

ÉTUDE DE MARCHÉ



La poule qui chante

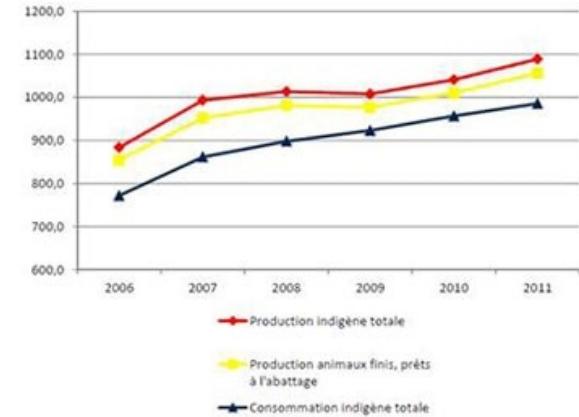
CONTEXTE

ENJEUX & OBJECTIFS

Bilan production / consommation de poulet en 2011 en milliers de tonnes

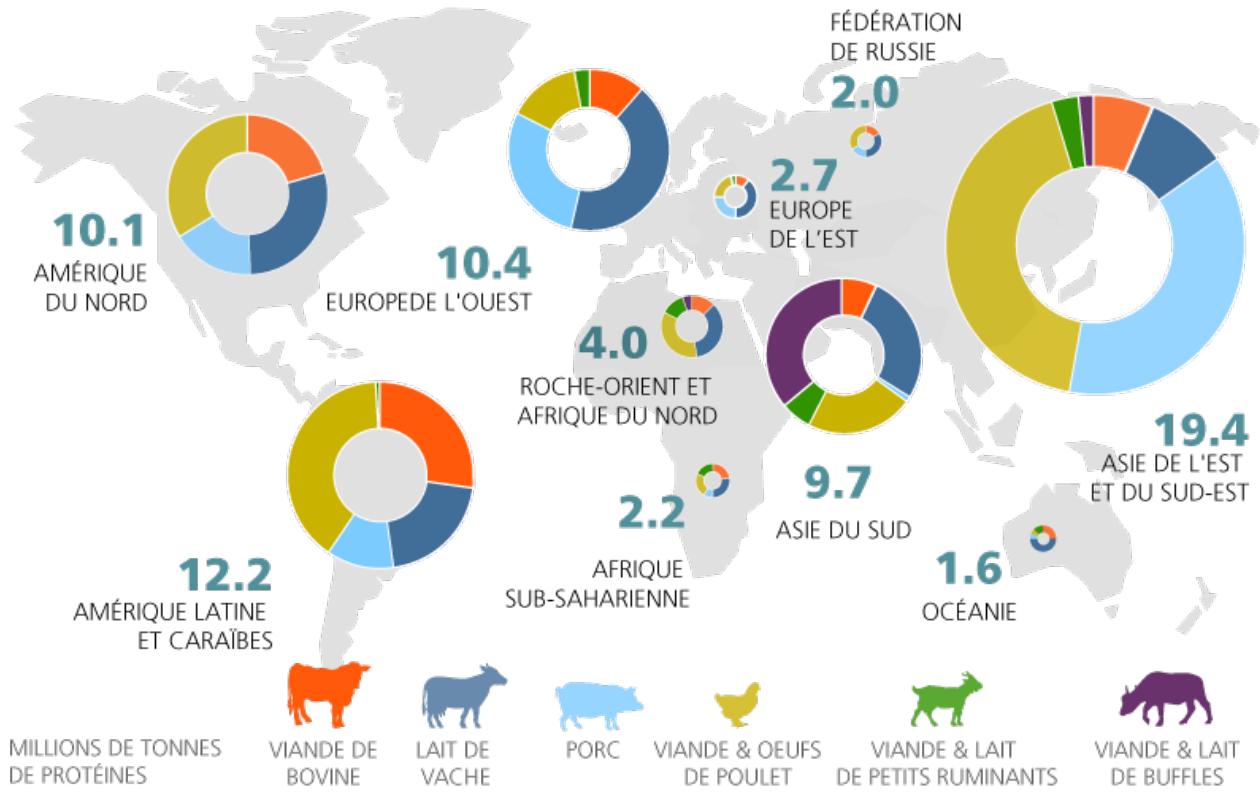
	2006	2007	2008	2009	2010	2011
Production indigène totale	883,6	993,3	1013,8	1008,0	1040,9	1089,1
Exportations animaux vivants adultes (-)	32,5	43,8	37,6	37,4	6,7	34,6
Importations animaux vivants adultes (+)	3,1	2,7	5,0	6,2	37,2	1,7
Production animaux finis, prêts à l'abattage	854,2	952,2	981,2	976,8	1010,4	1056,2
Exportations animaux morts (-)	327,9	376,8	403,9	403,6	437,1	486,8
Importations animaux morts (+)	234,6	293,8	319,9	345,7	381,5	419,9
Variation de stocks (-)	-10,8	7,6	-1,0	-4,2	-0,9	3,7
Consommation indigène totale	771,7	861,6	898,1	923,1	956,7	985,6

Source : CIPC d'après Agreste



- La société *La poule qui chante* a pour objectif d'exporter ses poulets français à l'international, afin de s'adresser à une clientèle plus vaste et sortir d'une zone déjà saturée par la concurrence (cf graphique ci-dessus, tenons également compte du fait que la tendance montrée par ledit graphique n'a pas connu de changement au cours des années suivantes) ;
- Notre but est de cibler un certain nombre de pays pouvant potentiellement être intéressants dans le cadre de cette internationalisation afin d'affiner par la suite l'étude de marché.

RÉPARTITION DE LA PRODUCTION DE PROTÉINES D'ORIGINE ANIMALE PAR RÉGION



La carte présente les niveaux de production régionaux et leurs profils. La production de viande exprimée en protéine a été calculée en utilisant des données sur les rendements d'abattage, la proportion de viande désossés et la teneur moyenne en protéine de la viande sans os. Par ailleurs, le lait de toutes les espèces a été converti en lait corrigé pour les matières grasses et les protéines. La production d'œufs est également exprimée en termes de protéines.

Le poulet et ses dérivés occupent une part de marché très importante.

Source : FAO

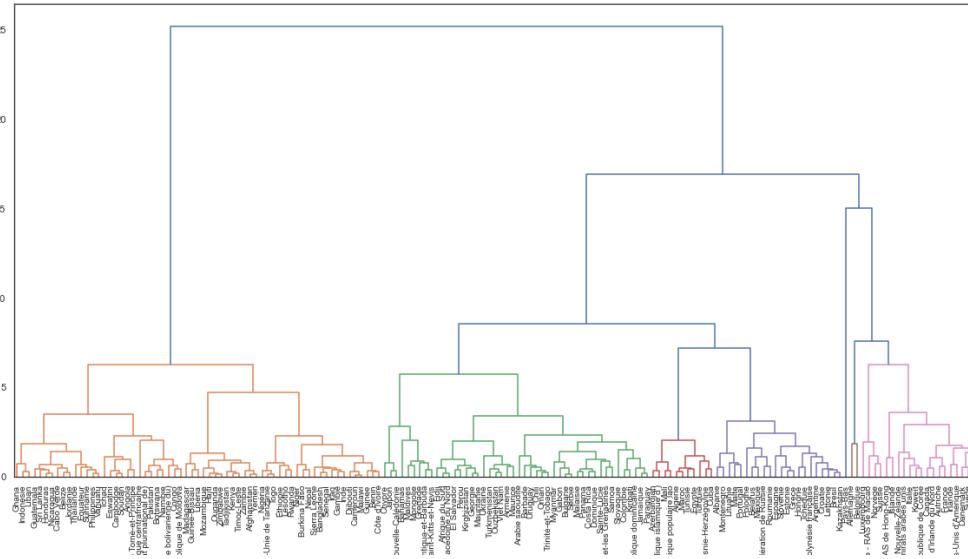
DÉMARCHE

CRÉATION D'UNE BASE DE DONNÉES SUR LAQUELLE TRAVAILLER

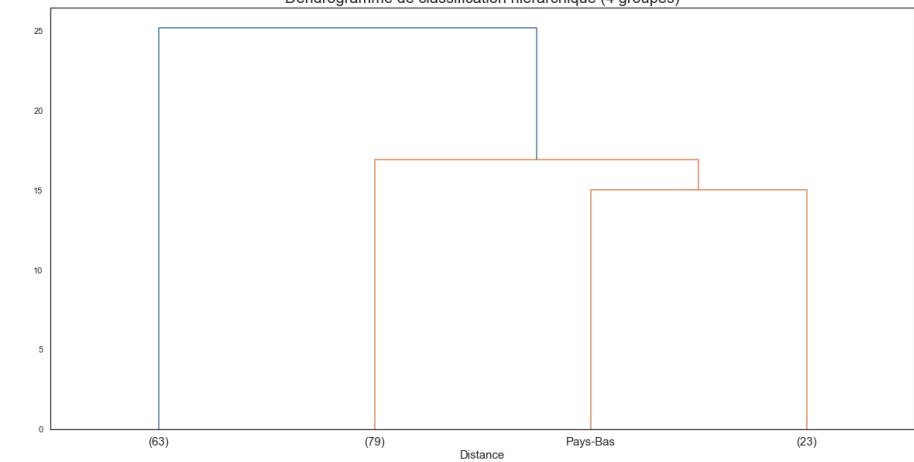
Nous avons choisi de recueillir un certain nombre de variables nous permettant d'effectuer efficacement notre analyse. Ces variables sont les suivantes :

- Disponibilité alimentaire, exprimée en calories par habitant ;
- Disponibilité en protéine, exprimée en grammes par habitant ;
- Proportion de protéines animales par rapport aux protéines totales ;
- Taux d'évolution de la population (moyenne calculée sur 5 ans) ;
- PIB par habitant, exprimé en dollars américains (moyenne calculée encore une fois sur 5 ans) ;
- Quantité de poulets importés, exprimée en dollars américains (moyenne calculée sur 5 ans).

MÉTHODE 1 : CLASSIFICATION ASCENDANTE HIÉRARCHIQUE



Dendrogramme de classification hiérarchique (4 groupes)



Nous avons commencé par effectuer une classification ascendante hiérarchique ou CAH. Pour ce faire, il nous a fallu réaliser un dendrogramme afin de visualiser le découpage (ici en 4 groupes).

VISUALISATION DES CARACTÉRISTIQUES DES 4 CLASSES

Groupe	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Proportion protéines animales (%)	Taux évolution moyen population (%)	PIB moyen/habitant (\$)	Quantité moyenne poulets importée (\$)
1	2442.396825	61.222222	27.586190	2.129048	2425.834127	3.424445e+06
2	3016.810127	88.763291	49.717975	0.823544	11018.273291	9.240657e+06
3	3451.695652	108.817391	59.272174	0.946957	53528.310000	2.926756e+07
4	3251.000000	104.430000	65.540000	0.260000	49023.290000	6.149210e+08

Nous avons affiché les centroïdes (valeurs moyennes) des classes obtenues afin d'avoir un aperçu de la façon dont les pays ont été répartis dans lesdites classes. La visualisation des centroïdes nous permet également d'avoir une idée des classes contenant les pays les plus susceptibles d'être sélectionnés pour y exporter.

MÉTHODE 2 : UTILISATION DE L'ALGORITHME DES K-MEANS

Après exclusion de la France, nous avons obtenu une liste de 23 pays (se trouvant dans les groupes n°3 et 4) pouvant potentiellement être ciblés par l'entreprise.

```
Entrée [21]: # Affichage des pays qui pourraient potentiellement être ciblés
print(data_gp[data_gp['Groupe'] == 3]['Zone'].unique())
print(data_gp[data_gp['Groupe'] == 4]['Zone'].unique())
['Allemagne' 'Autriche' 'Belgique' 'Canada' 'Chine - RAS de Hong-Kong'
 'Chine - RAS de Macao' 'Danemark' 'Finlande' 'France' 'Irlande' 'Islande'
 'Israël' 'Italie' 'Koweït' 'Luxembourg' 'Norvège' 'Nouvelle-Zélande'
 "Royaume-Uni de Grande-Bretagne et d'Irlande du Nord"
 'République de Corée' 'Suisse' 'Suède' 'Émirats arabes unis'
 'États-Unis d'Amérique"]
 ['Pays-Bas']
```

Afin d'affiner notre sélection nous avons utilisé l'algorithme des K-Means ou *algorithme des centres mobiles*. Cette méthode, très utilisée en clustering, possède de nombreux avantages dont les suivants :

- Simplicité et rapidité d'utilisation et de compréhension ;
- Flexibilité et efficacité

VISUALISATION DES CARACTÉRISTIQUES DES 5 CLUSTERS

	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Proportion protéines animales (%)	PIB moyen/habitant (\$)	Quantité moyenne poulets importée (\$)
0	1.297155	1.451013	1.073333	1.872180	0.274487
1	-0.914703	-0.983253	-0.940508	-0.596924	-0.185418
2	0.319381	0.320595	0.411465	-0.164120	-0.085660
3	0.861422	1.157241	1.425962	1.835476	11.154556

Encore une fois, l'affichage des centroïdes permet une meilleure lecture des caractéristiques des clusters. Ici par exemple, nous pouvons clairement voir que les clusters n°0 et n°3 se démarquent par des valeurs nettement supérieures à celles des autres classes. Nous obtenons cette fois une liste de 26 pays (toujours après exclusion de la France).

SÉLECTION DE NOTRE ÉCHANTILLON FINAL

En utilisant la méthode de la classification ascendante hiérarchique, nous avions obtenu une liste de 23 pays. En utilisant l'algorithme des K-Means, nous en avons obtenu une contenant 26 pays. Nous avons donc pris le parti de créer un nouvel échantillon ne contenant que les pays présents dans les listes que nous avions précédemment obtenues. Nous nous sommes ainsi retrouvés avec un nouvel échantillon de 22 pays.

```
Entrée [45]: echantillon.shape
```

```
Out[45]: (22, 15)
```

```
Entrée [46]: pays_commun = [echantillon['Zone'].unique()]
print("Liste des pays communs = ", pays_commun)
```

```
Liste des pays communs = [array(['Allemagne', 'Autriche', 'Belgique', 'Canada',
'Chine - RAS de Hong-Kong', 'Chine - RAS de Macao', 'Danemark',
'Finlande', 'Irlande', 'Islande', 'Israël', 'Italie', 'Koweït',
'Luxembourg', 'Norvège', 'Nouvelle-Zélande', 'Pays-Bas',
'Royaume-Uni de Grande-Bretagne et d'Irlande du Nord', 'Suisse',
'Suède', 'Émirats arabes unis', "États-Unis d'Amérique"]),
dtype=object)]
```

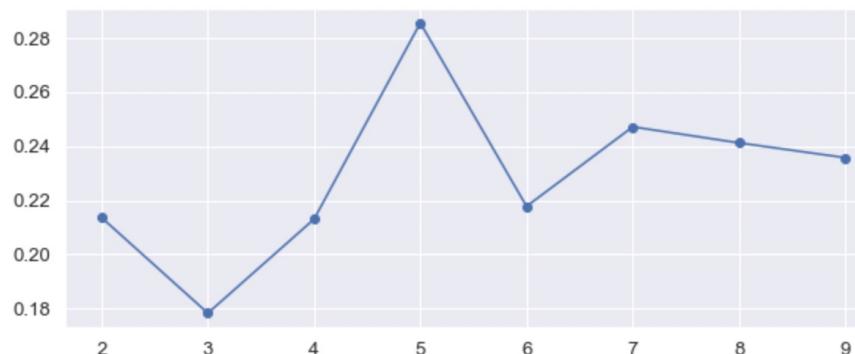
RÉSULTATS

UTILISATION DE L'ALGORITHME DES K-MEANS SUR LE NOUVEL ÉCHANTILLON

```
Entrée [53]: # Établissement du nombre de clusters optimal à l'aide de la métrique "silhouettes"
```

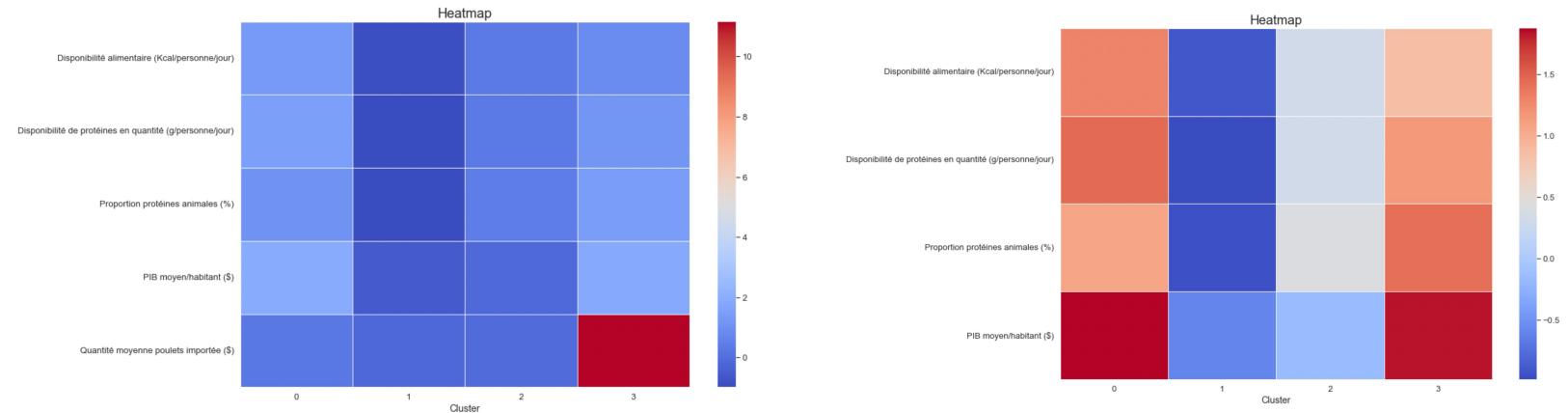
```
silhouettes = []
for num_clusters in range(2, 10):
    cls = KMeans(n_clusters=num_clusters, random_state=1512)
    cls.fit(X3_scaled)
    silh = metrics.silhouette_score(X3_scaled, cls.labels_)
    silhouettes.append(silh)

plt.plot(range(2, 10), silhouettes, marker='o')
plt.show()
```



Nous avons une nouvelle fois utilisé l'algorithme des K-Means mais cette fois sur notre échantillon de 22 pays afin d'affiner davantage la sélection. La métrique Silhouette nous expose un bon équilibre pour un découpage en 4 ou 5 clusters. Nous avons décidé de partir sur un découpage en 4 clusters.

VISUALISATION DES CARACTÉRISTIQUES DES 4 CLUSTERS AINSI QUE DES PAYS LES COMPOSANT



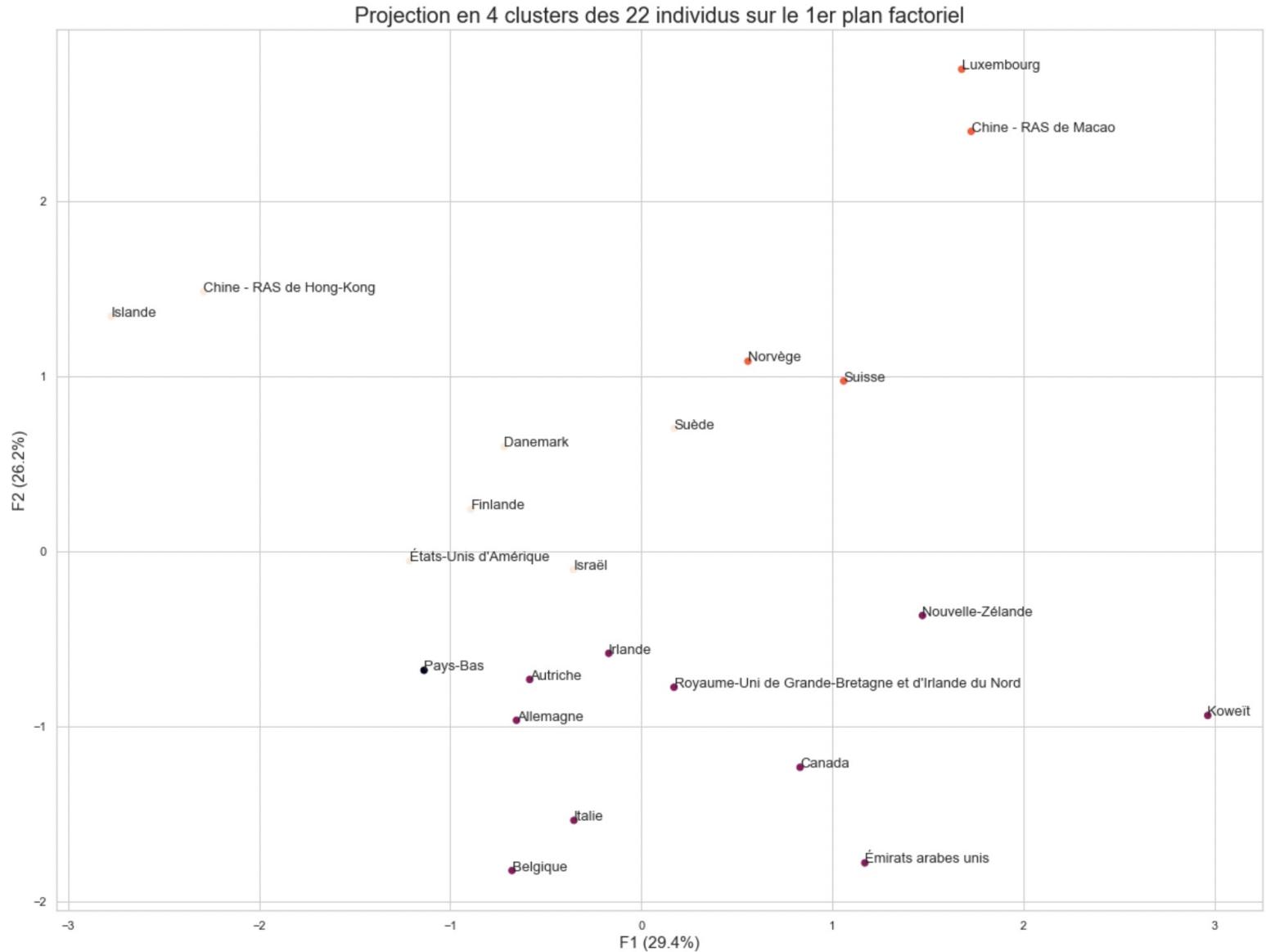
Nous avons réalisé des heatmaps afin de représenter visuellement les caractéristiques des différents clusters. En fonction de la politique de l'entreprise, il sera maintenant plus facile de prendre des décisions concernant les pays dans lesquels elle souhaite exporter ses volailles.

```
Entrée [65]: # Il semblerait que le cluster 2 ne soit composé que de très peu de pays par rapport aux autres clusters
# Vérification de la composition de chacun des clusters

print("Le cluster numéro 0 est composé du/des " + str(len(data_cls4b[data_cls4b['Cluster'] == 0]['Zone'])) + " pays suivant(s) :")
print("Le cluster numéro 1 est composé du/des " + str(len(data_cls4b[data_cls4b['Cluster'] == 1]['Zone'])) + " pays suivant(s) :")
print("Le cluster numéro 2 est composé du/des " + str(len(data_cls4b[data_cls4b['Cluster'] == 2]['Zone'])) + " pays suivant(s) :")
print("Le cluster numéro 3 est composé du/des " + str(len(data_cls4b[data_cls4b['Cluster'] == 3]['Zone'])) + " pays suivant(s) :")

Le cluster numéro 0 est composé du/des 1 pays suivant(s) :
['Pays-Bas']
Le cluster numéro 1 est composé du/des 10 pays suivant(s) :
['Allemagne' 'Autriche' 'Belgique' 'Canada' 'Irlande' 'Italie' 'Koweït'
 'Nouvelle-Zélande' "Royaume-Uni de Grande-Bretagne et d'Irlande du Nord"
 'Émirats arabes unis']
Le cluster numéro 2 est composé du/des 4 pays suivant(s) :
['Chine - RAS de Macao' 'Luxembourg' 'Norvège' 'Suisse']
Le cluster numéro 3 est composé du/des 7 pays suivant(s) :
['Chine - RAS de Hong-Kong' 'Danemark' 'Finlande' 'Islande' 'Israël'
 'Suède' 'États-Unis d'Amérique']
```

PROJECTION DES 22 PAYS EN 4 CLUSTERS SUR LE 1ER PLAN FACTORIEL



RECOMMANDATIONS

CHOIX DES PAYS À CIBLER

Comme indiqué précédemment, en fonction de la politique de la société, le choix des pays peut être effectué de diverses manières :

- D'après nos heatmaps, il serait pertinent de s'intéresser en priorité aux pays des clusters 0 et 3, à savoir la Chine (plus particulièrement la région administrative spéciale de Hong-Kong), le Danemark, la Finlande, l'Islande, l'Israël, les Pays-Bas (qui se démarquent des autres pays par la quantité de volailles importées), la Suède et les États-Unis d'Amérique ;
- Les pays pourraient aussi être favorisés de diverses autres manières (états membres de l'UE, proximité géographique, pays ayant des accords commerciaux avec la France, ...) ;
- Nous aurions également pu ajouter d'autres variables afin d'en tenir compte lors de l'analyse telles que la stabilité politique ou encore les quantités de poulets élevés.