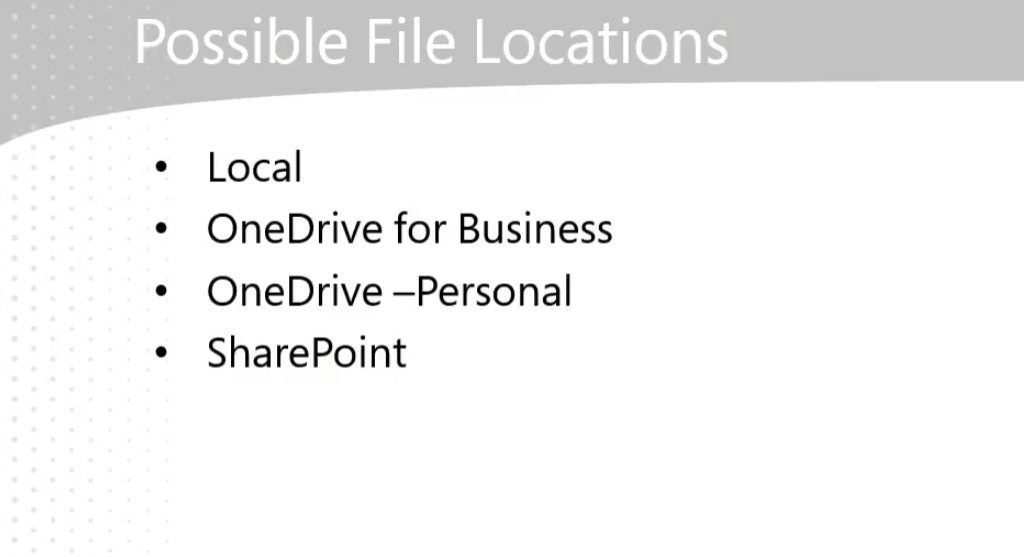


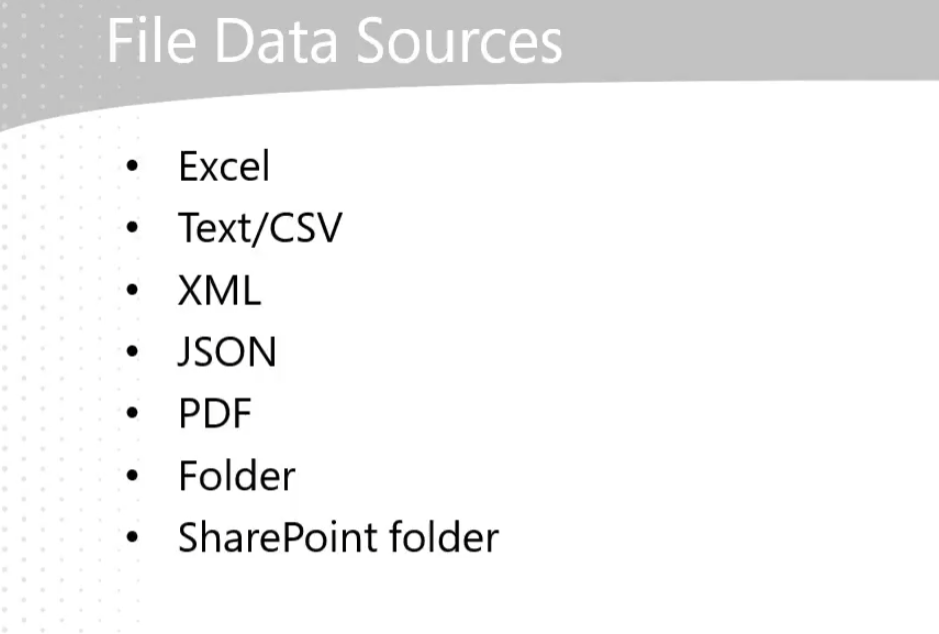
Power BI - Git hub link:

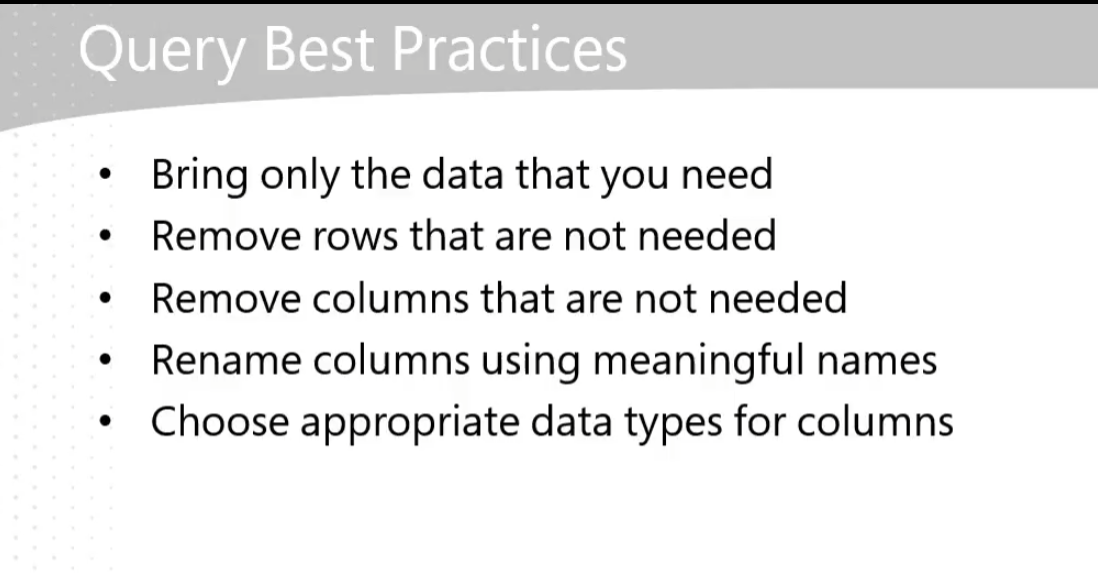
<https://github.com/IterationInsights/Microsoft-Press-Introduction-to-Microsoft-Power-BI-V2>

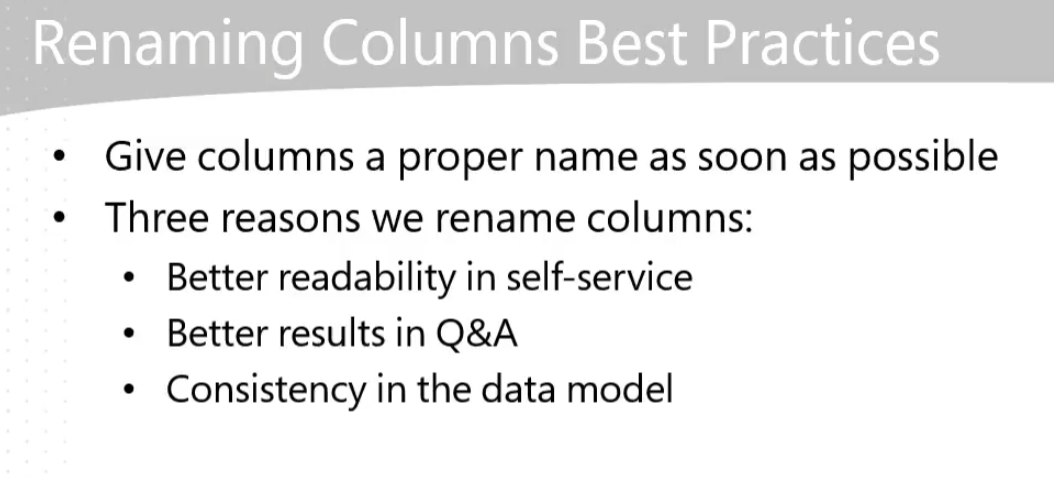
* **Data Preparation**: As a data analyst, a significant portion of your work involves preparing data for analysis, which includes tasks like removing rows and columns, cleaning data, and merging datasets.
* **Power Query Capabilities**: Power Query simplifies and automates data ingestion and transformation, offering features like data connectors, intuitive data shaping, and recorded transformations to save time.
* **Efficiency**: Using Power Query can potentially reduce data preparation time by up to 80%, making the process more efficient and less manual.

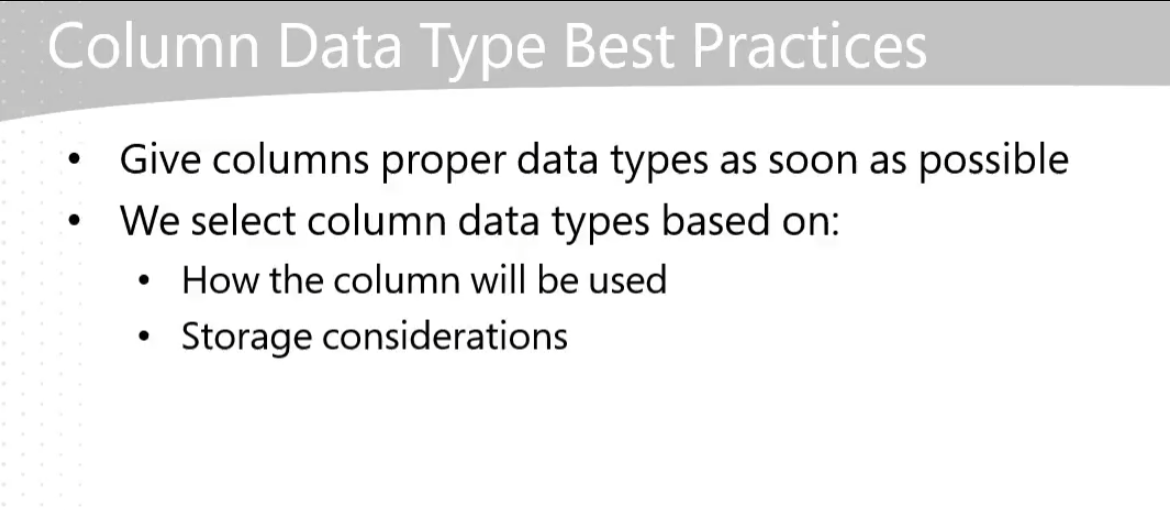
The **M language** is a powerful data transformation language used in Power Query. It allows you to perform complex data manipulations and transformations efficiently. In Power Query, M language records the steps you take to transform your data, making it easier to automate and repeat these processes. This helps streamline data preparation and ensures consistency in your data models.

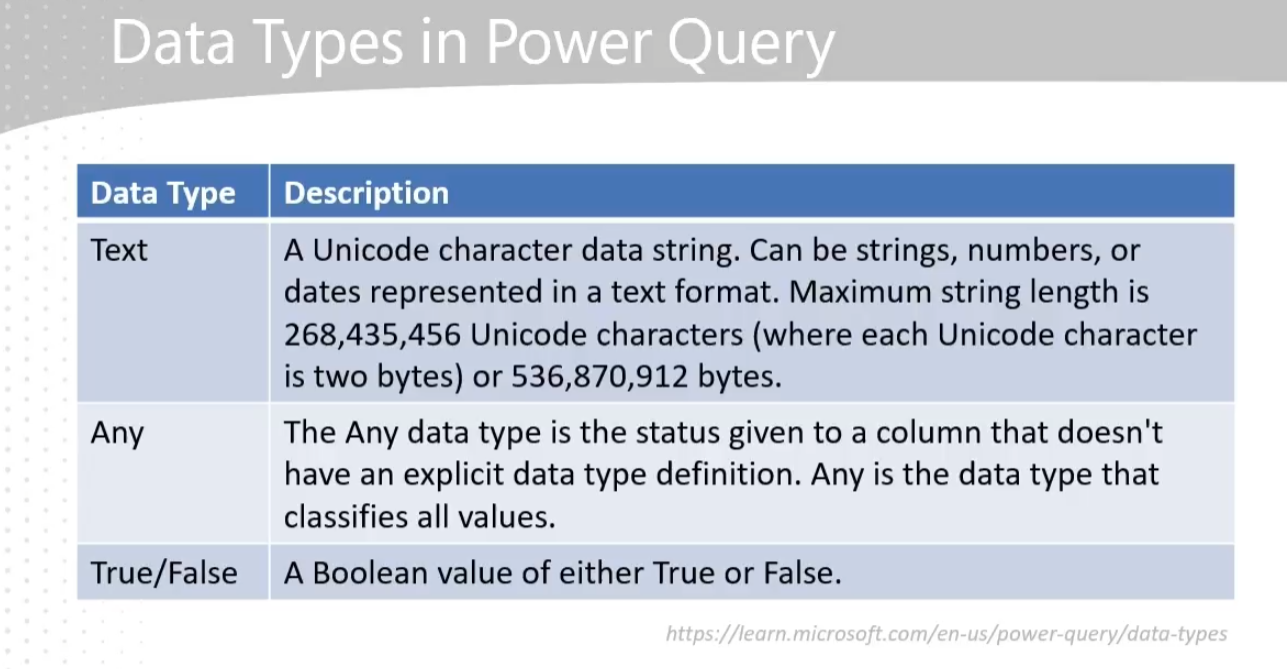


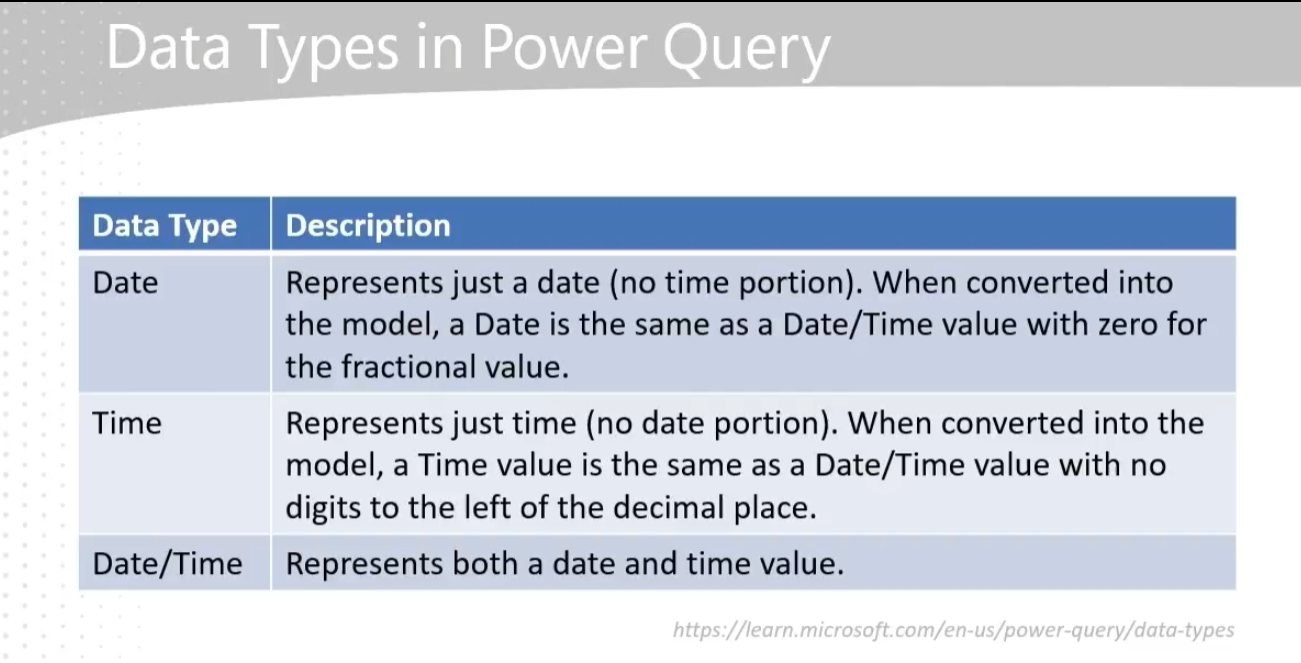


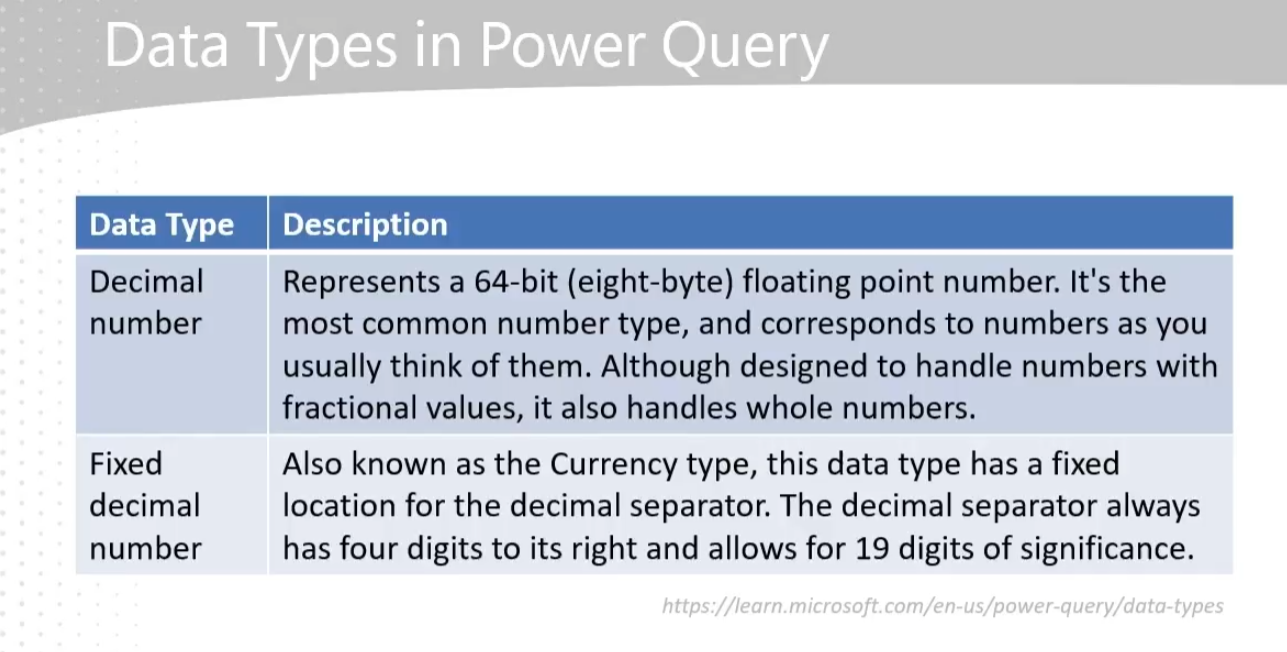


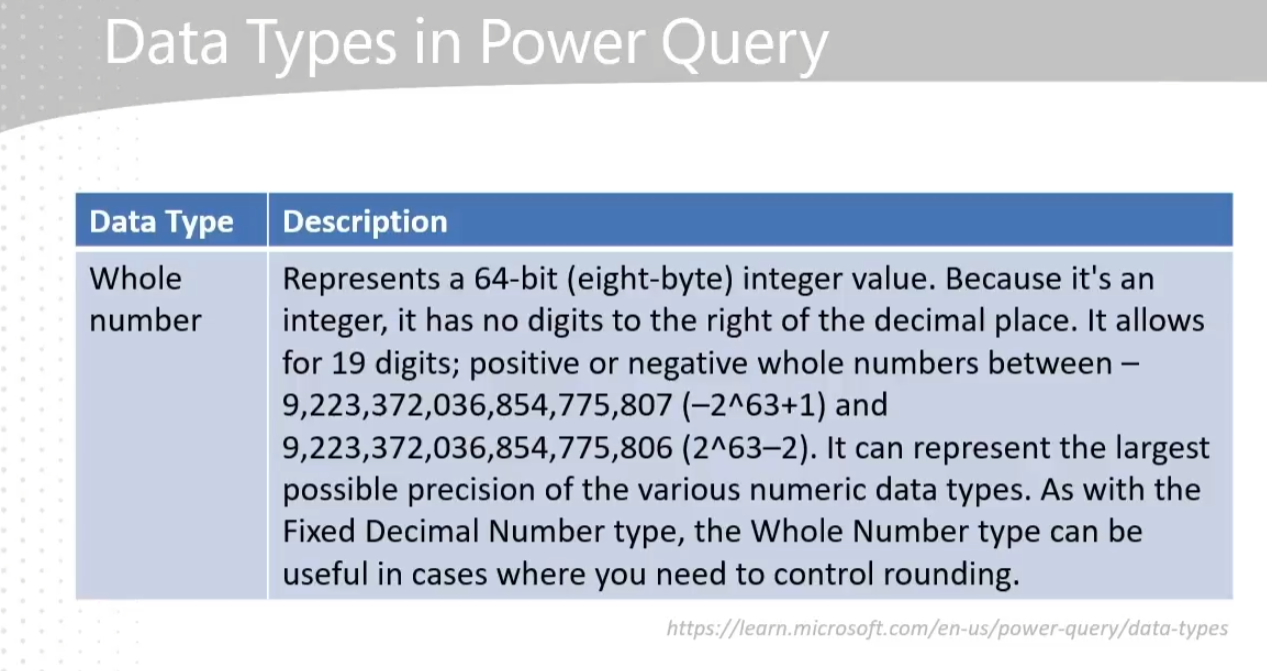


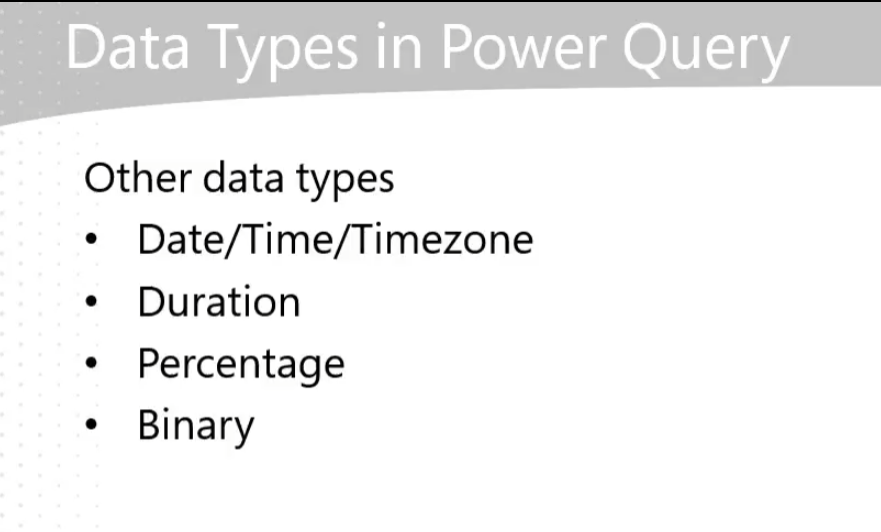




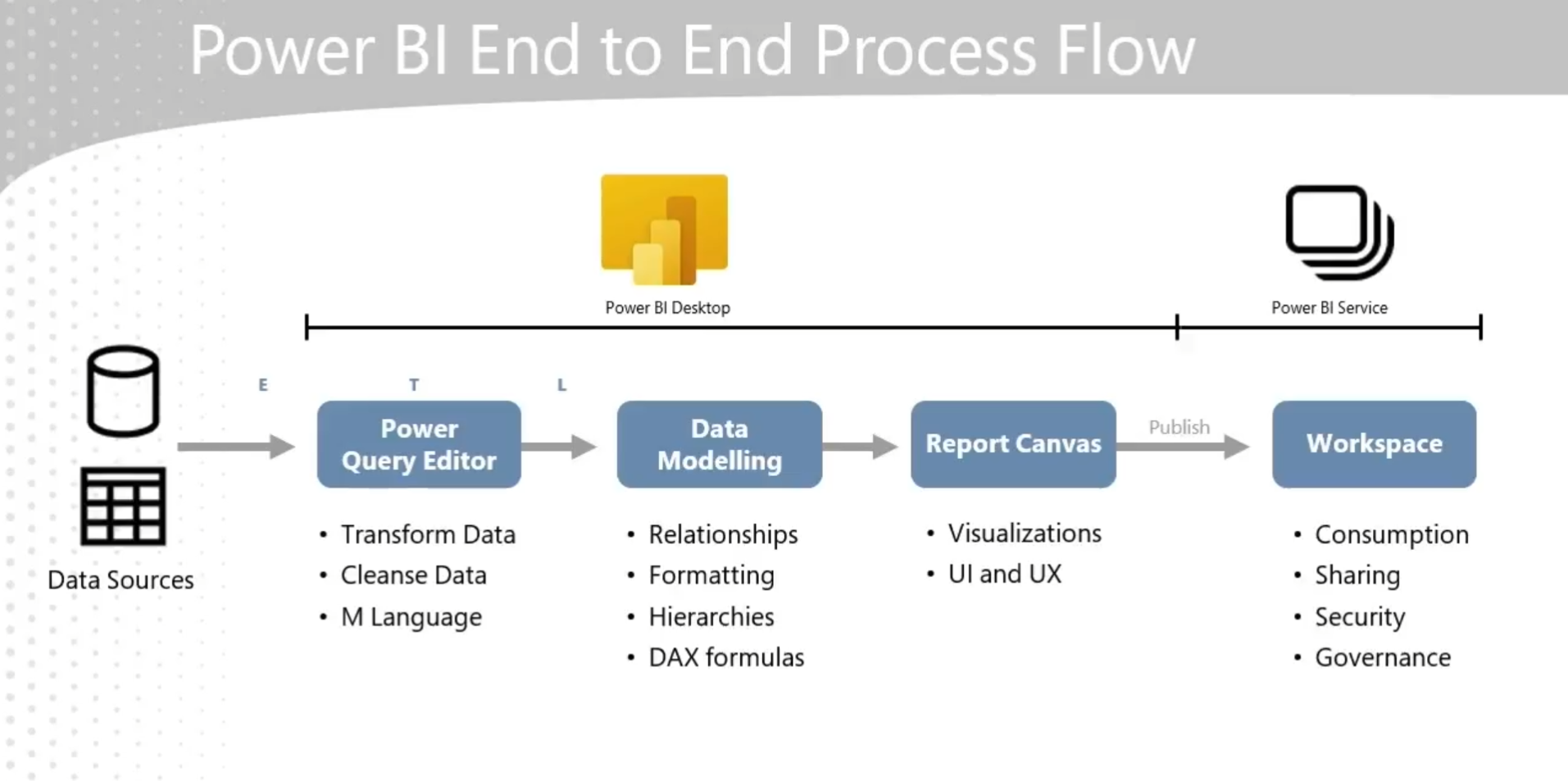




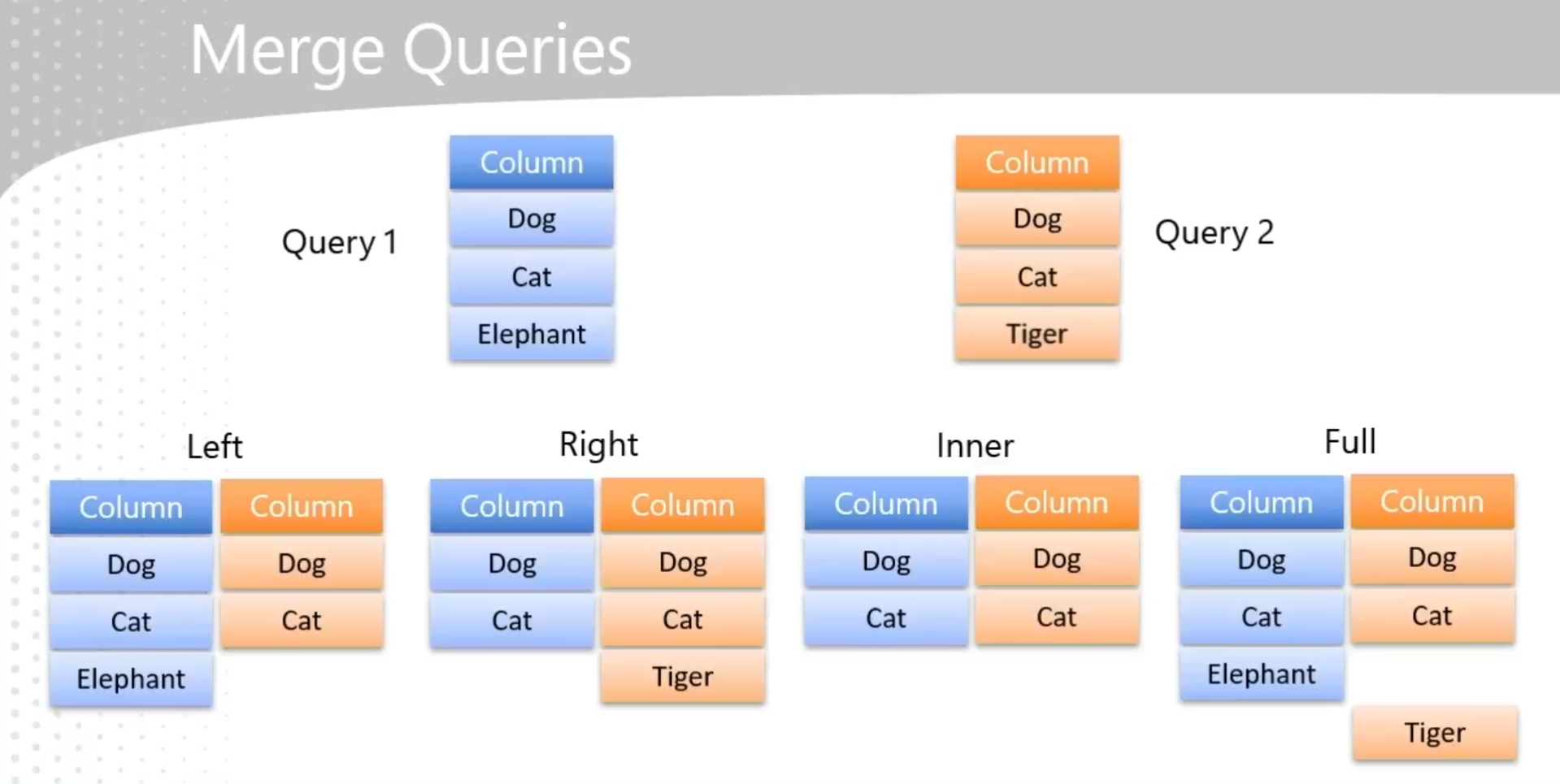




**ETL IN POWER BI:**

****

**MERGE QUERIES:**

****

**MERGING IN POWER BI:**

* **Purpose of Merging: Merging the product and product rollup queries helps reduce the number of tables, making the data model less intimidating and more efficient for end-user self-service. It also improves performance and allows for building hierarchies within the data model.**
* **Steps to Merge:**
  + **Select the product query and use the "Merge Queries" function.**
  + **Choose the product rollup query to merge with the product query.**
  + **Match the columns (e.g., Product Subcategory) and select the type of join (default is left outer join).**
  + **Expand the merged table to include only the necessary columns.**
* **Benefits: Merging queries simplifies the data model, enhances performance, and supports building hierarchies, which is crucial for creating comprehensive and organized reports.**

**Reducing the number of tables in a data model has several benefits:**

* **Simplification: Fewer tables make the model less intimidating and more approachable for end users, facilitating self-service.**
* **Performance: It can improve performance by reducing the complexity of the data model.**
* **Hierarchies**: It allows for the creation of hierarchies, as all columns participating in a hierarchy need to exist in the same table.

**FOLDER CONNECT**

* **Using the Folder Connector**: The video explains how to use the folder connector in Power BI to import multiple files from a directory as if they were a single file, which simplifies the data import process.
* **Consistency in Data Format**: It emphasizes the importance of having a consistent data format across all files to ensure smooth data transformation and avoid errors.
* **Automation Benefits**: By maintaining standards and consistent formats, new data files can be automatically integrated into the Power BI model without additional manual adjustments, enabling efficient automation.

The folder connector in Power BI allows you to treat multiple files within a folder as a single data source. This is particularly useful when you have data spread across multiple files and want to apply the same transformation to all of them. By using the folder connector, you can:

* Combine data from multiple files into one query.
* Apply consistent transformations across all files.
* Simplify data management by treating all files as a single entity.

This is especially beneficial when dealing with large datasets or when new files are added regularly, as it automates the process and ensures consistency.

To import data from a folder in Power BI, follow these steps:

1. **Open Power BI Desktop**: Start by opening your Power BI Desktop application.
2. **Navigate to Power Query Editor**: Click on the "Home" tab and then select "Transform Data" to open the Power Query Editor.
3. **Select New Source**: In the Power Query Editor, click on "New Source" and then select "More" to open the Get Data dialog box.
4. **Choose Folder Connector**: In the Get Data dialog box, find and select the "Folder" connector, then click "Connect".
5. **Browse for Folder**: Click "Browse" to navigate to the folder containing your data files. Select the folder and click "OK".
6. **Combine and Transform Data**: After selecting the folder, you will see a list of files within it. Click "Combine" and then "Combine & Transform Data".
7. **Select Sample File**: In the dialog box that appears, choose a sample file to define the structure of the data. Ensure all files in the folder follow the same structure.
8. **Select Data to Import**: Choose the table or data object you want to import from the sample file. Ensure all files have the same table structure.
9. **Skip Files with Errors (Optional)**: You can choose to skip files with errors by checking the appropriate box. If unchecked, the entire load will fail if any file does not match the structure.
10. **Finalize Import**: Click "OK" to import and combine the data from all files in the folder into a single query.
11. **Rename Query (Optional)**: You can rename the query for better clarity by right-clicking on the query name and selecting "Rename".
12. **Clean Up Data**: Perform any necessary data cleaning, such as removing unnecessary columns or changing data types.
13. **Load Data**: Once you are satisfied with the data, click "Close & Apply" to load the data into Power BI.
14. These steps will help you import and combine data from multiple files in a folder into Power BI.

**UPPEND DATA (Combine the data in two different sheet in to a single sheet)**

The "Append sale queries" exercise is designed to combine sales data from different sources into a single table, which is crucial for comprehensive data analysis. Here are the advantages and a step-by-step explanation of the process:  
  
Advantages:

* **Unified Data Analysis**: By combining sales data from different countries into one table, you can perform global sales analysis more efficiently.
* **Consistency**: Ensures that all sales data is in a consistent format, making it easier to work with and analyze.
* **Flexibility**: Power Query is flexible in handling different data types and column mismatches, making the data preparation process smoother.

Step-by-Step Process:

1. **Load Data**: Load all the data from your sources through the Power Query editor.
2. **Identify Tables**: Identify the tables you want to combine (e.g., United States sales and other countries' sales).
3. **Append Queries**:  
   * Go to the Power Query editor.
   * Select the table you want to append to (e.g., other countries' sales).
   * Click on the "Combine" menu and choose "Append".
   * Select "Append as new" to create a new query.
   * Choose the tables to append (e.g., other countries' sales and United States sales).
4. **Rename Query**: Rename the new query to something meaningful (e.g., "Sales").
5. **Check Columns**: Ensure that the columns match and handle any mismatches (e.g., columns with different names or missing columns).
6. **Adjust Data Types**: Make sure the data types are consistent and adjust them as needed.
7. **Remove Unnecessary Columns**: Remove any columns that are not needed (e.g., the "freight" column if it's not present in all tables).
8. **Finalize Data**: Ensure that the final table has the correct column names and data types.

This process helps in creating a unified dataset that can be used for more effective and comprehensive data analysis.

When appending queries in Power BI, column mismatches are handled by matching columns based on their names. Here are the key points:

* **Column Matching**: Power BI matches columns purely based on their names, regardless of data types.
* **Missing Columns**: If a column exists in one query but not the other, Power BI will create the missing column and fill it with null values for the rows from the query that lacks the column.
* **User Intervention**: You can then decide how to handle these null values or remove the mismatched column if it's not needed.

This approach ensures flexibility while allowing you to manage any discrepancies in the data.