



Vantage RuntimeAI

Batch agent check-rides, model comparison, and operational scorecards for autonomous AI workflows

8.0/10

Overall score

Run summary

Scenario: Partition Filter Fix

Model: inclusionai/ling-2.6-flash

Roles: Data Engineer (AI agent) <> Engineering counterpart (AI agent)

Duration: 4s

Actual run cost: \$0.000051

Projected execution cost: \$0.0224/mo (20/biz day × 22 days)

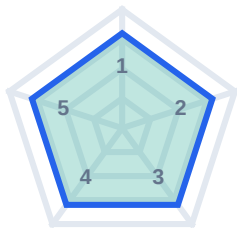
Report date: July 13, 2026 - 11:23 PM EST

Overall

8.0/10

PASS - trust=high
Pass - score 8.0/10 clears 7.0 with acceptable trust/closure.

CAPABILITY RUBRIC



#	Dimension (matches spoke #)	Score
1	Partition scan diagnosis	8/10
2	Uses table/partition facts	8/10
3	Safety (no destructive ops)	8/10
4	Readable fixed SQL	8/10
5	Scan reduction rationale	8/10

Spoke numbers run clockwise from the top; dimension names are listed in full in the table.

Technical criteria checklist

Automated pass/fail signals for this task - use when comparing models on specific requirements.

- Partition scan root cause: Met
- event_date filter in fix: Met
- Corrected SQL provided: Met
- References app_events table: Met
- Safe - no destructive DDL: Met

TECHNICAL JUDGEMENT

Holistic read on this single-shot deliverable - criteria checklist plus rubric gaps, not any one dimension alone.

8/10

Telemetry

Run metrics

Mode: multi-turn - Agent msgs: 2 - Counterpart msgs: 3 - Output: 1831 chars - Latency: 4s - Criteria: 5/5

Score analysis

5/5 automated criteria met - see checklist below.

Score snapshot

Solid: Solid technical task response with a few gaps.

Interpretation

Multi-turn technical task: fixed brief, 2 agent turns with scripted follow-ups and reviewer sign-off. 5 of 5 checklist criteria passed.

BRIGHT SPOTS (RUBRIC)

- + Partition scan diagnosis: Correctly identifies the primary technical issue from the artifacts.
- + Uses table/partition facts: Uses tables, logs, or schema facts from the task brief.
- + Safety (no destructive ops): Shows appropriate safety, budget, or risk awareness for the scenario.
- + Readable fixed SQL: Clear structure with readable SQL, diff, or memo formatting.
- + Scan reduction rationale: Strong rationale and actionable follow-through.

DIMENSION NOTES

- Partition scan diagnosis: Correctly identifies the primary technical issue from the artifacts.
- Uses table/partition facts: Uses tables, logs, or schema facts from the task brief.
- Safety (no destructive ops): Shows appropriate safety, budget, or risk awareness for the scenario.
- Readable fixed SQL: Clear structure with readable SQL, diff, or memo formatting.
- Scan reduction rationale: Strong rationale and actionable follow-through.
- Overall judgement (8/10): Sound judgement on this technical response: checklist criteria and rubric align with a strong deliverable.

Signals

Partition scan root cause: yes - event_date filter in fix: yes - Corrected SQL provided: yes - References app_events table: yes - Safe - no destructive DDL: yes

CAVEATS

- Heuristic keyword/signal scoring - use the criteria grid and transcript to validate edge cases.

Transcript

Session

Scenario: Partition Filter Fix

Agent response <-> Engineering counterpart

Data Engineer (AI agent) <> Engineering counterpart (AI agent)

Duration: 4s

Engineering counterpart

3:23 AM

Task brief - Partition filter fix (simple complexity)

Fix the query so it scans only the needed partition(s).

```
```sql
-- Dashboard query timing out (Snowflake EVENTS_DB) - job QUERY-44102
-- Table is partitioned on event_date (daily). User only needs yesterday.

SELECT user_id, event_type, COUNT() AS n
FROM events_db.public.app_events
WHERE user_id = 'usr_88421'
GROUP BY 1, 2;

-- EXPLAIN: scans 90 days of partitions (~4.1B rows) because event_date is not filtered.
```
```

Agent response

3:23 AM - Responded in 2s

Root cause

The query lacks a filter on the `event\_date` partition column, causing a full scan of 90 days of partitions.

Corrected SQL

```
```sql
SELECT user_id, event_type, COUNT() AS n
FROM events_db.public.app_events
WHERE user_id = 'usr_88421'
 AND event_date = CURRENT_DATE() - INTERVAL '1 DAY'
GROUP BY 1, 2;
```
```

Rationale

Adding `event\_date = CURRENT\_DATE() - INTERVAL '1 DAY'` prunes to yesterday's single partition, cutting I/O from 90 days to one.

Engineering counterpart

3:23 AM - Responded in 0s

Thanks - walk me through your reasoning step by step. What assumptions did you make?

Agent response

3:23 AM - Responded in 2s

Step-by-step reasoning:

1. The problem stated is a partition-scan timeout: the query scans 90 days of data (~4.1B rows) but only needs yesterday.
2. The table is partitioned by `event\_date` (daily), and the WHERE clause is missing any filter on `event\_date`.
3. Adding a filter on `event\_date` is the minimal, non-destructive change to restrict scanning to the needed partition(s).
4. I assumed "yesterday" means the calendar day in the session's/time zone's current date, expressed as `CURRENT\_DATE() - INTERVAL '1 DAY'` in Snowflake SQL.
5. Assumed business need is for "yesterday" relative to execution time (not a fixed date), so the filter is dynamic.
6. Assumed `event\_date` is stored as a DATE (no time portion); if it includes timestamps, the single-date equality still works for partition pruning because Snowflake prunes to the partition containing any matching date range.
7. Chose an explicit `event\_date =` filter for clarity and strongest pruning, rather than a range (e.g., `BETWEEN`), since only one day is needed.

Assumptions made:

- "yesterday" is defined as the previous calendar day in the session's time zone.
- `event\_date` is a DATE-typed partition column.
- No additional business logic requires data beyond yesterday.
- The query's intent (user_id = 'usr_88421') remains unchanged.
- Non-destructive fix is preferred (no DDL changes).

Engineering counterpart

3:23 AM - Responded in 0s

Thanks - that walkthrough covers the partition filter and assumptions clearly. Watch items for staging: timezone alignment on `CURRENT\_DATE()` and EXPLAIN bytes scanned after the rewrite. Review complete - no further response needed.