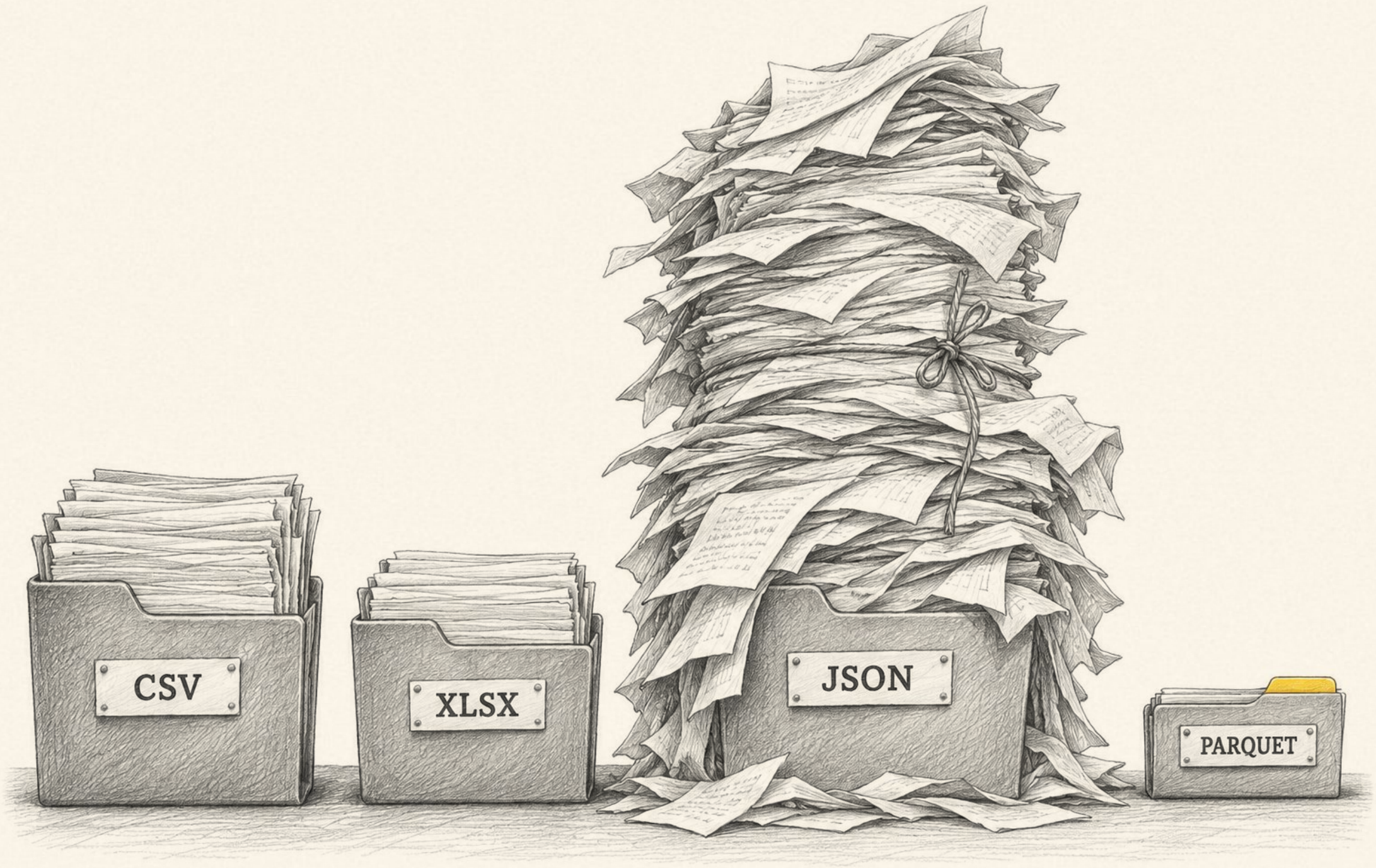
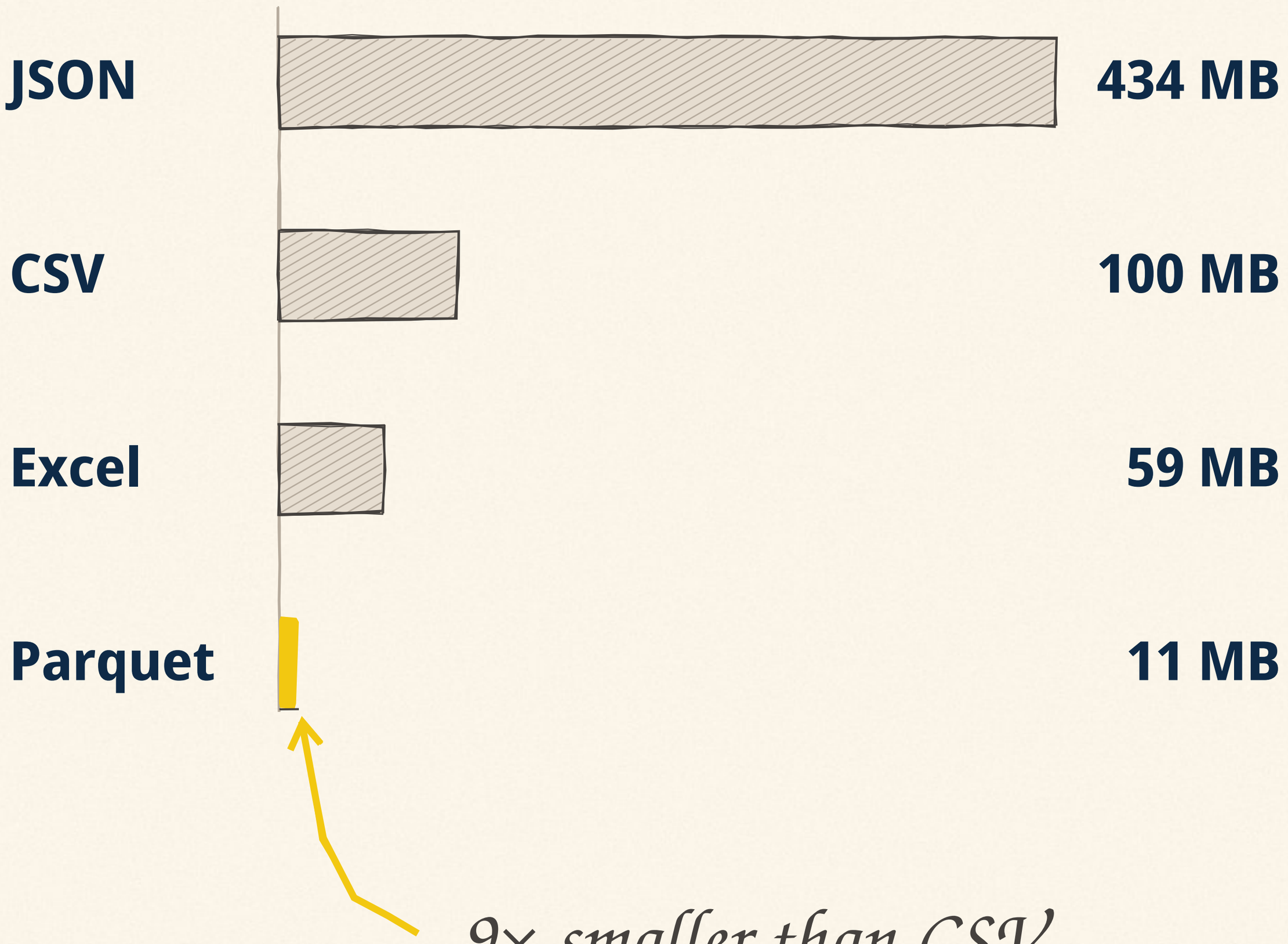
1M GL LINES

CSV, Excel, JSON or Parquet?

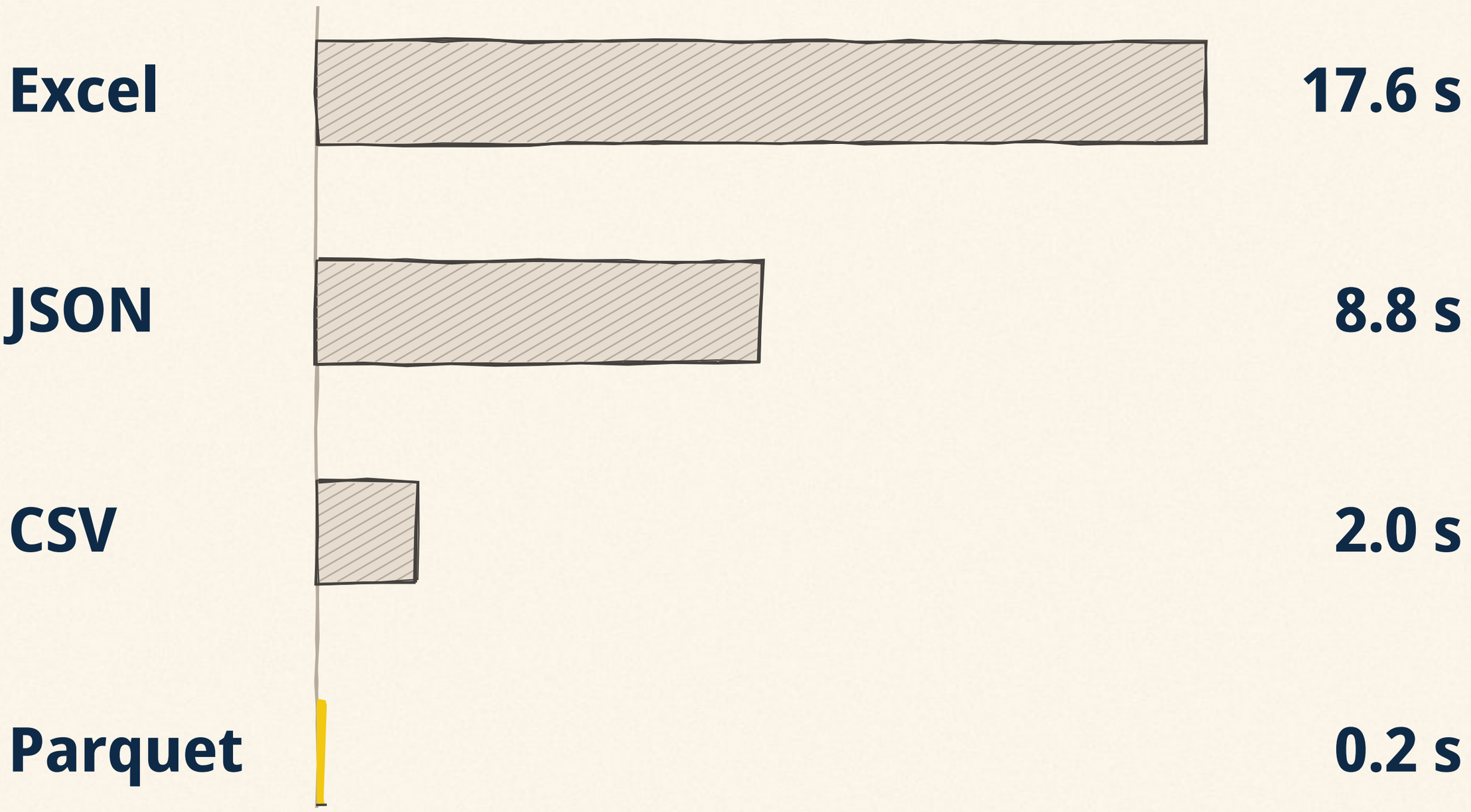
We saved 1,000,000 general-ledger lines in 11 formats. Here's what happened.



Same data.
Very different files.



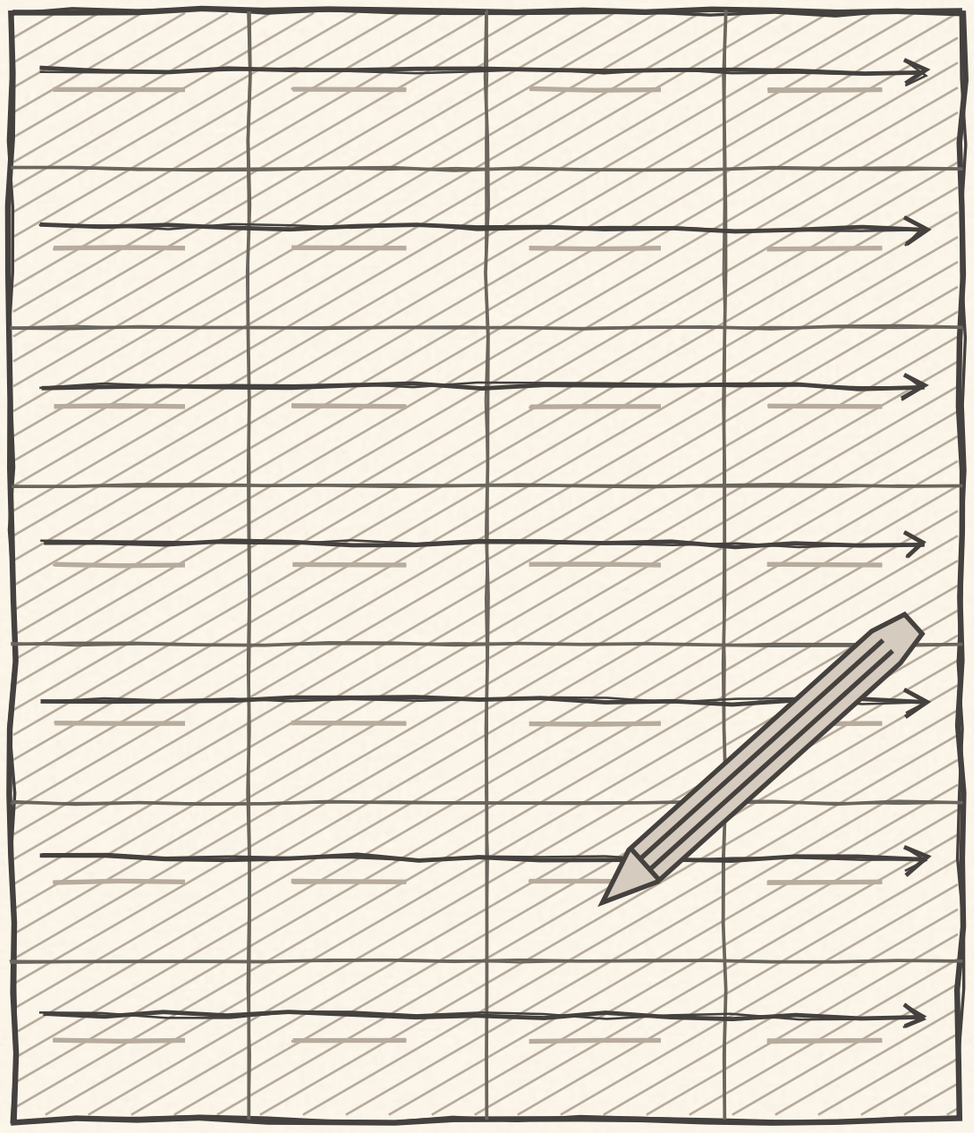
Time to open 1 million lines



Excel took 2 min 23 s just to SAVE the file.

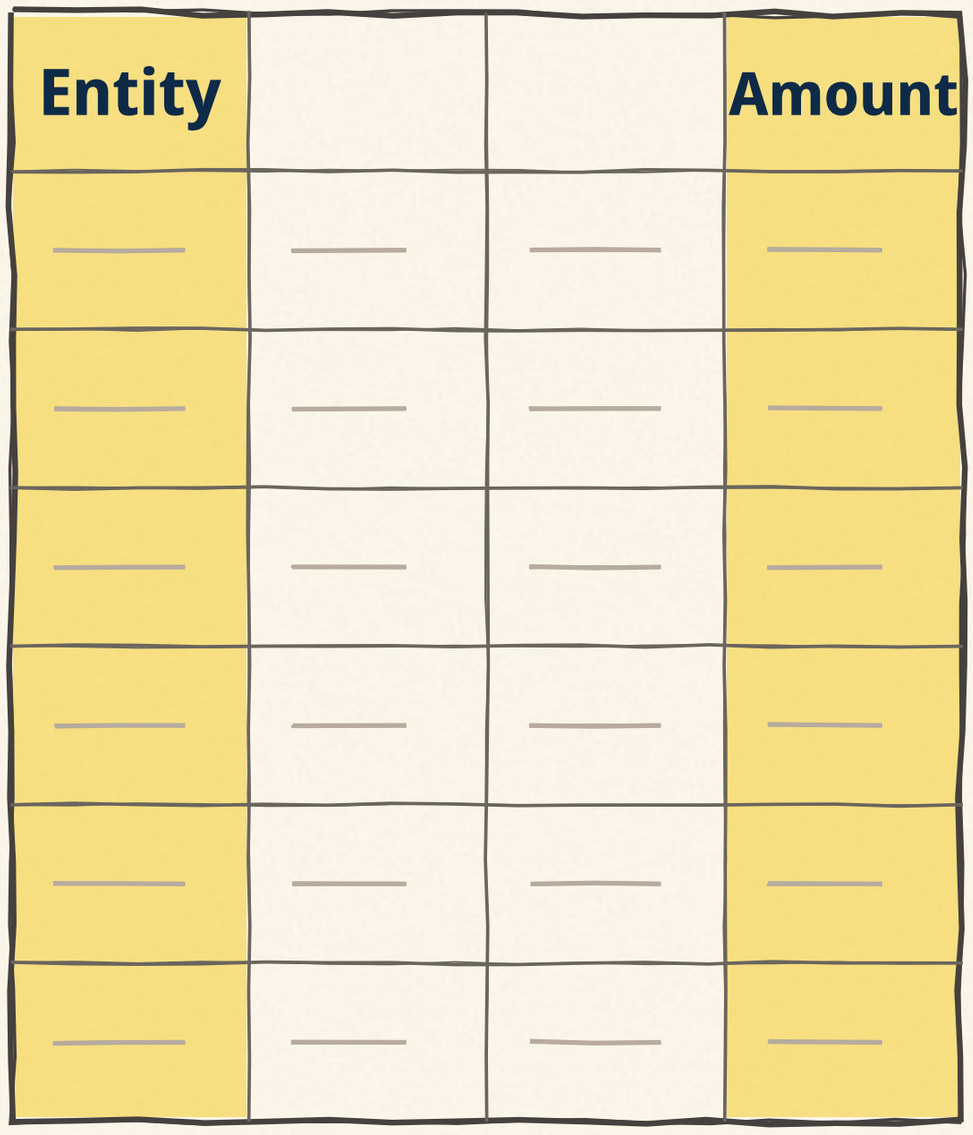
Why Parquet is fast: rows vs columns

CSV



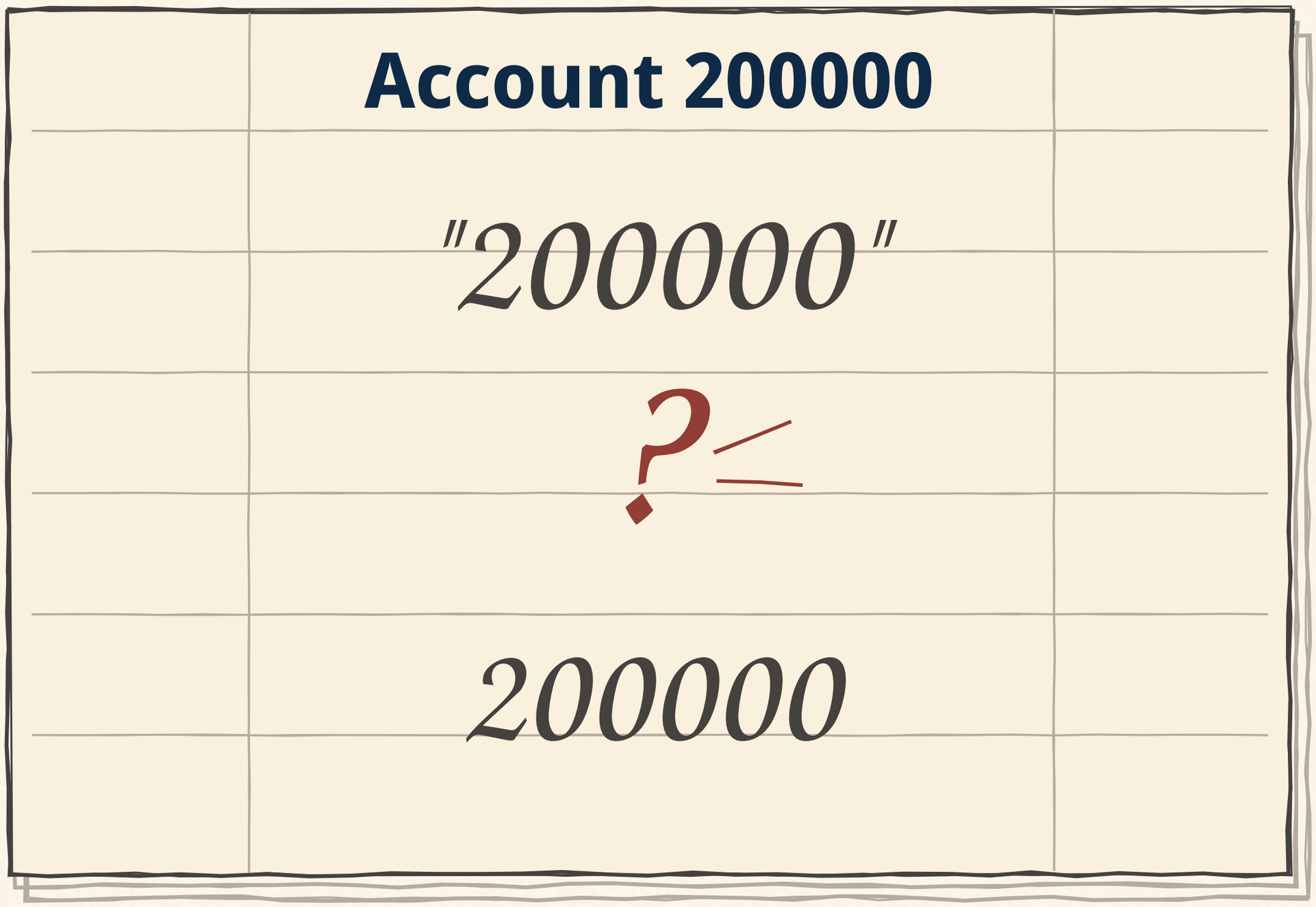
reads everything

Parquet



reads only what you ask for

The CSV trap



Read back with default settings, the same account code came back as text in part of the file and as a number in the rest. Your lookup quietly misses rows. Parquet keeps the type you saved.

50 years of data files

1972	**CSV (IBM Fortran)**
1985	**Excel .xls (Mac 1985, Windows 1987)**
2001	**JSON**
2007	**Excel .xlsx (1,048,576 rows)**
2013	**Parquet (Twitter + Cloudera)**
2016	**Apache Arrow**
2019	**Delta Lake**
Today:	**Microsoft Fabric**

Where Power BI and Fabric fit

Power BI

Get data

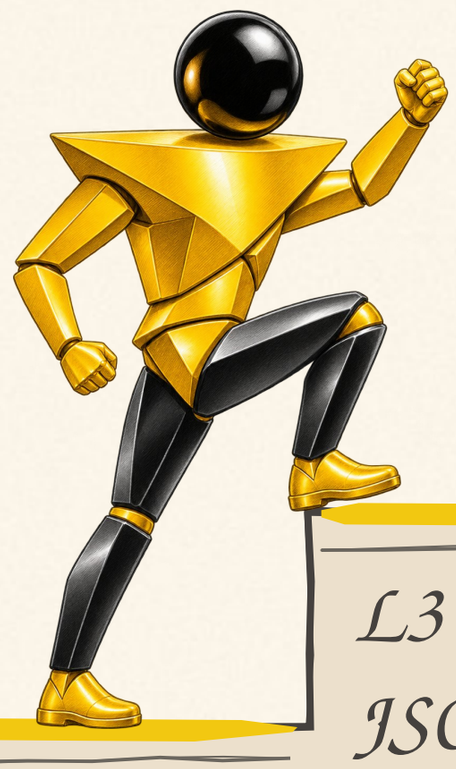
- Excel
- CSV
- JSON
- Parquet



Fabric OneLake — every table stored as Delta Parquet

Learn Parquet now — it's what Fabric runs on.

Practise every format on real finance data



L1 · Excel workbook

L2 · CSV tables

L3 · Parquet + JSON feeds

L4–L5 · millions of lines in Parquet

DATA DARE — real ledgers, graded questions

Which format does your team still live in?

- **Full test + charts:** findi360.com/data-resources/csv-vs-parquet-excel-json
- **Practise each format:** findi360.com/data-dare

findi360 — a dedicated platform for finance professionals to learn data and finance.