

## Anomaly Detection in Credit Card Transactions using Power BI

### About:-

Anomaly detection in credit card transactions refers to the process of identifying unusual or fraudulent activities in credit card transactions. It involves applying statistical, machine learning and Power BI techniques to detect patterns and deviations from normal behaviour, helping to identify potential fraudulent transactions in real-time.

### Project Overview:

The objective of this project is to develop a Power BI dashboard for anomaly detection in credit card transactions. Anomaly detection is crucial for detecting fraudulent activities and ensuring the security of credit card transactions. By leveraging Power BI's data visualisation and analytical capabilities, we can create an interactive dashboard that provides insights into transaction patterns and identifies potential anomalies.

### Project Steps:

---

- **Dataset Info:**

- ❖ **step** - maps a unit of time in the real world. In this case 1 step is 1 hour of time. Total steps 744 (30 days simulation).
- ❖ **type** - CASH-IN, CASH-OUT, DEBIT, PAYMENT and TRANSFER.
- ❖ **amount** - amount of the transaction in local currency.
- ❖ **nameOrig** - customer who started the transaction
- ❖ **oldbalanceOrig** - initial balance before the transaction
- ❖ **newbalanceOrig** - new balance after the transaction
- ❖ **nameDest** - customer who is the recipient of the transaction
- ❖ **oldbalanceDest** - initial balance recipient before the transaction. Note that there is no information for customers that start with M (Merchants).
- ❖ **newbalanceDest** - new balance recipient after the transaction. Note that there is no information for customers that start with M (Merchants).
- ❖ **isFraud** - This is the transactions made by the fraudulent agents inside the simulation. In this specific dataset the fraudulent behaviour of the agents aims to profit by taking control of customers accounts and try to empty the funds by transferring to another account and then cashing out of the system.

Import the Dataset given and follow the steps to accomplish the task

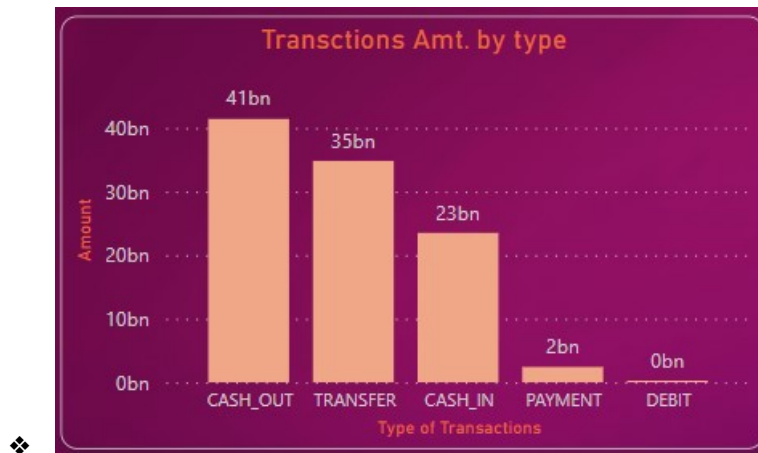
- **Data Preprocessing:**

- ❖ Performed data cleaning and transformed the data ensuring suitable data types.

- **DAX Function:**

Perform all the below questions using DAX functions:

- ❖ What is the average transaction amount for normal transactions versus fraudulent transactions?
  - `Average_Transactions_amt = AVERAGEX(FILTER('Fraud', 'Fraud'[isFraud]= 0), 'Fraud'[amount])`
  - `Average_Fraud_Transactions_amt = AVERAGEX(FILTER('Fraud', 'Fraud'[isFraud]= 1), 'Fraud'[amount])`
- ❖ How many credit card transactions were recorded in the dataset? And How many fraudulent credit card transactions were recorded in the dataset?
  - `Total_Transactions = COUNTROWS('Fraud')`
  - `Fraudulent_Transactions = CALCULATE(COUNTROWS('Fraud'), 'Fraud'[isFraud] = 1)`
- ❖ What is the highest Fraud transaction amount recorded?
  - `HighestFraudTransactionAmount = MAXX(FILTER('Fraud', 'Fraud'[isFraud] = 1), 'Fraud'[amount])`
- ❖ Is there a significant difference in the maximum transaction amount for normal transactions compared to fraudulent transactions?
  - `diff_max_TransactionAmount = CALCULATE(max(Fraud[amount]), Fraud[isFraud] = 1) - CALCULATE(max(Fraud[amount]), Fraud[isFraud] = 0)`
- ❖ What is the percentage of fraudulent transactions in the dataset?
  - `%age_Fraud_Transactions = DIVIDE([Fraudulent_Transactions], [Total_Transactions], 0)`
- ❖ What is the distribution of transaction amounts? (using Clustered column chart)



- **Anomaly Visualisation:**

Develop visualisations that highlight potential anomalies in the credit card transactions.

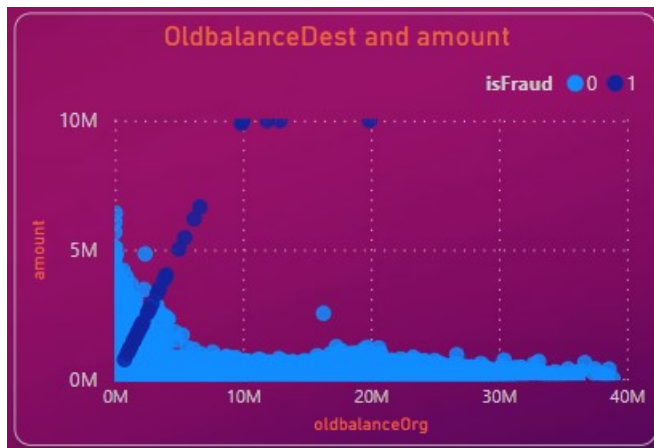
Use line charts, scatter plots, or heat maps to display transaction patterns and identify outliers.

- ❖ Which merchants have the highest number of transactions? (Only Top 10)

nameDest	Count of amount
C985934102	95
C97730845	74
C665576141	85
C451111351	76
C248609774	87
C2083562754	85
C1883840933	60
C1590550415	83
C1360767589	79
C1286084959	89
<b>Total</b>	<b>813</b>

A table with two columns, one with the name and other with the count of occurrence will be good to visualize the top 10 merchants as tables are structured and easily visualize making it a good option.

- ❖ Create a scatter plot to visualise the relationship between 'oldbalanceOrg' and 'amount' columns.



The scatter chart is plotted to visualize the relationship between 'oldbalanceOrg' and 'amount'. As it's easy to look for any unusual patterns or outliers in this plot and numerous data can be plotted at once in a scatter chart.

- ❖ Use a line chart to plot the transaction amount over time (step) to identify any unusual spikes or drops in transaction amounts.



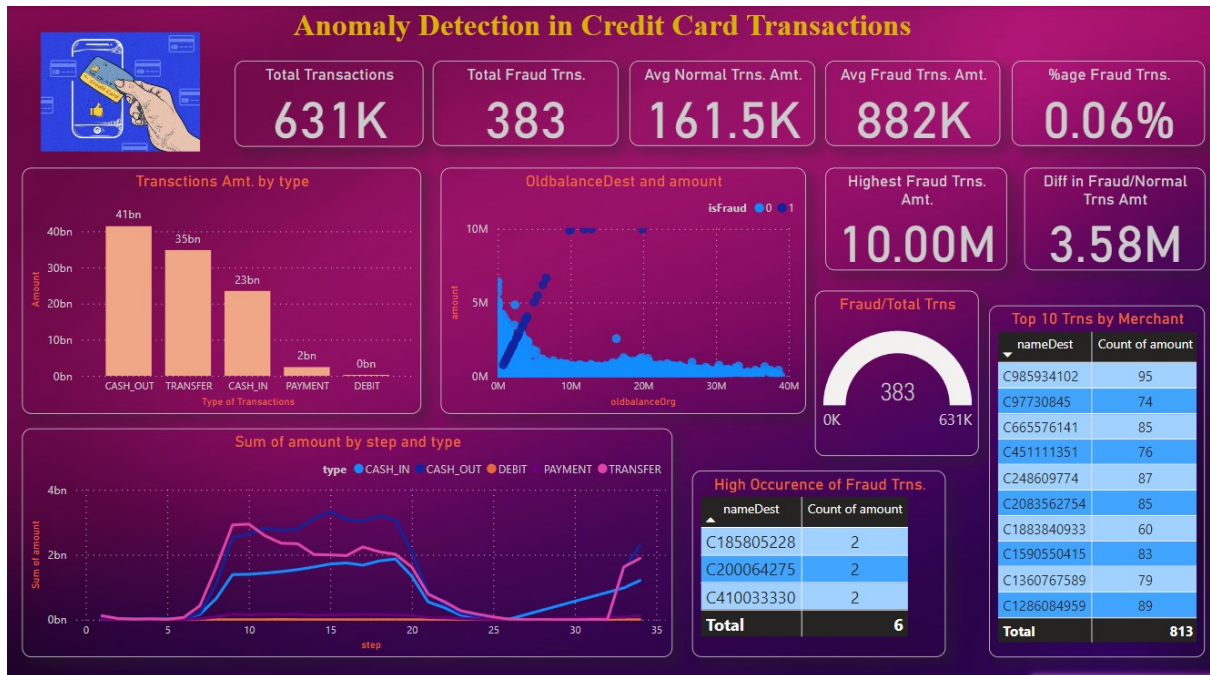
This chart can help identify unusual spikes or drops in transaction amounts at specific times. The sum of amount is on the y-axis and step is on the x-axis, with all types of transactions occurring.

- ❖ Are there any merchants with a high occurrence of fraudulent transactions?

High Occurrence of Fraud Trns.	
nameDest	Count of amount
C185805228	2
C200064275	2
C410033330	2
<b>Total</b>	<b>6</b>

This chart can visualize the merchants with a high occurrence of fraudulent transactions easily with the count.

- **Documentation and Deployment:**



## ANALYSIS

- As per the provided dataset, It was analyzed that there are multiple fraudulent transactions present.
- In the given dataset, there are 631K (6,30,895) transactions in total, out of which 383 transactions are fraudulent transactions, making it 0.06% of total. This indicates the proportion of fraud transactions in the dataset.
- Insights into the typical transaction amounts for normal and fraudulent transactions. The average Normal transactions amount is 161K, however if we look into the fraudulent transactions it was 882K as can be visualized which is substantial and hence, concerning.
- Additionally, the maximum amount identified for fraudulent is 10.00M which is very high. Moreover, the difference between the fraud transactions and normal transactions is observed as 3.58M which also is high indicating fraud transactions are of very high amounts.
- Most of the fraudulent transactions were either done as Cash\_outs and Transfers indicating that this two transaction types requires more vigilance to ensure proper safety measures against fraud.
- By analyzing the line chart and clustered column chart, it can be observed that the most used type of transactions are cash\_out and transfers and the least used transactions type is debits. The line chart can also help identify unusual spikes or drops in transaction amounts at specific times.
- Top 10 Merchants identified as conducting highest number of transactions with their count of transactions are visualized in the dashboard.
- Also, top 3 Identified Merchants with high occurrence of fraudulent transactions are C185805228, C200064275 and C41003330 with the count as 2 (twice) making these potential anomalies.
- Visualizing the distribution of transaction amounts helps to identify common patterns and potential anomalies. Further the dashboard can be visualized for any more information on the provided dataset.