

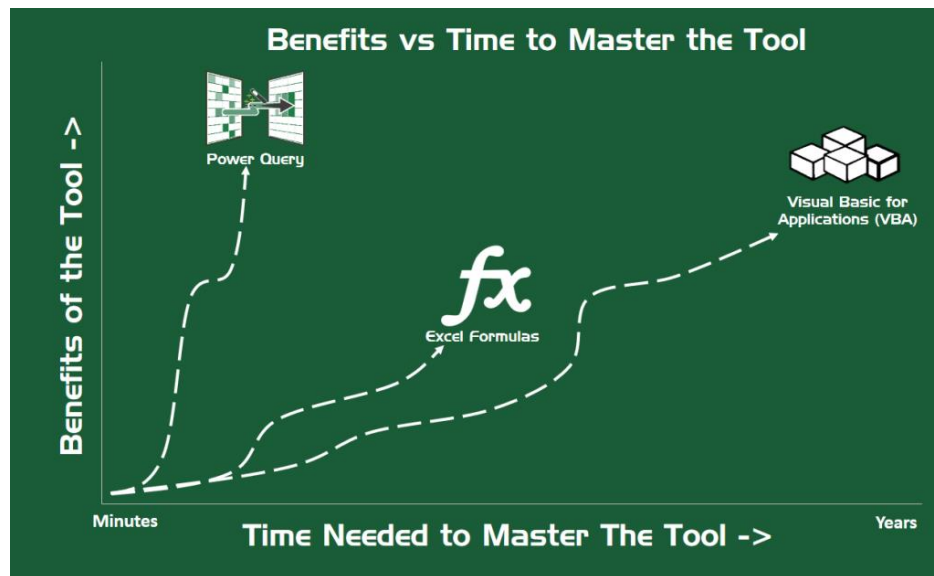
Power Query

Created by:	Marek Prazak	Reviewed by:	Marek Prazak	Create Date:	4 April 2019
--------------------	--------------	---------------------	--------------	---------------------	--------------

1.0	Purpose	2
2.0	Definition	2
3.0	Content	Error! Bookmark not defined.
3.1	Getting started	3
3.1.1	Single source.....	3
3.1.2	Multiple source files	4
3.1.3	Uploading the PQ result	7
3.2	Navigating PQ	8
3.3	Functions of PQ	11
3.4	Tips and tricks.....	13
4.0	Closing Exercise.....	15
5.0	Additional knowledge	16

1. Purpose

- The Power Query helps you refine your data for later use in the VBA or just to create a table for the pivots and the reports. Below you can see why we would recommend to get familiar with the PQ.



- The lesson should take approximately 1 hour to finish, as that is not nearly enough to teach you all there is about PQ, this lesson should give you only solid base to start from.

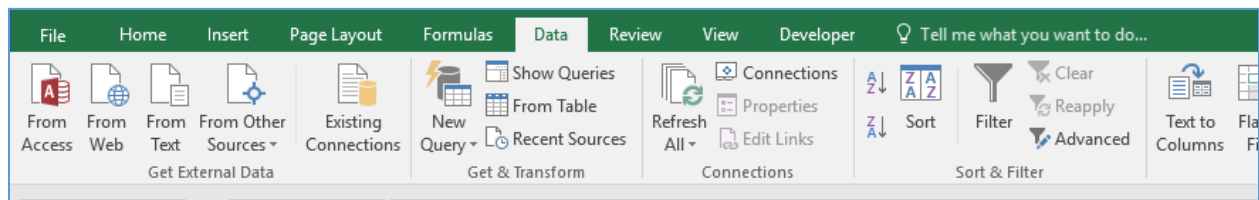
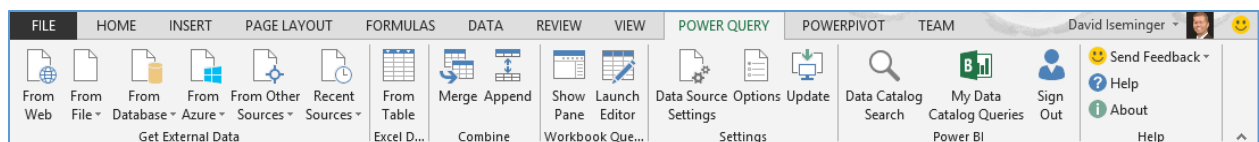
2. Definition

- **Power Query** is a data connection technology that enables you to discover, connect, combine, and refine your data sources to meet your needs, you'd notice the advantage of PQ the most in cases where you need to get an overview of employees from tracker, give them for example addresses from employee detail report and in addition calculate additional values and dates. To do all this in formulas would result in many hours of your time spend on creation of correct formulas and cleaning the data as well as encountering the limits of excel such as having tens of thousands of formulas in one workbook that might result in crash of excel. On the other hand, done through PQ it takes up to few min.
- Features in Power Query are available in Excel and Power BI Desktop.

3. Content

3.1 GETTING STARTED

- You can create queries that are as simple or complex as you need. And since Power Query uses the **M Language** (make sure to remember that this is Power Query M when searching for answers online as there is also Power Query BI) to record and carry out its steps, you can create queries from scratch (or tweak them manually) to harness the power and flexibility of data scripting, all within Power Query.
- You can access Power Query from the Power Query ribbon in Excel or from Data tab Get & Transform option.



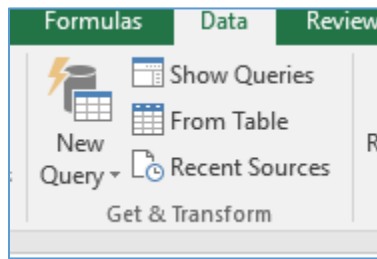
- You can use **Power Query** to connect to a single data source, such as an Excel workbook, or you can connect to multiple data sources.
- You connect to data sources from the **Power Query** ribbon, in the **Get External Data** section. Data sources include data from the **Web**, **File**, **Database**, **Azure**, **Other Sources**, or even **Tables** in an Excel workbook. In this lesson we will focus on Single and multiple excel sources.

3.1.1 SINGLE SOURCE

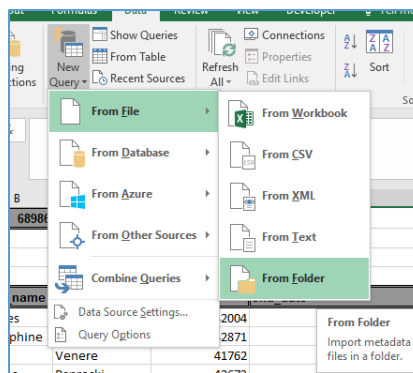
- As mentioned above there are multiple source options, the ones that you'll encounter the most is From Table and From Workbook

HRS Transformation

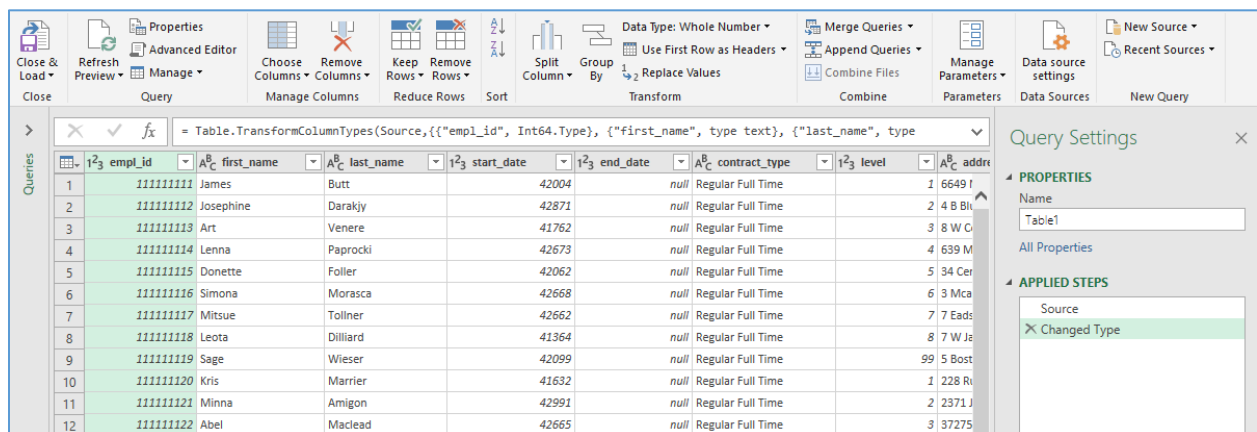
- From Table – Transform your source range into a table and then click on From Table.



- From Workbook** – click on New Query and from the menu select From Workbook



- Result:**

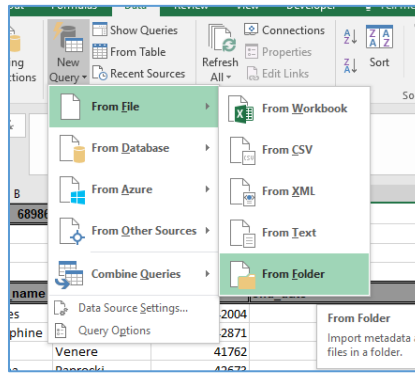


	empl_id	first_name	last_name	start_date	end_date	contract_type	level	address
1	111111111	James	Butt	42004	null	Regular Full Time	1	6649 I
2	111111112	Josephine	Darakjy	42871	null	Regular Full Time	2	4 B Blt
3	111111113	Art	Venere	41762	null	Regular Full Time	3	8 W C
4	111111114	Lenna	Paprocki	42673	null	Regular Full Time	4	639 M
5	111111115	Donette	Foller	42062	null	Regular Full Time	5	34 Cer
6	111111116	Simona	Morasca	42668	null	Regular Full Time	6	3 Mca
7	111111117	Mitsue	Tollner	42662	null	Regular Full Time	7	7 Eads
8	111111118	Leota	Dilliard	41364	null	Regular Full Time	8	7 W Je
9	111111119	Sage	Wieser	42099	null	Regular Full Time	99	5 Bost
10	111111120	Kris	Marrier	41632	null	Regular Full Time	1	228 R
11	111111121	Minna	Amigon	42991	null	Regular Full Time	2	2371 J
12	111111122	Abel	Maclead	42665	null	Regular Full Time	3	37275

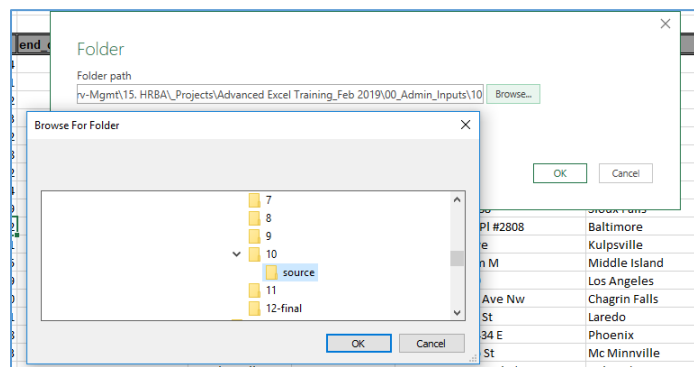
3.1.2 MULTIPLE SOURCE FILES

- Click on **From Folder** option.

HRS Transformation



- Select location of your folder.



- Click **OK**
- You will see all files in the folder that will be included in the query.
- Click on **Edit**

Content	Name	Extension	Date accessed	Date modified	Date created	Attributes	Folder Path
Binary	source1 - Copy (2).xlsx	.xlsx	4/3/2019 11:01:27 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:27 AM	Record	\\ant\dept-eu\PRG10\
Binary	source1 - Copy (3).xlsx	.xlsx	4/3/2019 11:01:27 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:27 AM	Record	\\ant\dept-eu\PRG10\
Binary	source1 - Copy.xlsx	.xlsx	4/3/2019 11:01:27 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:27 AM	Record	\\ant\dept-eu\PRG10\
Binary	source1.xlsx	.xlsx	4/3/2019 11:01:12 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:12 AM	Record	\\ant\dept-eu\PRG10\

Combine Load Edit Cancel

- Now you can set up rules in case your folder contains more files that you don't want to include.
 - For example, I will apply filter to not include any file that contains "2" (Does not contain) in the name and even though currently these files are .xlsx(Excel) I will still apply filter for this data type, just in case.

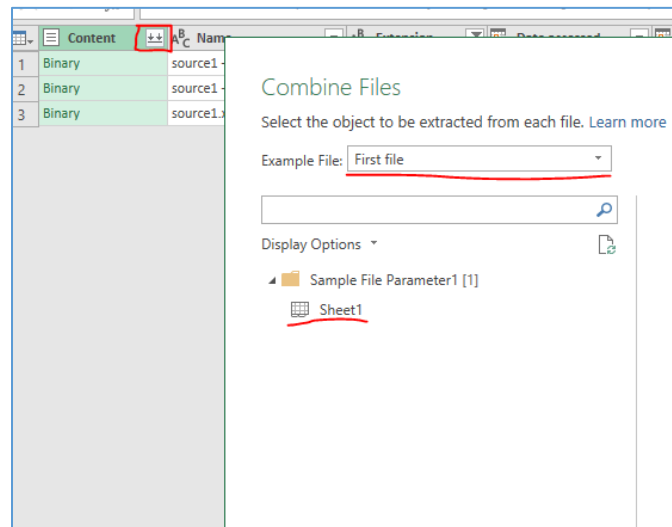
	Content	Name	Extension	Date accessed	Date modified	Date created	Attrib
1	Binary	source1 - Copy (2).xlsx	.xlsx	4/3/2019 11:01:27 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:27 AM	Record
2	Binary	source1 - Copy (3).xlsx	.xlsx	4/3/2019 11:01:27 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:27 AM	Record
3	Binary	source1 - Copy.xlsx	.xlsx	4/3/2019 11:01:27 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:27 AM	Record
4	Binary	source1.xlsx	.xlsx	4/3/2019 11:01:12 AM	4/3/2019 11:01:12 AM	4/3/2019 11:01:12 AM	Record

- You can see that when I applied these filters my steps have been recorded in the Applied steps.

Content	Name	Extension	Date accessed
1	Binary	source1 - Copy (3).xlsx	4/3/2019 11:01:27 AM
2	Binary	source1 - Copy.xlsx	4/3/2019 11:01:27 AM
3	Binary	source1.xlsx	4/3/2019 11:01:12 AM

Query Settings	
PROPERTIES	
Name	source
All Properties	
APPLIED STEPS	
Source	✱
Filtered Rows	✱
✕ Filtered Rows1	✱

- Click on Combine (**two arrows**)
- You can select the example file (based on example file, header and formatting will be set), but we will leave the first file in this case
- Select Sheet1 that will be taken from all the files.
- Click OK

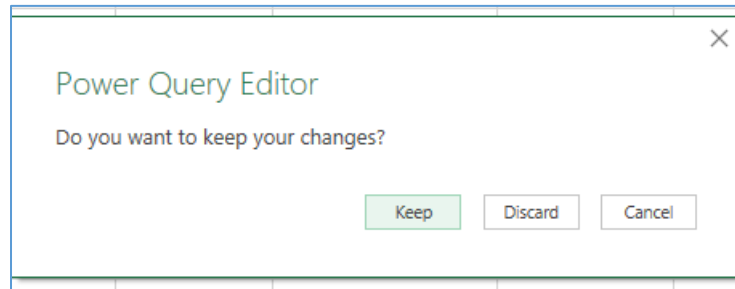
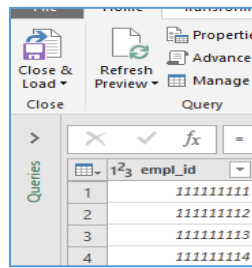


- Compared to the individual source you can see that this query additionally contains a column with name of the source file.

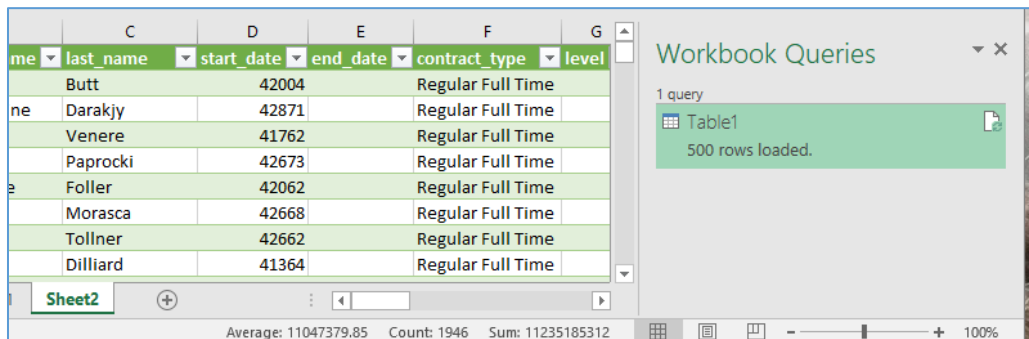
= Table.TransformColumnTypes("#Expanded Table Column1",				
	Source.Name	empl_id	first_name	last_name
1	source1 - Copy (3).xlsx	111111111	James	Butt
2	source1 - Copy (3).xlsx	111111112	Josephine	Darakijy
3	source1 - Copy (3).xlsx	111111113	Art	Venere
4	source1 - Copy (3).xlsx	111111114	Lenna	Paprocki
5	source1 - Copy (3).xlsx	111111115	Donette	Foller
6	source1 - Copy (3).xlsx	111111116	Simona	Morasca
7	source1 - Copy (3).xlsx	111111117	Mitsue	Tollner
8	source1 - Copy (3).xlsx	111111118	Leota	Dilliard
9	source1 - Copy (3).xlsx	111111119	Sage	Wieser
10	source1 - Copy (3).xlsx	111111120	Kris	Marrier
11	source1 - Copy (3).xlsx	111111121	Minna	Amigon
12	source1 - Copy (3).xlsx	111111122	Abel	Maclead
13	source1 - Copy (3).xlsx	111111123	Kiley	Caldarera
14	source1 - Copy (3).xlsx	111111124	Graciela	Ruta
15	source1 - Copy (3).xlsx	111111125	Cammy	Albares
16	source1 - Copy (3).xlsx	111111126	Mattie	Poquette

3.1.3 UPLOADING THE PQ RESULT

- Once you're satisfied with your PQ you can either use **Close & Load** option or you can close the window and confirm **Keep** which will load the PQ table.

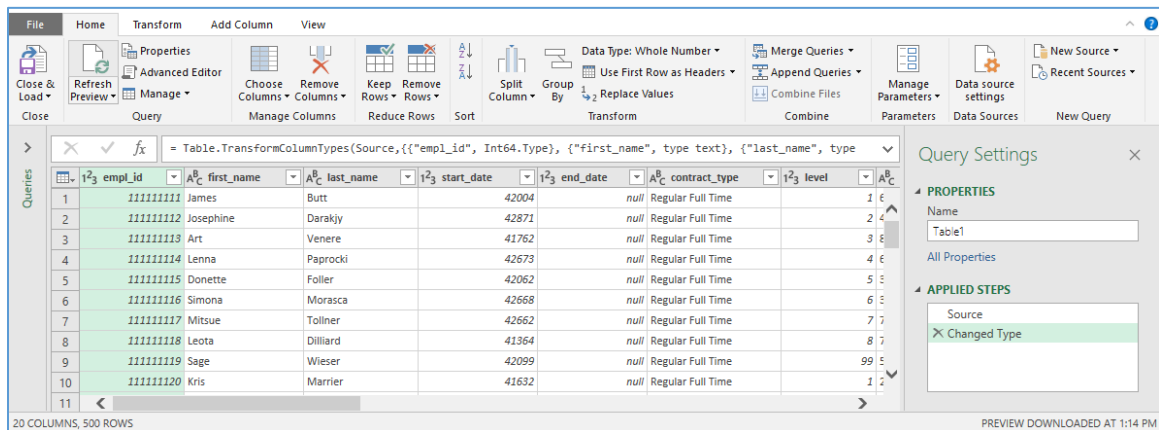


- New sheet will be generated with table containing outcome of the PQ.



3.2 NAVIGATING PQ

- Query settings



- **Properties** show you the **Name** of the query, which you should keep relative to the purpose of the query.
- **Applied steps** shows you all the steps that you've taken with this query. In each new query you will see Source that contains path to the source file and Changed Type where the PQ will choose adequate formatting for each column (ABC-text, 123-numbers, ABC123-combination, etc.). You can change this formatting as you see fit, but take into account that text will not be recognized as numbers. (Meaning, if you have a word in column of numbers, after the formatting to numbers this word will show up as an error and in resulting table it will disappear.)
- **Query preview** displays only limited amount of rows to make the editor faster. In bottom left corner you will see amount of displayed data.
- **Formula bar** displays code for the selected step.
- **Queries** bar on the left is minimized, but when enlarged it will help you navigate all queries in this workbook.
- When you click on any **cell**, it's **content** will be displayed in the bottom of the PQ

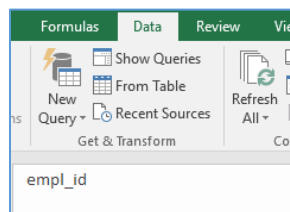
3	111111113	Art	Venere
4	111111114	Lenna	Paprocki
5	111111115	Donette	Foller
6	111111116	Simona	Morasca
7	111111117	Mitsue	Tollner

Foller

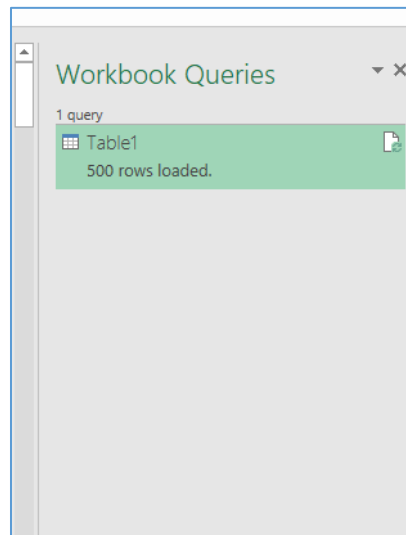
- When you highlight any **row** it's **content** will be displayed in the bottom of the PQ

3	11111113	Art	Venere
4	11111114	Lenna	Paprocki
5	11111115	Donette	Foller
6	11111116	Simona	Morasca
7	11111117	Mitsue	Tollner
8	<		
empl_id		11111115	
first_name		Donette	
last_name		Foller	
start_date		42062	
end_date		null	
0 COLUMNS, 500 ROWS			
11111127 Mooghan Garufi			

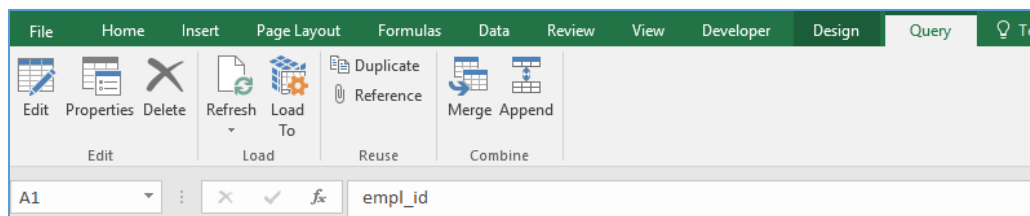
- Outside of PQ editor you can view and access all the PQ by navigating to **DATA** tab and clicking **Show Queries**



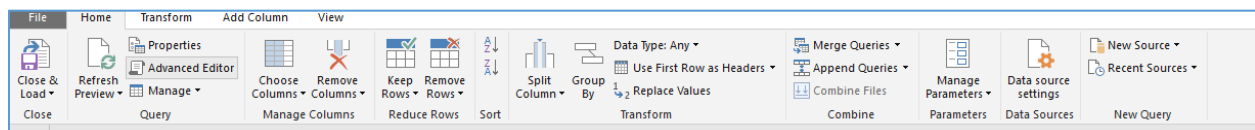
- You can rename, refresh, delete or do other basic commands for the individual PQ from this view.



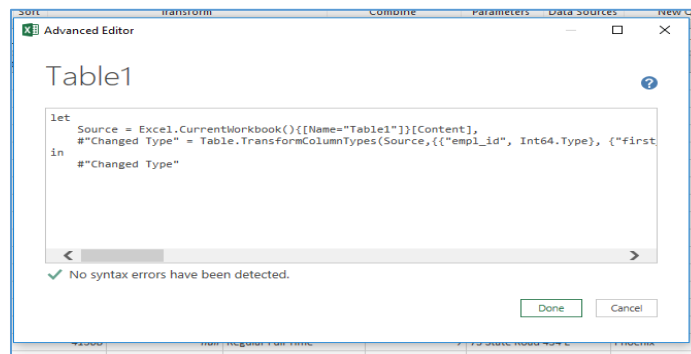
- In **Query** tab you have option to **merge**, **append** query. Create **reference** query(copy without previous steps) or **duplicate** the whole query(copy with all steps)



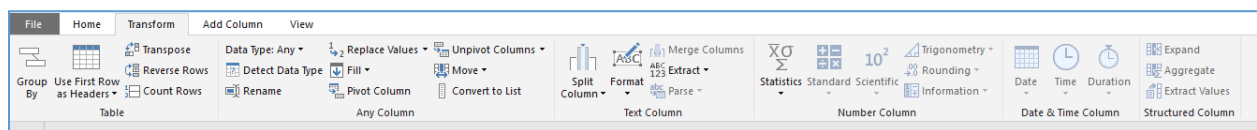
3.3 FUNCTIONS OF PQ



- **Home Tab**
 - **Advanced Editor** shows the code for the whole power query until currently selected step.

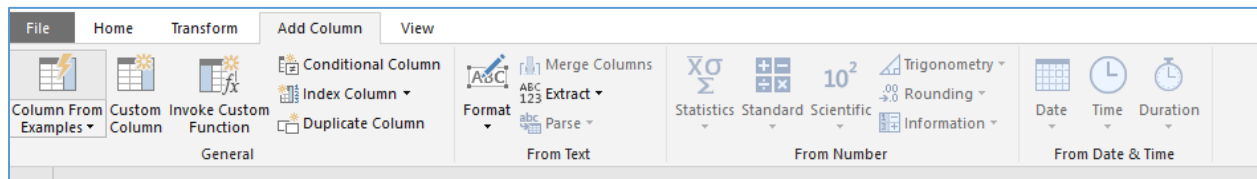


- **Choose Columns** allows you to select columns that you want to keep, all other will be deleted.
- **Remove Columns** deletes columns that you select.
- **Keep/Remove rows** has same logic as similar function for columns.
- **Sort** based on name.
- **Split columns** based on preselected rules.
- **Group** rows according to certain rules.
- Often done automatically **Use First Row as headers** will take first row of your source and use it as a header.
- **Replace value** in whole or selected region of PQ with different value (null to N/A, etc.)
- **Merge** two or more **Queries** together (similar to vlookup).
- **Append Queries** will put together queries with the same column

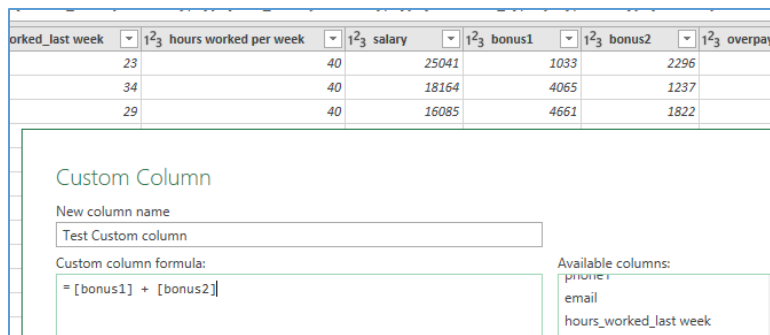


• Transform Tab

- **Transpose** table will switch rows for columns and columns for rows
- **Convert to List** will take one column and deletes all else.
- **Format** will change the formatting of selected column
- **Extract** part of column based on selected choice.
- **Number Column** section and **Date & Time Column** section allows you to do mathematical and date adjustment to the columns



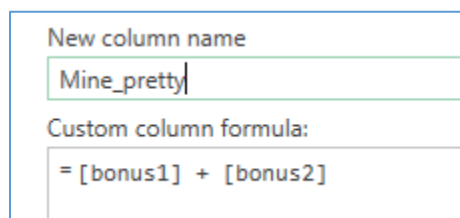
- **Add Column**
 - **Custom column** lets you create column with implemented coding (if, sum and other excel like formulas)



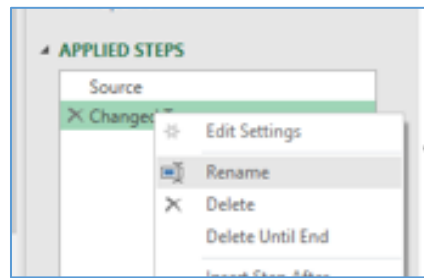
- **Conditional Column** allows you to create column with option for if condition

3.4 TIPS AND TRICKS

- It is easier and faster to name your custom columns when you're creating them as it will spare you extra steps.



- Renaming Applied steps might help you with navigating the steps the longer they get. In addition, it is much easier to understand the steps for anyone who comes after you.



- Queries when created are loaded into new sheet and table which they are connected to. If you have multiple queries that are connected to each other (merged, append, duplicated, etc.), but only one is resulting query, it is recommended to delete the sheet and table of the supporting queries as loading each single one into your excel will make the size of workbook larger and it will slow down the loading time of queries.

	A	B	C	D	E	F
1	empl_id	first_name	last_name	start_date	end_date	contract_type
2	111111111	James	Butt	42004		Regular Full Time
3	111111112	Josephine	Darakjy	42871		Regular Full Time
4	111111113	Art	Venere	41762		Regular Full Time
5	111111114	Lenna	Paprocki	42673		Regular Full Time
6	111111115	Donette	Foller	42062		Regular Full Time
7	111111116	Simona	Morasca	42668		Regular Full Time
8	111111117	Mitsue	Tollner	42662		Regular Full Time
9	111111118	Leota	Dilliard	41364		Regular Full Time
10	111111119	Sage	Wieser	42099		Regular Full Time

Query Name	Status
support	Connection only.
result	500 rows loaded.

- Query that is used as support does not have the outcome tab and is listed as connection only, resulting query has outcome table and is listed as loaded.
- If you want to load Connection only query into table, click on query with right button and choose load to.

4. Closing Exercise

- 1) Create new excel
- 2) Load query from excel named les_7_source1. You have to adjust the sheet in the query to load only the main table not the details above it. (Query will be run weekly so it is not possible to adjust the source directly as it would create unnecessary work)
- 3) Amend the formatting for address, each word will have first letter capital
Login to be small letters
- 4) Filter out AAs who are higher than lvl3 and 9lvl
- 5) Create column of amazon emails
- 6) Create column that would add bonus1 to bonus2
- 7) if the AA has worked more hours per week in last week than is supposed to add 10% of his salary to the column from step 7
- 8) Use les_7_source2 and create mapping table in new sheet to allow users to always update it, this. (One column would contain full name of the state and second abbreviation of the state name)
- 9) Create connect only power query out of the state table.
- 10) In main query Create column of consolidated addresses which are separated by comma (Address, City, County, State, Zip) and remove the original columns of these fields. (State will be full name not just abbreviation.
- 11) Create VBA code that would refresh the queries when the excel is first opened and informed the user by msgbox that file is ready.
- 12) Create macro button that would refresh the file.
- 13) Create macro that would create new workbook(.csv) which will copy the result of your query and past it into the CSV. Make sure that data from query are pasted without formatting and formulas.

5. Additional knowledge

- <https://www.powerquery.io/>
- <https://docs.microsoft.com/en-us/powerquery-m/power-query-m-reference>