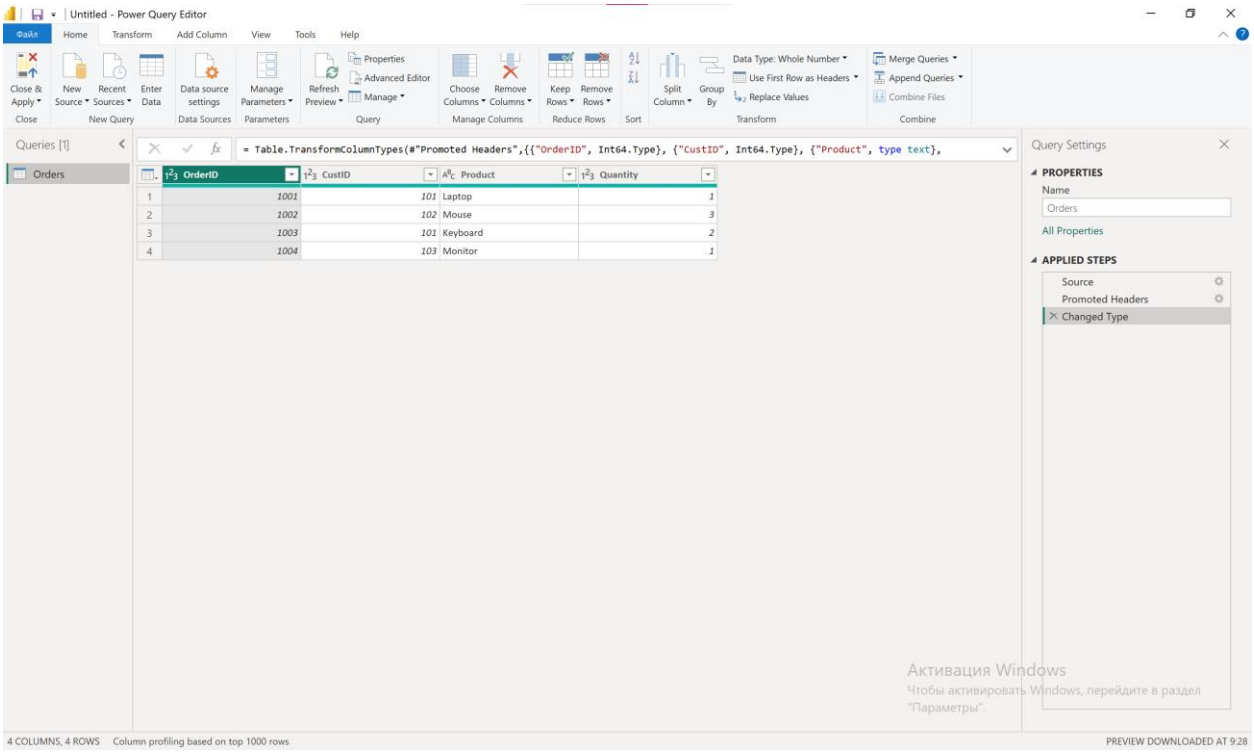
Lesson 4

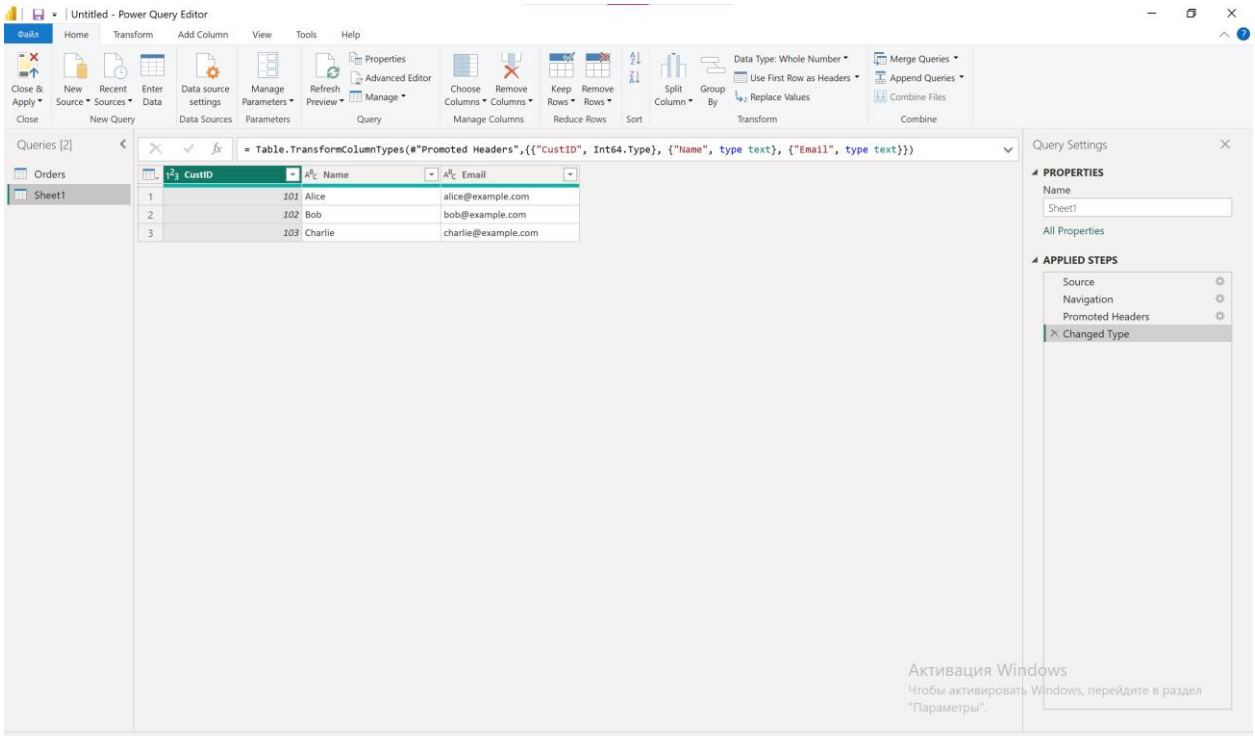
Topic: Data Transformation with Power Query (Part 2)

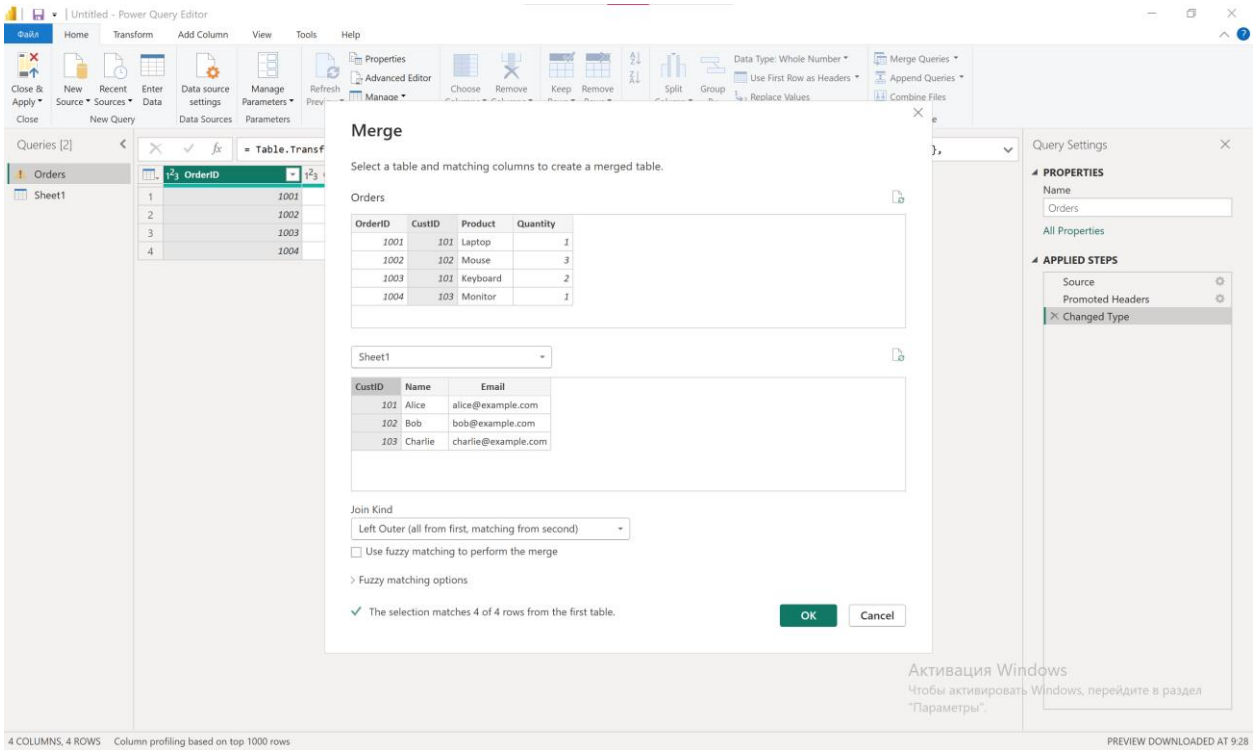
Prerequisites: Download Customers.xlsx, Orders.csv

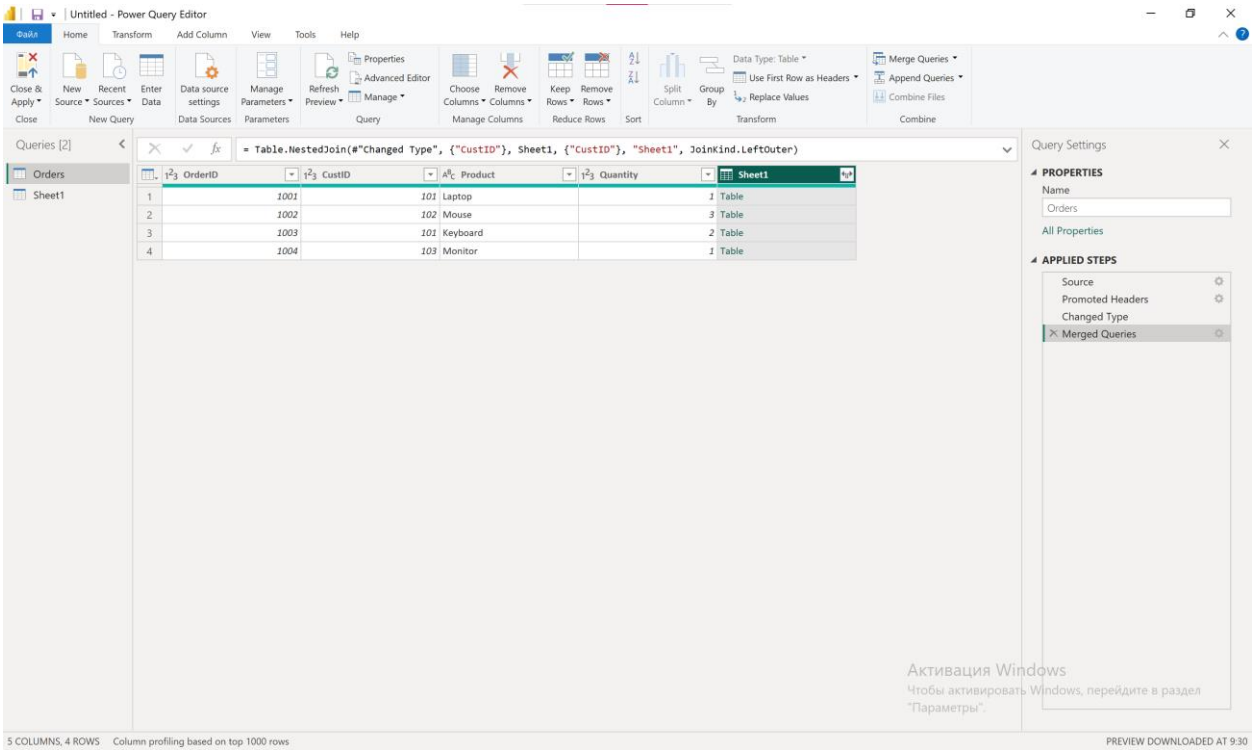
1. What is the difference between "Merge" and "Append" in Power Query?

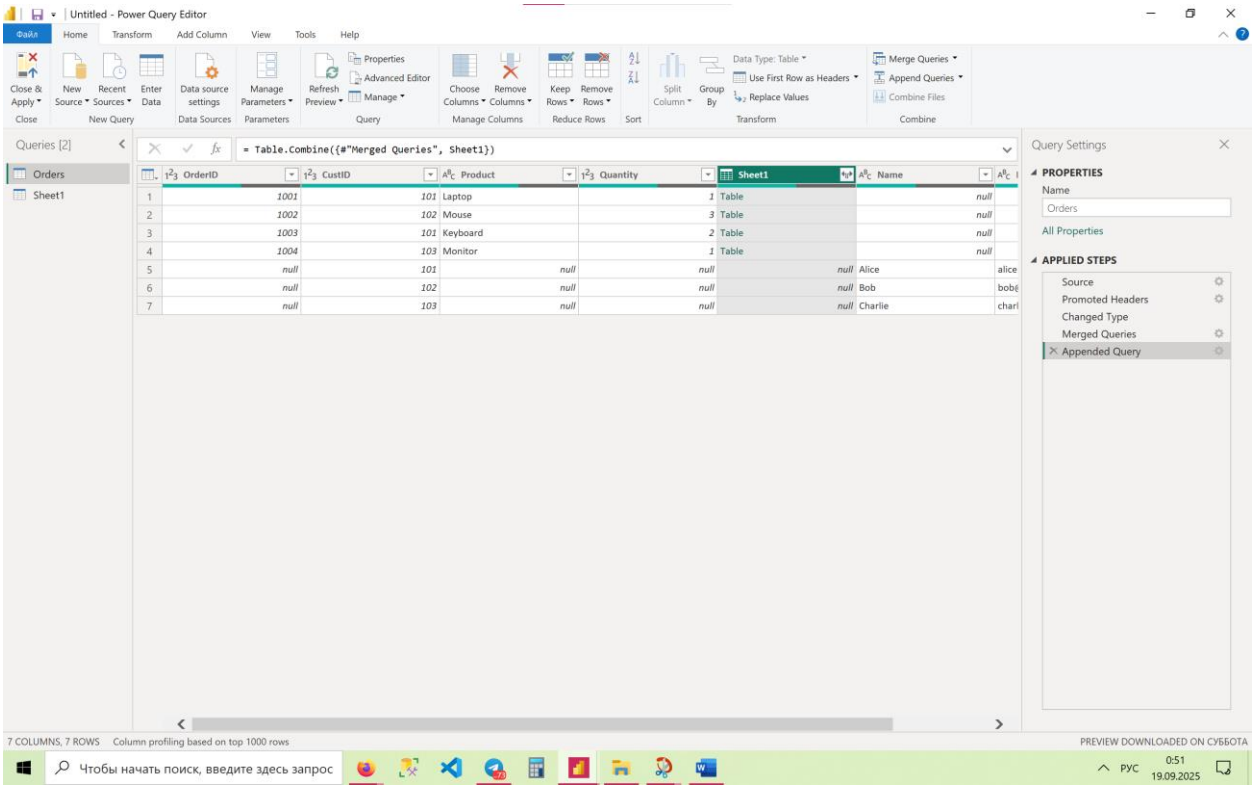
- **Merge** joins two tables **side by side** based on a key column (like SQL JOIN).
- **Append** stacks two or more tables **on top of each other** (row union).




Untitled - Power Query Editor
Файл
Home
Transform
Add Column
View
Tools
Help
Close & Apply
New Source
Recent Sources
Enter Data
Data source settings
Manage Parameters
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Whole Number
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Close
New Query
Data Sources
Parameters
Query
Manage Columns
Reduce Rows
Sort
Transform
Combine
Queries [2]
Orders
Sheet1
= Table.TransformColumnTypes(#"Promoted Headers",{{"CustID", Int64.Type}, {"Name", type text}, {"Email", type text}})
CustID
Name
Email
1 101 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
Активация Windows
Чтобы активировать Windows, перейдите в раздел "Параметры".
3 COLUMNS, 3 ROWS
Column profiling based on top 1000 rows
PREVIEW DOWNLOADED AT 9:30


Untitled - Power Query Editor
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
Queries [2]
Orders
Sheet1
OrderID
1001
1002
1003
1004
Query Settings
PROPERTIES
Name
Orders
All Properties
APPLIED STEPS
Source
Promoted Headers
Changed Type
Активация Windows
Чтобы активировать Windows, перейдите в раздел "Параметры".
4 COLUMNS, 4 ROWS
Column profiling based on top 1000 rows
PREVIEW DOWNLOADED AT 9:28





---

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

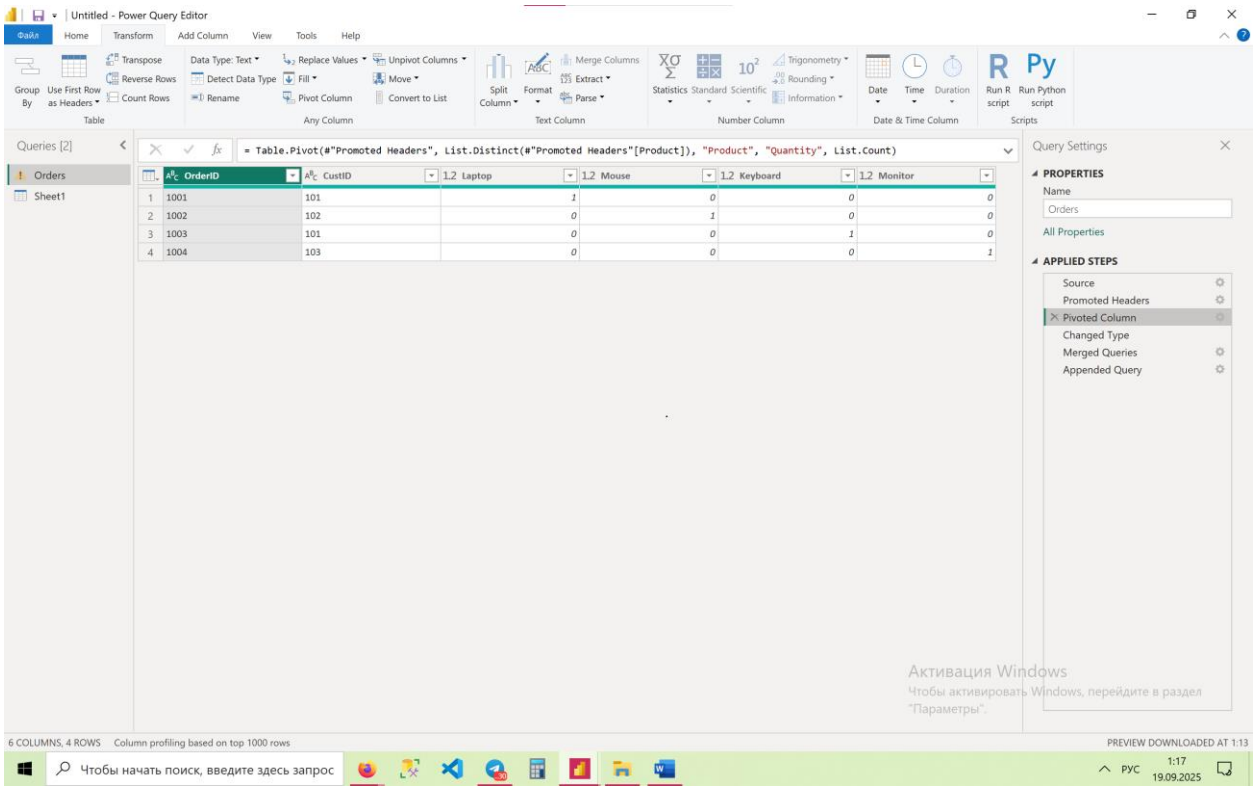
Use **Split Column → By Delimiter → Space**.

- 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**

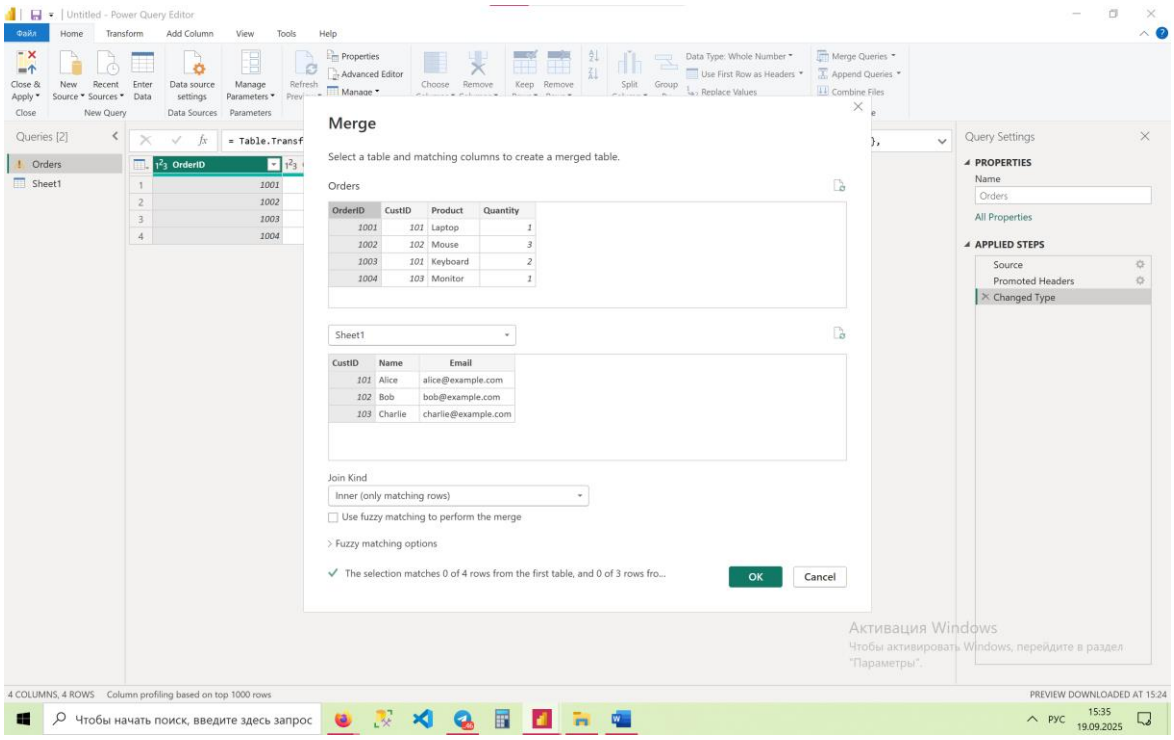
---

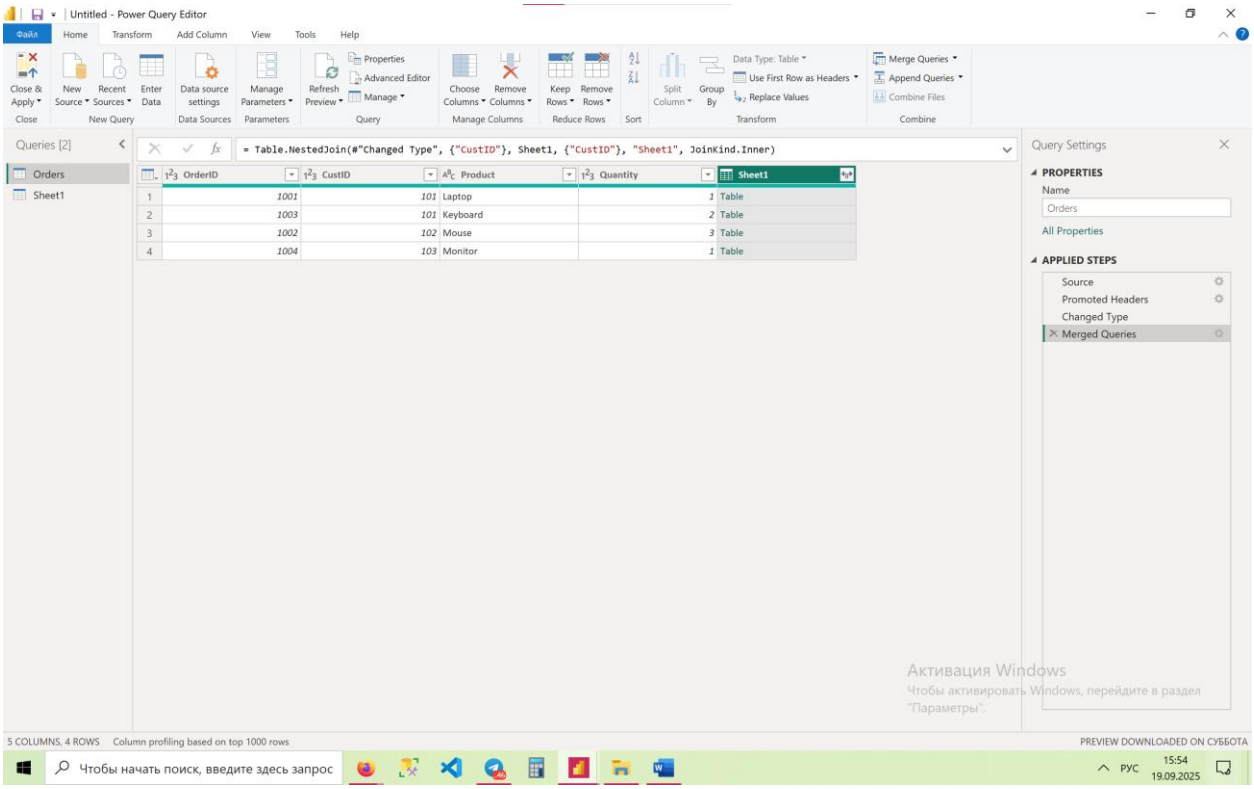
5. What is the purpose of "Reference" vs. "Duplicate" in queries?

- **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).

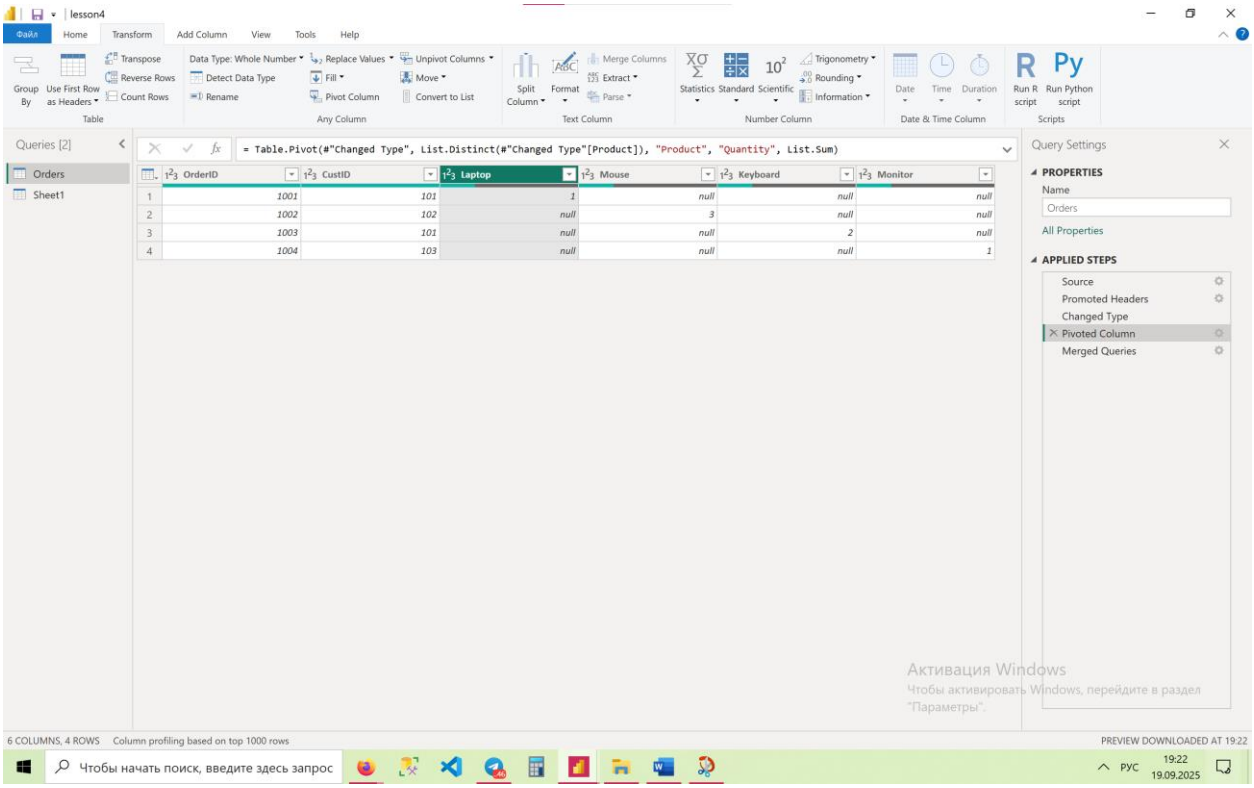




7. Pivot Product column

Transform → Pivot Column → Values Column = Quantity → Aggregation = Sum

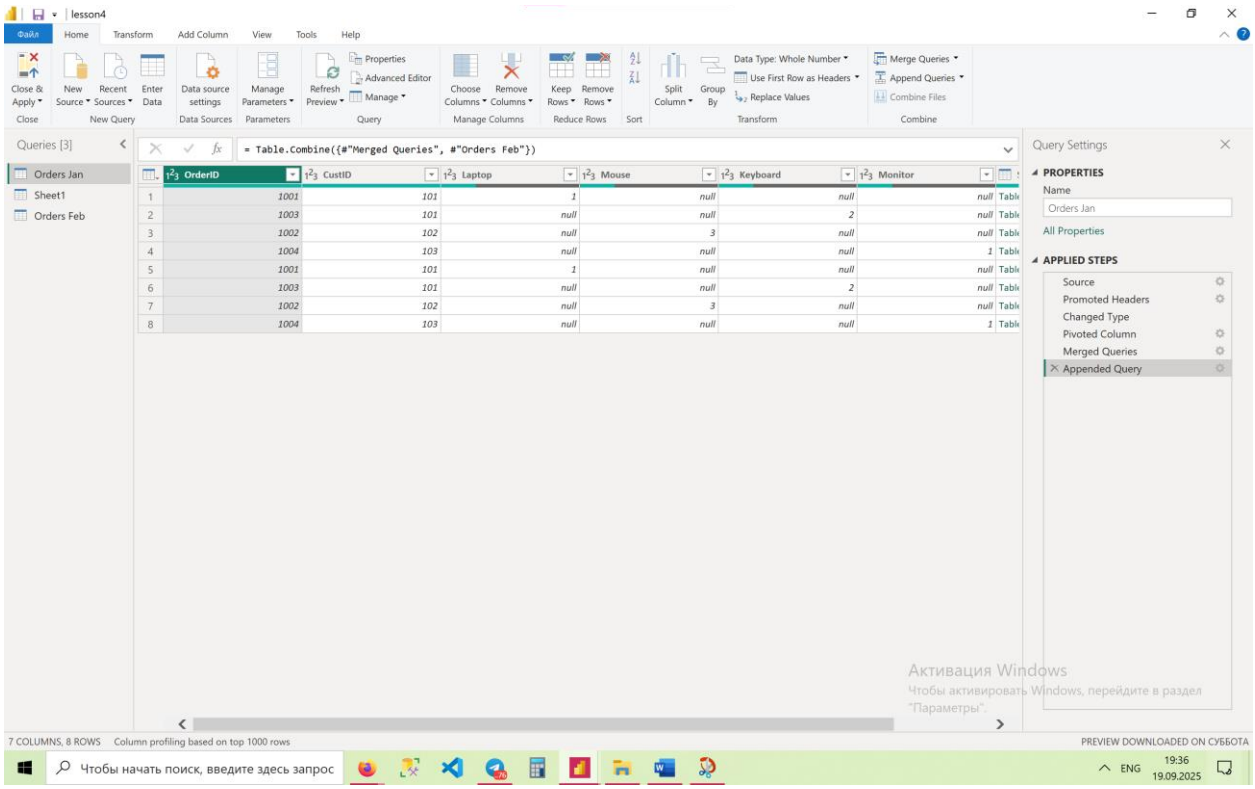
Result: Laptop, Mouse, Keyboard, Monitor become columns, values = total Quantity.



8. Append Orders_Jan + Orders_Feb

Home → Append Queries → Select both tables → OK

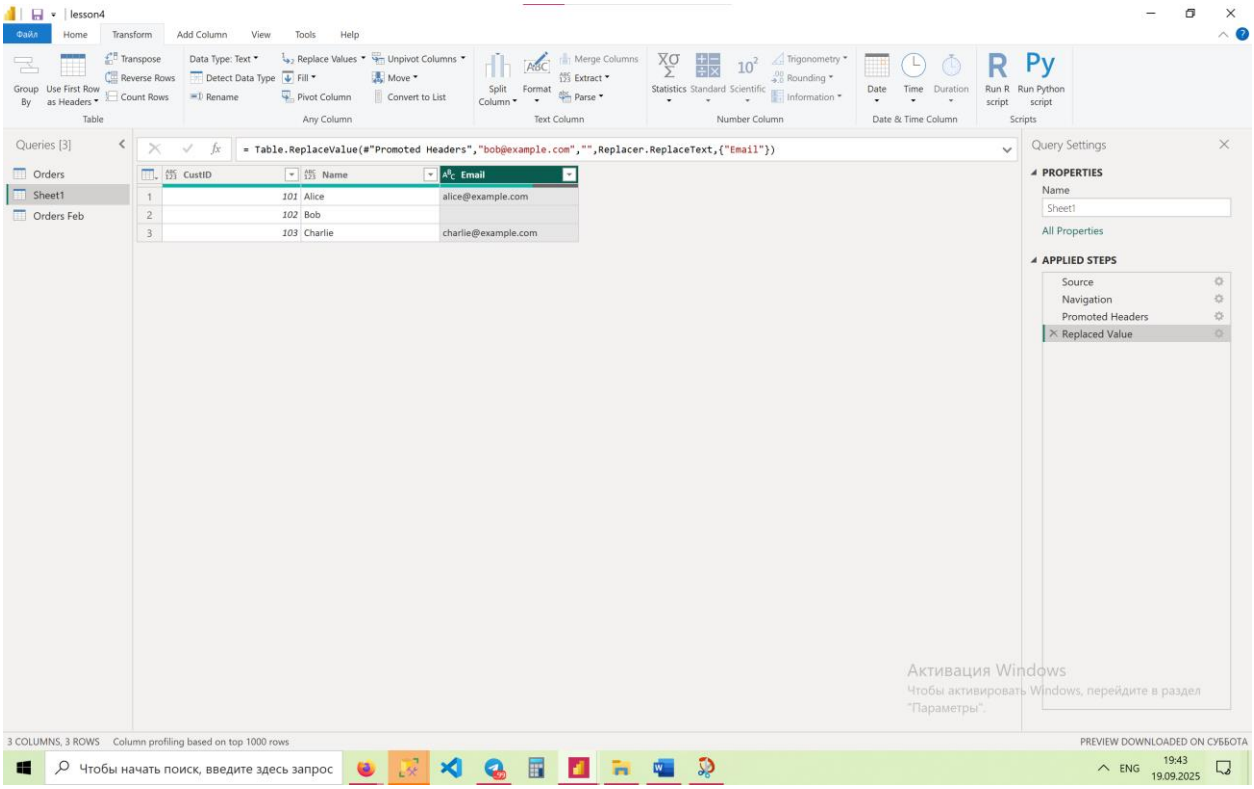
All rows from both tables are combined into one.

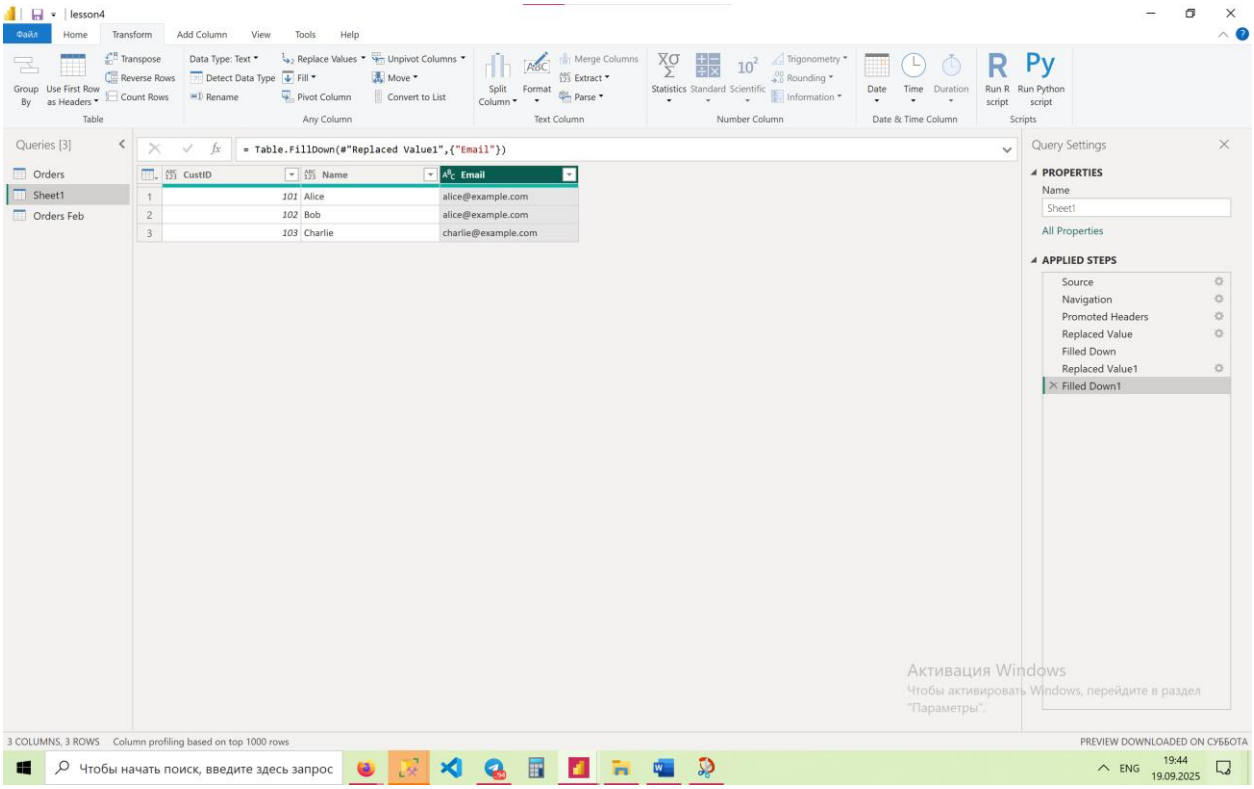


9. Fill Down Email column

Transform → Fill → Down

Replaces null with the previous value in the column.

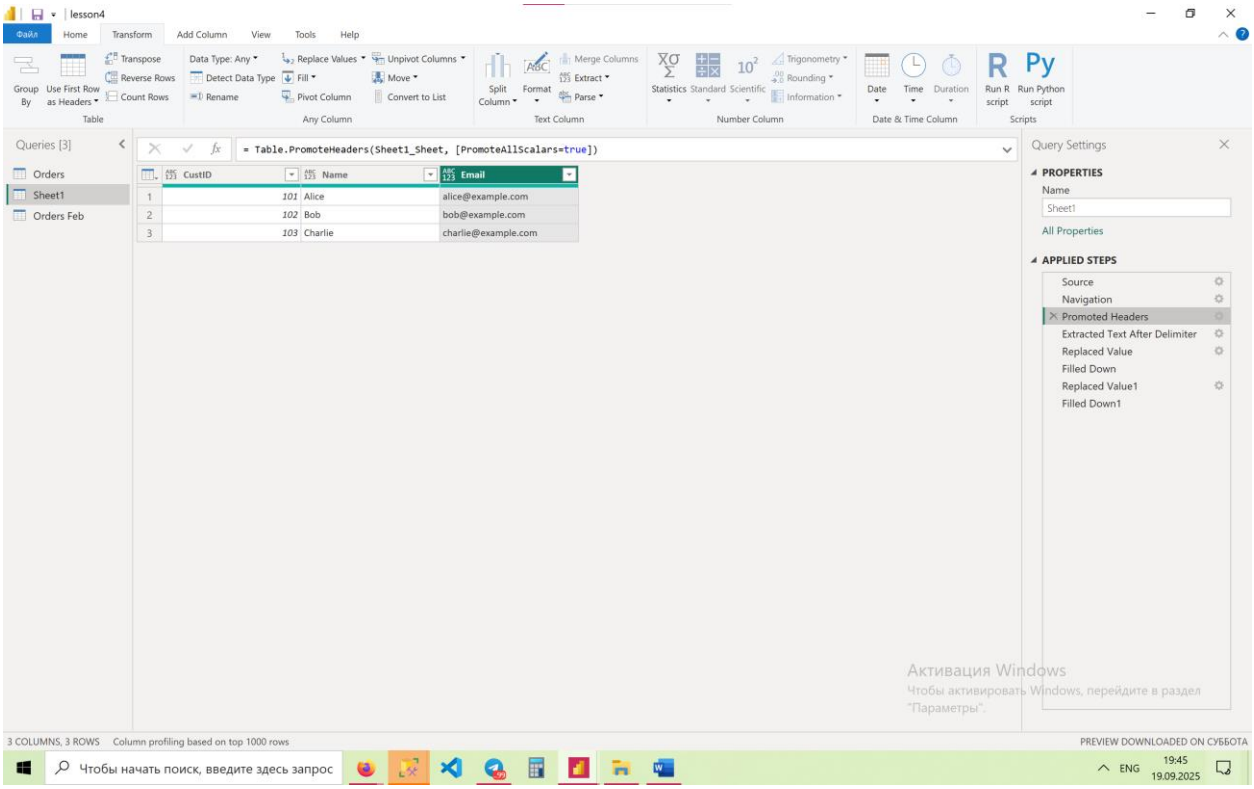


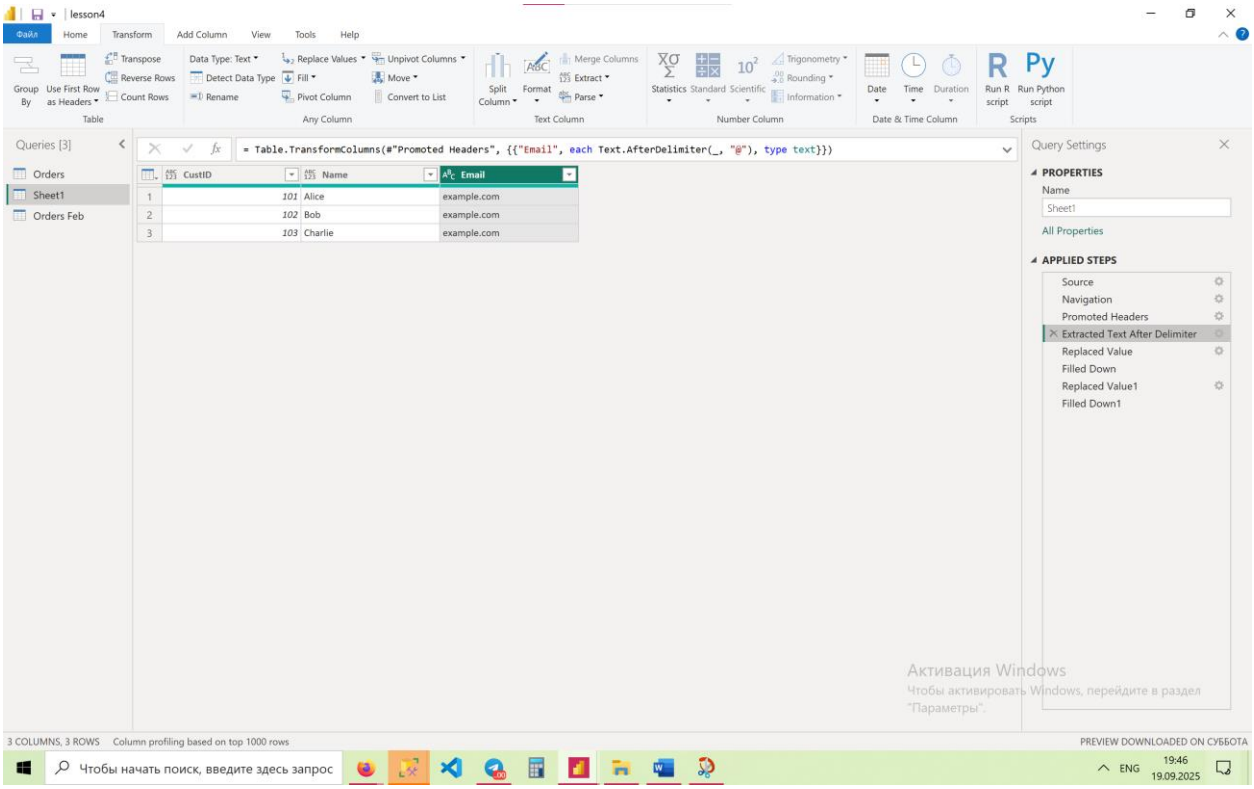


10. Extract Email domain

Transform → Extract → Text After Delimiter (@)

alice@example.com





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