Pattern Identification Report

Igor Reis

October 15, 2025

Overview

The transaction dataset reveals distinct spending dynamics across categories. Some, such as travel or electronics, show irregular high-value purchases followed by quiet periods, while others, like rent or subscriptions, remain stable and predictable. This difference reflects two behavioral regimes: discretionary and recurrent spending. Analyzing these variations helps identify where volatility, anomalies, or structural changes in spending occur.

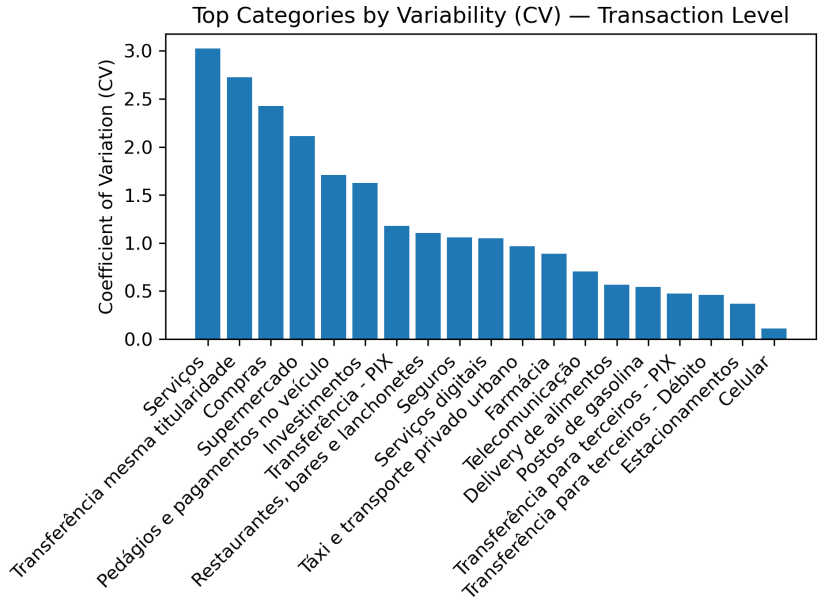

Figure 1: Coefficient of Variation (CV) by category at the transaction level.

Volatility Metrics

Spending variability was quantified using two complementary measures. The Causal Coefficient of Variation:

$$CV_t = \frac{\sigma_{\text{past months}}}{\mu_{\text{past months}}},$$

is computed recursively per category using only historical data, providing a realistic, time-consistent indicator of volatility. A Robust CV substitutes the mean and standard deviation with the median and MAD:

$$CV_t^R = \frac{1.4826 \times \text{MAD}_{\text{past months}}}{\text{Median}_{\text{past months}}},$$

offering stability under outliers, zeros, or irregular distributions. Together, these metrics describe both smooth fluctuations and erratic patterns, enabling adaptive thresholds tailored to each user's spending habits.

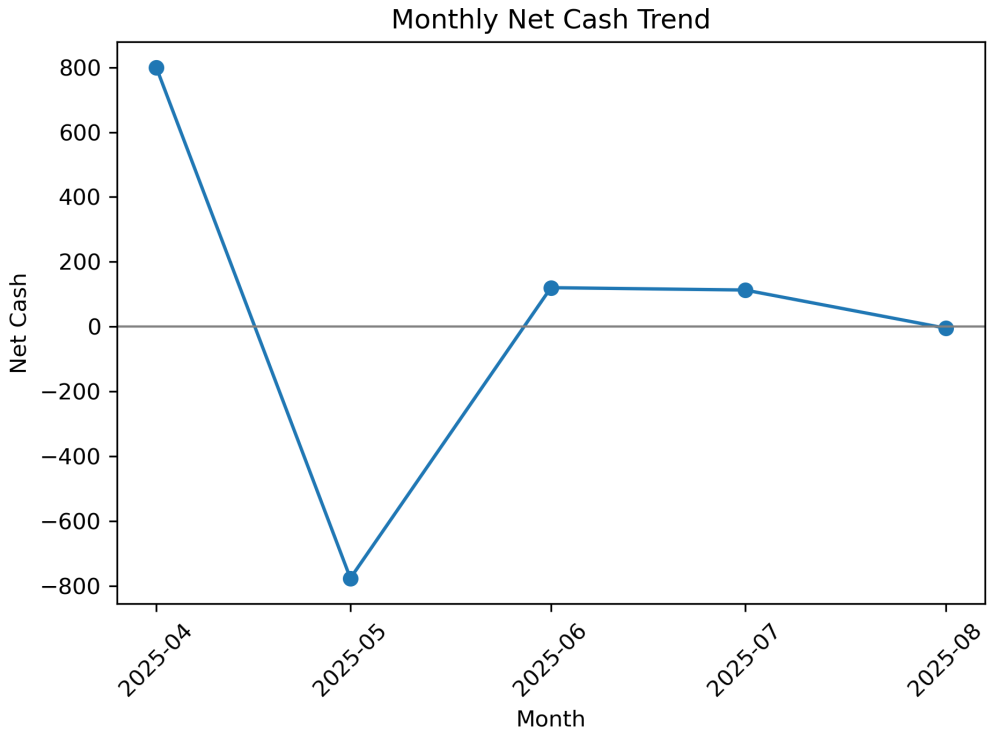

Figure 2: Monthly Net Cash Trend — overall inflows and outflows through time.

Adaptive Outlier Detection

To detect abnormal transactions, a causal rule combines classical and robust scores. For each transaction with effective outflow x, it is flagged as anomalous if:

$$|z| > \text{thr}_z \quad \vee \quad |z| > 3.5 \quad \vee \quad x > Q_3 + 1.5(Q_3 - Q_1).$$

Here, $z = (x - \mu_{\text{prev}})/\sigma_{\text{prev}}$, and thr_z adapts between 2.5 (stable) and 4.0 (volatile) according to the historical coefficient of variation. All computations are causal, using only past data. Refunds are paired with original transactions to form an effective outflow, and quickly reversed charges ($\geq$90% refunded within 14 days) are excluded to prevent false positives.

Conclusion

This framework detects both abrupt anomalies and long-term shifts in financial behavior, providing a compact, data-driven foundation for adaptive personal finance monitoring.