

## Série de TD 5

### Exercice 3

Considérer un repère euclidien avec les 9 points suivants :

$A_1(2, 10)$  ;  $A_2(4, 9)$  ;  $A_3(5, 8)$  ;  $A_4(5, 9)$  ;  $A_5(8, 5)$  ;  $A_6(8, 4)$  ;  $A_7(7, 5)$  ;  $A_8(6, 4)$  ;  $A_9(7, 4)$  . où  $(x, y)$  représente les coordonnées du point.

3) Avec  $r = 1$  unité, déterminer le voisinage de chacun des points donnés précédemment.

On va utiliser la distance de Manhattan.

|          | A1 (2,10) | A2 (4,9) | A3 (5,8) | A4 (5,9) | A5 (8,5) | A6 (8,4) | A7 (7,5) | A8 (6,4) | A9 (7,4) |
|----------|-----------|----------|----------|----------|----------|----------|----------|----------|----------|
| A1(2,10) | 0         | 3        | 5        | 4        | 11       | 12       | 10       | 10       | 11       |
| A2(4,9)  | 3         | 0        | 2        | 1        | 9        | 9        | 7        | 7        | 8        |
| A3(5,8)  | 5         | 2        | 0        | 1        | 6        | 7        | 5        | 5        | 6        |
| A4(5,9)  | 4         | 1        | 1        | 0        | 7        | 8        | 6        | 6        | 7        |
| A5(8,5)  | 11        | 9        | 6        | 7        | 0        | 1        | 1        | 3        | 2        |
| A6(8,4)  | 12        | 9        | 7        | 8        | 1        | 0        | 2        | 2        | 1        |
| A7(7,5)  | 10        | 7        | 5        | 6        | 1        | 2        | 0        | 2        | 1        |
| A8(6,4)  | 10        | 7        | 5        | 6        | 3        | 2        | 2        | 0        | 1        |
| A9(7,4)  | 11        | 8        | 6        | 7        | 2        | 1        | 1        | 1        | 0        |

voisinage( $A_1$ ) = { $A_1$ }

voisinage( $A_2$ ) = { $A_2$  ,  $A_4$ }

voisinage( $A_3$ ) = { $A_3$  ,  $A_4$ }

voisinage( $A_4$ ) = { $A_4$  ,  $A_2$  ,  $A_3$ }

voisinage( $A_5$ ) = { $A_5$  ,  $A_6$  ,  $A_7$ }

voisinage( $A_6$ ) = { $A_6$  ,  $A_5$  ,  $A_9$ }

voisinage( $A_7$ ) = { $A_7$  ,  $A_5$  ,  $A_9$ }

voisinage( $A_8$ ) = { $A_8$  ,  $A_9$ }

voisinage( $A_9$ ) = { $A_9$  ,  $A_6$  ,  $A_7$  ,  $A_8$ }

4) Avec  $MinPts = 2$ , appliquer l'algorithme DBSCAN sur les 9 points.

Etape 1: On choisit un point de D : soit  $A_1$

$A_1$  est marqué comme visité.

Etape 2: On calcule le voisinage de  $A_1 \Rightarrow PtsVoisins = \{A_1\}$

Etape 3:  $Taille(PtsVoisins) = 1 < MinPts \Rightarrow A_1.cluster = Bruit$

|       |                                |
|-------|--------------------------------|
| Bruit | Reste des points               |
| A1    | A2, A3, A4, A5, A6, A7, A8, A9 |

Étape 4: On choisit un autre point de D : soit  $A_2$

A2 est marqué comme visité.

| Bruit     | Reste des points                       |
|-----------|--|
| <b>A1</b> | <b>A2</b> , A3, A4, A5, A6, A7, A8, A9 |

Étape 2: On calcule le voisinage de A2  $\Rightarrow$  PtsVoisins = {A<sub>2</sub>, A<sub>4</sub>}

Étape 3: Taille(PtsVoisins) = 2  $\geq$  MinPts  $\Rightarrow$  C=1 et on étend le cluster.

Étape 4: A2.cluster = C1

| Bruit     | C1        | Reste des points           |
|-----------|-----------|----------------------------|
| <b>A1</b> | <b>A2</b> | A3, A4, A5, A6, A7, A8, A9 |

Étape 5: La boucle sur chaque point de PtsVoisins :

A2 déjà visité et déjà clusterisé.

A4 est marqué comme visité.

| Bruit     | C1        | Reste des points                   |
|-----------|-----------|------------------------------------|
| <b>A1</b> | <b>A2</b> | A3, <b>A4</b> , A5, A6, A7, A8, A9 |

Étape 6: On calcule le voisinage de A4  $\Rightarrow$  PtsVoisins' = {A<sub>2</sub>, A<sub>4</sub>, A<sub>3</sub>}

Étape 7: Taille(PtsVoisins') = 3  $\geq$  MinPts

$\Rightarrow$  PtsVoisins = {A<sub>2</sub>, A<sub>4</sub>}  $\cup$  {A<sub>2</sub>, A<sub>4</sub>, A<sub>3</sub>} = {A<sub>2</sub>, A<sub>4</sub>, A<sub>3</sub>}

A4 : n'appartient à aucun cluster : A4.cluster = C1

| Bruit     | C1            | Reste des points       |
|-----------|---------------|------------------------|
| <b>A1</b> | <b>A2, A4</b> | A3, A5, A6, A7, A8, A9 |

Étape 8: La boucle sur chaque point de PtsVoisins :

A3 : est marqué comme visité.

| Bruit     | C1            | Reste des points               |
|-----------|---------------|--------------------------------|
| <b>A1</b> | <b>A2, A4</b> | <b>A3</b> , A5, A6, A7, A8, A9 |

Étape 9: On calcule le voisinage de A3  $\Rightarrow$  PtsVoisins' = {A<sub>4</sub>, A<sub>3</sub>}

Étape 10: Taille(PtsVoisins') = 2  $\geq$  MinPts

$\Rightarrow$  PtsVoisins = {A<sub>2</sub>, A<sub>4</sub>, A<sub>3</sub>}  $\cup$  {A<sub>4</sub>, A<sub>3</sub>} = {A<sub>2</sub>, A<sub>4</sub>, A<sub>3</sub>}

A3 : déjà visité, mais n'appartient à aucun cluster : A3.cluster = C1

| Bruit     | C1                | Reste des points   |
|-----------|-------------------|--------------------|
| <b>A1</b> | <b>A2, A4, A3</b> | A5, A6, A7, A8, A9 |

-----Fin de la boucle dans `etendreCluster()` car nous avons parcourus toute la liste `PtsVoisin`.

Étape 11: On choisit un autre point de D : soit A5

A5 est marqué comme visité.

| Bruit     | C1                | Reste des points          |
|-----------|-------------------|---------------------------|
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6, A7, A8, A9</b> |

Étape 12: On calcule le voisinage de A5  $\Rightarrow$  `PtsVoisins` = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>}

Étape 13: `Taille(PtsVoisins)` = 3  $\geq$  `MinPts`  $\Rightarrow$  C=2 et on étend le cluster.

Étape 14: A5.cluster = C2

| Bruit     | C1                | C2        | Reste des points |
|-----------|-------------------|-----------|------------------|
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5</b> | A6, A7, A8, A9   |

Étape 15: La boucle sur chaque point de `PtsVoisins` :

A5 déjà visité et déjà clusterisé.

A6 est marqué comme visité.

| Bruit     | C1                | C2        | Reste des points      |
|-----------|-------------------|-----------|-----------------------|
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5</b> | <b>A6, A7, A8, A9</b> |

Étape 16: On calcule le voisinage de A6  $\Rightarrow$  `PtsVoisins'` = {A<sub>5</sub>, A<sub>6</sub>, A<sub>9</sub>}

Étape 17: `Taille(PtsVoisins')` = 3  $\geq$  `MinPts`

$\Rightarrow$  `PtsVoisins` = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>}  $\cup$  {A<sub>5</sub>, A<sub>6</sub>, A<sub>9</sub>} = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>, A<sub>9</sub>}

A6 : n'appartient à aucun cluster : A6.cluster = C2

| Bruit     | C1                | C2            | Reste des points |
|-----------|-------------------|---------------|------------------|
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6</b> | A7, A8, A9       |

A7 : est marqué comme visité.

| Bruit     | C1                | C2            | Reste des points |
|-----------|-------------------|---------------|------------------|
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6</b> | A7, A8, A9       |

Étape 18: On calcule le voisinage de A7  $\Rightarrow$  `PtsVoisins'` = {A<sub>5</sub>, A<sub>7</sub>, A<sub>9</sub>}

Étape 19: `Taille(PtsVoisins')` = 3  $\geq$  `MinPts`

$\Rightarrow$  `PtsVoisins` = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>, A<sub>9</sub>}  $\cup$  {A<sub>5</sub>, A<sub>7</sub>, A<sub>9</sub>} = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>, A<sub>9</sub>}

A7 : n'appartient à aucun cluster : A7.cluster = C2

| Bruit | C1 | C2 | Reste des points |
|-------|----|----|------------------|
|       |    |    |                  |

|           |                   |                   |               |
|-----------|-------------------|-------------------|---------------|
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6, A7</b> | <b>A8, A9</b> |
|-----------|-------------------|-------------------|---------------|

A9 : est marqué comme visité.

|           |                   |                   |                  |
|-----------|-------------------|-------------------|------------------|
| Bruit     | C1                | C2                | Reste des points |
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6, A7</b> | <b>A8, A9</b>    |

Étape 20: On calcule le voisinage de A9  $\Rightarrow$  PtsVoisins' = {A<sub>6</sub>, A<sub>7</sub>, A<sub>8</sub>, A<sub>9</sub>}

Étape 21: Taille(PtsVoisins') = 3  $\geq$  MinPts

$\Rightarrow$  PtsVoisins = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>, A<sub>9</sub>} U {A<sub>6</sub>, A<sub>7</sub>, A<sub>8</sub>, A<sub>9</sub>} = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>, A<sub>9</sub>, A<sub>8</sub>}

A9 : n'appartient à aucun cluster : A9.cluster = C2

|           |                   |                       |                  |
|-----------|-------------------|-----------------------|------------------|
| Bruit     | C1                | C2                    | Reste des points |
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6, A7, A9</b> | <b>A8</b>        |

A8 : est marqué comme visité.

|           |                   |                       |                  |
|-----------|-------------------|-----------------------|------------------|
| Bruit     | C1                | C2                    | Reste des points |
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6, A7, A9</b> | <b>A8</b>        |

Étape 22: On calcule le voisinage de A8  $\Rightarrow$  PtsVoisins' = {A<sub>8</sub>, A<sub>9</sub>}

Étape 23: Taille(PtsVoisins') = 2  $\geq$  MinPts

$\Rightarrow$  PtsVoisins = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>, A<sub>9</sub>, A<sub>8</sub>} U {A<sub>8</sub>, A<sub>9</sub>} = {A<sub>5</sub>, A<sub>6</sub>, A<sub>7</sub>, A<sub>9</sub>, A<sub>8</sub>}

A8 : n'appartient à aucun cluster : A8.cluster = C2

|           |                   |                           |                  |
|-----------|-------------------|---------------------------|------------------|
| Bruit     | C1                | C2                        | Reste des points |
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6, A7, A9, A8</b> | /                |

-----Fin de la boucle dans etendreCluster() car nous avons parcourus toute la liste PtsVoisin.

-----Fin de DBSCAN car nous avons parcouru tous le Dataset D.

|           |                   |                           |
|-----------|-------------------|---------------------------|
| Bruit     | Cluster : C1      | Cluster : C2              |
| <b>A1</b> | <b>A2, A4, A3</b> | <b>A5, A6, A7, A9, A8</b> |

6) Déterminer le bruit engendré par DBSCAN.

Le bruit engendré par DBSCAN est le point A1.

## Exercice 4

Méthode agglomérative dite AGNES (AGglomerative NESTing).

2) Appliquer l'algorithme sur les 9 points  $A_1(2, 10)$  ;  $A_2(4, 9)$  ;  $A_3(5, 8)$  ;  $A_4(5, 9)$  ;  $A_5(8, 5)$  ;  $A_6(8, 4)$  ;  $A_7(7, 5)$  ;  $A_8(6, 4)$  ;  $A_9(7, 4)$ .

On choisit la méthode : Centroid linkage: (Lien centroïde) La distance entre deux clusters est définie comme la distance entre le centroïde du cluster C1 et le centroïde du cluster C2. et comme formule de calcul de distance entre 2 instances : la distance euclidienne.

**Etape 1 :** affecter chaque point à un cluster différent :

|               |              |              |              |              |              |              |              |              |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A1<br>(2, 10) | A2<br>(4, 9) | A3<br>(5, 8) | A4<br>(5, 9) | A5<br>(8, 5) | A6<br>(8, 4) | A7<br>(7, 5) | A8<br>(6, 4) | A9<br>(7, 4) |
| C1<br>(2, 10) | C2<br>(4, 9) | C3<br>(5, 8) | C4<br>(5, 9) | C5<br>(8, 5) | C6<br>(8, 4) | C7<br>(7, 5) | C8<br>(6, 4) | C9<br>(7, 4) |

**Etape 2 :** calculer la distance entre chaque pair de clusters

|          | C1<br>(2,10) | C2<br>(4,9) | C3<br>(5,8) | C4<br>(5,9) | C5<br>(8,5) | C6<br>(8,4) | C7<br>(7,5) | C8<br>(6,4) | C9<br>(7,4) |
|----------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| C1(2,10) | 0            | $\sqrt{5}$  | $\sqrt{13}$ | $\sqrt{10}$ | $\sqrt{61}$ | $\sqrt{72}$ | $\sqrt{50}$ | $\sqrt{52}$ | $\sqrt{61}$ |
| C2(4,9)  |              | 0           | $\sqrt{2}$  | $\sqrt{1}$  | $\sqrt{32}$ | $\sqrt{41}$ | $\sqrt{25}$ | $\sqrt{29}$ | $\sqrt{34}$ |
| C3(5,8)  |              |             | 0           | $\sqrt{1}$  | $\sqrt{18}$ | $\sqrt{25}$ | $\sqrt{13}$ | $\sqrt{17}$ | $\sqrt{20}$ |
| C4(5,9)  |              |             |             | 0           | $\sqrt{25}$ | $\sqrt{34}$ | $\sqrt{20}$ | $\sqrt{26}$ | $\sqrt{29}$ |
| C5(8,5)  |              |             |             |             | 0           | $\sqrt{1}$  | $\sqrt{1}$  | $\sqrt{5}$  | $\sqrt{2}$  |
| C6(8,4)  |              |             |             |             |             | 0           | $\sqrt{2}$  | $\sqrt{4}$  | $\sqrt{1}$  |
| C7(7,5)  |              |             |             |             |             |             | 0           | $\sqrt{2}$  | $\sqrt{1}$  |
| C8(6,4)  |              |             |             |             |             |             |             | 0           | $\sqrt{1}$  |

**Etape 3 :** fusionner les paires de distance minimale → nouveaux clusters:

Distance minimale = 1 ( $\sqrt{1}$ )

Les paires concernées sont : C2-C4 ; C3-C4 ; C5-C6 ; C5-A7 ; C6-C9 ; C7-C9 ; C8-C9

Donc on va fusionner C2-C4 ; C5-C6 ; C7-C9 ;

**Etape 4 :** calculer les nouveaux centroïdes :

|           |          |                     |                     |          |                     |
|-----------|----------|---------------------|---------------------|----------|---------------------|
| A1(2, 10) | A3(5, 8) | A2(4, 9) ; A4(5, 9) | A5(8, 5) ; A6(8, 4) | A8(6, 4) | A7(7, 5) ; A9(7, 4) |
| C1(2, 10) | C2(5, 8) | C3(4.5, 9)          | C4(8, 4.5)          | C5(6,4)  | C6(7, 4.5)          |

-----Et on recommence les étapes de 2 à 4 jusqu'à former 1 seul cluster

**Etape 2'** : calculer la distance entre chaque pair de clusters

|              | C1 (2 , 10) | C2 (5 , 8)  | C3 (4.5 , 9)  | C4 (8 , 4.5)   | C5 (6 , 4)     | C6 (7 , 4.5)   |
|--------------|-------------|-------------|---------------|----------------|----------------|----------------|
| C1 (2 , 10)  | 0           | $\sqrt{13}$ | $\sqrt{7.25}$ | $\sqrt{66.25}$ | $\sqrt{52}$    | $\sqrt{55.25}$ |
| C2 (5 , 8)   |             | 0           | $\sqrt{1.25}$ | $\sqrt{21.25}$ | $\sqrt{17}$    | $\sqrt{16.25}$ |
| C3 (4.5 , 9) |             |             | 0             | $\sqrt{32.5}$  | $\sqrt{27.25}$ | $\sqrt{26.5}$  |
| C4 (8 , 4.5) |             |             |               | 0              | $\sqrt{4.25}$  | $\sqrt{1}$     |
| C5 (6 , 4)   |             |             |               |                | 0              | $\sqrt{1.25}$  |

**Etape 3'** : fusionner les paires de distance minimale → nouveaux clusters:

Distance minimale = 1 ( $\sqrt{1}$ )

La paire concernée est : C4-C6

**Etape 4'** : calculer les nouveaux centroïdes :

|           |          |                      |   |          |
|-----------|----------|----------------------|---|----------|
| A1(2, 10) | A3(5, 8) | A2 (4, 9) ; A4(5, 9) | A5(8, 5) ; A6(8, 4) ; A7(7, 5) ; A9(7, 4) | A8(6, 4) |
| C1(2, 10) | C2(5, 8) | C3(4.5 , 9)          | C4(7.5 , 4.5)                             | C5(6,4)  |

**Etape 2''** : calculer la distance entre chaque pair de clusters

|                | C1 (2 , 10) | C2 (5 , 8)  | C3 (4.5 , 9)  | C4 (7.5 , 4.5) | C5 (6 , 4)     |
|----------------|-------------|-------------|---------------|----------------|----------------|
| C1 (2 , 10)    | 0           | $\sqrt{13}$ | $\sqrt{7.25}$ | $\sqrt{60.5}$  | $\sqrt{52}$    |
| C2 (5 , 8)     |             | 0           | $\sqrt{1.25}$ | $\sqrt{18.5}$  | $\sqrt{17}$    |
| C3 (4.5 , 9)   |             |             | 0             | $\sqrt{29.25}$ | $\sqrt{27.25}$ |
| C4 (7.5 , 4.5) |             |             |               | 0              | $\sqrt{2.5}$   |

**Etape 3''** : fusionner les paires de distance minimale → nouveaux clusters:

Distance minimale =  $\sqrt{1.25}$

Le paire concernée est : C2-C3

**Etape 4''** : calculer les nouveaux centroïdes :

|           |                                 |   |          |
|-----------|---------------------------------|---|----------|
| A1(2, 10) | A3(5, 8) ; A2 (4, 9) ; A4(5, 9) | A5(8, 5) ; A6(8, 4) ; A7(7, 5) ; A9(7, 4) | A8(6, 4) |
| C1(2, 10) | C2(4.66, 8.66)                  | C3(7.5 , 4.5)                             | C4(6,4)  |

**Etape 2'''** : calculer la distance entre chaque pair de clusters

|                  | C1 (2 , 10) | C2 (4.66 , 8.66) | C3 (7.5 , 4.5) | C4 (6 , 4)     |
|------------------|-------------|------------------|----------------|----------------|
| C1 (2 , 10)      | 0           | $\sqrt{8.87}$    | $\sqrt{60.5}$  | $\sqrt{52}$    |
| C2 (4.66 , 8.66) |             | 0                | $\sqrt{25.37}$ | $\sqrt{23.51}$ |

|                       |  |  |   |              |
|-----------------------|--|--|---|--------------|
| <b>C3 (7.5 , 4.5)</b> |  |  | 0 | $\sqrt{2.5}$ |
|-----------------------|--|--|---|--------------|

**Etape 3'''** : fusionner les paires de distance minimale → nouveaux clusters:

Distance minimale =  $\sqrt{2.5}$

Le paire concernée est : C3-C4

**Etape 4'''** : calculer les nouveaux centroïdes :

|           |                                 |  |
|-----------|---------------------------------|--|
| A1(2, 10) | A3(5, 8) ; A2 (4, 9) ; A4(5, 9) | A5(8, 5) ; A6(8, 4) ; A7(7, 5) ; A9(7, 4) ; A8(6, 4) |
| C1(2, 10) | C2(4.66, 8.66)                  | C3(7.2 , 4.4)  |

**Etape 2''''** : calculer la distance entre chaque pair de clusters

|                         |                    |                         |                       |
|-------------------------|--------------------|-------------------------|-----------------------|
|                         | <b>C1 (2 , 10)</b> | <b>C2 (4.66 , 8.66)</b> | <b>C3 (7.2 , 4.4)</b> |
| <b>C1 (2 , 10)</b>      | 0                  | $\sqrt{8.87}$           | $\sqrt{58.4}$         |
| <b>C2 (4.66 , 8.66)</b> |                    | 0                       | $\sqrt{24.6}$         |

**Etape 3''''** : fusionner les paires de distance minimale → nouveaux clusters:

Distance minimale =  $\sqrt{8.87}$

Le paire concernée est : C1-C2

**Etape 4''''** : calculer les nouveaux centroïdes :

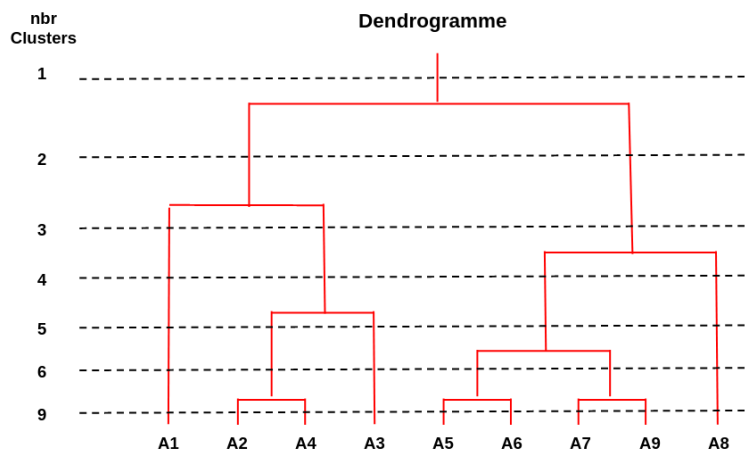
|   |  |
|---|--|
| A1(2, 10) ; A3(5, 8) ; A2 (4, 9) ; A4(5, 9) | A5(8, 5) ; A6(8, 4) ; A7(7, 5) ; A9(7, 4) ; A8(6, 4) |
| C1(4 , 9)                                   | C2(7.2 , 4.4)  |

**il ne reste que 2 clusters, on les fusionne pour donner le cluster global.**

**Fin du clustering.**

4) Dessiner le dendrogramme des clusters.

Ci-dessous le dendrogramme avec **le nombre de cluster** représenté sur la gauche



5) Spécifier les clusters qui sont à une distance de séparation supérieure à  $\sqrt{5}$ .

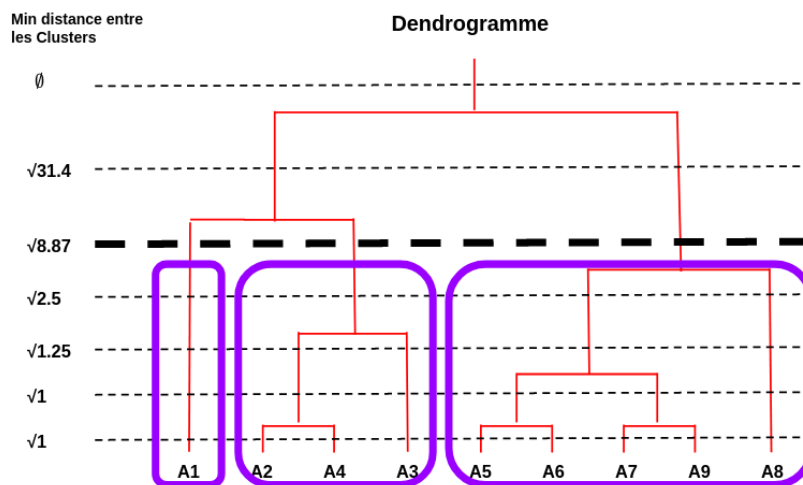
Cela signifie qu'on coupe là où la distance minimal entre 2 clusters était supérieure à  $\sqrt{5}$  ce qui correspond à l'avant dernière étape de fusion ou la distance minimale entre clusters est de  $\sqrt{8.87}$  (donc on aura 3 clusters)

**C1 = {A1}**

**C2 = {A2 , A3 , A4}**

**C3 = {A5 , A6 , A7 , A8 , A9}**

Ci-dessous le dendrogramme avec la **distance minimale** entre clusters représenté sur la gauche.



6) Appliquer l'algorithme DIANA sur les 9 points  $A_1(2, 10)$  ;  $A_2(4, 9)$  ;  $A_3(5, 8)$  ;  $A_4(5, 9)$  ;  $A_5(8, 5)$  ;  $A_6(8, 4)$  ;  $A_7(7, 5)$  ;  $A_8(6, 4)$  ;  $A_9(7, 4)$ .

On choisit la formule de calcul de distance entre 2 instances : la distance de manhattan.

**Etape 1 :** affecter tous les points à un cluster global C1 :

|               |              |              |              |              |              |              |              |              |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A1<br>(2, 10) | A2<br>(4, 9) | A3<br>(5, 8) | A4<br>(5, 9) | A5<br>(8, 5) | A6<br>(8, 4) | A7<br>(7, 5) | A8<br>(6, 4) | A9<br>(7, 4) |
| C1            | C1           | C1           | C1           | C1           | C1           | C1           | C1           | C1           |

**Etape 2 :** Calculer l'instance la plus dissimilaire de C1 :

|          | A1<br>(2,10) | A2<br>(4,9) | A3<br>(5,8) | A4<br>(5,9) | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) | A8<br>(6,4) | A9<br>(7,4) |
|----------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| A1(2,10) | 0            | 3           | 5           | 4           | 11          | 12          | 10          | 10          | 11          |
| A2(4,9)  | 3            | 0           | 2           | 1           | 9           | 9           | 7           | 7           | 8           |
| A3(5,8)  | 5            | 2           | 0           | 1           | 6           | 7           | 5           | 5           | 6           |



|                 |           |           |           |           |           |           |           |           |           |
|-----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <b>A4 (5,9)</b> | 4         | 1         | 1         | 0         | 7         | 8         | 6         | 6         | 7         |
| <b>A5 (8,5)</b> | 11        | 9         | 6         | 7         | 0         | 1         | 1         | 3         | 2         |
| <b>A6 (8,4)</b> | 12        | 9         | 7         | 8         | 1         | 0         | 2         | 2         | 1         |
| <b>A7 (7,5)</b> | 10        | 7         | 5         | 6         | 1         | 2         | 0         | 2         | 1         |
| <b>A8 (6,4)</b> | 10        | 7         | 5         | 6         | 3         | 2         | 2         | 0         | 1         |
| <b>A9 (7,4)</b> | 11        | 8         | 6         | 7         | 2         | 1         | 1         | 1         | 0         |
|                 | <b>66</b> | <b>46</b> | <b>37</b> | <b>40</b> | <b>40</b> | <b>42</b> | <b>34</b> | <b>36</b> | <b>37</b> |

Point avec la plus grande dissimilarité est  $A1(2,10) = 66$

Mettre A1 dans un new cluster C2

|               |              |              |              |              |              |              |              |              |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A1<br>(2, 10) | A2<br>(4, 9) | A3<br>(5, 8) | A4<br>(5, 9) | A5<br>(8, 5) | A6<br>(8, 4) | A7<br>(7, 5) | A8<br>(6, 4) | A9<br>(7, 4) |
| <b>C2</b>     | C1           | C1           | C1           | C1           | C1           | C1           | C1           | C1           |

**Etape 3 :** Cinder le cluster C1 entre C1 et C2

|           |                          | <b>A1<br/>(2,10)</b> | <b>A2<br/>(4,9)</b> | <b>A3<br/>(5,8)</b> | <b>A4<br/>(5,9)</b> | <b>A5<br/>(8,5)</b> | <b>A6<br/>(8,4)</b> | <b>A7<br/>(7,5)</b> | <b>A8<br/>(6,4)</b> | <b>A9<br/>(7,4)</b> |
|-----------|--------------------------|----------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| <b>C2</b> | <b>A1(2,10)</b>          | 0                    | 3                   | 5                   | 4                   | 11                  | 12                  | 10                  | 10                  | 11                  |
|           | <b>Dist(C2)</b>          | -                    | <b>3</b>            | <b>5</b>            | <b>4</b>            | <b>11</b>           | <b>12</b>           | <b>10</b>           | <b>10</b>           | <b>11</b>           |
| <b>C1</b> | <b>A2 (4,9)</b>          | 3                    | 0                   | 2                   | 1                   | 9                   | 9                   | 7                   | 7                   | 8                   |
|           | <b>A3 (5,8)</b>          | 5                    | 2                   | 0                   | 1                   | 6                   | 7                   | 5                   | 5                   | 6                   |
|           | <b>A4 (5,9)</b>          | 4                    | 1                   | 1                   | 0                   | 7                   | 8                   | 6                   | 6                   | 7                   |
|           | <b>A5 (8,5)</b>          | 11                   | 9                   | 6                   | 7                   | 0                   | 1                   | 1                   | 3                   | 2                   |
|           | <b>A6 (8,4)</b>          | 12                   | 9                   | 7                   | 8                   | 1                   | 0                   | 2                   | 2                   | 1                   |
|           | <b>A7 (7,5)</b>          | 10                   | 7                   | 5                   | 6                   | 1                   | 2                   | 0                   | 2                   | 1                   |
|           | <b>A8 (6,4)</b>          | 10                   | 7                   | 5                   | 6                   | 3                   | 2                   | 2                   | 0                   | 1                   |
|           | <b>A9 (7,4)</b>          | 11                   | 8                   | 6                   | 7                   | 2                   | 1                   | 1                   | 1                   | 0                   |
|           | <b>Dist(C1)</b>          | -                    | <b>43</b>           | <b>32</b>           | <b>36</b>           | <b>29</b>           | <b>30</b>           | <b>24</b>           | <b>26</b>           | <b>26</b>           |
|           | <b>dist(C1)-dist(C2)</b> | -                    | <b>40</b>           | <b>27</b>           | <b>32</b>           | <b>18</b>           | <b>18</b>           | <b>14</b>           | <b>16</b>           | <b>15</b>           |

L'instance ayant la plus forte différence  $\text{dist}(C1)-\text{dist}(C2)$  est A2 avec une différence de 40.

Migrer A2 vers C2

|               |              |              |              |              |              |              |              |              |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A1<br>(2, 10) | A2<br>(4, 9) | A3<br>(5, 8) | A4<br>(5, 9) | A5<br>(8, 5) | A6<br>(8, 4) | A7<br>(7, 5) | A8<br>(6, 4) | A9<br>(7, 4) |
| <b>C2</b>     | <b>C2</b>    | C1           | C1           | C1           | C1           | C1           | C1           | C1           |

|                   |          | A1<br>(2,10) | A2<br>(4,9) | A3<br>(5,8) | A4<br>(5,9) | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) | A8<br>(6,4) | A9<br>(7,4) |
|-------------------|----------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| C2                | A1(2,10) | 0            | 3           | 5           | 4           | 11          | 12          | 10          | 10          | 11          |
|                   | A2(4,9)  | 3            | 0           | 2           | 1           | 9           | 9           | 7           | 7           | 8           |
|                   | Dist(C2) | -            | -           | 7           | 5           | 20          | 21          | 17          | 17          | 19          |
| C1                | A3(5,8)  | 5            | 2           | 0           | 1           | 6           | 7           | 5           | 5           | 6           |
|                   | A4(5,9)  | 4            | 1           | 1           | 0           | 7           | 8           | 6           | 6           | 7           |
|                   | A5(8,5)  | 11           | 9           | 6           | 7           | 0           | 1           | 1           | 3           | 2           |
|                   | A6(8,4)  | 12           | 9           | 7           | 8           | 1           | 0           | 2           | 2           | 1           |
|                   | A7(7,5)  | 10           | 7           | 5           | 6           | 1           | 2           | 0           | 2           | 1           |
|                   | A8(6,4)  | 10           | 7           | 5           | 6           | 3           | 2           | 2           | 0           | 1           |
|                   | A9(7,4)  | 11           | 8           | 6           | 7           | 2           | 1           | 1           | 1           | 0           |
| Dist(C1)          | -        | -            | 30          | 35          | 20          | 21          | 17          | 19          | 18          |             |
| dist(C1)-dist(C2) | -        | -            | 23          | 30          | 0           | 0           | 0           | 2           | -1          |             |

L'instance ayant la plus forte différence  $\text{dist}(C1) - \text{dist}(C2)$  est A4 avec une différence de 30.  
Migrer A4 vers C2

| A1<br>(2, 10) | A2<br>(4, 9) | A3<br>(5, 8) | A4<br>(5, 9) | A5<br>(8, 5) | A6<br>(8, 4) | A7<br>(7, 5) | A8<br>(6, 4) | A9<br>(7, 4) |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| C2            | C2           | C1           | C2           | C1           | C1           | C1           | C1           | C1           |

|                   |          | A1<br>(2,10) | A2<br>(4,9) | A3<br>(5,8) | A4<br>(5,9) | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) | A8<br>(6,4) | A9<br>(7,4) |
|-------------------|----------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| C2                | A1(2,10) | 0            | 3           | 5           | 4           | 11          | 12          | 10          | 10          | 11          |
|                   | A2(4,9)  | 3            | 0           | 2           | 1           | 9           | 9           | 7           | 7           | 8           |
|                   | A4(5,9)  | 4            | 1           | 1           | 0           | 7           | 8           | 6           | 6           | 7           |
| Dist(C2)          | -        | -            | 8           | -           | 27          | 29          | 23          | 23          | 26          |             |
| C1                | A3(5,8)  | 5            | 2           | 0           | 1           | 6           | 7           | 5           | 5           | 6           |
|                   | A5(8,5)  | 11           | 9           | 6           | 7           | 0           | 1           | 1           | 3           | 2           |
|                   | A6(8,4)  | 12           | 9           | 7           | 8           | 1           | 0           | 2           | 2           | 1           |
|                   | A7(7,5)  | 10           | 7           | 5           | 6           | 1           | 2           | 0           | 2           | 1           |
|                   | A8(6,4)  | 10           | 7           | 5           | 6           | 3           | 2           | 2           | 0           | 1           |
|                   | A9(7,4)  | 11           | 8           | 6           | 7           | 2           | 1           | 1           | 1           | 0           |
| Dist(C1)          | -        | -            | 29          | -           | 13          | 13          | 11          | 13          | 11          |             |
| dist(C1)-dist(C2) | -        | -            | 21          | -           | -14         | -16         | -12         | -10         | -15         |             |

L'instance ayant la plus forte différence  $\text{dist}(C1) - \text{dist}(C2)$  est A3 avec une différence de 21.  
Migrer A3 vers C2

|               |              |              |              |              |              |              |              |              |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| A1<br>(2, 10) | A2<br>(4, 9) | A3<br>(5, 8) | A4<br>(5, 9) | A5<br>(8, 5) | A6<br>(8, 4) | A7<br>(7, 5) | A8<br>(6, 4) | A9<br>(7, 4) |
| C2            | C2           | C2           | C2           | C1           | C1           | C1           | C1           | C1           |

|    |                   | A1<br>(2,10) | A2<br>(4,9) | A3<br>(5,8) | A4<br>(5,9) | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) | A8<br>(6,4) | A9<br>(7,4) |
|----|-------------------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| C2 | A1(2,10)          | 0            | 3           | 5           | 4           | 11          | 12          | 10          | 10          | 11          |
|    | A2(4,9)           | 3            | 0           | 2           | 1           | 9           | 9           | 7           | 7           | 8           |
|    | A4(5,9)           | 4            | 1           | 1           | 0           | 7           | 8           | 6           | 6           | 7           |
|    | A3(5,8)           | 5            | 2           | 0           | 1           | 6           | 7           | 5           | 5           | 6           |
|    | Dist(C2)          | -            | -           | -           | -           | 33          | 36          | 28          | 28          | 32          |
| C1 | A5(8,5)           | 11           | 9           | 6           | 7           | 0           | 1           | 1           | 3           | 2           |
|    | A6(8,4)           | 12           | 9           | 7           | 8           | 1           | 0           | 2           | 2           | 1           |
|    | A7(7,5)           | 10           | 7           | 5           | 6           | 1           | 2           | 0           | 2           | 1           |
|    | A8(6,4)           | 10           | 7           | 5           | 6           | 3           | 2           | 2           | 0           | 1           |
|    | A9(7,4)           | 11           | 8           | 6           | 7           | 2           | 1           | 1           | 1           | 0           |
|    | Dist(C1)          | -            | -           | -           | -           | 7           | 6           | 6           | 8           | 5           |
|    | dist(C1)-dist(C2) | -            | -           | -           | -           | -26         | -30         | -22         | -20         | -27         |

Toutes les différences de distances sont négatives ou nulles. Fin de la séparation.  
Réitérer le processus sur les deux clusters obtenus C1 et C2.

**Etape 2 :** Calculer l'instance la plus dissimilaire de C1 :

|    |          | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) | A8<br>(6,4) | A9<br>(7,4) |
|----|----------|-------------|-------------|-------------|-------------|-------------|
| C1 | A5(8,5)  | 0           | 1           | 1           | 3           | 2           |
|    | A6(8,4)  | 1           | 0           | 2           | 2           | 1           |
|    | A7(7,5)  | 1           | 2           | 0           | 2           | 1           |
|    | A8(6,4)  | 3           | 2           | 2           | 0           | 1           |
|    | A9(7,4)  | 2           | 1           | 1           | 1           | 0           |
|    | Dist(C1) | 7           | 6           | 6           | 8           | 5           |

**Etape 3 :** Cinder le cluster C1 entre C1 et C3

Point avec la plus grande dissimilarité est  $A8(6,4) = 8$

Mettre A8 dans un new cluster C3

|         |         |         |         |         |
|---------|---------|---------|---------|---------|
| A5(8,5) | A6(8,4) | A7(7,5) | A8(6,4) | A9(7,4) |
| C1      | C1      | C1      | C3      | C1      |

|                   |          | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) | A8<br>(6,4) | A9<br>(7,4) |
|-------------------|----------|-------------|-------------|-------------|-------------|-------------|
| C3                | A8 (6,4) | 3           | 2           | 2           | 0           | 1           |
| C1                | Dist(C3) | 3           | 2           | 2           | -           | 1           |
|                   | A5 (8,5) | 0           | 1           | 1           | 3           | 2           |
|                   | A6 (8,4) | 1           | 0           | 2           | 2           | 1           |
|                   | A7 (7,5) | 1           | 2           | 0           | 2           | 1           |
|                   | A9 (7,4) | 2           | 1           | 1           | 1           | 0           |
|                   | Dist(C1) | 4           | 4           | 4           | -           | 4           |
| dist(C1)-dist(C3) |          | 1           | 2           | 2           | -           | 3           |

L'instance ayant la plus forte différence  $\text{dist}(C1)-\text{dist}(C3)$  est A9 avec une différence de 3.  
Migrer A9 vers C3

| A5 (8,5) | A6 (8,4) | A7 (7,5) | A8 (6,4) | A9 (7,4) |
|----------|----------|----------|----------|----------|
| C1       | C1       | C1       | C3       | C3       |

|                   |          | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) | A8<br>(6,4) | A9<br>(7,4) |
|-------------------|----------|-------------|-------------|-------------|-------------|-------------|
| C3                | A8 (6,4) | 3           | 2           | 2           | 0           | 1           |
|                   | A9 (7,4) | 2           | 1           | 1           | 1           | 0           |
|                   | Dist(C3) | 5           | 3           | 3           | -           | -           |
| C1                | A5 (8,5) | 0           | 1           | 1           | 3           | 2           |
|                   | A6 (8,4) | 1           | 0           | 2           | 2           | 1           |
|                   | A7 (7,5) | 1           | 2           | 0           | 2           | 1           |
|                   | Dist(C1) | 2           | 3           | 3           | -           | -           |
| dist(C1)-dist(C3) |          | -3          | 0           | 0           | -           | -           |

Toutes les différences de distances sont négatives ou nulles. Fin de la séparation.  
Réitérer le processus sur les deux clusters obtenus C1 et C3.

**Il est clair qu'à l'appel suivant C3 contenant A8 et A9 seront séparés dans 2 clusters différents C3 et C5, respectivement.**

**Etape 2 :** Calculer l'instance la plus dissimilaire de C1 :

|    |          | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) |
|----|----------|-------------|-------------|-------------|
| C1 | A5 (8,5) | 0           | 1           | 1           |
|    | A6 (8,4) | 1           | 0           | 2           |
|    | A7 (7,5) | 1           | 2           | 0           |
|    | Dist(C1) | 2           | 3           | 3           |

**Etape 3 :** Cinder le cluster C1 entre C1 et C7

Point avec la plus grande dissimilarité est  $A6(8,4) = 3$

Mettre A6 dans un new cluster C7

|          |          |          |
|----------|----------|----------|
| A5 (8,5) | A6 (8,4) | A7 (7,5) |
| C1       | C7       | C1       |

|    |                   | A5<br>(8,5) | A6<br>(8,4) | A7<br>(7,5) |
|----|-------------------|-------------|-------------|-------------|
|    | A6 (8,4)          | 1           | 0           | 2           |
| C1 | Dist(C7)          | 1           | 0           | 2           |
|    | A5 (8,5)          | 0           | 1           | 1           |
|    | A7 (7,5)          | 1           | 2           | 0           |
|    | Dist(C1)          | 1           | -           | 1           |
|    | dist(C1)-dist(C7) | 0           | -           | -1          |

Toutes les différences de distances sont négatives ou nulles. Fin de la séparation.

Réitérer le processus sur les deux clusters obtenus C1 et C7.

**Il est clair qu'à l'appel suivant C1 contenant A5 et A7 seront séparés dans 2 clusters différents C1 et C9, respectivement.**

/\*\*\*\*\*/

**Etape 2 :** Calculer l'instance la plus dissimilaire de C2 :

|    |          | A1 (2,10) | A2 (4,9) | A3 (5,8) | A4 (5,9) |
|----|----------|-----------|----------|----------|----------|
| C2 | A1(2,10) | 0         | 3        | 5        | 4        |
|    | A2 (4,9) | 3         | 0        | 2        | 1        |
|    | A3 (5,8) | 5         | 2        | 0        | 1        |
|    | A4 (5,9) | 4         | 1        | 1        | 0        |
|    | Dist(C2) | 12        | 6        | 8        | 6        |

**Etape 3 :** Cinder le cluster C2 entre C2 et C4

Point avec la plus grande dissimilarité est  $A1(2,10) = 12$

Mettre A1 dans un new cluster C4

|            |           |           |           |
|------------|-----------|-----------|-----------|
| A1 (2, 10) | A2 (4, 9) | A3 (5, 8) | A4 (5, 9) |
| C4         | C2        | C2        | C2        |

|  |  | A1 (2,10) | A2 (4,9) | A3 (5,8) | A4 (5,9) |
|--|--|-----------|----------|----------|----------|
|--|--|-----------|----------|----------|----------|

|                          |                 |   |          |           |           |
|--------------------------|-----------------|---|----------|-----------|-----------|
| <b>C4</b>                | <b>A1(2,10)</b> | 0 | 3        | 5         | 4         |
|                          | <b>Dist(C4)</b> | - | <b>3</b> | <b>5</b>  | <b>4</b>  |
| <b>C2</b>                | <b>A2 (4,9)</b> | 3 | 0        | 2         | 1         |
|                          | <b>A3 (5,8)</b> | 5 | 2        | 0         | 1         |
|                          | <b>A4 (5,9)</b> | 4 | 1        | 1         | 0         |
|                          | <b>Dist(C2)</b> | - | <b>3</b> | <b>3</b>  | <b>2</b>  |
| <b>dist(C2)-dist(C4)</b> |                 | - | <b>0</b> | <b>-2</b> | <b>-2</b> |

Toutes les différences de distances sont négatives ou nulles. Fin de la séparation.

Réitérer le processus sur les deux clusters obtenus C2 et C4.

**Etape 2 :** Calculer l'instance la plus dissimilaire de C2 :

|           |                 |                 |                 |                 |
|-----------|-----------------|-----------------|-----------------|-----------------|
|           |                 | <b>A2 (4,9)</b> | <b>A3 (5,8)</b> | <b>A4 (5,9)</b> |
| <b>C2</b> | <b>A2 (4,9)</b> | 0               | 2               | 1               |
|           | <b>A3 (5,8)</b> | 2               | 0               | 1               |
|           | <b>A4 (5,9)</b> | 1               | 1               | 0               |
|           | <b>Dist(C2)</b> | <b>3</b>        | <b>3</b>        | <b>2</b>        |

**Etape 3 :** Cinder le cluster C2 entre C2 et C6

Point avec la plus grande dissimilarité est  $A2(4,9) = 3$

Mettre A2 dans un new cluster C6

|           |                  |                  |                  |
|-----------|------------------|------------------|------------------|
|           | <b>A2 (4, 9)</b> | <b>A3 (5, 8)</b> | <b>A4 (5, 9)</b> |
| <b>C6</b> |                  | <b>C2</b>        | <b>C2</b>        |

|                          |                 |                 |                 |                 |
|--------------------------|-----------------|-----------------|-----------------|-----------------|
|                          |                 | <b>A2 (4,9)</b> | <b>A3 (5,8)</b> | <b>A4 (5,9)</b> |
| <b>C6</b>                | <b>A2 (4,9)</b> | 0               | 2               | 1               |
|                          | <b>Dist(C4)</b> | -               | <b>2</b>        | <b>1</b>        |
| <b>C2</b>                | <b>A3 (5,8)</b> | 2               | 0               | 1               |
|                          | <b>A4 (5,9)</b> | 1               | 1               | 0               |
|                          | <b>Dist(C2)</b> | -               | <b>1</b>        | <b>1</b>        |
| <b>dist(C2)-dist(C6)</b> |                 | -               | <b>-1</b>       | <b>0</b>        |

Toutes les différences de distances sont négatives ou nulles. Fin de la séparation.

Réitérer le processus sur les deux clusters obtenus C2 et C6.

**Il est clair qu'à l'appel suivant sur C2 contenant A3 et A4 seront séparés dans 2 clusters différents C2 et C8, respectivement.**

**Fin de l'algo lorsque tous les clusters ne contiennent qu'une seule instance.**