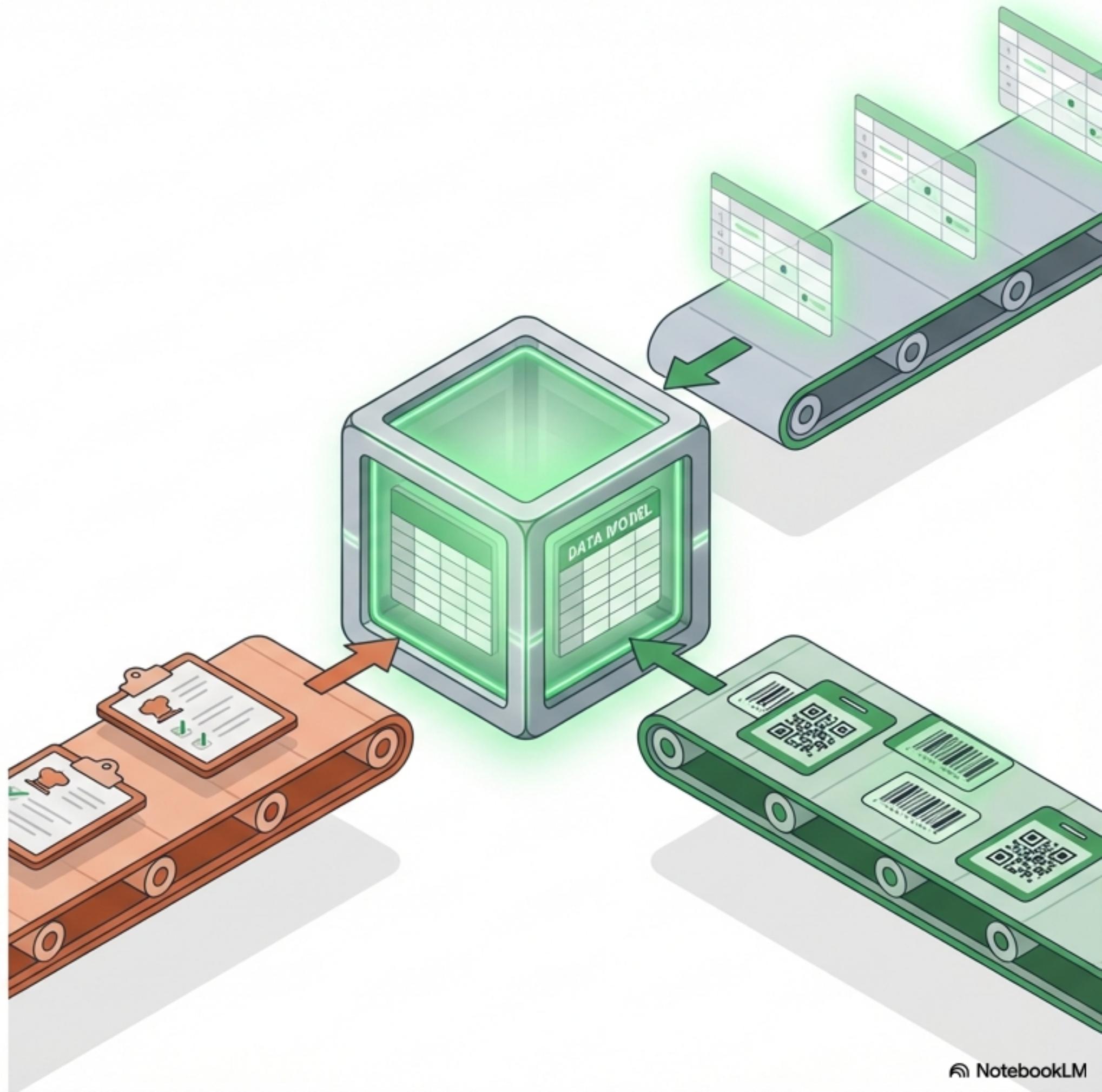
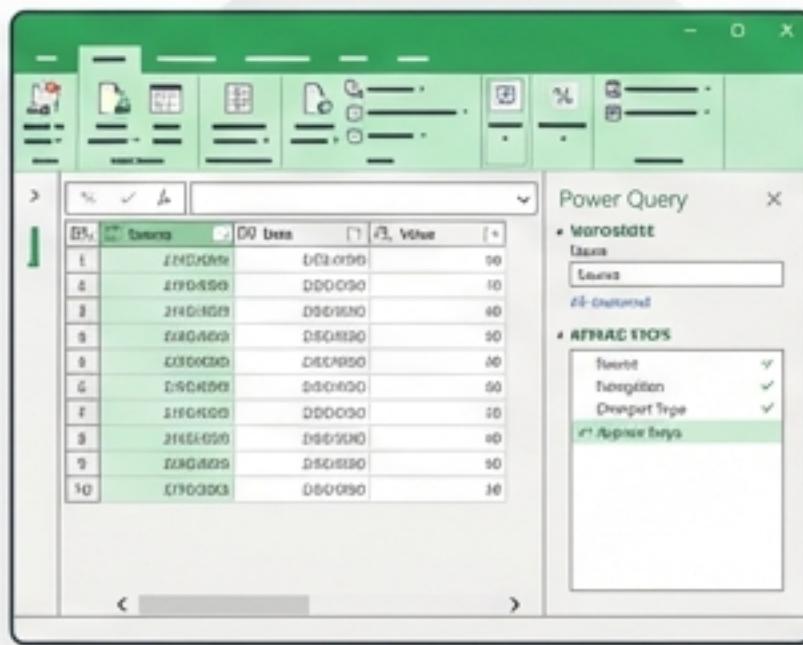
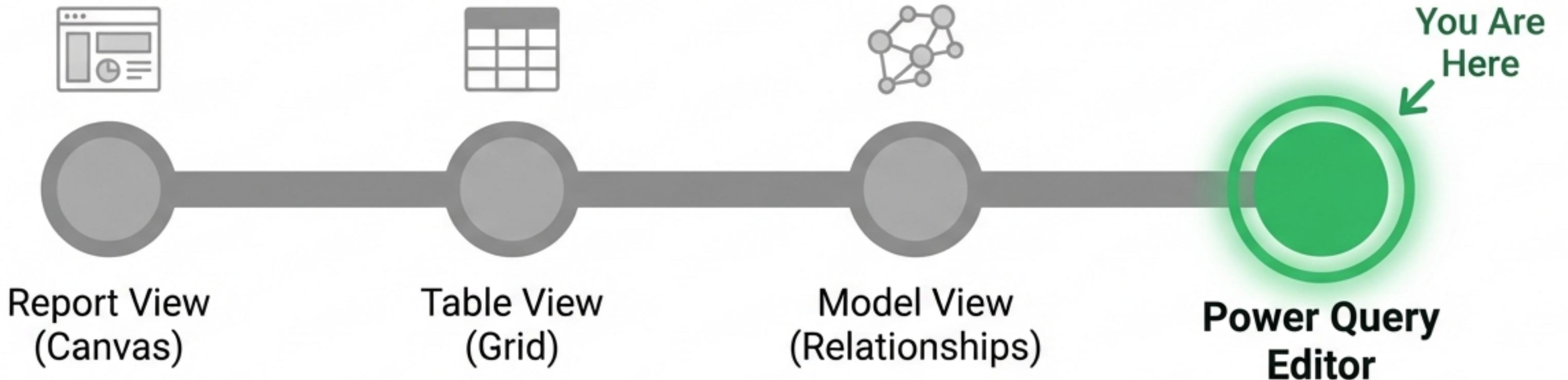


The final assembly in Power Query

Combining queries, configuring loads, and the final handoff to the data model.

Part II: Profiling, Cleaning,
and Transforming Data





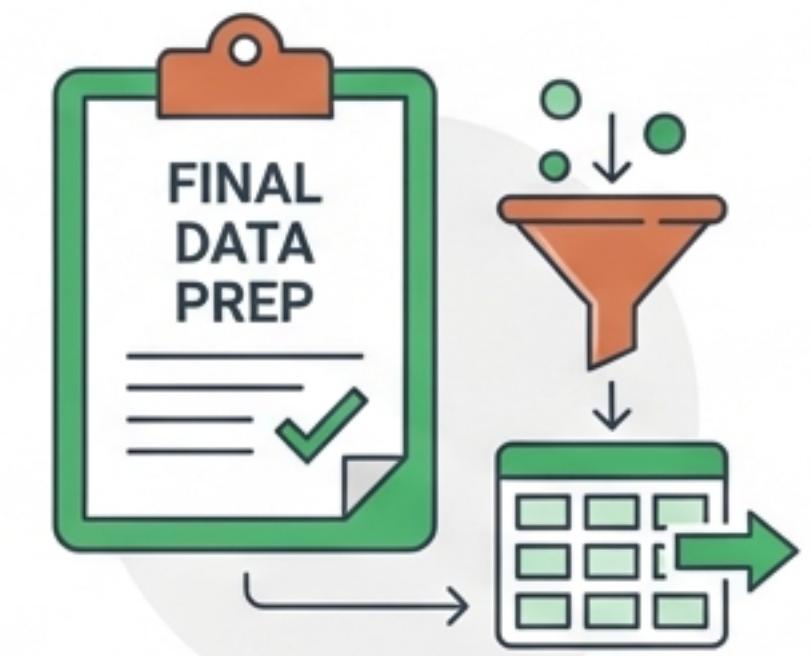
Our home base remains the Power Query Editor

Current Location:

We are still in the separate window with the green ribbon and Applied Steps pane.

The Mission:

This chapter is our farewell to the Editor. We are preparing the data for its final exit.

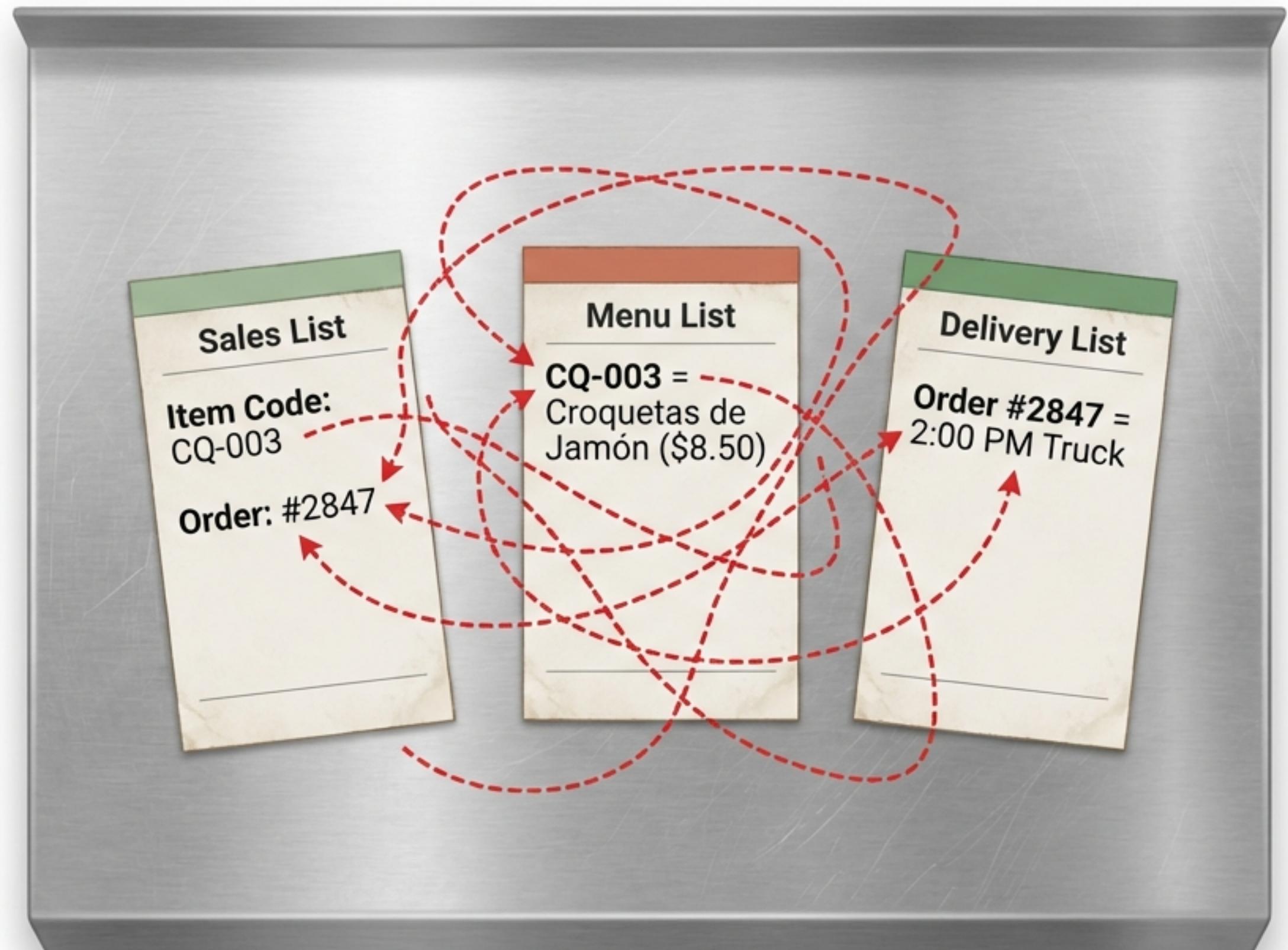


Clean data is useless if it lives in isolation

The Scenario: A customer asks, "What time will my croquetas arrive, and how much do I owe?"

The Friction: Sofia has to mentally connect three separate codes across three separate lists to answer one question.

The Goal: Imagine doing this for 200 orders. We need a single, connected dataset.



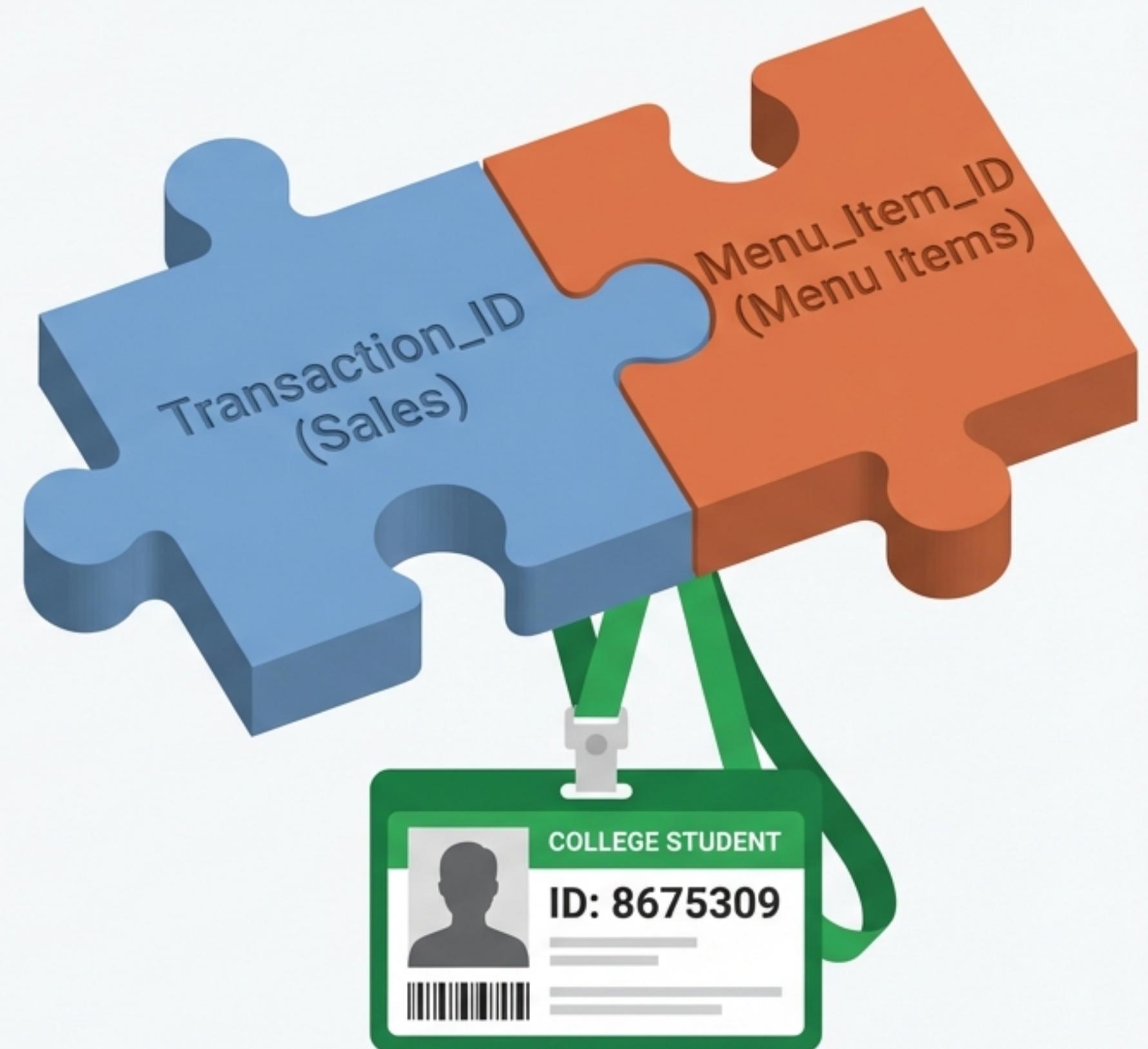
Keys are the glue that connect isolated tables

Primary Key (The ID Card):

Uniquely identifies every row in a table. No duplicates. No blanks.
(e.g., Menu_Item_ID in the Menu table).

Foreign Key (The Bridge):

A column in another table that points back to the primary key.
(e.g., Menu_Item_ID inside the Sales table).



Two fundamentally different ways to combine your data

Merge Queries

Merge adds columns.

Connects tables horizontally using a matching key.
Like scanning a wristband to look up a meal plan.



SQL Equivalent: JOIN ... ON
Excel: VLOOKUP

Append Queries

Append adds rows.

Stacks tables vertically with matching columns. Like stacking morning and afternoon ventanita shift logs.



SQL Equivalent: UNION ALL
Excel: Copy-paste below

The Left Join automatically matches what it can find

The Action:

A Left Outer join keeps ALL rows from your first table (Sales) and brings in matching information from the second (Menu Items).

What if there is no match?

If an ingredient has no price listed, it stays on the card with a blank space (null). Nothing gets deleted. The blanks simply tell you where your price list has gaps.



Deciding who makes the final cut

Table A: Your Family's List

Left Outer (The Default)

Everyone on YOUR family list, even if the venue hasn't confirmed them yet. Keep all of A, matching B.

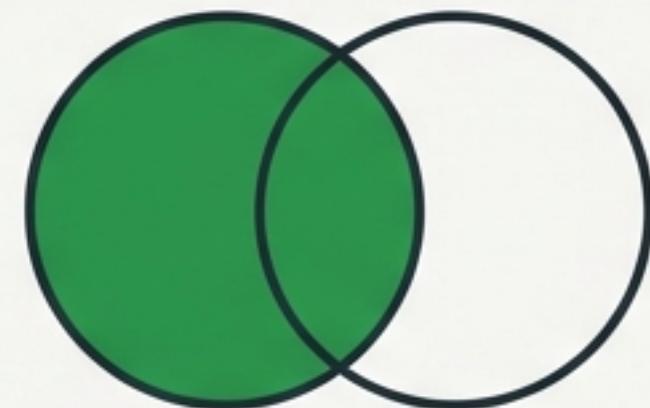
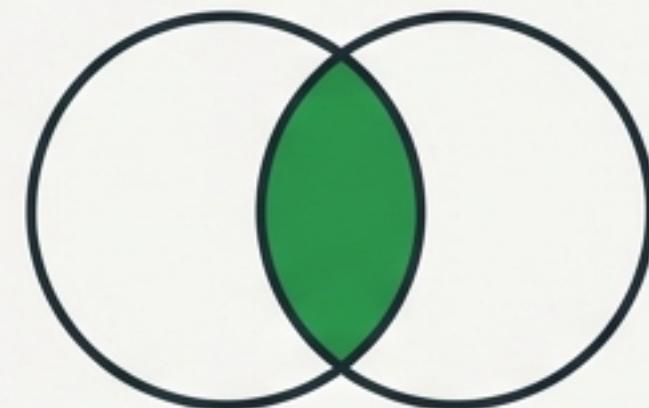


Table B: The Venue's List

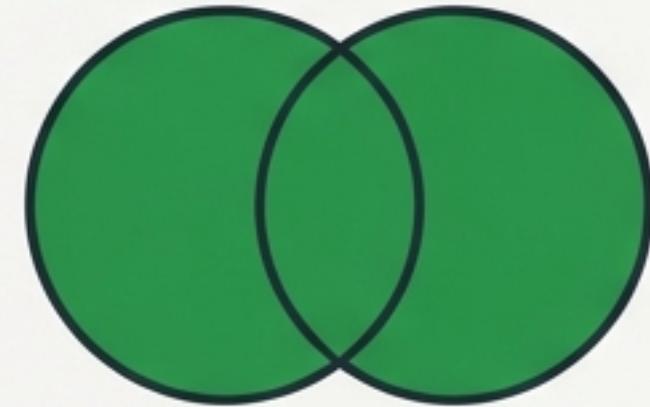
Inner

Only people who appear on BOTH lists. Strict matching.



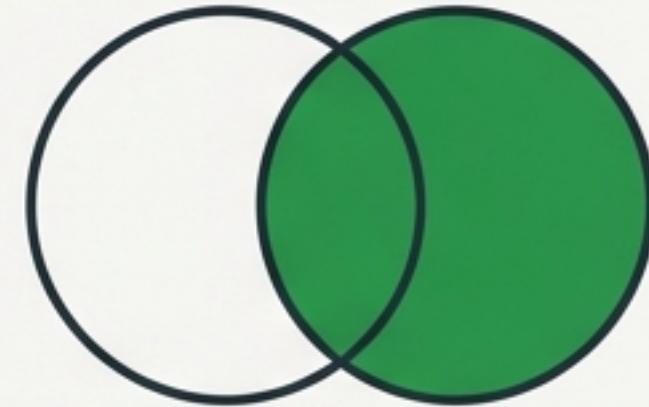
Full Outer

Everyone from both lists. No one is left out.



Right Outer

Everyone the venue approved, even if they aren't on your family list.



Use Anti Joins to diagnose hidden data problems

The Concept:

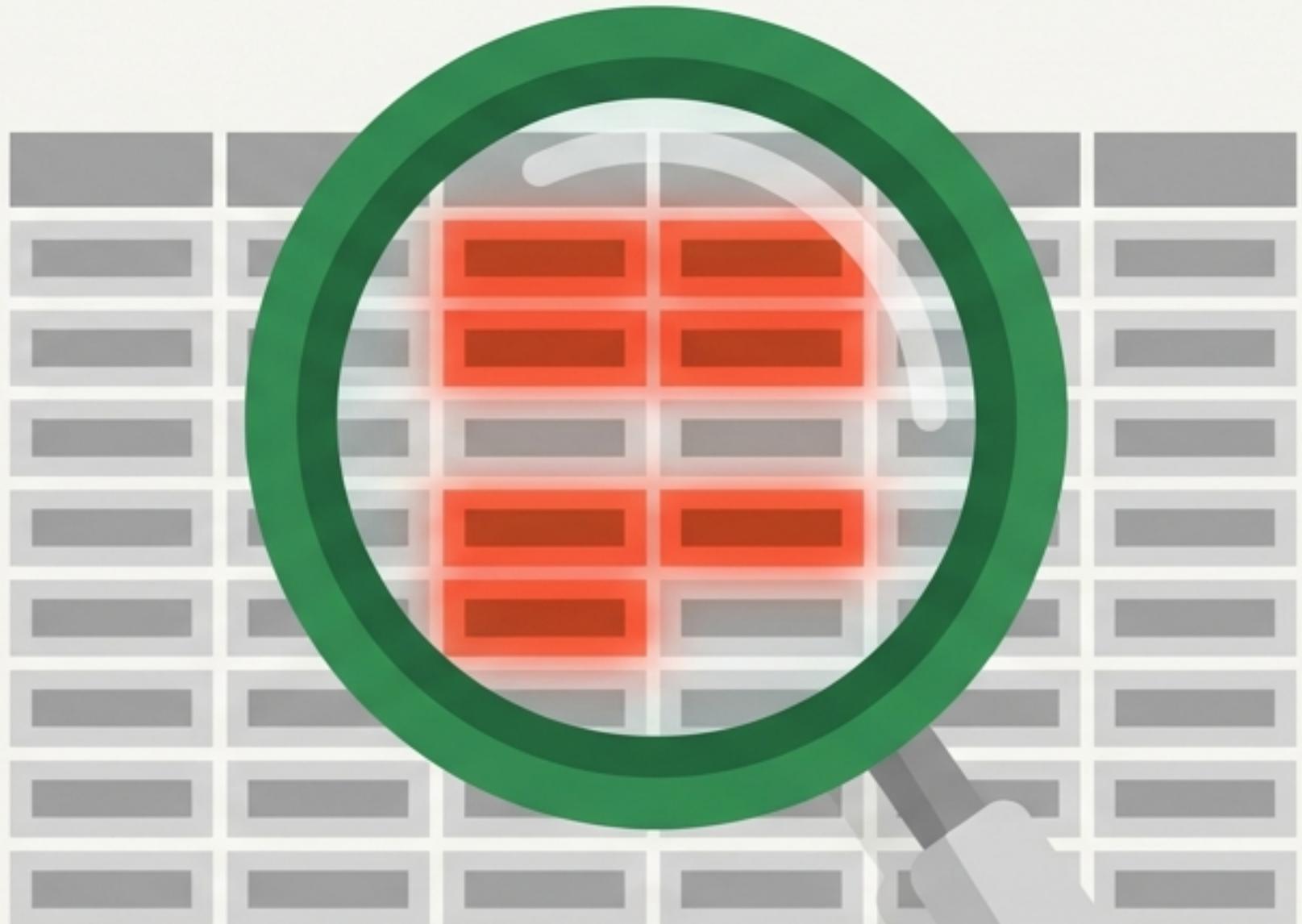
An Anti Join returns only rows that have NO match in the other table.

Left Anti Join:

People on your family list who the venue explicitly rejected.

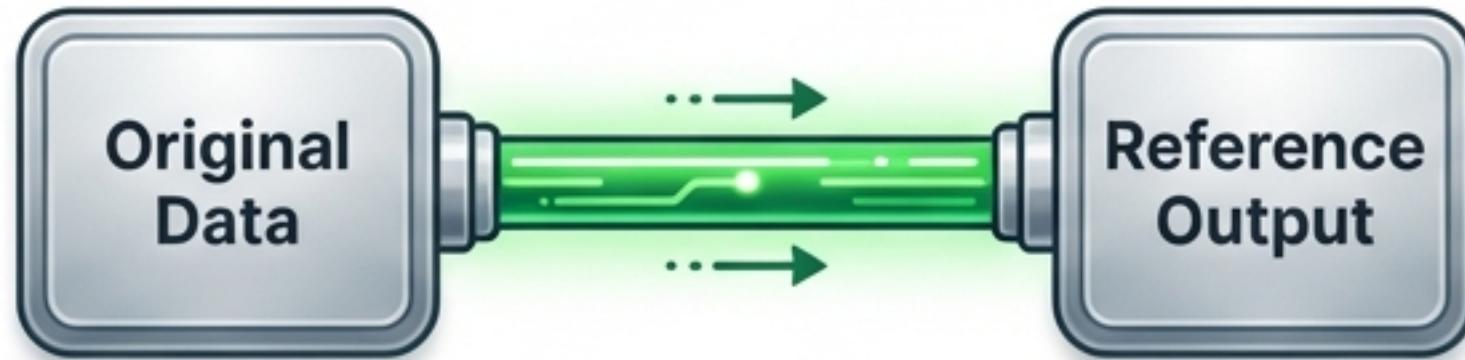
The Business Value:

Instantly find data quality issues. Use a Left Anti Join to find sales transactions referencing a **Menu_Item_ID** that does not exist in your system.



Managing your query paths: Reference vs. Duplicate

Reference Query



- A live link. Starts from the original's final output. If you change a step in the original, the Reference updates automatically.
- **Best for:** Building summaries (like Total Sales per Truck) on top of cleaned data.

Duplicate Query



- A frozen snapshot. Copies every single applied step independently. Changes to the original will NOT affect the duplicate.
- **Best for:** Creating a separate starting point that you don't want to change.

Keep your prep bowls out of the dining room

Staging Queries:

Do the intermediate work (cleaning, reshaping, summarizing) behind the scenes.

The Toggle:

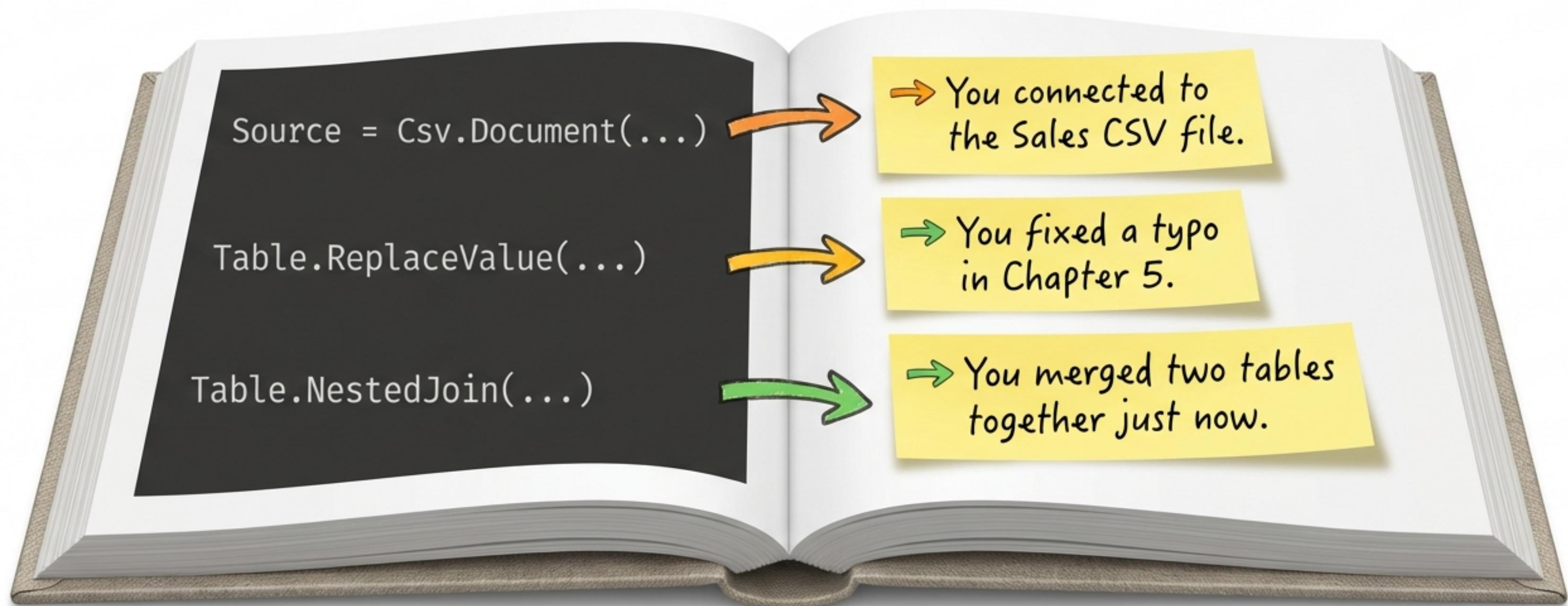
Right-click a query and uncheck Enable Load. It stays in the kitchen, feeding other queries, but never appears in the final report.



← Unchecked items become *italicized* staging queries.

You are reading a diary, not writing a program

Every click in Power Query writes a line of M code in the Advanced Editor. You don't need to write it from scratch; you just need to read the history.



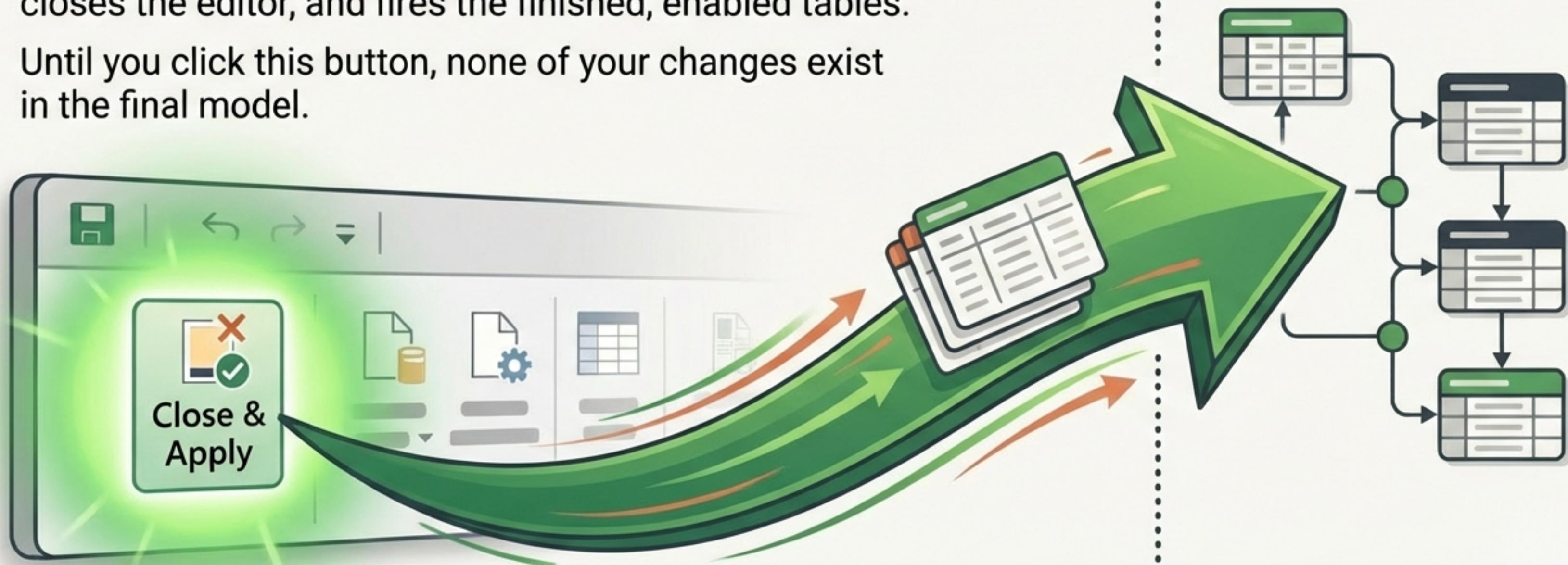
Passing the ticket through the serving window

The final action in Power Query.

- Clicking Close & Apply seals your transformations, closes the editor, and fires the finished, enabled tables.
- Until you click this button, none of your changes exist in the final model.

Power Query
Editor

Power BI
Data Model





Power Query Editor



Report View
(Canvas)



Model View
(Relationships)

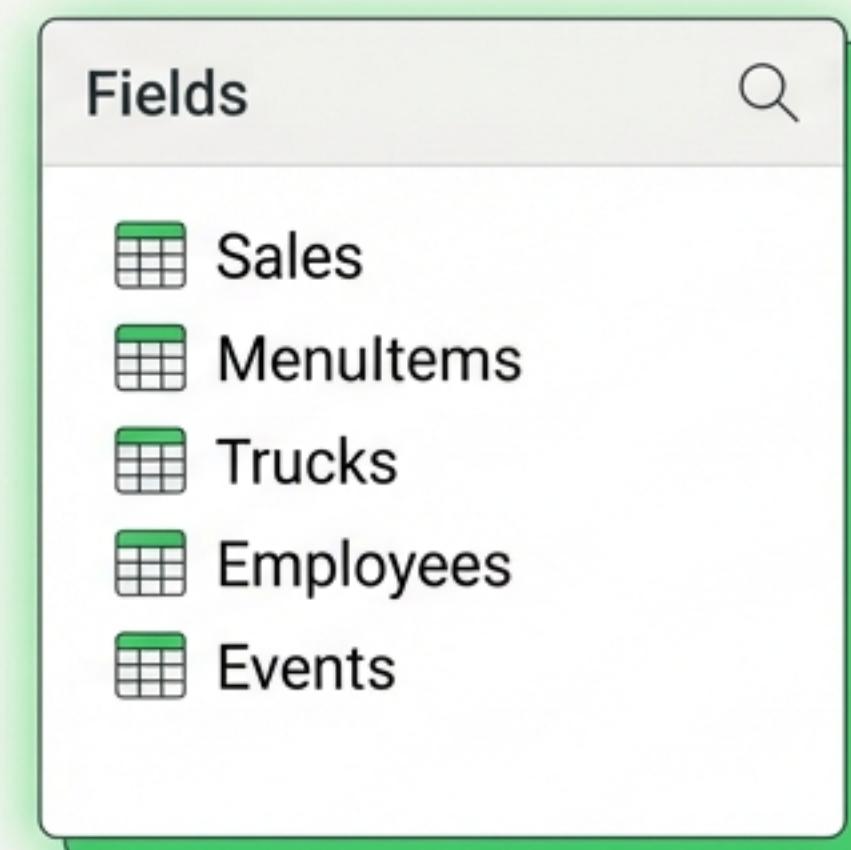


Data Modeling

250,000 clean rows, ready for calculation

The Payoff:

Every food truck transaction from 2024 is now cleaned of errors, enriched with item names, and loaded into a single model.



What's Next:

The data is prepared. Next, we define the **relationships** (Data Modeling) and write custom DAX calculations to extract the insights.

