

# Etude de marché

# La poule qui chante

Projet de développement à l'internationale pour  
l'exportation de volailles

# Contexte

Spécialisation dans la production et l'exportation de volailles.

Projet de développement à l'international.

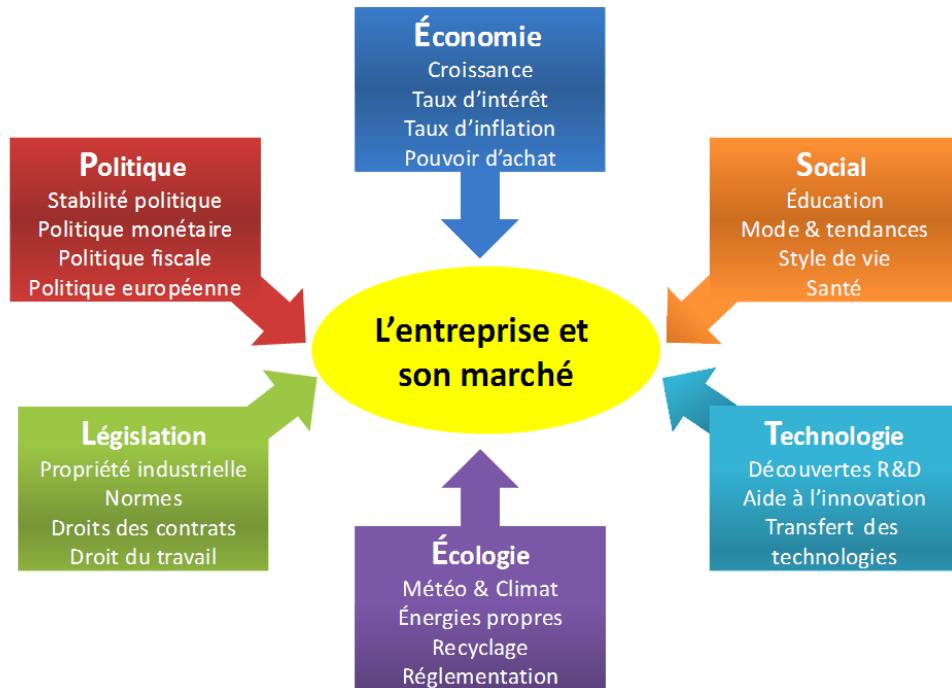
Pour le moment, aucune piste de pays ni même de continent cible !



**Objectif →** A l'issue d'une première analyse, proposer un groupe de pays pouvant être des cibles potentielles pour l'exportation de nos poulets.

# Analyse PESTEL

Le but de l'analyse PESTEL est d'identifier les facteurs qui pourraient avoir un impact sur l'entreprise.



Dans le contexte de notre projet de développement actuel, en plus des éléments propre à notre activité (production de volaille, consommation, importation, etc.), les composantes de l'analyse qui vont nous concerner sont les suivantes :

- Politique
  - Stabilité politique
- Economique
  - PIB
  - RNB
  - Inflation
- Sociologique
  - Démographie

# Présentation des données

186 pays

Nous nous intéressons principalement aux données de disponibilités alimentaire, production, import/export de volailles. Nous regardons également les indicateurs de l'analyse PESTEL que nous avons retenus. Voici la liste des variables :

```
Disponibilité alimentaire (Kcal)_Viande de Volailles_kcal
Disponibilité alimentaire (Kcal/personne/jour)_Viande de Volailles_Kcal/personne/jour
Disponibilité alimentaire en quantité (kg/personne/an)_Viande de Volailles_kg
Disponibilité de matière grasse en quantité (g)_Viande de Volailles_g
Disponibilité de matière grasse en quantité (g/personne/jour)_Viande de Volailles_g/personne/jour
Disponibilité de protéines en quantité (g)_Viande de Volailles_g
Disponibilité de protéines en quantité (g/personne/jour)_Viande de Volailles_g/personne/jour
Disponibilité intérieure_Viande de Volailles_Milliers de tonnes
Exportations - Quantité_Viande de Volailles_Milliers de tonnes
Food price inflation_%
Importations - Quantité_Viande de Volailles_Milliers de tonnes
Nourriture_Viande de Volailles_Milliers de tonnes
Pertes_Viande de Volailles_Milliers de tonnes
Population totale_Population-Estimations_1000 personnes
Production_Viande de Volailles_Milliers de tonnes
Résidus_Viande de Volailles_Milliers de tonnes
Traitement_Viande de Volailles_Milliers de tonnes
Valeur US dollars par habitant_Produit Intérieur Brut_US dollars
Valeur US dollars par habitant_Revenu national brut_US dollars
Valeur_Stabilité politique et absence de violence/terrorisme (indice)_indice
Variation de stock_Viande de Volailles_Milliers de tonnes
```

# Méthodologie

## 1. ACP

- Comprendre l'influence des variables
- Eventuellement en réduire le nombre

## 2. Mise en évidence des outliers

- Utilisation de IsolationForest pour isoler les pays éloignés de autres
- Déterminer si oui ou non ils ont leur place dans notre analyse

## 3. Clustering pour créer des groupes de pays

### a) ACP

- On refait une ACP après réduction des variables et suppressions des outliers

### b) Kmeans

### c) CAH

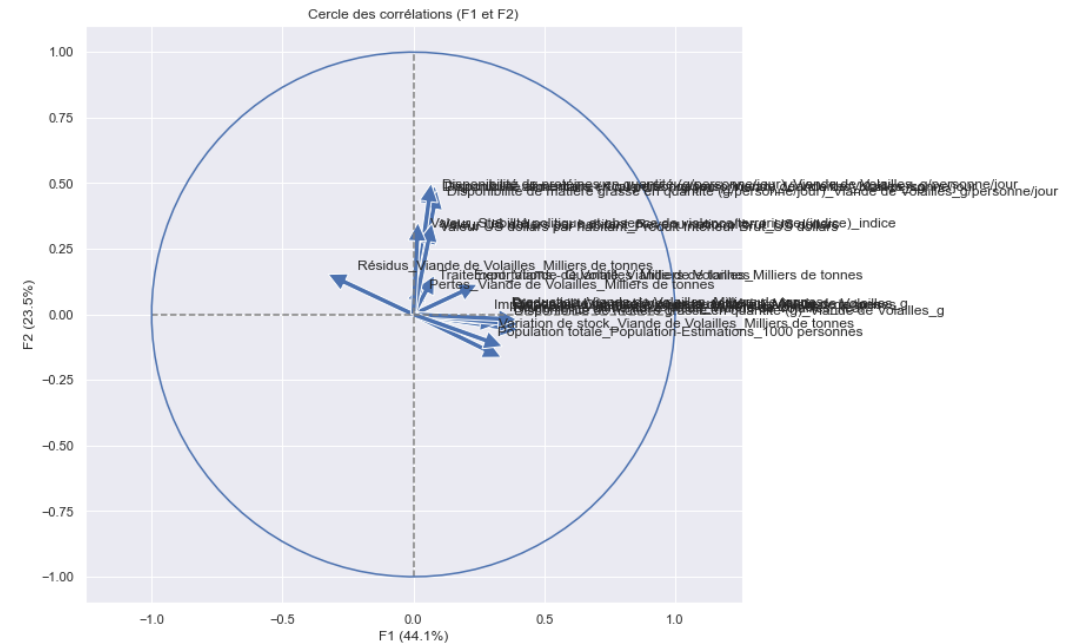
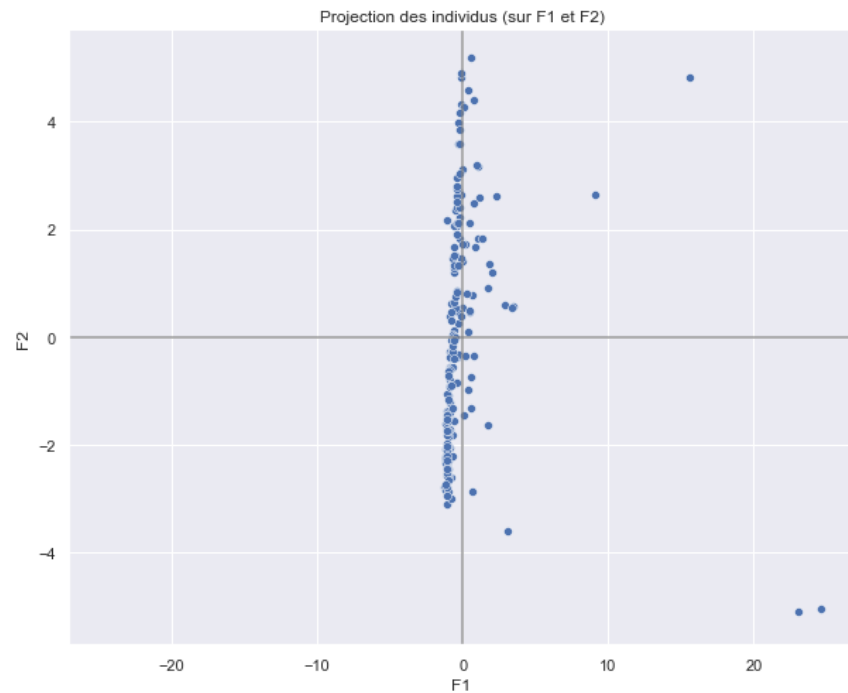
### d) Comparaison des deux méthodes de clustering

- En s'appuyant sur les résultats de l'ACP
- Compréhension des groupes de pays et du lien entre les variables et les individus

# 1. ACP

Analyse globale des données à l'aide de l'ACP

- ➔ Comprendre l'importance et l'implication des variables
- ➔ Simplifier le jeu de données en éliminant les variables non pertinentes et réduire les groupes de variables très corrélées entre elles



- ➔ **Conclusion** : On voit sur le cercle de corrélation que certaines variables sont très corrélées entre elles. Nous pouvons en savoir plus en regardant la matrice de corrélation.

Disponibilité alimentaire (Kcal)_Viande de Volailles_kcal	0.33	-0.04	-0.03	0.05	-0.02	-0.05	-0.02	0.00	-0.04	-0.17
Disponibilité alimentaire (Kcal/personne/jour)_Viande de Volailles_Kcal/personne/jour	0.07	0.42	-0.24	-0.04	-0.03	0.08	0.07	0.03	-0.00	0.03
Disponibilité alimentaire en quantité (kg/personne/an)_Viande de Volailles_kg	0.06	0.42	-0.23	-0.05	-0.06	0.09	0.09	0.01	0.02	0.01
Disponibilité de matière grasse en quantité (g)_Viande de Volailles_g	0.33	-0.05	-0.03	0.02	-0.03	-0.05	-0.03	-0.00	-0.07	-0.13
Disponibilité de matière grasse en quantité (g/personne/jour)_Viande de Volailles_g/personne/jour	0.08	0.40	-0.27	-0.03	-0.06	0.09	0.04	0.03	0.00	0.02
Disponibilité de protéines en quantité (g)_Viande de Volailles_g	0.33	-0.02	-0.01	0.13	-0.01	-0.04	-0.01	0.01	0.02	-0.24
Disponibilité de protéines en quantité (g/personne/jour)_Viande de Volailles_g/personne/jour	0.06	0.43	-0.18	-0.05	0.03	0.06	0.11	0.04	-0.02	0.06
Disponibilité intérieure_Viande de Volailles_Milliers de tonnes	0.33	-0.03	-0.02	0.11	-0.01	-0.04	-0.01	0.00	0.02	-0.23
Exportations - Quantité_Viande de Volailles_Milliers de tonnes	0.18	0.08	0.06	0.66	-0.08	-0.05	-0.19	-0.21	-0.08	0.59
Importations - Quantité_Viande de Volailles_Milliers de tonnes	0.26	-0.03	0.12	-0.30	0.12	0.25	0.17	-0.69	0.23	-0.06
Nourriture_Viande de Volailles_Milliers de tonnes	0.33	-0.03	-0.02	0.10	-0.03	-0.04	-0.01	0.01	0.02	-0.23
Pertes_Viande de Volailles_Milliers de tonnes	0.01	0.05	-0.16	0.05	0.95	-0.24	-0.11	0.02	-0.01	0.00
Population totale_Population-Estimations_1000 personnes	0.27	-0.13	-0.02	-0.16	0.00	-0.03	0.06	0.49	0.71	0.33
Production_Viande de Volailles_Milliers de tonnes	0.33	-0.02	-0.02	0.21	-0.03	-0.06	-0.05	0.02	-0.02	-0.10
Résidus_Viande de Volailles_Milliers de tonnes	-0.26	0.12	0.04	0.34	0.03	-0.02	0.07	-0.35	0.60	-0.10
Traitement_Viande de Volailles_Milliers de tonnes	0.05	0.09	0.40	0.17	0.25	0.81	-0.03	0.26	-0.05	-0.05
Valeur US dollars par habitant_Produit Intérieur Brut_US dollars	0.05	0.27	0.51	-0.04	0.02	-0.32	0.25	0.08	-0.06	0.00
Valeur US dollars par habitant_Revenu national brut_US dollars	0.06	0.28	0.51	-0.03	0.02	-0.28	0.24	0.05	-0.03	0.00
Valeur_Stabilité politique et absence de violence/terrorisme (indice)_indice	0.01	0.28	0.23	-0.26	-0.08	-0.08	-0.87	-0.05	0.14	-0.07
Variation de stock_Viande de Volailles_Milliers de tonnes	0.27	-0.10	0.03	-0.37	0.06	0.05	0.02	-0.18	-0.19	0.55
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10

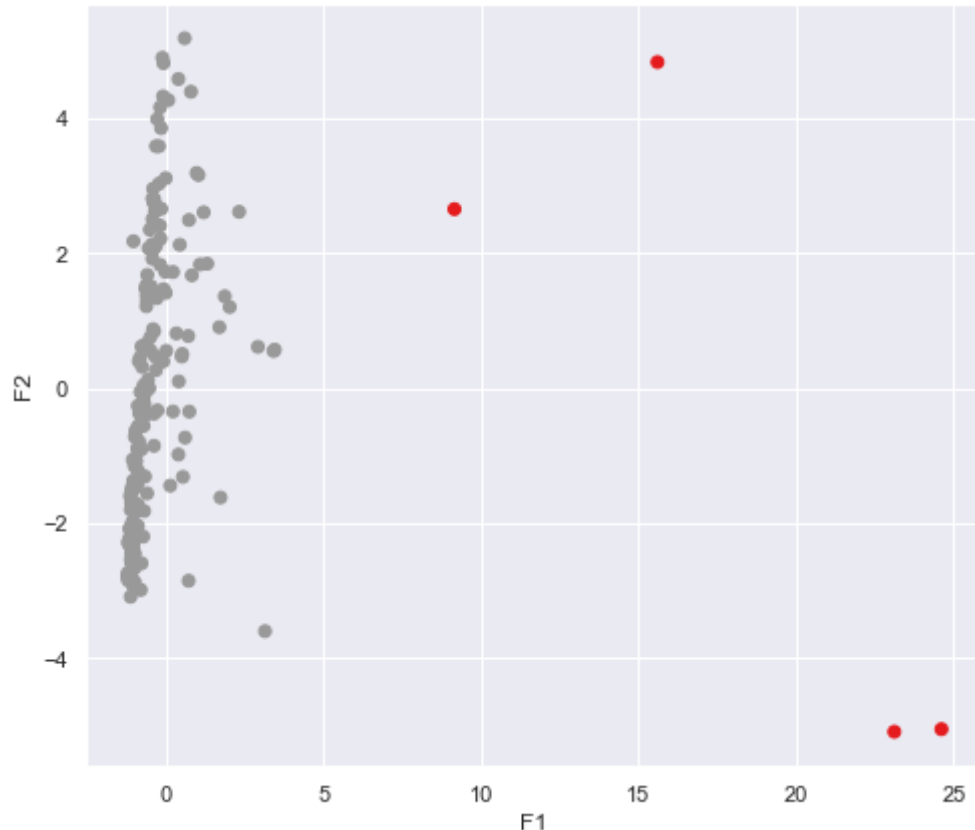
Suite à la lecture de la matrice corrélation de l'ACP, nous pouvons réduire le nombre de variables et obtenir ainsi la liste finale de variables suivante :

Disponibilité alimentaire en quantité  
(kg/personne/an)\_Volailles\_kg  
Disponibilité intérieure\_Volailles\_Milliers de tonnes  
Exportations - Quantité\_Volailles\_Milliers de tonnes  
Importations - Quantité\_Volailles\_Milliers de tonnes  
Pertes\_Volailles\_Milliers de tonnes

Production\_Volailles\_Milliers de tonnes  
Valeur US dollars par habitant\_Produit Intérieur Brut\_US dollars  
Stabilité politique et absence de violence/terrorisme  
Variation de stock\_Volailles\_Milliers de tonnes

## 2. Mise en évidence des outliers

Je veux mettre en évidence les 2% de mes pays qui sont les plus éloignés des autres



Les outliers apparaissent en rouge.

Ces pays sont les Etats-Unis, la Chine, la Chine continentale et le Brésil

Je regarde au cas par cas si ce sont des pays à conserver ou à exclure de l'analyse

➔ Chacun de ces 4 pays se distinguent du reste de notre jeu de données par certaines variables, mais le point le plus notable est leur forte production de volaille à tous les 4.

➔ Ils ne nous intéressent donc pas pour l'exportation de nos poulets

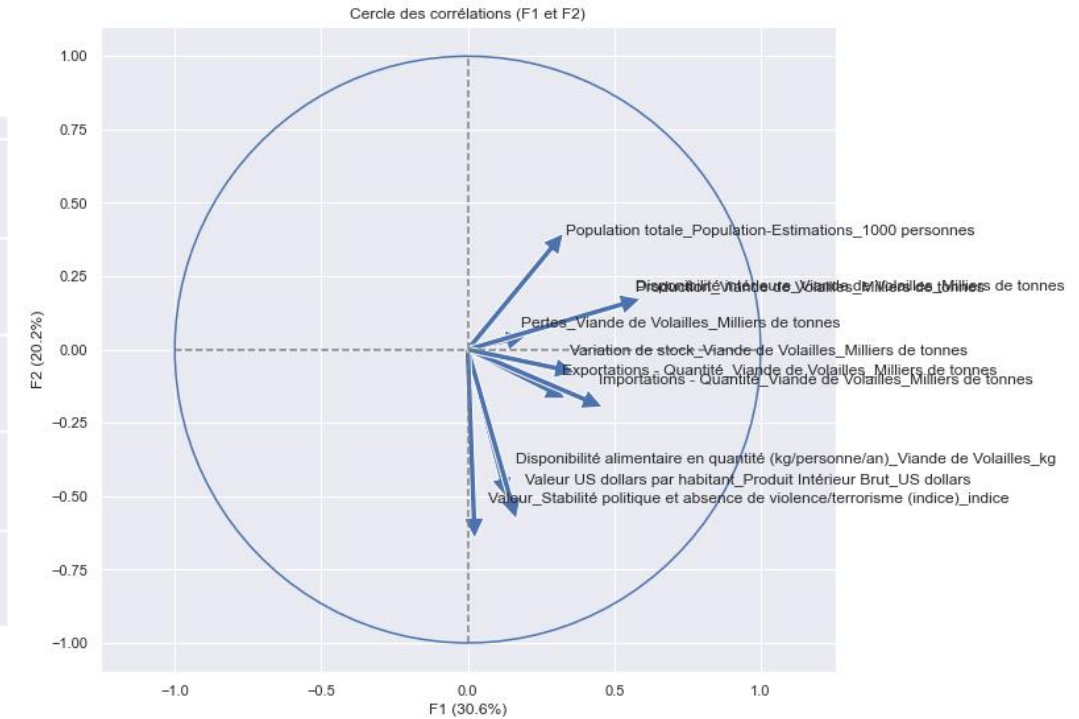
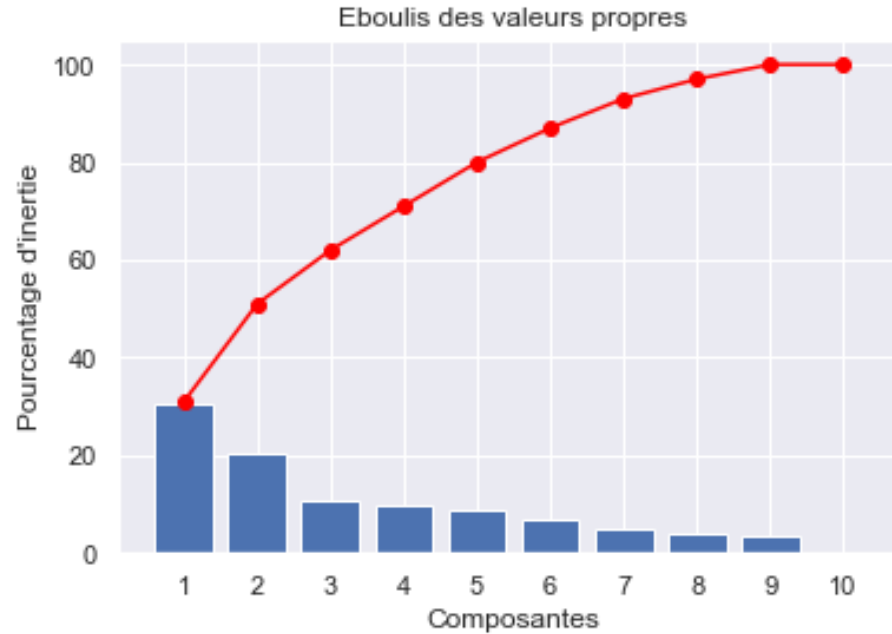
➔ On les exclut de l'analyse



# 3. Clustering

## a) ACP

Résultats de l'ACP  
avec variables  
réduites et outliers  
supprimés

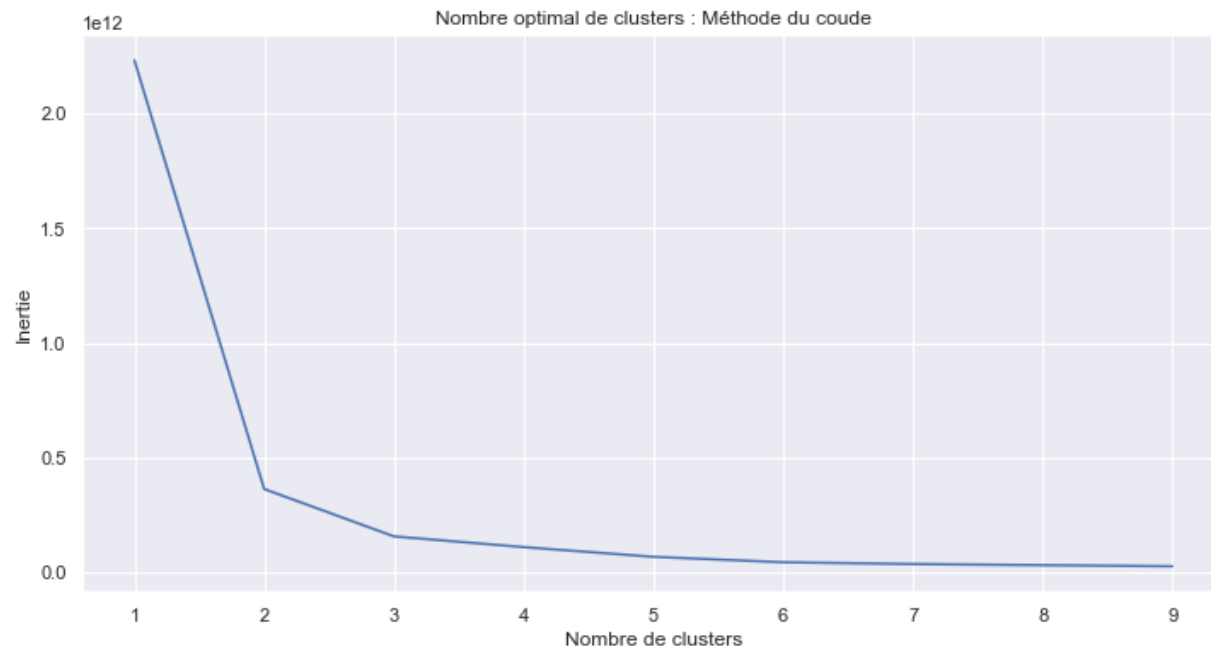


	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Disponibilité alimentaire en quantité (kg/personne/an)_Viande de Volailles_kg	0.12	-0.44	0.47	-0.17	0.20	-0.47	0.23	0.32	0.36	0.00
Disponibilité intérieure_Viande de Volailles_Milliers de tonnes	0.52	0.15	0.15	-0.14	0.07	-0.16	-0.14	-0.04	-0.36	0.69
Exportations - Quantité_Viande de Volailles_Milliers de tonnes	0.27	-0.14	-0.21	0.22	-0.79	-0.07	0.35	0.06	0.20	0.16
Importations - Quantité_Viande de Volailles_Milliers de tonnes	0.40	-0.17	-0.32	0.20	0.17	-0.28	-0.45	-0.44	0.39	-0.14
Pertes_Viande de Volailles_Milliers de tonnes	0.13	0.03	0.63	0.65	-0.01	0.36	-0.11	-0.08	0.11	0.00
Population totale_Population-Estimations_1000 personnes	0.28	0.34	0.09	-0.50	0.02	0.47	0.12	-0.09	0.55	-0.00
Production_Viande de Volailles_Milliers de tonnes	0.52	0.15	0.15	-0.11	-0.11	-0.10	0.06	0.08	-0.40	-0.69
Valeur US dollars par habitant_Produit Intérieur Brut_US dollars	0.15	-0.51	-0.11	-0.16	-0.11	0.44	-0.50	0.47	-0.06	0.00
Valeur_Stabilité politique et absence de violence/terrorisme (indice)_indice	0.02	-0.57	0.10	-0.20	0.03	0.27	0.31	-0.62	-0.25	-0.00
Variation de stock_Viande de Volailles_Milliers de tonnes	0.30	-0.06	-0.40	0.33	0.53	0.23	0.48	0.27	-0.03	0.04

## b) Kmeans

Méthode du coude pour déterminer le nombre de clusters

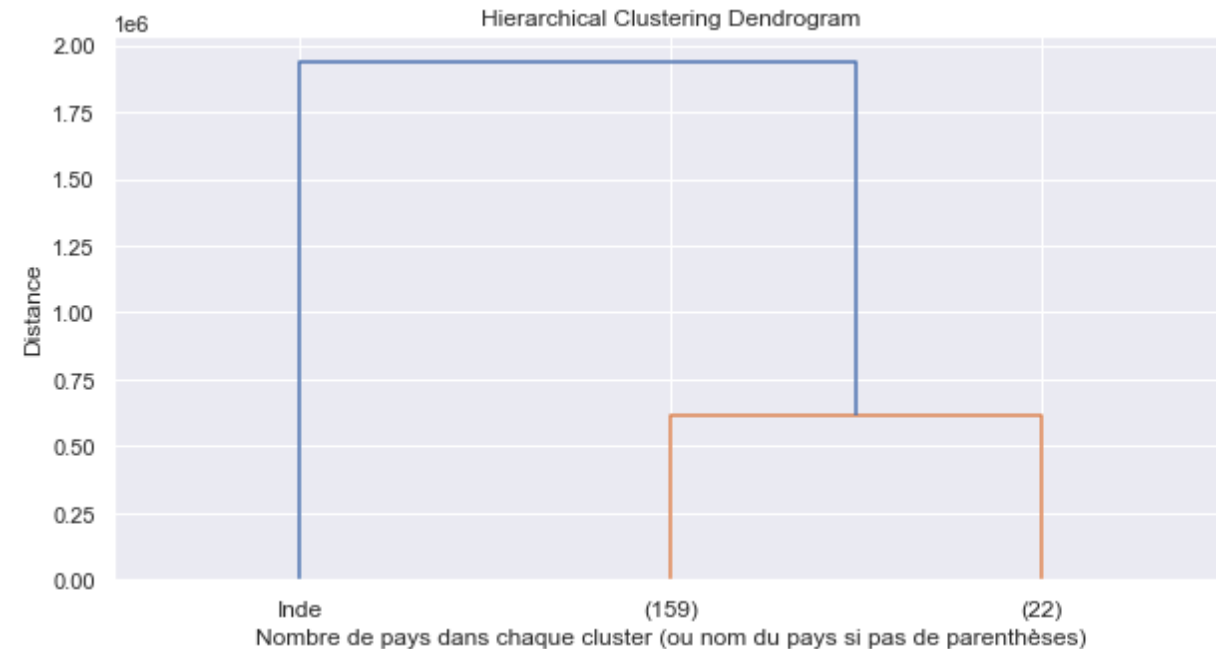
➔ 3 clusters de 1, 15 et 166 pays



## c) Classification hiérarchique

Visualisation sous forme de dendrogramme

➔ 3 clusters de 1, 22 et 159 pays

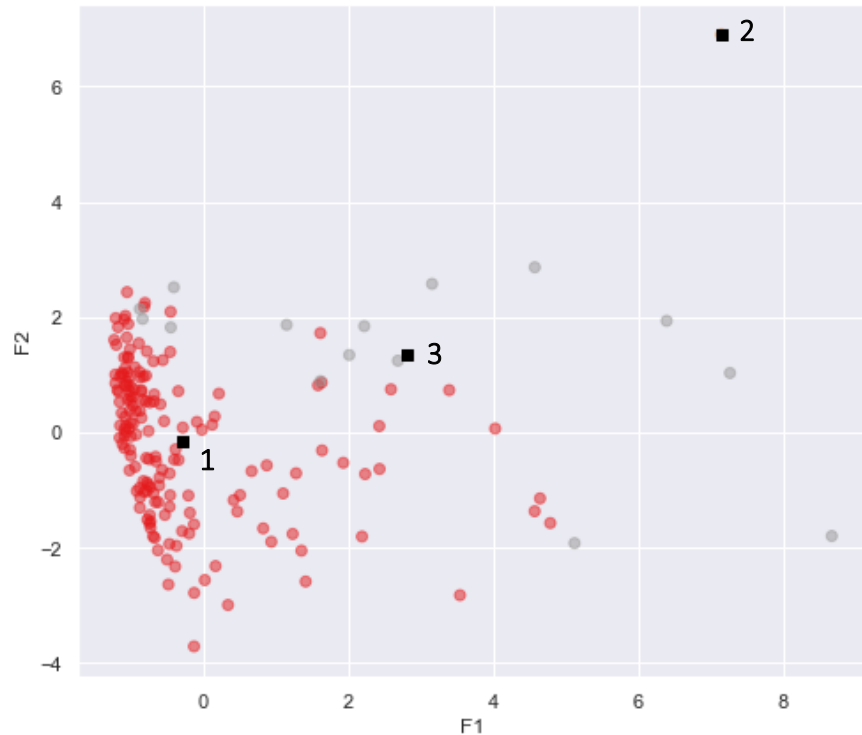


Au vue de la variation de l'inertie et de la répartition des pays sur le dendrogramme, on sait que notre clustering ne nous permettra pas d'obtenir des groupes homogènes.

On va donc chercher à éliminer les clusters isolés avant de refaire un analyse.

## d) Comparaison des deux méthodes de clustering

### Kmeans



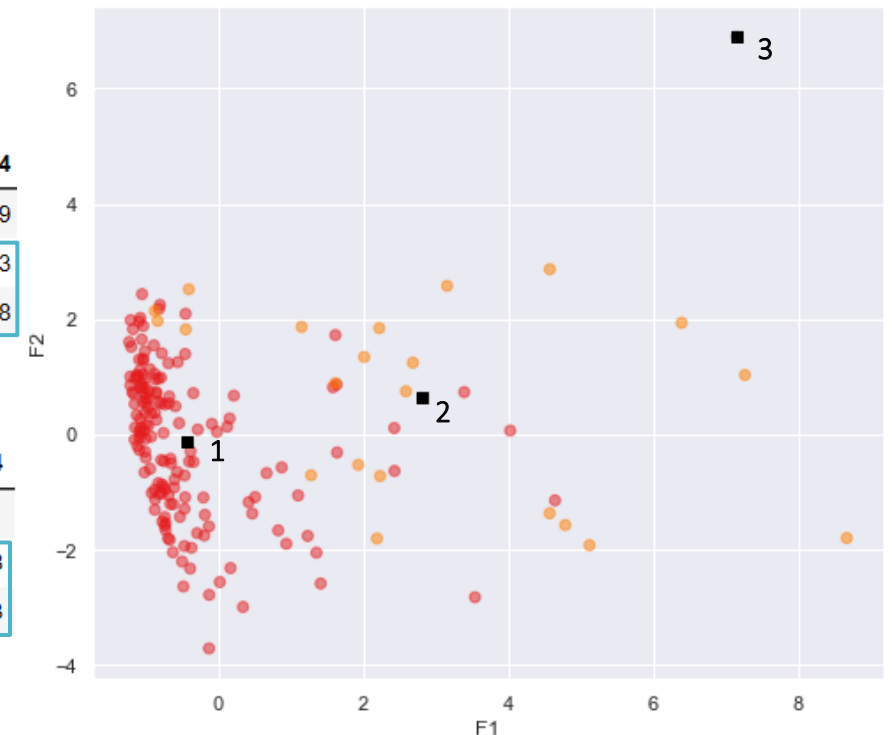
Coordonnées des centroïdes Kmeans

	F1	F2	F3	F4
Cluster 1	-0.297870	-0.163890	0.025332	0.050029
Cluster 2	7.150048	6.899961	2.029512	-7.136513
Cluster 3	2.819759	1.353722	-0.415641	-0.077888

Coordonnées des centroïdes CAH

	F1	F2	F3	F4
Cluster 1	-0.434148	-0.133593	0.040301	0.047701
Cluster 2	2.812700	0.651880	-0.383520	-0.020358
Cluster 3	7.150048	6.899961	2.029512	-7.136513

### Classification hiérarchique



La différence principale entre les deux méthodes réside dans la répartition des pays sur l'axe F2

F2 ➔ Forte corrélation négative avec la stabilité politique et le PIB, deux variables pour lesquelles nous voulons garder les valeurs les plus élevées

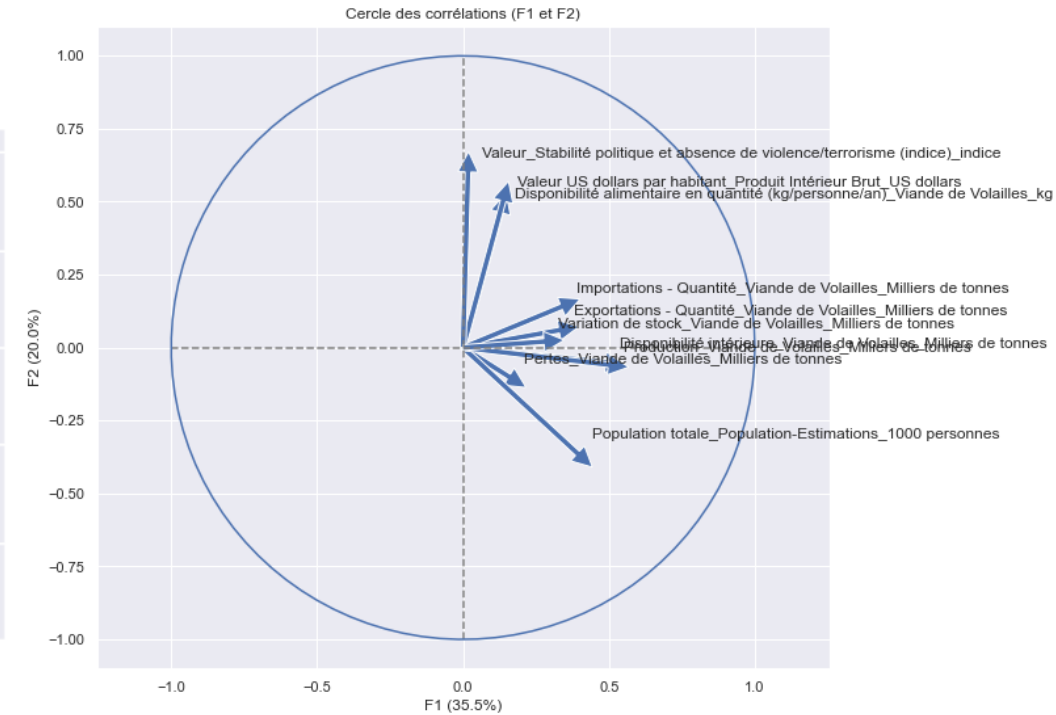
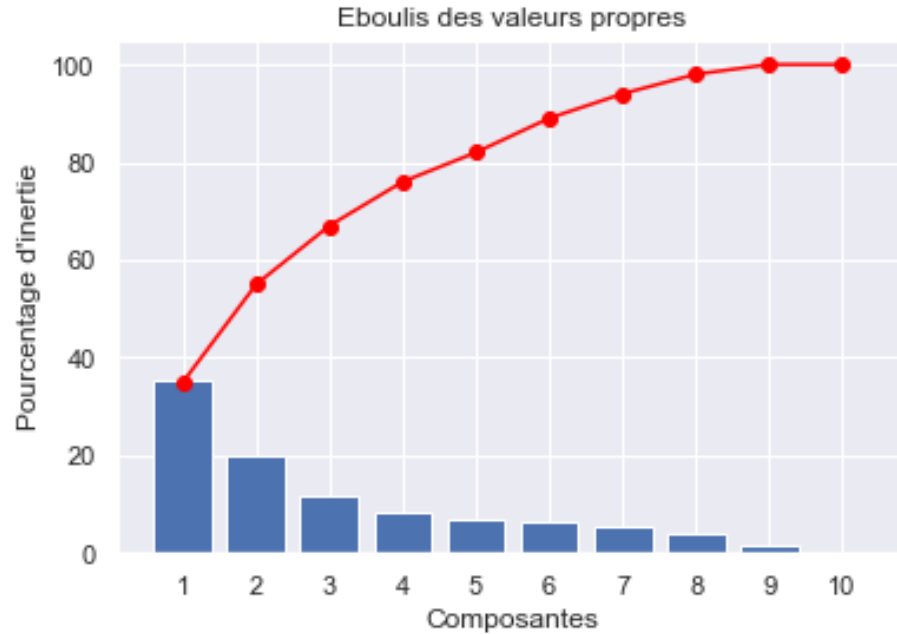
➔ On conserve les pays qui sont placés plutôt bas sur cet axe

➔ DONC on élimine le cluster des Kmeans dont le centroïde est le plus haut sur l'axe de F2

# 3bis. Clustering

## a) ACP

Résultats de l'ACP après suppression des deux clusters mis en évidence dans la première analyse

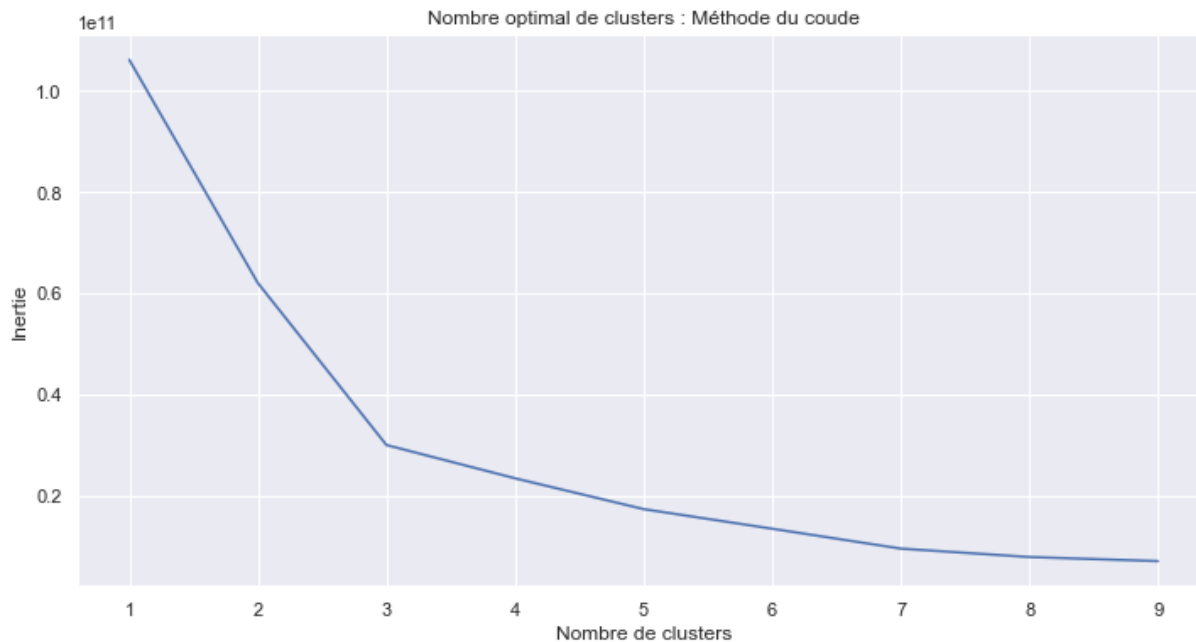


	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Disponibilité alimentaire en quantité (kg/personne/an)_Viande de Volailles_kg	0.13	0.46	0.38	-0.40	0.27	-0.33	0.16	0.39	0.33	-0.00
Disponibilité intérieure_Viande de Volailles_Milliers de tonnes	0.49	-0.05	0.18	-0.17	0.20	-0.01	-0.18	-0.17	-0.42	0.65
Exportations - Quantité_Viande de Volailles_Milliers de tonnes	0.33	0.06	-0.24	0.61	-0.20	-0.45	0.30	0.23	0.10	0.24
Importations - Quantité_Viande de Volailles_Milliers de tonnes	0.34	0.14	-0.31	-0.02	0.43	0.44	0.59	-0.16	-0.01	-0.17
Pertes_Viande de Volailles_Milliers de tonnes	0.16	-0.10	0.72	0.24	-0.36	0.43	0.27	0.02	0.11	-0.00
Population totale_Population-Estimations_1000 personnes	0.39	-0.36	-0.08	-0.00	0.14	0.06	-0.35	-0.15	0.74	-0.00
Production_Viande de Volailles_Milliers de tonnes	0.50	-0.06	0.14	0.04	-0.02	-0.27	-0.20	-0.03	-0.35	-0.70
Valeur US dollars par habitant_Produit Intérieur Brut_US dollars	0.14	0.50	-0.15	0.25	-0.01	0.47	-0.50	0.41	-0.03	0.00
Valeur Stabilité politique et absence de violence/terrorisme (indice)_indice	0.02	0.60	0.04	0.09	-0.21	-0.12	-0.08	-0.73	0.16	-0.00
Variation de stock_Viande de Volailles_Milliers de tonnes	0.28	0.02	-0.33	-0.55	-0.69	0.10	0.12	0.11	0.03	0.04

## b) Kmeans

Méthode du coude pour déterminer le nombre de clusters

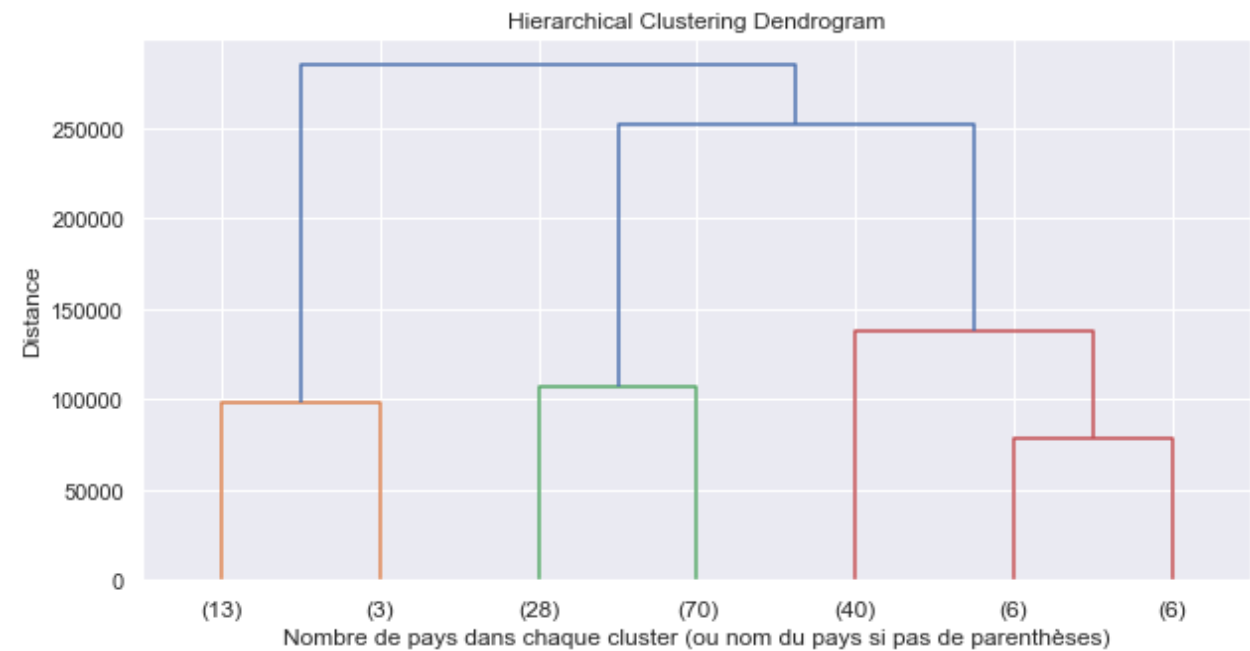
➔ 7 clusters



## c) Classification hiérarchique

On va ici aussi chercher 7 clusters

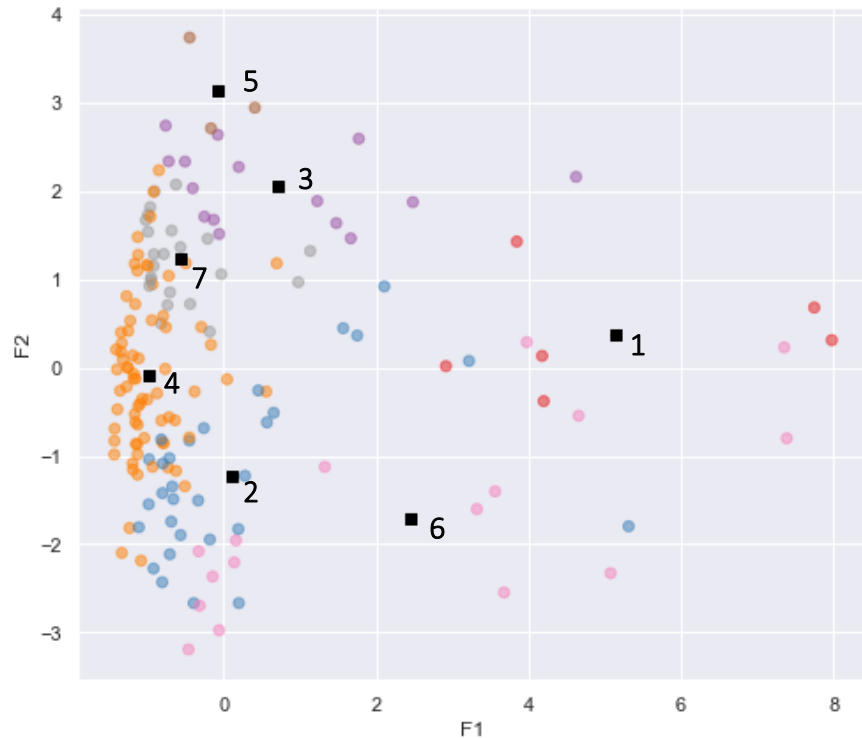
Visualisation sous forme de dendrogramme



Au vue de la variation de l'inertie et de la répartition des pays sur le dendrogramme, on voit que nos groupes de pays sont relativement homogènes, il y a donc des chances que les résultats soient exploitables.

## d) Comparaison des deux méthodes de clustering

### Kmeans



### Coordonnées des centroïdes Kmeans

	F1	F2	F3	F4
Cluster 1	5.143858	0.362585	-0.803991	-0.401404
Cluster 2	0.118252	-1.226786	0.118278	-0.032970
Cluster 3	0.705016	2.053740	-0.474420	0.699492
Cluster 4	-0.972011	-0.089912	0.077186	-0.073549
Cluster 5	-0.065275	3.125954	-0.780083	1.227242
Cluster 6	2.459856	-1.707164	0.010461	0.038359
Cluster 7	-0.573522	1.231404	0.214356	-0.261792

### Coordonnées des centroïdes CAH

	F1	F2	F3	F4
Cluster 1	0.774072	2.045487	-0.497640	0.796363
Cluster 2	-0.065275	3.125954	-0.780083	1.227242
Cluster 3	-0.530615	1.253972	0.214287	-0.214317
Cluster 4	-0.982547	-0.130539	0.055968	-0.074997
Cluster 5	0.580021	-1.311673	0.054857	0.039060
Cluster 6	3.284068	-1.941882	0.253570	-0.322959
Cluster 7	5.143858	0.362585	-0.803991	-0.401404

### Classification hiérarchique



F1 ➔ Principalement corrélée à la production et la disponibilité en volaille.

F2 ➔ Principalement corrélée à la population, au PIB et à la stabilité politique.

**Ce que l'on cible :** Les pays avec des valeurs basses sur l'axe F1 et des valeurs élevées sur l'axe F2 ➔ 3 clusters d'intérêt (encadrés en bleu)

Centroïdes Kmeans : Plus à gauche sur l'axe de la deuxième composante

➔ DONC, on retient les trois clusters obtenus par la méthode des Kmeans.

# Conclusion

Voici une liste de 40 pays qui constituent de potentielles cibles commerciales pour l'exportation de nos poulets du fait de leur taux de production faible, de leur PIB élevé, de leur population élevée, et de leur bonne stabilité politique :

Antigua-et-Barbuda  
Australie  
Autriche  
Bahamas  
Bahreïn  
Barbade  
Belgique  
Chine - RAS de Hong-Kong  
Chine - RAS de Macao  
Chypre  
Croatie  
Danemark  
Estonie  
Finlande

Grèce  
Hongrie  
Irlande  
Islande  
Israël  
Koweït  
Lettonie  
Lituanie  
Luxembourg  
Malte  
Nauru  
Norvège  
Nouvelle-Calédonie  
Nouvelle-Zélande

Pays-Bas  
Polynésie française  
Portugal  
Qatar  
Suède  
Suisse  
Saint-Kitts-et-Nevis  
Slovaquie  
Slovénie  
Tchéquie  
Trinité-et-Tobago  
Uruguay