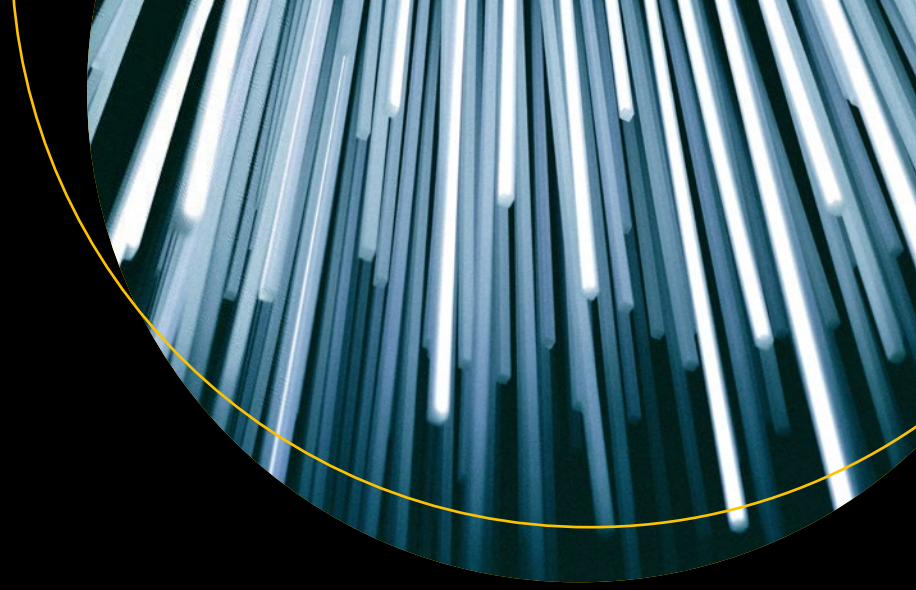


Pro Data Mashup for Power BI

Powering Up with Power Query and the M Language to Find, Load, and Transform Data

—

Adam Aspin

Apress®

Pro Data Mashup for Power BI

Powering Up with Power Query and the M Language to Find, Load, and Transform Data

Adam Aspin

Apress®

Pro Data Mashup for Power BI: Powering Up with Power Query and the M Language to Find, Load, and Transform Data

Adam Aspin
STAFFORD, UK

ISBN-13 (pbk): 978-1-4842-8577-0 ISBN-13 (electronic): 978-1-4842-8578-7
https://doi.org/10.1007/978-1-4842-8578-7

Copyright © 2022 by Adam Aspin

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Jonathan Gennick
Development Editor: Laura Berendson
Coordinating Editor: Jill Balzano

Cover photo by Christopher Burns on Unsplash

Distributed to the book trade worldwide by Springer Science+Business Media New York, 1 New York Plaza, Suite 4600, New York, NY 10004-1562, USA. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail booktranslations@springernature.com; for reprint, paperback, or audio rights, please e-mail bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at http://www.apress.com/bulk-sales.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub via the book's product page, located at www.apress.com. For more detailed information, please visit http://www.apress.com/source-code.

Printed on acid-free paper

Table of Contents

About the Author xiii

About the Technical Reviewer xv

Acknowledgments xvii

Introduction xix

Chapter 1: Discovering and Loading Data with Power BI Desktop 1

The Data Load Process 2

Understanding Data Load 7

The Navigator Dialog 8

Searching for Usable Data 10

Display Options 11

Refresh 11

The Navigator Data Preview 12

Modifying Data 13

Data Sources 14

The Power BI Desktop Screen 16

The Get Data Dialog 16

Conclusion 19

Chapter 2: Discovering and Loading File-Based Data with Power BI Desktop 21

File Sources 21

Loading Data from Files 23

CSV Files 23

Text Files 29

Text and CSV Options 31

Simple XML Files 32

Excel Files 34
Microsoft Access Databases 37
PDF Files 39
JSON Files 41
Conclusion 41

Chapter 3: Loading Data from Databases and Data Warehouses 43

Relational Databases 44
SQL Server 47
Automatically Loading Related Tables 52
Database Options 52
Searching for Tables 56
Oracle Databases 63
Other Relational Databases 68
Microsoft SQL Server Analysis Services Data Sources 71
Analysis Services Cube Tools 74
SSAS Tabular Data Warehouses 77
Import or Connect Live/DirectQuery 80
Other Database Connections 80
Conclusion 80

Chapter 4: DirectQuery and Connect Live 83

DirectQuery and Connect Live 83
Microsoft SQL Server Data 86
SQL Server Analysis Services Dimensional Data 93
Microsoft SQL Server Analysis Services Tabular Data Sources 96
DirectQuery with Non-Microsoft Databases 99
DirectQuery and In-Memory Tables 100
DirectQuery and Refreshing the Data 101
DirectQuery Optimization 101

Modifying Connections 102
Changing Permissions 105
Conclusion 108

Chapter 5: Loading Data from the Web and the Cloud 109

Web and Cloud Services 110
Web Pages 110
Online Services 110
Microsoft Azure 112
Web Pages 113
Advanced Web Options 116
Table View or Web View 117
Salesforce 118
Loading Data from Salesforce Objects 118
Salesforce Reports 125
Microsoft Dynamics 365 126
Google Analytics 131
OData Feeds 134
OData Options 135
Azure SQL Database 136
Azure Synapse Analytics 142
Connecting to SQL Server on an Azure Virtual Machine 145
Azure Blob Storage 147
Azure Databricks 151
Azure Security 154
Conclusion 154

Chapter 6: Loading Data from Other Data Sources 155

Other Sources 155
Power BI Datasets 156
Power BI Dataflows 158

R Scripts 161
R Options 165
Python Scripts to Load Data 166
Python Options 169
Dataverse 170
ODBC Sources 171
Refreshing Data 181
Refreshing the Entire Data in the Power BI Desktop In-Memory Model 181
Refreshing an Individual Table 182
Adding Your Own Data 183
Conclusion 185

Chapter 7: Power Query 187

Power BI Desktop Queries 188
Editing Data After a Data Load 189
Transforming Data Before Loading 192
Transform or Load? 194
Power Query 195
The Applied Steps List 197
Power Query Ribbons 197
The View Ribbon 205
The Tools Ribbon 207
The Help Ribbon 208
Viewing a Full Record 208
Power Query Context Menus 209
Conclusion 211

Chapter 8: Structuring Data 213

Dataset Shaping 214
Renaming Columns 215
Reordering Columns 215
Removing Columns 217

Choosing Columns 218
Merging Columns 220
Going to a Specific Column 223
Removing Records 225
Rows 226
Removing Duplicate Records 230
Sorting Data 231
Reversing the Row Order 233
Undoing a Sort Operation 233
Filtering Data 234
Selecting Specific Values 235
Finding Elements in the Filter List 236
Filtering Text Ranges 238
Filtering Numeric Ranges 238
Filtering Date and Time Ranges 239
Filtering Data 241
Applying Advanced Filters 243
Excluding Rows Where a Value Is Missing 245
Grouping Records 246
Simple Groups 247
Complex Groups 249
Saving Changes in Power Query 252
Exiting Power Query 252
Conclusion 253

Chapter 9: Shaping Data 255

Merging Data 256
Adding Data 256
Aggregating Data During a Merge Operation 261
Types of Join 265
Joining on Multiple Columns 266
Fuzzy Matching 268

Fuzzy Matching Options 269
Merge As New Query 270
Preparing Datasets for Joins 271
Correct and Incorrect Joins 271
Examining Joined Data 272
Appending Data 274
Adding the Contents of One Query to Another 274
Appending the Contents of Multiple Queries 277
Changing the Data Structure 279
Unpivoting Tables 279
Pivoting Tables 282
Transposing Rows and Columns 284
Data Quality Analysis 284
Column Quality 285
Column Distribution 285
Column Profile 286
Profiling the Entire Dataset 288
Correcting Anomalies 289
Data Transformation Approaches 291
Conclusion 291

Chapter 10: Data Cleansing 293

Using the First Row As Headers 294
Changing Data Type 295
Detecting Data Types 298
Data Type Indicators 298
Switching Data Types 300
Data Type Using Locale 301
Replacing Values 302
Transforming Column Contents 305
Text Transformation 305
Adding a Prefix or a Suffix 307

Removing Leading and Trailing Spaces 307
Removing Nonprinting Characters 308
Number Transformations 309
Conclusion 319

Chapter 11: Data Transformation 321

Filling Down Empty Cells 322
Extracting Part of a Column's Contents 326
Advanced Extract Options 327
Duplicating Columns 330
Splitting Columns 330
Splitting Column by a Delimiter 331
Advanced Options for Delimiter Split 334
Splitting Columns by Number of Characters 335
Splitting Columns by Character Switch 337
Merging Columns 337
Creating Columns from Examples 339
Adding Conditional Columns 341
Index Columns 344
Conclusion 345

Chapter 12: Complex Data Structures 347

Adding Multiple Files from a Source Folder 348
Filtering Source Files in a Folder 351
Displaying and Filtering File Attributes 355
The List Tools Transform Ribbon 357
Parsing XML Data from a Column 358
Parsing JSON Data from a Column 361
Complex JSON Files 363
Complex XML Files 365
Python and R Scripts 369
Using Python Scripts to Modify Data 370

Using R Scripts to Modify Data 373
Convert a Column to a List 375
Query Folding 376
Reusing Data Sources 380
Pinning a Data Source 383
Copying Data from Power Query 384
Conclusion 385

Chapter 13: Organizing, Managing, and Parameterizing Queries 387

Managing the Transformation Process 388
Modifying a Step 389
Renaming a Step 390
Deleting a Step or a Series of Steps 390
Modifying an Existing Step 392
Adding a Step 395
Altering Process Step Sequencing 395
An Approach to Sequencing 395
Error Records 397
Removing Errors 397
Managing Queries 398
Organizing Queries 398
Grouping Queries 399
Duplicating Queries 402
Referencing Queries 402
Documenting Queries 405
Adding a Column As a New Query 406
Enabling Data Load 408
Enabling Report Refresh 409
Pending Changes 410
Parameterizing Queries 410
Creating a Simple Parameter 411
Creating a Set of Parameter Values 413

Creating a Query-Based Parameter 416
Modifying a Parameter 419
Applying a Parameter When Filtering Records 420
Modifying the Current Value of a Parameter 422
Applying a Parameter in a Search and Replace 423
Applying a Parameter to a Data Source 424
Other Uses for Parameters 427
Using Parameters in the Data Source Step 428
Applying a Parameter to a SQL Query 429
Query Icons 431
Power BI Templates with Parameters 431
Conclusion 432

Chapter 14: The M Language 433

What Is the M Language? 434
M and the Power Query Editor 435
Modifying the Code for a Step 436
M Expressions 438
Writing M by Adding Custom Columns 439
The Advanced Editor 441
Expressions in the Advanced Editor 442
The Let Statement 443
Modifying M in the Advanced Editor 444
Syntax Checking 445
Advanced Editor Options 446
Basic M Functions 447
Text Functions 448
M or DAX? 450
Number Functions 450
Date Functions 452
Time Functions 454

Duration Functions 454
M Concepts 455
M Data Types 456
M Values 458
Defining Your Own Variables in M 458
Writing M Queries 459
M Autocomplete 460
Lists 462
Creating Lists Manually 462
Generating Sequences Using Lists 464
Accessing Values from a List 465
List Functions 465
Records 466
Tables 468
Other Function Areas 470
Custom Functions in M 470
Adding Comments to M Code 472
Single-Line Comments 473
Multiline Comments 473
Conclusion 473

Appendix A: Sample Data 475

Sample Data 475
Downloading the Sample Data 475

Index 477

About the Author

Adam Aspin is an independent Business Intelligence consultant based in the United Kingdom. He has worked with SQL Server for over 25 years. During this time, he has developed several dozen reporting and analytical systems based on the Microsoft data and analytics product suite.

A graduate of Oxford University, Adam began his career in publishing before moving into IT. Databases soon became a passion, and his experience in this arena ranges from dBase to Oracle, and Access to MySQL, with occasional sorties into the world of DB2. He is, however, most at home in the Microsoft universe when using the Microsoft data platform—both in Azure and on-premises.

Business Intelligence has been Adam's principal focus for the last 20 years. He has applied his skills for a range of clients in finance, banking, utilities, leisure, luxury goods, and pharmaceuticals. Adam is a frequent contributor to SQLServerCentral.com and Simple Talk. He is a regular speaker at events such as Power BI user groups, SQL Saturdays, and SQLBits. A fluent French speaker, Adam has worked in France and Switzerland for many years.

Adam is the author of *SQL Server 2012 Data Integration Recipes*; *Business Intelligence with SQL Server Reporting Services*; *High Impact Data Visualization in Excel with Power View, 3D Maps, Get & Transform and Power BI*; *Data Mashup with Microsoft Excel Using Power Query and M*; and *Pro Power BI Theme Creation*—all with Apress.

About the Technical Reviewer

Ed Freeman has been a data engineer ever since graduating with a mathematics degree at UCL in 2017. Throughout his career, he has been implementing intelligent cloud and data platforms built on Azure for clients of all sizes and industries. Ed is particularly passionate about Power BI and Microsoft data platform services as a whole edits *Power BI Weekly*, a free weekly newsletter to help the community keep up with the latest and greatest developments from the Power BI ecosystem.

Acknowledgments

Writing a technical book can prove to be a daunting challenge. So I am all the more grateful for all the help and encouragement that I have received from so many friends and colleagues.

First, my heartfelt thanks go, once again, to Jonathan Gennick, the commissioning editor of this book. Throughout the publication process, Jonathan has been an exemplary mentor. He has shared his knowledge and experience selflessly and courteously and provided much valuable guidance.

My deepest gratitude goes yet again to the Apress coordinating team for managing this volume through the rocks and shoals of the publication process. Particular thanks go to Silembarasan Panneerselvam for making a potentially stress-filled trek into a pleasant and stress-free journey.

When delving into the arcane depths of technical products, it is all too easy to lose sight of the main objectives of a book. Fortunately, my good friend Ed Freeman, the technical reviewer, has worked unstintingly to help me retain focus on the objectives of this book. He has also shared his considerable experience of Power Query and the Microsoft Data Platform in the enterprise and has helped me immensely with his comments and suggestions.

Finally, my deepest gratitude has to be reserved for the two people who have given the most to this book. They are my wife and son, who have always encouraged me to persevere while providing all the support and encouragement that anyone could want. I am very lucky to have both of them.

Introduction

Analytics has become one of the buzzwords that define an age. Managers want their staff to deliver meaningful insight in seconds; users just want to do their jobs quickly and well. Everyone wants to produce clear, telling, and accurate analysis with tools that are intuitive and easy to use.

Microsoft recognized these trends and needs a few short years ago when they extended Excel with an add-in called Power Query. Once a mere optional extension to the world's leading spreadsheet, Power Query is now an integral part of Power BI–Microsoft's world-beating analytics application. It allows a user to take data from a wide range of sources and transform them into the base data that they can build on to add metrics, instant analyses, and KPIs to project their insights in eye-catching dashboards.

With Power Query, the era of self-service data access and transformation has finally arrived.

What Is Power Query?

Power Query is a tool that is used to carry out ETL. This acronym stands for **E**xtract, **T**ransform, **L**oad. This is the sequential process that covers:

- Connecting to source data and accessing all or part of the data that you need to bring into Power BI Desktop and Power Query. This is the *extract* phase of ETL.
- Reshaping the data (the "data mashup" process) so that the resulting data is in a form that can be used by Power BI Desktop. Essentially this means ensuring that the data is in a coherent, structured, and complete tabular format. This is the *transform* phase of ETL.
- Returning the data into a Power BI Desktop data model. This is the *load* phase of ETL.

These three phases make up the data ingestion process. So it is worth taking a short look at what makes up each one of them.

Connecting to Source Data

Gone are the days when you manually entered all the data you needed into a spreadsheet. Today's data are available in a multitude of locations and formats and are too voluminous to re-key.

This is where Power Query's ability to connect instantly to 40-odd standard data sources is simply invaluable. Is your accounting data in MS Dynamics? Just connect. Is your CRM data in Salesforce? Just connect. Is your organization using a Data Lake?... You can guess the reply.

Yet this is only a small part of what Power Query can do to help simplify your analyses. For not only can it connect to a multitude of data sources (many of which are outlined in Chapters 1 through 6), it does this via a unified interface that makes connecting to data sources brilliantly simple. On top of this, you can use Power Query to preview the source data and ensure that you are loading exactly what you need. Finally, to top it all, the same interface is used for just about all of the available source data connections. This means that once you have learned to set up *one* connection, you have learned how to *connect to virtually all of the available data sources*.

Data Transformation

Once you have established a connection to a data source, you may need to tweak the data in some way. Indeed, you may even need to reshape it entirely. This is the data mashup process, and it is the area where Power Query shines.

Power Query can carry out the simplest data transformation tasks to the most complex data restructuring challenges in a few clicks. You can:

- Filter source data so that you only load exactly the rows and columns you need
- Extend the source data with calculations or data extracted from existing columns of data
- Cleanse and rationalize the data easily and quickly in a multitude of ways
- Join or split source tables to prepare a logical set of data tables for each specific analytical requirement

- Group and aggregate source data to reduce the quantity of data loaded into Excel
- Prepare source data tables to become a usable data model

This list merely scratches the surface of all that Power Query can do to mash up your data. It is, without hyperbole, unbelievably powerful at transforming source data. Indeed, it can carry out data ingestion and transformation tasks that used to be the preserve of expensive products that required complex programming skills and powerful servers.

All of this can now be done using a code-free interface that assists you in taking the messiest source data and delivering it to Excel as limpid tables of information ready to work with.

The aim of this short book is to introduce the reader to this brave new world of user-driven data integration. This will not involve a complete tour of Power BI Desktop, however. The product is simply too vast for that. Consequently, this book concentrates on data mashup using Power BI Desktop and Power Query. If you need to learn further aspects of the Power BI Desktop ecosystem, then two companion volumes are available:

- *Pro DAX and Data Modeling in Power BI*
- *Pro Power BI Dashboard Creation*

The first explains how to meld disparate data sources into a unified Power BI data model that you then extend using DAX—the built-in analytics language. The second guides you through the process of dashboard creation using all the range of available visuals and techniques.

Although a basic knowledge of the MS Office suite will help, this book presumes that you have little or no knowledge of Power BI Desktop. This product is therefore explained from the ground up with the aim of providing the most complete coverage possible of the way data can be discovered and loaded into Power BI as the basis for user-driven dashboards. Hopefully, if you read the book and follow the examples given, you will arrive at a level of practical knowledge and confidence that you can subsequently apply to your own data ingestion requirements. This book should prove invaluable to business intelligence developers, MS Office power users, IT managers, and finance experts—indeed anyone who wants to deliver efficient and practical business intelligence to their colleagues. Whether your aim is to develop a proof of concept or to deliver a fully fledged BI system, this book can, hopefully, be your guide and mentor when it comes to ensuring that the data you base your analytics on is clean and well structured.

If you wish, you can read this book from start to finish as it is designed to be a progressive tutorial that will help you to learn Power Query. However, as Power Query is composed of four main areas, this book is broken down into four sets of chapters that focus on the various key areas of the product. It follows that you can, if you prefer, focus on individual topics in Power Query without having to take a linear approach to reading this book.

- Chapters 1 through 6 show you how to connect to a range of varied data sources and bring this data into Excel using Power Query. Depending on the source data that you need to use, you may only need to dip into parts of these chapters to find guidance on how to use a specific source data type.
- Chapters 7 through 11 explain how to transform and clean data so that you can use it for analysis. These data transformations range from the extremely simple to the potentially complex. Indeed, they are as potentially vast as data itself. You may never need to apply all of the extensive range of data modification and cleansing techniques that Power Query can deliver, but just about everything that it can do is explained in detail in these chapters.
- Chapters 12 and 13 explain how to tame the real world of data loading and transformation. Here, you will learn how to organize and manage your queries, as well as how to add parameters to make them more interactive and resilient.
- Chapter 14 introduces you to M—the language that Power Query uses to transform your data. Using M, you can push your data ingestion and transformation routines to new heights that are simply not possible using just the Power Query interface.

Inevitably, not every question can be answered, and not every issue can be resolved in one book. I truly hope that I have answered many of the essential Power Query questions that you will face when ingesting data into Power BI Desktop. Equally I hope that I have provided ways of solving a reasonable number of the challenges that you may encounter. I wish you good luck in using Power Query and Power BI Desktop to prepare and deliver your insights. And I sincerely hope that you have as much fun with it as I had writing this book.

CHAPTER 1

Discovering and Loading Data with Power BI Desktop

Before you can use Power BI Desktop to present any analysis or discover new insights, you need data. Your sources could be in many places and in many formats. Nonetheless, you need to access them, look at them, select them, and quite possibly restructure them or clean them up to some extent. You may also need to join many separate data sources before you shape the data into a coherent model that you can use as the foundation for your dashboards and reports. The amazing thing is that you can do all of this using Power BI Desktop without needing any other tools or utilities.

Discovering, loading, cleaning, and modifying source data are some of the many areas where Power BI Desktop really shines. It allows you to accomplish the following:

- *Data discovery*: Find and connect to a myriad of data sources containing potentially useful data. This can be from both public and private data sources. This is the subject of Chapters 1 through 6.
- *Data loading*: Select the data you have examined and load it into Power BI Desktop for shaping. This, too, is handled in Chapters 1 through 6.
- *Data modification*: Modify the structure of each table that you have imported and then filter and clean the data itself (we will look at this in detail in Chapters 7 through 13).
- *Programming data transformations*: This is described in Chapter 14.

© Adam Aspin 2022
A. Aspin, *Pro Data Mashup for Power BI*, https://doi.org/10.1007/978-1-4842-8578-7_1

- *Data shaping*: Join tables to create a clear, unified, and accessible data model. As this is a fairly complex subject, it is covered in the companion volume "Pro Dax and Data Modeling with Power BI Desktop".

Although I have outlined these five steps as if they are completely separate and sequential, the reality is that they often blend into a single process. Indeed, there could be many occasions when you will examine the data *after* it has been loaded into Power BI Desktop—or clean data *before* you load it. The core objective will, however, always remain the same: find some data and then load it into Power BI Desktop where you can tweak, clean, and shape it.

This process could be described simplistically as "First, catch your data." In the world of data warehousing, the specialists call it ETL, which is short for **E**xtract, **T**ransform, and **L**oad. Indeed, it is also known as ELT (**E**xtract, **L**oad, and **T**ransform), which can also be handled by Power BI Desktop. Despite the reassuring confidence that the acronym brings, this process is rarely a smooth, logical progression through a clear-cut series of steps. The reality is often far messier. You may often find yourself importing some data, cleaning it, importing some more data from another source, combining the second table with the first one, removing some rows and columns, and then repeating these operations, as well as many others, several times over.

In this book, I will show you how the process can work in practice using Power BI Desktop. I hope that this will make the various steps that comprise an ETL process clearer. All I am asking is that you remain aware that the range of options that Power BI Desktop includes make it a multifaceted and tremendously capable tool. The science is to know *which* options to use. The art is to know *when* to use them.

The Data Load Process

Let's begin with a rapid overview of what you need to do to get some data into Power BI Desktop. This requires you to follow the download instructions in Appendix A, to prepare the source files in the `C:\PowerBIDesktopSamples` folder.

Once you have launched Power BI Desktop, you are faced with the splash screen that looks something like the one that you can see in Figure 1-1.

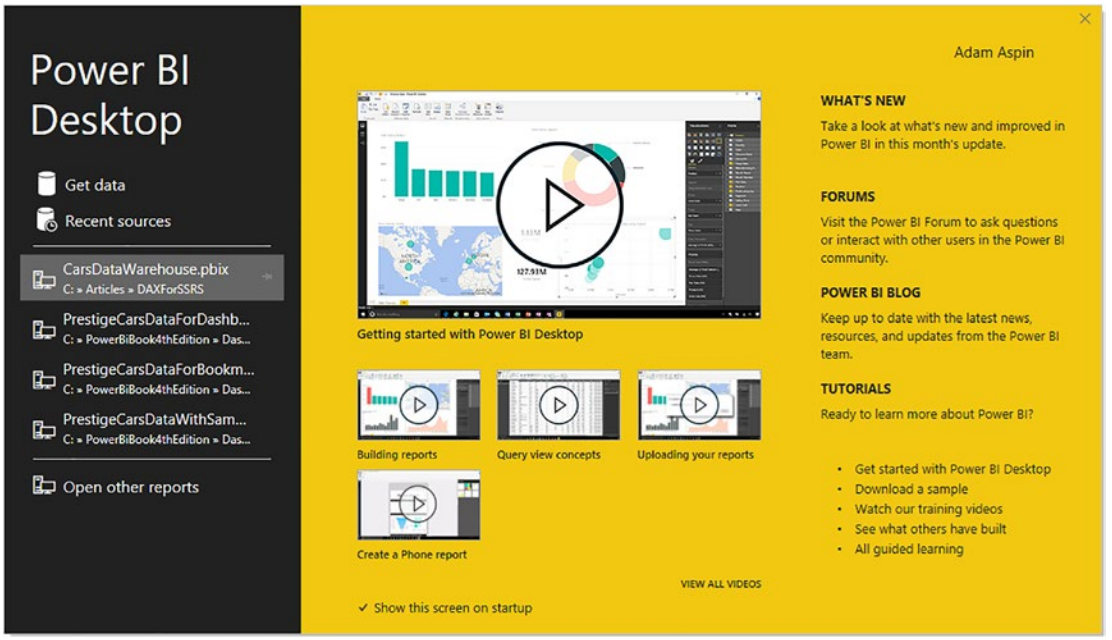


Figure 1-1. *The splash screen*

Given that you are working with an application that lives and breathes data, it is not really surprising that the first step in a new analytical challenge is to find and load some data. So the following explains what you have to do (assuming that you have downloaded the sample data that accompanies this book from the Apress website—as is explained in Appendix A):

1. Click Get data in the splash screen. The Get Data dialog will appear, as shown in Figure 1-2.

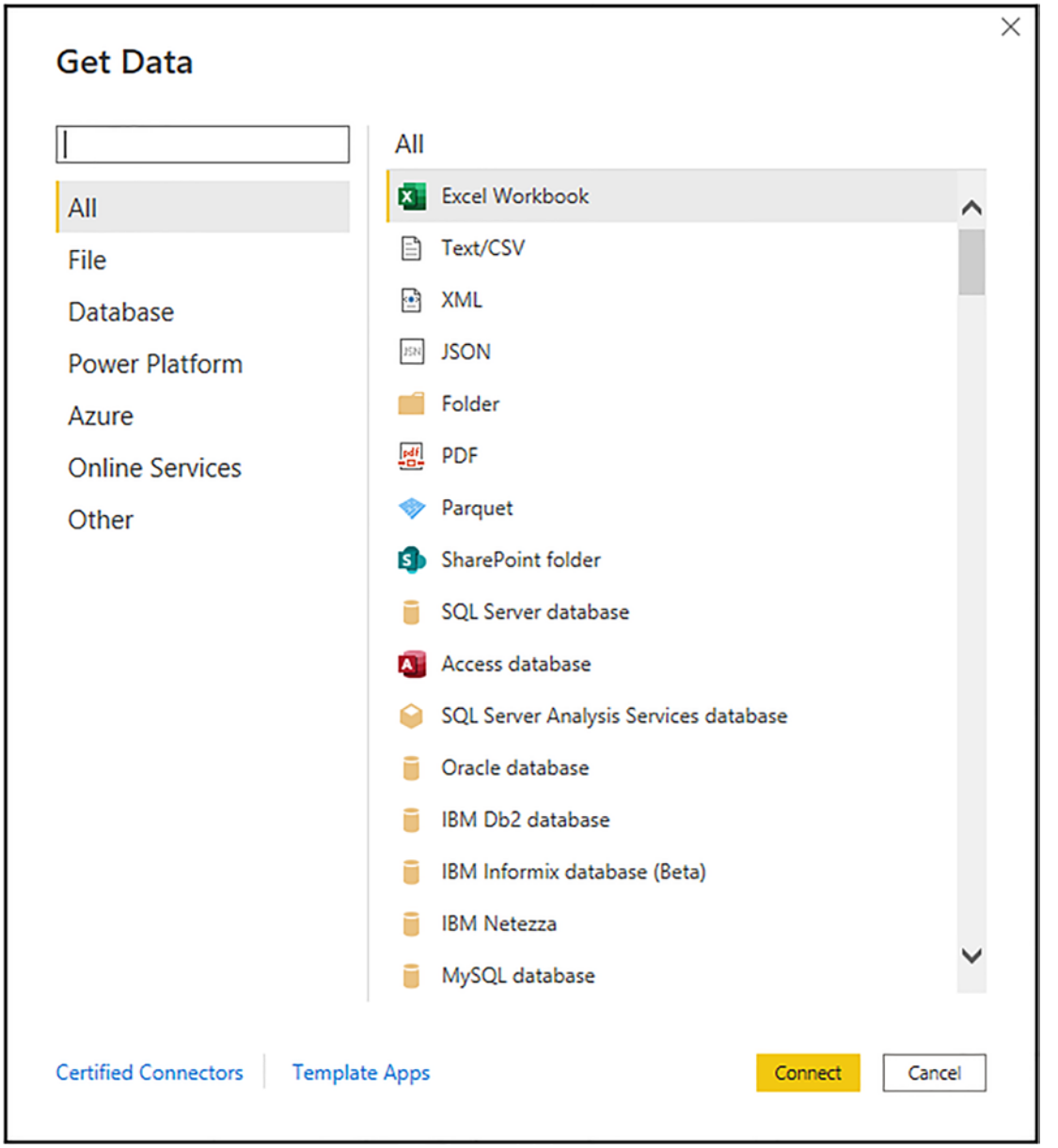


Figure 1-2. *The Get Data dialog*

2. In the list of all the possible data sources on the right of this dialog, select Excel Workbook, and then click Connect. The Windows Open File dialog will appear.

3. Click the file `C:\PowerBiDesktopSamples\BrilliantBritishCars.xlsx`. The Windows Open dialog will look like the one in Figure 1-3.

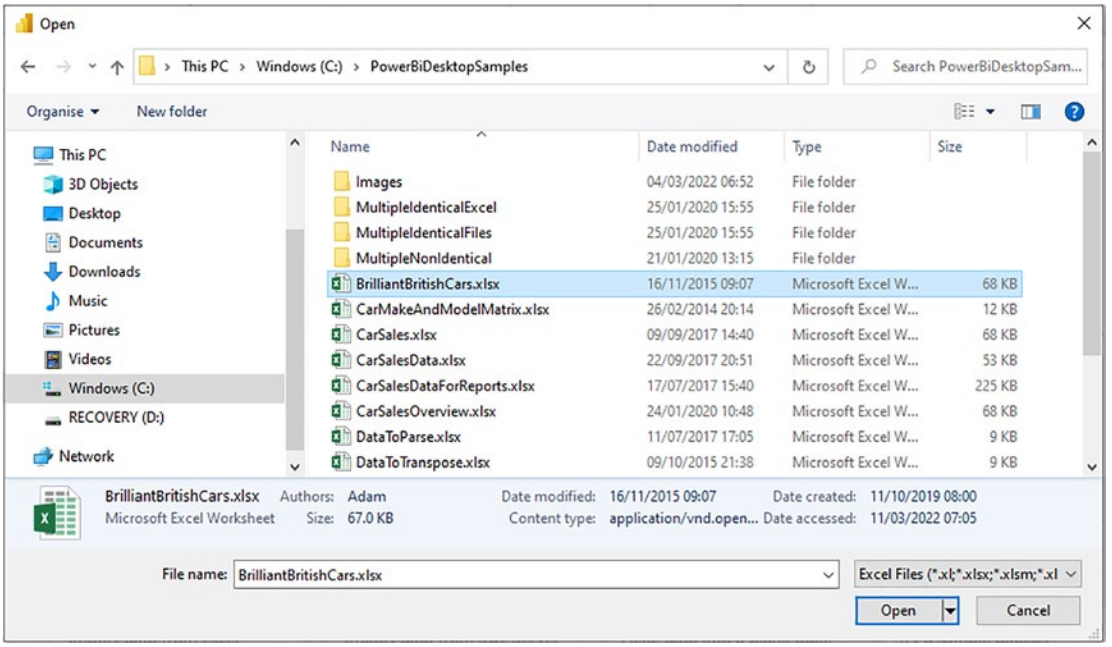


Figure 1-3. *The Windows Open File dialog when loading data from a file source*

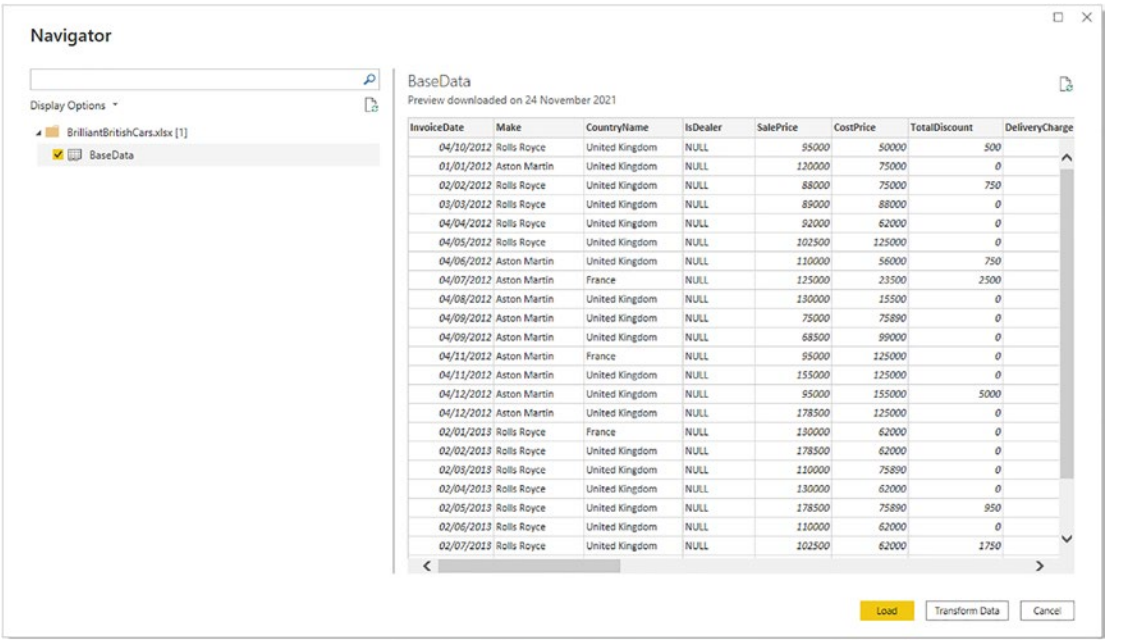
4. Click the Open button. The Connecting dialog will appear for a second or two (it looks like Figure 1-4), and then the Navigator dialog will appear.

Figure 1-4. *The Connecting dialog*

5. You will see that the BrilliantBritishCars.xlsx file appears on the left of the Navigator dialog and that any workbooks, named ranges, or data tables that it contains are listed under the file name.

6. Click the BaseData worksheet name that is on the left. The contents of this workbook will appear in the data pane on the right of the Navigator dialog.

7. Click the check box for the BaseData worksheet on the left. The Load and Transform Data buttons will be activated. The Navigator dialog should look like Figure 1-5.



InvoiceDate	Make	CountryName	IsDealer	SalePrice	CostPrice	TotalDiscount	DeliveryCharge
04/10/2012	Rolls Royce	United Kingdom	NULL	95000	50000	500	
01/01/2012	Aston Martin	United Kingdom	NULL	120000	75000	0	
02/02/2012	Rolls Royce	United Kingdom	NULL	88000	75000	750	
03/03/2012	Rolls Royce	United Kingdom	NULL	89000	88000	0	
04/04/2012	Rolls Royce	United Kingdom	NULL	92000	62000	0	
04/05/2012	Rolls Royce	United Kingdom	NULL	102500	125000	0	
04/06/2012	Aston Martin	United Kingdom	NULL	110000	56000	750	
04/07/2012	Aston Martin	France	NULL	125000	23500	2500	
04/08/2012	Aston Martin	United Kingdom	NULL	130000	15500	0	
04/09/2012	Aston Martin	United Kingdom	NULL	75000	75890	0	
04/09/2012	Aston Martin	United Kingdom	NULL	68500	99000	0	
04/11/2012	Aston Martin	France	NULL	95000	125000	0	
04/11/2012	Aston Martin	United Kingdom	NULL	155000	125000	0	
04/12/2012	Aston Martin	United Kingdom	NULL	95000	155000	5000	
04/12/2012	Aston Martin	United Kingdom	NULL	178500	125000	0	
02/01/2013	Rolls Royce	France	NULL	130000	62000	0	
02/02/2013	Rolls Royce	United Kingdom	NULL	178500	62000	0	
02/03/2013	Rolls Royce	United Kingdom	NULL	110000	75890	0	
02/04/2013	Rolls Royce	United Kingdom	NULL	130000	62000	0	
02/05/2013	Rolls Royce	United Kingdom	NULL	178500	75890	950	
02/06/2013	Rolls Royce	United Kingdom	NULL	110000	62000	0	
02/07/2013	Rolls Royce	United Kingdom	NULL	102500	62000	1750	

Figure 1-5. *The Navigator dialog with data selected*

8. Click Load. The data will be loaded from the Excel file into Power BI Desktop.

You will see the Power BI Desktop report window, like the one shown in Figure 1-6. This is the canvas where you will add visuals to create dashboards. If you expand the BaseData table in the Fields pane on the right of the Power BI Desktop application, you can see all the columns from the Excel worksheet are now fields.

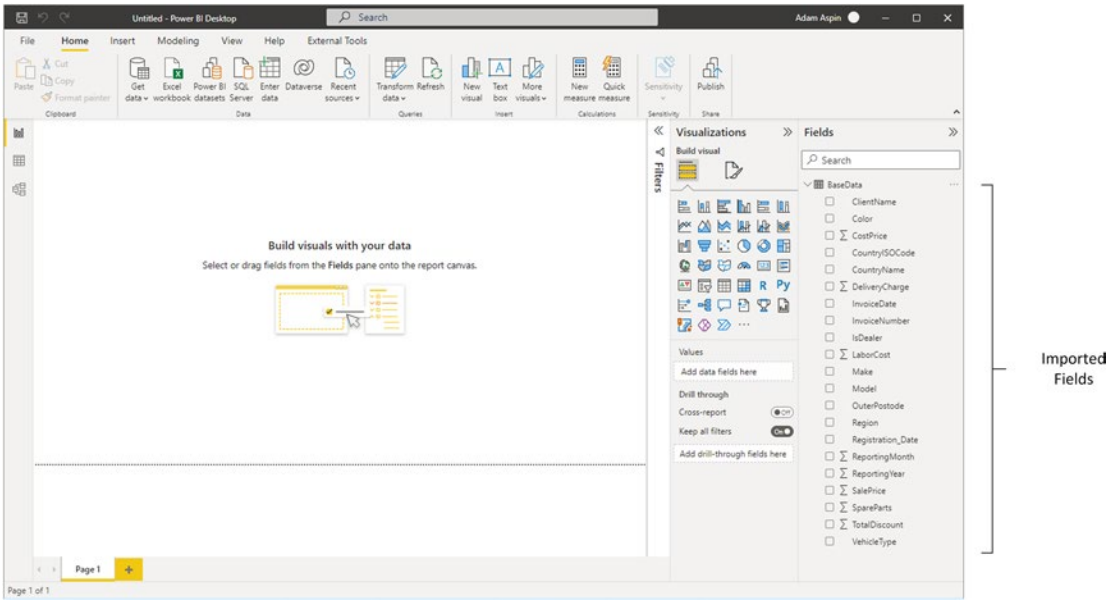


Figure 1-6. *Data available in Power BI Desktop*

I imagine that loading this data took under a minute. Yet you now have a fully operational data model in Power BI Desktop that is ready to feed data into your dashboard.

So if you are itching to race ahead and actually create a couple of tables, charts, and visuals, then you can always use the companion volume *Pro Power BI Dashboard Creation* to guide you through the process of building charts and dashboards.

However, as this book is all about getting and transforming data in Power BI Desktop, I would like to pause for an instant and explain exactly what you have seen so far.

Note The Get Data dialog contained a reference to certified connectors. These are developed by third parties but certified and distributed by Microsoft.

Understanding Data Load

What you have seen so far is an extremely rapid dash through a Power BI Desktop data load scenario. In reality, this process can range from the blindingly simple (as you just saw) to the more complex where you join, filter, and modify multiple tables from

different sources (as you will discover in Chapters 7 through 14). However, loading data will always be the first step in any data analysis scenario when you are using Power BI Desktop.

In this short example, you nonetheless saw many of the key elements of the data load process. These included

- Accessing data that is available in any of the source formats that Power BI Desktop can read
- Taking a first look at the data before loading it into Power BI Desktop

What you did not see here is how Power BI Desktop can transform the source data in Query Editor. This aspect of data manipulation is covered extensively in Chapters 7 through 12.

The Navigator Dialog

One key aspect of the data load process is using the Navigator dialog correctly. You saw this dialog in Figure 1-5. The Navigator window appears when connecting to many, but not all, data sources. It allows you to

- Take a quick look at the available data tables in the data source
- Look at the data in individual tables
- Select one or more data tables to load into Power BI Desktop

---

Note If you want to display more sample data, you can increase the height and width of the Navigator dialog. Alternatively, you can scroll through the window on the right of the Navigator.

Depending on the data source to which you have connected, you might see only a few data tables in the Navigator window, or hundreds of them. In any case, what you can see are the structured tables that Power BI Desktop can recognize and is confident that it can import. Equally dependent on the data source is the level of complexity of what you will see in the Navigator window. If you are looking at a database server, for instance, then you may start out with a list of databases and you may need to dig deeper into the arborescence of the data by expanding databases to list the available data tables and views.

The more you work with Power BI Desktop, the more you will use the Navigator dialog. So it seems appropriate to explain at this early juncture some of the tricks and techniques that you can apply to make your life easier when delving into potential sources of data.

Let's start by taking a closer look at the available options. These are highlighted in Figure 1-7.

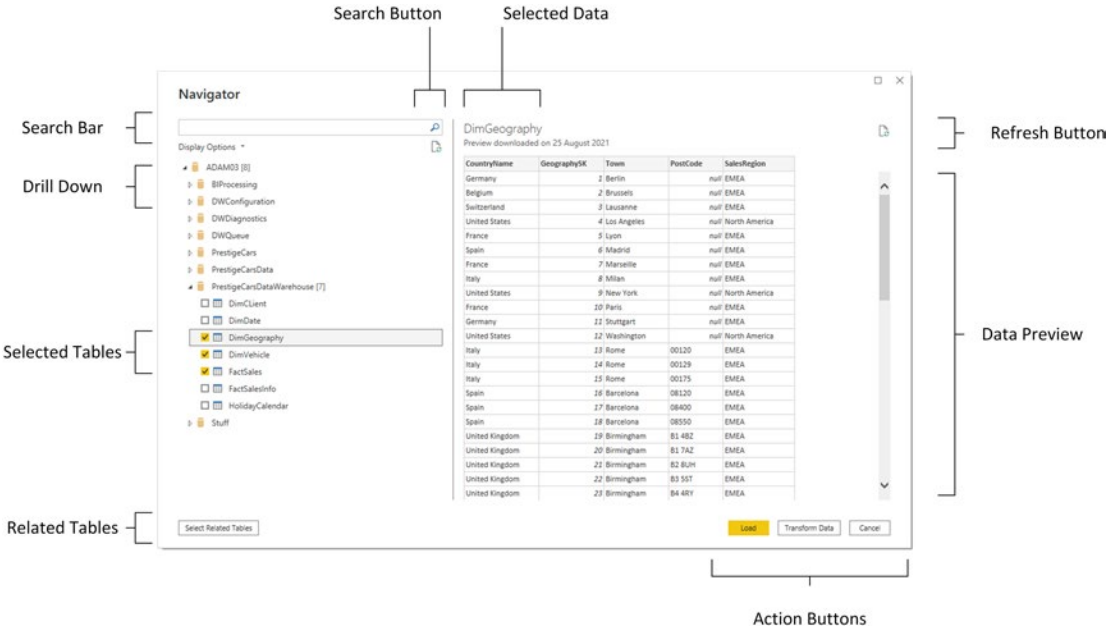


Figure 1-7. *The Navigator dialog*

The Navigator dialog is essentially in two parts:

- *On the left*: The hierarchy of available data sources. These can consist of a single table or multiple tables, possibly organized into one or many folders.
- *On the right*: A preview of the data in the selected element.

The various Navigator dialog options are explained in the following sections.

Searching for Usable Data

You will, inevitably, come across cases where the data source that you are connecting to will contain hundreds of potential datasets that you could use. These datasets are generally referred to as tables. This is especially true for databases. Fortunately, Power BI Desktop lets you filter the tables that are displayed extremely easily.

1. In the Navigator dialog, click inside the Search box.
2. Enter a part of a table name that you want to isolate.
3. Click the magnifying glass icon at the right of the Search box. The list of tables will be filtered to show only tables containing the text that you entered. You can see an example of this in Figure 1-8.

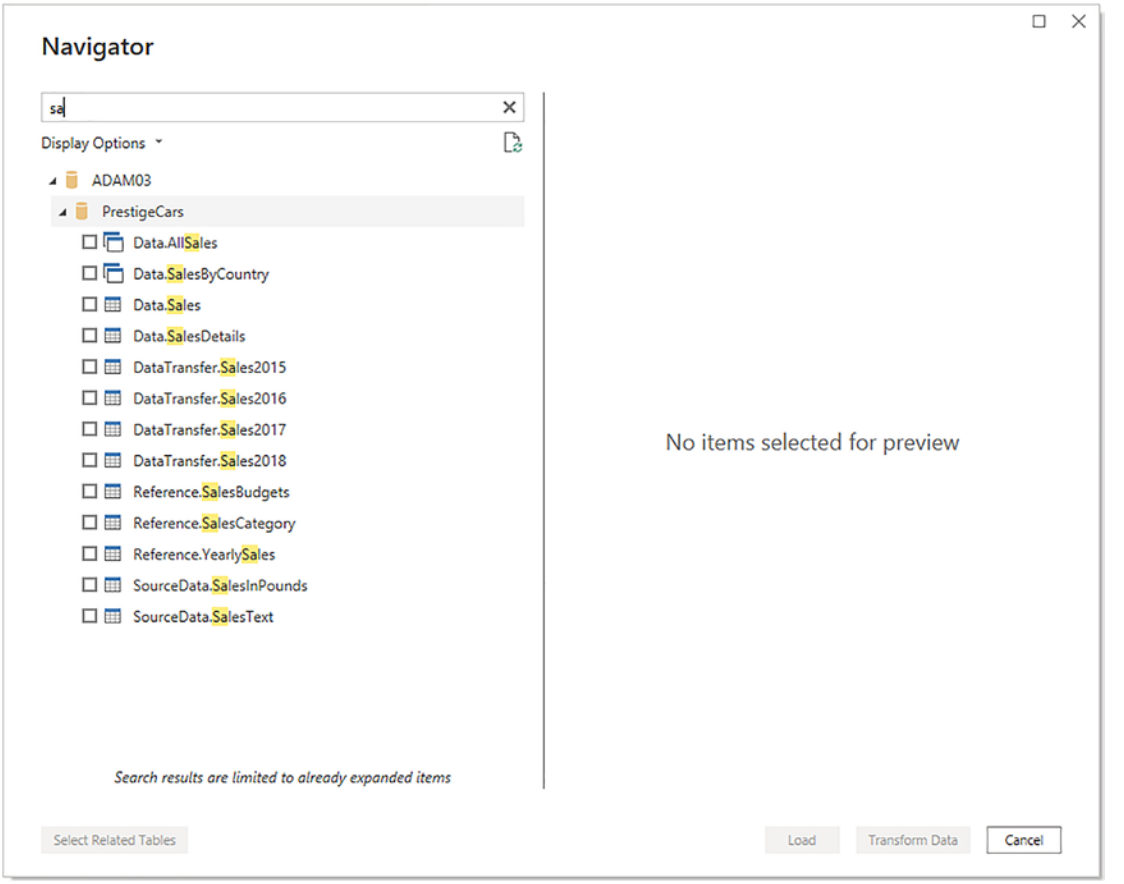


Figure 1-8. *Table search in the Navigator dialog*