

PROJET 5 – PRODUIRE UNE ÉTUDE DE MARCHÉ



Quelles sont les pays les plus propices
pour accueillir notre développement à
l'international ?



Sommaire

1. Introduction aux enjeux

- Explications de l'étude de marché

2. Les données utilisées

- Nos données devront répondre à ces questions
- La composition de notre échantillon de travail
- Corrélations entre nos variables

3. Restriction du choix des pays

- Découpage des pays selon les variables de notre dataframe
- Explication des différents clusters
- Comment les pays les plus intéressants ont-ils été choisis ?
- Résumé

4. Les tests statistiques

- Nos valeurs suivent-elles une loi normale ?
- Nos clusters sont-ils distincts ?

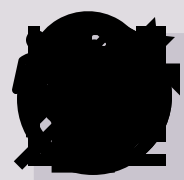
5. Conclusion



A person wearing a light blue button-down shirt is seated at a dark wooden table. Their right hand is positioned over a tablet computer, with the index finger pointing at the screen. To the right of the tablet, a pair of glasses with dark frames and red temples lies on the table. Further right, a white ceramic cup sits on a matching saucer. The background is softly blurred, showing a white cushion and a window with light streaming in. The overall lighting is soft and natural, creating a calm, professional atmosphere.

INTRODUCTION AUX ENJEUX





Les enjeux & la méthode de notre développement à l'international

Les enjeux

Promouvoir les volailles françaises à l'international

Rechercher la croissance

Augmentation du chiffre d'affaire

La méthode

Optimiser le choix des pays pour notre développement

Minimiser le risque d'échec

Ne rien laisser au hasard



LES DONNÉES UTILISÉES





Mes données
devront répondre à
ces questions :



Quels sont les pays qui ont le plus de
disponibilité alimentaire ?

Quels pays sont les plus gros consommateurs de
viande de volailles ?

Quels pays sont les plus gros importateurs de
viande de volaille ?

Quels sont les pays au PIB/hab le plus élevé ?

La composition de notre échantillon de travail

pays	variance de la population 2014-2018	PIB/hab en dollar	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Taux de disponibilité en protéines animales par rapport aux total	Taux de disponibilité en protéines de type volaille par rapport aux total	Taux de dépendance à l'importation de viande de volaille par rapport au total	Taux d'autosuffisance de viande de volaille par rapport au total
Afghanistan	11.39	483.89	2038.0	55.50	19.44	0.92	47.17	54.72
Afrique du Sud	5.96	6369.23	2895.0	84.37	42.72	18.06	23.11	79.08
Albanie	-0.47	5254.38	3361.0	115.67	53.38	3.04	44.44	59.26
Algérie	8.49	4153.96	3321.0	91.82	26.92	2.31	0.00	100.00
Allemagne	2.06	47684.76	3554.0	105.15	60.89	7.07	52.28	94.30

- Des variables qui vont répondre à nos questions précédentes
 - Des valeurs qui correspondent à chaque pays
 - Source des données : www.fao.org

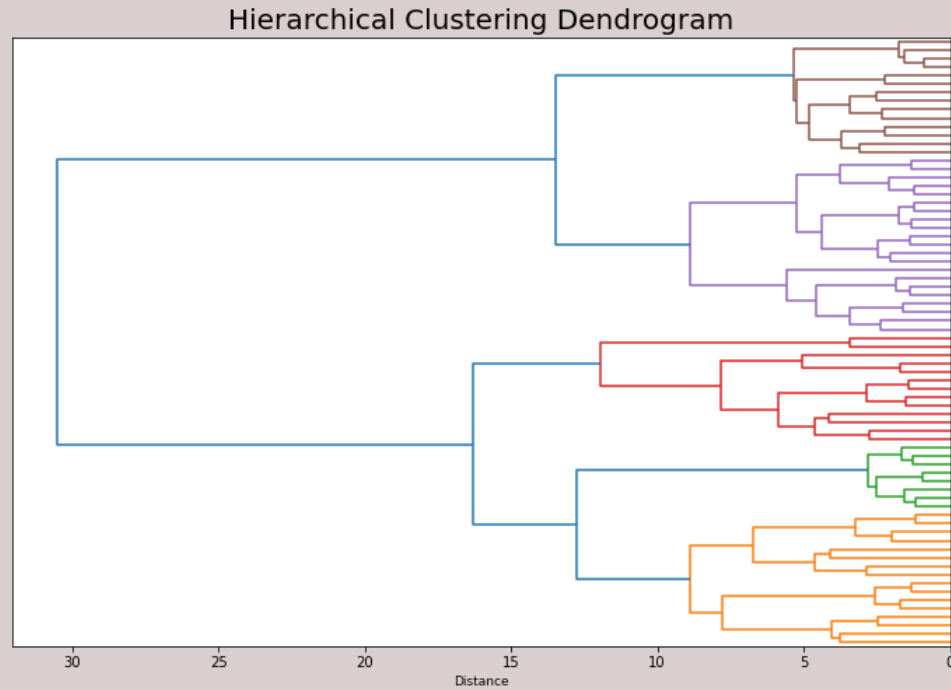
A stylized globe centered on the African continent, rendered in shades of green and blue. The landmasses are dark green, while the oceans are a lighter, teal blue. The globe is slightly tilted, showing parts of Europe, Asia, and Australia.

RESTRICTION DU CHOIX DES PAYS

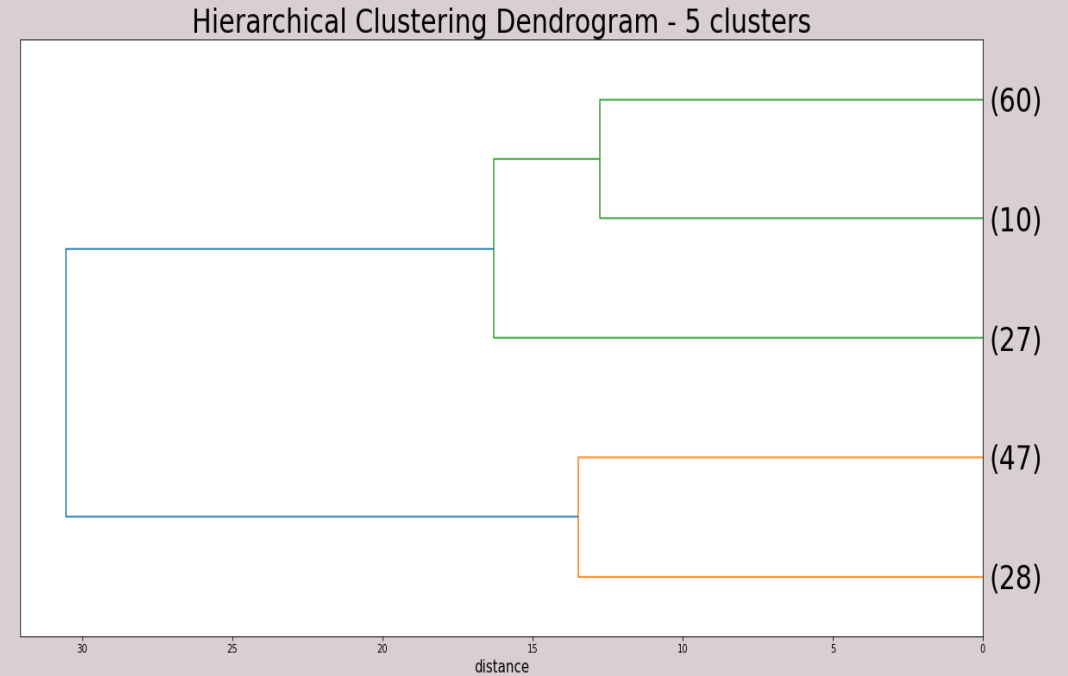


Restriction du choix des pays

Partitionnement – Classification hiérarchique



Création de 5 clusters

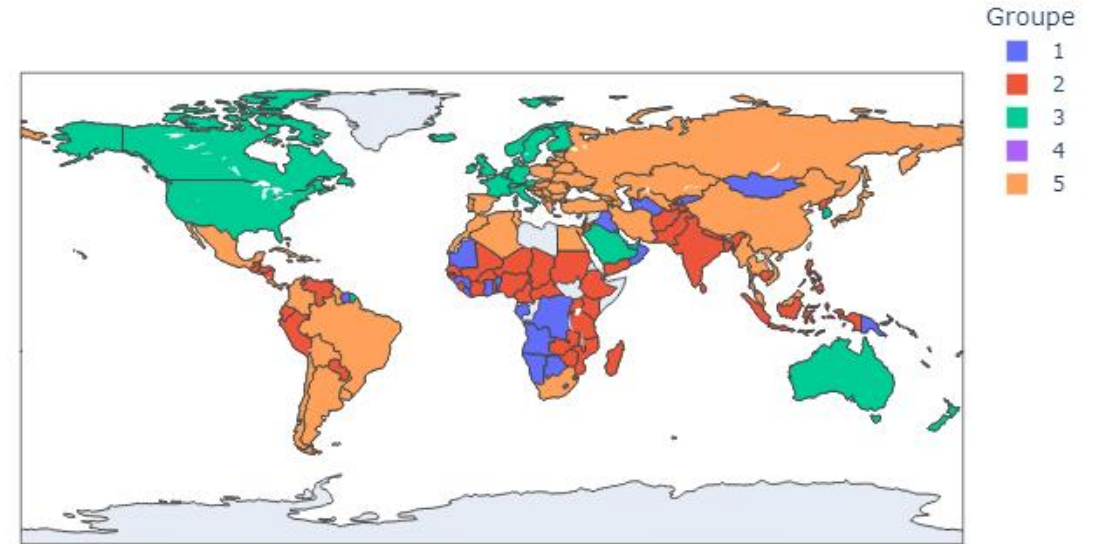


- Clustering selon un algorithme de 'minimisation de la variance'
- 5 différents groupes ont bien été crée

Restriction du choix des pays

Quelles sont les caractéristiques de nos clusters ?

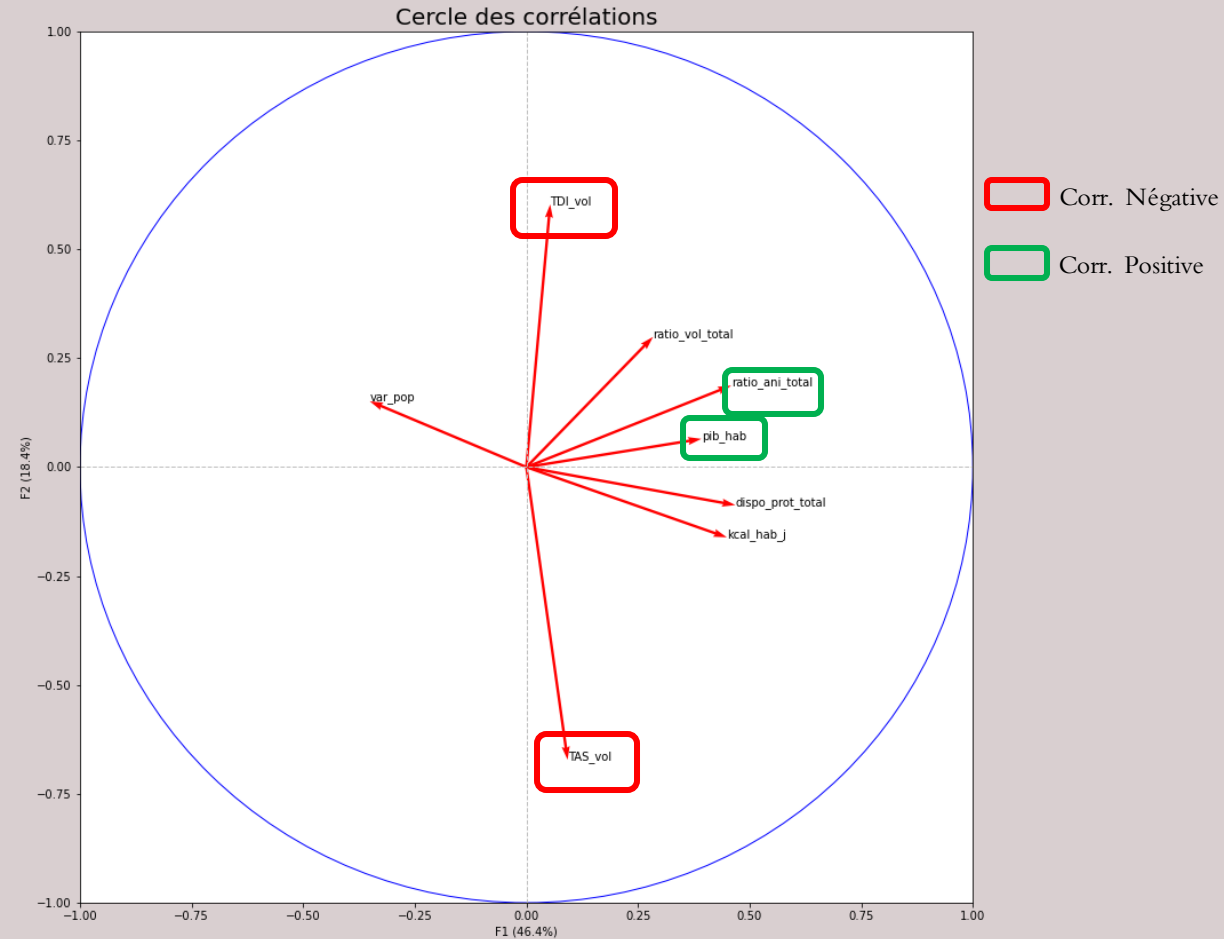
- **Cluster 1** : variance de population élevée et un PIB/hab faible
- **Cluster 2** : similaire au cluster 1 et un taux d'autosuffisance élevé
- **Cluster 3** : un PIB/hab élevé et un régime alimentaire basé sur la viande de volaille convenable
- **Cluster 4** : Similaire au cluster 3 avec un PIB/hab plus faible mais un régime alimentaire basé sur la viande de volaille très élevé
- **Cluster 5** : un taux d'autosuffisance trop élevé

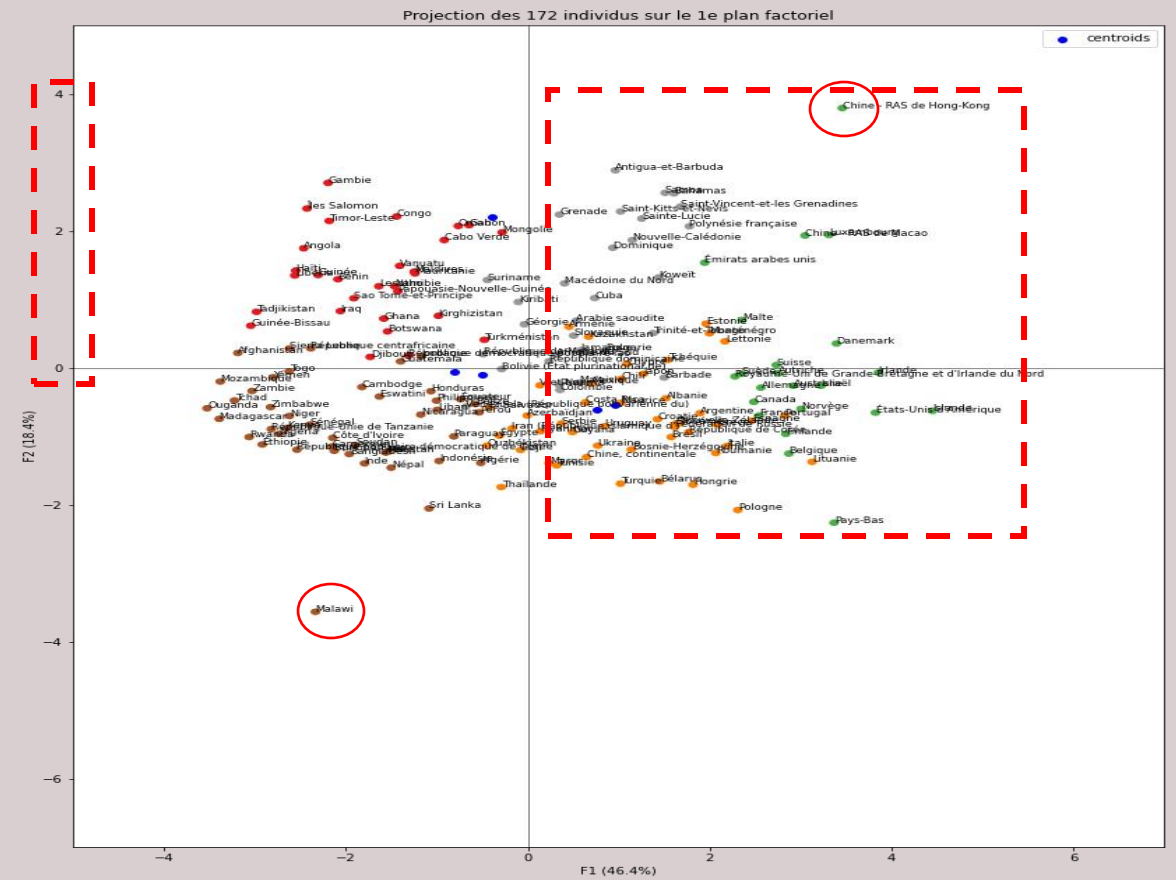
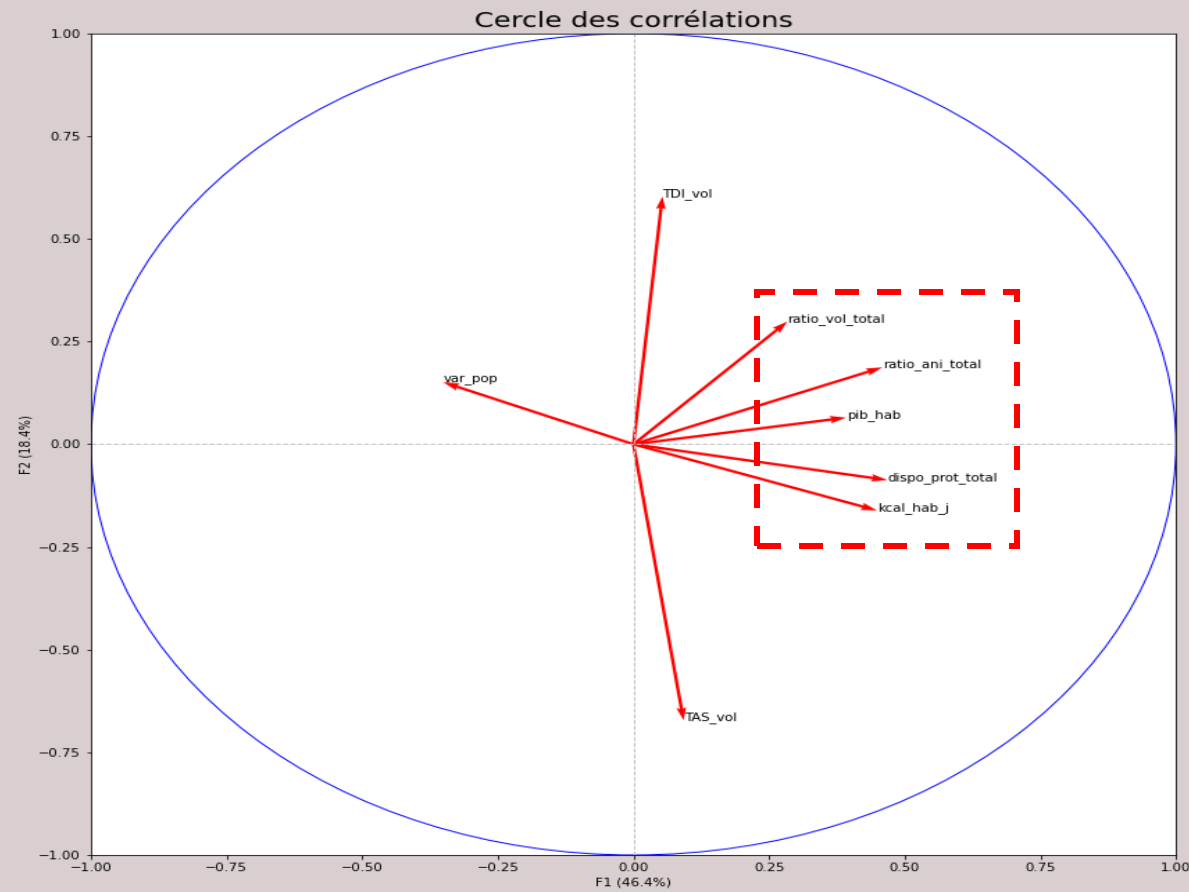


Corrélations entre nos variables



- Permet de visualiser la corrélations entre nos variables
- Permet de conserver et d'expliquer le plus de variance possible de nos individus
- L'ACP permet de réduire le nombre de dimension



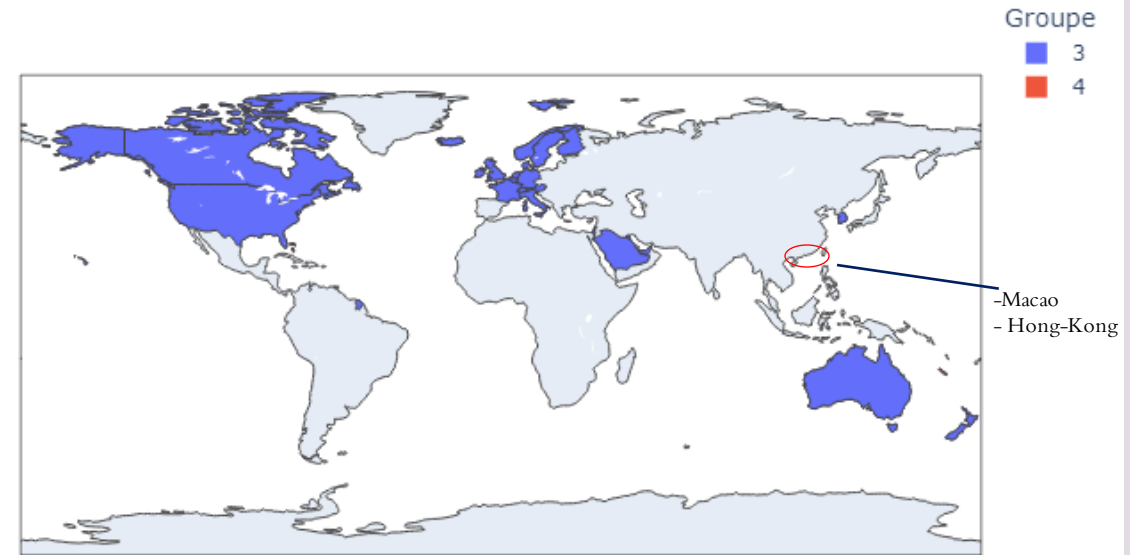


- Les pays situés le plus à droite ont une disponibilité alimentaire et un PIB/hab plus important
- Les pays situés le plus haut ont une dépendance à l'importation plus élevée

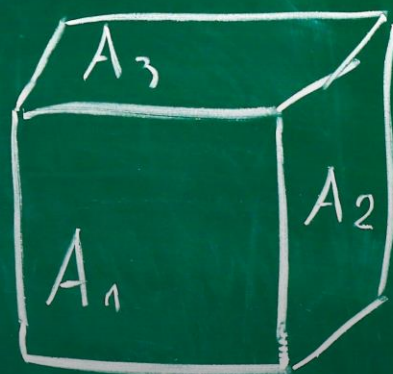
Restriction du choix des pays

Résumé

- Un total de 37 pays
- Les clusters 3 & 4 ont été retenus comme candidats
- Taux de dépendance à l'importation élevé
- Consommation de viande de volailles élevée
- PIB/hab élevé



$12=3$
 $3+4=7$
 60
 74
 105
 55



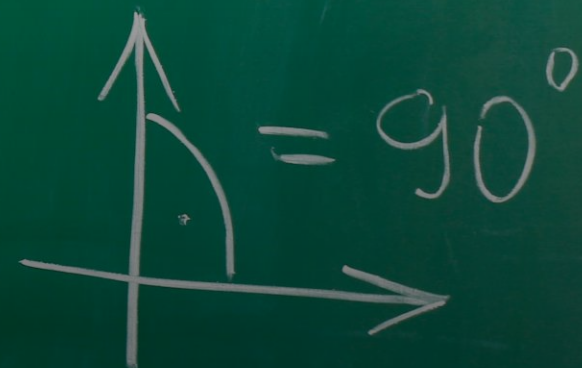
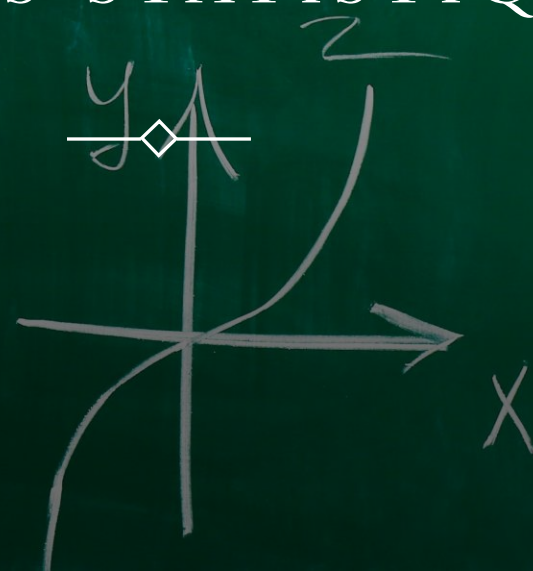
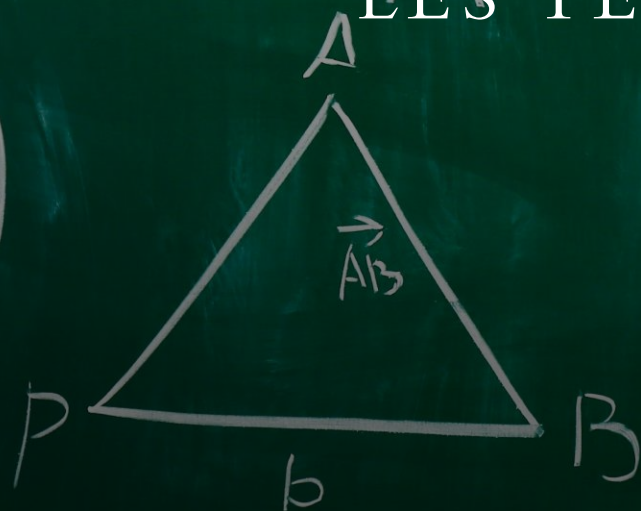
$A_1 + u = c$
 $A_2 + b = c$
 $A_3 = u \cdot b$

$\frac{2}{1} = 2$

$4x^5 - x^4 + 2x^3 + x^3 - 1) : (x^2 + 1) = 4x^3 - x^2 + 2x \frac{2x-3}{x+1}$

LES TESTS STATISTIQUES

$P_1 = \begin{pmatrix} 8 \\ -3 \\ 2 \end{pmatrix}$



$6x + 3 - 1 + 2x = (4 \cdot 2x) + x - 7$
 $6x + 3 - 1 + 2x = 8x + x - 7$

Test d'adéquation de Shapiro-Wilk

Recommandé pour tester la normalité dans le cas de petits échantillons.

- H_0 : Ma variable suit une loi normale
- H_1 : Ma variable ne suit pas une loi normale

```
Entrée [43]: # Test plus optimal concernant des petits échantillons allant jusqu'à 5000 échantillon(source : https://lemakistatheux.com)
stat, p = shapiro(df_analyse_cluster['pib_hab'])

print('Statistics=%.3f, p=%.3f' % (stat, p))

#Interprétation
alpha = 0.05
if p > alpha:
    print('On ne peut pas rejeter  $H_0$  pour des niveaux de test de 5%')
else:
    print('H0 est rejetée pour des niveaux de test de 5%')

Statistics=0.958, p=0.170
On ne peut pas rejeter  $H_0$  pour des niveaux de test de 5%
```

Le Test de Shapiro-Wilk est plus précis que celui de Kolmogorov-Smirnov, et également plus adapté dans notre cas de petit échantillonnage. La variable 'pib/hab' suit une loi normale.

- Notre variable suit bien une loi normale

Test de l'égalité de la variance de bartlett

La variable 'pib_hab' suit une loi normale et sera par conséquent choisie pour le test.

- H0 : Mon cluster 3 a le même niveau de PIB/hab que le cluster 4
- H1 : Mon cluster 3 n'a pas le même niveau de PIB/hab que le cluster 4

```
Entrée [44]: # Sélection de deux différents clusters pour les comparées dans un cas gaussien
cluster_test3 = df_analyse_cluster[df_analyse_cluster['Groupe'] == 3]['pib_hab']
cluster_test4 = df_analyse_cluster[df_analyse_cluster['Groupe'] == 4]['pib_hab']
```

```
Entrée [45]: #On teste tout d'abord l'égalité des variances à l'aide de la commande
stat, p = bartlett(cluster_test3, cluster_test4)
print('Statistics=%.3f, p=%.3f' % (stat, p))

#Interprétation
alpha = 0.05
if p > alpha:
    print('On ne rejette donc pas H0, l'égalité des variances au niveau de test 5%')
else:
    print('H0 est rejetée au niveau de test 5%')
```

```
Statistics=4.213, p=0.040
H0 est rejetée au niveau de test 5%
```

- Nos clusters ont des variances distinct

Test de l'égalité de la moyennes de Student

La variable 'pib_hab' suit une loi normale et sera par conséquent choisie pour le test.

- H0 : Mon cluster 3 a la même moyenne dans ma variable PIB/hab que le cluster 4
- H1 : Mon cluster 3 n'a pas la même moyenne dans ma variable PIB/hab que le cluster 4

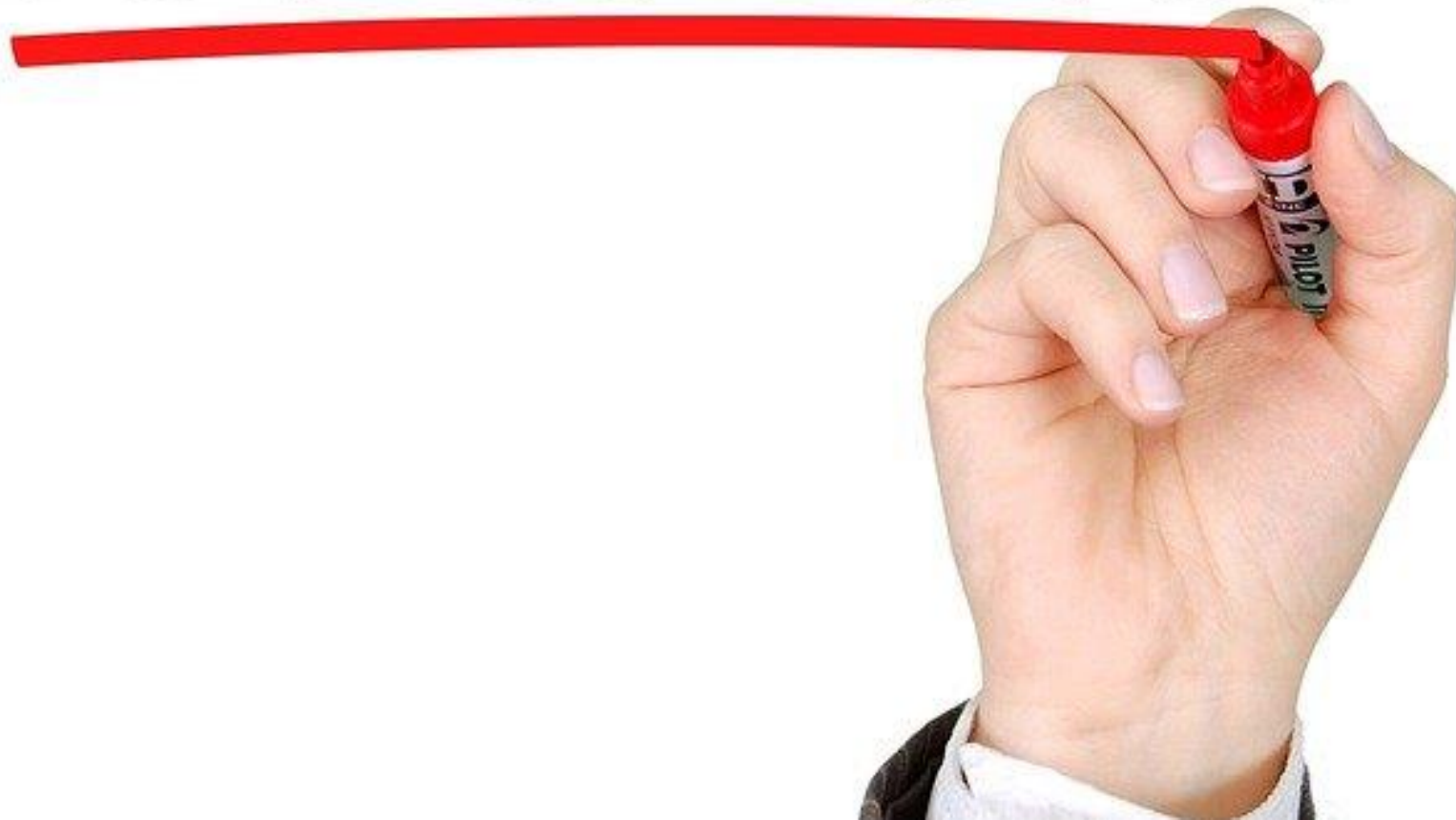
```
#On teste ensuite l'égalité des moyennes à l'aide de la commande(t-test)
stat, p = ttest_ind(cluster_test3, cluster_test4, equal_var=True)
print('Statistics=%.3f, p=%.9f' % (stat, p))

#Interprétation
alpha = 0.05
if p > alpha:
    print('On ne rejette donc pas H0, l'égalité des moyennes de nos 2 clusters au niveau de test 5%')
else:
    print("H0 l'hypothèse d'égalité des moyennes est rejetée au niveau de test 5%")
```

```
Statistics=5.419, p=0.000004503
H0 l'hypothèse d'égalité des moyennes est rejetée au niveau de test 5%
```

- La moyenne de nos clusters sont bien distinct

CONCLUSION





Conclusion

	pays	d_i	Groupe
0	Pays-Bas	67.584329	3
1	Chine - RAS de Hong-Kong	33.894540	3
2	Luxembourg	33.815931	3
3	Belgique	29.408971	3
4	Islande	25.919180	3
5	Malawi	25.435774	2
6	Irlande	20.442021	3
7	Chine - RAS de Macao	20.418656	3
8	Gambie	19.808839	1
9	États-Unis d'Amérique	17.812332	3
10	Israël	16.156967	3
11	Norvège	16.057685	3
12	Saint-Vincent-et-les Grenadines	15.517083	4
13	Suisse	15.507045	3
14	Ouganda	15.272202	2
15	Lituanie	15.020342	5
16	Danemark	13.592463	3
17	Îles Salomon	13.487001	1
18	Madagascar	13.338935	2
19	Samoa	13.091356	4



- 1. Pays-Bas
- 2. Luxembourg
- 3. Belgique
- 4. Irlande
- 5. Danemark

- 1. Hong-Kong
- 2. Islande
- 3. Macao
- 4. États-Unis d'Amérique
- 5. Israël