

## Power BI Desktop

### Datenimport und Datentransformation

Unsere Daten sind von der fiktiven US-amerikanischen Firma VanArsdel. Die Daten sind daher in US-englischem Format. Daher müssten Sie das **Gebietsschema** zu Beginn in «English (United States)» umändern.

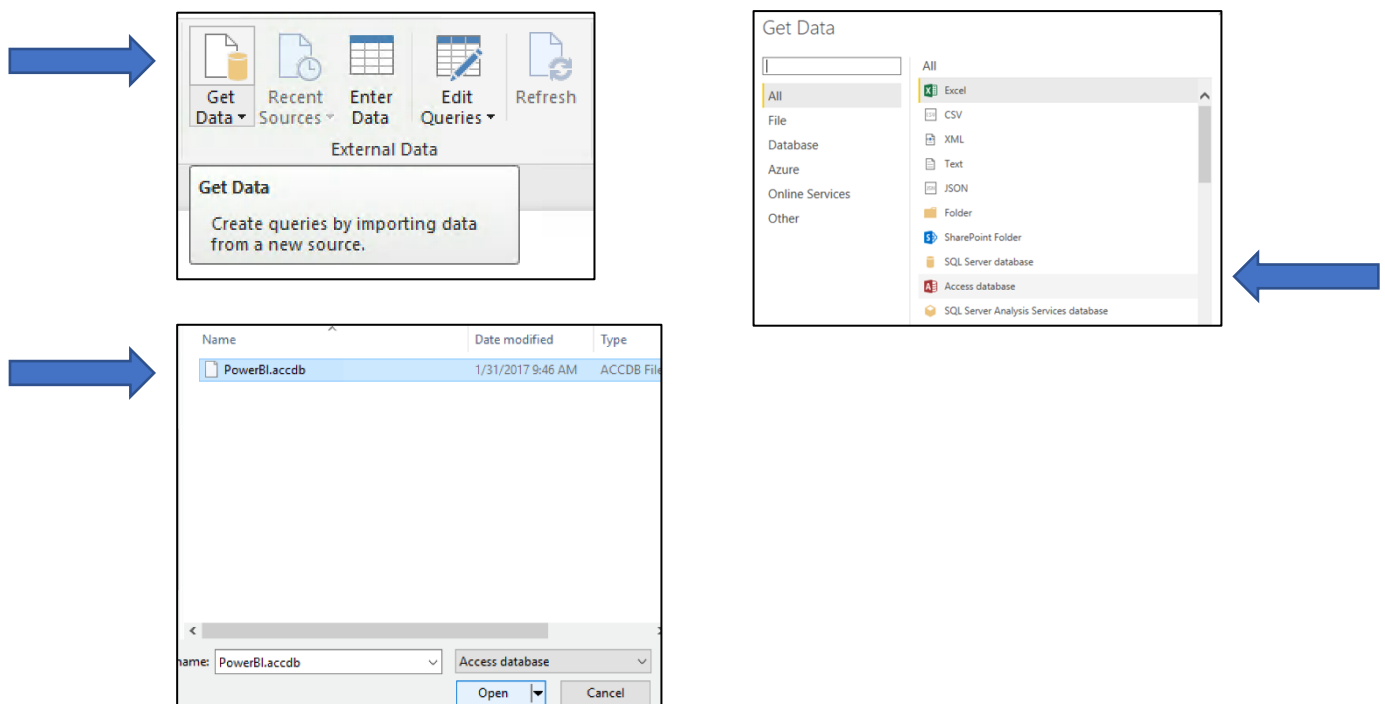
Falls Sie noch keinen Access-Datenbank Engine auf Ihrem Computer installiert haben, müssen Sie diesen vorgängig noch herunterladen und installieren:

<https://www.microsoft.com/de-de/download/details.aspx?id=13255>

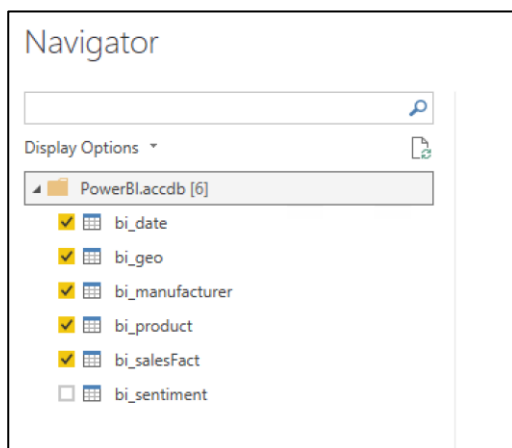
Speichern Sie die Übungsdaten im C: Laufwerk ab (C:\pbi01\daten) und entzippen Sie alle Dateien, die darin enthalten sind. Danach können Sie mit den Übungen starten.

### Access-Datenbank

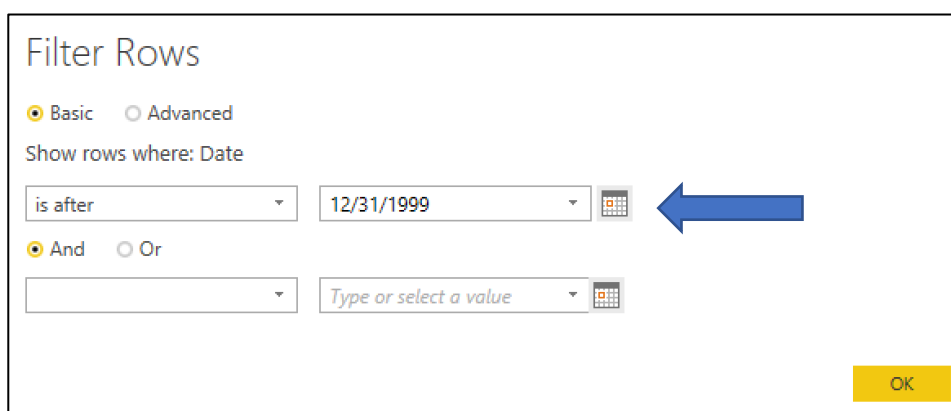
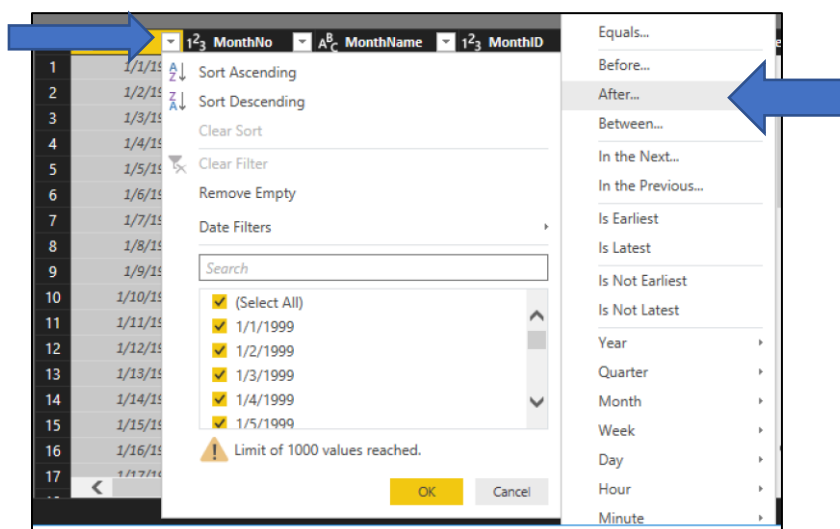
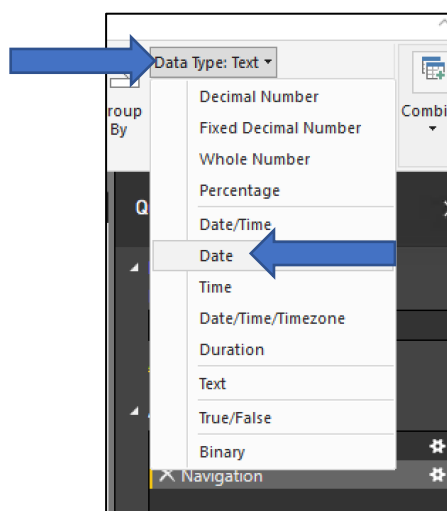
1. Starten Sie Power BI Desktop mit einem leeren Dokument
2. Verbinden Sie die Access Datenbank mit Power BI. Dazu klicken Sie auf Get Data und wählen Access Database als Quelle. Dann wählen Sie das gewünschte Access File.



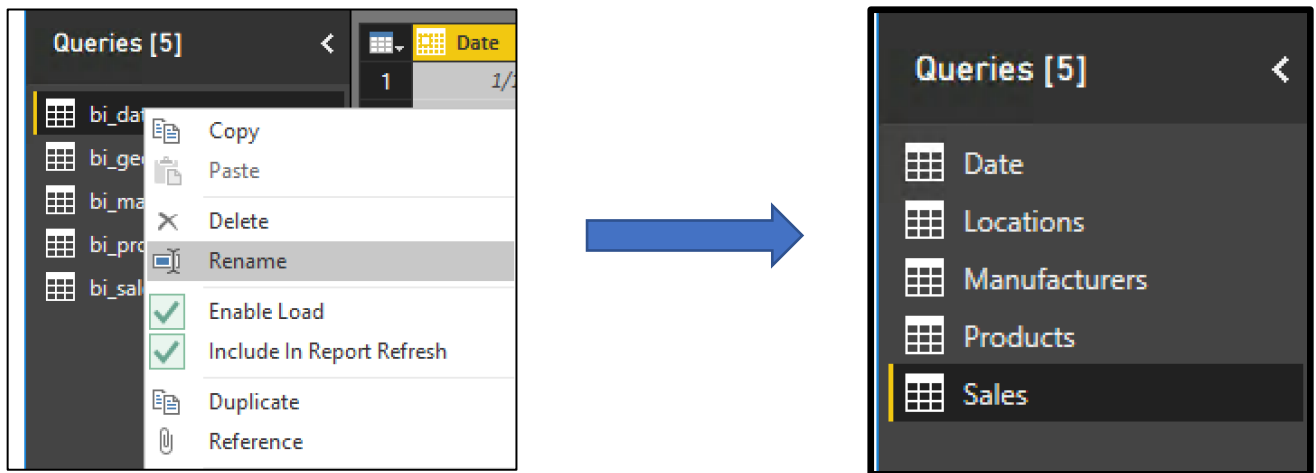
3. Wählen Sie folgende Ordner zum Import aus: **bi\_date**, **bi\_geo**, **bi\_manufacturer**, **bi\_product**, und **bi\_salesFact**. Klicken Sie auf Edit bevor Sie die Daten laden.



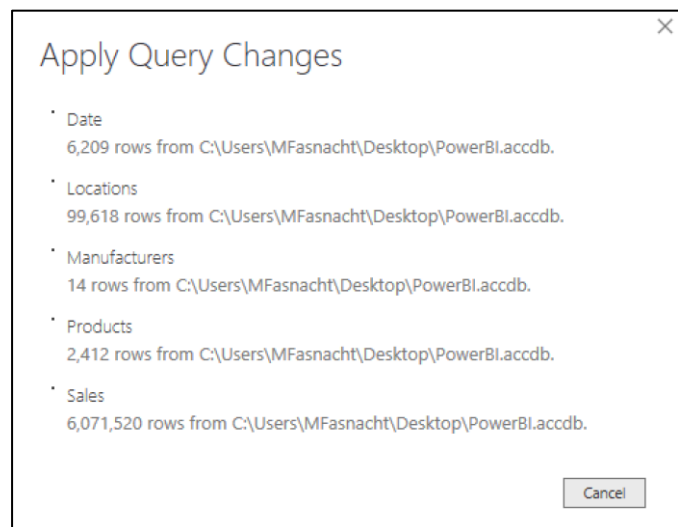
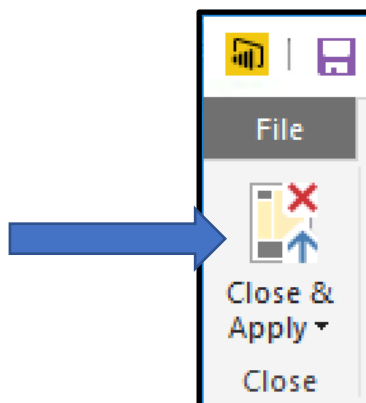
4. Filtern Sie die Reihen der bi\_salesFact Tabelle und der bi\_date Tabelle um Daten vom 1. Januar 2000 zu erhalten. Ändern Sie dazu den Datentyp von Date zu Date und führen Sie einen Datumsfilter ein der nur Daten nach 31. Dezember 1999 anzeigt.



5. Mit einem Rechtsklick auf die Tabellen können deren Namen geändert werden. Ändern Sie die Namen der Tabellen wie folgt:
- bi\_date: **Date**
  - bi\_geo: **Locations**
  - bi\_manufacturer: **Manufacturers**
  - bi\_product: **Products**
  - bi\_salesFact: **Sales**



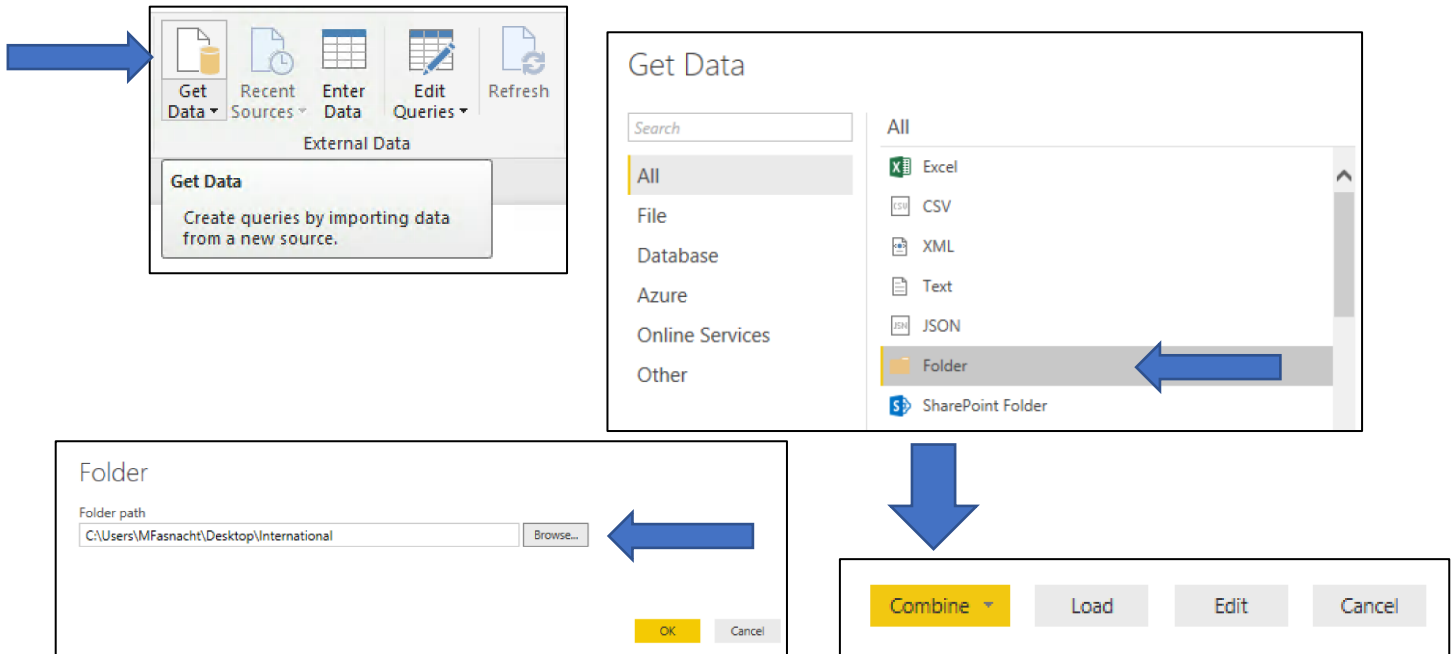
6. Laden Sie die Daten in Power BI indem Sie auf Close & Apply klicken.



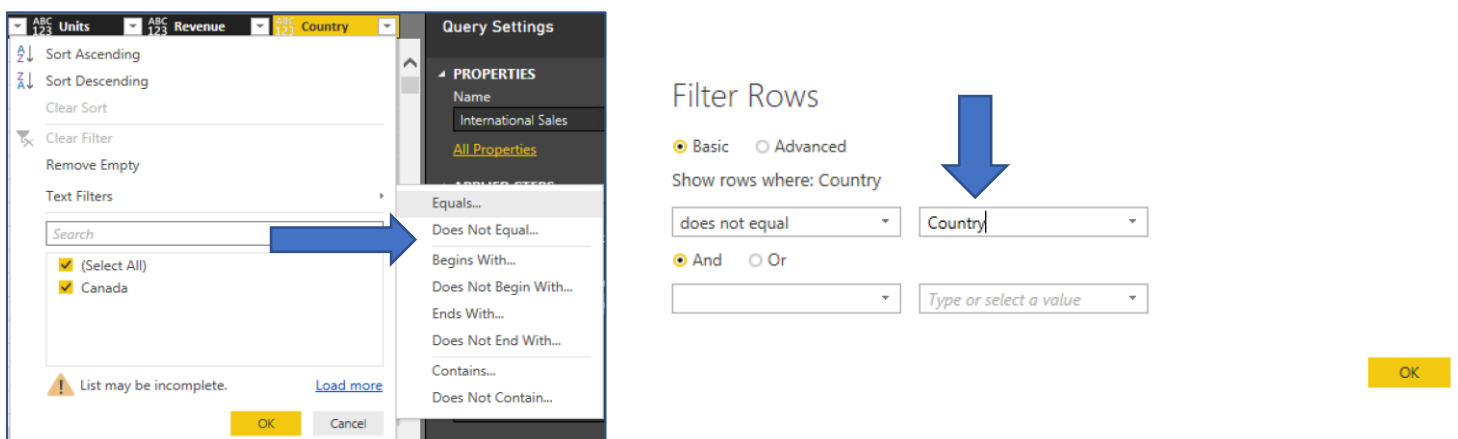
Speichern Sie Ihren Fortschritt ab und machen Sie mit der nächsten Übung weiter.

## Ordner mit CSV-Dateien

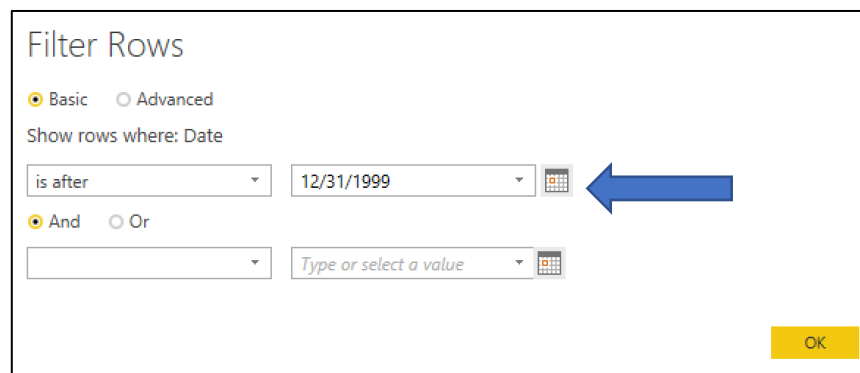
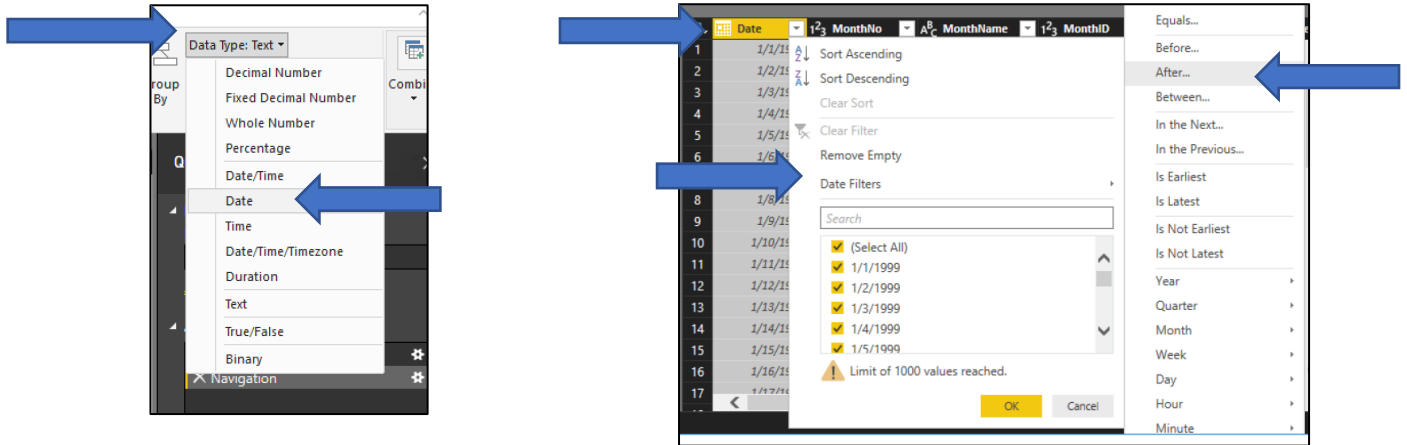
1. Starten Sie Power BI mit dem Dokument, welches Sie in der ersten Übung erstellt haben.
2. Laden Sie die CSV-Dateien in Power BI. Klicken Sie dazu auf Get Data, wählen Folder, dann den gewünschten Ordner mit den CSV Dateien (International) und klicken auf Combine und anschliessend auf «Combine & Edit».



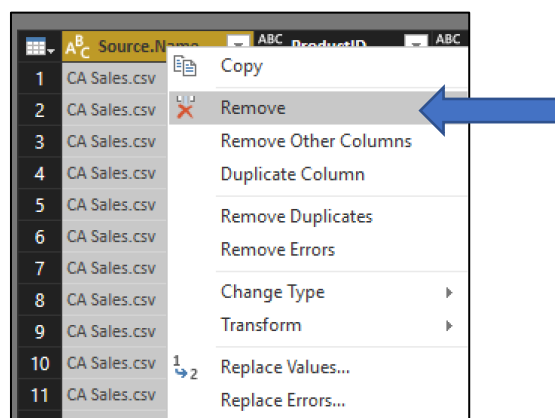
3. Führen Sie im Query Editor folgende Aktionen durch:
  - a. Filtern Sie die Reihen heraus die aus den Titel der CSV Files entstanden sind. Ein möglicher Weg dazu ist alle Einträge mit „Country“ aus der Country Spalte zu filtern. Dazu auf die Spalte Country klicken und „Text Filters“ wählen. Dann klicken Sie auf „Does Not Equal...“ und geben den Wert „Country“ ein.



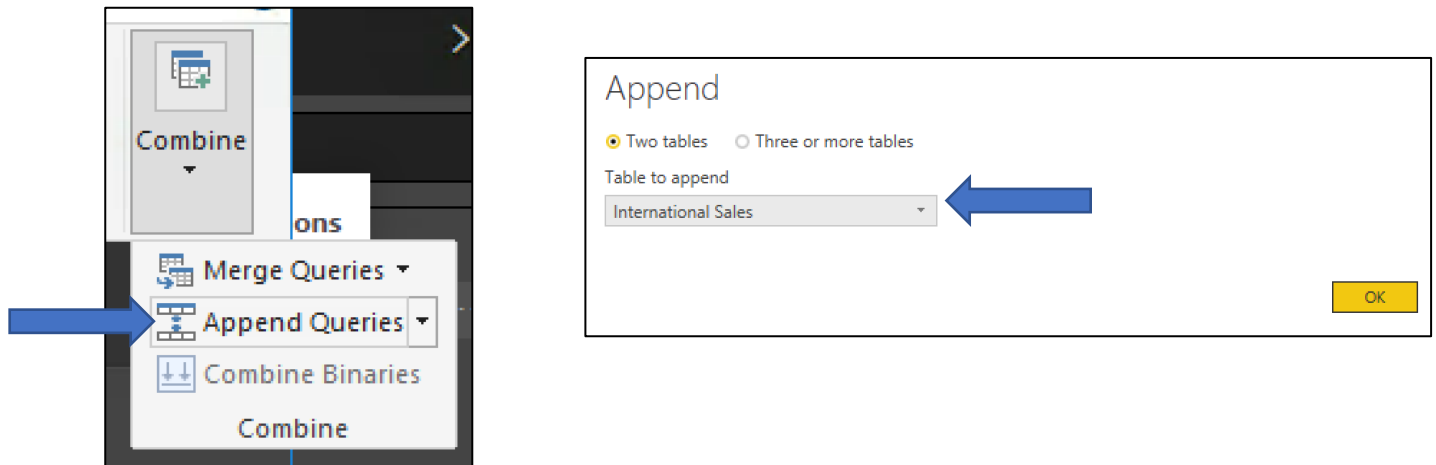
- b. Filtern Sie alle Einträge vor dem 31. Dezember 1999 heraus, so dass nur noch Einträge nach dem 31. Dezember 1999 enthalten sind. Dazu muss wieder zuerst der Datentyp der Date Spalte in Date geändert werden. Anschliessend soll ein Datumsfilter eingesetzt werden.



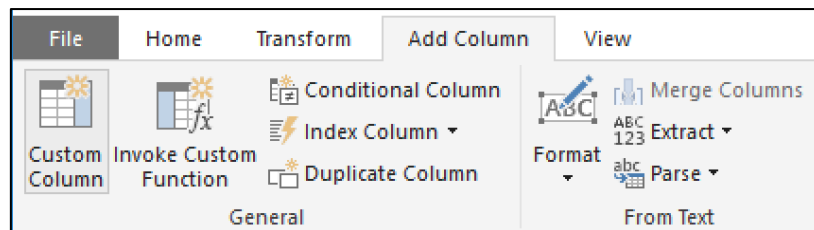
4. Entfernen Sie die erste Spalte des Datasets. Dazu genügt ein Rechtsklick auf den Titel und auf Remove.



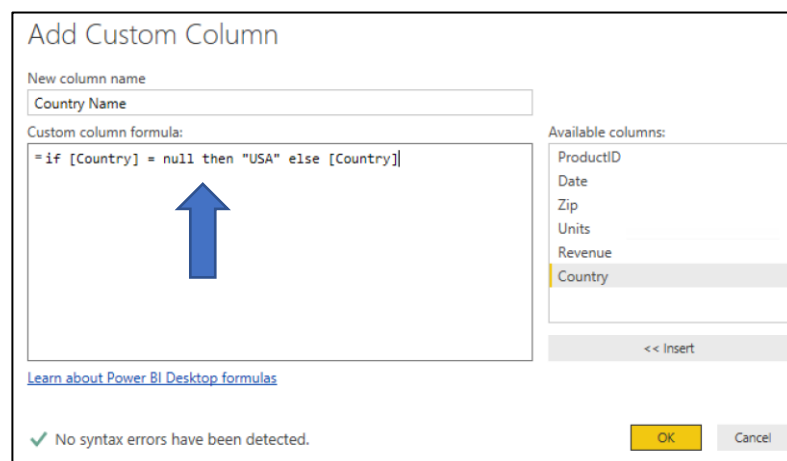
5. Hängen Sie die Daten aus der International Sales Tabelle an die Tabelle mit den US-Sales (Tabelle aus der vorherigen Übung). Dazu auf die US-Sales Tabelle klicken und oben unter Combine auf „Append Queries“ klicken. Unter Append die gewünschte Tabelle auswählen und mit OK bestätigen. Die International Sales wurden nun der Tabelle mit den US Sales hinzugefügt.



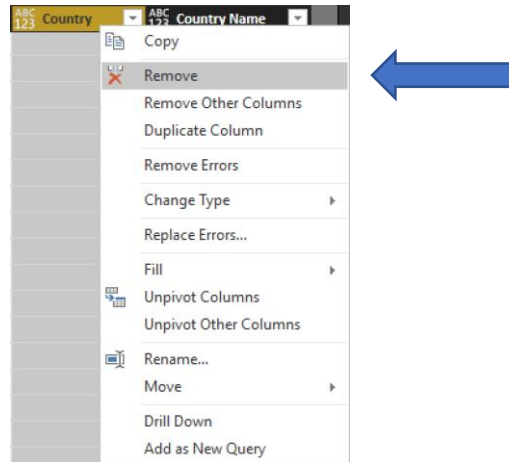
6. In der US-Sales Tabelle soll nun eine neue Spalte mit dem Namen des Landes entstehen.
  - a. Dazu unter „Add Column“ auf „Custom Column“ klicken um eine neue Spalte erstellen.



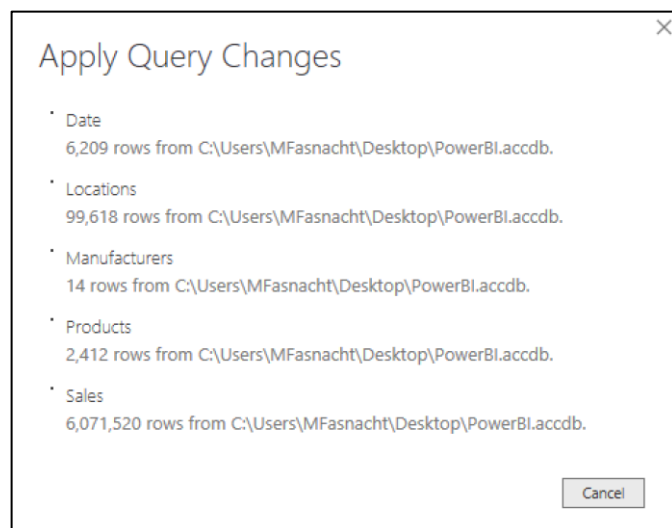
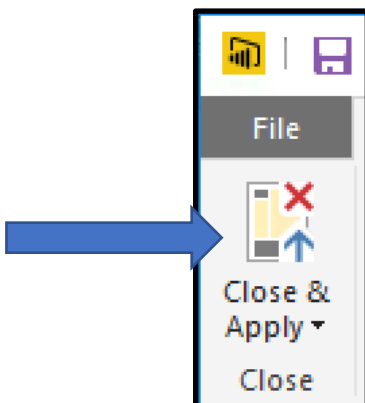
- b. Unter „Add Custom Column“ kann jetzt mit einer kurzen Formel die Spalte erstellt werden. Überall wo unter Country kein Eintrag (null) vorhanden ist, soll „USA“ eingefügt werden und bei allen anderen soll der Name des Landes erscheinen. Dazu folgenden Text in die Formelbox schreiben und mit Ok bestätigen.



- c. Anschliessend soll die andere Spalte „Country“ gelöscht werden. Dazu nach einem Rechtsklick auf den Namen auf „Remove“ klicken.



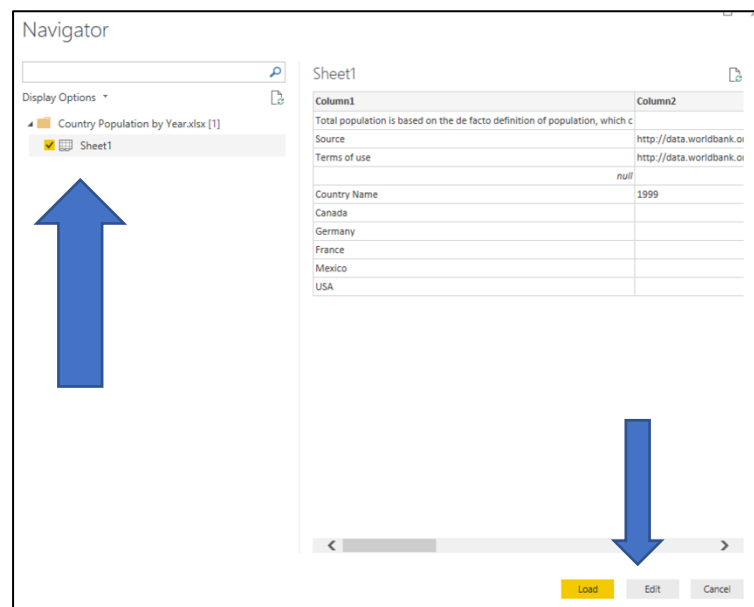
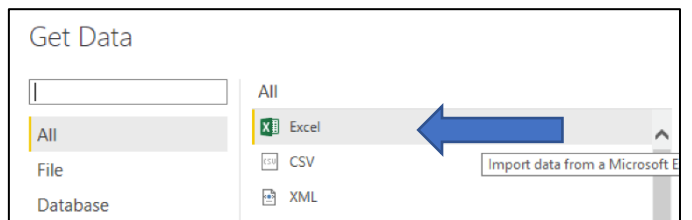
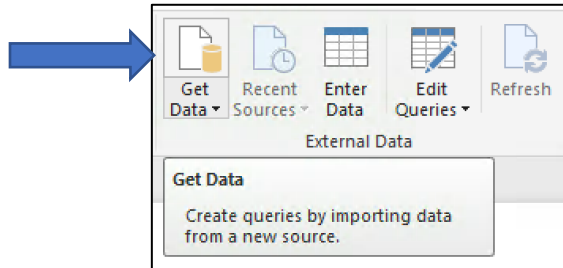
7. Laden Sie die Daten in Power BI indem Sie auf Close & Apply klicken. Der Vorgang könnte einige Minuten dauern da Sie etwas mehr als 10 Mio. Zeilen laden.



Speichern Sie Ihren Fortschritt ab und machen Sie mit der nächsten Übung weiter.

## Weniger strukturierte Excel-Datei

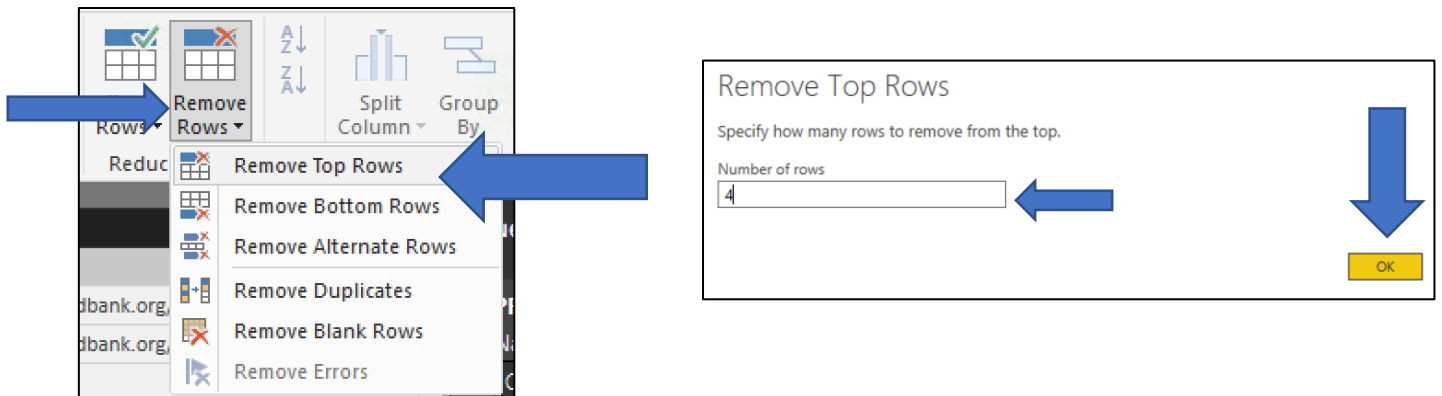
1. Starten Sie Power BI mit dem Dokument, welches Sie in der ersten Übung erstellt haben.
2. Laden Sie die Excel-Datei in Power BI. Klicken Sie dazu auf Get Data, wählen Excel, dann das gewünschte Dokument (Country Population By Year.xlsx) und klicken auf Edit.



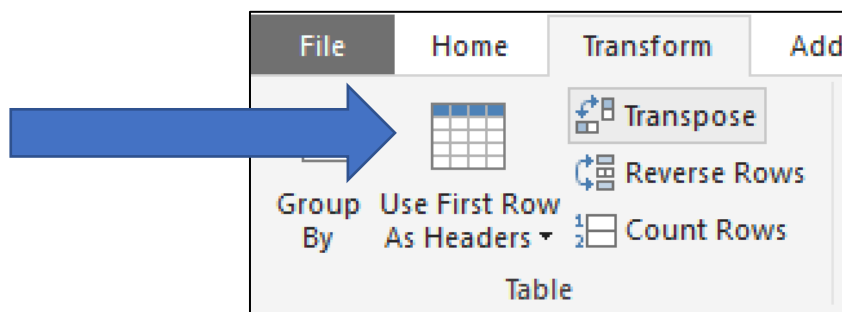


3. Führen Sie im Query Editor folgende Aktionen aus:

- a. Benennen Sie das Query in „Country Population“ um (Doppelklick auf den Namen)
- b. Entfernen Sie die ersten vier Zeilen der Tabelle.  
Dazu klicken Sie oben auf „Remove Rows“, wählen Remove „Top Rows“ und geben die gewünschte Anzahl Reihen ein. Mit Ok wird die Eingabe bestätigt.




- c. Die oberste Zeile der Tabelle enthält die Namen der Spalten. Um diese als Namen anzuzeigen klicken Sie im Transform Ribbon auf „Use First Row As Headers“.



- d. Sie möchten für alle fünf Einträge eine eigene Zeile. So dass bei jedem Land zuerst alle Jahre angezeigt werden bevor dann das nächste Land folgt. Dazu markieren Sie den gewünschten Bereich und klicken im Transform Ribbon auf „Unpivot Columns“.

Das Resultat sollte so aussehen:

Das Resultat sollte so aussehen:



1 2 Replace Values ▾ Unpivot Columns ▾

Fill ▾

Pivot Column

Any Column

Move ▾

Convert To List

Split Column ▾

Format

Text Column

Merge Columns

ABC 123 Extract ▾

Statistics Standard Scientific

10<sup>2</sup>

Trigon

Round

Inform

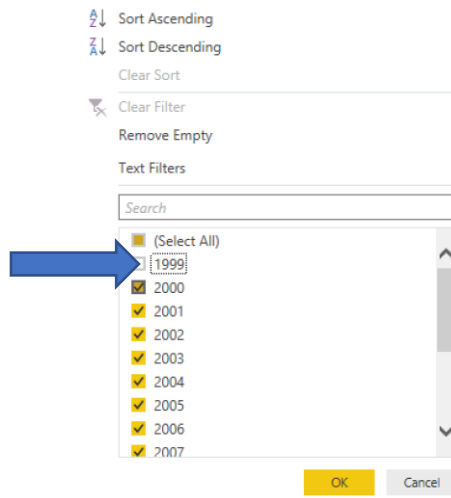
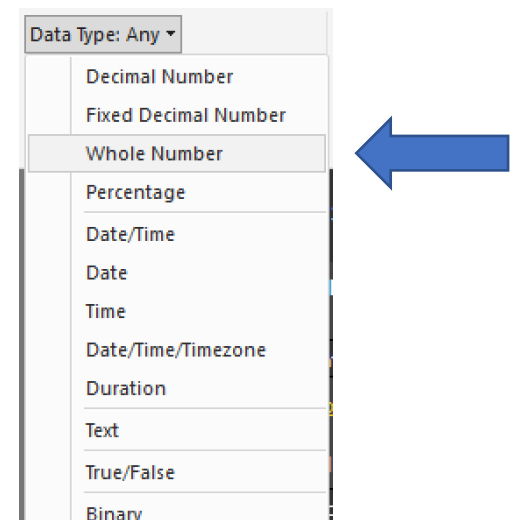
	1999	2000	2001	2002	2003	2004	2005
	30499200	30769700	31081900	31362000	31676000	31995000	32312000
	82100243	82211508	82349925	82488495	82534176	82516260	82469422
	60496715	60912498	61357431	61805267	62244884	62704897	63179356
	101266570	102808590	104239563	105578297	106888418	108257822	109747906
	279040000	282162411	284968955	287625193	290107933	292805298	295516599

	Country Name	Attribute	Value
1	Canada	1999	30499200
2	Canada	2000	30769700
3	Canada	2001	31081900
4	Canada	2002	31362000
5	Canada	2003	31676000
6	Canada	2004	31995000
7	Canada	2005	32312000
8	Canada	2006	32570505
9	Canada	2007	32887928
10	Canada	2008	33245773
11	Canada	2009	33628571
12	Canada	2010	34005274
13	Canada	2011	34342780
14	Canada	2012	34754312
15	Canada	2013	35158304
16	Canada	2014	35540419
17	Germany	1999	82100243

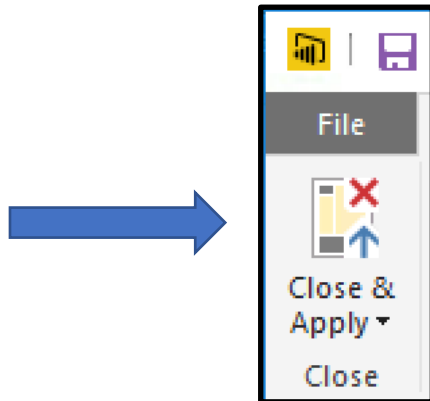
- e. Benennen Sie die neu entstandenen Spalten mit „Year“ beziehungsweise „Population“. Dazu genügt ein Doppelklick auf den Namen der Spalte.

Year	Population
1999	30499200
2000	30769700
2001	31081900

- f. Filtern Sie das Jahr 1999 heraus. Dazu auf den Namen der Spalte klicken und den Hacken bei 1999 entfernen. Setzen Sie danach die Datentypen von Year und Population auf „Whole Number“

4. Laden Sie die Daten in Power Bi indem Sie auf Close & Apply klicken.



Speichern Sie Ihren Fortschritt ab.  
Sie sind am Ende des ersten Übungsblockes angelangt.