**Improving data model performance**

**Overview**

In the exercise *Improving Data Model Performance*, you were asked to fix a slow-loading Power BI report at Adventure Works. To complete this task, you had to identify bottlenecks in the unoptimized data model and implement changes to improve its efficiency and the overall performance of the report.

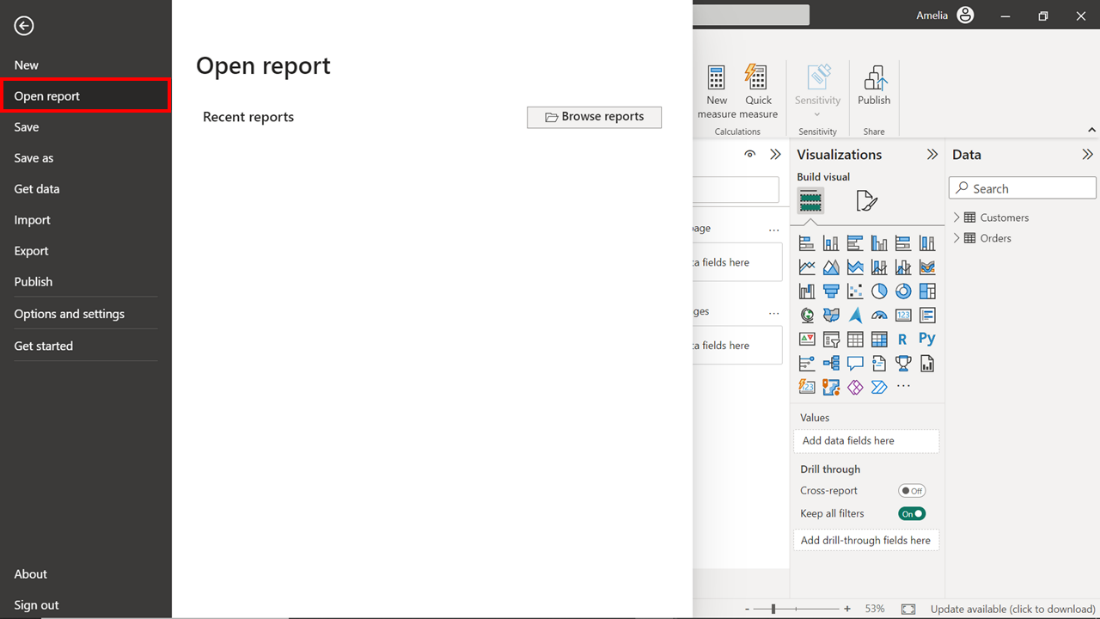
More specifically, you were asked to:

* Download the Adventure Works Power BI report titled *AdventureWorksSales.pbix* and open it in Power BI Desktop.
* Switch to **Data** view to observe the records contained in the report and understand the scope of data that the model is dealing with.
* Switch to **Model** view and change the relationship between the **Customers** and **Orders** tables from **Many-to-many** to **One-to-many** to simplify the relationship and improve performance.
* Save the changes and check to ensure that they have been implemented correctly and that the updated data model aligns with the adjustments.

This reading provides you with a guide that you can use as a benchmark for your solution. You can also refer to the videos *Resolving performance issues in the data model.*

**Step 1: Open Your Project**

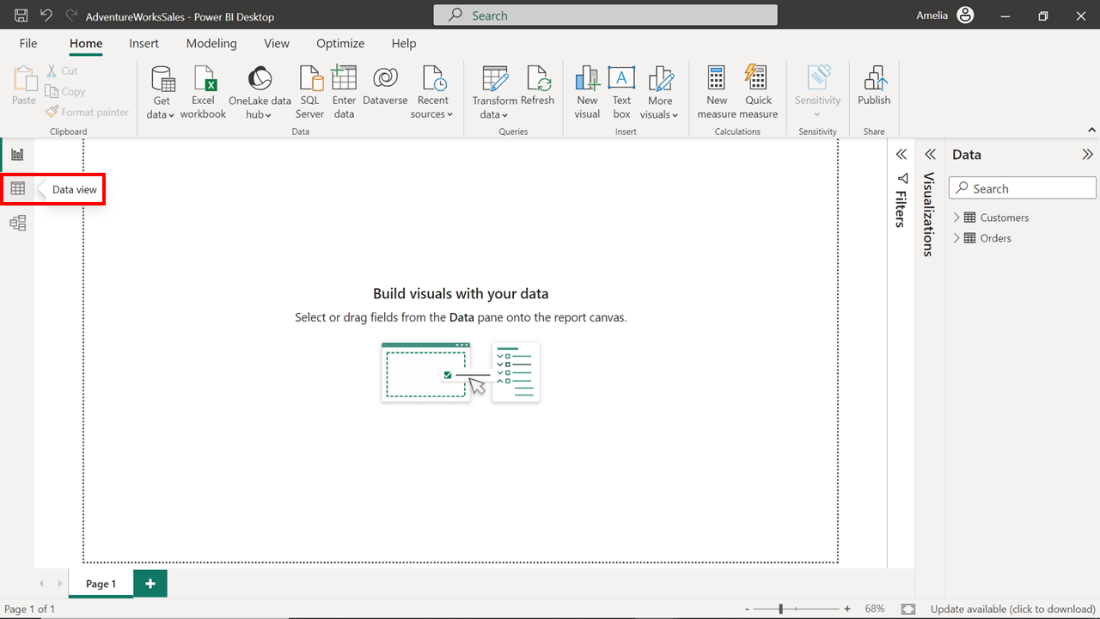
1. In **Power BI Desktop**, select **File** in the top left corner. On the **File** menu, select **Open Report**.



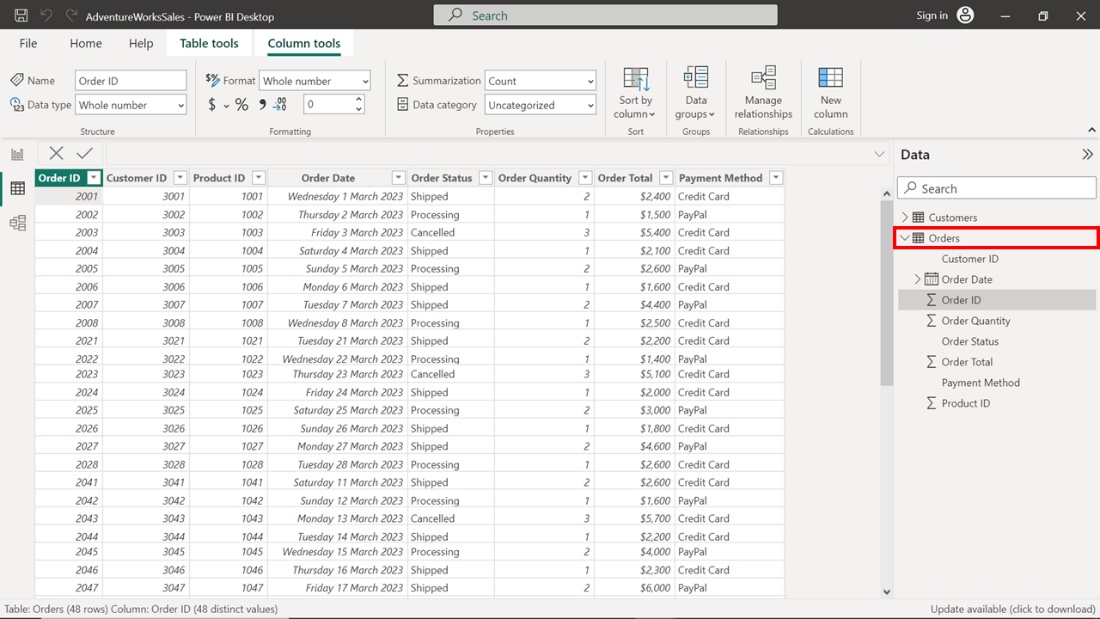
1. A dropdown menu will then appear where you select **Browse Reports**. Selecting this causes the file explorer window to open. Navigate to the location where you’re the downloaded file *AdventureWorksSalesReport.pbix* is saved.
2. Select the file and click **Open** in the **File explorer** window. This action opens the saved project in the **Power BI Desktop** application.

**Step 2: Go to Data View**

1. In the **Power BI Desktop** window, you'll find a vertical toolbar with different icons on the left-hand side. The second icon from the top resembles a table and is the **Data** view icon. Select this icon to switch to **Data v**iew.

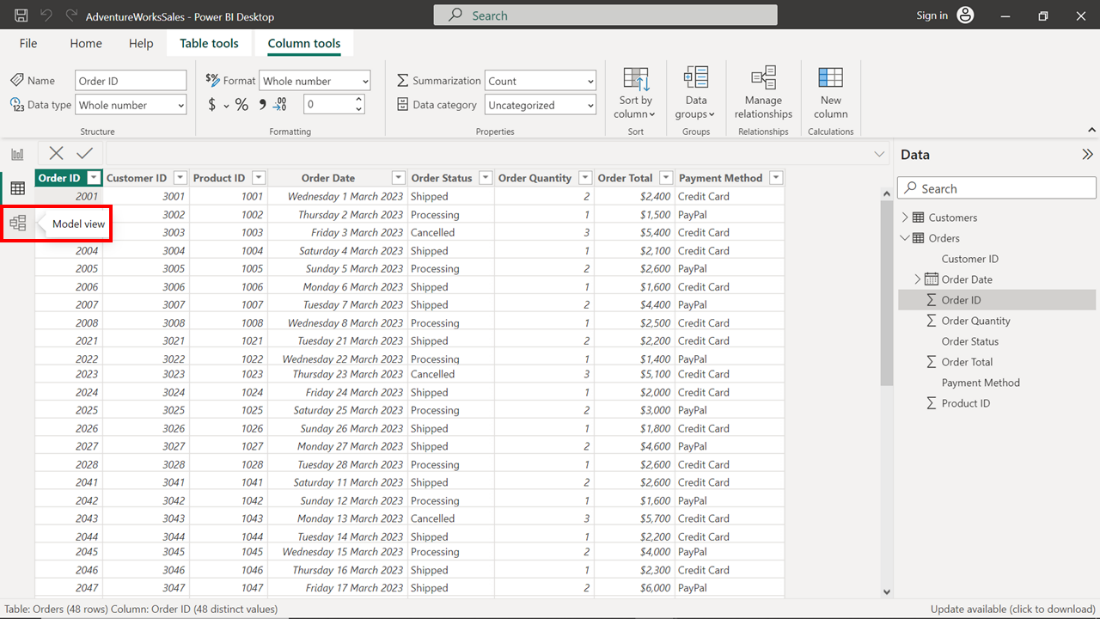


1. Power BI switches to **Data view** and displays the data contained in the project. Viewing this data can help you to understand the granularity and structure of the data at the most basic level. This information about the data and its structures is crucial for any kind of data analysis or data modeling. Select the **Orders** table on the right of the screen and take a moment to observe the first ten records. The order identified by **Order ID 2003** carries the highest **Order Total** value of **$5400** amongst the first ten records. This could result from various factors such as the quantity of goods ordered, the individual cost of each item, and the application of any taxes, fees, or discounts. It also suggests efficient transaction processing and could indicate a high-value customer.



**Step 3: Navigate to Model View**

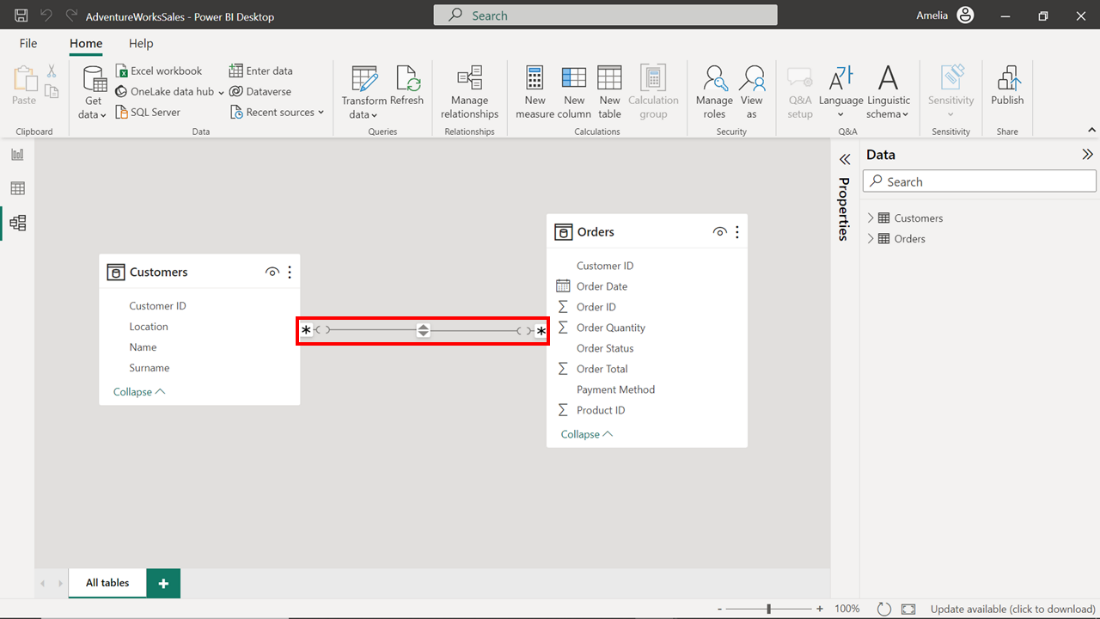
1. The next step asks that you switch to **Model view.** To do this, select a different icon in the vertical toolbar on the left side of the Power BI interface.  Choosing the third icon from the top switches to **Model view**. This icon shows three tables linked with connectors. Select this icon.



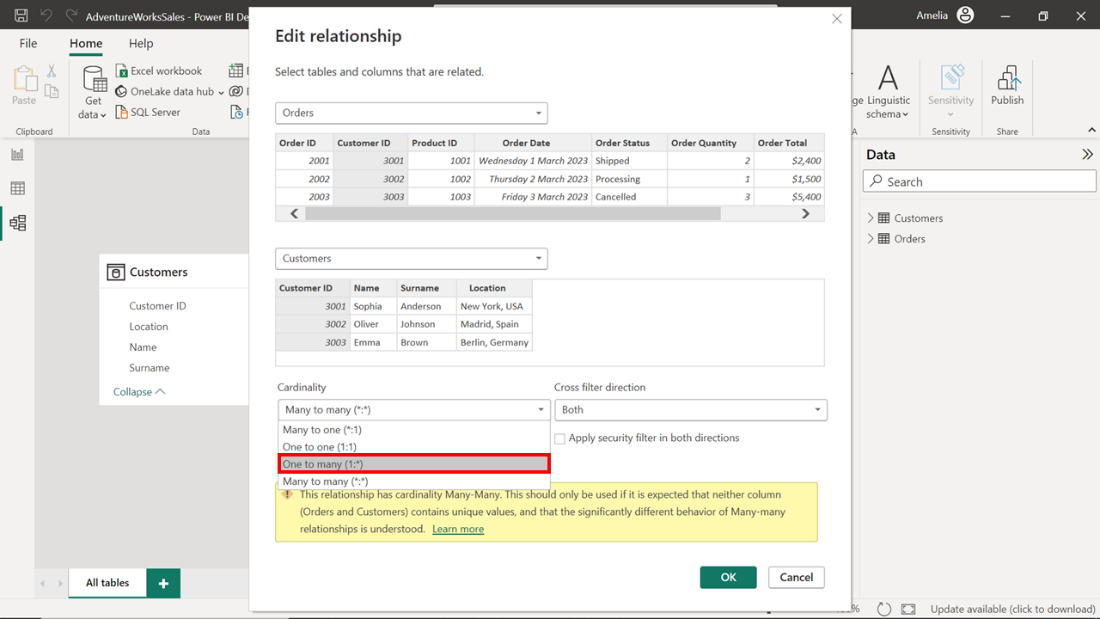
1. **Model view** displays a graphical representation of the tables in your report and the relationships between them. You can use this high-level overview to quickly identify and understand how different tables are interconnected. This is crucial when making modifications to improve performance.

**Step 4: Select Relationships**

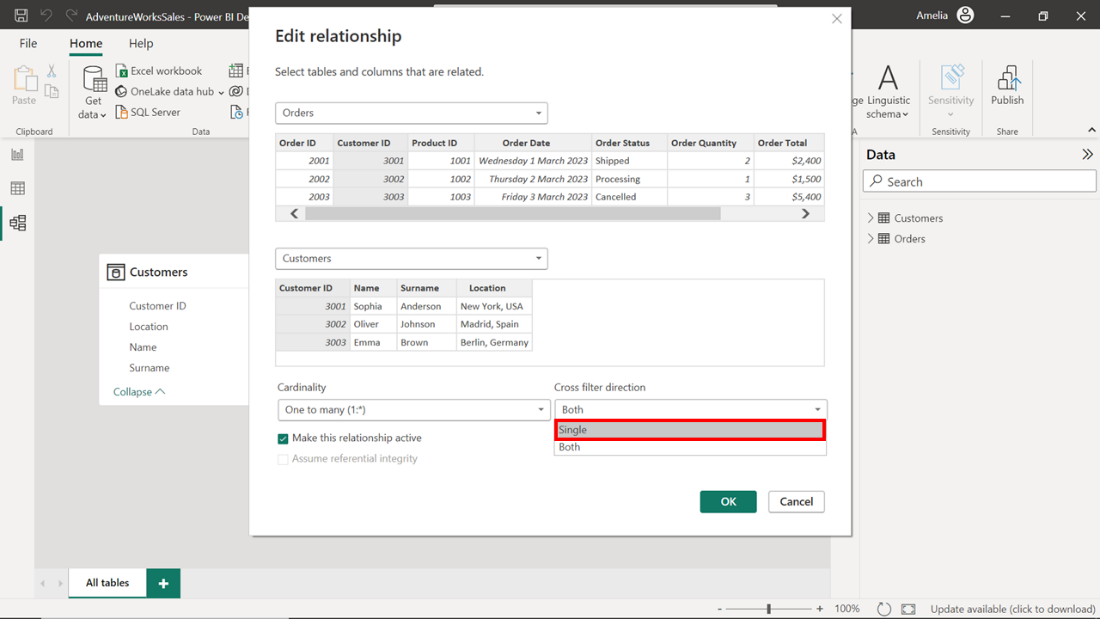
1. **Model view** displays a network of tables connected by lines that represent the relationships between the tables. The **Customers** and **Orders** tables line shows a many-to-many relationship (depicted by asterisks on both sides of line) that you must modify. Modification is important because it lets you configure relationships that are simpler and faster for Power BI to navigate when loading data and calculating results.



1. To modify the relationship, double-click on the line to open the **Edit relationship** dialog. The dialog box displays the properties of the two linked tables, **Customers** and **Orders**, and provides options to edit various aspects of the relationship, including **Cross filter direction** and **Cardinality**. To optimize the data model, you'll need to adjust both properties.
2. In the **Cardinality** drop-down, change the selection from its current state of **Many-to-many** to **One-to-many**. Adjusting the **Cardinality** from **Many-to-many** to **One-to-many** can improve performance because **One-to-many** relationships are simpler and faster for Power BI to navigate when loading data and calculating results.

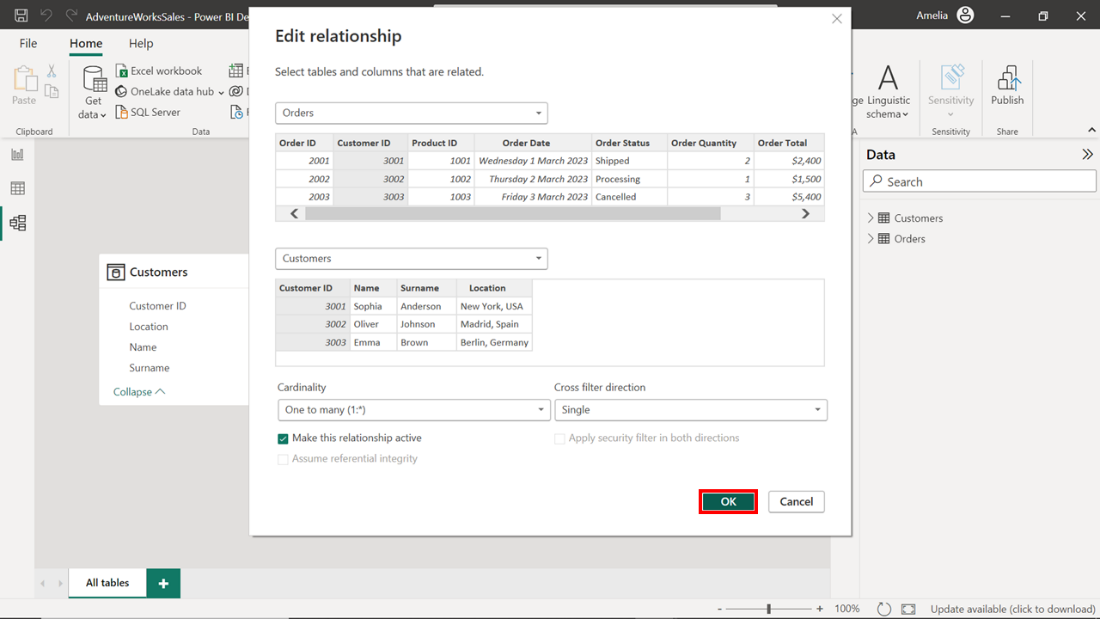


1. In the **Cross-filter direction** drop-down, choose the **Single** option to reduce the complexity of the model and limit the direction in which filters are applied. The logic behind these changes is simple: one customer can have many orders, but each order can belong to only one customer. Adjusting these settings to reflect the actual business relationships will improve the efficiency and accuracy of your data model.

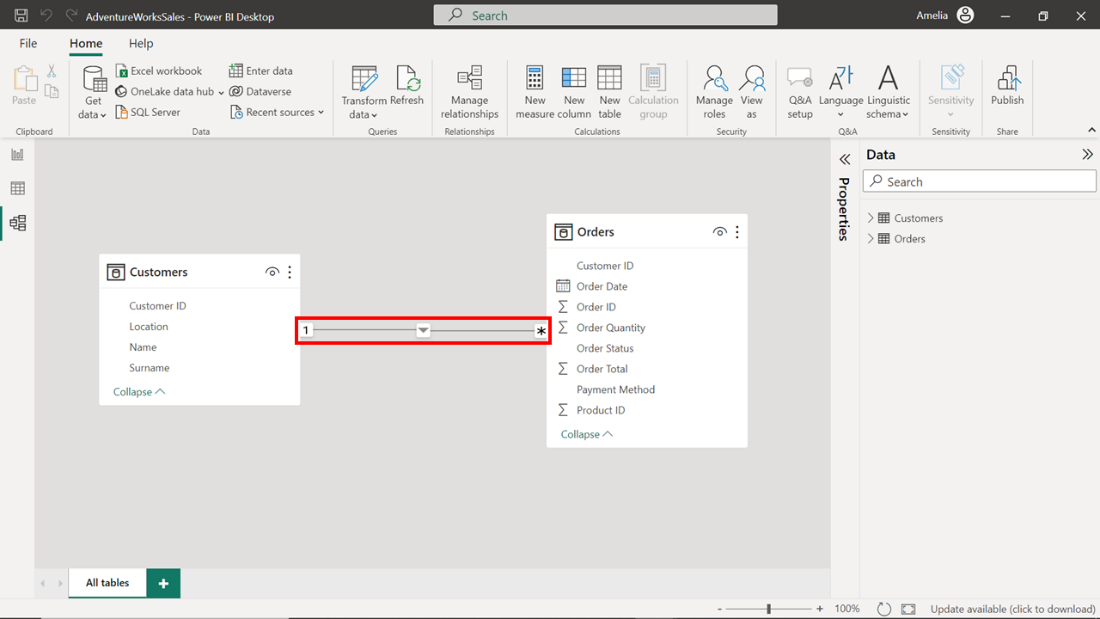


**Step 5: Save the Changes**

1. Once you have modified the relationship properties for **Customers** and **Orders** you must save these changes for them to take effect. To save the changes made to the data model, select the **OK** button located at the bottom right corner of the dialog box.



1. Once the changes are saved, review and confirm that they have been implemented by checking the relationships in the **Model view.** The line connecting **Customers** to **Orders** should now display the number one (**1**) attached to the **Customers** table and an asterisk (**\***) attached to the **Orders** table. This confirms that the relationship has been set to **One-to-many** as intended.



**Conclusion**

The changes you have made to the data model in this project file will improve the quality and efficiency of the report. Streamlining the relationships in the model allows Power BI to process the data efficiently and generate visuals more quickly. This improved performance will benefit your colleagues in the Sales department who first flagged the poorly performing report. Well done!

**Optimizing columns and metadata**

**Introduction**

Microsoft Power BI contains various features that allow you to shape, transform, and model data to create immersive and interactive insights. However, if the data model is large and complex, the data manipulations may seem slower than usual. The solution? You need to optimize your columns and metadata.

This reading focuses on optimizing columns and metadata in Power BI and explores the features provided in Microsoft Power BI for this purpose. It explains how to inspect tables and columns, rename columns, categorize data, change column data types, and add descriptions to fields in Power BI.

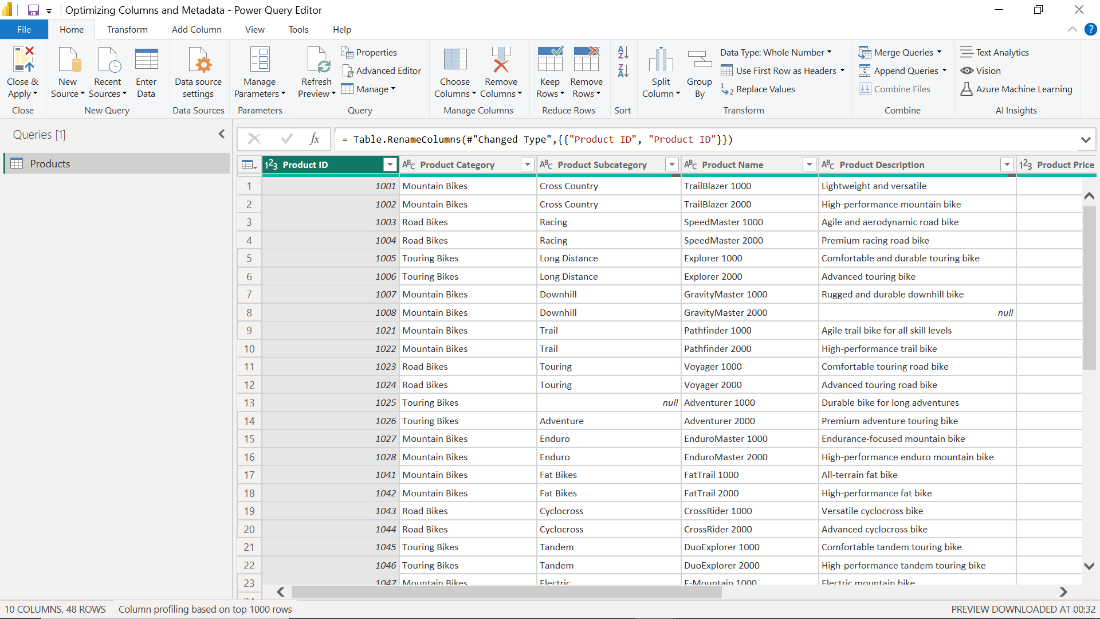
**Optimizing columns and metadata**

Metadata, in the simplest terms, is data about data. It provides information about other data, making it easier to retrieve, manipulate, and manage. Some examples of metadata In Microsoft Power BI, would be table names, column names, relationships, or data types. Metadata acts as a map guiding you through the data landscape, making navigating vast volumes of information easier. Columns are fundamental components of tables in Power BI, which hold data for various attributes or variables. Each column has a specific data type that determines what kind of values it can store. The data type could be text, whole number, decimal number, or date/time.

Optimizing columns and metadata is important. In Power BI, every piece of data stored consumes memory, and the efficiency with which data is stored and retrieved can significantly impact the performance of your data model. An optimized data model ensures a faster, smoother Power BI experience and uses memory more efficiently, contributing to cost-effectiveness. Power BI offers multiple ways to optimize columns and metadata.

**Data Preview**

Understanding the metadata of your columns can significantly improve the quality of your analysis, as it helps identify errors, outliers, patterns, and distributions that might influence your data interpretations. The **Power Query Editor** offers robust options for previewing your data and analyzing metadata, including **Column Quality**, **Column Distribution**, and **Column Profile.**



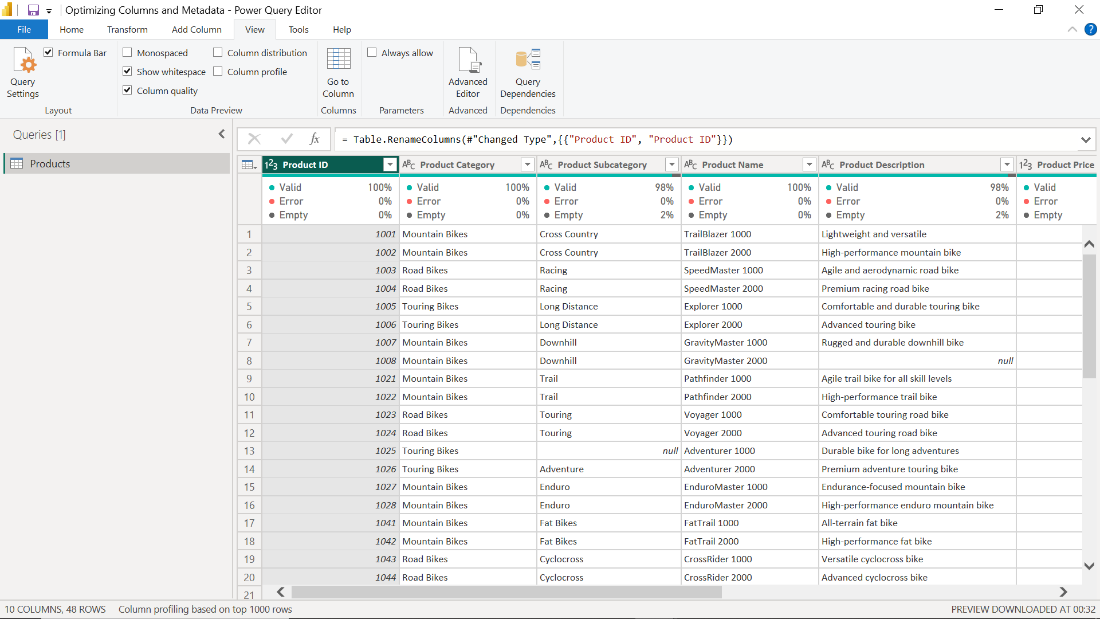
**Column Quality**

This feature allows you to quickly assess the data quality in your columns. It quantifies the percentage of valid, erroneous, or empty entries in your column, giving you a quick overview of potential issues that might need to be addressed. The steps to check **Column Quality** are as follows:

1. Select the **View** tab in the ribbons in the top portion of the **Microsoft Power Query Editor** window.



1. In the **Data Preview** group, select the **Column Quality** checkbox to enable the feature.



1. A **Column Quality** pane opens below each column header in the data preview. It displays a column's percentage of valid, error, and empty entries. A **Valid** percentage of less than 100% indicates the presence of errors or empty cells that need addressing.

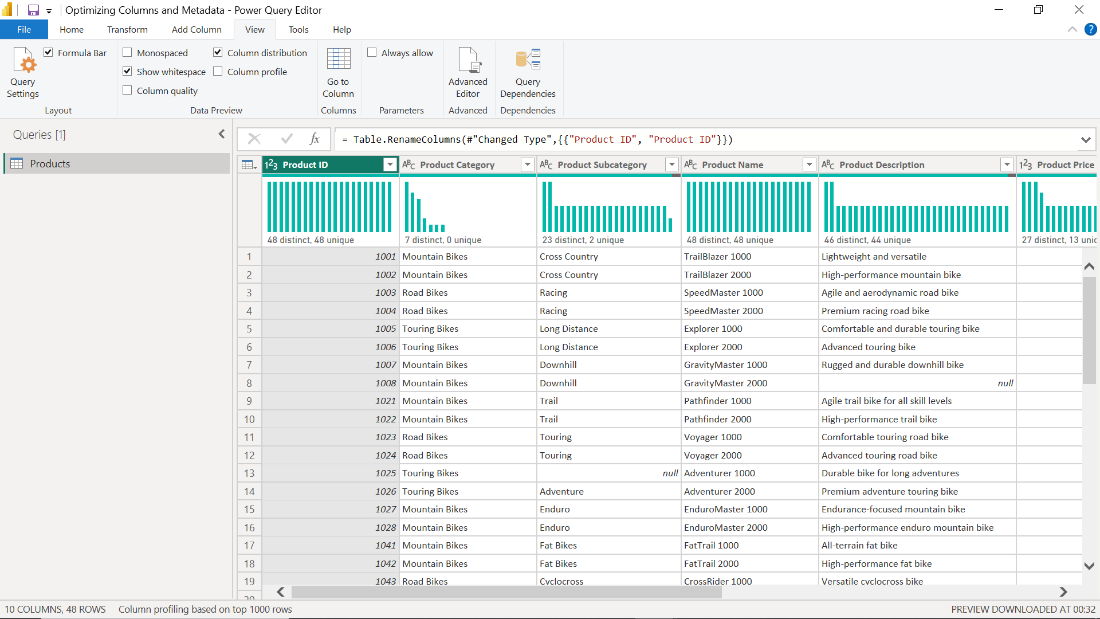
**Column Distribution**

This feature shows the frequency and distribution of values within columns, visualized as a histogram. It helps you understand the data's range and frequency, which is crucial for data analysis. The steps to enable this feature are as follows:

1. Select the **View** tab in the ribbons in the top portion of the **Power Query Editor** window. In the **Data Preview** group, select the **Column Distribution** checkbox.



1. The **Column Distribution** pane opens under the column headings in the data preview. It displays a histogram that represents the frequency of values within your column.



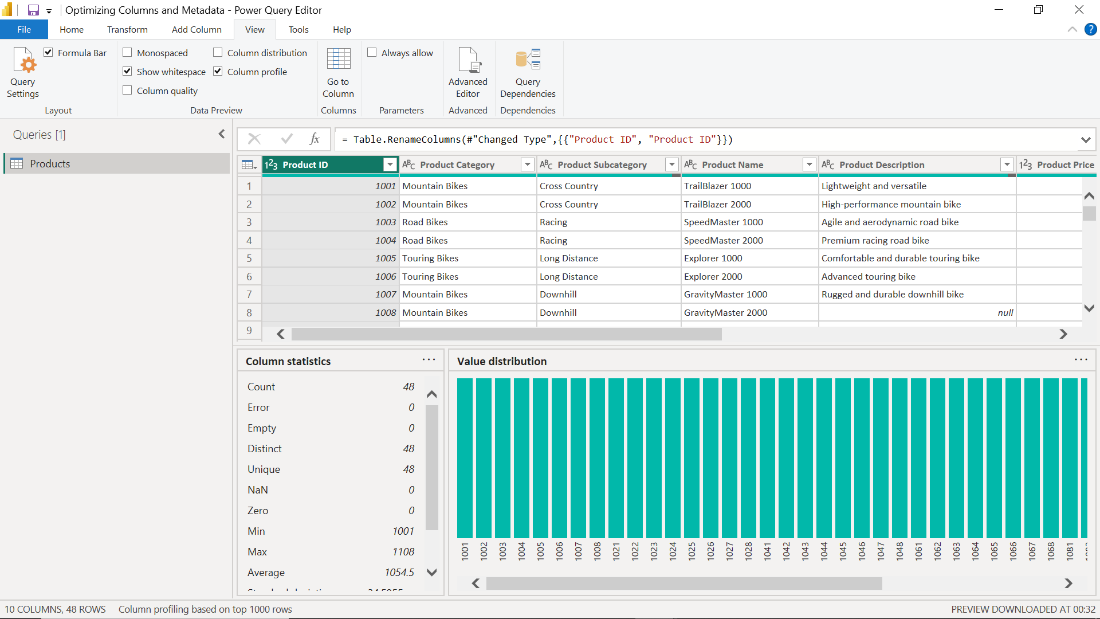
**Column Profile**

The **Column Profile** feature provides a detailed overview of your columns, showing statistical measures such as count, unique count, min, max, and average. It also displays a distribution chart. This feature is highly beneficial when understanding your data and helps identify issues such as outliers. The steps to enable this feature are as follows:

1. Select the **View** tab in the ribbons in the top portion of the **Power Query Editor** window. In the **Data Preview** group, select the **Column Profile** checkbox.

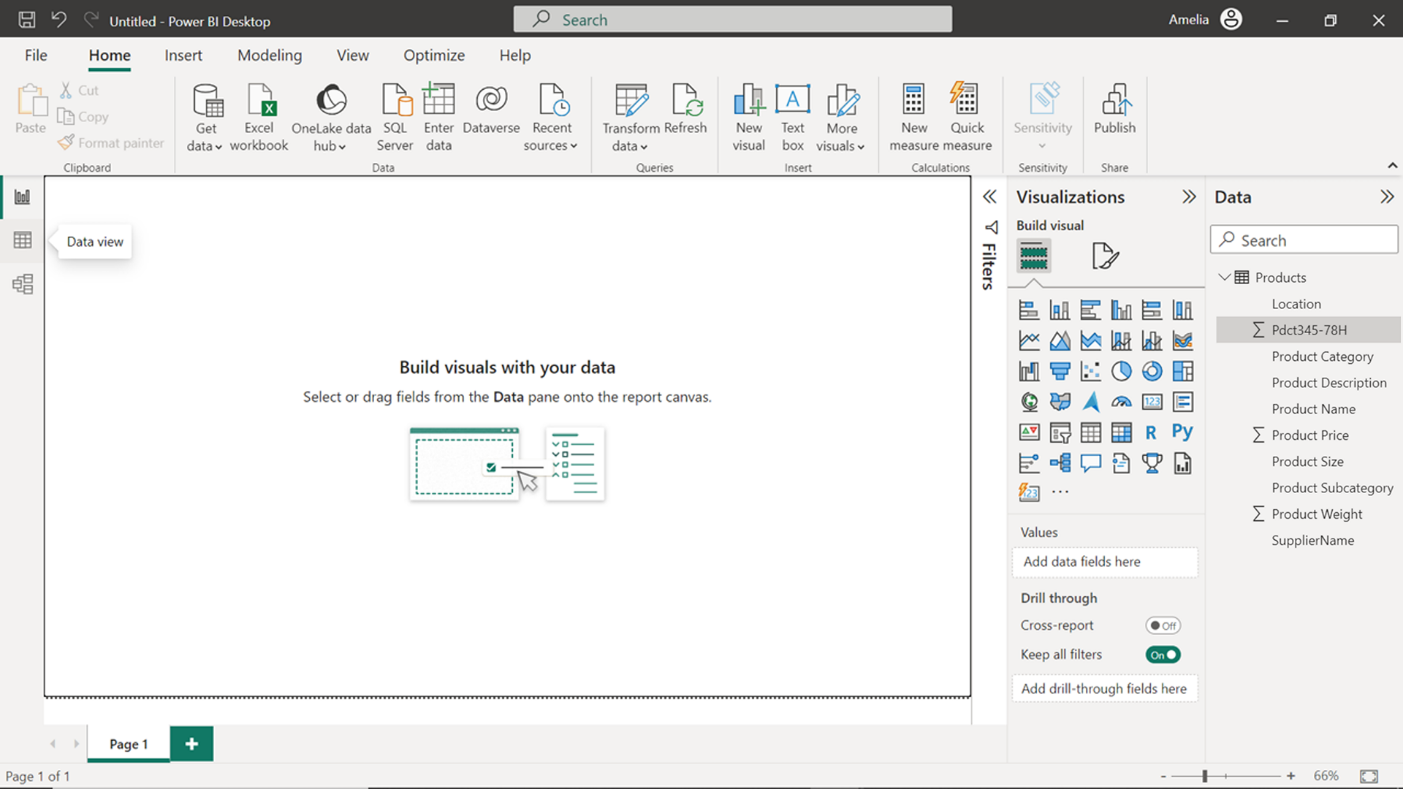


1. In the **Data preview** window, select the column you want to inspect. An additional statistical pane opens underneath the data preview. The statistical pane is divided into two sub-sections: **Column Statistics** and **Value Distribution**. It displays a range of statistical measures and a distribution chart, providing a rich, detailed profile of your selected column.



**Inspect the Tables and Columns**

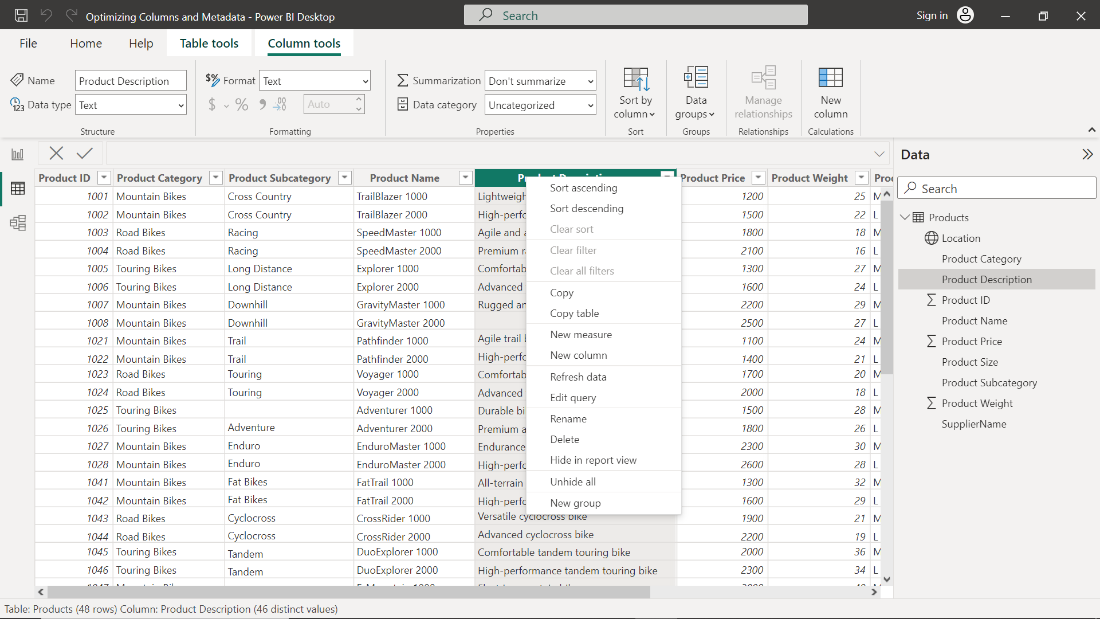
To look at your report's current structure, navigate to the **Data view**. Here, you can see all your tables in the **Data** pane on the right. Selecting a table expands it to reveal its columns. You'll see the column names, data types, and records, which will give you an idea about the data contained in each.



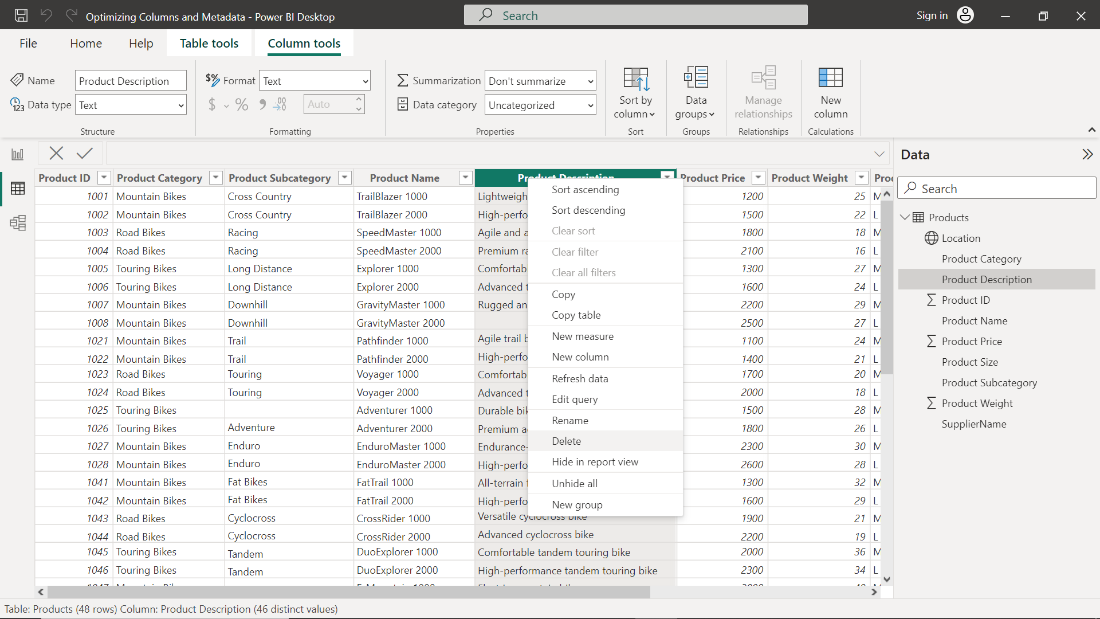
**Removing Columns**

When you gain a deeper knowledge of your data, you may notice that some columns are unnecessary for your analysis. Unnecessary columns take up memory and can slow down queries. Remember, less is more when it comes to data modeling, as you want your data model to include only what's necessary for the analysis. The steps to remove unnecessary columns are as follows:

1. First, select the table name in the **Data** pane to display a list of its columns. Then, right-click on the column that you want to remove.



1. From the menu that appears, select **Delete**. This action will remove the column from your dataset, helping you keep your data model lean, improving its performance, and simplifying its structure.



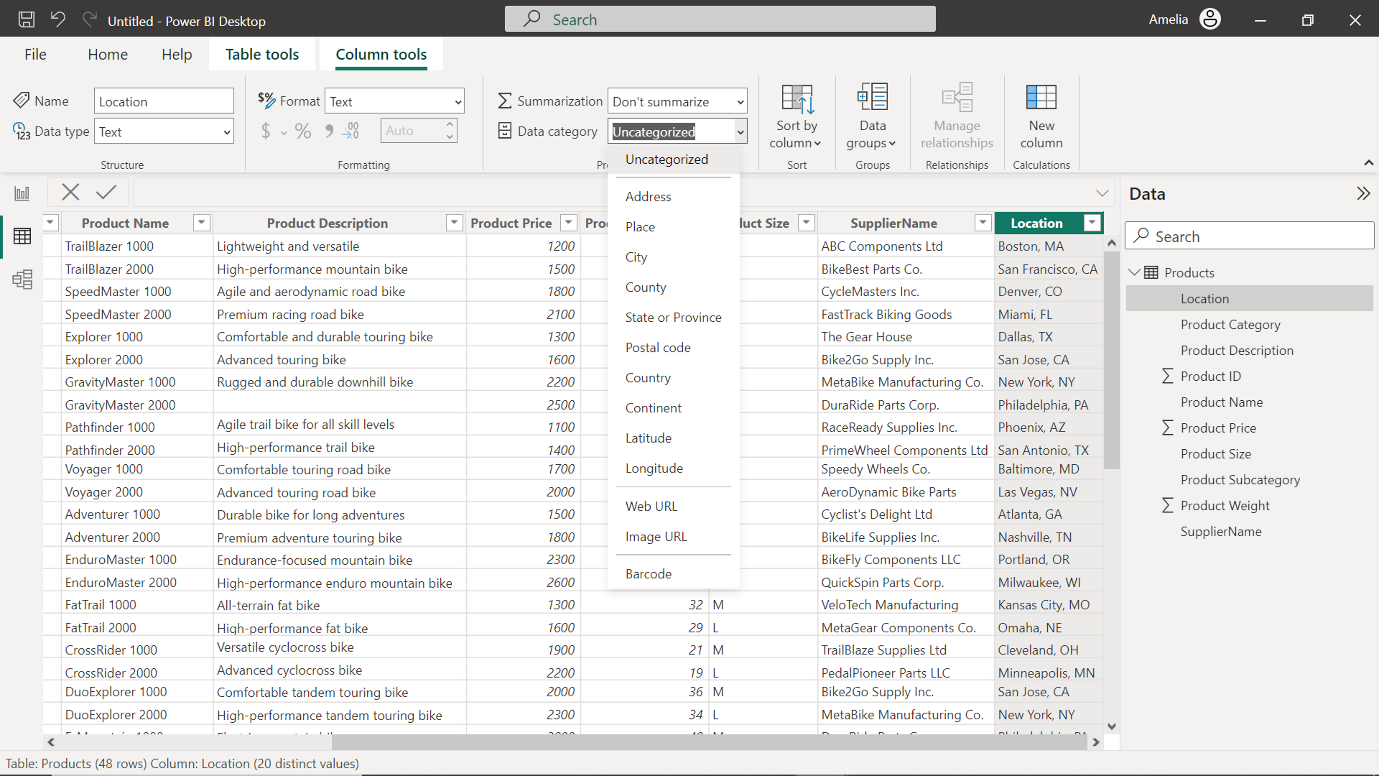
**Categorizing Data**

Once you have removed unnecessary columns, you should then categorize your data. Categorizing your columns in Power BI can help the system understand your data better and offer more relevant features and insights. For example, if a column is categorized as **City**, Power BI will know that this column can be used in geographical analysis. The steps to categorize columns are as follows:

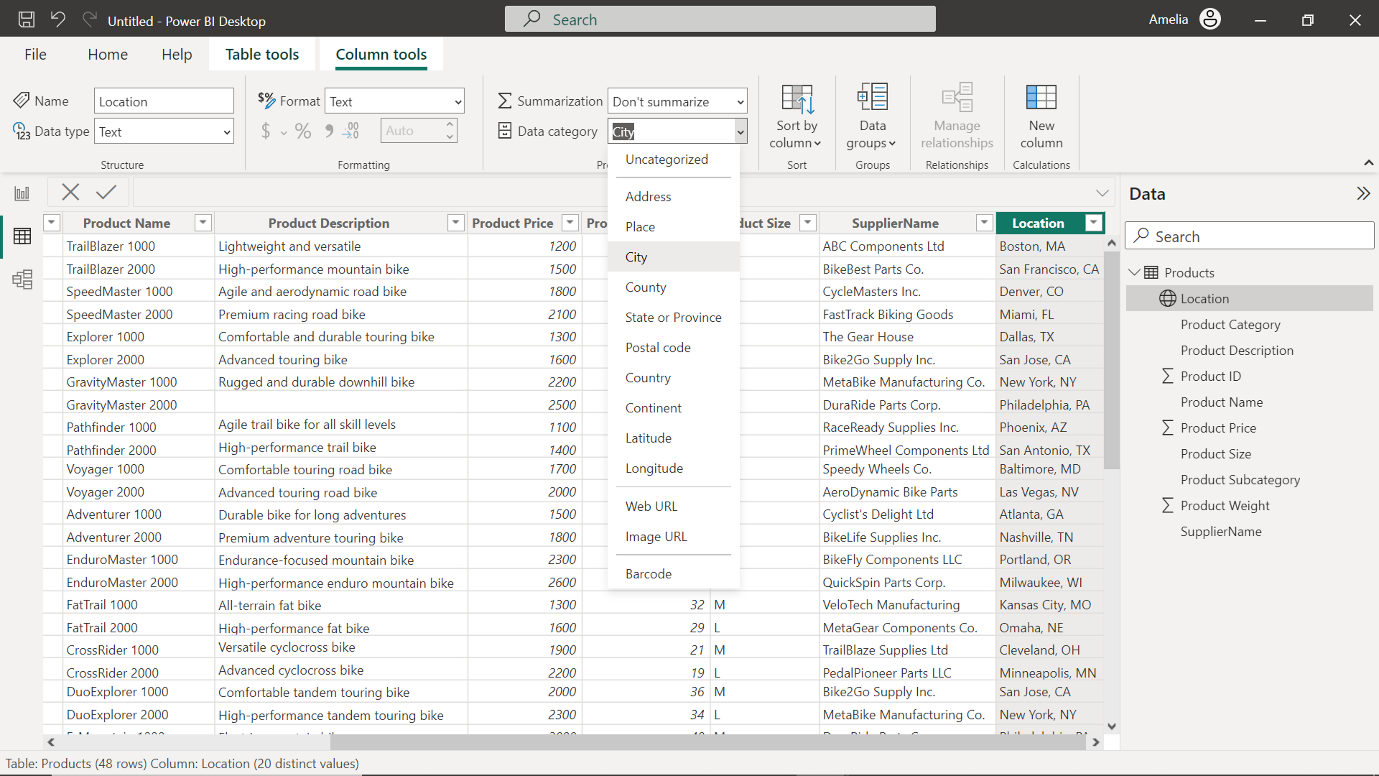
1. In **Data view**, select the column you want to categorize.



1. In the ribbon area, select the **Column tools** tab. In the **Properties** group, select the **Data category** drop-down.



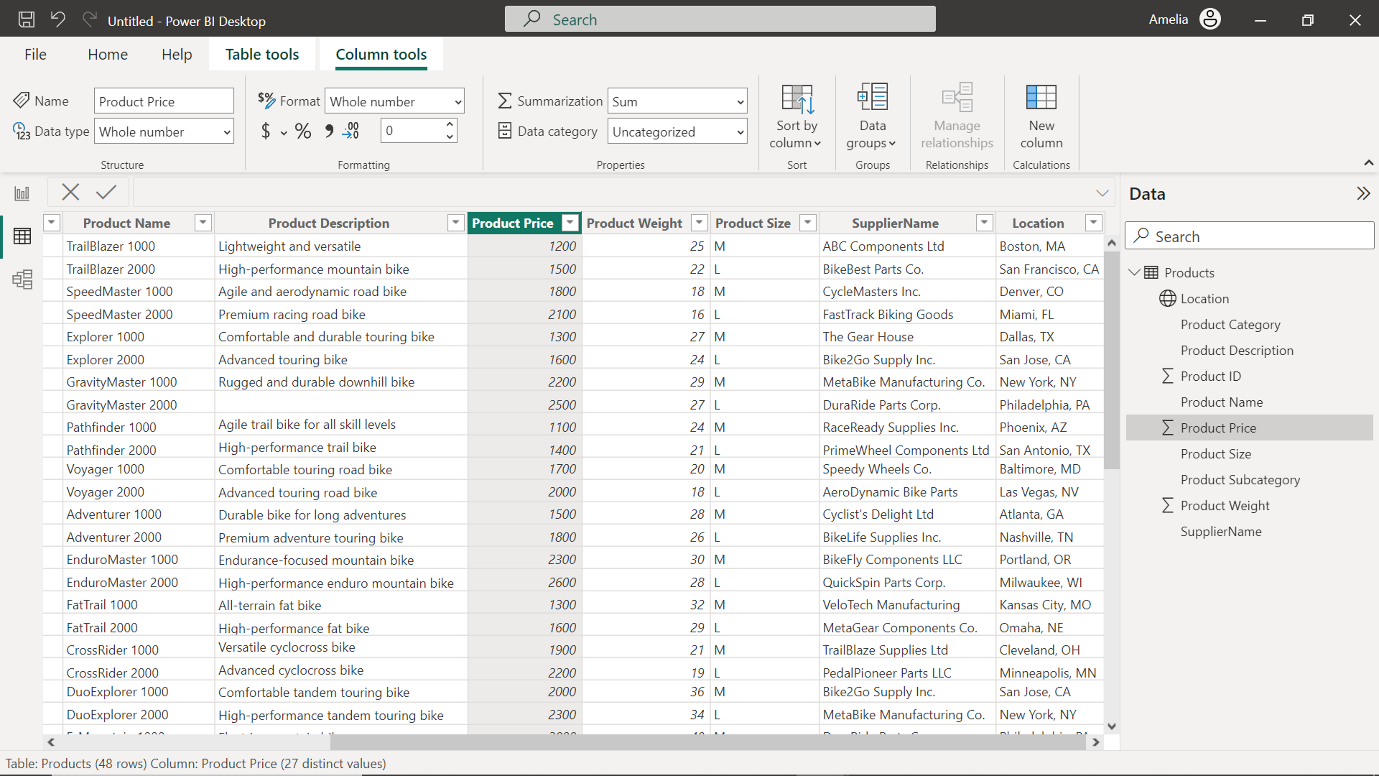
1. A dropdown menu opens, which contains categories such as **City**, **Country**, **Image URL,** and others. Select the category that best fits the data in your column.



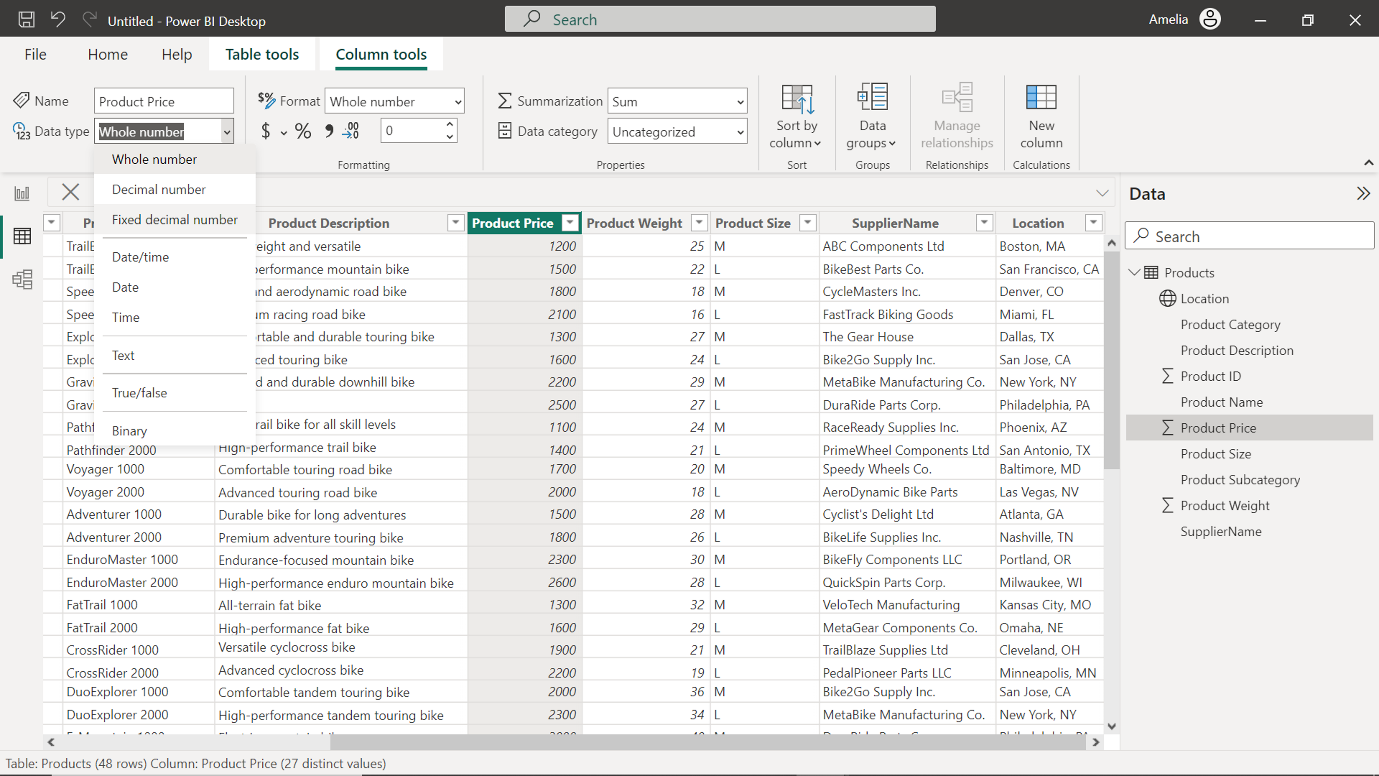
**Change Column Data Types**

Microsoft Power BI tries to guess the data type of each column when you first import data. Sometimes, it may assign a data type that isn't the most efficient. For instance, if a column of numbers is assigned a **Text** data type, it can take up more memory than necessary. To change the data type, right-click the column name and select **Change Type**. A list of data types will appear. Select the most appropriate one. For example, a column that contains whole numbers can be changed from **Text** to **Whole number**, which is a more memory-efficient data type. The steps to define data types for columns are as follows:

1. In **Data view**, select the column for which you want to define the data type.



1. In the ribbon area select the **Column tools** tab. In the **Properties** group select the **Data type** drop-down.



1. A dropdown menu opens, displaying various data types such as **Text**, **Whole number**, **Decimal number**, **Date/Time,** and others. Choose the data type that corresponds with the data in your column.

**Conclusion**

Optimizing columns and metadata in Microsoft Power BI can make a data model more efficient and intelligible. Optimizing the data model improves the technical performance of Power BI and paves the way for more precise, meaningful, and impactful data analysis. The changes you've implemented today can lead to greater insights tomorrow, and these insights can influence the strategic decisions made by Adventure Works.