

Code de partage avec Capytale : 645d-9007032

On utilisera les bibliothèques suivantes :

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

### Exercice 1 - PIB et population urbaine

On va étudier ici l'évolution de deux variables de 1960 à 2021 en Norvège : le pourcentage de population urbaine et le PIB par habitant (données issues du World Bank Group), et leur possible corrélation.

Dans un premier temps, il faut importer le jeu de données. On commence donc par exécuter la commande suivante :

```
import pandas as pd
donnees=pd.read_csv('tp11_norvege.csv', delimiter=';')
```

1. Représenter le nuage de points des deux variables étudiées (pourcentage de population urbaine en fonction du PIB par habitant) et commenter le résultat.
2. Que fait le programme suivant ? interpréter le résultat.

```
np.mean([(x[i]-x.mean())*(y[i]-y.mean()) for i in x.index])/(x.std()*y.std())
```

3. Représenter le nuage de points  $(\ln(x), y)$
4. Calculer le coefficient de corrélation linéaire de  $y$  et  $\ln(x)$
5. Déterminer l'équation de la droite de régression de  $y$  en fonction de  $\ln(x)$
6. En déduire qu'on peut supposer que la dépendance entre  $y$  et  $x$  est de la forme  $y = a \ln(x) + b$
7. Représenter le nuage de points initial avec lequel on fera apparaître la courbe d'équation  $y = a \ln(t) + b$