

Fraud Detection & Risk Analysis

Project Overview:

- **Project Title:** Fraud Detection & Risk Analysis – Critical Thinking Analysis
- **Tools Used:** Python, SQL (MySQL), Power BI
- **Dataset:** Financial Transaction Dataset

Abstract:

This project analyzes transaction data to detect fraudulent activities and assess financial risk. Python was used for data cleaning and preparation, SQL was used for fraud KPI analysis and pattern detection, and Power BI was used for interactive dashboard visualization. The dashboard provides insights into fraud trends, high-risk regions, and suspicious transaction behaviors to support proactive fraud prevention.

Problem Statement:

Businesses process thousands of transactions daily, making it difficult to detect fraud manually. Fraudulent activities often hide within normal transaction patterns, causing financial losses and operational risk.

This project answers :

"Which transactions show fraud patterns and what risk factors contribute most to fraud?"

By analyzing transaction behavior, location, device type, and amount patterns, this project identifies hidden fraud risks and supports data-driven fraud prevention strategies.

Objectives:

- Identify fraudulent transactions using pattern analysis
- Analyze fraud trends across time, location, and device type
- Measure fraud rate and financial risk exposure
- Detect high-risk transaction behaviors
- Develop an interactive Power BI dashboard for fraud monitoring

Technologies & Skills Used:

SQL (MySQL):

Data extraction, fraud filtering, aggregations, and KPI calculations

Power BI:

Interactive dashboards, KPI cards, trend charts, and filters

DAX:

Custom measures for Fraud Rate %, Fraud Count, and Risk Amount

Python (Pandas, NumPy, Matplotlib, Seaborn):

Data cleaning, EDA, fraud pattern visualization

Data Analysis:

Fraud trend analysis, risk assessment, behavioral pattern detection

Business Intelligence:

Fraud monitoring framework and decision support

Methodology:

1. Data Preparation (Python)

- Cleaned and processed transaction dataset
- Handled missing values and removed duplicates
- Standardized data formats (date, amount, categories)
- Created calculated metrics: fraud_flag, transaction_month, risk_level
- Exported structured data to MySQL database

2. SQL Analysis (MySQL)

Performed comprehensive analysis using SQL queries:

- Overall fraud KPI calculations (Fraud Rate %, Fraud Count, Amount at Risk)
- Monthly fraud trend analysis
- Location-wise fraud comparison
- Device-type fraud analysis
- High-risk transaction identification
- Merchant/category fraud ranking

3. Visualization (Power BI)

Built interactive dashboard featuring:

- KPI cards for key fraud metrics
- Monthly fraud trend analysis
- Fraud distribution by device and location
- Category-wise fraud comparison
- Dynamic filters for date and category

SQL Analysis & Key Findings

Overall Fraud Metrics

Core KPIs analyzed to understand fraud exposure.

	Fraud_Label	Total_Transactions	Percentage
▶	0	33933	67.87
	1	16067	32.13

Observation:

Fraud rate represents a small percentage of total transactions but contributes to significant financial risk.

Location-Based Fraud Analysis

Fraud distribution analyzed across locations

	Location	Total_Transactions	Fraud_Count	Avg_Transaction_Amount
▶	Tokyo	3315	3315	99.83
	New York	3221	3221	100.96
	Sydney	3194	3194	98.03
	London	3181	3181	100.94
	Mumbai	3156	3156	98.64

Observation:

Certain locations show higher fraud frequency, indicating geographic risk patterns

Fraud by Authentication Method

Fraud cases were analyzed across different authentication methods to identify which security mechanisms are most vulnerable.

	Authentication_Method	Total_Transactions	Fraud_Cases	Fraud_Rate	Avg_Transaction_Amount
▶	OTP	12366	4021	32.52	99.92
	Password	12457	4030	32.35	98.41
	Biometric	12591	4038	32.07	99.42
	PIN	12586	3978	31.61	99.9

Observation:

Certain authentication methods show noticeably higher fraud rates, indicating weaker verification processes. Transactions without strong multi-factor authentication are more prone to fraud. This suggests the need for stronger authentication controls and additional verification layers.

Fraud Distribution by Card Type

Fraud distribution was evaluated across card types to understand which user segments face higher risk.

	Card_Type	Total_Transactions	Fraud_Count	Avg_Card_Age	Avg_Amount
►	Mastercard	12693	4038	120	99.23
	Amex	12419	4021	120	100.62
	Discover	12328	4006	120	99.26
	Visa	12560	4002	120	98.56

Observation:

Some card types show higher fraud counts and higher average transaction amounts, indicating targeted fraud behavior. Fraudsters tend to focus on card categories associated with higher spending limits. This highlights the need for enhanced monitoring on high-value card segments.

Account Balance Distribution for Fraud Cases

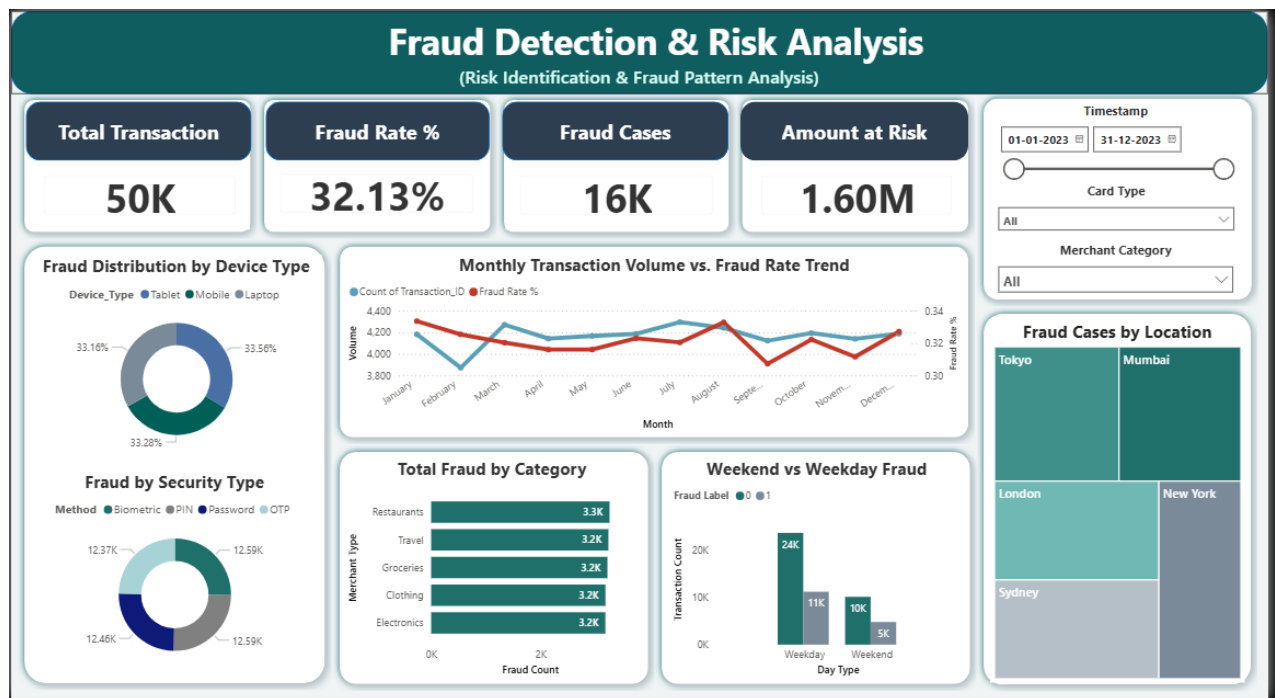
Fraud cases were analyzed based on customer account balance categories to understand whether fraud is concentrated among low, medium, or high-balance accounts.

	balance_category	transaction_count	fraud_count	avg_transaction_amount
►	High (50K-100K)	25166	8095	99.24
	Medium (10K-50K)	19970	6382	100.14
	Low (<10K)	4864	1590	97.33

Observation:

Fraud is not limited to low-balance accounts; it also appears in medium and high-balance categories. Higher-balance accounts often show larger average transaction amounts, making them attractive targets for fraudsters. This indicates that fraud risk increases with potential financial gain.

Power BI Dashboard:



Interactive Features:

- Time and location filters
- Real-time fraud KPI tracking
- Fraud trend analysis
- Device-wise breakdown
- Risk-level segmentation

Dashboard Purpose:

Enables organizations to monitor fraud patterns and take preventive actions based on data insights.

Key Findings

- Overall Performance Metrics
- Total Transaction: 50K
- Fraud Rate %: 32.13%
- Fraud Cases: 16K
- Amount at risk: 1.60M

Business Recommendations

Strengthen Monitoring in High-Risk Areas

Focus fraud detection systems on high-risk locations and devices.

Flag High-Value Transactions

Apply additional verification for large transactions.

Improve Real-Time Alerts

Set automated alerts for suspicious patterns.

Track Fraud KPIs Regularly

Monitor fraud rate and risk exposure monthly.

Conclusion:

This project demonstrates that fraud detection requires proactive data analysis. Using Python, SQL, and Power BI, the analysis uncovered hidden fraud patterns across locations, devices, and transaction values.

The key insight is clear :

Even a low fraud rate can lead to major financial losses if ignored.

By implementing monitoring systems and data-driven fraud prevention strategies, organizations can reduce risk and improve transaction security. This project provides a practical framework for identifying and minimizing fraud risk using analytics.