Take-home test for Card Fraud Analyst.

Assignment I:

You are in the process of reviewing card Fraud related activity and in particular merchants activity and possible trends, with analysis on transactions at select merchants over the last 30 days shown below. Define and explain the next steps you would take to prevent further losses at the following merchants, and indicate which merchants you would prioritise (rank them!) for preventative measures.

False Positive Ratio = Number of False Positives : Number of True Positives

Merchant Name	Transaction Type	Fraud Loss ($)	False Positive Ratio	Total Number of Transactions (last 30 days)	Fraud Prevention Action & Reasoning
Tickedek	ECOM	15,000	89:1	1350	
Sp* Natilaus	POS	3,500	2:1	9	
Shinet.com	ECOM	6,000	12:1	156	
Bet350	ECOM	14,000	5:1	18	

My solution:

Here's how I would prioritize the merchants:

1. Tickedek, 2. Bet35, 3. Shinet.com, 4. Sp* Natilaus

Now I'll explain why I chose this order and explain further:

1. **Tickedek**

This merchant has the highest fraud loss and the highest false positive ratio, which means it should be prioritised in the first place.
I would suggest optimising our fraud detection algorithms since they're overly sensitive (a very high false positive ratio) but still miss a significant amount of fraud. For that we might use machine learning systems, which can analyse data and identify fraud patterns. Also I would consider the implementation of MFA for high-risk transactions, for example, those higher than 100$ or from unusual locations, if it isn't implemented yet.

2. **Bet350**

Almost the same amount of fraud (14000$ vs 15000$) but a much lesser amount of transactions, meaning a significant per-transaction loss. False positive ratio is moderate, but still misses critical frauds. I would suggest conducting a thorough transaction pattern analysis to better identify fraud and since there are far fewer transactions, it would be easier to implement strict transactions monitoring. We could review account creation processes: the merchant suffered significant loss despite having a few transactions. This suggests that fraudsters may be creating accounts specifically to commit fraud, and Bet350 current account creation process may not have sufficient safeguards to prevent this. Also, additional verification such as 3D Secure for high-risk transactions, or even tightening transaction rules and limits, could be considered.

3. **Shinet.com**

A relatively high false positive ratio and medium fraud loss. As in the Tickedek case, we need to reduce false positives and fraud by improving our fraud detection model through the analysis of fraud patterns. We might also recommend improving customer verification mechanisms or adding additional security layers.

4. **Sp* Natilaus**

It has the lowest fraud loss and good false positive ratio, but a relatively high fraud loss per transaction, so things could be improved. We might implement basic transaction monitoring. Since this is POS, we could recommend training staff on fraud detection and verify their POS security protocols such as POS software updates and EMV compliance

Assignment II:

You will find tables commonly used by our Card Fraud Team in the appendix (**pages 4-6**). The **REPORTS.PLASTIC_TRANSACTION** table contains information on all card transactions, and the **REPORTS.DISPUTED_TRANSACTIONS** table contains information on all disputed card transactions.

Please use these tables in your responses to the following questions:

a) You have noticed a wave of valid fraud disputes being reported at the ecommerce merchant **"Choco Life"** over the last week. The transactions are largely being reported on our MasterCard UK cards.

Write a sample SQL query you would use to calculate the fraud density at the merchant, with the focus on reducing false positives and creating a prevention strategy. You are free to decide on an appropriate lookback time frame for calculating the fraud density at this merchant.

b) While you are investigating the fraud density at **Choco Life**, an agent reports that the team has observed that a number of these fraudulent transactions have occurred approximately 5 minutes after zero-dollar card testers were made at merchant "**TokenEx**".

Write a sample SQL query you would use to calculate the fraud density for transactions at **Choco Life** occurring after zero-dollar card testers were made at **TokenEx**.

Query a)

```
-- Looking back 30 days to get a comprehensive view
WITH disputed_transactions AS (
    -- Select valid disputed transactions
    SELECT
        DATE(d.DATETIME_TRANSACTION_CREATION) AS tx_date,
        d.TRANSACTION_ID
    FROM REPORTS.DISPUTED_TRANSACTIONS d
    WHERE d.MERCHANT_NAME = 'Choco Life'
      AND d.DISPUTE_STATE = ‘VALID'
      AND d.DATETIME_TRANSACTION_CREATION >= CURRENT_DATE - INTERVAL '30 days'
    ),
    all_transactions AS (
    -- Select all MasterCard UK transactions at Choco Life in the last 30 days
    SELECT
        DATE(pt.CREATION_TIME) AS tx_date,
        pt.TRANSACTION_ID
    FROM REPORTS.PLASTIC_TRANSACTION pt
    WHERE pt.MERCHANT_NAME = 'Choco Life'
        AND pt.CREATION_TIME >= CURRENT_DATE - INTERVAL '30 days'
        AND pt.PROGRAM = 'MC_DEBIT_UK'  -- Filtering MasterCard UK
        AND pt.DETAILS_TYPE = ‘ECOM_PURCHASE’ — Optional
    )
-- Calculate fraud density
SELECT
    at.tx_date,
```

```
    (CAST(COUNT(DISTINCT dt.TRANSACTION_ID) AS FLOAT) / NULLIF(COUNT(DISTINCT
at.TRANSACTION_ID), 0)) AS fraud_density

FROM all_transactions at

LEFT JOIN disputed_transactions dt ON at.TRANSACTION_ID = dt.TRANSACTION_ID

GROUP BY at.tx_date

ORDER BY at.tx_date;
```

I've grouped transactions by date and chose 30-day time frame to look for trends, but if there are too few transactions per day, we can either change the grouping interval or remove the grouping at all. Usually, there are no duplicate transaction IDs, so «DISTINCT» may be also unnecessary.

Query b)

```
WITH token_ex_testers AS (

    -- Find zero-dollar transactions at TokenEx

    SELECT

        pt.CARD_TOKEN,

        pt.CREATION_TIME AS tester_time

    FROM REPORTS.PLASTIC_TRANSACTION pt

    WHERE MERCHANT_NAME = 'TokenEx'

        AND pt.AMOUNT = 0

        AND pt.CREATION_TIME >= now() - INTERVAL '7 days'

    ),

    choco_life_tx AS (

    -- Find Choco Life transactions occurring within 5 mins after a TokenEx tester on the same card

    SELECT

        pt.TRANSACTION_ID,

        pt.CARD_TOKEN

    FROM REPORTS.PLASTIC_TRANSACTION pt

    JOIN token_ex_testers tet ON pt.CARD_TOKEN = tet.CARD_TOKEN

    WHERE
```

```
        pt.MERCHANT_NAME = 'Choco Life'

        -- Check if Choco Life transaction time is between tester time and 5 mins after

        AND pt.CREATION_TIME > tet.tester_time

        AND pt.CREATION_TIME <= tet.tester_time + INTERVAL '5 minutes'

        ),

        fraud_choco_life AS (

        -- Find which of the Choco Life post-tester transactions were fraudulent

        SELECT

                dt.TRANSACTION_ID

        FROM REPORTS.DISPUTED_TRANSACTIONS dt

        JOIN choco_life_tx clt ON dt.TRANSACTION_ID = clt.TRANSACTION_ID

        WHERE dt.DISPUTE_STATE = ‘VALID'

        )

-- Calculate fraud density

SELECT

        (CAST(COUNT(DISTINCT fcl.TRANSACTION_ID) AS FLOAT) / NULLIF(COUNT(DISTINCT
clt.TRANSACTION_ID), 0)) AS fraud_density

FROM choco_life_tx clt

LEFT JOIN fraud_choco_life fcl ON clt.TRANSACTION_ID = fcl.TRANSACTION_ID;
```

This query should show the fraud density for the required transactions for the past 7 days.

Assignment III

You notice that our card fraud prevention alerting system is producing high levels of false positive alerts around a specific rule to prevent card-present fraud, meaning that Customers are unnecessarily inconvenienced with restricted access to their cards until the alert is manually dealt with by an agent. Once you flag this issue, you are asked to provide a Quality Control review of this alert creation procedure.

What focus areas and possible action items would your investigation suggest in your final write up?

My solution:

I'd focus my Quality Control review on these key areas:

1. Customer impact

Volume of affected customer, customer satisfaction metrics related to fraud prevention

2. Rule logic, design and parameters

Check the current thresholds and trigger conditions. Study the past performance of this rule and any recent changes that could have impacted it.

3. Operational impact on Wise's Fraud Team

Review agents workflow, average resolution time for false alerts, agent feedback on alert quality

4. Data quality

Review input data accuracy, completeness and latency

Possible Action Items

1. **Adjust or refine rules**

We can modify the alert parameters from the analysis we conduct or replace the rule altogether. The implementation of static rules which analyze only transaction amount or location produces elevated rates of false positives. Multiple conditions should be applied when analyzing transactions by combining amount with frequency and merchant type to improve detection accuracy.

2. I**ntegrate machine learning**

AI systems learn to detect new fraudulent patterns while adapting to user behavioural changes which leads to decreased false positive rates.

3. **Improve data quality**

We might acquire more user information to enhance our understanding of transaction contexts. The frequency of data checks combined with updates will help preserve data quality which leads to better results in reducing false positives.

4. **Optimization of work processes**

The manual workload could be decreased through automation of specific tasks, agents may provide feedback to enhance rules or train models. The implementation of scheduled rule assessments will enable timely necessary modifications.

Appendix: Tables for Assignment II

REPORTS.PLASTIC_TRANSACTION			
Column name	**Type**	**Description**	**Example**
CREATION_TIME	TIMESTAMP_	When the transaction was cr	2020-03-01 18:18:18.47385
TRANSACTION_ID	INTEGER	plastic transaction id	52682368
CARD_TOKEN	TEXT	the tokenized card	8ecc3c28-5057-49c4-b2ef-68f2
MERCHANT_COUNTR	TEXT	where the merchant is alloc	GB
DETAILS_TYPE	TEXT	type of transaction	ECOM_PURCHASE, POS_PURCHASE
POS_ENTRY_MODE	TEXT	how was the entry mode in the	CONTACTLESS
DETAILS_CARD_OWNERSHI	TEXT	How the ownership identity was	PIN
TERMINAL_ID	TEXT		28897203
CARD_PRESENCE	TEXT	Info whether it was a card present present transaction	Present
MERCHANT_NAME	TEXT	Merchant Name	DIGITALOCEAN COM
MERCHANT_ID	TEXT	Merchant ID	7600016737
MERCHANT_CITY	TEXT	City where the merchant	NEW YORK CITY
MERCHANT_ZIP	TEXT	Zipcode	75018
MCC	TEXT	four-digit number listed in ISO 182 financial services	9399
USER_ID	INTEGER	Users' user id	7379887
PROGRAM	TEXT	Program from where the user regis	MC_DEBIT_UK, MC_DEBIT_US, MC
AMOUNT_CURRENC	TEXT	authorization info	EUR
AMOUNT	INTEGER	authorization info	6
AMOUNT_GBP	DECIMAL	authorization info	5.49959214
DECLINE_REASON	TEXT	reason why auth transaction was	INSUFFICIENT_FUNDS
PAYMENT_TOKEN	TEXT	PAYMENT_TOKEN	2ae46a4d2b4fe3c491475604406e71a3 0fb84f8d330f7740c19

REPORTS.DISPUTED_TRANSACTIONS			
Column name	**Type**	**Description**	**Example**
TRANSACTION_I	INTEGER	Transaction ID	128619274
DATETIME_TRANSACTION_	TIMESTAMP_N	When the transaction was	2020-03-01 18:18:18.473850
CARD_TOKEN	TEXT	Tokenized card	867b2ab2-007d-4941-aa6d-ab17
USER_ID	INTEGER	Users's user id	5512413
MERCHANT_NAM	TEXT	Merchant name	Amazon Music
MCC	TEXT	four-digit number listed in ISO 1 financial services	8211
CASE_ID	INTEGER	the dispute's case i	125898
DATETIME_DISPUTE_CR	TIMESTAMP_N	when the transaction was	2020-03-01 18:18:18.473850
DISPUTE_STATE	TEXT	if the dispute is valid o	VALID, INVALID
SAFE_CATEGOR	TEXT	Type of dispute	00 Lost Fraud