Payment Fraud & AML Risk Analysis

SQL Case Study

Candidate: Navina M

Domain: Payments, Transaction Monitoring, AML

Tools & Skills: SQL Server, CTEs, Window Functions, Risk Calibration

Target Roles: AML Analyst | AML Team Lead | Risk Analyst | Compliance Analyst | Data Analyst

Executive Summary

This case study presents an end-to-end **transaction monitoring and AML risk analysis** using SQL on payment transaction data.

The objective was to design **realistic AML monitoring rules**, identify suspicious behavioral patterns, and **prioritize high-risk users** while controlling false positives.

From a dataset of **500 users and 4,900 transactions**, a calibrated rule framework identified **40 high-risk users (8%)**, aligning with realistic industry alert rates.
The analysis focuses on **failed transactions, velocity fraud, high-value anomalies, geographic risk exposure**, and **composite risk scoring**, reflecting how modern AML teams operate in production environments.

---

1. Business Problem

Payment platforms face several AML challenges:

- High transaction volumes across regions
- Identifying suspicious behavior without over-flagging
- Detecting velocity and high-value anomalies
- Prioritizing investigations based on risk severity

The goal of this case study is to demonstrate **how SQL can be used to build explainable, defensible AML monitoring logic** that balances detection effectiveness with operational efficiency.

---

2. Dataset Overview

METRIC	COUNT
Users	500
Transactions	4,900
Flagged Users	40
Flag Rate	8%

Tables Used

- **users_payment_v3** – User demographics and country
- **transactions_payment_v3** – Transaction amount, status, date
- **risk_flags_payment_v3** – Final AML risk flags

The dataset is synthetic but **calibrated to reflect realistic payment behavior**.

---

3. AML Monitoring Framework

The analysis is structured around four AML dimensions:

- ### Transaction Behavior
  - Failed vs successful transactions
  - User-level failure rates
  - Transaction amount distribution
- ### Velocity Risk
  - Multiple transactions within short time windows
  - Same-day transaction bursts
- ### Geographic Risk
  - Country-wise failure rates
  - Transaction volume exposure
- ### Risk Prioritization
  - Composite risk scoring
  - Severity-based investigation focus

---

4. Key Analytical Findings

Failed Transaction Analysis

- Users with repeated failed transactions show elevated risk
- Failure **rate** is a stronger signal than failure count

Insight:
High transaction volume alone is not suspicious without abnormal failure behavior.

Transaction Volume & Amount Patterns

- Top users generated **₹0.6 – 0.7 million** in successful transactions
- High-value transactions were rare and treated as **high-severity risk indicators**

Insight:
Thresholds were scaled to prevent flagging legitimate high-volume users.

Velocity Fraud Detection

- Users with ≥3 transactions on the same day showed higher risk
- Velocity alone was insufficient without volume context

Insight:
Velocity rules must be combined with transaction volume and value.

---

5. Advanced Risk Insights

Time-Based Fraud Patterns

Results Messages

	txn_week	failed_txn_count
1	4	228
2	3	210
3	2	209
4	5	141
5	1	124

Observation

- Failed transactions cluster around specific weeks
- Week 4, 3, and 2 show the highest failure concentration

Interpretation

- Indicates coordinated fraud attempts or systematic payment testing

AML Relevance

- Fraud activity often occurs in bursts, not randomly

Recommendation

- Apply time-based adaptive thresholds
- Increase monitoring sensitivity during high-risk periods

Country-Level Risk Exposure

Results Messages

	country	total_txn_count	failed_txn_count	failure_rate	total_amt_Millions	failed_amt_Millions
1	Canada	2084	406	19.48	52.21	10.33
2	US	834	154	18.47	20.70	4.08
3	India	1334	241	18.07	33.30	5.62
4	UK	648	111	17.13	16.59	2.93

COUNTRY	FAILURE RATE (%)	TOTAL AMOUNT (MILLIONS)
Canada	**19.48% (Highest)**	52.21
US	18.47%	20.70
India	18.07%	33.30
UK	**17.13% (Lowest)**	16.59

Insights

- High transaction volume does not equal high risk
- Canada exhibits elevated failure behavior
- UK shows comparatively stable transaction patterns

Recommendation

- Apply country-specific risk thresholds
- Avoid one-size-fits-all monitoring logic

Composite User Risk Scoring

Results | Messages

	user_id	failed_txns	high_value_txns	total_txns	risk_score
1	10308	7	18	28	68
2	10409	6	18	29	66
3	10405	4	19	25	65
4	10404	10	14	24	62
5	10407	7	16	27	62
6	10008	7	15	28	59
7	10108	5	16	28	58
8	10204	6	15	24	57
9	10207	4	16	27	56
10	10107	8	13	27	55

Query... | NAVINA\SQLEXPRESS (16.0 RTM) | NAVINA\ASUS (53) | payments_fraud_risk | 00:00:00 | Row: 1, Col: 1 | 104 rows

Methodology

Risk score calculated using:

- Failed transactions (weight = 2)
- High-value transactions ≥ ₹25,000 (weight = 3)
- Total transaction count

Outcome

- Only high-severity users crossed the risk threshold
- Prevented over-flagging of normal high-volume users

Key Insight

Risk emerges when **failure behavior, transaction value, and activity volume intersect**.

AML Relevance

- Enables risk-based queue prioritization
- Supports efficient analyst workload management

---

6. Risk Calibration & False Positive Control

Initial rule design resulted in excessive alerts.
Through iterative calibration:

- Fixed thresholds were replaced with scaled logic
- Multi-condition gating was applied
- Alert volume stabilized at **8% of users**

Key Takeaway

Effective AML systems prioritize precision over raw detection volume.

---

7. Final Risk Outcomes

RISK CATEGORY	USERS FLAGGED
High Transaction Amount	17
Velocity Fraud	15
Multiple Failed Transactions	8
Total Flagged Users	**40**

- One flag per user
- Balanced distribution
- Realistic alert volume

---

8. Operational Impact & Alert Management

Effective AML systems must balance **risk detection accuracy** with **operational feasibility**. Over-flagging leads to analyst fatigue, delayed investigations, and increased false positives, while under-flagging increases regulatory and financial risk.

In this analysis:

- **40 users (8%)** were flagged out of 500 total users
- Flags were distributed across **three distinct risk categories**, preventing concentration bias
- Each user was assigned **a single dominant risk flag**, simplifying investigation workflows

Operational Implications

- At an alert rate of **8%**, the framework produces a **manageable investigation queue**
- Risk-based prioritization ensures analysts focus on **high-severity users first**
- Composite risk scoring enables:
    - Efficient case triage
    - SLA-driven investigation workflows
    - Better allocation of analyst capacity

Assuming a scaled environment, this approach supports **linear growth in transaction volume without proportional increases in analyst workload**.

Key Takeaway

AML effectiveness is measured not by the number of alerts generated, but by the **quality, explainability, and actionability of those alerts**.

This design aligns with industry best practices by ensuring **investigation efficiency**, **regulatory defensibility**, and **analyst sustainability**.

---

9. Recommendations

1. Use percentile-based thresholds instead of fixed values
2. Combine velocity signals with transaction context
3. Implement composite risk scoring for prioritization
4. Continuously monitor alert volumes
5. Recalibrate rules as transaction scale evolves

---

10. Limitations

- Rule-based approach (no ML)
- Synthetic dataset
- Limited historical depth

Despite this, the framework closely mirrors **production AML monitoring systems**.

---

11. Conclusion

This case study demonstrates how SQL can be used to design **realistic, explainable AML transaction monitoring systems**.
By balancing detection effectiveness with false-positive control, the analysis reflects best practices used by modern AML and risk teams.
