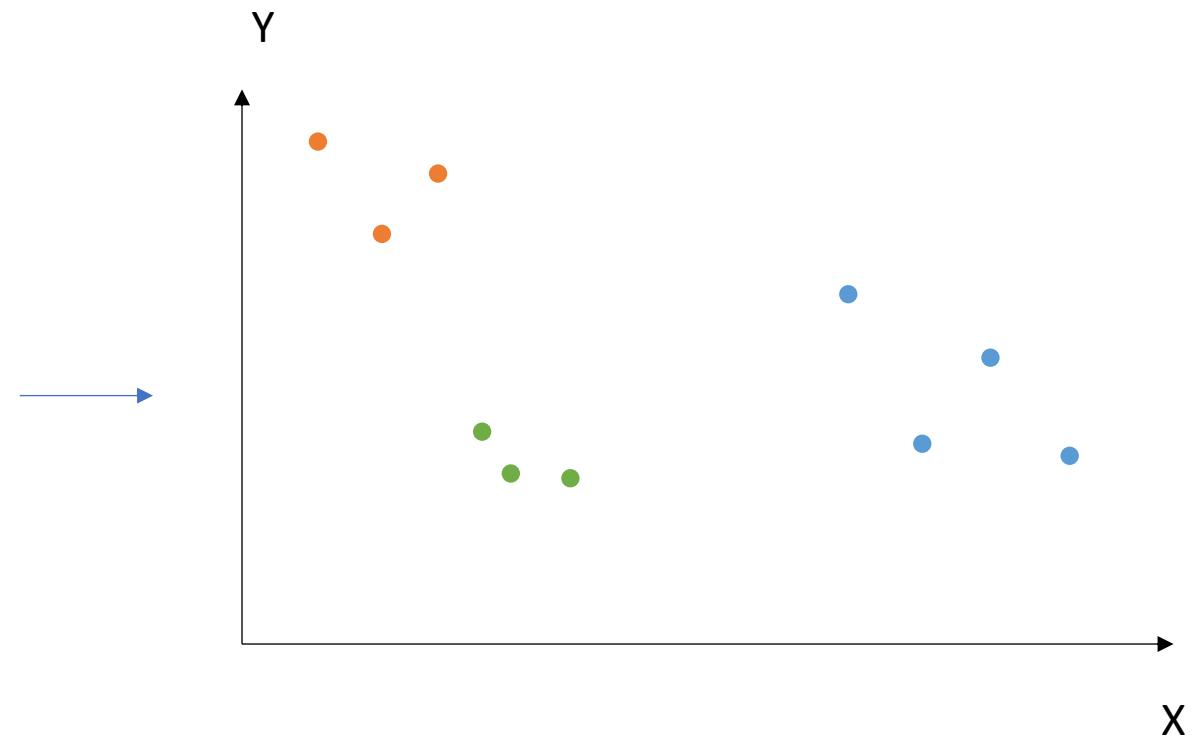
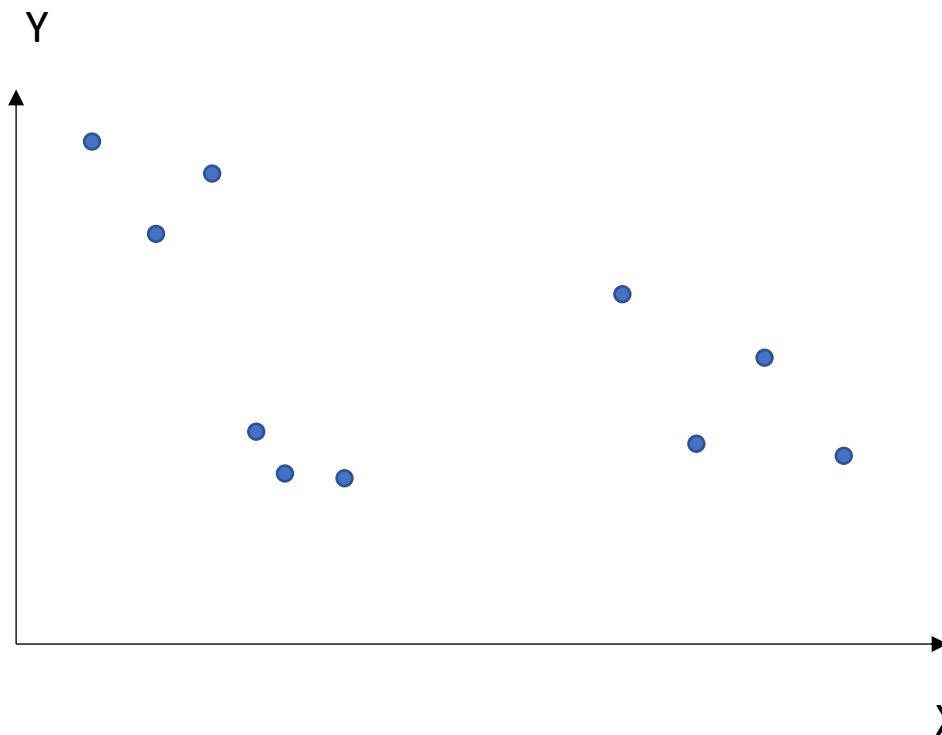
- Tenemos muchas observaciones, y queremos identificar qué datos son más cercanos a otros: formar grupos

Identificador	Dato útil 1	Dato útil 2
Alfredo	12	7432
Juliana	16	9375

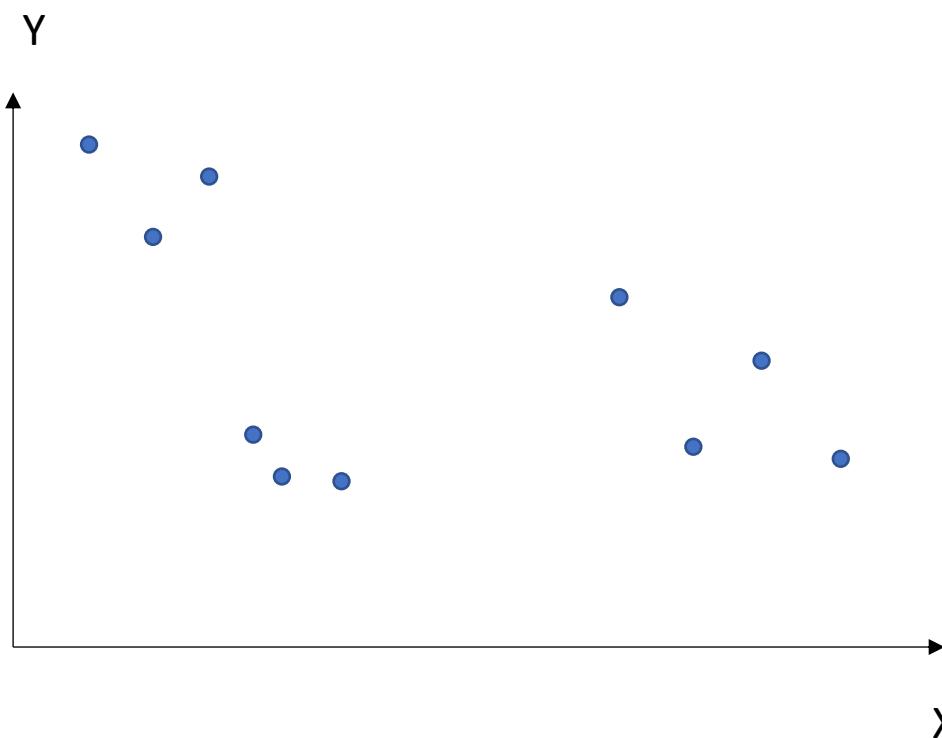


Identificador	Dato útil 1	Dato útil 2
Alfredo	12	7432
Juliana	16	9375

El reto

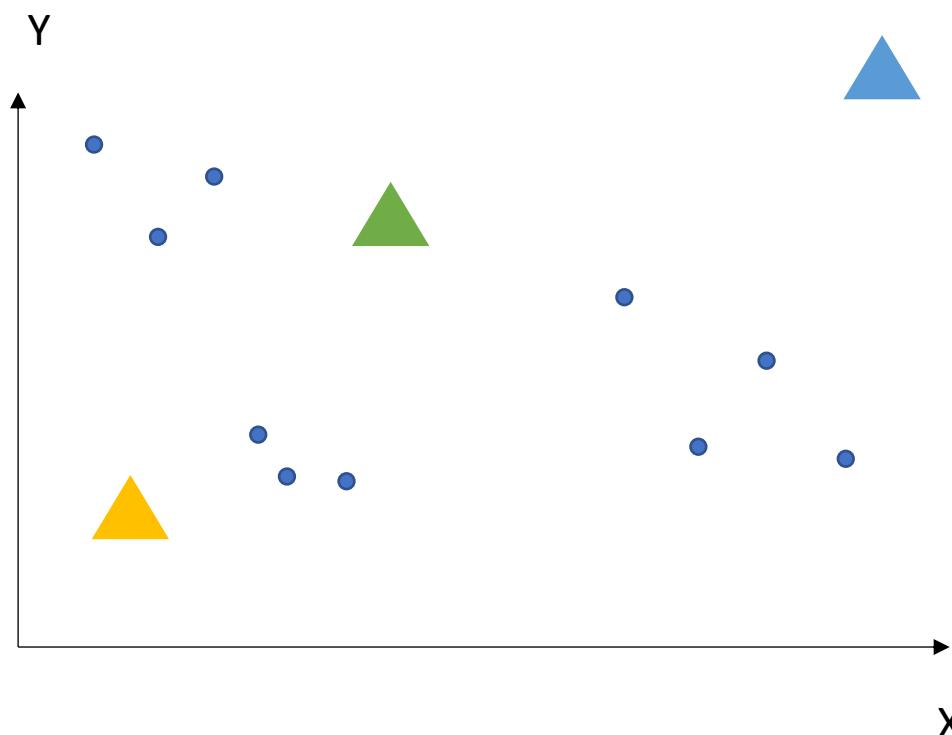


Clustering de K-medias



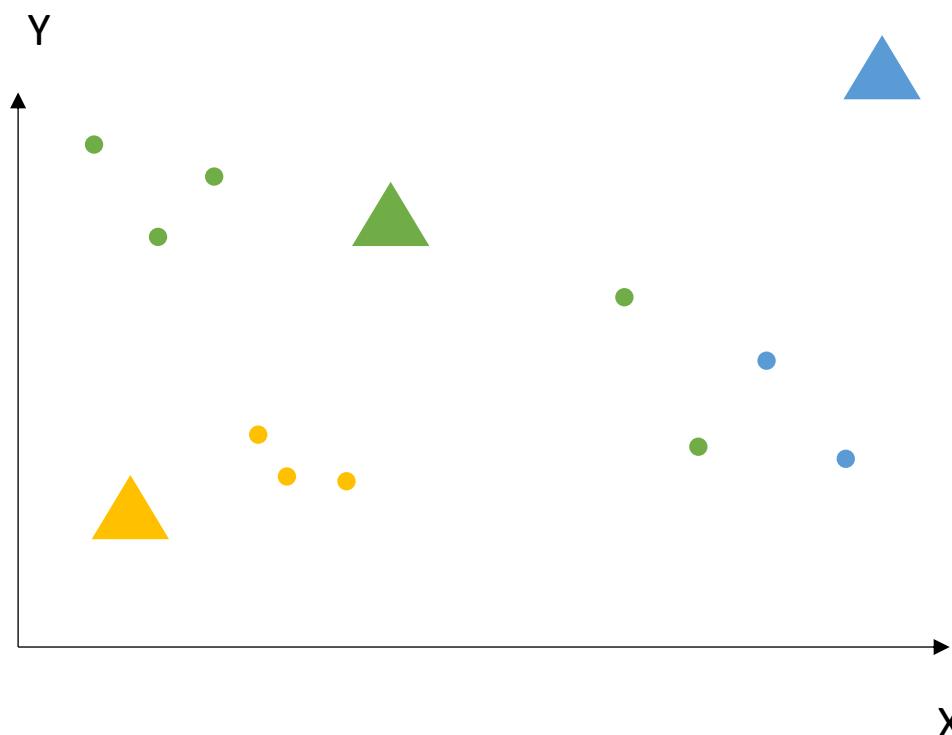
- Digamos que tenemos datos dispersos.
- ¿Cuántos clusters (a ojo) hay?

Clustering de K-medias



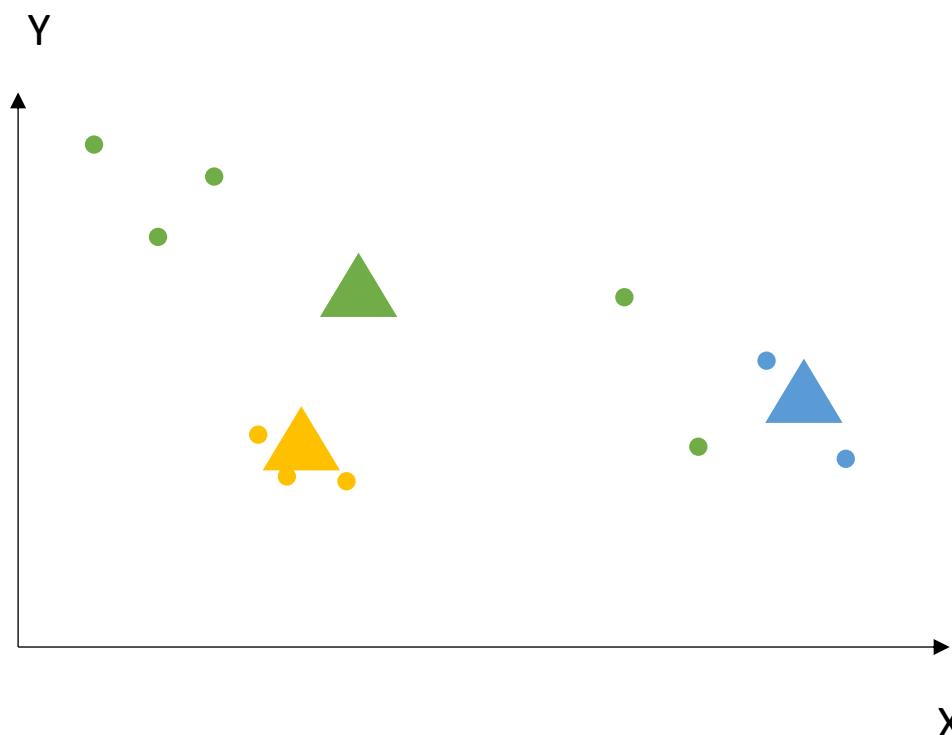
- Digamos que tenemos datos dispersos.
- ¿Cuántos clusters (a ojo) hay?
- Ponemos tres “centros” en lugares aleatorios

Clustering de K-medias



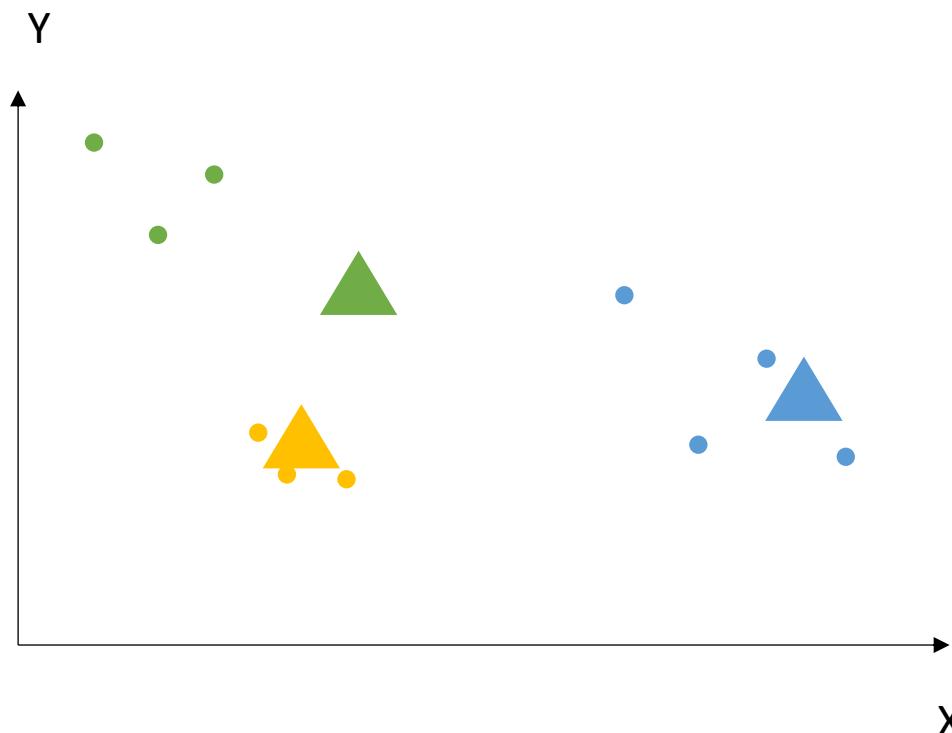
- Digamos que tenemos datos dispersos.
- ¿Cuántos clusters (a ojo) hay?
- Ponemos tres “centros” en lugares aleatorios
- Identificamos a qué centro está mas cerca cada dato

Clustering de K-medias



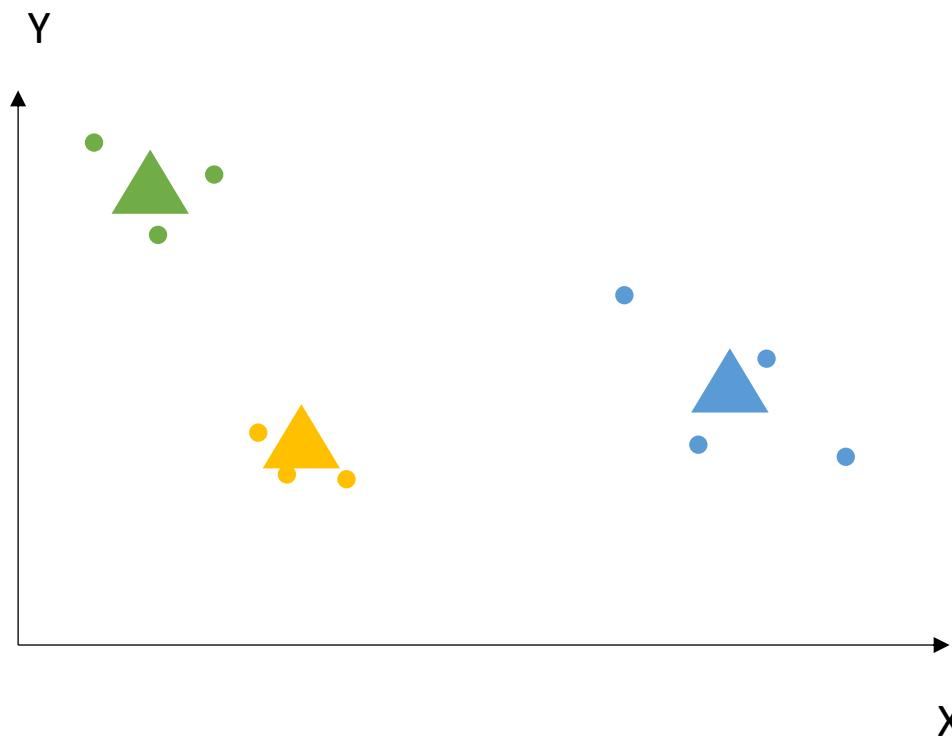
- Digamos que tenemos datos dispersos.
- ¿Cuántos clusters (a ojo) hay?
- Ponemos tres “centros” en lugares aleatorios
- Identificamos a qué centro está mas cerca cada dato
- Ubicamos los centros en el centro de sus grupos

Clustering de K-medias



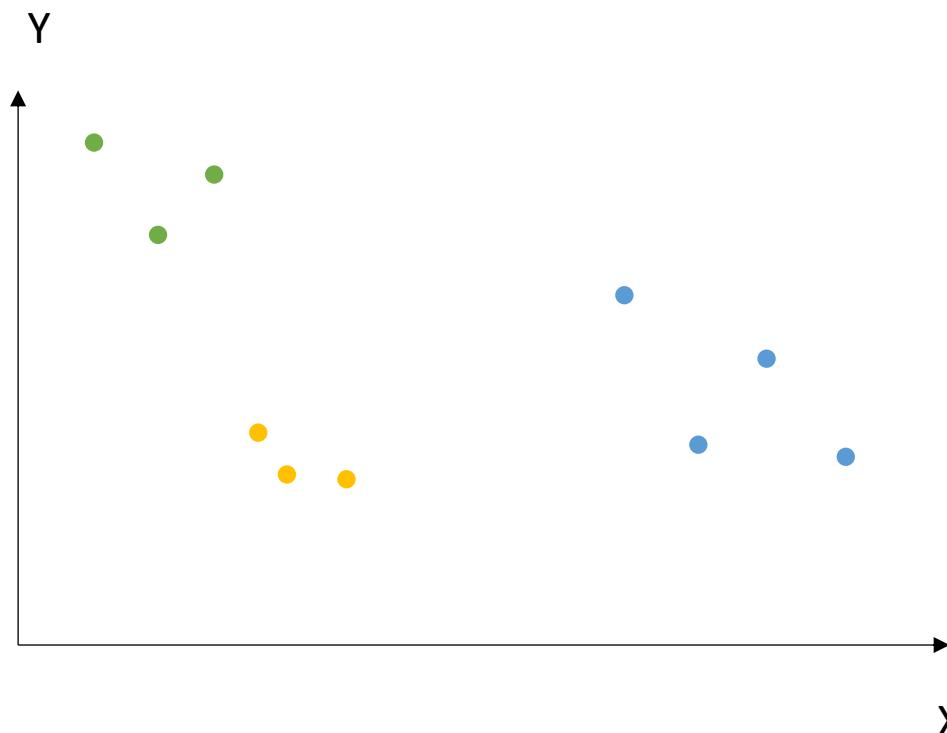
- Digamos que tenemos datos dispersos.
- ¿Cuántos clusters (a ojo) hay?
- Ponemos tres “centros” en lugares aleatorios
- Identificamos a qué centro está mas cerca cada dato
- Ubicamos los centros en el centro de sus grupos
- Identificamos cercanías de nuevo

Clustering de K-medias



- Digamos que tenemos datos dispersos.
- ¿Cuántos clusters (a ojo) hay?
- Ponemos tres “centros” en lugares aleatorios
- Identificamos a qué centro está mas cerca cada dato
- Ubicamos los centros en el centro de sus grupos
- Identificamos cercanías de nuevo

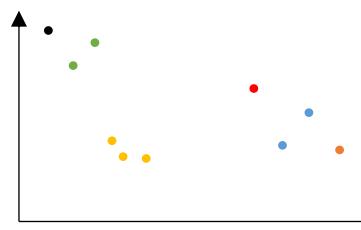
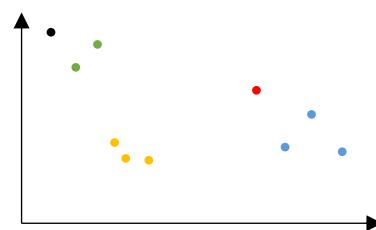
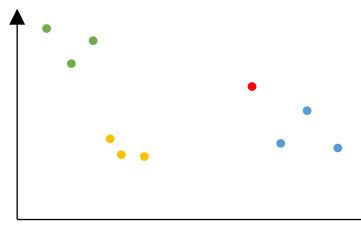
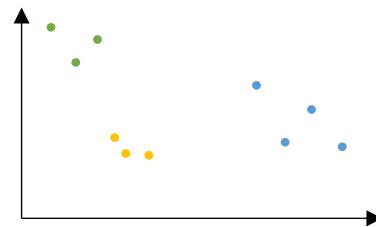
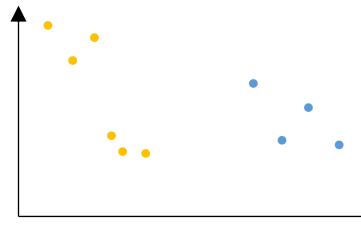
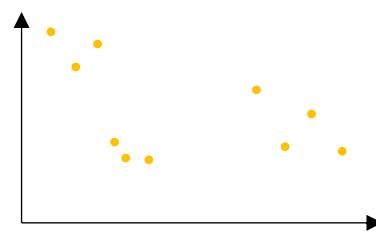
Clustering de K-medias



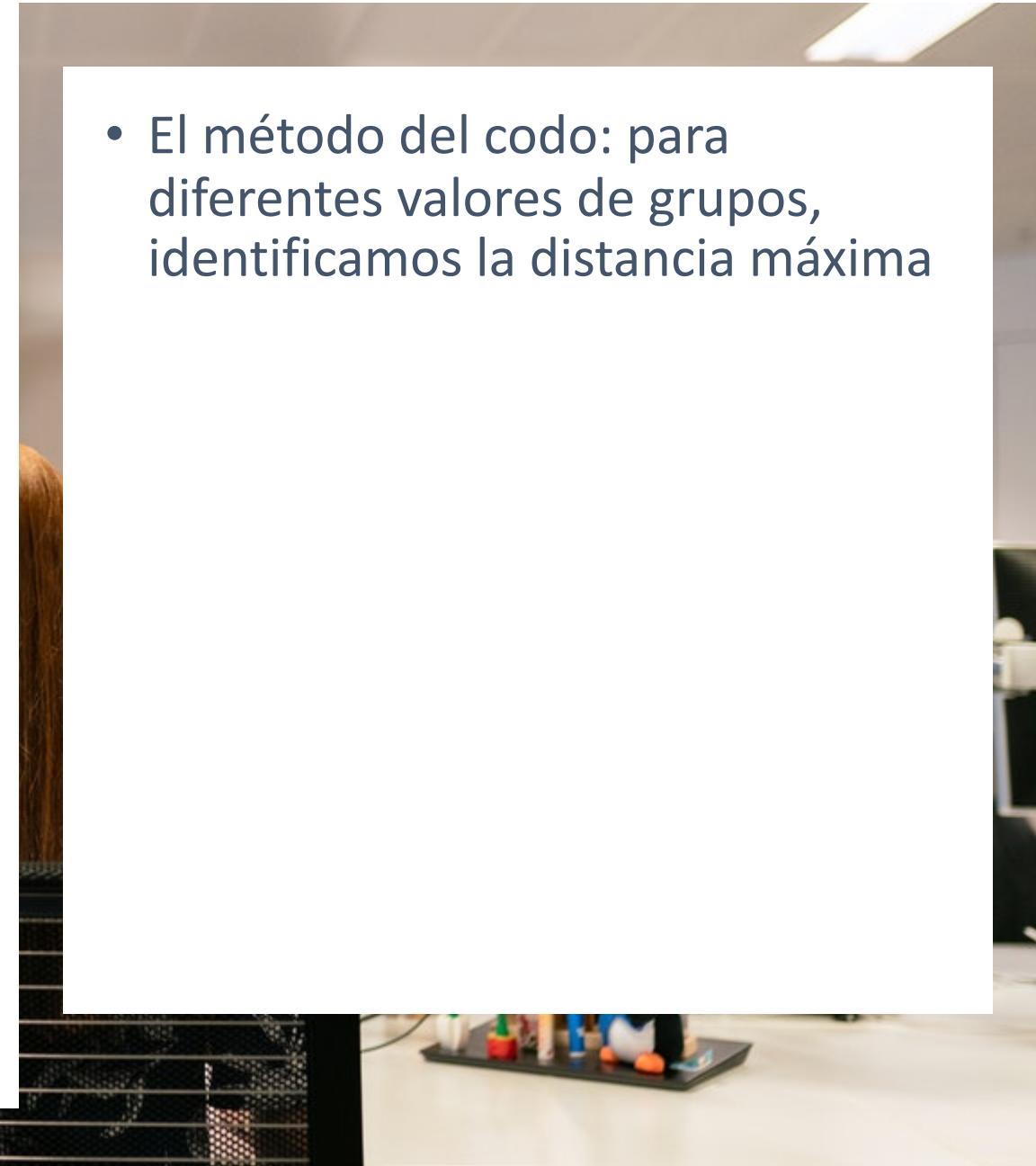
- Digamos que tenemos datos dispersos.
- ¿Cuántos clusters (a ojo) hay?
- Ponemos tres “centros” en lugares aleatorios
- Identificamos a qué centro está mas cerca cada dato
- Ubicamos los centros en el centro de sus grupos
- Identificamos cercanías de nuevo

¿Pero cuántos grupos escogemos?

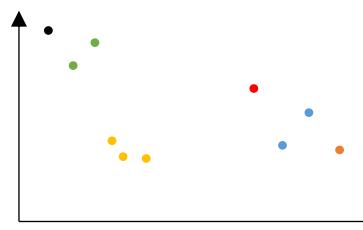
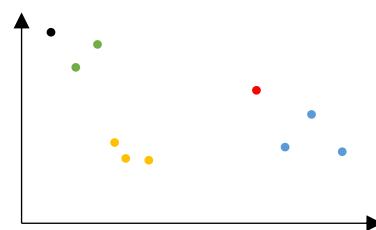
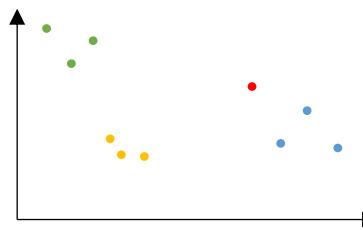
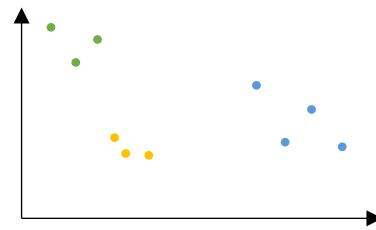
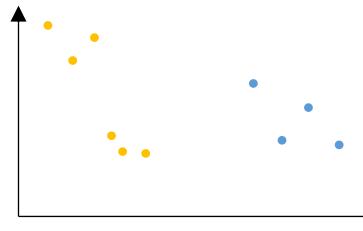
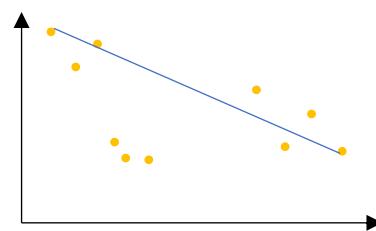
Clustering de K-medias



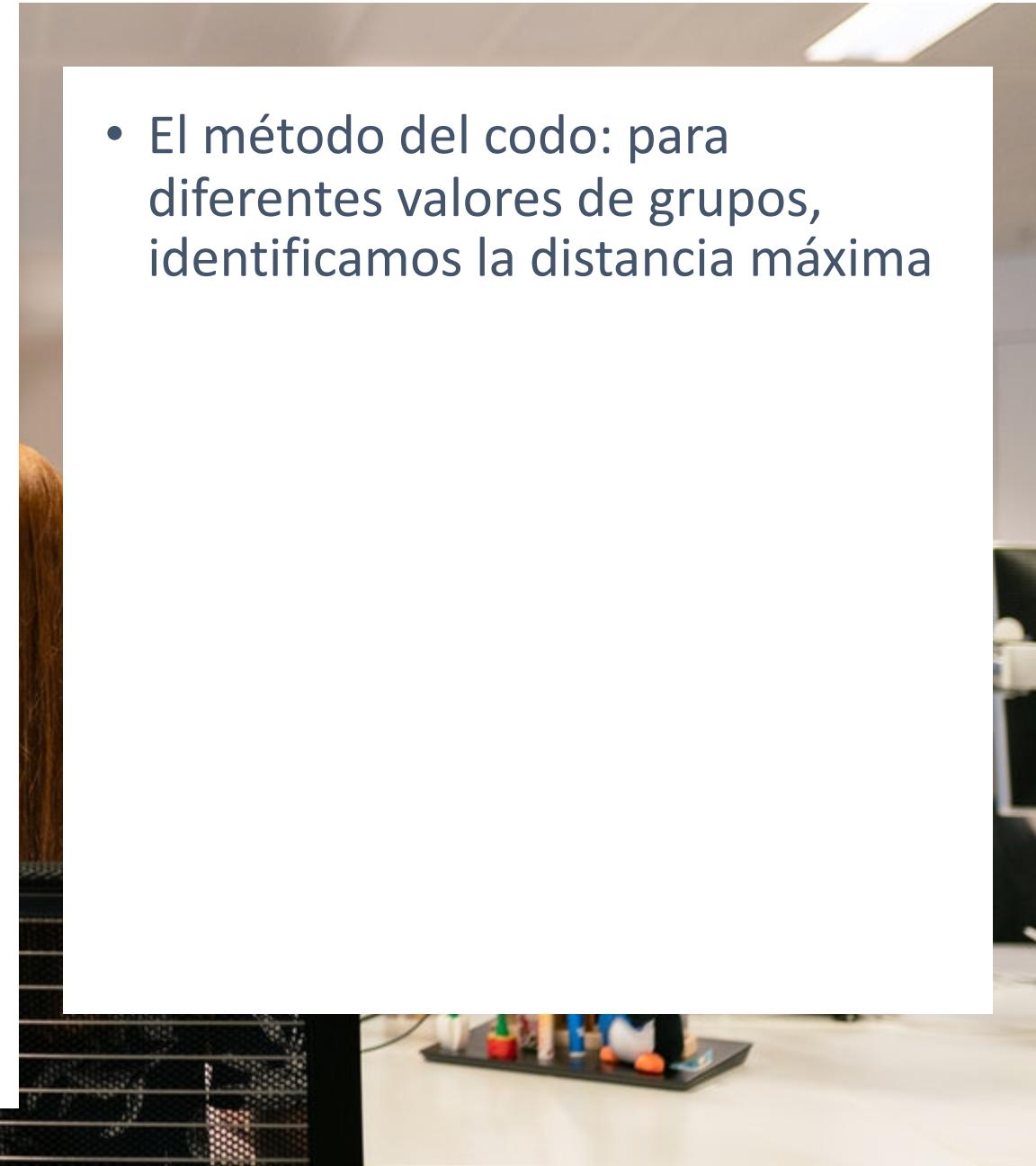
- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima



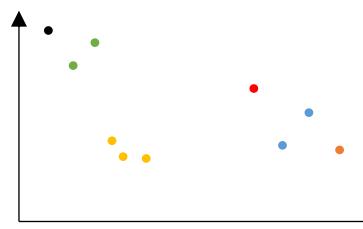
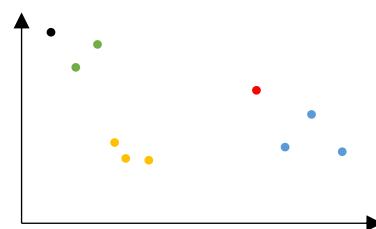
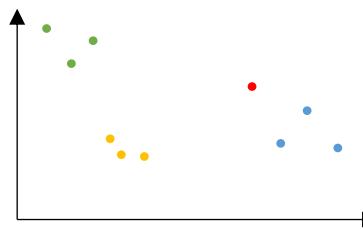
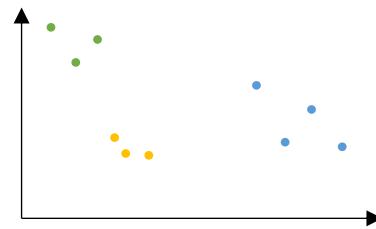
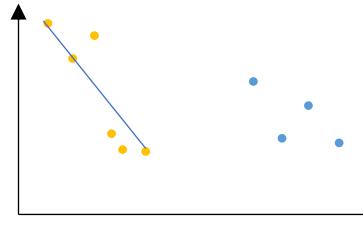
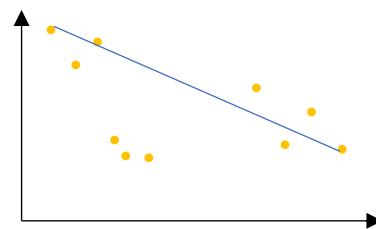
Clustering de K-medias



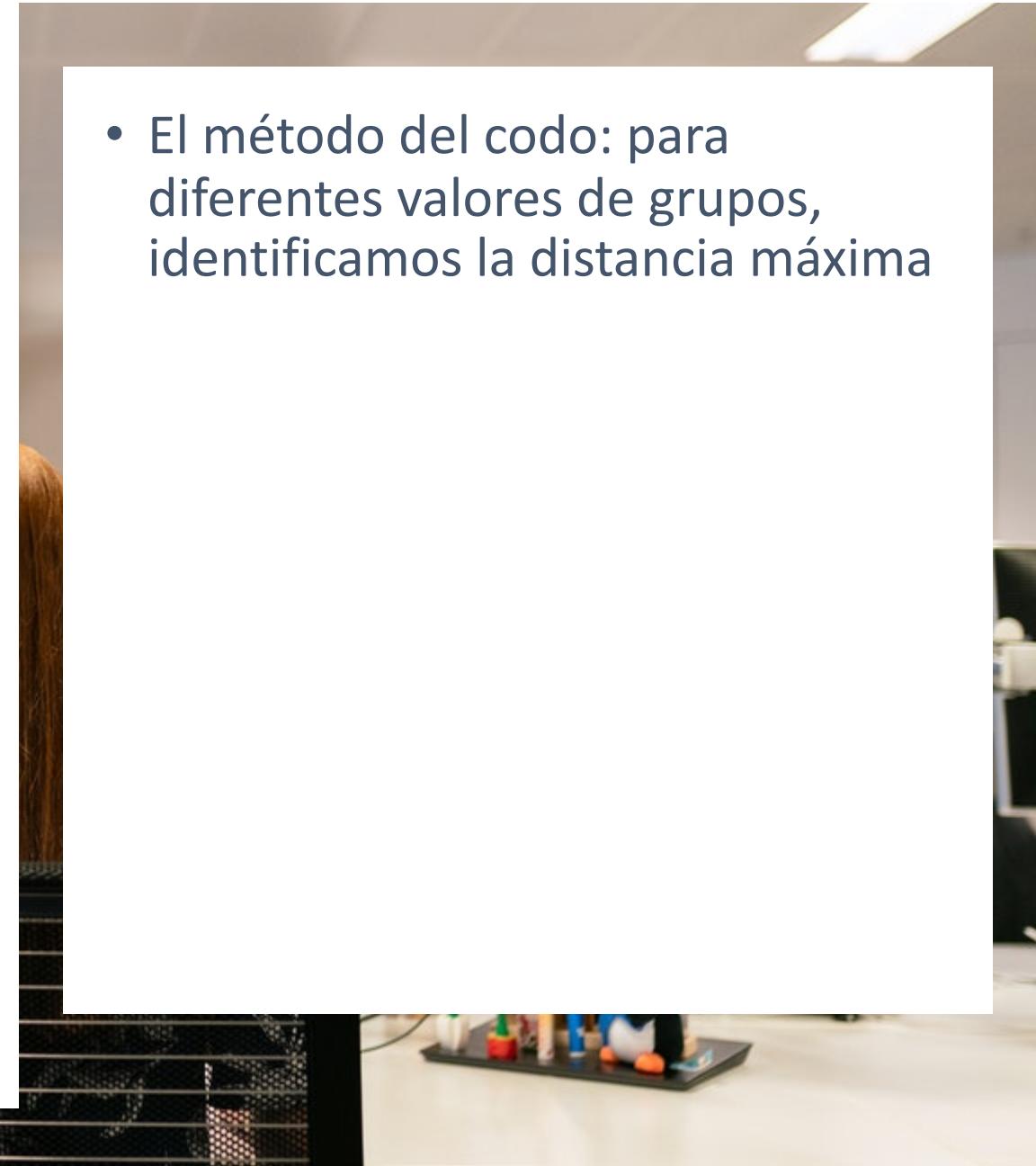
- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima



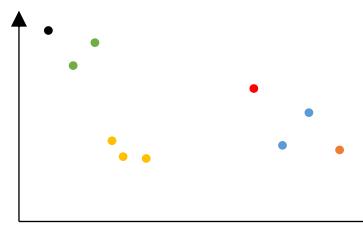
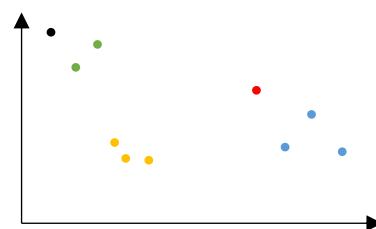
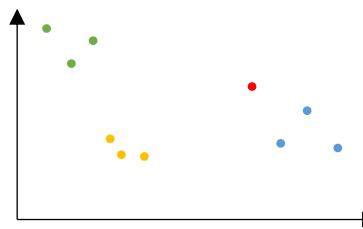
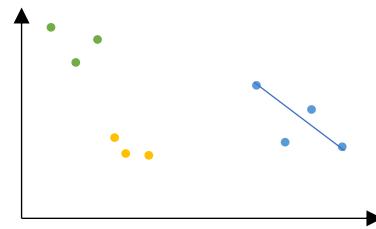
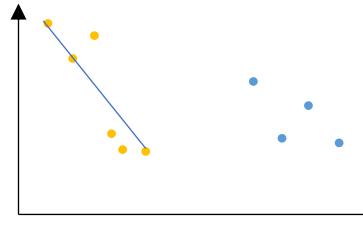
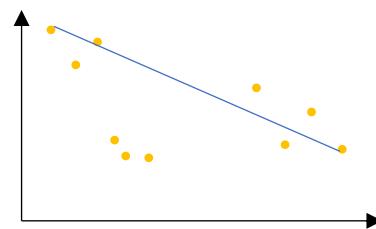
Clustering de K-medias



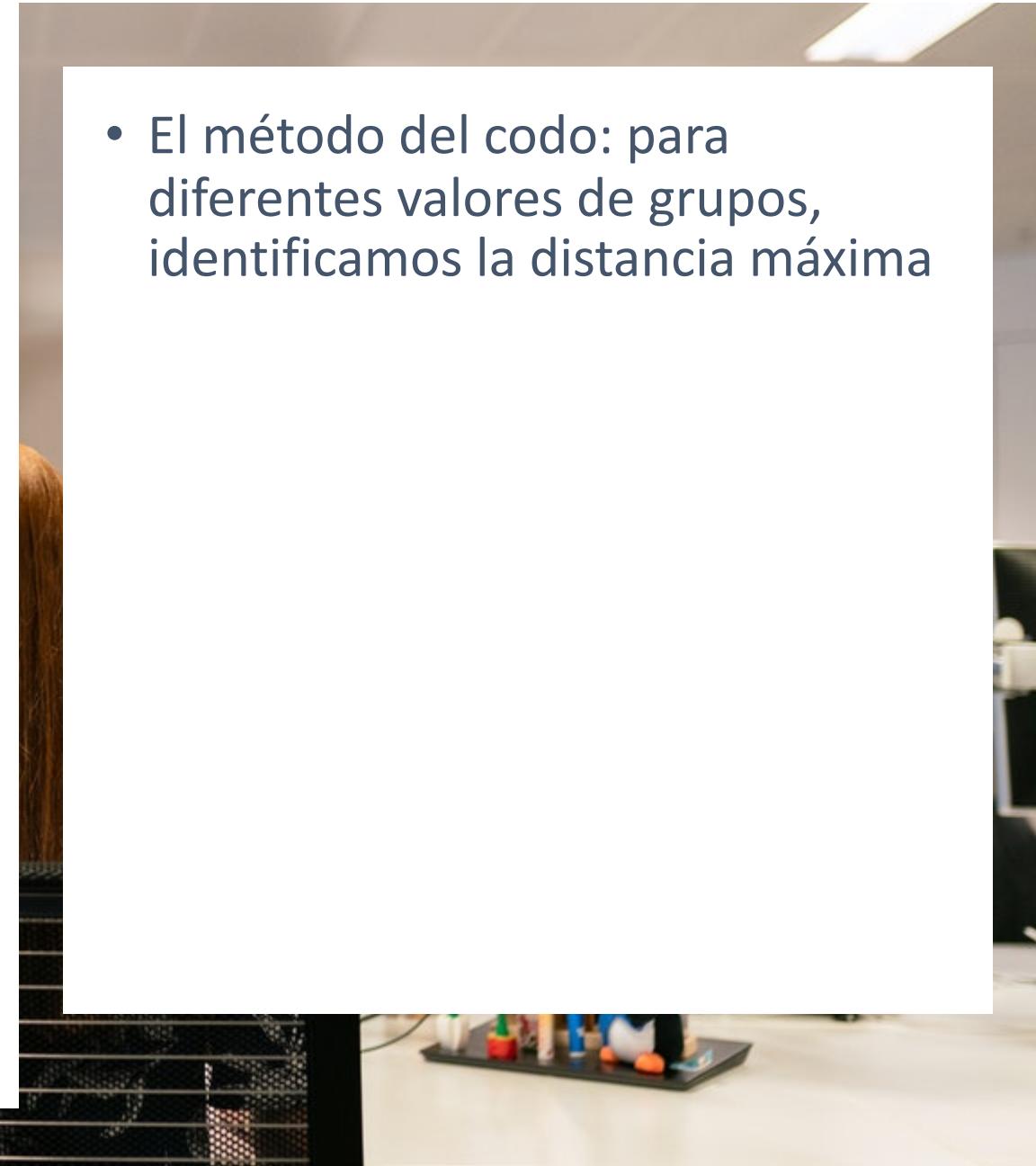
- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima



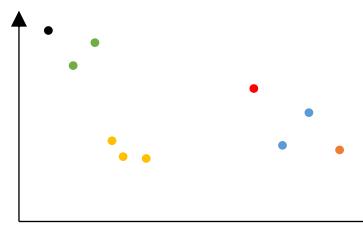
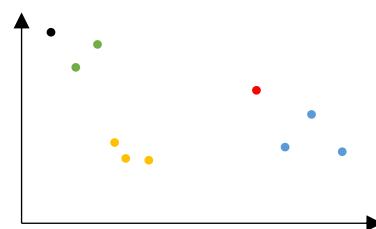
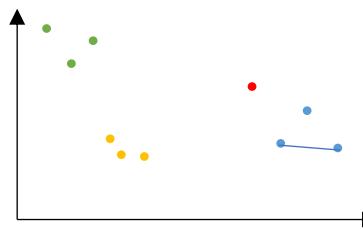
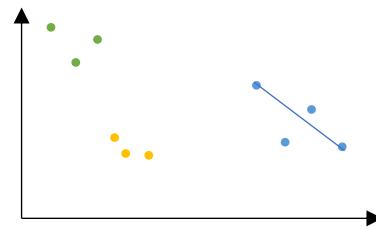
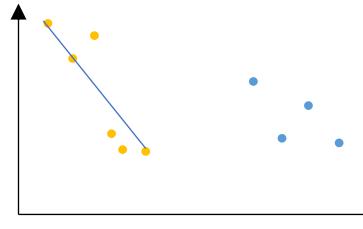
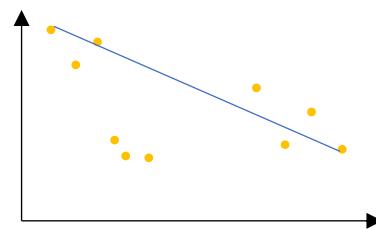
Clustering de K-medias



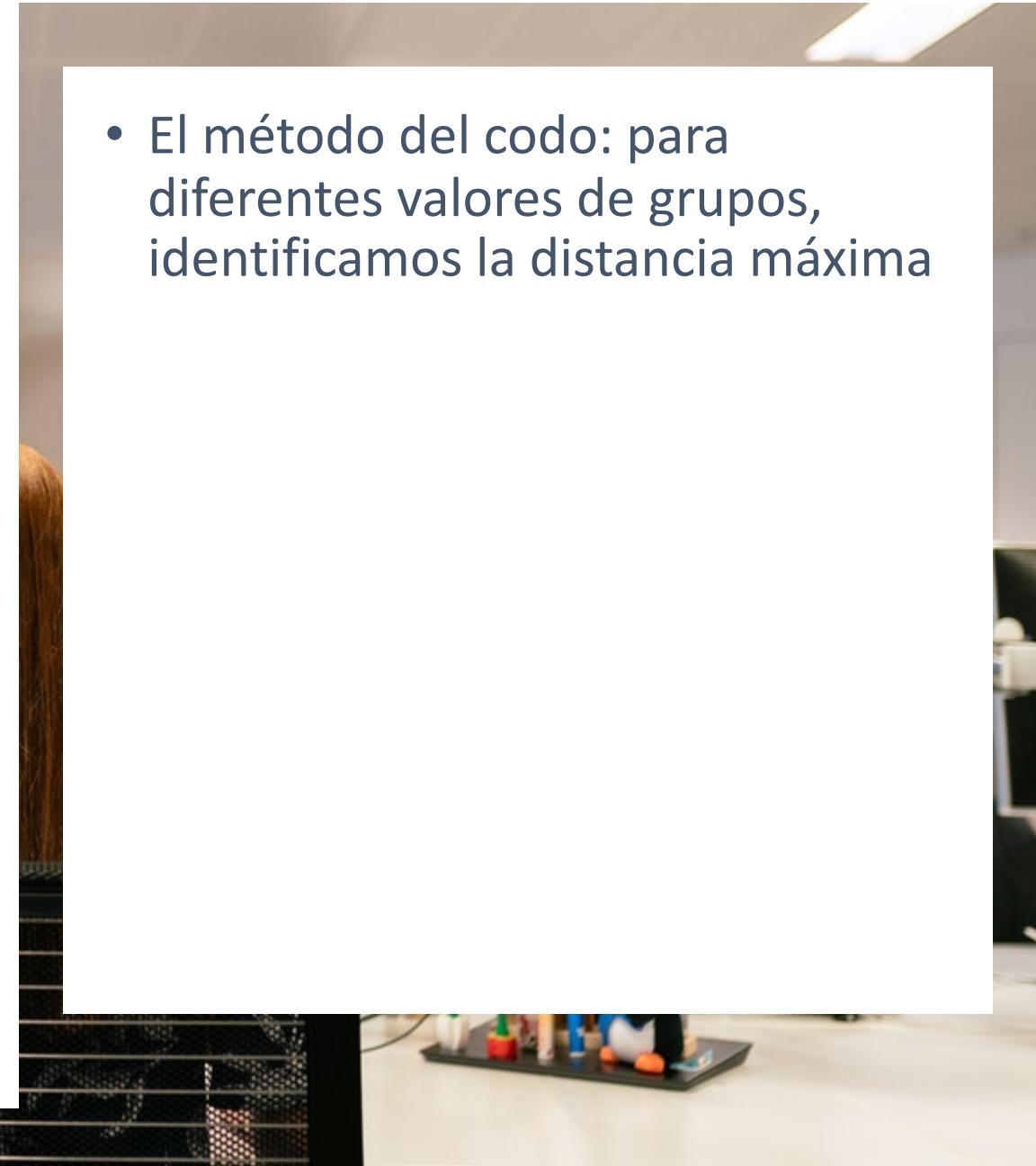
- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima



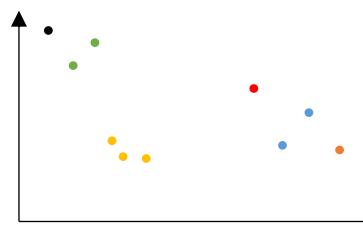
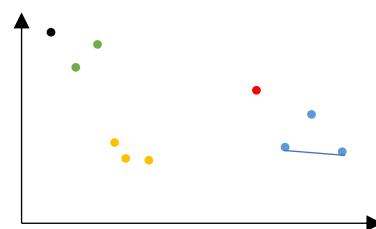
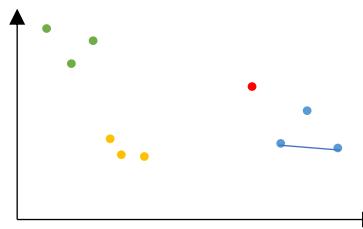
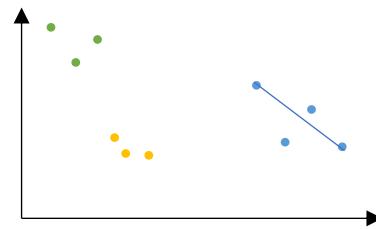
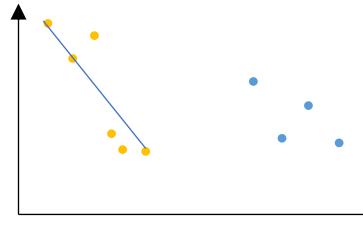
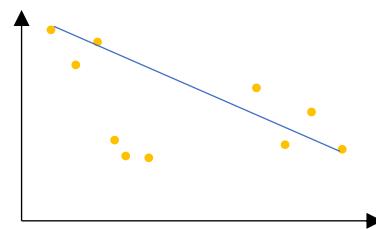
Clustering de K-medias



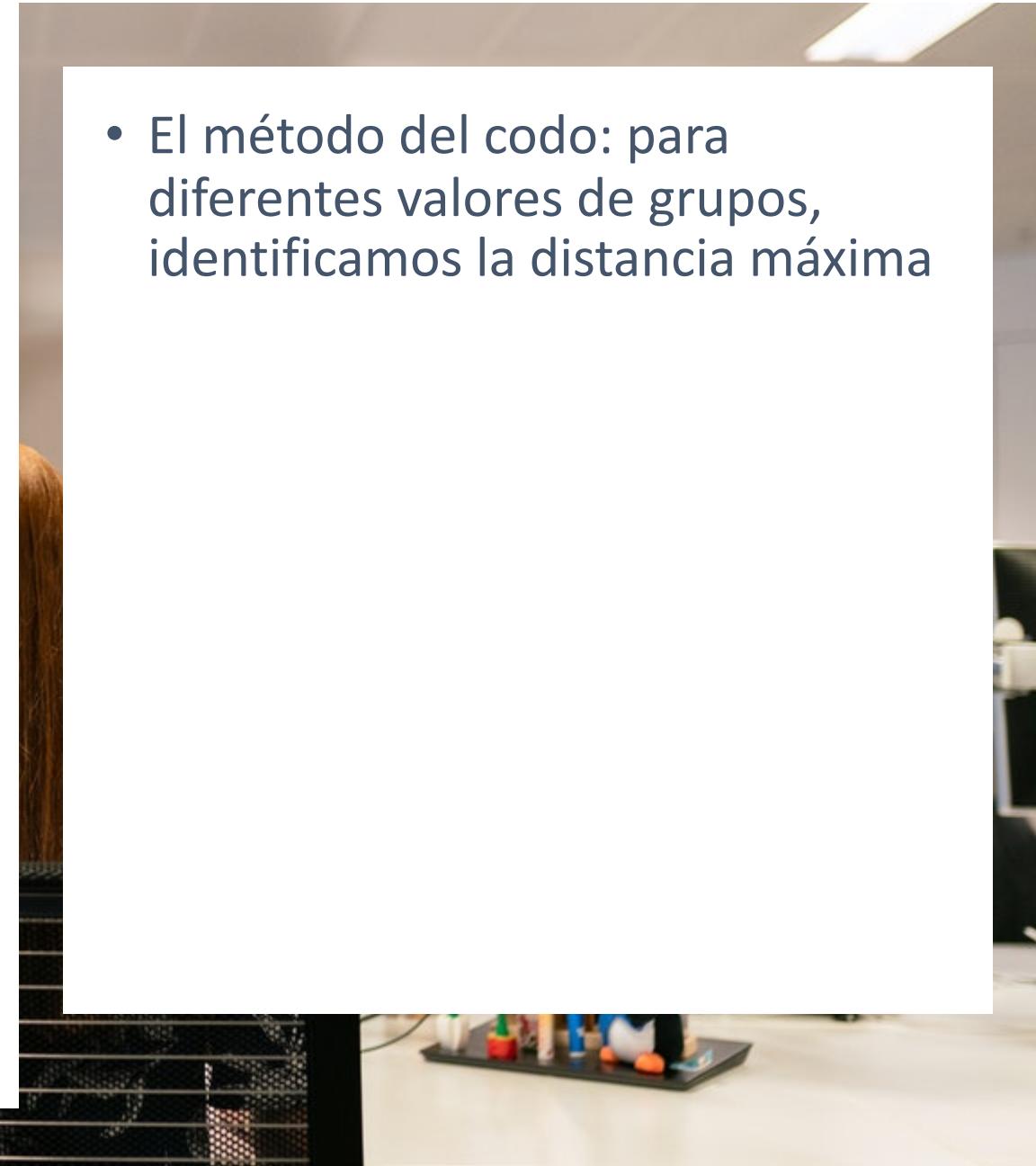
- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima



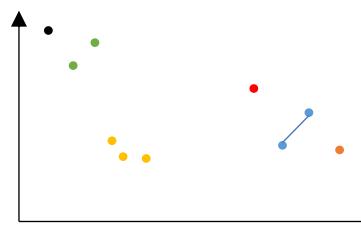
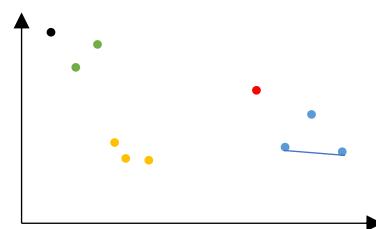
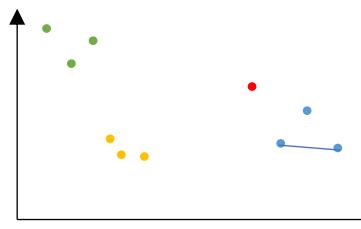
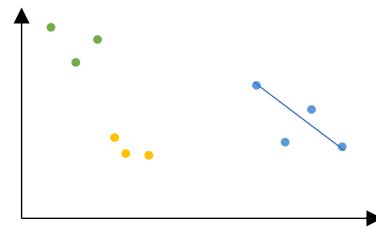
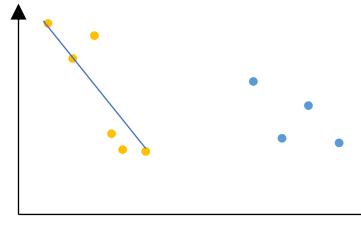
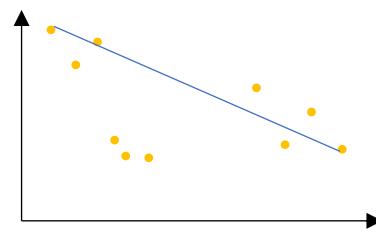
Clustering de K-medias



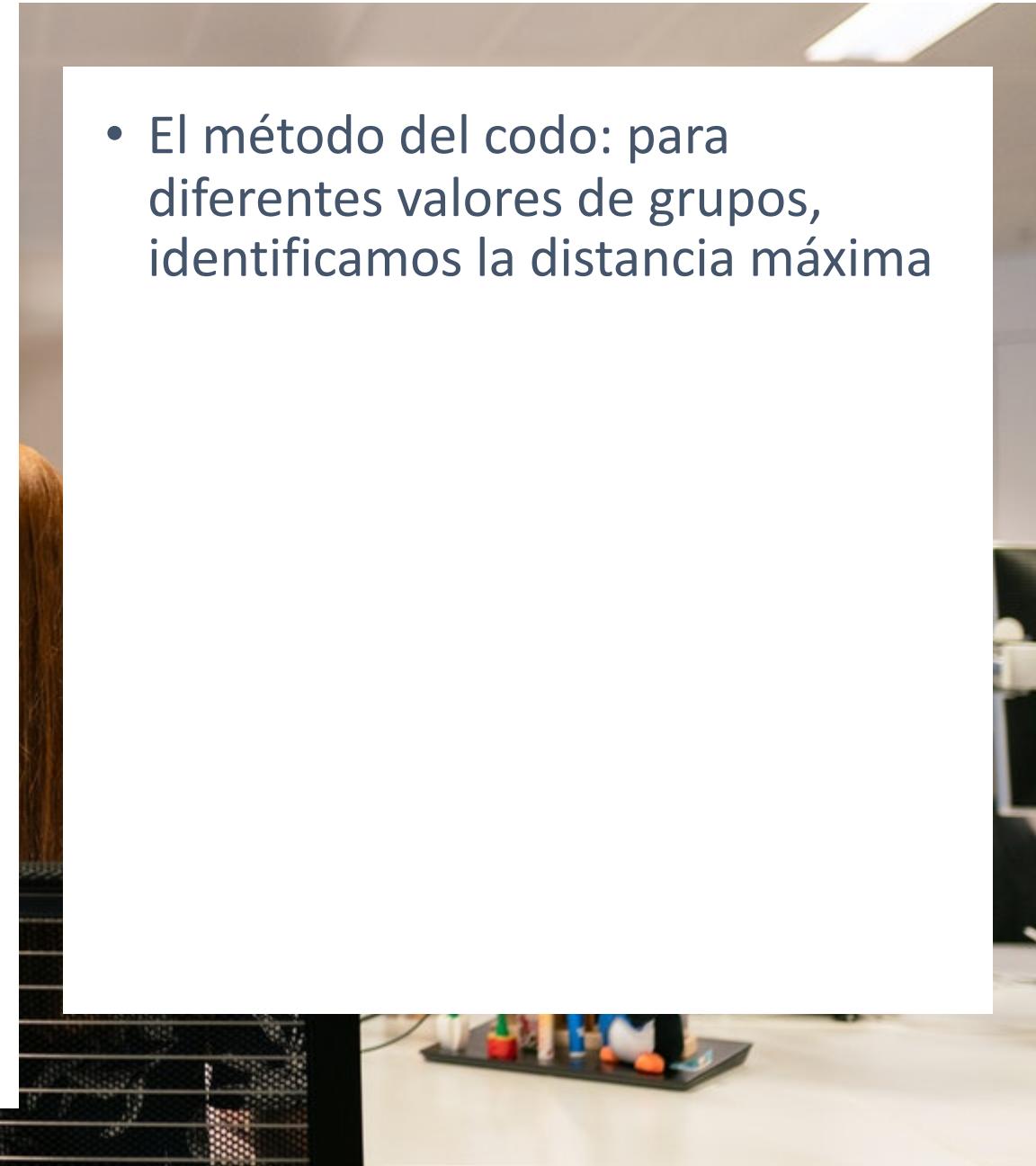
- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima



Clustering de K-medias

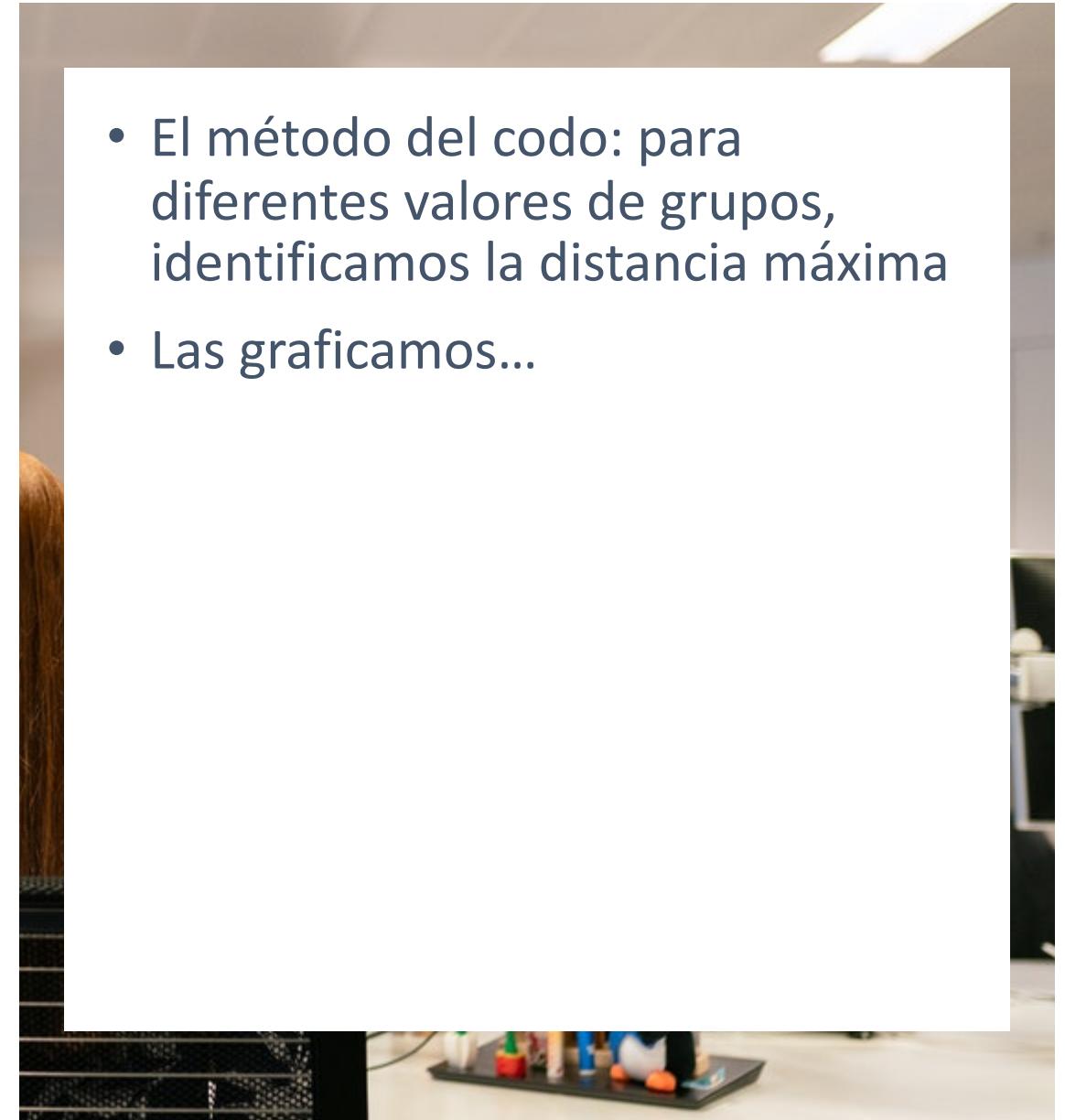


- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima

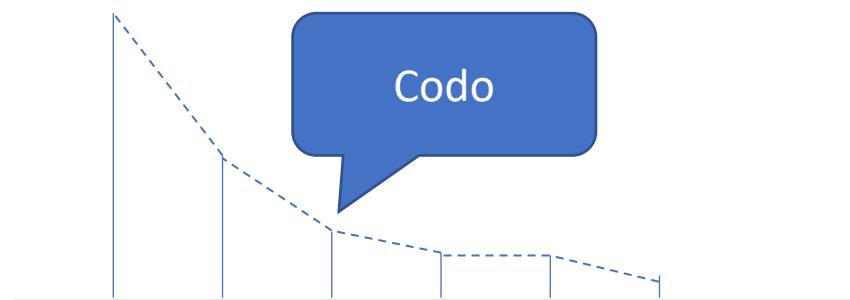


Clustering de K-medias

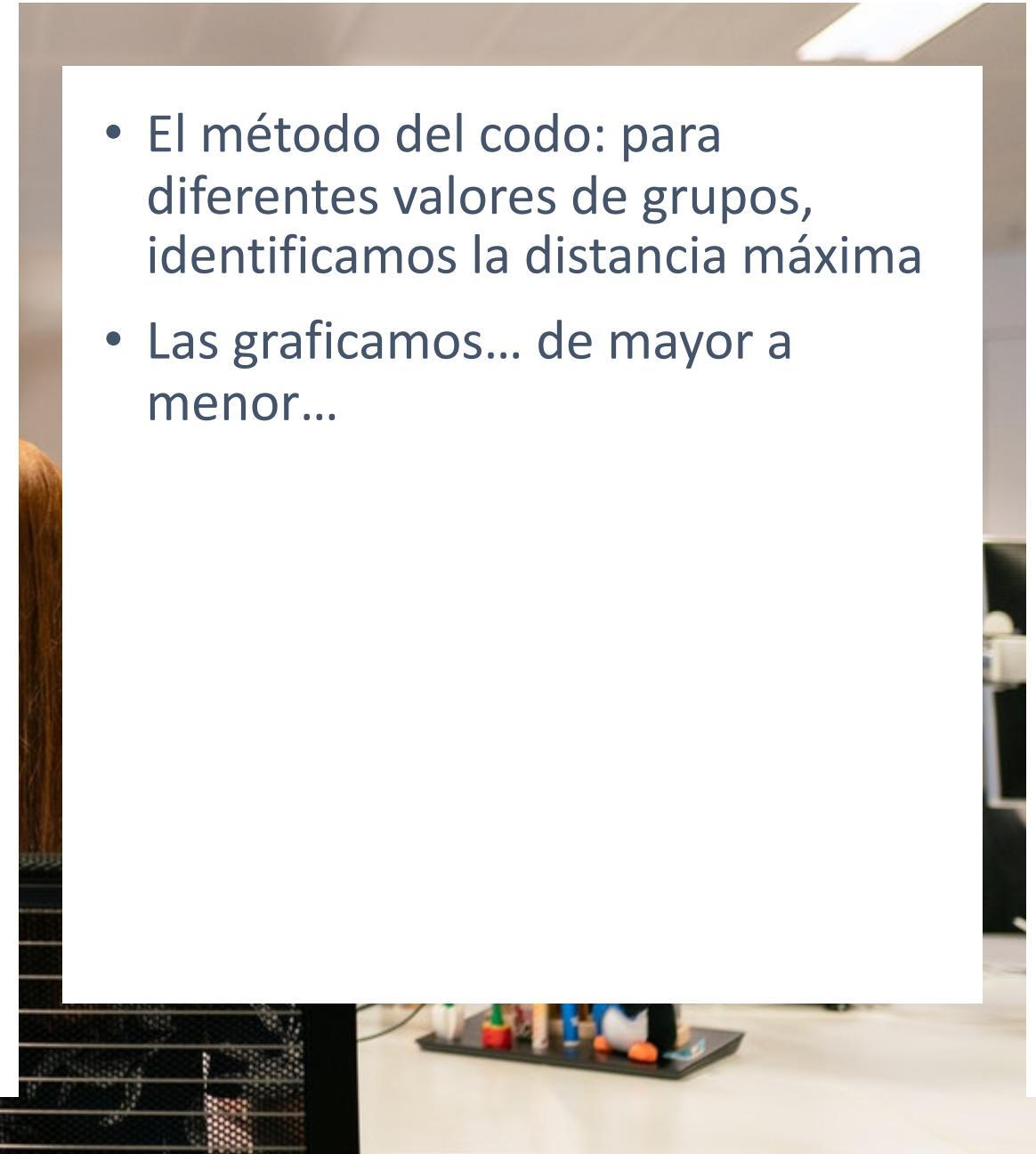
- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima
- Las graficamos...



Clustering de K-medias



- El método del codo: para diferentes valores de grupos, identificamos la distancia máxima
- Las graficamos... de mayor a menor...



Clustering

Clustering jerárquico

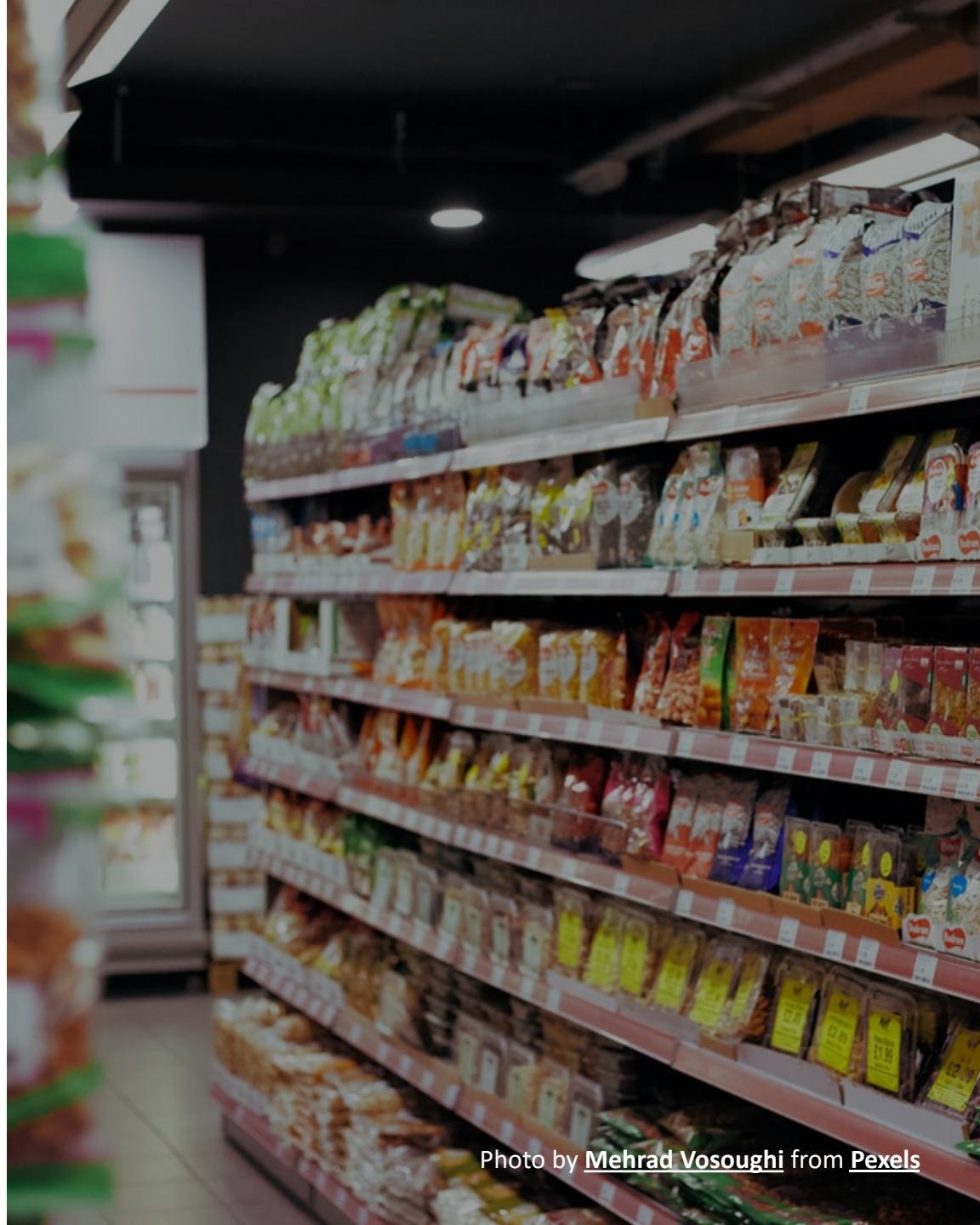
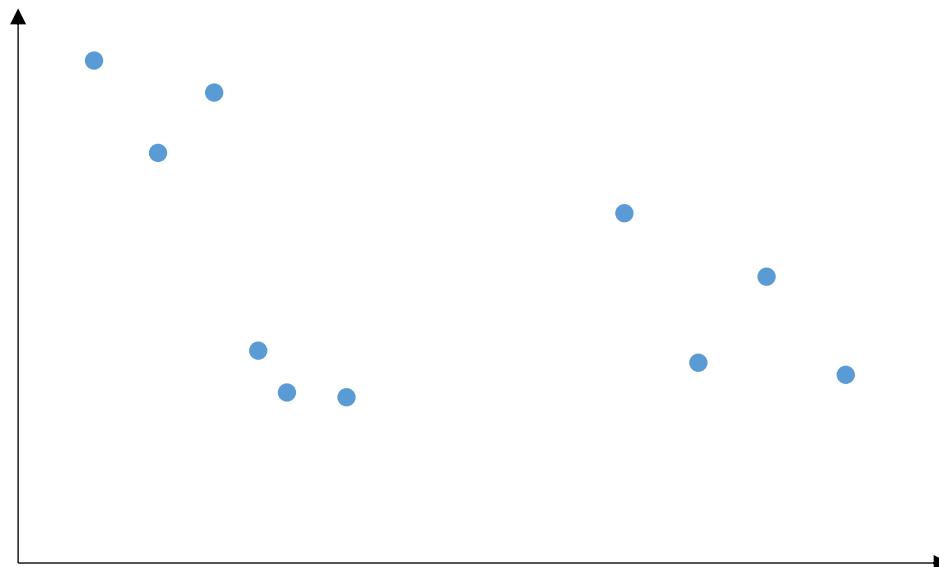
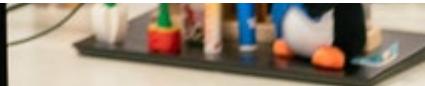


Photo by [Mehrad Vosoughi](#) from [Pexels](#)

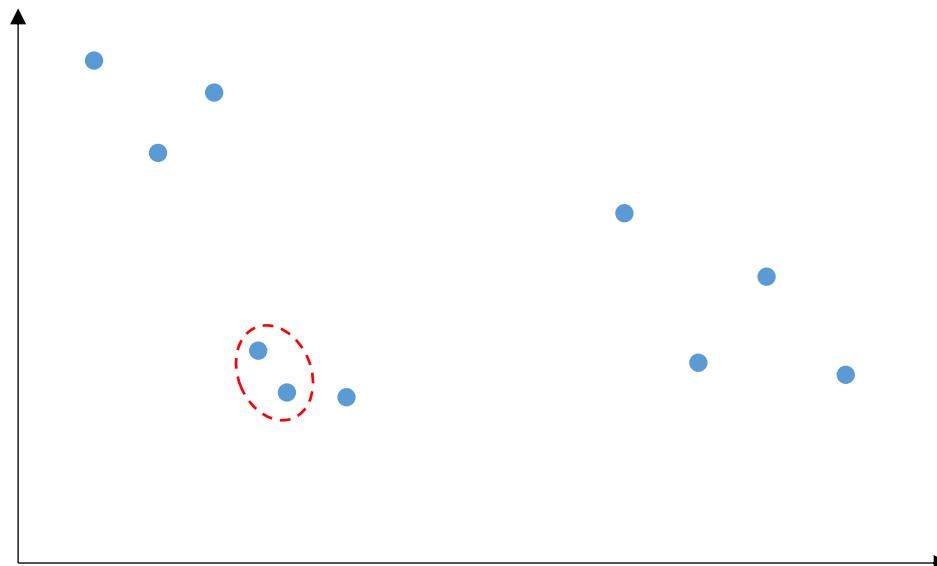
Clustering Jerárquico



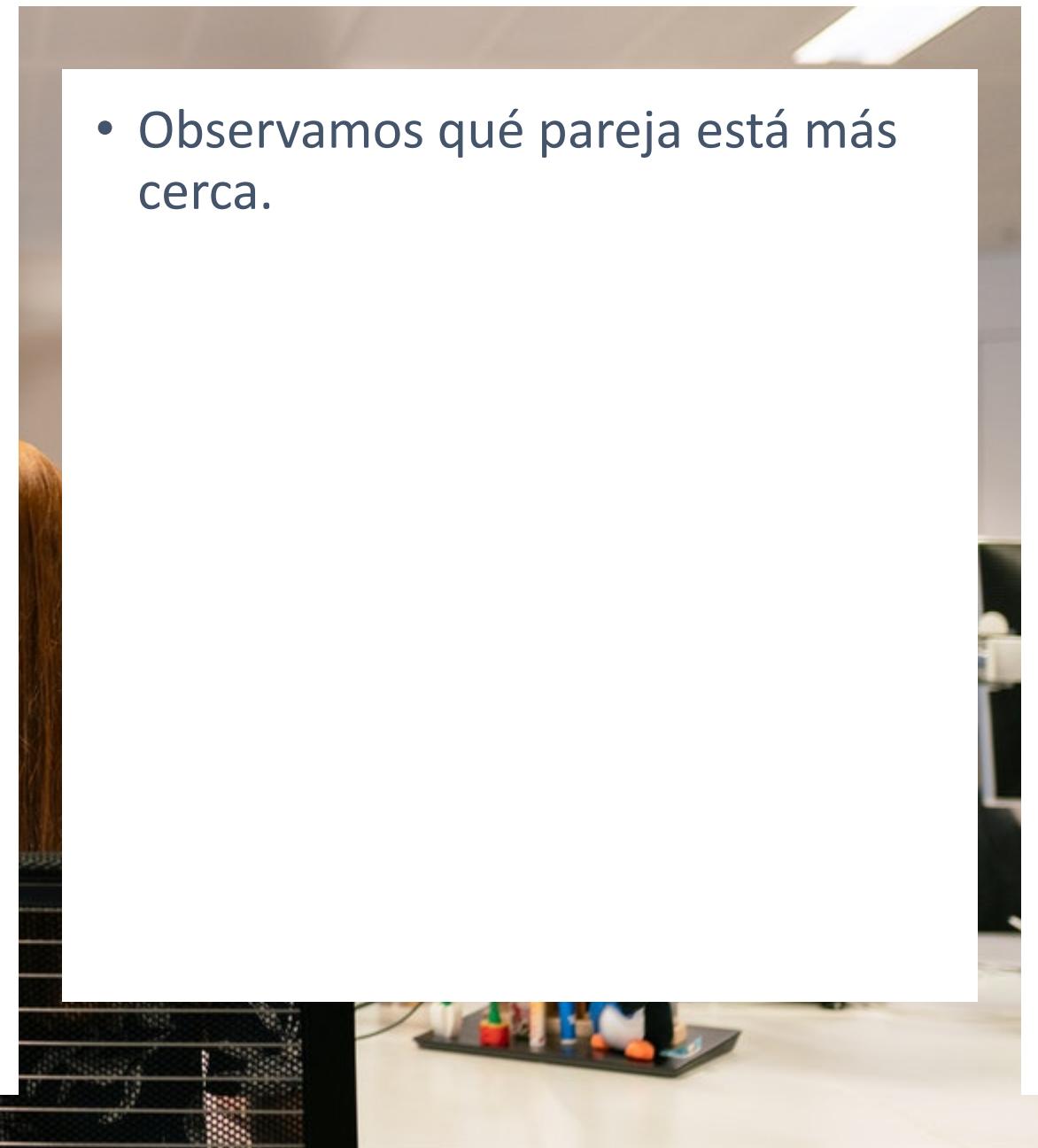
- Observamos qué pareja está más cerca.



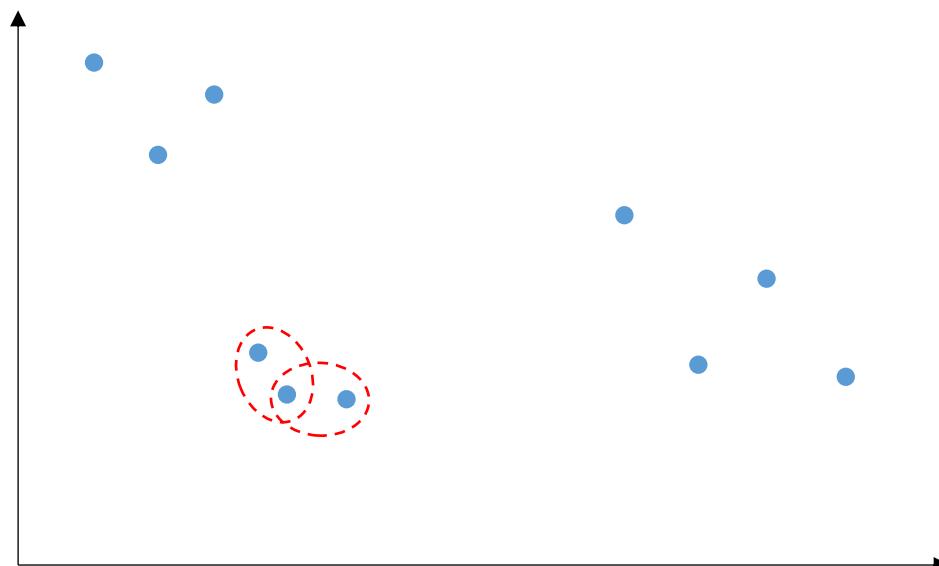
Clustering Jerárquico



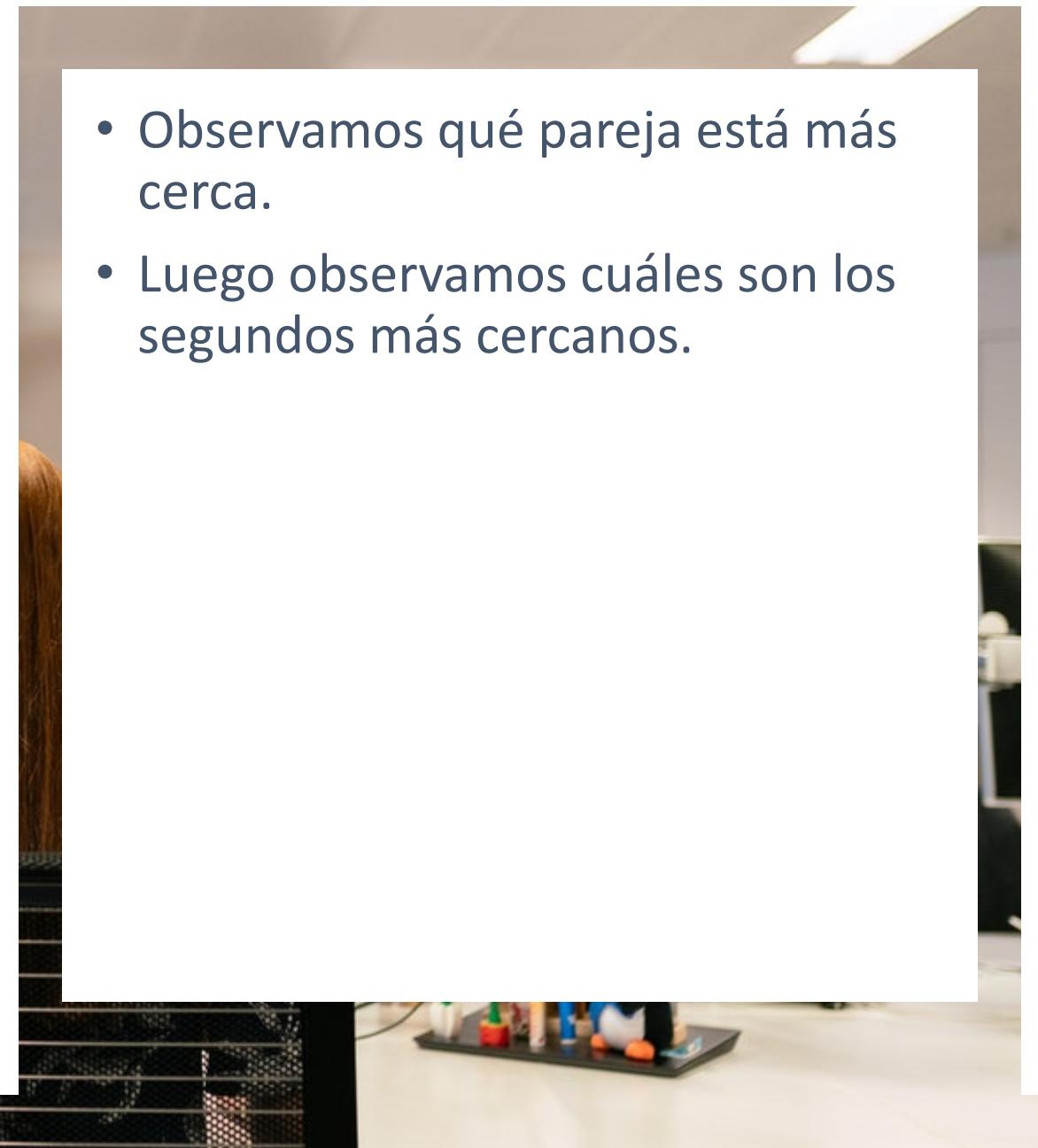
- Observamos qué pareja está más cerca.



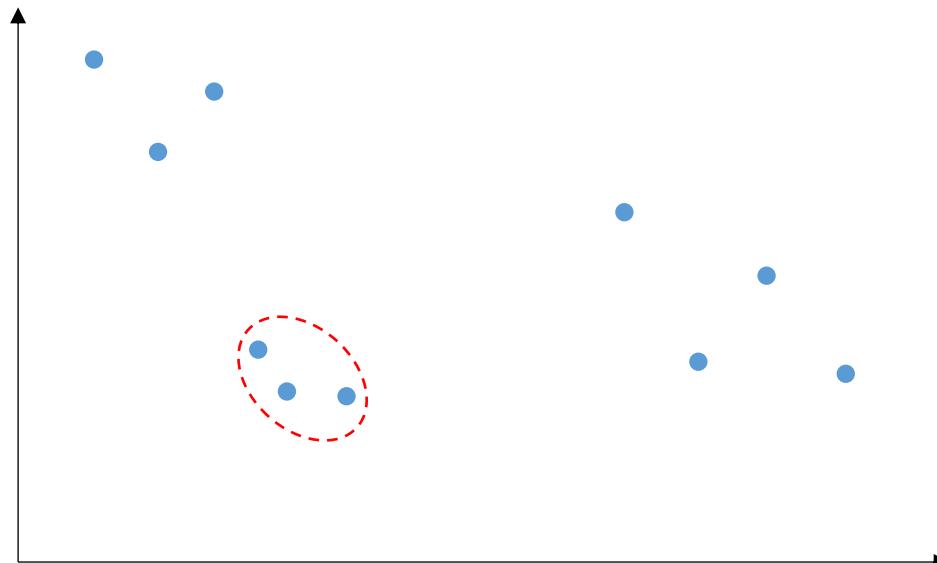
Clustering Jerárquico



- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.

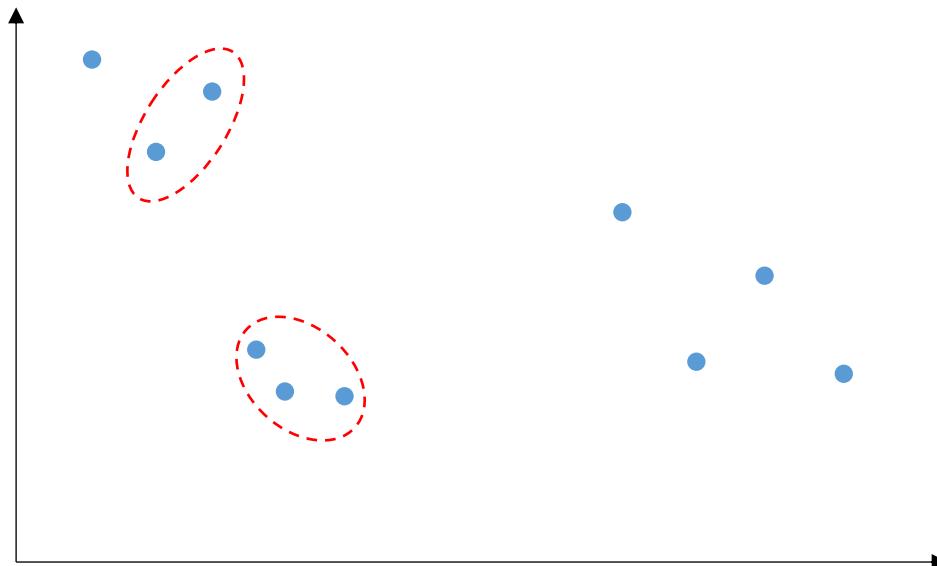


Clustering Jerárquico



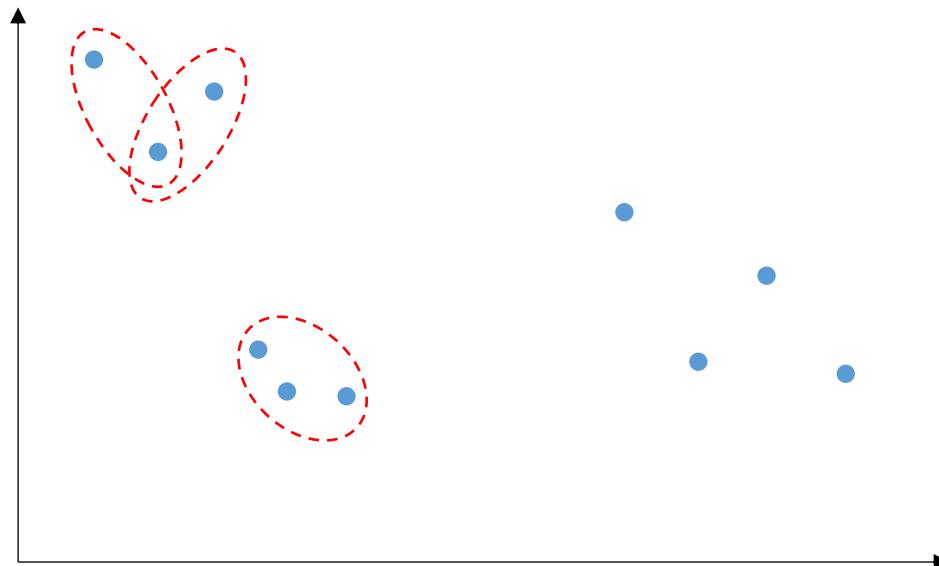
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.

Clustering Jerárquico



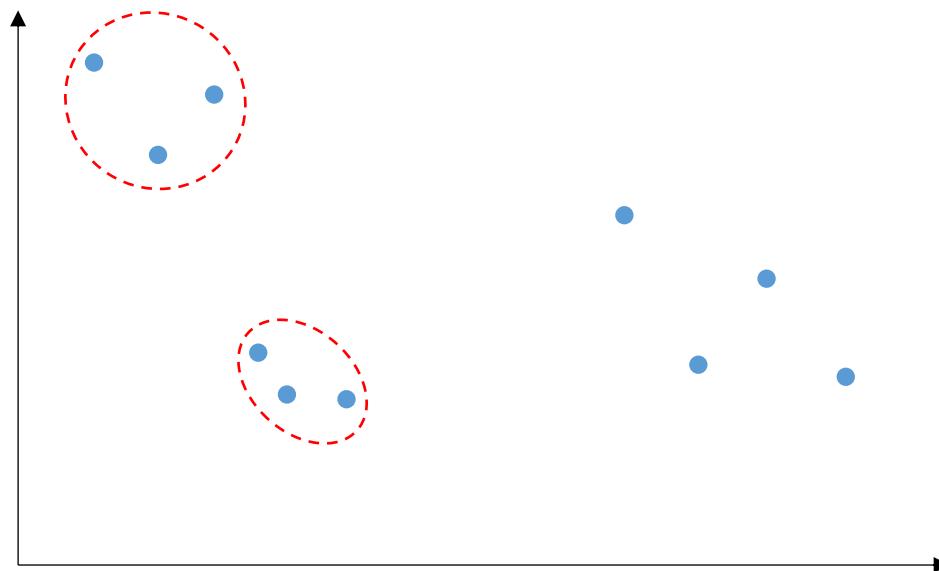
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.

Clustering Jerárquico



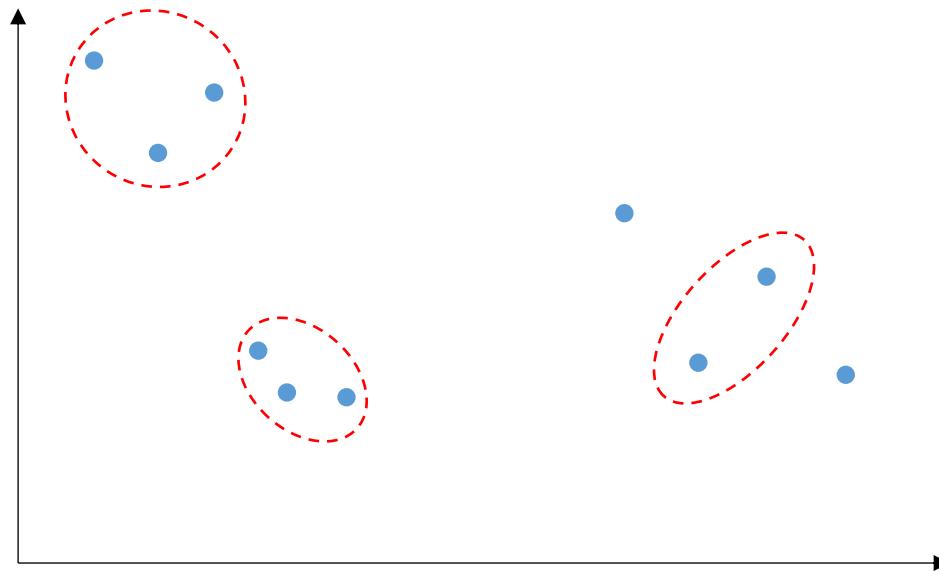
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.

Clustering Jerárquico



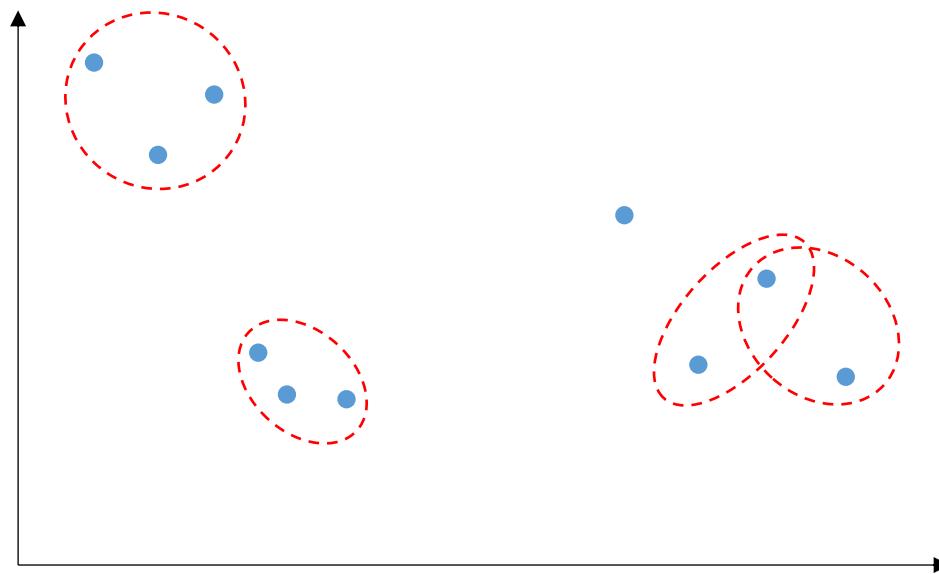
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

Clustering Jerárquico



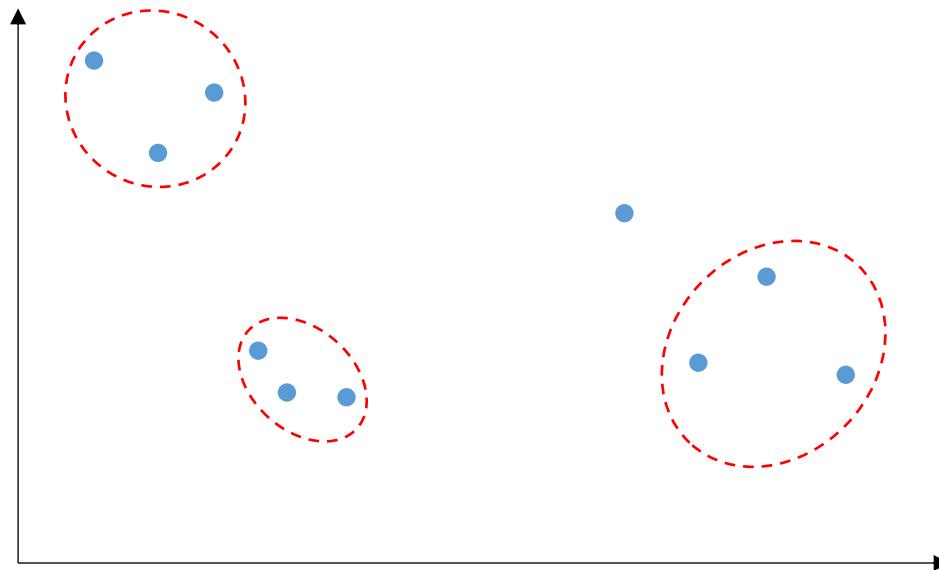
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

Clustering Jerárquico



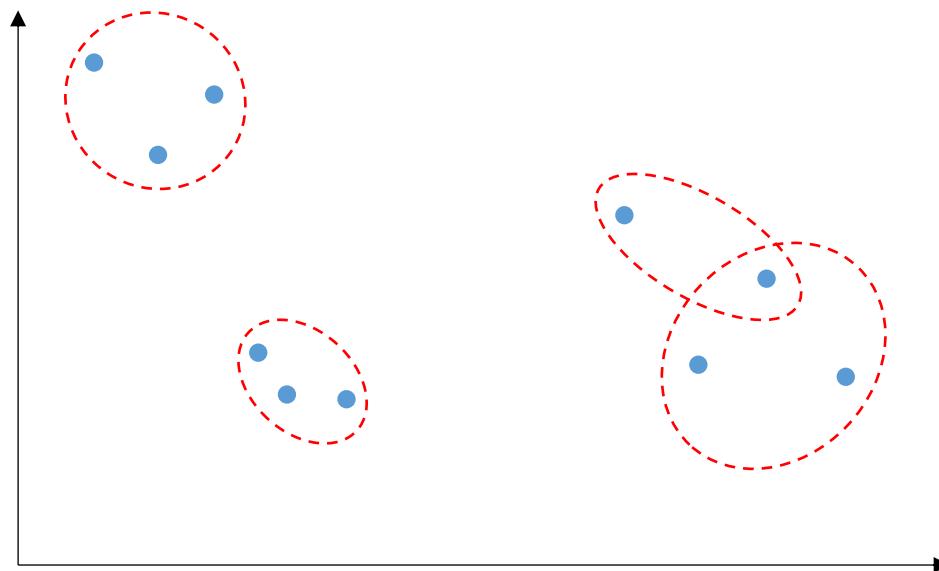
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

Clustering Jerárquico

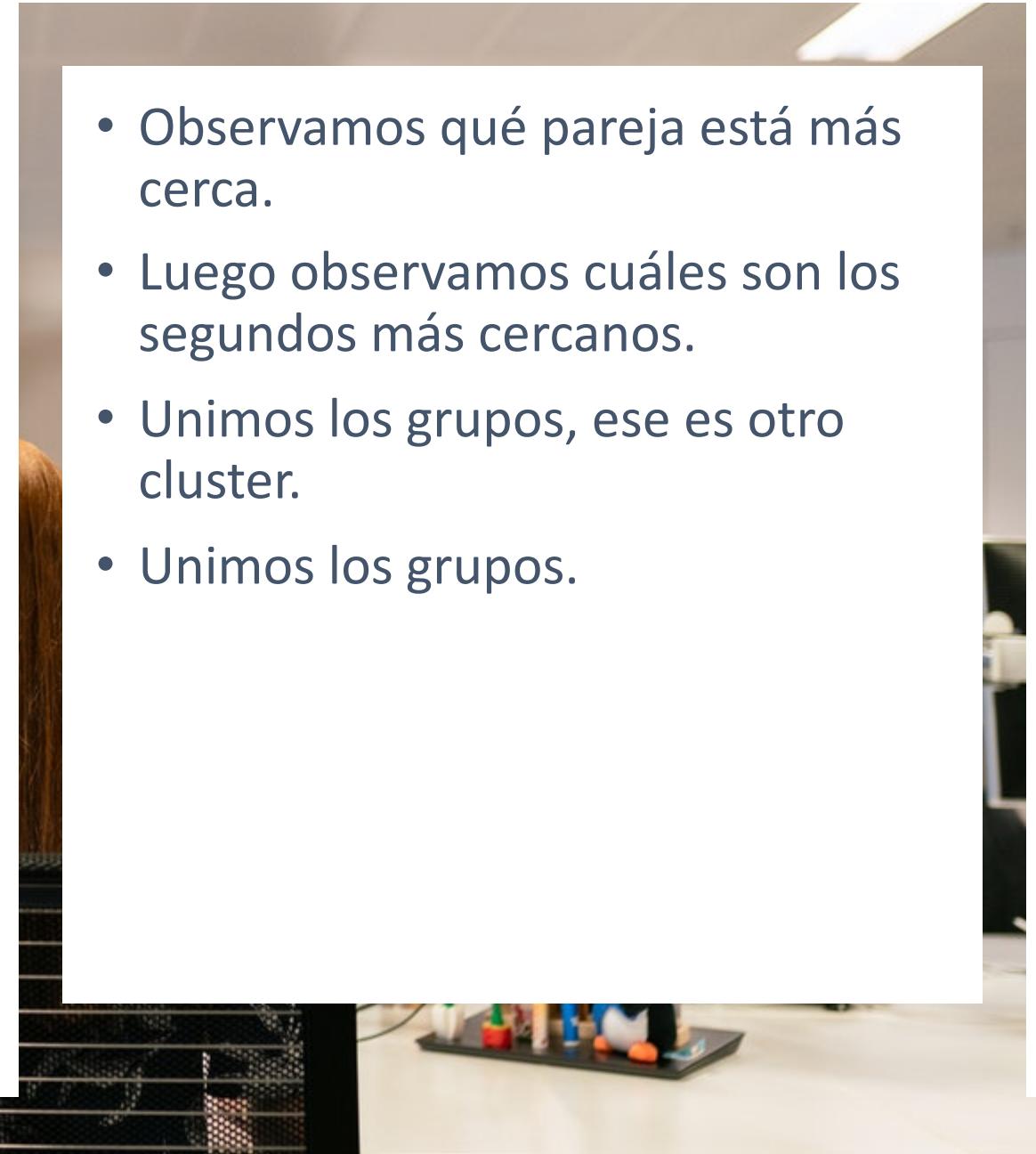


- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

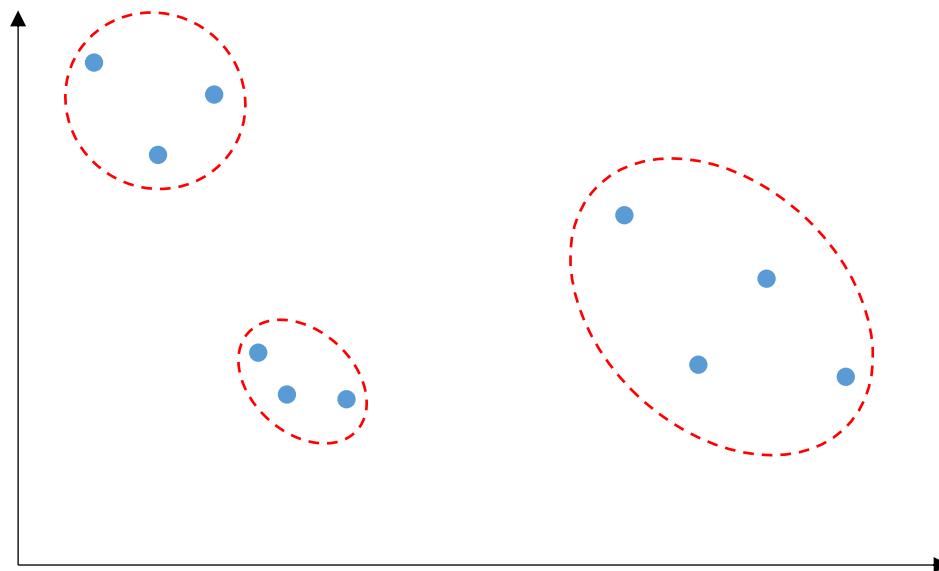
Clustering Jerárquico



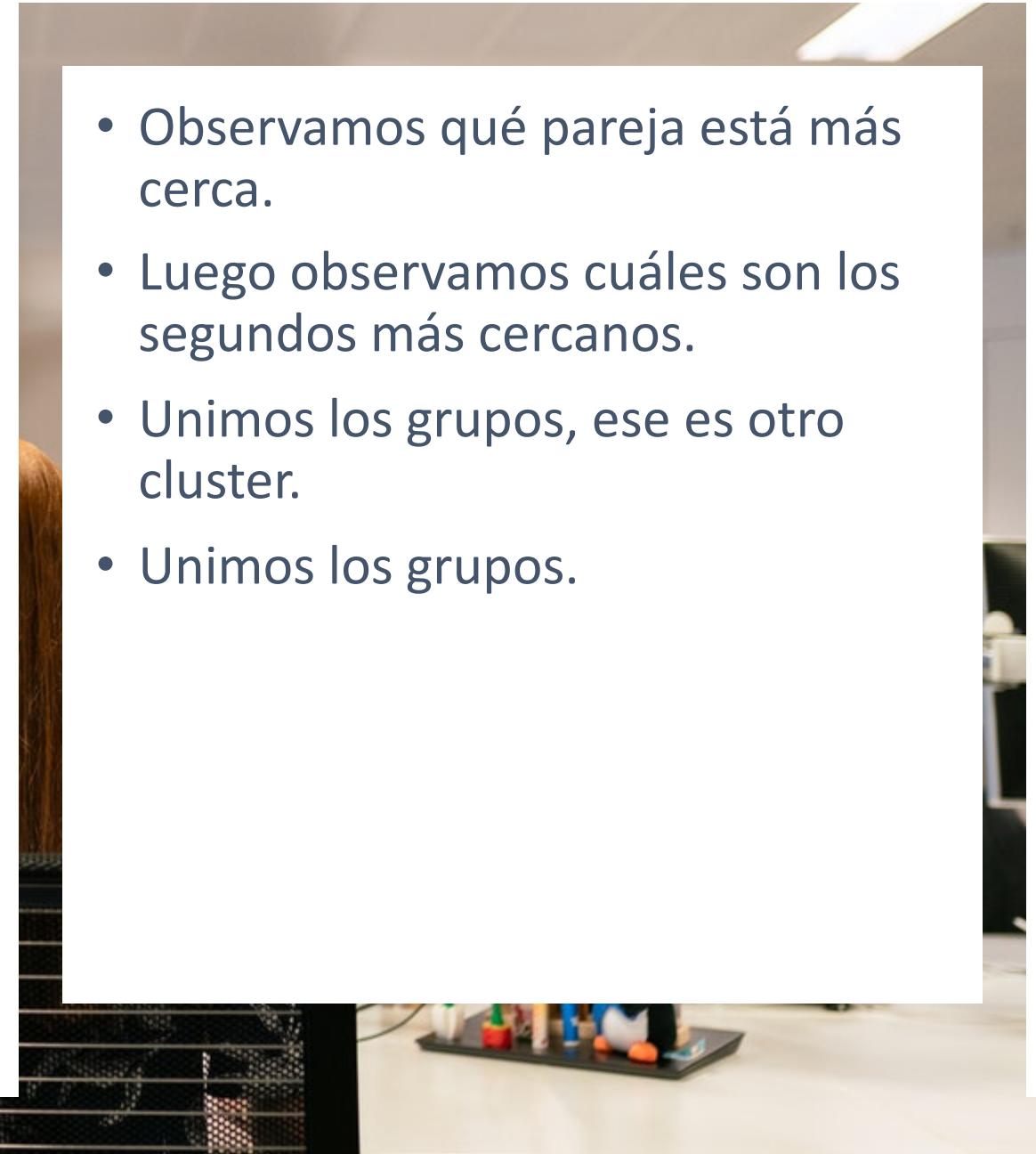
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.



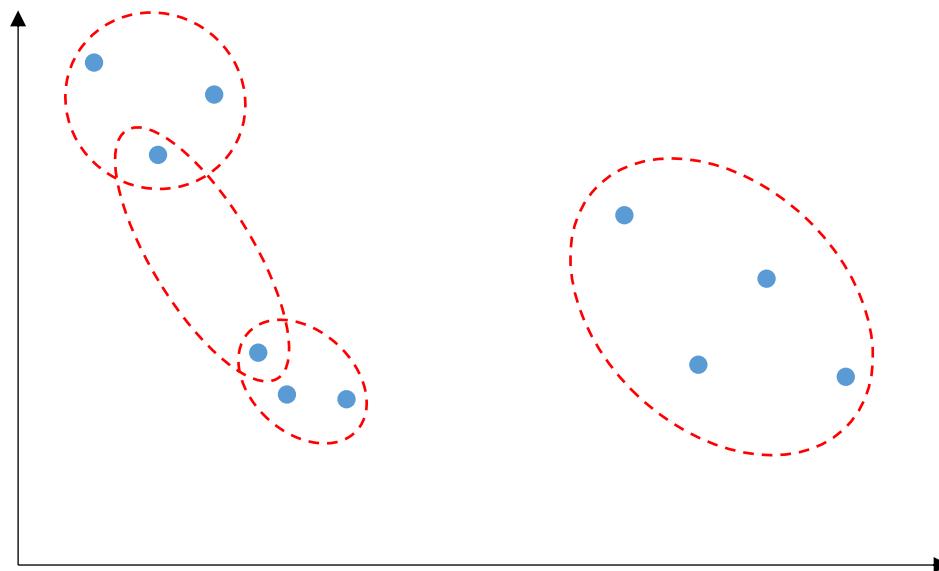
Clustering Jerárquico



- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

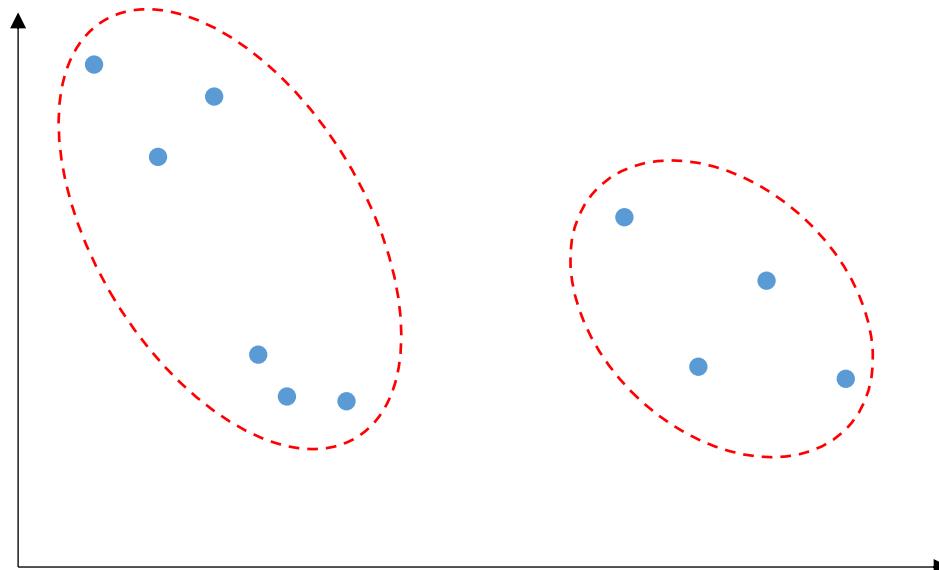


Clustering Jerárquico



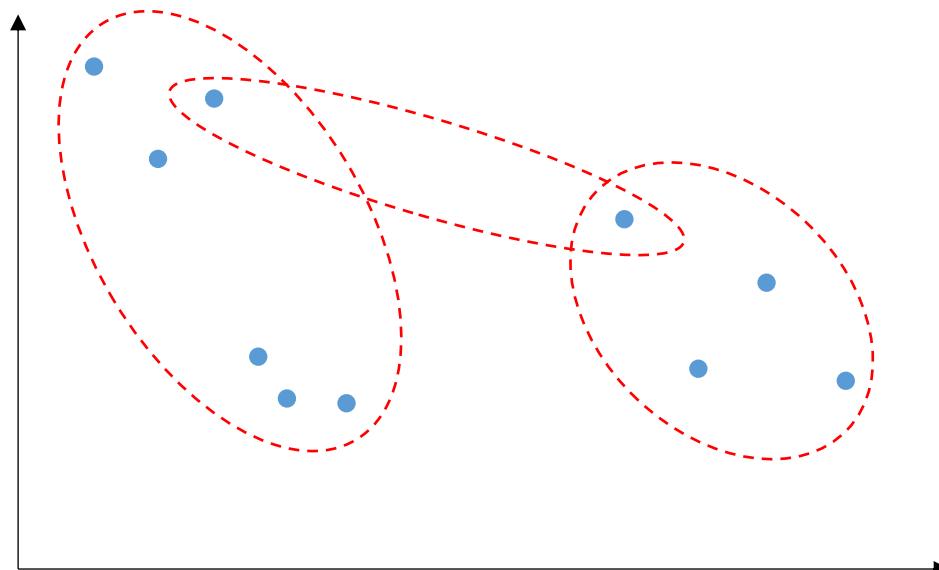
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

Clustering Jerárquico



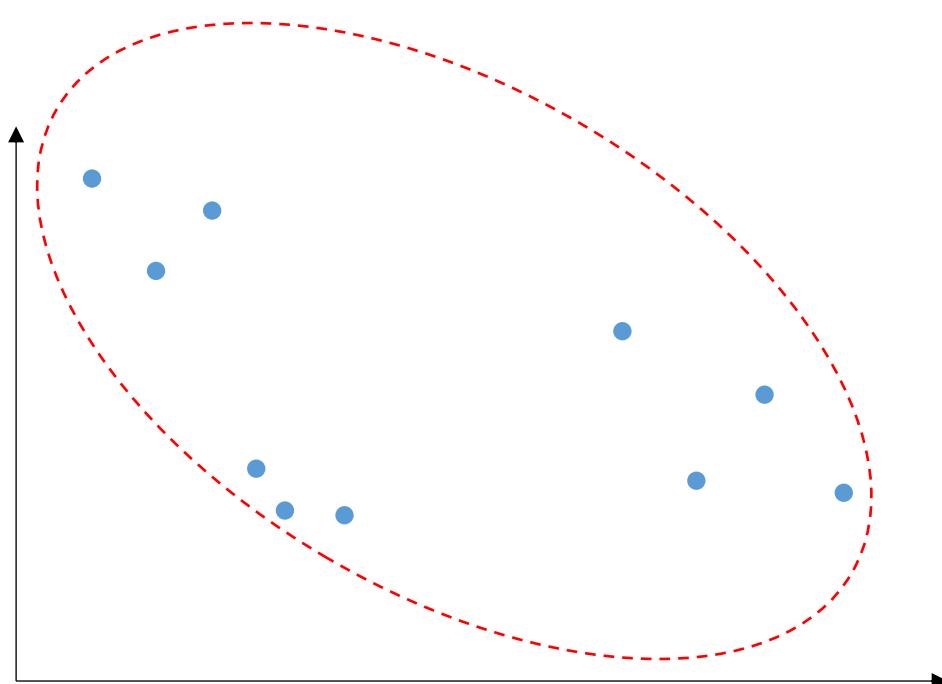
- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

Clustering Jerárquico



- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

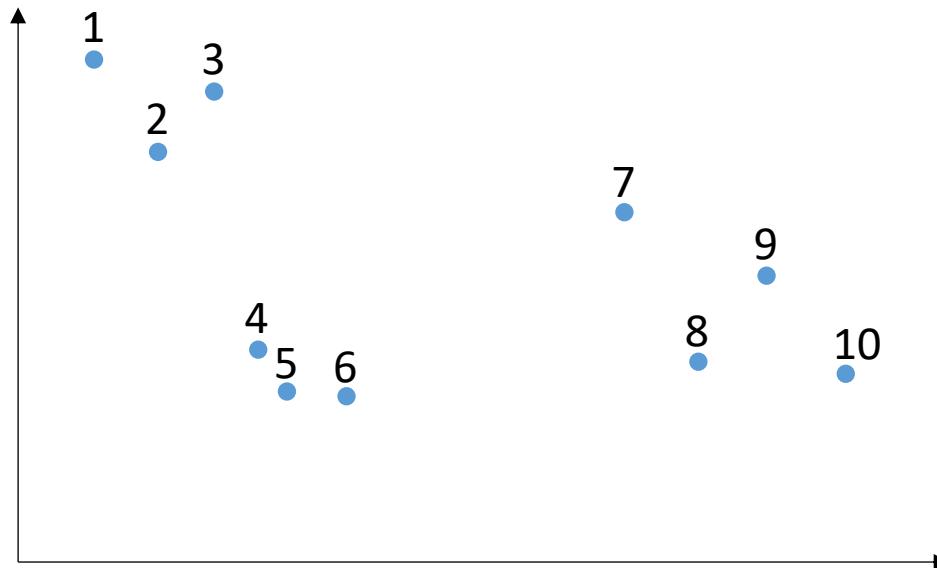
Clustering Jerárquico



- Observamos qué pareja está más cerca.
- Luego observamos cuáles son los segundos más cercanos.
- Unimos los grupos, ese es otro cluster.
- Unimos los grupos.

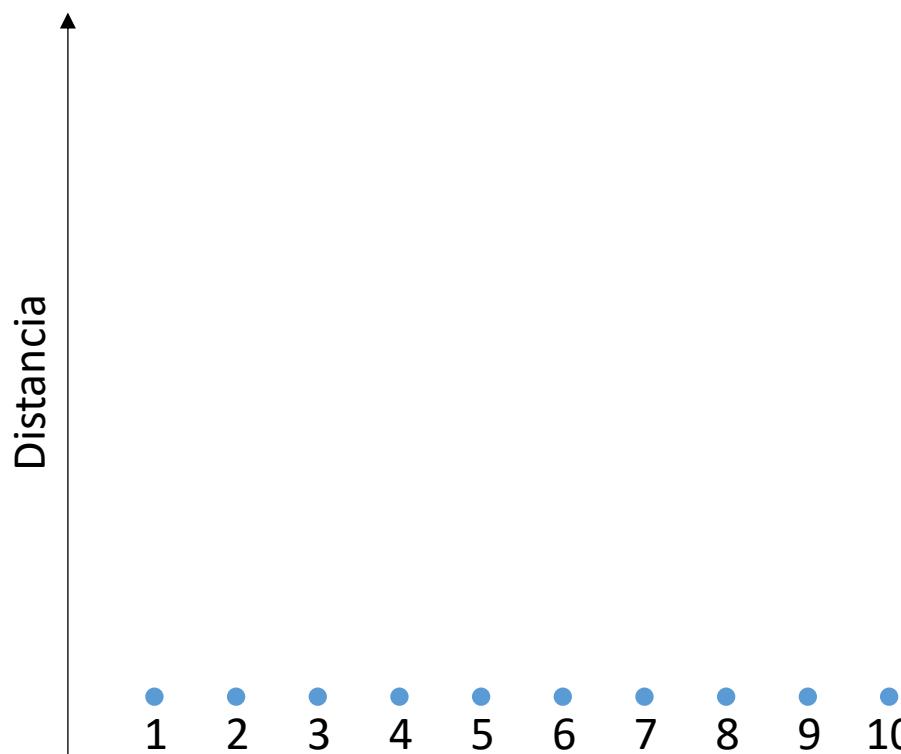
¿Pero cuántos grupos o qué distancia escogemos?

Clustering Jerárquico

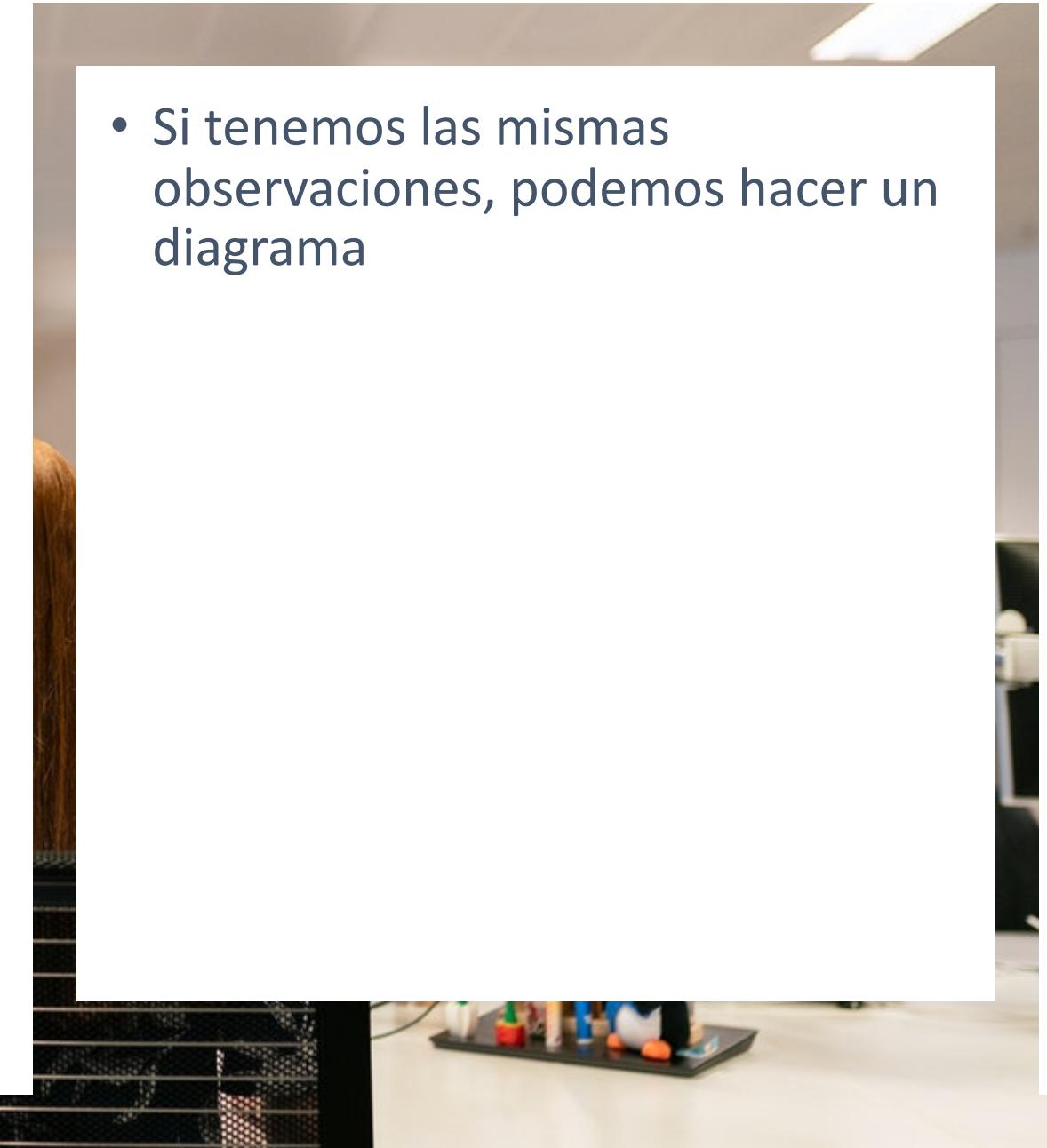


- Si tenemos las mismas observaciones, podemos hacer un diagrama

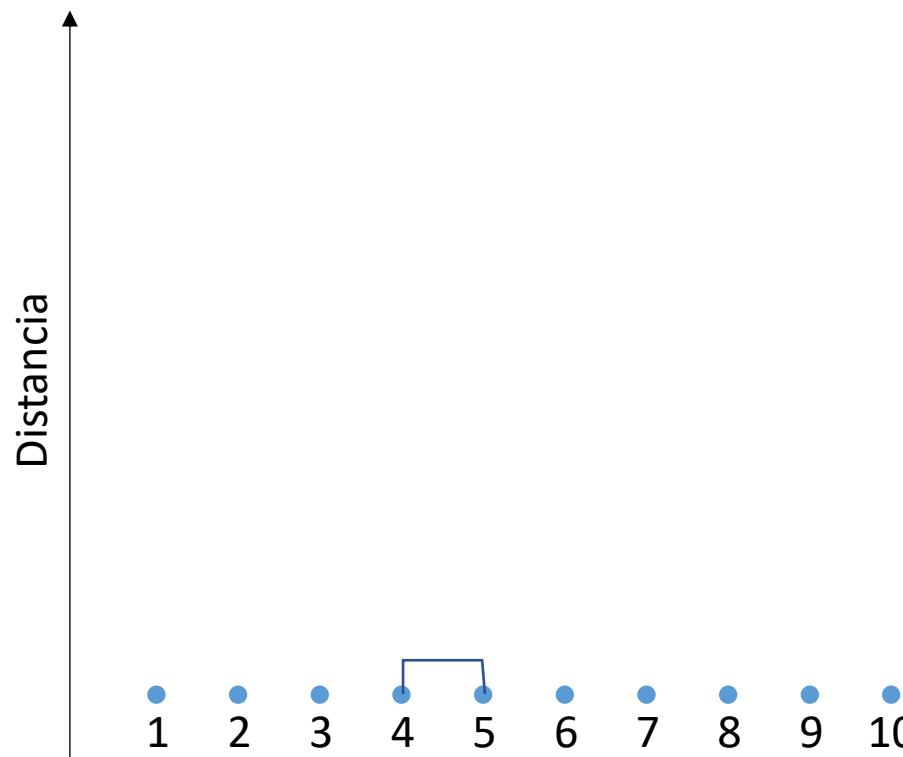
Clustering Jerárquico



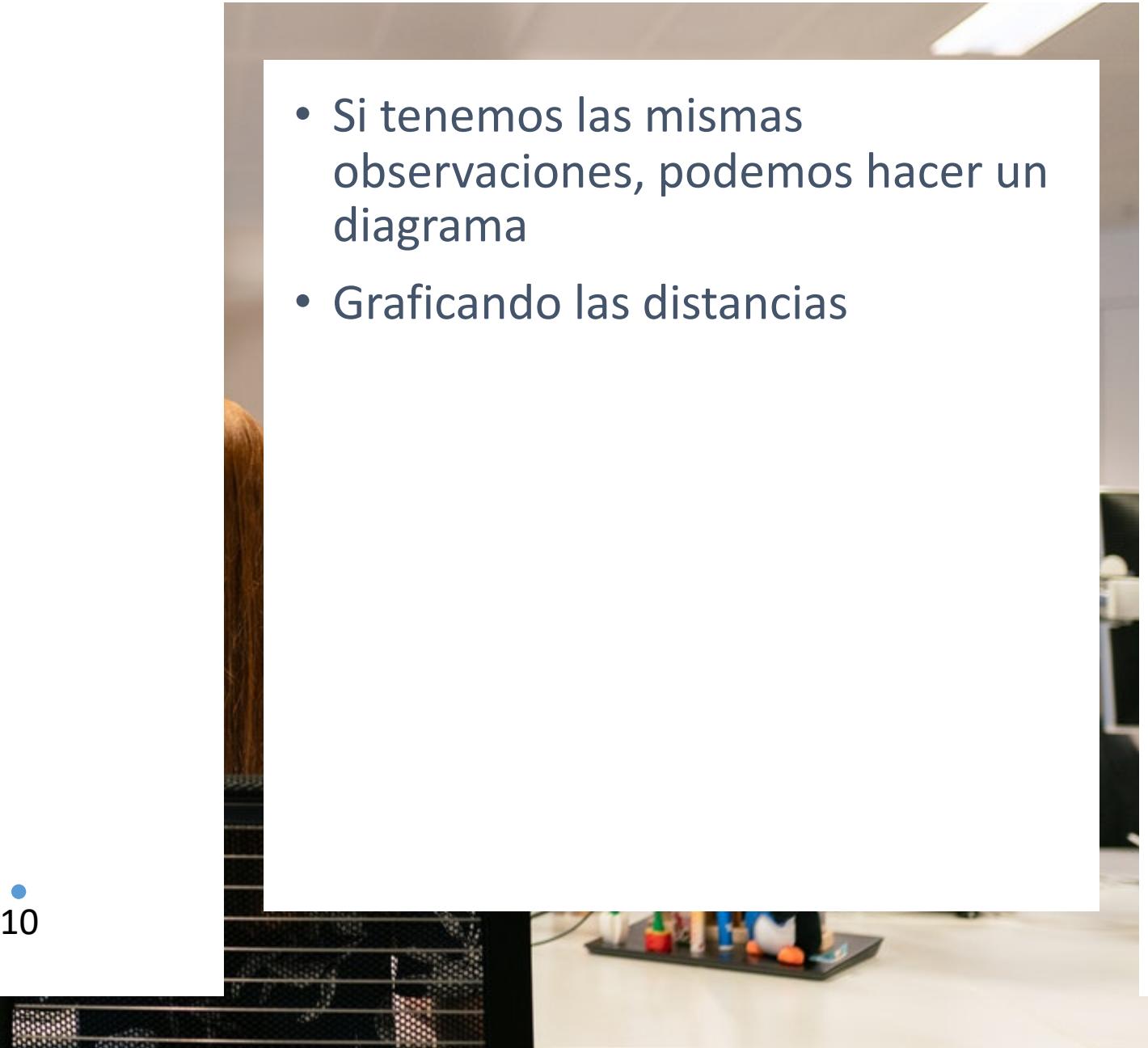
- Si tenemos las mismas observaciones, podemos hacer un diagrama



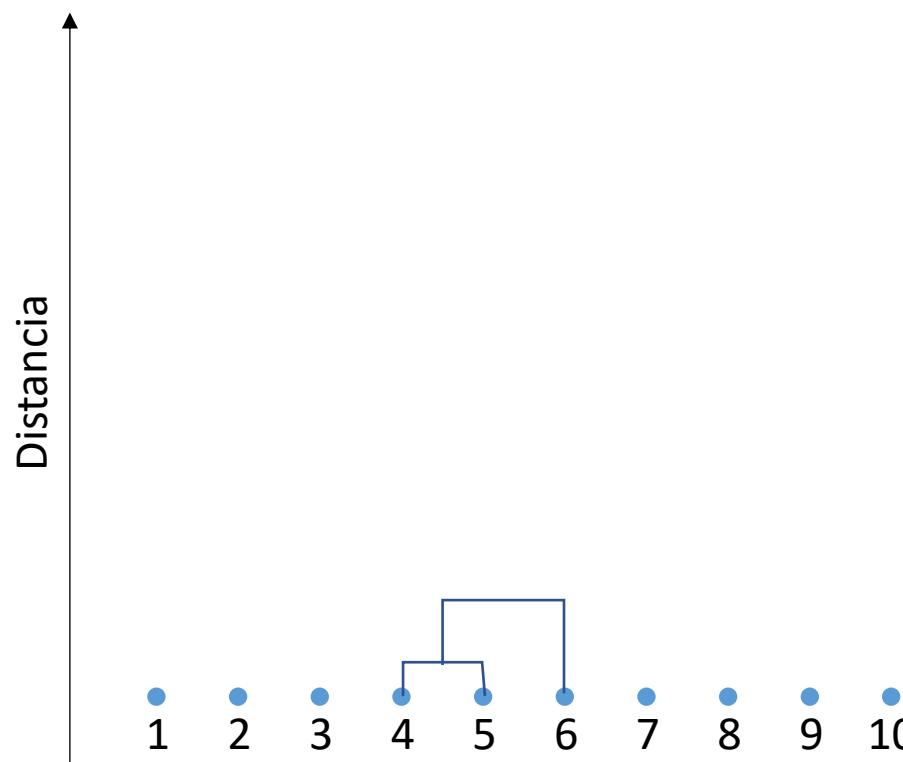
Clustering Jerárquico



- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

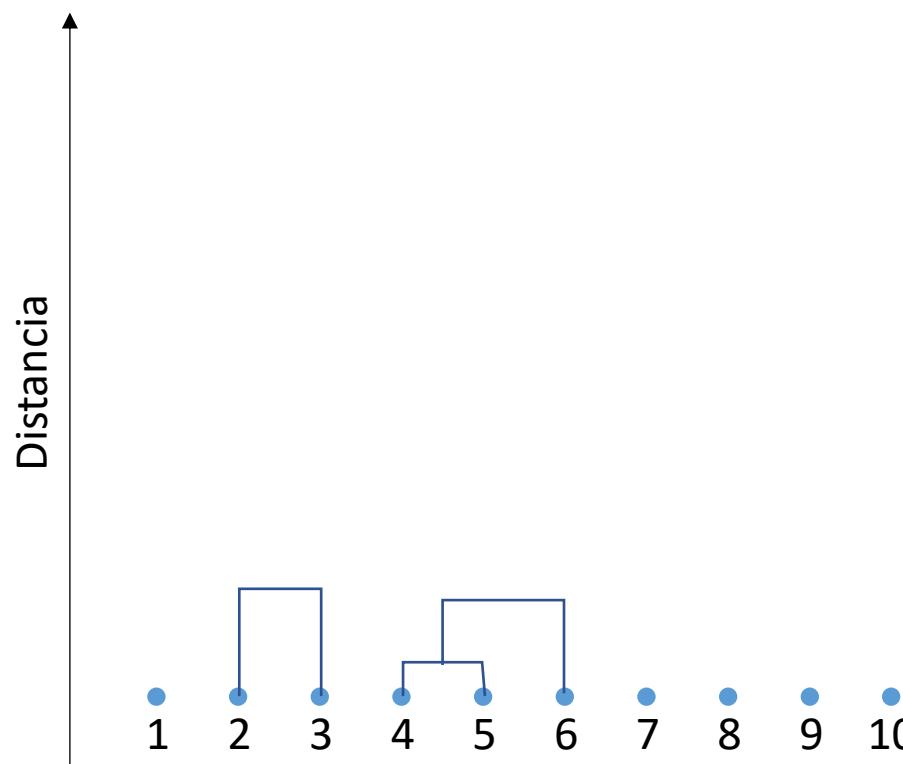


Clustering Jerárquico



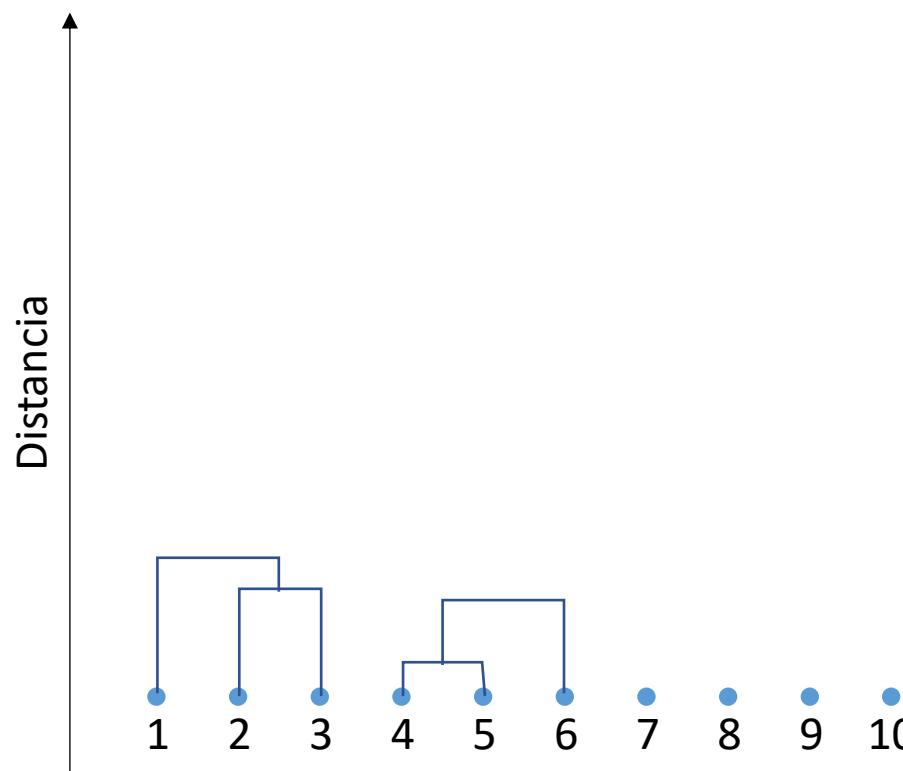
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

Clustering Jerárquico



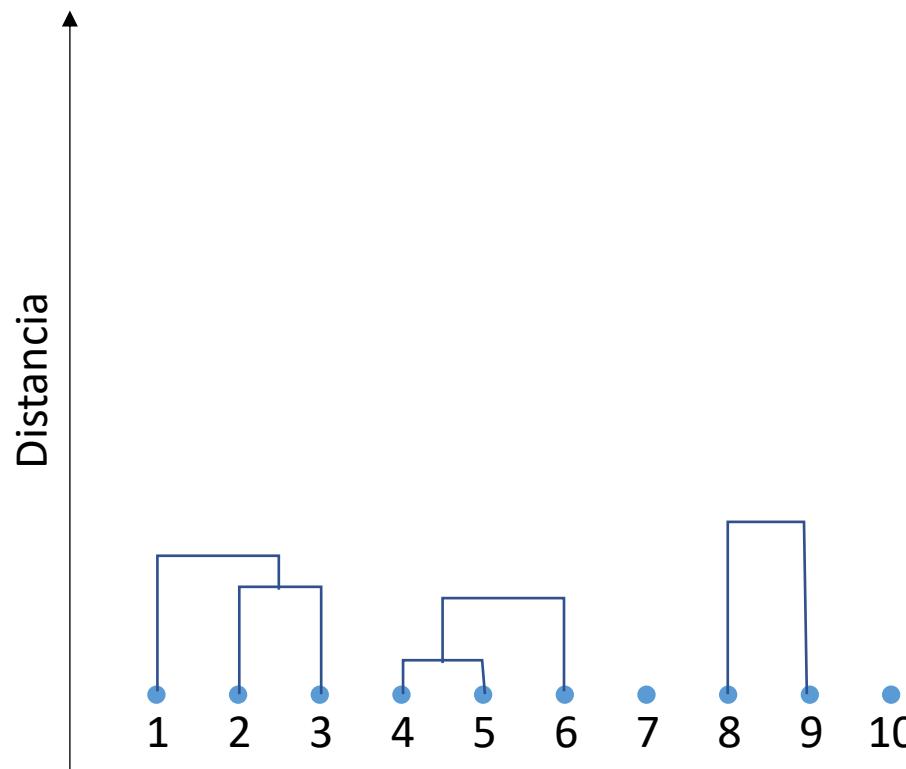
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

Clustering Jerárquico

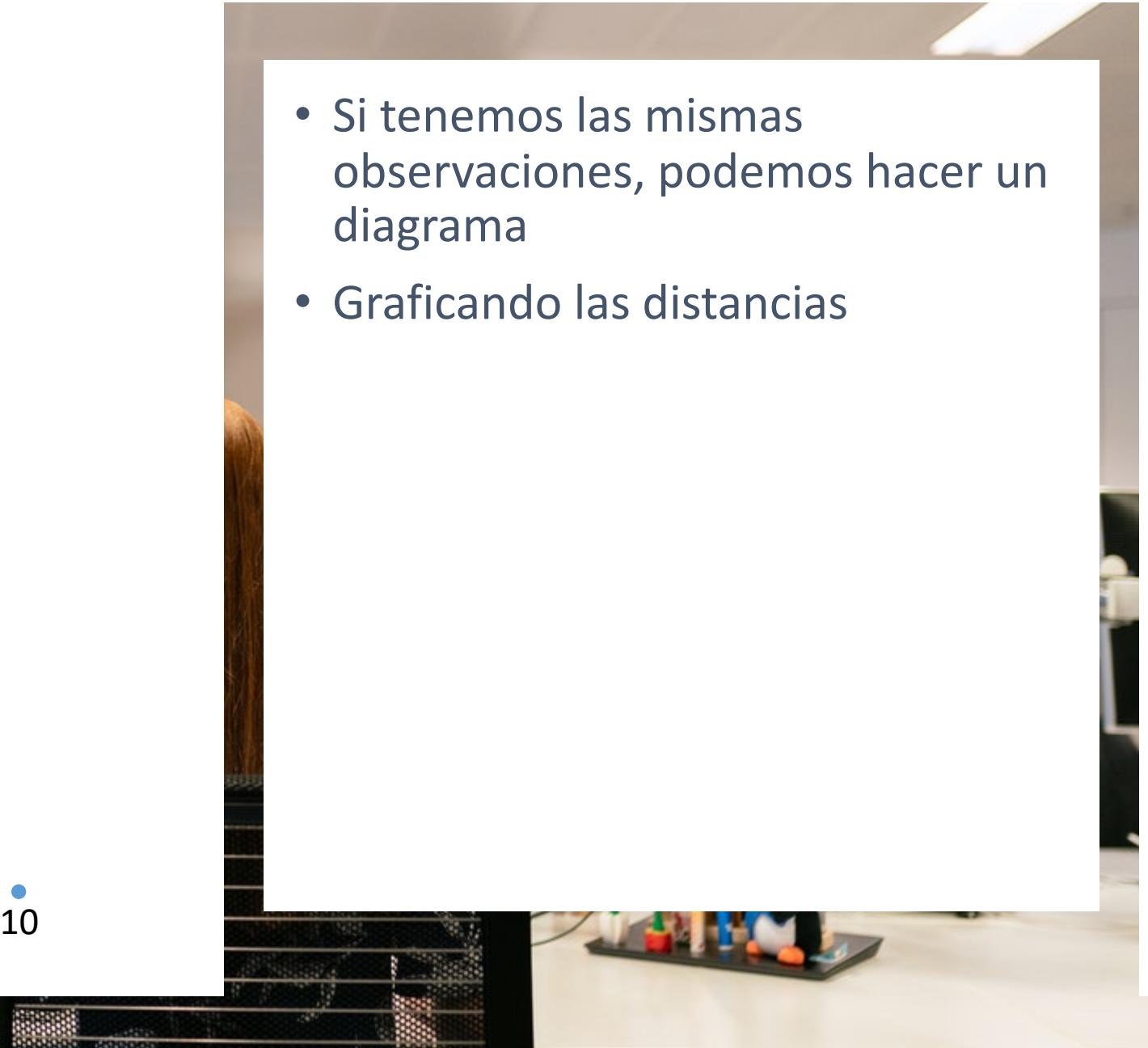


- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

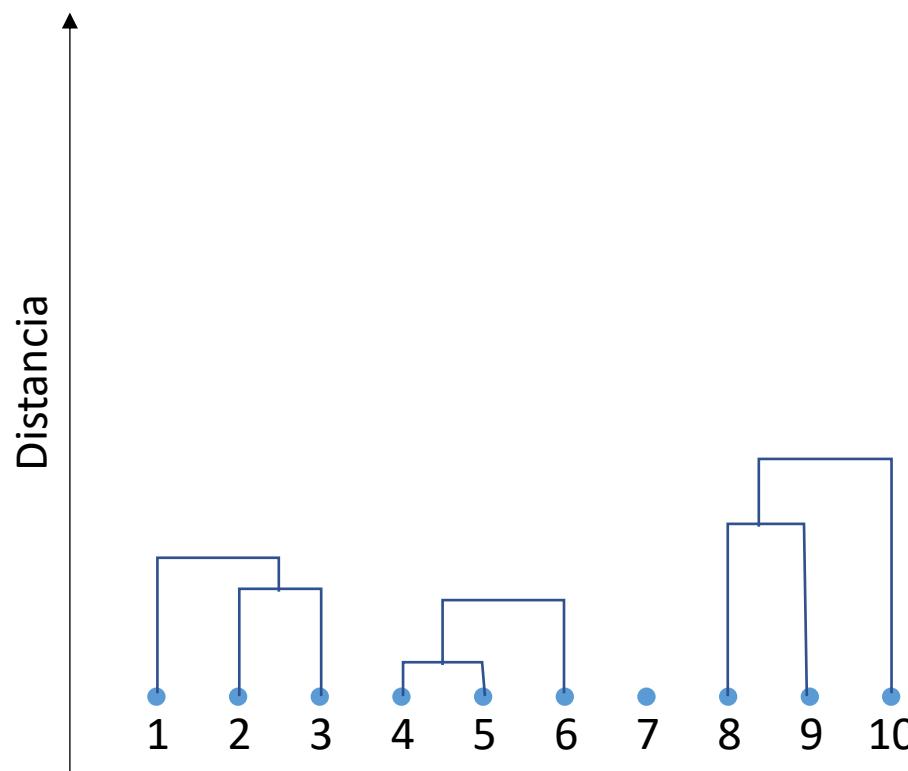
Clustering Jerárquico



- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

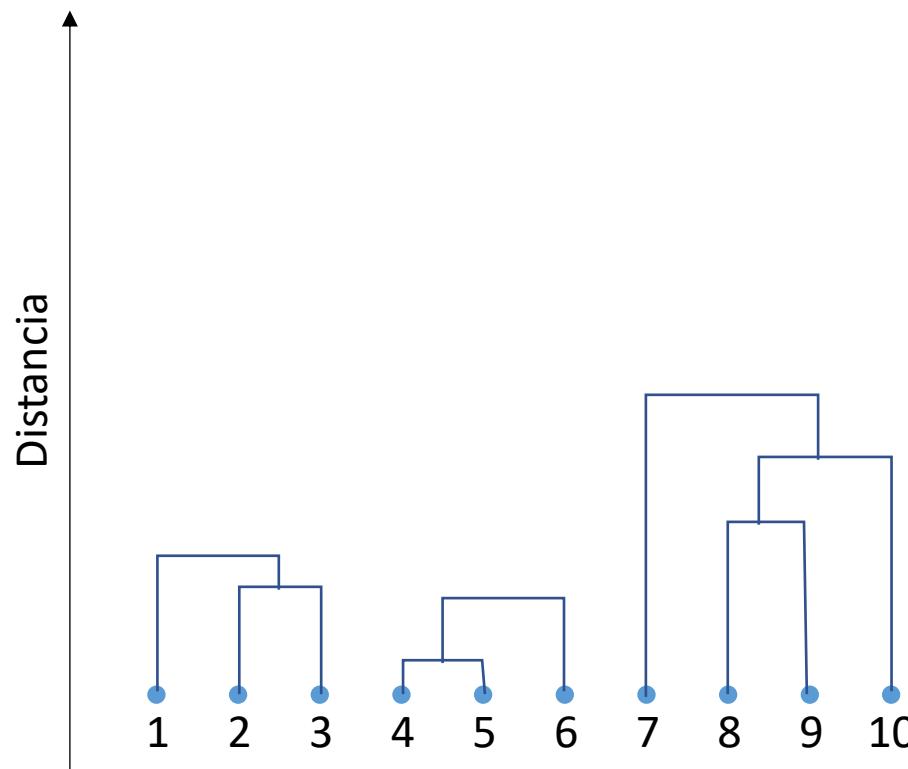


Clustering Jerárquico



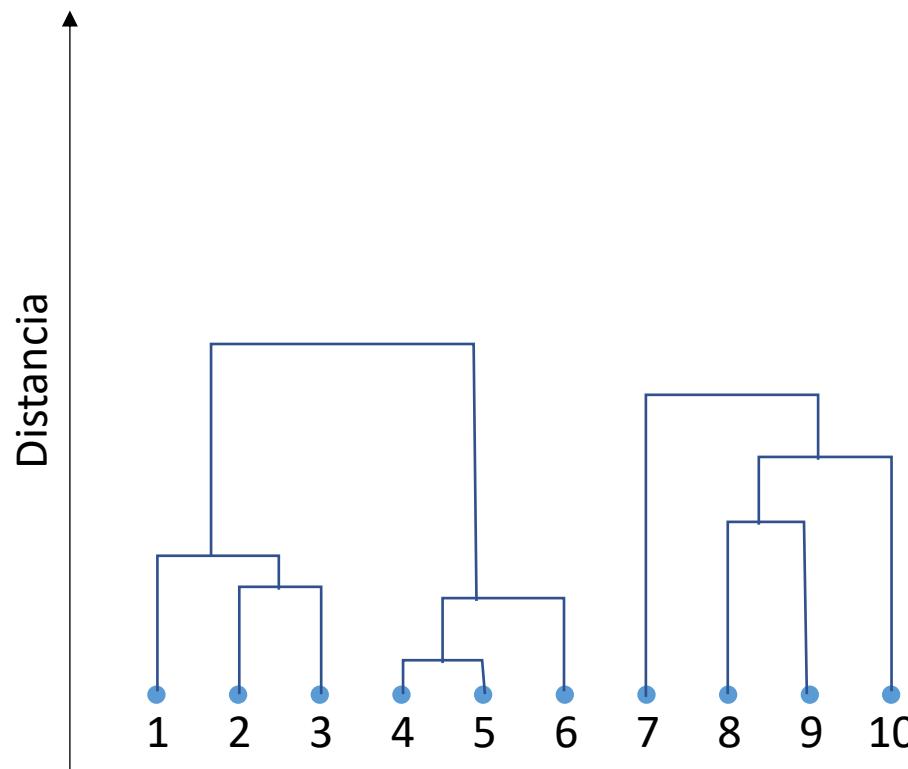
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

Clustering Jerárquico



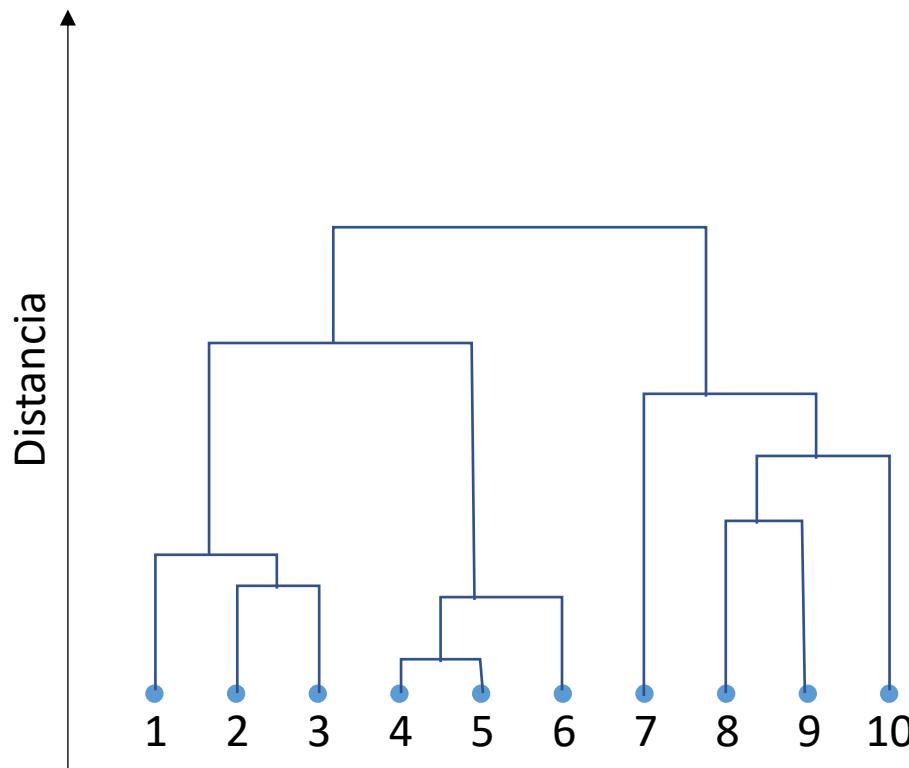
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

Clustering Jerárquico



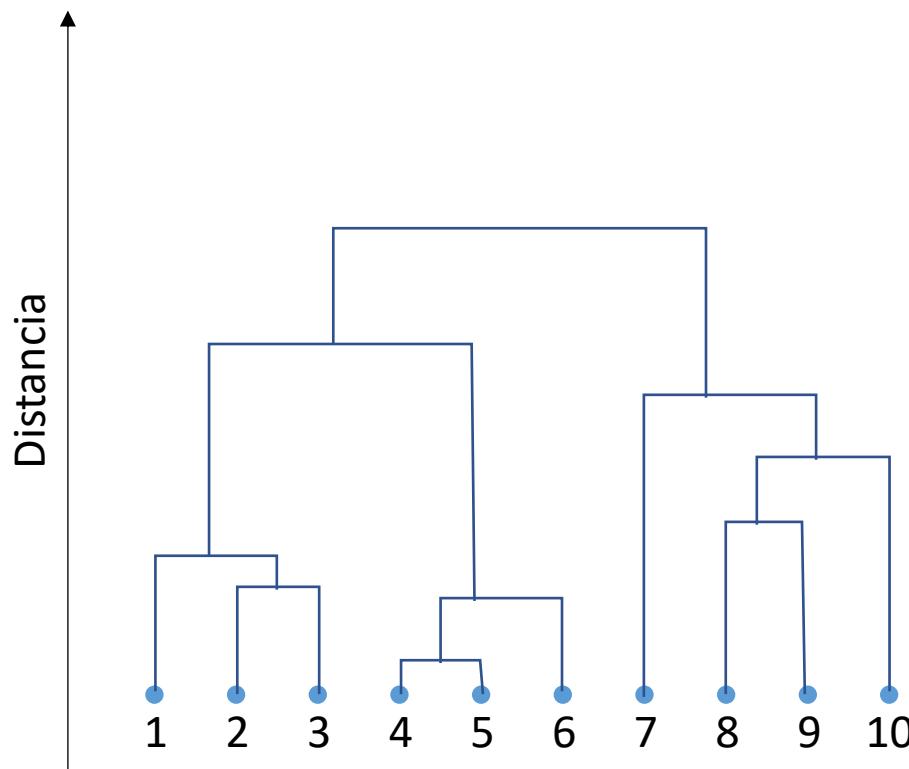
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias

Clustering Jerárquico



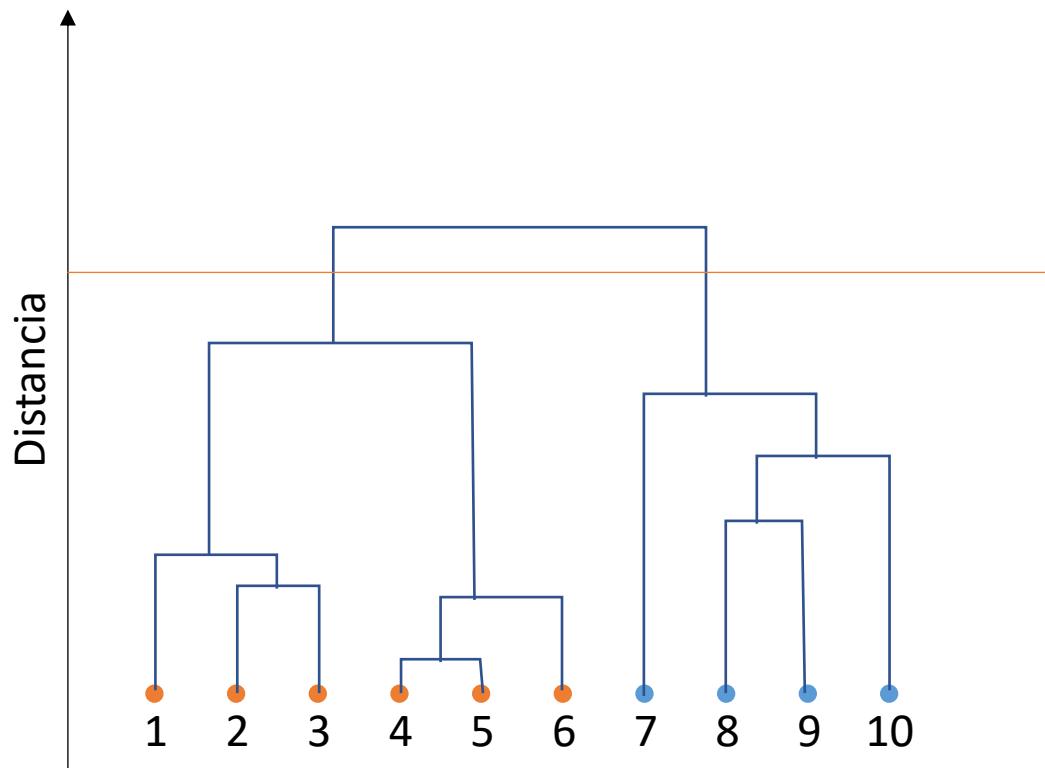
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias: dendrograma

Clustering Jerárquico



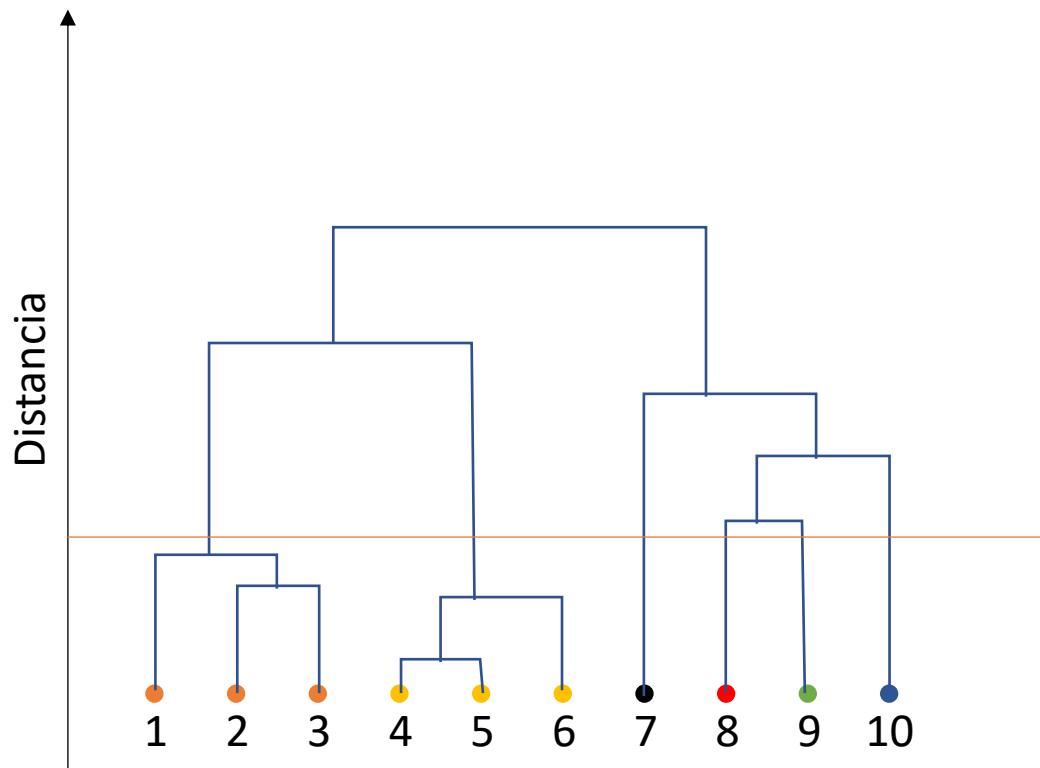
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias: dendrograma
- Podemos escoger dónde cortar

Clustering Jerárquico



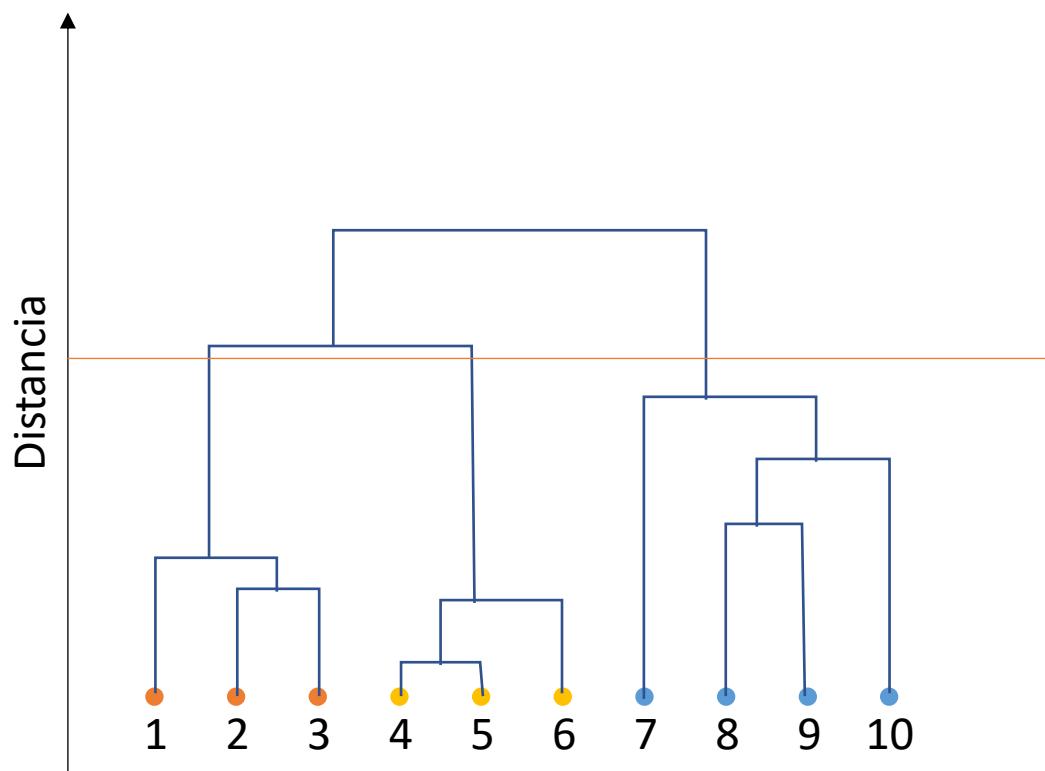
- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias: dendrograma
- Podemos escoger dónde cortar

Clustering Jerárquico



- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias: dendrograma
- Podemos escoger dónde cortar

Clustering Jerárquico



- Si tenemos las mismas observaciones, podemos hacer un diagrama
- Graficando las distancias: dendrograma
- Podemos escoger dónde cortar

Los tipos de *clustering*

- Centroides (como k-medias)
- Densidad (según cercanía secuencial de puntos)
- Distribución (según probabilidades)
- Conectividad (como jerárquico)
- ¡Diferentes técnicas nos van a dar diferentes clusters!

IA puede...

Propiedades de la inteligencia artificial que la hacen útil, inspiradas en aplicaciones previas

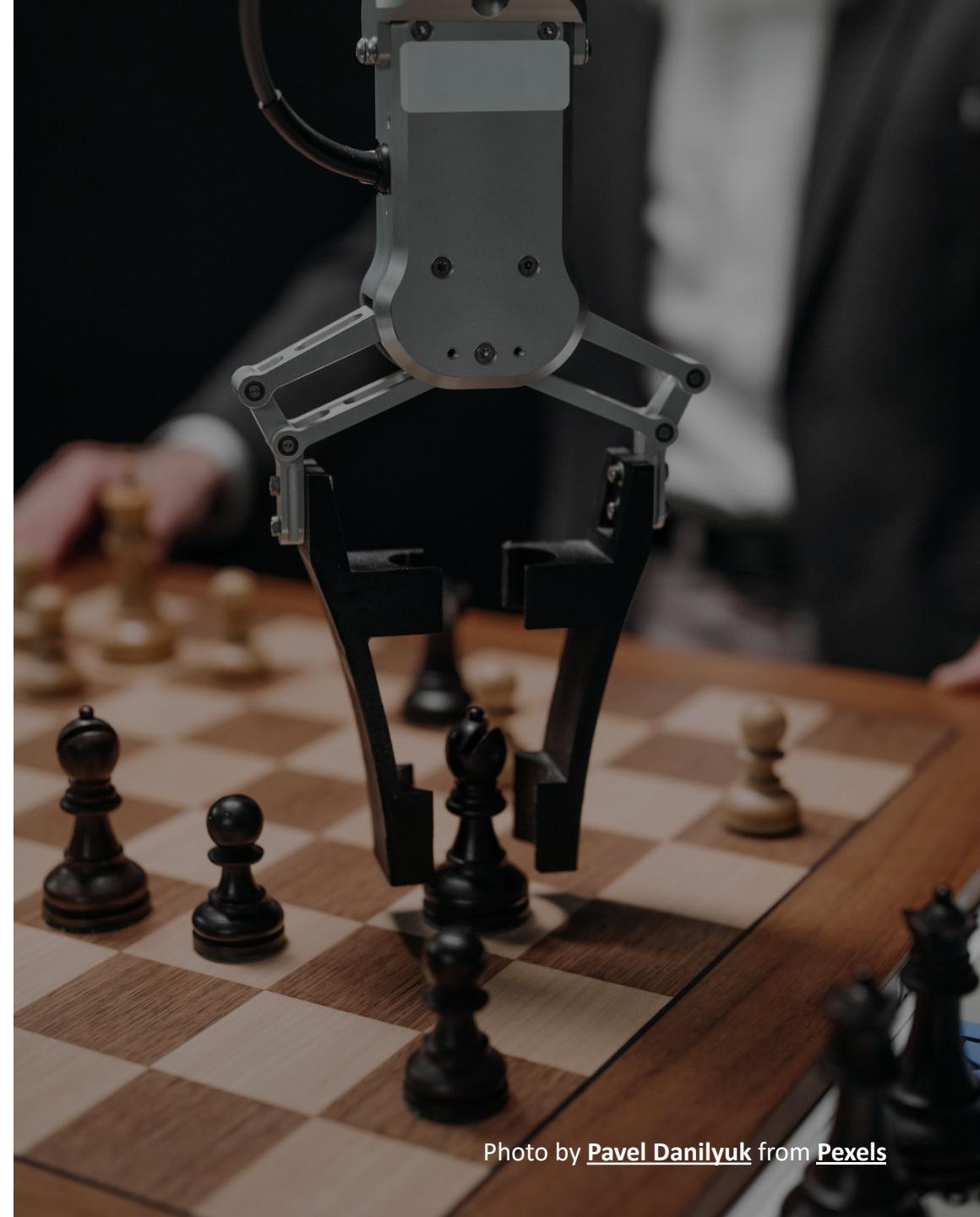


Photo by [Pavel Danilyuk](#) from [Pexels](#)

La IA tiene propiedades

- Las técnicas estudiadas se pueden utilizar en múltiples contextos.
- Existen razones por las cuales la inteligencia artificial es útil.
- Estudiar las propiedades de la inteligencia artificial y el aprendizaje de máquinas nos puede servir para encontrar aplicaciones.

IA es precisa

- Una vez una máquina aprende un proceso, puede tomar decisiones con altísima precisión:
 - Robot neuro cirujano
 - Carros autónomos



IA automatiza

- Caso de los drones de combate militar que despegan y aterrizan solos de forma remota.

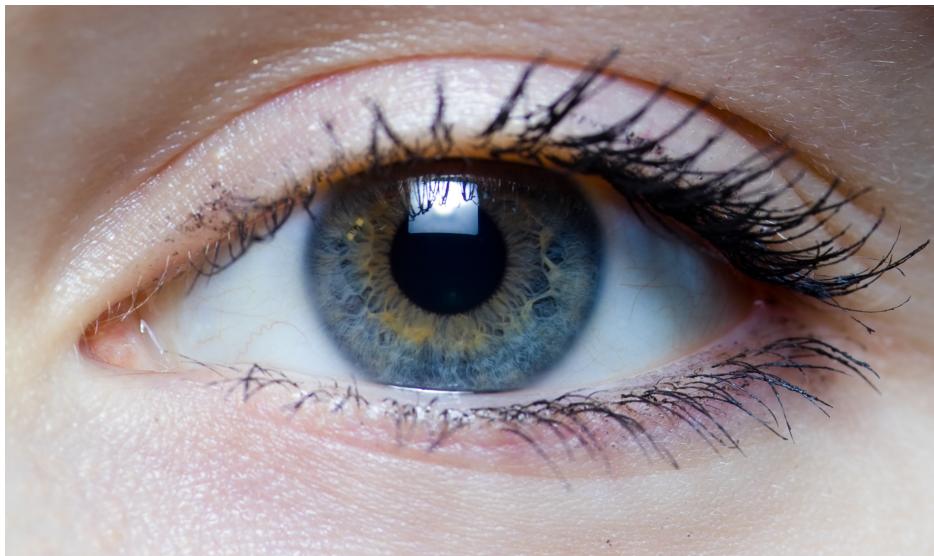


IA predice el futuro



- La fecha en la que una cosecha estará lista es clave para mantener los costos y la calidad agrícolas.
- Muy tarde: la cosecha pierde frescura.
- Muy temprano: hay que pagar los costos de maquinaria o trabajadores para la cosecha.

IA procesa información



- En India, hay más pacientes que oftalmólogos. Algunos procedimientos son sencillos de realizar si se diagnostican a tiempo, pero de lo contrario conducen a la pérdida de visión.
- IA lee resultados de exámenes y asigna prioridad de cita a quienes presentan las patologías.

Ética de la IA

Lo que hacemos puede afectar a las personas

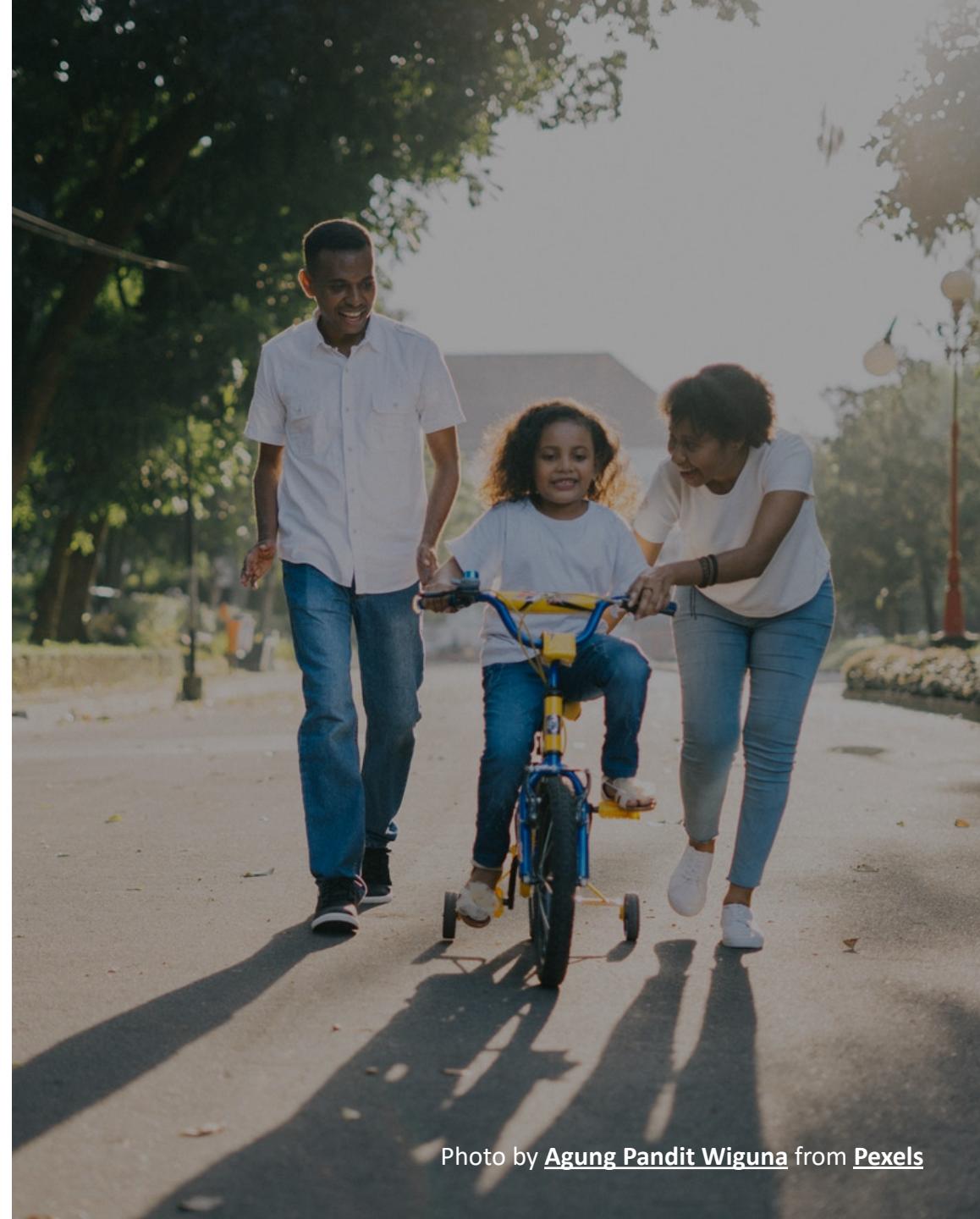


Photo by [Agung Pandit Wiguna](#) from [Pexels](#)

Las máquinas fallan

- Las predicciones son adivinanzas informadas en datos.
- Pueden errar (y lo hacen frecuentemente)
- Aunque... a veces se equivocan menos que los humanos...

Podemos afectar a otras personas con máquinas que fallan

- Existen tres grandes consideraciones éticas de la inteligencia artificial en este momento (sintetiza Sandel):
- Privacidad y vigilancia
- Sesgo y discriminación
- El rol del juicio humano en la sociedad



Privacidad y vigilancia



- Ahora podemos analizar una cantidad masiva de datos en tiempo real
- Riesgos asociados a la estabilidad de los sistemas políticos
- Modelo de China: disyuntiva vigilancia vs seguridad
- Además, nuestros datos dicen mucho de nosotros: caso *Cambridge Analytics*

Sesgo y discriminación

- A primera vista las máquinas parecen “objetivas”
- Pero se ha identificado que pueden aprender y acentuar los sesgos preexistentes en la sociedad
- Si un sistema presenta discriminación contra un grupo, puede que esta discriminación se manifieste en los datos
- Las máquinas aprenden que estos son comportamientos “naturales del sistema”
- Ej. Sistemas de admisión a universidades, aprobaciones de crédito, etc.



¿Nos reemplazarán las máquinas?



- En este momento sólo tenemos IA débil (especializada en una tarea)
- La automatización es real: tareas cada vez más creativas están siendo aprendidas por máquinas
- Pero no es nueva: revolución industrial
- Algunos argumentan que por cada trabajo desplazado por las máquinas se crearán otros

Enfoque centrado en humanos para la IA

BRIEFING



EU guidelines on ethics in artificial intelligence: Context and implementation

SUMMARY

The discussion around artificial intelligence (AI) technologies and their impact on society is increasingly focused on the question of whether AI should be regulated. Following the call from the European Parliament to update and complement the existing Union legal framework with guiding ethical principles, the EU has carved out a 'human-centric' approach to AI that is respectful of European values and principles. As part of this approach, the EU published its guidelines on ethics in AI in April 2019, and European Commission President-elect, Ursula von der Leyen, has announced that the Commission will soon put forward further legislative proposals for a coordinated European



1/80

Enfoque centrado en humanos para la IA

- Siete principios que derivan en los lineamientos generales del Parlamento Europeo.

The key EU requirements for achieving trustworthy AI

- human agency and oversight
- robustness and safety
- privacy and data governance
- transparency
- diversity, non-discrimination and fairness
- societal and environmental well-being
- accountability

Taller diferenciado

Grupos

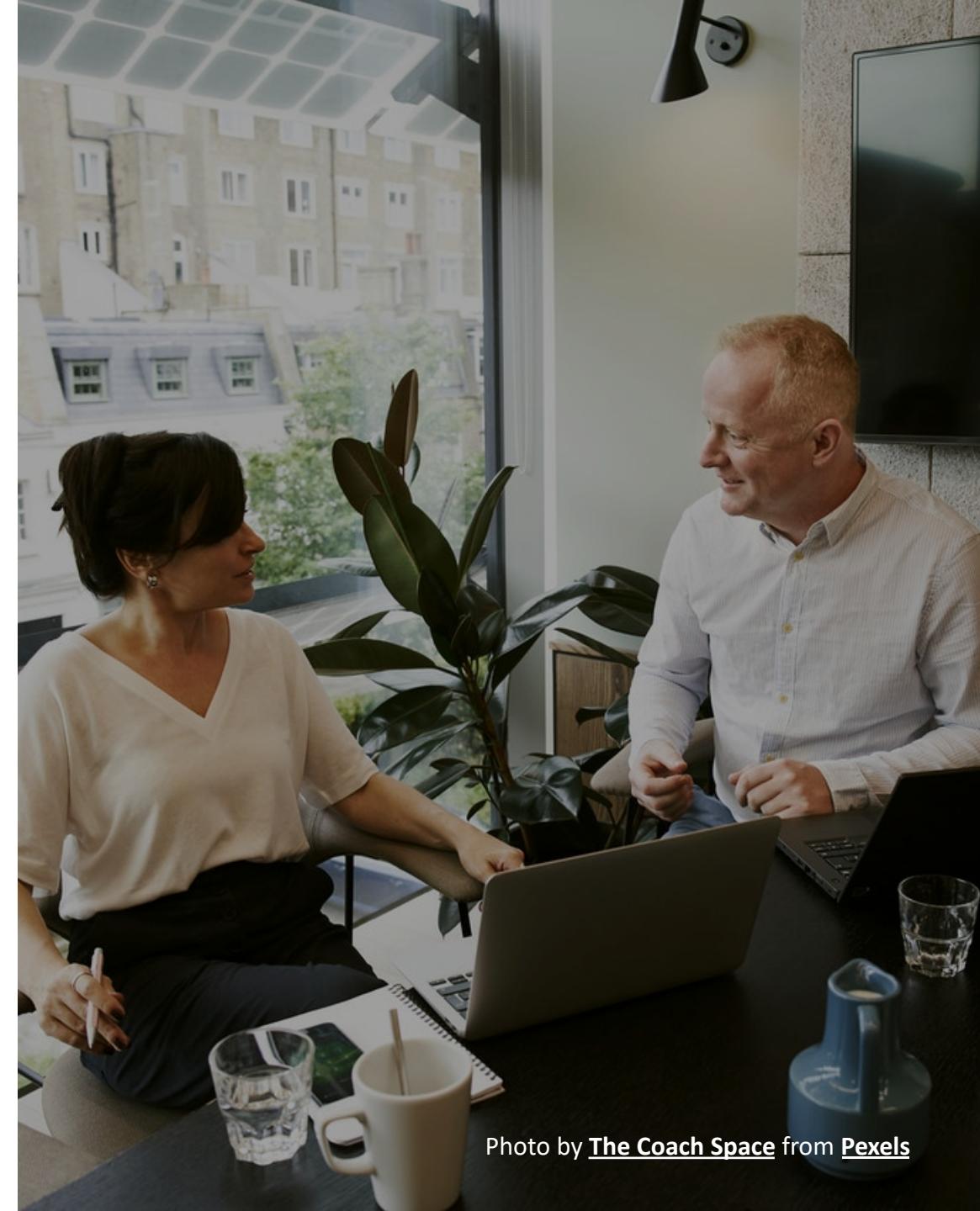
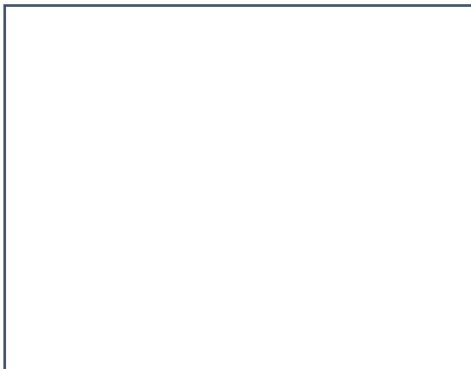


Photo by [The Coach Space](#) from [Pexels](#)

Retos potenciales:

Primera exploración

Clasificación



Regresión



No supervisado



1. ¿Qué tipos de datos manejamos en la CGR?
¿Qué información tienen?
2. ¿Están estructurados o desestructurados?
3. ¿Qué decisiones se toman a partir de datos / información en nuestros departamentos?
4. ¿Hay cantidades particulares que son de interés en nuestros departamentos?
5. ¿Qué filas (observaciones) suelen tener nuestras bases de datos? ¿Por ejemplo entidades, municipios, sectores, años, etc.?

Retos potenciales:

Segunda exploración

1. ¿Qué tipo de situaciones en la Contraloría tienen estas características?
2. Piensen en tareas, procesos cotidianos o periódicos, etc.

Necesitamos saber el futuro

Alguien puede decidir, pero requiere mucha precisión de análisis

Es repetitivo, engorroso, consume tiempo... ¿se podría automatizar?

Requiere procesar grandes cantidades de información

