

Task 1: Data Analysis (what I noticed in the sample extracts)

I reviewed the sample Finance payments extract against the Delivery events extract to get familiar with the data and identify why deeper analysis is needed before I lock "truth" into the warehouse. The mismatches fall into a few distinct types:

1) **Duplicate / replayed finance payments (double-count risk)**

- Example: 'O5006/PK9006' appears twice as PAID for 800 ETB ('P1006' and 'P1021').
- Example: 'O5014/PK9014' appears twice as PAID for 530 ETB ('P1014' and 'P1022').

If these are retries, corrections, or system replays, I need a deduplication rule (and a definition of a "unique payment").

2) **Missing / incomplete join keys (unmatchable records)**

- Example: 'P1023' ('O5021') is PAID for 500 ETB but has a blank 'package_id'.

Without consistent keys (order/package/customer), Finance rows cannot reliably link to Delivery.

3) **Lifecycle / coverage gaps between systems**

- Example: Finance has 'O5023/PK9023' PAID 750 ETB ('P1025'), but there is no corresponding Delivery event in the provided Delivery extract.
- Example: Finance shows 'O5022/PK9022' as REFUNDED 600 ETB ('P1024'), but there is no Delivery event for PK9022 in the extract.

This could be incomplete extracts, late-arriving delivery events, cancellations, or true process exceptions.

4) **Definition mismatches (PAID != DELIVERED)**

- Example: 'O5016/PK9016' is PAID 770 ETB but Delivery is 'FAILED' (D2016).
- Example: 'O5020/PK9020' is PAID 460 ETB but Delivery is 'RETURNED' (D2020).

I need client agreement on how reporting should treat failed/returned shipments and how refunds map to operational outcomes.

5) **Timing / reporting-window differences**

- Example: payment recorded 2026-01-10 for 'O5006/PK9006', but delivery event is 2026-02-01 (D2006).

This will create month-end drift if Finance is reported by 'payment_date' and Delivery is reported by 'event_time'.

Task 2: Client-ready email (<=400 words, contents only)

Subject: Request to time-box Finance vs Delivery reconciliation before warehouse metrics

Hi [Client Team],

I'm building the centralized warehouse and ran an initial reconciliation between the Finance payments extract and the Delivery events extract. Totals and counts do not fully tie out between "paid" activity in Finance and "delivered" outcomes in Delivery.

I do not yet have a confirmed root cause. Based on the sample extracts, I am seeing several specific issue types that could each explain part of the gap:

* **Duplicate/replayed payments:** the same order/package appears twice as PAID with the same

amount/date (for example 'O5006/PK9006' shows 800 ETB PAID in both 'P1006' and 'P1021'; 'O5014/PK9014' shows 530 ETB PAID in 'P1014' and 'P1022'). I need your guidance on what constitutes a "unique payment" and how retries/corrections should be treated.

* ***Missing join keys:** 'P1023' ('O5021') is PAID 500 ETB but has a blank 'package_id', which makes it unmatchable to Delivery without a mapping rule.

* ***Operational outcome differences:** some PAID items are not delivered (e.g., 'O5016/PK9016' is 'FAILED'; 'O5020/PK9020' is 'RETURNED'). I need agreement on definitions: whether reporting should count these as revenue, exclude them, or treat them as exceptions pending re-ship/refund.

* ***Timing differences:** at least one case is paid in January but delivered in February ('O5006/PK9006' paid 2026-01-10; delivered 2026-02-01), which affects period-based reporting.

To avoid embedding incorrect assumptions into executive dashboards, I recommend a **time-boxed 1-2 business day reconciliation** to confirm drivers and lock reporting rules. Deliverable: a short decision memo covering (1) definitions, (2) deduplication + join rules, and (3) how exceptions (failed/returned/refunded/late) will appear in metrics.

Options:

* ***Option A (recommended):** approve the 1-2 day time-box now.

* ***Option B:** proceed on schedule with provisional rules (higher risk of later rework and reporting disputes).

Please confirm whether I can proceed with **Option A**, and nominate one Finance and one Delivery point person for a short working session.

Thanks,
Lemlem