Credit Card Fraud Detection

Computer Security Project

PROJECT GUIDE

UPLOAD DATASET

MODELING ML

CONCLUTION

PREPROCESSING

TEST RESULTS
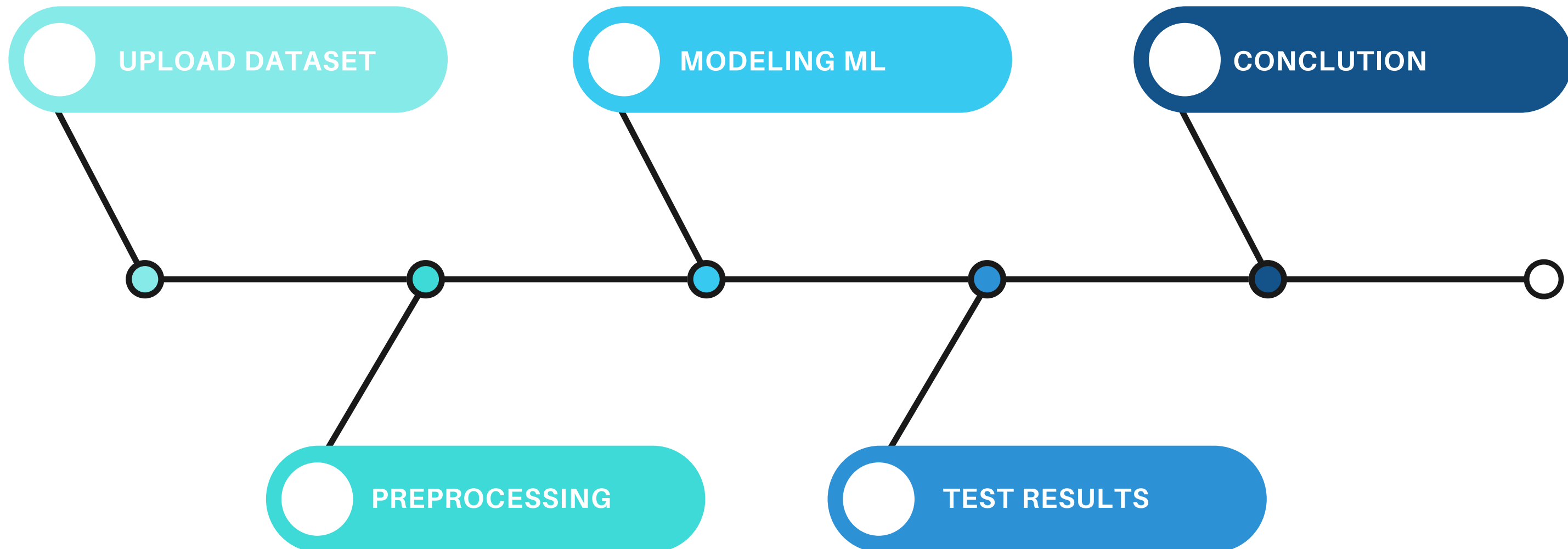
DATASET

RECORDS

FEATURES

24.4 M

15

DATA SAMPLE

	User	Card	Year	Month	Day	Time	Amount	Use Chip	Merchant Name	Merchant City	Merchant State	Zip	MCC	Errors?	Is Fraud?
Type	int64	int64	int64	int64	int64	object	object	object	int64	object	object	float64	int64	object	object

User	Card	Year	Month	Day	Time	Amount	Use Chip	Merchant Name	Merchant City	Merchant State	Zip	MCC	Errors?	Is Fraud?
661	2	2012	5	13	20:01	$45.64	Swipe Transaction	8919682822789039434	Frankfort	IN	46041.0	5814	NaN	No
1034	1	2013	9	21	11:47	$7.60	Swipe Transaction	-6766443224559424591	Alcoa	TN	37701.0	5813	NaN	No
1018	0	2013	8	31	17:44	$2.41	Swipe Transaction	1184165243450200853	Clifton	NJ	7012.0	4121	NaN	No
1423	3	2019	8	22	18:51	$37.17	Chip Transaction	8454598841454856436	Hamptonville	NC	27020.0	5812	NaN	No
1979	3	2017	7	21	11:36	$20.08	Swipe Transaction	309536160951779042	Roanoke	TX	76262.0	5812	NaN	No

DATA DESCRIPTION

	count	unique	top	freq	mean	std	min	25%	50%	75%
User	24386900.000000				1001.019335	569.461157	0.000000	510.000000	1006.000000	1477.000000
Card	24386900.000000				1.351366	1.407154	0.000000	0.000000	1.000000	2.000000
Year	24386900.000000				2011.955170	5.105921	1991.000000	2008.000000	2013.000000	2016.000000
Month	24386900.000000				6.525064	3.472355	1.000000	3.000000	7.000000	10.000000
Day	24386900.000000				15.718123	8.794073	1.000000	8.000000	16.000000	23.000000
Amount	24386900.000000				41.987522	78.688927	-500.000000	9.200000	30.139999	65.059998
Use Chip	24386900	3	Swipe Transaction	15386082						
Merchant Name	24386900.000000				-476922962766468352.000000	4758939870561625088.000000	-9222899435637403648.000000	-4500542936415012352.000000	-794676495118551552.000000	3189517333335617024.000000
Merchant City	24386900	13429	ONLINE	2720821						
Merchant State	21666079	223	CA	2591830						
Zip	21508765.000000						501.000000	28368.000000	46752.000000	
MCC	24386900.000000				5561.171253	879.315433	1711.000000	5300.000000	5499.000000	5812.000000
Errors?	388431	23	Insufficient Balance	242783						
Is Fraud?	24386900	2	No	24357143						
Hour	24386900.000000				12.414200	5.065882	0.000000	9.000000	12.000000	16.000000
Minutes	24386900.000000				29.585951	17.318349	0.000000	15.000000	30.000000	45.000000

MISSING VALUES

MERCHANT STATE

A

2.7 M   11.1%
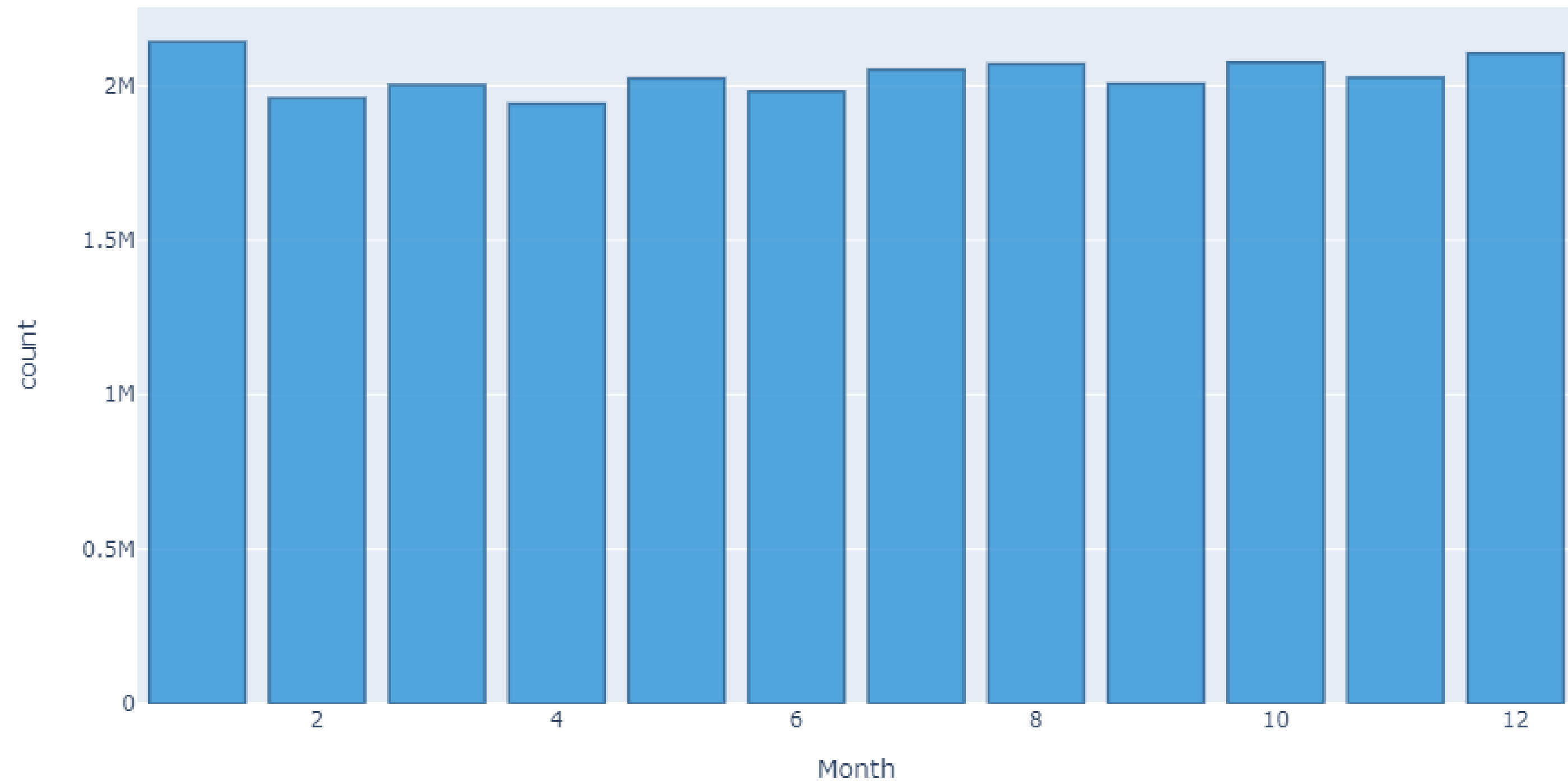
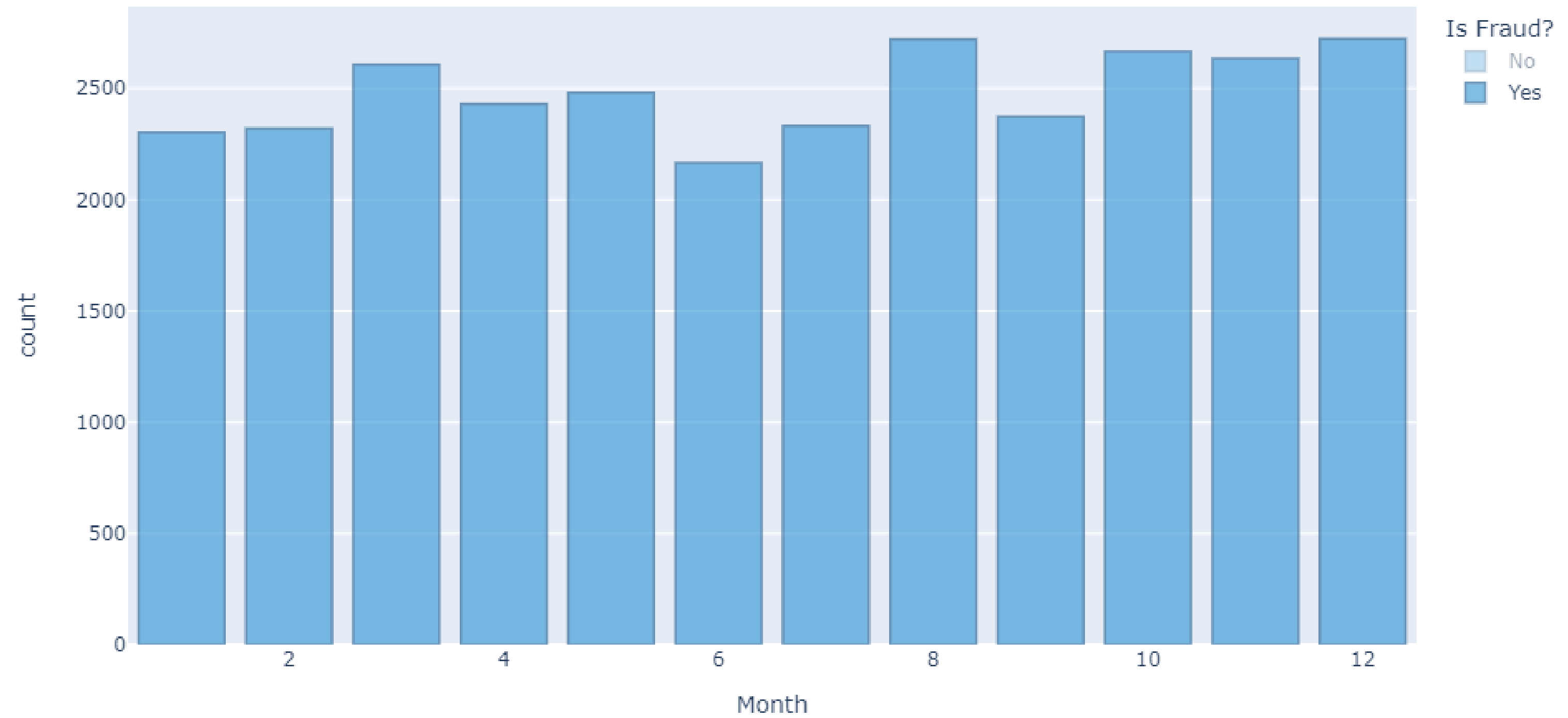**ZIP**

A

2.8 M   11.8%
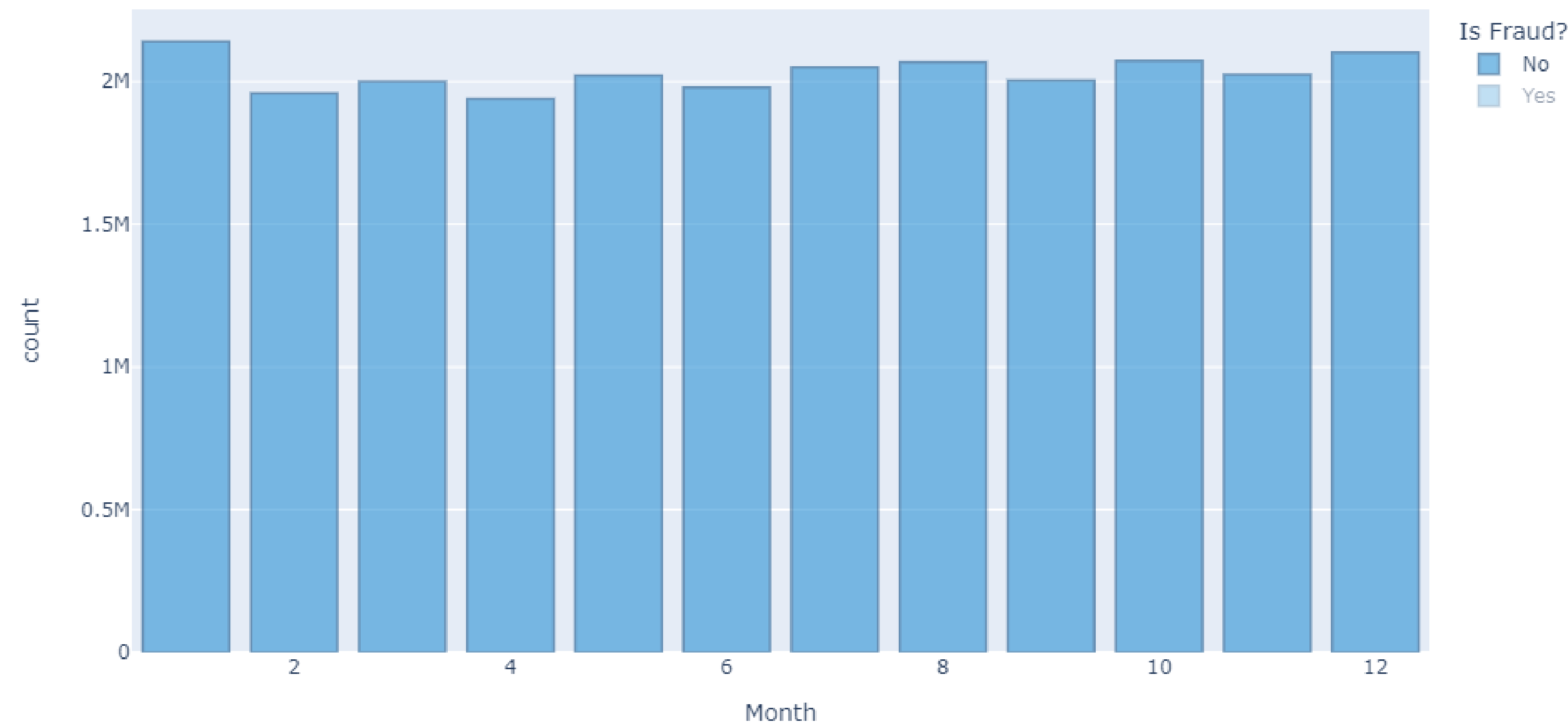
ERRORS

A

23.9 M   98.4%
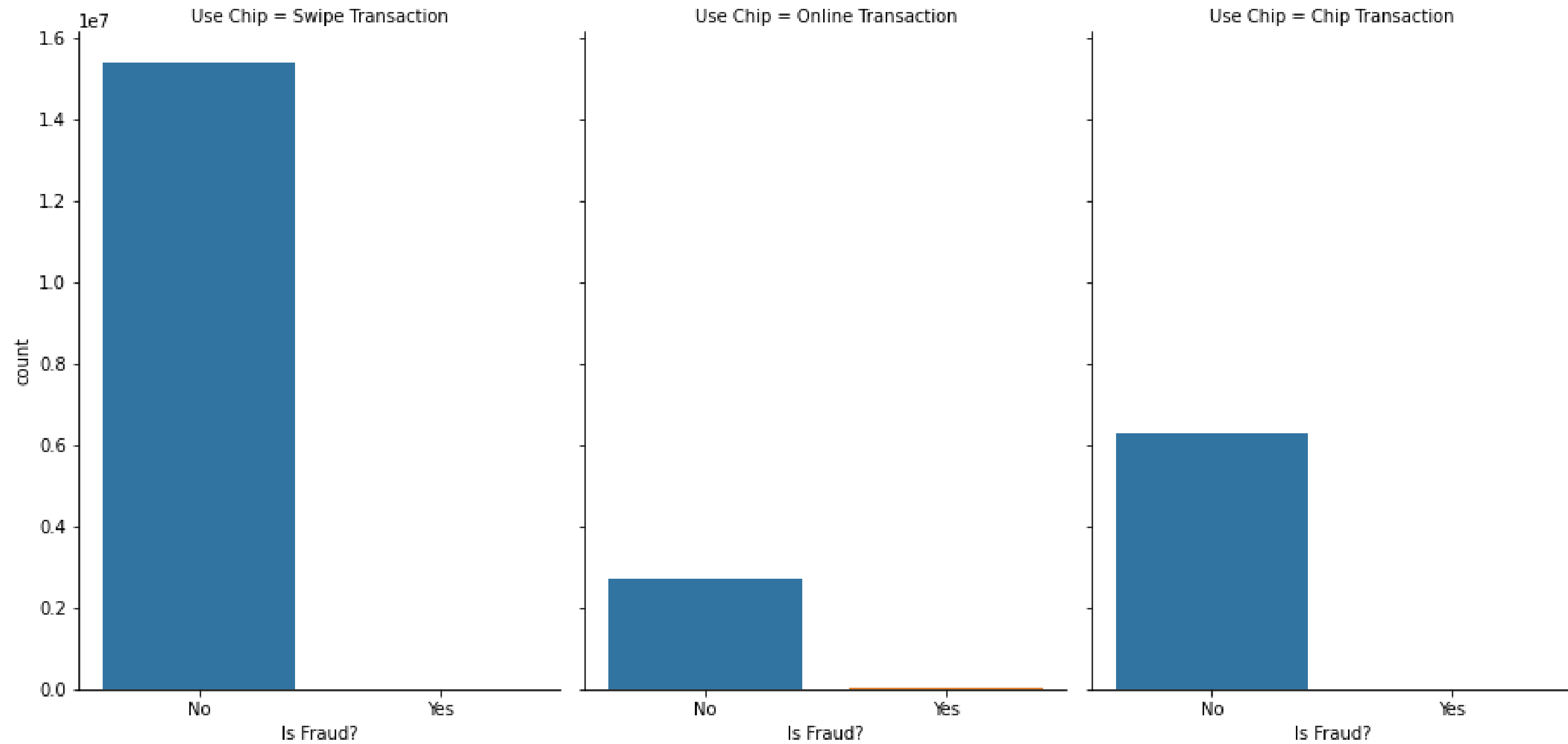
Transactions Distribution over the months

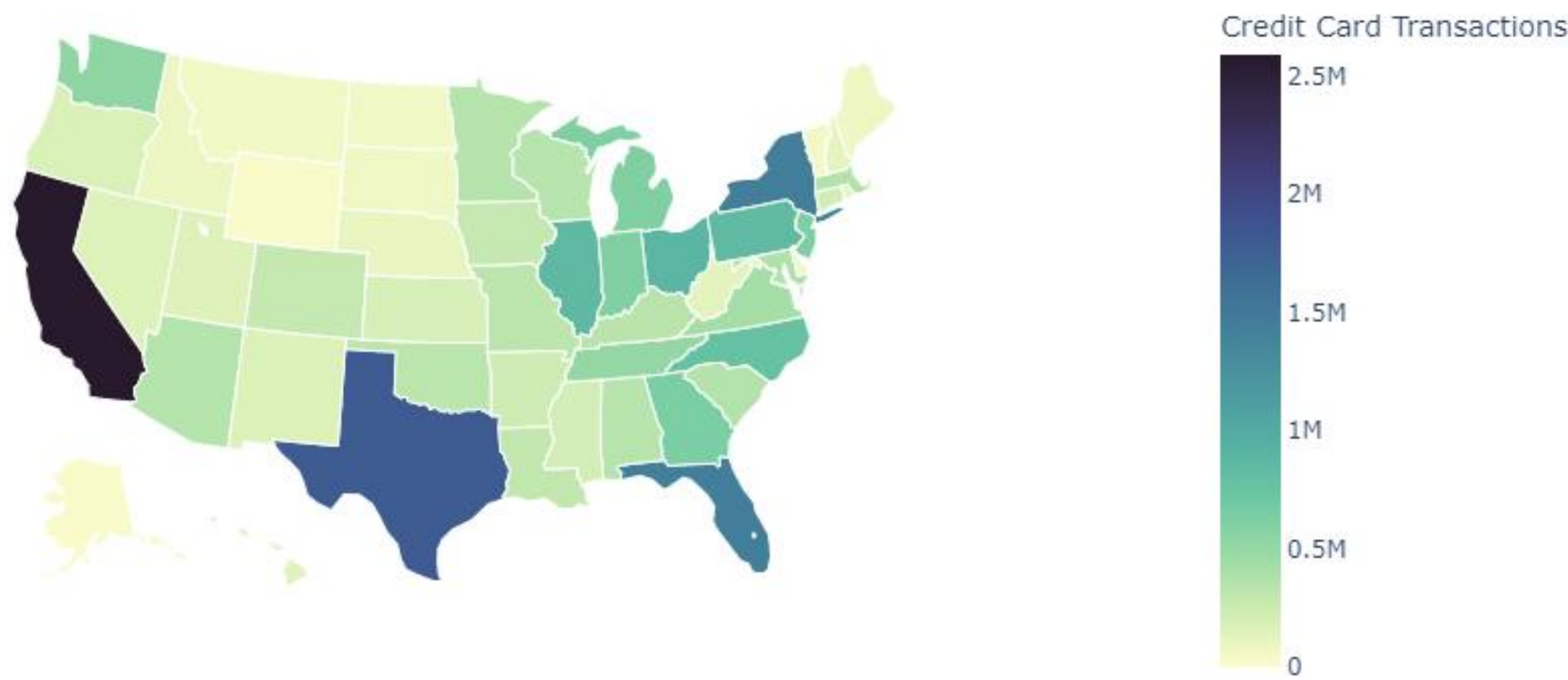# Transactions Distribution over the months

# Transactions Distribution over the months

Card Type and Fraud

Credit Card Transactions per US state

DATA SAMPLE AFTER CLEANING

	User	Card	Year	Month	Day	Amount	Use Chip	Merchant Name	Merchant City	Merchant State	Zip	MCC	Errors?	Is Fraud?	Hour	Minutes
Type	uint16	uint8	int16	uint8	uint8	float32	object	int64	object	object	float16	int64	object	object	uint8	uint8

Card	Year	Month	Day	Amount	Use Chip	Merchant Name	Merchant City	Merchant State	Zip	MCC	Errors?	Is Fraud?	Hour	Minu
4	2010	6	9	166.690002	Online Transaction	-8338381919281017248	ONLINE	NaN	NaN	4899	NaN	No	16	
0	2011	2	11	17.680000	Online Transaction	-6160036380778658394	ONLINE	NaN	NaN	4121	NaN	No	1	
2	2010	8	17	15.640000	Swipe Transaction	4722913068560264812	Everett	PA	15536.0	5411	NaN	No	13	
0	2010	6	23	222.389999	Swipe Transaction	748788744921077435	Alpharetta	GA	30000.0	4900	NaN	No	4	
0	2004	10	16	42.330002	Swipe Transaction	8080934608468946743	Niles	OH	44448.0	5310	NaN	No	18	

DATA CORRELATION
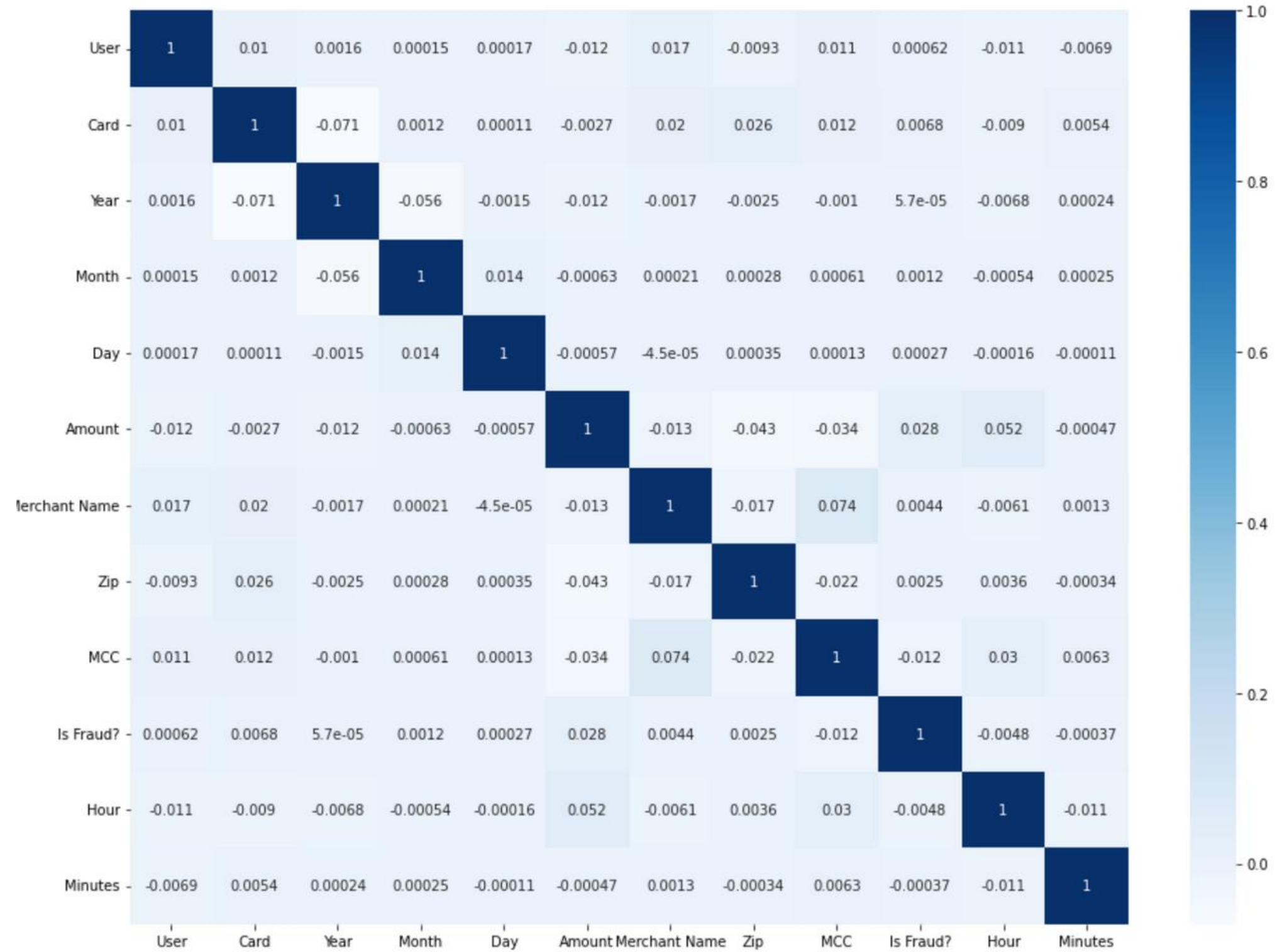
APPLY THE ML ALGORITHM

DATA SPLIT

A

66.7%
TRAIN

33.3%
TEST
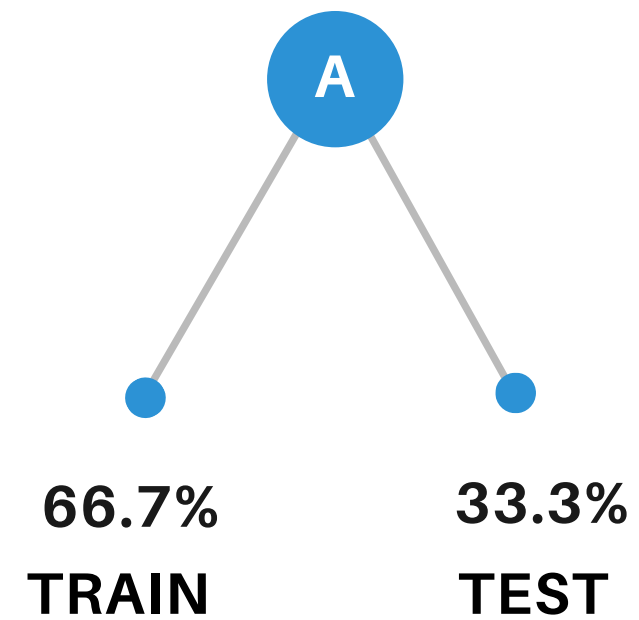
LIGHTGBM
CLASSIFIER

CONFUSION MATRIX

TRAINING SCORE

99.9%

ACCURACY

99.3%
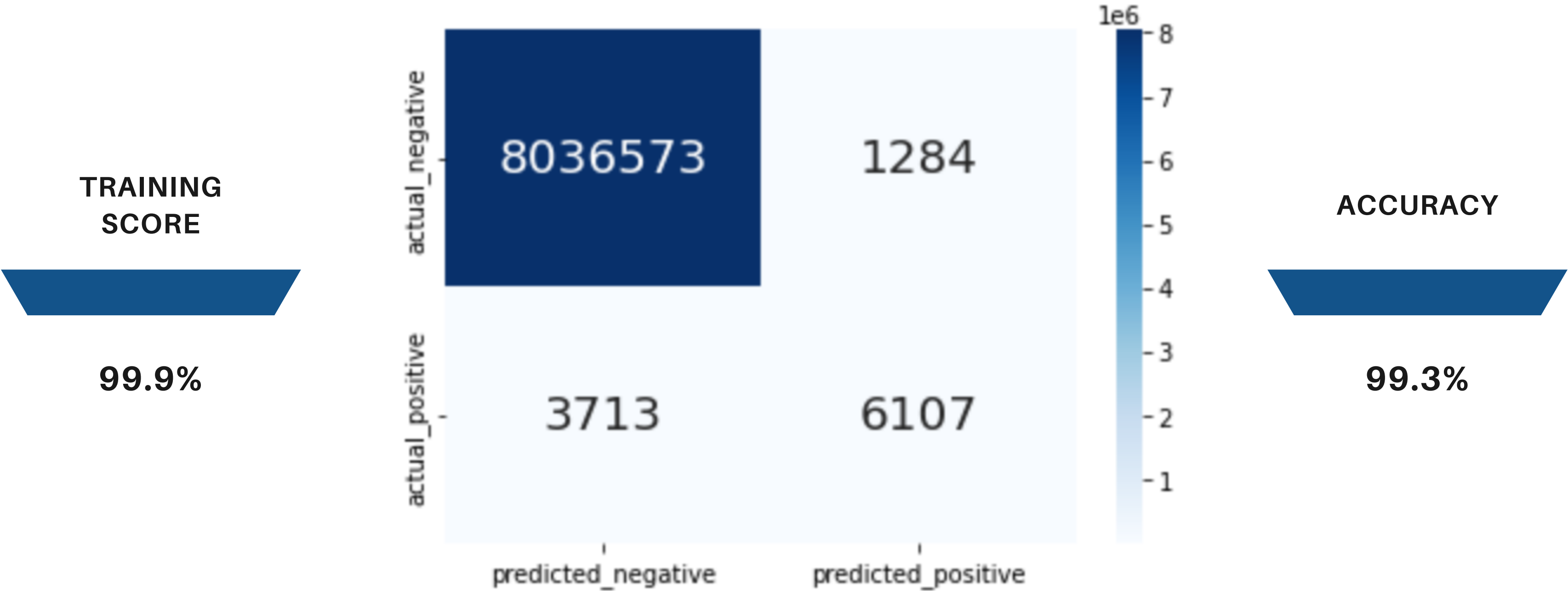
CLASSIFICATION REPORT

	PRECISION	RECALL	F1-SCORE	SUPPORT
NO	1.00	1.00	1.00	8037857
YES	0.83	0.62	0.71	9820
ACCURACY			1.00	8047677
MACRO AVG	0.91	0.81	0.85	8047677
WEIGHTED AVG	1.00	1.00	1.00	8047677

Thanks

- **Abdulaziz Alkhonain**
- **Hamad Alhajlan**
- **Mohammed Banajh**

Computer Security Project