

Power BI - Advance Data Transformation

PIVOTING & UNPIVOTING

Pivoting describes the process of turning distinct row values into columns, and unpivoting describes the process of turning distinct columns into rows

The diagram illustrates the relationship between two tables. The top table is a pivoted version of the bottom table. The bottom table has three columns: Date, Product Category, and Region. The top table has four columns: Date, Product Category, North Region, Central Region, and South Region. Arrows labeled "PIVOTING" point from the bottom table to the top table, and arrows labeled "UNPIVOTING" point from the top table back to the bottom table.

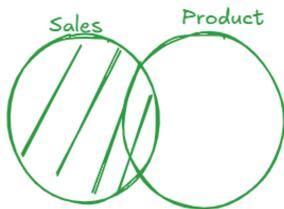
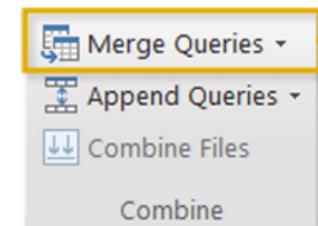
| | Date | A% Product Category | 1 ² 3 North Region | 1 ² 3 Central Region | 1 ² 3 South Region |
|---|----------|---------------------|-------------------------------|---------------------------------|-------------------------------|
| 1 | 7/1/2022 | Bikes | 10 | 19 | 25 |
| 2 | 7/1/2022 | Components | 14 | 31 | 16 |
| 3 | 7/1/2022 | Clothing | 35 | 32 | 46 |

| | Date | A% Product Category | A% Region | 1 ² 3 Quantity Sold |
|---|----------|---------------------|----------------|--------------------------------|
| 1 | 7/1/2022 | Bikes | North Region | 10 |
| 2 | 7/1/2022 | Bikes | Central Region | 19 |
| 3 | 7/1/2022 | Bikes | South Region | 25 |
| 4 | 7/1/2022 | Components | North Region | 14 |
| 5 | 7/1/2022 | Components | Central Region | 31 |
| 6 | 7/1/2022 | Components | South Region | 16 |
| 7 | 7/1/2022 | Clothing | North Region | 35 |
| 8 | 7/1/2022 | Clothing | Central Region | 32 |
| 9 | 7/1/2022 | Clothing | South Region | 46 |

Imagine the table on a hinge; pivoting rotates it from vertical to horizontal, and unpivoting rotates it from horizontal to vertical

NOTE: Transpose works very similarly, but doesn't recognize unique values; instead, the entire table is transformed so that each row becomes a column and vice versa

MERGING QUERIES



Left Join - All the final result would be coming up in left table.

Merge

Select a table and matching columns to create a merged table.

Sales Data

| Order Date | Product Key | Customer Key | Order Quantity | Index | Stock Date | Order Number | Territory |
|------------|-------------|--------------|----------------|-------|------------|--------------|-----------|
| 6/25/2022 | 214 | 14719 | 1 | 55115 | 4/20/2022 | SO73780 | |
| 10/8/2021 | 214 | 21990 | 1 | 14247 | 7/2/2021 | SO55746 | |
| 12/30/2021 | 214 | 22098 | 1 | 26322 | 11/10/2021 | SO61052 | |
| 6/29/2022 | 214 | 22748 | 1 | 55740 | 4/9/2022 | SO74069 | |

Product Lookup

| Product Key | Product Subcategory Key | Product S K U | Product Name | Model Name | |
|-------------|-------------------------|---------------|-------------------------|---------------------|------------------|
| 214 | 31 | HL-US09-R | Sport-100 Helmet, Red | Sport-100 | Universal fit, v |
| 215 | 31 | HL-US09 | Sport-100 Helmet, Black | Sport-100 | Universal fit, v |
| 218 | 23 | SO-B909-M | Mountain Bike Socks, M | Mountain Bike Socks | Combination c |
| 219 | 23 | SO-B909-L | Mountain Bike Socks, L | Mountain Bike Socks | Combination c |

Join Kind
Left Outer (all from first, matching from second)

Use fuzzy matching to perform the merge

> Fuzzy matching options

✓ The selection matches 56046 of 56046 rows from the first table.

OK Cancel

Merging queries allows you to join tables based on a common column (like a lookup in Excel)

In this case we're merging the Sales Data table with the Product Lookup table, which share a common Product Key column

NOTE: Merging adds columns to an existing table/query

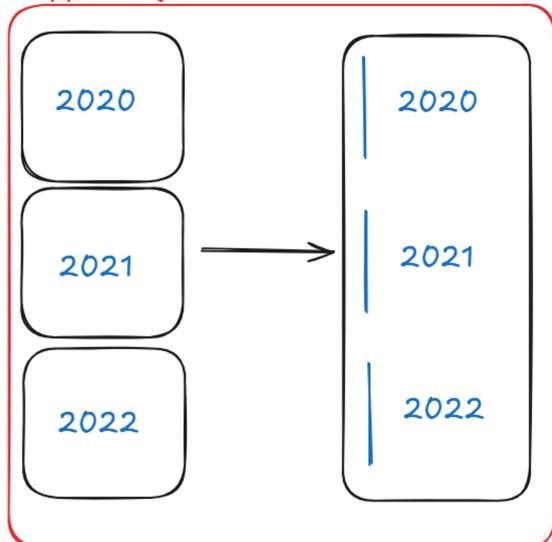
Just because you can merge tables, doesn't mean you should!

In many cases, it's better to keep tables separate and define relationships between them in the data model (more on that soon!)

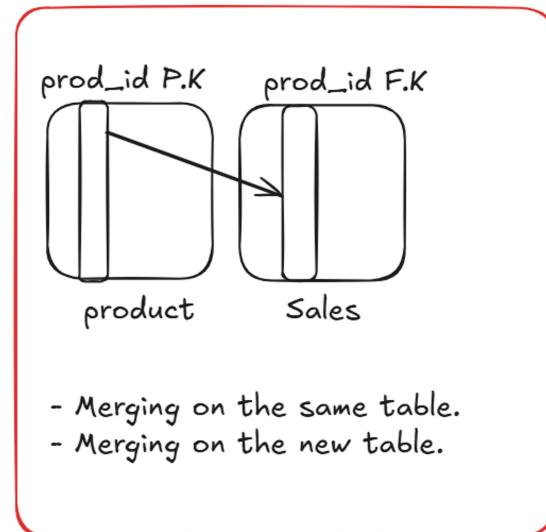
Table.ColumnName

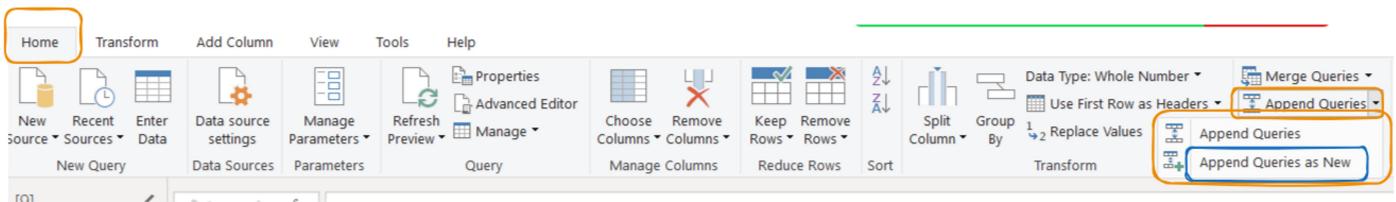
Append Vs Merging

Appending - Vertical Direction



Merge - Horizontal Direction with the help of P.K & F.K





Append

Concatenate rows from two tables into a single table.

Two tables Three or more tables

First table

Sales Data 2020

Second table

OK

Cancel

Append

Concatenate rows from three or more tables into a single table.

Two tables Three or more tables

Available tables

Calendar Lookup
Product Categories Lookup
Product Subcategories Lookup
Product Lookup
Customer Lookup
Sales Data 2020
Sales Data 2021
Sales Data 2022
Product Category Sales (Unpivot Demo)

Tables to append

Sales Data 2020
Sales Data 2021
Sales Data 2022

Add >>

Sales Data 2020
Sales Data 2021
Sales Data 2022
Product Category Sales (...)
Append1

OK

Cancel

Queries [10]

| Index | OrderDate | StockDate | Avg OrderNumber | ProductKey | CustomerKey | TerritoryKey |
|-------|------------|------------|-----------------|------------|-------------|--------------|
| 1 | 01-01-2020 | 21-09-2019 | SO45080 | 332 | 14657 | |
| 2 | 01-01-2020 | 05-12-2019 | SO45079 | 312 | 29255 | |
| 3 | 01-01-2020 | 29-10-2019 | SO45082 | 350 | 11455 | |
| 4 | 01-01-2020 | 16-11-2019 | SO45081 | 338 | 26782 | |
| 5 | 02-01-2020 | 15-12-2019 | SO45083 | 312 | 14947 | |
| 6 | 02-01-2020 | 12-10-2019 | SO45084 | 310 | 29143 | |

Query Settings

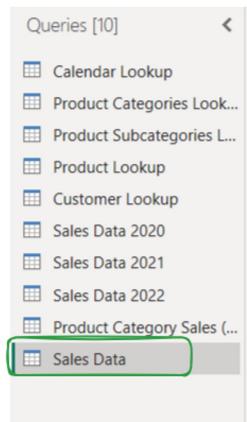
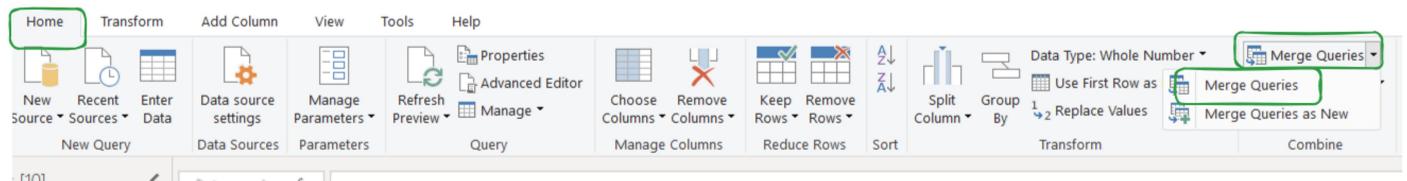
PROPERTIES
Name: Sales Data
All Properties

APPLIED STEPS
Source

$P.\text{Price} * \text{Quantity} = \text{Sales}$
 $P.\text{Cost} * \text{Quantity} = \text{Cost to company.}$
 $\text{Profit} = \text{Sales} - \text{Cost To Company.}$
 $\text{SUM of Sales} = \text{Revenue.}$

Merging Tables

We have to bring Product Cost & Product Price from Product Lookup that merges with Sales Transactional Record.



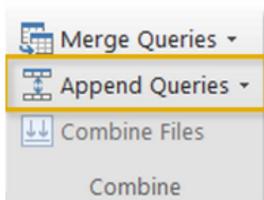
This represent we wanted to merge new columns into Sales Data.

The screenshot shows the 'Merge' dialog box. It displays two tables: 'Sales Data' and 'Product Lookup'. The 'Sales Data' table has columns: Index, OrderDate, StockDate, OrderNumber, ProductKey, CustomerKey, TerritoryKey, and OrderLineItem. The 'Product Lookup' table has columns: ProductKey, ProductSubcategoryKey, ProductSKU, ProductName, ModelName, and several rows of data. A green box highlights the 'ProductKey' column in both tables. The 'Merge' settings at the bottom include: 'Join Kind: Left Outer (all from first, matching from second)', 'Use fuzzy matching to perform the merge', 'Fuzzy matching options', and a note: 'The selection matches 45303 of 45303 rows from the first table.' The 'OK' button is also highlighted with a green box.

The screenshot shows the 'Merge Queries' dialog box. It displays the 'Product Lookup' table with columns: ProductKey, ProductSubcategoryKey, ProductSKU, ProductName, ModelName, and several rows of data. A green box highlights the 'ProductKey' column. On the right, there is a list of columns to expand: 'Search Columns to Expand' (Quantity Type), 'Expand' (radio button selected), 'Aggregate' (radio button), and a list of columns: '(Select All Columns)', 'ProductKey', 'ProductSubcategoryKey', 'ProductSKU', 'ProductName', 'ModelName', 'ProductDescription', 'ProductColor', 'ProductSize', 'ProductStyle', 'ProductCost' (checkbox checked), 'ProductPrice' (checkbox checked), 'SKU type', and 'SKU Category'. The 'OK' button is highlighted with a green box.

Rename the column

APPENDING QUERIES



Appending queries allows you to combine or stack tables sharing the exact same column structure and data types

Append

Concatenate rows from two tables into a single table.

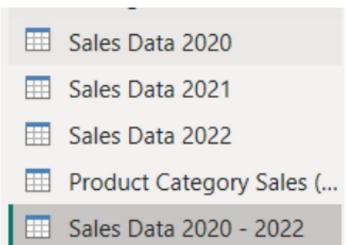
- Two tables
- Three or more tables

First table

AdventureWorks Sales Data 2020

Here we're appending the Adventure Works Sales 2020 table to the Adventure Works Sales 2021 table, which is valid since they share identical table structures.

NOTE: Appending adds rows to an existing table/query



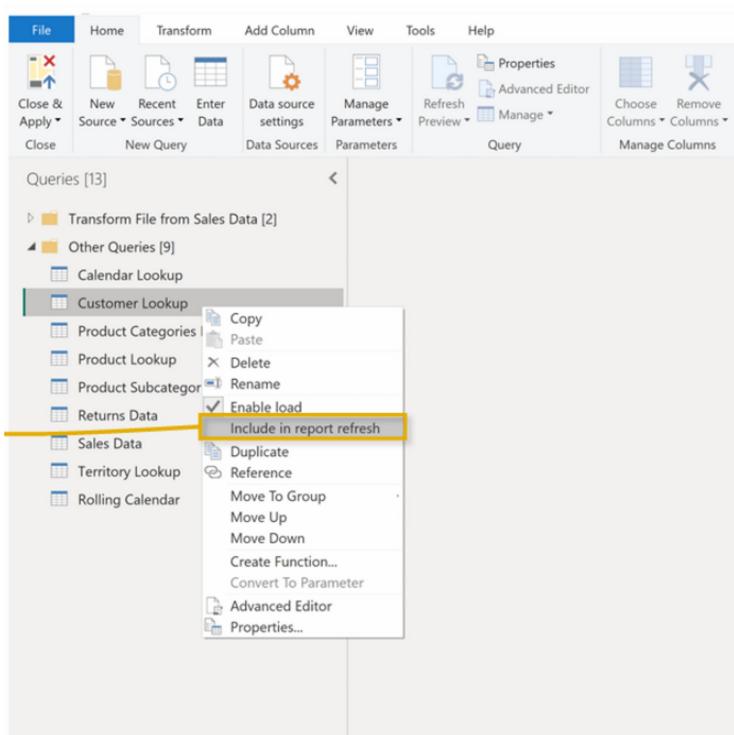
REFRESHING QUERIES



- By default, all queries will refresh when you use the Refresh command from the Home tab

From the Query Editor, uncheck Include in report refresh to exclude individual queries from the refresh

PRO TIP: Exclude queries from refresh that don't change often (like lookups or static data tables)



POWER QUERY BEST PRACTICES

Get organized before connecting and loading data

- Define clear and intuitive table/query names from the start, and establish an organized file/folder structure if you are working with local flat files to avoid changes to file names or paths.

Disable report refresh for any static data sources

- There's no need to constantly refresh data sources that don't change, like lookups or static data tables

When working with large tables, only load the data you need

- Don't include hourly data when you only need daily, or transaction-level data when only need a product-level summary (extra data will only slow your report down!)

Time Complexity
Space Complexity