

## Sujet d'étude: Effet de l'émission de gaz à effet de serre sur la température

Sujet présenté par

Yaya CISSE

En vue d'obtenir le Diplôme de niveau 6 (BAC +3/4)

Parcours de Data Analyst sur OpenClassrooms

Spécialisation : préparation, analyse et modélisation des données.

### Table des matières

Sujet d'étude: Effet de l'émission de gaz à effet de serre sur la température.....	1
Introduction.....	1
I. Problématique.....	2
II. Présentation des données.....	2
1) Importation des dataframes.....	2
2) Nettoyage des données et jointure des tables.....	4
III. Analyse univariée.....	6
1) Déforestation.....	6
2) CH <sub>4</sub> .....	7
3) N <sub>2</sub> O.....	7
4) CO <sub>2</sub> .....	8
5) Température.....	9
IV. Classification hiérarchique ascendante.....	10
1) Dendrogramme.....	10
2) Les clusters.....	11
V. Analyse bivariée.....	12
1) Les boxplots.....	12
2) Les pairplots.....	17
VI. Analyse en composantes principales (ACP).....	18
1) Éboulis des valeurs propres.....	18
2) Cercle de corrélations et projections des individus.....	19
Conclusion.....	20
Tableau des dataframes.....	20
Tableau des illustrations.....	21

### Introduction

De nos jours, le débat sur le réchauffement climatique de la planète est un sujet primordial.

Vu l'importance du sujet, nous avons choisi d'étudier minutieusement les données liées aux changements climatiques.

Notre travail consistera à poser la problématique liée aux variations de la température, ensuite on présentera les données des dataframes. Après cela, nous analyserons de manière univariée et multivariée nos informations.

De plus, nous effectuerons une classification hiérarchique des données et enfin nous ferons une analyse en composantes principales des clusters.

## **I. Problématique**

Beaucoup de personnes sont sceptiques sur la réalité du réchauffement climatique dans le monde.

Du coup, on a décidé de nous focaliser sur l'impact de la déforestation plus précisément de l'émission des gaz à effet de serre suite aux feux de forêts.

Avec cette étude, l'objectif est de démontrer l'ampleur du réchauffement climatique de la planète

## **II. Présentation des données**

La définition de certains concepts est très importante avant d'entrer dans le vif du sujet.

Ainsi, le réchauffement climatique<sup>1</sup> comprend à la fois le réchauffement induit par les émissions de gaz à effet de serre d'origine humaine et les changements de régimes météorologiques à grande échelle qui en résultent.

Pour mener à bien notre analyse sur l'effet de l'émission des gaz à effets de serre sur la variation de la température, nous importerons les tables nécessaires.

### **1) Importation des dataframes**

Avant de commencer l'importation des données, il faut trouver les bonnes sources.

Dans un premier temps, nous chargerons la déforestation, les émissions de gaz à effet de serre et la variation des températures.

#### **1.a) La Déforestation**

La déforestation<sup>2</sup> est un phénomène irréversible de réduction de la surface de forêts. Elle est causée par de multiples facteurs naturels comme humains et a des conséquences irréversibles sur l'environnement.

Les données de la déforestation<sup>3</sup> sont issues de Kaggle.

La table contient 4 colonnes:

Year: année

Code: code iso des pays

Country: les pays

Net forest conversion: expansion ou régression des forêts en hectares

---

<sup>1</sup> Le réchauffement climatique  
[https://fr.wikipedia.org/wiki/R%C3%A9chauffement\\_climatique](https://fr.wikipedia.org/wiki/R%C3%A9chauffement_climatique)

<sup>2</sup> La définition de la déforestation  
<https://climate.selectra.com/fr/comprendre/deforestation>

<sup>3</sup> Source des données de la déforestation  
<https://www.kaggle.com/chiticariucristian/deforestation-and-forest-loss>

	Year	Country	Deforestation
0	1990-01-01	Africa	4096000.00
248	1990-01-01	Madagascar	70500.00
116	1990-01-01	Djibouti	0.00
249	1990-01-01	Malawi	42000.00
253	1990-01-01	Maldives	0.00

Table 1: dataframe de la déforestation

### 1.b) Les émissions de gaz dues aux feux de forêts

Les gaz à effet de serre<sup>4</sup> sont Gaz d'origine naturelle (vapeur d'eau) ou anthropique (liée aux activités humaines) absorbant et réémettant une partie des rayons solaires (rayonnement infrarouge), phénomènes à l'origine de l'effet de serre.

Les principaux gaz à effet de serre (GES) liés aux activités humaines sont le dioxyde de carbone (CO<sub>2</sub>), le méthane (CH<sub>4</sub>), l'oxyde nitreux ou protoxyde d'azote (N<sub>2</sub>O) et des gaz fluorés.

Pour cette partie, les émissions de GES proviennent des données de la FAO<sup>5</sup>.

Dans cette table nous avons les pays, les années et les émissions de gaz en kilotonnes.

	Domain Code	Domain	Area Code (FAO)	Area	Element Code	Element	Item Code	Item	Year Code	Year	Source Code	Source	Unit	Value	Flag	Flag Description	Note
0	GT	Emissions Totals	2	Afghanistan	7225	Emissions (CH <sub>4</sub> )	1707	LULUCF	1990	1990	3050	FAO TIER 1	kilotonnes	0.01	Fc	Calculated data	NaN
1	GT	Emissions Totals	2	Afghanistan	7230	Emissions (N <sub>2</sub> O)	1707	LULUCF	1990	1990	3050	FAO TIER 1	kilotonnes	0.00	Fc	Calculated data	NaN
2	GT	Emissions Totals	2	Afghanistan	7273	Emissions (CO <sub>2</sub> )	1707	LULUCF	1990	1990	3050	FAO TIER 1	kilotonnes	-2388.80	Fc	Calculated data	NaN
3	GT	Emissions Totals	2	Afghanistan	7225	Emissions (CH <sub>4</sub> )	1707	LULUCF	1991	1991	3050	FAO TIER 1	kilotonnes	0.01	Fc	Calculated data	NaN
4	GT	Emissions Totals	2	Afghanistan	7230	Emissions (N <sub>2</sub> O)	1707	LULUCF	1991	1991	3050	FAO TIER 1	kilotonnes	0.00	Fc	Calculated data	NaN

Table 2: dataframe des émissions des gaz à effet de serre

### 1.c) Dataframe de la variation de la température

La variation de la température désigne ici la différence entre la température annuelle et la température de référence entre 1951 et 1980.

La table est tirée des données de changements climatiques<sup>6</sup> de la FAO.

<sup>4</sup> Les gaz à effet de serre

<https://www.insee.fr/fr/metadonnees/definition/c1855>

<sup>5</sup> Les données des émissions de gaz à effet de serre

<https://www.fao.org/faostat/en/#data/GT>

<sup>6</sup>The FAOSTAT Temperature Change

<https://www.fao.org/faostat/en/#data/ET/metadata>

Sur cette table, on trouve les variables: Year, Country et Temperature\_change.

	Domain Code	Domain	Area Code (FAO)	Area	Element Code	Element	Months Code	Months	Year Code	Year	Unit	Value	Flag	Flag Description
0	ET	Temperature change	2	Afghanistan	7271	Temperature change	7020	Meteorological year	1961	1961	?C	-0.12	Fc	Calculated data
1	ET	Temperature change	2	Afghanistan	7271	Temperature change	7020	Meteorological year	1962	1962	?C	-0.17	Fc	Calculated data
2	ET	Temperature change	2	Afghanistan	7271	Temperature change	7020	Meteorological year	1963	1963	?C	0.84	Fc	Calculated data
3	ET	Temperature change	2	Afghanistan	7271	Temperature change	7020	Meteorological year	1964	1964	?C	-0.78	Fc	Calculated data
4	ET	Temperature change	2	Afghanistan	7271	Temperature change	7020	Meteorological year	1965	1965	?C	-0.25	Fc	Calculated data

Table 3: dataframe sur les variations des températures

## 2) Nettoyage des données et jointure des tables

Les tables importées sont des données brutes pour bien étudier ces dernières, il est nécessaire de bien nettoyer les tables.

### 2.a) Nettoyage des données

- Exemple de la table température

Pour cela, on va effectuer une projection sur la table de la variation de la température sur les colonnes Year, Area et Value.

Ces dernières seront renommées Area en Country et Value en Temperature\_change.

Quant aux valeurs manquantes, on les corrigera par une imputation par la moyenne.



	Year	Country	Temperature_change
0	1961-01-01	Afghanistan	-0.12
1	1962-01-01	Afghanistan	-0.17
2	1963-01-01	Afghanistan	0.84
3	1964-01-01	Afghanistan	-0.78
4	1965-01-01	Afghanistan	-0.25

Table 4: Projection, nettoyage et imputation de la table température

- Exemple de la table des émissions de gaz

Concernant cette table, il faut effectuer d'abord un pivot de la table avec comme index Area, Year et Unit, les groupes de valeur de la colonne Element seront transformées en colonnes et la variable Values sera agrégée.

Ensuite je renomme Area en Country et Value en Temperature\_change, Emissions (CH4) en CH4, Emissions (N2O) en N2O et Emissions (CO2) en CO2.

Après cela, on effectue la restriction uniquement sur les pays.

Enfin, on projette sur les variables Year, Country, Unit, CH4, CO2, N2O.

Element	Country	Year	Unit	CH4	CO2	N2O
0	Afghanistan	1990-01-01	kilotonnes	0.01	-2388.80	0.00
1	Afghanistan	1991-01-01	kilotonnes	0.01	-2388.80	0.00
2	Afghanistan	1992-01-01	kilotonnes	0.01	-2388.80	0.00
3	Afghanistan	1993-01-01	kilotonnes	0.01	-2388.80	0.00
4	Afghanistan	1994-01-01	kilotonnes	0.01	-2388.80	0.00

Table 5: Pivot, renommage, restriction et projection sur la table d'émissions de gaz

-Exemple de la table de la déforestation

On fait une projection sur les variables Year, Country et Deforestation.

	Year	Country	Deforestation
0	1990-01-01	Africa	4096000.00
248	1990-01-01	Madagascar	70500.00
116	1990-01-01	Djibouti	0.00
249	1990-01-01	Malawi	42000.00
253	1990-01-01	Maldives	0.00

Table 6: Projection sur la table déforestation

## 2. b) Jointure des tables

Dorénavant, on a les 3 tables la température, la déforestation et les émissions de gaz.

De ce fait, je fait une double jointure à gauche en premier lieu entre la déforestation et les émissions de gaz sur les clés Country et Year et en second lieu entre la 1ère jointure et la température sur les clés Country et Year.

On fait un Groupby par Country et par Year avec une fréquence par année.

Et on supprime les valeurs manquantes.

	Year	Country	Deforestation	Unit	CH4	CO2	N2O	Temperature_change
0	1990-01-01	Algeria	14800.00	kilotonnes	1.76	573.79	0.13	1.31
1	2000-01-01	Algeria	14500.00	kilotonnes	9.63	573.79	0.53	0.86
2	2010-01-01	Algeria	1400.00	kilotonnes	0.18	-2405.50	0.02	2.29
3	2015-01-01	Algeria	5400.00	kilotonnes	0.07	-625.42	0.01	1.15
4	1990-01-01	Argentina	213600.00	kilotonnes	26.66	44315.78	1.23	0.39

Table 7: jointure des tables de la déforestation, des émissions de gaz à effet de serre et des variations de température

### III. Analyse univariée

L'analyse univariée est une analyse statistique sur une seule variable.

Cette analyse se basera sur 5 pays le Brésil, la Norvège, la Mozambique, l'Indonésie et le Groenland.

#### 1) Déforestation

Les déforestations qu'on a eu à observer de 1990 à 2015 se sont passées au Brésil se sont les pires.

	Year	Country	Deforestation	Unit	CH4	CO2	N2O	Temperature_change
42	2000-01-01	Brazil	5129300.00	kilotonnes	87.01	1038487.71	2.58	0.55
41	1990-01-01	Brazil	4254800.00	kilotonnes	222.14	1038468.19	14.50	0.44
180	1990-01-01	Indonesia	2526000.00	kilotonnes	2590.61	684395.58	4.93	0.12
43	2010-01-01	Brazil	1867800.00	kilotonnes	326.58	1101118.52	29.82	1.11
44	2015-01-01	Brazil	1695700.00	kilotonnes	183.88	261986.42	16.79	1.60

Table 8: les pires déforestations de 1990 à 2015

Pour la représentation graphique, comme les valeurs sont trop grandes, on utilise le logarithme de la déforestation pour mieux les présenter sans trahir l'allure des graphes.

Le logarithme des déforestations du Brésil passe de 15,2 à 15,3 de 1990 à 2000 qui est la plus grande augmentation.

De 2000 à 2015, il diminue de 15,3 à 14,3.

La Norvège a l'évolution du logarithme de la déforestation la meilleure entre les 5 pays choisis.

De 1990 à 2000, la Norvège passe de 8,5 à 8,8.

Et de 2000 à 2015, le logarithme décroît de 8,8 à 8,6.

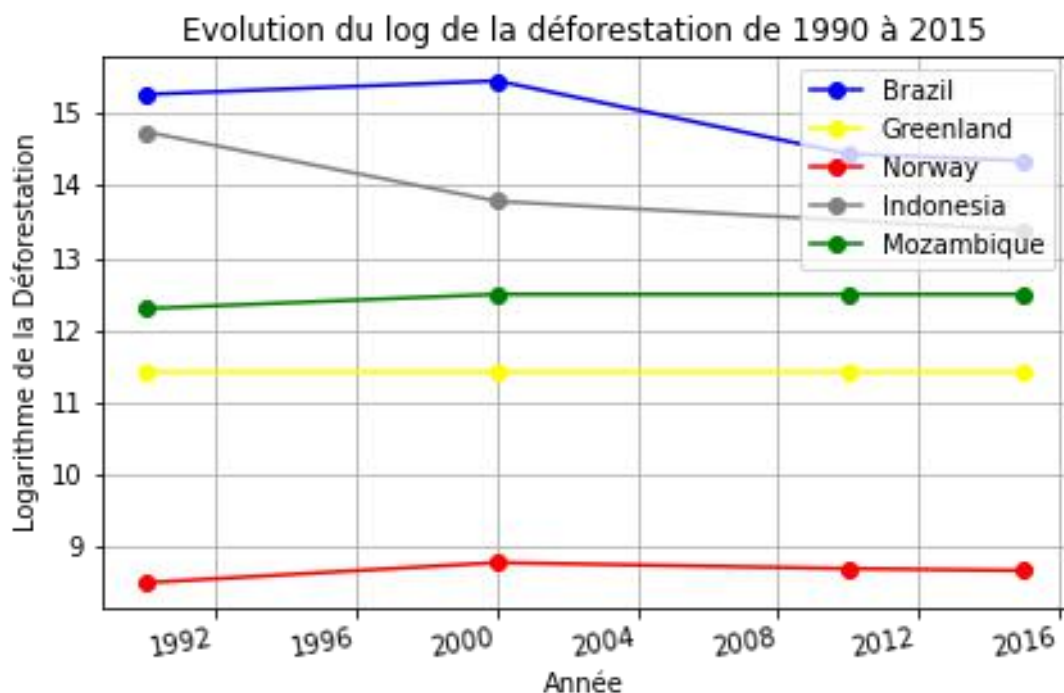


Figure 1: évolution du logarithme de la déforestation au cours du temps



## 2) CH4

En 2015, on a la plus forte émission de CH4 qui est de 5295 kilotonnes en Indonésie.

	Year	Country	Deforestation	Unit	CH4	CO2	N2O	Temperature_change
182	2015-01-01	Indonesia	650000.00	kilotonnes	5295.79	1083923.70	6.44	0.79
180	1990-01-01	Indonesia	2526000.00	kilotonnes	2590.61	684395.58	4.93	0.12
262	2010-01-01	Mozambique	267030.00	kilotonnes	521.13	48927.67	47.58	0.90
394	2015-01-01	Zambia	189710.00	kilotonnes	433.64	32840.42	39.59	1.86
263	2015-01-01	Mozambique	267030.00	kilotonnes	390.79	45580.68	35.68	0.97

Table 9: évolution de l'émission de CH4 au cours du temps

Pour l'Indonésie, l'émission de CH4 décroît fortement de 2600 à 300 kilotonnes, de 1990 à 2000. De 2000 à 2015, l'émission de CH4 augmente drastiquement de 300 à 5700 kilotonnes.

En ce qui concerne la Norvège, l'évolution est assez stable à 0 de 1990 à 2015.

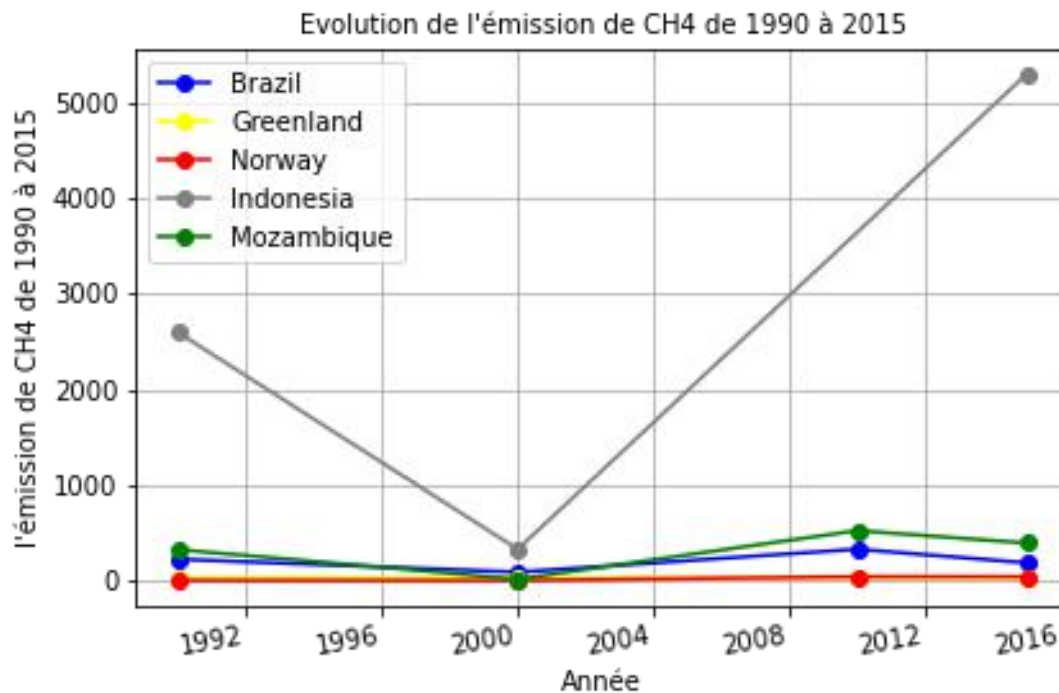


Figure 2: évolution de l'émission de CH4 au cours du temps

## 3) N2O

La Mozambique a la plus grande émission de N2O se situe à 47,58 kilotonnes en 2010.

	Year	Country	Deforestation	Unit	CH4	CO2	N2O	Temperature_change
262	2010-01-01	Mozambique	267030.00	kilotonnes	521.13	48927.67	47.58	0.90
394	2015-01-01	Zambia	189710.00	kilotonnes	433.64	32840.42	39.59	1.86
263	2015-01-01	Mozambique	267030.00	kilotonnes	390.79	45580.68	35.68	0.97
64	2010-01-01	Central African Republic	30000.00	kilotonnes	379.40	13967.33	34.64	1.30
393	2010-01-01	Zambia	189680.00	kilotonnes	355.89	17443.14	32.49	1.45

Table 10: les plus grandes émissions de N2O

De 1990 à 2000, l'émission de N2O la Mozambique diminue grandement de 29 à 0 kilotonne.

De 2000 à 2010, elle croit fortement de 0 à 48 kilotonnes.

Et de 2010 à 2015, l'émission décroît de 48 à 36 kilotonnes.

Pour la Norvège, on constate une faible augmentation de l'émission de N2O en passant de 0 à 2 kilotonnes de 1990 à 2015.

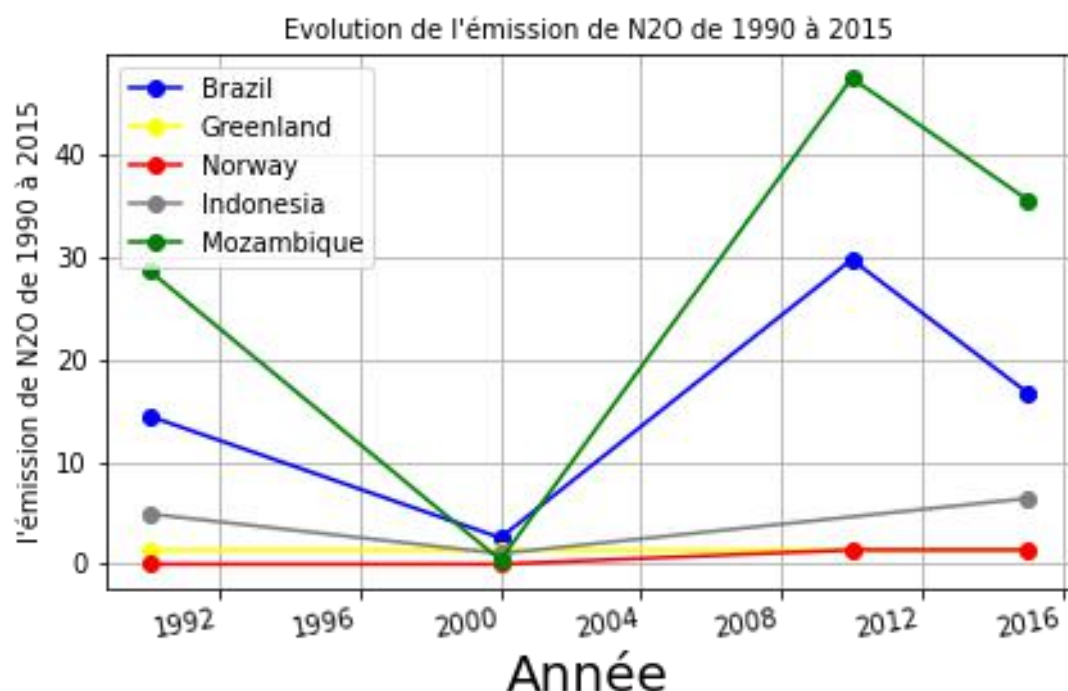


Figure 3: les plus grandes émissions de N2O au cours du temps

#### 4) CO2

Le Brésil produit la plus grande quantité de CO2 qui est de 1 million de kilotonnes en 2010.

	Year	Country	Deforestation	Unit	CH4	CO2	N2O	Temperature_change
43	2010-01-01	Brazil	1867800.00	kilotonnes	326.58	1101118.52	29.82	1.11
182	2015-01-01	Indonesia	650000.00	kilotonnes	5295.79	1083923.70	6.44	0.79
42	2000-01-01	Brazil	5129300.00	kilotonnes	87.01	1038487.71	2.58	0.55
41	1990-01-01	Brazil	4254800.00	kilotonnes	222.14	1038468.19	14.50	0.44
180	1990-01-01	Indonesia	2526000.00	kilotonnes	2590.61	684395.58	4.93	0.12



Table 11: évolution de l'émission de CO2 au cours du temps

L'émission de CO2 du Brésil passe de 1,2 à 1,8 million de kilotonnes de 1990 à 2010.

Elle passe de 1,8 million de kilotonnes 280 000 kilotonnes de 2010 à 2015.

Pour la Norvège, l'émission de CO2 stagne autour de -100 000 kilotonnes de 1990 à 2015.

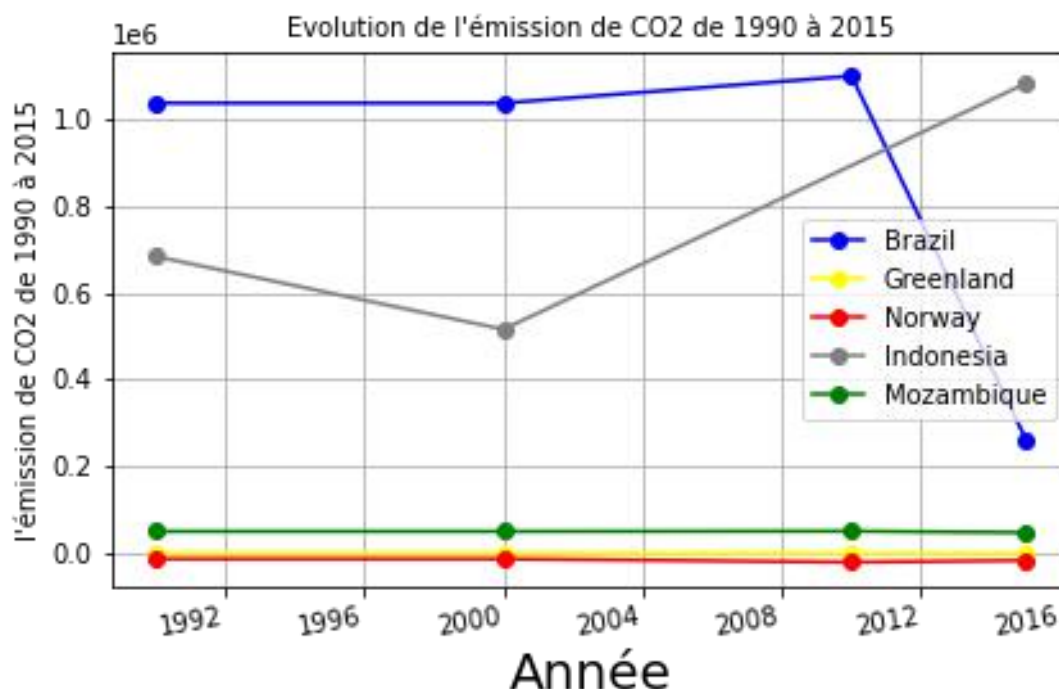


Figure 4: évolution de l'émission de CO2 au cours du temps

## 5) Température

Les pires changements de la variation de la température se sont passés au Groenland en 2010 est de 3°C.

	Year	Country	Deforestation	Unit	CH4	CO2	N2O	Temperature_change
149	2010-01-01	Greenland	93854.99	kilotonnes	40.47	-0.02	1.44	3.04
57	2010-01-01	Canada	41300.00	kilotonnes	7.35	301250.49	0.67	2.92
202	2010-01-01	Kuwait	93854.99	kilotonnes	40.47	-23.00	1.44	2.55
185	2010-01-01	Iraq	93854.99	kilotonnes	40.47	-51.70	1.44	2.54
117	2015-01-01	Estonia	5010.00	kilotonnes	0.00	-6152.65	0.00	2.35

Table 12: pires changements de la variation de la température

Pour la Norvège, de 1990 à 2010, la température passe de 1,6 à - 0,4°C.

De 2010 à 2015, elle passe de - 0,4°C jusqu'à 1,8°C.

Pour la Groenland, la température augmente drastiquement de - 0,47°C à 3,1°C de 1990 à 2010.

La température diminue grandement de 3,1°C à 0,2°C de 2010 à 2015.

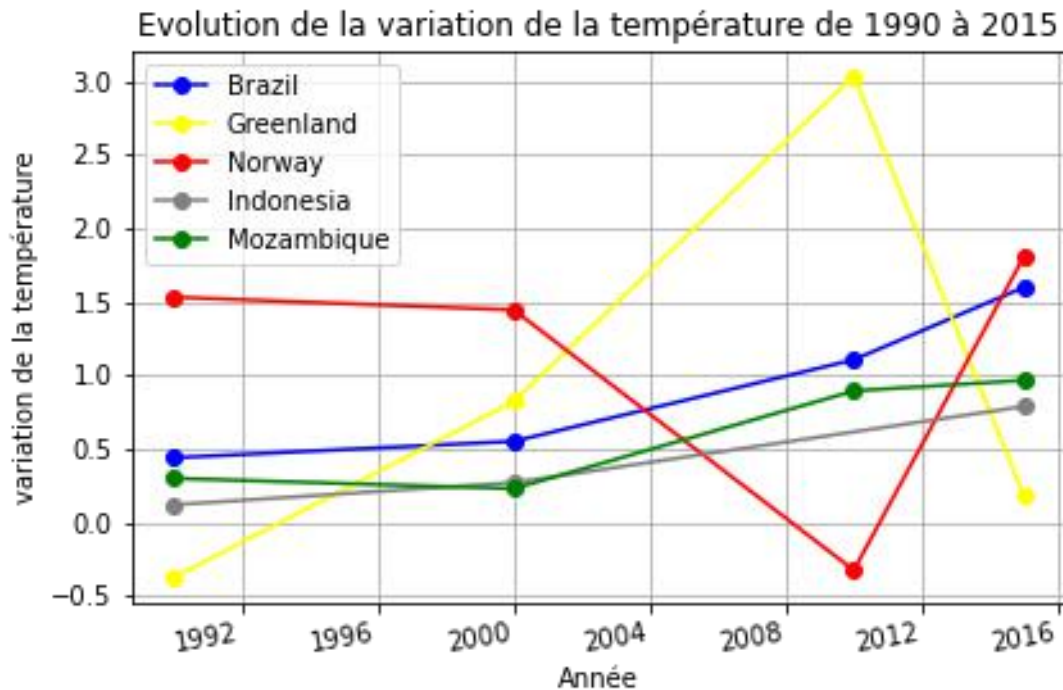


Figure 5: évolution de la variation des températures au cours du temps

#### IV. Classification hiérarchique ascendante

L'approche ascendante<sup>7</sup>, aussi appelée clustering agglomératif : on considère tout d'abord que chaque point est un cluster. Il y a donc autant de clusters que de points. Ensuite, on cherche les deux clusters les plus proches, et on les agglomère en un seul cluster. On répète cette étape jusqu'à ce que tous les points soient regroupés en un seul grand cluster.

##### 1) Dendrogramme

On obtient donc une arborescence qui a un cluster tout en haut, et qui se divise petit à petit jusqu'à avoir autant de clusters que de points. On appelle cette arborescence un dendrogramme.

<sup>7</sup> Classification hiérarchique ascendante

<https://openclassrooms.com/fr/courses/4525281-realisez-une-analyse-exploratoire-de-donnees/5177936-effectuez-une-classification-hierarchique>

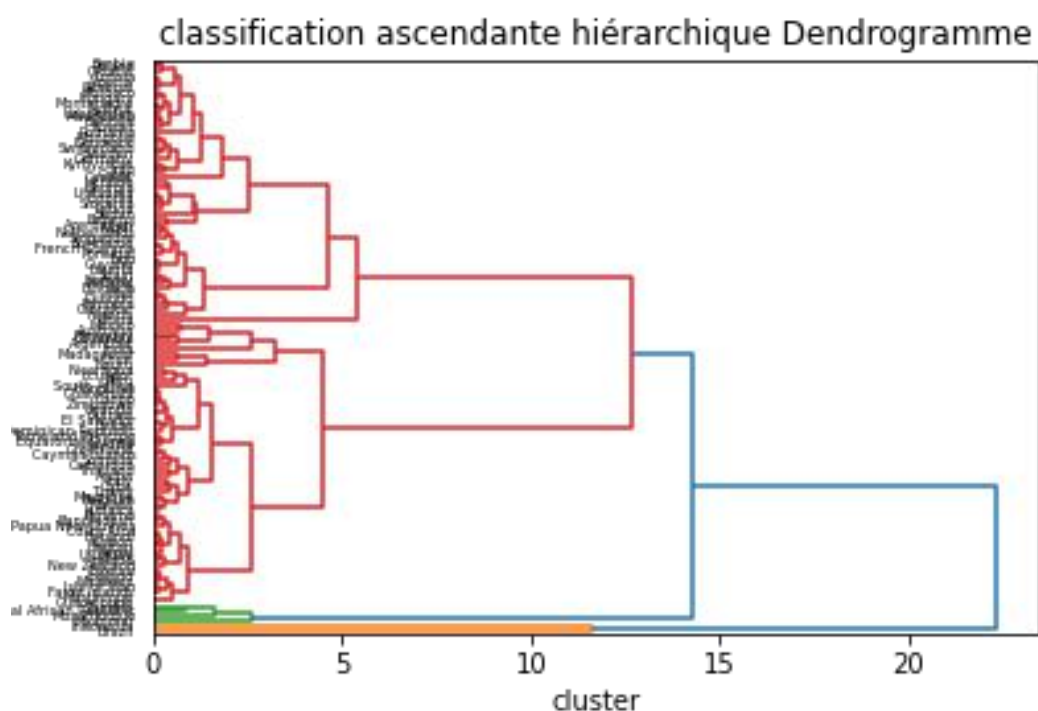


Figure 6: dendrogramme

## 2) Les clusters

Suite à la classification hiérarchique ascendante, on obtient 3 clusters avec des tailles différentes.

	Country	Deforestation	Temperature_change	CH4	N2O	CO2	cluster
0	Algeria	9025.00	1.40	2.91	0.17	-470.83	3
1	Argentina	229550.00	0.28	25.61	1.11	62261.59	3
2	Aruba	91337.05	0.70	39.38	1.40	-0.04	3
3	Australia	558446.67	0.42	62.71	3.60	43383.50	3
4	Austria	5712.50	1.48	19.70	0.70	-10044.28	3

Table 13: dataframe principal avec les clusters

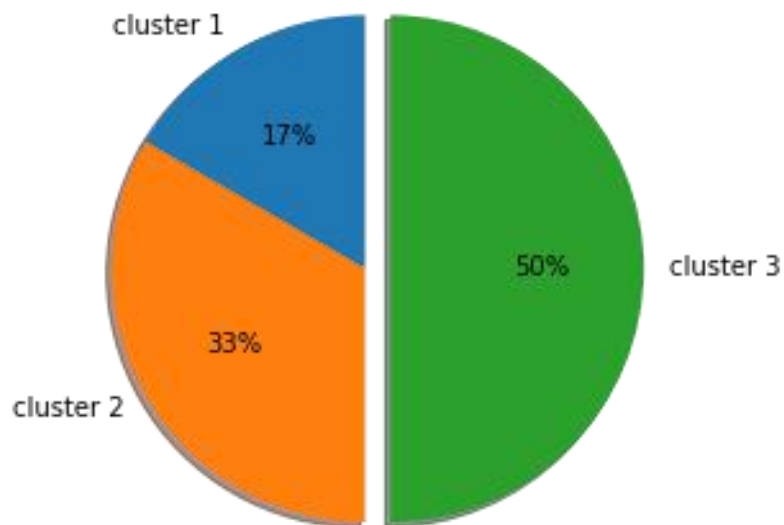


Figure 7: répartition des pays par cluster

Le cluster 1 représente 17% des pays

Le cluster 2 représente 33% des pays

Le cluster 3 représente 50% des pays

## V. Analyse bivariée

L'analyse bivariée<sup>8</sup> est une analyse menée entre plusieurs variables.

Elle permet d'établir des recommandations métier pertinentes sur les individus à partir de la compréhension du comportement d'une variable par rapport à une autre.

### 1) Les boxplots

Un diagramme en boîtes <sup>9</sup>à partir des colonnes de DataFrame. Un boxplot résume le minimum, le premier quartile, la médiane, le troisième quartile et la valeur maximale des données.

#### 1.a) Répartition de la déforestation

Le cluster 1 a un minimum de 1,4 million d'hectares et un maximum 3,8 millions d'hectares. La médiane et la moyenne valent 2,4 millions d'hectares.

Le cluster 3 a un minimum de 0 hectare et un maximum de 250 000 hectares.

La médiane est 100 000 hectares.

<sup>8</sup> L'analyse bivariée

<https://openclassrooms.com/fr/courses/7410486-nettoyez-et-analysez-votre-jeu-de-donnees/7428247-comprenez-les-enjeux-de-l-analyse-bivariee>

<sup>9</sup> Diagramme en boîtes

<https://www.delftstack.com/fr/api/python-pandas/pandas-dataframe-dataframe.boxplot-function/>

Nous constatons des outliers entre 250 000 et 600 000 hectares.

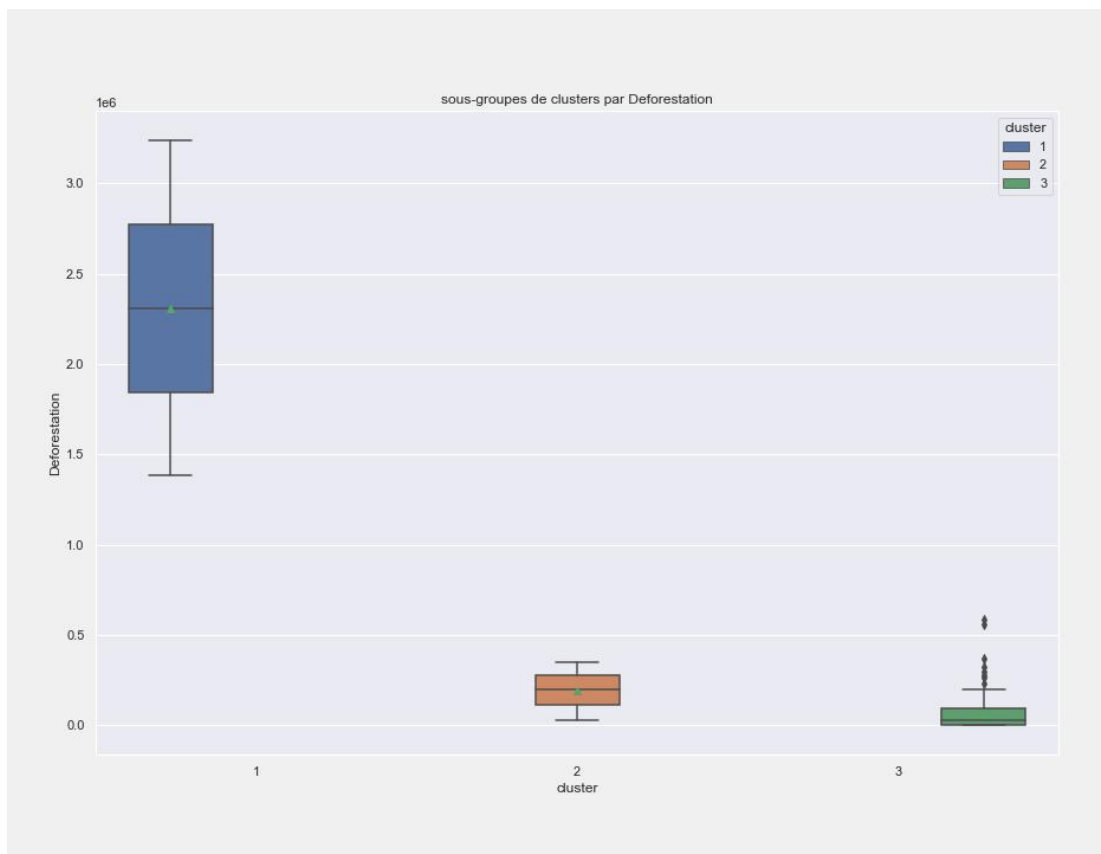


Figure 8: boîtes à moustaches de la déforestation selon les clusters

### 1. b) L'émission de CH<sub>4</sub> par cluster

Le cluster 1 a un minimum de 200 kilotonnes et un maximum de 2 850 kilotonnes d'émission de CH<sub>4</sub>. Ce cluster 1 a une médiane de 1 500 kilotonnes.

Le cluster 3 possède un minimum de 0 et un maximum de 180 kilotonnes de CH<sub>4</sub>.

La moyenne du cluster 3 est 10 kilotonnes de CH<sub>4</sub>.



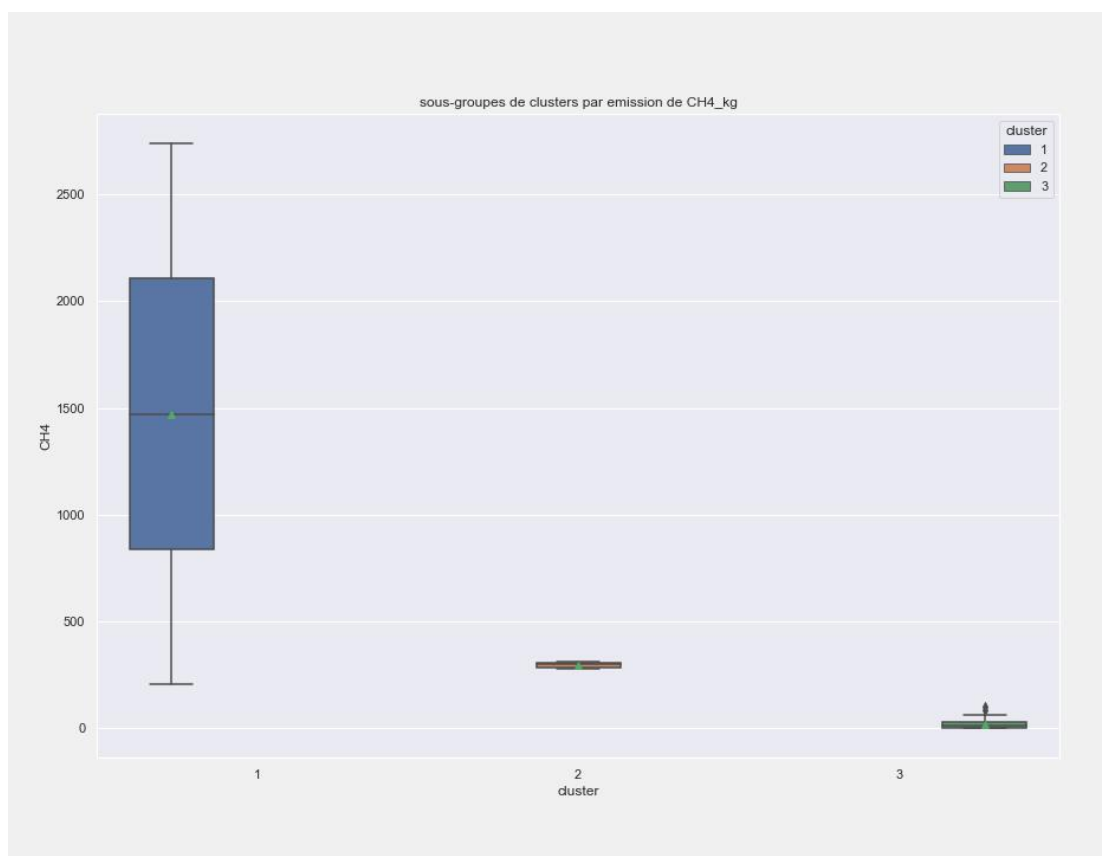


Figure 9: boîtes à moustaches de l'émission de CH4 selon les clusters

### 1.c) L'émission de N2O par cluster

Le cluster 2 a un minimum de 17 kilotonnes et un maximum de 28 kilotonnes d'émission de N2O. Ce cluster 2 a une médiane de 24 kilotonnes.

Le cluster 3 possède un minimum de 0 et un maximum de 4 kilotonnes de N2O.

La moyenne du cluster 3 est 1 kilotonne de N2O

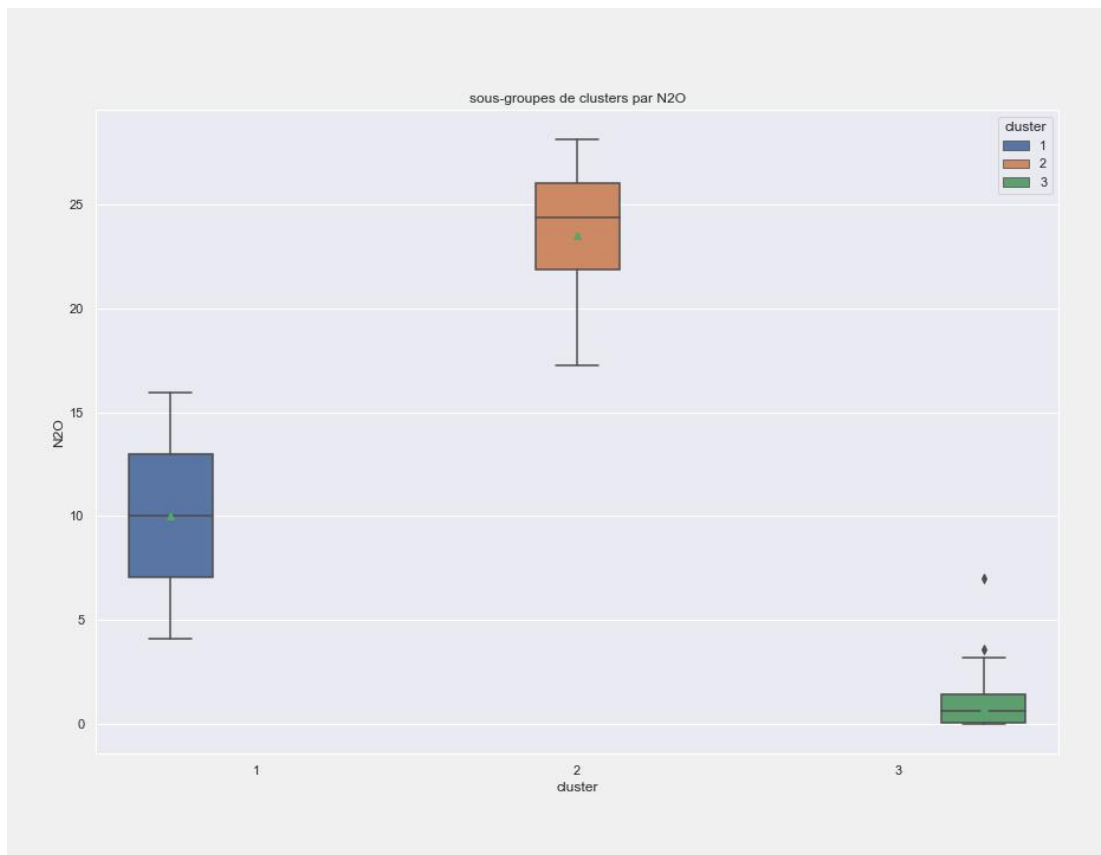


Figure 10: boîtes à moustaches de l'émission de N2O selon les clusters

#### 1.d) L'émission de CO2 par cluster

Le cluster 1 a un minimum de 780 000 kilotonnes et un maximum de 900 000 kilotonnes d'émission de CO2. Ce cluster 1 a une médiane de 820 000 kilotonnes.

Le cluster 3 possède un minimum de -20 000 et un maximum de 20 000 kilotonnes de CO2. La moyenne du cluster 3 est 0 kilotonne de CO2

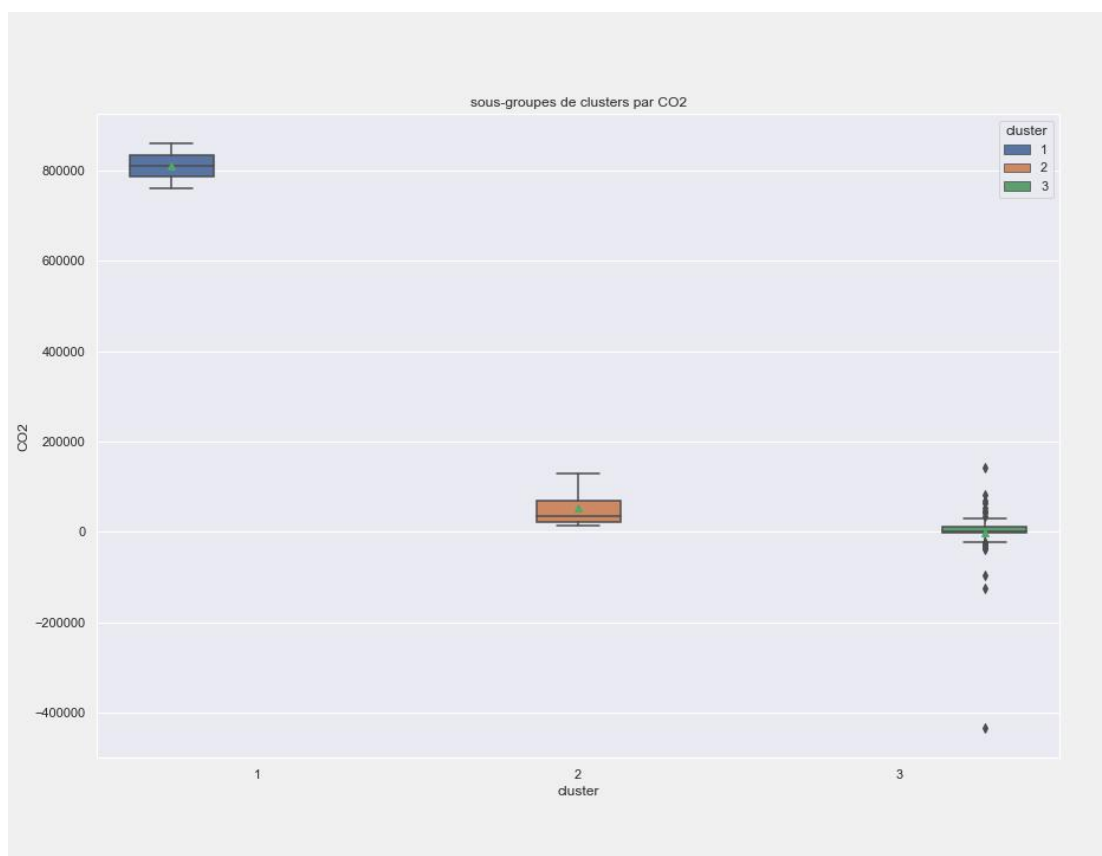


Figure 11: boîtes à moustaches de l'émission de CO2 selon les clusters

### 1.e) La variation des températures selon les clusters

Le cluster 3 a un minimum de -0,4°C et un maximum de 1,9°C.

Ce cluster 3 a une médiane de 0,9°C.

Le cluster 1 possède un minimum de 0,4°C et un maximum de 0,9°C.

La moyenne du cluster 1 est 0,7°C.

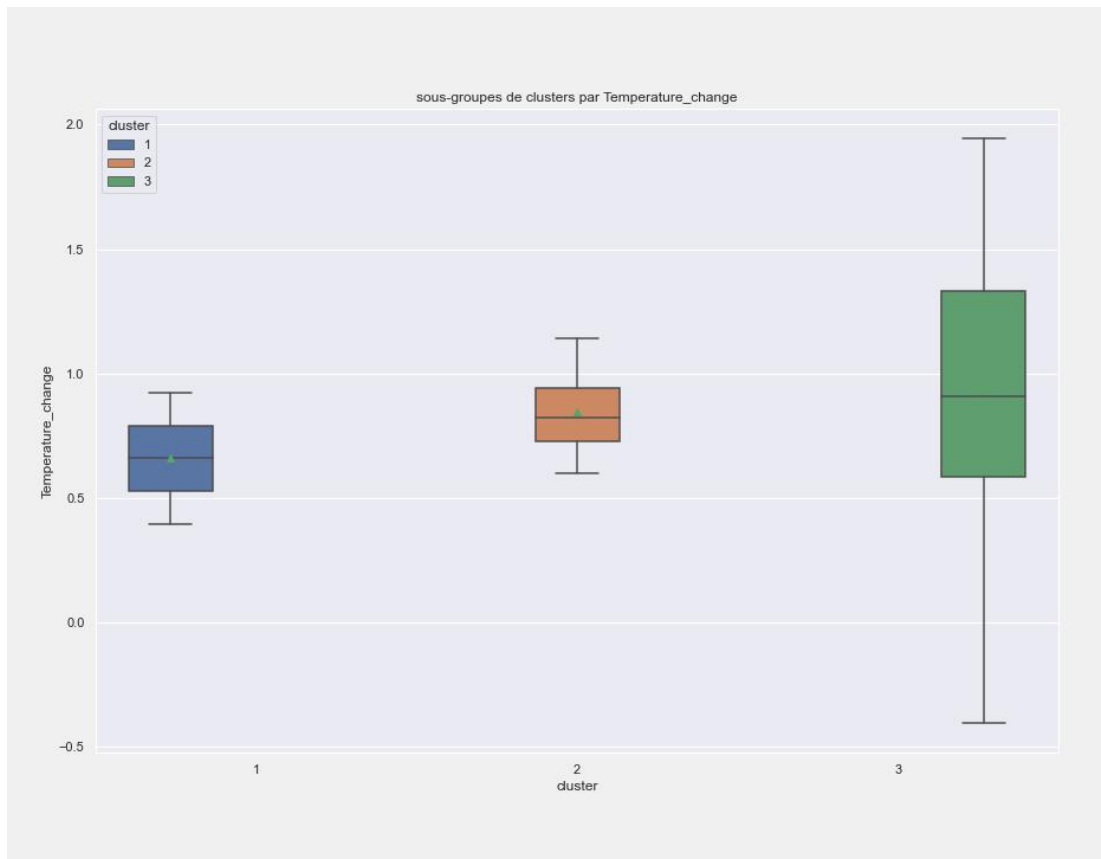


Figure 12: boîtes à moustaches de la variation de température selon les clusters

## 2) Les pairplots

Pour tracer plusieurs distributions bivariées par paires dans un ensemble de données, vous pouvez utiliser la fonction `pairplot()`<sup>10</sup>. Cela montre la relation pour (n, 2) combinaison de variable dans un DataFrame sous forme de matrice de tracés et les tracés diagonaux sont les tracés univariés.

<sup>10</sup> Les pairplots  
<https://fr.acervolima.com/python-methode-seaborn.pairplot/>

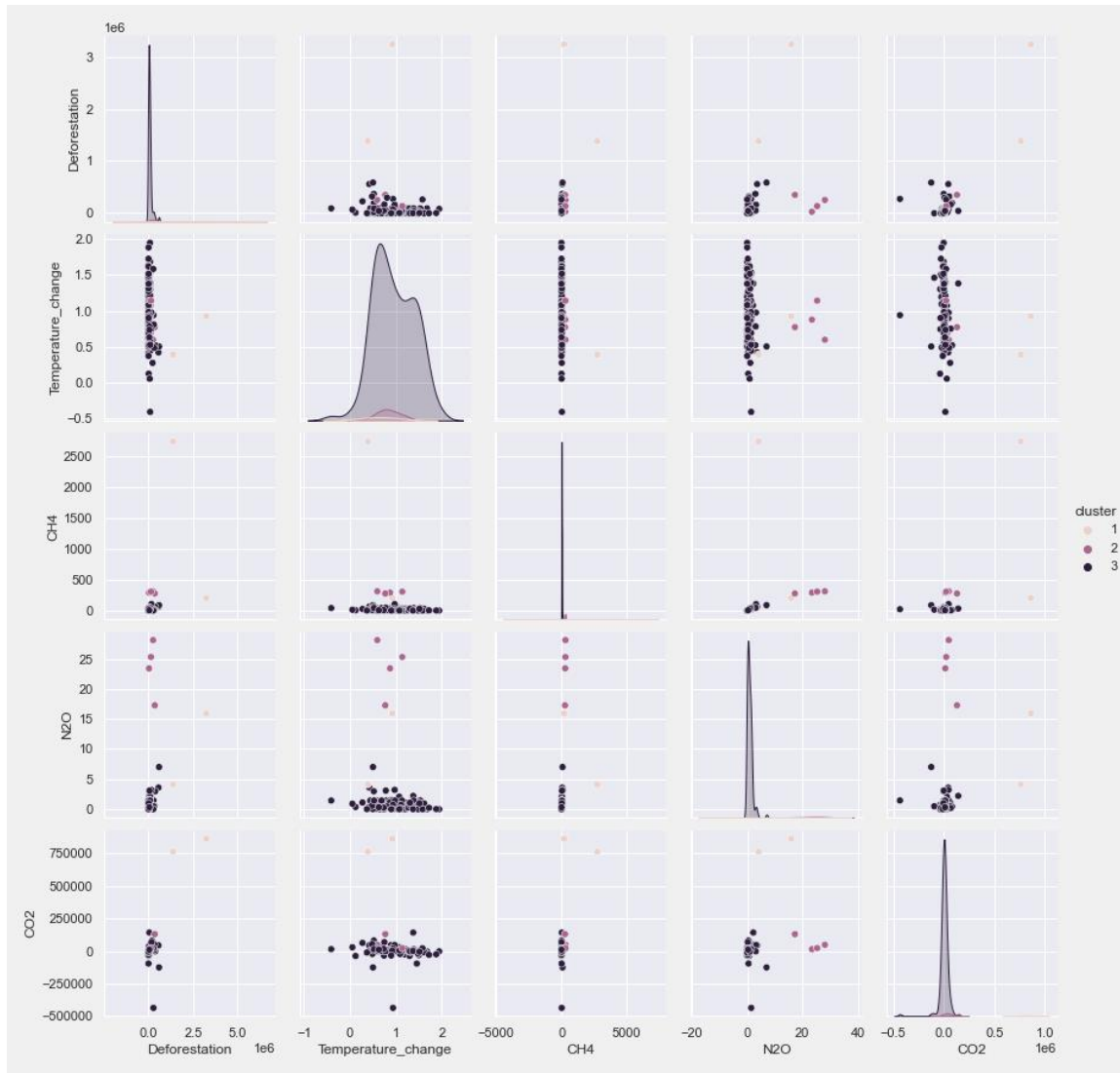


Figure 13: pairplot

## VI. Analyse en composantes principales (ACP)

L'analyse en composantes principales (ACP)<sup>11</sup> est une technique exploratoire très populaire. Selon les points de vue, on peut la considérer : comme une technique descriptive où l'on essaie de résumer les données dans ses dimensions les plus importantes ; comme une technique de visualisation où l'on essaie de préserver les proximités entre les individus dans un espace de représentation réduit ; comme une technique de compression de l'information ; etc.

### 1) Éboulis des valeurs propres

Quand on additionne les inerties associées à tous les axes, on obtient l'inertie totale du nuage des individus. On peut donc afficher un diagramme qui décrit le pourcentage d'inertie totale associé à chaque axe. On appelle ce diagramme l'éboulis des valeurs propres. Les 2 axes représentent plus de 80% de l'axe d'inertie.

<sup>11</sup>Analyse en composantes principales (ACP)

<http://tutoriels-data-mining.blogspot.com/2012/06/acp-avec-tanagra-nouveaux-outils.html>



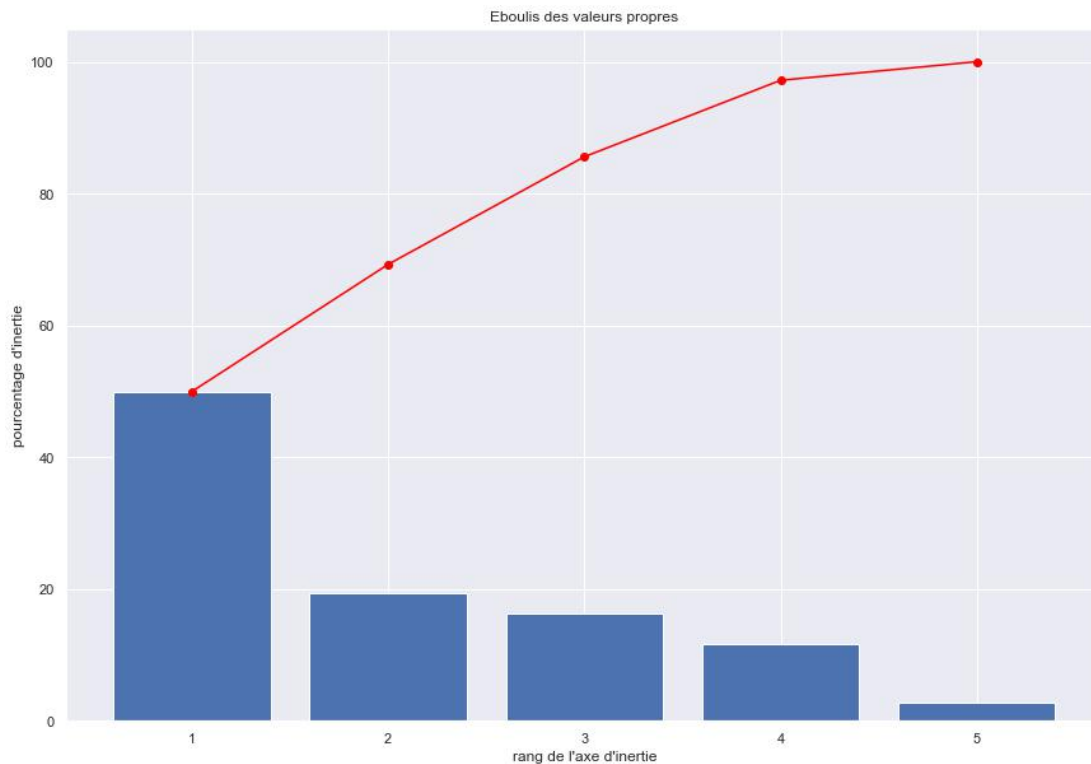
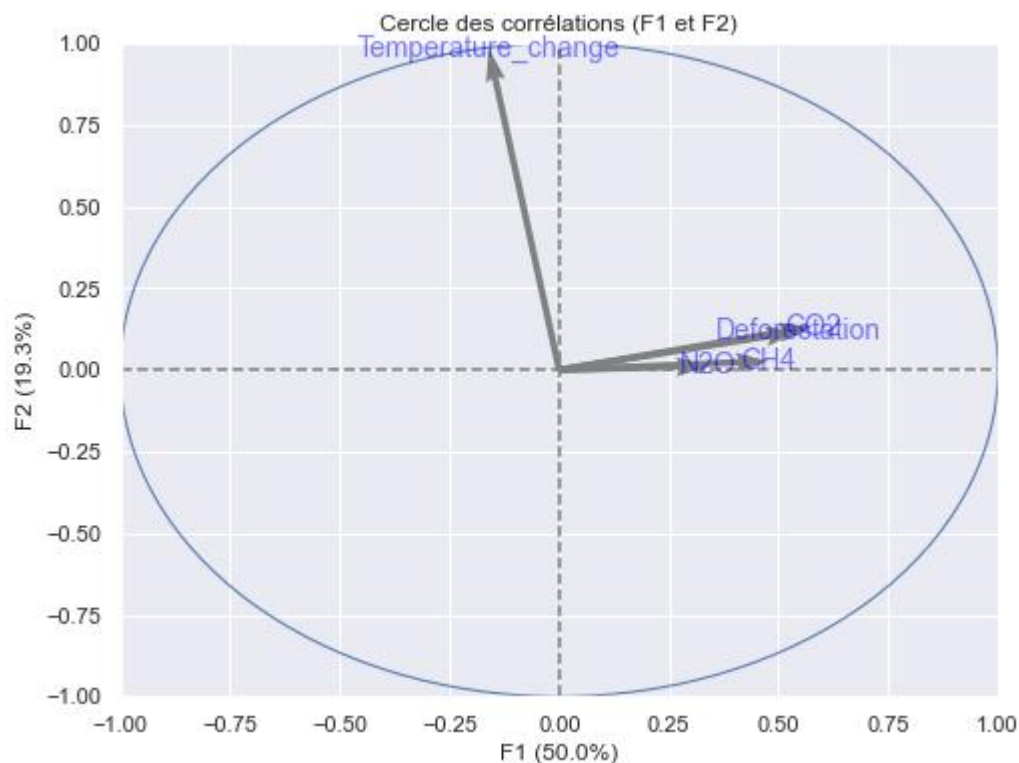


Figure 14: Éboulis des valeurs propres

## 2) Cercle de corrélations et projections des individus

Il suffit de savoir une chose : la projection de la flèche (représentant la variable Deforestation) sur F1 correspond au coefficient de corrélation entre Deforestation et F1. Rappelons qu'un coefficient de corrélation est compris entre -1 et 1. Cela tombe bien : le cercle est justement de rayon



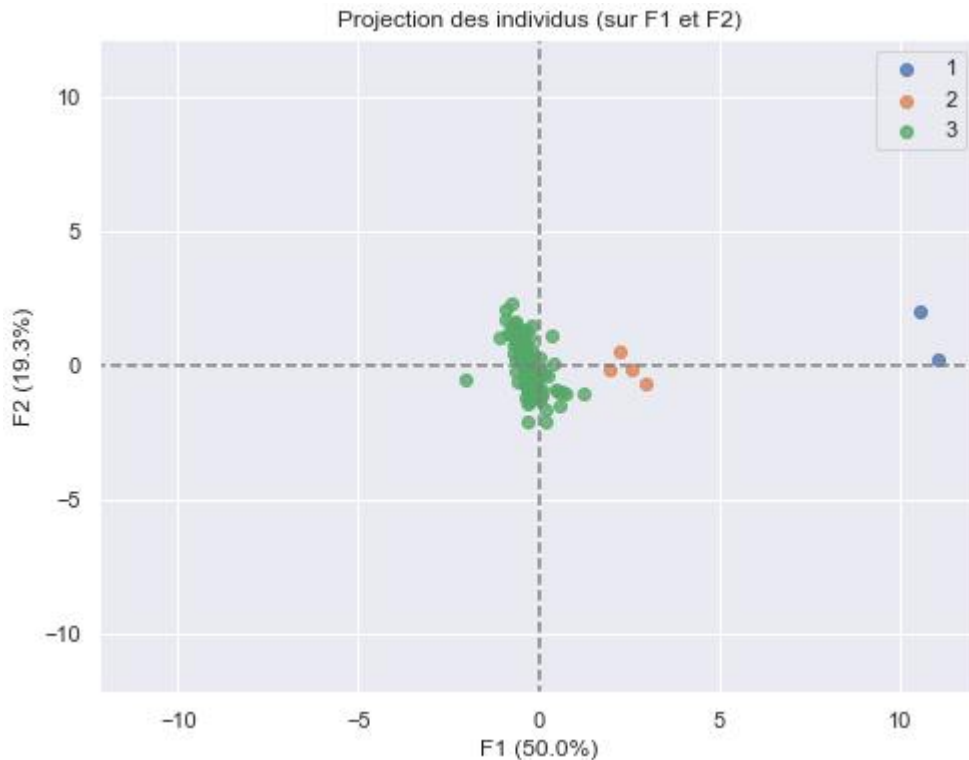


Figure 15: cercles de corrélations et projection des individus sur les 2 axes

Deforestation : sa projection sur F1 vaut 0.6.

Deforestation : sa projection sur F2 vaut -0.16.

Temperature\_change: sa projection sur F1 vaut -0.18.

Temperature\_change: sa projection sur F2 vaut 0.9.

L'interprétation des variables projetées sur F1 est que plus la déforestation est grande et plus les pays émettent des gaz à effet de serre.

Les 3 clusters sont nettement séparées, on peut parler de dissemblance

## **Conclusion**

L'analyse de la déforestation sur les variations de la température nous a permis de constater une corrélation entre la déforestation et les émissions de gaz à effet de serre.

De plus, l'augmentation des gaz à effet de serre implique une croissance de la variation de la température.

## **Tableau des dataframes**

Table 1 : dataframe de la déforestation.....	3
Table 2 : dataframe des émissions des gaz à effet de serre.....	3
Table 3 : dataframe sur les variations des températures.....	4
Table 4 : Projection, nettoyage et imputation de la table température.....	4
Table 5 : Pivot, renommage, restriction et projection sur la table d'émissions de gaz.....	5
Table 6 : Projection sur la table déforestation.....	5
Table 7 : jointure des tables de la déforestation, des émissions de gaz à effet de serre et des variations de	

température.....	5
Table 8 : les pires déforestations de 1990 à 2015.....	6
Table 9 : évolution de l'émission de CH <sub>4</sub> au cours du temps.....	7
Table 10 : les plus grandes émissions de N <sub>2</sub> O.....	8
Table 11 : évolution de l'émission de CO <sub>2</sub> au cours du temps.....	9
Table 12 : pires changements de la variation de la température.....	9
Table 13 : dataframe principal avec les clusters.....	11

### **Tableau des illustrations**

Figure 1 : évolution du logarithme de la déforestation au cours du temps.....	6
Figure 2 : évolution de l'émission de CH <sub>4</sub> au cours du temps.....	7
Figure 3 : les plus grandes émissions de N <sub>2</sub> O au cours du temps.....	8
Figure 4 : évolution de l'émission de CO <sub>2</sub> au cours du temps.....	9
Figure 5 : évolution de la variation des températures au cours du temps.....	10
Figure 6 : dendrogramme.....	11
Figure 7 : répartition des pays par cluster.....	12
Figure 8 : boîtes à moustaches de la déforestation selon les clusters.....	13
Figure 9 : boîtes à moustaches de l'émission de CH <sub>4</sub> selon les clusters.....	14
Figure 10 : boîtes à moustaches de l'émission de N <sub>2</sub> O selon les clusters.....	15
Figure 11 : boîtes à moustaches de l'émission de CO <sub>2</sub> selon les clusters.....	16
Figure 12 : boîtes à moustaches de la variation de température selon les clusters.....	17
Figure 13 : pairplot.....	18
Figure 14 : Éboulis des valeurs propres.....	19
Figure 15 : cercles de corrélations et projection des individus sur les 2 axes.....	20