

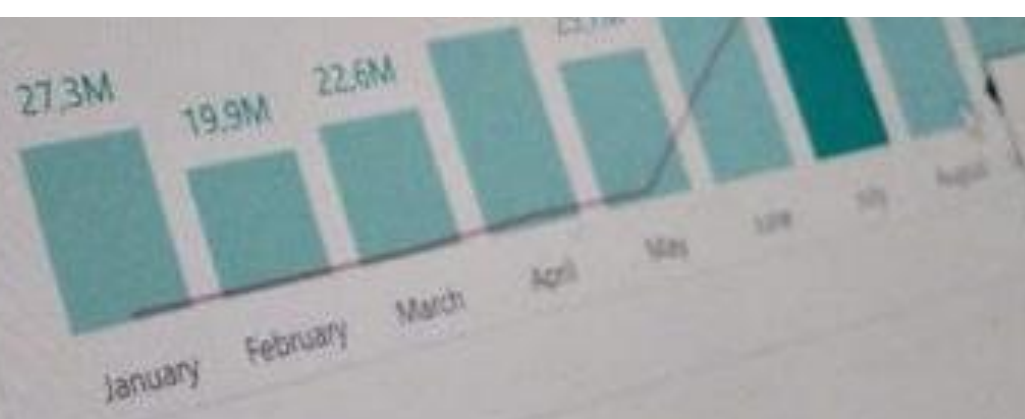
- Year
- 2015
 - 2016
 - 2017
 - 2018

Sales

● Sales ● Sales LY

Power BI Data Analyst I

- Time period
- by period
 - year-to-date
 - 12-mth rolling total



About: Michiel Rozema



Mathematician with 25 years of experience in IT
8 years Data Insight Lead at Microsoft, NL
Power BI expert, consultant, trainer, coach,
author

Founder & CEO of Quanto, *Collective Analytics*
Founder of Dutch Power BI User Group
Organizer of the Power BI Summer School

michiel.rozema@quanto.eu

www.linkedin.com/in/MichielRozema

www.quanto.eu

www.quanto-training.eu

www.powerbisummerschool.eu



Power BI
Summer School



Quanto
collective analytics

Agenda – dag 1

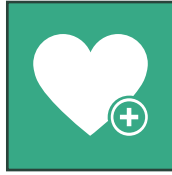
- The Five-layer model for BI & Power BI overview
- Working with Power BI reports
- Data transformation with Power Query

The Five-Layer model for Business Intelligence



Share

Collaborate and distribute insights to others



Visualize

View results, get insights, and play with reports and dashboards



Analyze

Define business views and implement calculations in analytical models



Prepare

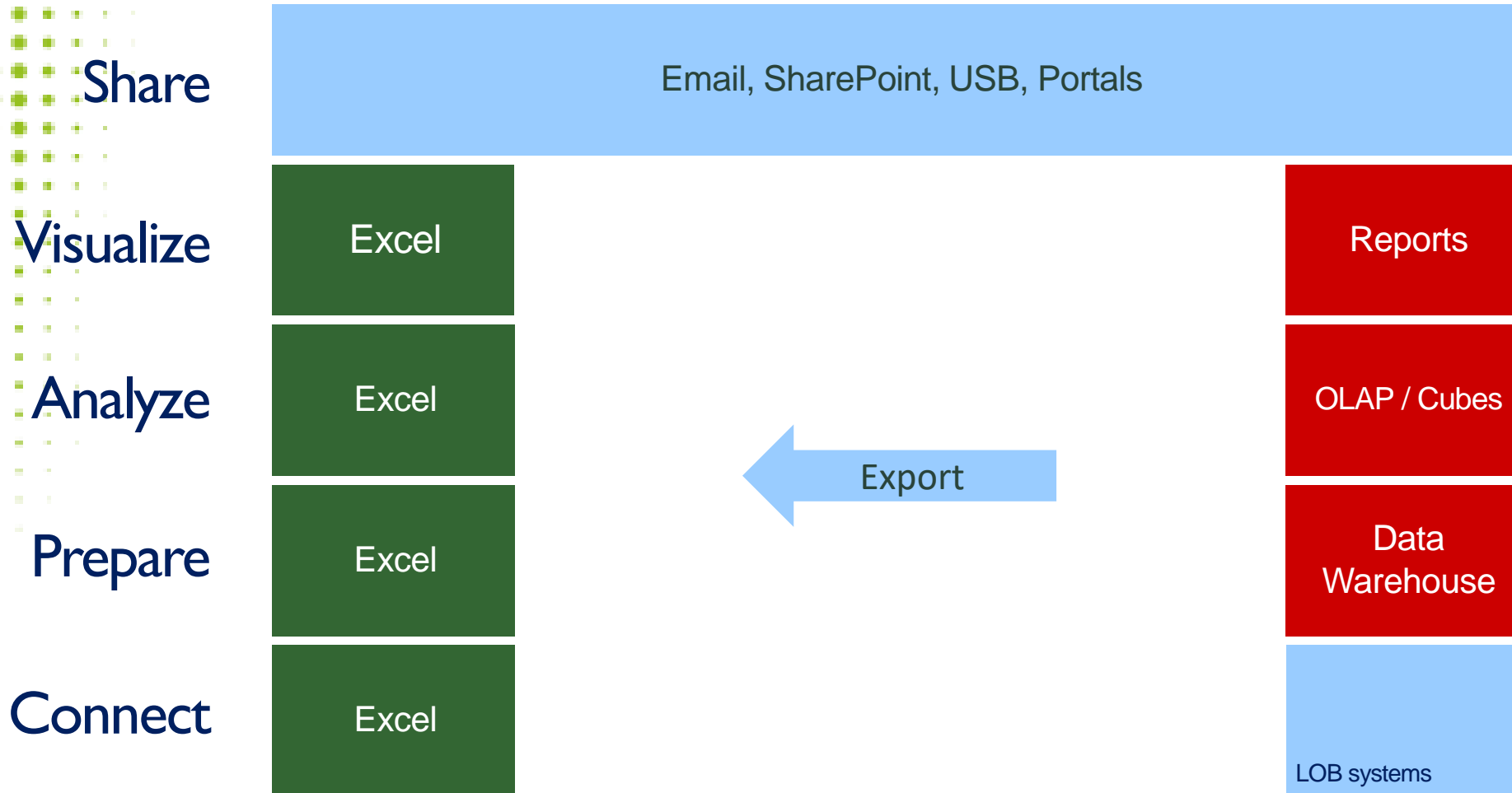
Transform, cleanse, and combine data



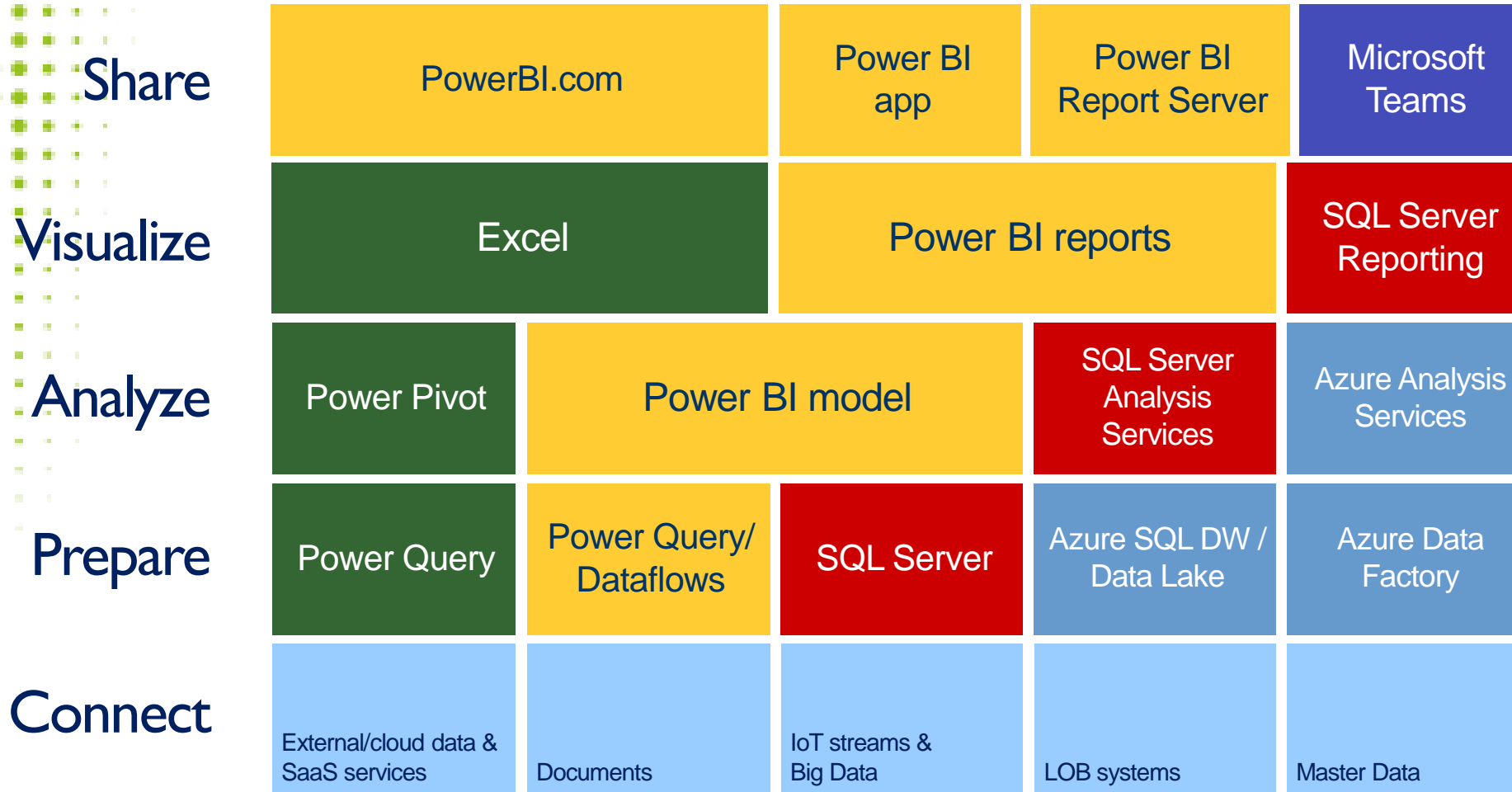
Connect

Discover data and extract data from source

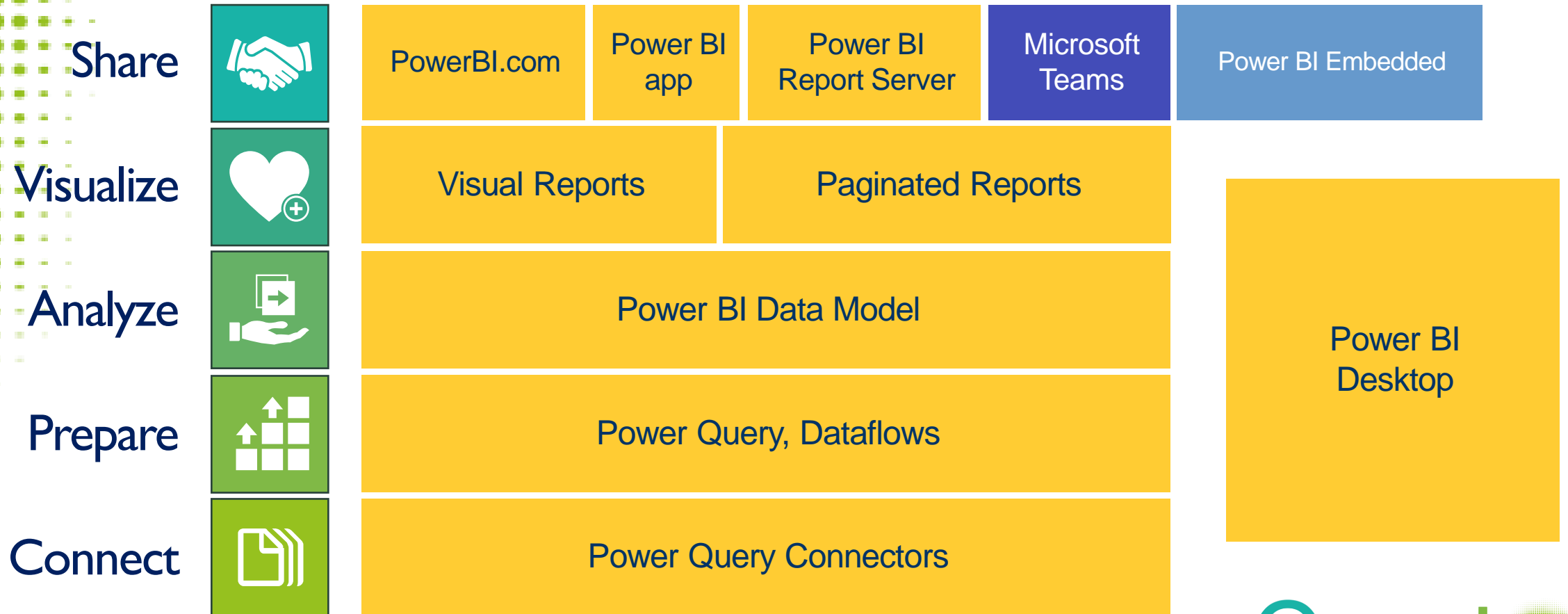
The Traditional BI landscape



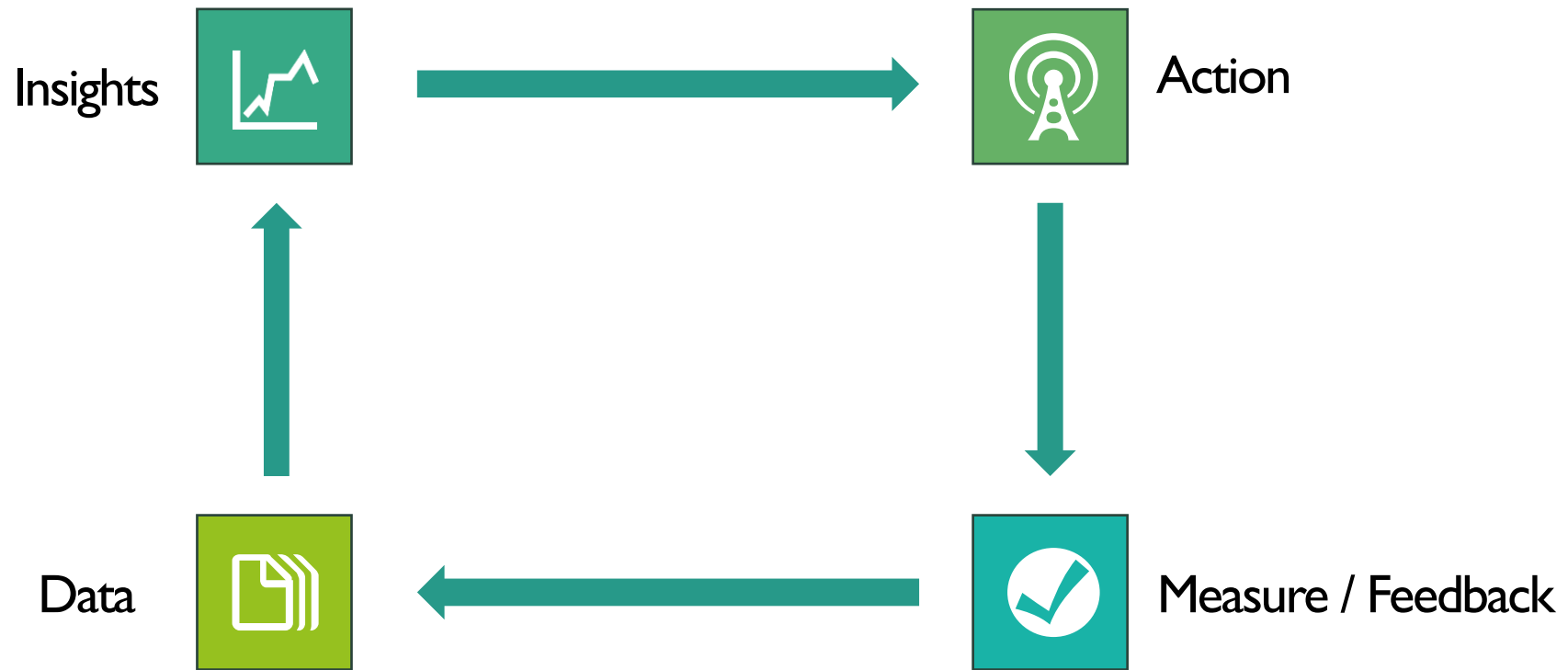
The Modern Microsoft BI platform



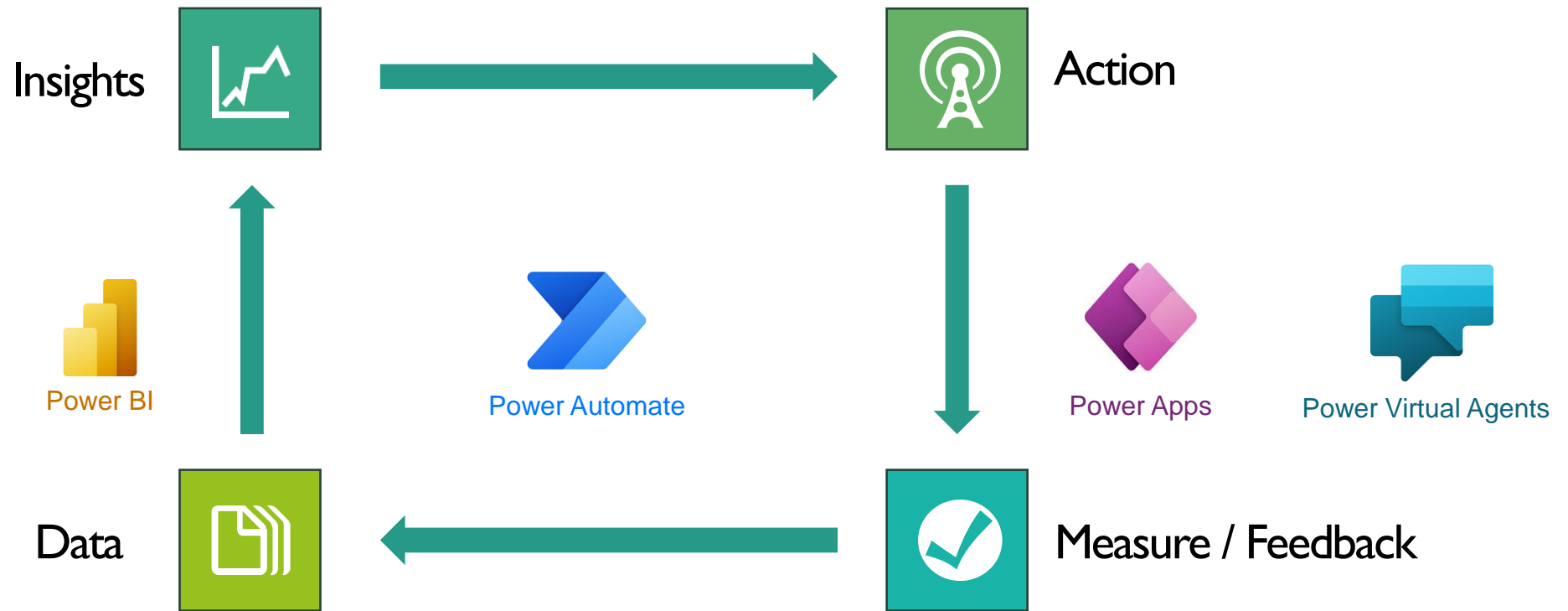
Power BI Platform Components



Data-driven business transformation



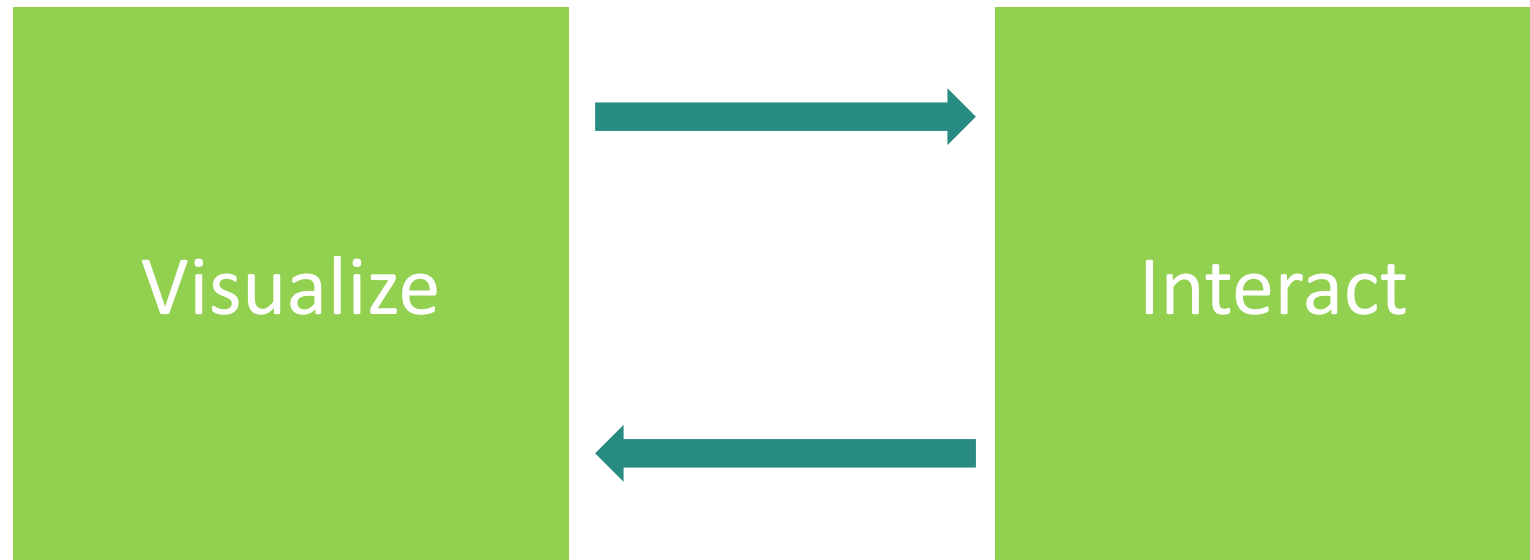
The Power Platform





Power BI Reports

From data to insights: visual reporting





Demo

<https://quanto.eu/powerbireport2>



Exercises

<https://quanto.eu/powerbireport2>

<https://quanto.eu/trainingdata>

Exercise

Working with Power BI reports – I

- *What was the year-to-date sales for team Asia Pacific 1 in July 2020? What was the year-on-year growth%?*
- *What was the sales on clothing in October 2019?*
- *What was the year-to-date sales growth % of team Europe 1 in September 2019, for bikes and components combined?*
- *Which employee in Europe had the highest sales on components in 2018?*
- *What is Giulia Moore's birthday?*
- *What was Giulia Moore's year-to-date sales number in November 2017?*

Exercise

Working with Power BI reports –2

- *What net profit percentage did Dasha Gonzalez realize on sales of the Road-550-W Yellow 44 bike to direct customers, year-to-date as of July, 2019?*
- *What was the total sales number on bikes in April, 2020 in the region NordRhein-Westfalen?*



Power Query / Dataflows

What is Power Query

- **Extract** data from data sources
- **Transform** data into a convenient format
- **Load** data into a Power BI model

- ETL instructions are defined working on a data **preview**
- ... are stored in a script called a **query**, based on the M language
- ... are re-applied when (automatically) **refreshing** data

What is Dataflows

- “Power Query in the cloud”
- A Dataflows query defines a data table, called a Dataflow
- Dataflows can be shared within Power BI workspaces
- Dataflows help to **reuse** data transformations

- Dataflows stores data as tables in Azure Data Lake storage
- These tables can be used outside of Power BI as well (if configured that way)

Where Power Query / M can be used

Power Query
(Power BI)

Get &
Transform
(Excel)

Dataflows
(Power BI)

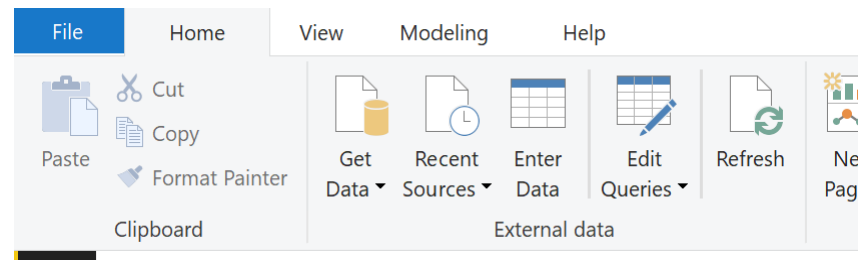
Dataflows
(PowerApps)

Dataflows
(Azure Data
Factory)

Analysis
Services
(SQL Server)

Working with Power Query in Power BI Desktop

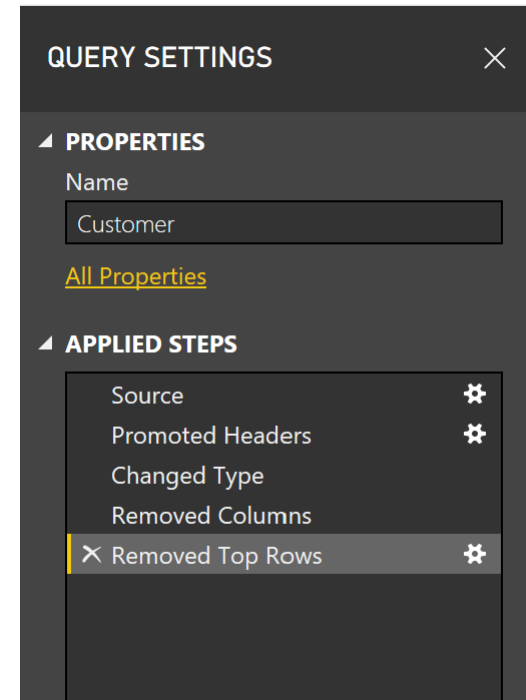
- Start from the **Home** tab in the ribbon, section **External Data**
- Choose your data source or click **Edit Queries**
- Work in the **Power Query Editor** window to form your dataset
 - You can work on multiple queries at the same time
- When ready, load the data through the **Close & Apply** button



Query Editor basics

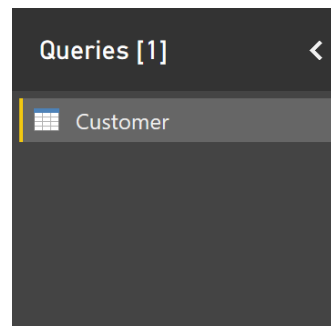
- Query Settings panel

The gear icon means the step can be reconfigured later



Query script / steps*

- Queries panel



* Query steps can be renamed, this may be used to make a script more readable

What you can do on the query level

By right-clicking in the Queries panel, you can

- Duplicate a query
- Reference a query: use it as a starting point for a new query
- Set a query to not load data: use this when it is only used for referencing by other queries
- Set a query not to refresh data
 - Note: this works only in Power BI Desktop, not in the Power BI Service!



Data Transformations

Removing Columns and Rows

- Removing columns
 - **Remove Columns:** cannot be reconfigured
 - **Choose Columns:** can be reconfigured
- Removing rows
 - **Remove Rows:** Top, Bottom, Alternate, Duplicates, Blanks, with Errors
 - **Keep Rows:** Top, Bottom, Range, Duplicates, with Errors
 - Some of these depend on selected columns – select multiple columns with CTRL or SHIFT

Splitting Columns

- By **Delimiter**, once from left or right, or all
- By **Number of Characters**, once from left or right, or all
- By **Uppercase to Lowercase**, or v.v.
- By **Digit to Non-Digit**, or v.v.

Changing Column Names and Data Types

- To change a column name, double-click the column name
- To change a data type, click the data type icon before the column name
- The **Using Locale...** option forces Power Query to read values using specific regional settings
- Note: the thin line below the column name indicates the status of all values in the column: OK, blank, or error

Other transformations

- **Use First Row as Headers**
 - Replaces the column names by the values in the first row
- **Replace Values**
 - For numeric columns: Replaces values in a column by another value
 - For text columns: Replaces part of a text by another text (optionally, only replace entire values)
- **Group by**
 - Summarizes a table based on one or more columns
 - New columns can be defined that aggregate values in the original table

Transforming or Adding Columns

- Many basic transformation features are available in both the **Transform** and the **Add Column** ribbon tab.
- In general, the relevant features are enabled after selecting one or more columns in the table
- **Add Column – Column From Examples** allows to type in what you need and Power Query will guess what transformation logic should be applied

Transpose, Pivot, Unpivot

- **Transpose**

- Use columns as rows, and vice versa
- Be careful: tables can have many rows!

- **Pivot**

- Use values in selected column to create new columns
- Must provide aggregation to fill columns with values
- This is like a pivot table in Excel!

- **Unpivot**

- The reverse of **Pivot**
- Either do this on selected columns (**Unpivot Columns**) or unselected columns (**Unpivot Other Columns**)

Indexes, Conditional Columns

- **Index Column** adds a column with unique numbers
- **Conditional Column** adds a column with values depending on certain criteria (IF – THEN – ELSE)



Exercises

<https://quanto.eu/trainingdata>

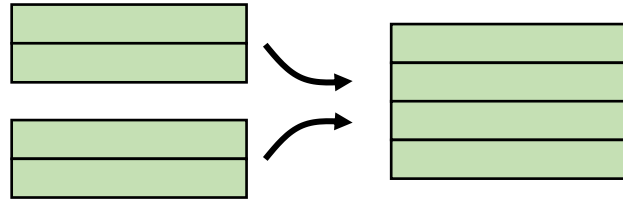
Exercises PQ 1



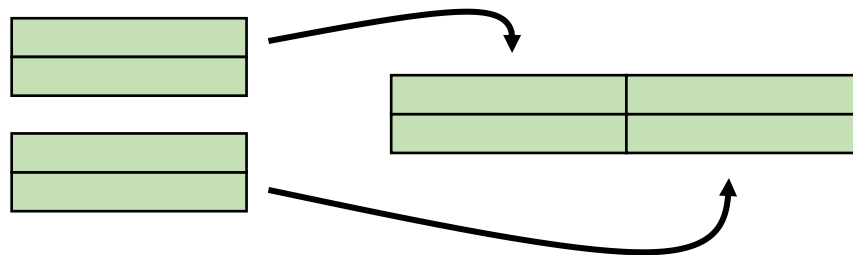
Combining Data Sources

Append and Merge

- **Append** combines two queries by adding rows.
 - Columns that have the same name will be combined in one



- **Merge** combines two queries by matching rows based on equal values in columns (in database terms, joins)
 - Choose one or more columns in each query to match on
 - Choose join type
 - The columns from the second query will be included in a single column, which should be expanded



Privacy Levels

- When combining queries from different sources, Power Query will ask to specify the *Privacy Level* of each source
- The reason: Power Query will often try to use data from one source to filter the data from the other source *before it is retrieved* (this is called Query Folding)
- Available privacy levels:
 - Public
 - Organizational
 - Private
- For data security reasons, Query Folding is only done from sources with a less restrictive privacy level

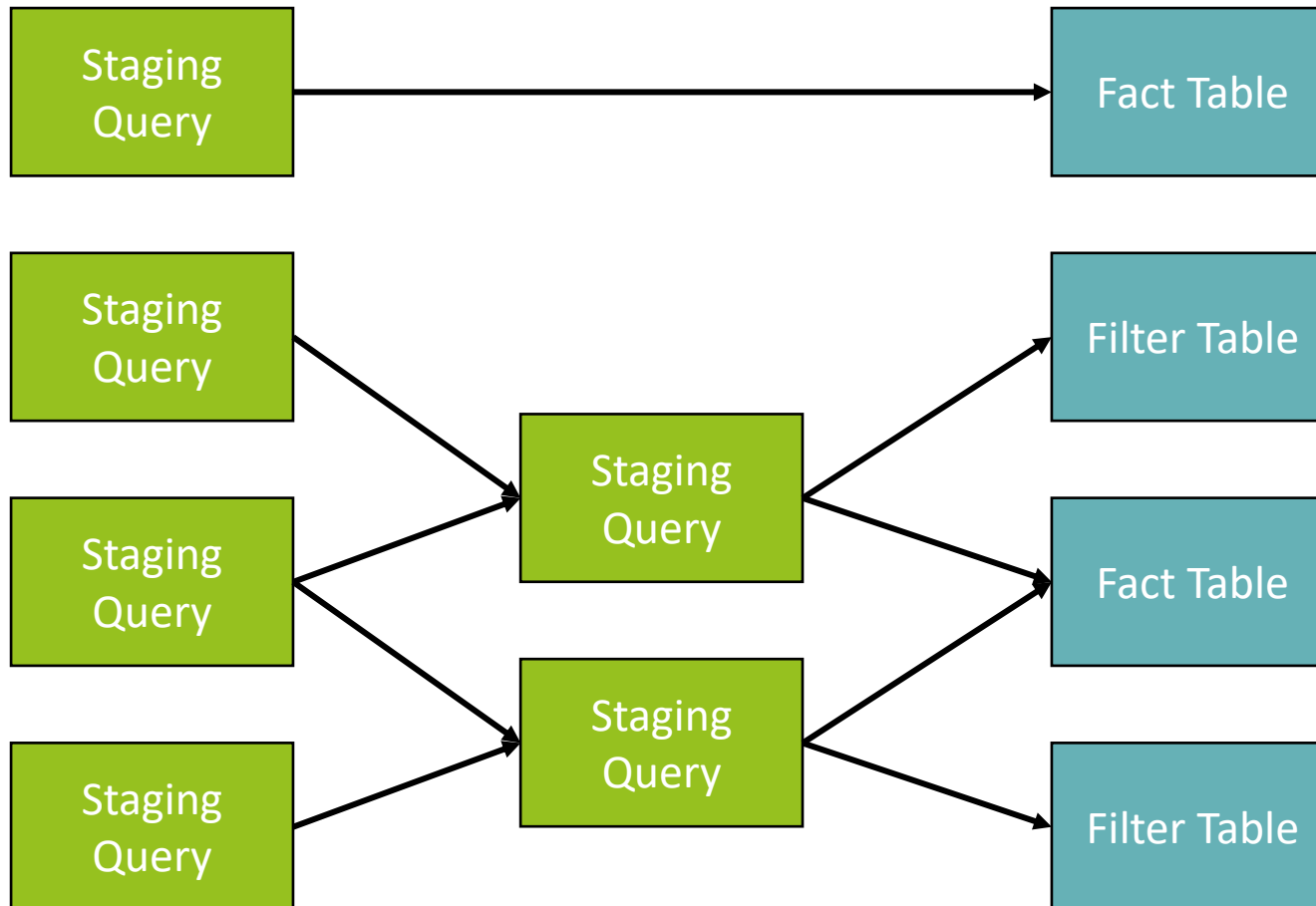
Best Practice: Staging Queries

- Power Query has trouble combining data directly from the source in one query

Formula.Firewall: Query references other queries or steps, so it may not directly access a data source. Please rebuild this data combination

- The solution: create separate queries that connect to the data sources, and combine these queries
- We call these ‘Staging Queries’
- Staging queries are typically not loaded into the data model
- Staging queries identified with the prefix ‘s’

Staging Queries (2)



Combining Files from a Folder

- Data source 'Folder'
- Select a folder, and choose **Combine and Edit**
- On the background, a lot of things will happen (we will dive into those later) resulting in loading data from all the files in the folder
- Choosing **Edit** allows for more precise querying, like selecting the files in the folder



Exercises

<https://quanto.eu/trainingdata>

Exercises PQ 2



Parameters & Query Grouping

Working with Parameters

- Parameters are predefined values
- Can be used as component of a query
- Can be changed in the Power BI service after publishing

Typical uses:

- Manage file location of local files
- Manage location of online resources
- Set development and production data sources
- Change data volumes between development and production

Creating a Parameter

- Click the **Manage Parameters** button
- Naming convention: start the parameter name with a lowercase 'p' as a prefix
- Parameters have a data type
- Parameters can be set to allow values:
 - Any value
 - Value from a fixed list
 - Value from a dynamic list coming from a query (advanced)

Parameters

New

1 ² 3	pCalendarYearFrom	
1 ² 3	pCalendarYearTo	
A ^B C	pCustomerFile	X

Name: pCustomerFile

Description: The name of the customer file

Required

Type: Text

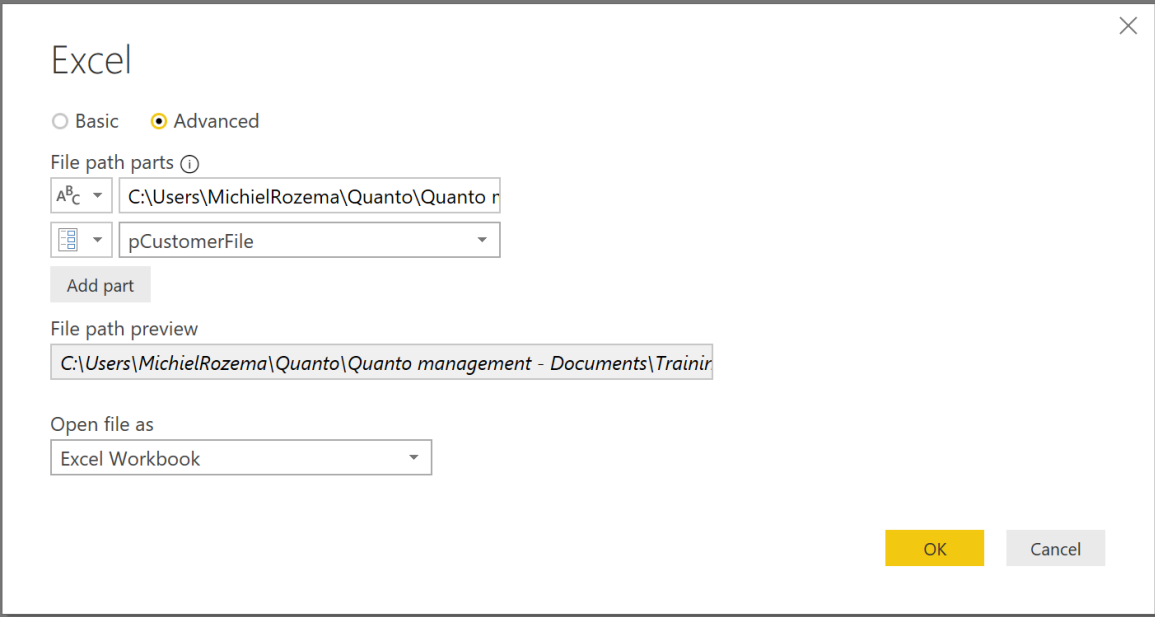
Suggested Values: Any value

Current Value: Customer.xlsx

OK Cancel

How to use a Parameter

- Many Power Query features allow for parameters, but not as the default
- For e.g. Excel files, load the file first, then click the gear icon in the first query step and select **Advanced**



Excel

Basic Advanced

File path parts ⓘ

A^BC ▾ C:\Users\MichielRozema\Quanto\Quanto r

▾ pCustomerFile ▾

Add part

File path preview

C:\Users\MichielRozema\Quanto\Quanto management - Documents\Trainir

Open file as

Excel Workbook ▾

OK Cancel

How to use a Parameter (2)

- Parameters can also be used:
 - In custom columns as a constant value
 - In M code

Custom Column ×

Add a column that is computed from the other columns.

New column name

Custom column formula ⓘ

Available columns

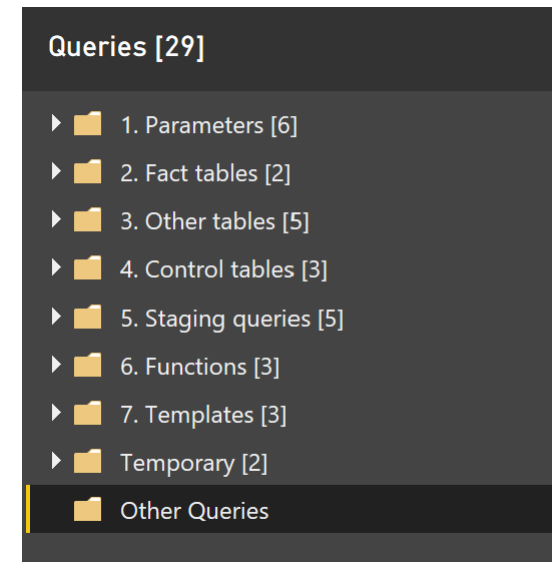
- EmpNr
- MngrNr
- Role
- SSN

Grouping queries

- Query groups help to keep track of all queries
- Right-click in the Queries pane and select **New Group...**
- After creating a group, any query is part of a group (default is *Other Queries*)
- Groups can contain other groups
- Groups can be moved – but *Other Queries* is always the last group

Best Practice: use a standard Group structure

- 1. Parameters
- 2. Fact tables
- 3. Other tables (*Filter tables*)
- 4. Control tables
- 5. Staging queries
- 6. Functions
- 7. Templates (*including folder query groups*)
- ...
- Other Queries





Exercise

<https://quanto.eu/trainingdata>

Exercises PQ 3



Introduction to M

The Advanced Editor

- Each query is really a script written in the M language
- To view the M code for each step, enable the **View – Formula Bar** option
- To view the complete script, click **Advanced Editor** in the ribbon
- Each feature in Power Query has its equivalent function in M – but there are more functions in M
- With M, you can create more advanced queries and better optimized queries than what the graphical editor offers
- To get a list of available functions, create a blank query and type **= #shared** in the formula bar

The M language

M script structure:

let

A = Function(<something>),

B = Function(A),

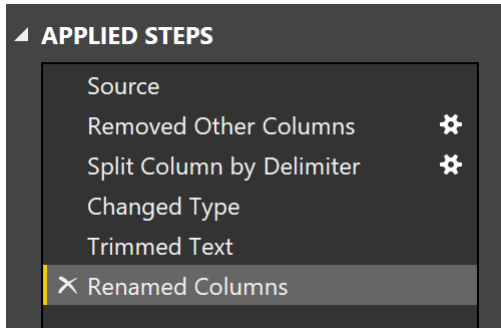
C = Function(B)

in

C

Warning: M is (annoyingly) case-sensitive!

M language: example



```
let
    Source = sCustomer,
    #"Removed Other Columns" = Table.SelectColumns(Source, {"Nr", "Customer", "Type", "City", "Sales", "From"}),
    #"Split Column by Delimiter" = Table.SplitColumn("#Removed Other Columns", "Type", Splitter.SplitTextByDelimiter(":", QuoteStyle.Csv), {"Type.1", "Type.2"}),
    #"Changed Type" = Table.TransformColumnTypes("#Split Column by Delimiter", {"Type.1", type text}, {"Type.2", type text}),
    #"Trimmed Text" = Table.TransformColumns("#Changed Type", {"Type.2", Text.Trim, type text}),
    #"Renamed Columns" = Table.RenameColumns("#Trimmed Text", {"Type.1", "Type"}, {"Type.2", "Category"})
in
    #"Renamed Columns"
```

What are Dataflows?

- Separation between data transformation logic and data model
 - Enables reuse of data transformation logic
- Like Power Query, but with intermediate storage for data
 - Azure Data Lake tables
- A dataflow is a group of entities
 - Entities are tables with data
 - Entities are created through queries