DATA SOLUTIONS PLATFORM

DATA ANALYTICS

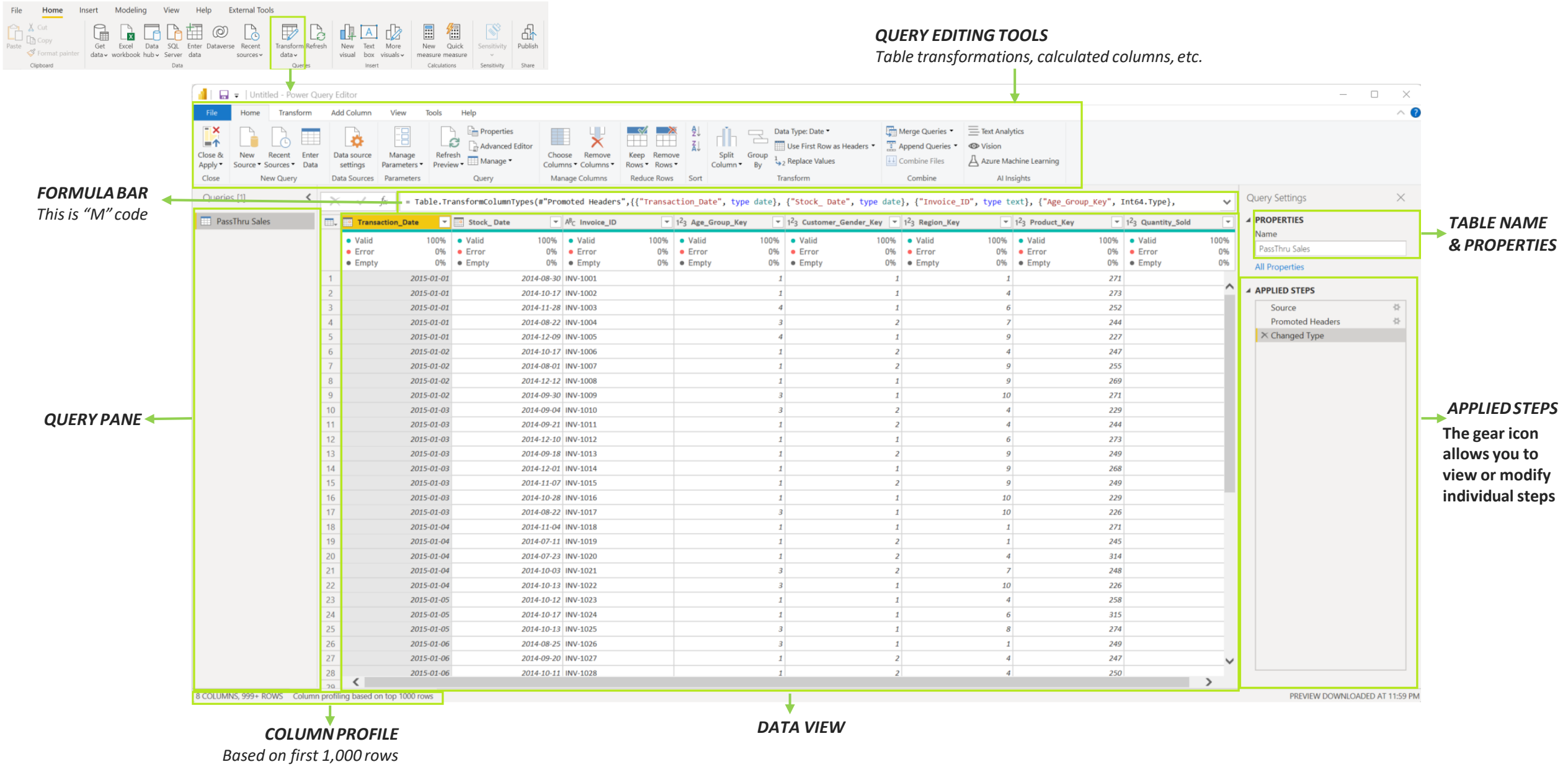
POWER BI DATA ANALYSIS

With Power BI Expert…. ***Benjamin Taiwo***

TODAY'S CONTENT

- The Query Editor
- Meet the Query Editing Tools
- The Home Tab / General Settings
- Basic Table Transformations
- Transform Tab / Table Tools
- Column Tools
- Text Column Tools
- Number Column Tools
- Date & Time Column Tools
- Add Column Tab / General Tools
- The View Tab
- The View Tab / Profiling Tools
- The Help Tab
- M Code / Advanced Editor
- M Code / Syntax
- M Code / Common Functions
- Q & A

THE QUERY EDITOR



QUERY EDITING TOOLS
Table transformations, calculated columns, etc.

FORMULA BAR
This is "M" code

QUERY PANE

TABLE NAME & PROPERTIES

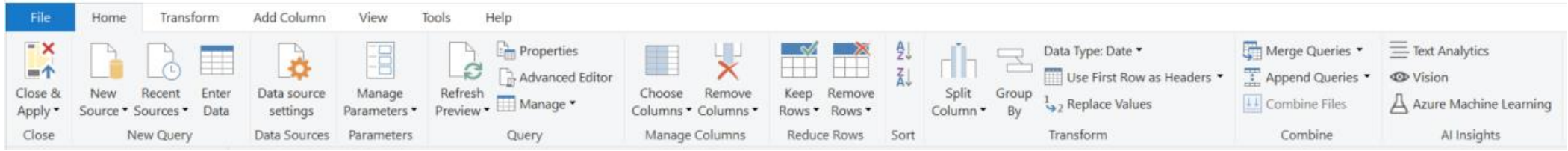
APPLIED STEPS
The gear icon allows you to view or modify individual steps

COLUMN PROFILE
Based on first 1,000 rows

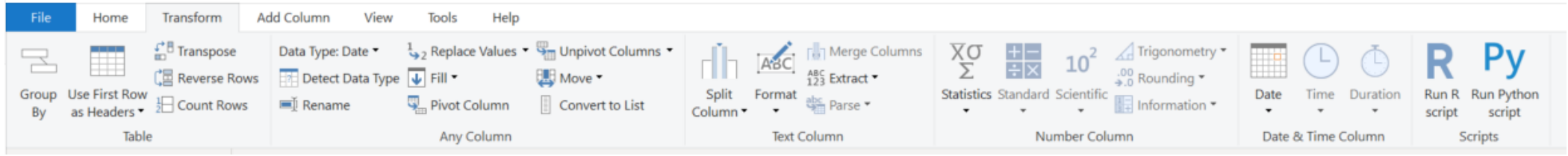
DATA VIEW

MEET THE QUERY EDITING TOOLS

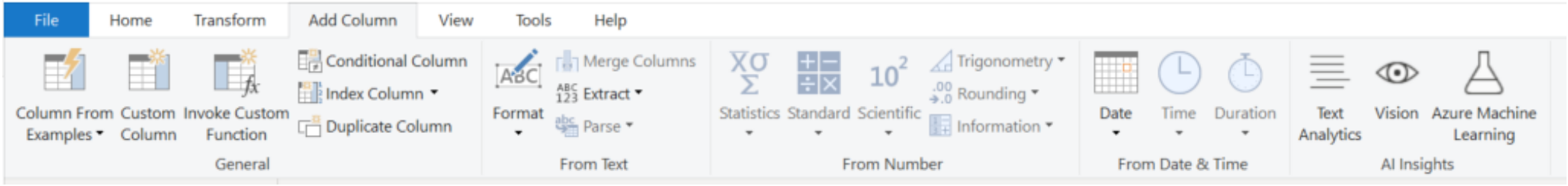
The HOME tab includes general settings and common table transformation tools



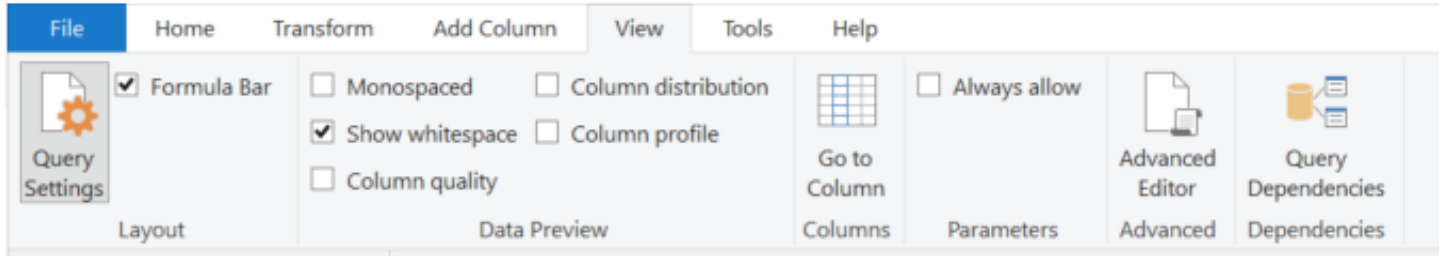
The TRANSFORM tab includes tools to modify existing columns (splitting/grouping, transposing, extracting text, etc)



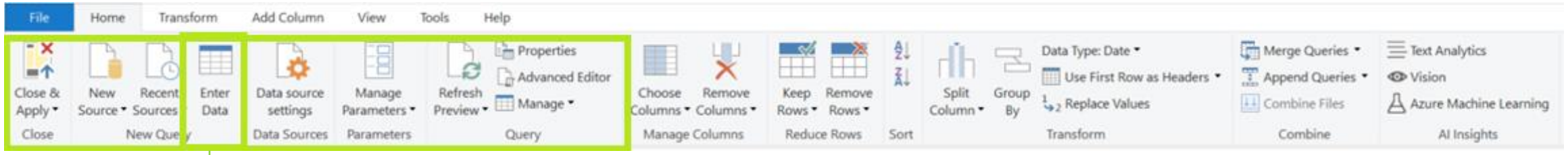
The ADD COLUMN tools create new columns (based on conditional rules, text operations, calculations, dates, etc)



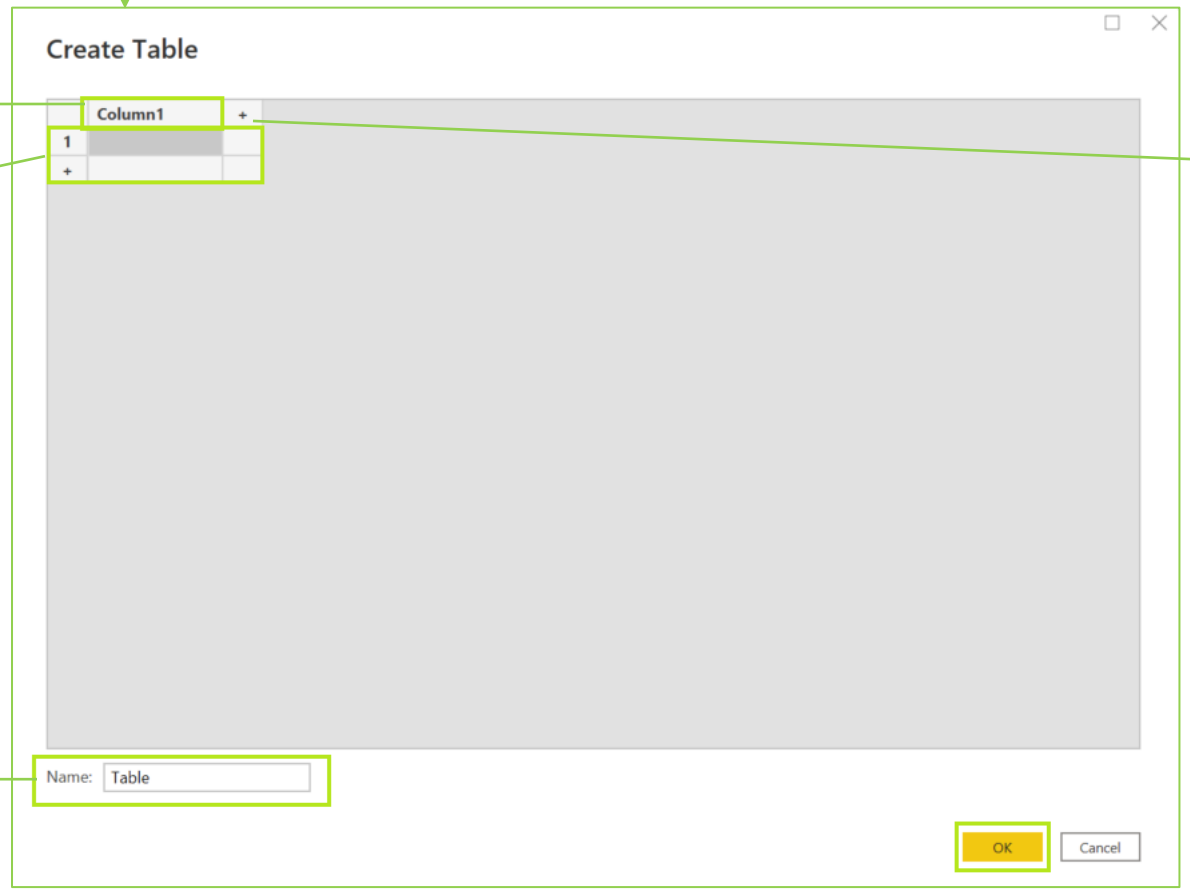
The VIEW tab includes data preview and other advanced editing tools



THE HOME TAB / GENERAL SETTINGS



Create New Table



Rename Column Name

Add Rows and Values

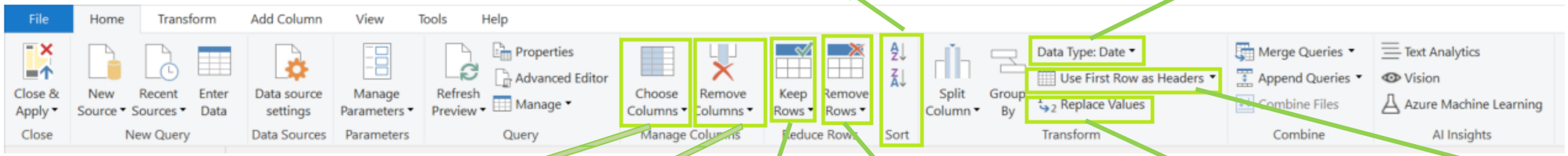
Add Columns

Rename Table Name

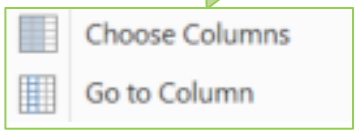
BASIC TABLE TRANSFORMATIONS

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

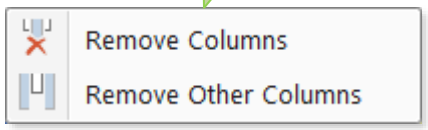
Change data type (*date, $, %, text, etc.*)



Promote header row

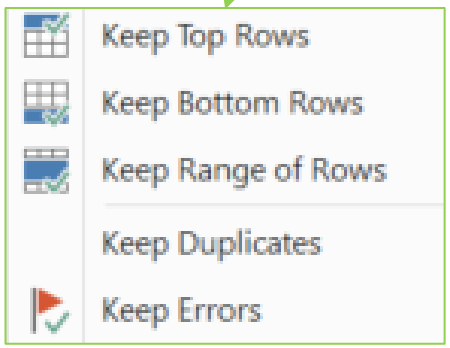


Choose or Go to columns

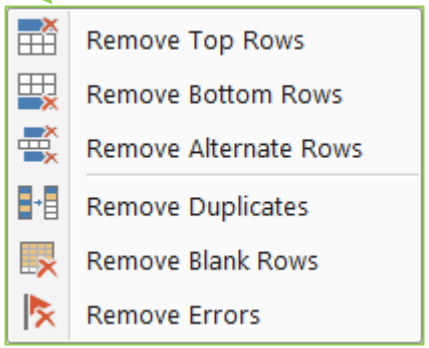


Remove columns

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



Keep rows or duplicates

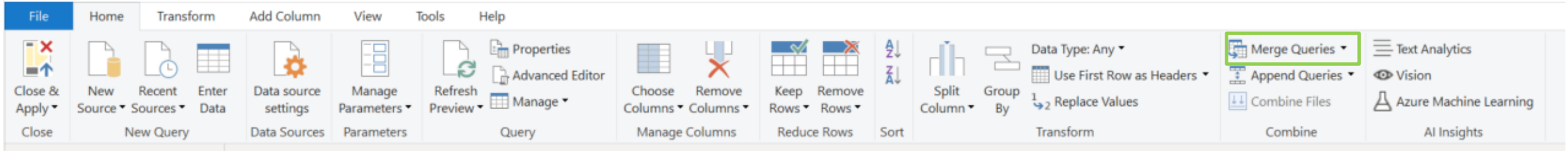


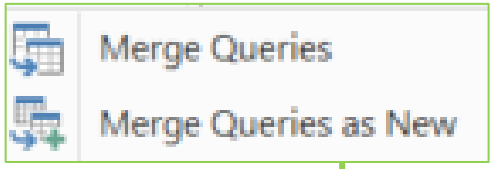
Remove rows or duplicates

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

Replace values in the selected column

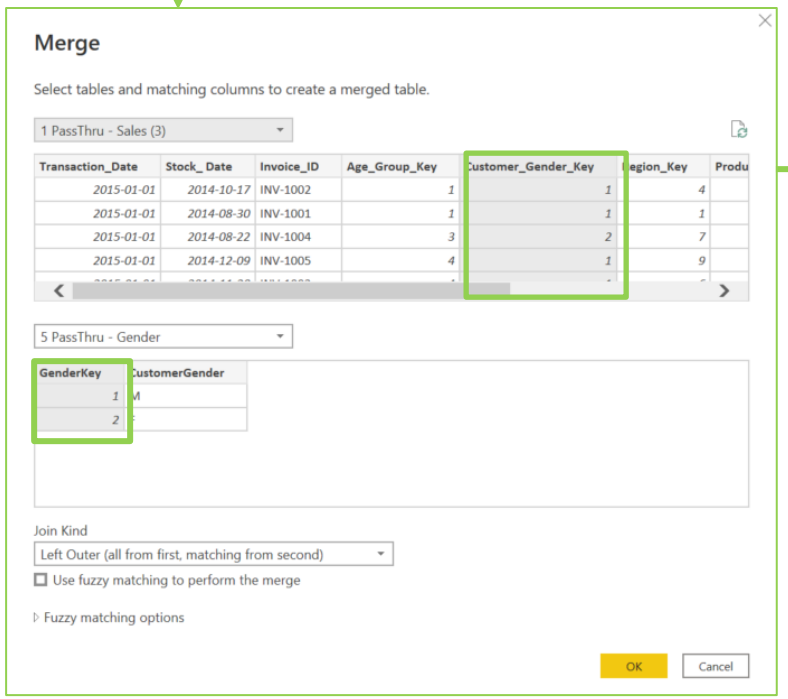
MERGE QUERIES

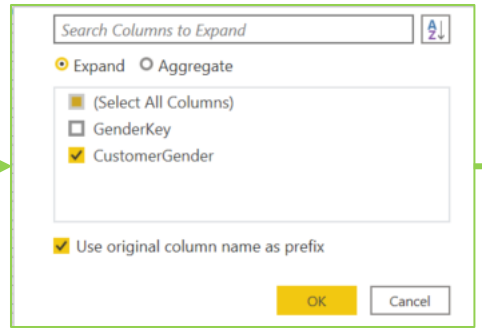




Merging queries allows you to add columns to an existing table by joining tables based on a common column (Primary/Secondary Key).

Merge Queries as New create a new table with the merged tables.

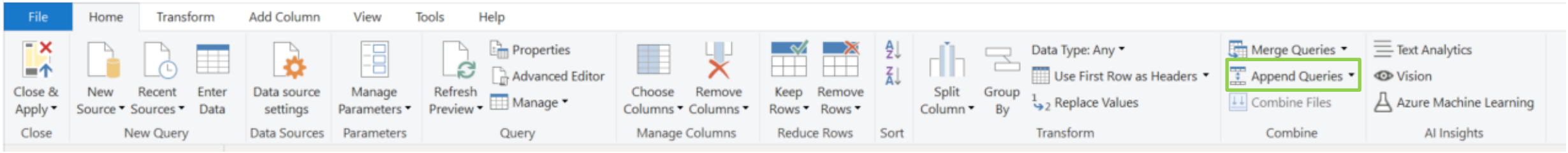


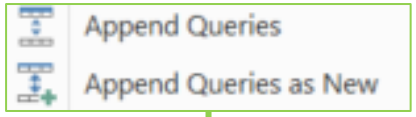


Note:

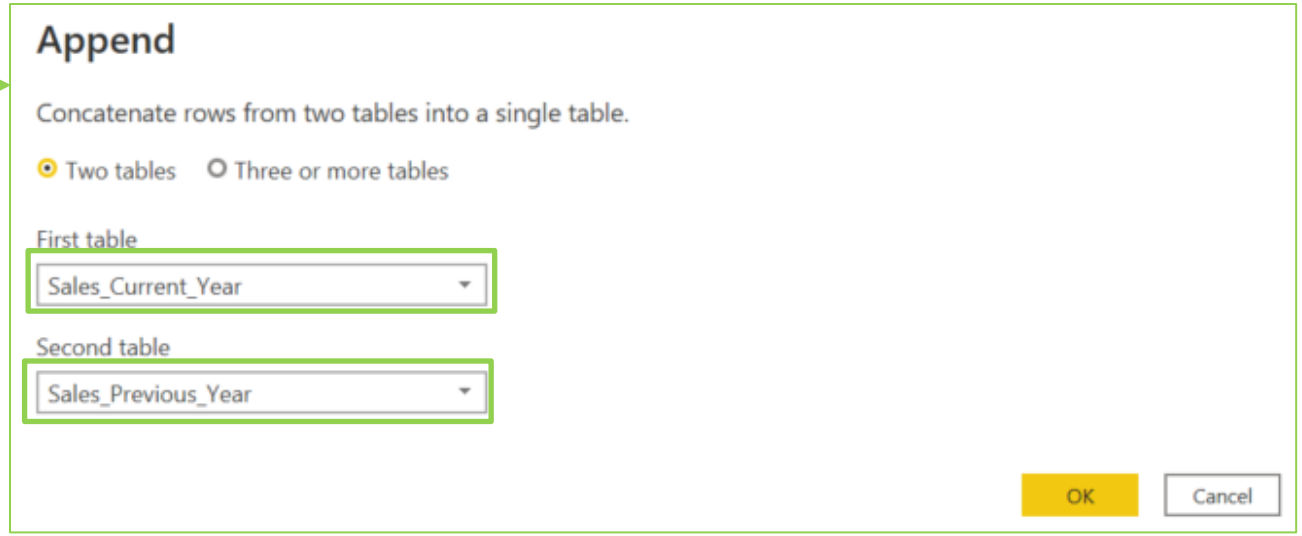
In general, it's better to keep tables separate and define relationships between them in your data model

APPEND QUERIES





Append Queries - allows you to add **rows** to an existing table that share the exact same column structure and data types
Append Queries as New creates new table with the appended tables.

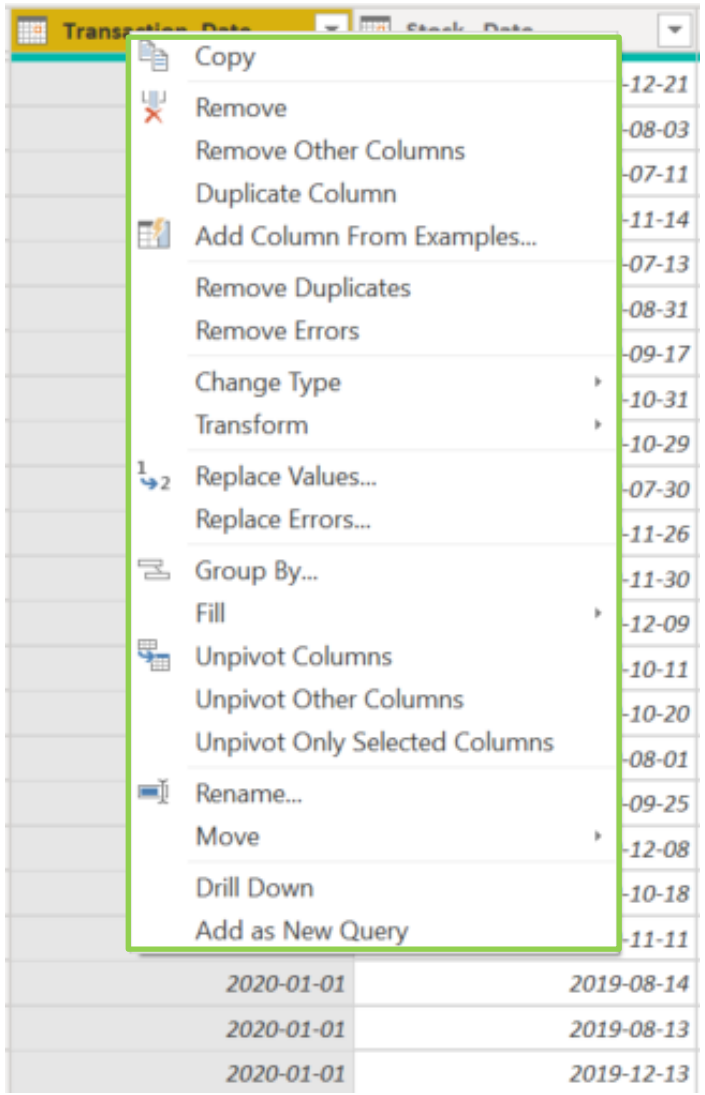


Note:

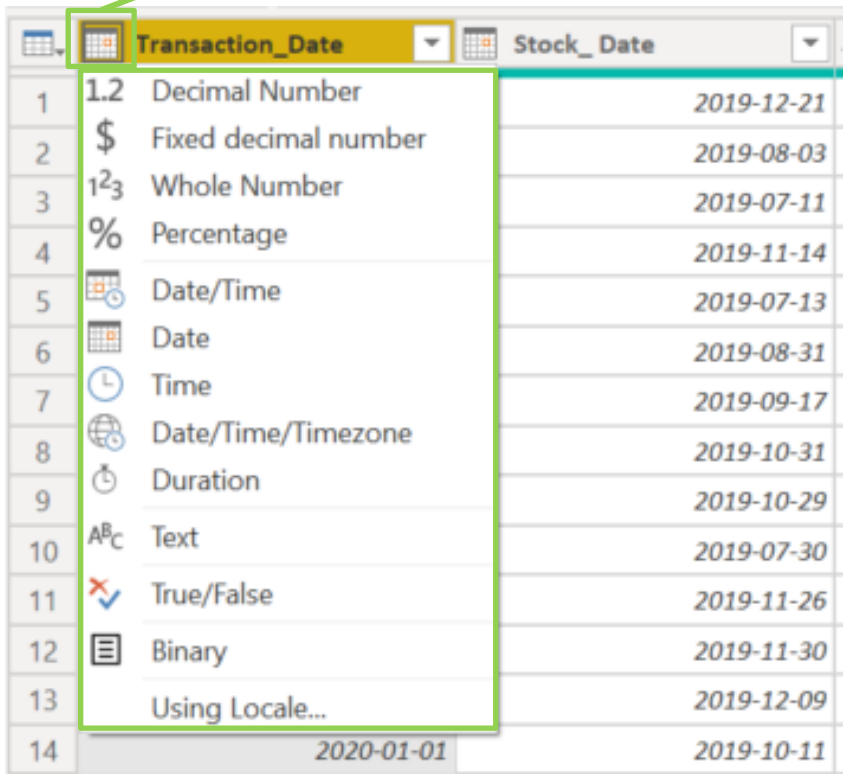
Use the "Folder" option (Get Data > More > Folder) to append all files within a folder (assuming they share the same structure); as you add new files, simply refresh the query and they will automatically append!

THE DATA VIEW TAB

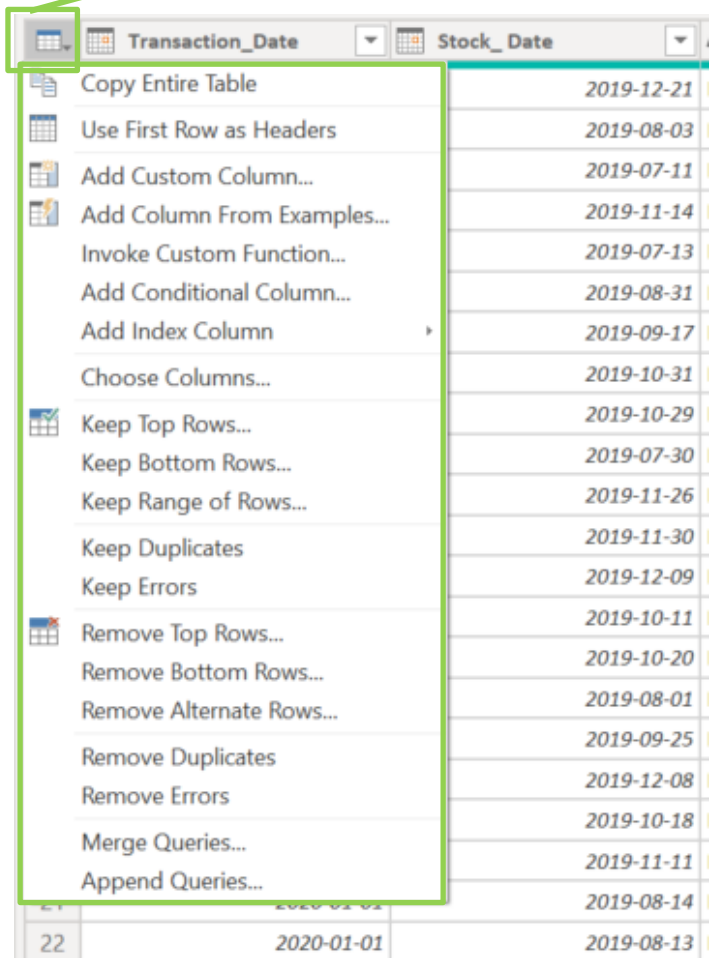
Right-click the column header



Click the column header icon



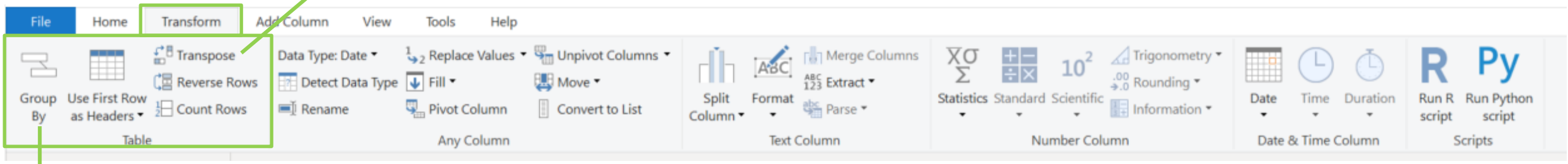
Click the table icon



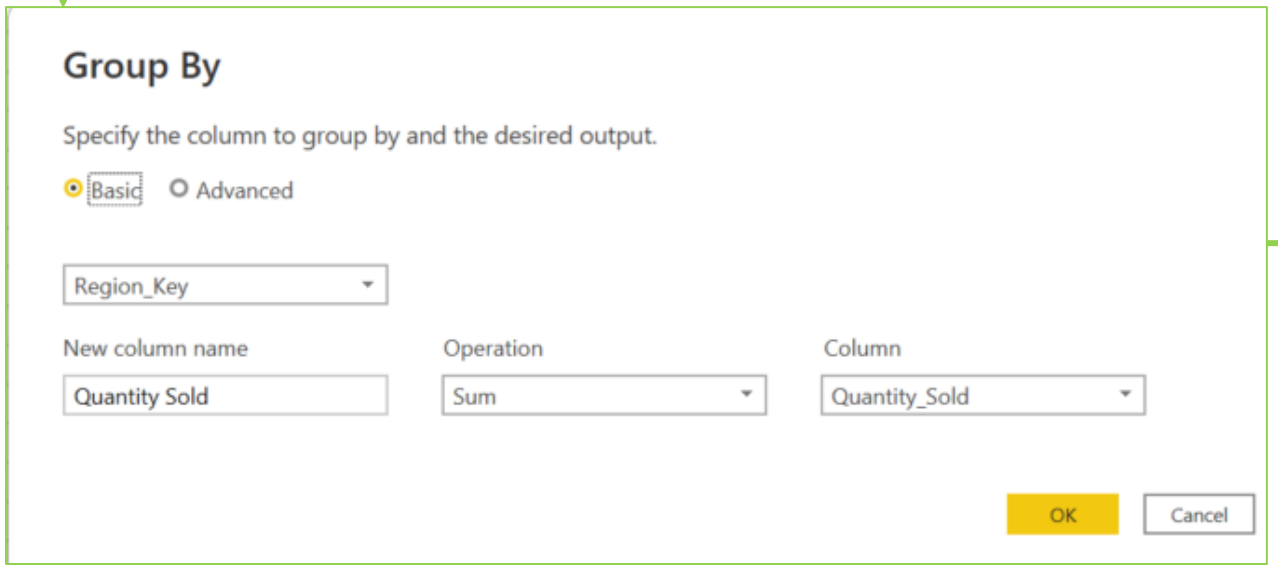
Follow the clicks to access common table and column tools

TRANSFORM TAB / TABLE TOOLS

Treat rows as columns and vice versa



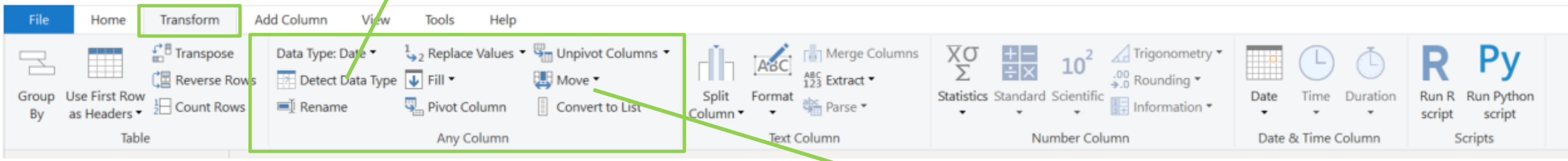
Group by allows you to aggregate your data at a different level. For example: Roll up transaction-level data by Region



	Region_Key	Quantity Sold
1	1	322867
2	4	428061
3	6	300002
4	7	204445
5	9	418375
6	10	249906
7	8	199134
8	5	992
9	2	446
10	3	1850

TRANSFORM TAB / COLUMN TOOLS

Automatically detect data type of selected column(s)

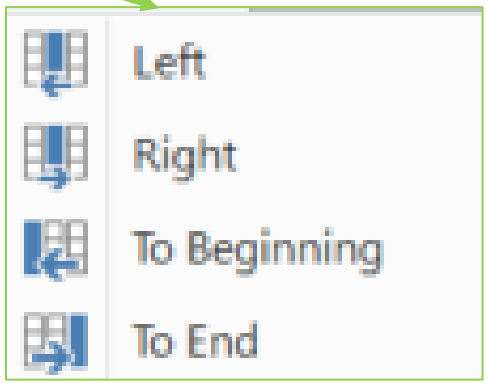


	Month Name	Quantity Sold by Month
1	January	177428
2	February	168089
3	March	178502
4	April	190329
5	May	202393
6	June	206919
7	July	111982
8	August	165052
9	September	164644
10	October	173338
11	November	173677
12	December	213725

Pivoting uses distinct values in the selected column to create new columns while **Unpivot** converts selected columns into attribute-value pairs.

Note:

Transpose works similarly, but differently. It doesn't recognize unique values, instead, the entire table is transformed.

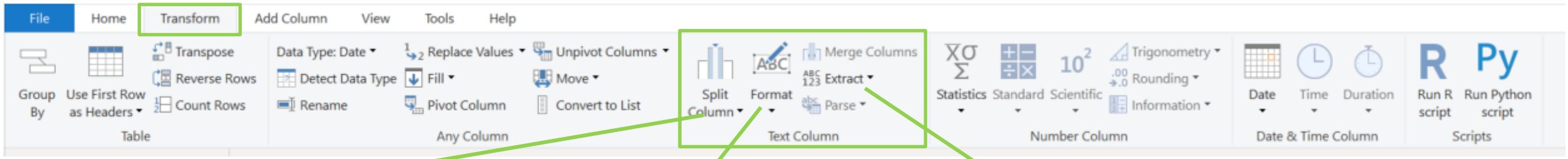


PIVOTING

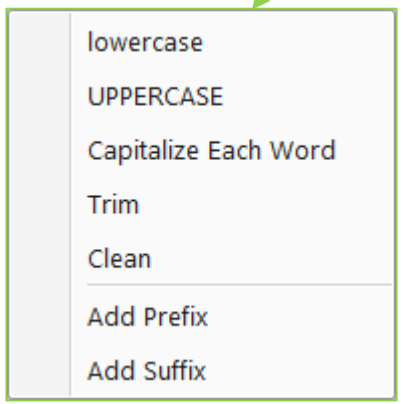
UNPIVOTING

	January	February	March	April	May	June	July	August	September	October	November	December
1	177428	168089	178502	190329	202393	206919	111982	165052	164644	173338	173677	213725

TEXT COLUMN TOOLS

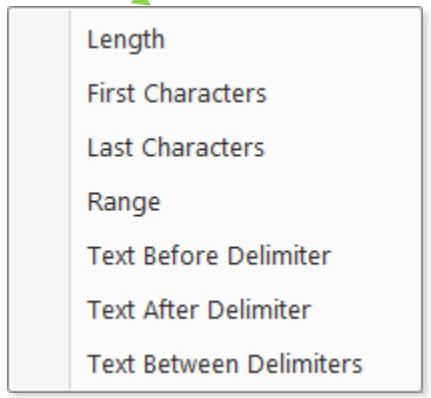


Split a text column based on either a specific delimiter or a number of characters or positions or Upper/Lower cases or digits



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*



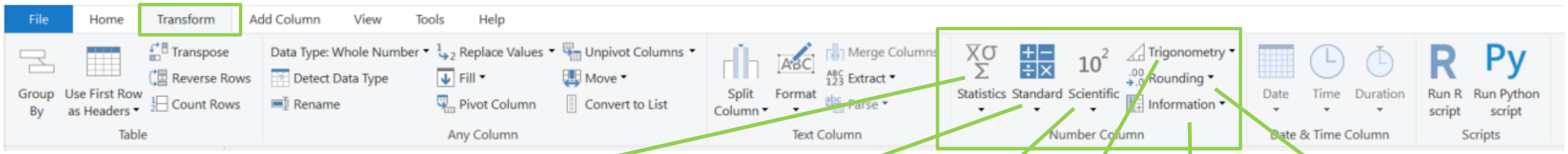
Extract characters from a text column based on fixed lengths, first/last, ranges or delimiters

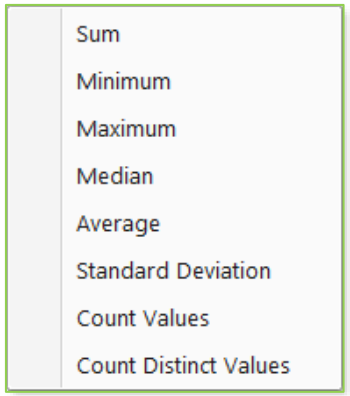
Tip*: Select two or more columns to* ***merge*** *(or* ***concatenate****) fields*

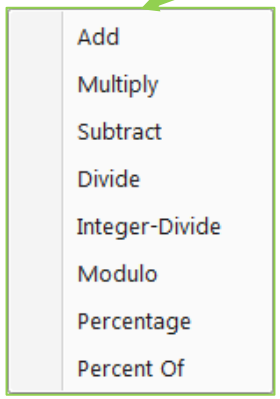
Note:

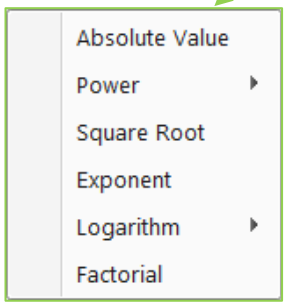
You can access many of these tools in both the "**Transform**" and "**Add Column**" menus -- the difference is whether you want to ***add a new column*** or ***modify an existing one***

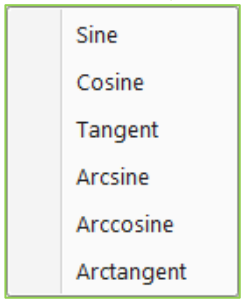
NUMBER COLUMN TOOLS

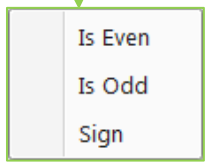












Rounding tools allow you to round up or down values in a selected column

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

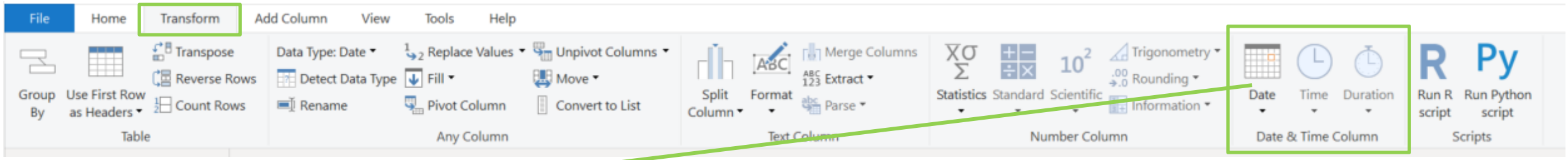
Note: *These tools return a SINGLE value, and are commonly used to explore a table rather than prepare it for loading*

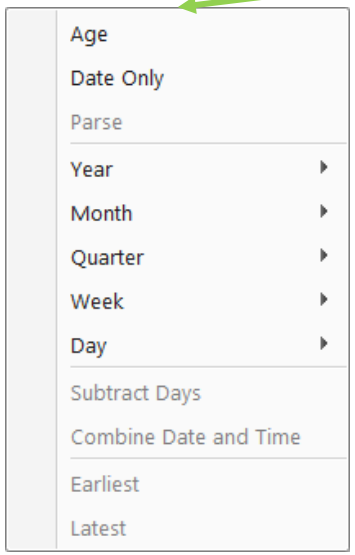
Standard, Scientific and **Trigonometry** tools allow you to apply standard operations (addition, multiplication, division, etc.) or more advanced calculations (power, logarithm, sine, tangent, etc) to each value in a column

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



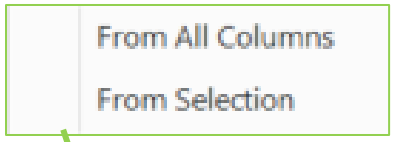


Date & Time tools are relatively straight-forward, and include the following options:

- **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)

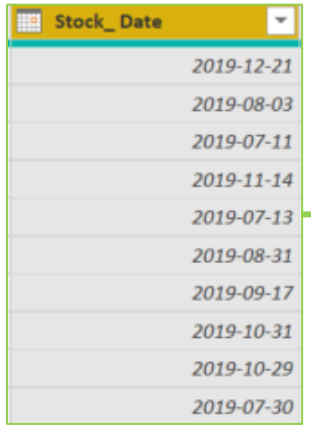
Note: *You will almost always want to perform these operations from the "Add Column" menu to build out new fields, rather than transforming an individual date/time column*

ADD COLUMNS FROM EXAMPLE

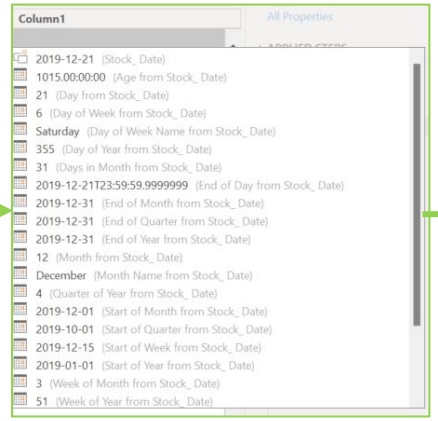


Columns from examples: Use examples to create a new column in the selected table

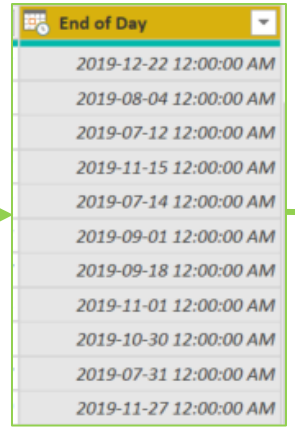
Selected Column in the active Table



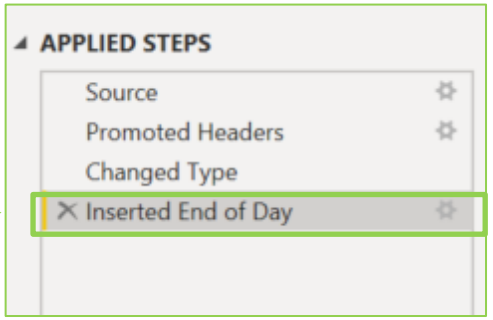
Column to insert the actual example



New column created based on example



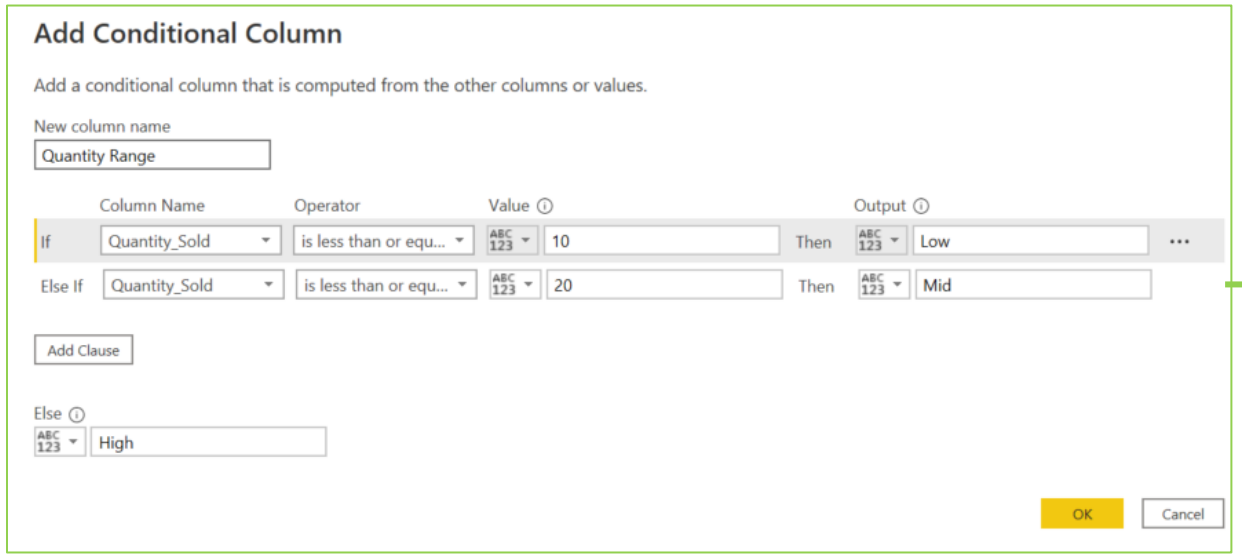
Creates column using text before delimiter transformation

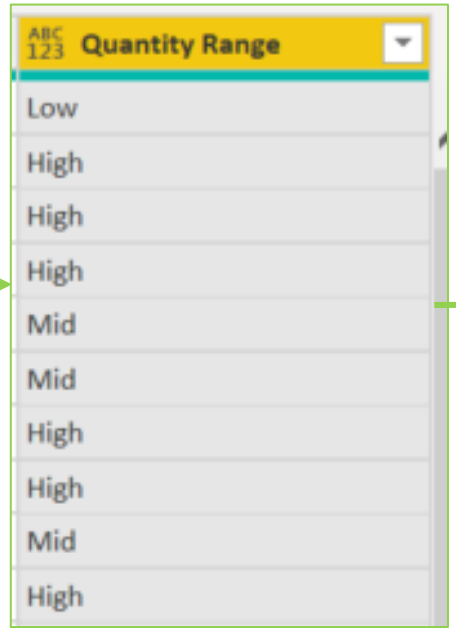


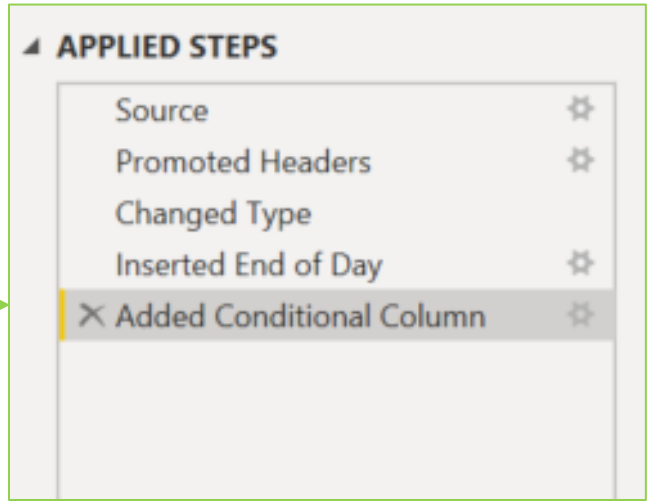
Note: This is very useful when you know the outcome you want but don't know which transformation, or group of transformations, to use

ADD CONDITIONAL COLUMNS

Conditional columns allow you to create new column using logical statements *(IF statements)*







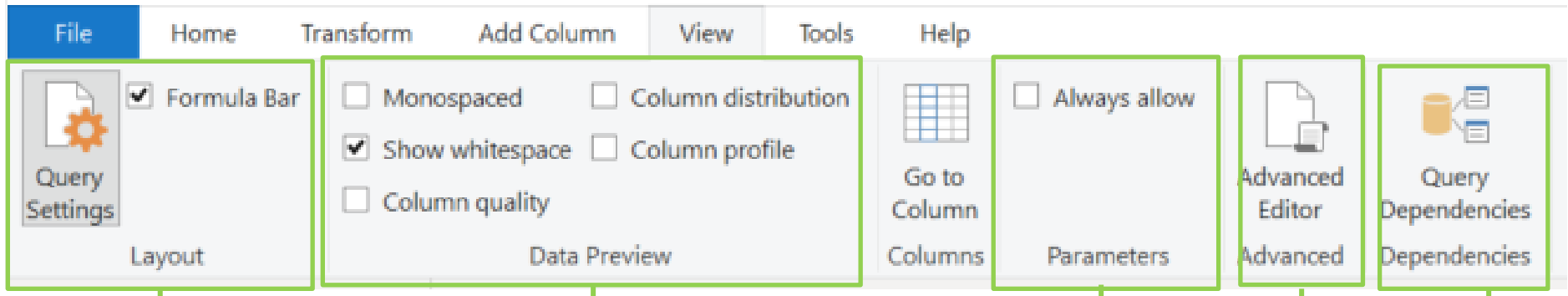
ADD INDEX COLUMNS

Index columns create a new column with a list of sequential values from 0 or 1 or custom number to identify each unique row in a table.

Note: They are often used to create unique IDs and form relationships between tables

	Index	Transaction_Date	Stock_ Date	Invoice_ID
1	0	2020-01-01	2019-12-21	INV-84639
2	1	2020-01-01	2019-08-03	INV-84640
3	2	2020-01-01	2019-07-11	INV-84641
4	3	2020-01-01	2019-11-14	INV-84642
5	4	2020-01-01	2019-07-13	INV-84643
6	5	2020-01-01	2019-08-31	INV-84644
7	6	2020-01-01	2019-09-17	INV-84645
8	7	2020-01-01	2019-10-31	INV-84646
9	8	2020-01-01	2019-10-29	INV-84647
10	9	2020-01-01	2019-07-30	INV-84648
11	10	2020-01-01	2019-11-26	INV-84649

VIEW MENU



Turn off or on the ***Applied Steps*** *menu and* ***M code formula bar***

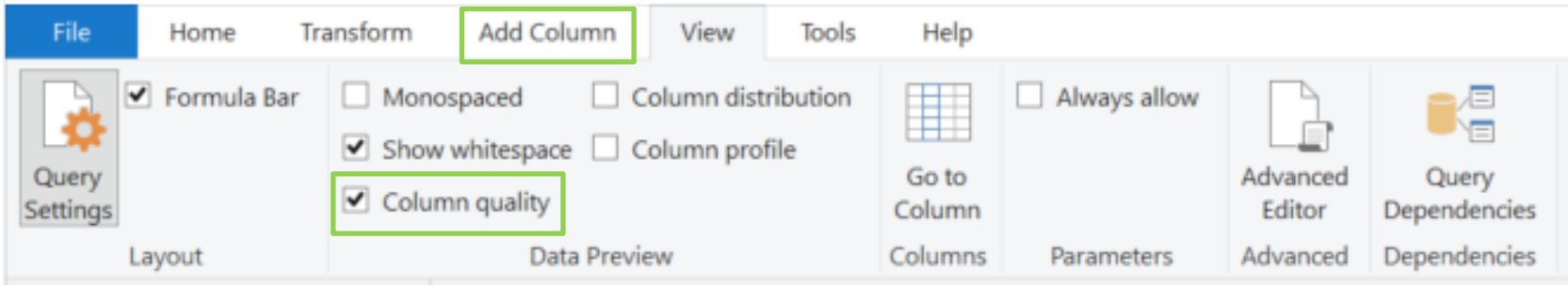
Data profiling tools *like column quality, distribution, and profile*

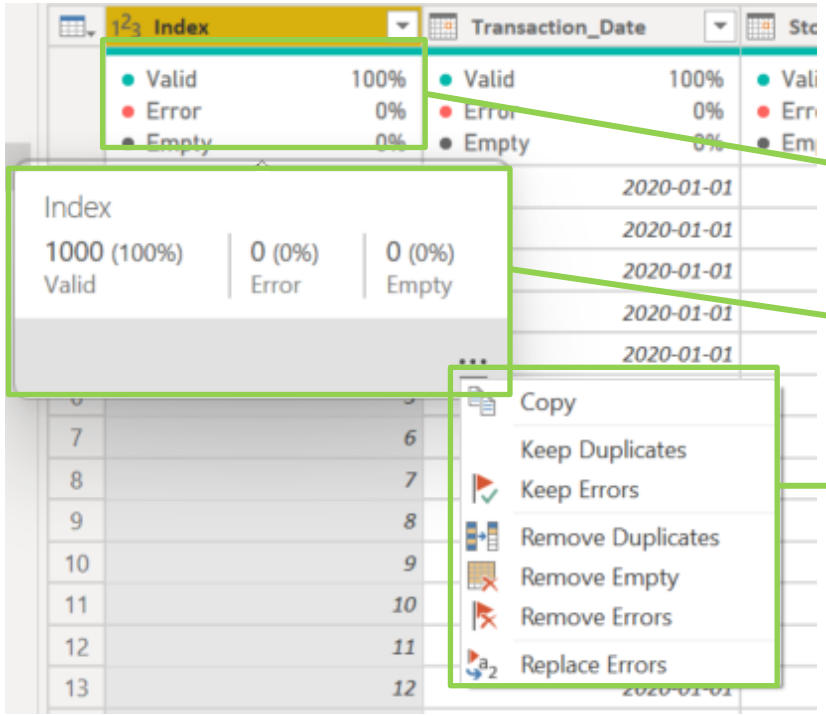
Enable ***parameter creation***

Open the ***M code advanced editor***

View dependency tree of all queries

DATA PROFILING (COLUMN QUALITY)





Data profiling tools such as *column quality*, *column distribution*, and *column profile* provide a visual way to explore data and get a sense of your dataset composition

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

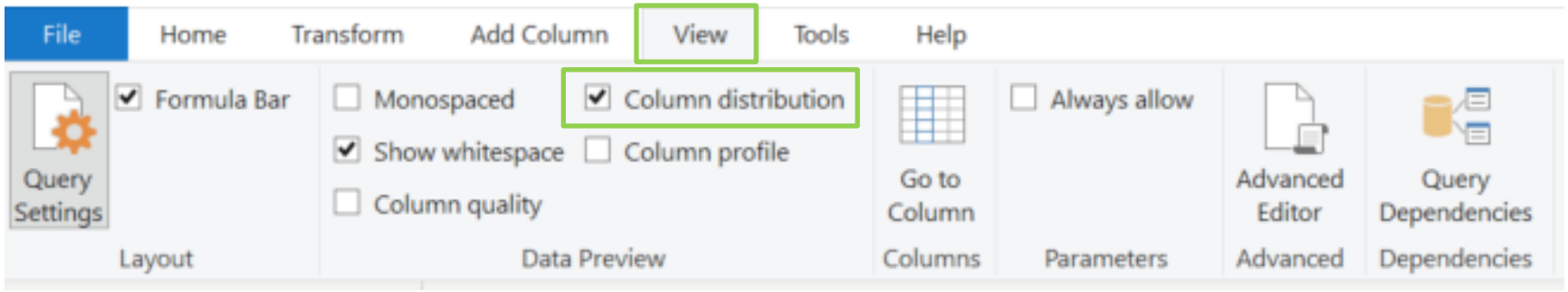
Hover over the column quality box to reveal a ***contextual menu***

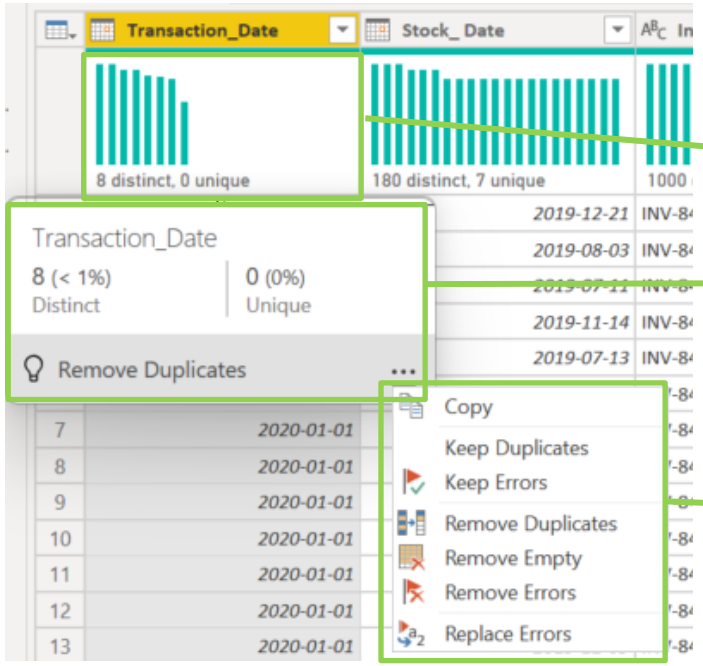
Use the ***options menu*** *to clean duplicates, errors & empty values*

Note:

Use the contextual menu to solve column quality issues instead of having to right-click or find the right Power Query option

DATA PROFILING (COLUMN DISTRIBUTION)





Column distribution Shows column value distribution in the dataset.

Shows the unique (total number of values that appear only once in a column) and distinct (total number of different values regardless how many times it appears in the dataset) values in a query

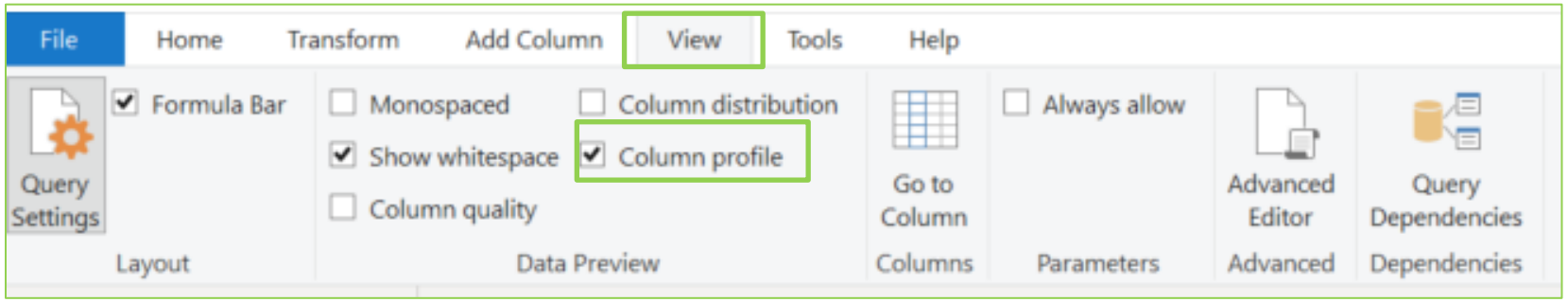
Hover over the column distribution box to reveal a contextual menu

Suggested action based on column distribution results

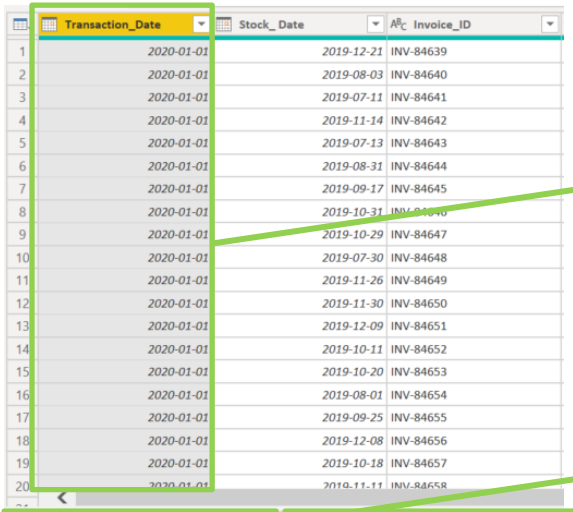
Note:

Use column distribution to identify primary keys within your lookup tables

DATA PROFILING (COLUMN PROFILE)

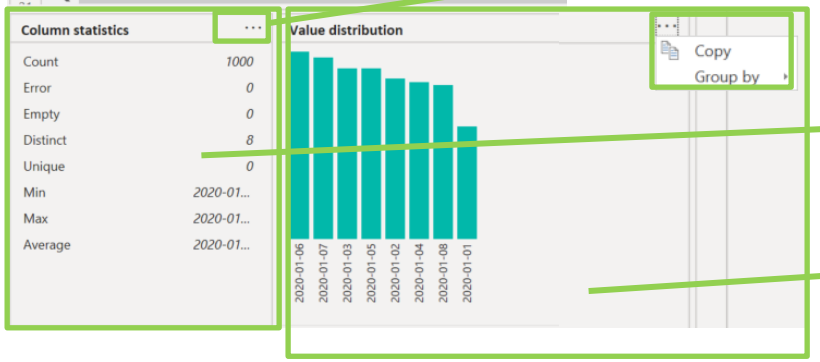


Column profile *provides a more holistic view of data within a column*
providing sample distribution of the data and column statistics



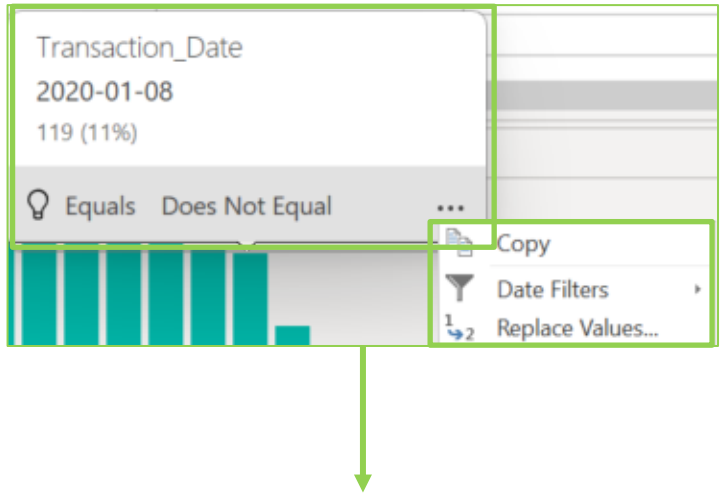
The Selected Column for profiling

Copy option allows you to copy the column statistic



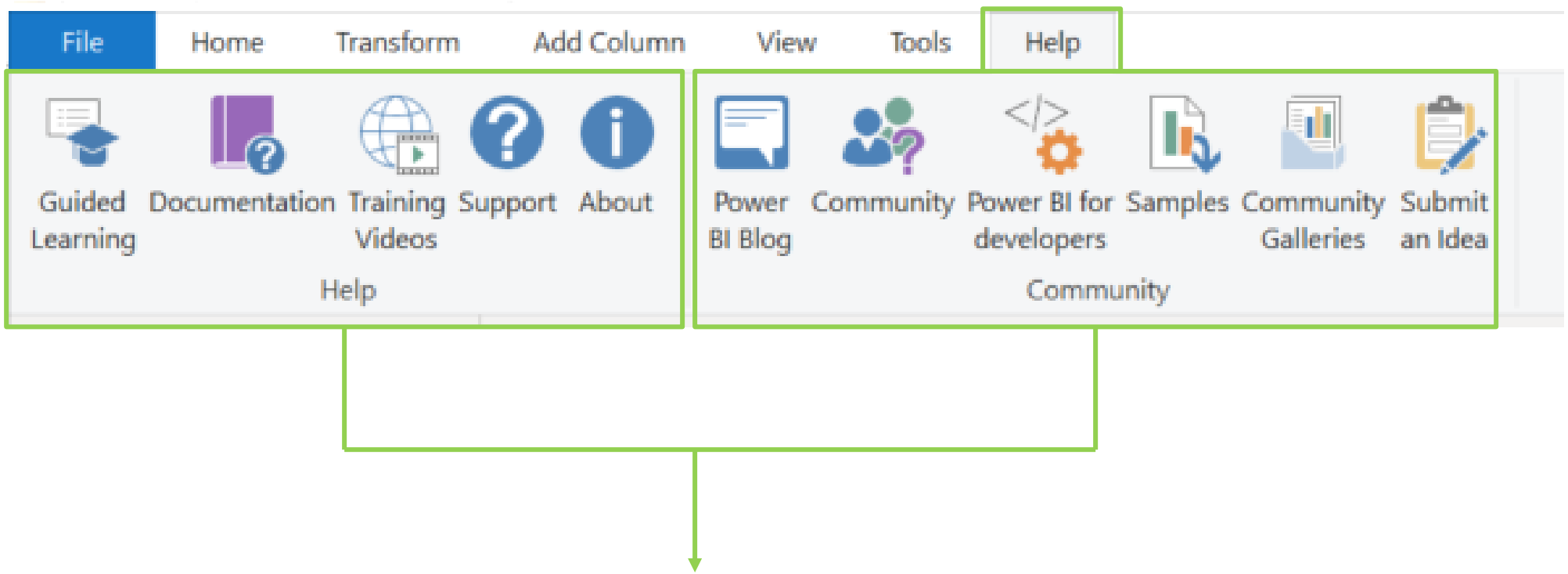
Column statistics shows the count, error, empty, distinct, unique, minimum, maximum and average values in the selected column

Value distribution shows the snapshot of the values in the selected column



Hover over value distribution bar for possible transformations and additional options

THE HELP TAB

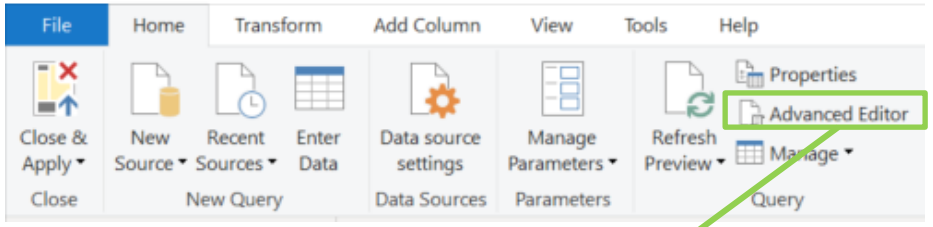


The Help and Community tools provides additional resources for getting assistance with any issues with Power BI. Check them out.

INTRODUCTION TO M LANGUAGE (ADVANCED EDITOR)

Advanced editor opens the M code that creates the query. This consists of two statements:

- Let statement: the definition of all variables
- In statement: the output of the query



let statement that defines all the variables

Variable names that match the individual steps in the Applied Steps pane

```
let
    Source = Csv.Document(File.Contents("C:\Users\benja\OneDrive\Documents\Education\Information Technology\Power BI\Training\Course\Raw Data
    #"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),
    #"Changed Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"Transaction_Date", type date}, {"Stock_ Date", type date}, {"Invoice_
    #"Sorted Rows" = Table.Sort(#"Changed Type",{{"Transaction_Date", Order.Ascending}}),
    #"Inserted Year" = Table.AddColumn(#"Sorted Rows", "Year", each Date.Year([Transaction_Date]), Int64.Type),
    #"Inserted Month Name" = Table.AddColumn(#"Inserted Year", "Month Name", each Date.MonthName([Transaction_Date]), type text),
    #"Grouped Rows" = Table.Group(#"Inserted Month Name", {"Month Name", "Year"}, {{"Quantity Sold by Month", each List.Sum([Quantity_Sold]),
    #"Pivoted Column" = Table.Pivot(#"Grouped Rows", List.Distinct(#"Grouped Rows"[#"Month Name"]), "Month Name", "Quantity Sold by Month", L
    #"Unpivoted Columns" = Table.UnpivotOtherColumns(#"Pivoted Column", {"Year"}, "Attribute", "Value")
in
    #"Unpivoted Columns"
```

in statement that is the final output of all variables in the query

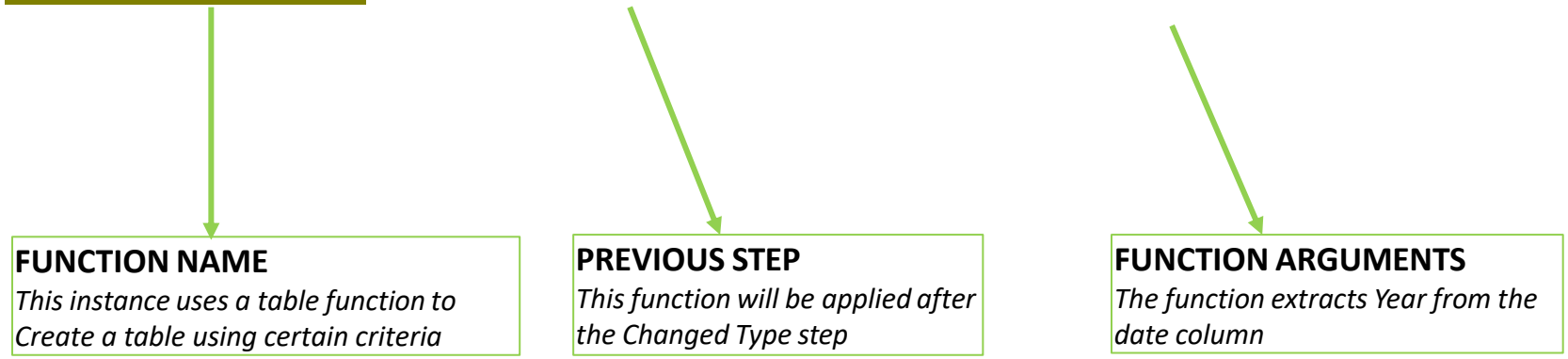
Final **output** variable

Function *called that reorders columns within the table*

M LANGUAGE SYNTAX

We will consider a statement in M language and break it down for our understanding of the syntax.

```
= Table.AddColumn(#"Changed Type", "Year", each Date.Year([Transaction_Date]), Int64.Type)
```



COMMON FUNCTION CATEGORIES

TABLE Functions

Functions to create and manipulate table values

Common Examples:

- Table.FromList
- Table.ToList
- Table.IsEmpty
- Table.FindText
- Table.RemoveColumns
- Table.Contains

Table function categories:

- Table construction
- Conversion
- Information
- Row operations
- Column operations
- Membership

LIST Function

Functions to create and manipulate list values

Common Examples:

- List.Select
- List.Contains
- List.Union
- List.Median
- List.Numbers

List function categories:

- Selection
- Membership
- Set operations
- Ordering
- Generators

TEXT Function

Functions to create and manipulate text values

Common Examples:

- Text.Length
- Text.From
- Text.Middle
- Text.Contains
- Text.Remove
- Text.BeforeDelimiter

Text function categories:

- Information
- Text comparisons
- Extraction
- Membership
- Modification
- Transformations

DATE Function

Functions to create and manipulate date, datetime, and datetimezone values

Common Examples:

- Date.EndOfMonth
- Date.EndOfQuarter
- Date.Day
- Date.StartOfWeek
- Date.StartOfMonth

THANK YOU FOR LISTENING

Q & A