

# Étude de marché pour l'exportation de poulets biologiques

---

Entreprise : La Poule Qui Chante

Oussama BENMAHAMMED | 09/2024 | Data Analyst



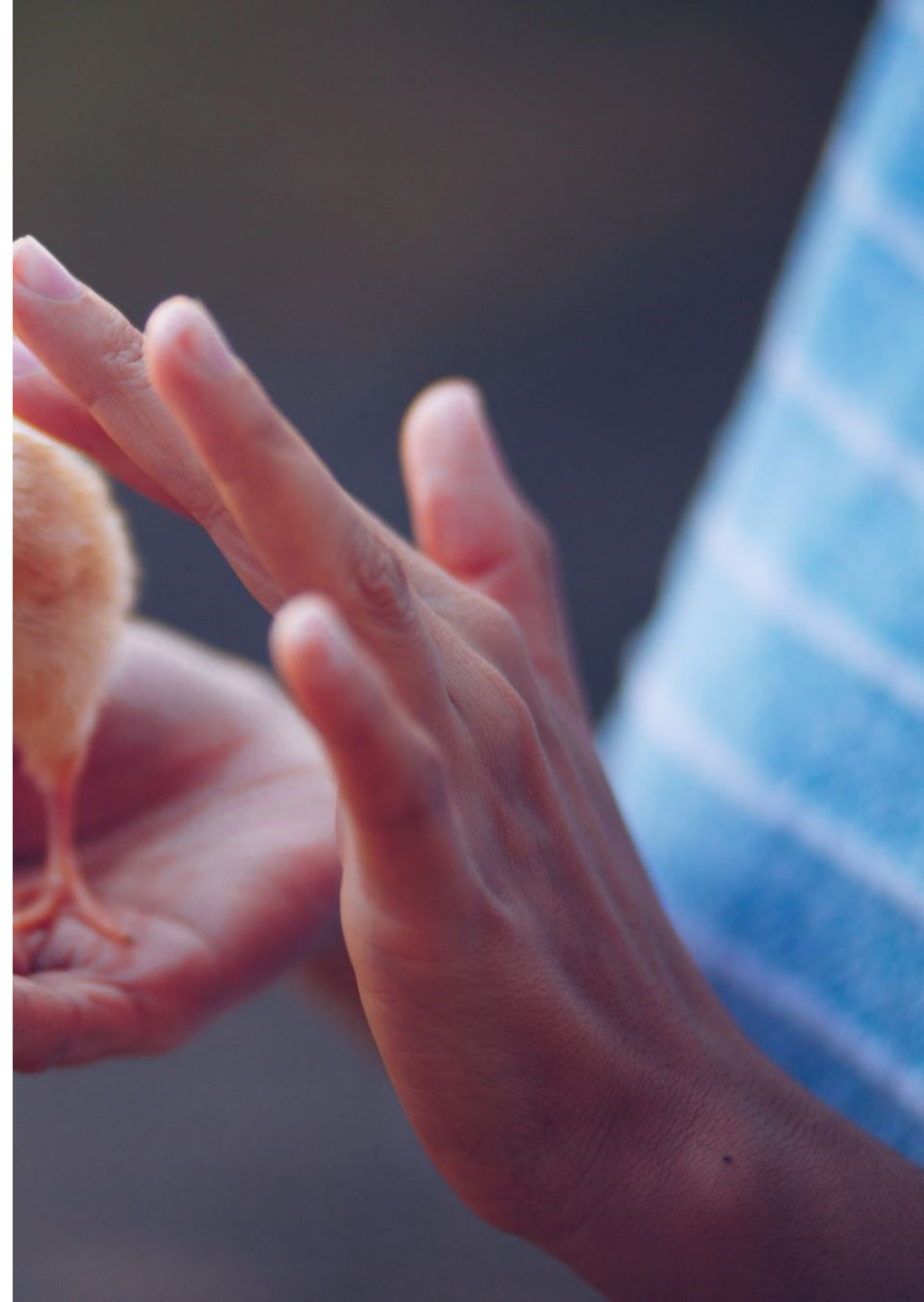
# Contexte du projet

- L'entreprise Française La Poule Qui Chante souhaite étendre ses activités internationales en exportant des poulets biologiques.
- L'objectif principal est de cibler des marchés internationaux présentant un fort potentiel basé sur des facteurs économiques, sociaux, et politiques.



# Objectifs et Problématique

1. Identifier les pays présentant un fort potentiel d'importation de poulets biologiques.
2. Proposer une stratégie d'exportation pour des marchés prioritaires.



# Démarche et Méthodologie



1. **Collecte des données** : Données provenant de la FAO, Banque Mondiale, et d'autres sources fiables sur la consommation et production de volaille.

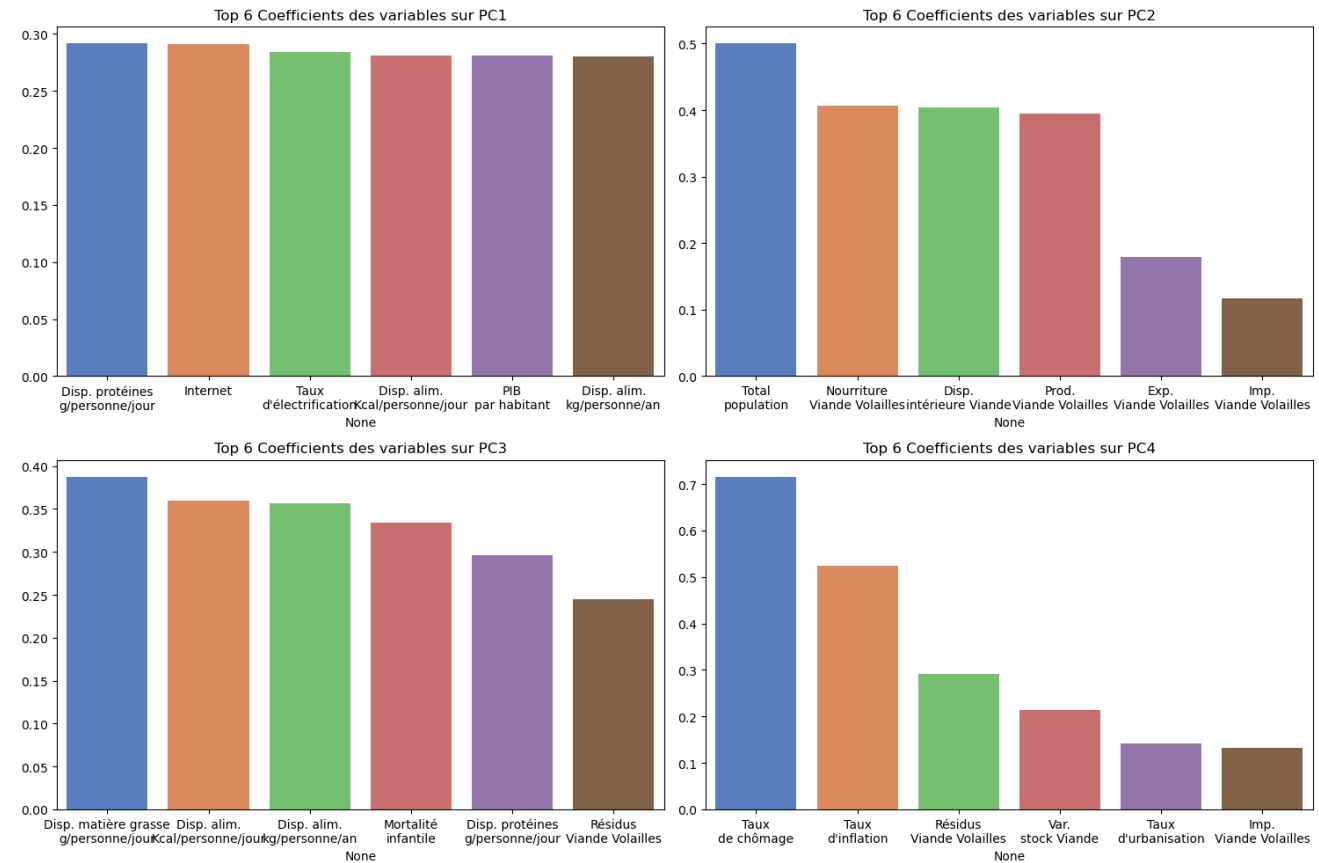
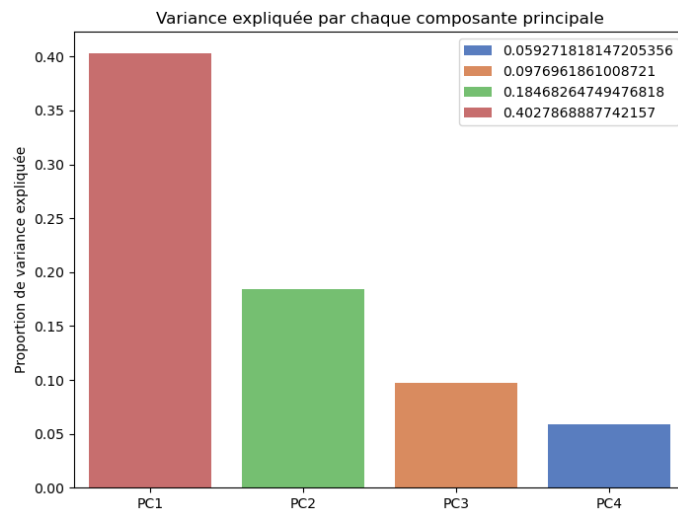


2. **Nettoyage des données** : Gestion des valeurs manquantes, transformation Yeo-Johnson pour les variables non paramétriques.



3. **Analyse exploratoire** : Utilisation de l'ACP pour réduire les dimensions et identifier des clusters via K-means et CAH.

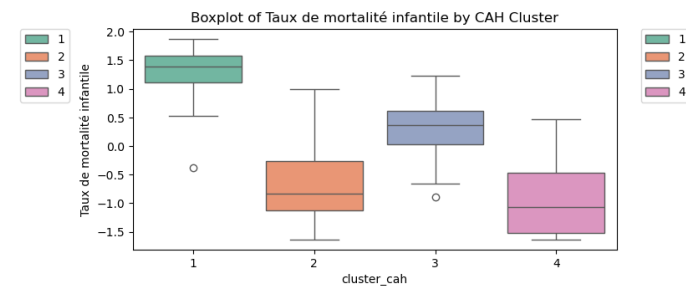
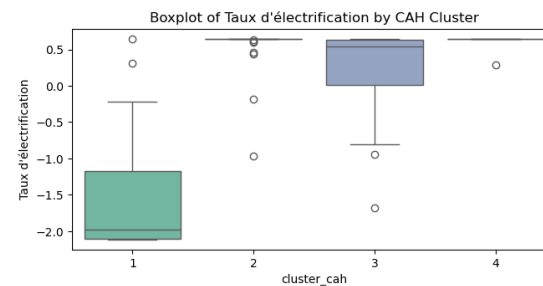
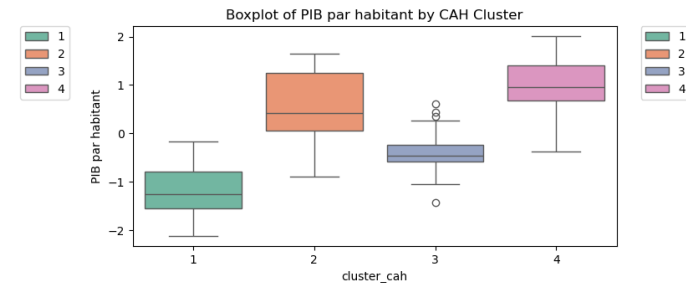
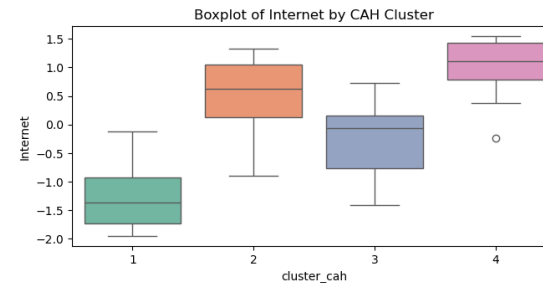
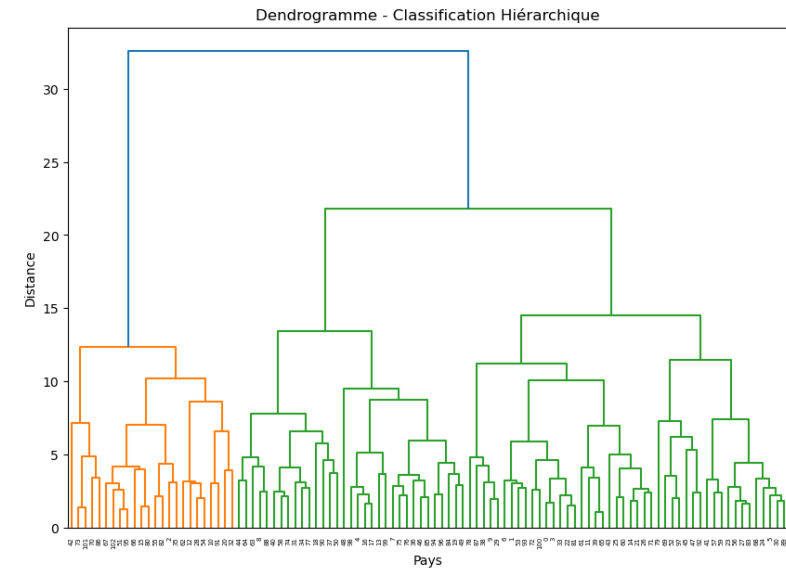
# Analyse en Composantes Principales (ACP)



- PC1 et PC2 expliquent de **60% de la variance totale** des données.
- **PC1** : Représente le niveau de **développement économique** et les **infrastructures**.
- **PC2** : Représente la **production de volaille** et la **population totale**.

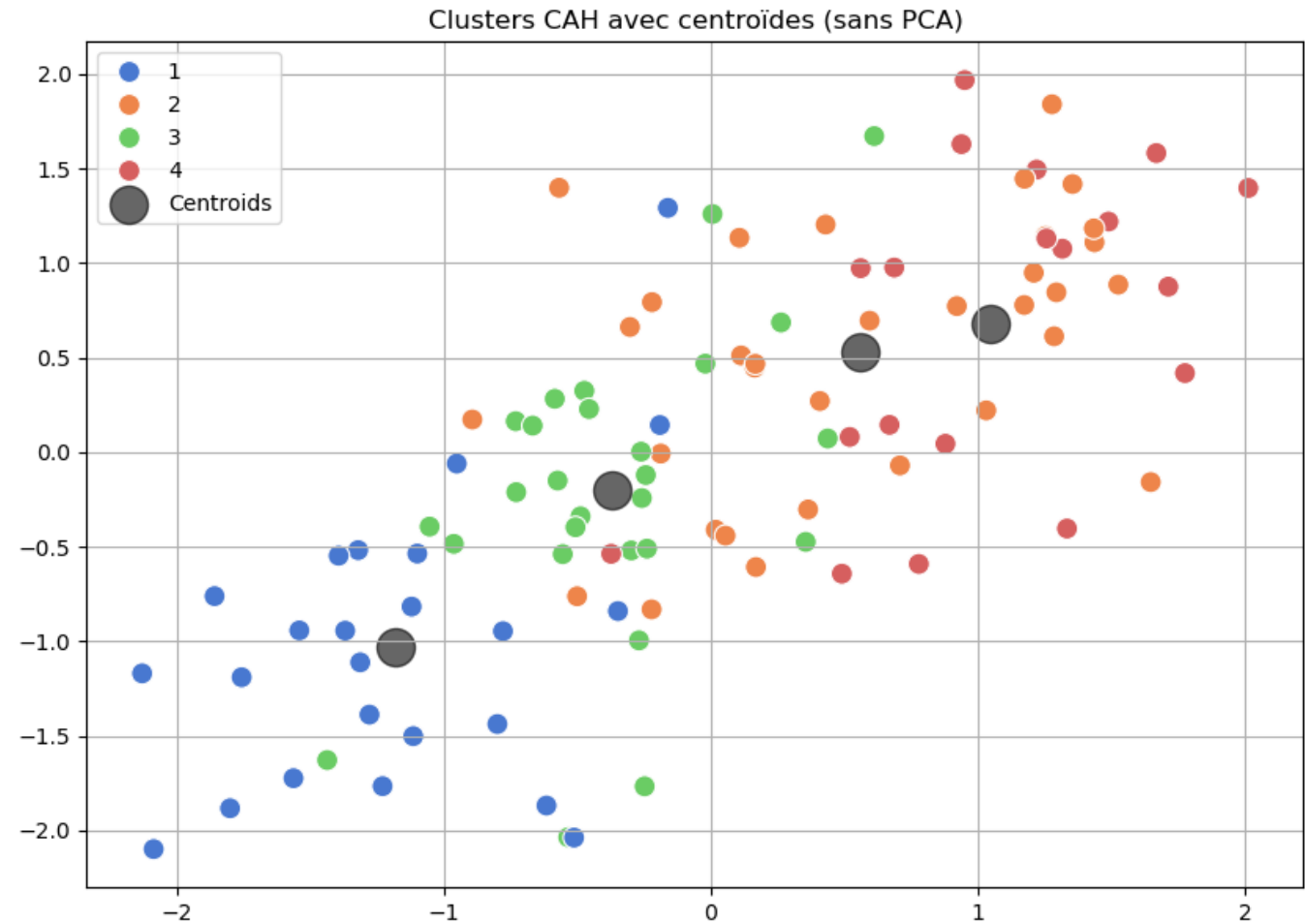
# Clustering CAH

- Le clustering hiérarchique (CAH) regroupe les pays selon des distances entre les variables.
- **Cluster 1** : Pays en développement avec un faible accès aux infrastructures.
- **Cluster 2** : Pays économiquement développés.
- **Cluster 3** : Pays en transition.
- **Cluster 4** : Pays riches avec de fortes infrastructures.



# Clustering CAH

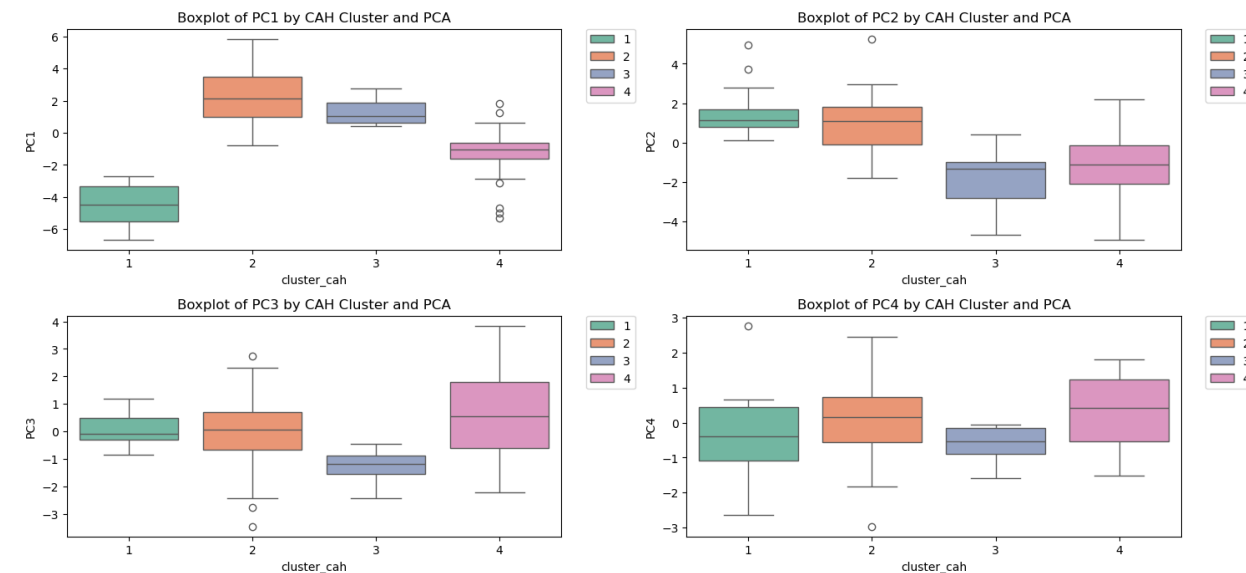
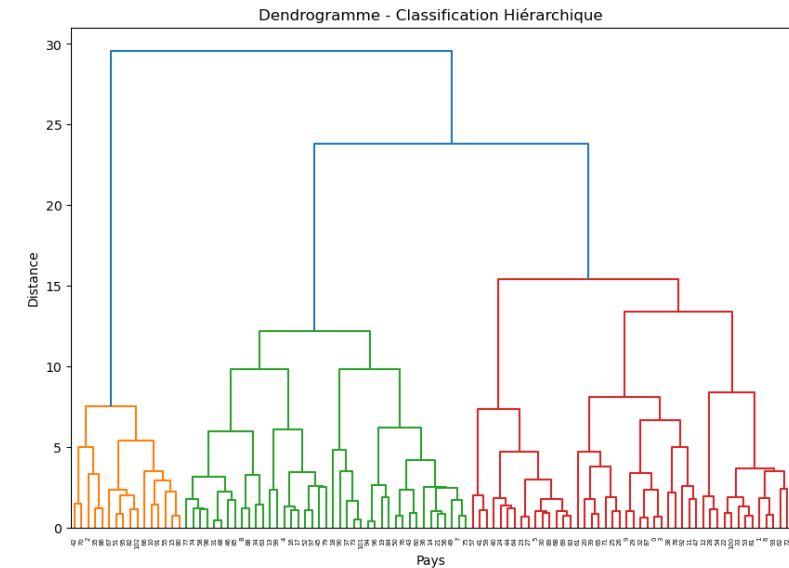
- Le clustering hiérarchique (CAH) méthode hiérarchique qui permet de visualiser les relations progressives entre les pays en fonction de toutes les variables disponibles.
- Les clusters révèlent des groupes bien distincts, de pays très développés à ceux qui luttent avec des infrastructures de base, et des pays en transition avec un potentiel de croissance.



Cluster cah		Areas (Pays/Régions)
0	1	angola, benin, botswana, cameroon, congo, eswatini, gabon, ghana, india, kenya, lesotho, liberia, namibia, niger, nigeria, pakistan, philippines, senegal, sierra leone, sri lanka, togo, uganda, vietnam, zambia
1	2	australia, belarus, belgium, brazil, canada, chile, china, colombia, france, germany, greece, guatemala, hungary, ireland, italy, japan, jordan, kazakhstan, malaysia, netherlands, new zealand, poland, portugal, romania, russia, south africa, spain, sweden, thailand, turkey, ukraine, united kingdom, united states
2	3	albania, algeria, armenia, azerbaijan, belize, bolivia, bulgaria, costa rica, croatia, ecuador, el salvador, fiji, georgia, guyana, honduras, iraq, lebanon, morocco, myanmar, nicaragua, panama, paraguay, samoa, serbia, suriname, tunisia, uruguay
3	4	austria, cyprus, denmark, estonia, finland, iceland, israel, jamaica, kuwait, lithuania, luxembourg, malta, norway, oman, saudi arabia, slovenia, switzerland, trinidad and tobago, united arab emirates

# Clustering CAH avec ACP

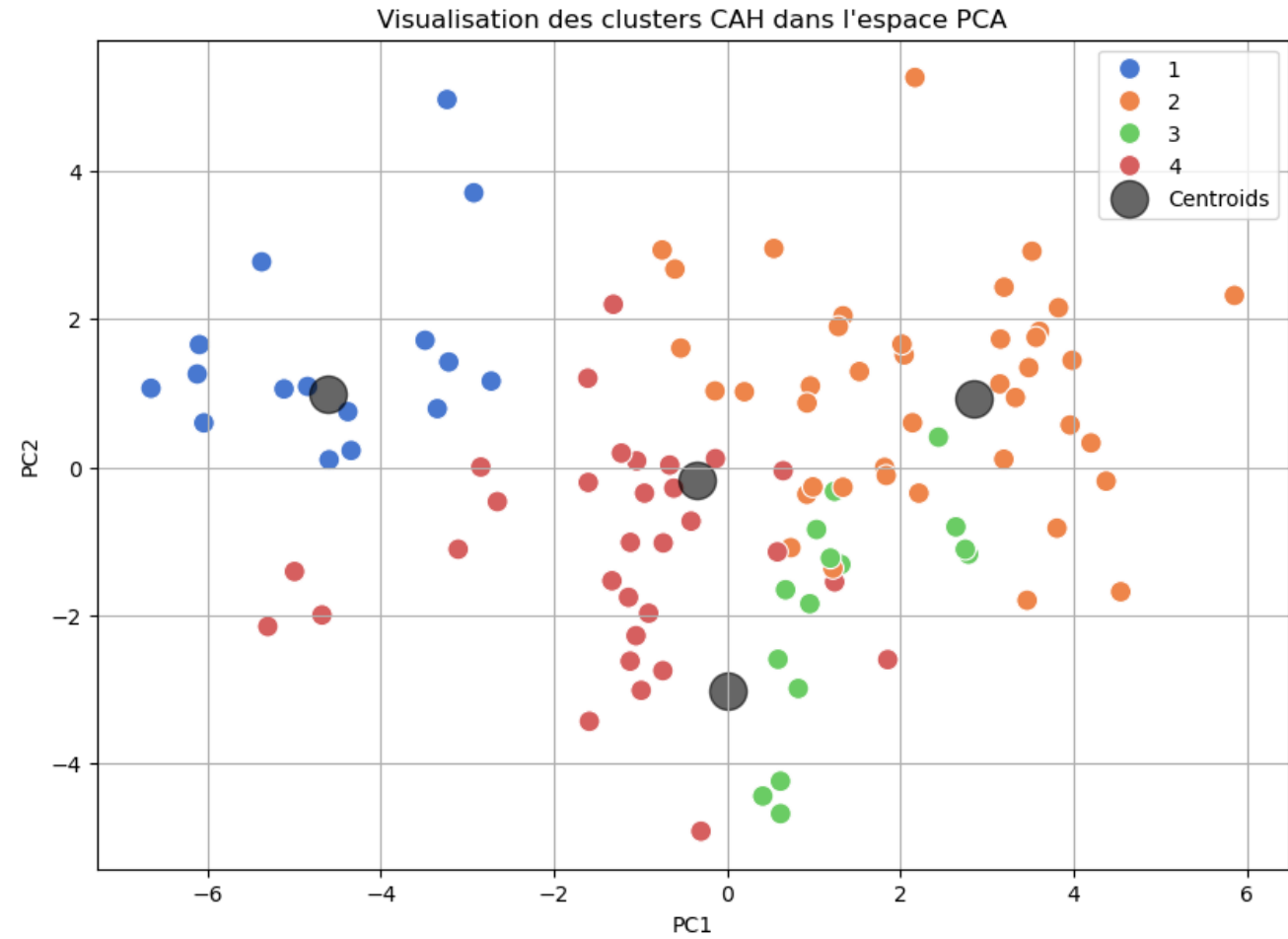
- L'ACP permet d'améliorer la cohérence des clusters en simplifiant les données.
- Les clusters sont plus homogènes et les différences entre les pays sont mieux accentuées.
- **Cluster 1** : Pays à faible développement avec des infrastructures limitées
- **Cluster 2** : Pays développés avec des infrastructures solides, une économie forte, et une grande capacité à fournir des ressources alimentaires). Ces pays dominent les composantes principales.
- **Cluster 3** : Pays en transition ou émergents qui affichent des caractéristiques mixtes, oscillant entre pays développés et en développement (ex : Bolivie, Serbie). Ces pays ont des infrastructures moyennes mais montrent un potentiel de croissance.
- **Cluster 4** : Petits pays à forte richesse, souvent avec une population réduite mais un accès exceptionnel aux ressources et à la technologie





# Clustering CAH avec ACP

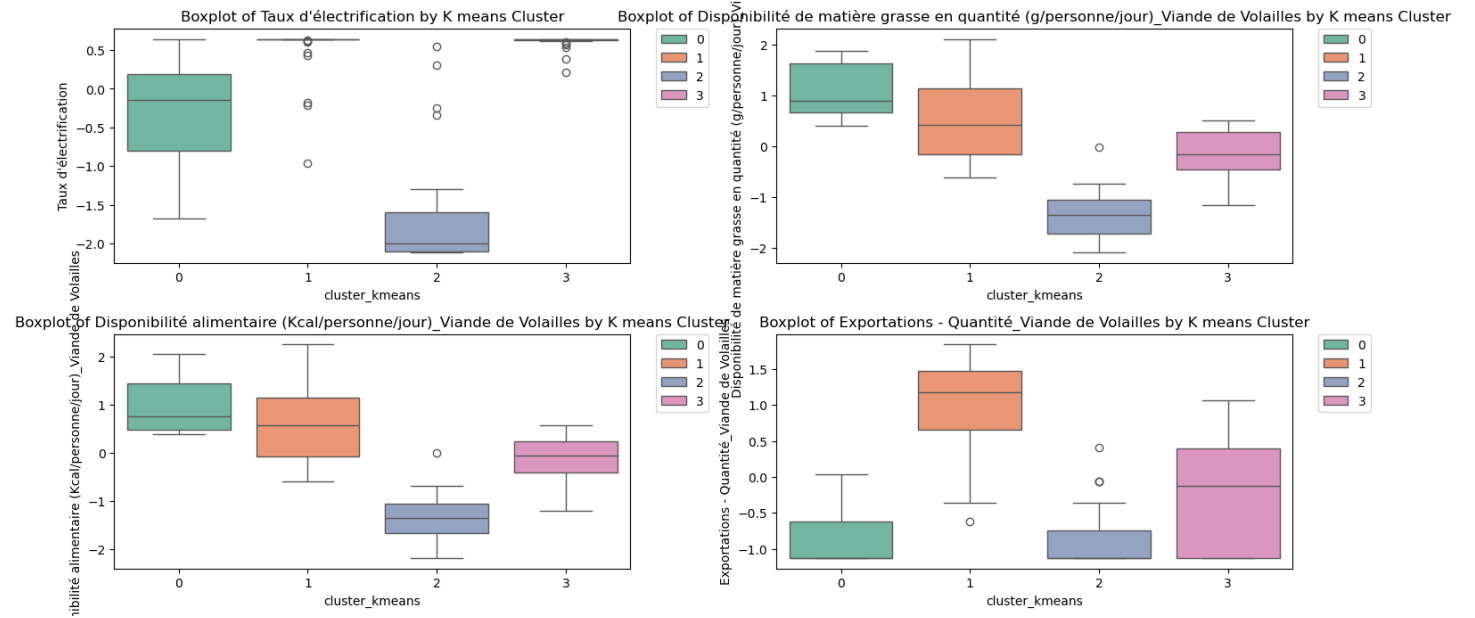
- Avec ACP les différences entre les pays sont mieux accentuées.



Cluster cah - PCA		Areas (Pays/Régions)
0	1	angola, benin, cameroon, ghana, india, kenya, liberia, niger, nigeria, pakistan, senegal, sierra leone, sri lanka, togo, uganda, zambia
1	2	australia, belarus, belgium, brazil, bulgaria, canada, chile, china, colombia, costa rica, france, germany, greece, guatemala, iraq, israel, italy, japan, jordan, kazakhstan, kuwait, lithuania, malaysia, morocco, netherlands, philippines, poland, portugal, romania, russia, saudi arabia, south africa, spain, sweden, thailand, turkey, ukraine, united arab emirates, united kingdom, united states, vietnam
2	3	austria, cyprus, denmark, estonia, finland, hungary, iceland, ireland, luxembourg, malta, new zealand, norway, oman, slovenia, switzerland
3	4	albania, algeria, armenia, azerbaijan, belize, bolivia, botswana, congo, croatia, ecuador, el salvador, eswatini, fiji, gabon, georgia, guyana, honduras, jamaica, lebanon, lesotho, myanmar, namibia, nicaragua, panama, paraguay, samoa, serbia, suriname, trinidad and tobago, tunisia, uruguay

# Clusters K-means

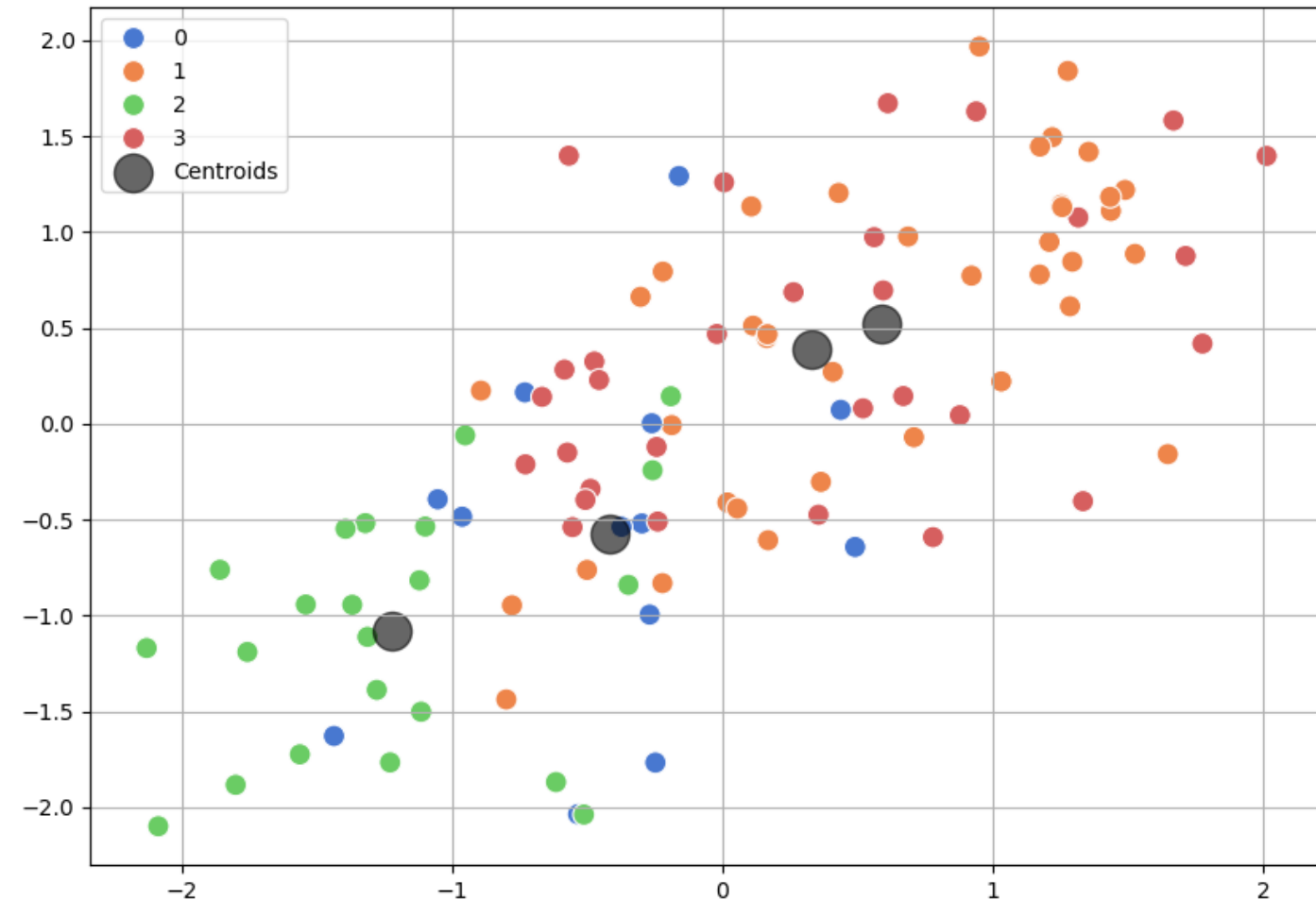
- Les pays sont regroupés directement sur la base des variables socio-économiques d'origine (PIB, taux d'électrification, mortalité infantile, etc.). Les résultats des 4 clusters regroupant des pays ayant des profils similaires sur l'ensemble de ces variables. Par exemple :
- Cluster 0** : Pays en développement avec un PIB plus bas et une électrification limitée.
- Cluster 1** : Pays avancés économiquement.
- Cluster 2** : Pays à faible développement avec une forte mortalité infantile et un accès limité aux infrastructures.
- Cluster 3** : Pays en transition avec des indicateurs mixtes.



# Clusters K-means

- Clustering effectué directement sur les données sans réduction de la dimensionnalité.
- 4 clusters sont identifiés, distinguant les pays par leur niveau de développement, production de volaille et infrastructures.

Visualisation des clusters Kmeans dans l'espace sans PCA



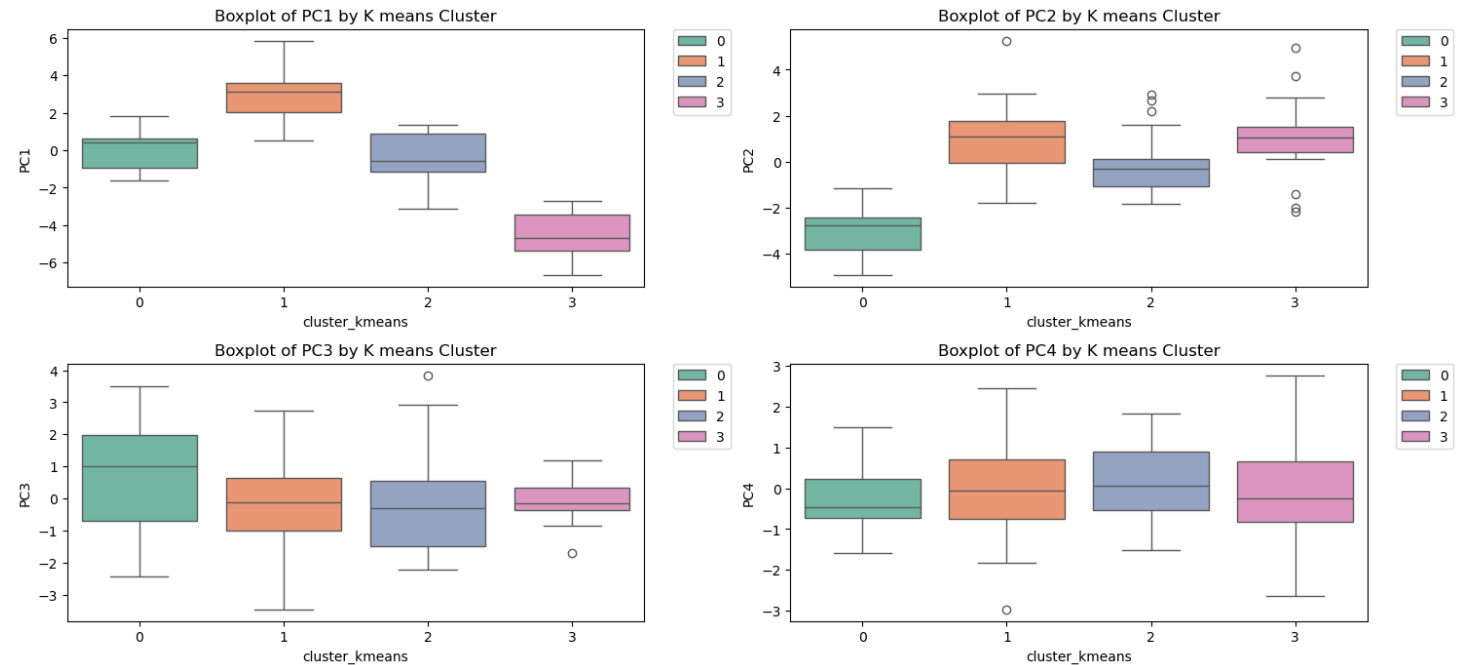
## Cluster K-Means

### Areas (Pays/Régions)

0	0	belize, bolivia, fiji, gabon, guyana, honduras, jamaica, myanmar, nicaragua, panama, samoa, suriname, trinidad and tobago
1	1	australia, belarus, belgium, brazil, canada, chile, china, colombia, denmark, france, germany, guatemala, hungary, ireland, israel, italy, japan, kazakhstan, kuwait, malaysia, netherlands, new zealand, philippines, poland, portugal, romania, russia, saudi arabia, south africa, spain, sweden, thailand, turkey, ukraine, united arab emirates, united kingdom, united states, vietnam
2	2	angola, benin, botswana, cameroon, congo, eswatini, ghana, india, kenya, lesotho, liberia, namibia, niger, nigeria, pakistan, paraguay, senegal, sierra leone, sri lanka, togo, uganda, zambia
3	3	albania, algeria, armenia, austria, azerbaijan, bulgaria, costa rica, croatia, cyprus, ecuador, el salvador, estonia, finland, georgia, greece, iceland, iraq, jordan, lebanon, lithuania, luxembourg, malta, morocco, norway, oman, serbia, slovenia, switzerland, tunisia, uruguay

# Clusters K-means avec ACP

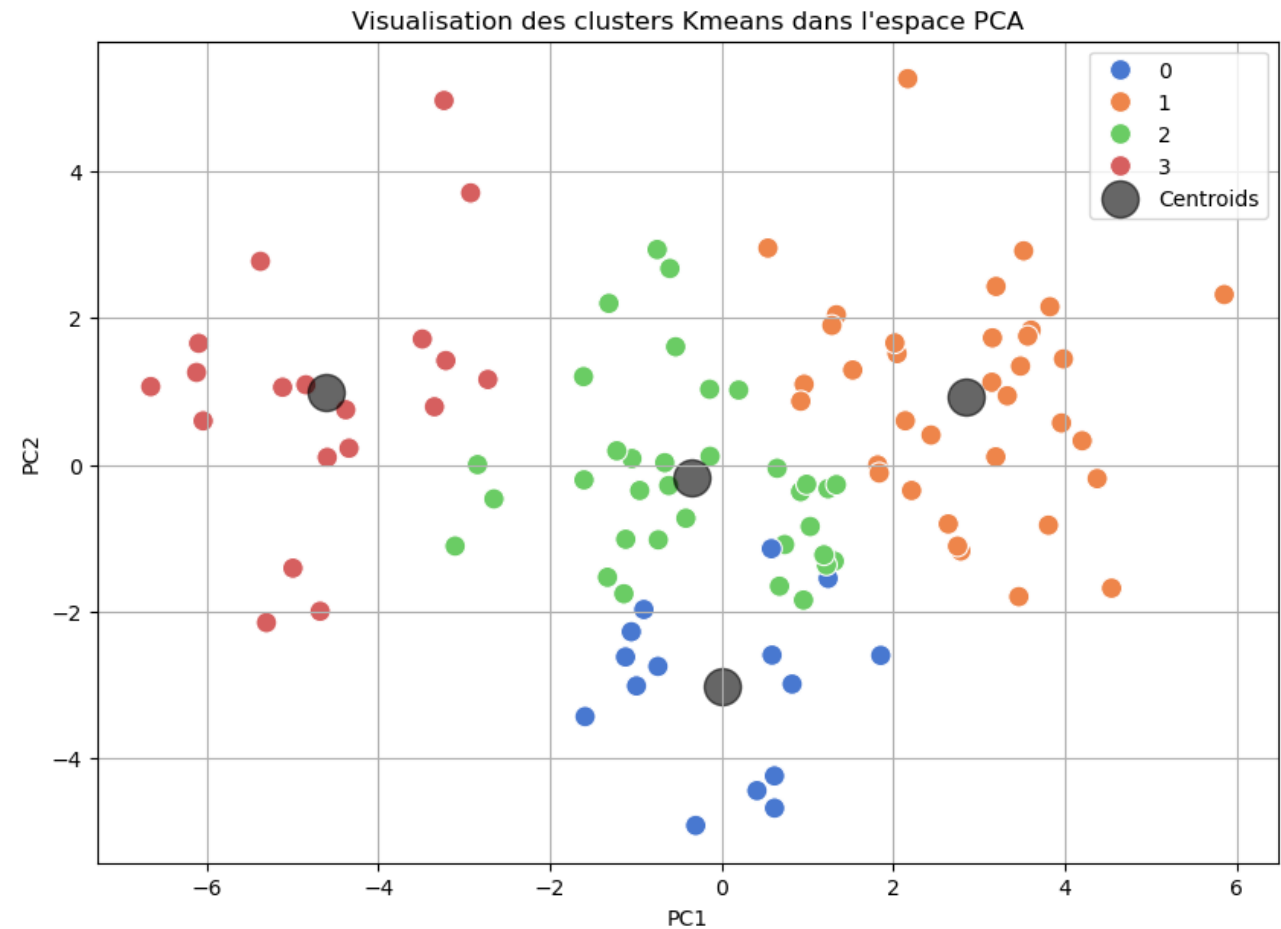
- Les clusters sont similaires en termes de segmentation, mais plus faciles à interpréter grâce à la réduction de dimensions.
- **Cluster 0** : Petits pays ou îles avec des niveaux de développement variables.
- **Cluster 1** : Pays développés avec un fort accès aux infrastructures.
- **Cluster 2** : Pays en développement ou en transition.
- **Cluster 3** : Pays à faible développement socio-économique, principalement en Afrique subsaharienne et en Asie du Sud, avec des infrastructures limitées et des indicateurs tels que la mortalité infantile élevée et un faible accès à l'électrification.






# Clusters K-means avec ACP

- L'ACP a permis de simplifier les données avant le clustering.
- Les clusters sont plus homogènes et faciles à interpréter car ils se basent sur les dimensions les plus significatives (PC1 et PC2).



Cluster K-Means_pca		Areas (Pays/Régions)
0	0	albania, belize, cyprus, estonia, fiji, gabon, guyana, iceland, jamaica, luxembourg, malta, panama, samoa, suriname, trinidad and tobago
1	1	australia, belarus, belgium, brazil, canada, chile, china, colombia, denmark, france, germany, hungary, ireland, israel, italy, japan, kazakhstan, kuwait, malaysia, netherlands, new zealand, poland, portugal, romania, russia, saudi arabia, south africa, spain, sweden, thailand, turkey, ukraine, united arab emirates, united kingdom, united states
2	2	algeria, armenia, austria, azerbaijan, bolivia, bulgaria, congo, costa rica, croatia, ecuador, el salvador, finland, georgia, greece, guatemala, honduras, iraq, jordan, lebanon, lithuania, morocco, myanmar, namibia, nicaragua, norway, oman, paraguay, philippines, serbia, slovenia, switzerland, tunisia, uruguay, vietnam
3	3	angola, benin, botswana, cameroon, eswatini, ghana, india, kenya, lesotho, liberia, niger, nigeria, pakistan, senegal, sierra leone, sri lanka, togo, uganda, zambia



# Recommandations Stratégiques pour l'exportation du poulet bio

---

**Priorité immédiate :** Ciblez les pays développés (**États-Unis, Allemagne, Japon**) identifiés dans les clusters avec et sans PCA. Ces pays disposent des infrastructures nécessaires et d'une demande élevée pour des produits biologiques de haute qualité.

**Opportunités à moyen terme :** Surveillez les pays émergents (**Inde, Nigeria, Pakistan**), qui montrent une croissance rapide avec des infrastructures en développement. Ces pays pourraient devenir des marchés intéressants à moyen terme à mesure que leur classe moyenne croît.

**Opportunités à long terme :** Pays avec des infrastructures limitées (**Angola, Ghana, Kenya**) pourraient représenter un potentiel à long terme, mais nécessitent des améliorations économiques et infrastructurelles avant de pouvoir être ciblés pour des exportations significatives de produits biologiques.

# Conclusion

- **Analyse avec PCA** est plus parlante car elle réduit le bruit causé par des variables fortement corrélées, tout en capturant les principales sources de variance.
  - En d'autres termes, elle permet de se concentrer sur les dimensions les plus significatives, comme les infrastructures et le niveau de développement, sans être distrait par des variables redondantes.
  - De plus, elle améliore la cohérence des clusters et facilite l'interprétation stratégique.
- **Sans PCA**, les résultats peuvent être influencés par la forte dispersion des données, rendant plus difficile l'identification des groupes homogènes.



Merci pour  
votre  
attention

Questions ?

