

# La poule qui chante

Jérôme Parnasse 14/09/2023



# Sommaire

## **Présentation**

- Introduction
- Contexte du projet
- Objectifs du projet

## **Acquisitions des données**

1. Découverte des données
2. Préparation des données

## **Techniques d'analyse exploratoires**

1. Analyse univarié et bivariée
2. Analyse multivarié en composante principale

## **Méthodes de segmentation :**

1. Clustering k-means
2. Clustering de classification ascendante hiérarchique
3. Comparaisons des méthodes de clustering
4. Interprétation des résultats

## **Conclusion**

- Recommandations

# Introduction



## En 2020 :

- Importance du Poulet
- Consommation Globale
- Croissance en Asie
- Tendance en Afrique
- Consommation aux États-Unis
- Potentiel de Marché
- Défis et Opportunités
- Tendances Diététiques
- Comparaison avec d'autres Viandes



# La poule qui chante

Le contexte

- Entreprise
- Projet d'Expansion
- Analyse de Marché Cible
- Focus sur les Tendances de Consommation

# Acquisitions des données

## Découverte des données



FOOD AND AGRICULTURE  
ORGANIZATION  
OF THE UNITED NATIONS

### Sources des Données

- Données recueillies auprès de la Food and Agriculture Organization (FAO) des Nations Unies
- Assurance de la qualité et de la précision

### Les indicateurs :

#### Bilan Alimentaire et Consommation de Poulet :

- Quantité de poulet consommée
- Identification des principaux pays consommateurs
- Compréhension des tendances de consommation

#### Indicateurs Économiques et Démographiques :

- Population
- PIB par habitant
- Production de poulets vifs
- Commerce de poulets

# Acquisitions des données

## Préparation et nettoyage des données

index		Zone	Dispo_intérieur en Tonne	Dispo_Alim_kcal_pers_jour_volaille	Dispo_Alim_kg_pers_volaille	Dispo_Prot_g_pers_jour_volaille
0	0	Afghanistan	67000	5.98	1.71	0.61
1	1	Afrique du Sud	2067000	140.22	34.86	13.78
2	2	Albanie	40000	71.27	13.94	5.41
3	3	Algérie	280000	22.04	6.38	1.96
4	4	Allemagne	1626000	66.50	18.34	7.53



Filtrage des Données: Sélection de l'année d'intérêt et filtrage du DataFrame pour ne conserver que les données correspondant à cette année.



Sélection et Pivotelement: Sélection des colonnes pertinentes et pivotelement du DataFrame pour organiser les données de manière plus lisible.



Renommage des Colonnes: Modification des noms de colonnes pour une meilleure compréhension.



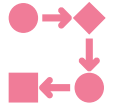
Création et Conversion de Variables: Création d'un nouveau DataFrame avec les colonnes nécessaires et conversion des unités si nécessaire.



Fusion de Données: Fusion de plusieurs DataFrames contenant différentes sources de données pour créer un DataFrame final consolidé.



Gestion des Valeurs Manquantes et Infinies: Remplacement des valeurs infinies par NaN, puis imputation des valeurs manquantes en utilisant la méthode KNN.



Finalisation: Réinitialisation de l'index et concaténation des données imputées pour créer le DataFrame final.

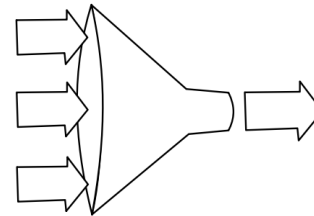
# Analyses exploratoire

# Analyse exploratoire

## 1.0 Analyse univarié et bivariée

Filtrage des Données pour la selection des pays:

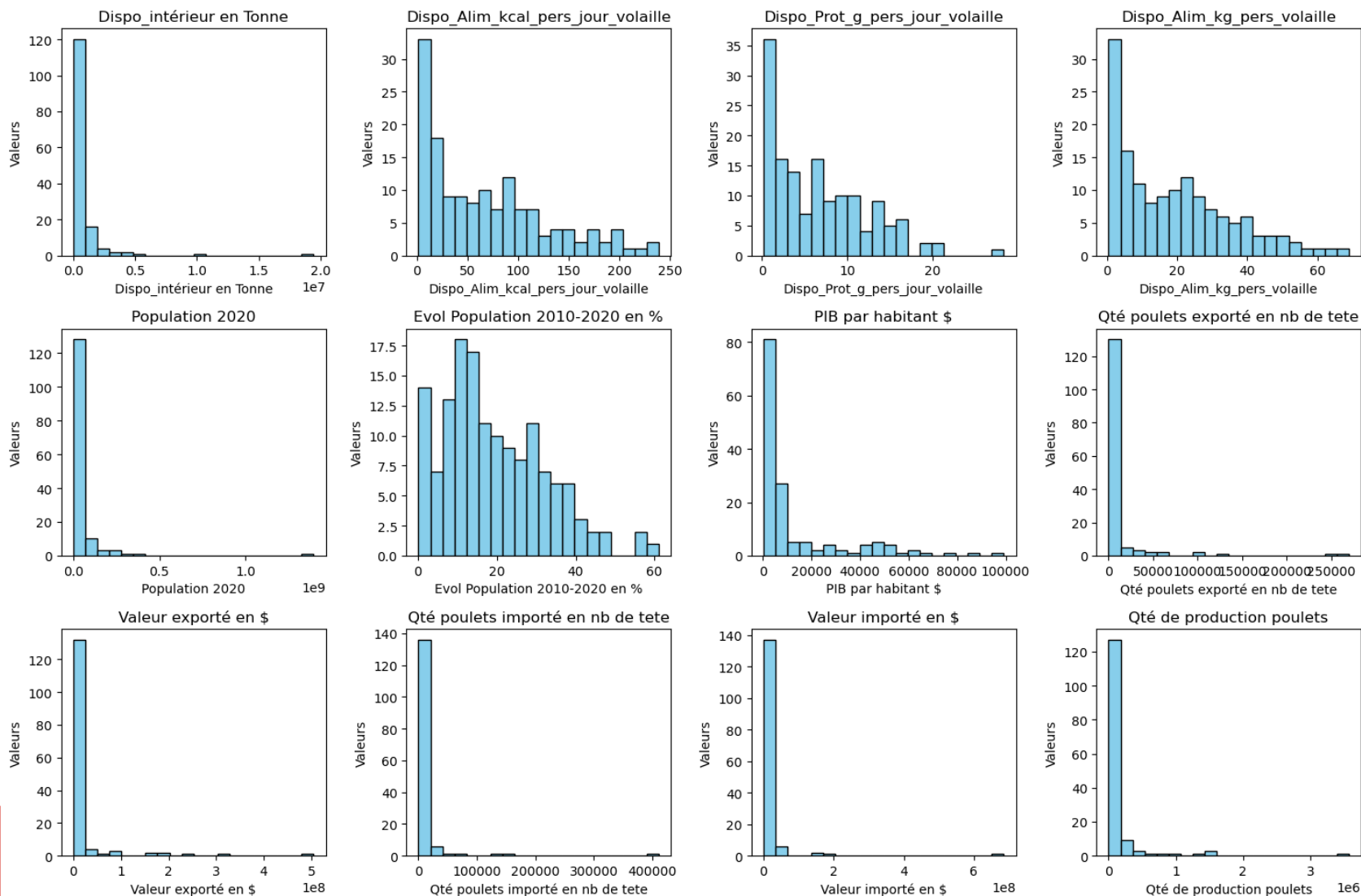
- Cible les pays avec croissance démographique positive
- Population supérieure à 500 000 habitants en 2020
- Concentration sur les marchés pertinents pour la consommation de poulet
- Élimination des pays moins significatifs pour réduire le "bruit"
- Mise en évidence des zones potentiellement rentables pour l'industrie de la volaille





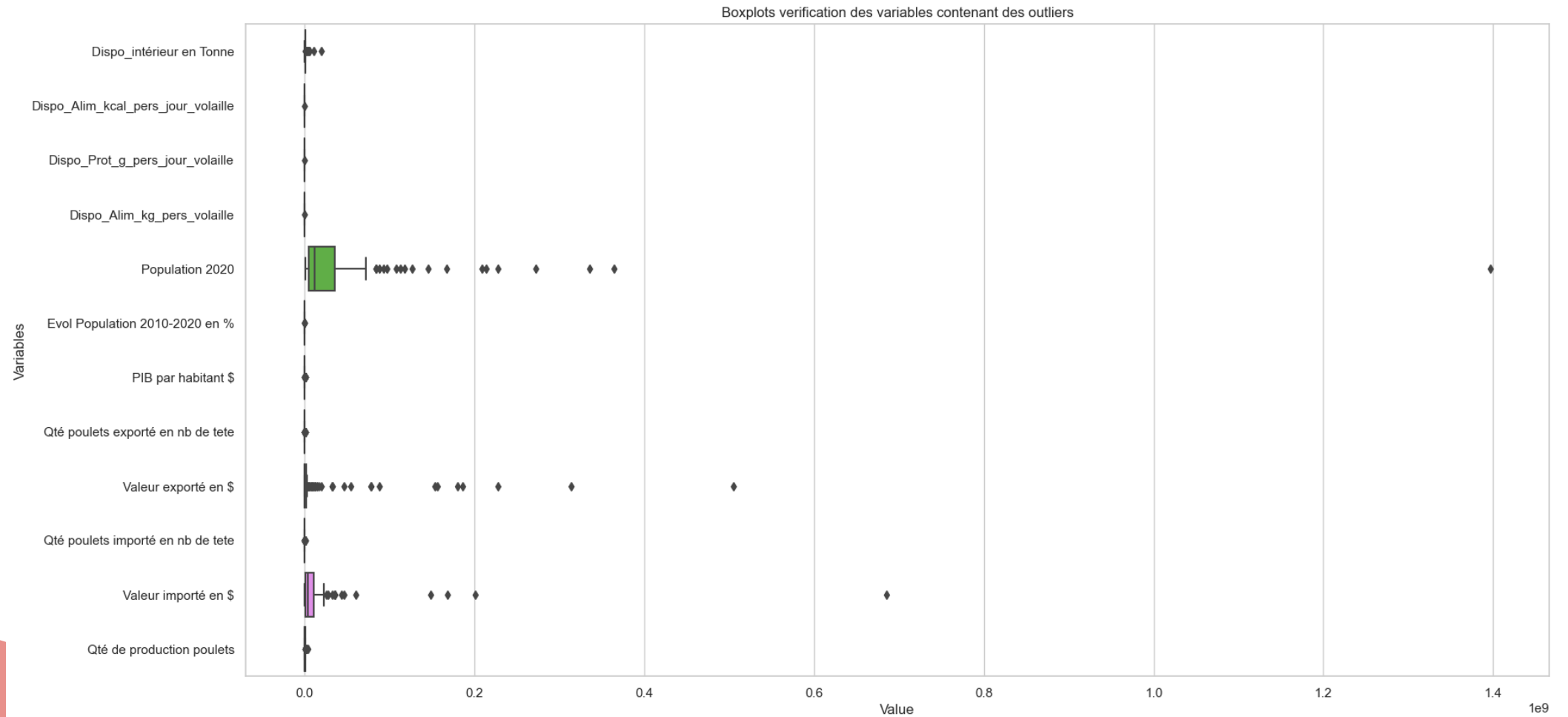
# Analyse exploratoire

## 1.1a Analyse univarié



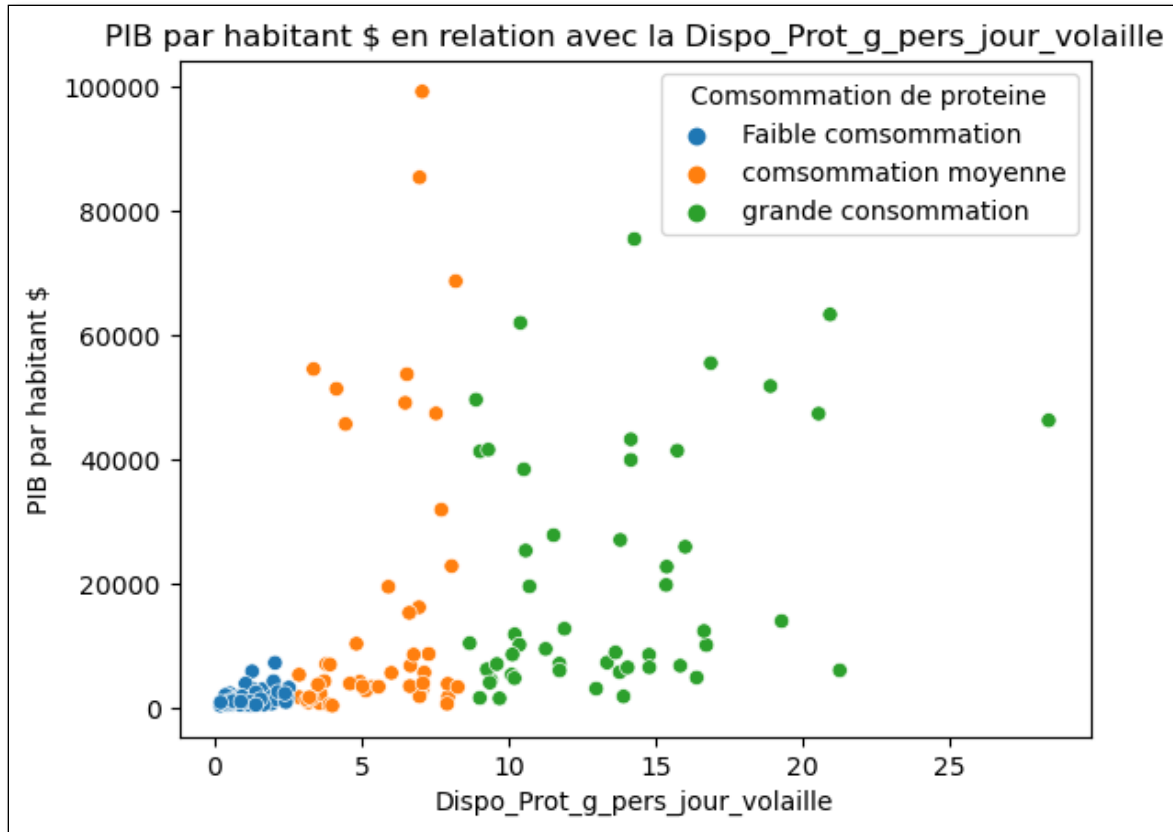
# Analyse exploratoire

## 1.1 b Analyse univarié



# Analyse exploratoire

## 1.2 Analyse bivariée

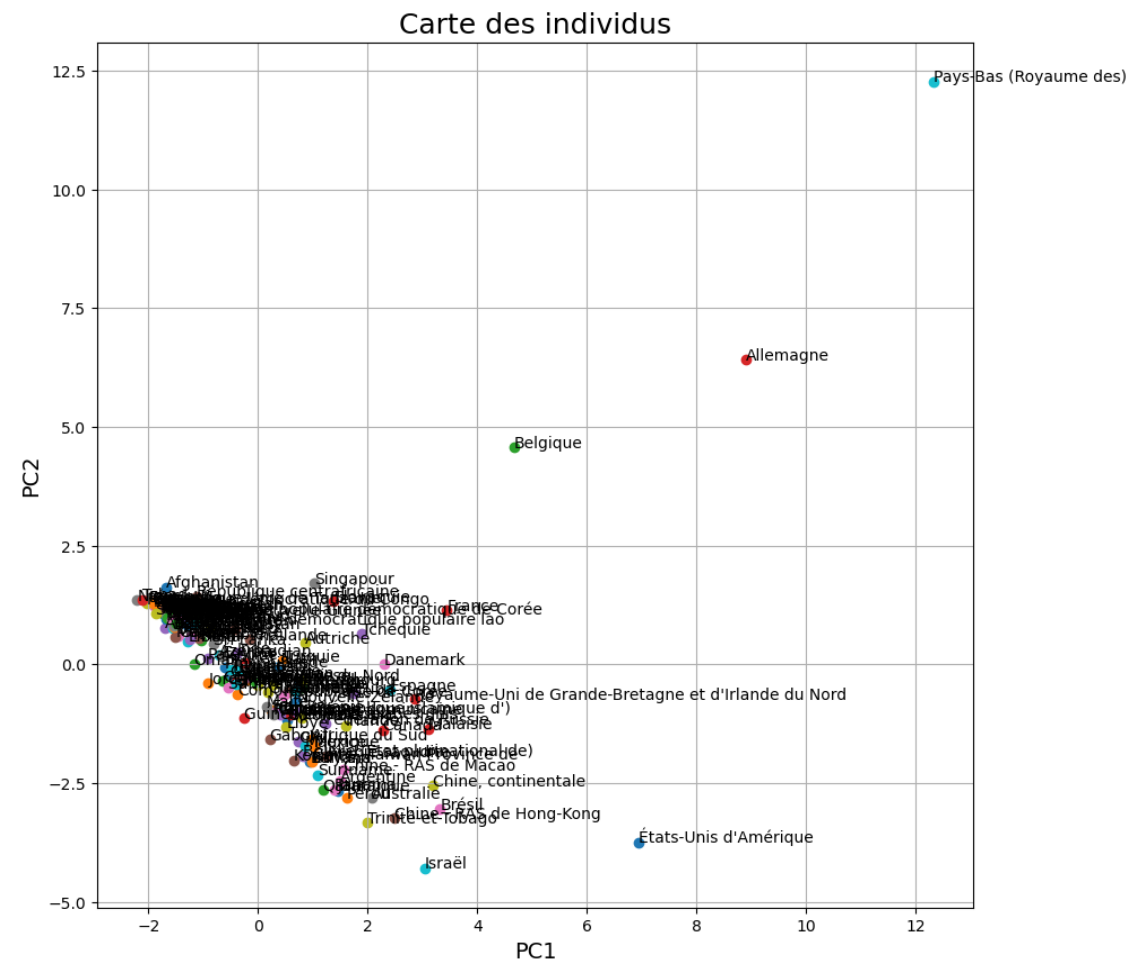
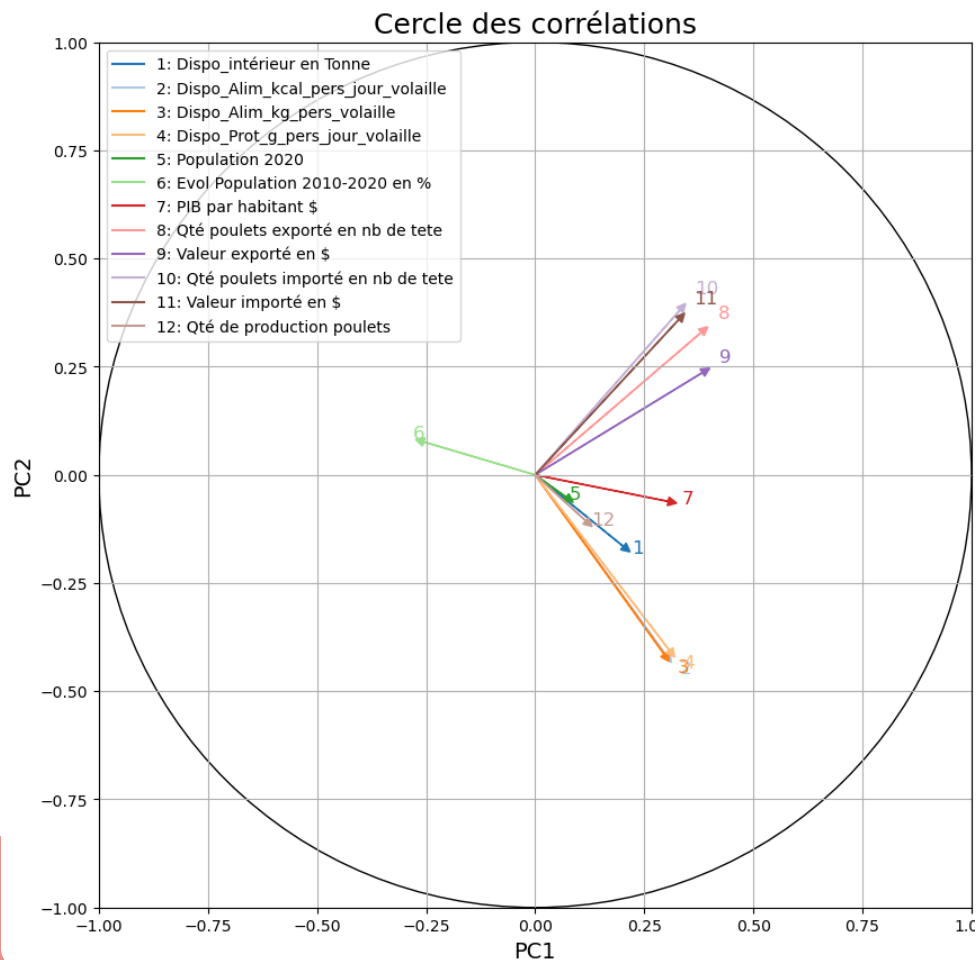


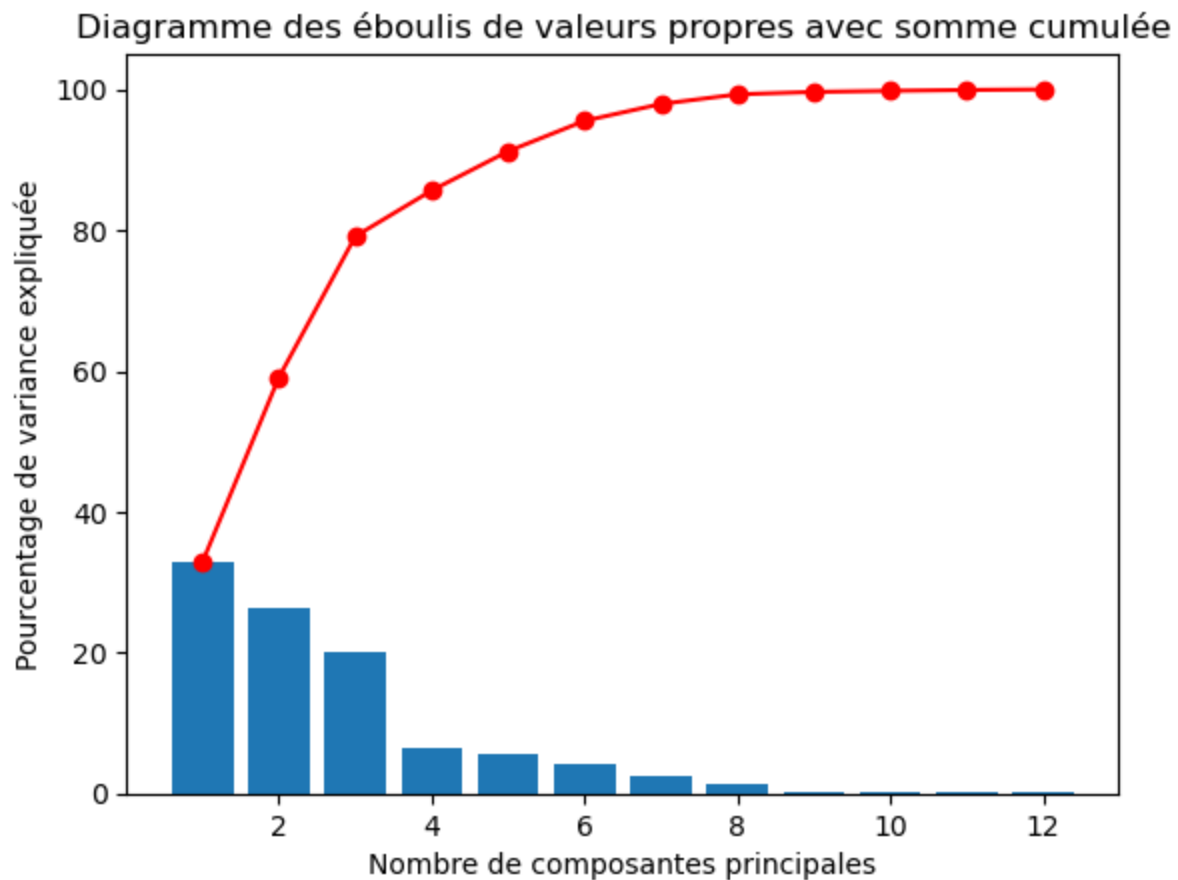
Matrice de corrélation

Dispo_intérieur en Tonne	1	0.49	0.48	0.47	0.71	0.37	0.4	0.42	0.31	0.61	0.8
Dispo_Aliment_kcal_pers_jour_volaille	0.49	1	0.99	0.99	-0.17	0.69	0.18	0.2	0.1	0.22	0.16
Dispo_Aliment_kg_pers_jour_volaille	0.48	0.99	1	0.99	-0.19	0.69	0.18	0.2	0.093	0.21	0.15
Dispo_Prot_g_pers_jour_volaille	0.47	0.99	0.99	1	-0.2	0.72	0.2	0.22	0.096	0.21	0.13
Population 2020	0.71	-0.17	-0.19	-0.2	1	-0.17	0.27	0.28	0.21	0.47	0.82
PIB par habitant \$	0.37	0.69	0.69	0.72	-0.17	1	0.38	0.4	0.24	0.36	0.086
Qté poulets exporté en nb de tete	0.4	0.18	0.18	0.2	0.27	0.38	1	0.98	0.39	0.46	0.29
Valeur exporté en \$	0.42	0.2	0.2	0.22	0.28	0.4	0.98	1	0.37	0.48	0.31
Qté poulets importé en nb de tete	0.31	0.1	0.093	0.096	0.21	0.24	0.39	0.37	1	0.76	0.23
Valeur importé en \$	0.61	0.22	0.21	0.21	0.47	0.36	0.46	0.48	0.76	1	0.54
Qté de production poulets	0.8	0.16	0.15	0.13	0.82	0.086	0.29	0.31	0.23	0.54	1

# Analyse exploratoire

## 2. Multivarié en composante principale (ACP)





	PC	Noms de Variables	Variance Expliquée (%)	Variance Cumulative (%)
0	PC1	Dispo_intérieur en Tonne	33.81%	33.81%
1	PC2	Dispo_Aliment_kcal_pers_jour_volaille	26.48%	60.29%
2	PC3	Dispo_Prot_g_pers_jour_volaille	15.74%	76.03%
3	PC4	Dispo_Aliment_kg_pers_volaille	6.64%	82.67%
4	PC5	Population 2020	5.68%	88.35%
5	PC6	Evol Population 2010-2020 en %	4.64%	92.99%
6	PC7	PIB par habitant \$	4.06%	97.05%
7	PC8	Qté poulets exporté en nb de tete	2.25%	99.30%
8	PC9	Valeur exporté en \$	0.34%	99.64%
9	PC10	Qté poulets importé en nb de tete	0.16%	99.80%
10	PC11	Valeur importé en \$	0.12%	99.92%
11	PC12	Qté de production poulets	0.08%	100.00%

# Méthodes de segmentation

## Choix du meilleur K

```
### Résumé de l'Analyse de Segmentation par KMeans
```

```
#### Utilisation du graphique du coude pour déterminer k
```

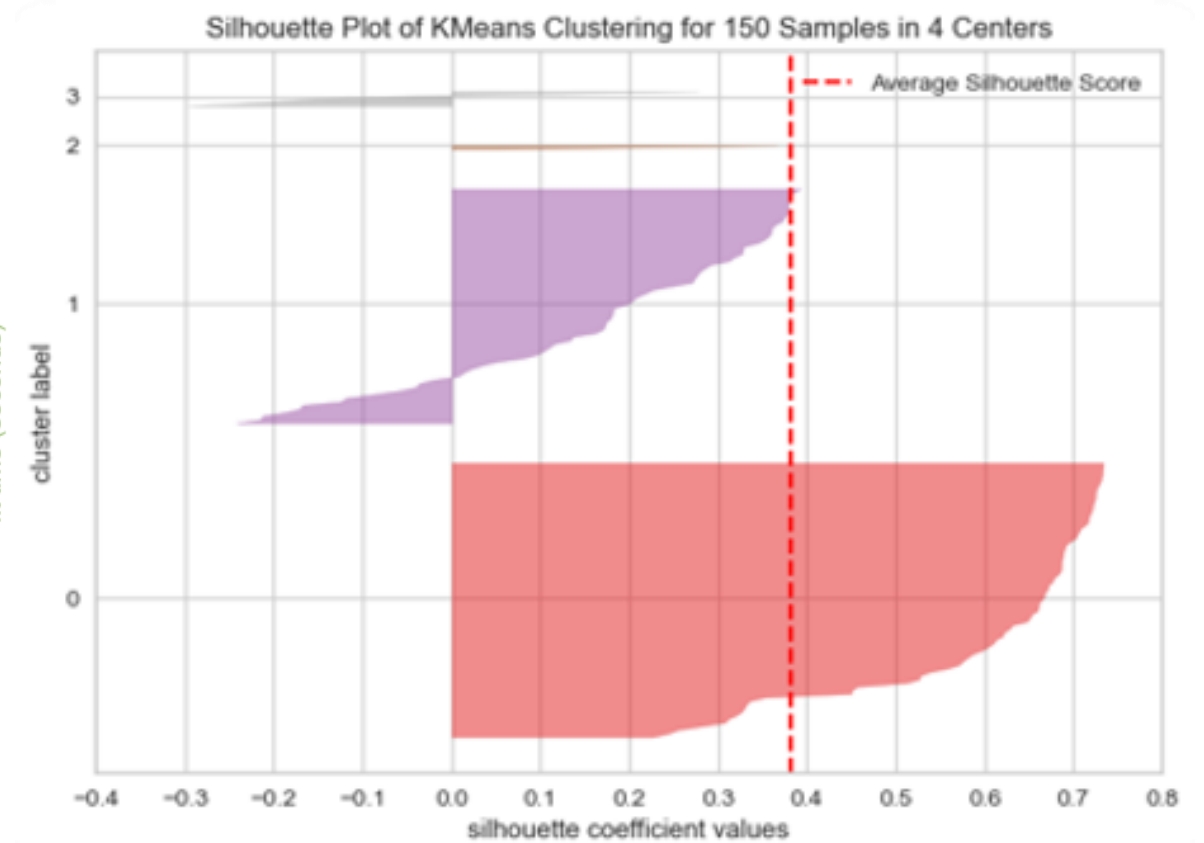
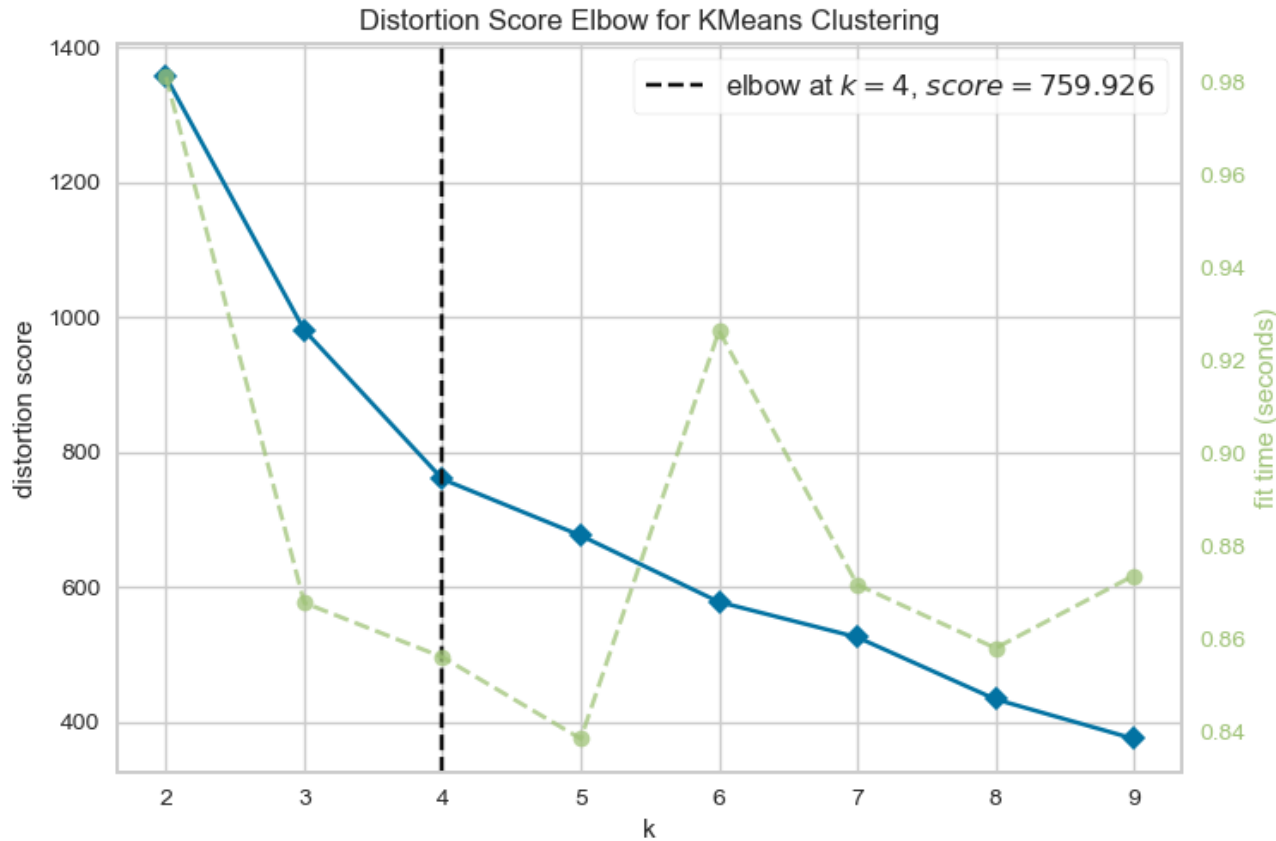
```
Le meilleur k selon le graphique du coude est 4.
```

```
#### Inertie et Score de Silhouette pour différentes valeurs de k
```

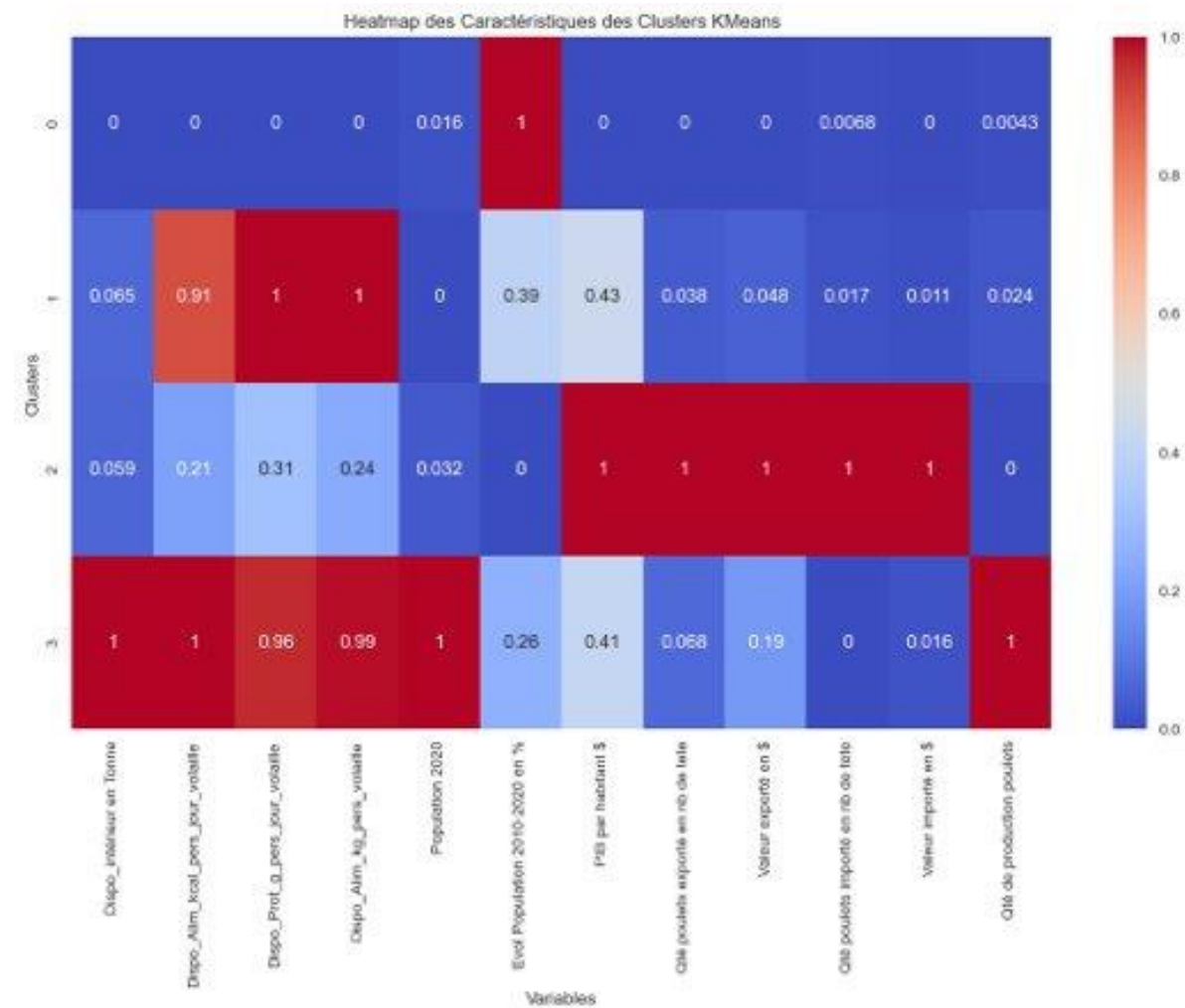
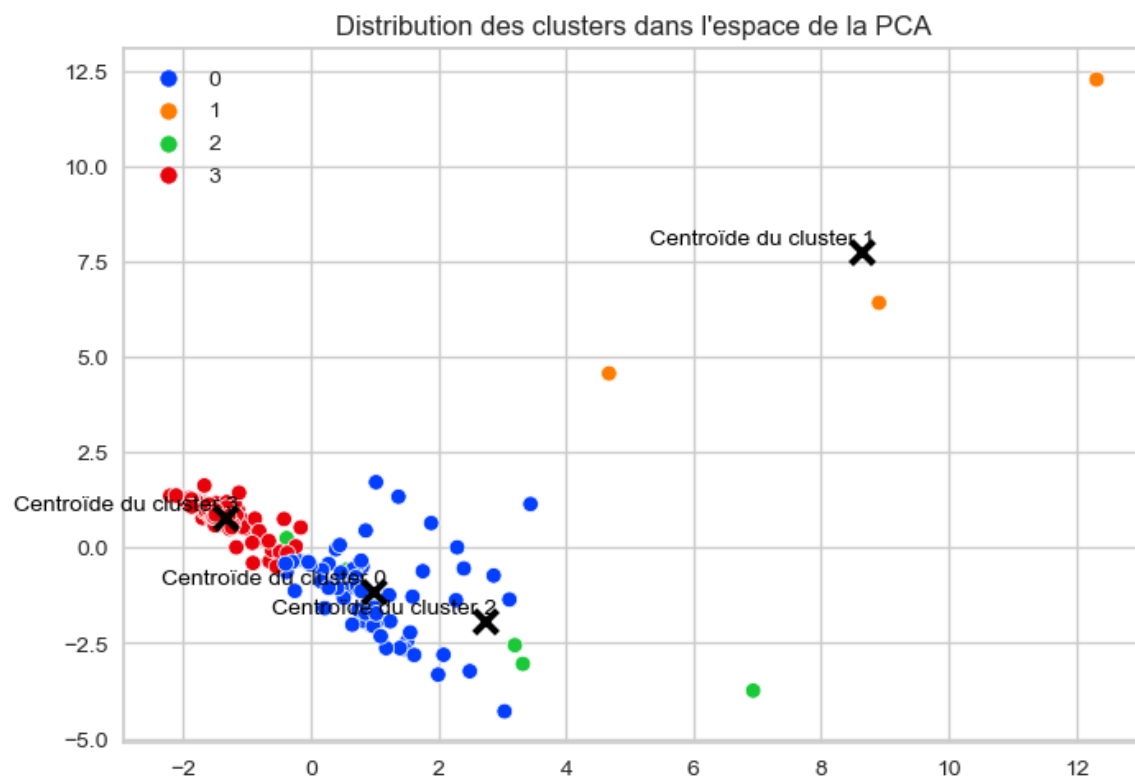
```
- k = 3 : Inertie = 981.9424873660778, Score de Silhouette = 0.37730858320587646  
- k = 4 : Inertie = 760.2523423438734, Score de Silhouette = 0.37696179383109785  
- k = 5 : Inertie = 667.3685966077737, Score de Silhouette = 0.40157719410702525  
- k = 6 : Inertie = 601.8225270044753, Score de Silhouette = 0.3988897126426677  
- k = 7 : Inertie = 509.8935103761287, Score de Silhouette = 0.40059936301904486  
- k = 8 : Inertie = 433.4237450176321, Score de Silhouette = 0.3390789197178579  
- k = 9 : Inertie = 375.7881235060469, Score de Silhouette = 0.3415180652504426  
- k = 10 : Inertie = 340.58128404595425, Score de Silhouette = 0.3513019464963659
```

# Segmentation

## 1. Clustering kmean







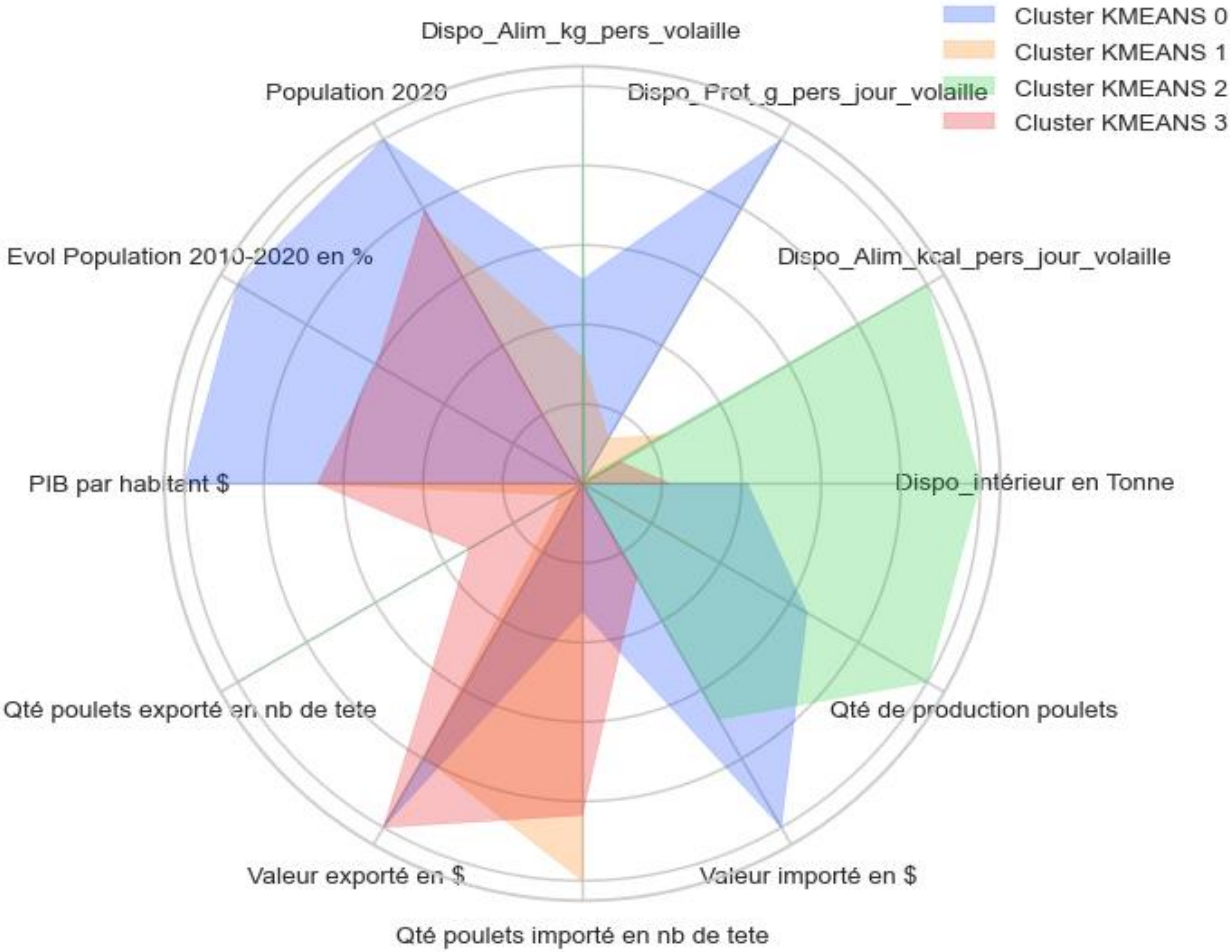
Liste des pays par cluster:

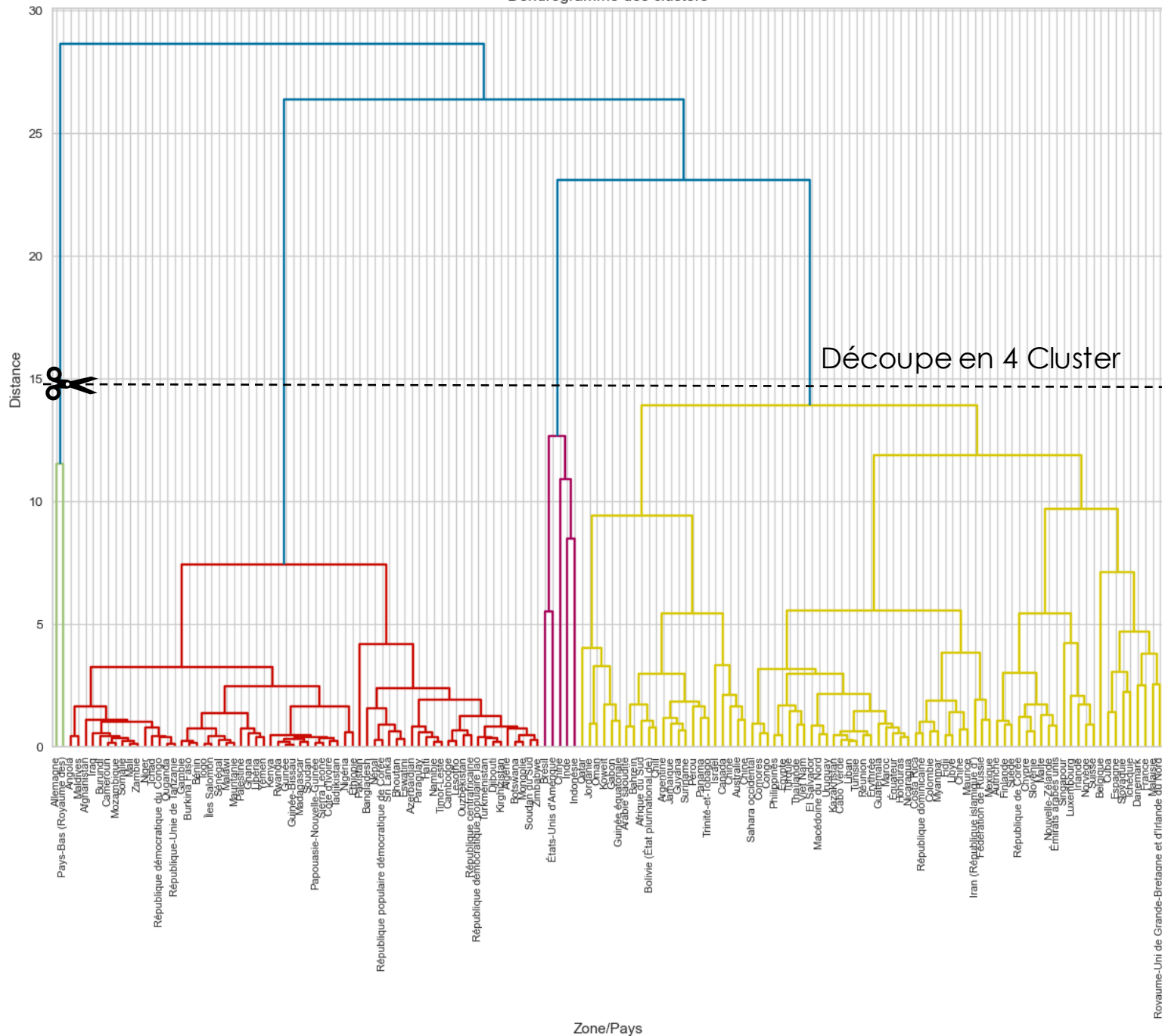
**Cluster 0:** 'Algérie', 'Singapour', 'Thaïlande', 'Zimbabwe', 'Égypte'

**Cluster 1:** 'Autriche', 'Danemark', 'Fédération de Russie', 'Iran', 'Irlande', 'Israël'

**Cluster 2:** 'Allemagne', 'Belgique', 'Pays-Bas'

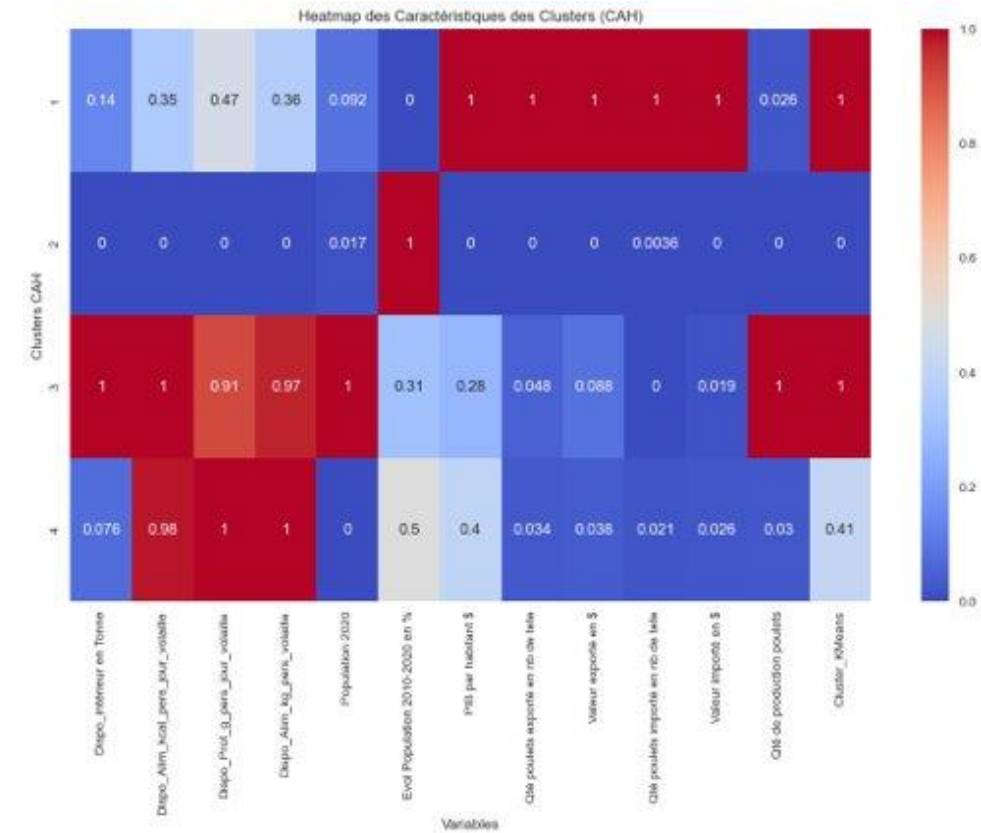
**Cluster 3:** 'Brésil', 'Chine', 'Inde', 'Indonésie', 'États-Unis d'Amérique'





# Segmentation

## 2.CAH (Classification Ascendante Hiérarchique)



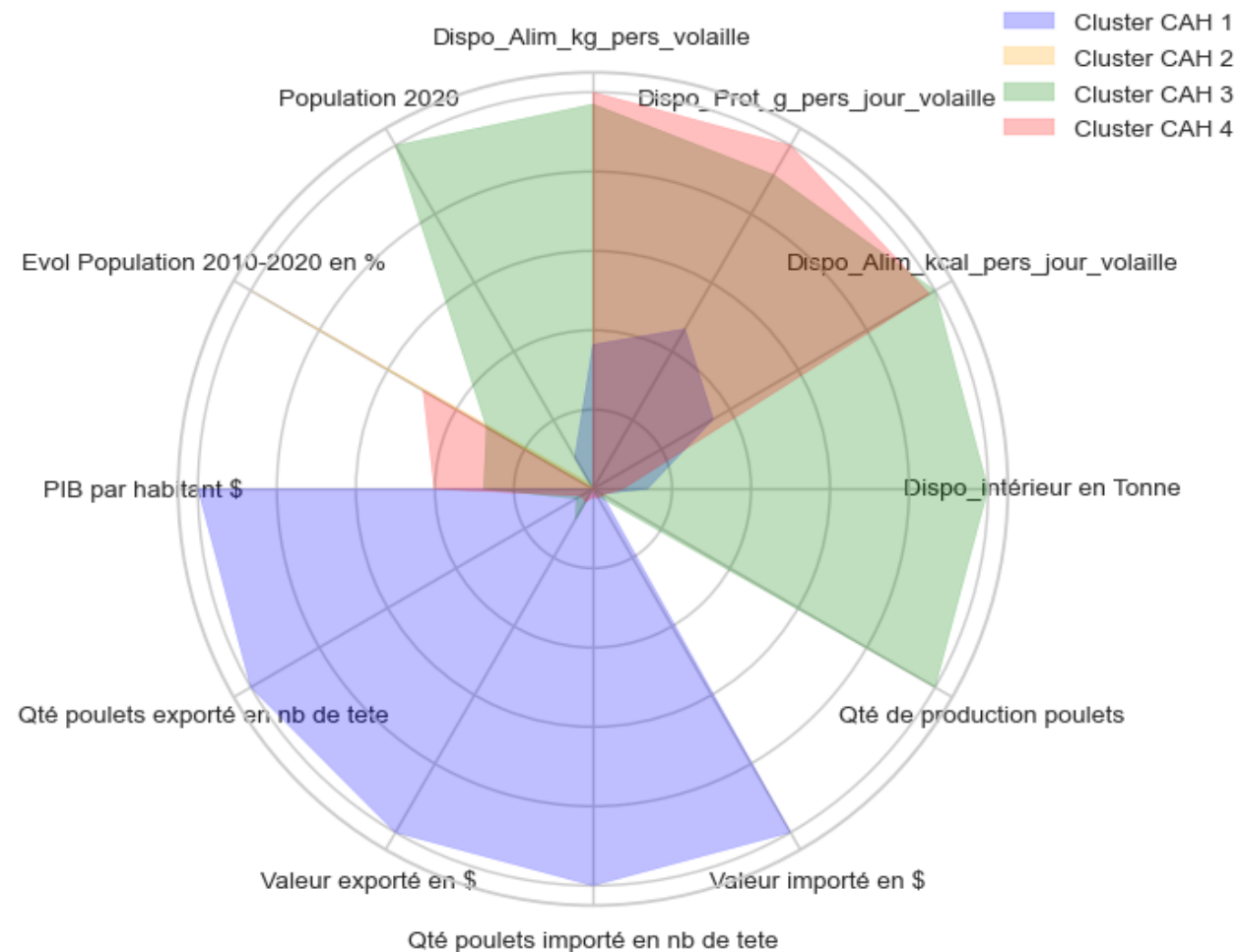
## Liste des pays par cluster:

**Cluster 1:** 'Allemagne', 'Pays-Bas'

**Cluster 2:** 'Algérie', 'Tadjikistan', 'Tchad', 'Timor-Leste', 'Éthiopie'

**Cluster 3:** 'Brésil', 'Chine', 'Fédération de Russie', 'Inde', 'Indonésie', 'Iran', 'Mexique', 'Pakistan', 'États-Unis d'Amérique'

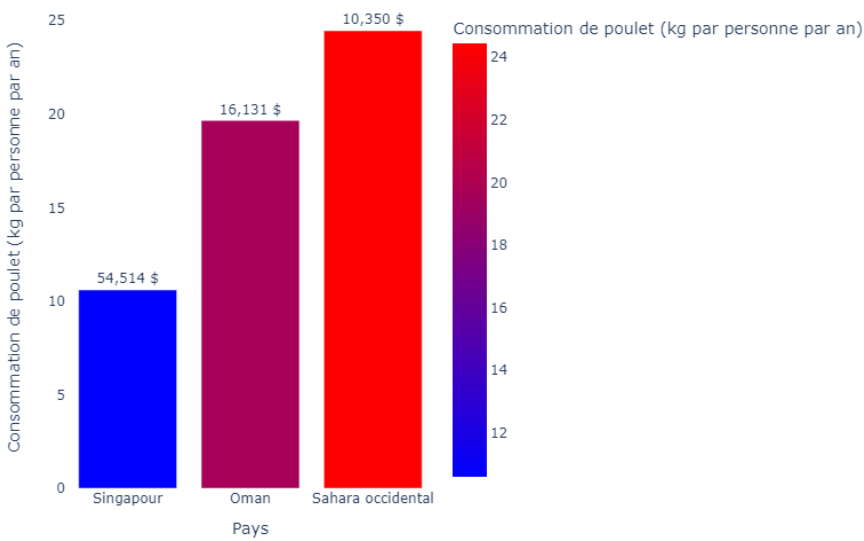
**Cluster 4:** 'Autriche', 'Belgique', 'Irlande', 'Israël', 'Koweït', 'Liban'



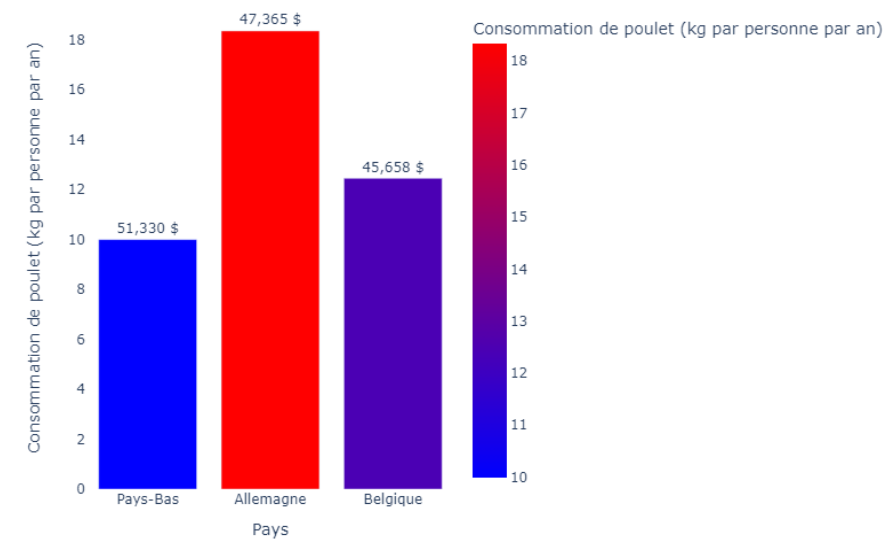
# Comparaisons des méthodes de clustering

Modèle	Silhouette Coefficient	Davies-Bouldin Index	Calinski-Harabasz Index
KMeans	0.388928	1.01572	67.961682
CAH	0.354264	0.98246	64.362513

Groupe 0: Top 3 des pays en consommation de poulet et PIB par habitant



Groupe 2: Top 3 des pays en consommation de poulet et PIB par habitant



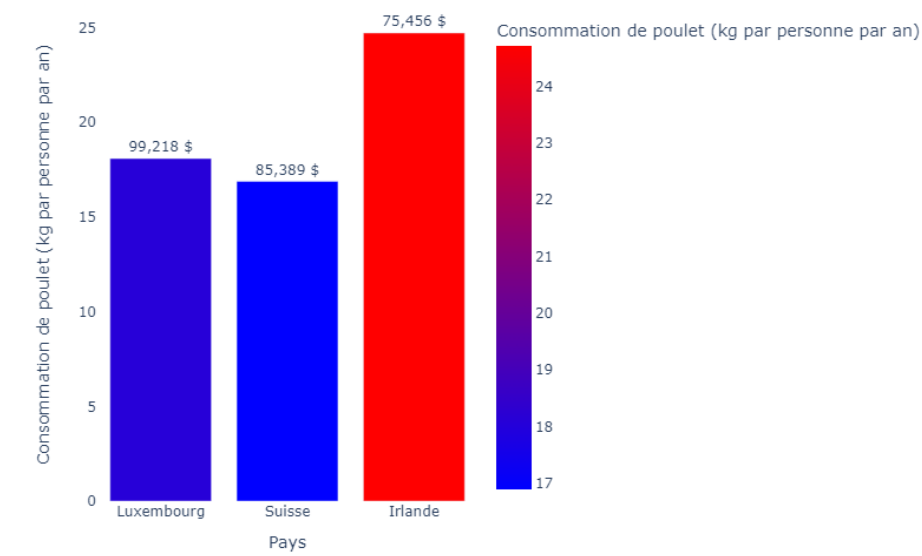
**Cluster 0 :**  
Population : Petite mais croissante rapidement (25%).  
PIB par habitant : Faible.  
Consommation et production de volaille : Réduites.  
Opportunités d'exportation : Potentiel pour les marchés émergents. Peut nécessiter des investissements pour développer le marché.

**Cluster 1 :**  
Population : Moyenne avec une croissance modérée.  
PIB par habitant : Plus élevé.  
Consommation et production de volaille : Plus importantes.  
Opportunités d'exportation : Marché plus développé avec potentiel pour l'exportation haut de gamme.

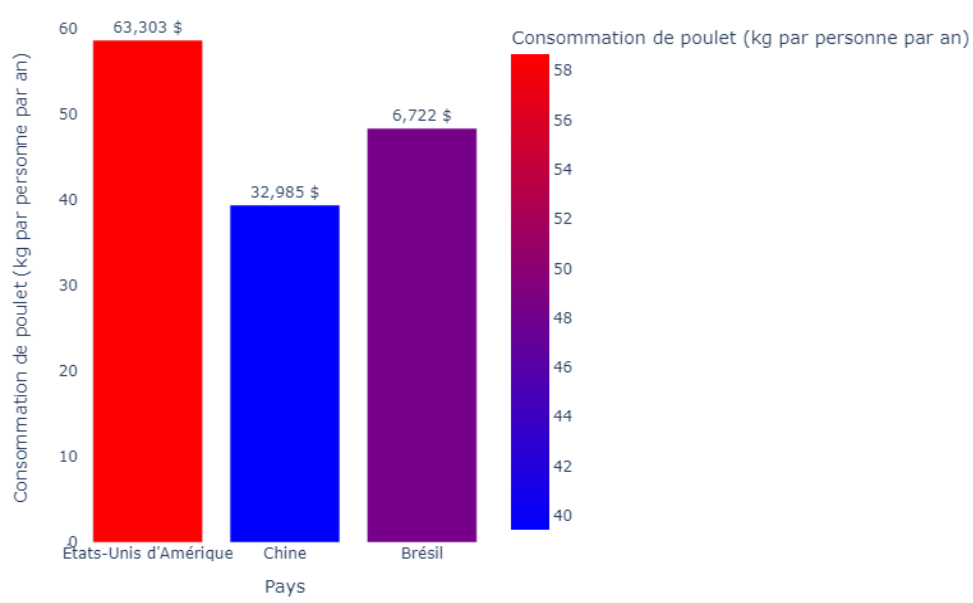
**Cluster 2 :**  
Population : Similaire au Cluster 1, mais croissance plus faible.  
PIB par habitant : Encore plus élevé.  
Importation de volaille : Très élevée.  
Opportunités d'exportation : Marché riche avec une forte demande. Peut-être le plus attrayant pour l'exportation si l'on peut rivaliser en termes de prix et de qualité.

**Cluster 3 :**  
Population : Très grande avec une croissance modérée.  
PIB par habitant : Intermédiaire.  
Consommation et production de volaille : Très élevées.  
Opportunités d'exportation : Marché vaste avec potentiel, mais peut nécessiter une stratégie ciblée pour se différencier.

Groupe 1: Top 3 des pays en consommation de poulet et PIB par habitant



Groupe 3: Top 3 des pays en consommation de poulet et PIB par habitant









# Conclusion

## Marché Européen

- **Allemagne:** Avec un PIB par habitant élevé de 47 366 USD et des importations de volaille de 167,6 millions de dollars, l'Allemagne pourrait être un marché très attrayant.
- **Autriche:** Bien que la disponibilité intérieure de volaille soit relativement faible (156 000 tonnes), l'Autriche a un PIB par habitant élevé de 49 076 USD et importe pour 43,4 millions de dollars de volaille.
- **Belgique:** Avec un PIB par habitant de 45 659 USD et des importations de volaille de 200,4 millions de dollars, la Belgique présente également un potentiel élevé.
- **Danemark:** Le Danemark a le PIB par habitant le plus élevé parmi les pays sélectionnés (61 961 USD) et importe pour 9,8 millions de dollars de volaille.
- **Espagne:** Bien que son PIB par habitant soit plus bas (26 994 USD), l'Espagne a une grande disponibilité intérieure de volaille (1 584 000 tonnes) et importe pour 34,9 millions de dollars de volaille.
- **Irlande:** Avec le PIB par habitant le plus élevé de la liste (75 456 USD) et des importations de 8,8 millions de dollars, l'Irlande pourrait être un marché niche.
- **Suède:** La Suède a un PIB par habitant de 53 677 USD et importe pour 6,8 millions de dollars de volaille.

## Marché Mondiale

- **Fédération de Russie, Iran, Israël, Mexique, Pakistan:** Ces pays sont également prometteur et pourraient représenter des marchés secondaires intéressants à explorer.


## Grand Marché Mondiale

- **Brésil, Chine, Inde, Indonésie, États-Unis:** Ces pays sont présents dans les cluster les plus prometteurs selon les deux méthodes de clustering. Ils ont une grande population et une grande disponibilité intérieure de volaille, ce qui pourrait indiquer un marché robuste pour les produits de volaille.

## Conseil:

- Il serait également bénéfique de considérer d'autres facteurs tels que la réglementation locale, la stabilité politique, les préférences culturelles, les canaux de distribution existants, les barrières tarifaires, etc. Une analyse plus approfondie en tenant compte de ces facteurs pourrait vous aider à prendre une décision plus éclairée.





Merci pour  
votre  
attention

Avez vous des questions ?