



La poule qui chante

Étude de marché

1^{ère} analyse pour le lancement à l'international

Sélection, nettoyage et préparation des données

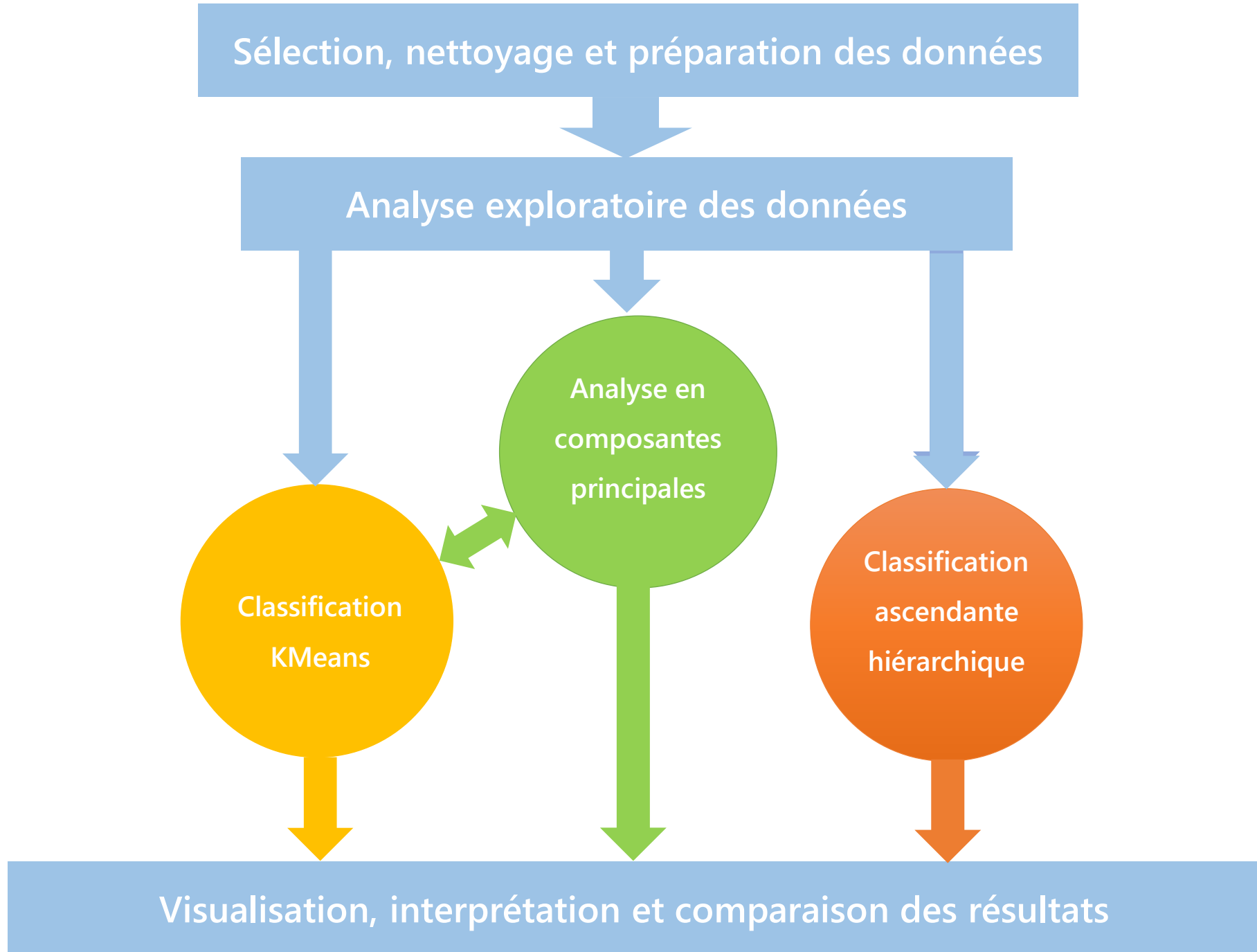
Analyse exploratoire des données

Analyse en
composantes
principales

Classification
KMeans

Classification
ascendante
hiérarchique

Visualisation, interprétation et comparaison des résultats



Sélection des données

Variables	Signification
Dispo_protein (g/pers/j)	Disponibilité en protéine en g/personne/jour
Exportations (kt)	Exportation de volaille en milliers de tonne
Importations (kt)	Importation de volaille en milliers de tonne
Production (kt)	Production de volaille en milliers de tonne
Var_stock (kt)	Variation de stock en milliers de tonne
Population_2017 (k)	Population année 2017 en millier
evo_popu_2008_17	Evolution de la population entre 2008 et 2017 en millier
PIB/hab (\$)	PIB par habitant année 2017 en dollars
Index_stability	Indice de stabilité politique
Infla_jan_2017 (%)	Inflation début 2017 des produits alimentaires
Distance (km)	Distance des autres capitales en kilomètre
Evo_pib_2008_17 (%)	Evolution du PIB entre 2008 et 2017

Données sur la volaille (170 éléments)

	Zone	Dispo_protein (g/pers/j)	Disponibilité intérieure	Exportations (kt)	Importations (kt)	Production (kt)	Var_stock (kt)
0	Afghanistan	0.54	57.0	0.0	29.0	28.0	0.0
1	Afrique du Sud	14.11	2118.0	63.0	514.0	1667.0	-0.0



Données sur la démographie (229 éléments)

	Zone	Population_2017 (k)	evo_popu_2008_17
0	Afghanistan	36296.113	0.31
1	Afrique du Sud	57009.756	0.15



data (169 éléments)

	Zone	Dispo_protein (g/pers/j)	Disponibilité intérieure	Exportations (kt)	Importations (kt)	Production (kt)	Var_stock (kt)	Population_2017 (k)	evo_popu_2008_17
0	Afghanistan	0.54	57.0	0.0	29.0	28.0	0.0	36296.113	0.31
1	Afrique du Sud	14.11	2118.0	63.0	514.0	1667.0	-0.0	57009.756	0.15

	Zone	ISO3	PIB/hab (\$)
0	Aruba	ABW	29326.708058
2	Afghanistan	AFG	530.149831

PIB par habitant 2017 (255 éléments)

	ISO3	Index_stability
0	AFG	-2.801084
1	ALB	0.378475

Indice de stabilité 2017 (224 élément)

	ISO3	Infla_jan_2017 (%)
0	AFG	15.442461
2	ZAF	3.373015

Inflation janvier 2017 (204 éléments)

	ISO3	Distance (km)
0	ABW	7685.884
1	AFG	5590.381

Distance capitales (224 éléments)

	ISO3	Evo_pib_2008_17 (%)
0	ABW	0.087640
1	AFE	0.442064

Evolution PIB 2008-2017 (253 éléments)

	Zone	ISO3	PIB/hab (\$)	Index_stability	Infla_jan_2017 (%)	Distance (km)	Evo_pib_2008_17 (%)
0	Aruba	ABW	29326.708058	1.335038	1.643545	7685.8840	0.087640
1	Afghanistan	AFG	530.149831	-2.801084	15.442461	5590.3810	0.843588

Fusion 1 (181 éléments)

data (169 éléments)									
	Zone	Dispo_protein (g/pers/j)	Disponibilité intérieure	Exportations (kt)	Importations (kt)	Production (kt)	Var_stock (kt)	Population_2017 (k)	evo_popu_2008_17
0	Afghanistan	0.54	57.0	0.0	29.0	28.0	0.0	36296.113	0.31
1	Afrique du Sud	14.11	2118.0	63.0	514.0	1667.0	-0.0	57009.756	0.15

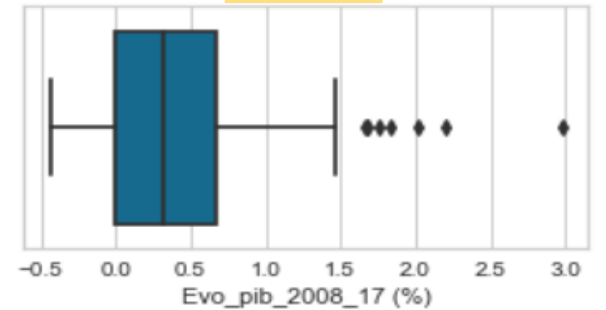
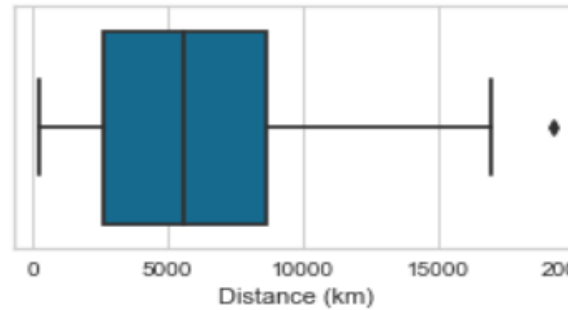
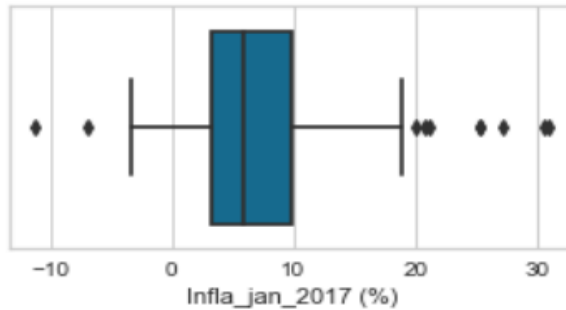
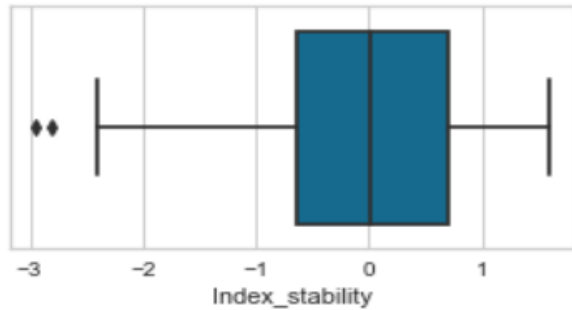
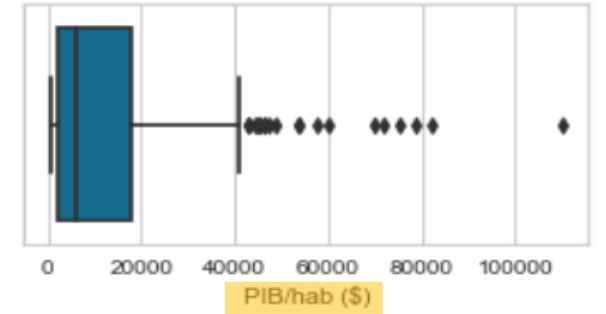
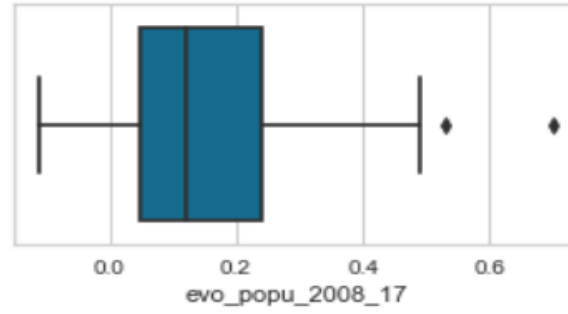
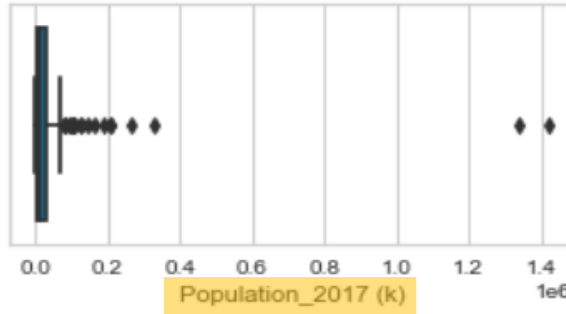
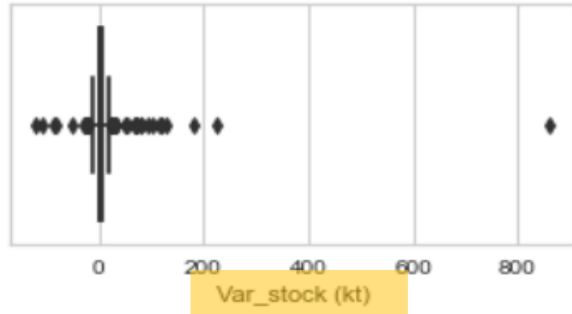
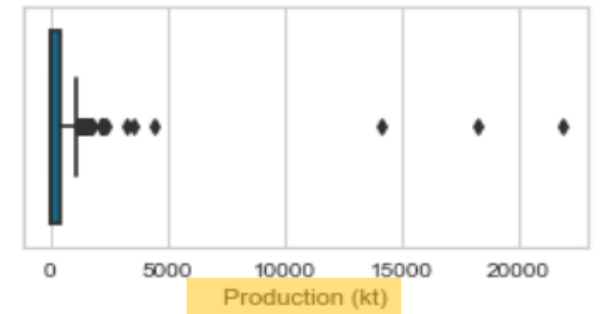
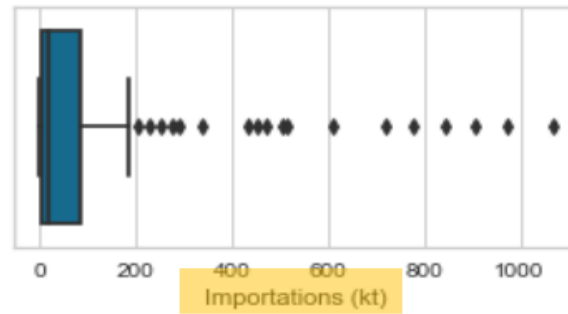
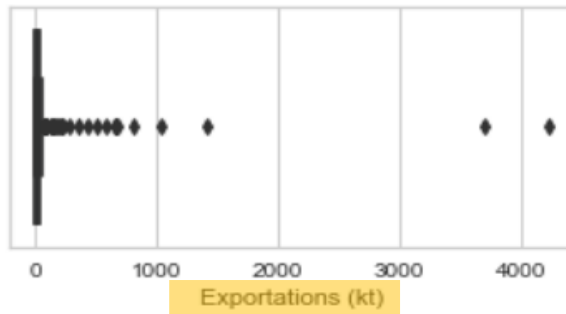
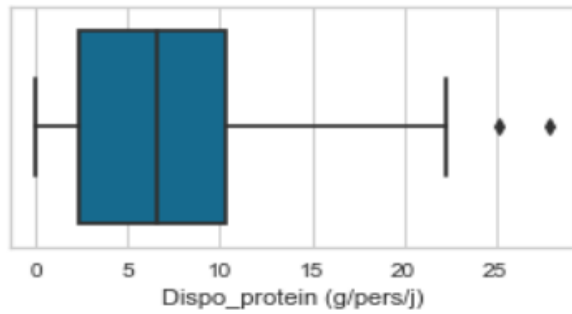


Fusion1 (181 éléments)							
	Zone	ISO3	PIB/hab (\$)	Index_stability	Infla_jan_2017 (%)	Distance (km)	Evo_pib_2008_17 (%)
0	Aruba	ABW	29326.708058	1.335038	1.643545	7685.8840	0.087640
1	Afghanistan	AFG	530.149831	-2.801084	15.442461	5590.3810	0.843588



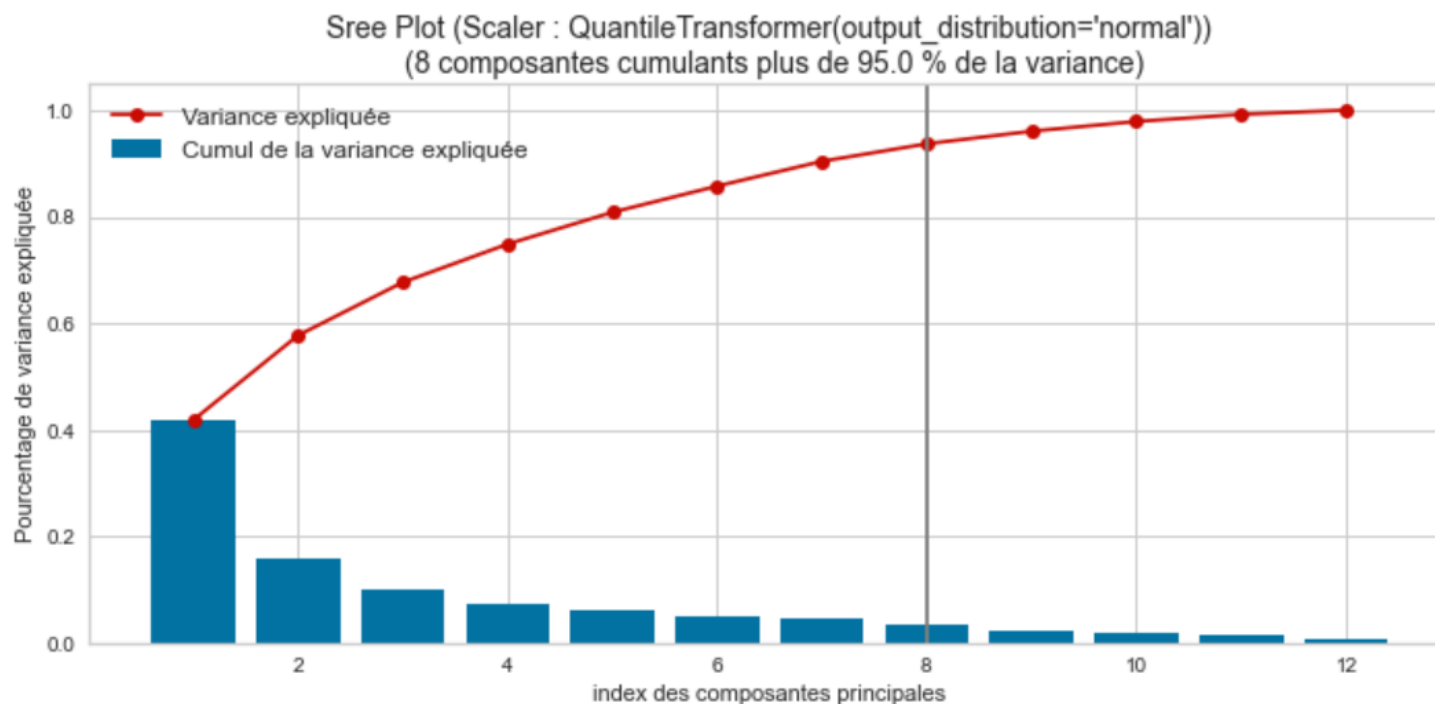
							data (156 éléments)							
	Zone	Dispo_protein (g/pers/j)	Exportations (kt)	Importations (kt)	Production (kt)	Var_stock (kt)	Population_2017 (k)	evo_popu_2008_17	ISO3	PIB/hab (\$)	Index_stability	Infla_jan_2017 (%)	Distance (km)	Evo_pib_2008_17 (%)
151	Émirats arabes unis	14.80	94.0	433.0	48.0	-26.0	9487.203	0.34	ARE	43063.967477	0.618641	3.859943	5249.535	0.237871
152	Équateur	6.15	0.0	0.0	340.0	-1.0	16785.361	0.15	ECU	6246.404252	-0.074854	4.877023	9367.098	0.688656
153	États-Unis	19.93	3692.0	123.0	21914.0	80.0	325084.756	0.07	USA	59907.754261	0.291817	2.101637	6169.154	0.318722

Analyse exploratoire des données



Analyse en composantes principales

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Variance expliquée	11.20	4.27	2.68	1.92	1.60	1.32	1.25	0.88	0.64	0.49	0.36	0.21
% variance expliquée	41.79	15.91	9.98	7.17	5.98	4.92	4.66	3.28	2.37	1.83	1.33	0.79
% cumul variance expliquée	41.79	57.69	67.67	74.84	80.82	85.74	90.40	93.68	96.05	97.88	99.21	100.00



Interprétation des axes

Coefficient de corrélation entre variables/axes

Dispo_protein (g/pers/j)	0.14	-0.25	0.081	-0.1	0.32	0.0036	-0.48	0.39
Exportations (kt)	0.86	0.18	0.29	0.34	-0.14	0.054	0.066	0.014
Importations (kt)	0.29	-0.44	-0.81	0.18	0.031	-0.085	0.072	-0.12
Production (kt)	0.23	0.5	-0.28	-0.47	0.37	-0.05	-0.086	0.11
Var_stock (kt)	0.083	-0.16	0.03	-0.22	0.15	0.35	0.69	0.28
Population_2017 (k)	0.091	0.42	-0.24	-0.1	0.093	-0.14	0.038	-0.14
evo_popu_2008_17	-0.14	0.14	-0.085	0.39	-0.065	-0.52	0.094	0.65
PIB/hab (\$)	0.18	-0.26	0.14	-0.17	0.16	-0.27	-0.27	-0.054
Index_stability	0.09	-0.35	0.25	-0.064	0.34	-0.12	0.039	-0.14
Infla_jan_2017 (%)	-0.06	0.11	-0.18	0.28	-0.03	0.67	-0.41	0.15
Distance (km)	-0.089	0.041	0.04	0.35	0.67	0.15	0.15	0.13
Evo_pib_2008_17 (%)	-0.13	0.18	0.05	0.41	0.35	-0.12	0.017	-0.49
		F2 (15.9%)		F4 (7.2%)	F5 (6.0%)	F6 (4.9%)	F7 (4.7%)	F8 (3.3%)

Poids des variables/axes

Dispo_protein (g/pers/j)	0.02	0.063	0.0066	0.011	0.1	1.3e-05	0.23	0.15
Exportations (kt)	0.74	0.034	0.083	0.11	0.021	0.0029	0.0043	0.00018
Importations (kt)	0.087	0.19	0.65	0.034	0.00098	0.0072	0.0052	0.015
Production (kt)	0.051	0.25	0.076	0.22	0.14	0.0025	0.0074	0.013
Var_stock (kt)	0.0069	0.025	0.00089	0.048	0.022	0.12	0.47	0.079
Population_2017 (k)	0.0083	0.18	0.057	0.01	0.0087	0.021	0.0014	0.021
evo_popu_2008_17	0.019	0.02	0.0073	0.16	0.0043	0.27	0.0088	0.42
PIB/hab (\$)	0.034	0.066	0.02	0.03	0.025	0.074	0.073	0.0029
Index_stability	0.0081	0.12	0.061	0.0041	0.11	0.015	0.0016	0.02
Infla_jan_2017 (%)	0.0036	0.011	0.031	0.077	0.00088	0.44	0.17	0.023
Distance (km)	0.0078	0.0017	0.0016	0.12	0.44	0.024	0.024	0.016
Evo_pib_2008_17 (%)	0.018	0.031	0.0025	0.17	0.12	0.016	0.00029	0.24
		F2 (15.9%)		F4 (7.2%)	F5 (6.0%)	F6 (4.9%)	F7 (4.7%)	F8 (3.3%)

- La variable "Exportation (kt)" est positivement corrélée au premier axe et elle pèse 74% dans l'explication de ce dernier.
- La variable "Importation (kt)" est négativement corrélée au troisième axe et elle pèse 65% dans l'explication de ce dernier.

Analyse des axes:

- le premier axe oppose à droite les pays exportateurs et à gauche les pays non exportateurs.
- le deuxième axe oppose en bas les pays importateurs et en haut les pays non importateurs

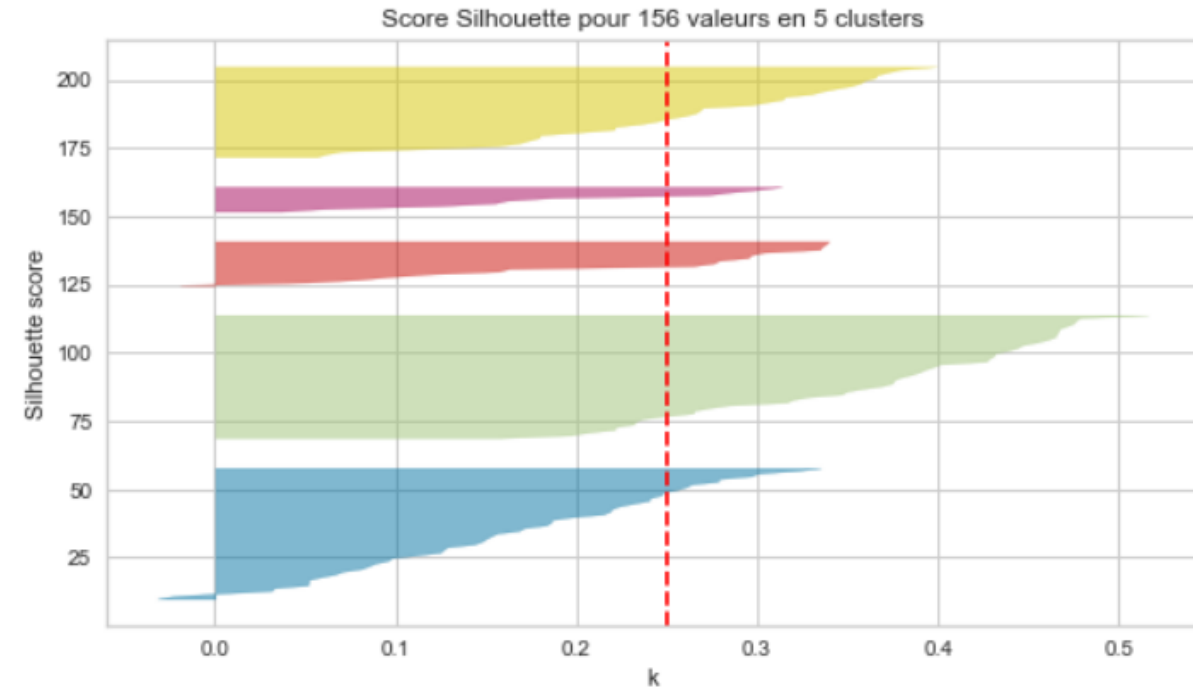
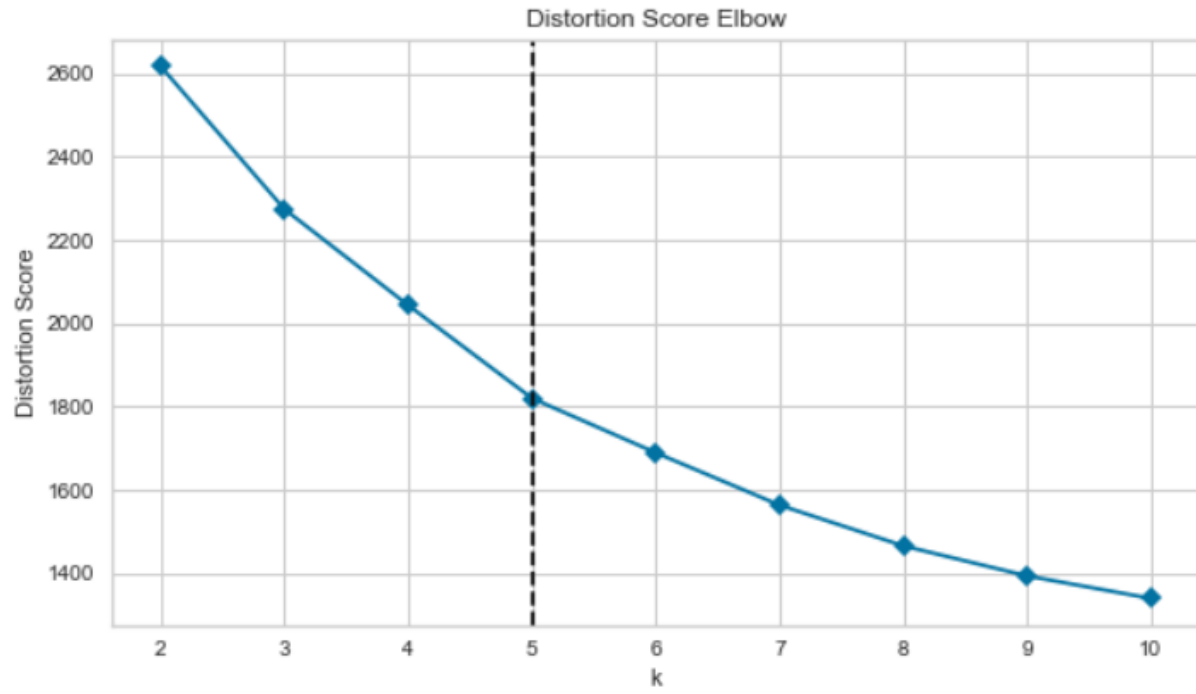
Analyse des individus:

Deux grands groupes se distingues à gauche et à droite du premier axe, on peut supposé que ce sont les pays exportateurs et non exportateur.

Les individus qui expliquent le mieux les axes sont situés aux extrémités et proches du cercle, exemple : Brésil, États-Unis, Japon, Rwanda, Belize ...

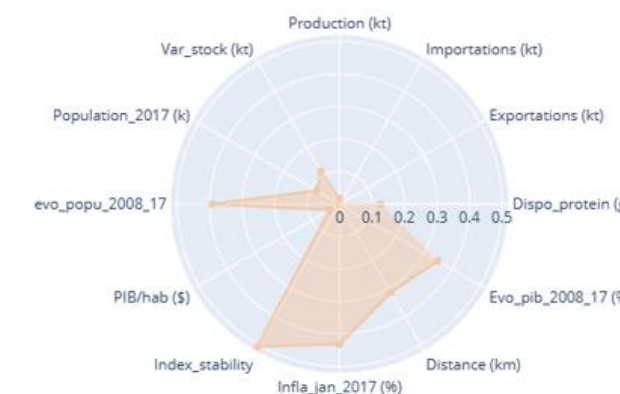
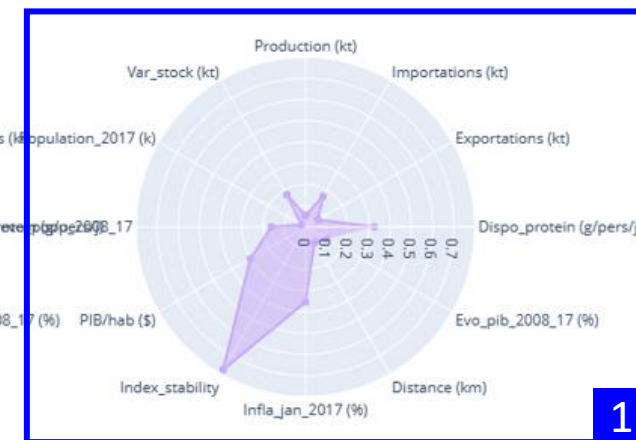
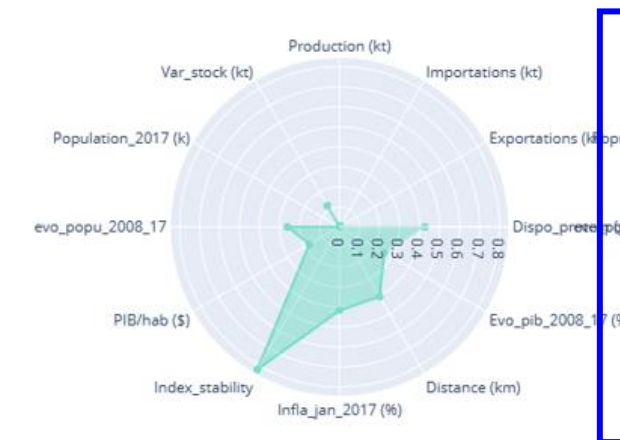
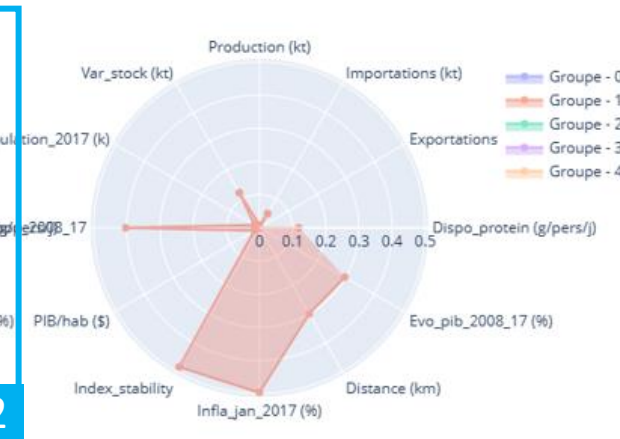
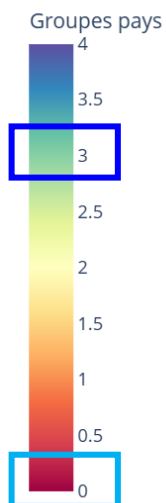
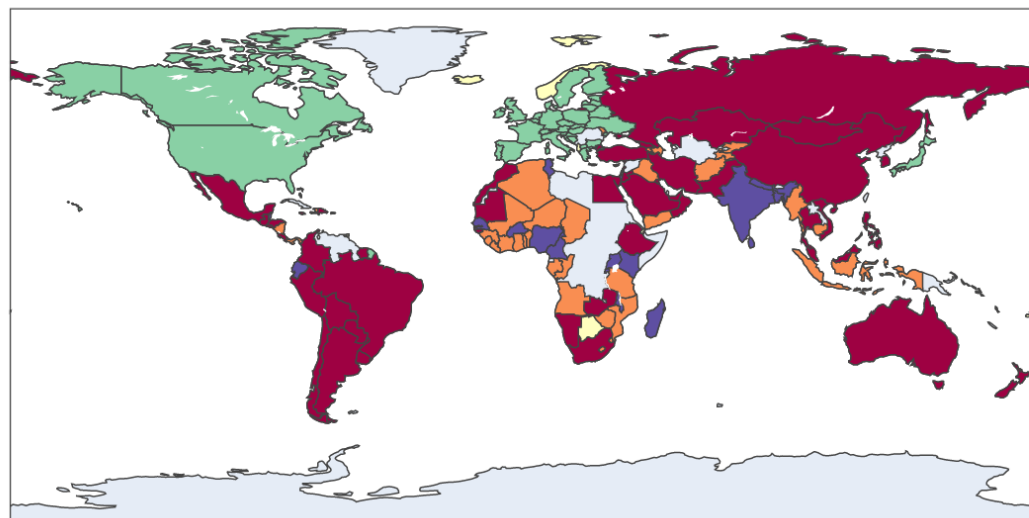


Classification KMeans



- Nombre optimal de clusters : 5
- Score silhouette moyen ≈ 0.25

Visualisation des clusters



kmeans_clusters	Dispo_protein (g/pers/j)	Exportations (kt)	Importations (kt)	Production (kt)	Var_stock (kt)	Population_2017 (k)	evo_popu_2008_17	PIB/hab (\$)	Index_stability	Infla_jan_2017 (%)
0	0.279274	0.036074	0.119121	0.060451	0.129591	0.046266	0.354815	0.093832	0.610676	0.476857
1	0.118653	0.000000	0.049777	0.007892	0.121956	0.017452	0.402918	0.017335	0.484328	0.496734
2	0.427014	0.000023	0.008330	0.000621	0.123040	0.000587	0.261611	0.174696	0.823374	0.416791
3	0.325651	0.068962	0.165602	0.054354	0.175811	0.021038	0.161199	0.300722	0.776456	0.354488
4	0.127074	0.001086	0.000000	0.016055	0.116745	0.081955	0.392883	0.036501	0.507213	0.431793

Groupe 1:

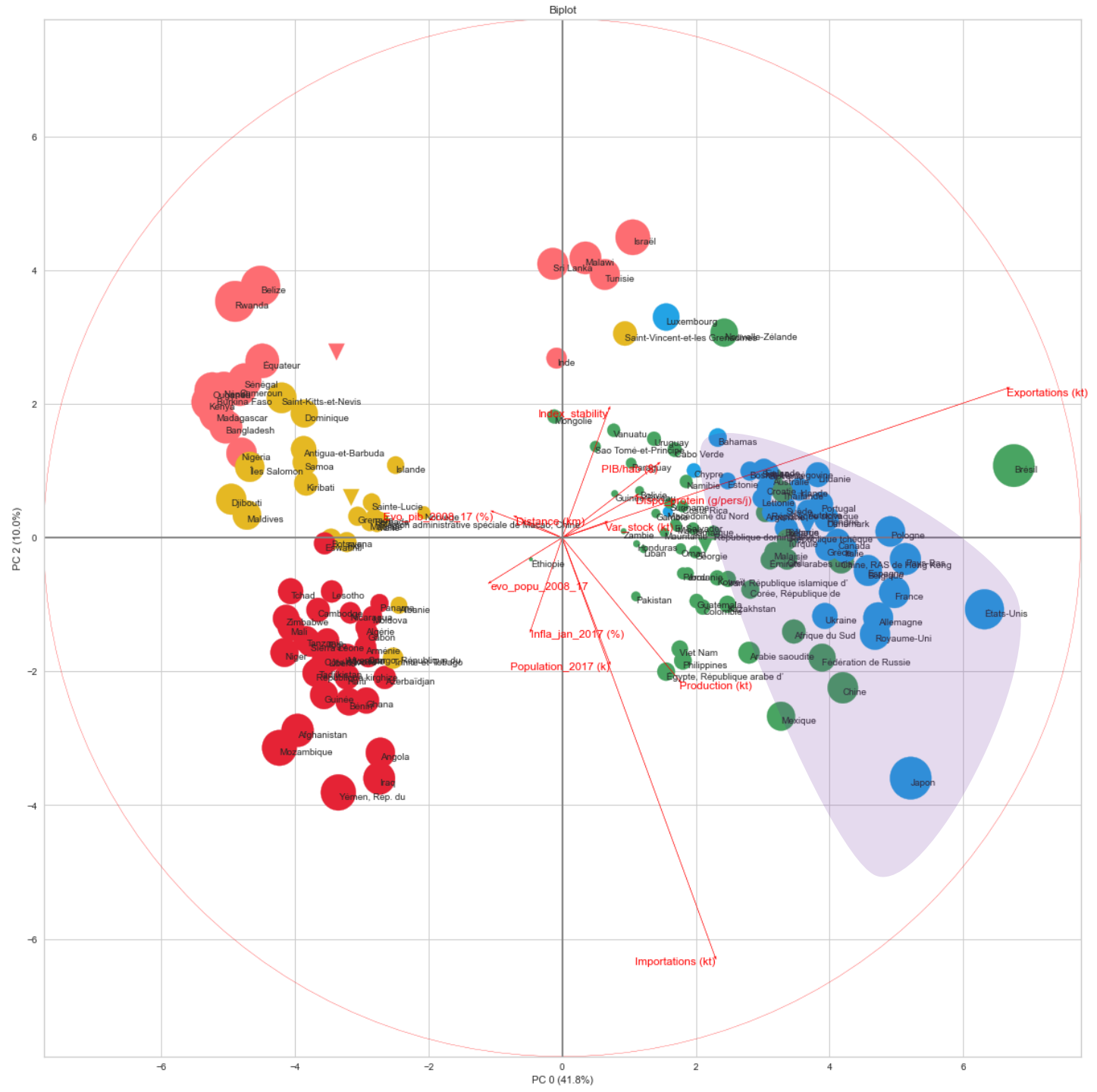
- * Pays politiquement stables
- * Peu distants
- * Inflation maitrisée
- * PIB par habitant élevé
- * Faible évolution de la population (population vieillissante)
- * Une demande en volaille
- * Producteurs et exportateurs

Groupe 1 : ['Allemagne' 'Autriche' 'Bahamas' 'Belgique' 'Bosnie-Herzégovine' 'Bulgarie' 'Bélarus' 'Canada' 'Chypre' 'Croatie' 'Danemark' 'Espagne' 'Estonie' 'Finlande' 'France' 'Grèce' 'Hongrie' 'Irlande' 'Italie' 'Japon' 'Lettonie' 'Lituanie' 'Luxembourg' 'Macédoine du Nord' 'Pays-Bas' 'Pologne' 'Portugal' 'Royaume-Uni' 'République slovaque' 'Slovénie' 'Suisse' 'Suède' 'République tchèque' 'Ukraine' 'États-Unis']

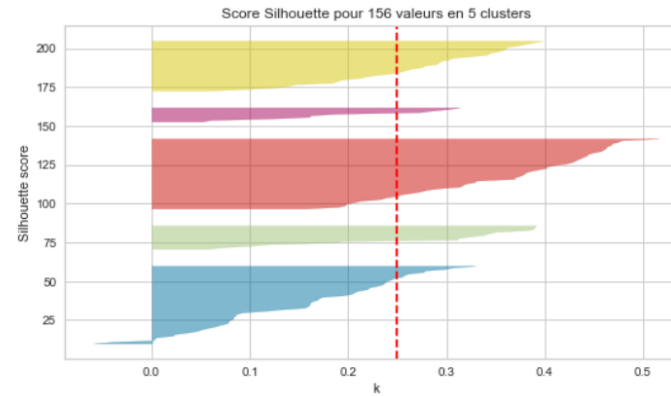
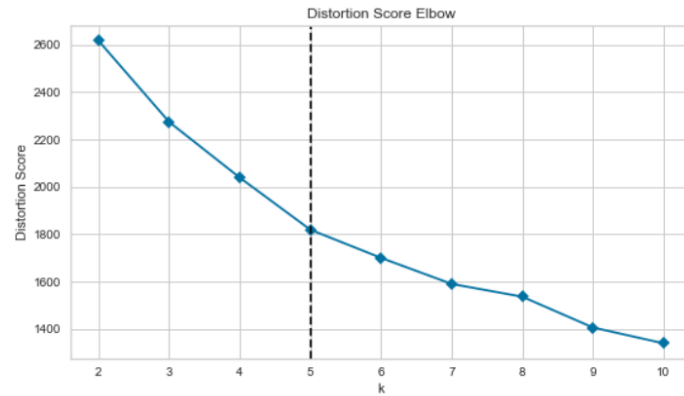
Groupe 2:

- * Pays politiquement stables
- * distants
- * Inflation élevée
- * PIB par habitant faible
- * Population jeune
- * Une demande en volaille
- * Producteurs

Groupe 2: ['Afrique du Sud' 'Arabie saoudite' 'Argentine' 'Australie' 'Bolivie' 'Brésil' 'Cabo Verde' 'Chili' 'Chine, RAS de Hong Kong' 'Chine' 'Colombie' 'Costa Rica' 'El Salvador' 'Fédération de Russie' 'Gambie' 'Guatemala' 'Guinée-Bissau' 'Géorgie' 'Honduras' 'Iran, République islamique d'' 'Jamaïque' 'Jordanie' 'Kazakhstan' 'Koweït' 'Liban' 'Malaisie' 'Maroc' 'Mauritanie' 'Mexique' 'Mongolie' 'Namibie' 'Nouvelle-Zélande' 'Oman' 'Pakistan' 'Paraguay' 'Philippines' 'Pérou' 'Corée, République de' 'République dominicaine' 'Sao Tomé-et-Principe' 'Suriname' 'Thaïlande' 'Turquie' 'Uruguay' 'Vanuatu' 'Viet Nam' 'Zambie' 'Égypte, République arabe d'' 'Émirats arabes unis' 'Éthiopie']

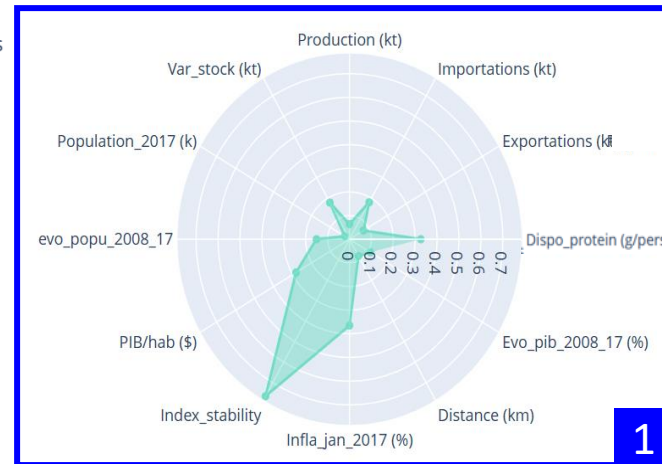
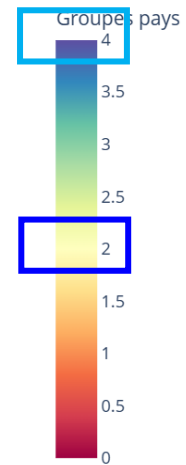
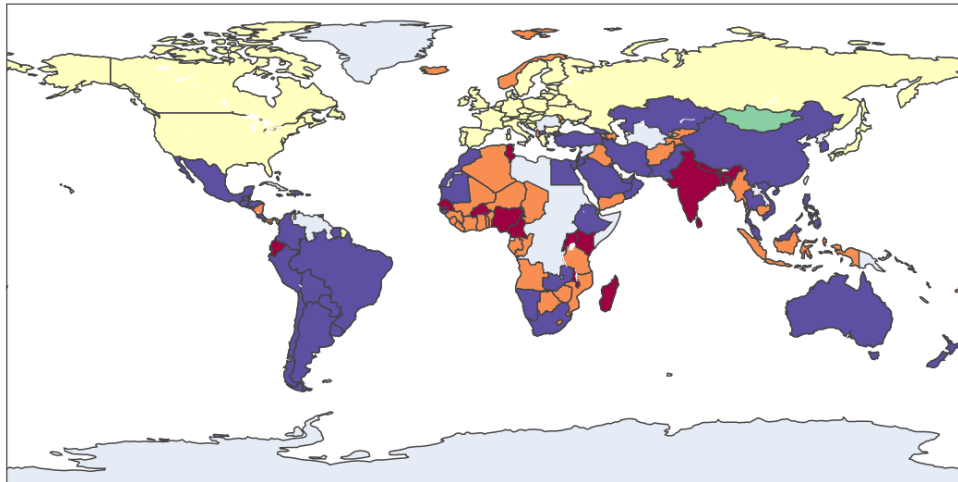


Classification KMeans

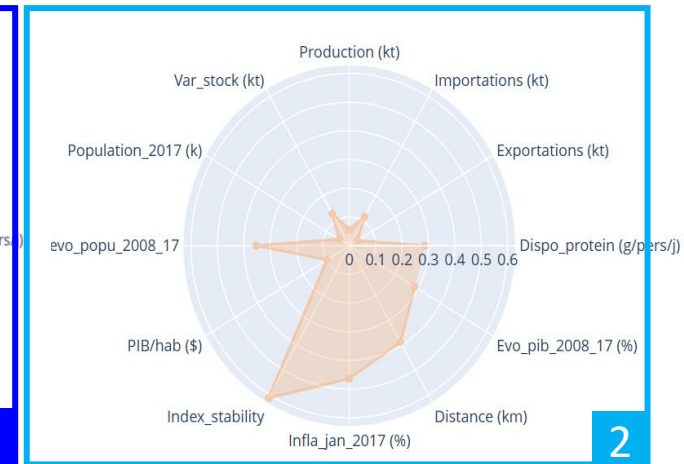


Nombre optimal de clusters : 5

Score silhouette moyen ≈ 0.25

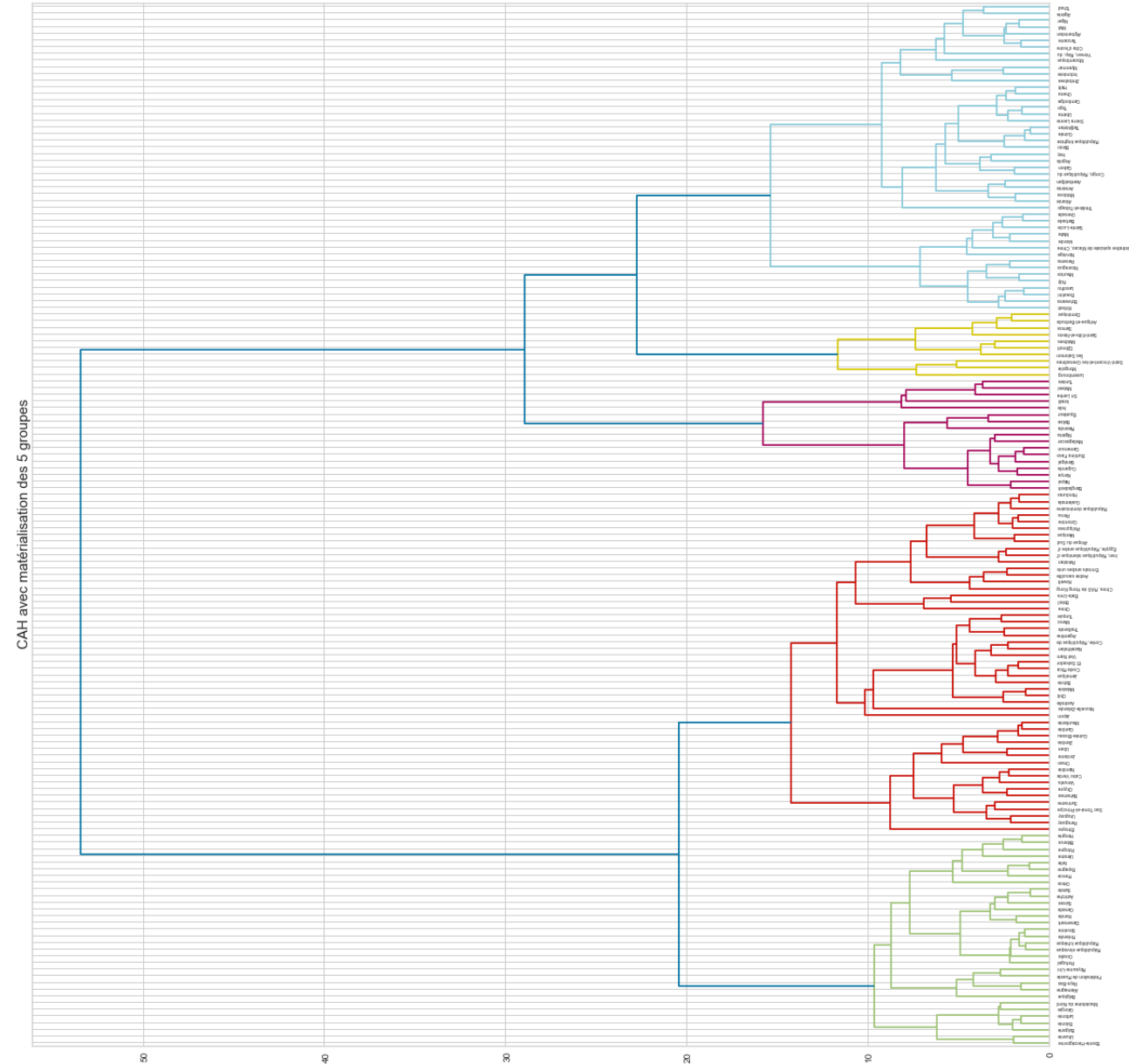
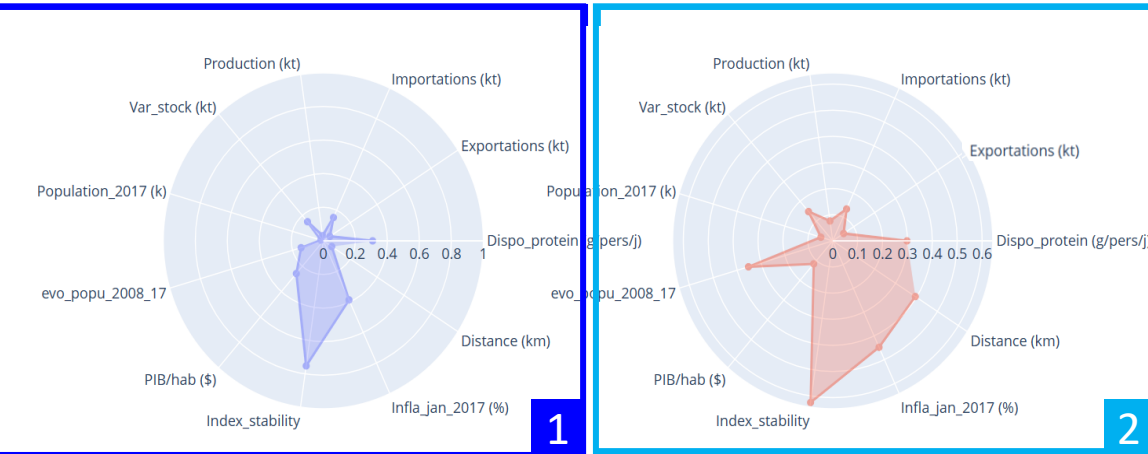
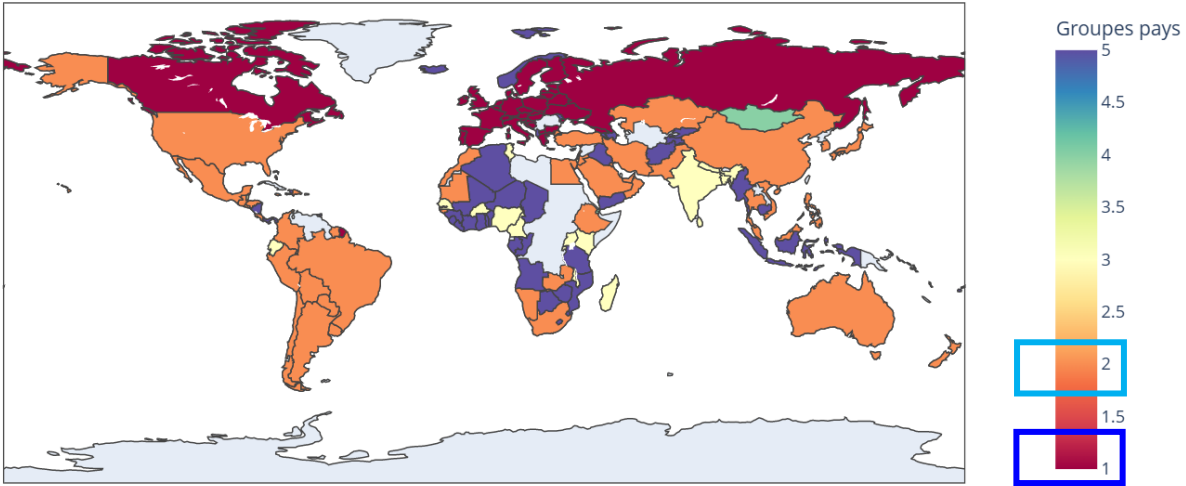


1

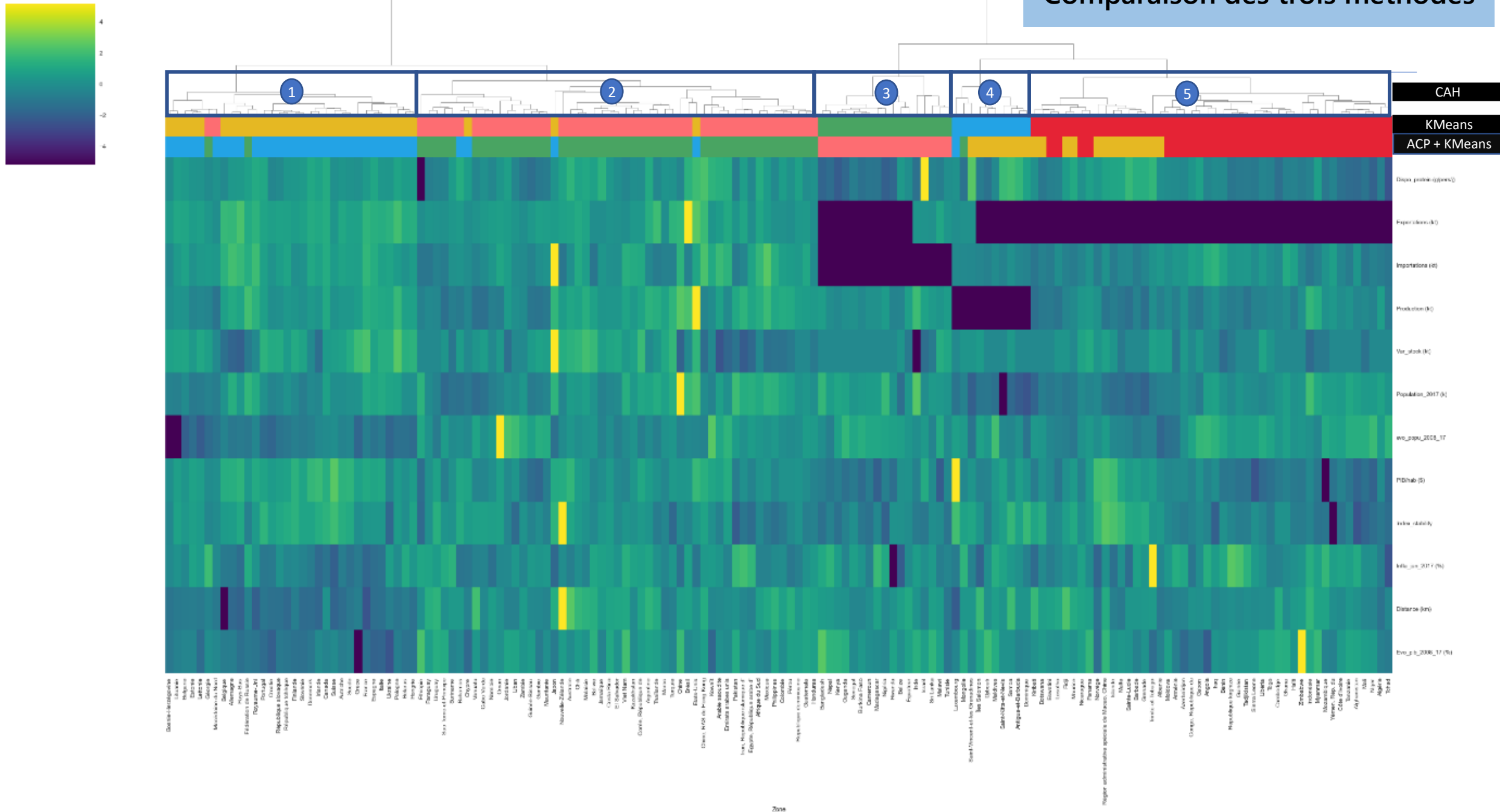


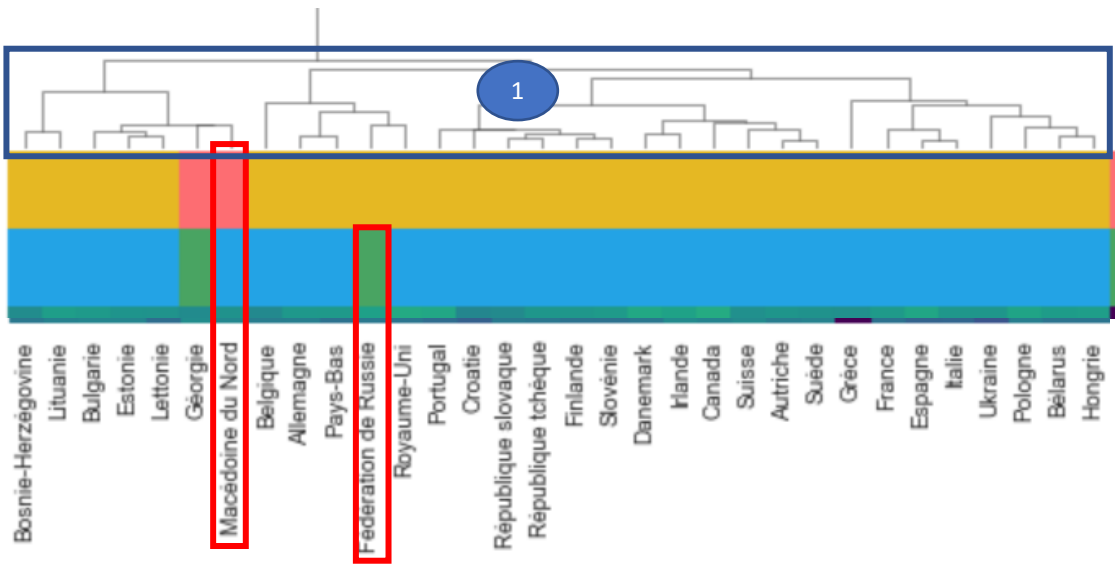
2

Classification CAH



Comparaison des trois méthodes



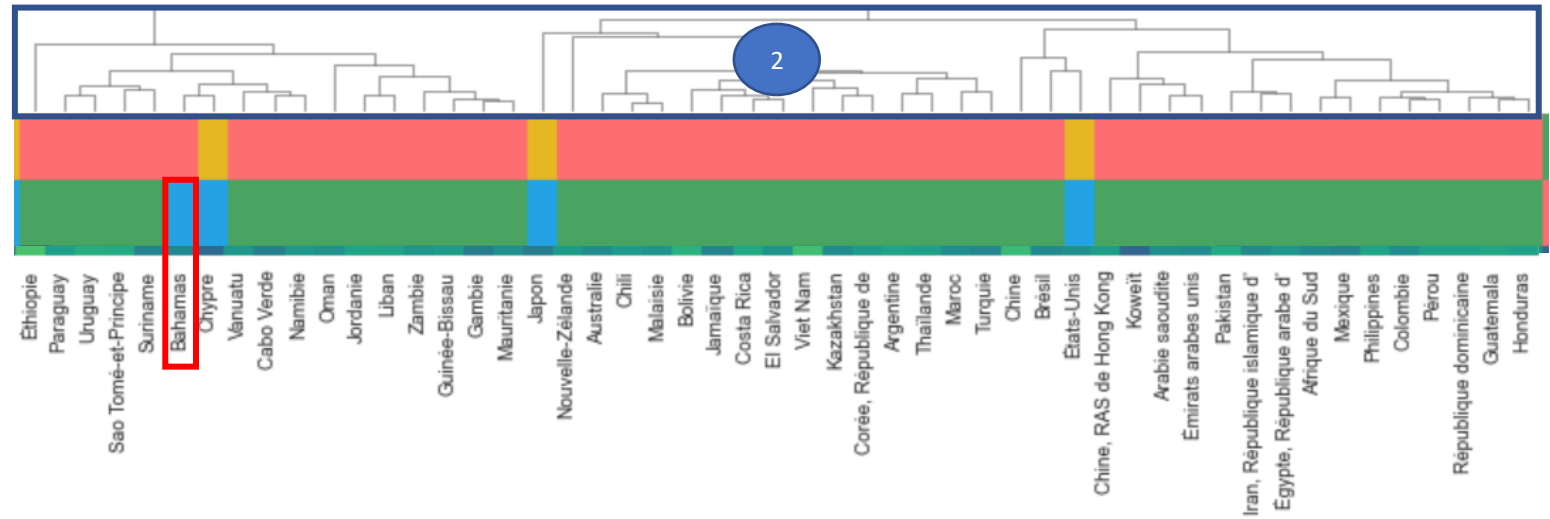


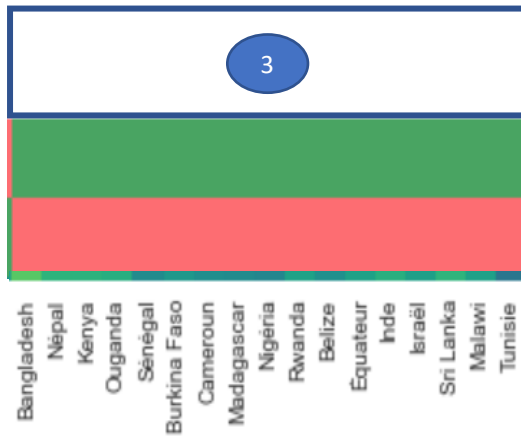
1^{er} cluster :

Les trois méthodes donnent le même résultat à quelques exceptions près.

2nd cluster :

Les trois méthodes donnent le même résultat à quelques exceptions près.



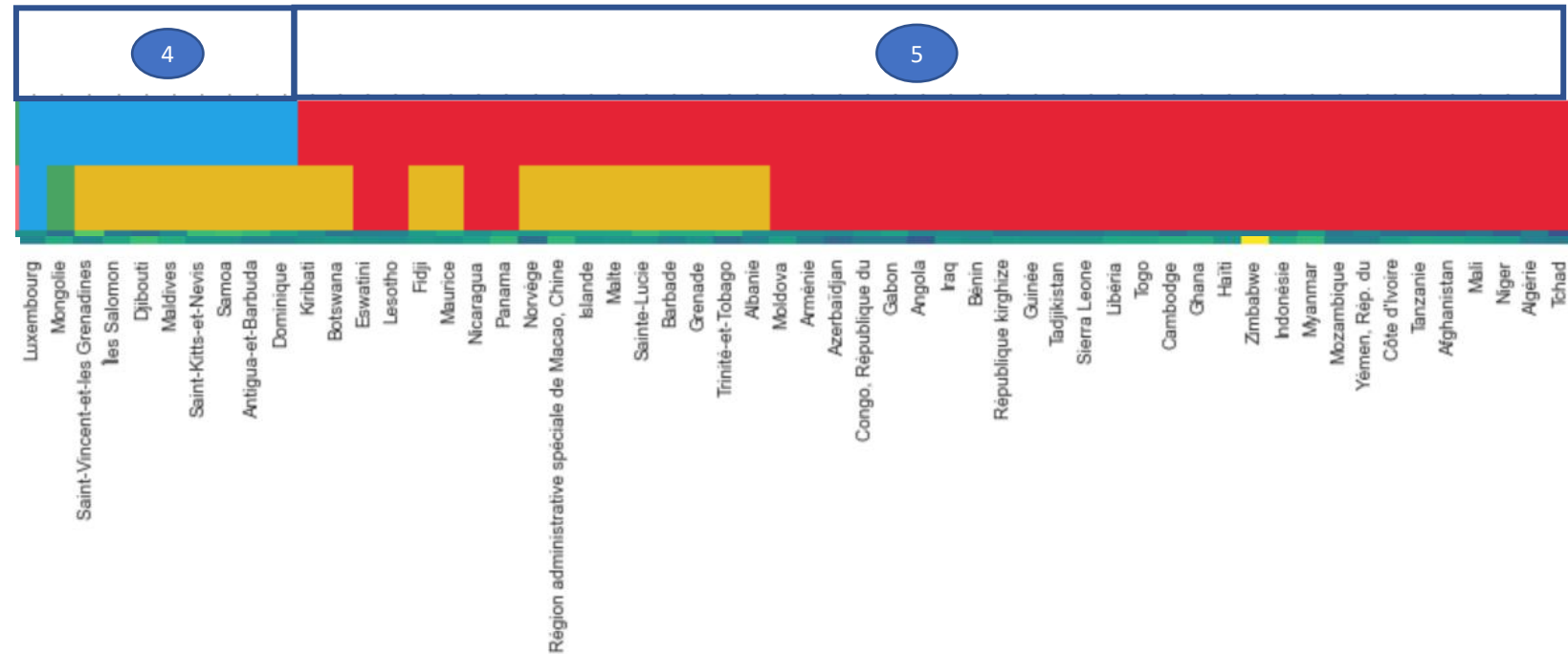


3^{ème} cluster:

Les trois méthodes donne le même résultat.

4^{ème} et 5^{ème} clusters :

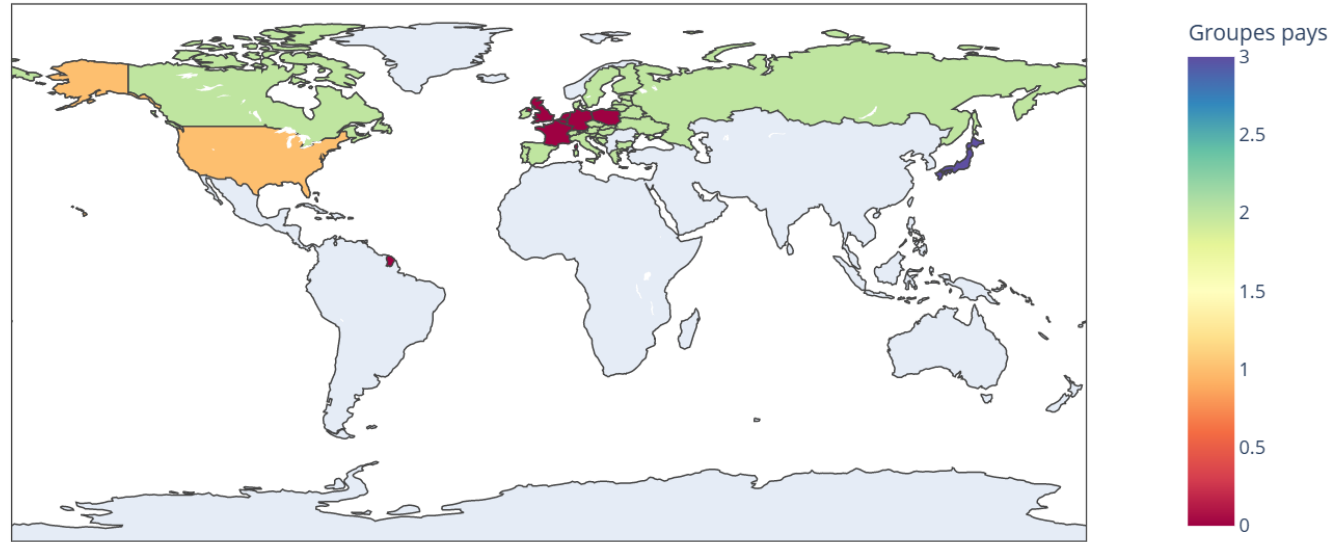
- Le clustering KMeans est identique au clustering CAH.
- Le clustering Kmeans sur données réduites donne des résultats très différents des autres méthodes.



Qualité de clustering :

- Le score silhouette du combiné PCA/KMeans est de : 0.24
- Le score silhouette de la méthode KMeans est de : 0.25
- Le score silhouette de la méthode CAH est de : 0.23

Seconde Classification



Les pays suivants représentent un premier groupe d'intérêt pour le développement à exportation:

- L'Allemagne
- La Belgique
- Les Pays-Bas
- La Pologne
- Le Royaume-Uni
- Le Japon

Conclusion