



Etude de marché Développement à l'international

Analyses et segmentation
Réalisé par Aurélie RIVIERE
03/2025



Contexte & Objectif

- ✓ **La poule qui chante** : Entreprise française d'agroalimentaire
- ✓ **Elevage et vente de poulets** sous le label "*Poulet Agriculture Biologique*"
- ✓ **Marché** actuellement uniquement français
- ✓ **Volonté de développement** à l'international

Objectif

Déterminer des pays avec un potentiel de marché pour l'exportation.

1

• Analyse des données

2

• Segmentation des pays

3

• Analyse des marchés

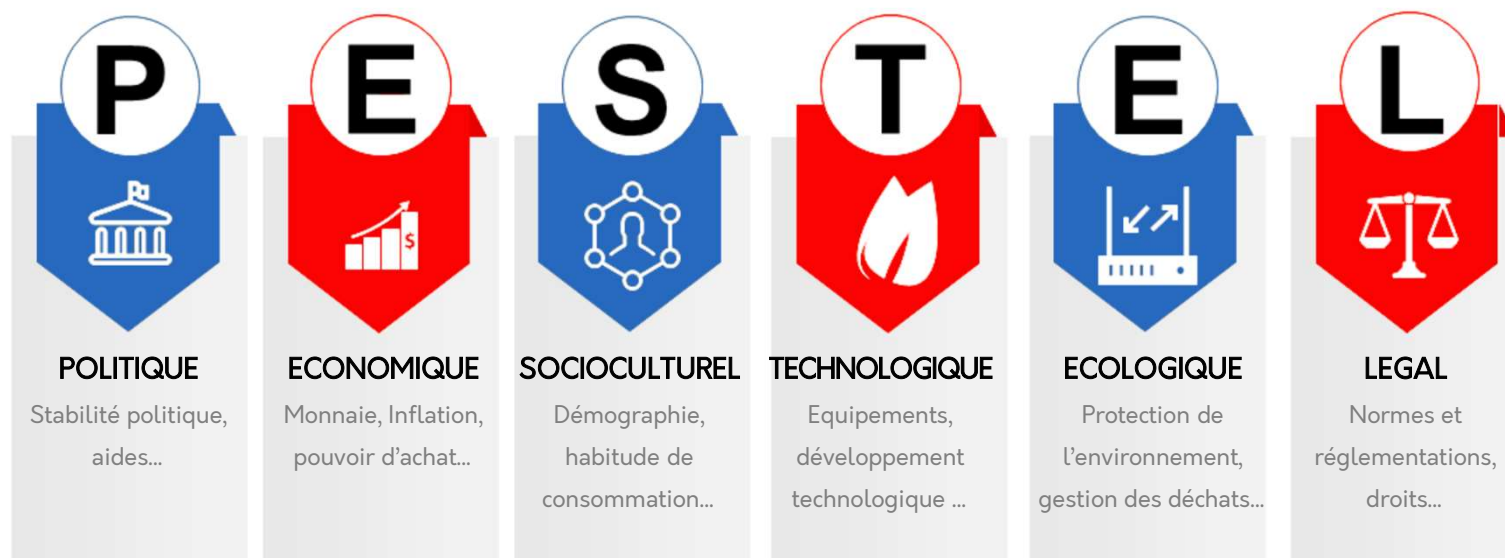
4

• Recommandations des pays à cibler



Analyse PESTEL & Sélection des données

Outil de diagnostic stratégique permettant de comprendre l'influence extérieure qui peut impacter une activité.



- Sélection des variables ayant un impact sur l'exportation de volaille permettant de regrouper les pays.
- Recherche des données : Données opendata FAO (Food and Agriculture Organization)
- Exportation des jeux de données
- Sélection de 8 variables représentatives.

Variables sélectionnées

- ✓ **Disponibilité alimentaire de la volaille (en kg/ par personne / an)** : quantité disponible par habitant.
- ✓ **Population** : Taille de la population du pays et indice de la taille du marché
- ✓ **Stabilité Politique du pays** : indice de stabilité du marché
- ✓ **Pib par habitant** : indice de pouvoir d'achat

Création de nouvelles variables via feature engineering

- ✓ **Croissance démographique** : Indicateur de la croissance moyenne annuelle du pays. Indicateur de l'évolution à venir du marché.
- ✓ **Taux de stockage** : Mesure la propension d'un pays à stocker ou consommer rapidement sa production et ses importations. $Taux\ de\ stockage = Variation\ de\ stock / Disponibilité\ Intérieure$
- ✓ **Taux d'autosuffisance (%)** : Indique si un pays produit suffisamment pour couvrir sa consommation. $Taux\ d'autosuffisance = Production / Disponibilité\ intérieure$
- ✓ **Taux de dépendance aux importations (%)** : Mesure la dépendance d'un pays aux importations pour sa consommation de volaille. $Taux\ de\ dépendance\ aux\ importations = Importations / Disponibilité\ intérieure$



Spécifications & Nettoyage des données



4 jeux de données ont été collectés (FAO, Opendata)

- ✓ DispoAlimentairesVolailles
- ✓ Pib par Habitant
- ✓ Political Stability
- ✓ Population



Traitement des données dans un notebook Jupyter en Python

Nettoyage et préparation des jeux de données :

- ✓ Analyse exploratoire des fichiers
- ✓ Nettoyages des données inutiles ou redondantes
- ✓ Mise à l'échelle des unités
- ✓ Gestion des valeurs manquantes (complétée ou supprimée en fonction de l'intérêt)
- ✓ Analyse des outliers

Fusion des jeux de données

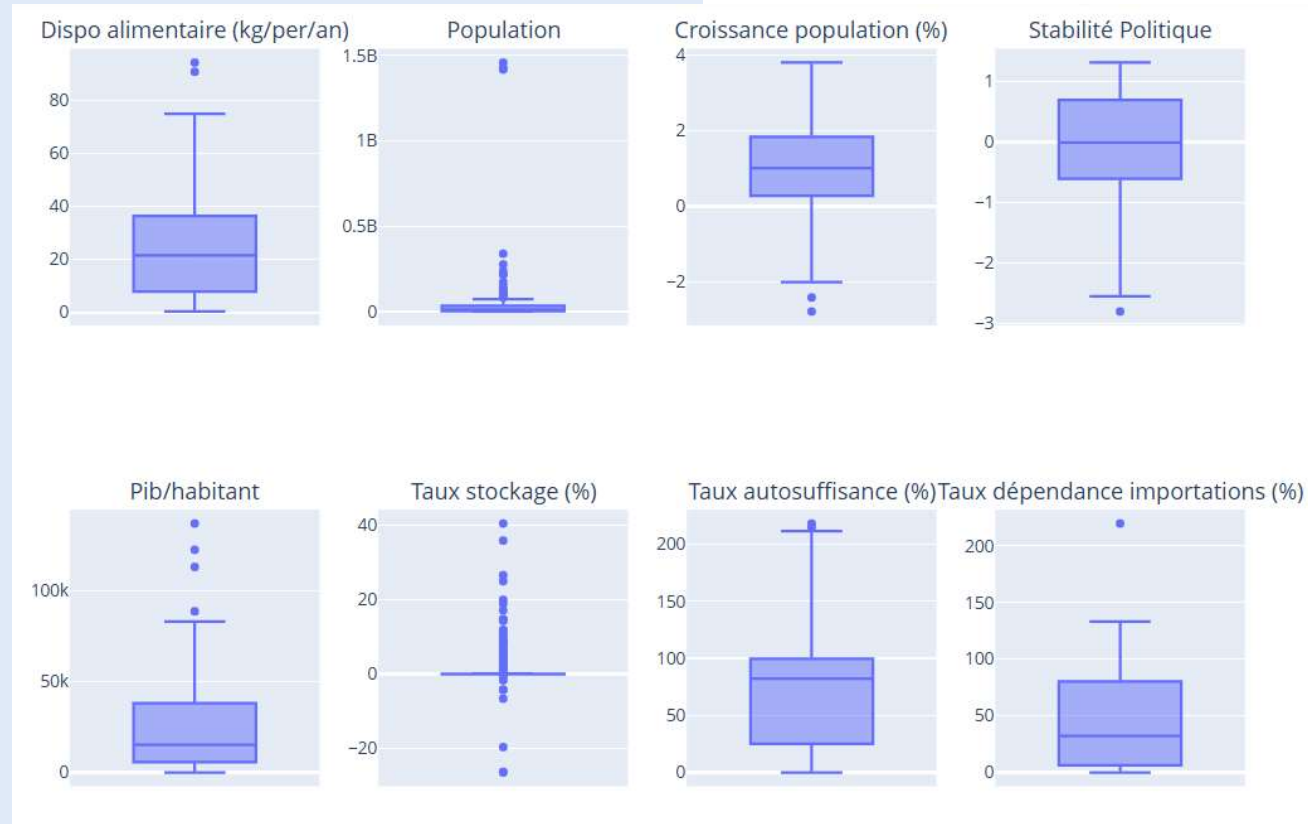


✓ **Création des
4 variables calculés**

✓ Création d'un fichier
de données avec :
**8 variables sélectionnées
et 183 pays.**

Analyse des variables et dispersion des valeurs

- ✓ Analyse et vérification du jeu de données complet
- ✓ Création de graphique en boîte à moustache pour analyser la distribution des variables
- ✓ Répartition équilibrée
- ✓ Outliers cohérents



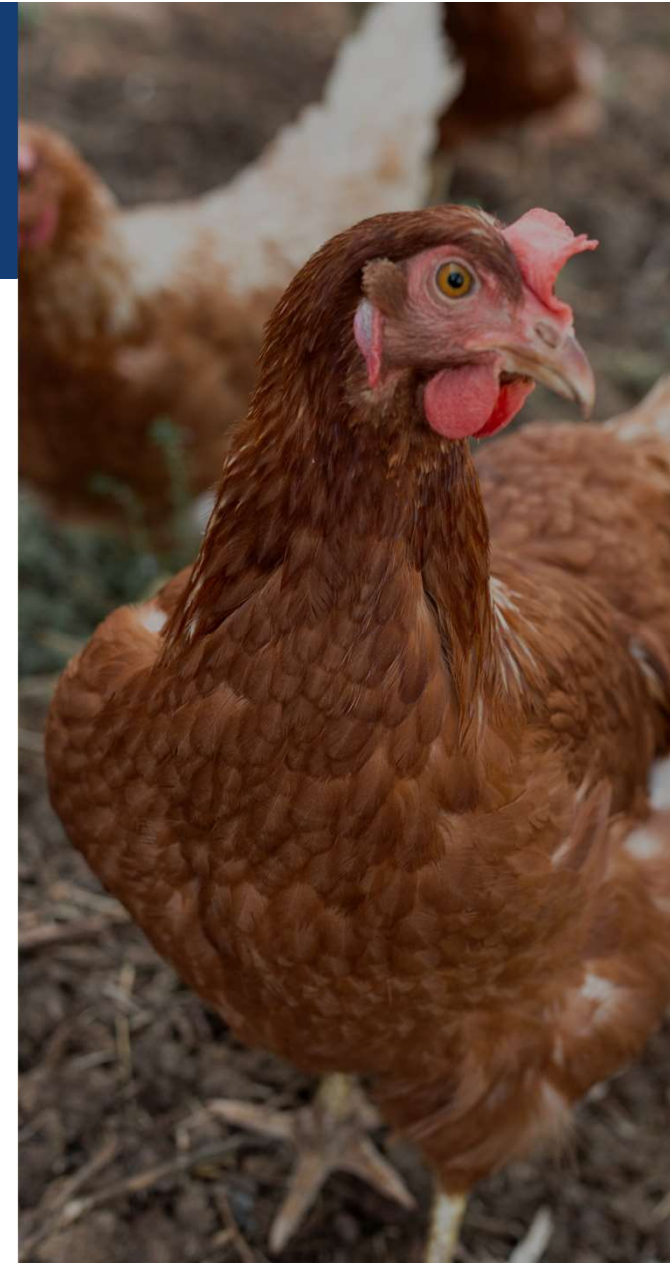
Analyse en composantes principales

Réduction des variables

Technique de réduction de variables, permet de **regrouper les variables corrélées** en un nombre restreint de **composantes principales**.

Les composantes principales sont des nouvelles variables qui sont des combinaisons linéaires des variables initiales.

- ✓ Evite la redondance lorsqu'on a beaucoup de variables
- ✓ Réduit la complexité de lecture en conservant l'essentiel de l'information
- ✓ Facilite la visualisation des individus et des relations sur un plan en 2d ou 3d
- ✓ Améliore les performances des algorithmes de clustering



Mise en place de l'ACP

Préparation des variables

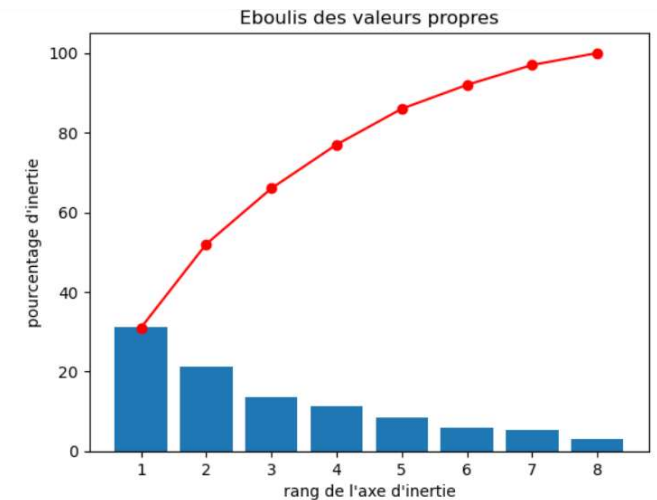
- ✓ Sélection des variables numériques
- ✓ Standardisation et centrage des données (pour mise à l'échelle).

Entraînement de l'ACP

- ✓ Entraînement de l'ACP avec 8 composantes (autant que de variables)
- ✓ Calcul de la matrice de covariance : mesure les relations entre les variables.
- ✓ Décomposition en valeurs propres : On extrait les vecteurs propres et valeurs propres pour déterminer les directions principales.

Sélection des composantes

- ✓ Eboulis des valeurs
- ✓ Sélection des composantes principales : Sélection des composantes qui expliquent le plus de variance.
- ✓ Réduction à 4 composantes



Eboulis des valeurs

- La première composante explique 31% de la variance. Elle est la plus importante et capte une grande part d'information.
- Avec 4 composantes, on capture 77% de la variance.

Ce nombre de composantes semble un bon compromis afin de conserver le maximum d'informations tout en limitant la complexité de l'analyse.

Interprétations des composantes

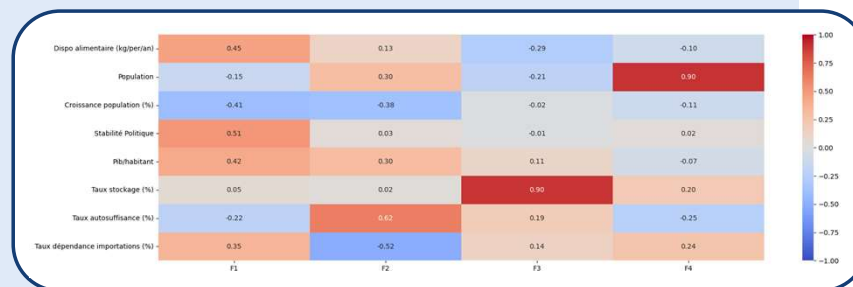
Cercles des corrélations & Heatmap

Cercle des corrélations :

- ✓ Permet de visualiser la relation entre les variables initiales et les nouvelles composantes.
- ✓ Permet de donner un sens à chaque composante en fonction de l'impact des variables initiales.

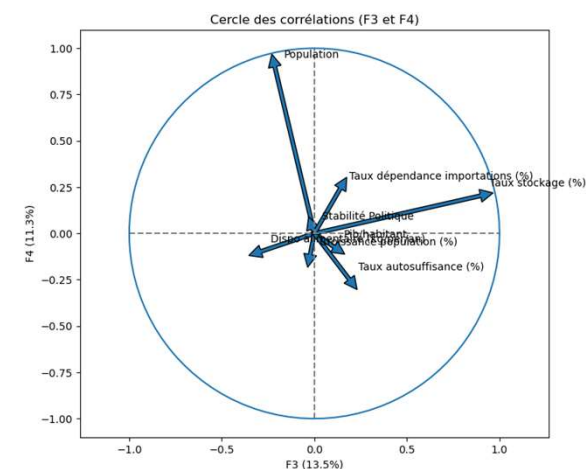
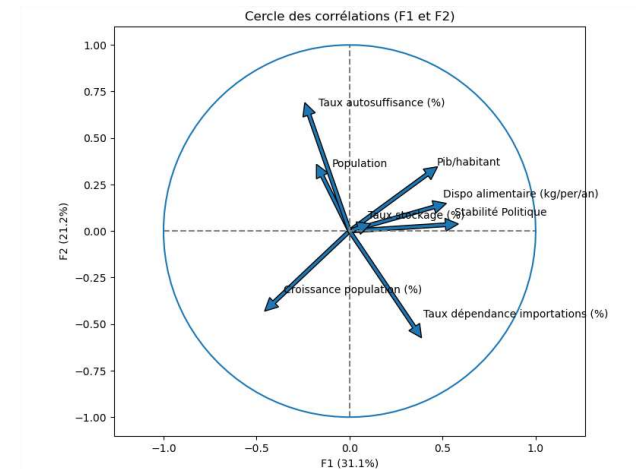
Heatmap des corrélations

- ✓ Permet de visualiser les implications de chaque variable dans les nouvelles composantes.



Interprétations des composantes

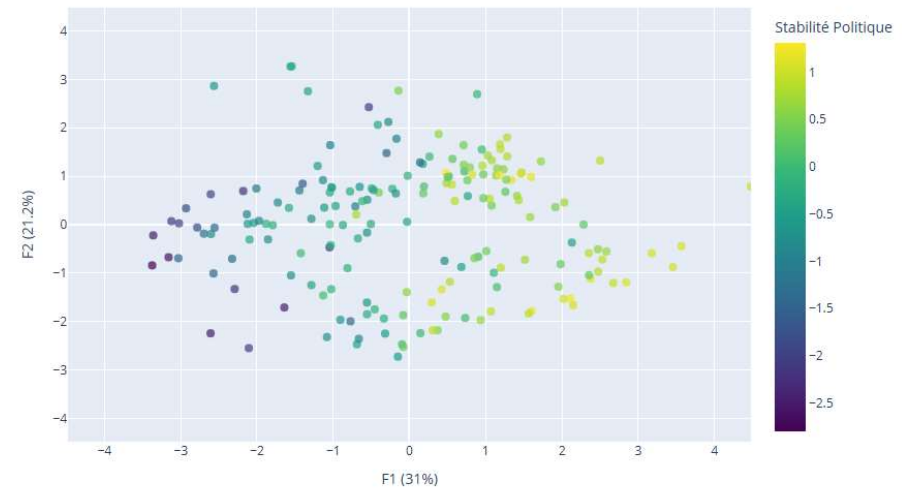
- ✓ **F1** : corrélation forte avec la stabilité politique, le pib/habitant et la disponibilité alimentaire → **Richesse du pays**
- ✓ **F2** : corrélation positive forte au taux d'autosuffisance et corrélation négative forte aux taux de dépendance aux importations → **Autosuffisance**
- ✓ **F3** : forte corrélation au taux de stockage alimentaire → **Capacité de stockage**
- ✓ **F4** : Forte corrélation avec la population, une dimension démographique. → **Population**



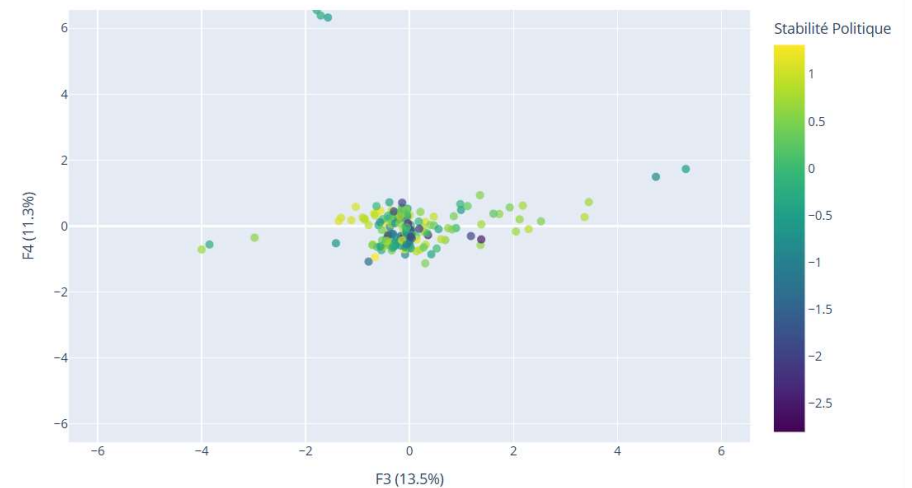
Projection des individus dans l'espace des nouvelles composantes

- ✓ permet de visualiser la répartition des pays dans l'espace des axes de nos nouvelles composantes.
- ✓ sert à **identifier des structures** dans les données et à **détecter des regroupements** entre individus.
- ✓ Vérification de la répartition des individus avec segmentation.
- ✓ Première observation des tendances.

Projection des individus (F1 et F2)



Projection des individus (F3 et F4)



CAH Classification Ascendante Hiérarchique

Méthode de clustering qui permet de regrouper des observations en fonction de leur similarité.

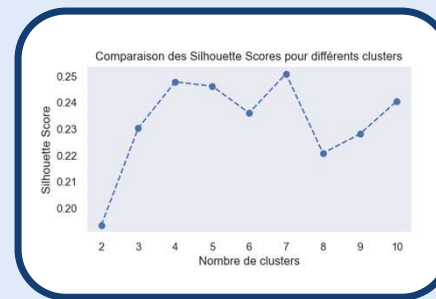
CAH

- ✓ Ne nécessite pas de définir un nombre de clusters à l'avance
- ✓ **Dendrogramme**, qui permet de visualiser les regroupements.

Test de score silhouette

- ✓ Permet de mesurer la qualité du clustering en évaluant la cohésion interne des clusters et leur séparation

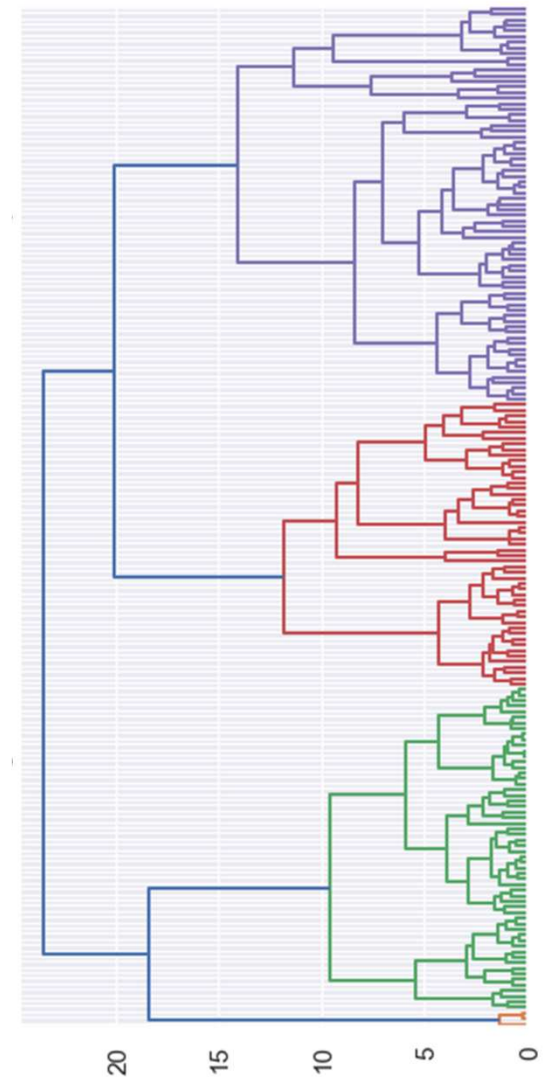
Sélection de 7 clusters



Méthode Scikit-Learn

- ✓ Nombre de cluster fixé à l'avance
- ✓ Segmentation avec des données numériques structurées

Attribution de chacun des pays à un cluster



K-means Algorithme de clustering

Regroupe des observations en groupes homogènes

K-Means

- ✓ Algorithme de clustering
- ✓ Permet de regrouper des observations en **groupes** homogènes en minimisant la variance intra-cluster.

Méthode du coude

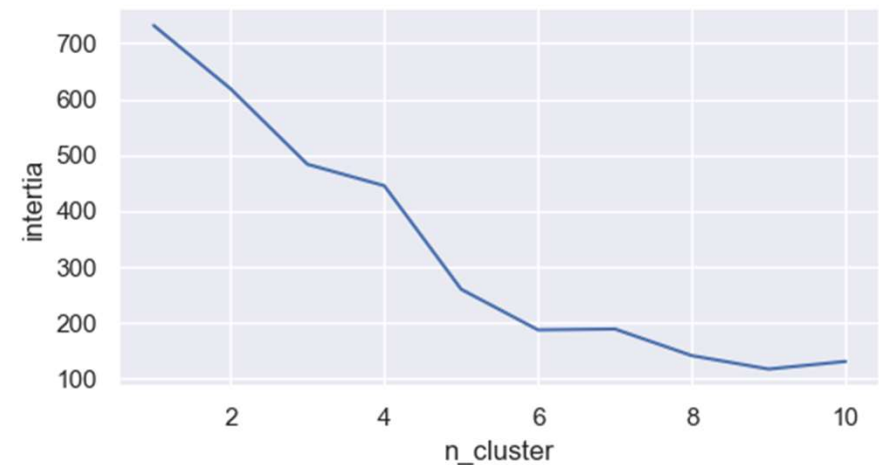
- ✓ L'observation de la courbe et de son coude, c'est à dire le moment où la tendance se réduit permet de déterminer un nombre pertinent de cluster.

Ici la courbe montre une diminution rapide de l'inertie jusqu'à environ $K=6$,

Silhouette score

- ✓ Permet de mesurer la qualité du clustering en évaluant la cohésion interne des clusters et leur séparation.

Les Silhouette Score montrent les meilleurs résultats autour de $K = 6, 7$ ou 8 . Au delà on observe une diminution.



Au vu de l'observation de la méthode du coude et de l'observation des scores de silhouette

$k=6$ semble offrir un bon compromis entre l'inertie et la séparation des clusters, tout en permettant une meilleure interprétation des clusters.

CAH ou Kmeans

Quelle méthode de clustering utiliser ?

CAH : 7 clusters / Kmeans : 6 clusters

- **Observation** : Segmentation et répartition des pays quasi similaire dans un cas comme dans l'autre

Kmeans

- ✓ Choix plus pertinent des clusters grâce à la méthode du coude
- ✓ Meilleur silhouette score
- ✓ Plus flexible

Sélection du modèle kmeans à 6 clusters



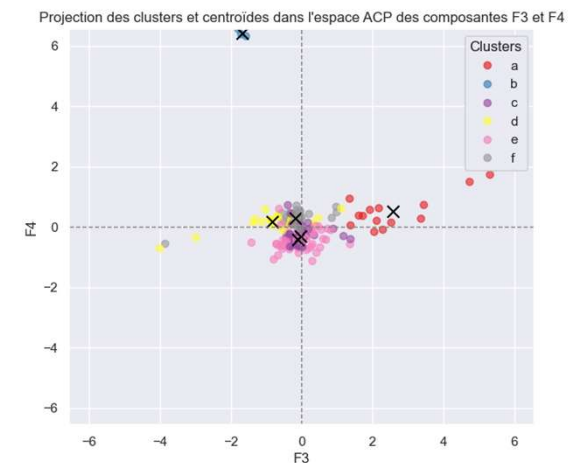
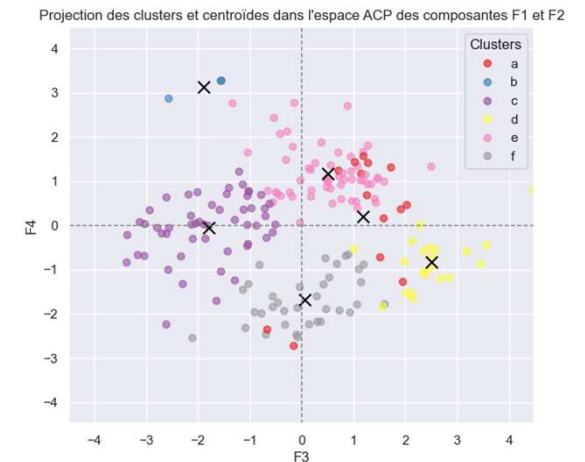
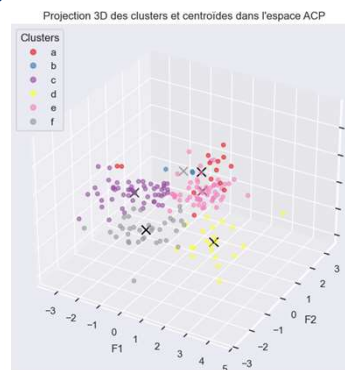
Visualisation des centroïdes

- ✓ Projection des pays dans l'espace des composantes produites par l'ACP
- ✓ Ajout des centres des clusters
- ✓ Colorisation de chaque pays en fonction de son cluster

Projection F1-F2 : bonne séparation des clusters. Ces axes capturent une grande partie de la variance.

Projection F3-F4 : concentration des points plus forte autour du centre, F3 et F4 capturent moins de variance.

La visualisation en 3D confirme que les centroïdes sont bien distinct

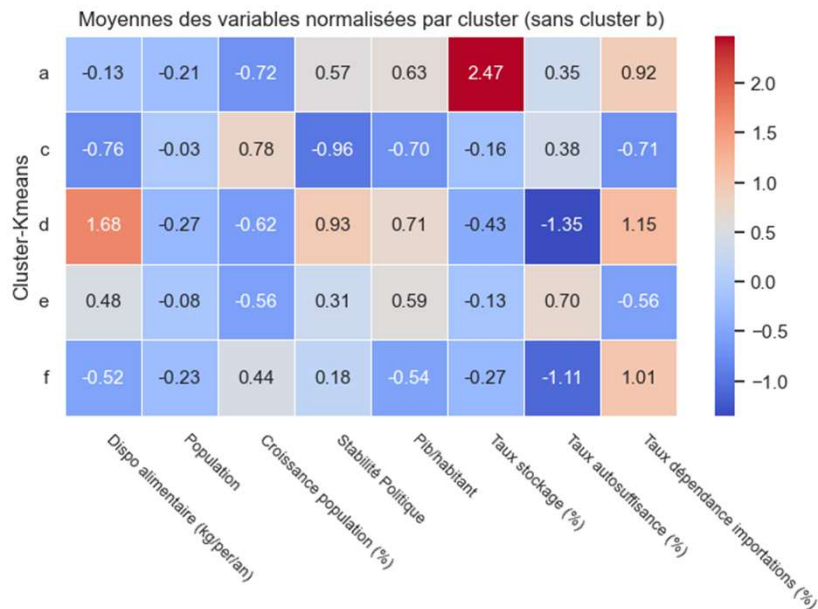


Segmentation

Observations des profils de nos clusters

Afin d'interpréter nos clusters, on observe :

- ✓ La liste des pays par clusters
- ✓ Les projections des clusters sur l'espace des nouvelles composantes
- ✓ La heatmap des moyennes des variables pour chaque cluster.



Cluster A → Pays développés : pays économiquement stables ou émergents, faible croissance démographique, forte capacité de stockage alimentaire. Pays relativement dépendants des importations.

Cluster B → Pays à forte population : Pays autosuffisants caractérisés par une très forte population.

Cluster C → Pays en développement instable : pays souvent instables politiquement, forte croissance démographique, PIB faible, autosuffisance alimentaire modérée.

Cluster D → Petits États riches importateurs : pays riches à faible population., dépendent fortement des importations..







Cluster E → Puissances agricoles et industrielles : pays équilibrés avec un PIB moyen, autosuffisance alimentaire, croissance démographique faible.

Cluster F → Petits pays en développement importateurs : PIB faible mais une forte dépendance aux importations.

Analyse des clusters & Pertinence de marché

Analyse des clusters les plus adaptés à notre marché, en tenant compte des critères stratégiques :

- ✓ Demande en volaille
- ✓ Capacité à payer un produit (premium dans notre cas avec le label biologique)
- ✓ Autosuffisance faible ou ouverture à l'importation pour un produit de niche (premium)
- ✓ Accès au marché

Cluster	Analyse	Pertinence	Exemple de pays
A : Pays développés	Pays dépendants des importations avec un marché stable et des pays avec un bon pouvoir d'achat. Principalement pays européens, marché établi.	 Pertinent	Allemagne, Bulgarie, Estonie, Finlande...
B : Pays à forte population	Economiquement instable et marchés lointains.	 Non pertinent	Chine, Inde...
C : Pays en développement instables	Faible pouvoir d'achat, risques politiques	 Non pertinent	Afghanistan, Algérie, Azerbaïdjan, Bangladesh, Burkina Faso...
D : Petits États riches importateurs	Fort pouvoir d'achat, forte dépendance aux importations, demande pour le premium	 Pertinent	Luxembourg, Qatar, Émirats arabes unis...
E : Puissances agricoles et industrielles	Pays producteur et autosuffisant mais avec des marchés importants. Potentiellement concurrent mais peut être ouvert pour du marché premium.	 Potentiellement intéressant	Afrique du Sud, Arabie saoudite, Argentine, Australie, Autriche, Japon, République de Corée...
F : Petits pays en développement importateurs	Marché trop faible.	 Non pertinent	Angola, Arménie, Bhoutan, Botswana, Congo, Djibouti, Gabon, Gambie...

Recommandations

Calcul d'un score d'opportunité

Afin d'identifier les pays les plus pertinents pour l'exportation de poulet bio, mise en place **un score d'opportunité** basé sur des critères stratégiques.

Ce score est une moyenne pondérée de plusieurs facteurs influençant le potentiel d'exportation :

- ✓ **PIB par habitant (15%)** → Un pouvoir d'achat élevé favorise la consommation de produits bio.
- ✓ **Stabilité politique (15%)** → Un environnement stable est essentiel pour sécuriser les échanges.
- ✓ **Taux d'autosuffisance alimentaire (-15%)** → Un pays trop autosuffisant aura moins besoin d'importer.
- ✓ **Taux de dépendance aux importations (15%)** → Plus un pays importe, plus il est habitué aux échanges internationaux.
- ✓ **Dispo alimentaire (kg/per/an) (10%)** → Pays où la volaille est déjà bien consommée
- ✓ **Croissance démographique (5%)** → Un marché en expansion peut offrir de nouvelles opportunités.
- ✓ **Population totale (5%)** → Un marché plus grand peut signifier plus de clients potentiels.

Cluster A : Pays-Bas, Oman, Seychelles, Estonie, Mauritanie

Cluster D : Luxembourg, Qatar, Chine, Émirats arabes unis

Cluster E : Irlande, Suisse, Autriche, Islande, Suède



Merci pour votre attention