

Power BI - Data Transformation - p2

NUMERICAL TOOLS

The screenshot shows the Power BI Transform ribbon with several numerical tools highlighted:

- Statistics:** Sum, Minimum, Maximum, Median, Average, Standard Deviation, Count Values, Count Distinct Values.
- Standard:** Add, Multiply, Subtract, Divide, Integer-Divide, Modulo, Percentage, Percent Of.
- Trigonometry:** Sine, Cosine, Tangent, Arcsine, Arccosine, Arctangent.
- Number Column:** Statistics, Standard, Scientific, Trigonometry, Rounding, Information.

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 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

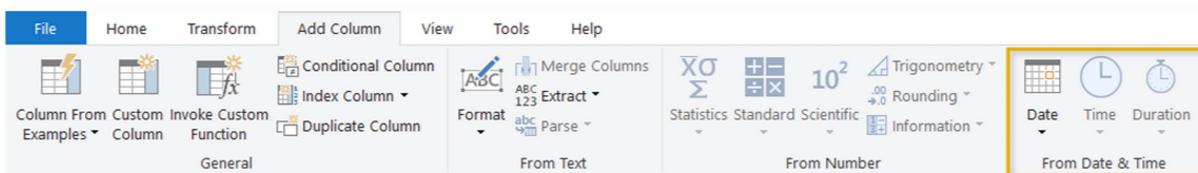
Note: Unlike the Statistics tools, these are applied to each row in the table

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

NUMERICAL TOOLS : Assignment

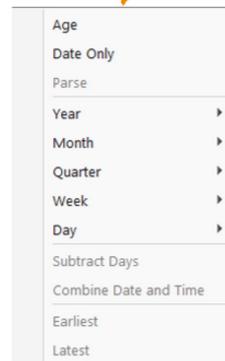
- What is our average product cost? 413.66
- How many colors do we sell our products in? 10
- How many distinct customers do we have? 18018 18148
- What is the maximum annual customer income? 170000
- Return the tables to their original state

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)

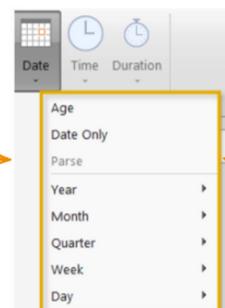


Note: You will almost always want to perform these operations from the "Add Column" menu to build out new fields, rather than transforming an individual date/time column

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

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

ASSIGNMENT: CALENDAR TABLES



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")

INDEX COLUMNS

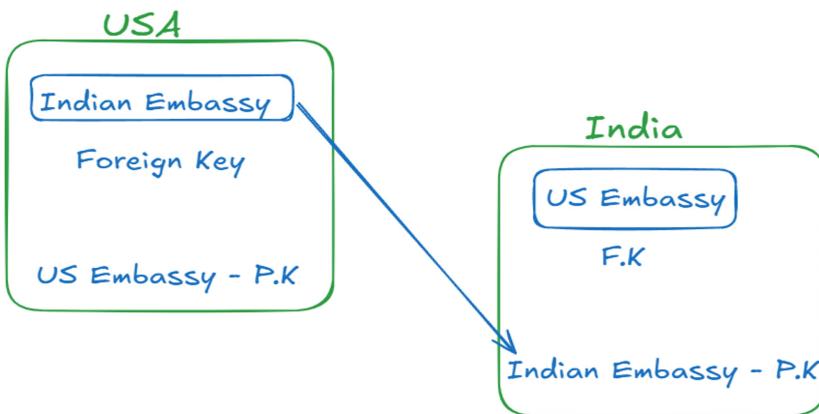


The screenshot shows the Power BI ribbon with the 'Add Column' tab selected. Below the ribbon, there are several icons and buttons. A yellow box highlights the 'Index Column' button, which is located under the 'General' section of the 'Add Column' dropdown menu.

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!)

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



Product Table



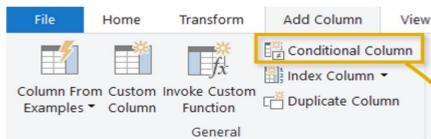
Customer Table



Prod Key	Cust Key
101	1021
121	1021
111	1021
222	1022
222	9212
222	4233
222	3243

Transactional Record

CONDITIONAL COLUMNS



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

The dialog box is titled "Add Conditional Column" and contains the following instructions:

Add a conditional column that is computed from the other columns or values.

New column name:

Column Name Operator Value Output

If Order Quantity	equals	1	Then Single Item
Else If Order Quantity	is greater than	1	Then Multiple Items
Add Clause			
Else Other			

OK Cancel

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"