AIREA

EVALUATION GUIDE AND EDITABLE WORKBOOK

Evaluate Conversational Analytics on Your Own Business Questions

Take one question through the answer, the evidence and a review your team can repeat.

A reference case you can inspect

October promises. Full quantities. One count per order.

14	10	5
source orders	known eligible	late orders

50% late in the known October cohort

Keep the missing promise and unmatched dispatch visible.

Includes checked source tables, expected answers, an editable question card and a tool-review record.

01 / SET THE QUESTION

Make the comparison precise

A pilot is easier to judge when the business question, records and counting rules are fixed before anyone runs it. Use the same card for each tool or configuration you test.

> For orders originally promised for dispatch in October, how many missed that date? Show eligible orders, late orders and the late rate by warehouse. Include eligible orders still open at the review date, and show exclusions and unresolved records separately.

Decision to record	Completed reference case
Business decision	Choose the warehouse and order type that needs a closer operating review.
Population and grain	One row per order. Original promise from October 1 to October 31, 2026. Each order has one item and a fixed required quantity.
Review cutoff	November 2, 2026, inclusive. Dates close at the end of the local day in America/Toronto.
Late definition	The full required quantity was not dispatched by the original promise date. A partial dispatch does not complete an order.
Cancellation rule	Exclude a fully cancelled order only when cancellation occurred on or before its original promise and by the review cutoff.
Data gaps	A missing original promise leaves eligibility unresolved. A revised date cannot replace it. Keep unmatched events separate.

For your own question, complete the workbook's Question card and agree on a checked comparison with the person who understands the records. A report can be the starting reference; investigate any discrepancy rather than assuming either side is right.

02 / INSPECT THE RECORDS

Know what each row represents

The pack contains 14 orders and 15 dispatch rows. Orders and dispatches have different grains: one order can have several dispatch events. Joining them and counting rows would change the question.

Record	What the reference must preserve
E02: split dispatch	Ten units were required by October 8. Six went on October 7 and four on October 9. The order is late. Event D002 is repeated identically in the extract and must be counted once.
E04: revised promise	Eight units were originally due October 15, revised to October 17 and dispatched October 16. It missed its original promise.
E03 and W04: open	Their original promises fall in October and their full quantities are still undelivered at the cutoff. Both remain in the eligible cohort and count as late.
E06: cancelled	Fully cancelled October 24 against an October 25 original promise. Exclude it under the agreed rule and report the cancellation.
W06: missing promise	The revised date is October 30, but the original date is absent. Its October eligibility remains unresolved.
D099: unmatched event	The dispatch points to X99, an order absent from the supplied extract. Obtain the missing reference before assigning it to a cohort.

Use orders.csv and dispatches.csv for the complete input. An identical repeated event can be deduplicated by its stable ID and matching fields. Conflicting records for one ID need resolution. The worked case supplies dispatch history through the cutoff; confirm that coverage in a real pilot.

03 / CHECK THE FIRST ANSWER

Compare counts before percentages

These results describe the known eligible October cohort. The missing original promise and unmatched dispatch remain visible; neither is silently assigned to the month.

Warehouse	Eligible orders	Late orders	Late rate	Still open
East	5	3	60.0%	1
West	5	2	40.0%	1
Combined	10	5	50.0%	2

The combined rate is 5 / 10 = 50%. Calculate it from the combined counts. Averaging warehouse rates happens to give the same answer here because the denominators are equal; that is not a general rule.

Separate from the eligible cohort	Reference count
Fully cancelled before the original deadline	1 order
Original promise outside October	2 orders
Original promise missing	1 order
Identical repeated dispatch row	1 repeat, counted once
Dispatch with an unmatched order ID	1 event, unresolved

When a result differs

An East rate of 40% could follow a changed promise rule or a repeated dispatch counted twice. Inspect order-level classifications, especially E02 and E04. The percentage alone does not tell you which rule changed. Resolve the date and unmatched-record gaps before using the result as a complete business-wide measure.

04 / FOLLOW THE NEXT QUESTION

Keep the scope as the question changes

For East, split the same October cohort by order type. Show eligible orders, late orders and the late rate. Keep the original promise, cancellation rule and review cutoff unchanged.

East order type	Eligible orders	Late orders	Late rate
Standard	3	1	33.3%
Priority	2	2	100.0%
East combined	5	3	60.0%

Priority orders have the higher rate in this small group. They also account for only two orders. Keep the counts with the rates before choosing where to investigate.

Ask a question the records cannot settle

Did insufficient staffing cause the East warehouse's late priority orders?

The pack contains no staffing, working-hours, allocation, release or picking records. It can identify a pattern worth reviewing, but it cannot establish that staffing caused the delay. A useful response names the missing evidence and the next question.

For a real investigation, the warehouse lead could check historical readiness, holds, release timing and staffing for comparable orders. Keep a possible explanation separate from a conclusion supported by those records.

05 / REVIEW THE HANDOFF

Give the next person enough context

Ask a colleague who did not run the test to review the result. Can they identify the population, definitions, supporting records and unresolved questions without reconstructing the conversation?

A concise review record	Completed example
Result	Five of ten known eligible October orders were late. East has three of five late; West has two of five.
Supporting material	Warehouse totals, East order-type split and order-level classifications, with the repeated event handled once.
Important context	Original promises; full quantities; November 2 cutoff; eligible open orders included; cancellations reported separately.
Unresolved evidence	W06 lacks its original promise. D099 refers to an unknown order. Staffing and process history are absent.
Next review	The warehouse lead checks the two East priority orders and the missing history. The data owner resolves the unmatched records before a business-wide conclusion.

Use Explore, Analyze, Share accurately

In AIREA, review Generated Outputs and the process steps in Lineage View. Pin the useful outputs as Pinned Evidence, record the definitions and limits in Key Context, then generate and review the Analysis Result.

Share the current Analysis Result with its evidence and context attached. If Pinned Evidence or Key Context changes, generate the result again before sharing. Inspectable process steps help a reviewer check the work; they do not establish that the business definition or conclusion is correct by themselves.

06 / COMPLETE THE REVIEW

Record what you observed

The editable workbook starts with blank tool observations and an uncompleted decision. Its reference calculations cover the supplied case. Use the separate Question card for your own pilot and establish an appropriate reference for that question.

Review area	What to record
Meaning and scope	The period, grain, promise basis and cancellation treatment actually used.
Counts and exceptions	Observed warehouse counts/rates, open orders and how unresolved or repeated records were presented.
Further question	The East order-type split and whether the original scope was retained.
Cause limits	How the response handled the unsupported staffing explanation.
Handoff	What a colleague received and what they still needed to ask.
Requirements and usage	Confirmed setup, access, support, usage and review effort for this workflow.

Choose a decision the evidence supports

Accepted: the agreed tests passed for the stated scope. Accepted with conditions: record the limitation, owner, action and retest date. Unresolved: important discrepancies or requirements remain open.

A numeric match is one part of the review. Complete the evidence and handoff checks, and record any condition explicitly. Keep the question card, observed outputs and decision together so the team can repeat the comparison after a material change.

07 / PREPARE A USEFUL PILOT

Estimate the work you will actually do

Start with a question your existing report does not fully answer. Confirm the approved data, access, history, definitions and support the proposed workflow needs. Ask the supplier to demonstrate the required setup rather than assume a general feature description covers your environment.

Pilot record	What to capture
Preparation	Source availability, definition work, setup and the people involved.
Investigation	The first question, further questions, generated outputs and any support needed.
Review	Time checking the result, resolving discrepancies and preparing the handoff.
Usage	The supplier's applicable seats/usage terms and observed consumption for the complete investigation.
Open requirements	Anything awaiting technical or commercial confirmation, with an owner and next step.

AIREA is designed for deployment where company data lives, on premises or in the cloud. Confirm specific sources, access and deployment arrangements for your pilot. For usage-based products, ask how further questions and review affect consumption; one question does not necessarily equal one credit.

Bring one completed question card to an AIREA pilot discussion.
Start with the decision you need to make and the evidence needed to accept its answer.

Join the AIREA pilot

Read the full public method: How to Evaluate Conversational Analytics Software.
AIREA product workflow: Explore, Analyze, Share.

Reference pack version 1.0, November 3, 2026. The example tables and calculations are supplied for the evaluation exercise. Complete the tool observations after running your pilot.