Microsoft

Prepare data in Power BI Desktop

aka.ms/PL300-2

© Copyright Microsoft Corporation. All rights reserved.

Agenda

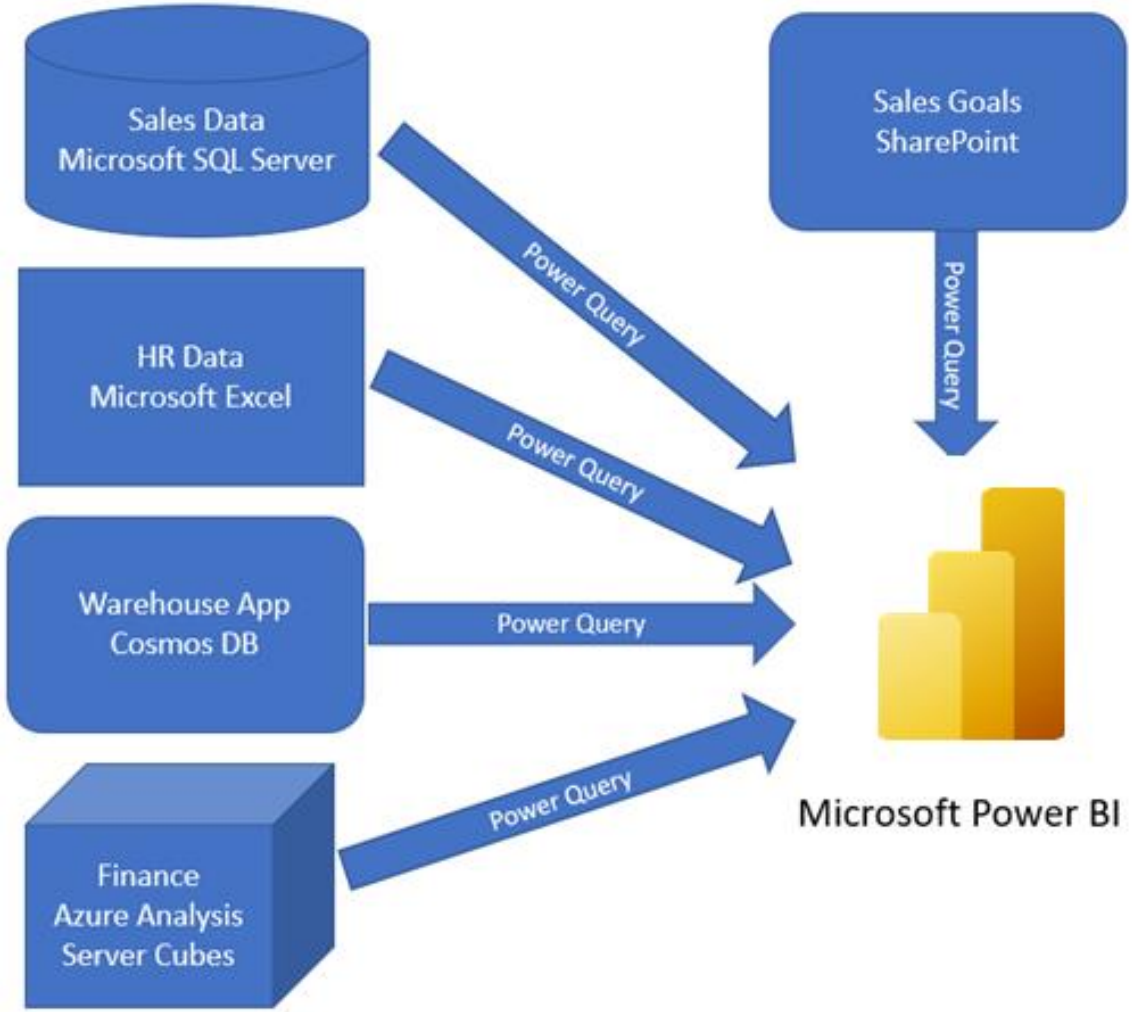
- Get data in Power BI
- Clean, transform, and load data in Power BI

© Copyright Microsoft Corporation. All rights reserved.

Get data in Power BI

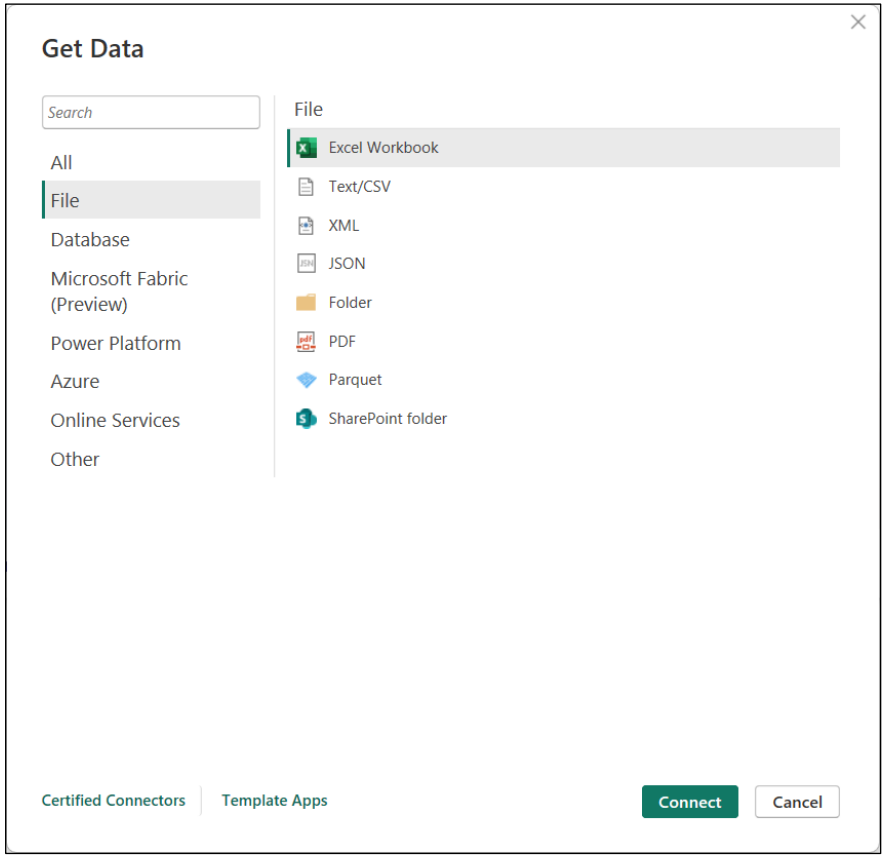
© Copyright Microsoft Corporation. All rights reserved.

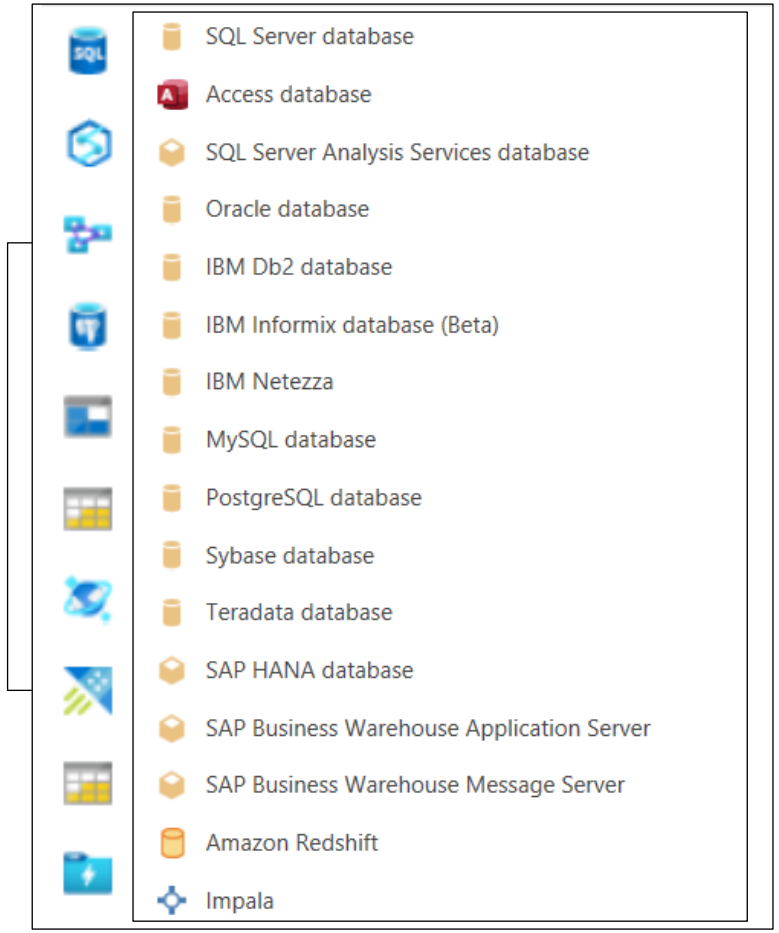
Combine all data into a single semantic model



© Copyright Microsoft Corporation. All rights reserved.

Get data from a wide variety of sources





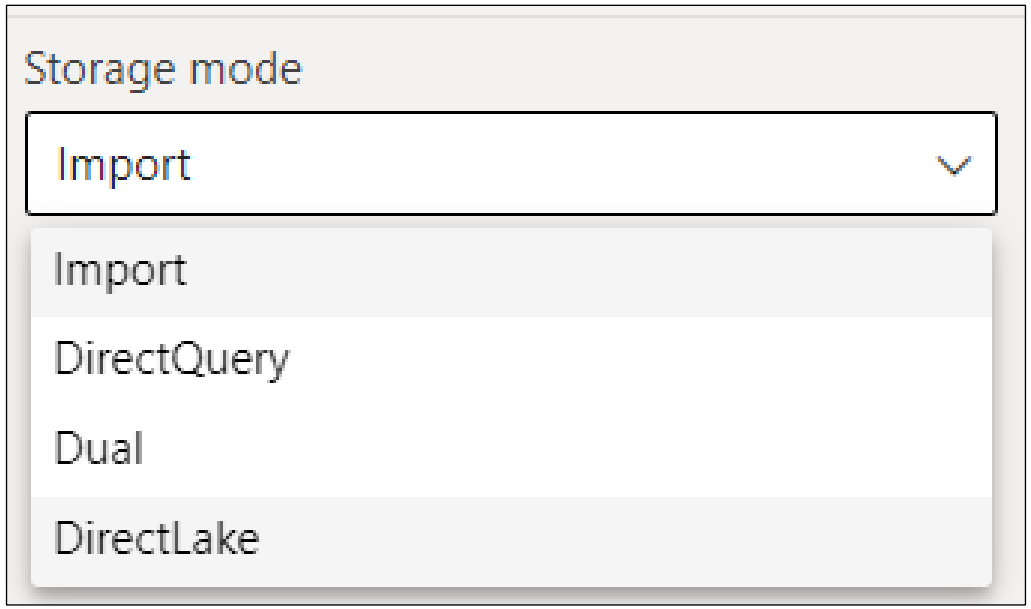
© Copyright Microsoft Corporation. All rights reserved.

Select a storage mode

Storage mode affects

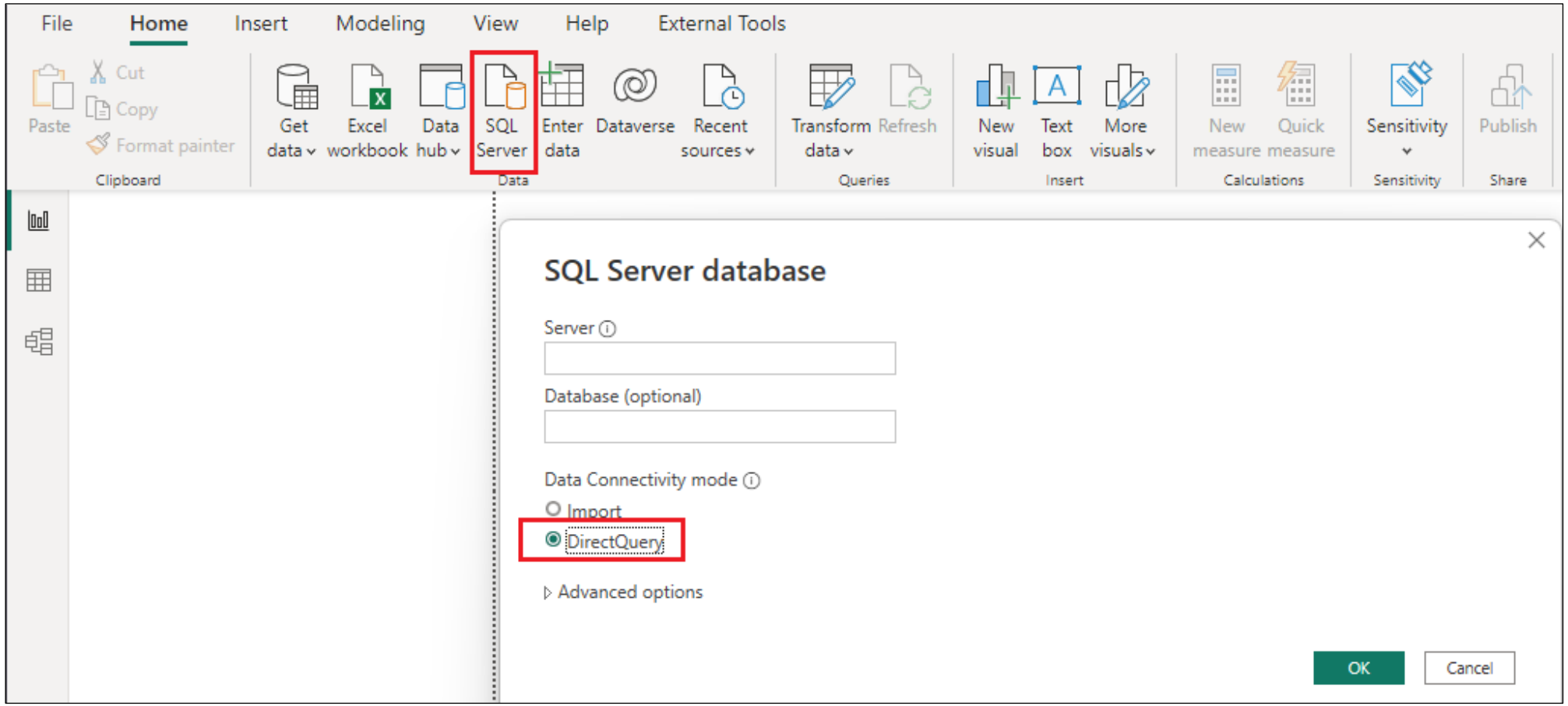
- Available transformations
- Report performance

**Not all sources support all modes*



© Copyright Microsoft Corporation. All rights reserved.

Introduction to DirectQuery



© Copyright Microsoft Corporation. All rights reserved.

Implications of using DirectQuery

Benefits

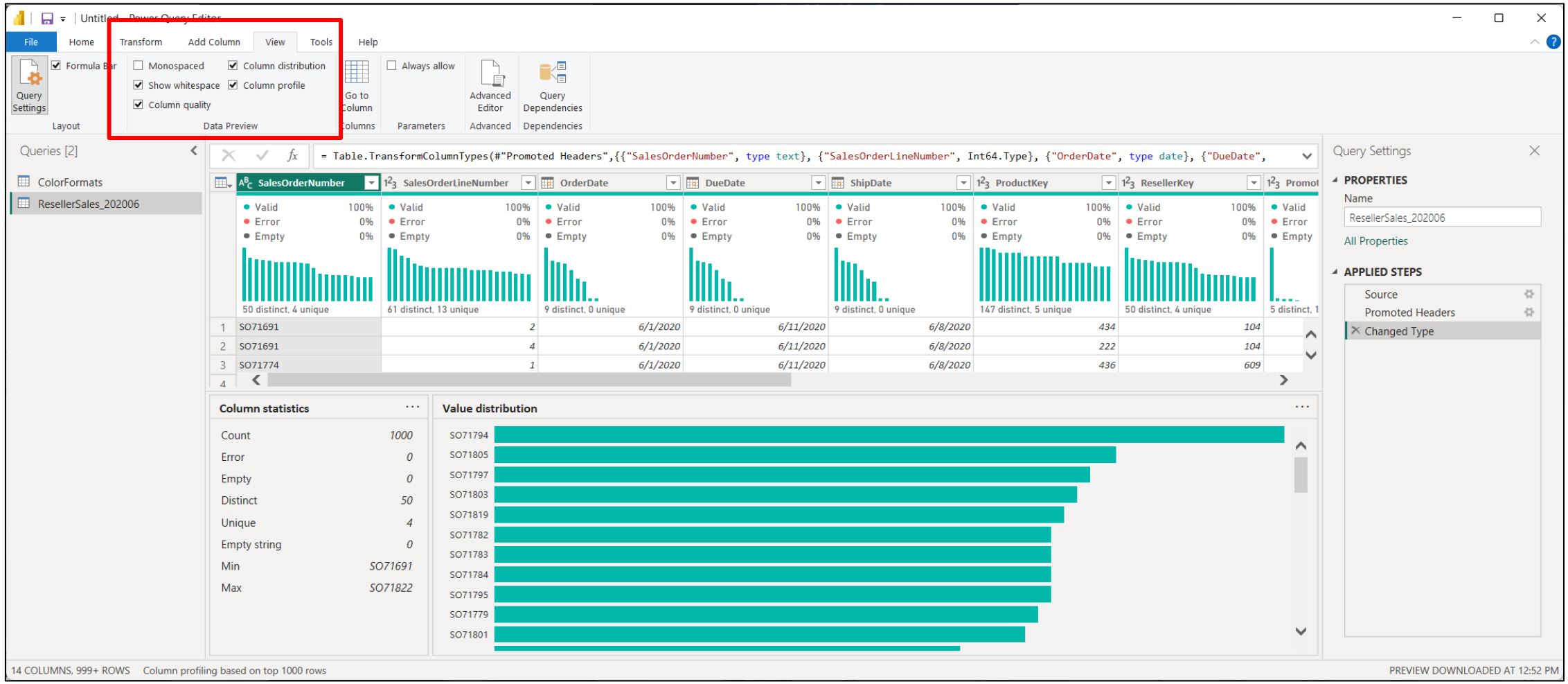
- Frequently changing data
- Need near real-time
- Large data volumes
- Multi-dimensional data

Limitations

- Dependent on data source performance
- Security between source and destination
- Limited modeling capabilities
- Limited transformation features

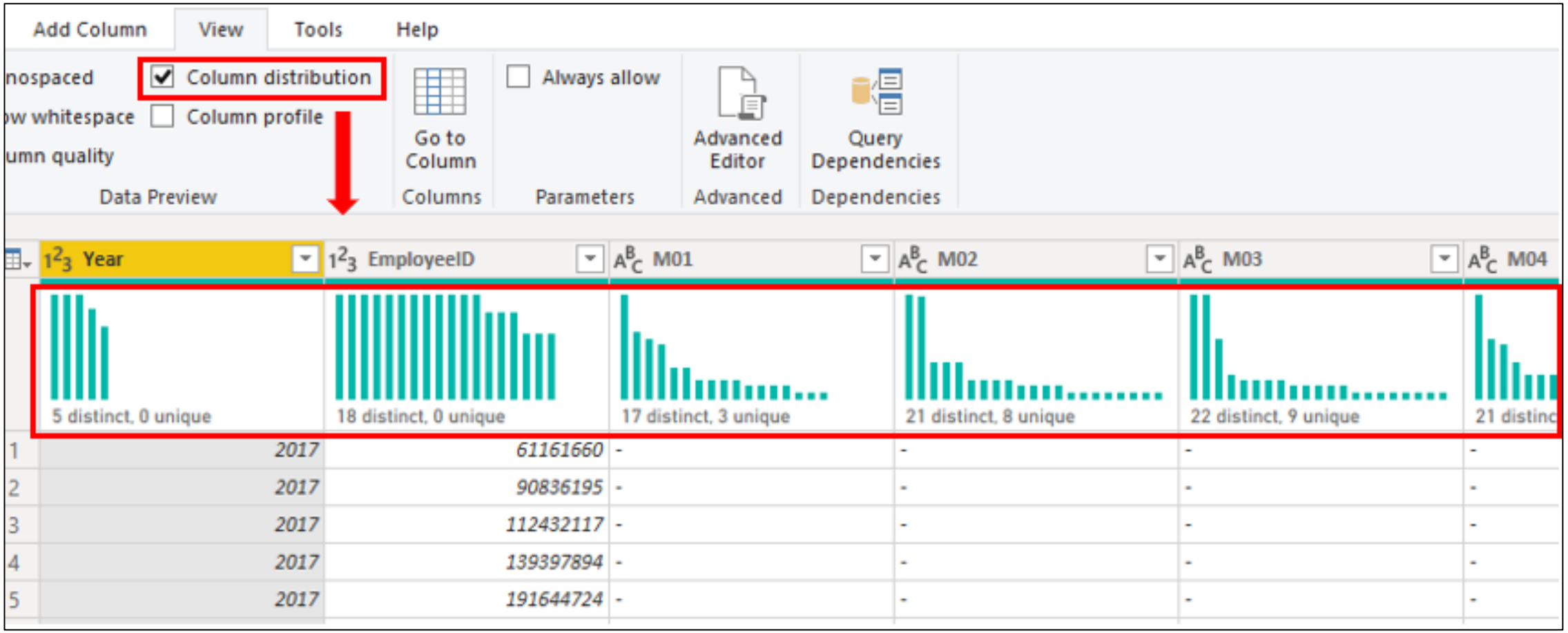
© Copyright Microsoft Corporation. All rights reserved.

Data profiling options in Power Query Editor



© Copyright Microsoft Corporation. All rights reserved.

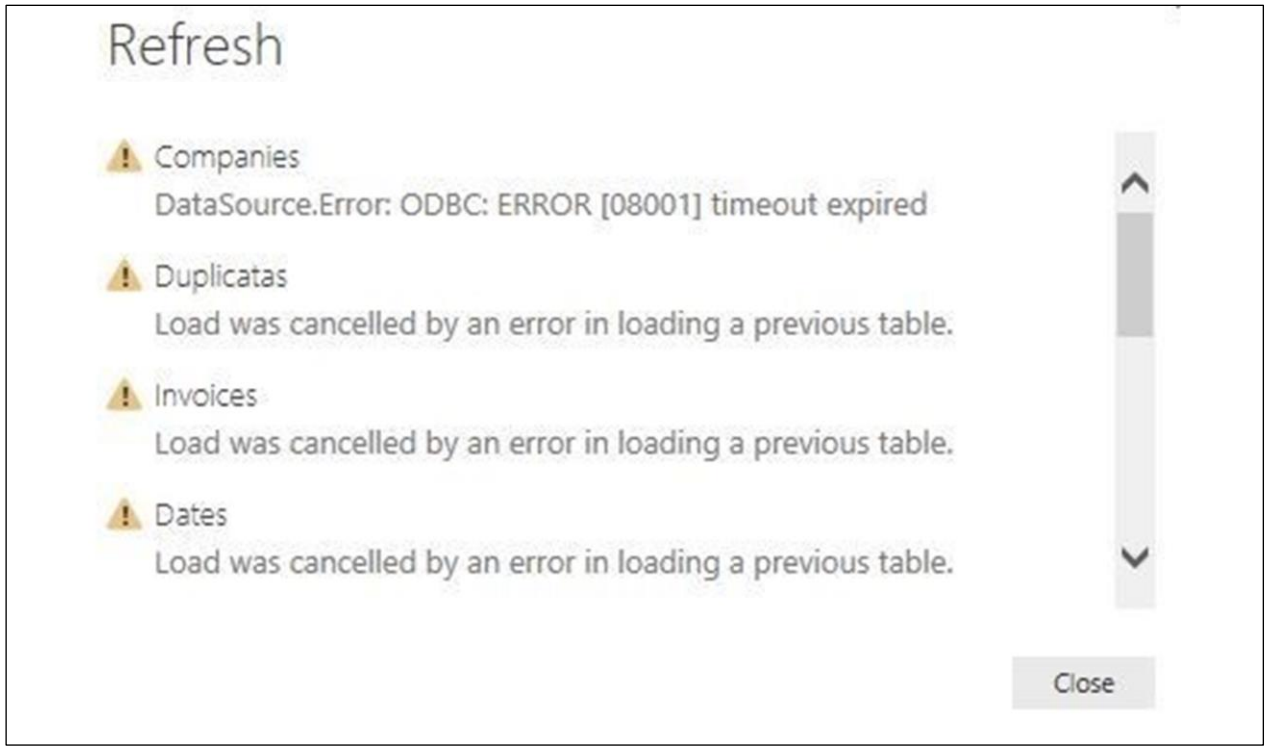
Determine cardinality of a column



© Copyright Microsoft Corporation. All rights reserved.

Data import errors

- Possible data load errors:
  - Query Timeout
  - Couldn’t find data formatted as a table
  - Could not find file
  - Data type errors



© Copyright Microsoft Corporation. All rights reserved.

Lab: Prepare data in Power BI Desktop (30 minutes)

Prepare Data | GitHub Exercise

This lab is designed to introduce you to Power BI Desktop application and how to connect to data and how to use data preview techniques to understand the characteristics and quality of the source data.

- Open Power BI Desktop
- Connect to source data
- Preview source data
- Use data profile tools

© Copyright Microsoft Corporation. All rights reserved.

Knowledge check: get data

When connecting to a SQL Server database to get data, what language should you use to extract data?

- ☐ DAX
- ☑ T-SQL
- ☐ MDX

How many rows are sampled in Power Query Editor by default?

- ☐ All
- ☑ 1000
- ☐ 100

What can you do to improve performance when you're getting data in Power BI?

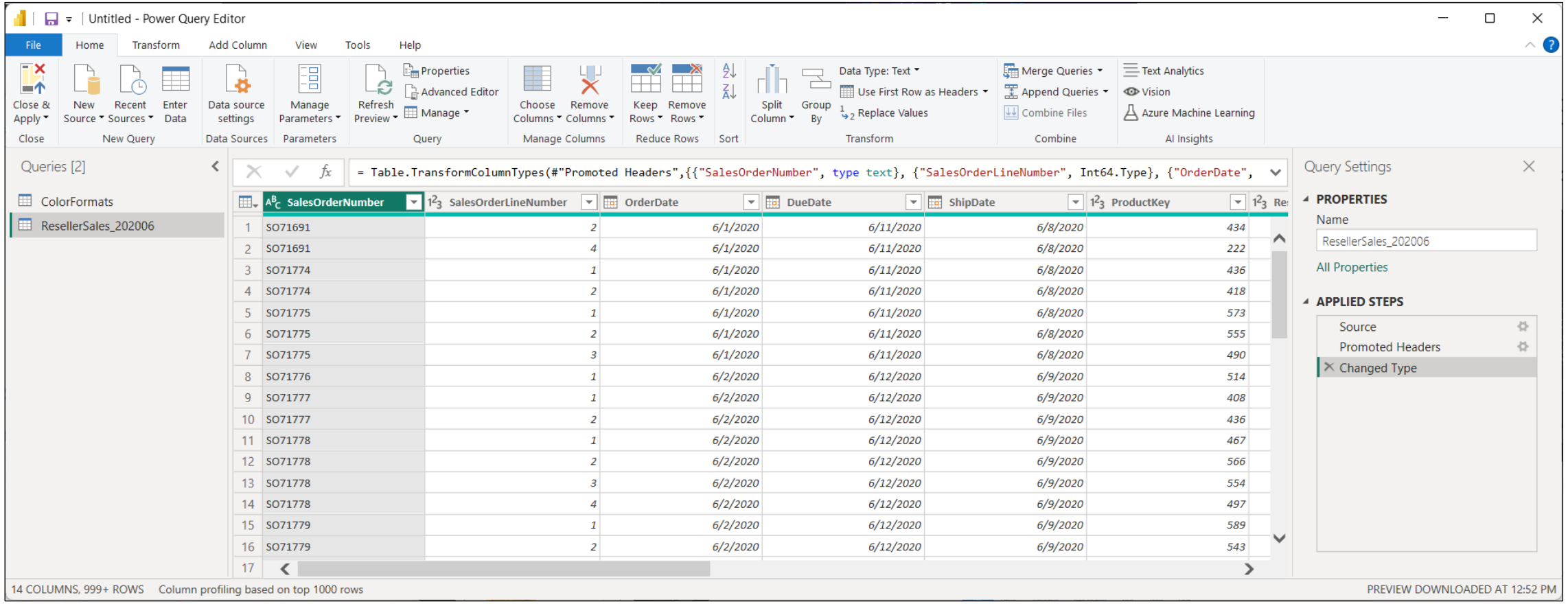
- ☑ Remove unnecessary columns and rows
- ☐ Export database files to CSV to load
- ☐ Combine date and time columns into a single column

© Copyright Microsoft Corporation. All rights reserved.

Clean, transform, and load data

© Copyright Microsoft Corporation. All rights reserved.

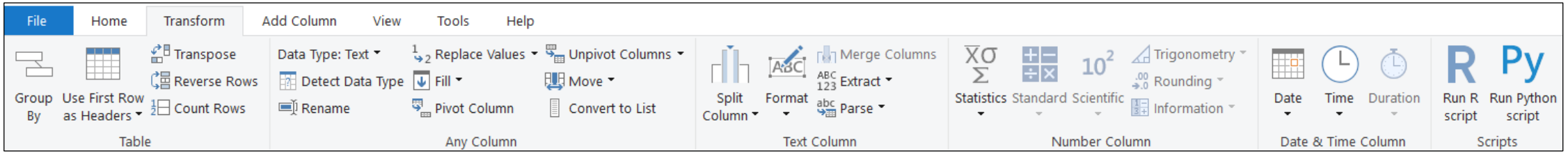
Transform data with Power Query Editor

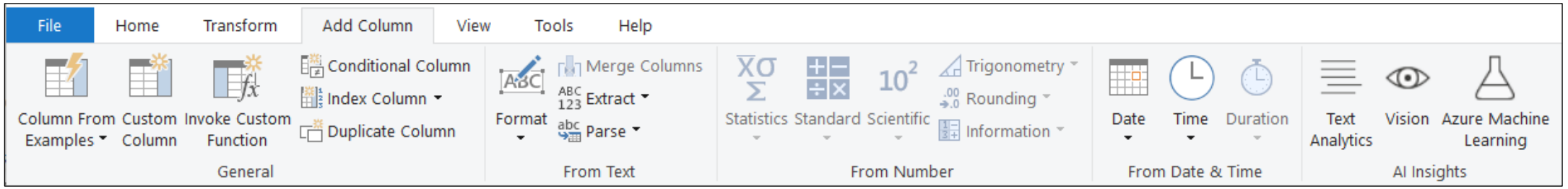


© Copyright Microsoft Corporation. All rights reserved.

Common transformations

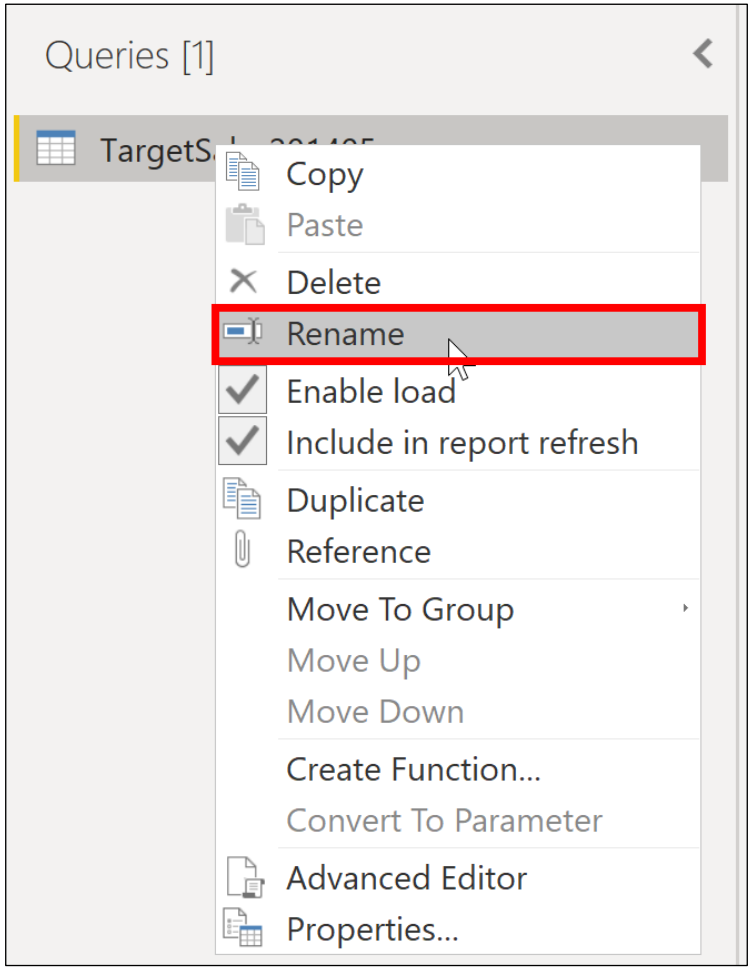
Transform columns, add new, split, extract, and more

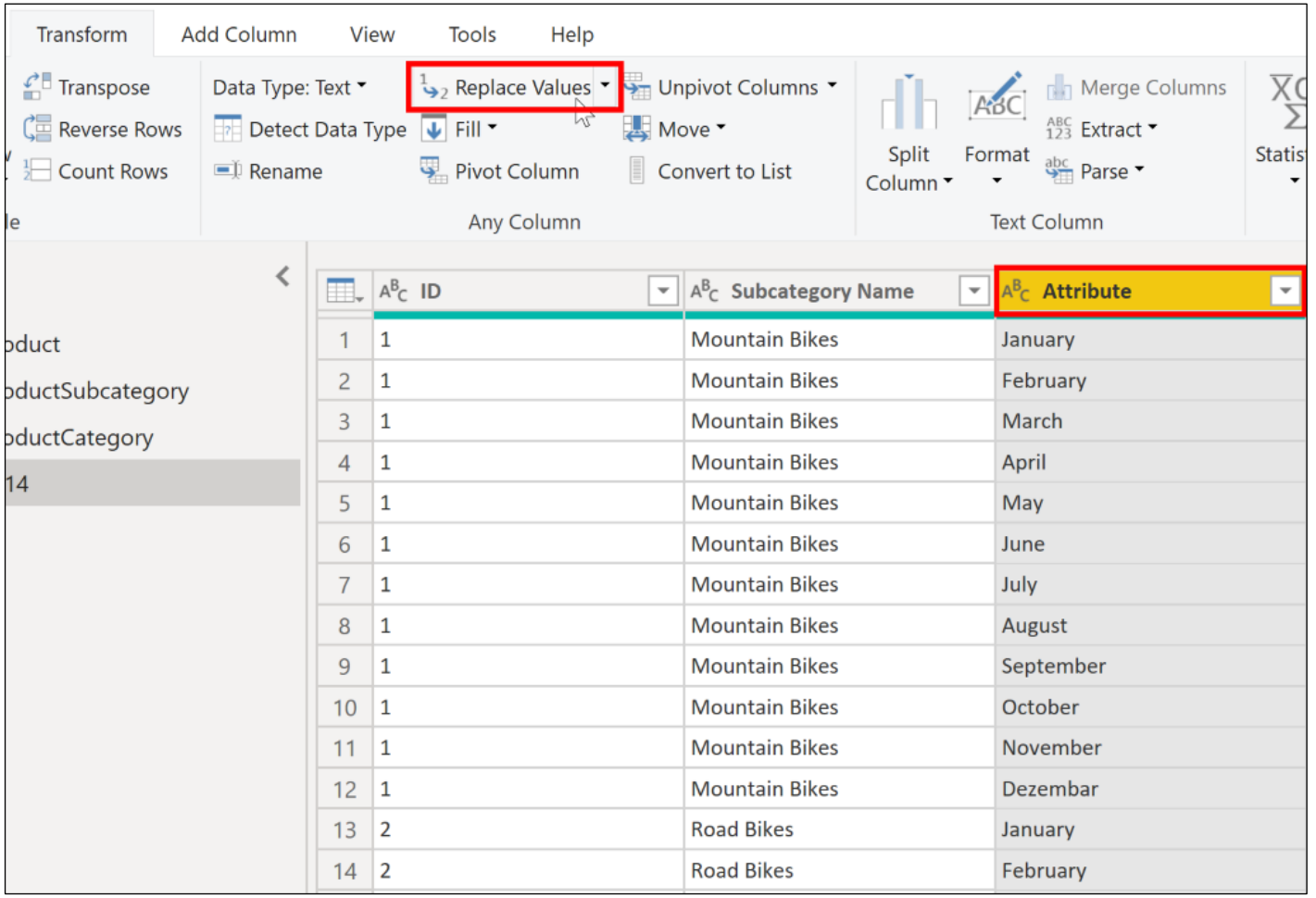




© Copyright Microsoft Corporation. All rights reserved.

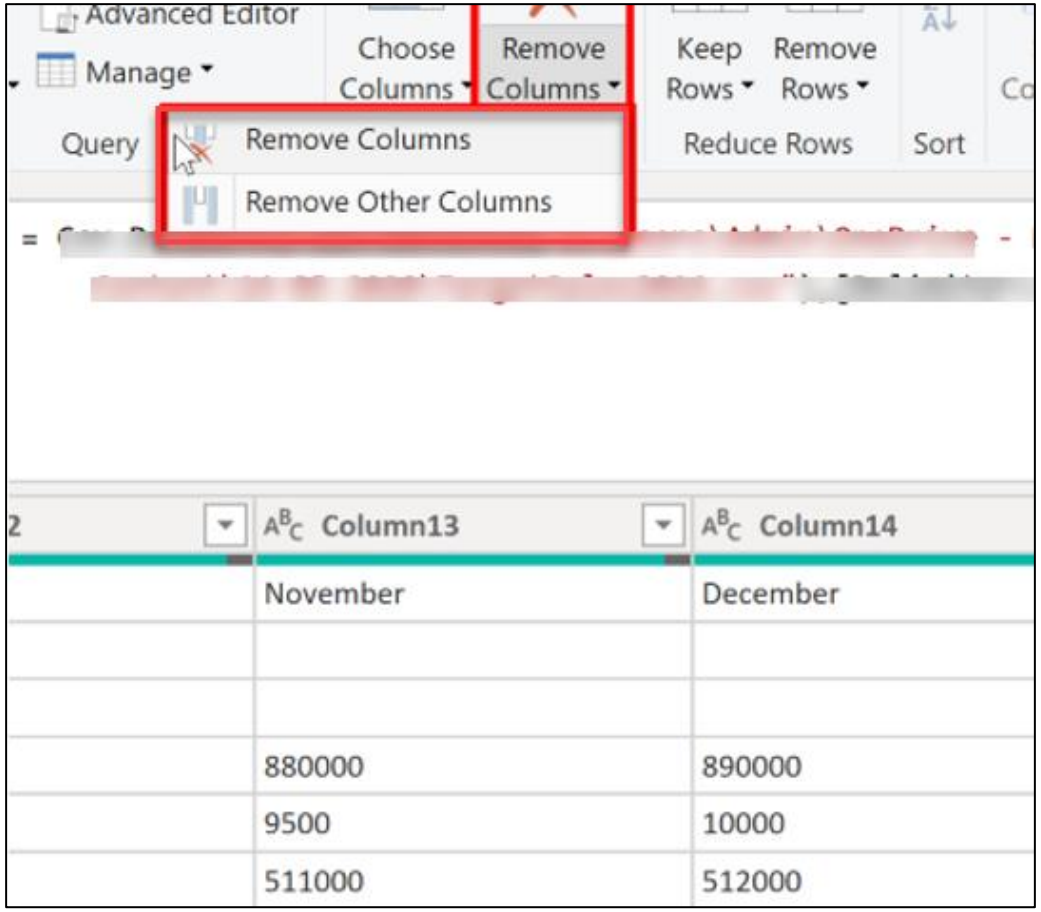
Choose user-friendly values

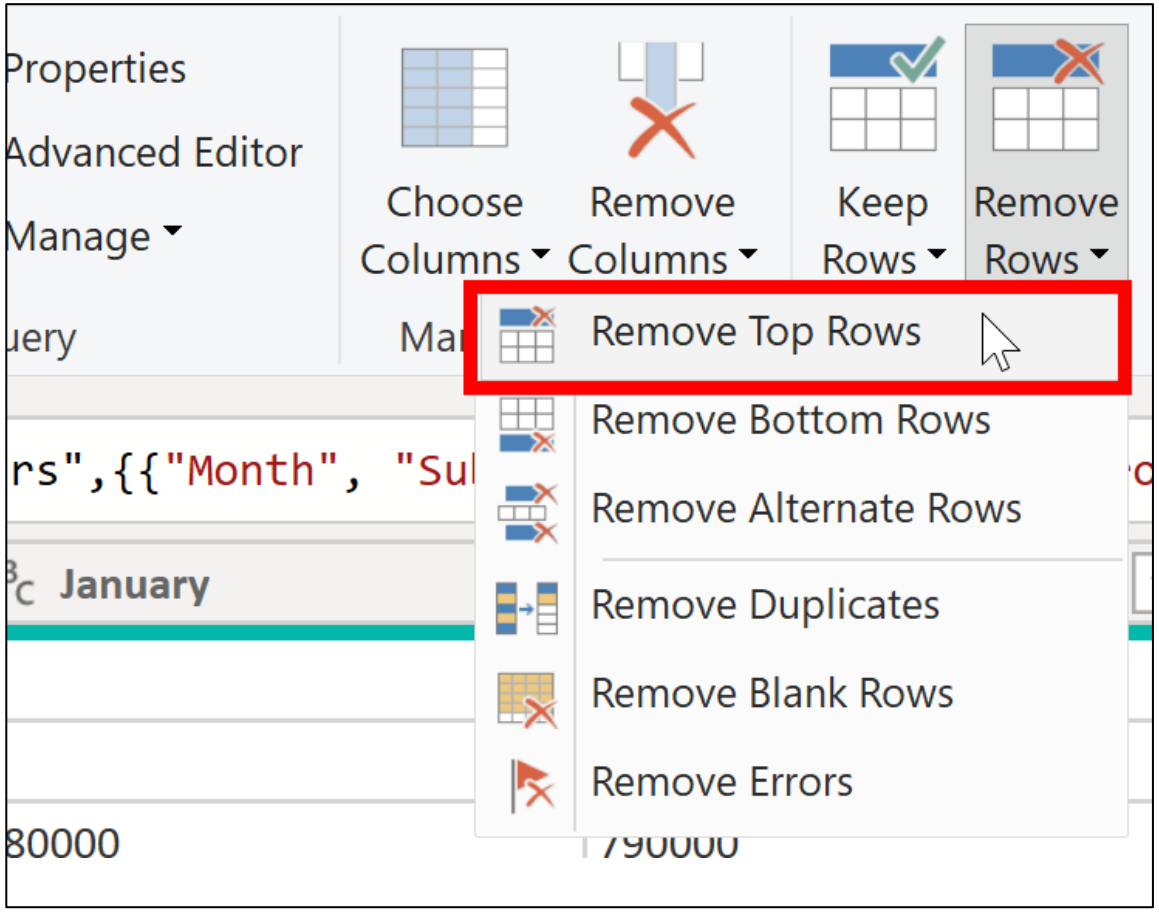




© Copyright Microsoft Corporation. All rights reserved.

Shaping table structure

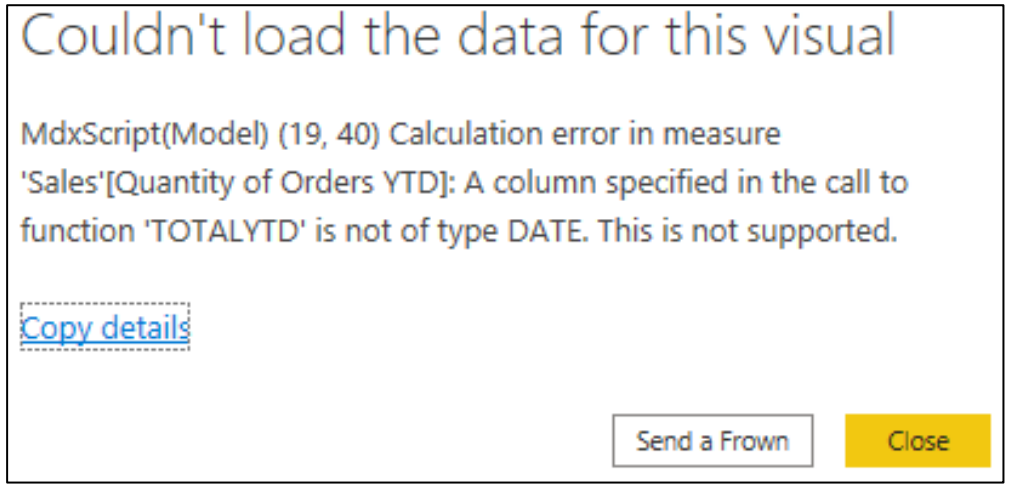




© Copyright Microsoft Corporation. All rights reserved.

Evaluate and change column data types

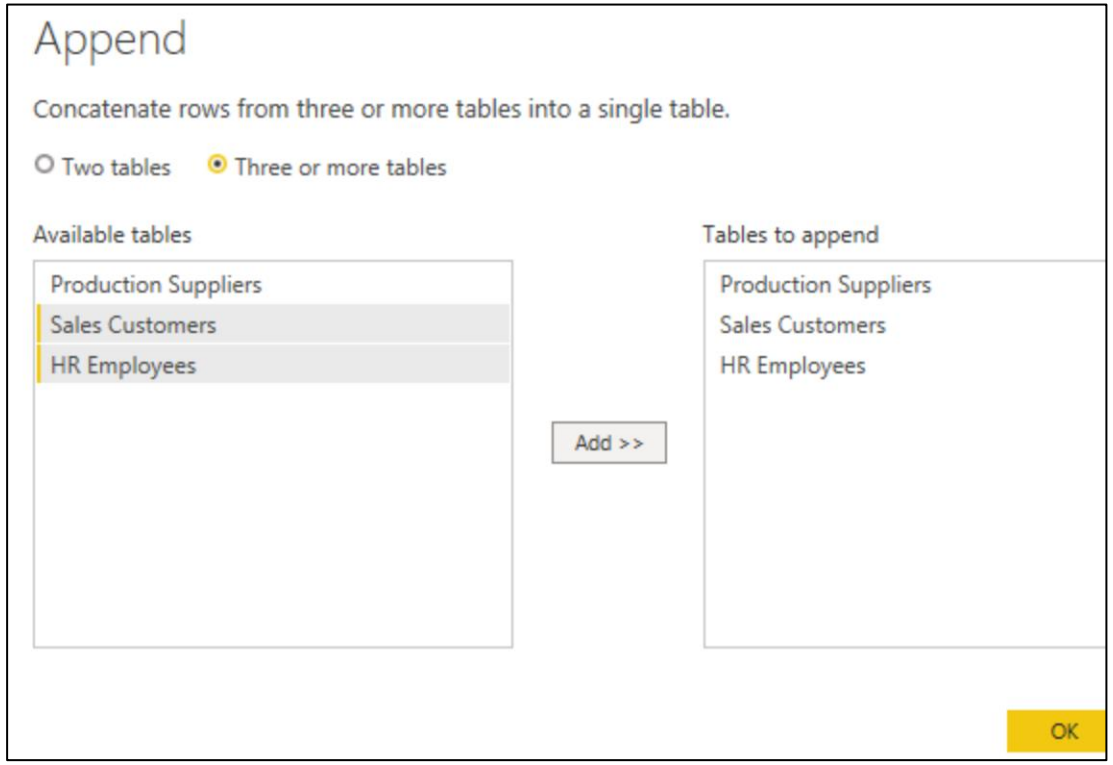
SalesOrderNumber	OrderDate	TotalCost
SO43659	7/1/2017	2024.99
SO43659	7/1/2017	6074.97
SO43659	7/1/2017	2024.99
SO43659	7/1/2017	2039.99
SO43659	7/1/2017	2039.99
SO43659	7/1/2017	4079.98
SO43659	7/1/2017	2039.99
SO43659	7/1/2017	86.52
SO43659	7/1/2017	28.84
SO43659	7/1/2017	34.2



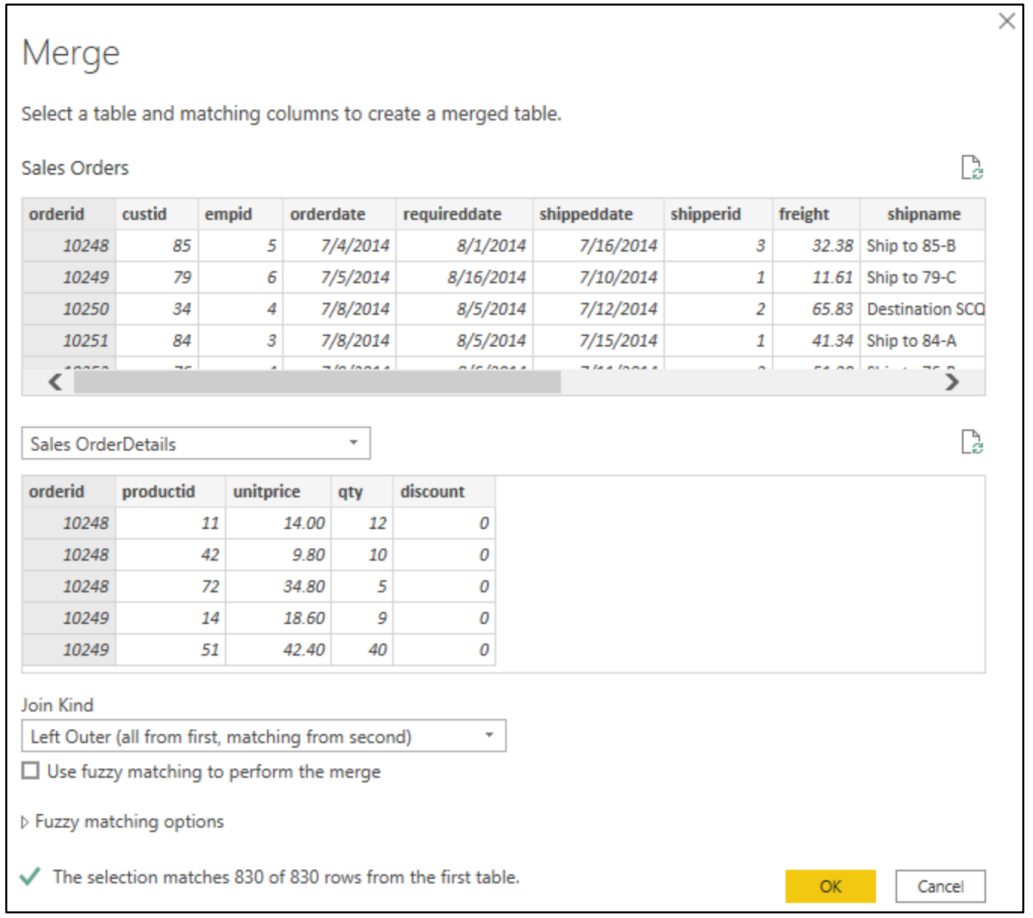
© Copyright Microsoft Corporation. All rights reserved.

Combine multiple queries into one

Append



Merge



© Copyright Microsoft Corporation. All rights reserved.

Unpivot or pivot columns

Add or remove table structure to meet your aggregation needs.

	Category Name	Subcategory Name
1	Bikes	Mountain Bikes
2	Bikes	Road Bikes
3	Bikes	Touring Bikes
4	Clothing	Bib-Shorts
5	Clothing	Caps
6	Clothing	Gloves
7	Clothing	Jerseys
8	Clothing	Shorts
9	Clothing	Socks
10	Clothing	Tights
11	Clothing	Vests
12	Accessories	Bike Racks
13	Accessories	Bike Stands
14	Accessories	Bottles and Cages

	1.2 Bikes	1.2 Components	1.2 Clothing	1.2 Accessories
1	*3*	*14*	*8*	*12*

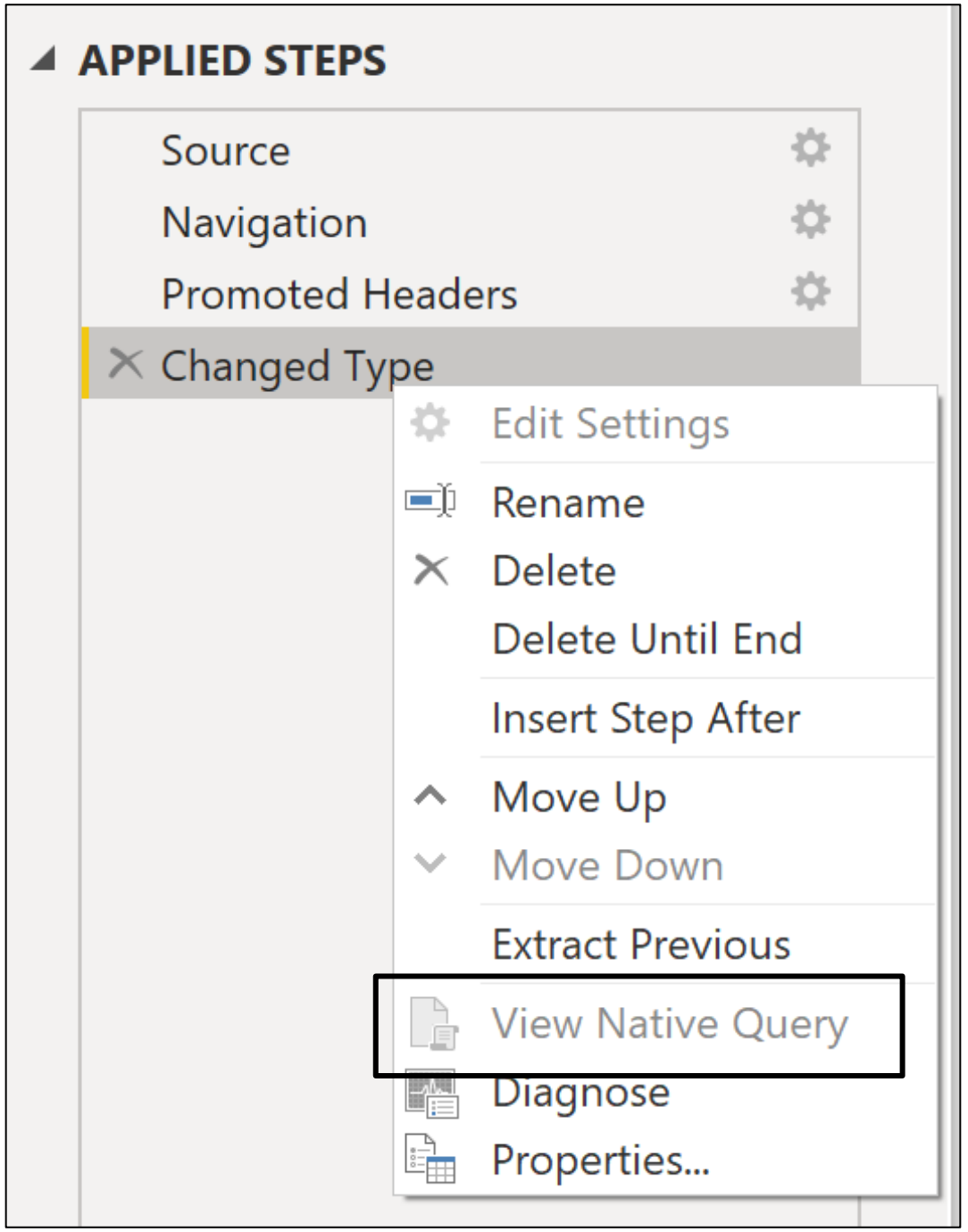
© Copyright Microsoft Corporation. All rights reserved.

Query folding

Pushes data transformations to the source for better performance and efficiency.

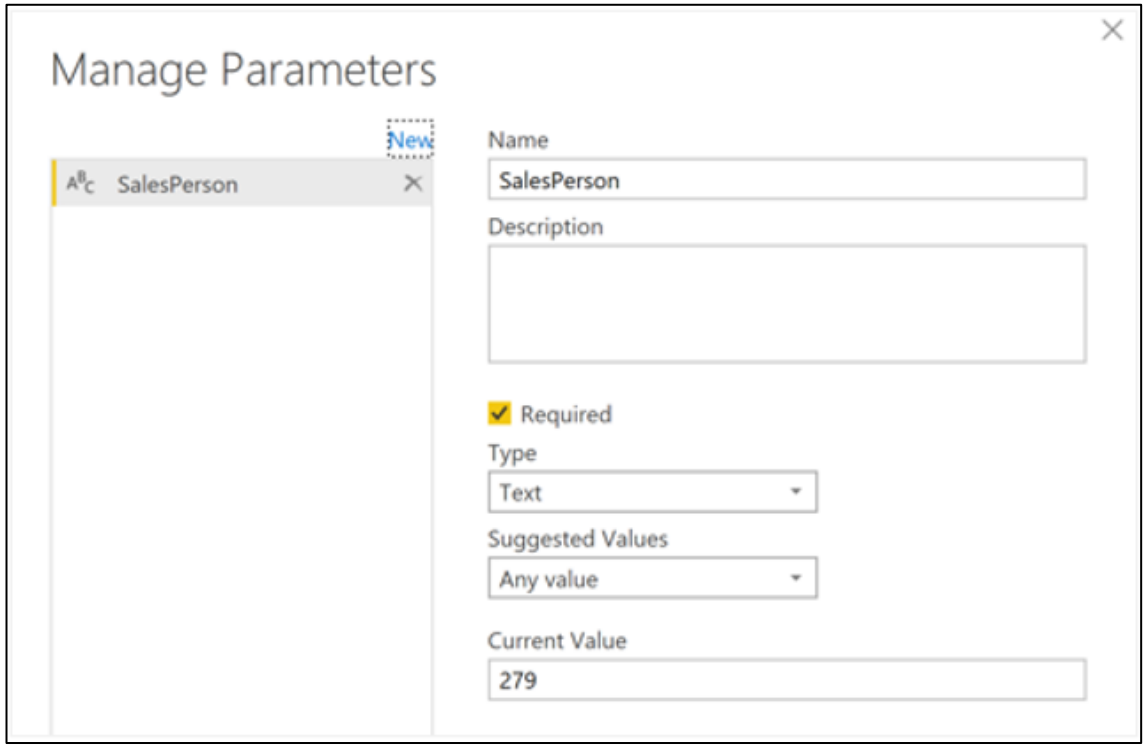
Supported sources

- Relational databases
- OData feeds
- Active Directory



© Copyright Microsoft Corporation. All rights reserved.

Dynamic reports with parameters



	SalesOrderNumber	SalesOrderID	SalesPersonID
1	SO43659	43659	279
2	SO43660	43660	279
3	SO43681	43681	279
4	SO43684	43684	279
5	SO43685	43685	279
6	SO43694	43694	279
7	SO43695	43695	279
8	SO43696	43696	279

© Copyright Microsoft Corporation. All rights reserved.

Performance recommendations

- Only keep necessary data
- Check data types
- Reduce cardinality
- Disable query load
- Use parameters

© Copyright Microsoft Corporation. All rights reserved.

Lab: Load data in Power BI Desktop (45 minutes)

Load Data | GitHub Exercise

In this lab, you'll apply transformations to each of the queries created in the previous lab. You'll then apply the queries to load each as a table to the data model.

In this lab you learn how to:

- Apply various transformations
- Apply queries to load them to the data model

© Copyright Microsoft Corporation. All rights reserved.

Knowledge check: clean and transform data

What is a risk of having null values in a numeric column?

- ☐ That function SUM of data are incorrect.
- ☐ That function MAX of data will be incorrect.
- ☑ That function AVERAGE of data will be incorrect.

If you have two queries that have different data but the same column headers, and you want to combine both tables into one query with all the combined rows, which operation should you perform?

- ☑ Append
- ☐ Merge
- ☐ Combine

© Copyright Microsoft Corporation. All rights reserved.

Recap

In this section, we covered:

Connecting to various data sources.

Determining the right storage mode.

Profiling data for accuracy and insights.

Cleaning, transforming, and loading data.

Using Query Folding for appropriate sources.

Recommendations for more performant reports.

© Copyright Microsoft Corporation. All rights reserved.

Microsoft

Thanks

Resources

Prepare data for analysis

Clean, transform, and load data in Power BI