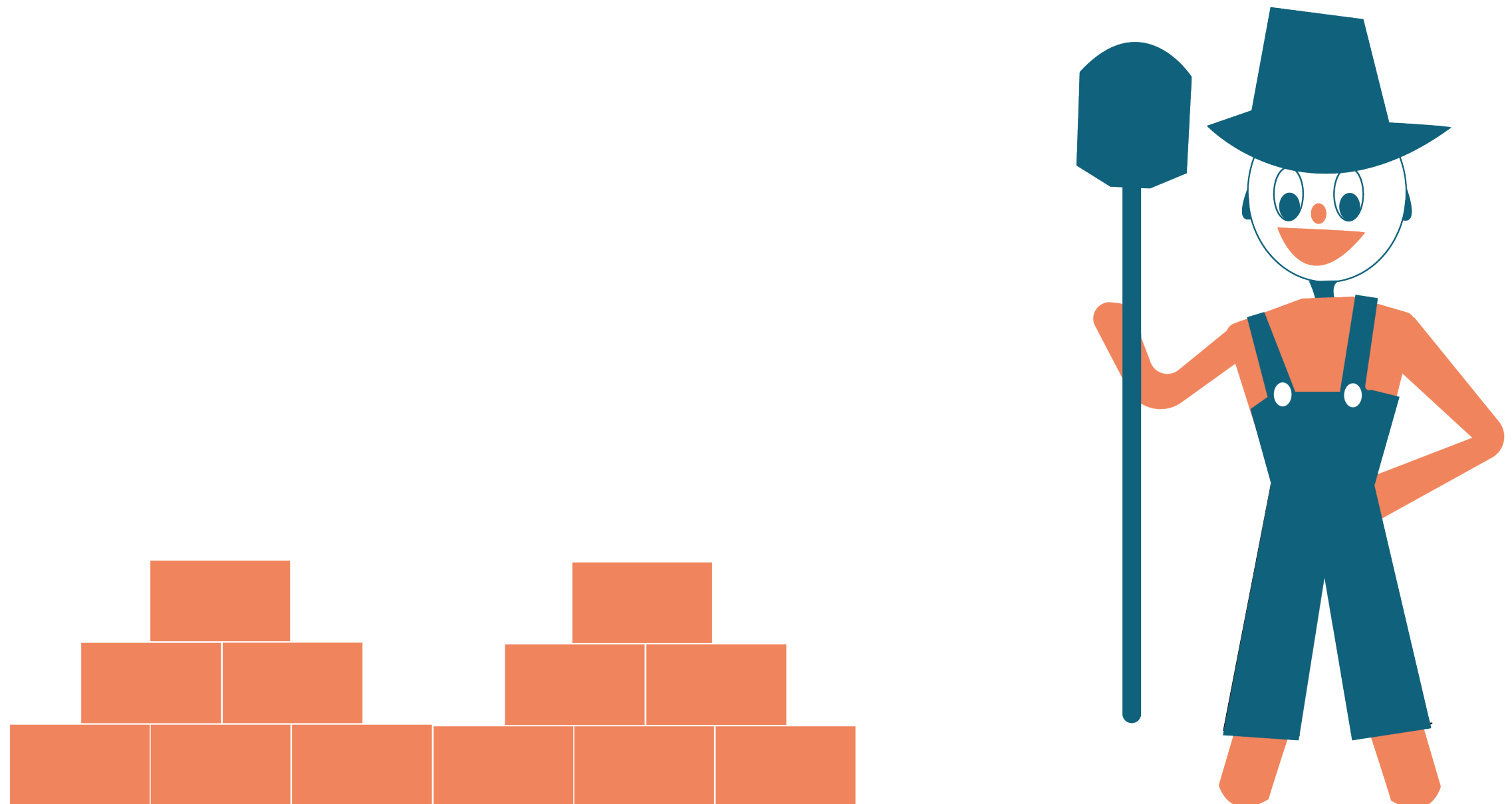


ÉTUDE DE MARCHÉ



OBJECTIF : IDENTIFIER LES PAYS PROPICES À
UNE INSERTION DANS LE MARCHÉ DU POULET.

CONSTRUCTION DES DONNÉES



SOURCE DES DONNÉES : SITE DE LA FAO

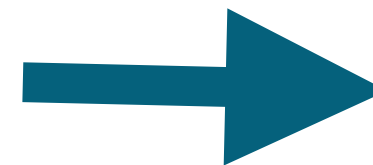
IMPORTÉES DE L'INTERFACE DU SITE DE LA FAO



Organisation des Nations Unies
pour l'alimentation et l'agriculture

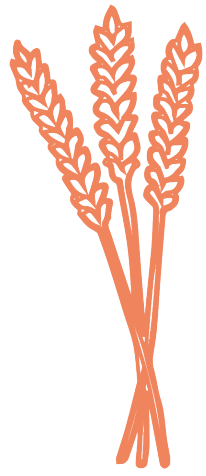


POPULATION 2012 ET 2013
PAR PAYS EN MILLIERS



pourcentage_dif_pop_2012_2013	
	3.260873
	1.430575
	-0.036348
	2.015503
	0.244768

CONSTRUCTION DE DF POPULATION AVEC CALCULE DU POURCENTAGE
DE DIFFÉRENCE DE POPULATION ENTRE L'ANNÉE 2013 ET L'ANNÉE 2012



PROTÉINES
VÉGÉTALE



PROTÉINES
ANIMALE

CRÉATION DE LA VARIABLE
« RATIO PROTÉINES ANIMAL
PROTÉINES TOTAL »



	pays	Disponibilité_protéines(g/personne/jour)	Disponibilité_alimentaire(Kcal/personne/jour)	ratio_prot_ani_prot_total
0	Afghanistan	58.26	2087.0	20.957775
1	Afrique du Sud	85.33	3020.0	42.646197
2	Albanie	111.37	3188.0	53.344707
3	Algérie	91.92	3293.0	27.175805
4	Allemagne	101.39	3503.0	60.647007



REPLISSAGE DES VALEURS MANQUANTES PAR LA MOYENNE

DATAFRAME FINAL

	pays	Disponibilité_protéines(g/personne/jour)	Disponibilité_alimentaire(Kcal/personne/jour)	ratio_prot_ani_prot_total	pourcentage_dif_pop_2012_2013
0	Afghanistan	58.26	2087.0	20.957775	3.260873
1	Afrique du Sud	85.33	3020.0	42.646197	1.430575
2	Albanie	111.37	3188.0	53.344707	-0.036348
3	Algérie	91.92	3293.0	27.175805	2.015503
4	Allemagne	101.39	3503.0	60.647007	0.244768

DATAFRAME AVEC LES VARIABLES PIB EN EUROS, EXPORTATIONS ET IMPORTATIONS

	PIB_en_euros	exportations	importations	Disponibilité_protéines(g/personne/jour)	Disponibilité_alimentaire(Kcal/personne/jour)	ratio_prot_ani_prot_total	pourcentage_dif_pop_2012_2013
count	230.000000	230.000000	230.000000	230.000000	230.000000	230.000000	230.000000
mean	15537.993279	6796.021739	6217.313043	81.299598	2847.367816	42.768671	1.272012
std	22537.577484	31587.853827	25770.472779	17.352631	381.172327	14.039801	1.396311
min	122.139890	0.000000	0.000000	37.660000	1879.000000	9.825674	-3.087262
25%	2150.888539	0.000000	3.000000	69.757500	2624.000000	34.369379	0.384537
50%	7245.224014	0.000000	310.500000	81.299598	2847.367816	42.768671	1.176439
75%	16273.413131	304.250000	2135.250000	90.285000	3065.000000	53.136857	2.218281
max	157910.012504	330181.000000	27111.000000	133.060000	3770.000000	73.138607	6.817495

CORRESPOND À LA QUANTITÉ DES POULETS EXPORTÉS ET IMPORTÉS PAR CHAQUE PAYS



CLASSIFICATION ASCENDANTE HIÉRARCHIQUE

MISE EN ŒUVRE D'UNE CLASSIFICATION :

CHOIX DE LA MESURE D'ÉLOIGNEMENT (DISTANCE EUCLIDIENNE) ENTRE INDIVIDUS

CHOIX DU CRITÈRE D'HOMOGÉNÉITÉ DES CLASSES À OPTIMISER (INERTIE).

CHOIX DE LA MÉTHODE UTILISÉE

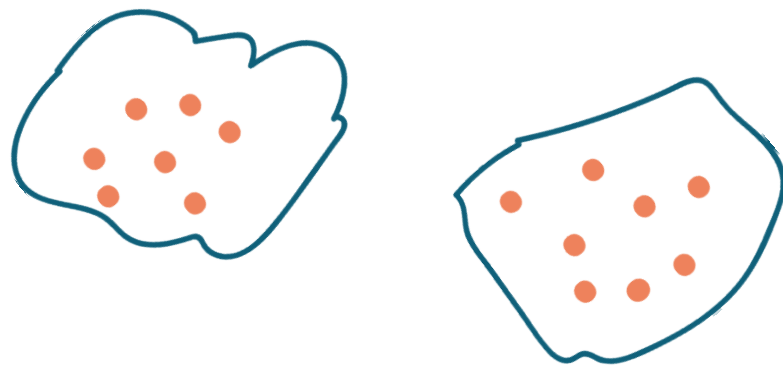
MESURE DE LA QUALITÉ DE LA CLASSIFICATION

CHOIX DU NOMBRE DE CLASSES ET LEUR INTERPRÉTATION

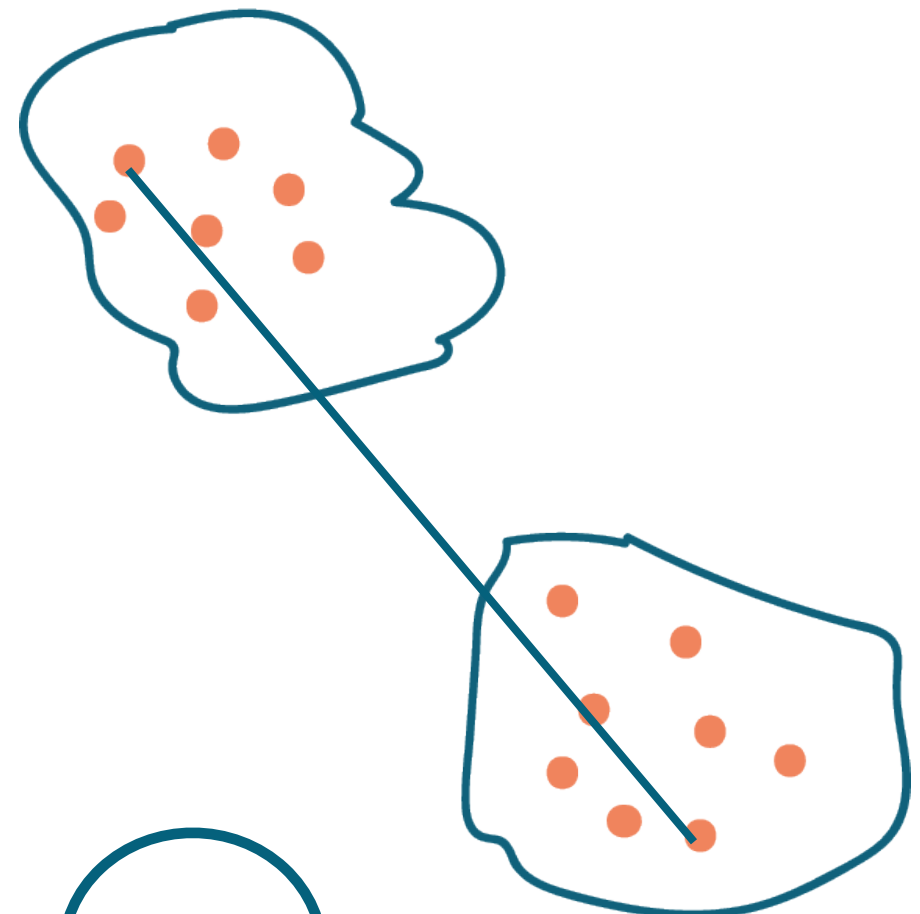
STRATÉGIE D'AGRÉGATION

RESSEMBLANCE ENTRE
GROUPES D'INDIVIDUS :
CRITÈRE DE WARD

RESSEMBLANCE ENTRE INDIVIDUS :
DISTANCE EUCLIDIENNE

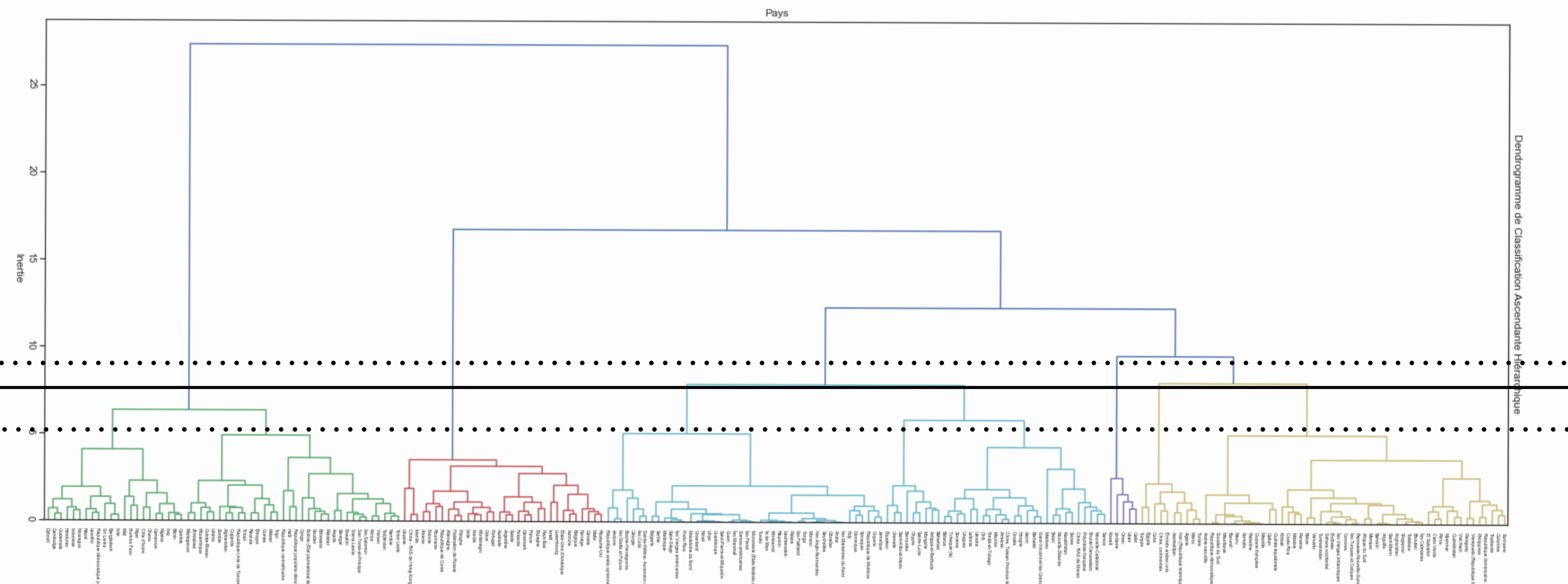


ON AGRÈGE À CHAQUE ITÉRATION LES
CLASSES DONT L'AGRÉGATION FAIT
PERDRE LE MOINS D'INERTIE
INTERCLASSE



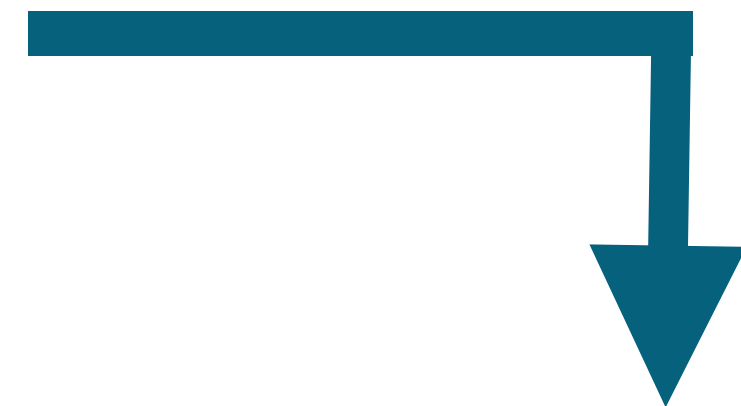
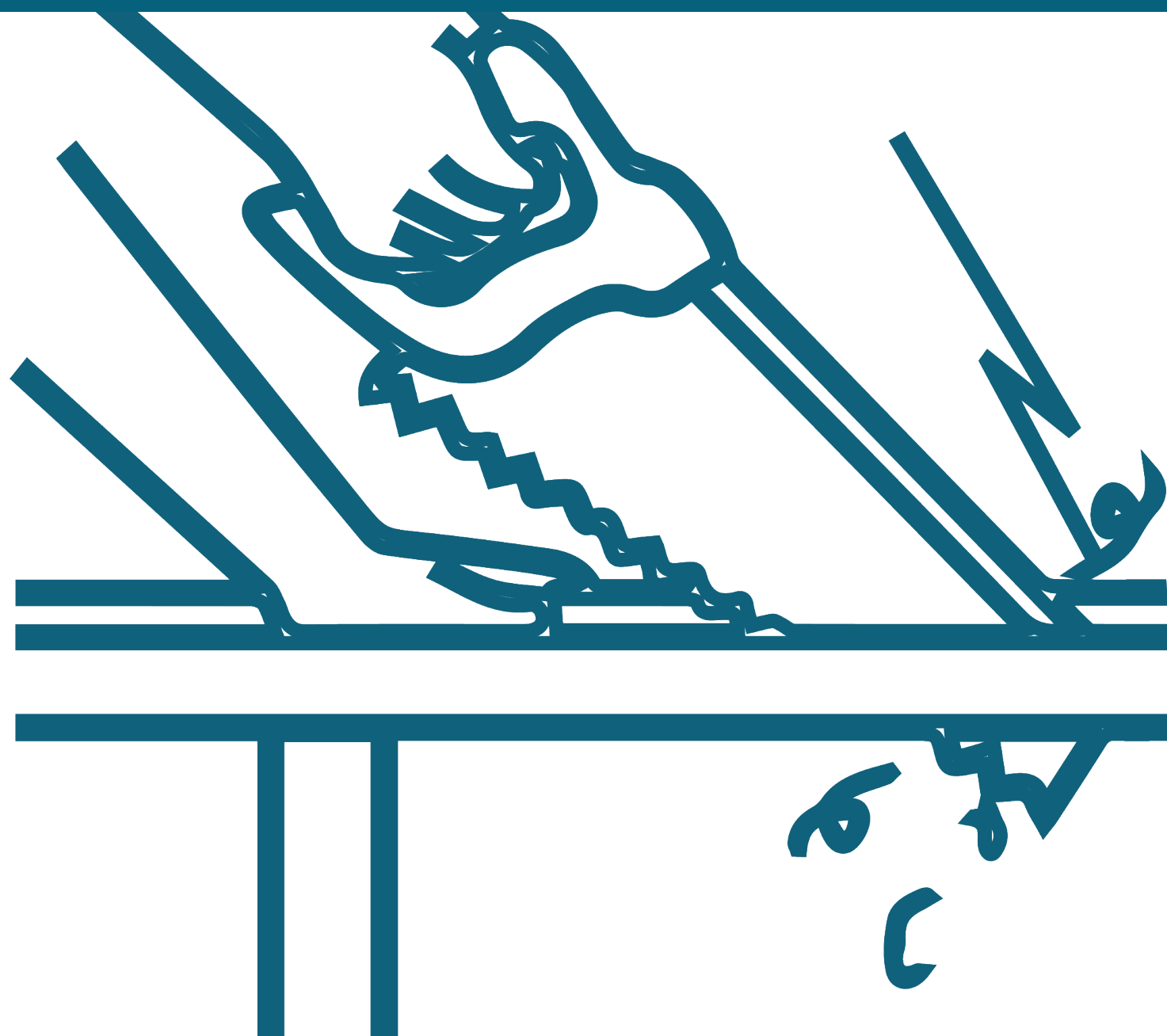
```
z = sch.linkage(X ,method='ward',metric='euclidean')
```

VISUALISATION DU DENDROGRAMME



LE DENDROGRAMME SUGGÈRE UN DÉCOUPAGE EN 5 GROUPES
ON AURAIT PU ENVISAGER AUSSI UN DÉCOUPAGE EN 4 GROUPES
SEULEMENT OU EN 8 GROUPES

DÉCOUPAGE EN CLASSES – MATÉRIALISATION DES GROUPES



	cluster	pays
0	1	Afghanistan
1	1	Nicaragua
2	1	Namibie
3	1	Mozambique
4	1	Mali

DÉCOUPAGE DU DENDROGRAMME ET
CRÉATION DATAFRAME AVEC CLUSTERS



INTERPRETATION DE CLASSES

L'IDÉE EST DE COMPARER LES MOYENNES DES VARIABLES ACTIVES CONDITIONNELLEMENT AUX GROUPES

MOYENNES

CARACTÉRISATION DES CINQ GROUPES

	PIB	EXPO	IMPO	DIS. PROTÉINES	DIS. ALIM.	RATIO PROTÉINES	DIS. PROTÉINES
GROUPE 1	↓	↓	↓	↓	↓	↓	↑
GROUPE 2	↑	↑	↑	↑	↑	↑	↓
GROUPE 3	↑	↓	↓	↑	↑	↑	↓
GROUPE 4	↑	↓	↓	↓	↑	↑	↑
GROUPE 5	↓	↓	↓	↑	↑	↓	↑

POUR CHAQUE GROUPE ÉCART À LA MOYENNE PAR VARIABLE

CARACTÉRISTIQUES DU GROUPE DE PAYS CIBLÉS

1. PIB SUPÉRIEUR À LA MOYENNE
2. IMPORTATIONS DES POULETS SUPÉRIEUR À LA MOYENNE
3. CROISSANCE DE LA POPULATION POSITIVE

GROUPE 2

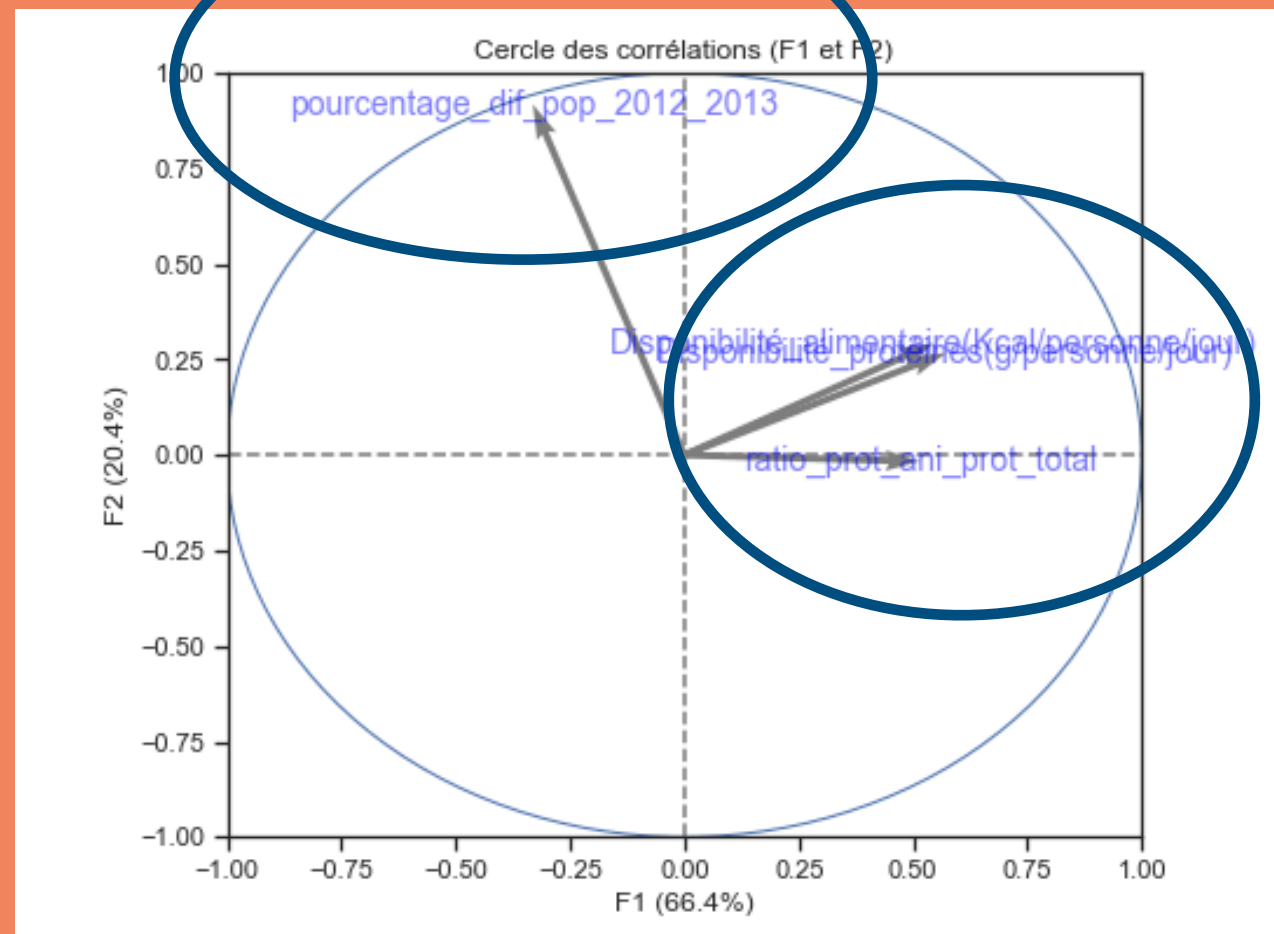
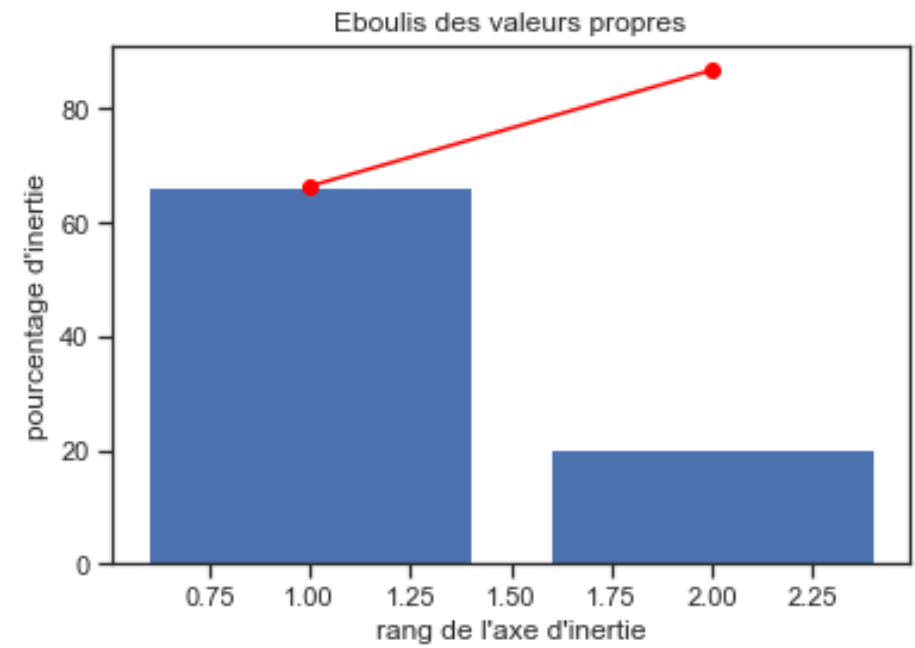
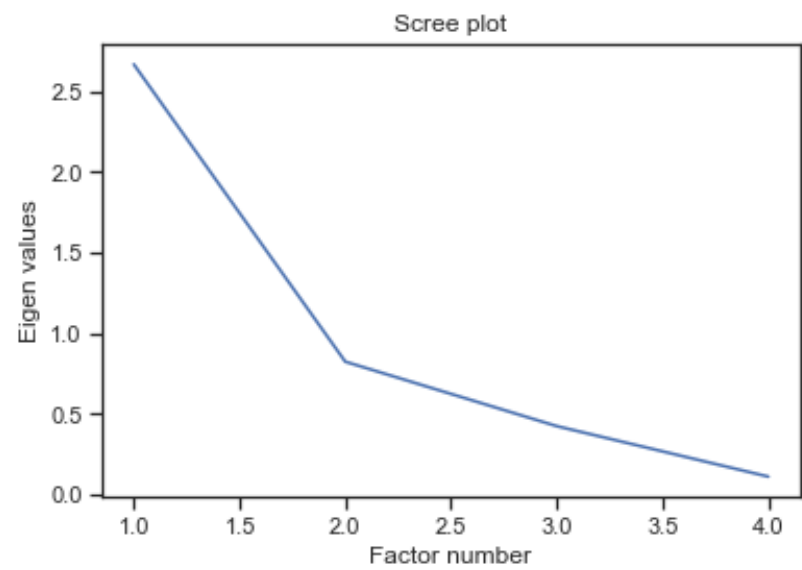
'ALBANIE', 'ALLEMAGNE', 'ARGENTINE', 'AUSTRALIE',
'AUTRICHE', 'BELGIQUE', 'CANADA', 'CHINE - RAS DE HONG-
KONG', 'DANEMARK', 'ESPAGNE', 'ESTONIE', 'ÉTATS-UNIS
D'AMÉRIQUE', 'FÉDÉRATION DE RUSSIE', 'FINLANDE',
'FRANCE', 'GRÈCE', 'IRLANDE', 'ISLANDE', 'ISRAËL', 'ITALIE',
'LITUANIE', 'LUXEMBOURG', 'MALTE', 'MONTÉNÉGRO',
'NORVÈGE', 'PAYS-BAS', 'POLOGNE', 'PORTUGAL',
'RÉPUBLIQUE DE CORÉE', 'ROUMANIE', 'ROYAUME-UNI',
'SUÈDE'



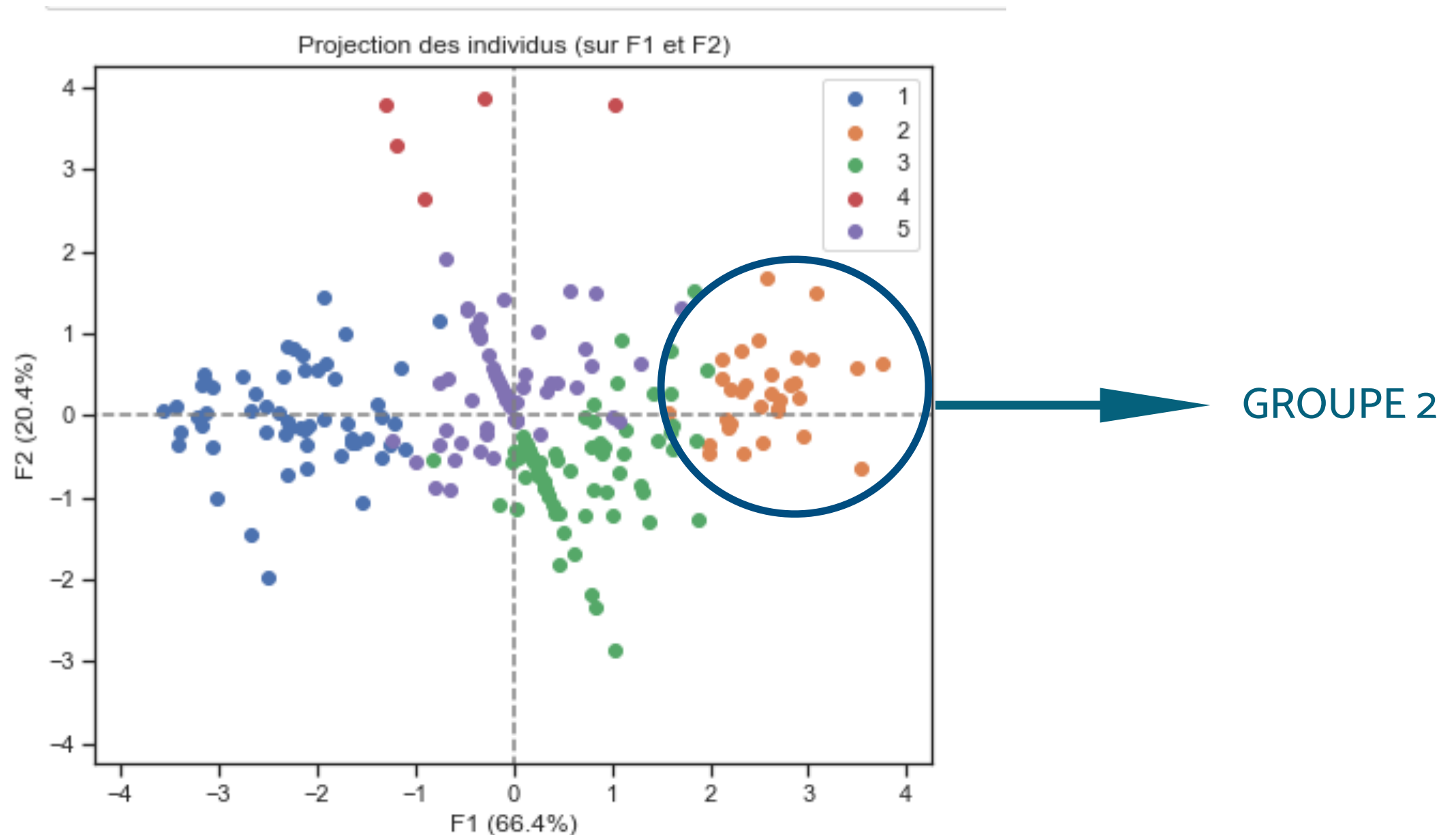


ANALYSE DE COMPOSANT PRINCIPALE

REDUCTION DE VARIABLES



VISUALISATION DES PARTITIONS DANS LE PREMIER PLAN FACTORIEL OBTENU PAR ACP



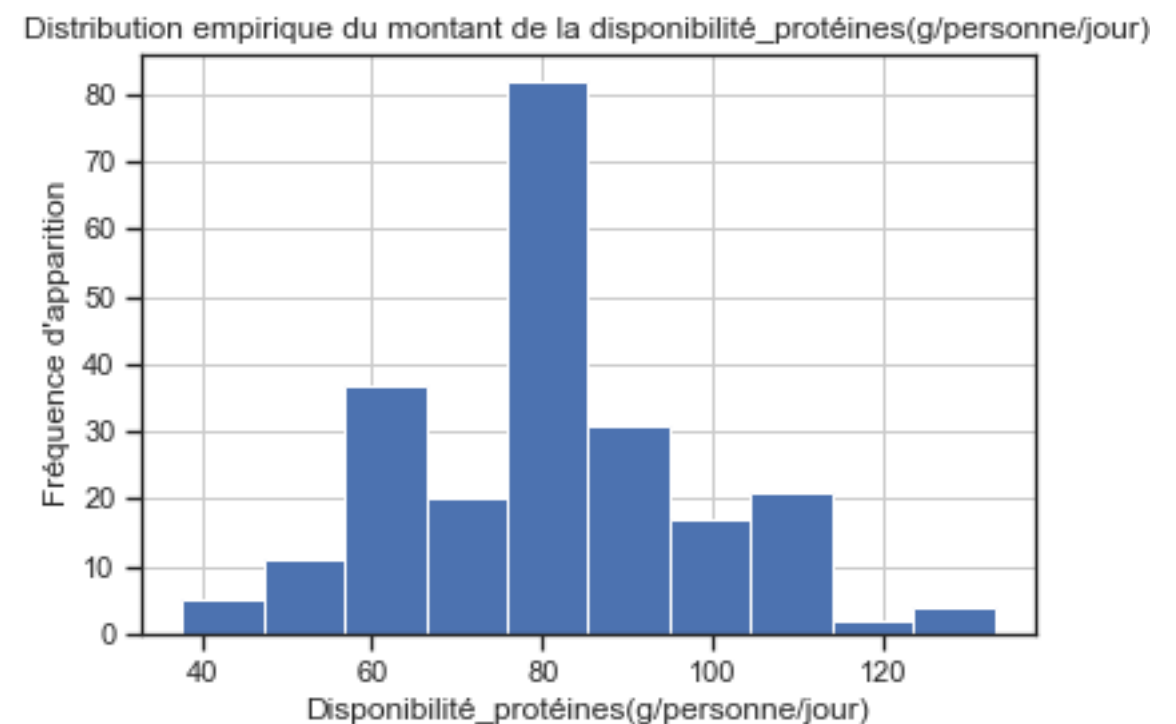
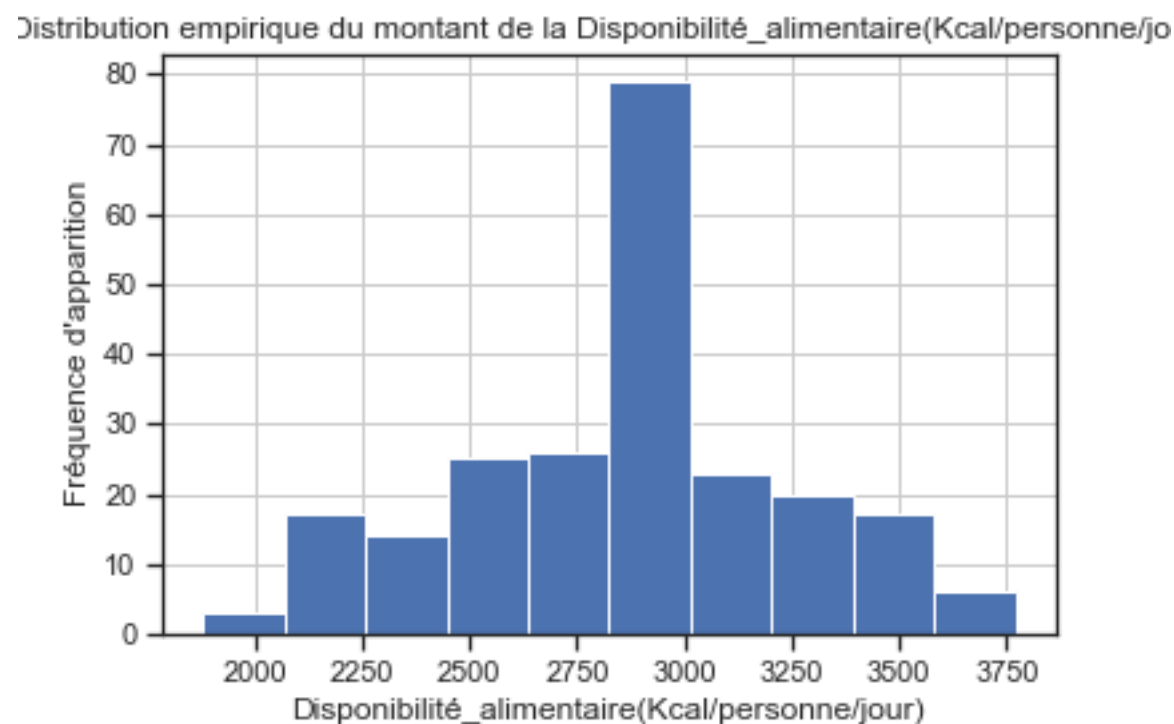
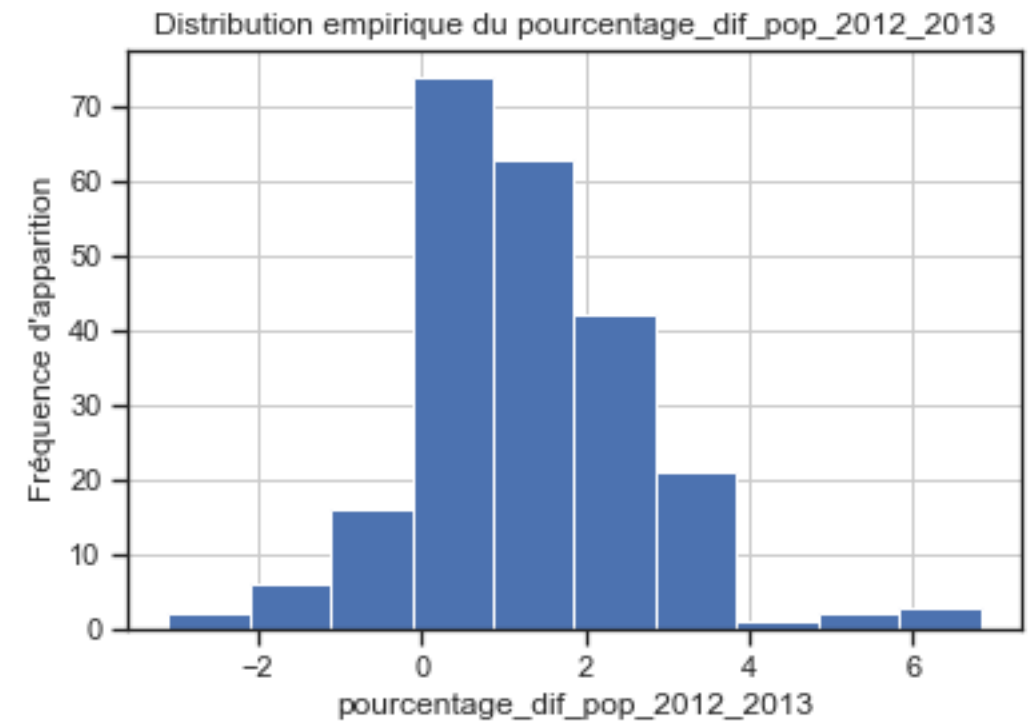
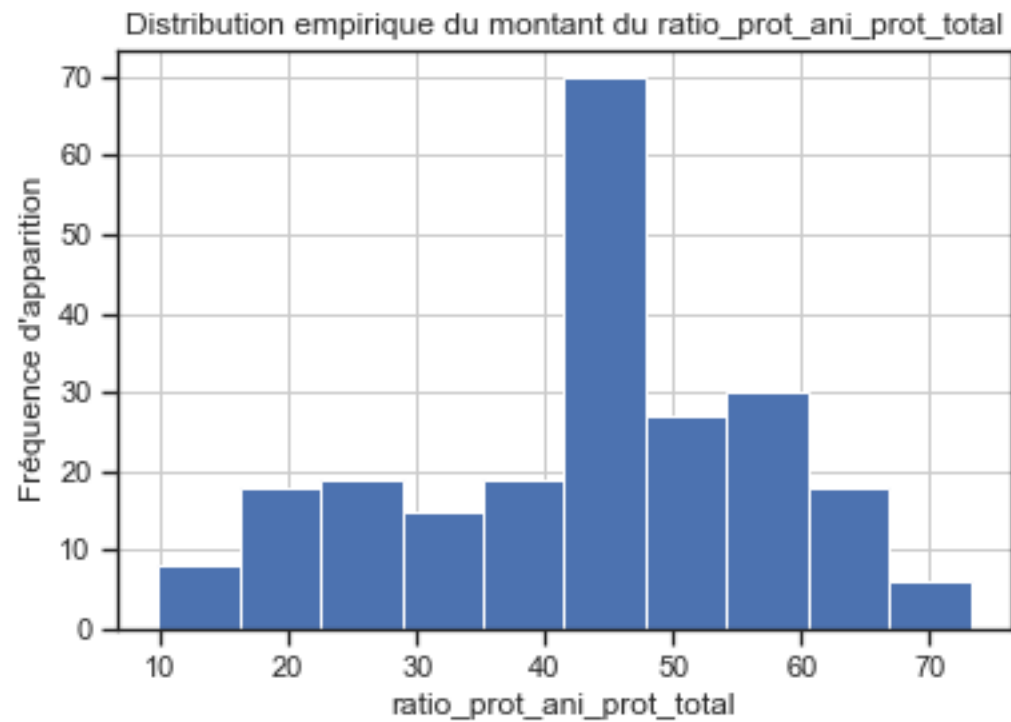
COMPOSANT 1: VARIABLES NUTRITIONNELLES

COMPOSANT 2: VARIABLE POPULATION



A LA RECHERCHE D'UNE VARIABLE GAUSSIENNE

HISTOGRAMME DES VARIABLES



LES TEST DE NORMALITÉ

H₀ LA VARIABLES EST GAUSSIENNE.

H₁ LA VARIABLE N'EST PAS GAUSSIENNE.

NOUS FIXONS NOTRE SEUIL DE SIGNIFICATION À P_VALEUR =0,01,

VARIABLE	P_VALEUR	REJET H ₀
DISPONIBILITÉ_PROTÉINES	0.24	NON
DISPONIBILITÉ_ALIMENTAIRE	0.94	NON
RATIO_PROT_ANI_PROT_TOTAL	0.05	NON
POURCENTAGE_DIF_POP_2012_2013	0.00	OUI



LES TESTS STATISTIQUES

VÉRIFIER QUE NOS GROUPES SOIENT SIGNIFICATIVEMENT DIFFÉRENTS

TEST DE COMPARAISON DE DEUX POPULATIONS

HYPOTHÈSE NULLE : LES DEUX VARIABLES SONT INDÉPENDANTES
HYPOTHÈSE ALTERNATIVE : LES DEUX VARIABLES SONT DÉPENDANTES
NOUS FIXONS NOTRE SEUIL DE SIGNIFICATION À P_VALEUR = 0,01

ANOVA

REJET H_0

DISPONIBILITÉ_PROTÉINES	P-VALUE = 0.00	OUI
DISPONIBILITÉ_ALIMENTAIRE	P-VALUE = 0.00	OUI
RATIO_PROT_ANI_PROT_TOTAL	P-VALUE = 0.00	OUI

TEST DE COMPARAISON DE DEUX POPULATIONS

POPULATION 1 = CLUSTER 2

POPULATION 1 = CLUSTER 3

VÉRIFIONS SI LES VARIANCES SONT SIMILAIRES.

HYPOTHÈSE NULLE : LES VARIANCES SONT SIMILAIRES.

HYPOTHÈSE ALTERNATIVE : LES VARIANCES SONT DIFFÉRENTES.

NOUS FIXONS NOTRE SEUIL DE SIGNIFICATION À P_VALEUR = 0,01,

TEST D'ADÉQUATION DE LA VARIANCE

REJET H_0

DISPONIBILITÉ_PROTÉINES

P-VALUE = 0.49

NON

DISPONIBILITÉ_ALIMENTAIRE

P-VALUE = 0.00

OUI

RATIO_PROT_ANI_PROT_TOTAL

P-VALUE = 0.00

OUI

TEST DE COMPARAISON DE DEUX POPULATIONS

POPULATION 1 = CLUSTER 2

POPULATION 1 = CLUSTER 3

VÉRIFIONS SI LES MOYENNES SONT SIMILAIRES.

HYPOTHÈSE NULLE : LES MOYENNES SONT SIMILAIRES.

HYPOTHÈSE ALTERNATIVE : LES MOYENNES SONT DIFFÉRENTES.

NOUS FIXONS NOTRE SEUIL DE SIGNIFICATION À P_VALEUR = 0,01,

TEST DE STUDENT - TEST D'ADÉQUATION DES MOYENNES

REJET H_0

DISPONIBILITÉ_ALIMENTAIRE

P-VALUE = 0.00

OUI

DISPONIBILITÉ_PROTÉINES

P-VALUE = 0.00

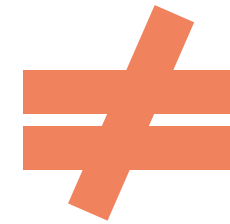
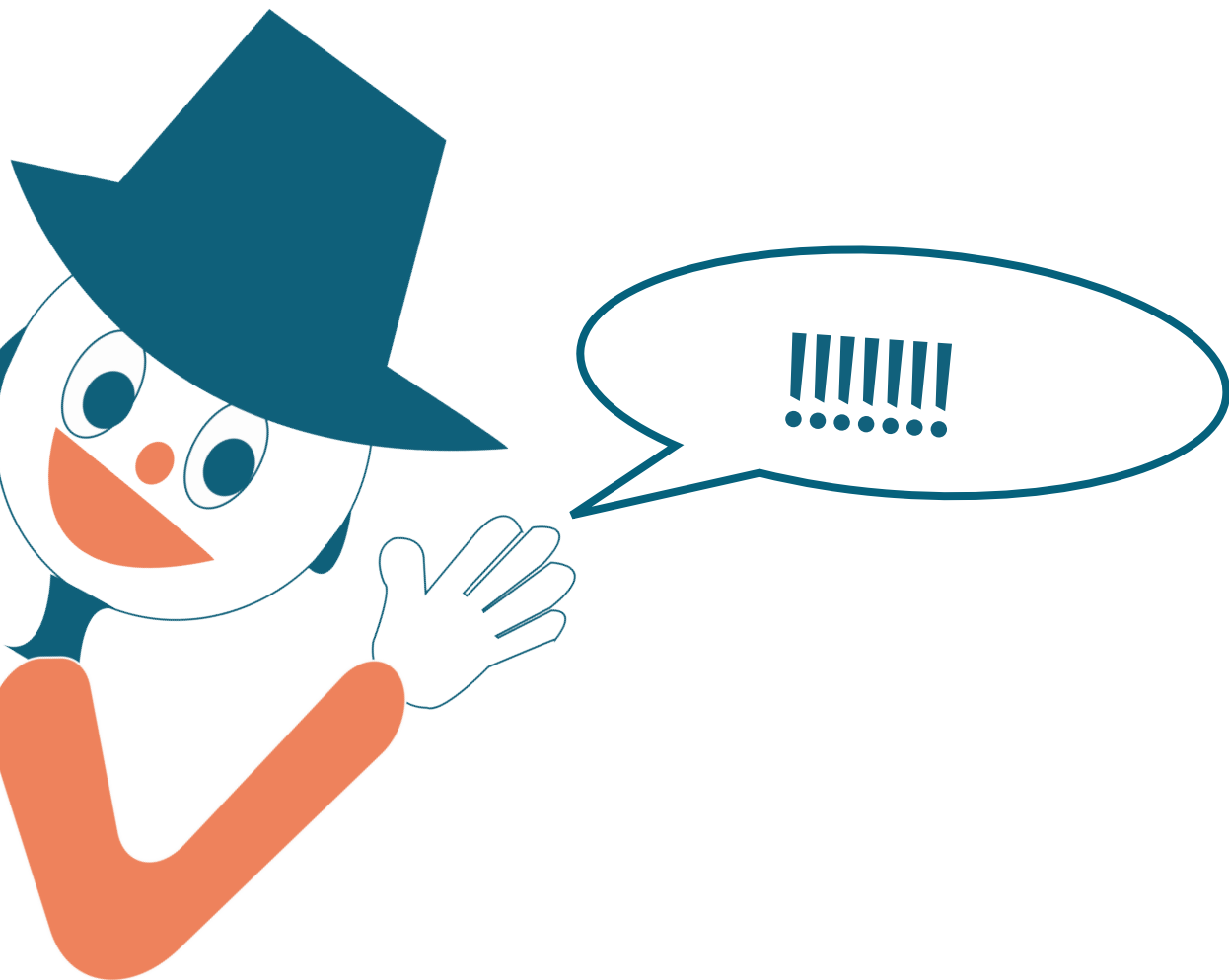
OUI

RATIO_PROT_ANI_PROT_TOTAL

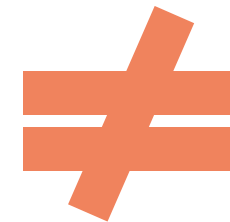
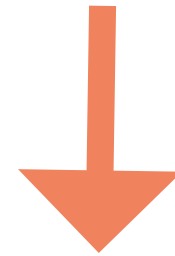
P-VALUE = 0.00

OUI

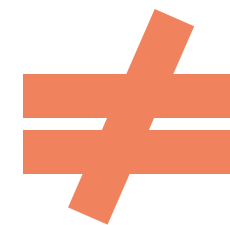
ON CONFIRME QUE CHAQUE CLUSTER EST SIGNIFICATIVEMENT DIFFÉRENT



CLUSTERS



VARIANCE

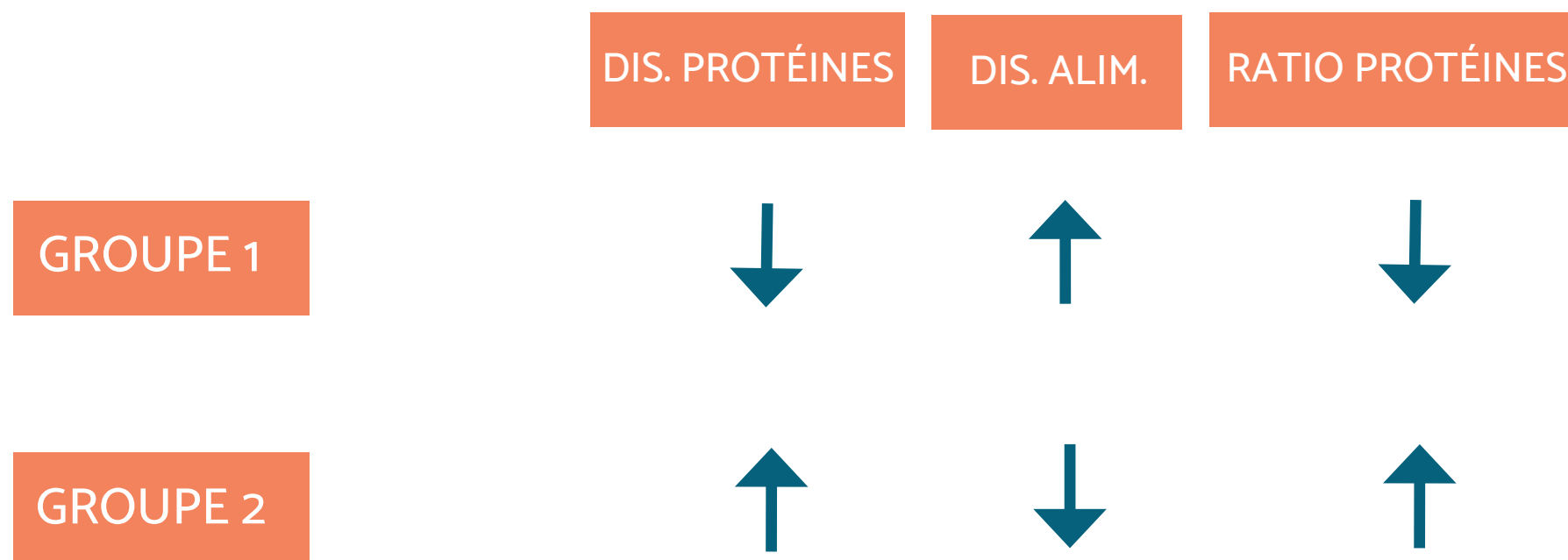


MOYENNES



RÉDUCTION DE PAYS

RELATION DE MOYENNES DE GROUPE PAR RAPPORT AUX MOYENNES GENERALES



CARACTÉRISTIQUES DU GROUPE DE PAYS CIBLÉS

1. FAIBLE RATIO PROTÉINES TOTALES / PROTÉINES ANIMALES

GROUPE 1

ROUMANIE- RÉPUBLIQUE DE CORÉE- PORTUGAL -
POLOGNE - NORVÈGE - MONTÉNÉGRO - MALTE - ITALIE -
IRLANDE -ROYAUME-UNI- FÉDÉRATION DE RUSSIE -
GRÈCE - ALLEMAGNE - ESTONIE - AUTRICHE - BELGIQUE -
CANADA - ÉTATS-UNIS D'AMÉRIQUE





...FIN