



Anomaly Detection for Fraud Risk in Banking Data

[PL300 - 149] - Group 1 - Đào Minh Thuấn

Mentor: Cô Nguyễn Thị Lê Trinh

November 16, 2025





Banking Transaction

City

All

2020

2021

2022

2023

2024

2025



Total Amount

\$14.89M

PY: \$12.52M

18.99% ▲ vs PY



Total Transaction

50.00K

PY: 42.10K

18.75% ▲ vs PY

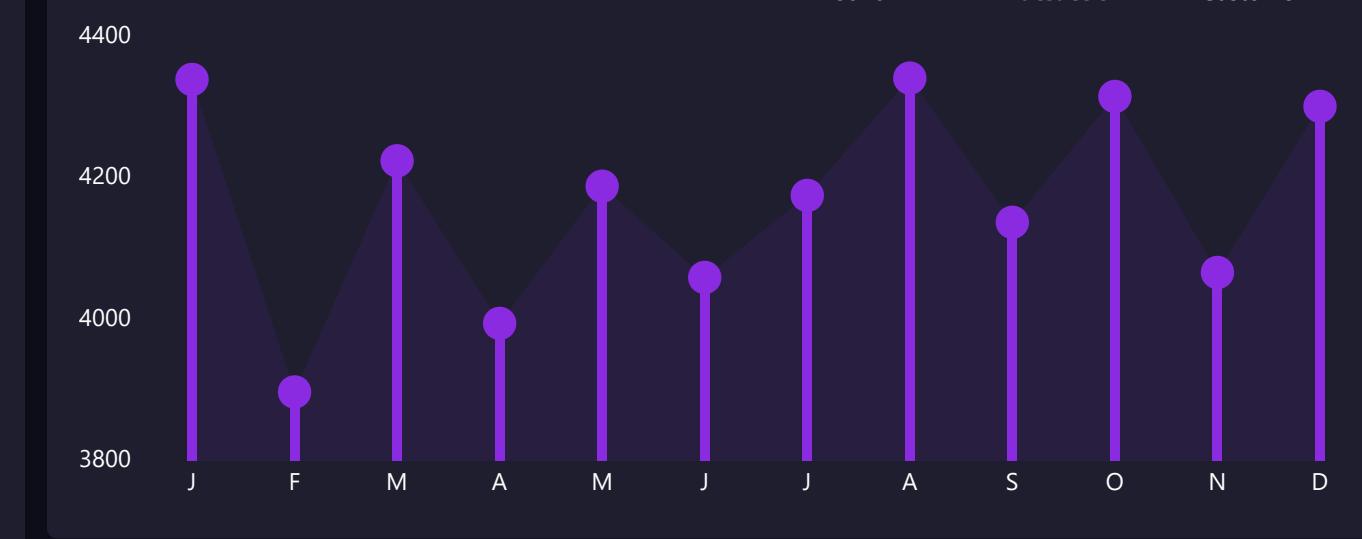


Top 20 Amount by Merchant

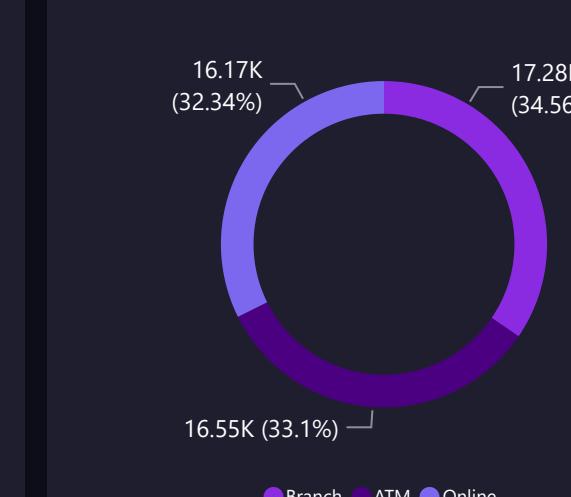
M026	\$0.30M
M009	\$0.23M
M066	\$0.22M
M081	\$0.22M
M064	\$0.22M
M060	\$0.21M
M013	\$0.21M
M028	\$0.21M
M032	\$0.20M
M054	\$0.20M
M038	\$0.20M
M088	\$0.20M
M048	\$0.19M
M014	\$0.19M
M007	\$0.19M
M042	\$0.19M
M018	\$0.19M
M071	\$0.19M
M015	\$0.19M
M068	\$0.19M



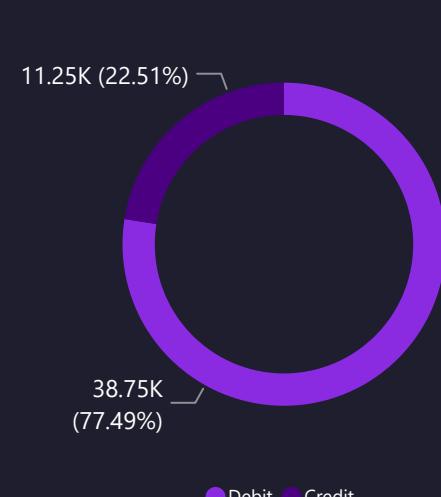
Tracsaction by Month



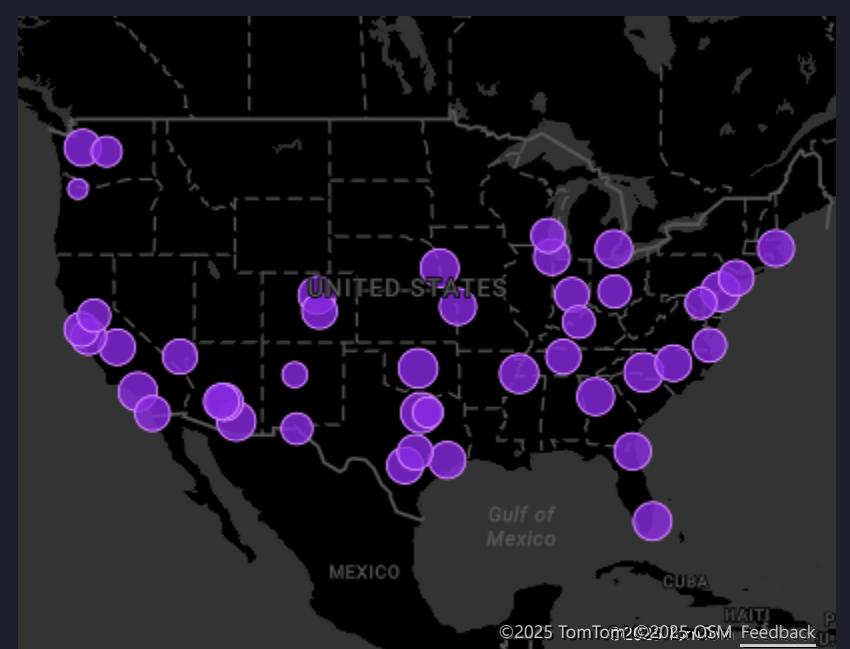
Tracsaction by Channel



Tracsaction by TransactionType



Tracsaction by Location



City [Top 5]	Amount %	Total Amount
Austin	20.72%	\$453,525.00
Memphis	20.36%	\$445,660.00
Oklahoma City	20.31%	\$444,552.00
San Antonio	19.32%	\$423,001.00
Fort Worth	19.29%	\$422,300.00



Banking Transaction

MTD

QTD

YTD

City

All

2020

2021

2022

2023

2024

2025



Amount MTD

\$215.47K

PY: \$191.07K

12.77% ▲ vs PY



Transaction MTD

724

PY: 633

14.38% ▲ vs PY



Avg Trans Value MTD

Avg Trans Value MTD

\$297.61

PY: \$301.84

-1.40% ▼ vs PY



Avg Cus Value MTD

\$605.26

PY: \$557.04

8.66% ▲ vs PY



Amount xTD and Amount xTD PY by Month



Top 20 ACV MTD by Customer

Top 20 ACV MTD by Customer

AC00265	\$3.4K
AC00362	\$3.1K
AC00257	\$2.7K
AC00304	\$2.6K
AC00318	\$2.6K
AC00179	\$2.4K
AC00418	\$2.4K
AC00078	\$2.3K
AC00136	\$2.3K
AC00407	\$2.3K
AC00156	\$2.1K
AC00464	\$2.0K
AC00086	\$2.0K
AC00460	\$2.0K
AC00431	\$1.9K
AC00272	\$1.9K
AC00028	\$1.8K
AC00046	\$1.8K
AC00366	\$1.8K
AC00357	\$1.7K



Top 10 Amount MTD by Device

Top 10 Amount MTD by Device

D000019	\$2.8K
D000563	\$2.7K
D000032	\$2.5K
D000030	\$2.5K
D000022	\$2.3K
D000008	\$2.2K
D000208	\$2.2K
D000246	\$2.0K
D000443	\$2.0K
D000326	\$2.0K

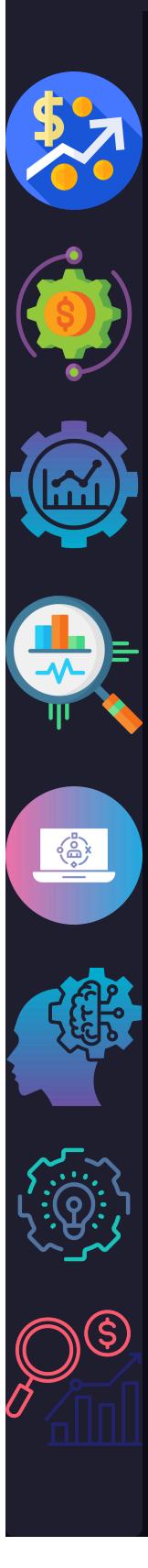
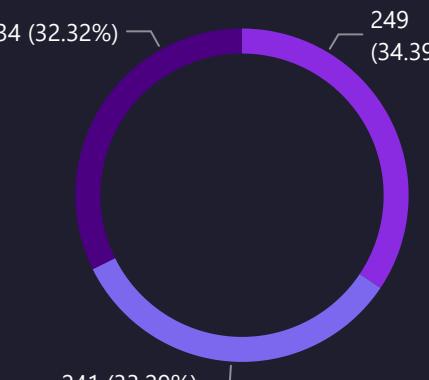


City [Top 7] Trans Growth % xTD

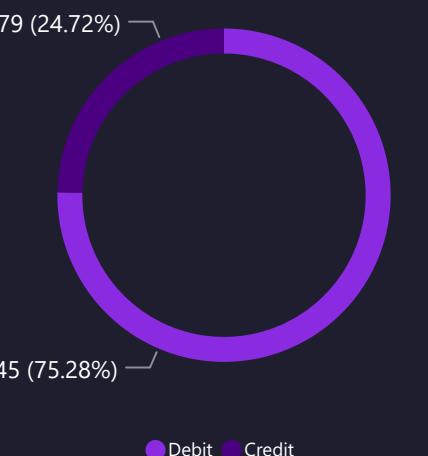
City [Top 7]	Trans Growth % xTD
Baltimore	47.06%
Albuquerque	37.93%
Dallas	37.50%
Austin	36.84%
Kansas City	31.71%
San Francisco	58.33%
Boston	74.29%



Transaction MTD



Transaction MTD





Banking Transaction

City

All

2020

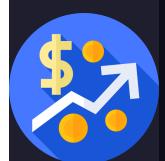
2021

2022

2023

2024

2025



Total Customer
495

PY: 495
0.00% ▼ vs PY



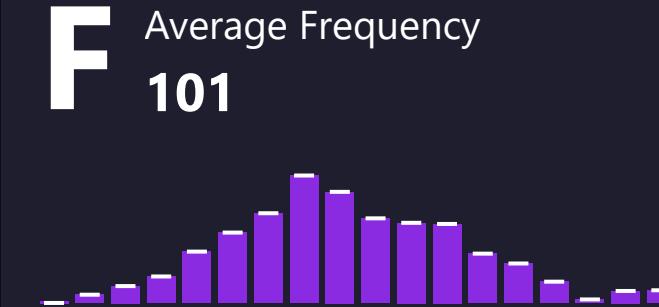
Total Customer by Segment Customer



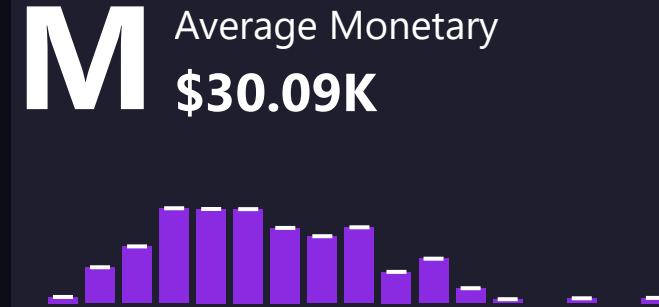
R Average Recency
29



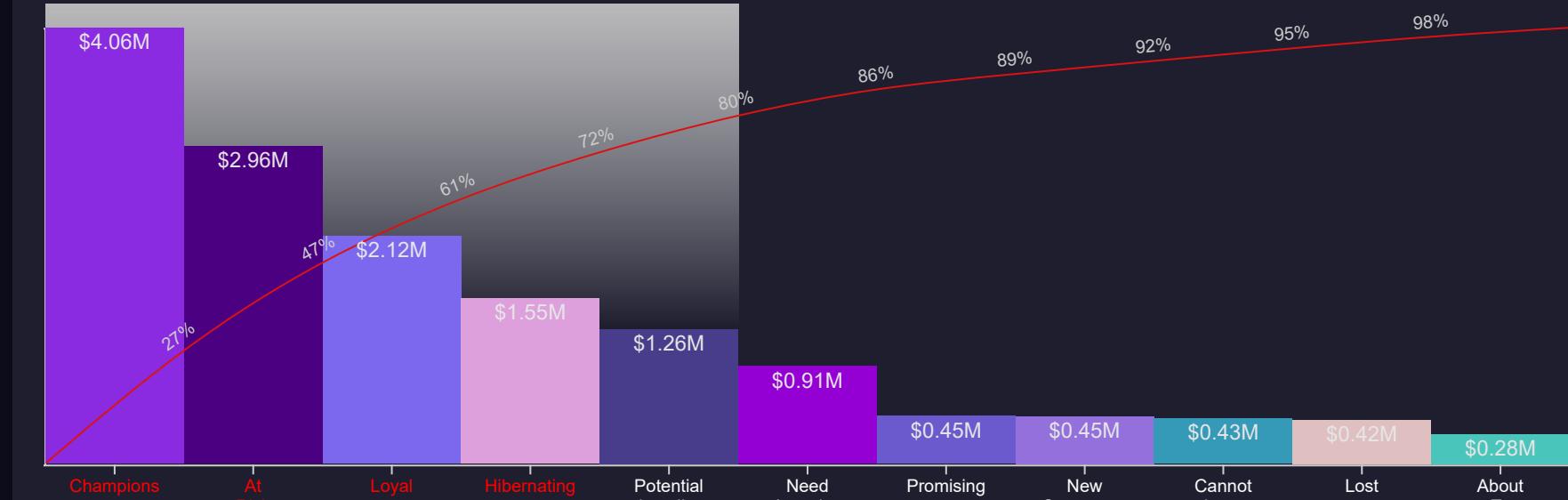
F Average Frequency
101



M Average Monetary
\$30.09K



Pareto Analysis of Customer Segments



AccountID	Segment Customer	RFM_Score	Recency	Frequency	MonetaryTuan
AC00014	Champions	555	3	176	\$64,202
AC00030	Champions	555	4	184	\$61,908
AC00032	Champions	555	3	187	\$72,499
AC00035	Champions	555	4	190	\$48,228
AC00086	Champions	555	2	146	\$62,681
AC00095	Champions	555	3	179	\$81,857
AC00136	Champions	555	1	159	\$77,692
AC00204	Champions	555	2	153	\$54,651



Banking Transaction

2020

2021

2022

2023

2024

City

All



Retention Customer by Month after First Transaction Month

First Transaction Month	2020-Jan	2020-Feb	2020-Mar	2020-Apr	2020-May	2020-Jun	2020-Jul	2020-Aug	2020-Sep	2020-Oct	2020-Nov	2020-Dec	2021-Jan	2021-Feb	2021-M
	▲														
01/01/2020	350	262	261	265	264	255	248	269	259	270	261	279	264	235	210
02/01/2020		79	50	48	53	47	49	49	43	42	52	57	51	44	30
03/01/2020			33	20	20	20	14	21	17	21	17	20	18	16	10
04/02/2020				14	7	6	10	7	7	5	9	3	9	5	3
05/04/2020					7	2	5	3	3	4	4	1	2	3	1
06/04/2020						4	3	2	3	1		2	4	2	1
07/09/2020							3		1	1			1		1
11/09/2020										1					
12/23/2020											2			1	
02/09/2021												1			
04/08/2021															
Total	350	341	344	347	351	334	332	351	333	344	344	364	349	307	3

Summary

- **Clear early churn:** most customers leave within the first 1-3 periods.
- **Large cohort (2020-01)** dominates overall trends; smaller cohorts retain fewer and vanish quickly.
- **Retention steadily declines** over time, with many cohorts dropping to zero after a few periods.
- **Churn concentrates in early stages.**

Retention Rate by Month after First Transaction Month

First Transaction Month	2020-Jan	2020-Feb	2020-Mar	2020-Apr	2020-May	2020-Jun	2020-Jul	2020-Aug	2020-Sep	2020-Oct	2020-Nov	2020-Dec	2021-Jan	2021-Feb	2021-M
	▲														
01/01/2020	100.00%	74.86%	74.57%	75.71%	75.43%	72.86%	70.86%	76.86%	74.00%	77.14%	74.57%	79.71%	75.43%	67.14%	76.00%
02/01/2020		100.00%	63.29%	60.76%	67.09%	59.49%	62.03%	62.03%	54.43%	53.16%	65.82%	72.15%	64.56%	55.70%	56.90%
03/01/2020			100.00%	60.61%	60.61%	60.61%	42.42%	63.64%	51.52%	63.64%	51.52%	60.61%	54.55%	48.48%	60.60%
04/02/2020				100.00%	50.00%	42.86%	71.43%	50.00%	50.00%	35.71%	64.29%	21.43%	64.29%	35.71%	57.14%
05/04/2020					100.00%	28.57%	71.43%	42.86%	42.86%	57.14%	57.14%	14.29%	28.57%	42.86%	42.86%
06/04/2020						100.00%	75.00%	50.00%	75.00%	25.00%		50.00%	100.00%	50.00%	25.00%
07/09/2020							100.00%		33.33%	33.33%			33.33%		66.67%
11/09/2020									100.00%						
12/23/2020										100.00%			50.00%	50.00%	
02/09/2021											100.00%		100.00%	100.00%	





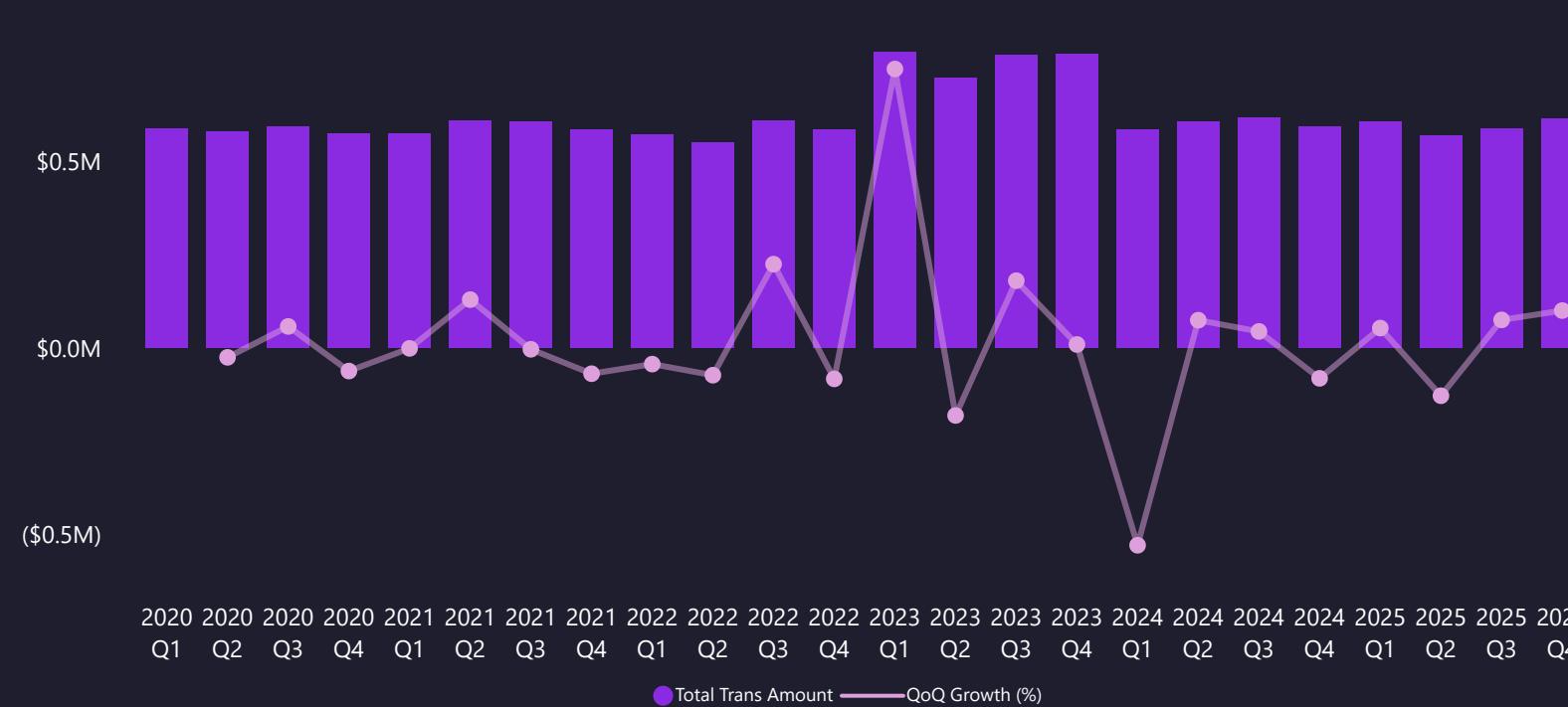
Banking Transaction

City

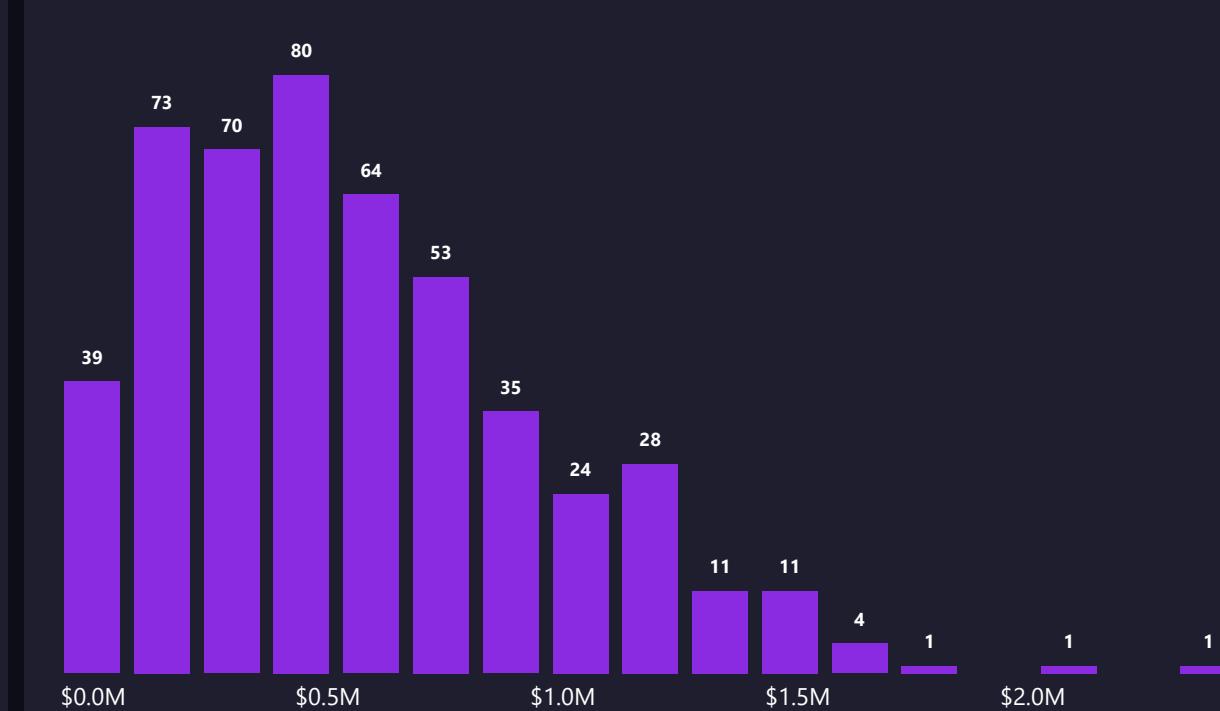
All



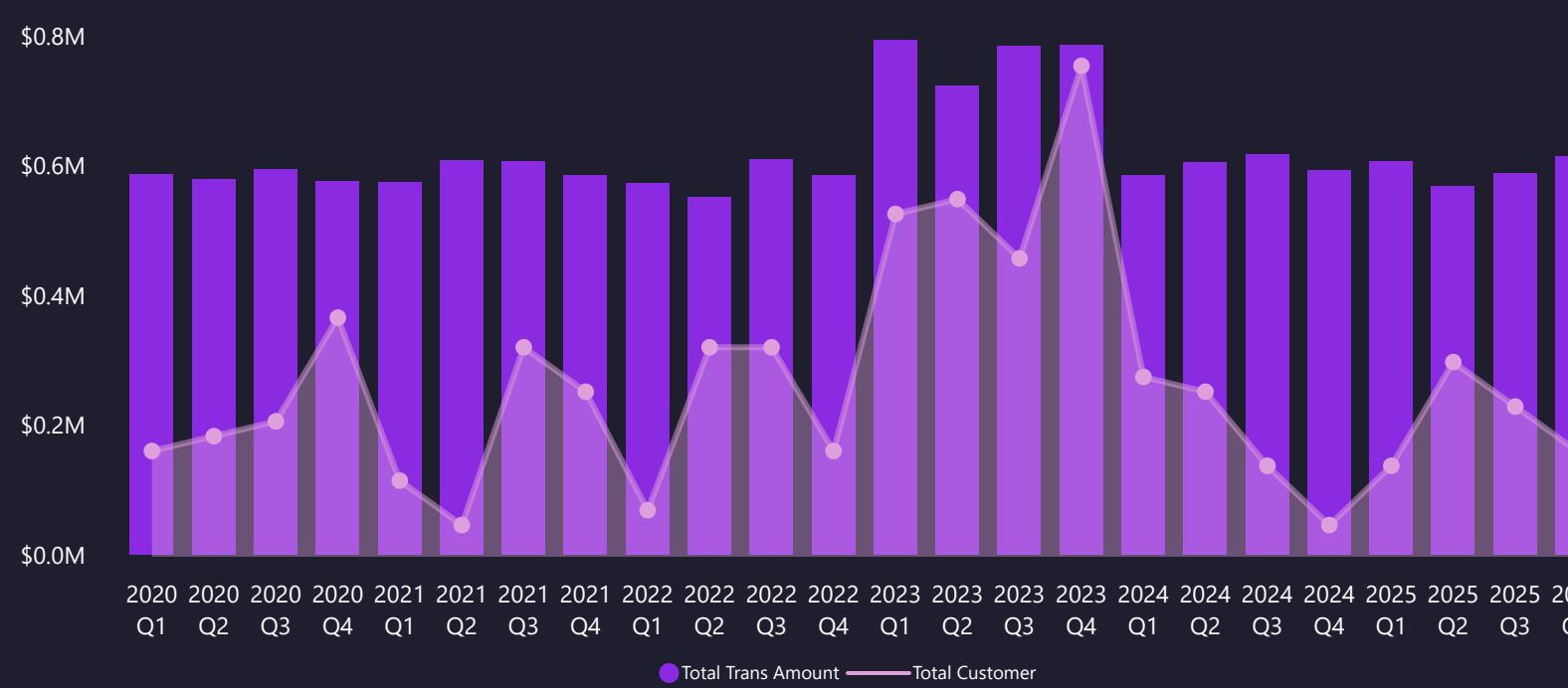
Quarter-over-Quarter Transaction Growth



Distribution of Customer Lifetime Value



Revenue & Account Count by Quarter (with QoQ Growth)



CLV Segment Distribution



CLV [Top 11]

CLV	Value
AC00460	\$2,378,294.10
AC00363	\$2,200,737.30
AC00337	\$1,793,822.10
AC00095	\$1,743,554.10
AC00021	\$1,684,830.00
AC00020	\$1,656,735.30
AC00136	\$1,654,839.60
AC00358	\$1,622,463.60
AC00202	\$1,612,899.90
AC00439	\$1,591,259.10
AC00257	\$1,546,763.40



Banking Transaction

2020

2021

2022

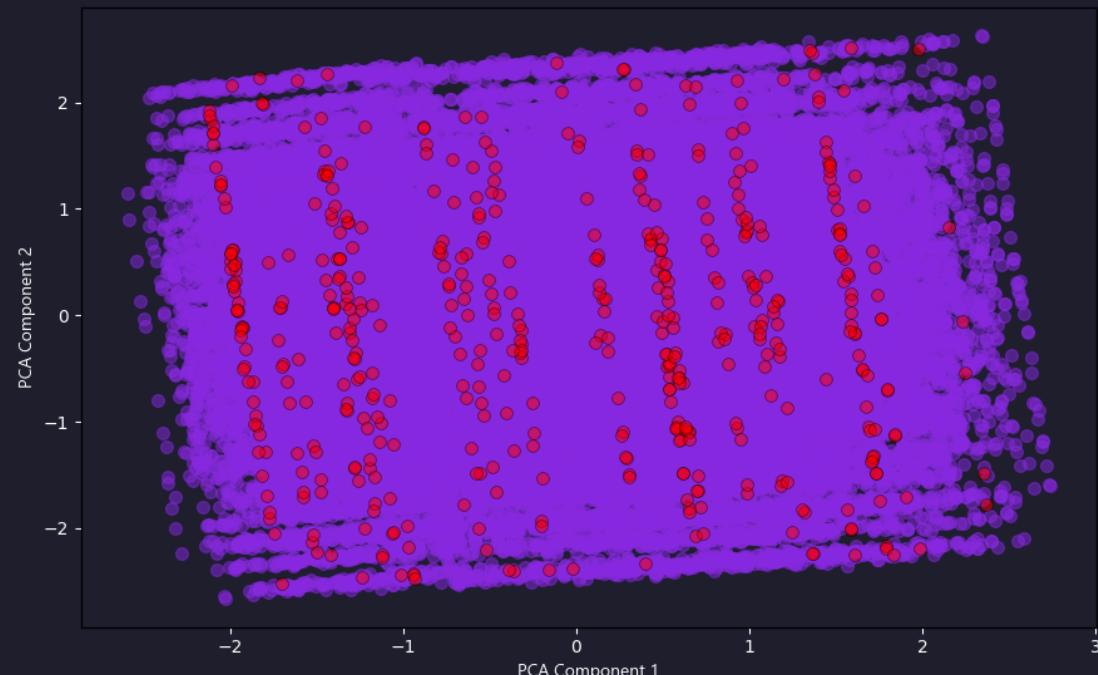
2023

2024

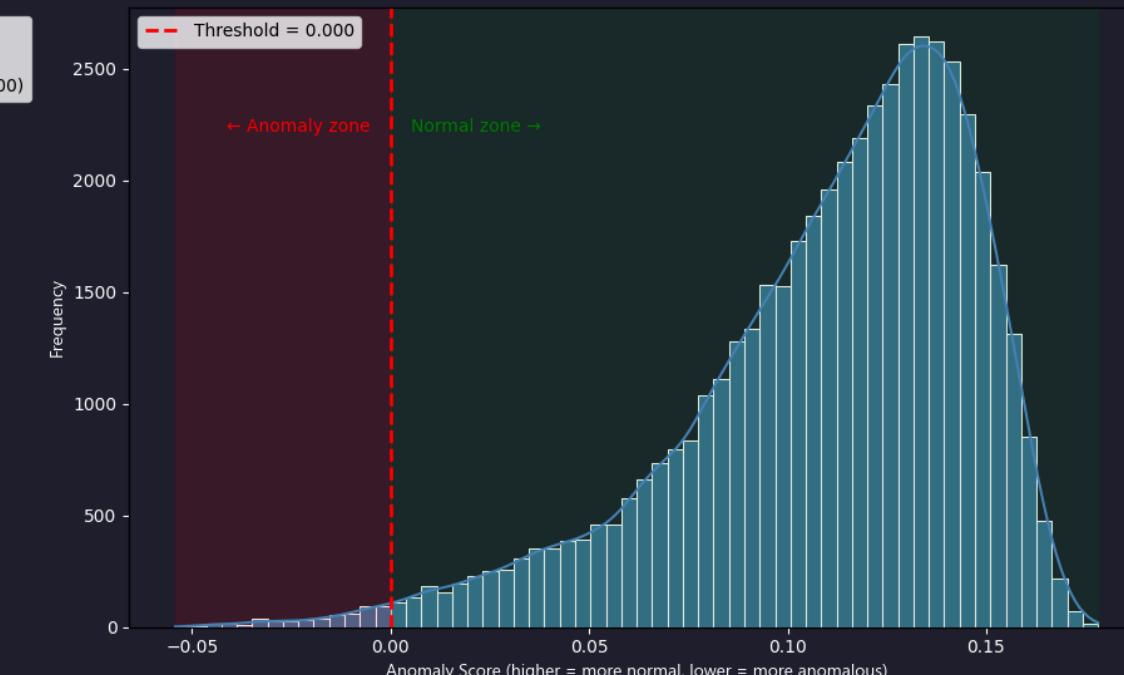
2025



Isolation Forest — Potential Frauds vs Normal (PCA 2D)



Distribution of Isolation Forest Anomaly Scores



TransactionID	AccountID	CustomerAge	CustomerOccupation	TransactionAmount	Location	Channel	TransactionType	TransactionDuration	AnomalyLabel	Segment	Customer
TX004749	AC00071	18	Student	\$124	Mesa	Branch	Debit	22	Potential Fraud	Champions	
TX005872	AC00071	18	Student	\$132	Mesa	Branch	Debit	22	Potential Fraud	Champions	
TX012221	AC00393	18	Student	\$150	Los Angeles	Online	Credit	176	Potential Fraud	Cannot Lose Them	
TX014762	AC00071	18	Student	\$123	Mesa	Branch	Debit	22	Potential Fraud	Champions	
TX021184	AC00393	18	Student	\$146	Los Angeles	Online	Credit	176	Potential Fraud	Cannot Lose Them	
TX023916	AC00071	18	Student	\$115	Mesa	Branch	Debit	22	Potential Fraud	Champions	
TX031608	AC00071	18	Student	\$118	Mesa	Branch	Debit	22	Potential Fraud	Champions	
TX031817	AC00071	18	Student	\$128	Mesa	Branch	Debit	22	Potential Fraud	Champions	
TX032465	AC00393	18	Student	\$151	Los Angeles	Online	Credit	176	Potential Fraud	Cannot Lose Them	
TX033669	AC00071	18	Student	\$122	Mesa	Branch	Debit	22	Potential Fraud	Champions	



Banking Transaction

2020

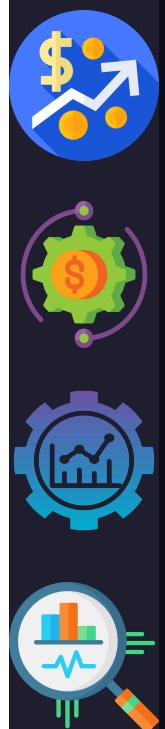
2021

2022

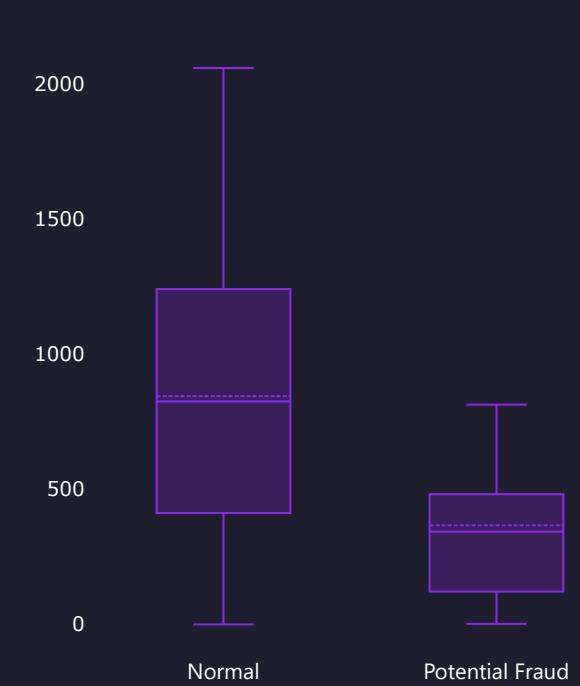
2023

2024

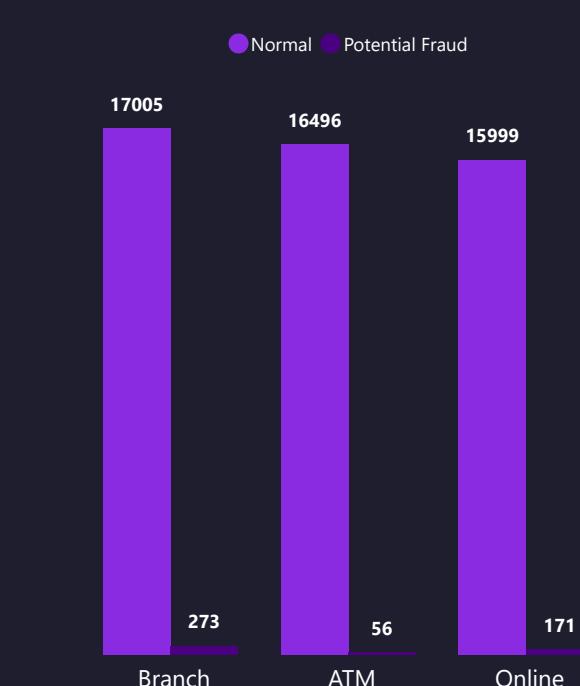
2025



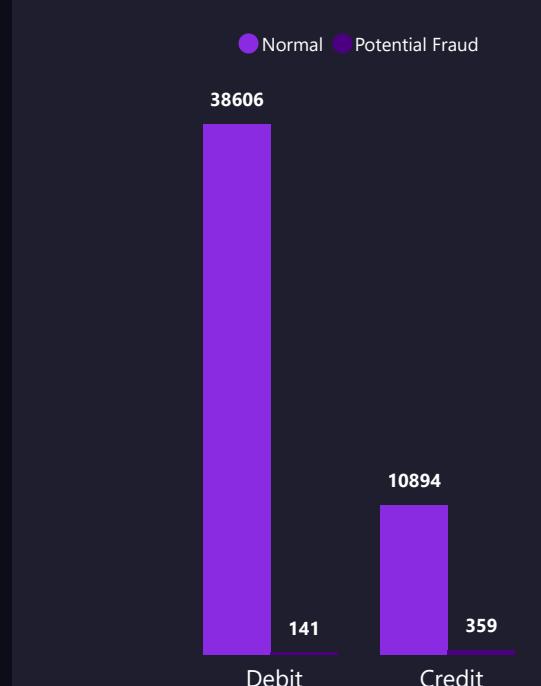
Trans Amount Distribution by Anomaly Label



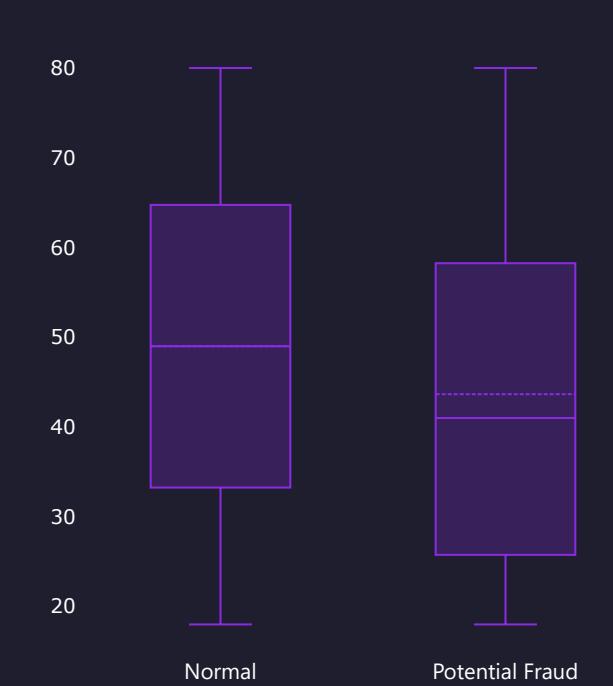
Fraud Pattern by Channel



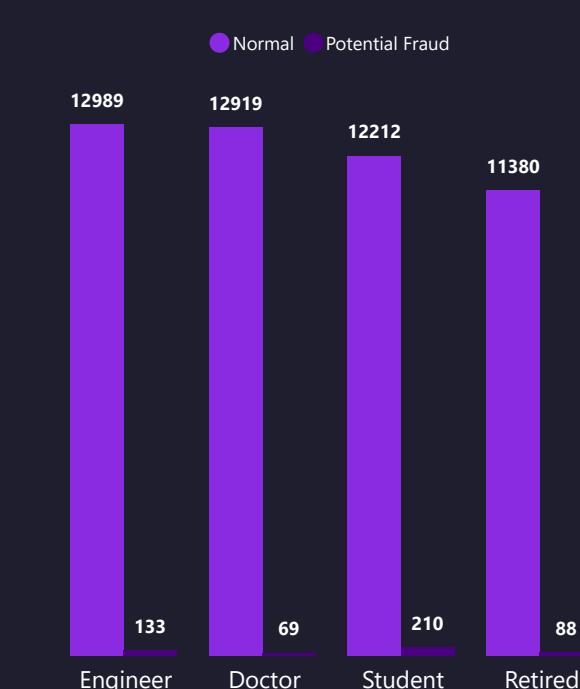
Fraud Pattern by Transaction Type



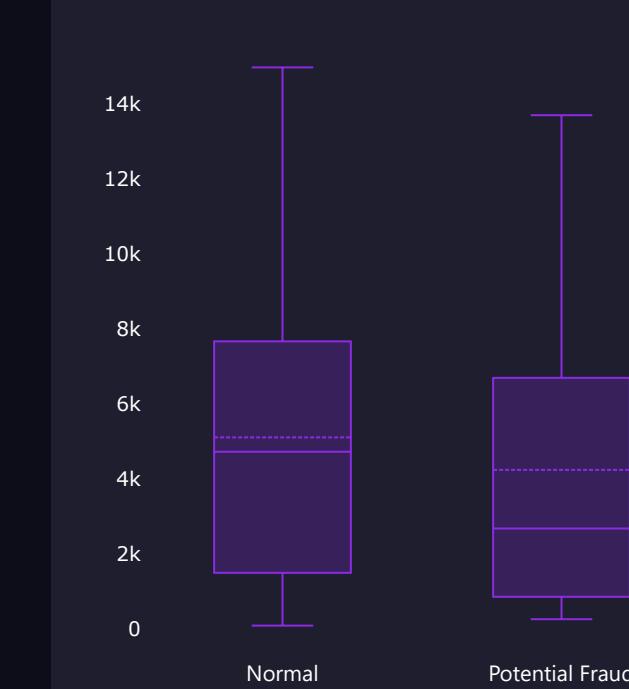
Customer Age Distribution by Anomaly Label



Fraud Pattern by Customer Occupation



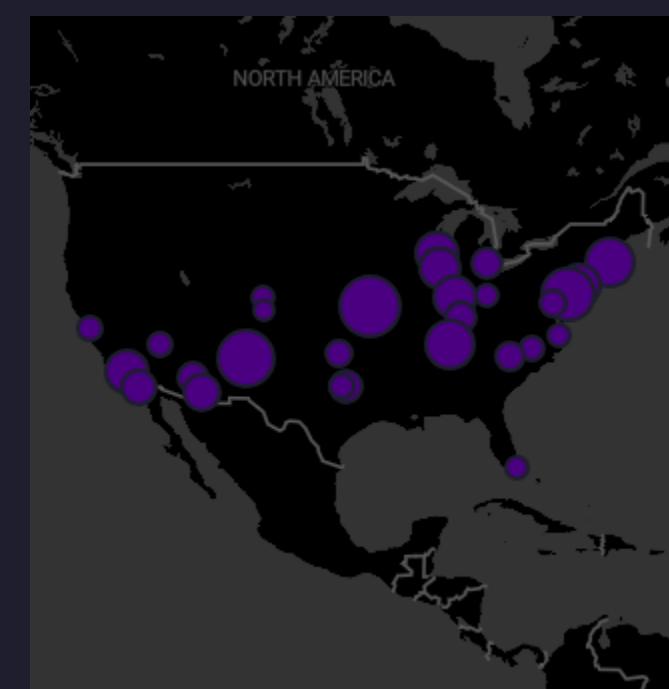
Acc Balance Distribution by Anomaly Label



Trans Duration Distribution by Anomaly Label



Geographic Distribution of Fraud





Banking Transaction

2020

2021

2022

2023

2024

2025



Monthly Trend of Average Transaction Amount by Anomaly Label

\$700

\$600

\$500

\$400

\$300

\$200

\$100

\$0

J

F

M

A

M

J

J

A

S

O

N

D

● Normal ● Potential Fraud

2020 | 2021 | 2022 | 2023 | 2024 | 2025