PRO TIP: STORAGE & CONNECTION MODES

Power BI Desktop supports several types of **storage** and **connection modes**:

- **Import**: Tables are stored in-memory within Power BI and queries are fulfilled by cached data *(default)*
- **DirectQuery**: Tables are connected directly to the source and queries are executed on-demand at the data source
- **Composite Model (Dual)**: Tables come from a mix of Import and DirectQuery modes, or integrate multiple DirectQuery tables
- **Live Connection**: Connect to pre-published Power BI datasets in Power BI Service or Azure Analysis Services

Import

- ✓ Dataset is less than 1GB (after compression) & fast performance
- ✓ Source data does not change frequently
- ✓ No restrictions on Power Query, data modeling, and DAX functions

DirectQuery

- ✓ Dataset is too large to be stored in-memory
- ✓ Source data changes frequently and reports must reflect changes
- ✓ Company policy states that data can only be accessed from the original source

Composite Model

- ✓ Boost performance by setting appropriate storage for each table
- ✓ Combine a DirectQuery model with additional imported data
- ✓ Create a single model from two or more DirectQuery models

Live Connection

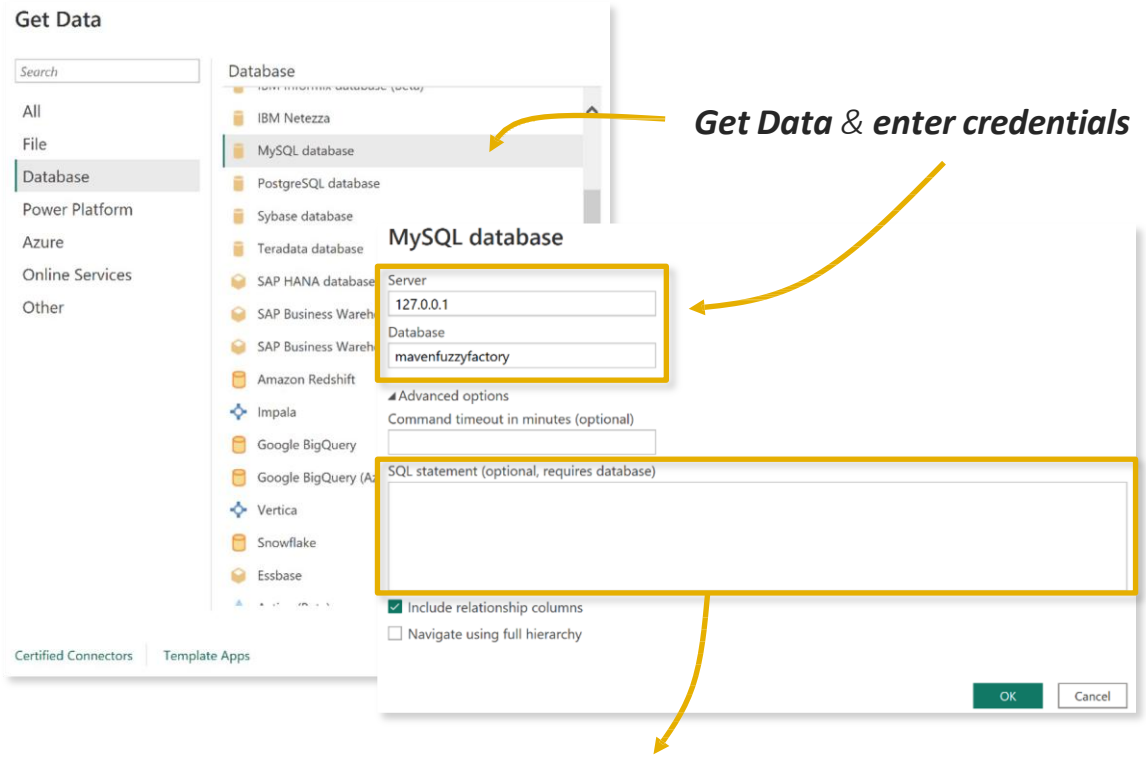
- ✓ Create one dataset that serves as a central source of truth
- ✓ Analyst teams can create different reports from the same source
- ✓ Multi-developer teams where one user builds the model and another works on visualization

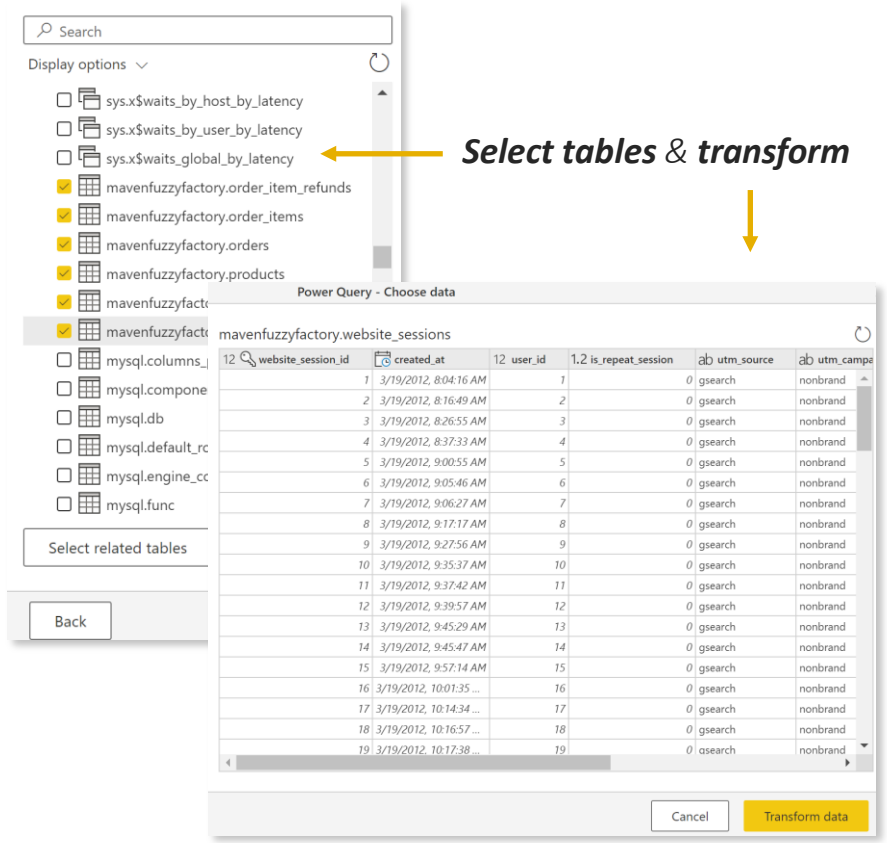
Learn more: *https://learn.microsoft.com/en-us/power-bi/connect-data/service-dataset-modes-understand*

*Copyright Maven Analytics, LLC

CONNECTING TO A DATABASE

Power Query can connect to data from various **database sources** including SQL Server, MS Access, MySQL, PostgreSQL, Oracle, SAP, and more



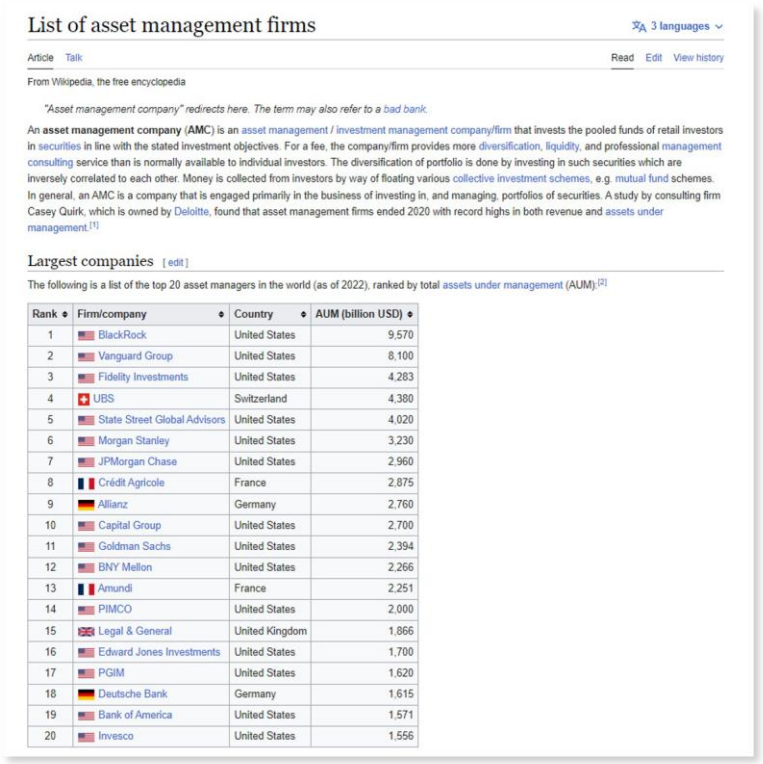


Write custom or advanced queries *with SQL statements (optional)*

*Copyright Maven Analytics, LLC

EXTRACTING DATA FROM THE WEB

Power Query includes a native **Web connector** for importing web-hosted files (csv, xlsx, etc.) or scraping URLs for anything that Power Query can identify as a structured table



List of asset management firms

3 languages

Article Talk | Read Edit View history

From Wikipedia, the free encyclopedia

"Asset management company" redirects here. The term may also refer to a bad bank.

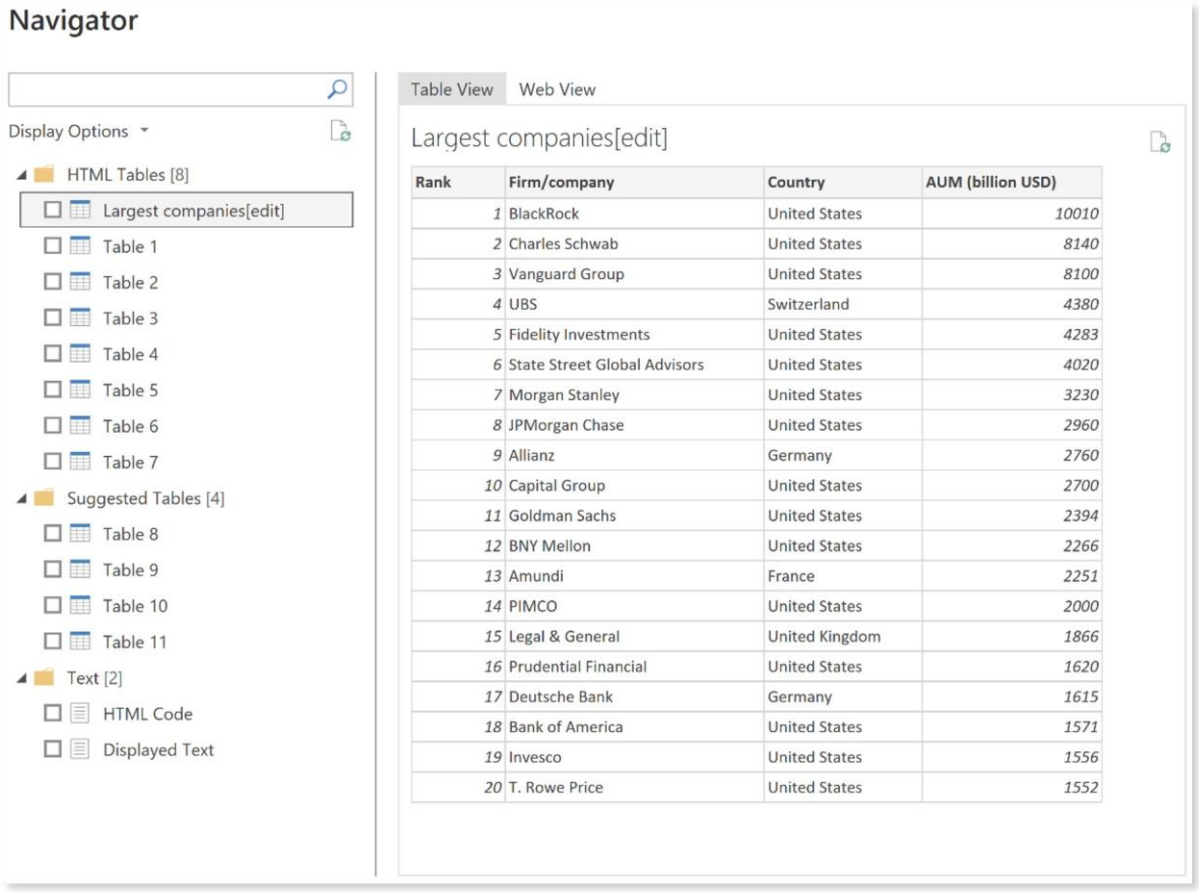
An **asset management company** (**AMC**) is an asset management / investment management company/firm that invests the pooled funds of retail investors in securities in line with the stated investment objectives. For a fee, the company/firm provides more diversification, liquidity, and professional management consulting service than is normally available to individual investors. The diversification of portfolio is done by investing in such securities which are inversely correlated to each other. Money is collected from investors by way of floating various collective investment schemes, e.g. mutual fund schemes. In general, an AMC is a company that is engaged primarily in the business of investing in, and managing, portfolios of securities. A study by consulting firm Casey Quirk, which is owned by Deloitte, found that asset management firms ended 2020 with record highs in both revenue and assets under management.[1]

Largest companies [edit]

The following is a list of the top 20 asset managers in the world (as of 2022), ranked by total assets under management (AUM):[2]

Rank	Firm/company	Country	AUM (billion USD)
1	BlackRock	United States	9,570
2	Vanguard Group	United States	8,100
3	Fidelity Investments	United States	4,283
4	UBS	Switzerland	4,380
5	State Street Global Advisors	United States	4,020
6	Morgan Stanley	United States	3,230
7	JPMorgan Chase	United States	2,960
8	Crédit Agricole	France	2,875
9	Allianz	Germany	2,760
10	Capital Group	United States	2,700
11	Goldman Sachs	United States	2,394
12	BNY Mellon	United States	2,266
13	Amundi	France	2,251
14	PIMCO	United States	2,000
15	Legal & General	United Kingdom	1,866
16	Edward Jones Investments	United States	1,700
17	PGIM	United States	1,620
18	Deutsche Bank	Germany	1,615
19	Bank of America	United States	1,571
20	Invesco	United States	1,556

https://en.wikipedia.org/wiki/List_of_asset_management_firms



Navigator

Display Options

- HTML Tables [8]
  - Largest companies[edit]
  - Table 1
  - Table 2
  - Table 3
  - Table 4
  - Table 5
  - Table 6
  - Table 7
- Suggested Tables [4]
  - Table 8
  - Table 9
  - Table 10
  - Table 11
- Text [2]
  - HTML Code
  - Displayed Text

Table View | Web View

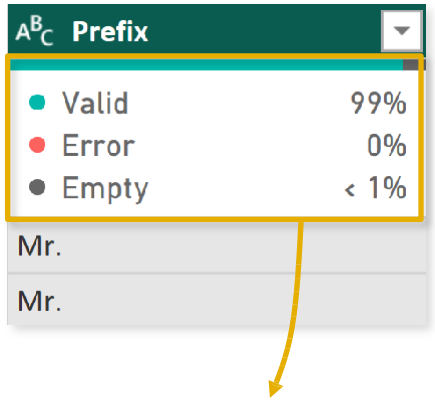
Largest companies[edit]

Rank	Firm/company	Country	AUM (billion USD)
1	BlackRock	United States	10010
2	Charles Schwab	United States	8140
3	Vanguard Group	United States	8100
4	UBS	Switzerland	4380
5	Fidelity Investments	United States	4283
6	State Street Global Advisors	United States	4020
7	Morgan Stanley	United States	3230
8	JPMorgan Chase	United States	2960
9	Allianz	Germany	2760
10	Capital Group	United States	2700
11	Goldman Sachs	United States	2394
12	BNY Mellon	United States	2266
13	Amundi	France	2251
14	PIMCO	United States	2000
15	Legal & General	United Kingdom	1866
16	Prudential Financial	United States	1620
17	Deutsche Bank	Germany	1615
18	Bank of America	United States	1571
19	Invesco	United States	1556
20	T. Rowe Price	United States	1552

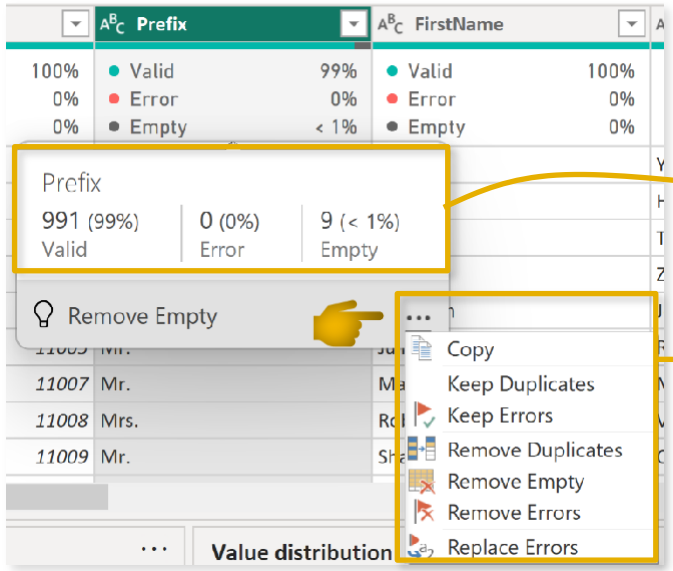
*Copyright Maven Analytics, LLC

DATA PROFILING: COLUMN QUALITY

Profiling tools like **column quality**, **column distribution**, and **column profile** allow you to explore the quality, composition, and distribution of your data before loading it into the Power BI front-end



Column quality *shows the percentage of values within a column that are* ***valid****, contain* ***errors****, or are* ***empty***



Hover over the column quality box to see the ***number of records*** *in each category*

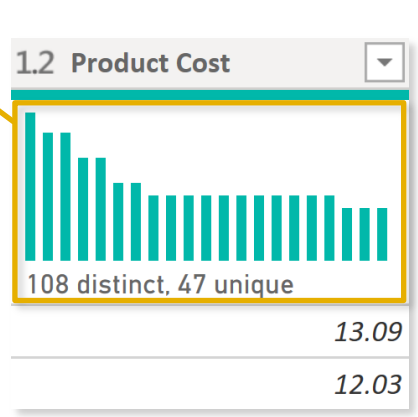
Click the ***options menu*** *to remove duplicates, errors or empty values*

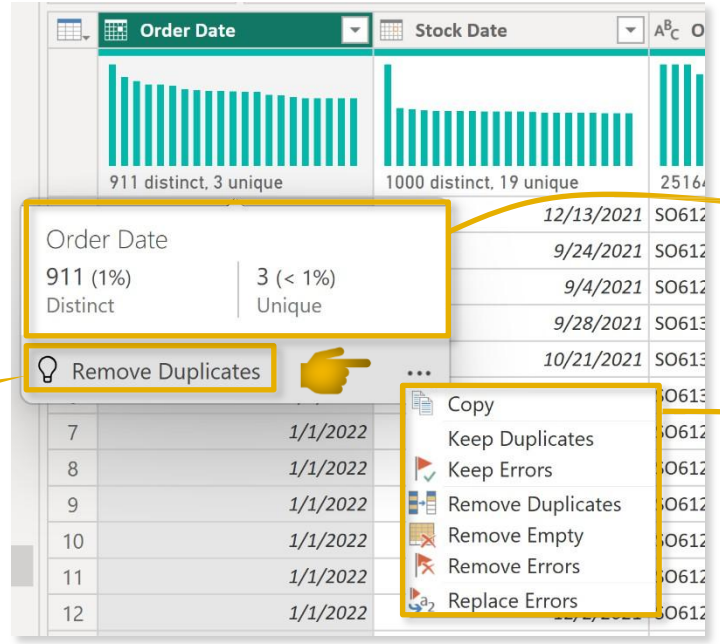
PRO TIP: Profiling tools are a great way to **quickly find and address common data quality issues in one place**, instead of having to manually apply multiple tools or filters

*Copyright Maven Analytics, LLC

DATA PROFILING: COLUMN DISTRIBUTION

Column distribution *provides a sample distribution of the data in a column*





Suggested action *based on column distribution results*

Hover over the column quality box to see the ***number of distinct & unique records***

Click the ***options menu*** *to remove duplicates, errors or empty values*

*Copyright Maven Analytics, LLC

DATA PROFILING: COLUMN PROFILE

Column profile *provides a more holistic view of the data in a column, including a sample distribution and profiling statistics*

Column statistics *provide more detailed profiling metrics, including:*

Count = **293**

(total number of values in column)

Distinct Count = **119**

(total number of distinct values, whether they appear once or multiple times)

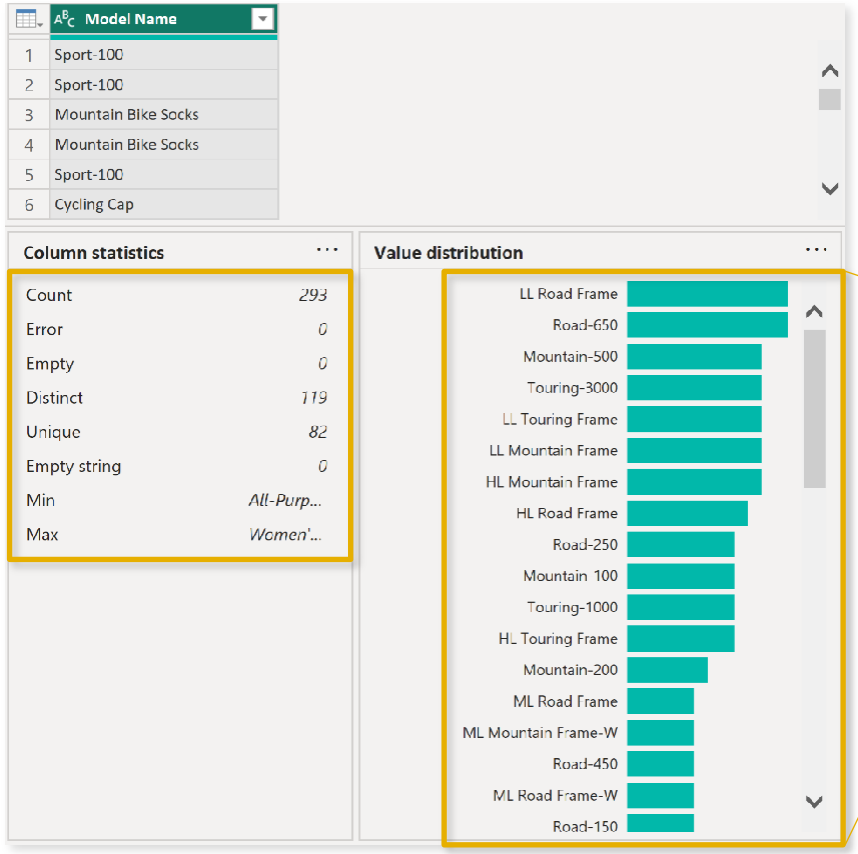
Unique = **82**

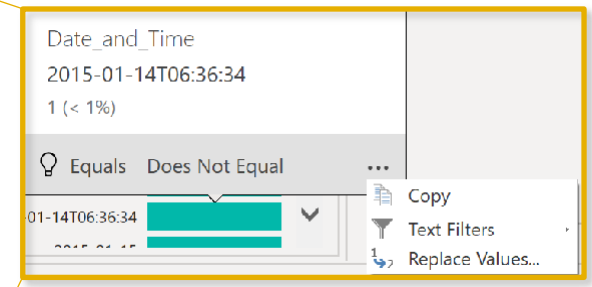
(total number of values that appear exactly once)

Min & Max

(lowest and highest observed values)

Note: *Typically only useful for numerical values*

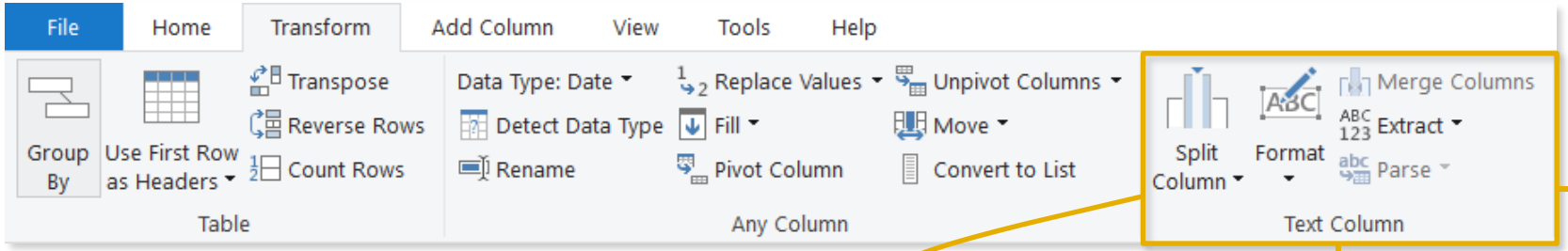




Hover over the value distribution bar for ***suggested transformations*** *and additional options*

*Copyright Maven Analytics, LLC

TEXT TOOLS



Split a text column based on a specific delimiter, number of characters, or other attributes

- By Delimiter
- By Number of Characters
- By Positions
- By Lowercase to Uppercase
- By Uppercase to Lowercase
- By Digit to Non-Digit
- By Non-Digit to Digit

- Length
- First Characters
- Last Characters
- Range
- Text Before Delimiter
- Text After Delimiter
- Text Between Delimiters

Extract characters from text based on fixed lengths, first/last characters, ranges or delimiters

- lowercase
- UPPERCASE
- Capitalize Each Word
- Trim
- Clean
- Add Prefix
- Add Suffix

Format a text column to upper, lower or proper case, or add a prefix or suffix

Tip: *Use "Trim" to eliminate leading & trailing spaces, or "Clean" to remove non-printable characters*

HEY THIS IS IMPORTANT!

You can access many tools from both the **Transform** and **Add Column** menus - the difference is whether you want to **ADD** a new column or **OVERWRITE** an existing one

*Copyright Maven Analytics, LLC

ASSIGNMENT: TEXT TOOLS

NEW MESSAGE

From: **Ethan T. Langer** *(Analytics Manager)*

Subject: **Customer domains**

Hi!

We're looking to better understand where our customers may be coming from, based on their email domains.

Could you please create a new column in the customer table that will allow us do this?

Thanks!
-ETL

Reply Forward

Key Objectives

1. Duplicate the email address column and name it "**Domain Name"**
2. In the new column, remove all text/characters except for the domain name
3. Use transformation steps to clean up and capitalize the domain names (i.e. "**Adventure Works**")
4. Save & Apply changes

*Copyright Maven Analytics, LLC

SOLUTION: TEXT TOOLS

NEW MESSAGE

From: **Ethan T. Langer** *(Analytics Manager)*

Subject: **Customer domains**

Hi!

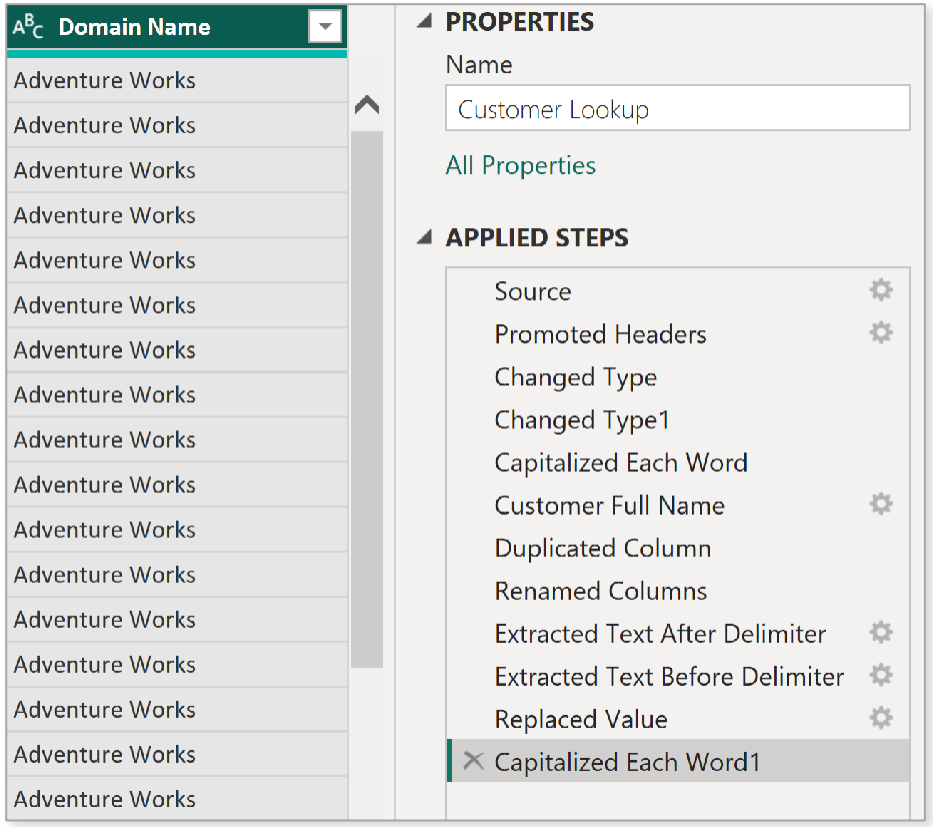
We're looking to better understand where our customers may be coming from, based on their email domains.

Could you please create a new column in the customer table that will allow us do this?

Thanks!
-ETL

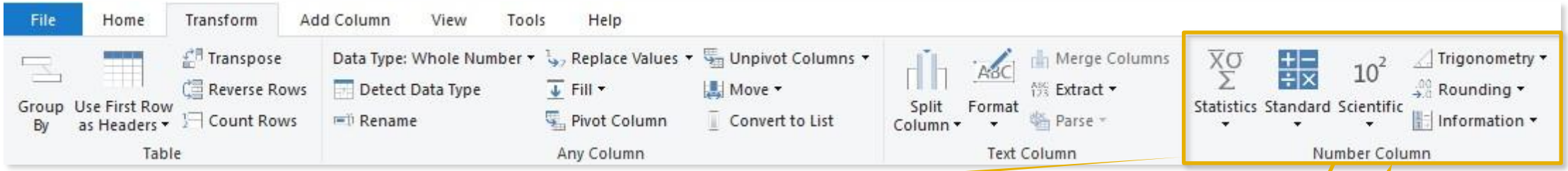
Reply Forward

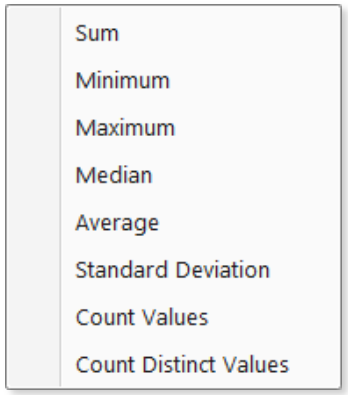
Solution Preview



*Copyright Maven Analytics, LLC

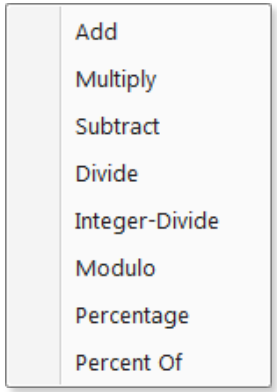
NUMERICAL TOOLS



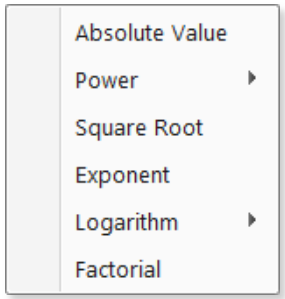


Statistics functions allow you to evaluate basic stats for a selected column (sum, min/max, average, count, count distinct, etc.)

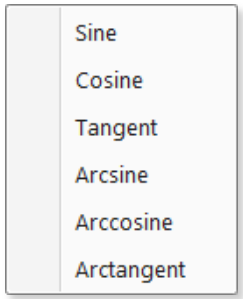
Note: *These tools return a SINGLE value, and are commonly used to explore a table rather than prepare it for loading*



Standard



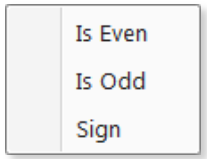
Scientific



Trigonometry

Standard, **Scientific** and **Trigonometry** tools allow you to apply standard operations (addition, multiplication, division, etc.) or more advanced calculations (power, logarithm, sine, tangent, etc.) to each value in a column

Note*: Unlike the Statistics tools, these are applied to each row in the table*



Information tools allow you to define binary flags (*1/0 or TRUE/FALSE*) to mark rows as even, odd, positive or negative

*Copyright Maven Analytics, LLC

ASSIGNMENT: NUMERICAL TOOLS

NEW MESSAGE

From: **Ethan T. Langer** *(Analytics Manager)*

Subject: **Need some stats for leadership**

Hi again,

Leadership is asking us to validate some high-level stats about our products and customers. Can you please help me answer the following questions?

We don't really need to store these values anywhere, so make sure to restore the tables back to their original state once you're done pulling the stats.

Thank you!
-ETL

Reply Forward

Key Objectives

1. What is our average product cost?
2. How many colors do we sell our products in?
3. How many distinct customers do we have?
4. What is the maximum annual customer income?
5. Return the tables to their original state

*Copyright Maven Analytics, LLC

SOLUTION: NUMERICAL TOOLS

NEW MESSAGE

From: **Ethan T. Langer** *(Analytics Manager)*

Subject: **Need some stats for leadership**

Hi again,

Leadership is asking us to validate some high-level stats about our products and customers. Can you please help me answer the following questions?

We don't really need to store these values anywhere, so make sure to restore the tables back to their original state once you're done pulling the stats.

Thank you!
-ETL

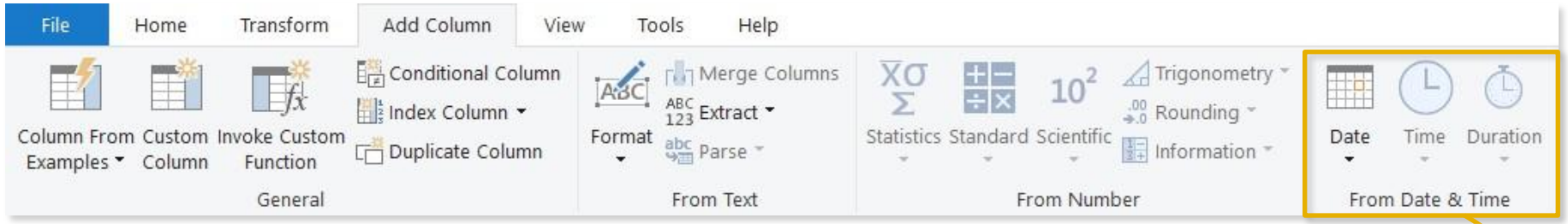
Reply Forward

Solution Preview

1. What is our average product cost? **($413.66)**
2. How many colors do we sell our products in? **(10)**
3. How many distinct customers do we have? **(18,148)**
4. What is the maximum annual customer income? **($170k)**
5. Return the tables to their original state

*Copyright Maven Analytics, LLC

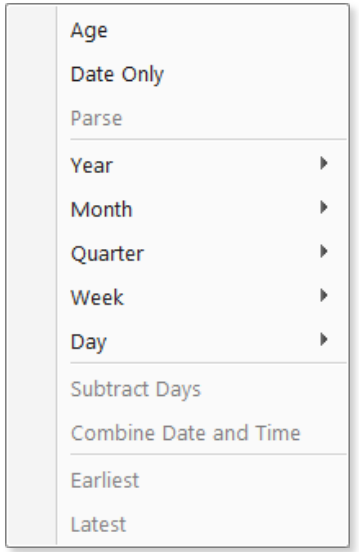
DATE & TIME TOOLS



Date & Time tools are relatively straight-forward, and include the following options:

- **Age**: Difference between the current date and the date in each row
- **Date Only**: Removes the time component from a date/time field
- **Year/Month/Quarter/Week/Day**: Extracts individual components from a date field (time-specific options include Hour, Minute, Second, etc.)
- **Earliest/Latest**: Evaluates the earliest or latest date from a column as a single value (can only be accessed from the "Transform" menu)

***Note**: You will almost always want to perform these operations from the "Add Column" menu to build out new fields, rather than transforming an individual date/time column*

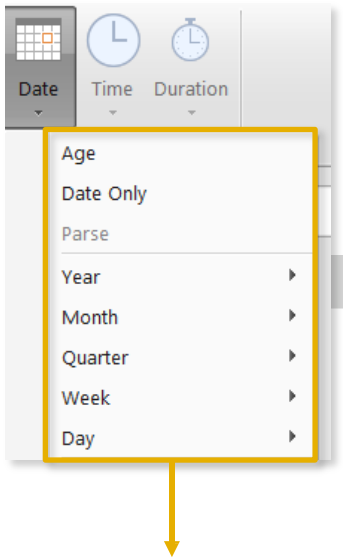


PRO TIP: Load up a table containing a **single date column** and use Date tools to build out an **entire calendar table**

*Copyright Maven Analytics, LLC

CREATING A CALENDAR TABLE

	Date
1	1/1/2020
2	1/2/2020
3	1/3/2020
4	1/4/2020
5	1/5/2020
6	1/6/2020
7	1/7/2020
8	1/8/2020
9	1/9/2020
10	1/10/2020
11	1/11/2020
12	1/12/2020
13	1/13/2020
14	1/14/2020
15	1/15/2020
16	1/16/2020
17	1/17/2020
18	1/18/2020
19	1/19/2020
20	1/20/2020
21	1/21/2020
22	1/22/2020
23	1/23/2020
24	1/24/2020
25	1/25/2020
26	1/26/2020
27	1/27/2020
28	1/28/2020

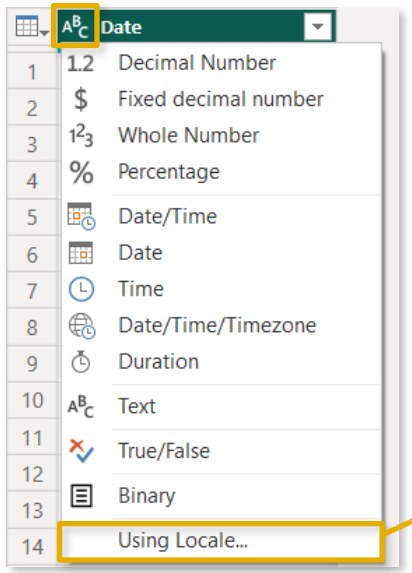


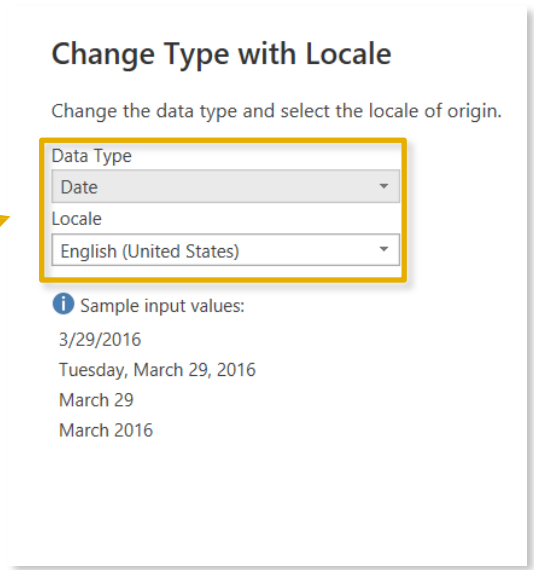
Use the **Date** options in the **Add Column** menu to quickly build out an entire calendar table from a list of dates

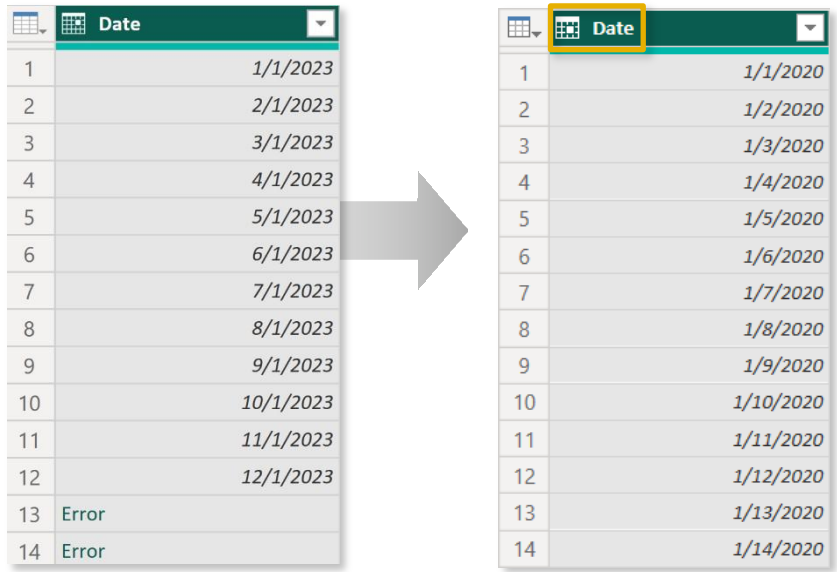
	Date	Day Name	Start of Week	Start of Month	Month Name
1	1/1/2020	Wednesday	12/29/2019	1/1/2020	January
2	1/2/2020	Thursday	12/29/2019	1/1/2020	January
3	1/3/2020	Friday	12/29/2019	1/1/2020	January
4	1/4/2020	Saturday	12/29/2019	1/1/2020	January
5	1/5/2020	Sunday	1/5/2020	1/1/2020	January
6	1/6/2020	Monday	1/5/2020	1/1/2020	January
7	1/7/2020	Tuesday	1/5/2020	1/1/2020	January
8	1/8/2020	Wednesday	1/5/2020	1/1/2020	January
9	1/9/2020	Thursday	1/5/2020	1/1/2020	January
10	1/10/2020	Friday	1/5/2020	1/1/2020	January
11	1/11/2020	Saturday	1/5/2020	1/1/2020	January
12	1/12/2020	Sunday	1/12/2020	1/1/2020	January
13	1/13/2020	Monday	1/12/2020	1/1/2020	January
14	1/14/2020	Tuesday	1/12/2020	1/1/2020	January
15	1/15/2020	Wednesday	1/12/2020	1/1/2020	January
16	1/16/2020	Thursday	1/12/2020	1/1/2020	January
17	1/17/2020	Friday	1/12/2020	1/1/2020	January
18	1/18/2020	Saturday	1/12/2020	1/1/2020	January
19	1/19/2020	Sunday	1/19/2020	1/1/2020	January
20	1/20/2020	Monday	1/19/2020	1/1/2020	January
21	1/21/2020	Tuesday	1/19/2020	1/1/2020	January
22	1/22/2020	Wednesday	1/19/2020	1/1/2020	January
23	1/23/2020	Thursday	1/19/2020	1/1/2020	January
24	1/24/2020	Friday	1/19/2020	1/1/2020	January
25	1/25/2020	Saturday	1/19/2020	1/1/2020	January
26	1/26/2020	Sunday	1/26/2020	1/1/2020	January
27	1/27/2020	Monday	1/26/2020	1/1/2020	January
28	1/28/2020	Tuesday	1/26/2020	1/1/2020	January

*Copyright Maven Analytics, LLC

CHANGE TYPE WITH LOCALE







1) Left click the data type icon in the column header and select the **Using Locale** option

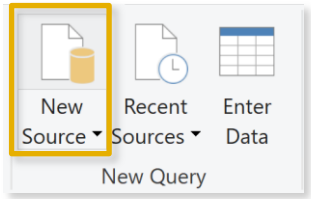
2) Select **Date** as the data type **and English (United States)** as the locale for all datasets in this course (regardless of your actual location)

3) Confirm that **the data type is correctly recognized**. You should see a calendar icon next to the column name in the header and no errors in the column

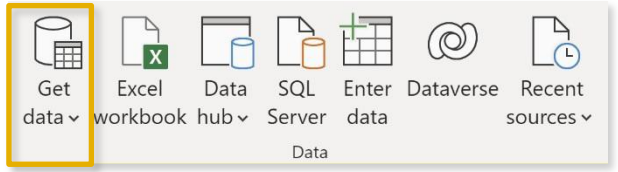
*Copyright Maven Analytics, LLC

PRO TIP: ROLLING CALENDARS

1 Create a new **blank query** & name it "***Rolling Calendar***"

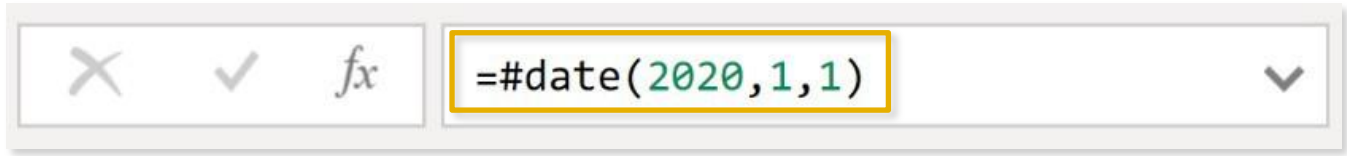


Power Query: *New Source > Blank Query*



Front end: *Get Data > Blank Query*

2 In the formula bar, type a **"literal"** to generate a start date:



Format as: *YYYY, MM, DD*

3 Click the fx icon to **add a custom step**, and enter the following formula to generate a list of dates between the start date and the current day:

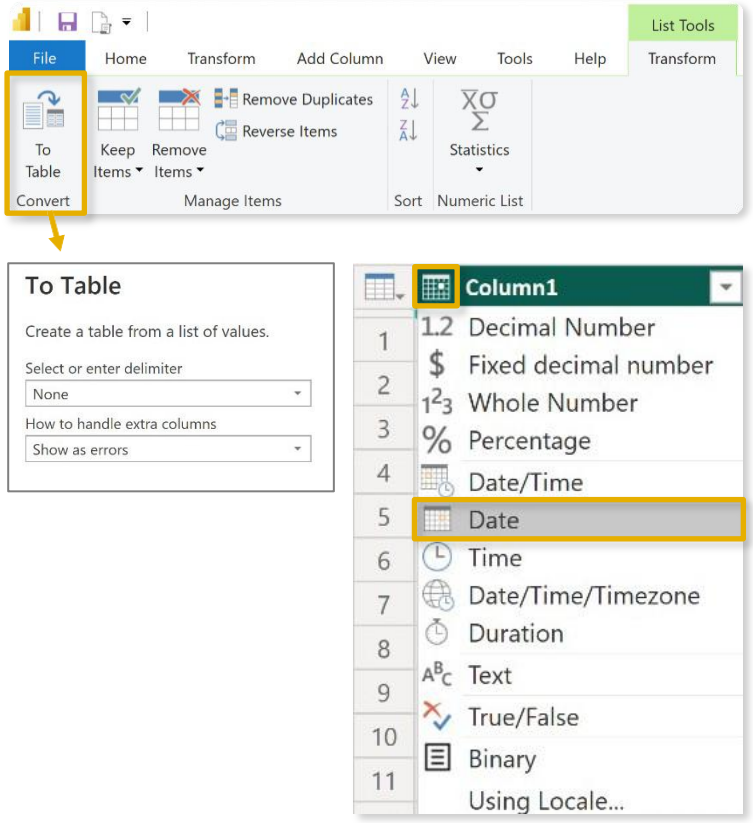
```
= List.Dates(
    Source,
    Number.From(DateTime.LocalNow()) - Number.From(Source),
    #duration(1, 0, 0, 0)
  )
```

Note: *If your first applied step is named something other than "**Source**", use that name in your formula (this is common for non-US users)*

*Copyright Maven Analytics, LLC

PRO TIP: ROLLING CALENDARS

4 Convert the resulting list into a **Table** and set the data type as a **Date**



5 Rename the column to "**Date**" and add calculated date columns (year, month, quarter, etc.) using the **Add Column** tools

	Date	Year	Start of Quarter	Start of Month
1	1/1/2020	2020	1/1/2020	1/1/2020
2	1/2/2020	2020	1/1/2020	1/1/2020
3	1/3/2020	2020	1/1/2020	1/1/2020
4	1/4/2020	2020	1/1/2020	1/1/2020
5	1/5/2020	2020	1/1/2020	1/1/2020
6	1/6/2020	2020	1/1/2020	1/1/2020
7	1/7/2020	2020	1/1/2020	1/1/2020
8	1/8/2020	2020	1/1/2020	1/1/2020
9	1/9/2020	2020	1/1/2020	1/1/2020
10	1/10/2020	2020	1/1/2020	1/1/2020
11	1/11/2020	2020	1/1/2020	1/1/2020
12	1/12/2020	2020	1/1/2020	1/1/2020
13	1/13/2020	2020	1/1/2020	1/1/2020
14	1/14/2020	2020	1/1/2020	1/1/2020
15	1/15/2020	2020	1/1/2020	1/1/2020
16	1/16/2020	2020	1/1/2020	1/1/2020
17	1/17/2020	2020	1/1/2020	1/1/2020
18	1/18/2020	2020	1/1/2020	1/1/2020
19	1/19/2020	2020	1/1/2020	1/1/2020
20	1/20/2020	2020	1/1/2020	1/1/2020
21	1/21/2020	2020	1/1/2020	1/1/2020

*Copyright Maven Analytics, LLC

ASSIGNMENT: CALENDAR TABLES

NEW MESSAGE

From: **Ethan T. Langer** *(Analytics Manager)*

Subject: **New date fields**

Hi,

We need to add a few fields to our calendar table to help us analyze sales trending over time.

Could you please add the following columns when you get a chance?

Thanks!
-ETL

Reply Forward

Key Objectives

Add the following columns to the calendar table:

1. **Month Name** (e.g. “January”)
2. **Month Number** (e.g. “1”)
3. **Start of Year** (e.g. “1/1/2020”)
4. **Year** (e.g. “2020”)

*Copyright Maven Analytics, LLC

SOLUTION: CALENDAR TABLES

NEW MESSAGE

From: **Ethan T. Langer** *(Analytics Manager)*

Subject: **New date fields**

Hi,

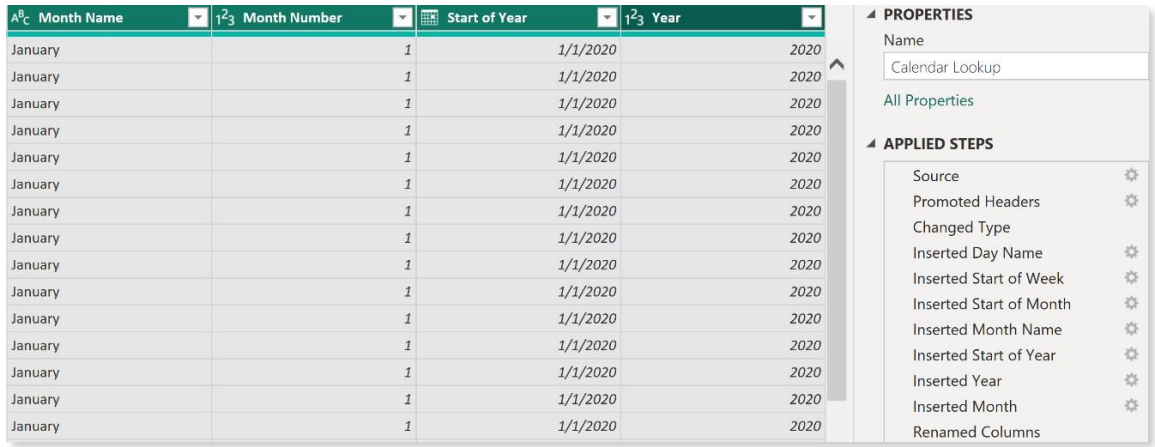
We need to add a few fields to our calendar table to help us analyze sales trending over time.

Could you please add the following columns when you get a chance?

Thanks!
-ETL

Reply Forward

Solution Preview



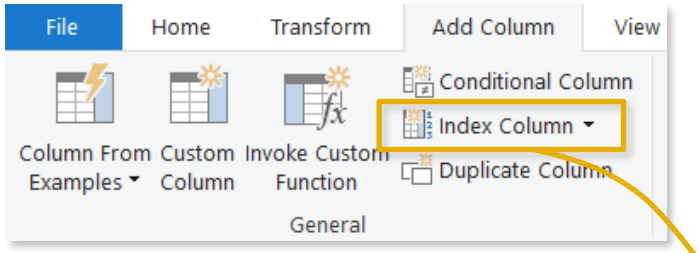
Month Name	Month Number	Start of Year	Year
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020
January	1	1/1/2020	2020

PROPERTIES
Name
Calendar Lookup
All Properties

APPLIED STEPS
Source
Promoted Headers
Changed Type
Inserted Day Name
Inserted Start of Week
Inserted Start of Month
Inserted Month Name
Inserted Start of Year
Inserted Year
Inserted Month
Renamed Columns

*Copyright Maven Analytics, LLC

INDEX COLUMNS



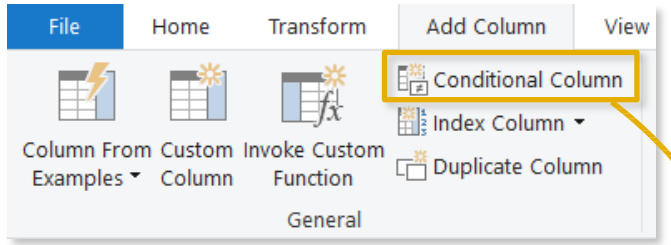
Index Columns contain a list of sequential values that can be used to identify each unique row in a table (*typically starting from 0 or 1*)

These are often used to create **unique IDs** that can be used to form relationships between tables (*more on that later!*)

	Index	Order Date	Stock Date	Order Number	Product Key
1	1	1/1/2020	9/21/2019	SO45080	332
2	2	1/1/2020	12/5/2019	SO45079	312
3	3	1/1/2020	10/29/2019	SO45082	350
4	4	1/1/2020	11/16/2019	SO45081	338
5	5	1/2/2020	12/15/2019	SO45083	312
6	6	1/2/2020	10/12/2019	SO45084	310
7	7	1/2/2020	12/18/2019	SO45086	314
8	8	1/2/2020	10/9/2019	SO45085	312
9	9	1/3/2020	10/3/2019	SO45093	312
10	10	1/3/2020	9/29/2019	SO45090	310
11	11	1/3/2020	12/11/2019	SO45088	345
12	12	1/3/2020	10/24/2019	SO45092	313
13	13	1/3/2020	12/16/2019	SO45089	351
14	14	1/3/2020	10/26/2019	SO45091	314
15	15	1/3/2020	9/11/2019	SO45087	350
16	16	1/3/2020	9/11/2019	SO45094	310
17	17	1/4/2020	10/30/2019	SO45096	312
18	18	1/4/2020	10/30/2019	SO45097	313
19	19	1/4/2020	9/15/2019	SO45098	310
20	20	1/4/2020	12/7/2019	SO45095	344

*Copyright Maven Analytics, LLC

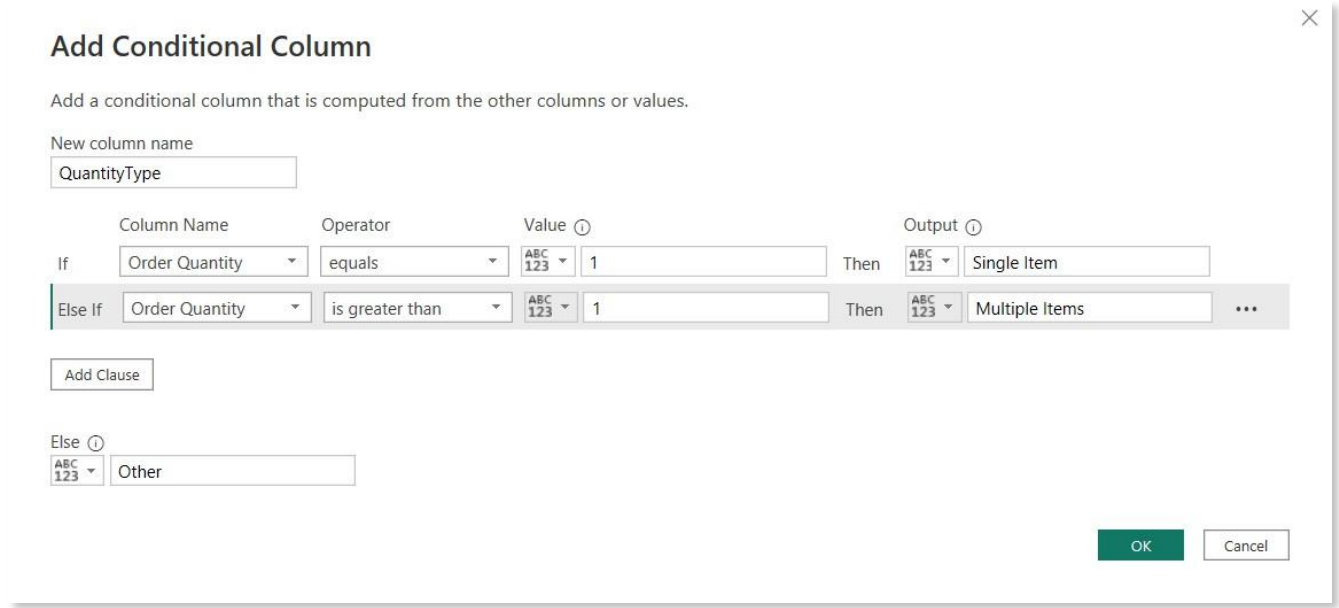
CONDITIONAL COLUMNS



Conditional Columns allow you to define new fields based on logical rules and conditions (IF/THEN statements)

Here we're creating a conditional column named **Quantity Type**, which is based on **Order Quantity**:

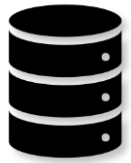
- *If Order Quantity **=1**, Quantity Type = "**Single Item**"*
- *Else If Order Quantity **>1**, Quantity Type = "**Multiple Items**"*
- *Else; Quantity Type = "**Other**"*



*Copyright Maven Analytics, LLC

CALCULATED COLUMN BEST PRACTICES

As a best practice, table transformations and column calculations should ideally happen **as close to the original data source as possible**, to optimize performance and speed



Data Source

Power Query

Power BI Front-End

Published Reports

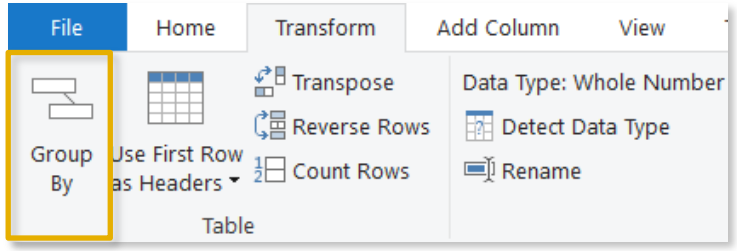
UPSTREAM → **DOWNSTREAM**

HEY THIS IS IMPORTANT!

This is not a strict rule or requirement but can significantly impact performance for very large or complex data models. Where you define calculations often depends on several factors *(accessibility, complexity, business requirements, etc.)*, so we will practice creating columns using both Power Query and the Power BI front-end (DAX) throughout this course

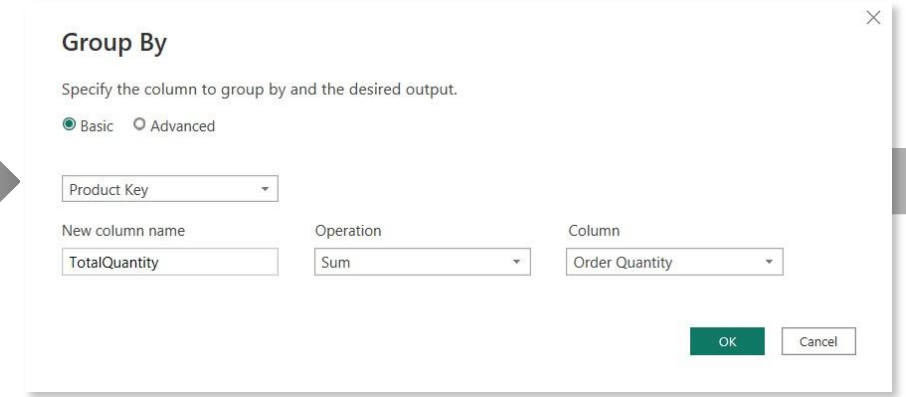
*Copyright Maven Analytics, LLC

GROUPING & AGGREGATING



Group By allows you to aggregate data at a different level or "grain" (*i.e. group daily records into monthly, aggregate transactions by store, etc.*)

	Order Date	Product Key	Customer Key	Order Quantity
1	6/25/2022	214	14719	1
2	10/8/2021	214	21990	1
3	12/30/2021	214	22098	1
4	6/29/2022	214	22748	1
5	8/16/2021	214	27821	1
6	10/9/2021	214	15685	1
7	8/9/2021	214	14951	1
8	1/19/2022	214	23101	1
9	9/23/2021	214	17158	1
10	1/19/2022	214	24196	1
11	6/29/2022	214	12963	1
12	9/13/2021	214	12715	1
13	10/2/2021	214	14846	1
14	7/31/2021	214	11290	1
15	11/24/2021	214	22103	1
16	8/1/2021	214	16982	1
17	10/12/2021	214	20410	1
18	9/10/2021	214	14217	1
19	10/22/2021	214	19642	1
20	8/11/2021	214	11666	1



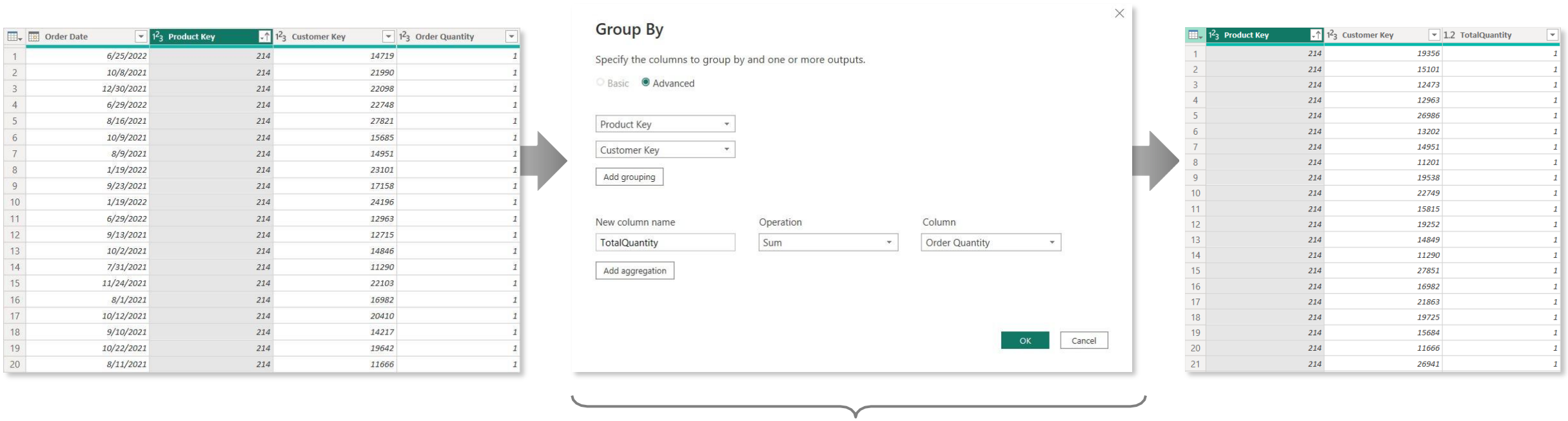
	Product Key	TotalQuantity
1	214	2099
2	215	1940
3	220	1995
4	223	4151
5	226	392
6	229	408
7	232	424
8	235	381
9	310	169
10	311	139
11	312	179
12	313	168
13	314	157
14	320	65
15	322	39
16	324	72
17	326	65

*Here we're transforming a daily, transaction-level table into a summary of **Total Quantity** by **Product Key***

***NOTE:** Any fields not specified in the Group By settings are lost*

*Copyright Maven Analytics, LLC

GROUPING & AGGREGATING



	Order Date	Product Key	Customer Key	Order Quantity
1	6/25/2022	214	14719	1
2	10/8/2021	214	21990	1
3	12/30/2021	214	22098	1
4	6/29/2022	214	22748	1
5	8/16/2021	214	27821	1
6	10/9/2021	214	15685	1
7	8/9/2021	214	14951	1
8	1/19/2022	214	23101	1
9	9/23/2021	214	17158	1
10	1/19/2022	214	24196	1
11	6/29/2022	214	12963	1
12	9/13/2021	214	12715	1
13	10/2/2021	214	14846	1
14	7/31/2021	214	11290	1
15	11/24/2021	214	22103	1
16	8/1/2021	214	16982	1
17	10/12/2021	214	20410	1
18	9/10/2021	214	14217	1
19	10/22/2021	214	19642	1
20	8/11/2021	214	11666	1

	Product Key	Customer Key	TotalQuantity
1	214	19356	1
2	214	15101	1
3	214	12473	1
4	214	12963	1
5	214	26986	1
6	214	13202	1
7	214	14951	1
8	214	11201	1
9	214	19538	1
10	214	22749	1
11	214	15815	1
12	214	19252	1
13	214	14849	1
14	214	11290	1
15	214	27851	1
16	214	16982	1
17	214	21863	1
18	214	19725	1
19	214	15684	1
20	214	11666	1
21	214	26941	1

*This time we're transforming the daily, transaction-level table into a summary of **Total Quantity** grouped by both **Product Key** and **Customer Key** (using the "Advanced" option)*

***NOTE:** This is like creating a PivotTable in Excel and pulling in **Sum of Order Quantity** with **Product Key** and **Customer Key** as row labels*

*Copyright Maven Analytics, LLC

PIVOTING & UNPIVOTING

Pivoting describes the process of turning **distinct row values into columns**, and **unpivoting** describes the process of turning **distinct columns into rows**

	Date	Product Category	North Region	Central Region	South Region
1	7/1/2022	Bikes	10	19	25
2	7/1/2022	Components	14	31	16
3	7/1/2022	Clothing	35	32	46

PIVOT

UNPIVOT

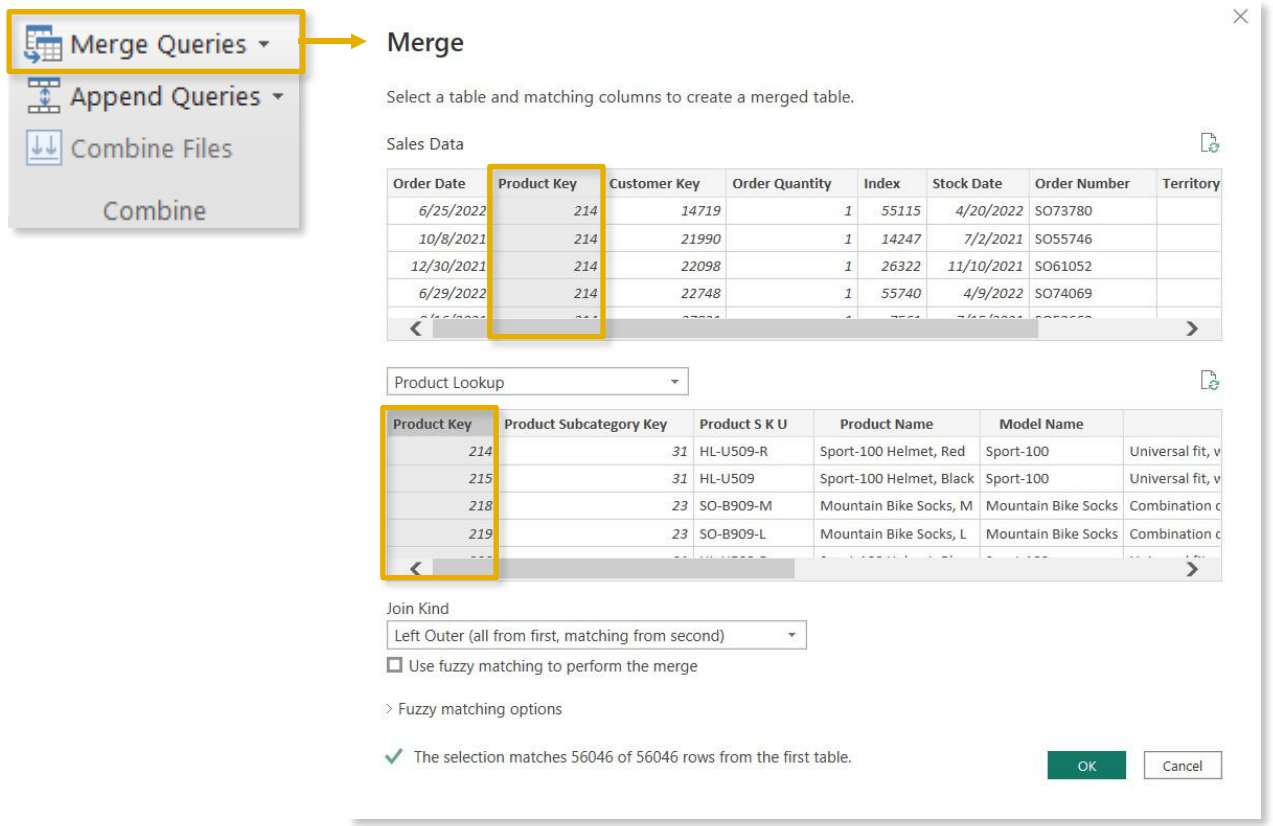
	Date	Product Category	Region	Quantity Sold
1	7/1/2022	Bikes	North Region	10
2	7/1/2022	Bikes	Central Region	19
3	7/1/2022	Bikes	South Region	25
4	7/1/2022	Components	North Region	14
5	7/1/2022	Components	Central Region	31
6	7/1/2022	Components	South Region	16
7	7/1/2022	Clothing	North Region	35
8	7/1/2022	Clothing	Central Region	32
9	7/1/2022	Clothing	South Region	46

*Imagine the table on a hinge; **pivoting** rotates it from **vertical** to **horizontal**, and **unpivoting** rotates it from **horizontal** to **vertical***

***NOTE: Transpose** works very similarly, but doesn't recognize unique values; instead, the entire table is transformed so that each row becomes a column and vice versa*

*Copyright Maven Analytics, LLC

MERGING QUERIES



Merging queries allows you to **join tables** based on a common column (like a lookup in Excel)

*In this case we're merging the **Sales Data** table with the **Product Lookup** table, which share a common **Product Key** column*

***NOTE**: Merging **adds columns** to an existing table/query*

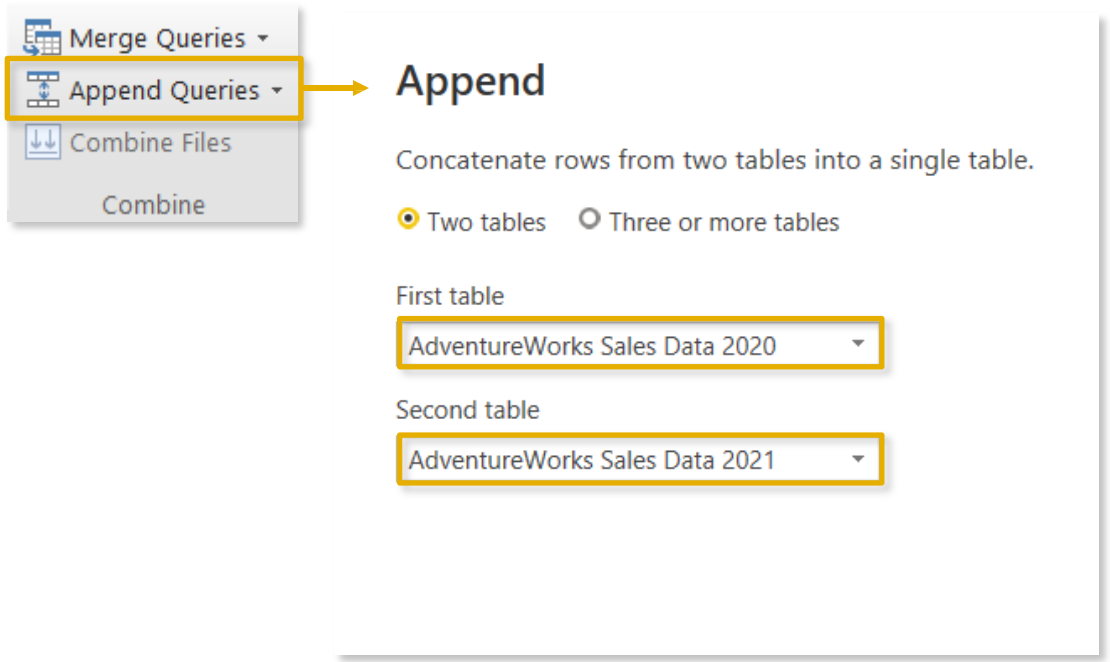
HEY THIS IS IMPORTANT!

Just because you can merge tables, doesn't mean you should!

In many cases, it's better to keep tables separate and define **relationships** between them in the data model (*more on that soon!*)

*Copyright Maven Analytics, LLC

APPENDING QUERIES



Appending queries allows you to **combine** or **stack** tables sharing the exact same column structure and data types

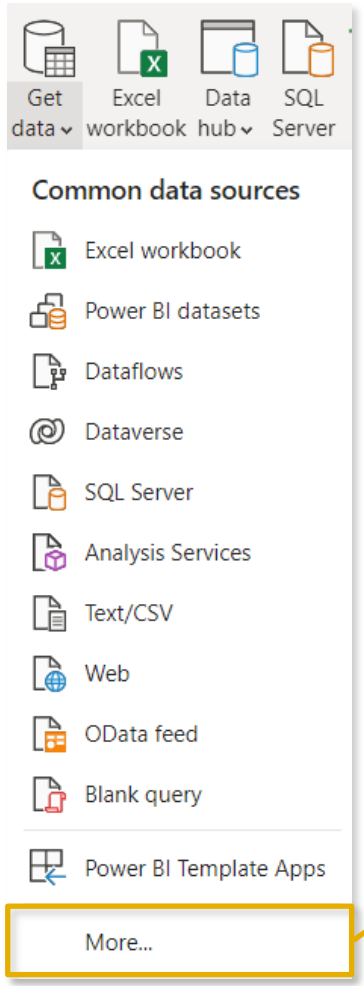
*Here we're appending the **AdventureWorks Sales 2020** table to the **AdventureWorks Sales 2021** table, which is valid since they share identical table structures*

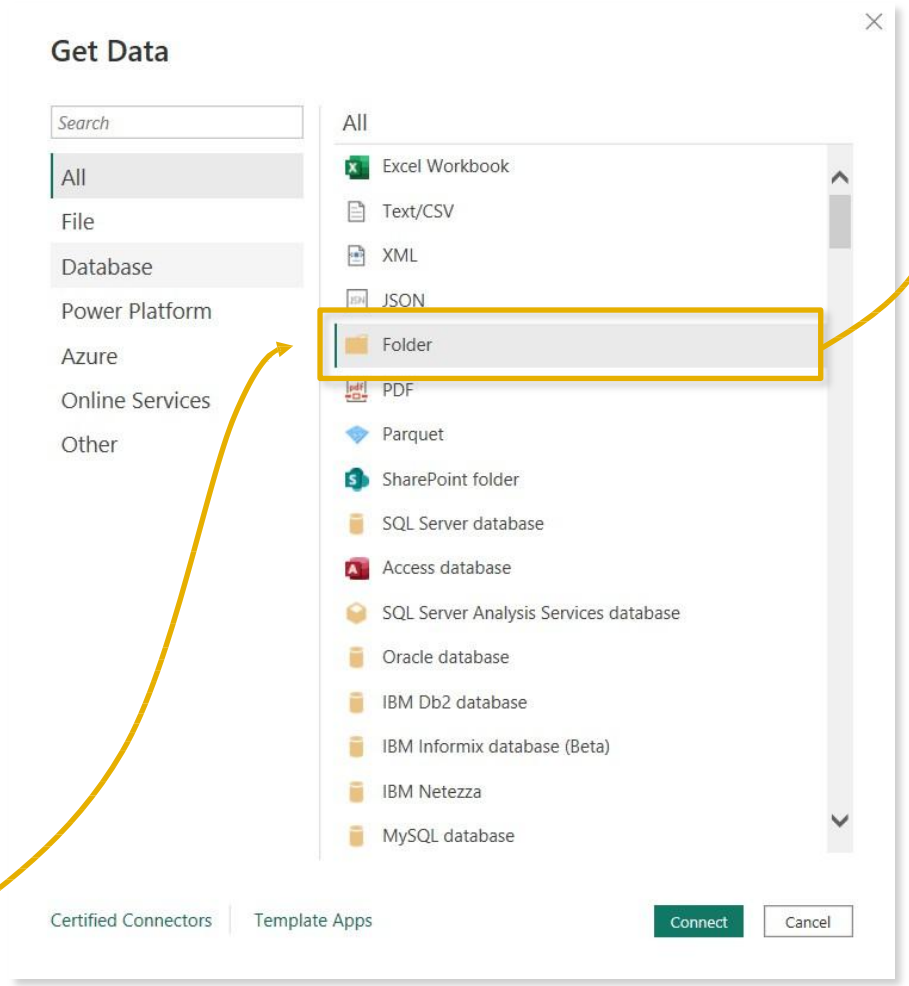
***NOTE:** Appending **adds rows** to an existing table/query*

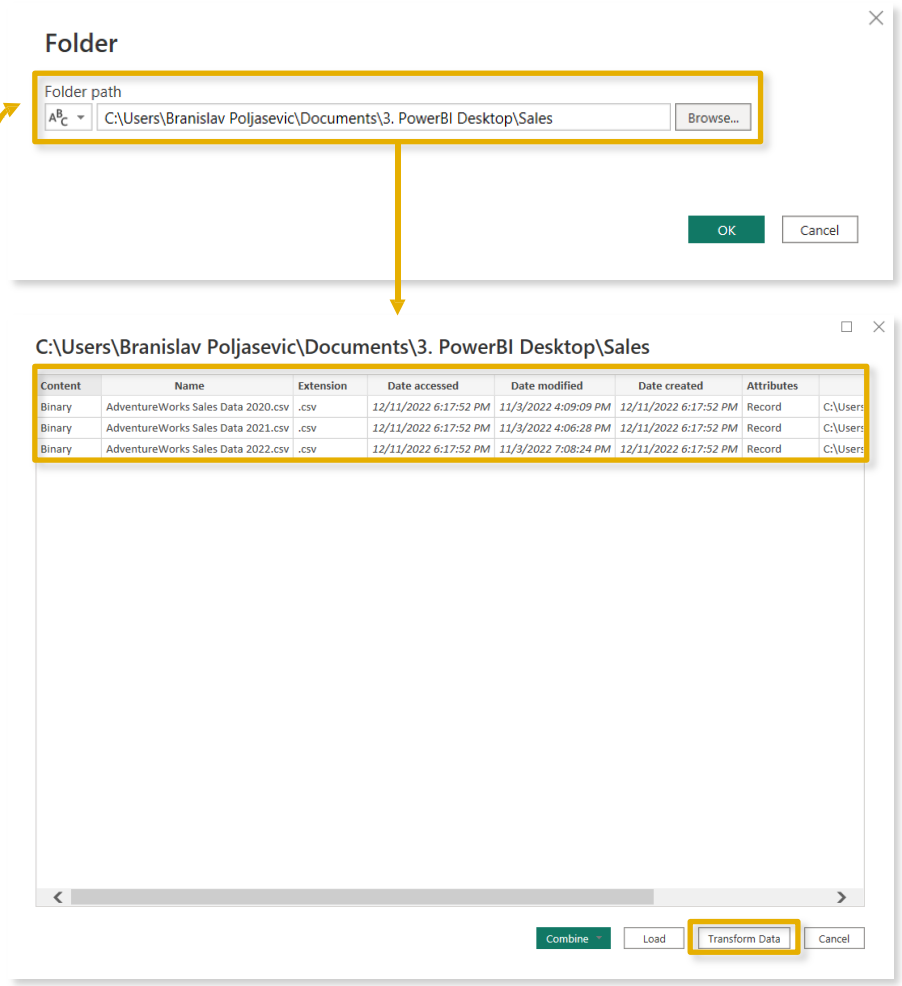
PRO TIP: *Use the **Folder** option (Get Data > More > Folder) to **append all files within a specified folder** (assuming they share the same structure); as you add new files, simply refresh the query and they will automatically append!*

*Copyright Maven Analytics, LLC

PRO TIP: APPENDING FILES FROM A FOLDER



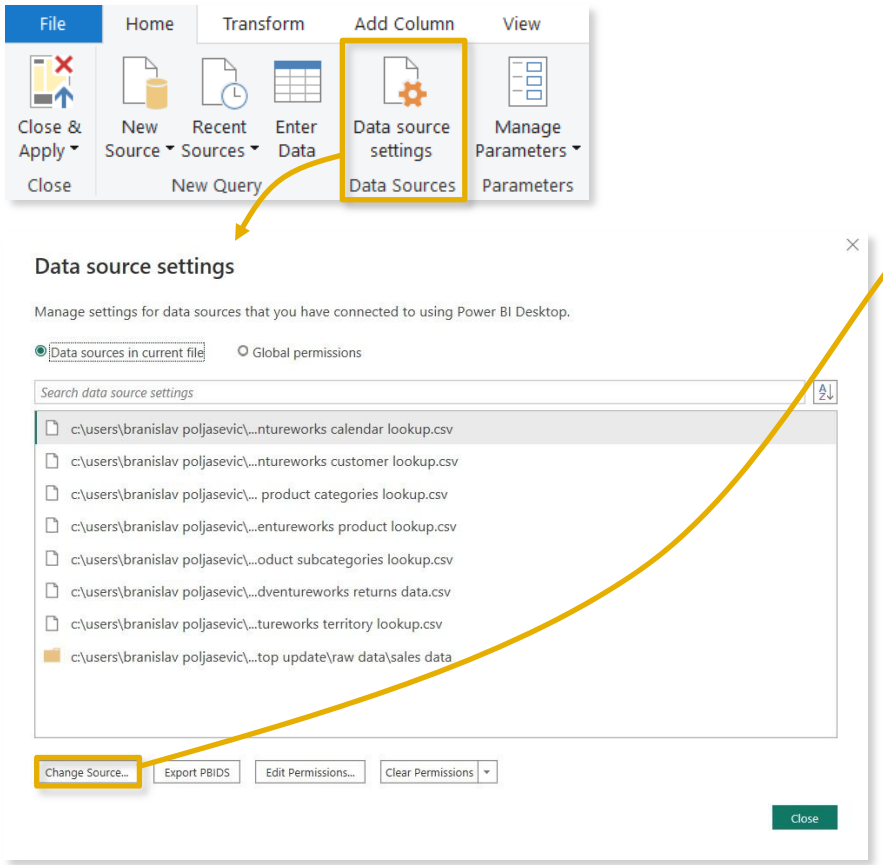


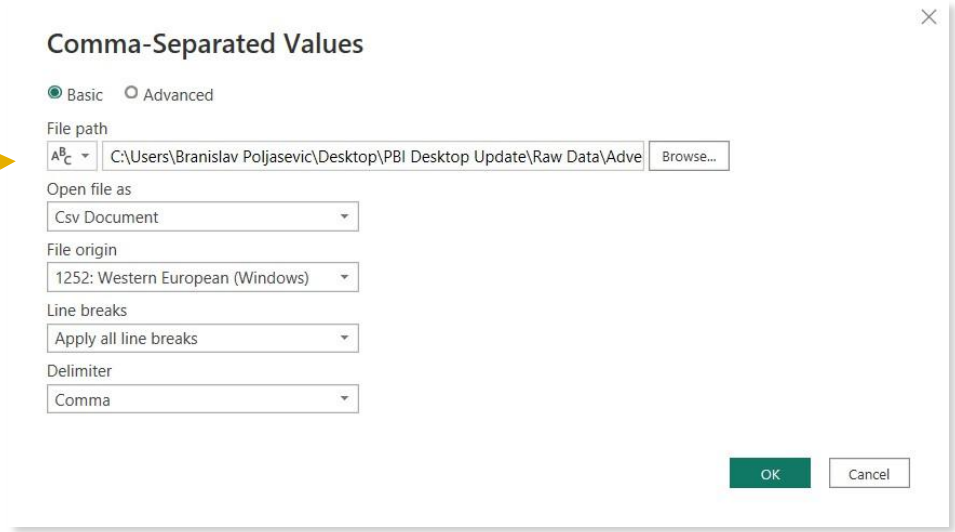


*Copyright Maven Analytics, LLC

DATA SOURCE SETTINGS

Data Source Settings allow you to manage existing data connections, file paths and permissions





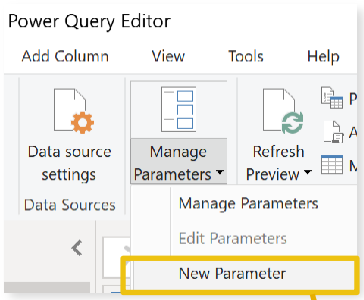
HEY THIS IS IMPORTANT!

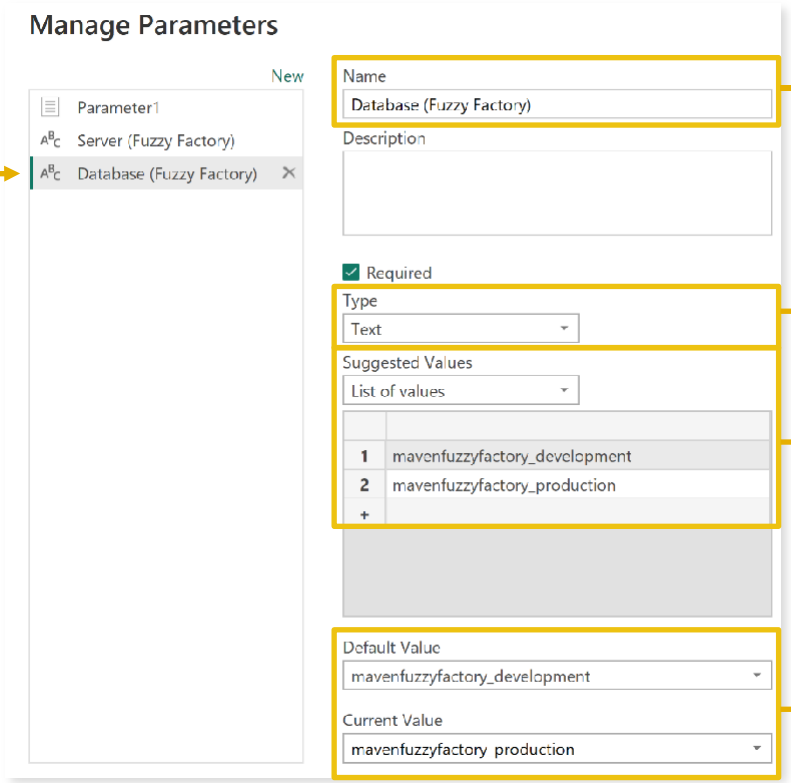
Connections to local files reference the **exact file path**, so if the file name or location changes you will need to update your data source settings

*Copyright Maven Analytics, LLC

PRO TIP: DATA SOURCE PARAMETERS

Use **parameters** to dynamically manage and update connection paths in the Power Query editor



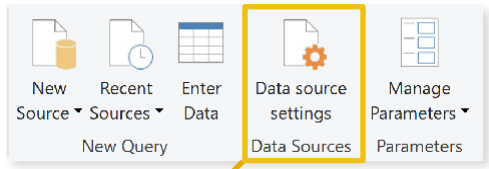


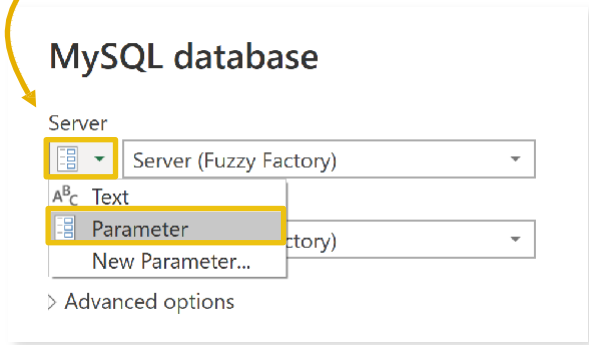
*Parameter **name***
(Name of the query/table)

*Parameter **type***
(Any value, text, date, etc.)

*Parameter **value***
(Any value, list, query)

*Parameter **type***
(Default & current)

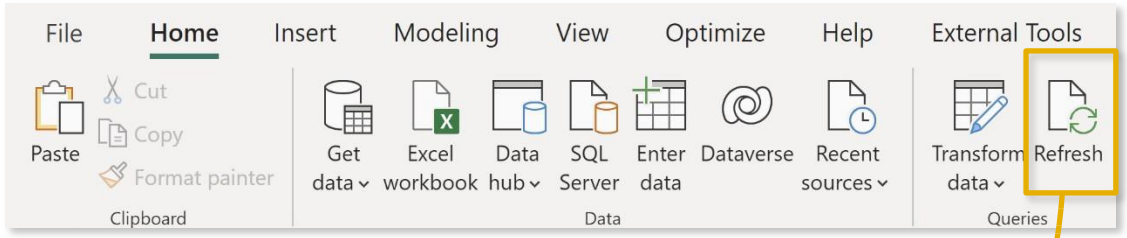




Update** Server & Database connection text values with **parameters

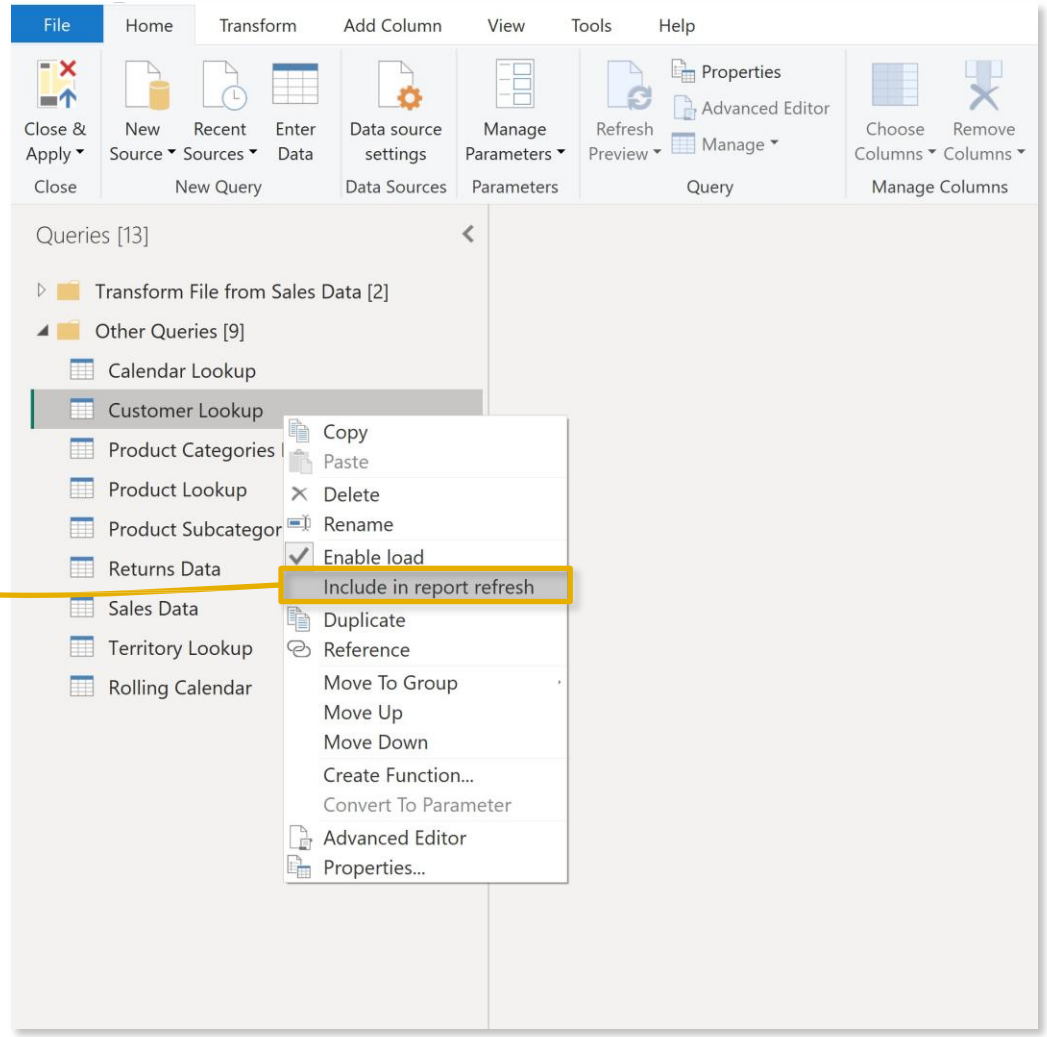
*Copyright Maven Analytics, LLC

REFRESHING QUERIES



By default, ***all queries*** will refresh when you use the **Refresh** command from the **Home** tab

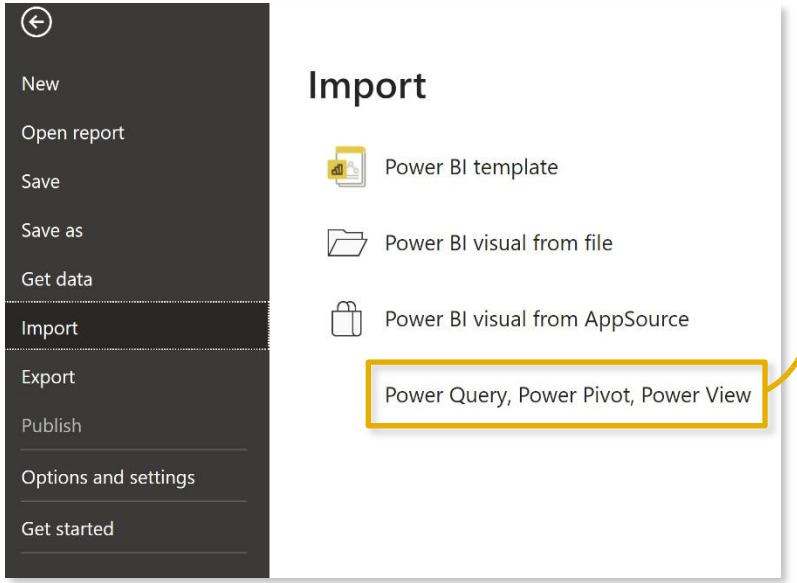
From the Query Editor, uncheck **Include in report refresh** to exclude individual queries from the refresh



PRO TIP: Exclude queries from refresh that don't change often (like lookups or static data tables)

*Copyright Maven Analytics, LLC

PRO TIP: IMPORTING EXCEL MODELS



Already have a fully-built model in Excel?

You can import models built in Excel directly into Power BI Desktop using: ***Import > Power Query, Power Pivot, Power View***

Imported models retain the following:

- Data source **connections** and **queries**
- Query editing procedures and **applied steps**
- Table **relationships**, **hierarchies**, **field settings**, etc.
- All **calculated columns** and **DAX measures**

PRO TIP: If you are more comfortable working in Excel, build your models there first then import to Power BI!

**In older versions of Power BI, this import option may be called "Excel Workbook Contents"*

**Copyright Maven Analytics, LLC*

POWER QUERY BEST PRACTICES

Get organized before connecting and loading data

- *Define clear and intuitive table/query names from the start, and establish an organized file/folder structure if you are working with local flat files to avoid changes to file names or paths*

Disable report refresh for any static data sources

- *There's no need to constantly refresh data sources that don't change, like lookups or static data tables*

When working with large tables, only load the data you need

- *Don't include hourly data when you only need daily, or transaction-level data when only need a product-level summary (extra data will only slow your report down!)*

*Copyright Maven Analytics, LLC