IterationInsights

Power BI Dataflows

Follow Along Guide ➡️
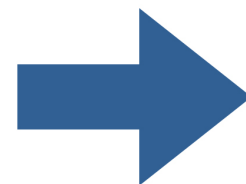
Introduction to Power BI Dataflows

Dataflows are an important tool within the Power BI ecosystem for streamlining data preparation that also provide easier scalability than what is found within regular datasets.

There are several ways in which they can be used but with the right knowledge can be set up extremely quickly and provide more control over our data pipelines.

Follow along as we explore what Power BI Dataflows are and provide a comprehensive guide on how to use them effectively.

Dataflows vs Datasets

In most cases a dataset can do everything a dataflow can do, so why shouldn't we just use datasets? Well dataflows allow for us to store transformations within the data further upstream.

This means we can have any additional logic attached to this one source that other datasets will pull from instead of having to copy that logic across to multiple datasets.
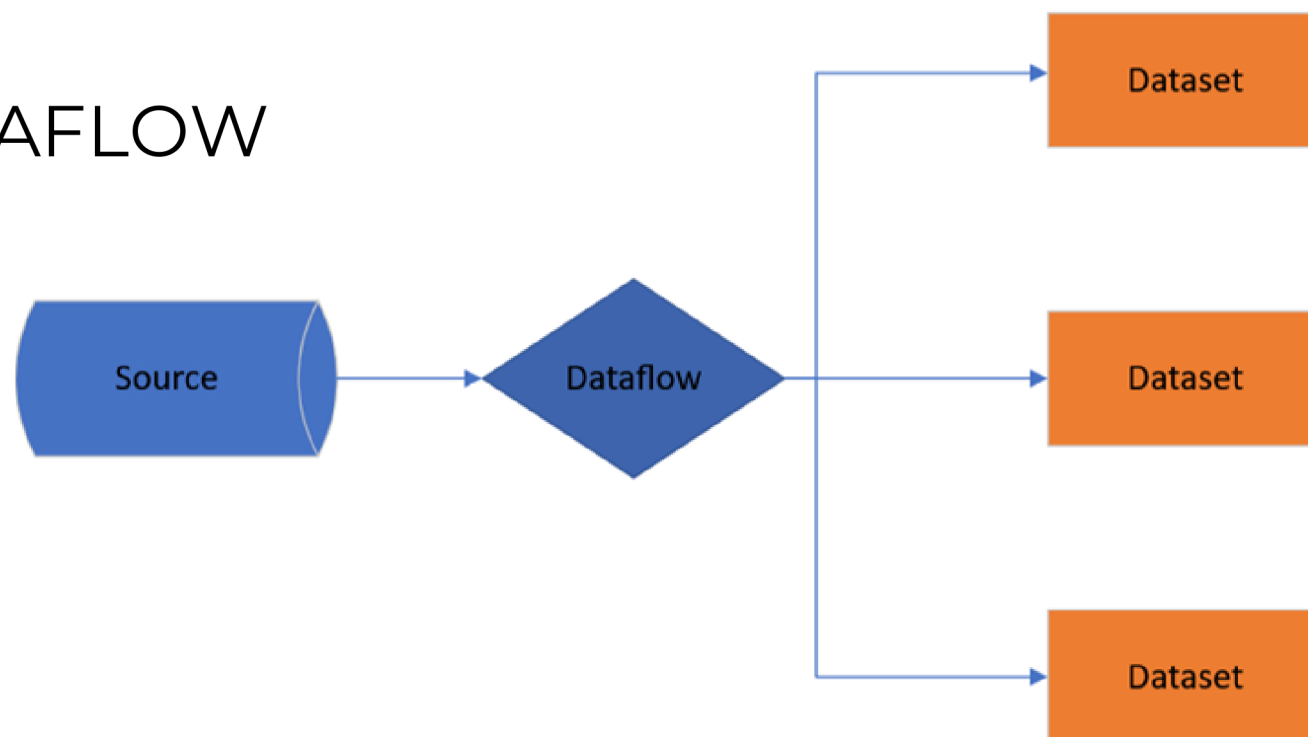
Having this centralized means that instead of our reports hitting a source multiple times and refreshing off it, we are only using a single thread through our dataflow. This also simplifies the creation of new reports that need to pull from a main source within an organization.
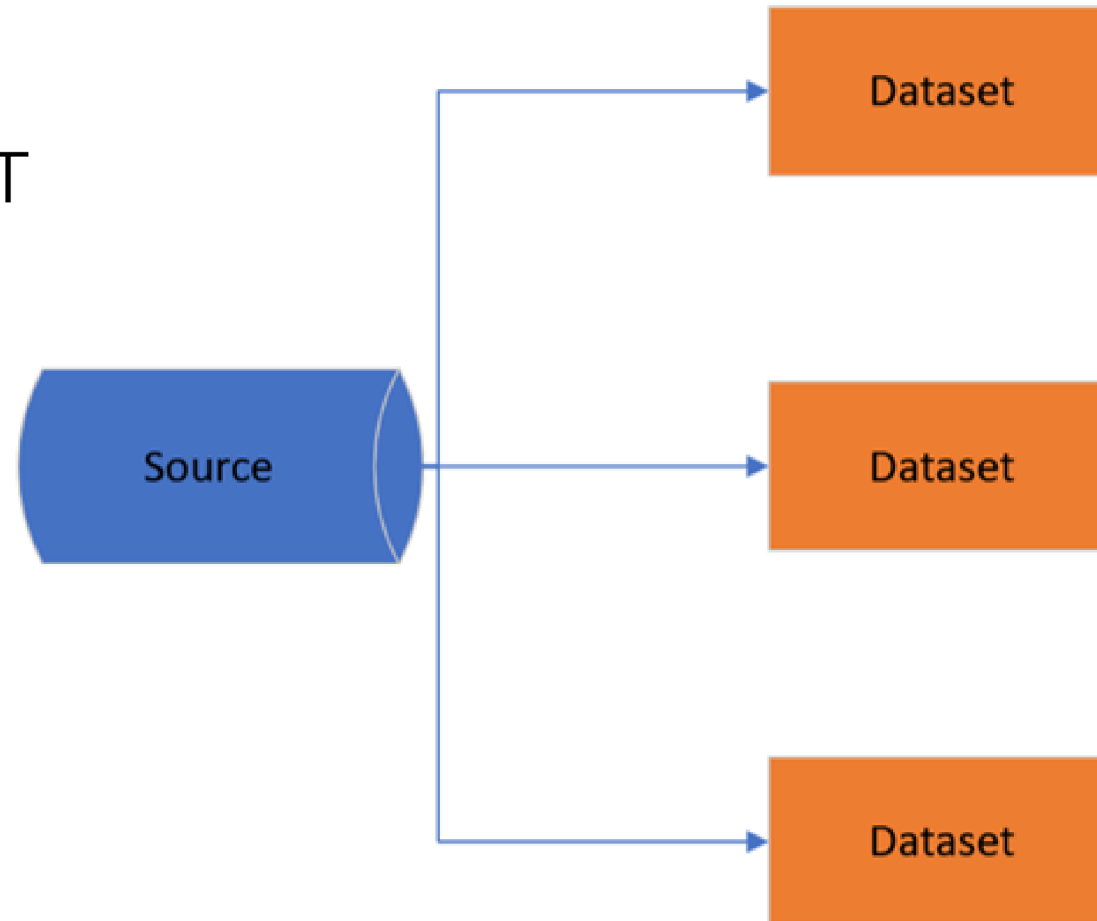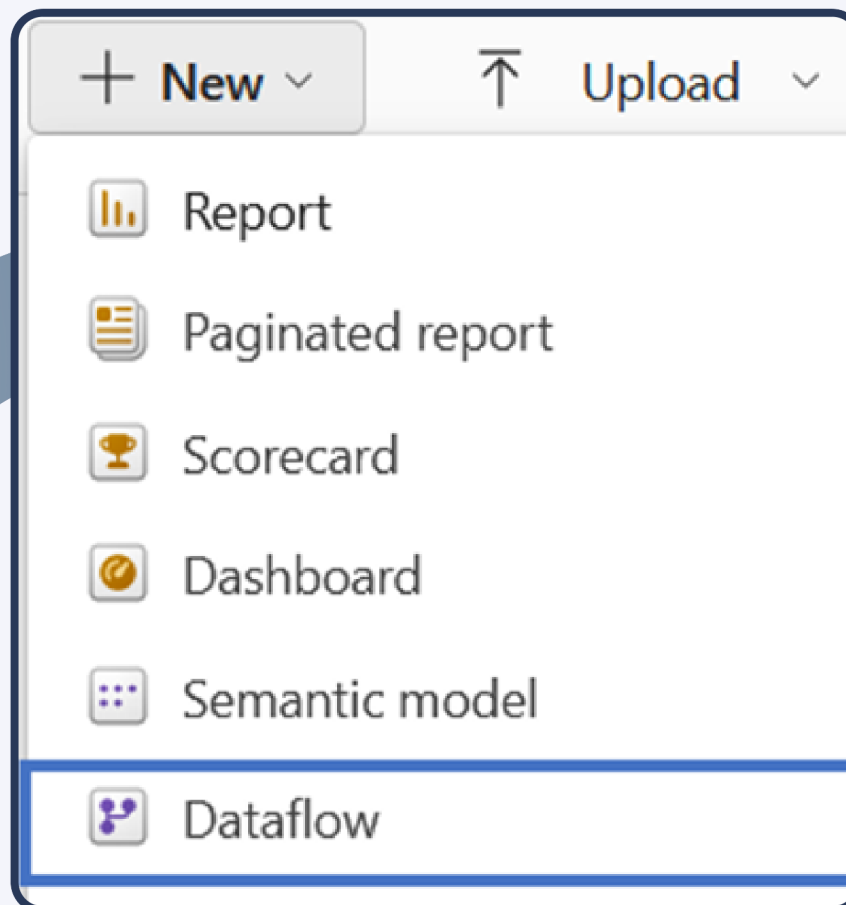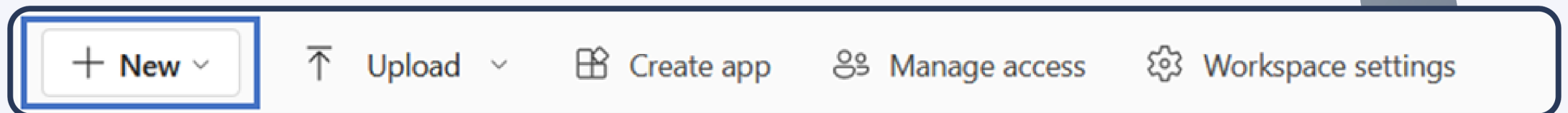
Getting Started

1 **Accessing Power BI Dataflows:**

- To start using Dataflows, navigate to the Power BI service.
- Open the Power BI workspace where you want to create the Dataflow.
- Select the new dropdown and dataflows

+ New ∨	↑ Upload ∨	🎁 Create app	😃 Manage access	⚙ Workspace settings

+ New ∨	↑ Upload ∨

📊 Report

📇 Paginated report

🏆 Scorecard

◐ Dashboard

⠿ Semantic model

🔀 Dataflow

2 **Creating a Blank Dataflow:**

- The next part of the process is determined by how you want the dataflow to be set up, in this case we will be starting fresh, but dataflows allow the ability to connect from other workspaces or import from previous dataflows.
- Select "Define new tables"

Start creating your dataflow

Define new tables

Choose a data source to define the tables for your dataflow. You can map your data to standard Common Data Model tables, or define custom tables instead.

Learn more

Add new tables

Link tables from other dataflows

Linking to tables from other dataflows reduces duplication and helps maintain consistency across your organization.

Learn more

Add linked tables

Import Model

Choose a dataflow model to import into your workspace.

Learn more

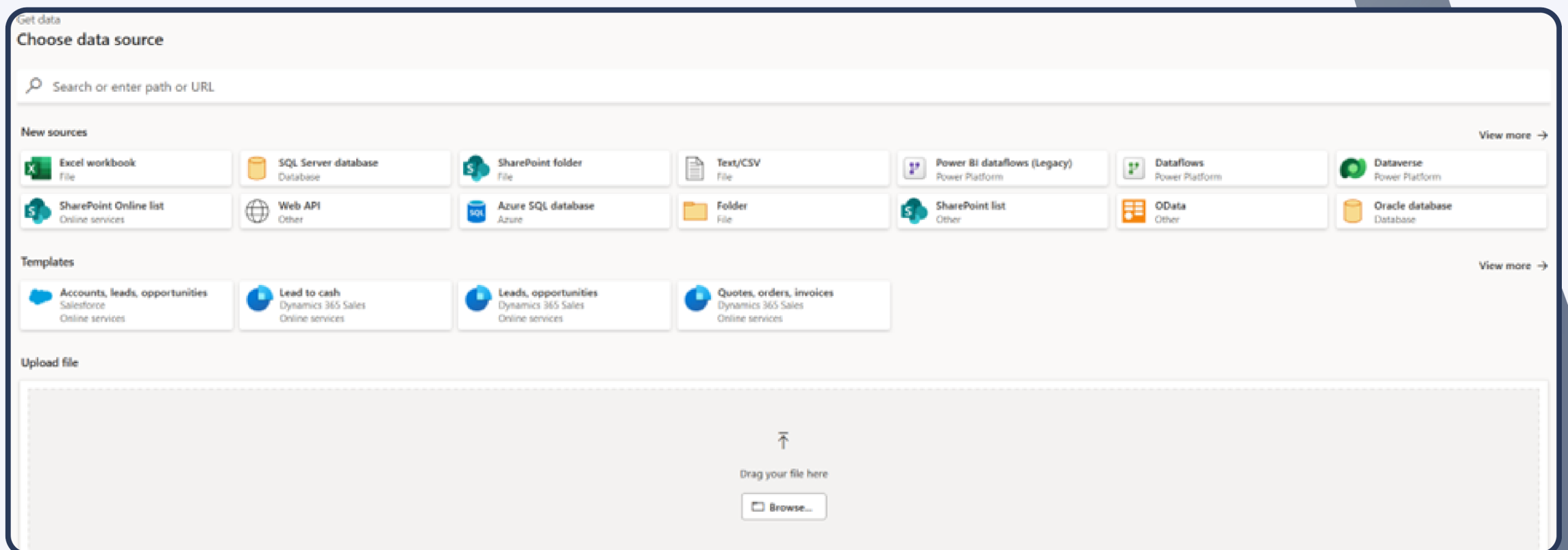Import model

Attach a Common Data Model folder (preview)

Attach a Common Data Model folder from your Azure Data Lake Storage Gen2 account to a new dataflow, so you can use it in Power BI.

Learn more

Create and attach

3 Setting up the data source:

- oDataflows offer the same connection types found within datasets, so any existing data source should be easily connected to in the same way.
- If there is a gateway connection available, given the correct configuration it will automatically populate that gateway in the "Connection" field.

Connect to data source

SQL Server database
Database
Learn more

Connection settings

Server * ⓘ

Example: test.corp.contoso.com

Database

Example: Contoso_DB

> Advanced options

Connection credentials

Will autofill with gateway given a matching connection established above

Connection

Create new connection   ⌄   ↻

Connection name

Connection

Data gateway

(none)   ⌄   ↻

Authentication kind

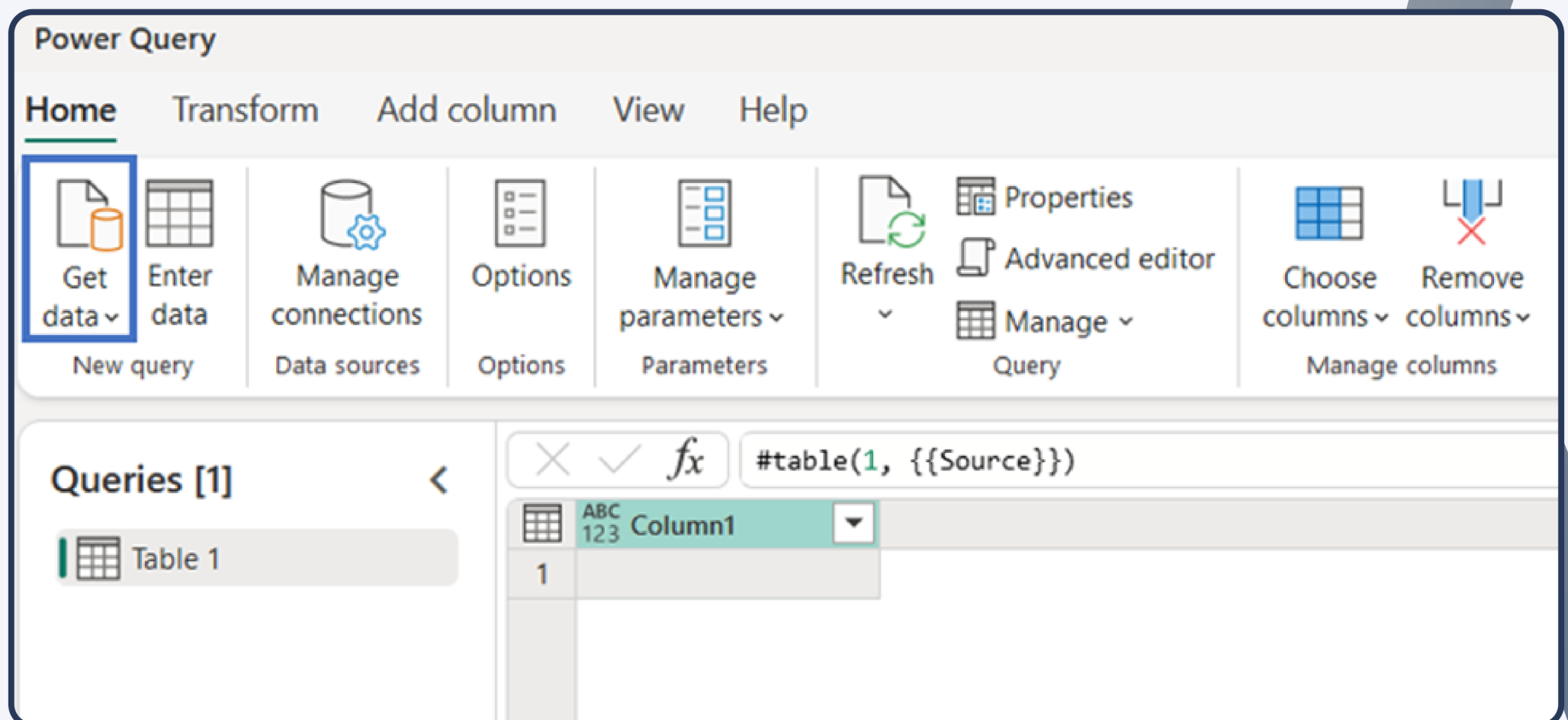Basic   ⌄

Username

Password

4 Transforming data:

- Dataflows allow for the same power query transformations that can be made in a dataset, so it is very easy to transfer the knowledge over to this tool.

Power Query

| Home | Transform | Add column | View | Help |

Get data ⌄ | Enter data | Manage connections | Options | Manage parameters ⌄ | Refresh ⌄ | Properties | Advanced editor | Manage ⌄ | Choose columns ⌄ | Remove columns ⌄

New query | Data sources | Options | Parameters | Query | Manage columns

fx | #table(1, {{Source}})

Queries [1] ‹

Table 1

ABC 123 Column1 ⌄

1

Once a table from a source has been added into power query additional sources or additional tables can also be added in a similar method to datasets, by getting data, duplicating, or entering data.

4 Transforming data:

- oAfter the steps have all been applied and tables completed, we can continue with **"Save & close"** at the bottom of the page
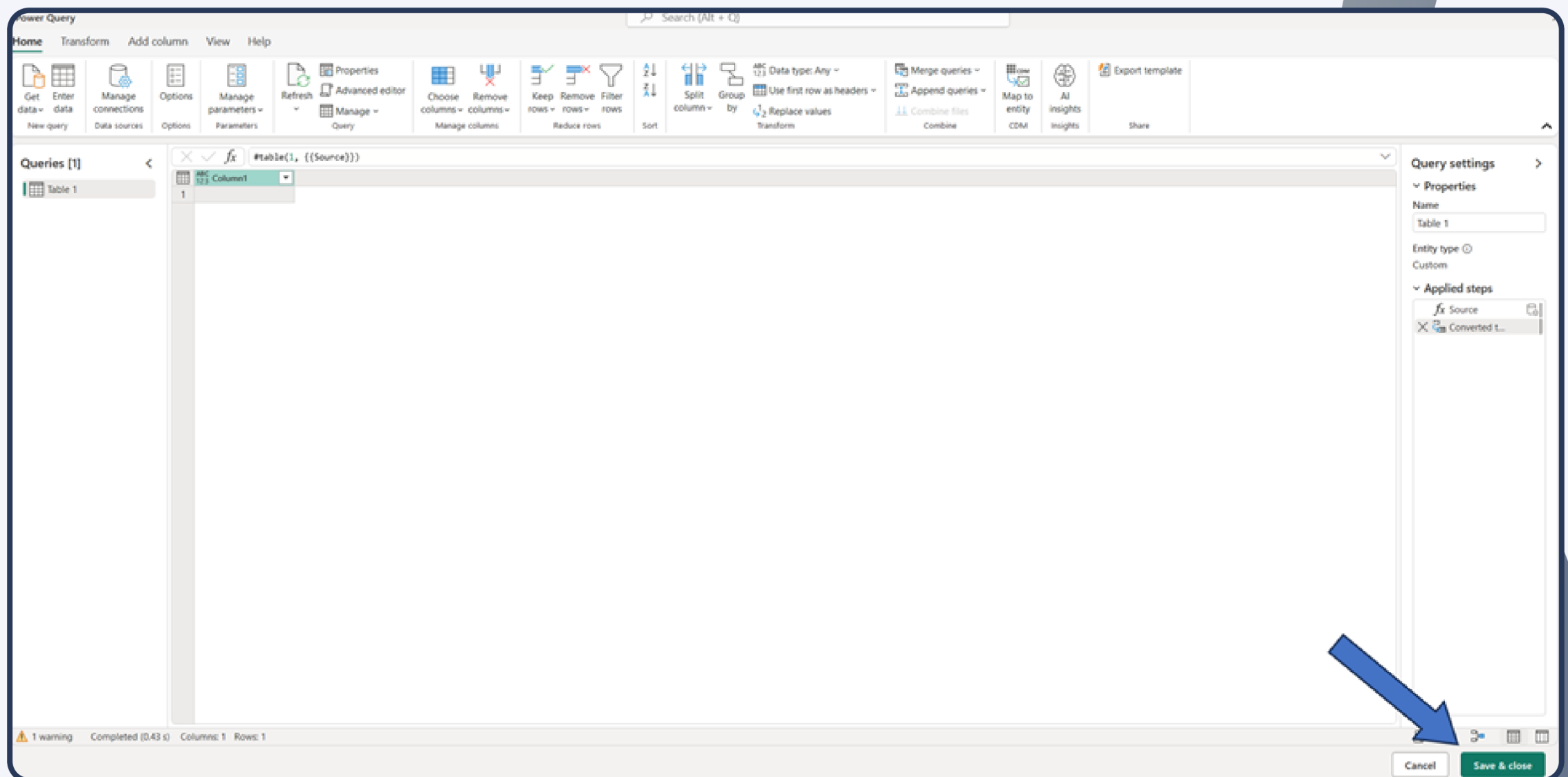
5 **Dataflow created:**

- Now that the transformations have been made, we have some options in the use of the dataflow.

1. **Incremental refresh –** Can be implemented to only refresh new data within the specified period. This is useful if there are not many changes in historical data and helps to cut down on the refresh times/usage.

2. **Scheduled refresh –** Same as on datasets can configure set times to have the dataflow refresh. With the way the datasets will be pulling from this as their source it is recommended that the refresh times are triggered before datasets so that newer data flows into them.
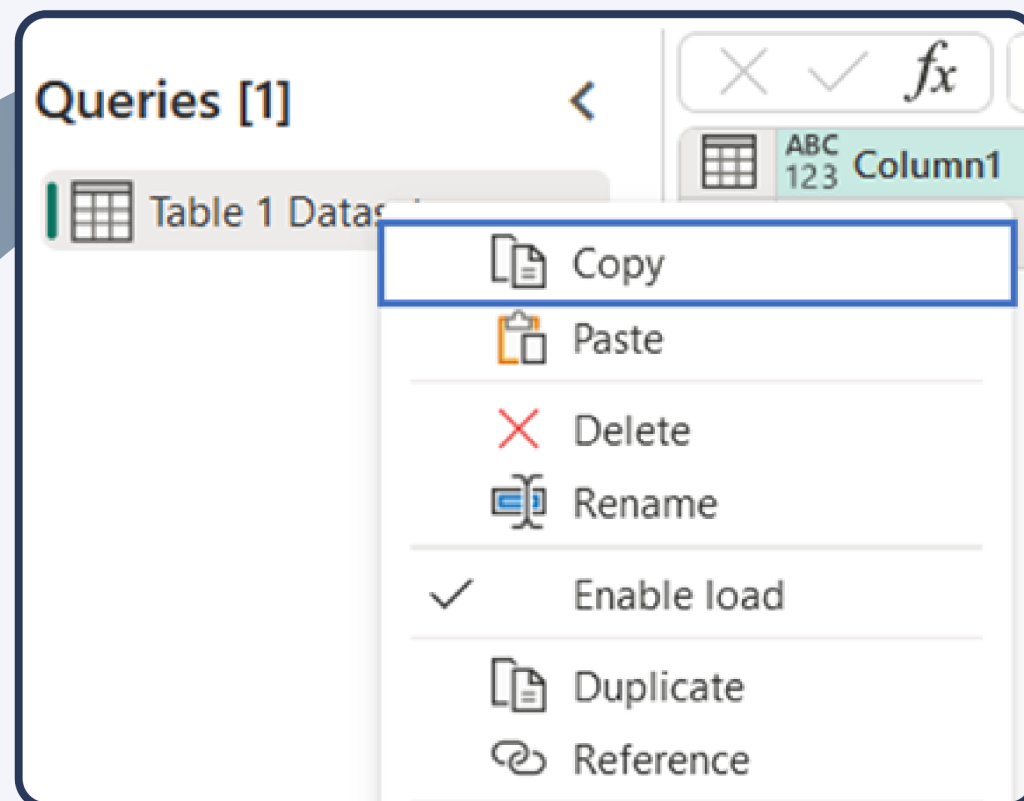
Easy Setup

An excellent feature within Power BI is the ability to easily copy over a transformed table from a Power BI dataset into a dataflow.

Since dataflows use the same Power Query as datasets we can just copy and paste tables from open report files within the Power Query editor over to the dataflows. All the applied steps of the table will automatically be brought over and if it is connected to the previous source you will see all the same data.
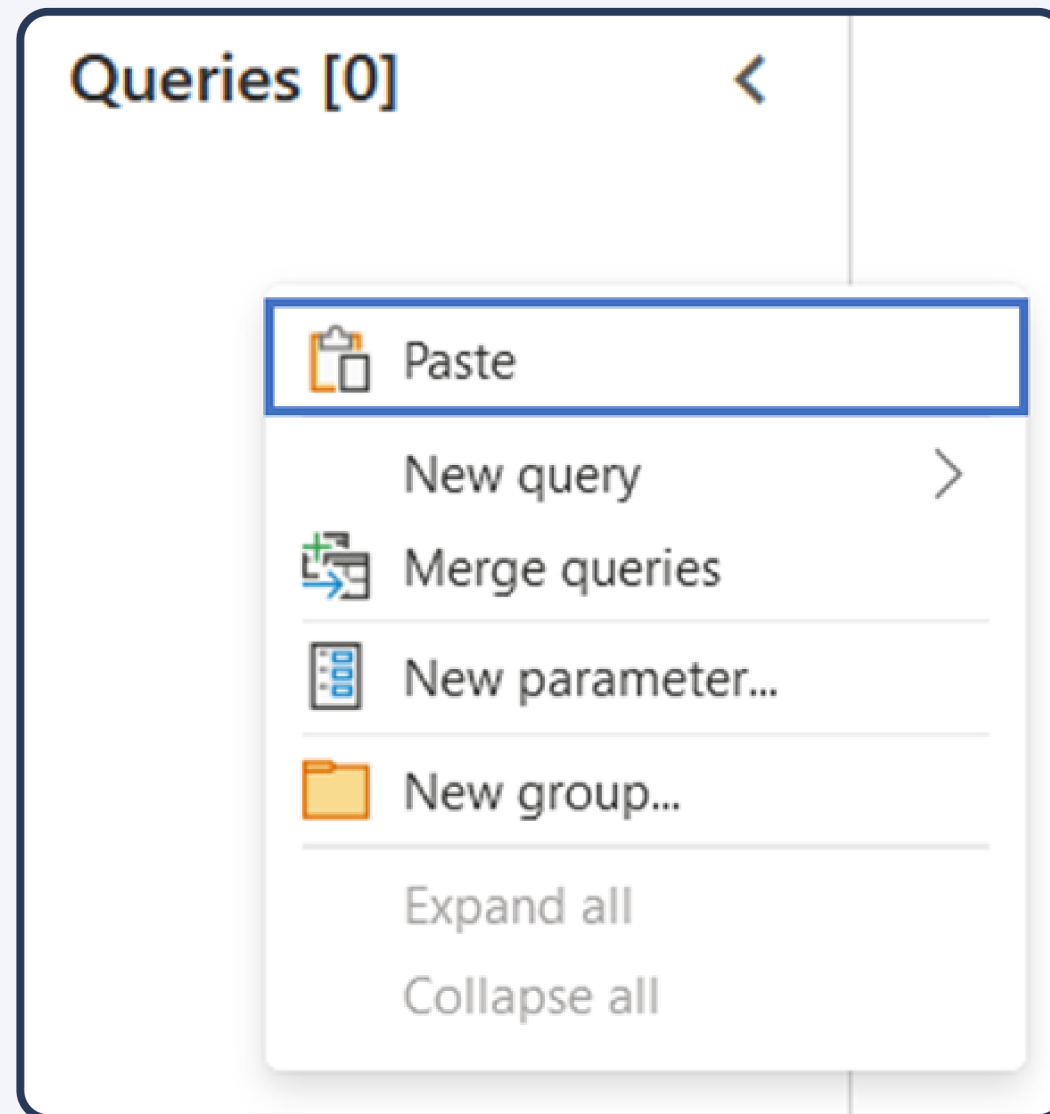
DATASET

Easy Setup

This feature makes it extremely easy to set up dataflows based on previously constructed datasets and get started using them. This method can also be used between dataflows or of course datasets.