**Activity: Import a dataset with best practices**

[Product](https://d3c33hcgiwev3.cloudfront.net/1LjV1xvPTPCmHYNCYCE9Qw_adab8ec5230d435c8f76cec01761cce1_Product.csv?Expires=1709856000&Signature=QGsbJ0Myq5WaUzdUABjAzxc5S9-TwbfbqIrACh-mC3xr0LvTnLjUWmguVKZL4rlN6NamiuBIbrq39wcMX0u3SWfvILadVq~zooF~gMLTpwYjMXOATSdcRqjOlLRiW9UV0gpt0Fhe~~vhSB1RdDvL9GOrANlaB9fqJBlMBHpPY1U_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[CSV File](https://d3c33hcgiwev3.cloudfront.net/1LjV1xvPTPCmHYNCYCE9Qw_adab8ec5230d435c8f76cec01761cce1_Product.csv?Expires=1709856000&Signature=QGsbJ0Myq5WaUzdUABjAzxc5S9-TwbfbqIrACh-mC3xr0LvTnLjUWmguVKZL4rlN6NamiuBIbrq39wcMX0u3SWfvILadVq~zooF~gMLTpwYjMXOATSdcRqjOlLRiW9UV0gpt0Fhe~~vhSB1RdDvL9GOrANlaB9fqJBlMBHpPY1U_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[Reseller](https://d3c33hcgiwev3.cloudfront.net/BWROa5BIQ3u_s9EITkF-fA_575d5a8c76a14b8fb0d2a614cb1ef1e1_Reseller.csv?Expires=1709856000&Signature=ct6Yg7dHyjhUBamQ9JgFFmnvumurMsFq6btYtDFZ8w~fvb1XnI5XaZkwXnXVb~gQhJBWpyYEOChD-P1yqqQkH9AIQ2Z2bn~z1xqQ5DFk6Sb2DsgH1HF3jtHC1NqfaqBGx8wYsXu89O8R~16FSQsv09mq4lLdf7T8mThqeT-Fe5M_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[CSV File](https://d3c33hcgiwev3.cloudfront.net/BWROa5BIQ3u_s9EITkF-fA_575d5a8c76a14b8fb0d2a614cb1ef1e1_Reseller.csv?Expires=1709856000&Signature=ct6Yg7dHyjhUBamQ9JgFFmnvumurMsFq6btYtDFZ8w~fvb1XnI5XaZkwXnXVb~gQhJBWpyYEOChD-P1yqqQkH9AIQ2Z2bn~z1xqQ5DFk6Sb2DsgH1HF3jtHC1NqfaqBGx8wYsXu89O8R~16FSQsv09mq4lLdf7T8mThqeT-Fe5M_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[Sales](https://d3c33hcgiwev3.cloudfront.net/R2Tr4TzkR7aY152ZFgaikQ_f768951bd2ef4d43909b4aac69c4c3e1_Sales.csv?Expires=1709856000&Signature=XaInULJgu7NrX1YekguSj4F-jjMK9RW5GpkOzZZwaT32QhX-FW8FHoN4topNGbsSedPXMGwqra716oLIUHQ-9LpmzgOKlwd3gRn1GRelwGbyNxrRsyniMjj5NhK84ObZ-M9KHjlHaAboPSFocJ2YfpVUAyLlk8U-V~IIpC1bPBU_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[CSV File](https://d3c33hcgiwev3.cloudfront.net/R2Tr4TzkR7aY152ZFgaikQ_f768951bd2ef4d43909b4aac69c4c3e1_Sales.csv?Expires=1709856000&Signature=XaInULJgu7NrX1YekguSj4F-jjMK9RW5GpkOzZZwaT32QhX-FW8FHoN4topNGbsSedPXMGwqra716oLIUHQ-9LpmzgOKlwd3gRn1GRelwGbyNxrRsyniMjj5NhK84ObZ-M9KHjlHaAboPSFocJ2YfpVUAyLlk8U-V~IIpC1bPBU_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

**Introduction**

As a data analyst, you play a vital role in analyzing large datasets to gain insights and improve business operations. Microsoft Power Query is an essential tool in your workflow, allowing you to transform and integrate data from various sources. At this point, you should have a good understanding of the best practices when using Microsoft Power Query and its practical applications. You gained valuable knowledge on planning, documentation, connectors, data profiling, performance, error handling, using groups, and optimizing data types. In this step-by-step exercise, you have the opportunity to apply best practices when importing datasets in Power BI.

**Case study**

You are working as a data analyst at Adventure Works which sells various consumer products. As the business expands, the management team wants to gain deeper insights into their sales performance across different regions. You use Power BI to import and analyze the sales data.

The company uses CSV files containing the Adventure Works company sales, reseller and product data. You will be applying the best practices for importing data.

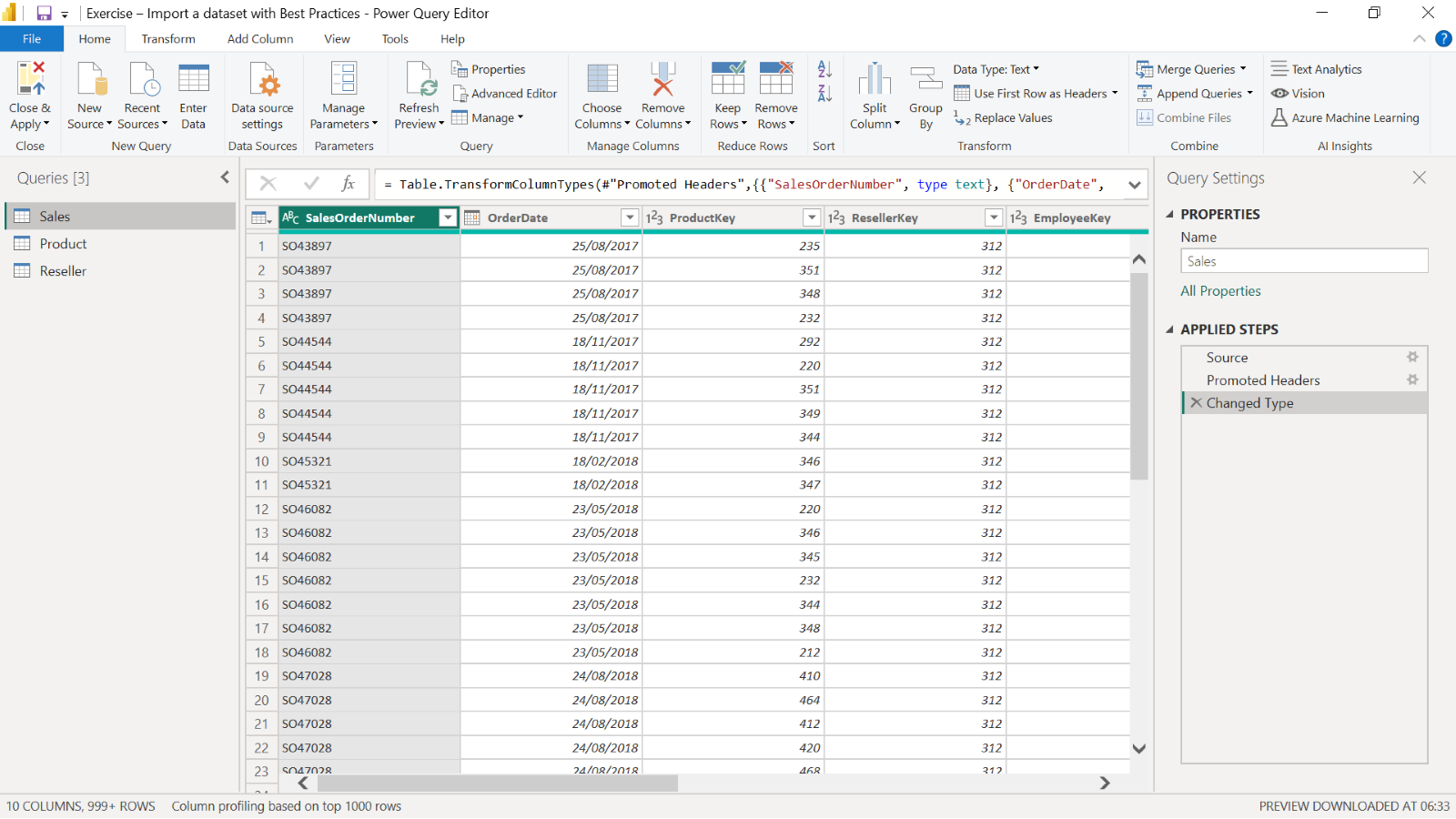
By the end of this activity, you’ll have a good understanding of how to apply best practices when importing datasets into Power BI.

**Instructions**

Create a new Power BI project called *Exercise – Import a dataset with Best Practices*. Follow the prompts below to complete the exercise.

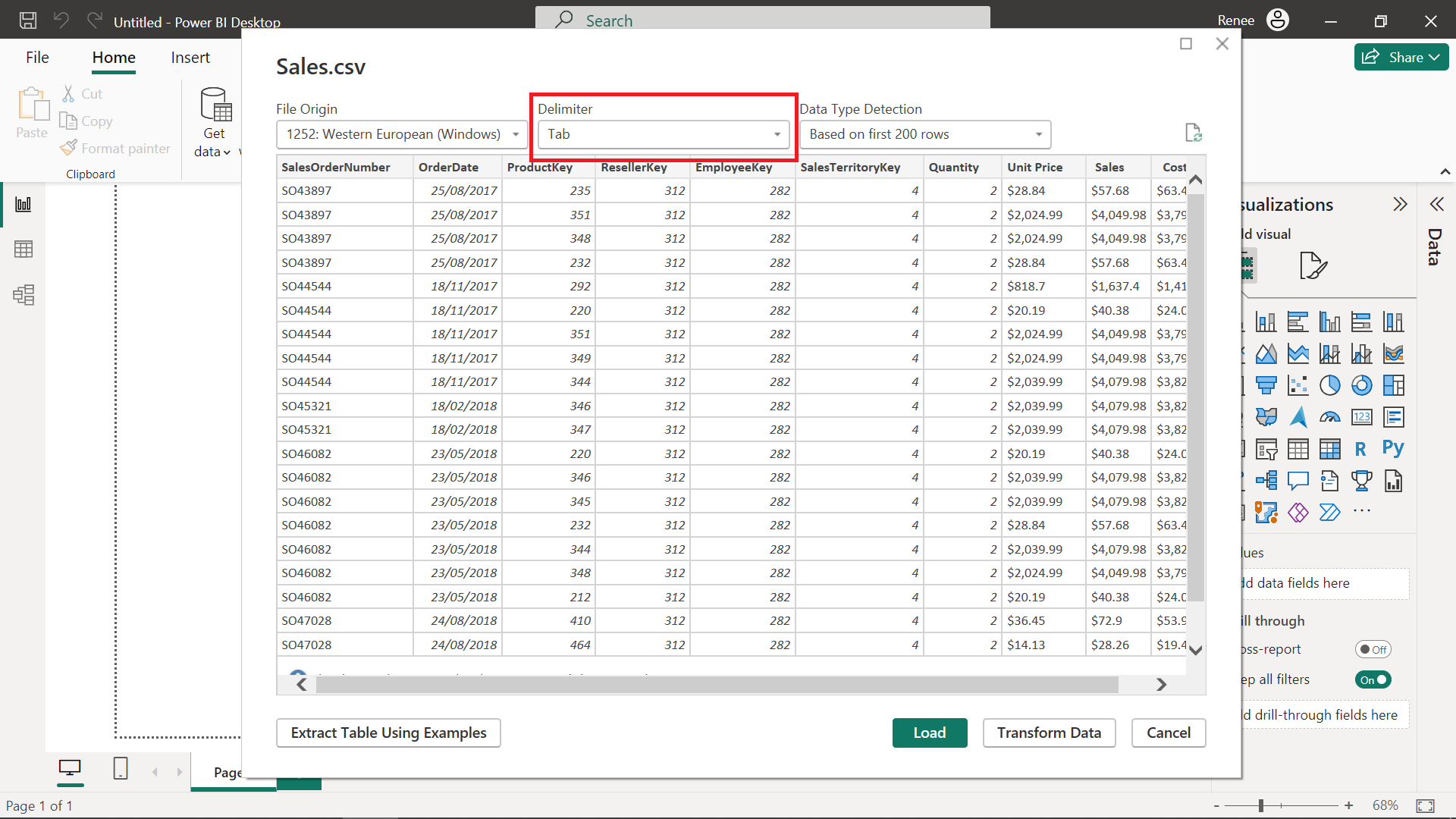
**Step 1: Download the CSV files**

* Download the *Sales.csv*, *Product.csv* and *Reseller.csv* files, which will be used in this exercise. These files are available at the top of this article.



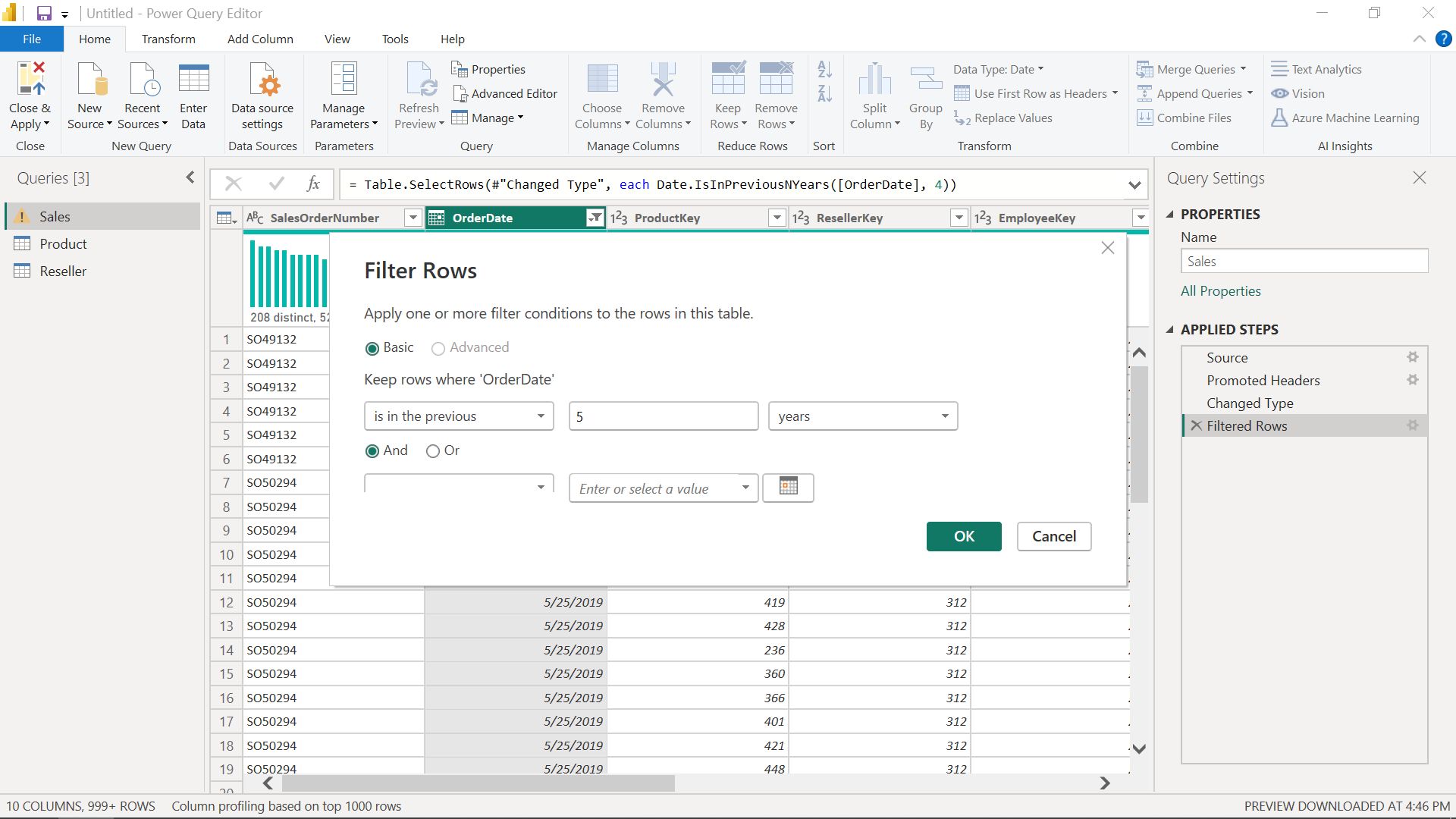
**Step 2: Open the Power Query editor**

* Open the Power Query Editor and import your datasets, *Sales*, *Product*, and *Reseller* using the Text/CSV option. **Tip:** If the columns do not appear correctly, ensure that the Delimiter is set to Tab.



**Step 3: Optimize data extraction**

1. To optimize data extraction, apply filters to retrieve the sales data for the past five years only.
2. Select the **OrderDate** column, select the down arrow at the right side of the column and select the **Date Filters** menu.
3. Then select **Custom Filter** on the bottom of the list to select the last five years' records.
4. In the *Keep rows where* the '**OrderDate**' field boxes, select “is in the previous,” “5” and “years” in the respective sections to filter the dataset.



**Step 4: Transform and clean data**

Using **Power Query,** transform and clean the data.

1. Find the empty rows in the **OrderDate** column
2. Find the error rows in the **ProductKey** column
3. Remove errors in these columns.

Note the results of performing these operations.

**Step 5: Profile data**

1. Examine the data in distribution, quality and profile.
2. Find error and empty rows, calculate minimum, maximum, average, count and other statistics, and compare distinct and unique values in different sales columns.

There is no error and empty values everything is valid

1. Find the minimum, maximum, and average values of the **UnitPrice** column

A screen shot of a computer

Description automatically generated

A white background with black text

Description automatically generated

1. Find distinct and unique values in **the Product Name** column. Compare the distinct and unique values and find the reason for that.

A screenshot of a bar code

Description automatically generated

1. Find the **Count value** of the **SalesOrderID** column.

1000

1. Write down the values that you found.

**Conclusion**

By following these best practices, you successfully import Adventure Works sales data into Power BI, enabling you to analyze and profile sales performance across different regions. You gain best practices for optimizing data extraction, transforming, cleansing, and profiling data.