Projet

Power BI

Encadré par :

Pr. Kawtar Younsi Dahbi

Réalisé par :

Mlle. BENHLIMA Douae

Table des matières

[1 Transformation des données 3](#_Toc131400314)

[1.1 Entrées par jour 3](#_Toc131400315)

[1.2 Séances par mois et entrées par Origine 8](#_Toc131400316)

[1.3 Fréquentation cinéma global 10](#_Toc131400317)

[1.4 Retraitement de données 11](#_Toc131400318)

[2 Gestion du Data Model 14](#_Toc131400319)

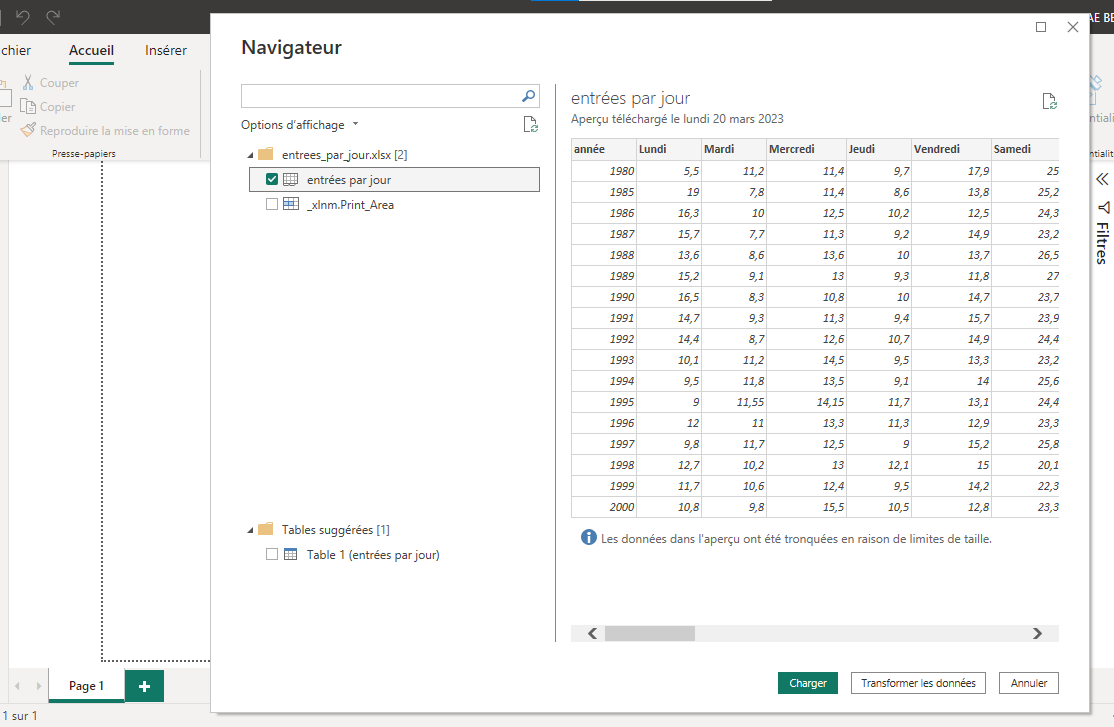
[3 Création de rapport 15](#_Toc131400320)

[4 Langage DAX 20](#_Toc131400321)

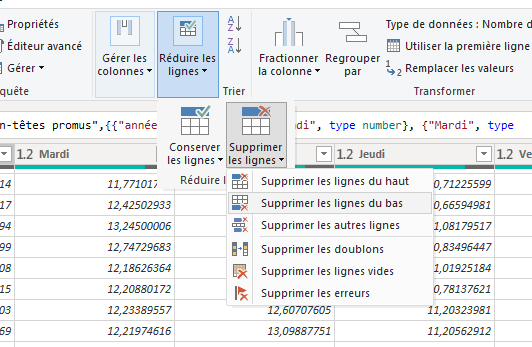
# Transformation des données

## Entrées par jour

* Importez la table des entrées par jour



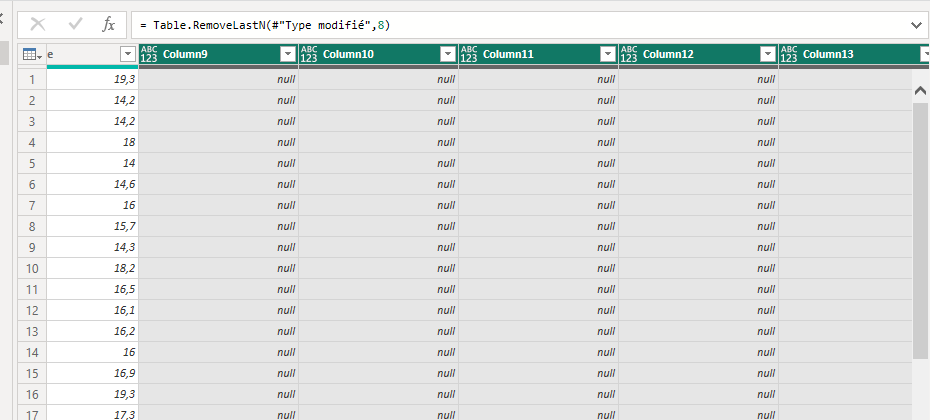
* Des lignes vides apparaissent en bas de tableau, supprimez-les



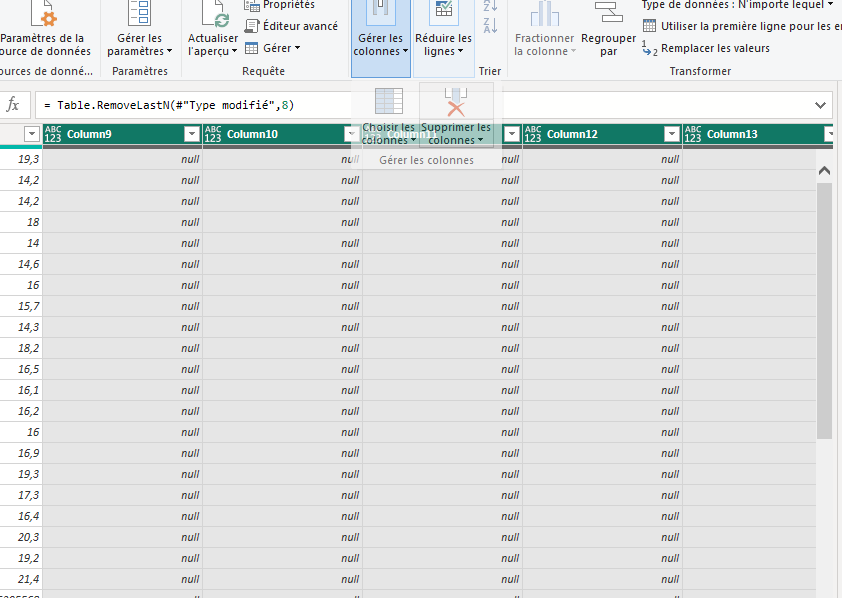
Après choisir nombre des lignes à supprimé égale à 8.



* Des colonnes vides apparaissent également en fin de tableau, conservez uniquement celles qui contiennent des données

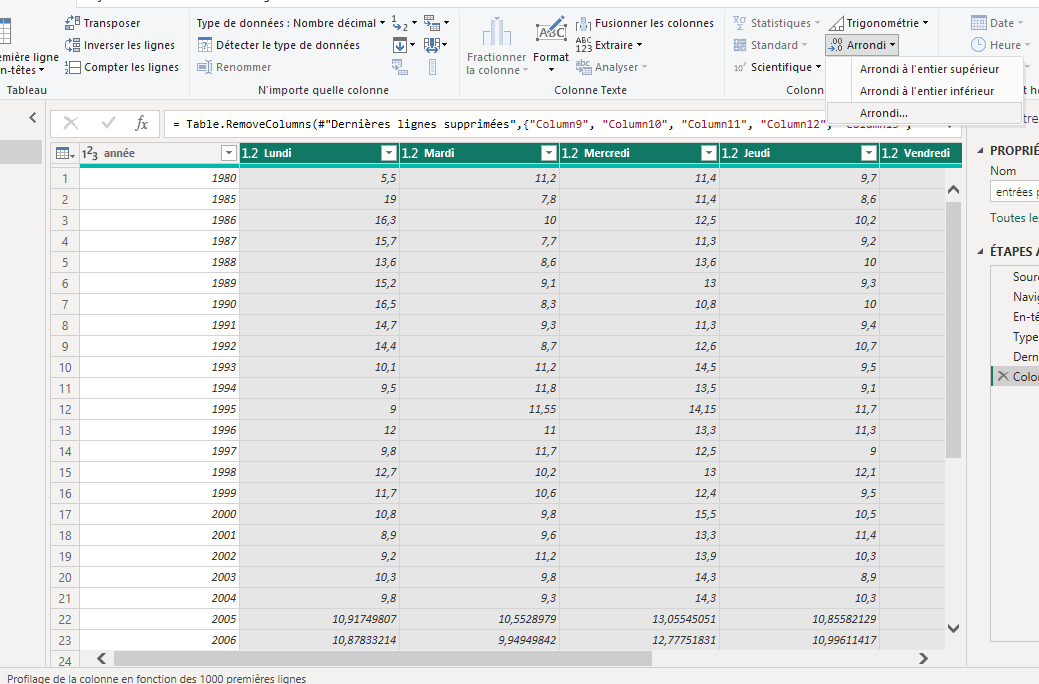


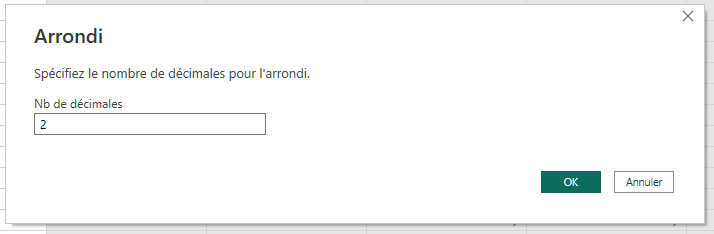
Sélectionner les colonnes à supprimé



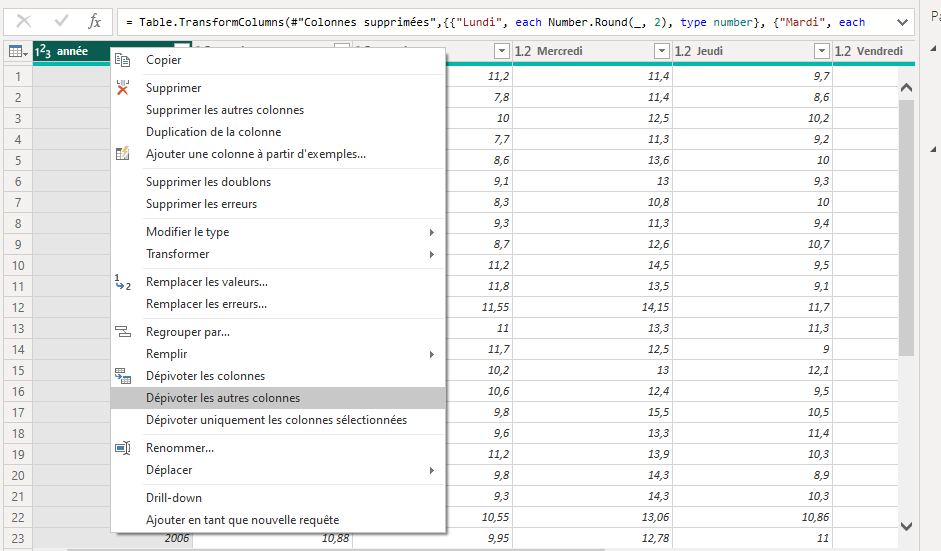


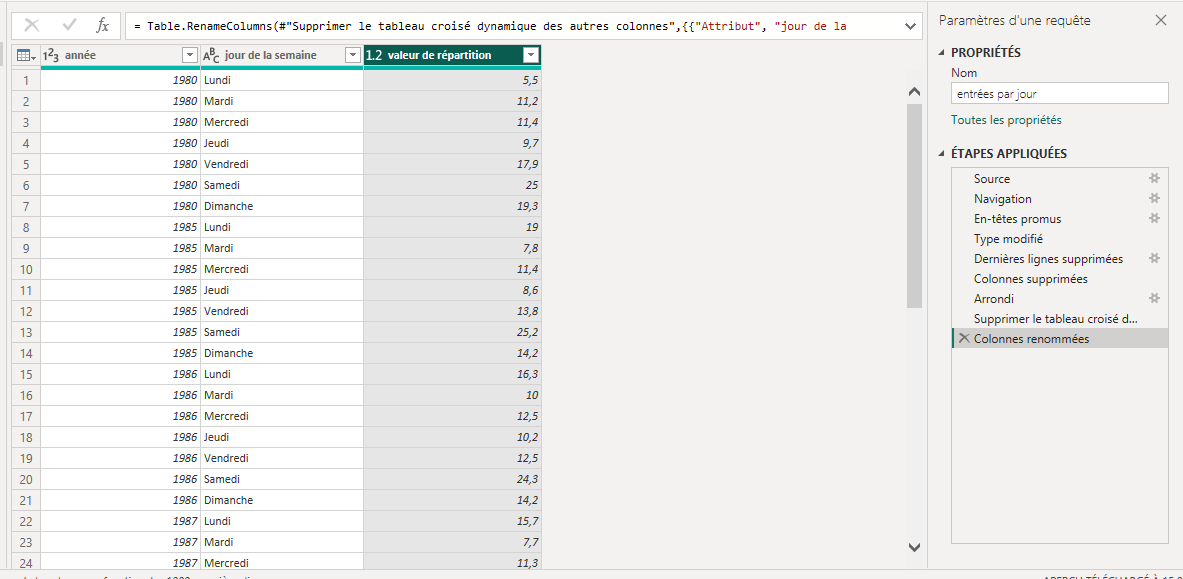
* On voit que le nombre de décimales n’est pas le même selon les lignes. Unifiez cela en arrondissant toutes les valeurs des jours de la semaine à 2 décimales après la virgule



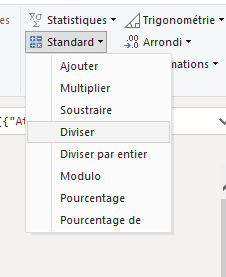
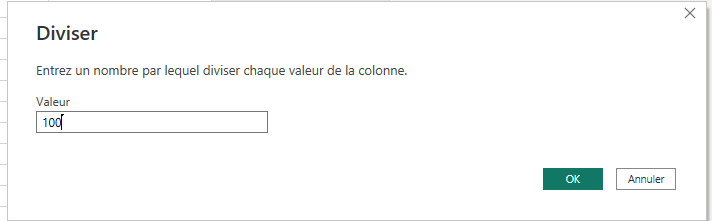
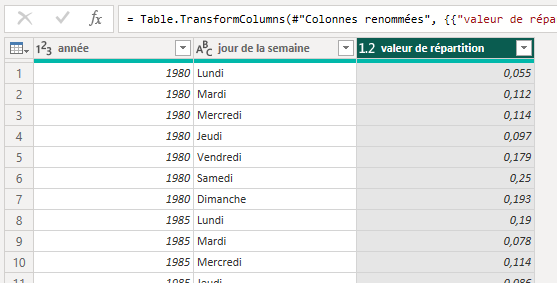


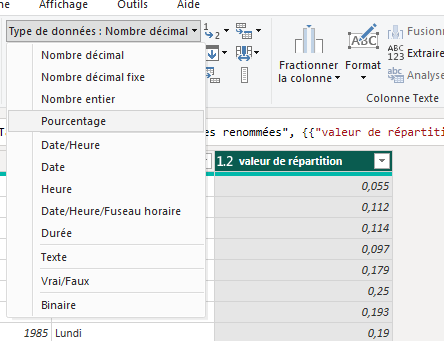
* Dépivoter le tableau afin de ne plus avoir que 3 colonnes : l’année, le jour de la semaine, les valeurs de répartitions

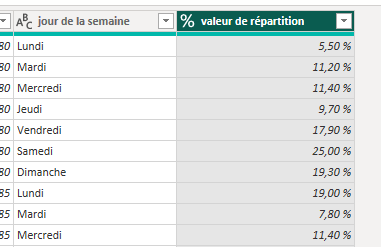




* Les valeurs sont en réalité des pourcentages. Divisez la colonne des valeurs par 100, puis passez la colonne au format pourcentage

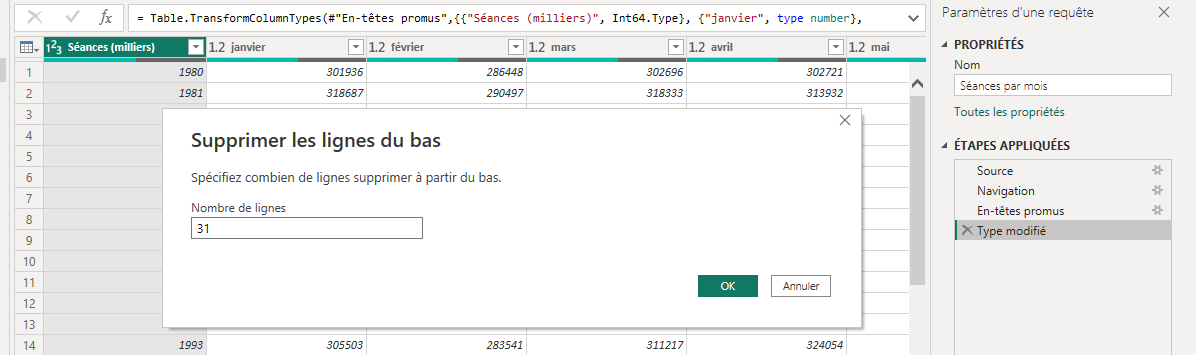
****



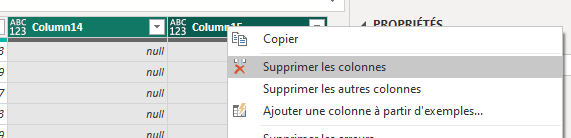


## Séances par mois et entrées par Origine

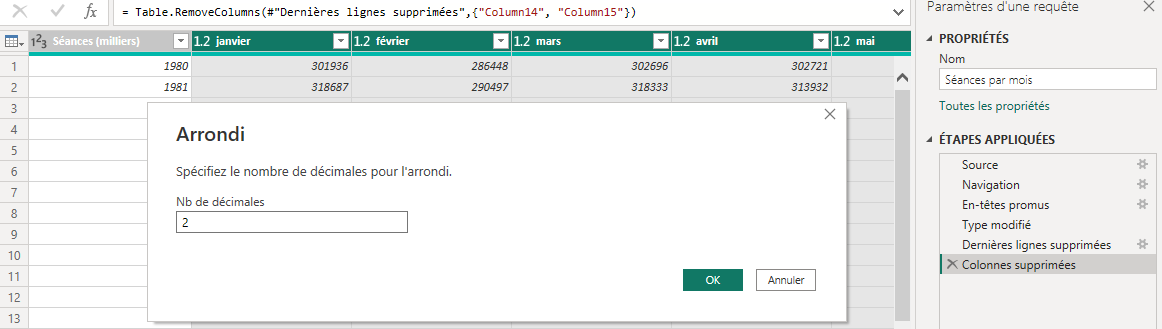
* A partir des modifications réalisées sur la table précédente, déduisez l’ensemble des retraitements à appliquer sur ces tables
* Pour la table Séances par mois :
* Importez la table
* Des lignes vides apparaissent en bas de tableau, supprimez-les



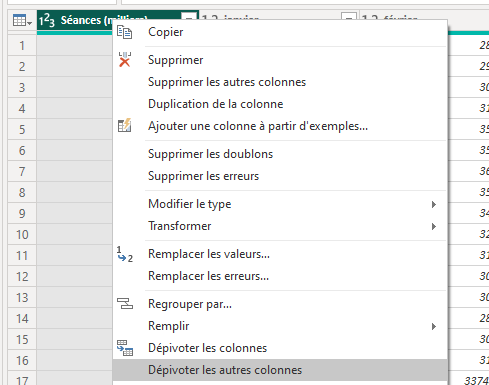
* Des colonnes vides apparaissent également en fin de tableau, conservez uniquement celles qui contiennent des données

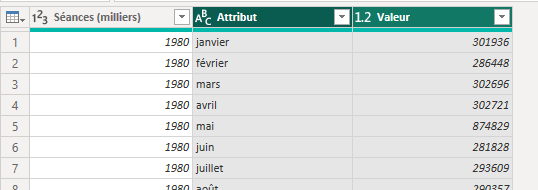


* On voit que le nombre de décimales n’est pas le même selon les lignes. Unifiez cela en arrondissant toutes les valeurs des jours de la semaine à 2 décimales après la virgule

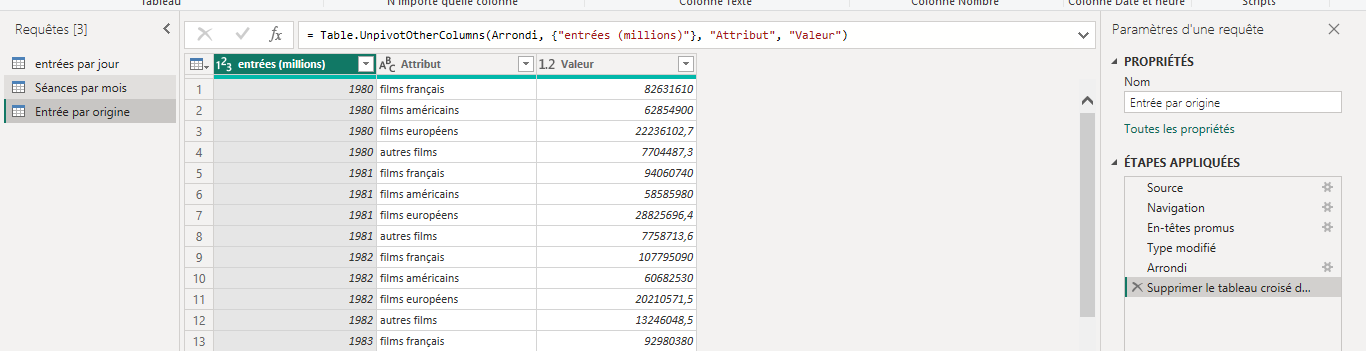


* Dépivoter le tableau afin de ne plus avoir que 3 colonnes : l’année, le jour de la semaine, les valeurs de répartitions



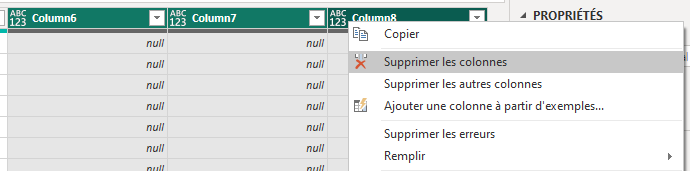


* Les valeurs sont en réalité des pourcentages. Divisez la colonne des valeurs par 100, puis passez la colonne au format pourcentage ???????????????????
* **Entrées par Origine**

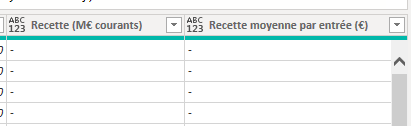


## Fréquentation cinéma global

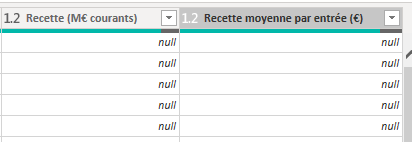
* Supprimez les colonnes vides en fin de tableau

****

* Changez le format des colonnes en fonction de ce qui a été proposé par power query

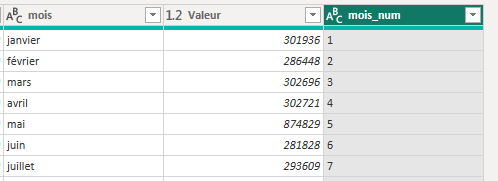
****Pour ces colonnes on doit modifier le type de la colonne en décimal

* Il y a un problème sur certaines colonnes, qui contiennent des « -« , à la place d’une valeur null. Remplacez cette valeur afin de pouvoir convertir la colonne dans un format numérique

Il faut remplacer les erreurs par null

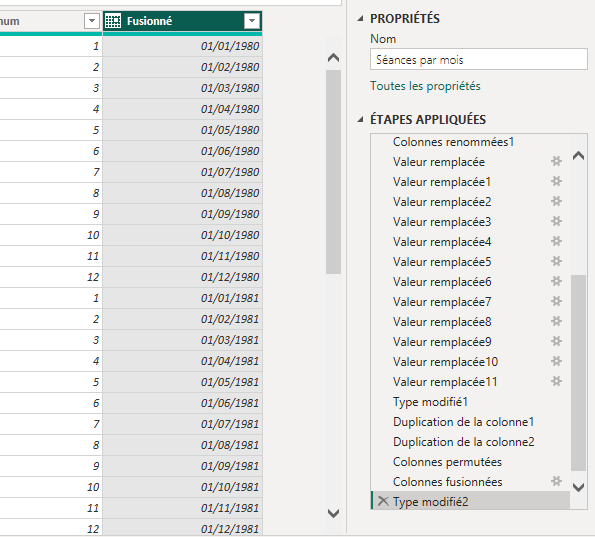
## Retraitement de données

* Dans la table « séance par mois » créez une nouvelle colonne permettant d’avoir le mois en format nombre

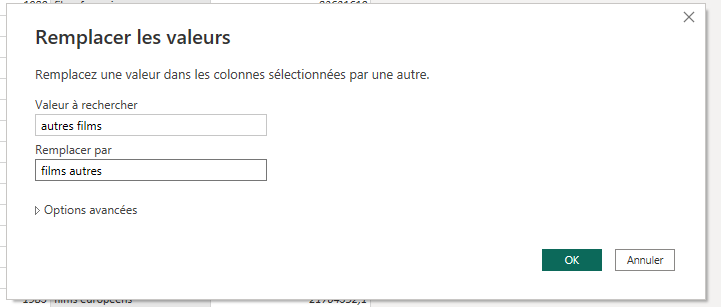


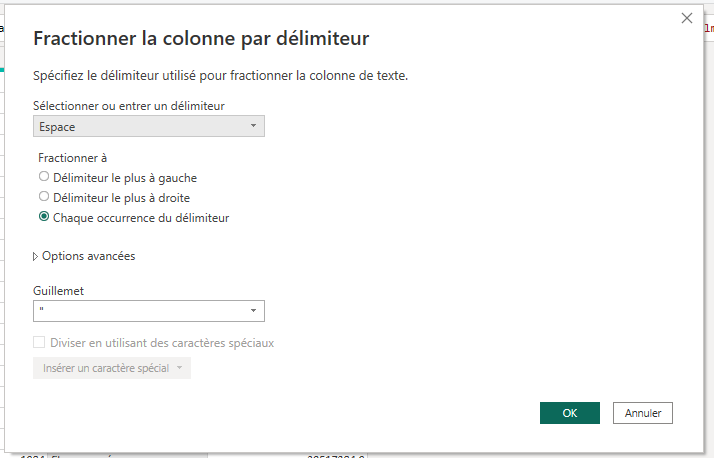
Il faut dupliquer la colonne de mois après remplacer la valeur de chaque mois par son nombre.

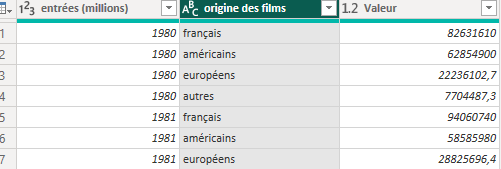
* Dans la table « séance par mois » créez une nouvelle colonne permettant d’obtenir une date avec la combinaison de l’année et du mois (le jour sera systématiquement le premier jour du mois)



* Dans la table « Entrée par origine », le terme « films » apparait systématiquement dans la colonne « origine des films » ce qui n’est pas pertinent. Supprimez les mentions « films » dans les valeurs de la colonne. Selon la méthode utilisée, il est possible que vous vous retrouviez avec des espaces inutiles dans les valeurs, trouvez une solution pour corriger ça.

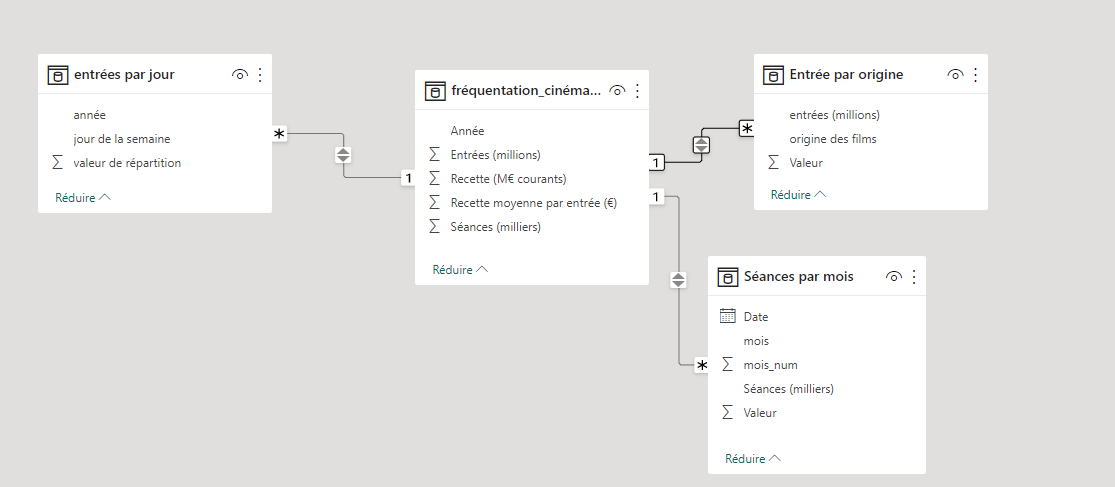




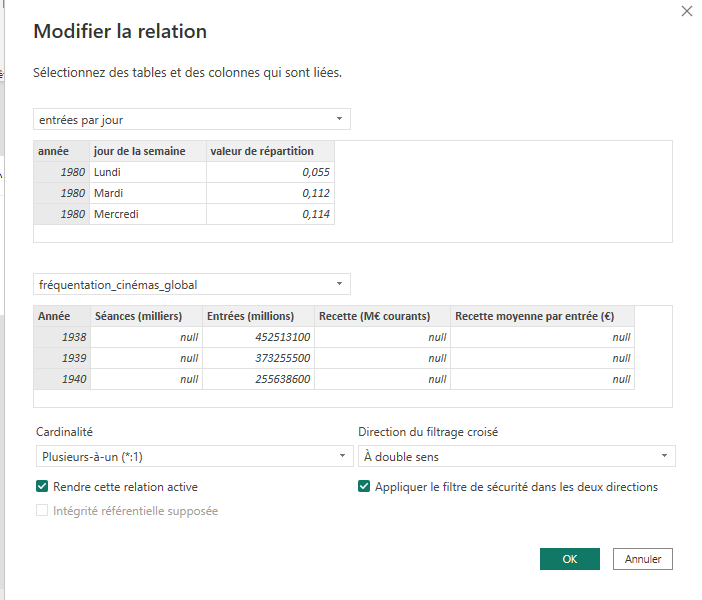


# Gestion du Data Model

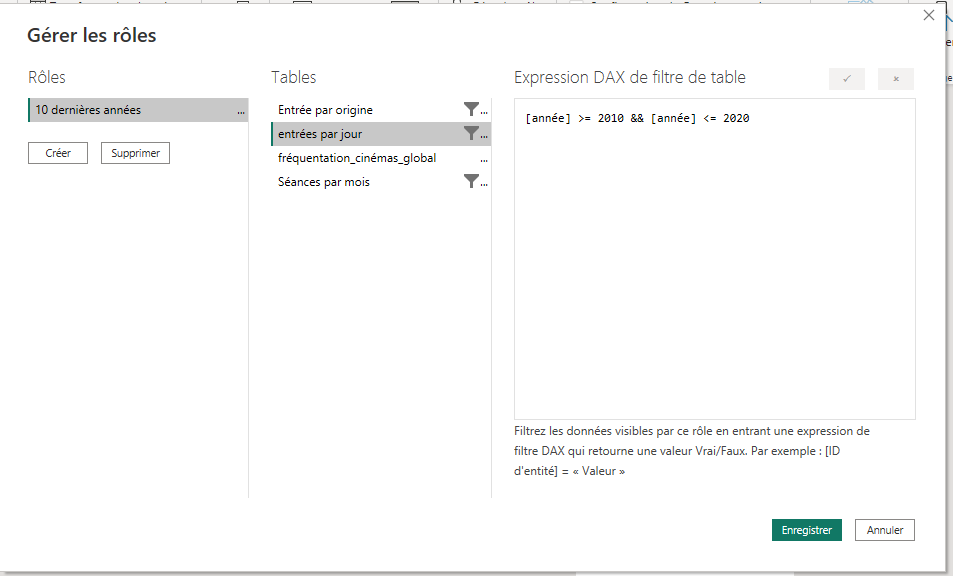
* Faites en sorte que l’ensemble de vos tables soient reliées à la table centrale



* Configurez les liaisons pour le filtrage croisé soient systématiquement dans les 2 sens

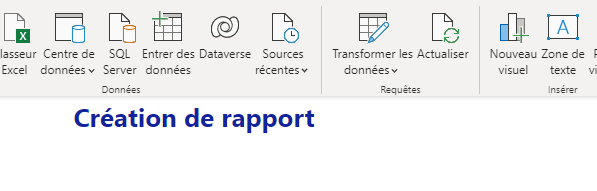


* Créez un rôle « 10 dernières années » qui ne permettra d’accéder qu’au données des années 2010 à 2020

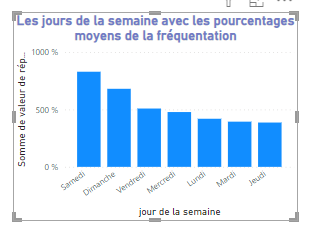


# Création de rapport

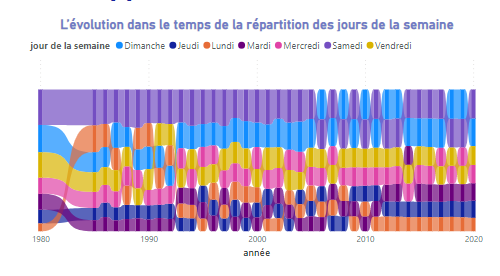
* Ajoutez un titre en haut de votre page de rapport avec le texte de votre choix



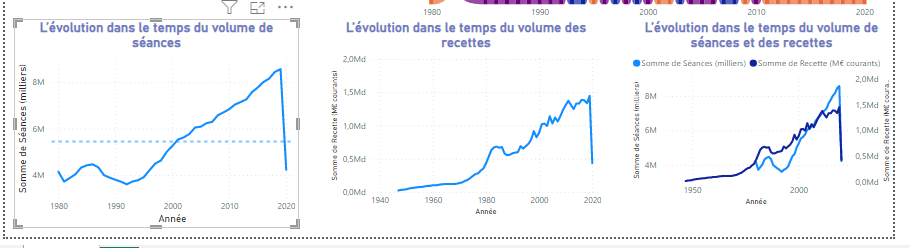
* A partir de la table « Entrées par jour », Créez un histogramme vertical affichant les jours de la semaine avec les pourcentages moyens de la fréquentation



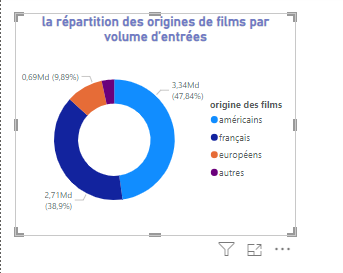
* A partir de la table « Entrées par jour », Créez un graphique à rubans affichant l’évolution dans le temps de la répartition des jours de la semaine



* A partir de la table « Fréquentation cinéma global », créez un graphique à courbes affichant l’évolution dans le temps du volume de séances et des recettes. Ajoutez la ligne de la moyenne pour la courbe des séances



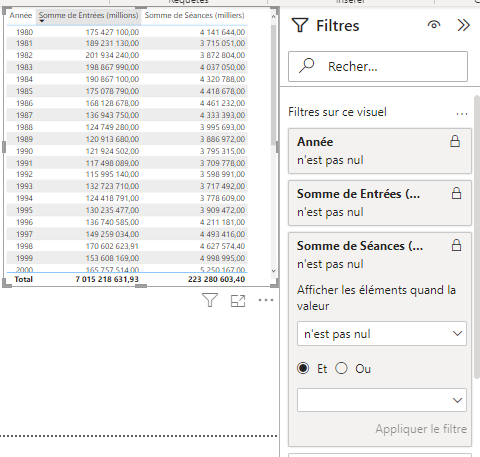
* A partir de la table « Entrées par origine », créez un graphique en anneau pour afficher la répartition des origines de films par volume d’entrées. Retirez la légende, ajoutez l’origine des films sur l’étiquette des données et placez-les étiquettes à l’intérieur de l’anneau.

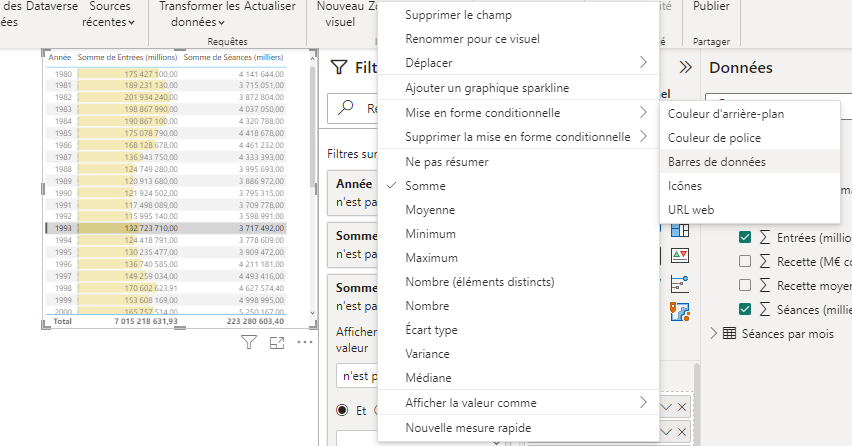


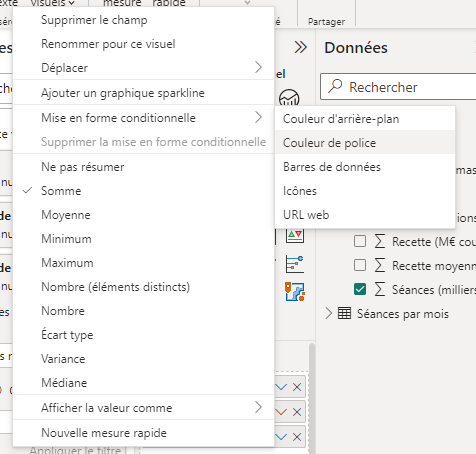
* A partir de la table « Fréquentation cinémas global », créez un tableau avec les années, les recettes et les séances. Filtrez sur le tableau pour qu’une valeur vide n’apparaisse.

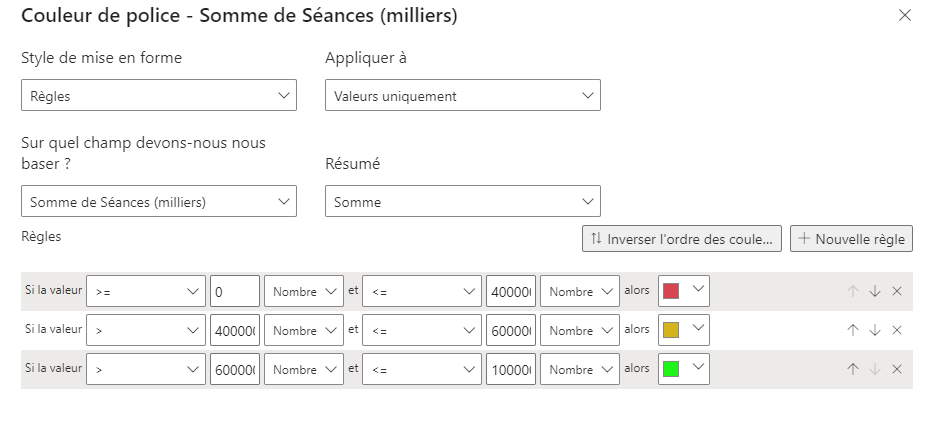
Ajoutez des barres de données sur la colonne Entrées. Ajoutez une couleur de police sur la colonne Séances avec les critères suivants :

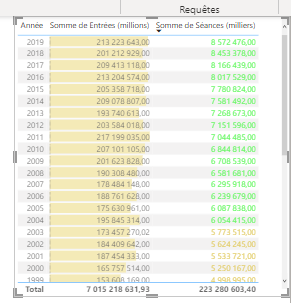
* + Si la valeur est entre zéro et 4 millions, alors la valeur est rouge
  + Si la valeur est entre 4 millions et 6 millions, alors la valeur est jaune
  + Si la valeur est supérieure à 6 millions, alors la valeur est verte
* Ajoutez un filtre global basé sur les années dans l’entête de votre rapport











# Langage DAX

* A partir de la table « Entrées par origine », créez une mesure retournant le nombre d’entrées réalisés par les films Européens (Français + Européen)

Entrées européennes = SUMX(FILTER('Entrée par origine', 'Entrée par origine'[origine des films] = "européens"), 'Entrée par origine'[Valeur])

****

Entrées euro\_ou\_fr = SUMX(

    FILTER(

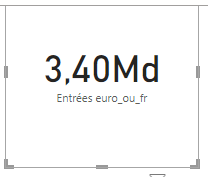
        'Entrée par origine',

        'Entrée par origine'[origine des films] = "européens" || 'Entrée par origine'[origine des films] = "français"

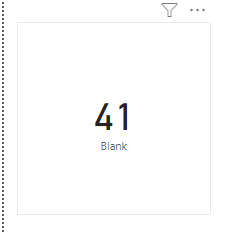
    ),

    'Entrée par origine'[Valeur]

)

****

* A partir de la table « Fréquentation Cinémas global », créez une mesure retournant le nombre de valeurs manquantes dans la colonne « Séances »
* Blank = COUNTBLANK('fréquentation\_cinémas\_global'[Séances (milliers)])

****

* A partir de la table « Entrées par jour », créez une nouvelle colonne indiquant si la ligne est sur un jour de la semaine (du lundi au vendredi) ou sur un jour de week-end (samedi et dimanche)

spec = IF(OR([jour de la semaine]="Samedi",[jour de la semaine]="Dimanche"),"Week-end","Semaine")

