Real Time Finance Data Extraction

Adam Cooke

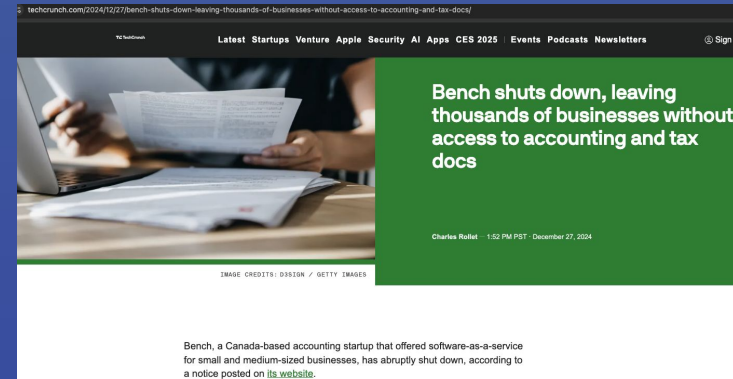adamcooke@cloudslurp.com

A simple.. but annoying problem.

Financial Institution Logins:
1. Average Accounts Person: 2 to 4
2. Average business: 5 to 10
3. Larger businesses: 10+

Average session duration: 8 to 12 minutes

Average 300 sessions per year.
People want quick customised answers they trust. **Usually expressed via 'I want it in Excel'**

techcrunch.com/2024/12/27/bench-shuts-down-leaving-thousands-of-businesses-without-access-to-accounting-and-tax-docs/

TC TechCrunch   Latest Startups Venture Apple Security AI Apps CES 2025 | Events Podcasts Newsletters   Sign In

Bench shuts down, leaving thousands of businesses without access to accounting and tax docs

Charles Rollet · 1:52 PM PST · December 27, 2024

IMAGE CREDITS: D3SIGN / GETTY IMAGES

Bench, a Canada-based accounting startup that offered software-as-a-service for small and medium-sized businesses, has abruptly shut down, according to a notice posted on its website.

- Cloud != Backup (downtime is common)

Cloudslurp

r/ireland · 2 yr. ago
whatThisOldThrowAway

Has anyone found a not-torturous way to export your transaction history from Bank of Ireland to a useful format (CSV, Excel, etc)

I've got a BOI current account with 10+ years of spending history.

I want to do some basic analytics on my spending and saving history: What have I historically spent my money on, how has that changed over time, what affects saving, etc. Basic stuff.

r/SwissPersonalFinance · 1 yr. ago
Nepopotamus

Quality of account transaction exports from online banking - am I the only one wondering?

I am a big fan of occasionally analyzing personal expenses in Excel, because only there you can really calculate your net worth across all assets, can create pivots to find tax relevant stuff etc.

Just wondering if banks provide bad exports on purpose? Whats the deal with these "Sammelbelastungen", where banks group transactions into one big block? Do they hate customers?

r/Bookkeeping · 2 yr. ago
joyfulstocks

Automating bank transactions into Google Sheets, Plaid or something else?

Software

I have a small bookkeeping business for friends and family. I have been pulling bank account CSVs for a while now and importing it into my system I have in google sheets. However, I am interested in pursuing an automated way to have bank transactions pull into a google sheet.

I have come across the idea of paying for a service (Tiller), but I see that most of these services use Plaid. It seems to me that it would be possible to have Plaid set up in google sheets for about 20-50 accounts for free. Would it be possible? If so, what are the first steps to setting it up?

If you have any other solutions, I would love to learn what is working for you!

r/excel · 1 yr. ago
blancjua

Is it possible to download live bank feeds into Excel?

unsolved

Say I want to build my own database and pull live data from a checking account and credit cards, is this possible?

reddit   r/plaintextaccounting   Search

r/plaintextaccounting · 10 mo. ago
Budget_Putt8393

Automated downloads of bank data

I have 6 banks. Manually exporting the ofx files is getting annoying.

Are there any options to skip the manual process of exporting bank info.

r/singapore · 2 yr. ago
kirso

How do you extract your bank data for expense/investment tracking?

Discussion

Curious if anyone figured out a way to extract bank data into Google Sheets or Microsoft Excel for own budgeting or other purposes?

I know there are some apps like Spendee and YNAB but I don't want to give them my data and prefer to do it in the sheets (which is also customizable). Also sick of manual entry, its 2022 and banks still don't have open APIs to do this.

r/Bookkeeping · 2 yr. ago
joyfulstocks

Automating bank transactions into Google Sheets, Plaid something else?

Software

I have a small bookkeeping business for friends and family. I have been pulling bank account and importing it into my system I have in google sheets. However, I am interested in pursuing to have bank transactions pull into a google sheet
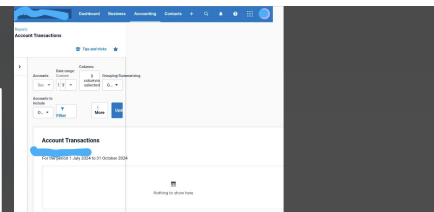
r/xero · 4 mo. ago
Reginleif69

Does anyone else find Xero almost completely unusable outside of using an actual computer?

r/xero · 3 yr. ago
Hungry-Moose-121

Xero frustrations

Longtime Xero bookkeeper here - I have a bookkeeping firm and use Xero for the bulk of my clients. Over the past few years, I've found that the customer support has gotten really bad, and now the bank feeds are glitchier for local banks than I'm comfortable with (Note that I would never recommend my clients change banks to a bigger one).

THE PAIN IS REAL

Cloudslurp...What is it?

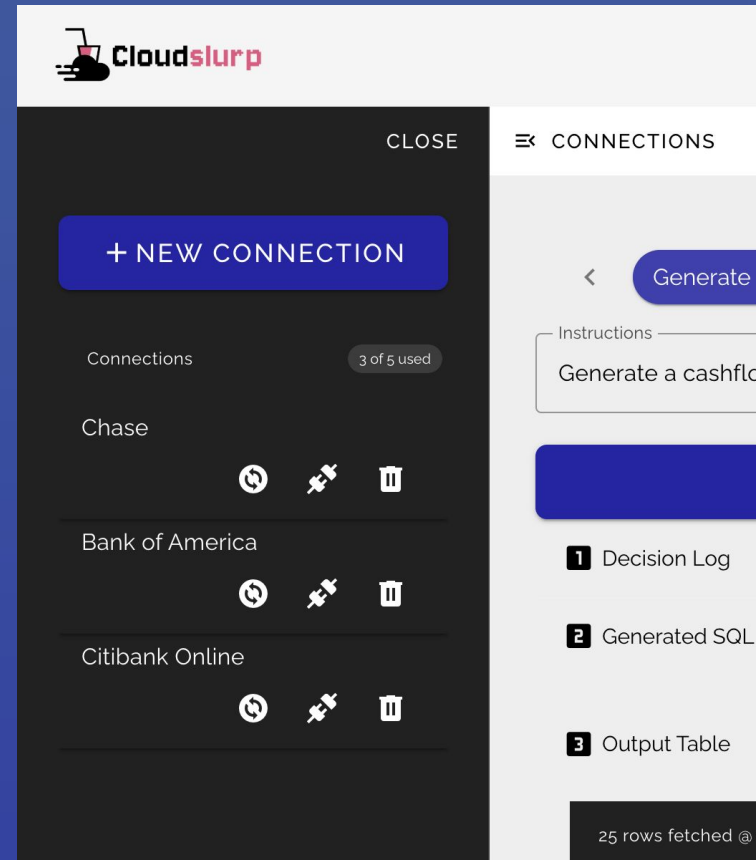Generates 'live' data tables you can embed

1. Bank transactional data sync
2. Combines transactions into one table
3. Provides third party backup
4. Instructions to SQL Generator (Web UI + Auditor Agent)
5. 'Slurps' Data API

SLURP - 'A template of a SQL query' that is run on every API call (fetching the data from source when possible).

- Complexity is fine as long as its explainable and logged
- You should always own your own data

Combine and Backup
Data Across
Institutions and SAAS tools

- Build up long history of transactions
- Backed up location
- Combine accounts from multiple people

Cloudslurp

CLOSE   CONNECTIONS

+ NEW CONNECTION

< Generate

Instructions
Generate a cashflo

Connections   3 of 5 used

Chase

Bank of America

Citibank Online

1 Decision Log

2 Generated SQL

3 Output Table

25 rows fetched @

Instant Customised Financial Statements

Cloudslurp

| ‹ | Generate a Cashflow Statement | List All Transactions | Categorise Expenses | Add Extra Columns | Do Complex Formatting | › |

Instructions

Generate a cashflow statement

⬇ **GENERATE DATA**

1 Decision Log   ⌄

2 Generated SQL   ⌄

3 Output Table   ⌃

25 rows fetched @ 2025-01-30 18:51:17   ⬇ 🅧 ⊞ API ◎

month	inflow	outflow	net_flow
2023-01	30.4	0	30.4
2023-02	6381.12	1012.6600000000001	5368.46

Use Case: Complex Formatting

list the top 5 largest inflow transactions and round the numbers to 0 decimal point. Also capitalise the first all the transaction names and add a column for category and make them upper case. Include the date and format it in dd/mm/yyyy

date	transaction_name	amount	category
28/01/2025	AUTOMATIC PAYMENT - THANK	2079	"GENERAL_MERCHANDISE_OTHER_GENERAL_MERCHANDISE"
29/12/2024	AUTOMATIC PAYMENT - THANK	2079	"GENERAL_MERCHANDISE_OTHER_GENERAL_MERCHANDISE"

T & Cs

Use Case: Data Enrichment

Instructions

Classify each month by high medium low based on revenue

month	revenue	revenue_category
2023-01	30.4	low
2023-02	5368.46	medium

Full SQL Editor Access (♡ SQL)

Cloudslurp

<   Generate a Cashflow Statement   List All Transactions   Categorise Expenses   Add Extra Columns   Do Complex Formatting   >

Instructions
Generate a cashflow statement

⬇ GENERATE DATA

1 Decision Log   ⌄

2 Generated SQL   ⌃

```sql
SELECT
    strftime('%Y-%m', t.date) AS month,
    SUM(CASE WHEN t.amount > 0 THEN t.amount ELSE 0 END) AS inflow,
    SUM(CASE WHEN t.amount < 0 THEN -t.amount ELSE 0 END) AS outflow,
    SUM(t.amount) AS net_flow
FROM
    Transactions t
GROUP BY
    strftime('%Y-%m', t.date)
ORDER BY
```

Decision Log (For Auditors)

Cloudslurp

| < | Generate a Cashflow Statement | List All Transactions | Categorise Expenses | Add Extra Columns | Do Complex Formatting | > |

Instructions

Generate a cashflow statement

↓ GENERATE DATA

1 Decision Log ⌃

{'time':'2025-01-30 19:35:49,103', 'name': 'aimodel',   'level': 'DEBUG', 'details': 'RUN 69ed242b-9ca7-494c-80af-6710738c16e6: THOUGH
T Thought: I need to know the schema of the Transactions table to write a valid query. Action: transaction_table_schema Action Input: na '}
{'time':'2025-01-30 19:35:49,773', 'name': 'aimodel',   'level': 'INFO', 'details': 'RUN 69ed242b-9ca7-494c-80af-6710738c16e6: START AC
TION transaction_table_schema'}
{'time':'2025-01-30 19:35:53,911', 'name': 'aimodel',   'level': 'DEBUG', 'details': 'Observation: ('account_id', 'VARCHAR', 'YES', Non
e, None, None)('account_owner', 'VARCHAR', 'YES', None, None, None)('amount', 'DOUBLE', 'YES', None, None, None)('authorized_date', 'DATE', '
YES', None, None, None)('authorized_datetime', 'TIMESTAMP', 'YES', None, None, None)('check_number', 'VARCHAR', 'YES', None, None, None)('cou
nterparties', 'STRUCT(confidence_level VARCHAR, entity_id VARCHAR, logo_url VARCHAR, "name" VARCHAR, phone_number VARCHAR, "type" VARCHAR, we
bsite VARCHAR)', 'YES', None, None, None)('date', 'DATE', 'YES', None, None, None)('datetime', 'TIMESTAMP', 'YES', None, None, None)('iso_cur
rency_code', 'VARCHAR', 'YES', None, None, None)('location', 'STRUCT(address VARCHAR, city VARCHAR, country VARCHAR, lat FLOAT, lon FLOAT, po
stal_code VARCHAR, region VARCHAR, store_number VARCHAR)', 'YES', None, None, None)('logo_url', 'VARCHAR', 'YES', None, None, None)('merchant
_entity_id', 'VARCHAR', 'YES', None, None, None)('merchant_name', 'VARCHAR', 'YES', None, None, None)('name', 'VARCHAR', 'YES', None, None, N
one)('payment_channel', 'VARCHAR', 'YES', None, None, None)('payment_meta', 'STRUCT(by_order_of VARCHAR, payee VARCHAR, payer VARCHAR, paymen

Embed Real-Time Data

Cloudslurp

Download | Excel | **Google Sheet** | API | ChatGPT

Live Data in Google Sheets

1. Copy/Paste this formula in a cell.
2. Live data will sync at up to 1 hour intervals.

```excel
=importData("https://yit0goz324.execute-api.
```

Duck Examination

In the browser for local query execution (duckdb-wasm)

In the cloud for pre-processing and aggregation (duckdb-node-neo)

Embedded Duckdb inside a python process as tooling for our LLM. (We are currently using Llama, also works with Gemini and Claude).

3 ducks for the implementation price of 1

Theoretical path to be 100% offline and give users tech vendor independence, no external API calls apart from inbound data. Browser based SQL + Inference.

- Less security risk
- Less compute cost
- More ownership
- Less pipeline administration

Duckdb - Helpful Features

1.Node neo type bindings

2.Support for S3 secret & secretmanager

4.Parallel union_by_name

5. Export statement different format support

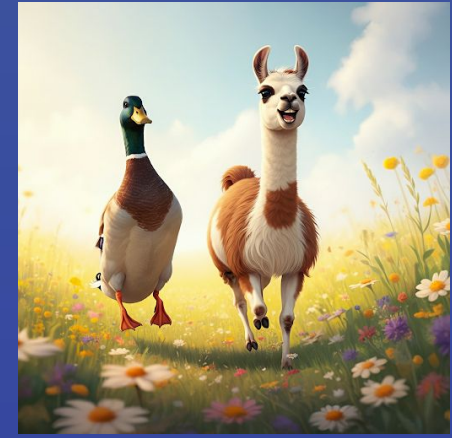6. The text based readme docs (Please keep them up to date :)

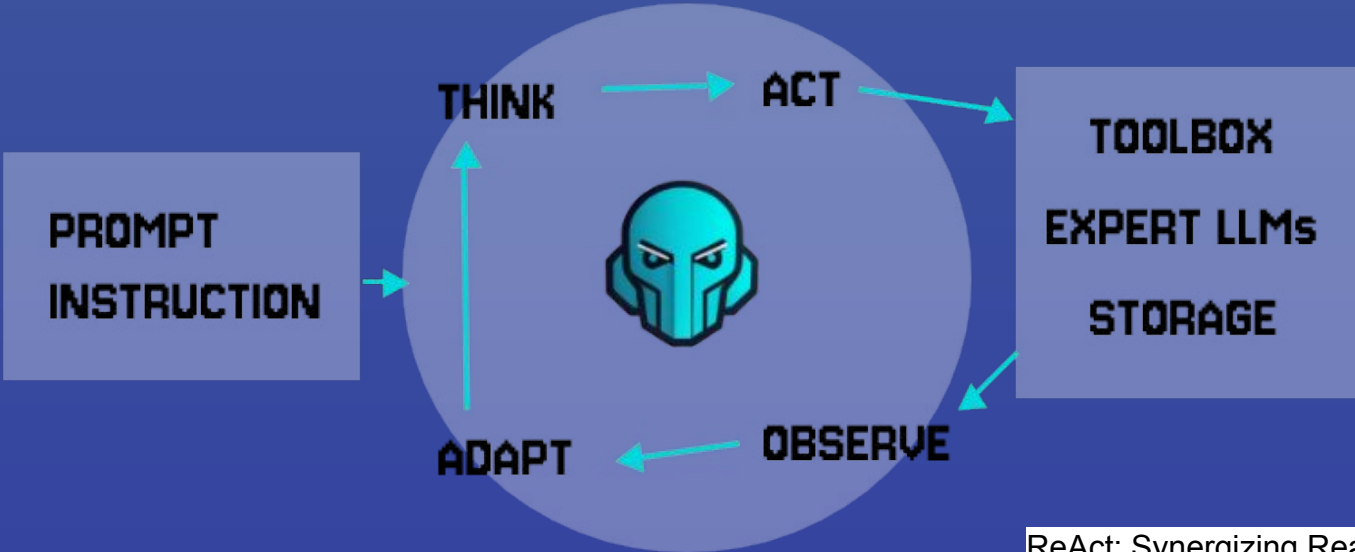Some suggestions

1. Direct JSON object support in nodejs (like it works in WASM
2. union_by_name = true .. have some way to match types to simplify ad hoc analysis for where you don't know field names and types in advance (like being generated in a step by an llm)

Going beyond text-to-sql toward 'big instructions'

"For my transactions at food restaurants categorize them as 'healthy' or 'unhealthy' and how much have I spent on healthy vs unhealthy food this week in usd then send me an email."

ReAct: Synergizing Reasoning and Acting white paper:2022

FIN.

Cloudslurp

adamcooke@cloudslurp.com