

Dans un nouveau rapport PBI, charger des données excel, et sélectionner le fichier “Dataset_AdventureWorks” fourni dans le dossier “Exercices DAX”.

Sélectionner toutes les tables SAUF “DimProductCategory1”, “dimProductCategory” et “dimProductSubCategory”.

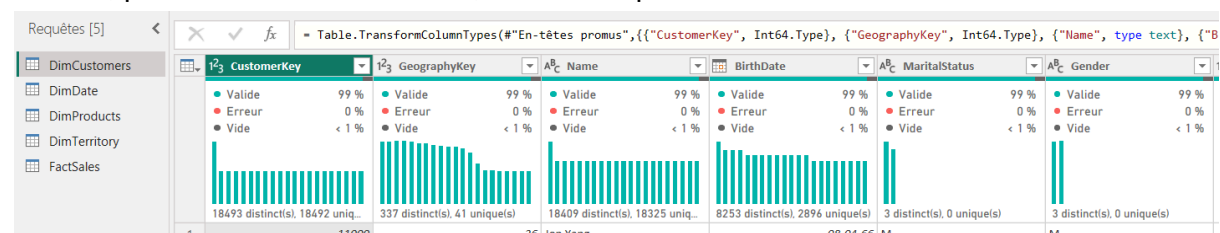
The screenshot shows the Power BI Navigator window. On the left, under 'Options d'affichage', the 'Dataset_AdventureWorks.xlsx [8]' folder is expanded. The following tables are listed with checkboxes: DimCustomers (checked), DimDate (checked), dimProductCategory (unchecked), DimProducts (checked), dimProductSubCategory (unchecked), DimTerritory (checked), FactSales (checked and highlighted), and dimProductCategory1 (unchecked). On the right, the 'FactSales' table is previewed. The preview title is 'FactSales' and it says 'Aperçu téléchargé le mardi 15 octobre 2024'. The table has 6 columns: ProductKey, OrderDate, ShipDate, OrderDateKey, CustomerKey, and SalesTr. The data shows a list of sales transactions with dates ranging from 29-09-03 to 27-10-03. At the bottom right, there are three buttons: 'Charger', 'Transformer les données', and 'Annuler'.

ProductKey	OrderDate	ShipDate	OrderDateKey	CustomerKey	SalesTr
528	29-09-03	02-10-03	20030929	16115	
528	01-10-03	04-10-03	20031001	15307	
528	05-10-03	08-10-03	20031005	16003	
528	06-10-03	09-10-03	20031006	15883	
528	08-10-03	11-10-03	20031008	15368	
528	08-10-03	11-10-03	20031008	15273	
528	08-10-03	11-10-03	20031008	11499	
528	08-10-03	11-10-03	20031008	12960	
528	10-10-03	13-10-03	20031010	15875	
528	11-10-03	14-10-03	20031011	16238	
528	11-10-03	14-10-03	20031011	11695	
528	12-10-03	15-10-03	20031012	23728	
528	13-10-03	16-10-03	20031013	11694	
528	14-10-03	17-10-03	20031014	26335	
528	14-10-03	17-10-03	20031014	14399	
528	15-10-03	18-10-03	20031015	11807	
528	16-10-03	19-10-03	20031016	14903	
528	17-10-03	20-10-03	20031017	16143	
528	19-10-03	22-10-03	20031019	15344	
528	21-10-03	24-10-03	20031021	14920	
528	23-10-03	26-10-03	20031023	15305	
528	24-10-03	27-10-03	20031024	15946	

Puis sélectionner “Transformer les données” afin d’accéder à Power Query.

Commencer par demander un profilage sur l’ensemble du jeu de données plutôt que sur les 1000 premières lignes de la table (option en bas de page dans Query).

Ensuite, pour l’ensemble des tables, vérifier la “qualité de colonne”



Ainsi on constate, pour “DimCustomers” par exemple, que certaines lignes sont complètement vides (si vous n’avez pas accès à la qualité de colonne, aller dans l’onglet affichage de power query et cocher la case “qualité de la colonne”).

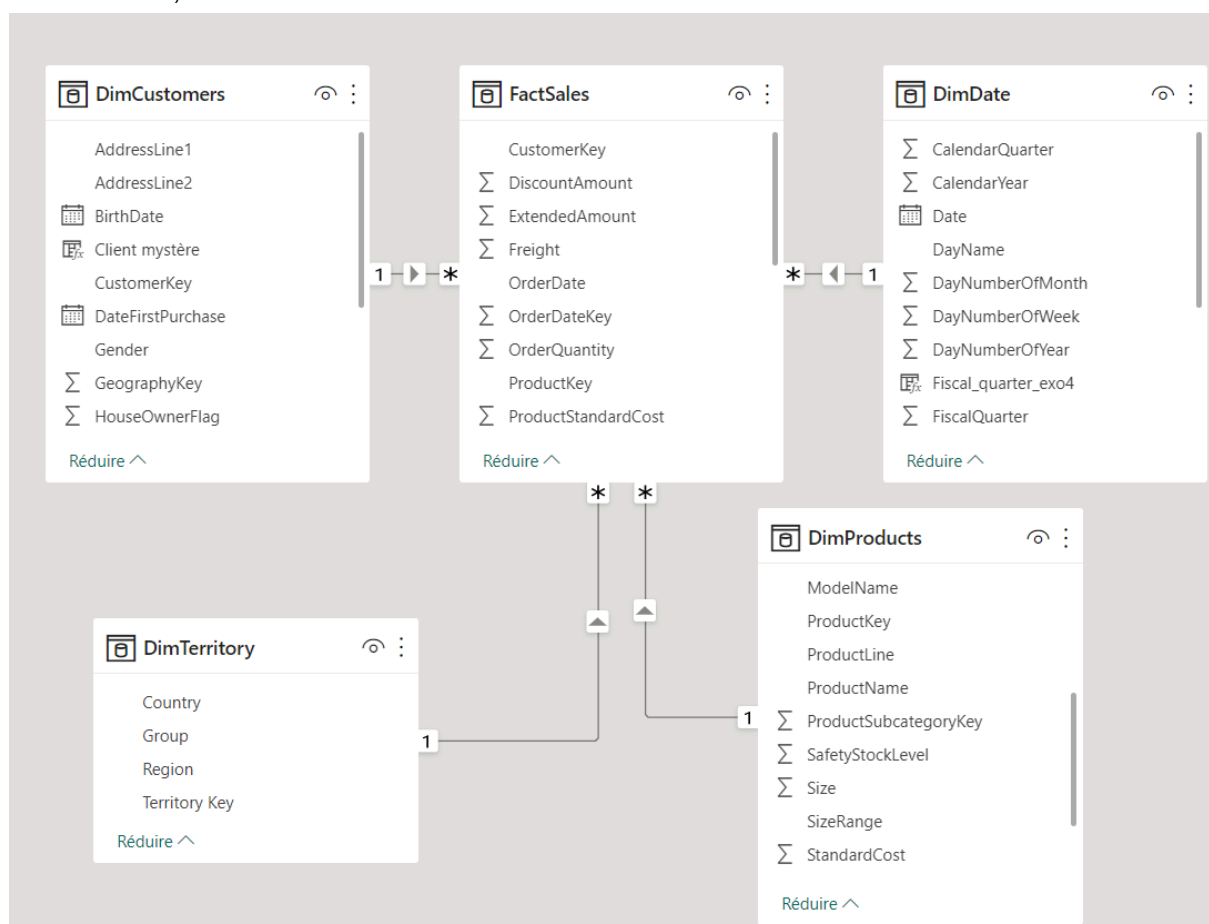
Pour retirer ces lignes vides, dans l'onglet "Accueil", partie "Réduire les lignes" -> "Supprimer les lignes" -> "Supprimer les lignes vides". Cela doit être fait sur toutes les tables dimensions.

Par ailleurs, dans la table "DimCustomers", sélectionner la colonne "CustomerKey" et retirer également les doublons. ("Réduire les lignes" -> "Supprimer les lignes" -> "Supprimer les doublons").

Une fois ces opérations effectuées, vous pouvez fermer et appliquer (sans vous tracasser des erreurs signalées au sein de la table "DimProduct") pour retourner sur Power BI.

Cela fait, sélectionner la vue de modèle afin de relier la tables de faits au dimensions qui n'auraient pas été liées automatiquement par le logiciel.

Vous devriez, à la fin avoir une vue de modèle semblable à celle-ci :



Si certaines dimensions restent isolées, il "suffit" de drag & drop la colonne qui sert de référence au sein de la table "FactSales" sur la colonne "key" de la table de dimension non reliée.

Les noms des colonnes servant aux différentes relations sont les suivants :

Entre FactSales et DimDate : "OrderDate" côté FactSales et "Date" côté DimDate.

Entre FactSales et DimCustomers : "CustomerKey" côté FactSales et "CustomerKey" côté DimCustomers.

Entre FactSales et DimTerritory : "SalesTerritoryKey" côté FactSales et "TerritoryKey" côté DimTerritory.

Entre FactSales et DimProducts : "ProductKey" côté FactSales et "ProductKey" côté DimProducts.