



La poule qui chante

Etude de marché du commerce de Poulet.



La poule qui chante

Objectif de l'entreprise : développer son marché à l'étranger

Mission : Déterminer les meilleurs pays d'accueil pour y exporter du Poulet

Les indicateurs (variables) utilisés dans cette étude

P	E	S	T	E	L
Politique	Economie	Social	Technologie	Environnement	Legal
- Stabilité politique (indice)	- PIB par habitant (\$) - Taux de croissance (par hab. %annuel) - Investissement étrangers directs (% PIB)	- Population totale - Population active (%)	- Accès Electricité (%) - Performance logistique (indice)	- Emission de CO2	- Solidité des garanties juridiques (indice)

- Le niveau de consommation (t) , de production (t), et d'importation (t) de Poulet
- Année : 2019

150 pays comportant toutes les valeurs des différentes variables



- Quels sont les pays susceptibles d'importer le poulet de la Poule qui Chante ?

Méthode

Stratégies : S'installer dans un pays :

- 1) Majoritairement importateur de Poulet (forte demande, forte concurrence)
- 2) Qui importe peu de poulet (faible concurrence, plus de risque)

Critères de Sélection : Importation de poulet importante et de bons indicateurs socio-économiques

Méthode de détermination du groupe des pays cibles :

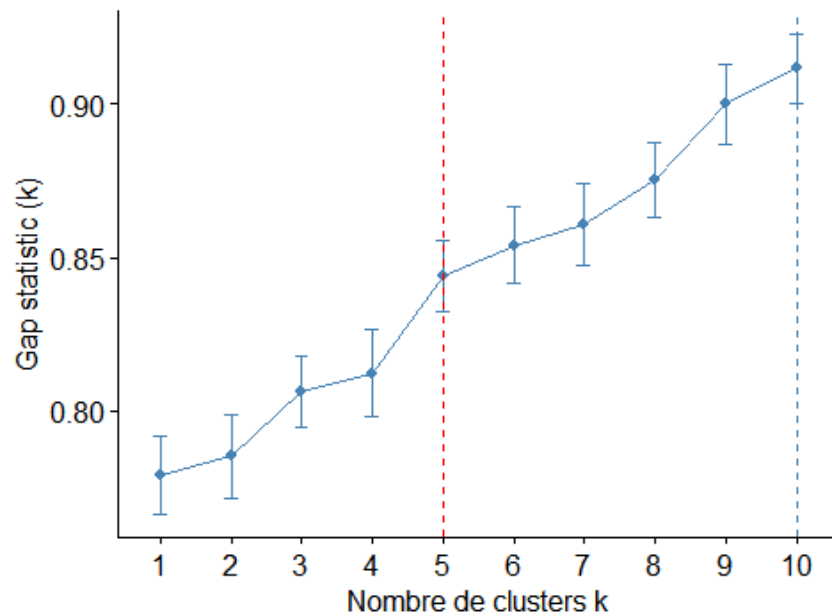
	Analyse en composantes principales	Variables d'origine
Classification Ascendante hiérarchique	Groupe 1	Groupe 2
Kmeans	Groupe 3	Groupe 4

Clustering : La méthode des kmeans

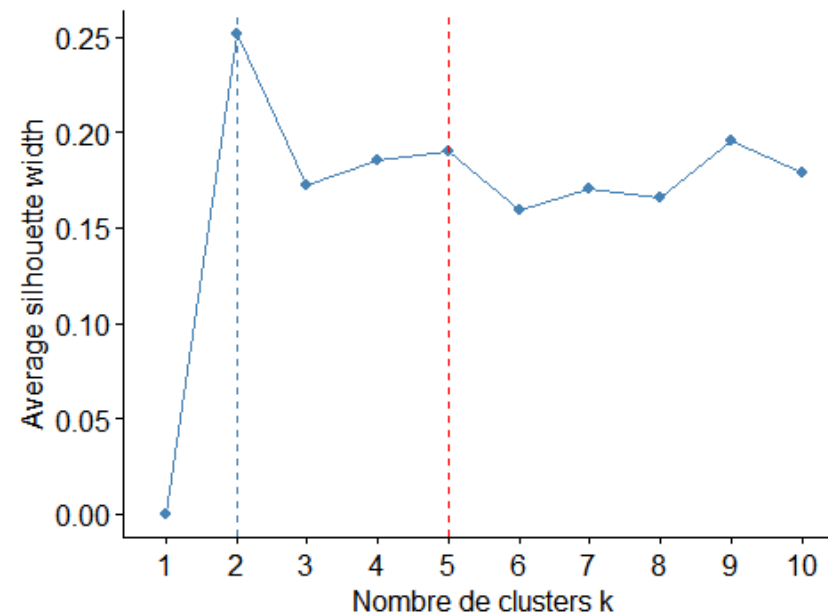
Principe : Il s'agit de trouver des groupes homogènes et compacts. Minimiser l'inertie intraclasse au sein d'un nombre de groupe donné k par un algorithme itératif.

- 1) Placement de k centroïdes au hasard (une observation)
- 2) Attribution d'un groupe aux observations les plus proches des centroïdes
- 3) Calculer le centre de gravité du groupe
- 4) Déplacer le centroïde > Etape 2

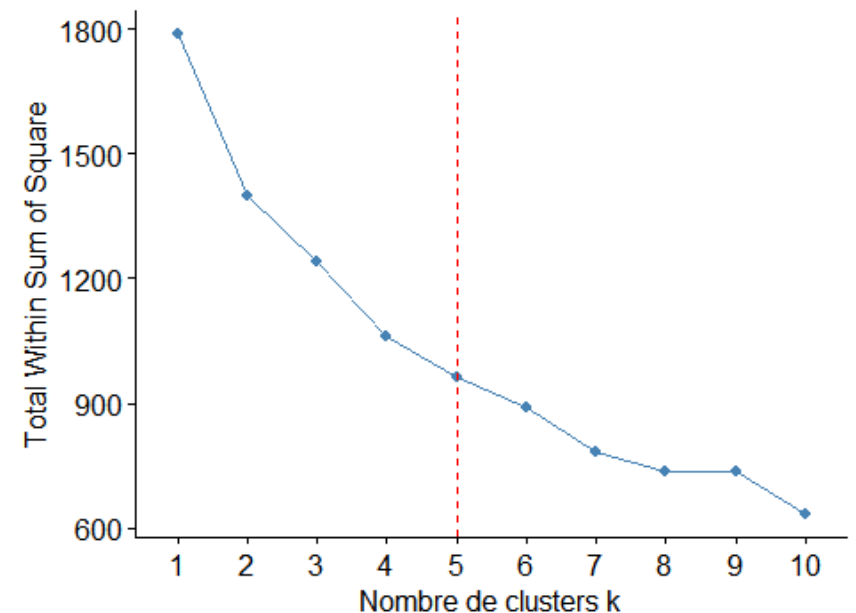
A Nombre de clusters optimal : gap statistic



B Nombre de clusters optimal : méthode des silhoues

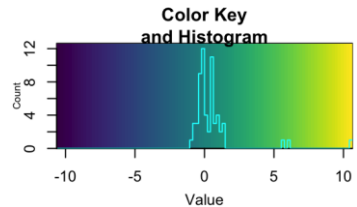


C Nombre de clusters optimal : méthode du coude



Clustering : La méthode des kmeans

Heatmap de la position des centroids
sur les différentes variables



1.47	0.68	0.56	-0.6	-0.29
0.96	0.77	-0.05	-0.81	0.17
-0.13	-0.29	6	-0.12	-0.12
0.61	0.62	0.59	-0.98	0.47
1.19	0.52	0.93	-0.67	-0.11
-0.11	0.37	0.59	-0.49	0.43
1.41	0.35	1.3	-0.8	-0.13
-0.2	10.64	-0.21	-0.15	0.07
-0.23	0.85	0.69	-0.51	0.5
0.6	0.56	-0.17	-0.39	0.02
0.02	-0.31	5.56	-0.22	-0.08
0.92	-0.41	1.09	-0.36	-0.22
1	2	3	4	5

Clusters (kmeans)

Parmi les 5 clusters :

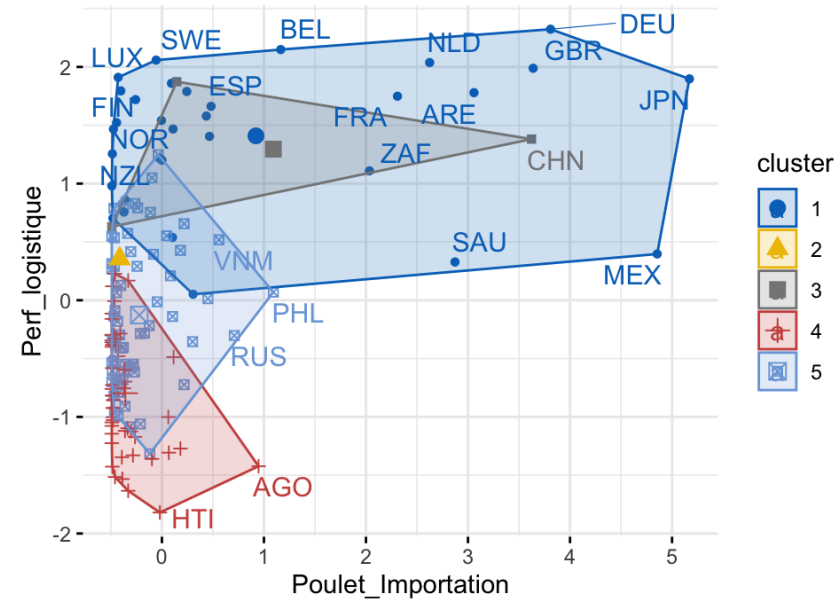
- L'importation de poulet est importante dans le cluster 1 et 3
- Cluster 1 : PIB, population active, et performance logistique
- Cluster 3 : effet de la population ?

Variables

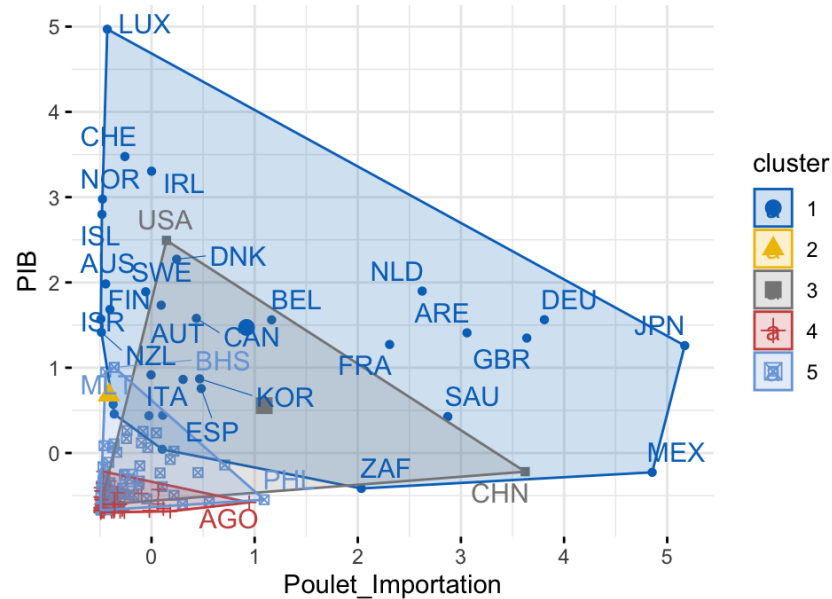
PIB
Stabilite
Population
Acces_electricite
Emission_CO2
Indice_juridique
Perf_logistique
Investissement_etranger
taux_croissance
Taux_pop_active
Poulet_Consommation
Poulet_Importation

Clustering : La méthode des kmeans

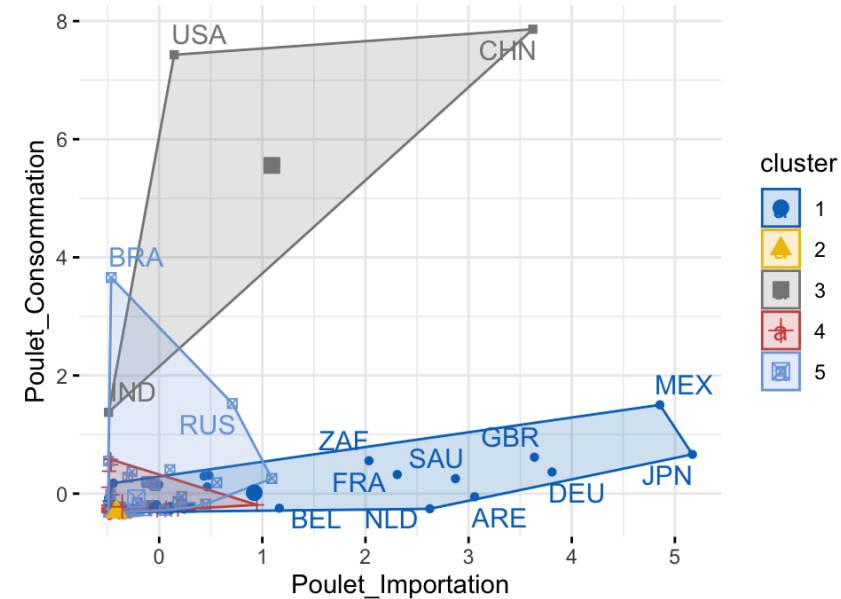
A Clustering sur les variables : Kmeans



B



C



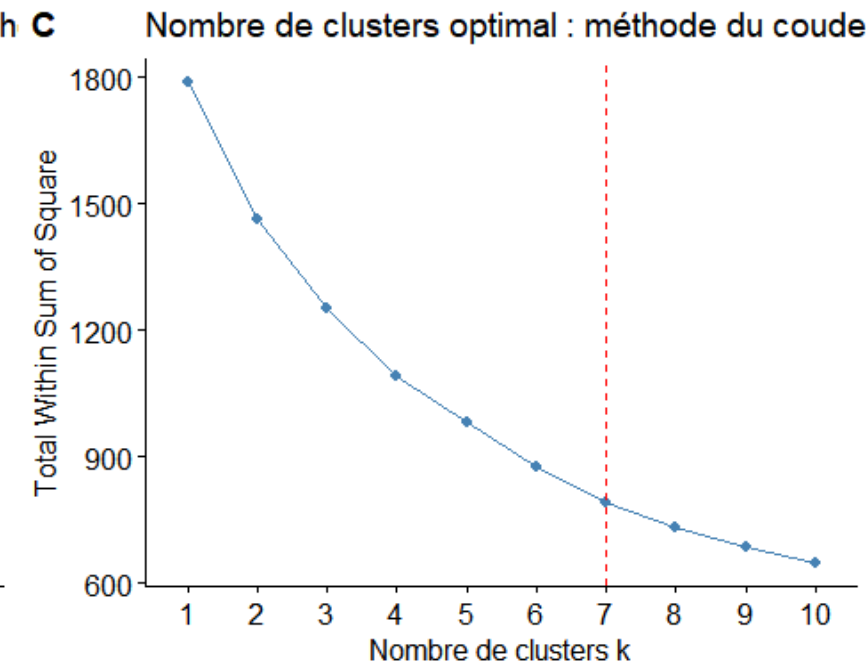
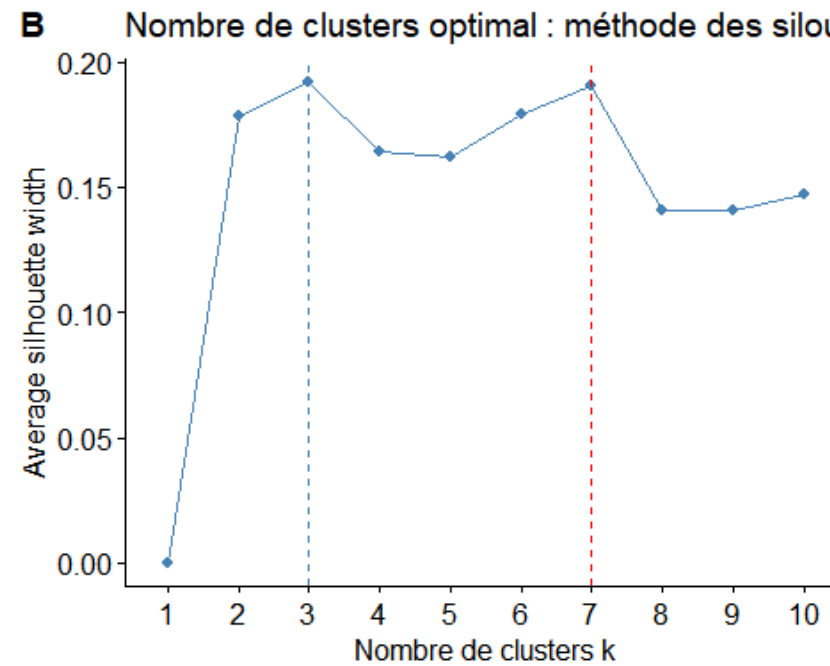
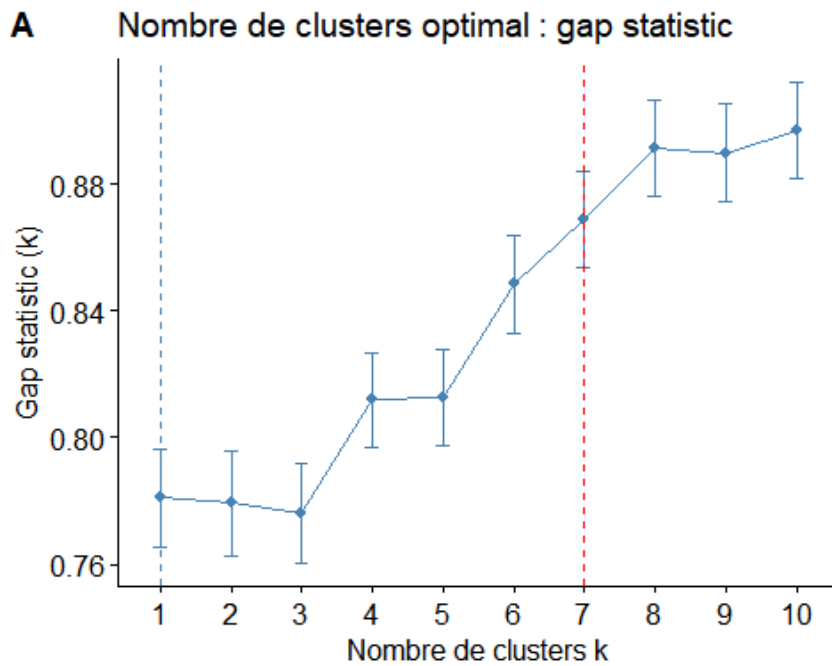
Parmi les 5 clusters :

- Le groupe n°1 sera celui d'intérêt.
- Il comporte 32 pays.

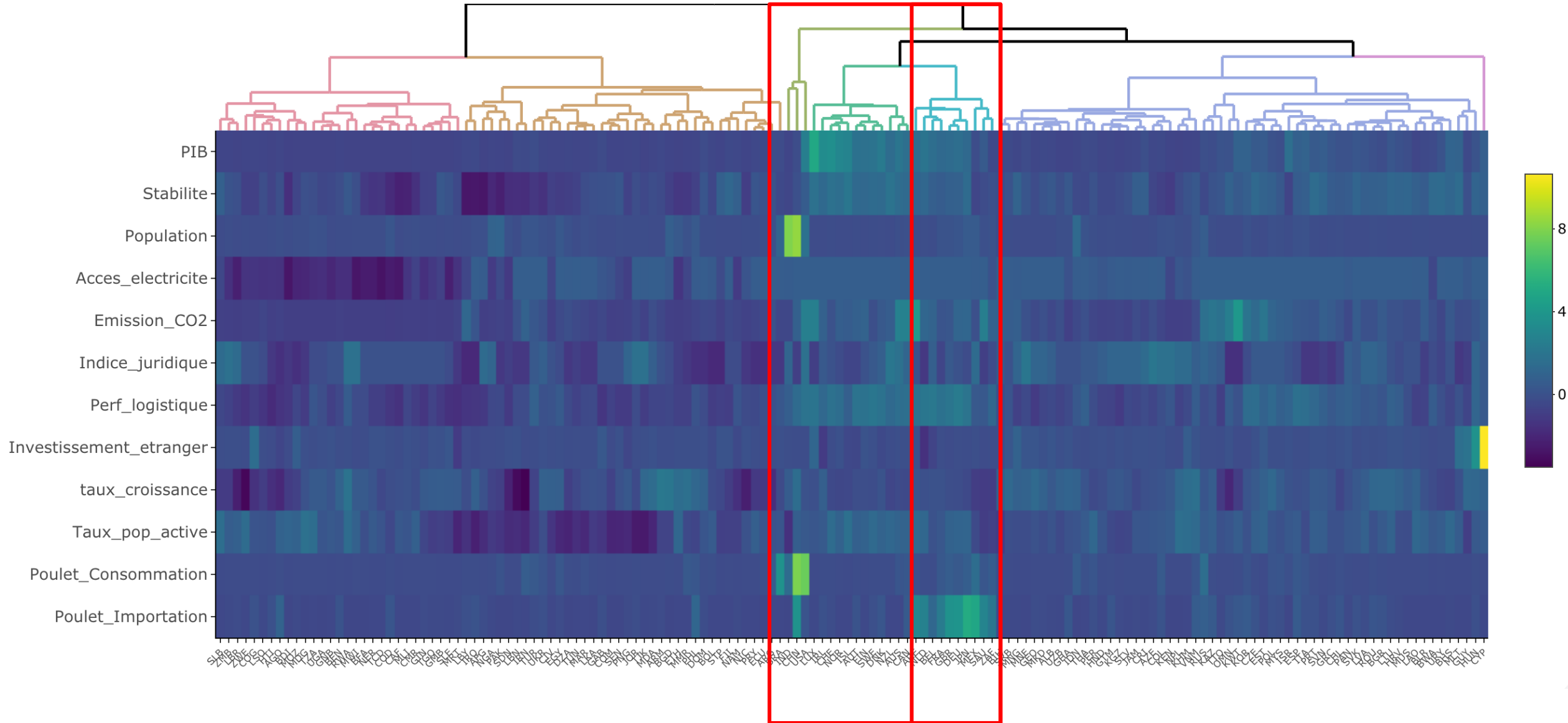
Clustering : La méthode de la Classification Ascendante Hierarchique (CAH)

Principe : Rassembler les observations deux à deux selon un critère de distance

- 1) Attribution d'un groupe à chaque observation
- 2) Détermination des deux groupes les plus proches (méthode de Ward)
- 3) Agrégation en un groupe > Etape 2.

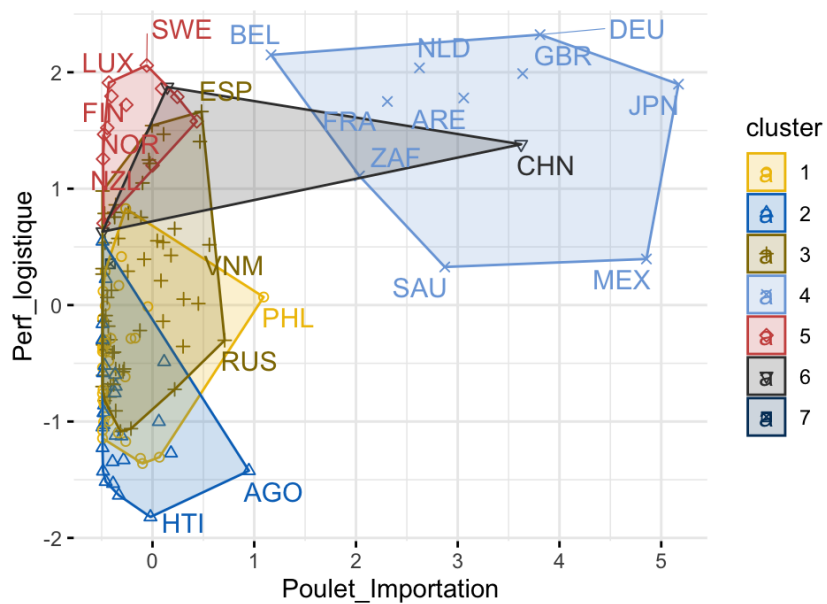


Clustering : La méthode de la Classification Ascendante Hierarchique (CAH)

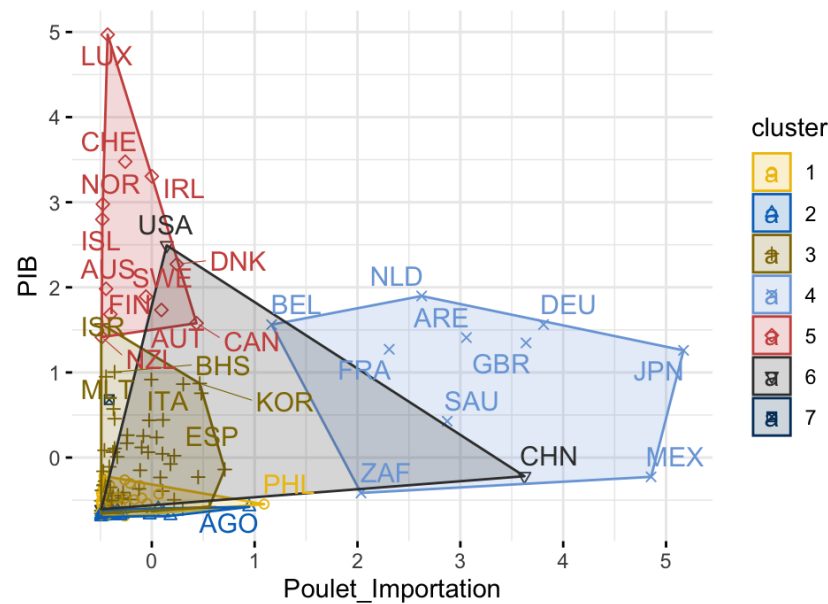


Comparaison des clusters des kmeans et de la CAH

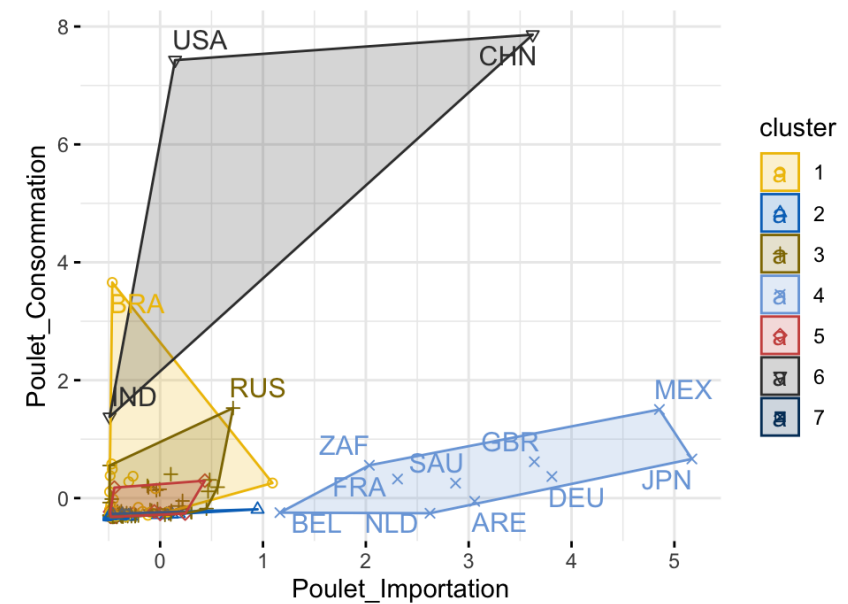
D Clustering sur les variables : CAH



E



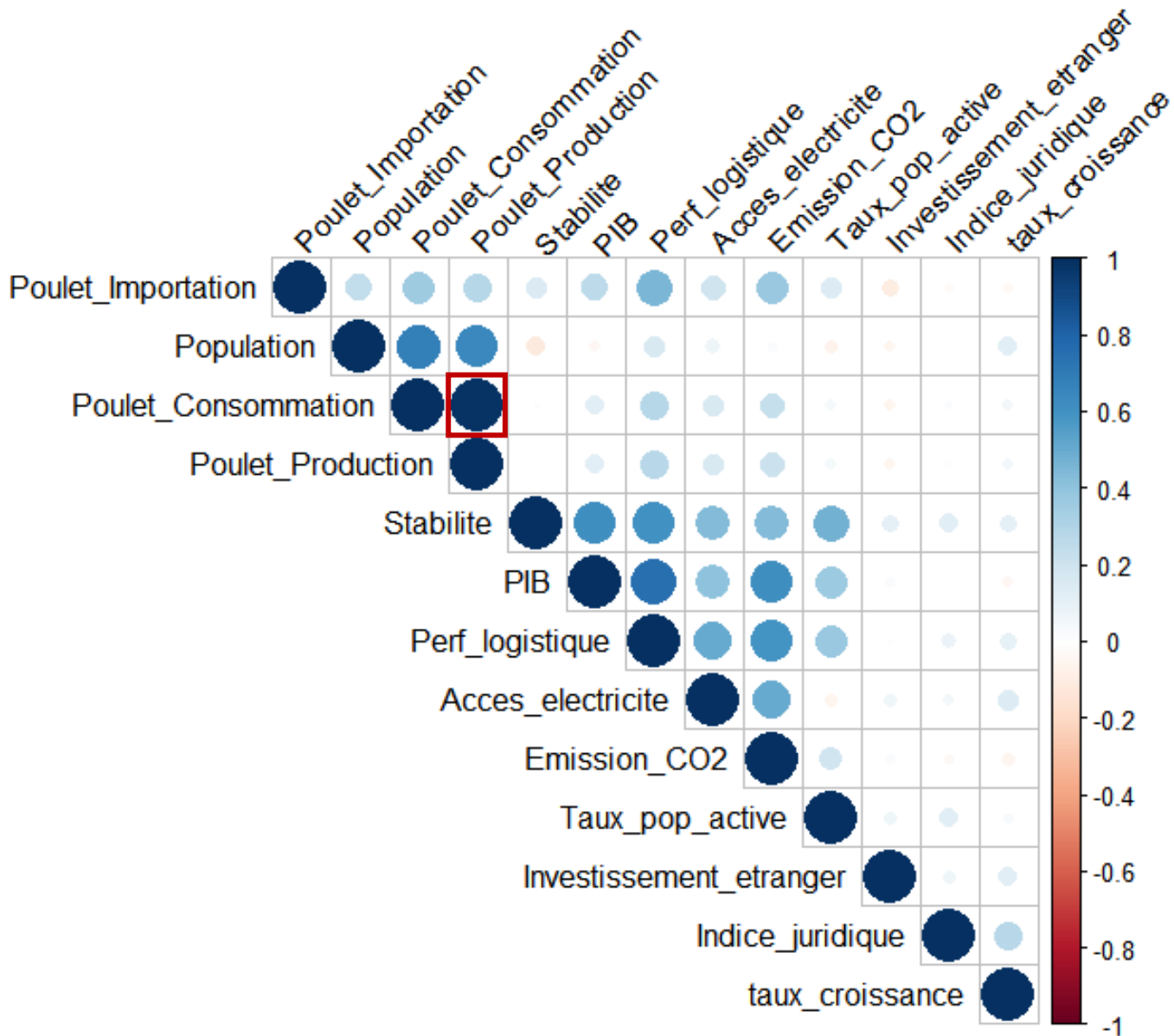
F



Parmi les 7 clusters :

- Le groupe n°4 sera celui d'intérêt.
- Il comporte 10 pays.

Analyse de la corrélation entre les variables



Certaines variables sont corrélées entre elles (positivement ou négativement) et peuvent biaiser l'analyse.

Afin de faciliter l'analyse, une étape possible est l'analyse en composante principale.

Etape 1 : Réduire le nombre de variable à analyser par la méthode d' **Analyse en Composante Principale**.

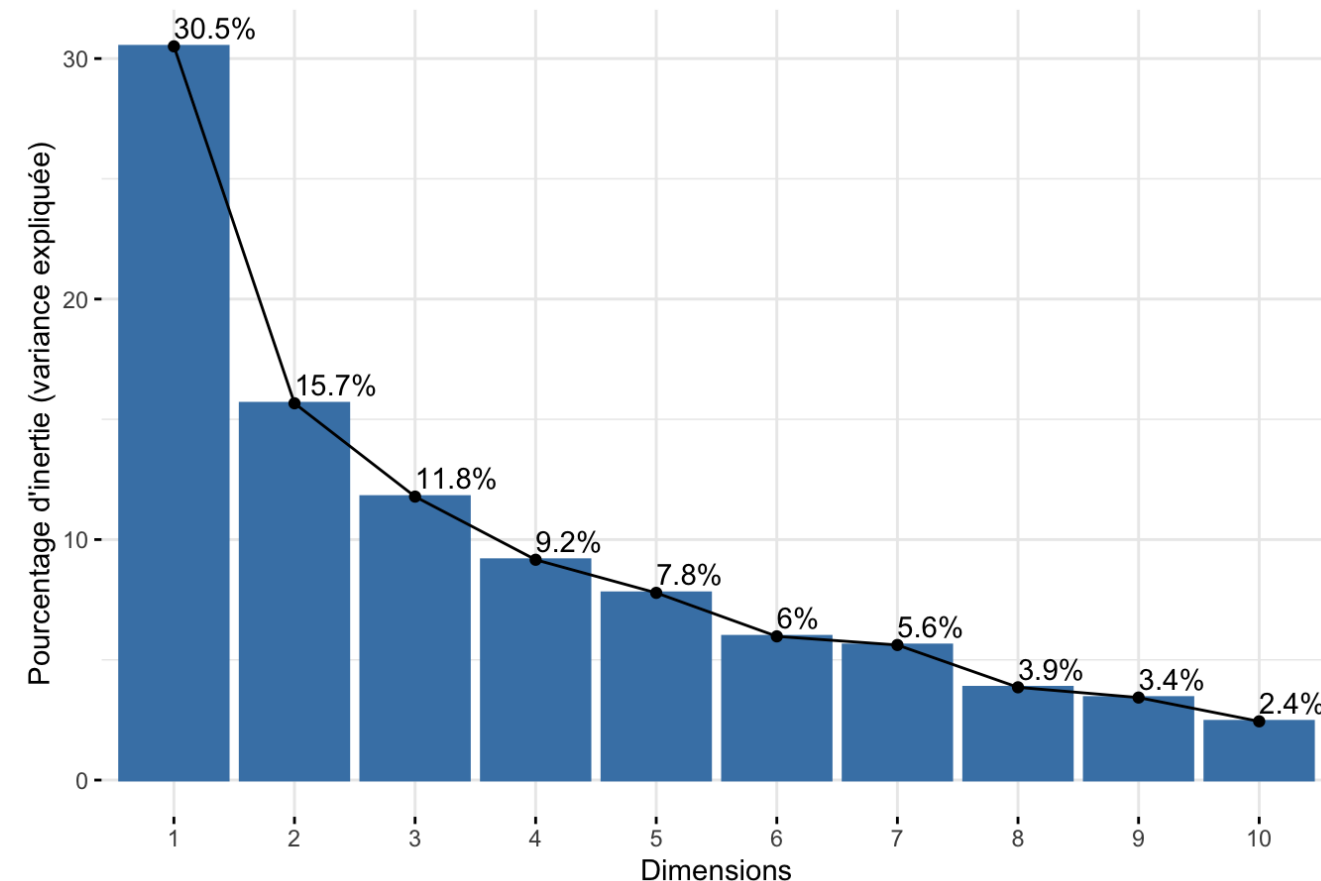
Etape 2 : Effectuer les analyses de regroupement (clustering).

Analyse en Composantes Principales

Principe : Résumer l'information qui est contenue dans de nombreuses variables en un certain nombre d'axes synthétiques (Composantes principales) en gardant le plus d'information possible.

- > Observer les corrélations entre les variables d'origine et les axes de l'ACP
- > Observer les observations (pays) sur les axes synthétiques.

Détermination du nombre d'axes d'intérêts :



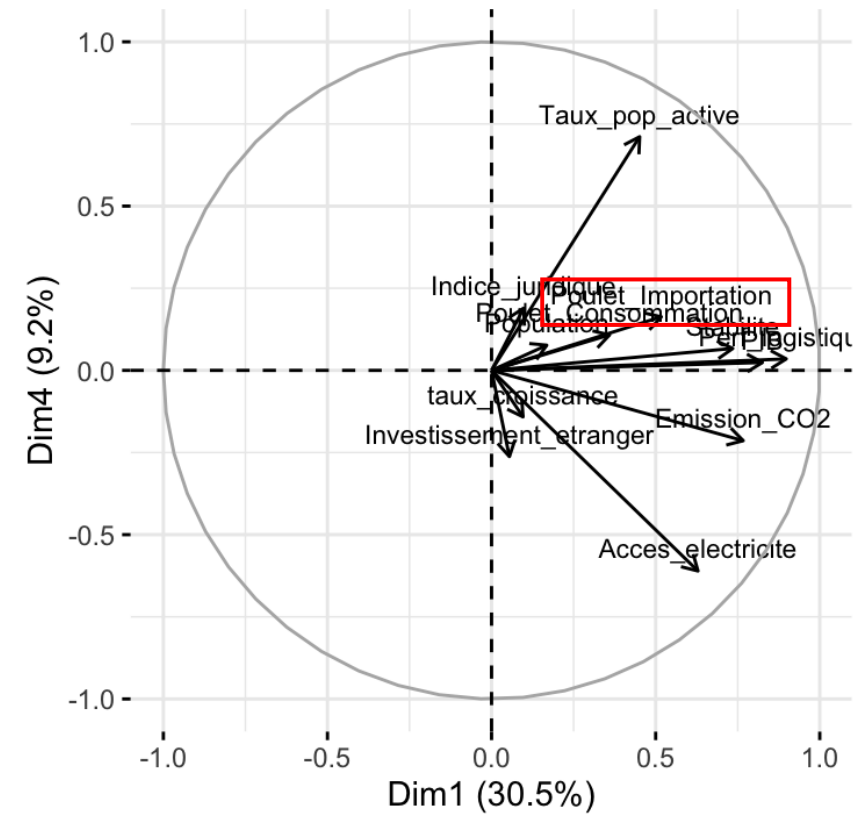
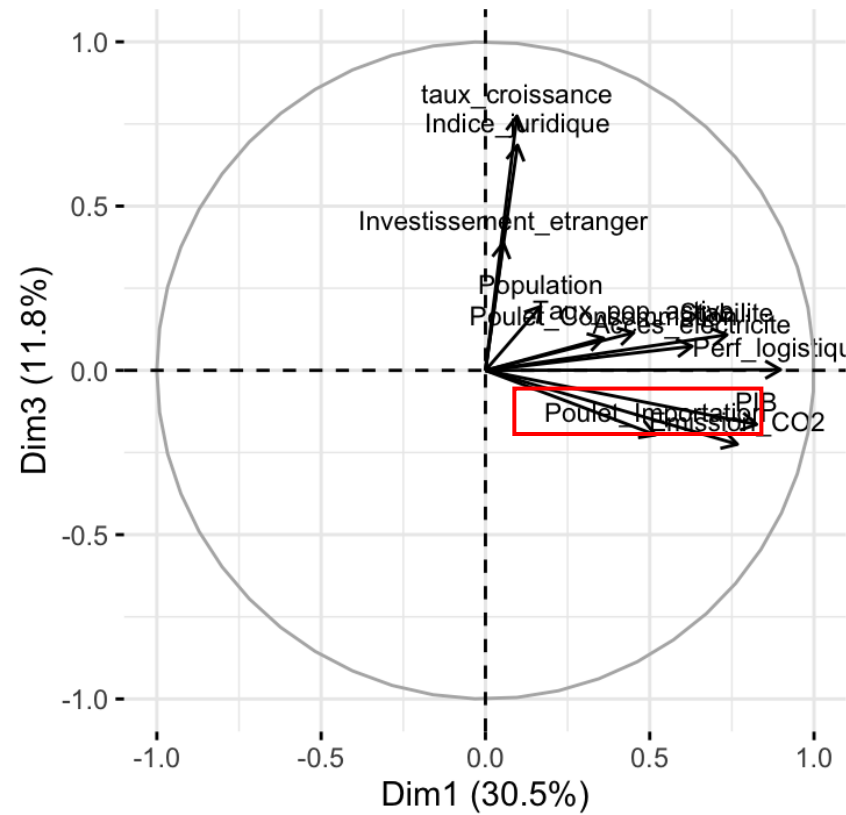
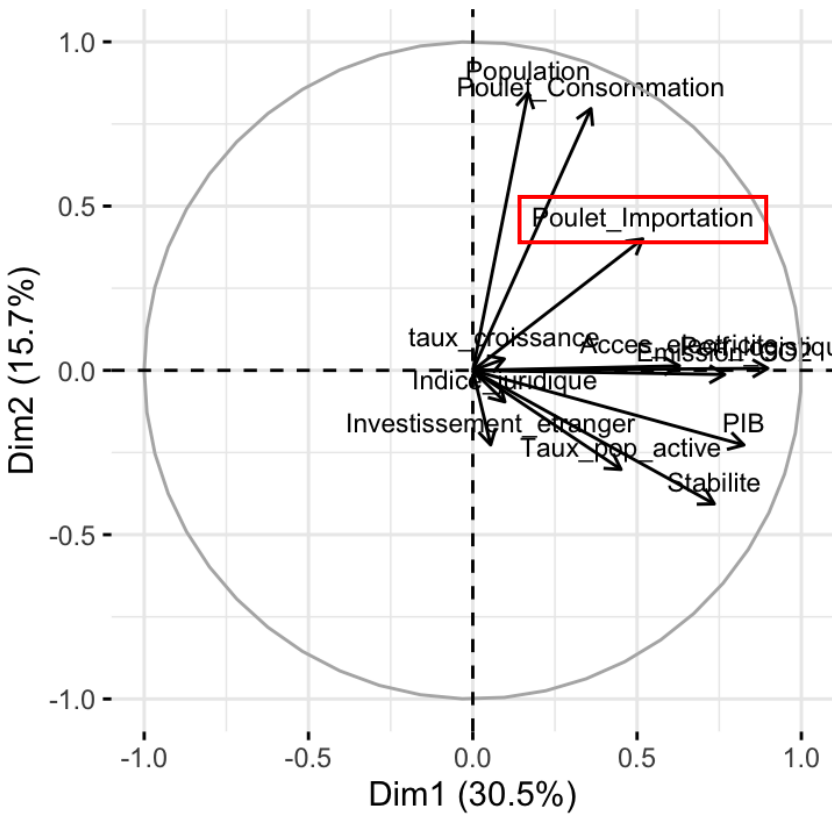
➤ **Critère de Kayser** : on ne garde que les composantes $> (100/p)\%$ où p est le nombre de variables. Ici $(100/p)\% = 8,33$

➤ **Méthode du coude**

Nous analyserons les **4 premières dimensions**, expliquant 67% de la variance totale.

Analyse en Composantes Principales : Cercle des corrélations

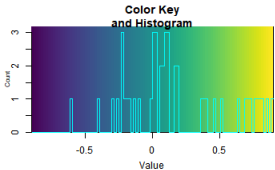
ACP - Cercle des corrélations



➤ L'importation de poulet semble corrélée à la dimension 1 et 2.

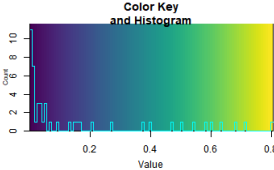
ACP : Corrélation et représentation des variables

Corrélation entre les variables et les composantes principales



Dim.1	Dim.2	Dim.3	Dim.4	Variables
0.83	-0.23	-0.16	0.02	PIB
0.74	-0.41	0.11	0.07	Stabilite
0.17	0.85	0.19	0.08	Population
0.63	0.01	0.07	-0.61	Acces_electricite
0.77	-0.01	-0.22	-0.21	Emission_CO2
0.1	-0.1	0.69	0.19	Indice_juridique
0.9	0.01	0	0.04	Perf_logistique
0.05	-0.23	0.39	-0.26	Investissement_etrange
0.1	0.03	0.77	-0.14	taux_croissance
0.45	-0.3	0.11	0.71	Taux_pop_active
0.36	0.8	0.1	0.11	Poulet_Conso
0.52	0.4	-0.2	0.16	Poulet_Importation

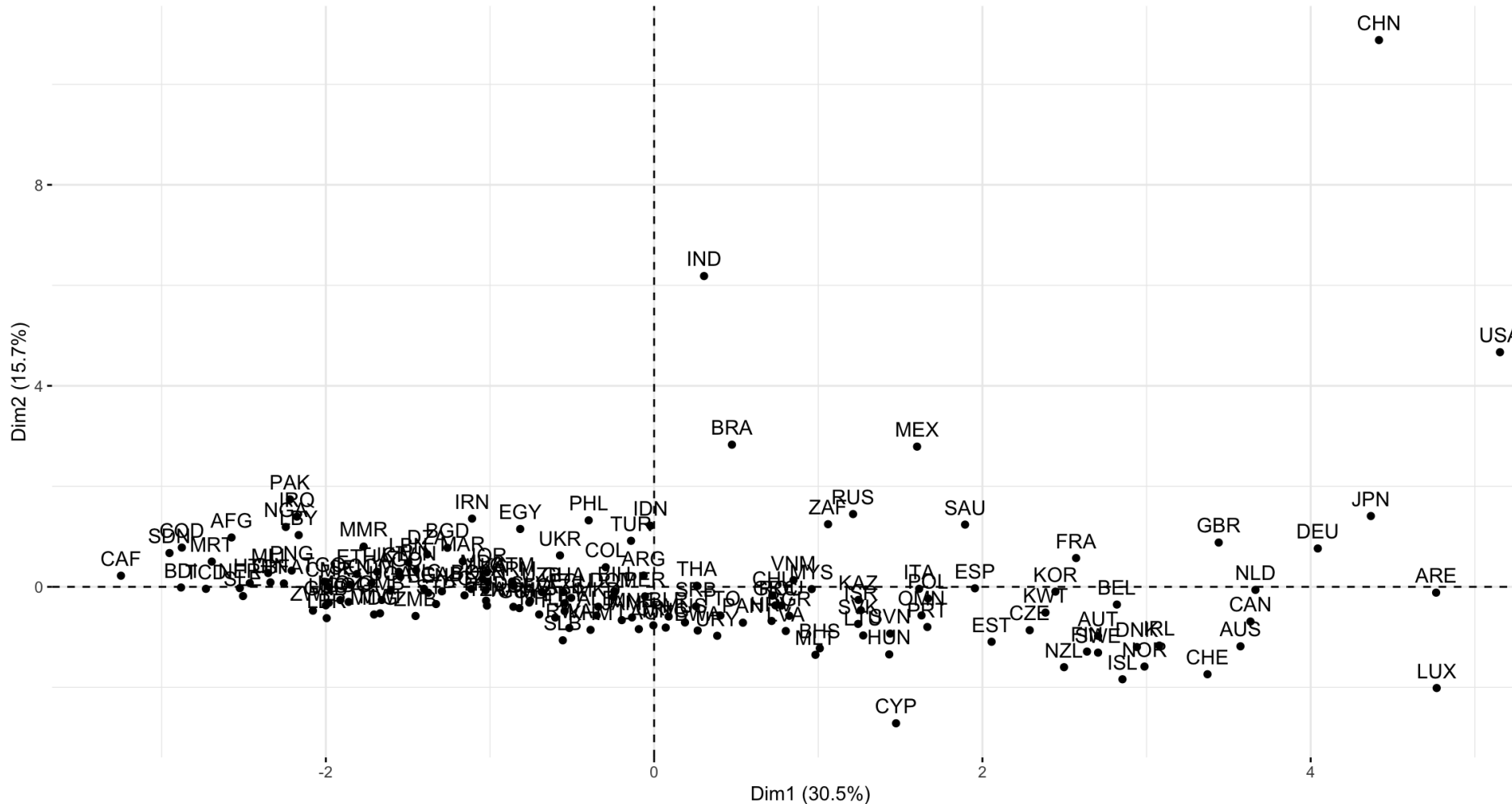
Représentation des variables dans les composantes principales



Dim.1	Dim.2	Dim.3	Dim.4	Variables
0.68	0.05	0.03	0	PIB
0.54	0.17	0.01	0	Stabilite
0.03	0.72	0.04	0.01	Population
0.4	0	0.01	0.37	Acces_electricite
0.59	0	0.05	0.05	Emission_CO2
0.01	0.01	0.47	0.04	Indice_juridique
0.81	0	0	0	Perf_logistique
0	0.05	0.15	0.07	Investissement_etrange
0.01	0	0.6	0.02	taux_croissance
0.2	0.09	0.01	0.51	Taux_pop_active
0.13	0.63	0.01	0.01	Poulet_Conso
0.27	0.16	0.04	0.03	Poulet_Importation

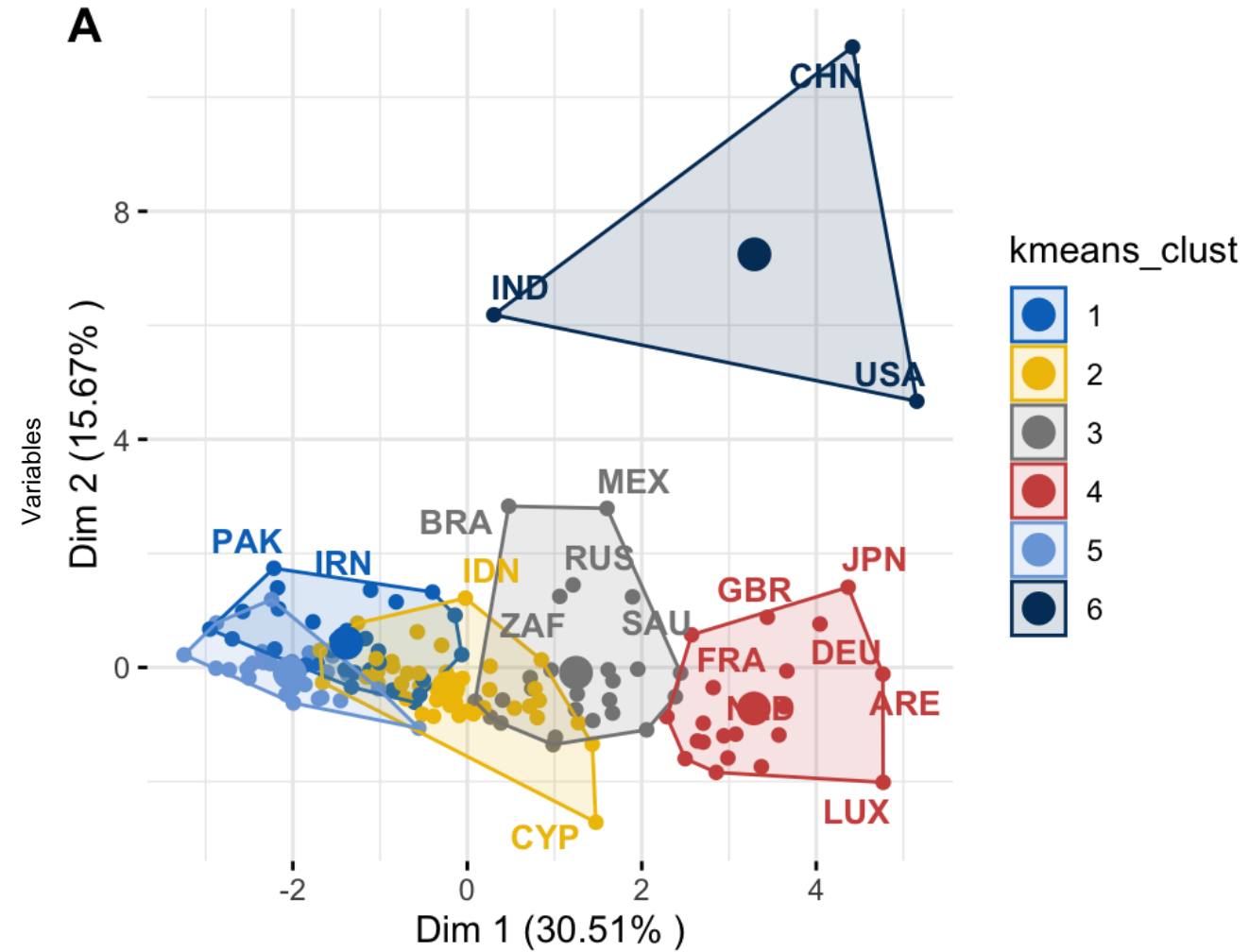
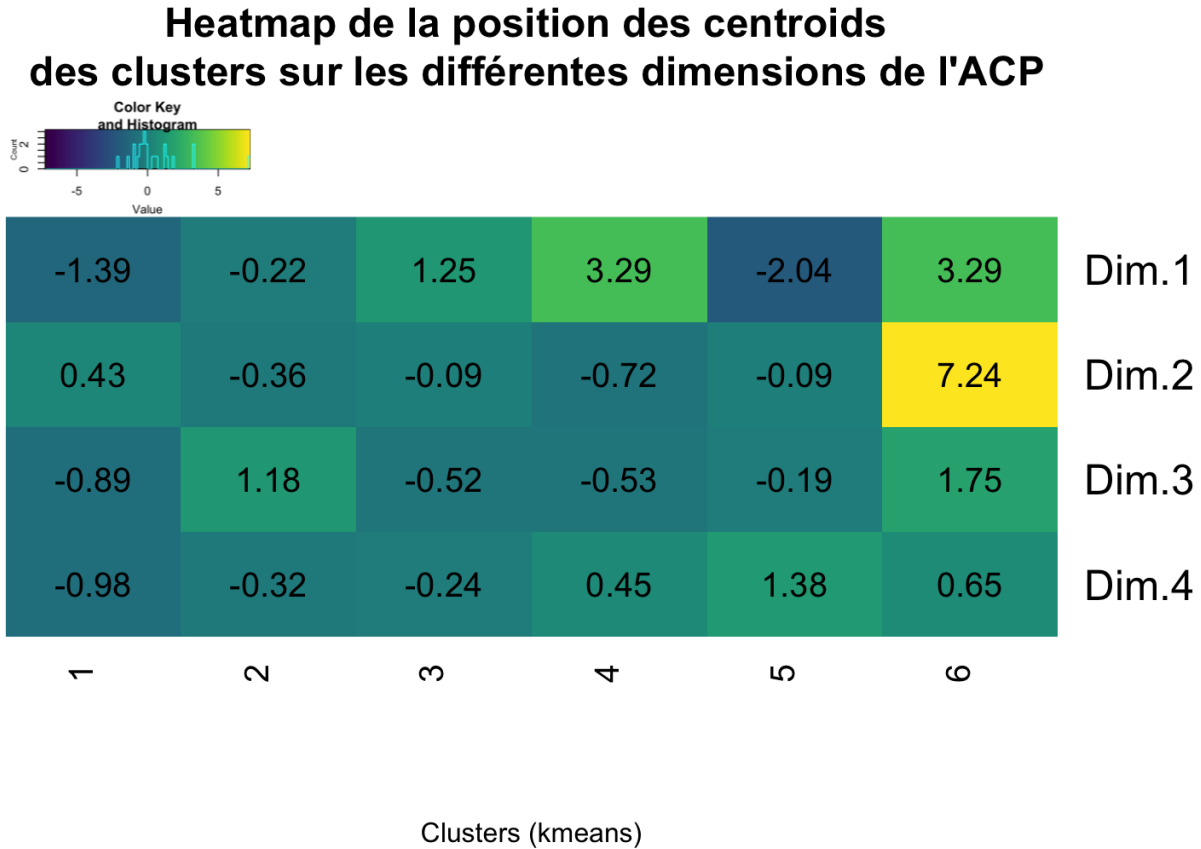
ACP : Représentation des individus (pays)

ACP - Individus



➤ **Kmeans et CAH pour déterminer les groupes de pays cibles**

Méthode des kmeans sur les axes de l'ACP

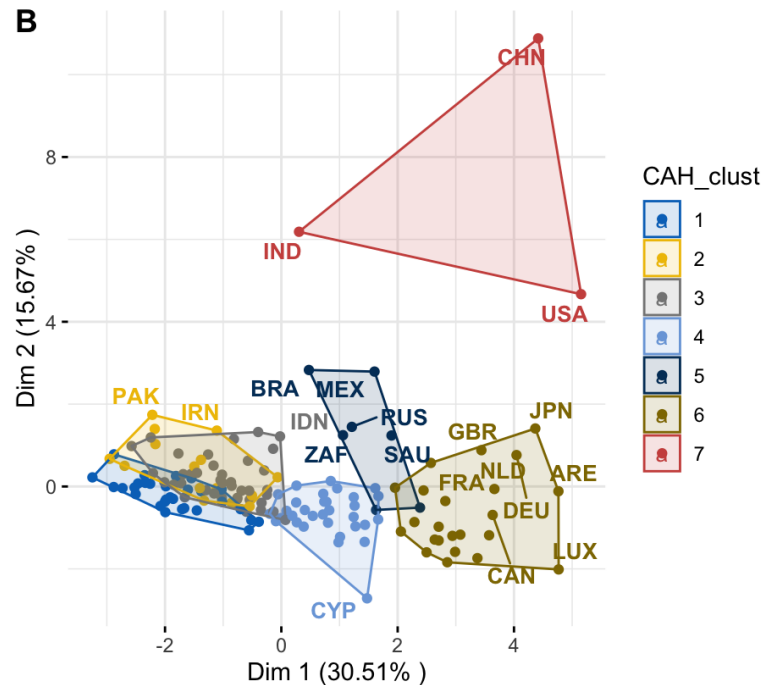
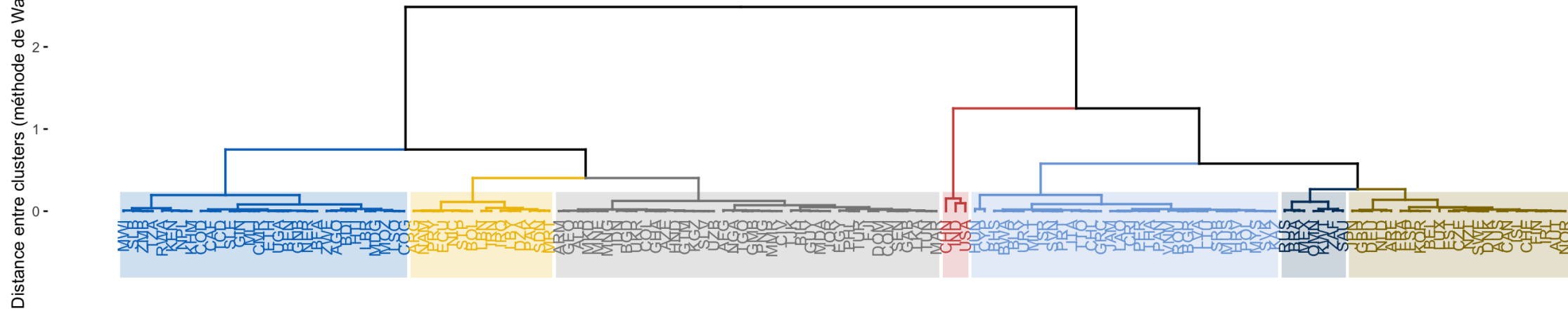


Parmi les 6 clusters :

- Le groupe n°4 sera celui d'intérêt.
- Il comporte 20 pays.

Méthode de la CAH sur les axes de l'ACP

A Dendrogramme



Parmi les 7 clusters :

- Le groupe n°6 sera celui d'intérêt.
- Il comporte 23 pays.

Bilan : Liste des pays cibles communs aux différentes méthodes de clustering

Critère de Sélection : Importation de poulet importante et de bons indicateurs socio-économiques (PIB, stabilité Politique, Perf. Logistique)

	Analyse en composante principale	Variables d'origine
Classification Ascendante hiérarchique	23 pays	10 pays
Kmeans	20 pays	32 pays

6 pays cibles appartenant au même groupe que la France

- United Arab Emirates
- Belgium
- Germany
- **France**
- United Kingdom of Great Britain and Northern Ireland
- Japan
- Netherlands

Bilan : Liste des pays cibles communs aux différentes méthodes de clustering



6 pays cibles appartenant au même groupe que la France

- United Arab Emirates
- Belgium
- Germany
- **France**
- United Kingdom of Great Britain and Northern Ireland
- Japan
- Netherlands

Bonus 2 : Indicateurs des différents pays cibles

