

Nom, prénom : .....

date : .....

5

## Fiche activités SNT Données structurées – Unité 2-4

*Institut Florimont*

Ce document se base sur le travail de D. Roche. Il est publié sous licence libre Creative Commons



### Exercice 1 — Format de données ouvert (1 pt)

/1

On trouve énormément de données sur internet. Une partie de ces données sont publiques, par exemple le site <https://opendata.swiss/fr/> recense un grand nombre de données publiques. Ces données sont librement réutilisables.



### À vous de jouer !

Afin de découvrir ce qu'est "l'open data", allez sur le site <https://opendata.swiss/fr/>. Cliquez sur "En savoir plus sur opendata.swiss".

1. Résumez en quelques lignes ce que vous avez appris en lisant cette page.

/0.5

.....  
.....  
.....

2. En utilisant votre livre p. 90, indiquez l'intérêt d'un format de données ouvert (ou libre).

/0.5

.....  
.....  
.....

### Exercice 2 — Le format de fichier CSV (1 pt)

/1

Le format *CSV* est très courant pour stocker des données, nous allons l'étudier en premier.

Voici ce que nous dit Wikipédia sur le format CSV :

*Comma-separated values, connu sous le sigle CSV, est un format informatique ouvert représentant des données tabulaires sous forme de valeurs séparées par des virgules.*

*Un fichier CSV est un fichier texte, par opposition aux formats dits « binaires ». Chaque ligne du texte correspond à une ligne du tableau et les virgules correspondent aux séparations entre les colonnes. Les portions de texte séparées par une virgule correspondent ainsi aux contenus des cellules du tableau.*

Voici un exemple du contenu d'un fichier CSV :

```
nom,prenom,date_naissance
Durand,Jean-Pierre,23/05/1985
Dupont,Christophe,15/12/1967
Terta,Henry,12/06/1978
```

Je pense qu'il est évident pour vous que nous avons ici 3 personnes :

- Jean-Pierre Durand qui est né le 23/05/1985
- Christophe Dupont qui est né le 15/12/1967
- Henry Terta qui est né le 12/06/1978

"nom", "prenom" et "date\_naissance" sont appelés des *descripteurs* alors que, par exemple, "Durand", "Dupont" et "Terta" sont les valeurs du descripteur "nom".



### À vous de jouer !

Pour mieux comprendre ce format de fichier très utilisé, nous allons en manipuler un :

1. Recherchez sur le site <https://opendata.swiss/fr/> les données portant sur les Résultats de la votation populaire du 12 février 2017 à Genève.
2. Téléchargez le fichier .csv en cliquant sur Aller à la ressource.
3. Ouvrez le fichier à l'aide du tableur de LibreOffice.



### Attention

Au moment de l'ouverture du fichier, le tableur demande de préciser le séparateur utilisé dans le fichier. Dans notre cas, il faut choisir le **point-virgule (;)**. Afin de ne pas se tromper LibreOffice permet de prévisualiser le résultat en fonction du séparateur choisi.

4. Quel est le rôle du séparateur dans le fichier CSV ?

/0.5

.....

.....

Vous pouvez constater que les données sont bien "rangées" dans un tableau avec des lignes et des colonnes (voilà pourquoi on parle de données tabulaires).

5. En utilisant le fichier téléchargé, trouvez le pourcentage de réponse positive à la question 1 dans la ville du Petit-Lancy ?

/0.5

.....

.....

**Exercice 3 — Le format JSON (3 pts)**

/3

Autre format de données très courant sur le "web", le JSON (JavaScript Object Notation). Le JSON fonctionne avec un système de paire *clé/valeur*.

Un "objet" est encadré par des accolades :

```
1 {cle_1 : val_1, cle_2 : val_2, cle_3 : val_3}
```

souvent pour une question de lisibilité on écrira :

```
1 {  
2   cle_1 : val_1,  
3   cle_2 : val_2,  
4   cle_3 : val_3  
5 }
```

Un fichier au format JSON peut regrouper un grand nombre d'objets :

```
1 [{  
2   "nom" : "Durand",  
3   "prenom" : "Jean-Pierre",  
4   "date_naissance" : "23/05/1985"  
5 },  
6 {  
7   "nom" : "Dupont",  
8   "prenom" : "Christophe",  
9   "date_naissance" : "15/12/1967"  
10 },  
11 {  
12   "nom" : "Terta",  
13   "prenom" : "Henry",  
14   "date_naissance" : "12/06/1978"  
15 }]
```

Ci-dessus, nous avons une liste (délimité par [ ]) contenant 3 objets (délimité par { }).

La "valeur" d'une paire "clé/valeur" peut être une liste :

```
1 {  
2   "nom" : "Durand",  
3   "prenom" : "Jean-Pierre",  
4   "date_naissance" : "23/05/1985"  
5   "sport" : ["tennis", "football", "pétanque"]    # <- une liste de sports  
6 }
```

ou même un autre objet :

```

1 {
2   "nom" : "Durand",
3   "prenom" : "Jean-Pierre",
4   "date_naissance" : "23/05/1985"
5   "adresse" : {"num":6, "rue":"impasse du rossignol",
6     ↪  "ville":"Nogent-le-Rotrou", "cp":"28400"} # <- un autre objet
7 }

```

Comme vous pouvez le constater, il est possible d'obtenir des structures de données très complexes avec le format JSON.

De nombreux sites web proposent des services basés sur des API (Application Programming Interface). Ces sites sont capables de fournir des données aux formats JSON sur "simple demande". Souvent, ces "demandes" sont effectuées par l'intermédiaire d'une *URL* (si vous avez besoin de vous rafraîchir la mémoire, jeter un coup d'oeil au cours sur le Web vu précédemment).

Nous allons illustrer ce propos en utilisant l'API d'un site qui fournit des informations météo au format JSON. Vous trouverez ce site à l'adresse suivante : [openweathermap.org/api](https://openweathermap.org/api)

Pour profiter de ce service, il est nécessaire d'obtenir une clé (API key). Cette clé vous sera fournie par votre enseignant.



À vous de jouer !

1. Ouvrez votre navigateur préféré et copiez-collez l'url suivante dans la barre d'adresse du navigateur :

<http://api.openweathermap.org/data/2.5/weather?q=Sydney,AU&APPID=XXXXXXXXXXXXXXXXXXXX>



**Attention**

il faut remplacer les "X" par la clé (API key) qui vous est donnée sur la page Moodle du cours

Au lieu d'obtenir, comme d'habitude, une page web, vous devriez obtenir uniquement quelque chose qui ressemblera à ceci :

```

1 {"coord":{"lon":151.21,"lat":-33.87},"weather":[{"id":803,
2   "main":"Clouds","description":"broken clouds","icon":"04n"}]},
3 "base":"stations","main":{"temp":293.97,"feels_like":291.58,"temp_min":293.15,
4   "temp_max":295.15,"pressure":1015,"humidity":64},"visibility":10000,
5   "wind":{"speed":5.1,"deg":170},"clouds":{"all":75},"dt":1578735986,
6   "sys":{"type":1,"id":9600,"country":"AU","sunrise":1578682510,
7   "sunset":1578733782},"timezone":39600,"id":2147714,"name":"Sydney","cod":200}

```

Vous l'aurez compris, vous avez récupéré les données météo actuelle d'un endroit dans le monde au format JSON. Comme les données sont réactualisées relativement souvent, il est normal que vous n'obteniez pas exactement la même chose que moi !



À vous de jouer !

1. De quel endroit, avez-vous récupéré les données météo ? /0.5  
.....
2. Dans l'*URL*, de la question précédente, encadrer les 2 informations qui permettent de préciser le lieu (Ville + Pays). /0.5
3. Chercher maintenant les données météo actuelles de Genève.
4. Quels sont les coordonnées géographiques (longitude et latitude) de la ville ? /0.5  
Longitude : .....  
Latitude : .....
5. Quel est la valeur actuelle de la pression atmosphérique ? /0.5  
*Pour information l'unité de la valeur renvoyée est l'hectopascal (hpa).*  
.....
6. Citer 4 autres informations renvoyées par la requête. /1  
.....  
.....