

Les données structurées - Exercice 1

Création d'un fichier CSV

Vous allez créer un fichier csv nommé films.csv avec une dizaine de films en utilisant les descripteurs suivants :

id	titre	annee	longueur	action	animation	comedie	drame	documentaire	romance	science-fiction	fantastique
1	un acte d'amour	1953	108	0	0	0	1	0	1	0	0
2	un air de famille	1996	110	0	0	1	0	0	0	0	0
3	un air si pur...	1997	106	0	0	1	1	0	0	0	0
4	un aller simple	1971	118	0	0	0	1	0	0	0	0

MÉTHODE :

On le crée dans libre office calc

On l'enregistre au format csv avec la virgule comme séparateur

Python et les fichiers csv

Nous travaillerons avec Spyder ou EduPython

Écrire ce script et enregistrez-le dans un dossier qui contient votre fichier films.csv

```
import csv
liste=[]
with open("films.csv",'r') as data:
    lecteur=csv.reader(data,delimiter=",")
    for ligne in lecteur:
        liste.append(ligne)
print(liste)
```

? EXERCICE 1 :

1. Quelle est la structure de la variable liste?
2. Que contient : liste[0] ?
3. Peut-on faire la même opération avec un fichier dont le séparateur serait un point-virgule?
4. Comment faire afficher la 3^{ème} ligne du fichier?
5. Comment faire afficher le titre et l'année du film de la 3^{ème} ligne du fichier?
6. Comment sait-on dans quel genre est répertorié un film?

Voici un script qui permet d'afficher les titres de films dont l'année de sortie est supérieure à 2000

```
# chargement de la bibliothèque
import csv
liste=[] #liste vide
#chargement du fichier csv dans une liste
with open("films.csv", 'r') as data:
    lecteur=csv.reader(data,delimiter=",")
    for ligne in lecteur:
        liste.append(ligne)
    print(liste[0]) #affichage de la première ligne
#suppression de la première ligne (sinon la boucle plante...)
liste.pop(0)
# affiche les films produits après 2000
for ligne in liste:
    if int(ligne[2])>2000:
        print(ligne[1])
```

? EXERCICE 2 :

↳ Inspirez-vous de ce script pour faire afficher les films par genre

Avec une liste de dictionnaires

Ce script charge une liste de dictionnaires dans la variable liste

```
import csv
liste=[] #liste vide
#chargement du fichier csv dans une liste de dictionnaires
with open("films.csv", 'r') as data:
    lecteur=csv.DictReader(data,delimiter=",")
    liste=[dict(row) for row in lecteur]

print(liste[0]) #affichage de la première ligne
```

? EXERCICE 3 :

1. Quelles sont les clés de ces dictionnaires?
2. Comment atteindre le titre du film dans le 3^{ème} dictionnaire?
3. Quel code faut-il écrire pour vérifier si un film est présent dans la liste?
4. Quel code faut-il écrire pour faire afficher tous les films dont la durée est supérieure à 2h?

Pour aller plus loin

Créer un second fichier `realisateur.csv`, contenant les réalisateurs des films du fichiers `films.csv`. Ces deux fichiers devront avoir un descripteur commun (le titre). Imaginez et réalisez un programme qui affiche pour une année donnée les films et leurs réalisateurs.