

# Data Quality Report

## 1) File Description

- a. Dataset name: Credit Card transactions data
- b. Description: This dataset contains information of credit card transactions that have occurred in USA. It is a supervised dataset with Fraud labels.
- c. Time period: 01/01/2006 – 12/31/2006
- d. No. of records: 96753
- e. No. of fields: 10

## 2) Summary Statistics

### Categorical variables:

Column Name	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch ZIP	Transtype	Fraud
# of records that have a value	96,753	96,753	93,378	96,753	95,558	92,097	96,753	96,753
% of record populated	100%	100%	96.51%	100%	98.76%	95.19%	100%	100%
# of records with null value	0	0	3375	0	1,195	4,656	0	0
# of unique values	96,753	1,645	13,092	13,126	228	4,568	4	2
Most common field value	1	514248452	930090121224	GSA-FSS-ADV	TN	38118	P	0

### Numeric variables:

Amount	
# of records that have a value	96,753
% of record populated	100%
# of records with null value	0
Mean value	427.88
Standard deviation	10,006.14
Minimum value	0.01
Maximum value	3,102,045.53

### Date variables:

Date	
# of records that have a value	96,753
% of record populated	100%
# of records with null value	0
Minimum value	2006-01-01
Maximum value	2006-12-31
Most Common	2006-02-28

### 3) Data Field Exploration

#### a. Field 1

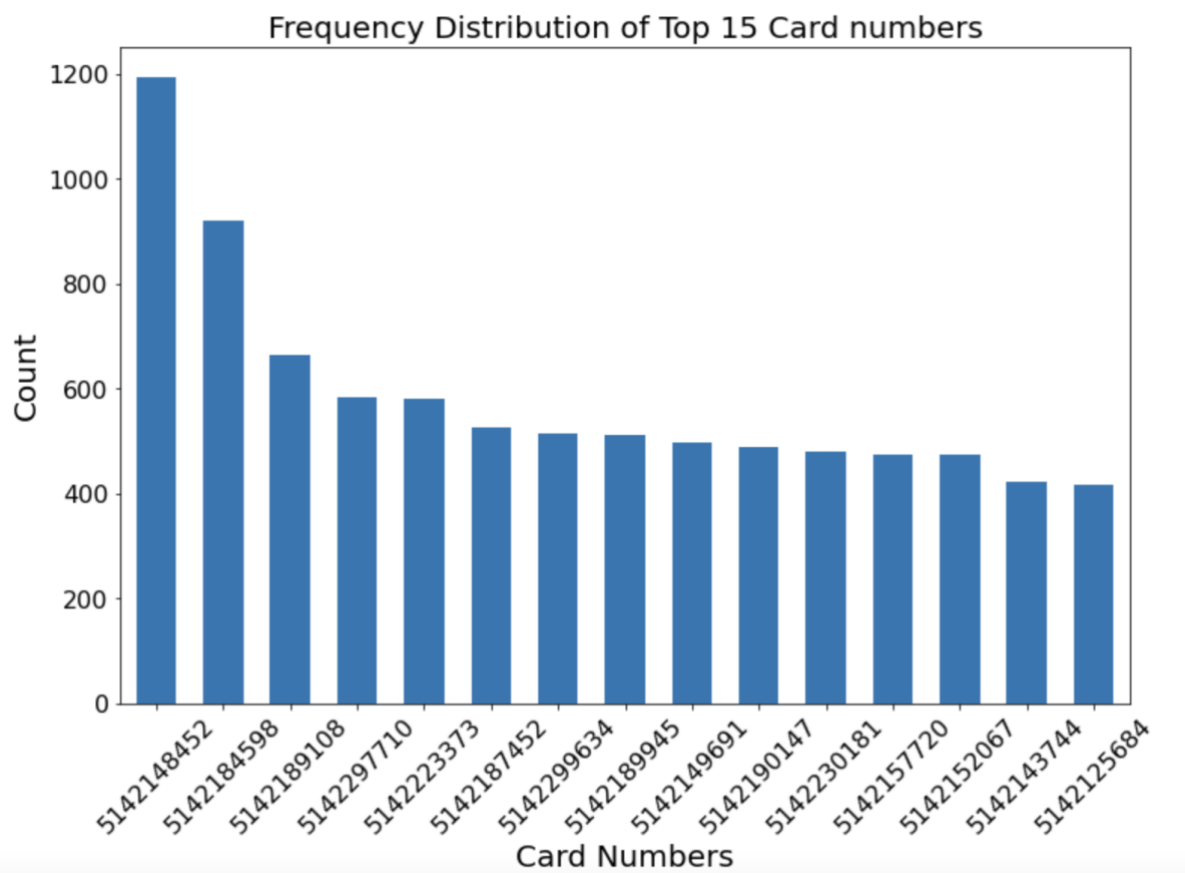
Name: Recnum

Description: Unique record number identifier for the transaction. All record numbers are unique.

#### b. Field 2

Name: Cardnum

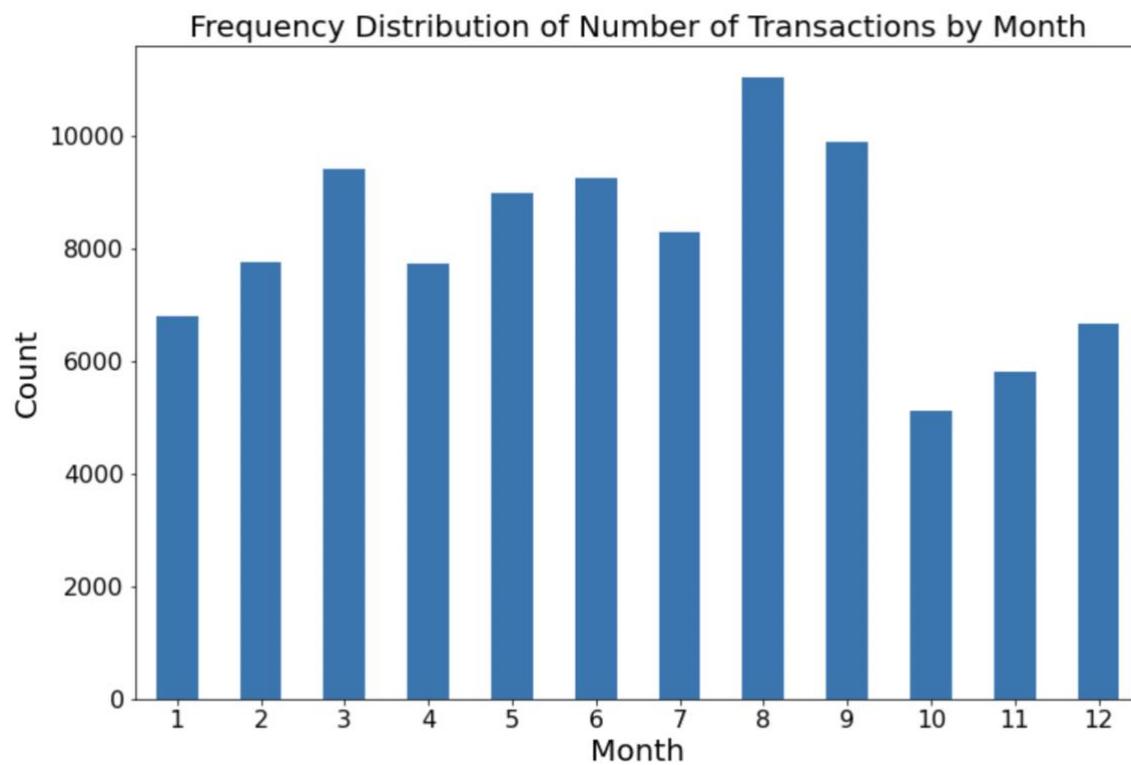
Description: Credit card number used for the transaction



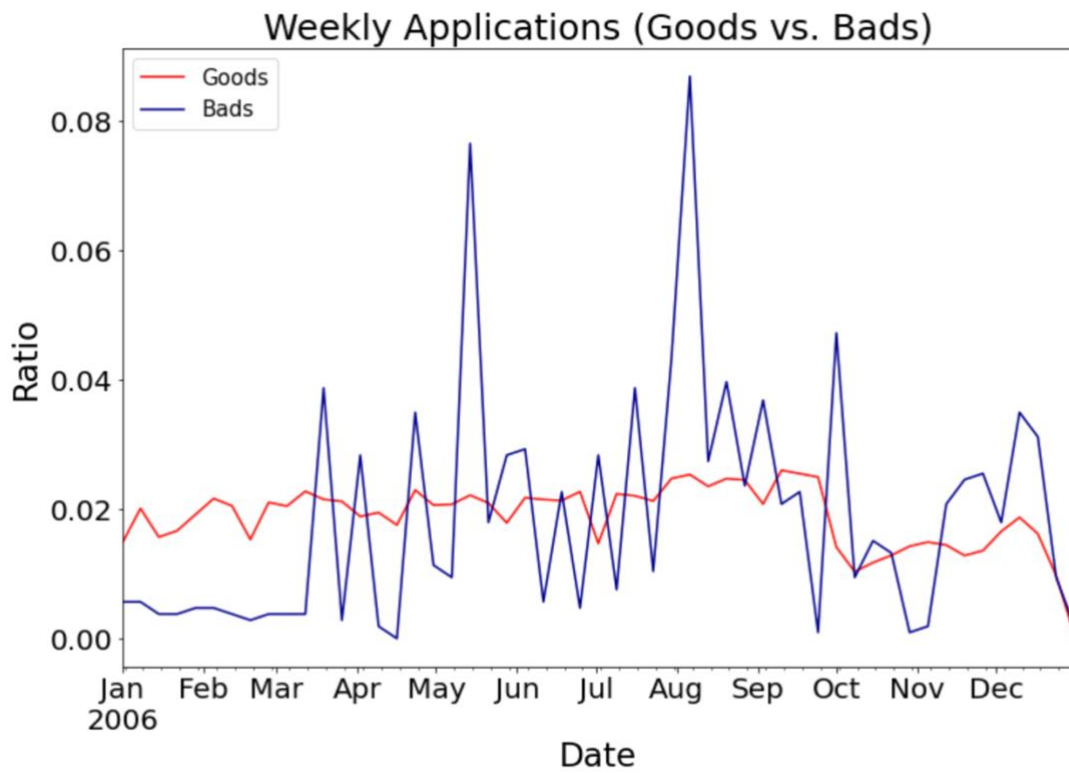
**c. Field 3**

Name: Date

Description: Contains date of transaction



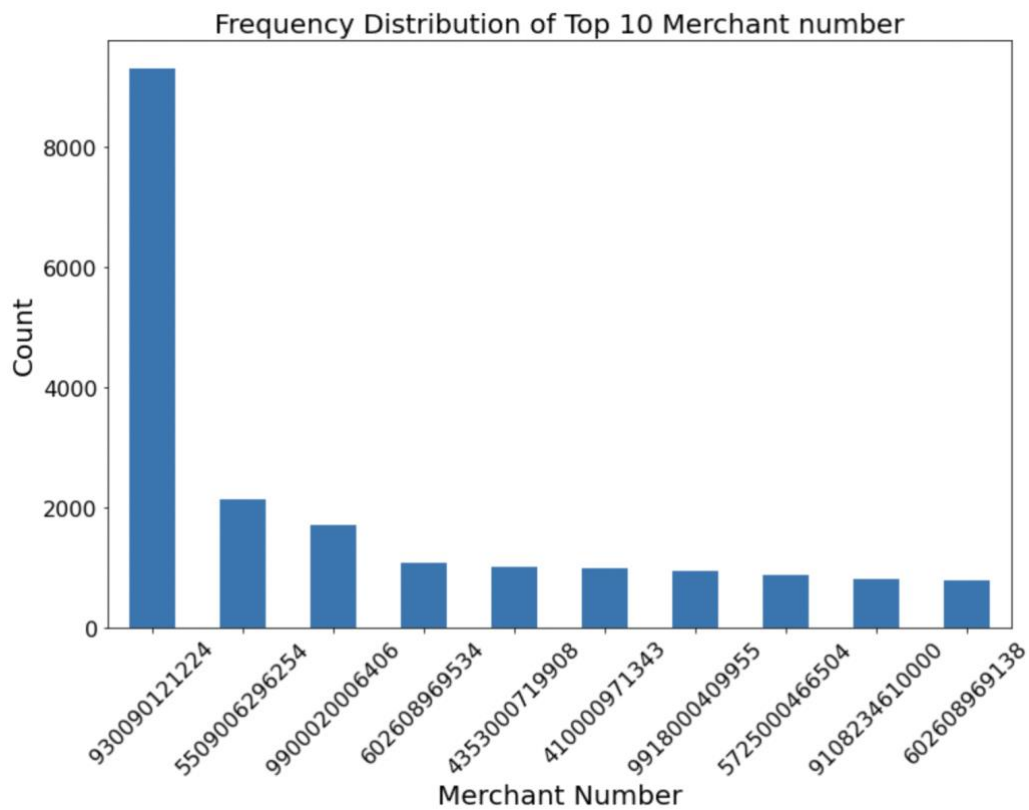
Month	Count
1	6,810
2	7,756
3	9,421
4	7,731
5	8,982
6	9,249
7	8,296
8	11,050
9	9,895
10	5,109
11	5,801
12	6,653



**d. Field 4**

Name: Merchnum

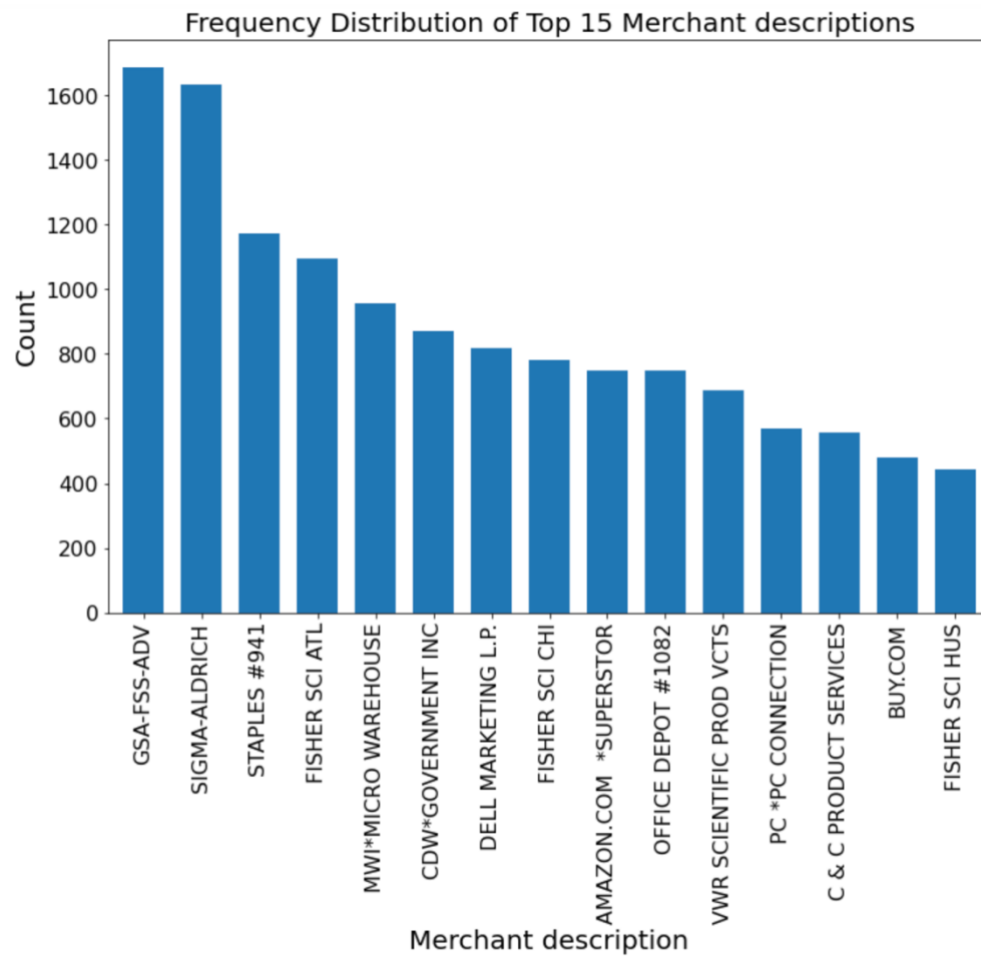
Description: Contains unique identifier of merchant involved in the transaction



**e. Field 5**

Name: Merch description

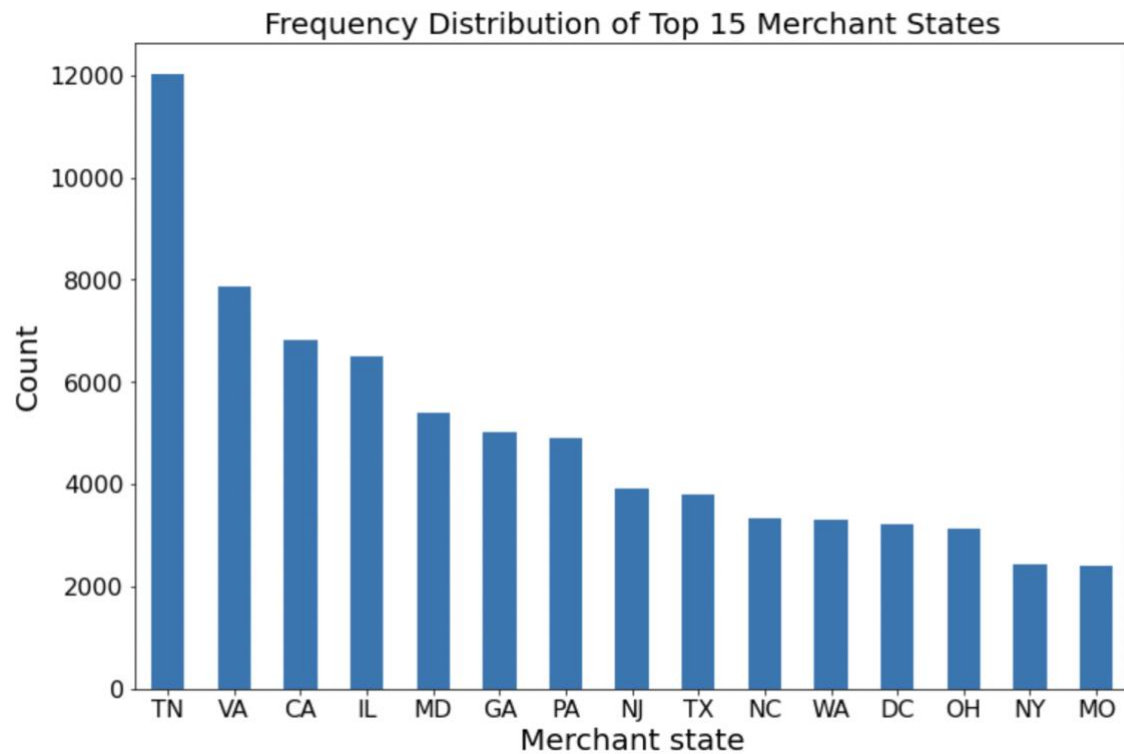
Description: Contains description of merchant involved in transaction



**f. Field 6**

Name: Merch state

Description: Contains state of merchant involved in transaction

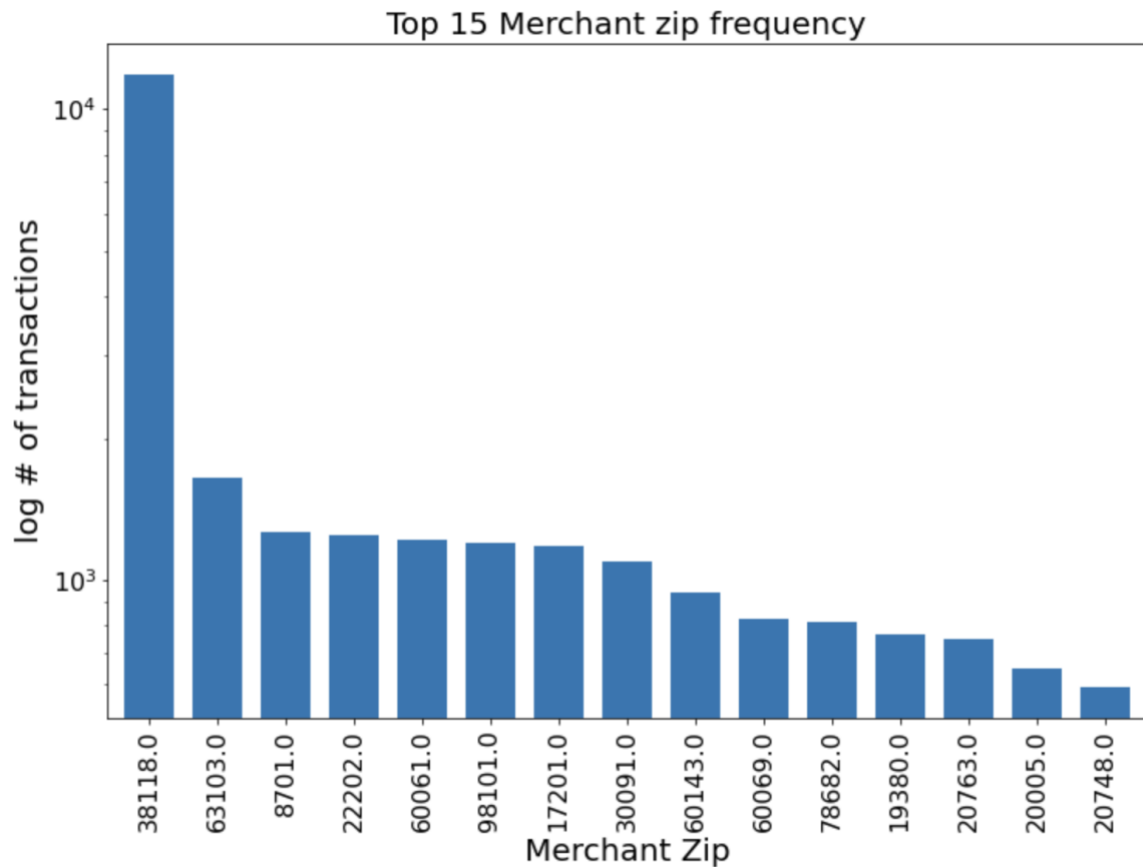


State	Count
TN	12,035
VA	7,872
CA	6,817
IL	6,508
MD	5,398
GA	5,025
PA	4,899
NJ	3,912
TX	3,790
NC	3,322
WA	3,300
DC	3,208
OH	3,131
NY	2,430
MO	2,420

**g. Field 7**

Name: Merch zip

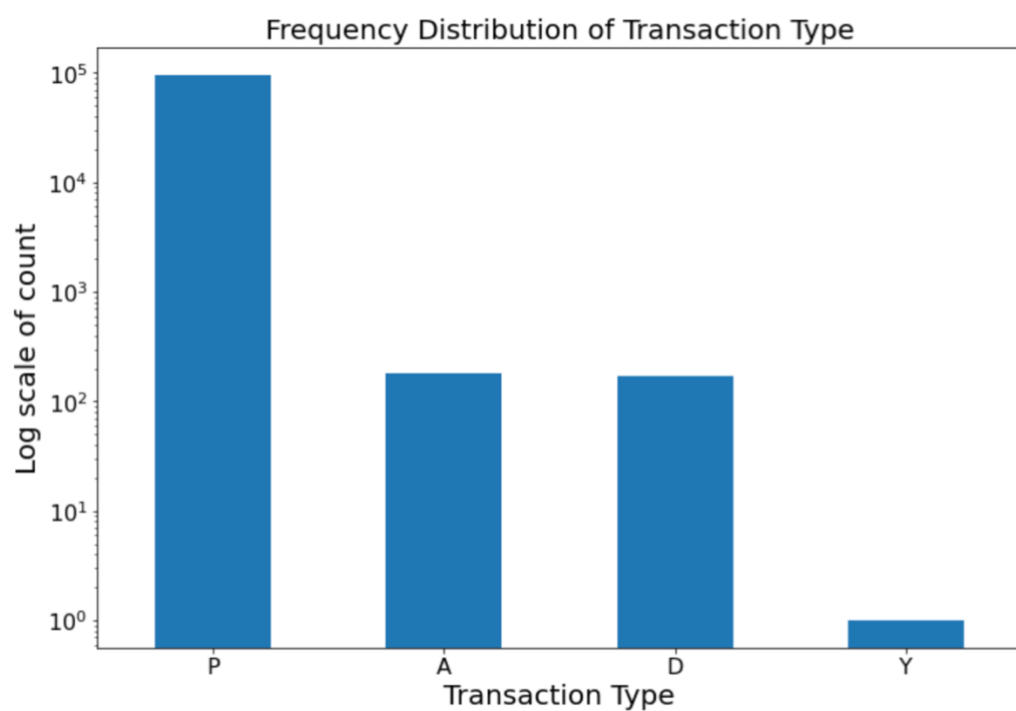
Description: Contains ZIP code of merchant involved in transaction



**h. Field 8**

Name: Transtype

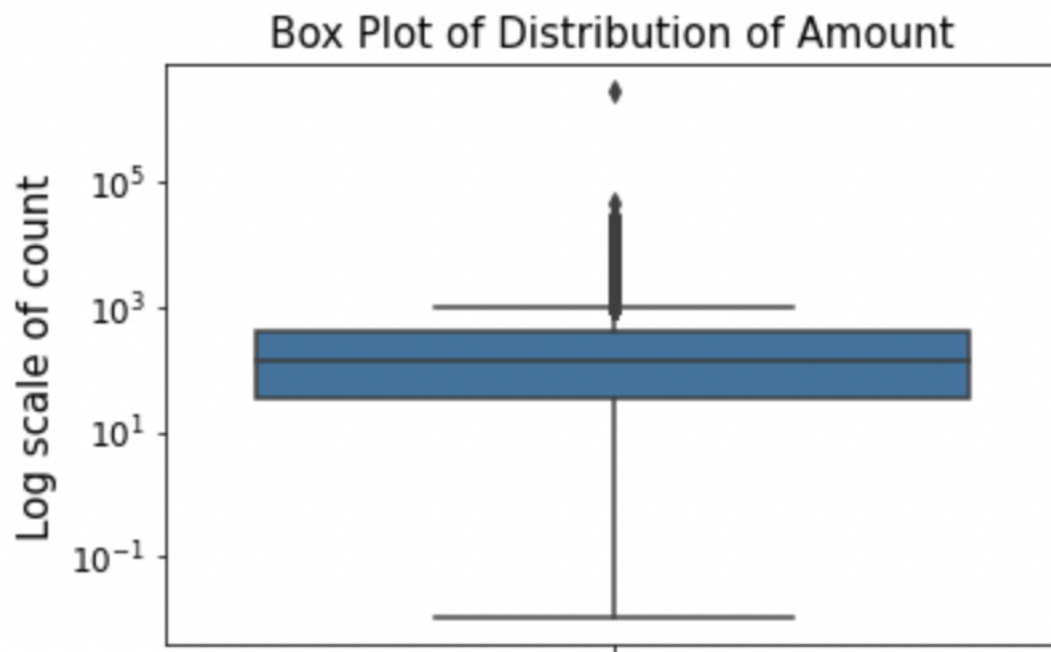
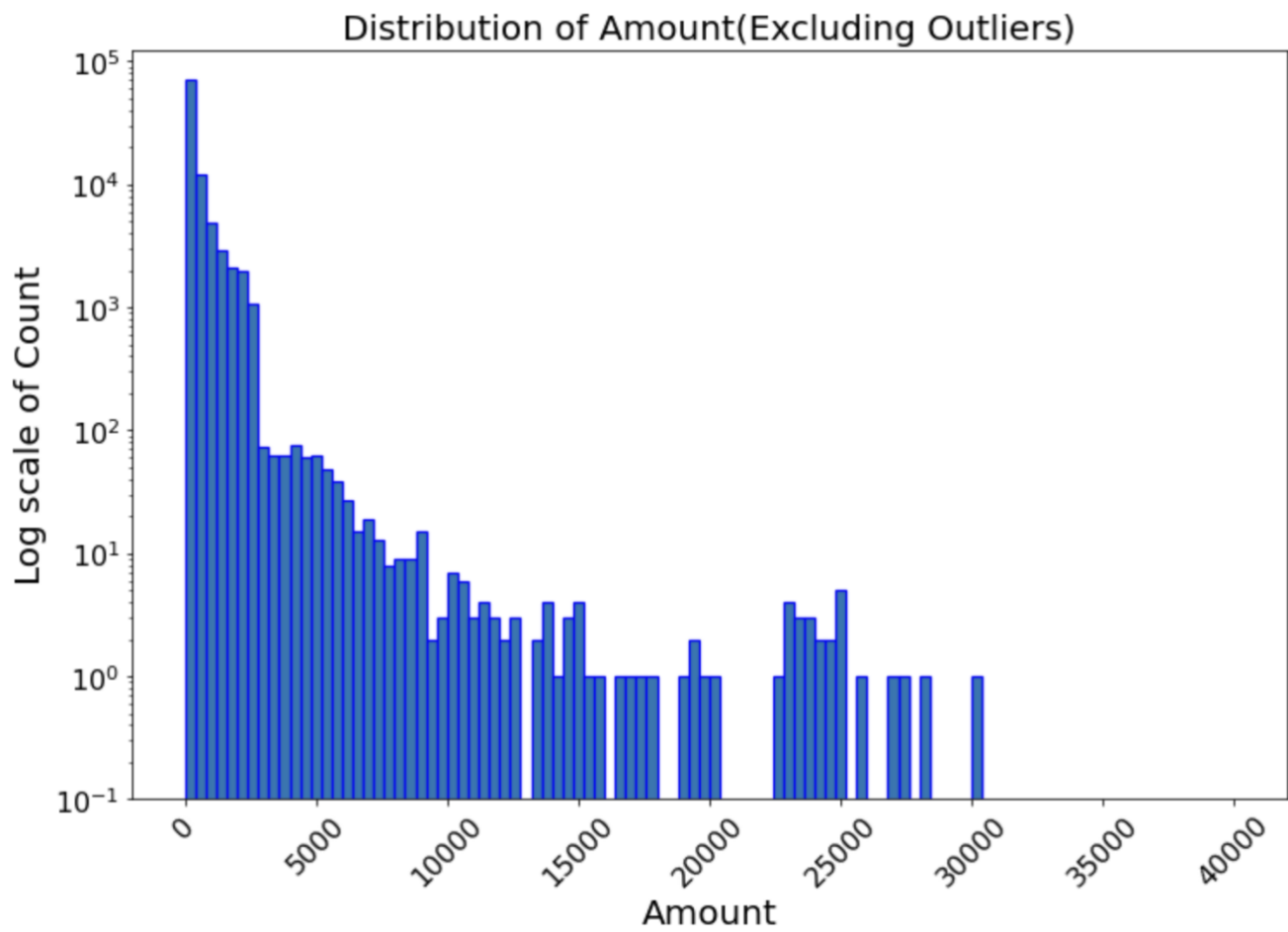
Description: Contains the type of transaction – P, A, D, Y



**i. Field 9**

Name: Amount

Description: Contains the total value of the transaction in \$





j. **Field 10**

Name: Fraud

Description: Target field containing '1' or '0' for each application.

0 – non-Fraud transaction

1 – Fraud transaction



Fraud Label	Count
0 (Non-Fraud Applicant)	95,694
1 (Fraud Applicant)	1,059