



La poule qui chante

# Développement à l'international

---

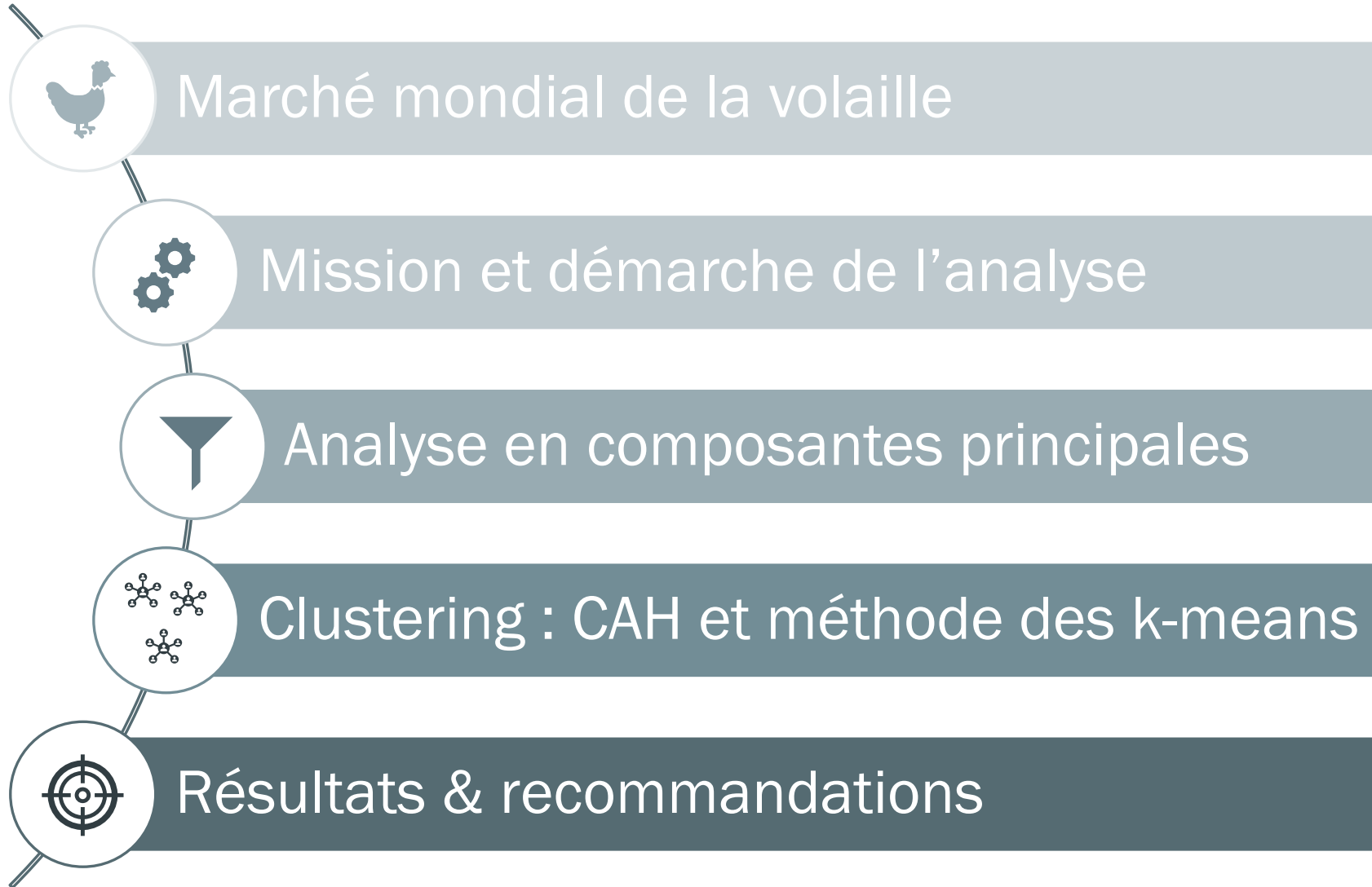
Projet P9 - Produisez une étude de  
marché avec R ou Python



PARCOURS DATA ANALYST\_V2

ADELIN LE RAY

# SOMMAIRE



# Le marché mondial de la volaille

**N°1 des viandes  
consommées  
dans le monde**

**Coût de revient** ↗

Normes et exigences des consommateurs (qualité de la viande, environnement, bien-être animal) de plus en plus strictes

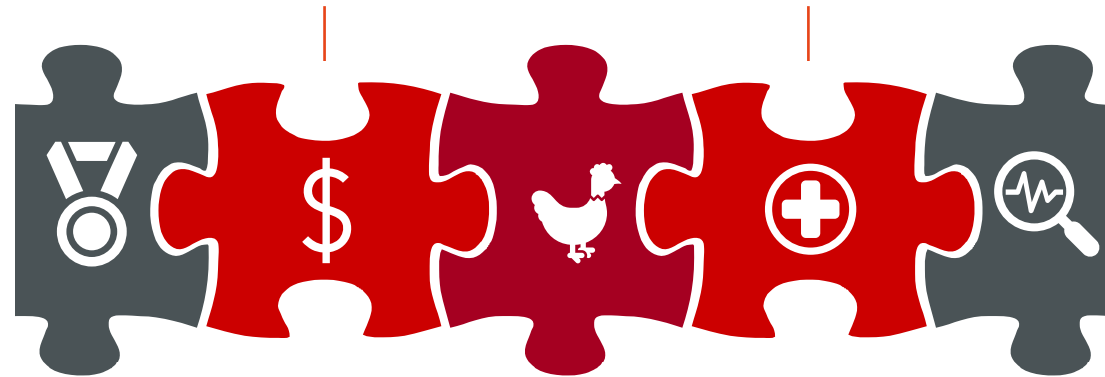
**Grippe aviaire**

Impacte tous les pays producteurs de volaille et décime les élevages

**Prévisions pour la volaille**

47% du total des protéines de viande consommées dans le monde d'ici 2031

Taux de croissance annuel de 5,45% de 2022 à 2028



**Poulet : marché en pleine  
croissance**

86% des importations de viande de volaille en valeur  
Importations mondiales de poulet en progression de 30% depuis 2018

# Mission de data analyse



Première analyse des groupements de pays cibles pour l'export de poulets.



Données du site de la FAO



Année 2017



**La poule qui chante** souhaite se développer à l'international

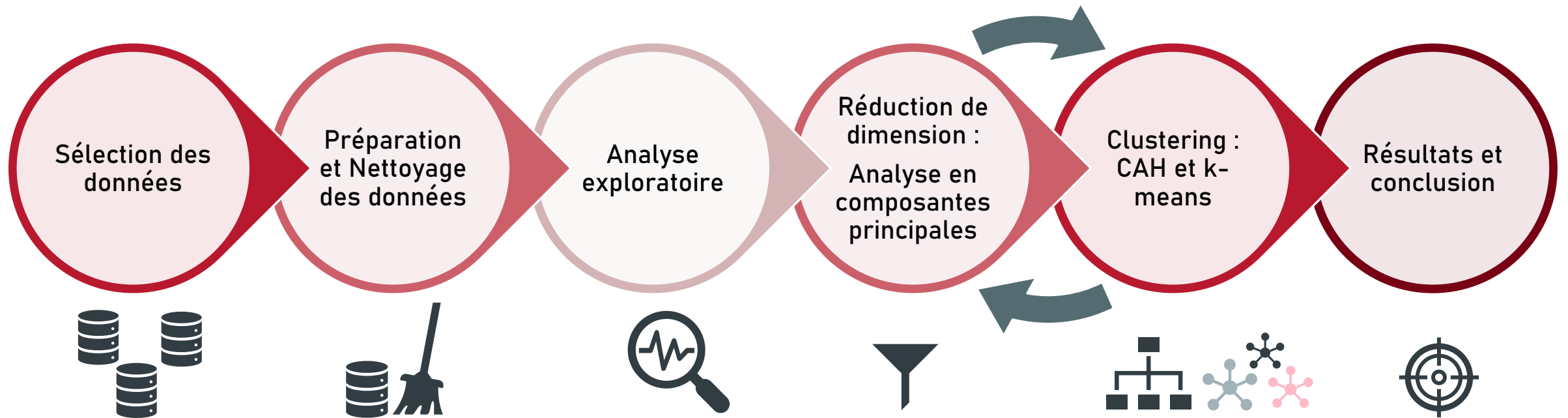


## Pays cibles :

- Dépendant des importations
- Stable politiquement
- Bon pouvoir d'achat

# Démarche de l'analyse

---

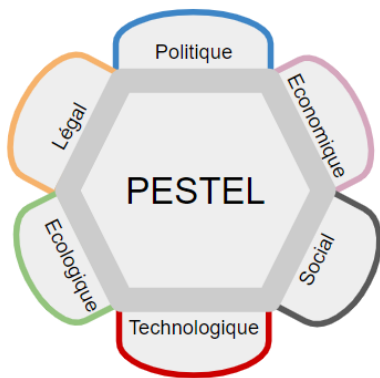








# 14 variables retenues



## Variables issues des bilans alimentaires

TAS - Taux d'autosuffisance	<i>Production / Dispo intérieure</i>
TDI - Taux de dépendance aux importations	<i>Importations / Dispo intérieure</i>
TDC - Taux de dépendance aux commerces	<i>(Production - Importations) / Dispo intérieure</i>
TCC - Taux de couverture de la consommation	<i>(Production - Exportations) / Dispo intérieure</i>
TEX - Taux d'exportation	<i>Exportations / (Production + Importations)</i>
Disponibilité alimentaire (kg/pers/an)	/
Part de la volaille dans les apports en protéine d'origine animale	<i>Dispo protéine (volaille) / Total dispo protéine animale</i>



 POLITIQUE	 ECONOMIQUE	 SOCIOLOGIQUE	 TECHNOLOGIQUE	 ENVIRONNEMENT	 LEGAL
Stabilité politique	PIB / habitant	Population		Emissions CO2 / habitant	
	Croissance PIB	Evolution de la population	X		X
		Abordabilité de la nourriture			

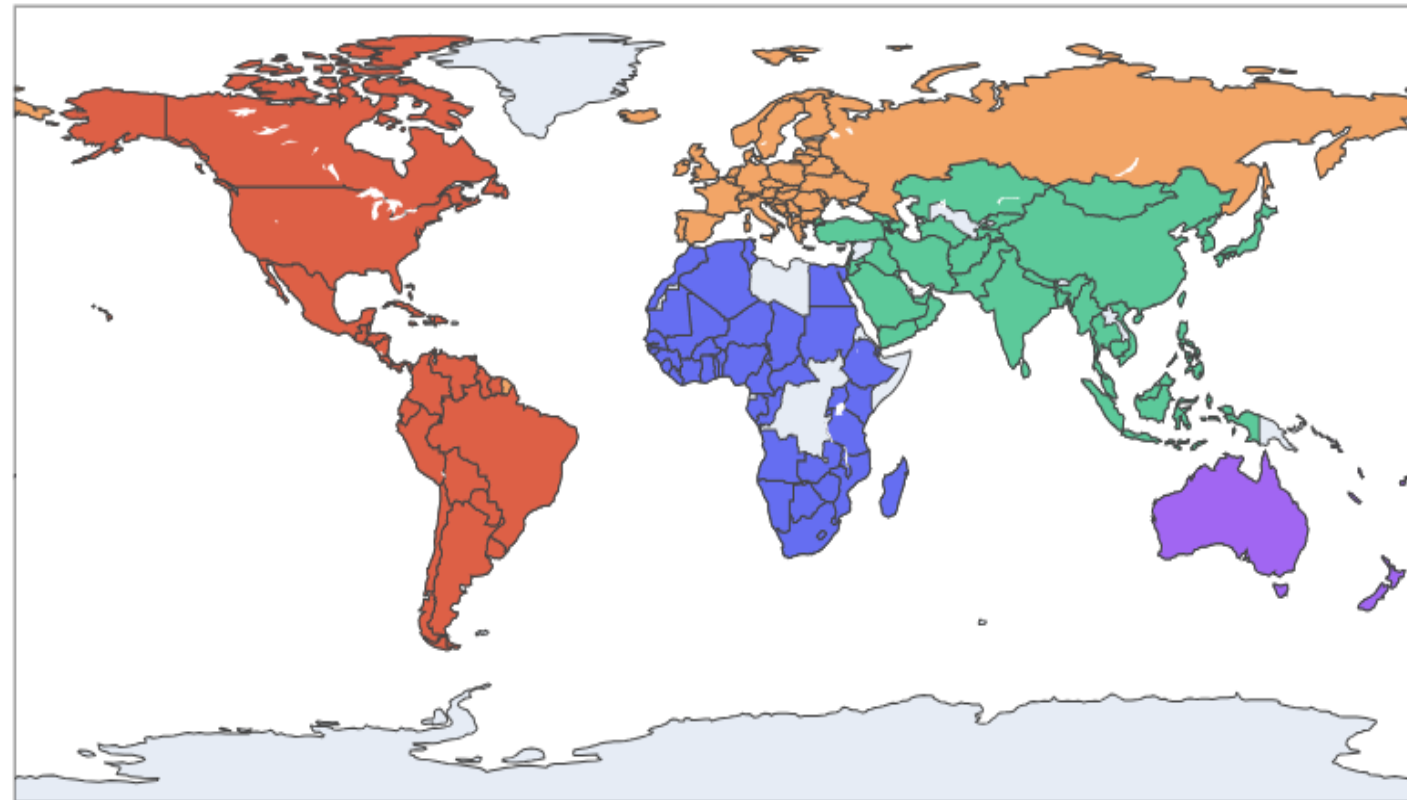
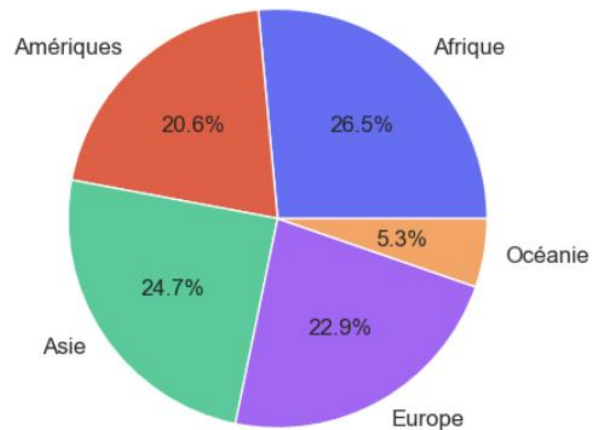
# 170 pays retenus

Nombre d'individus (pays) retenus après nettoyage :

- 170 pays sur 236 pays

*(172 pays dans les bilans alimentaires)*

Répartition des pays retenus par continent (%)

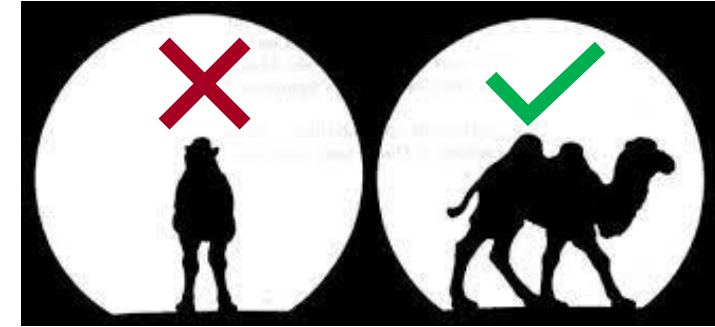


# Analyse en Composantes Principales

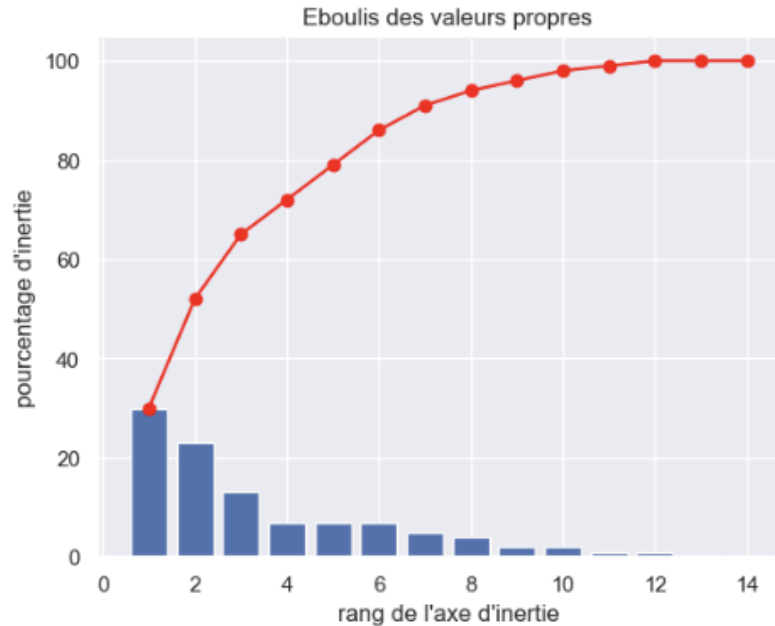
- Objectif : Réduction de dimension
- Principe : Recherche de la projection permettant de visualiser au mieux les données
- + : Visualisation des données, Améliorer l'apprentissage, Gain en stockage et temps de traitement par la suite
- - : Perte d'informations

Plus la projection du nuage de points est étalée, plus faible est la perte d'informations.

- Projection : vue 2D définie par 2 axes orthogonaux (1 axe = 1 composante principale)
- Composante principale = combinaison linéaire des variables initiales



Notre étude :



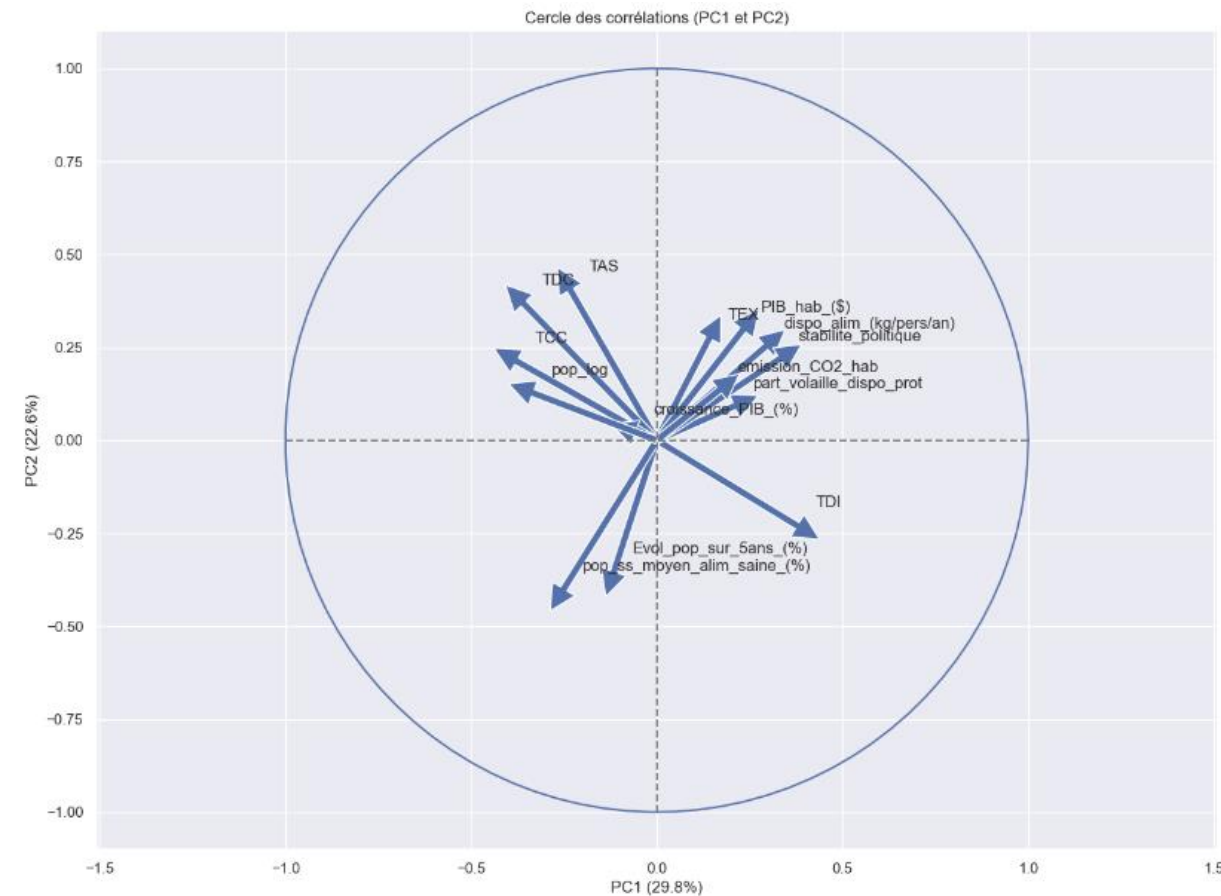
14 variables = 100% des informations



6 composantes = 86% des informations



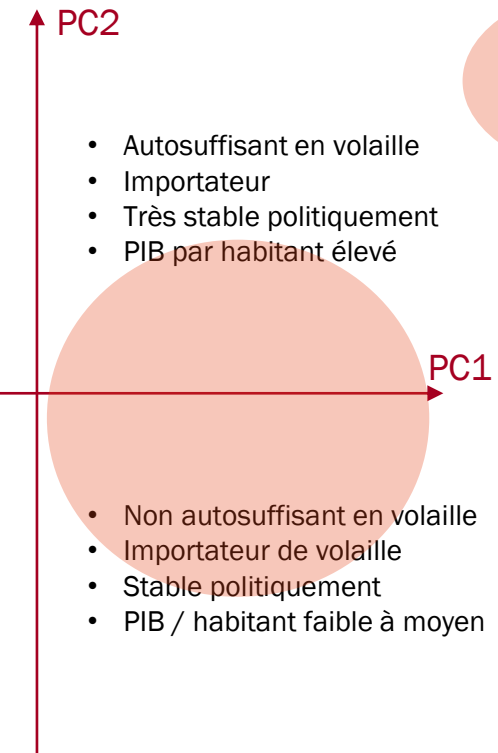
# Analyse en composantes principales



- Autosuffisant en volaille
- Non importateur
- Stabilité politique faible
- PIB par habitant faible à moyen

- Non Autosuffisant en volaille
- Non Importateur
- Instable politiquement
- PIB par habitant très faible

Autosuffisance en volaille ↗  
Abordabilité de la nourriture ↗  
PIB par habitant ↗



**Pays cibles**

- Autosuffisant en volaille
- Importateur
- Très stable politiquement
- PIB par habitant élevé

Dépendance aux importations ↗  
Stabilité politique ↗  
Dispo alim / pers ↗

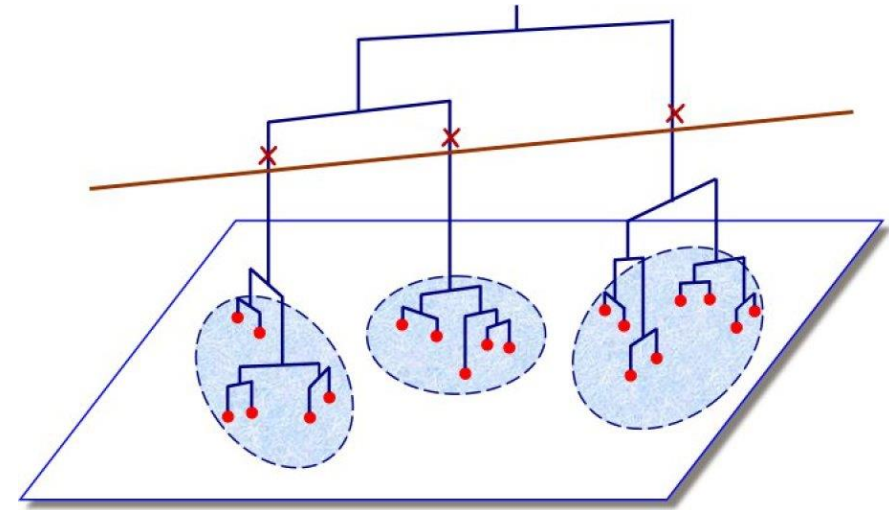
- Non autosuffisant en volaille
- Importateur de volaille
- Stable politiquement
- PIB / habitant faible à moyen

10

# Classification ascendante hiérarchique (CAH)

La première méthode utilisée pour partitionner les pays est la classification ascendante hiérarchique :

- Initialisation : 1 individu = 1 cluster
- Itérations : Regroupement avec les individus les + proches en distances (les plus similaires)



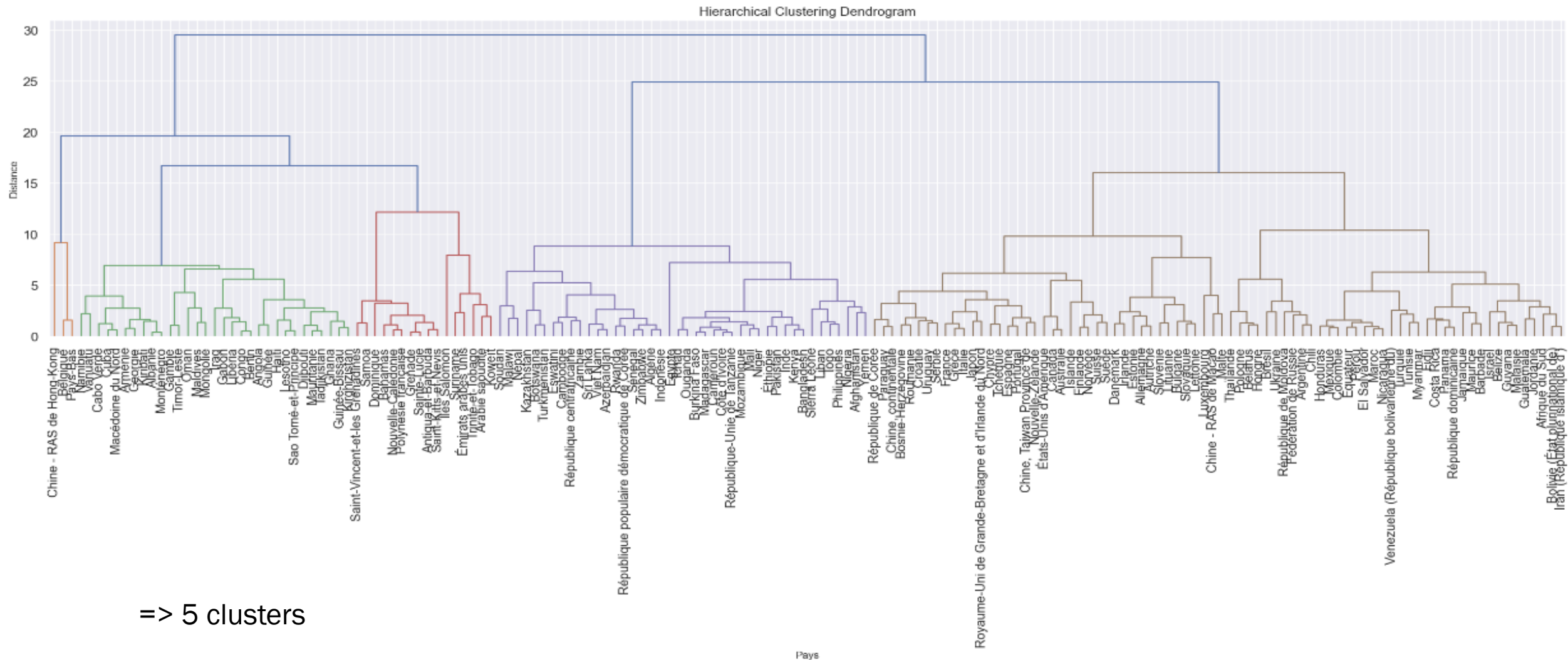
**Un dendrogramme** représente la hiérarchie sous la forme d'un arbre :

- Individus = feuilles
- Branches = liaisons entre les individus

**La distance entre les individus** est représentée par la longueur de la branche : plus la branche est courte, plus les individus sont similaires et inversement.

**Les classes**, ou *clusters*, sont obtenues en coupant l'arbre à une distance choisie (*souvent à l'endroit où les branches s'allongent*).

# Classification hiérarchique ascendante : Dendrogramme

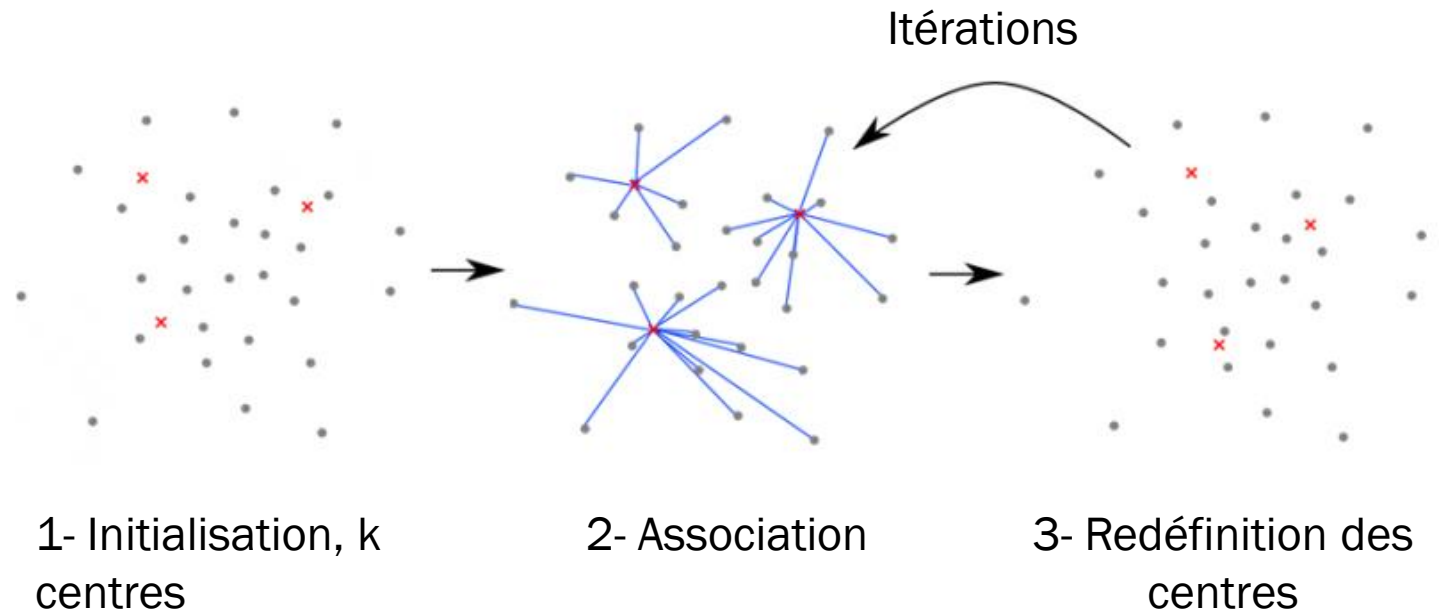


=> 5 clusters

# Méthode des k-means

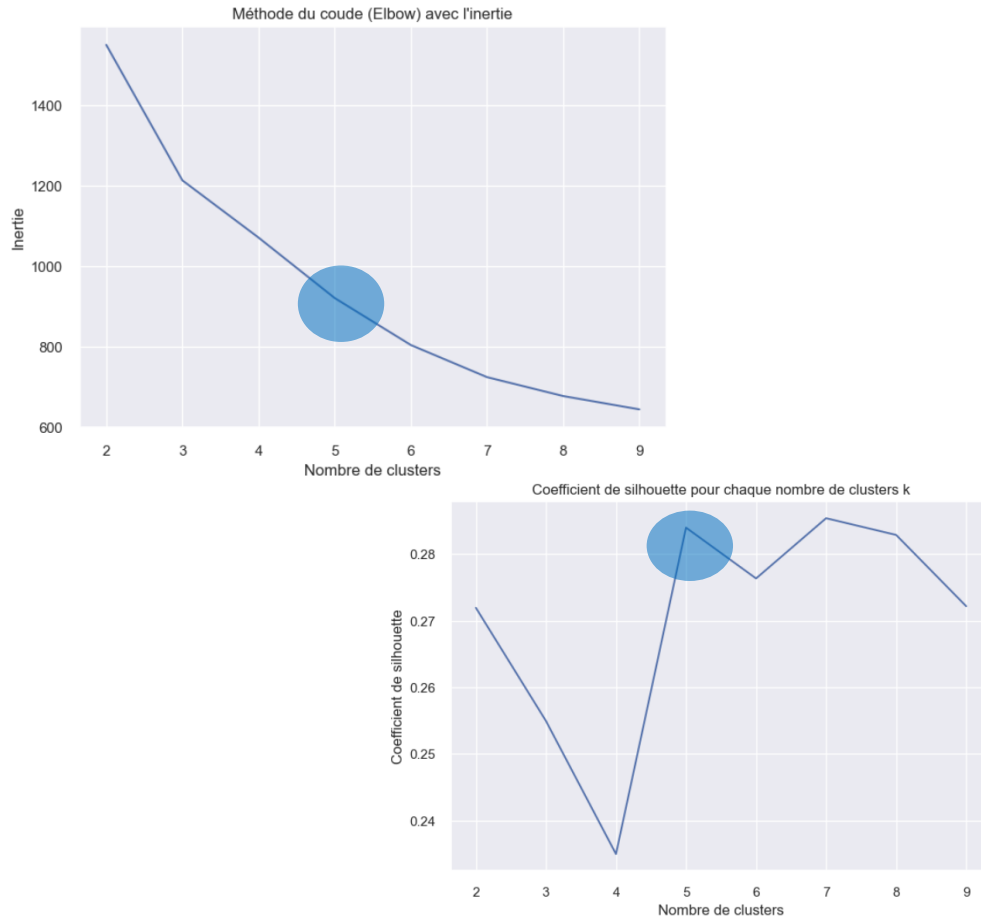
La deuxième méthode utilisée pour partitionner les pays est la méthode des k-means:

- Choix du nb de clusters : k
- Initialisation : k centres placés aléatoirement au sein du nuage de points
- Itérations :
  - Association des points au centre le plus proche,
  - Calcul des nouveaux centres de gravité
  - Association des points aux nouveaux centres
  - Etc... jusqu'à ce que les centres ne bougent plus (= convergence)

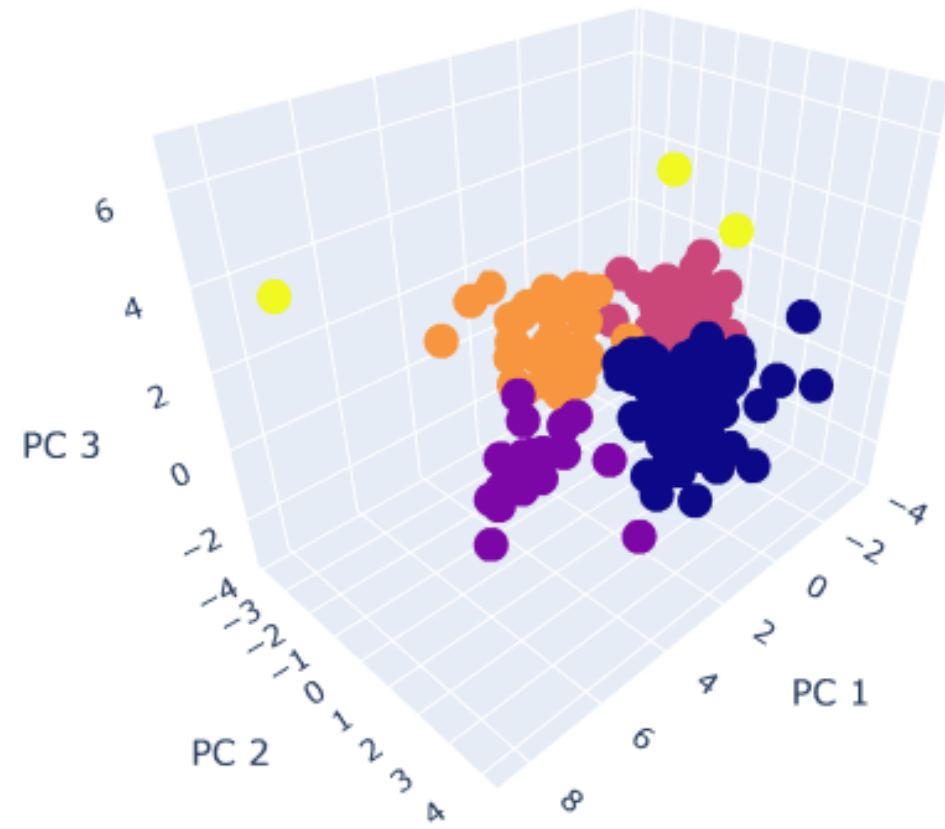


# Méthode du k-means

## Choix de k = 5 clusters



## Projection des 5 clusters sur les 3 premières composantes de l'ACP



# Comparaison entre les résultats des 2 méthodes

- Les résultats de la CAH et du k-means sont très similaires.

cluster_kmeans	1	2	3	4	5
cluster_CAH					
1	71.0	3.0	4.0	NaN	NaN
2	NaN	15.0	NaN	1.0	NaN
3	2.0	NaN	39.0	1.0	NaN
4	NaN	NaN	NaN	31.0	NaN
5	NaN	NaN	NaN	NaN	3.0

## Projection des clusters sur le premier facteuriel de l'ACP

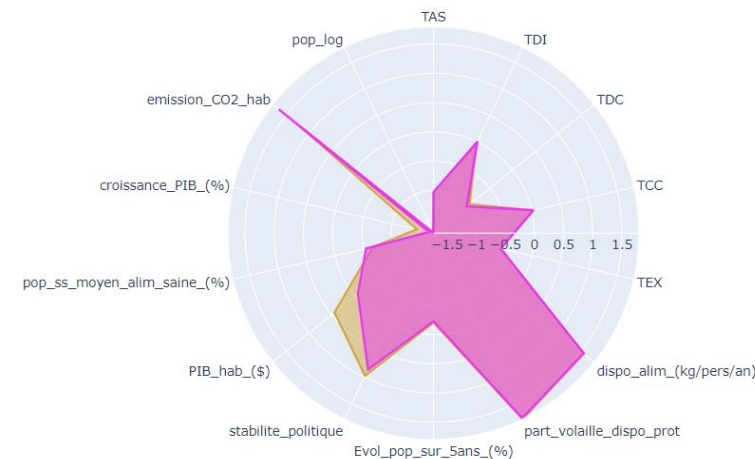
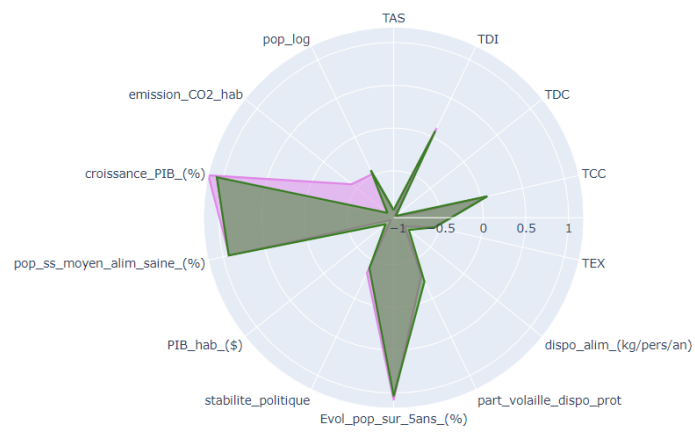
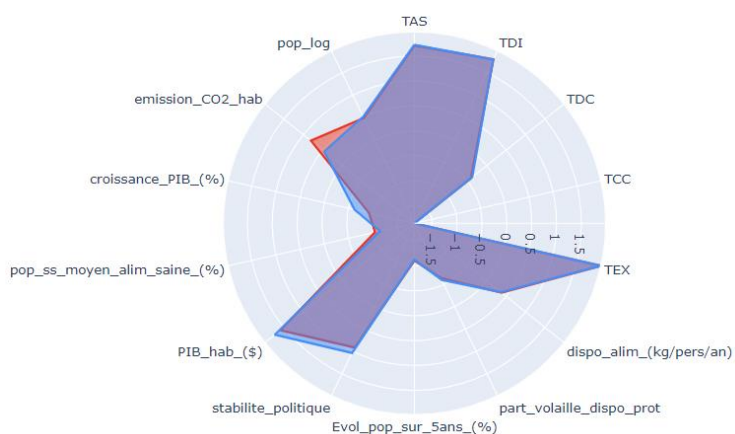


Pays cibles

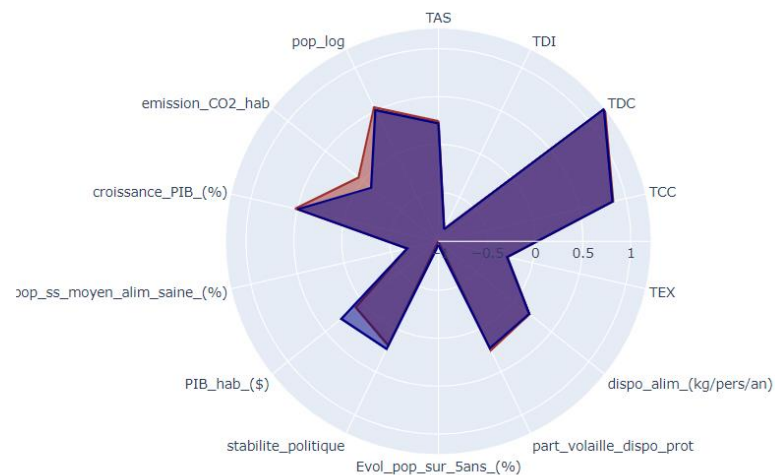
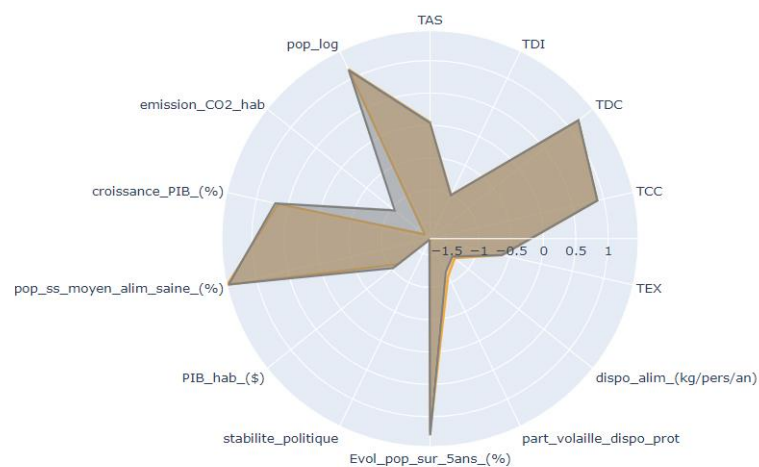




# Comparaison des centroïdes CAH et k-means



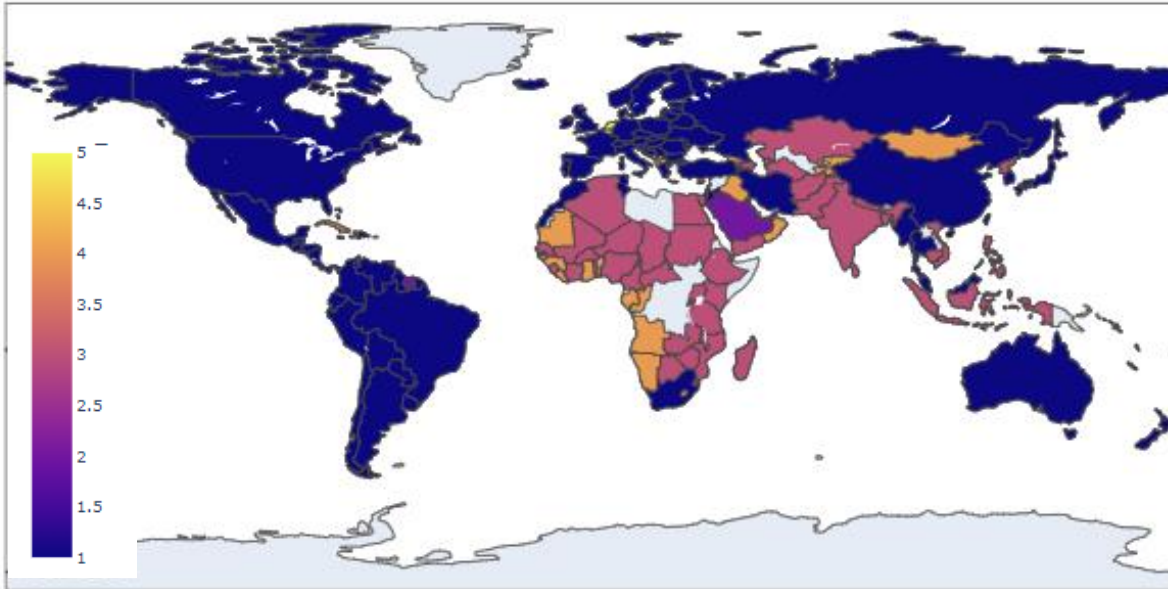
**Similarité dans les caractéristiques des centroïdes**



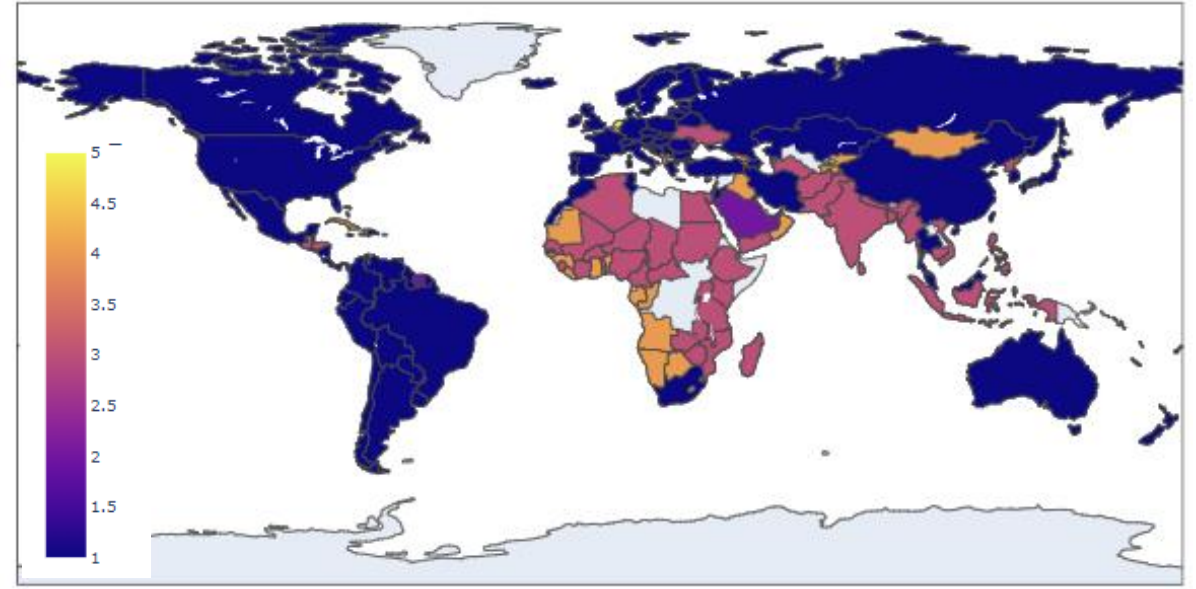


# Les pays des clusters

Clusters CAH



Clusters k-means



Cluster	Taille	Continents
1	71 à 73 pays	Europe et Amériques
2	16 à 18 pays	Majorité d'îles
3	42 à 43 pays	Afrique et Asie
4	31 à 33 pays	Afrique et Asie
5	3 pays	Europe et Asie

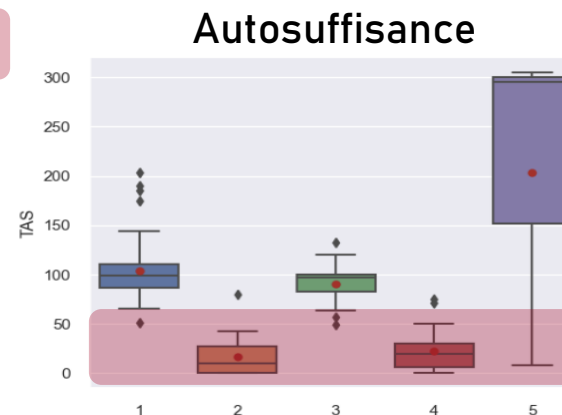
# Analyse des clusters

## Caractéristiques des pays cibles :

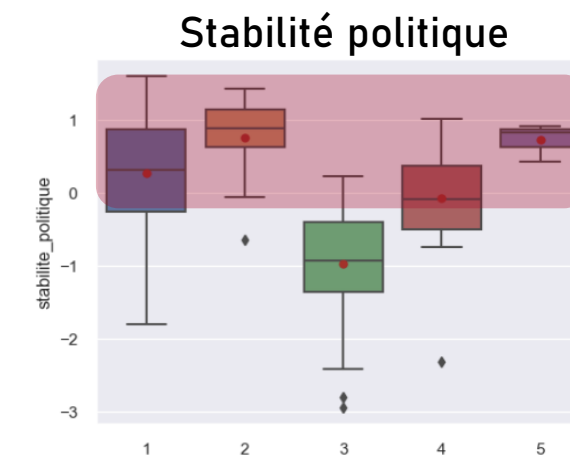
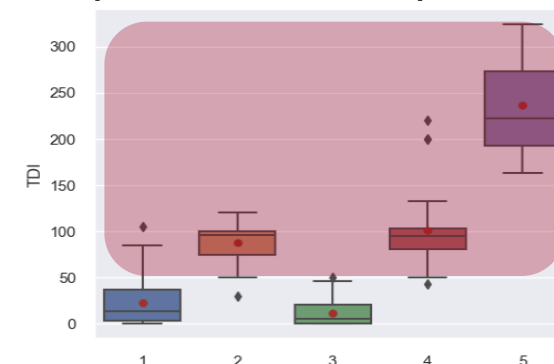
- Dépendant des importations (TAS <100%, TDI ~ 100%)
- Stable politiquement
- Bon pouvoir d'achat

## Analyse réalisée sur les résultats de la méthode k-means

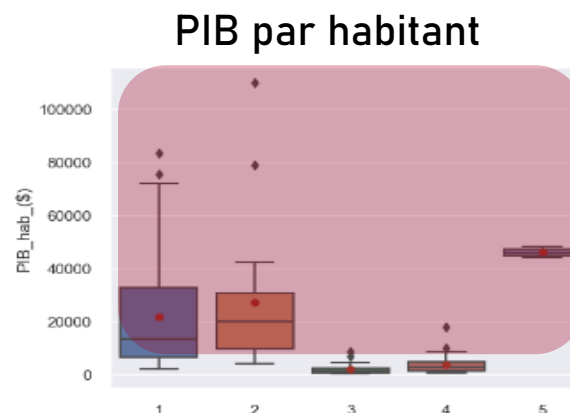
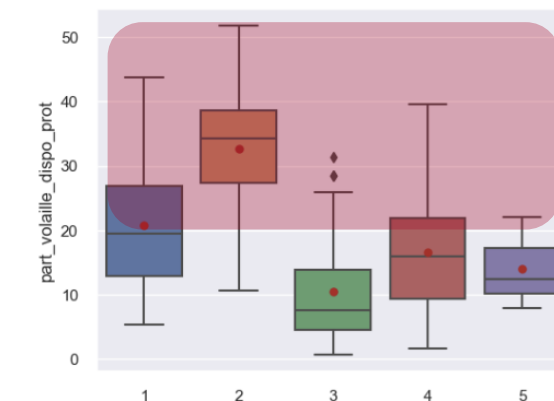
Cible



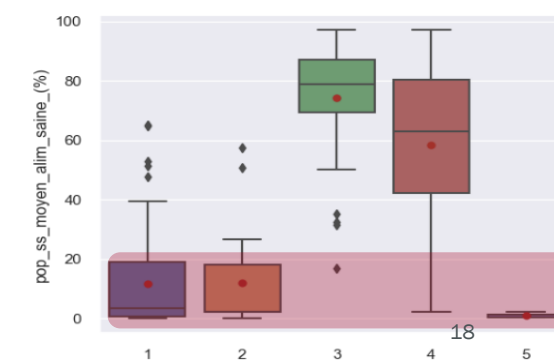
### Dépendance aux importations



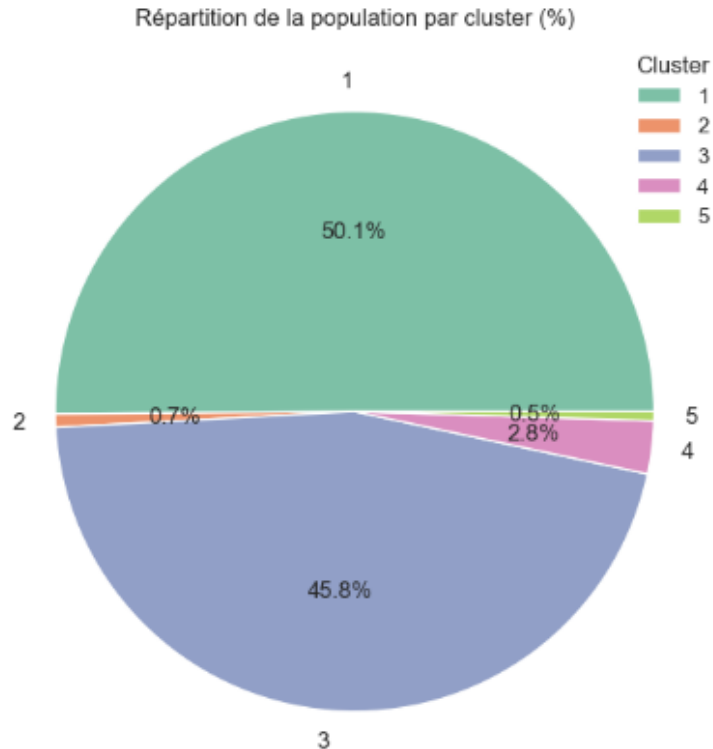
### Part de la volaille dans les apports en protéine animale



### Abordabilité de la nourriture



# Définition des clusters



Cluster	Taille	Description
1	73 pays	pays riches, stables politiquement, autosuffisants ou faible dépendance aux importations
2	18 pays	pays à revenu moyen supérieur et haut, très stable politiquement, très dépendants des importations
3	43 pays	pays instables, à faible/moyen revenu, autosuffisant en volaille
4	33 pays	pays à faible/moyen revenu, stabilité politique moyenne, très dépendants des importations
5	3 pays	pays qui font de l'import / export, stables politiquement, à revenu moyen supérieur, 1 pays sur les 3 non autosuffisant

cible

exclus



La poule qui chante

## Conclusion & recommandations



### NEXT STEP :

- **Identifications d'autres pays cibles** au sein des clusters 1 (poulet de bonne qualité) et 4 (poulet à bas prix)
- **Etude de marché :**
  - Distance (coût de transport, conservation produit frais, impact écologique)
  - Accords de libre-échange



### 19 PAYS CIBLES (Cluster 2 + 1 pays du cluster 5)

```
['Suriname',  
'Arabie saoudite',  
'Émirats arabes unis',  
'Koweït',  
'Chine - RAS de Hong-Kong',  
'Chine - RAS de Macao',  
'Antigua-et-Barbuda',  
'Bahamas',  
'Dominique',  
'Grenade',  
'Sainte-Lucie',  
'Saint-Kitts-et-Nevis',  
'Saint-Vincent-et-les Grenadines',  
'Trinité-et-Tobago',  
'Malte',  
'Luxembourg',  
'Nouvelle-Calédonie',  
'Polynésie française',  
'Samoa']
```

Répartition des pays cibles par continent (%)

