

Projet Traitement Numérique

Rédacteur :

Yanis HAMMACI
Réda ARBANE
Akram CHAABNIA

14/04/2023

1 Dendrogramme de la CAH

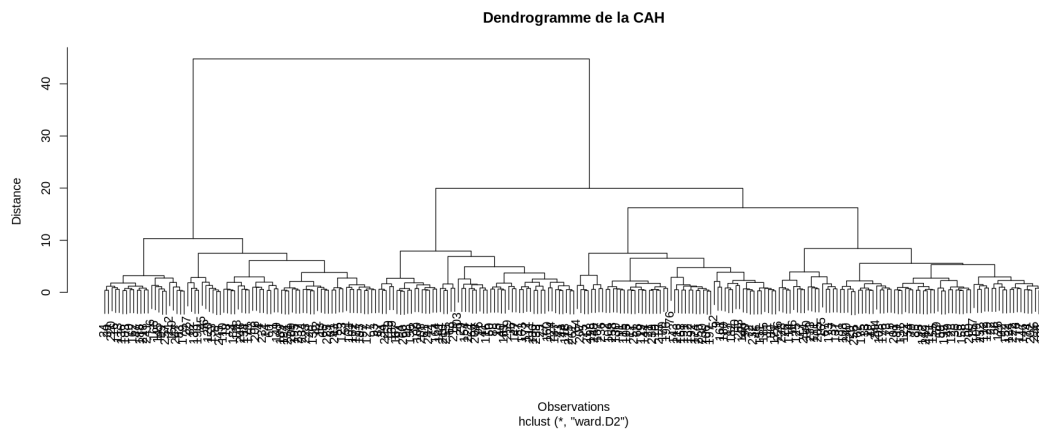


Figure 1: Dendrogramme de la CAH

2 Graphique de l'Inertie

2.1 Explication du graphique de l'inertie

Le graphique de l'inertie (Figure 2) présente l'évolution de l'inertie totale à mesure que le nombre de clusters augmente. L'inertie représente la variabilité intra-classe, c'est-à-dire la dispersion des points de données à l'intérieur de chaque cluster.

Diminution importante de l'inertie de 1 à 4 clusters: Cette observation indique que les regroupements successifs de 1 à 4 clusters permettent de capturer une part significative de la variabilité des données. En d'autres termes, la formation de ces clusters regroupe efficacement les points de données qui se ressemblent, réduisant ainsi la variabilité globale.

Diminution lente de l'inertie au-delà de 4 clusters: Ce changement de tendance suggère que les regroupements supplémentaires de 5 clusters et plus n'apportent pas d'amélioration majeure à la structure des données. En effet, la variabilité intra-classe restante est déjà relativement faible, et l'ajout de clusters supplémentaires ne permet pas de la réduire de manière significative.

Point d'inflexion ("coude"): Le point d'inflexion observé sur la courbe correspond approximativement à 4 clusters. Ce point marque le moment où la diminution de l'inertie devient moins importante, indiquant que c'est à ce niveau que le regroupement des données atteint un bon équilibre entre la réduction de la variabilité et la complexité du modèle.

En conclusion, le graphique de l'inertie fournit un outil utile pour déterminer le nombre optimal de clusters en analysant l'évolution de la variabilité intra-classe en fonction du nombre de regroupements. Le point d'inflexion ("coude") permet d'identifier le nombre de clusters qui maximise la réduction de la variabilité tout en limitant la complexité du modèle.

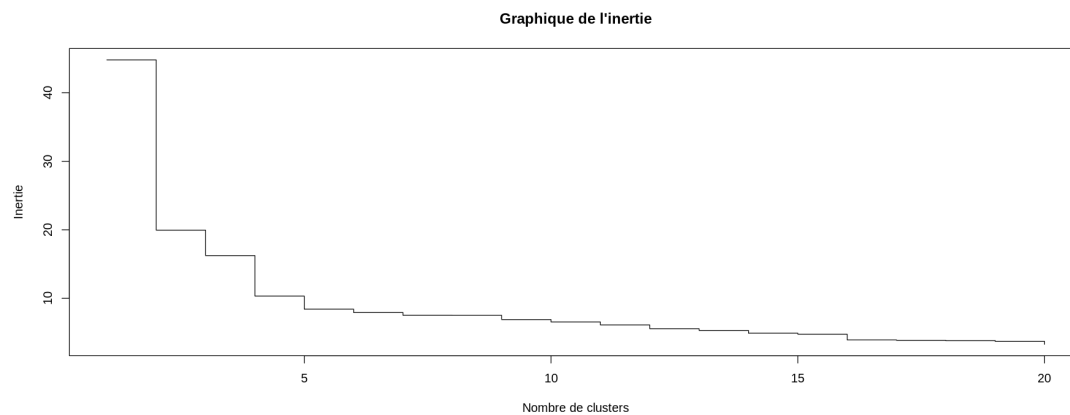


Figure 2: Graphique de l'Inertie

3 Inertie du dendrogramme avec plus grands gains entourés

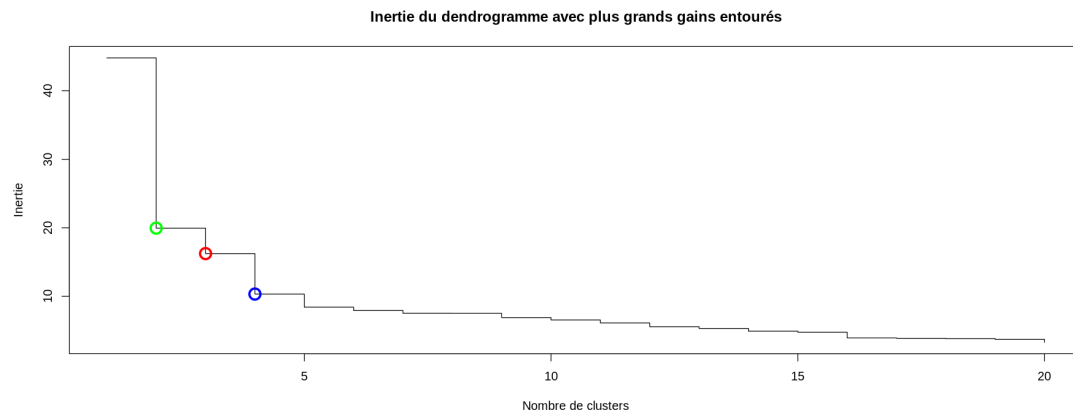


Figure 3: Inertie du dendrogramme avec plus grands gains entourés

4 Figure 4: Partitions du Dendrogramme en 2, 3 ou 4 classes

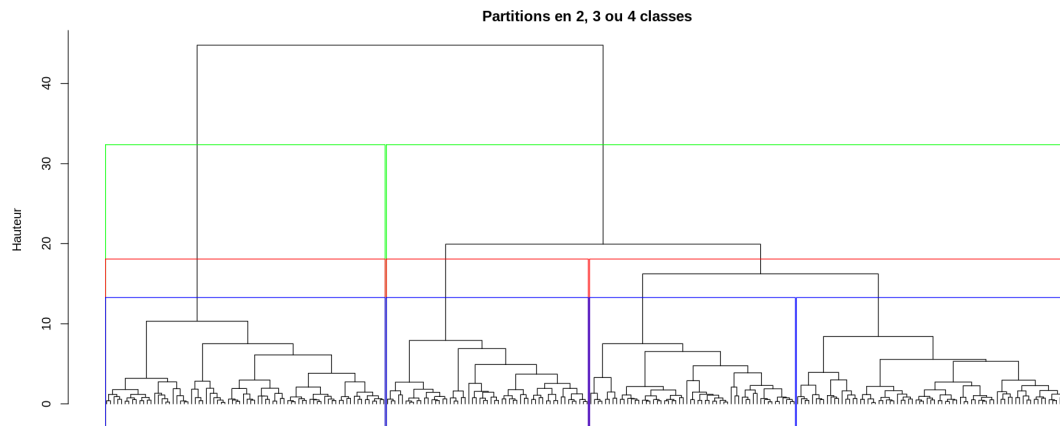


Figure 4: Partitions du Dendrogramme en 2, 3 ou 4 classes

5 Figure 5: Partitions des voitures sur les différentes classes

```
Cut = 2
      clusters
factor    1    2
Europe.   0.00 100.00
Japan.    0.00 100.00
US.       46.91  53.09

Cut = 3
      clusters
factor    1    2    3
Europe.   0.00 93.75  6.25
Japan.    0.00 90.20  9.80
US.       46.91 24.07 29.01

Cut = 4
      clusters
factor    1    2    3    4
Europe.   0.00 22.92 70.83  6.25
Japan.    0.00 43.14 47.06  9.80
US.       46.91 14.20  9.88 29.01
```

Figure 5: Partitions des voitures sur les différentes classes