



# MARKET SEGMENTATION

A Machine learning based market  
segmentation project.

PROJECT BY: Rakib Mathakiya



**ABOUT PROJECT**



# WHAT IS MARKET SEGMENTATION?

## Market Segmentation

In marketing, market segmentation is the process of dividing a broad consumer or business market, normally consisting of existing and potential customers, into sub-groups of consumers based on some type of shared characteristics.



## OBJECTIVE

A case requires to develop a customer segmentation to give recommendations like saving plans, loans, wealth management, etc. on target customers groups.

# OBJECTIVE





**ABOUT DATASET**



# Dataset

The sample Dataset summarizes the usage behavior of about 9000 active credit card holders during the last 6 months. The file is at a customer level with 18 behavioral variables.

## Variables of Dataset

- Balance
- Balance Frequency
- Purchases
- One-off Purchases
- Installment Purchases
- Cash Advance
- Purchases Frequency
- One-off Purchases Frequency
- Purchases Installments Frequency
- Cash Advance Frequency
- Cash Advance TRX
- Purchases TRX
- Credit Limit
- Payments
- Minimum Payments
- PRC Full payment
- Tenure
- Cluster

A large green arrow pointing to the left, centered horizontally on a dark gray background. The arrow has a white text label inside it.

**TOOLKIT AND ALGORITHMS**



# TECH TOOLKITS USED

## Python

Python is the programming language which is used to code the project. Various algorithms used are coded in Python



## Streamlit

Streamlit is an open source framework for Machine learning and Data Science. It is used for prototype and deployment purpose of the models.



## Jupyter Notebook

Jupyter notebook is the effective IDE used for coding in Python. It is very easy to use and widely used over the industry.



# Algorithms used in the project





# **STEPS AND PROCESS**

# PROJECT PROCESS

1

## Research and business understanding

The first thing you have to do before you solve a problem is to define exactly what it is. You need to be able to translate data questions into something actionable.

2

## Data pre-processing

Data preprocessing can refer to manipulation or dropping of data before it is used in order to ensure or enhance performance, and is an important step in the data mining process.

3

## Exploratory Data analysis

Exploratory data analysis is an approach of analyzing data sets to summarize their main characteristics, often using statistical graphics and other data visualization methods.

4

## Model Building

Model building process where different machine learning algorithms are used to make different machine learning models for various applications.

5

## Model Deployment

Model Deployment is the process where various ML algorithms are deployed on various platforms like flask, streamlit, various open source platforms, etc. Here we have used Streamlit to deploy our ML project.



**INSIGHTS GENERATED**

# OUTPUT SCREEN

app · Streamlit

localhost:8501

Apps edX New Tab Business Name Gen... GitHub - Ishtiyag18... ATH | Analytica Tre... All Modules - Statis... Online Statistics Ed... Statistics for Machi... Binomial And Poiss... Other bookmarks Reading list

☰

## Prediction

Balance

0.000000 - +

Balance Frequency

0.000000 - +

Purchases

0.00 - +

OneOff\_Purchases

0.00 - +

Installments Purchases

0.00 - +

Windows Taskbar: Windows Start, Task View, Edge, File Explorer, Task Manager, VS Code, PC, Clock, WhatsApp, Telegram, Signal, WhatsApp, Word, PowerPoint, Excel, Outlook, Teams, Spotify, Zoom, Photoshop, Premiere Pro

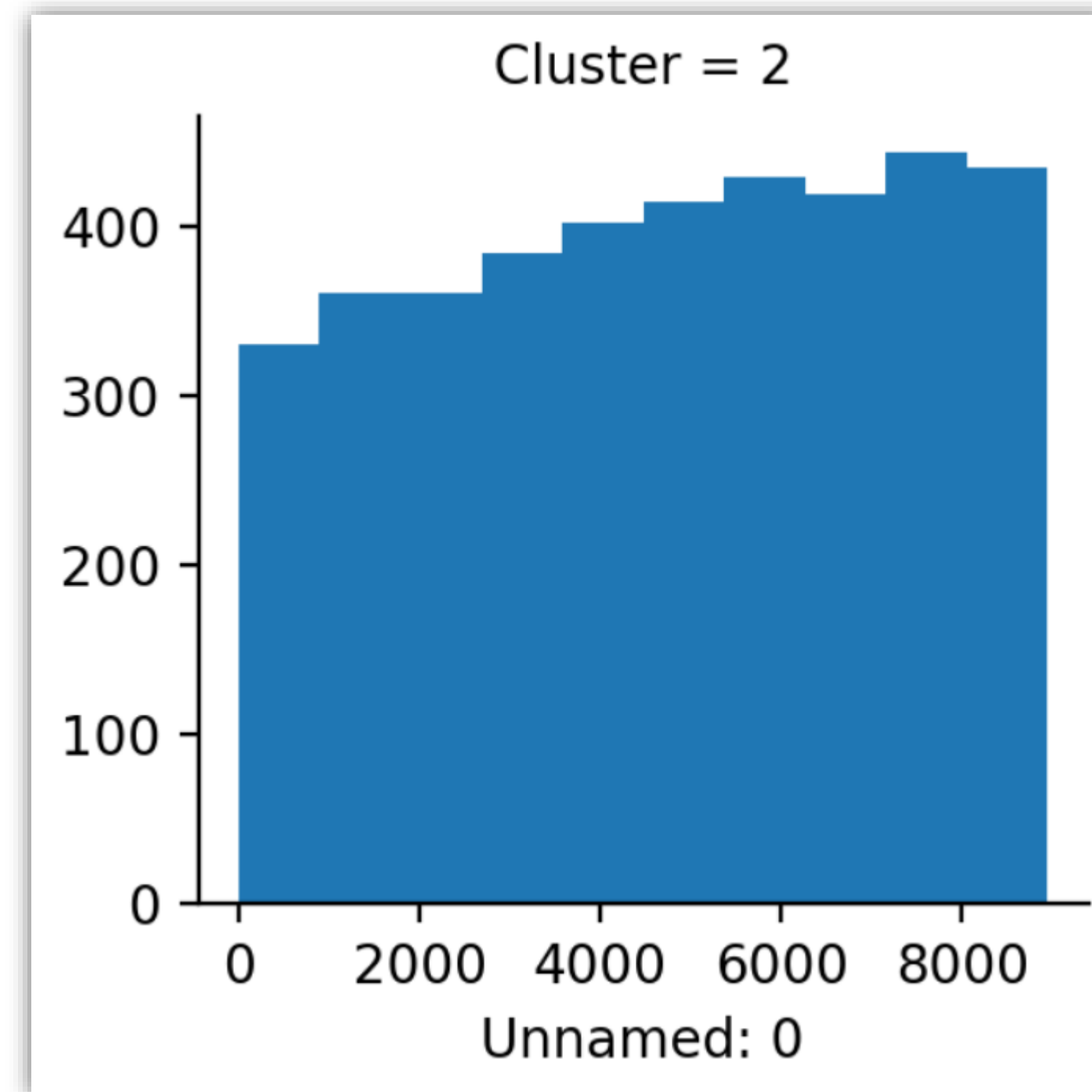
System Tray: Network, Volume, ENG 18:59

## INPUTS

- Balance – 40.900749
- Balance Frequency – 0.818182
- Purchases – 95.40
- One-off purchases – 0.00
- Installment Purchases – 95.40
- Cash Advance – 0.000
- Purchases Frequency – 0.166667
- One-off Purchases Frequency – 0.00
- Purchases Installments Frequency – 0.083333
- Cash advance frequency – 0.000
- Cash Advance TRX – 0
- Purchases TRX – 2
- Credit Limit – 1000
- Payments – 201.802084
- Minimum Payments – 139.509787
- PRC Full Payments – 0
- Tenure – 12

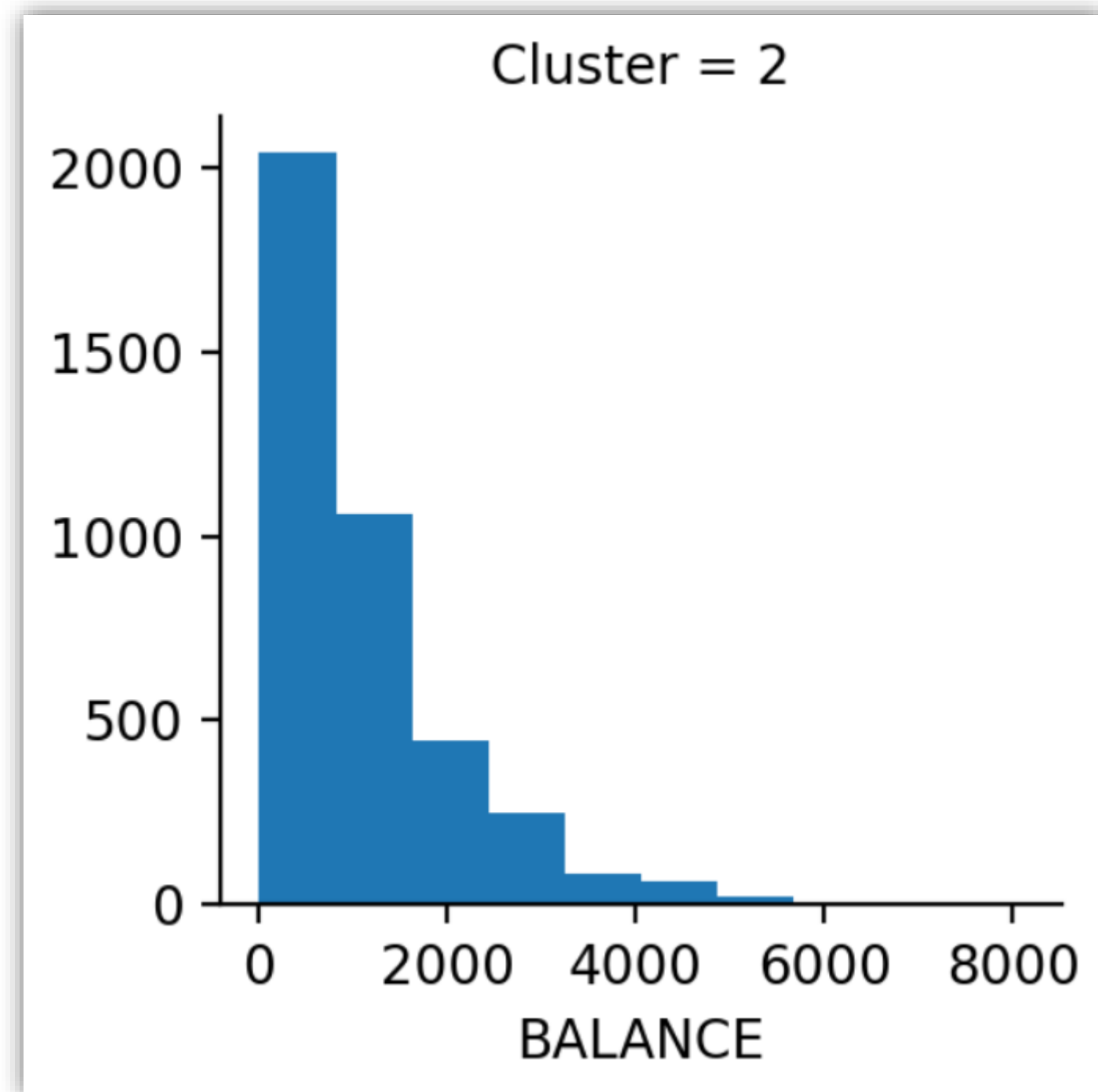
# OVERALL ANALYSIS

The visual shows the overall analysis and it belongs to Cluster 2 and the highest frequency is for 7000 and lowest for 0.

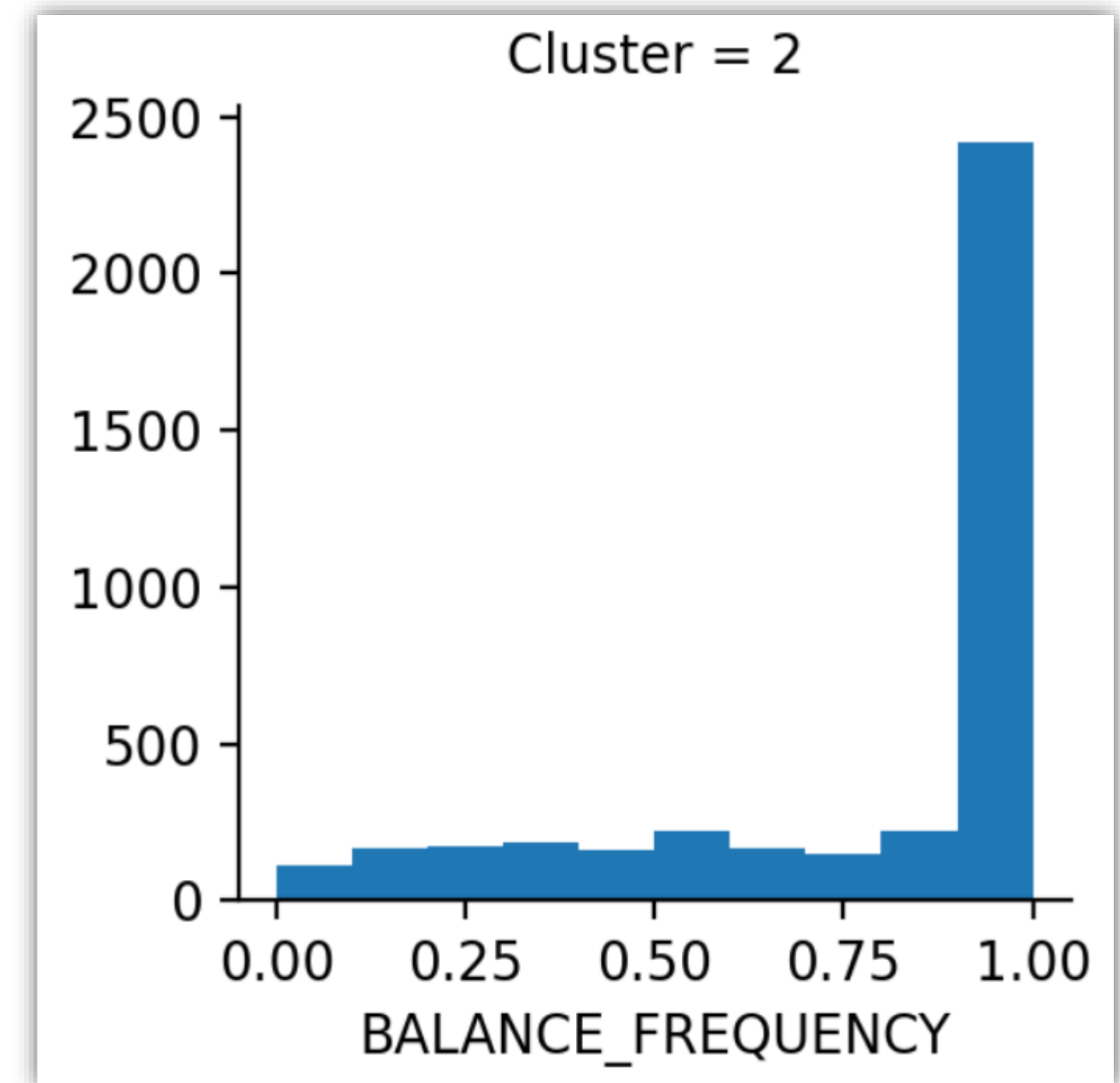




# BALANCE & BALANCE FREQUENCY

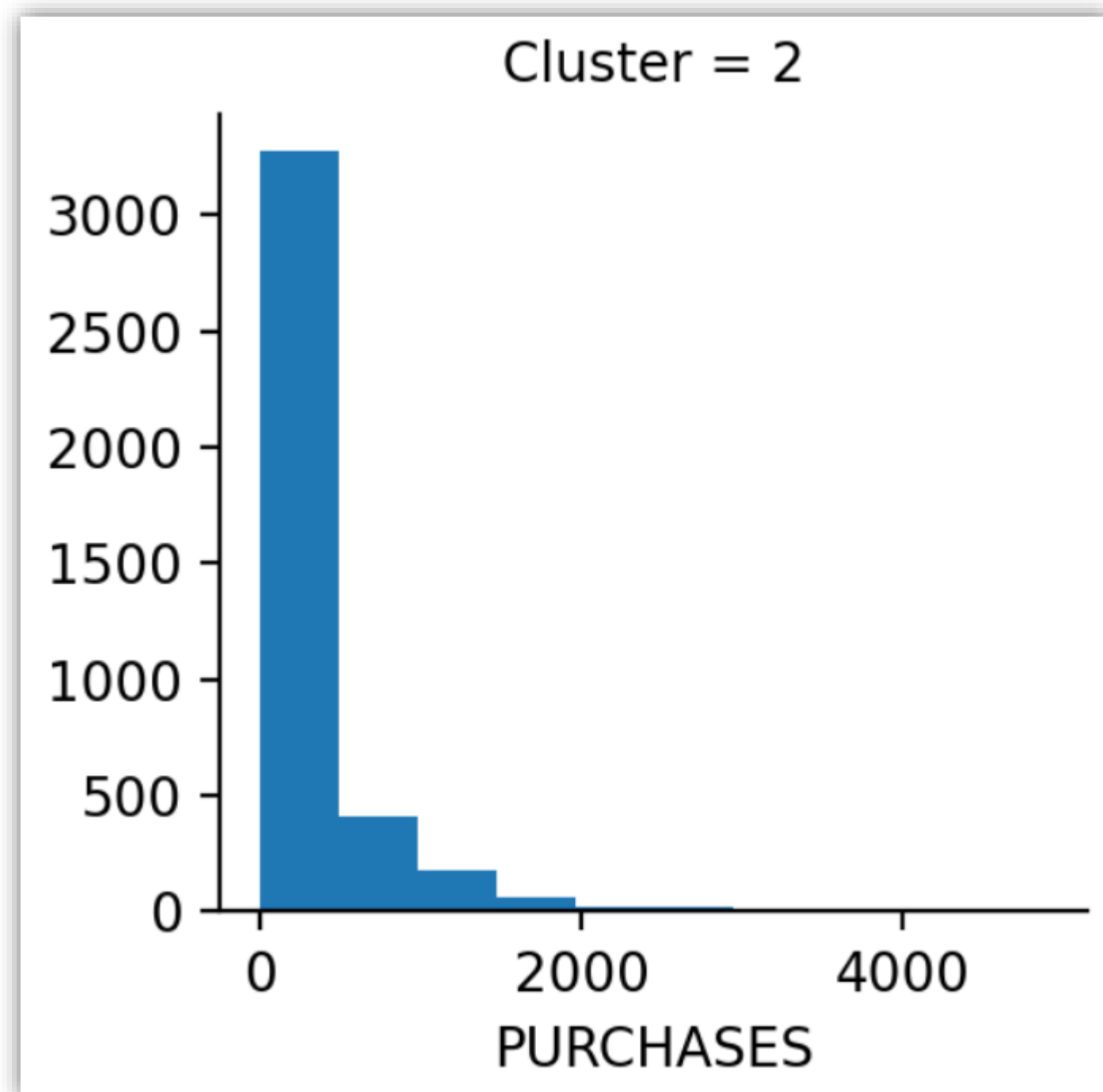


The visual shows that it belongs to Cluster 2 and the highest balance is 2000.

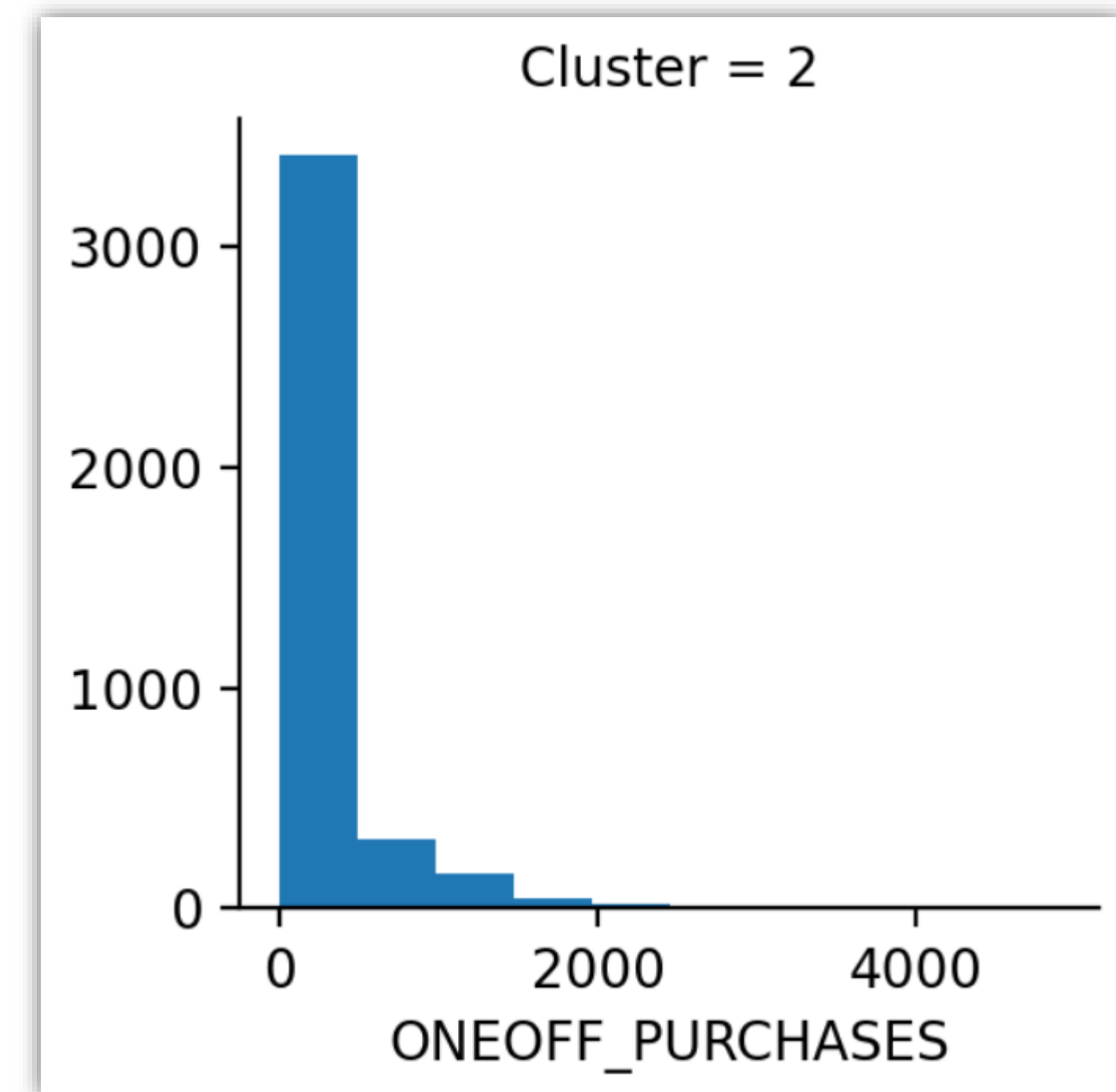


The visual shows the Balance Frequency which has highest frequency for 1.00 and lowest for 0.75

## PURCHASES & ONE-OFF PURCHASES

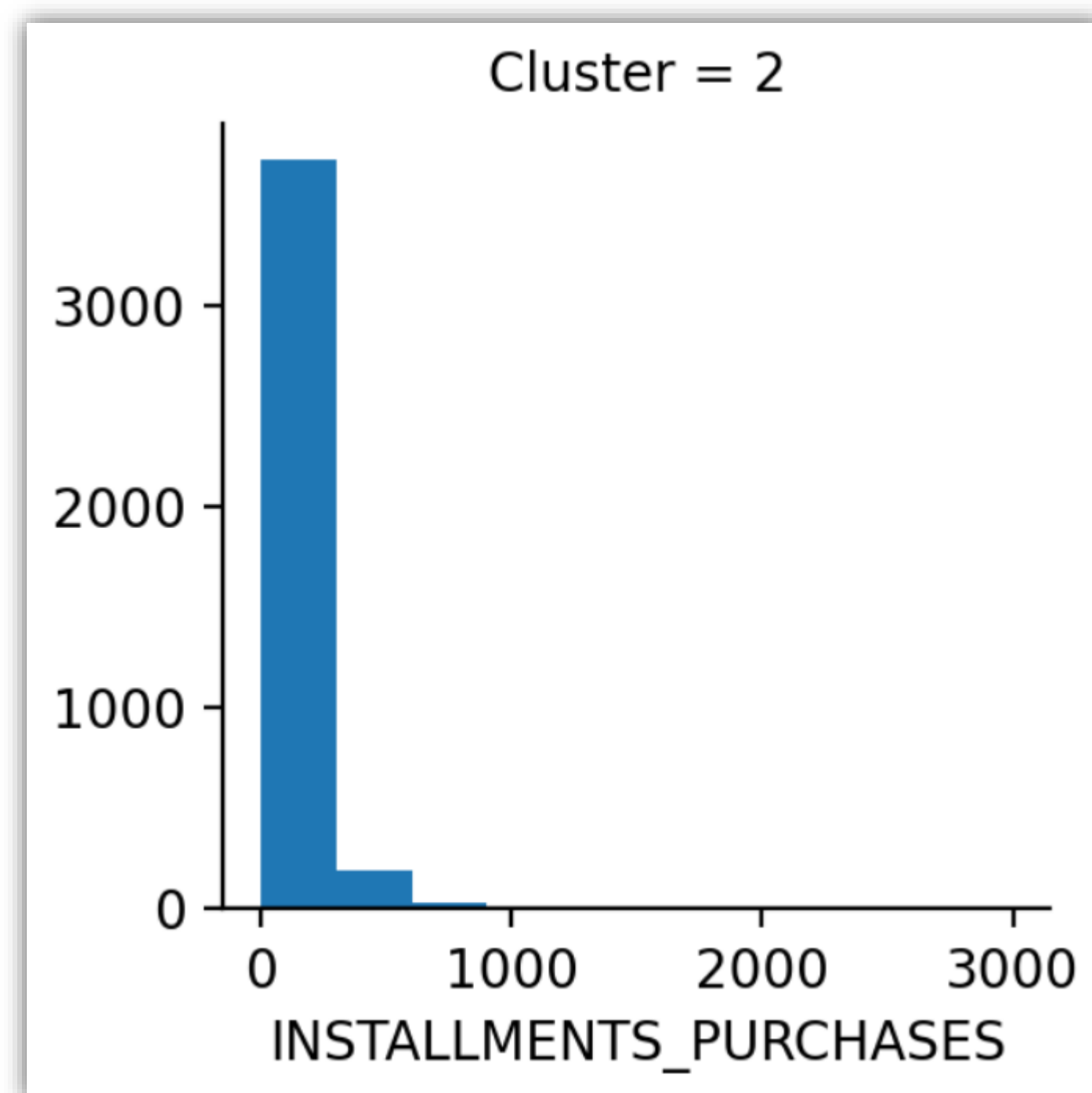


The visual shows Purchases that it belongs to Cluster 2 and has highest frequency of 3000.

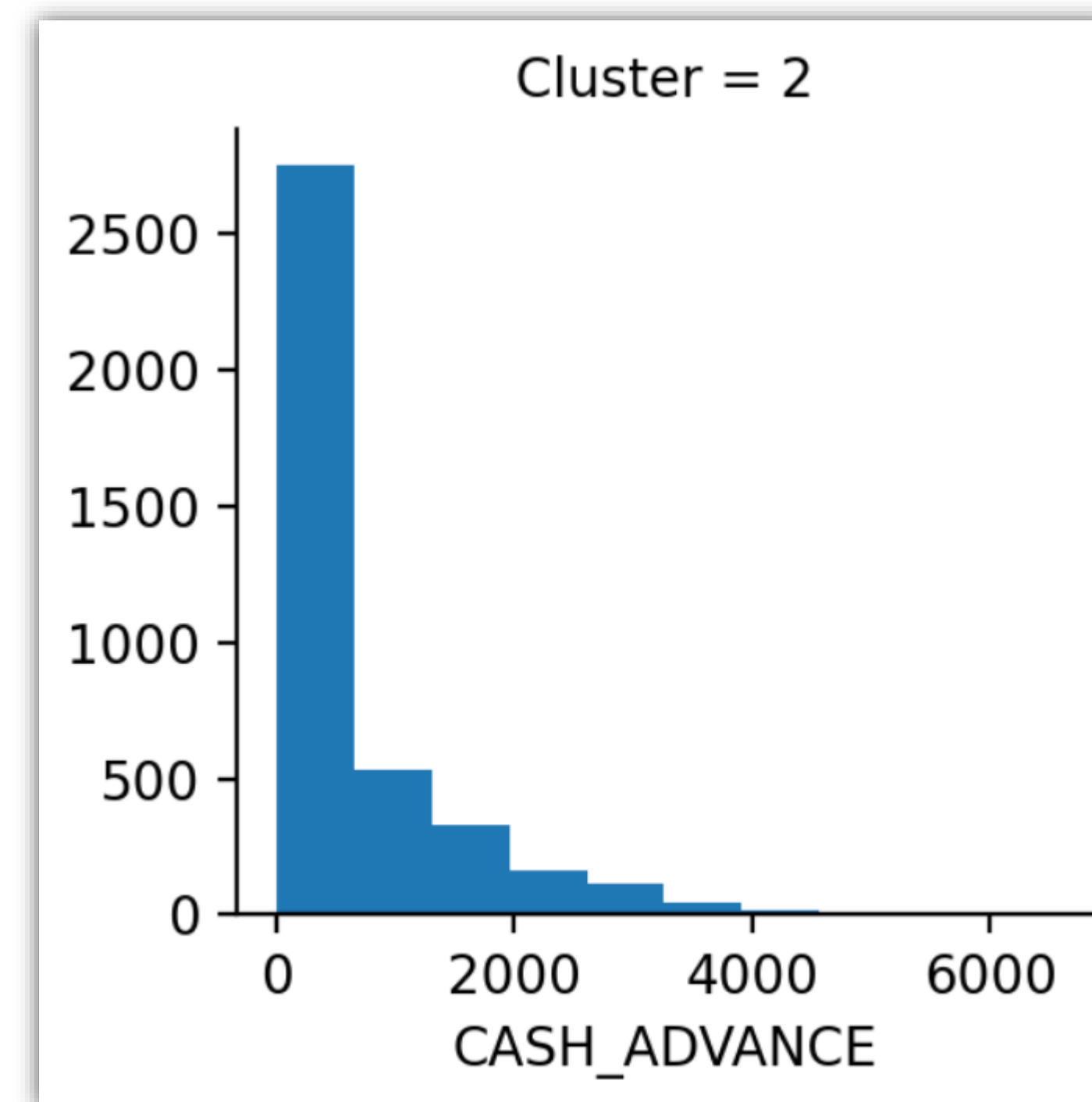


The visual shows the One-off Purchases which has highest frequency of 3000.

# INSTALLMENTS PURCHASES & CASH ADVANCE

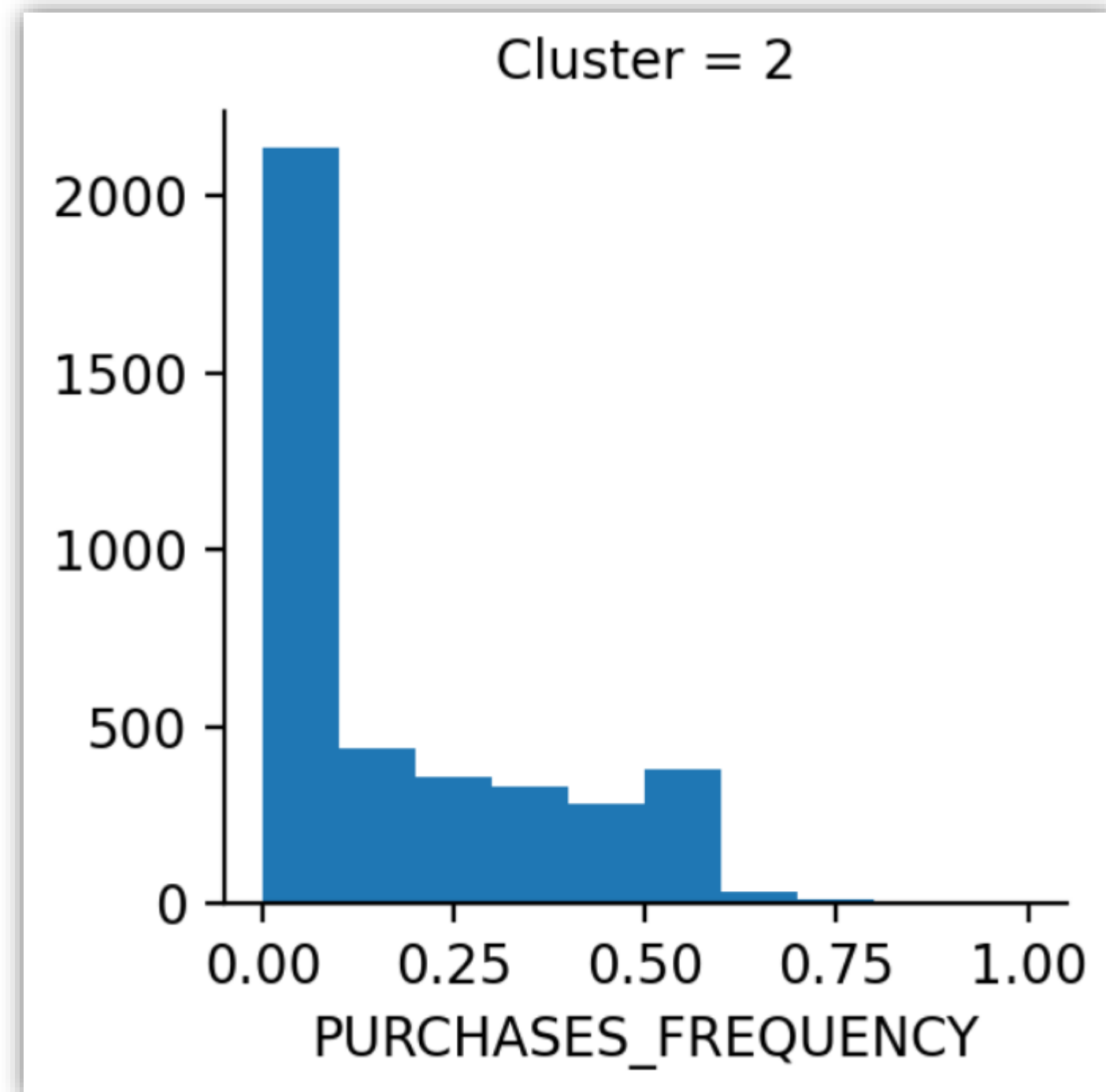


The visual shows Purchases that it belongs to Cluster 2 and has highest frequency of 3000.

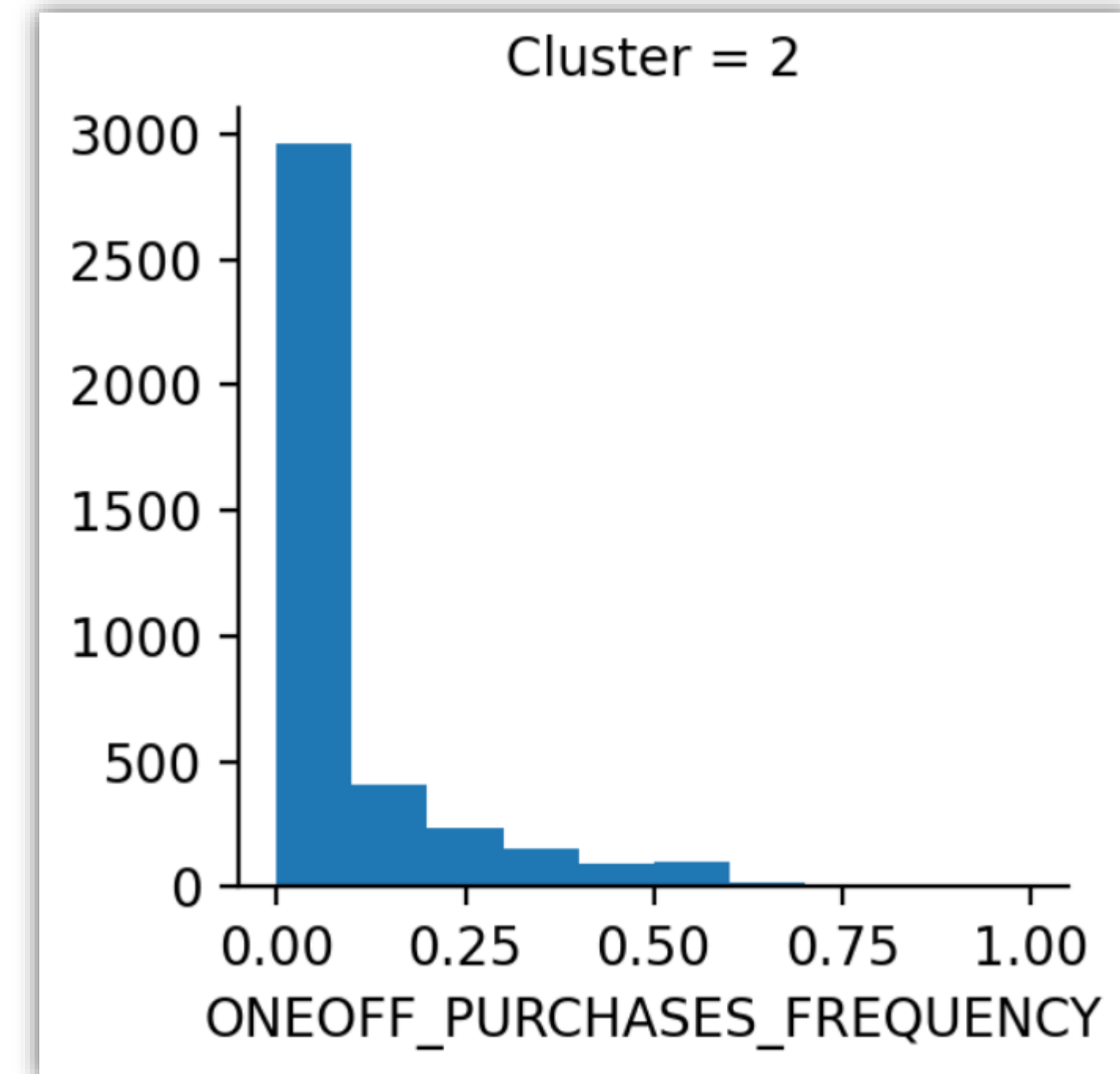


The visual shows the Cash Advance that belongs to Cluster 2 and has highest frequency of 2500.

# PURCHASES FREQUENCY & ONE-OFF PURCHASES FREQUENCY

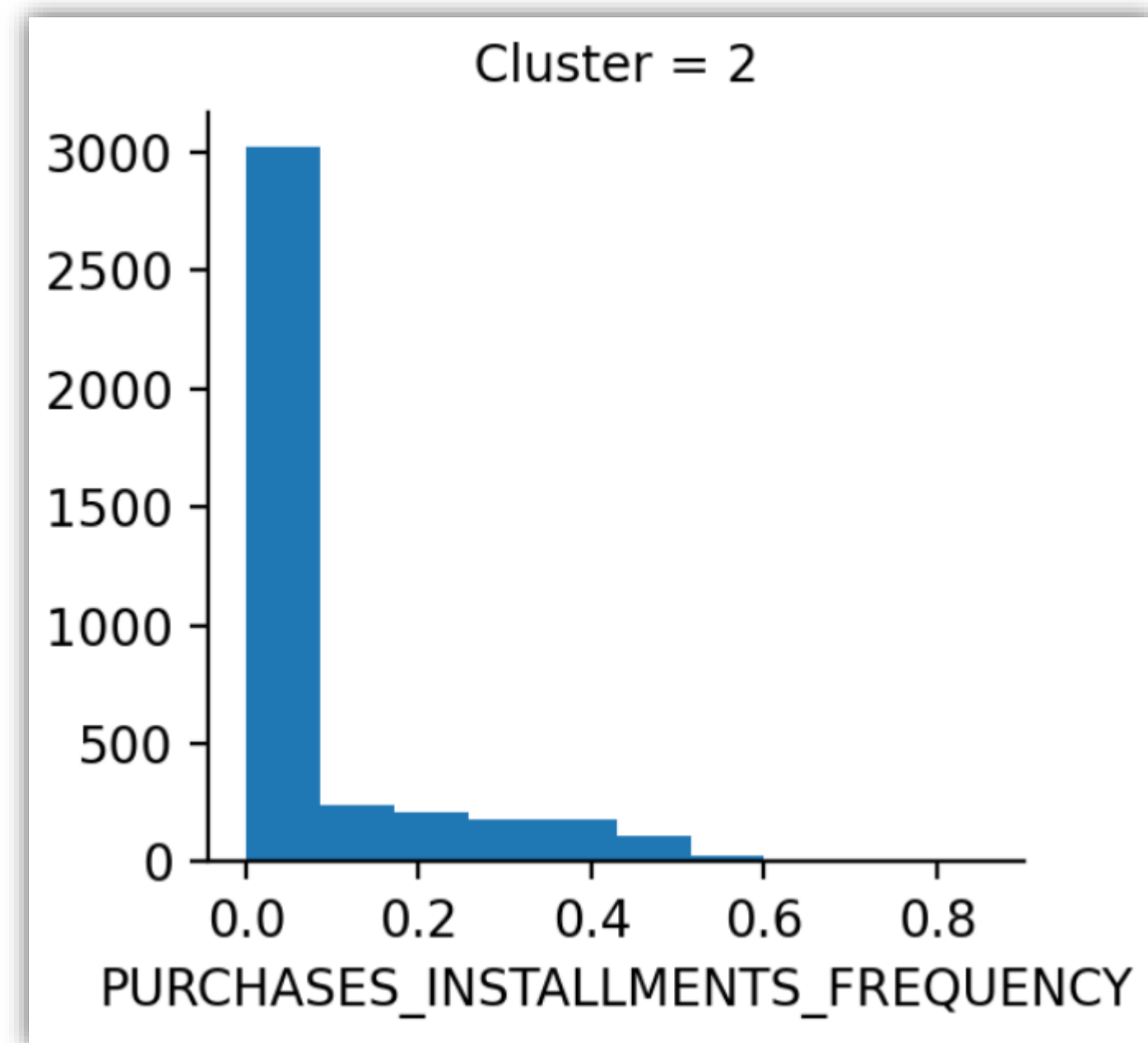


The visual shows Purchases Frequency that it belongs to Cluster 2 and has highest frequency of 2000.

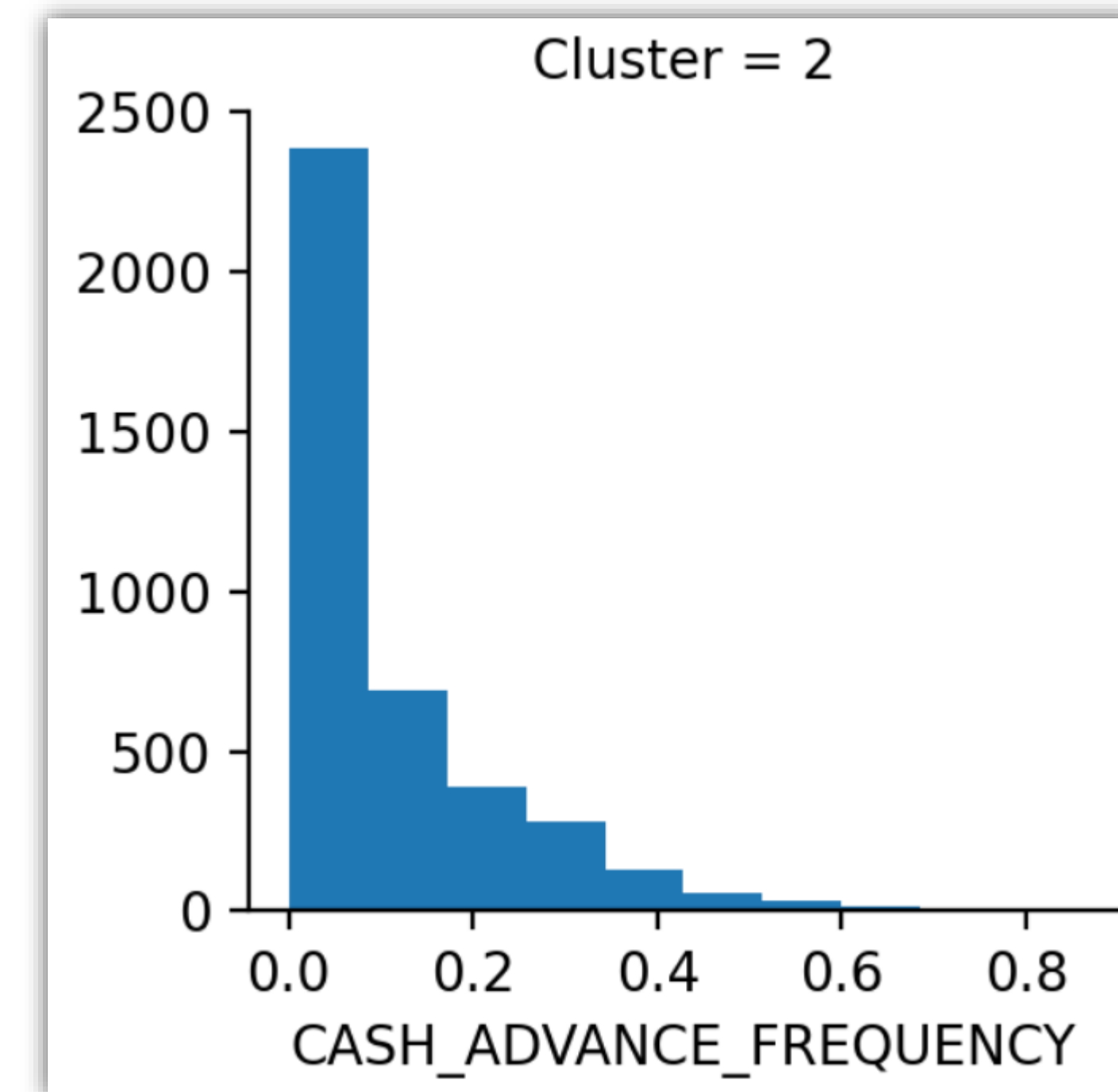


The visual shows the One-off Purchase Frequency which has highest frequency of 3000.

# PURCHASES INSTALLMENTS FREQUENCY & CASH ADVANCE FREQUENCY

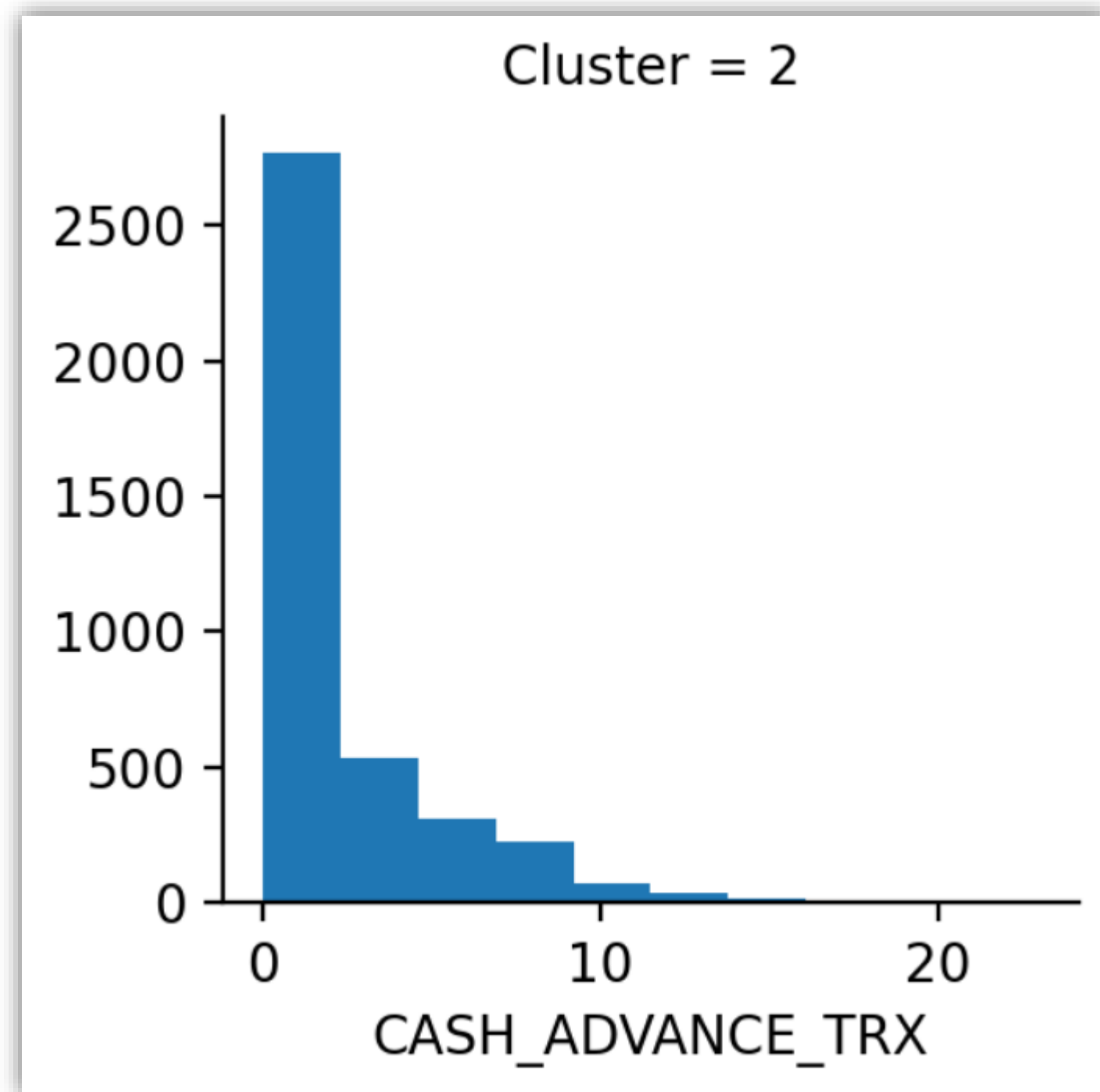


The visual shows Purchases Installments Frequency that it belongs to Cluster 2 and has highest frequency of 3000.



The visual shows the Cash Advance Frequency which belong to Cluster 2 which has highest frequency of 2500.

# CASH ADVANCE TRX & PURCHASES TRX

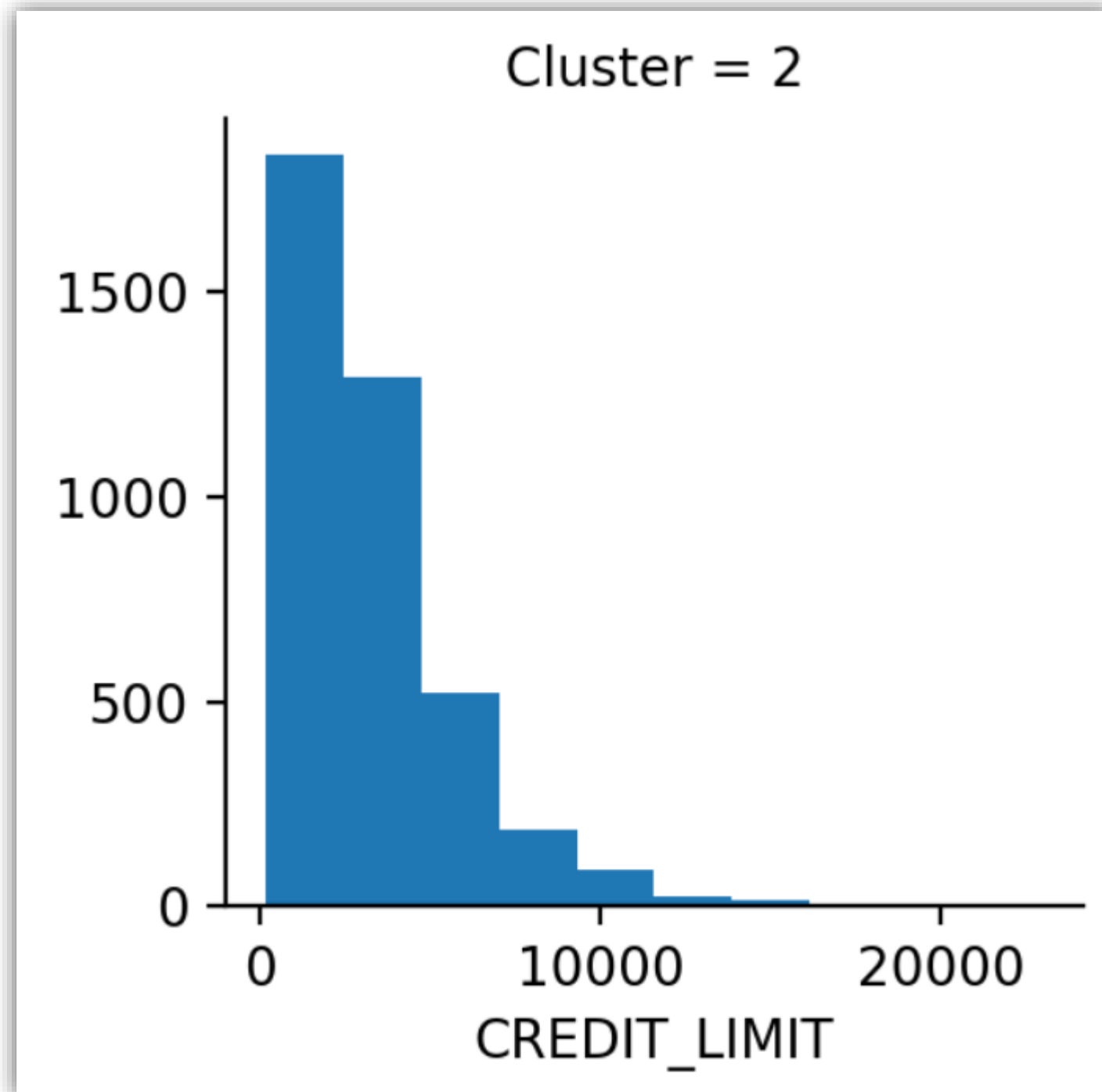


The visual shows Cash Advance TRX that it belongs to Cluster 2 and has highest frequency of 2500.

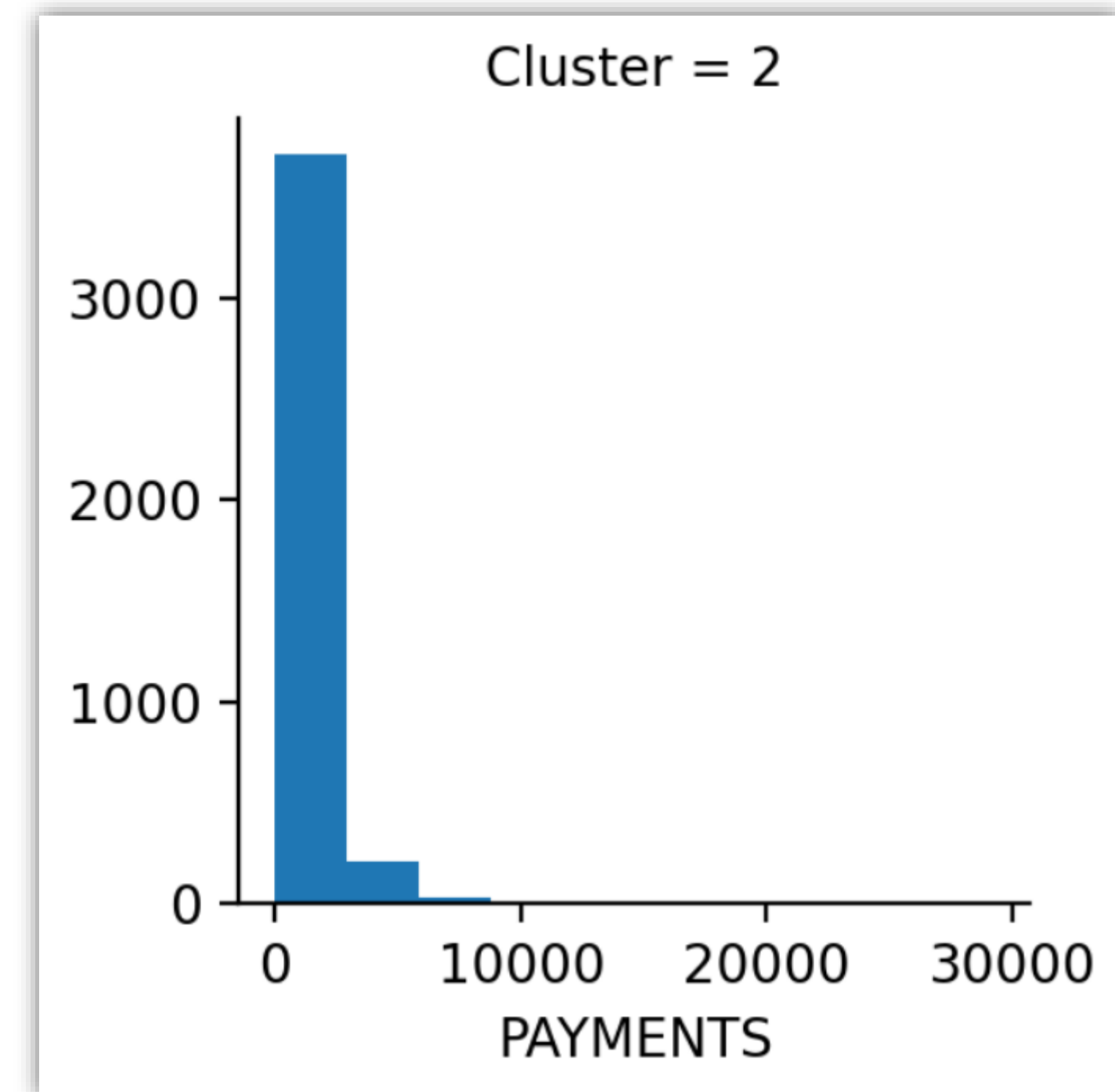


The visual shows the Purchases TRX which has highest frequency of 3000.

# CREDIT LIMIT & PAYMENTS



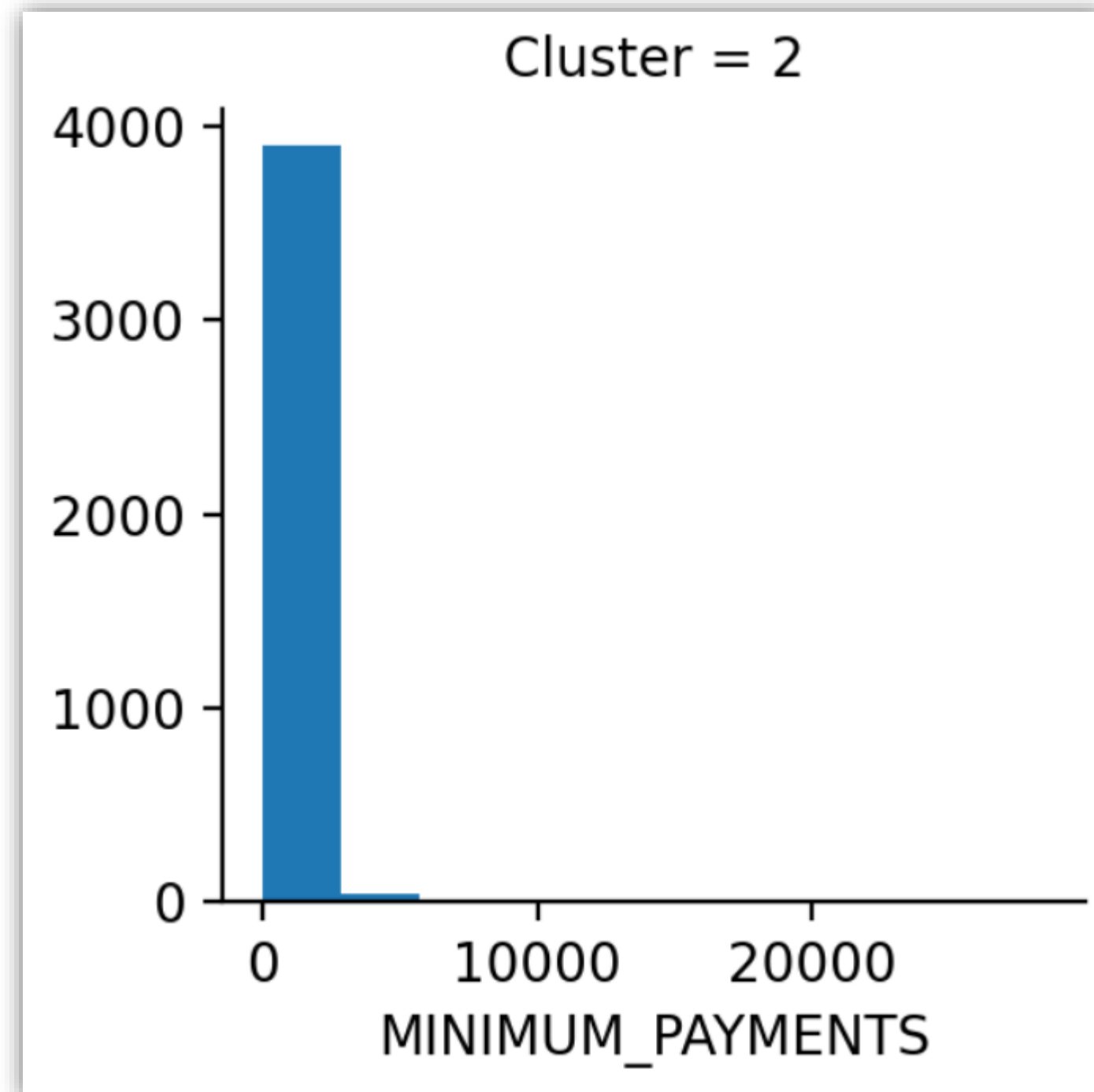
The visual shows Credit Limit that it belongs to Cluster 2 and has highest frequency of 1500.



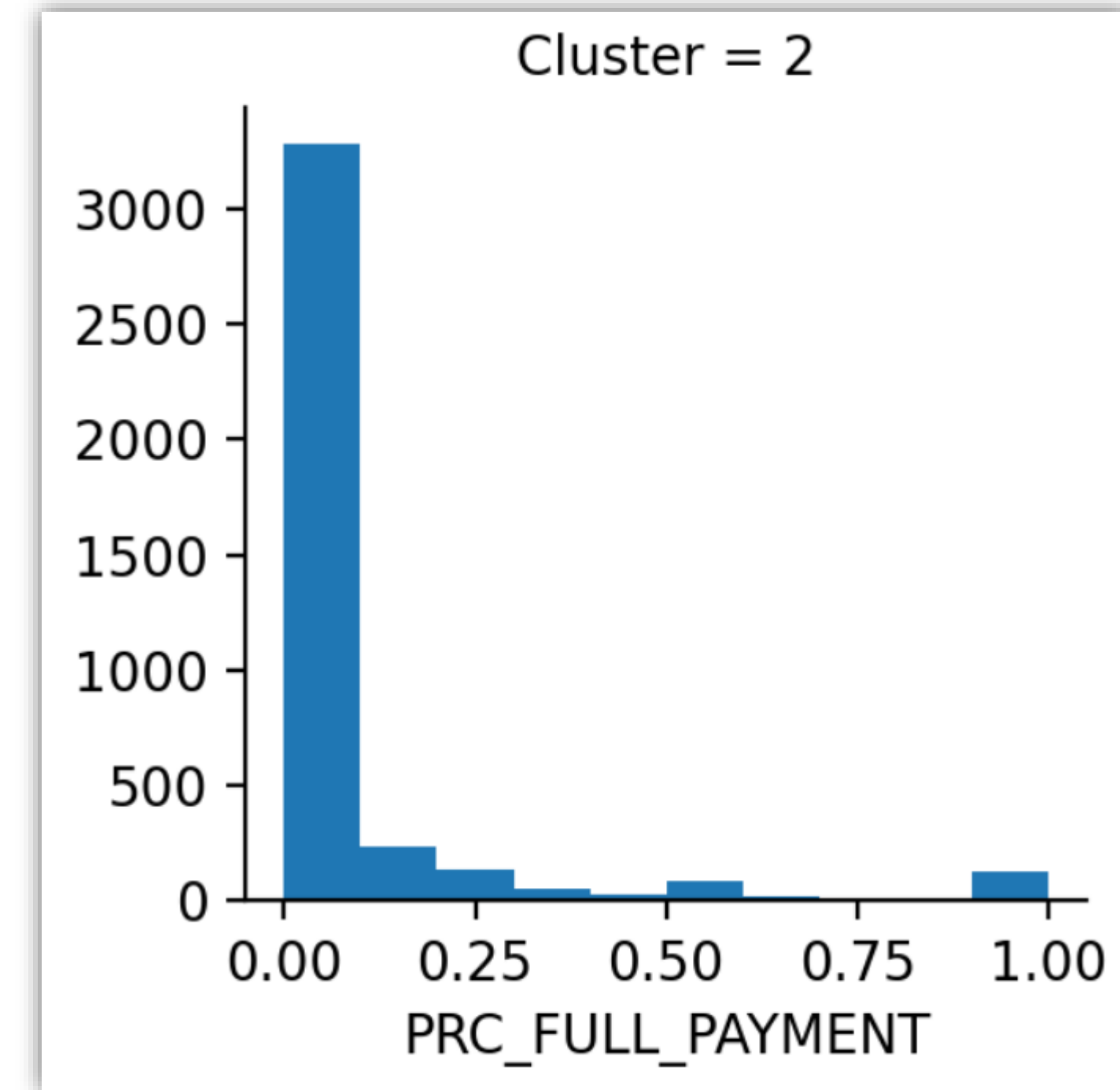
The visual shows the Payments for Cluster 2 which has highest frequency of 3000.



# MINIMUM PAYMENTS & PRC FULL PAYMENT



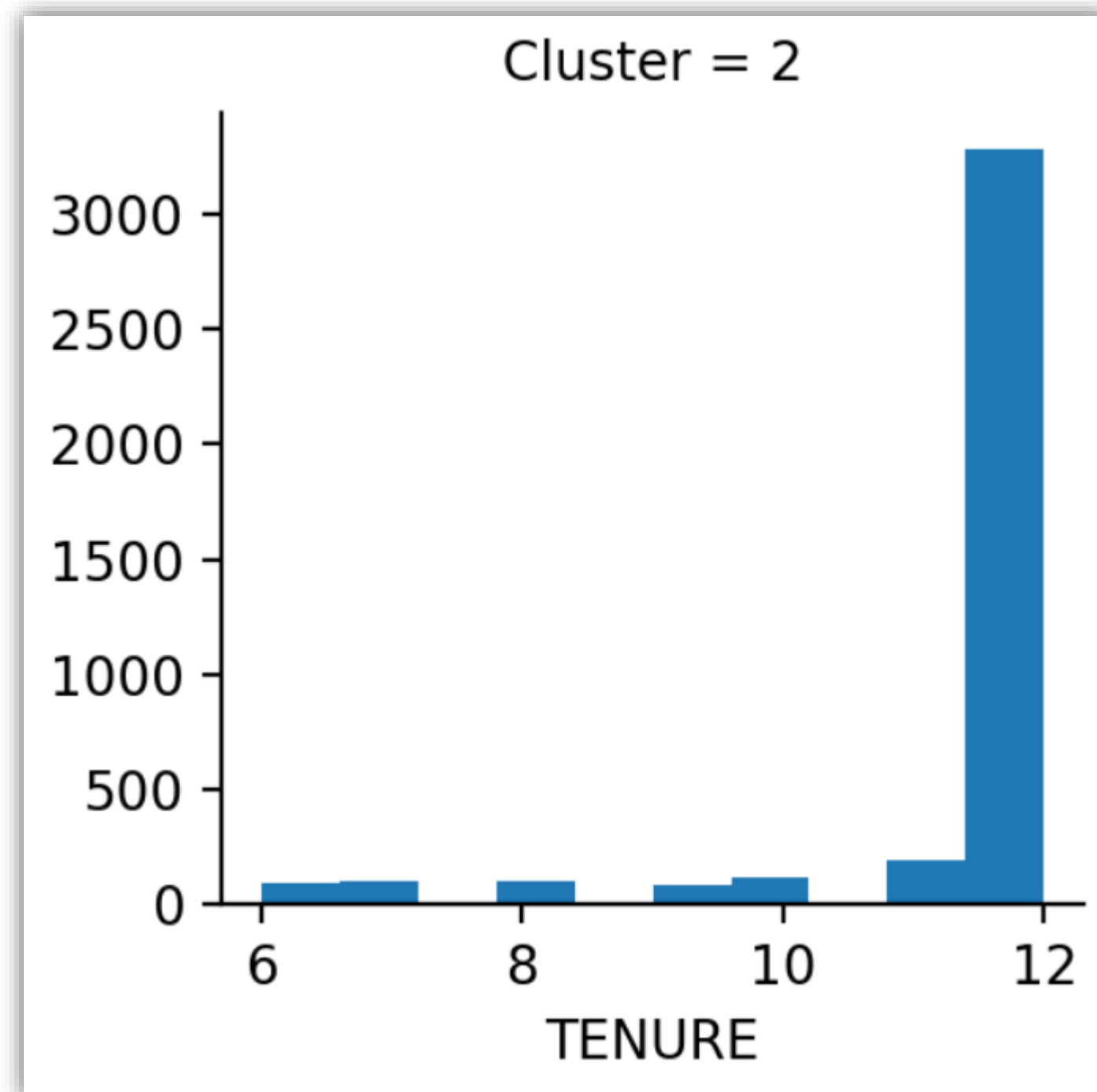
The visual shows Minimum Payments that it belongs to Cluster 2 and has highest frequency of 4000.



The visual shows the PRC Full Payment which has highest frequency of 3000.



# TENURE



The visual shows the Tenure for cluster 2 which has highest frequency of 3000.



## DEPLOYMENT

We have created a Streamlit Application based on this clustering technique, where we are taking the customer details & identifying which cluster the customer belongs to.

THANK YOU

# MARKET SEGMENTATION

A Machine learning based market  
segmentation project.

Type your text

PROJECT BY: Rakib Mathakiya