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
  - ❖ *oldbalanceOrg* - 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

- **Power BI Dashboard Creation:**

  - ❖ Launch Power BI and connect to the preprocessed credit card transaction dataset.
  - ❖ Design an intuitive and visually appealing dashboard layout with appropriate charts, tables, and filters.
  - ❖ Create visualisations that provide an overview of transaction statistics, such as total transactions, average transaction amount, and transaction frequency.

- **Data Preprocessing:**

  - ❖ Perform data cleaning tasks, such as handling missing values and duplicates if needed.
  - ❖ Transform the data into a format suitable for Power BI, ensuring proper 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 normal amount = CALCULATE(AVERAGE(Fraud[amount]),Fraud[isFraud]=0)`
  - ❖ `average Fraud amount = CALCULATE(AVERAGE(Fraud[amount]),Fraud[isFraud]=1)`

  - ❖ How many credit card transactions were recorded in the dataset? And How many fraudulent credit card transactions were recorded in the dataset?

  - ❖ Count of Total Transaction = COUNTROWS(Fraud)

  - ❖ Count Of Fraud Transactions= CALCULATE(COUNTROWS(Fraud),Fraud[isFraud]=1)
  - ❖ What is the highest Fraud transaction amount recorded?
  - ❖ Max Fraud Amount = CALCULATE(MAX(Fraud[amount]),Fraud[isFraud]=1)
  - ❖ Is there a significant difference in the maximum transaction amount for normal transactions compared to fraudulent transactions?
  - ❖ Max Fraud Amount = CALCULATE(MAX(Fraud[amount]),Fraud[isFraud]=1)
  - ❖ Max Normal Amount = CALCULATE(MAX(Fraud[amount]),Fraud[isFraud]=0)
  - ❖ (Max)Diff Normal and Fraud = CALCULATE([Max Fraud Amount]-[Max Normal Amount])

- ❖ What is the percentage of fraudulent transactions in the dataset?
- ❖ % of Fraud Transaction = [Count Of Fraud Transactions]/[Count of Total Transaction]*100
- ❖ What is the distribution of transaction amounts? (using Clustered column chart)


Sum of amount by type and isFraud

- ❖
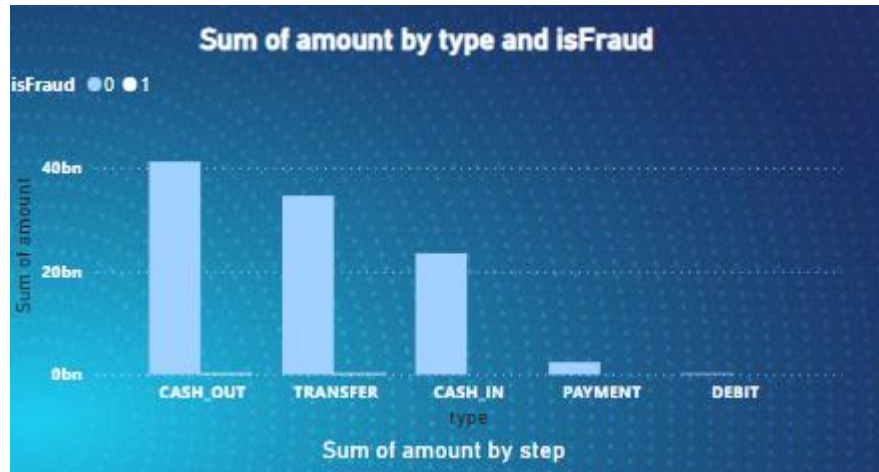
● **Anomaly Visualisation:**

- ❖ Developed visualisations that highlight potential anomalies in the credit card transactions. Used line charts, scatter plots, and tables to display transaction patterns and identify outliers.

  ❖ Merchants with highest number of transactions.(Only Top 10) The tile comprises two columns(name, count of highest transaction) in a table. A table is used here so that the data is compared easily among different rows.

Top 10 Transaction amount	
nameDest	Count of amount
C985934102	95
C1286084959	89
C248609774	87
C2083562754	85
C665576141	85
C1590550415	83
C977093101	80
Total	**630894**

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

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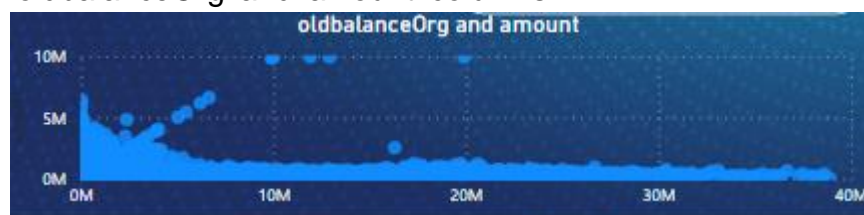
- ❖
- ❖ Use a line chart to plot the transaction amount over time (step) to identify any unusual spikes or drops in transaction amounts.

Sum of amount by step
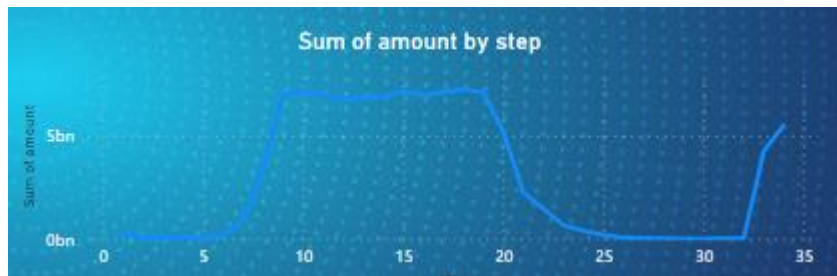
❖
❖ Are there any merchants with a high occurrence of fraudulent transactions?

nameDest	Total fraud transactions
C185805228	2
C200064275	2
C410033330	2
C1002031672	1
C1007251739	1
C1009459055	1
Total	**383**

Merchant with Fraud Transactions

❖