

Daten aus mehreren Quellen kombinieren



Der Versuch, Wissen und Macht zu verbinden, war nur selten erfolgreich, und wenn, dann nur für eine kurze Zeit.

– Albert Einstein

In diesem Kapitel werden Sie

- Tabellen aus mehreren angegebenen Datenquellen anfügen
- Mehrere Excel-Arbeitsmappen aus einem Ordner anfügen
- Mehrere Tabellenblätter aus einer Excel-Arbeitsmappe anfügen

Während Albert Einsteins Zitat richtig sein mag und Weisheit und Macht nur selten erfolgreich kombiniert wurden, haben Sie mit Power Query die Macht, mehrere Quellen zu kombinieren, um so dauerhafte Weisheit zu erlangen. In diesem Kapitel erfahren Sie, wie Sie mehrere Tabellen in verschiedenen Komplexitätsstufen kombinieren, um Einblicke aus den kombinierten Datensets zu gewinnen.

In Kapitel 2, »Grundlegende Herausforderungen bei der Datenvorbereitung«, werden verschiedene Szenarien schlecht formatierter und nicht vorbereiteter Daten erläutert. Oft sind die Daten, mit denen Sie umgehen müssen, gut formatiert, aber auf mehrere Datenquellen verteilt. In solchen Szenarien besteht die größte Herausforderung darin, die Daten zusammenzuführen, um aus den einzelnen Silos eine einzige Quelle der Weisheit werden zu lassen. Das Anfügen mehrerer Tabellen steht im Mittelpunkt dieses Kapitels.

Dieses Kapitel beginnt mit einfachen Szenarien für das Zusammenführen von einer kleinen Anzahl von Tabellen, die ein gemeinsames Format aufweisen, und geht dann reale Herausforderungen aus dem Alltag an, wie beispielsweise sich ständig ändernde Datenquellen.

Ein paar Tabellen anfügen

Eines der häufig verwendeten Features in Power Query ist *Abfragen anfügen*, das Sie auf der Registerkarte *Home* des Power Query-Editors finden. Sie können einen vordefinierten Satz von Tabellen anfügen. Der Einfachheit halber beginnt dieses Kapitel mit dem Anfügen von zwei Tabellen, die das gleiche Format besitzen (das heißt, beide Tabellen verwenden die gleichen Spaltennamen).

Zwei Tabellen anfügen

Bevor wir zu der Art und Weise springen, wie zwei Tabellen angefügt werden, überprüfen wir die Gründe dafür, warum die beiden Tabellen für die Durchführung der Analyse angefügt werden, anstatt sie getrennt zu halten. Warum sollten Sie nicht einfach die beiden Tabellen kopieren und jede Tabelle auf einem eigenen Tabellenblatt in eine neue Excel-Arbeitsmappe einfügen und dann ein drittes Tabellenblatt für die erforderlichen Berechnungen verwenden? Während dieser Ansatz für Tabellen mit wenigen Zeilen und Spalten einfach sein kann, ist das Durchführen solcher Berechnungen für größere und sich ständig ändernde Tabellen eine völlig andere Sache.

So gehen Sie diese Herausforderung in Excel ohne Power Query an: Sie müssen die Tabellen kopieren und einfügen und diese dann manuell in einer einzigen Tabelle zusammenführen. Wenn sich die Daten ändern, müssen Sie die Tabelle erneut kopieren und einfügen. Alternativ müssen Sie komplexe Berechnungen erstellen, um die Bereiche in den verschiedenen Tabellenblättern zusammenzufassen. Wenn sich die Größe des Datensets ändert, müssen Sie die Zellbereiche in den Formeln aktualisieren. Glücklicherweise können Sie Power Query verwenden, um diesen aufwendigen Prozess zu automatisieren. Wenn sich beim Erstellen des nächsten Berichts die Daten in Ihren Tabellen geändert haben, reicht eine einfache Aktualisierung aus, und die Arbeit wird für Sie erledigt.

Übung 3-1: Beispiel Fahrräder und Zubehör

Diese Übung umfasst zwei Tabellen des fiktiven Unternehmens AdventureWorks, das in Kapitel 2 vorgestellt wurde. Die Daten stammen wie auch in Kapitel 2 aus der Beispieldatenbank von Microsoft. Für die Begleitdateien dieses Kapitels wurden die Spaltenüberschriften übersetzt; die eigentlichen Daten wurden unverändert übernommen.

Um das Problem dieser Übung in einen geschäftlichen Kontext zu stellen, stellen Sie sich vor, dass AdventureWorks, dessen Hauptproduktlinie Fahrräder sind, ein anderes Unternehmen übernimmt, dass das Produktportfolio um Zubehörprodukte bereichert. Als Chefanalyst von AdventureWorks werden Sie gebeten, einen neuen Bericht basierend auf den vereinheitlichten Produktlinien zu erstellen (siehe Abbildung 3-1).

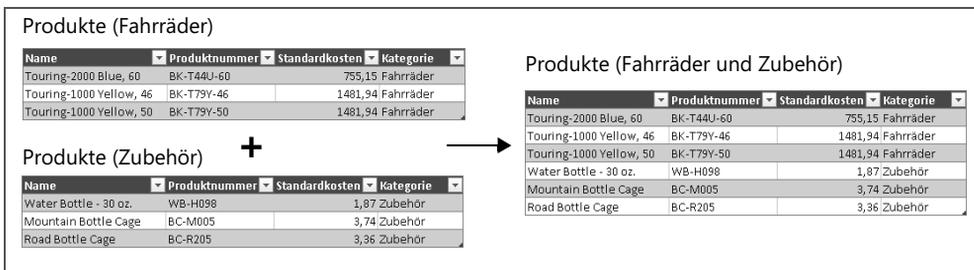


Abbildung 3-1 Ihr Geschäftsproblem besteht darin, die Tabellen für Fahrräder und Zubehör in einer einzigen Produkttabelle zusammenzubringen.

Um die Tabellen *Fahrräder* und *Zubehör* zusammenzufügen, können Sie mit einer neuen Excel-Arbeitsmappe oder einem Power BI Desktop-Bericht beginnen und die beiden Tabellen als separate Abfragen importieren. Sie können die Schritte dieser Übung ausführen, nachdem Sie die Dateien *K03U01 - Fahrräder.xlsx* und *K03U01 - Zubehör.xlsx* aus den Begleitdateien zu diesem Buch in den Ordner *C:\Daten\K03* kopiert haben.

1. Öffnen Sie eine leere neue Excel-Arbeitsmappe oder einen Power BI-Bericht.
In Excel: Wählen Sie auf der Registerkarte *Daten* den Befehl *Daten abrufen | Aus Datei | Aus Arbeitsmappe*.
In Power BI Desktop: Wählen Sie *Daten abrufen | Excel | Verbinden*.
2. Wählen Sie die Datei *K03U01 - Fahrräder.xlsx* aus und klicken Sie auf *Importieren*.
3. Wählen Sie im Navigator *Fahrräder* aus und klicken Sie auf *Daten transformieren*.
4. Wählen Sie im Power Query-Editor *Schließen & laden in*. Wenn Sie Power BI Desktop verwenden, wählen Sie *Schließen & übernehmen* und überspringen Sie Schritt 5.
5. Wählen Sie im Dialogfeld *Daten importieren* die Option *Tabelle* aus und klicken Sie auf *OK*. Sie haben jetzt alle Daten zu den Fahrrädern in Ihre neue Arbeitsmappe oder in Power BI Desktop importiert. Wenn sich die Daten in *K03U01 - Fahrräder.xlsx* ändern und der Produktlinie neue Fahrräder hinzugefügt werden, können Sie die neuesten Daten abrufen, indem Sie auf der Registerkarte *Daten* von Excel auf *Alle aktualisieren* klicken oder **[Strg] + [Alt] + [F5]** drücken. Wenn Sie Power BI Desktop verwenden, können Sie einfach auf der Registerkarte *Home* die Option *Aktualisieren* anklicken.
Es ist an der Zeit, die Tabelle *Zubehör* in die Arbeitsmappe oder den Power BI-Bericht zu importieren.
6. **In Excel:** Wählen Sie auf der Registerkarte *Daten* den Befehl *Daten abrufen | Aus Datei | Aus Arbeitsmappe*.
In Power BI Desktop: Wählen Sie *Daten abrufen | Excel | Verbinden*.
7. Wählen Sie die Datei *K03U01 - Zubehör.xlsx* aus und klicken Sie auf *Importieren*.
8. Wählen Sie im Navigator *Fahrräder* aus und klicken Sie auf *Daten transformieren*.
9. Um im Power Query-Editor *Fahrräder* an *Zubehör* anzufügen, stellen Sie sicher, dass die Abfrage *Zubehör* ausgewählt ist, und wählen Sie auf der Registerkarte *Home* den Befehl *Kombinieren | Abfragen anfügen | Abfragen anfügen* aus. Beachten Sie, dass das Dropdownmenü der Schaltfläche *Abfragen anfügen* auch den Befehl *Abfragen als neu anfügen* enthält, der in Übung 3-2 beschrieben wird.
10. Wenn das Dialogfeld *Anfügen* geöffnet wird, wählen Sie *Fahrräder* als anzufügende Tabelle aus und klicken dann auf *OK*.
11. Achten Sie darauf, dass der Bereich *Abfragen* (der Bereich an der linken Seite des Power Query-Editors) erweitert ist.
Beachten Sie, dass sich im Bereich *Abfragen* zwei Abfragen befinden: *Fahrräder* und *Zubehör*. Da Sie die Abfrage *Fahrräder* soeben an *Zubehör* angefügt haben, spiegelt der Name der Abfrage nicht mehr die enthaltenen Daten wider. Sie sollten daher die Abfrage *Zubehör* umbenennen und sicherstellen, dass diese Abfrage und nicht *Fahrräder* in die Excel-Arbeitsmappe oder den Power BI-Bericht geladen wird. Um die Abfrage umzubenennen, kli-

cken Sie mit der rechten Maustaste auf die Abfrage *Zubehör*, wählen Sie *Umbenennen* aus und benennen Sie die Abfrage in **Produkte** um. Sie können die Abfrage auch im Abschnitt *Eigenschaften* des Bereichs *Abfrageeinstellungen* umbenennen.

12. Führen Sie diese Schritte durch, um *Fahrräder* zu entladen:

In Excel:

- a. Wählen Sie *Schließen & laden in*.
- b. Markieren Sie im Dialogfeld *Daten importieren* die Option *Tabelle* und klicken Sie auf *OK*.
- c. Um *Fahrräder* aus der Arbeitsmappe zu entladen, klicken Sie im Bereich *Abfragen & Verbindungen* mit der rechten Maustaste auf die Abfrage *Fahrräder* und wählen Sie *Laden in*.
- d. Wählen Sie im Dialogfeld *Daten importieren* die Option *Nur Verbindung erstellen* aus und klicken Sie auf *OK*.
- e. Wenn das Meldungsfeld *Möglicher Datenverlust* geöffnet wird, klicken Sie auf *OK*, um die Abfrage zu entladen und die Tabelle zu löschen.
- f. Löschen Sie schließlich *Tabelle2*, wenn Sie hier die Tabelle *Fahrräder* ursprünglich geladen haben, und benennen Sie das letzte Blatt in **Produkte** um.

In Power BI Desktop: Klicken Sie im Bereich *Abfragen* die Abfrage *Fahrräder* mit der rechten Maustaste an, deaktivieren Sie die Option *Laden aktivieren* und wählen Sie *Schließen & übernehmen*.

Sie haben jetzt eine kombinierte Tabelle *Produkte*, die die kombinierten Daten von *Fahrräder* und *Zubehör* enthält. Wenn Sie jedoch den Bereich *Abfragen* genauer betrachten, werden Sie die asymmetrische Natur des Anfügeschritts bemerken. Sie haben jetzt die Abfrage *Fahrräder* und die kombinierte Abfrage *Produkte*. Aber wo ist *Zubehör*? Erinnern Sie sich daran, dass Sie direkt nach dem Import von *Zubehör* die Daten aus *Fahrräder* an diese Abfrage angehängt und dann in *Produkte* umbenannt haben. Wenn Sie in Zukunft einige Änderungen an der Tabelle *Zubehör* vornehmen möchten, wird es schwierig sein, den richtigen Schritt zum Ändern zu finden, da Sie jetzt eine Abfrage mit den kombinierten Transformationsschritten von *Zubehör* und der durch Anfügen erstellten Tabelle *Produkte* haben. Ein sauberer Ansatz wäre, *Zubehör* als separate Abfrage zu verwenden – und aus diesem Grund verfügt die Transformation *Abfragen anfügen* über eine weitere Transformation mit dem Namen *Abfragen als neu anfügen*.

Übung 3-2, Teil 1: Abfragen als neu anfügen verwenden

Auf der Registerkarte *Home* des Power Query-Editors können Sie in der Gruppe *Kombinieren* Abfragen auf mehreren Wegen anfügen: mithilfe von *Abfragen anfügen* und mit *Abfragen als neu anfügen* (siehe Abbildung 3-2).

In der vorherigen Übung haben Sie zwei Tabellen mit der Transformation *Abfragen anfügen* angefügt; hierbei wird eine zweite Abfrage an die aktuelle angefügt. *Abfragen als neu anfügen* erstellt eine dritte Abfrage, die die beiden Abfragen kombiniert, und hierdurch bleiben die beiden Ausgangsabfragen davon unberührt.

Abfragen	Name	Produktnummer	Farbe	Standardkosten	Listenpreis
1	Water Bottle - 30 oz.	WB-H098		null	1,8663
2	Mountain Bottle Cage	BC-M005		null	3,7363
3	Road Bottle Cage	BC-R205		null	3,3623
4	Patch Kit/8 Patches	PK-7098		null	0,8565
5	Hitch Rack - 4-Bike	RA-H123		null	44,88
6	Bike Wash - Dissolver	CL-9009		null	2,9733
7	Fender Set - Mountain	FE-6654		null	8,2205
8	All-Purpose Bike Stand	ST-1401		null	59,466
9	Hydration Pack - 70 oz.	HY-1023-70	Silver		20,5663
10	Mountain Tire Tube	TT-M928		null	1,8663
11	Road Tire Tube	TT-R982		null	1,4923
12	Touring Tire Tube	TT-T092		null	1,8663
13	LL Mountain Tire	TI-M267		null	9,3463
14	ML Mountain Tire	TI-M602		null	11,2163
15	HL Mountain Tire	TI-M823		null	13,09
16	LL Road Tire	TI-R092		null	8,0373
17	ML Road Tire	TI-R628		null	9,3463

Abbildung 3-2 Die Transformation *Abfragen als neu anfügen* ermöglicht es Ihnen, für die angefügte Tabelle und ihre Teile eigene Abfragen zu haben.

Anstatt *Zubehör* so umzuwidmen, dass dort alle Produkte aufgenommen werden, bleibt in dieser Übung *Zubehör* als separate Abfrage bestehen und Sie erstellen eine neue für die Produkte.

1. Führen Sie die Schritte 1 bis 8 in Übung 3-1 aus, um die Abfragen *Fahrräder* und *Zubehör* in den Power Query-Editor zu laden.
2. Achten Sie darauf, dass im Power Query-Editor die Abfrage *Zubehör* ausgewählt ist, und wählen Sie auf der Registerkarte *Home* im Dropdownmenü der Schaltfläche *Abfragen anfügen* den Befehl *Abfragen als neu anfügen* aus.
3. Wenn das Dialogfeld *Anfügen* geöffnet wird, wählen Sie *Fahrräder* als Tabelle aus, die an die primäre Tabelle angefügt werden soll, und klicken Sie auf *OK*. Sie haben jetzt eine neue Abfrage, *Append1*, mit den Kombinationen der Daten aus *Fahrräder* und *Zubehör*.
4. Benennen Sie die neue Abfrage in **Produkte** um und führen Sie Schritt 12 aus Übung 3-1 aus, um *Fahrräder* zu entladen. Um *Zubehör* zu entladen, können Sie wieder Schritt 12 ausführen, aber wenden Sie die Anweisungen dieses Mal auf *Zubehör* und nicht auf *Fahrräder* an.
5. Speichern Sie Ihre Arbeitsmappe oder Ihren Power BI-Bericht. Als Nächstes lernen Sie eine andere Methode kennen, um *Fahrräder* und *Zubehör* an eine neue Abfrage anzufügen, und zwar mit dem Befehl *Verweis*. Sie vergleichen außerdem den M-Ausdruck, der den beiden Methoden zugrunde liegt.



Wenn Sie einen kleinen Satz von Tabellen anfügen, wird empfohlen, *Abfragen als neu anfügen* anstelle von *Abfragen anfügen* zu verwenden, damit Sie in Zukunft die separaten Bausteine der kombinierten Tabelle problemlos bearbeiten können.

Übung 3-2, Teil 2: Abfrageabhängigkeiten und Verweise

In Übung 3-2 Teil 1, haben Sie *Abfragen als neu anfügen* angewendet, was zu einer neuen Abfrage führte, in der die Daten aus *Fahrräder* und *Zubehör* kombiniert waren. Die neue Abfrage referenzierte die beiden anderen Abfragen und verwendete deren Ausgabe als Eingabe. Wenn eine Abfrage auf eine andere Abfrage verweist, wird eine Abhängigkeit zwischen den Abfragen erstellt. Wenn eine Abfrage von anderen Abfragen abhängt, können Sie sie nicht löschen. (Wenn Sie zum Beispiel versuchen, *Fahrräder* zu löschen, wird eine Meldung angezeigt, die Sie daran hindert, da *Produkte* von *Fahrräder* abhängig ist.)

Um die Abhängigkeiten zwischen Abfragen anzuzeigen, öffnen Sie die Excel-Arbeitsmappe oder den Power BI-Bericht, die/den Sie in Übung 3-2, Teil 1 erstellt haben, und starten Sie den Power Query-Editor. Wählen Sie auf der Registerkarte *Ansicht* den Befehl *Abfrageabhängigkeiten*. Das Dialogfeld *Abfrageabhängigkeiten* wird geöffnet, wie in Abbildung 3-3 dargestellt.

Das Dialogfeld *Abfrageabhängigkeiten* zeigt ein Flussdiagramm, das von den Abfragen *Fahrräder* und *Zubehör* als Datenquellen ausgeht, durch die Zwischenabfragen *Fahrräder* und *Zubehör* fließt und schließlich bei *Produkte* endet, in der Sie die beiden Tabellen kombiniert haben. Diese Ansicht ist nützlich, wenn Sie an komplexen Datenvorbereitungsszenarien arbeiten.

Nachdem Sie nun wissen, wie Sie die Abhängigkeiten zwischen Abfragen anzeigen können, werden Sie mehr über Verweise (Referenzen) erfahren, mit denen Sie umfangreiche und wiederverwendbare Datentransformationsflüsse erstellen und die Herausforderungen bei realen Daten bewältigen können. Ein Verweis kann verwendet werden, um eine neue Abfrage zu erstellen, deren Datenquelle, also die Eingabedaten, die Ausgabe einer anderen Abfrage ist.

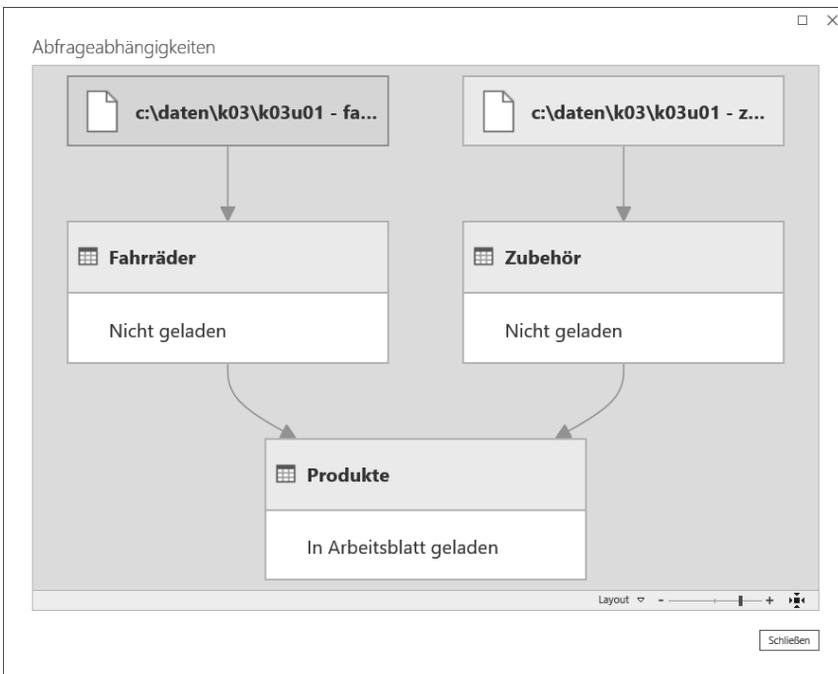


Abbildung 3-3 Das Dialogfeld *Abfrageabhängigkeiten* zeigt die Verweise und Abhängigkeiten zwischen den Abfragen eines Berichts an.

Weiter vorne haben Sie in Übung 3-2, Teil 1, *Fahrräder* und *Zubehör* als Abfragen geladen und *Abfragen als neu anfügen* angewendet, um eine neue Abfrage mit den kombinierten Ergebnissen zu erstellen. Eine alternative Methode, um das gleiche Ergebnis zu erzielen, ist die Verwendung einer Kombination aus *Abfragen anfügen* und *Verweis*:

1. Führen Sie die Schritte 1 bis 8 aus Übung 3-1 aus, um die Abfragen *Fahrräder* und *Zubehör* in den Power Query-Editor zu laden.
2. Öffnen Sie den Bereich *Abfragen* und klicken Sie mit der rechten Maustaste auf *Zubehör*. Wählen Sie *Verweis*. Beachten Sie, dass Sie jetzt eine neue Abfrage haben, *Zubehör (2)*, die mit der Ausgabe von *Zubehör* beginnt.
3. Benennen Sie *Zubehör (2)* in **Produkte** um.
4. Achten Sie darauf, dass die Abfrage *Produkte* ausgewählt ist, und wählen Sie auf der Registerkarte *Home* den Befehl *Abfragen anfügen*. Wenn das Dialogfeld *Anfügen* geöffnet wird, wählen Sie *Fahrräder* als anzufügende Tabelle aus und klicken Sie auf *OK*.

Nun können Sie untersuchen, wie Power Query die Schritte *Verweis* und *Abfrage anfügen* in den zugrunde liegenden M-Formeln verarbeitet hat, und die Ausgabe der Kombination aus den Befehlen *Verweis* und *Abfrage anfügen* mit der Ausgabe von Übung 3-2, Teil 1, vergleichen. Betrachten wir den zugrunde liegenden M-Code.



Rufen Sie sich in Erinnerung, dass in Kapitel 1, »Einführung in Power Query«, kurz der M-Code erwähnt wird, der von Power Query generiert wird, um die Transformationsschritte auszuführen. Jeder Transformationsschritt, den Sie im Power Query-Editor erstellen, wird in eine Codezeile übersetzt, die in der Bearbeitungsleiste angezeigt werden kann.

Schalten Sie, falls erforderlich, auf der Registerkarte *Ansicht* die *Bearbeitungsleiste* ein und wählen Sie dann im Abschnitt *Angewendete Schritte* im Bereich *Abfrageeinstellungen* den Eintrag *Quelle* aus. Hier ist die M-Formel, die Sie sehen:

```
= Zubehör
```

Dieser Code repräsentiert die Ausgabe der Abfrage *Zubehör*, die als Eingabe für die neue Abfrage verwendet wird. Wie Sie vielleicht vermuten, wurde dieser Code durch Anwenden von *Verweis* auf die Abfrage *Zubehör* generiert.

Klicken Sie auf der Registerkarte *Home* auf *Erweiterter Editor*, um sich den M-Ausdruck anzusehen:

```
let
    Quelle = Zubehör,
    #"Angefügte Abfrage" = Table.Combine({Quelle, Fahrräder})
in
    #"Angefügte Abfrage"
```

Kapitel 9, »Einführung in die Power Query-Formelsprache M«, erläutert diese Syntax ausführlicher und erklärt die Bedeutung von *let* und *in*. Konzentrieren Sie sich daher im Moment nur auf die beiden Zeilen zwischen *let* und *in*.

In der ersten Zeile nach *let* wird die Ausgabe der Abfrage *Zubehör* in *Quelle* platziert. (Sie können sehen, dass *Quelle* auch der erste Schritt im Bereich *Angewendete Schritte* ist.) Die zweite Zeile verwendet *Table.Combine*, um *Quelle* und *Fahrräder* zu kombinieren – hierbei wird *Fahrräder* an *Zubehör* angefügt. (Sie haben möglicherweise bemerkt, dass die Namen der Abfragen in geschweiften Klammern eingeschlossen sind, was im nächsten Abschnitt erläutert wird. Hinweis: Geschweifte Klammern werden verwendet, um eine Liste zu erstellen.)

Wenn Sie sich nun den M-Code ansehen, der generiert wurde, als Sie in Übung 3-2, Teil 1, den Befehl *Abfragen als neu anfügen* verwendet haben, finden Sie den folgenden Code im erweiterten Editor:

```
let
    Source = Table.Combine({Zubehör, Fahrräder})
in
    Source
```

Wenn Sie sich diese beiden Codeabschnitte genau ansehen, werden Sie feststellen, dass die beiden Ausdrücke die gleiche Ausgabe liefern: die kombinierten Tabellen *Zubehör* und *Fahrräder*.

Zum Abschluss dieser Übung können Sie eine Kombination aus *Verweis* und *Abfragen anfügen* verwenden, mit der Sie dieselben Ergebnisse erzielen wie bei der Verwendung von *Abfragen als neu anfügen*. Während die letztgenannte Technik zu weniger Codezeilen führt und in der Benutzeroberfläche nur einen einzigen Schritt anstelle von zweien erfordert, gibt es keinen wirklichen Unterschied zwischen den beiden Methoden. Warum habe ich also den Befehl *Verweis* eingeführt? Wie Sie in späteren Kapiteln sehen werden, kann *Verweis* sehr hilfreich beim Erstellen einer Pipeline von Transformationsabfragen oder mehreren Transformationszweigen derselben Basisabfrage sein.

Drei oder mehr Tabellen anfügen

Bisher haben Sie in diesem Kapitel gelernt, wie Sie zwei Tabellen verketteten. Es ist an der Zeit, zu einer allgemeineren Herausforderung überzugehen: wie man drei oder mehr Tabellen verkettet. Ein Ansatz besteht darin, zunächst zwei Tabellen zu kombinieren und dann die Ausgabe der beiden Tabellen an eine dritte Tabelle anzuhängen. Als Nächstes können Sie die Ausgabe dieser Schritte mit einer vierten Tabelle kombinieren und so weiter. Obwohl dieser Ansatz gut funktionieren kann, ist er schwer zu pflegen, da er die Erstellung von Zwischenabfragen erfordert. Wenn Sie drei oder vier Tabellen kombinieren müssen, ist dies kein signifikanter Aufwand, aber für eine größere Anzahl von Abfragen ist ein besserer Ansatz erforderlich.

Übung 3-2, Teil 3: Fahrräder + Zubehör + Komponenten

Kehren wir zu unserer fiktiven Geschichte zurück. Nach der erfolgreichen Integration der Zubehör-Produktlinie entschließt sich AdventureWorks, ein neues Unternehmen zu akquirieren und sein Portfolio auf Komponenten auszudehnen. Ihr Ziel als Chefanalytiker ist, alle Produktlinien in einer einzigen Tabelle zu kombinieren.

In dieser Übung werden Sie die folgenden drei Arbeitsmappen verwenden, die sich in den Begleitdateien zu diesem Buch befinden. Kopieren Sie diese drei Dateien in den Ordner *C:\Daten\K03*:

- K03U01 - Fahrräder.xlsx
- K03U01 - Zubehör.xlsx
- K03U01 - Komponenten.xlsx

Öffnen Sie eine neue leere Excel-Arbeitsmappe oder einen neuen Power BI-Bericht und importieren Sie die drei Excel-Arbeitsmappen in den Power Query-Editor. Jede Arbeitsmappe enthält ein einzelnes Tabellenblatt mit einem entsprechenden Namen (das heißt ein Tabellenblatt *Fahrräder* in *K03U01 - Fahrräder.xlsx*, ein Tabellenblatt *Zubehör* in *K03U01 - Zubehör.xlsx* und ein Tabellenblatt *Komponenten* in *K03U01 - Komponenten.xlsx*). Wenn Sie nicht sicher sind, wie Sie diese Arbeitsmappen importieren können, finden Sie unter Übung 3-1 die Schritte zum Laden von *Fahrräder* und *Zubehör*. Verwenden Sie dann die entsprechenden Schritte, um *Komponenten* zu importieren.

1. Wählen Sie im Bereich *Abfragen* des Power Query-Editors *Komponenten* aus.
2. Wählen Sie auf der Registerkarte *Home* den Befehl *Abfragen als neu anfügen*.
3. Wenn das Dialogfeld *Anfügen* geöffnet wird, schalten Sie das Optionsfeld *Drei oder mehr Tabellen* ein. Die Listen *Verfügbare Tabelle(n)* und *Anzufügende Tabellen* werden angezeigt, wie in Abbildung 3-4 dargestellt. *Fahrräder* und *Zubehör* finden Sie unter *Verfügbare Tabelle(n)*.
4. Fügen Sie *Fahrräder* und *Zubehör* der Liste *Anzufügende Tabellen* hinzu und klicken Sie auf *OK*.

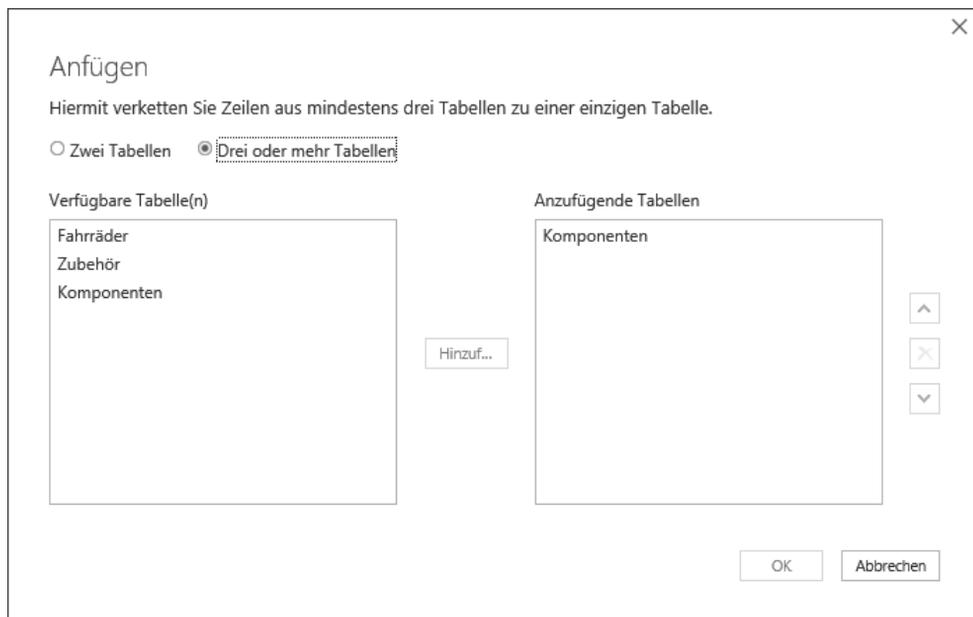


Abbildung 3-4 Im Dialogfeld *Anfügen* können Sie mit der Option *Drei oder mehr Tabellen* mehrere Abfragen auswählen, die angefügt werden sollen.

Sie haben jetzt eine neue Abfrage, *Append1*, mit der kombinierten Produktlinie. Von hier aus können Sie diese Abfrage in **Produkte** umbenennen und die Zwischenabfragen entladen, wie Sie es in Übung 3-1 gesehen haben.



Im Dialogfeld *Anfügen*, das Sie in Abbildung 3-4 sehen, können Sie die Reihenfolge der Tabellen in der kombinierten Tabelle steuern. Sie können eine Tabelle auswählen und ihre Reihenfolge in der Liste ändern, indem Sie auf die Aufwärts- und Abwärtspeile auf der rechten Seite von *Anzufügende Tabellen* klicken. Sie können sogar eine Tabelle mehrmals in die Liste *Anzufügende Tabellen* einfügen und so die Daten in der kombinierten Tabelle duplizieren.

Nachdem Sie nun *Fahrräder*, *Zubehör* und *Komponenten* zusammengefügt haben, können Sie sehen, wie einfach es ist, dieselbe Transformation in der Bearbeitungsleiste des Power Query-Editors anzuwenden. Wenn die Abfrage *Produkte* ausgewählt ist, wird die folgende Formel angezeigt:

```
= Table.Combine({Komponenten, Fahrräder, Zubehör})
```

Kommt Ihnen das bekannt vor? Als Sie weiter vorne zwei Tabellen angefügt haben, enthielt die Bearbeitungsleiste den folgenden Code:

```
= Table.Combine({Zubehör, Fahrräder})
```

An der Syntax dieser Formel können Sie erkennen, dass *Table.Combine* die Funktion ist, die zum Ausführen der Anfügetransformation mehrerer Tabellen verwendet wird. Der Funktion können in den geschweiften Klammern mehrere Tabellen übergeben werden.



In **M** werden geschweifte Klammern verwendet, um eine Liste zu definieren. Die einzelnen Elemente der Liste werden durch Kommata getrennt. Weitere Beispiele für Listen finden Sie in späteren Kapiteln. Mit Listen können Sie leistungsstarke Transformationen erstellen und sogar Ihre Abfrage zur Laufzeit bearbeiten.

Da Sie nun wissen, wie Sie drei Tabellen anfügen, können Sie eine vierte Tabelle hinzufügen, aber dieses Mal werden Sie dies nur mithilfe von Code tun. (Lassen Sie Ihre Excel-Arbeitsmappe oder Ihren Power BI-Bericht für die nächste Übung geöffnet.)

Übung 3-2, Teil 4: Fahrräder + Zubehör + Komponenten + Kleidung

Kehren wir zu unserem fiktiven Szenario zurück. Mit der Erweiterung der Produktlinien beschließt der Chief Operating Officer von AdventureWorks, den Geschäftsbereich Bekleidung mit den Kern-Business-Units zu konsolidieren. Als Chefanalytiker erhalten Sie eine vierte Tabelle mit der Bekleidungs-Produktlinie. In Teil 4 dieser Übung erfahren Sie, wie Sie die vier Tabellen kombinieren – diesmal nur mit M und ohne das Dialogfeld *Anfügen*.

In dieser Übung verwenden Sie zusätzlich die Arbeitsmappe *K03U01 - Kleidung.xlsx*, die sich in Begleitdateien zu diesem Buch befindet. Speichern Sie auch diese Datei im Ordner *C:\Daten\K03* ab. Öffnen Sie die Excel-Arbeitsmappe oder den Power BI Desktop-Bericht aus Übung 3-2, Teil 3, und importieren Sie *Kleidung* als neue Abfrage. (Siehe Übung 3-1 für die Schritte zum Laden von *Fahrräder* und *Zubehör*. Wenden Sie dann ähnliche Schritte für *Kleidung* an.) Benennen Sie die neue Abfrage in **Kleidung** um.

Wählen Sie im Power Query-Editor die Abfrage *Produkte* aus. Sehen Sie sich die Bearbeitungsleiste an. Sie sollten dort die folgende Formel sehen:

```
= Table.Combine({Komponenten, Fahrräder, Zubehör})
```

Sie müssen diese Formel ändern, indem Sie irgendwo in den geschweiften Klammern *Kleidung* hinzufügen. Stellen Sie sicher, dass alle vier Tabellen innerhalb der geschweiften Klammern durch Kommata getrennt sind. Wenn Sie fertig sind, drücken Sie die -Taste. Ihre Formel kann wie folgt aussehen:

```
= Table.Combine({Kleidung, Komponenten, Fahrräder, Zubehör})
```

oder:

```
= Table.Combine({Komponenten, Kleidung, Fahrräder, Zubehör})
```

oder:

```
= Table.Combine({Komponenten, Fahrräder, Kleidung, Zubehör})
```

oder:

```
= Table.Combine({Komponenten, Fahrräder, Zubehör, Kleidung})
```

Um zu überprüfen, ob *Table.Combine* funktioniert hat, klicken Sie auf das Filtersteuerelement in der Überschrift der Spalte *ÜbergeordneteProduktkategorieName*. Wenn der Bereich *Filter* geöffnet wird, sollten vier Werte angezeigt werden: *Accessoires*, *Bikes*, *Components* und *Clothing*. Das Überprüfen der Werte im Filterbereich ist eine einfache und praktische Methode zum Überwachen Ihrer Transformationsschritte.

Nachdem Sie nun *Kleidung* an die *Produkte* angefügt haben, können Sie zu neuen Szenarien wechseln und erfahren, wie Sie mehrere Tabellen in größerem Maßstab anhängen.

Tabellen in größerem Maßstab anhängen

Im vorherigen Abschnitt haben Sie gelernt, wie Sie mehrere Tabellen anfügen, indem Sie im Power Query-Editor *Abfragen anfügen* und *Abfragen als neu anfügen* verwenden. Diese Transformationen sind zwar nützlich, wenn Sie über einen vordefinierten Satz von Tabellennamen verfügen, aber in einigen häufig vorkommenden Szenarien können Sie nicht im Voraus wissen, welche Tabellennamen verfügbar sind, und Sie müssen häufig einen Bericht über die kombinierten Daten erstellen. Daher kann die manuelle Bearbeitung der Abfragen, um bei jeder Änderung die aktuellen Tabellennamen einzubeziehen, eine zeitaufwendige Aktion sein, die Sie besser vermeiden sollten.

Tabellen aus einem Ordner anfügen

Das am häufigsten vorkommende Szenario und seine Herausforderung lassen sich an diesem Beispiel verdeutlichen: Stellen Sie sich vor, Sie haben mehrere Tabellen, die Sie kombinieren müssen. Jede Tabelle befindet sich in einer separaten Excel-Arbeitsmappe, deren Name ein anderes Jahr und einen anderen Monat enthält:

- Jan2018.xlsx
- Feb2018.xlsx
- Mrz2018.xlsx
- Apr2018.xlsx

Jeden Monat wird eine neue Excel-Arbeitsmappe in einem freigegebenen Ordner gespeichert, und Sie werden aufgefordert, einen aktualisierten Bericht vorzubereiten, der alle Tabellen aus dem Ordner enthält. Angenommen, Sie haben bereits einen Bericht erstellt, in dem jede der Tabellen separat importiert und mit der weiter vorne beschriebenen Technik angefügt wurde. Wenn eine neue Datei hinzugefügt wird (zum Beispiel *Mai2018.xlsx*), müssen Sie eine Reihe von Schritten manuell ausführen, um die Abfragen zu ändern. Zuerst importieren Sie die Excel-Datei als neue Abfrage mit dem Namen *Mai2018*. Anschließend bearbeiten Sie in *Angewendete Schritte* im Power Query-Editor den Anfügeschritt, um die neue Abfrage zu ergänzen. Sie können die Abfrage in der folgenden Formel hinzufügen (wobei die Änderung fett hervorgehoben ist):

```
= Table.Combine({Jan2018, Feb2018, Mar2018, Apr2018, Mai2018})
```

Dieser Ansatz ist letztendlich sehr mühsam, wenn Sie dies jeden Monat wiederholen müssen. Wäre es nicht besser, eine verbesserte Abfrage zu erstellen, die automatisch alle Arbeitsmappen aus einem Ordner anfügt – mit einer einzigen Aktualisierung der Arbeitsmappe und ohne manuelles Bearbeiten der Abfrage? Dies ist machbar, und zwar dank der Möglichkeit, Daten aus einem Ordner abzurufen.

Übung 3-3: AdventureWorks-Produkte aus einem Ordner anfügen

Kehren wir zu AdventureWorks zurück, um die Herausforderung zu veranschaulichen. Angenommen, ein VBA-Entwickler hat ein Skript erstellt, das eine Excel-Datei aktualisiert, wenn ein neues Produkt veröffentlicht wird. Jedes Jahr wird eine neue Excel-Datei erstellt, die im laufenden Betrieb mit neu veröffentlichten Produkten aktualisiert wird. Die Produktdateien werden in einem freigegebenen Ordner aggregiert, wie in Abbildung 3-5 dargestellt. Als Chefanalytiker möchten Sie einen Bericht erstellen, der alle Produkte kombiniert.

Um die Schritte in dieser Übung auszuführen, können Sie die Dateien *KOUE03 - 2015.xlsx*, *KOUE03 - 2016.xlsx* und *KOUE03 - 2017.xlsx* verwenden, die Sie in den Begleitdateien zu diesem Buch finden. Kopieren Sie diese Arbeitsmappen in den Ordner *C:\Daten\K03\K03U03 - Produkte*.

1. Erstellen Sie eine leere neue Excel-Arbeitsmappe oder einen neuen Power BI Desktop-Bericht.

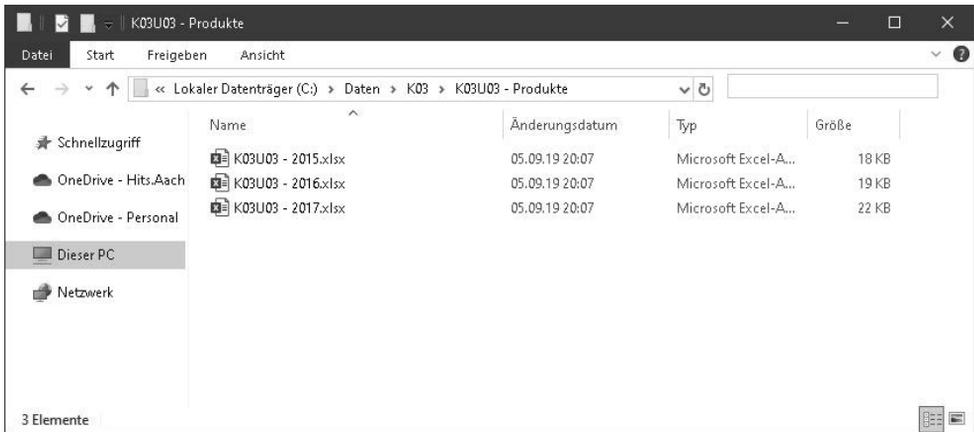


Abbildung 3-5 In diesem Beispiel sammelt AdventureWorks in einem freigegebenen Ordner Excel-Arbeitsmappen mit Produkten; die Produkte, die in einem bestimmten Jahr veröffentlicht wurden, befinden sich jeweils in einer eigenen Arbeitsmappe.

In Excel: Wählen Sie auf der Registerkarte *Daten* den Befehl *Daten abrufen | Aus Datei | Aus Ordner*.

In Power BI Desktop: Wählen Sie auf der Registerkarte *Home* den Befehl *Daten abrufen | Datei | Ordner* und klicken Sie auf *Verbinden*.

2. Navigieren Sie im geöffneten Dialogfeld *Ordner* zu dem zuvor erstellten Ordner (siehe Abbildung 3-5) oder kopieren Sie den Pfad und fügen Sie ihn in das Textfeld *Ordnerpfad* ein und klicken Sie auf *OK*.
3. Klicken Sie im Dialogfeld, das dann geöffnet wird (und das den Namen des Ordners, den Sie in Schritt 2 ausgewählt haben, als Titel verwendet), auf *Kombinieren* und *Daten kombinieren und transformieren*.



Es wird nicht empfohlen, den Befehl *Kombinieren und laden* zu verwenden, der alle Dateien kombiniert und das angehängte Ergebnis in den Bericht lädt ohne die Möglichkeit, es zu bearbeiten. Wenn Sie nicht sicher sind, welche Dateien kombiniert werden sollen, wählen Sie im Dialogfeld, das in Schritt 3 beschrieben wird, die Option *Daten transformieren* aus. Der Power Query-Editor wird mit einer Vorschau des Ordners geöffnet und ermöglicht es Ihnen, Filter auf die Dateien anzuwenden, die Sie kombinieren möchten. Wenn sich zum Beispiel andere Dokumenttypen im Ordner befinden, können Sie die Dateien mit der Erweiterung *.xlsx* filtern, um nur die Excel-Dateien zu kombinieren und die anderen Typen zu ignorieren. Anschließend können Sie die Dateien im Power Query-Editor kombinieren, indem Sie in der Überschrift der Spalte *Content* die kleine Schaltfläche *Dateien kombinieren* anklicken.

4. Markieren Sie im Dialogfeld *Dateien kombinieren* (siehe Abbildung 3-6) *Tabelle1* und klicken Sie auf *OK*.



Das Dialogfeld *Dateien kombinieren* entspricht dem Dialogfeld *Navigator*, das beim Laden einer einzelnen Excel-Arbeitsmappe geöffnet wird. Im Dialogfeld *Navigator* können Sie das Tabellenblatt oder die Tabelle auswählen, das/die bearbeitet oder in den Bericht geladen werden soll; im Dialogfeld *Dateien kombinieren* können Sie das Tabellenblatt oder die Tabelle auswählen, das/die aus allen im Ordner vorhandenen Dateien verwendet werden soll. Sie müssen Ihre Auswahl auf eine Beispielarbeitsmappe und ein Tabellenblatt anwenden; diese Auswahl wird dann für alle anderen Arbeitsmappen im Ordner verwendet.

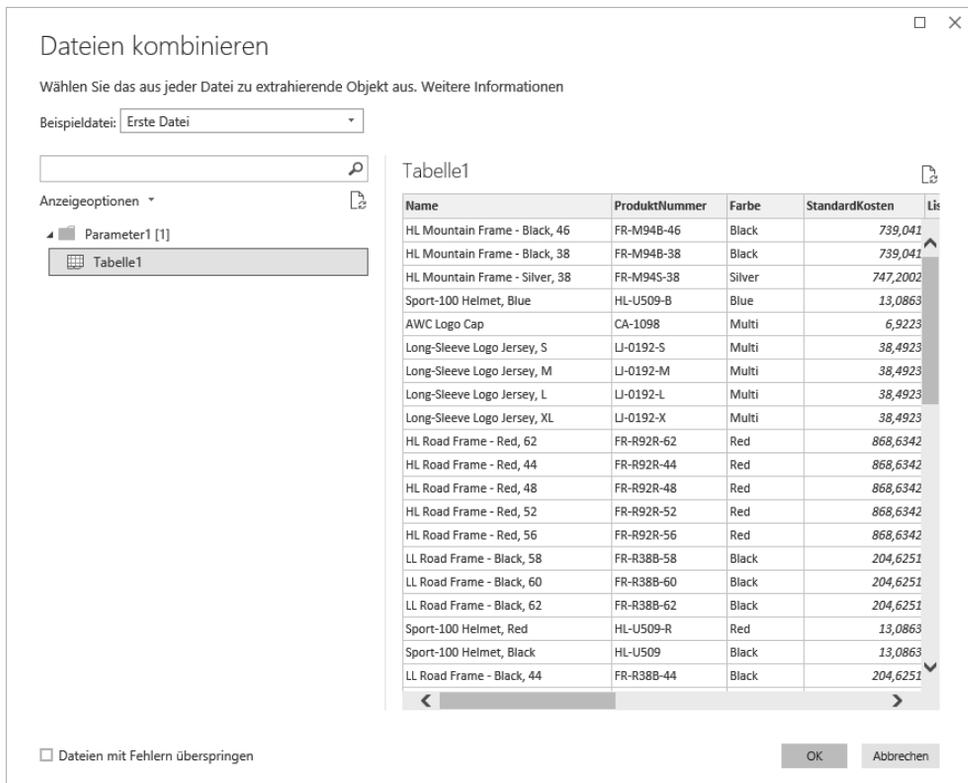


Abbildung 3-6 Im Dialogfeld *Dateien kombinieren* können Sie das zu kombinierende Tabellenblatt oder die zu kombinierende Tabelle auswählen. Nachdem Sie die Auswahl vorgenommen haben, werden alle Tabellenblätter beziehungsweise Tabellen mit demselben Namen aus jeder Excel-Arbeitsmappe im Ordner kombiniert.

5. Werfen Sie einen Blick in den Vorschaubereich des Power Query-Editors, der nun geöffnet wird. In der ersten Spalte, *Source.Name*, befinden sich die Dateinamen. Klicken Sie das Filtersteuerelement in der Überschrift *Source.Name* an. Im Filterbereich sehen Sie, dass die Werte in dieser Spalte mit den Dateinamen im Ordner identisch sind.

Die Spalte *Source.Name* kann manchmal sehr wertvoll sein. Im vorherigen Beispiel enthält die Spalte *VerkaufsstartDatum* das Datum, ab dem das Produkt bei AdventureWorks verkauft wird. Das Veröffentlichungsjahr kann aus dieser Spalte extrahiert werden, wenn es im Bericht benötigt wird. In vielen Fällen gibt es jedoch in den Quelldaten keine solche Spalte und diese kontextbezogenen Informationen sind nur im Dateinamen verfügbar. Glücklicherweise können Sie diese kontextbezogenen Informationen aus *Source.Name* extrahieren.

Um aus dem Dateinamen *K03U03 - 2017.xlsx* in *Source.Name* das Jahr zu extrahieren, können Sie verschiedene Transformationstechniken anwenden. Kapitel 2 enthält eine ausführliche exemplarische Vorgehensweise zu ähnlichen Szenarien, daher weisen wir in diesem Zusammenhang nur kurz auf einige hin. Die Dateinamen in diesem Fall enthalten das Präfix *K03U03* - und das Suffix *.xlsx*, und dazwischen befindet sich das Jahr. Sie können das Jahr aus *Source.Name* auf zwei Arten extrahieren:

- Ersetzen Sie *K03U03* - durch eine leere Zeichenfolge und ersetzen Sie dann *.xlsx* ebenfalls durch eine leere Zeichenfolge. In Kapitel 2 haben Sie gelernt, wie Sie Werte ersetzen.
- Alternativ können Sie *Source.Name* ab dem 9. Zeichen von links teilen (indem Sie den Befehl *Spalte teilen | Nach Anzahl von Zeichen* verwenden) und dann die zweite Spalte ab dem Punkt teilen oder alternativ die zweite Spalte nach Position teilen und dann bei 4 Zeichen von rechts teilen. Entfernen Sie die erste und dritte Spalte und behalten Sie die zweite Spalte bei, die die extrahierten Jahre enthält.

Sie können nun die kombinierte Produkttabelle in Ihre Excel-Arbeitsmappe oder Ihren Power BI-Bericht laden. Sie können die Lösungsdateien *K03U03 - Lösung.xlsx* und *K03U03 - Lösung.pbix* einsehen, die in den Begleitdateien zu diesem Buch verfügbar sind.

Ein paar Gedanken zum Importieren aus Ordnern

Es gibt drei wichtige Punkte zu bedenken, wenn Sie mit der Option *Aus Ordner* Dateien kombinieren:

- Das Importieren mehrerer Tabellen aus einem Ordner ist sehr leistungsstark. Wenn eine neue Datei hinzugefügt wird, muss die zuvor erstellte Abfrage nicht aktualisiert werden, um die neue Datei in die Anfügetransformation zu integrieren. Um dies zu überprüfen, können Sie einen kleinen Test ausführen. Kopieren Sie eine der vorhandenen Quellarbeitsmappen und fügen Sie diese unter dem Namen *K03U03 - 2018.xlsx* in den gleichen Ordner ein. Aktualisieren Sie nun die Arbeitsmappe oder den Power BI-Bericht, die/den Sie in Übung 3-3 erstellt haben, und bestätigen Sie, dass es neue Produkte für das Jahr 2018 gibt. Ist dies nicht eine bemerkenswerte Funktion?
- Nach Schritt 5 in Übung 3-3 haben Sie vielleicht bemerkt, dass der Bereich *Abfragen* neue Artefakte wie Gruppen und verschiedenen Arten von Abfragen enthält. Sie können diese Elemente vorerst ignorieren. Wir werden später darauf zurückkommen, wenn wir uns Abfragefunktionen und anderen fortgeschrittenen Themen widmen.
- Sie können mehrere Dateien aus freigegebenen Ordnern in einem Netzwerk oder auf SharePoint-Websites kombinieren. Mithilfe freigegebener Ordner können Sie eine Excel-Ar-

beitsmappe oder einen Power BI-Bericht für andere Berichtsaoren freigeben, die dann auch Aktualisierungen durchführen können. Wenn Sie Power BI verwenden, können Sie den Bericht im Power BI-Dienst veröffentlichen und so die Vorteile von geplanten Aktualisierungen genießen. Wenn sich der Ordner in SharePoint Online oder OneDrive for Business befindet, hat der Power BI-Dienst direkten Zugriff auf die Daten und kann so konfiguriert werden, dass er regelmäßig aktualisiert wird. Wenn der freigegebene Ordner lokal ist, können Sie das lokale Power BI Gateway verwenden, sodass der Power BI-Dienst Zugriff auf den Ordner hat, um regelmäßige Aktualisierungen durchzuführen.

Tabellenblätter aus einer Arbeitsmappe anfügen

Sehen wir uns ein letztes Szenario für das Anfügen von Tabellen im großen Maßstab an: das Kombinieren von Tabellen aus mehreren Tabellenblättern. Im vorherigen Beispiel haben Sie AdventureWorks-Produkte kombiniert, die in unterschiedliche Arbeitsmappen aufgeteilt waren. Jede Arbeitsmappe hatte ihre eigenen Produkte und die Arbeitsmappen waren anhand des Veröffentlichungsjahrs organisiert. Stellen Sie sich jedoch vor, dass dieselben Tabellen in separaten Tabellenblättern in derselben Excel-Arbeitsmappe gespeichert werden. Können Sie alle Tabellenblätter in einer einzigen Tabelle anfügen, wobei der Kontext des Verkaufsstarts für jedes Produkt beibehalten wird, wie in Abbildung 3-7 gezeigt? Können Sie auch sicherstellen, dass Sie die Abfrage nicht jedes Jahr aktualisieren müssen, wenn ein neues Tabellenblatt hinzugefügt wird? Wie können Sie sicherstellen, dass die Produkte des nächsten Jahres automatisch in Ihrem Bericht kombiniert werden, wenn Sie auf *Aktualisieren* klicken?

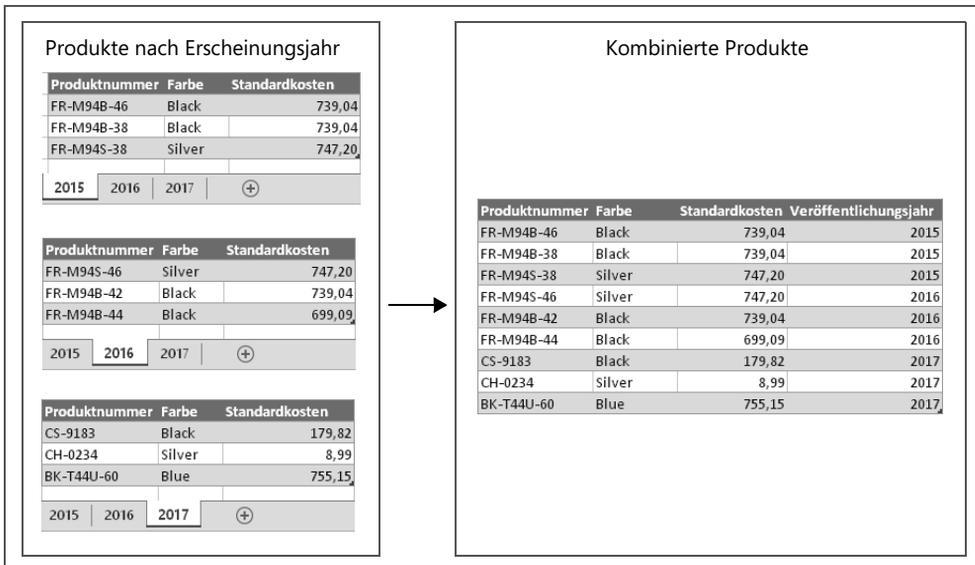


Abbildung 3-7 Ihr Ziel für die AdventureWorks-Produkte, die nach ihrem Veröffentlichungsjahr in separate Tabellenblätter unterteilt sind, besteht darin, sie zu kombinieren, gleichzeitig den Kontext des Veröffentlichungsjahrs zu erhalten und außerdem für zukünftige Jahre weitere Änderungen an der Abfrage zu vermeiden.

Übung 3-4: Tabellenblätter anfügen – Die Lösung

Um die Schritte für die Lösung dieser Herausforderung selbst durchzuführen, können Sie die Datei *K03U04 - Jahr pro Arbeitsblatt.xlsx* aus den Begleitdateien zu diesem Buch verwenden.

1. Öffnen Sie eine leere neue Excel-Arbeitsmappe oder einen neuen Power BI Desktop-Bericht.
In Excel: Wählen Sie auf der Registerkarte *Daten* den Befehl *Daten abrufen | Aus Datei | Aus Arbeitsmappe*.
In Power BI Desktop: Klicken Sie auf der Registerkarte *Home* auf *Daten abrufen*, wählen Sie *Datei | Excel* und klicken Sie dann auf *Verbinden*.
2. Wählen Sie die Datei *K03U04 - Jahr pro Arbeitsblatt.xlsx* aus und klicken Sie auf *Importieren*.
3. Markieren Sie im Dialogfeld *Navigator* die Zeile mit dem Ordnersymbol (in diesem Fall *K03U04 - Jahr pro Arbeitsblatt.xlsx*).



Wählen Sie im Dialogfeld *Navigator* keines der Tabellenblätter aus. Obwohl Sie dies tun und dann die Tabellenblätter wie in Übung 3-1 gezeigt anfügen können, führt diese Lösung langfristig zu unerwünschten Ergebnissen. Wenn neue Tabellenblätter in die Arbeitsmappe eingefügt werden, müssen Sie die Abfragen ändern und die neuen Tabellenblätter hinzufügen. Stattdessen haben Sie in Schritt 3 die gesamte Arbeitsmappe ausgewählt und können so eine robuste Lösung erstellen.

In Excel: Klicken Sie auf *Daten transformieren*.

In Power BI Desktop: Falls die Schaltfläche *Daten transformieren* deaktiviert ist, klicken Sie *K03U04 - Jahr pro Arbeitsblatt.xlsx* mit der rechten Maustaste an und klicken Sie dann im Kontextmenü auf *Daten transformieren*.

4. Wenn der Power Query-Editor geöffnet wird, benennen Sie die Abfrage in **Produkte** um. Im Vorschaubereich sehen Sie eine Tabelle, in der jede der Zeilen ein Tabellenblatt enthält. Der tatsächliche Inhalt jedes Tabellenblatts wird in der Spalte *Data* gekapselt. Bevor Sie die Tabellenblätter kombinieren, können Sie die folgenden drei Spalten entfernen: *Item*, *Kind* und *Hidden*.



Wenn sich in Ihrer Quellarbeitsmappe ausgeblendete Tabellenblätter oder bestimmte Tabellenblätter mit Daten befinden, die Sie nicht verwenden wollen, können Sie diese herausfiltern, bevor Sie in Schritt 5 die Tabellenblätter kombinieren. Dies ist ein guter Zeitpunkt, um die erforderlichen Filter anzuwenden. Sie können zum Beispiel einen Filter auf die Spalte *Hidden* anwenden, um nur die Zeilen beizubehalten, die den Wert *FALSE* enthalten; auf diese Weise können Sie ausgeblendete Tabellenblätter ausschließen.

-
5. Klicken Sie in der Überschriftenzeile der Spalte *Data* auf die Schaltfläche *Erweitern*.

6. Wenn der Bereich *Erweitern* geöffnet wird, in dem die Spalten angezeigt werden, die Sie erweitern können, klicken Sie auf *OK*, um alle Spalten zu erweitern. Wenn Sie die Spalte *Data* erweitern, enthält die kombinierte Tabelle bedeutungslose Überschriften (*Data.Column1*, *Data.Column2* und so weiter). Außerdem enthält die Tabelle sowohl die Überschriftenzeilen als auch die Datenzeilen aus jedem Tabellenblatt, wie in der ersten Tabelle in Abbildung 3-8 gezeigt.



Wenn es sich bei den Produkttabellen in jedem Tabellenblatt um Tabellenobjekte handeln würde, können Sie in Schritt 6 die tatsächlichen Produktspaltennamen erweitern, anstatt *Data.Column1*, *Data.Column2* und so weiter zu verwenden. (Um in Excel einen Zellbereich in eine Tabelle zu konvertieren, wählen Sie den Bereich oder eine einzelne Zelle in einem Bereich aus und klicken Sie dann auf der Registerkarte *Einfügen* auf *Tabelle* oder drücken Sie **[Strg]+[T]**.)

Das Transformationsflussdiagramm in Abbildung 3-8 zeigt die drei Transformationsschritte, die zum Bereinigen der kombinierten Tabelle erforderlich sind. Im ersten Schritt verschieben Sie die erste Zeile, die die Überschriftenzeile der Tabelle 2015 war, nach oben, damit diese Zeile die Überschriftenzeile der kombinierten Tabelle wird. Im nächsten Schritt entfernen Sie dann alle Überschriftenzeilen der Tabellenblätter 2016 und 2017. Schließlich benennen Sie im dritten Schritt die Spalte 2015 um, damit diese das Veröffentlichungsjahr jedes Produkts enthält. Ich könnte im gesamten Buch diese Art von Diagramm verwenden, um die Abfolge der wichtigsten Transformationsschritte darzustellen und um die jeweilige Zielsetzung für die Lösung fortgeschrittener Datenvorbereitungsprobleme zu beschreiben. Als Nächstes betrachten wir die Schritte im Detail.

7. Wählen Sie auf der Registerkarte *Transformieren* den Befehl *Erste Zeile als Überschriften verwenden*. Auf diese Weise können Sie die Spaltenüberschriften aus dem ersten Tabellenblatt als Spaltenüberschriften für die kombinierte Tabelle beibehalten.
8. Beachten Sie im Vorschaubereich, dass die Spaltenüberschriften der Tabellenblätter 2016 und 2017 als Zeilen in der angefügten Tabelle enthalten sind. Diese Zeilen enthalten den Wert *Name* in der Spalte *Name* und sollten aus der Tabelle entfernt werden. Um diese Zeilen herauszufiltern, klicken Sie auf das Filtersteuerelement auf der linken Seite der Überschrift *Name*. Wenn der Filterbereich geöffnet wird, geben Sie **Name** in das Suchfeld ein und deaktivieren Sie das Kontrollkästchen *Name*. Klicken Sie dann auf *OK*.
9. Doppelklicken Sie auf die erste Spaltenüberschrift mit dem Namen 2015 und benennen Sie sie in **Veröffentlichungsjahr** um.



Die Schritte 7 bis 9 können auf lange Sicht fehlschlagen. Sie werden in Kürze erfahren, warum und wie Sie es lösen können.

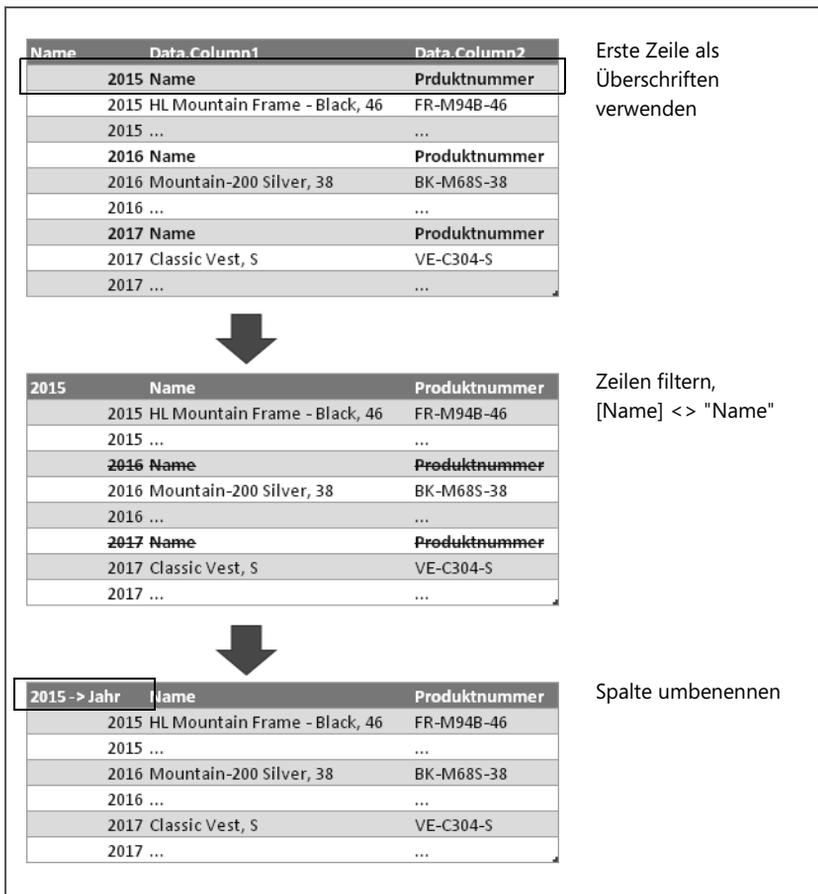


Abbildung 3-8 Dieses Transformationsflussdiagramm zeigt, wie Sie kombinierte Tabellenblätter bereinigen.

Sie können die kombinierte Tabelle jetzt in Ihre Excel-Arbeitsmappe oder Ihren Power BI-Bericht laden. Sie haben die drei Tabellenblätter für die drei Veröffentlichungsjahre kombiniert und das Veröffentlichungsjahr jedes Produkts aus den Tabellenblattnamen extrahiert. Es ist an der Zeit, die Skalierbarkeit der Lösung zu testen und zu prüfen, ob sie korrekt aktualisiert wird, wenn ein viertes Tabellenblatt mit den neu veröffentlichten Produkten hinzugefügt wird.

10. Während Ihre Excel-Arbeitsmappe oder Ihr Power BI-Bericht mit den kombinierten Produkten noch geöffnet ist, öffnen Sie die Excel-Quelldatei *K03U04 - Jahr pro Arbeitsblatt.xlsx*, die in dieser Übung verwendet wurde. Duplizieren Sie das Tabellenblatt *2017* als viertes Tabellenblatt (fügen Sie also die Kopie am Ende ein) und benennen Sie es in **2018** um.



Wenn Sie den Power Query-Editor in Excel (und nicht in Power BI Desktop) geöffnet haben, haben Sie möglicherweise bemerkt, dass Sie nicht auf andere geöffnete Arbeitsmappen zugreifen oder über die Excel-Registerkarte *Datei* weitere Arbeitsmappen öffnen können, wenn der Power Query-Editor geöffnet ist. Daher müssen Sie in Schritt 10 möglicherweise den Power Query-Editor schließen, bevor Sie die Quellarbeitsmappe öffnen. Um diese Einschränkung zu umgehen, können Sie über die Taskleiste oder das Startmenü eine neue Excel-Instanz starten und dort *KO3U04 - Jahr pro Arbeitsblatt.xlsx* öffnen. Die neue Excel-Instanz wird vom aktuellen Power Query-Editor-Fenster nicht blockiert.

11. Aktualisieren Sie die Abfrage *Produkte* in Ihrer Excel-Arbeitsmappe oder Ihrem Power BI-Bericht und überprüfen Sie, ob die duplizierten Produkte aus dem Tabellenblatt *2018* jetzt mit dem Wert 2018 als Veröffentlichungsdatum angefügt wurden. Sie sollten feststellen, dass es funktioniert hat!

Fügen Sie als erstes Tabellenblatt ein Blatt mit historischen Daten für *Produkte* hinzu, die 2014 veröffentlicht wurden, und nennen Sie dieses Blatt **2014**. Dazu können Sie eines der vorhandenen Tabellenblätter duplizieren und es als erstes Tabellenblatt in *KO3U04 - Jahr pro Arbeitsblatt.xlsx* hinzufügen. Wenn das erste, also am weitesten links stehende Tabellenblatt *2014* und nicht *2015* ist, schlägt die Aktualisierung mit diesem Fehler fehl: *Expression.Error: Die Spalte '2015' der Tabelle wurde nicht gefunden*.

12. Um diesen Fehler kurzfristig zu beheben, starten Sie den Power Query-Editor, wählen Sie die Abfrage *Produkte* aus und markieren Sie in *Angewendete Schritten* den Schritt *Geänderter Typ*. In der Bearbeitungsleiste wird die folgende Formel angezeigt:

```
= Table.TransformColumnTypes("#Höher gestufte Header",{{"2015", Int64.Type}, {"Name", type text}, {"Produktnummer", type text}, {"Farbe", type text}, {"Standardkosten", type any}, {"Listenpreis", type any}, {"Größe", type text}, {"Gewicht", type any}, {"ÜbergeordneteProduktkategorieName", type text}, {"ProduktkategorieName", type text}})
```

Suchen Sie in obiger Formel nach *2015* und ersetzen Sie diesen Text durch **2014**.

13. Wählen Sie im Abschnitt *Angewendete Schritte* den Schritt *Umbenannte Spalten* aus. In der Bearbeitungsleiste wird die folgende Formel angezeigt:

```
= Table.RenameColumns("#Gefilterte Zeilen",{{"2015", "Veröffentlichungsjahr"}})
```

Suchen Sie in obiger Formel nach dem Text *2015* und ersetzen Sie diesen durch **2014**.

14. Schließen Sie den Power Query-Editor und aktualisieren Sie den Bericht. Die in 2014 veröffentlichten Produkte werden nun korrekt mit Ihrem Bericht kombiniert.

Sie können sich diese Schritte in den Lösungsdateien *KO3U04 - Lösung 01.xlsx* oder *KO3U04 - Lösung 01.pbix* ansehen, die Sie in den Begleitdateien zu diesem Buch finden.

Wenn Sie davon ausgehen können, dass das erste Tabellenblatt immer *2015* oder jetzt, nach der Korrektur, *2014* sein wird, und wenn Sie sicherstellen können, dass neue Produkte immer als neue Tabellenblätter auf der rechten Seite hinzugefügt werden, dann haben Sie Ihre Mission erfüllt. Aber was tun Sie, wenn Sie nicht davon ausgehen können, dass diese Annahmen