



# ÉTUDE DE MARCHÉ

## EXPORT À L'INTERNATIONAL

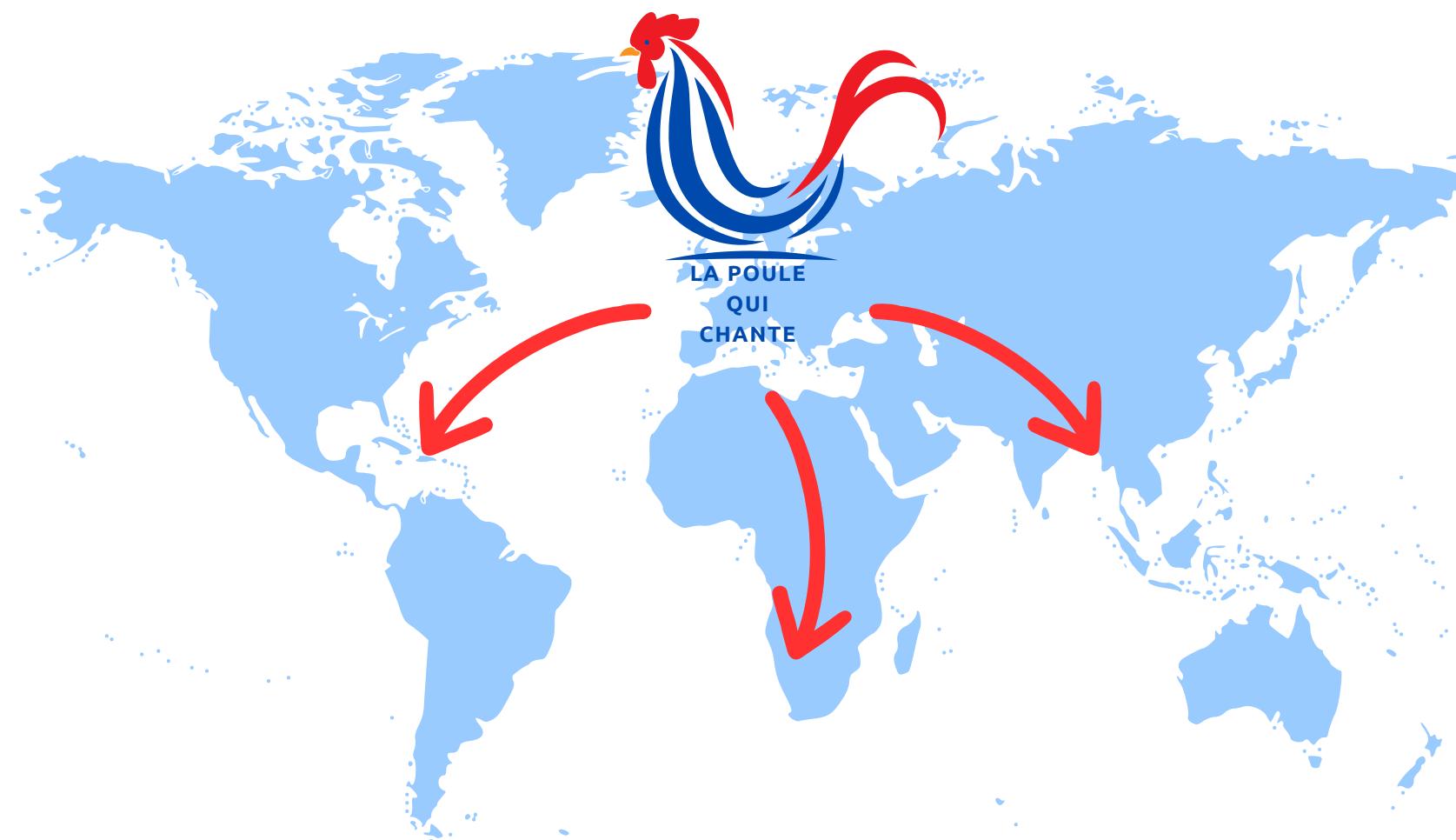
Elodie Mendes

# CONTEXTE



- ▶ **Une entreprise française d'agroalimentaire.**
- ▶ **Son activité principale est l'élevage et la vente de poulets sous le label "Poulet Agriculture Biologique".**

# OBJECTIF



- Le PDG de l'entreprise souhaite évaluer la possibilité de se développer à l'international.
- Identifier des groupes de pays similaires selon des critères économiques, démographiques et commerciaux
- Fournir à l'entreprise une vision claire des segments de pays pour orienter sa stratégie d'exportation.

# PRÉSENTATION DES DONNÉES

## DISPONIBILITÉ ALIMENTAIRE

- ▶  Organisation des Nations Unies pour l'alimentation et l'agriculture
- ▶ Disponibilité de poulets
- ▶ Période: 2010 à 2022
- ▶ 190 pays recensés

## REVENU PAR HABITANT

- ▶  GROUPE DE LA BANQUE MONDIALE
- ▶ Revenu national brut par habitant (PPA USD constant)
- ▶ Période: 2010 à 2023
- ▶ 210 pays recensés

## INDICE STABILITÉ POLITIQUE

- ▶  GROUPE DE LA BANQUE MONDIALE
- ▶ Indicateur de risque de violence ou terrorisme
- ▶ Période: 2010 à 2024
- ▶ 265 pays recensés

# PRÉPARATION DES DONNÉES

## DISPONIBILITÉ ALIMENTAIRE (POULETS)

```
dispo_alimentaire.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15725 entries, 0 to 15724
Data columns (total 12 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Code Domaine      15725 non-null   object  
 1   Domaine          15725 non-null   object  
 2   Code zone (ISO3)  15725 non-null   object  
 3   Zone              15725 non-null   object  
 4   Code Élément      15725 non-null   int64  
 5   Élément           15725 non-null   object  
 6   Code Produit (CPC) 15725 non-null   object  
 7   Produit           15725 non-null   object  
 8   Code année        15725 non-null   int64  
 9   Année              15725 non-null   int64  
 10  Unité              15725 non-null   object  
 11  Valeur             15725 non-null   float64
dtypes: float64(1), int64(3), object(8)
```



```
dispo_alimentaire.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2383 entries, 0 to 2382
Data columns (total 12 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Code zone (ISO3)  2383 non-null   object  
 1   Zone              2383 non-null   object  
 2   Année             2383 non-null   int64  
 3   Exportations - quantité  2383 non-null   float64
 4   Importations - quantité  2383 non-null   float64
 5   Population totale    2383 non-null   float64
 6   Production         2383 non-null   float64
 7   Variation de stock   2383 non-null   float64
 8   Disponibilité Intérieure  2383 non-null   float64
 9   Dispo_kg_pers/an     2383 non-null   float64
 10  tx_dépendance       2383 non-null   float64
 11  Balance commerciale (tonnes) 2383 non-null   float64
dtypes: float64(9), int64(1), object(2)
```

# PRÉPARATION DES DONNÉES

## DISPONIBILITÉ ALIMENTAIRE (INDICATEURS CLÉS)

### ► Disponibilité Intérieure : Importations + Production – Exportations + Variation de stock- Pertes

Cet indicateur permet d'évaluer quelle quantité de poulets est disponible dans un pays.

### ► Balance commerciale (tonnes): Exportations – Importations

Si le résultat est positif => excédentaire donc un pays qui exporte plus qu'il importe.

Si le résultat est négatif => déficitaire donc un pays qui importe plus qu'il exporte.

### ► Taux de dépendance aux importations (%): Quantité importée/ Disponibilité intérieure

Part des importations sur la disponibilité intérieure.

Plus le % est élevé, plus le pays est dépendant des importations de poulets.

# PRÉPARATION DES DONNÉES

## REVENU PAR HABITANT

```
revenu_habitant.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2940 entries, 0 to 2939
Data columns (total 12 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Code Domaine      2940 non-null    object  
 1   Domaine          2940 non-null    object  
 2   Code zone (ISO3)  2940 non-null    object  
 3   Zone              2940 non-null    object  
 4   Code Élément      2940 non-null    int64  
 5   Élément          2940 non-null    object  
 6   Code Produit      2940 non-null    int64  
 7   Produit          2940 non-null    object  
 8   Code année        2940 non-null    int64  
 9   Année             2940 non-null    int64  
 10  Unité             2940 non-null    object  
 11  Valeur            2940 non-null    float64
dtypes: float64(1), int64(4), object(7)
```



```
revenu_habitant.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2940 entries, 0 to 2939
Data columns (total 3 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Code zone (ISO3)  2940 non-null    object  
 1   Année             2940 non-null    int64  
 2   Revenu par habitant (US$) 2940 non-null    float64
dtypes: float64(1), int64(1), object(1)
```

# PRÉPARATION DES DONNÉES

## INDICE DE STABILITÉ POLITIQUE

```
stabilite_politique.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 266 entries, 0 to 265
Data columns (total 18 columns):
 #   Column           Non-Null Count  Dtype  
---  --  
 0   Country Name    264 non-null    object  
 1   Country Code    266 non-null    object  
 2   Indicator Name  266 non-null    object  
 3   2010             204 non-null    float64 
 4   2011             205 non-null    float64 
 5   2012             205 non-null    float64 
 6   2013             205 non-null    float64 
 7   2014             205 non-null    float64 
 8   2015             205 non-null    float64 
 9   2016             205 non-null    float64 
 10  2017             205 non-null    float64 
 11  2018             205 non-null    float64 
 12  2019             205 non-null    float64 
 13  2020             205 non-null    float64 
 14  2021             205 non-null    float64 
 15  2022             205 non-null    float64 
 16  2023             205 non-null    float64 
 17  2024             0 non-null     float64 
dtypes: float64(15), object(3)
```



```
stabilite_politique.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3990 entries, 0 to 3989
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
---  --  
 0   index            3990 non-null   int64  
 1   Country Name    3960 non-null   object  
 2   Code zone (ISO3) 3990 non-null   object  
 3   Indicator Name  3990 non-null   object  
 4   Année            3990 non-null   object  
 5   Indice de stabilité politique 2869 non-null   float64 
dtypes: float64(1), int64(1), object(4)
```

# PRÉPARATION DES DONNÉES

## TABLE DE REFÉRENCE

La table 'disponibilité alimentaire' a été la table de référence pour effectuer les fusions des 2 autres fichiers (revenu et stabilité politique).

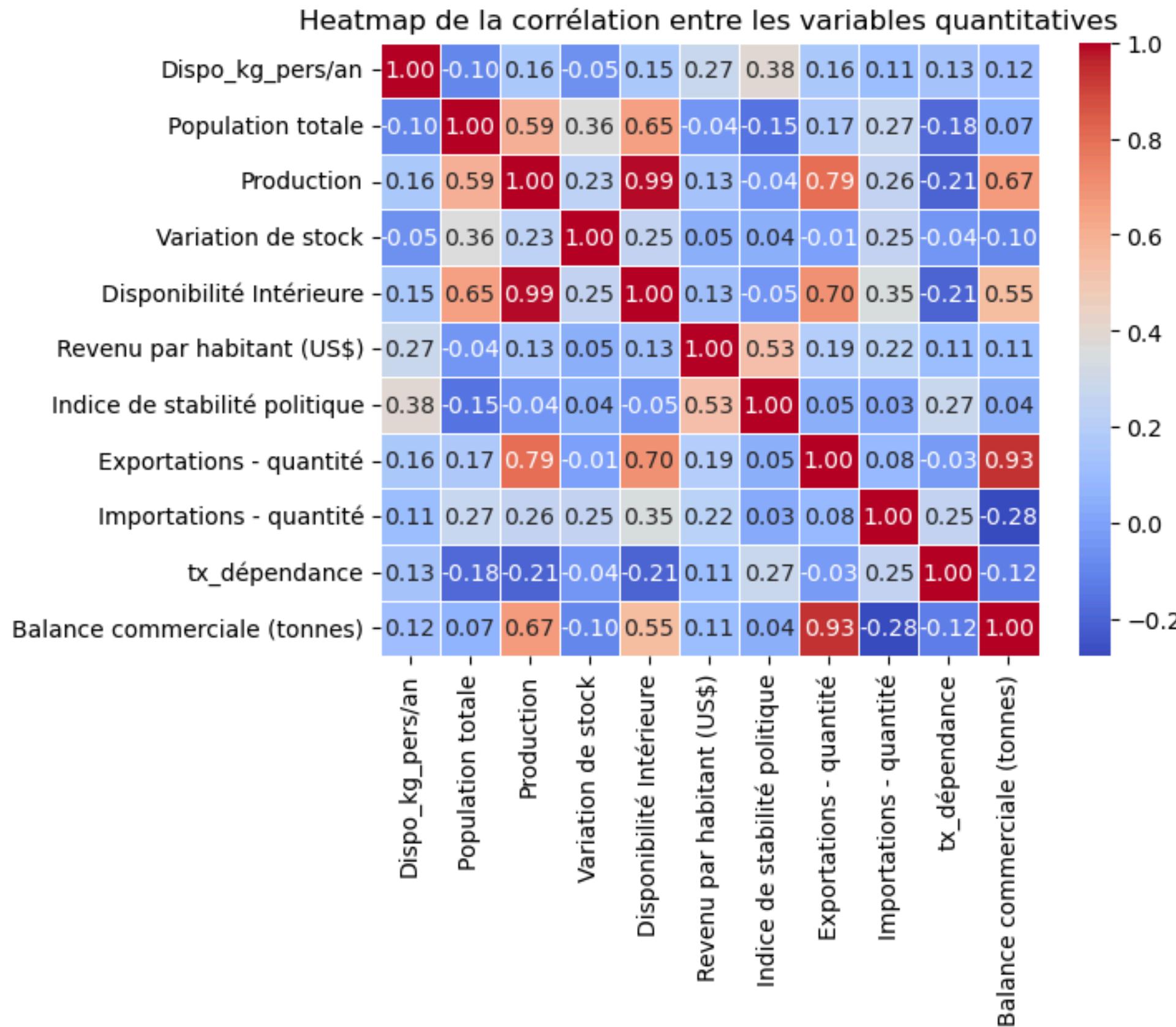
## CLÉ DE JOINTURE: CODE ISO + ANNÉE

J'ai gardé le nom des pays uniquement sur le fichier des disponibilités pour éviter les redondances et j'ai utilisé les clés primaires composé du code iso de chaque pays et l'année

`df_merge.info()`

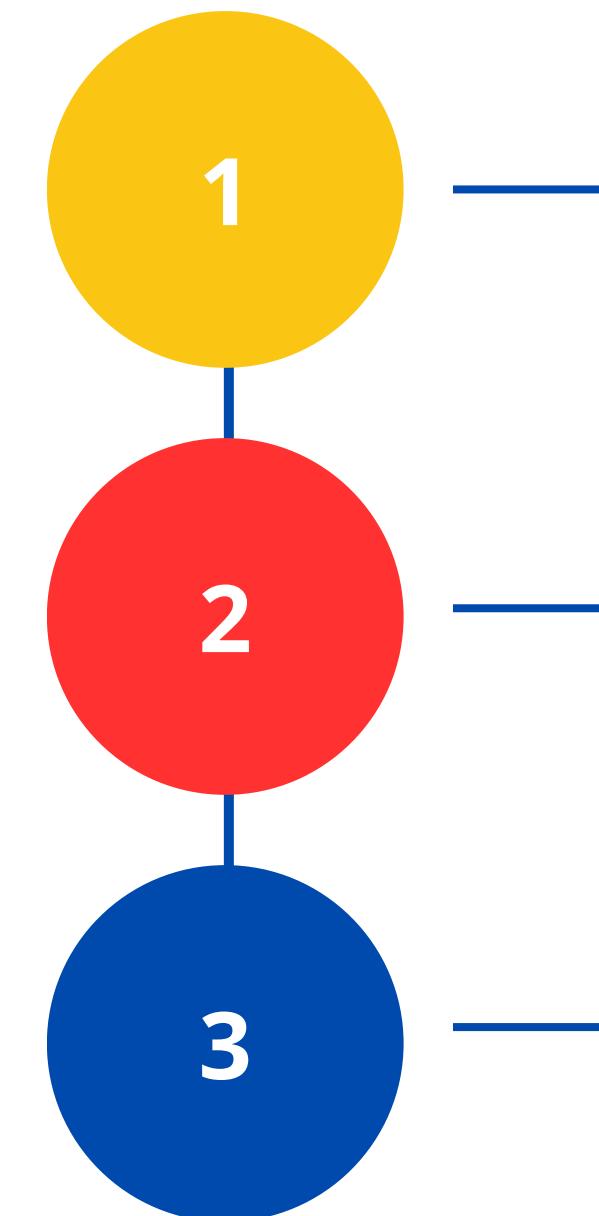
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 188 entries, 0 to 187
Data columns (total 13 columns):
 #   Column           Non-Null Count   Dtype  
--- 
 0   Zone             188 non-null     object  
 1   Code zone (ISO3) 188 non-null     object  
 2   Balance commerciale (tonnes) 188 non-null     float64 
 3   Population totale 188 non-null     float64 
 4   Production       188 non-null     float64 
 5   Disponibilité Intérieure 188 non-null     float64 
 6   Dispo_kg_pers/an 188 non-null     float64 
 7   tx_dépendance    188 non-null     float64 
 8   Revenu par habitant (US$) 188 non-null     float64 
 9   Indice de stabilité politique 188 non-null     float64 
 10  Variation de stock 188 non-null     float64 
 11  Exportations - quantité 188 non-null     float64 
 12  Importations - quantité 188 non-null     float64 
dtypes: float64(11), object(2)
```

# ANALYSE DE LA CORRÉLATION DES VARIABLES QUANTITATIVES



# ANALYSE EN COMPOSANTES PRINCIPALES

L'ACP est une méthode de réduction de la dimensionnalité qui est souvent utilisée pour transformer un grand ensemble de variables en un ensemble plus petit qui contient encore la plupart des renseignements dans le grand ensemble.



Chaque variable est réduite et centrée pour les ramener sur une même échelle d'unité. Ainsi, les variables pourront être synthétiser en un axe factoriel (combinaison linéaire).

Choisir le nombre d'axe factoriel qui va maximiser le nombre d'informations disponibles (variance expliquée) grâce à des méthodes comme la règle du coude.

Représentation du cercle de corrélation avec les pays projetés dessus.

# CHOIX DU NOMBRE D'AXE FACTORIEL

Variance expliquée par dimension

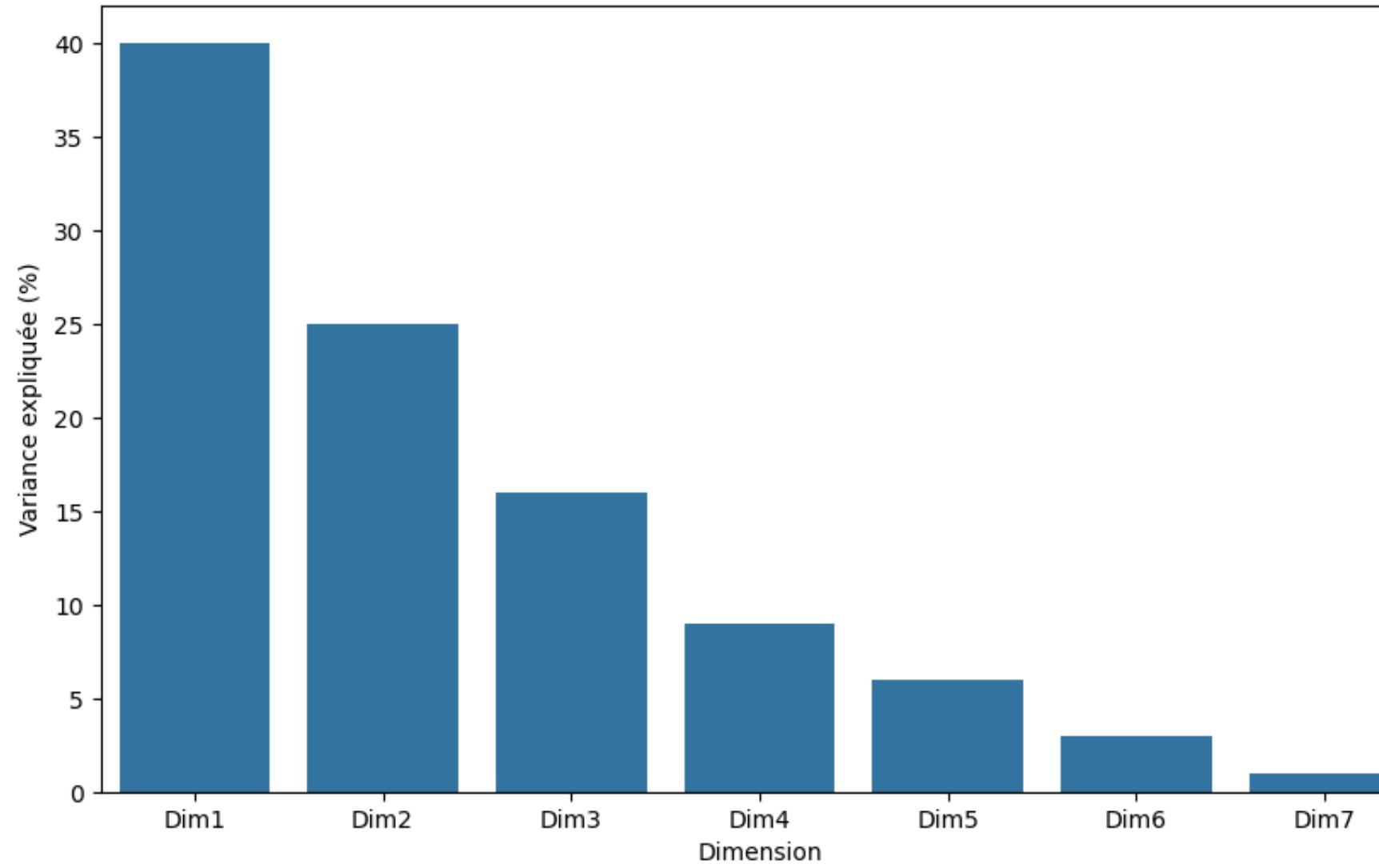
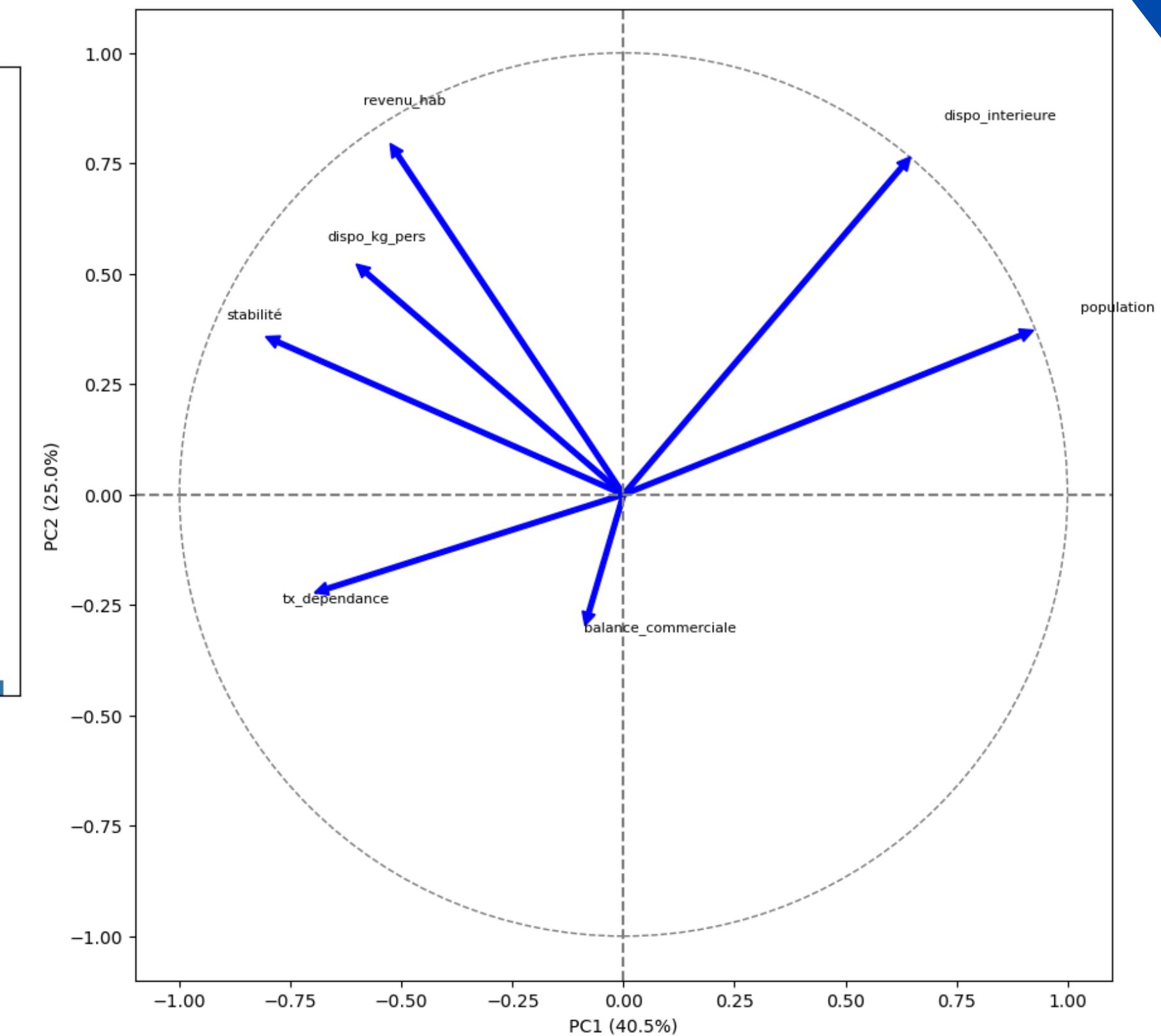


Tableau des cos<sup>2</sup> (qualité de représentation des variables)

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	Qualité (PC1+PC2)
balance_commerciale	0.006	0.071	0.822	0.009	0.091	0.000	0.000	0.078
revenu_hab	0.258	0.589	0.004	0.052	0.004	0.091	0.001	0.847
stabilité	0.603	0.119	0.018	0.141	0.048	0.072	0.000	0.722
tx_dépendance	0.443	0.045	0.277	0.012	0.217	0.006	0.000	0.488
dispo_kg_pers	0.334	0.251	0.019	0.383	0.002	0.009	0.002	0.585
dispo_interieure	0.391	0.547	0.000	0.000	0.033	0.010	0.017	0.938
population	0.798	0.129	0.003	0.021	0.016	0.012	0.021	0.927

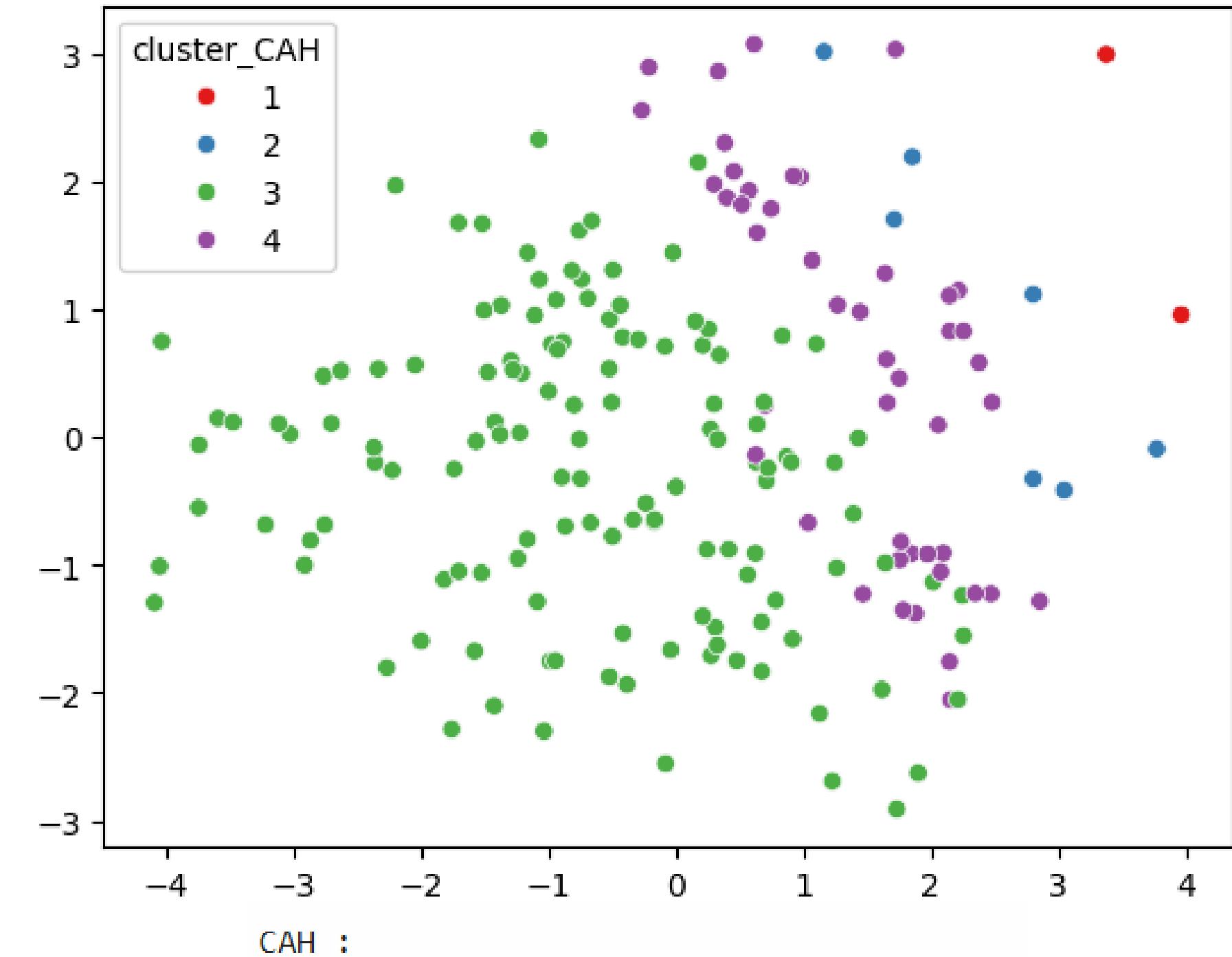
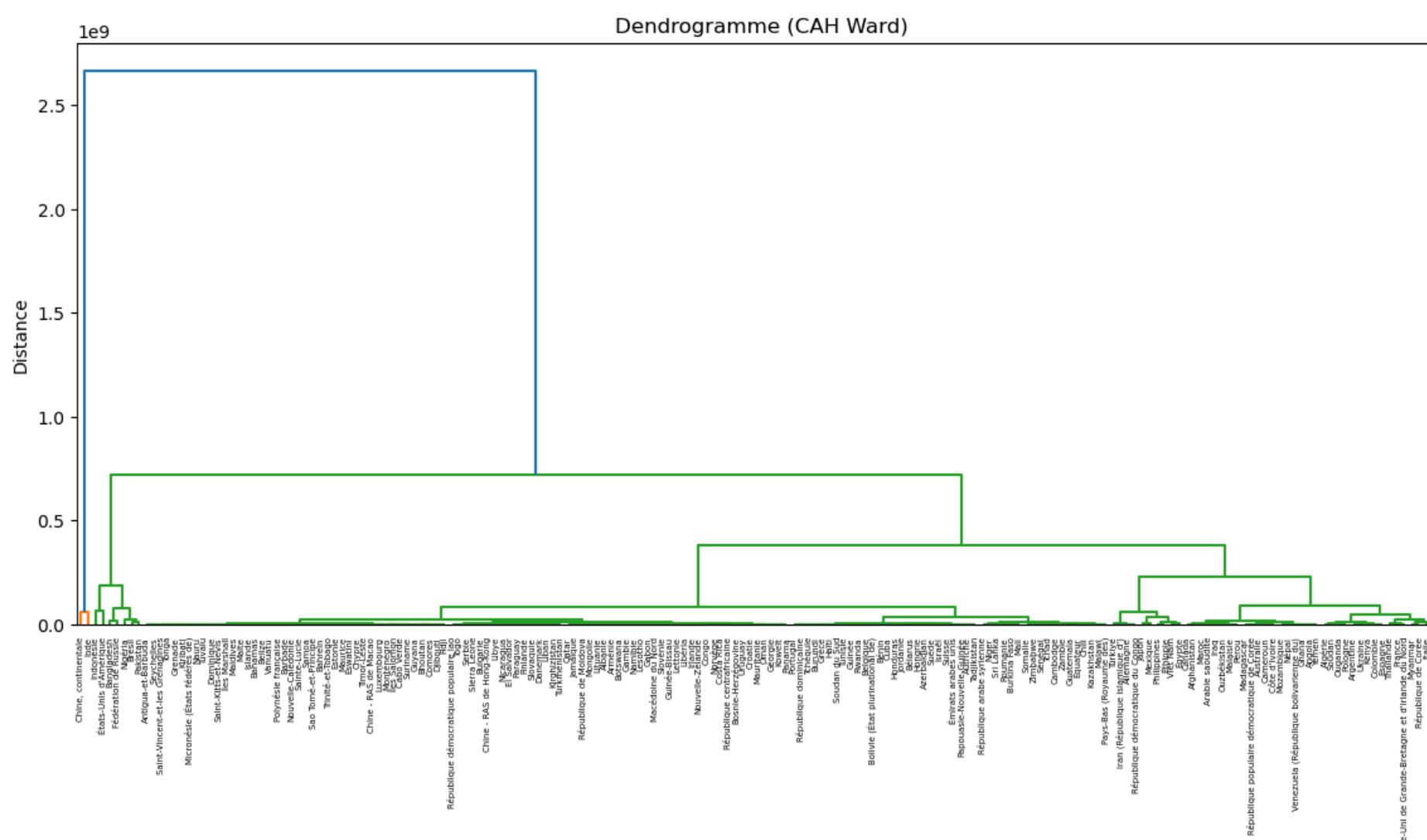
# CERCLE DE CORRÉLATION

Cercle de corrélation



# CLUSTERING PAR CLASSIFICATION ASCENDANTE HIÉARCHIQUE

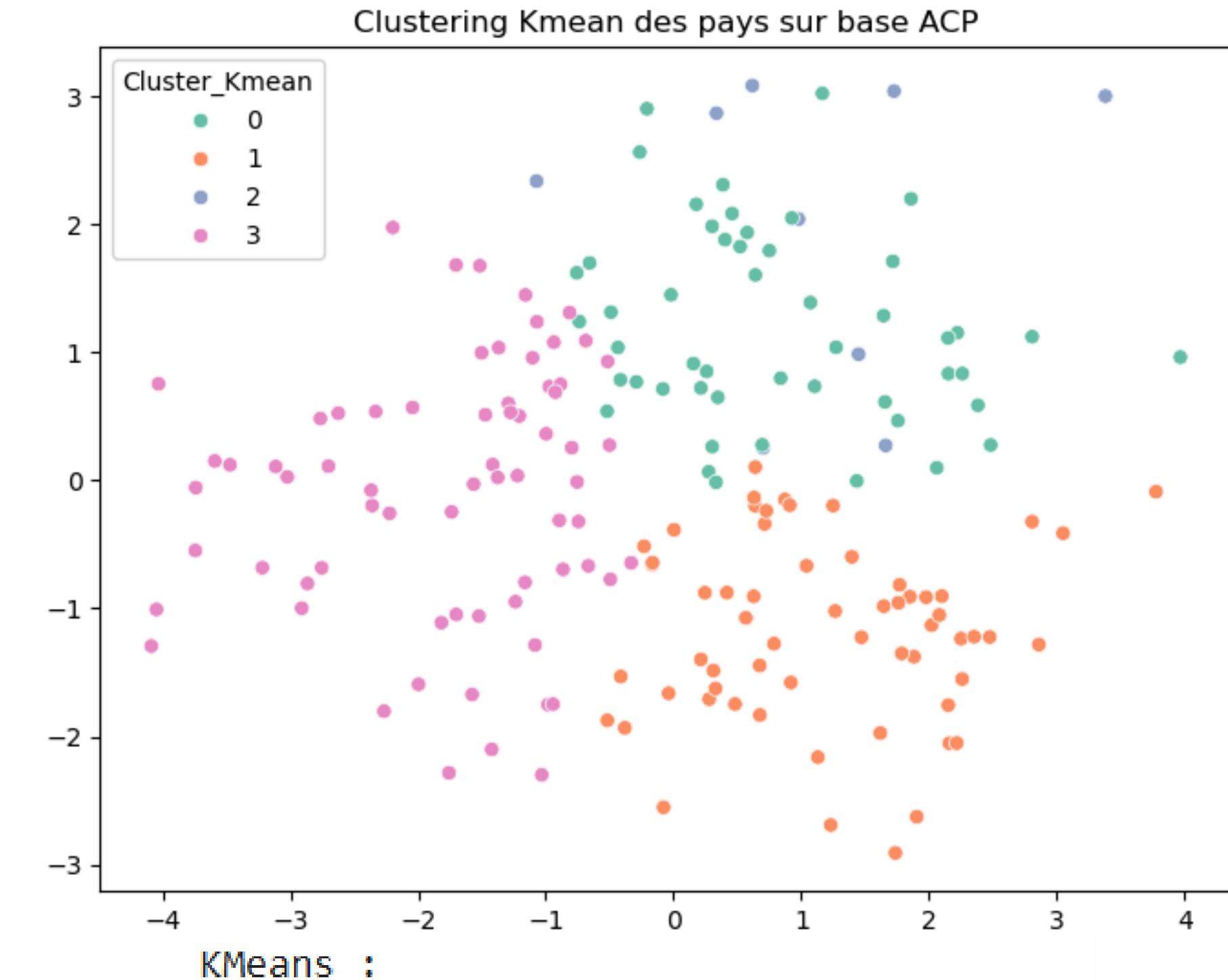
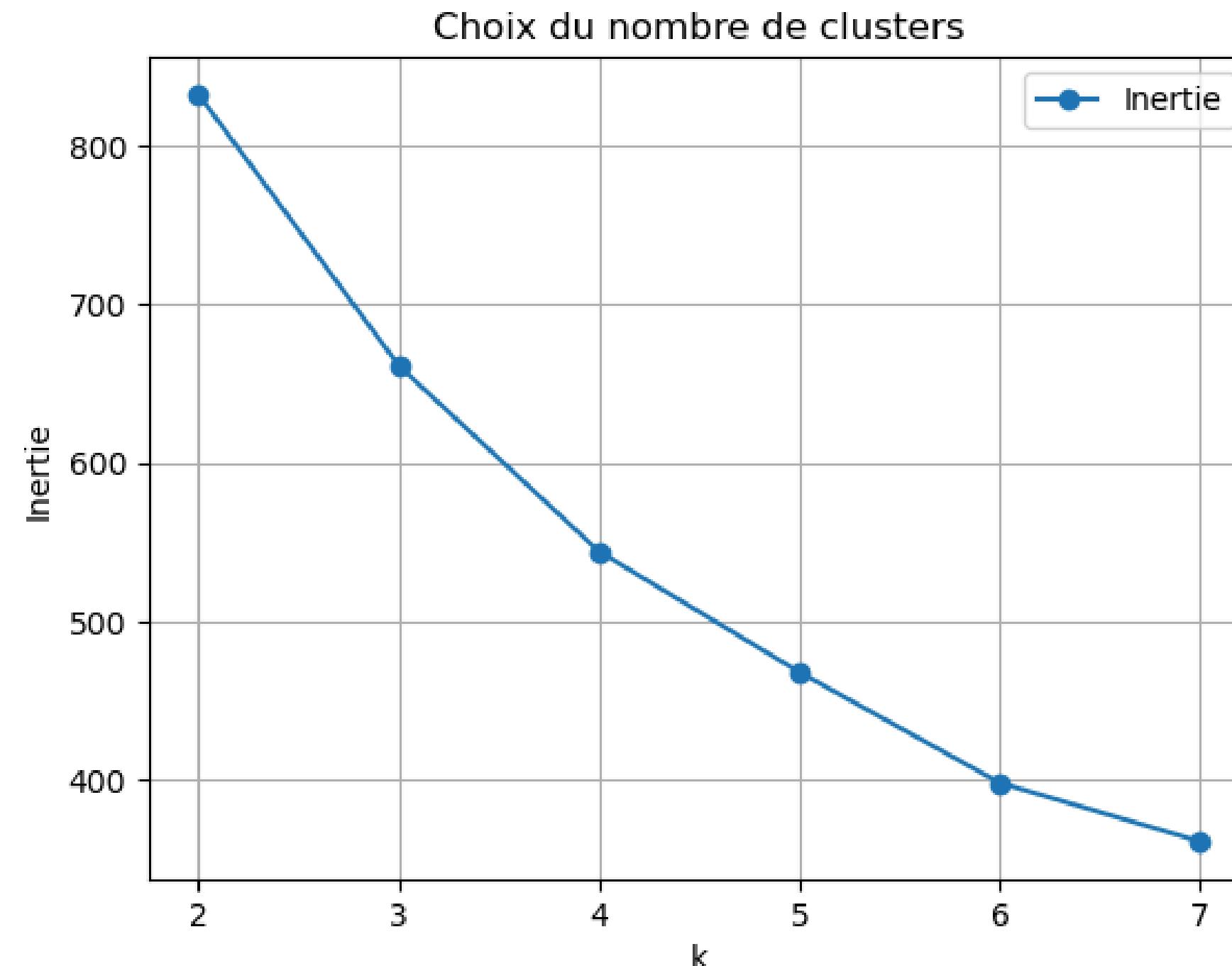
Clustering CAH des pays sur base ACP



Silhouette = 0.22832600700728414

Calinski-Harabasz = 52.304217097682596

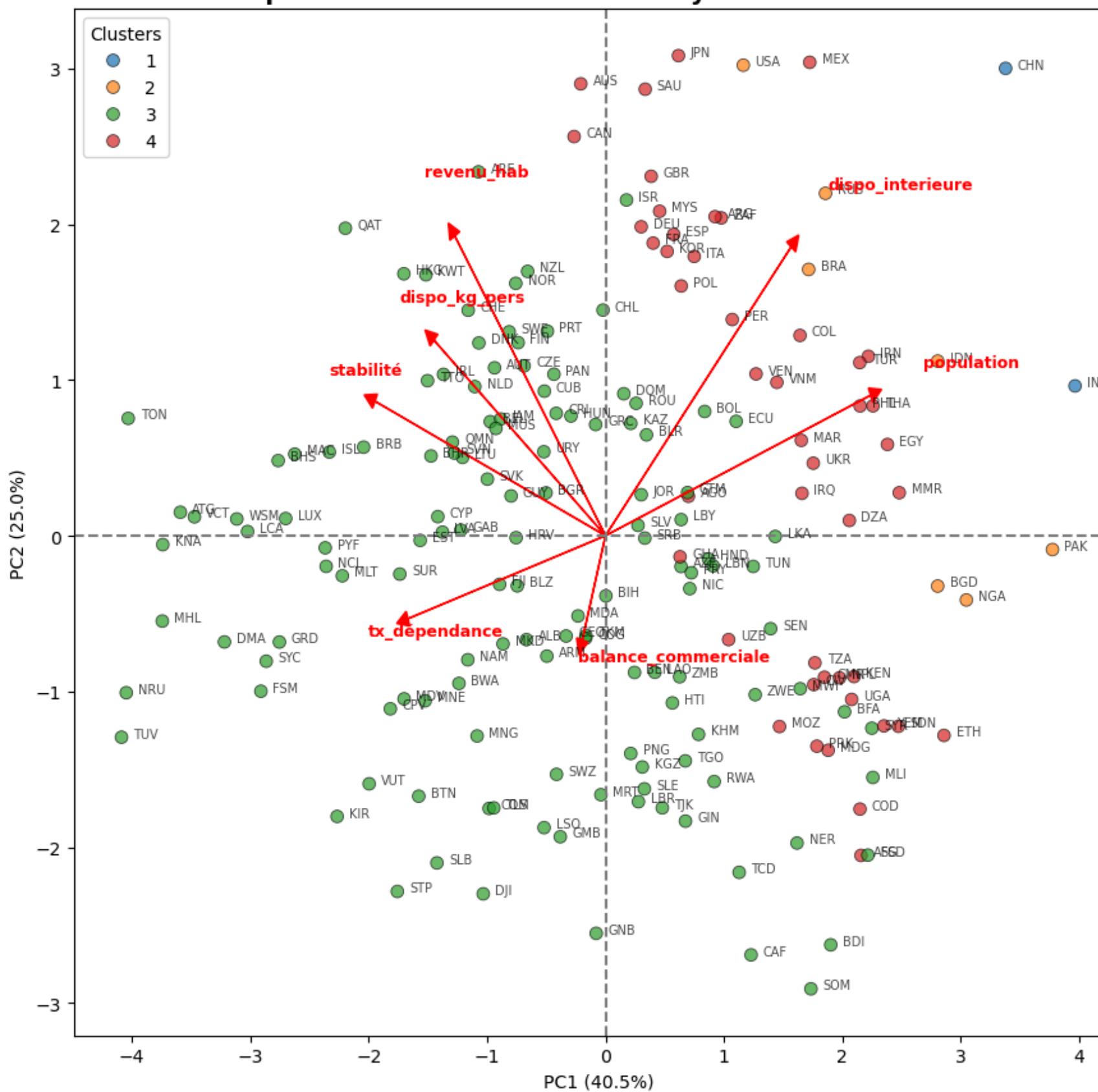
# CLUSTERING PAR K-MEANS



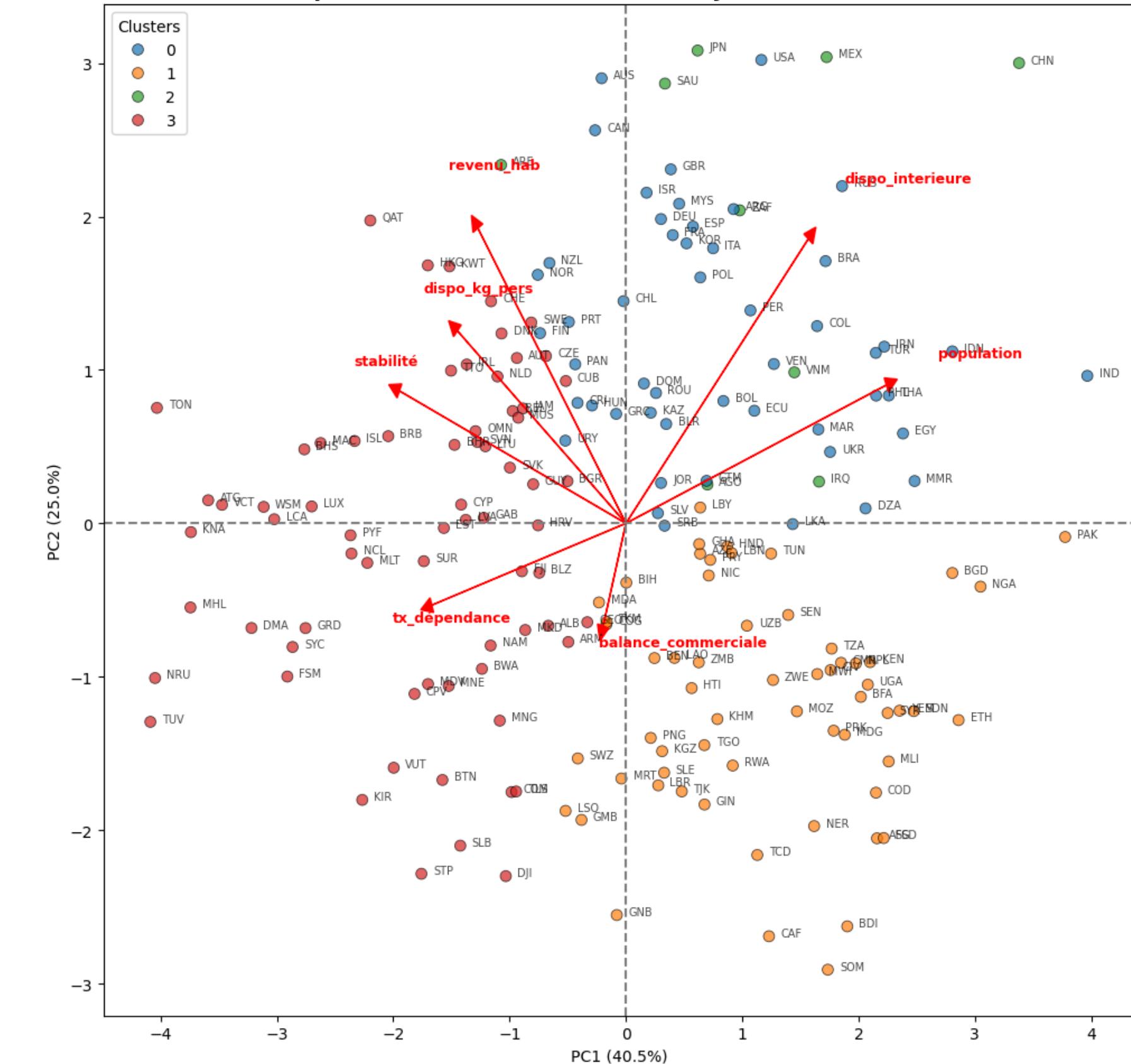
Silhouette = 0.24500683620882133

Calinski-Harabasz = 58.66450878733886

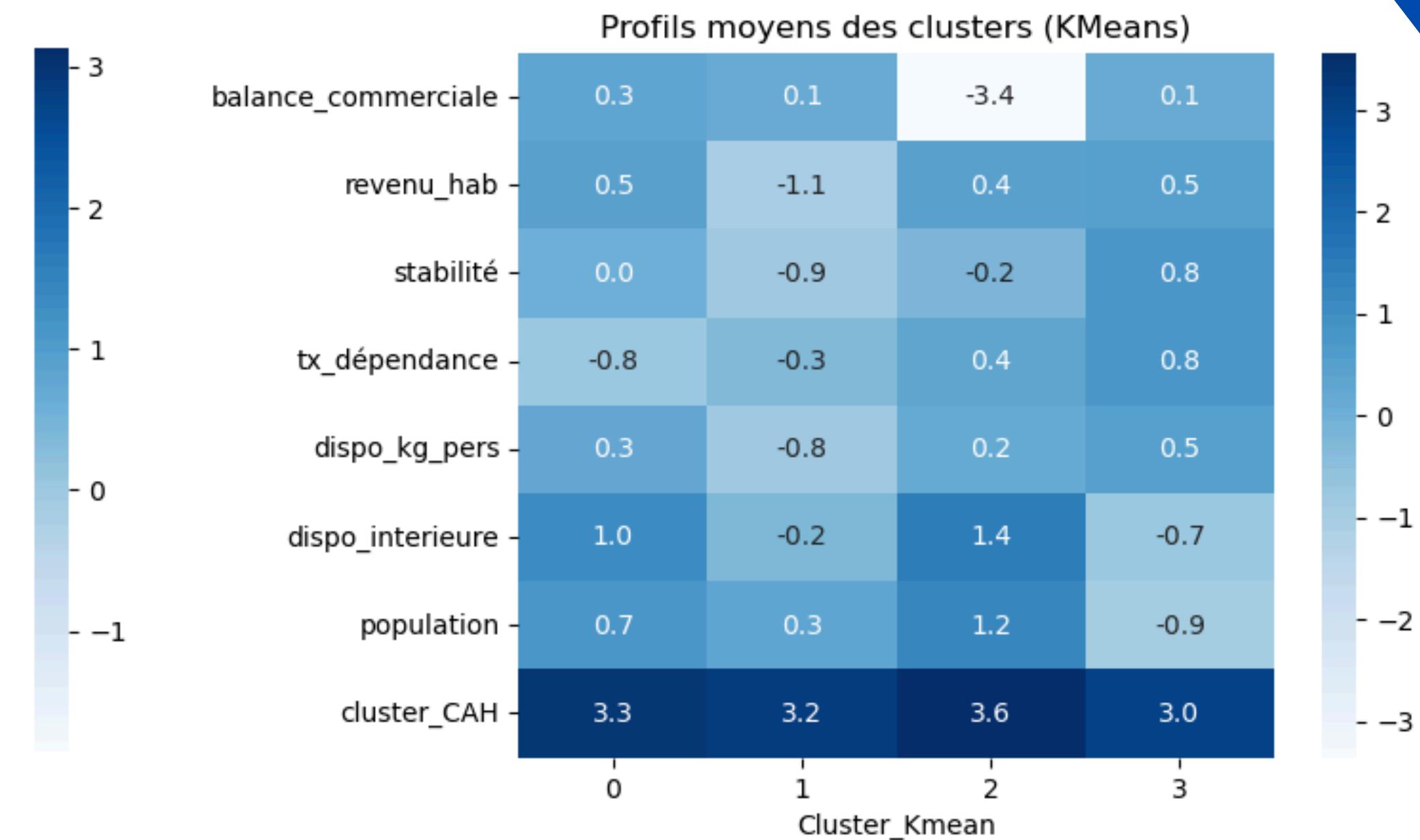
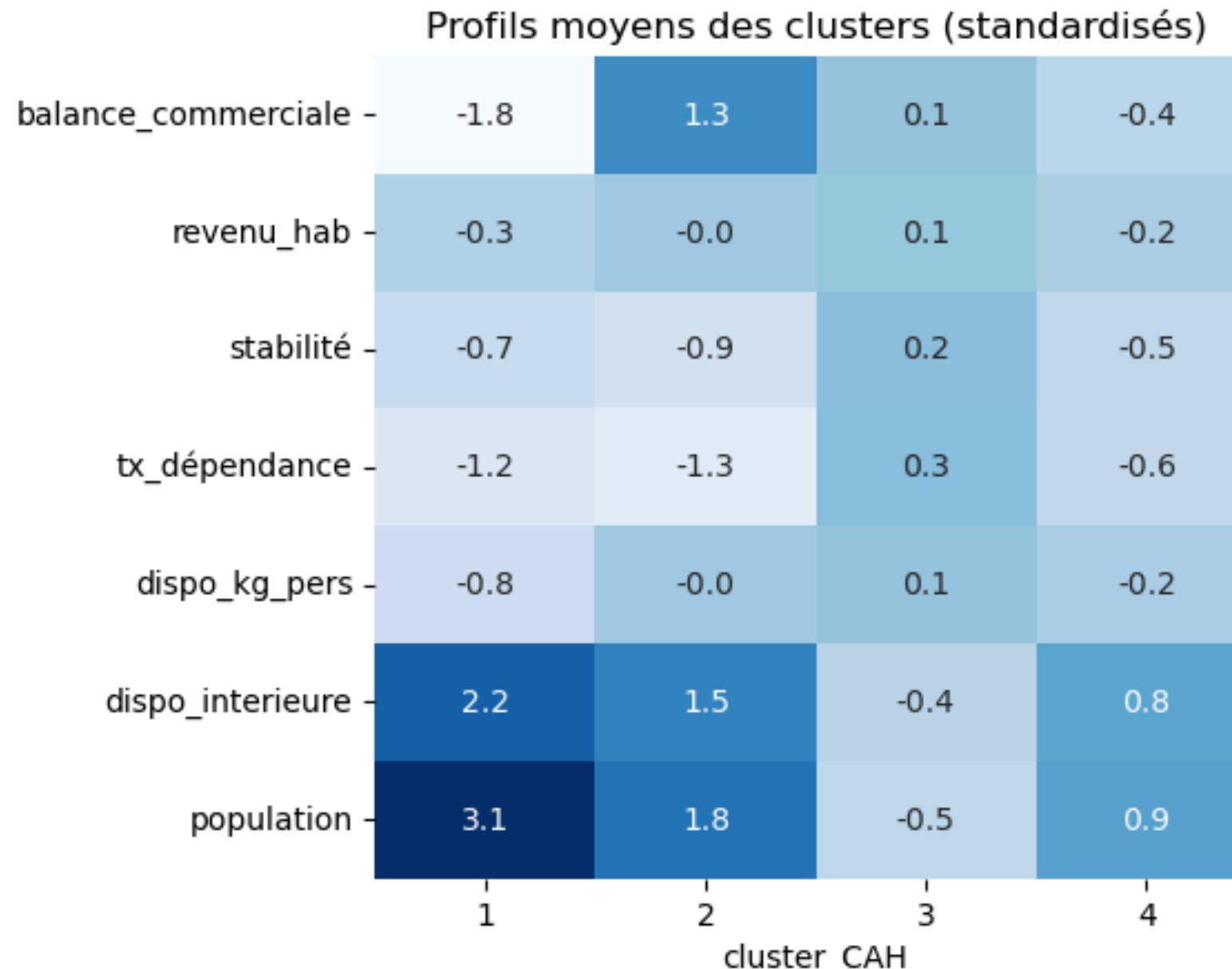
**Biplot ACP avec cluster CAH : Pays et variables clés**



**Biplot ACP avec cluster Kmean: Pays et variables clés**



# PROFIL MOYEN DES CLUSTERS



Pays	
cluster_CAH	
1	2
2	7
3	133
4	46
cluster_CAH	
1	[Chine, continentale, Inde]
2	[Bangladesh, Brésil, Fédération de Russie, Ind...]
3	[Albanie, Antigua-et-Barbuda, Arménie, Autrich...]
4	[Afghanistan, Afrique du Sud, Algérie, Allemag...

Pays	
Cluster_Kmean	
0	50
1	60
2	9
3	69
Cluster_Kmean	
0	[Algérie, Allemagne, Argentine, Australie, Bol...
1	[Afghanistan, Azerbaïdjan, Bangladesh, Bosnie-...
2	[Afrique du Sud, Angola, Arabie saoudite, Chin...
3	[Albanie, Antigua-et-Barbuda, Arménie, Autrich...

# TABLE DES SCORES D'ATTRACTIVITÉ POUR L'EXPORT

1

Définir les critères d'export clés

```
# Sélection des variables pertinentes
variables = ["tx_dépendance", "balance_commerciale", "revenu_hab", "stabilité","population"]
```

2

Evaluer chaque critère d'export par une note pondérée

```
df_merge["score"] = (
    0.5 * df_merge["tx_dépendance"] +
    0.5 * df_merge["revenu_hab"] +
    0.4 * df_merge["stabilité"] +
    0.1 * df_merge["population"] -
    0.2 * df_merge["balance_commerciale"]
)
```

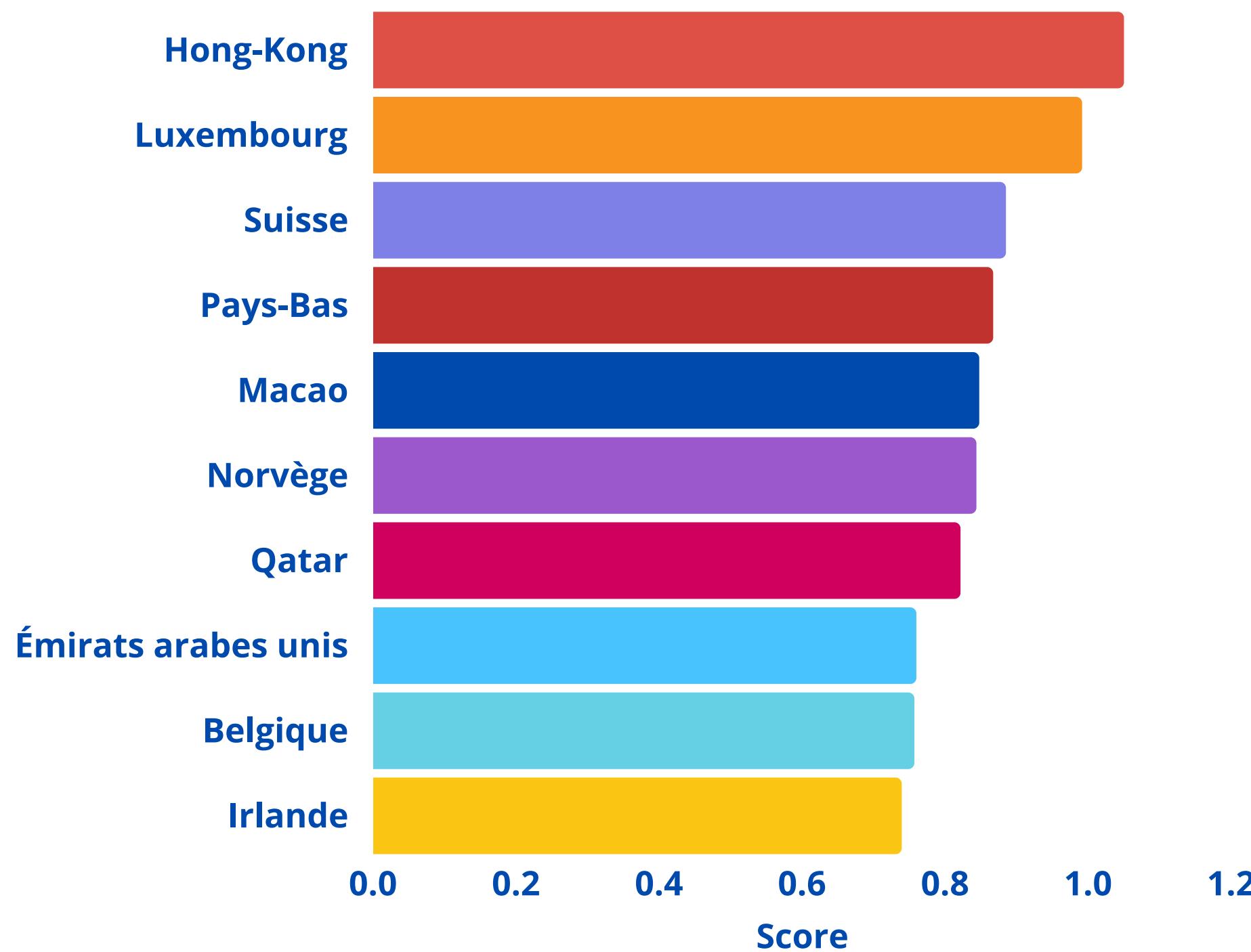
3

Table des scores d'attractivité à l'export

	cluster_CAH	Cluster_Kmean	score
Pays			
Afghanistan	3	3	0.096572
Afrique du Sud	4	2	0.296446
Albanie	3	1	0.397480
Algérie	3	0	0.160862
Allemagne	4	0	0.654739
...	...	...	...
Équateur	3	0	0.245639
États-Unis d'Amérique	1	0	0.535608
Éthiopie	3	3	0.086729
Îles Marshall	3	1	0.585249
Îles Salomon	3	1	0.464411

188 rows × 3 columns

# TOP 10 DES PAYS ATTRACTIFS POUR L'EXPORT



# RECOMMANDATIONS



# CONCLUSION

Analyse multivariée des critères d'export grâce à L'ACP afin d'identifier des corrélations entre certains critères et la variabilité des pays.

Création de clustering pour rassembler les pays qui ont des similitudes et les séparer de ceux qui sont différents. Puis identifier les clusters des pays attractifs.

Score final pour établir un top 10 des pays qui ont réunis au mieux tous les critères pertinents pour l'exportation des produits de l'entreprise.





# MERCI