

Activité les données structurées CSV et JSON

Les applications utilisées quotidiennement manipulent des données qui sont structurées pour être exploitées efficacement.

Le format de fichier CSV

Le format **csv** (Comma-separated values) est un format simple d'**échange de données numérique** au format texte entre des applications.

- la **première ligne** contient le **descripteur** de chaque donnée, séparée par un **point-virgule**,
- chaque **ligne suivante** représente un **objet** avec les **valeurs** de chaque descripteur séparées par un **point-virgule**;
- l'ensemble de ces données constituent une **collection de données**.

Le contenu du fichier `eleves.csv` au format `csv`

```
nom;prenom;date_naissance
Durand,Jean-Pierre;23/05/2004
Dupont,Christophe;15/12/2003
Terta;Henry;12/06/2002
```

- **Question 1** : Donne la liste des descripteurs de cette collection de données. * **Question 2** : Les données enregistrées concernent combien d'élèves ? * **Question 3** : Donne les différentes valeurs du descripteur "date_naissance". </WRAP> D'autres séparateurs peuvent être utilisés comme la virgule à la place du point-virgule. Les tableurs (Calc - Libre Office; Excel - Microsoft), sont capables de lire les fichiers au format CSV. ===== Je modifie le fichier `eleves.csv` ===== * Je télécharge le fichier `eleve.csv` * puis je l'ouvre avec un tableur ; Je visualise une table avec des colonnes de données.

Question 4 :

- J'ajoute un **descripteur** pour pouvoir ajouter à chaque élève sa **taille**,
- **j'ajoute** une ligne avec mon **nom**, **mon prénom**, **ma date de naissance** et **ma taille** puis je le **sauvegarde**;
- **j'ouvre** le fichier sauvegardé avec le logiciel de **traitement de texte** pour **visualiser** la ligne rajoutée et l'utilisation du **point-virgule (;)** comme séparateur.
- **j'écris** la ligne obtenue.

===== Le format de fichier JSON =====

Le format de fichier **JSON** (JavaScript Object Notation) permet de représenter des **données plus complexes** que des tables et est souvent utilisé pour échanger des données sur le **Web** et avec les **apps des mobiles**.

* une **collection** de données est **délimitée par les crochets [et]**, * chaque objet est délimité par les accolades **{ et }**, * les **descripteurs** et les **valeurs** sont toujours entourés de **guillemets** et séparés par des **deux-points**.
 ===== Le contenu du fichier `eleves` au format JSON ===== <code> [{"nom": "Durand", "prenom": "Jean-Pierre", "datenaissance": "23/05/2004"}, {"nom": "Dupont", "prenom": "Christophe", "datenaissance": "15/12/2003"}, {"nom": "Terta", "prenom": "Henry", "date_naissance": "12/06/2002"}]</code>

Question 5 :

- En m'aidant de l'exemple du contenu du fichier `eleve.json`, **j'écris** un nouvel objet de données avec mon **nom**, **mon prénom**, **ma date de naissance** et **ma taille**

===== Les métadonnées =====

Les **métadonnées** ou données à propos des données, donnent des informations sur les données. Les formats de fichiers permettent d'enregistrer des métadonnées comme la date de création du fichier.

* je consulte consulte les **propriétés** du fichier **elevés.csv**.

Question 6 :

- Quels sont les **métadonnées disponibles** sur le fichier **elevés.csv** ?

==== Les données ouvertes (Open Data) ==== Il est possible de trouver sur Internet des données dans de nombreux domaines. Ces données sont soit payantes soit en accès libre. L'**ouverture des données publiques sur le web** a été décidée en 2009 aux Etats-Unis par le président Obama dans le cadre de l'initiative **DATA.GOV**. En **Europe**, c'est la **directive de 2013** qui encadre le processus. La France diffuse des données publiques sur le site www.data.gouv.fr Les données ouvertes sont utiles pour les chercheurs, les scientifiques mais aussi pour les citoyens pour que chacun puisse les analyser et les utiliser selon ses besoins. ==== Utiliser les données de la SNCF ==== * J'accède au site de la SNCF www.sncf.com , * je recherche **Open Data** puis je clique sur le lien **Explorer les données** , * je recherche ensuite le jeu de données **Liste des gares** (Liste des gares du Réseau Ferré National). * je consulte les métadonnées dans l'**onglet Informations**.

Question 7 :

- Quels sont les informations sur le producteur de ces données, la licence d'utilisation et la date de dernière modification ?

* Je clique ensuite sur l'onglet **Export**.

Question 8 :

- Quels sont les différents **formats de fichiers plats** proposés ?

* Je télécharge le **jeu de données entier** au format **CSV** et je l'ouvre dans un tableur. * j'utilise les fonctionnalités de **tri**, de **recherche** ou de **filtre** du tableur pour répondre aux questions suivantes :

Question 9 :

- Quelles sont les **gares SNCF** de la ville de **Limoges** ?
- Quelle est la **localisation GPS** de la gare de **Limoges-Puy-Imbert** ?

==== Activité à faire à la maison ====

Je **consulte** la vidéo pour répondre aux questions suivantes :

- Quelles sont les informations qui ne peuvent pas être rendues publiques ?

https://youtu.be/aHxv_2BMjfw



Video

==== Je continue ... ====

Je reviens à l'accueil SNT du thème [Les données structurées et leur traitement](#)

===== Les données structurées =====

From:

/ - **Les cours du BTS SIO**

Permanent link:

[/doku.php/snt/donnee/activite4?rev=1568272453](#)

Last update: **2019/09/12 09:14**

