Power Query Editor

Objective

This chapter

- Loading data from folder
- Understanding Power Query in detail
- Promote header, split to limiter, Add columns, append, merge queries etc.

Power Query Editor

Power Query Editor is a data transformation and cleansing tool available in Power BI. It is used to connect to various data sources, transform and shape data, and load the resulting data into the Power BI data model.

Power Query Editor provides a graphical interface for performing data transformations using a series of intuitive steps. You can use it to combine data from multiple sources, filter data, split columns, remove duplicates, change data types, and more.

Once the data has been transformed, you can preview the changes and load the data into Power BI. The resulting data can then be used to create visualizations, reports, and dashboards. Power Query Editor can be accessed by selecting the "Transform Data" option in the Home tab of the Power BI Desktop ribbon.

With Power BI Desktop you can connect to the world of data, create compelling and foundational reports, and share your efforts with others – who can then build on your work, and expand their business intelligence efforts.

Power BI Desktop has three views:

- **Report** view – You can use queries that you create to build compelling visualizations, arranged as you want them to appear, and with multiple pages, that you can share with others.
- **Data** view – See the data in your report in data model format, where you can add measures, create new columns, and manage relationships.
- **Model** view – Get a graphical representation of the relationships that are established in your data model, and manage or modify them as needed.

Access these views by selecting one of the three icons along the left side of Power BI Desktop. In the following image, **Report** view is selected, indicated by the yellow band beside the icon.
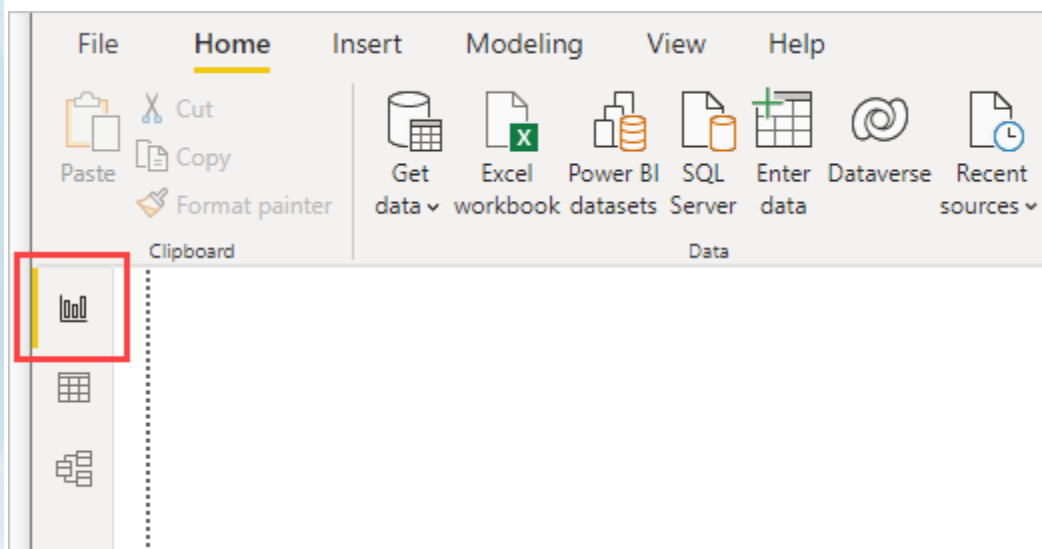

Power BI Desktop also comes with Power Query Editor. Use Power Query Editor to connect to one or many data sources, shape and transform the data to meet your needs, then load that model into Power BI Desktop.

This article provides an overview of the work with data in the Power Query Editor, but there's more to learn. At the end of this article, you'll find links to detailed guidance about supported data types. You'll also find guidance about connecting to data, shaping data, creating relationships, and how to get started.
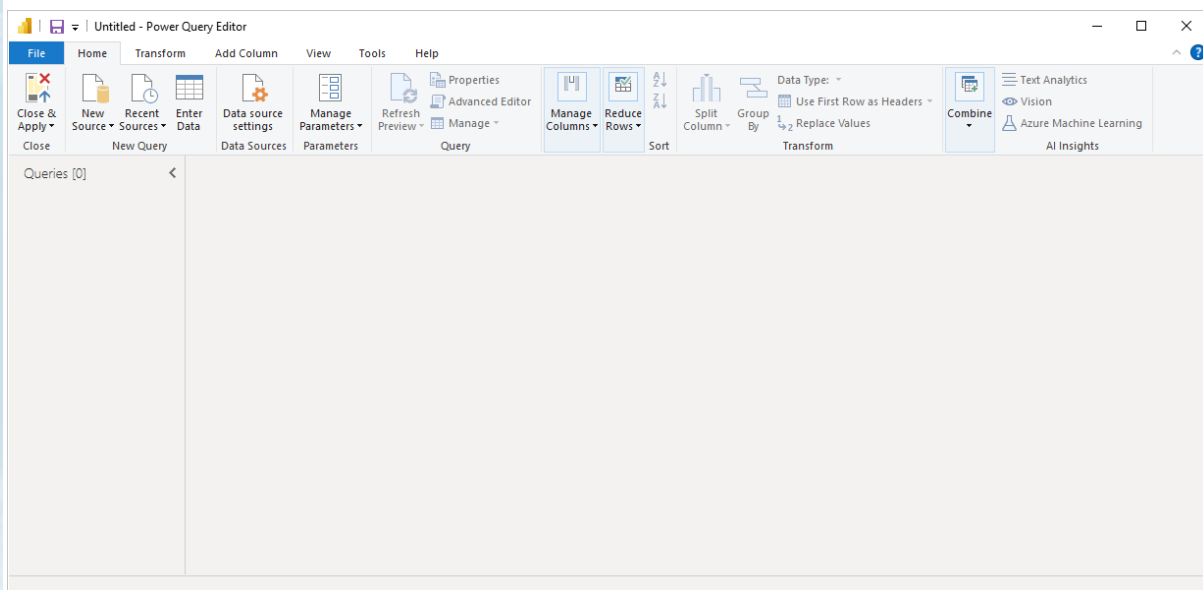
But first, let's see get acquainted with Power Query Editor.

Power Query Editor

To get to Power Query Editor, select **Transform data** from the **home** tab of Power BI Desktop.
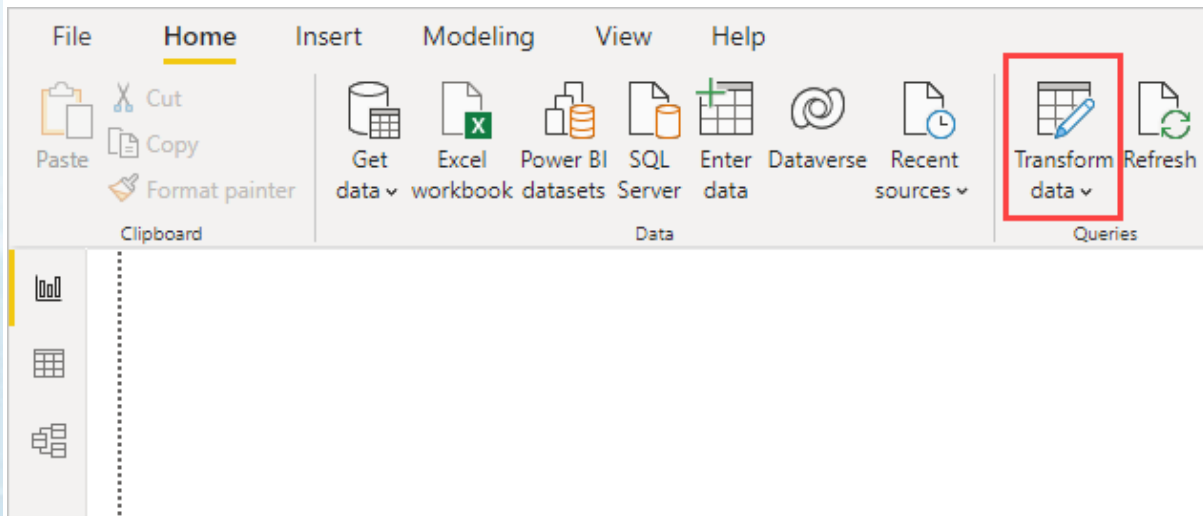
With no data connections, Power Query Editor appears as a blank pane, ready for data.


After a query is loaded, Power Query Editor view becomes more interesting. If you connect to the following Web data source, Power Query Editor loads information about the data, which you can then begin to shape:

Here's how Power Query Editor appears after a data connection is established:

1. In the ribbon, many buttons are now active to interact with the data in the query.
2. In the left pane, queries are listed and available for selection, viewing, and shaping.
3. In the centre pane, data from the selected query is displayed and available for shaping.

4. The **Query Settings** pane appears, listing the query's properties and applied steps.
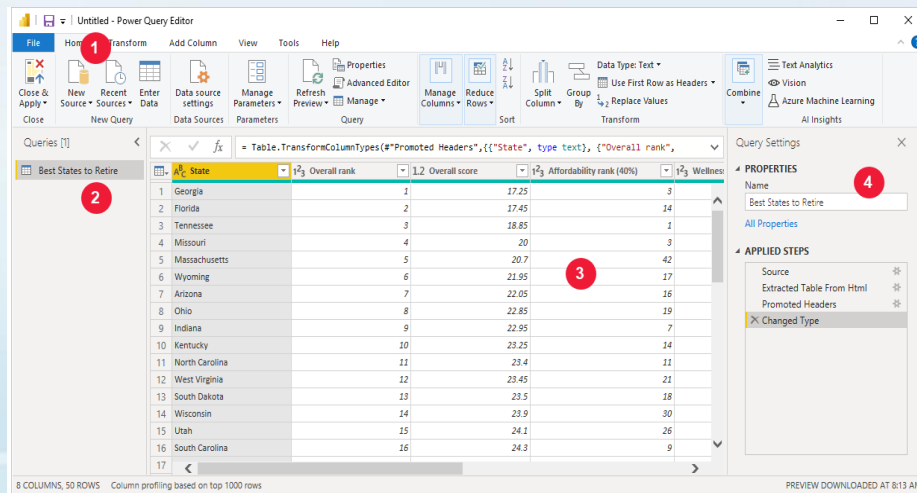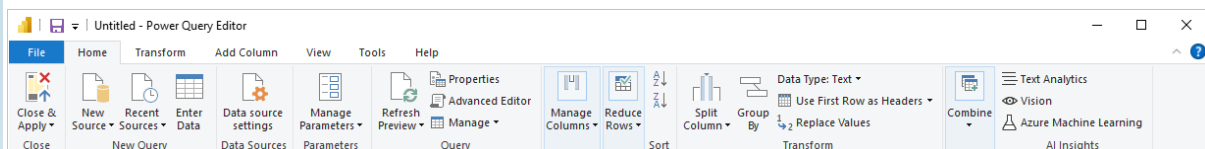

Each of these four areas will be explained later: the ribbon, the Queries pane, the Data view, and the Query Settings pane.
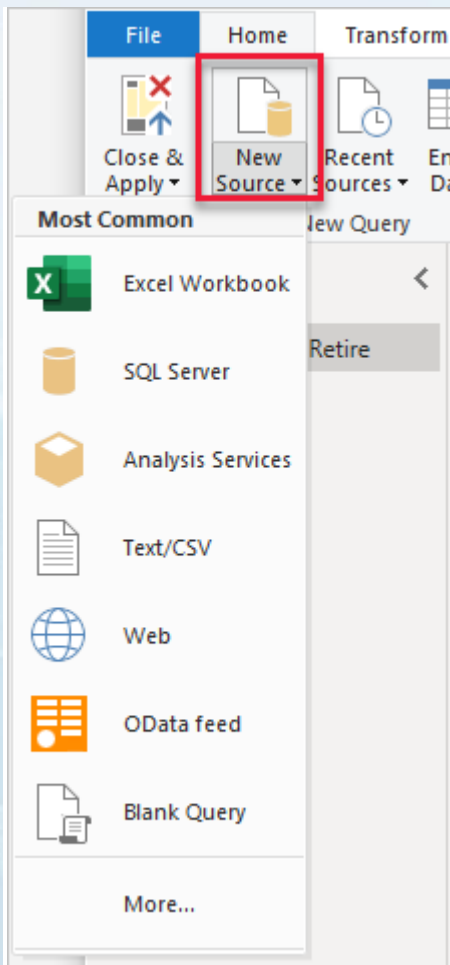
The query ribbon

The ribbon in Power Query Editor consists of four tabs: **Home**, **Transform**, **Add Column**, **View**, **Tools**, and **Help**.

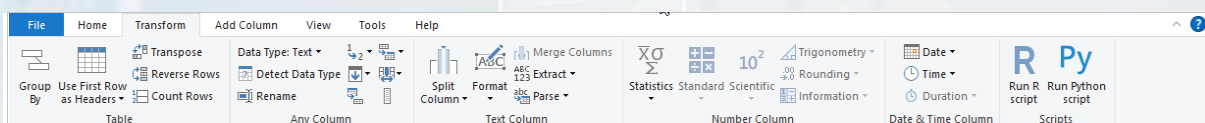The **Home** tab contains the common query tasks.

To connect to data and begin the query building process, select **New Source**. A menu appears, providing the most common data sources.
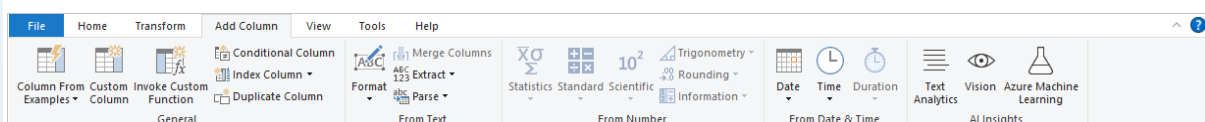
The **Transform** tab provides access to common data transformation tasks, such as:

- Adding or removing columns
- Changing data types
- Splitting columns
- Other data-driven tasks

The **Add Column** tab provides more tasks associated with adding a column, formatting column data, and adding custom columns. The following image shows the **Add Column** tab.


The **View** tab on the ribbon is used to toggle whether certain panes or windows are displayed. It's also used to display the Advanced Editor. The following image shows the **View** tab.
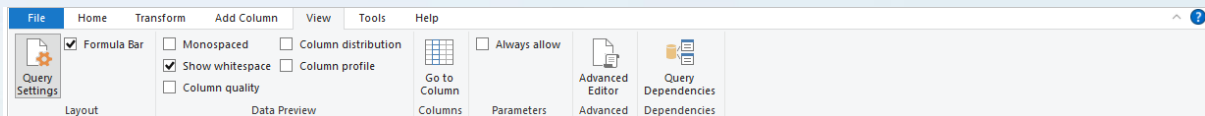
It's useful to know that many of the tasks available from the ribbon are also available by right-clicking a column, or other data, in the centre pane.
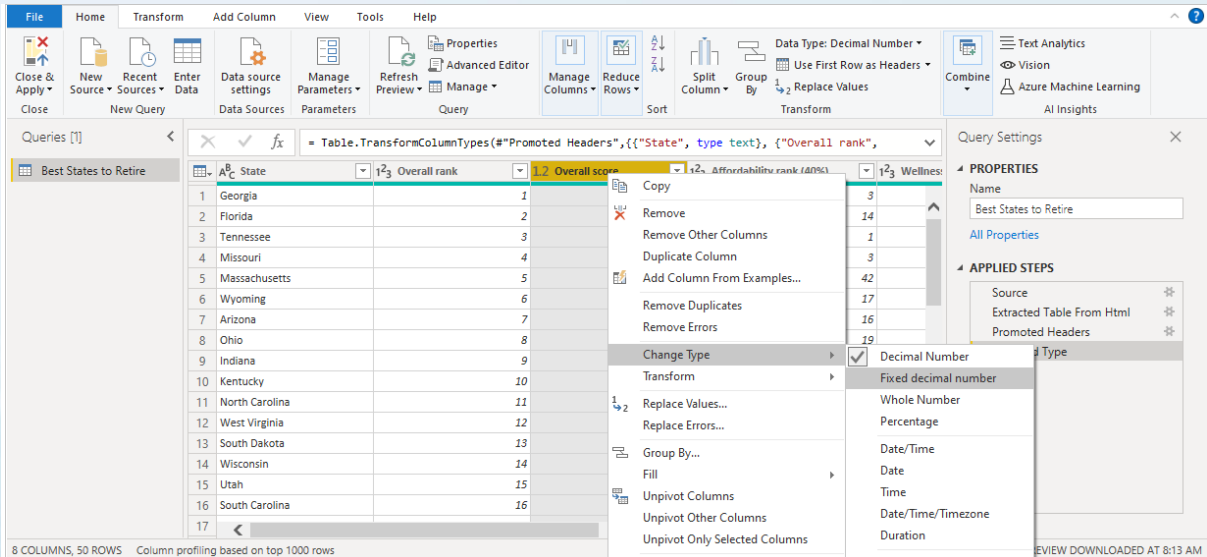
The left (Queries) pane

The left pane, or **Queries** pane, displays the number of active queries and the name of the query. When you select a query from the left pane, its data is displayed in the centre pane, where you can shape and transform the data to meet your needs. The following image shows the left pane with a query.
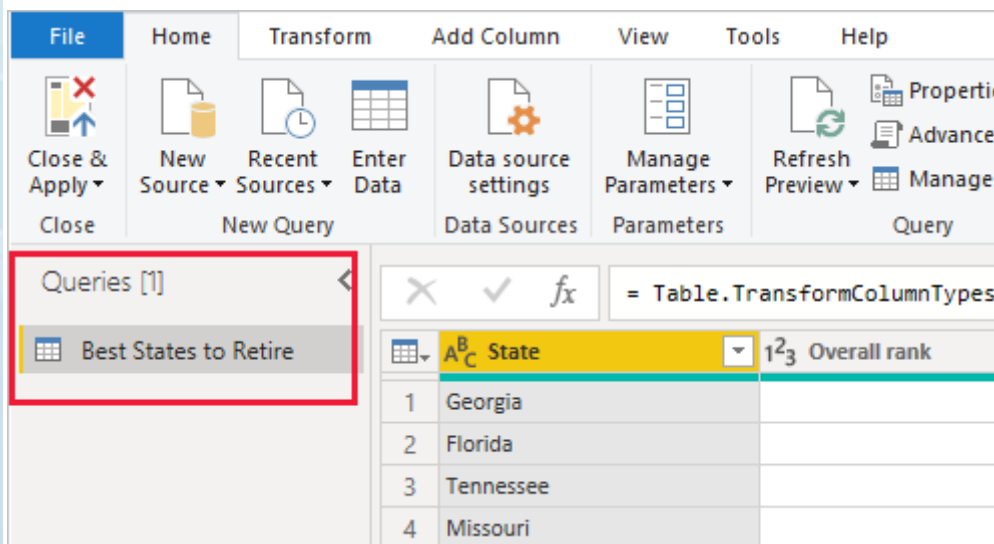
The centre (Data) pane

In the centre pane, or **Data** pane, data from the selected query is displayed. This pane is where much of the work of the **Query** view is accomplished.

The following image shows the Web data connection established earlier. The **Overall score** column is selected, and its header is right-clicked to show the available menu items. Notice that many of these items in the right-click menu are the same as buttons in the ribbon tabs.

When you select a right-click menu item (or a ribbon button), the query applies the step to the data. It also saves step as part of the query itself. The steps are recorded in the **Query Settings** pane in sequential order, as described in the next section.

The right (Query Settings) pane

The right pane, or **Query Settings** pane, is where all steps associated with a query are displayed. For example, in the following image, the **Applied Steps** section of the **Query Settings** pane reflects the fact that we just changed the type of the **Overall score** column.
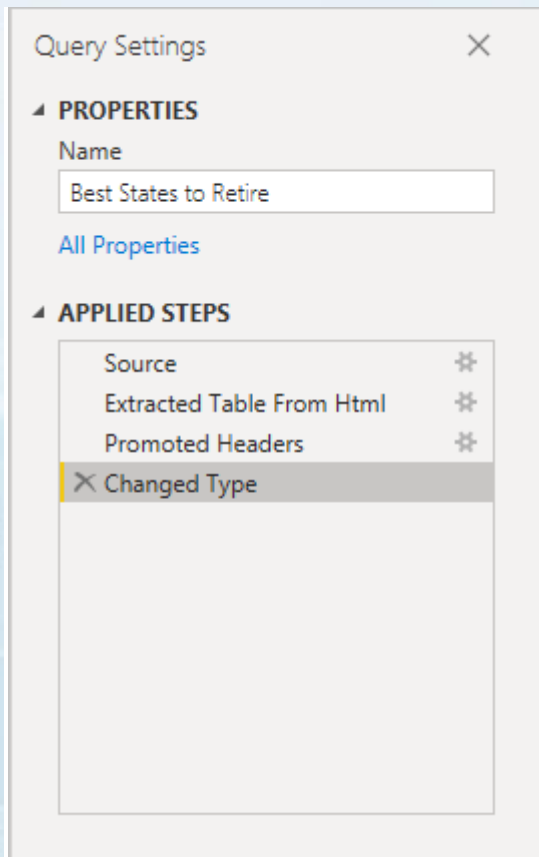
As more shaping steps are applied to the query, they're captured in the **Applied Steps** section.

It's important to know that the underlying data *isn't* changed. Rather, Power Query Editor adjusts and shapes its view of the data. It also shapes and adjusts the view of any interaction with the underlying data that occurs based on Power Query Editor's shaped and modified view of that data.

In the **Query Settings** pane, you can rename steps, delete steps, or reorder the steps as you see fit. To do so, right-click the step in the **Applied Steps** section, and choose from the menu that appears. All query steps are carried out in the order they appear in the **Applied Steps** pane.
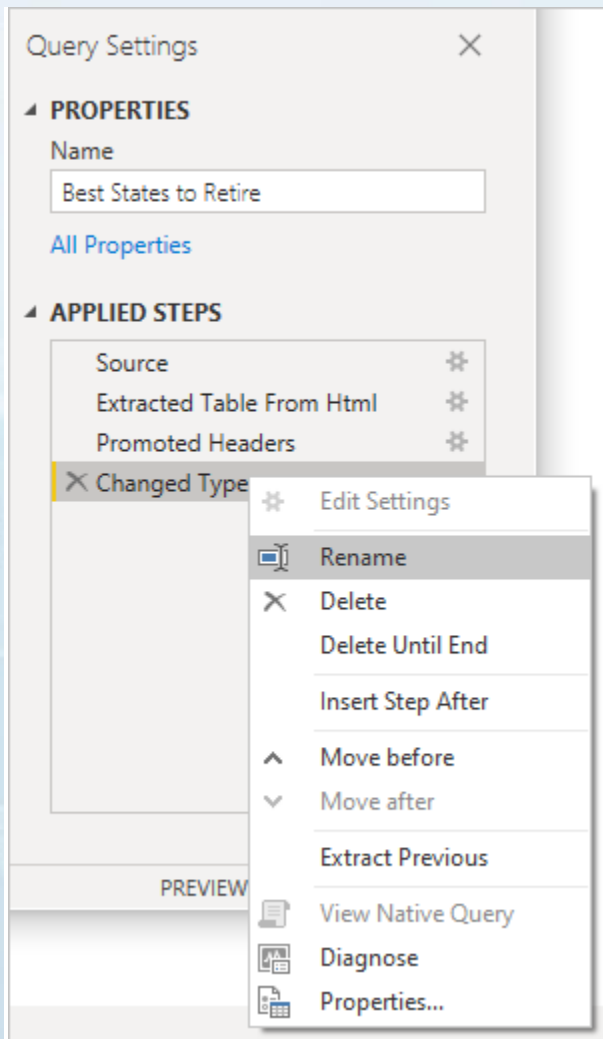
Advanced Editor

The **Advanced Editor** lets you see the code that Power Query Editor is creating with each step. It also lets you create your own shaping code. To launch the advanced editor, select **View** from the ribbon, then select **Advanced Editor**. A window appears, showing the existing query code.
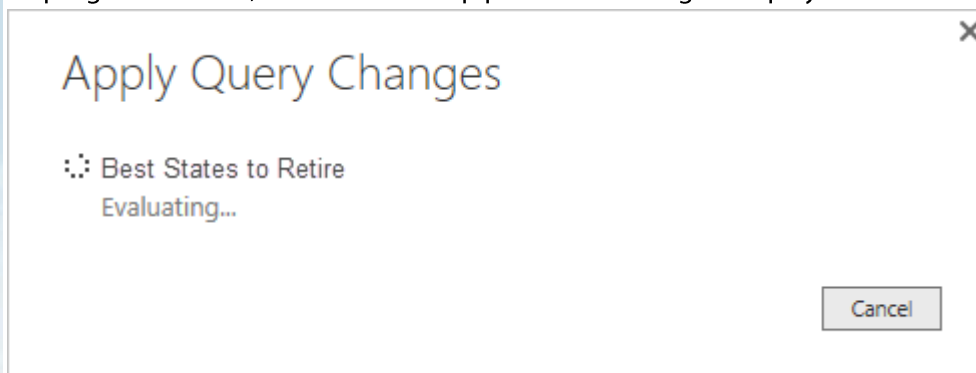
You can directly edit the code in the **Advanced Editor** window. To close the window, select the **Done** or **Cancel** button.

Saving your work

When your query is where you want it, select **Close & Apply** from Power Query Editor's **File** menu. This action applies the changes and closes the editor.
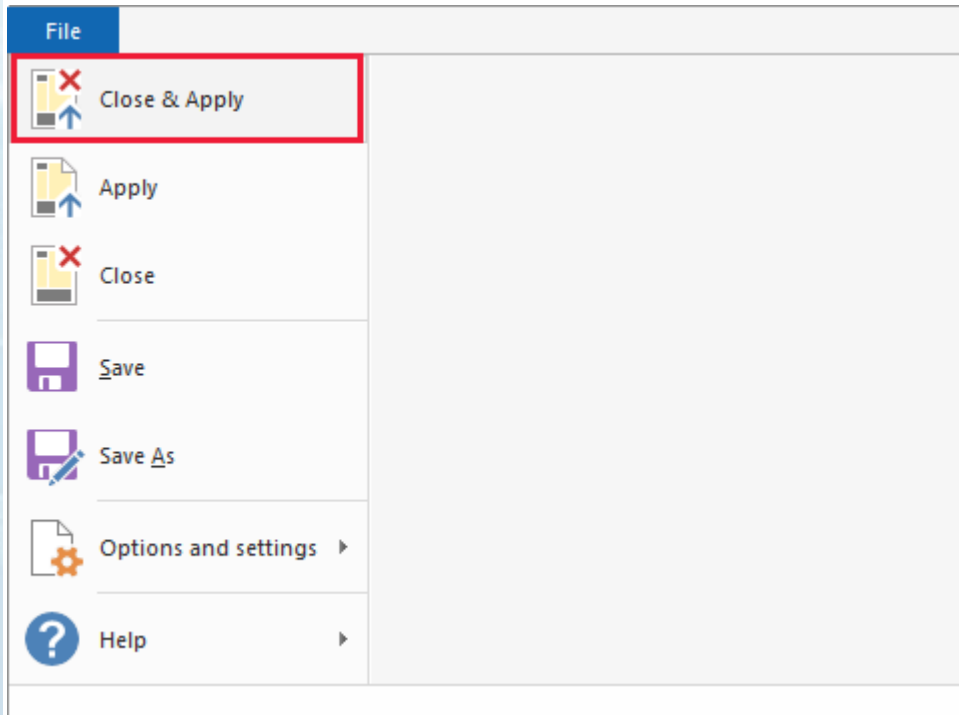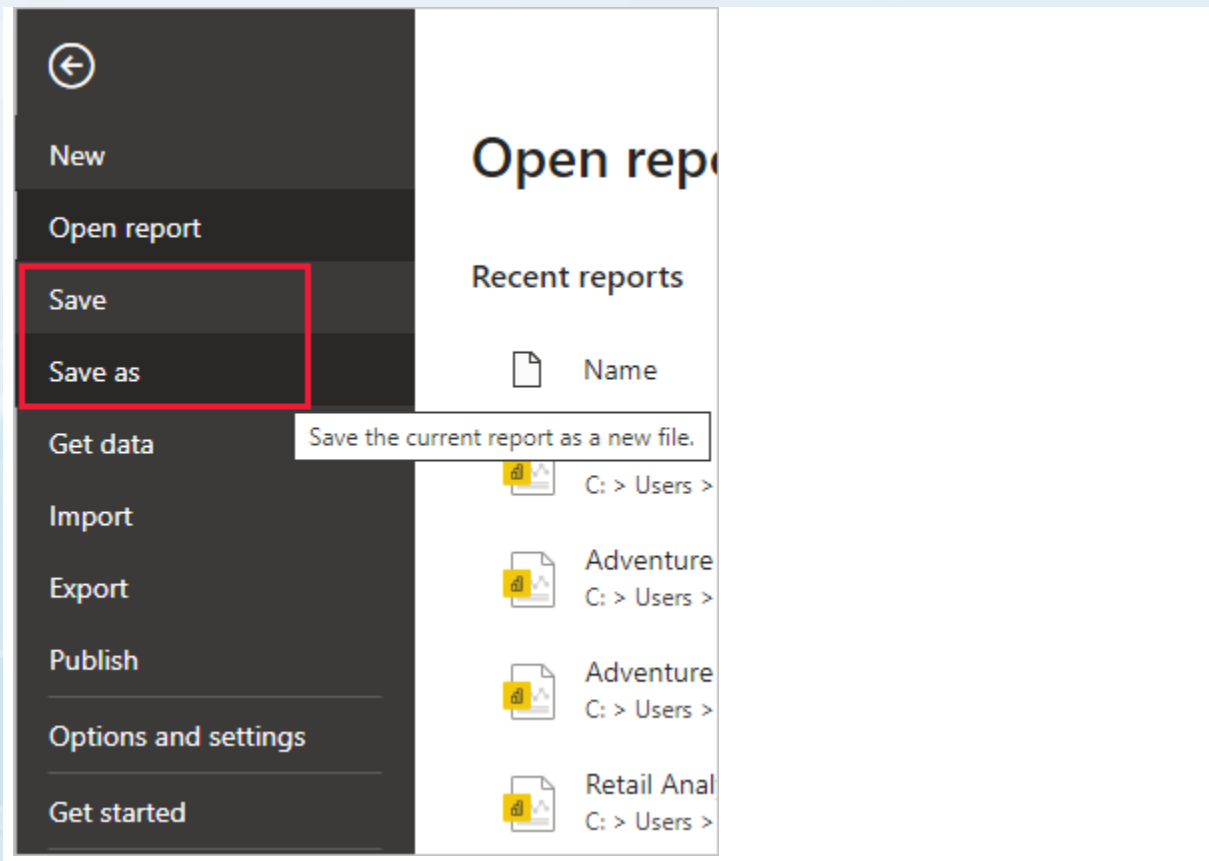
As progress is made, Power BI Desktop provides a dialog to display its status.

When you're ready, Power BI Desktop can save your work in the form of a *.pbix* file.

To save your work, select **File** > **Save** (or **File** > **Save As**), as shown in the following

Image

Uses of Power Query Editor in Power BI

Power Query Editor is a powerful data transformation and cleaning tool in Power BI that helps in preparing data for analysis. Some of the main uses of Power Query Editor in Power BI are:

Data cleaning and transformation: Power Query Editor allows for data cleaning and transformation by removing unwanted columns, splitting and merging columns, changing data types, filtering rows, and more. This ensures that the data is in a usable format for analysis.

Data merging and consolidation: Power Query Editor allows for data merging and consolidation from multiple sources. This helps in combining data from different sources into a single dataset for analysis.

Data filtering and sorting: Power Query Editor allows for data filtering and sorting by applying specific conditions to the data. This helps in narrowing down the data to specific subsets for analysis.

Data shaping and modelling: Power Query Editor allows for data shaping and modelling by creating new columns, aggregating data, and pivoting data. This helps in creating new insights and trends from the data.

Data enrichment and augmentation: Power Query Editor allows for data enrichment and augmentation by adding new columns or data from external sources. This helps in creating more comprehensive and insightful data models.
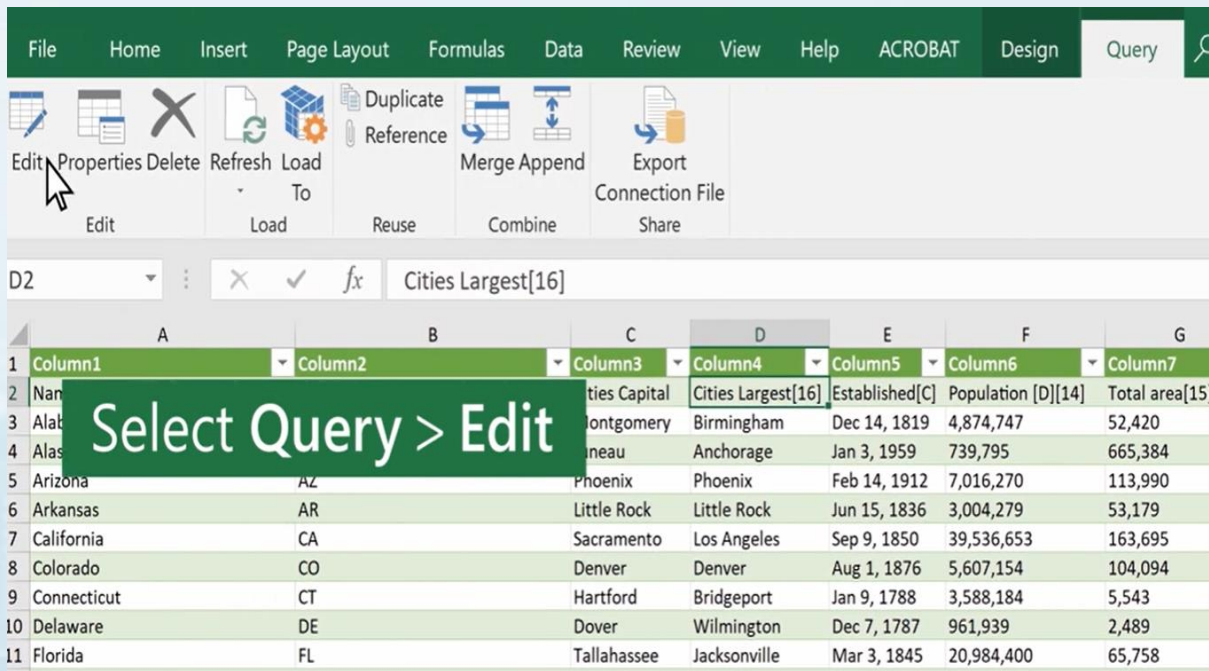
Overall, Power Query Editor in Power BI is a powerful tool for data transformation and cleaning, and it enables users to prepare data for analysis in an efficient and effective way.

Here are the steps to perform the Promote Header, Split to Limiter, Add Columns, Append, and Merge Queries in Power Query Editor in Power BI:

❑ **Promote Header:** If your table has a header row, you can promote it to a table header. To do this, select the header row by clicking on it, and then select the "Transform" tab in the Power Query Editor ribbon. From the "Any Column" dropdown, select "Use First Row As Headers". This will promote the header row to a table header.

Procedure

1. To open a query, locate one previously loaded from the Power Query Editor, select a cell in the data, and then select **Query** > **Edit**.

2. Do one of the following:
   - To promote the first row to column headers, select **Home** > **Use First Row As Headers**.
   - To demote column headers to the first row, select **Home,** select the arrow next to **Use First Row As Headers**, and then select **Use Headers as First Row**.

- ❑ **Split to Limiter:** If you have a column with values separated by a delimiter, you can split them into separate columns using the "Split Column" option. To do this, select the column to be split, and then select the "Transform" tab in the Power Query Editor ribbon. From the "Split Column" dropdown, select "By Delimiter" and specify the delimiter. This will split the column into separate columns based on the delimiter.

- ❑ **Add Columns:** You can add new columns to a table using the "Add Column" option. To do this, select the "Add Column" tab in the Power Query Editor ribbon, and then select the type of column you want to add. You can choose to add a custom column, a conditional column, or a calculated column. This will add a new column to the table based on your selection.
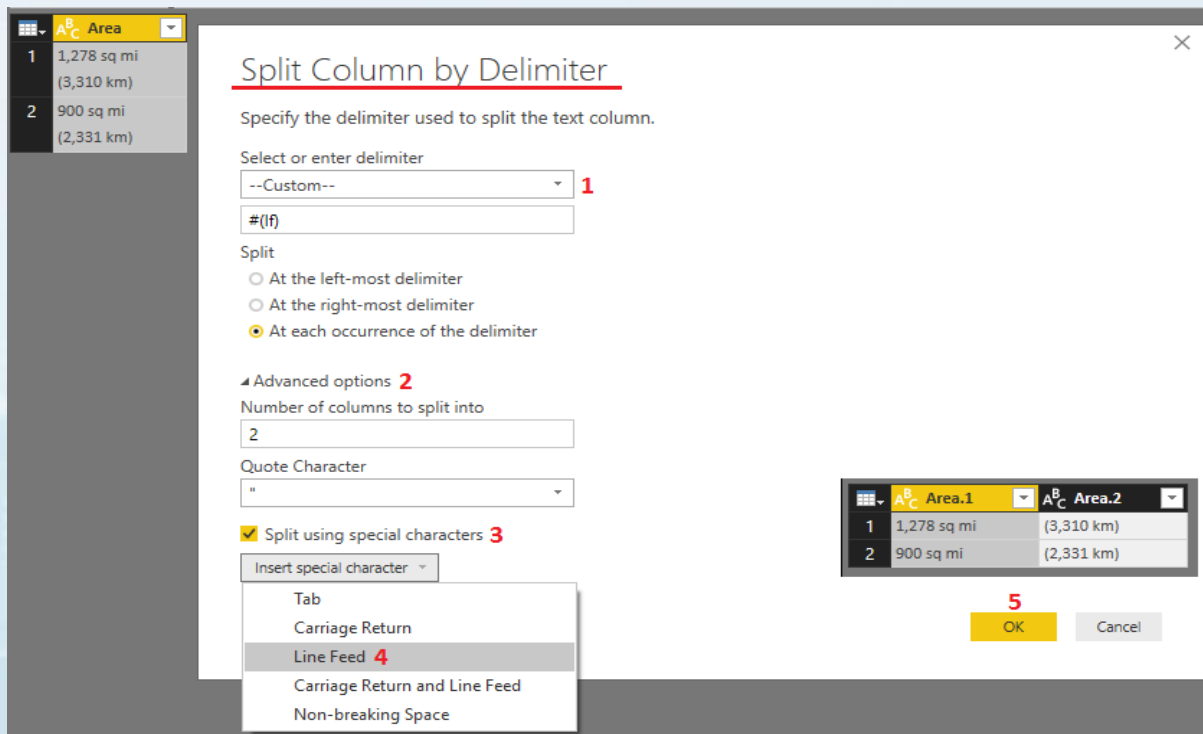
You can split a column with a text data type into two or more columns by using a common delimiter character. For example, a Name column that contains values written as <LastName>, <FirstName> can be split into two columns using the comma (,) character.

1. To open a query, locate one previously loaded from the Power Query Editor, select a cell in the data, and then select **Query** > **Edit**.

2. Select the column you want to split. Ensure that it is a text data type.

3. Select **Home** > **Split Column** > **By Delimiter**. The **Split a column by delimiter** dialog box appears.

4. In the **Select or enter a delimiter** drop-down, select **Colon**, **Comma**, **Equals Sign**, **Semicolon**, **Space**, **Tab**, or **Custom**. You can also select **Custom** to specify any character delimiter.

5. Select a **Split at** option:

   a. **Left-most delimiter**   If there are several delimiters, the first split column is based on the delimiter farthest to the left and the second split column is based on the rest of the characters on its right.

   b. **Right-most delimiter**   If there are several delimiters, the second split column is based on the delimiter farthest to the right and the first split column is based on the rest of the characters on its left.

   c.   **Each occurrence of the delimiter**   If there are several delimiters, split each column by the delimiter. For example, if you have three delimiters, you end up with four columns.

6.   Select **Show advanced options**, and the enter the number of columns or rows to split into.

7.   If you choose **Custom** in **Select or enter a delimiter** drop-down list, you may need to enter an alternative quote character or a special character.

8.   Select **OK**.

9.   You may want to rename the new columns to more meaningful names.

❑   **Append:** You can append tables vertically using the "Append Queries" option. To do this, select the tables to be appended, and then select the "Combine" tab in the Power Query Editor ribbon. From the "Combine" dropdown, select "Append Queries". This will append the tables vertically, creating a new table with all the data.
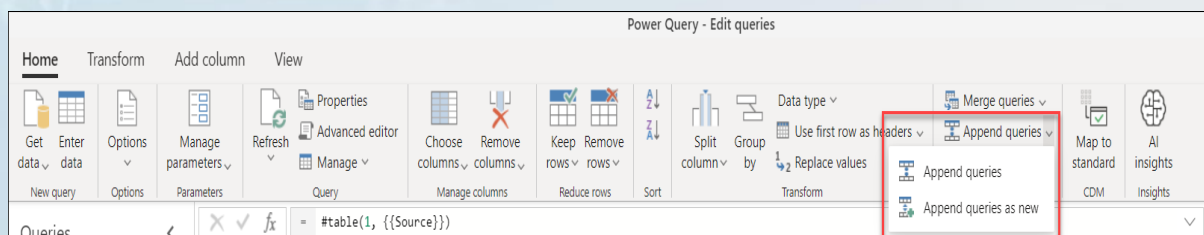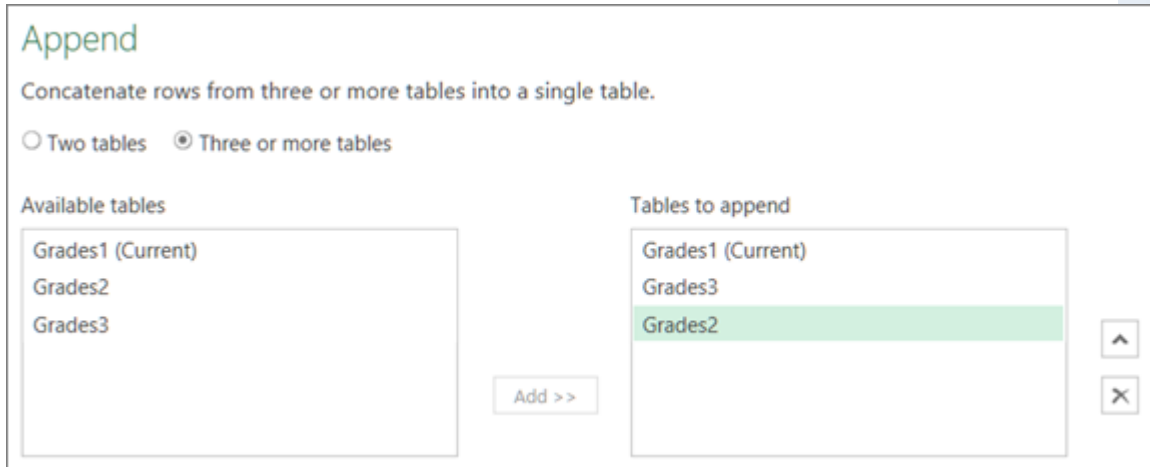
1.   To open a query, locate one previously loaded from the Power Query Editor, select a cell in the data, and then select **Query** > **Edit**.

2.   Select **Home** > **Append Queries**. The default action is to do an inline append. To do an intermediate append, select the arrow next to the command, and then select **Append Queries as New**.
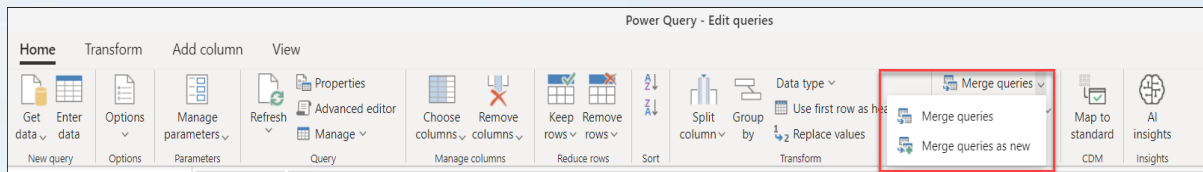
The **Append** dialog box appears.

3. Decide the number of tables you want to append:

   a. Select **Two tables**, and then select the second table in the drop down list box to append.

   b. Select **Three or more tables**. From the **Available tables** box, add the tables you want to append to the **Tables to append**. Use the arrows on the right of that box to change sequence.

4. Select **OK**.

   Result

   ○ If you chose to do an inline append in step 2, a new step in the current query is created. You can continue adding steps to the same query to append additional queries.

   ○ If you chose to do an intermediate append in step 2, a new query is created. You can continue creating additional queries.

❑ **Merge Queries:** You can merge tables horizontally using the "Merge Queries" option. To do this, select the tables to be merged, and then select the "Combine" tab in the Power Query Editor ribbon. From the "Combine" dropdown, select "Merge Queries". Specify the join type and the join columns, and then select the columns to be included in the merged table. This will merge the tables horizontally, creating a new table with the merged data.

A merge query creates a new query from two existing queries. One query result contains all columns from a primary table, with one column serving as a single column containing a relationship to a secondary table. The related table contains all rows that match each row from a primary table based on a common column value. An Expand operation adds columns from a related table into a primary table.

There are two types of merge operations:

- **Inline Merge**   You merge data into your existing query until you reach a final result. The result is a new step at the end of the current query.

- **Intermediate Merge**   You create a new query for each merge operation.

To see a visual representation of the relationships in the **Query Dependencies** dialog box, select **View** > **Query Dependencies**. At the bottom of the dialog box, select the **Layout** command to control the diagram orientation.
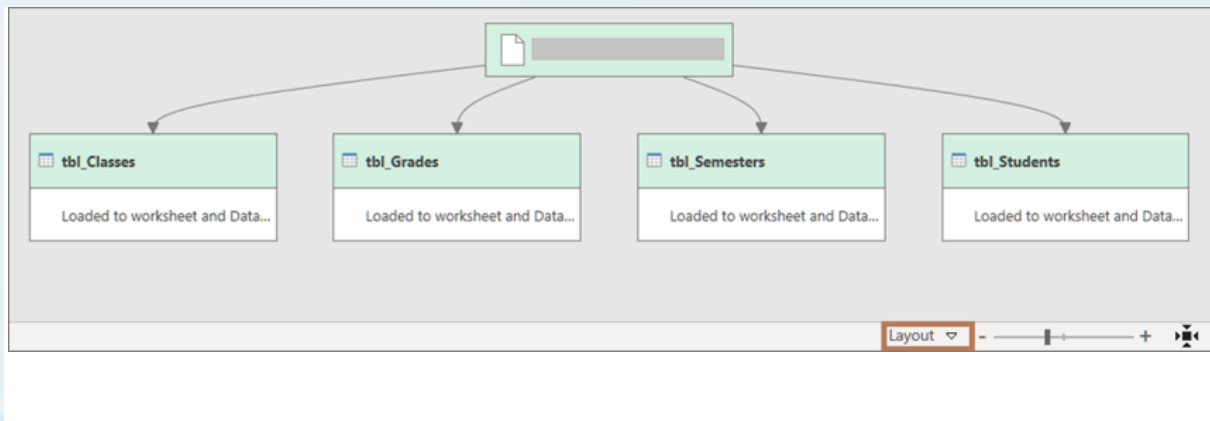
Perform a Merge operation

You need at least two queries that can be merged and that have at least one or more columns to match in a join operation. They can come from different types of external data sources. The following example uses Products and Total Sales.

1. To open a query, locate one previously loaded from the Power Query Editor, select a cell in the data, and then select **Query** > **Edit**.

2. Select **Home** > **Merge Queries**. The default action is to do an inline merge. To do an intermediate merge, select the arrow next to the command, and then select **Merge Queries as New**.

   The **Merge** dialog box appears.

3. Select the primary table from the first drop-down list, and then select a join column by selecting the column header.

4. Select the related table from the next drop-down list, and then select a matching column by selecting the column header.

   Ensure that you select the same number of columns to match in the preview of the primary and related or secondary tables. Column comparison is based on the order of selection in each table. Matching columns must be the same data type, such as **Text** or **Number**. You can also select multiple columns to merge.

5. After you select columns from a primary table and related table, Power Query displays the number of matches from a top set of rows. This action validates whether the **Merge** operation was correct or whether you need to make changes to get the results you want. You can either select different tables or columns.

6. The default join operation is an inner join, but from the **Join Kind** drop down list, you can select the following types of join operations:

Inner join   Brings in only matching rows from both the primary and related tables.

Left outer join   Keeps all the rows from the primary table and brings in any matching rows from the related table.

Right outer join   Keeps all the rows from the related table and brings in any matching rows from the primary table.

Full outer   Brings in all the rows from both the primary and related tables.

Left anti join   Brings in only rows from the primary table that don't have any matching rows from the related table.

Right anti join   Brings in only rows from the related table that don't have any matching rows from the primary table.

Cross join   Returns the Cartesian product of rows from both tables by combining each row from the primary table with each row from the related table.

7. If you want to do a fuzzy match, select **Use fuzzy matching to perform the merge** and select from the **Fuzzy Matching options**.

8. To include only those rows from the primary table that match the related table, select **Only include matching rows**. Otherwise, all rows from the primary table are included in the resulting query.
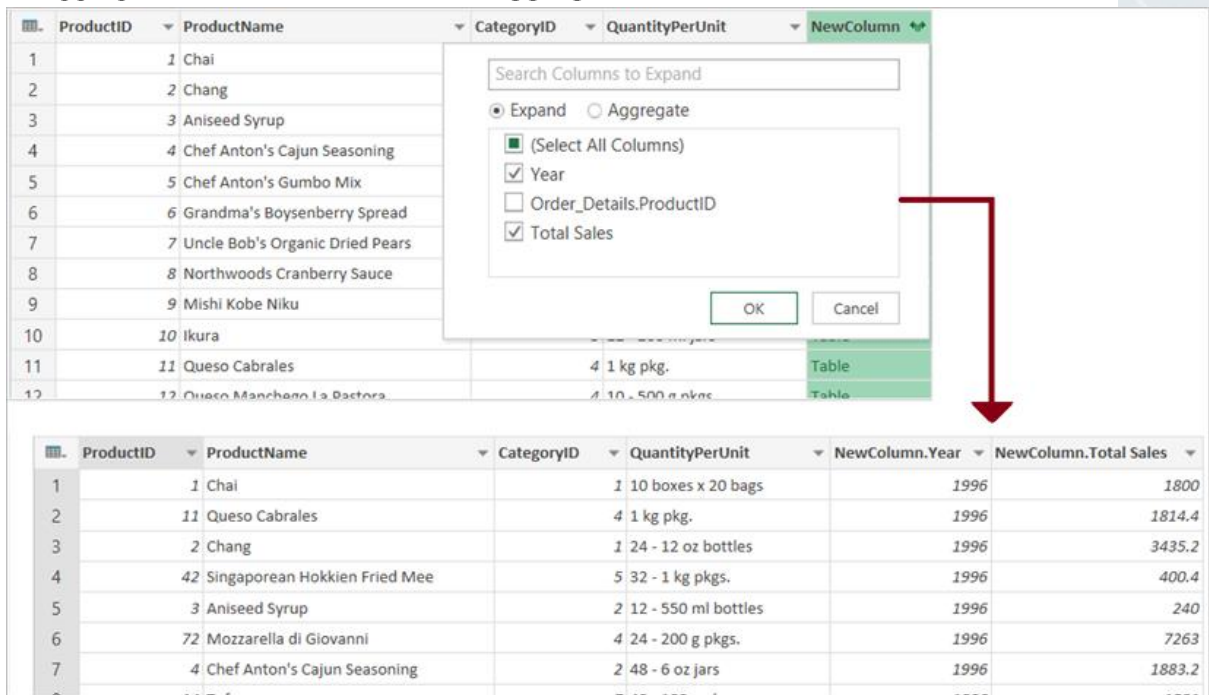
9. Select **OK**.

Result

fx = Table.NestedJoin(Products,{"ProductID"},#"Total Sales",{"Order_Details.ProductID"},"NewColumn")

ProductID	ProductName	CategoryID	QuantityPerUnit	NewColumn
1	Chai	1	10 boxes x 20 bags	Table
2	Chang	1	24 - 12 oz bottles	Table
3	Aniseed Syrup	2	12 - 550 ml bottles	Table
4	Chef Anton's Cajun Seasoning	2	48 - 6 oz jars	Table
5	Chef Anton's Gumbo Mix	2	36 boxes	Table
6	Grandma's Boysenberry Spread	2	12 - 8 oz jars	Table

Expand the Table column

After a **Merge** operation, you can expand the **Table** structured column to add columns from the related table into the primary table. Once a column is expanded into the primary table, you can apply filters and other transform operations.

1. In the Data Preview, select the **Expand** 〜 icon next to the **New Column** header.

2. In the **Expand** drop-down box, select or clear the columns to display the results you want. To aggregate the column values, select **Aggregate.**

3. You may want to rename the new columns.