# Modélisation : Bilan semaine 1

* Choix du groupe
* Choix du sujet
* Description des données
* Analyse univariée – Aspect graphique
* Réflexion sur la façon de réunir les informations des deux feuilles dans le sujet 1 ou dans le sujet 2 (les étudiants ayant choisi de travailler sur d’autres données doivent aussi réflechir à la question)
* Réflexion sur la création éventuelle de nouvelles variables
* Questions sur les données

## Membres du groupe

|  |  |  |
| --- | --- | --- |
| Tp n°2 | Nom | Prénom |
| Murillo | Arthur |
| Avellan | Xabi |
|  |  |

## Sujet choisi

Liste des films ayant réalisé plus d’un million d’entrées entre 2013 ET 2018

## Description des données

Population étudiée : Les films ayant réalisé plus d’un million d’entrées au cinéma en France

Date des données : 2003 à 2018

Fournisseur des données : Data.gouv.fr (Plateforme ouverte des données publiques françaises).

Nombre d’individus : 787

Nombre de variables : 6

* Qualitatives nominales : 3
* Qualitatives ordinales : 2
* Quantitatives discrètes : /
* Quantitatives continues : 1

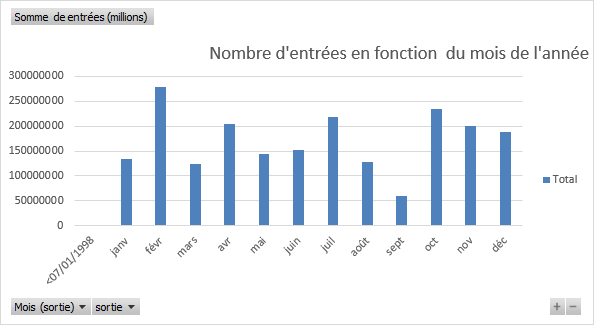
Tableau récapitulatif

|  |  |  |
| --- | --- | --- |
| Nom de Variable | Type | Ensemble ou Intervalle de valeurs. Liste des modalités (si pas trop longue) |
| Titre | Qualitative nominale | / |
| Nationalité | Qualitative nominale | (FR , US , GB , AUS , BE , CA , MAROC , CHINE , REPUBLIQUE TCHEQUE , LUXEMBOURG ) |
| Sortie | Qualitative ordinale | [08/01/03 ; 19/12/18] |
| Entrées (en millions) | Quantitative continues | [1 ; 20,44] |
| Année | Qualitative ordinale | [2013 ; 2018] |
| Type | Qualitative nominale | (Comédie, Action, Aventure, Science-fiction, Comédie musicale, Comédie dramatique, Fantasy, Animation, Biographie, Drame, Horreur) |

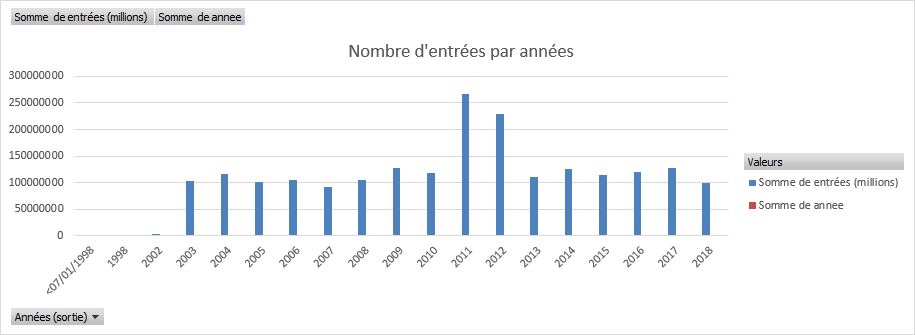
## Analyse univariée – Représentations graphiques

On ne peut pas représenter graphiquement la variable « titre ».

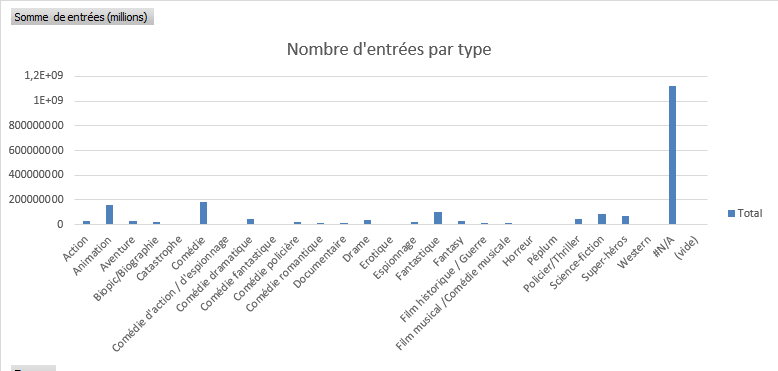
***Sortie***



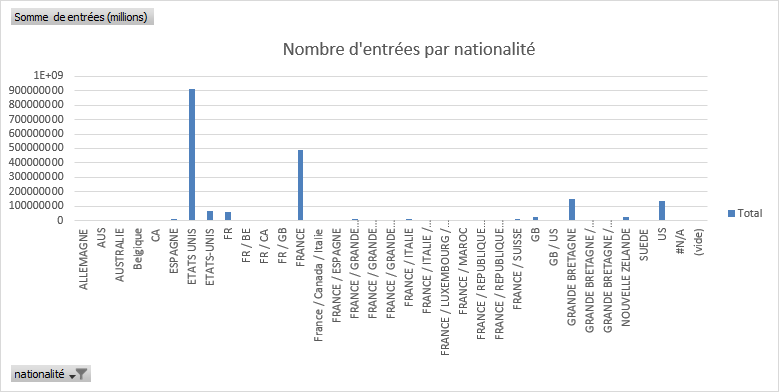
***Année***

******

***Type***

******

***Nationalité***

******

## Réflexion sur la façon de réunir les informations des deux feuilles dans le sujet 1 ou dans le sujet 2

Pour fusionner 2 tableaux Exel, comme celui des films, on peut utiliser la formule :

=RECHERCHEV(@[Nom du film cherché];[Table de données ‘Type’];[Colonne cherchée dans la table ‘Type’;[Type de recherche]).

Ainsi pour le premier film par exemple, cela donne =RECHERCHEV(B2;Types!A2:B766;2;0).

## Réflexion sur la création éventuelle de nouvelles variables

* Le nombre d’entrées dans d’autres pays/dans le monde : pour voir le succès global du film et le comparer entre les différentes régions du monde.
* Le producteur/le distributeur : pour voir si les plus grands succès sont produits par les mêmes studios.
* Les avis (une note sur 5 par exemple) sur chaque film.

## Questions sur les données

* Est-ce que les entrées sont les entrées totales ou pour une période précise suivant la sortie du film ?
* Est-ce que les entrées d’un film sorti en fin d’année sont comptées jusqu’au 31 décembre ou comprennent les entrées de l’année suivante ?