

R Exchange 2026



TE TAI ŌHANGA
THE TREASURY

Taxmate

An R-Based Solution to Streamline Treasury Processes

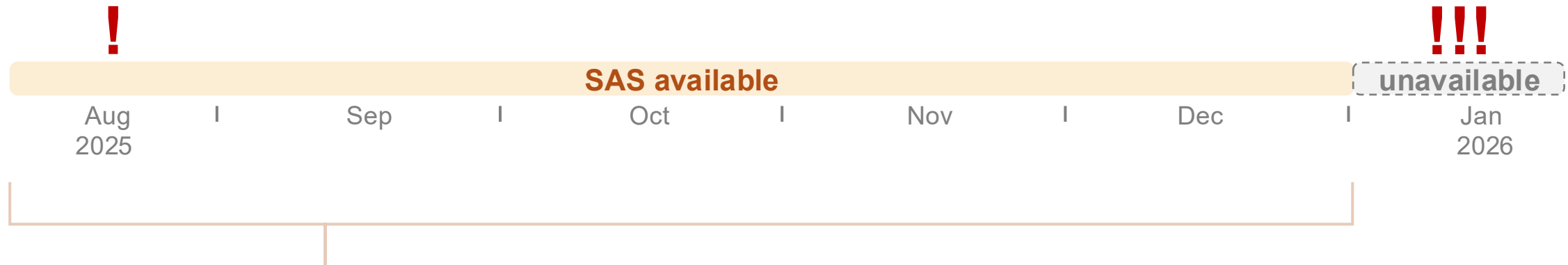
Yang Hu

SAS was used for tax-related work at the Treasury for decades.

(a commercial software)

In August 2025, we were informed that:

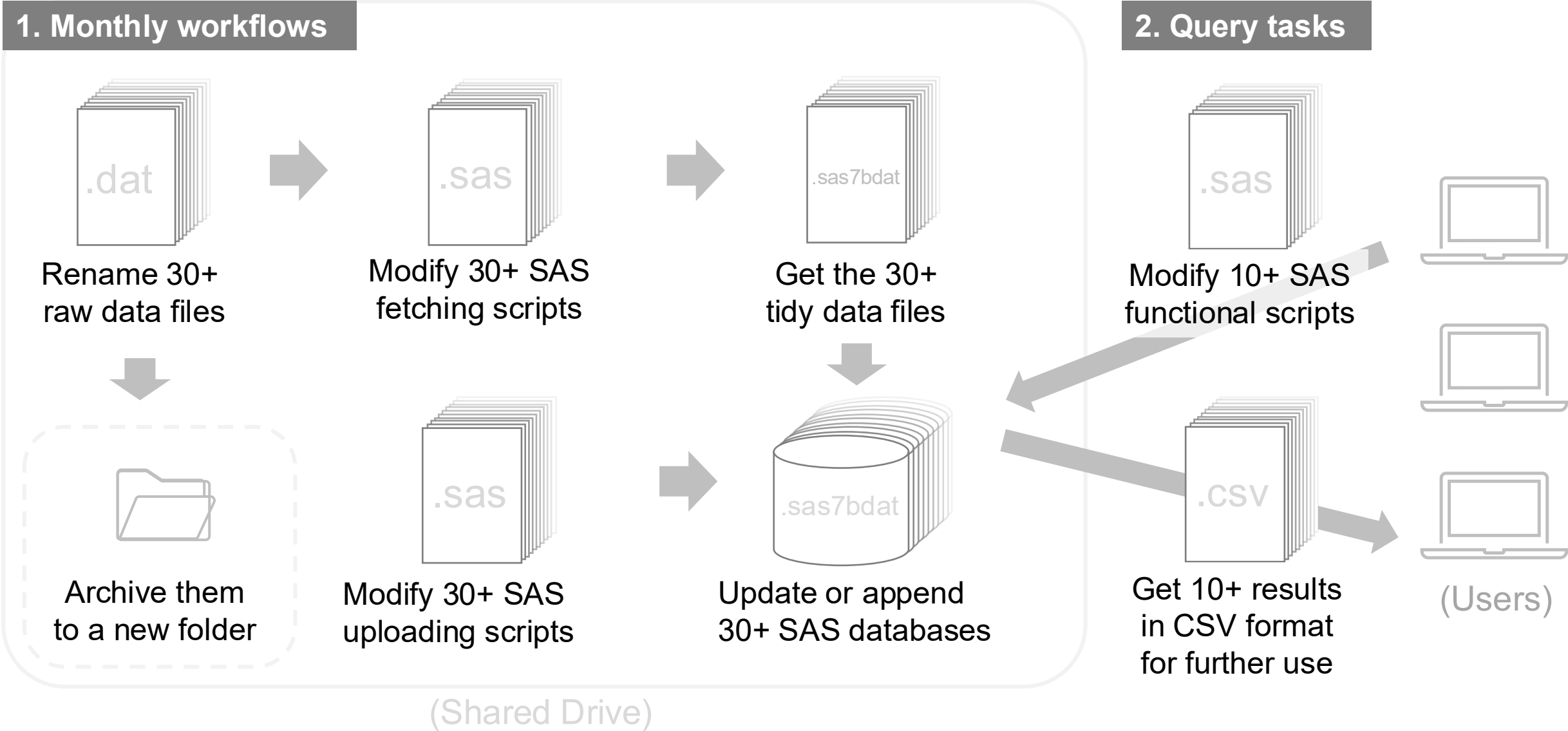
- The licence would **expire** at the end of the year (December 31st).
- From 2026 onwards, the yearly licence fee would have **tripled**.



Given the **short window**, a tough decision had to be made:

- **Accept the new quote, or**
- **Explore an alternative.**

Internal evaluation shows that SAS was used for **two specific types of work**:



SAS is used as a **tool** rather than a **platform**; this suggests an alternative is feasible:

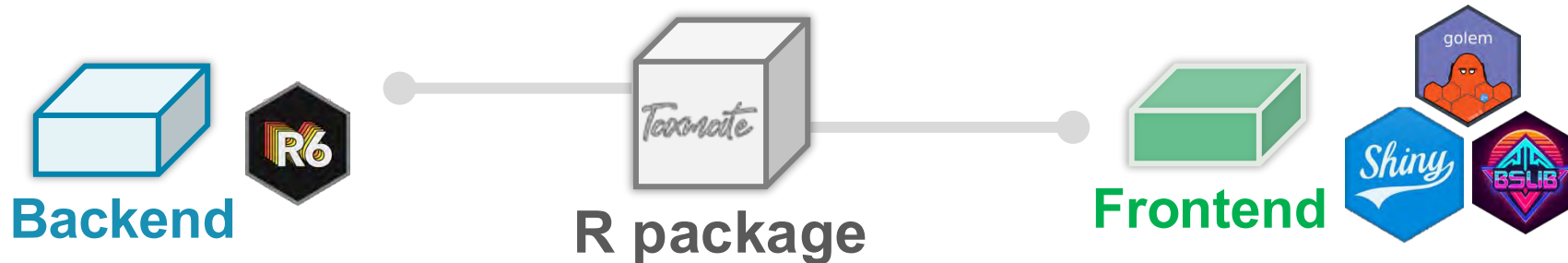
- Use **DuckDB** as the database



OLAP database
(Online Analytical Processing)

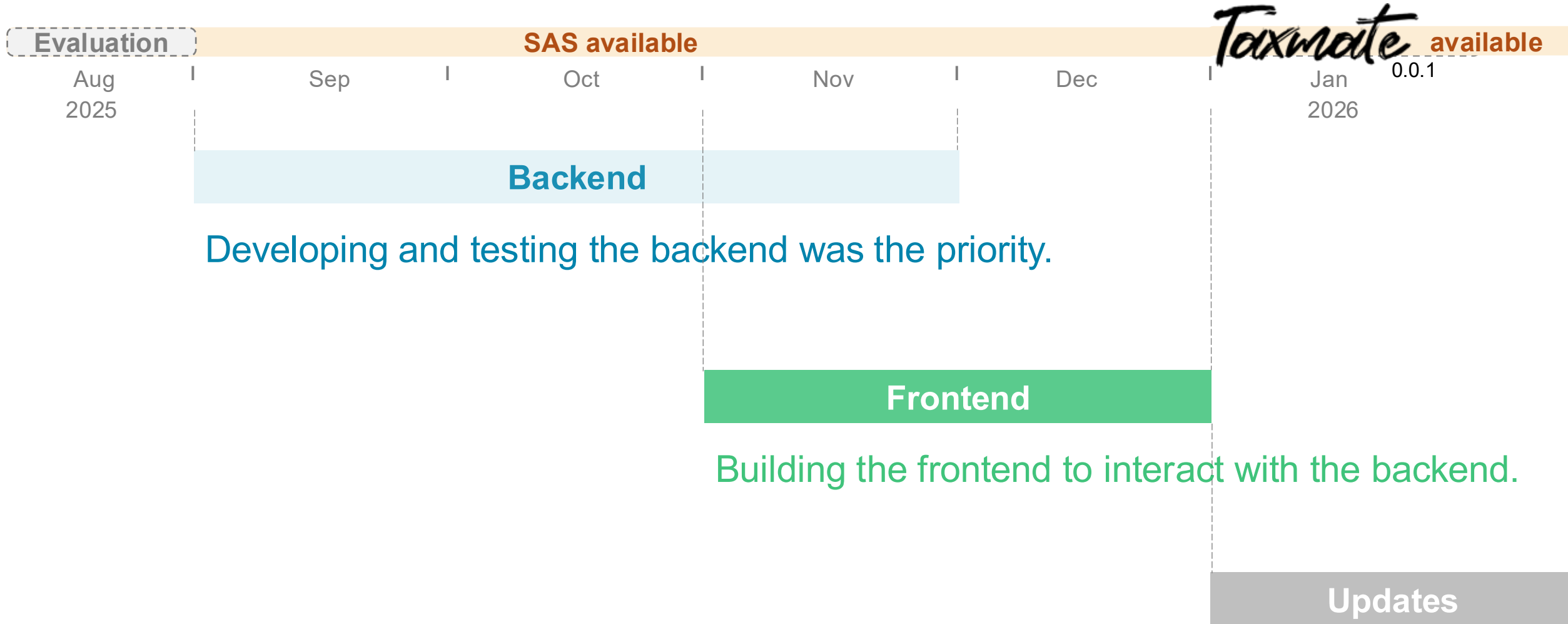
- ✓ Fast for analytics
- ✓ Efficient storage
- ✓ Handle data larger than memory

- Build an R package backend structured using the **R6** system



- Develop the **Shiny** app frontend using the **golem** framework with **Bootstrap 5**

However, the **development time** is very limited:

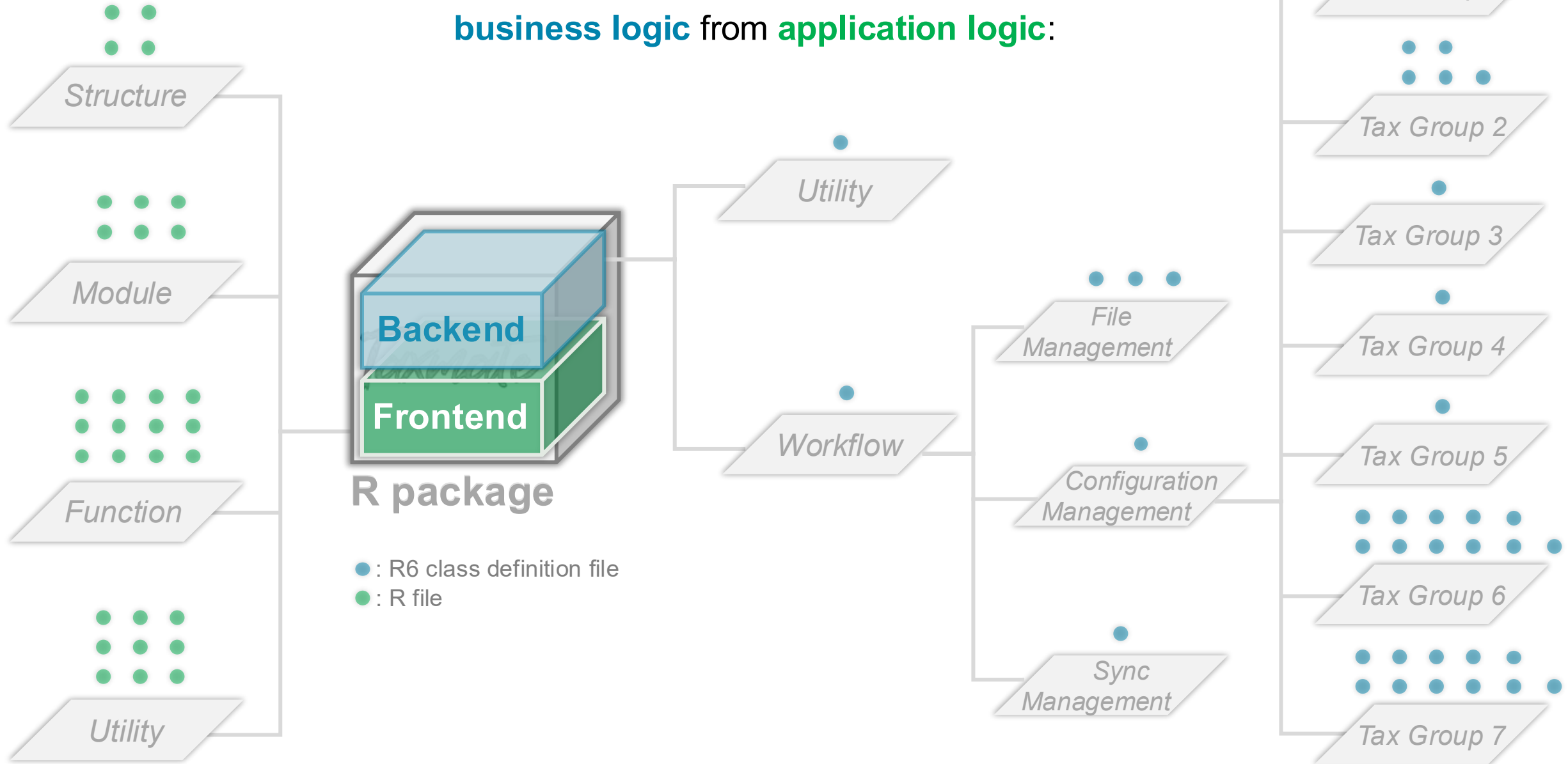


Developing and testing the backend was the priority.

Building the frontend to interact with the backend.

More functionality will be made available through package updates.

A 16,000-line R codebase is organised by role, separating **business logic** from **application logic**:



```
> library(taxmat|
```

taxmate

Import, Manage, and Explore Tax Data

taxmate is a system designed for automated workflows and database synchronization among users, based on DuckDB databases that provide extremely fast query speeds. Frequently used functions and SQL ports are integrated into the Shiny app, allowing users to complete tasks with simple clicks or use SQL to generate custom outputs.

Press F1 for additional help

```
> library(taxmate)
```

≡ NZ Treasury R Series ≡

— **taxmate¹: Import, Manage, and Explore Tax Data** —

Simply enter ``run_app()`` to launch the app

¹The package (version 0.0.4) is developed for the Forecasting and Monitoring team by the Analytics Insights and Modelling team. The repository is: https://dev.azure.com/nztreasury/Analytics_Insights/_git/taxmate. Please refer to the documentation in the repository for a complete introduction. Contact the maintainer Yang Hu [Yang.Hu@treasury.govt.nz] with any questions or feedback.

```
> |
```

G'day!

Workflow Automation 🤖

It renames new IRD files, extracts data, updates DuckDB, archives used files by year and month, and keeps the Received folder tidy.

No more sweaty manual work — it's all done with one click. Sit back with a coffee (or hojicha), watch the progress bar, and focus on what truly matters.

DuckDB Execution Engine 🐥

DuckDB is an in-process SQL database management system designed for fast analytical queries on large datasets.

Its parallel execution can process larger-than-memory workloads, allowing tax data to be queried at blazing-fast speeds.

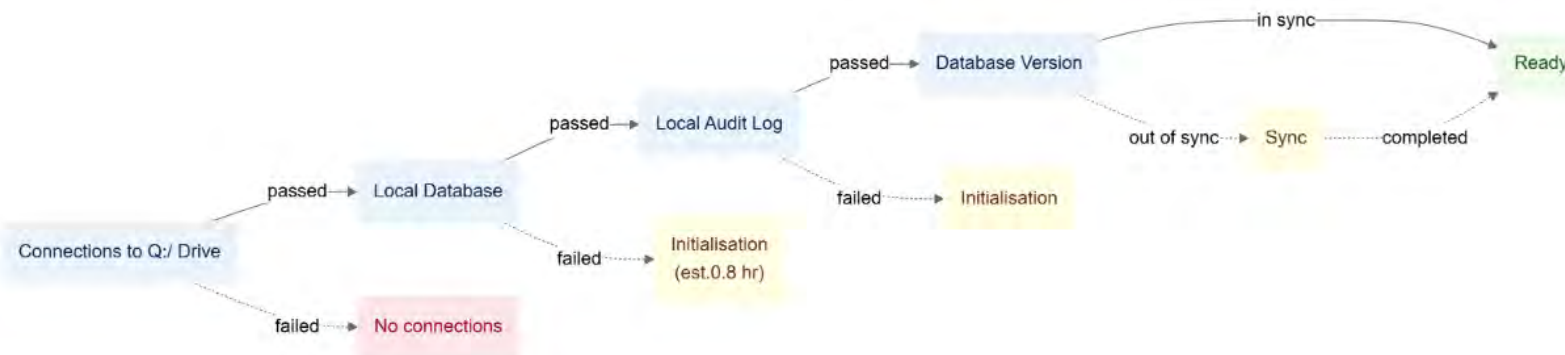
Decentralised Framework 🏠

Databases sit securely on each mate's laptop, while monthly snapshots and the audit log stay on the shared drive for optimal performance.

New files update the mate's local database who starts the automation, and the latest snapshot syncs for everyone else to keep databases aligned.

Pre-launch Checks ⚡

- First time using {taxmate}? Don't worry, {taxmate} will sort it out.¹
- Every time you press the start button, the backend checks that the local database, audit log, and snapshots are ready.
- Please keep R and the app open, avoid locking your screen, and wait until everything is done.



¹ Depending on the speed of the shared Q:/ drive, this may take up to an hour. You may need to start the app a few times to complete setup.



Start

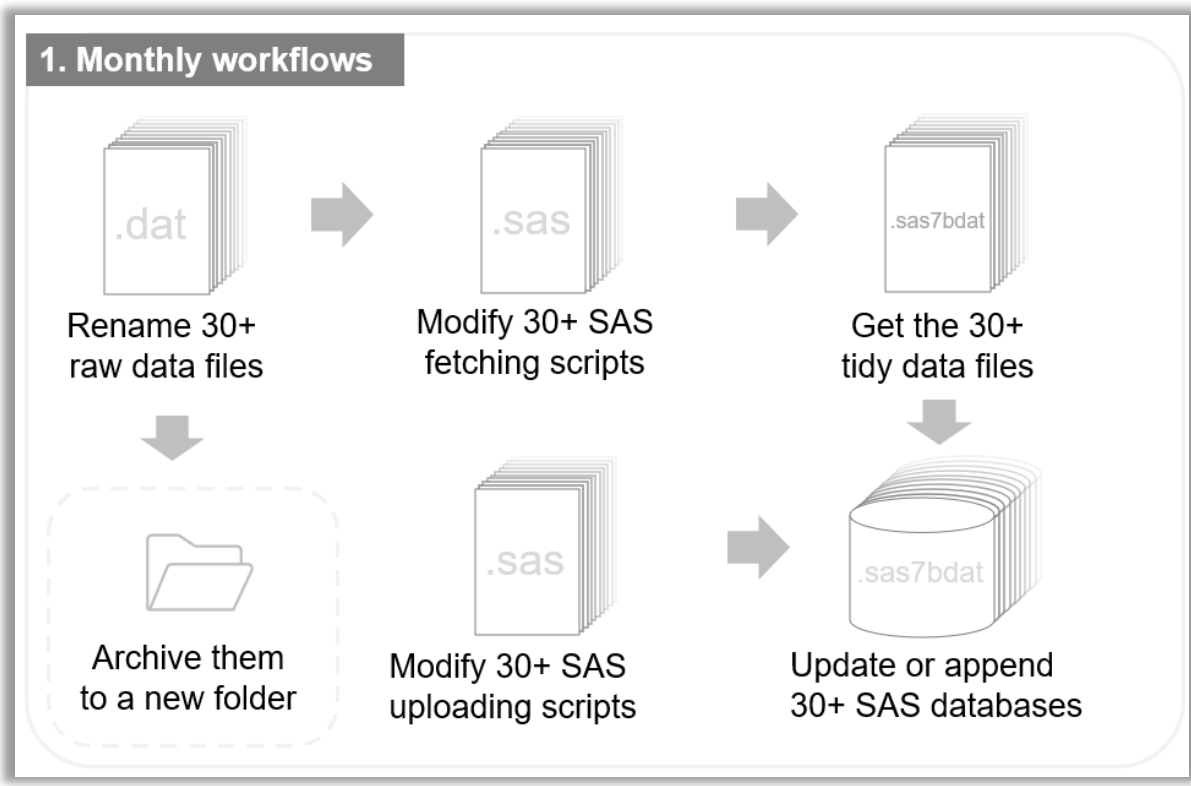
Corporate Tax Annual Returns Summary (2019 onwards)

2025

Hide raw names Rounded values

Run Est. time: <3s

IR4



New IR Data


Update Available

35 new files detected in the Received folder.

Process

Last checked at 09:29



 Workflow Automation

What's happening?

Taxmate is processing IRD files. Progress is shown in the bottom-right corner.

While running:

- Keep this app open
- Keep R running
- Prevent your laptop from sleeping.

What happens next?

This window will close once the task finishes. You can then use Taxmate as usual with the updated database.

 [2/14] Processing 'ail' (3.3 KB)

 Syncing database...

Task Query

Corporate Tax Annual Returns Summary (2019 onwards)

IR4

Missing files - Message (HTML)

File Message Insert Draw Options Format Text Review Help Tell me

The following recipient is outside your organization: [redacted]@ird.govt.nz

Send this email during your work hours: Tue, Apr 28 at 12:00 AM [Schedule send](#) | [Dismiss](#)

To: [redacted]@ird.govt.nz

Cc:

Subject: Missing files

Hi [redacted],

the following files are missing this month:

AIL (Withhold), NRTP (Withhold)

Cheers

Run Est. time: <10s

New IR Data

Up to Date

The local database is up to date. No action is needed.

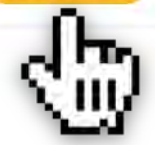
Last checked at 10:36

Missing Files

2 files are missing in the **Received** folder:

AIL (Withhold), NRTP (Withhold)

Send E-mail



Corporate Tax Annual Returns Summary (2019 onwards)

Company

- Corporate Tax Annual Returns Summary (2019 onwards)
- NZ Superannuation Fund Returns Summary
- Residual Income Tax by Industry (2019 onwards)

GST

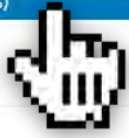
- GST Monthly Growth Summary

2025

Hide raw names

Run

Est. time: ~10s



Corporate Tax Annual Returns Summary (2019 onwards)

IR4

- 2025
- 2020
- 2021
- 2022
- 2023
- 2024
- 2025
- 2026

Hide raw names

Run
Est. time: <10s



Corporate Tax Annual Returns Summary (2019 onwards)

2025

Hide raw names

Run

Time taken: 2s

Download

IR4

Preview output

Show 10 entries

Search: _____

Name	Other	March	June	September	December	All
All	All	All	All	All	All	All

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Information Withheld

Showing 1 to 10 of 32 entries

Task Query

Corporate Tax Annual Returns Summary (2019 onwards)

2025

Hide raw names

Run Time taken: 2s

Download

IR4

Show 10 entries

Search:

Name	Other	March	June	September	December	All
All	All	All	All	All	All	All

1
2
3
4
5
6
7
8
9
10

Information Withheld

Showing 1 to 10 of 32 entries

Previous 1 2 3 4 Next

Task Query

Corporate Returns Summary (2019 onwards)

Company

- Corporate Tax Annual Returns Summary (2019 onwards)
- NZ Superannuation Fund Returns Summary
- Residual Income Tax by Industry (2019 onwards)

GST

- GST Monthly Growth Summary

2025 Hide raw names

Run
Est. time: ~10s

✓ Existing workflows are covered

But what about new tasks or database exploration?

Database: Table:

Run

```
Enter SQL
```

Column

Show 10 entries Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Database

1. Company
aim
ir10
ir4
ir44
ir4j

Table

aim



Run

Enter SQL

2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Previous 1 2 3 4 5 ... 10 Next

Database: aim Table: aim

Run

Column

Show 10 entries Search:

Variable	Type	Description
1	VARCHAR	IRD number
2	DOUBLE	Location
3	DATE	Last day of return period YYYY-MM-DD
4	retype	Return type
5	taxtype	Tax type
6	name	Name
7	balmth	Balance month (1-12)
8	balday	Balance day (1-31)
9	rsdncy	Residency status (R/N)
10	sic	ANZSIC code

```
Enter SQL
```

Database: Table:

Run

Column

Show 10 entries Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

```
SELECT COUNT (*)
AS total_rows
FROM aim;
```

Database: aim Table: aim

Column

Show 10 entries Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Copy SQL to clipboard

```
SELECT COUNT (*) AS total_rows FROM aim;
```

Run

Database: Table:

Run

```
SELECT COUNT (*)
AS total_rows
FROM aim;
```

Column

Show 10 entries Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Database: aim Table: aim

Column

Show 10 entries Search:

Variable	Type	Description
1 ird	VARCHAR	IRD number
2 loc	DOUBLE	Location
3 retperd	DATE	Last day of return period YYYY-MM-DD
4 retype	VARCHAR	Return type
5 taxtype	VARCHAR	Tax type
6 name	VARCHAR	Name
7 balmth	DOUBLE	Balance month (1-12)
8 balday	DOUBLE	Balance day (1-31)
9 rsdncy	VARCHAR	Residency status (R/N)
10 sic	VARCHAR	ANZSIC code

Run

```
SELECT COUNT (*)  
AS total_rows  
FROM aim;
```

Yang Hu [TSY] (You)

SELECT COUNT (*)
AS total_rows
FROM aim;

Shift+Enter starts a new line.

Database

aim

Table

aim

Column

Show 10 entries

Search: _____

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

  [Run](#)
[Download SQL as a TXT file](#)

```
SELECT COUNT (*)  
AS total_rows  
FROM aim;
```



Task Query

Database: aim Table: aim

Run

```
SELECT COUNT (*) AS total_rows FROM aim;
```

Column

Show 10 entries Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Previous 1 2 3 4 5 ... 10 Next

Database: aim Table: aim

Column: Show 10 entries Search:

Variable	Type	Description
1 ird	VARCHAR	IRD number
2 loc	DOUBLE	Location
3 retperd	DATE	Last day of return period YYYY-MM-DD
4 retype	VARCHAR	Return type
5 taxtype	VARCHAR	Tax type
6 name	VARCHAR	Name
7 balmth	DOUBLE	Balance month (1-12)
8 balday	DOUBLE	Balance day (1-31)
9 rsdncy	VARCHAR	Residency status (R/N)
10 sic	VARCHAR	ANZSIC code

Run

```
SELECT COUNT (*)  
AS total_rows  
FROM aim;
```

```
File Edit Format View Help  
SELECT COUNT (*)  
AS total_rows  
FROM aim;
```

Database: aim Table: aim

Run

```
SELECT COUNT (*) AS total_rows FROM aim;
```

Column

Show 10 entries Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Show 10 entries

1

Search:

total_rows

121278

Download



Database

aim

Table

aim

Column

Show 10 entries

Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Run

```
SELECT COUNT (*)
AS total_rows
FROM aim;
```

Show 10 entries

Search:

1

total_rows

121278

Download

Database: aim Table: aim

Run

Download

```
SELECT COUNT (*)  
AS total_rows  
FROM aim;
```

Show 10 entries

Search:

Column
Show 10 entries Search:

Variable	Type	Description
1 ird	VARCHAR	IRD number
2 loc	DOUBLE	Location
3 retperd	DATE	Last day of return period YYYY-MM-DD
4 retype	VARCHAR	Return type
5 taxtype	VARCHAR	Tax type
6 name	VARCHAR	Name
7 balmth	DOUBLE	Balance month (1-12)
8 balday	DOUBLE	Balance day (1-31)
9 rsdncy	VARCHAR	Residency status (R/N)
10 sic	VARCHAR	ANZSIC code

1

total_rows

121278

Next

```
SELECT COUNT (*)  
AS total_rows  
FROM aim;
```

People make mistakes — will a typo break the app?

Previous 1 2 3 4 5 ... 10 Next

Database Table

aim aim

Column

Show 10 entries Search:

Variable	Type	Description
1 ird	VARCHAR	IRD number
2 loc	DOUBLE	Location
3 retperd	DATE	Last day of return period YYYY-MM-DD
4 retype	VARCHAR	Return type
5 taxtype	VARCHAR	Tax type
6 name	VARCHAR	Name
7 balmth	DOUBLE	Balance month (1-12)
8 balday	DOUBLE	Balance day (1-31)
9 rsdncy	VARCHAR	Residency status (R/N)
10 sic	VARCHAR	ANZSIC code

Previous 1 2 3 4 5 ... 10 Next

Run

```
SELECTT COUNT (*)
AS total_rows
FROM aim;
```

Error: Parser Error: syntax error at or near "SELECTT"

LINE 1: SELECTT COUNT (*)
^

LINE 1: SELECTT COUNT (*)
^

[34mi [39m Context: rapi_prepare
[34mi [39m Error type: PARSER

Why I got this error: Parser Error: syntax error at or near "SELECTT" LINE 1: SELECTT COUNT (*) ^ LINE 1: SELECTT COUNT (*) ^ [34mi][39m Context: rapi_prepare [34mi][39m Error type: PARSER?

The error is simply a typo — you wrote SELECTT instead of SELECT (double "T").

```
sql
-- ✘ Wrong
SELECTT COUNT (*)

-- ✔ Correct
SELECT COUNT(*)
```

Reply...

+



One last thing

Task Query

Database

aim

Table

aim

Run

Download

Column

Show 10 entries

Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

Previous 1 2 3 4 5 ... 10 Next

```
SELECT COUNT (*)  
AS total_rows  
FROM aim;
```

Show 10 entries

Search:

1

total_rows

121278

Previous 1 Next

Database: Table:

Column

Show entries Search:

	Variable	Type	Description
1	ird	VARCHAR	IRD number
2	loc	DOUBLE	Location
3	retperd	DATE	Last day of return period YYYY-MM-DD
4	retype	VARCHAR	Return type
5	taxtype	VARCHAR	Tax type
6	name	VARCHAR	Name
7	balmth	DOUBLE	Balance month (1-12)
8	balday	DOUBLE	Balance day (1-31)
9	rsdncy	VARCHAR	Residency status (R/N)
10	sic	VARCHAR	ANZSIC code

[Run](#)

```
SELECT COUNT (*)
AS total_rows
FROM aim;
```

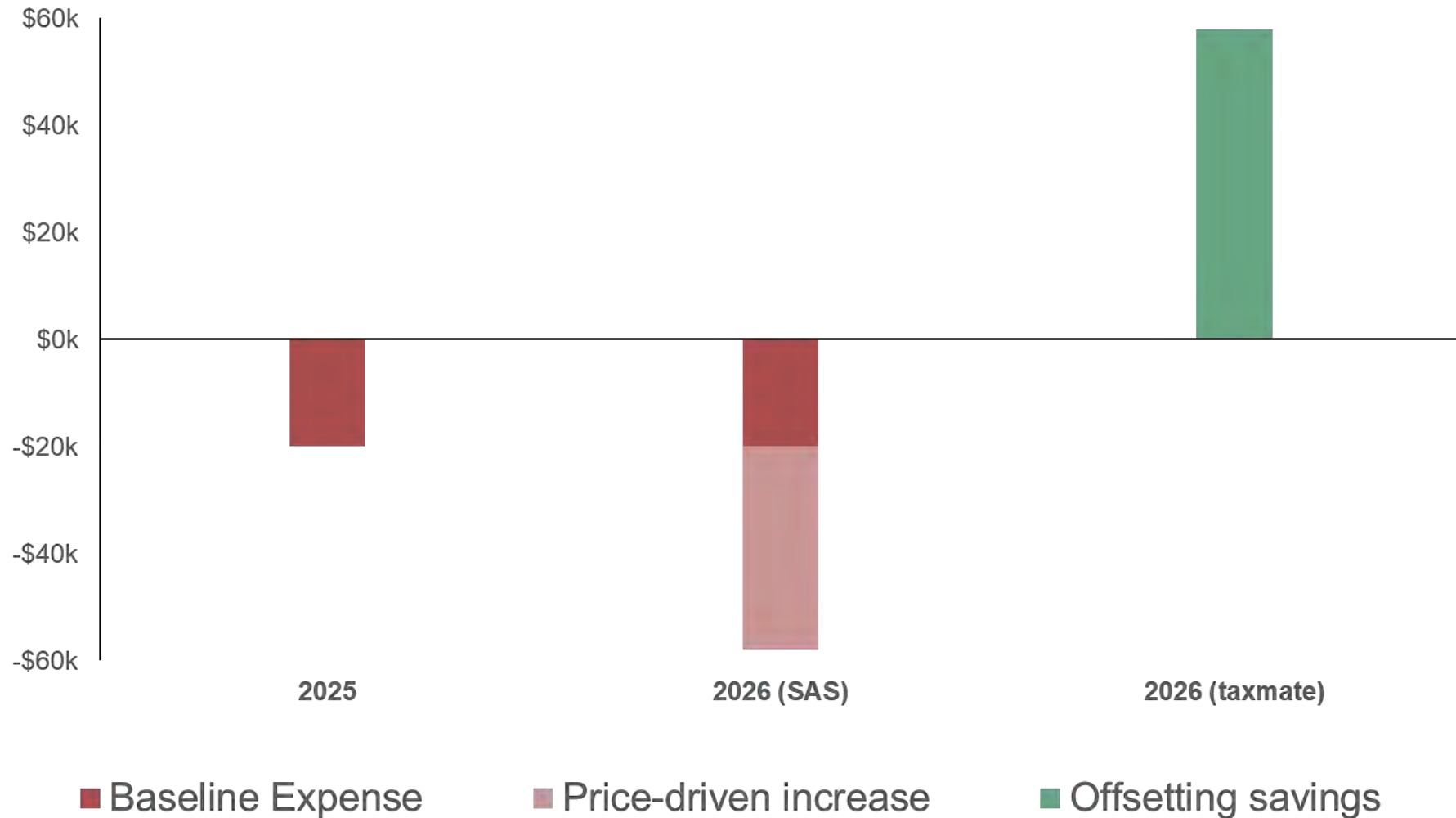
Show entries

Search:

	total_rows
1	121278

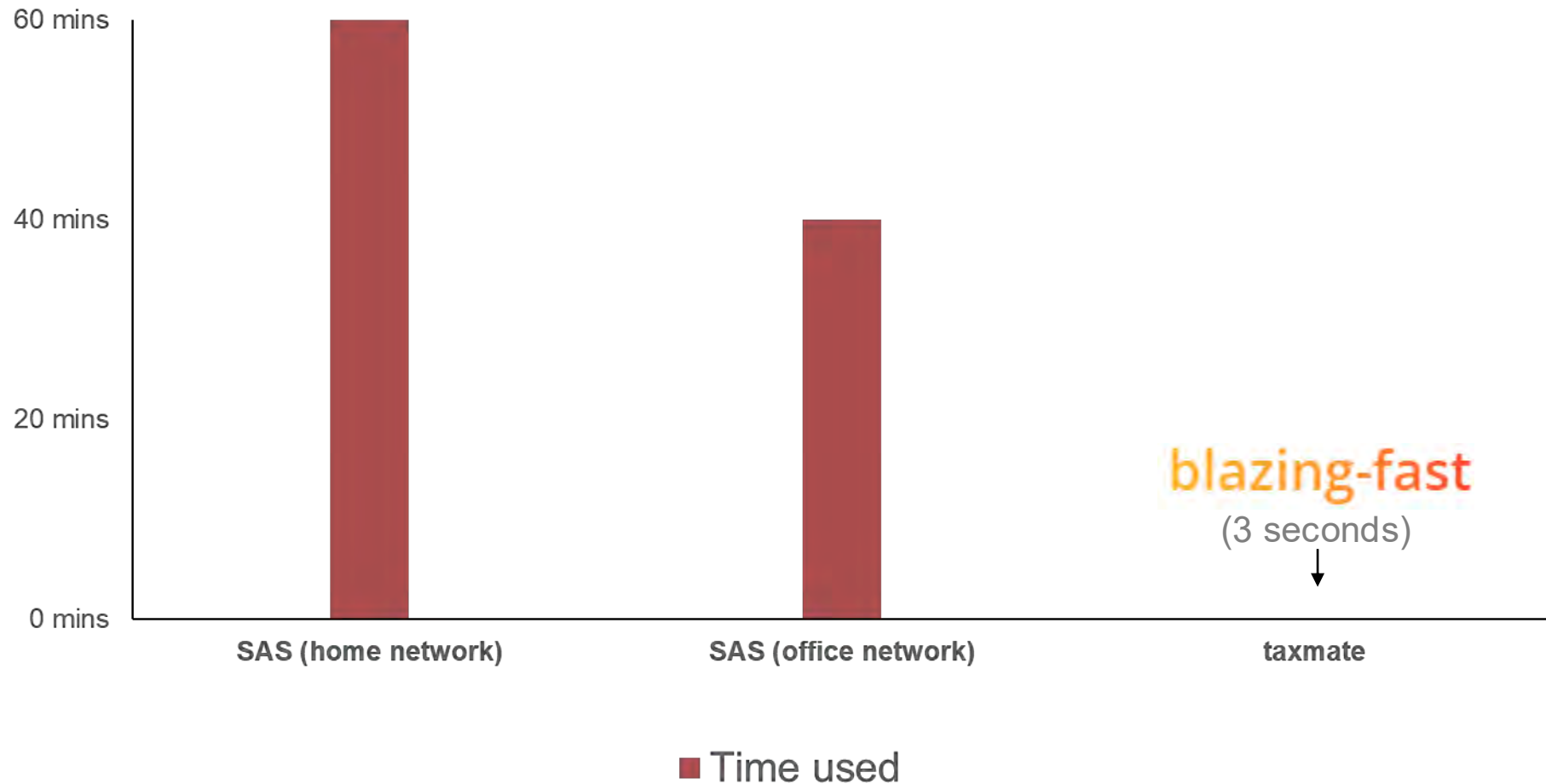
[Download](#)

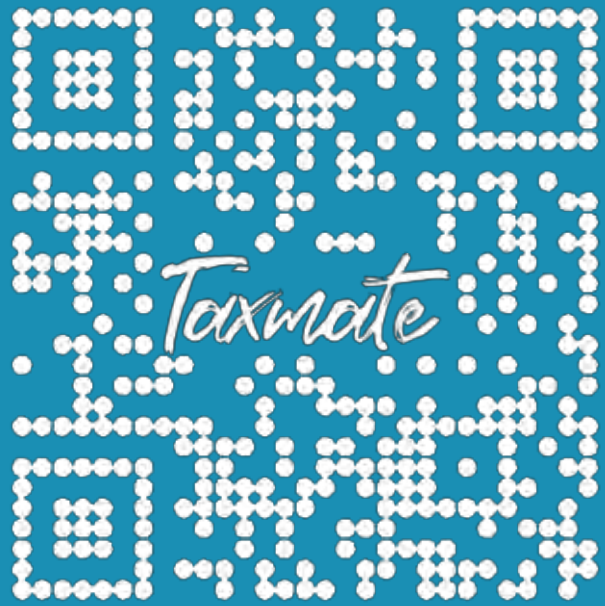
Taxmate saves money each year. A lot of money.



Taxmate saves time for users. A lot of time.

Same Query on a One Billion-row Database





Scan for more resources in GitHub



TE TAI ŌHANGA
THE TREASURY

