

# PANDAS

## AIDE MÉMOIRE

La bibliothèque *Pandas* est une bibliothèque Python qui permet de manipuler des données tabulaires.

Pour charger la bibliothèque *Pandas* :

```
import pandas as pd # Charger la bibliothèque
```

### Charger les données

```
df = pd.read_csv("fichier.csv") # Charger un fichier CSV
df = pd.read_csv("https://url.com/fichier.csv") # Charger un fichier CSV depuis une URL
```

### Afficher les données

```
df.head() # Affiche les 5 premières lignes
df.tail() # Affiche les 5 dernières lignes

df.shape # Affiche le nombre de lignes et de colonnes
df.columns # Affiche les noms des colonnes
df.info() # Affiche les informations sur les colonnes

df["colonne"] # Affiche la colonne
df[["colonne1", "colonne2"]]] # Affiche les colonnes
```

### Filtrer les données

```
df[df["colonne"] == "valeur"] # Filtre les lignes en fonction d'une valeur
df[df["colonne"].isin(["valeur1", "valeur2"])] # Filtre les lignes en fonction d'une liste de valeurs
df[df["colonne"].str.contains("valeur")] # Filtre les lignes en fonction d'une chaîne de caractères
df[df["colonne"] == df["colonne"].max()] # Filtre les lignes en fonction de la valeur maximale d'une colonne
df[df["colonne"] == df["colonne"].min()] # Filtre les lignes en fonction de la valeur minimale d'une colonne
```

### Manipuler les données

```
df["nouvelle_colonne"] = df["colonne1"] + df["colonne2"] # Additionne deux colonnes existantes
df["nouvelle_colonne"] = df["colonne1"] * 2 # Multiplie une colonne par un nombre
df["nouvelle_colonne"] = df["colonne1"].str.replace("valeur1", "valeur2") # Remplace une valeur par une autre
```