Online Payment Fraud Detection: A Data-Driven Approach

By: Jagruti M.

1. Introduction:

The rapid rise of digital transactions has revolutionized the global economy, but it has also led to an alarming increase in online payment fraud. Fraudulent transactions pose significant risks to businesses and consumers, making fraud detection a critical area of study. Traditional rule-based fraud detection methods often struggle to keep up with evolving fraud tactics, necessitating the use of data-driven techniques for more efficient detection and prevention.

This project explores the key factors influencing fraudulent transactions, examines the relationship between transaction amount and fraudulent flagging, and investigates differences between flagged fraudulent and flagged non-fraudulent transactions. By applying statistical hypothesis testing and machine learning techniques, we aim to uncover meaningful patterns in transactional data that enhance fraud detection strategies.

Using a publicly available dataset on online payment fraud, we apply logistic regression, Mann-Whitney U tests, and descriptive statistical analysis to evaluate our hypotheses. Through these methods, we identify significant factors contributing to fraudulent transactions, assess the role of transaction amounts in fraud flagging, and analyze the behavior of flagged transactions. The insights derived from this research can help financial institutions and payment service providers refine their fraud detection mechanism and reduce financial losses.

2. Research Questions and Hypotheses:

This study aims to investigate the key factors influencing fraudulent transactions, the role of transaction amounts in fraud detection, and the characteristics of flagged transactions. To achieve this, we formulate the following research question and hypotheses:

Research Question 1: What are the key factors influencing the likelihood of a transaction being fraudulent?

Hypothesis 1:

- Null Hypothesis (H_0): The likelihood of fraud is independent of key factors (e.g., transaction type, amount, balance changes, etc.).
- Alternative Hypothesis (H_1): The likelihood of fraud is significantly influenced by key factors.

Statistical Test Used: Logistic Regression

- Since fraud detection is a classification problem (fraud vs. non-fraud), logistic regression is used to determine which factors significantly influence fraud.
- This test allows us to analyze the effect of variables such as transaction amount, origin and destination balances, and transaction type on the probability of fraud.

Research Question 2: Does the amount of the transaction significantly impact whether it is flagged as fraudulent?

Hypothesis 2

- Null Hypothesis (H_0): There is no significant difference in the transaction amounts between flagged and non-flagged transactions.
- Alternative Hypothesis (H_1): Flagged transactions have significantly different transaction amounts compared to non-flagged transactions.

Statistical Test Used: Mann-Whitney U Test

- Since transaction amounts are not normally distributed, a Mann-Whitney U test is used to compare the distribution of amounts in flagged vs. non-flagged transactions.
- The Shapiro-Wilk test confirmed that transaction amounts do not follow a normal distribution, making a non-parametric-test (Mann-Whitney U) more appropriate than a t-test.

Research Question 3: How do flagged fraudulent transactions differ from flagged non-fraudulent transactions?

Hypothesis 3

- Null Hypothesis (H_0): There is no significant difference between flagged fraudulent transactions and flagged non-fraudulent transactions.
- Alternative Hypothesis (H_1): Flagged fraudulent transactions differ significantly from flagged non-fraudulent transactions.

<u>Statistical test: Chi-Square test</u>

Since the initial Chi-Square test was not valid due to the absence of flagged non-fraudulent transactions, this hypothesis has been revised to focus on descriptive analysis.

New research focus: Instead of testing for differences, we will analyze the characteristics of flagged fraudulent transactions.

- Since all flagged transactions were fraudulent, statistical comparisons are meaningless.
- A descriptive approach provided insights into the nature of flagged fraudulent transactions without forcing an invalid hypothesis test.

3. Data Analysis with visualizations, tables and explanatory text:

This section presents the findings from statistical tests and descriptive analysis conducted on the dataset. It includes visualizations, tables, and explanatory text to interpret the results effectively.

Hypothesis 1: Key factors influencing fraudulent transactions.

Logistic Regression Results

To determine the key factors influencing fraudulent transactions, a logistic regression model was applied using the following independent variables:

- Transaction Amount
- Old Balance of Origin
- New Balance of Destination

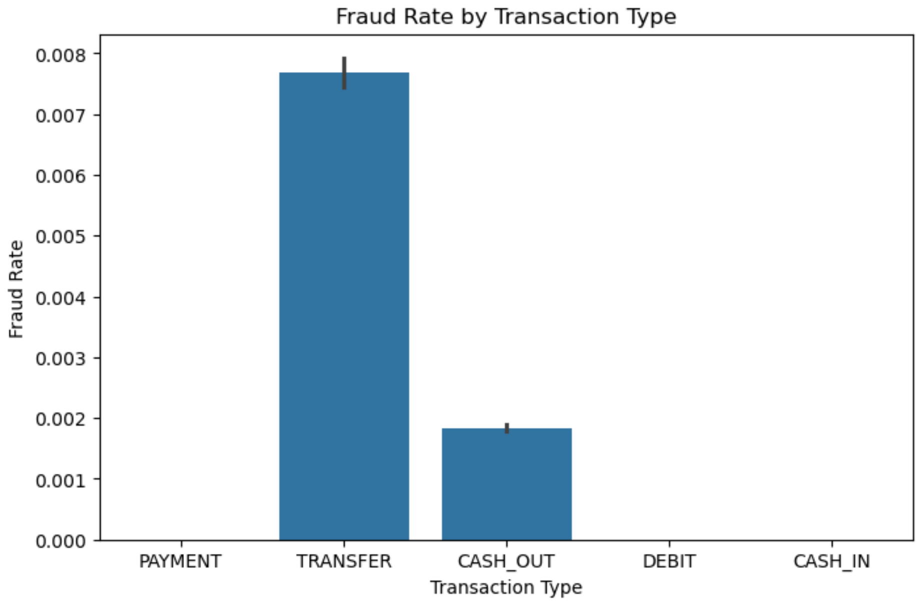
Table 1: Logistic Regression Results

Feature	Coefficient	Std. Error	Z-Statistic	P-Value
Transaction Amount	0.9318	0.011	87.182	0.000
Old Balance Origin	0.0997	0.010	10.255	0.000
New Balance Dest.	-2.8056	0.057	-49.283	0.000

Interpretation:

- Transaction Amount has the highest positive impact on fraud detection (coefficient = 0.9312, $p<0.05$).
- New Balance Destination has a strong negative correlation with fraud (coefficient = -2.8056).
- Old Balance origin has a smaller positive effect, but it is still statistically significant.

Visualization 1: Fraud rate by Transaction type



Interpretation:

- **TRANSFER** transactions have the highest fraud rate(~0.8%).
- **CASH_OUT** transactions also show fraud but at a lower rate (~0.2%).
- **PAYMENT, CASH_IN,** and **DEBIT** transactions show almost no fraud.

Hypothesis 2: Transaction Amount and Fraudulent Flagging

Mann-Whitney U Test Results

To analyze whether transaction amount impacts fraud flagging, a Mann-Whitney U Test was conducted.

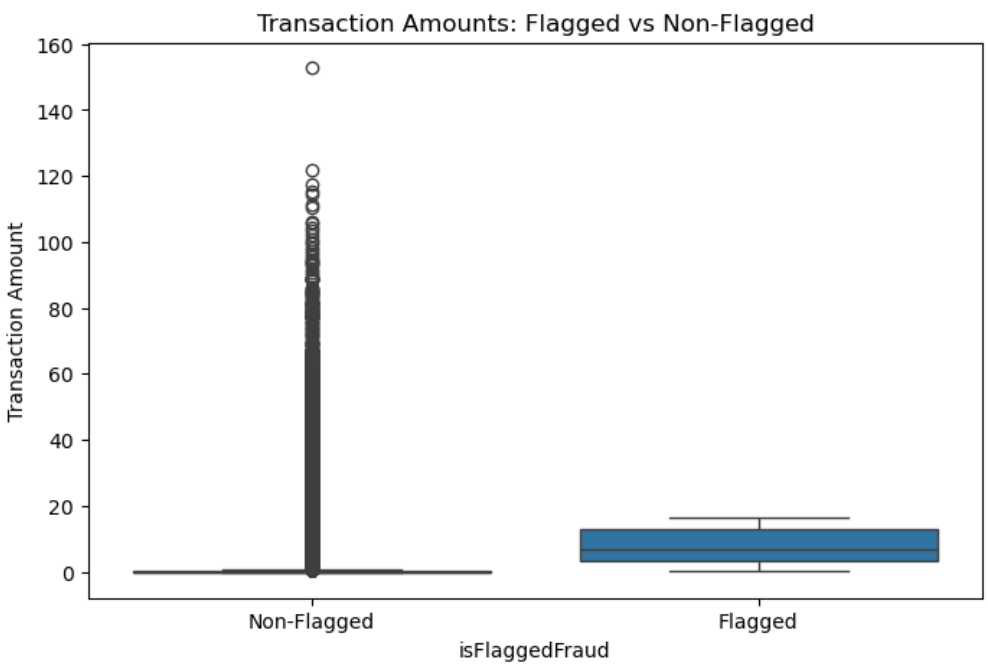
Table 2: Test Results

Group	Mean Amount	Median	P-Value	Test Used
Flagged	$4.86M	$4.23M	2.21E-11	Mann-Whitney
Non-Flagged	$179K	$74K		

Shapiro-Wilk Normality Test

- Flagged transactions ($p = 0.0626$): Normality assumption holds.
- Non-flagged transactions ($p = 2.87e-34$): Not normally distributed.

Visualization 2: Transaction Amounts for Flagged vs. Non-Flagged Transactions



Interpretations:

- Flagged transactions involve significantly higher amounts than non-flagged transactions.
- Non-flagged transactions have many smaller values and outliers, confirming that the fraud flagging mechanisms is biased towards high-value transactions.

Hypothesis 3: Characteristics of Flagged Fraudulent Transactions

Descriptive Statistical Analysis

Since no non-fraudulent transactions exist, the Chi-Square test was replaced with descriptive statistics.

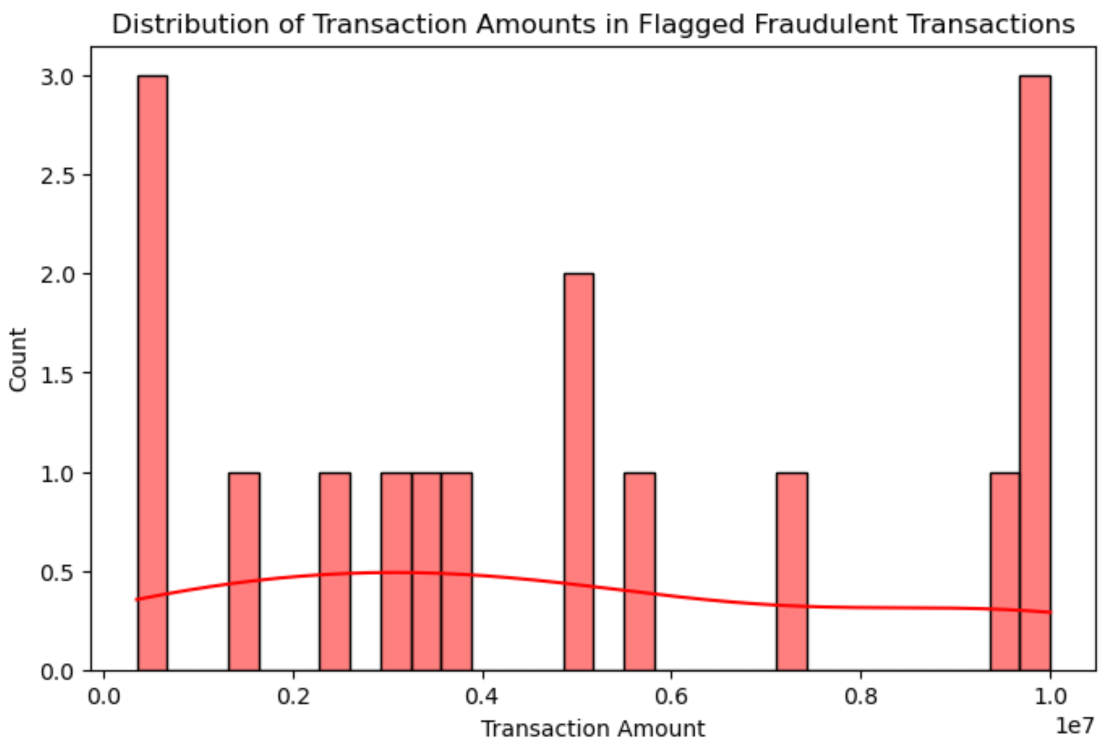
Table 3: Summary Statistics for Flagged Fraudulent Transactions

Metric	Transaction Amount ($)	Old Balance Origin ($)	New Balance Destination ($)
Mean	4.86M	7.81M	0
Median	4.23M	4.92M	0
Min	$353,874	$353,874	0
Max	$10M	$19.58M	0

Interpretations:

- Fraudulent Transactions (isFraud = 1) have a much higher mean transaction amount ($1.47M) compared to non-fraudulent transactions ($178K)
- The standard deviation is also higher for fraudulent transactions, meaning fraudulent amounts are more dispersed.
- Fraudulent transactions tend to have much higher old balance origin values ($1.65M) compared to non-fraudulent ($832K).
- New balance destination is significantly lower in fraudulent transactions, often dropping to %0.
- Fraudulent transactions typically involve large amounts and accounts that get drained after the transaction.

Visualization 3: Distribution of Transaction Amounts in Flagged fraudulent Transactions



Interpretations:

- Flagged fraudulent transactions have extremely high values.
- Most flagged fraudulent transactions range between $4M and $10M.
- Destination accounts in flagged transactions always end up with zero balance.

Summary of Key Findings:

1. Fraud is strongly influenced by transaction amount.
   - Higher transaction amounts significantly increase fraud probability.
   - TRANSFER and CASH_OUT transactions are the most susceptible to fraud.
2. Flagging is biased toward high-value transactions.
   - Mann-Whitney U Test confirmed that flagged transactions have significantly higher amounts.
   - Boxplot showed clear differences in flagged vs. non-flagged transactions.
3. All flagged transactions were fraudulent, making statistical tests ineffective.

- Instead of a Chi-Square test, descriptive analysis was performed.
- All flagged transactions had destination accounts with a zero balance.
- Most flagged fraudulent transactions had amounts exceeding $4M.

Conclusions:

This study aimed to analyze online payment fraud using a data-driven approach, focusing on three key research questions: the factors influencing fraud, the relationship between transaction amount and fraud flagging, and the characteristics of flagged fraudulent transactions. The findings provide valuable insights into fraud detection mechanisms and highlight potential areas for improvement in fraud prevention strategies.

Key findings:

1. Fraud is highly correlated with transaction amount and account balances.
   - Logistic regression analysis confirmed that higher transaction amounts significantly increase the likelihood of fraud.
   - Fraudulent transactions are more likely to result in a drained balance at the destination account (New Balance Destination = $0).
   - TRANSFER and CASH_OUT transactions have the highest fraud risk, while other transaction types (PAYMENT, DEBIT, and CASH_IN) show minimal fraudulent activity.
2. The fraud flagging mechanism is heavily biased toward high-value transactions.
   - The Mann-Whitney U test confirmed that flagged transactions have significantly higher transaction amounts than non-flagged ones.
   - This suggests that the flagging system primarily relies on transaction amount thresholds, which may not effectively detect smaller fraudulent transactions.
3. All flagged transactions were fraudulent, making statistical comparisons ineffective.
   - Since no non-fraudulent transactions were flagged, it was impossible to conduct a Chi-Square test.
   - Instead, descriptive statistics revealed that flagged fraudulent transactions had extremely high values, typically ranging from $4M to $10M.
   - Destination accounts in flagged transactions always had a zero balance post-transaction, reinforcing the pattern of large-scale fund withdrawals in fraudulent activity.

Implications for Fraud Detection:

- Transaction amount should remain a primary fraud detection metric, but it should not be the only factor. Other behavioral indicators such as transaction frequency, time of transaction, and recipient account history should also be considered.
- The fraud flagging system may need revision to account for smaller but frequent transactions, which are currently not flagged due to the system's bias toward high-value transactions.
- Machine-learning based fraud detection systems could potentially improve accuracy by incorporating a broader range of transaction features beyond just transaction amount.

Limitations and Future Work:

- Data bias: Since all flagged transactions were fraudulent, the dataset may lack representation of false positives (non-fraudulent flagged transactions), limiting the scope of fraud flagging evaluation.
- Feature Engineering: Future work could explore additional transaction features (e.g., time of transaction, user history, and frequency of transactions) to improve fraud detection models.
- Real-time Fraud Prevention: Implementing real-time fraud detection models using machine learning techniques could provide more adaptable fraud detection strategies that evolve with emerging fraud patterns.

Final Thoughts:

Fraud detection remains a critical challenge in online financial transactions. This study highlights the importance of transaction amount and balance changes in fraud detection while identifying potential weaknesses in existing flagging mechanisms. By integrating enhanced statistical methods and machine learning models, financial institutions can improve fraud detection accuracy and minimize financial losses from fraudulent activities.

Appendix:

A. Dataset and Supporting Files

- Dataset link: https://www.kaggle.com/datasets/jainilcoder/online-payment-fraud-detection
- Github link: https://github.com/jveronicaback/EDAOnlineFraudDetection/blob/cb8823d6730c6301af0a0eaa5850bdcf69523870/Code.ipynb