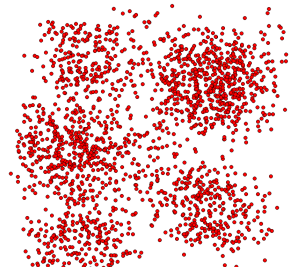
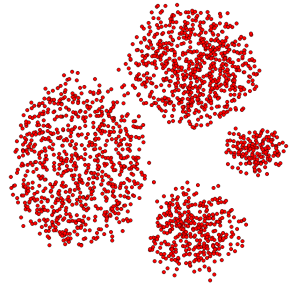


Agrupamiento

Tipos de algoritmos de agrupamiento

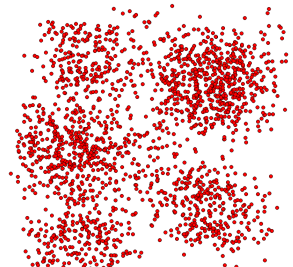
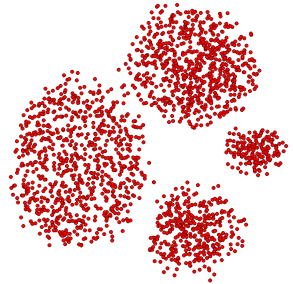
- ▶ Basados en particiones
- ▶ Jerárquicos
- ▶ Espectrales
- ▶ Basados en densidad
- ▶ Probabilísticos



Agrupamiento

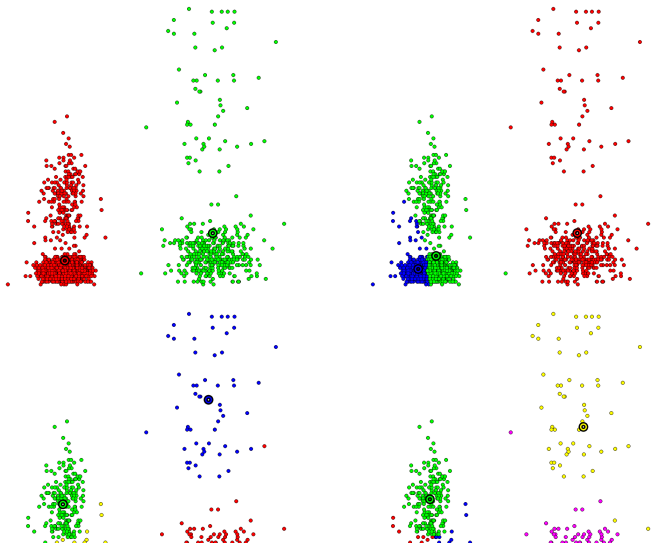
Tipos de algoritmos de agrupamiento

- ▶ Basados en particiones
- ▶ **Jerárquicos**
- ▶ Espectrales
- ▶ Basados en densidad
- ▶ Probabilísticos



Agrupamiento

Elegir el número de clústeres (K)

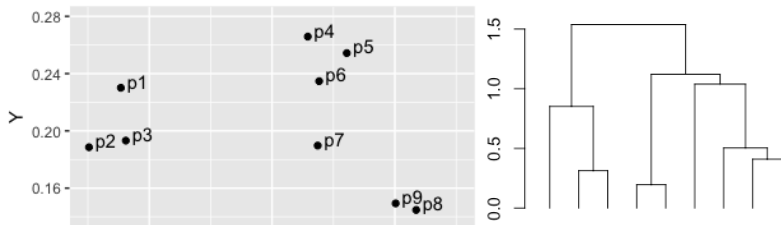


Agrupamiento Jerárquico

Un continuo de particiones de los datos

Se particiona el dataset desde $K = 1$ hasta $K = n$

** ¿Cuál es la mejor partición?



Agrupamiento Jerárquico

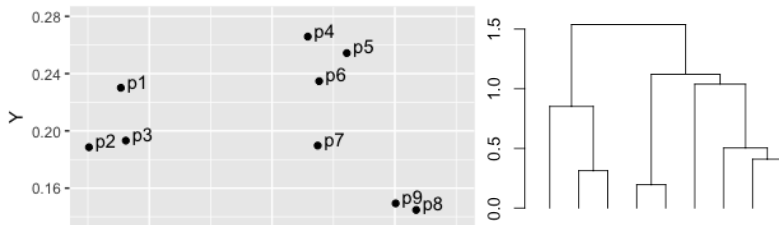
Un continuo de particiones de los datos

Se particiona el dataset desde $K = 1$ hasta $K = n$

** ¿Cuál es la mejor partición?

Algoritmos:

- ▶ Aglomerativo
- ▶ Divisivo

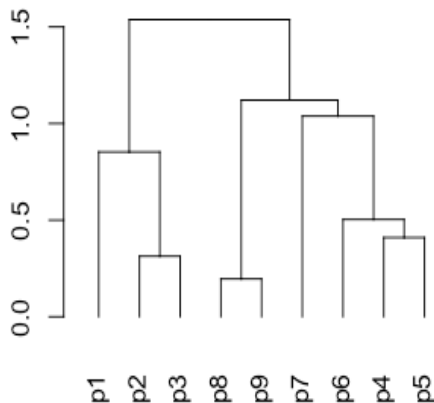


Agrupamiento Jerárquico

Dendrograma

Representación gráfica de un agrupamiento jerárquico

- ▶ Cada nodo, es un conjunto de ejemplos (clúster)
- ▶ Los clústeres se van uniendo/separando según criterios de distancia
- ▶ La longitud de las líneas verticales indica la distancia entre los clústeres que se unen/separan



Agrupamiento Jerárquico

Intuición

Si no conozco cuántos grupos/clústeres hay, de entrada no voy a elegir el número K

Los clústeres se forman de ejemplos que están cercanos entre ellos

El concepto de cercanía puede ser relativo:

1. **Términos absolutos:** La similitud entre estos dos clústeres es...
2. **Términos relativos:** Los dos clústeres más similares entre sí son...

**** De manera equivalente, podemos hablar de lejanía/diferencia**

Agrupamiento Jerárquico

Aglomerativo

Aglomeración

Partiendo de $K = n$, se van uniendo iterativamente pares de clústeres hasta $K = 1$ de manera voraz

0. Al principio, cada ejemplo tiene su propio clúster

1. Tras la primera unión, existen $K = n - 1$ clústeres
(todos unitarios, menos uno clúster que tiene 2 elementos)

...

i. Tras la i -ésima unión, existen $K = n - i$ clústeres

...

n-1. El algoritmo acaba cuando $K = 1$

(se unen los dos últimos clústeres en un clúster con todos los ejemplos)

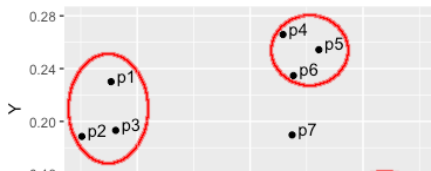
Agrupamiento Jerárquico

Aglomerativo

Dos cuestiones

A medida que avanza el algoritmo...

¿qué dos clústeres se deben unir en cada paso?



Agrupamiento Jerárquico

Aglomerativo

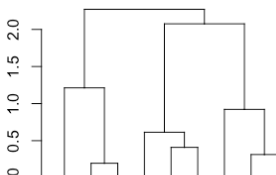
Dos cuestiones

A medida que avanza el algoritmo...

¿qué dos clústeres se deben unir en cada paso?

Al final del algoritmo, si queremos una partición concreta,

¿con qué partición nos quedamos?



Agrupamiento Jerárquico

Aglomerativo

Primera cuestión

A medida que avanza el algoritmo...

¿qué dos clústeres se deben unir en cada paso?

El par de clústeres, S_A^* y S_B^* , con menor disimilitud interclúster:

$$\{S_A^*, S_B^*\} = \arg \min_{\{S_A, S_B\}} d(S_A, S_B)$$

Agrupamiento Jerárquico

Aglomerativo

Primera cuestión

A medida que avanza el algoritmo...

¿qué dos clústeres se deben unir en cada paso?

El par de clústeres, S_A^* y S_B^* , con menor disimilitud interclúster:

$$\{S_A^*, S_B^*\} = \arg \min_{\{S_A, S_B\}} d(S_A, S_B)$$

¿cómo se mide la disimilitud interclúster?

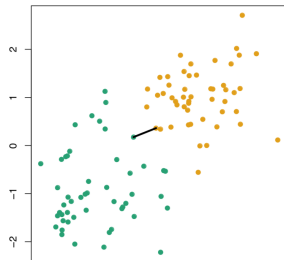
Agrupamiento Jerárquico

Aglomerativo

Criterios de unión

$$d(S_A, S_B) = \min_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud mínima



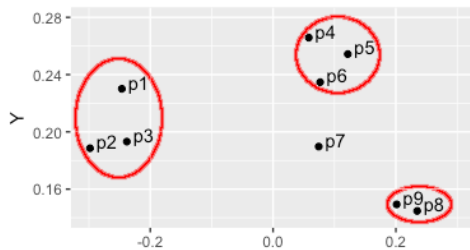
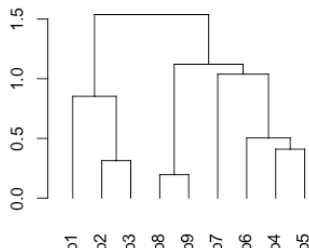
Agrupamiento Jerárquico

Aglomerativo

Criterios de unión

$$d(S_A, S_B) = \min_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud mínima



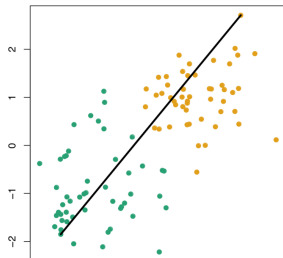
Agrupamiento Jerárquico

Aglomerativo

Criterios de unión

$$d(S_A, S_B) = \max_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud máxima



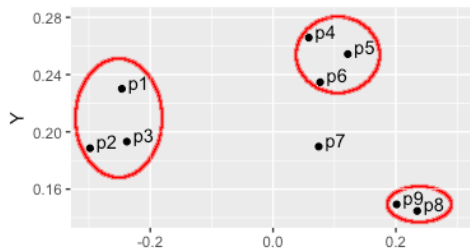
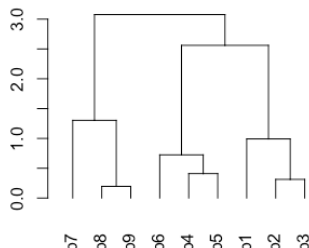
Agrupamiento Jerárquico

Aglomerativo

Criterios de unión

$$d(S_A, S_B) = \max_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud máxima



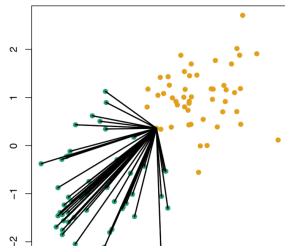
Agrupamiento Jerárquico

Aglomerativo

Criterios de unión

$$d(S_A, S_B) = \frac{1}{|S_A| \cdot |S_B|} \sum_{x_a \in S_A} \sum_{x_b \in S_B} d(x_a, x_b)$$

Disimilitud media



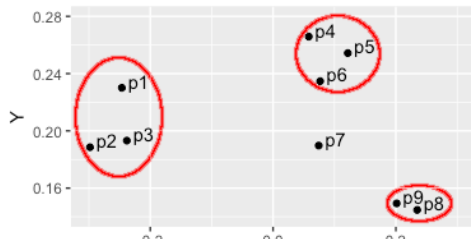
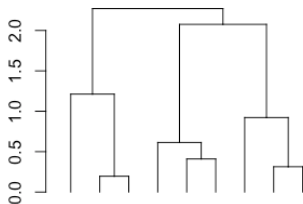
Agrupamiento Jerárquico

Aglomerativo

Criterios de unión

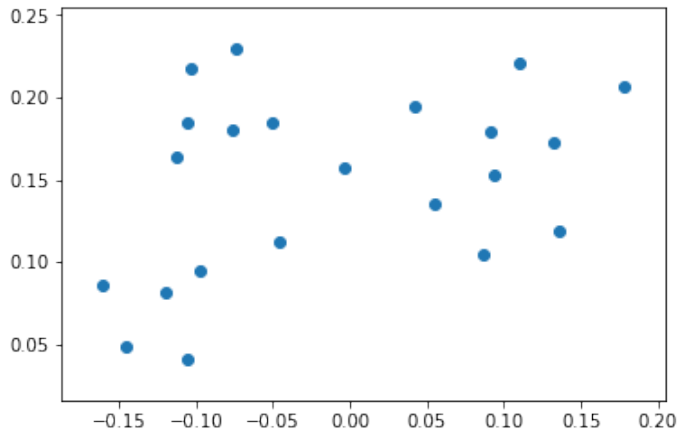
$$d(S_A, S_B) = \frac{1}{|S_A| \cdot |S_B|} \sum_{x_a \in S_A} \sum_{x_b \in S_B} d(x_a, x_b)$$

Disimilitud media



Agrupamiento Jerárquico

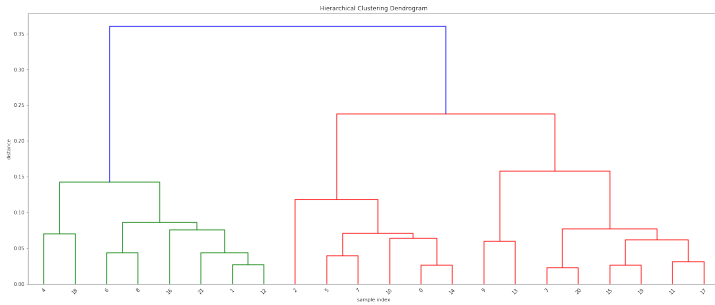
Aglomerativo



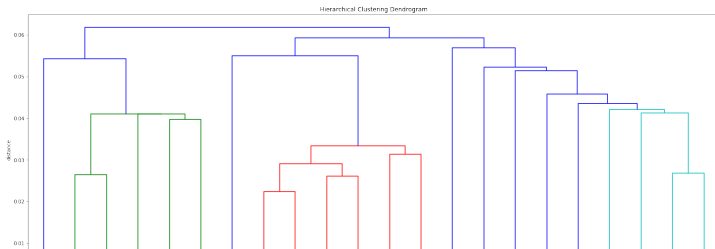
Agupamiento Jerárquico

Aglomerativo

Máximo



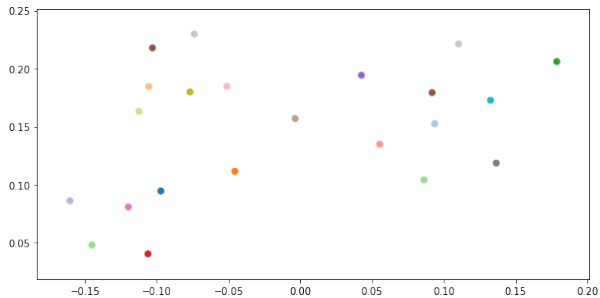
Mínimo



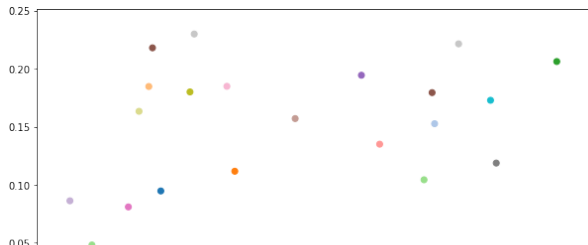
Agrupamiento Jerárquico

Aglomerativo

Máximo



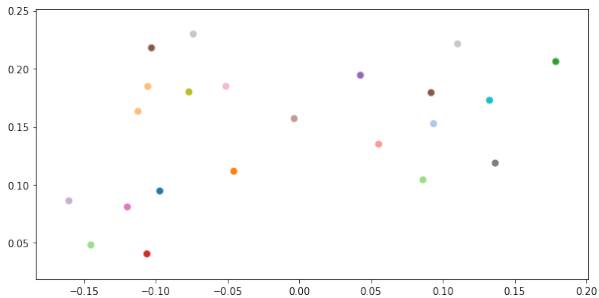
Mínimo



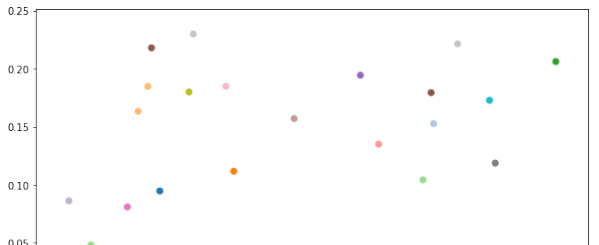
Agrupamiento Jerárquico

Aglomerativo

Máximo



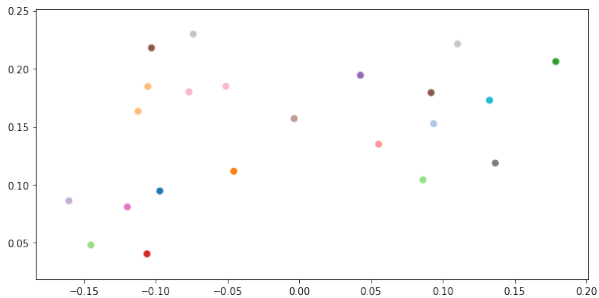
Mínimo



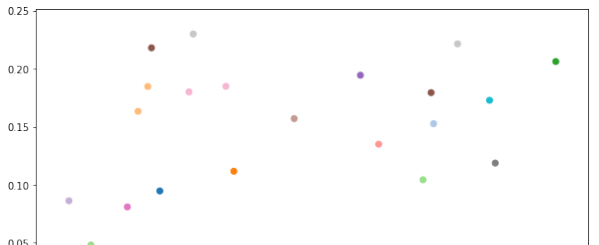
Agrupamiento Jerárquico

Aglomerativo

Máximo



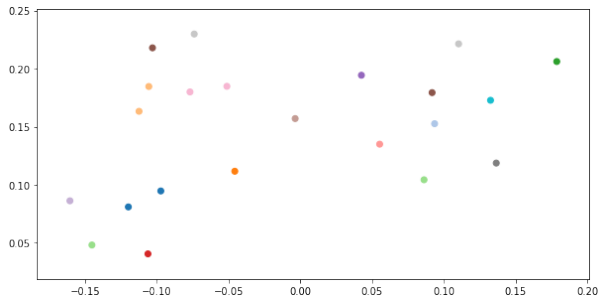
Mínimo



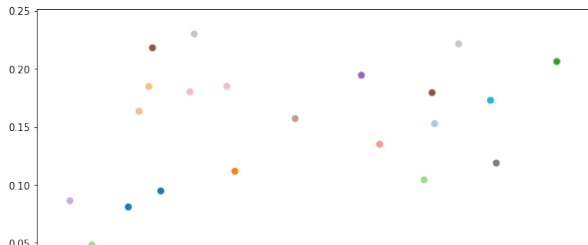
Agrupamiento Jerárquico

Aglomerativo

Máximo



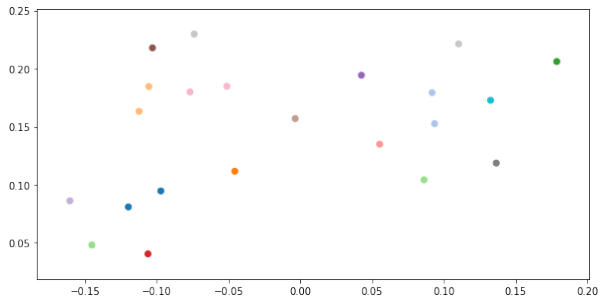
Mínimo



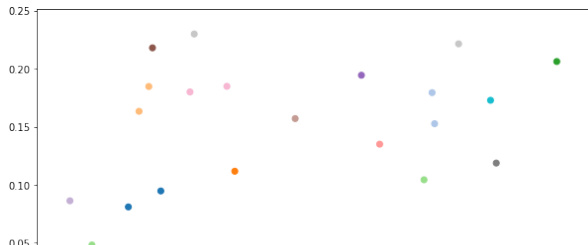
Agrupamiento Jerárquico

Aglomerativo

Máximo



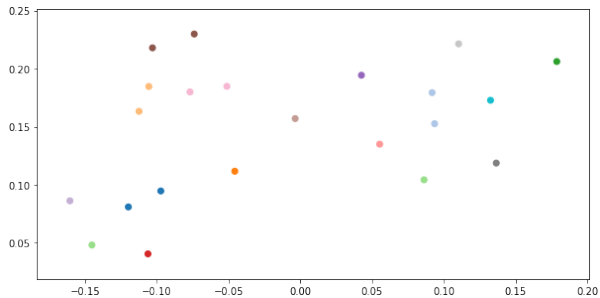
Mínimo



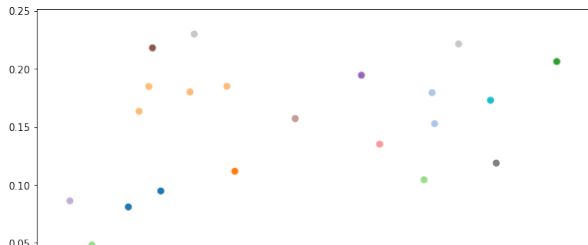
Agrupamiento Jerárquico

Aglomerativo

Máximo



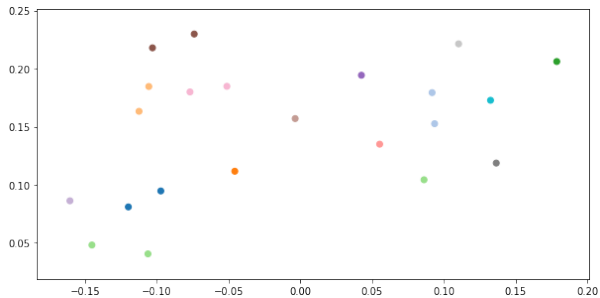
Mínimo



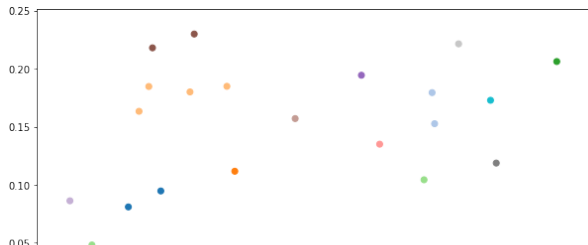
Agrupamiento Jerárquico

Aglomerativo

Máximo



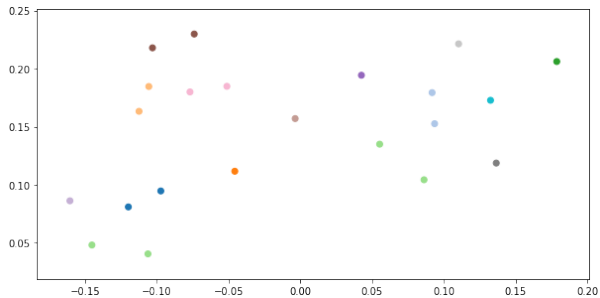
Mínimo



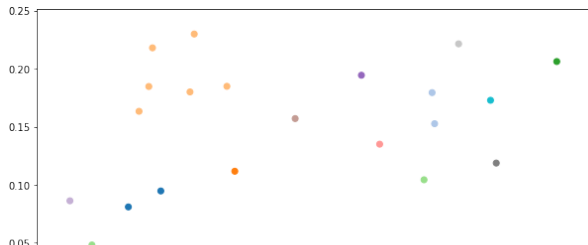
Agrupamiento Jerárquico

Aglomerativo

Máximo



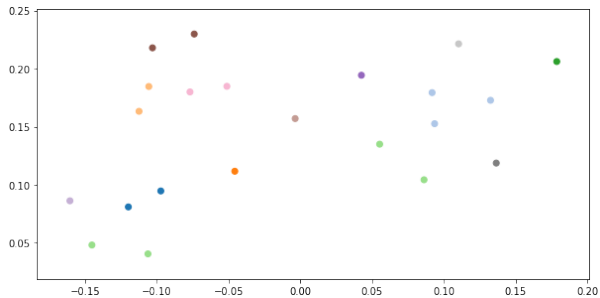
Mínimo



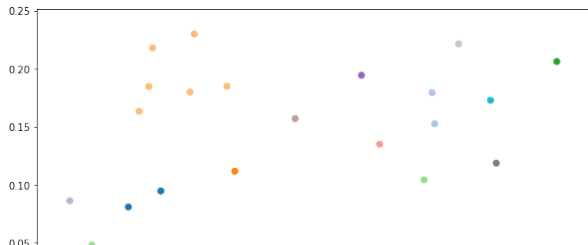
Agrupamiento Jerárquico

Aglomerativo

Máximo



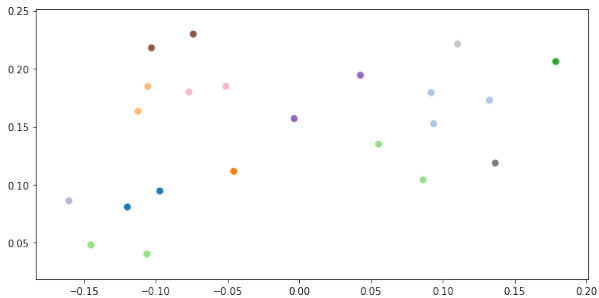
Mínimo



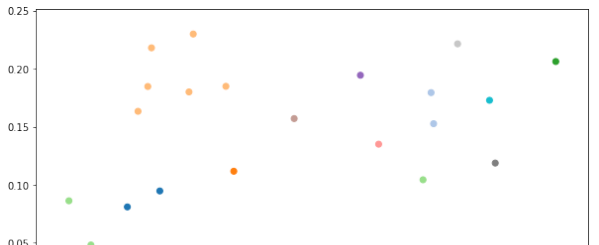
Agrupamiento Jerárquico

Aglomerativo

Máximo



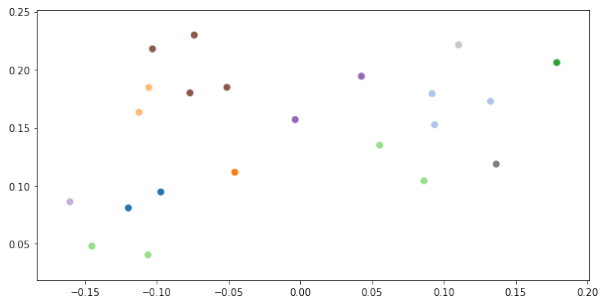
Mínimo



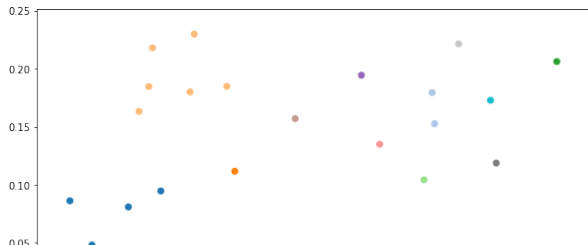
Agrupamiento Jerárquico

Aglomerativo

Máximo



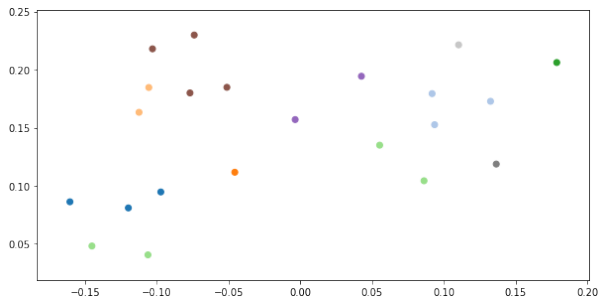
Mínimo



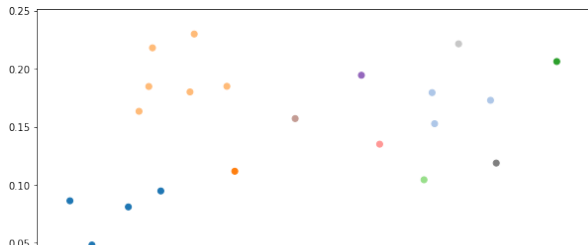
Agrupamiento Jerárquico

Aglomerativo

Máximo



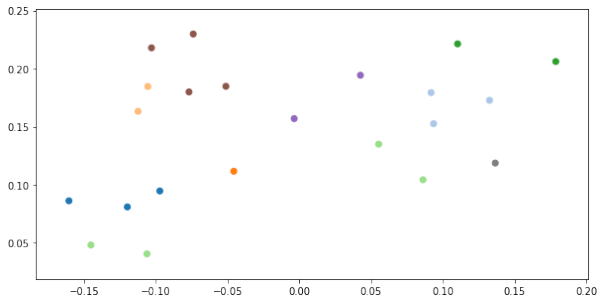
Mínimo



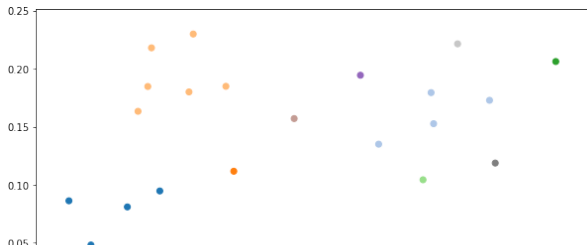
Agrupamiento Jerárquico

Aglomerativo

Máximo



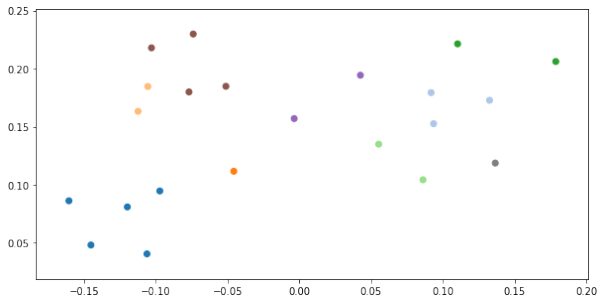
Mínimo



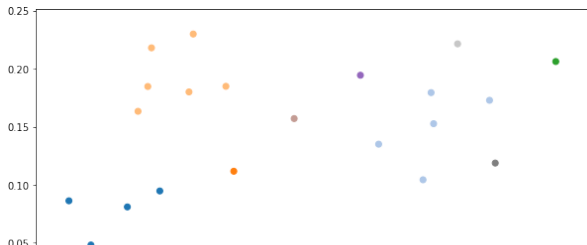
Agrupamiento Jerárquico

Aglomerativo

Máximo



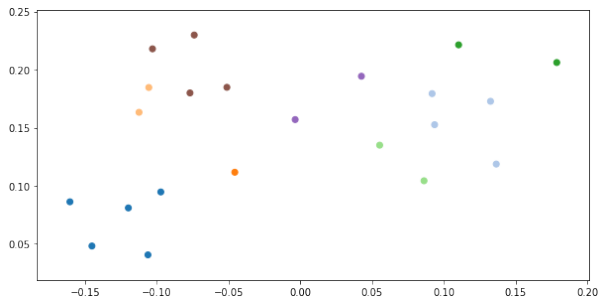
Mínimo



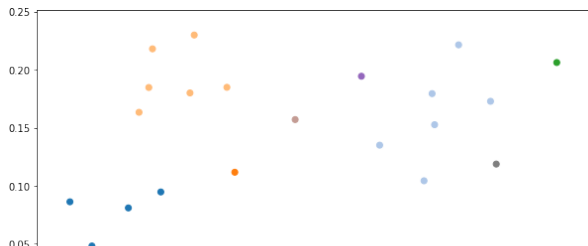
Agrupamiento Jerárquico

Aglomerativo

Máximo



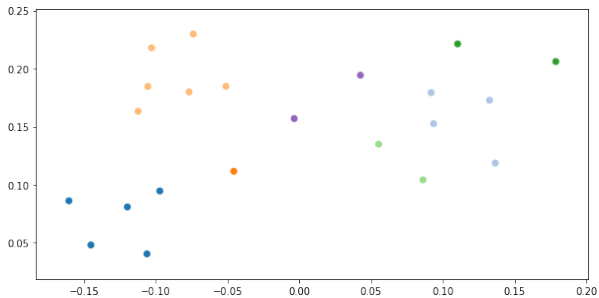
Mínimo



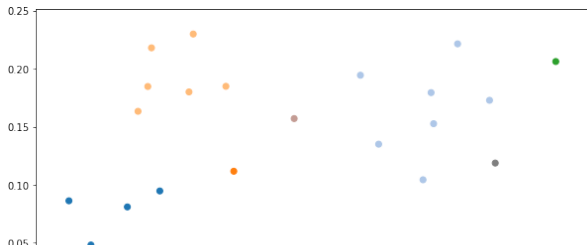
Agrupamiento Jerárquico

Aglomerativo

Máximo



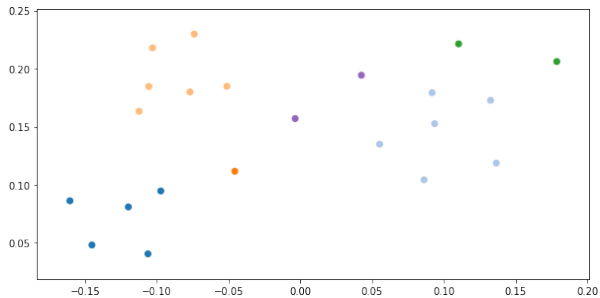
Mínimo



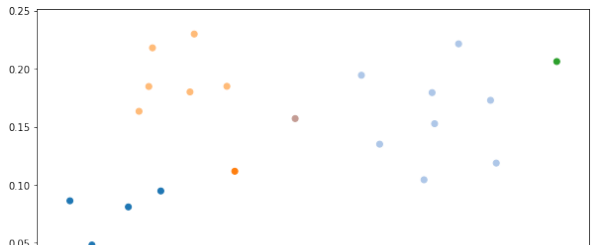
Agrupamiento Jerárquico

Aglomerativo

Máximo



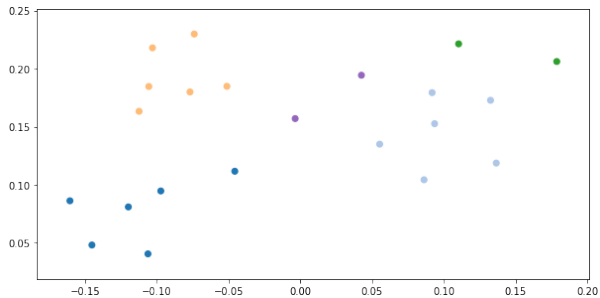
Mínimo



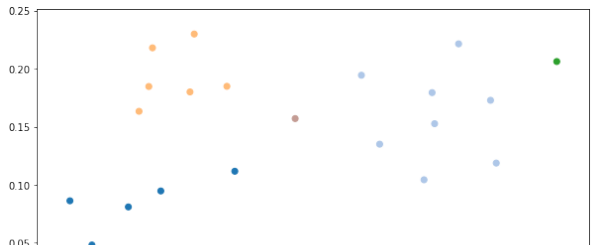
Agrupamiento Jerárquico

Aglomerativo

Máximo



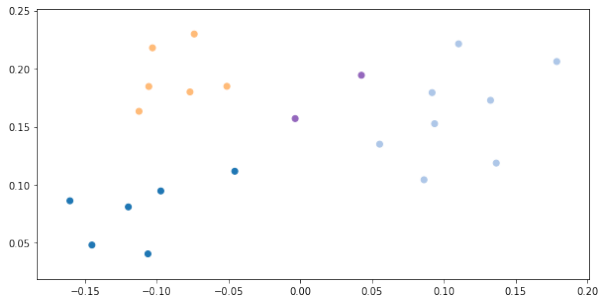
Mínimo



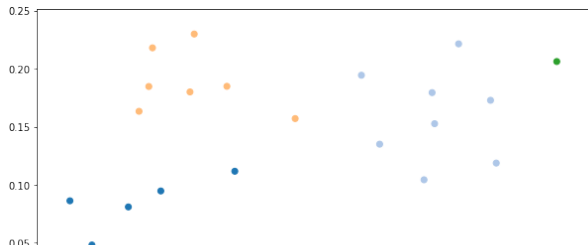
Agrupamiento Jerárquico

Aglomerativo

Máximo



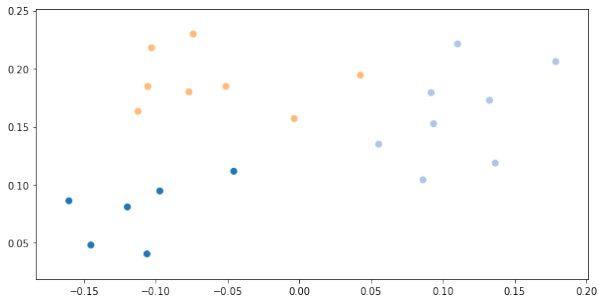
Mínimo



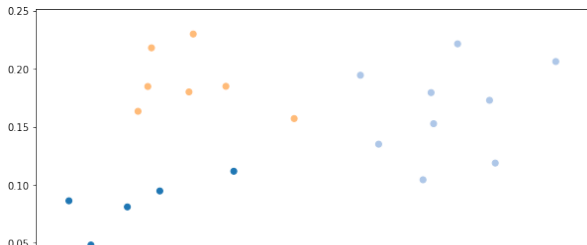
Agrupamiento Jerárquico

Aglomerativo

Máximo



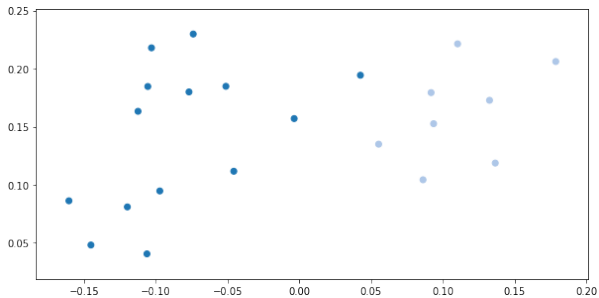
Mínimo



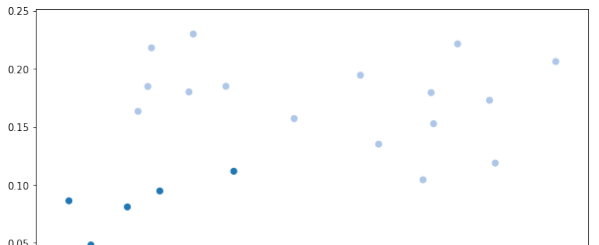
Agrupamiento Jerárquico

Aglomerativo

Máximo



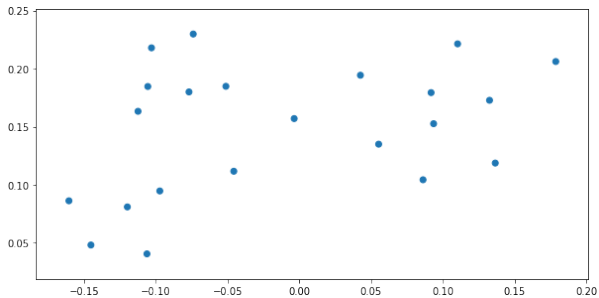
Mínimo



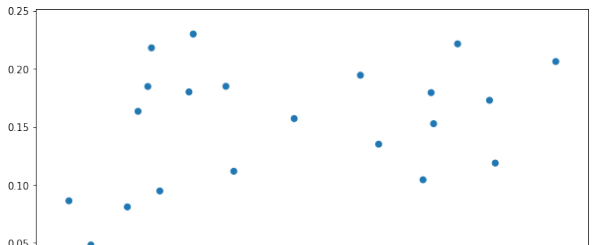
Agrupamiento Jerárquico

Aglomerativo

Máximo



Mínimo



Agrupamiento Jerárquico

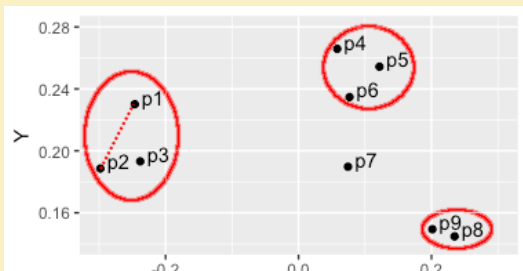
Aglomerativo

Tipos de clústeres obtenidos según criterio de unión

Definamos el concepto de **diámetro** de un clúster, S_K :

$$d(S_K) = \max_{x_i, x_j \in S_K} d(x_i, x_j)$$

Disimilitud máxima entre dos elementos del clúster S_K



Agrupamiento Jerárquico

Aglomerativo

Tipos de clústeres obtenidos según criterio de unión

Disimilitud mínima:

$$d(S_A, S_B) = \min_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud máxima:

$$d(S_A, S_B) = \max_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud media:

$$d(S_A, S_B) = \frac{1}{|S_A| + |S_B|} \sum_{x_a \in S_A} \sum_{x_b \in S_B} d(x_a, x_b)$$

Agrupamiento Jerárquico

Aglomerativo

Tipos de clústeres obtenidos según criterio de unión

Disimilitud mínima:

$$d(S_A, S_B) = \min_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

- Clústeres de ejemplos similares que pueden no formar una unidad compacta

Idea de la cadena

- El diámetro puede salir perjudicado

Disimilitud máxima:

$$d(S_A, S_B) = \max_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud media:

$$d(S_A, S_B) = \frac{1}{|S_A| + |S_B|} \sum_{x_a \in S_A} \sum_{x_b \in S_B} d(x_a, x_b)$$

Agrupamiento Jerárquico

Aglomerativo

Tipos de clústeres obtenidos según criterio de unión

Disimilitud mínima:

$$d(S_A, S_B) = \min_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud máxima:

$$d(S_A, S_B) = \max_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

- ▶ Clústeres compactos con diámetro reducido
- ▶ Se minimiza el diámetro, precisamente

La disimilitud máxima intraclúster es, tras la unión, el diámetro del nuevo clúster

- ▶ Puede separar en clústeres diferentes a ejemplos *muy* similares

Disimilitud media:

$$d(S_A, S_B) = \frac{1}{|S_A| + |S_B|} \sum_{x_a \in S_A} \sum_{x_b \in S_B} d(x_a, x_b)$$

Agrupamiento Jerárquico

Aglomerativo

Tipos de clústeres obtenidos según criterio de unión

Disimilitud mínima:

$$d(S_A, S_B) = \min_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud máxima:

$$d(S_A, S_B) = \max_{x_a \in S_A; x_b \in S_B} d(x_a, x_b)$$

Disimilitud media:

$$d(S_A, S_B) = \frac{1}{|S_A| + |S_B|} \sum_{x_a \in S_A} \sum_{x_b \in S_B} d(x_a, x_b)$$

- ▶ Escenario intermedio
- ▶ Clústeres relativamente compactos
- ▶ Junta elementos no necesariamente muy similares

Agrupamiento Jerárquico

Aglomerativo

Ventajas

- ▶ Intuitivo
- ▶ Conceptualmente sencillo
- ▶ Funciona con clústeres de diferente tamaño
- ▶ Una decisión de entrenamiento: criterio de unión
- ▶ Diferentes criterios
- ▶ Puede funcionar con diferentes medidas de distancia

Agrupamiento Jerárquico

Aglomerativo

Desventajas

- ▶ Lento
- ▶ Problemas al lidiar con clústeres de diferente densidad

Agrupamiento Jerárquico

Aglomerativo

Desventajas

- ▶ Lento
- ▶ Problemas al lidiar con clústeres de diferente densidad
- ▶ ¿Qué partición elegir?

Agrupamiento Jerárquico

Aglomerativo

Elección de una partición

Elegir una altura en la jerarquía donde cortar

- ▶ Número de clústeres concreto (fijando K)
- ▶ Máxima distancia en la unión de clústeres

