

PROJET 5

# ÉTUDE DE MARCHÉ

Identification des pays propices à une  
insertion dans le marché de poulet

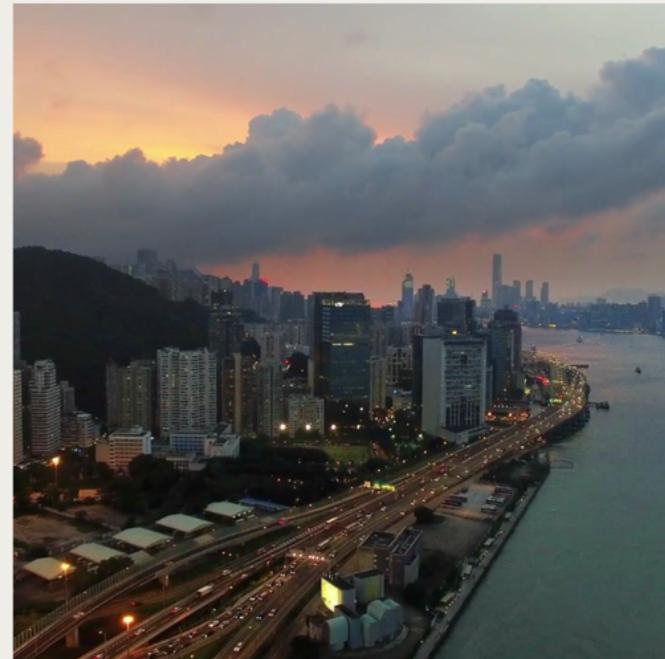
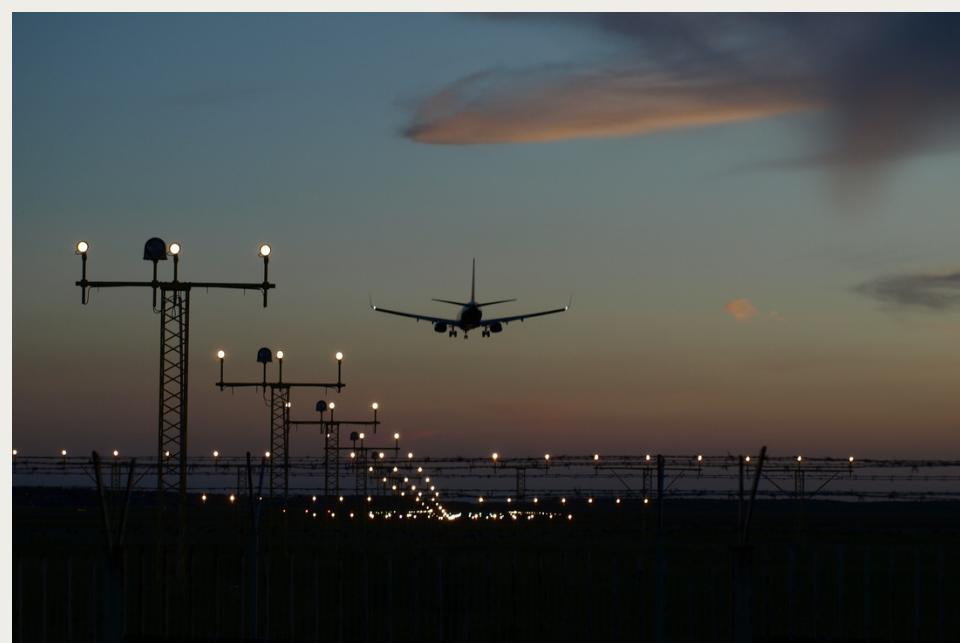
Data Analyst, *Katrina JUMADIAO* | OpenClassrooms



# SOMMAIRE

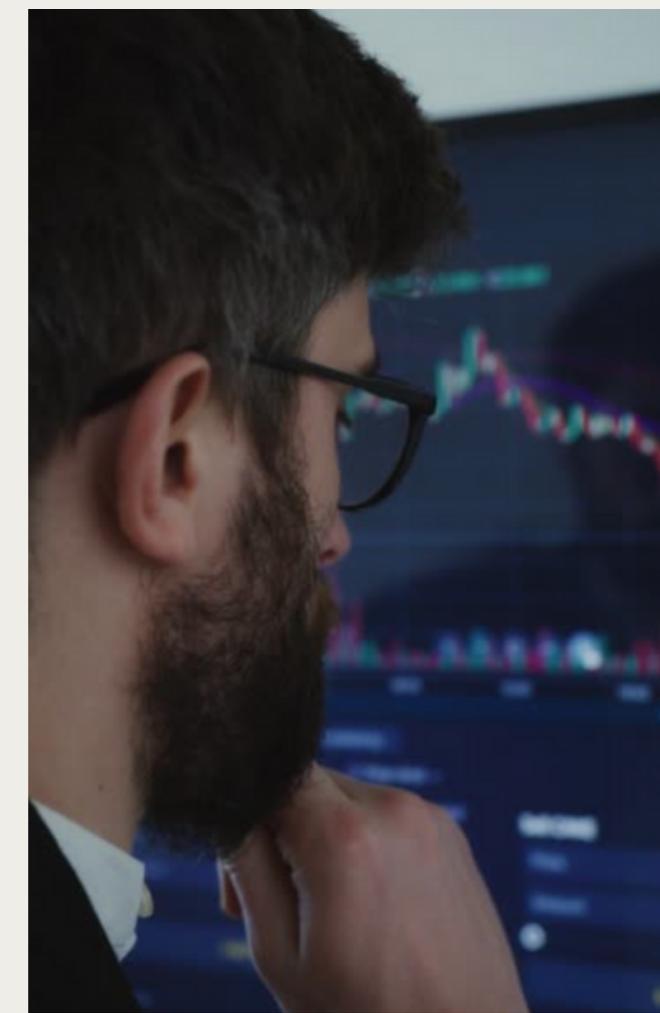
## Axes principaux du projet

### PHASE 3 Conclusion



### PHASE 1

Aperçu globale du projet



### PHASE 2 Analyse des données



# PHASE 1

Aperçu du projet

## Partie A

Autour de l'entreprise

## Partie B

Données utilisées

## Partie C

Méthodologie



# AUTOUR DE L'ENTREPRISE

## Enjeux

Entreprise agroalimentaire bien implanté en France avec envie d'étendre son marché à l'international

## Objectif

Se diriger davantage dans l'exportation et pas d'implantation, ni production locale



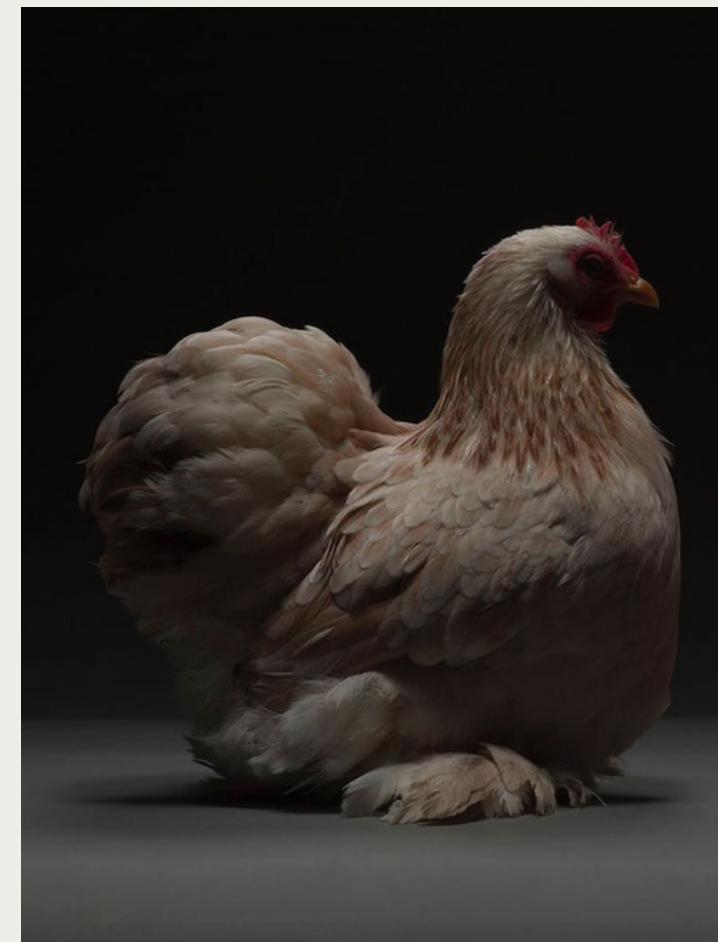
## Problématique

Quels pays choisir pour se développer à l'étranger?

# DONNÉES UTILISÉES POUR L'ÉTUDE DE MARCHÉ



POPULATION



ANIMAUX



VÉGÉTAUX



PIB

*source FAO*

# DÉMARCHES ET MÉTHODOLOGIE



## NETTOYAGE

Vérification des données aberrantes ou manquantes



## SELECTION

Tri des variables nécessaires pour la construction de l'échantillon



## FINALISATION

Création des nouvelles variables pour la data set finale



# Jeu de donnée finale

	pays	croissance_pop_(en %)	prot_animal_(en %)	dispo_prot_hab_(en g)	dispo_kcal_hab_(en kcal)
0	Arménie	0.24	49.04	35525.45	1119820.0
1	Afghanistan	2.41	18.54	21027.65	827455.0
2	Albanie	-0.05	53.95	41175.65	1202675.0
3	Algérie	2.03	26.91	32959.50	1233335.0
4	Angola	3.33	30.97	19647.95	892425.0

L'échantillon final utilisé pour notre étude de marché

# PHASE 2

Analyse des données

**Partie A**

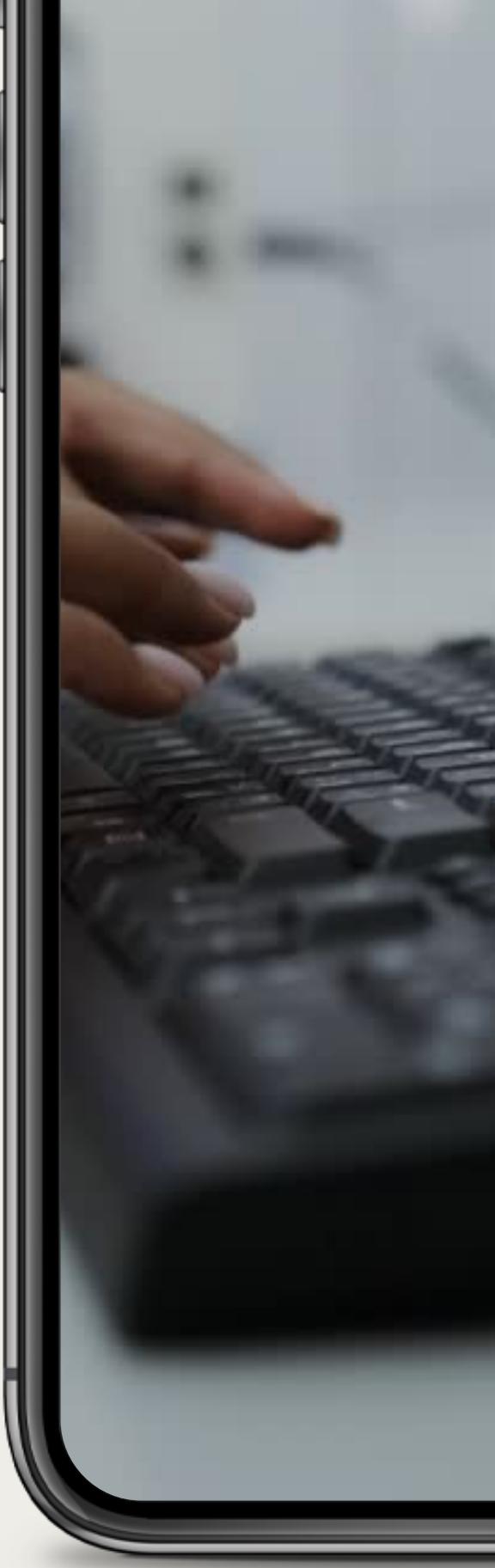
Clustering

**Partie B**

Analyse par  
Composantes  
Principales

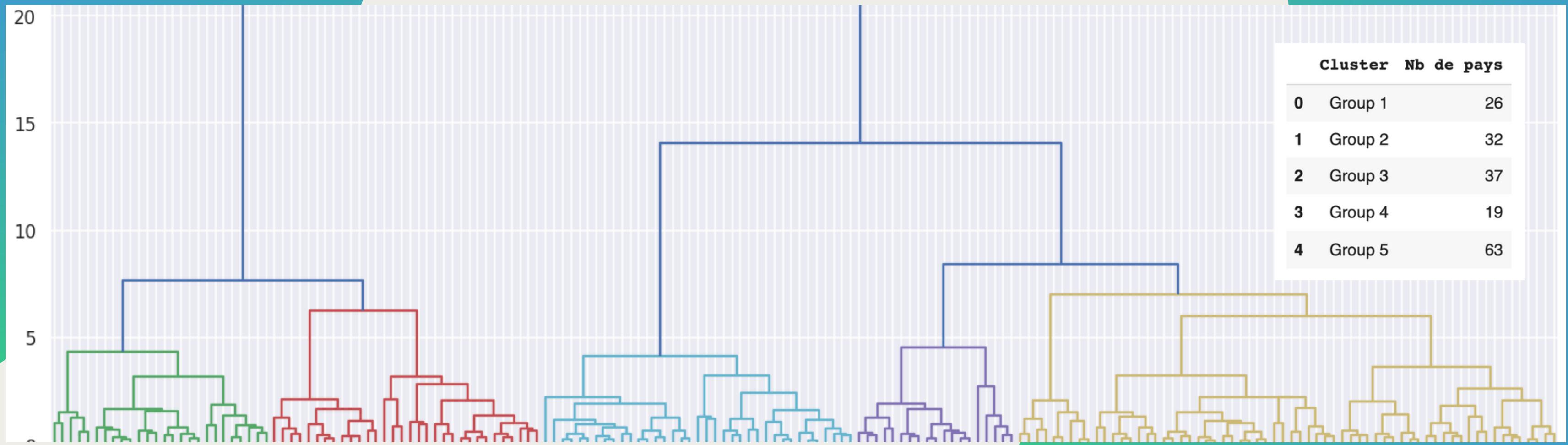
**Partie C**

Test statistiques



# CLUSTERING

*La classification hiérarchique : Dendrogramme*



## But

Diviser les pays en fonction de leurs caractéristiques afin de former 5 groupes de pays distincts

## MÉTHODE "Ward"

Maximise l'inertie entre les groupes

## Quel groupe choisir?

Utilisation de l'Analyse en Composante Principale et des statistiques descriptives

# Caractéristiques descriptives des clusters

CLUSTER	Croissance Population	% de protéines animales	Disponibilité alimentaire en protéines	Disponibilité alimentaire en calories
Group 1	2.39	22.99	18,767.74	797,384.62
Group 2	2.53	32.51	25,554.45	951,760.31
Group 3	0.50	60.42	40,264.43	1,255,442.16
Group 4	0.70	31.60	25,141.39	991,551.32
Group 5	0.75	49.32	31,740.28	1,099,965.16



# ACP

## LES SUJETS COUVERT

- 01 Éboulis des valeurs propres
- 02 Cercle de corrélation
- 03 Projection des individus
- 04 Projections des centroïdes
- 05 Test d'adéquation
- 06 Test de comparaison de 2 populations

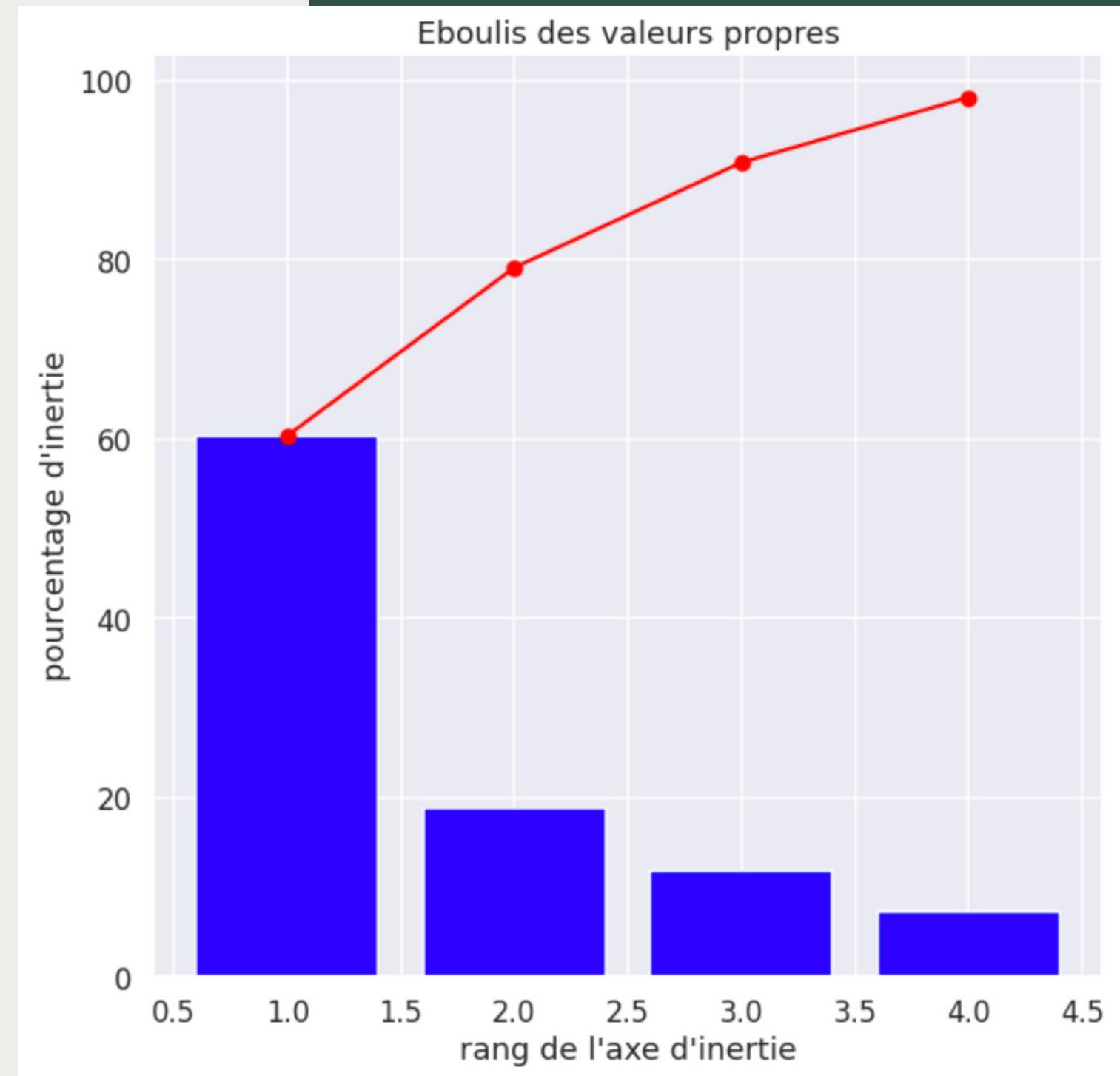
# Ébouli des valeurs propres

## 5 variables quantitative

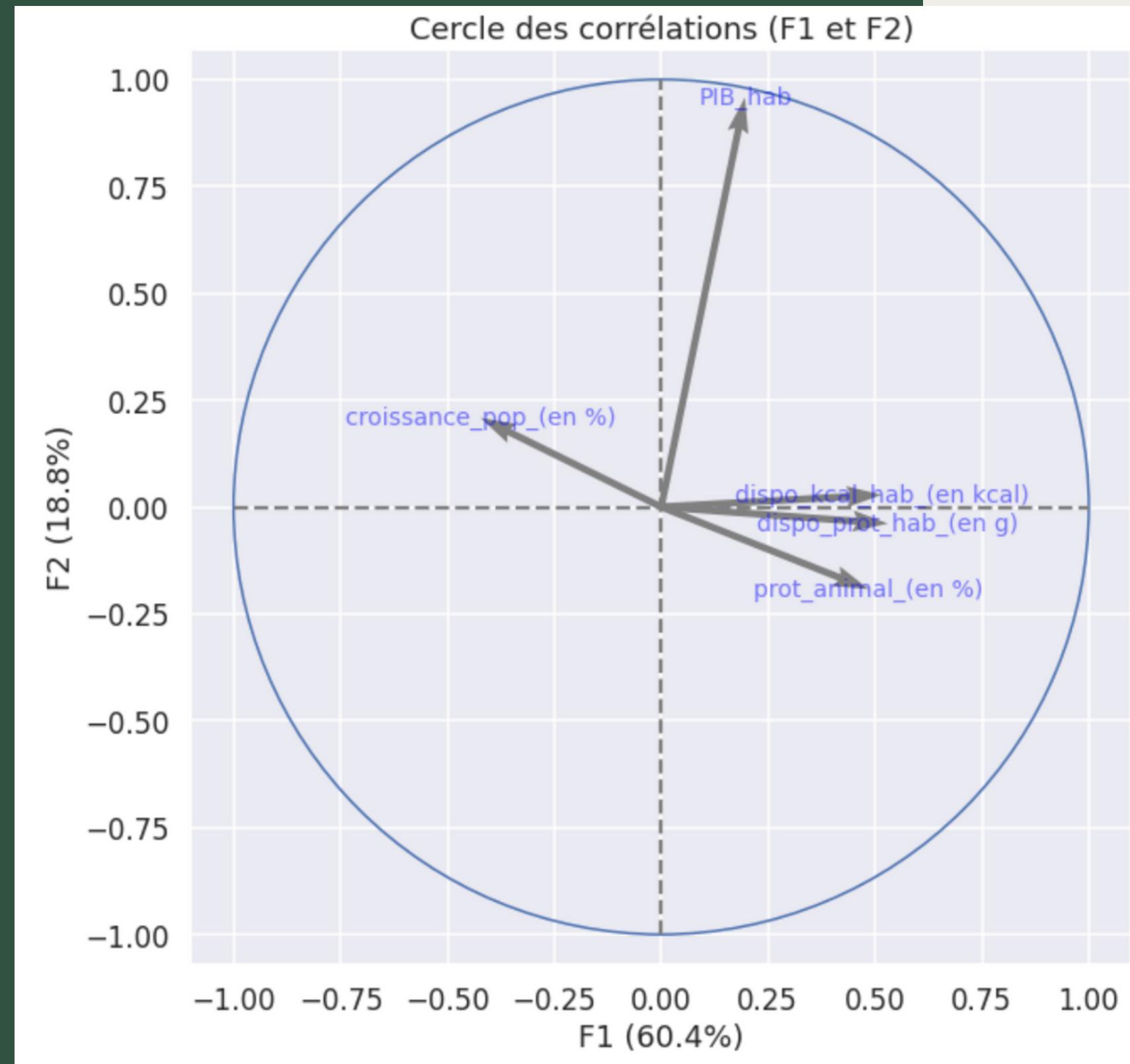
- croissance de population
- proportion de protéines animales
- disponibilité alimentaire en protéines
- disponibilité alimentaire en calories

## Méthode du coude

- 2 composantes conservées



# Cercle de corrélation

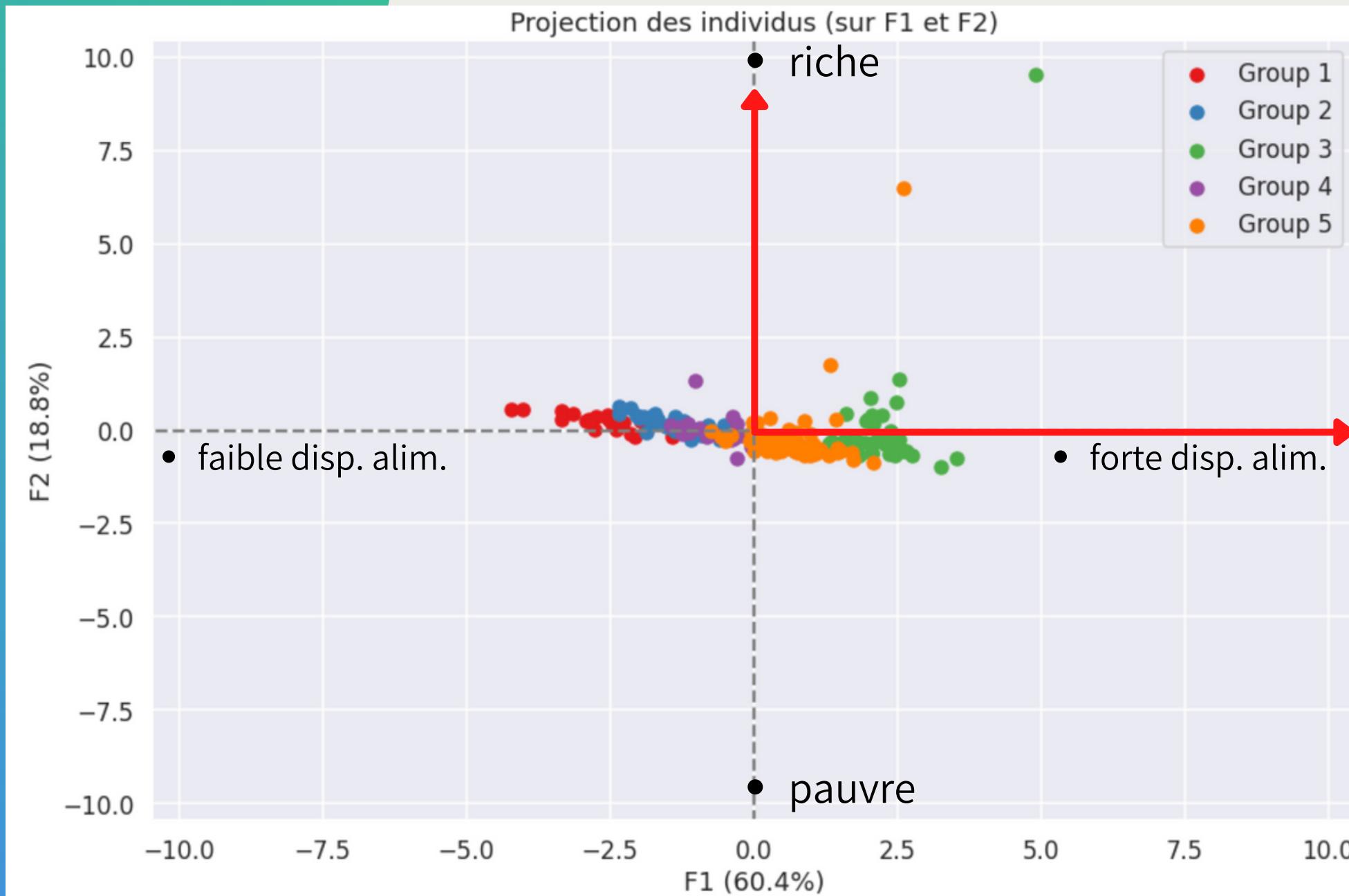


COMPOSANTE	EVR*
F1	Alimentation
F2	Économique

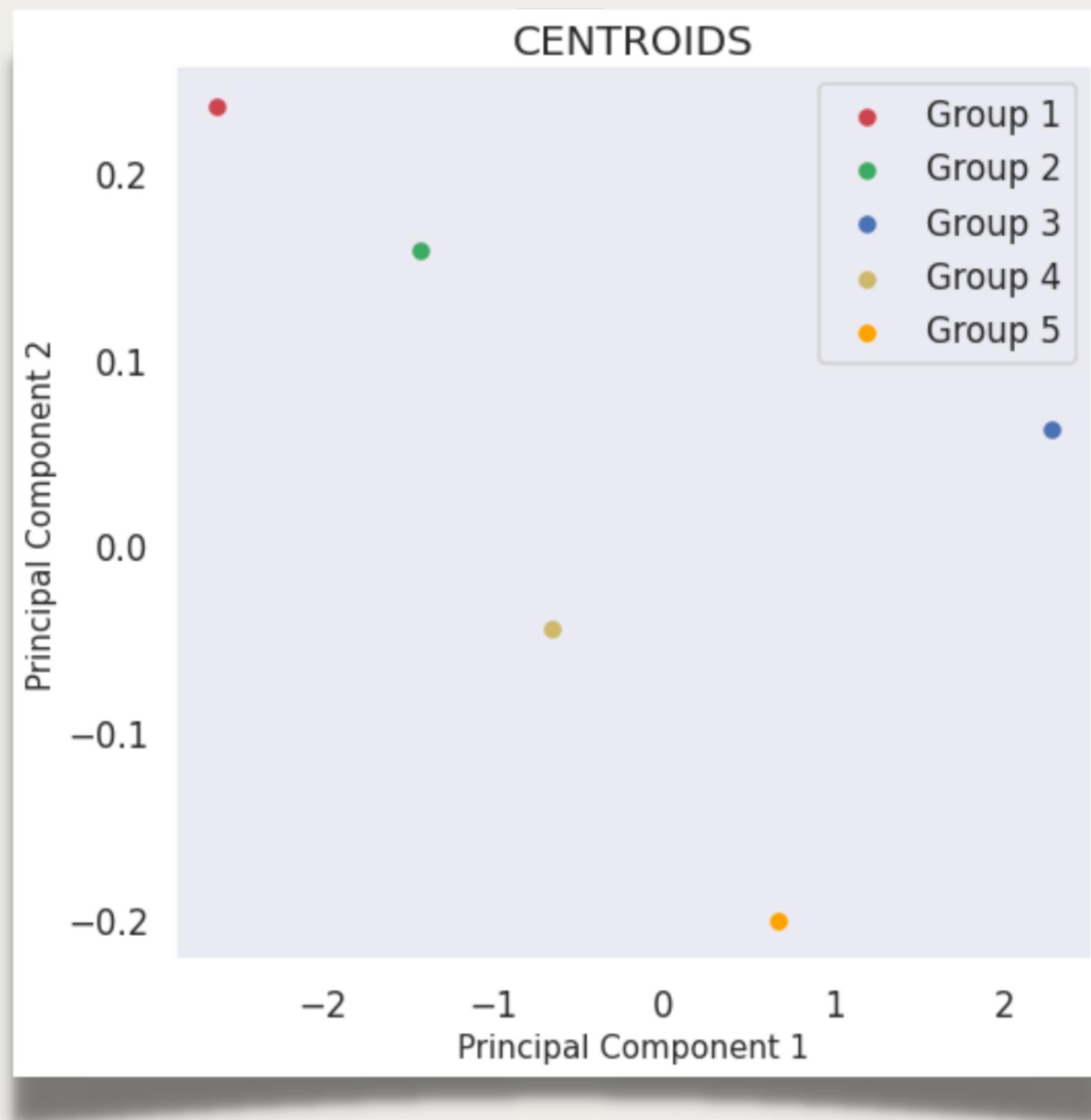
\*Explained variance ratio

# Projection des individus

Projection des 177 pays (groupé) sur nos 2 composantes

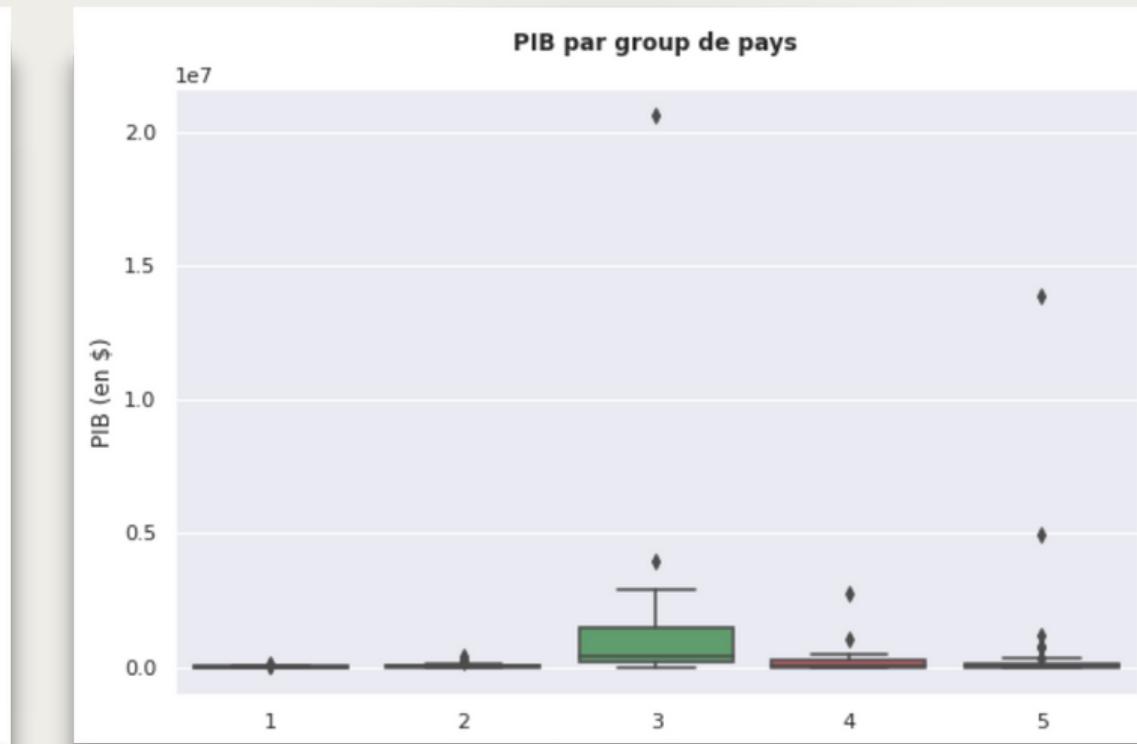
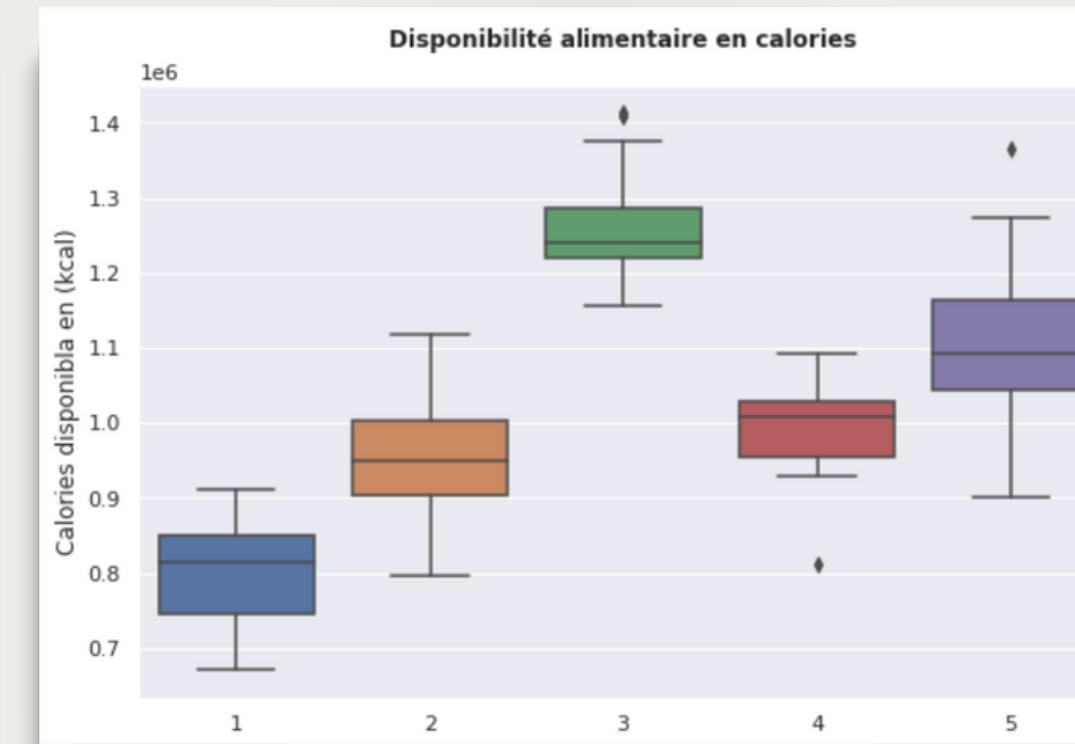
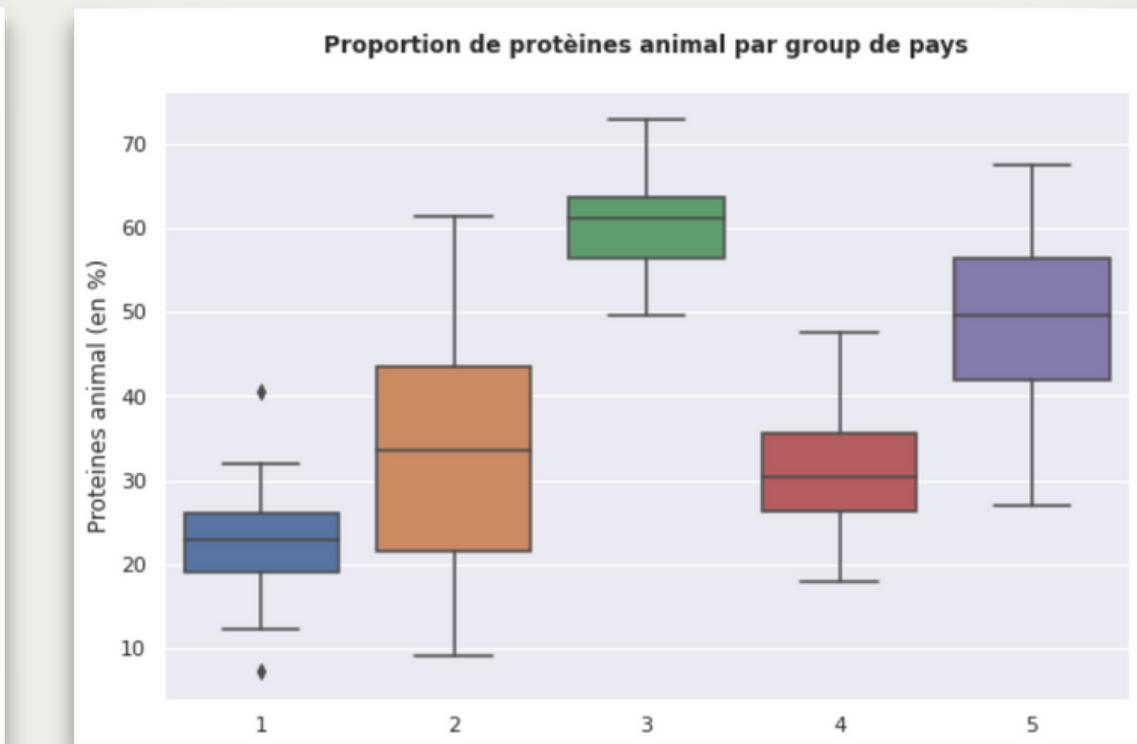
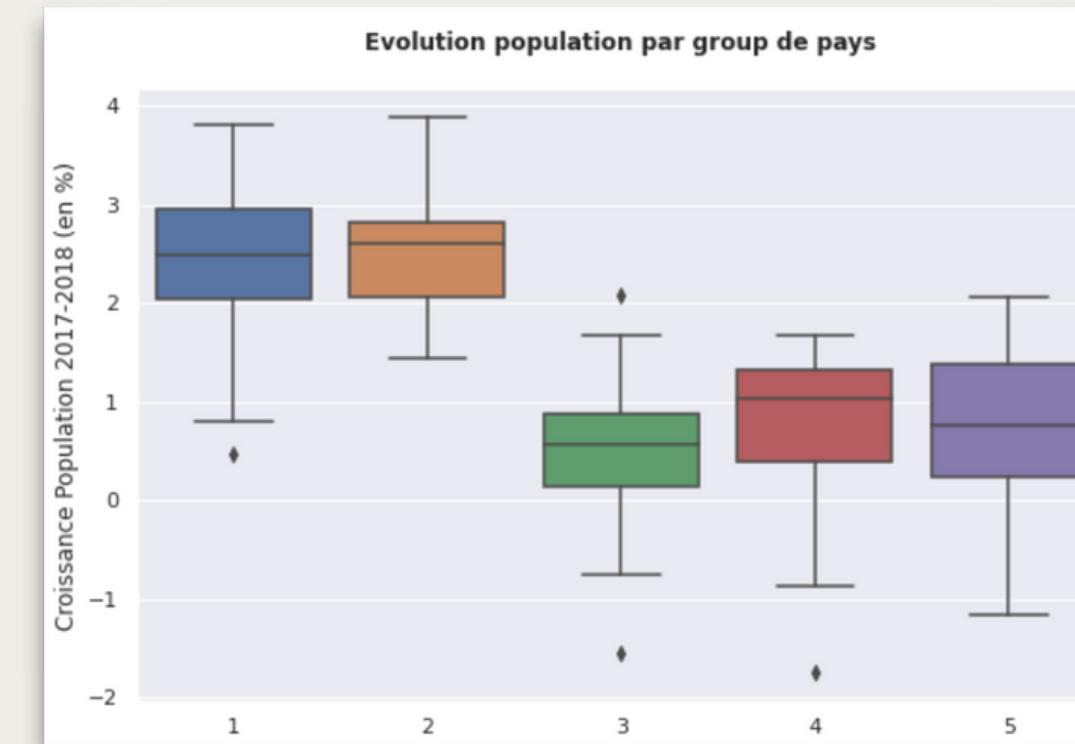
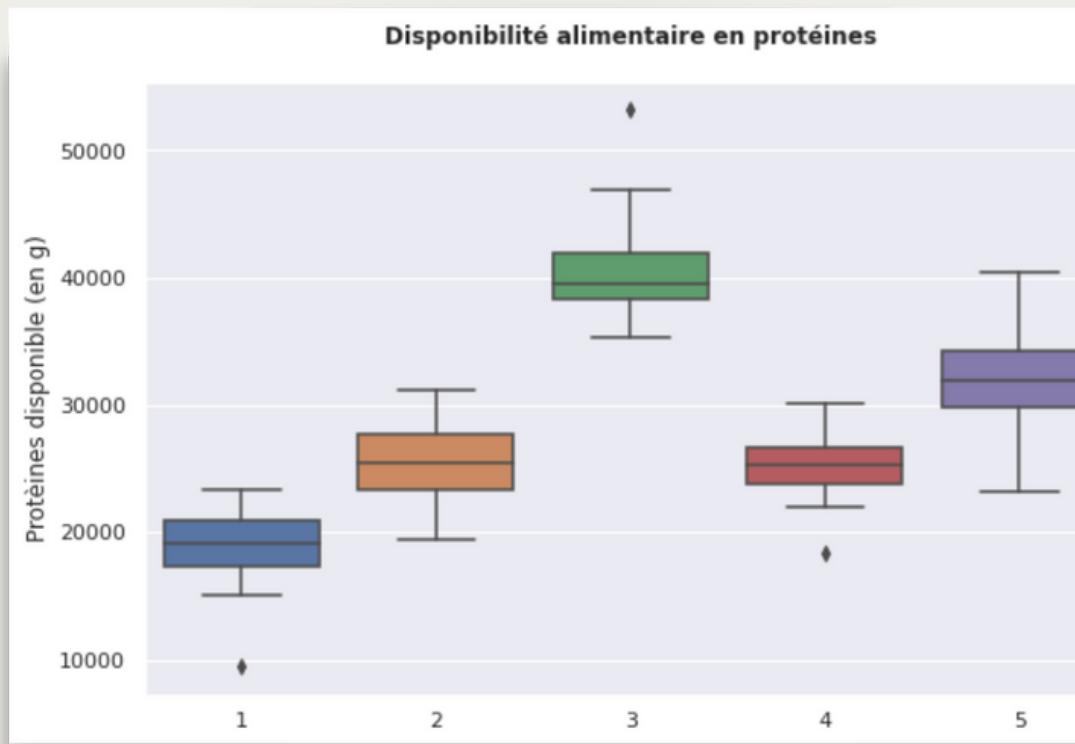


# POSITION DES CENTROÏDES



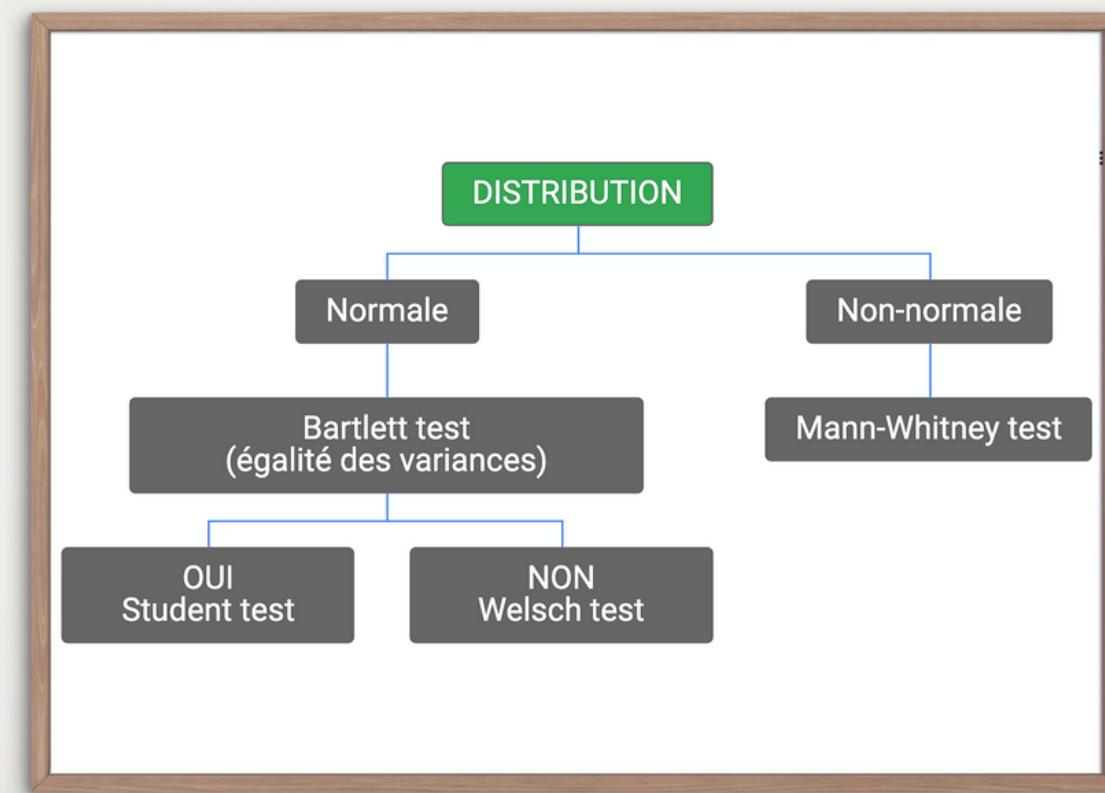
cluster	principal component 1	principal component 2
Group 1	-2.623488	0.235451
Group 2	-1.432403	0.158046
Group 3	2.273539	0.062262
Group 4	-0.652776	-0.044756
Group 5	0.671895	-0.200517

# MOYENNES DES CLUSTERS

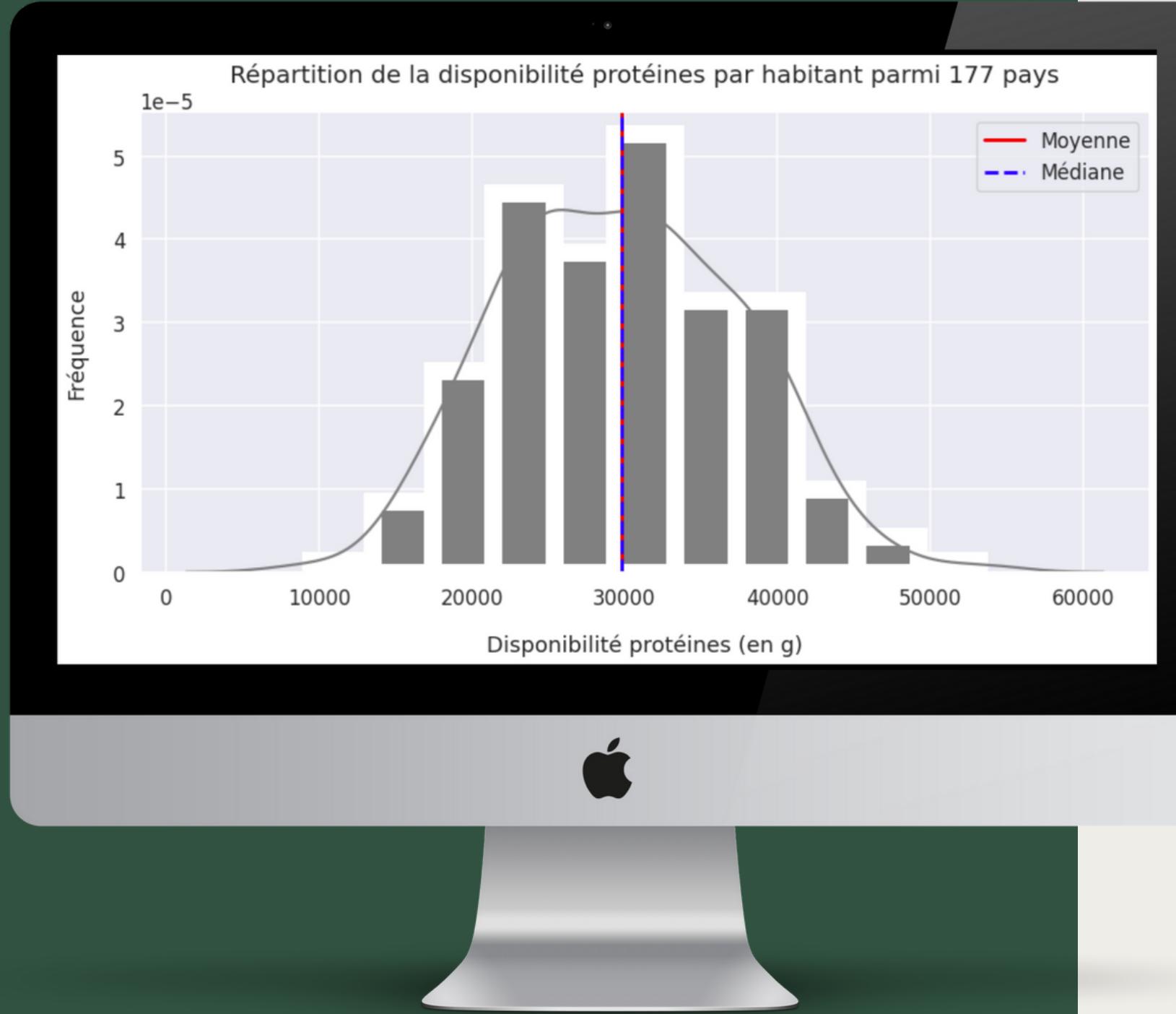


# MÉTHODOLOGIE

*pour les comparaisons des moyennes*



# Test d'adéquation de Kolmogorov-Smirnov



Non-rejet de l'hypothèse nul  $H_0$ ,  
la distribution suit une loi normale



Vérification d'une variable dont la loi est normale.  
Hypothèse testée:

$$\begin{cases} H_0: F = F_0 \\ H_1: F \neq F_0 \end{cases}$$

$D = -1.89$     $N = 177$    **critical value = 0.10**  
 $p\text{-value} = 0.99$     $a = 0.05$    **KS table = 1.36**



"La variable  
Disponibilité alimentaire  
en protéines

suit une loi normale et sera par conséquent choisi pour ce test."

- *clusters réellement distinct?*

---

Test de comparaison de 2 populations  
dans le cas gaussien

# TEST STATISTIQUE

avec *Bartlett et Student T-test*

## Bartlett Test

Pour vérifier l'égalité des variances des 2 groupes

$H_0$  : égalité variance      seuil : 0.05  
 $H_1$  : pas d'égalité      p-value : 0.95

## Résultat :

On ne peut pas rejeter  $H_0$  au niveau de test à 5%



# TEST STATISTIQUE

avec *Bartlett et Student T-test*

## Student T-test

Pour vérifier l'égalité des moyennes des 2 groupes

$H_0$  : égalité moyenne

seuil : 0.05

$H_1$  : pas d'égalité

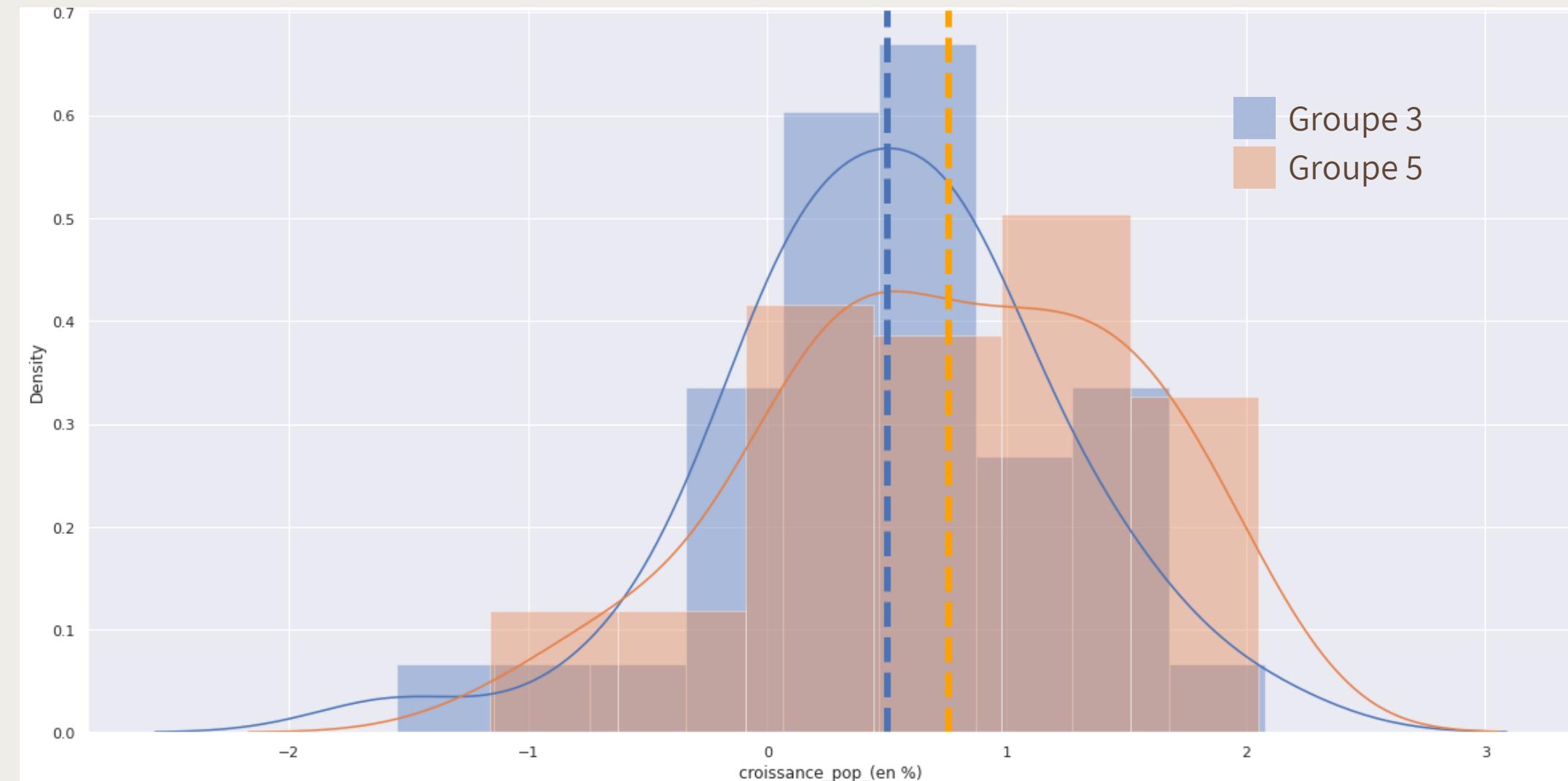
p-value : 1.4781e-19

## Résultat :

P-value extrêmement petit, on peut rejeter  $H_0$

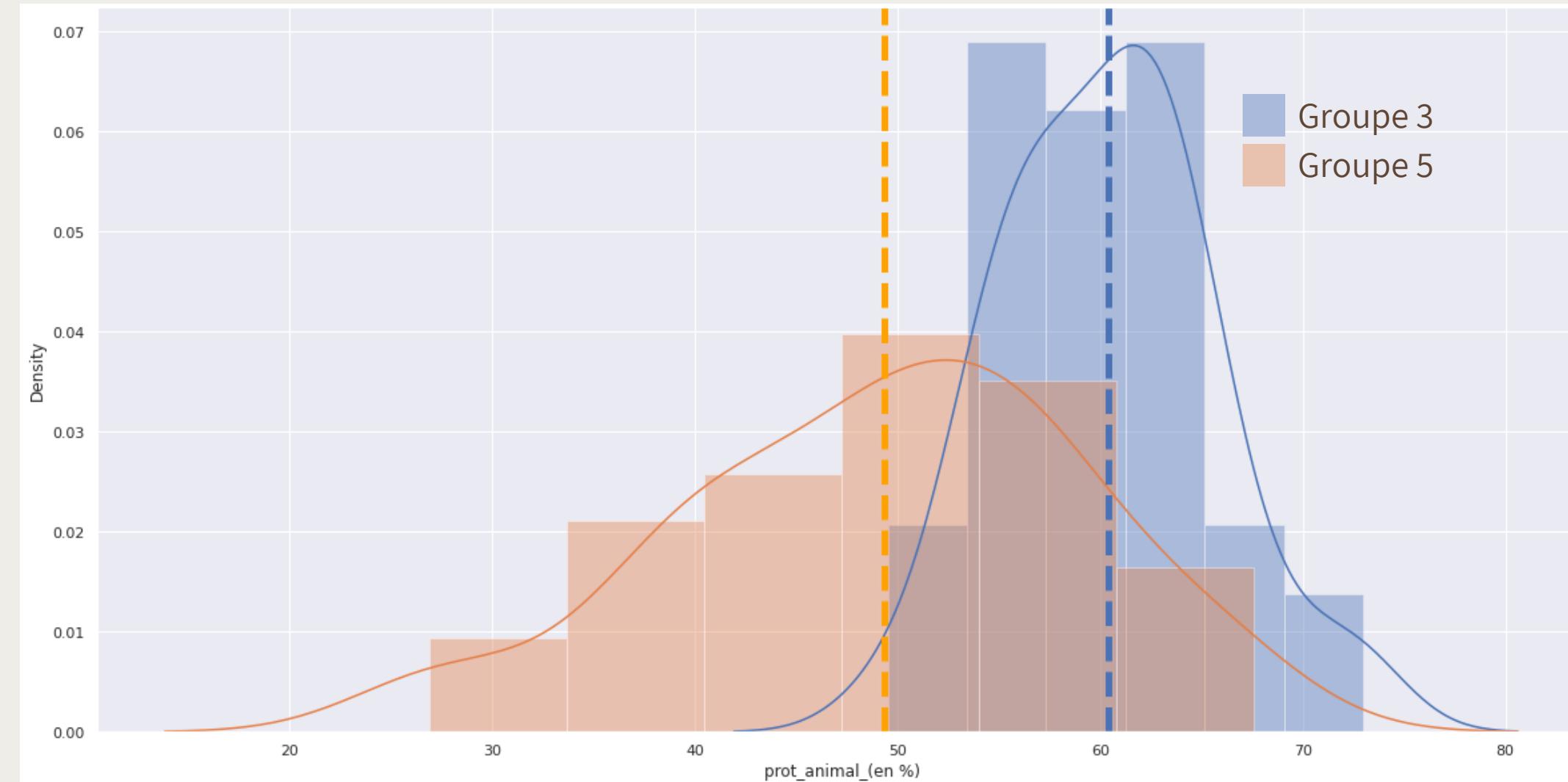


# Comparaison Moyenne



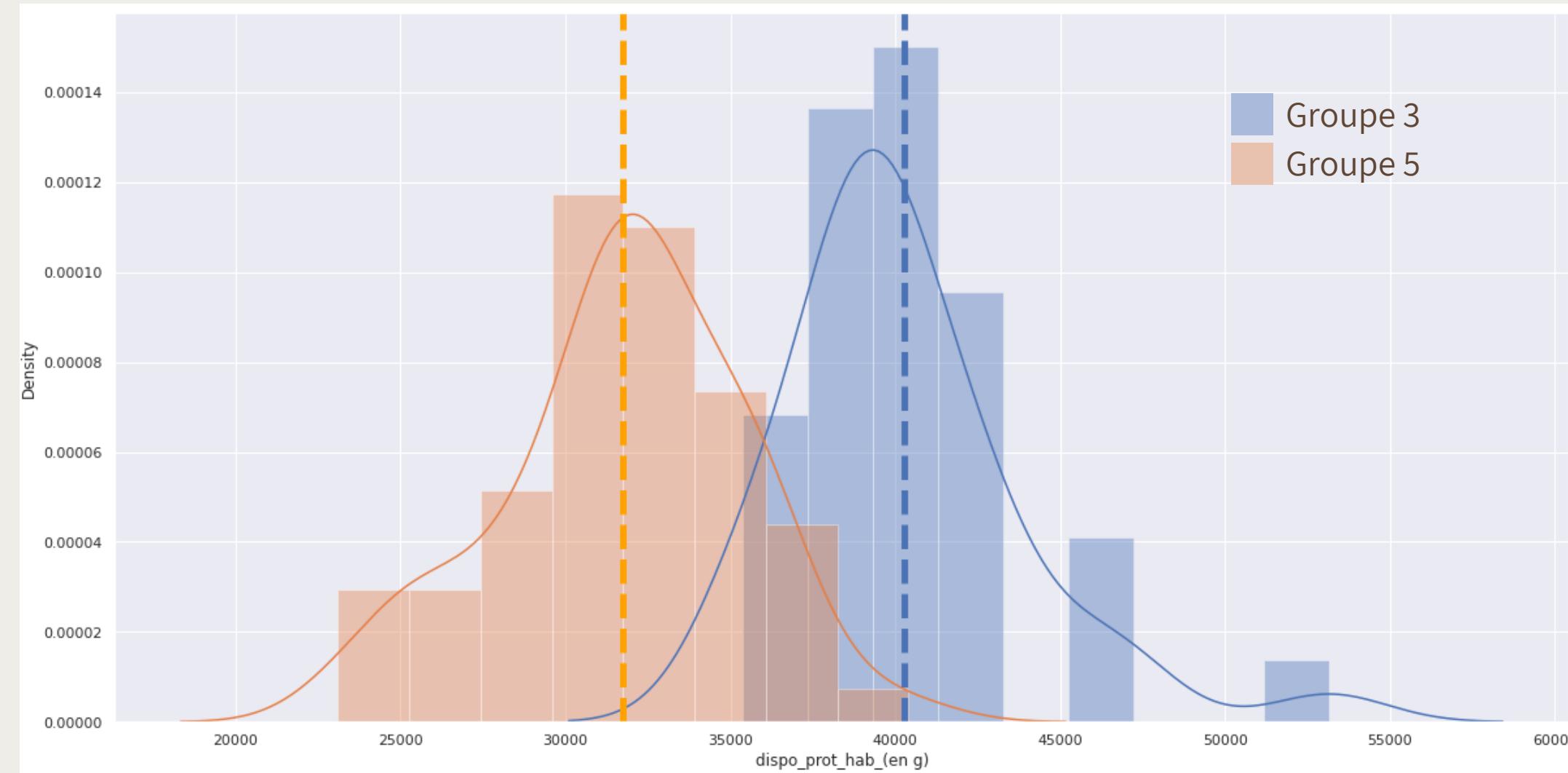
*Évolution de la population 2017-2018*

# Comparaison Moyenne



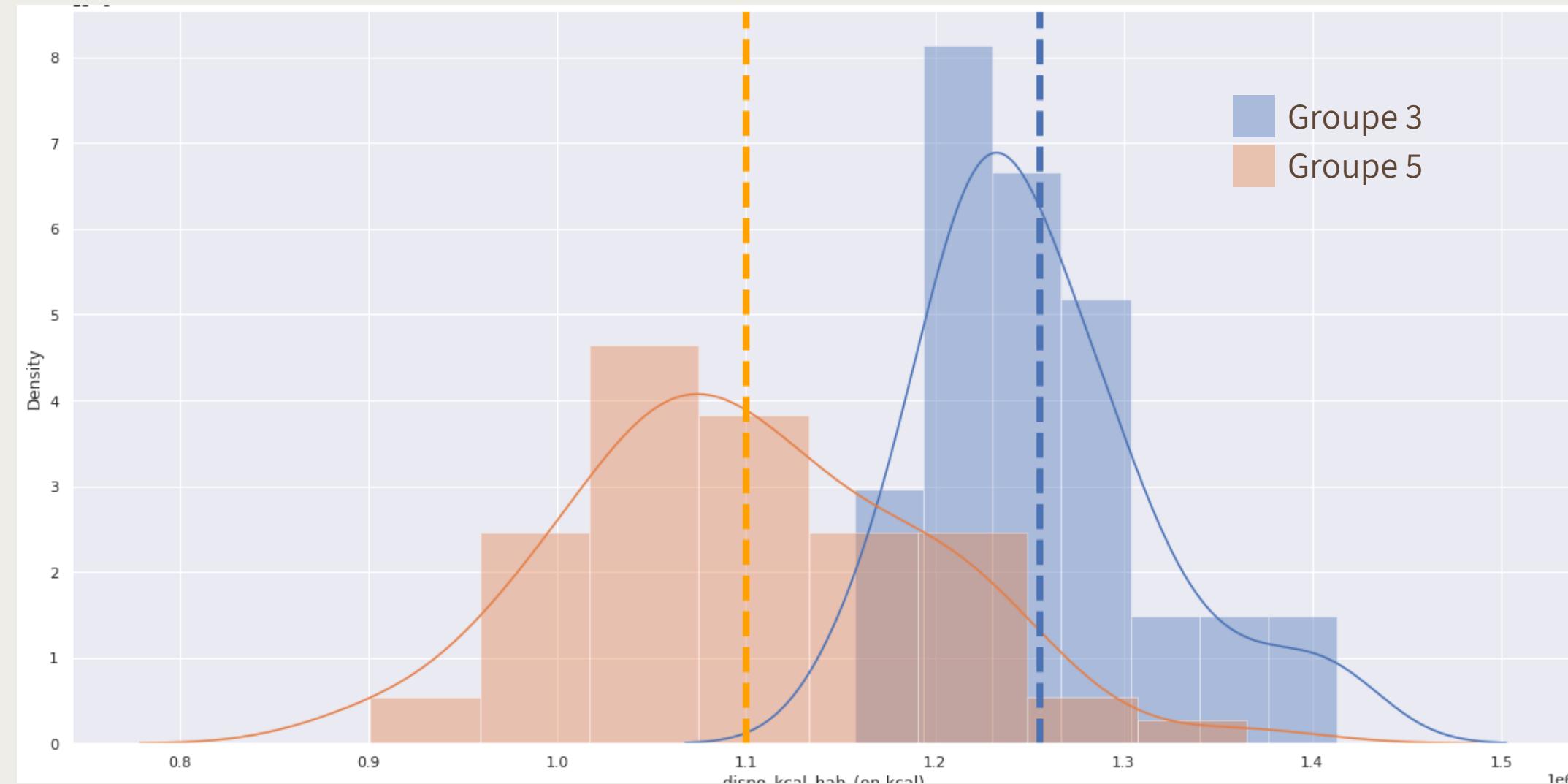
*Proportion de protéines d'origine animale*

# Comparaison Moyenne

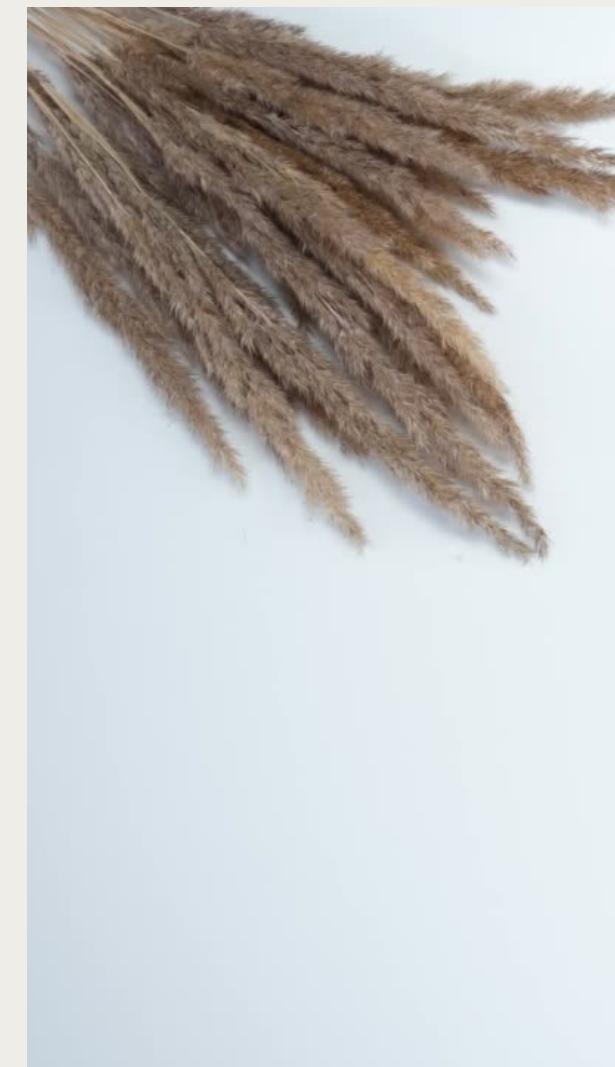
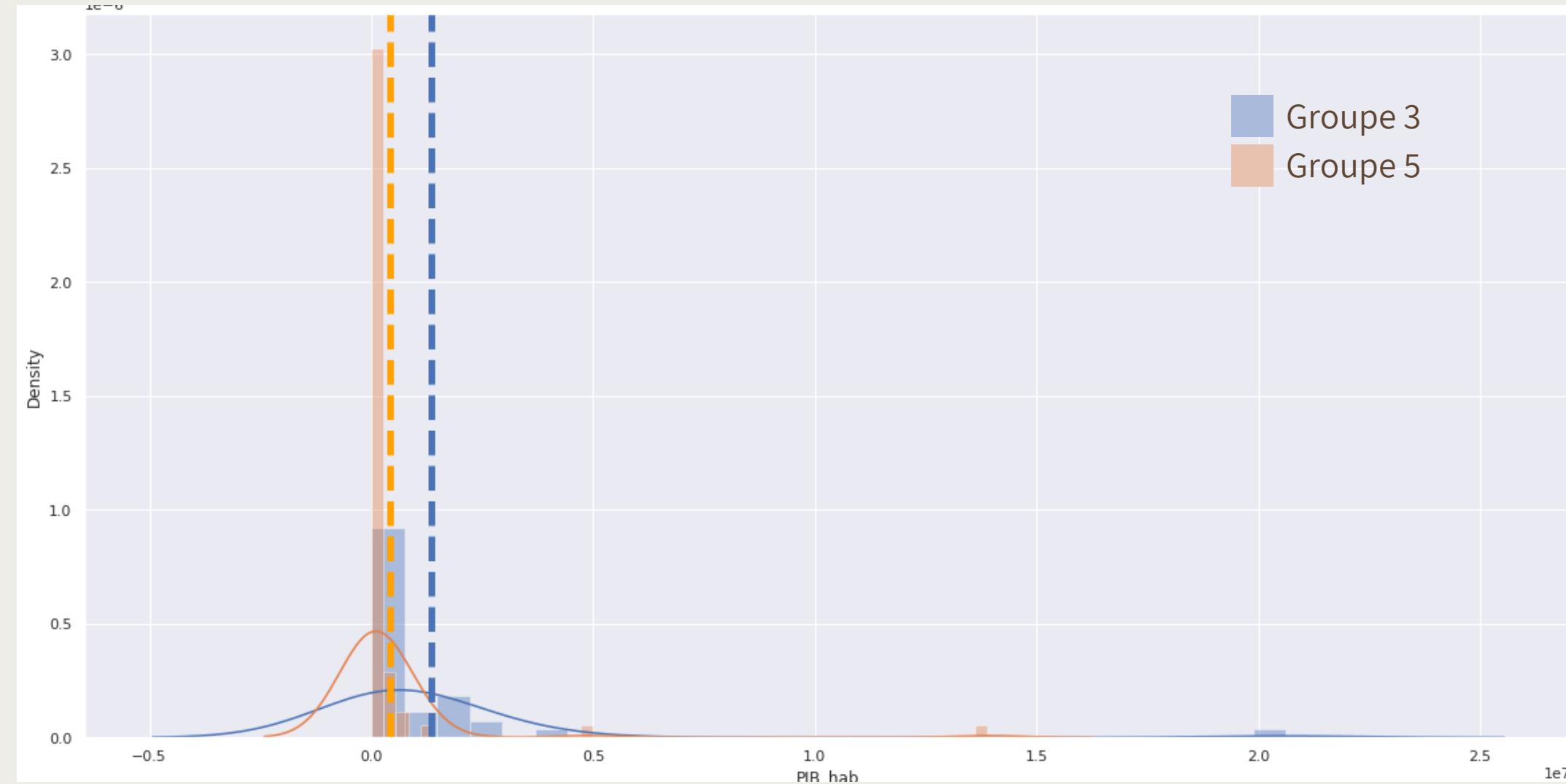


*Disponibilité alimentaire en protéines*

# Comparaison Moyenne



# Comparaison Moyenne



*PIB*

# Choix

- 01 Restriction aux pays du cluster 3
- 02 Disponibilité alimentaire forte
- 03 Indicateurs économiques intéressants

# LISTE DES PAYS

*Groupe 3*

## EUROPE

---

Albanie	Lituanie
Allemagne	Luxembourg
Autriche	Malte
Belgique	Monténégro
Danemark	Norvège
Espagne	Pays-Bas
Estonie	Pologne
Finlande	Portugal
France	Roumanie
Fédération de Russie	Royaume-Uni de Grande-Bretagne et d'Irlande du Nord
Grèce	Suisse
Irlande	Suède
Islande	
Italie	

## AMÉRIQUE DU NORD

---

Canada  
États-Unis d'Amérique

## AMÉRIQUE DU SUD

---

Argentine  
Brésil

## OCÉANIE

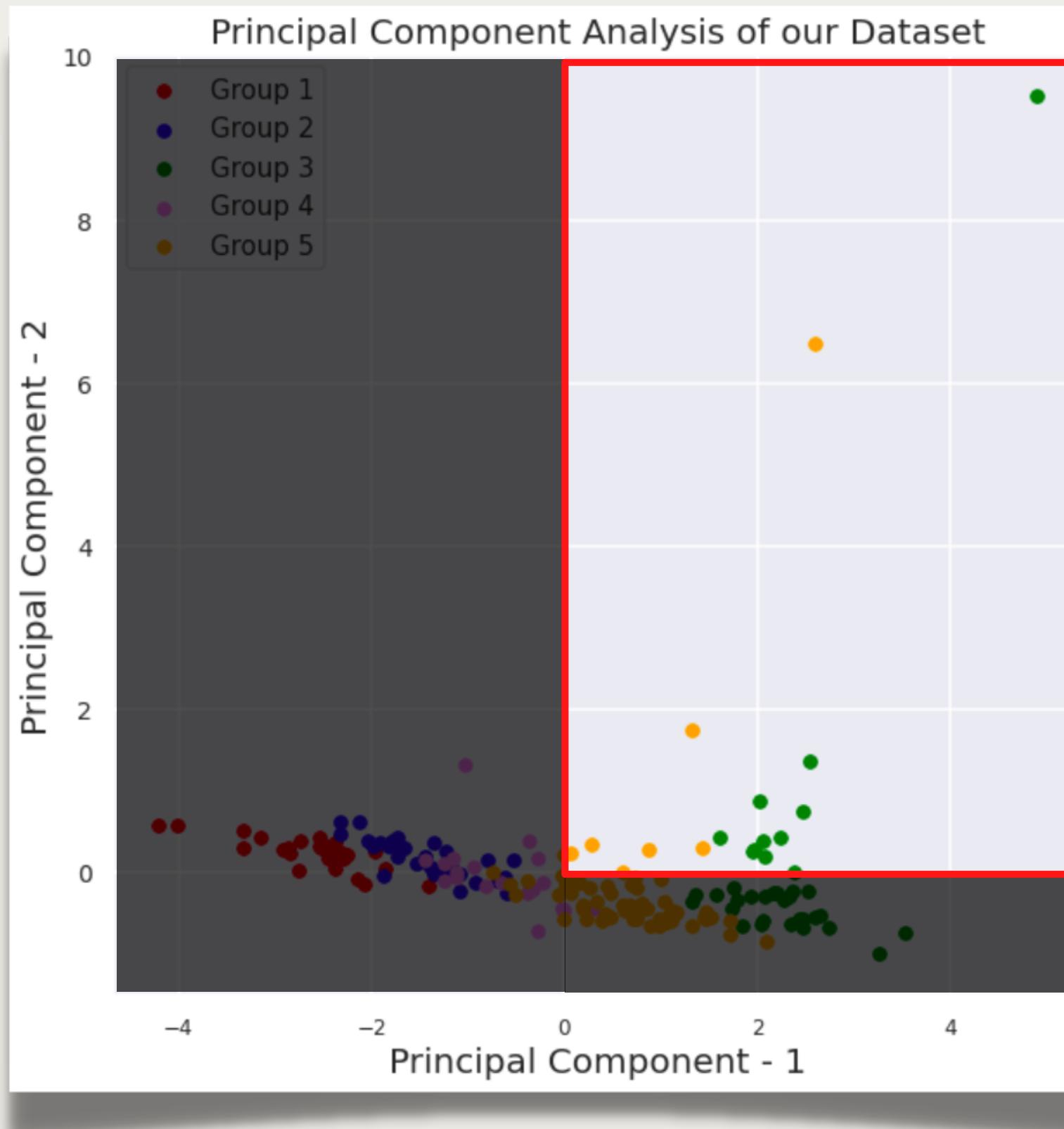
---

Australie  
Nouvelle-Zélande

## ASIE

---

Hong-Kong  
Macao  
Israël  
Kazakhstan  
République de Corée



# GROUPE 3

## CHOIX DES PAYS

---

Australie	01
Brésil	02
Canada	03
France	04
Allemagne	05
Italie	06
République de Corée	07
Fédération de Russie	08
Royaume-Uni de Grande-Bretagne et d'Irlande du Nord	09
États-Unis d'Amérique	10

# Conclusion

- 01 Clusters par classification hiérarchique
- 02 Analyse sur 2 composantes
- 03 2 groupes sélectionnés distincts
- 04 10 pays à cibler

# RECOMMANDATION

## Pays à cibler

Envisager les 10 pays choisi dans le cluster 3

## Exportation

Première exportation à l'Europe, ensuite se développer hors l'Union Européenne

## Autre possibilité

Possibilité de se développer dans des pays à faible disponibilité alimentaire et/ou PIB



# IMPLANTATION

## Cibles envisageables (Groupe 2):

		pays	cluster
12	Bolivie (État plurinational de)	Group 2	
13	Botswana	Group 2	
15	Belize	Group 2	
30	Congo	Group 2	
40	Équateur	Group 2	
50	Gabon	Group 2	
62	Honduras	Group 2	
78	Jordanie	Group 2	
96	Maldives	Group 2	
106	Namibie	Group 2	
120	Papouasie-Nouvelle-Guinée	Group 2	

## Cibles envisageables (Groupe 4):

		pays	cluster
22	Cabo Verde	Group 4	
24	Sri Lanka	Group 4	
42	El Salvador	Group 4	
87	Liban	Group 4	
105	République de Moldova	Group 4	
113	Nicaragua	Group 4	
146	Eswatini	Group 4	
149	République arabe syrienne	Group 4	
152	Thaïlande	Group 4	
165	Venezuela (République bolivarienne du)	Group 4	

Groupe des pays avec une disponibilité alimentaire et PIB faible qui sont plutôt intéressants pour une implantation.



# FIN.

Avis | Questions | Réponses

**Katrina JUMADIAO**

Data Analyst

**OpenClassrooms**

Session 2021-2022

**Étude de marché**

Projet 5



