

Spécialité
Numérique et sciences informatiques
Classe de première
Feuille de révision pour le contrôle n° 2
sur le traitement des données en table
du mardi 8 mars 2022

Savoir et savoir-faire

- Revoir les techniques fondamentales de manipulation des tableaux et les dictionnaires en Python : chapitres C1, C2 et C3.

E1 Indexation de table

- Notions associées à la représentation de données en table :
 - * nom et type des attributs (les colonnes) ;
 - * ligne d'une table ;
 - * valeur d'un attribut dans une ligne.
- Nature et organisation d'un fichier CSV
 - * caractère de séparation EOL (« end of line ») et EOF (« end of file ») ;
 - * information délivrée par la première ligne ;
 - * information délivrée par les lignes suivantes.
- Structures de données pouvant accueillir l'information contenue dans un fichier csv :
 - * tableau de tableaux ;
 - * tableau de dictionnaires.
- Opérations d'ouverture et d'écriture d'un fichier en Python :
 - * que signifie « ouvrir » et « fermer » un fichier dans un programme ?
 - * signification de l'instruction `F = open(...)` et de ses différents arguments : notamment de l'argument indiquant mode d'ouverture `'r'`, `'w'` ou `'a'` (on se limitera à l'ouverture d'un fichier texte)
 - * l'instruction `F.close()` ;
 - * l'instruction de lecture `F.read()`.
- Quelques notions sur le module `csv` et le résultat de l'instruction
`contenu = csv.DictReader(F)`

E2 Recherche dans un ensemble de données tabulées

On suppose que l'on dispose d'un tableau de dictionnaires représentant le contenu d'un fichier `csv`.

- Tester la présence d'une ligne vérifiant une certaine condition.
- Compter le nombre de lignes vérifiant un certaine condition.
- Sélection récupérer toutes les lignes vérifiant un certaine condition pour former un - tableau de dictionnaires extrait du premier.
- Projection : former un tableau de dictionnaires qui ne contient qu'une partie des attributs du premier dictionnaire.
- Combiner une sélection et une projection.

E3 Opérations complémentaires sur des données tabulées

- Trier un tableau de dictionnaires selon un certain attribut (une certaine colonne) à l'aide de la méthode `sort()` et de l'argument `key = . . .` par ordre croissant ou décroissant.
- Trier selon deux critères hiérarchisés.
- Notion de contraintes d'intégrité et programmation d'un test de cohérence.
- Recherche et élimination des doublons.
- Fusion de deux tables .

Pour se préparer

Revoir et bien comprendre le code des différents TD sur le traitement des données en table :

- sur l'indexation du fichier `commande.csv` ;
- sur la recherche de données dans `commande.csv` ;
- sur le tri à partir du fichier `pays.csv` ;
- sur la recherche et l'élimination de doublons avec le fichier `doublons.py` ;
- sur la fusion des fichiers `prenomsRennes2015.csv` et `prenomsRennes2016.csv`.