

# 3

## Modellierung von Daten in Power BI

# 3.1

## Entwerfen eines Datenmodells in PowerBI

## Vorteile eines guten Datenmodells

- Durchsuchen von Daten erfolgt schneller
- Aggregationen können einfacher erstellt werden
- Berichte sind genauer
- Schreiben von Berichten nimmt weniger Zeit in Anspruch
- Berichte sind in Zukunft einfacher zu verwalten

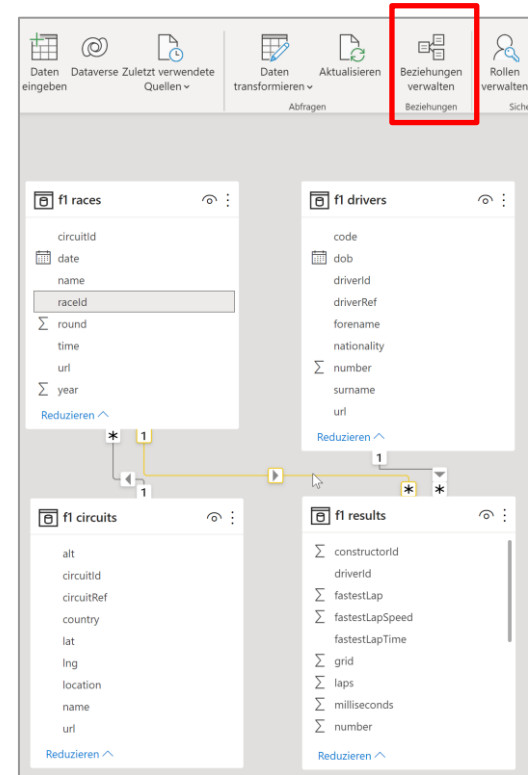
## Allgemeine Regeln

- Generell sind kleinere Datenmodelle mit weniger Tabellen und weniger Spalten in jeder Tabelle besser
- Sternschema (
  - Faktentabellen mit den beobachteten Daten
  - Dimensionstabellen mit den Informationen zu bestimmten Entitäten
- Beziehungen zwischen den Tabellen sind 1:n, sofern möglich

In der Registerkarte Modell doppelt auf eine bestehende Verbindung klicken oder im Menüband auf *Beziehungen verwalten*.

Die Beziehungen bestimmen, ob sich das Datenmodell z.B. bei Filterungen korrekt verhält. Es ist also sehr wichtig, dass diese korrekt gesetzt sind.

Generell sind keine "Schleifen" möglich. Dann muss das Datenmodell geändert werden.



Bei einer Beziehung werden die identifizierenden Spalten der beiden Tabellen festgelegt.

Zusätzlich wird die **Kardinalität** (1:1, 1:n, n:1, n:m) bestimmt und die **Kreuzfilterrichtung**. Bei *einfach* werden Filterungen in verknüpften Tabellen nur in eine Richtung angepasst, bei *beide* in beide. Beide ist immer mit Bedacht einzusetzen.

### Beziehung bearbeiten

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

f1 results

resultid	raceld	driverid	constructorid	number	grid	position	positionText	positionOrder
2230	125	22	6	2	3	null	R	14
3121	166	30	6	3	1	null	R	16
3361	177	56	6	4	4	null	R	14

f1 races

raceld	year	round	circuitid	name	date	time	
90	2004	1	1	Australian Grand Prix	Sonntag, 7. März 2004	null	http://en.wikipedia.org/w
91	2004	2	2	Malaysian Grand Prix	Sonntag, 21. März 2004	null	http://en.wikipedia.org/w
92	2004	3	3	Bahrain Grand Prix	Sonntag, 4. April 2004	null	http://en.wikipedia.org/w

Kardinalität

n:1 (\*:1)

Kreuzfilterrichtung

Einfach

☒ Diese Beziehung aktivieren

☐ Sicherheitsfilter in beide Richtungen anwenden

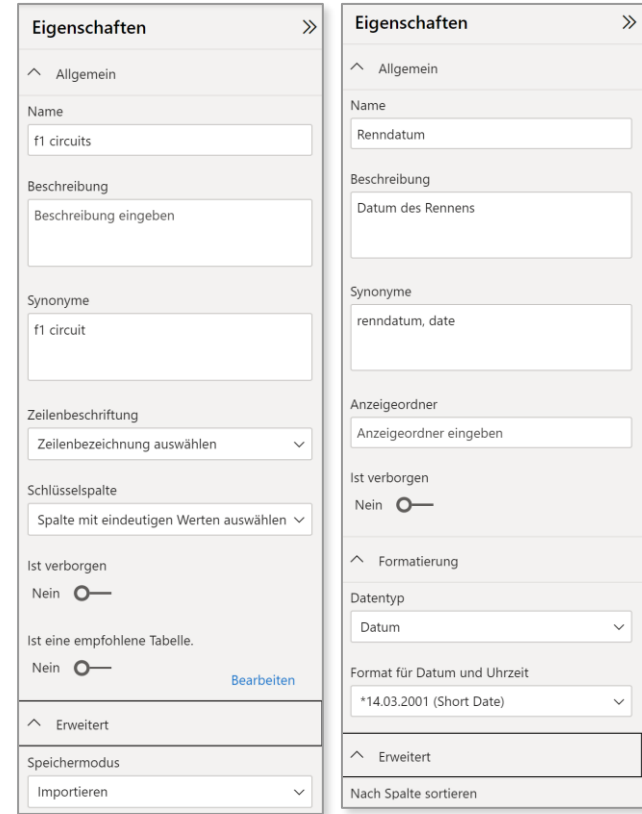
☐ Referenzielle Integrität voraussetzen

OK

Abbrechen

Im Tab Eigenschaften können  
**Tabelleneigenschaften** (z.B. Name, Speichermodus, Schlüsselspalte/Primary Key) als auch **Spalteneigenschaften** (z.B. Name, Beschreibung, Formatierung, Verbergen der Spalte) gesetzt werden

Auswahl mehrerer Tabellen oder Spalten/Felder mit Shift bzw. Strg



The image displays two side-by-side screenshots of the 'Eigenschaften' (Properties) tab in the DataCraft application, showing settings for a table and a column.

**Left Screenshot (Table Properties):**

- Allgemein**
  - Name: f1 circuits
  - Beschreibung: Beschreibung eingeben
  - Synonyme: f1 circuit
  - Zeilenbeschriftung: Zeilenbezeichnung auswählen
  - Schlüsselspalte: Spalte mit eindeutigen Werten auswählen
  - Ist verborgen: Nein
  - Ist eine empfohlene Tabelle: Nein
  - [Bearbeiten](#)
- Erweitert**
  - Speichermodus: Importieren

**Right Screenshot (Column Properties):**

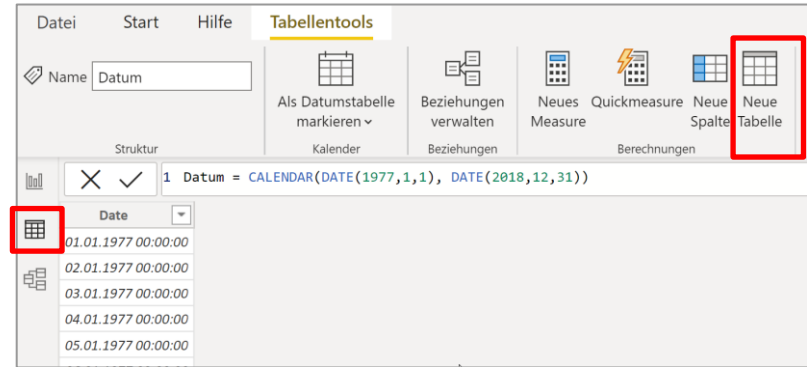
- Allgemein**
  - Name: Renndatum
  - Beschreibung: Datum des Rennens
  - Synonyme: renndatum, date
  - Anzeigereihenfolge: Anzeigereihenfolge eingeben
  - Ist verborgen: Nein
  - Formatierung
    - Datentyp: Datum
    - Format für Datum und Uhrzeit: \*14.03.2001 (Short Date)
- Erweitert**
  - Nach Spalte sortieren

Es wird häufig eine Datumstabelle benutzt, um z.B. zwischen Kalender- und Geschäftsjahr zu unterscheiden.

Es gibt drei Möglichkeiten, diese zu erstellen:

- Erstellen in den Quelldaten (csv, Datenbank)
- DAX (Data Analysis Expression): Funktionen CALENDAR() oder CALENDARAUTO()
- Power Query

- In der Tabellenansicht eine neue Tabelle erstellen
- Die Datumsspalte mittels CALENDAR() definieren
- Weitere neue Spalten:  
Jahr = year(Datum[Datum])  
Monat = MONTH(Datum[Datum])  
KW = WEEKNUM(Datum[Datum])  
Wochentag = FORMAT(Datum[Datum], "DDDD")
- Markieren als Datumstabelle



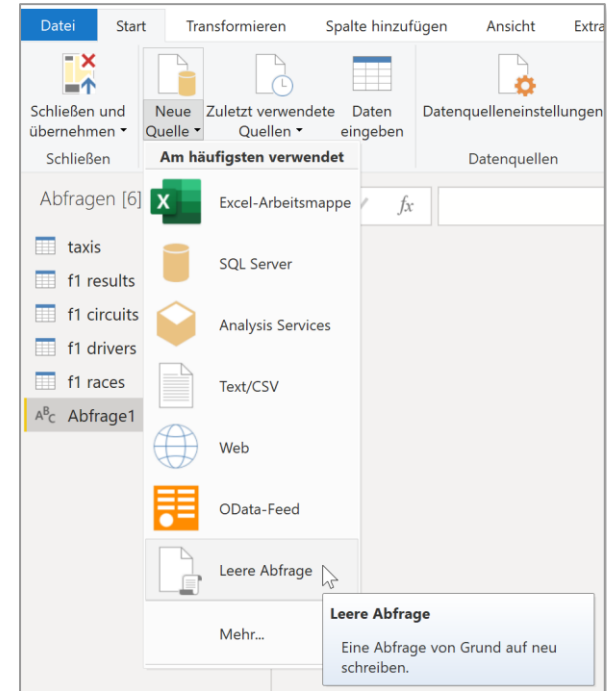


Start Menüband -> Daten transformieren -> Neue Quelle -> Leere Abfrage

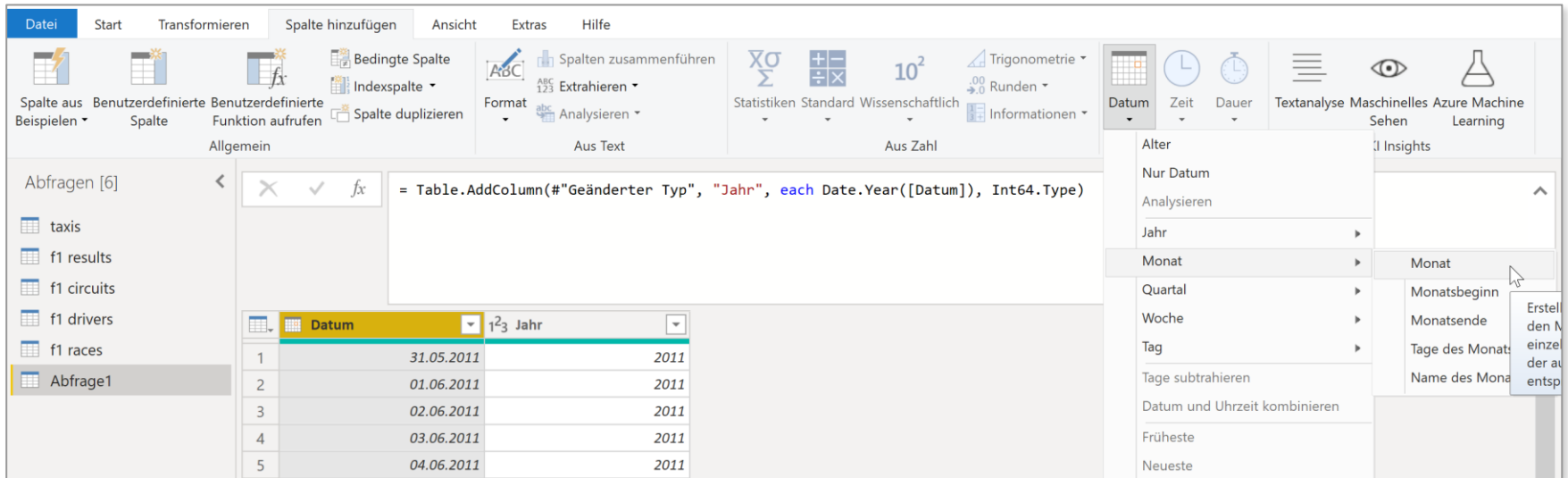
M-Formel:

```
= List.Dates(  
    #date(2011,05,31),           // Startdatum  
    365*10,                     // Anzahl  
    #duration(1,0,0,0))         // jeden Tag
```

Die erzeugte Liste muss noch in eine Tabelle transformiert werden.



## Hinzufügen von weiteren Spalten über Spalte hinzufügen -> Datum



The screenshot shows the Power Query Editor interface. The 'Spalte hinzufügen' (Add Column) tab is active, and the 'Datum' (Date) submenu is open. The main area displays a table with columns 'Datum' and 'Jahr'. The formula bar shows the M code: `= Table.AddColumn("#Geänderter Typ", "Jahr", each Date.Year([Datum]), Int64.Type)`.

**Spalte hinzufügen (Add Column) Submenu:**

- Alter
- Nur Datum
- Analysieren
- Jahr
  - Monat
    - Monat
    - Monatsbeginn
    - Monatsende
    - Tage des Monats
    - Name des Monats
  - Quartal
  - Woche
  - Tag
  - Tage subtrahieren
  - Datum und Uhrzeit kombinieren
  - Früheste
  - Neueste

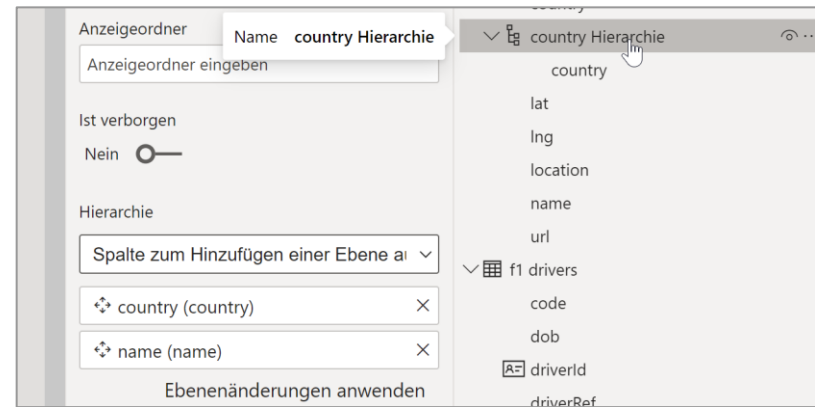
**Table Data:**

	Datum	Jahr
1	31.05.2011	2011
2	01.06.2011	2011
3	02.06.2011	2011
4	03.06.2011	2011
5	04.06.2011	2011

Beim Typ Datum wird automatisch eine Hierarchie erstellt. Für andere Felder können aber ebenfalls Hierarchien erstellt werden, z.B. Kategorie, Unterkategorie, ..., Produkt

Rechtsklick auf Feld -> Hierarchie erstellen. Leider lassen sich nicht mehr Spalten hineinziehen. Im Datenmodell können weitere Ebenen hinzugefügt werden. *Ebenenänderungen anwenden* nicht vergessen.

In einigen Visuals kann dann ein Drilldown durchgeführt werden.



Erstellen von mehreren Spalten für eine einfachere Navigation.

- Neue Spalte mit Pfad erstellen

Path = PATH(Employee[Employee ID], Employee[Manager ID])

Employee ID	Manager ID	Employee	Manager	Pfad
1010		Roy F		1010
1011	1010	Pam H	Roy F	1010 1011
1012	1010	Guy L	Roy F	1010 1012
1013	1011	Roger M	Pam H	1010 1011 1013
1014	1011	Kaylie S	Pam H	1010 1011 1014
1015	1012	Mike O	Guy L	1010 1012 1015
1016	1012	Rudy Q	Guy L	1010 1012 1016

- Hierarchie-Level erstellen

Level 1 = PATHITEM(Employee[Path],1)

Level 2 = PATHITEM(Employee[Path],2)

Level 3 = PATHITEM(Employee[Path],3)

Employee ID	Name	Manager	Manager ID	Path	Level 1	Level 2	Level 3	Level 4
1000	Quincy Howard		1000	1000	1000			
1001	Mallory Yang	Quincy Howard	1000	1000 1001	1000	1001		
1002	Donovan Maynard	Quincy Howard	1000	1000 1002	1000	1002		
1003	Giselle McClain	Mallory Yang	1001	1000 1001 1003	1000	1001	1003	
1004	Melvin Marsh	Mallory Yang	1001	1000 1001 1004	1000	1001	1004	
1005	Ria Snow	Giselle McClain	1003	1000 1001 1003 1005	1000	1001	1003	1005
1006	Callie Savage	Giselle McClain	1003	1000 1001 1003 1006	1000	1001	1003	1006

- Jetzt neue Hierarchie mit den entsprechenden Levels anlegen

Die Granularität der Daten beschreibt, wie detailliert Ihre Daten dargestellt werden. D.h., je mehr Granularität Ihre Daten besitzen, desto detaillierter sind sie. Die Wahl der richtigen Datengranularität hat große Auswirkungen auf die Leistung und Nutzbarkeit von Power BI.

- Je gröber die Daten sind (z.B. schon auf Wochen-/Tagesebene aggregiert), desto schneller und ressourcenschonender ist PowerBI
- Je feiner die Daten sind, desto tiefer kann ein Drilldown durchgeführt werden

## Verknüpfung von verschiedenen Auflösungen mittels Datumstabelle und geeigneten Spalten

- Tabelle Budget ist auf Jahr-Monats-Basis
- Tabelle Sales ist auf Datumsbasis
- Hinzufügen einer Spalte JahrMonat in Budget und Date, um eine Beziehung zwischen den beiden Tabellen anzulegen (beim Einlesen/PowerQuery Editor oder neue Spalte hinzufügen)

