

Projet 5 : Produisez une étude de marché

Année 2018-2019



СОВЕРШЕНСТВОВАНИЕ

Sommaire



Mise en contexte



Problématique :

Identifier les pays propices à une insertion dans le marché du poulet ?

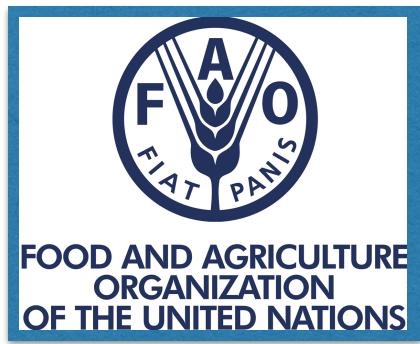


Notre entreprise



Son secteur : agroalimentaire

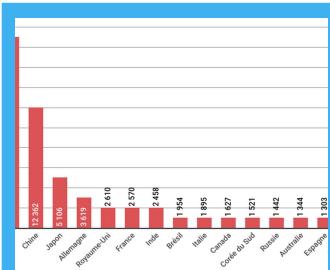
Les jeux de données



Population



Bilan alimentaire



PIB



L'évolution de la population

« Comment
évolue les futurs
consommateurs
? »



Population

```
population=population.fillna(0)
```

1)On remplace les valeurs manquantes

Code Domaine	Domaine	Code zone	Zone	Code Élément	Élément	Code Produit	Produit	Code année	Année	Unité	Valeur	Symbol	Description du Symbole
OA	Séries temporelles annuelles	2	Afghanistan	511	Population totale	3010	Population-Estimations	2013	2013	1000 personnes	31731.688	X	Source internationale sûre
OA	Séries temporelles annuelles	2	Afghanistan	511	Population totale	3010	Population-Estimations	2014	2014	1000 personnes	32758.020	X	Source internationale sûre
OA	Séries temporelles annuelles	202	Afrique du Sud	511	Population totale	3010	Population-Estimations	2013	2013	1000 personnes	53767.396	X	Source internationale sûre

2)Notre jeu de données

	Zone	Année	Valeur	Code Élément	Code zone
0	Afghanistan	2013	31731688.0	511	2
1	Afghanistan	2014	32758020.0	511	2
2	Afrique du Sud	2013	53767396.0	511	202
3	Afrique du Sud	2014	54539571.0	511	202
4	Albanie	2013	2918978.0	511	3
5	Albanie	2014	2920775.0	511	3
6	Algérie	2013	38338562.0	511	4

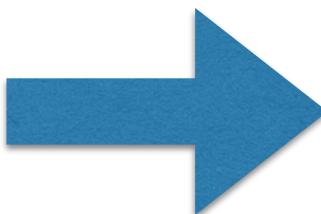
3)On découpe nos données

Population

```
population=population.pivot_table(values="Valeur", index=["Zone"],  
columns="Année",aggfunc='sum')
```

1)On applique un pivot

Année	Zone	2013	2014
0	Afghanistan	31731688.0	32758020.0
1	Afrique du Sud	53767396.0	54539571.0
2	Albanie	2918978.0	2920775.0
3	Algérie	38338562.0	39113313.0
4	Allemagne	81265139.0	81489660.0



	Pays	2013	2014	Evolution_pop
0	Afghanistan	31731688.0	32758020.0	3.234407
1	Afrique du Sud	53767396.0	54539571.0	1.436140
2	Albanie	2918978.0	2920775.0	0.061563
3	Algérie	38338562.0	39113313.0	2.020814
4	Allemagne	81265139.0	81489660.0	0.276282

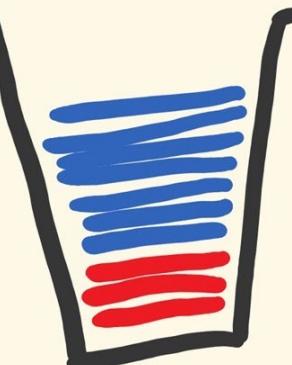
2)Puis la formule :
 $((2014-2013):2013)*100$

Ratio animal

« Si un pays est consommateur de viande ? »

RATIOS

3:7



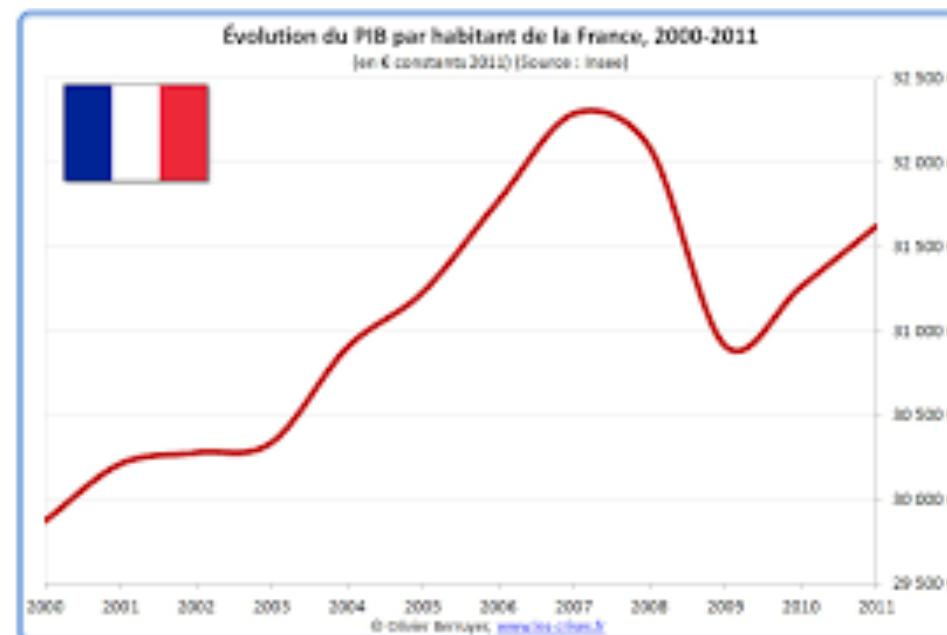
Disponibilité alimentaire en protéine et en calorie

« Déterminer le niveau de consommation d'un habitant par pays. »



PIB

« Indicateur de santé de l'économie d'un pays »



Jointure

230 lignes

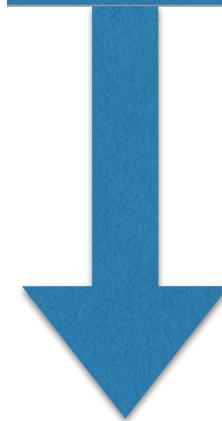
	Pays	2013	2014	Evolution_pop
0	Afghanistan	31731688.0	32758020.0	3.234407
1	Afrique du Sud	53767396.0	54539571.0	1.436140
2	Albanie	2918978.0	2920775.0	0.061563
3	Algérie	38338562.0	39113313.0	2.020814
4	Allemagne	81265139.0	81489660.0	0.276282

174 lignes

	DA_pro	DA_cal
	58.25	2089.0
	85.32	3022.0
	111.42	3193.0
	91.96	3296.0

188 lignes

	Pays	Valeur_PIB
0	Afghanistan	1848.7
1	Afrique du Sud	12339.7



167 lignes

	Pays	Evolution_pop	Ratio_Animal	DA_pro	DA_cal	Valeur_PIB
0	Afghanistan	3.234407	0.209785	58.25	2089.0	1848.7
1	Afrique du Sud	1.436140	0.426395	85.32	3022.0	12339.7
2	Albanie	0.061563	0.533297	111.42	3193.0	10492.8
3	Algérie	2.020814	0.271749	91.96	3296.0	13253.6

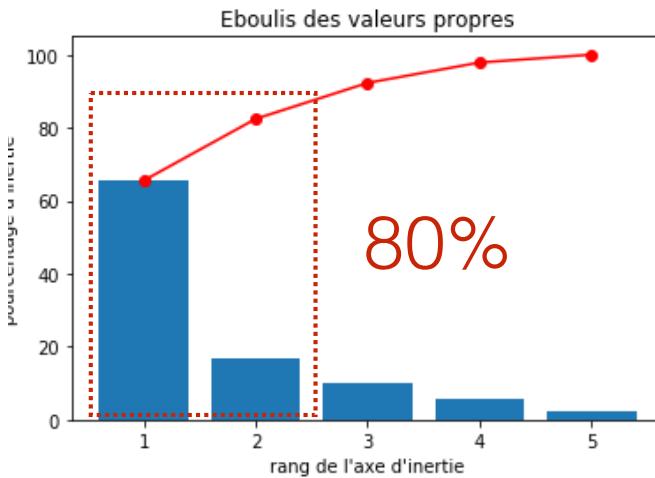


ACP

Analyse en composante principale

Evolution_pop

Valeur PIB



Eboulis des valeurs propres

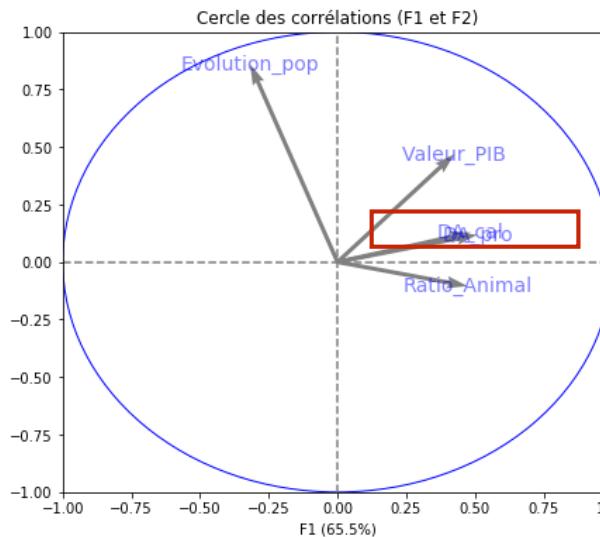
-Déterminer les axes d'inerties

ACP
Analyse en composante principale

DA_pro

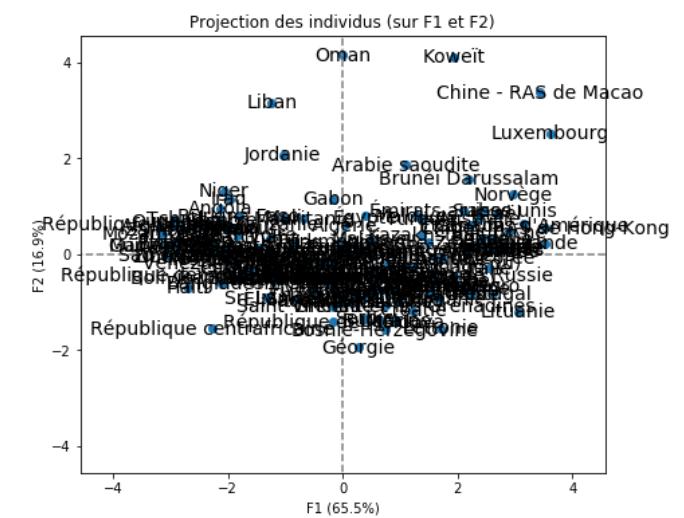
DA_cal

Ratio Animal



Cercle des corrélations

-Liaison entre les variables

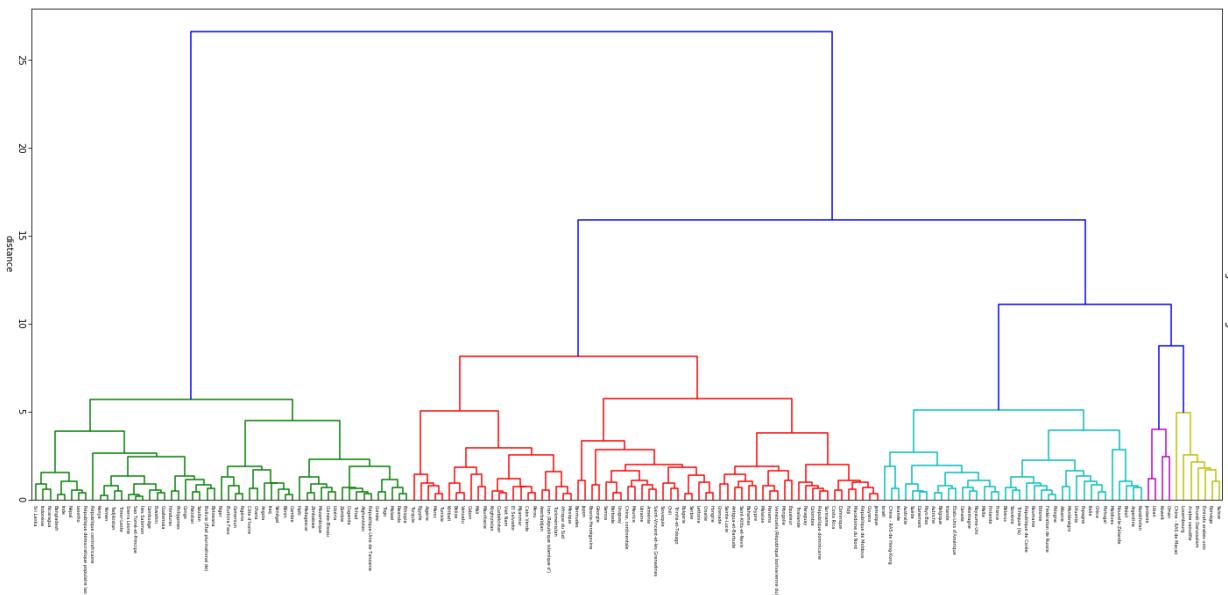


Projection des individus

-Etudier la variabilité

Clustering

01 5 clusters



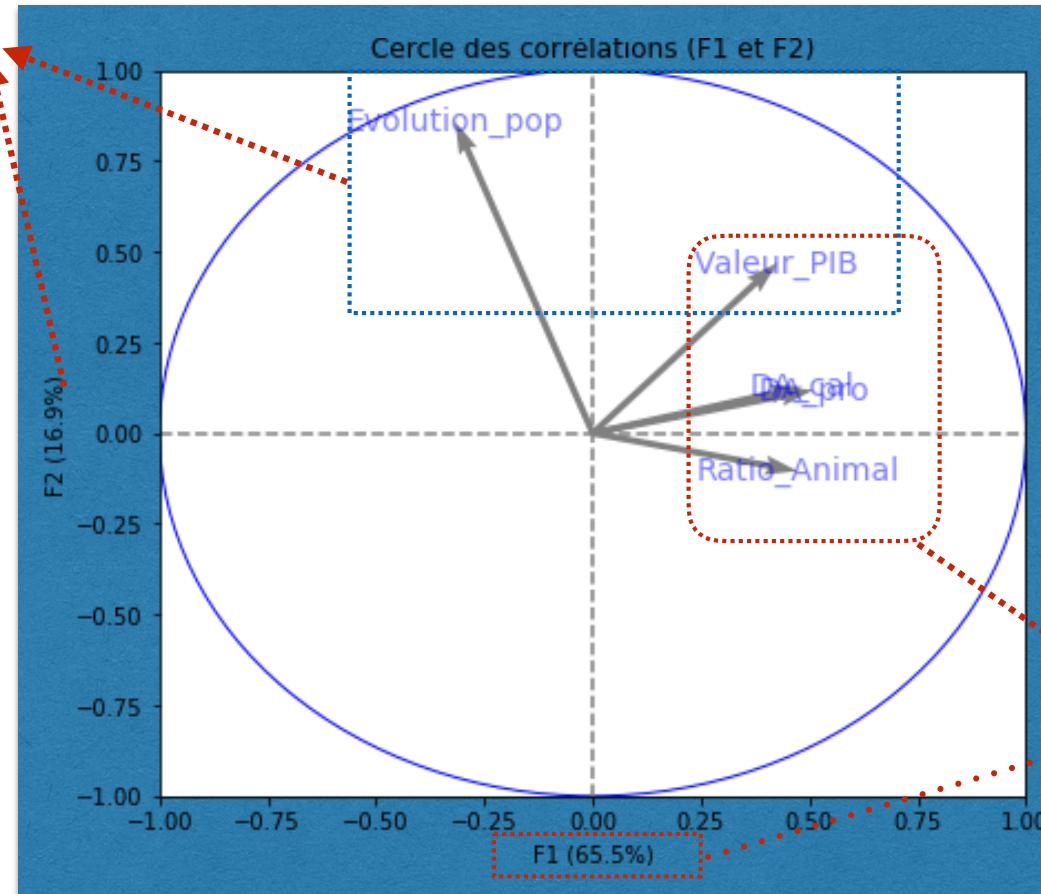
Hierarchical clustering
dendrogram

-Divise successivement

-Les points sont des pays

Analyse des axes

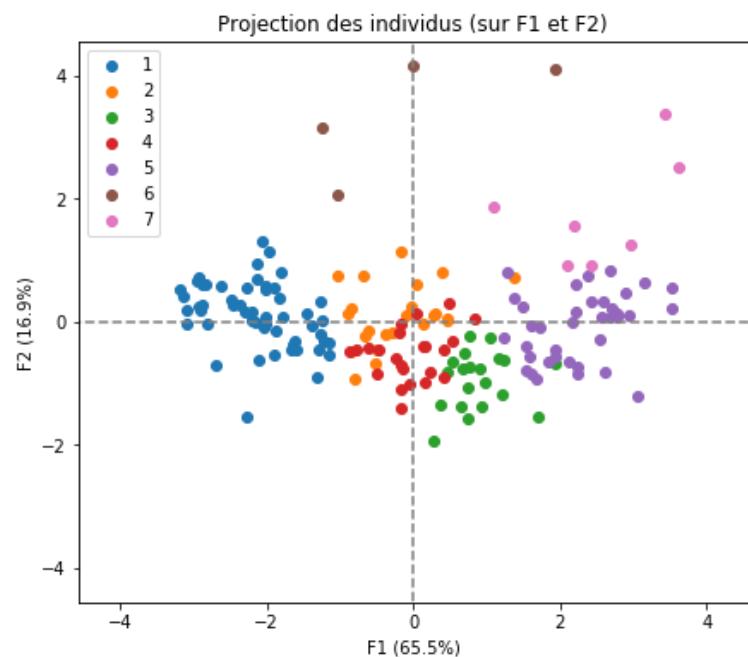
L' évolution d'un pays,
la santé d'un pays



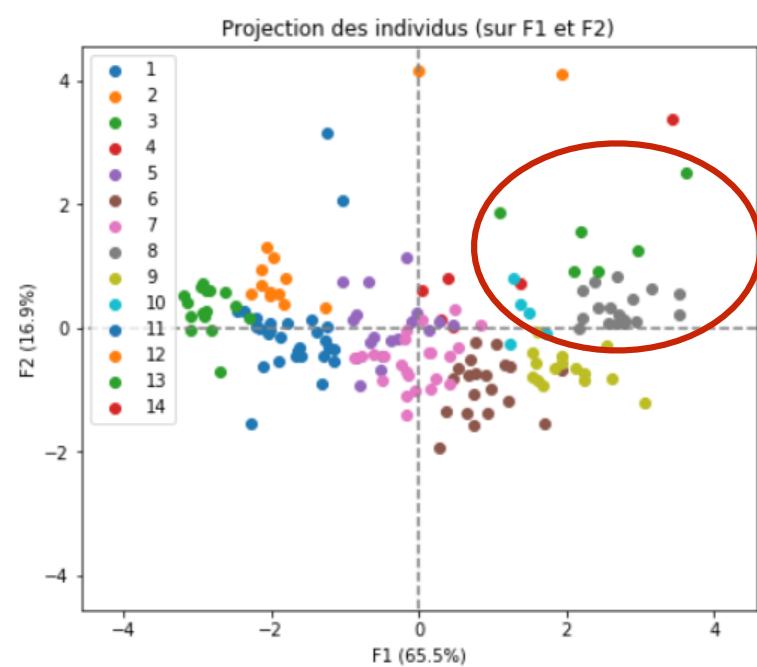
La qualité de vie
du citoyen
, sa consommation

01

Etude par centroïde

**7 classes de cluster****02**

Etude visuelle

**14 classes de cluster**

centroïde

Moyenne des centroïdes par classe

	Evolution_pop	Ratio_Animal	DA_pro	DA_cal	Valeur_PIB	cluster
1	1.748382	0.305171	61.693462	2398.115385	5042.188462	1
2	3.006708	0.216629	65.680909	2652.272727	4075.445455	2
3	2.734780	0.177884	53.393125	2202.875000	1655.337500	3
4	1.705363	0.284506	99.966000	3454.600000	12458.720000	4
5	1.714016	0.388774	81.678333	2841.555556	8401.744444	5
6	-0.127697	0.510580	86.586000	2970.650000	18879.895000	6
7	0.941040	0.529303	73.903043	2688.826087	14383.295652	7
8	0.636924	0.621752	111.721176	3456.000000	41869.258824	8
9	-0.164414	0.549543	104.396000	3344.066667	24340.206667	9
10	1.475718	0.618439	101.798000	3125.000000	20963.660000	10
11	5.452988	0.354667	79.650000	3082.500000	11578.650000	11
12	5.918043	0.506372	98.245000	3322.000000	58004.650000	12
13	1.589718	0.545275	101.255000	3322.833333	66730.766667	13
14	2.247148	0.644311	94.830000	2915.000000	135319.000000	14

Cluster	Evolution_pop	Ratio_Animal	DA_Pro	DA_Cal	Valeur_PIB
8	0.636924	0.621752	111.721176	3456	41869.2588
13	1.589718	0.545275	101,255	3322.8333	66730.76

Analyse sur les groupes de pays

```
69]: print(final.loc[(final['cluster']==13),:])
```

Pays	Evolution_pop	Ratio_Animal	DA_pro	DA_cal	Valeur_PIB
Arabie saoudite	2.779297	0.435774	91.63	3256.0	49537.6
Brunéi Darussalam	1.475909	0.566856	93.41	2985.0	79070.2
Luxembourg	2.128613	0.633298	113.88	3539.0	90656.4
Norvège	1.245002	0.595041	110.90	3486.0	62896.3
Suisse	1.192166	0.642136	93.08	3391.0	56535.7
Émirats arabes unis	0.717323	0.398547	104.63	3280.0	61688.4

Cluster 13

Analyse des pays

Taille du pays (km ²)	
Arabie Saoudite	2,15 millions
Brunéi Darussalam	5 765
Luxembourg	2 586
Norvège	385 203
Suisse	41 285
Emirats Arabes Unis	83 600

Comparaison Pays



4 pays



PIB



Affaires Instabilités Populations

Comparaison Pays

	ARABIE SAOUDITE	ÉMIRATS ARABES UNIS	SUISSE	NORVÈGE
POPULATION	32,6 millions	10,1 millions	8,4 millions	5,3 millions
PIB PAR HABITANT	21 096 \$US	37 733 \$US	80 637 \$US	75 389 \$US
EVALUATION DES RISQUES PAYS	C	A3	A1	A1
EVALUATION ENVIRONNEMENT DES AFFAIRES	B	A2	A1	A1
	A1	A2	A3	A4
	B	C	D	E

Test d'adéquation

01 Test sur la valeur PIB

H_0 =une loi normale H_1 ! = une loi normale

Si p-value est inférieur au seuil qui est de 5% on rejette l'hypothèse

```
74]: ks_2samp(final["Valeur_PIB"], list(numpy.random.normal(numpy.mean(final["Valeur_PIB"]),
    numpy.std(final["Valeur_PIB"]),
    1000)))|
```

```
Ks_2sampResult(statistic=0.173, pvalue=0.0003163469115481845)
```

p_value < 5%

02 Test l'évolution population

```
73]: ks_2samp(final["Evolution_pop"],
    list(numpy.random.normal(numpy.mean(final["Evolution_pop"]),
        numpy.std(final["Evolution_pop"]),
        1000)))|
```

```
Ks_2sampResult(statistic=0.053047904191616764, pvalue=0.8051402307148661)
```

p_value > 5%

Evolution population suit une loi normale

Test de comparaison de deux populations

```
75]: cluster8=final.loc[(final['cluster']==8),'Evolution_pop']

76]: cluster5=final.loc[(final['cluster']==5),'Evolution_pop']

77]: bartlett(cluster5,cluster8) #variance
BartlettResult(statistic=6.87103474926213, pvalue=0.008760391318214268)

78]: ttest_ind(cluster5,cluster8, equal_var=False) #moyenne
Ttest_indResult(statistic=5.455960391414747, pvalue=1.0867171774825219e-05)
```

