**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? And How many fraudulent credit card transactions were recorded in the dataset?
* 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)
* Create a scatter plot to visualise the relationship between 'oldbalanceOrg' and 'amount' columns.
* Use a line chart to plot the transaction amount over time (step) to identify any unusual spikes or drops in transaction amounts.
* Are there any merchants with a high occurrence of fraudulent transactions?

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

* **Documentation and Deployment:**

**Key Performance Indicators (KPIs)**

Our anomaly detection system employs a range of Key Performance Indicators (KPIs) to assess transaction behavior. KPIs include average transaction amounts, transaction frequency, geographic origin, and the time of day transactions occur. These metrics, when visualized through interactive dashboards, provide a comprehensive view of transaction trends and help us identify unusual patterns.

**DAX Formulas**

To facilitate our analysis, we've employed Data Analysis Expressions (DAX) to create custom measures and calculated columns. These DAX formulas enable us to calculate statistical metrics, such as moving averages, standard deviations, and z-scores, which play a critical role in the identification of anomalies. By using DAX, we can compute values based on these formulas and incorporate them into our detection algorithms.

Our anomaly detection system's effectiveness in protecting credit card transactions is driven by the utilization of these KPIs and DAX formulas, as well as the integration of machine learning models. This multifaceted approach allows us to continuously monitor and adapt to emerging fraud patterns, ensuring the security and trustworthiness of credit card transactions.

* **Measures:**
  + We have developed a set of essential measures to gain insights into our credit card transactions. The "% of Fraud Transactions" and "% of Normal Transactions" measures provide a clear overview of the distribution of fraudulent and legitimate transactions within our dataset.
  + "Average Fraud Transaction Amount" and "Average Normal Transaction Amount" measures help us understand the typical transaction values for these two categories.
  + For a more in-depth analysis, we also utilize "Max Fraud Transaction Amount" and "Max Normal Transaction Amount" measures to identify the highest transaction values for both fraud and normal transactions.
  + "Total Credit Card Transactions" serves as a fundamental measure that provides the total count of transactions in our dataset, forming the basis of various analyses.
* **Column Chart:**
  + We employ a dynamic column chart to visualize the sum of transaction amounts by transaction type (fraud or normal). This chart offers a quick visual comparison of the total transaction values in each category, facilitating easy pattern recognition.
* **Bar Chart - Top 10 Merchants:**
  + Our bar chart showcasing the "Top 10 Merchants" is a valuable resource for identifying the most active and influential merchants within our transaction data. It allows us to pinpoint which merchants engage in the most significant number of transactions.
* **Line Chart - Old Balance vs. Amount:**
  + The "Old Balance vs. Amount" line chart provides insights into how account balances relate to transaction amounts over time. This dynamic visualization helps us track changes in account balances, allowing for the identification of unusual patterns.
* **Line Chart - Amount Over Time:**
  + The "Amount Over Time" line chart offers a step-by-step representation of transaction amounts over time. This helps us monitor transaction trends and spot any significant deviations or anomalies in the data.

Bottom of Form