



La poule qui chante

ETUDE DE MARCHÉ

LES PAYS PROPICES À L'EXPORTATION

BRÈVE INTRODUCTION

La poule qui chante est une entreprise française d'agroalimentaire. Elle souhaite se développer à l'international. Le but est de déterminer les pays potentiels que l'on peut cibler pour l'exportation de poulets.

On cherche à cibler les pays avec les critères suivant :

Importation élevée
de poulet

Population
croissante

Niveau correct
d'alimentation

Ratio de protéines
correct

Consommation
de poulet correcte

Fort PIB

Stabilité politique

LES POINTS DE DISCUSSION



- Données utilisées
- Classification Ascendante Hiérarchique
- Classification K-Means
- Comparatif des 2 méthodes
- Analyse en Composante Principale
- Sélection de pays cibles

PREPARATION DES DONNÉES



LES DONNÉES UTILISÉES

SOURCE FAO - ORGANISATION DES NATIONS UNIES POUR L'ALIMENTATION ET L'AGRICULTURE

Population

2013 et 2018

	Domain	Area	Element	Item	Year	Unit	Value
0	Annual population	Afghanistan	Total Population - Both sexes	Population - Est. & Proj.	2013	1000 persons	32,269.59
1	Annual population	Afghanistan	Total Population - Both sexes	Population - Est. & Proj.	2018	1000 persons	37,171.92
2	Annual population	Albania	Total Population - Both sexes	Population - Est. & Proj.	2013	1000 persons	2,903.79
3	Annual population	Albania	Total Population - Both sexes	Population - Est. & Proj.	2018	1000 persons	2,882.74
4	Annual population	Algeria	Total Population - Both sexes	Population - Est. & Proj.	2013	1000 persons	38,140.13

Indicateur de PIB et de stabilité politique

PIB et PIB par habitant

	Domain	Area	Element	Item	Year	Unit	Value
0	ssi	Afghanistan	Value	gross_domestic_PPP	2013	I\$	2,116.50
1	ssi	Afghanistan	Value	gross_domestic_PPP	2018	I\$	2,033.80

Bilans alimentaires

Disponibilités en calories et protéines

Disponibilités en viande de poulet

Importations et exportations

	Domain	Area	Element	Item	Year	Unit	Value
0	Food Balances (2014-)	Afghanistan	Food supply (kcal/capita/day)	Grand Total	2018	kcal/capita/day	2,040.00
1	Food Balances (2014-)	Afghanistan	Protein supply quantity (g/capita/day)	Grand Total	2018	g/capita/day	55.52
2	Food Balances (2014-)	Afghanistan	Food supply (kcal/capita/day)	Vegetal Products	2018	kcal/capita/day	1,849.00
3	Food Balances (2014-)	Afghanistan	Protein supply quantity (g/capita/day)	Vegetal Products	2018	g/capita/day	44.73
4	Food Balances (2014-)	Afghanistan	Food supply (kcal/capita/day)	Animal Products	2018	kcal/capita/day	191.00

CONCATÉNATION ET PIVOT

Element	Area	Domestic supply quantity	Export Quantity	Import Quantity	Production	Value						food_supply_kcal_capita_day				protein_supply_g_capita_day				total_pop	
Item		Poultry Meat	Poultry Meat	Poultry Meat	Poultry Meat	energy_required		gross_domestic_PPP		political_stability		Animal Products	Grand Total	Poultry Meat	Vegetal Products	Animal Products	Grand Total	Poultry Meat	Vegetal Products	pop_est	
Year		2018	2018	2018	2018	2013	2018	2013	2018	2013	2018	2018	2018	2018	2018	2018	2018	2018	2018	2013	2018
0	Afghanistan	53.00	0.00	25.00	29.00	2,097.00	2,142.00	2,116.50	2,033.80	-2.52	-2.75	191.00	2,040.00	5.00	1,849.00	10.79	55.52	0.51	44.73	32,269,589.00	37,171,921.00
1	Albania	27.00	NaN	12.00	16.00	2,494.00	2,488.00	11,361.30	13,323.80	0.09	0.38	1,003.00	3,360.00	50.00	2,357.00	61.75	115.74	3.52	54.00	2,903,790.00	2,882,740.00
2	Algeria	295.00	NaN	0.00	295.00	2,313.00	2,284.00	11,319.10	11,642.20	-1.20	-0.83	390.00	3,322.00	24.00	2,932.00	24.73	91.83	2.12	67.10	38,140,133.00	42,228,408.00
3	American Samoa	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.93	1.20	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	55,713.00	55,465.00
4	Andorra	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.28	1.42	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	80,774.00	77,006.00

- Nettoyage des données
- Filtre et calcul des variables.
- On garde les pays pour lesquels, les données sont renseignées.



	chicken_prod	chicken_import_quantity	food_supply_kcal	food_supply_protein	ratio_animal_protein	pop_evo	PIB
country_code							
AFG	0.78	0.67	744,600.00	20,264.80	70.94	15.19	2,033.80
ALB	5.55	4.16	1,226,400.00	42,245.10	194.74	-0.72	13,323.80
DZA	6.99	0.00	1,212,530.00	33,517.95	98.30	10.72	11,642.20
AGO	0.88	10.87	870,525.00	19,363.25	112.22	18.43	6,933.50
ATG	0.00	72.70	892,425.00	29,269.35	235.50	5.21	21,030.40

Total de 157 pays

- Aucun doublon
- Aucune valeur manquante

CENTRAGE ET REDUCTION

Les variables sont exprimées dans des unités différentes.
On applique un centrage et une réduction sur ces dernières afin que leurs moyennes respectives soient égales à 0 avec un écart-type de 1.
Cela permettra aux variables d'influer de manière égale nos calcul.

Librairie sklearn

Bibliothèque libre Python destinée
à l'apprentissage automatique

```
X = df_main.values
```

```
std_scale = preprocessing.StandardScaler().fit(X)
```

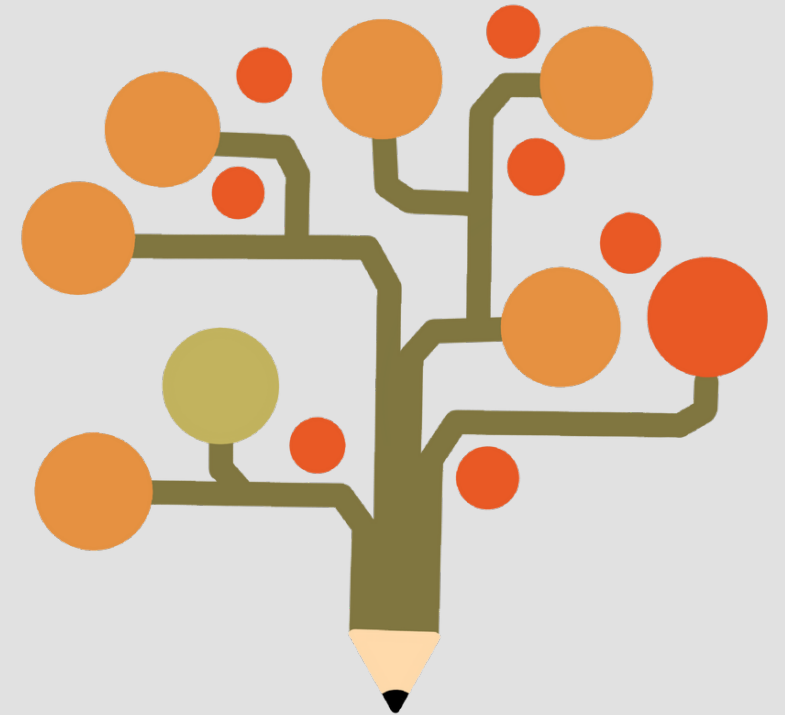
```
X_scaled = std_scale.transform(X)
```

Analyse sur
157 pays

	chicken_prod	chicken_import_quantity	food_supply_kcal	food_supply_protein	ratio_animal_protein	pop_evo	PIB	stability
country_code								
AFG	0.78	0.67	744,600.00	20,264.80	70.94	15.19	2,033.80	-2.75
ALB	5.55	4.16	1,226,400.00	42,245.10	194.74	-0.72	13,323.80	0.38
DZA	6.99	0.00	1,212,530.00	33,517.95	98.30	10.72	11,642.20	-0.83
AGO	0.88	10.87	870,525.00	19,363.25	112.22	18.43	6,933.50	-0.32
ATG	0.00	72.70	892,425.00	29,269.35	235.50	5.21	21,030.40	0.85

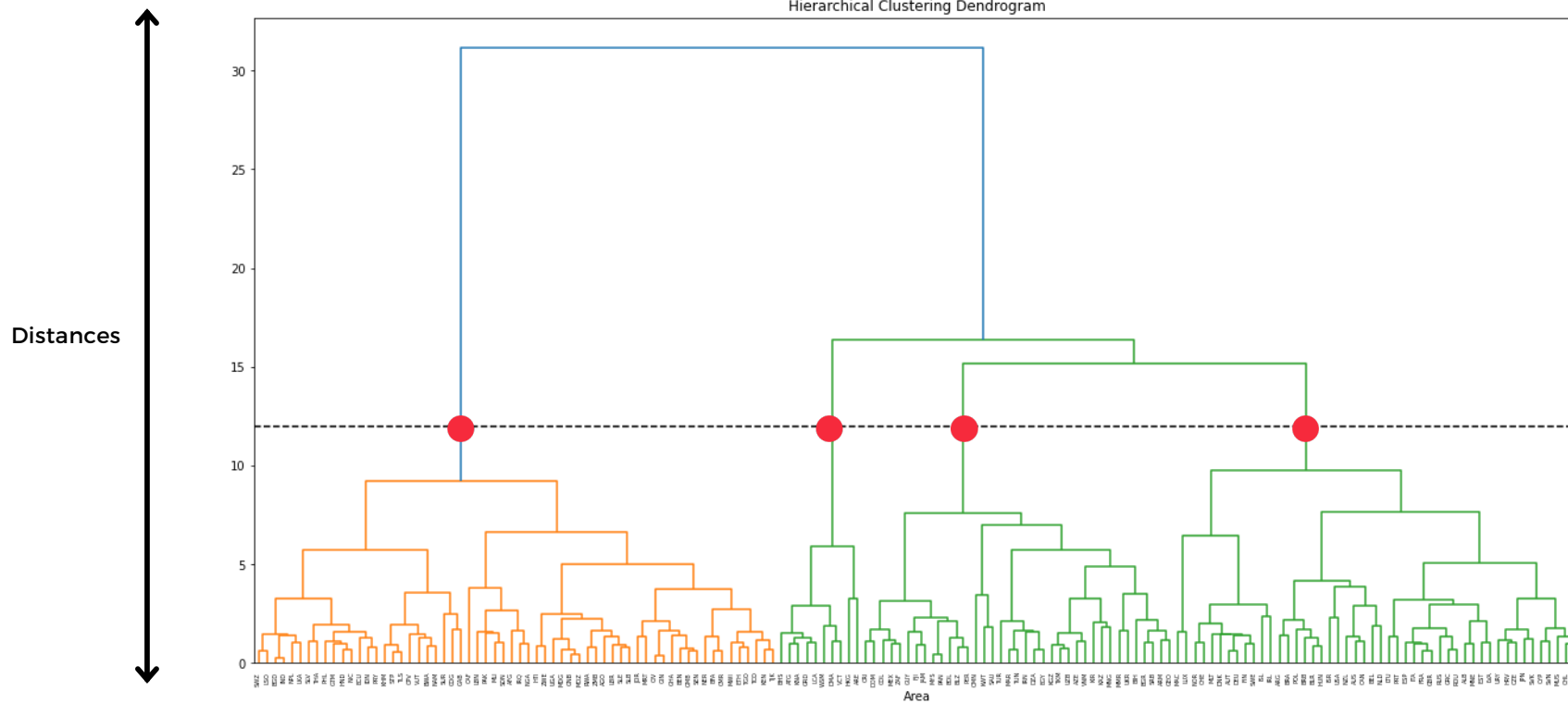
	chicken_prod	chicken_import_quantity	food_supply_kcal	food_supply_protein	ratio_animal_protein	pop_evo	PIB	stability
count	157.00	157.00	157.00	157.00	157.00	157.00	157.00	157.00
mean	17.46	10.93	1,054,673.31	30,116.96	157.11	6.94	20,928.03	-0.05
std	18.20	18.36	165,311.23	7,686.96	58.23	6.10	21,538.14	0.89
min	0.00	0.00	651,890.00	15,479.65	36.04	-6.90	933.10	-2.75
25%	2.18	0.39	940,970.00	24,020.65	105.27	2.60	5,033.50	-0.60
50%	10.96	4.16	1,047,915.00	30,149.00	161.87	6.20	13,323.80	0.02
75%	27.07	11.18	1,203,405.00	35,981.70	205.12	11.37	30,036.10	0.75
max	78.27	111.37	1,418,025.00	53,337.45	266.77	28.28	132,417.00	1.54

CLASSIFICATION HIÉRARCHIQUE ASCENDANT



DENDROGRAMME

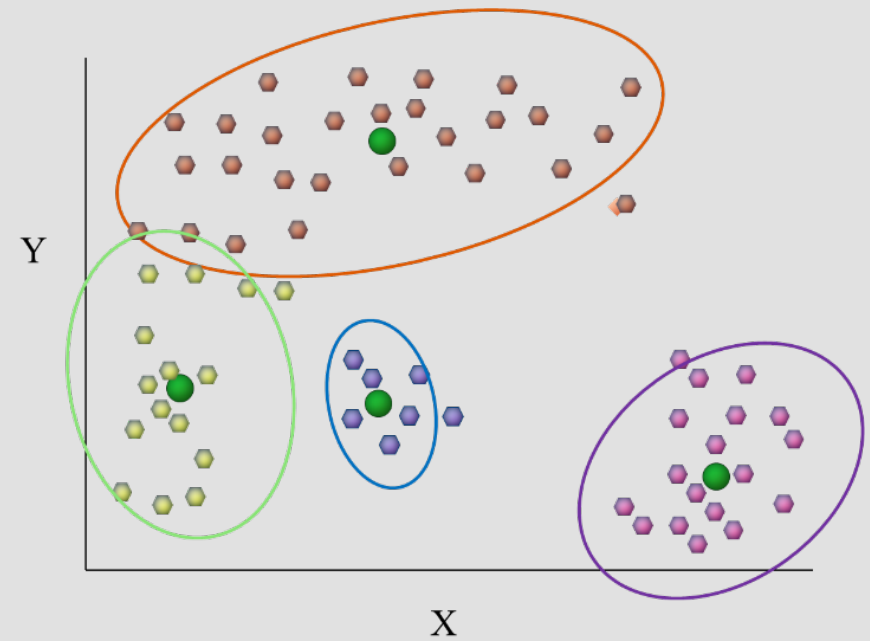
Effectifs			chicken_prod	chicken_import_quantity	food_supply_kcal	food_supply_protein	ratio_animal_protein	pop_evo	PIB	stability
0	62	0	5.39	4.69	904,552.42	22,770.47	101.40	11.54	5,667.38	-0.63
1	10	1	4.65	70.99	1,030,504.50	32,520.41	219.86	3.38	27,437.21	0.91
2	37	2	23.01	6.34	1,097,870.68	31,571.71	159.03	6.59	16,338.50	-0.25
3	48	3	31.44	10.03	1,220,316.67	37,984.11	214.50	2.02	42,821.40	0.65



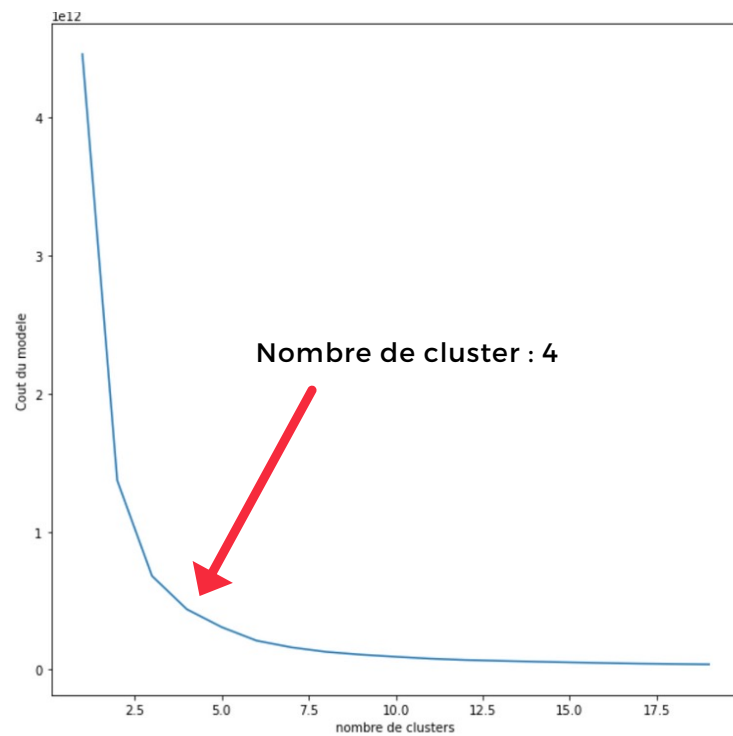
Librairie sklearn
Méthode **fcluster** de
scipy.cluster.hierarchy

Calcul des distances
euclidienne entre les pays
Classification par méthode de
Ward

CLASSIFICATION MÉTHODE DES K-MEANS



K-MEANS CLUSTERS



Librairie sklearn

Méthode **cluster**

Clustering par K-means

```
model = cluster.KMeans(n_clusters=n_clust)
```

```
model.fit(X_scaled)
```

```
# Récupération des clusters attribués à chaque individu
```

```
k_means_clusters = model.labels_
```

Effectifs

0	10
1	39
2	59
3	49

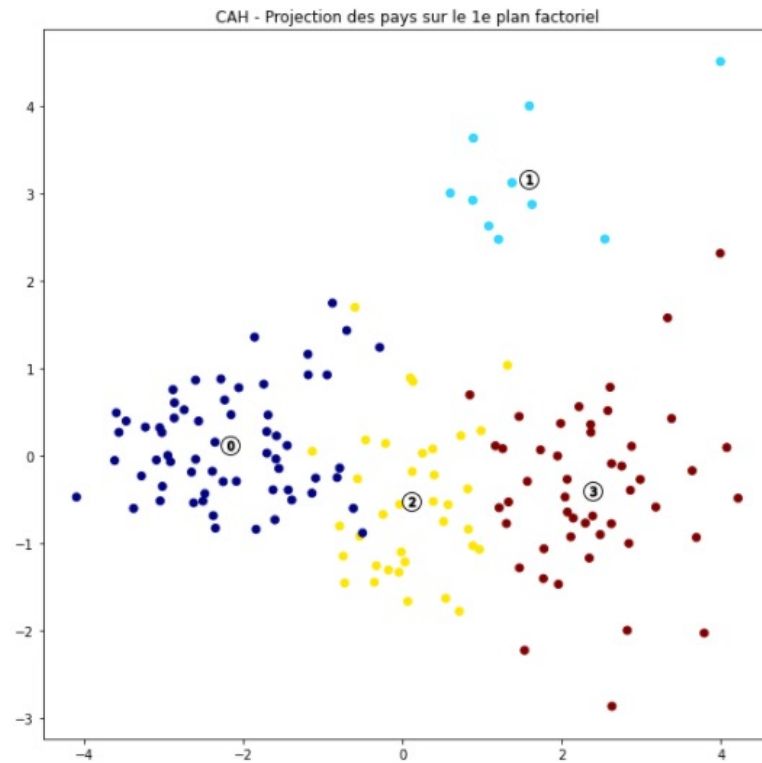
	chicken_prod	chicken_import_quantity	food_supply_kcal	food_supply_protein	ratio_animal_protein	pop_evo	PIB	stability	clusters
0	4.65	70.99	1,030,504.50	32,520.41	219.86	3.38	27,437.21	0.91	0
1	29.39	11.53	1,243,957.44	39,361.69	218.67	2.35	47,764.26	0.69	1
2	4.06	4.72	901,655.17	22,896.94	99.39	12.35	5,334.86	-0.64	2
3	26.71	5.68	1,093,197.35	30,961.91	164.80	4.82	17,015.63	-0.13	3

COMPARATIF DES 2 METHODES

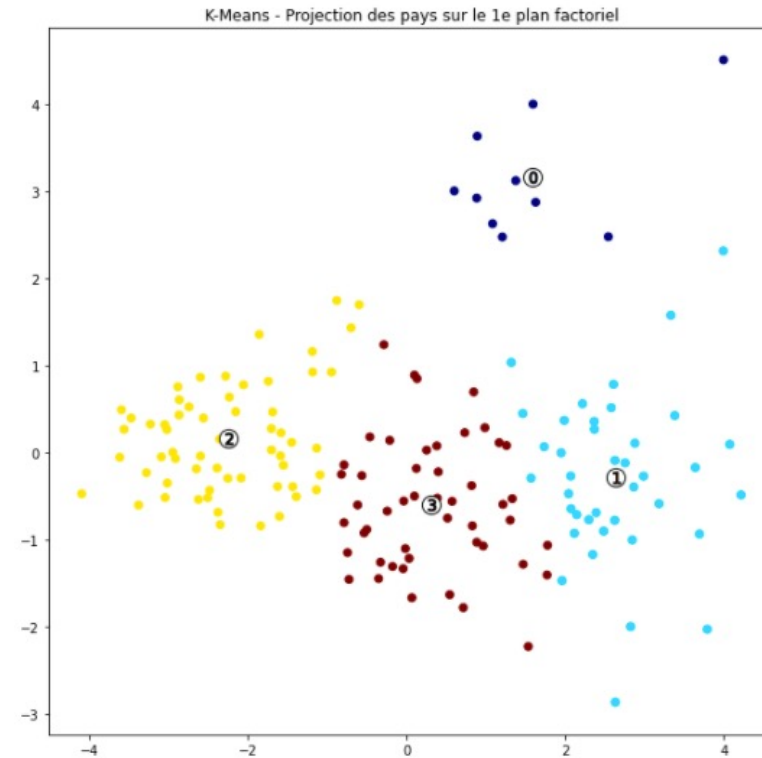


VISUALISATION PAR PROJECTION

Classification Ascendante Hiérarchique



Méthode des K-Means



SILHOUETTE SCORE

Classification Ascendante Hiérarchique

Silhouette de SkLearn :

Analyser les inerties intra et inter-classes des deux méthodes.

Nous cherchons à avoir :

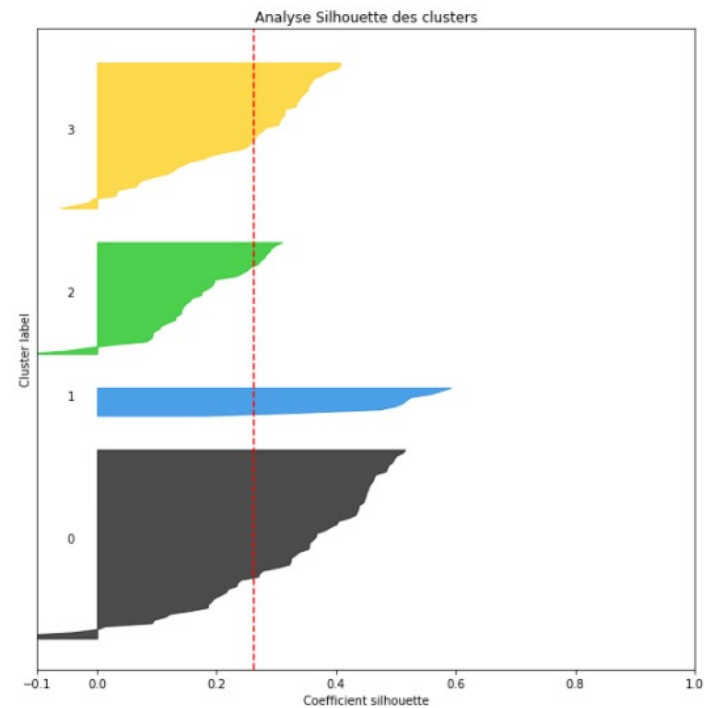
- la variation intra-classe la plus petite
- la variation inter-classe la plus grande

Score varie de -1 à 1.

Si le score est de 1, le cluster est dense et bien séparé par rapport aux autres clusters.

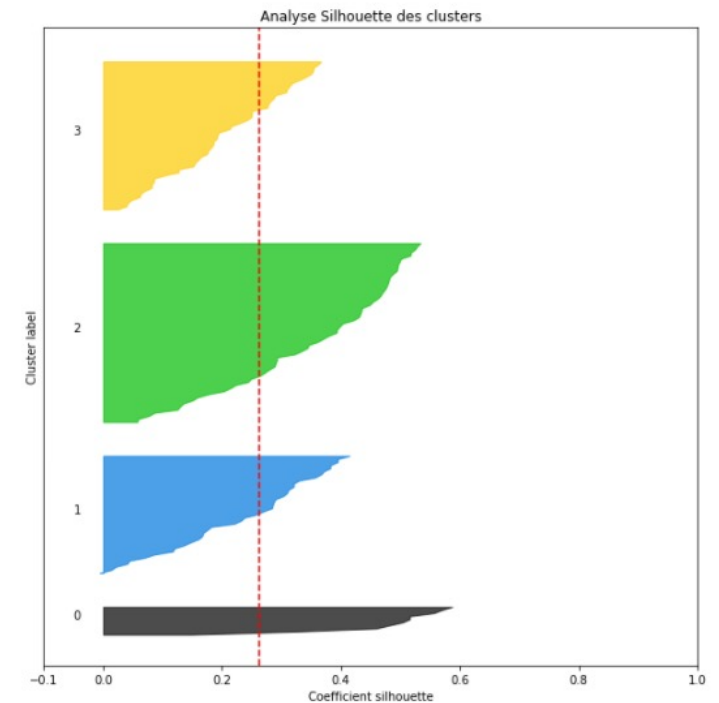
Une proche de 0 représente des clusters qui se chevauchent avec des échantillons très proches de la frontière de valeur des clusters voisins.

The average CAH silhouette_score is : 0.2629630345373933



Méthode des K-Means

The average K-Means silhouette_score is : 0.2819083543906498



GROUPE DE PAYS

Choix de la méthode

On choisit la méthode des K-Means.

Groupe de pays

```
0
[ 'ATG' 'BHS' 'HKG' 'DMA' 'GRD' 'KNA' 'LCA' 'VCT' 'WSM' 'ARE' ]
1
[ 'ALB' 'ARG' 'AUS' 'AUT' 'BEL' 'CAN' 'MAC' 'HRV' 'CZE' 'DNK' 'EST' 'FIN'
  'FRA' 'DEU' 'GRC' 'HUN' 'ISL' 'IRL' 'ISR' 'ITA' 'JPN' 'KWT' 'LVA' 'LTU'
  'LUX' 'MLT' 'MNE' 'NLD' 'NZL' 'NOR' 'POL' 'PRT' 'ROU' 'SVN' 'ESP' 'SWE'
  'CHE' 'GBR' 'USA' ]
2
[ 'AFG' 'AGO' 'BGD' 'BEN' 'BWA' 'BFA' 'CPV' 'KHM' 'CMR' 'CAF' 'TCD' 'COG'
  'CIV' 'SWZ' 'ETH' 'GAB' 'GMB' 'GHA' 'GTM' 'GIN' 'GNB' 'HTI' 'HND' 'IND'
  'IDN' 'IRQ' 'JOR' 'KEN' 'KGZ' 'LBN' 'LSO' 'LBR' 'MDG' 'MWI' 'MLI' 'MRT'
  'MOZ' 'NAM' 'NPL' 'NIC' 'NER' 'NGA' 'OMN' 'PAK' 'PHL' 'RWA' 'STP' 'SEN'
  'SLE' 'SLB' 'LKA' 'SDN' 'TJK' 'TLS' 'TGO' 'UGA' 'VUT' 'ZMB' 'ZWE' ]
3
[ 'DZA' 'ARM' 'AZE' 'BRB' 'BLR' 'BLZ' 'BOL' 'BIH' 'BRA' 'BGR' 'CHL' 'COL'
  'CRI' 'CYP' 'DOM' 'ECU' 'EGY' 'SLV' 'FJI' 'GEO' 'GUY' 'IRN' 'JAM' 'KAZ'
  'KIR' 'MYS' 'MUS' 'MEX' 'MNG' 'MAR' 'MMR' 'PAN' 'PRY' 'PER' 'RUS' 'SAU'
  'SRB' 'SVK' 'ZAF' 'SUR' 'THA' 'TTO' 'TUN' 'TUR' 'TKM' 'UKR' 'URY' 'UZB'
  'VNM' ]
```

GROUPE DE PAYS

Groupe 0 :

Production de poulet faible, importation importante, fort ratio en proteines, fort taux d'évolution du nombre d'habitants. Ce sont de bons candidats pour nos exportations.

Groupe 1 :

PIB très élevé avec une production et importation de poulet importante, disponibilité en kcal et proteines élevées. Ce sont de bons candidats pour nos exportations.

Groupe 2 :

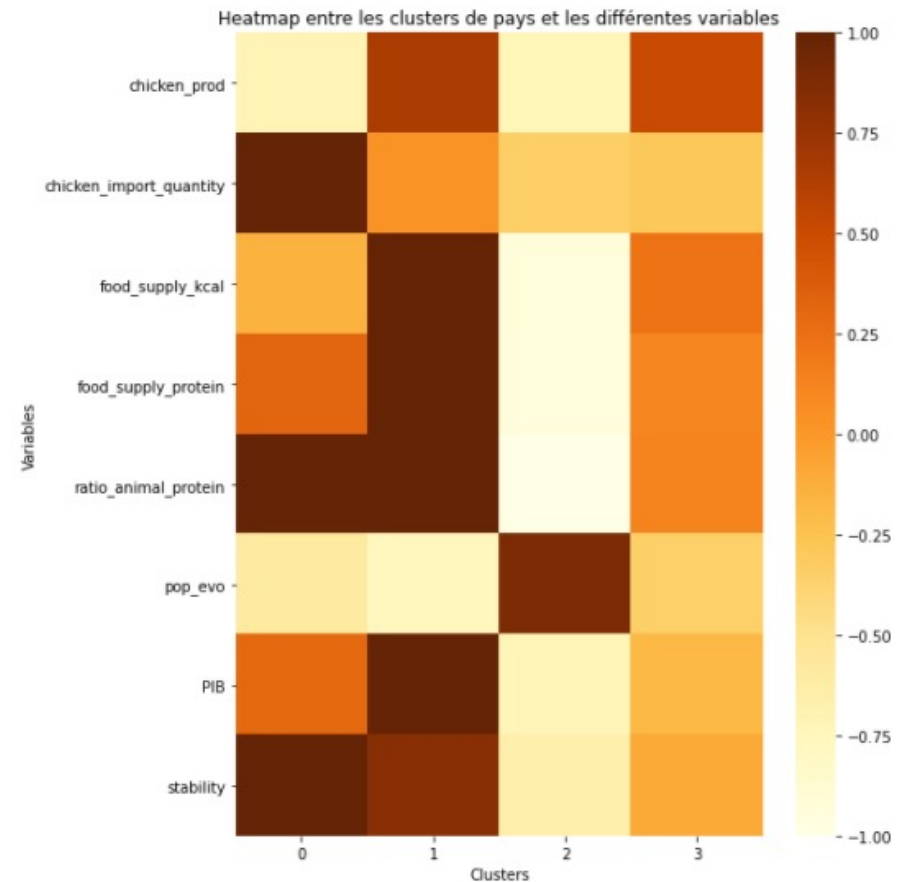
Production et importations très faibles, ratio de proteines le plus faible, PIB bas. Ils ne seraient pas de bons candidats pour notre exportation.

Groupe 3 :

Les pays de ce groupe sont producteurs de poulet. La quantité importée est faible et le PIB est moyen.

Première approche

A première vue, on ciblera les pays des groupes 0 et 1.



ANALYSE PAR COMPOSANTES PRINCIPALES



ACP

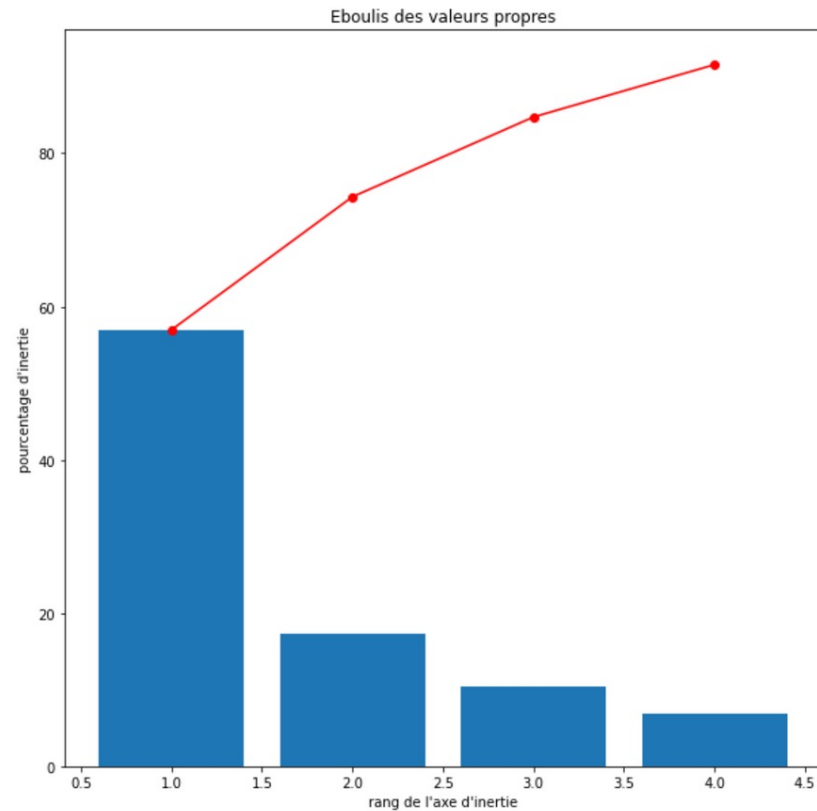
Intérêt et utilisation

- Vérifier la première approche de classification
- Utilisation des variables réduites et centrées
- Visualiser les corrélations
- Réduire les variables à x dimensions

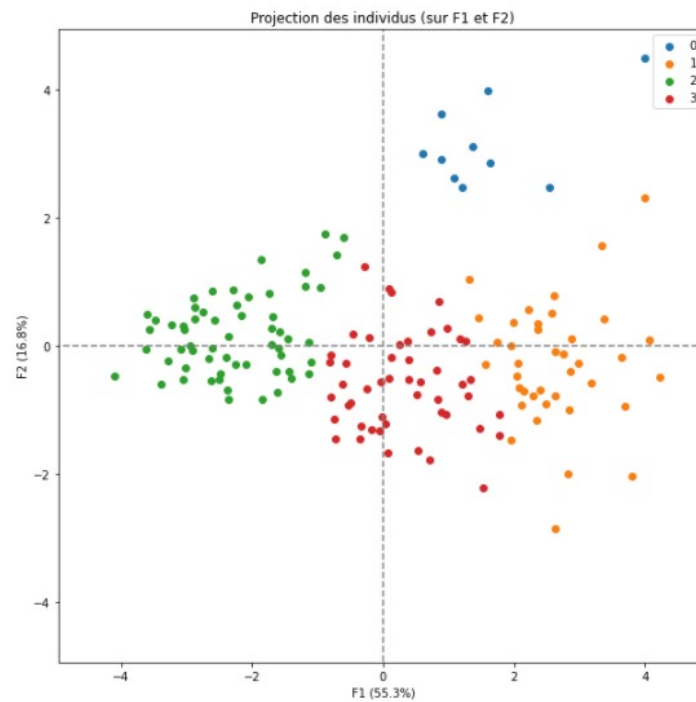
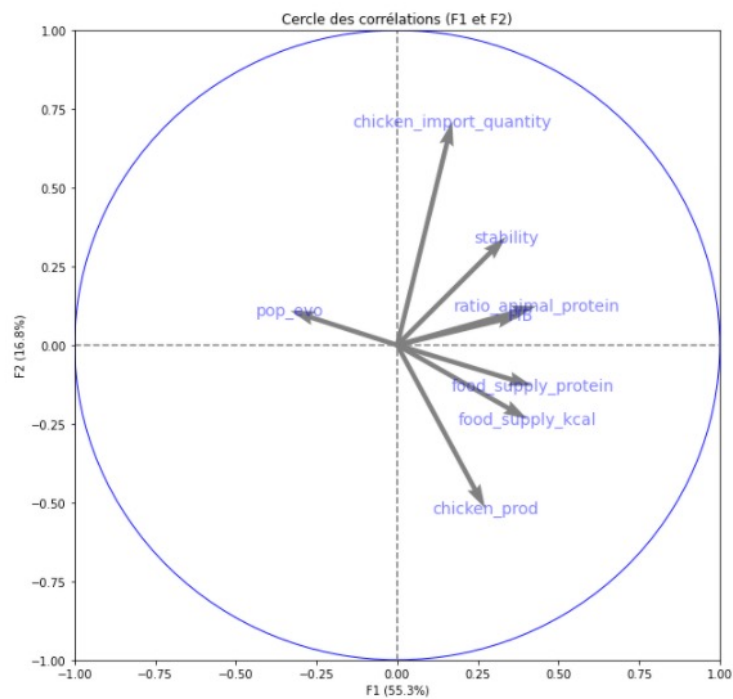
Rang de classe d'inertie choisie : 4

2 composantes : 74,3% des variables exprimées

4 composantes : 91,5% des variables exprimées



COMPOSANTES F1 ET F2



Composante F1 :

57% de l'inertie totale. Elle représente l'indice du besoin alimentaire des pays par les variables :

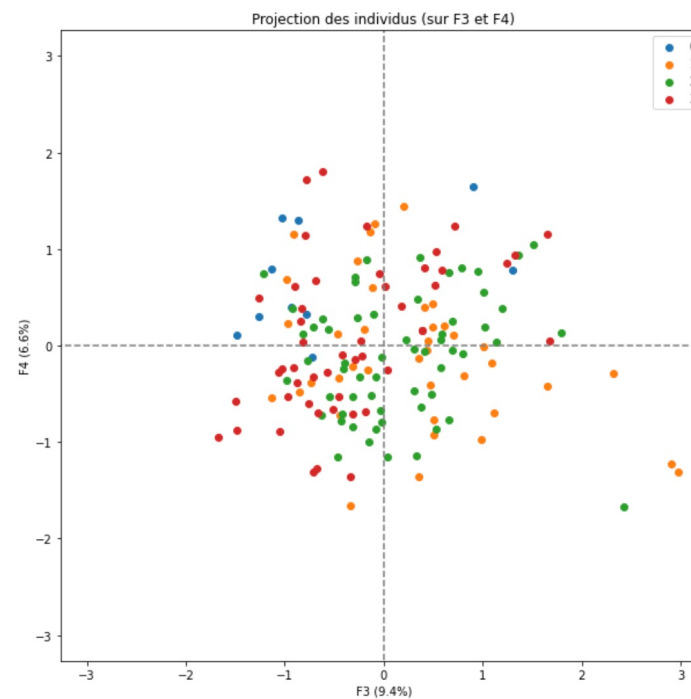
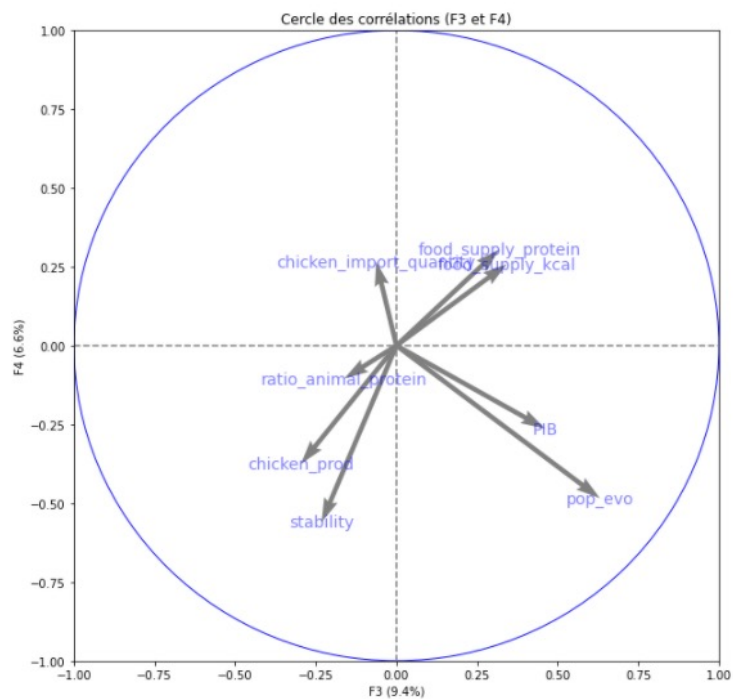
- food_supply_kcal
- food_supply_protein
- ratio_animal_protein

Plus l'indice est élevé, plus le pays consommera de la nourriture.

Composante F2 :

17% de l'inertie totale. Elle représente le manque de viande de poulet des pays.

COMPOSANTES F3 ET F4



Composante F3 :

10% de l'inertie totale. Elle représente l'indice économique des pays par les variables :

- pop_evo
- PIB

Composante F4 :

7% de l'inertie totale. Elle représente la stabilité politique.

PAYS À CIBLER



LISTE DES PAYS

Première approche

A première vue, on ciblera les pays des groupes 0 et 1.

	chicken_prod	chicken_import_quantity	food_supply_kcal	food_supply_protein	ratio_animal_protein	pop_evo	PIB	stability	clusters
0	4.65	70.99	1,030,504.50	32,520.41	219.86	3.38	27,437.21	0.91	0
1	29.39	11.53	1,243,957.44	39,361.69	218.67	2.35	47,764.26	0.69	1
2	4.06	4.72	901,655.17	22,896.94	99.39	12.35	5,334.86	-0.64	2
3	26.71	5.68	1,093,197.35	30,961.91	164.80	4.82	17,015.63	-0.13	3

Calcul de score avec les parapètre de l'ACP

SCORE = (F1 x 0,569) + (F2 x 0,174) + (F3 x 0,104) + (F4 x 0,68)

country_code	area	population	pop_evo	chicken_prod	chicken_import_quantity	food_supply_kcal	food_supply_protein	nal_protein	PIB	stability	dist_from_FRA	clusters	F1	F2	F3	F4	f_score
MAC	China, Macao SAR	631,636.00	9.40	6.33	41.16	1,214,355.00	39,956.55	236.43	132,417.00	1.26	9,614.59	1	3.99	2.31	2.90	-1.22	2.94
HKG	China, Hong Kong SAR	7,371,730.00	3.99	3.39	111.37	1,192,455.00	46,913.45	266.77	61,062.90	0.79	9,645.86	0	4.00	4.50	0.90	1.65	2.93
LUX	Luxembourg	604,245.00	11.37	0.00	19.86	1,264,725.00	39,657.25	226.66	113,590.00	1.37	287.97	1	3.33	1.57	2.97	-1.31	2.46
IRL	Ireland	4,818,690.00	4.40	31.13	21.79	1,418,025.00	42,916.70	223.54	84,303.30	1.03	780.84	1	4.07	0.09	1.65	-0.42	2.43
ISL	Iceland	336,713.00	2.91	26.73	2.97	1,333,710.00	53,337.45	262.22	56,722.10	1.41	2,235.75	1	4.22	-0.49	1.09	-0.18	2.35
NLD	Netherlands	17,059,560.00	1.28	62.96	34.29	1,203,405.00	38,941.85	238.83	56,060.90	0.86	430.89	1	3.64	-0.18	-0.45	-0.72	1.98
DNK	Denmark	5,752,126.00	2.02	27.12	24.51	1,241,365.00	42,734.20	248.56	56,281.30	0.95	1,028.88	1	3.38	0.42	0.44	-0.05	1.97
ARE	United Arab Emirates	9,630,959.00	4.71	4.98	63.23	1,209,610.00	40,244.90	167.04	66,968.20	0.71	5,251.80	0	2.54	2.47	1.30	0.78	1.88
USA	United States of America	327,096,265.00	3.38	68.17	0.39	1,380,430.00	41,515.10	235.80	61,585.80	0.47	5,852.74	1	3.79	-2.03	0.51	-0.92	1.85
LTU	Lithuania	2,801,264.00	-6.90	47.84	16.06	1,245,015.00	46,227.25	227.24	35,427.50	0.75	1,701.10	1	3.69	-0.94	-0.98	0.68	1.74

CONCLUSION



La Classification Hiérarchique Ascendante et la méthode de classification K-Means, nous ont permis de créer des groupes de pays.

A partir des groupes les plus propices à notre importation, nous avons sélectionné des pays grâce à l'Analyse Principale des Composantes des données.

Nous devrions donc cibler les pays des groupes 0 et 1 et plus particulièrement les pays :

- MAC
- HKG
- LUX



La poule qui chante

MERCI

Avez vous des questions ?