Credit Card Transaction Analytics: An Enterprise Data Analytics Case Study

Executive Summary

This case study examines the application of data analytics to credit card transaction data for a leading US card issuer FinPay. By leveraging large-scale transaction data, spanning over 1.29 million records with information from 983 customers and 693 merchants. This report derives insights into customer behavior, merchant performance, category trends, and fraud detection. The analysis reveals vital opportunities for growth, improved risk controls, and actionable business strategies.

Table of Contents

- Executive Summary
- Organization Background
- Data Landscape and Methodology
- Customer Analytics
- Merchant Performance Insights
- Category-Level Insights
- Fraud Detection and Risk Analysis
- Time-Series and Growth Trends
- Strategic Recommendations
- Conclusion
- References

Organization Background

The subject of this study is a major US credit card issuer FinPay with a nationwide merchant acceptance network. The business objectives are:

- Grow card usage through increased spend per active customer
- Strengthening the merchant ecosystem for broad and profitable acceptance
- Detect and prevent fraud to reduce financial exposure

The company processes millions of transactions annually across point-of-sale (POS) and e-commerce channels. Merchants are categorized by industry segment (e.g., Grocery POS, Shopping POS, Utilities, Travel).

Data Landscape and Methodology

Dataset Overview

- **Transactions:** 1,290,000+
- **Customers:** 983
- **Merchants:** 693
- **Categories:** 14
- **Dates Represented:** 434 (2019-01-01 to 2020-06-21)

Data Model Structure

The data is organized into a star schema:

- Fact table: transactions
- Dimension tables: customers, merchants, categories, date

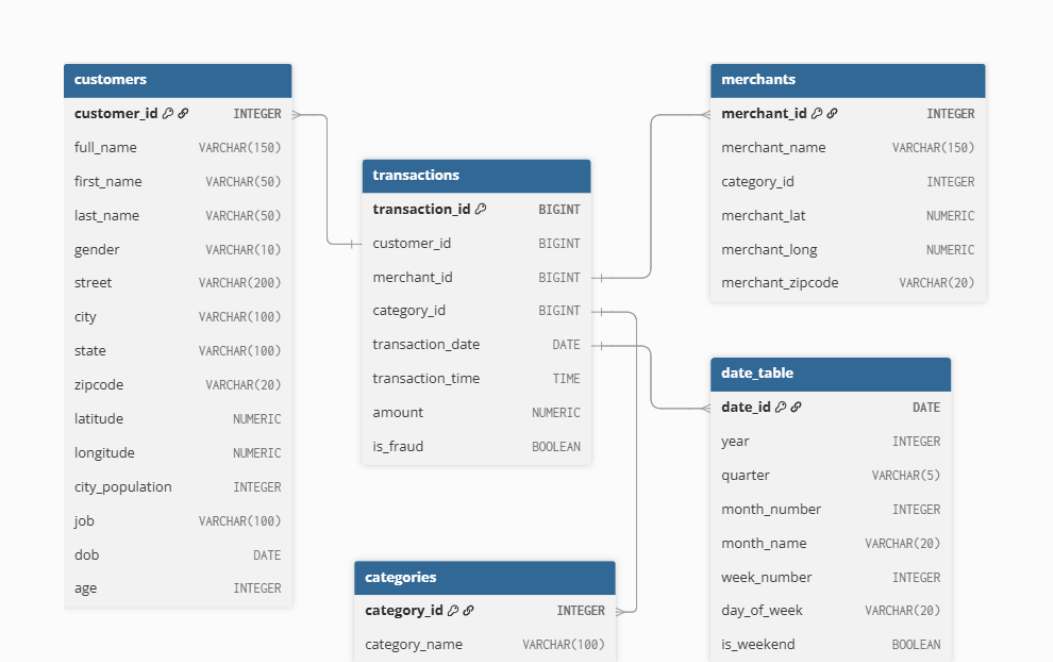

Figure 1: Star Schema Used in the Analysis

Processing Steps

1. Raw CSV ingested via Python
2. Data cleaning: nulls/malformed records removed, identifiers validated
3. Data normalization into fact and dimension tables
4. Loaded into PostgreSQL for exploratory data analysis and SQL querying
5. Aggregated results exported to Python for visualization and modeling

Tools Used

- PostgreSQL (data storage, SQL analytics)
- Python (pandas for data processing, matplotlib & seaborn for visualization)

Customer Analytics

Overview

- **Customers analyzed:** 983 (with 94% active in the last 90 days)
- **Average revenue per customer:** $92,800, with heavy repeat transaction behavior
- **High-value segment:** Top customers each contribute $275K–$296K

Metric	Value	Comment
Active Customers (90 days)	924	94% retention
Avg. Revenue per Customer	$92,800	High lifetime value
VIP Segment Revenue	$275K--$296K	Top contributors
Max. Transactions per User	3,100+	Repeat users

Table 1: Key Customer Metrics Overview

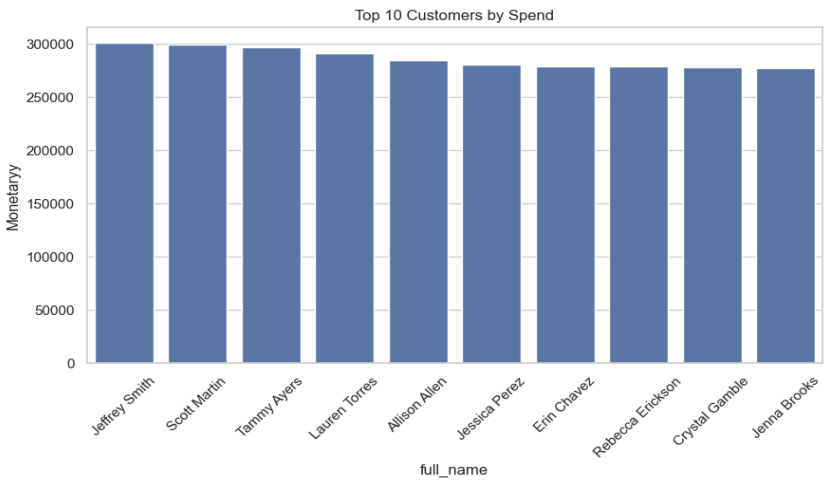

Figure 2: Top 10 Customer by spend

RFM (Recency, Frequency, Monetary) Summary

- Average Recency: 3 days
- Average Frequency: 3 purchases
- Average Monetary: $3,000 (estimated)

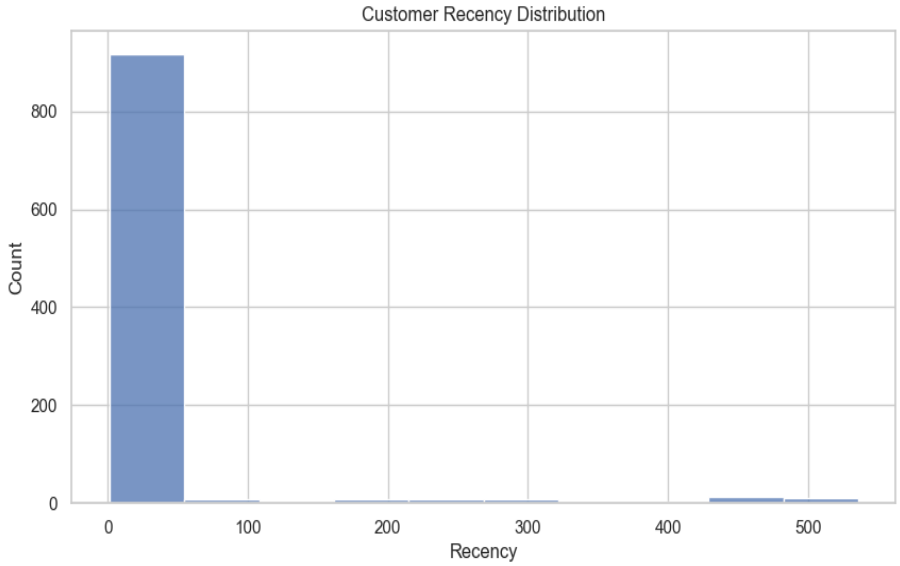

Figure 3: Customer Recency Distribution

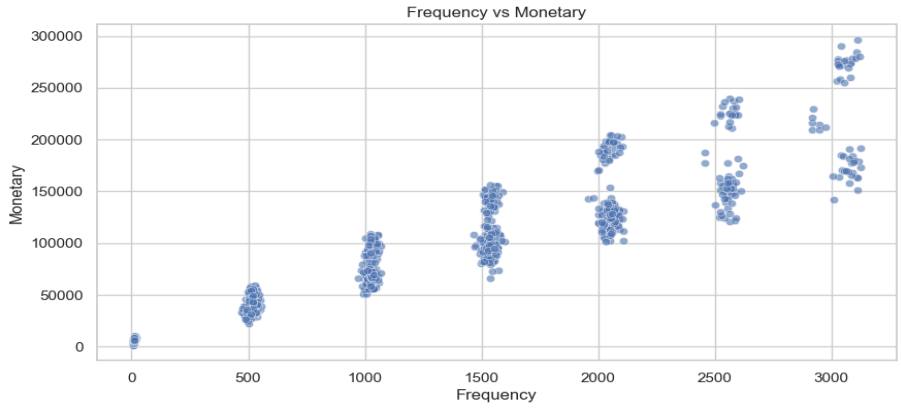

Figure 4: Frequency vs Monetary Scatterplot

Insights

- Frequent usage suggests strong product-market fit and integration in users' daily spend
- Churn risk: small minority inactive over last 90 days—key opportunity for targeted recovery campaigns

Merchant Performance Insights

- **Total Merchants:** 693
- **Revenue Distribution:** Highly skewed—Top 10 merchants each generate $295K–$391K
- **Pareto Pattern:** <2% of merchants account for a majority of sales

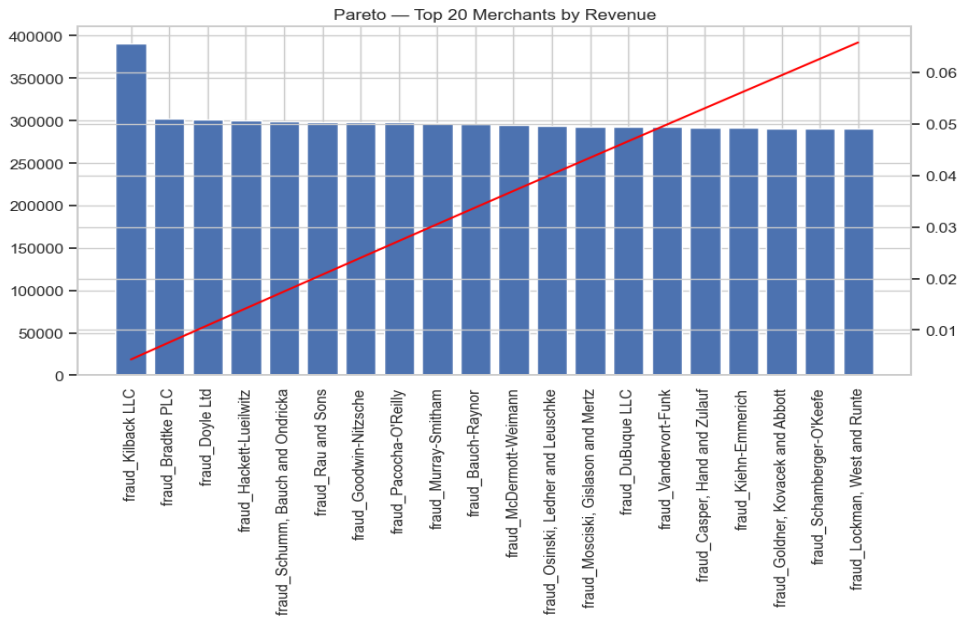

Figure 5: Top 20 Merchants by Revenue

- **Activeness:** 100% of merchants have processed transactions in the last 90 days

Merchant Group	Revenue Range	Notable Traits
Top 10 Merchants	$295K--$391K	High volume, strategic partners
Volume Leaders	3,500+ tx	Lower AOV, frequent needs categories
High-Ticket Merchants	$129--$166/order	Upsell targets

Table 2: Merchant Performance Segmentation

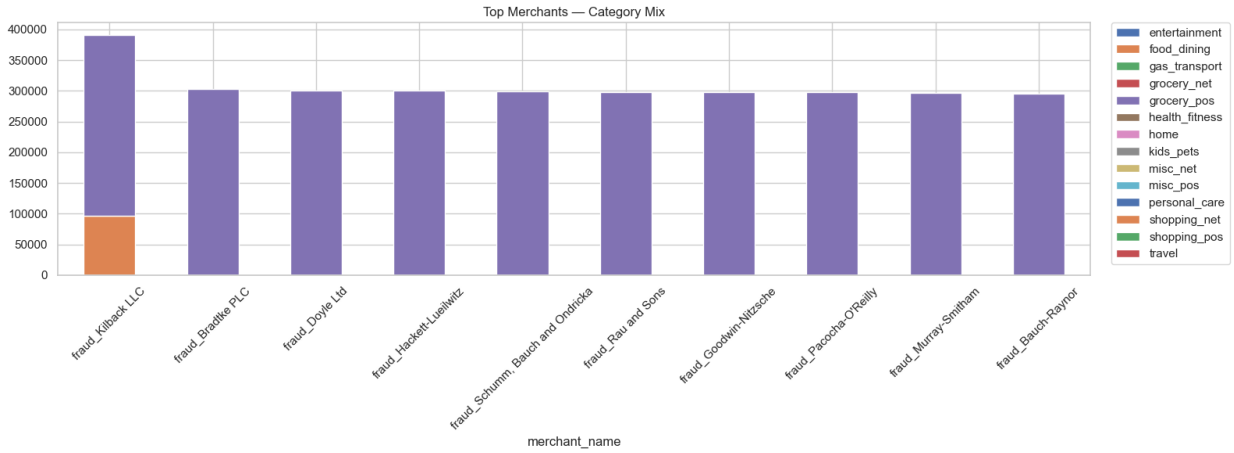

Figure 6: Top Merchants category Breakdown

Fraud Patterns Among Merchants

- Fraud is concentrated: Some merchants see fraud rates of 2–2.57%
- These should be priority for advanced risk controls and targeted audits

Category-Level Insights

- Retail (POS) categories dominate revenue and volume, led by Grocery POS
- **Top Revenue Categories:**
    - Grocery POS: 15.85%
    - Shopping POS: 10.20%
    - Shopping Net: 9.46%
    - Gas/Transport: 9.16%
    - Home, Kids/Pets: ☐15% combined

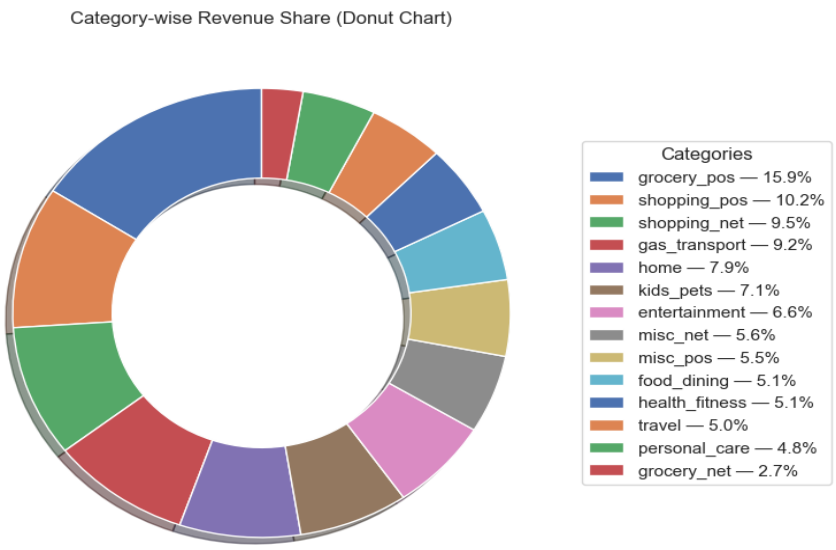

Figure 7: Category-Wise Revenue Share

- **High-value categories:** Travel and Grocery POS (high AOV)
- **Fraud risk clusters:** Shopping Net, Misc Net, and Grocery POS have elevated fraud rates (1.76% or higher)

Category	Revenue Share (%)	Risk/Opportunity
Grocery POS	15.85	High, essential spend; fraud risk
Shopping Net	9.46	Discretionary, high fraud risk
Travel	4.95	High AOV, target for campaigns
Gas/Transport	9.16	Recurring spend anchor
Personal Care	4.77	Low value, steady revenue stream

Table 3: Category-Level Revenue and Risk Overview

Fraud Detection and Risk Analysis

- **Overall fraud rate:** 0.58% (7,506 of 1.3M transactions)
- **Temporal clusters:** Fraud spikes during holiday shopping and lockdown peaks
- **Channel risk:** Card-not-present (online) and high-velocity categories are most affected
- **Merchant hotspots:** Handful of partners show 2%+ fraud—possible systemic weaknesses

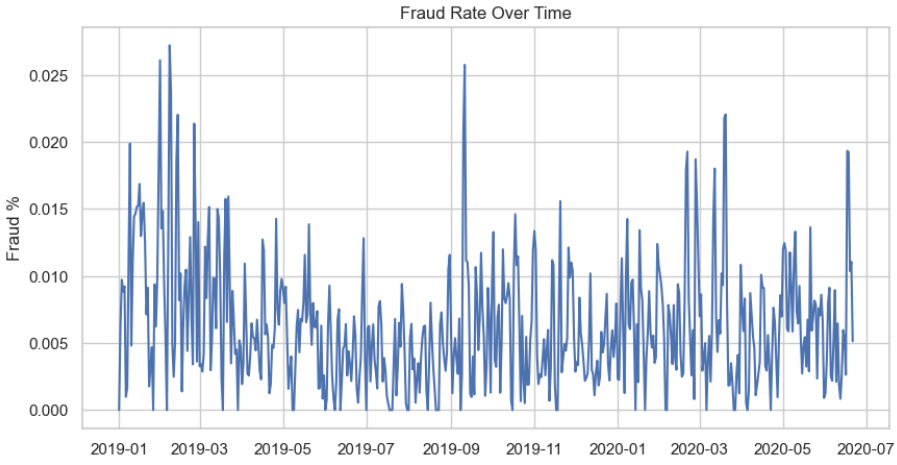

Figure 8: Fraud Rate over Time

High-Risk Behavior Highlights

- Late-night transactions (midnight–3AM) have above-average fraud rates
- A few customers appear in 15–19 separate fraudulent events, suggesting account takeover risks
- Tighten rules for late-night, high-velocity activity
- Audit and monitor high-exposure merchants intensively
- Migrate toward real-time, behavior-based fraud alerts rather than static threshold rules

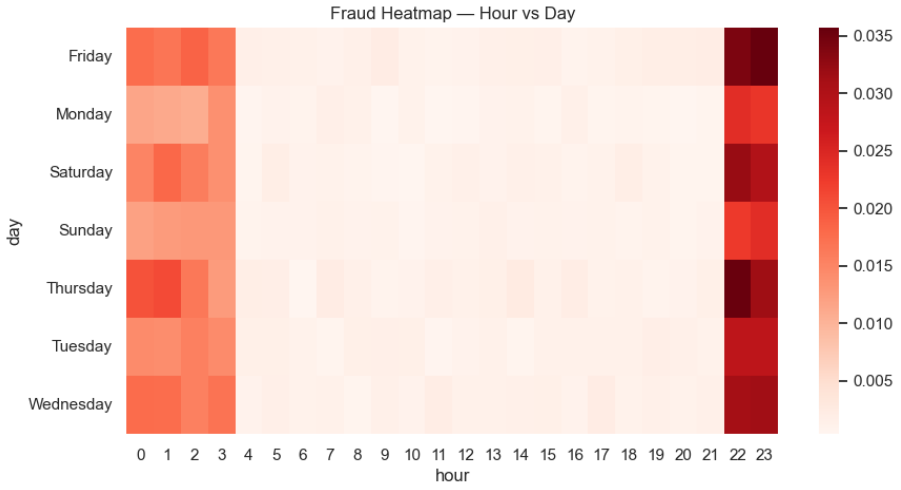

Figure 9: Fraud Heatmap

Time-Series and Growth Trends

Revenue Growth

- Month-over-month revenue trends highlight both seasonal spikes (Dec 2019: +10%) and short-term dips (Jan 2020: --6.3%)
- Week-on-week transaction growth supports seasonal demand and promotional period impact

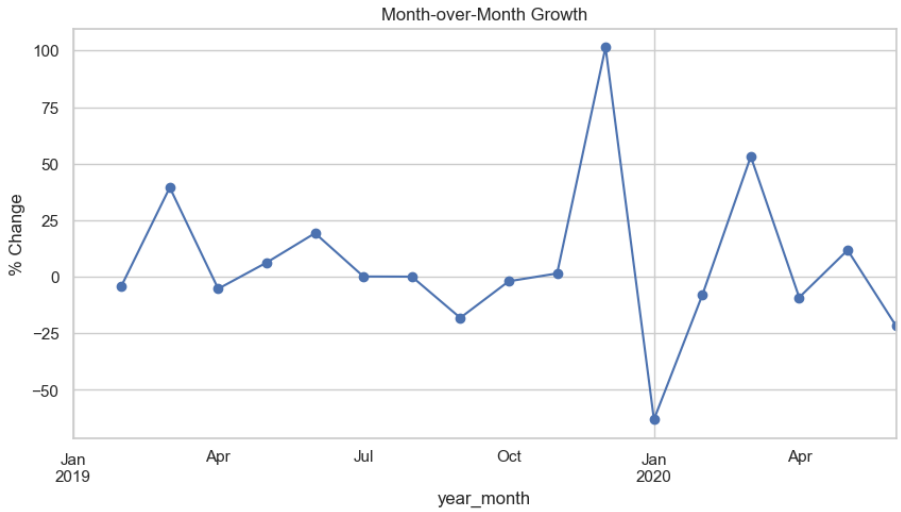

Figure 10: Month-over-Month Growth

Rolling Metrics

- 3-month and 6-month moving averages smooth volatility for clearer trend analyses
- Rolling 7-day fraud rates reveal short-term vulnerabilities, supporting daily operational risk management

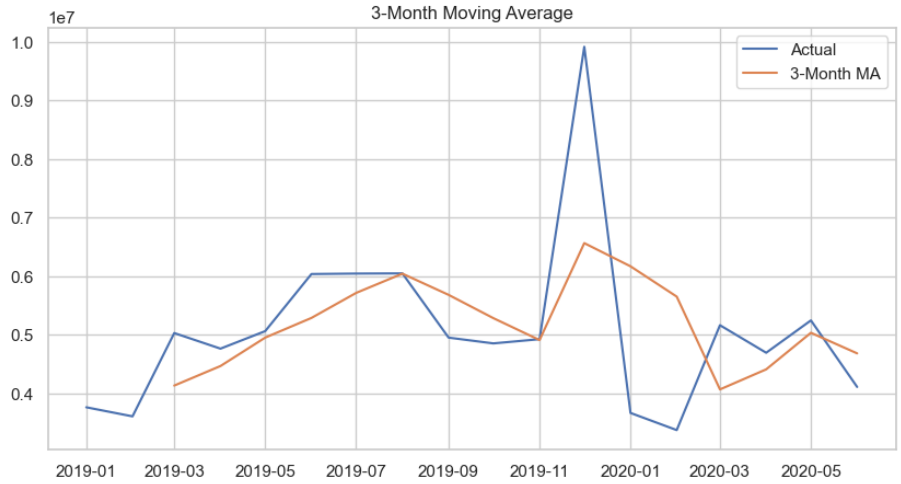

Figure 11: 3 Month Moving Average

Cumulative Performance

- 2020 YTD (Year-to-Date) revenue: $26.24M
- Q2 (Quarter-to-Date) revenue: $14.04M

Strategic Recommendations

1. **Customer Growth:** Expand retention programs for mid-tier, high-loyalty customers; launch win-back campaigns for near-churned users
2. **VIP Program:** Strengthen engagement for top-performing customers with exclusive incentives
3. **Merchant Strategy:** Prioritize relationship management with top revenue and frequency leaders; tie incentives to both revenue and risk controls
4. **Category Push:** Invest in high-potential categories (Home, Kids/Pets) and steady anchors (Grocery, Gas)
5. **Fraud Management:** Employ advanced analytics for real-time fraud detection; focus on high-risk categories, periods, and merchant clusters
6. **Operational Analytics:** Integrate moving averages and rolling windows into dashboards for proactive decision making
7. **Promotion Optimization:** Time promotional activity around known seasonal/resource spikes to optimize marketing ROI

Conclusion

This data analytics project delivers a comprehensive view of customer spending, merchant dynamics, category-specific trends, and fraud risks within a major credit card transaction ecosystem. By applying SQL-driven analytics and visualization best practices, the business can strengthen customer relationships, extract more value from key merchant partnerships, optimize its category mix, and safeguard against evolving fraud threats. Adoption of these recommendations will support continued growth, profitability, and operational resilience.

References

[1] SQL Project Enterprise Case Study Credit Card Transaction Analytics. (2025). Internal report.

[2] Omdena. (2025). Data Analytics Case Study Examples for Businesses 2025. https://www.omdena.com/blog/data-analytics-case-study-examples-businesses-2025

[3] Modern Analyst. (2024). Writing a Good Data Analysis Report: 7 Steps. https://www.modernanalyst.com/Resources/Articles/tabid/115/ID/6070/Writing-a-Good-Data-Analysis-Report-7-Steps.aspx

[4] Contentsquare. (2025, November 5). How Marketers Can Write Data Analysis Reports (+ Examples). https://contentsquare.com/guides/data-analytics/report/

[5] Metabase Learn. (2015). Best practices for writing SQL queries. https://www.metabase.com/learn/sql/working-with-sql/sql-best-practices

[6] Datawithsarah.com. (2025, November 25). SQL Best Practices for Writing Clean, Reliable Queries. https://datawithsarah.com/post/sql-best-practices-for-writing-clean-reliable-queries/