**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?
* How many credit card transactions were recorded in the dataset?

**Total Transaction = COUNTROWS(FRAUD)**

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

**Fraudulent transaction Count = COUNTROWS(FILTER('FRAUD',FRAUD[isFraud] = 1))**

* What is the highest Fraud transaction amount recorded?
* Is there a significant difference in the maximum transaction amount for normal transactions compared to fraudulent transactions?
* What is the percentage of fraudulent transactions in the dataset?
* 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)

C985934102,C1286084959, C248609774, C2083562754, C66557614, C1590550415, C1360767589, C451111351, C97730845, C1883840933.

In the top ten, Merchant C985934102 leads with the highest number of transactions (95), while Merchant C1883840933 has the lowest number of transactions (60). The table chart simplifies the process of identifying the top ten and enables a clear comparison of transaction numbers.

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

There is a significant difference in transactions between 'oldbalanceOrg' and 'amount.' Utilizing a scatter plot facilitates the easy identification of fraudulent transactions, where a distinct diagonal line is evident in the chart.

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

After the 8th step (8 hours), there is an unusual spike in transaction amounts, totalling $3,439,602,407.35, predominantly in the form of transfers and cash withdrawals. The amount experiences a significant drop after the 20th step (20 hours), reaching $5,173,930,961.93. By utilizing a line chart, one can easily observe the distinct spikes and drops in the transaction amounts.

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

These are the top five merchants, identified by their frequent occurrence in fraudulent transactions: C9901046, C991505714, C185805228, C200064275, and C410033330. A table chart provides a straightforward means of displaying the names of these top five merchants.

Also, Explain why do you think the above mention charts are best to visualise the problem

* **Documentation and Deployment:**
* Create summary sections or report pages that provide key insights derived from the transaction data.
* Document the project details, including data sources, preprocessing steps, and dashboard features.
* Remember to comply with data privacy and security regulations while working with credit card transaction data

By following these steps, you will be able to develop a comprehensive Power BI project on anomaly detection in credit card transactions, empowering users to identify and respond to potential fraud efficiently.

**Insights**

1. There are 6.42 million normal transactions.
2. A total of 631k transactions were conducted using credit cards.
3. The bank recorded 10 million fraudulent transactions.
4. Merchants accounted for 383 fraudulent transactions.
5. The highest fraudulent transaction amount reached 3.58 million.
6. The graphical chart illustrates various transaction types, including cash in, cash out, transfer, payment, and debit. Notably, most fraudulent transactions occur in cash out and transfer categories.
7. Debit card transactions have the lowest frequency of fraud, indicating a lower occurrence of both fraudulent and normal transactions.
8. On average, fraudulent transactions make up 84.48%, while normal transactions constitute 15.52%. This suggests a higher frequency of fraudulent transactions.
9. The top ten merchants with the highest transaction volumes are C985934102, C1286084959, C248609774, C2083562754, C665576141, C1590550415, C1360767589, C451111351, C97730845, and C1883840933.
10. The top 3 merchants with the highest number of fraudulent transactions, particularly in cash out and transfer modes, are C9901046, C991505714 and C185805228.
11. The addition of a slicer allows for the easy classification of transactions as either 0 (no fraud) or 1 (fraud), enhancing visualization and analysis capabilities.