Clean, Map and Reconcile a Finance Dataset

Excel & Financial Modeling | Foundation | 45 minutes (estimate)

Learning outcomes

Clean a small ledger export while preserving every exception. Calculate the result, reconcile it and communicate its limitations.

Method

Keep raw rows intact, then create a normalized layer with explicit parsing and duplicate rules. A repeated transaction ID is not enough to delete a row if its contents differ. Reconcile raw totals, excluded duplicates and accepted totals; keep unmapped or invalid rows visible instead of silently dropping them.

Smaller worked example

Raw amounts 100, 100 and 50 include one exact duplicated 100 record. Raw total 250 less excluded duplicate 100 equals accepted total 150. An unmapped account still belongs in accepted totals until its treatment is resolved.

Business context and rules

Fictional Birch Agency, September 2026, USD. The companion CSV has seven rows: T01 sales 1,200.00; T02 expense (250.00); an exact duplicate T02; T03 sales 800.00; T04 expense -100.00 to unmapped account 6999; T05 expense -50.00 with invalid date 2026-09-31; T06 sales 300.00. Dates use ISO format. Account 4000 maps to Revenue, 6100 to Operating expense; 6999 is unresolved. Parentheses mean negative. Exclude only the exact duplicated T02; retain T04 and T05 in the reconciliation but flag them. Accept the signed amount of T05 for reconciliation without inventing a corrected date. Raw signed control total is 1,650.

Original synthetic inputs

Input	Value	Unit
T01 normalized amount	1200	USD
T02 normalized amount	-250	USD
Duplicate T02 normalized amount	-250	USD
T03 normalized amount	800	USD
T04 unmapped amount	-100	USD
T05 invalid-date amount	-50	USD
T06 normalized amount	300	USD
Raw signed control	1650	USD

Required deliverables

Import the raw CSV as text, preserve it, build typed amount / valid-date / account-mapping columns, and provide an exception log for duplicate, invalid date and unmapped account. The normalized Inputs tab is a reconciliation aid; submitting its totals without cleaning the raw CSV does not satisfy the task.

Use formulas for derived amounts and preserve source data. Put narrative deliverables in the workbook response area; expand it as needed. Compare amounts within 0.01 of the stated unit and percentages within 0.1 percentage point. No unsupported balancing plugs.

Required raw dataset

Download XFM-01_raw_v1.csv from the case download panel. It deliberately contains duplicate and invalid records. The workbook Inputs sheet contains numeric control aids, not a substitute for cleaning the raw file.