blue mantis

Case study

How a National Bank Built a Predictive Fraud Detection Foundation on 5 Billion Transactions

Modern data engineering. Faster fraud detection. Lower analytical warehouse costs.

Client Overview

A national bank processes about 60 million credit and debit card transactions per day and needs to continuously integrate new data with historical data, which totals roughly 5 billion transaction records every 90 days. To better predict credit and debit card fraud, analysis had to be done across this massive volume of data, which is one of the top big data use cases in financial services.

Challenges

- With the exponential growth in digital transactions, the bank was grappling with efficiently handling and making sense of over 5 billion credit and debit card transaction records stored in their transactional database
- Transactional data needed to be integrated into the data warehouse for real-time analysis and securely archived in the data lake for future reference and compliance
- Using Power BI to connect directly to the entire dataset hit the memory threshold, causing reload failures
- Using direct query required a larger warehouse, leading to significant cost escalation if used consistently throughout the day

60M / Day

Transactions Processed for Fraud Analysis

After Partnering With Blue Mantis, the Bank Achieved:

- Discovered actionable insights by visualizing anomalies in the data
- Narrowed from 5 billion records down to specific accounts with unusual spending patterns
- Significant reduction in analytical warehousing expenses
- Faster time to insight for fraud investigations

Solutions

Blue Mantis designed a modern data engineering solution to harness 3 months of transaction data for more precise fraud identification, laying the groundwork for predictive fraud detection while improving report performance, speed, and reliability.

The technical team used Azure Data Factory to move data from the transactional database (Azure SQL) to Azure Data Lake Storage, then partitioned the data, stored it in the data lake, and loaded it into Snowflake. Complex data transformations were performed within Snowflake to prepare a data mart for analysis, and Power BI direct queried the transformed data mart in Snowflake to deliver fast, reliable insights.