

# LES DONNÉES STRUCTURÉES E01

Merci à S.Colomban pour cette activité

## Situation

On trouve énormément de données sur l'internet :

- Une partie de ces données est publique c'est-à-dire qu'elle est accessible librement. Par exemple le site <https://www.data.gouv.fr> regroupe des données librement consultables et téléchargeables sur de nombreux domaines (agriculture, santé, territoires, économie, emploi,...)
- Une partie de ces données est privée c'est-à-dire qu'elle n'est théoriquement pas accessible. Par exemple, les données bancaires ou militaires sont des données sensibles. Elles sont privées et sécurisées.

## EXERCICE N°1 Exemple d'étude de données publiques à l'aide d'un tableur

Le but de cette partie est d'étudier un fichier de données publiques présent sur <https://www.data.gouv.fr> afin d'en extraire des informations.

### Remarque n°1.

Bien noter les réponses numériques au fur et à mesure, elles seront demandées en fin de travail.

### 1) Récupération et ouverture du fichier :

- Lancer un navigateur et se rendre sur le site <https://www.data.gouv.fr>
- Rechercher "films entrées" dans le champ de recherche.
- Télécharger le fichier "films ayant réalisés plus d'un million d'entrées" au format .xlsx
- Double-cliquer sur ce fichier afin de l'ouvrir avec le tableau Excel.

### Remarque n°2.

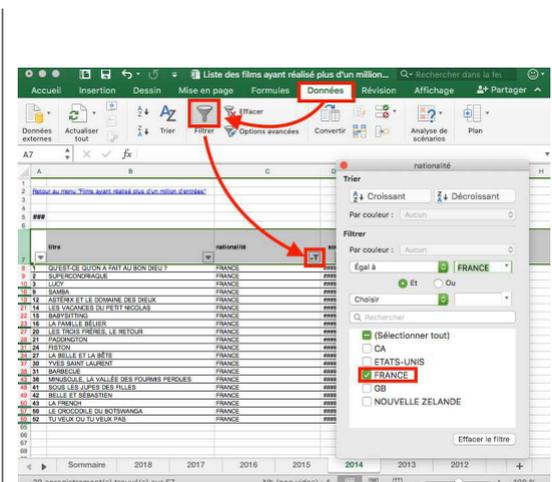
Le format .xlsx est une extension de nom de fichier Office Open XML créée par Microsoft. Ce format étant propriétaire, nous essayerons à l'avenir, de ne pas l'utiliser au profit de formats libres comme par exemple .csv

### 2) Analyse sommaire du fichier :

- Déterminer le film ayant fait le plus d'entrée en 2018.
- Déterminer le film français ayant fait le plus d'entrée en 2018.
- Déterminer le film ayant fait le plus d'entrée en 2017.
- Déterminer le film français ayant fait le plus d'entrée en 2017.

### 3) Analyse approfondie du fichier.

- Se placer sur l'année 2014.
- Cliquer sur la gauche de la ligne 7 afin de la sélectionner entièrement.
- Aller sur l'onglet données puis choisir filtrer.
- Filtrer par nationalité et cocher uniquement la case FRANCE.
- Combien de films français ont fait plus d'un million d'entrée en 2014 ?
- Saisir la formule « =SOUS.TOTAL(9;E8 :E59) » dans la cellule E65 afin de calculer le nombre d'entrées que cumulent les 20 films français pour l'année 2014.
- Par un travail analogue, déterminer le nombre total d'entrées qu'ont réalisé les films US pour l'année 2014.



1) Récupération et ouverture du fichier :

- 1.a) Récupérer le fichier de données [films\\_2003-2019.csv](#)
- 1.b) Faire un clic droit sur ce fichier et l'ouvrir avec LibreOffice.
- 1.c) A l'ouverture cocher UNIQUEMENT le point-virgule pour option de séparateur.

2) Analyse sommaire du fichier :

- 2.a) Déterminer le film ayant fait le plus d'entrée sur toute la période.
- 2.b) Déterminer le nombre de films ayant fait plus de dix millions d'entrées.

3) Création d'un filtre sur les données :

- 3.a) Comme à la partie précédente, surligner la première ligne.
- 3.b) Aller sur l'onglet données puis choisir autofiltre.

4) Analyse de certains filtres :

- 4.a) À l'aide d'un filtre sur les entrées déterminer le film ayant fait le plus d'entrées.
- 4.b) Filtrer les films afin de n'afficher que les films français. Pour cela :

- Sélectionner la cellule nationalité.
- Choisir Filtre Standard .
- Puis faire utiliser

Nom du champ	condition	valeur
nationalité	contient	fr

afin de filtrer les films français.

5)

5.a) À l'aide d'un sous total comme à la partie précédente, déterminer le nombre total d'entrées réalisées par les films français .

5.b) À l'aide d'un autre sous total bien choisi, calculer le nombre de films que cela représente.

### Remarque n°3.

On pourra commencer par lire la documentation de l'instruction SOUS.TOTAL en saisissant "sous.total avec libreoffice" dans un moteur de recherche.

Créer une page web constituée des fichiers **cinema.html** et **decocinema.css** qui peut ressembler à cela par exemple :

### Meilleures entrées au cinéma

Par Ella Fitzgerald

#### Films ayant fait le plus d'entrées en 2017 et 2018 :

##### Côté international :

- Pour l'année 2018, il s'agit du film .....
- Pour l'année 2017, il s'agit du film .....

##### Côté français :

- Pour l'année 2018, il s'agit du film .....
- Pour l'année 2017, il s'agit du film .....

#### Films ayant fait le plus d'entrées entre 2003 et 2019 :

##### Records :

- .... films ont réalisé plus d'un million d'entrées.
- Le film ayant fait le plus d'entrées est : .....
- A noter que .... films ont réalisés plus de dix millions d'entrées !!!

##### Côté français :

- Le nombre total d'entrées sur la période est .....
- Cela représente .... films.