Data story · DuckDB vs MongoDB

One stores your app.
One analyses your ledger.

DuckDB vs MongoDB: two databases, two different jobs — and which one a finance team needs.

```
{
"order": 1042,
"customer": "Ali",
"items": [ ... ],
"gift_note": "..."
}
```

MongoDB: a document

Entity	Period	Amount

DuckDB: ledger columns

reads only what it needs



Where they came from

Two origin stories

MongoDB

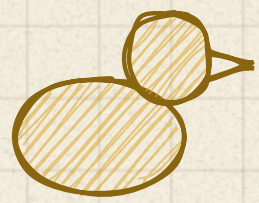
New York, 2007. 10gen set out to build a cloud platform. The database it needed did not exist — so it built one, kept it, and dropped the platform.

Open source: Feb 2009. Renamed MongoDB Inc. in 2013.

DuckDB

Amsterdam, 2018. Two researchers at CWI — Mark Raasveldt and Hannes Mühleisen — built it for data people who "loved data but hated databases".

v0.1: 2019. Named after Hannes's pet duck, Wilbur.



Wilbur

2009 → web apps needed flexible records

2018 → analysts needed SQL without a server

Sources on page 9

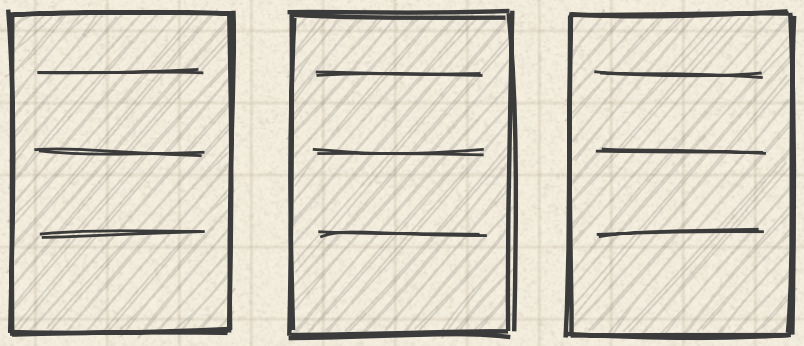
What MongoDB solved

Records that don't fit a table

```
{ "user": "Sara",
"city": "Lahore" }
```

```
{ "user": "Omar",
"city": "Riyadh",
"loyalty": "gold",
"devices": 3 }
```

new fields — no redesign



spread across servers

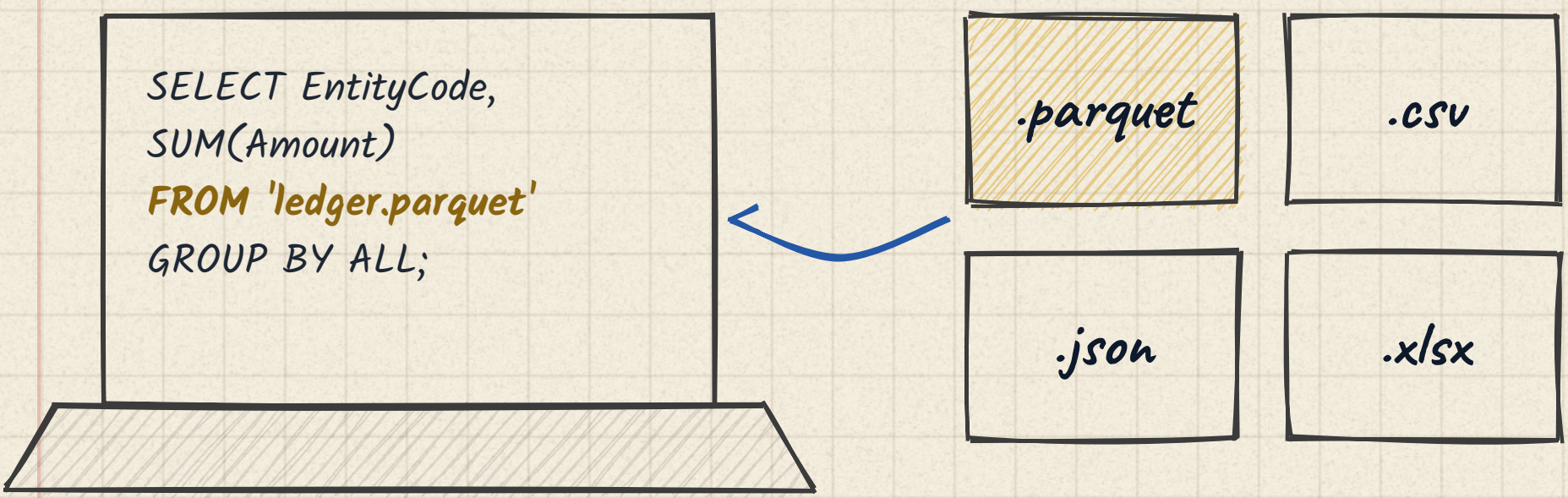
Web apps change every week. MongoDB stores each record as a flexible document — a new field needs no redesign — and spreads the data across many servers.

Built for: *live orders, users, app events.*
Harder for: *joins and big analysis across documents.*

Licence: SSPL since 16 Oct 2018 (not OSI-approved)

What DuckDB does now

SQL on your files. No server.



no import · no server · no setup

DuckDB runs inside your laptop's Python, R or command line and reads CSV, Parquet, JSON and Excel files directly — only the columns each question needs.

***2.0 (21 Oct 2026):** biggest release yet · **Aug 2026:** its company joins AWS, stays open source · **1M+ downloads a day**.*
***Not for:** being an app's live database.*

The evolution

From tables to ducks

1970 Relational model (Codd)

2000 SQLite — a database inside the program

← *big data went to clusters*

2006 Hadoop — big data on clusters

2009 MongoDB open source — the NoSQL wave

2018 MongoDB → SSPL licence · DuckDB started at CWI

2019 DuckDB v0.1 at SIGMOD

2024 DuckDB 1.0

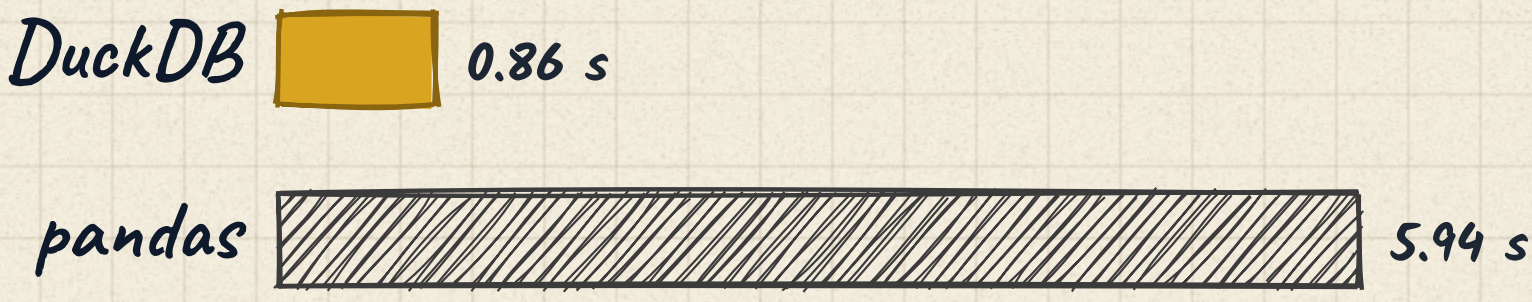
2026 DuckDB joins AWS (Aug) · DuckDB 2.0 (Oct)

...and came back to the laptop

The findi360 test · 6,884,744 GL lines

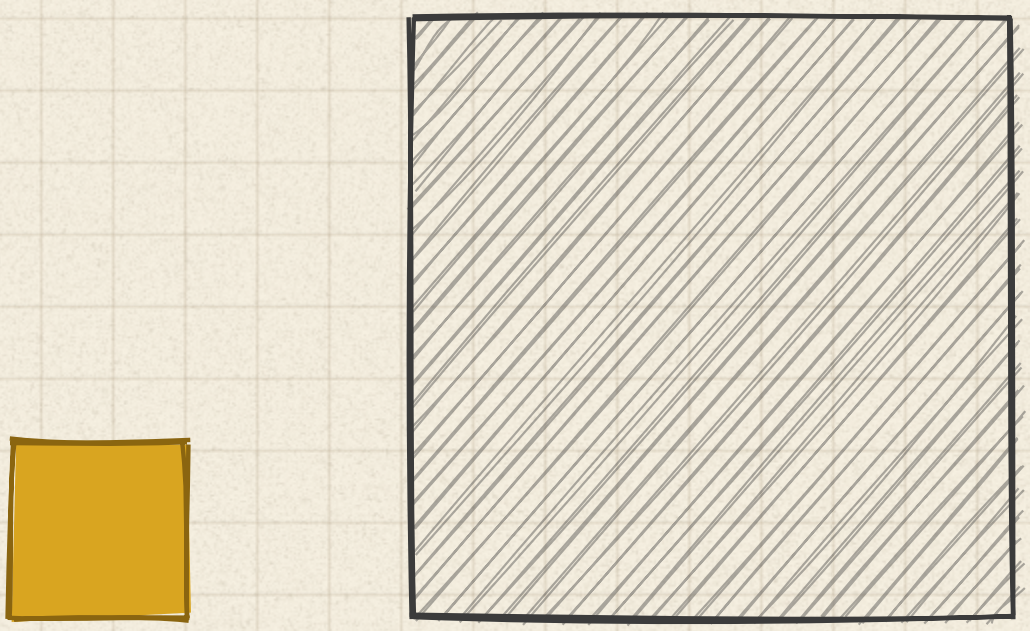
5 finance queries. Same answers.

Time for all 5 queries (seconds)



pandas includes 0.99 s to load the file first

Peak memory (areas to scale)



≈ 7× faster

≈ 12× less memory

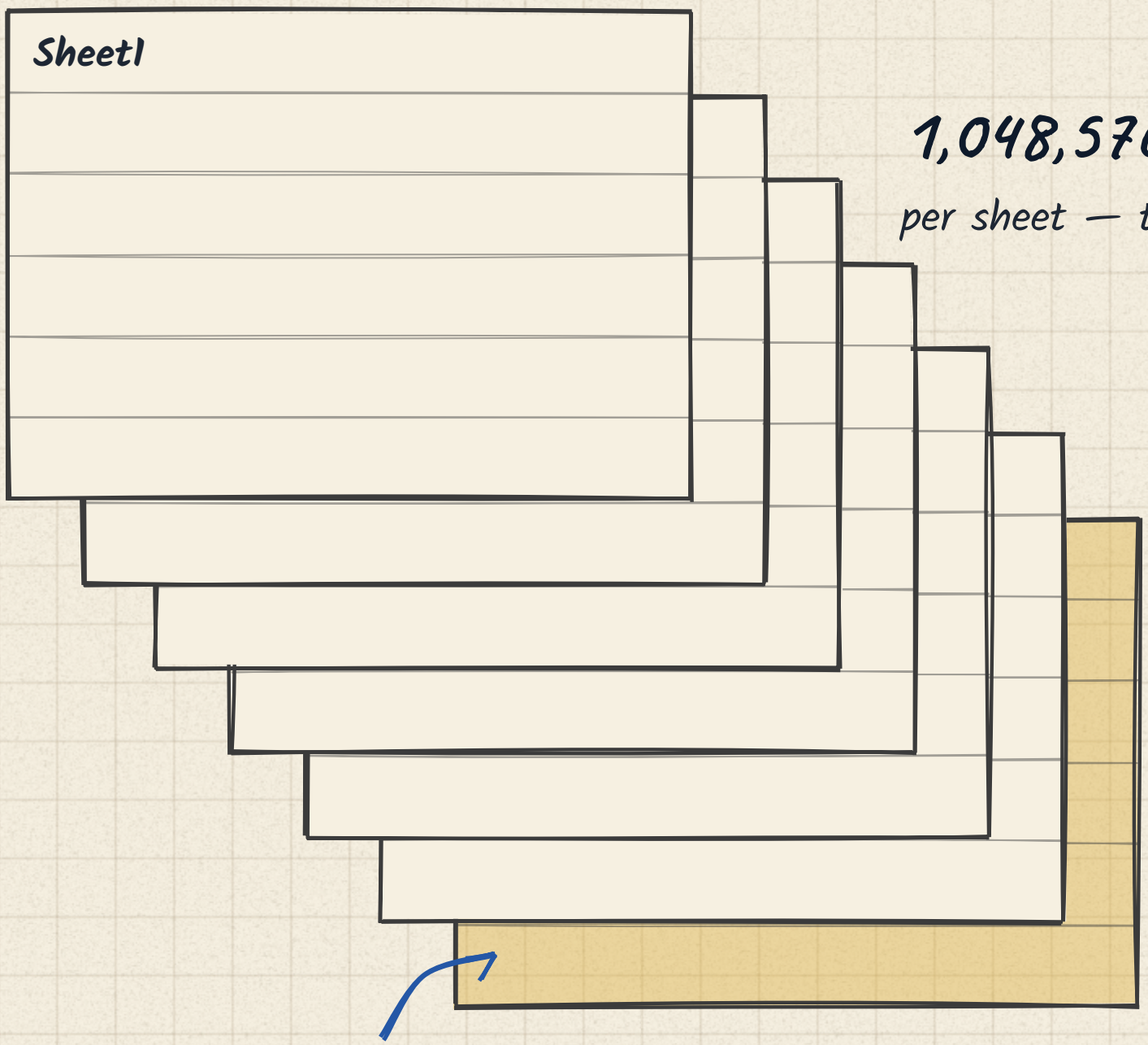
all 5 answers agree ✓

DATA DARE DD-CON-004 (Level 4). Median of 5 runs, each engine in a fresh process. 2 vCPU, 7 GB RAM — weaker than most office laptops. DuckDB 2.0.0 vs pandas 3.0.5.

Your laptop, your numbers: notebook in the data pack

And Excel?

This ledger needs 7 sheets



↑ the 7th sheet still holds 593,294 lines

A worksheet stops at **1,048,576 rows**. 6,884,744 lines → 7 sheets, and every formula across them. Excel's Data Model (Power Pivot) holds more — the same engine as Power BI.

Microsoft: Excel specifications and limits

Which one does a finance team need?

Each tool, its job

DuckDB	analyse big ledgers & files on a laptop
MongoDB	store live app records (a source system)
Excel	small models, what-ifs, presenting
Power BI	governed reports for the team
Python	automate and analyse — DuckDB runs inside it

Is your finance team still analysing in Excel — what's stopping you?

The full story, the data pack and the discussion:
findi360.com/data-resources/duckdb-vs-mongodb

Comment below, or answer on findi360

Sources & dates

Where the facts come from

DuckDB origin 2018 at CWI; v0.1 at SIGMOD 2019; joining AWS (26 Aug 2026)	duckdb.org/history	2026
DuckDB 2.0.0 LTS	duckdb.org	21 Oct 2026
DuckDB 1.0.0	duckdb.org	3 Jun 2024
1M+ downloads a day; stays open source	Help Net Security	Aug 2026
"Loved data but hated databases"	MotherDuck (Hannes Mühleisen)	6 Mar 2023
Named after the duck Wilbur	CWI	23 May 2023
10gen 2007; MongoDB open source Feb 2009; renamed 2013	Wikipedia — MongoDB Inc.	2026
MongoDB SSPL from 16 Oct 2018	MongoDB press release	16 Oct 2018
Excel: 1,048,576 rows per worksheet	Microsoft Support	2026
Test: 6,884,744 GL lines (DATA DARE DD-CON-004)	findi360 data dictionary	Oct 2026
Timings and memory (DuckDB 2.0.0, pandas 3.0.5)	findi360 test	16 Nov 2026



Visit findi360 to learn data and finance — www.findi360.com