

# Lesson 3: Data Transformation with Power Query (Part 1)

## 1. What is the purpose of the Applied Steps pane in Power Query?

**Applied Steps** показывает все шаги трансформации данных, которые были выполнены в Power Query (удаление строк, фильтры, переименование столбцов и т.д.).

Каждый шаг:

- выполняется последовательно;
  - может быть отредактирован, удалён или переупорядочен;
  - автоматически записывается в М-код.
- 

## 2. How do you remove duplicate rows in Power Query?

1. Выделить один или несколько столбцов
2. Вкладка **Home** → **Remove Rows** → **Remove Duplicates**

Power Query оставит только уникальные строки.

---

## 3. What does the Filter icon do in Power Query?

Иконка фильтра:

- отбирает строки по условию (текст, число, дата);
  - позволяет исключать/включать значения;
  - работает аналогично фильтрам в Excel.
- 

## 4. How would you rename a column from "CustID" to "CustomerID"?

Способы:

- Двойной клик по заголовку столбца → ввести новое имя
  - Правый клик → **Rename**
- 

## 5. What happens if you click Close & Apply in Power Query?

- Все изменения **сохраняются**
  - Данные загружаются в модель Power BI
  - Power Query закрывается
- 

## 6. Remove all rows where Quantity < 2

Шаги:

1. Нажать фильтр в столбце **Quantity**
  2. **Number Filters** → **Greater Than or Equal To** → 2
- 

## 7. Split the OrderDate column into Year, Month, Day

1. Выделить **OrderDate**
  2. **Transform** → **Date** → **Year / Month / Day**
  3. Создаются три новых столбца
- 

## 8. Replace all "Mouse" with "Computer Mouse" in Product

1. Выделить столбец **Product**
  2. **Transform** → **Replace Values**
  3. Find: `Mouse`  
Replace with: `Computer Mouse`
- 

## 9. Sort the table by OrderDate (newest first)

- Нажать стрелку в столбце **OrderDate**
  - Выбрать **Sort Descending**
- 

## 10. How would you handle null values in the Price column?

Варианты:

- **Replace Values** → **null** → 0
- Удалить строки с null
- Заменить средним/минимальным значением

Чаще всего в продажах: замена на 0

---

## 11. Custom M-code: TotalSpent = Quantity \* Price

```
= Table.AddColumn(  
    Source,  
    "TotalSpent",  
    each [Quantity] * [Price],  
    type number  
)
```

---

## 12. Group by CustID to show total spending per customer

1. **Transform** → **Group By**
2. Group by: `CustID`
3. Operation: **Sum**
4. Column: `TotalSpent`

Результат: общий расход по каждому клиенту.

---

## 13. Fix inconsistent date formats in OrderDate

Решение:

1. Выделить **OrderDate**
2. **Change Type** → **Using Locale**
3. Тип: **Date**
4. Locale: нужный регион (например, English – US)

Power Query приведёт даты к единому формату.

---

## 14. Conditional column: High Value if Price > 100

1. **Add Column** → **Conditional Column**
2. Условие:
  - If `Price > 100` → "High Value"
  - Else → "Normal"

## 15. Optimize the query to reduce refresh time

Лучшие практики:

- Удалять **ненужные столбцы как можно раньше**
  - Удалять лишние строки до сложных вычислений
  - Минимизировать количество шагов
  - Избегать сложных вычислений на больших таблицах
- 

## Исходные данные (пример)

### Orders

CustID	Name	OrderDate	Product	Quantity	Price
101	Alice	2023-01-10	Laptop	1	1200

<b>CustID</b>	<b>Name</b>	<b>OrderDate</b>	<b>Product</b>	<b>Quantity</b>	<b>Price</b>
102	Bob	2023-01-15	Mouse	3	25
101	Alice	2023-01-20	Keyboard	2	80
103	Charlie	2023-01-25	Monitor	1	300

### **Customer\_Orders**

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