POWER BI

SHARED BY: THUNNM

COURSE OUTLINE

1. **Introducing Power BI Desktop**
   - *Installing Power BI, exploring the Power BI workflow, comparing Power BI vs. Excel, etc.*
2. **Connecting & Shaping Data**
   - *Connecting to source data, shaping and transforming tables, editing, merging and appending queries, etc.*
3. **Creating a Data Model**
   - *Building relational models, creating table relationships, understanding cardinality, exploring filter flow, etc.*
4. **Adding Calculated Fields with DAX**
   - *Understanding DAX syntax, adding calculated columns and measures, writing common formulas & functions, etc.*
5. **Visualizing Data with Reports**
   - *Inserting charts and visuals, customizing formats, editing interactions, applying filters and bookmarks, etc.*

INTRODUCING THE COURSE

1. **THE BRIEF**

- Your client needs a way to **track KPIs** (*sales, revenue, profit, returns*), **compare regional performance, analyze product-level trends and forecasts,** and **identify high-value customer.**

2. **THE OBJECTIVE**

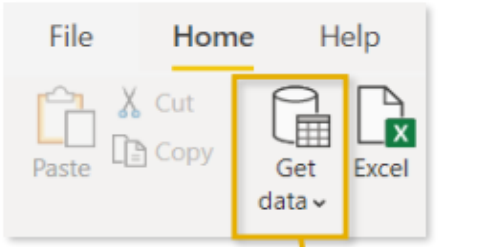
- **Use Power BI Desktop to:**
  - ✓ Connect and transform the raw data
  - ✓ Build a relational data model
  - ✓ Create new calculated columns and DAX measures
  - ✓ Design an interactive report to analyze and visualize the data

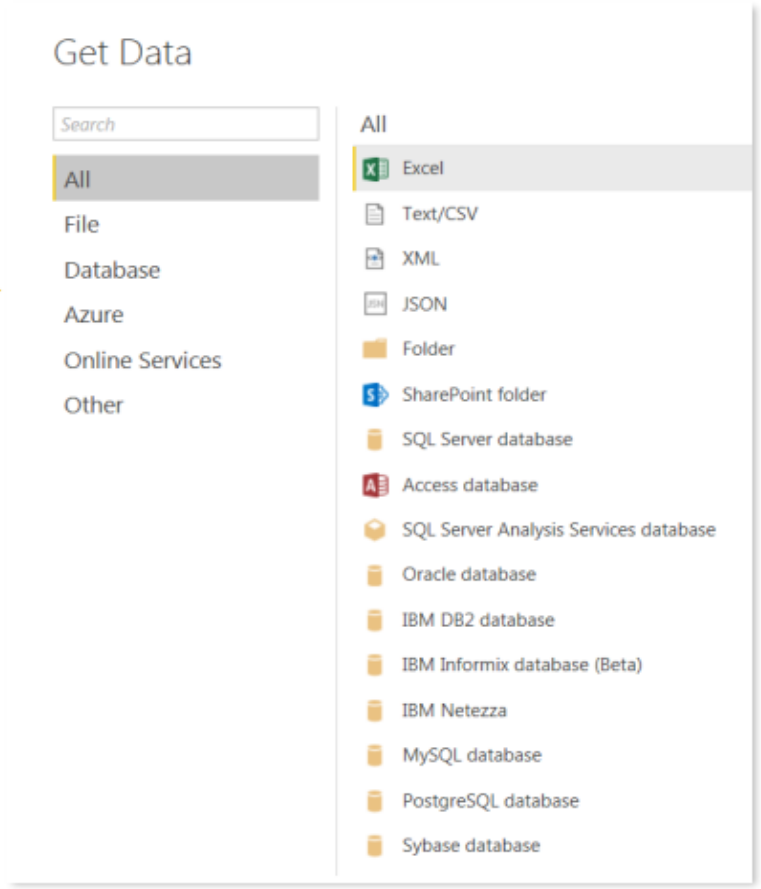
DAY 2 – CONNECTING & SHAPING DATA

OBJECTIVE

- *Connecting to source data, shaping and transforming tables, editing, merging and appending queries, etc*

TYPES OF DATA CONNECTORS

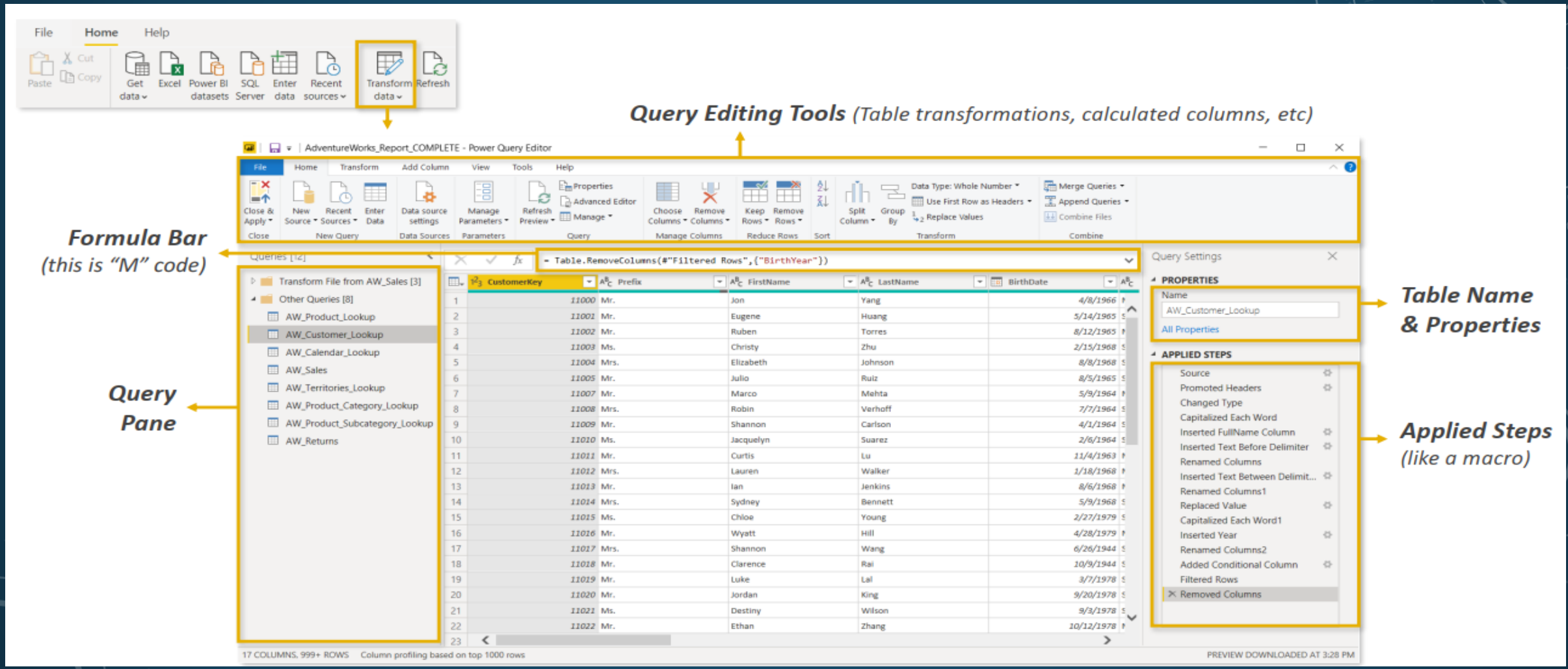




Power BI can connect to virtually **any** type of source data, including (*but not limited to*):

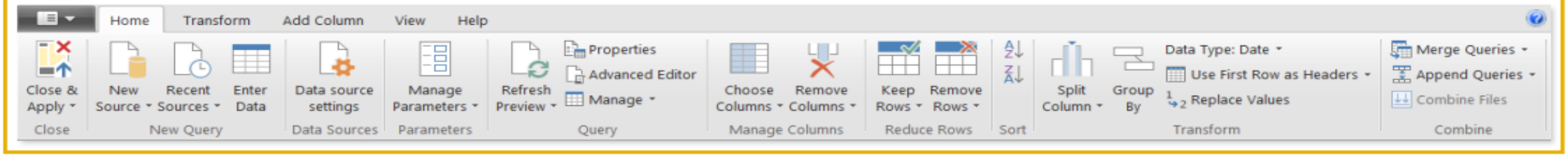
- **Flat files & Folders** (*csv, text, xls, etc*)
- **Databases** (*SQL, Access, Oracle, IBM, Azure, etc*)
- **Online Services** (*Sharepoint, GitHub, Dynamics 365, Google Analytics, Salesforce, Power BI Service, etc*)
- **Others** (*Web feeds, R scripts, Spark, Hadoop, etc*)

THE QUERY EDITOR

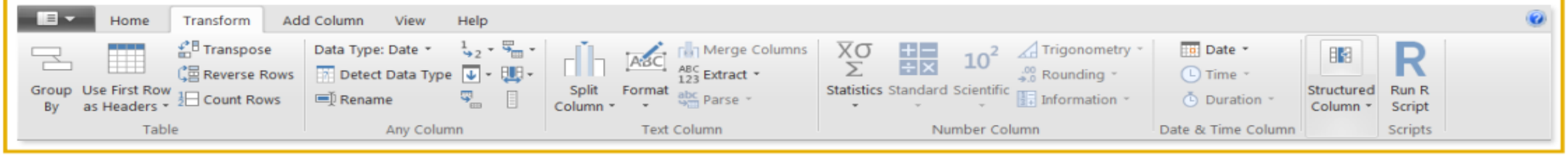


QUERY EDITING TOOLS

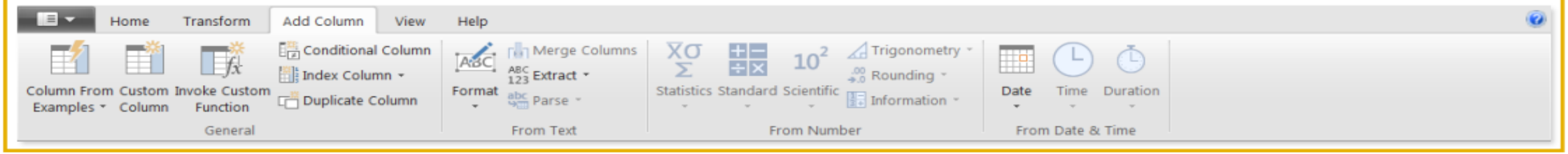
*The **HOME** tab includes **general settings** and **common table transformation tools***



*The **TRANSFORM** tab includes tools to **modify existing columns** (splitting/grouping, transposing, extracting text, etc)*



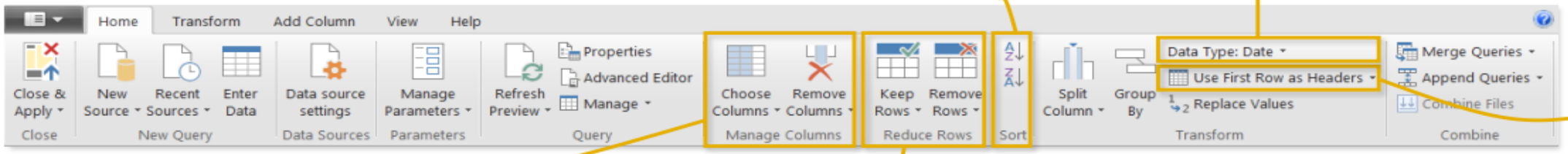
*The **ADD COLUMN** tools **create new columns** (based on conditional rules, text operations, calculations, dates, etc)*



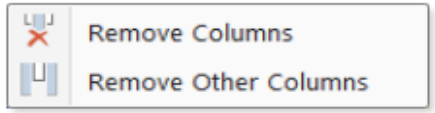
BASIC TABLE TRANSFORMATIONS

Sort values (*A-Z, Low-High, etc.*)

Change data type (*date, $, %, text, etc.*)

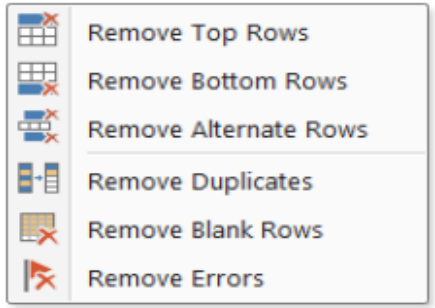


Promote header row



Choose or remove columns

***Tip**: use the "Remove Other Columns" option if you always want a specific set*

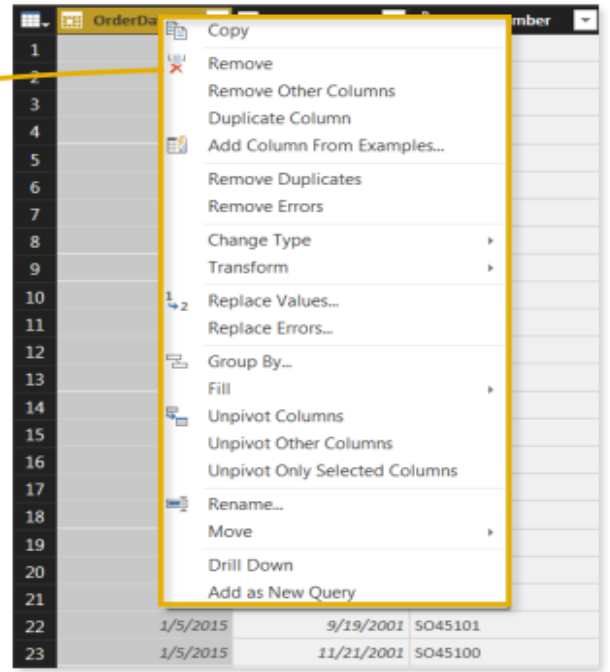


Keep or remove rows

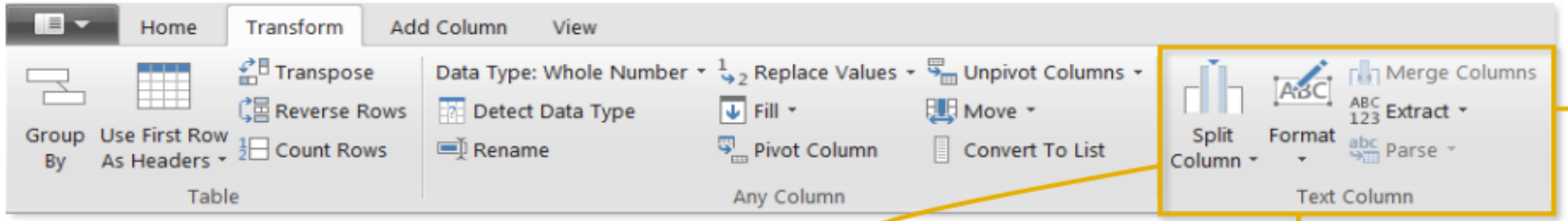
***Tip**: use the "Remove Duplicates" option to create a new lookup table from scratch*

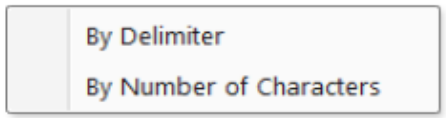
Duplicate, move & rename columns

***Tip**: Right-click the column header to access common tools*

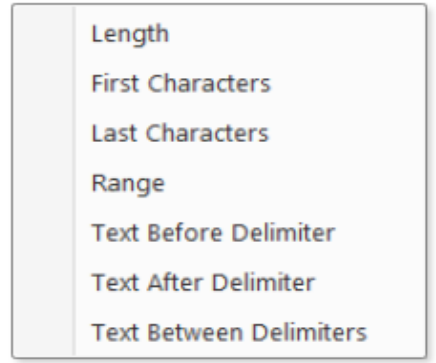


TEXT-SPECIFIC TOOLS



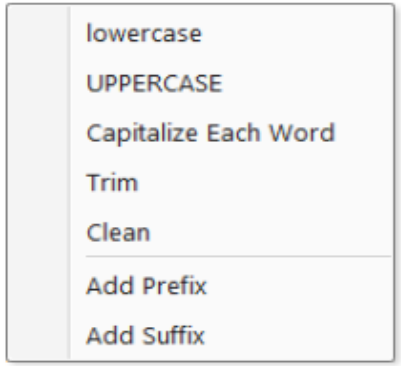


Split a text column based on either a specific delimiter or a number of characters



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

***Tip**: Select two or more columns to **merge** (or **concatenate**) fields*



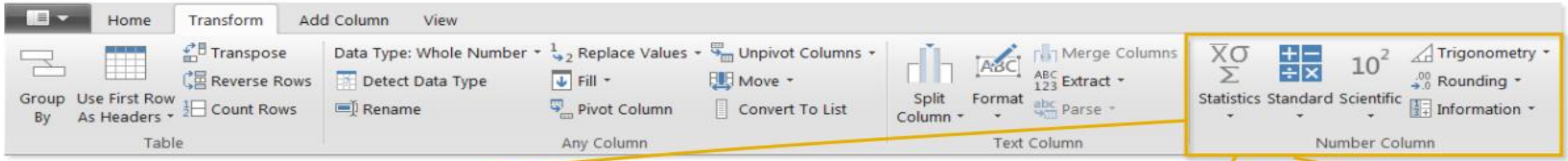
Format a text column to upper, lower or proper case, or add a prefix or suffix

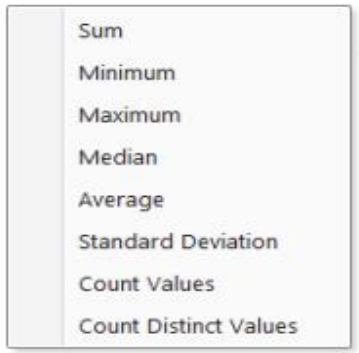
***Tip**: Use "Trim" to eliminate leading & trailing spaces, or "Clean" to remove non-printable characters*

HEY THIS IS IMPORTANT!

You can access many of these tools in both the "**Transform**" and "**Add Column**" menus -- the difference is whether you want to ***add a new column*** or ***modify an existing one***

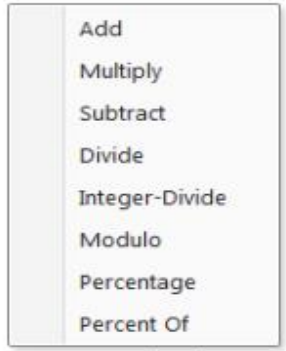
NUMBER-SPECIFIC TOOLS





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

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



Standard

Scientific

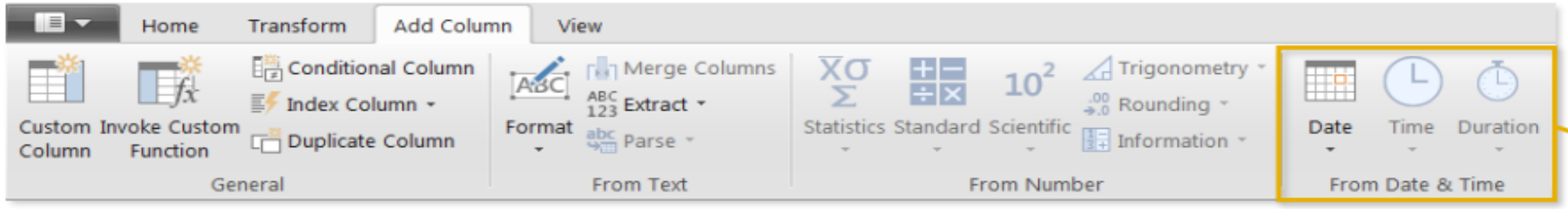
Trigonometry

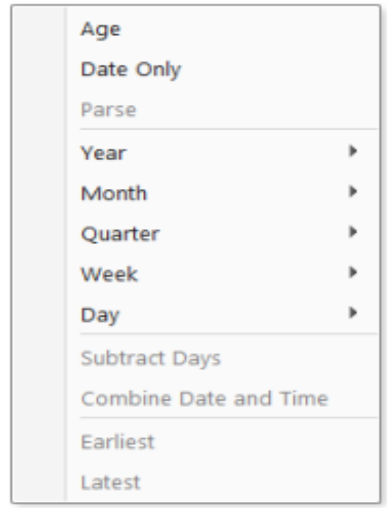
Standard, Scientific and **Trigonometry** tools allow you to apply standard operations (addition, multiplication, division, etc.) or more advanced calculations (power, logarithm, sine, tangent, etc) to each value in a column

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

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

DATE-SPECIFIC TOOLS





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

- **Age**: Difference between the current time and the date in each row
- **Date Only:** Removes the time component of 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)

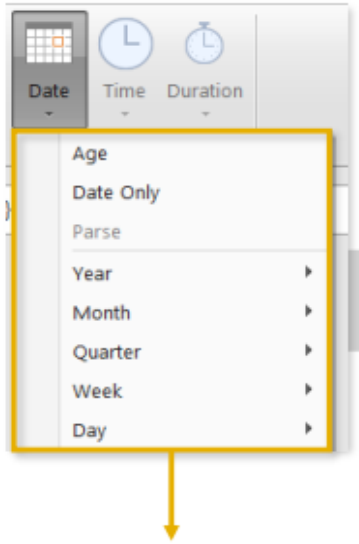
***Note**: You will almost always want to perform these operations from the "Add Column" menu to build out new fields, rather than transforming an individual date/time column*

PRO TIP:

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

CREATING A BASIC CALENDAR TABLE

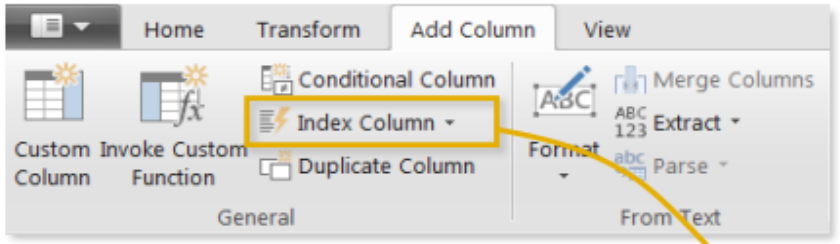
	Date
1	1/1/2015
2	1/2/2015
3	1/3/2015
4	1/4/2015
5	1/5/2015
6	1/6/2015
7	1/7/2015
8	1/8/2015
9	1/9/2015
10	1/10/2015
11	1/11/2015
12	1/12/2015
13	1/13/2015
14	1/14/2015
15	1/15/2015
16	1/16/2015
17	1/17/2015
18	1/18/2015
19	1/19/2015
20	1/20/2015
21	1/21/2015
22	1/22/2015
23	1/23/2015
24	1/24/2015
25	1/25/2015



Use pre-defined **Date** options in the "**Add Column**" menu to quickly build out a calendar table from a list of dates

	Date	Day	Day of Week	Day Name	Start of Week	Month
1	1/1/2015	1	4	Thursday	12/28/2014	1
2	1/2/2015	2	5	Friday	12/28/2014	1
3	1/3/2015	3	6	Saturday	12/28/2014	1
4	1/4/2015	4	0	Sunday	1/4/2015	1
5	1/5/2015	5	1	Monday	1/4/2015	1
6	1/6/2015	6	2	Tuesday	1/4/2015	1
7	1/7/2015	7	3	Wednesday	1/4/2015	1
8	1/8/2015	8	4	Thursday	1/4/2015	1
9	1/9/2015	9	5	Friday	1/4/2015	1
10	1/10/2015	10	6	Saturday	1/4/2015	1
11	1/11/2015	11	0	Sunday	1/11/2015	1
12	1/12/2015	12	1	Monday	1/11/2015	1
13	1/13/2015	13	2	Tuesday	1/11/2015	1
14	1/14/2015	14	3	Wednesday	1/11/2015	1
15	1/15/2015	15	4	Thursday	1/11/2015	1
16	1/16/2015	16	5	Friday	1/11/2015	1
17	1/17/2015	17	6	Saturday	1/11/2015	1
18	1/18/2015	18	0	Sunday	1/18/2015	1
19	1/19/2015	19	1	Monday	1/18/2015	1
20	1/20/2015	20	2	Tuesday	1/18/2015	1
21	1/21/2015	21	3	Wednesday	1/18/2015	1
22	1/22/2015	22	4	Thursday	1/18/2015	1
23	1/23/2015	23	5	Friday	1/18/2015	1
24	1/24/2015	24	6	Saturday	1/18/2015	1
25	1/25/2015	25	0	Sunday	1/25/2015	1

ADDING INDEX COLUMNS

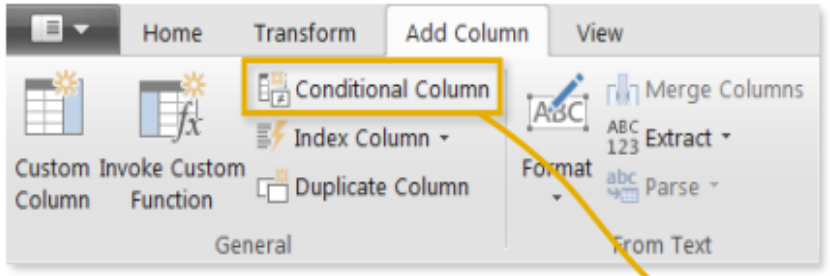


Index Columns contain a list of sequential values that can be used to identify each unique row in a table (*typically starting from 0 or 1*)

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

	Index	OrderDate	StockDate	OrderNumber	ProductKey	CustomerKey
1	1	1/1/2015	9/21/2001	SO45080	332	14657
2	2	1/1/2015	12/5/2001	SO45079	312	29255
3	3	1/1/2015	10/29/2001	SO45082	350	11455
4	4	1/1/2015	11/16/2001	SO45081	338	26782
5	5	1/2/2015	12/15/2001	SO45083	312	14947
6	6	1/2/2015	10/12/2001	SO45084	310	29143
7	7	1/2/2015	12/18/2001	SO45086	314	18747
8	8	1/2/2015	10/9/2001	SO45085	312	18746
9	9	1/3/2015	10/3/2001	SO45093	312	18906
10	10	1/3/2015	9/29/2001	SO45090	310	29170
11	11	1/3/2015	12/11/2001	SO45088	345	11398
12	12	1/3/2015	10/24/2001	SO45092	313	18899
13	13	1/3/2015	12/16/2001	SO45089	351	25977
14	14	1/3/2015	10/26/2001	SO45091	314	18909
15	15	1/3/2015	9/11/2001	SO45087	350	11388
16	16	1/3/2015	9/11/2001	SO45094	310	22785
17	17	1/4/2015	10/30/2001	SO45096	312	12483
18	18	1/4/2015	10/30/2001	SO45097	313	29151

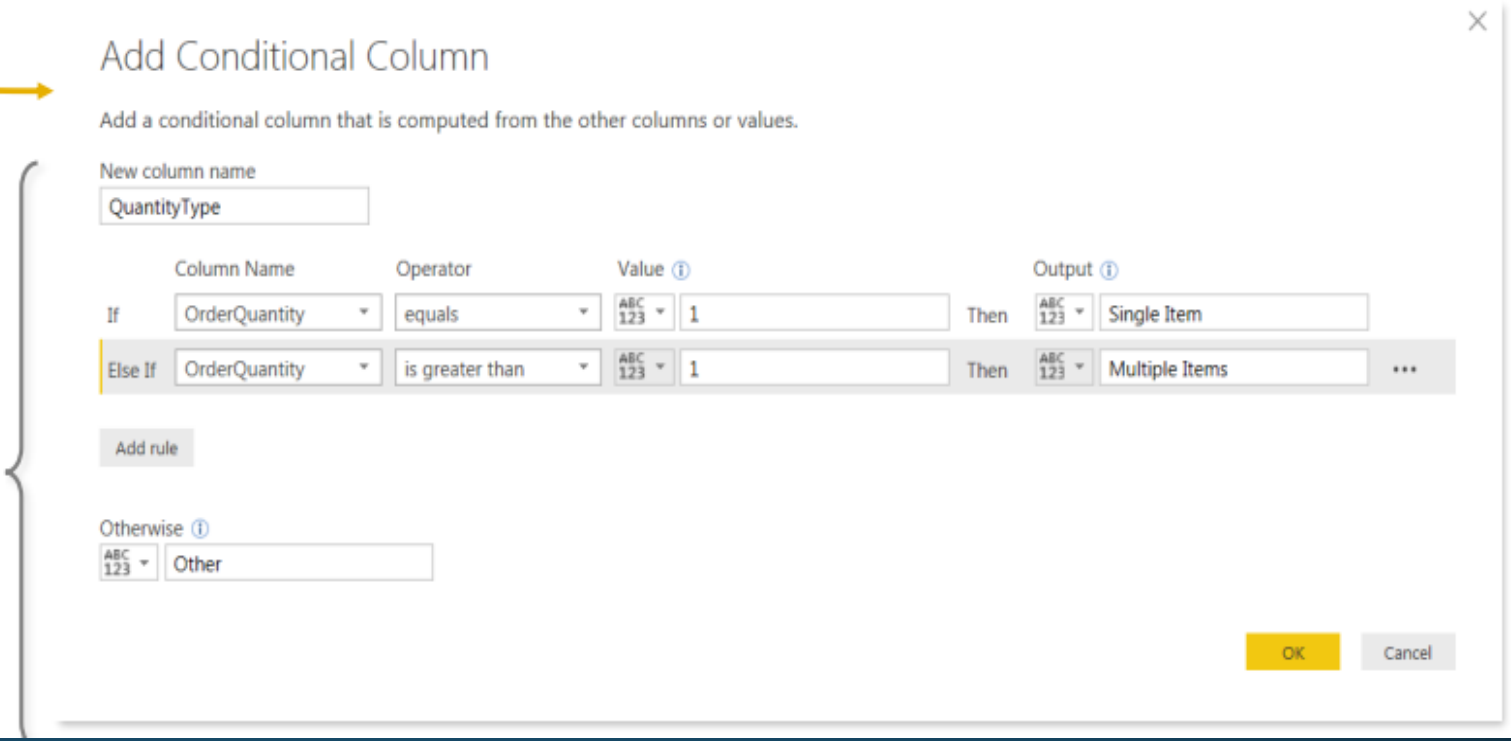
ADDING CONDITIONAL COLUMNS



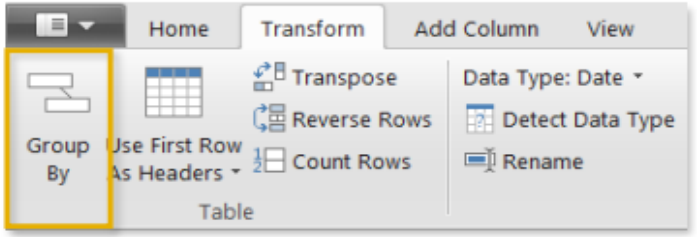
Conditional Columns allow you to define new fields based on logical rules and conditions (*IF/THEN statements*)

*In this case we're creating a new conditional column called "**QuantityType**", which depends on the values in the "**OrderQuantity**" column, as follows:*

- *If OrderQuantity **=1**, QuantityType = "**Single Item**"*
- *If OrderQuantity **>1**, QuantityType = "**Multiple Items**"*
- *Otherwise QuantityType = "**Other**"*

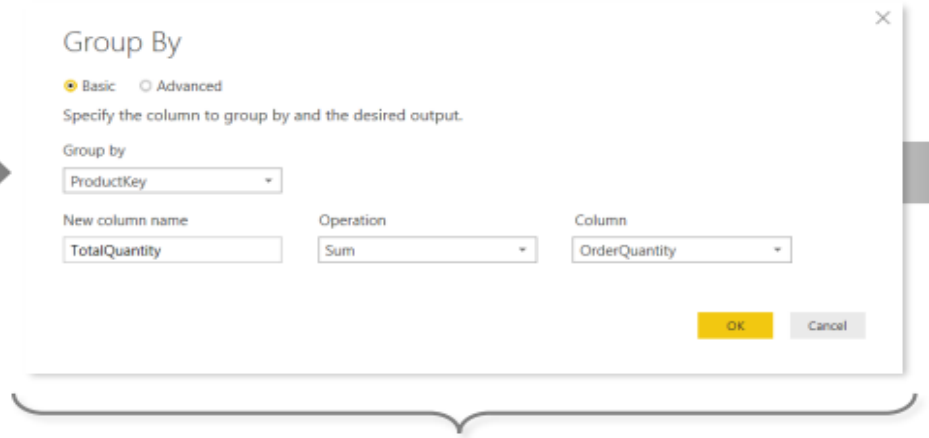


GROUPING & AGGREGATING DATA



Group By allows you to aggregate your data at a different level
(*i.e. transform daily data into monthly, roll up transaction-level data by store, etc*)

	OrderDate	ProductKey	CustomerKey	OrderQuantity
1	6/25/2017	214	14719	1
2	7/16/2016	214	11243	1
3	12/31/2016	214	21452	1
4	6/29/2017	214	22748	1
5	10/6/2016	214	25025	1
6	10/7/2016	214	16504	1
7	10/13/2016	214	13043	1
8	1/19/2017	214	23101	1
9	9/7/2016	214	24900	1
10	1/19/2017	214	24196	1
11	6/29/2017	214	12963	1
12	11/6/2016	214	14570	1
13	11/13/2016	214	16999	1
14	7/31/2016	214	12281	1
15	10/9/2016	214	15685	1
16	8/1/2016	214	16982	1
17	12/4/2016	214	12835	1

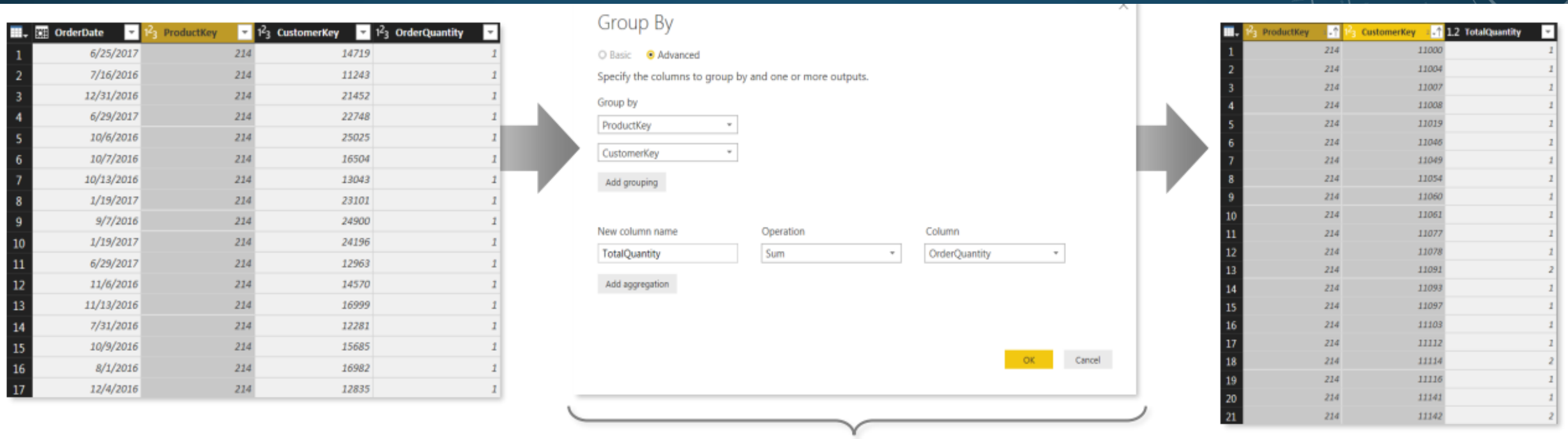


	ProductKey	TotalQuantity
1	214	2099
2	217	1940
3	222	1995
4	225	4151
5	228	392
6	231	408
7	234	424
8	237	381
9	310	169
10	311	139
11	312	179
12	313	168
13	314	157
14	320	10
15	321	55
16	322	5
17	323	34

*In this case we're transforming a daily, transaction-level table into a summary of "**TotalQuantity**" rolled up by "**ProductKey**"*

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

GROUPING & AGGREGATING DATA (CONT)



	OrderDate	ProductKey	CustomerKey	OrderQuantity
1	6/25/2017	214	14719	1
2	7/16/2016	214	11243	1
3	12/31/2016	214	21452	1
4	6/29/2017	214	22748	1
5	10/6/2016	214	25025	1
6	10/7/2016	214	16504	1
7	10/13/2016	214	13043	1
8	1/19/2017	214	23101	1
9	9/7/2016	214	24900	1
10	1/19/2017	214	24196	1
11	6/29/2017	214	12963	1
12	11/6/2016	214	14570	1
13	11/13/2016	214	16999	1
14	7/31/2016	214	12281	1
15	10/9/2016	214	15685	1
16	8/1/2016	214	16982	1
17	12/4/2016	214	12835	1

	ProductKey	CustomerKey	TotalQuantity
1	214	11000	1
2	214	11004	1
3	214	11007	1
4	214	11008	1
5	214	11019	1
6	214	11046	1
7	214	11049	1
8	214	11054	1
9	214	11060	1
10	214	11061	1
11	214	11077	1
12	214	11078	1
13	214	11091	2
14	214	11093	1
15	214	11097	1
16	214	11103	1
17	214	11112	1
18	214	11114	2
19	214	11116	1
20	214	11141	1
21	214	11142	2

*This time we're transforming the daily, transaction-level table into a summary of "**TotalQuantity**" aggregated by both "**ProductKey**" and "**CustomerKey**" (using the advanced option in the dialog box)*

***NOTE**: This is similar to creating a PivotTable in Excel and pulling in "**Sum of OrderQuantity**" with **ProductKey** and **CustomerKey** as row labels*

PIVOTING & UNPIVOTING

"Pivoting" is a fancy way to describe the process of turning **distinct row values** into **columns** (*"pivoting"*) or turning **columns** into **rows** (*"unpivoting"*)

	1.2 1994	1.2 1995	1.2 1996	1.2 1997	1.2 1998
1	286322	253787	155483	246491	130602

PIVOT

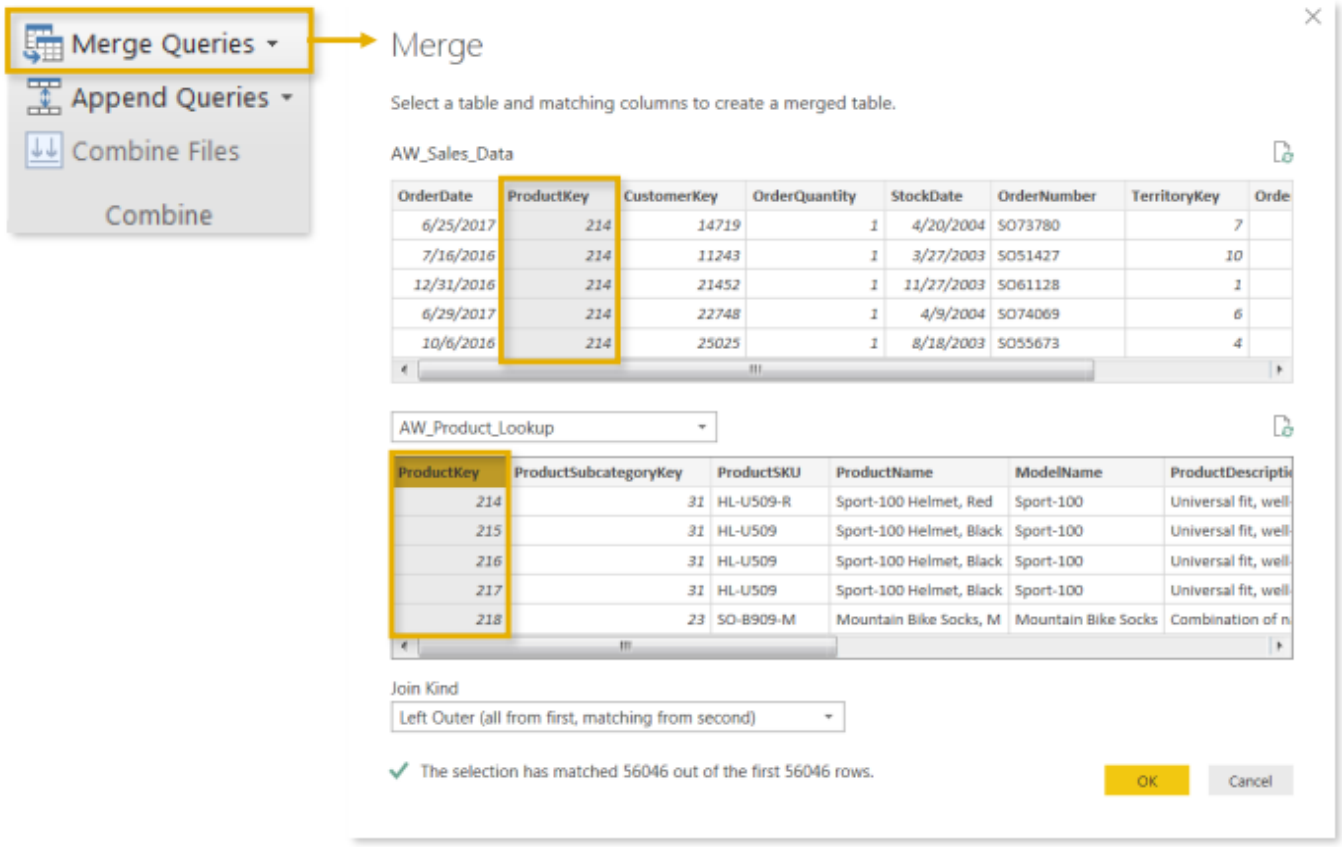
UNPIVOT

	Year	Unit Sales
1	1994	286322
2	1995	253787
3	1996	155483
4	1997	246491
5	1998	130602

*Imagine that the table is on a hinge; pivoting is like rotating it from a **vertical** to a **horizontal** layout, and unpivoting is like rotating 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*

MERGING QUERIES



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

In this case we're merging the **AW_Sales_Data** table with the **AW_Product_Lookup** table, which share a common "*ProductKey*" column

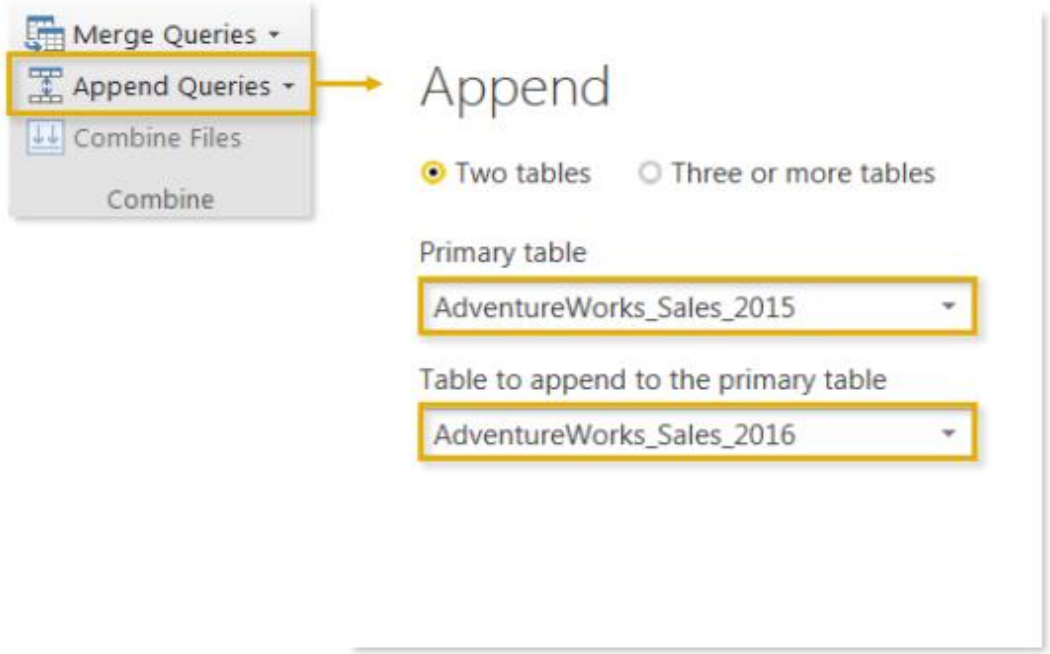
NOTE: Merging ***adds columns*** to an existing table

HEY THIS IS IMPORTANT!

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

In general, it's better to keep tables separate and define **relationships** between them (*more on that later!*)

APPENDING QUERIES



Appending queries allows you to **combine** (or ***stack***) tables that share the exact same column structure and data types

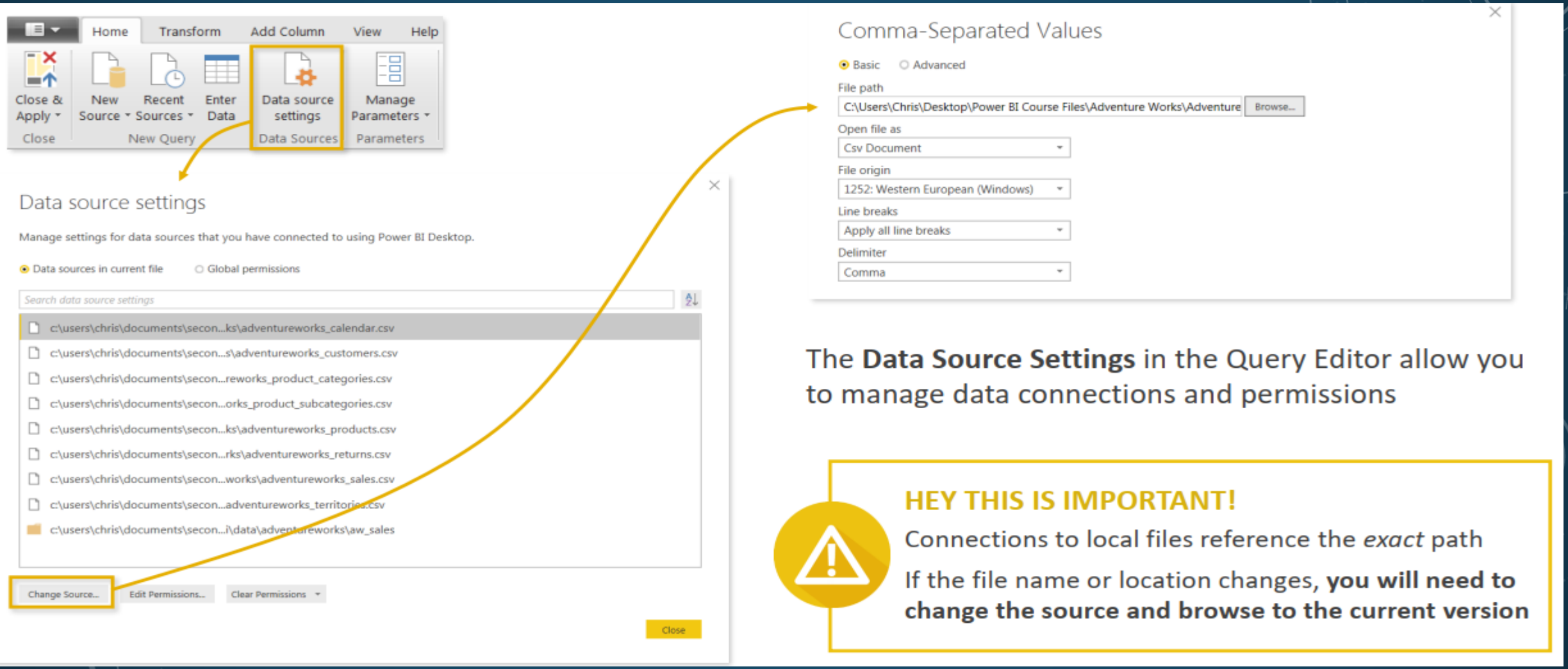
In this case we're appending the **AdventureWorks_Sales_2015** table to the **AdventureWorks_Sales_2016** table, which is valid since they share identical table structures

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

PRO TIP:

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

DATA SOURCE SETTINGS



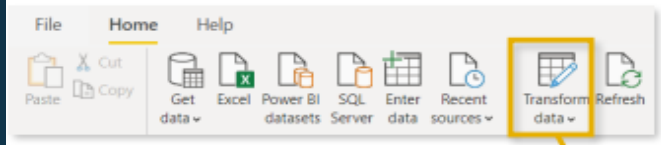
The **Data Source Settings** in the Query Editor allow you to manage data connections and permissions

HEY THIS IS IMPORTANT!

Connections to local files reference the *exact* path

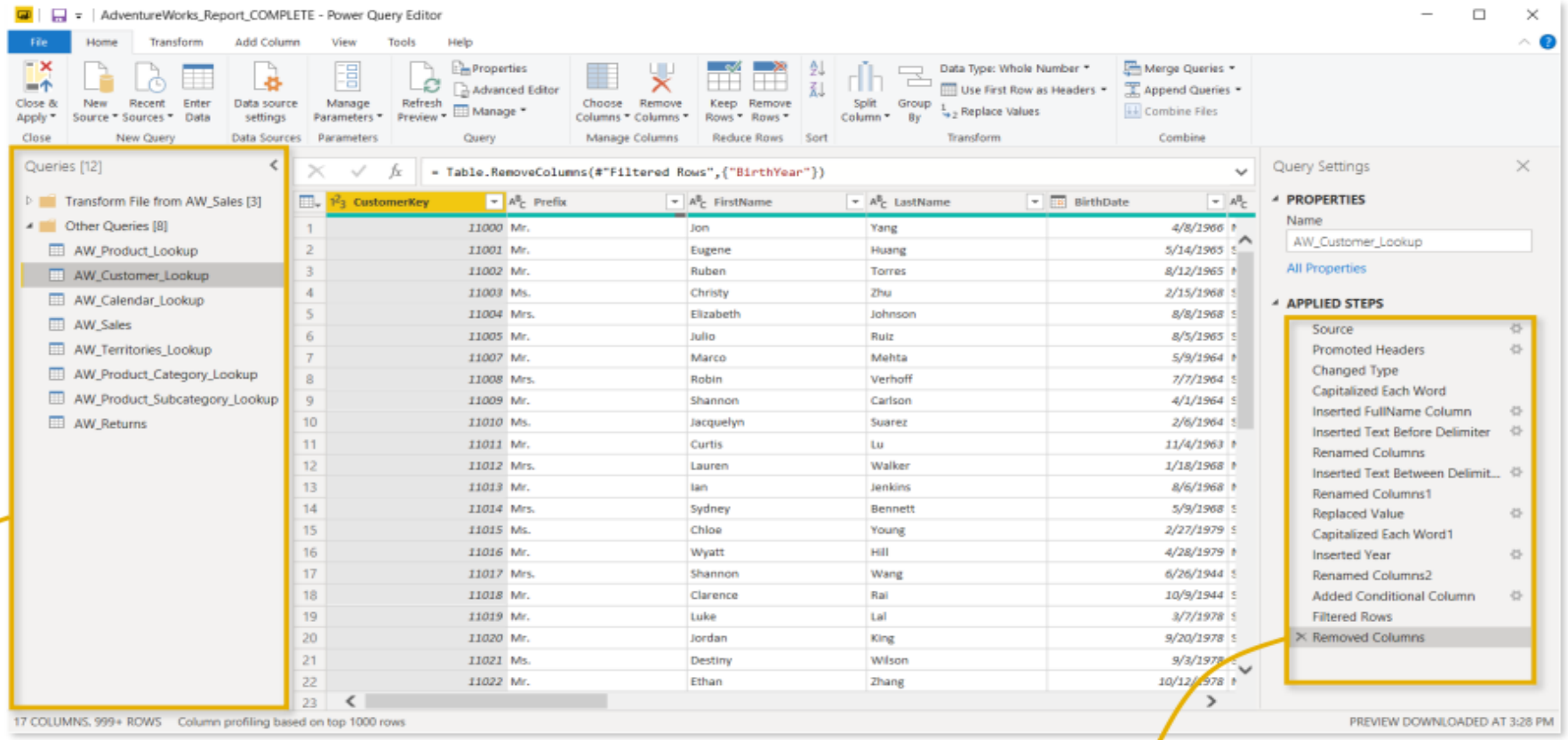
If the file name or location changes, **you will need to change the source and browse to the current version**

MODIFYING QUERIES



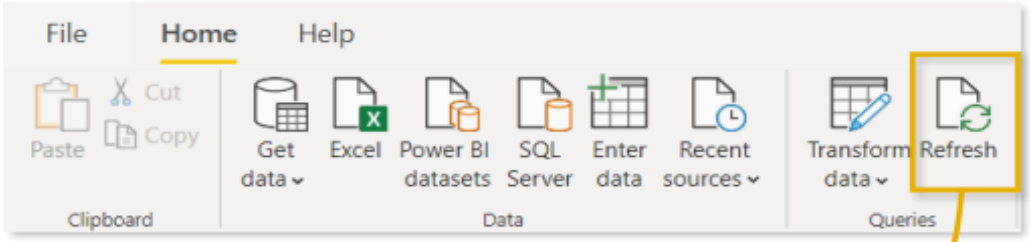
Select **Transform Data*** from the **Home** tab to launch the Query Editor

Within the editor, view or modify existing queries in the "**Queries**" pane



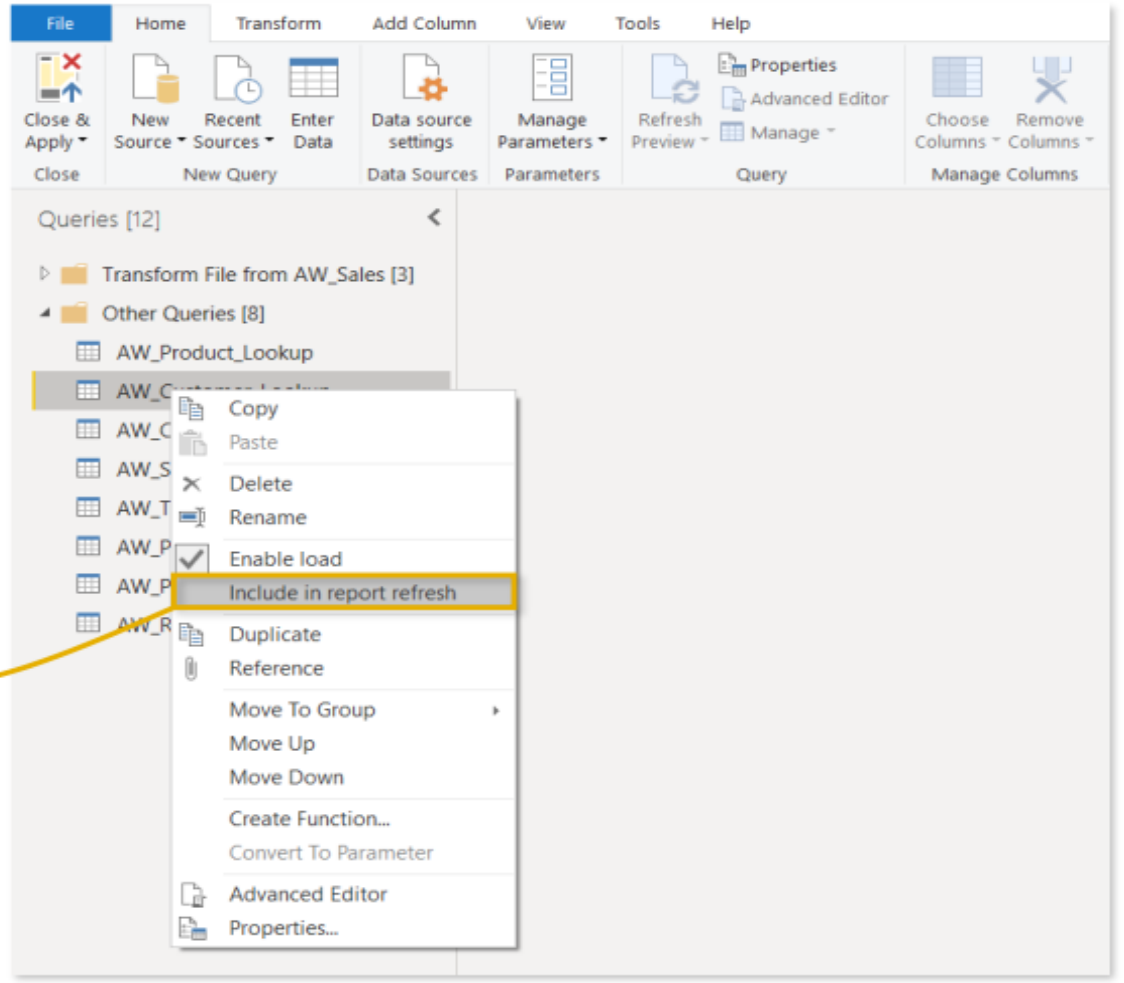
Within each query, you can click each item within the "**Applied Steps**" pane to view each stage of the transformation, add new steps or delete existing ones, or modify individual steps by clicking the gear icons

REFESHING QUERIES



By default, **ALL** queries in the model will refresh when you use the "*Refresh*" command from the **Home** tab

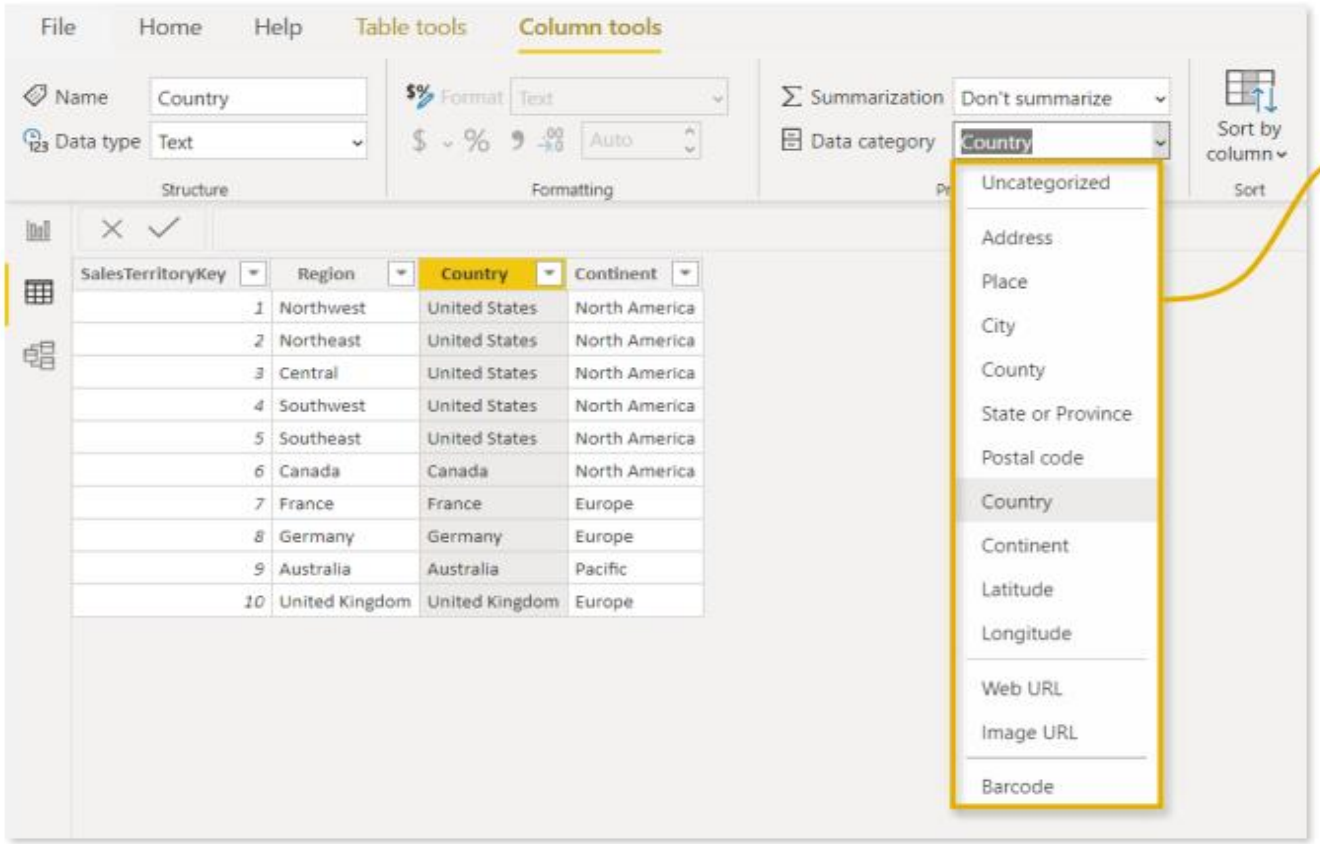
From the Query Editor, uncheck "***Include in report refresh***" to exclude individual queries from the refresh



PRO TIP:

Exclude queries that don't change often, like lookups or static data tables

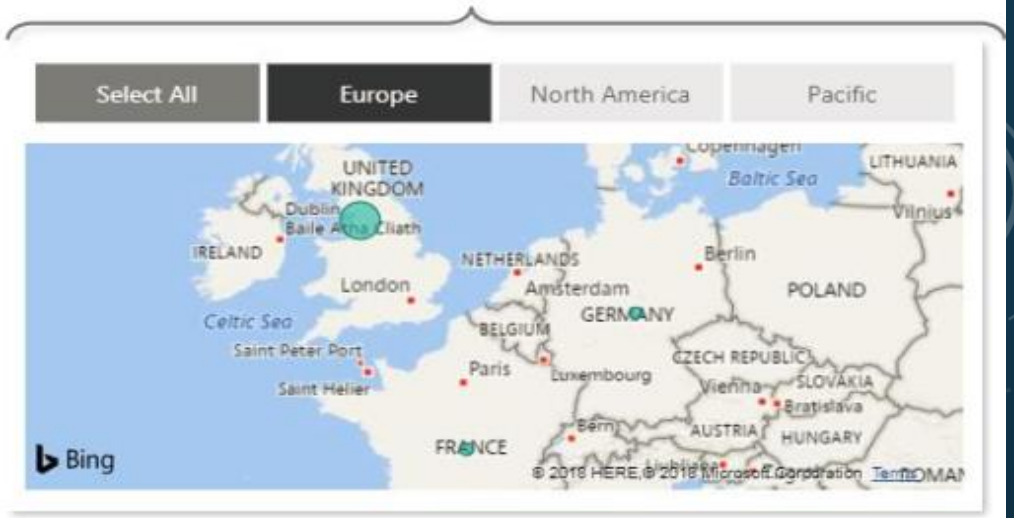
DEFINING DATA CATEGORIES



SalesTerritoryKey	Region	Country	Continent
1	Northwest	United States	North America
2	Northeast	United States	North America
3	Central	United States	North America
4	Southwest	United States	North America
5	Southeast	United States	North America
6	Canada	Canada	North America
7	France	France	Europe
8	Germany	Germany	Europe
9	Australia	Australia	Pacific
10	United Kingdom	United Kingdom	Europe

Select a column in the **Data** view to access **Column Tools**, where you can edit field properties to define specific categories*

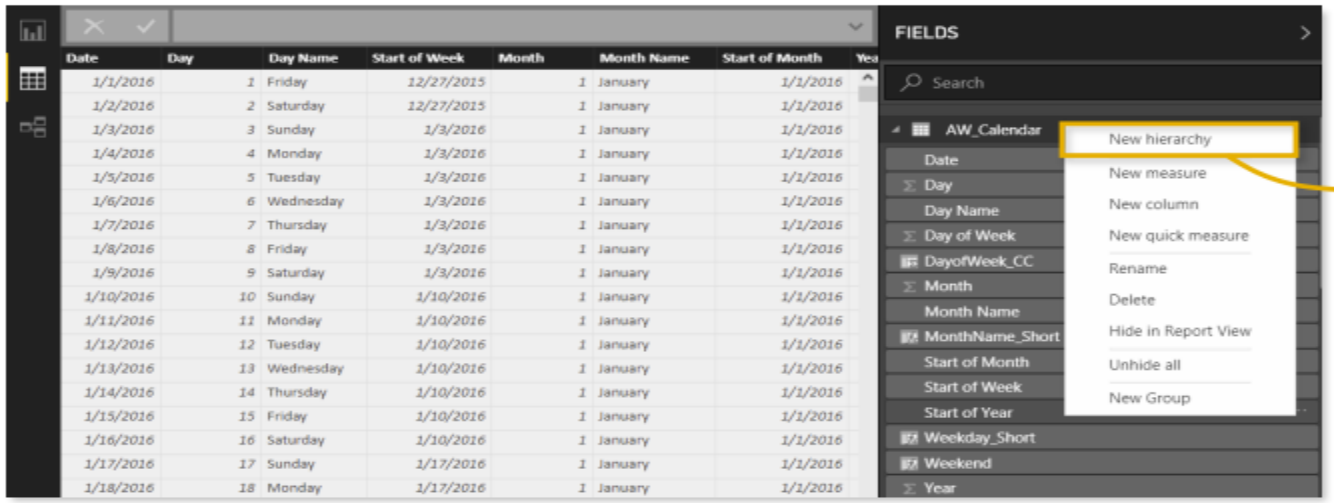
*This is commonly used to help Power BI accurately map location-based fields like **addresses**, **countries**, **cities**, **latitude/longitude coordinates**, **zip codes**, etc.*



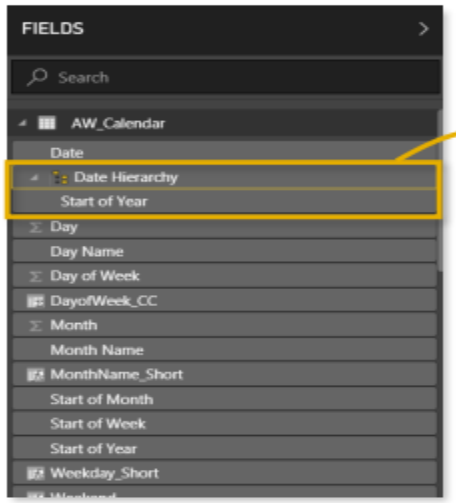
DEFINING HIERARCHIES

Hierarchies are groups of nested columns that reflect multiple levels of granularity

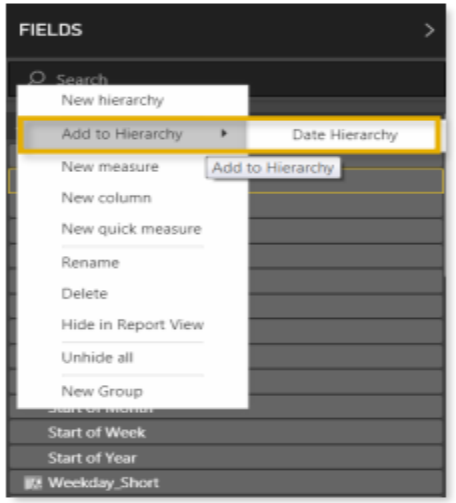
- For example, a "**Geography**" hierarchy might include **Country**, **State**, and **City** columns
- Each hierarchy can be treated as a **single item** in tables and reports, allowing users to "drill up" and "drill down" through different levels of the hierarchy in a meaningful way



1) From within the **Data** view, right-click a field (or click the ellipsis) and select "**New hierarchy**" (here we've selected "*Start of Year*")



2) This creates a hierarchy field containing "*Start of Year*", which we've renamed "**Date Hierarchy**"



3) Right-click other fields (like "*Start of Month*") and select "**Add to Hierarchy**"

THANK YOU

Q&A