

WEEK 10&11 EXERCISE

Transaction Data Analysis Report

-Table of Contents

1. Introduction
 - Overview of the report and objectives.
2. Dataset Overview
 - Description of dataset attributes.
 - Summary statistics of the dataset.
3. Data Cleaning
 - Handling missing values, duplicates, and outliers.
 - Summary of cleaned data.
4. Exploratory Data Analysis (EDA)
 - 4.1 Distribution of Transaction Amounts
 - 4.2 Proportion of Transaction Types
 - 4.3 Transaction Amount vs Fraud Flag
 - 4.4 Comparison of Transaction Types by Status
 - 4.5 Device Usage Analysis
5. Analyzing Feature Interactions
 - 5.1 Average Transaction Amount for Each Transaction Type
 - 5.2 Transaction Amounts by Time of Day
 - 5.3 Correlation Between Device Used and Transaction Amount
 - 5.4 Trends in Fraud Occurrences Over Time
 - 5.5 Top Sender and Receiver Accounts by Average Transaction Amount
6. Summary
 - Key findings from the analysis.
7. Conclusion
 - Actionable insights and recommendations.

Introduction

This report analyzes a dataset containing financial transaction details, including transaction types, amounts, statuses, fraud flags, geolocation, and device usage. The analysis aims to identify patterns, correlations, and actionable insights that can assist in improving transaction security and efficiency.

1.Dataset Overview

The data set only contains the Transactions data of the date 17-1-2025

The dataset contains 1,000 rows and 10 columns with the following attributes:

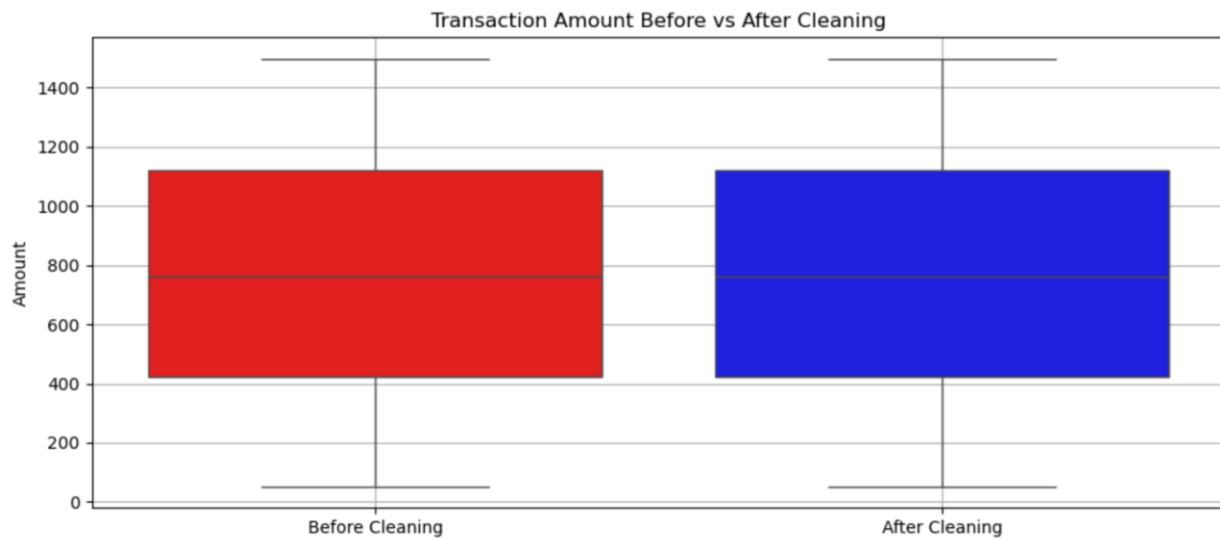
- Transaction ID: Unique identifier for each transaction.
- Sender/Receiver Account IDs: Account IDs involved in transactions.
- Transaction Amount: Monetary value of the transaction.
- Transaction Type: Categorized as Deposit, Withdrawal, or Transfer.
- Timestamp: Date and time of transaction occurrence.
- Transaction Status: Indicates whether the transaction was successful or failed.
- Fraud Flag: Boolean flag indicating fraudulent transactions.
- Geolocation: Latitude/Longitude of transaction initiation.
- Device Used: Device type (Desktop/Mobile).

2..Data Cleaning

1. Missing Values:
 - No missing values were detected in the dataset.
2. Duplicate Records:
 - No duplicate rows were found.
3. Outliers:
 - Extreme values in the **Transaction Amount** column were analyzed but retained for insights into high-value transactions.

Summary of cleaned data:

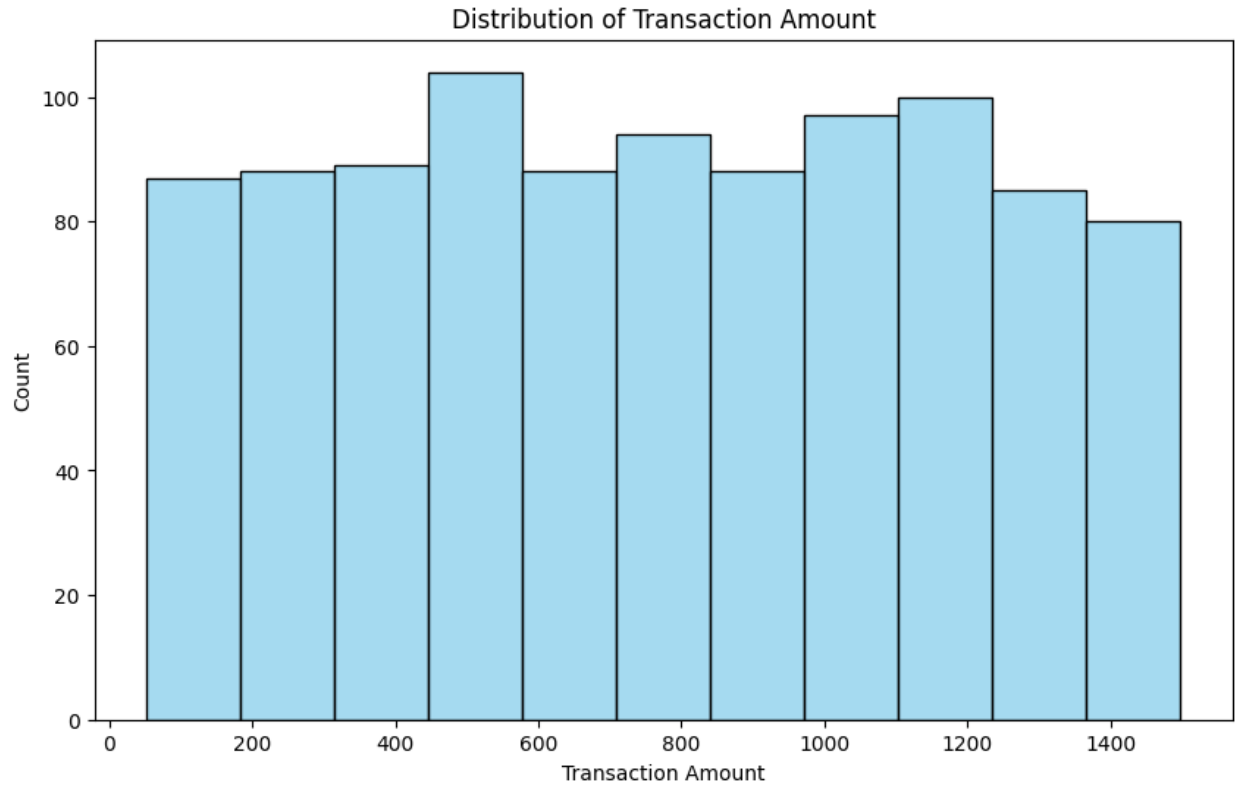
- Total Rows: 1,000
- Total Columns: 10
- No major change in any category of the data set was observed which is demonstrated by the box given below



3.Exploratory Data Analysis

3.1.Distribution of Transaction Amounts

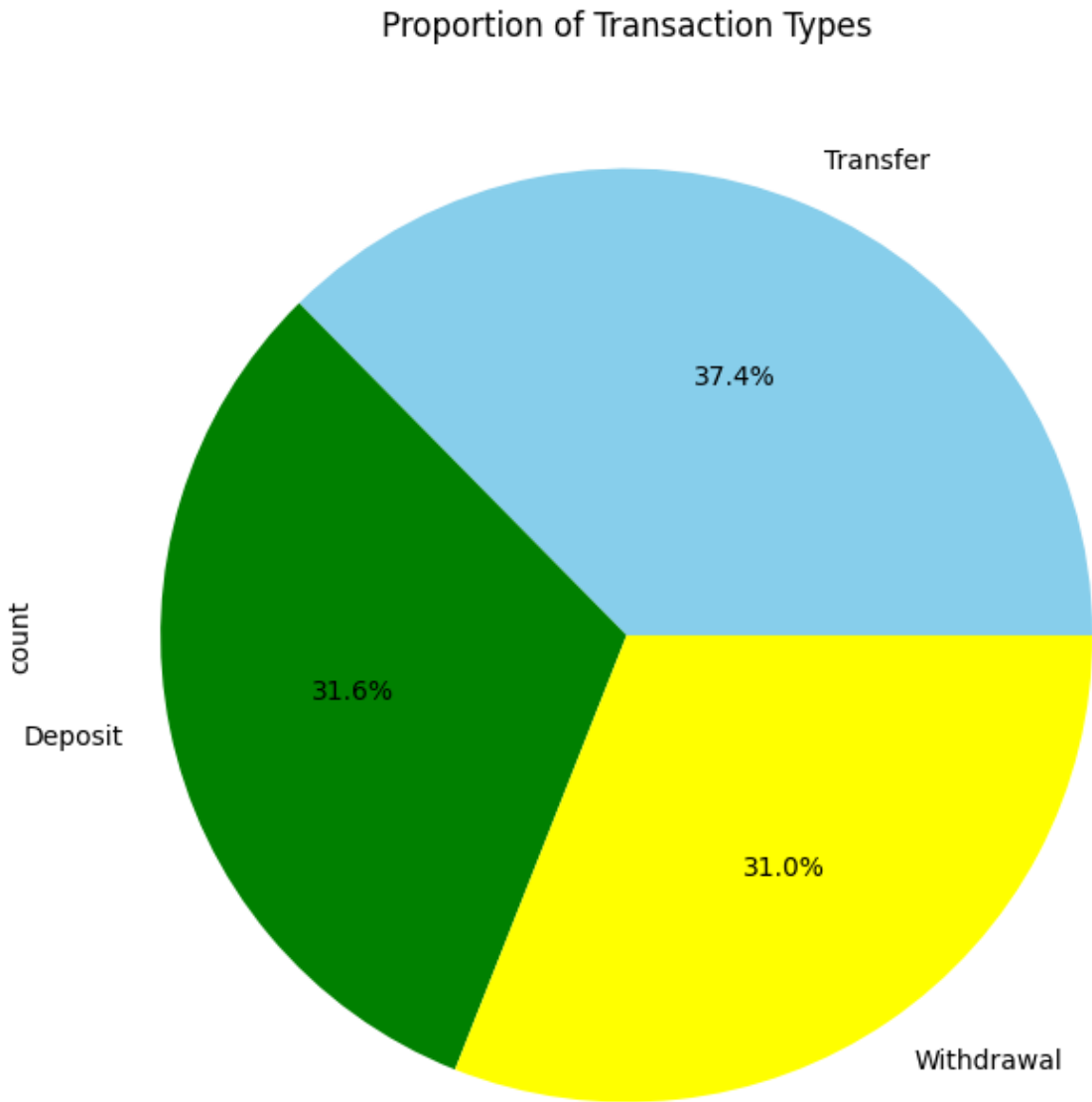
The histogram below shows the distribution of transaction amounts. Most transactions fall between 200 and 1,200 units. Outliers exist at both ends of the spectrum.



3.2 Proportion of Transaction Types

The bar chart below highlights the proportions of different transaction types:

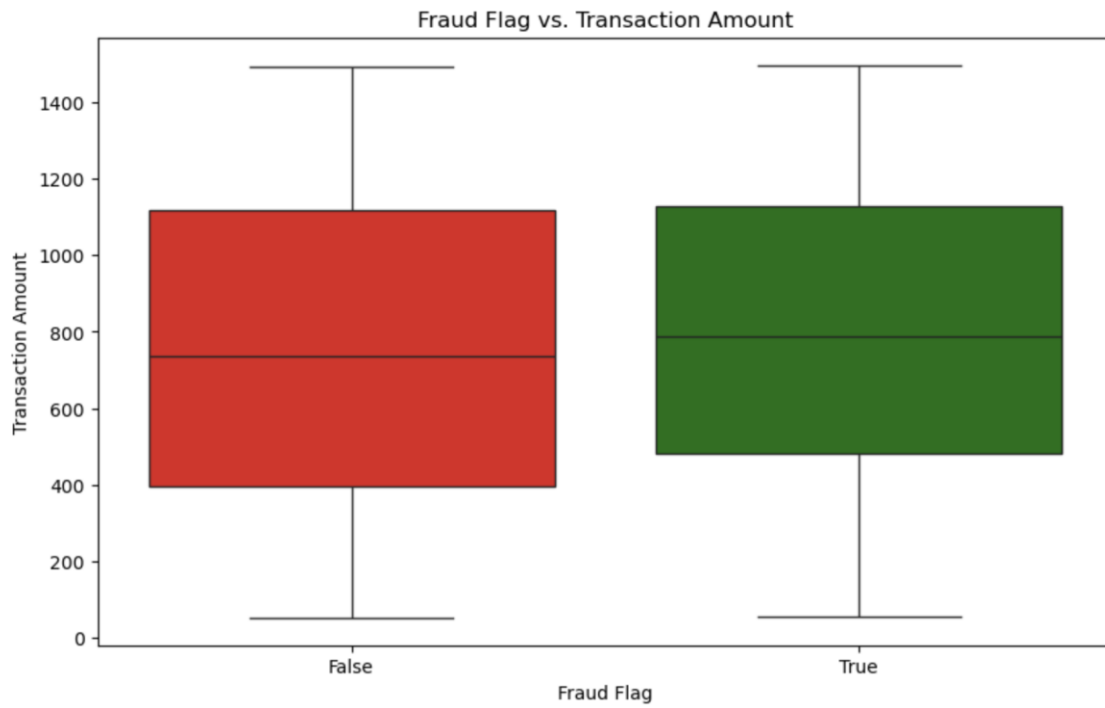
- Transfers dominate with 37.4% of transactions.
- Deposits and withdrawals are evenly distributed at around 31% each.

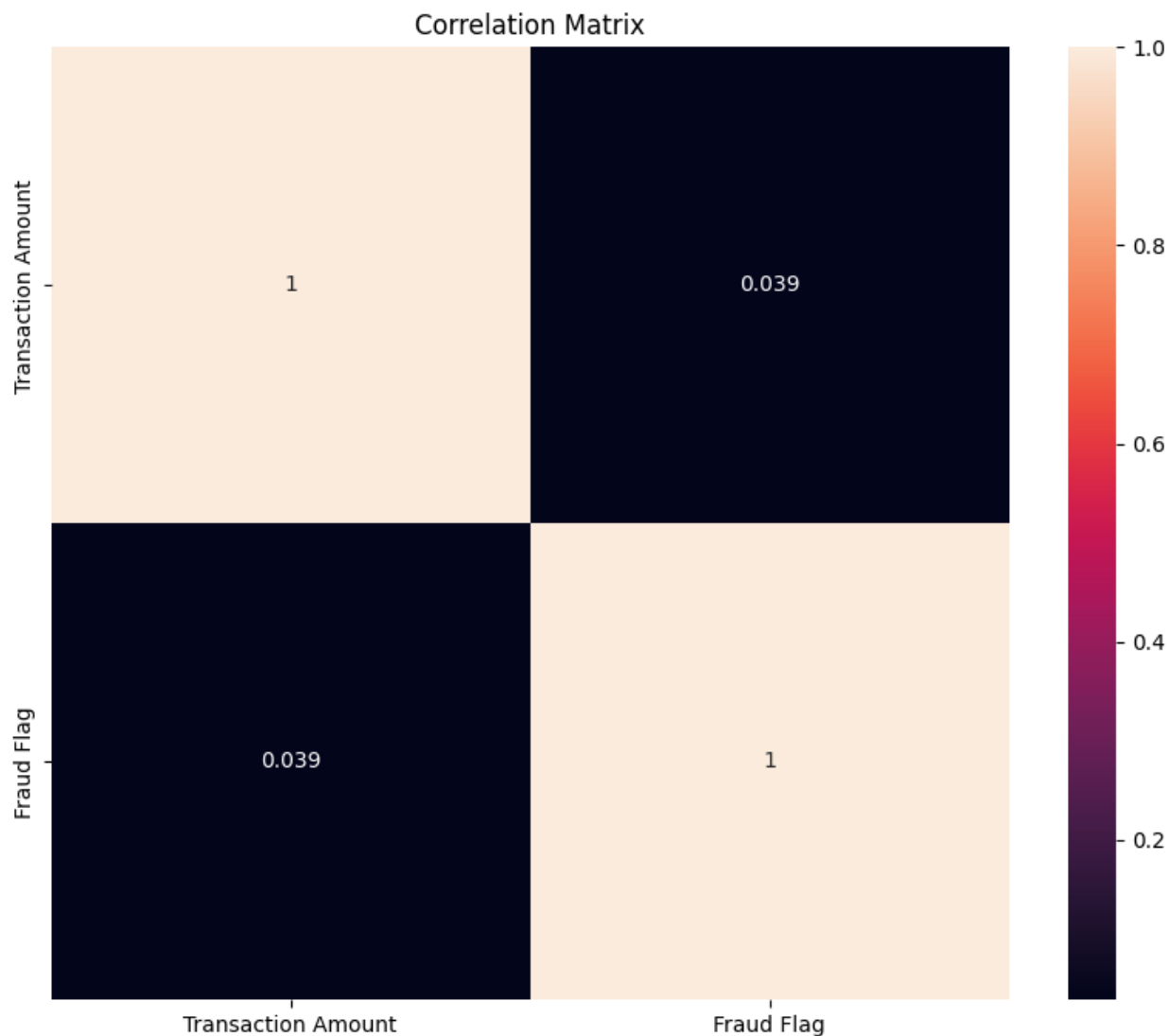


Approximately 48.10% of transactions are flagged as fraudulent.

3.3 Transaction Amount vs Fraud Flag:

The Heat Map and the box plot below demonstrates how Higher transaction amounts show a slightly higher likelihood of being flagged as fraudulent.

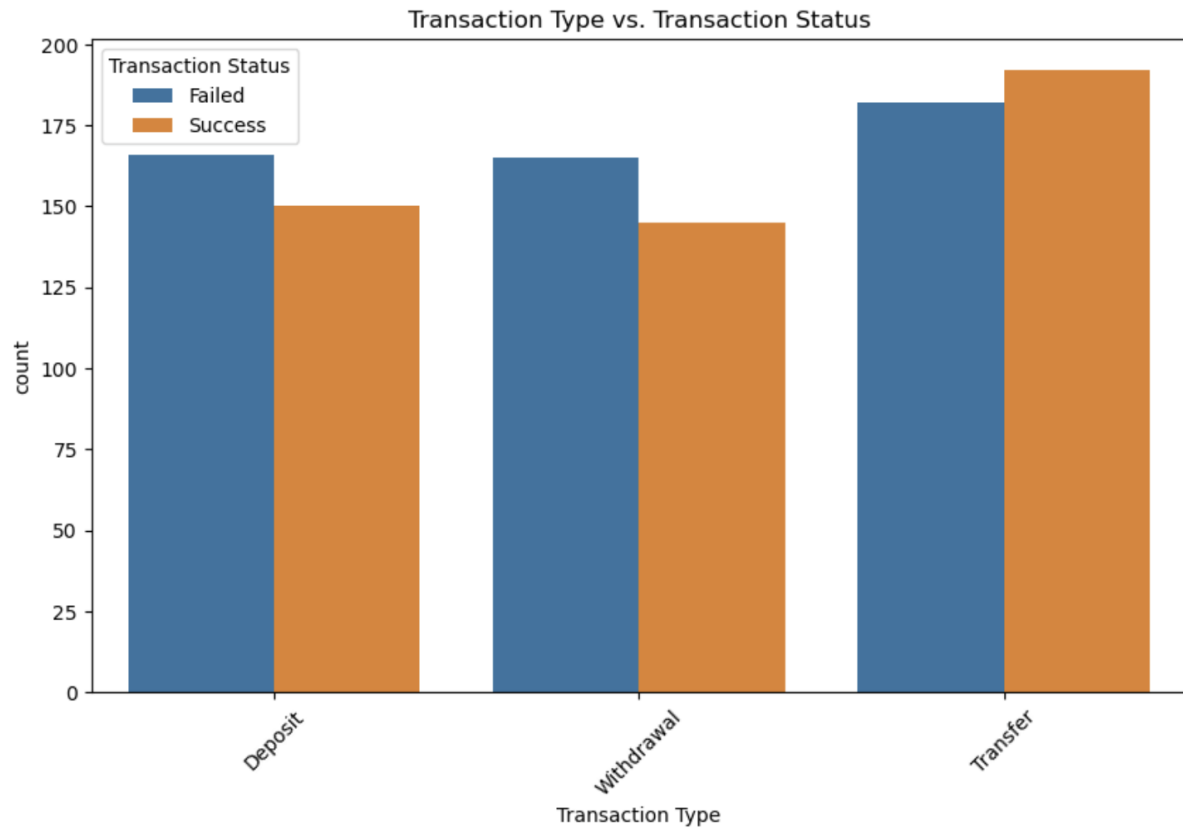




3.4 Comparison of Transaction Types by Status:

The graph given below demonstrates the distribution of transaction statuses (Failed and Success) across three types of transactions: Deposit, Withdrawal, and Transfer. Key insights include:

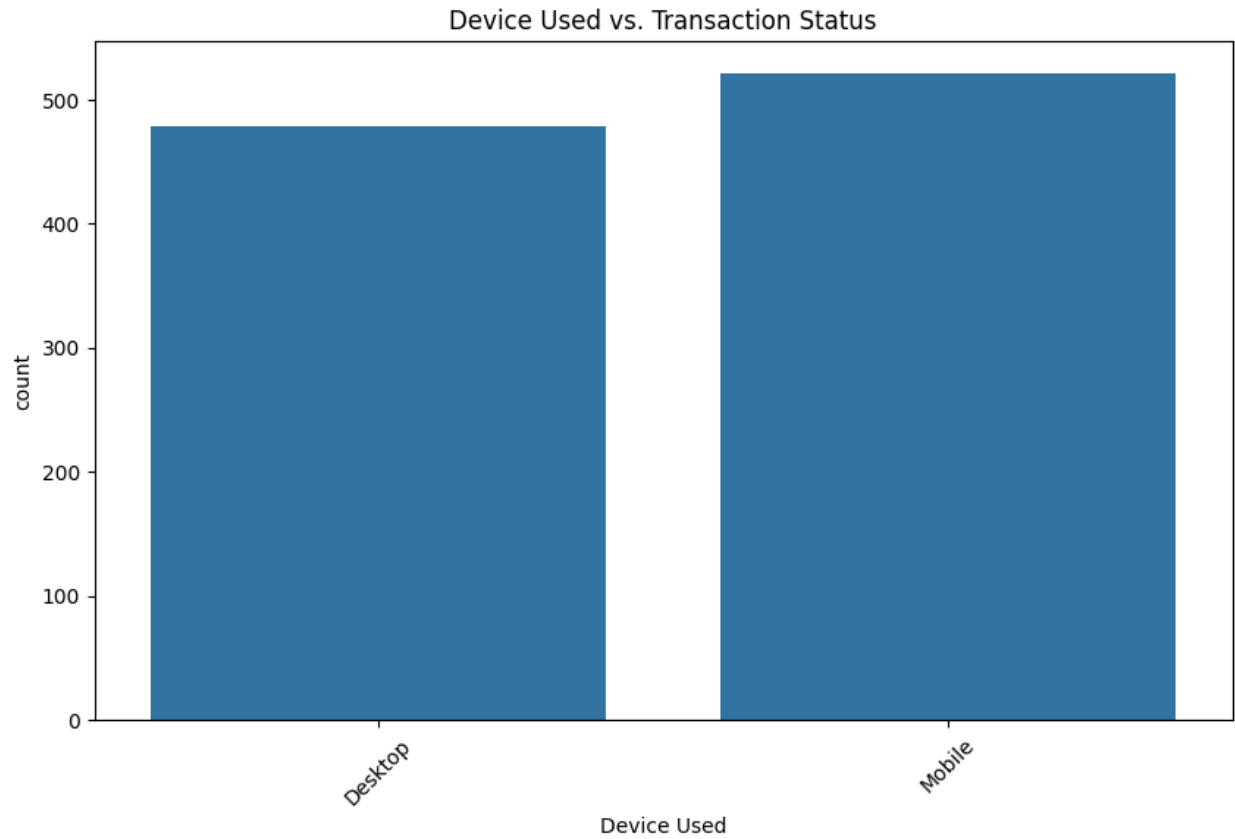
1. **Deposit:** The number of failed transactions slightly exceeds successful transactions, indicating potential issues with deposit processes.
2. **Withdrawal:** Similar to deposits, withdrawals have more failed transactions than successful ones, suggesting challenges in this transaction type.
3. **Transfer:** Transfers show a higher success rate compared to failures, making it the most reliable transaction type among the three.



3.5 Device Usage Analysis

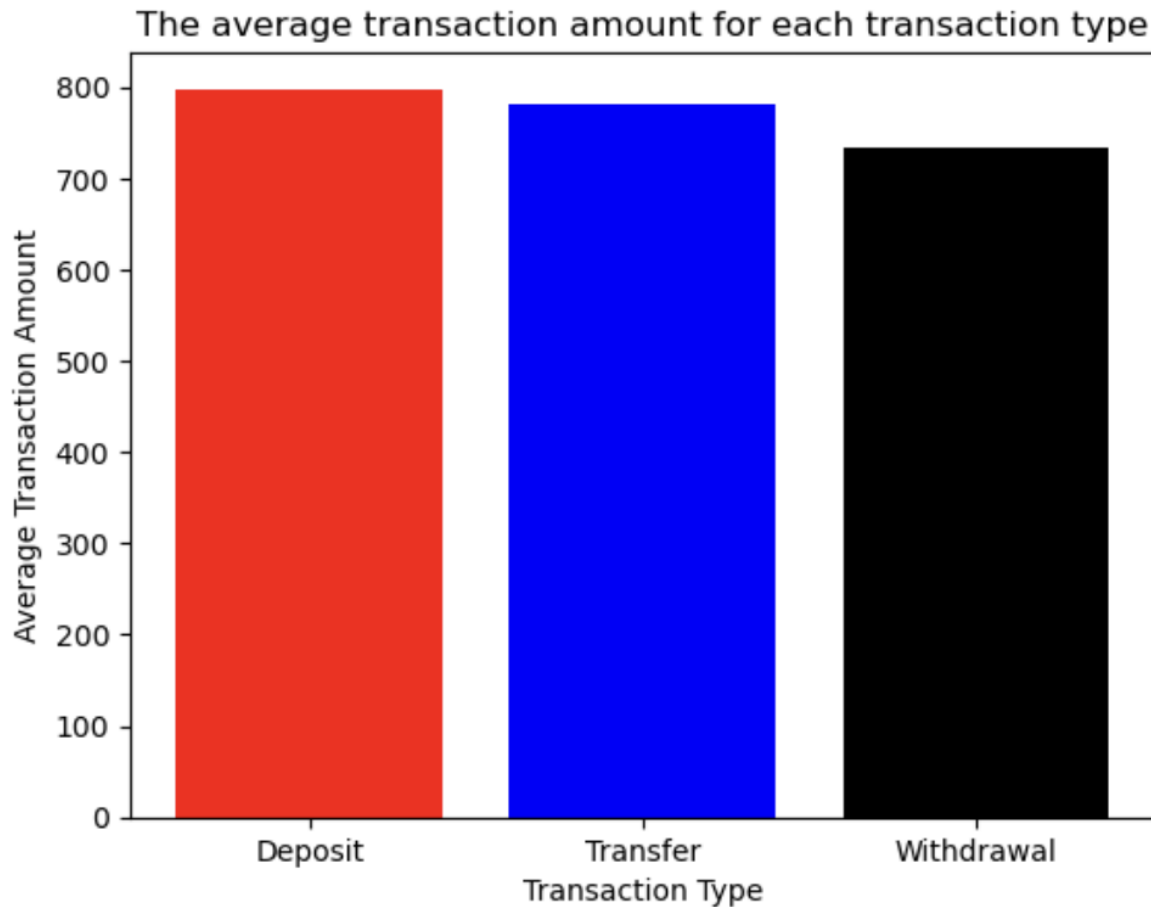
Mobile users account for a higher proportion of failed transactions compared to desktop users demonstrated in the given graph below.

Both device types have high and nearly equal transaction volumes, with **mobile usage just exceeding desktop** demonstrated in the graph given below.



4. Analyzing Feature Interactions

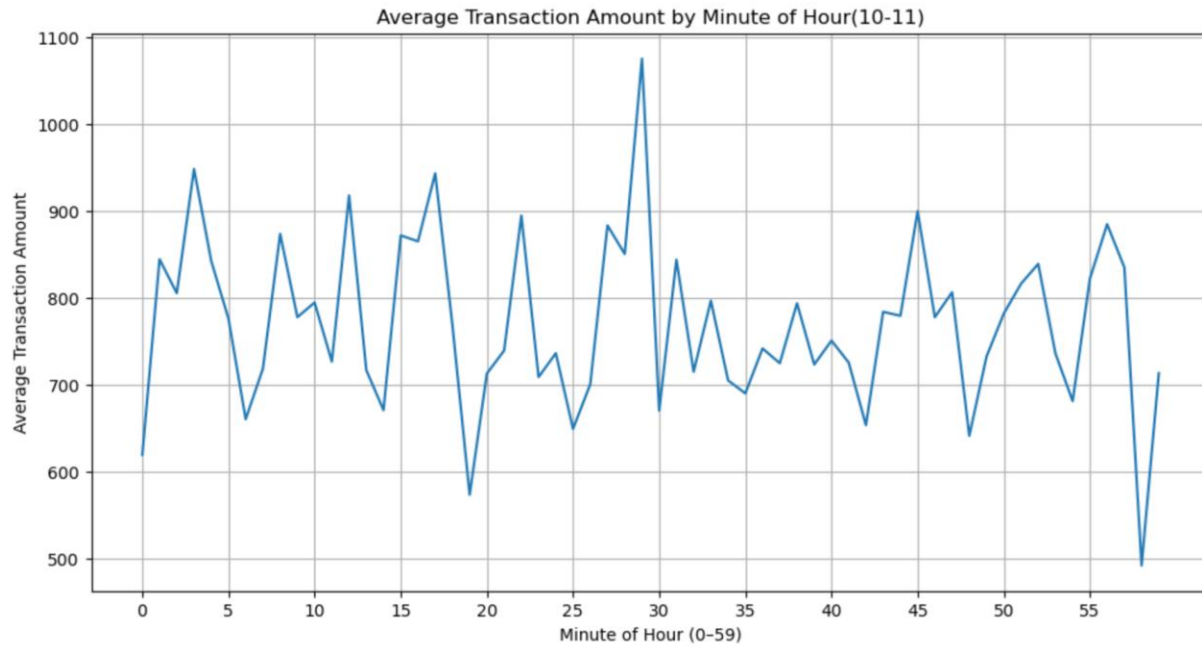
4.1. Analysis of The average transaction amount for each transaction type



- **Deposits** have the **highest average transaction amount**, indicating users typically deposit larger sums compared to other transaction types.
- **Transfers** come next with a slightly lower average than deposits, reflecting frequent but moderately sized peer-to-peer or account-to-account movements.
- **Withdrawals** show the **lowest average amount**, which could imply that users tend to withdraw money in smaller, more controlled increments—possibly for day-to-day needs.

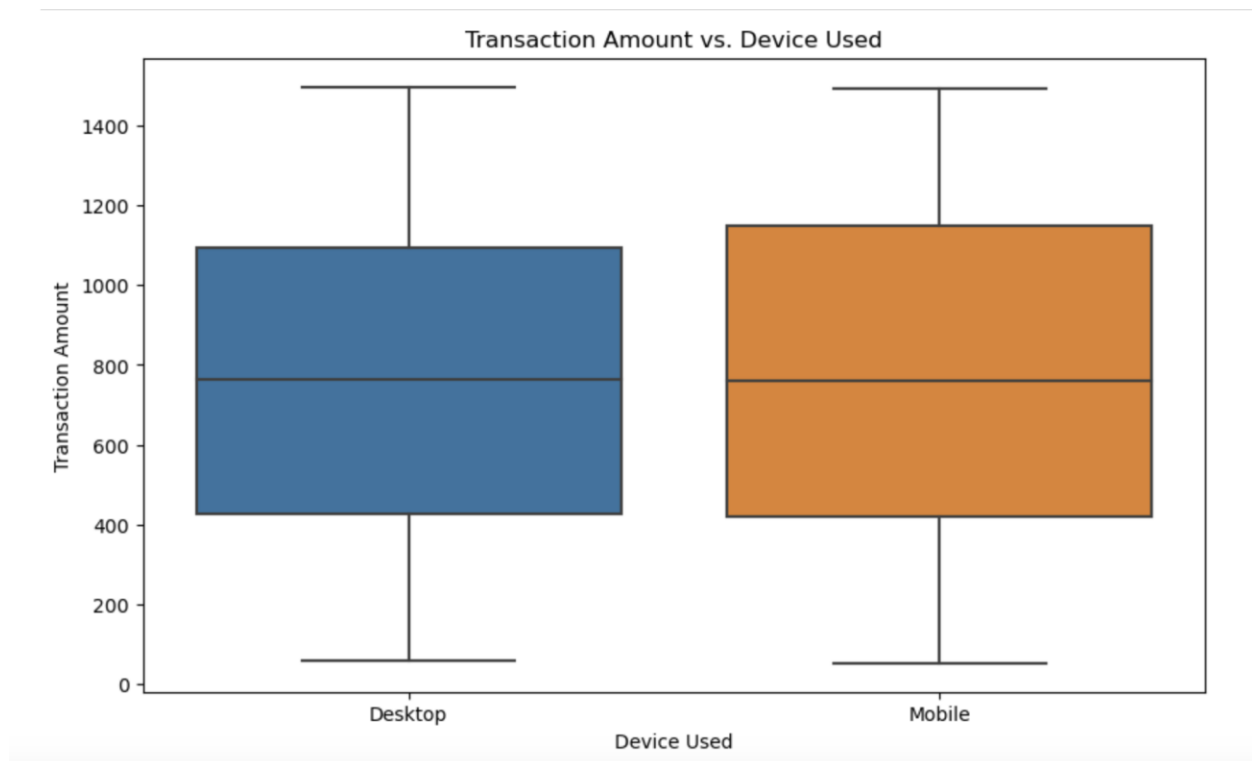
4.2. Analysis of transaction amounts by time of day

The values bounce between approximately **500 to 1100** units and around the time 10.30 the average transaction amount peaks close to 1100 .



4.3. Correlation between device used and the transaction amount:

There is **no significant difference** in the distribution of transaction amounts between **Desktop and Mobile users**. This suggests that **device type does not heavily influence the size of transactions** demonstrated in the graph given below



4.4. Analyze trends in fraud occurrences over time:

It shows a **sharp drop** in the number of fraudulent transactions from **472 at 10 AM** to only **9 at 11 AM**.

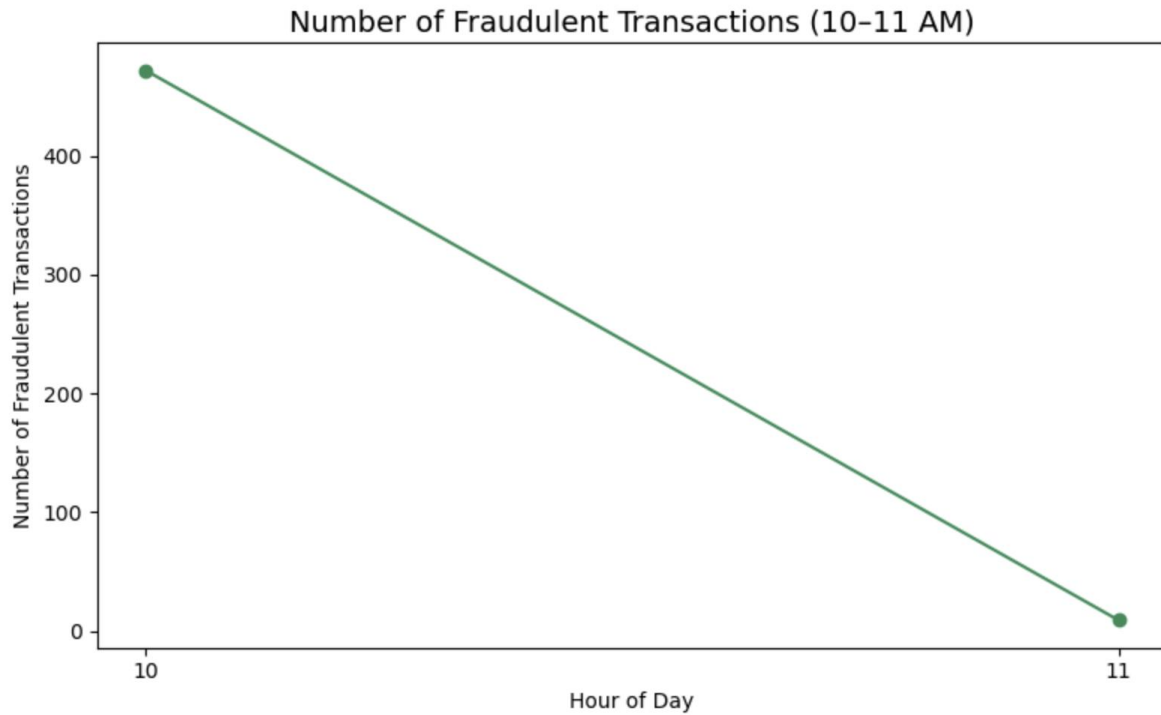
At **10 AM**, there were **472** fraudulent transactions.

At **11 AM**, the number fell sharply to just **9**.

This drastic decline may suggest that **fraudulent activities are heavily concentrated at a specific time window**, possibly due to scheduled attack patterns, system vulnerabilities, or timing strategies used by fraudsters.

The sharp contrast could also indicate that **preventive measures or monitoring protocols** may have kicked in around 11 AM, deterring further fraud.

But it is hard to conclude anything specifically because of very less data provided



4.5 Analyzing the average transaction amount for each sender and receiver account. Identify the top 10 sender and receiver accounts with the highest average transaction amounts.

Top 10 Senders by Average Transaction Amount:

Sender Account ID

ACC44804	1497.76
ACC67128	1495.01
ACC76875	1494.20
ACC14340	1494.03
ACC47545	1494.01
ACC76710	1493.86
ACC52161	1493.27
ACC75685	1488.93
ACC28678	1485.37

ACC16531 1485.22

Top 10 Receivers by Average Transaction Amount:

Receiver Account ID

ACC26269 1497.76

ACC77786 1495.01

ACC22724 1494.20

ACC33467 1494.03

ACC63108 1494.01

ACC62664 1493.86

ACC74234 1493.27

ACC92454 1488.93

ACC21230 1485.37

ACC41695 1485.22

- The top sender and receiver accounts all exhibit **average transaction amounts above 1485 units**, indicating that they frequently engage in high-value transactions.
- Notably, the top-ranked sender **ACC44804** and receiver **ACC26269** both share the exact highest average of **1497.76**, which might suggest a pattern or direct relationship in their transactions.
- These accounts may be key nodes in the transaction network and could warrant **further investigation** for financial behavior analysis, fraud detection, or priority customer profiling.

5.Summary

Key findings from the analysis include:

1. Transfers are the most frequent transaction type.
2. Fraudulent transactions account for 48.10% of all transactions.
3. Mobile devices are associated with higher failure rates compared to desktops.
4. Larger transaction amounts show a slightly higher likelihood of being flagged as fraudulent.
5. Withdrawals have a higher failure rate compared to other types.

6.Conclusion

This analysis provides actionable insights into transaction patterns that can help optimize fraud detection mechanisms and improve user experience across devices. Recommendations include enhancing security measures for high-value transfers and improving mobile transaction reliability.

BY-

Name: Rijan Poudel

Section: CA

Roll Number: 61

Registration Number : 240962035