Luno Exchange Take-Home Assignment

Luno Context

Luno operates a platform that makes it easy for customers to buy, sell, send and receive digital currencies such as Bitcoin. To facilitate the accounting process, a customer is provided with a dedicated account for each currency that they are legally allowed to transact in, namely a digital currency account and a fiat-denominated account. Before buying their first crypto, a customer typically first makes a fiat deposit payment (e.g. South African rands - ZAR) that is credited to her/his fiat wallet. They can then either choose to buy crypto using the instant buy/sell feature (broker) or by making a trade bid on the exchange. The bought crypto is credited to their respective crypto wallet, and their fiat wallet debited with an equivalent value, i.e., double-entry bookkeeping. Transaction charges are debited to the appropriate account where applicable. In the same way, a customer can decide to sell their stored crypto using either the instant buy/sell feature (broker) or by offering to sell these on the exchange market. Note that when trading on the exchange, a customer's buy/sell order may be partially filled, completely filled or not filled at all, depending on offers from counterparties.

Other transactions include:

- Sending/receiving crypto (Bitcoin, etc.)
- Withdrawing fiat
- Placing/cancelling limit orders (bids and asks)

Your Task

Using the provided CSV files, analyse monthly customer trade transaction activity and customer behaviour from `ledger_entries.csv`. Produce a clear, concise report and include all code (formatted as production-ready, merge-request quality).

Deliverables:

- A report answering the business questions below, including any charts and visualisations with associated interpretations, with enough context for technical and non-technical stakeholders.

- All code used to generate results, organised and commented as if for a production codebase.

Business Questions:

1. What is the distribution of customer statuses (`New`, `Returning`, `Churned`, `Reactivated`) each month?
2. What was the churn rate for February and March? Churn rate can be defined as the % of customers that did not trade in a given month but had traded in the month prior.
3. Do different customer transacting segments have different propensities to trade in specific markets? *Ie are there trends that show new customers typically buy X crypto pair etc*
4. What is the average trade volume per user per month? How does this differ by transacting segments?

Data Files:

- `ledger_entries.csv`: One row per account transaction (transaction value indicated by `balance_delta`).
- `accounts.csv`: Can be used to map `id` to `account_id` in `ledger_entries`.
- `rates.csv`: 15-minute snapshots of USD exchange rates (USD per 1 unit of currency)
- `trades.csv`: Trade logs, including `bid_user_id`, `ask_user_id`, `base_currency`, `counter_currency` and `volume.`

Data Preparation

You will need to create an analysis-ready dataset that links each trade to its user, timestamp, market pair, and USD volume. You may decide on the join strategies, rate-matching approach, and any cleaning or transformations. Clearly document your methodology so it can be reproduced by others.

Define Transacting Status:

As part of your data preparation, you will need to define a new field, namely `transacting_status` that segments customers into one of the following mutually-exclusive groups per month:

- **New:** First transaction in the dataset occurs in this month.
  Returning: User had at least one transaction last month and at least one this month.
  Churned: User had transactions last month but none this month.
- **Reactivated:** User has a transaction this month after being marked "Churned" in the prior month.

Ensure each user-month receives exactly one status.

Final Dataset Structure:
 Your final table (or DataFrame) should include these columns:

- `transaction_timestamp` (UTC)
- `month`
- `user_id`
- `transacting_status` *(New, Returning, Churned, Reactivated)*
- `volume_usd`
- `market_pair`

Analysis & Storytelling

Data storytelling is a core skill we value; how you present your findings can make or break stakeholder engagement. Take care to:

- Frame insights with clear context so both technical and non-technical audiences understand the impact.
- Use visuals strategically (charts, tables) to highlight key trends without overwhelming details.
- Document assumptions, methodology, and edge cases so your analysis is fully reproducible.
- Craft a narrative that connects your metrics to business outcomes.

(The above list is illustrative; we are looking for thoughtful communication, concise explanations, and strong justification for your design choices.)

Notes

- Treat all timestamps as UTC.
- Use the closest available rate for USD conversion. (Should this prove difficult, you can use an hourly or daily averaged rate as an approximation.)
- You will need to create the `market_pair` field by concatenating `base_currency` and `counter_currency` in the trades table. This can be mapped to `ledger_entries.csv` by matching `foreign_id` with the trades.csv `id`.
- `balance_delta` represents the transaction amount in its native currency.
- Since `ledger_entries.csv` records both the debit *and* credit legs of every trade, simply summing all usd equivalent `balance_delta` values for a market will always net out to zero. You will need to account for this in your calculation of trade volume usd.

Good luck and feel free to ask any clarifying questions at any time!