

L'algorithme de agglomération **Complete-linkage** : à chaque étape, on fusionne deux clusters C_i et C_j si la dissimilarité $D(C_i, C_j)$ est la *plus petite* valeur parmi toutes les paires des clusters, où le $D(C_i, C_j)$ est défini comme montrant dans (1).

$$D(C_i, C_j) = \max_{x_i \in C_i, x_j \in C_j} d(x_i, x_j) \quad (1)$$

Exercice de la note de cours :

	o_1	o_2	o_3	o_4	o_5
o_1	-	4	9	7	11
o_2	4	-	1	6	3
o_3	9	1	-	8	6
o_4	7	6	8	-	4
o_5	11	3	6	4	-

Matrice de dissimilarité

	o_1	o_2	o_3	o_4	o_5
o_1	-	4	9	7	11
o_2	4	-	1	6	3
o_3	9	1	-	8	6
o_4	7	6	8	-	4
o_5	11	3	6	4	-

Étape 1

	o_1	o_2, o_3	o_4	o_5
o_1	-	9	7	11
o_2, o_3	9	-	8	6
o_4	7	8	-	4
o_5	11	6	4	-

Matrice Fusionné¹

	o_1	o_2, o_3	o_4	o_5
o_1	-	9	7	11
o_2, o_3	9	-	8	6
o_4	7	8	-	4
o_5	11	6	4	-

Étape 2

	o_1	o_2, o_3	o_4, o_5
o_1	-	9	11
o_2, o_3	9	-	8
o_4, o_5	11	8	-

Matrice Fusionné

	o_1	o_2, o_3	o_4, o_5
o_1	-	9	11
o_2, o_3	9	-	8
o_4, o_5	11	8	-

Étape 3

	o_1	o_2, o_3, o_4, o_5
o_1	-	11
o_2, o_3, o_4, o_5	11	-

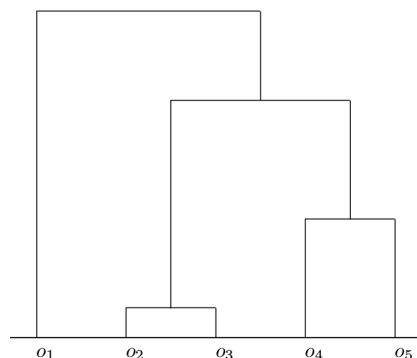
Matrice Fusionné

	o_1	o_2, o_3, o_4, o_5
o_1	-	11
o_2, o_3, o_4, o_5	11	-

Étape 4

	o_1, o_2, o_3, o_4, o_5
o_1, o_2, o_3, o_4, o_5	-

Matrice Fusionné



La solution avec un dendrogramma

1. Pour mettre à jour la matrice, nous conservons la distance maximale entre le cluster et les clusters fusionnés. Par exemple, pour la singleton o_1 , la dissimilarité au nouveau groupe fusionné $\{o_2, o_3\}$ est $\max\{d(o_1, o_2), d(o_1, o_3)\}$.

L'algorithme de agglomération **Single-linkage** : à chaque étape, on fusionne deux clusters C_i et C_j si la dissimilarité $D(C_i, C_j)$ est *la plus petite valeur* parmi toutes les paires des clusters, où le $D(C_i, C_j)$ est défini comme montrant dans (2).

$$D(C_i, C_j) = \min_{x_i \in C_i, x_j \in C_j} d(x_i, x_j) \quad (2)$$

	o_1	o_2	o_3	o_4	o_5
o_1	-	4	9	7	11
o_2	4	-	1	6	3
o_3	9	1	-	8	6
o_4	7	6	8	-	4
o_5	11	3	6	4	-

Matrice de dissimilarité

	o_1	o_2	o_3	o_4	o_5
o_1	-	4	9	7	11
o_2	4	-	1	6	3
o_3	9	1	-	8	6
o_4	7	6	8	-	4
o_5	11	3	6	4	-

Étape 1

	o_1	o_2, o_3	o_4	o_5
o_1	-	4	7	11
o_2, o_3	4	-	6	3
o_4	7	6	-	4
o_5	11	3	4	-

Matrice de fusionné¹

	o_1	o_2, o_3	o_4	o_5
o_1	-	4	7	11
o_2, o_3	4	-	6	3
o_4	7	6	-	4
o_5	11	3	4	-

Étape 2

	o_1	o_2, o_3, o_5	o_4
o_1	-	4	7
o_2, o_3, o_5	4	-	4
o_4	7	4	-

Matrice de fusionné¹

	o_1	o_2, o_3, o_5	o_4
o_1	-	4	7
o_2, o_3, o_5	4	-	4
o_4	7	4	-

Étape 3

	o_1, o_2, o_3, o_5	o_4
o_1, o_2, o_3, o_5	-	4
o_4	4	-

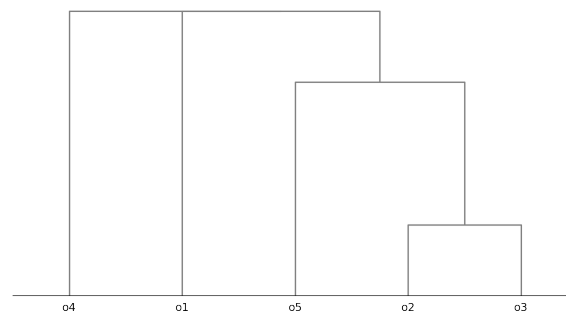
Matrice Fusionné

	o_1, o_2, o_3, o_5	o_4
o_1, o_2, o_3, o_5	-	4
o_4	4	-

Étape 4

	o_1, o_2, o_3, o_4, o_5
o_1, o_2, o_3, o_4, o_5	-

Matrice Fusionné



La solution avec un dendrogramma

1. Pour mettre à jour la matrice, nous conservons la distance minimale entre le cluster et les clusters fusionnés. Par exemple, pour la singleton o_1 , la dissimilarité au nouveau groupe fusionné $\{o_2, o_3\}$ est $\min\{d(o_1, o_2), d(o_1, o_3)\}$.