Patou Tips #24

Start with PowerBI

"From Rookie to Rock"

Quick Guide PowerQuery
Connection & normalisation
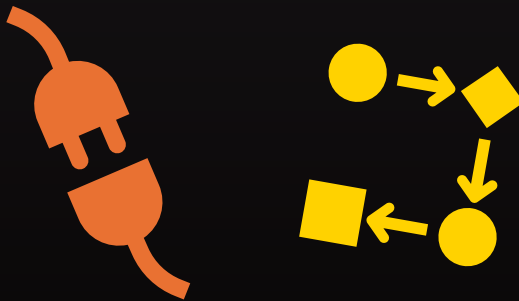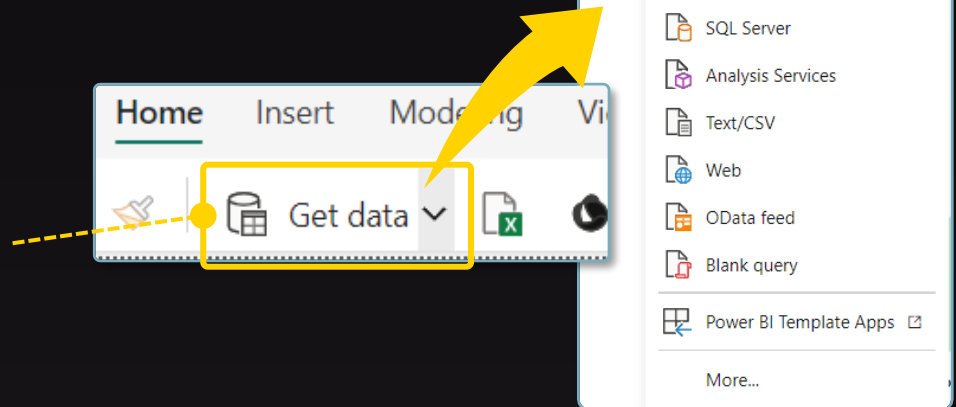
5 really useful connectors

There are so many possibilities to connect with data source types in PowerBI.

To get started...

✓ **Excel file**
✓ **CSV file**
✓ **Folder of files**
✓ **Web site source**
✓ **SQL source**

In the "home" menu...
Selecting a connector gives access to PowerQuery
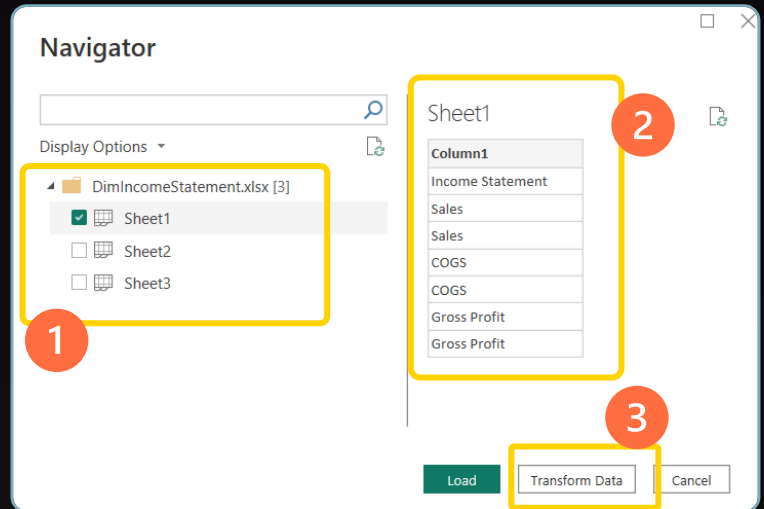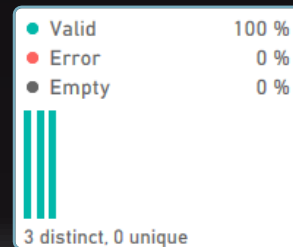
Select the connector « [Excel workbook] »

1 All sheets in the Excel file are displayed here. Select one (or more) sheet(s).

2 The contents of the selected sheet are displayed here.

💡 Patou Tip: The "transform data" button is really

3 interesting to check the quality of your data for yourself. See "Patou Tips #16 (step 2)" for more information.

Navigator

DimIncomeStatement.xlsx [3]
- ☑ Sheet1
- ☐ Sheet2
- ☐ Sheet3

Display Options ▾

1

Sheet1

Column1
Income Statement
Sales
Sales
COGS
COGS
Gross Profit
Gross Profit

2

Load   Transform Data   Cancel

3

● Valid	100 %
● Error	0 %
● Empty	0 %

3 distinct, 0 unique

Select the connector « 📄 Text/CSV »

products.csv

	File Origin	①	Delimiter	②	Data Type Detection		
	65001: Unicode (UTF-8) ▾		Comma ▾		Based on first 200 rows ▾		

productID	productName	quantityPerUnit	unitPrice	discontinued	categoryID
1	Chai	10 boxes x 20 bags	18	0	1
2	Chang	24 - 12 oz bottles	19	0	1
3	Aniseed Syrup	12 - 550 ml bottles	10	0	2
4	Chef Anton's Cajun Seasoning	48 - 6 oz jars	22	0	2
5	Chef Anton's Gumbo Mix	36 boxes	21.35	1	2
6	Grandma's Boysenberry Spread	12 - 8 oz jars	25	0	2
7	Uncle Bob's Organic Dried Pears	12 - 1 lb pkgs.	30	0	7
8	Northwoods Cranberry Sauce	12 - 12 oz jars	40	0	2
9	Mishi Kobe Niku	18 - 500 g pkgs.	97	1	6

③

Extract Table Using Examples   Load   **Transform Data**   Cancel

① **Patou Tip:** Use the format "65001: Unicode (UTF-8)", when you have data with accentuation like French, Spanish languages...

② Select the delimiter (could be a coma "," or a semicolon ";")

3 **Connection with a folder of files (1/2)**

Select the connector « 📁 Folder **»**
Note: This connector works with many file formats; Excel, CSV, PDF…

1 Select your folder

Folder

1

Folder path

C:\Users\Personnel\OneDrive\Documents\Patou Tips (Start PowerBI)   Browse...

2 The contents of the folder (files) is displayed here.

C:\Users\Personnel\OneDrive\Documents\Patou Tips\Ch...

Content	Name	Extension	Date accessed	Date modified	Date created
Binary	orders.csv	.csv	17/11/2024 18:07:42	08/05/2023 15:12:56	17/11/2024 18
Binary	order_details.csv	.csv	17/11/2024 18:07:42	08/05/2023 15:12:56	17/11/2024 18
Binary	products.csv	.csv	17/11/2024 17:28:07	08/05/2023 15:12:56	17/11/2024 17
Binary	shippers.csv	.csv	17/11/2024 18:07:42	08/05/2023 15:12:56	17/11/2024 18

2

3 Combine ▾ | Load | Transform Data | Cancel

3 Select the "combine and transform data" button

Combine ▾ | Load | Trans

Combine & Transform Data
Combine & Load

Connection with data

1 See tip #2, "Connection with CSV file"

Combine Files

Specify the settings for each file. Learn more

Sa___le:

File Origin
1252: Western European (Windows)

Delimiter
Comma

Data Type Detection
Based on first 200 rows

orderID	customerID	employeeID	orderDate	requiredDate	shippedDate	shipperID	freight
10248	VINET	5	2013-07-04	2013-08-01	2013-07-16	3	32.38
10249	TOMSP	6	2013-07-05	2013-08-16	2013-07-10	1	11.61
10250	HANAR	4	2013-07-08	2013-08-05	2013-07-12	2	65.83
10251	VICTE	3	2013-07-08	2013-08-05	2013-07-15	1	41.34
10252	SUPRD	4	2013-07-09	2013-08-06	2013-07-11	2	51.3

OK   Cancel

Patou Tips #22

Start with PowerBI
"From Rookie to Rock"

PowerBI projects with **Teams**

Store data → T → Connect data
Publish report

Discover "Patou Tips #22" to connect PowerBi with Teams (or Sharepoint)

Start with PowerBI
"From Rookie to Rock"

Connection with data

Select the connector « 🌐 Web »

From Web

○ Basic ○ Advanced

URL

https://editorial.rottentomatoes.com/guide/best-movies-of-all-time/

1 Paste the Web site address. Here a Web site with a list of "300 best movies of all the time".

Navigator

Table View | Web View

Display Options ▾

📁 HTML Tables [3]
   ☑ ▦ Table 1
   ☐ ▦ Table 2
   ☐ ▦ Table 3
▾ 📁 Suggested Tables [4]
   ☐ ▦ Table 4
   ☐ ▦ Table 5
   ☐ ▦ Table 6
   ☐ ▦ Table 7
▾ 📁 Text [2]
   ☐ ▦ HTML Code
   ☐ ▦ Displayed Text
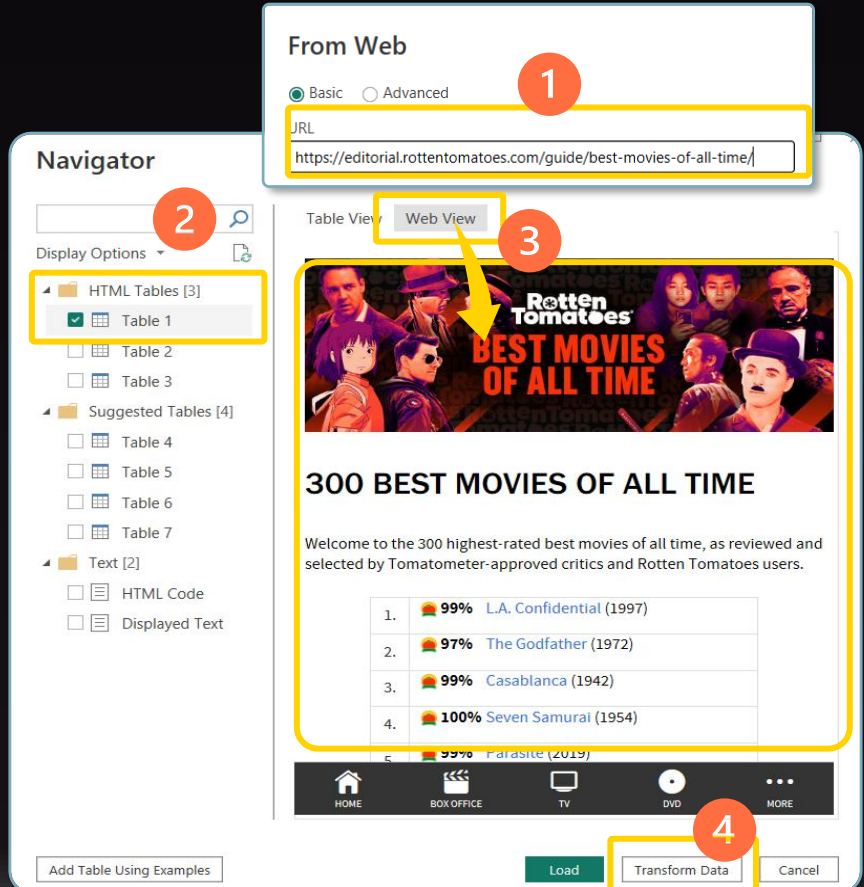
2 Select the object to display

Rotten Tomatoes

BEST MOVIES OF ALL TIME

300 BEST MOVIES OF ALL TIME

Welcome to the 300 highest-rated best movies of all time, as reviewed and selected by Tomatometer-approved critics and Rotten Tomatoes users.

1. 🍅 **99%** L.A. Confidential (1997)
2. 🍅 **97%** The Godfather (1972)
3. 🍅 **99%** Casablanca (1942)
4. 🍅 **100%** Seven Samurai (1954)
5. 🍅 **99%** Parasite (2019)

HOME | BOX OFFICE | TV | DVD | MORE

3 The contents of selected object is displayed here. Great it's the list that we wanted!

4 Select the "transform data" button

Add Table Using Examples | Load | Transform Data | Cancel

Connection with data

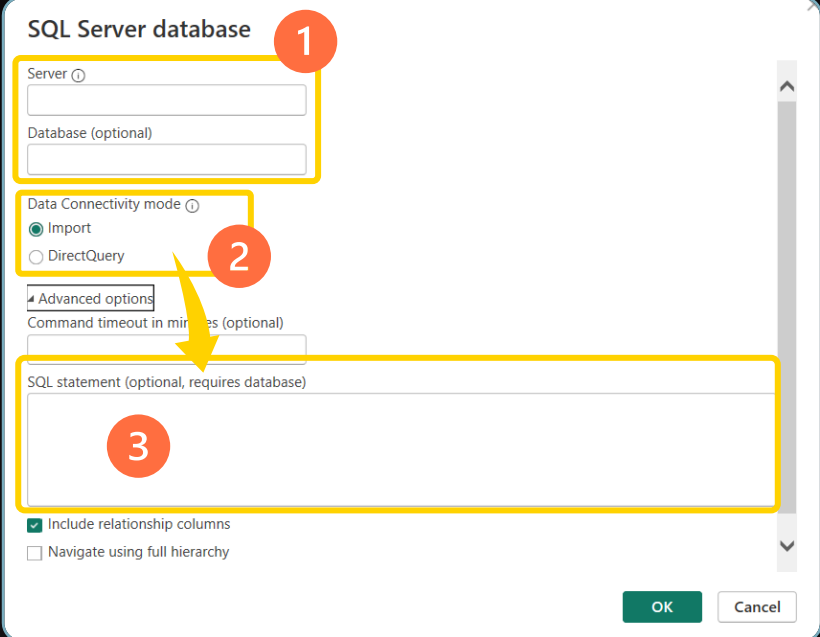Select the connector « [SQL Server] »

1 Specify the SQL Server address

2 Select the connection mode. Select "import" to always keep the data in PowerBI or "DirectQuery" for continuous refresh.

3 It is better to create a SQL statement to limit each query.

SQL Server database

1

Server ⓘ

Database (optional)

Data Connectivity mode ⓘ
● Import
○ DirectQuery

2

▲ Advanced options
Command timeout in minutes (optional)

SQL statement (optional, requires database)

3

☑ Include relationship columns
☐ Navigate using full hierarchy

OK   Cancel

To go further...

6 great ressources!

Wise Cat Noro Boris
@wisecatnoroboris · 21,1 k abonnés · 481 vidéos
Hello, content de vous voir par ici. ...plus
wisecat.tv et 2 autres liens
S'abonner

Access Analytic
@AccessAnalytic · 95,6 k abonnés · 275 vidéos
Amazing Excel and Power BI Solutions. Perth,
accessanalytic.com.au et 4 autres liens
S'abonner

Curbal •
@CurbalEN · 139 k abonnés · 1 k vidéos
Curb your Power BI skils! ...plus
curbal.com et 1 autre lien
S'abonner   Rejoindre

Goodly
@GoodlyChandeep · 99,9 k abonnés · 563
A lot of people think that my name is Good
goodly.co.in/learn-m-powerquery et 4 autr
S'abonner   Rejoindre

Le CFO masqué
@lecfomasque · 5,16 k abonnés · 269 vidéos
La référence francophone en informatique d
lecfomasque.com et 3 autres liens
S'abonner

Dataseito
@Dataseito · 20,2 k abonnés · 80 vidéos
Je partage mes expériences dans le mon
dataseito.substack.com et 2 autres liens
S'abonner
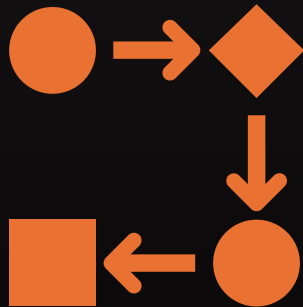
Start with PowerBI
"From Rookie to Rock"

Connection with data
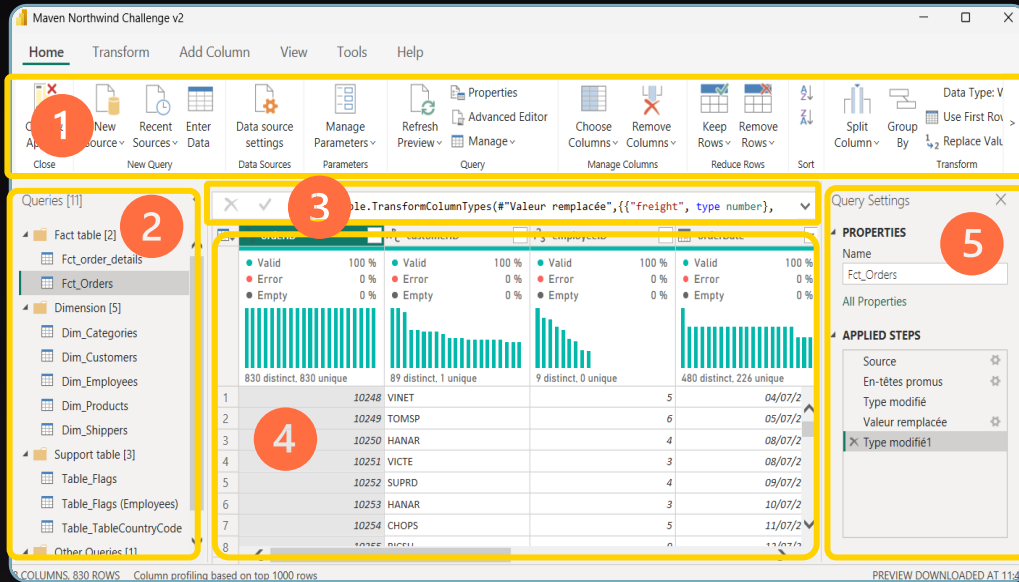
Start with PowerBI

"From Rookie to Rock"

Data normalization

About this step...

After connecting the data, to prepare and clean the project data, we start again in PowerQuery...

PowerQuery is a powerful ETL (Treat, Transform, Load) inside PowerBI (and Excel)

1	Functions pane	
2	Query pane	
3	Formula window	
4	Data window	
5	Transformation history	

Start with PowerBI

"From Rookie to Rock"

Data normalization

5 data normalization rules

💡 Before preparing your data in Power Query, it is necessary **to understand the business needs and the business codes of your future users** (HR, finance, sales, etc.).

✓ 1 → Understand the data profiling and use the right click

✓ 2 → Select necessary columns and remove duplicate rows

✓ 3 → Rename columns and change data types

✓ 4 → Replace Null values and filter data

✓ 5 → Rename and organize your queries

💡 « Power Patou Tip » → **Create the « N value »**

Start with PowerBI

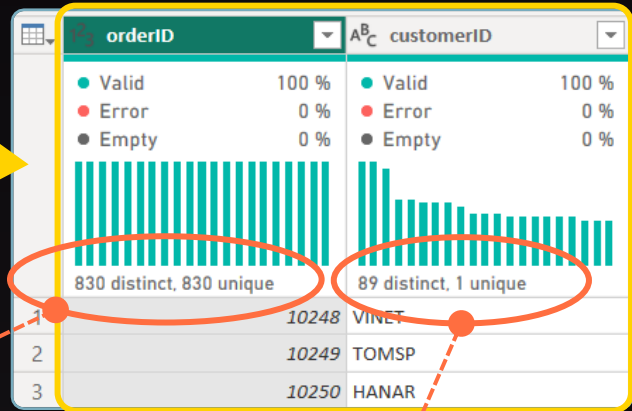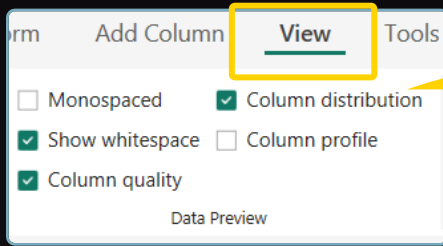"From Rookie to Rock"

Data normalization

Data profiling provides you with important information about data quality and distribution.

On View Menu

rm	Add Column	View	Tools
☐ Monospaced	☑ Column distribution		
☑ Show whitespace	☐ Column profile		
☑ Column quality			

Data Preview

⊞ ▾	⟨⟩ orderID ▾	ABC customerID ▾
	● Valid 100 %	● Valid 100 %
	● Error 0 %	● Error 0 %
	● Empty 0 %	● Empty 0 %
	830 distinct, 830 unique	89 distinct, 1 unique
1	10248	VINET
2	10249	TOMSP
3	10250	HANAR

830 distinct, 830 unique

830 different values and 830 values appear once, so there are no duplicate values. In this example, this means that the order IDs are unique, great, that's what we want!

89 distinct, 1 unique

89 different values and 1 value only appears once. In this example, this means that 88 customers ordered multiple times.

Data normalization

Right-clicking on the upper left corner or on a column name gives you **access to almost all the tools useful** for cleaning and transforming data.

Right-clicking on the upper left corner

	Copy Entire Table
	Use First Row as Headers
	Add Custom Column...
	Add Column From Examples...
	Invoke Custom Function...
	Add Conditional Column...
	Add Index Column
	Choose Columns...
	Keep Top Rows...
	Keep Bottom Rows...
	Keep Range of Rows...
	Keep Duplicates
	Keep Errors
	Remove Top Rows...
	Remove Bottom Rows...
	Remove Alternate Rows...
	Remove Duplicates
	Remove Errors
	Merge Queries...
	Append Queries...

Right-clicking on a column name

customerID

Copy
Remove
Remove Other Columns
Duplicate Column
Add Column From Examples...
Remove Duplicates
Remove Errors
Change Type
Transform
Replace Values...
Replace Errors...
Split Column
Group By...
Fill
Unpivot Columns
Unpivot Other Columns
Unpivot Only Selected Columns
Rename...
Move
Drill Down
Add as New Query

VINET
TOMSP
HANAR
VICTE
SUPRD
HANAR
CHOPS
RICSU
WELLI
HILAA
ERNSH
CENTC
OTTIK
QUEDE
RATTC

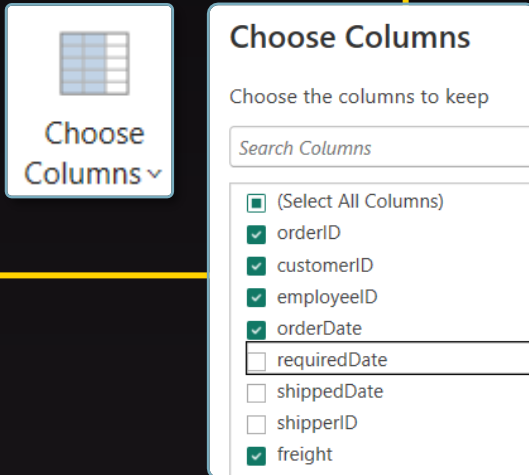89 distinct, 1 unique

Data normalization

2 **Select necessary columns and remove duplicate rows**

These two actions help to keep the necessary data and remove duplicate rows. By reducing the number of columns and rows you will **increase the performance of your project.**
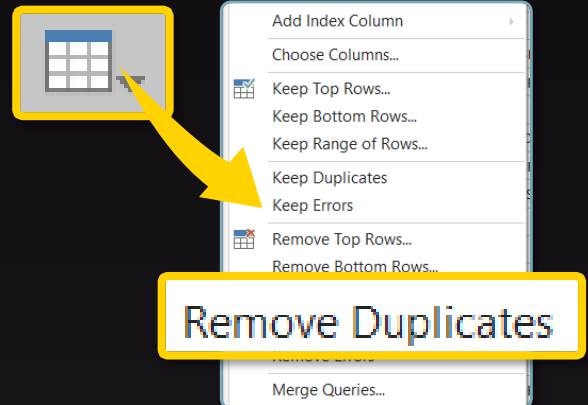
Select necessary columns

On Home Menu

Choose Columns

Choose the columns to keep

Search Columns

- ☑ (Select All Columns)
- ☑ orderID
- ☑ customerID
- ☑ employeeID
- ☑ orderDate
- ☐ requiredDate
- ☐ shippedDate
- ☐ shipperID
- ☑ freight

Choose Columns ⌄

Remove duplicate rows

Right-clicking on the upper left corner

Add Index Column
Choose Columns...
Keep Top Rows...
Keep Bottom Rows...
Keep Range of Rows...
Keep Duplicates
Keep Errors
Remove Top Rows...
Remove Bottom Rows...

Remove Duplicates

Merge Queries...

Data normalization
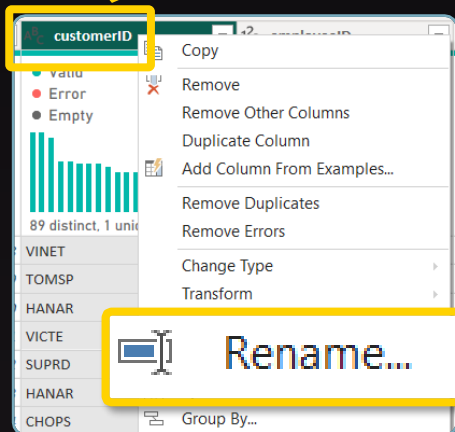
These two actions allow you to name the data in the business sense and to assign them the appropriate format to **facilitate data manipulation.**

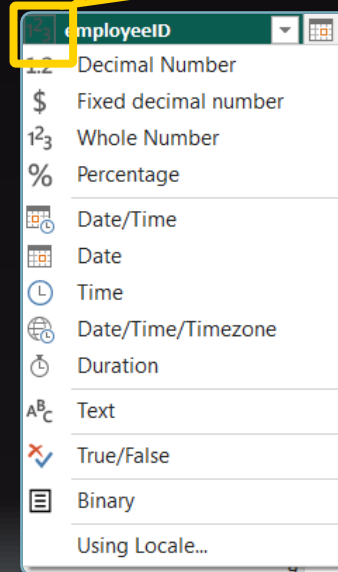### Rename columns

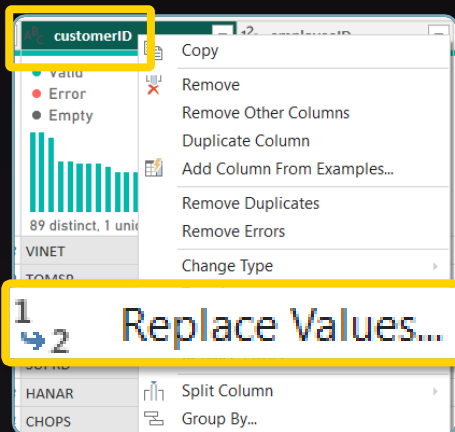**Right-clicking** on a column name

Change data types

Click on the logo on the left of the column name

These two actions allow you to reduce the number of data by their value and **improve the performance of your project.**

Replace Null values

Right-clicking on a column name

Number values

Replace Values

Replace one value with another in the selec

Value To Find

null

Replace With

0

customerID
• Valid
• Error
• Empty
89 distinct, 1 uniq
VINET
TOMSP

Copy
Remove
Remove Other Columns
Duplicate Column
Add Column From Examples…
Remove Duplicates
Remove Errors
Change Type
1→2 **Replace Values…**
Split Column
Group By…

Text values

Replace Values

Replace one value with another in the selec

Value To Find

null

Replace With

ND ND = Not Determinated

HANAR
CHOPS

Filter data

Click on the logo on the right of the column name

$^{A^B_C}$ customerID ▼
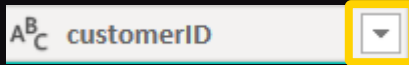
Text filters

Equals...
Does Not Equal...

Greater Than...
Greater Than Or Equal To...
Less Than...
Less Than Or Equal To...

Between...

Number filters

Equals...
Does Not Equal...

Begins With...
Does Not Begin With...
Ends With...
Does Not End With...

Contains...
Does Not Contain...

Date filters

Equals...

Before...
After...
Between...

In the Next...
In the Previous...

Is Earliest
Is Latest

Is Not Earliest
Is Not Latest

Year   ►
Quarter   ►
Month   ►
Week   ►
Day   ►

Rename queries

Right-clicking on a query name

```
Queries [11]                          <
▲ 📁 Fact table [2]
     ▦  Fct_order_details
     ▦  Fct_Orders
▲ 📁 Dimension [5]
     ▦  Dim_Categories
     ▦  Dim_Customers
     ▦  Dim_Employees
     ▦  Dim_Products
     ▦  Dim_Shippers
▲ 📁 Support table [3]
     ▦  Table_Flags
     ▦  Table_Flags (Employees)
```
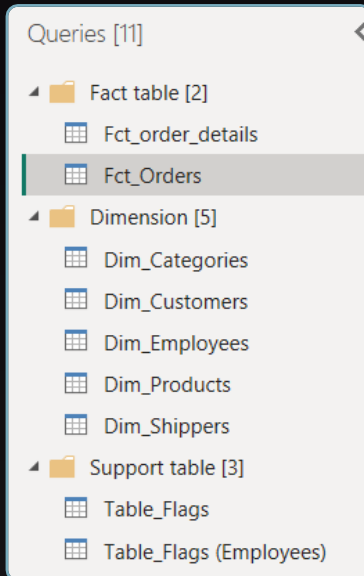
Fct = Fact
Dim = Dimension

Organize queries

Right-clicking on a query name
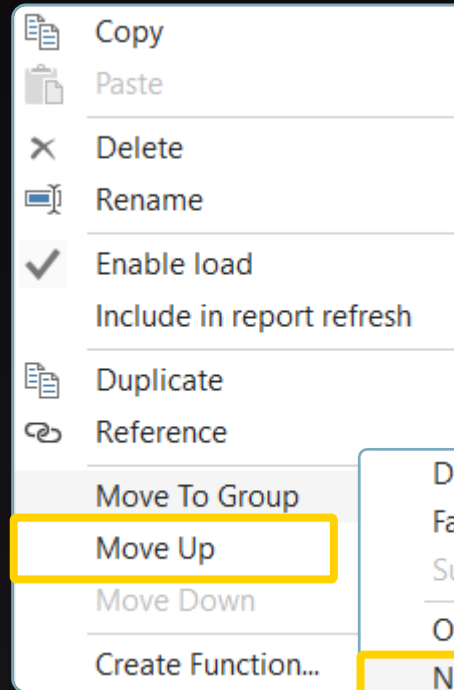
```
📋  Copy
📋  Paste
✕   Delete
▥   Rename
✓   Enable load
    Include in report refresh
📋  Duplicate
🔗  Reference
    Move To Group          Dimension
    Move Up                Fact table
    Move Down              Support table
    Create Function...     Other Queries
                           New Group...
```

Data normalization

This technique allows you to group all the variables of type "fact table" in the same field. It will be **useful and easier to create measures in PowerBi** with the Dax language.

Variables

Unpivot Only Selected Columns

Start with PowerBI

"From Rookie to Rock"

Data normalization

To go further...

6 great ressources!

RADACAD
@RADACAD · 49,6 k abonnés · 568 vidéos

RADACAD is all about helping YOU to get more i

radacad.com et 6 autres liens

S'abonner

Access Analytic
@AccessAnalytic · 95,6 k abonnés · 275 vidéos

Amazing Excel and Power BI Solutions. Perth,

accessanalytic.com.au et 4 autres liens

S'abonner

Augustin de la Fouchardiere · 1er
Fondateur de MYPE, Consultant & formateur Data
Paris

Expérience : MYPE, ESSEC Business School et 7 autre(s)

14 relations en commun

Message **Voir le profil complet**

Clément Pruvost · 1er
Responsable Data - Ofilae Family Office
Greater Lille Metropolitan Area

8 k abonnés

14 relations en commun

Message **Voir le profil complet**

Enzo's Data Expedi
@EnzoDataExpedition · 210 abonnés · 11 vidéos

Bienvenue sur Enzo's Data Expedition, la chaîne Y

linkedin.com/in/enzo-rideau

S'abonner

Melissa de Korte in · 2e
Power Platform Enthusiast | Microsoft MVP | Super User | Enterp
The Randstad, Netherlands

8 k abonnés

23 relations en commun

Message **Voir le profil complet**

Start with PowerBI

"From Rookie to Rock"

Data normalization

Don't forget!
This isn't the truth, it's just my truth!

Patou Tips

Follow me
Like me
Share me