

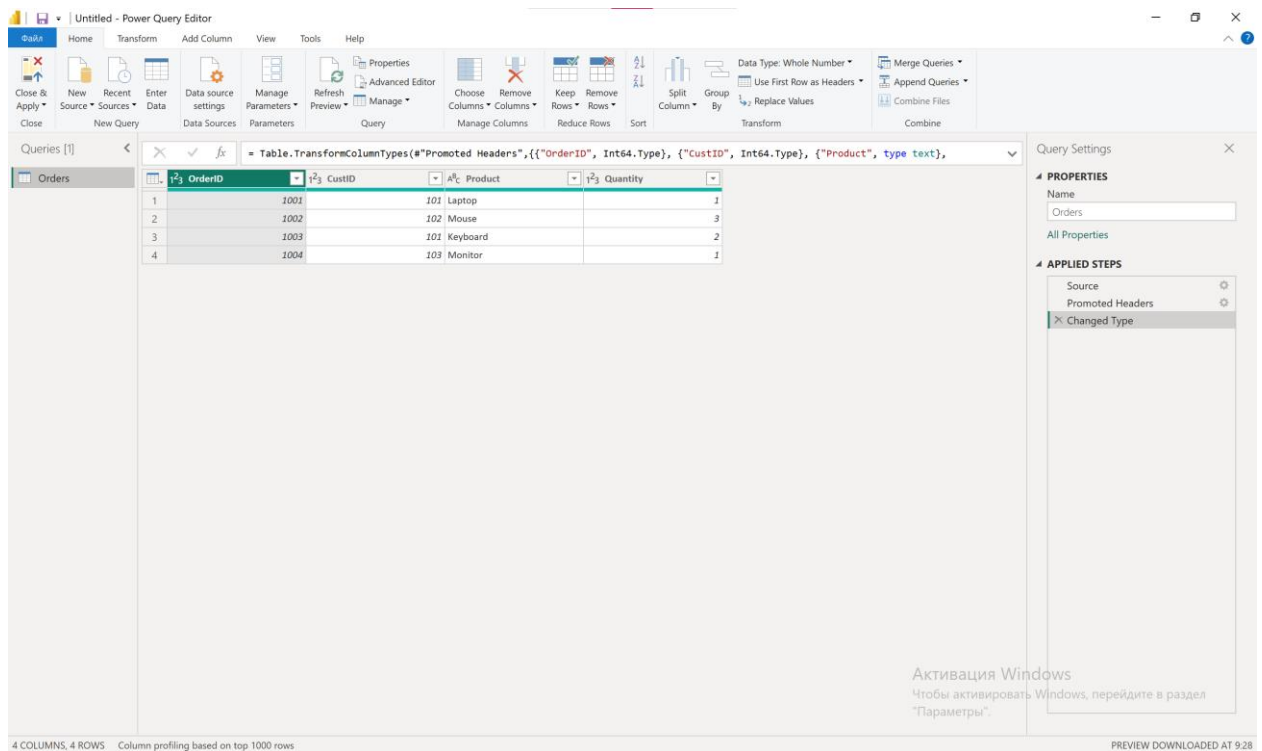
Lesson 4

Topic: Data Transformation with Power Query (Part 2)

Prerequisites: Download Customers.xlsx, Orders.csv

1. What is the difference between "Merge" and "Append" in Power Query?

- **Merge** joins two tables **side by side** based on a key column (like SQL JOIN).
- **Append** stacks two or more tables **on top of each other** (row union).



Untitled - Power Query Editor

Close & Apply, New Source, Recent Sources, Enter Data, Data source settings, Manage Parameters, Refresh Preview, Properties, Advanced Editor, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Data Type: Whole Number, Merge Queries, Append Queries, Combine Files, Combine

Queries [2] < Sheet1

1 CustID 101 Alice alice@example.com
2 102 Bob bob@example.com
3 103 Charlie charlie@example.com

Query Settings

PROPERTIES
Name: Sheet1
All Properties

APPLIED STEPS
Source, Navigation, Promoted Headers, Changed Type

3 COLUMNS, 3 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 9:30

Untitled - Power Query Editor

Close & Apply, New Source, Recent Sources, Enter Data, Data source settings, Manage Parameters, Refresh Preview, Properties, Advanced Editor, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Data Type: Whole Number, Merge Queries, Append Queries, Combine Files, Combine

Queries [2] < Orders, Sheet1

1 OrderID 1001
2 1002
3 1003
4 1004

101 Alice alice@example.com
102 Bob bob@example.com
103 Charlie charlie@example.com

Merge

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

Orders

OrderID	CustID	Product	Quantity
1001	101	Laptop	1
1002	102	Mouse	3
1003	101	Keyboard	2
1004	103	Monitor	1

Sheet1

CustID	Name	Email
101	Alice	alice@example.com
102	Bob	bob@example.com
103	Charlie	charlie@example.com

Join Kind
Left Outer (all from first, matching from second)
☐ Use fuzzy matching to perform the merge
> Fuzzy matching options
☒ The selection matches 4 of 4 rows from the first table.

OK Cancel

Query Settings

PROPERTIES
Name: Orders
All Properties

APPLIED STEPS
Source, Promoted Headers, Changed Type

4 COLUMNS, 4 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 9:28

Untitled - Power Query Editor

Queries [2] `= Table.NestedJoin("#Changed Type", {"CustID"}, Sheet1, {"CustID"}, "Sheet1", JoinKind.LeftOuter)`

OrderID	CustID	Product	Quantity	Sheet1
1	1001	101 Laptop	1	Table
2	1002	102 Mouse	3	Table
3	1003	101 Keyboard	2	Table
4	1004	103 Monitor	1	Table

5 COLUMNS, 4 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 9:30

Активация Windows
Чтобы активировать Windows, перейдите в раздел "Параметры".

Untitled - Power Query Editor

Queries [2] `= Table.Combine({"Merged Queries", Sheet1})`

OrderID	CustID	Product	Quantity	Sheet1	Name
1	1001	101 Laptop	1	Table	null
2	1002	102 Mouse	3	Table	null
3	1003	101 Keyboard	2	Table	null
4	1004	103 Monitor	1	Table	null
5	null	101	null	null	alice
6	null	102	null	null	bob
7	null	103	null	null	charlie

7 COLUMNS, 7 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED ON СЫББОТА

Чтобы начать поиск, введите здесь запрос

0:51
19.09.2025

2. How do you split a "Full Name" column into "First Name" and "Last Name"?

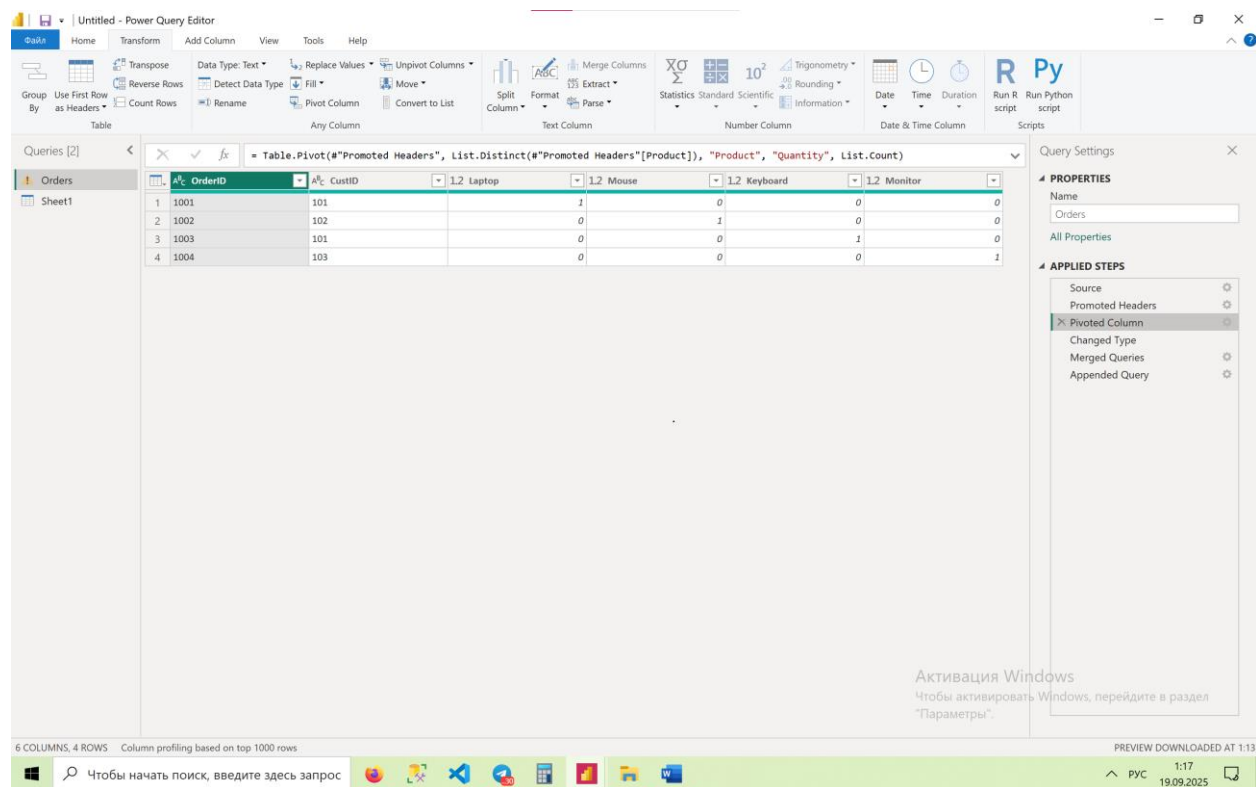
Use **Split Column → By Delimiter → Space**.

- First part → First Name
- Second part → Last Name

3. What is "Pivot Columns" used for?

It transforms row values into new column headers.

Example: Products (Laptop, Mouse, Monitor) become columns with Quantity as values.



4. How do you undo a step in Power Query?

In **Applied Steps** pane:

- Click the **X** to remove a step, or

- Right-click → **Delete**

5. What is the purpose of "Reference" vs. "Duplicate" in queries?

- **Duplicate** = full independent copy of a query.
 - **Reference** = linked copy, changes in the original flow through.
-

6. Merge Orders.csv and Customers.xlsx on CustID (inner join).

The screenshot shows the Power Query Editor interface. On the left, the 'Queries' pane lists 'Orders' and 'Sheet1'. The 'Orders' query is selected, and its data is displayed in the main area:

OrderID	CustID	Product	Quantity
1001	101	Laptop	1
1002	102	Mouse	3
1003	101	Keyboard	2
1004	103	Monitor	1

The 'Merge' dialog box is open, showing the 'Orders' query as the first table and 'Sheet1' as the second table. The 'Join Kind' is set to 'Inner (only matching rows)'. The dialog also shows the data for 'Sheet1':

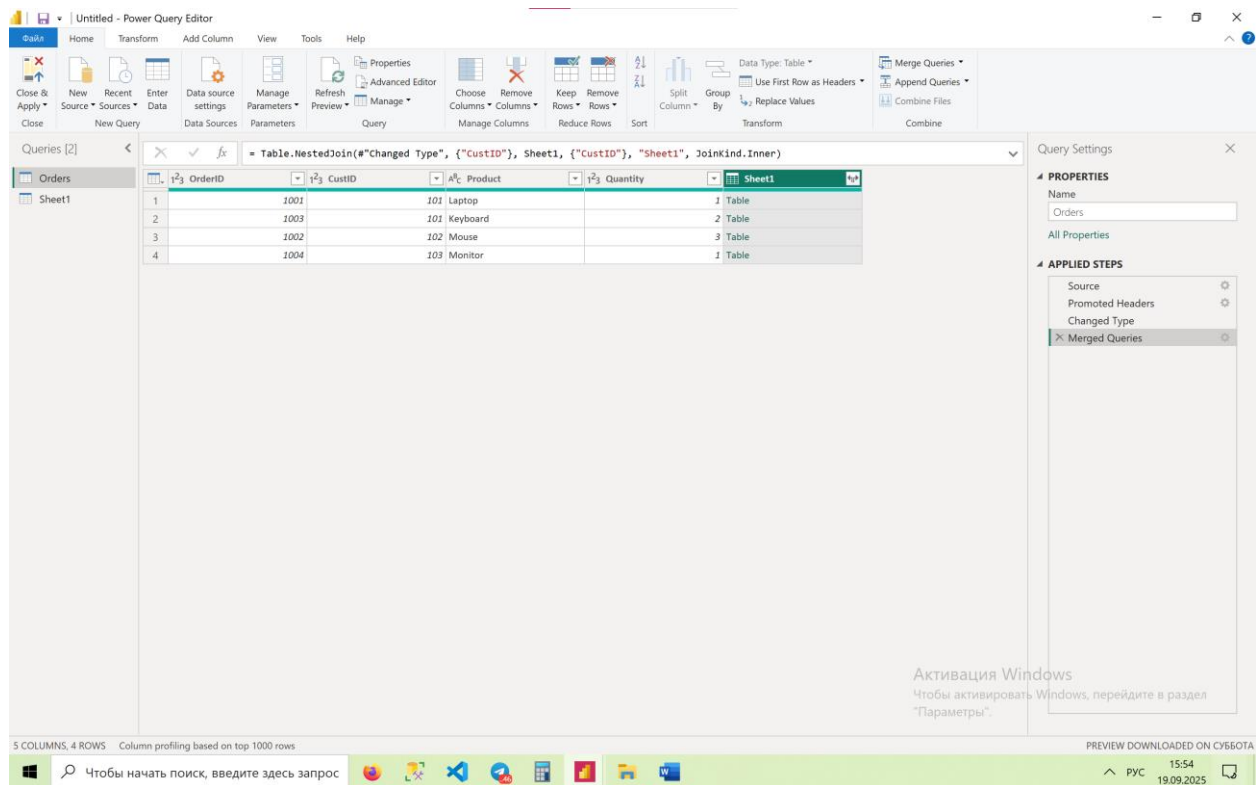
CustID	Name	Email
101	Alice	alice@example.com
102	Bob	bob@example.com
103	Charlie	charlie@example.com

The 'Merge' dialog box also shows the 'Properties' and 'Applied Steps' panes on the right. The 'Properties' pane shows the 'Name' as 'Orders'. The 'Applied Steps' pane shows the steps: 'Source', 'Promoted Headers', and 'Changed Type'.

At the bottom of the dialog, it states: 'The selection matches 0 of 4 rows from the first table, and 0 of 3 rows from the second table.' The 'OK' button is highlighted.

Windows activation watermark: 'Активация Windows. Чтобы активировать Windows, перейдите в раздел "Параметры".'

Taskbar: 'Чтобы начать поиск, введите здесь запрос' and 'PREVIEW DOWNLOADED AT 15:24'.

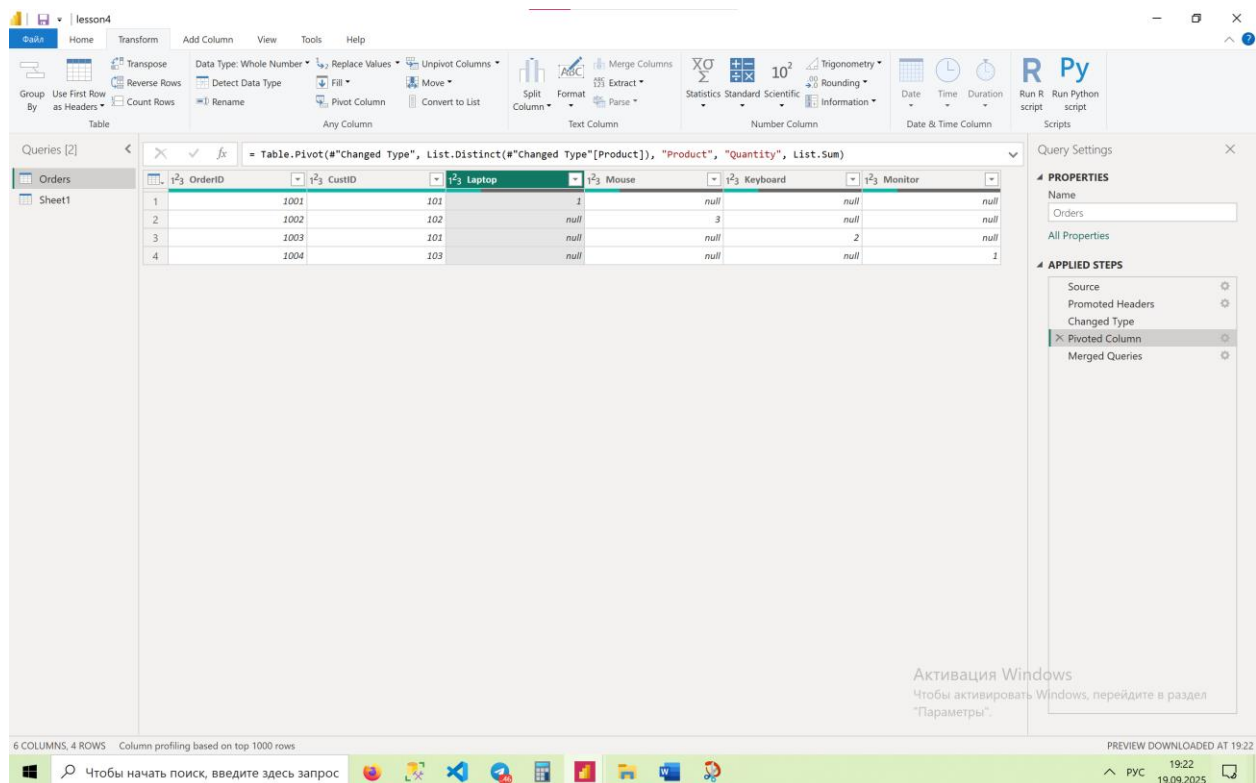


7. Pivot Product column

Transform → Pivot Column → Values Column = Quantity →

Aggregation = Sum

Result: Laptop, Mouse, Keyboard, Monitor become columns, values = total Quantity.



8. Append Orders_Jan + Orders_Feb

Home → Append Queries → Select both tables → OK
All rows from both tables are combined into one.

Table: Table.Combine({#"Merged Queries", #"Orders Feb"})

	OrderID	CustID	Laptop	Mouse	Keyboard	Monitor
1	1001	101	1	null	null	null
2	1003	101	1	null	2	null
3	1002	102	1	3	null	1
4	1004	103	1	null	null	2
5	1001	101	1	null	null	null
6	1003	101	1	null	2	1
7	1002	102	1	3	null	1
8	1004	103	1	null	null	1

7 COLUMNS, 8 ROWS - Column profiling based on top 1000 rows

Активация Windows
Чтобы активировать Windows, перейдите в раздел "Параметры".

PREVIEW DOWNLOADED ON СУББОТА 19:36 19.09.2025

9. Fill Down Email column

Transform → Fill → Down

Replaces null with the previous value in the column.

lesson4

Home Transform Add Column View Tools Help

Group By Use First Row as Headers Count Rows

Table

Any Column

Text Column

Number Column

Date & Time Column

Scripts

Queries [3]

Orders

Sheet1

Orders Feb

Table

Table.ReplaceValue(#Promoted Headers,"bob@example.com","",Replacer.ReplaceText,{"Email"}))

CustID	Name	Email
1	Alice	alice@example.com
2	Bob	
3	Charlie	charlie@example.com

Query Settings

PROPERTIES

Name

Sheet1

All Properties

APPLIED STEPS

Source

Navigation

Promoted Headers

Replaced Value

Активация Windows

Чтобы активировать Windows, перейдите в раздел "Параметры".

3 COLUMNS, 3 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED ON CYBOTA

ENG 19:43 19.09.2025

lesson4

Home Transform Add Column View Tools Help

Group By Use First Row as Headers Count Rows

Table

Any Column

Text Column

Number Column

Date & Time Column

Scripts

Queries [3]

Orders

Sheet1

Orders Feb

Table

Table.FillDown(#Replaced Value1,{"Email"})

CustID	Name	Email
1	Alice	alice@example.com
2	Bob	alice@example.com
3	Charlie	charlie@example.com

Query Settings

PROPERTIES

Name

Sheet1

All Properties

APPLIED STEPS

Source

Navigation

Promoted Headers

Replaced Value

Filled Down

Replaced Value1

Filled Down1

Активация Windows

Чтобы активировать Windows, перейдите в раздел "Параметры".

3 COLUMNS, 3 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED ON CYBOTA

ENG 19:44 19.09.2025

10. Extract Email domain

Transform → **Extract** → **Text After Delimiter (@)**

alice@example.com

The screenshot shows the Power Query Editor interface. The main area displays a table with the following data:

CustID	Name	Email
1	Alice	alice@example.com
2	Bob	bob@example.com
3	Charlie	charlie@example.com

The formula bar shows the function: `= Table.PromoteHeaders(Sheet1_Sheet, [PromoteAllScalars=true])`. The right-hand pane shows the 'Query Settings' for 'Sheet1', with the 'APPLIED STEPS' list containing: Source, Navigation, Promoted Headers, Extracted Text After Delimiter, Replaced Value, Filled Down, Replaced Value1, and Filled Down1.

The screenshot shows the Power Query Editor interface. The main area displays a table with the following data:

CustID	Name	Email
1	Alice	example.com
2	Bob	example.com
3	Charlie	example.com

The formula bar shows the function: `= Table.TransformColumns(#"Promoted Headers", {"Email", each Text.AfterDelimiter(_, "@", type text)})`. The right-hand pane shows the 'Query Settings' for 'Sheet1', with the 'APPLIED STEPS' list containing: Source, Navigation, Promoted Headers, Extracted Text After Delimiter, Replaced Value, Filled Down, Replaced Value1, and Filled Down1.

11. Write M-code to merge queries dynamically based on a parameter (e.g., JoinType = "Inner").

Home → Manage Parameters → New Parameter.

let

```

SourceOrders = Orders,
SourceCustomers = Customers,
JoinType = "Inner", // параметр можно поменять
Merged = Table.NestedJoin(
    SourceOrders, {"CustID"},
    SourceCustomers, {"CustID"},
    "CustomerData",
    JoinKind.Inner
)
in
Merged

```

12. Unpivot a table with columns like "Jan_Sales," "Feb_Sales" into a "Month" and "Sales" format.

Transform → Unpivot Columns → Select Jan_Sales, Feb_Sales, etc.
Result:

- Column "Attribute" → Month
- Column "Value" → Sales

13. Handle errors in a custom column (e.g., division by zero) using try...otherwise.

= try [Numerator] / [Denominator] otherwise 0

14. Create a function in Power Query to clean phone numbers (e.g., remove dashes).

(phone as text) =>

let

Clean = Text.Remove(phone, {"-", " ", "(", ")"})

in

Clean

15. Optimize a query with 10+ steps—identify bottlenecks and simplify.

- Remove unnecessary steps (especially repeated Change Type).
- Combine filters into a single step.
- Use **Buffer()** for temporary tables.
- Remove unused columns early.
- Use **Reference** instead of Duplicate when possible