**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.

**Data Preprocessing:**

Perform data cleaning tasks, such as handling missing values and duplicates if needed - done

Transform the data into a format suitable for Power BI, ensuring proper data types - done

**DAX Function:**

Perform all the below questions using DAX functions:

**What is the average transaction amount for normal transactions versus fraudulent transactions?**

* average transaction amount for normal transactions = 161.50k
* average transaction amount for fraudulent transactions = 881.59K

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

* credit card transactions were recorded in the dataset =631K
* fraudulent credit card transactions were recorded in the dataset = 383

**What is the highest Fraud transaction amount recorded?**

* 10Million

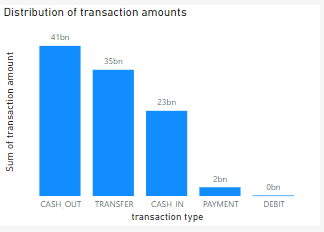
**Is there a significant difference in the maximum transaction amount for normal transactions compared to fraudulent transactions?**

* Yes.3.58 million

**What is the percentage of fraudulent transactions in the dataset?**

* 0.06%

**What is the distribution of transaction amounts? (using Clustered column chart)**

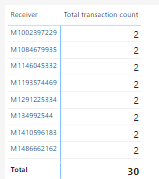


**Anomaly Visualisation:**

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

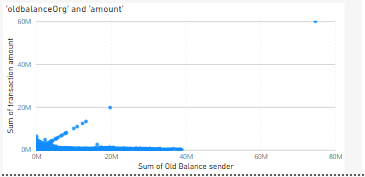
Use line charts, scatter plots, or heat maps to display transaction patterns and identify outliers.

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



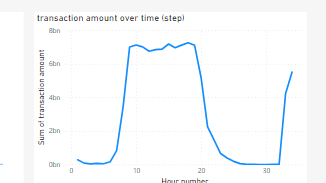
* Clustered column charts allow us to easily compare the number of transactions for different merchants side by side. Each column represents a different merchant, making it straightforward for users to compare the transaction counts.

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



* Scatter plots are designed for visualizing relationships between two continuous variables. In this case, both 'oldbalanceOrg' and 'amount' are numeric, continuous variables, making a scatter plot an appropriate choice.

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



* Line charts are particularly effective for displaying time series data, where we have data points recorded at regular time intervals (e.g., daily, weekly, monthly). Each data point is connected by a line, which allows us to visualize trends and patterns over time.

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

* I cannot see such observation

**Documentation and Deployment:**

**Create summary sections or report pages that provide key insights derived from the transaction data.**

* CASHOUT is he transaction type having maximum count of transactions
* 0.06% of Transactions are fraud
* CASHOUT accounted for 40.54% of Sum of Transactions

**Document the project details, including data sources, preprocessing steps, and dashboard features.**

1. **Data Loading:**

First, we loaded the CSV file into Power BI using the "Load Data" option. This step brought the external data into our Power BI project.

2. **Data Transformation:**

After loading the data, we proceeded to the Power Query Editor to transform and clean it. This involved tasks such as data type conversions, splitting columns, and merging tables to prepare the data for analysis.

3. **Data Quality Check:**

Fortunately, the data was relatively clean, with no error values or missing data points. This ensured that we had a reliable dataset to work with.

4. **Column Name Modification:**

To enhance the understandability of the data, we decided to change some column names. Renaming columns can make the dataset more user-friendly and align it with our analysis objectives.

5. **Measures Creation:**

To extract specific insights from the data, we created various measures. Measures are calculations that summarize and aggregate data, allowing us to gather meaningful information. These measures were tailored to our specific analysis needs.

6. **Dashboard Creation:**

Finally, in the report view of Power BI, we utilized the measures and data visualizations to create a dashboard. This dashboard showcased different trends and patterns observed in the data, making it easier for users to grasp and interpret the insights we derived.Remember to comply with data privacy and security regulations while working with credit card transaction data