Financial Analysis

By Jay Ganesh Charole

Objective

Financial institutions around the world are turning to data science to combat crime and manage compliance due to the changing nature of crime and a quickly expanding regulatory landscape.

The global financial crisis of 2008 altered the course of history. It had an impact not only on the financial industry, but also on other industries and enterprises around the world. The crisis exposed ineffective policies that resulted in severe fractures that threatened to bring the global financial system to its knees.

Objective

Technological advancements, and new capabilities to understand enormous volumes of data can help to analyze and formulate the best approach to identify flaws and appropriate interventions techniques to reduce financial crime.

AI, machine learning, and automation, among other advanced analytics and cognitive techniques, can help to filter out false positives and improve inefficiencies in existing investigation processes. Data and analytics have the potential to not only improve efficiencies and save operating costs, but also help identify intelligence-led and data-driven approaches to combating financial crime.

About Dataset

There are 3 datasets mentioned here: alerts, transactions and accounts.

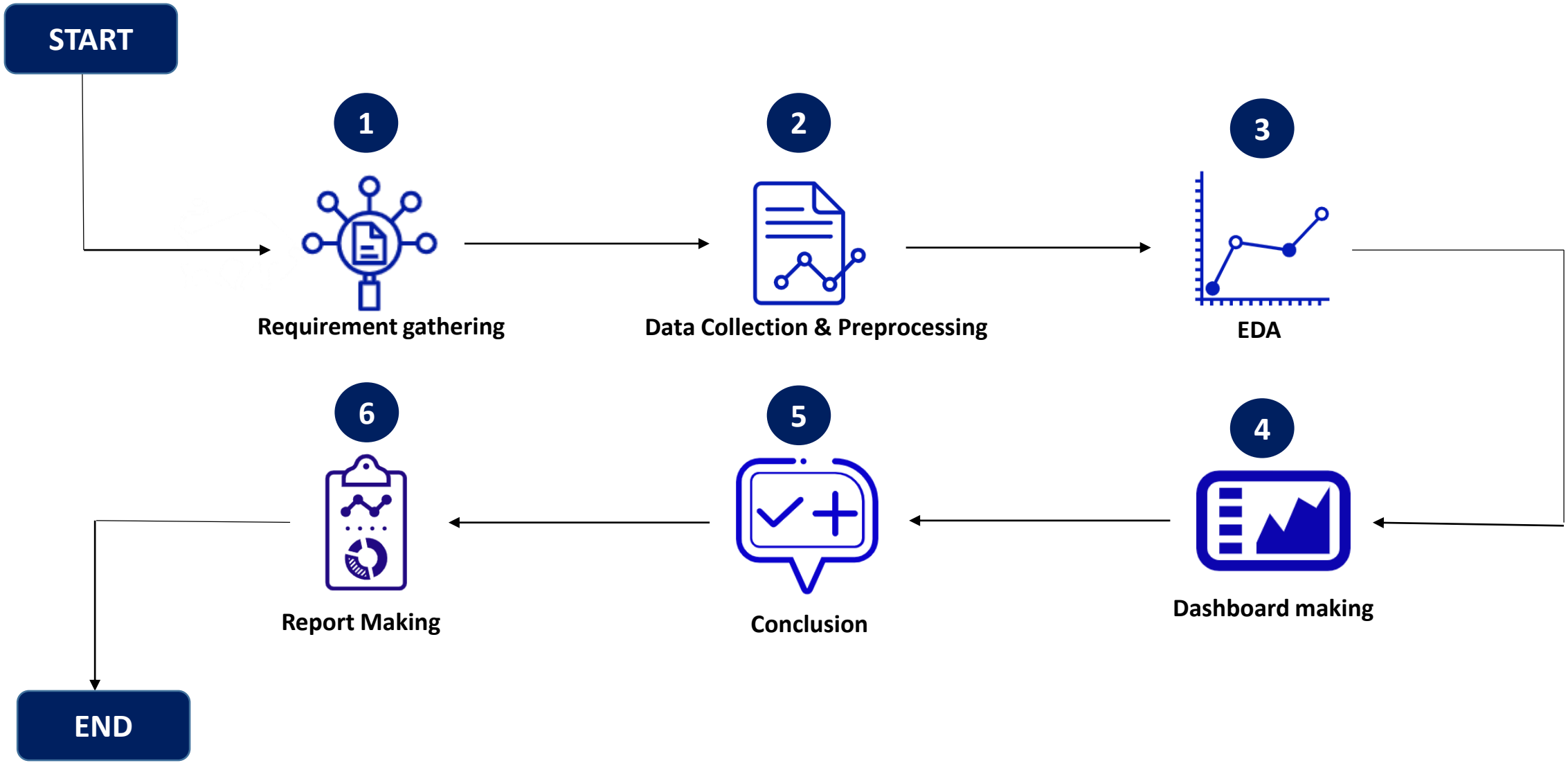
- **Accounts dataset:** Contains the information about all the bank accounts whose transactions are monitored.
- **Alerts dataset:** Contains the transactions which triggered an alert according to AML guidelines.
- **Transactions dataset:** Contains the list of all the transactions with information about sender and receiver accounts.

Tool Stack

Python is a high-level, interpreted, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Python is dynamically-typed and garbage-collected.

Tableau is a leading data visualization tool used for **data analysis and business intelligence**. Gartner's Magic Quadrant classified Tableau as a leader for analytics and business intelligence.

Project Process



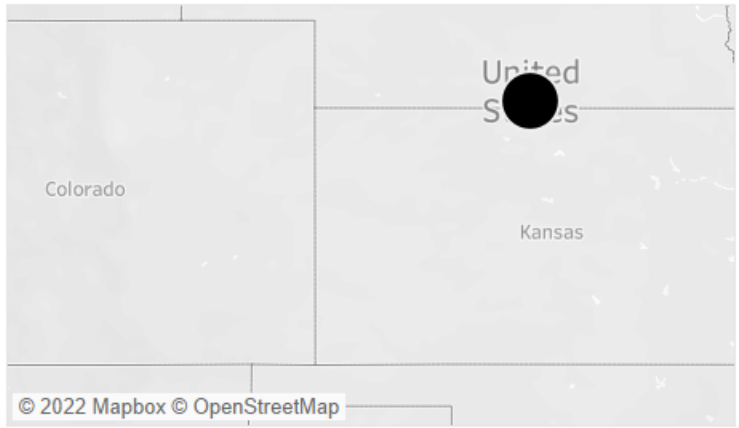
Dashboard

FINANCIAL CRIME ANALYSIS DASHBOARD

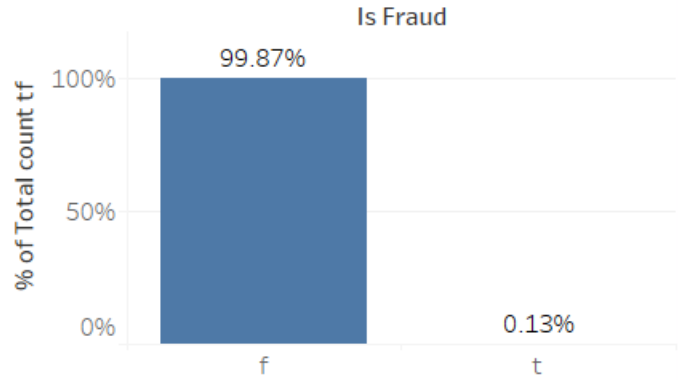
TX amount per alert Id

Alert Id	C ALERT ID	Tx Amount
-1	1,321,515	153,479,489,422
0	4	18
1	4	16
2	4	14
3	4	12
4	4	13
5	4	14
6	4	20
7	4	19
8	4	16
9	4	19
10	4	19

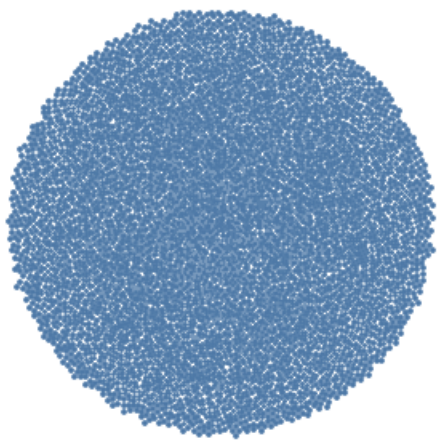
Analysed Country



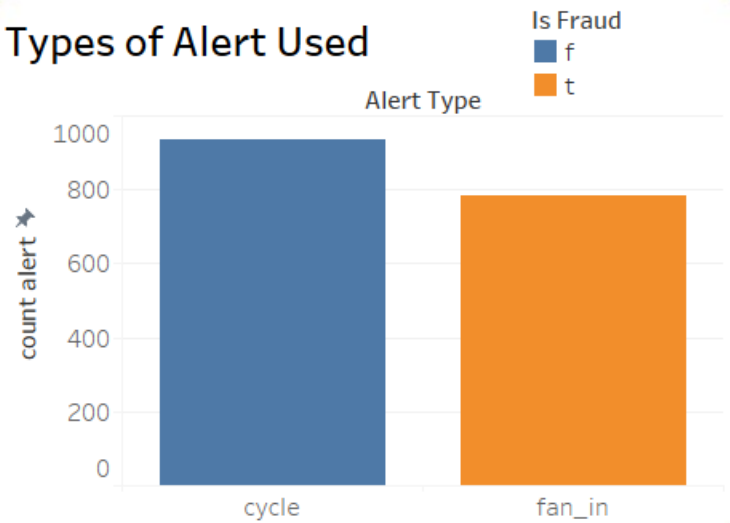
Percentage of Frauds



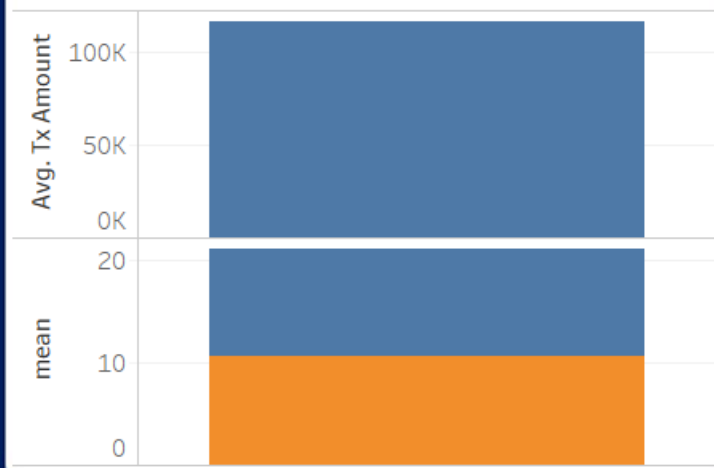
Timestamps as per Reciver account ID



Types of Alert Used

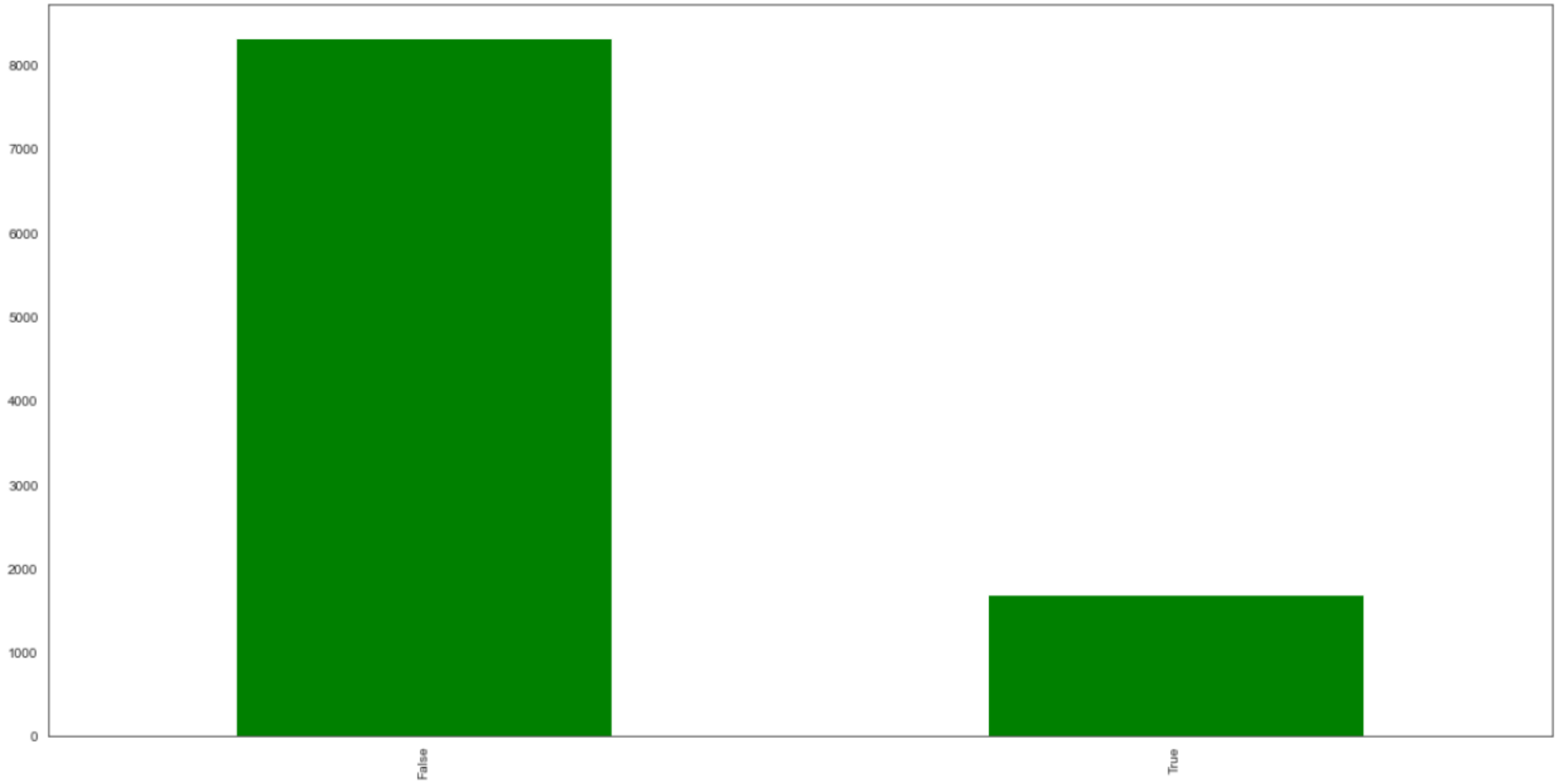


Tx amount mean n avg



Insights on Accounts

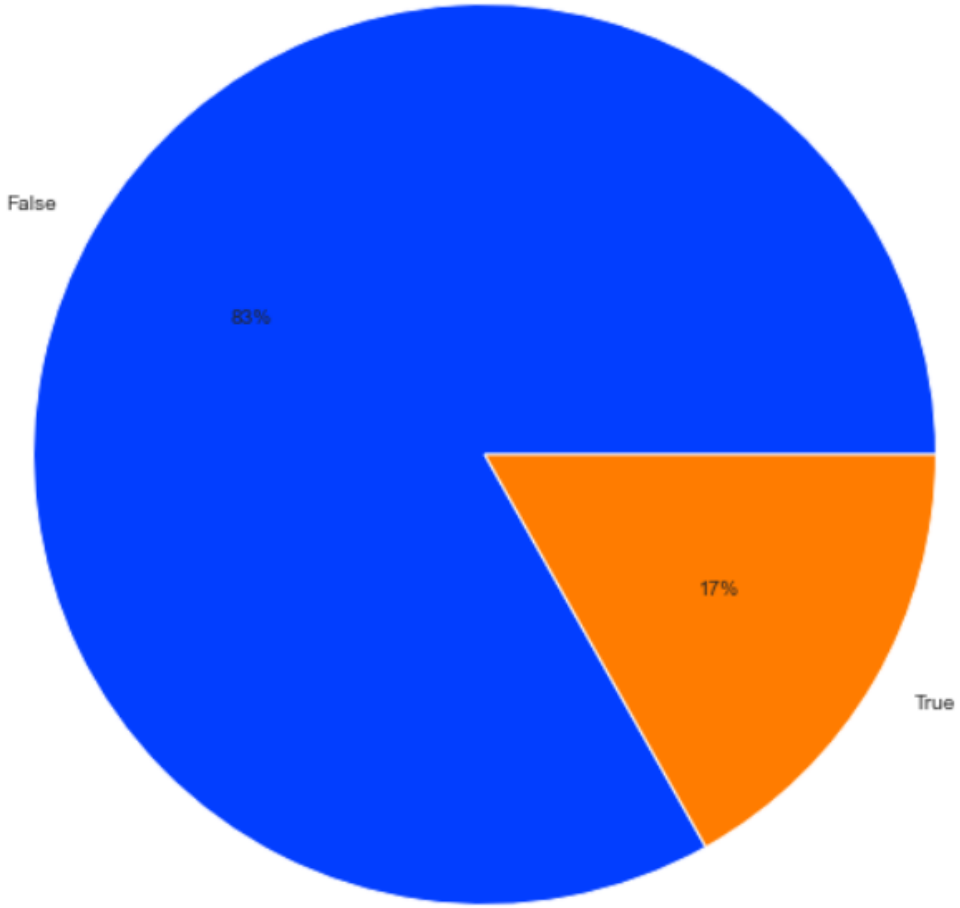
Distribution over Alert generated or not



Here we can say that more than 8315 transaction records did not generated any kind of alert while 1685 generated alert while getting transacted. according this info we can say that the risk of going through a fraud is less.

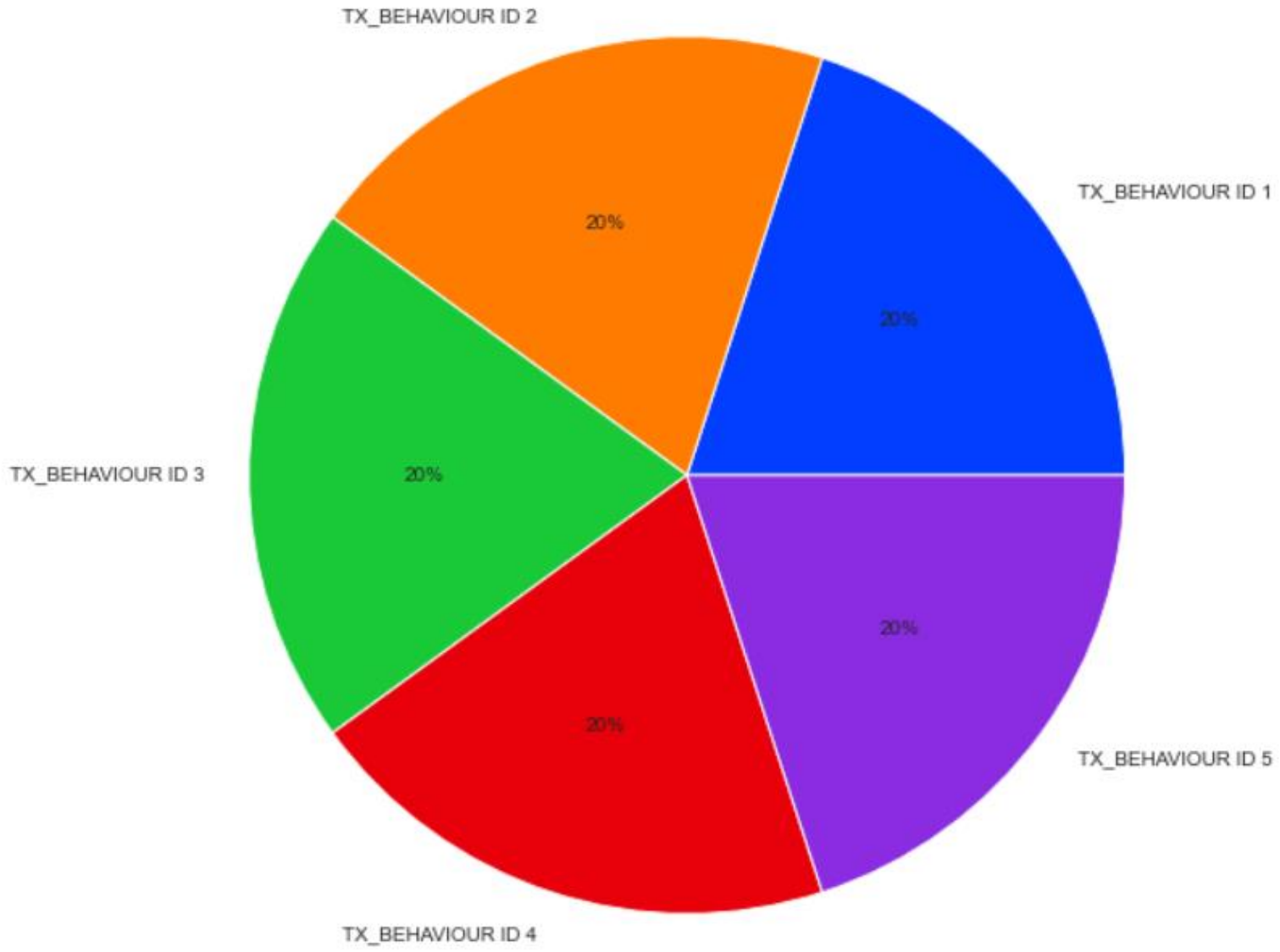
Committed and Attempted Analysis

There are only 17% chance that a transaction can encounter fraud rest 83% times its safe to transact amount from one account to another

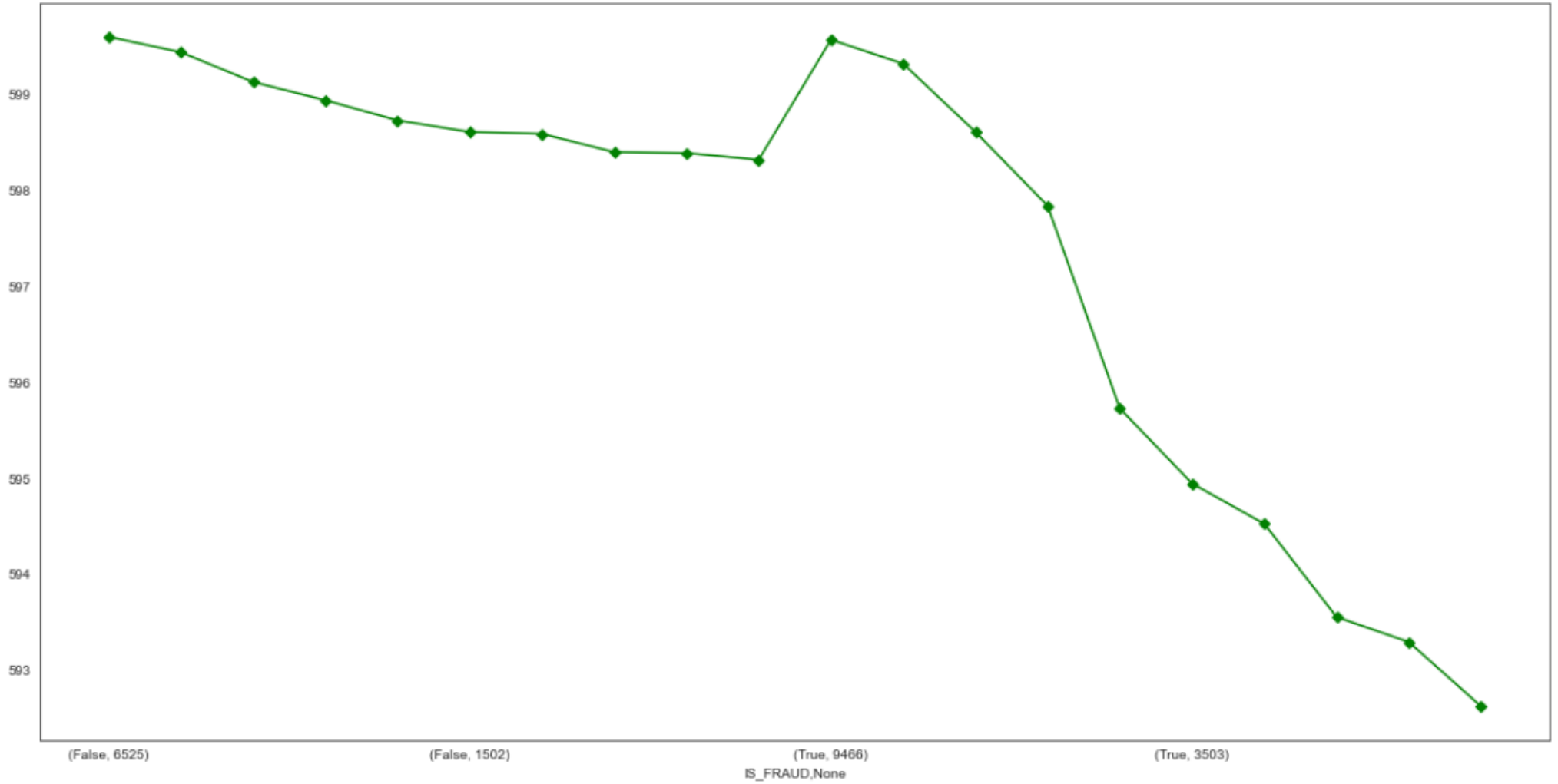


TX_BEHAVIOR analysis

According to insight we gained we noticed that each TX_ID is effecting in on data in a equally contributed manner. for over all data each TX_ID contributes 20%

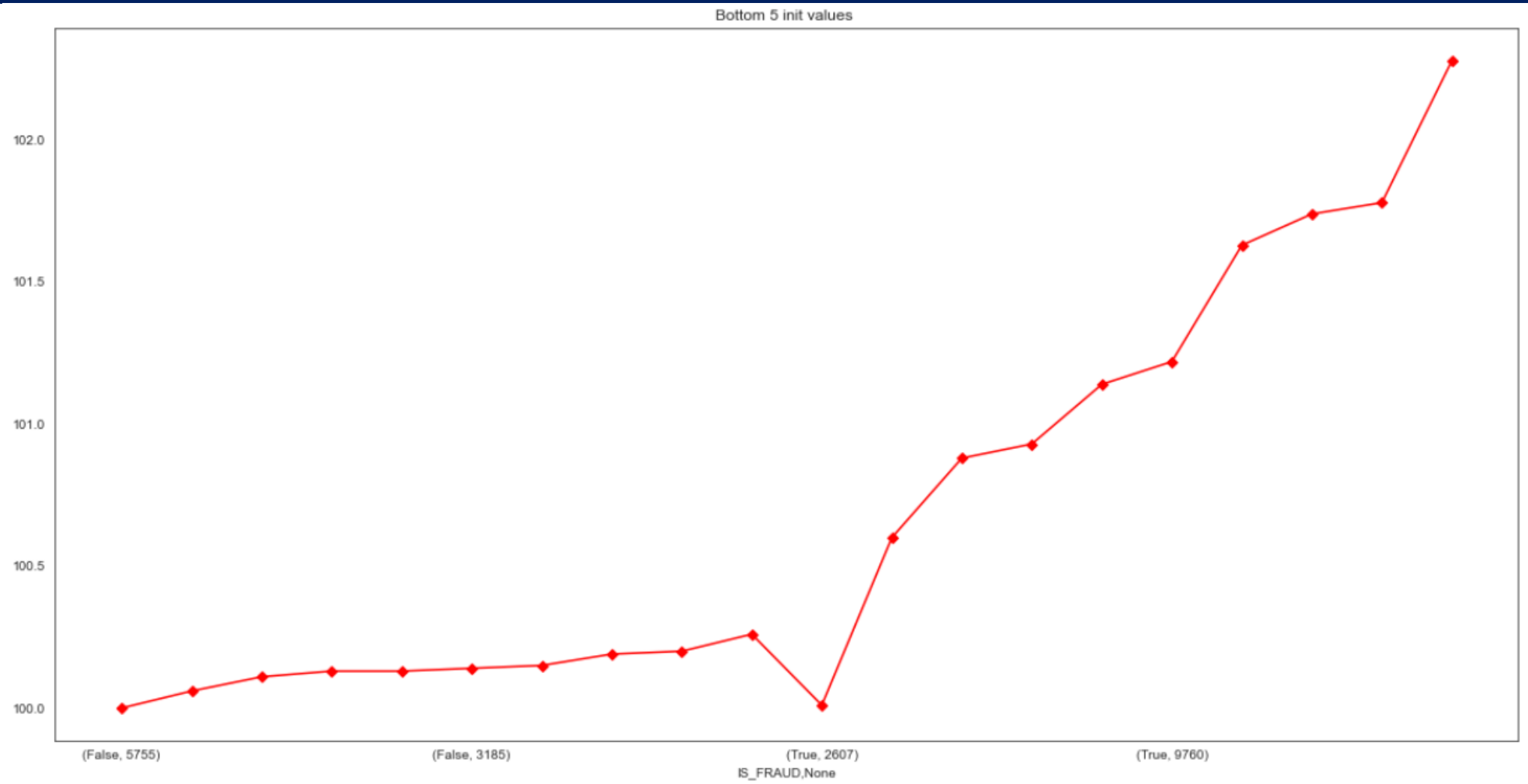


Analyzing pattern between fraud (true) and not Fraud (false)



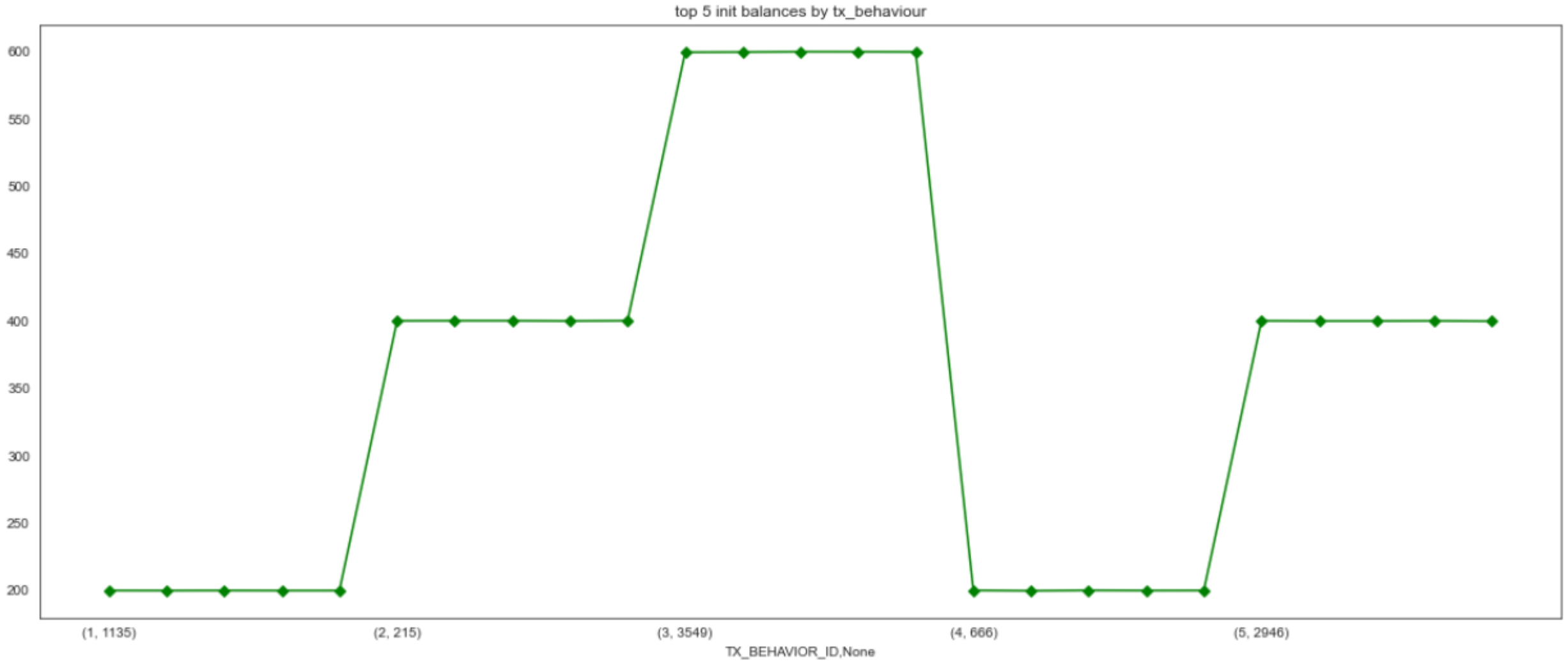
Those 20 points are plotted there from max in it balances those are top 10 of fraud alert is true and 10 from top alert is false. as we can see over here only top 4 points are having the approx same range of in it balance kept in account. where as fraud are more likely to happen when amount kept is less.

What are the chances that if i keep less in it balance then I am more likely to encounter a fraud?



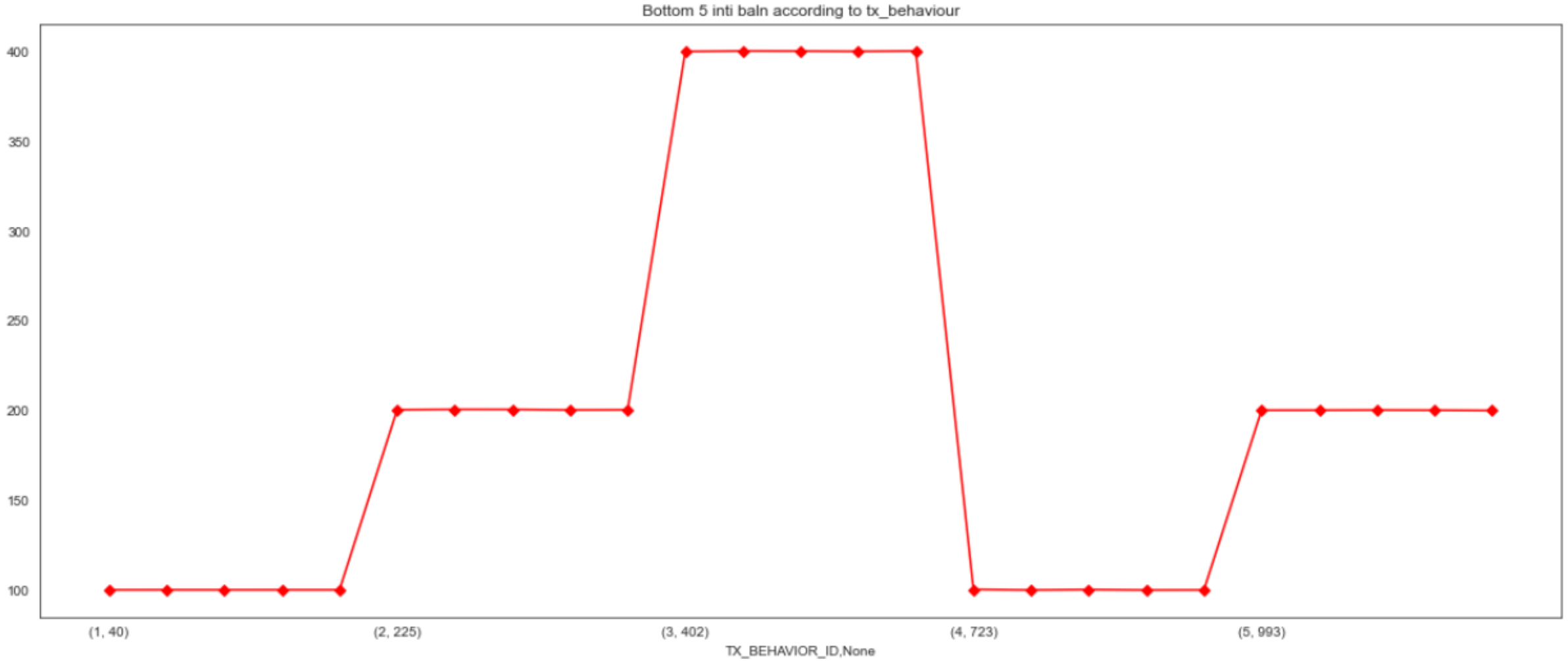
So we checked for min values of in it balances kept in the account and we came to know that as we go higher in init balance kept is the risk of getting fraud is also high. We can conclude that the amount above 100 and less 500 is more likely to get a fraud.

Checking pattern in top five largest values of INIT_BALANCE in each TX_BEHAVIOUR



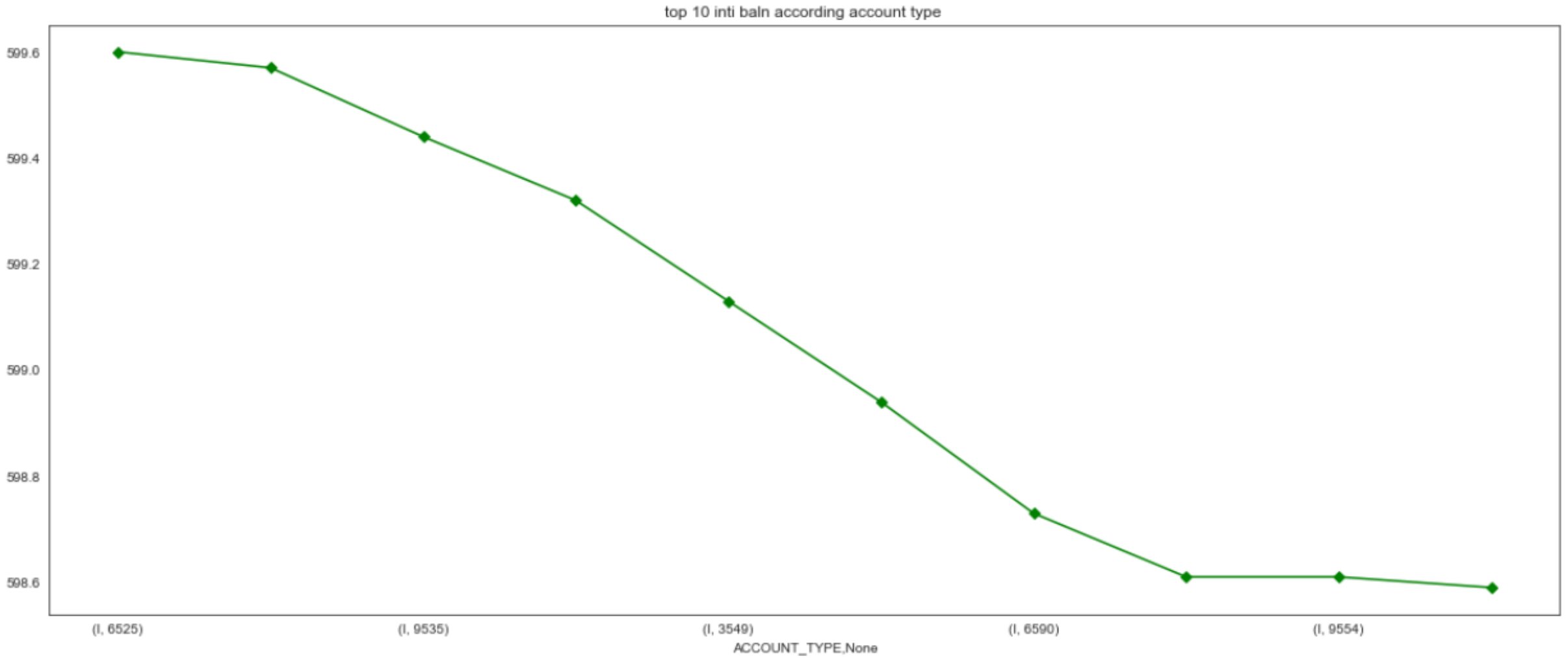
Checked behavior of each tx_id between tx_id 3 4 peoples keep highest in it balance in the account

Checking pattern in bottom five largest values of INIT_BALANCE in each TX_BEHAVIOUR



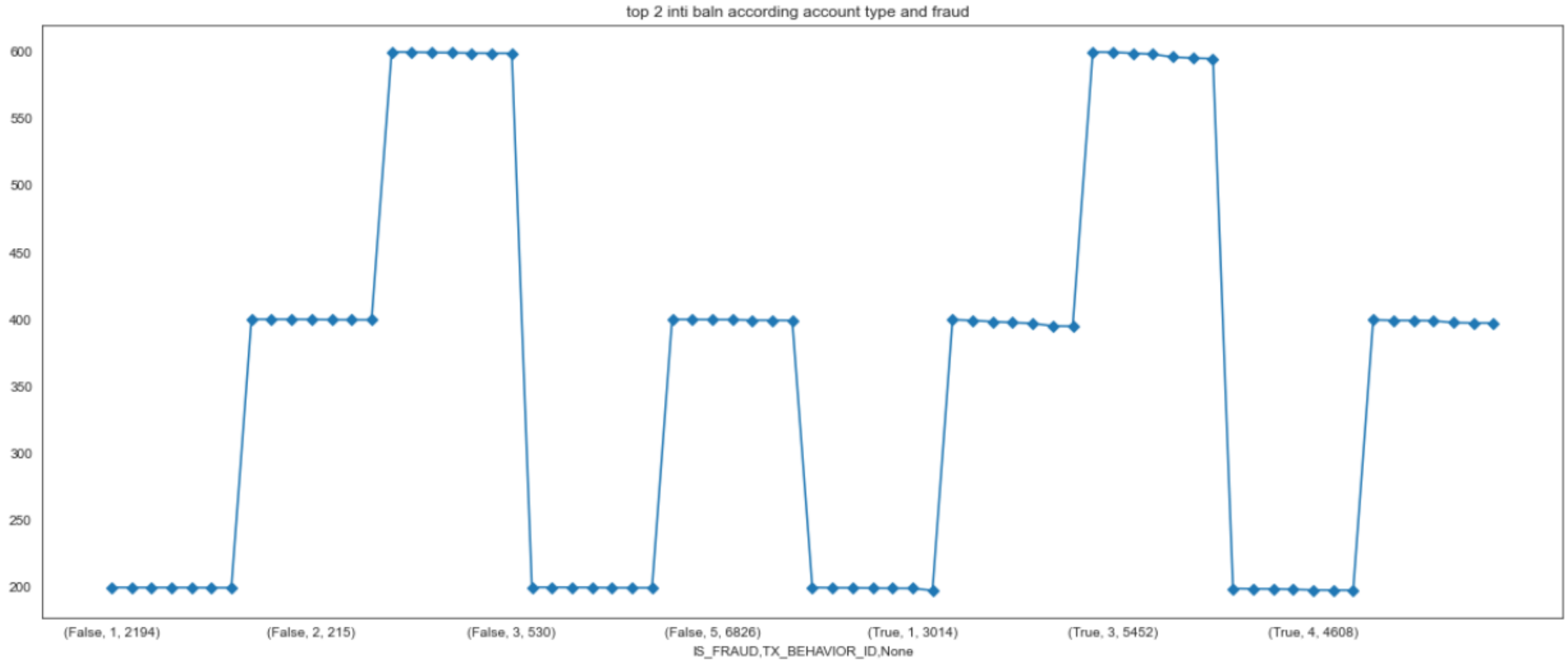
In the bottom values same pattern is generated where as the amount kept reduced and no of peoples gets reduced in it. There is no drastic change because of account type cause data provided is of single kind of account

Checking pattern in bottom five largest values of INIT_BALANCE in each TX_BEHAVIOUR



Checking this for top in it balances this is kind of interesting that in both conditions when an account go through fraud or it dose not still in both conditions the tx id 3 to 4 is at peek and they kind of posses same kind of pattern.

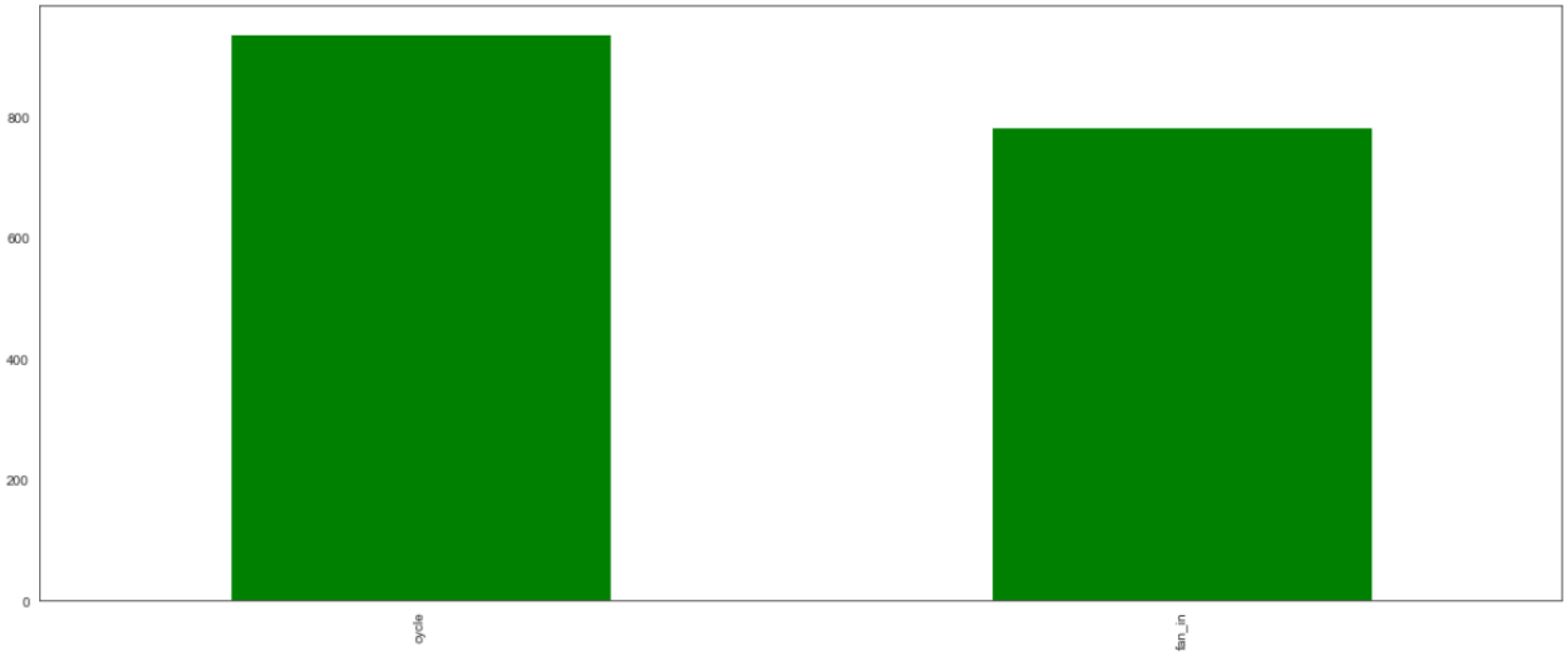
Checking pattern in bottom five largest values of INIT_BALANCE in each TX_BEHAVIOUR



While checking for bottom in in balances we get a different pattern from top values but get a symmetric pattern in between fraud and no fraud.

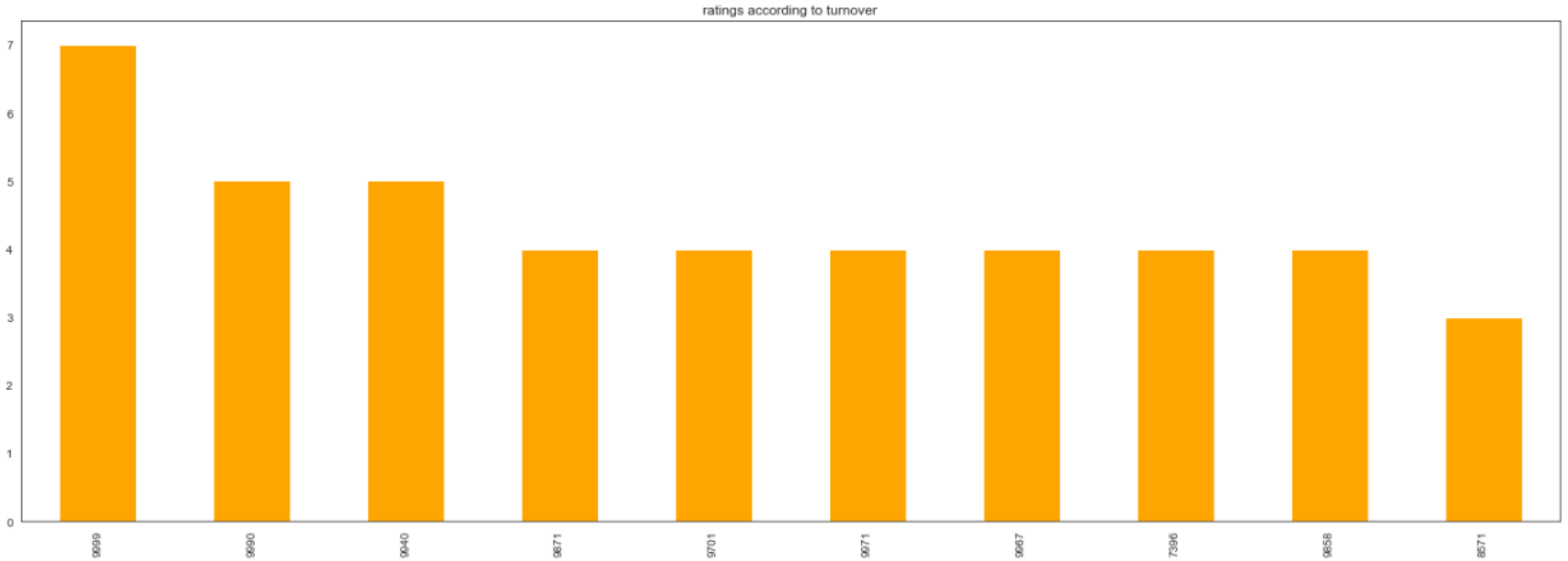
Insights on Alerts

Type of Alert



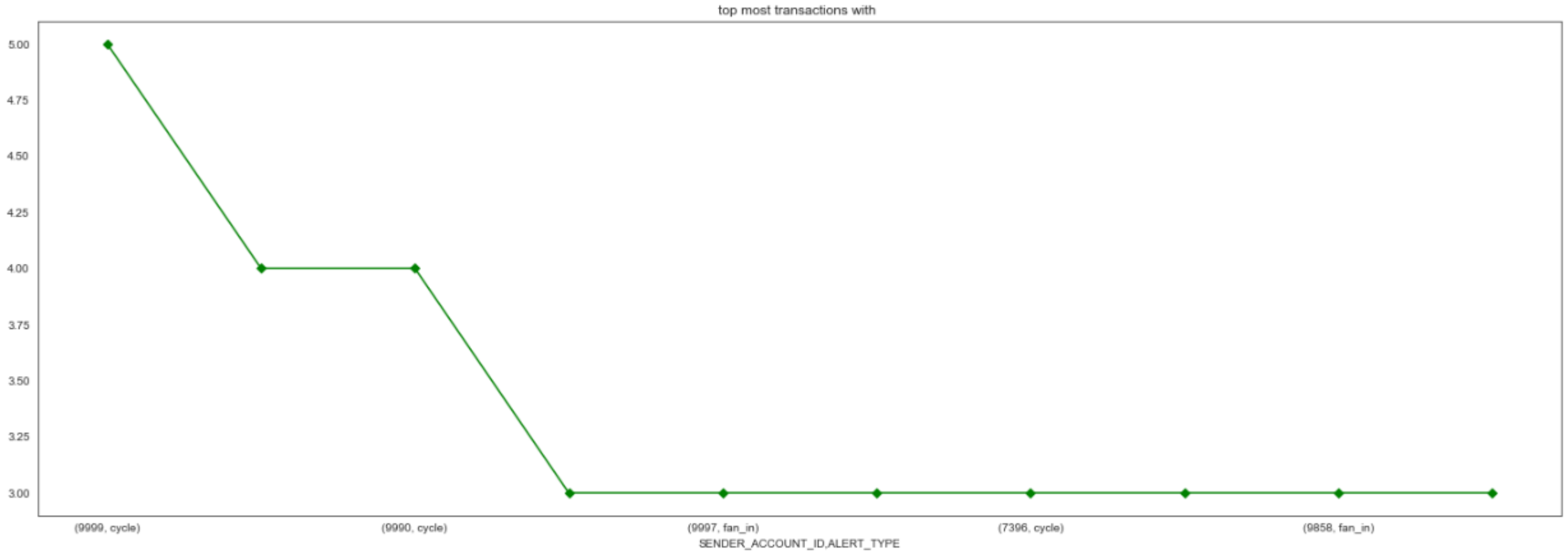
Cycle has the most distribution in type of alert.

Ratings according to turnover



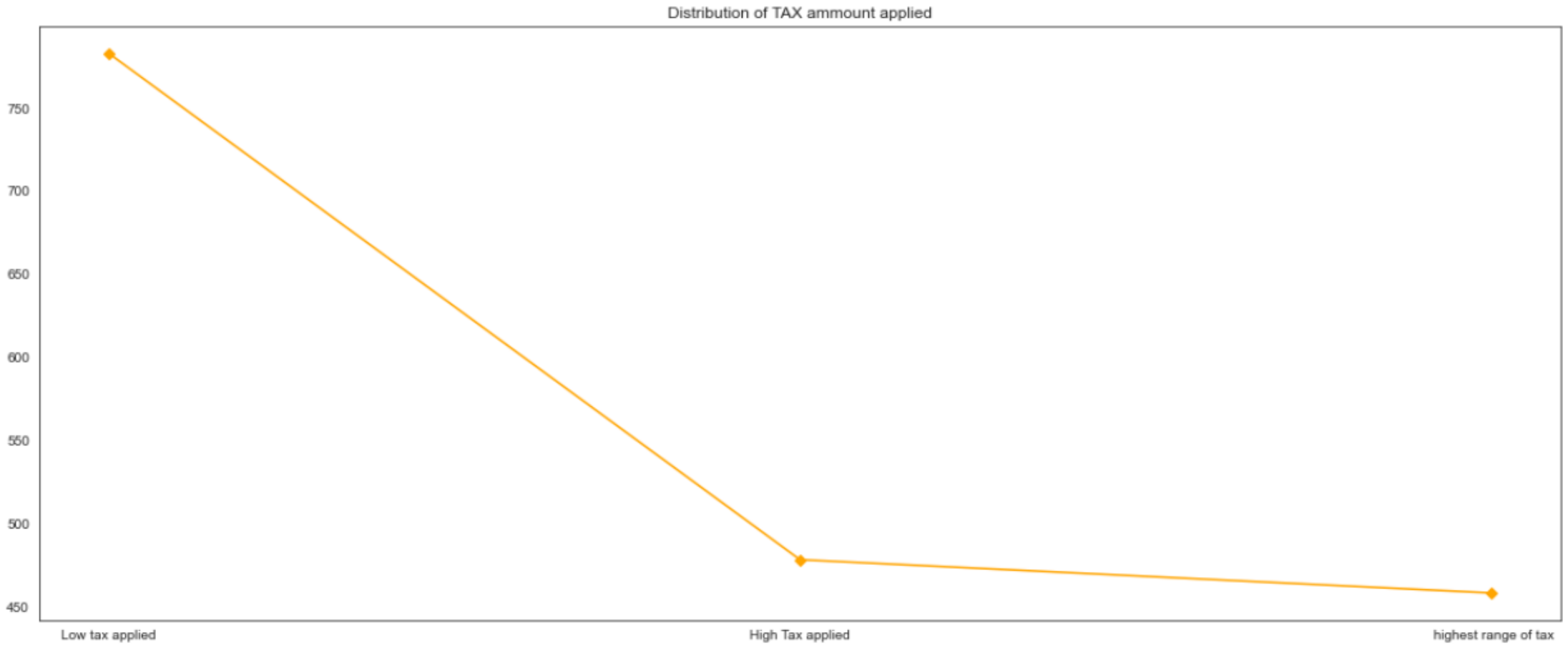
The highest rating is for the turnover of 9999.

Top most transactions



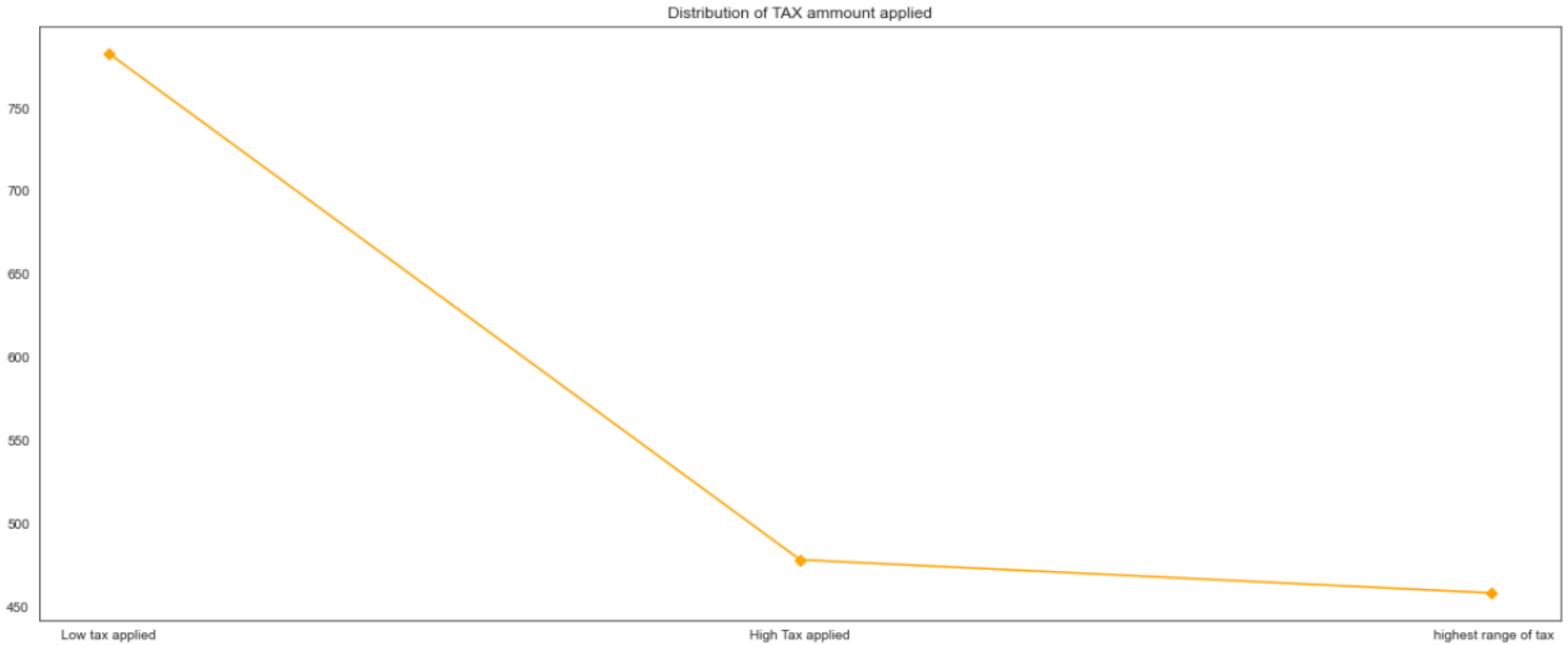
Can clearly say that all top transaction is getting alert type cyclic.

Ratings according to turnover



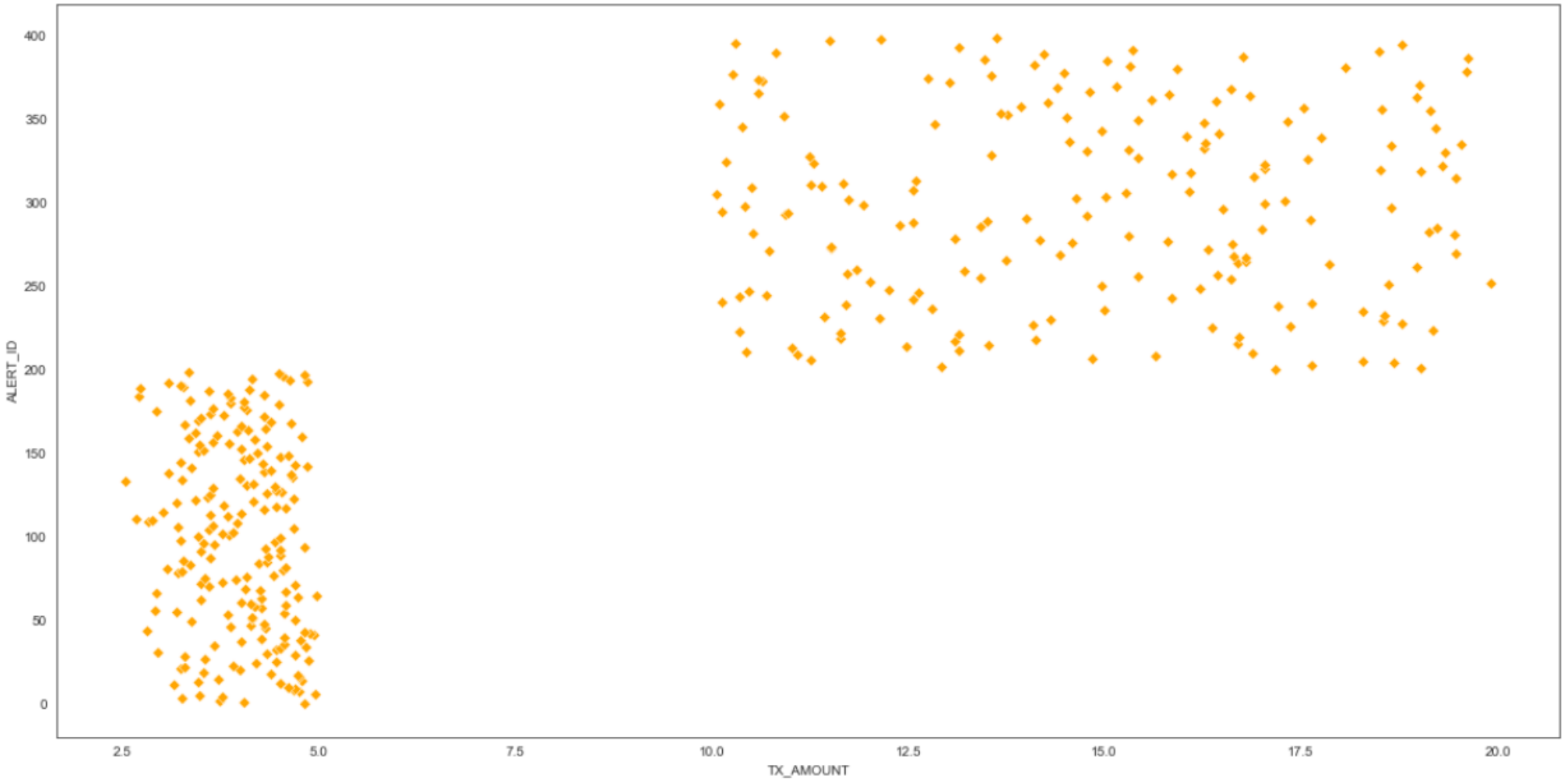
Most of the time the tax applied on the transaction is less than 5 as the tax applied on a transaction get higher the no of peoples doing that kind of transaction becomes lesser. where as from low to high amount of tax applied the no of people doing transaction changes with slop tan(-45)

Cluster according alert and TX applied



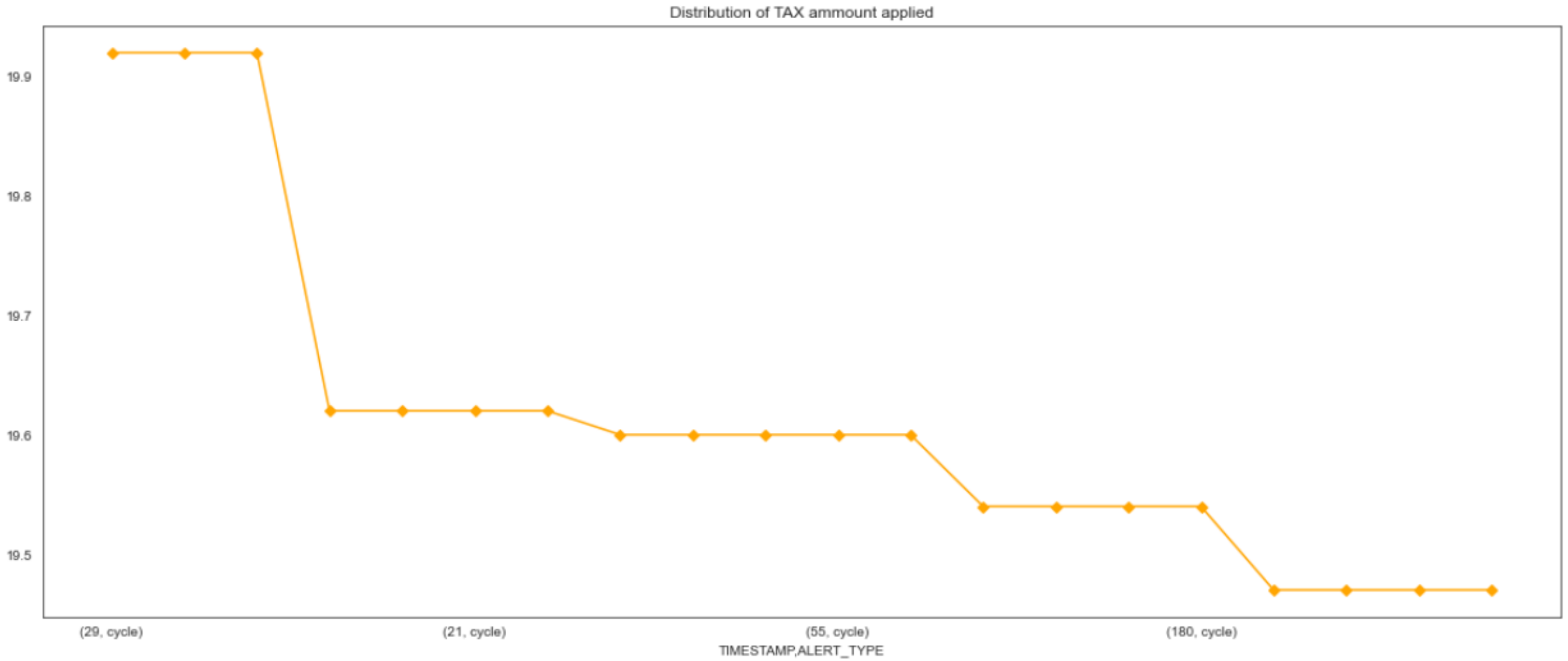
Most of the time the tax applied on the transaction is less than 5 as the tax applied on a transaction get higher the no of peoples doing that kind of transaction becomes lesser. where as from low to high amount of tax applied the no of people doing transaction changes with slop tan(-45)

Cluster according alert and TX applied



From this it is clear that the cluster with less amount of tax directly belongs to the fan_in type of alert and the cluster with more amount of tax belongs to the Cyclic type of alert

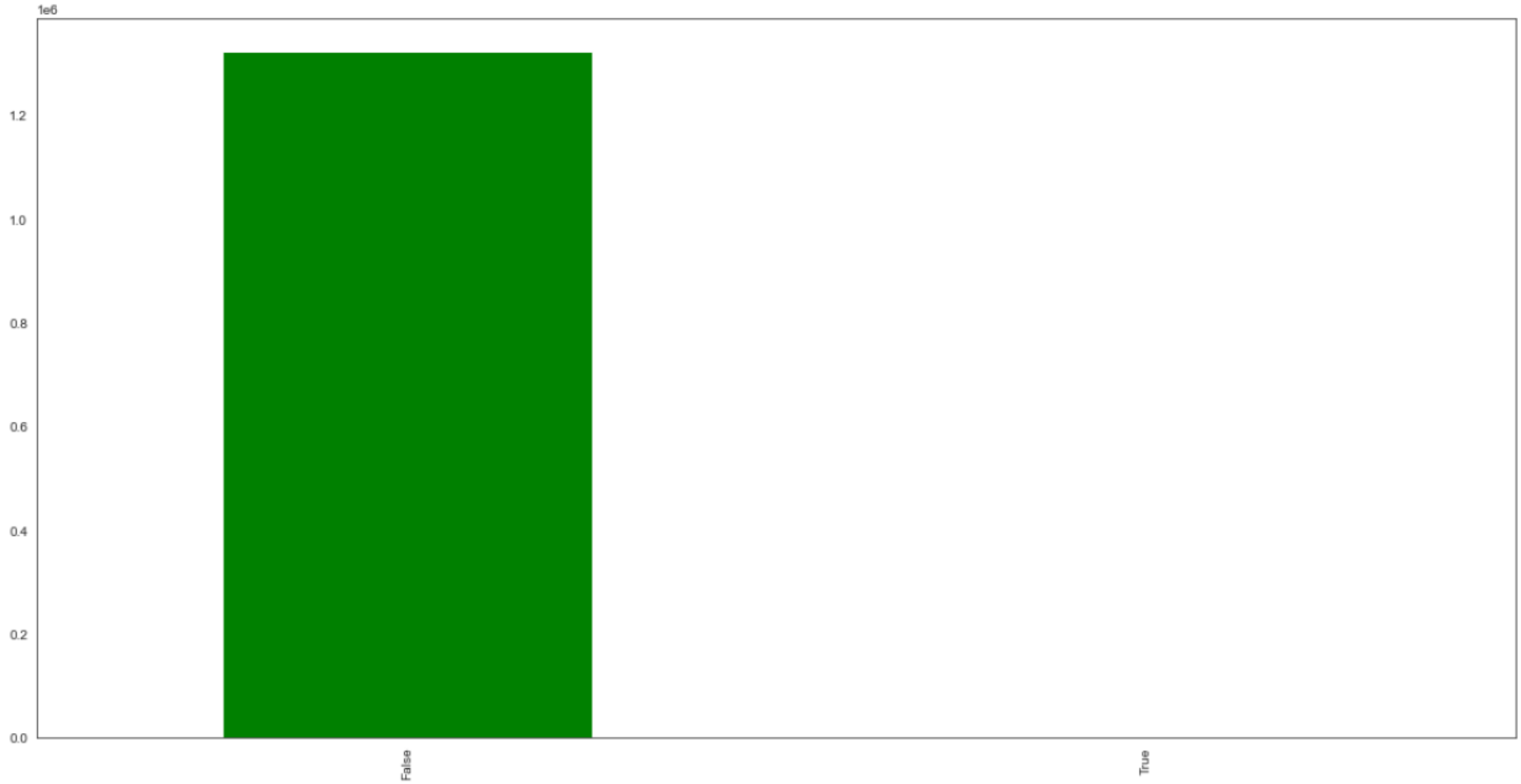
Top tx amount applied for a particular time stamp and what type of alert was there



When ever tx amount applied high transaction gets the alert type of cyclic checked top 20 transactions and each time alert type shown is cyclic.

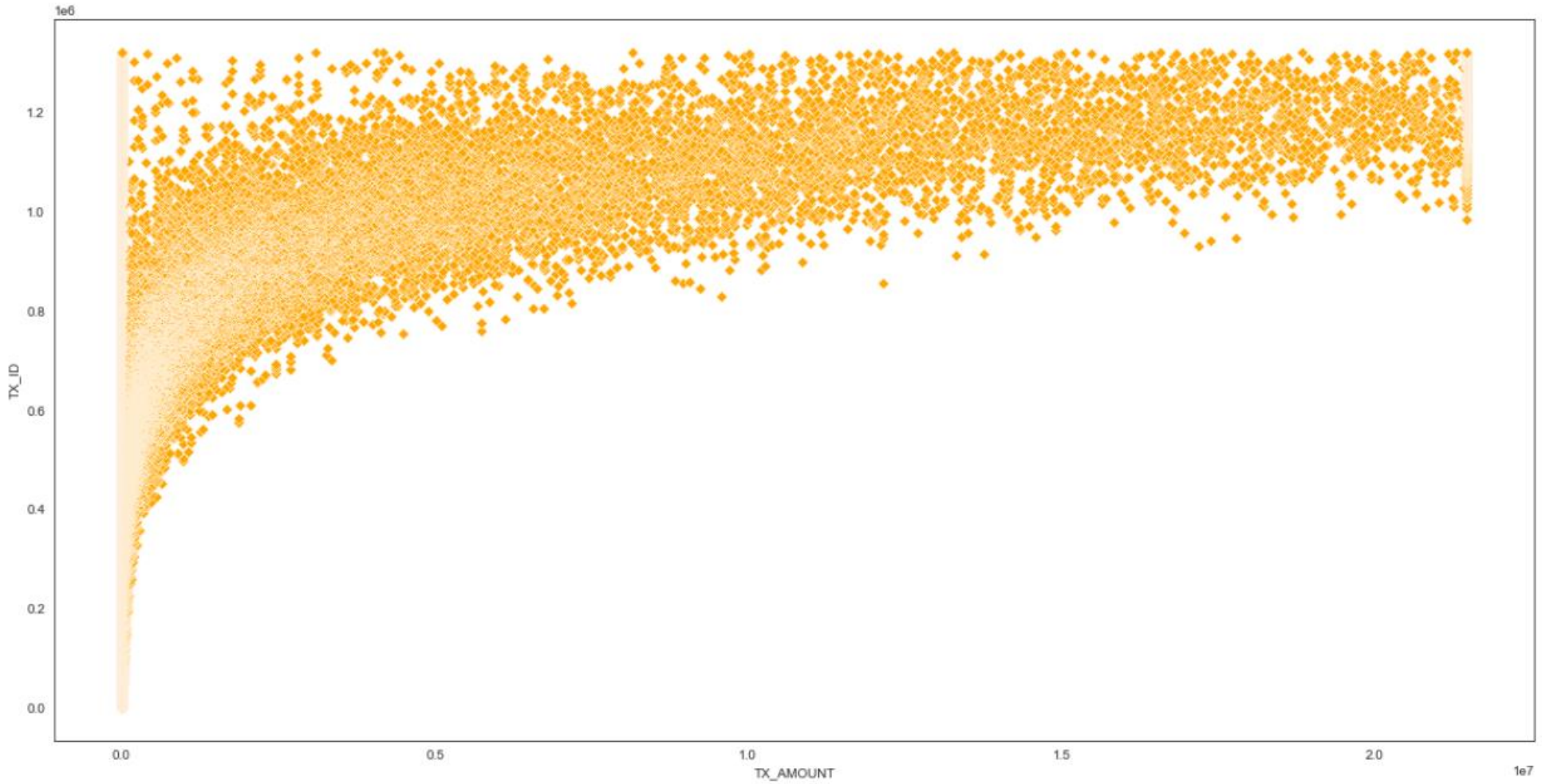
Insights on Transactions

There is no true value of fraud



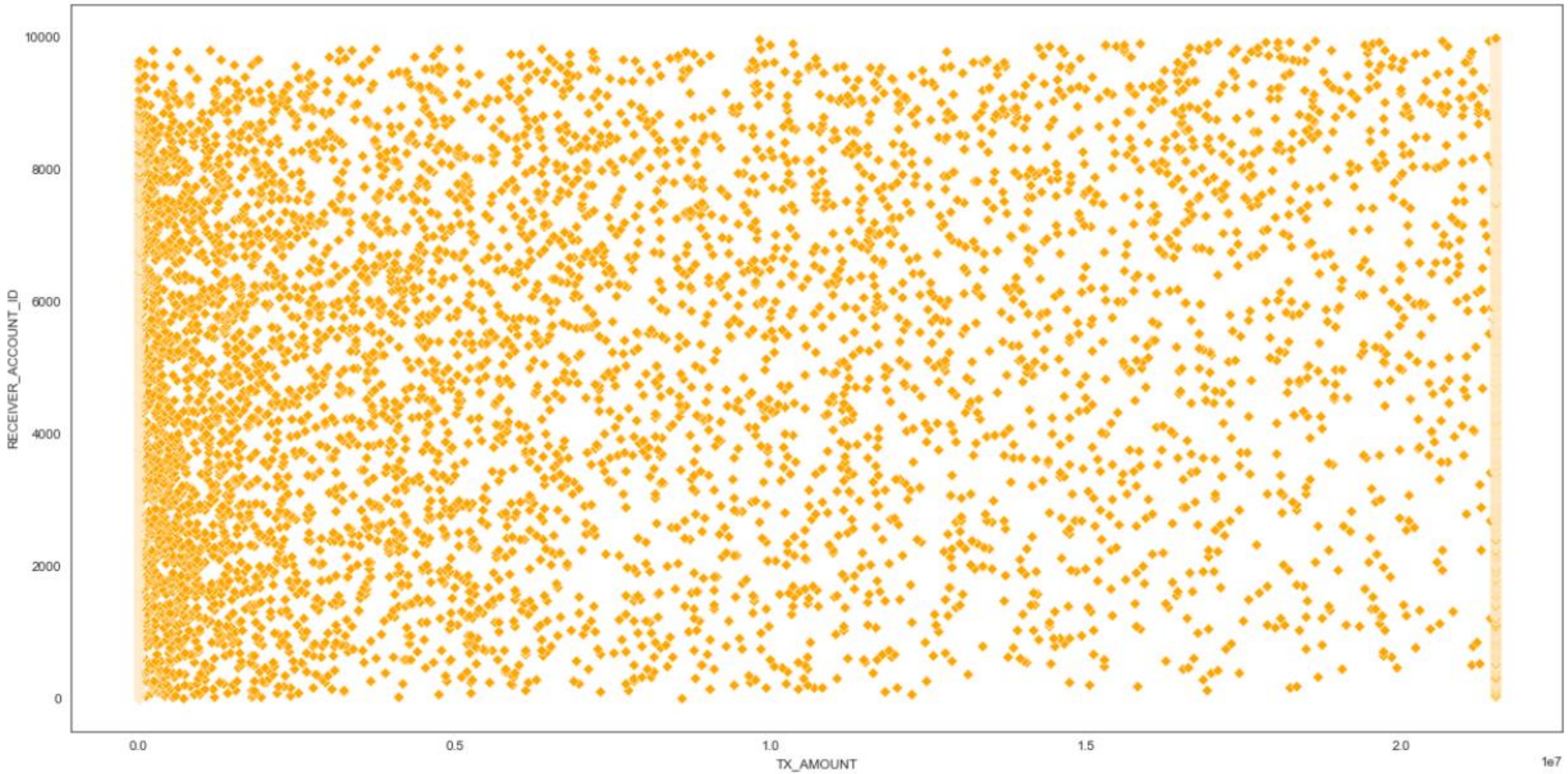
That meanse there is no fraude detection over here checking for the top largest transaction values there dose not exist any fraud detection so everything is clear

There is no true value of fraud



It is clear that tax amount is showing an exponential nature over tx_id.

There is no true value of fraud



As we can see as the no of tax amount is less no of transactions received is more as the tax amount grew the no of transactions becomes less.

THANK YOU !!