

Lab02 : Charger et transformer les données dans Power BI Desktop

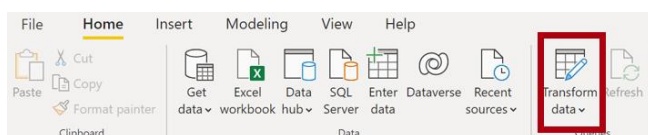
Dans cet atelier, vous commencerez à appliquer des **transformations** à chacune des requêtes créées dans l'atelier précédent. Vous appliquerez ensuite les requêtes pour charger chacune en tant que table dans le modèle de données.

Dans cet atelier, vous apprendrez à :

- Appliquer diverses transformations
- Appliquer des requêtes pour les charger dans le modèle de données

Editeur Power Query

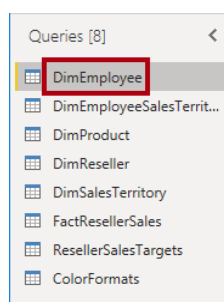
1. Pour ouvrir la fenêtre de l'éditeur **Power Query**, dans l'onglet **Accueil** du ruban, depuis le groupe **Requêtes**, cliquez sur l'icône **Transformer les données**.



Configurer la requête du Salesperson

Dans cette tâche, vous allez configurer la requête **Salesperson**.

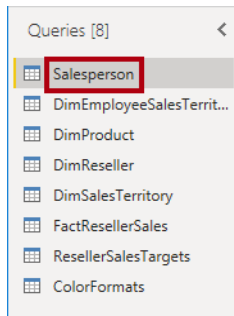
2. Dans la fenêtre de l'éditeur **Power Query**, dans le volet à gauche **Requêtes**, sélectionnez la requête **DimEmployee**.



3. Pour renommer la requête, dans le volet **Paramètres** de la requête (situé à droite), dans la zone **Nom**, remplacez le nom par défaut de la requête par **Salesperson**, puis appuyez sur **Entrée**.

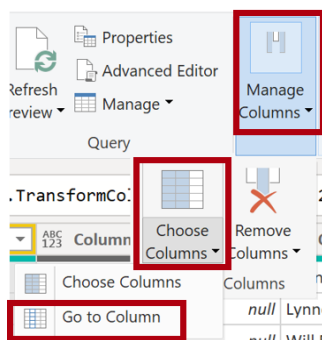
Le nom de la requête déterminera le nom de la table modèle. Il est recommandé de définir des noms concis, mais conviviaux.

4. Dans le volet **Requêtes**, vérifiez que le nom de la requête a été mis à jour.



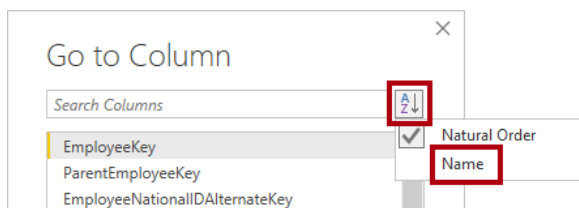
Vous allez maintenant filtrer les lignes de la requête pour ne récupérer que les employés qui sont des commerciaux.

5. Pour localiser une colonne spécifique, dans l'onglet **Accueil** du ruban, cliquez sur la flèche vers le bas **Gérer les colonnes**, cliquez sur la flèche vers le bas **Choisir les colonnes**, puis sélectionnez **Accéder à la colonne**.

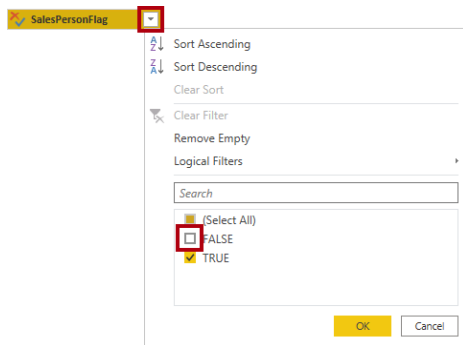


Conseil : Cette technique est utile lorsqu'une requête (une table) contient de nombreuses colonnes. S'il n'y a pas trop de colonnes, vous pouvez simplement faire défiler horizontalement pour localiser la colonne qui vous intéresse.

6. Dans la fenêtre **Aller à la colonne**, pour trier la liste par nom de colonne, cliquez sur le bouton de tri **AZ**, puis sélectionnez **Nom**.

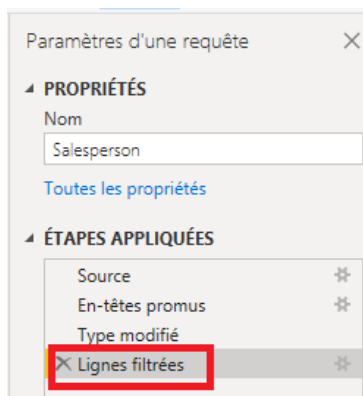


7. Sélectionnez la colonne **SalesPersonFlag**, puis cliquez sur **OK**.
8. Pour filtrer la requête, dans l'en-tête de colonne **SalesPersonFlag**, cliquez sur la flèche vers le bas, puis décochez **FALSE**. Cette personne permet de spécifier si l'employé est un commercial ou pas.



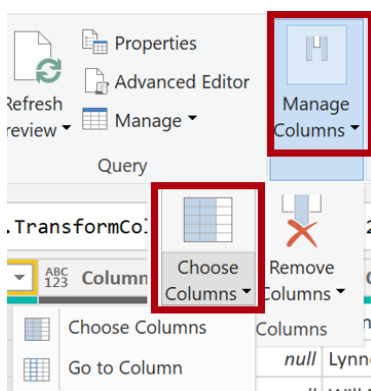
9. Cliquez sur **OK**.

10. Dans le volet **Paramètres de la requête**, dans la liste **Étapes appliquées**, notez l'ajout de l'étape **Lignes filtrées**.

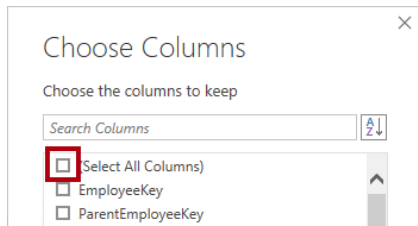


Chaque transformation que vous créez entraîne une logique d'étape supplémentaire. Il est possible de modifier ou de supprimer des étapes. Il est également possible de sélectionner une étape pour pré-visualiser les résultats de la requête à ce stade de la transformation de la requête.

11. Pour supprimer des colonnes, dans l'onglet **Accueil** du ruban, cliquez sur le groupe **Gérer les colonnes**, cliquez sur l'icône **Choisir les colonnes**.



12. Dans la fenêtre **Choisir les colonnes**, pour décocher toutes les colonnes, décochez l'élément **(Sélectionner toutes les colonnes)**.

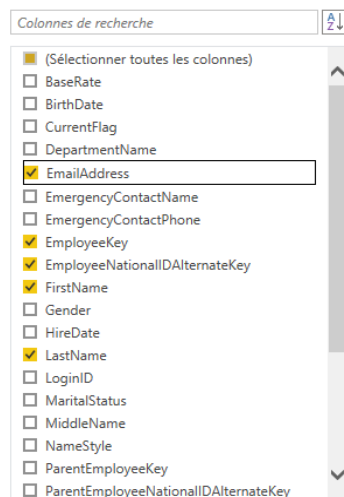


13. Pour inclure des colonnes et ne garder que les colonnes disponibles au niveau de la liste suivante, cochez les six colonnes suivantes. Vous pouvez faire un tri sur la liste des attributs et suivre l'ordre de la liste :

- EmailAdress
- EmployeeKey
- EmployeeNationalIDAlternateKey
- FirstName
- LastName
- Title

Choisir les colonnes

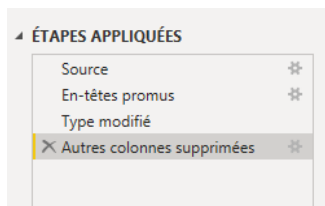
Choisir les colonnes à conserver



OK Annuler

14. Cliquez sur **OK**.

15. Dans la liste **Étapes appliquées**, notez l'ajout d'une autre étape de requête.

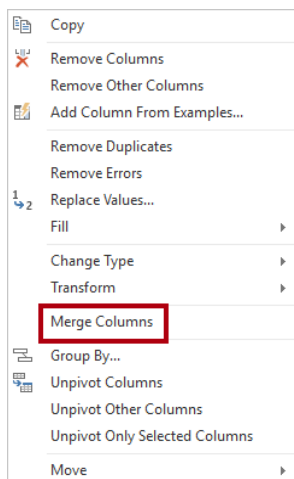


16. Pour créer une colonne de nom unique, sélectionnez d'abord l'en-tête de colonne **FirstName**.

17. Tout en appuyant sur la touche **Ctrl**, sélectionnez la colonne **LastName**.

EmployeeNationalIDAlternateKey	FirstName	LastName	Title
502097814	Stephen	Jiang	North American Sales Manager
112432117	Brian	Welcker	Director of Sales
841560125	Michael	Blythe	Sales Representative

18. Cliquez avec le bouton droit sur l'un des en-têtes de colonne sélectionnés, puis dans le menu contextuel, sélectionnez **Fusionner les colonnes**.



De nombreuses transformations courantes peuvent être appliquées en cliquant avec le bouton droit sur l'en-tête de colonne, puis en les choisissant dans le menu contextuel.

19. Dans la fenêtre **Fusionner les colonnes**, dans la liste déroulante **Séparateur**, sélectionnez **Espace**.

20. Dans la zone Nouveau nom de colonne, remplacez le texte par **Salesperson**.

Merge Columns

Choose how to merge the selected columns.

Separator: Space

New column name (optional): Salesperson

21. Cliquez sur **OK**.

22. Les deux colonnes ont été remplacées par une nouvelle colonne **Salesperson**.

Salesperson	Title
Stephen Jiang	North American Sales Manager
Brian Welcker	Director of Sales
Michael Blythe	Sales Representative
Linda Mitchell	Sales Representative
Jillian Carson	Sales Representative
Garrett Vargas	Sales Representative
Tsvi Reiter	Sales Representative
Pamela Ansman-Wolfe	Sales Representative
Shu Ito	Sales Representative
José Saraiva	Sales Representative
David Campbell	Sales Representative
Amy Alberts	European Sales Manager
Jae Pak	Sales Representative
Ranjit Varkey Chudukatil	Sales Representative
Tete Mensa-Annan	Sales Representative
Syed Abbas	Pacific Sales Manager
Rachel Valdez	Sales Representative
Lynn Tsofilas	Sales Representative

23. Pour renommer la colonne **EmployeeNationalIDAlternateKey**, double-cliquez sur l'en-tête de cette colonne.

24. Remplacez le texte par **EmployeeID**, puis appuyez sur **Entrée**.

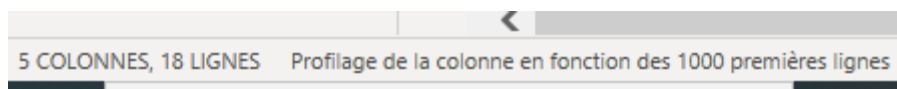
rees-{{EmployeeNationalIDAlternateKey}}, EmployeeID

EmployeeID	Salesperson
502097814	Stephen Jiang
112432117	Brian Welcker
841560125	Michael Blythe
191644724	Linda Mitchell
615389812	Lillian Carson
234474252	Garrett Vargas
716374314	Yvonne Reiter
61161660	Pamela Ansman-Wolfe
139397894	Shu Ito
399771412	José Saraiva
987554265	David Campbell
982310417	Amy Alberts
668991357	Chae Pak
134219713	Ranjit Varkey Chudakatil
90836195	Pete Mensa-Annan
481044938	Syed Abbas
954276278	Rachel Valdez
758596752	Lynn Tsofilias

Important : Lorsque vous êtes invité à renommer les colonnes, il est important de les renommer exactement comme décrit car on va les utiliser par la suite.

25. Utilisez les étapes précédentes pour renommer la colonne **EmailAddress** en **UPN** pour l'acronyme de User Principal Name.

26. En bas à gauche, dans la barre d'état, vérifiez que la table **SalesPerson** comporte cinq colonnes et 18 lignes.

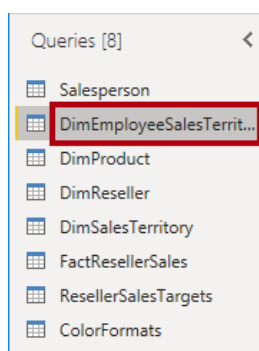


Important : Il est important que vous ne poursuiviez pas si votre requête ne produit pas le bon résultat. Il ne sera pas possible de terminer les ateliers ultérieurs. Si les colonnes ou les lignes de la requête ne correspondent pas, reportez-vous aux étapes de cette tâche pour résoudre les problèmes.

Configurer la requête SalespersonRegion

Dans cette partie, vous allez configurer la requête **SalespersonRegion**.

27. Dans le volet Requêtes, sélectionnez la requête **DimEmployeeSalesTerritory**.

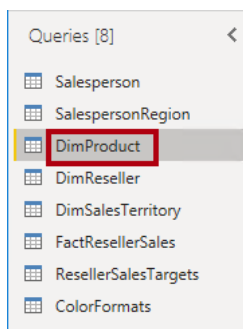


28. Dans le volet **Paramètres de la requête**, renommez la requête en **SalespersonRegion**.
29. A quoi correspond cette table ?
30. Quelle cardinalité on peut attribuer aux colonnes disponibles au niveau de cette table à savoir **EmployeeKey** et **SalesTerritoryKey**.
31. Dans la barre d'état, vérifiez que la requête comporte deux colonnes et 39 lignes.

Configurer la requête Produit

Dans cette partie, vous allez configurer la requête **Product**.

32. Sélectionnez la requête **DimProduct**.



33. Renommez la requête en **Product**.
34. Localisez la colonne **FinishedGoodsFlag**, puis filtrez la colonne pour récupérer les produits qui sont des produits finis (c'est-à-dire **TRUE**).
35. Supprimez toutes les colonnes, à l'exception des colonnes suivantes :

- ProductKey
- EnglishProductName
- StandardCost
- Color
- EnglishProductSubcategoryName
- EnglishProductCategoryName

36. Renommez les quatre colonnes suivantes :

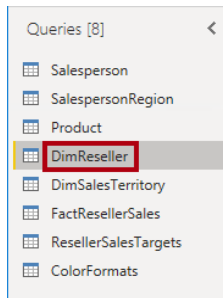
- **EnglishProductName** to **Product**
- **StandardCost** to **Standard Cost** (include a space)
- **EnglishProductSubcategoryName** to **Subcategory**
- **EnglishProductCategoryName** to **Category**

37. Dans la barre d'état, vérifiez que la requête comporte six colonnes et 397 lignes.

Configurer la requête DimReseller

Dans cette tâche, vous allez configurer la requête **DimReseller**.

38. Sélectionnez la requête **DimReseller**.



39. Renommez la requête **Reseller**.

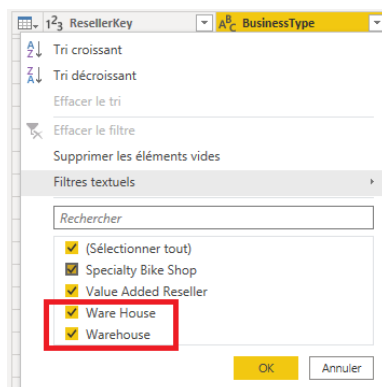
40. Supprimez toutes les colonnes, à l'exception des suivantes :

- ResellerKey
- BusinessType
- ResellerName
- DimGeography.City
- DimGeography.StateProvinceName
- DimGeography.EnglishCountryRegionName

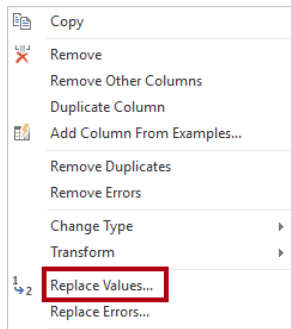
41. Renommer les colonnes suivantes :

- DimGeography.City → City.
- DimGeography.StateProvinceName → StateProvinceName
- DimGeography.EnglishCountryRegionName → EnglishCountryRegionName

42. Dans l'en-tête de colonne **BusinessType**, cliquez sur la flèche vers le bas, puis passez en revue les valeurs de colonne distinctes et notez l'orthographe incorrecte de **Warehouse**.

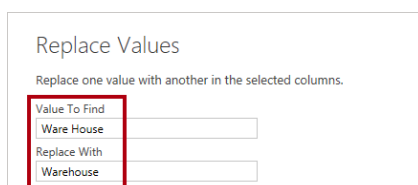


43. Cliquez avec le bouton droit sur l'en-tête de colonne **BusinessType**, puis sélectionnez **Remplacer les valeurs**.



44. Dans la fenêtre **Remplacer les valeurs**, configurez les valeurs suivantes :

- Dans la zone **Valeur à rechercher**, entrez **Ware House**
- Dans la zone **Remplacer par**, entrez **Warehouse**



45. Cliquez sur **OK**.

46. Renommez les quatre colonnes suivantes :

- BusinessType to Business Type (include a space)
- ResellerName to Reseller
- StateProvinceName to State-Province
- EnglishCountryRegionName to Country-Region

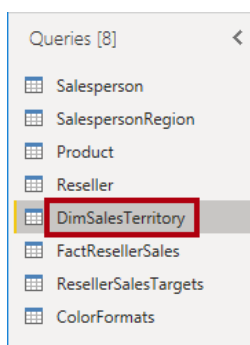
47. Dans la barre d'état, vérifiez que la requête comporte six colonnes et 701 lignes.

6 COLUMNS, 701 ROWS Column profiling based on top 1000 rows

Configurer la requête Region

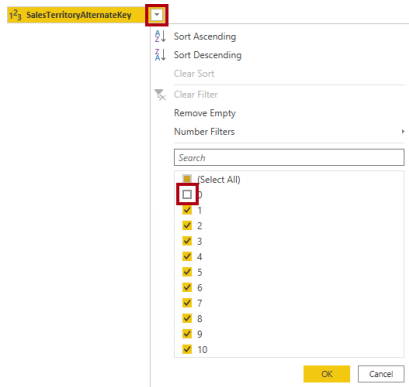
Dans cette tâche, vous allez configurer la requête **Region**.

48. Sélectionnez la requête **DimSalesTerritory**.



49. Renommez la requête en **Region**.

50. Appliquez un filtre à la colonne **SalesTerritoryAlternateKey** pour supprimer la valeur 0 (zéro).



51. Supprimez toutes les colonnes, à l'exception des suivantes :

- SalesTerritoryKey
- SalesTerritoryRegion
- SalesTerritoryCountry
- SalesTerritoryGroup

52. Renommez les trois colonnes suivantes :

- SalesTerritoryRegion → Region
- SalesTerritoryCountry → Country
- SalesTerritoryGroup → Group

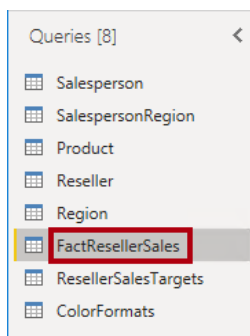
53. Dans la barre d'état, vérifiez que la requête comporte quatre colonnes et 10 lignes.

4 COLUMNS, 10 ROWS Column profiling based on top 1000 rows

Configurer la requête Sales

Dans cette tâche, vous allez configurer la requête Sales.

54. Sélectionnez la requête **FactResellerSales**.



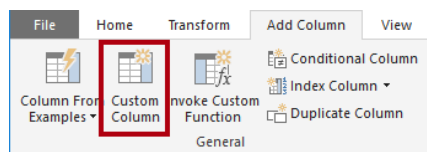
55. Renommez la requête en **Sales**.

56. Supprimez toutes les colonnes, à l'exception des suivantes :

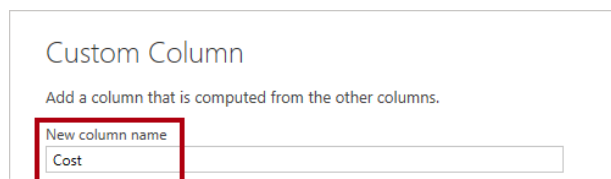
- SalesOrderNumber

- OrderDate
- ProductKey
- ResellerKey
- EmployeeKey
- SalesTerritoryKey
- OrderQuantity
- UnitPrice
- TotalProductCost
- SalesAmount
- StandardCost

57. Pour créer une **Colonne personnalisée**, dans l'onglet de ruban **Ajouter une colonne**, depuis le groupe **Général**, cliquez sur **Colonne personnalisée**.



58. Dans la fenêtre **Colonne personnalisée**, dans la zone **Nouveau nom de colonne**, remplacez le texte par **Cost**.



59. Dans la zone **Formule de colonne personnalisée**, saisissez l'expression suivante (après le symbole égal) :

```
if [TotalProductCost] = null then [OrderQuantity] * [StandardCost]
else [TotalProductCost]
```

Cette expression teste si la valeur **TotalProductCost** est manquante. Si c'est le cas, produit une valeur en multipliant la valeur **OrderQuantity** par la valeur **StandardCost** ; sinon, il utilise la valeur **TotalProductCost** existante.

60. Cliquez sur **OK**.

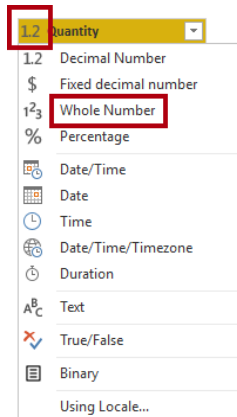
61. Supprimez les deux colonnes suivantes :

- TotalProductCost
- StandardCost

62. Renommez les trois colonnes suivantes :

- OrderQuantity to Quantity
- UnitPrice to Unit Price (include a space)
- SalesAmount to Sales

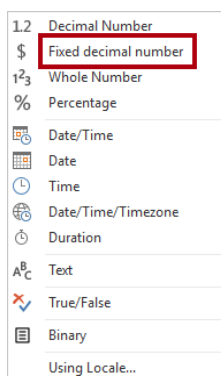
63. Pour modifier le type de données de colonne, dans l'en-tête de colonne **Quantity**, à gauche du nom de la colonne, cliquez sur l'icône **1.2**, puis sélectionnez **Nombre entier**.



La configuration du type de données correct est importante. Lorsque la colonne contient une valeur numérique, il est également important de choisir le type correct si vous prévoyez d'effectuer des calculs mathématiques.

64. Modifiez les trois types de données de colonne suivants en **nombre décimal fixe**.

- Unit Price
- Sales
- Cost



Le type de données à nombre décimal fixe stocke les valeurs avec une précision totale et nécessite donc plus d'espace de stockage que le nombre décimal. Il est important d'utiliser le type de nombre décimal fixe pour les valeurs financières ou les taux (comme les taux de change).

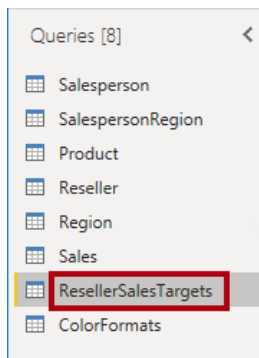
65. Dans la barre d'état, vérifiez que la requête comporte 10 colonnes et plus de 999 lignes.

Un maximum de 1 000 lignes seront chargées en tant que données d'aperçu pour chaque requête.

Configurer la requête Targets

Dans cette tâche, vous allez configurer la requête **Targets**.

66. Sélectionnez la requête **ResellerSalesTargets**.



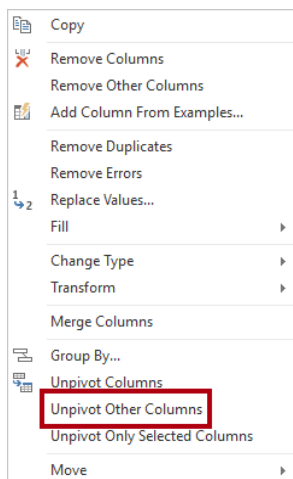
67. Renommez la requête en **Targets**.

68. Pour annuler le pivot des colonnes de 12 mois (**M01-M12**), sélectionnez d'abord les en-têtes de colonne **Year** et **EmployeeID**.

The image shows a data table with columns: Year, EmployeeID, M01, and M02. The first two columns are highlighted with a red rectangle. The data rows show values for Year (2017) and EmployeeID (81161680, 90836195, 112432117, 139397894, 191644724, 234474252). The M01 and M02 columns contain dashes (-).

	1 ² Year	1 ² EmployeeID	A ^B M01	A ^B M02
1	2017	81161680	-	-
2	2017	90836195	-	-
3	2017	112432117	-	-
4	2017	139397894	-	-
5	2017	191644724	-	-
6	2017	234474252	-	-

69. Cliquez avec le bouton droit sur l'un des en-têtes de colonne de sélection, puis dans le menu contextuel, sélectionnez **Dépivoter les autres colonnes**.



70. Notez que les noms de colonne apparaissent maintenant dans la colonne **Attributes** et que les valeurs apparaissent dans la colonne **Value**.

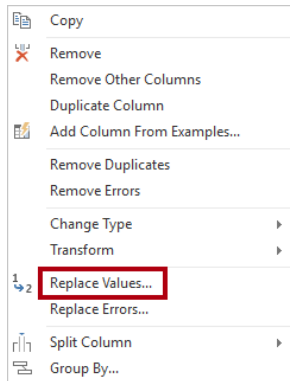
71. Appliquez un filtre à la colonne **Value** pour supprimer le trait d'union (-). Vous vous souviendrez peut-être que le trait d'union a été utilisé dans le fichier CSV source pour représenter zéro (0).

72. Renommez les deux colonnes suivantes :

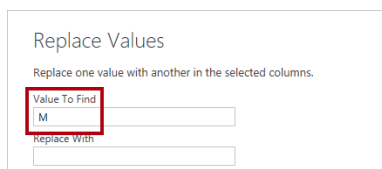
- Attribute → **MonthNumber**
- Value → **Target**

Vous allez maintenant appliquer des transformations pour produire une colonne de date. La date sera dérivée des colonnes **Year** et **MonthNumber**. Vous allez créer la colonne à l'aide de la fonctionnalité **Colonnes à partir d'exemples**.

73. Pour préparer les valeurs de la colonne **MonthNumber**, cliquez avec le bouton droit sur l'en-tête de la colonne **MonthNumber**, puis sélectionnez **Remplacer les valeurs**.

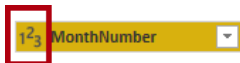


74. Dans la fenêtre **Remplacer les valeurs**, dans la zone **Valeur à rechercher**, saisissez **M**.

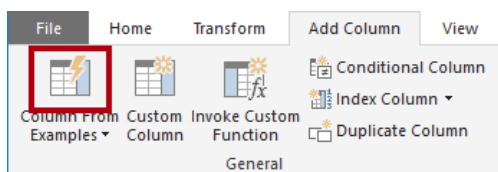


75. Cliquez sur **OK**.

76. Remplacez le type de données de la colonne **MonthNumber** par **Whole Number**.



77. Dans l'onglet de ruban **Ajouter une colonne**, à partir du groupe **Général**, cliquez sur l'icône **Colonne à partir d'exemples**.



78. Notez que la première ligne correspond à l'année **2017** et au mois numéro **7**.

79. Dans la colonne **Colonne1**, dans la première cellule de la grille, commencez à saisir **01/7/2017**, puis appuyez sur **Entrée**.

80. Notez que les cellules de la grille sont mises à jour avec les valeurs prédites.

La fonctionnalité a prédit avec précision que vous combinez les valeurs des colonnes **Year** et **MonthNumber**.

81. Pour renommer la nouvelle colonne, double-cliquez sur l'en-tête de colonne **Fusionné**.

82. Renommez la colonne **TargetMonth**.

TargetMonth
7/1/2017
8/1/2017
9/1/2017
10/1/2017
11/1/2017
12/1/2017
7/1/2017

83. Pour ajouter la nouvelle colonne, cliquez sur **OK**.

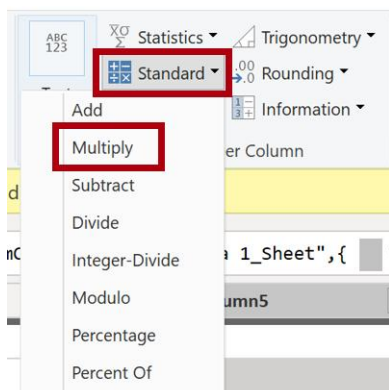
84. Supprimez les colonnes suivantes :

- Year
- MonthNumber

85. Modifiez les types de données de colonne suivants :

- Target → fixed decimal number
- TargetMonth → date

86. Pour multiplier les valeurs **Target** par 1 000, sélectionnez l'en-tête de colonne **Target**, puis dans l'onglet **Transformer** du ruban, à l'intérieur du groupe **Colonne numérique**, cliquez sur **Standard**, puis sélectionnez **Multiplier**.



87. Dans la fenêtre **Multiplier**, dans la zone Valeur, entrez **1000**.

 A screenshot of the 'Multiply' dialog box in Power Query. The title bar says 'Multiply'. Below the title, it says 'Enter a number by which to multiply each value in the column.' There is a text input field labeled 'Value' which contains the number '1000'. The 'Value' label and the input field are highlighted with a red box.

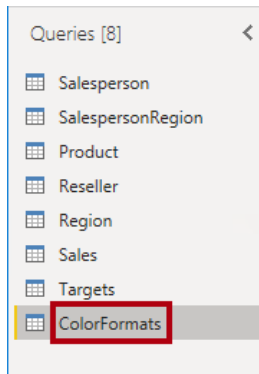
88. Cliquez sur **OK**.

89. Dans la barre d'état, vérifiez que la requête comporte trois colonnes et 809 lignes.

Configurer la requête ColorFormats

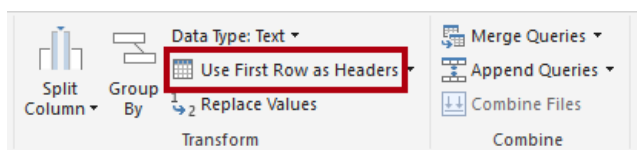
Dans cette tâche, vous allez configurer la requête **ColorFormats**.

90. Sélectionnez la requête **ColorFormats**.



91. Notez que la première ligne contient les noms de colonne.

92. Dans l'onglet **Accueil** du ruban, depuis le groupe **Transformer**, cliquez sur **Utiliser la première ligne comme en-têtes**.

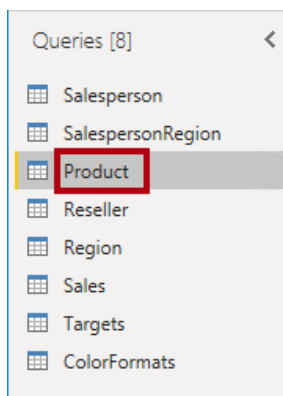


93. Dans la barre d'état, vérifiez que la requête comporte trois colonnes et 10 lignes.

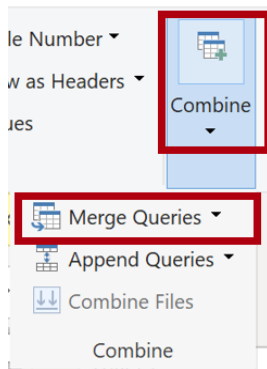
Mettre à jour la requête Product

Dans cette tâche, vous allez mettre à jour la requête **Product** en fusionnant la requête **ColorFormats**.

94. Sélectionnez la requête **Product**.



95. Pour fusionner la requête **ColorFormats**, dans l'onglet du ruban **Accueil**, cliquez sur la flèche vers le bas **Combiner**, puis cliquez sur **Fusionner les requêtes**.



La fusion des requêtes permet d'intégrer des données, en l'occurrence de différentes sources de données (SQL Server et un fichier CSV).

96. Dans la fenêtre **Fusionner**, dans la grille de requête **Product**, sélectionnez l'en-tête de colonne **Color**.

Merge

Select a table and matching columns to create a merged table.

Product

ProductKey	Product	Standard Cost	Color	Subcategory	Category
210	HL Road Frame - Black, 58	868.63	Black	Road Frames	Components
211	HL Road Frame - Red, 58	868.63	Red	Road Frames	Components
212	Sport-100 Helmet, Red	12.03	Red	Helmets	Accessories
213	Sport-100 Helmet, Red	13.88	Red	Helmets	Accessories
214	Sport-100 Helmet, Red	13.09	Red	Helmets	Accessories

97. Sous la grille de requête **Product**, dans la liste déroulante, sélectionnez la requête **ColorFormats**.

Merge

Select a table and matching columns to create a merged table.

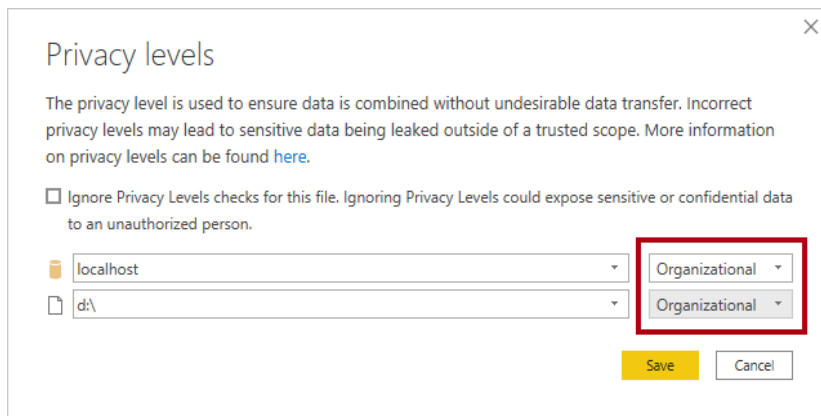
Product

ProductKey	Product	Standard Cost	Color	Subcategory	Category
210	HL Road Frame - Black, 58	868.63	Black	Road Frames	Components
211	HL Road Frame - Red, 58	868.63	Red	Road Frames	Components
212	Sport-100 Helmet, Red	12.03	Red	Helmets	Accessories
213	Sport-100 Helmet, Red	13.88	Red	Helmets	Accessories
214	Sport-100 Helmet, Red	13.09	Red	Helmets	Accessories

ColorFormats

98. Dans la grille de requête **ColorFormats**, sélectionnez l'en-tête de colonne **Color**.

99. Lorsque la fenêtre **Niveaux de confidentialité** s'ouvre, pour chacune des deux sources de données, dans la liste déroulante correspondante, sélectionnez **Organisationnel**.



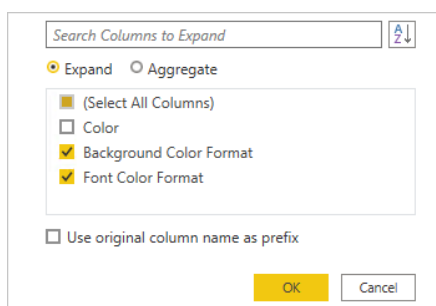
Les niveaux de confidentialité peuvent être configurés pour la source de données afin de déterminer si les données peuvent être partagées entre les sources. Définir chaque source de données comme **Organisational** leur permet de partager des données, si nécessaire. Notez que les sources de données privées ne peuvent jamais être partagées avec d'autres sources de données.

100. Cliquez sur **Enregistrer**.

101. Dans la fenêtre de fusion, utilisez le type de jointure par défaut - en conservant la sélection de **Left Outer** et cliquez sur **OK**.

102. Développez la colonne **ColorFormats** pour inclure les deux colonnes suivantes :

- Background Color Format
- Font Color Format

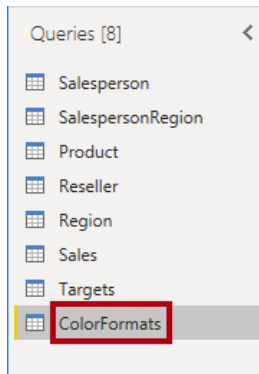


103. Dans la barre d'état, vérifiez que la requête comporte maintenant huit colonnes et 397 lignes.

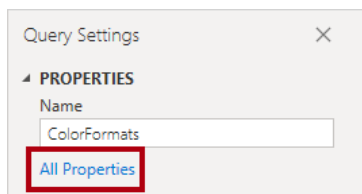
Mettre à jour la requête ColorFormats

Dans cette partie, vous mettrez à jour le **ColorFormats** pour désactiver son chargement.

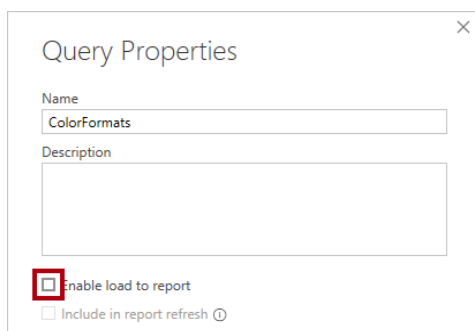
104. Sélectionnez la requête **ColorFormats**.



105. Dans le volet **Paramètres de la requête**, cliquez sur le lien **Toutes les propriétés**.



106. Dans la fenêtre **Propriétés de la requête**, décochez la case **Activer le chargement dans le rapport**.



La désactivation du chargement signifie qu'il ne se chargera pas en tant que table dans le modèle de données. Cela est dû au fait que la requête a été fusionnée avec la requête **Product**, qui est activée pour se charger dans le modèle de données.

107. Cliquez sur **OK**.

Terminer

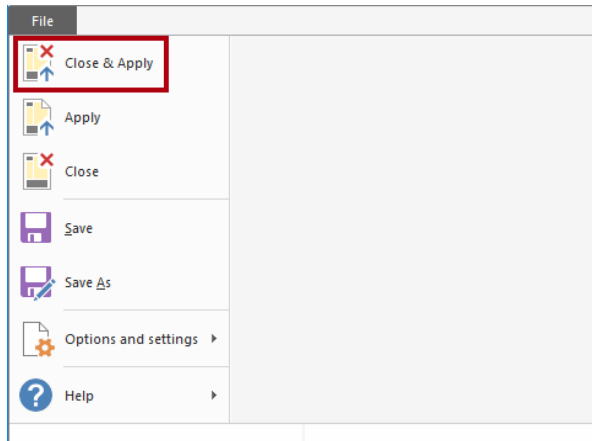
Dans cette tâche, vous terminerez le laboratoire.

108. Vérifiez que vous disposez de huit requêtes correctement nommées comme suit :

- Salesperson
- SalespersonRegion
- Product
- Reseller
- Region
- Sales
- Targets

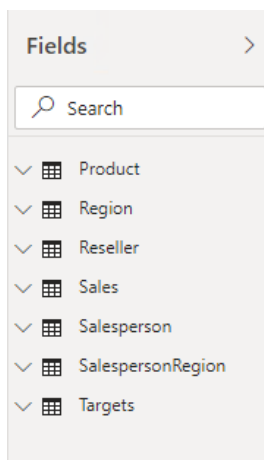
- ColorFormats (which will not load to the data model)

109. Pour charger le modèle de données, dans la vue Backstage du fichier, sélectionnez **Fermer et appliquer**.



Toutes les requêtes activées pour le chargement sont désormais chargées dans le modèle de données.

110. Dans le volet **Champs** (situé à droite), notez les sept tables chargées dans le modèle de données.



111. Enregistrez le fichier Power BI Desktop.