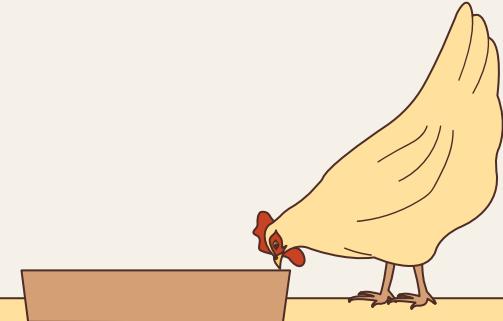
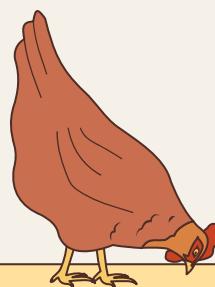
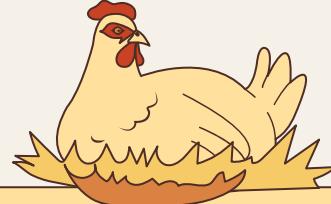
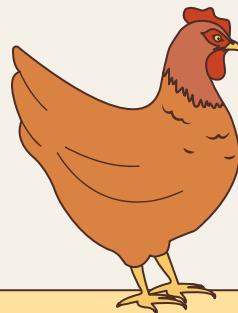


LA POULE QUI CHANTE

ETUDE DE MARCHE



LA POULE QUI CHANTE

Entreprise agroalimentaire française spécialisée dans l'élevage et la vente de poulet.

Son objectif est désormais de s'étendre à l'international.



LA POULE QUI CHANTE - LES DONNEES

Données FAO

**Population
De 2000
À 2017**

Informations fournies par Food and Agriculture Association

**Disponibilités
Alimentaires
2017**

**Stabilité
Politique
2017**

Evalue l'impact des facteurs politiques, économiques, socioculturels, technologiques, environnementaux et légaux sur une entreprise.

Analyse PESTEL

**PIB
Par habitant
2017**



PRE-TRAITEMENT DES DONNEES

1

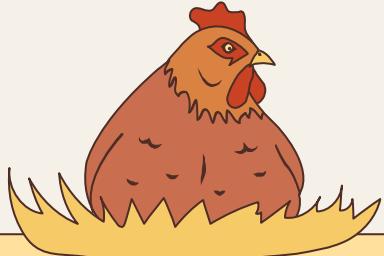
Identification et gestion des
valeurs manquantes

2

Identification et gestion des
valeurs atypiques

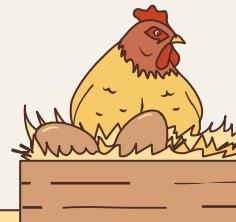
3

Fusion des données



**8 variables PESTEL
151 Pays soit 76% de la population mondiale**

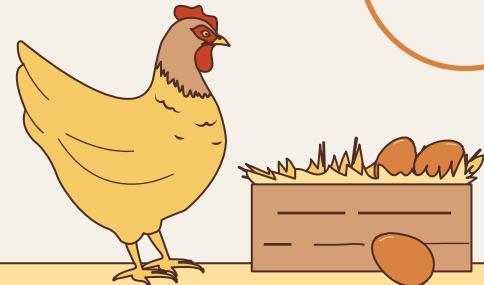
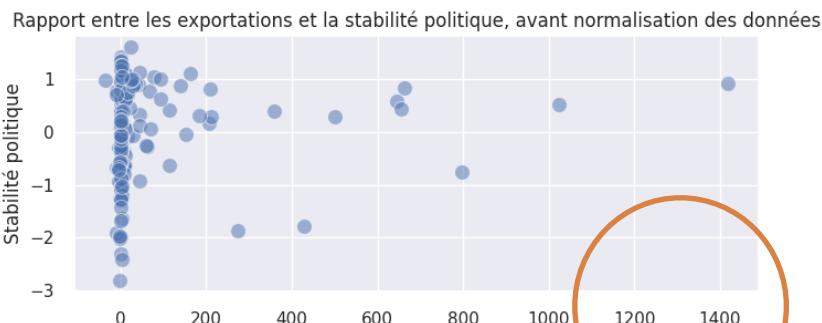
ANALYSE CLUSTERING VISUALISATIONS



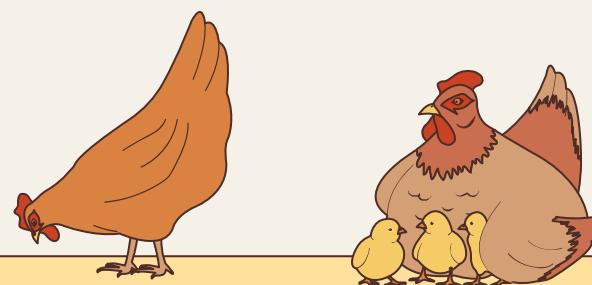
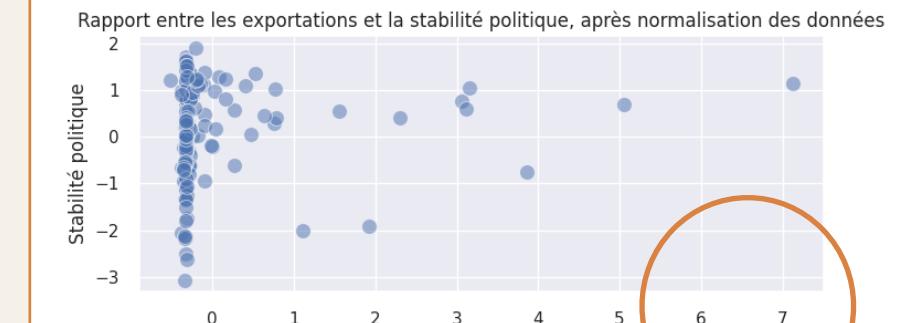
NORMALISATION DES DONNEES

On standardise les valeurs pour obtenir une moyenne à 0 et un écart type à 1.
Ce processus est appelé " Scaling ", ou mise à l'échelle en français.

AVANT



APRES



CLASSIFICATION ASCENDANTE HIERARCHIQUE

01

Première étape

On considère chaque individu de notre échantillon comme un cluster à part entière.

02

Deuxième étape

On identifie puis regroupe les clusters dit « les plus proches » par deux, afin d'en créer de nouveaux.

03

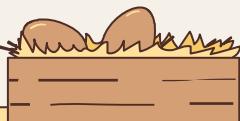
Methode Ward

Ici, on utilise la méthode « *ward* » qui tente de minimiser l'augmentation de la variance totale.

04

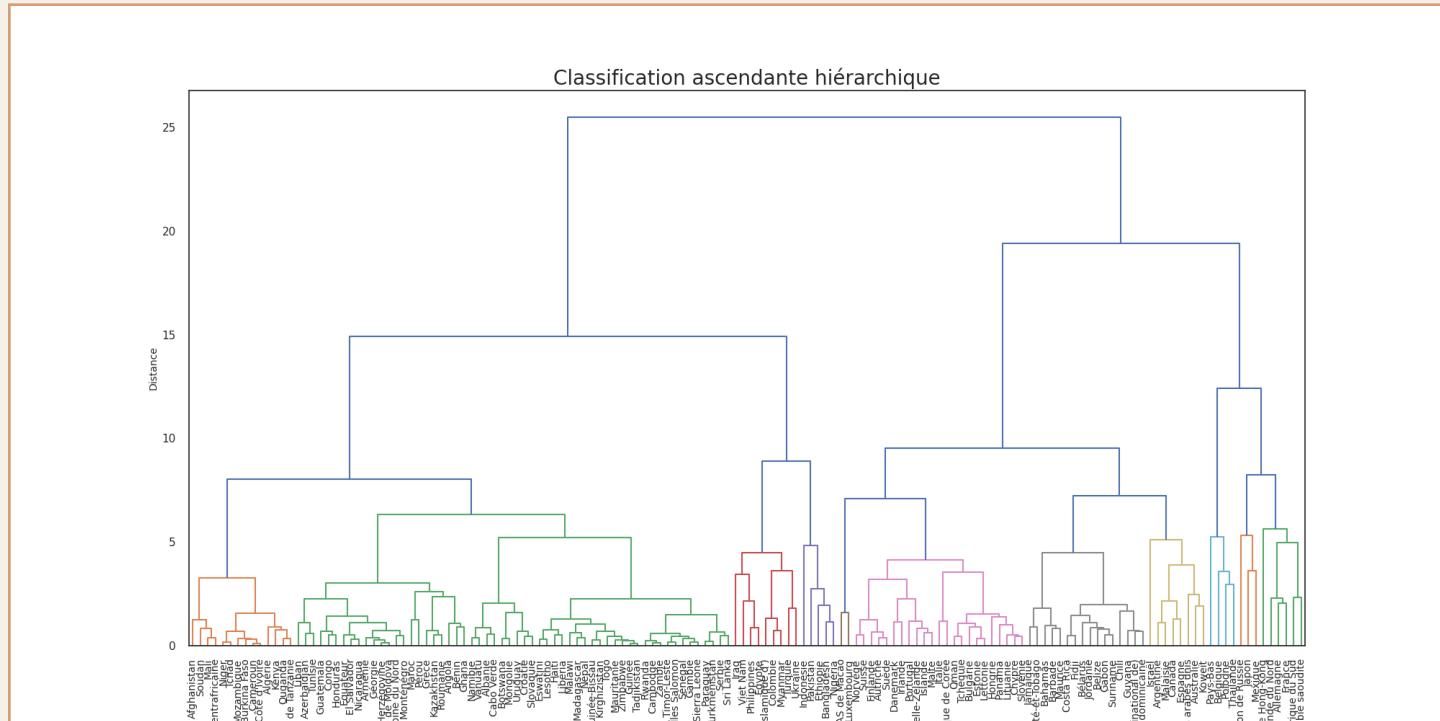
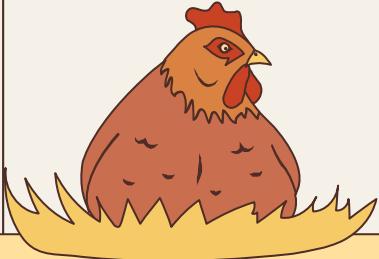
Dernière étape

On recommence l'opération en boucle afin d'obtenir qu'un seul cluster englobant l'intégralité de nos données.

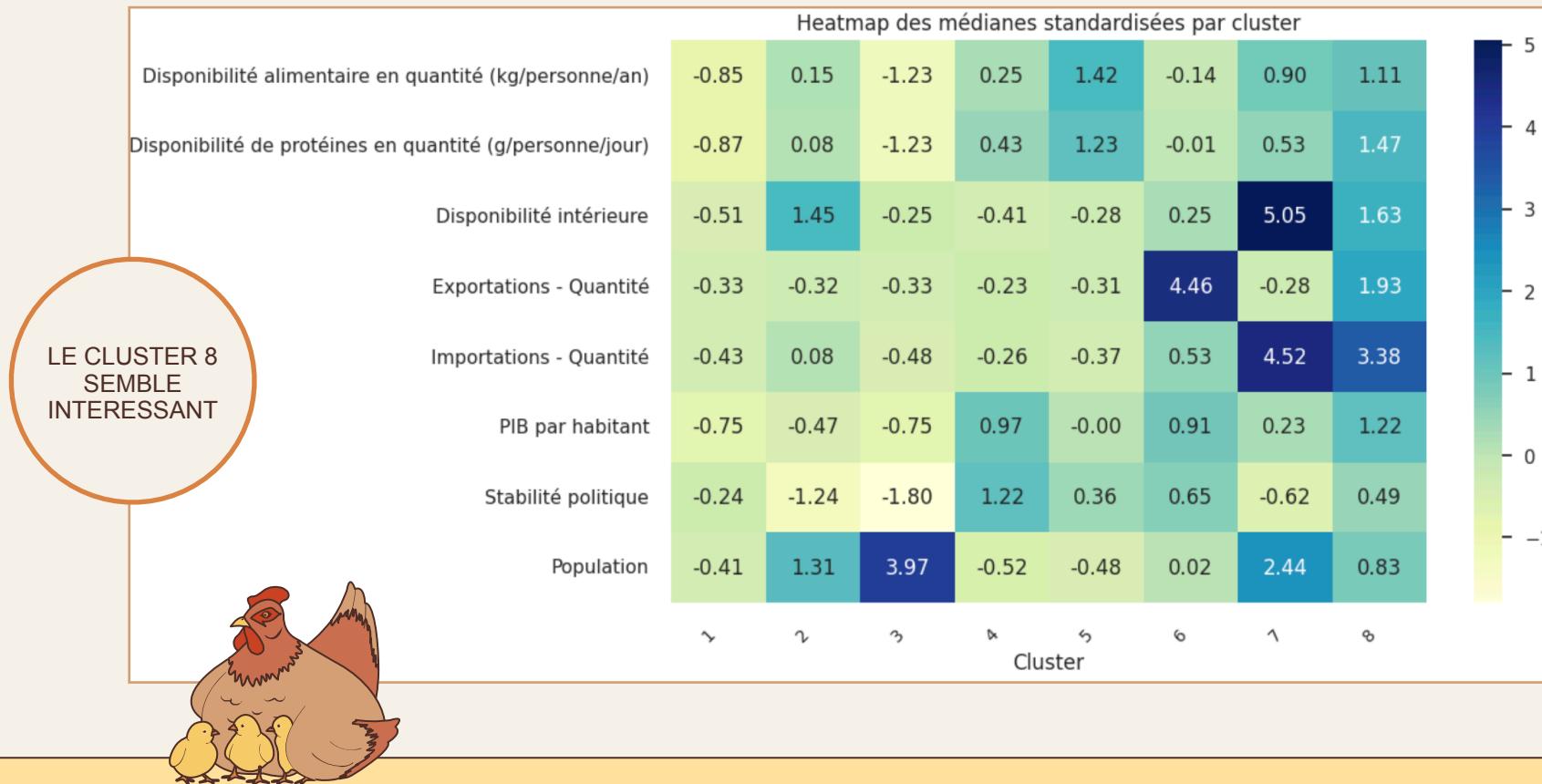


CLASSIFICATION ASCENDANTE HIERARCHIQUE

IDENTIFICATION
DE 8 GROUPES
POTENTIELS

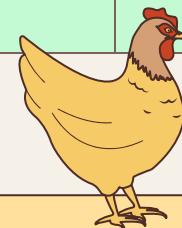


ANALYSE DES CLUSTERS OBTENUS



CLUSTERS CLASSIFICATION HIERARCHIQUE

	PIB	STABILITE POLITIQUE	IMPORTATIONS	DISPONIBILITE PROTEINE
AFRIQUE DU SUD	\$13 850	-0,28	514	14g/Jour
ALLEMAGNE	\$53 071	0,59	842	7g/Jour
ARABIE SAHOUDITE	\$74 306	-0,64	722	15g/Jour
HONG-KONG	\$59 842	0,83	907	22g/Jour
FRANCE	\$44 577	0.28	506	8g/Jour
ROYAUME-UNIS	\$46 372	0.39	779	13g/Jour



ANALYSE KMEANS

01

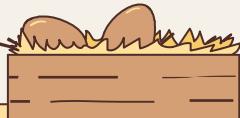
Methode Kmeans

On souhaite garder les points dans un même cluster aussi proches les uns des autres que possible mais aussi garder les clusters aussi éloignés les uns des autres que possible.

03

On recommence !

Une fois chaque individu associé à un cluster, on recommence en calculant la distance à partir du nouveau centre obtenu.



02

Processus

On place aléatoirement K points comme centre initiaux des clusters et on calcule la distance entre nos individus et ces derniers.

04

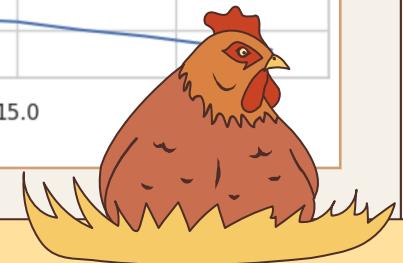
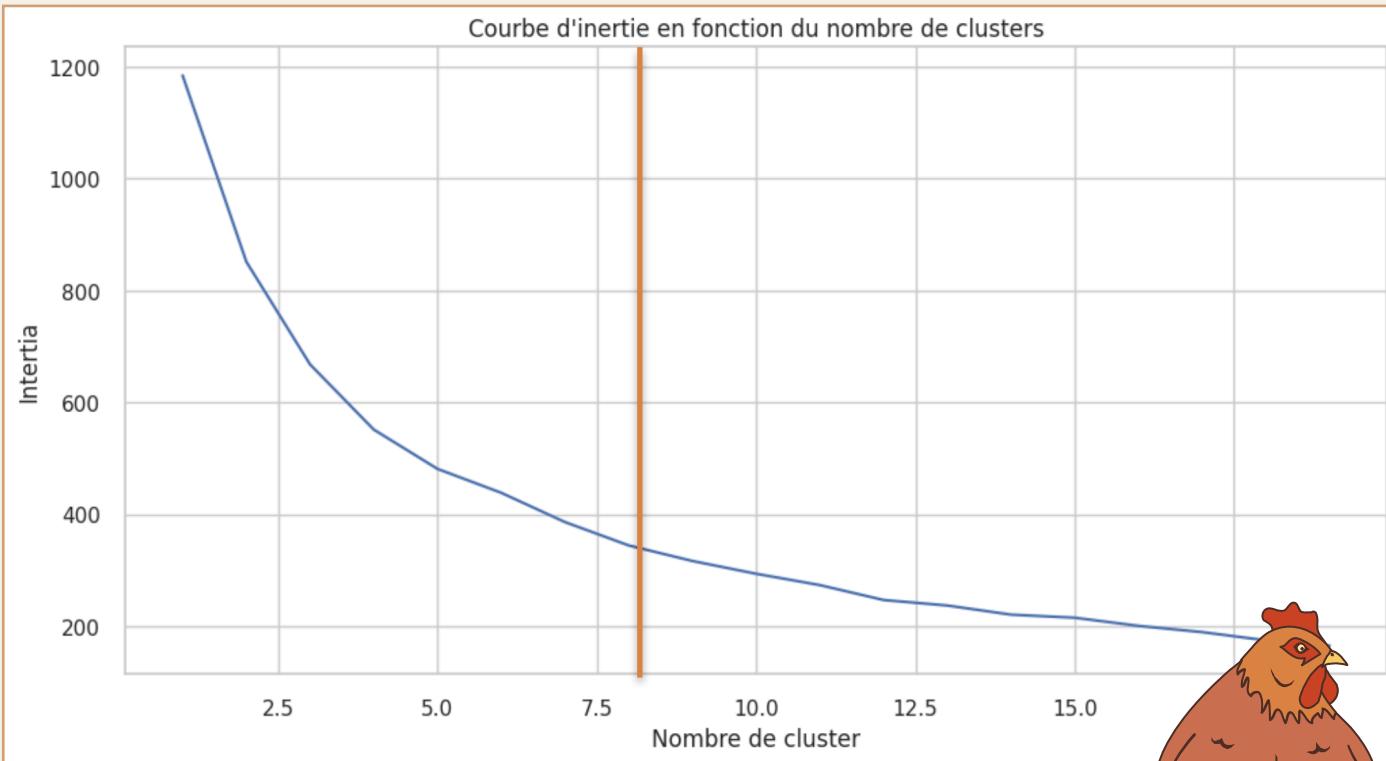
Convergence

On finit par retrouver les mêmes centres, et donc les clusters finaux.

COURBE D'INERTIE

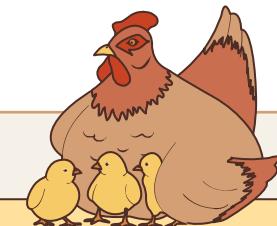
On trace la variation de l'inertie intra-cluster en fonction de K.

Courbe d'inertie pour choisir le nombre optimal K



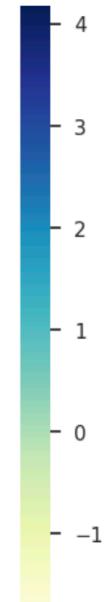
ANALYSE DES CLUSTERS KMEANS

LE CLUSTER 2 SEMBLE INTERESSANT



Heatmap des médianes standardisées par cluster

	0	1	2	3	4	5	6	7
Disponibilité alimentaire en quantité (kg/personne/an)	0.02	1.61	0.13	-1.08	-1.08	0.99	0.02	0.42
Disponibilité de protéines en quantité (g/personne/jour)	-0.06	1.42	0.38	-1.07	-1.08	1.10	0.08	0.43
Disponibilité intérieure	-0.48	-0.06	0.59	-0.52	0.44	2.52	1.45	-0.43
Exportations - Quantité	-0.32	-0.31	3.15	-0.33	-0.33	-0.14	-0.32	-0.25
Importations - Quantité	-0.39	-0.36	2.12	-0.45	-0.48	3.38	0.01	-0.26
PIB par habitant	-0.35	0.28	1.38	-0.84	-0.74	0.60	-0.43	1.06
Stabilité politique	0.10	0.44	0.70	-0.54	-1.74	-0.41	-1.12	1.24
Population	-0.53	-0.42	0.28	-0.28	4.17	1.67	0.83	-0.52



CLUSTERS KMEANS



	PIB	STABILITE POLITIQUE	IMPORTATIONS	DISPONIBILITE PROTEINE
ALLEMAGNE	\$53 071	0,59	842	7g/Jour
BELGIQUE	\$50 442	0,43	338	4g/Jour
HONG-KONG	\$59 842	0,83	907	22g/Jour
FRANCE	\$44 577	0,28	506	8g/Jour
PAYS-BAS	\$55 088	0,92	608	8g/Jour
POLOGNE	\$30 064	0,52	55	12g/Jour
THAÏLANDE	\$17 423	-0,75	2	4g/Jour

ANALYSE EN COMPOSANTES PRINCIPALES

01

Méthode ACP

On souhaite réduire les dimensions de nos données (synthétiser) tout en conservant un maximum d'informations.

02

Matrice de covariance

On mesure les relations linéaires entre nos différentes variables une fois les données centralisées, que l'on regroupe dans une matrice.

03

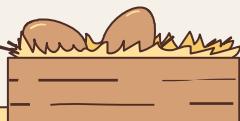
Calcul des vecteurs et valeurs propres

On utilise la matrice de corrélation pour calculer les axes principaux et leur importance dans l'explication des variations de nos données.

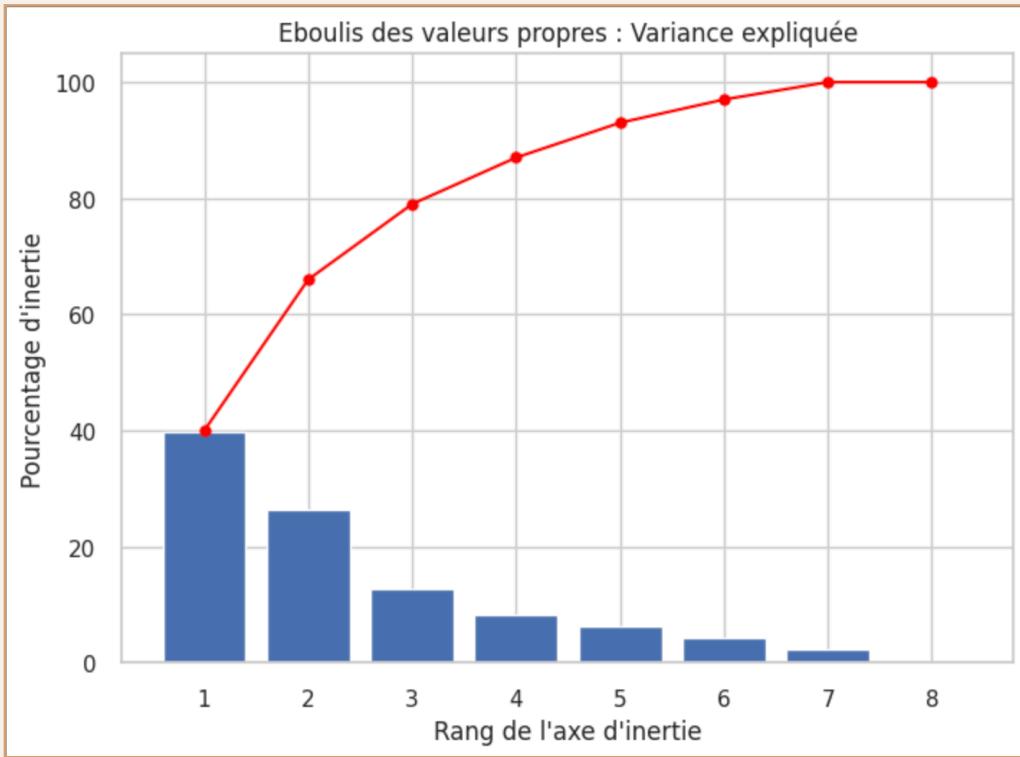
04

Eboulis des valeurs propres

On choisit le nombre de composantes principales optimales en fonction du pourcentage de variations expliquées.



EBOULIS DES VALEURS PROPRES



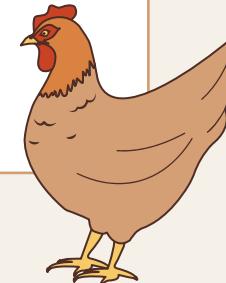
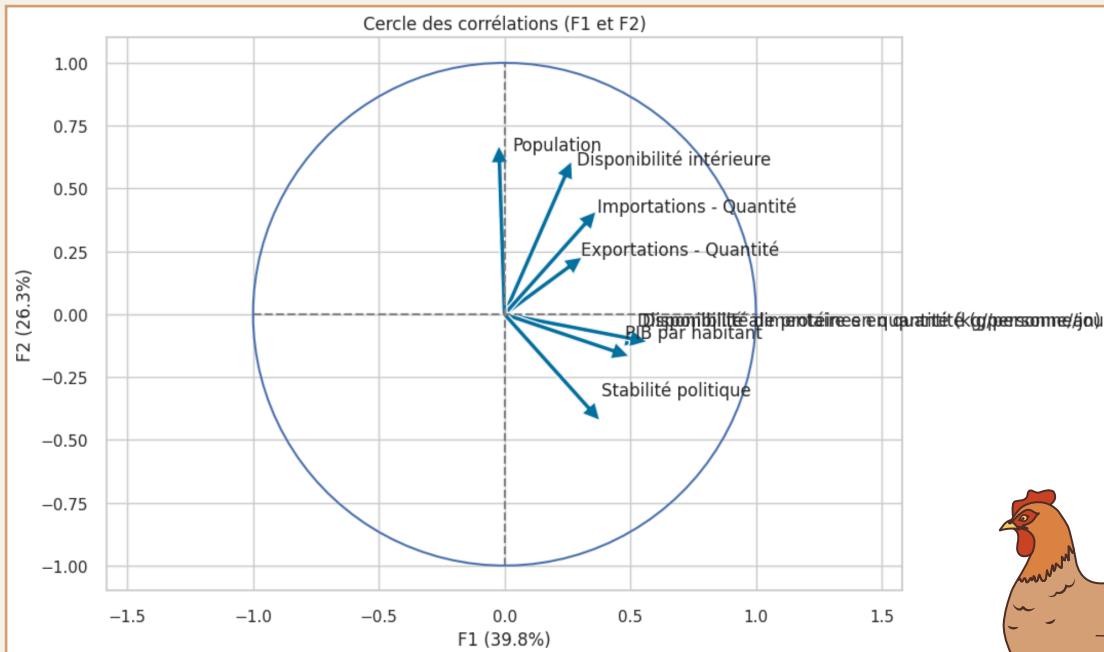
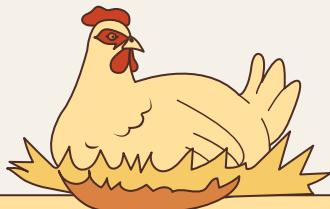
Seulement
trois composantes
permettent d'expliquer
80 % de l'inertie.



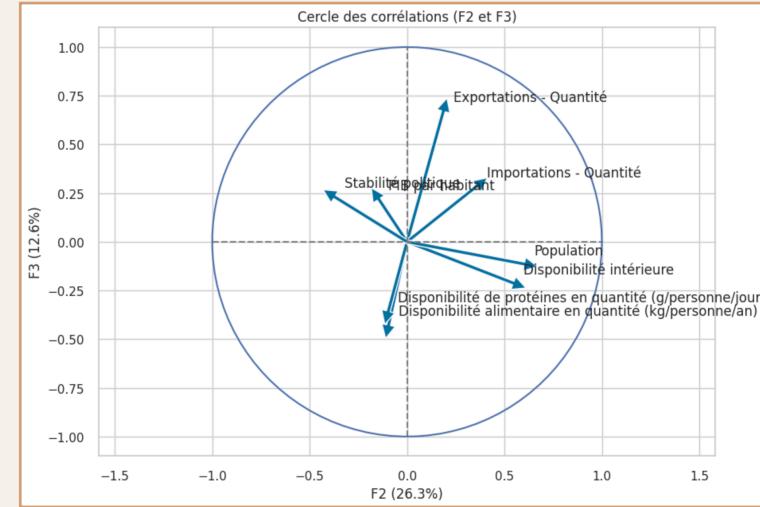
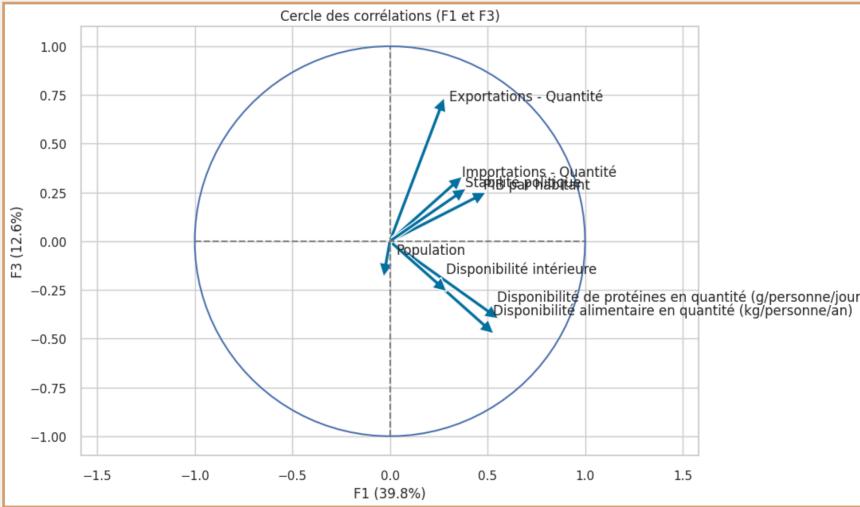
CERCLE DE CORRELATION

PERMET
D'ANALYSER LES
CORRELATION
ENTRE NOS
VARIABLES...

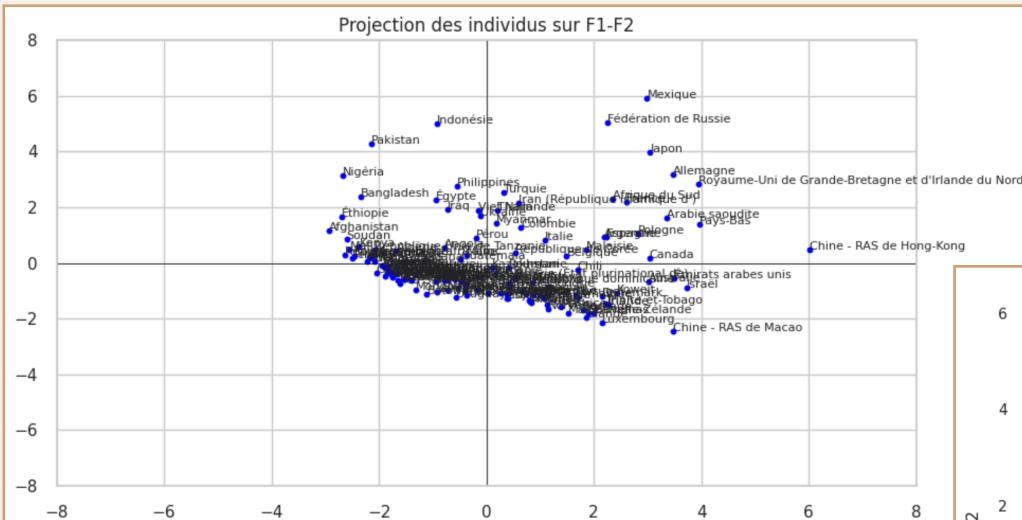
... AVEC LES
COMPOSANTES
PRINCIPALES, MAIS
AUSSI ENTRES
ELLES.



CERCLE DE CORRELATION



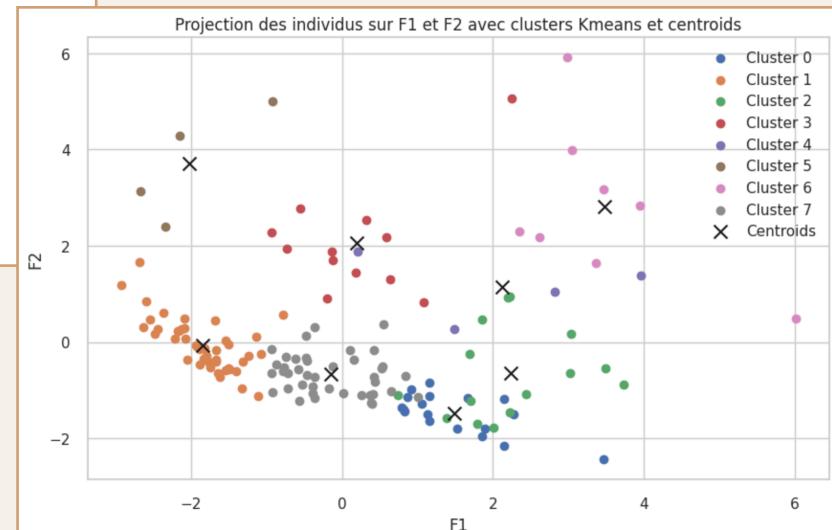
PROJECTION DES INDIVIDUS



ON PEUT DONC
ANALYSER LE
COMPORTEMENT
DE NOS
INDIVIDU...

... MAIS
AUSSI DE NOS
CLUSTERS EN
FONCTION DES
COMPOSANTES
PRINCIPALES
OBTENUES

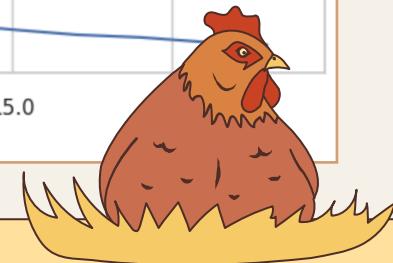
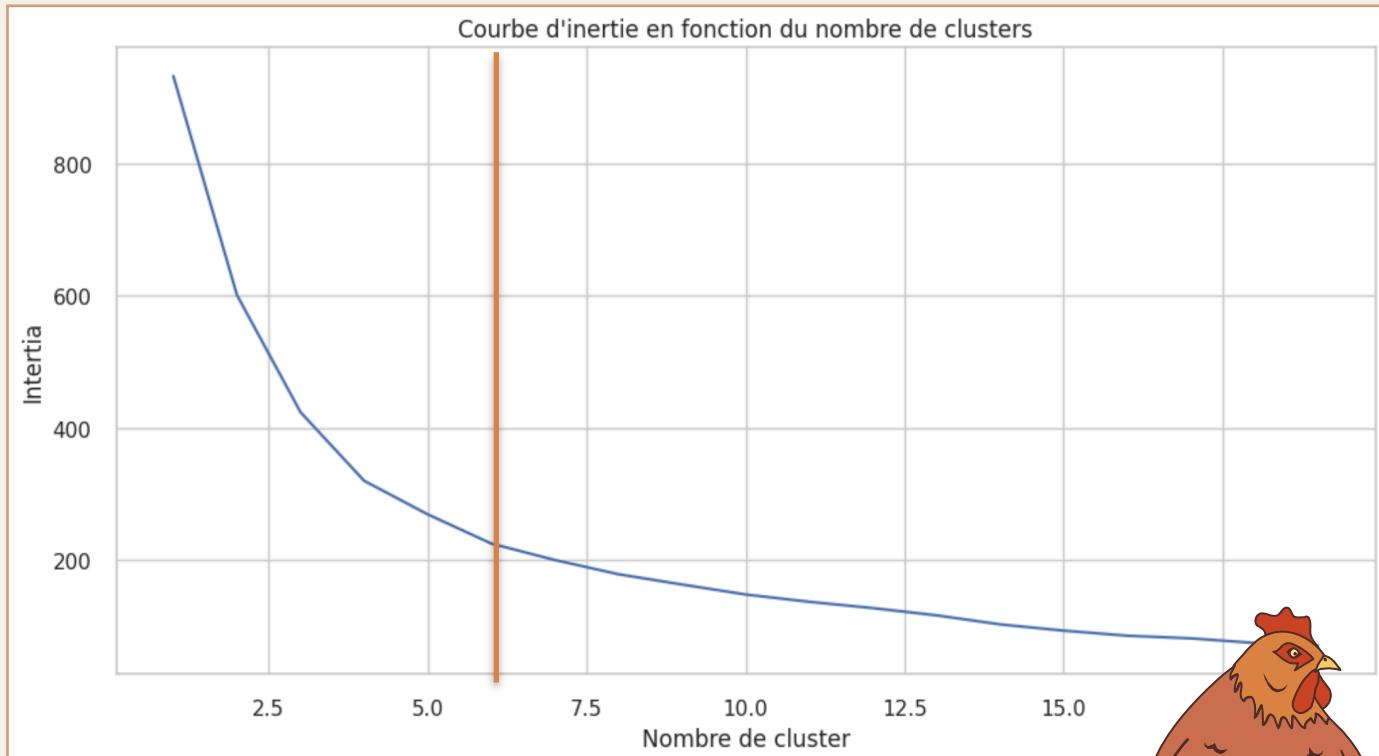
IL EST MÊME POSSIBLE
DE FAIRE UNE PROJECTION 3D
EN AJOUTANT UNE TROISIÈME COMPOSANTE !



COURBE D'INERTIE APRES ACP

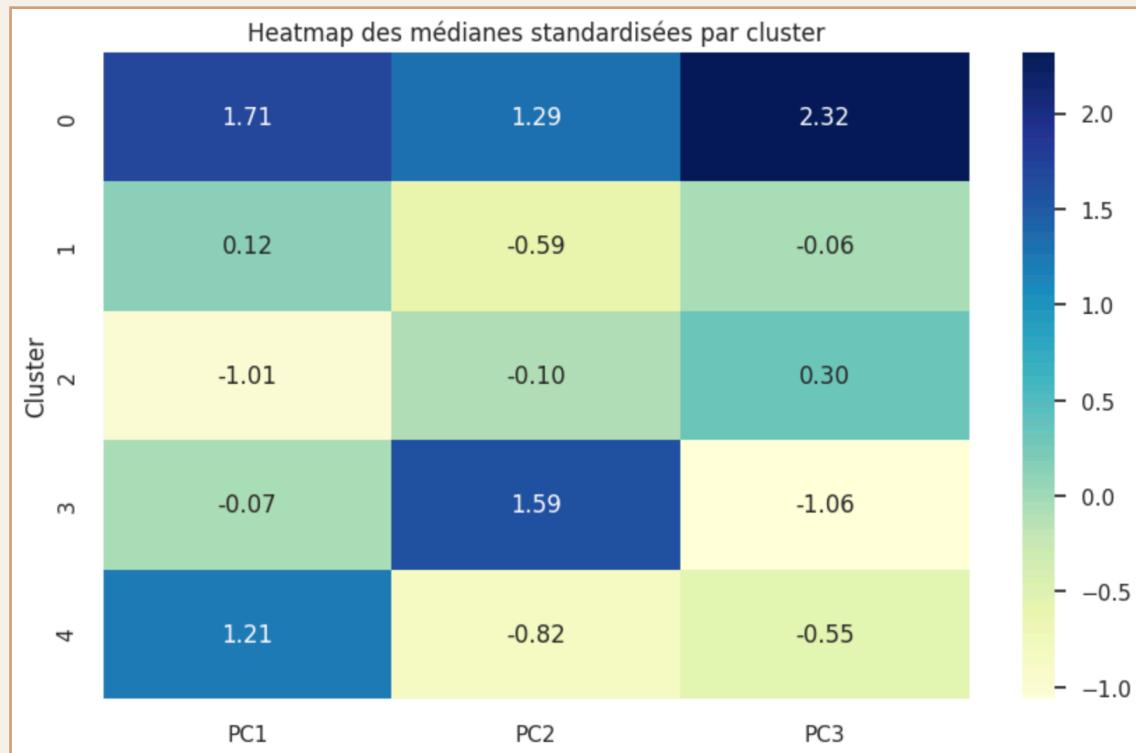
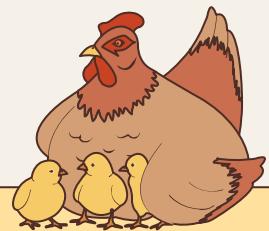
On trace la variation de l'inertie intra-cluster en fonction de K.

Courbe d'inertie pour choisir le nombre optimal K



ANALYSE DES CLUSTERS KMEANS APRES ACP

LE CLUSTER 0 SEMBLE INTERESSANT



ANALYSE DES RESULTATS APRES ACP

On retrouve encore l'Allemagne, la Belgique, Hong-Kong, la France, les Pays-Bas, la Thaïlande, le Royaume-Unis, la Pologne... Mais aussi le Japon.



PRINCIPAUX CANDIDATS

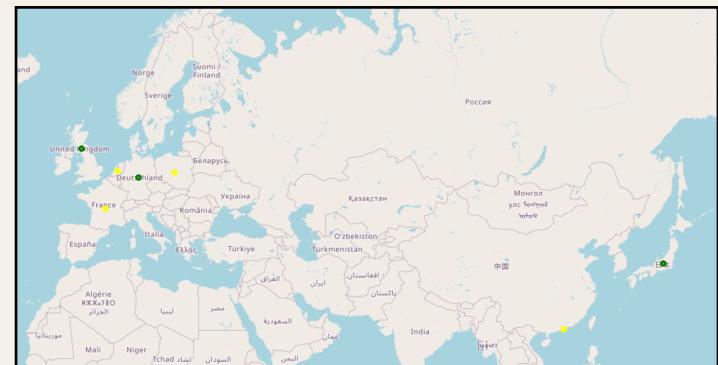
Résultats obtenus aux regards des différents clusters obtenus et des valeurs respectives de chaque pays aux variables initiales.

PAYS	
ALLEMAGNE	1
ROYAUME-UNIS	2
JAPON	3
BELGIQUE	4
PAYS-BAS	5

Conclusion

- L'Allemagne semble être un candidat idéal de par sa proximité, la stabilité de son économie et ses besoins en protéines.
- Le Royaume-Unis et le Japon sont des candidats potentiels, de même que la Belgique et les Pays-Bas.

Locations



CONCLUSION

Il est recommandé de s'implanter en Allemagne, afin de pouvoir par la suite continuer l'expansion au Royaume-Unis ou en Belgique.

Le Japon semble être une alternative cohérente, de part son besoin en protéine et sa proximité avec Hong-Kong.

Toutefois, la proximité avec la Chine oblige à faire des analyses plus abouties pour ce continent, notamment en terme de stabilité politique (Inde, Thaïlande, Chine, etc..)

