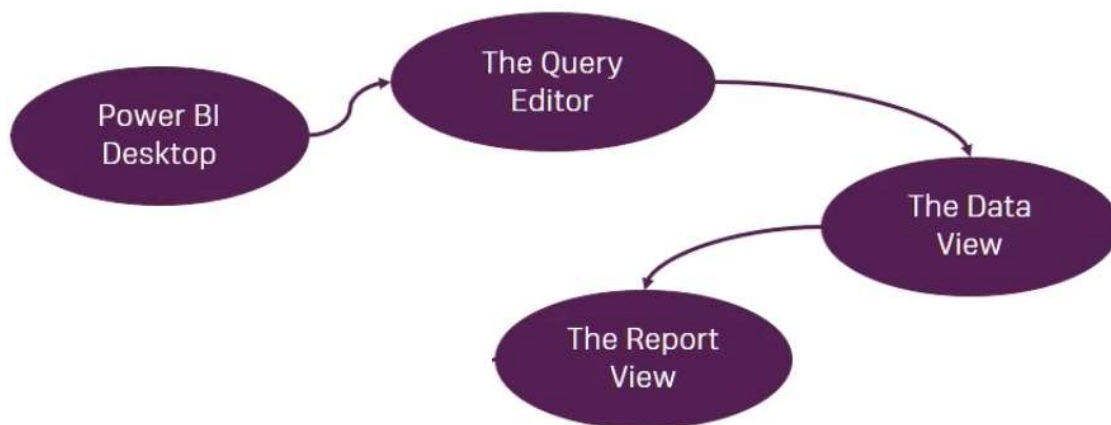


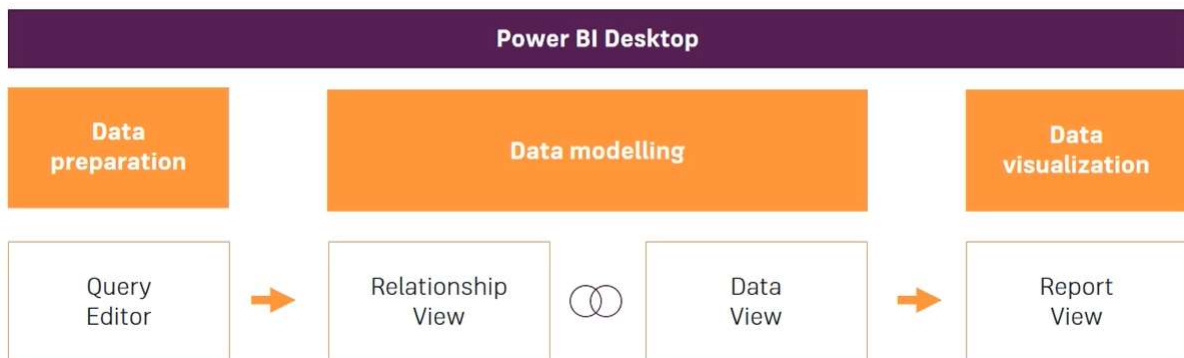
CONNEXION ET MISE EN FORME DE DONNÉES

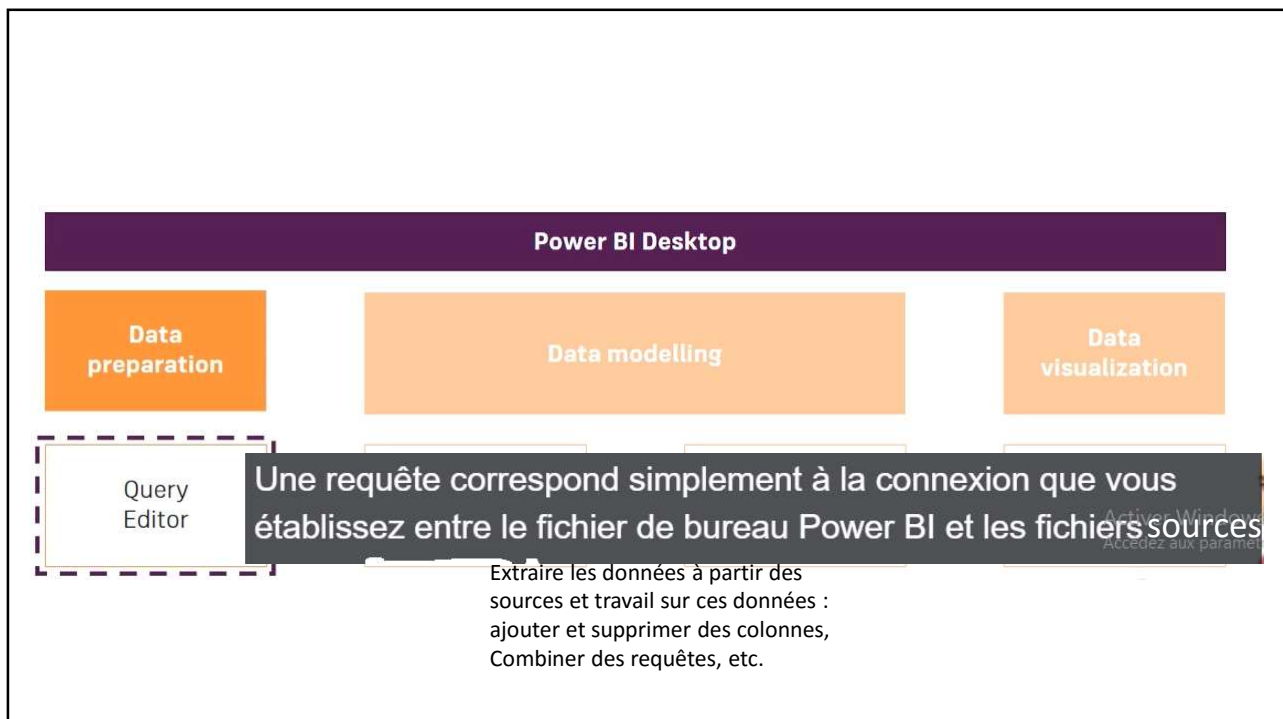
- Connexion aux données source
- Mise en forme et transformation des tables
- Edition, fusion et ajout de requêtes

Flux de travail de Power BI desktop

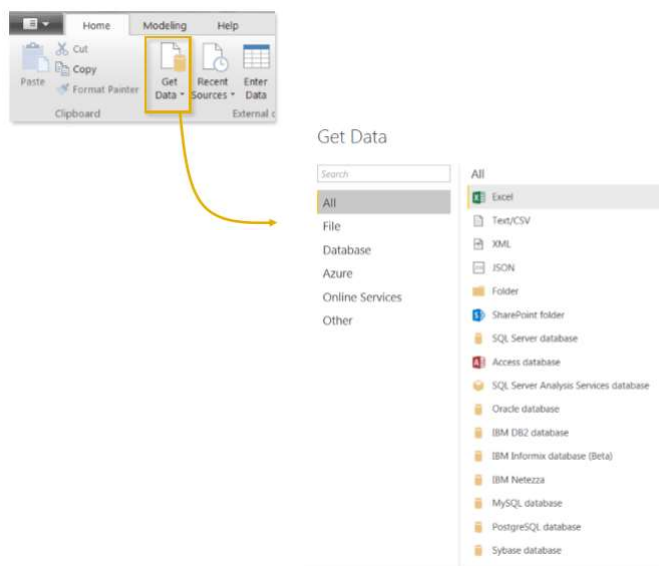


Workflow of Power BI Desktop





TYPES DE CONNECTEURS DE DONNÉES



Power BI peut se connecter à **n'importe quel type de données source**, y compris (mais sans s'y limiter) :

- **Fichiers plats et dossiers** (csv, texte, xls, etc.)
- **Bases de données** (SQL, Access, Oracle, IBM, Azure, etc.)
- **Services en ligne** (Sharepoint, GitHub, Dynamics 365, Google Analytics, Salesforce, Power BI Service, etc.)
- **Autres** (flux Web, scripts R, Spark, Hadoop, etc.)

L'éditeur de requêtes

Outils d'édition de requête (transformations de table, colonnes calculées, etc.)

Barre de formule
(c'est le code «M»)

Liste de requêtes

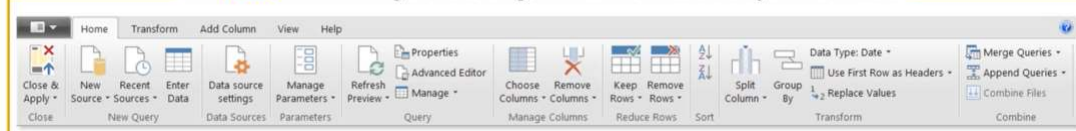
Nom & propriétés de la table

Étapes appliquées
(comme une macro)

Activer Windows

OUTILS D'ÉDITION DE REQUÊTES

The **HOME** tab includes **general settings** and **common table transformation tools**



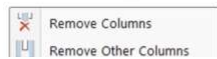
The **TRANSFORM** tab includes tools to **modify existing columns** (splitting/grouping, transposing, extracting text, etc)



The **ADD COLUMN** tools **create new columns** (based on conditional rules, text operations, calculations, dates, etc)



TRANSFORMATIONS DE BASE DES TABLES



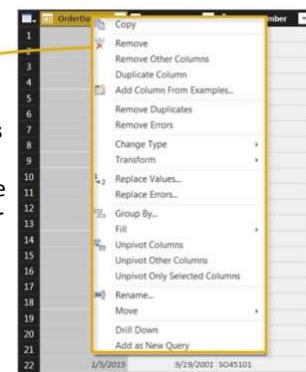
Choisir ou supprimer des colonnes

Astuce: utilisez l'option «Supprimer les autres colonnes» si vous souhaitez toujours un ensemble spécifique

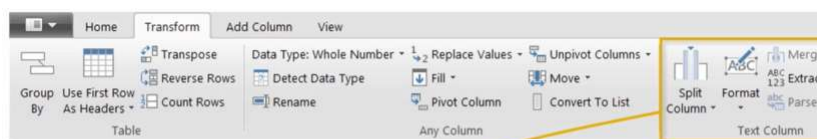


Conservier ou supprime des lignes

Dupliquer, déplacer et renommer des colonnes
Conseil: cliquez avec le bouton droit sur l'en-tête de colonne pour accéder aux outils courants



OUTILS SPÉCIFIQUES AU TEXTE



By Delimiter
By Number of Characters

Fractionner une colonne de texte en fonction d'un délimiteur spécifique ou d'un nombre de caractères

lowercase
UPPERCASE
Capitalize Each Word
Trim
Clean
Add Prefix
Add Suffix

Length
First Characters
Last Characters
Range
Text Before Delimiter
Text After Delimiter
Text Between Delimiters

Extraire des caractères d'une colonne de texte en fonction de longueurs fixes, premier / dernier, plages ou délimiteurs
Conseil: sélectionnez deux colonnes ou plus pour fusionner (ou concaténer) les champs

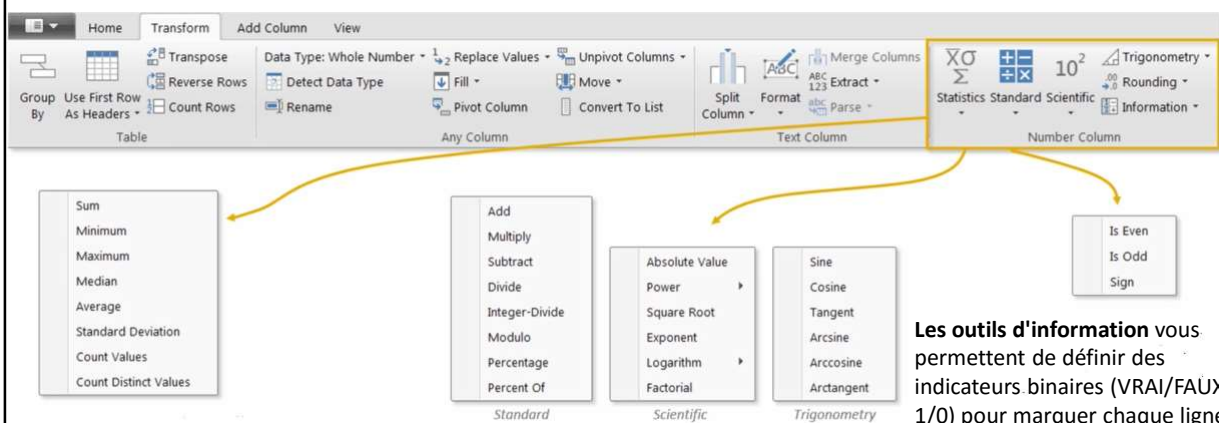


IMPORTANT!

Vous pouvez accéder à plusieurs de ces outils dans les menus "**Transform**" et "**Add Column**" - la différence est de savoir si vous voulez **ajouter une nouvelle colonne** ou en **modifier une existante**

Mettre en forme une colonne de texte en majuscules, en minuscules ou en première lettre en majuscules, ou ajouter un préfixe ou un suffixe
Astuce: Utilisez « **Trim** » pour éliminer les espaces de fin ou « **Clean** » pour supprimer les caractères non imprimables

OUTILS SPÉCIFIQUES AU NOMBRE



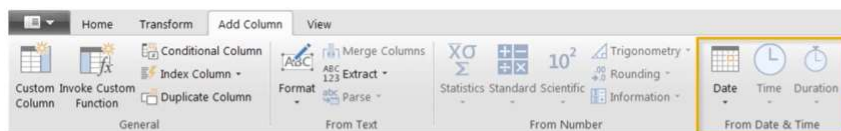
statistiques de base pour la colonne sélectionnée (somme, min/max, moyenne, compte, countdistinct, etc.) ou des calculs plus avancés (puissance, logarithme, sinus, tangente, etc.) à chaque valeur d'une colonne.

Remarque : Ces outils renvoient une valeur unique pour toute la colonne, et non pas pour chaque ligne de la table, ce qui les rend couramment utilisés pour explorer une table plutôt que de la préparer pour le chargement.

Remarque : contrairement aux options Statistiques, ces outils sont appliqués à chaque ligne individuelle de la table.

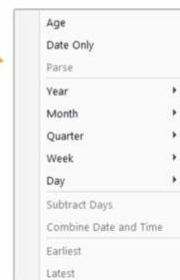
Les outils d'information vous permettent de définir des indicateurs binaires (VRAI/FAUX ou 1/0) pour marquer chaque ligne d'une colonne comme paire, impaire, positive ou négative.

OUTILS SPÉCIFIQUES AUX DATES



Les outils **date et d'heure** incluent les options suivantes :

- **Âge** : différence entre l'heure actuelle et la date dans chaque ligne
- **Date uniquement** : supprime la composante horaire d'un champ date/heure
- **Année/mois/trimestre/semaine/jour** : extrait les composants individuels d'un champ date (les options spécifiques aux temps incluent Heure, Minute, Seconde, etc.)
- **Premier/Dernier** : évalue la date la plus ancienne ou la plus récente d'une colonne comme une seule valeur (accessible uniquement depuis la menu « Transform »)



PRO TIP:

Chargez une table contenant une seule colonne de date et utilisez les outils de date pour créer une table de calendrier complet
colonne de date/heure individuelle

CRÉER UNE TABLE DE CALENDRIER DE BASE

Utilisez les options de **date** prédéfinies dans le menu «**Ajouter une colonne**» pour créer rapidement une table de calendrier à partir d'une liste de dates

	Date	Day	Day of Week	Day Name	Start of Week	Column
1	1/1/2015	1	4	Thursday	12/28/2014	1
2	1/2/2015	2	5	Friday	12/28/2014	1
3	1/3/2015	3	6	Saturday	12/28/2014	1
4	1/4/2015	4	0	Sunday	1/4/2015	1
5	1/5/2015	5	1	Monday	1/4/2015	1
6	1/6/2015	6	2	Tuesday	1/4/2015	1
7	1/7/2015	7	3	Wednesday	1/4/2015	1
8	1/8/2015	8	4	Thursday	1/4/2015	1
9	1/9/2015	9	5	Friday	1/4/2015	1
10	1/10/2015	10	6	Saturday	1/4/2015	1
11	1/11/2015	11	0	Sunday	1/11/2015	1
12	1/12/2015	12	1	Monday	1/11/2015	1
13	1/13/2015	13	2	Tuesday	1/11/2015	1
14	1/14/2015	14	3	Wednesday	1/11/2015	1
15	1/15/2015	15	4	Thursday	1/11/2015	1
16	1/16/2015	16	5	Friday	1/11/2015	1
17	1/17/2015	17	6	Saturday	1/11/2015	1
18	1/18/2015	18	0	Sunday	1/18/2015	1
19	1/19/2015	19	1	Monday	1/18/2015	1
20	1/20/2015	20	2	Tuesday	1/18/2015	1
21	1/21/2015	21	3	Wednesday	1/18/2015	1
22	1/22/2015	22	4	Thursday	1/18/2015	1
23	1/23/2015	23	5	Friday	1/18/2015	1
24	1/24/2015	24	6	Saturday	1/18/2015	1
25	1/25/2015	25	0	Sunday	1/25/2015	1

CONSEIL DE PRO: CRÉATION D'UN CALENDRIER ROULANT

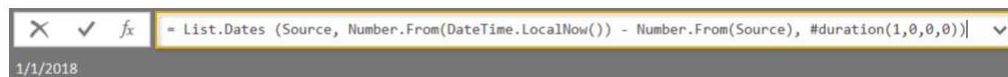
- **Situation** : vous avez créé un rapport hebdomadaire que vous souhaitez actualiser avec de nouvelles données chaque semaine et partager avec vos collègues, votre patron ou client plutôt que de créer un calendrier fixe et de le mettre à jour dans le temps.
- Utilisez un peu de code. Nous allons utiliser ce code pour créer un calendrier qui se mettra à jour en fonction du moment actuel. Il agira donc comme un calendrier glissant qui évoluera avec le temps.

1) Créer une nouvelle requête vide(**Get Data > Blank Query** or **New Source > Blank Query**)

2) Dans la barre de formule, générez une date de début en entrant un «littéral» (au format AAAA, MM, JJ):



3) Cliquez sur l'icône fx pour ajouter une nouvelle étape personnalisée et entrez la formule suivante exactement comme indiqué :



4) Convertissez la liste résultante en tableau (List Tools > To Table) et formatez la colonne en date

5) Add calculated Date columns (Year, Month, Week, etc.) as necessary using the Add Column tools

AJOUT DE COLONNES D'INDEX

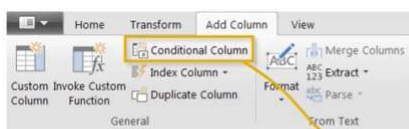


Les **colonnes d'index** contiennent une liste de valeurs séquentielles qui peuvent être utilisées pour **identifier chaque ligne unique** dans une table (généralement à partir de 0 ou 1)

Ces colonnes sont souvent utilisées pour créer des **ID uniques** qui peuvent être utilisés pour former des relations entre les tables (plus à ce sujet plus tard!)

	Index	OrderDate	StockDate	OrderNumber	ProductKey	CustomerKey
1	1	1/1/2015	9/21/2001	SO45080	332	14657
2	2	1/1/2015	12/5/2001	SO45079	312	29255
3	3	1/1/2015	10/29/2001	SO45082	350	11455
4	4	1/1/2015	11/16/2001	SO45081	338	26782
5	5	1/2/2015	12/15/2001	SO45083	312	14947
6	6	1/2/2015	10/12/2001	SO45084	310	29143
7	7	1/2/2015	12/18/2001	SO45086	314	18747
8	8	1/2/2015	10/9/2001	SO45085	312	18746
9	9	1/3/2015	10/3/2001	SO45093	312	18906
10	10	1/3/2015	9/29/2001	SO45090	310	29170
11	11	1/3/2015	12/11/2001	SO45088	345	11398
12	12	1/3/2015	10/24/2001	SO45092	313	18899
13	13	1/3/2015	12/16/2001	SO45089	351	25977
14	14	1/3/2015	10/26/2001	SO45091	314	18909
15	15	1/3/2015	9/11/2001	SO45087	350	11388
16	16	1/3/2015	9/11/2001	SO45094	310	22785
17	17	1/4/2015	10/30/2001	SO45096	312	12483
18	18	1/4/2015	10/30/2001	SO45097	313	29151

AJOUT DE COLONNES CONDITIONNELLES



Les **colonnes conditionnelles** vous permettent de définir de nouveaux champs en fonction de règles et conditions logiques (instructions IF/THEN)

Dans ce cas, nous créons une nouvelle colonne conditionnelle appelée "**QuantityType**", qui dépend des valeurs de colonne "**OrderQuantity**", comme suit :

- Si OrderQuantity = 1,
QuantityType = "**Single Item**"
- Si OrderQuantity > 1,
QuantityType = "**Multiple Items**"
- Sinon QuantityType = "**Other**"

Add Conditional Column

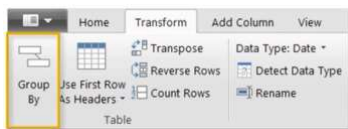
Add a conditional column that is computed from the other columns or values.

New column name
QuantityType

Column Name	Operator	Value	Output
If OrderQuantity	equals	1	Single Item
Else If OrderQuantity	is greater than	1	Multiple Items
Otherwise			Other

OK Cancel

Regroupement et agrégation de données



Group By vous permet d'agréger vos données à différents niveaux (i.e., transformer les données quotidiennes en mensuelles, regrouper les données au niveau des transactions par magasin, par région, etc.)

	OrderDate	ProductKey	CustomerKey	OrderQuantity
1	4/25/2017	214	14719	1
2	7/16/2016	214	11243	1
3	12/31/2016	214	23452	1
4	6/28/2017	214	22748	1
5	10/6/2016	214	25025	1
6	10/7/2016	214	16504	1
7	10/18/2016	214	13043	1
8	1/19/2017	214	23101	1
9	9/7/2016	214	24900	1
10	1/18/2017	214	24196	1
11	6/29/2017	214	12363	1
12	11/6/2016	214	14570	1
13	11/13/2016	214	16999	1
14	7/11/2016	214	12282	1
15	10/9/2016	214	15685	1
16	8/1/2016	214	16982	1
17	12/4/2016	214	12835	1

Quantités commandées par date de commande, clé de produit et clé



	ProductKey	TotalQuantity
1	214	2059
2	217	1940
3	222	1995
4	225	4151
5	226	892
6	231	406
7	234	424
8	237	381
9	230	169
10	211	135
11	212	179
12	213	168
13	214	151
14	199	10
15	198	53
16	222	5
17	223	34

Dans ce cas nous transformons une table avec des transactions quotidiennes en un résumé de la **TotalQuantity** regroupée par **ProductKey**.
REMARQUE les champs non spécifiés dans les paramètres Group By sont perdus

Regroupement et agrégation de données (avancé)

OrderDate	ProductKey	CustomerKey	OrderQuantity
6/25/2017	214	14719	1
7/24/2016	214	11243	1
11/21/2016	214	21452	1
6/28/2017	214	22748	1
10/6/2016	214	25025	1
10/7/2016	214	16504	1
10/13/2016	214	13043	1
1/19/2017	214	23301	1
9/7/2016	214	24900	1
1/18/2017	214	24196	1
6/29/2017	214	12963	1
11/6/2016	214	14570	1
11/13/2016	214	16999	1
7/11/2016	214	12281	1
10/9/2016	214	15685	1
6/2/2016	214	16982	1
12/4/2016	214	12835	1

ProductKey	CustomerKey	TotalQuantity
214	11243	1
214	13043	1
214	14719	1
214	16504	1
214	16999	1
214	17049	1
214	18054	1
214	21452	1
214	22748	1
214	23301	1
214	24196	1
214	24900	1
214	25025	1
214	25685	1
214	25982	1
214	26335	1

Cette fois, nous transformons la table des transactions quotidiennes au en un résumé de «**TotalQuantity**» agrégé à la fois par «**ProductKey**» et «**CustomerKey**» (en utilisant l'option avancée de la boîte de dialogue)

REMARQUE: cela revient à créer un tableau croisé dynamique dans Excel et à extraire «Sum of OrderQuantity» avec ProductKey et CustomerKey comme étiquettes de ligne

PIVOTING & UNPIVOTING

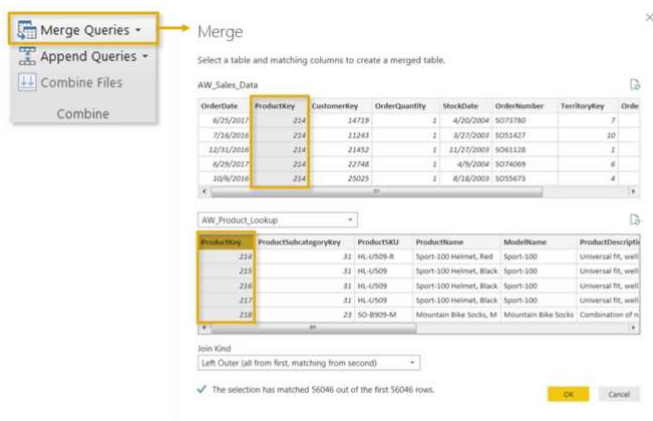
«**Pivoter**» est un moyen sophistiqué de décrire le processus de conversion de **valeurs de ligne distinctes en colonnes** («pivoting») ou de transformer des **colonnes en lignes** («unpivoting»)



Imaginez que la table est sur une charnière; pivoter, c'est comme le faire pivoter d'une disposition **verticale** à une **horizontale**, et dépivoter, c'est comme le faire pivoter d'une horizontale à la verticale

REMARQUE: **Transpose** fonctionne de manière très similaire, mais ne reconnaît pas les valeurs uniques; la table entière est transformée de sorte que chaque ligne devienne une colonne et vice versa

FUSION DE REQUETES



La **fusion de requêtes** vous permet de **joindre des tables** basées sur une colonne commune (comme VLOOKUP)

Dans ce cas, nous fusionnons la table **AW_Sales_Data** avec la table **AW_Product_Lookup**, qui partagent une colonne commune, "ProductKey"

REMARQUE: la fusion **ajoute des colonnes** à une table existante

IMPORTANT!

En général, il est préférable de garder les tables séparées et de définir des relations entre elles

Les types de fusion de requêtes

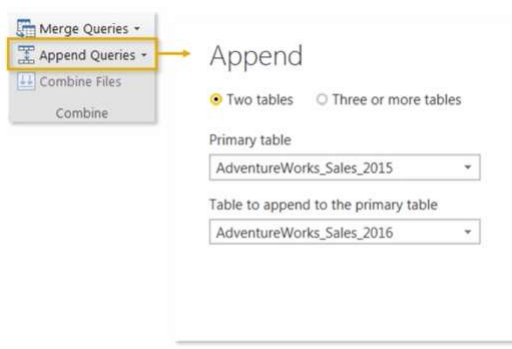
Separate Queries

Query 1 LEFT		Query 2 RIGHT	
ID	Sales	ID	Region
A	10	A	USA
B	50	BB	Europe
C	20	C	Asia

Merged Queries

	LEFT			RIGHT			FULL		
	ID	Sales	Region	ID	Region	Sales	ID	Sales	Region
	A	10	USA	A	USA	10	A	10	USA
	B	50	n/a	BB	Europe	n/a	B	50	n/a
Outer	C	20	Asia	C	Asia	20	C	20	Asia
							BB	n/a	Europe
Anti	ID	Sales	Region	ID	Region	Sales			
	B	50	n/a	BB	Europe	n/a			
Inner	ID	Sales	Region						
	A	10	USA						
	C	20	Asia						

AJOUT DE REQUETES



L'ajout de requêtes vous permet de **combiner** (ou empiler) des tables qui partagent exactement la même structure de colonnes et les mêmes types de données

Dans ce cas, nous ajoutons la table **AdventureWorks_Sales_2015** à la table **AdventureWorks_Sales_2016**, qui est valide car ils partagent des structures de table identiques

REMARQUE : *Append* ajoute des lignes à une table existante



PRO TIP:

Utilisez l'option «**Folder**» (Obtenir les données> Plus> Dossier) pour ajouter tous les fichiers dans un dossier (en supposant qu'ils partagent la même structure) ; lorsque vous ajoutez de nouveaux fichiers, actualisez simplement la requête et ils s'ajouteront automatiquement!

PARAMÈTRES DE SOURCE DE DONNÉES

Data source settings

Manage settings for data sources that you have connected to using Power BI Desktop.

★ Data sources in current file ○ Global permissions

Search data source settings

- c:\users\chris\documents\secon...s\adventureworks_calendar.csv
- c:\users\chris\documents\secon...s\adventureworks_customers.csv
- c:\users\chris\documents\secon...reworks_product_categories.csv
- c:\users\chris\documents\secon...orks_product_subcategories.csv
- c:\users\chris\documents\secon...ks\adventureworks_products.csv
- c:\users\chris\documents\secon...rks\adventureworks_returns.csv
- c:\users\chris\documents\secon...works\adventureworks_sales.csv
- c:\users\chris\documents\secon...adventureworks_territories.csv
- c:\users\chris\documents\secon...data\adventureworks\aw_sales

Change Source... Edit Permissions... Clear Permissions...

Comma-Separated Values

★ Basic ○ Advanced

File path
c:\users\chris\Desktop\Power BI Course Files\Adventure Works\Adventure Browse...

Open file as
Csv Document

File origin
1252: Western European (Windows)

Line breaks
Apply all line breaks

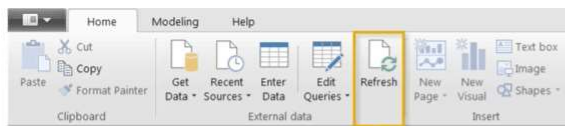
Delimiter
Comma

IMPORTANT!

Les connexions aux fichiers locaux font référence au chemin exact. Si le nom ou l'emplacement du fichier change, vous devrez **changer la source et accéder à la version actuelle**

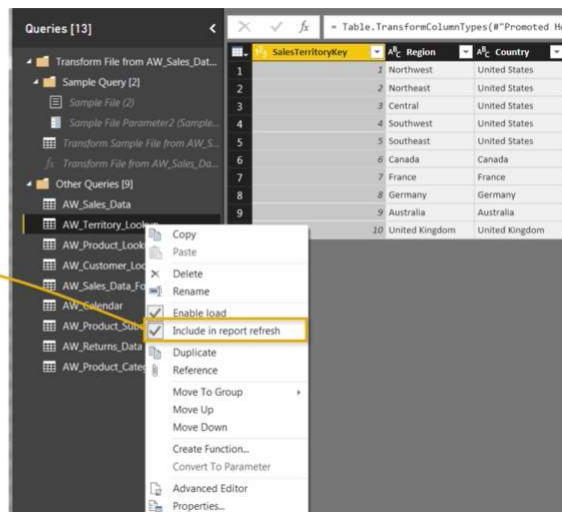
Les **paramètres de source de données** dans l'éditeur de requête vous permettent de gérer les connexions de données et les autorisations

RAFRAÎCHISSEMENT DE REQUÊTES



Par défaut, **TOUTES** les requêtes du modèle sont actualisées, lorsque vous utilisez l'option «Actualiser» (Refresh) de l'onglet Accueil

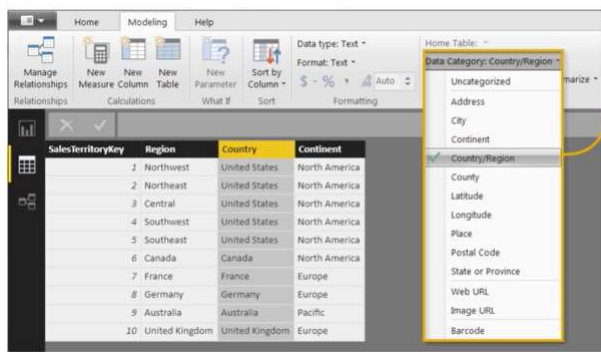
Dans l'éditeur de requêtes, décochez «Inclure dans l'actualisation du rapport» pour exclure les requêtes individuelles de l'actualisation



PRO TIP:

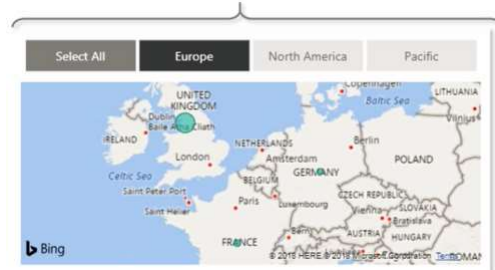
Exclure les requêtes qui ne changent pas souvent, comme des tables de lookups ou de données statiques

DÉFINITION DES CATÉGORIES DE DONNÉES



Dans l'onglet «**Modélisation**» de la vue Données, vous pouvez modifier les propriétés des champs pour définir des catégories spécifiques

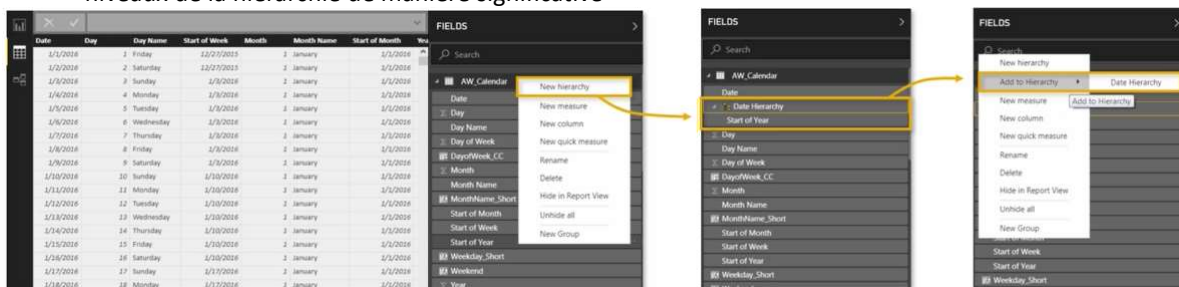
Ceci est couramment utilisé pour aider Power BI à cartographier avec précision les champs basés sur la localisation tels que les adresses, les pays, les villes, les coordonnées de latitude/longitude, les codes postaux, etc.



DÉFINITION DES HIÉRARCHIES

Les **hiérarchies** sont des groupes de colonnes imbriquées qui reflètent plusieurs niveaux de granularité

- Par exemple, une hiérarchie «**Géographie**» peut inclure des colonnes **Pays**, **État** et **Ville**
- Chaque hiérarchie peut être traitée comme un champs unique dans les tables et les rapports, permettant aux utilisateurs de «forer» en haut et en bas ("drill up" and "drill down") à différents niveaux de la hiérarchie de manière significative

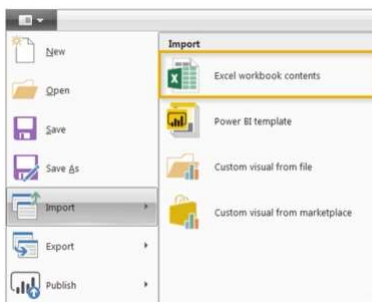


1) Dans la vue Données, cliquez avec le bouton droit sur un champ (ou cliquez sur les points de suspension) et sélectionnez "Nouvelle hiérarchie" (ici, nous avons sélectionné "Début de l'année")

2) Cela crée un champ de hiérarchie contenant "Start of Year", que nous avons renommé "**Date Hierarchy**"

3) Faites un clic droit sur d'autres champs (comme «Début du mois») et sélectionnez «Ajouter à la hiérarchie»

CONSEIL DE PRO: IMPORTATION DE MODÈLES D'EXCEL



Vous avez déjà un modèle entièrement construit dans Excel ?

Importez des fichiers créés avec Power Query / Power Pivot directement dans Power BI Desktop à l'aide de **Importer > Contenu du classeur Excel**

Les modèles importés conservent les éléments suivants :

- **Connexions et requêtes de source de données**
- **Procédures d'édition de requête et étapes appliquées**



PRO TIP:

Power Pivot comprend certaines fonctionnalités que Power BI ne propose pas (options de filtrage, aide sur les fonctions DAX, etc.) ; si vous êtes plus à l'aise dans l'environnement Excel, créez-y vos modèles, puis importez-les dans Power BI!

BONNES PRATIQUES : CONNEXION & MISE EN FORME



Organisez-vous avant de charger les données dans Power BI

- Définissez des noms de table clairs et intuitifs (pas d'espaces!) Dès le début; les mettre à jour plus tard peut être un casse-tête, en particulier si vous les avez référencés à plusieurs endroits.

- Établissez une structure de fichiers / dossiers qui a du sens dès le départ, pour éviter d'avoir à modifier les paramètres de la source de données si les noms ou les emplacements des fichiers changent



Désactivation de l'actualisation des rapports pour toutes les sources statiques

- Il n'est pas nécessaire d'actualiser en permanence les sources qui ne sont pas mises à jour fréquemment (ou pas du tout), comme les tables de lookups ou de données statiques ; activer uniquement l'actualisation pour les tables qui seront modifiées



Lorsque vous travaillez avec de grandes tables, chargez uniquement les données dont vous avez besoin

- des données supplémentaires ne feront que vous ralentir

HOMEWORK Connexion et mise en forme des données