DATA QUALITY REPORT

For Card Transaction Data

Prepared by: Chen, Siqi

Prepared Time:02/27/2019

Dataset Summary

What the Data is:
Card Transaction Data

Where the Data came from:
Internet

The Time the Data Covers:
In this file, the card transactions happened from 2010-01-01 to 2010-12-31.

Number of Fields
There are 10 fields in the data, 8 categorical fields, 1 numeric field, and 1 date/time field.

Number of Records:
There are 96753 records in the data.

Field Summary

For Categorical and Date/Time fields:

Field name	Field type	# records that have a value	% populated	# unique values	# records with a value zero	Most common field value
Recnum	Categorical	96753	100.00%	96753	0	1
Cardnum	Categorical	96753	100.00%	1645	0	5142148452
Date	Date/Time	96753	100.00%	365	0	2/28/2010 00:00:00
Merchnum	Categorical	93378	96.51%	13092	0	930090121224
Merch description	Categorical	96753	100.00%	13126	0	GSA-FSS-ADV
Merch state	Categorical	95558	98.76%	228	0	TN
Merch zip	Categorical	92097	95.19%	4568	0	38118
Transtype	Categorical	96753	100.00%	4	0	P
Fraud	Categorical	96753	100.00%	2	95694	0

For Numeric Field:

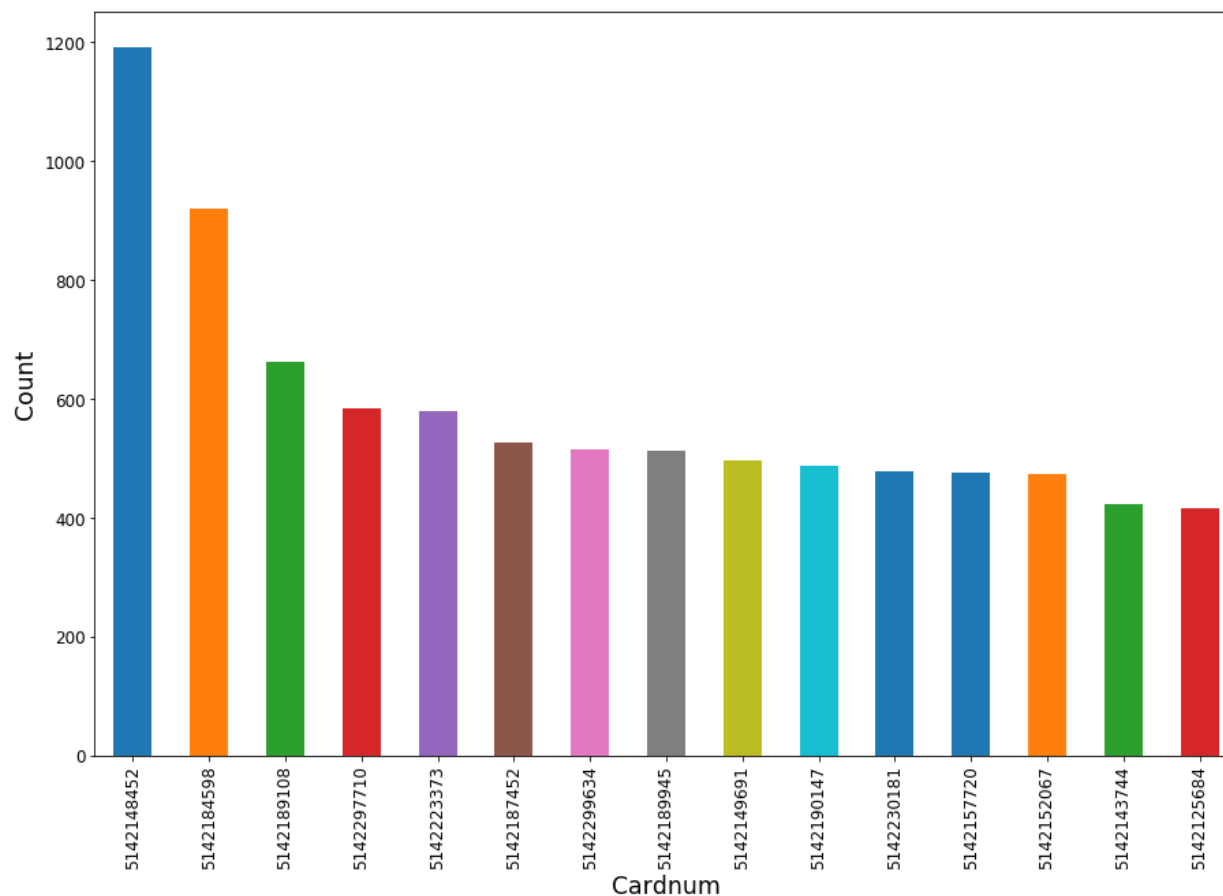| Field name	Field type	# records that have a value	% populated	# unique values	# records with a value zero	mean	SD	min	max
Amount	Numeric	96753	100.00%	34909	0	427.8857	10006.14	0.01	3102046

Step by Step Description

I. Recnum

Recnum represents the numbering for each record. There are 1070994 unique values in this field. Each record has a corresponding Recnum.
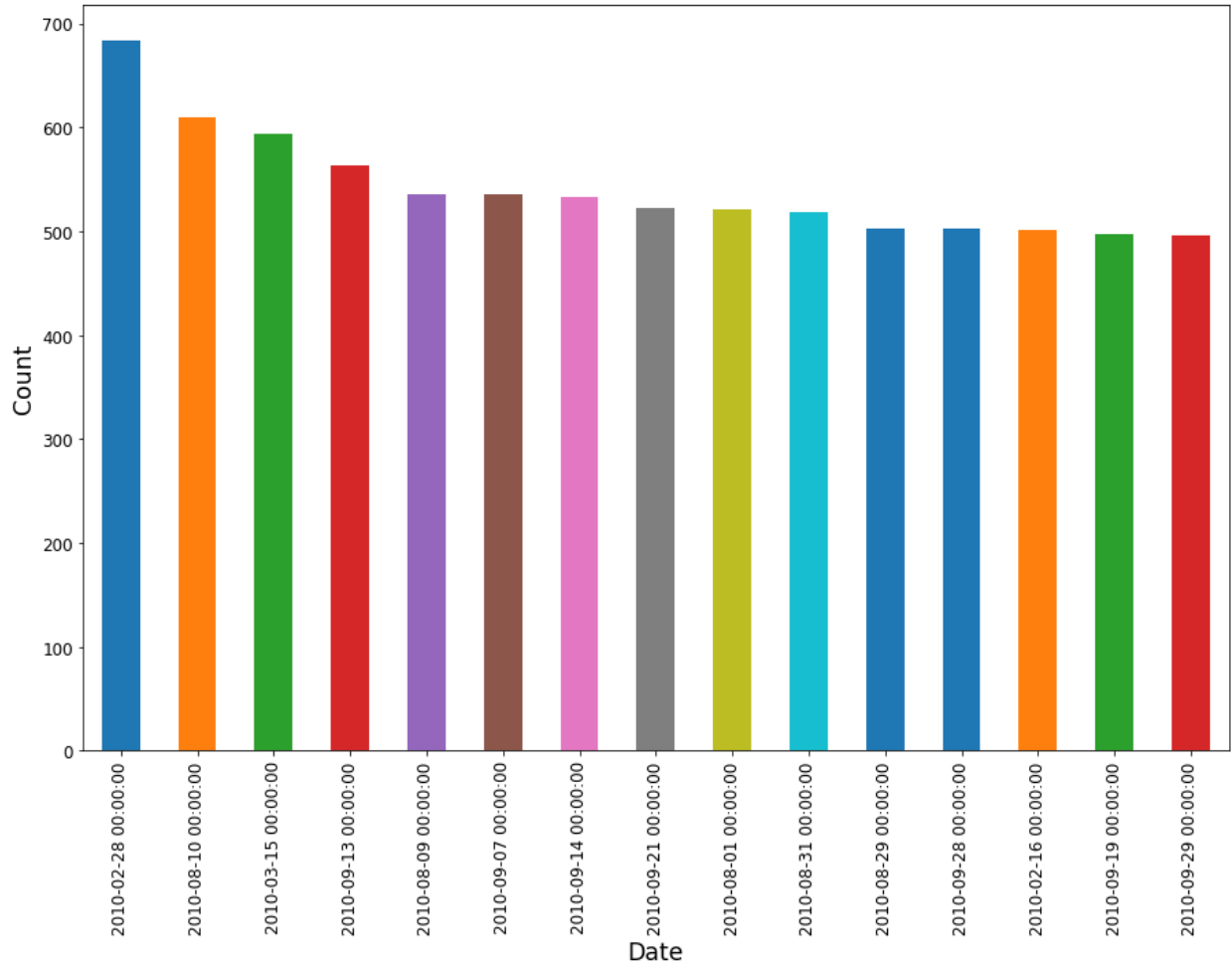
II. Cardnum

Cardnum is the card number of the card which is used to pay the transaction. The bar chart below shows the count of top 15 most common values in Cardnum.
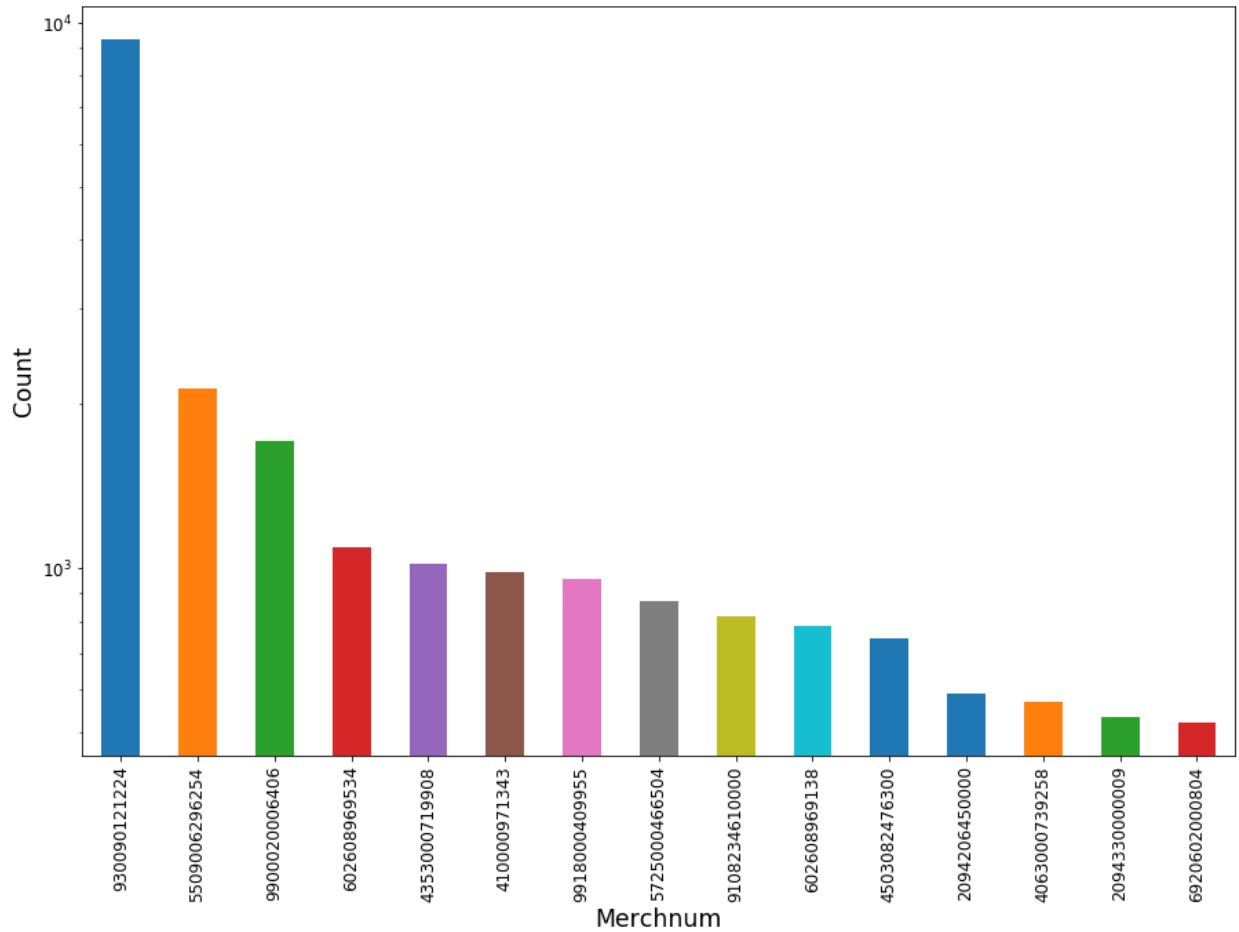
III. Date

Date is the data the credit card transaction happened. The bar chart below shows the count of top 15 most common values in Date.
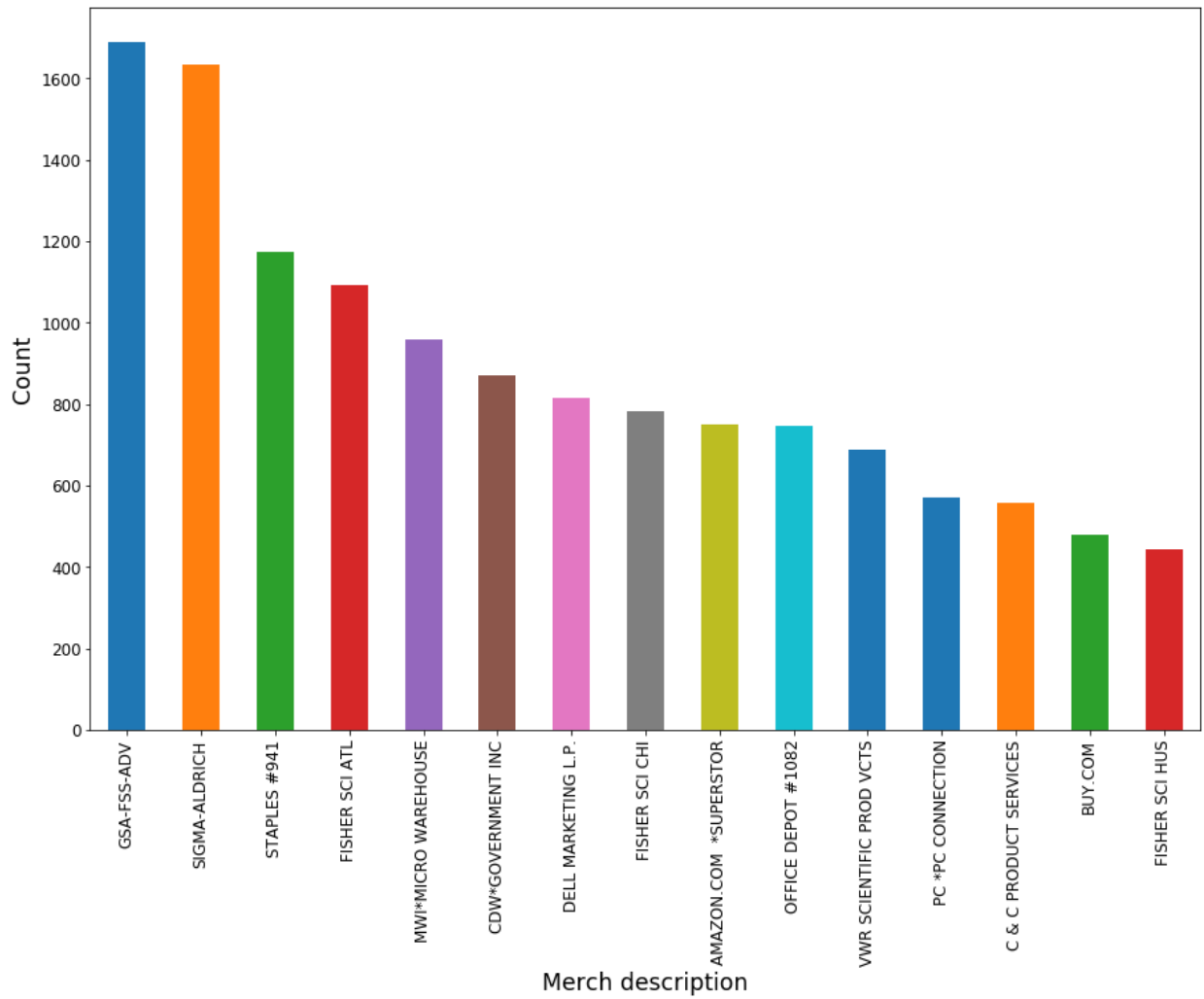
IV. Merchnum

Merchnum shows at which merchant the credit card transaction happened. The bar chart below shows the log scale count of top 15 most common values in Merchnum.
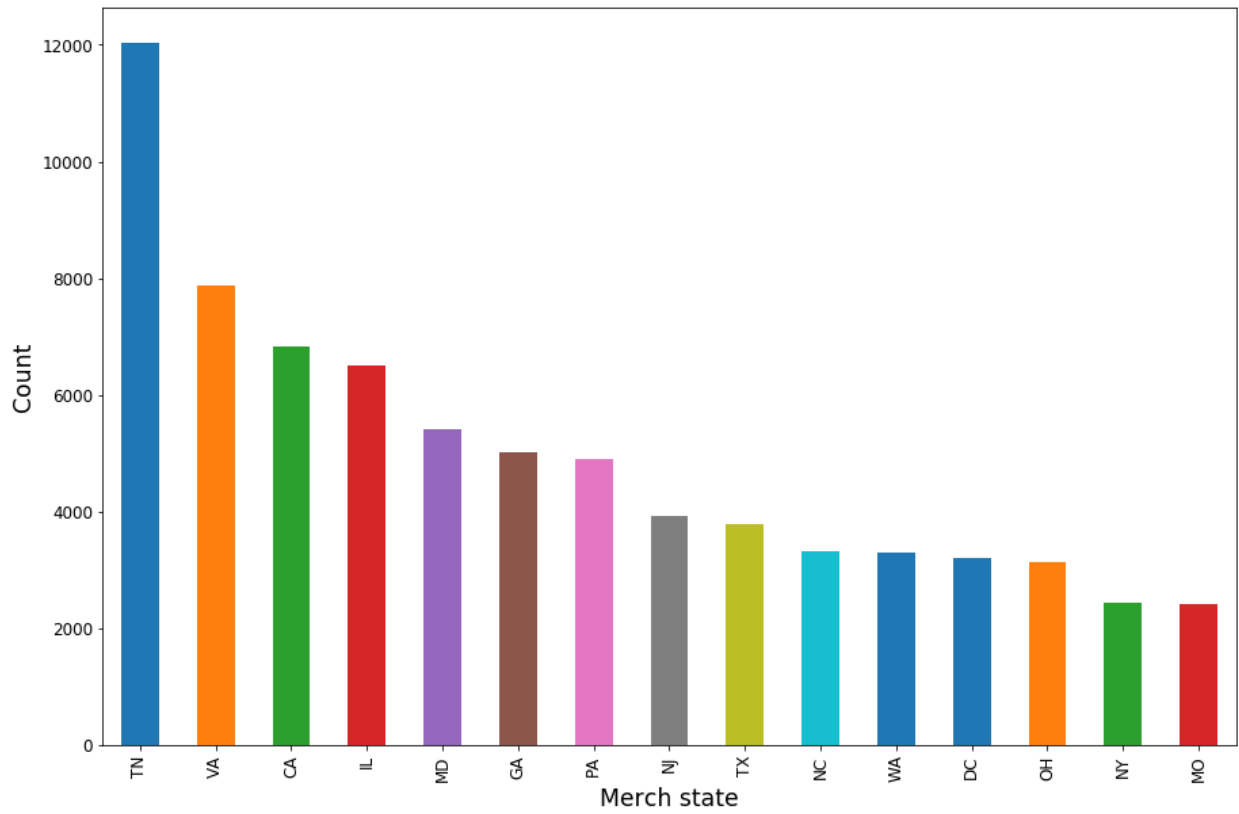
V. Merch description

Merch description is a brief description of the merchant. The bar chart below shows the count of top 15 most common values in Merch description.
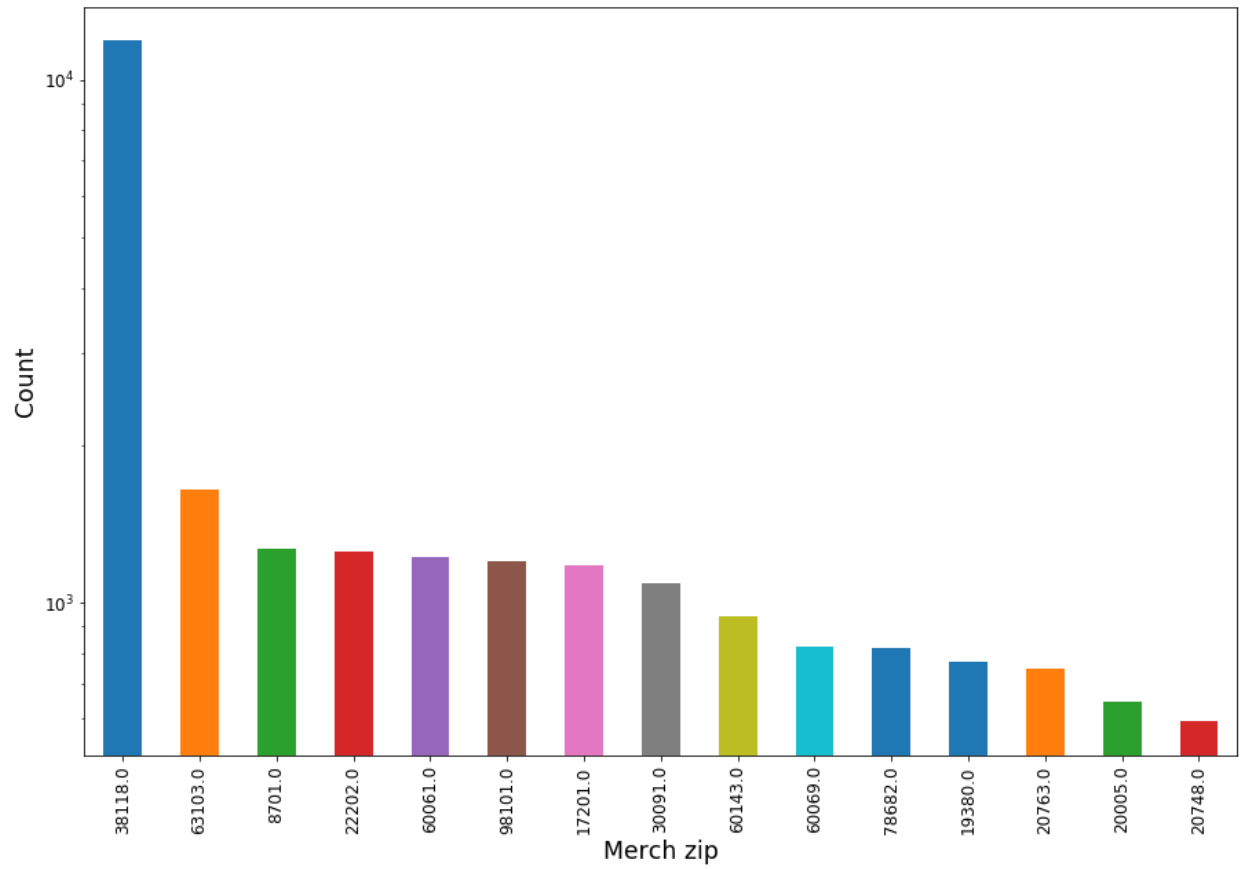
VI. Merch state

Merch state is the state which merchant locates in. The bar chart below shows the count of top 15 most common values in Merch state.
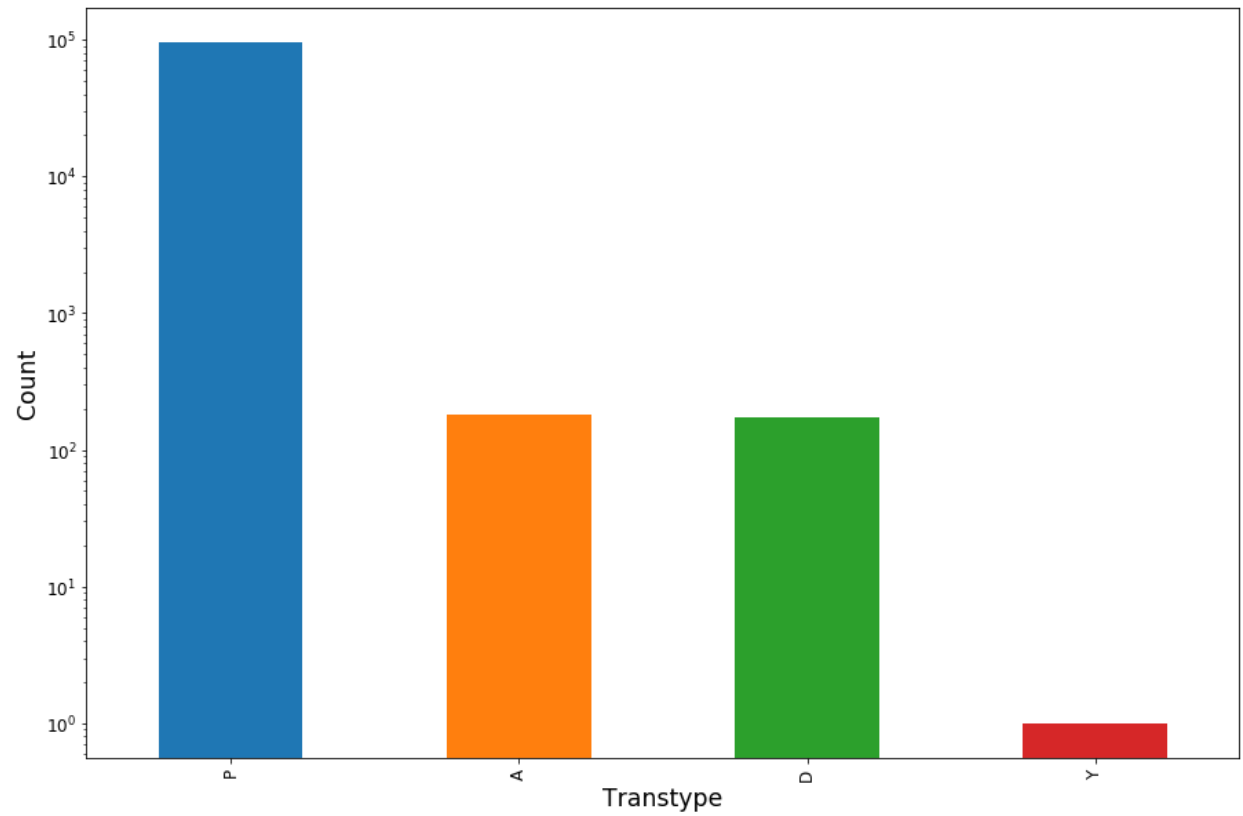
VII.   Merch zip

Merch zip represents the zip code of merchant. The bar chart below shows the log scale count of top 15 most common Merch zip.
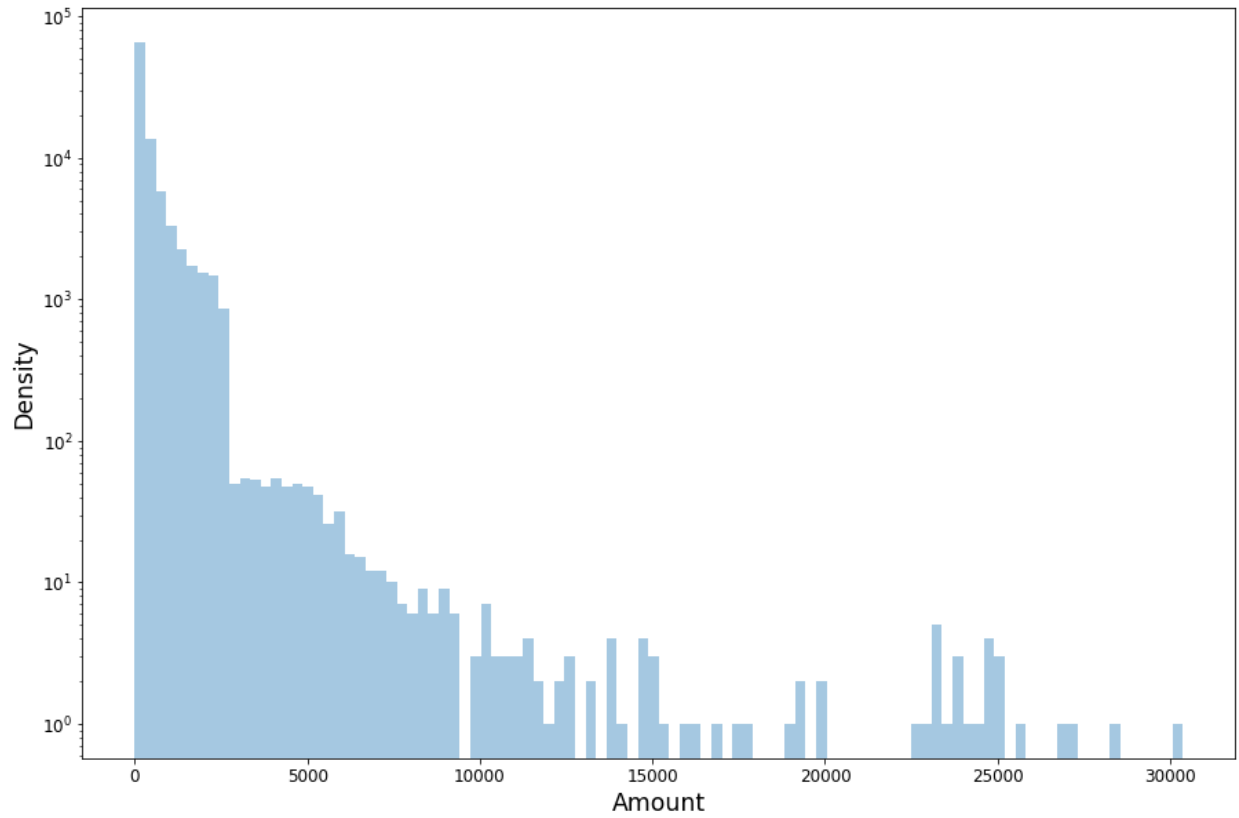
VIII.   Transtype

Transtype is the type of credit card transaction. The bar chart below shows the log scale count of each Transtype.

IX. Amount

Amount is amount of money paid in the credit card transaction. The bar chart below shows the log scale distribution of Amount with value less than 32000.

X. Fraud

Fraud indicates whether the credit card transaction is fraud or not. The bar chart below shows the log scale count of all categories in Fraud.