

Étude de marché

Nassima Rahmani

Problématique : Quels sont les pays propices à l'exportation des poulets ?

- 1. Présentation des données :**
 - pays et variables.**
- 2. Créer des groupes de pays similaires :**
 - CAH, caractérisation des clusters.**
 - choix d'un groupe de pays.**
- 3. Analyse en composantes principales :**
 - 1.- les corrélations entre les variables .**
 - 2.- la variabilité entre les pays.**
- 4. Tests d'adéquations et de comparaisons**
- 5. Réduction de la liste des pays à cibler.**

Les individus : pays de tous les continents

Les variables :

- **Evolution de la population**
- **Proportion de protéines animales par rapport à la quantité totale de protéines**
- **Quantité protéines par habitant**
- **Quantité de calorie par habitant**

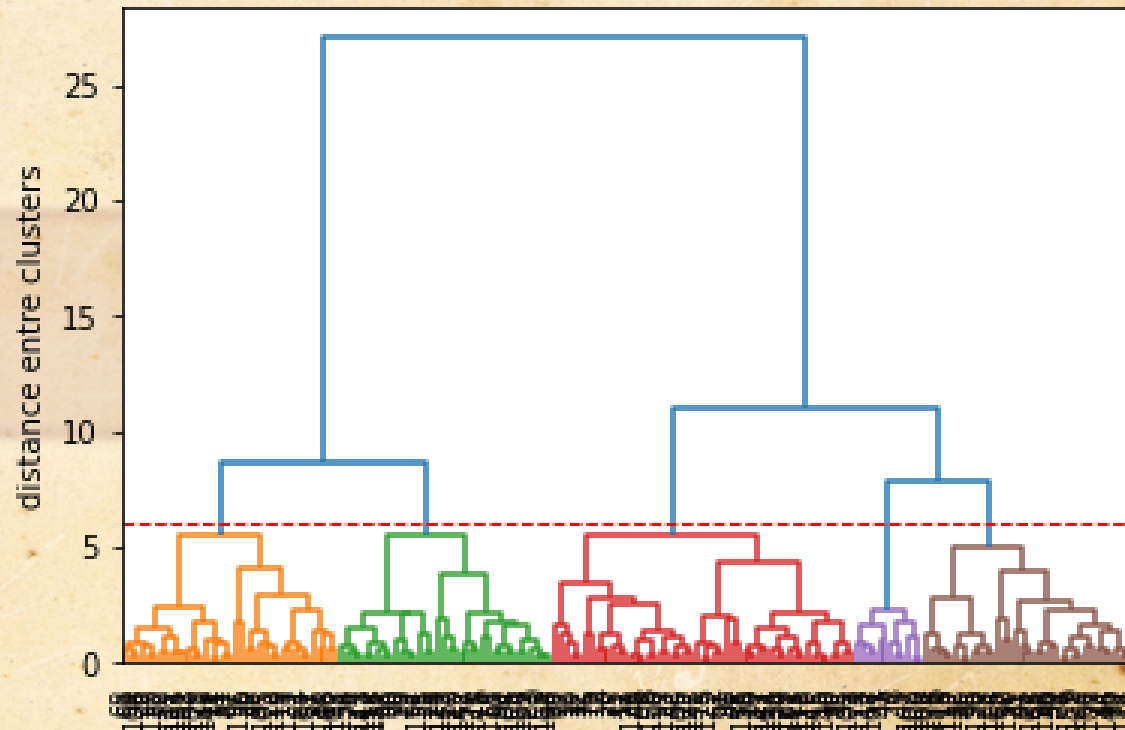
171 pays


| | Zone | diff_popu_pourc | Prop_prot | prot_hab | Calo_hab |
|-----|--|-----------------|-----------|----------|----------|
| 0 | Afghanistan | 8.015197 | 0.194344 | 0.05552 | 2040 |
| 1 | Afrique du Sud | 4.344302 | 0.427048 | 0.08437 | 2899 |
| 2 | Albanie | -0.268914 | 0.533523 | 0.11574 | 3360 |
| 3 | Algérie | 6.293751 | 0.269302 | 0.09183 | 3322 |
| 4 | Allemagne | 1.634734 | 0.607495 | 0.10540 | 3554 |
| ... | ... | ... | ... | ... | ... |
| 166 | Venezuela (République bolivarienne du) | -3.971537 | 0.453572 | 0.05417 | 2120 |
| 167 | Viet Nam | 3.095572 | 0.402313 | 0.09858 | 3025 |
| 168 | Yémen | 7.550768 | 0.190389 | 0.05452 | 2154 |
| 169 | Zambie | 9.272080 | 0.226334 | 0.06053 | 2002 |
| 170 | Zimbabwe | 4.518203 | 0.262514 | 0.04415 | 1908 |

171 rows × 5 columns

Créer des groupes de pays similaires :
- CAH, caractérisation des clusters.

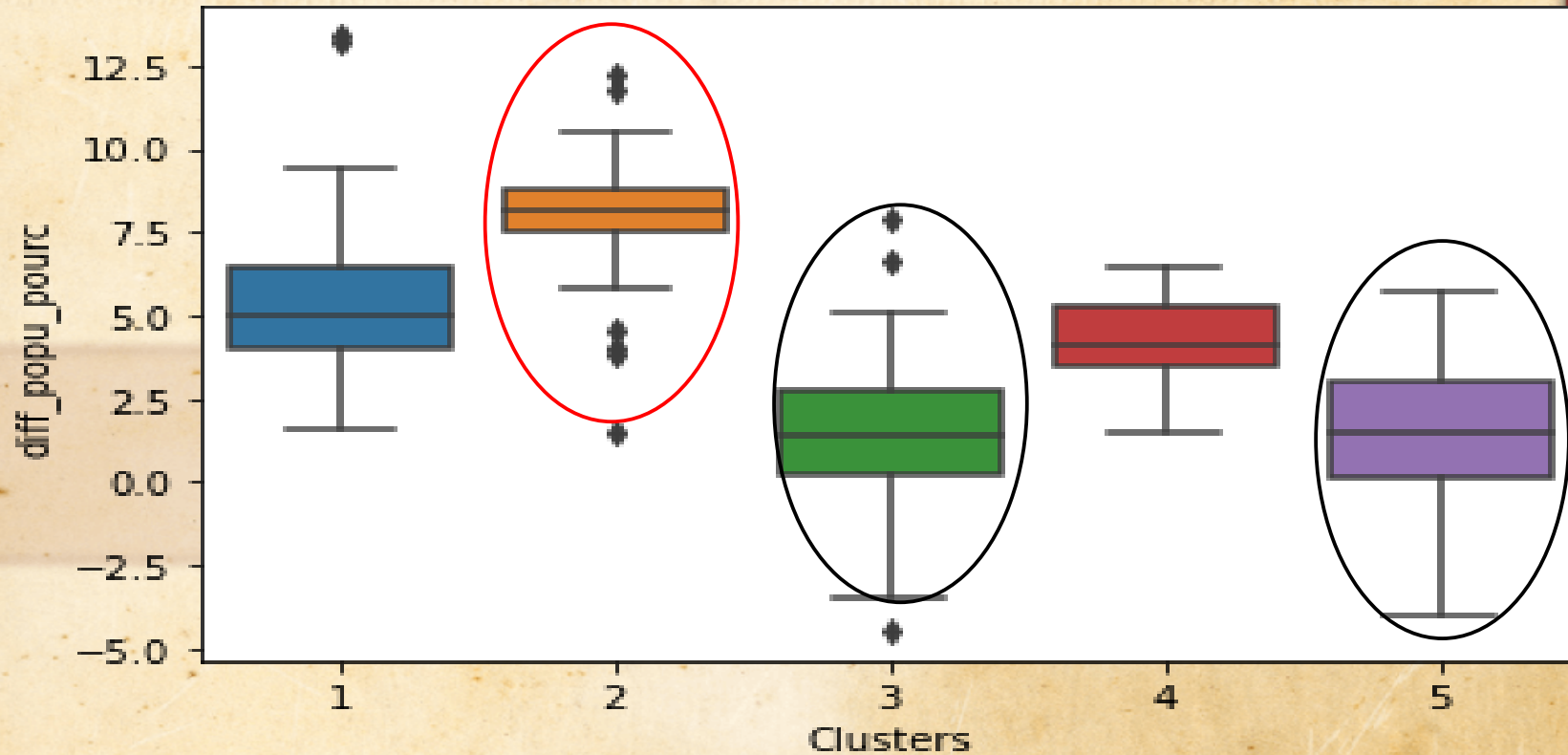
Un dendrogramme





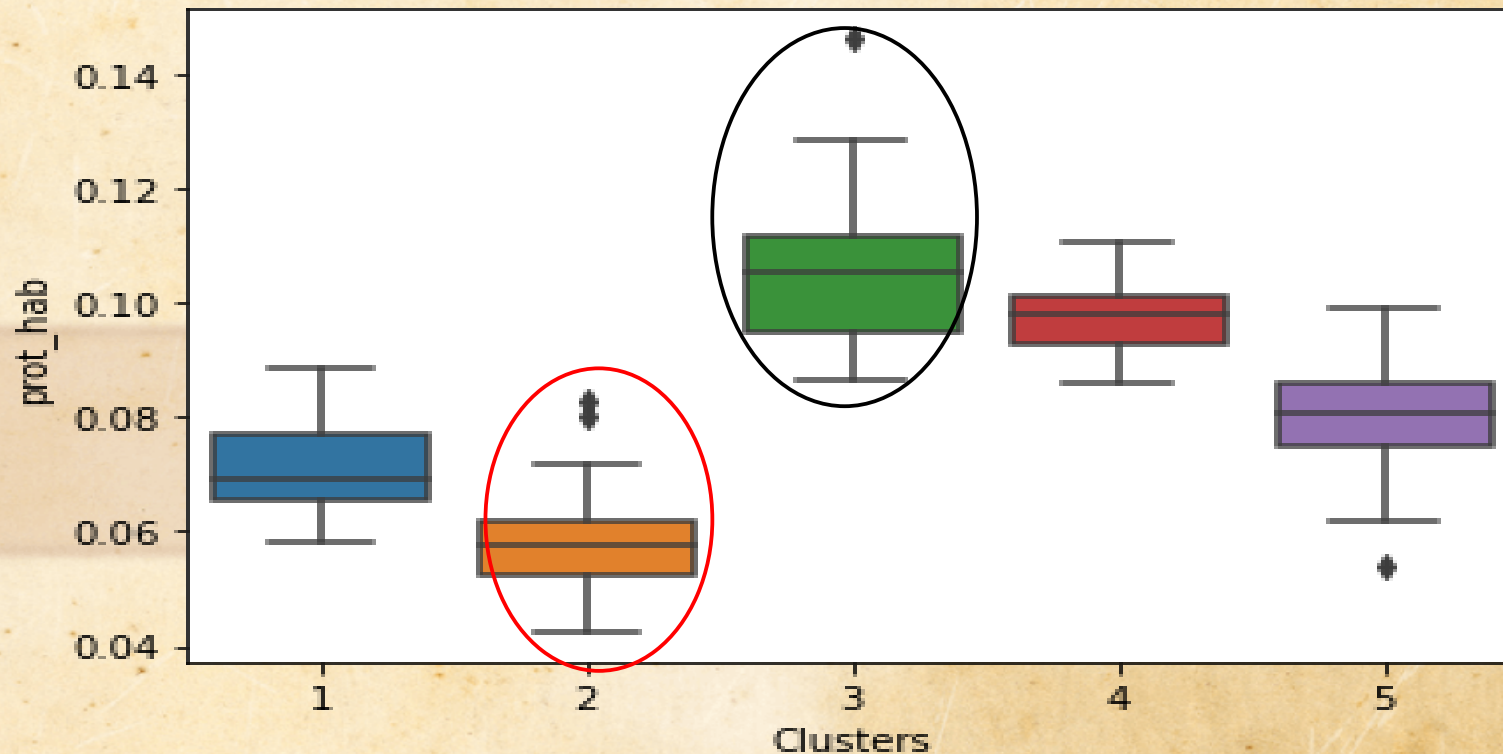
| Clusters | Nombre de pays |
|----------|----------------|
| 1 | 37 |
| 2 | 36 |
| 3 | 51 |
| 4 | 12 |
| 5 | 35 |

Evolution de la population pour les 5 clusters



Cluster 2: évolution de la population est la plus forte
Cluster 3 et 5: plus faible

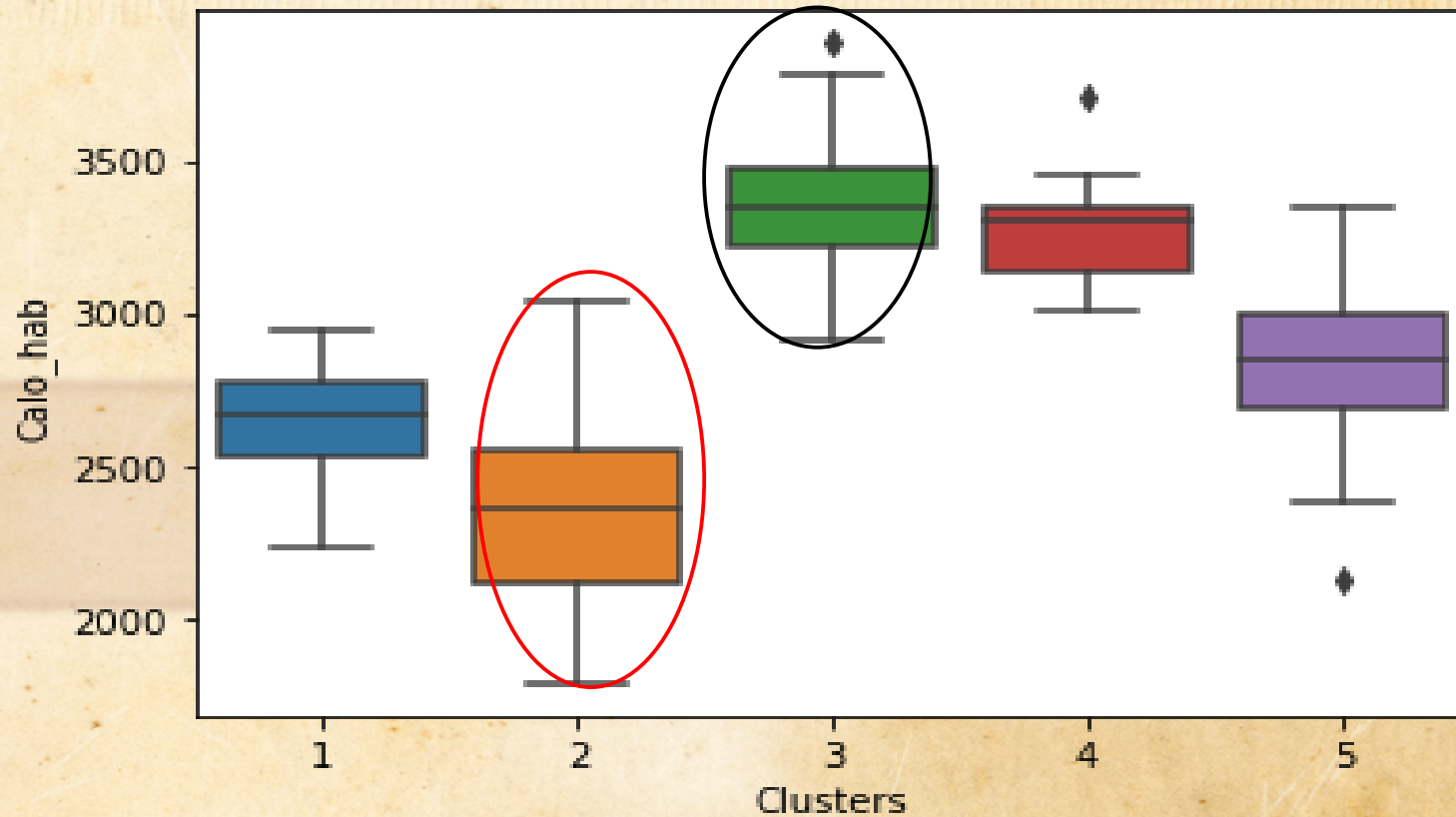
Protéines par habitant



Cluster 2 : est plus faible

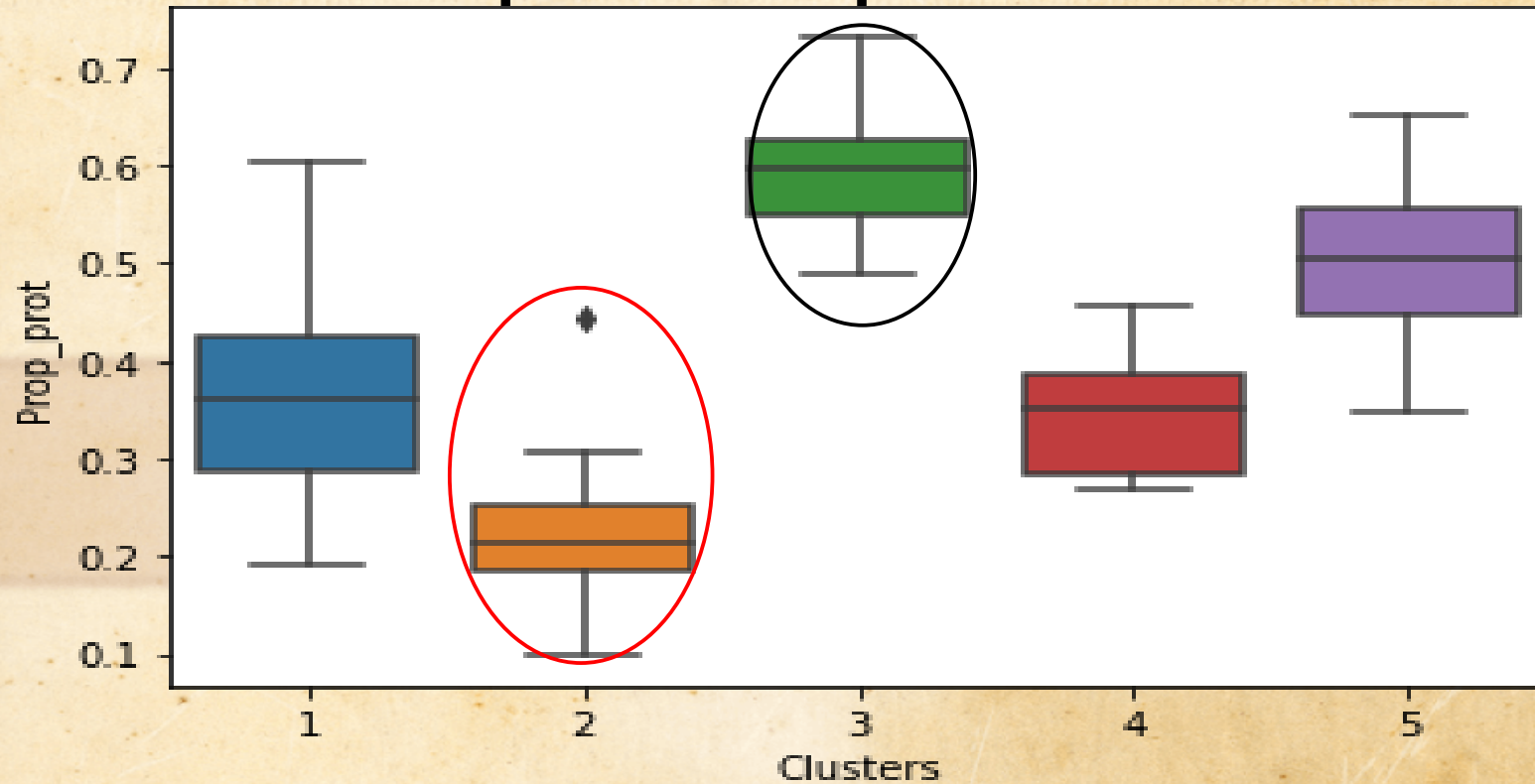
Cluster 3: la plus forte

Calories par habitant



Cluster 2 : est plus faible
Cluster 3 : est plus forte

Proportion de protéines animales



Cluster 2 : la plus faible
Cluster 3 : la plus forte

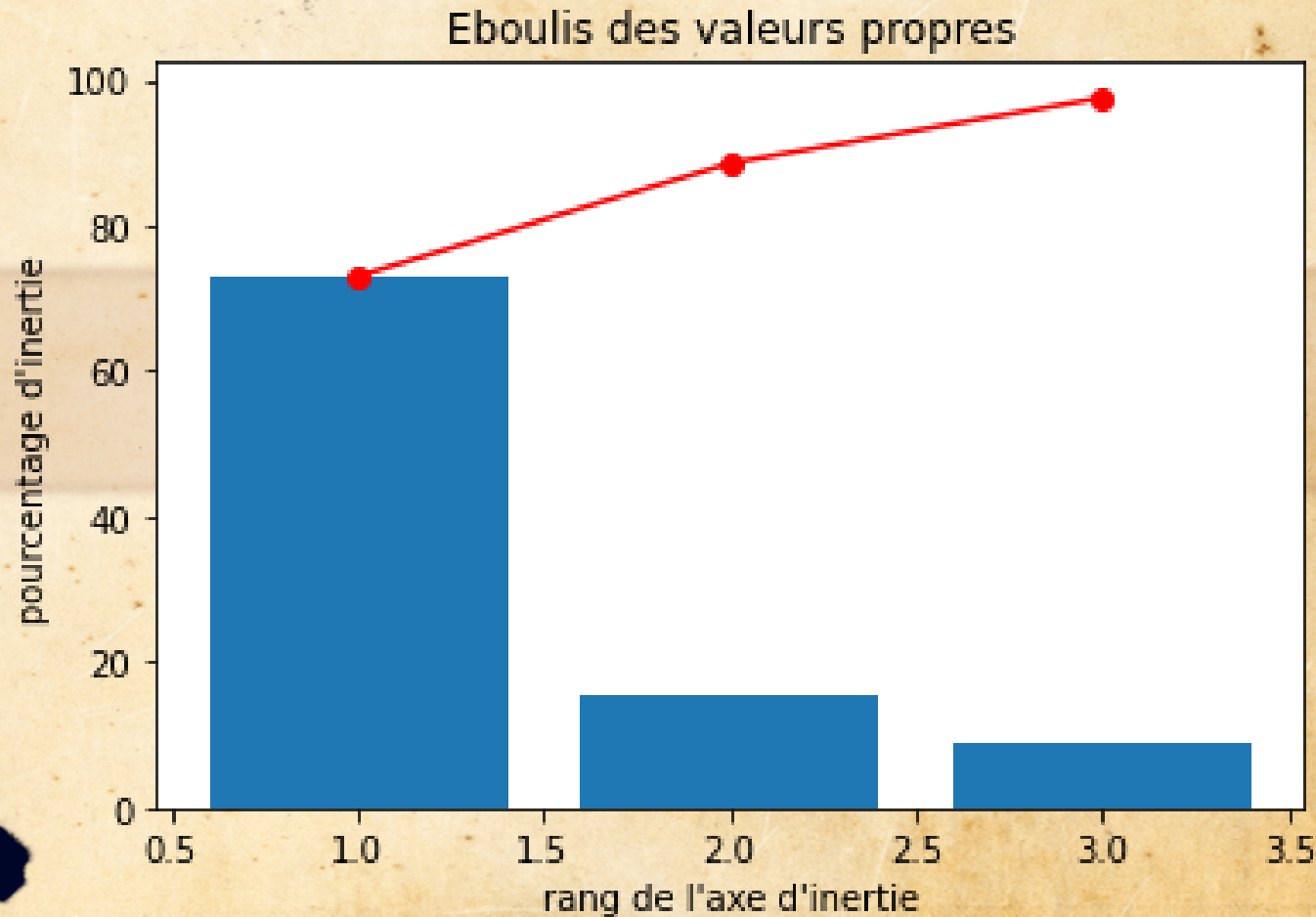
Moyennes des variables par clusters

| Variables Clusters | 1 | 2 | 3 | 4 | 5 |
|-------------------------------|--------------|--------------|--------------|--------------|--------------|
| diff_popu_pourc | 5,67 | 7,865 | 1,499 | 4,368 | 1,329 |
| Prop_prot | 0,359 | 0,217 | 0,591 | 0,342 | 0,504 |
| Prot_hab | 0,070 | 0,057 | 0,104 | 0,097 | 0,079 |
| Calo_hab | 2637 | 2337 | 3346 | 3273 | 2827 |

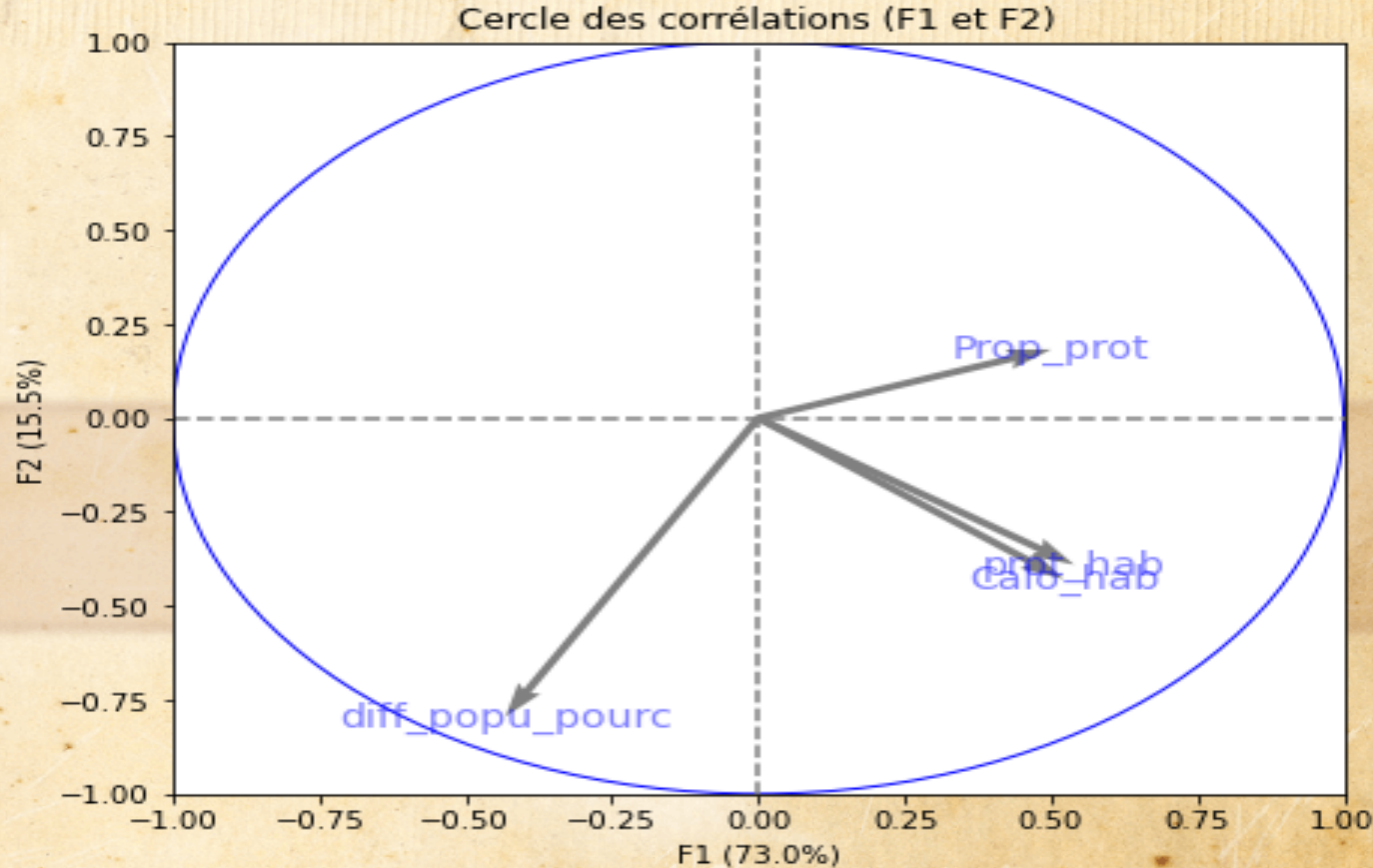
Choix du cluster : Cluster3 (51 pays)

| | | | | |
|-----------------|----------------------------|-------------------|----------------------|---------------------------------|
| Albanie | Allemagne | Argentine | Australie | Autriche |
| Barbade | Bélarus | Belgique | Brésil | Canada |
| Chili | Hong_Kong | Macao | Chypre | Croatie |
| Danemark | Espagne | Estonie | USA | Russie |
| Finlande | Grèce | Hongrie | Irlande | Islande |
| Israël | Italie | Kazakhstan | Koweït | Lettonie |
| Lituanie | Luxembourg | Malte | Mexique | Monténégro |
| Norvège | Nouvelle- Zélande | Pays-Bas | Pologne | Polynésie |
| Portugal | République de Corée | Roumanie | Royaume_ Unis | StVincent-les grenadines |
| Samoa | Slovénie | Suède | Suisse | Tchéquie |
| Uruguay | | | | |

Analyse en composantes principales



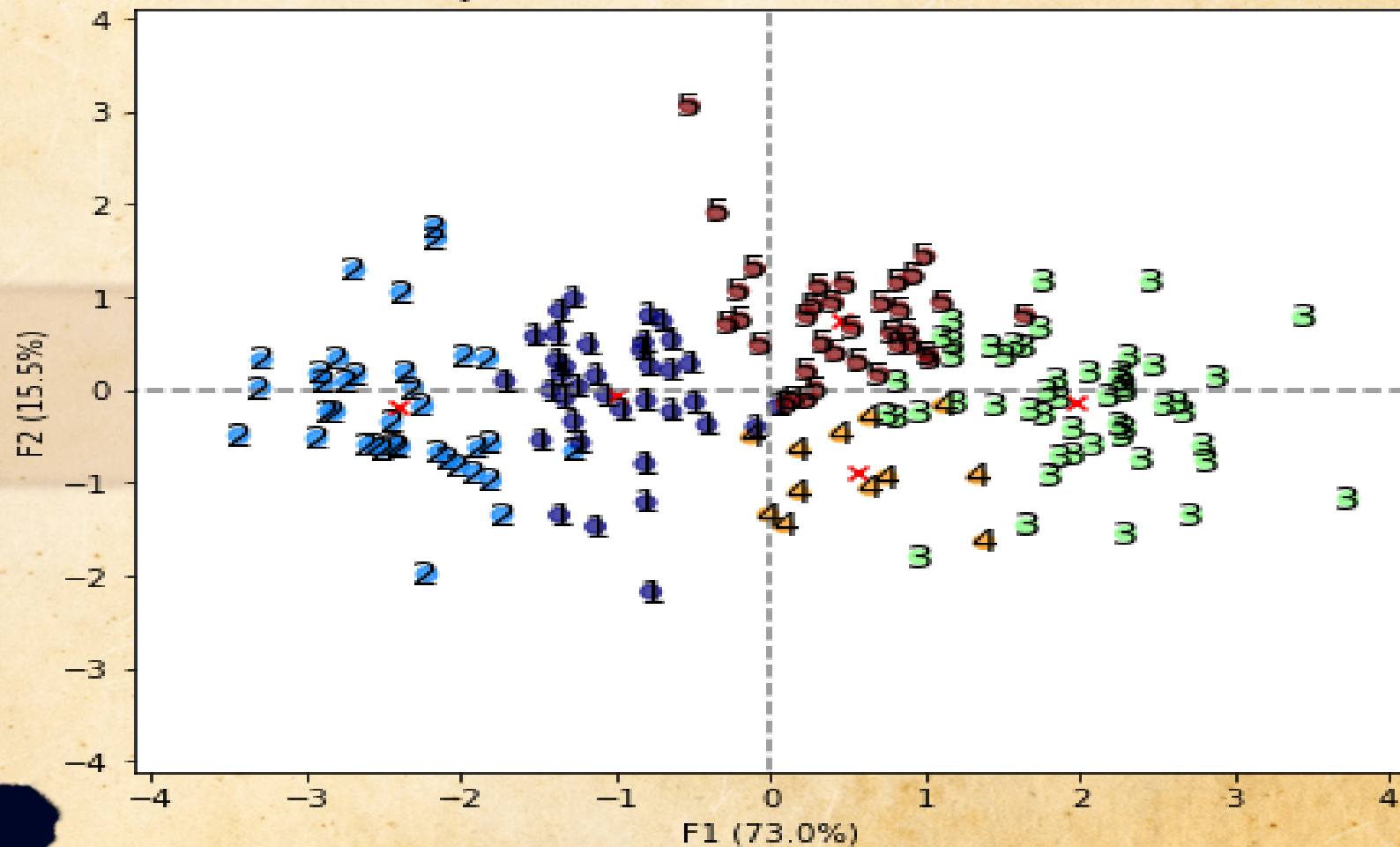
Analyse en composantes principales



Forte corrélation positive des 3 variables avec 1^{er} axe (73%)

Corrélation négative de la variable évolution population avec 1^{er} axe

Projection des individus (sur F1 et F2)



Test d'adéquation 4 variables

Test de shapiro

| Variables | P_value | Niveau de confiance : 0,05 |
|-----------------|-----------|----------------------------|
| diff_popu_pourc | 0,118 | > 0,05 |
| Prop_prot | 3,054E-05 | <0,05 |
| prot_hab | 0,03 | <0,05 |
| Cal_hab | 0,102 | >0,05 |

Deux variables gaussiennes : diff_popu_pourc
Cal_hab

Comparer deux échantillons

Cluster 3 et cluster 4 (évolution de la population)

Comparaison des deux variances :

Test de Levene : $P_value = 0,2219 > 0,05$, les deux variances sont égales

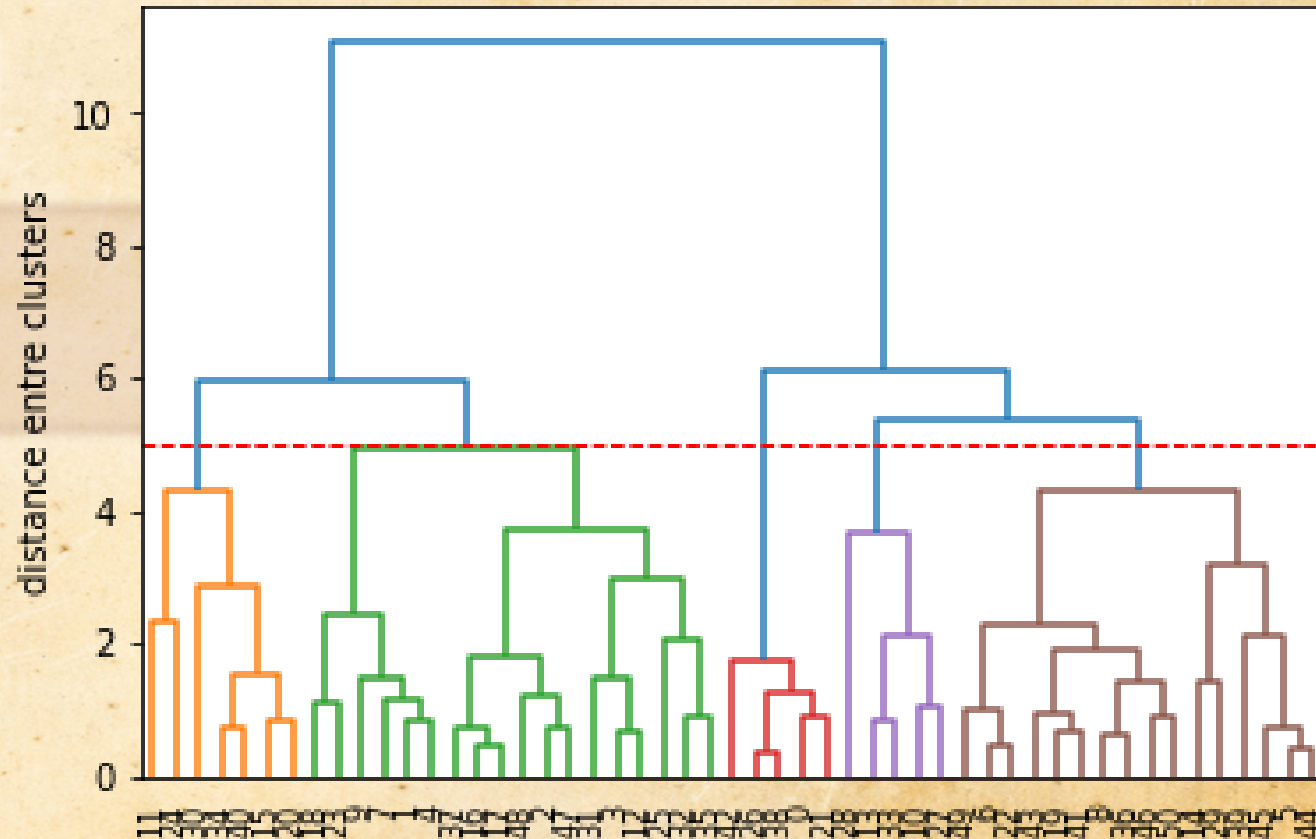
Comparaison des moyennes :

Test de Student : $P_value = 0,0001 < 0,05$, les deux moyennes sont différentes.

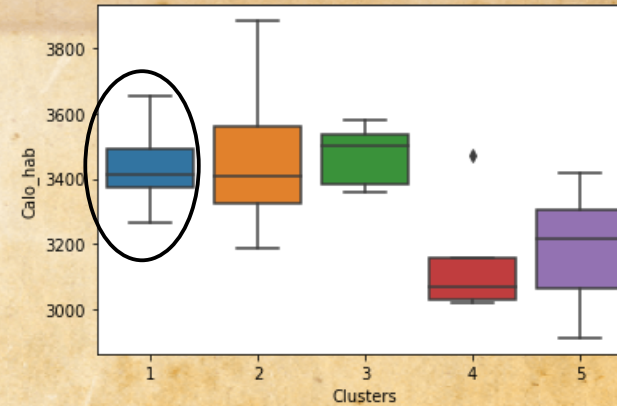
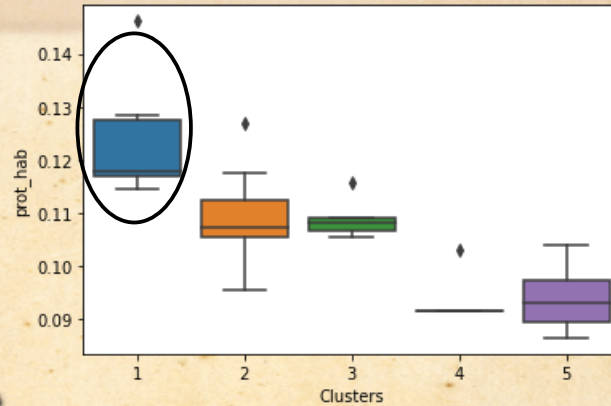
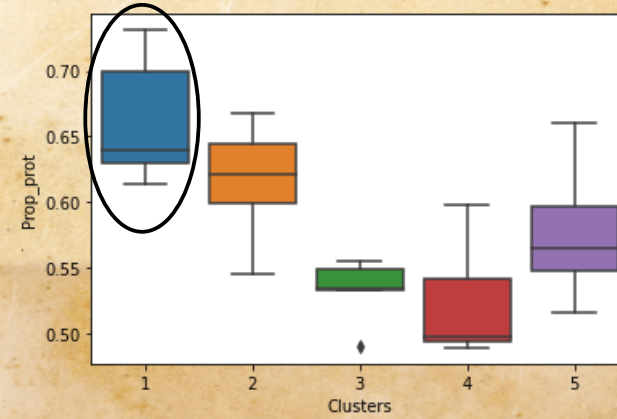
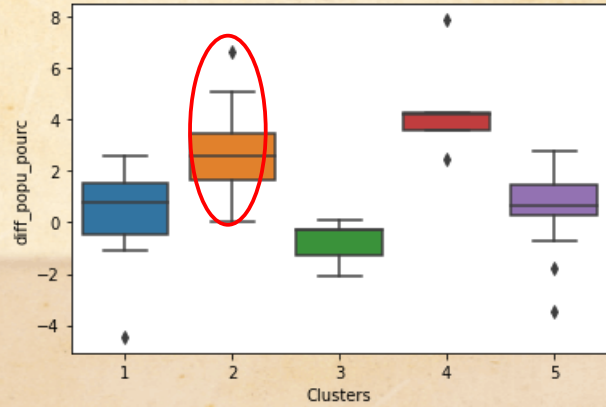
Les clusters 3 et 4 sont différents

Réduction de la liste des pays à cibler

CAH sur le cluster 3



Étude des variables et les clusters



Cluster 1 : 7 pays

Chine-RAS de Hong-Kong

Danemark

Finlande

Islande

Lituanie

Monténégro

Portugal

Conclusion

- Base de données :

171 pays

choix des variables :

Evolution de la population et disponibilité en
protéines animales

- 1^{ère} classification hiérarchique :

partition par groupe d'individus similaires

Choix du cluster : 51 pays (majorité européens)

- Analyse en composantes principales :

⇒ les 3 variables : `prot_hab`, `Prop_prot` et `Calo_hab` sont
bien corrélées à l'axe 1 (73% d'inertie)
et la variable `diff_popu_pourc` est négativement
corrélée à l'axe 1

⇒ Pour les individus : la répartition des clusters sur l'axe 1
suit celle des variables

- 2^{ème} classification hiérarchique : 7 pays

- 2^{ème} classification hiérarchique 7 Pays :

Chine-RAS de Hong-Kong

Danemark

Finlande

Islande

Lituanie

Monténégro

Portugal