

## SAÉ : Gestion de fichiers

### ➤ Contexte et jeu de données

Jeu de données à utiliser : `comorbidites.json` (disponible sur Updago).

La structure du fichier Json est visualisable sur le net par simple copie/coller sur l'outil <http://jsonviewer.stack.hu/> ou <https://jsoneditoronline.org/>

Parmi les patients pris en charge pour une pathologie donnée (ou traitement chronique ou épisode de soins), certains peuvent également être pris en charge pour une ou plusieurs autres pathologies, appelées comorbidités. Ce jeu de données informe sur les comorbidités associées aux pathologies (ou traitements chroniques ou épisodes de soins) repérées par la Caisse nationale de l'Assurance Maladie (CNAM) en France. À noter : il n'est pas intéressant de sommer les fréquences ou les effectifs des comorbidités, un même patient pouvant être pris en charge pour plusieurs comorbidités.

### Description du fichier

- *annee*
- *patho\_niv1*: groupe de pathologies (ou traitements chroniques ou épisodes de soins)
- *patho\_niv2*: sous-groupe de pathologies (ou traitements chroniques ou épisodes de soins)
- *patho\_niv3*: sous-sous-groupe détaillé de pathologies (ou traitements chroniques ou épisodes de soins). Information la plus pertinente !
- *top*: libellé technique de la pathologie (ou traitement chronique ou épisode de soins)
- *comorbidite*: libellé technique de la comorbidité associée
- *libelle\_comorbidite*: libelle\_comorbidite[text] comorbidité associée
- *region*: 99 : France (niveau national)
- *dept*: 999 : France (niveau national)
- *Ncomorb*: effectif de patients pris en charge pour la comorbidité associée
- *Ntop*: effectif de patients pris en charge pour la pathologie (ou traitement chronique ou épisode de soins) dont il est question.
- *Proportion\_comorb*: proportion\_comorb[double] effectif de patients pris en charge pour la comorbidité identifiée rapporté à l'effectif de patients pris en charge pour la pathologie dont il est question (Ncomorb / Ntop)
- *patho\_niv3\_comorb*: même raisonnement que pour la pathologie
- *patho\_niv2\_comorb*:
- *patho\_niv1\_comorb*:
- *Niveau prioritaire*:

### ➤ Travail à faire

L'objectif de cette Saé est de transformer le fichier Json fourni pour obtenir un fichier csv avec les colonnes suivantes et **uniquement pour les lignes où ces colonnes sont complètes et sans données aberrantes** :

PathologieLaPlusPrécise ; NomComorbidité ; Année ; EffectifPathologie ;  
EffectifComorbidité ; Proportion ; niveauPrioritaire

**Remarques :**

- Le séparateur du fichier csv sera le point-virgule ;
- Attention, aux accents, le fichier csv devra être lisible en l'ouvrant dans Excel ;
- Le fichier devra être trié sur les proportions, de la plus élevée à la plus petite. Le 2<sup>ème</sup> critère de tri sera sur les pathologies, puis sur la comorbidité et enfin sur l'année.

Vous devrez donc réaliser un script Python `transformerJsonToCsv.py` permettant de réaliser cette transformation.

➤ **Restitution**

Le ou les fichiers Python commentés sont à rendre sur Updago avant le **dimanche 10 décembre 2023 (23h59 dernier délai)**

➤ **Evaluation**

Elle portera sur la qualité technique de la production : opérationnalité, optimisation, bonnes pratiques (*commentaires, indentation, découpage modulaire si besoin, ...*) et ergonomie.

➤ **Apprentissages critiques concernés par cette Saé** (*à mentionner dans votre portfolio du S2*)

- Correctement interpréter et prendre en compte le besoin du commanditaire ou du client
- Respecter les formalismes de notation
- Connaître la syntaxe des langages et savoir l'utiliser
- Mesurer l'importance de maîtriser la structure des données à exploiter
- Comprendre les structures algorithmiques de base et leur contexte d'usage
- Prendre conscience de l'intérêt de la programmation