

Extraheren en Transformeren van data

- **Benodigde tijd:** 40 a. 60 minuten
- **Beschrijving:** Tijdens deze oefeningen maak je een nieuw Power BI Desktop project en sla je het op als een PBIX bestand. Vervolgens leer je hoe je met de Power Query mogelijkheden van Power BI Desktop data kunt extraheren uit een SQL Azure Data Warehouse om deze vervolgens te transformeren en in een datamodel in te laden. Deze oefeningen vormen de basis voor verdere oefeningen. Dat wil zeggen dat alle oefeningen op elkaar voortbouwen.

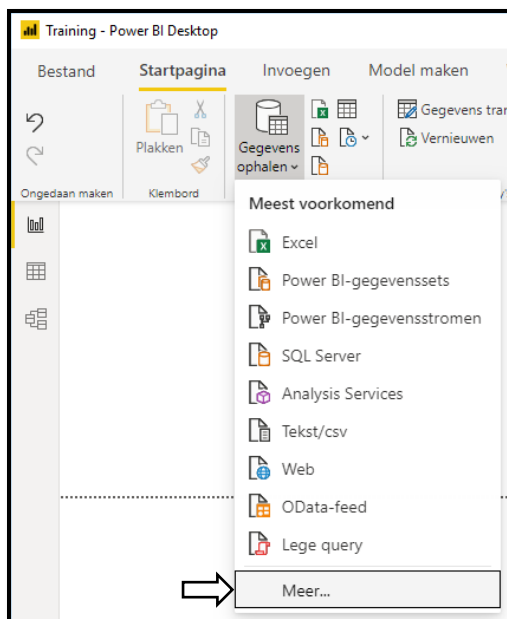
Oefening 1: Importeren Data vanuit Azure SQL Data Warehouse

In deze oefening creëer je een nieuw Power BI Desktop project. Daarna maak je met Power Query een verbinding met Azure SQL Data Warehouse om data te importeren in Power BI Desktop.

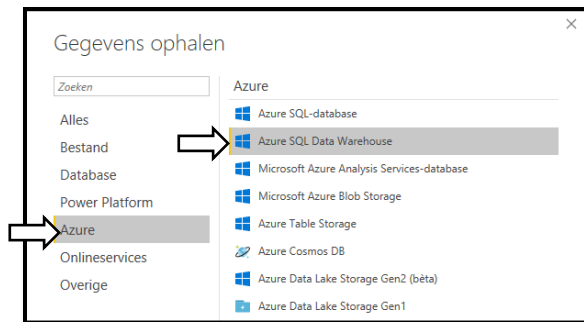
1. Open Power BI Desktop en begin een nieuw project.
2. Sla het nieuw project op als **AdventureWorks InternetSales Analysis.pbix** in het volgende pad.

C:\Power BI\AdventureWorks InternetSales Analysis.pbix

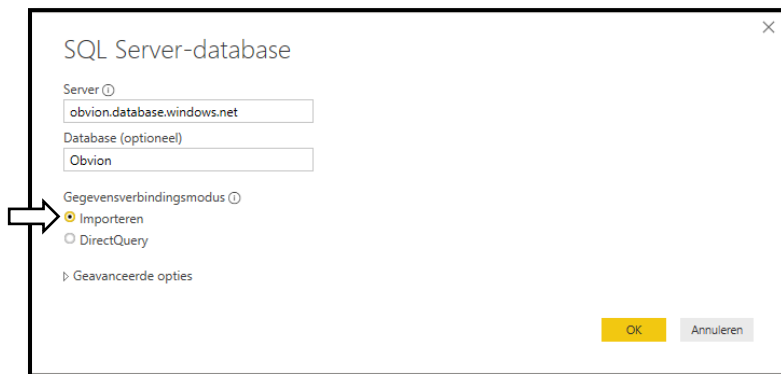
3. Gebruik de menuknop **Gegevens ophalen** om de lijst van bronnen op te halen en klik op **Meer ...**.



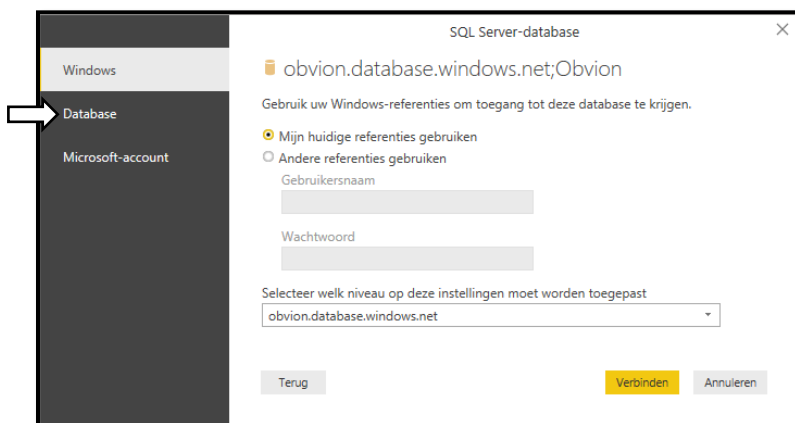
4. In het **Gegevens ophalen** venster, selecteer **Azure** aan de linkerkant. Daarna, selecteer **Azure SQL Data Warehouse** aan de rechterkant en druk op de **Verbinding maken** knop.



5. In het **SQL Server-database** scherm, voer de volgende taken uit.
- Voer onder **Server** de waarde **obvion.database.windows.net** in.
 - Voer onder **Database (optioneel)** de waarde **Obvion** in.
 - Laat de waarde onder **Gegevensverbindingsmodus** op **Importeren** staan.
 - Druk op de knop **OK** om verder te gaan.



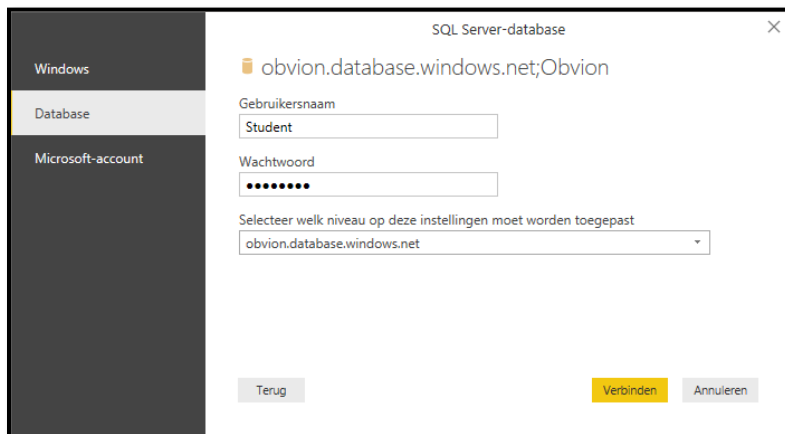
6. In het volgende **SQL Server-database** is het van belang dat aan de linkerkant van het dialoogscherf de optie **Database** wordt geselecteerd. Dit zorgt ervoor dat de database authenticatie in plaats van Windows authenticatie wordt gebruikt.



Let op! Als deze **Windows** niet wordt vervangen door **Database** dan werken de vervolgstappen niet correct.

7. Gebruik de volgende inloggegevens om verbinding te maken met een SQL gebruikersaccount met leesrechten op de database.
- Gebruikersnaam: **Student**
 - Wachtwoord: **P@\$w0rd**

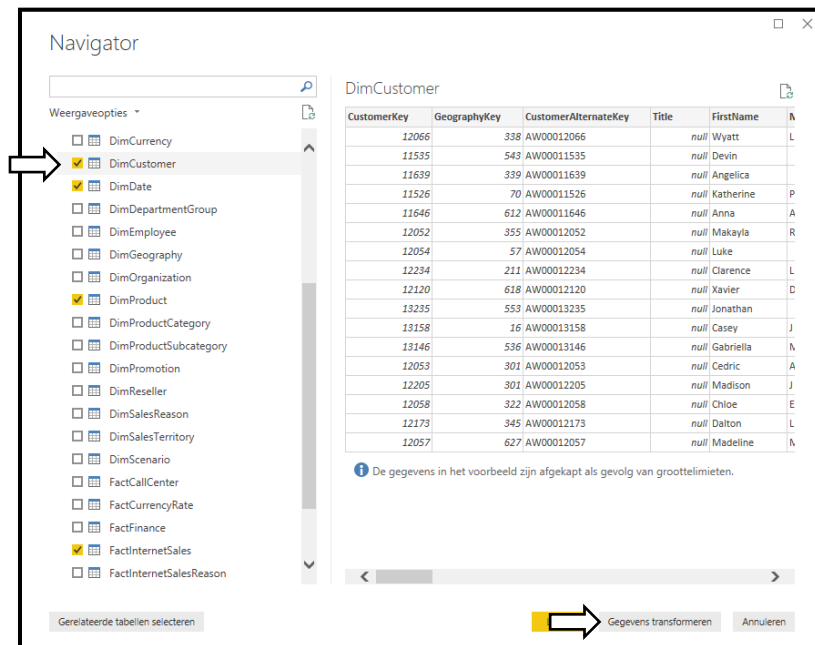
8. Zodra de inloggegevens in het **SQL Server-database** dialoogschermb zijn ingevuld, druk op de knop **Verbinden** om verder te gaan.



Als alles correct is uitgevoerd moet Power BI Desktop in staat zijn een verbinding te maken met de database en het **Navigator** dialoogschermb tonen. Het **Navigator** dialoogschermb stelt je in staat om tabellen te selecteren die je wilt gaan gebruiken in je PBIX project.

9. In het **Navigator** dialoogschermb, selecteer de volgende vier tabellen:
- **FactInternetSales**
 - **DimCustomer**
 - **DimDate**
 - **DimProduct**

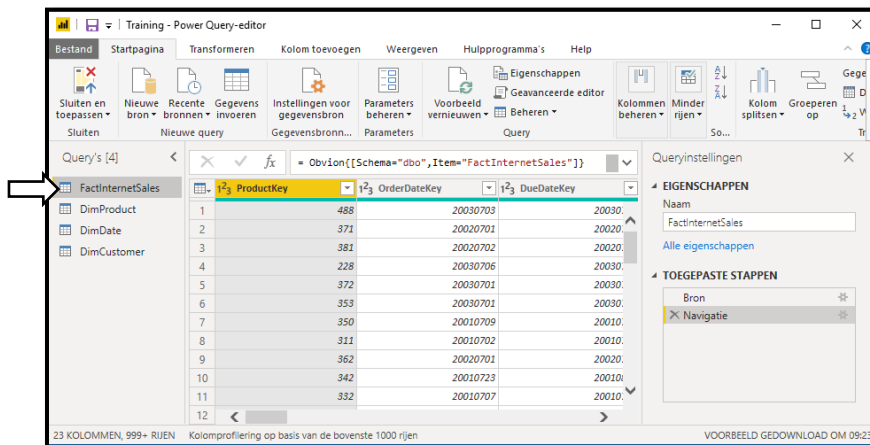
Zodra alle tabellen aan de linkerkant geselecteerd zijn, druk op de knop **Gegevens transformeren** om een query voor iedere tabel te maken en het **Power Query-editor** dialoogschermb te openen.



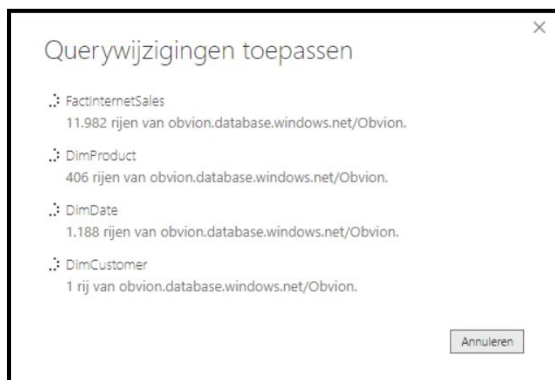
CustomerKey	GeographyKey	CustomerAlternateKey	Title	FirstName	LastName
12066	338	AW00012066	null	Wyatt	L
11535	543	AW00011535	null	Devin	
11639	339	AW00011639	null	Angelica	
11526	70	AW00011526	null	Katherine	P
11646	612	AW00011646	null	Anna	A
12052	355	AW00012052	null	Makayla	R
12054	57	AW00012054	null	Luke	
12234	211	AW00012234	null	Clarence	L
12120	618	AW00012120	null	Xavier	G
12325	553	AW00013235	null	Jonathan	
13158	16	AW00013158	null	Casey	J
13146	536	AW00013146	null	Gabriella	N
12053	301	AW00012053	null	Cedric	A
12205	301	AW00012205	null	Madison	J
12058	322	AW00012058	null	Chloe	E
12173	345	AW00012173	null	Dalton	L
12057	627	AW00012057	null	Madeline	N

10. Wanneer het **Power Query-editor** dialoogschermb wordt bekeken dan zie je aan de linkerkant dat Power BI Desktop een query heeft aangemaakt voor ieder van de vier geselecteerde tabellen in het **Navigator** dialoogschermb. In deze oefening gaan we de query's nog niet aanpassen, dit doen we pas in de volgende

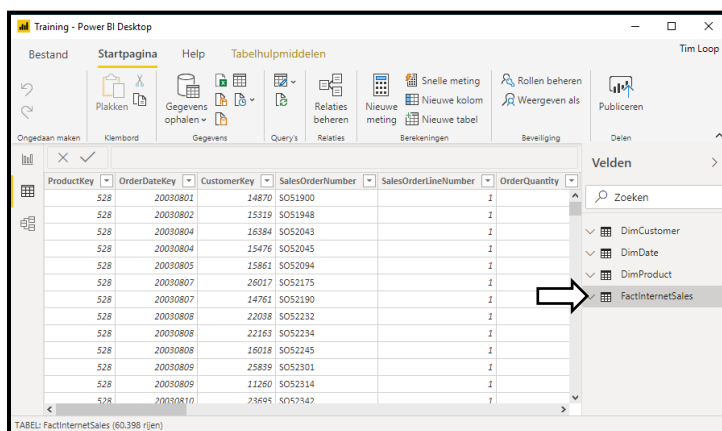
oefening. Inspecteer iedere query door op de querynaam in de **Query's** lijst aan de linkerkant te klikken. Hierdoor zie je dat iedere query initieel alle kolommen van de tabel ophaalt.



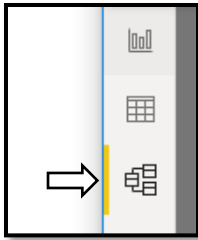
11. Klik op de knop **Sluiten en toepassen** in het **Power Query-editor** dialoogscherf om het extractieproces van de data uit de Azure SQL Data Warehouse te starten. Het **Querywijzigingen toepassen** dialoogvenster laat de voortgang van dit proces zien.



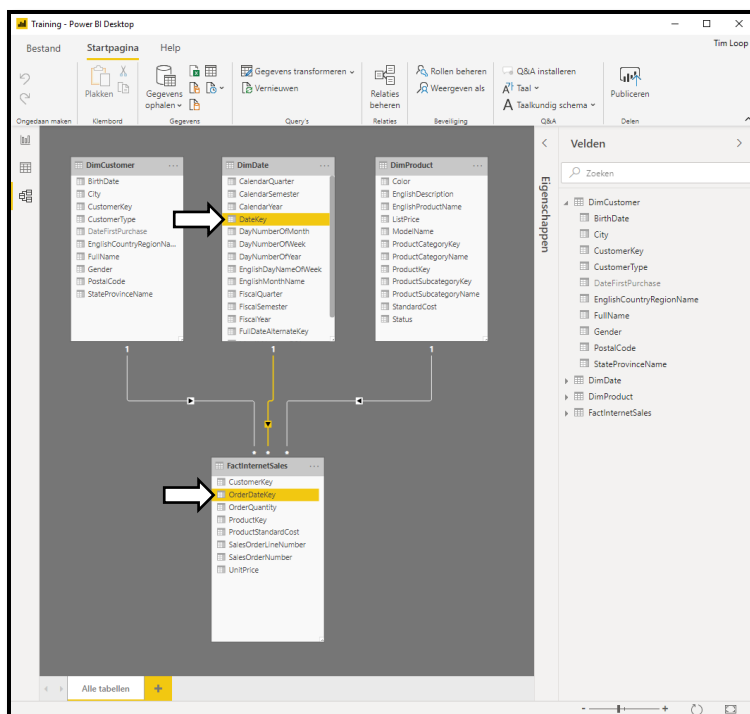
12. Wacht tot de data van alle vier query's is geïmporteerd. Zodra dit is gebeurd kan je in de **Velden** lijst aan de rechterkant van het Power BI Desktop dialoogvenster zien dat de query's vier tabellen hebben aangemaakt.



13. Inspecteer het datamodel in de Model weergave.
 - a) Druk op de Model weergave knop aan de linkerkant van het navigatiescherf.



- b) Als het goed is zie je zeven tabellen weergegeven.
- c) Power BI Desktop heeft de meeste relaties automatisch gedetecteerd en gemaakt tijdens de import van de data in het datamodel.
- d) Gebruik je muis om het formaat van de tabellen te wijzigen zodat de meeste velden zichtbaar zijn, en positioneer de Dim Tabellen rondom de Fact Tabel. Deze structuur heet een stermodel.
- e) Creëer een relatie tussen **DimDate** en **FactInternetSales**. Sleep hiervoor het veld **DateKey** van de tabel **DimDate** naar het veld **OrderDateKey** van de tabel **FactInternetSales**.

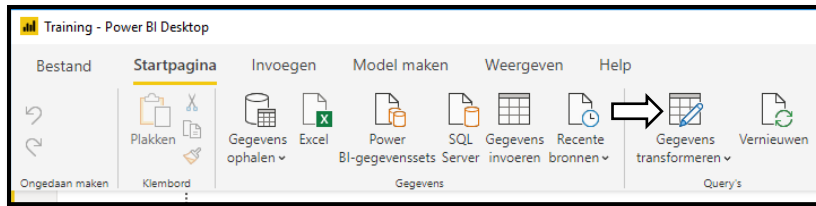


14. Sla je werk op door de toetsencombinatie **ctrl + s** of door linksboven in het Power BI Desktop dialoogvenster op **Bestand** te klikken en op de optie **Opslaan** te klikken.

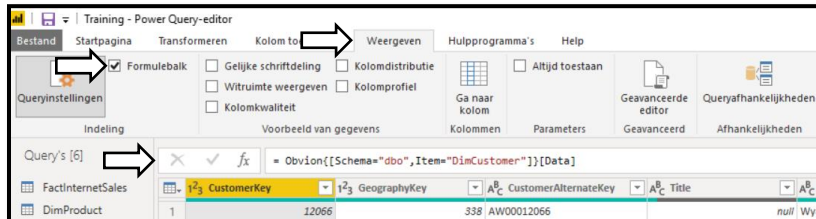
Oefening 2: Bewerken query's om klantdata te transformeren

In deze oefening gebruik je het **Power Query-editor** navigatiescherm om de **DimCustomer** query aan te passen om wijzigingen in de klantdata voordat deze in het model wordt ingelezen aan te brengen.

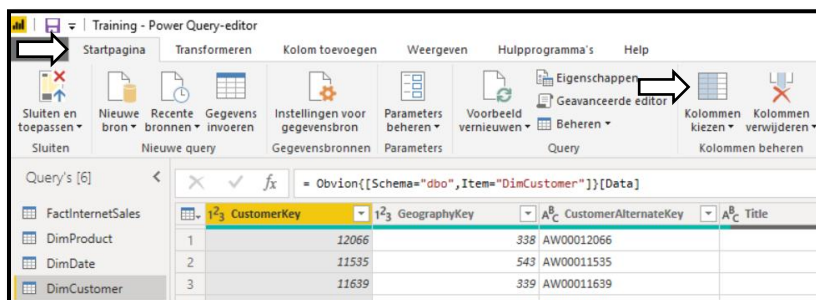
1. Zorg ervoor dat het **AdventureWorks InternetSales Analysis.pbix** project uit de vorige oefening open is.
2. Druk in de lint op de knop **Gegevens transformeren** om naar het Power Query-editor navigatiescherm te navigeren.



3. Zorg ervoor dat de Formulebalk zichtbaar is in het Power Query-editor navigatieschermb.
 - a) In de Power Query-editor, navigeer naar het **Weergeven** tabblad en vink de **Formulebalk** optie aan.
 - b) Je kan nu de query formulebalk met de M code voor de huidige stap in het queryproces zien.



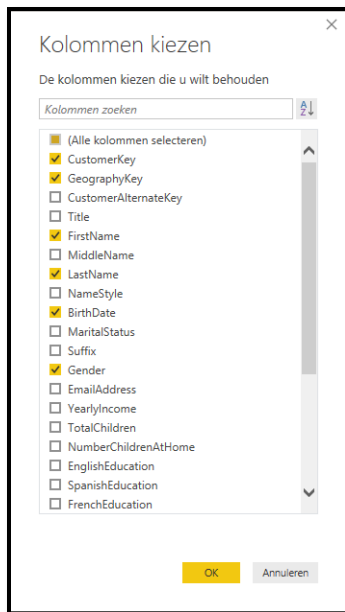
4. Selecteer de gewenste kolommen vanuit de **DimCustomer** tabel.
 - a) Zorg ervoor dat de **DimCustomer** query in de **Query's** lijst aan de linkerkant van het Power Query-editor navigatieschermb is geselecteerd.
 - b) Navigeer naar het **Startpagina** tabblad en druk op de knop **Kolommen kiezen** om het **Kolommen kiezen** navigatieschermb te openen.



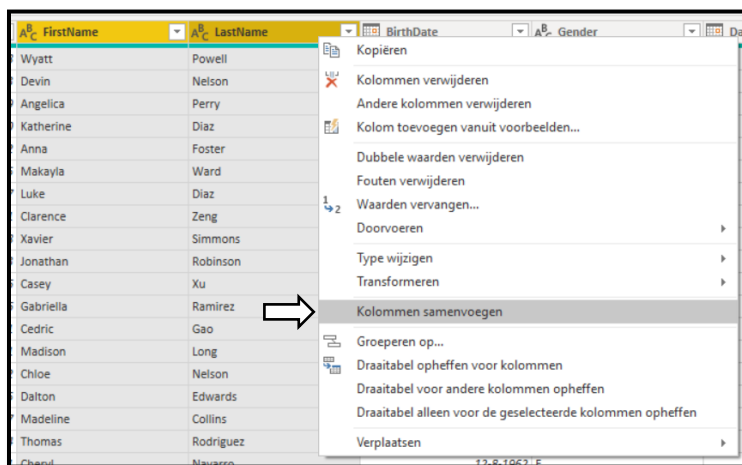
- c) Klik in het **Kolommen kiezen** navigatieschermb op het selectieveld (**Alle kolommen selecteren**) om alle kolommen uit te vinken. Klik daarna op het selectieveld van de volgende kolommen om ze toe te voegen.
 - I) **CustomerKey**
 - II) **GeographyKey**
 - III) **FirstName**
 - IV) **LastName**
 - V) **BirthDate**
 - VI) **Gender**
 - VII) **DateFirstPurchase**

GeographyKey bevat de sleutel voor een dimensie die we later aan deze query zullen toevoegen om de kolommen **City**, **StateProvinceName**, **EnglishCountryRegionName** en **PostalCode** aan de tabel toe te voegen.

- d) Druk, zodra alle kolommen geselecteerd zijn, op de **OK** knop om het **Kolommen kiezen** navigatieschermb te sluiten en de onderliggende query aan te passen.



- e) In Power Query-editor zie je als het goed is nu alleen de kolommen die je net geselecteerd hebt.
5. In deze stap voegen we de kolommen **FirstName** en **LastName** samen in een kolom genaamd **FullName**.
- Selecteer de kolom **FirstName** door op de kolomnaam te klikken.
 - Houdt de **CTRL** toets ingedrukt en selecteer de kolom **LastName** door op de kolomnaam te klikken.
 - Rechtsklik op de geselecteerde kolomnamen en klik op het **Kolommen samenvoegen** commando.



- d) In het **Kolommen samenvoegen** dialoog, gebruik de drop down onder **Scheidingsteken** en selecteer de optie **Spatie**. Vul onder **Nieuwe kolomnaam (optioneel)** de tekst **FullName** in en druk op de knop **OK** om deze transformatie aan de onderliggende query toe te voegen.

Kolommen samenvoegen

Kiezen hoe de geselecteerde kolommen worden samengevoegd.

Scheidingsteken
Spatie

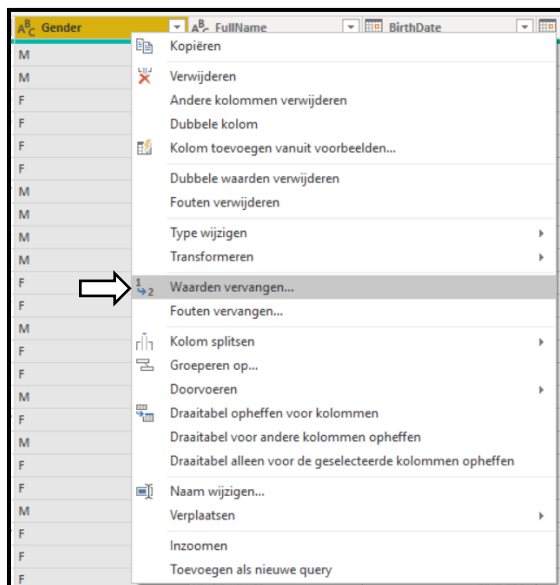
Nieuwe kolomnaam (optioneel)
FullName

OK Annuleren

- e) Je ziet nu dat de kolommen **FirstName** en **LastName** zijn vervangen door een enkele kolom genaamd **FullName**.

	1 ² CustomerKey	1 ² GeographyKey	A ³ FullName	BirthDate	A ⁴
1	12066	338	Wyatt Powell	12-8-1944	M
2	11535	543	Devin Nelson	3-7-1958	M
3	11639	339	Angelica Perry	23-12-1972	F
4	11526	70	Katherine Diaz	6-8-1956	F
5	11646	612	Anna Foster	19-3-1969	F

6. Modificeer de kolom **Gender** zodat deze de waarden **Male** en **Female** in plaats van **M** en **F** bevat.
- Vind de kolom **Gender** in de **DimCustomer** tabel.
 - Rechtsklik op de kolomnaam van de kolom **Gender** en klik op het **Waarden vervangen...** commando.



- In het **Waarden vervangen** dialoogvenster, vul onder **Te zoeken waarde** de waarde **M** in. Vul daarna onder **Vervangen door** de waarde **Male** in. Druk op de knop **OK** om deze transformatie aan de onderliggende query toe te voegen.

Waarden vervangen

Een waarde door een andere waarde vervangen in de geselecteerde kolommen.

Te zoeken waarde
M

Vervangen door
Male

Geavanceerde opties

OK Annuleren

- d) Je ziet nu dat alle waarden **M** in de **Gender** kolom zijn vervangen door de waarde **Male**.

Gender
Male
Male
F
F
F
F
Male

- e) Rechtsklik op de kolomnaam van de kolom **Gender** en klik op het **Waarden vervangen...** commando.
- f) In het **Waarden vervangen** dialoogvenster, vul onder **Te zoeken waarde** de waarde **F** in. Vul daarna onder **Vervangen door** de waarde **Female** in. Druk op de knop **OK** om deze transformatie aan de onderliggende query toe te voegen.

Waarden vervangen

Een waarde door een andere waarde vervangen in de geselecteerde kolommen.

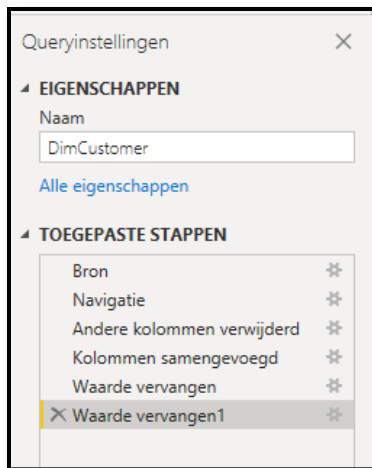
Te zoeken waarde
F

Vervangen door
Female

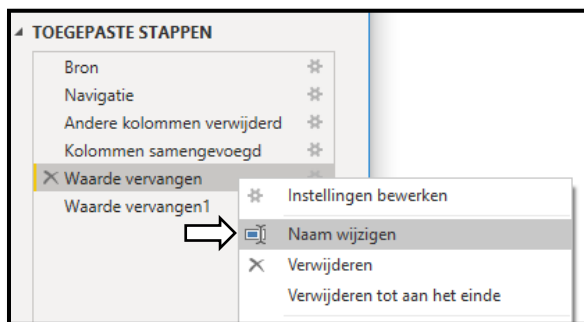
Geavanceerde opties

OK Annuleren

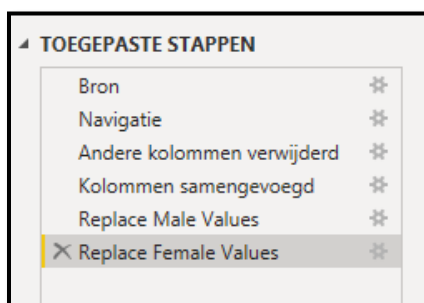
- g) Je zit nu dat alle waarden in de kolom **Gender** zijn vervangen met de waarden **Male** of **Female**.
- h) Bekijk de lijst onder **TOEGEPASTE STAPPEN** aan de rechterkant van het Power Query-editor navigatiescherm. De laatste twee stappen aan het einde moeten de generieke namen **Waarde vervangen** en **Waarde vervangen1** bevatten.



- i) Om voor een betere beheersbaarheid te zorgen is het verstandig om generieke namen zoals **Waarde vervangen** en **Waarde vervangen1** te vervangen. Hernoem de stap **Waarde vervangen** door op deze stap te rechtsklikken en het commando **Naam wijzigen** te selecteren. Pas de naamwaarde aan naar **Replace Male Values**.



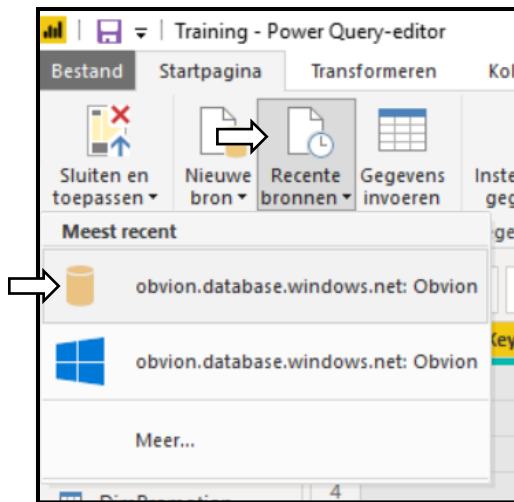
- j) Gebruik dezelfde techniek om de naam van de stap **Waarde vervangen1** aan te passen naar **Replace Female Values**.



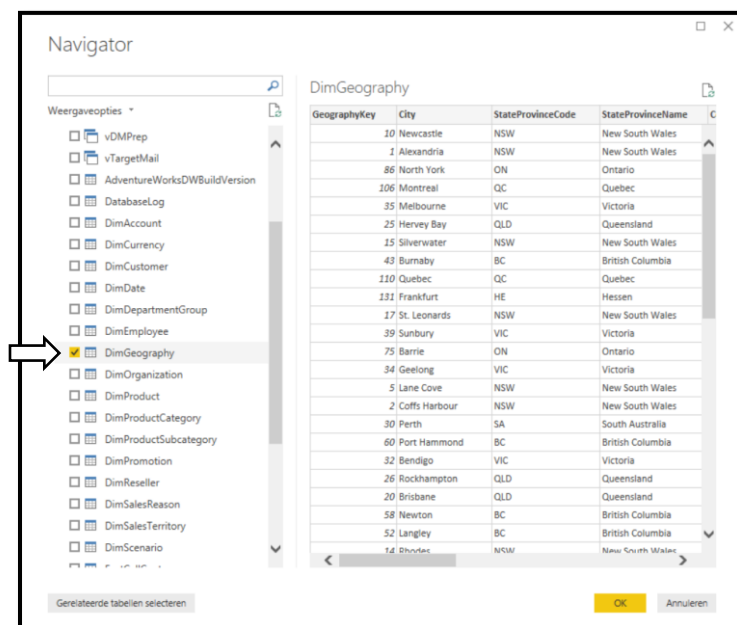
Je hebt nu gezien hoe je een query stap kunt hernoemen. Deze oefening wordt in de verdere materie niet herhaald. Echter, wanneer je query's in een grotere business context gaat maken dan is het verstandig om betekenisvolle namen aan de stappen te geven, zodat anderen je logica makkelijker kunnen lezen en begrijpen.

7. Gebruik de kolom **GeographyKey** om de waarden **City**, **StateProvinceName**, **EnglishCountryRegionName** en **PostalCode** aan de tabel toe te voegen.

- a) In de Power Query-editor klik op **Recente bronnen** en selecteer de connectiestring die in oefening 1 is ingesteld. Deze actie zorgt ervoor dat hetzelfde navigatievenster als uit een eerdere oefening wordt getoond.



- b) In de **Navigators** selecteer de tabel **DimGeography** en klik op **OK** om verder te gaan en om het **Verbindingsinstellingen** navigatiescherm te tonen.



- c) In het **Verbindingsinstellingen** navigatiescherm, controleer of de optie **Importeren** is geselecteerd en druk op **OK** om verder te gaan.
- d) Controleer of de nieuwe query **DimGeography** wordt getoond in de lijst onder **Query's** aan de linkerkant van de Power Query-editor.
- e) Selecteer de query **DimCustomer** en in het tabblad **Startpagina** in het lint aan de bovenkant van de Power Query-editor klik in de sectie **Combineren** op **Query's samenvoegen** om het dialoogvenster **Samenvoegen** te tonen.
- f) In het dialoogvenster **Samenvoegen** klik op de kolomnaam **GeographyKey** onder **DimCustomer**. Gebruik de lege drop down onder DimCustomer om de waarde **DimGeography** te selecteren, en klik hier wederom op de kolom met de naam **GeographyKey**. Laat alle overige opties op default staan. Druk op OK om de wijziging aan de onderliggende query toe te voegen.

Samenvoegen

Selecteer een tabel en overeenkomende kolommen om een samengevoegde tabel te maken.

DimCustomer

CustomerKey	GeographyKey	FullName	BirthDate	Gender	DateFirstPurchase
12066	338	Wyatt Powell	12-8-1944	Male	2-9-2003
11535	543	Devin Nelson	3-7-1958	Male	5-3-2002
11639	339	Angelica Perry	23-12-1972	Female	12-5-2004
11526	70	Katherine Diaz	6-8-1956	Female	7-8-2003
11646	612	Anna Foster	19-3-1969	Female	16-1-2004

DimGeography

GeographyKey	City	StateProvinceCode	StateProvinceName	CountryRegionCode	EnglishCountryRegion
10	Newcastle	NSW	New South Wales	AU	Australia
1	Alexandria	NSW	New South Wales	AU	Australia
86	North York	ON	Ontario	CA	Canada
106	Montreal	QC	Quebec	CA	Canada
35	Melbourne	VIC	Victoria	AU	Australia

Type join
Left outer (alle uit de eerste, overeenkomende uit de t...)

☐ Gebruik fuzzy overeenkomst om de samenvoeging uit te voeren

Opties voor fuzzy overeenkomsten

✓ De selectie komt overeen met 18484 van 18484 rijen in de eerste tabel.

OK Annuleren

- g) In de nieuwe kolom **DimGeography** klik in de kolomnaam op het symbool voor **Uitvouwen** en in het **Uitvouwen** dialoogscherm klik op **(Alle kolommen selecteren)** om alle kolommen uit de selectie te halen en vink tevens de optie **Oorspronkelijke kolomnaam gebruiken als voorvoegsel** uit.
- h) Vervolgens, selecteer de kolommen **City**, **StateProvinceName**, **EnglishCountryRegionName** en **PostalCode** en druk op **OK** om verder te gaan.

graphKey"}, "DimGeography", JoinKind.LeftOuter)

Gender DateFirstPurchase DimGeography

Kolommen zoeken om uit te breiden

☒ Uitvouwen ☐ Samenstellen

☒ (Alle kolommen selecteren)

☐ GeographyKey

☒ City

☐ StateProvinceCode

☒ StateProvinceName

☐ CountryRegionCode

☒ EnglishCountryRegionName

☐ SpanishCountryRegionName

☐ FrenchCountryRegionName

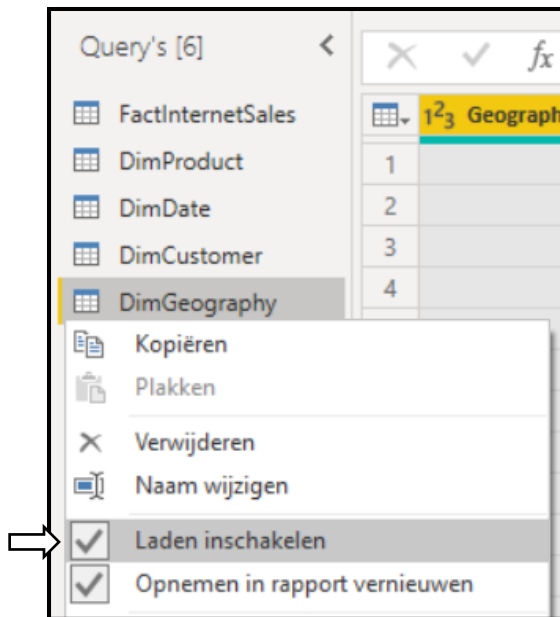
☒ PostalCode

☐ SalesTerritoryKey

☐ Oorspronkelijke kolomnaam gebruiken als voorvoegsel

OK Annuleren

- i) Tot slot, in de **DimCustomer** query, ga terug naar het **Kolommen kiezen** dialoogvenster en vink de kolom **GeographyKey** uit. Druk vervolgens op **OK** om de wijziging aan de onderliggende query toe te voegen.
- j) Rechtsklik aan de linkerkant van de Power Query-editor op de querynaam **DimGeography** en vink de optie **Laden inschakelen** uit. Dit voorkomt dat straks **DimGeography** een aparte tabel vormt in het Power BI datamodel.

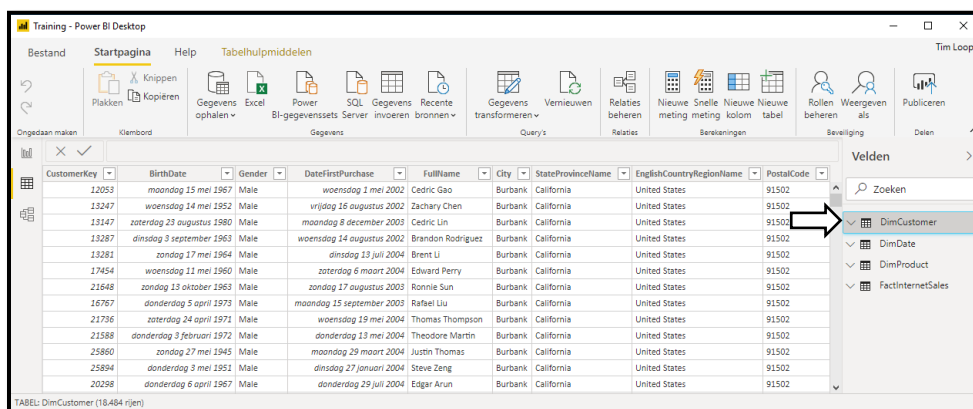


8. Voer de **DimCustomer** query uit om de resultaten te kunnen bekijken in Power BI Desktop.
- Druk op **Sluiten en toepassen** in het **Startpagina** tabblad van het lint om de **DimCustomer** query uit te voeren.



Power BI Desktop toont het **Querywijzigingen toepassen** dialoogscherm tijdens het importeren en transformeren van data vanuit Azure SQL Datawarehouse naar het datamodel.

- Nadat de **DimCustomer** query is toegepast moeten alle resultaten zichtbaar zijn in de **Gegevens** weergave van Power BI Desktop wanneer de **DimCustomer** tabel wordt geselecteerd in de **Velden** lijst aan de rechterkant.



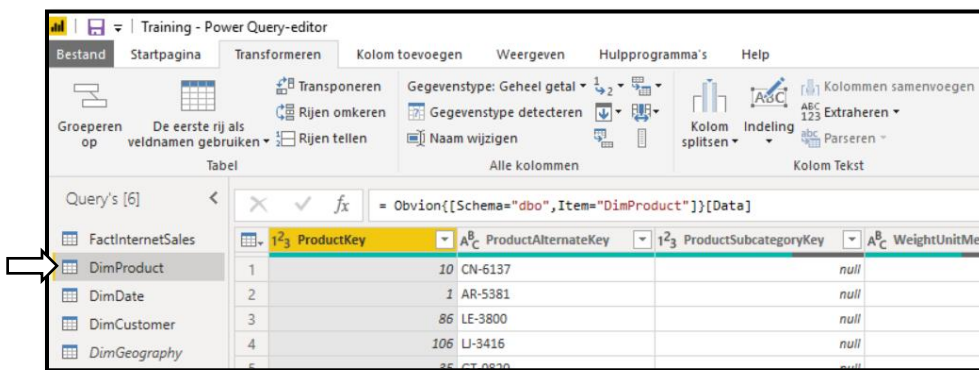
Je bent nu klaar met werken in de **DimCustomer** query.

9. Sla je werk op door in Power BI Desktop op **Bestand** en op **Opslaan** te klikken, of door de toetsenbordcombinatie **ctrl + s**.

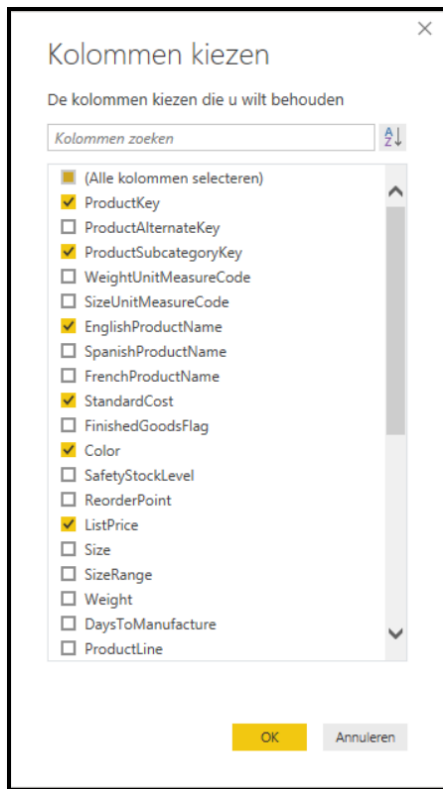
Oefening 3: Bewerken query's om productdata te transformeren

In deze oefening gebruik je het **Power Query-editor** navigatiescherm om de **DimProduct** query aan te passen om wijzigingen in de productdata voordat deze in het model wordt ingelezen aan te brengen.

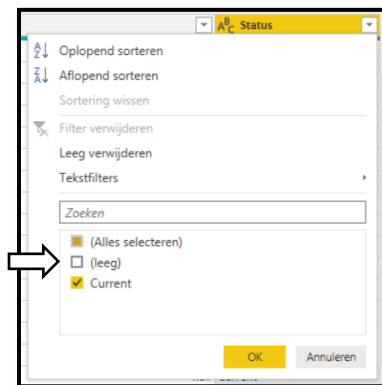
1. Open de **DimProduct** query in bewerken modus.
 - a) Klik op **Gegevens transformeren** in het lint van het **Startpagina** tabblad om het **Power Query Editor** navigatiescherm te tonen.
 - b) Zorg ervoor dat de **DimProduct** query geselecteerd is in de **Query's** lijst aan de linkerkant van de **Power Query-editor**.



2. Verwijder de kolommen die niet nodig zijn in de **DimProduct** query resultaten.
 - a) Klik op de **Kolommen kiezen** knop in het lint om het **Kolommen kiezen** dialoogvenster te openen.
 - b) In het **Kolommen kiezen** dialoogvenster, begin met te klikken op **(Alle kolommen selecteren)** om de selectie van alle kolommen ongedaan te maken.
 - c) Selecteer de volgende kolommen:
 - I) **ProductKey**
 - II) **ProductSubcategoryKey**
 - III) **EnglishProductName**
 - IV) **StandardCost**
 - V) **Color**
 - VI) **ListPrice**
 - VII) **ModelName**
 - VIII) **EnglishDescription**
 - IX) **Status**



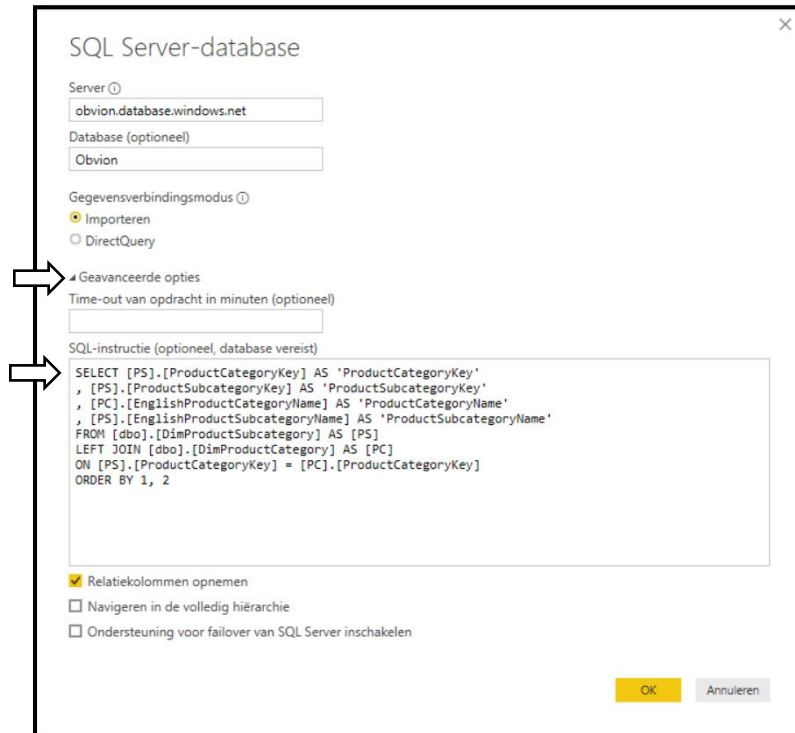
- d) Klik op **OK** om het **Kolommen kiezen** dialoogvenster te sluiten en de onderliggende query aan te passen.
3. Filter de resultaten om alleen actuele productinformatie te tonen.
- a) Klik op het filtericoon rechts van de kolomnaam **Status** om het filterdialoogscherm te tonen.



- b) Vink de optie **(leeg)** uit en druk op **OK** om de resultaten te filteren.
4. Voeg **ProductCategory** en **ProductSubcategory** toe aan de resultaten.
- a) Voeg een nieuwe **Azure SQL Data Warehouse** bron toe via de knop **Nieuwe bron** in het lint van het **Startpagina** tabblad van de Power Query-editor.
- b) Vul dezelfde informatie in als bij **oefening 1.5** (vergeet **Database (optioneel)** niet in te vullen!) en klik op **Geavanceerde opties**.
- c) Vul onder **SQL-instructie (optioneel, database vereist)** onderstaande SQL-code in.

```
SELECT [PS].[ProductCategoryKey] AS 'ProductCategoryKey'  
, [PS].[ProductSubcategoryKey] AS 'ProductSubcategoryKey'  
, [PC].[EnglishProductCategoryName] AS 'ProductCategoryName'
```

```
, [PS].[EnglishProductSubcategoryName] AS 'ProductSubcategoryName'
FROM [dbo].[DimProductSubcategory] AS [PS]
LEFT JOIN [dbo].[DimProductCategory] AS [PC]
ON [PS].[ProductCategoryKey] = [PC].[ProductCategoryKey]
ORDER BY 1, 2
```



SQL Server-database

Server ①
obvion.database.windows.net

Database (optioneel)
Obvion

Gegevensverbindingmodus ①
☒ Importeren
☐ DirectQuery

Geavanceerde opties
Time-out van opdracht in minuten (optioneel)
[]

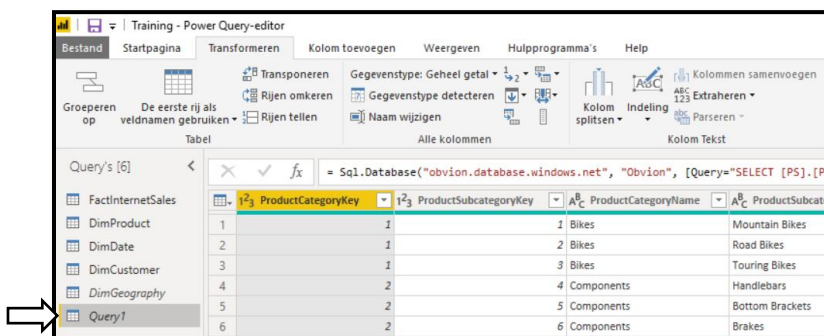
SQL-instructie (optioneel, database vereist)
SELECT [PS].[ProductCategoryKey] AS 'ProductCategoryKey',
[PS].[ProductSubcategoryKey] AS 'ProductSubcategoryKey',
[PC].[EnglishProductCategoryName] AS 'ProductCategoryName',
[PS].[EnglishProductSubcategoryName] AS 'ProductSubcategoryName'
FROM [dbo].[DimProductSubcategory] AS [PS]
LEFT JOIN [dbo].[DimProductCategory] AS [PC]
ON [PS].[ProductCategoryKey] = [PC].[ProductCategoryKey]
ORDER BY 1, 2

☒ Relatiekolommen opnemen
☐ Navigeren in de volledige hiërarchie
☐ Ondersteuning voor failover van SQL Server inschakelen

OK Annuleren

d) Druk ok **OK** om de query te starten.

e) Druk in het volgende dialoogvenster wederom op **OK** om **Query1** in te laden in de Power Query-editor.



Training - Power Query-editor

Bestand Startpagina Transformeren Kolom toevoegen Weergeven Hulpprogramma's Help

Groeperen op De eerste rij als veldnamen gebruiken Rijen omkeren Rijen tellen

Tabel

Gegevenstype: Geheel getal 1 2
Gegevenstype detecteren
Naam wijzigen

Alle kolommen

Kolom splitsen Indeling
Kolom Tekst

Query's [6]
FactInternetSales
DimProduct
DimDate
DimCustomer
DimGeography
Query1

Query1 = Sql.Database("obvion.database.windows.net", "Obvion", [Query="SELECT [PS].[Pr...])

	1 ² ProductCategoryKey	1 ² ProductSubcategoryKey	A ^B ProductCategoryName	A ^B ProductSubcategoryName
1	1	1	Bikes	Mountain Bikes
2	1	2	Bikes	Road Bikes
3	1	3	Bikes	Touring Bikes
4	2	4	Components	Handlebars
5	2	5	Components	Bottom Brackets
6	2	6	Components	Brakes

f) Selecteer **DimProduct** in de **Query's** lijst aan de linkerkant en klik op de optie **Query's samenvoegen** in de sectie **Combineren** om het **Samenvoegen** dialoogvenster te tonen.

g) In het **Samenvoegen** dialoogvenster klik onder **DimProduct** op de kolomnaam **ProductSubcategoryKey** om deze te gebruiken voor het samenvoegen.

h) Gebruik de drop down onder het gedeelte **DimProduct** om de optie **Query1** te selecteren en klik hier tevens op de kolomnaam **ProductSubcategoryKey** om deze te gebruiken voor het samenvoegen.

Samenvoegen

Selecteer een tabel en overeenkomende kolommen om een samengevoegde tabel te maken.

DimProduct

ProductKey	ProductSubcategoryKey	EnglishProductName	StandardCost	Color	ListPrice	ModelName
10	null	Chainring Nut	null	Silver	null	ni
1	null	Adjustable Race	null	NA	null	ni
86	null	External Lock Washer 7	null	NA	null	ni
106	null	Thin-Jam Lock Nut 16	null	NA	null	ni
35	null	LL Grip Tape	null	NA	null	ni

Query1

ProductCategoryKey	ProductSubcategoryKey	ProductCategoryName	ProductSubcategoryName
1	1	Bikes	Mountain Bikes
1	2	Bikes	Road Bikes
1	3	Bikes	Touring Bikes
2	4	Components	Handlebars
2	5	Components	Bottom Brackets

Type join
Left outer (alle uit de eerste, overeenkomende uit de t...)

☐ Gebruik fuzzy overeenkomst om de samenvoeging uit te voeren

Opties voor fuzzy overeenkomsten

✓ De selectie komt overeen met 197 van 406 rijen in de eerste tabel.

OK Annuleren

- Klik op **OK** om de query's samen te voegen.
- In de kolomnaam **Query1** druk op het **Uitvouwensymbool** om het uitvouwen dialoogvenster te tonen.
- In het uitvouwen dialoogvenster vink de optie **Oorspronkelijke kolomnaam gebruiken als voorvoegsel uit**.

Status Query1

Kolommen zoeken om uit te breiden

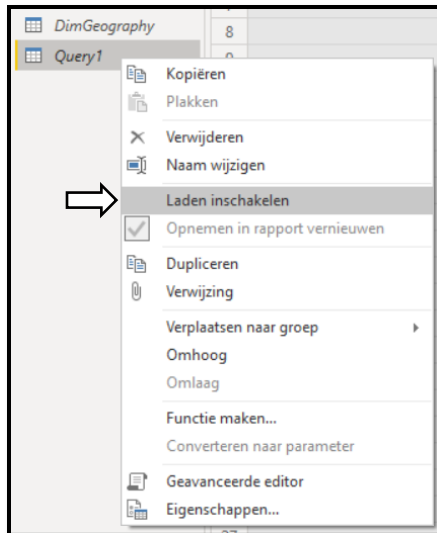
☒ Uitvouwen ☐ Samenstellen

- ☒ (Alle kolommen selecteren)
- ☒ ProductCategoryKey
- ☒ ProductSubcategoryKey
- ☒ ProductCategoryName
- ☒ ProductSubcategoryName

☐ Oorspronkelijke kolomnaam gebruiken als voorvoegsel

OK Annuleren

- Klik op **OK** om het uitvouwen uit te voeren.
- Klik op de kolomnaam **ProductSubcategoryKey.1** om deze te selecteren.
- Klik in het lint van het tabblad **Startpagina** op de optie **Kolommen verwijderen** om de kolom **ProductSubcategoryKey.1** te verwijderen.
- Selecteer **Query1** in de lijst **Query's** aan de linkerkant van de Power Query-editor om vervolgens via een rechtsklik de optie **Laden inschakelen** uit te schakelen.



5. Voer de **DimProduct** query uit om de resultaten te kunnen bekijken in Power BI Desktop.
 - a) Druk op **Sluiten en toepassen** in het **Startpagina** tabblad van het lint om de **DimProduct** query uit te voeren.
 - b) Nadat de **DimProduct** query is toegepast moeten alle resultaten zichtbaar zijn in de **Gegevens** weergave van Power BI Desktop wanneer de **DimProduct** tabel wordt geselecteerd in de **Velden** lijst aan de rechterkant.

Je bent nu klaar met werken in de **DimProduct** query.

6. Sla je werk op door in Power BI Desktop op **Bestand** en op **Opslaan** te klikken, of door de toetsenbordcombinatie **ctrl + s**.

Oefening 4: Bewerken query's om internetsalesdata te transformeren

In deze oefening gebruik je het **Power Query-editor** navigatieschermbom de **FactInternetSales** query aan te passen om wijzigingen in de internetsalesdata voordat deze in het model wordt ingelezen aan te brengen. Deze tabel vormt het centrale punt van ons gegevensmodel.

1. Open de **FactInternetSales** query in bewerken modus.
 - a) Klik op **Gegevens transformeren** in het lint van het **Startpagina** tabblad om het **Power Query Editor** navigatieschermbom te tonen.
 - b) Zorg ervoor dat de **FactInternetSales** query geselecteerd is in de **Query's** lijst aan de linkerkant van de **Power Query-editor**.

Training - Power Query-editor

Bestand Startpagina Transformeren Kolom toevoegen Weergeven Hulpprogramma's Help

Groeperen op De eerste rij als veldnamen gebruiken Rijen omkeren Rijen tellen

Gegevenstype: Geheel getal Gegevenstype detecteren Naam wijzigen

Kolom splitsen Indeling

Query's [6]

FactInternetSales

DimProduct

DimDate

DimCustomer

DimGeography

Query1

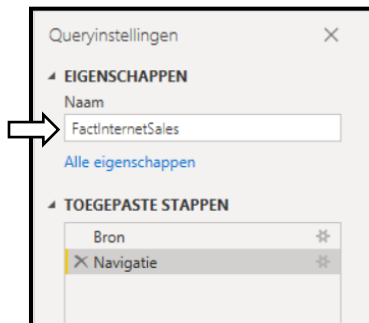
Tabel

ProductKey OrderDateKey DueDateKey

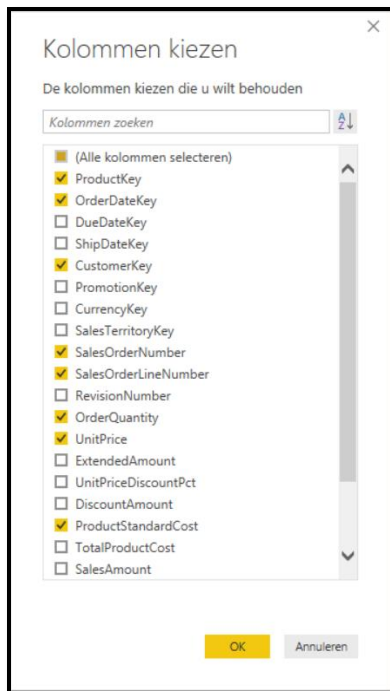
	ProductKey	OrderDateKey	DueDateKey
1	488	20030703	200307
2	371	20020701	200207
3	381	20020702	200207
4	228	20030706	200307
5	372	20030701	200307
6	353	20030701	200307

= Obvion([Schema="dbo", Item="FactInternetSales"])[Data]

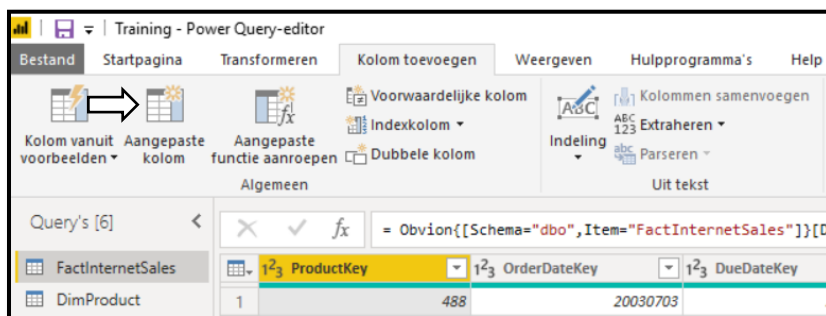
Een belangrijk punt van flexibiliteit in het importproces is de mogelijkheid van het hernoemen van een querystap, en hiermee tegelijkertijd de naam van de tabel in het datamodel aanpassen. Net als in het hernoemen van de individuele querystappen, probeer gebruik te maken van intuïtieve en makkelijk te begrijpen namen bij het benoemen van de query's.



2. Verwijder de kolommen die niet nodig zijn in de **FactInternetSales** query resultaten.
 - a) Klik op de **Kolommen kiezen** knop in het lint om het **Kolommen kiezen** dialoogvenster te openen.
 - b) In het **Kolommen kiezen** dialoogvenster, begin met te klikken op **(Alle kolommen selecteren)** om de selectie van alle kolommen ongedaan te maken.
 - c) Selecteer de volgende kolommen:
 - I) **ProductKey**
 - II) **OrderDateKey**
 - III) **CustomerKey**
 - IV) **SalesOrderNumber**
 - V) **SalesOrderLineNumber**
 - VI) **OrderQuantity**
 - VII) **UnitPrice**
 - VIII) **ProductStandardCost**



- d) Klik op **OK** om het **Kolommen kiezen** dialoogvenster te sluiten en de onderliggende query aan te passen.
3. Bereken met M een nieuwe kolom **TotalProductCost** waarin de waarde van de kolom **OrderQuantity** wordt vermenigvuldigd met de waarde van de kolom **ProductStandardCost**
- a) Klik in het **Kolom toevoegen** tabblad van de Power Query-editor op de knop **Aangepaste kolom** om het **Aangepaste kolom** dialoogvenster te tonen.



- b) In het **Aangepaste kolom** dialoogvenster vul als waarde voor **Nieuwe kolomnaam** de waarde **TotalProductCost** in.
- c) In het **Aangepaste kolomformule** gedeelte vul onderstaande formule in.

[OrderQuantity] * [ProductStandardCost]

- d) Druk op de knop **OK** om de berekening uit te voeren en de nieuwe kolom toe te voegen.

Aangepaste kolom

Een kolom toevoegen die wordt berekend op basis van andere kolommen.

Nieuwe kolomnaam
TotalProductCost

Aangepaste kolomformule ⓘ
= [OrderQuantity] * [ProductStandardCost]

Beschikbare kolommen

- ProductKey
- OrderDateKey
- CustomerKey
- SalesOrderNumber
- SalesOrderLineNumber
- OrderQuantity
- UnitPrice
- ProductStandardCost

<< Invoegen

Informatie over Power BI Desktop-formules

✓ Er zijn geen syntaxisfouten gedetecteerd.

OK Annuleren

- e) Bevestig dat de nieuwe kolom **TotalProductCost** de vermenigvuldigde waarden van de kolommen **OrderQuantity** en **ProductStandardCost** bevat.

1.2 OrderQuantity	\$ ProductStandardCost	ABC 123 TotalProductCost
1	41,5723	41,5723
1	1320,6838	1320,6838
1	605,6492	605,6492
1	38,4923	38,4923
1	1554,9479	1554,9479

4. Bereken met M een nieuwe kolom **TotalSales** waarin de waarde van de kolom **OrderQuantity** wordt vermenigvuldigd met de waarde van de kolom **UnitPrice**.
- Klik in het **Kolom toevoegen** tabblad van de Power Query-editor op de knop **Aangepaste kolom** om het **Aangepaste kolom** dialoogvenster te tonen.
 - In het **Aangepaste kolom** dialoogvenster vul als waarde voor **Nieuwe kolomnaam** de waarde **TotalSales** in.
 - In het **Aangepaste kolomformule** gedeelte vul onderstaande formule in.

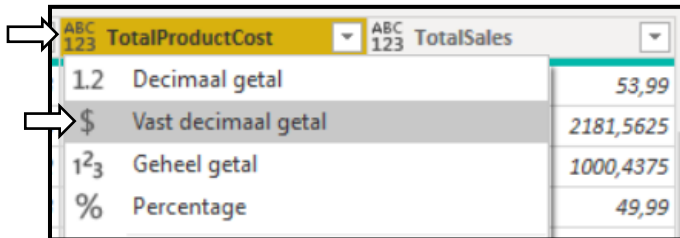
[OrderQuantity] * [UnitPrice]

- Druk op de knop **OK** om de berekening uit te voeren en de nieuwe kolom toe te voegen.
- Bevestig dat de nieuwe kolom **TotalSales** de vermenigvuldigde waarden van de kolommen **OrderQuantity** en **UnitPrice** bevat.

1.2 OrderQuantity	\$ UnitPrice	ABC 123 TotalSales
1	53,99	53,99
1	2181,5625	2181,5625
1	1000,4375	1000,4375
1	49,99	49,99
1	2443,35	2443,35

Wanneer kolommen in je query numerieke valutawaarden bevat dan is het verstandig het datatype aan te passen naar een **Vast decimaal getal**.

5. Wijzig de kolomtypen van de kolommen **TotalProductCost** en **TotalSales** naar het type **Vast decimaal getal**.
 - a) Gebruik de kolomtype drop down om het type van de kolom **TotalProductCost** aan te passen naar een **Vast decimaal getal**.



- b) Gebruik de kolomtype drop down om het type van de kolom **TotalSales** aan te passen naar een **Vast decimaal getal**.

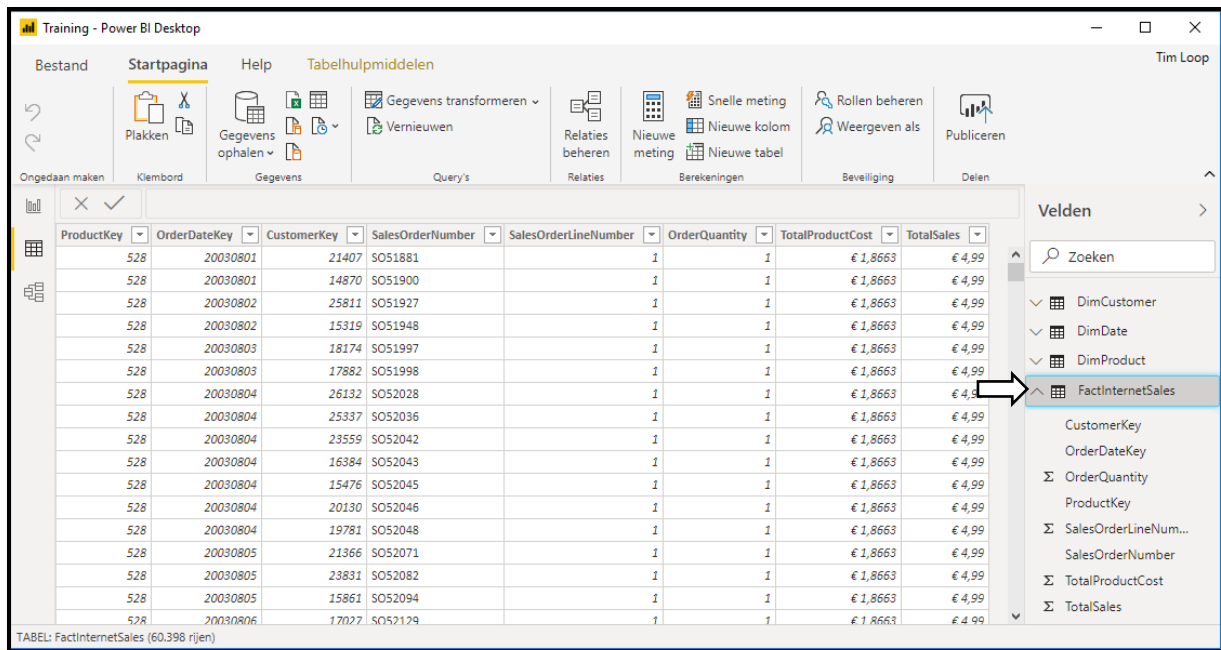
\$ TotalProductCost	\$ TotalSales
41,5723	53,99
1320,6838	2181,5625
605,6492	1000,4375
38,4923	49,99
1554,9479	2443,35

Zodra **TotalProductCost** en **TotalSales** zijn berekend zijn de kolommen **ProductStandardCost** en **UnitPrice** niet meer nodig.

6. Verwijder de kolommen **ProductStandardCost** en **UnitPrice**.
 - a) Verwijder de kolom **ProductStandardCost** door op de kolomnaam te rechtsklikken en het **Verwijderen** commando te selecteren.



- b) Verwijder de kolom **UnitPrice** door op de kolomnaam te rechtsklikken en het **Verwijderen** commando te selecteren.
7. Voer de **DimProduct** query uit om de resultaten te kunnen bekijken in Power BI Desktop.
 - a) Druk op **Sluiten en toepassen** in het **Startpagina** tabblad van het lint om de **DimProduct** query uit te voeren.
 - b) Nadat de **DimProduct** query is toegepast moeten alle resultaten zichtbaar zijn in de **Gegevens** weergave van Power BI Desktop wanneer de **DimProduct** tabel wordt geselecteerd in de **Velden** lijst aan de rechterkant.



The screenshot shows the Power BI Desktop interface. The main view displays the 'FactInternetSales' table with columns: ProductKey, OrderDateKey, CustomerKey, SalesOrderNumber, SalesOrderLineNumber, OrderQuantity, TotalProductCost, and TotalSales. The Fields pane on the right shows the 'FactInternetSales' table selected, with its columns listed below it.

ProductKey	OrderDateKey	CustomerKey	SalesOrderNumber	SalesOrderLineNumber	OrderQuantity	TotalProductCost	TotalSales
528	20030801	21407	SOS1881	1	1	€ 1,8663	€ 4,99
528	20030801	14870	SOS1900	1	1	€ 1,8663	€ 4,99
528	20030802	25811	SOS1927	1	1	€ 1,8663	€ 4,99
528	20030802	15319	SOS1948	1	1	€ 1,8663	€ 4,99
528	20030803	18174	SOS1997	1	1	€ 1,8663	€ 4,99
528	20030803	17882	SOS1998	1	1	€ 1,8663	€ 4,99
528	20030804	26132	SOS2028	1	1	€ 1,8663	€ 4,99
528	20030804	25337	SOS2036	1	1	€ 1,8663	€ 4,99
528	20030804	23559	SOS2042	1	1	€ 1,8663	€ 4,99
528	20030804	16384	SOS2043	1	1	€ 1,8663	€ 4,99
528	20030804	15476	SOS2045	1	1	€ 1,8663	€ 4,99
528	20030804	20130	SOS2046	1	1	€ 1,8663	€ 4,99
528	20030804	19781	SOS2048	1	1	€ 1,8663	€ 4,99
528	20030805	21366	SOS2071	1	1	€ 1,8663	€ 4,99
528	20030805	23831	SOS2082	1	1	€ 1,8663	€ 4,99
528	20030805	15861	SOS2094	1	1	€ 1,8663	€ 4,99
528	20030806	17027	SOS2129	1	1	€ 1,8663	€ 4,99

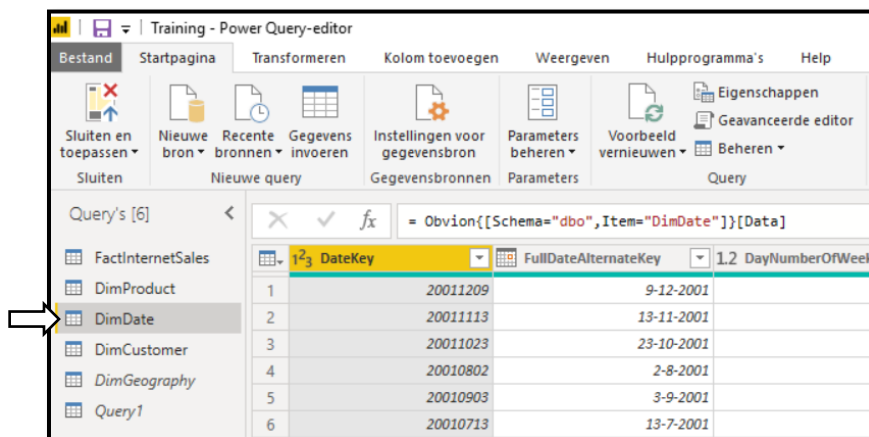
Je bent nu klaar met werken in de **FactInternetSales** query.

8. Sla je werk op door in Power BI Desktop op **Bestand** en op **Opslaan** te klikken, of door de toetsenbordcombinatie **ctrl + s**.

Oefening 5: Bewerken query's om datumdata te transformeren

In deze oefening gebruik je het **Power Query-editor** navigatiescherm om de **DimDate** query aan te passen om wijzigingen in de datumdata voordat deze in het model wordt ingelezen aan te brengen.

1. Open de **DimDate** query in bewerken modus.
 - a) Klik op **Gegevens transformeren** in het lint van het **Startpagina** tabblad om het **Power Query Editor** navigatiescherm te tonen.
 - b) Zorg ervoor dat de **DimDate** query geselecteerd is in de **Query's** lijst aan de linkerkant van de **Power Query-editor**.

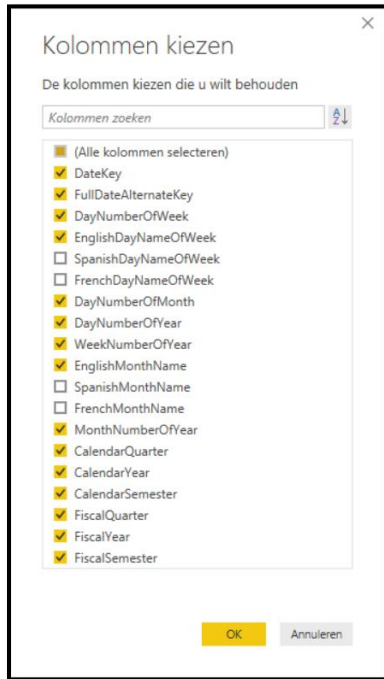


The screenshot shows the Power Query Editor interface. The 'Query's' list on the left has 'DimDate' selected. The main view shows the query formula and a preview of the data.

	123 DateKey	FullDateAlternateKey	1.2 DayNumberOfWeek
1	20011209	9-12-2001	
2	20011113	13-11-2001	
3	20011023	23-10-2001	
4	20010802	2-8-2001	
5	20010903	3-9-2001	
6	20010713	13-7-2001	

2. Verwijder de kolommen die niet nodig zijn in de **DimDate** query resultaten.

- a) Klik op de **Kolommen kiezen** knop in het lint om het **Kolommen kiezen** dialoogvenster te openen.
- b) In het **Kolommen kiezen** dialoogvenster, vink de volgende kolommen uit:
 - I) **SpanishDayNameOfWeek**
 - II) **FrenchDayNameOfWeek**
 - III) **SpanishMonthName**
 - IV) **FrenchMonthName**



- c) Klik op **OK** om het **Kolommen kiezen** dialoogvenster te sluiten en de onderliggende query aan te passen.
3. Voer de **DimDate** query uit om de resultaten te kunnen bekijken in Power BI Desktop.
- a) Druk op **Sluiten en toepassen** in het **Startpagina** tabblad van het lint om de **DimProduct** query uit te voeren.
 - b) Nadat de **DimDate** query is toegepast moeten alle resultaten zichtbaar zijn in de **Gegevens** weergave van Power BI Desktop wanneer de **DimDate** tabel wordt geselecteerd in de **Velden** lijst aan de rechterkant.

Je bent nu klaar met werken in de **DimDate** query.

4. Sla je werk op door in Power BI Desktop op **Bestand** en op **Opslaan** te klikken, of door de toetsenbordcombinatie **ctrl + s**.

Je hebt alle oefeningen gemaakt. Terwijl wij nog geen data modelleren hebben behandeld heb je wel alvast de tabellen verwerkt in een stermodel. Een stermodel is een veelgebruikte en bewezen methode voor het werken in data analyse projecten.