

## • Example report

Kitebrook is a fictional company. Every name, number and finding in this document is invented example data, produced to show the Baked Labs audit format.

DATA FOUNDATION AUDIT • BL-AUD-0007-EX

# Kitebrook's numbers don't match. Here is why, and the 90 days that fix it.

An independent review of Kitebrook's warehouse, dbt project, orchestration and BI, with findings ranked by severity and a sequenced roadmap.

PREPARED FOR

Kitebrook leadership  
team

FIELDWORK

14–25 September 2026

ISSUED

2 October 2026

PREPARED BY

Baked Labs

## ● TL;DR

# Kitebrook cannot yet defend its own revenue number. The fixes are known, cheap and fit in 90 days.

**11%**

spread between the three ARR figures leadership uses

**43**

BI users who can read unmasked customer PII

**4%**

of dbt models have any test (9 of 212)

**3.2×**

rise in Snowflake spend in 12 months, to \$11.4k/mo

## What we found

- **No single source of truth for ARR.** Board, finance and sales report \$6.80M, \$6.42M and \$6.12M. A Series B investor will rebuild this from Stripe and find the gap.
- **Customer data is over-exposed.** PII is unmasked for 43 BI users, 31 of whom have no data role, and one admin credential is shared.
- **Errors reach leadership unnoticed.** With 4% test coverage, a broken churn model under-reported churn for 19 days.
- **Cost is drifting.** Spend tripled on 1.3× more queries; about \$4.2k a month is avoidable.

## What to do first

- 01 This fortnight:** mask PII and rotate the shared credential (R2).
- 02 This month:** agree one ARR definition with the CFO and freeze it (R1).
- 03 Within 90 days:** test the revenue path, add CI and cost guardrails (R3, R4).

### OVERALL MATURITY

**1.8** / 5 · Emerging

Target after the 90-day roadmap: **3.2 (Defined)**.

### FINDINGS BY SEVERITY

|            |   |
|------------|---|
| ● Critical | 2 |
| ● High     | 2 |
| ● Medium   | 3 |
| ● Low      | 1 |

### RECOMMENDED NEXT STEP

A 4–6 week Foundation Sprint covering R1 and R3, with R2 and R4 done in-house in the first fortnight.

## 01 · Context

# A Series A company preparing for diligence on a data stack built for a seed-stage team.

## The company

|              |                                                                                       |
|--------------|---------------------------------------------------------------------------------------|
| WHAT THEY DO | Scheduling and dispatch software for field-service teams (HVAC, plumbing, electrical) |
| STAGE        | Series A, \$14M raised in January 2026                                                |
| ARR          | \$6.12M–\$6.80M, depending on who you ask                                             |
| PEOPLE       | 74 employees; data team of 2 (1 analytics engineer, 1 analyst)                        |
| STACK        | Fivetran → Snowflake → dbt Core → Metabase                                            |
| WHY NOW      | Series B process expected in 9–12 months                                              |

## Why leadership asked for an audit

In the July board meeting, Kitebrook's CEO presented ARR of \$6.80M. A board member asked for net revenue retention by cohort; the team needed nine days to produce it, and the answer changed twice. The CFO asked us to find out why the numbers move and what it would take to make them hold up under Series B diligence.

### SCOPE

- Snowflake account: warehouses, roles, query history (90 days)
- dbt Core project: 212 models, 9 tests, macros, run logs
- Fivetran: 6 connectors and sync history
- Metabase: 148 dashboards, 1,310 saved questions, permissions
- Interviews: CEO, CFO, VP Sales, Head of CS, the data team

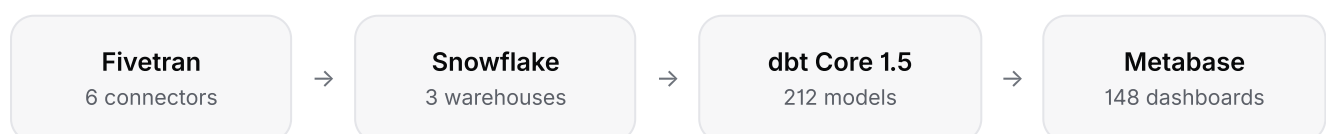
### OUT OF SCOPE

Building new models or pipelines; application databases; security outside the data stack.

### METHOD

- 01 Read-only access via a scoped service role
- 02 Automated scans of metadata and query logs
- 03 Rebuild of ARR from Stripe invoices
- 04 Scoring against the Baked Labs maturity model

### DATA FLOW AS FOUND



## 02 • Scorecard

# Overall maturity is 1.8 out of 5. Testing, metrics and access are the weakest links.

| DOMAIN                                                                            | MATURITY (1-5)                                                     | SCORE |
|-----------------------------------------------------------------------------------|--------------------------------------------------------------------|-------|
| <b>Ingestion &amp; freshness</b><br>Connectors run; nobody is told when they stop | <div><div></div><div></div><div></div><div></div><div></div></div> | 2.5   |
| <b>Modelling (dbt)</b><br>212 models, no layering convention                      | <div><div></div><div></div><div></div><div></div><div></div></div> | 2.0   |
| <b>Testing &amp; quality</b><br>9 of 212 models tested; none on revenue           | <div><div></div><div></div><div></div><div></div><div></div></div> | 1.0   |
| <b>Metric definitions</b><br>Three ARR figures in circulation                     | <div><div></div><div></div><div></div><div></div><div></div></div> | 1.5   |
| <b>BI &amp; adoption</b><br>148 dashboards, 61% unopened in 90 days               | <div><div></div><div></div><div></div><div></div><div></div></div> | 2.0   |
| <b>Access &amp; governance</b><br>PII open to 43 BI users                         | <div><div></div><div></div><div></div><div></div><div></div></div> | 1.5   |
| <b>Cost &amp; performance</b><br>Spend up 3.2x on 1.3x more queries               | <div><div></div><div></div><div></div><div></div><div></div></div> | 2.0   |

## HOW TO READ THE SCALE

- 1 • Ad hoc** Works by heroics; no repeatable process
- 2 • Emerging** Some practice exists, unevenly applied
- 3 • Defined** Agreed standards, mostly followed
- 4 • Managed** Measured, monitored, owned
- 5 • Optimized** Improved continuously from data

## HOW TO READ SEVERITY

- Critical** Fix now. Material risk to reported numbers, security or compliance.
- High** Fix this quarter. Recurring cost, incidents or loss of trust.
- Medium** Plan it. Slows the team; becomes High at the next stage.
- Low** Hygiene. Improve when you next touch the area.

## READING THIS PAGE

Scores are our judgement against the criteria in the appendix, based on evidence collected during fieldwork. A 3 is a healthy target for a Series A company; 4 and above is usually not worth the cost at this stage.

## 03 · What we found

# Eight findings. Two need action this month.

| ID   | SEVERITY | FINDING                                                                                                                                                                                                                         | DOMAIN                |
|------|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| F-01 | Critical | <b>Three different ARR numbers are in circulation</b><br>Board, finance and sales each report a different ARR: \$6.80M, \$6.42M and \$6.12M. The \$0.68M spread is 11% of the lowest figure.                                    | Metric definitions    |
| F-02 | Critical | <b>Customer PII is readable by 43 BI users</b><br>Names, emails, phone numbers and service addresses are unmasked in Metabase. 31 of the 43 users have no data role. One admin credential is shared by 3 services and 2 people. | Access & governance   |
| F-03 | High     | <b>Revenue models have no tests, and failures go unnoticed</b><br>9 of 212 dbt models (4%) carry any test. fct_churn broke silently for 19 days in August; leadership saw 1.2% monthly churn against a recalculated 2.1%.       | Testing & quality     |
| F-04 | High     | <b>Warehouse spend tripled while usage barely moved</b><br>Snowflake spend reached \$11.4k/month, up 3.2× year on year on 1.3× more queries. One hourly full-refresh model drives 38% of credits.                               | Cost & performance    |
| F-05 | Medium   | <b>Nobody is alerted when data goes stale</b><br>No freshness checks exist. The Stripe connector lagged more than 48 hours twice in Q3; a HubSpot sync failed for 6 days before anyone noticed.                                 | Ingestion & freshness |
| F-06 | Medium   | <b>Dashboard sprawl hides the few that matter</b><br>90 of 148 dashboards (61%) were not opened in 90 days. Four competing 'Pipeline' dashboards exist. 44% of saved questions are raw SQL that bypasses dbt.                   | BI & adoption         |
| F-07 | Medium   | <b>Changes ship to production without review</b><br>No CI and no pull-request checks; commits go straight to main. dbt Core 1.5 is past end of support.                                                                         | Modelling (dbt)       |
| F-08 | Low      | <b>Models have no owners and little documentation</b><br>12% of models are documented and none has a named owner. The analyst hired in June took 6 weeks to ship a first model.                                                 | Access & governance   |

## 03 · Critical findings

# The two issues to fix this month

F-01 • Critical METRIC DEFINITIONS

## Three different ARR numbers are in circulation

Each team calculates ARR its own way. Finance counts contracts once signed, including 11 not yet live. Salesforce subtracts churn only when a rep flags it. The Metabase dashboard annualises Stripe MRR and misses the 23 customers invoiced outside Stripe.

## EVIDENCE

ARR rebuilt from Stripe and the contract register; three SQL definitions compared line by line

## IMPACT

Board metrics cannot be reproduced; NRR and CAC payback inherit the error

## FIX

R1: one definition, one model, one metric

ARR BY SOURCE, \$M (EXAMPLE DATA)

|                            |             |         |
|----------------------------|-------------|---------|
| Board deck (finance sheet) | <div></div> | \$6.80M |
| Salesforce report          | <div></div> | \$6.42M |
| Metabase dashboard         | <div></div> | \$6.12M |

## WHY IT MATTERS

\$0.68M spread, 11% of the lowest figure. Diligence teams rebuild ARR from billing data first.

F-02 • Critical ACCESS &amp; GOVERNANCE

## Customer PII is readable by 43 BI users

Metabase connects to Snowflake as a single role that can read every raw table. Contact names, emails, phone numbers and service addresses for roughly 41,000 end customers are visible to anyone with a Metabase login.

## EVIDENCE

Snowflake grants export; Metabase permission audit; credential inventory

## FIX

R2: masking policies, per-team roles, rotate the shared credential

43

users can read PII

31

of them have no data role

1

admin credential,  
shared by 3  
services and 2  
people

6

users who actually  
need PII

## 03 • High findings

# Silent failures and a drifting bill

F-03 • High TESTING &amp; QUALITY

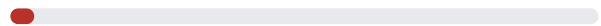
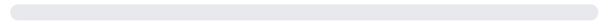
## Revenue models have no tests, and failures go unnoticed

Only 9 of 212 models carry any dbt test, and none of them sit on the revenue path. In August a Stripe schema change broke `fct_churn`; the model kept running on stale rows for 19 days.

**IMPACT** Leadership saw 1.2% monthly churn; the recalculated figure was 2.1%

**FIX** R3: tests on the 30 critical models, CI on every pull request

TEST COVERAGE

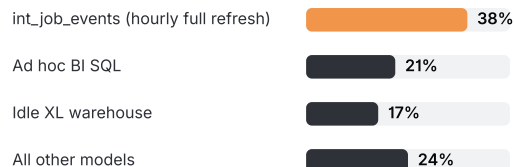
**4%** 9 of 212 models**0** of 30 critical revenue models tested

F-04 • High COST &amp; PERFORMANCE

## Warehouse spend tripled while usage barely moved

SNOWFLAKE SPEND PER MONTH, OCT 2025-SEP 2026  
(EXAMPLE DATA)

SHARE OF CREDITS, SEPTEMBER

**OPPORTUNITY**

Making one model incremental and suspending idle warehouses saves an estimated **\$4.2k a month**.

## 03 • Medium and low findings

# The habits that let problems in and keep them hidden.

F-05 • Medium INGESTION &amp; FRESHNESS

**Nobody is alerted when data goes stale**

No freshness checks exist. The Stripe connector lagged more than 48 hours twice in Q3; a HubSpot sync failed for 6 days before anyone noticed.

FIX R5: freshness checks on all 6 sources, alerts to Slack

F-06 • Medium BI &amp; ADOPTION

**Dashboard sprawl hides the few that matter**

90 of 148 dashboards (61%) were not opened in 90 days. Four competing 'Pipeline' dashboards exist. 44% of saved questions are raw SQL that bypasses dbt.

FIX R5: certify 12 dashboards, archive 90

F-07 • Medium MODELLING (DBT)

**Changes ship to production without review**

No CI and no pull-request checks; commits go straight to main. dbt Core 1.5 is past end of support.

FIX R3: GitHub Actions CI with dbt build on changed models; upgrade dbt

F-08 • Low ACCESS &amp; GOVERNANCE

**Models have no owners and little documentation**

12% of models are documented and none has a named owner. The analyst hired in June took 6 weeks to ship a first model.

FIX R2: owners in YAML, docs on critical models first

## NOTE

None of these is urgent alone. Together they explain why F-01 and F-03 happened: changes are unreviewed, nobody owns a model, and nothing alerts when data stops arriving.

04 • Why it matters

# The cost of doing nothing shows up in diligence, in risk and on the invoice.

01 • FUNDRAISING

## Diligence will find the gap

Investors rebuild ARR and NRR from billing data. An unexplained 11% difference becomes a price conversation, or a delay, at the worst moment.

02 • RISK

## Exposure without need

Unmasked PII for 41,000 end customers, readable by 43 people, is a privacy obligation Kitebrook cannot currently evidence it meets.

03 • MONEY AND TIME

## Paying twice

About \$50k a year in avoidable warehouse spend, plus an estimated 30% of the data team’s week spent reconciling numbers instead of answering questions.

FROM FINDING TO BUSINESS IMPACT

| IF NOTHING CHANGES                                                                                                          | LIKELY BY          | LINKED FINDINGS |
|-----------------------------------------------------------------------------------------------------------------------------|--------------------|-----------------|
| <b>Board questions a metric and the answer changes again</b><br>Each re-statement erodes trust in every other number.       | Next board meeting | F-01 F-03       |
| <b>A Series B data room shows inconsistent ARR</b><br>The reconciliation gets done under deadline, by the CFO.              | Q2–Q3 2027         | F-01 F-05       |
| <b>A customer or regulator asks who can see their data</b><br>There is currently no answer short of “anyone with Metabase”. | Any time           | F-02 F-08       |
| <b>Warehouse spend passes \$15k a month</b><br>At the current trend, before the end of Q1 2027.                             | Q1 2027            | F-04            |

THE GOOD NEWS

The foundations are recoverable. The ingestion layer works, the team knows the business well, and every fix in this report uses tools Kitebrook already pays for.

## 05 · What to do

# Five recommendations, ordered by risk removed per week of effort.

| ID | RECOMMENDATION                                             | FIXES      | OWNER                    | EFFORT | IMPACT    |
|----|------------------------------------------------------------|------------|--------------------------|--------|-----------|
| R1 | Agree one ARR definition and encode it once                | F-01       | CFO + analytics engineer | M      | Very high |
| R2 | Mask PII, rotate credentials, move to role-based access    | F-02, F-08 | Analytics engineer + IT  | S      | Very high |
| R3 | Test the revenue path and add CI to every change           | F-03, F-07 | Analytics engineer       | M      | High      |
| R4 | Put cost guardrails on Snowflake                           | F-04       | Analytics engineer       | S      | High      |
| R5 | Add freshness alerts and certify a small set of dashboards | F-05, F-06 | Analyst                  | M      | Medium    |

## R1 IN PRACTICE

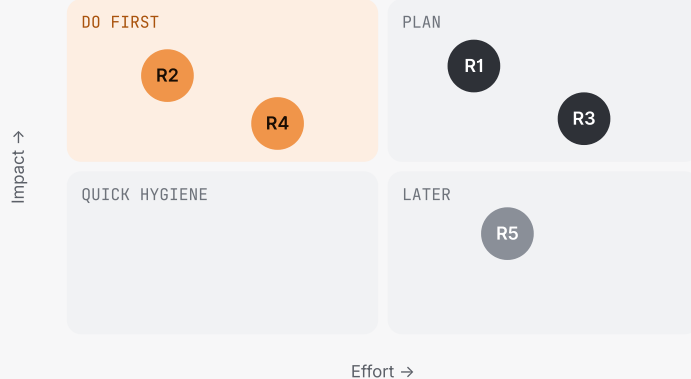
## One ARR, defined in writing

- Workshop with CFO, VP Sales and data team; one page definition signed off
- Recognise ARR at go-live, from Stripe plus the contract register
- Build `fct_revenue` and one ARR metric in the semantic layer
- Every dashboard and the board pack read from that metric only

## R4 IN PRACTICE

## Cost guardrails

- Make `int_job_events` incremental; run it every 6 hours, not hourly
- Auto-suspend all warehouses after 60 seconds; drop XL to M
- Resource monitor alerting at 80% of a \$7.5k monthly budget
- Expected saving: \$4.2k a month



## PRIORITISATION

## Start with the two small, high-impact fixes

R2 and R4 take under a week each and remove the privacy exposure and most of the avoidable spend. R1 and R3 are the larger build; they are what makes the numbers hold up in diligence.

Effort: S = under a week, M = 2–4 weeks of one engineer. Impact is our estimate of risk or cost removed.

06 · Roadmap

# 90 days from three numbers to one. Stop the bleeding, build the spine, earn back trust.

DAYS 0-30

Stop the bleeding

PII masked, shared credential gone, cost quick wins live, ARR definition signed.

DAYS 31-60

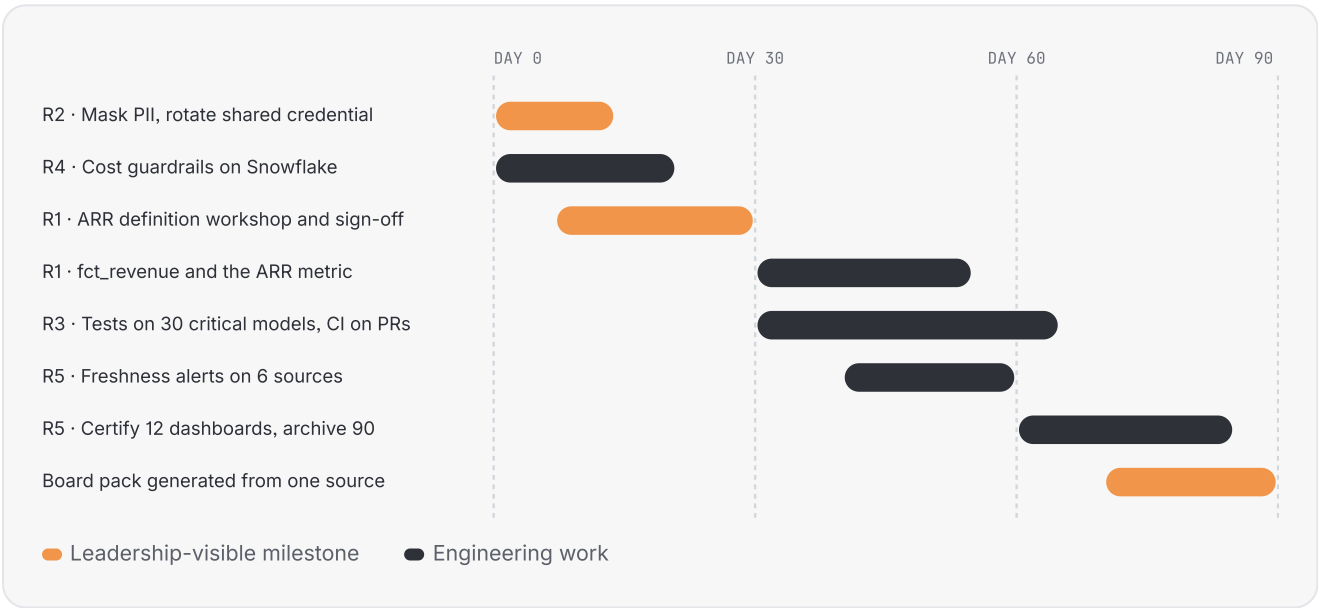
Build the spine

fct\_revenue and the ARR metric shipped; tests and CI on the revenue path; freshness alerts on.

DAYS 61-90

Earn back trust

12 certified dashboards; 90 archived; board pack generated from one source.



DAY-90 TARGETS

| MEASURE                             | TODAY   | DAY 90       |
|-------------------------------------|---------|--------------|
| ARR figures in circulation          | 3       | 1            |
| Critical models with tests          | 0 of 30 | 30 of 30     |
| Users with PII access               | 43      | 6            |
| Snowflake spend / month             | \$11.4k | ≤ \$7.2k     |
| Stale-data incidents found by users | 3 in Q3 | 0            |
| Overall maturity                    | 1.8     | 3.2 (target) |

RECOMMENDED NEXT STEP

**Foundation Sprint, 4-6 weeks.** Baked Labs builds R1 and R3 end to end in Kitebrook’s repo; the team handles R2 and R4 in-house with our runbooks.

DECISION NEEDED

CFO to own the ARR definition and attend the workshop in week 2.

## Appendix

# Method, definitions and limits

## Scoring criteria

Each domain is scored 1–5 against written criteria. Scores are rounded to the nearest 0.5; the overall score is the unweighted mean of the seven domains.

## Evidence collected

- 90 days of Snowflake query and metering history
- dbt manifest, run results and Git history (12 months)
- Fivetran sync logs; Metabase usage and permissions export
- Seven stakeholder interviews, 45 minutes each

## Glossary

|           |                                                                 |
|-----------|-----------------------------------------------------------------|
| ARR       | Annual recurring revenue: live subscription revenue, annualised |
| NRR       | Net revenue retention by starting cohort                        |
| PII       | Personally identifiable information                             |
| CI        | Automated checks run on every proposed change                   |
| FRESHNESS | Time since a source last delivered new data                     |

## Limits of this audit

We reviewed configuration, metadata and samples, not every row. Cost savings are estimates from 90 days of history. Nothing in this report is legal advice.

### ABOUT THIS DOCUMENT

**This is an example report.** Kitebrook is a fictional company invented by Baked Labs. All names, figures, findings and dates are illustrative and do not describe any real organisation.

### CONTACT

#### Baked Labs

bakedlabs.co · Data Foundation Audit  
Report BL-AUD-0007-EX · 2 October 2026