

# Traitement des données en tables

qkzk

2021/05/23

## 1. Reconnaître un fichier csv

nom	prenom	date_naissance
Dupont	Pierre	17/05/1987
Dupond	Catherine	18/07/1981
Haddock	Archibald	23/04/1998

1. Quels sont les descripteurs de ce tableau ?

Nom,Pays,Temps  
Camille Muffat,France,241.45  
FranckyVincent,France,222,30  
Geronomo,Apache, France,212.12

2. Rectifiez les erreurs dans le fichier csv ci-dessus.

## QCM

1. Qu'est-ce que le format de fichier CSV ?
  - A. un format de fichier mis au point par Microsoft pour Excel
  - B. un format de fichier pour décrire une base de données
  - C. un format de fichier où les données sont séparées par un caractère tel qu'une virgule
  - D. un format de fichier décrivant une page Web
2. Laquelle de ces affirmations est vraie ?
  - A. on ne peut accéder au contenu d'un fichier CSV que par l'intermédiaire d'un programme Python
  - B. CSV est un format de chiffrement des données
  - C. le format CSV a été conçu pour assurer la confidentialité d'une partie du code d'un programme
  - D. les fichiers CSV sont composés de données séparées par des caractères comme des virgules
3. Dans la plupart des fichiers CSV, que contient la première ligne ?
  - A. des notes concernant la table de données
  - B. les sources des données
  - C. les descripteurs des champs de la table de données
  - D. l'auteur de la table de données

## 1. Sélection et projections depuis une table donnée

On considère la table suivante :

```
repertoire = [
    {"nom": "Ducobu", "prenom": "Raoul", "tel": "0612345678"},
    {"nom": "Dreoul", "prenom": "Raymond", "tel": "0734567812"},
    {"nom": "Martin", "prenom": "Jacqueline", "tel": "0612341234"},
    {"nom": "Mersin", "prenom": "Tania", "tel": "0788990011"},
]
```

1. Donner la valeur de la liste construite dans chacune des sélections suivantes :

a.

```
liste = []
for enregistrement in repertoire:
    if enregistrement["nom"] == "Ducobu":
        liste.append(enregistrement)
```

b.

```
liste = [enre for enre in repertoire if "12" in enre["tel"]]
```

2. Construire l'enregistrement des personnes dont le nom comporte un "r".

3. Donner le résultat de la projection suivante :

```
proj = [enregistrement["prenom"] for enregistrement in repertoire]
```

4. Constuire la projection des téléphones du répertoire.

## 2. Importer et manipuler une table

On dispose du fichier « info.csv » donné ci-dessous :

```
nom, prenom, naissance, deces
lovelace, ada, 1815, 1852
von neumann, john, 1903, 1957
turing, alan, 1912, 1954
mccarthy, john, 1927, 2011
floyd, robert, 1935, 2001
```

Le programme ci-dessous nous permet de créer un tableau à partir de ce fichier.

```
file = open("info.csv", "r")
firstLine = file.readline()      # chargement de la ligne d'entête
tableau = [line.split(',') for line in file] # chargement des données
```

Détaillons ce programme :

1. On charge le fichier dans la variable `file`
2. on lit la première ligne, et on avance ainsi d'une ligne dans le fichier.
3. On crée un tableau à deux dimensions. Chaque ligne du tableau correspond à une ligne de valeurs du fichier. La première ligne est `["lovelace", "ada", "1815", "1852"]`.

Techniquement : `line` est un `str` et la méthode `.split(',')` va séparer cette chaîne en une liste de `str` autour de chaque virgule.

Remarquons que tous les éléments des enregistrements sont des `str` !

## Questions

1. Combien d'éléments comporte la liste `tableau` ?
2. Évaluez `tableau[2]`.
3. Évaluez `tableau[1][3]`
4. Comment accéder à l'année de décès de John McCarthy ?
5. Modifier l'année de naissance de Robert Floyd qui est né en 1936.
6. Réaliser la sélection des informaticiens décédés avant 1952 :
  - a. Directement en repérant les lignes intéressantes,
  - b. À l'aide d'une boucle et d'un test adaptable à un fichier beaucoup plus rempli. On prendra garde au *type* des données dans `tableau` !

## 3. Importer une table

On considère le fichier csv `films.csv` dont voici un extrait.

films

Pour chaque question on donnera une syntaxe Python permettant d'obtenir la réponse ainsi que la valeur de celle-ci lorsque l'information est connue.

1. Importer le fichier en Python vers un tableau de dictionnaires appelé `table`
2. Quel est le nombre d'enregistrements ?
3. Combien de champs différents comporte cette table ?
4. Quelles sont les différents champs ?

## 4. Exporter une table, sélections, projections

Les joueurs d'échecs sont classés selon les résultats de leurs parties. Chaque partie permet de gagner ou perdre des points élo et les joueurs sont classés selon leur score.

On a crée en Python une table dans une variable appelée `joueur` dont voici un extrait :

```
joueurs = [  
    {'nom': 'Carlsen, Magnus', 'pays': 'nor', 'elo': 2863},  
    {'nom': 'Caruana, Fabiano', 'pays': 'usa', 'elo': 2835},  
    {'nom': 'Ding, Liren', 'pays': 'chi', 'elo': 2791},  
    {'nom': 'Nepomniachtchi, Ian', 'pays': 'rus', 'elo': 2784},  
    {'nom': 'Vachier-Lagrave, Maxime', 'pays': 'fra', 'elo': 2778},  
    {'nom': 'Grischuk, Alexander', 'pays': 'rus', 'elo': 2777},  
    {'nom': 'Aronian, Levon', 'pays': 'arm', 'elo': 2773},  
    {'nom': 'So, Wesley', 'pays': 'usa', 'elo': 2770},  
    {'nom': 'Radjabov, Teimour', 'pays': 'aze', 'elo': 2765},  
    {'nom': 'Giri, Anish', 'pays': 'nl', 'elo': 2764},  
    ...  
]
```

## Enregistrer

1. Enregistrer cette table au format CSV.

## Parcourir les données

2. Comment accéder aux différents champs ?
3. Comment connaître le nombre d'enregistrements ?

## Sélections

2. Comment sont enregistrées les nationalités ?
3. Créer la table des joueurs français.
4. Créer la table des joueurs ayant plus de 2780 points élo.
5. Compter les joueurs de la table ayant entre 2700 et 2800 points élo.

## Projections

6. Créer la projection des noms de joueurs. Limiter ensuite cette table au top 5.
7. Créer la projection des scores élo. Calculer le score élo moyen des 10 premiers joueurs.
8. Créer une fonction qui prend la table en paramètre ainsi qu'un entier **nombre** et retourne le score moyen des “**nombre**” premiers joueurs.  
Si **nombre** est trop grand on se limitera à la taille du tableau.