

# Loco Pollo - nouveaux marchés

Projet 5 - Parcours Data Analyst - Hieu TRUONG Janvier 2020

- ***Objectifs***
- Périmètre de l'étude
- Détails de l'analyse
- Validation des resultats
- Conclusions

# Objectifs

Dans le cadre du développement à l'international pour l'exportation des poulets Loco Pollo, notre étude a pour objectifs:

- **Definir une strategie d'expansion**
- **Short-lister les pays vers lesquels cibler notre expansion**



- Objectifs
- ***Périmètre de l'étude***
- Détails de l'analyse
- Validation des resultats
- Conclusions

# L'étude

- Source: FAO
- Périmètre: 170+ pays
- Periode: 2013
- Types de données:
  - Croissance démographique (sur 5 ans)
  - Pib par habitant
  - Ratio proteines animales / proteines totales
  - Disponibilité alimentaire (kcal et proteines)
  - Ratio disponibilité alimentaire (kcal et proteines) en volailles
  - Balance imports exports en volailles

## Une etude portant sur:

- 170 pays
- 8 variables

**Viser une sélection d'une quinzaine de pays sur la base de variables composites**

- Objectifs
- Périmètre de l'étude
- ***Détails de l'analyse***
- Validation des resultats
- Conclusions

# Consolidation des données (1/2)

- Nettoyage des données
  - Chine (en doublon)
  - France (hors etude)
  - Corée du Nord: absence de données sur les imports exports
- Conservation de pays en décroissance démographique (présence du Japon ou de la Roumanie) mais qui restent des cibles potentielles
- Coherences des données
  - Verification des ratios (compris entre 0 et 1)



# Consolidation des données (2/2)

- Imputations
  - Exportations non définies remise à zéro (imputations par mediane)
  - Pib hab Taiwan (par moyenne des pib des territoires chinois)
- élimination des pays pour lesquels on n'a pas d'info sur les disponibilités animales
- Outliers identifiés par graphe des boxplots
  - Conservation des pays riches
  - Balance imports exports inférieurs à -1000
  - Etats-Unis et Brésil retirés car exports massifs de volailles

**Au final un jeu de données de 170 individus et 8 variables**

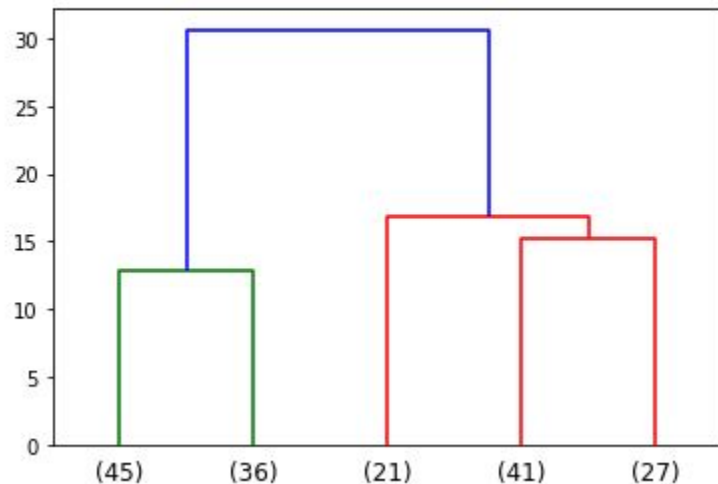


# Classification Ascendante Hierarchique

Regroupons les pays en fonction des variables données

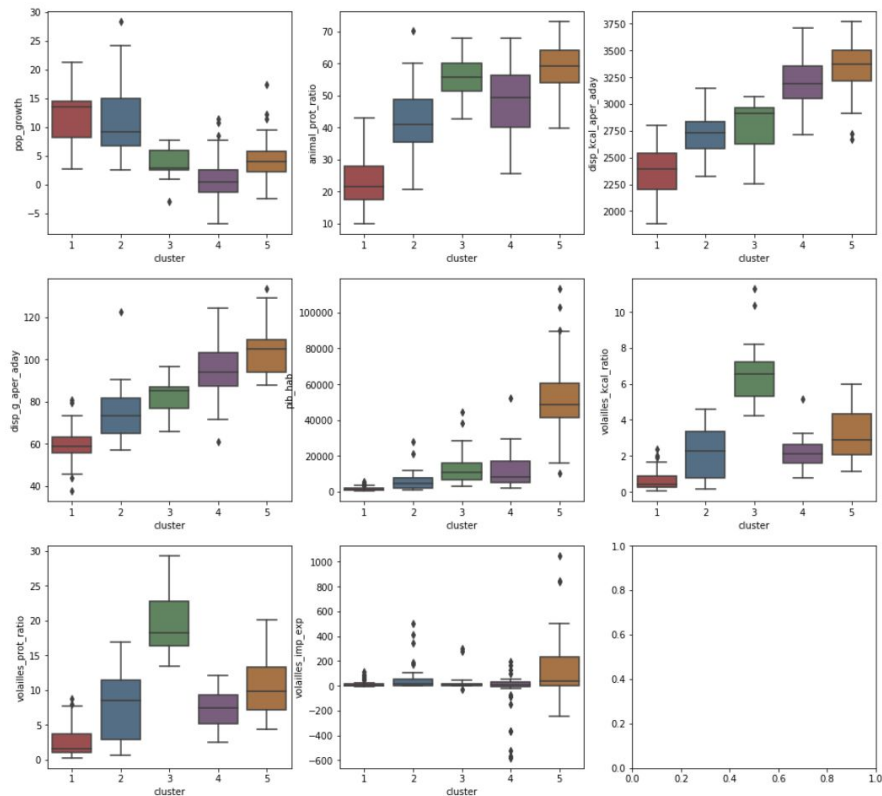
- Centrage et réduction des données
- Function linkage / ward

```
.]: 1 dendrogram(Z, truncate_mode='lastp', p=5)  
   2 plt.show()
```



**Cinq groupes équilibrés en taille**

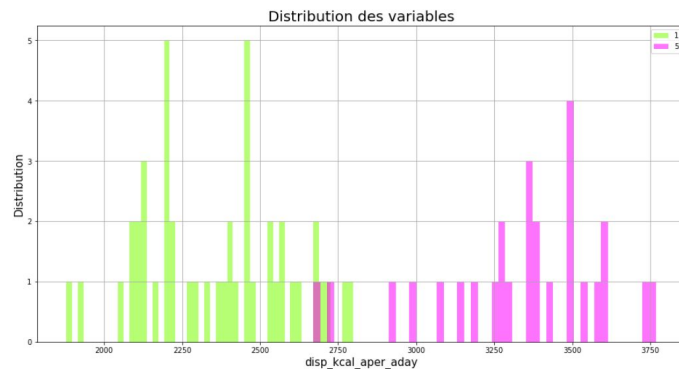
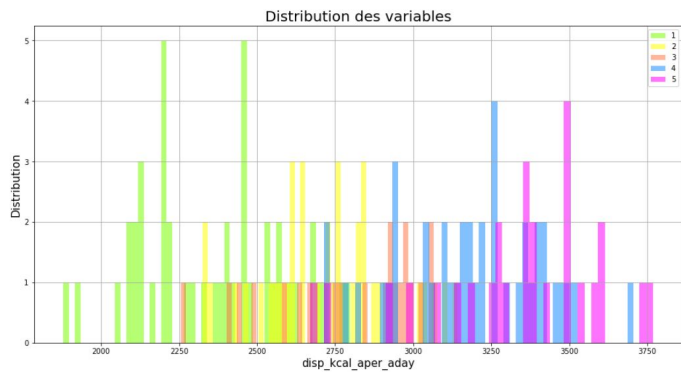
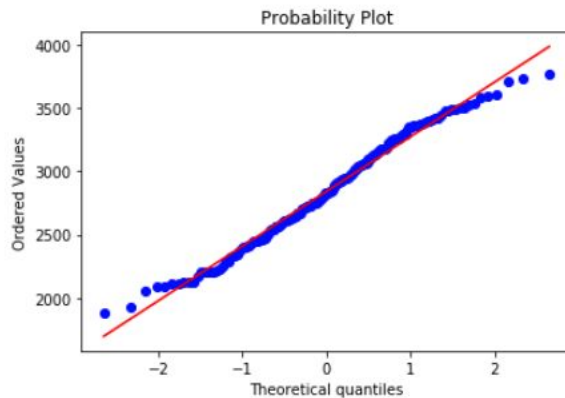
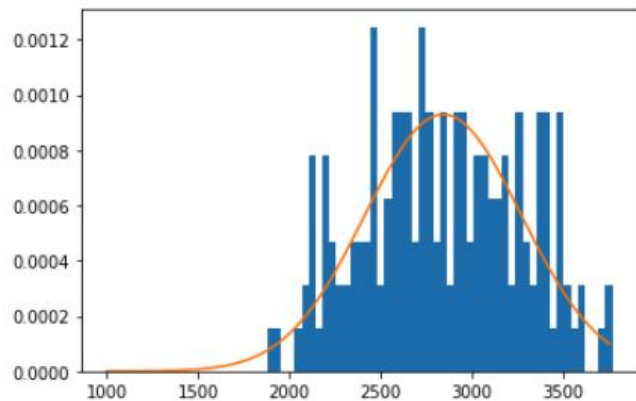
# Caractérisation des clusters



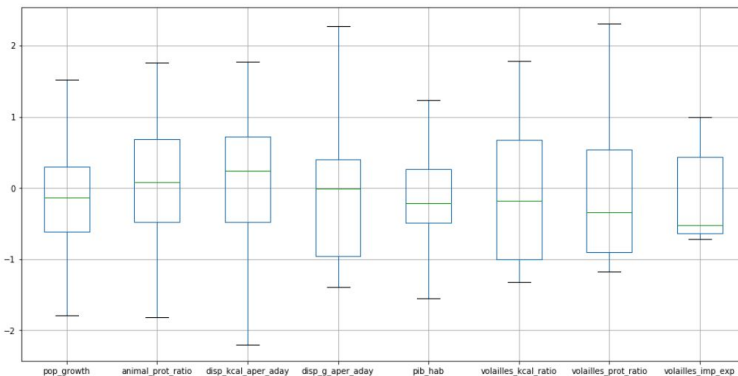
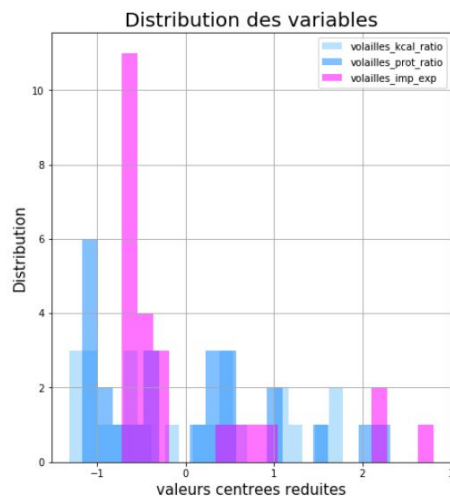
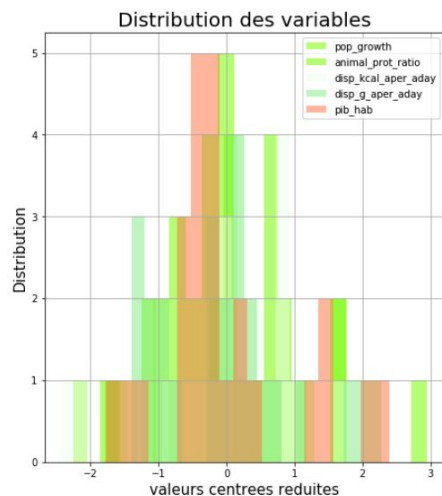
**Le cluster 5 nous intéresse pour les qualités suivantes (27 pays):**

- Pays riches
- Fortes importations
- Forte appétence pour la volaille

# Test d'indépendance

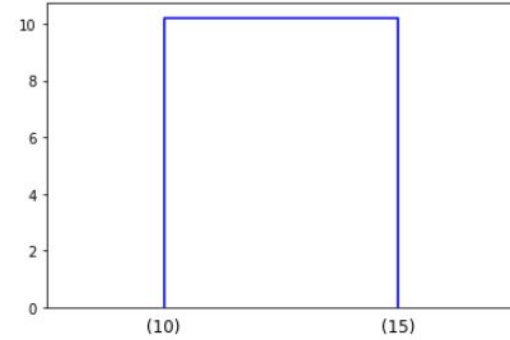
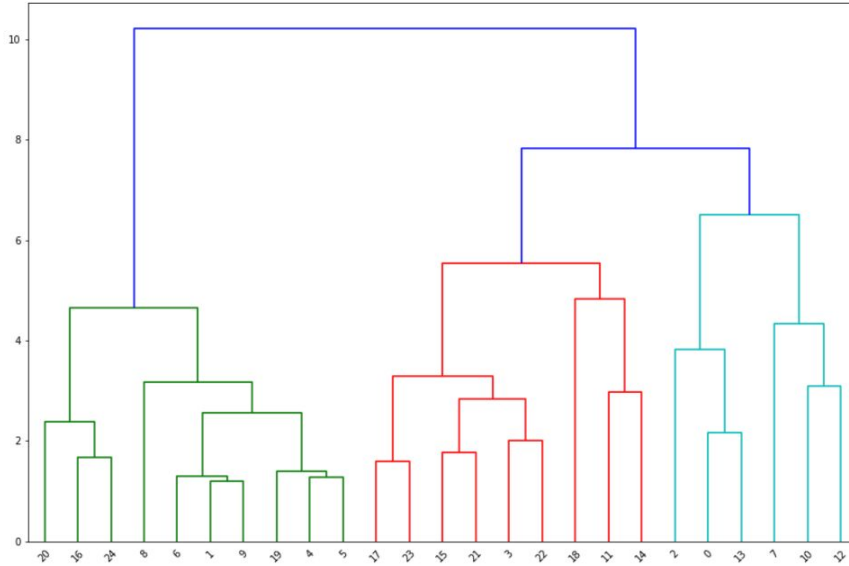


# Caracterisation du cluster 5



Le cluster 5 peut encore faire l'objet d'une classification

# Seconde classification



Deux groupes a valider

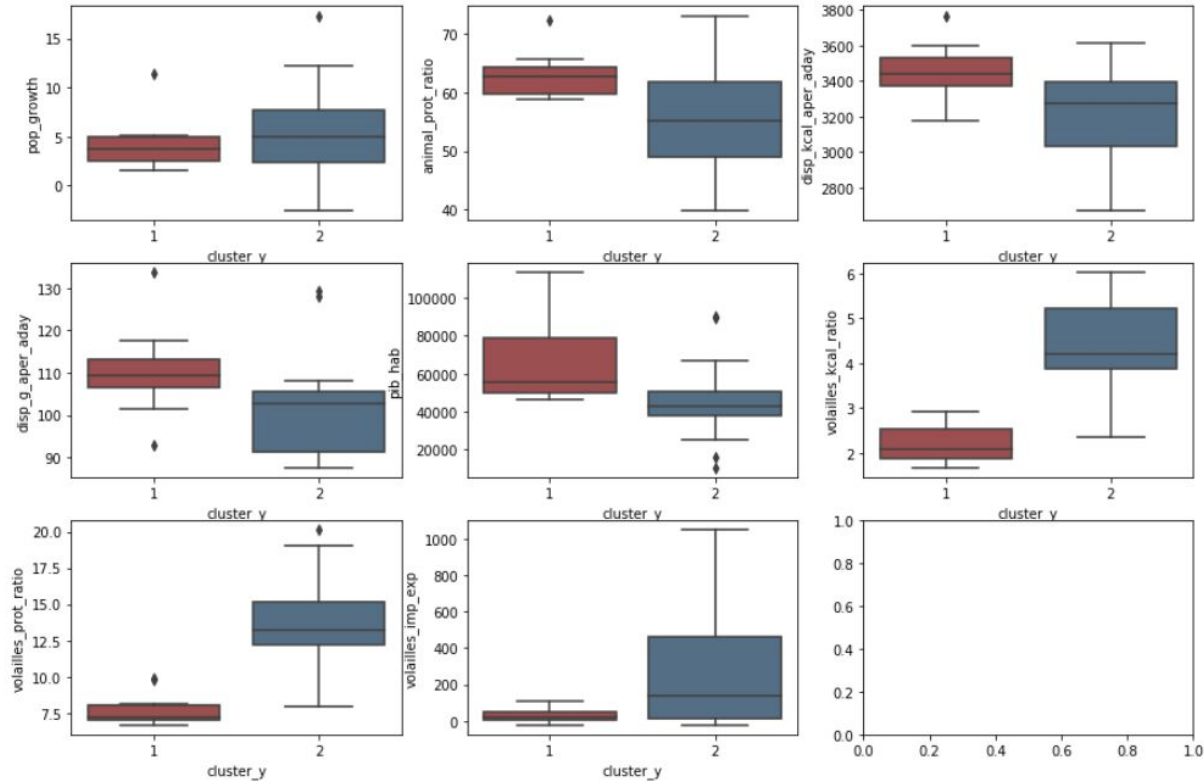
# Caracterisation du cluster 5

country	pop_growth	animal_prot_ratio	disp_kcal_aper_aday	disp_g_aper_aday	pi_b_hab	volailles_kcal_ratio	volailles_prot_ratio	volailles_imp_exp	clust
Islande	2.910000	72.250000	3380.000000	133.540000	49027.815665	2.720000	8.150000	1.000000	
Norvège	5.090000	59.500000	3485.000000	110.900000	103110.441896	2.070000	7.720000	2.000000	
Suisse	5.140000	64.210000	3391.000000	93.080000	84659.015399	1.680000	7.160000	52.000000	
Finlande	1.540000	62.040000	3368.000000	117.720000	49659.588178	2.110000	6.960000	3.000000	
Allemagne	2.400000	60.520000	3499.000000	101.590000	46176.177749	1.860000	7.150000	112.000000	
Autriche	3.920000	59.180000	3768.000000	106.210000	50137.519579	1.750000	7.220000	38.000000	
Luxembourg	11.370000	63.330000	3539.000000	113.880000	113341.237463	2.120000	6.720000	9.000000	
Danemark	2.020000	64.130000	3367.000000	108.880000	60942.805627	2.910000	9.770000	-21.000000	
Irlande	4.400000	58.920000	3600.000000	110.020000	51023.486060	2.670000	9.860000	27.000000	
Suède	3.680000	65.750000	3179.000000	107.720000	60254.372213	1.950000	6.670000	57.000000	
Chine - RAS de Macao	9.400000	64.420000	2915.000000	94.840000	89525.344880	4.970000	12.700000	17.000000	
Koweït	17.320000	49.500000	3501.000000	108.240000	48401.609794	5.110000	16.870000	136.000000	
Royaume-Uni	3.320000	56.470000	3424.000000	103.210000	42597.737474	3.270000	12.380000	429.000000	
Mexique	6.200000	46.400000	3072.000000	87.630000	10400.569944	3.710000	9.790000	839.000000	
Japon	-0.870000	55.240000	2726.000000	87.730000	40180.809978	2.350000	7.920000	1047.000000	
Bermudes	-2.530000	66.390000	2671.000000	90.790000	89968.618136	4.040000	13.330000	4.000000	
Chine, Taiwan Province de	1.300000	48.460000	2997.000000	89.740000	45013.000000	4.140000	11.910000	109.000000	
Israël	8.360000	56.560000	3610.000000	128.140000	37452.126321	6.010000	20.120000	-8.000000	
Arabie saoudite	12.150000	43.580000	3255.000000	91.630000	24934.386142	5.310000	19.020000	845.000000	
Fédération de Russie	0.980000	54.630000	3361.000000	102.840000	15997.000015	2.740000	9.010000	500.000000	
Nouvelle-Zélande	4.970000	59.230000	3137.000000	92.540000	42232.040825	4.460000	13.260000	-12.000000	
Émirats arabes unis	4.710000	39.860000	3280.000000	104.620000	43315.138419	4.210000	13.240000	330.000000	
Chine - RAS de Hong-Kong	3.990000	73.080000	3290.000000	129.180000	38566.712121	5.990000	15.160000	380.000000	
Canada	5.040000	52.120000	3494.000000	104.950000	52264.959967	4.150000	12.460000	36.000000	
Australie	7.070000	67.440000	3276.000000	106.280000	66544.670956	5.680000	15.110000	-21.000000	

On retire 2 pays pour lesquels la balance import export est tres negative

La seconde clusterisation se fait bien sur les 3 variables relatives aux volailles

# Seconde classification



# Selection finale via la balance import export

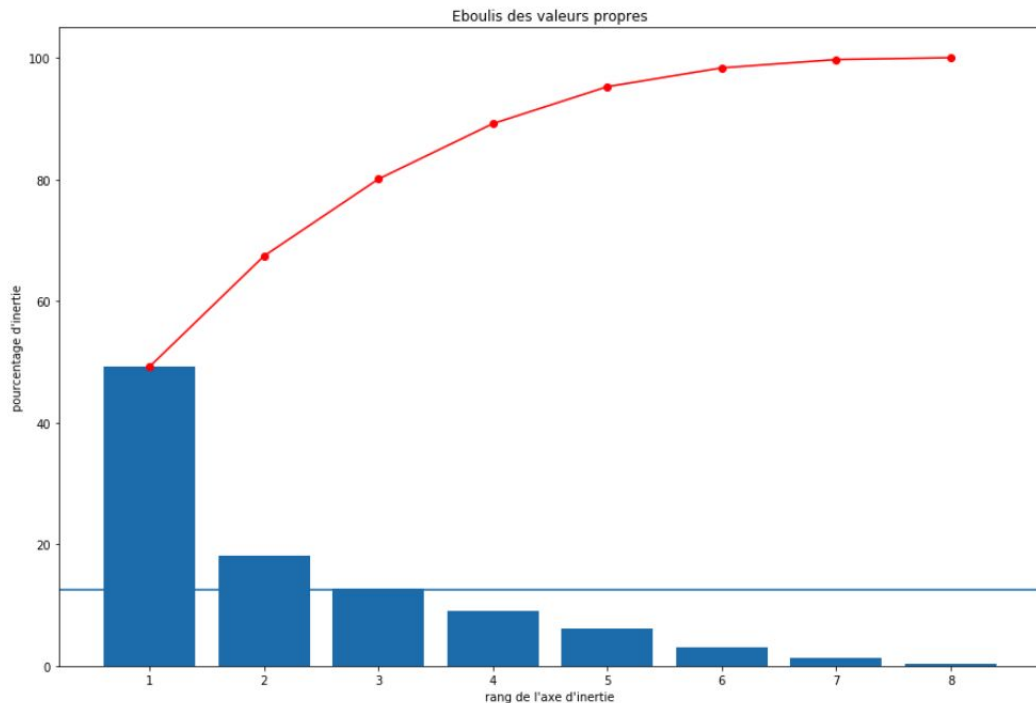
	country	pop_growth	animal_prot_ratio	disp_kcal_aper_aday	disp_g_aper_aday	pib_hab	volailles_kcal_ratio	volailles_prot_ratio	volailles_imp_exp
0	Australie	7.070000	67.440000	3276.000000	106.280000	66544.670956	5.680000	15.110000	-21.000000
2	Bermudes	-2.530000	66.390000	2671.000000	90.790000	89968.618136	4.040000	13.330000	4.000000
3	Canada	5.040000	52.120000	3494.000000	104.950000	52264.959967	4.150000	12.460000	36.000000
7	Chine - RAS de Hong-Kong	3.990000	73.080000	3290.000000	129.180000	38566.712121	5.990000	15.160000	380.000000
10	Israël	8.360000	56.560000	3610.000000	128.140000	37452.126321	6.010000	20.120000	-8.000000
11	Japon	-0.870000	55.240000	2726.000000	87.730000	40180.809978	2.350000	7.920000	1047.000000
12	Koweït	17.320000	49.500000	3501.000000	108.240000	48401.609794	5.110000	16.870000	136.000000
13	Chine - RAS de Macao	9.400000	64.420000	2915.000000	94.840000	89525.344880	4.970000	12.700000	17.000000
14	Mexique	6.200000	46.400000	3072.000000	87.630000	10400.569944	3.710000	9.790000	839.000000
15	Nouvelle-Zélande	4.970000	59.230000	3137.000000	92.540000	42232.040825	4.460000	13.260000	-12.000000
17	Fédération de Russie	0.980000	54.630000	3361.000000	102.840000	15997.000015	2.740000	9.010000	500.000000
18	Arabie saoudite	12.150000	43.580000	3255.000000	91.630000	24934.386142	5.310000	19.020000	845.000000
21	Chine, Taiwan Province de	1.300000	48.460000	2997.000000	89.740000	45013.000000	4.140000	11.910000	109.000000
22	Émirats arabes unis	4.710000	39.860000	3280.000000	104.620000	43315.138419	4.210000	13.240000	330.000000
23	Royaume-Uni	3.320000	56.470000	3424.000000	103.210000	42597.737474	3.270000	12.380000	429.000000

	country
3	Canada
7	Chine - RAS de Hong-Kong
11	Japon
12	Koweït
14	Mexique
17	Fédération de Russie
18	Arabie saoudite
21	Chine, Taiwan Province de
22	Émirats arabes unis
23	Royaume-Uni



- Objectifs
- Périmètre de l'étude
- Détails de l'analyse
- ***Validation des resultats***
- Conclusions

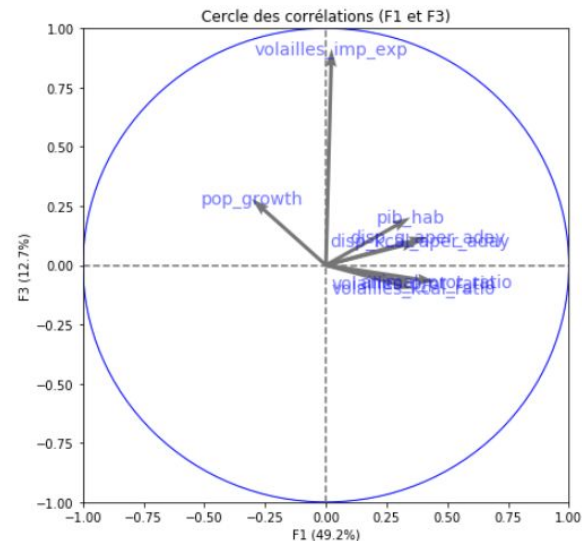
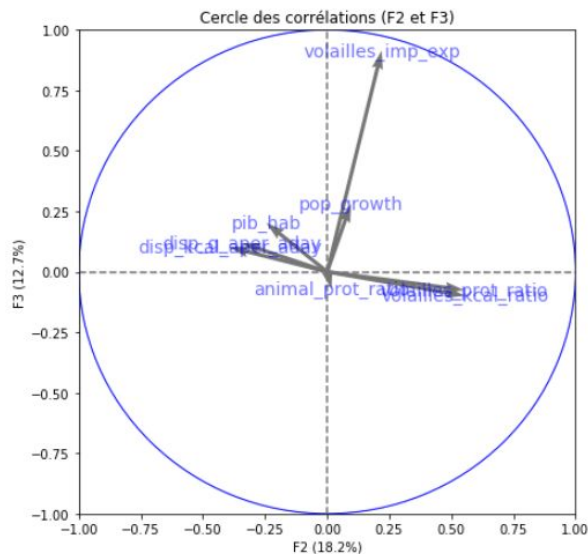
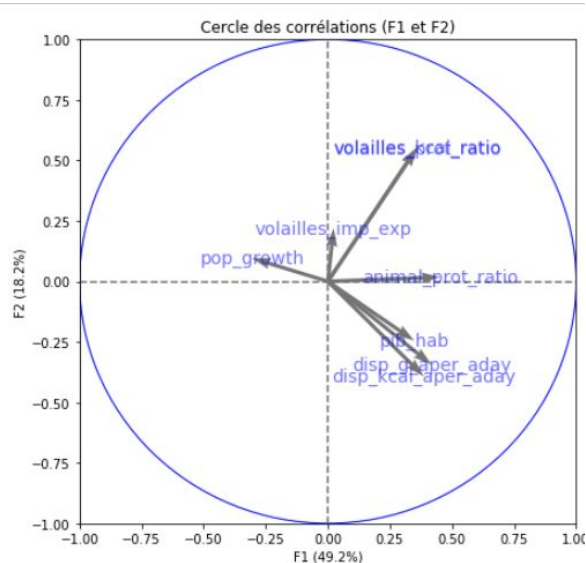
# ACP et projections



**Une ACP pour visualiser et caractériser les données et résultats obtenus**

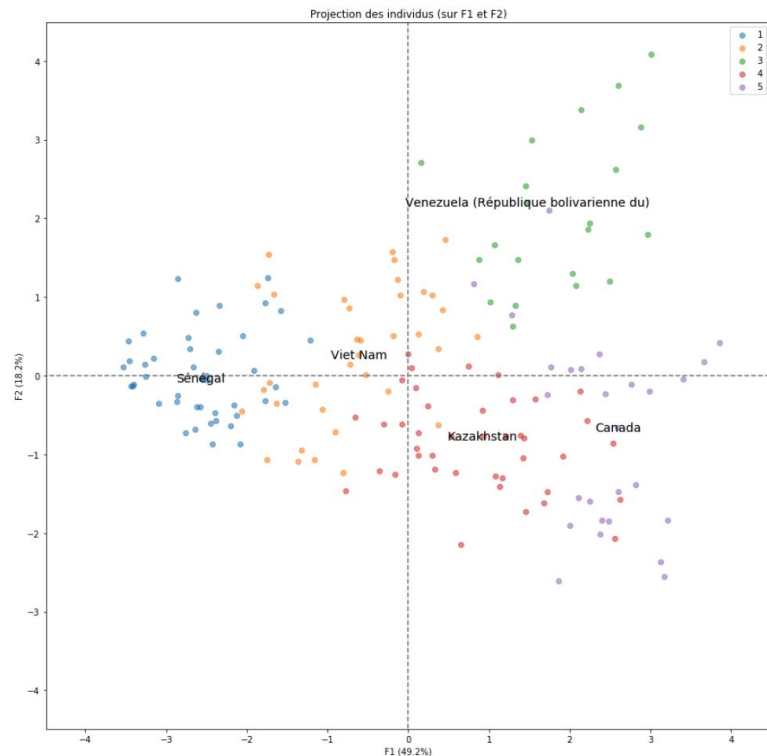
A priori les deux premiers plans suffisent le 4<sup>ème</sup> axe étant à moins de 12.5 pourcent, soit  $100 / p$

# ACP et projections



Notons que les axes 2 et 3 ont des poids similaires

# Projections sur le premier plan



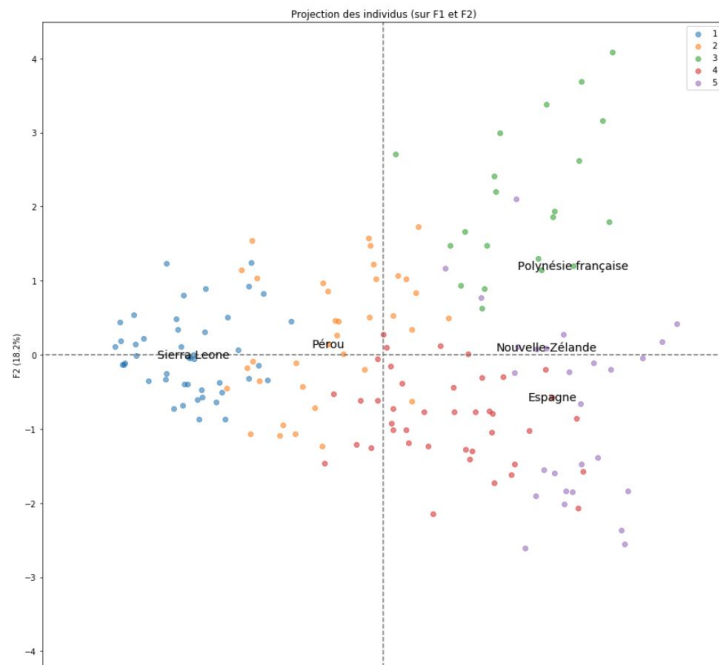
	country	Qual_F1	Qual_F2	cluster
132	Sénégal	0.943841	0.000168	1
160	Viet Nam	0.047857	0.008921	2
159	Venezuela (République bolivarienne du)	0.220267	0.489757	3
73	Kazakhstan	0.252576	0.175975	4
21	Canada	0.794085	0.052093	5

Les points les plus proches des  
centroids ne sont pas les mieux  
représentés sur l'axe

# Axe 1: Richesse, démographie et disponibilité

	country	Qual_F1	Qual_F2	cluster
3	Sierra Leone	0.982205	7.569817e-07	1
2	Pérou	0.864920	3.550458e-02	2
0	Polynésie française	0.764998	1.770302e-01	3
4	Espagne	0.862620	5.809796e-02	4
1	Nouvelle-Zélande	0.866134	1.679348e-03	5

	country	pop_growth	animal_prot_ratio	disp_kcal_aper_aday	disp_g_aper_aday	pi_b_hab	volailles_kcal_ratio	volailles_prot_ratio	volailles_imp_exp
133	Sierra Leone	11.45	26.49	2404.0	56.13	710.818700	0.50	2.74	14.0
117	Pérou	7.44	36.09	2700.0	74.92	6583.118622	2.00	7.78	28.0
45	Polynésie française	2.72	65.89	2927.0	96.85	22045.687291	5.88	17.41	16.0
137	Espagne	-0.51	62.12	3174.0	104.88	29163.288417	3.09	11.16	30.0
108	Nouvelle-Zélande	4.97	59.23	3137.0	92.54	42232.040825	4.46	13.26	-12.0

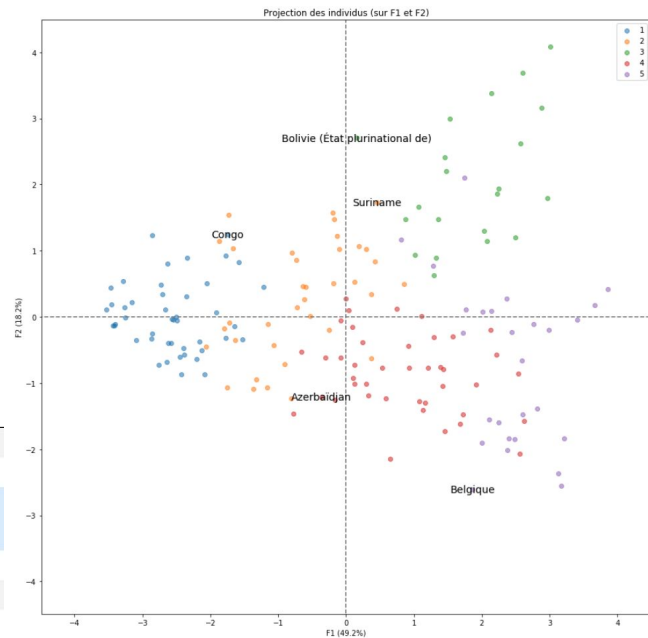


Selon l'axe 1 on recherche les pays ayant des valeurs positives

# Axe 2: appétence pour les produits volailles

	country	Qual_F1	Qual_F2	cluster
1	Congo	0.518454	0.263800	1
3	Suriname	0.051573	0.754847	2
0	Bolivie (État plurinational de)	0.002774	0.878081	3
2	Azerbaïdjan	0.056587	0.630000	4
4	Belgique	0.282885	0.554466	5

country	pop_growth	animal_prot_ratio	disp_kcal_aper_aday	disp_g_aper_aday	pib_hab	volailles_kcal_ratio	volailles_prot_ratio	volailles_imp_exp
Congo	13.45	43.07	2208.0	51.66	3747.359134	1.99	8.73	51.0
Suriname	5.24	45.09	2753.0	62.94	9484.379870	4.61	16.83	21.0
Bolivie (État plurinational de)	7.69	45.08	2256.0	66.01	2947.938522	6.56	16.53	-1.0
Azerbaïdjan	6.01	32.29	3118.0	93.15	7901.783462	1.12	3.82	5.0
Belgique	2.94	58.36	3733.0	99.59	46713.438406	1.15	4.38	-246.0

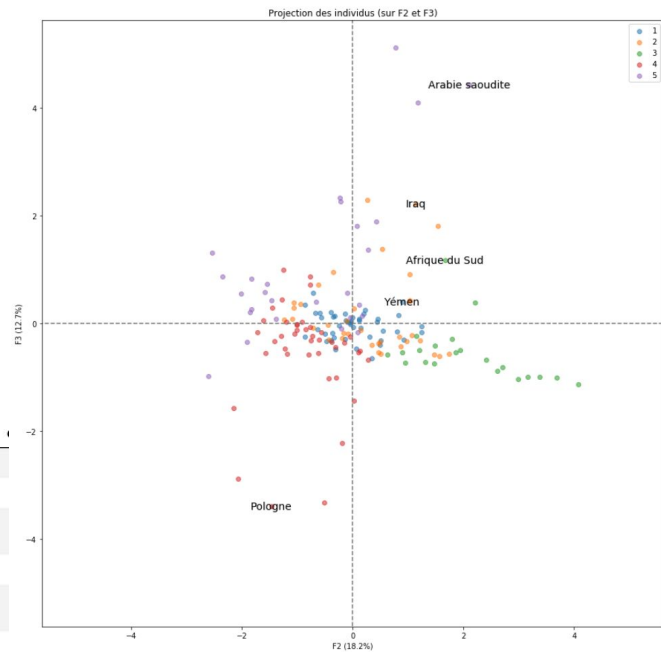


Selon l'axe 2 on recherche les pays ayant des valeurs positives

# Axe 3: balance des imports exports en volailles

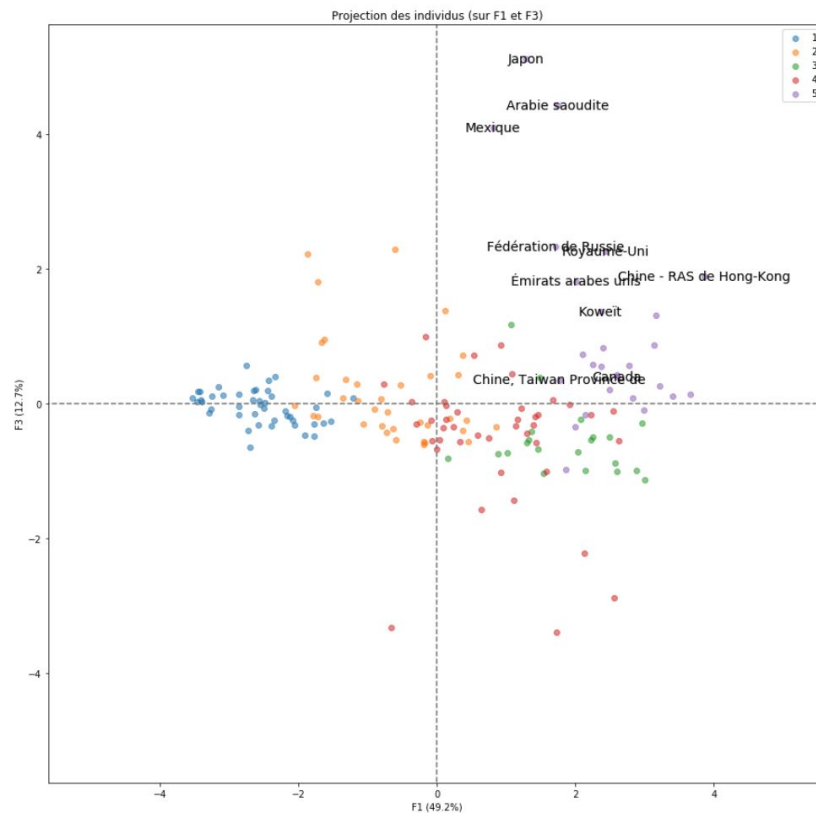
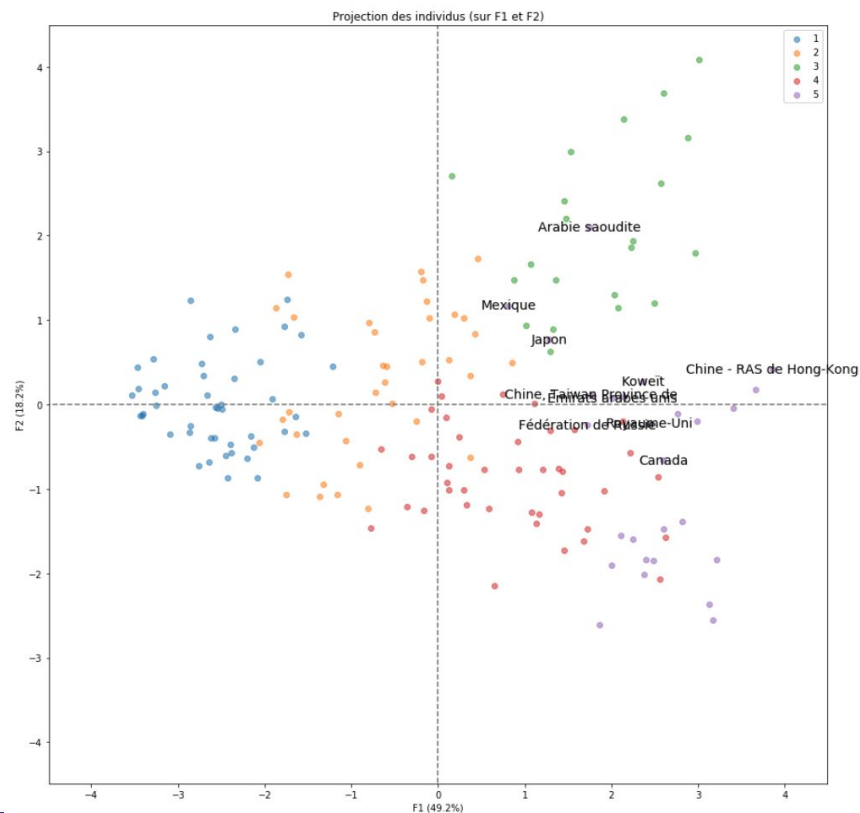
	country	Qual_F1	Qual_F2	Qual_F3	cluster
4	Yémen	0.824595	0.120137	0.023694	1
0	Iraq	0.349434	0.130101	0.490364	2
3	Afrique du Sud	0.184552	0.446770	0.219320	3
1	Pologne	0.169155	0.124058	0.655904	4
2	Arabie saoudite	0.108388	0.158471	0.697985	5

	country	pop_growth	animal_prot_ratio	disp_kcal_aper_aday	disp_g_aper_aday	pib_hab	volailles_kcal_ratio	volailles_prot_ratio	volailles_imp_exp
163	Yémen	13.33	23.55	2223.0	58.94	1358.885271	1.66	6.48	109.0
68	Iraq	15.91	20.71	2545.0	63.63	6112.878863	2.04	8.30	413.0
136	Afrique du Sud	7.65	42.63	3022.0	85.33	6819.093360	5.00	17.40	302.0
119	Pologne	-0.62	52.50	3451.0	101.47	13684.131318	2.75	10.93	-581.0
131	Arabie saoudite	12.15	43.58	3255.0	91.63	24934.386142	5.31	19.02	845.0



Selon l'axe 3 on recherche les pays ayant des valeurs positives

# Une sélection en ligne avec notre étude







Conclusions / autres

# Strategie

- Pays développés avec un fort PIB (mais pas les plus élevés)
- Pays ayant de fortes balances imports exports en produits de volailles
- Pays ayant une appétence moyenne pour les produits en volailles

	country
3	Canada
7	Chine - RAS de Hong-Kong
11	Japon
12	Koweït
14	Mexique
17	Fédération de Russie
18	Arabie saoudite
21	Chine, Taiwan Province de
22	Émirats arabes unis
23	Royaume-Uni