

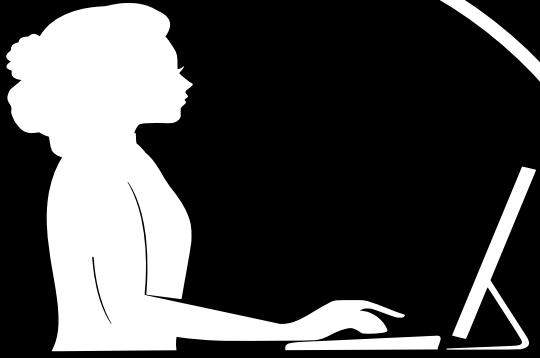
ETUDE DE MARCHÉ

Déploiement Export



SOMMAIRE



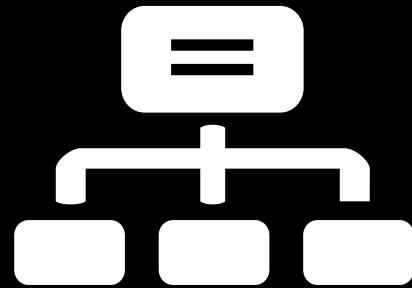


LES DONNEES

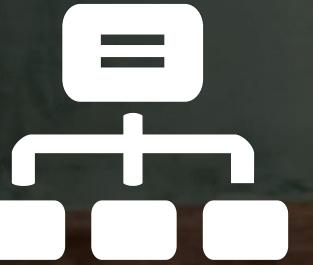
LES DONNEES

Préparation des données

- Fichiers utilisés
- Nettoyage
- Variables complémentaires
- Traitement des outliers
- Fichier final



CLUSTERING



CLUSTERING

Découpage le plus performant

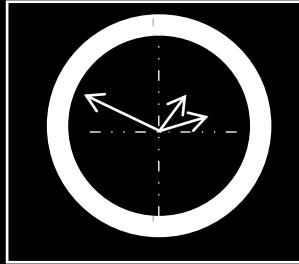
- Méthode CAH
- Méthode du coude
- Méthode Davies Bouldin
- Méthode Calinski Harabasz
- Méthode du Coef. De Silhouette

Comparaison CAH Vs K.Means

- Choix de découpage
- Comparaison des 2 partitionnements

F2

17%



F1

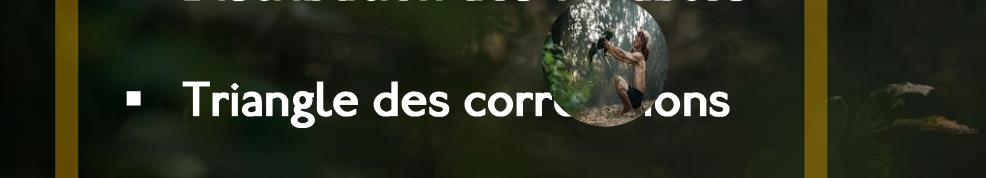
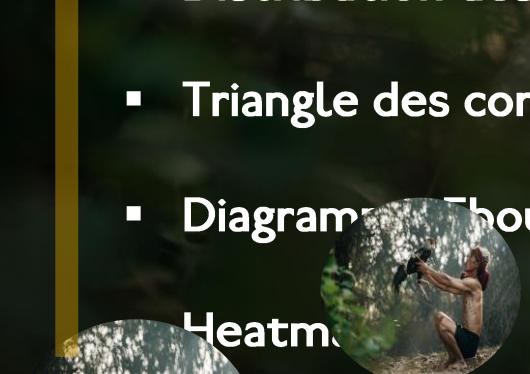
37%

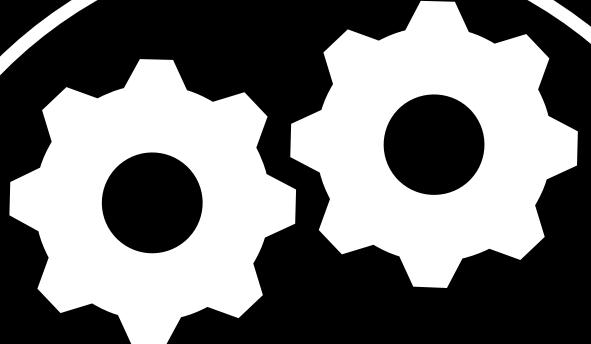
ACP

ACP

Choix des composantes principales

- Distribution des variables
- Triangle des corrélations
- Diagramme de Choulis
- Heatmaps





Analyse des
clusters

Analyse des clusters

Etude des clusters

- Projection sur les plans factoriels
- Moyenne des distributions
- Situation géographique



CONTEXTE

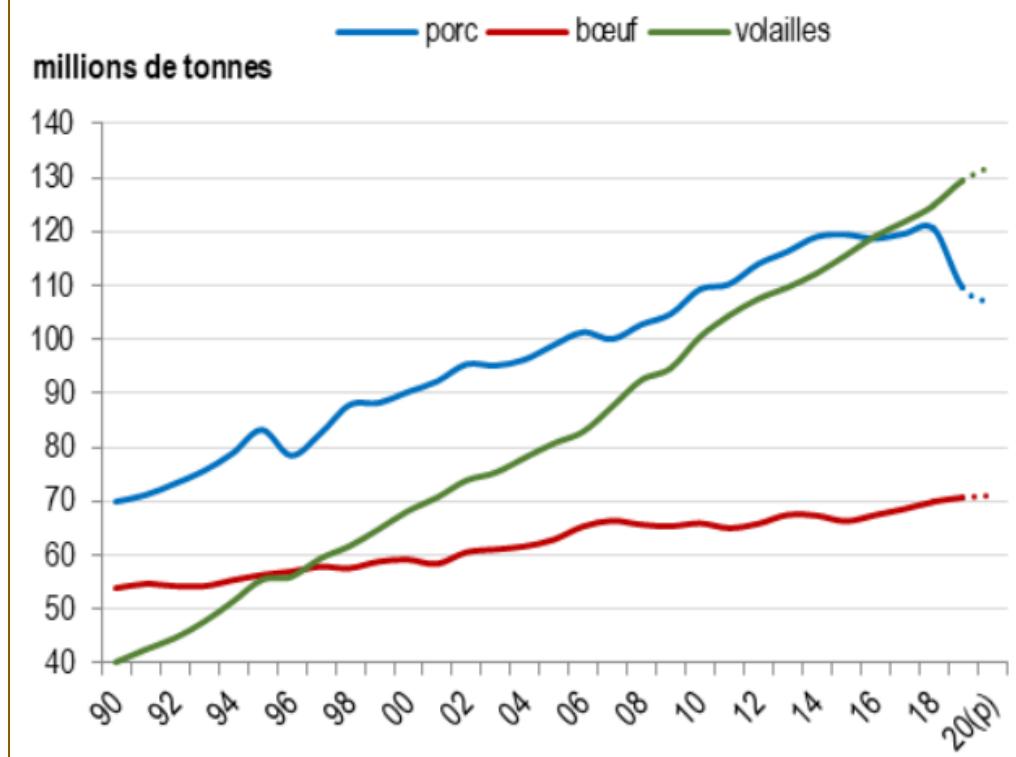


LE MARCHÉ DE LA VOLAILLE

- **N°1 des viandes vendues et consommée dans le monde :** Viande la plus consommée dans le monde
 - Première viande produite et consommée dans le monde en 2019 avec 129 millions de tonnes (Mt) devant la viande porcine (109 Mt) la viande bovine (70 Mt) et la viande ovine (15 Mt).
- **Plus forte croissance qui ouvre toujours plus de perspectives :**
 - Taux de croissance annuel moyen de 3,3 % depuis 2000 contre 1,4 % pour la viande porcine 1,5 % pour la viande ovine et 0,9 % pour la viande bovine.
 - D après l'OCDE sa consommation devrait continuer de croître de 1,5% par an jusqu' en 2027
 - Selon les projections de l'OCDE et de la FAO la volaille restera le principal moteur de la croissance de la production de viande pour la prochaine décennie (2019-2029)
- **Marché extérieur:**
 - Selon les perspectives OCDE-FAO à horizon 2029 la viande de volaille représentera avec la viande bovine la majeure partie de la croissance des importations mondiales de viande

(Sources : ITAVI/OCDE / FAO)

"Production mondiale de viandes sur la période 1990-2019"



source : ITAVI, d'après OCDE/FAO

OBJECTIFS

Sur la base des perspectives de marché vues ci-dessus, un positionnement sur le marché extérieur rapide afin de pouvoir s'assurer des parts de marché semble évident.

A ce titre afin de préparer notre projet de pénétration du marché extérieur j'ai été sollicité afin de pouvoir réaliser un premier regroupement de pays qui pourrait constituer une cible potentielle. Pour se faire j'ai utilisé différentes sources de données que j'ai transformé puis à l'aide de différentes méthodes de clustering j'ai généré des groupes de pays dont j'ai étudié les similitudes à l'appui d'une ACP.



LES DONNEES



LES FICHIERS UTILISÉS

Selon l'OCDE La consommation de viande est liée aux niveaux de vie, aux modes de consommation alimentaire, aux prix à la consommation, ainsi qu'à l'incertitude au niveau macroéconomique et aux chocs sur le PIB.

- Afin de pouvoir prendre en compte l ensemble de ces critères j ai choisi de travailler sur les fichier suivants:

Fichier Disponibilité alimentaire - Fichier Population - Fichier PIB - Fichier Inflation - Fichier Stabilité politique

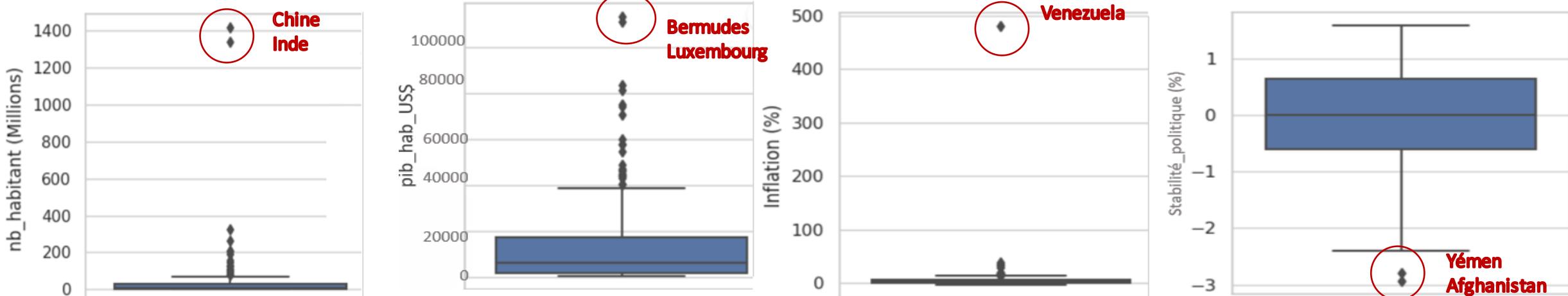
AJOUT DE VARIABLES

- A partir de ces fichier j ai calculé et ajouté des variables

NETTOYAGE

- Fichiers filtrés sur l année 2017 car il s agit de la source la plus complète
- Très peu de valeurs manquantes : 1,02% du Data frame final
- Les valeurs manquantes ont été ajoutée manuellement pour celles que j ai pu trouver, pour les autres celles-ci ont été remplacée par 0.

TRAITEMENT DES VALEURS EXTRÊMES



- Suppression des Pays suivants : Chine, Inde, Venezuela, Bermudes, Yémen, Afghanistan
- On conserve le Luxembourg car le pays est à 100% dépendant des importations de volaille
- Suppression des Pays Bas et la Belgique qui sont complètement autosuffisant dans la production de volaille et dont les Taux constituent également 2 outliers.

Fichier final

df.head()

	Zone	nb_habitant	Croiss.dem	Import_Volaille	TDI volaille (%)	TAS volaille (%)	dispo animale	dispo volaille	dispo Prot.animale	pib_hab_US\$	Inflation (%)	Stab_poltiq.
0	Afrique du Sud	57.01	26.78	514.0	24.27	78.71	11.67	6.41	31.09	6121.88	6.90	-0.28
1	Albanie	2.88	-7.83	38.0	80.85	27.66	4.23	1.30	17.54	4514.20	3.92	0.38
2	Algérie	41.39	33.33	2.0	0.72	99.28	2.43	0.80	7.68	4109.70	5.04	-0.92
3	Allemagne	82.66	1.54	842.0	48.42	87.06	9.46	2.06	28.30	44651.83	2.76	0.59
4	Angola	29.82	81.86	277.0	86.83	13.17	4.42	1.85	16.88	4095.81	28.91	-0.33

(166, 12)

Le fichier final comporte 166 Pays et 11 variables d'études dont 6 variables calculées

Voyons comment regrouper ses pays en fonction de leurs similitudes avec les différentes variables : clustering

Standardisation des données

Préalable à l'étude de nos données centrer et réduire celles-ci afin de faciliter leurs comparaisons

#Centrage / réduction des données

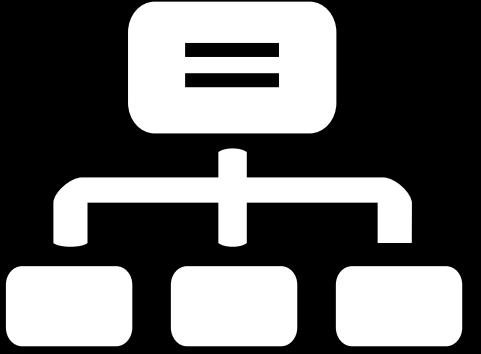
```
std_scale = preprocessing.StandardScaler().fit(X)
```

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

	0	1	2	3	4	5	6	7	8	9	10
mean	0.0	0.0	-0.0	0.0	-0.0	0.0	-0.0	0.0	0.0	0.0	-0.0
std	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0



Nos données sont bien toutes à la même échelle

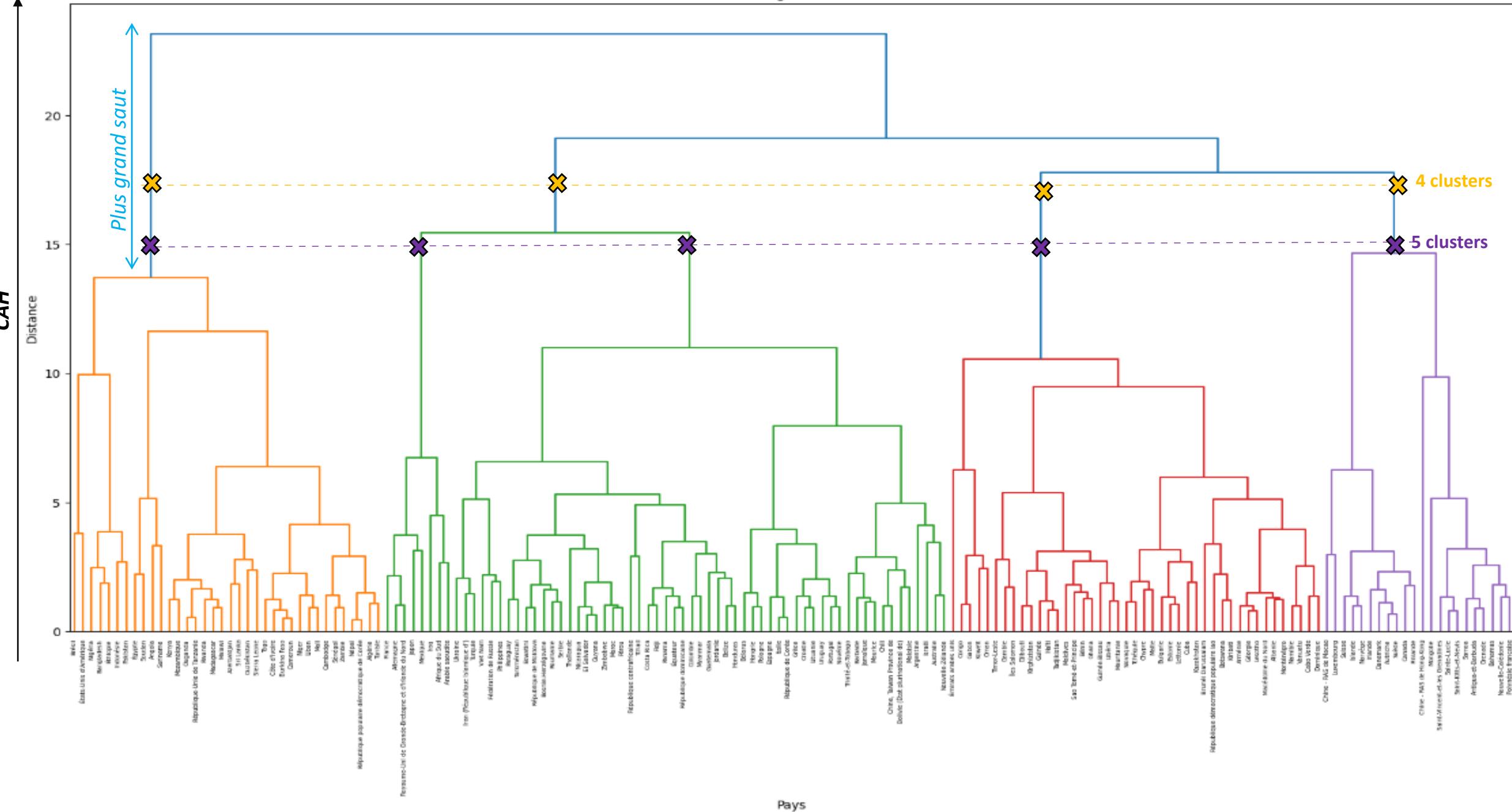


CLUSTERING

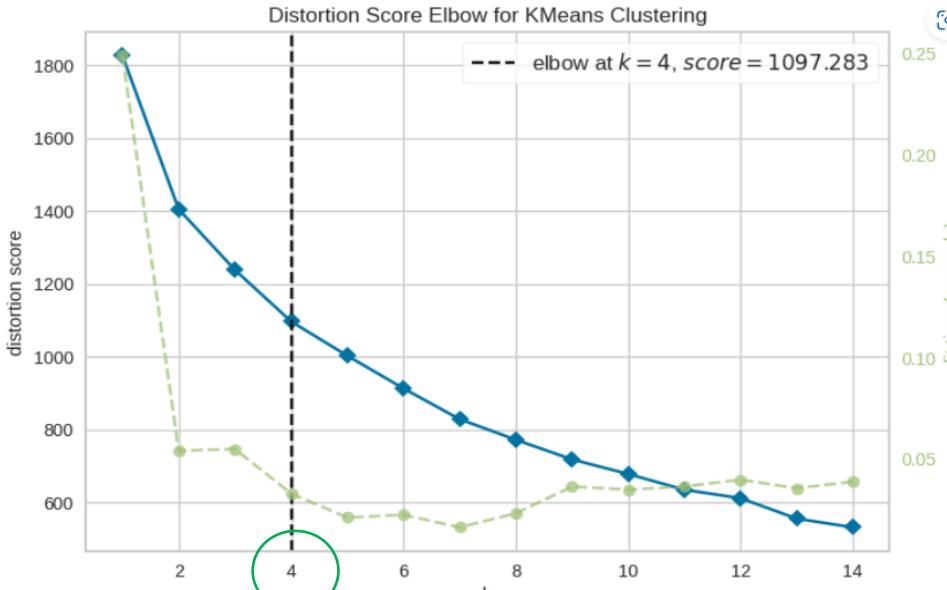


Méthode CAH

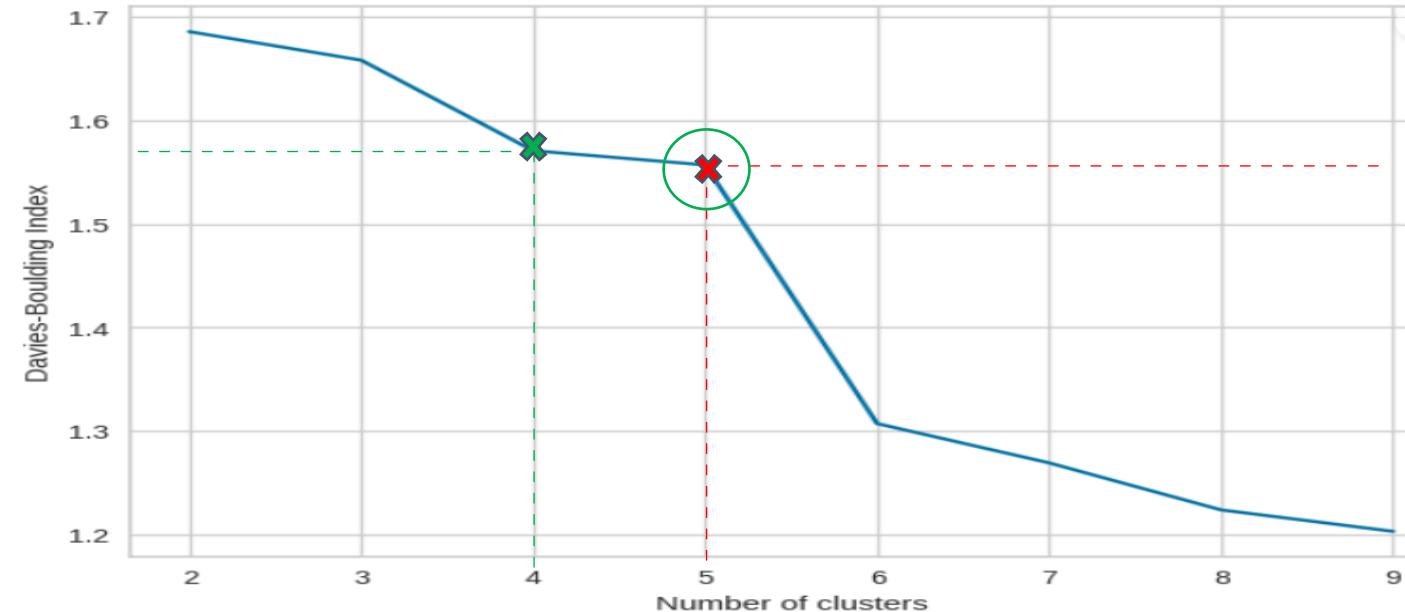
Dendrogram - CA



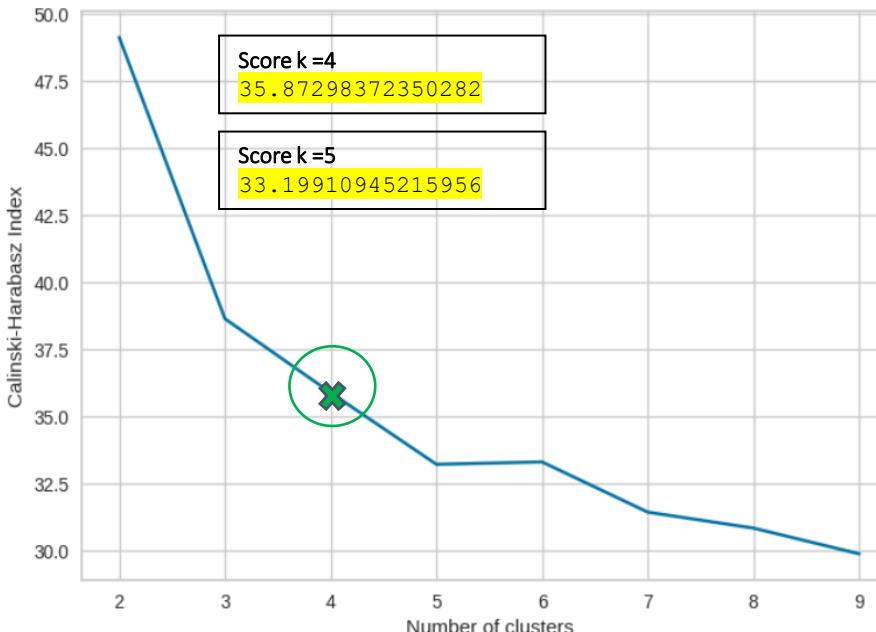
■ Méthode du coude (*Algorithme K.Means)



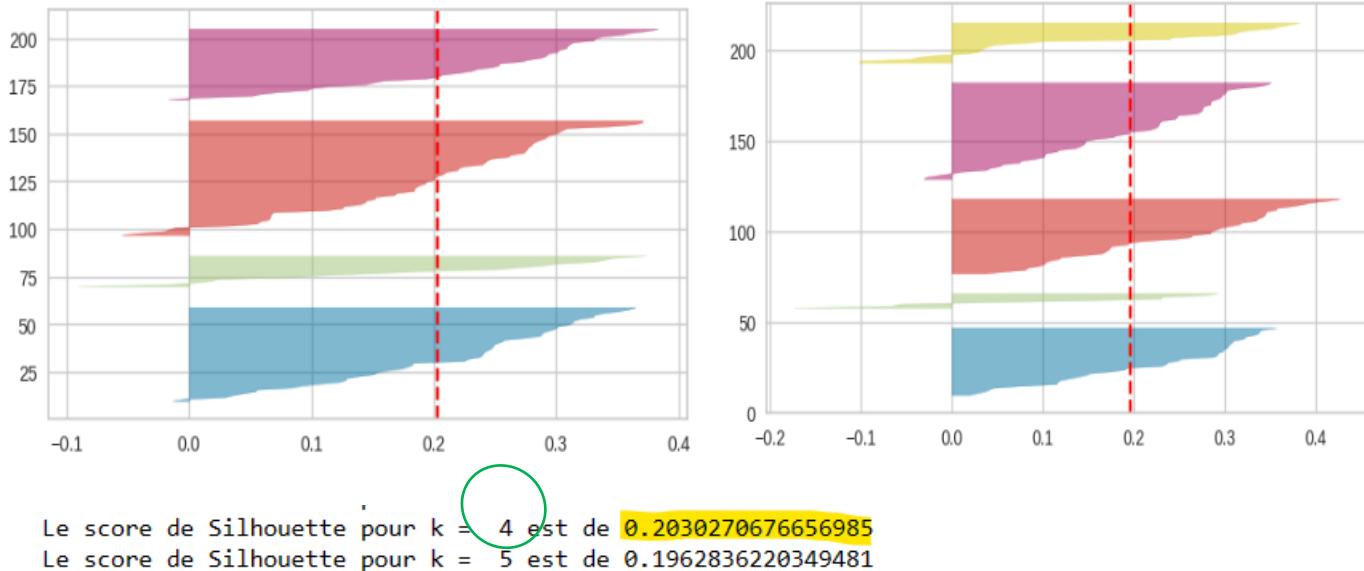
■ Méthode Davies Boulding (*Algorithme K.Means)



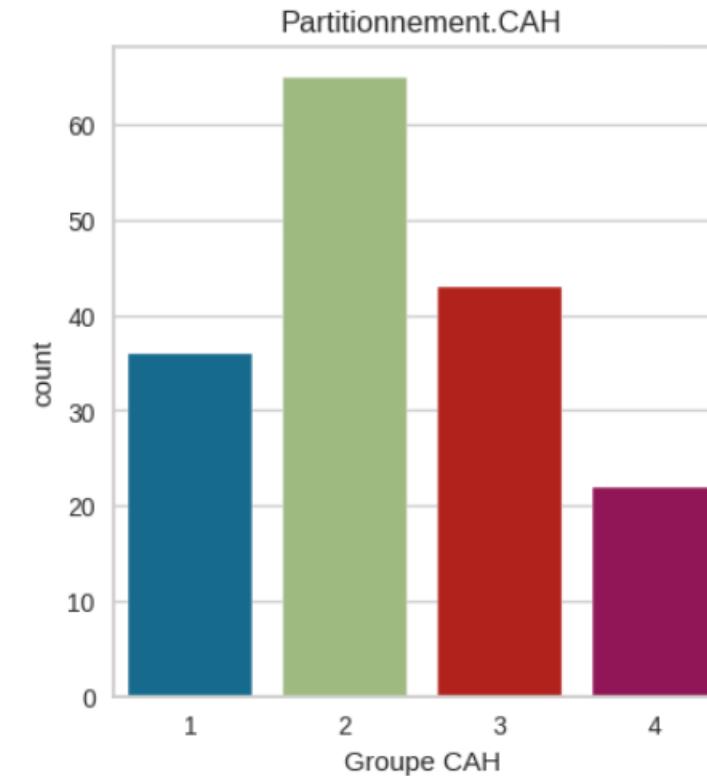
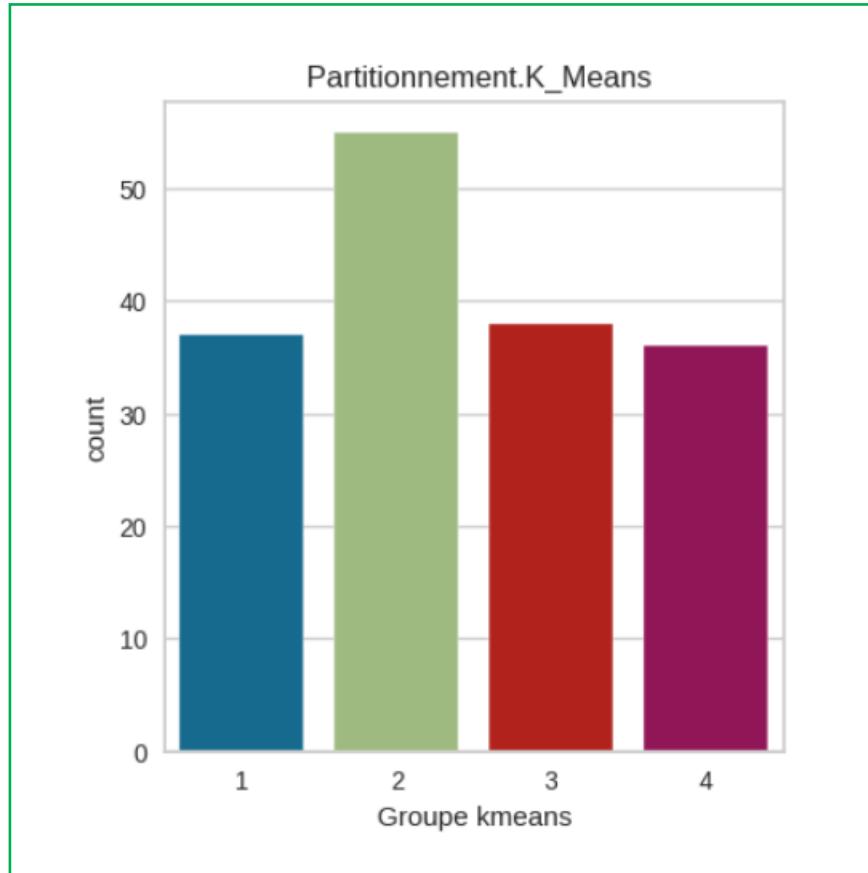
■ Méthode Calinski H  rabsz (*Algorithme K.Means)



■ Méthode du Coef. De Silhouette (*Algorithme K.Means)



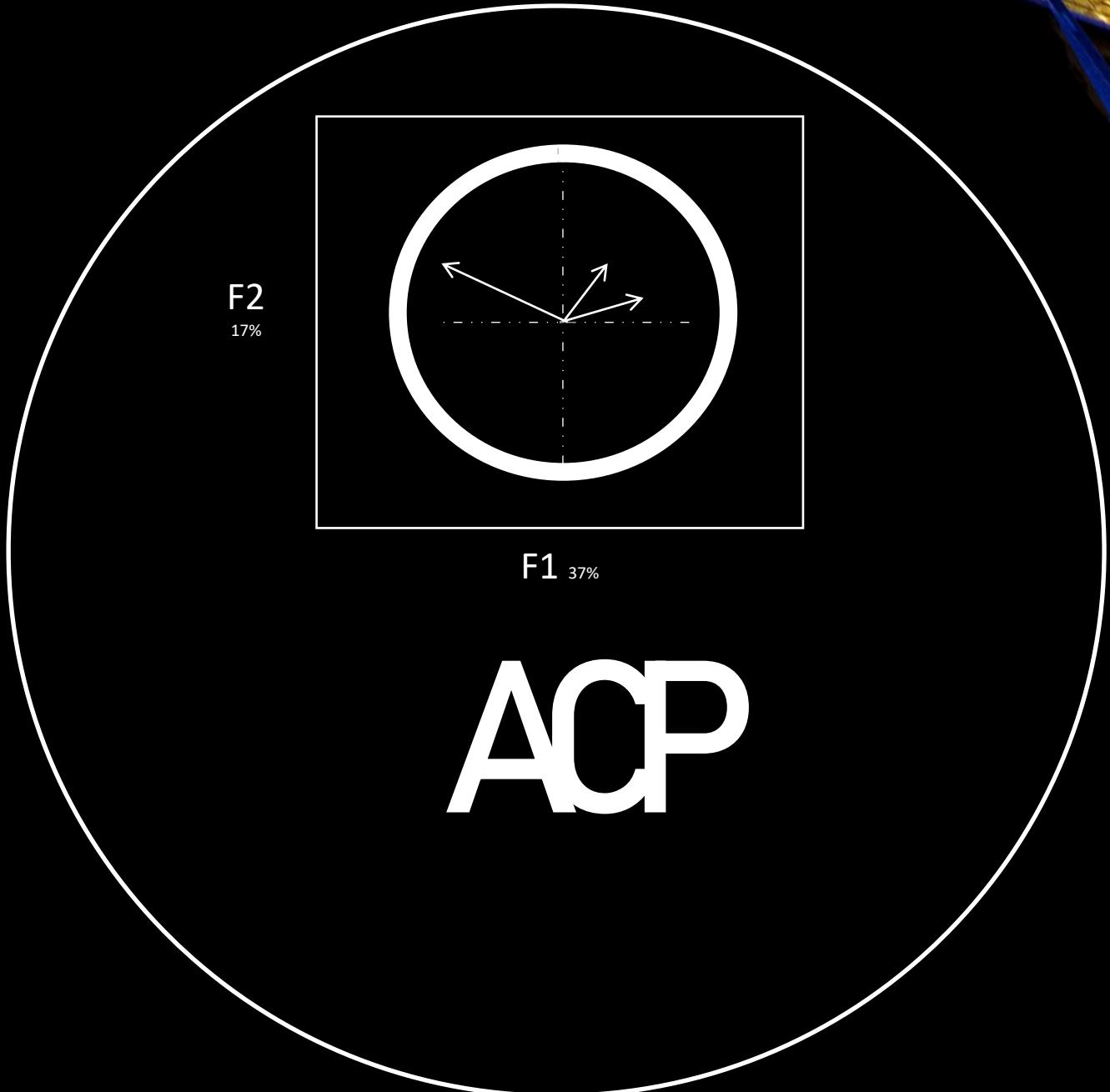
■ Comparaison des 2 partitionnements



Le partitionnement KMeans nous permet d'obtenir une répartition plus homogène. Nous choisissons donc de poursuivre l'étude sur un partitionnement Kmeans

■ Choix de découpage

- 4 Clusters
- Partitionnement K.Means



F2

17%

F1 37%

ACP



Distribution des variables

Test d Anderson (vérification de la distribution dite normale)

Critical values: [0.563 0.641 0.769 0.897 1.067]

Significance levels: [15. 10. 5. 2.5 1.]

A-D statistic for nb_habitant: 23.051000446092445

A-D statistic for Croissance démographique : 2.6800960705048453

A-D statistic for Import_Volaille: 30.21200388994717

A-D statistic for TDI volaille: 7.967812073815111

A-D statistic for dispo alimentaire d'origine animale: 0.9803973105296677

A-D statistic for dispo volaille: 5.315989934698877

A-D statistic for protéines d'origine animale: 0.5018032532552752

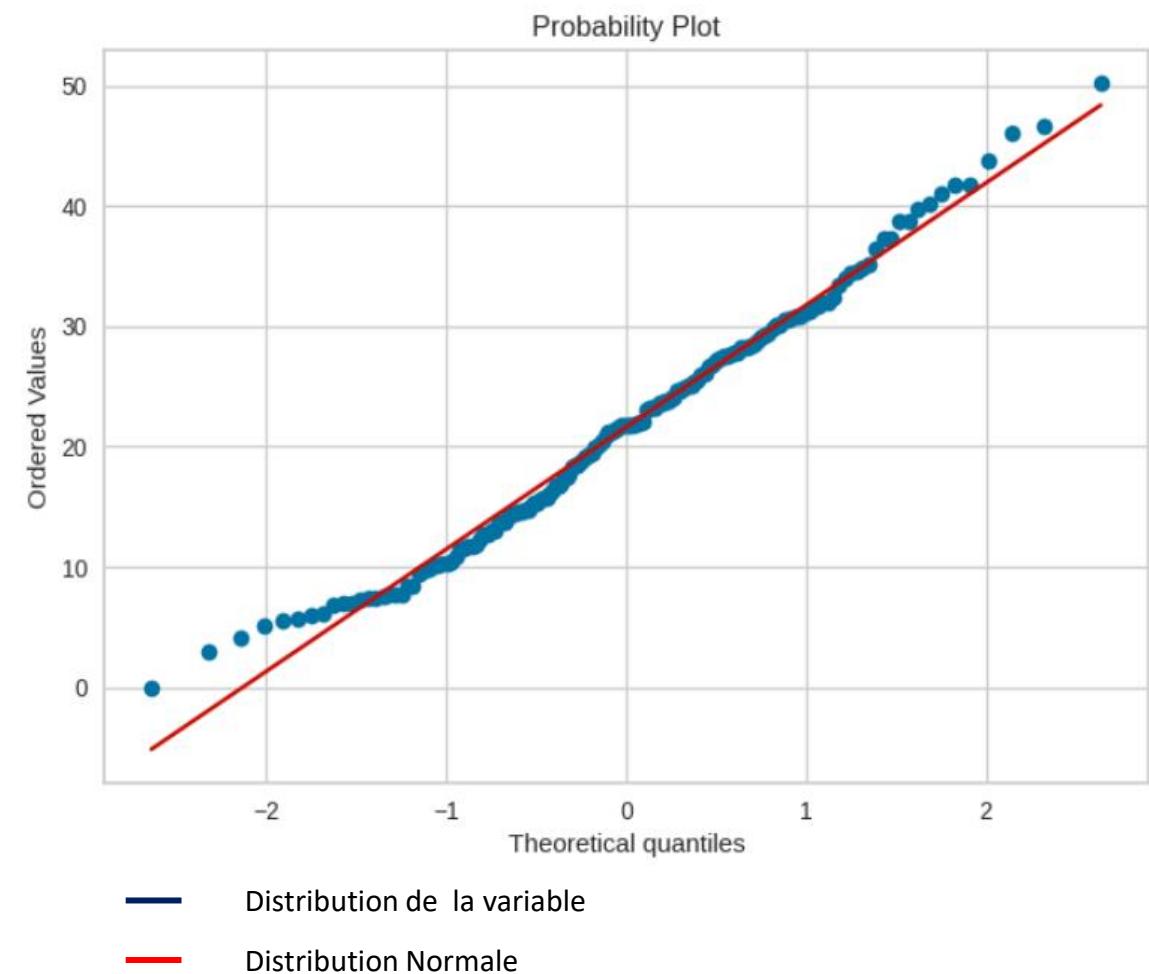
A-D statistic for PIB: 17.415015636480916

A-D statistic for Inflation: 11.446361649042927

A-D statistic for stabilité_politique: 0.9325782618847711

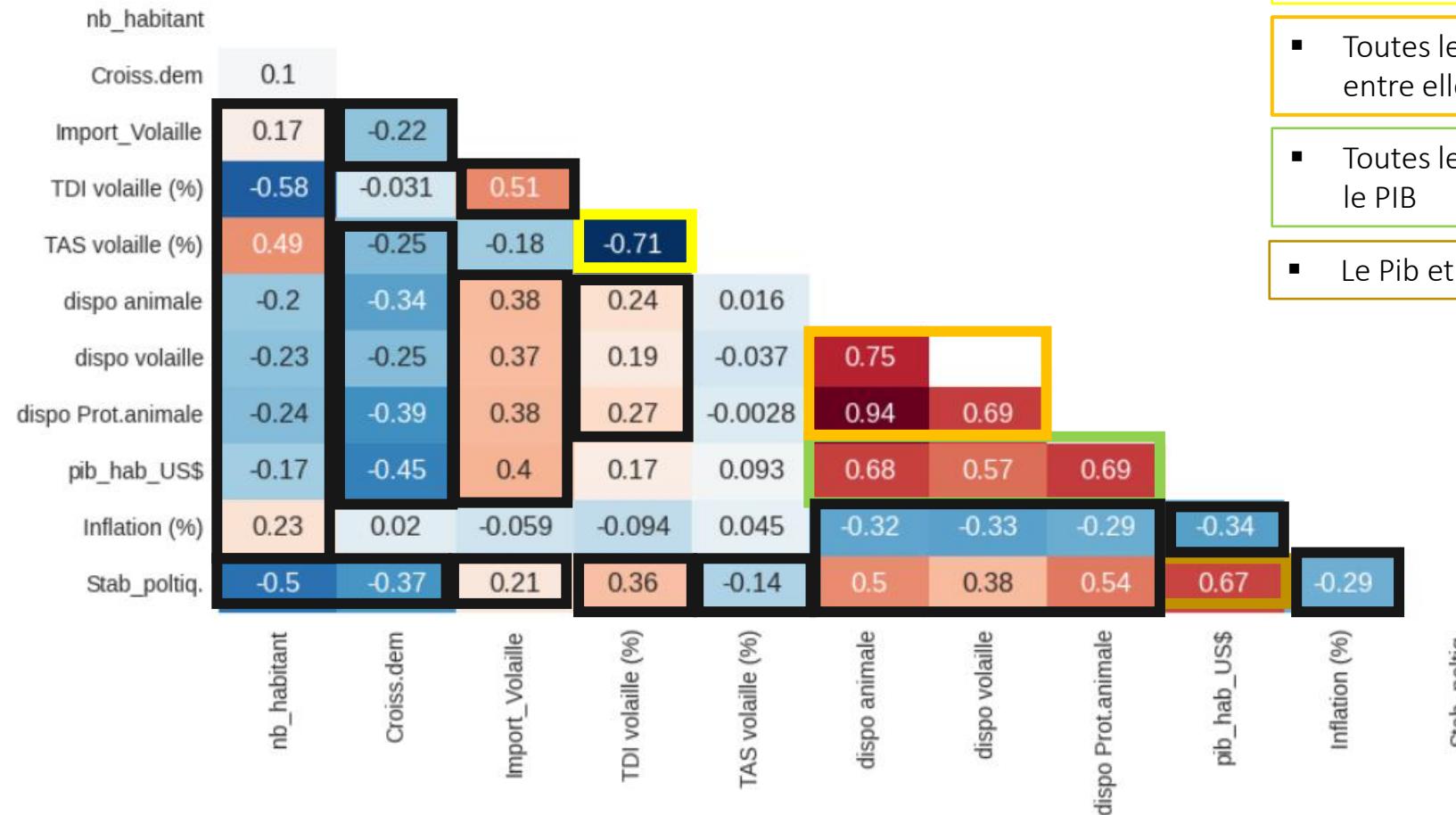
La variable dispo protéine animale ne suit pas une loi normale. Nous utiliserons donc une Heatmap basée sur la corrélation de Spearman (Test non paramétrique)

Distribution protéine d'origine animale



■ TRIANGLE DES CORRELATIONS

■ Triangle corrélation Spearman

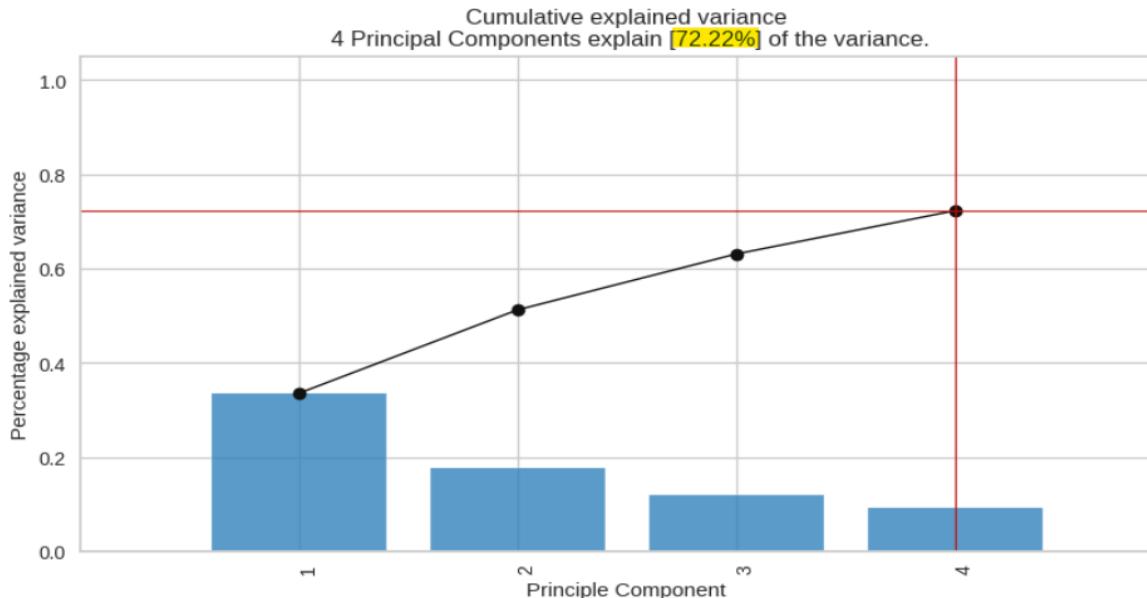
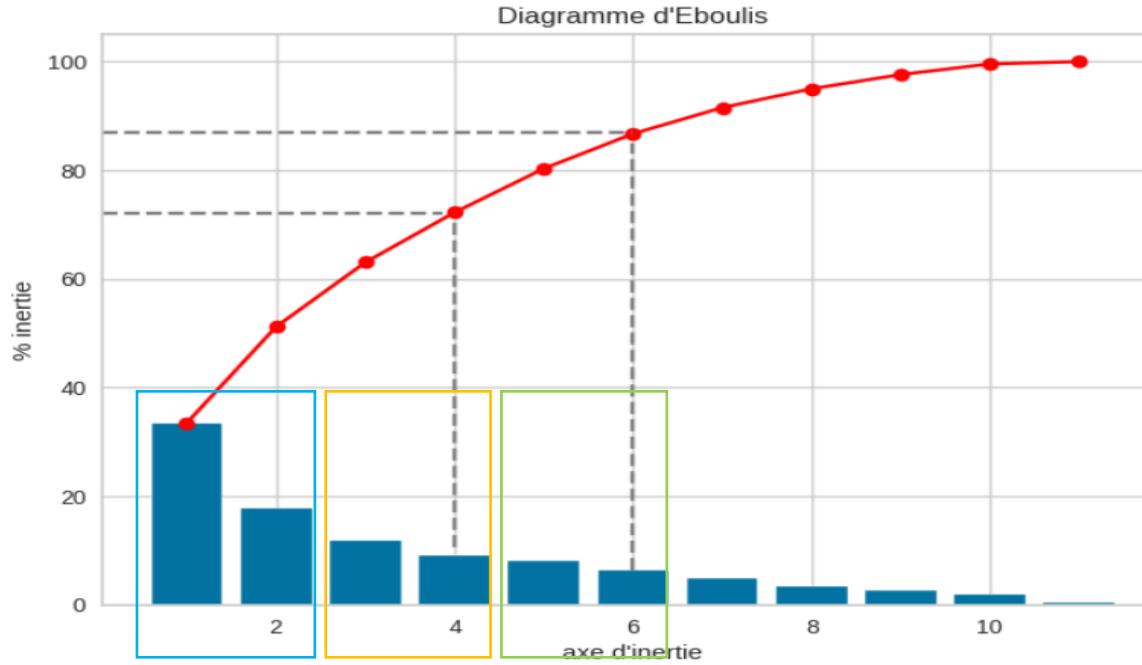


■ Interprétation

Il semblerait que :

- Le TDI et le TAS soient fortement corrélés
- Toutes les disponibilités soient fortement corrélées entre elles
- Toutes les disponibilités soient fortement corrélées avec le PIB
- Le Pib et la stabilité politique soient fortement corrélées

■ Diagramme Eboulis



■ Ratio de variance

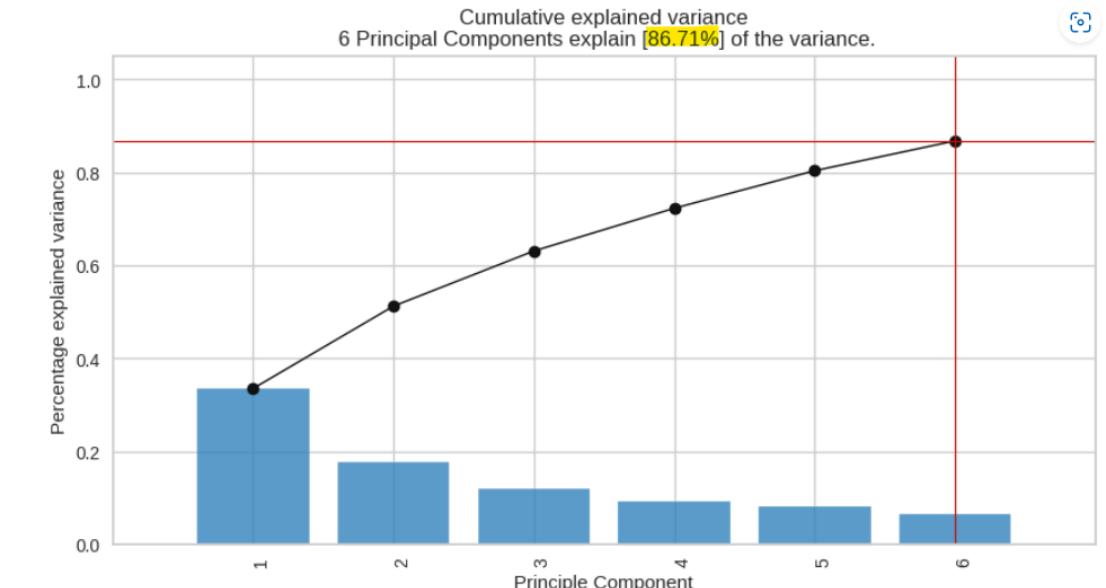
```
pca.explained_variance_ratio_
```

```
array([0.33466049, 0.17710023, 0.11862771, 0.09189025, 0.08073542,
       0.06417661, 0.0481892 , 0.03458319, 0.02619867, 0.01955787,
       0.00428036])
```

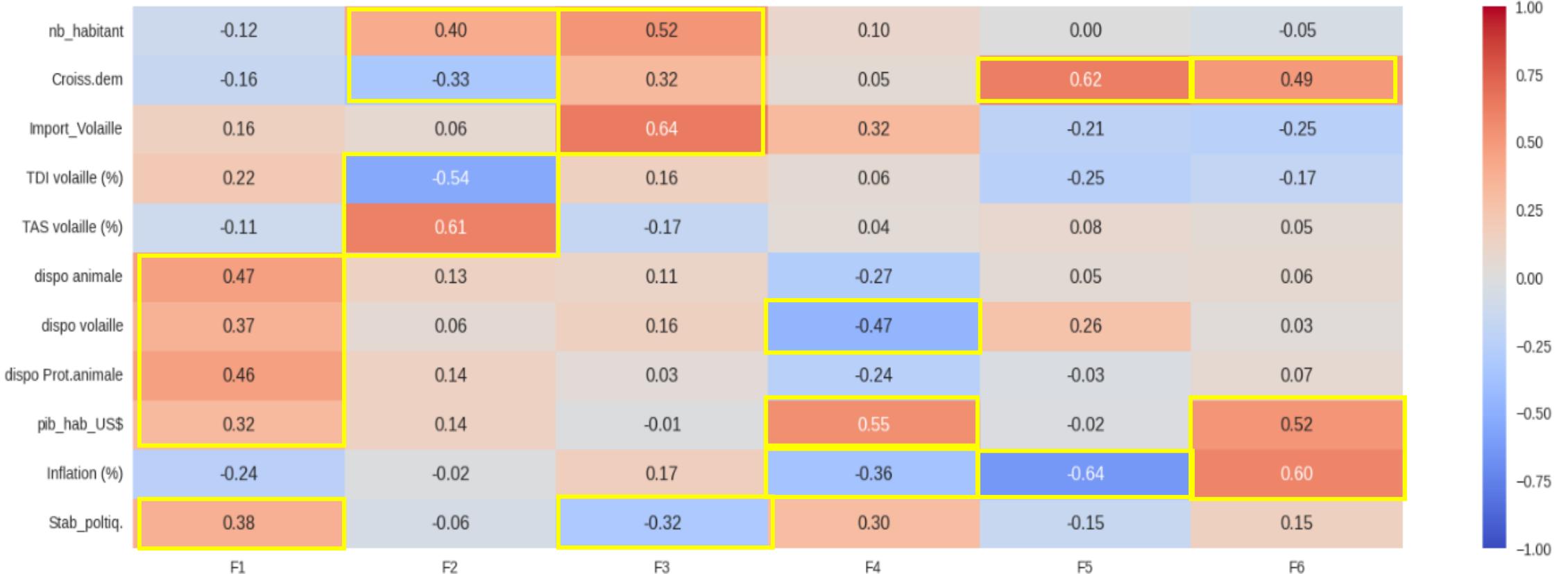
■ Ratio de variance cumulée

```
pca.explained_variance_ratio_.cumsum()
```

```
array([0.33466049, 0.51176072, 0.63038843, 0.72227868, 0.8030141 ,
       0.86719071, 0.91537991, 0.9499631 , 0.97616177, 0.99571964,
       1.])
```



■ HEATMAP COMPOSANTES PRINCIPALES



■ INTERPRETATION

F1 les variables suivantes lui soient fortement corrélées:

- Toutes les disponibilités le Pib et la stabilité politique (positivement)
- Inflation (négativement)

F2 les variables suivantes lui soient fortement corrélées:

- TAS volaille et Nombre d habitant (positivement)
- TDI volaille et croissance démographique (négativement)

F3 les variables suivantes lui soient fortement corrélées:

- Import de volaille Croissance démographique et nombre d habitant (positivement)
- Stabilité politique (négativement)

F4 les variables suivantes lui soient fortement corrélées:

- PIB Stabilité politique et import de volaille (positivement)
- Toutes les disponibilité (négativement)

F5 les variables suivantes lui soient fortement corrélées:

- Inflation (négativement)
- Croissance démographique (positivement)

F6 les variables suivantes lui soient fortement corrélées:

- Inflation Pib et Croissance démographique (positivement)
- Import de volaille (négativement)

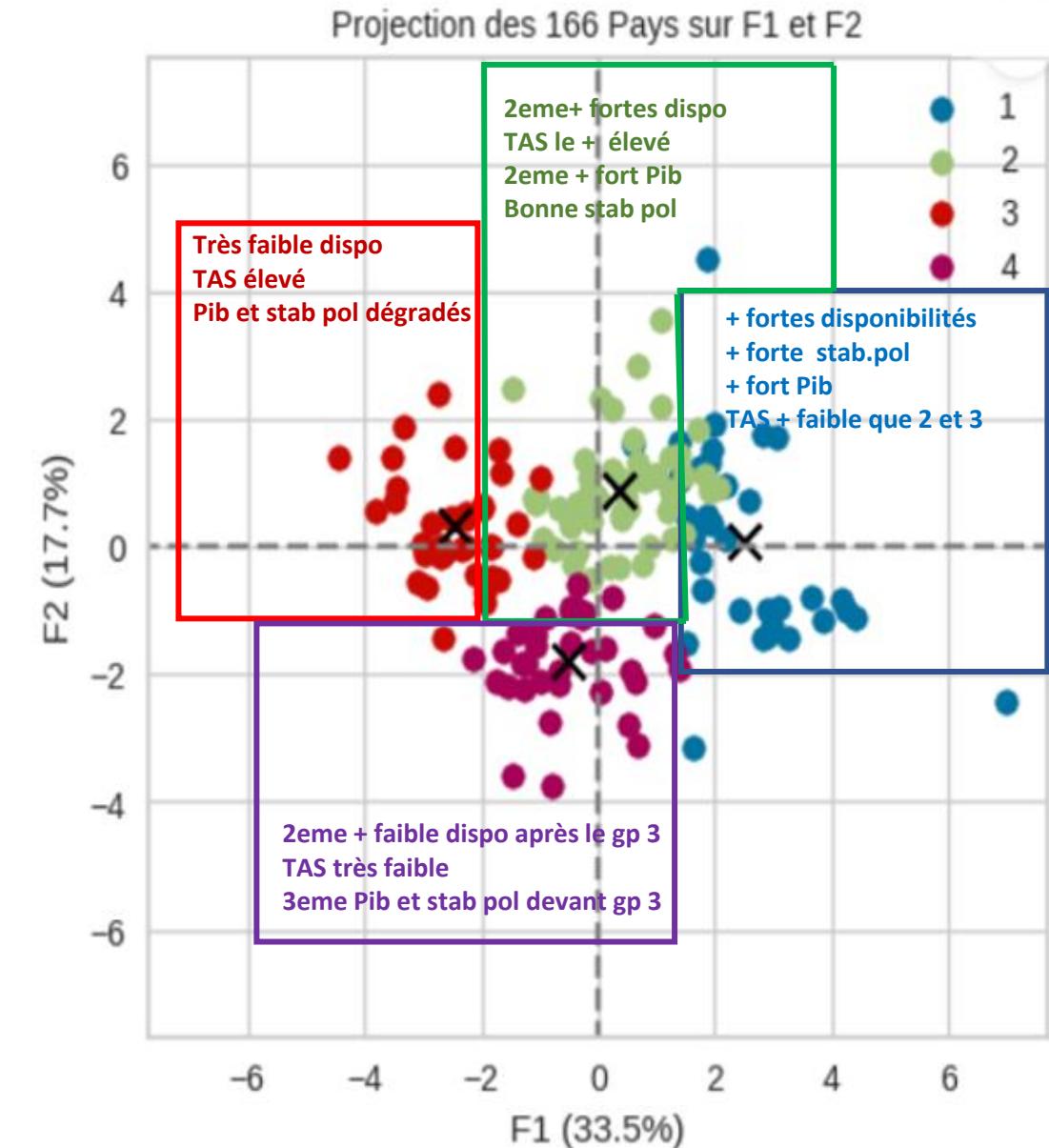
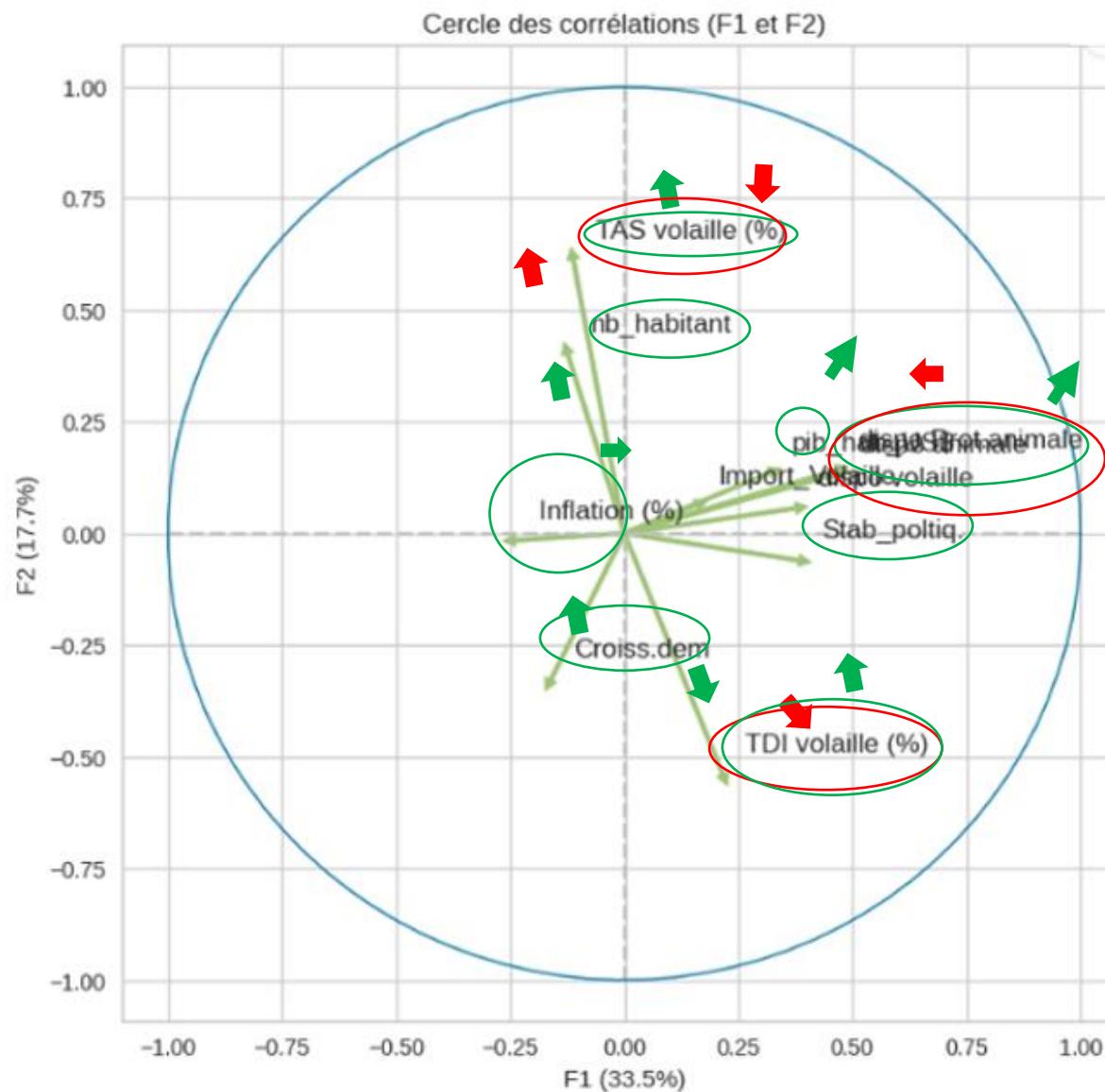


ANALYSE DES CLUSTERS

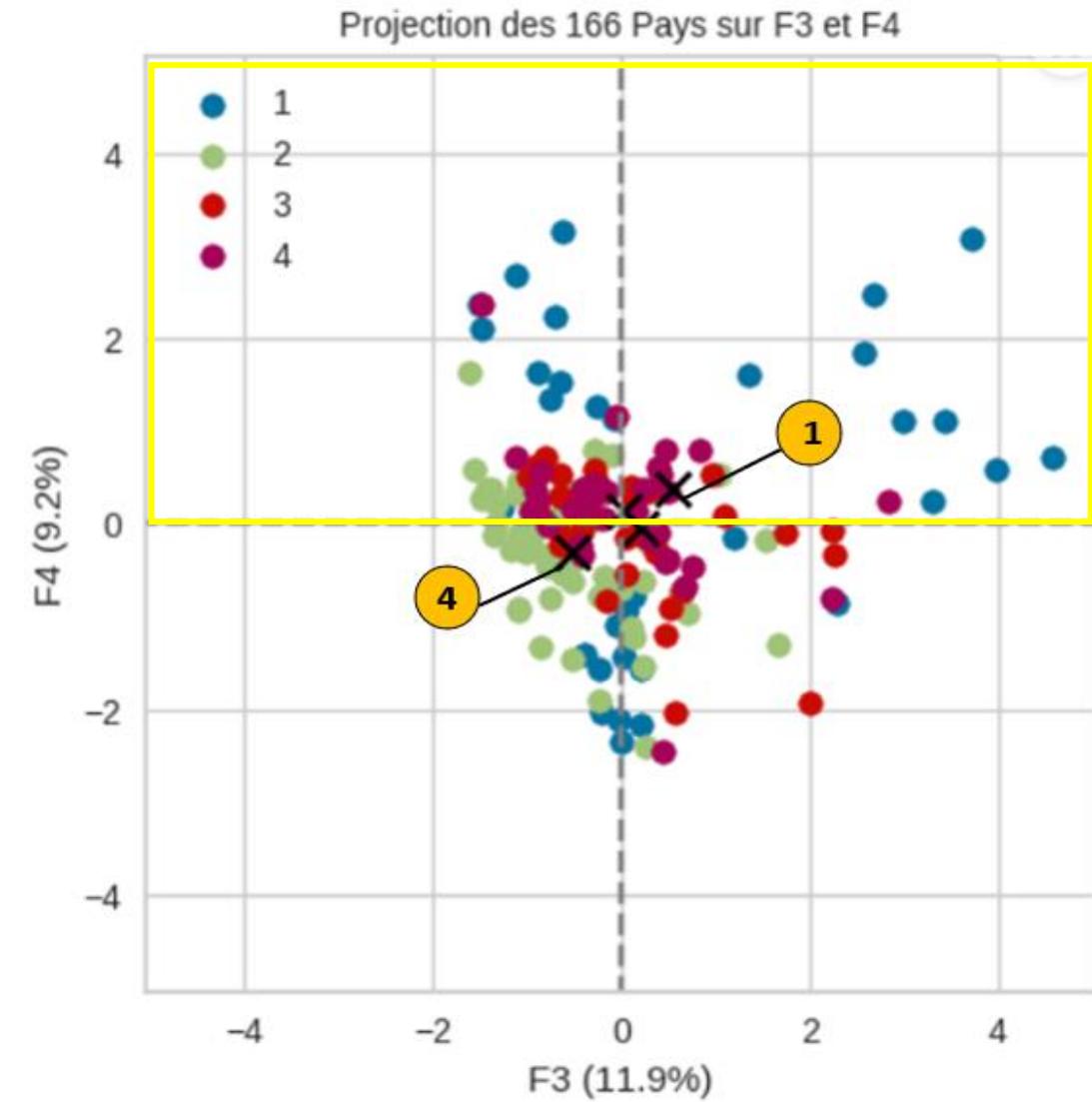
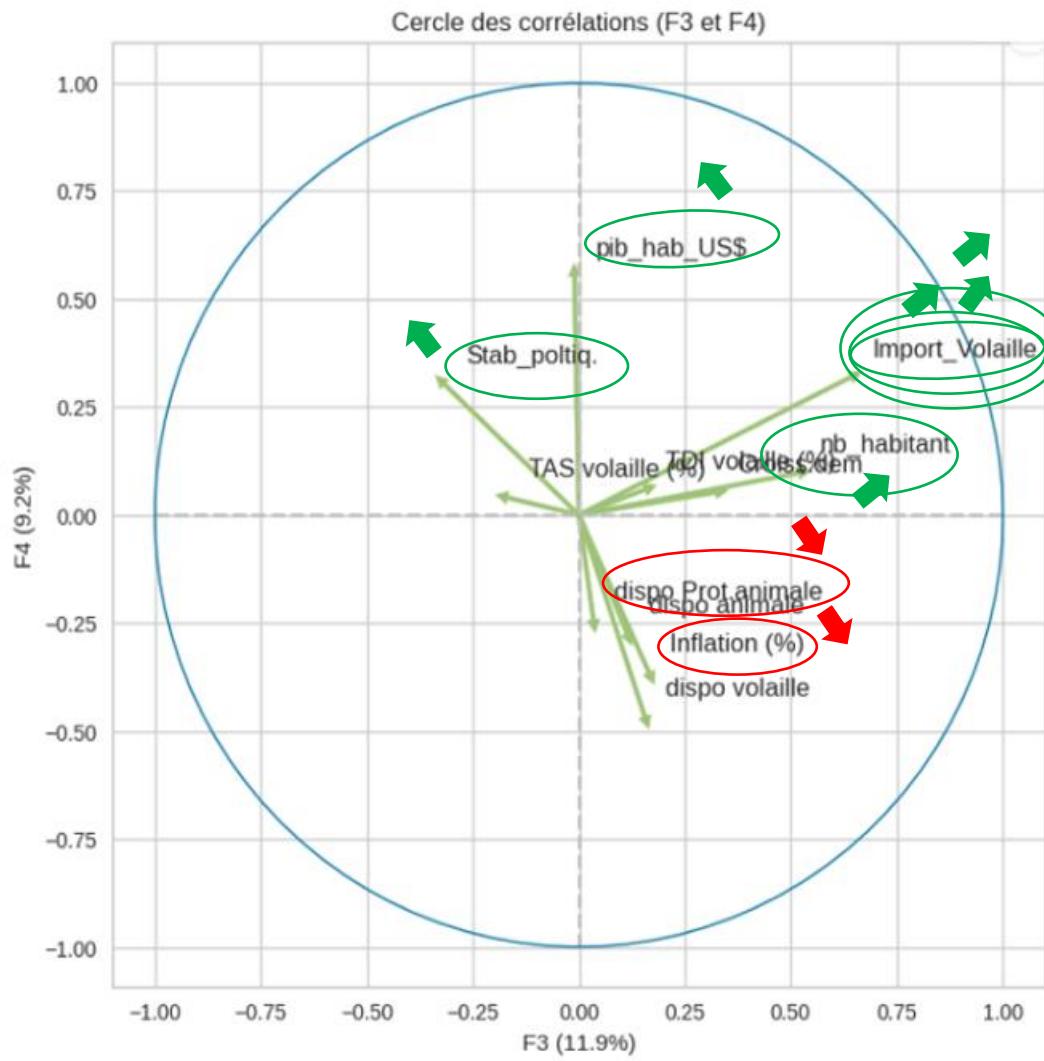


■ Projection sur les plans factoriels

- Cercles de corrélations et projection des clusters sur F1 F2

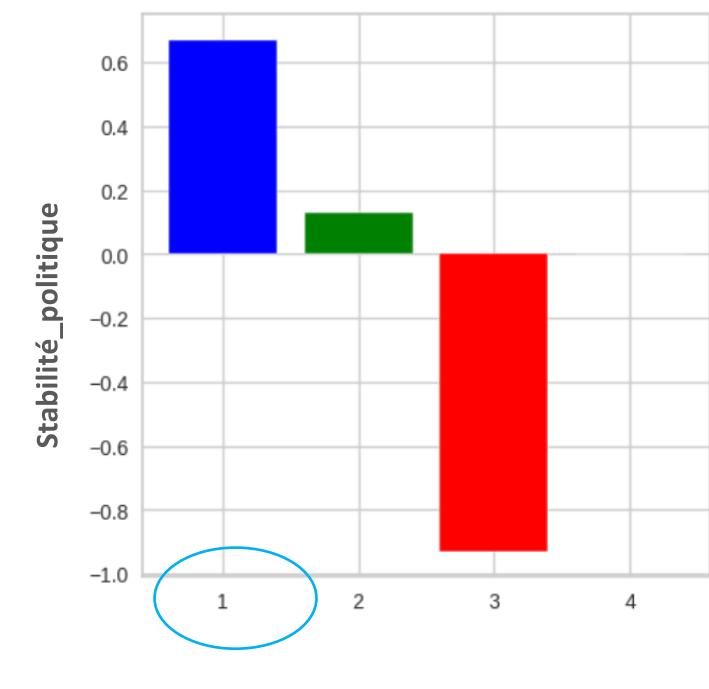
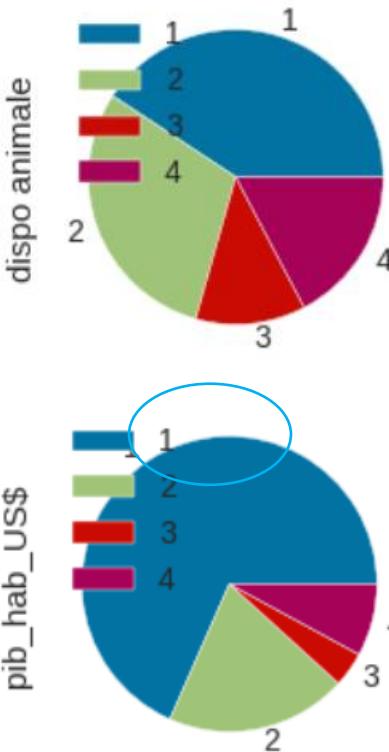
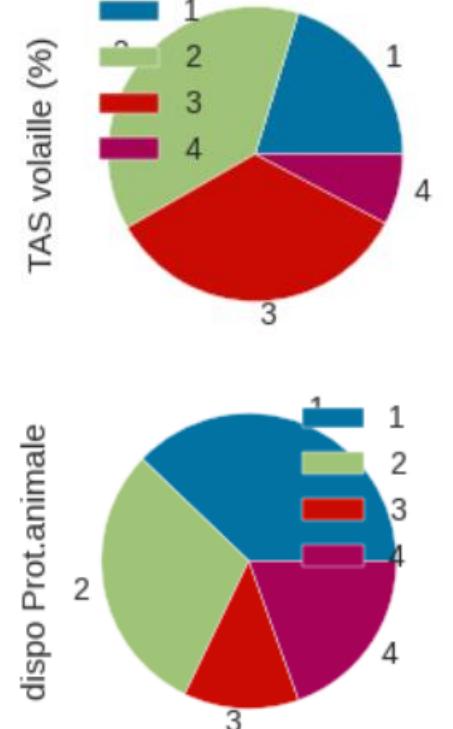
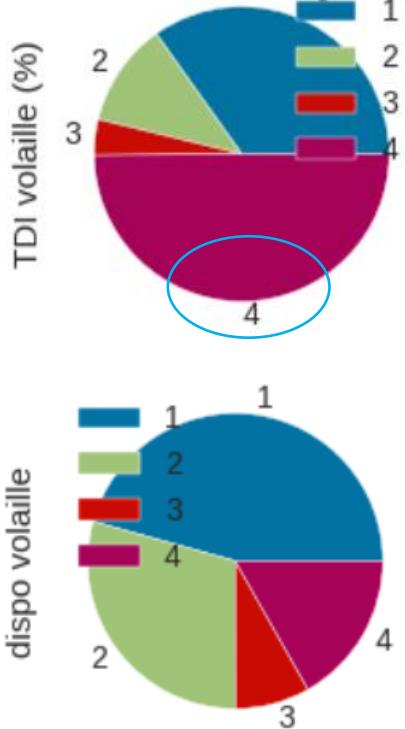
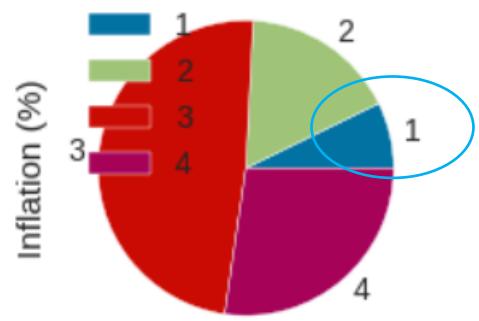
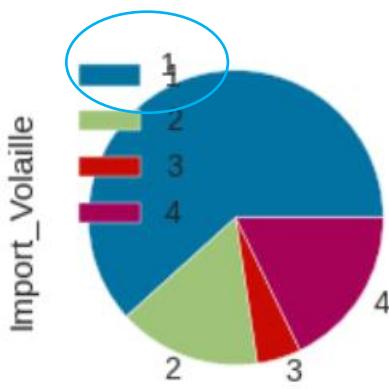
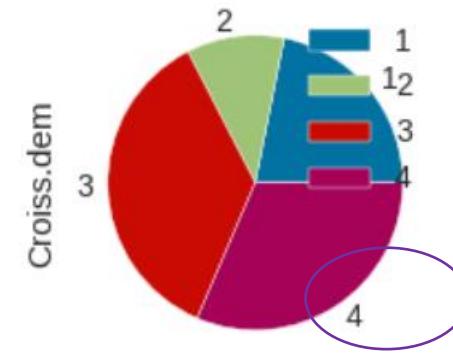
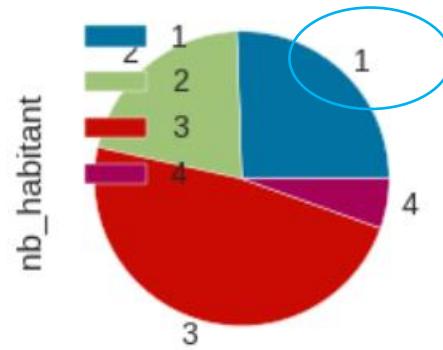


- Cercles de corrélations et projection des clusters sur F3 F4



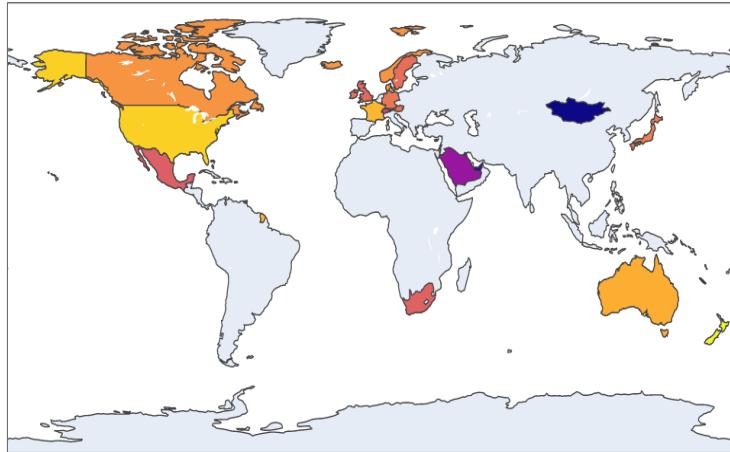
Moyenne des distributions

Pie chart moyenne des groupes par variables



Situation géographique

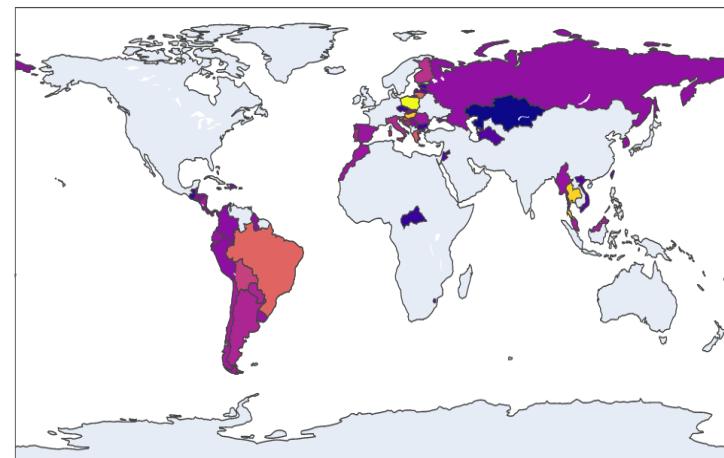
Situation géographique des pays du cluster n°1



Cluster n°1:

- 12 pays d'Europe
- 10 pays d'Amérique du Nord
- 9 pays d'Asie
- 5 pays Océanie
- 1 pays d'Afrique

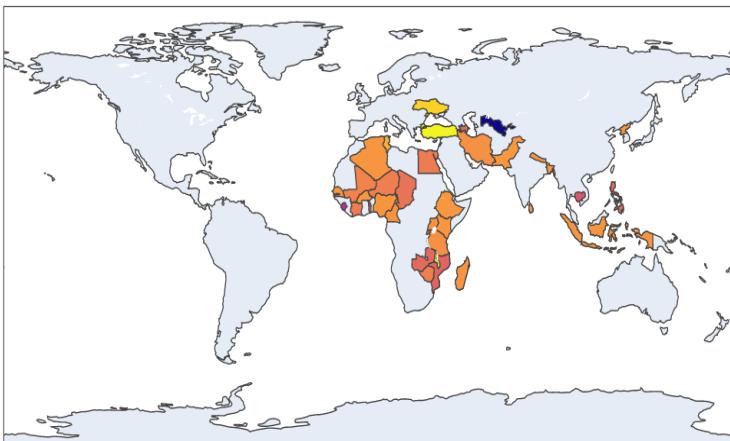
Situation géographique des pays du cluster n°2



Cluster n°2:

- 20 pays d'Europe
- 11 pays d'Amérique du Nord
- 10 pays d'Amérique du Sud
- 9 pays d'Asie
- 4 pays d'Afrique
- 1 pays Océanie

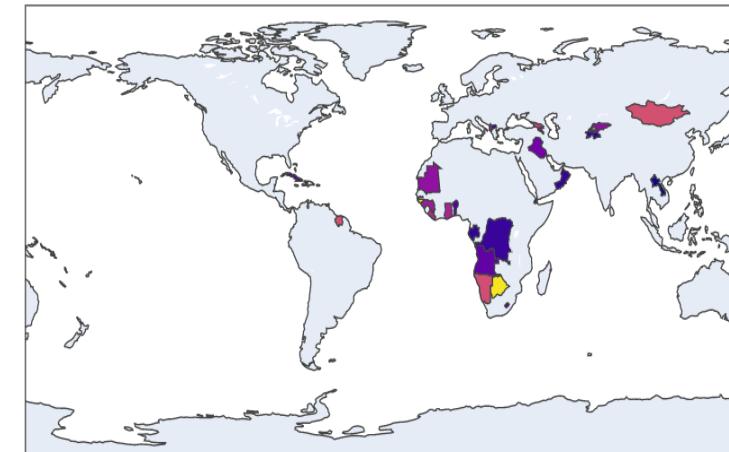
Situation géographique des pays du cluster n°3



Cluster n°3:

- 23 pays d'Afrique
- 14 pays d'Asie
- 1 pays d'Europe

Situation géographique des pays du cluster n°4



Cluster n°4:

- 16 pays d'Afrique
- 11 pays d'Asie
- 3 pays d'Amérique du Nord
- 1 pays d'Amérique du Sud
- 2 pays d'Europe
- 3 pays Océanie





SYNTHESE



■ Le Groupe n°2 : exclu de la selection

Le groupe 2 est complètement autonome dans la gestion de ses disponibilités animale le groupe priviléie la production plutôt que les importations. La quantité de volaille importée est plus faible que les groupes 1 et 2.

■ Le Groupe n°3 : exclu de la selection

Le groupe a les disponibilités les plus faibles le régime alimentaire est faible en protéine animale. Le Groupe est en autonomie et importe très peu de volaille. Le Pib et la stabilité politique sont dégradés celà a un impact sur les disponibilités qui s'amenuisent. De même l inflation forte a également un impact sur le niveau des disponibilités.

Les pays concernés par ce groupe sont majoritairement constitués de pays africains à hauteur de 61% et de pays asiatiques pour 37%.

■ Les groupes 1 et 4 sont les groupes qui offrent le plus de potentiel marché en effet la nécessité d importée est avérée (TAS faible TDI élevé) cependant le groupe 4 est marqué par une forte inflation et un Pib très faible ce qui peu avoir un impact sur la consommation et freiner les achats. Notre choix se portera donc sur le Groupe n°1 qui dispose d une inflation maîtrisée d une excellente stabilité politique et d un très bon niveau de PIB. De plus la quantité de volaille importée par ce groupe est 3.5fois plus élevée que celle du groupe 4

■ Choix du cluster n°1

- Afrique du Sud
- Allemagne
- Antigua-et-Barbuda
- Arabie saoudite
- Australie
- Autriche
- Bahamas
- Canada
- Chine - RAS de Hong-Kong
- Chine - RAS de Macao
- Chypre
- Danemark
- Émirats arabes unis
- États-Unis d Amérique
- Grenade
- Irlande
- Islande
- Israël
- Japon
- Koweït
- Luxembourg
- Malte
- Mexique
- Suisse
- Mongolie
- Norvège
- Nouvelle-Calédonie
- Nouvelle Zélande
- Polynésie française
- Royaume-Uni
- Sainte-Lucie
- Saint-Kitts-et-Nevis
- Saint-Vincent-et-les Grenadines
- Samoa
- Suède
- Trinité-et-Tobago

	Zone	Code alpha-3	Capitale	Continent	nb_habitant	Croiss.dem	Import_Volaille	TDI volaille (%)	TAS volaille (%)
78	Japon	JPN	Tokyo	Asie	127.50	-0.02	1069.0	44.27	91.72
101	Mexique	MEX	Mexico	Amérique du Nord	124.78	26.17	972.0	23.04	77.01
30	Chine - RAS de Hong-Kong	HKG	-	Asie	7.31	10.60	907.0	323.93	8.57
3	Allemagne	DEU	Berlin	Europe	82.66	1.54	842.0	48.42	87.06
133	Royaume-Uni de Grande-Bretagne et d'Irlande du...	GBR	Londres	Europe	66.73	13.24	779.0	34.87	81.20
6	Arabie saoudite	SAU	Riyad	Asie	33.10	60.19	722.0	50.31	42.93
0	Afrique du Sud	ZAF	Pretoria	Afrique	57.01	26.78	514.0	24.27	78.71
55	France	FRA	Paris	Europe	64.84	9.87	506.0	32.17	111.25
45	Émirats arabes unis	ARE	Abu Dhabi	Asie	9.49	202.71	433.0	105.10	11.65
28	Canada	CAN	Ottawa	Amérique du Nord	36.73	20.09	182.0	12.70	98.88