

Sources Used

1. Web Source (URL):

Sample structured data was sourced from a publicly available web page for learning and practice purposes:

[GDP by Country \(2025\) - Worldometer](#)

2. Generated Data Source:

Additional datasets (Sales data, Employee data, Monthly Excel files) were synthetically generated using ChatGPT exclusively for educational and project simulation purposes.

The data does not represent any real organization or individual and was created to demonstrate Power BI Power Query concepts such as data extraction, transformation, merging, appending, pivoting/unpivoting, and refresh simulation.

Transformation applied

The screenshot displays two Power BI data tables with their corresponding transformation settings.

Table 1: Monthly Sales (Generated Data Source)

This table contains sales data for various employees across different products and dates. The transformation steps include:

- Source
- Removed Other Columns
- Added Custom
- Expanded Custom
- Expanded Custom1
- Removed Blank Rows
- Removed Duplicates
- Removed Columns
- Promoted Headers
- Changed Type
- Trimmed Text
- Cleaned Text
- Split Column by Delimiter
- Changed Type1
- Renamed Columns
- Changed Type2
- Rounded Off
- Rounded Off1
- Added Custom1
- Added Custom2
- Changed Type3

	EmployeeID	SalesPerson Firstname	SalesPerson Lastname	OrderID	OrderDate	ProductName
1	1001	Vikas	Malhotra	50001	4/1/2023	Mobile
2	1002	Neha	Patel	50002	5/18/2023	Tablet
3	1003	Neha	Patel	50003	12/9/2023	Laptop
4	1004	Kiran	Rao	50004	3/25/2023	Keyboard
5	1005	Suresh	Iyer	50005	8/13/2023	Monitor
6	1006	Amit	Sharma	50006	8/3/2023	Monitor
7	1007	Pooja	Singh	50007	2/2/2023	Tablet
8	1008	Pooja	Singh	50008	6/19/2023	Tablet
9	1009	Anjali	Gupta	50009	8/21/2023	Keyboard
10	1010	Pooja	Singh	50010	3/14/2023	Keyboard
11	1011	Rahul	Verma	50011	9/14/2023	Printer
12	1012	Suresh	Iyer	50012	8/26/2023	Printer
13	1013	Anjali	Gupta	50013	7/15/2023	Laptop
14	1014	Neha	Patel	50014	1/24/2023	Mobile
15	1015	Pooja	Singh	50015	3/4/2023	Keyboard
16	1016	Suresh	Iyer	50016	1/3/2023	Printer
17	1017	Anjali	Gupta	50017	10/18/2023	Keyboard
18	1018	Vikas	Malhotra	50018	6/17/2023	Laptop
19	1019	Rahul	Verma	50019	2/14/2023	Printer

Table 2: GDP by Country (2025) (Web Source)

This table contains GDP data for various countries. The transformation steps include:

- Source
- Extracted Table From Html
- Promoted Headers
- Changed Type
- Removed Blank Rows
- Removed Duplicates
- Removed Errors

#	Country	GDP	\$ GDP (Full Value)	GDP Growth	\$ GDP per Capita
1	United States	\$30.62 trillion	30,615,743,000,000	2%	89,599.00
2	China	\$19.4 trillion	19,398,577,000,000	4.8%	13,806.00
3	Germany	\$5.01 trillion	5,013,574,000,000	0.2%	59,925.00
4	Japan	\$4.28 trillion	4,279,826,000,000	1.1%	34,713.00
5	India	\$4.13 trillion	4,125,213,000,000	6.6%	2,818.00
6	United Kingdom	\$3.96 trillion	3,958,780,000,000	1.3%	56,661.00
7	France	\$3.36 trillion	3,361,557,000,000	0.7%	48,982.00
8	Italy	\$2.54 trillion	2,543,677,000,000	0.5%	43,161.00
9	Russia	\$2.54 trillion	2,540,656,000,000	0.6%	17,446.00
10	Canada	\$2.28 trillion	2,283,599,000,000	1.2%	54,935.00
11	Brazil	\$2.26 trillion	2,256,910,000,000	2.4%	10,578.00
12	Spain	\$1.89 trillion	1,891,371,000,000	2.9%	38,040.00
13	Mexico	\$1.86 trillion	1,862,740,000,000	1%	13,967.00
14	South Korea	\$1.86 trillion	1,858,572,000,000	0.9%	35,962.00
15	Australia	\$1.83 trillion	1,829,508,000,000	1.8%	65,946.00
16	Turkey	\$1.57 trillion	1,565,471,000,000	3.5%	18,198.00
17	Indonesia	\$1.44 trillion	1,443,256,000,000	4.9%	5,074.00
18	Netherlands	\$1.32 trillion	1,320,635,000,000	1.4%	73,174.00
19	Saudi Arabia	\$1.27 trillion	1,268,535,000,000	4%	35,231.00
20	Poland	\$1.04 trillion	1,039,619,000,000	3.2%	28,485.00

The screenshot displays two Power Query Editor windows side-by-side.

Top Window:

```
= Table.ReorderColumns(#"Filtered Rows", {"EmployeeID", "SalesPerson Firstname", "SalesPerson Lastname", "OrderID", "OrderDate", "ProductName"} )
```

Bottom Window:

```
= Table.Group(#"Added Index1", {"Department"}, {{"Total Salary", each List.Sum([Salary]), type nullable number}})
```

Properties and Applied Steps:

- Top Window Properties:** Name: Monthly Sales, All Properties
- Top Window Applied Steps:** Reordered Columns
- Bottom Window Properties:** Name: Employees, All Properties
- Bottom Window Applied Steps:** Grouped Rows

Challenges Faced & How They Were Solved

1. Combining Multiple Excel Files from a Folder

Challenge:

While loading multiple monthly Excel files from a folder, issues such as column order mismatch and data duplication were encountered.

Solution:

The Get Data → Folder → Combine & Transform option in Power Query was used.

A sample file structure was standardized, and Append Queries was applied to combine all files into a single consolidated dataset.

2. Incorrect Data Types

Challenge:

Several date and numeric columns were incorrectly stored as text, causing errors during calculations.

Solution:

The Change Data Type and Change Type with Locale options in Power Query were used to convert columns into appropriate date and numeric formats.

3. Pivot and Unpivot Column Misalignment

Challenge:

After performing pivot and unpivot operations, column positions changed and caused confusion in the data structure.

Solution:

Columns were reordered after unpivoting, and meaningful column names were assigned to maintain a consistent and readable structure.

4. Creating Calculated and Conditional Columns

Challenge:

Creating calculated fields such as Profit and categorizing sales into High, Medium, and Low segments was initially complex.

Solution:

Custom Columns and Conditional Columns were created in Power Query using defined business rules

(e.g., Profit = Revenue – Cost).

5. Refresh Simulation with New Files

Challenge:

Ensuring that newly added Excel files in the folder automatically followed the same transformation logic during refresh.

Solution:

A dynamic folder connection was maintained with a consistent file structure.

On refresh, Power Query automatically applied all transformations to the new files.

6. Source Settings and Parameter Management

Challenge:

Changes in data source paths caused query refresh failures.

Solution:

Parameters were used to make folder paths dynamic, and Data Source Settings were properly configured to manage credentials.