

# Power Query vs. Python in Excel

## Which tool for which job?

Power Query and Python in Excel overlap more than people expect. Both can filter, reshape, merge, split columns, and parse dates, which is exactly what makes choosing between them feel harder than it should. The choice gets simple once you stop asking which tool can do the job and start asking which one fits the job in front of you.

### Power Query

#### GET THE DATA IN AND CLEAN IT UP

This is Excel's tool for the extract, transform, and load work. It connects to files, folders, databases, and web pages, then walks your data through a series of cleaning steps. Every one of those steps is recorded in the Applied Steps pane, so the work is easy to follow, audit, and run again next month. It also moves large tables without dragging the workbook down.

Where it runs out of room is analysis. Once you need statistics, modeling, or anything past a basic PivotChart, Power Query was not built to take you there.

### Python in Excel

#### RUN THE REAL ANALYSIS

This is where the heavier work lives. You get real statistics, from descriptive summaries to hypothesis tests and regression, proper charts through Matplotlib and Seaborn, and quick analysis across groups for things like segment averages and outliers.

Two things to plan around. It cannot pull in outside data yet, so you load the data first, usually with Power Query, then point Python at it. And because it recalculates inside the workbook, heavy scripts can slow things down. It also expects you to be comfortable writing a little code.

### WHEN TO USE WHICH

| WHEN YOU NEED TO...                            | USE             | BECAUSE   |
|--|-----------------|---|
| Pull data in from a file, database, or the web | Power Query     | It is the only one of the two that connects out |
| Clean and reshape before anything else         | Power Query     | Documented, repeatable steps                    |
| Keep an audit trail of every change            | Power Query     | Applied Steps records it all                    |
| Keep a heavy workbook responsive               | Power Query     | Cleaning runs outside the grid                  |
| Run statistics, regression, or a model         | Python in Excel | The libraries are built for it                  |
| Build a chart Excel cannot                     | Python in Excel | Matplotlib and Seaborn                          |
| Summarize or find outliers across groups       | Python in Excel | Fast, flexible grouping                         |
| Work with time series or a forecast            | Python in Excel | Built for trends and seasonality                |

### Use them together

You rarely have to pick just one. The strongest workflows use both. **Power Query** pulls the data in and does the first round of cleaning, then hands a tidy table to **Python in Excel** for the analysis and the visuals.

**Where to start:** new to both? Begin with Power Query. Get comfortable building cleaning workflows there, then add Python in Excel on top for the analysis Power Query was never meant to do.