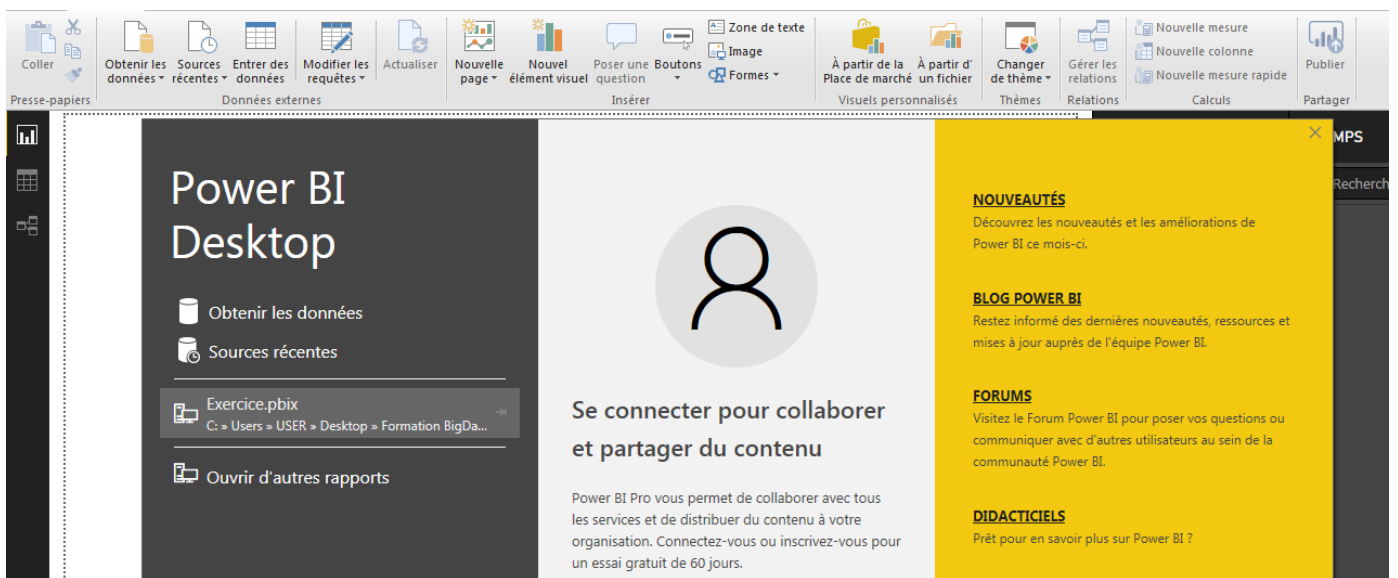

TP- Power BI

Objectifs du TP : Utilisation de « Power BI » pour l'extraction et la transformation et la restitution de données à partir et vers des sources différentes.

I. Installation et Démarrage

- Après avoir installé Power BI Desktop sur votre machine (<https://powerbi.microsoft.com/fr-fr/desktop/>) , le démarrer et créer un nouveau projet intitulé : Hébergement_Parisien.

Après la fermeture de la page de Bienvenue, la fenêtre qui s'affiche aura la forme suivante :



II. Préparation des sources de données

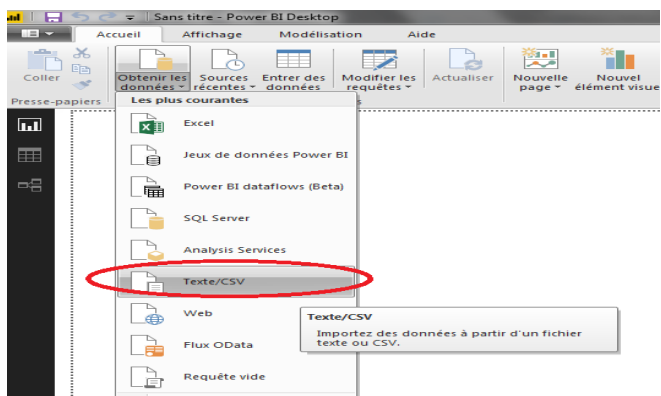
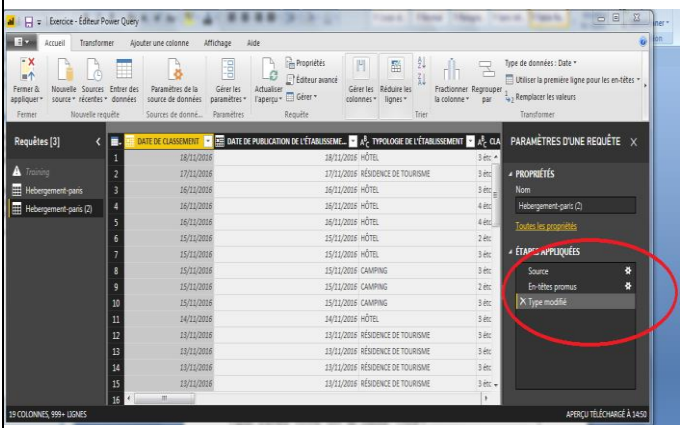
Dans ce TP, nous allons manipuler plusieurs sources de données pour en extraire les données, les transformer et les restituer sous forme de rapport.

La principale source de données avec laquelle on va travailler, est une base de données que nous avons récupérée sur un site web qui s'appelle classement.atout-france.fr, c'est le site officiel du classement des hébergements touristiques, et voyez hébergement-classe mais c'est un site donc dans lequel on peut faire des recherches et le résultat des recherches nous est fourni sous forme de CSV.

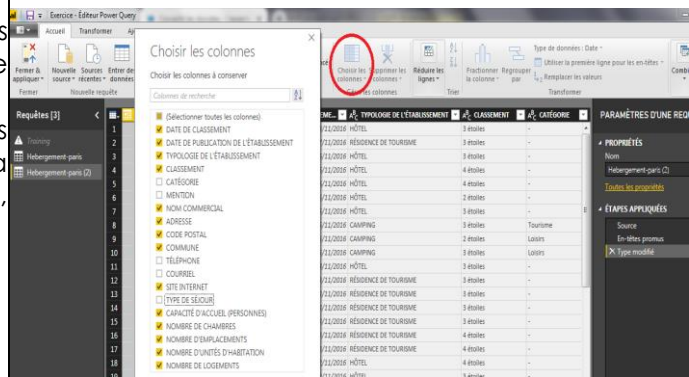
Remarque : Les fichiers que nous allons ajouter (Hebergement-paris.csv et fr_postal_codes.csv) vous ont été fournis avec le support de TP.

Extraction et transformation des données

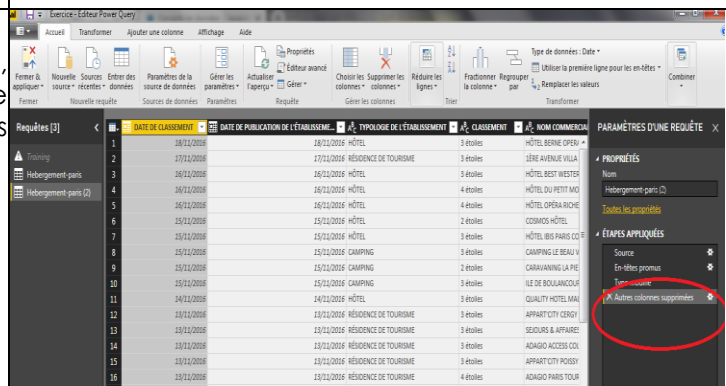
Pour faire cela, suivre les étapes suivantes :

<p>Dans l'onglet Accueil, développer la section <i>Obtenir les données</i>. Sélectionner par la suite le fichier <i>Hebergement-paris.csv</i> dans son emplacement. Une fois le fichier affiché, cliquer sur Modifier</p>	
<p>Commencer par vérifier si les étapes appliquées (étapes de transformation) systématiquement sur le fichier à son ouverture correspondent à vos attentes. Si ces transformations faites systématiquement par PowerBI ne vous correspondent pas alors vous pouvez les supprimer dans la fenêtre encadrée en rouge.</p>	

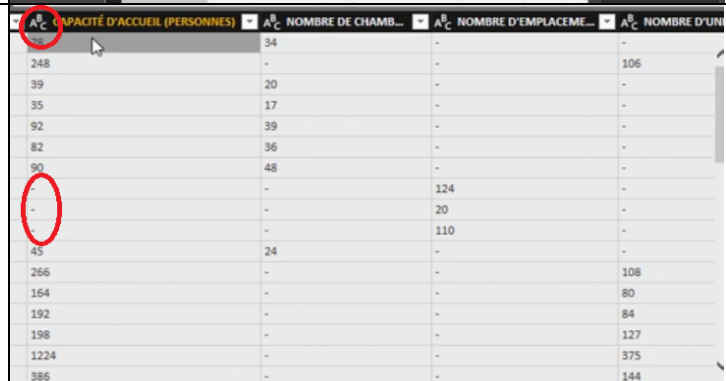
Il faut commencer par choisir les colonnes dont on a besoin pour faire notre analyse.
Il faut donc aller sur Choisir les colonnes et décocher les colonnes dont on n'a pas besoin à savoir : catégorie, mention, téléphone, courriel et type de séjour



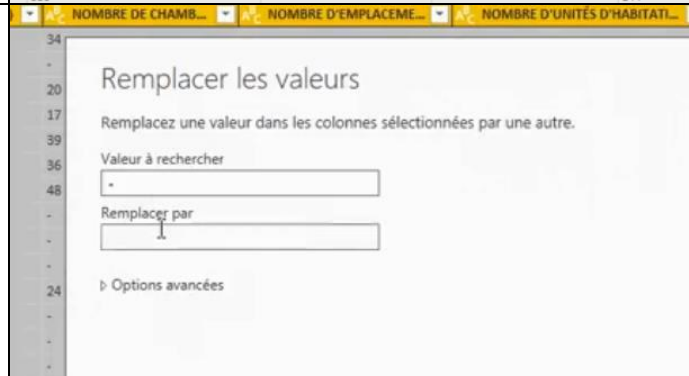
En validant l'étape précédente, remarquer que une nouvelle étapes de transformations s'est ajouté à suites des étapes déjà appliqués.



A cette étape, vous remarquez que toutes les colonnes de nombre, capacité sont en ABC, c'est à dire considérer comme caractères, nous allons donc les convertir en entier.
Afin d'éviter les erreurs, nous allons commencer tout d'abord par enlever tirais qui figurent dans certaines ligne de ces colonnes



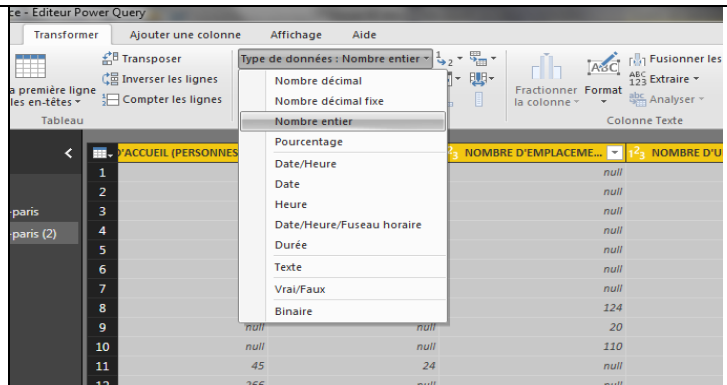
Sélectionner les colonnes en questions et cliquer droit et ensuite sur remplacer les valeurs et remplacer le « - » par un vide



Transformer le type d'une colonne

Aller sur l'onglet « Transformer » et ensuite sur « Type de données » et le mettre à « Nombre entier »

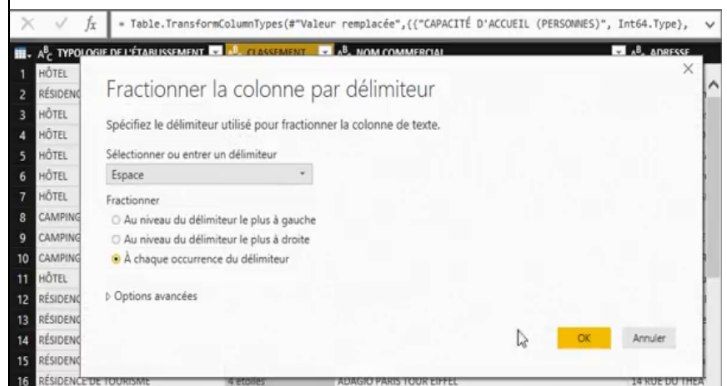
Rq : Les vides ont été convertis en Null



Fractionnement de colonne

Fractionner la colonne Classement en deux colonnes. Cliquer droit sur la colonne et choisir « fractionner la colonne » par délimiter et choisir l'espace comme délimiteur.

Supprimer la deuxième nouvelle colonne qui contient le mot étoile car elle ne nous servira à rien dans notre analyse.

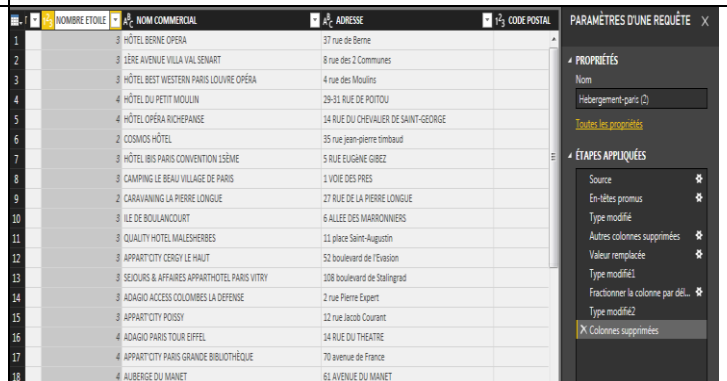


Renommer le nom d'une colonne

Renommer la première colonne « Nombre d'étoile » en cliquant sur le nom existant.

Rq :

- Notre liste d'étapes appliquées s'allonge
- Notre nouvelle colonne est de type entier



1. Importer le fichier « fr_postal_codes.csv »
2. Placer la première ligne comme entête de colonne
3. Ne garder que les colonnes : code postal, longitude et altitude
4. Fractionner la colonne code postal pour ne garder que le code postal (il faut choisir le délimiteur espace le plus à gauche pour n'avoir que 2 colonnes à la place de 3)
5. Supprimer la deuxième colonne et renommer la première en « Code Postal »
6. Changer les autres valeurs des deux colonnes restantes en remplaçant les points en virgules, les transformer par la suite en nombre décimal

Créer maintenant un nouveau groupe en cliquant droit dans la partie « requêtes », l'intituler « Tables temporaires ». Déplacer tous les fichiers dans ce dossier sauf le fichier « hébergement »

Ajouter une colonne personnalisée :

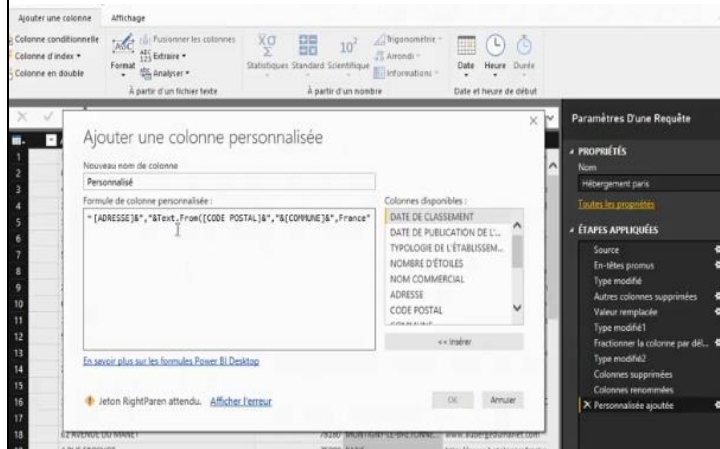
- Dans la table Hébergement, ajouter une nouvelle colonne dans l'angle « Colonne personnalisée » qui s'intitule « personnalisé »
- Renommer la nouvelle colonne « LOCALISATION »

Dans le champ formule taper :

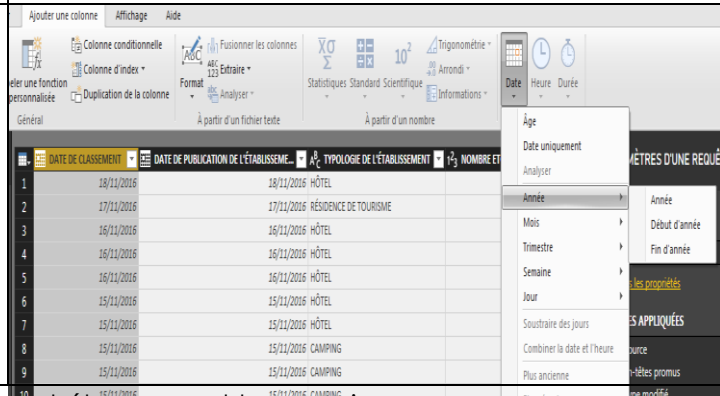
[ADRESSE]&"&"&Text.From([CODE POSTAL])&"&"&[COMMUNE]&"&"&[France]

Rq :

- Text.from est une fonction du langage M qui est le langage de l'éditeur de requête. Cette fonction transforme une valeur numérique en texte (afin de permettre la concaténation dans notre exemple).



- Ajouter une colonne date qui va contenir l'année de la colonne « date de classement ».
- Il faut sélectionner cette dernière colonne ensuite cliquer sur l'angle ajouter une colonne -> Date-> Année -> -> Année.
- Renommer cette colonne « Année »



Renommer la requête hébergement Paris en hébergement temporaire.

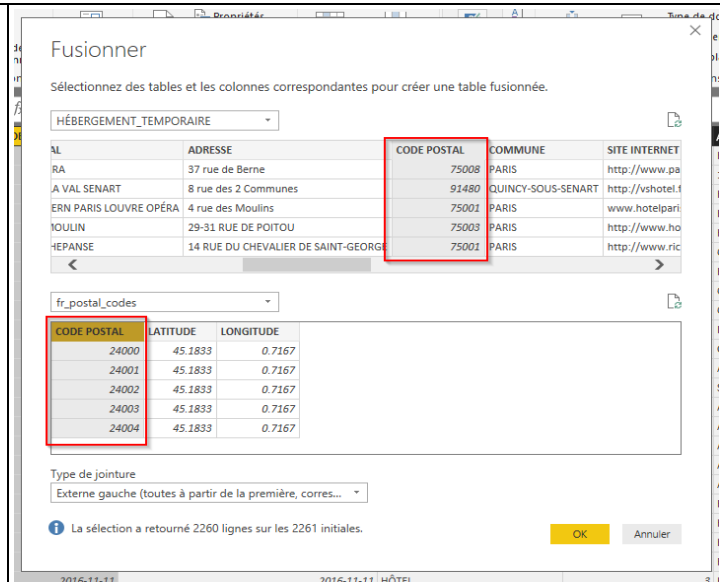
La déplacer dans « tables temporaires ». Rq : Il ne faut pas oublier de décocher l'option « activer le chargement ».

Créer une requête à partir d'autres requêtes

A fin de pouvoir chercher la longitude et les altitudes par code postale. Nous avons réaliser une fusion entre le fichier « fr-code postales » et « hébergement temporaire »

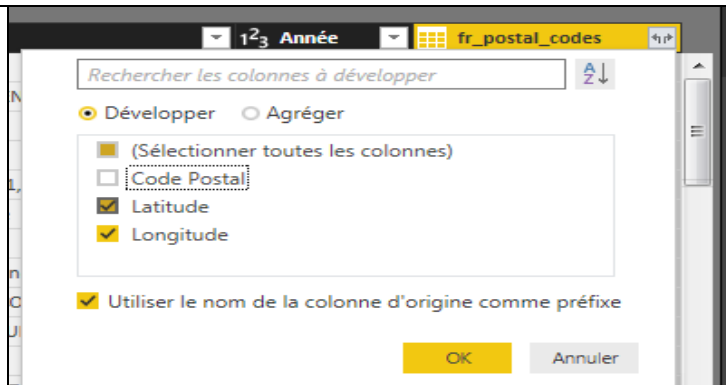
Aller sur Combiner et ensuite sur fusionner des requêtes et choisir « Fusionner les requêtes comme nouvelles »

Le message « La sélection a retourné 2660 lignes sur les 2261 lignes » en bas de la fenêtre signifie qu'un code postal dans le fichier hébergement qui n'a pas trouver un correspondant dans le fichier des codes postaux



TP : Power BI

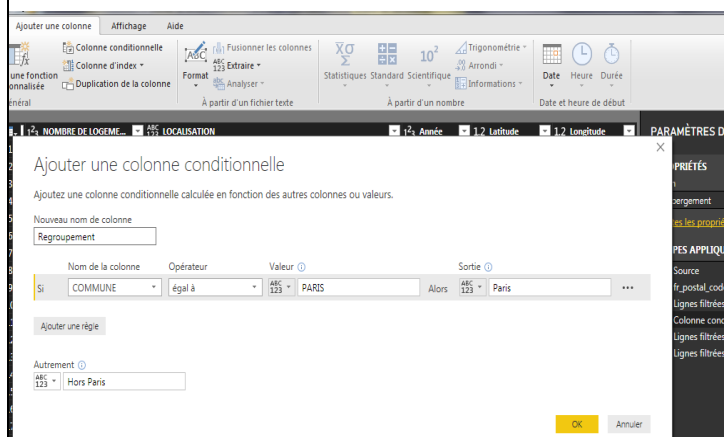
Récupérer la longitude et l'attitude de la colonne nouvellement créée en cliquant dessus.
Nommer la nouvelle requête « Hébergement »



Ajout de colonne conditionnelle :

Sélectionner la colonne commune et sous l'onglet « Ajouter une colonne » -> Colonne conditionnelle. Suivre l'image en face.

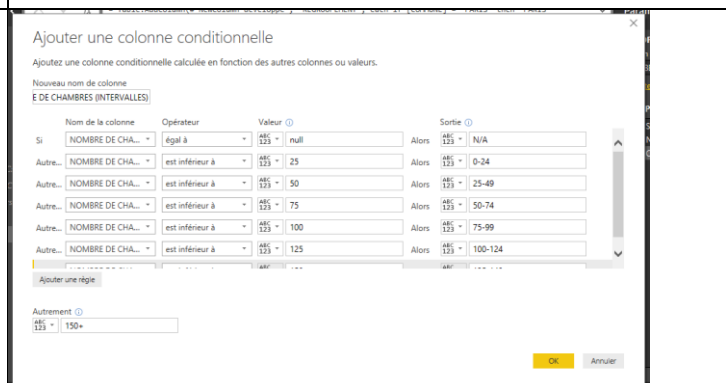
L'objectif c'est de créer deux groupes de commune Paris et Hors Paris
Transformer le type de colonne en texte



Ajout de colonne conditionnelle :

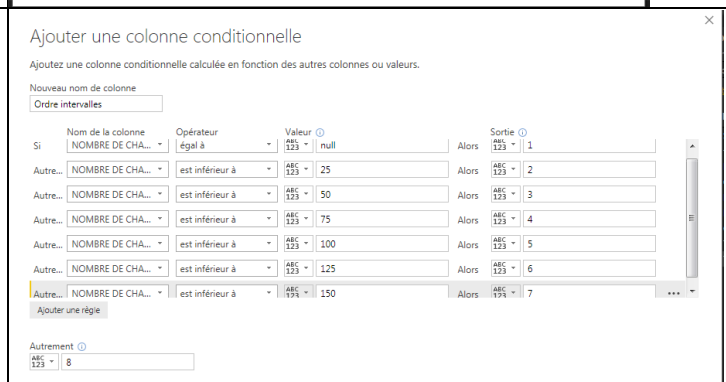
Sélectionner la colonne Nombre de chambre et sous l'onglet « Ajouter une colonne » -> Colonne conditionnelle. Attribuer le nom « Nombre de chambre intervalle ». Suivre l'image en face.

Transformer le type de colonne en texte



Ajouter une nouvelle colonne intitulé « Ordre intervalles » conditionnelle afin de pouvoir trier la colonne « Nombre de chambre ».

Transformer le type de colonne en nombre entier.



Nettoyage de données

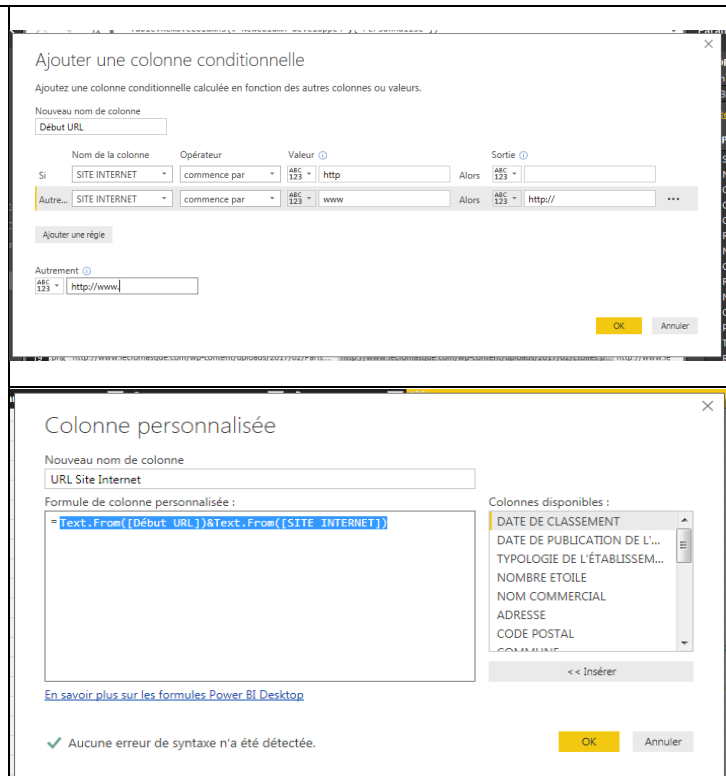
On va procéder au nettoyage de la colonne « SITE INTERNET ».

1. Créer une colonne conditionnelle, l'intituler « Début URL »

2. Créer une colonne personnalisée « URL Site Internet ».

Taper la formule suivante dans la fenêtre qui s'ouvre :

`Text.From([Début URL])&Text.From([SITE INTERNET])`



Clôturer l'éditeur de requête en cliquant sur l'onglet « Fermer & cliquer ».

Conclusion :

L'éditeur de requêtes permet :

- **D'importer des données de diverses sources**
 - Dans cet exemple, nous avons importé des données de fichiers .csv et des données de fichiers Excel
- **De transformer des données**
 - Dans cet exemple, nous avons notamment supprimé des colonnes, modifié des types de données, fractionné des colonnes, remplacé des valeurs, renommé des colonnes
- **D'optimiser et de bonifier des requêtes**
 - Dans cet exemple, nous avons ajouté des colonnes personnalisées, des colonnes conditionnelles et des colonnes de dates
- **De fusionner des requêtes**
 - Dans cet exemple, nous avons fusionné des requêtes à l'intérieur d'une requête existante et nous avons fusionné des requêtes en créant une nouvelle requête
- **D'organiser les requêtes**
 - Dans cet exemple, nous avons créé des dossiers, créé des paramètres et désactivé le chargement de requêtes intermédiaires

III. Création de rapport

- Créer le nom du rapport («Evaluation des établissements») en créant une zone de texte.