

Week 3: Analysis Methodology Selection

Bálint Décsi

February 6, 2026

1. Analysis Selection

Analysis 1: Retention Analysis (Cohort Analysis)

- **Question:** Do Express users sustain their higher engagement over time, or is it a one-off novelty?
- **Insight:** If D30 *and* rolling D30 retention curves for Express users stay parallel to and above Standard users, it suggests genuine long-term value. If they converge quickly, the effect is temporary/novelty-driven.
- **Limitation:** It does not control for the “wealthy professional” selection bias mentioned by the CFO.

Analysis 2: Ecosystem Analysis (Segmentation & Failure Analysis)

- **Question:** Is the performance difference driven by income or location (external factors) rather than the service itself?
- **Insight:** By comparing Express vs. Standard users within the same zip code or income bracket (using `census_data`), we can isolate the delivery speed’s impact from socioeconomic factors. It might shed light on failure in the different delivery processes.
- **Limitation:** We only have zip-level aggregates (`census_data`) and not individual income.

2. Data Validity Checks (Stop/Go Gates)

- **Check 1: Timestamp Consistency**
 - **Validation:** Ensure `actual_delivery_time > order_date` for all records in `orders` table. Also check if `actual_delivery_time` falls within the window promised by `delivery_window_selected`.
 - **Stop Condition:** If a significant % of “Express” orders were actually delivered late (effectively becoming Standard deliveries), the labels in our data are flawed and the analysis cannot proceed.
- **Check 2: Zip Code Granularity & NULL Rate**
 - **Validation:** Join `customers` with `census_data` on `zip_code`. Check for missing matches.
 - **Stop Condition:** If the high-performing Express users cluster in specific zip codes that are missing from the census data (NULL values after the JOIN operation), we cannot test the CFO’s hypothesis regarding income levels.

3. Causation vs. Correlation Challenge

Addressing Bias:

The CFO’s concern is valid: self-selection bias is likely. To address it, I would perform an experimental analysis using the available census and order attributes to create a control group, if available, of Standard users who look exactly like Express users (similar income area, order value, cohort) but chose Standard. In case they’re not present, I’d use the synthetic control group approach.

Establishing Causation:

Ideally, I would run an A/B Test (Randomized Control Trial). We could randomly upgrade a subset of Standard users to free Express delivery without them requesting it. If their reorder rate jumps to match the organic Express users, it strongly suggests the service *causes* the lift.

Stakeholder Communication:

If causation cannot be proven, I would try to turn the CFO into a champion of the analysis as it would prove their points made in the first place. I’d definitely use the *credibility* tactics to convince other stakeholders that the data

proves something that might be unintuitive: “We can’t guarantee causality without a fully-controlled test. However, even if it is purely correlation, these are clearly our most valuable customers.” To avoid converting the VP of Ops into a blocker, I’d suggest an alternative narrative to their original budget allocation: “The \$2M investment might be justified as a *retention play* to keep high-value segments (Power Users) from churning, rather than an *acquisition lever* intended to convert everyone.”

4. Counter-Metric

Metric: On-Time Delivery Rate (Reliability)

- **Why it matters:** Expanding Express capacity might strain the logistics network. If we promise 1-hour delivery to more people but start missing the window, we destroy trust. A “late Express” delivery is often perceived as a worse failure than a “on-time Standard” delivery because it breaks a premium promise. This is a guardrail to ensure long-term brand trust.

Reflection

This exercise highlighted the danger of taking high-level metrics (like the 45% lift) at face value. It reinforces that in complex business environments, the most obvious “win” is often confounded by selection bias. It taught me that I shouldn’t just present “the number” but must think out of the box and look for the variables that might explain it away before recommending major capital investments. The counter-metric ideation reminded me of the importance of the pre-mortem as well.

AI Declaration: I’ve used AI for this task according to our MS programs AI policy. This document’s structural skeleton (headings etc.) and typesetting were assisted by the Gemini CLI tool, using Gemini 3. The prompt used was: “Make a skeleton for `report/3` based on Week 3 task in `reports/3/task_3.md`. Stick to the outline of the task markdown file.”