



The king of chicken[©]

P5 : Produisez une étude de marché

Sommaire

1. Introduction
2. Démarche adoptée
3. Présentation des résultats
4. Recommandations
5. Conclusion

1. Introduction

- **Marché fort ≈ 30 Milliards de US\$**
- **Exportations en croissance en 2017**
- **Rebonds malgré les crises de 2008 (crise économique mondiale des subprimes) et 2015-2016 (grippe aviaire)**
- **Prévisions pour 2018 et 2019 en croissance**



1. Introduction

- Afin de se développer à l'international, nous avons besoin de cibler particulièrement certains pays afin d'approfondir l'étude de marché
- L'objectif est d'exporter dans les pays cibles dans un premier temps

2. Démarche adoptée

- Afin de cibler les pays, nous avons besoin d'étudier en priorité les régimes alimentaires de chaque pays => construction d'un jeu échantillon contenant l'ensemble des pays disponibles à partir des données FAO
- Chaque pays sera caractérisé par les variables suivantes :
 - différence de population entre l'année 2017 et l'année 2016, exprimée en pourcentage ;
 - proportion de protéines d'origine animale par rapport à la quantité totale de protéines dans la disponibilité alimentaire du pays ;
 - disponibilité alimentaire en protéines par habitant ;
 - disponibilité alimentaire en calories par habitant.
 - PIB par habitant
 - Autosuffisance en kg/hab qui est la différence entre production et la consommation de poulet de chaque pays

2. Démarche adoptée

- Une fois l'échantillon créé, nous allons utiliser des méthodes de statistiques descriptives et de classification pour atteindre notre objectif :
 - La classification ascendante hiérarchique(CAH) : classer des individus ayant un comportement similaire sur un ensemble de variables et choisir ensuite le nombre de groupes qui conviendra le mieux => méthode de Ward : minimiser l'inertie intraclasse
 - L'Analyse en Composantes Principales(ACP) : revenir à un espace de dimension réduite en déformant le moins possible la réalité. Il s'agit donc d'obtenir le résumé le plus pertinent des données initiales => Etudes des individus(ressemblances entre individus) et étude des variables (liaisons entre variables)

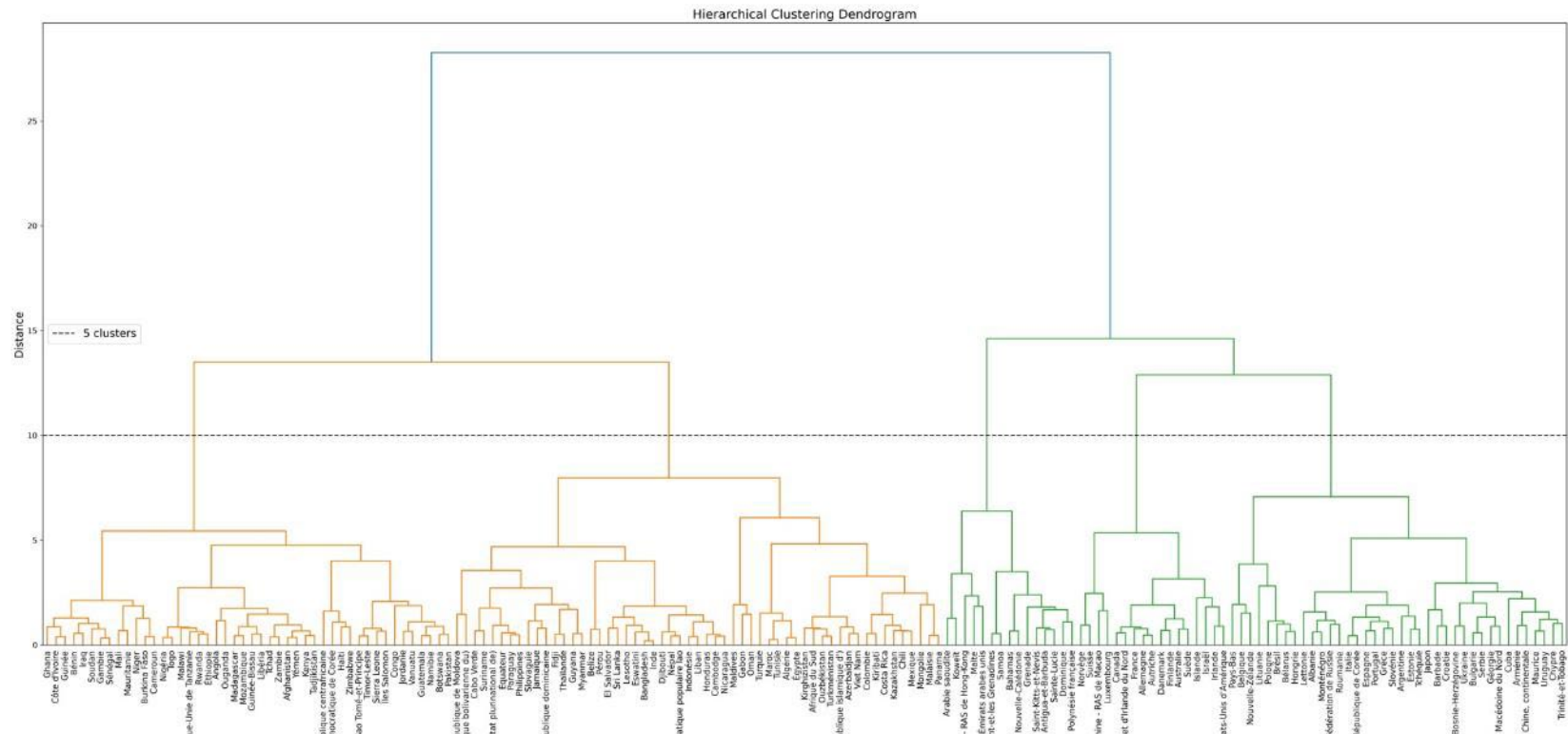
2. Démarche adoptée

Tableau de données nommé « df poulet » utilisé pour l'analyse

Zone	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Proportion de protéines animales (en %)	Diff pop. 2017-2016 (en %)	Autosuffisance (en kg/hab)	PIB par habitant (en US\$/hab)
Afghanistan	2050	56,16	19,84	2,52	-0,80	513
Afrique du Sud	2935	83,64	42,19	1,41	-7,95	6122
Albanie	3390	118,17	54,23	-0,08	-6,93	4514
Algérie	3342	92,82	28,78	2,02	-0,17	4110
Allemagne	3542	104,06	60,32	0,56	0,27	44552
Angola	2372	52,02	31,35	3,27	-9,29	4096
Antigua-et-Barbuda	2447	80,43	64,29	0,94	-52,40	15383
Arabie saoudite	3300	88,57	38,64	1,99	-21,72	20802
Argentine	3268	105,68	64,02	0,98	4,55	14628
Arménie	3072	97,71	46,68	0,29	-11,55	3915
Australie	3405	105,35	66,93	1,31	1,06	57629
Autriche	3687	107,42	59,53	0,82	-1,36	47306
Azerbaïdjan	3137	92,41	35,03	1,11	-2,54	4151
Bahamas	2644	81,62	64,30	1,00	-34,05	32719
Bangladesh	2562	60,68	20,73	1,07	-0,05	1538
Barbade	2911	87,30	57,03	0,15	10,48	17392
Belgique	3768	98,52	59,54	0,57	29,34	44020
Belize	2712	66,71	37,69	1,96	26,61	4888
Bolivie (état plurinational de)	2399	73,46	48,83	1,44	0,00	3351
Bosnie-Herzégovine	3273	97,07	37,31	-1,04	-0,60	5395
Botswana	2341	64,65	40,08	2,05	-0,91	7893
Brésil	3250	94,12	56,45	0,80	20,53	9925
Bulgarie	2826	85,48	48,34	-0,70	-7,04	8300
Burkina Faso	2721	77,67	14,38	2,85	0,00	735
Bélarus	3346	93,01	55,57	0,05	21,06	5791
Bénin	2757	64,23	20,27	2,71	-9,84	1137
Cabo Verde	2529	70,61	37,86	1,18	-14,88	3293
Cambodge	2502	66,20	29,09	1,52	-0,62	1385
Cameroun	2717	71,05	16,14	2,60	-0,04	1425
Canada	3500	104,71	52,94	0,95	-0,35	44907
Chili	3020	90,64	52,75	1,42	-1,79	14999
Chine - RAS de Hong-Kong	3270	135,91	73,29	0,86	-42,43	46705
Chine - RAS de Macao	3299	109,74	65,25	1,57	-27,31	81517
Chine	3179	101,04	39,96	0,49	0,00	8663
Chypre	3039	92,81	51,24	0,80	-4,24	26484

3. Présentation des résultats

- Afin de déterminer les différents clusters, on coupe les branches de l'arbre à une hauteur entre 2 nœuds relativement éloignés.
- On distingue 5 clusters
- Stratégie d'agrégation : Méthode de Ward



3. Présentation des résultats

Position des centroïdes :

Cluster n°	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Proportion de protéines animales (en %)	Diff pop_2017-2016 (en %)	Autosuffisance (en kg/hab)	PIB par habitant (en US\$/hab)
1	2382.1	60.1	24.7	2.5	-2.5	1680.0
2	2816.7	77.1	39.8	1.3	-0.8	5565.7
3	2959.7	93.4	58.1	0.9	-42.8	21950.6
4	3490.0	112.1	62.0	0.9	-1.3	60468.1
5	3215.8	96.9	53.2	-0.0	3.5	17139.1

Groupes 1 et 2 :

- consommation de protéines est relativement faible et en particulier la consommation de protéines animales,
- autosuffisance en poulet est proche de zéro ce qui indique un besoin assez faible en importation de poulet,
- PIB par habitant étant assez faible => ces pays sont généralement pauvres donc nous ne pourrions pas aligner nos prix d'exportation du poulet sur le niveau de vie local.

Groupes 4 et 5 :

- consomment beaucoup de protéines animales et qui ont des revenus élevés voire très élevés,
- Autosuffisants ou proches de l'autosuffisance pour la plupart ce qui indique un marché très concurrentiel et limité en terme part de marché.

Groupe 3 :

- Pays ayant une forte consommation de protéines animales, des revenus élevés et surtout une forte dépendance en matière de poulet, => Tous ces indicateurs nous orientent donc vers ce cluster.

3. Présentation des résultats

Liste préliminaire de pays à cibler :

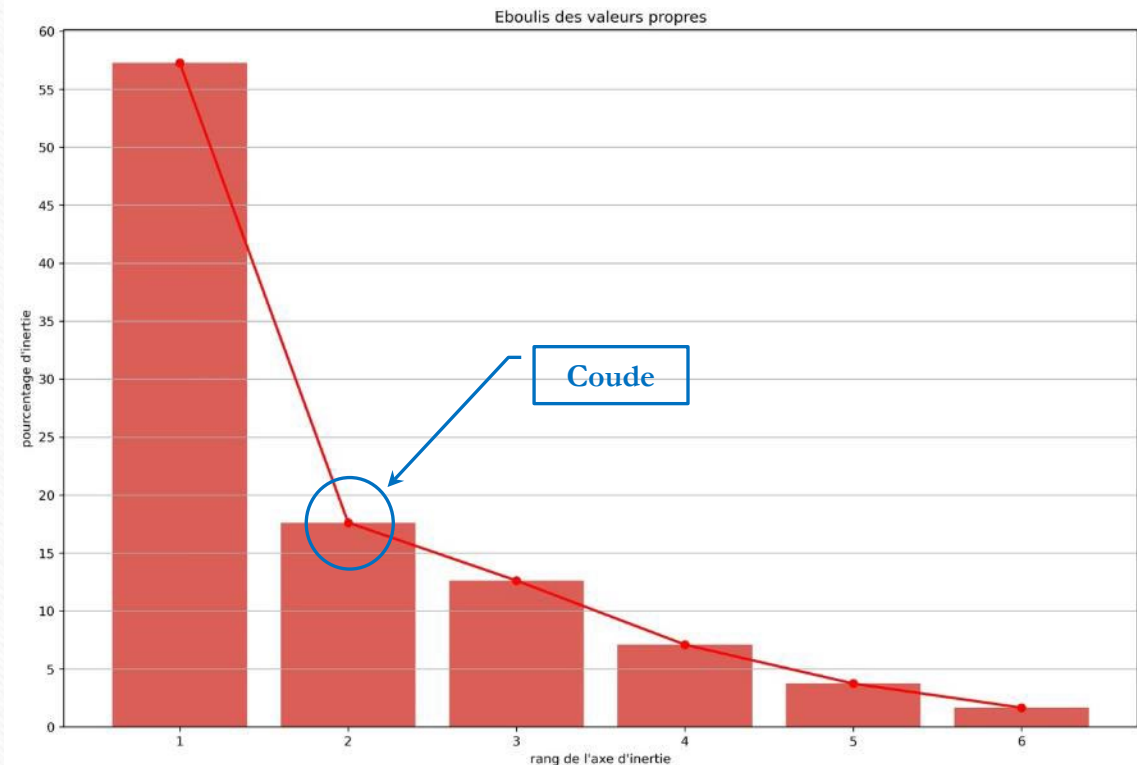
Zone	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Proportion de protéines animales (en %)	Diff pop_2017-2016 (en %)	Autosuffisance (en kg/hab)	PIB par habitant (en US\$/hab)	Cluster n°
Antigua-et-Barbuda	2447	80,43	64,29	0,94	-52,40	15383	3
Arabie saoudite	3300	88,57	38,64	1,99	-21,72	20802	3
Bahamas	2644	81,62	64,30	1,00	-34,05	32719	3
Chine - RAS de Hong-Kong	3270	135,91	73,29	0,86	-42,43	46705	3
Dominique	2973	82,08	57,88	0,21	-41,98	7275	3
Grenade	2364	72,03	57,34	0,55	-36,08	10153	3
Koweït	3444	100,42	48,11	2,45	-32,54	29755	3
Malte	3437	116,13	57,53	0,42	-18,27	29205	3
Nouvelle-Calédonie	2858	86,71	58,25	1,07	-32,47	34129	3
Polynésie française	2922	95,13	64,77	0,55	-43,46	20852	3
Saint-Kitts-et-Nevis	2744	85,11	65,02	0,81	-57,64	19155	3
Saint-Vincent-et-les Grenadines	2972	89,33	55,45	0,34	-72,84	7213	3
Sainte-Lucie	2638	86,04	60,59	0,51	-49,74	11047	3
Samoa	3040	89,20	60,12	0,42	-66,55	4220	3
Emirats arabes unis	3343	112,95	45,63	1,33	-39,32	40645	3

Problème : Un nombre encore élevé de pays

Solution : Ajouter une variable pour discriminer mes pays

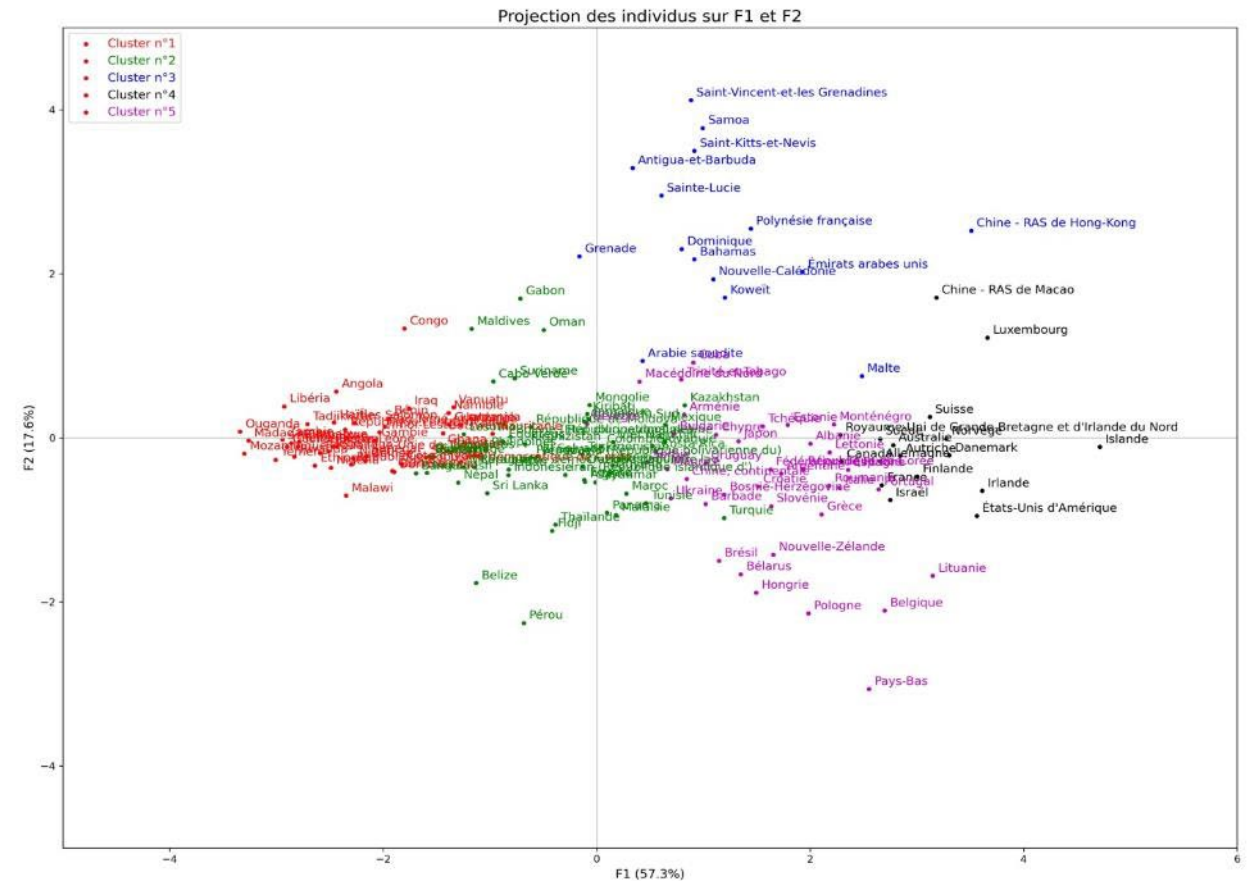
3. Présentation des résultats

- Choix du nombre de variable : méthode du coude
- La représentation des valeurs propres (nouvelles variables) montre que les variables 1 et 2 expliquent 75% de la dispersion du nuage d'individus. Ceci est donc très satisfaisant au regard de l'objectif de l'ACP.



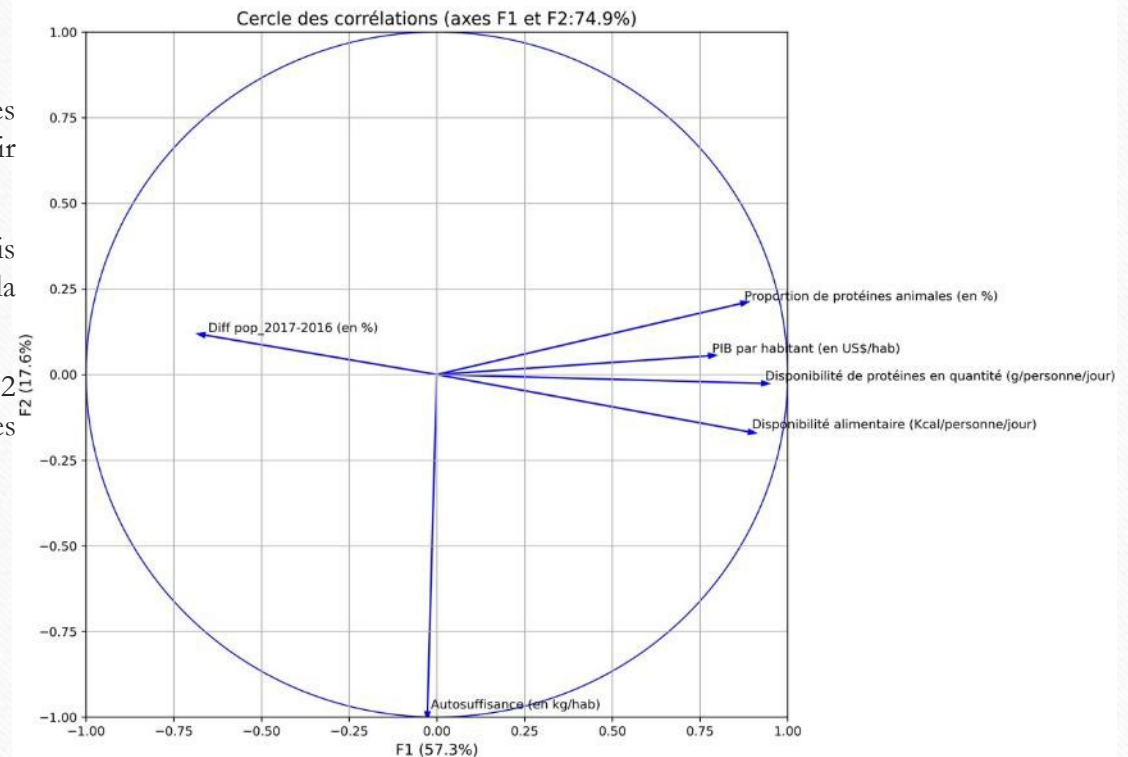
3. Présentation des résultats

- Cette projection traduit les ressemblances entre individus
- La plupart des pays s'étalent sur l'axe F1 : opposition pays du Sud/pays de Nord
- L'axe F2 traduit une opposition entre les pays insulaires et les pays continentaux
- Les pays en haut de l'axe F2 sont très dépendants des importations de poulet tandis que les pays en bas de l'axe sont très autonomes



3. Présentation des résultats

- Cette projection traduit les liaisons entre les variables
- La variable F1 synthétisent 5 variables initiales : elles sont très corrélées à F1. On peut traduire cette variable par le pouvoir d'achat des habitants de ces pays.
- La variable Diff_pop_2017_2018 est très corrélée à F1 mais négativement => Plus le pays est riche et plus la croissance de la population est faible
- L'axe variable Autosuffisance est très fortement corrélée à F2 négativement :=> Ceci traduit la dépendance ou l'autonomie des pays vis-à-vis des importations de poulet



3. Présentation des résultats

Liste finale de pays à cibler :

Zone	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Proportion de protéines animales (en %)	Diff pop_2017-2016. (en %)	Autosuffisance (en kg/hab)	PIB par habitant. (en US\$/hab)	ide (en % du PIB)
Antigua-et-Barbuda	2447	80,43	64,29	0,94	-52,40	15383	7,69
Bahamas	2644	81,62	64,30	1,00	-34,05	32719	7,22
Chine - RAS de Hong-Kong	3270	135,91	73,29	0,86	-42,43	46705	32,44
Grenade	2364	72,03	57,34	0,55	-36,08	10153	9,92
Malte	3437	116,13	57,53	0,42	-18,27	29205	26,64
Nouvelle-Calédonie	2858	86,71	58,25	1,07	-32,47	34129	5,23
Saint-Kitts-et-Nevis	2744	85,11	65,02	0,81	-57,64	19155	5,11
Saint-Vincent-et-les Grenadines	2972	89,33	55,45	0,34	-72,84	7213	12,39
Sainte-Lucie	2638	86,04	60,59	0,51	-49,74	11047	6,55
Emirats arabes unis	3343	112,95	45,63	1,33	-39,32	40645	2,69

- Afin d'affiner ma liste de pays à cibler, j'ai choisi d'ajouter une variable supplémentaire
- Cette variable, appelée « investissements directs étrangers (IDE) », est un bon indicateur de l'attractivité économique d'un pays
- Si on relève un taux élevé d'IDE, cela signifie que le pays séduit les investisseurs et qu'il est par conséquent attractif sur le plan économique
- J'ai donc cibler les 10 pays ayant l'IDE le plus élevé parmi la liste préliminaire

3. Présentation des résultats

- Tests d'adéquation de Kolmogorov Smirnov :
 - Afin de tester l'adéquation des variable du dataframe df_poulet à une loi normale, on réalise un test de Kolmogorv Smirnov.
 - On émet l'hypothèse H_0 selon laquelle la variable suit une loi normale.
 - On vérifie ensuite la statistique de test et la p-value.
 - On teste les 4 variables suivantes : « Disponibilité alimentaire (Kcal/personne/jour) », « Disponibilité de protéines en quantité (g/personne/jour) », « Proportion de protéines animales (en %) » et « Diff pop_2017-2016 (en %) »

3. Présentation des résultats

Résultats des tests :

	Statistic	P-value
Disponibilité alimentaire (Kcal/personne/jour)	0.0921578947368421	0.0921578947368421
Disponibilité de protéines en quantité (g/personne/jour)	0.06694152046783626	0.505489975442009
Proportion de protéines animales (en %)	0.09287719298245614	0.15000723285305284
Diff pop_2017-2016 (en %)	0.06637426900584795	0.5163468152656993

- Au vu de la statistique de test et de la p-value, on ne peut pas rejeter l'hypothèse H_0 de normalité au niveau de test 5%
- L'ensemble des 4 variables testées ci-dessus suivent une loi normale.

3. Présentation des résultats

- Test de comparaison de 2 populations :
 - On souhaite comparer deux échantillons gaussiens, pour cela il nous suffit de comparer leur moyenne μ_1 et μ_2 , et leur variance σ_1 et σ_2 .
 - On réalise d'abord un test d'égalité des variances puis un test d'égalité des moyennes.
 - Je choisis mes deux clusters 1 et 2 pour réaliser un test de comparaison sur ma variable gaussienne "Disponibilité de protéines en quantité (g/personne/jour)"

3. Présentation des résultats

- Test d'égalité des variances :
 - On réalise un test de Barlett pour comparer les variances des deux clusters,
 - On émet l'hypothèse H_0 selon laquelle les variances sont égales
 - On calcule les résultats du test : $\text{statistic}=1.965032083940408$ et $\text{pvalue}=0.1609761670678058$
 - On ne rejette pas l'hypothèse H_0 au niveau de test 5% puisque la p-value vaut 0.161

3. Présentation des résultats

- Test d'égalité des moyennes :
 - On réalise un test de Student pour comparer les moyennes des deux clusters.
 - On émet l'hypothèse H_0 selon laquelle les moyennes sont égales.
 - On calcule les résultats du test : $\text{statistic} = -7.266128716651409$ et $\text{pvalue} = 8.631672247860604 \times 10^{-11}$
 - On obtient une p-valeur égale à environ 8.63×10^{-11} . On constate donc que l'hypothèse d'égalité des moyennes de "Disponibilité de protéines en quantité" est rejetée à un niveau de test de 5%.

On rejette donc que la "Disponibilité de protéines en quantité" des clusters 1 et 2 suivent la même distribution car on a rejeté l'hypothèse d'égalité des moyennes

4. Recommandations

- S'orienter sur le cluster n°3 pour sélectionner nos pays
- Cibler dans un premier temps des pays proches géographiquement (Malte, Emirats arabes unis ?)
- Choisir des pays dont l'IDE est élevé afin de se projeter à moyen long terme sur une stratégie de production à l'étranger

5. Conclusion

- L'ACP et la CAH nous a permis de cibler un nombre restreint de pays malgré des variables nombreuses(6)
- Nous avons pu mettre en lumière les ressemblances entre pays et les liaisons entre variables afin d'en tirer des informations capitales avant l'étude de marché
- Nous avons gagner du temps et donc de l'argent en sélectionnant les pays les plus intéressant pour notre objectif commercial.