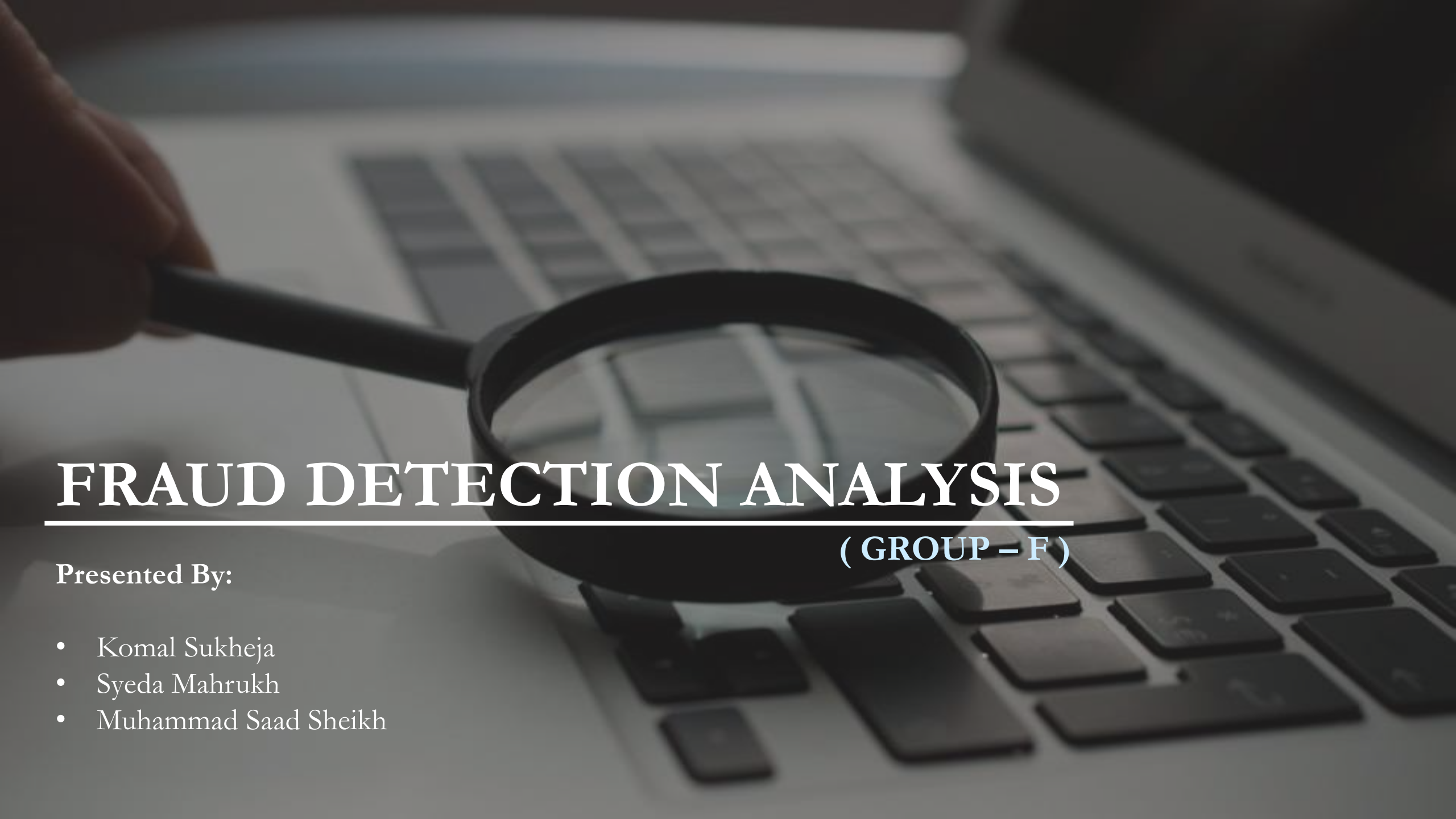


FRAUD DETECTION ANALYSIS

(GROUP – F)

Presented By:

- Komal Sukheja
- Syeda Mahrukh
- Muhammad Saad Sheikh

OVERVIEW OF DATA SET

The dataset provides a comprehensive collection of financial transaction records, allowing us to examine and understand the occurrences of fraudulent activities. By analyzing this data, we can uncover patterns, anomalies, and trends that indicate potential fraud attempts, ultimately contributing to the development of effective fraud detection methods and strategies..

STAKEHOLDER

Fraud Investigation Manager

We are presenting insights from the fraud detection dataset to the organization's Fraud Investigation Manager, who is tasked with identifying and addressing fraudulent activities. Our goal is to furnish them with actionable information to enhance their detection and investigation procedures.

PROBLEM STATEMENT

Develop a robust fraud detection system for the provided dataset, where the current fraud detection rate stands at 4.5%. The objective is to significantly enhance the accuracy of identifying fraudulent transactions while minimizing false positives.

KEY ATTRIBUTES & SIGNIFICANCE

- **Customers_data:** Includes customer profiles with information such as name, age, address, and contact details.
- **Account_activity**: Provides details of customer account activity, including account balance, transaction history, and account status.
- **Merchant_data**: Contains information about merchants involved in transactions.
- **Transaction_category_labels**: Provides category labels for different transaction types.
- **Transaction_records**: Contains transaction records with details such as transaction ID, date, amount, and customer ID.
- **Transaction_metadata**.: Contains additional metadata for each transaction.
- **Fraud_indicators** : Contains indicators of fraudulent patterns and suspicious activities.
- **Suspicious_activity**: Provides specific details of transactions flagged as suspicious.
- **Amount_data**: Includes transaction amounts for each transaction.
- **Anomaly_scores**: Provides anomaly scores for transaction amounts, indicating potential fraudulence.

OBJECTIVES

1. Identify Fraud Patterns:

Analyze the provided data to identify recurring patterns associated with fraudulent transactions.

2. Customer Analysis:

Analyze customers based on their transaction behavior, characteristics, and historical data.

3. Identify correlations:

Cross-reference Fraud indicators with other available data, such as customer profiles and transaction amounts, to identify correlations and build a more comprehensive understanding of potential fraud markers.

4. Age-based Analysis:

Investigating if certain age groups are more susceptible to fraudulent transactions and understanding the factors contributing to this trend.

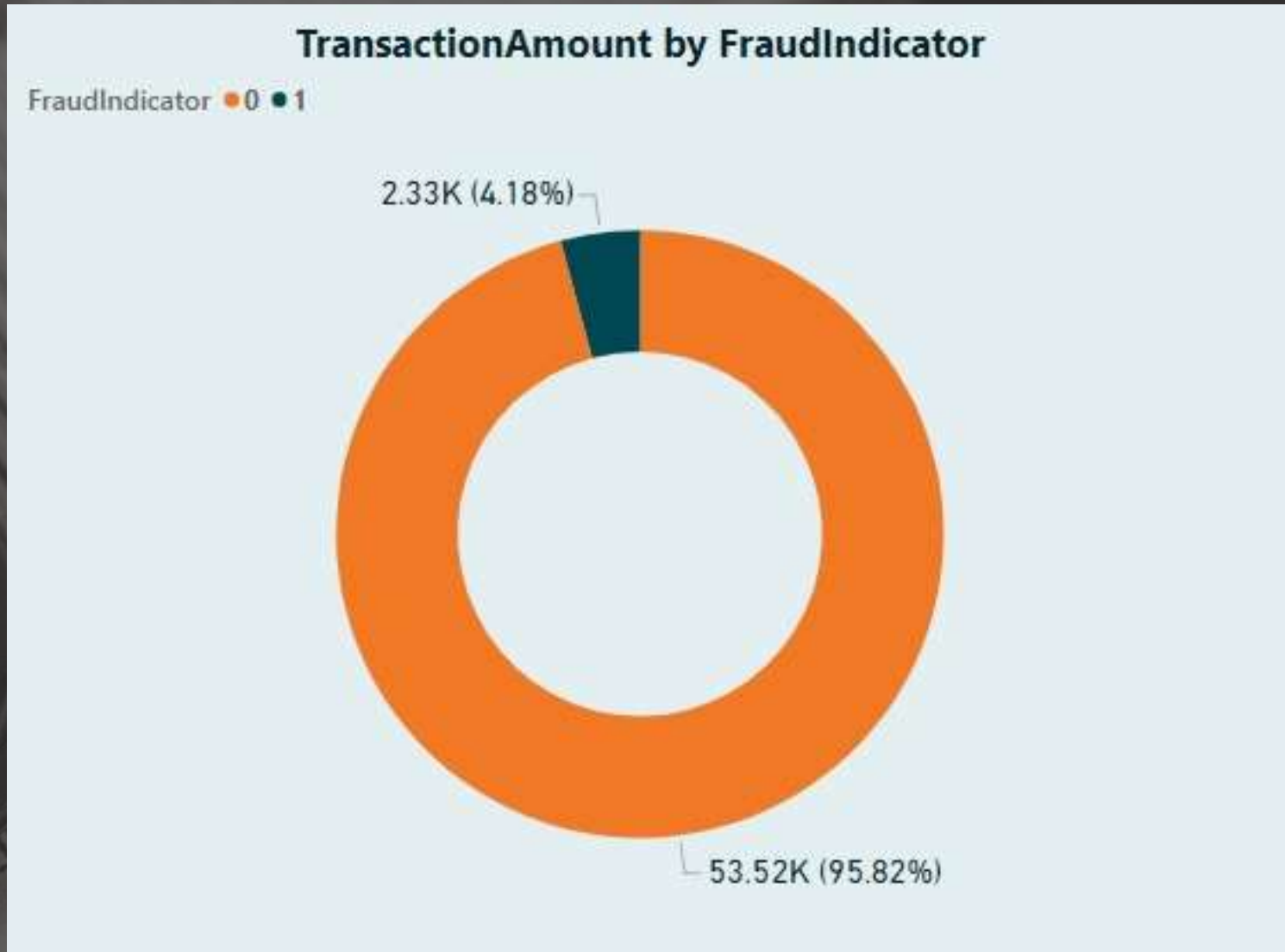
5. Temporal Trends:

Analyzing transaction patterns over time to uncover seasonal trends that could suggest coordinated fraudulent activities.

RESULTS

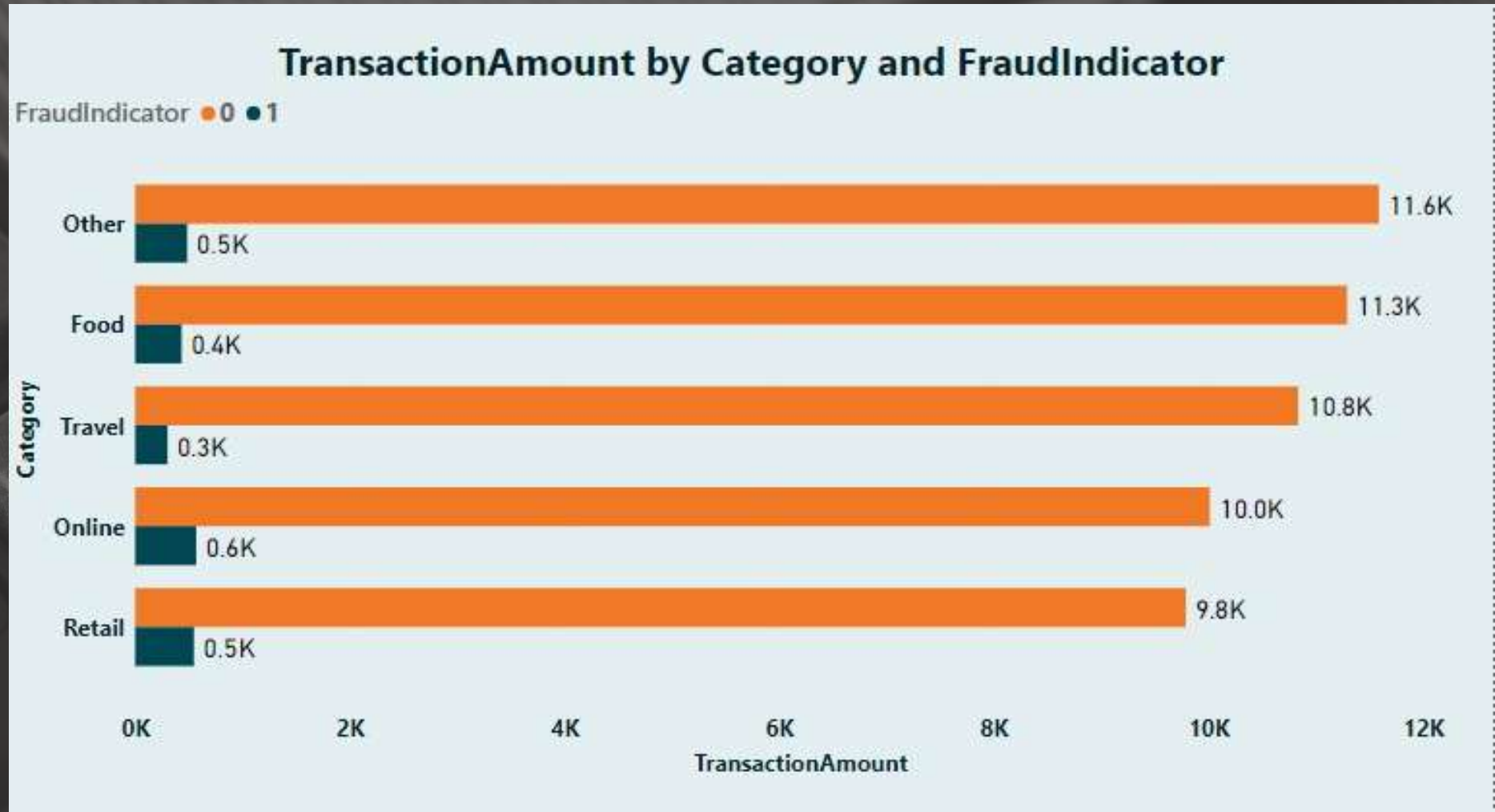
Transaction Amount By Fraud Indicator

- Fraudulent transactions makes up about 4.13% of total data set.
- The majority of transaction appears to be legitimate which is 95.82%.



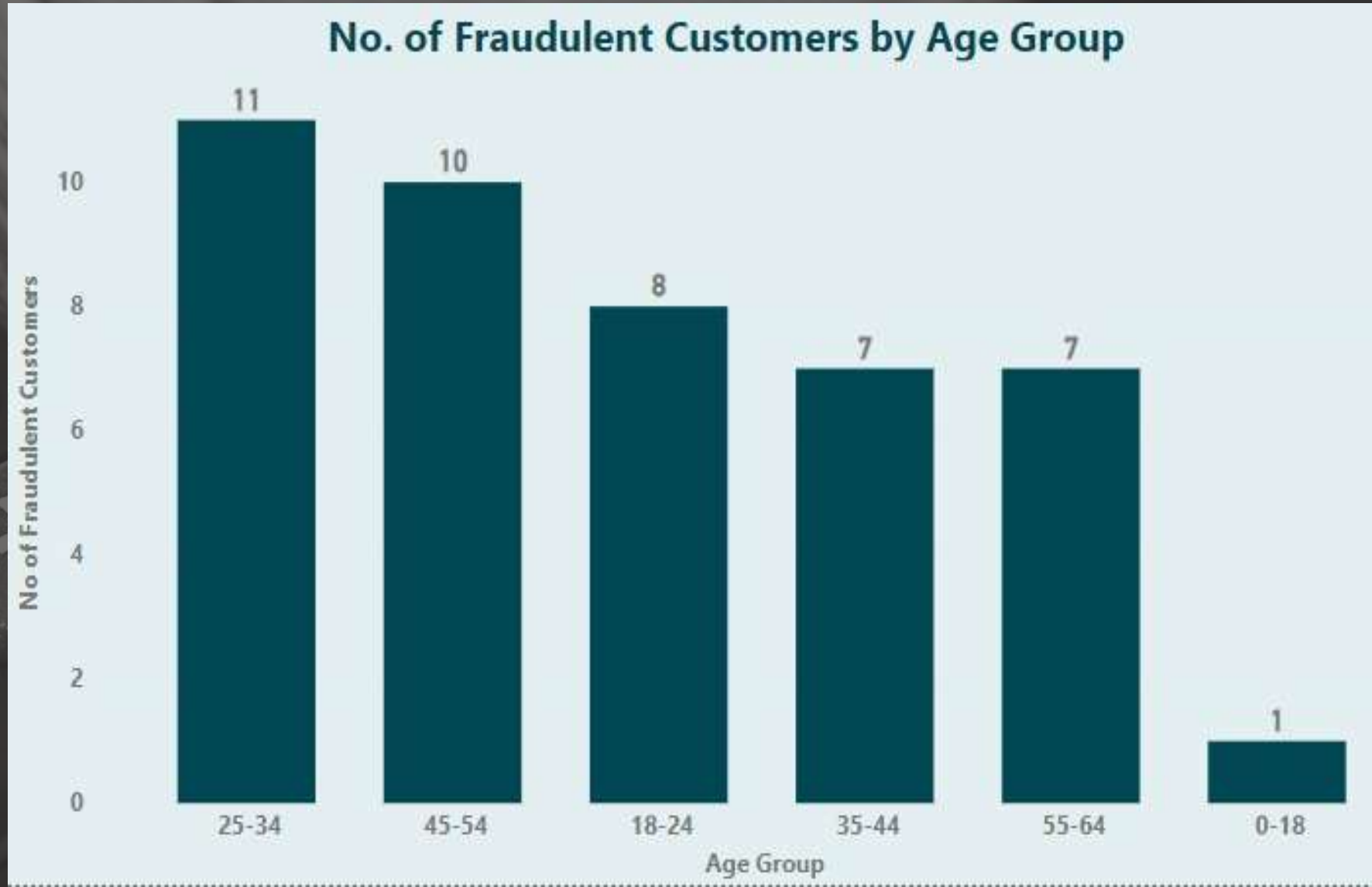
Transaction Amount Analysis

- Online and Retail has the highest Fraud transactions
- Travel category has the least amount of transactions



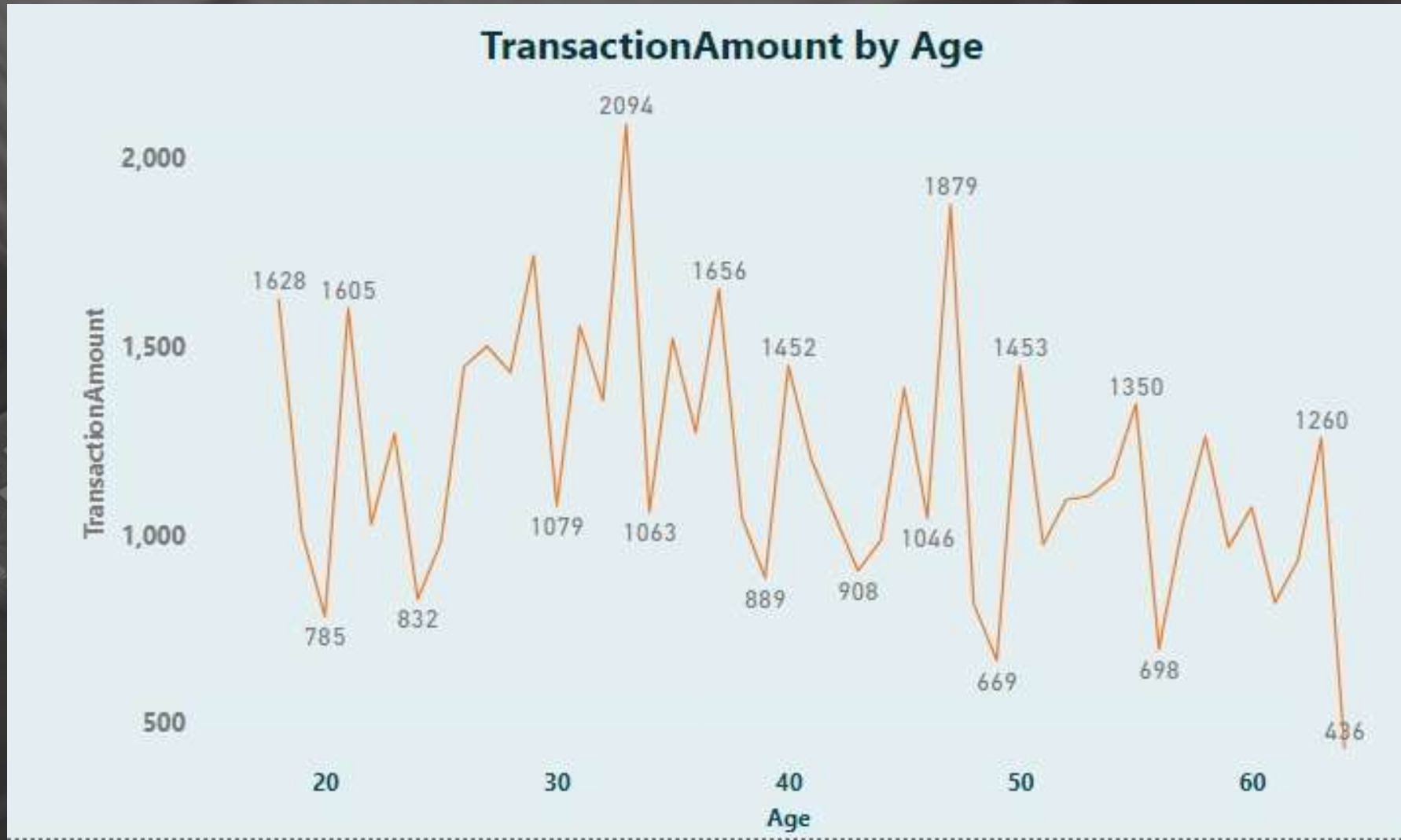
Age Group Analysis

- The age group 25-34 are more likely to be involved in fraudulent activities. .
- One fraud customer from 0-18 range.



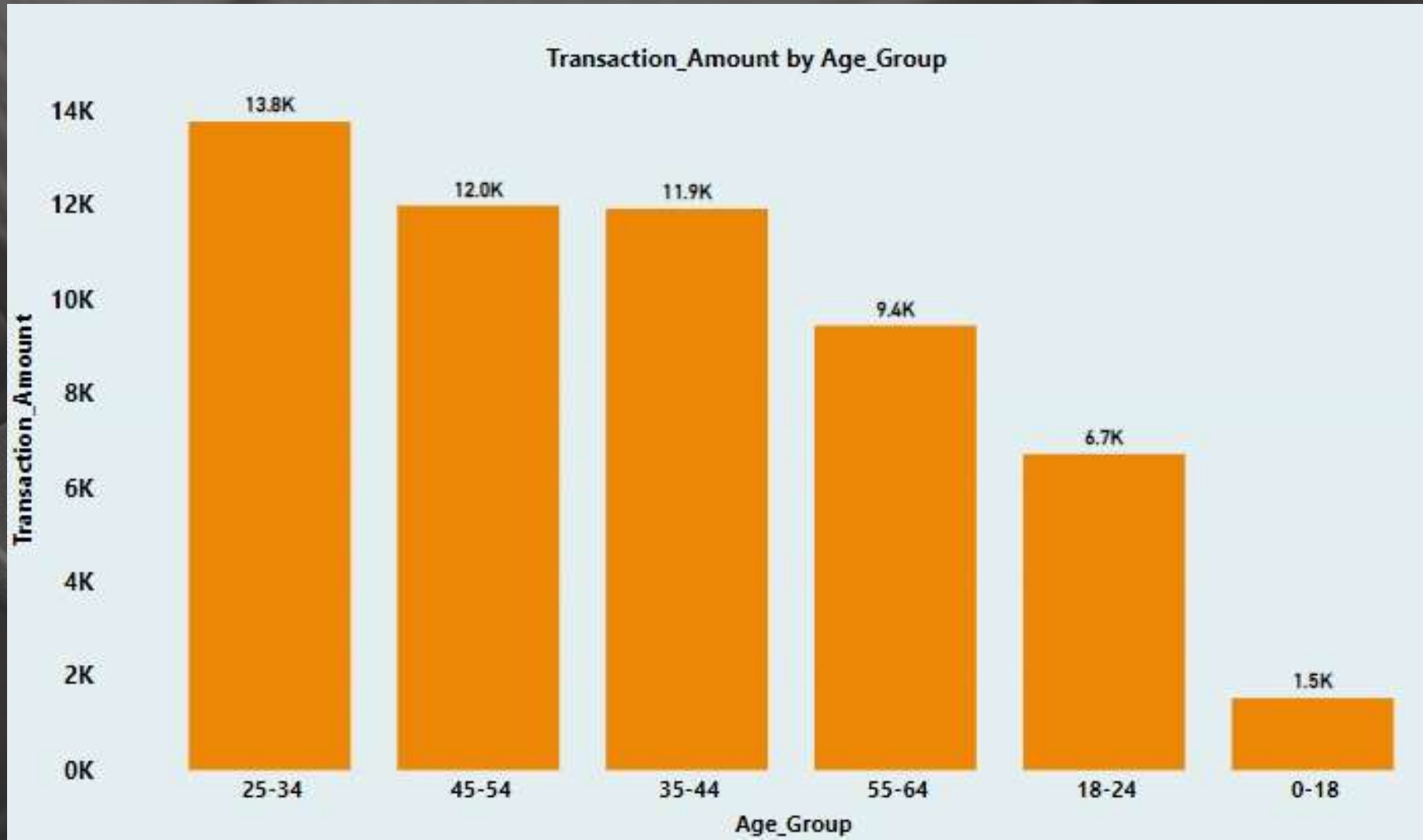
Transaction Behavior Analysis by Age

- No significant correlation between a customer's age and the transaction amount.



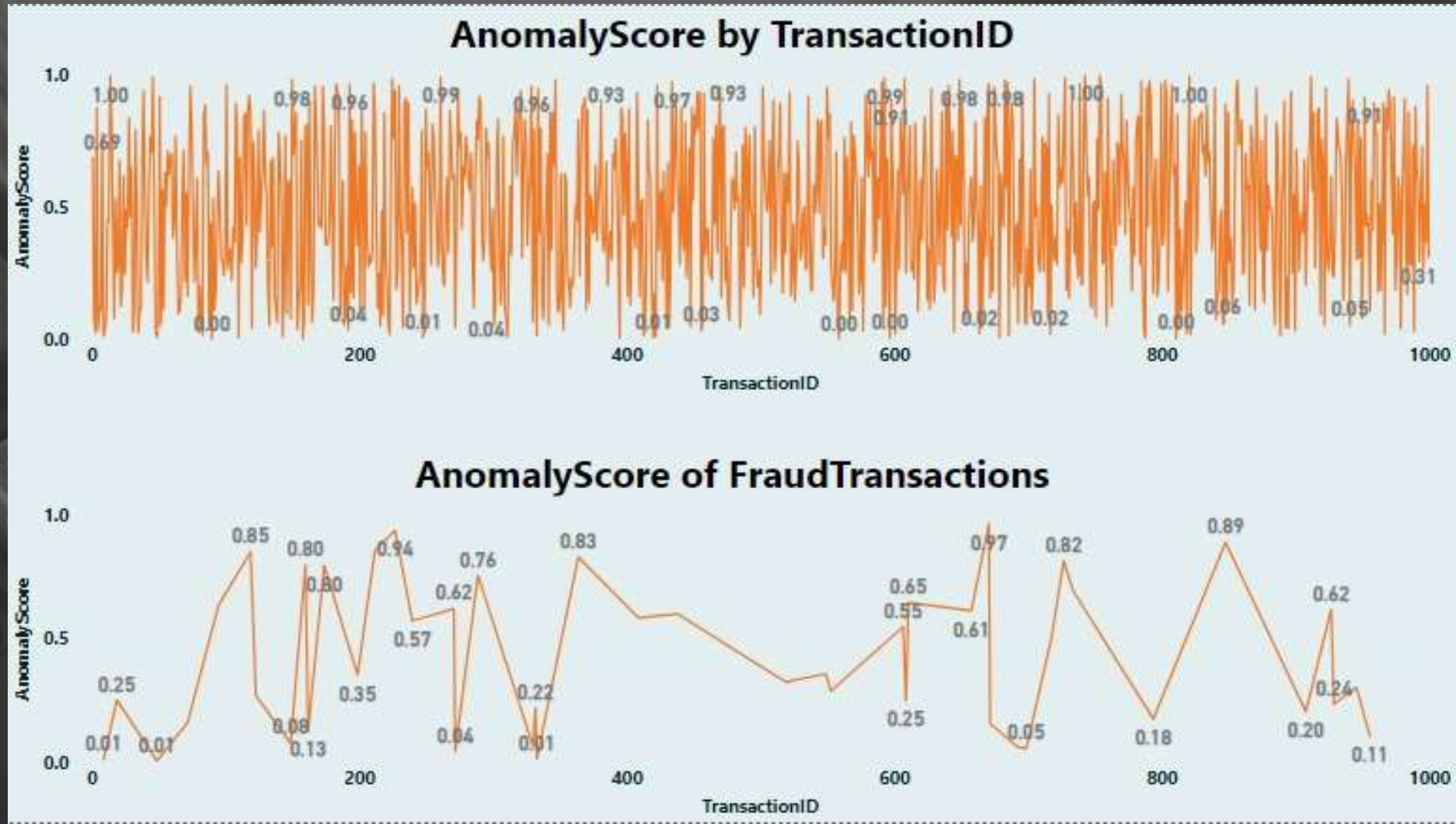
Transaction Behavior Analysis by Age Group

- Age Group 25-34 has the highest amount of transactions
- Age Group 0-18 has the lowest amount of transactions



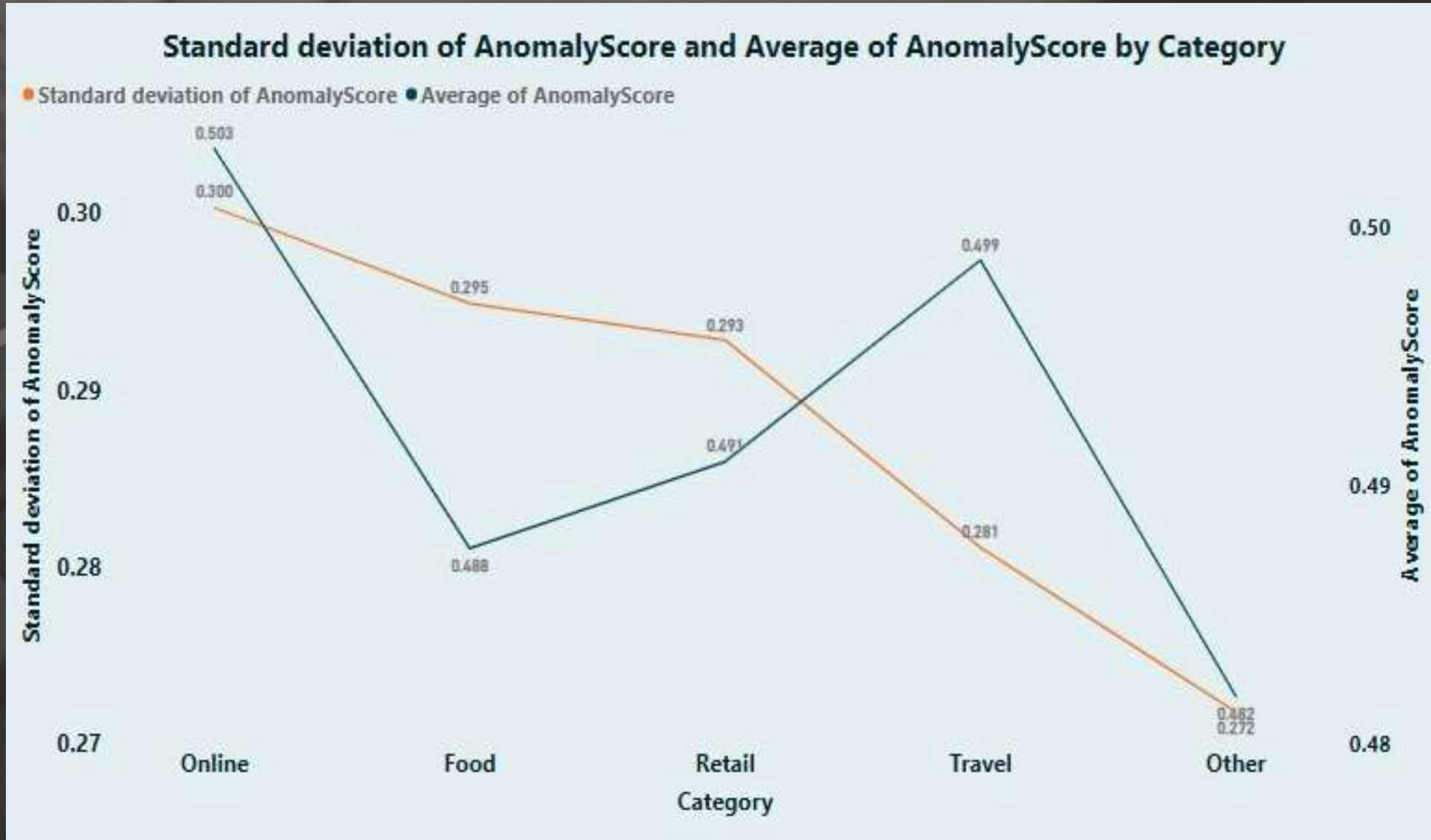
Anomaly scores Analysis

- In the dataset, we lack information about the pattern of the anomaly score and how it is calculated.



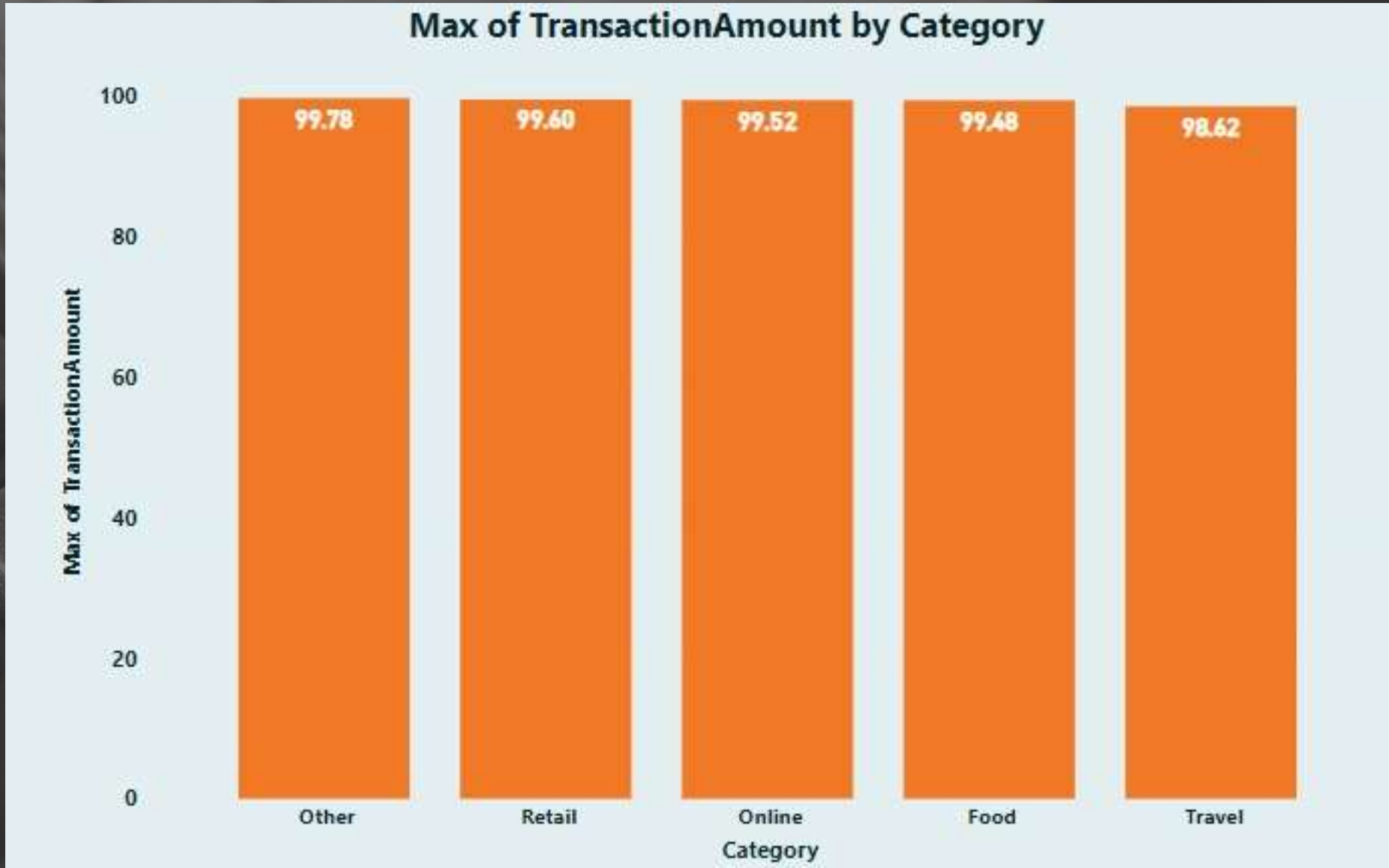
Anomaly scores Analysis

- Online and travel categories have the highest average anomaly scores
- Online category has the highest standard deviation anomaly scores.



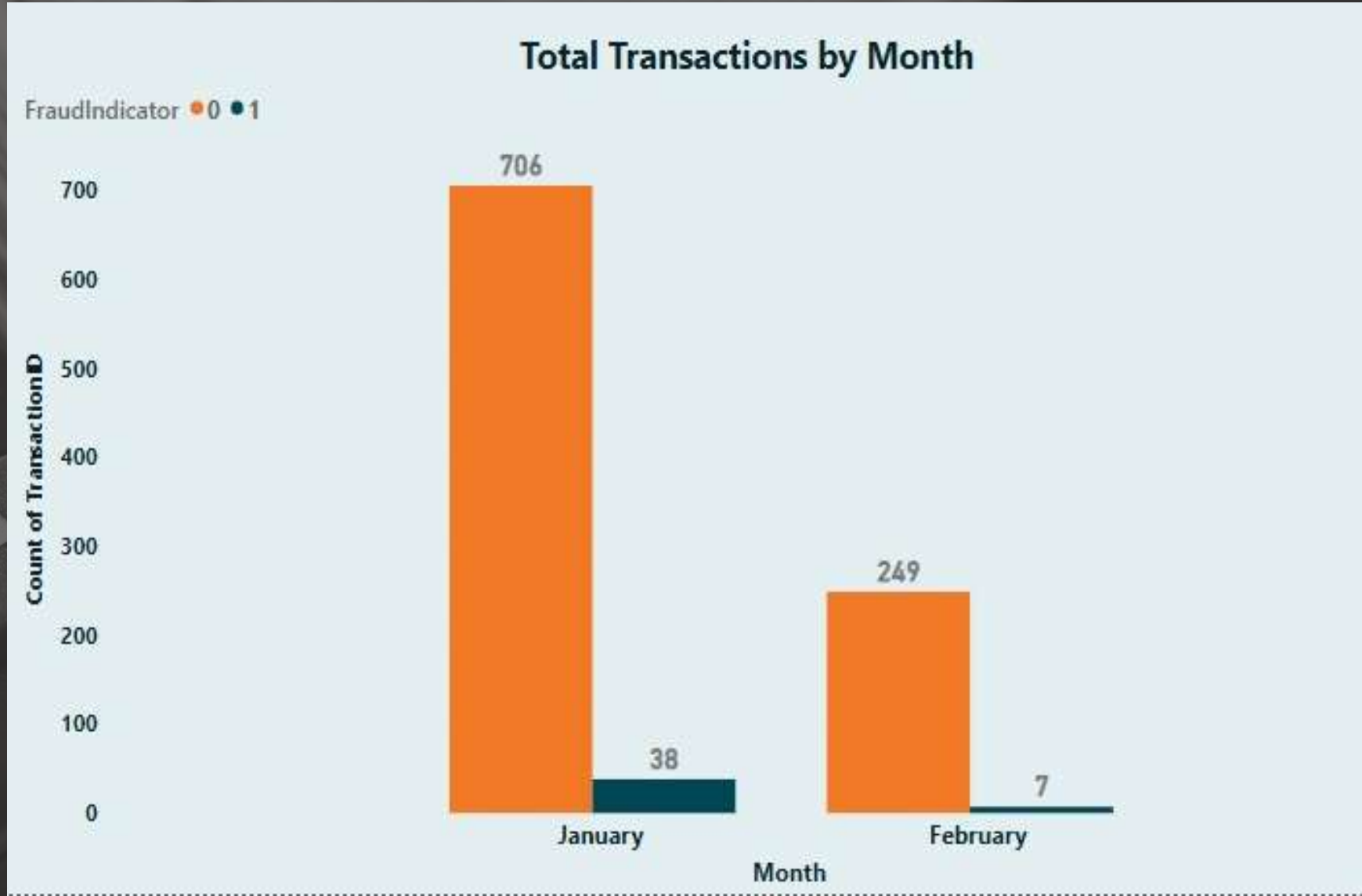
Highest Transaction Amounts

- No significant difference between each category.
- Retail and Online categories have Highest transaction amount of 99.60 and 99.52
- Travel has least 98.62 transaction amount.



Total Transaction By Month

- 38 Frauds in January and 7 in February
- Having data for only two months might restrict the ability to capture the complete range of fraudulent behaviors and patterns.



Fraud Transaction by Day

- Notable spikes in fraud occurrences are observed on Monday, Wednesday, and Friday, suggesting these days are more susceptible to fraudulent activities.

RECOMMENDATIONS

Transaction Amount By Fraud Indicator Insights

- Fraudulent transactions makes up about 4.18% of total data set.
- Most of the transactions appear to be legitimate which is 95.82%.

Recommendations:

- Refine fraud detection models for efficient management of a smaller percentage of fraudulent transactions, while

- Continuously updating them with advanced techniques and expert collaboration to enhance accuracy.

Transaction Amount Analysis Insights

- Online and Retail has the highest Fraud transactions
- Travel category has the least amount of transactions

Recommendations:

- Prioritize high-value categories(online and retail) for effective fraud detection

- Enhance methods to uncover potential low-value fraudulent activities, ensuring comprehensive prevention measures.

Age Group Analysis Insights

- The age group 25-34 are more likely to be involved in fraudulent activities.
- One fraud customer in 0-18.

Recommendations:

- Allocate extra resources for monitoring and investigating transactions for 25-34
- Implement strict verification for minors.

Transaction Behavior by Age Analysis Insights

- No significant correlation between a customer's age and the transaction amount.

Recommendations:

- Divide customers into age groups and analyze transaction behavior within each group to uncover potential insights for targeted fraud detection strategies.

Transaction Behavior Analysis by Age Group Insights

- Age Group 25-34 has the highest amount of transactions
- Age Group 0-18 has the lowest amount of transactions

Recommendations:

- Enhance fraud detection strategies for the highest transaction group (25-34) and consider tailored measures for the other age groups to address their distinct risk profiles

Anomaly scores **Analysis Insights**

- In the dataset, we lack information about the pattern of the anomaly score and how it is calculated.

Recommendations:

- Collaborate with experts to gain insight into the contributing factors and potential algorithmic methods used for calculating the score.

Anomaly Scores By Category Analysis Insights

- Online and travel categories have the highest average anomaly scores
- Online category has the highest standard deviation anomaly scores.

Recommendations:

- Give added focus to anomaly detection in "Online" and "travel" categories to prevent potential fraud
- Customize fraud detection models for the "Online" category to ensure accurate identification of transaction behaviors.

Highest Transaction Amounts Analysis Insights

- No significant difference between each category.
- Retail and Online categories have Highest transaction amount of 99.60 and 99.52
- Travel has least 98.52 transaction amount.

Recommendations:

- Employ biometric authentication for high-value transactions or sensitive account activities.
- Enhance fraud detection by integrating additional features and behaviors beyond transaction amounts, while ensuring thorough attention to all categories to avoid oversights.

Total Transaction By Month Analysis Insights

- 38 Frauds in January and 7 in February.
- Having data for only two months might restrict the ability to capture the complete range of fraudulent behaviors and patterns.

Recommendations:

- Continuously update and fine-tune fraud detection models using historical data and feedback loops
- Employ expert level techniques that can help to enhance the fraud detection accuracy despite the limited timeframe.

Fraud Transactions By Days Analysis Insights

- Notable spikes in fraud occurrences are observed on Monday, Wednesday, and Friday, suggesting these days are more susceptible to fraudulent activities.

Recommendations:

- Focus more resources and monitoring efforts on Monday, Wednesday, and Friday to prevent and quickly respond to potential fraudulent activities.


THANK
YOU


ANY QUERIES?