Anomaly Detection in Credit Card Transactions using Power BI

Objective:

Develop a Power BI dashboard to analyze credit card transactions, detect anomalies, and visualize transaction patterns.

Data Source:

Dataset:

- Source: Newton School Platform
- Features: step, type, amount, nameOrig, oldbalanceOrg, newbalanceOrig, nameDest, oldbalanceDest, newbalanceDest, isFraud

Data Preprocessing:

1. Data Cleaning:

- Looked for any missing values, duplicates, and ensured data integrity.

2. Data Transformation:

- Transformed the dataset into a format suitable for Power BI, addressing proper data types and structure.

Power BI Dashboard:

1. Overview Section:

- Transaction statistics, total transactions, average transaction amount, and frequency.

2. DAX Functions:

- Calculated average transaction amount for normal and fraudulent transactions.
- Counted total credit card transactions and total fraudulent transactions.

- Determined the highest fraud transaction amount.
- Compared maximum transaction amounts for normal and fraudulent transactions.
- Calculated the percentage of fraudulent transactions.
- Created a clustered column chart for the distribution of transaction amounts.

DAX Formulae Explanation:

a. DAX Function 1: Average Normal vs. Fraudulent Transaction Amount

```
Average Normal Transaction =
CALCULATE(AVERAGE(Fraud[amount]),Fraud[isFraud] = 0)

Average Fraudulent Transaction =
CALCULATE(AVERAGE(Fraud[amount]),Fraud[isFraud] = 1)
```

b. DAX Function 2: Count of Total Credit Card and Fraudulent Credit Card Transactions

```
Total Credit Card Transactions = COUNTROWS(ALL(Fraud))

Total Fraudulent Credit Card Transactions =
CALCULATE(COUNTROWS(Fraud),Fraud[isFraud] = 1)
```

c. DAX Function 3: Highest Fraud and Normal Transaction Amount

```
Highest Fraud Transaction Amount =
CALCULATE(MAX(Fraud[amount]),Fraud[isFraud] = 1)

Highest Normal Transaction Amount = MAXX(FILTER(Fraud,Fraud[isFraud] =
0),Fraud[amount])
```

d. DAX Function 4: Percentage of Fraudulent Transactions

```
Percentage Fraudulent Transaction = DIVIDE([Total Fraudulent Credit
Card Transactions],[Total Credit Card Transactions])
```

These DAX formulae provide calculations for the specified questions in the Power BI dashboard, enabling the generation of insightful visualizations and metrics.

3. Anomaly Visualizations:

- Implemented visualizations (cluster bar chart, scatter plot, line chart) to highlight potential anomalies and outliers in credit card transactions

Summary:

- Upon analyzing credit card transactions, a significant number of fraudulent activities have been uncovered within the dataset.

- The total recorded transactions are 6,30,894 with 383 fraud transactions, which is 0.06% of the total transactions.

- Transactions were categorized into five types: cash in, cash out, debit, payment, and transfer. Notably, fraudulent transactions were detected primarily within the 'transfer' and 'cash out' categories, with approximate values of $169M and $168M, respectively (Overall total fraud amount is $337.65M). This highlights the need for increased attention to ensure the safety of these transaction types.

- The 'cash out' type is the most frequently used, with 224,013 transactions totalling around $41 billion. Conversely, the 'debit' type is the least utilized, with only 4,769 transactions amounting to approximately $29M.

- The line chart reveals a concentration of transactions within the range of 10 to 20 steps, accounting for the highest total transaction volume, approximately $7.3 billion. Of this amount, a significant portion, totalling $3.3 billion, falls into the 'cash out' category, representing a notably high-value subset.

- Among top 10 merchants with highest number of transactions, 'M1964436639', 'M1098072792' and 'M1115616079' stands out with the highest top 3 transaction amount of $93K, $92K and $91K respectively.

In conclusion, the analysis emphasizes the importance of staying alert and proactive in preventing fraud, especially in 'transfer' and 'cash out' transactions. It also underscores how understanding the distribution of transaction volumes, spotting repeat offenders, and having strong fraud detection systems are vital in protecting financial transactions.

Deployment:

1. Power BI Deployment:

- Deployed the Power BI dashboard on GitHub.

2. Access Control:

- Implemented user authentication for secure access.

Documentation:

1. Project Details:

   - Defined project objectives, scope, and stakeholders.

2. Data Dictionary:

   - Described data features and their meanings.

3. Code Documentation:

   - Documented data preprocessing steps, DAX functions, and visualization logic.

4. Deployment Instructions:

   - Provided a step-by-step guide for accessing and using the Power BI dashboard.

5. Future Recommendations:

   - Suggested potential improvements and future iterations.