

Credit Card Usage Trend & Anomaly Detection

Author: Devin Richmond

Date: July 21, 2025

Project Overview

This project analyzes over **44,000+ credit card transactions** spanning from **December 2023 to May 2024**, across **8 countries**. The goal was to uncover customer behavior, spending patterns, and detect unusual activity or fraud using Python, SQL, and Tableau.

Dataset Description

The dataset included the following columns:

- transaction_id, customer_id, merchant
- date, amount, currency, category, country
- card_type, flagged, is_international

Currency was mixed (USD, TWD, JPY, etc.) and the data included missing values and typos.

Step 1: Data Cleaning (Python)


We cleaned the dataset using **Pandas**, focusing on:

- **Date Column:** Converted to datetime, resolved errors.
- **Amount:** Converted to float.
- **Booleans:** is_international and flagged converted to True/False.
- **Category/Card Type:** Fixed typos like "grocerie" to "Groceries".

Handling Missing Values (Storytelling)

I carefully evaluated missing values based on business importance. For example, a transaction with no amount or date isn't usable for analysis, so I dropped it. For categorical fields like category, I filled missing values with 'Unknown' to preserve row-level data.

- 279 missing in date → dropped
- 235 missing in amount → dropped
- 3,299 missing in category → filled as "Unknown"

 **Currency Normalization:** Created amount_usd by converting local currencies to USD using assumed exchange rates.

Step 2: Exploratory Data Analysis (Python)

Key Metrics (KPIs):

KPI	Value
Total Spend (USD)	\$1.3M
Average Transaction	\$54.28
Top Spender	Customer 1097 (\$5,832)
Most Used Card	AmEx (9,681 transactions)
% Flagged	0.7%
% International	20.5%

Monthly Spending Trend

What Happened?

Total credit card spending showed a clear upward trend from February to May 2024, increasing by over 17% in just 3 months (from ~\$199K in Feb to ~\$233K in May).

Why It Matters

Understanding seasonality helps the bank forecast demand and detect anomalies more accurately.

Recommendation

The bank should investigate category-level drivers of this growth and identify which user segments are fueling the increase, this insight can guide targeting for upcoming promotions or reward offers.


Average Transaction Insight

Storytelling

Average transaction size remained stable from December 2023 to May 2024, hovering between \$45–\$48, with a slight dip in February — likely due to smaller, frequent purchases post-holiday. A spike in April or May could indicate large-value categories like travel or electronics gaining traction.

Outlier Detection

- Used **IQR method** to flag high-value transactions
- Outliers defined above \$159.82
- **2,223 outlier transactions** identified

 **Why It Matters:** Outliers may indicate risk, fraud, or big spenders. Useful for fraud teams or marketing targeting.

Behavior Check

Several customers had repeated outlier activity (4+ transactions over threshold) — these customers may represent VIP clients or suspicious patterns.

Step 3: SQL Analysis (SQLite)

Data exported as credit_card_cleaned.sqlite for structured querying.

SQL Business Insights:

- Monthly spend by country:

"In December 2023, UK users had the highest credit card spend at \$67,629.24, followed by the US and Germany."

- Spending by card type:

"AmEx users spent the most overall — \$532,497.08, followed by MasterCard at \$503,708.29, and Visa at \$267,528.41."

- Spend by category
- Top 10 spenders by total
- International usage rates by country

Step 4: Visualization (Tableau Dashboard)

Dashboard Title: *Credit Card Analytics, Flag & Behavior*

KPI Cards:

- Total Spend
- Avg. Transaction Size
- % International
- % Flagged
- Top Spender (Customer ID + Amount)
- Most Used Card

Visuals:

Chart	Purpose
Monthly Spend Trend (Line)	Track seasonal behavior
Spend by Country & Month (Heatmap)	Spot regional volume
Spend by Card Type (Bar)	Compare card usage
Spend by Category (Bar)	Understand category-wise behavior
Flagged Over Time (Bar)	Watch monthly risk trends
% International by Country (Stacked Bar)	Compare global usage
Outlier Transactions (Box Plot)	Spot spending anomalies

Filters (Interactive):

- Card Type
 - Country
 - Date Range
 - Flagged Status
 - Customer ID
 - Is International
-

Business Questions Answered

Fraud / Risk Detection

- 2,223 high-value outliers > \$159.82
- 0.7% of all transactions flagged
- Several repeat-risk customers spotted

Customer Behavior

- Most activity from UK, USA, and Singapore
- Top categories: Electronics, Dining, Travel
- High spend periods: Jan & May 2024
- AmEx cards used most frequently

Finance/Product Insights

- Monthly volume tracked by country
- Breakdown of spend: AmEx (\$532K), MasterCard (\$503K), Visa (\$267K)
- International usage highest in Singapore and UK

Tools Used

- **Python:** pandas, matplotlib, seaborn
 - **SQL:** SQLite + Jupyter SQL magic
 - **Tableau:** KPI dashboard with filters, interactivity, and storytelling visuals
-

Final Result

A complete data analytics project demonstrating:

- ✓ Data cleaning & preparation
- ✓ Insightful storytelling from KPIs to anomalies
- ✓ SQL query logic for performance metrics
- ✓ Interactive Tableau dashboard with real-world use cases

This project is **portfolio-ready** and ideal for roles in **finance, marketing analytics, or risk detection**.