



La Qualité de l'air en Ile-de-France

Les stations d'observation de la qualité de l'air mesurent les concentrations de polluants atmosphériques et diffusent quotidiennement un indice de qualité de l'air calculée à partir de ces mesures.

Le site [data.gouv.fr](https://www.data.gouv.fr/fr/datasets/indices-qualite-de-l-air-citeair-journaliers-par-polluant-sur-lile-de-france-les-departements-les-communes-franciliennes-et-les-arondissements-parisiens/) met à disposition les indices de qualité de l'air calculés pour chacune des communes de la région IDF (Île de France).

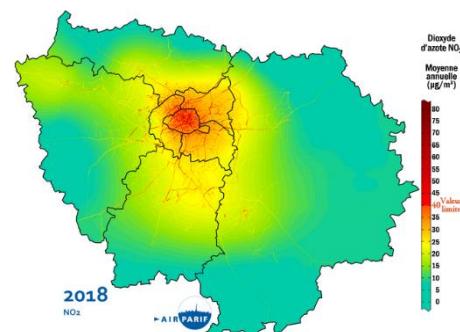
<https://www.data.gouv.fr/fr/datasets/indices-qualite-de-l-air-citeair-journaliers-par-polluant-sur-lile-de-france-les-departements-les-communes-franciliennes-et-les-arondissements-parisiens/> (CTRL+CLIC)

Ce fichier CSV contient près de 40000 lignes et est donc difficilement manipulable avec un tableur.

Il contient cinq séries de données :

- la date : variable « date »
- le code INSEE de la commune ou du département : « ninsee »
- l'indice NO₂ (dioxyde d'azote) : variable « no2 »
- l'indice O₃ (Ozone) : variable « o3 »
- l'indice PM10 (particules fines) : variable « pm10 »

Le programme python AnalyseScolair.py présenté ci-dessous ouvre le fichier « .csv » et affiche, pour la commune de Paris 16 (code INSEE 75116), la moyenne des indices quotidiens de NO₂ : « s_no2 ».



➤ Programme AnalyseScolair.py

Le programme ci-dessous est décrypté en page 3 !

```
import os
f = open('AirIDF2018.csv','r')
entete = f.readline()
s_o3, s_pm10, s_no2, n = 0, 0, 0, 0
ligne = f.readline()
while ligne != "":
    date, ninsee, no2, o3, pm10 = ligne.split(",")
    if ninsee == '75116':
        n += 1
        s_no2 += int(no2)
    ligne = f.readline()
print("Moyenne annuelle indice no2:", s_no2/n)
f.close()
```

Pour calculer avec uniquement les résultats du 16ème arrondissement, ce programme lit ligne à ligne tous les enregistrements, mais ne traite que ceux dont le code INSEE est égal 75116.



LES DONNÉES STRUCTURÉES

1. Copie le dossier complet « **Activité-3** » depuis « Echange/Diffusion/2-seconde/SNT/THEME_4 » vers ton dossier personnel « DOCUMENTS/SNT/THEME_4 ».
2. Lance PYZO depuis « Bureau/Autres raccourcis/SNT/Pyzo_SNT » puis ouvre le programme « Documents/SNT/THEME_4/activité-3/AnalyseScolair.py ».
3. Exécute le programme Python et visualise le résultat fourni.
4. Après avoir consulté le décryptage de ce programme en page 3, modifie le programme pour qu'il affiche aussi les moyennes annuelles des indices de O3 et de PM10. Tu peux aussi t'aider du logigramme de la page 4. **Fais vérifier ton programme par le professeur.**
5. Modifie le programme pour qu'il affiche les informations pour la commune dont le code INSEE est 75118. **Fais vérifier ton programme par le professeur.**
Enregistre ton programme puis ferme-le !



➤ La qualité de l'air dans une commune

L'analyse de données en grand nombre peut nécessiter de filtrer les données. L'objectif de cette activité est de lire ligne à ligne le fichier et de réécrire dans un nouveau fichier uniquement les lignes correspondant à la commune choisie. Pour cela, on dispose des instructions Python suivantes :

```
g= open('Paris75118.csv', 'w')
```

Cette ligne ouvre un nouveau fichier (*g*) en écriture ('w' : write) en lui donnant un nom (*Paris75118.csv*).

```
g.write(ligne)
```

Cette ligne écrit (write) dans le fichier ouvert (*g*) le contenu de la variable (*ligne*)

On devra aussi utiliser la fonction `g.close()` pour fermer le fichier après remplissage de son contenu.

6. Ouvre le programme « **CreaParis75118.py** ».
7. Modifie-le pour lire les données du fichier contenant l'ensemble des indices des communes d'Île de France, et réécrire uniquement les lignes concernant la commune dont le numéro INSEE = 75118.
8. Exécute le programme, puis regarde le dossier en cours pour voir si un nouveau fichier est apparu.
9. Ouvre le nouveau fichier avec un tableur pour vérifier son contenu ; il doit comporter 31 lignes !
10. Vérifie que l'entête (les descripteurs) figure bien en première ligne du nouveau fichier. Sinon, modifie le programme pour qu'il recopie l'entête depuis le fichier initial.

Fais vérifier ton programme par le professeur. Enregistre ton programme puis ferme-le !

➤ Histogramme de pollution

On prend comme source de données le fichier « **Paris75118.csv** ».

Ce fichier est beaucoup plus facile à manipuler car il ne contient que 31 lignes.

Pour interpréter ces données qualitativement, il faut savoir à quoi correspondent les valeurs des indices.

11. Ouvre le programme « **Paris75118-PM10.py** » afin de le modifier.
Ce programme parcourt les indices du fichier « **Paris75118.csv** », en comptant le nombre de jours où l'indice de pollution PM10 a été supérieur à 50.
Fais vérifier ton programme par le professeur. Enregistre ton programme puis ferme-le !
12. Complète ce programme afin de calculer les nombres de jours où l'indice de pollution PM10 a été très faible ou faible.
Fais vérifier ton programme par le professeur.

Niveau de pollution	Très faible	Faible	Moyenne
	0 à 25	26 à 50	51 à 75

Enregistre ton programme puis ferme-le !

Partage ton dossier « Activité-3 » !



➤ Décryptage du programme « AnalyseScolair.py »

```
import os
```

Charge le module "os" (Operating System)

```
f = open('AirIDF2018.csv','r')
```

Ouvre le fichier 'AirIDF2018.csv' en lecture 'r' → read

```
entete = f.readline()
```

Lit la première ligne du fichier et stocke le résultat dans la variable 'entete'

```
s_o3, s_pm10, s_no2, n = 0, 0, 0, 0
```

Déclare et initialise les variables

```
ligne = f.readline()
```

Lit la ligne N°2 et stocke le résultat dans la variable 'ligne'

```
while ligne != "":
```

Répète les instructions qui suivent tant que la variable 'ligne' n'est pas vide

```
date, ninsee, no2, o3, pm10 = ligne.split(",")
```

Stocke chaque donnée contenue dans la variable 'ligne' dans une variable différente sachant que ces données sont séparées par une virgule 'split(",")'

```
if ninsee == '75116':
```

Teste si la variable 'ninsee' contient le code INSEE '75116'

```
n += 1
```

Si c'est le cas, on augmente le nombre de fois que ce code postal a été trouvé. 'n+=1' est équivalent à $n = n+1$.

```
s_no2 += int(no2)
```

puis on augmente la valeur de 's_no2' de la valeur numérique de 'no2'
's_no2' est un mnémonique pour Somme de NO₂

```
ligne = f.readline()
```

Lit la ligne suivante et stocke le résultat dans la variable 'ligne'

```
print("Moyenne annuelle indice no2:", s_no2/n)
```

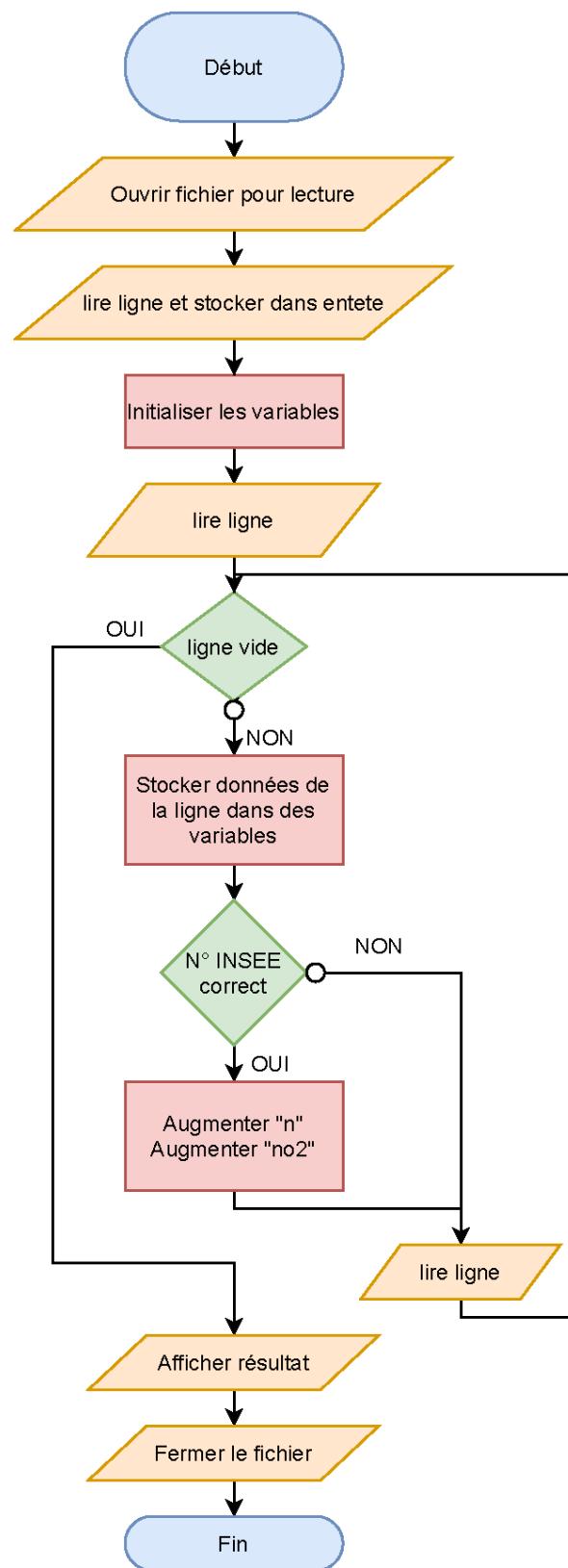
Affiche la moyenne de no2 relevée pendant l'année 2018

```
f.close()
```

Ferme le fichier 'AirIDF2018.csv'



➤ Logigramme du programme « AnalyseScolair.py »





➤ Logigramme du programme « AnalyseParis.py »

