

# VOIE GÉNÉRALE ET TECHNOLOGIQUE

2<sup>DE</sup>

1<sup>RE</sup>

1<sup>LE</sup>

## Sciences numériques et technologie

ENSEIGNEMENT  
COMMUN

# LES DONNÉES STRUCTURÉES ET LEUR TRAITEMENT - DONNÉES DES SALLES DE CINÉMA EN ÎLE-DE-FRANCE

## Contenus et capacités

Contenus	Capacités attendues
Données structurées	Identifier les différents descripteurs d'un objet. Distinguer la valeur d'une donnée de son descripteur.
Traitement de données structurées	Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.

## Activité

- Explorer les données d'un fichier CSV à l'aide d'opérations de tri et de filtre.
- Effectuer des calculs sur ces données.
- Réaliser une visualisation graphique des données.

## Note d'intention

Cette activité peut être découpée en plusieurs parties et proposée à différents moments de l'année scolaire puisqu'elle est transversale à deux thématiques (les données structurées et leur traitement/géolocalisation, cartographie et mobilité).

La première partie consiste à importer des données ouvertes. Il est possible de trouver facilement des données autres que celles des salles de cinéma d'Île de France.

La deuxième partie, permet de se questionner sur le format csv en l'ouvrant avec un logiciel de traitement de texte et non un tableur.

La troisième partie, permet de traiter les données avec un tableur sous la forme de questions.

La quatrième partie, permet de traiter les données en écrivant des programmes en Python. L'utilisation de la fonction print n'a qu'un but pédagogique pour permettre aux élèves de visualiser et comprendre les premiers programmes. Les programmes sont soit donnés et à étudier, soit à compléter.

Retrouvez éduscol sur :



## Les salles de cinéma en Île-de-France

### Partie A - importer des données ouvertes

1. À partir du site [data.gouv.fr](http://data.gouv.fr), télécharger des données ouvertes sur les salles de cinéma en Île-de-France. Nommer ce fichier « salles.csv ».

#### Les salles de cinéma en Île-de-France

Ce jeu de données provient d'un service public certifié

[WIKI-MODIFIER](#)

Liste des cinémas franciliens et leur adresse géocodée (WGS84)

(Source Centre national du cinéma et de l'image animée)

#### Ressources

##### Fichier CSV

CSV   Disponible

[TÉLÉCHARGER](#)

2. Indiquer quand a été créé ce fichier. (il est possible de cliquer sur Fichier csv)

### Partie B - fichier csv ?

1. Ouvrir le fichier « [salles.csv](#) » avec notepad++ ou notepad.
2. Décrire les 3 premières lignes.

#### Commentaires

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

Un fichier CSV est un fichier texte.

### Partie C - traitement des données avec un tableur

1. Dans un tableur, importer les données. Le séparateur utilisé dans cette table est le point-virgule.

#### Commentaires

Avec un tableur, les données sont représentées sous forme tabulaire.

2. Identifier les descripteurs de cette table de données.
3. Stella habite à Aubergenville et souhaite connaître le nombre de salles de cinéma de son département. Proposer la formule à écrire en F313 pour répondre à Stella.
4. Stella souhaite connaître le nombre de salle de cinéma à Paris. Proposer la formule à écrire en H312.

Retrouvez éducol sur :



## Partie D - traitement automatisé des données

Il est souhaitable de créer un dossier dans lequel on trouvera le fichier salles.csv sans les formules de la partie C et les différents programmes Python.

1. Écrire le programme suivant :

```
3 f = open("Salles.csv", "r")
4 Descripteurs = f.readline()
5 print(Descripteurs)
6 Ligne = f.readline()
7 print(Ligne)
```

- 1.1. Expliquer les deux lignes qui s'affichent lors de l'exécution du programme.

2. Écrire le programme suivant :

```
3 f = open("Salles.csv", "r")
4 Descripteurs = f.readline()
5 for i in range(10):
6     ligne = f.readline()
7     print(ligne)
```

- 2.1. Expliquer la ligne 5.
- 2.2. Expliquer l'affichage lors de l'exécution du programme.

3. Écrire le programme suivant :

```
3 f = open("Salles.csv", "r")
4 Descripteurs = f.readline()
5 for i in range(10):
6     ligne = f.readline()
7     donnees = ligne.split(";")
8     print(donnees[2])
```

- 3.1. Expliquer les lignes qui s'affichent lors de l'exécution du programme.
- 3.2. Modifier le programme précédent afin d'afficher les adresses des salles de cinéma.

4. Pour étudier toutes les données du fichier.

```
3 f = open("Salles.csv", "r")
4 Descripteurs = f.readline()
5 ligne = f.readline()
6 compteur = 0
7 while ligne != "":
8     donnees = ligne.split(";")
9     if donnees[...] == '...':
10         compteur = ...
11     ligne = f.readline()
12
13 print(compteur)
```

- 4.1. Expliquer l'utilisation de la boucle non bornée while ligne7.
- 4.2. Écrire le programme Python suivant.
- 4.3. Expliquer les lignes qui s'affichent lors de l'exécution du programme.

Retrouvez éducol sur :



5. Pour calculer le nombre de salles de cinéma dans le 77.

```
3 f = open("Salles.csv", "r")
4 Descripteurs = f.readline()
5 ligne = f.readline()
6 compteur = 0
7 while ligne != "":
8     donnees = ligne.split(";")
9     print(donnees[3])
10    ligne = f.readline()
11    compteur = compteur + 1
12
13 print(compteur)
```

5.1. Compléter le programme afin de calculer ce nombre.

Retrouvez éduscol sur :



## Partie E - géolocalisation des données

Carla souhaite voir sur une carte la position des salles de cinéma. Pour cela nous allons créer un programme qui utilise la latitude et longitude des salles de cinéma du fichier « [Salles.csv](#) » et crée une carte dans le fichier `carte_salles.html`.

1. Écrire le programme suivant. Enregistrer le : `salle.py`.

```
3 import folium
4
5 carte_salles = folium.Map(location=[42.97, 2.76], zoom_start=10)
6
7 f = open("Salles.csv", "r")
8 Descripteurs = f.readline()
9 ligne = f.readline()
10
11 while ligne != "":
12     donnees = ligne.split(";")
13     latitude = float(donnees[9])
14     longitude = float(donnees[10])
15     folium.Marker([latitude, longitude]).add_to(carte_salles)
16     ligne = f.readline()
17
18 carte_salles.save('carte_salles.html')
```

2. Ouvrir le fichier `carte_salles.html` avec votre navigateur.



- 2.1. Exécuter le programme.
- 2.2. Nommer le logiciel qui permet de visualiser le fichier `carte_salles.html` qui se trouve dans le même dossier que votre programme.

Malheureusement la carte n'est pas centrée sur Paris.

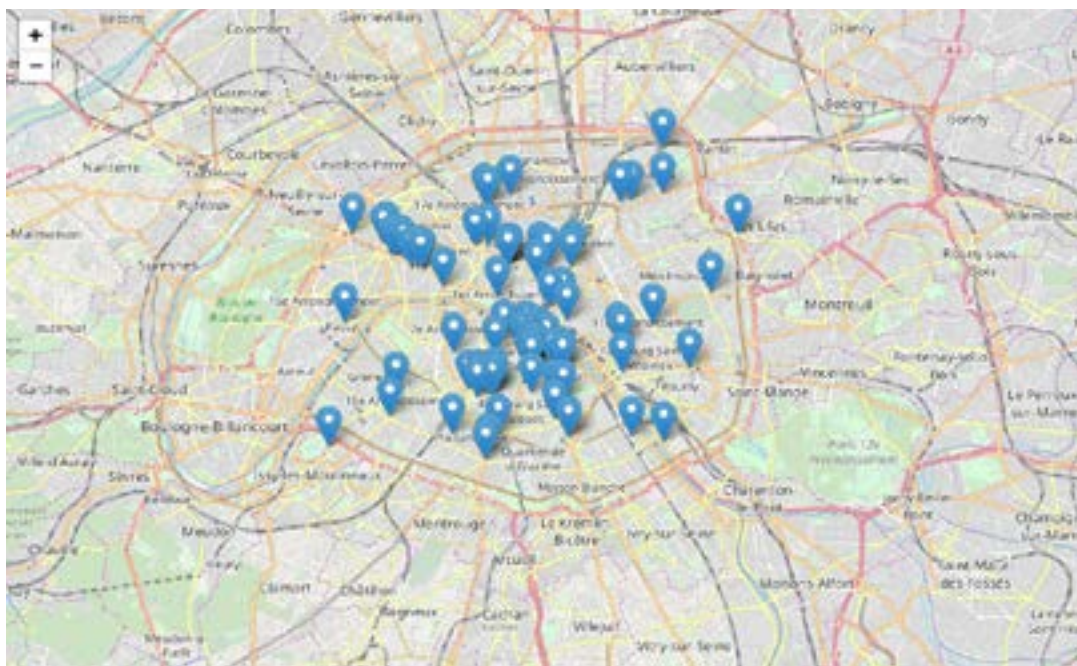
- 2.3. Modifier la ligne 7 du programme afin de centrer la carte sur Paris.
- 2.4. Expliquer les lignes 13, 14 et 15.

Retrouvez éducol sur :



3. Carla souhaite ne voir que les salles de cinéma du département 75. Compléter les lignes 15 et 16 du programme pour obtenir la carte ci-dessous.

```
3 import folium
4
5 carte_salles = folium.Map(location=[48.8534, 2.3488], zoom_start=12)
6
7 f = open("Salles.csv", "r")
8 Descripteurs = f.readline()
9 ligne = f.readline()
10
11 while ligne != "":
12     donnees = ligne.split(";")
13     latitude = float(donnees[9])
14     longitude = float(donnees[10])
15     dep = donnees[...]
16     if dep == "...":
17         folium.Marker([latitude, longitude]).add_to(carte_salles)
18     ligne = f.readline()
19
20 carte_salles.save('carte_salles.html')
```



Retrouvez éducol sur :

