Database Project

E-commerce Fraud Detection System

PowerBI Analysis & Insights

Data Understanding:

Can you explain the source of the data used in these dashboards?

Three Amazon datasets from a MySQL database

How did you preprocess and clean the data before visualizing it?

Using Power Query

Did you encounter any challenges with the data quality, and how did you address them?

Yes. I changed the columns' data type accordingly. Dealt with null values and split the columns to provide more features.

Dashboard Design:

What considerations did you take into account when designing the dashboards?

Provide necessary interactions to allow the admin to filter the data as he pleases and maintain consistency in design elements, such as color schemes, fonts, and layout, to create a cohesive and professional appearance.

How did you choose the visualizations to represent different aspects of fraud detection?

Depending on the type of metric I want to visualize: line plots for date & time, maps for countries and states,..etc

Key Metrics and Indicators:

What are the primary metrics you're monitoring for fraud detection?

For Credit Card Fraud:

1. Purchase Value
2. Country the card was used in
3. Transaction Date
4. Transaction Device / Browser

For Human / Bot Detection:

1. Account Creation Date
2. Account Language
3. Does the account use default settings
4. Does it has a membership / Geographical Visibility
5. Number of Followers / Friends
6. Number of Purchases

For Product / Content Moderation Fraud:

1. Product Category & Subcategory
2. Product Price
3. Product Company Origin
4. Product Company Logistics & Industry

Data Analysis and Insights:

1. **Credit Card Fraud Rate is similar across device type / browser**

   This insight indicates that credit card fraud is possible through many means and does not indicate a lack of security in a specific browser / device.

2. **Money Spent is significantly higher for fraudulent credit card**

   This insight indicates that if a purchase amount suddenly rose, there might be a credit card fraud in action.

3. **Credit Card Fraud is mostly common at the beginning and end of the year**

   The combination of increased spending, higher transaction volumes, seasonal promotions, and potential vulnerabilities in financial systems makes the beginning and end of the year prime times for credit card fraud.

4. **Having a default account increases the possibility of a bot account**

   This occurs because of ease of creation, lack of customization, and potential vulnerabilities in security measures

5. **Less popular languages are used in bot accounts**

   This may used to hide the true personality of the bot creator or lower his chances of getting a report

6. **Bots have higher purchases per day but a lower number overall**

   Bots often are programmed for a very specific task under very specific conditions. Since the probability of the condition is low, the overall number of purchases is low. However, when that condition is met, the spamming begins.

7. **Bots have low number of followers, but can have both low and high number of friends.**

8. **Fraudulent Products have a lower price than non-fraudulent products**

   This happens, probably, to speed up the selling process and lower the chance of getting caught.

9. **Fraudulent Products tend NOT to have a specific subcategory / specifics mentioned.**

10. **Number of reports / feedback and the presence of a company logo has little influence on a product's possibility of fraud.**