رواد مصر الرقمية

AMAN REPORT

Project Title: **Payment Security – Smart Fraud Detection & Analysis**
Author: Data Engineer Team
Affiliation: Data Department

Executive Summary

This project delivers a comprehensive analysis of a large-scale synthetic banking dataset comprising financial transactions. The primary objective is twofold: to **understand customer behavior** and to **detect, measure, and visualize fraudulent transaction patterns.**

A fully interactive Power BI dashboard was developed to empower banking institutions to effectively identify high-risk behaviors, understand spending paflerns, and highlight fraud hotspots. This analytical solution translates raw data into actionable intelligence, enabling proactive risk management and enhanced

.customer protection

Dataset Overview

Key Performance Indicator (KPI)	**Value**
Total Transactions	+1,000,000
Unique Customers	200,000
Total Valid Transaction Amount	Billion $9.41
Total Fraudulent Amount	Million $497.12
Fraudulent Transactions (Count)	(of Total 1%) 10,000
Fraud Rate by Value	**5.02%~**

The dataset contains granular aflributes, including TransactionID, CustomerID, Age, Gender, AccountBalance, TransactionAmount, MerchantCategory, DeviceType, Location, IsFraud, and

.transaction timestamps

Key Customer Insights

Demographic Distribution

Gender: A nearly equal split between **Male (50.45%)** and **Female (49.55%)** customers .suggests no significant gender bias in banking activity

Age Group: The **26–35 age group** is the most active in terms of transaction volume. However, the **56+ age group** records the highest average transaction amount. This indicates that younger customers transact more frequently, while older customers make higher-value purchases .

Account-Type Behavior

.**Savings Accounts** maintain the highest average balance

.**Business Accounts** drive the largest total spending

.**Current Accounts** exhibit steady, consistent daily activity

Critical Fraud Findings

Disproportionate Financial Impact

Although fraud constitutes only **1%** of total transaction volume, it accounts for over **5%** of the total monetary value lost. This makes a single fraudulent transaction, on average, **10 .times more damaging** than a legitimate one

Fraudulent vs. Valid Transaction Value

Average Fraudulent Transaction: **~$49,712**

Average Valid Transaction: **~$4,956**

This stark difference confirms that fraudsters strategically target high-value transfers to .maximize their gains

Temporal Fraud Patterns

A sharp spike in fraudulent activity was detected between **1:00 AM and 5:00 AM**. These ."low-monitoring hours" represent a critical risk window for the bank

High-Risk Geographic Hotspots

Unusually high fraud totals were recorded in specific locations, including **Kavaratti,**

.**Chandigarh, and Silvassa**, marking them as high-risk zones requiring closer scrutiny

Most Targeted Merchant Categories

The majority of fraud occurs in the **Clothing, Restaurants, Electronics, and Entertainment** categories. These sectors are characterized by high-volume, .quick-turnover purchases, which makes real-time fraud detection more challenging

Actionable Red Flags for Fraud Detection

Risk Factor	**Key Observation & Threshold**
Transaction Hour	Peak fraud occurs during the **1 AM – 5 AM** .window
Transaction Amount	Any single transaction exceeding **$30,000** has an extremely high likelihood of being .fraudulent
Victim Age Group	The **26–35 age group** is the most .frequently victimized demographic
Merchant Category	**Clothing & Restaurant** purchases warrant .heightened suspicion
Geographic Anomaly	Sudden transaction spikes in historically low-activity regions are a major warning .sign

Technical Enhancements to the Data Model

To enable this analysis, custom columns and DAX measures were engineered to enrich the :dataset

Calculated Columns

.AgeGroup: To segment customers into distinct age brackets for targeted analysis

Transaction_Hour: To extract the hour of the day from the timestamp, enabling .time-based paflern detection

Key DAX Measures

.TotalFraudTransactions: To count the total number of fraudulent events

.FraudPercent: To calculate the percentage of transactions that are fraudulent

AvgFraudAmount & AvgValidAmount: To compare the financial impact of fraudulent vs. .legitimate transactions

.FraudLossRatio: To quantify the proportion of total transacted value lost to fraud

These custom calculations serve as the analytical engine powering the visual insights on the .Power BI dashboard

Actionable Recommendations for the Bank

Enhance Real-Time Monitoring .1

.Implement **automated alerts** for all transactions exceeding a **$25,000** threshold

Enforce **Multi-Factor Authentication (MFA)** for transactions initiated during .(high-risk hours (1 AM – 5 AM

Develop Advanced Fraud-Scoring Models .2

Build a real-time risk-scoring model that uses **transaction amount, hour of the .day, location anomaly, and merchant category** as key input features

Launch Targeted Customer Protection Campaigns .3

Educate customers in the **26–35 age group** on best practices for preventing online .fraud

Provide instant SMS or push notification alerts to senior customers (**56+**) for any

.high-value transactions on their accounts

Strengthen Merchant & Category Monitoring .4

Apply more aggressive and frequent reviews for merchants in high-risk categories,

.particularly **Clothing and Restaurants**

Conclusion

This analysis conclusively demonstrates that while infrequent, fraudulent transactions pose a significant financial threat. The Power BI dashboard makes these paflerns clear and actionable by visualizing **high-risk hours, vulnerable age groups, suspicious transaction**

.values, and compromised merchant categories