

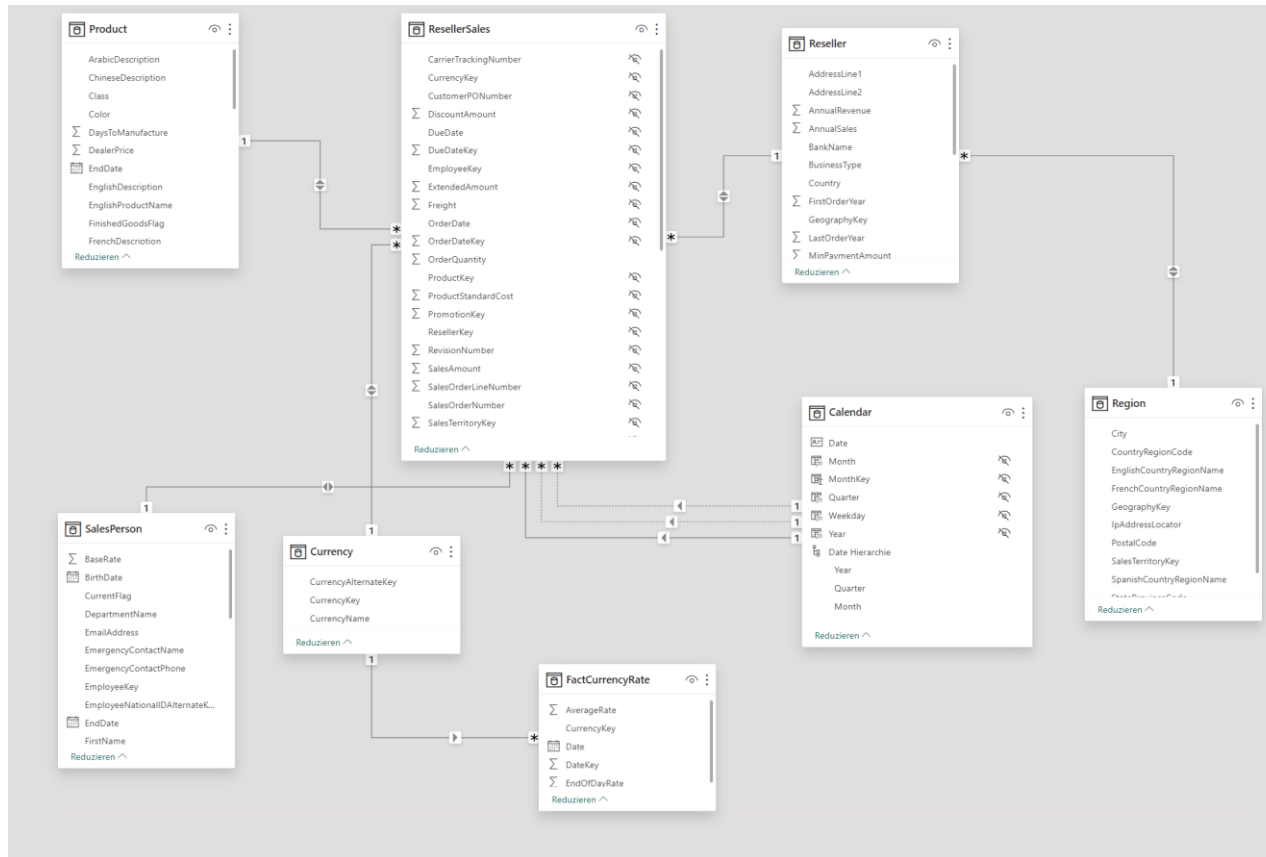
# Datenmodellierung

Datenmodellierung, Berechnungen, Measures und Neue Tabellen

# Datenmodell erstellen

- In dieser Übung werden Daten aus zwei unterschiedlichen Datenquellen geladen:
  - SQL Server und Excel
- Um ein Datenmodell zu erzeugen, kann eine Datenquelle oder bei Bedarf auch mehrere unterschiedliche Datenquellen geladen und zusammengeführt werden
- Nachfolgend wird eine SQL-Datenbank mitsamt mehrerer Tabellen als Relationen ins Datenmodell geladen
- Anschließend noch eine Excel-Datei die mit einer SQL Tabelle verbunden wird

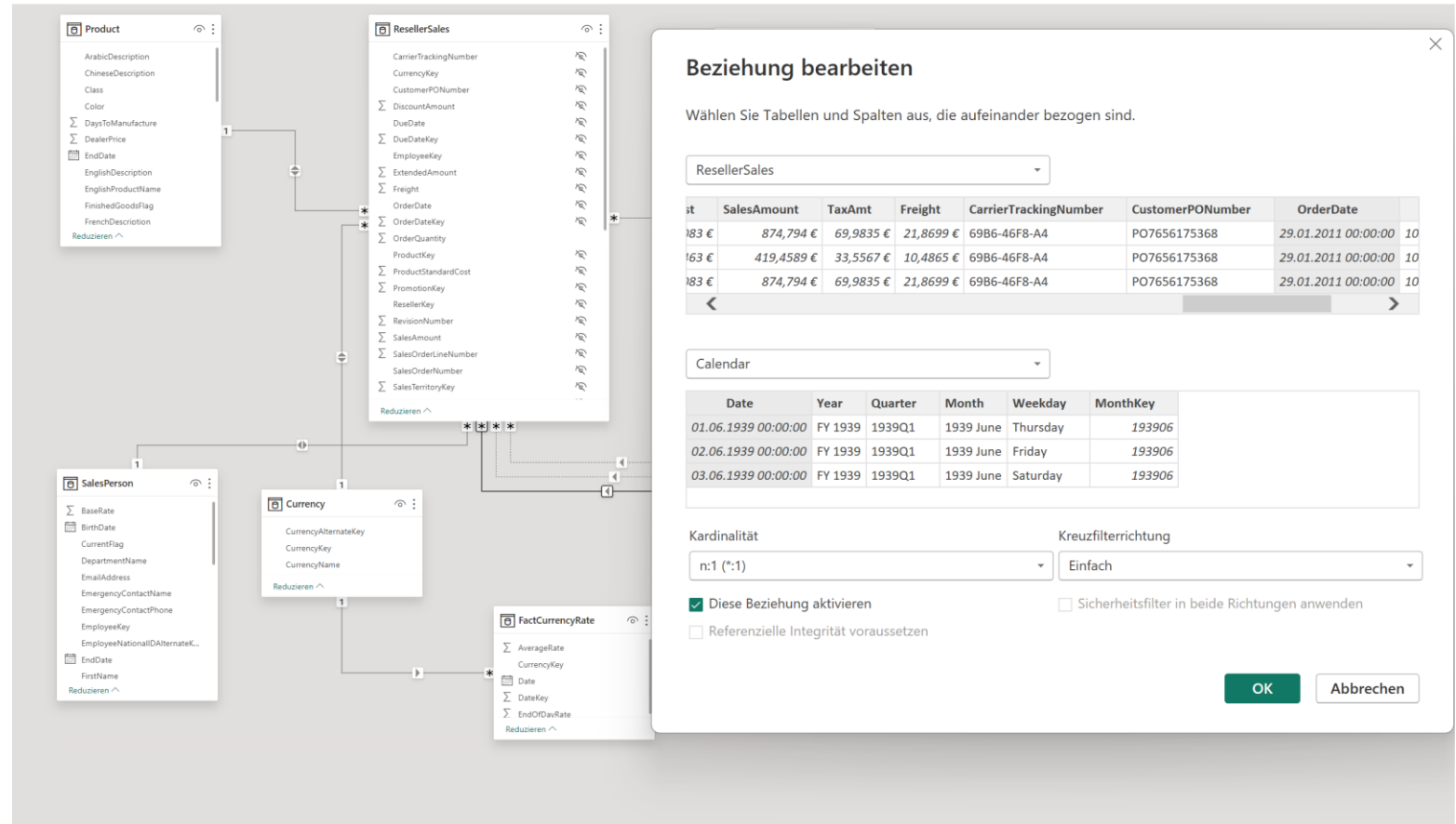
# Grundlegendes zu Sternschemas



- Im Sternschema haben wir eine Faktentabelle, die mit mehreren Dimensionstabellen verknüpft ist.
- Diese Trennung vereinfacht die Daten und bietet eine klare Hierarchie und einen klaren Kontext für die Analyse.
- Indem wir die Faktentabelle mit diesen Dimensionstabellen verknüpfen, richten wir Beziehungen ein und bestimmen die Detailebene in den Daten.

# Datenmodellierung

- Beziehungen sind vorhanden (automatische Erkennung)
- Beziehungen erstellen (Drag & Drop) von einer Spalte auf eine andere



**Beziehung bearbeiten**

Wählen Sie Tabellen und Spalten aus, die aufeinander bezogen sind.

ResellerSales

it	SalesAmount	TaxAmt	Freight	CarrierTrackingNumber	CustomerPONumber	OrderDate	
183 €	874,794 €	69,9835 €	21,8699 €	69B6-46F8-A4	PO7656175368	29.01.2011 00:00:00	10
163 €	419,4589 €	33,5567 €	10,4865 €	69B6-46F8-A4	PO7656175368	29.01.2011 00:00:00	10
183 €	874,794 €	69,9835 €	21,8699 €	69B6-46F8-A4	PO7656175368	29.01.2011 00:00:00	10

Calendar

Date	Year	Quarter	Month	Weekday	MonthKey
01.06.1939 00:00:00	FY 1939	1939Q1	1939 June	Thursday	193906
02.06.1939 00:00:00	FY 1939	1939Q1	1939 June	Friday	193906
03.06.1939 00:00:00	FY 1939	1939Q1	1939 June	Saturday	193906

Kardinalität: n:1 (\*:1)

Kreuzfilterrichtung: Einfach

☒ Diese Beziehung aktivieren

☐ Referenzielle Integrität voraussetzen

☐ Sicherheitsfilter in beide Richtungen anwenden

OK Abbrechen

# Daten mit relationalen Beziehungen

- Nun wird eine SQL Datenbank verbunden und mehrere in Beziehung zueinander stehende Daten geladen
- Daten abrufen -> SQL Server ->
  - Server angeben (z.B. win10-base) ->
  - Datenbank **leer** oder „**Northwind**“ eingeben ->
  - Datenkonnektivitätsmodus „**Importieren**“ wählen
- Zuerst wird die Tabelle „**Order Details**“ markiert -> „**Verknüpfte Tabellen auswählen**“ anklicken -> „**Orders**“ und „**Products**“ werden markiert -> „**Verknüpfte Tabellen auswählen**“ -> „**Categories**“, „**Customers**“, „**Employees**“, „**Shippers**“, „**Suppliers**“
- Dann noch manuell „**Territories**“, „**Region**“ und „**EmployeeTerritories**“ markieren
- Abschließend alle markierten Tabellen per Klick auf „**Laden**“ ins Datenmodell laden

# Beziehungen definieren

Beziehung bearbeiten

Wählen Sie Tabellen und Spalten aus, die aufeinander bezogen sind.

Orders

OrderID	CustomerID	EmployeeID	OrderDate	RequiredDate	ShippedDate	ShipVia	Freight
10249	TOMSP	6	05.07.1996 00:00:00	16.08.1996 00:00:00	10.07.1996 00:00:00	1	
10260	OTTIK	4	19.07.1996 00:00:00	16.08.1996 00:00:00	29.07.1996 00:00:00	1	
10267	FRANK	4	29.07.1996 00:00:00	26.08.1996 00:00:00	06.08.1996 00:00:00	1	

Zeitdimension

Datum	Jahr	Quartal	Monat	Monatsnummer	Tag
Montag, 1. Juli 1996	1996	3	Juli	7	1
Dienstag, 2. Juli 1996	1996	3	Juli	7	2
Mittwoch, 3. Juli 1996	1996	3	Juli	7	3

Kardinalität: Viele-zu-Eins (\*:1)

Kreuzfilterrichtung: Einfach

☒ Diese Beziehung aktivieren

☐ Sicherheitsfilter in beide Richtungen anwenden

☐ Referenzielle Integrität voraussetzen

OK Abbrechen

- Tabellen können verschiedene Beziehungen zu anderen Tabellen aufweisen, dabei spielen die „Kardinalität“ und die „Kreuzfilterrichtung“ eine Rolle
- Beziehungen können aktiviert oder wahlweise deaktiviert werden

# Beziehungen verändern

- Die geladenen Daten anzeigen per „**Modell**“ in der linken Leiste anzeigen
- Reiter Start -> „**Beziehungen verwalten**“ -> Kreuzfilterrichtung „**Beide**“ einstellen für alle Tabellen
- **Trainer Beispiel:** werden nicht alle Tabellen mit „Beide“ formatiert, kann es dazu führen, dass kumulierte Auswertungen über mehrere Tabellen hinweg nicht funktionieren werden!

# Beziehungen bearbeiten

## Beziehung bearbeiten

Wählen Sie Tabellen und Spalten aus, die aufeinander bezogen sind.

FactCurrencyRate

CurrencyKey	DateKey	AverageRate	EndOfDayRate	Date
100	20101229	1	1	29.12.2010 00:00:00
100	20101230	1	1	30.12.2010 00:00:00
100	20101231	1	1	31.12.2010 00:00:00

Currency

CurrencyKey	CurrencyAlternateKey	CurrencyName
1	AFA	Afghan
2	DZD	Algerian Dinar
3	ARS	Argentine Peso

Kardinalität

n:1 (\*:1)

Kreuzfilterrichtung

Einfach

☒ Diese Beziehung aktivieren

☐ Sicherheitsfilter in beide Richtungen anwenden

☐ Referenzielle Integrität voraussetzen

OK

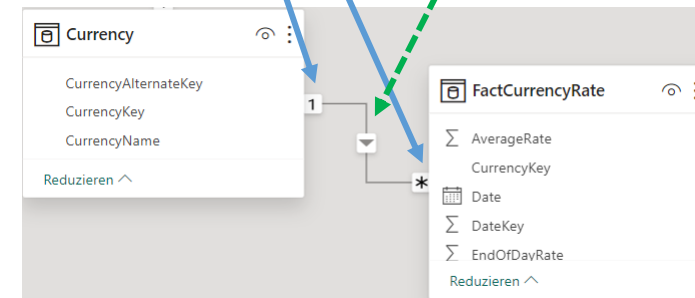
Abbrechen

Kardinalität

n:1 (\*:1)  
n:1 (\*:1)  
1:1 (1:1)  
1:n (1:\*)  
m:n (\*:\*)

Kreuzfilterrichtung

Einfach  
Einfach  
Beide

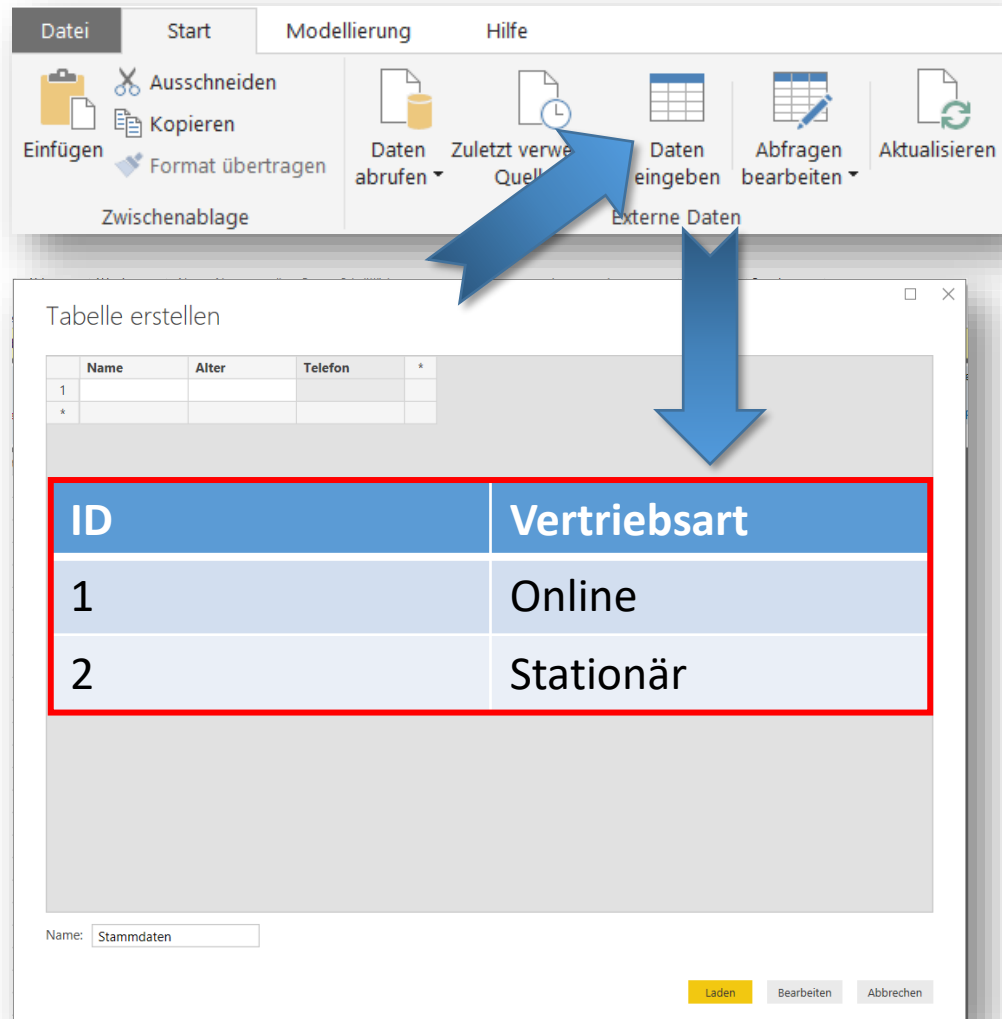




# Datenstrukturen in Power BI erstellen

- Daten werden üblicherweise ins Power BI Datenmodell hineingeladen
- Gleichwohl ist es möglich diese ebenso in Power BI manuell einzugeben, was zugegebenermaßen kaum sinnvoll wäre
- Aus einer anderen Datenquelle kopieren und in Power BI hineinkopieren – das könnte schon mehr Sinn machen
- Es kann Sinn machen, wenn z.B. eine kleine Dimensionstabelle mit zwei Spalten und z.B. vier Zeilen benötigt würde, um Daten nach Regionen zu gruppieren.
- Erstelle nach dem auf der nächsten Folie dargestelltem Vorgehen eine kleine Tabelle

# Daten manuell eingeben



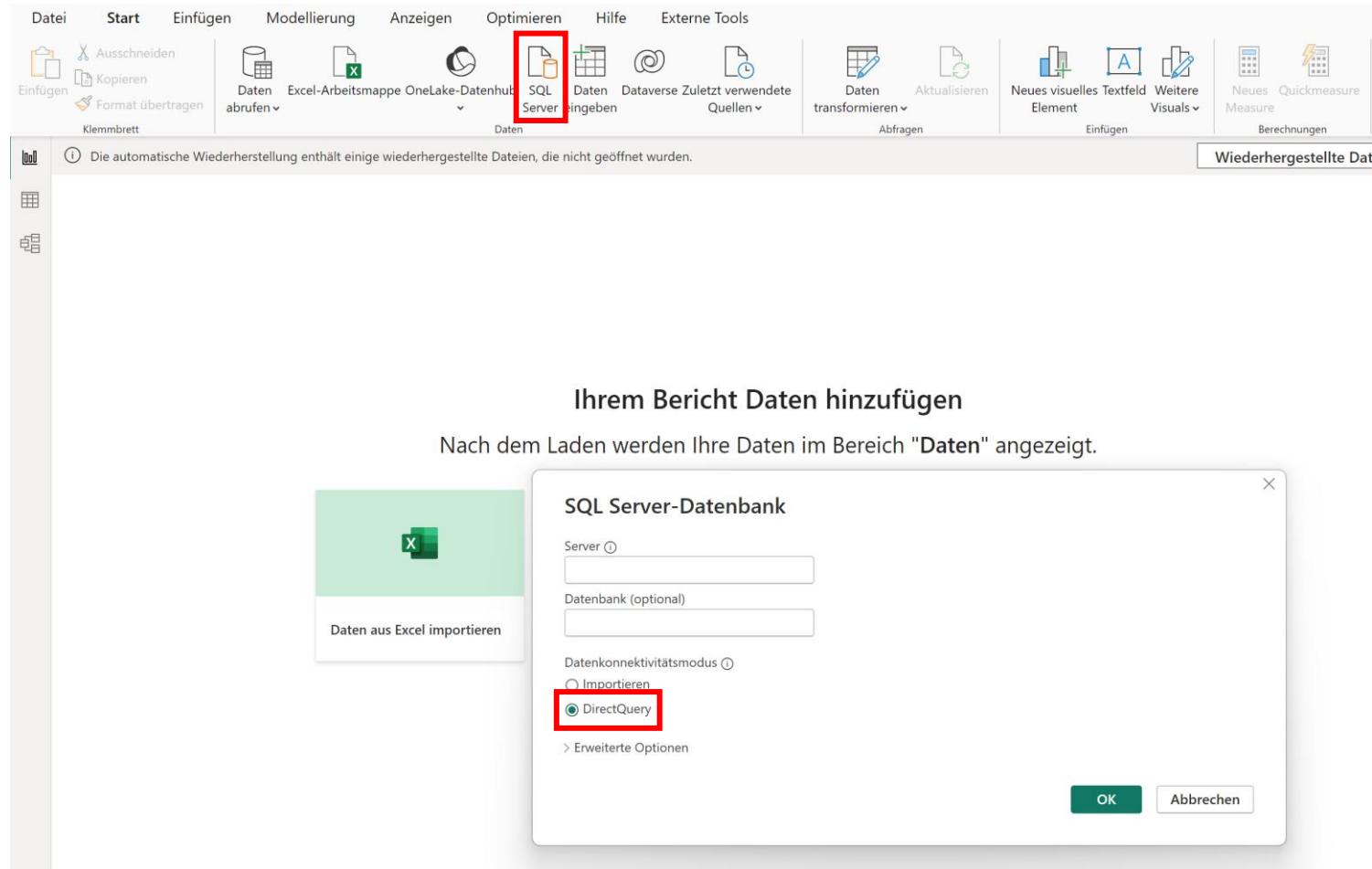
- Register „**Start**“
- Schaltfläche -> „**Daten eingeben**“
- Spaltenüberschriften definieren
- Inhalte händisch eingeben oder hineinkopieren, z.B. aus Excel
- Name der Tabelle -> „**Laden**“

Hinweis: sobald die Daten geladen worden sind, ist weiterhin eine nachträgliche Bearbeitung möglich!

# Einführung in DirectQuery

Daten abrufen und verarbeiten

# DirectQuery verwenden



The screenshot shows the Microsoft Power BI Desktop interface. The 'Optimieren' (Optimize) tab is selected in the ribbon, and the 'SQL Server' icon is highlighted with a red box. Below the ribbon, a message states: 'Die automatische Wiederherstellung enthält einige wiederhergestellte Dateien, die nicht geöffnet wurden.' (The automatic recovery contains some recovered files that were not opened.)

In the center, a dialog box titled 'Ihrem Bericht Daten hinzufügen' (Add data to your report) is displayed. It contains the text: 'Nach dem Laden werden Ihre Daten im Bereich "Daten" angezeigt.' (After loading, your data will be displayed in the "Data" area.)

On the left side of the dialog, there is a green button labeled 'Daten aus Excel importieren' (Import data from Excel). On the right side, there is a section titled 'SQL Server-Datenbank' (SQL Server database) with the following fields:

- Server: [Empty text box]
- Datenbank (optional): [Empty text box]
- Datenkonnektivitätsmodus (Data connectivity mode):
  - ☐ Importieren (Import)
  - ☒ DirectQuery (DirectQuery) - This option is highlighted with a red box.
- > Erweiterte Optionen (Advanced options)

At the bottom right of the dialog, there are two buttons: 'OK' and 'Abbrechen' (Cancel).

# Vorteile und Nachteile von DirectQuery

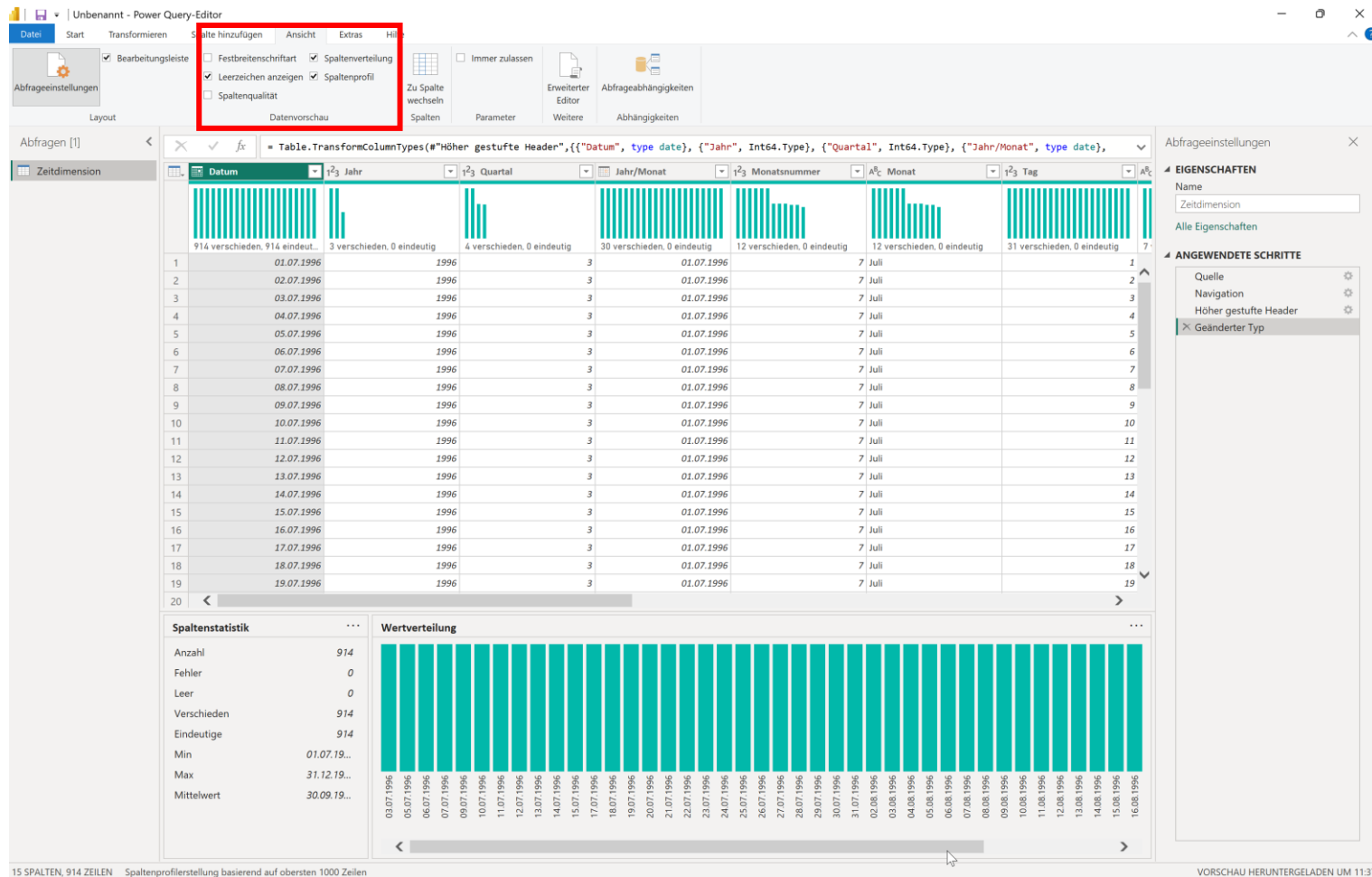
## Vorteile

- Bei häufig ändernden Daten
- Wenn Echtzeit erforderlich
- Große Datenmengen
- Mehrdimensionale Daten

## Nachteile

- Sind abhängig von der Datenquellenleistung
- Eingeschränkte Modellierungsfunktionen
- Eingeschränkte Transformationsfeatures
- DAX Funktionen eingeschränkt

# Datenprofilerstellung im Power Query-Editor



The screenshot shows the Power Query Editor interface. The 'Ansicht' (View) tab is selected in the ribbon, and the 'Spaltenqualität' (Column Quality) and 'Spaltenverteilung' (Column Distribution) options are checked. The main area displays a data table with columns: Datum, Jahr, Quartal, Jahr/Monat, Monatsnummer, Monat, and Tag. Each column has a corresponding bar chart showing the distribution of values. The bottom left pane shows the 'Spaltenstatistik' (Column Statistics) for the 'Datum' column, and the bottom right pane shows the 'Wertverteilung' (Value Distribution) for the 'Datum' column.

**Spaltenstatistik**

Statistik	Wert
Anzahl	914
Fehler	0
Leer	0
Verschieden	914
Eindeutige	914
Min	01.07.19...
Max	31.12.19...
Mittelwert	30.09.19...

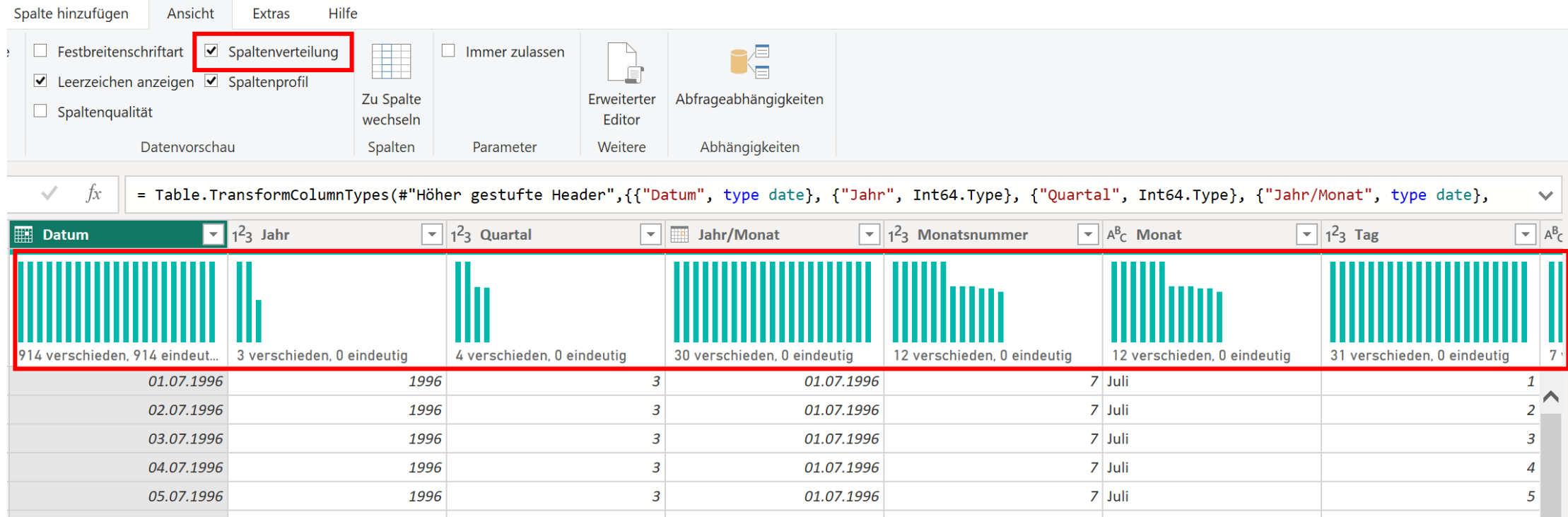
**Wertverteilung**

15 SPALTEN, 914 ZEILEN Spaltenprofilerstellung basierend auf obersten 1000 Zeilen

VORSCHAU HERUNTERGELADEN UM 11:37

- Register -> „Ansicht“
- Optionen einschalten
  - Spaltenqualität
  - Spaltenverteilung
  - Spaltenprofil

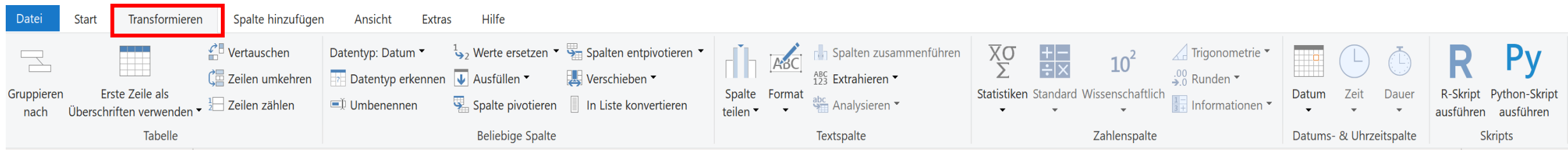
# Kardinalität einer Spalte bestimmen



- Anzahl der unterschiedlichen Werte: Gesamtzahl der verschiedenen Werte, die gefunden wurden (geringe Kardinalität)
- Anzahl der eindeutigen Werte: Gesamtzahl der Werte, die nur einmal vorkommen (hohe Kardinalität)
- Eine geringere Kardinalität führt zu einer besseren Leistung

# Daten transformieren

- Zu den häufigen Transformationen im Power Query-Editor zählen das Umbenennen oder Entfernen von Spalten „Spalte hinzufügen“ und Formatierungen im Menüband „Transformieren“.





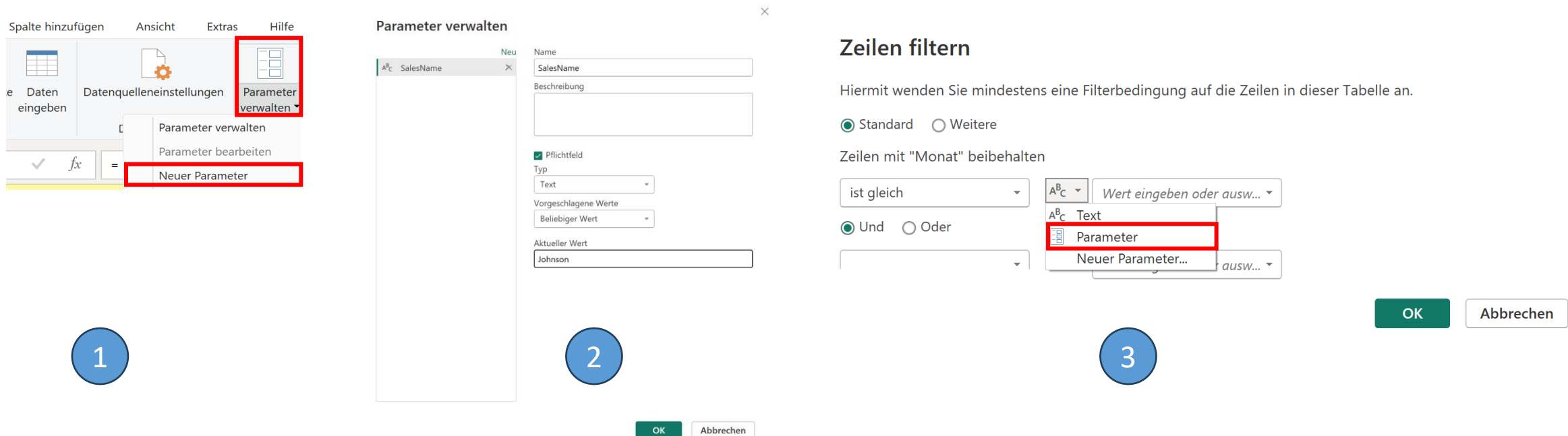
# Tabellen strukturieren

[illegible]

The screenshot shows the Microsoft Access ribbon with the 'Entfernen' (Remove) button highlighted in the 'Zeilen' (Rows) group. A context menu is open, listing options like 'Erste Zeilen entfernen' (Remove first rows), 'Untere Zeilen entfernen' (Remove bottom rows), 'Alternative Zeilen entfernen' (Remove alternate rows), 'Duplikate entfernen' (Remove duplicates), 'Leere Zeilen entfernen' (Remove empty rows), and 'Fehler entfernen' (Remove errors). The background shows a table with columns 'Quartal' and 'Jahr'.

# Parameter in Berichten verwenden

- Es können Parameter verwendet werden, indem Sie die Werte festlegen, für die Daten im Bericht angezeigt werden sollen. Daraufhin wird der Bericht durch Filtern der Daten entsprechend aktualisiert.



The screenshot illustrates the process of using parameters in reports, divided into three numbered steps:

- Step 1:** The 'Extras' menu is open, and the 'Parameter verwalten' option is highlighted with a red box. Below it, the 'Neuer Parameter' option is also highlighted with a red box.
- Step 2:** The 'Parameter verwalten' dialog box is shown. It contains fields for 'Name' (SalesName), 'Beschreibung', 'Typ' (Text), 'Vorgeschlagene Werte' (Beliebiger Wert), and 'Aktueller Wert' (Johnson). The 'Pflichtfeld' checkbox is checked. The 'OK' button is highlighted with a red box.
- Step 3:** The 'Zeilen filtern' dialog box is shown. It contains a dropdown menu for 'ist gleich' and a dropdown menu for 'Wert eingeben oder ausw...'. The 'Parameter' option is highlighted with a red box. The 'OK' button is highlighted with a red box.

Buttons: OK, Abbrechen

# Empfehlungen zur Leistung in Power BI

- Bewahren Sie nur erforderliche Daten auf.
- Überprüfen Sie die Datentypen.
- Verringern der Kardinalität
- Datenkonvertierung in PowerQuery statt Power BI per DAX
- Spaltenberechnungen (Indexspalte erstellen, Mehrwertsteuer berechnen) vor dem Laden in Power Query erstellen
- Parameter verwenden