

exercice

November 24, 2022

1 Sujet 1 : Concentration de CO2 dans l'atmosphère depuis 1958

```
[12]: %matplotlib inline
import os
from datetime import date
import requests
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats
```

1.1 Provenance des données

Les données sont disponibles du site Web de l'[institut Scripps](#). Nous les récupérons sous forme d'un fichier en format CSV dont chaque ligne correspond à une date et lui faire correspondre sa concentration CO2 en micro-mole de CO2 par mole (ppm) rapporté par la norme CIO de 2012 et les données ont été prises tous les midis.

1.2 Analyse des données

On souhaite tout d'abord supprimer toute l'entête du fichier ne contenant pas les données. On pourrait simplement ignorer les premières lignes, mais on a ici décidé d'enlever toutes les lignes commençant par le caractère “” car si le nombre de lignes à supprimer augmente le code ne fonctionnerait plus.

On ajoute aussi en première ligne “day,concentration” pour que pandas puissent les mettre en nom de colonne.

On garde aussi une copie des données utilisées car on ne modifie le fichier que s'il n'est pas présent dans le répertoire courant.

```
[13]: data_path='donnees.csv'

def addName(data_path):
    with open(data_path,"r") as fr:
        lines = fr.readlines()
        with open(data_path,"w") as fw:
```

```

        fw.write("day,concentration\n")
    for line in lines:
        if(line[0]!=''):
            fw.write(line)

if not(os.path.exists(data_path)):
    url='https://scrippsco2.ucsd.edu/assets/data/atmospheric/stations/
    ↳in_situ_co2/weekly/weekly_in_situ_co2_mlo.csv'
    data_path='donnees.csv'
    response = requests.get(url, allow_redirects=True)
    with open(data_path,'wb') as file:
        file.write(response.content)
    addName(data_path)

```

On va ici lire les informations contenues dans le fichier csv et ensuite convertir les dates en dates lisibles par Python. Pour pouvoir ensuite analyser les informations et faire une prédiction dans le futur, nous allons calculer pour chaque année le taux de CO2 moyen et mettre ces données dans deux tableaux distincts.

```

[14]: raw_data = pd.read_csv(data_path,parse_dates=True)

def convert_date(x):
    y = date.fromisoformat(x)
    return y

raw_data['day'] = [convert_date(yw) for yw in raw_data['day']]

nb = 0
somme_conc = 0
tab_year = []
tab_year2 = []
avg_concentration = []

for i in range(0,(len(raw_data['day'])-1)):
    if(raw_data['day'][i].year == raw_data['day'][i+1].year):
        somme_conc += raw_data['concentration'][i]
        nb = nb +1
    else:
        year2 = raw_data['day'][i]
        year = raw_data['day'][i].year
        avg_concentration.append(somme_conc/nb)
        tab_year.append(year)
        tab_year2.append(year2)

raw_data

```

```
[14]:
```

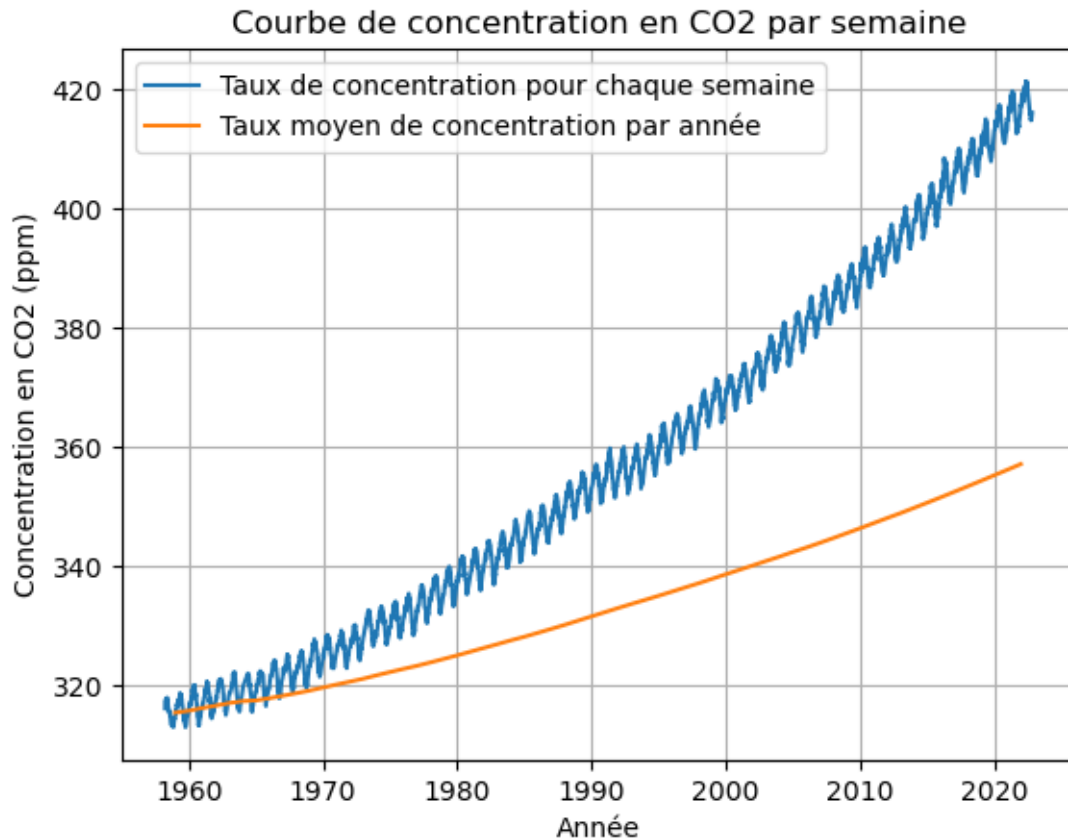
	day	concentration
0	1958-03-29	316.19
1	1958-04-05	317.31
2	1958-04-12	317.69
3	1958-04-19	317.58
4	1958-04-26	316.48
...
3294	2022-10-01	415.12
3295	2022-10-08	414.85
3296	2022-10-15	415.31
3297	2022-10-22	415.60
3298	2022-10-29	416.08

[3299 rows x 2 columns]

1.3 Affichage des données

Affichons désormais le graphe correspondant avec les informations que nous avons récupérés afin de pouvoir comparer l'augmentation avec différents points de vue.

```
[15]: axes = plt.axes()
x = raw_data['day']
y = raw_data['concentration']
axes.grid()
plt.plot(x,y, label="Taux de concentration pour chaque semaine")
plt.plot(tab_year2,avg_concentration, label="Taux moyen de concentration par_
↪année")
plt.title('Courbe de concentration en CO2 par semaine')
plt.xlabel("Année")
plt.ylabel("Concentration en CO2 (ppm)")
plt.legend()
plt.show()
```



1.4 Prédiction

Afin de facilement pouvoir faire une prédiction nous allons utiliser une droite de [régression linéaire](#).

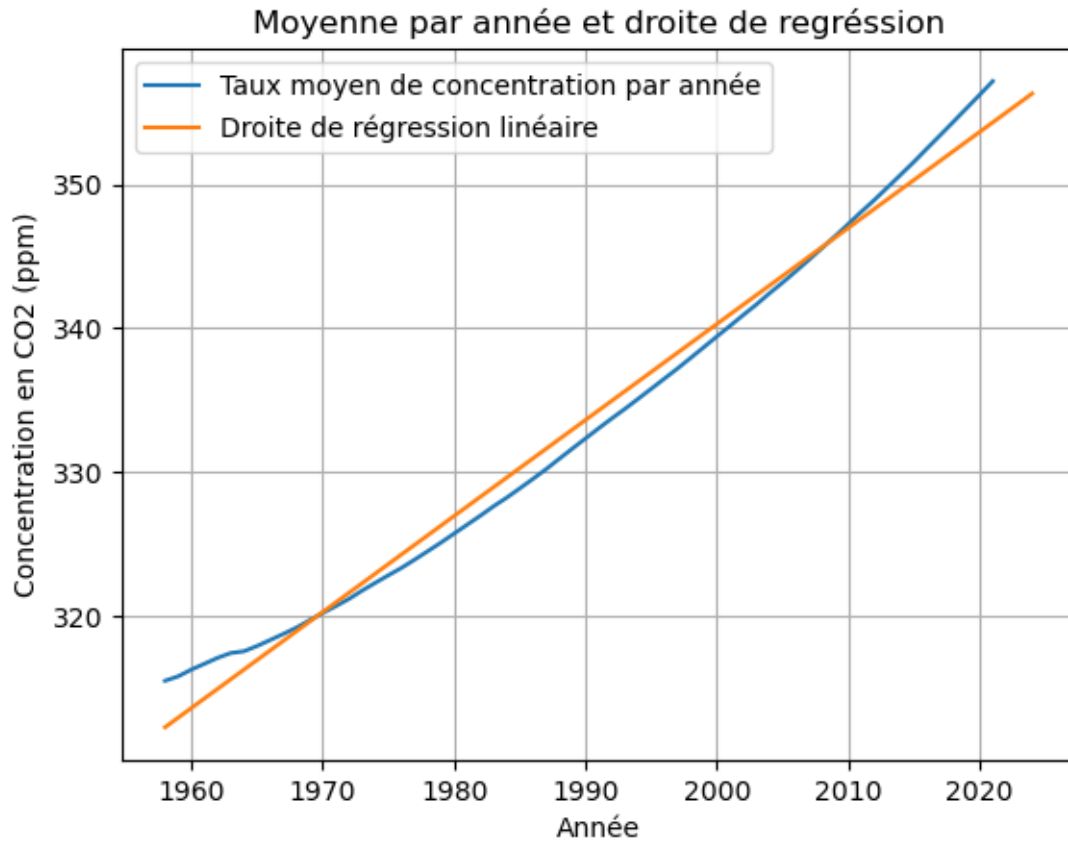
```
[16]: slope, intercept, r_value, p_value, std_err = stats.linregress(tab_year,
    ↪ avg_concentration)

def predict(x):
    return slope * x + intercept
```

Nous affichons maintenant la courbe des moyennes par années et la droite de régression linéaire qui nous permet de voir jusqu'à 2025.

```
[17]: axes = plt.axes()
axes.grid()
x = range(1958,2025)
plt.plot(tab_year,avg_concentration, label="Taux moyen de concentration par
    ↪ année")
plt.plot(x,predict(x), label="Droite de régression linéaire")
```

```
plt.title('Moyenne par année et droite de régression')
plt.xlabel("Année")
plt.ylabel("Concentration en CO2 (ppm)")
plt.legend()
plt.show()
```



Nous pouvons donc observer qu'il y a effectivement une augmentation significative du taux de CO2 chaque année.

```
[18]: print("Concentration moyenne de CO2 en 2025 (prédiction arrondie à l'unité) :
↪",round(predict(2025)),"ppm")
```

Concentration moyenne de CO2 en 2025 (prédiction arrondie à l'unité) : 357 ppm