

Ignaz A. SCHELS

2. Auflage

M

Daten abfragen und
verarbeiten mit
Excel und Power BI



Beispieldateien im Internet

HANSER

Schels

M



Bleiben Sie auf dem Laufenden!

Unser **Computerbuch-Newsletter** informiert Sie monatlich über neue Bücher und Termine. Profitieren Sie auch von Gewinnspielen und exklusiven Leseproben. Gleich anmelden unter:

www.hanser-fachbuch.de/newsletter



Ignaz A. Schels

M

Daten abfragen und verarbeiten
mit Excel und Power BI

2., überarbeitete Auflage

HANSER

Der Autor:

Ignaz A. Schels, München

Alle in diesem Werk enthaltenen Informationen, Verfahren und Darstellungen wurden nach bestem Wissen zusammengestellt und mit Sorgfalt getestet. Dennoch sind Fehler nicht ganz auszuschließen. Aus diesem Grund sind die im vorliegenden Werk enthaltenen Informationen mit keiner Verpflichtung oder Garantie irgendeiner Art verbunden. Autor und Verlag übernehmen infolgedessen keine juristische Verantwortung und werden keine daraus folgende oder sonstige Haftung übernehmen, die auf irgendeine Art aus der Benutzung dieser Informationen – oder Teilen davon – entsteht.

Ebenso wenig übernehmen Autor und Verlag keine Gewähr dafür, dass beschriebene Verfahren usw. frei von Schutzrechten Dritter sind. Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt also auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Buches, oder Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form (Fotokopie, Mikrofilm oder ein anderes Verfahren), auch nicht für Zwecke der Unterrichtsgestaltung – mit Ausnahme der in den §§ 53, 54 URG genannten Sonderfälle – reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

© 2023 Carl Hanser Verlag München, <http://www.hanser-fachbuch.de>

Lektorat: Brigitte Bauer-Schiewek

Copy editing: Petra Kienle, Fürstfeldbruck

Umschlagdesign: Marc Müller-Bremer, München, www.rebranding.de

Umschlagrealisation: Max Kostopoulos

Satz: Eberl & Koesel Studio, Kempten

Druck und Bindung: Hubert & Co. GmbH & Co. KG BuchPartner, Göttingen

Printed in Germany

Print-ISBN: 978-3-446-47760-5

E-Book-ISBN: 978-3-446-47811-4

E-Pub-ISBN: 978-3-446-47944-9

Inhalt

Vorwort	XI
1 Grundlagen der Abfrage-Logik	1
2 Abrufen: Verbindungen zu externen Quellen herstellen	3
2.1 Einfache Abfrage auf Excel-Datei	5
2.2 Aktualisieren der Abfrage	8
2.3 Abfragen automatisch aktualisieren	9
2.4 Kann man Abfragen separat speichern?	11
3 Transformieren: Aufbereiten der abgerufenen Daten	13
3.1 Der Power Query-Editor	13
3.2 Transformationsschritte löschen	15
3.3 Zeilen filtern	16
3.4 Überschriften (Header) einfügen	17
3.5 Spalten entfernen	17
3.6 Werte ersetzen	18
3.7 Datentypen bestimmen	18
3.8 Transformationsschritte prüfen und benennen	22
3.9 Abschließen der Transformation	23
3.10 Erneutes Bearbeiten der Abfrage	23
4 Erweiterte Transformationen	25
4.1 Nach unten bzw. oben ausfüllen	27
4.2 Zeilen und Spalten vertauschen (Transponieren)	28
4.3 Spalten verbinden	29
4.4 Entpivotieren	30
4.5 Spalten teilen und Abschluss der Transformation	31

5	Anpassen der Rohdaten	33
5.1	Texttransformationen	33
5.2	Datumstransformationen	36
5.3	Zahlentransformationen	37
5.4	Gruppieren	38
6	Abfragen auf Webseiten	40
6.1	Webabfrage auf eine Tabelle	40
6.2	Webabfrage ohne Tabellen	42
6.2.1	Das Element <i>Document</i>	43
6.2.2	Untersuchen des HTML-Codes	46
6.2.3	Den Text der Homepage direkt auswerten	47
7	Abfragen auf relationale Datenbanken	48
7.1	Beispiel mit Oracle-Datenbank	48
7.2	Query Folding	54
8	Mehrere Quellen kombinieren	58
8.1	Abfragen anfügen	58
8.2	Abfragen zusammenführen	61
8.3	Alle Dateien in einem Ordner einlesen	67
9	Spalten hinzufügen – erste Schritte mit M	72
9.1	Bedingte Spalten	72
9.2	Datums- und Rechenfunktionen	76
9.3	Benutzerdefinierte Spalten	76
9.3.1	Spaltenformeln für Berechnungen eingeben	77
9.3.2	Bedingungslogik mit if	78
9.3.3	M-Funktionen	80
10	Abfragecode bearbeiten mit M	82
10.1	Die Bearbeitungsleiste	82
10.2	Das Editor-Fenster	83
10.3	Erstellen einer leeren Abfrage	84
10.4	Die Grundstruktur des Abfrage-Codes	85
10.5	Schritte und Schrittnamen verstehen	86
10.6	Fehler im M-Code vermeiden	88
10.7	Schritte zusammenfassen	89
10.8	Kommentare	90
11	Werte und Datentypen	92
11.1	Einfache Datentypen	92
11.1.1	Null	92
11.1.2	Logical (true/false)	92

11.1.3	Number (Zahl)	94
11.1.4	Time (Zeit)	95
11.1.5	Date (Datum)	95
11.1.6	Datetime (Datum/Uhrzeit)	96
11.1.7	Datetimezone (Datum/Uhrzeit/Zeitzone)	96
11.1.8	Duration (Dauer)	97
11.1.9	Text	98
11.1.10	Binary (Binär)	101
11.2	Übergeordnete Datentypen	101
11.2.1	List (Liste)	101
11.2.2	Record (Datensatz)	103
11.2.3	Table (Tabelle)	104
11.3	Spezielle Datentypen	107
11.3.1	Type (Datentyp)	107
11.3.2	Function (Funktion)	109
12	Abfragen flexibel gestalten	111
12.1	Nicht-lineare Abfragen	111
12.2	Unterschiedliche Tabellenzeilen vergleichen	117
12.2.1	Zeilensprung über den Index	118
12.2.2	Schnellerer Zeilensprung mit Index und Join	120
12.3	Abfragen mit Parametern	122
12.3.1	Power-Query-Parameter	122
12.3.2	Parameter aus anderen Quellen	125
13	Fehlerbehandlung	129
13.1	Fehler vorbeugen	129
13.2	Error-Werte entfernen	130
13.3	Fehler abfangen mit try	132
13.4	Fehlermeldungen erzeugen	133
14	M-Funktionen	137
14.1	Das Syntax-Schema	137
14.2	Abruf der Funktionsliste	138
14.3	Listenfunktionen	142
14.3.1	Listenerstellung	142
14.3.2	Informationen über Listenelemente	145
14.3.3	Berechnungen	149
14.3.4	Ordnung und Reihenfolge	151
14.3.5	Auswahl	154
14.3.6	Transformationen	160
14.3.7	Vergleiche mehrerer Listen	167
14.4	Datensatzfunktionen	168
14.4.1	Erstellung und Umwandlung	168
14.4.2	Informationen über Datensätze	170

14.4.3	Auswahl	171
14.4.4	Transformationen	173
14.5	Tabellenfunktionen	176
14.5.1	Informationen über Tabellen	176
14.5.2	Umwandlungen	179
14.5.3	Spalten	181
14.5.4	Zeilen	188
14.5.5	Ordnung und Sortierung	201
14.5.6	Erstellung	204
14.5.7	Transformationen	208
14.5.8	Arbeiten mit mehreren Tabellen	217
14.5.9	Ähnlichkeiten erkennen (Fuzzy-Funktionen)	220
14.5.10	Sonstige	223
14.6	Textfunktionen	223
14.6.1	Erstellung und Konvertierung	223
14.6.2	Informationen über Texte	227
14.6.3	Auszüge von Textteilen	229
14.6.4	Modifikationen	231
14.6.5	Transformationen	232
14.7	Zahlenfunktionen	238
14.7.1	Informationen über Zahlen	238
14.7.2	Rechenoperationen	239
14.7.3	Rundung	243
14.7.4	Zufallszahlen	245
14.7.5	Trigonometrie	246
14.7.6	Konvertierung und Formatierung	248
14.8	Logical-Funktionen	254
14.9	Datumsfunktionen	255
14.9.1	Erzeugung und Umwandlung	255
14.9.2	Informationen über Datumswerte	257
14.9.3	Berechnungen	263
14.9.4	Vergleiche mit 'Jetzt'	264
14.10	DateTime-Funktionen	268
14.10.1	Erzeugung und Umwandlung	268
14.10.2	Vergleiche mit 'Jetzt'	271
14.11	DateTimeZone-Funktionen	273
14.12	Duration-Funktionen	278
14.13	Zeitfunktionen	281
14.14	Hilfsfunktionen	283
14.14.1	Comparer-Funktionen	283
14.14.2	Replacer-Funktionen	285
14.14.3	Combiner-Funktionen	286
14.14.4	Splitter-Funktionen	289

15	Eigene Funktionen erstellen	295
15.1	Funktionen innerhalb einer Abfrage	295
15.2	Das Schlüsselwort <i>each</i>	296
15.3	Funktionen als eigene Abfrage	297
15.4	Funktionen, die aus mehreren Schritten bestehen	301
15.5	Rekursive Funktionen	305
15.6	Funktionsbeschreibungen in der Metadata	306
15.7	Funktionen teilen	309
	15.7.1 Aktivierung und Vorbereitung	309
	15.7.2 Funktionserweiterungen	310
16	Die KI macht's möglich	313
16.1	Spaltenformeln aus Beispielen erstellen	313
16.2	KI-Funktionen (in Power BI mit Premium-Lizenz)	316
	16.2.1 Textanalyse mit KI-Insights	317
	16.2.2 Bildanalyse mit KI-Insights	319
17	M und VBA	321
17.1	Abfragen per VBA-Befehl aktualisieren	321
	17.1.1 Schaltflächen zum Aktualisieren nutzen	321
	17.1.2 Aktualisieren bei Ereignissen	323
17.2	Abfrage-Code per VBA anpassen	325
17.3	Abfrageergebnisse mit VBA auslesen	328
18	Excel-Datentypen aus Power Query	331
18.1	Eigenen Datentyp per Abfrage erzeugen	332
18.2	Datentypen aus Power BI-Datasets für die ganze Organisation	335
19	Datenschutz und Firewall	338
19.1	Datenschutz bei verschiedenen Datenquellen	338
19.2	Der Formula.Firewall-Fehler	339
20	Tipps für mehr Geschwindigkeit	342
20.1	Schnelleres Entwerfen der Abfrage	342
20.2	Schnelleres Aktualisieren	343
20.3	Messen der Geschwindigkeit	345
	20.3.1 Power BI: Nutzen der Diagnose-Werkzeuge	346
	20.3.2 Excel: Messen der Zeit über VBA	347
	20.3.3 Auswertung der Trace-Datei	347
	Index	350

Vorwort

Es mag ungewöhnlich erscheinen, ein Fachbuch zu schreiben, das nur einen Teilbereich von zwei so vielseitigen Programmen wie Excel und Power BI Desktop behandelt. Warum nicht stattdessen ein Buch, das entweder Excel oder Power BI in allen Facetten erklärt?

Wenn Sie dieses Buch gekauft haben, ahnen Sie zumindest, welches Potenzial in den Features schlummert, die unter dem Namen *Power Query* zusammengefasst werden. Excel-Nutzer verbringen unzählige Stunden damit, Daten aus den verschiedensten Quellen zusammenzukopieren und dabei händisch zu modellieren. Etliche Power-BI-Nutzer wissen zwar, wie Daten abgerufen werden, stoßen aber schnell an ihre Grenzen, wenn die Datenbasis nicht richtig aufbereitet ist. Wenn Sie nicht zu diesen Benutzern gehören wollen, haben Sie das richtige Buch gefunden.

Das Thema des vorliegenden Buchs

Was als Add-In für Excel begann, ist nun fester Teil von einer wachsenden Menge an Microsoft-Produkten, allen voran Excel und Power BI Desktop. Die Funktionsweise der Abfragen und die Benutzeroberfläche des Abfrage-Editors sind dabei bis auf wenige Ausnahmen programmübergreifend gleich. Deshalb ist es durchaus möglich, Power Query sowohl für Excel als auch Power BI Desktop zu erklären.

Aus genau diesem Grund findet aber auch eine starke Beschränkung der abgedeckten Themen statt: Das Buch behandelt nur das Abrufen von externen Daten und die Bearbeitung (Transformation) im Power Query-Editor, basierend auf der Abfragesprache M. Sie werden lernen, wie Sie Daten in Ihr Programm laden und dabei genau in die Form bringen, in der Sie sie benötigen. Die so aufbereiteten Tabellen können dann in Excel und Power BI nach Belieben interpretiert und visualisiert werden – doch das ist nicht mehr Thema des Buchs.

Unterschiede zwischen den Versionen

Die Grundlagen der Abfrage-Logik und der Formelsprache M haben sich seit den ersten Versionen von Power Query kaum verändert. Die Benutzeroberfläche und die Entwicklungsumgebung des Abfrage-Editors werden jedoch laufend optimiert.

Die regelmäßige Verbesserung eines Programms ist natürlich zu begrüßen, doch erschwert sie das Lernen und Lehren der Grundlagen. Es ist nicht auszuschließen, dass das eine oder andere Beispiel auf Ihrem Computer nicht genauso funktioniert wie hier beschrieben – sei

es, dass Sie eine ältere Version benutzen oder sei es, dass nach der Buchveröffentlichung weitere Neuerungen eingeführt wurden.

Seien Sie aber bitte nicht frustriert, wenn eine Beschreibung oder eine Abbildung nicht exakt zutrifft. In den meisten Fällen ist das entsprechende Werkzeug nur anders benannt oder befindet sich an einer anderen Stelle. Spätestens wenn es um das Schreiben von M-Code geht, sind Versionsunterschiede ohnehin kaum mehr von Belang.

Kontakt zum Autor

Auch bei einem so starren Medium wie einem Buch darf der Kontakt zwischen Leser und Autor nicht zu kurz kommen. Scheuen Sie sich nicht, mir zu schreiben, wenn Sie Fragen zu einem der behandelten Themen haben. Ein Software-Spezialist lebt von solchen Herausforderungen. Feedback jeder Art, egal ob Lob oder Kritik, ist natürlich auch immer willkommen. Schicken Sie einfach eine Mail an *ignaz.alexander@schels.de*.

Der Verlag und die Autoren haben sich mit der Problematik einer gendergerechten Sprache intensiv beschäftigt. Um eine optimale Lesbarkeit und Verständlichkeit sicherzustellen, wird in diesem Werk auf Gendersternchen und sonstige Varianten verzichtet; diese Entscheidung basiert auf der Empfehlung des Rates für deutsche Rechtschreibung. Grundsätzlich respektieren der Verlag und die Autoren alle Menschen unabhängig von Ihrem Geschlecht, ihrer Sexualität, ihrer Hautfarbe, ihrer Herkunft und ihrer nationalen Zugehörigkeit.

1

Grundlagen der Abfrage-Logik

Der ursprüngliche Name der hier behandelten Technologie, *Power Query*, wird in aktuellen Microsoft-Produkten fast nicht mehr verwendet. Stattdessen sieht man je nach Programm und Version meist verschiedene Varianten der Begriffsgruppe „Daten abrufen und transformieren“ (englisch „Get & Transform Data“). Diese Bezeichnung ist zwar etwas sperriger, beschreibt aber treffend die zwei Schritte, in die sich jede Abfrage aufteilen lässt:

1. *Abrufen*: Eine externe Datenquelle, z. B. eine Excel-Arbeitsmappe, eine Textdatei oder eine Datenbank, wird „angezapft“ und die enthaltenen Informationen werden eingelesen.
2. *Transformieren*: Wenn die Daten abgerufen sind, werden sie in die gewünschte Form gebracht. Dabei können zum Beispiel die benötigten Zeilen und Spalten ausgewählt, Formate angepasst oder Sortierungen vorgenommen werden. Auch Ersetzungen oder Berechnungen sind in diesem Schritt schon möglich.

Das Ergebnis einer Power-Query-Abfrage ist im Normalfall eine Tabelle. Diese kann wiederum als Basis für weitere Auswertungen oder Visualisierungen verwendet werden.

Wenn Sie bereits etwas Erfahrung mit dem Erstellen von Diagrammen, Pivot-Tabellen oder anderen Visualisierungen in Excel und Power BI haben, wissen Sie, dass eine saubere Datenbasis dabei unerlässlich ist. Damit eine Visualisierung das zeigt, was sie soll, müssen die Ausgangsdaten in der entsprechenden Form sein. Schon eine Zwischenüberschrift oder ein Zeilenumbruch kann alles durcheinanderbringen.

Die Möglichkeit, die eingelesenen Daten direkt zu transformieren, kann also ein wahrer Segen sein: Egal in welcher Form die Quelldaten vorliegen, Sie können in der Abfrage definieren, welche Daten Sie benötigen und wie sie angepasst werden sollen. Wenn sich die Quelldaten ändern, liest die Abfrage die neuen Daten ein und führt damit die festgelegten Transformierungsschritte durch.



Die Quelldaten werden durch eine Abfrage nicht verändert.

Falls Sie mehrere Datenquellen gemeinsam auswerten wollen, ist das auch kein Problem: Abfragen können miteinander verbunden werden. Das Ergebnis einer Abfrage kann auch die Basis einer weiteren Abfrage sein. Ihre Quelldaten müssen nicht einmal immer am gleichen Ort liegen: Sie können beispielsweise bestimmte Dateien innerhalb eines Ordners auswählen oder Parameter für die Quelladresse festlegen.

In den folgenden Kapiteln erhalten Sie einen Einblick in die vielen Möglichkeiten, die eine Abfrage mit Power Query bietet. Die meisten Techniken werden anhand von Beispielen gezeigt, die Sie Schritt für Schritt nachvollziehen können. Die zugehörigen Beispieldateien finden Sie zum Download unter dieser Adresse:

www.schels.de/publikationen

2

Abrufen: Verbindungen zu externen Quellen herstellen

Um zu sehen, welche Vielzahl an Datenquellen für die Abfrage zur Verfügung stehen, verschaffen Sie sich am besten selbst einen Überblick.

EXCEL:

Um eine neue Abfrage in Excel (ab Version 2016) zu erstellen, wählen Sie das Register **DATEN** und klicken Sie unter der Rubrik *Daten abrufen und Transformieren* auf **DATEN ABRUFEN**. Daraufhin können Sie Ihre Datenquelle auswählen.



Bei Excel-Versionen vor Januar 2018 ist die Schaltfläche mit **NEUE ABFRAGE** beschriftet.

Bei den Excel-Versionen vor Excel 2016 sind die Abfrage-Tools noch nicht integriert. Nutzer von Excel 2010 und 2013 müssen aber nicht das Nachsehen haben: Die Werkzeug-Rubrik *Abrufen und Transformieren* ist lediglich die Weiterentwicklung eines Add-Ins, das schon ab Excel 2010 verfügbar ist. Das Add-In *Power Query* basiert auf den gleichen grundlegenden Methoden wie die späteren, integrierten Werkzeuge. Es gibt jedoch Unterschiede, etwa bei den Bezeichnungen und Positionen einzelner Features. Mit leichten Abwandlungen können aber fast alle Beispiele in diesem Buch auch von Nutzern der älteren Excel-Versionen nachvollzogen werden.

Den Download-Link für das Add-In und Hinweise zur Installation gibt es auf der Seite von Microsoft:

<https://www.microsoft.com/en-us/download/details.aspx?id=39379>

POWER BI:

In Power BI Desktop haben Sie zwei Möglichkeiten, eine Abfrage zu erstellen: Gleich bei Programmstart erscheint ein kleines Begrüßungsfenster, in dem Sie auf der linken Seite **DATEN ABRUFEN** anklicken können. Alternativ finden Sie auch im Programmfenster auf dem Register **START** eine Schaltfläche **DATEN ABRUFEN**. In beiden Fällen gelangen Sie zu einem Fenster, in dem Sie Ihre Datenquelle auswählen können.

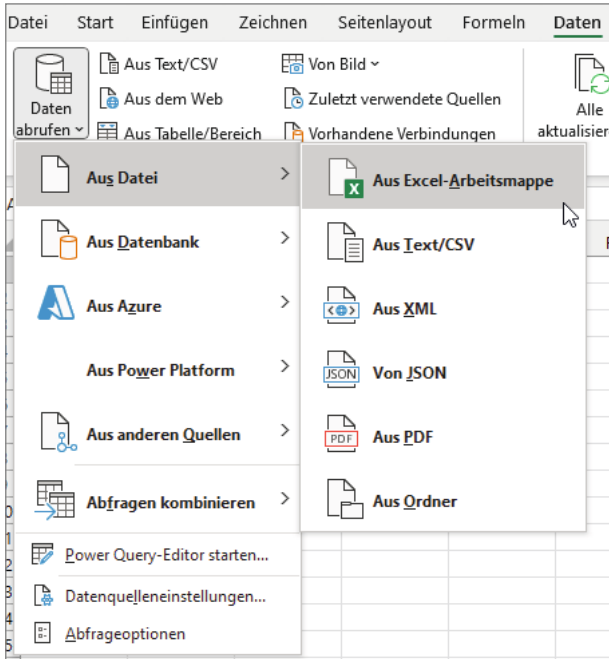


Bild 2.1 Erstellen einer Abfrage in Excel.

Die Auswahl der Datenquellen ist beträchtlich: Neben einfachen Excel- und Textdateien können Sie auf die meisten gängigen Datenbanksysteme zugreifen. Microsofts Sharepoint und Azure werden natürlich ebenfalls unterstützt. Sie können sogar Daten von Exchange-Servern oder Webseiten auslesen.

Das Vorgehen ist für die verschiedenen Datenquellen ähnlich und meist selbsterklärend: Zunächst wählen Sie den Speicherort oder die Adresse Ihrer Quelle aus. Je nach Art der Verbindung können zusätzlich noch weitere Informationen wie z.B. Zugangsdaten oder Teilbereiche der jeweiligen Datei oder Datenbank ausgewählt werden. Zum Abschluss haben Sie die Möglichkeit, die Daten unverändert einzulesen oder vorher im Abfrage-Editor zu bearbeiten. Der folgende Abschnitt zeigt eine einfache Abfrage ohne Transformation am Beispiel einer Excel-Datei.

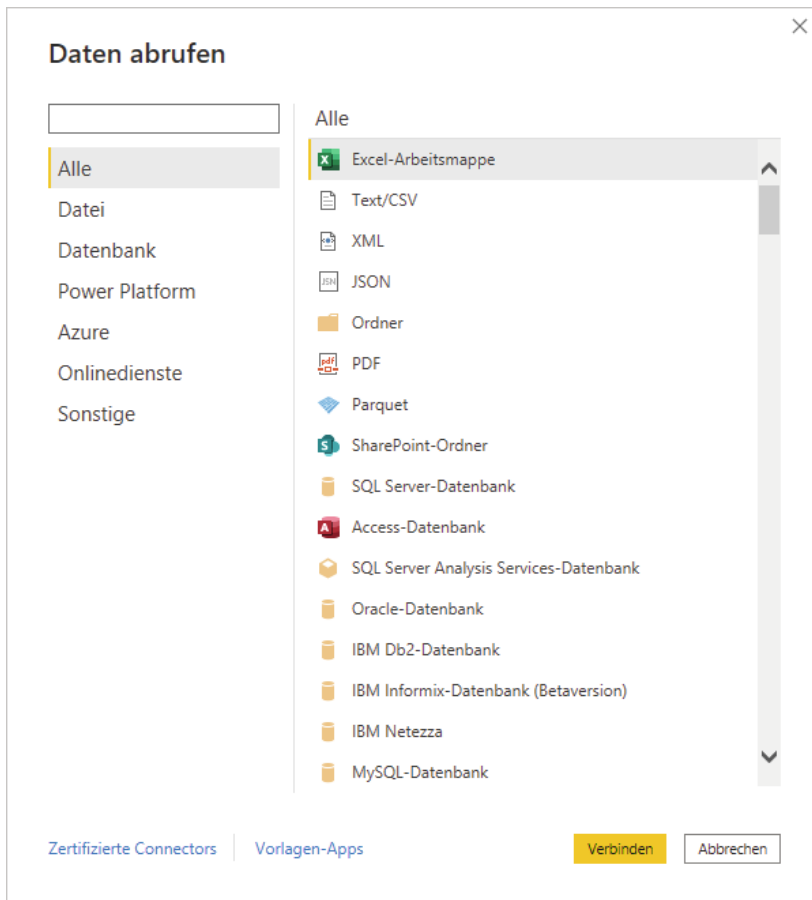


Bild 2.2 Auswahl der Datenquelle in Power BI Desktop.

■ 2.1 Einfache Abfrage auf Excel-Datei

Das erste Beispiel können Sie, wie fast alle Beispiele dieses Buchs, gleichermaßen mit Excel und Power BI Desktop nachvollziehen. Die Arbeitsschritte sind meistens für beide Programme die gleichen. Falls es Unterschiede gibt, ist dies im Text hervorgehoben.



Die Übungsdateien zu diesem Buch können Sie unter folgender Adresse herunterladen:

www.schels.de/publikationen

- **EXCEL:** Öffnen Sie eine neue Arbeitsmappe, wählen Sie das Register **DATEN** und klicken Sie unter der Rubrik *Daten abrufen und transformieren* auf **DATEN ABRUFEN**. Wählen Sie im Menü **AUS DATEI – AUS EXCEL-ARBEITSMAPPE**.
- **POWER BI:** Klicken Sie auf **DATEN ABRUFEN** und wählen Sie die Quelle **EXCEL**.

Sie haben nun die Möglichkeit, eine Excel-Arbeitsmappe auszuwählen. Wählen Sie aus den Übungsdateien zum Buch die Datei *2-01-Umsatz.xlsx*. Daraufhin erscheint das Navigator-Fenster, in dem Sie wählen können, welches Arbeitsblatt aus der Excel-Arbeitsmappe eingelesen werden soll. Wählen Sie *Umsatz 2015*.



Das Navigator-Fenster zeigt Ihnen bei Excel-Quelldateien nicht nur Arbeitsblätter, sondern, wenn verfügbar, auch Tabellen und benannte Bereiche zur Auswahl an.

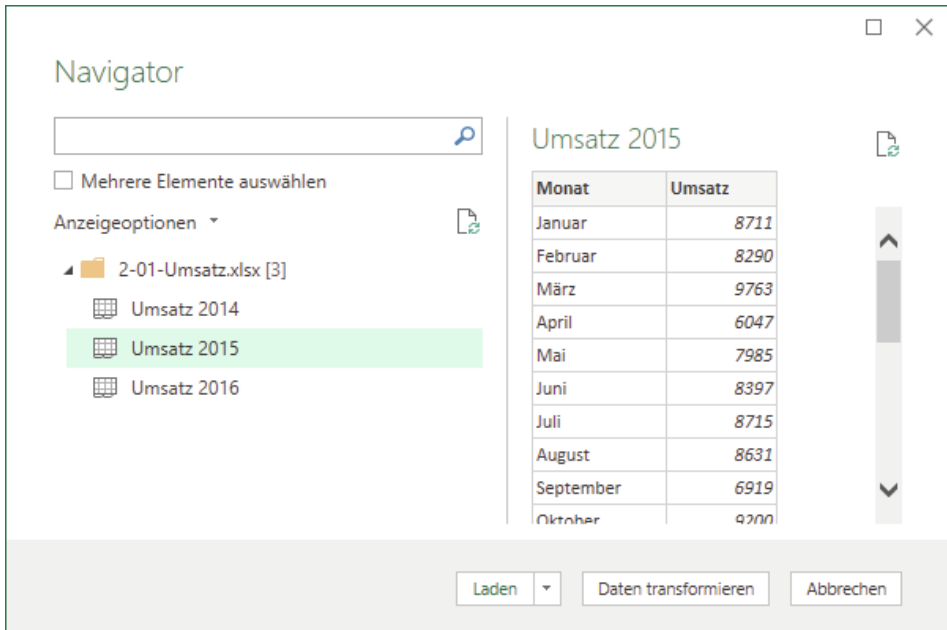


Bild 2.3 Im Navigator-Fenster können Sie das gewünschte Tabellenblatt auswählen.

In der rechten Fensterhälfte sehen Sie eine Vorschau der Tabelle, die importiert wird. Sie können nun bereits auf **LADEN** klicken, um die Abfrage abzuschließen.

EXCEL:

In Excel haben Sie mehrere Möglichkeiten, das Abfrageergebnis zu laden. Wenn Sie direkt auf die Schaltfläche **LADEN** klicken, wird eine Tabelle mit den abgerufenen Daten auf einem neuen Arbeitsblatt angelegt. Um weitere Optionen zu erhalten, klicken Sie auf den kleinen Pfeil neben der Schaltfläche und wählen Sie **LADEN IN ...**

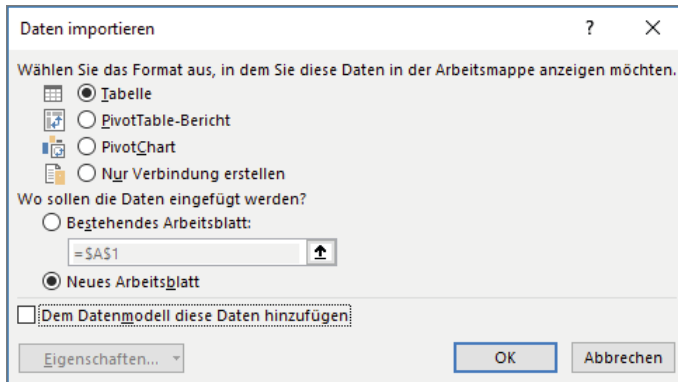


Bild 2.4 Über den Dialog *Laden in* wird der Zielort bestimmt.

Daraufhin erscheint ein neuer Dialog, in dem Sie mehrere Auswahlmöglichkeiten haben:

- *Tabelle*: Mit der Standard-Option werden die importierten Daten in einer Tabelle gespeichert. Weiter unten können Sie genau bestimmen, wo die neue Tabelle angelegt werden soll.
- *PivotTable-Bericht*: Die abgerufenen Daten bilden die Quelle einer Pivot-Tabelle.
- *PivotChart*: Die abgerufenen Daten dienen als Datenbasis eines Pivot-Charts.
- *Nur Verbindung erstellen*: Wenn Sie diese Option wählen, werden die importierten Daten nicht in einem Arbeitsblatt angezeigt. Dies kann nützlich sein, wenn Sie mehrere Abfragen kombinieren wollen (vgl. Kapitel 8) oder wenn Sie lediglich Ihr Datenmodell befüllen wollen (siehe nächster Punkt).
- *Dem Datenmodell diese Daten hinzufügen*: Falls Sie die importierten Daten in einer Pivot-Tabelle oder einem PivotChart verwenden wollen, aktivieren Sie diese Option. Die Daten stehen dann direkt zur Verfügung, ohne dass Sie die Quell-Tabelle angeben müssen. (Diese Option ist nicht in allen Excel-Versionen vorhanden.)

POWER BI:

In Power BI Desktop müssen Sie keinen Zielort für die abgefragten Daten angeben. Sobald Sie auf **LADEN** klicken, werden sie zu Ihrem Datenmodell hinzugefügt. Anschließend können Sie sie mit anderen Tabellen in Verbindung setzen oder direkt Visualisierungen daraus erstellen.

Um den Bericht schließlich in der Cloud auf Power BI Service verfügbar zu machen, klicken Sie im Register **START** auf **VERÖFFENTLICHEN**. Daraufhin können Sie online auf die Visualisierungen zugreifen oder den Bericht mit anderen Nutzern teilen.



Analog zur Option *Nur Verbindung erstellen* in Excel, haben Sie auch in Power BI Desktop die Möglichkeit, Abfragen so zu erstellen, dass sie nur für andere Abfragen verwendet werden können.

Auf der linken Seite des Abfrage-Editors befindet sich eine Liste mit allen Abfragen einer Datei. Klicken Sie mit der rechten Maustaste auf eine Abfrage und deaktivieren Sie die Option *Laden aktivieren*. Das bewirkt, dass die Tabelle nicht ins Datenmodell geladen wird.

■ 2.2 Aktualisieren der Abfrage

EXCEL:

Wenn Sie den Abfrage-Dialog ganz normal mit **LADEN** geschlossen haben, erhalten Sie eine einfache Tabelle mit den eingelesenen Daten. Das ist zunächst nicht spektakulär, denn dafür hätten Sie die Daten auch einfach per Copy & Paste in Ihre Arbeitsmappe kopieren können.

Doch nehmen Sie einmal an, die Quelldatei wird regelmäßig geändert. In vielen Firmen gibt es Excel-Listen mit Umsatzzahlen oder Personaldaten, die laufend angepasst werden. Mehrere Personen greifen auf dieselbe Datei zu, um jeweils unterschiedliche Auswertungen zu erstellen. Wer nicht in der Originaldatei arbeiten will oder darf, muss daher immer wieder die Originaldatei öffnen, die aktuellen Daten herauskopieren und in sein Auswertungsmodell übertragen.

Viel schneller geht es mit einer Abfrage, denn dabei müssen Sie nur einmal eine Verbindung zur Quelldatei herstellen. Anschließend können Sie sie jederzeit aktualisieren, um die neuesten Daten zu erhalten. Hierfür klicken Sie im Register **DATEN** auf **ALLE AKTUALISIEREN**. Noch schneller geht es mit der Tastenkombination **STRG + ALT + F5**.

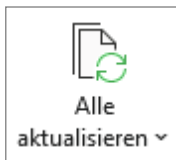


Bild 2.5

Aktualisieren Sie alle Abfragen mit einem Klick.

Falls Sie mehrere Abfragen oder Verbindungen in Ihrer Arbeitsmappe haben, möchten Sie vielleicht nur eine bestimmte Abfrage aktualisieren (z. B. um Zeit zu sparen). Auch das ist möglich: Falls am rechten Bildschirmrand nicht die Liste der Abfragen angezeigt wird, aktivieren Sie diese über **DATEN – ABFRAGEN UND VERBINDUNGEN**. Rechts neben den jeweiligen Abfragen in der Liste können Sie nun auf das Aktualisieren-Symbol klicken. Alternativ können Sie auch einen Rechtsklick auf die Abfrage machen und **AKTUALISIEREN** wählen.

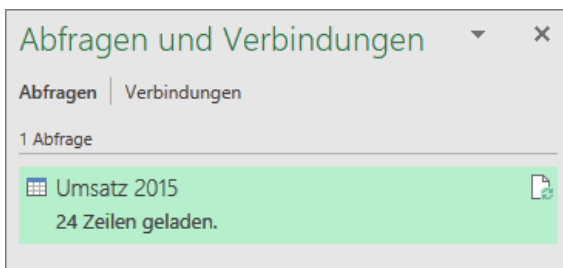


Bild 2.6

Die Liste der Abfragen.

POWER BI:

Auch bei Power BI kann es natürlich vorkommen, dass die Quelldaten sich ändern. Damit die aktuellsten Daten neu geladen werden, klicken Sie einfach auf die Schaltfläche **AKTUALISIEREN** im **START**-Register. Dadurch wird die Datenquelle erneut eingelesen.

Bei bestimmten Datenquellen haben Sie zusätzlich die Möglichkeit, eine Verbindung per *DirectQuery* herzustellen. Bild 2.7 zeigt, wie diese Option beim Erstellen einer Abfrage auf eine SQL-Server-Datenbank aussieht.

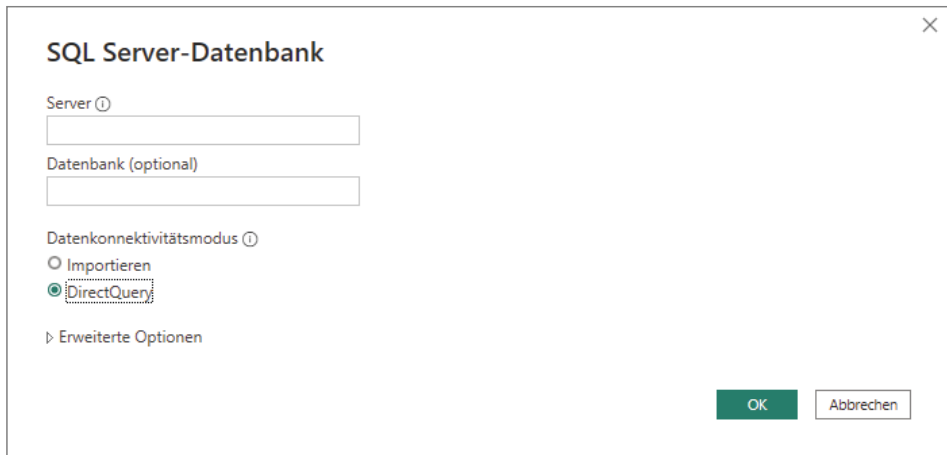


Bild 2.7 DirectQuery beim Erstellen einer SQL-Server-Abfrage.

Wenn Sie eine Abfrage mit *DirectQuery* erstellen, werden die abgefragten Daten nicht in der Power-BI-Datei gespeichert, wie es normalerweise der Fall wäre. Stattdessen wird eine direkte Verbindung zur Datenquelle hergestellt. Die benötigten Daten werden jedes Mal neu geladen, wenn Visualisierungen erstellt oder verändert werden. Wenn Sie beispielsweise durch Klick auf eine Visualisierung nach einem bestimmten Datenbereich filtern, werden diese Daten live abgerufen. Besonders bei großen Datenmengen kann diese Methode effizienter sein als eine permanente Speicherung der Daten, da immer nur der jeweils benötigte Teil gespeichert wird. Andererseits können jedoch Verzögerungen entstehen, wenn die Verbindung zum Server nicht schnell genug ist.

■ 2.3 Abfragen automatisch aktualisieren

EXCEL:

Excel bietet die Möglichkeit, Abfragen automatisch in bestimmten Zeitabständen oder beim Öffnen der Datei zu aktualisieren. Das kann von Vorteil sein, wenn sich die Quelldatei häufig ändert und Sie sicherstellen wollen, immer die neuesten Daten zu haben. Andererseits kann es aber zu störenden Verzögerungen führen. Es hängt also immer vom jeweiligen Anwendungsfall ab, ob eine automatische Aktualisierung sinnvoll ist.

Falls die Liste der Abfragen nicht sichtbar ist, öffnen Sie sie über **DATEN – ABFRAGEN UND VERBINDUNGEN**. Klicken Sie mit der rechten Maustaste auf eine Abfrage und wählen Sie **EIGENSCHAFTEN**.

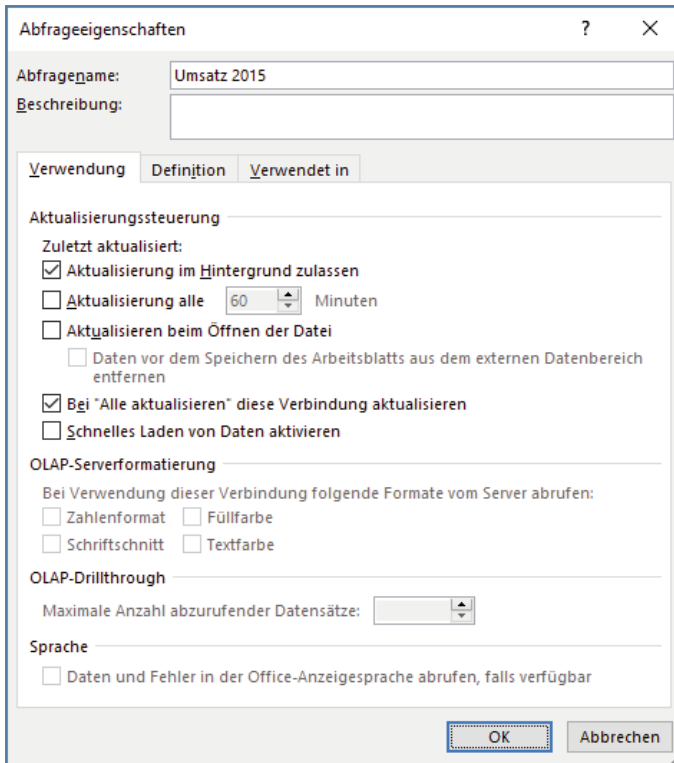


Bild 2.8 Die Abfrageeigenschaften.



Bei älteren Excel-Versionen (vor Februar 2018) finden Sie den entsprechenden Dialog unter **DATEN – VERBINDUNGEN – EIGENSCHAFTEN**.

Im Dialog *Abfrageeigenschaften* können Sie zum Beispiel einstellen, dass die Abfrage stündlich aktualisiert wird. Oder Sie können sie bei jedem Öffnen der Datei automatisch aktualisieren lassen. Wenn Sie diese Einstellung aktivieren, haben Sie auch die Möglichkeit, die Daten „vor dem Speichern des Arbeitsblatts aus dem externen Datenbereich (zu) entfernen“. Das bedeutet, dass die Tabelle mit den importierten Daten nicht mit der Arbeitsmappe gespeichert wird, sondern immer neu geladen werden muss. Dadurch reduziert sich die Dateigröße, insbesondere bei großen Tabellen. Ein weiterer Vorteil dieser Einstellung ist die Datensicherheit: Bei sensiblen Quelldaten innerhalb eines Firmennetzwerks ist die Datei mit der Abfrage unbrauchbar, solange sich der Benutzer außerhalb des Netzwerks befindet.

POWER BI:

Eine Option zum automatischen Aktualisieren gibt es in Power BI Desktop nicht. Klicken Sie also bei Bedarf auf die Schaltfläche **AKTUALISIEREN**, um die neuesten Daten einzulesen.

Etwas mehr Flexibilität bietet Power BI Service. Wenn Sie Ihre Daten dort veröffentlichen, können Sie regelmäßige Aktualisierungen einrichten. Die Voraussetzung ist natürlich, dass Ihre Datenquelle auch für Power BI Service zugänglich ist.



Befinden sich Ihre Quelldaten in einer lokalen Datei auf Ihrem Computer und Sie möchten die Abfrage auf Power BI Service mit anderen teilen, ist es das einfachste, die Quelldatei auf OneDrive hochzuladen. Auf diese Weise kann die Abfrage auch online jederzeit aktualisiert werden.

■ 2.4 Kann man Abfragen separat speichern?

Sowohl in Excel als auch in Power BI Desktop werden Abfragen zusammen mit der Datei gespeichert. Bei Power BI haben Sie zusätzlich noch die Möglichkeit, die Abfrage zusammen mit dem Datenmodell und dem Bericht in der Cloud zu veröffentlichen.

Vielleicht möchten Sie manchmal aber nur eine Abfrage speichern, unabhängig von anderen Tabellen, Diagrammen und sonstigen Daten. Eine Exportmöglichkeit oder eine Speichern-Schaltfläche im Fenster von Power Query gibt es dafür leider nicht. Aber es gibt ähnliche Alternativen.

EXCEL:

Sie können Excel-Abfragen als ODC-Dateien exportieren. ODC steht für Office Data Connection und ist ein Format, um Verbindungseigenschaften zu speichern. Es basiert auf den Formaten HTML und XML, Sie können es also mit einem beliebigen Texteditor öffnen, wenn Sie neugierig auf den gespeicherten Inhalt sind.

- Um eine Abfrage zu exportieren, öffnen Sie die Seitenleiste *Abfragen und Verbindungen* in Excel und klicken mit der rechten Maustaste auf eine Abfrage.
- Klicken Sie im Kontextmenü auf den Eintrag *Verbindungsdatei exportieren*.
- Es erscheint ein Dialogfeld, in dem Sie den Speicherort der Datei auswählen können. Der Standardordner hierfür ist der Ordner *Meine Datenquelle* in Ihrem Benutzer-Verzeichnis, aber Sie können natürlich auch einen anderen Ordner wählen.

Alternativ zu diesem Kontextmenü-Befehl, gibt es auch eine Schaltfläche **VERBINDUNGS-DATEI EXPORTIEREN** bei den Verbindungseigenschaften auf dem Register *Definition*.

Diese ODC-Datei enthält nun alle Informationen, um die Abfrage in einer anderen Excel-Datei zu verwenden. Sie können es ausprobieren, indem Sie im Windows Explorer einen Doppelklick auf die Datei machen.

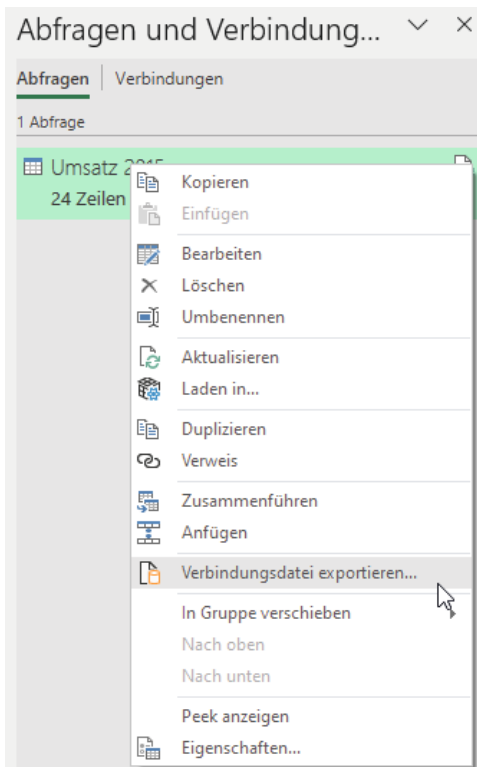


Bild 2.9 Erstellen einer Verbindungsdatei über das Kontextmenü der Abfrage.

Wenn Sie bereits eine Arbeitsmappe geöffnet haben, klicken Sie auf dem Register **DATEN** auf die Schaltfläche **VORHANDENE VERBINDUNGEN**. Die Verbindungen, die hier aufgelistet sind, stammen aus Ihrem Standard-Verbindungsordner, aber nach einem Klick auf die Schaltfläche **NACH WEITEREN ELEMENTEN SUCHEN**, können Sie eine beliebige ODC-Datei auswählen.

(*Hinweis:* In älteren Excel-Versionen finden Sie die Arbeitsmappenverbindungen im Register **DATEN** unter **VERBINDUNGEN**.)

POWER BI:

Eine Export-Funktion, die nur die Power-Query-Abfrage betrifft, gibt es in Power BI leider nicht. Eine einfache Alternative ist, den M-Code der Abfrage in einer Textdatei zu speichern und bei Bedarf im Erweiterten Editor einzufügen. In Kapitel 10 erfahren Sie mehr dazu. Eine professionellere Lösung ist das Erstellen eines eigenen Connectors. So werden die verschiedenen Datenquellenverbindungen genannt, die zur Auswahl stehen, wenn Sie auf **DATEN ABRUFEN** klicken. Es würde zu weit führen, an dieser Stelle zu erklären, wie benutzerdefinierte Connectors erstellt und zertifiziert werden, aber für fortgeschrittene Benutzer ist das durchaus möglich. In Kapitel 15 wird das Grundprinzip kurz gezeigt anhand eines Beispiels mit einer benutzerdefinierten Funktion.

3

Transformieren: Aufbereiten der abgerufenen Daten

Das Layout der Quelldaten ist nicht immer geeignet für die gewünschten Auswertungen. Oft muss erst eine Menge verschoben, gefiltert, ersetzt und gelöscht werden, bis eine brauchbare Analyse oder Visualisierung erstellt werden kann. Zudem benötigen Sie meist nur einen bestimmten Teil aus einer größeren Tabelle oder Datenbank. Je weniger unnötige Daten Sie ansammeln, umso kleiner und schneller ist Ihre Datei.

Aus diesen Gründen wird oft viel Arbeitszeit damit verbracht, Daten „in die richtige Form zu bringen“, bevor mit der eigentlichen Auswertung begonnen werden kann. Besonders frustrierend ist dies, wenn die gleichen Aktionen regelmäßig wiederholt werden müssen, z. B. wenn die gleiche Auswertung monatlich oder wöchentlich erstellt werden muss.

Der Power-Query-Abfrage-Editor kann Ihnen in solchen Fällen sehr viel Zeit sparen. Sie müssen die Daten nicht jedes Mal selbst aufbereiten, sondern nur einmal die nötigen Arbeitsschritte aufzeichnen. Die aufgezeichneten Transformationsschritte werden dann bei jeder Aktualisierung der Abfrage innerhalb weniger Sekunden ausgeführt.

■ 3.1 Der Power Query-Editor

Um beim Erstellen einer Abfrage zum Abfrage-Editor zu gelangen, klicken Sie im Navigator-Fenster nicht auf **LADEN** wie im letzten Beispiel, sondern stattdessen auf **DATEN TRANSFORMIEREN** oder **BEARBEITEN**.

Das folgende Beispiel verdeutlicht, wie der Editor funktioniert.

- Öffnen Sie Excel oder Power BI in einer neuen Datei und erstellen Sie eine Abfrage auf die Datei *3-01-Umsatz_Kwik-E-Mart.csv*. Da es sich um eine CSV-Datei handelt, wählen Sie den entsprechenden Dateityp aus.
- Es erscheint ein Fenster, in dem Sie die Text-Codierung und das Trennzeichen auswählen können. In diesem Fall wählen Sie UTF-8 und das Trennzeichen Komma. Der Text wird so in mehrere Spalten aufgeteilt. In der Vorschau sehen Sie direkt, ob Sie richtig gewählt haben.



Bei CSV-Dateien im deutschen Sprachraum wird üblicherweise das Komma als Dezimaltrennzeichen und das Semikolon als Markierung der Spalten-
grenzen verwendet. Bei der Beispieldatei handelt es sich um eine CSV-Datei,
wie sie international üblich ist: Das Komma trennt die Spalten und der Punkt
markiert die Dezimalstellen bei Zahlen.

Solche Dateien können in Excel bei deutscher Regionseinstellung nicht
richtig angezeigt werden – was in internationalen Unternehmen schon viele
Nerven gekostet hat. Eine Abfrage schafft hier Abhilfe.

Klicken Sie unten im Fenster auf die Schaltfläche **DATEN TRANSFORMIEREN** bzw. **BEARBEITEN**
(je nach Version).

Daraufhin öffnet sich der Power Query-Editor. Obwohl er auf den ersten Blick ähnlich wie
ein Excel-Fenster aussieht, ist sein Zweck ein ganz anderer. Sie können beispielsweise nicht
den Wert einzelner Zellen verändern. Auch Formatierungen oder Rahmen sind hier nicht
möglich.

Abfrage	Column1	Column2	Column3	Column4	Column5
1	Department:	8			
2	Branch:	12			
3	Revenue Nov. 2017:	115094.70			
4	Running Total 2017:	1221964.11			
5	Sales Mgr.:	Apu Nahasapeemepetton			
6					
7					
8	Item No	Amount	Refill	Price	Revenue
9	112-p04	10	no	7.72	77.20
10	124-p02	71	no	5.64	400.44
11	128-p00	6	yes	9.35	56.10
12	129-p01	40	no	7.96	318.40
13	130-p07	128	no	7.11	910.08
14	132-p03	157	no	1.29	202.53
15	134-p00	70	yes	4.38	306.60
16	134-p05	74	yes	2.31	170.94
17	135-p07	81	yes	3.24	262.44
18	137-p05	68	no	0.88	59.84
19	143-p00	149	no	1.14	169.86
20	144-p09	33	yes	5.95	196.35
21	145-p08	117	yes	8.80	1029.60
22	158-p03	24	yes	11.42	274.08
23	160-p04	187	no	1.74	325.38
24	161-p00	92	yes	3.31	304.52
25	167-p05	120	yes	6.92	830.40
26	168-p07	152	yes	5.82	884.64

Bild 3.1 Der Power Query-Editor in Excel.



Der Power Query-Editor sieht in Excel fast genauso aus wie in Power BI Desktop.

In diesem Fenster werden Transformationsschritte festgelegt, die auf die Quelldaten angewendet werden. Bei jeder Aktualisierung der Abfrage werden die Schritte wieder nacheinander abgearbeitet. Auf diese Weise können Sie auch aufwendige Aufgaben so automatisieren, dass sie künftig mit einem einzigen Klick erledigt werden.

Die Schritte werden am rechten Fensterrand unter *Angewendete Schritte* aufgelistet. Hier sind auch schon zwei Schritte angelegt:

1. *Quelle*: Das ist der erste Schritt bei fast jeder Abfrage. Er stellt die Verbindung zu den Quelldaten, also in diesem Beispiel der CSV-Datei, her.
2. *Geänderter Typ*: Dieser Schritt wird automatisch eingefügt, ist aber nicht immer sinnvoll. Power Query versucht hier, die Datentypen der einzelnen Spalten zu erkennen. In diesem Beispiel brauchen Sie diesen Schritt nicht, da wir die Datentypen später selbst festlegen werden. Sie können ihn also gleich entfernen.

■ 3.2 Transformationsschritte löschen

- Wenn Sie den letzten Schritt („Geänderter Typ“) anklicken, sehen Sie links daneben ein Kreuz. Klicken Sie darauf, um den Schritt zu entfernen.

Im Abfrage-Editor gibt es keine Rückgängig-Funktion wie in den meisten anderen Programmen. Doch jede Aktion, die Sie vornehmen, wird in einem Transformierungsschritt aufgezeichnet. Wenn Sie einen Fehler machen, können Sie also einfach den entsprechenden Schritt löschen.



Bild 3.2

Mithilfe des Kreuz-Symbols werden Schritte gelöscht.

Beachten Sie jedoch, dass jeder Schritt auch die nachfolgenden Schritte beeinflusst. Sollten bereits Arbeitsschritte nach dem entfernten Schritt existieren, kann es sein, dass diese nicht mehr wie gewünscht funktionieren oder Fehlermeldungen erzeugen. Nur beim Entfernen des letzten Schritts müssen Sie sich darüber keine Sorgen machen.

■ 3.3 Zeilen filtern

Wenn Sie sich die Vorschau-Tabelle ansehen, stellen Sie fest, dass sie eigentlich in mehrere Tabellen aufgeteilt ist. Oberhalb der Tabellenüberschriften stehen immer sieben Kopfzeilen, die zum Teil leer sind und zum Teil Zusatzinformationen enthalten, die für die Auswertung nicht relevant sind (z.B. „Department“). Diese Kopfzeilen sollten entfernt werden, damit sich am Ende nur eine zusammenhängende Tabelle ergibt.

- Die überflüssigen Zeilen haben die Gemeinsamkeit, dass sie ab Spalte drei leer sind. Nutzen Sie daher die Filterfunktion, indem Sie auf das Quadrat mit dem Filterpfeil rechts neben dem Spaltentitel *Column3* klicken.

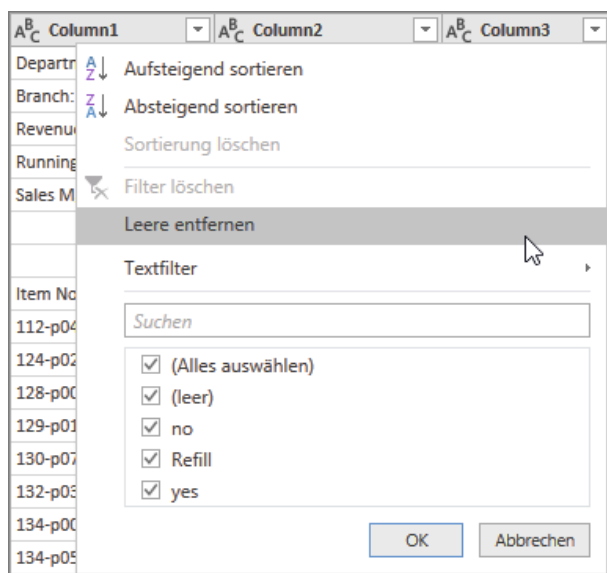


Bild 3.3 Mit der Filter-Funktion reduzieren Sie die Tabelle auf das Wesentliche.

In der Filter-Box sehen Sie eine Auswahl aus den verschiedenen Einträgen in der Spalte. Nur die angekreuzten Einträge verbleiben nach dem Filtern in der Tabelle. Standardmäßig werden bis zu 1000 verschiedene Einträge angezeigt. Sollte es mehr als 1000 geben, erscheint eine Schaltfläche **MEHR ANZEIGEN**. Mit dieser Schaltfläche können dann alle Einträge angezeigt werden (was jedoch je nach Tabellengröße eine Weile dauert).

Zusätzlich zu den Einträgen, die per Häkchen ausgewählt werden können, gibt es je nach Typ der Spalte noch spezielle Filter. Mit den Textfiltern können Sie beispielsweise festlegen, dass alle Einträge herausgefiltert werden, die eine bestimmte Zeichenfolge enthalten. Analog gibt es Zahlenfilter (z.B. alle Zahlen größer als 5) oder Datumsfilter (z.B. nur dieses Jahr), falls die Spalte den entsprechenden Datentyp hat.

- Klicken Sie auf **LEERE ENTFERNEN**. Alternativ können Sie auch das Häkchen neben dem Eintrag (*leer*) wegklicken und auf **OK** klicken.

■ 3.4 Überschriften (Header) einfügen

In den meisten Fällen werden Sie Überschriften für die Spalten Ihrer Tabellen benötigen. Bisher haben wir nur die generischen Standard-Überschriften *Column1*, *Column2* usw. Die eigentlichen Überschriften liegen in der ersten Zeile.

- Klicken Sie im Register **START** im Bereich *Transformieren* auf die Schaltfläche **ERSTE ZEILE ALS ÜBERSCHRIFTEN VERWENDEN**.
- Die erste Zeile ist nun in den Überschriften-Bereich gerutscht. Jedoch hatte anfangs jede der einzelnen Tabellen eine Überschriften-Zeile. Die Zeilen mit Zwischenüberschriften können nun auch entfernt werden. Nutzen Sie hierzu wieder die Filterfunktion in der dritten Spalte, entfernen Sie das Häkchen neben dem Wort *Refill* und klicken Sie auf **OK**.

Die Tabelle sollte nun nur noch die reinen Umsatzdaten enthalten. Sie können Spaltenüberschriften übrigens auch direkt ändern:

- Führen Sie einen Doppelklick auf die Spaltenüberschrift der letzten Spalte (*Revenue*) aus. Ersetzen Sie den Spaltentitel durch den neuen Titel *Umsatz* und bestätigen Sie mit der **ENTER**-Taste.

■ 3.5 Spalten entfernen

Sie haben gesehen, wie Sie Zeilen über die Filterfunktion entfernen. Das Entfernen von Spalten geht noch einfacher:

- Markieren Sie die Spalte *Refill*, indem Sie auf den Spaltentitel klicken.
- Machen Sie einen Rechtsklick auf die Spaltenüberschrift und wählen Sie **ENTFERNEN**. Noch schneller geht es mit der **ENTF**-Taste auf Ihrer Tastatur.
- Sie können auch gleich mehrere Spalten auf einmal entfernen. Wählen Sie hierzu die Spalten *Amount* und *Price* mit gedrückter **STRG**-Taste aus und entfernen Sie sie.



Unter dem Menüeintrag **ENTFERNEN** gibt es auch die Option **ANDERE SPALTEN ENTFERNEN**. Mit diesem Befehl entfernen Sie alle Spalten außer der/den markierten. Dies erspart Klickarbeit, wenn Sie viele Spalten entfernen wollen.

Außerdem kann diese Funktion nützlich sein, wenn Sie nicht sicher sind, dass das Layout der Quelltablelle unverändert bleibt. Bei **ENTFERNEN** werden nur die ausgewählten Spalten entfernt. Wenn Sie **ANDERE SPALTEN ENTFERNEN** benutzen, funktioniert die Transformation auch noch, wenn zu einer späteren Aktualisierung zusätzliche Spalten in der Tabelle auftauchen.

■ 3.6 Werte ersetzen

Die Quelldatei verwendet offenbar ein fremdes Zahlenformat: Anstatt mit Komma werden Dezimalzahlen mit einem Punkt angegeben. Damit die Zahlen im deutschen System korrekt angezeigt werden, müssen die Punkte durch Kommas ersetzt werden.

- Klicken Sie mit der rechten Maustaste auf die Spaltenüberschrift *Umsatz* und wählen Sie **WERTE ERSETZEN**.
- Geben Sie als zu suchenden Wert einen Punkt und als Ersatzwert ein Komma an.
- Bestätigen Sie mit **OK**.

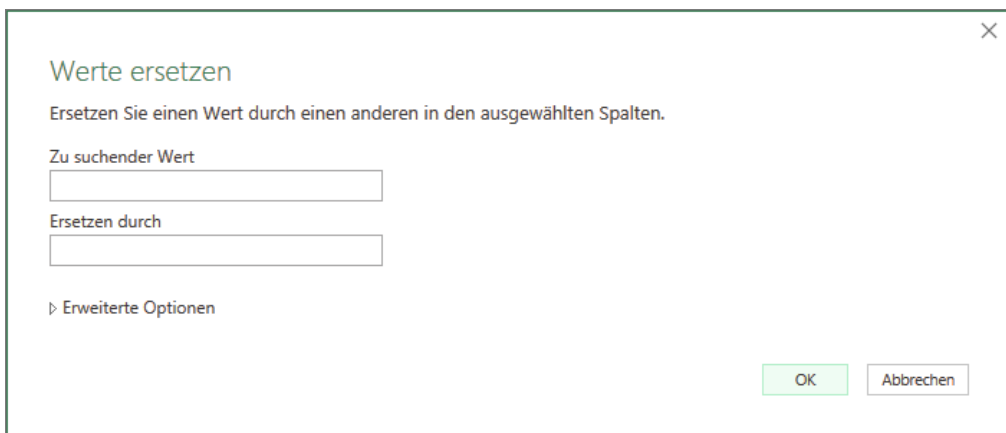


Bild 3.4 Mit *Werte ersetzen* ersetzen Sie Zeichenfolgen in einer Spalte.

■ 3.7 Datentypen bestimmen

In einer ordentlichen Tabelle enthält jede Spalte nur Daten eines bestimmten Typs. Im Abfrage-Editor können Sie daher festlegen, ob eine Spalte z. B. Zahlen, Texte oder Datums-
werte enthält.

- In den aktuellsten Versionen von Excel und Power BI sehen Sie neben den Spaltentiteln ein Symbol, das den Datentyp anzeigt. In unserem Beispiel sollte das „ABC“ für den Datentyp *Text* sein. Klicken Sie auf das ABC-Symbol links neben dem Spaltentitel *Umsatz*. Falls in Ihrer Version kein Typ-Symbol angezeigt wird, klicken Sie mit der rechten Maustaste auf den Spaltentitel und wählen Sie das Untermenü **TYP ÄNDERN**.
- Wählen Sie für die Umsatz-Spalte den Typ *Dezimalzahl*.

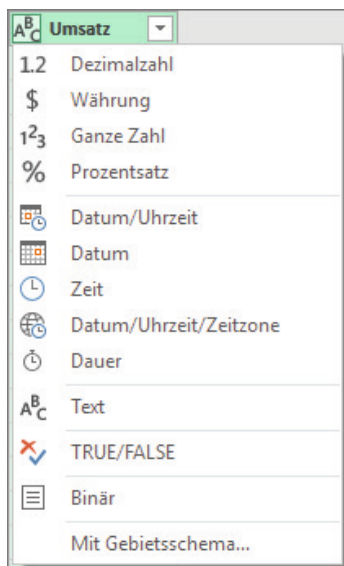


Bild 3.5
Auswahl des Datentyps.











Power Query versucht bei jeder Abfrage, die Datentypen automatisch zu bestimmen. Dies gelingt in vielen Fällen, doch Sie sollten sich nicht darauf verlassen. Für die Typbestimmung werden nur die ersten 200 Zeilen ausgewertet. Je nachdem, wie eindeutig die ersten 200 Zeilen aussehen und welche Transformationsschritte Sie noch durchführen wollen, kann die automatische Datentypänderung auch kontraproduktiv sein. Um immer auf der sicheren Seite zu sein, empfiehlt es sich daher, die Datentypen immer zu überprüfen und ggf. selbst zu bestimmen.




Das Einstellen des richtigen Datentyps hat mehrere Vorteile:

- Viele Transformationsfunktionen arbeiten nur mit bestimmten Datentypen. Beispielsweise können Berechnungen nur mit Zahlen-Spalten angestellt werden.
- Wie im Filter-Kapitel erwähnt, gibt es Spezialfilter, die nur bei bestimmten Typen möglich sind (Textfilter, Zahlenfilter, Datumsfilter, Zeitfilter, logische Filter).
- Alle Einträge werden entsprechend umgewandelt. Die Erkennungsfunktion ist dabei relativ flexibel. Beispielsweise wird der Text *1. Apr 17* beim Umwandeln in ein Datum richtig erkannt und dargestellt als *01.04.2017* (wobei die Anzeigart abhängig von der Einstellung des Betriebssystems ist).
- Umwandlungsfehler durch die automatische Datentyperkennung werden vermieden. So werden zum Beispiel bei Zahlen standardmäßig führende Nullen entfernt: Statt *001* wird nur *1* ausgegeben. Es gibt aber Fälle, in denen die führenden Nullen beibehalten werden sollen, z. B. bei Postleitzahlen, Telefonnummern oder Personalnummern. Verwenden Sie in diesen Fällen den Datentyp *Text*, damit die Nummern so belassen werden, wie sie sind.

In der folgenden Tabelle sehen Sie die verschiedenen Datentypen und ihren Anwendungsbereich.

Tabelle 3.1 Die Datentypen, die für Spalten möglich sind.

Symbol	Bezeichnung	Beschreibung	Beispiele
	Dezimalzahl	Dies ist der Standardtyp für Zahlenwerte und kann für ganze Zahlen sowie Dezimalzahlen mit sehr hoher Genauigkeit verwendet werden.	1,5 -7 0,00000152
	Feste Dezimalzahl/ Währung	Bei diesem Datentyp werden Zahlenwerte mit bis zu vier Nachkommastellen gespeichert. Alle weiteren Nachkommastellen werden speziell gerundet (siehe Hinweiskasten).	23,9887 12 -0,0001
	Ganze Zahl	Hiermit speichern Sie ganze Zahlen. Sollten die Zahlen Nachkommastellen haben, werden sie speziell gerundet (siehe Hinweiskasten).	0 -1 54888273
	Prozentsatz	Zahlenwerte mit diesem Datentyp werden als Prozentwerte interpretiert. Eine 1 in der Datenquelle entspricht 100%.	50% 99,99%
	Datum/ Uhrzeit	Zeitwerte mit Datum und Uhrzeit. Es werden alle Datumswerte ab 1900 bis zum Jahr 9999 unterstützt.	31.12.2000 14:00:01
	Datum	Datumswerte ohne Uhrzeit.	06.02.2017
	Zeit	Uhrzeit ohne zugehöriges Datum.	06:30 11:55
	Datum/ Uhrzeit/ Zeitzone	Zusätzlich zum Datum und der Uhrzeit wird noch die Zeitzone angezeigt. Die Zeitzone wird angegeben als der Unterschied zwischen der Ortszeit und der UTC-Zeit (in Deutschland +1:00 zur Winterzeit bzw. +2:00 zur Sommerzeit). Beim Laden der Abfrage werden die Werte jedoch ohne Zeitzone gespeichert und in den Datentyp <i>Datum/Uhrzeit</i> umgewandelt.	21.04.2001 08:30:00 +01:00
	Dauer	Ein Zeitraum, der z. B. bei der Addition oder Subtraktion von Datums- und Zeitwerten ermittelt wird. Er kann daher positiv oder negativ sein. Die Werte werden als Tage, Stunden, Minuten und Sekunden angezeigt, aber als Dezimalzahlen gespeichert. Eine 1 entspricht dabei einem Tag, 1,5 ist ein Tag und 12 Stunden.	1.12:00:00 (= ein- einhalb Tage) 0.00:30:00 (= eine halbe Stunde)
	Text	Zeichenketten, bestehend aus Buchstaben, Zeichen oder Zahlen. Eine einzelne Zeichenkette darf bis zu 268 Mio. Zeichen enthalten.	Das ist ein Text. 'XYZ{\$\$\$}' :-)

Symbol	Bezeichnung	Beschreibung	Beispiele
	TRUE/FALSE	Boolesche Wahrheitswerte, die für verschiedene M-Funktionen von Bedeutung sind, aber auch in Excel- oder DAX-Formeln verwendet werden können.	TRUE FALSE
	Binär	Ein einzelner Wert in einer Spalte kann kurioserweise auch eine Datei umfassen. Diese „Werte“ werden in der Vorschau-Tabelle nicht angezeigt, können aber in weiteren Transformationen verwendet werden. Hierfür wird der Datentyp <i>Binär</i> verwendet, der jede Art von Information im Binärformat speichert.	<i>Binary</i>
	Any	Spalten vom Typ <i>Any</i> können jede Art von Daten enthalten. Dieser Datentyp wird nach dem Einlesen einer Quelle angezeigt, wenn kein anderer Datentyp eingestellt wurde. <i>Any</i> ist eigentlich kein richtiger Datentyp, denn er kann nicht explizit zugewiesen werden, sondern bedeutet lediglich, dass kein Datentyp festgelegt wurde.	<i>(alles)</i>



Mathematisches Runden: Beim Umwandeln von Datentypen, z. B. von Dezimalzahlen auf ganze Zahlen, verwendet Power Query eine andere Rundungstechnik, als es die meisten Benutzer erwarten würden. So wird die Zahl 2,5 nicht aufgerundet zu 3, sondern abgerundet zu 2. Die Zahl 3,5 wird dagegen aufgerundet zu 4.

Es handelt sich dabei um die sogenannte *mathematische Rundung* (engl. *Round to Even*). Die Dezimalstellen 1, 2, 3, 4 werden wie gewohnt abgerundet und die Dezimalstellen 6, 7, 8, 9 werden wie gewohnt aufgerundet. Die 5er-Stellen werden allerdings nicht, wie bei der allgemein üblichen sogenannten *kaufmännischen Rundung* aufgerundet, sondern so auf- oder abgerundet, dass die gerundete Zahl bzw. Dezimalstelle immer gerade ist.

Beispiele:

- 1,1 wird zu 1.
- 1,5 wird zu 2.
- 1,8 wird zu 2.
- 2,4 wird zu 2.
- **2,5 wird zu 2.**
- 2,6 wird zu 3.
- 3,5 wird zu 4.
- **4,5 wird zu 4.**

Diese Art zu runden wird im Alltag zwar eher selten benutzt, hat aber tatsächlich Vorteile. Stellen Sie sich eine große Menge an Dezimalzahlen vor, die zu ganzen Zahlen gerundet werden. Die Zahlen, die auf ,5 enden, liegen

genau zwischen zwei ganzen Zahlen. Wenn sie immer aufgerundet werden, entsteht ein leichtes Ungleichgewicht. Der Durchschnitt oder die Summe aus den gerundeten Zahlen wird tendenziell größer sein als bei den ursprünglichen Dezimalzahlen. Beim mathematischen Runden wird davon ausgegangen, dass ungefähr genauso viele gerade wie ungerade Zahlen vorliegen. Wenn das der Fall ist, werden die 5er-Stellen genauso oft auf- wie abgerundet. Diese Rundung erzeugt somit weniger Abweichungen und wird daher auch symmetrische oder unverzerrte Rundung genannt.

Wenn Sie die herkömmliche Rundung verwenden wollen, benutzen Sie die Funktion `Number.Round` mit dem Rundungsmodus 2 (vgl. Kapitel 14.7.3).

■ 3.8 Transformations Schritte prüfen und benennen

Wenn Sie die Übung bis hierher gemacht haben, sehen Sie nun in der Liste auf der rechten Seite alle Transformationen, die Sie vorgenommen haben. Unter *Angewendete Schritte* können Sie die verschiedenen Schritte anklicken, um in der Vorschau das Ergebnis der jeweiligen Aktion zu sehen.

Achten Sie jedoch darauf, stets wieder den letzten Schritt zu markieren, bevor Sie eine neue Aktion durchführen – es sei denn, Sie möchten einen Schritt vor einem anderen Schritt einfügen.

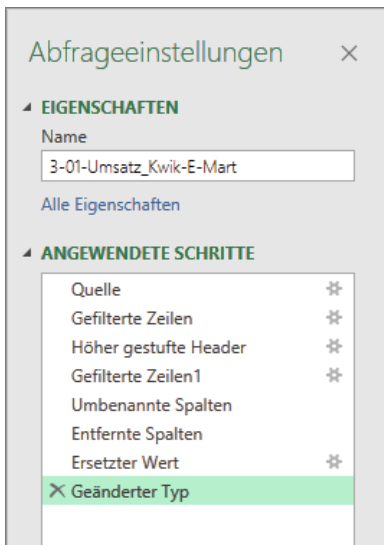


Bild 3.6
Alle Aktionen werden als Schritte aufgezeichnet.

Ihre Aktionen werden automatisch benannt, Sie können sie aber jederzeit umbenennen. Es empfiehlt sich, besonders bei umfangreichen Transformationen, den Schritten aussagekräftige Namen zu geben. Auf diese Weise können Sie später besser nachvollziehen, was genau gemacht wurde.

Bei der Benennung der Transformationsschritte sind Sie relativ frei. Es dürfen nur nicht zwei Schritte den gleichen Namen haben.

- Klicken Sie mit rechts auf den Schritt *Gefilterte Zeilen* und wählen Sie **UMBENENNEN**.
- Geben Sie den neuen Namen *Kopfzeilen entfernt* ein.
- Bestätigen Sie mit der **ENTER**-Taste.
- Ändern Sie auf die gleiche Weise den Namen des Schritts *Gefilterte Zeilen1* in *Zwischenüberschriften entfernt*.

■ 3.9 Abschließen der Transformation

Sie haben nun die Tabelle auf die relevanten Informationen reduziert, die Formate und Datentypen festgelegt sowie die Überschriften definiert. Sie ist nicht nur deutlich übersichtlicher als die Ausgangstabelle, sondern eignet sich nun für Visualisierungen in Power BI bzw. für Berechnungen und Auswertungen in Excel.

- Auf der rechten Fensterseite unter Abfrageeinstellungen sollte noch ein aussagekräftiger Name angegeben werden. Geben Sie den Namen *Umsatz November 2017* ein und bestätigen Sie mit der **ENTER**-Taste.
- Klicken Sie zum Abschließen der Bearbeitung in **Excel** auf **SCHLIESSEN & LADEN** bzw. in **Power BI** auf **SCHLIESSEN UND ÜBERNEHMEN**. Die Schaltflächen befinden sich jeweils links oben im Register **START**.

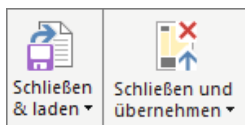


Bild 3.7

Hiermit beenden Sie den Editor (links: Excel, rechts: Power BI).

■ 3.10 Erneutes Bearbeiten der Abfrage

Die Transformationsschritte sowie die Verbindung zur Quelldatei sind nun in der Abfrage gespeichert. Bei jedem Aktualisieren der Abfrage wird die Quelldatei neu eingelesen und alle Transformationsschritte werden abgearbeitet.

Sie können die Schritte natürlich jederzeit anpassen.