

Data Quality Report (DQR)

Made by Yiyang Duanmu

Credit Card Transactions Data

The Credit Card Transactions Data is a dataset of actual credit card purchases from a US government organization. It covers the whole year of 2006 (2006-01-01 – 2006-12-31), consisting of 10 fields with 96,753 records.

SUMMARY TABLES

TABLE FOR NUMERIC FIELDS

Field Name	% Populated	Min	Max	Mean	Stdev	% Zero
Date	100	2006-01-01	2006-12-31	N/A	N/A	0
Amount	100	0.01	3,102,045.53	427.89	10,006.14	0

TABLE FOR CATEGORICAL FIELDS

Field Name	% Populated	# Unique Values	Most Common Value
Recnum	100	96,753	N/A
Cardnum	100	1,645	5142148452
Merchnum	96.51	13,091	930090121224
Merch description	100	13,126	GSA-FSS-ADV
Merch state	98.76	227	TN
Merch zip	95.19	4,567	38118
Transtype	100	4	P
Fraud	100	2	0

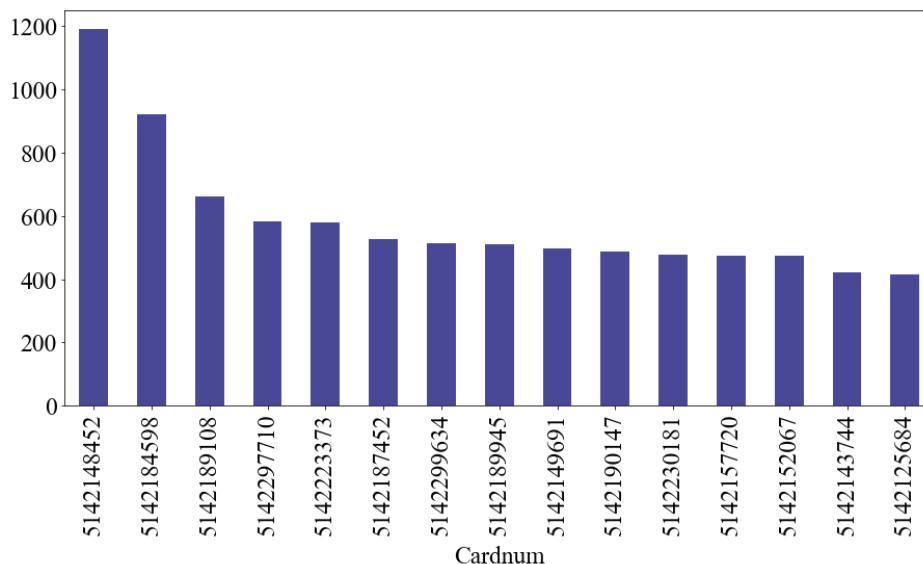
DESCRIPTION OF EACH FIELD

Recnum (Categorical) -- Each record has a unique number.

Cardnum (Categorical)

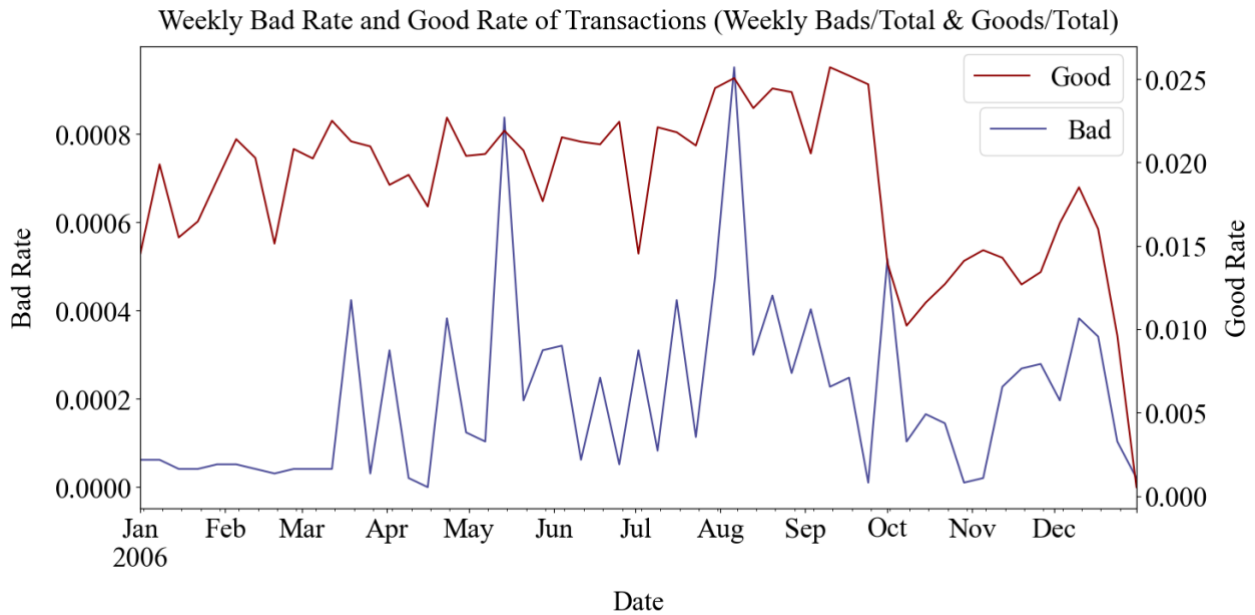
Card number. There are 1,645 unique card number and the most common value is 5142148452.

Bar Plot of Cardnum (Top 15)



Date (Numeric - Date)

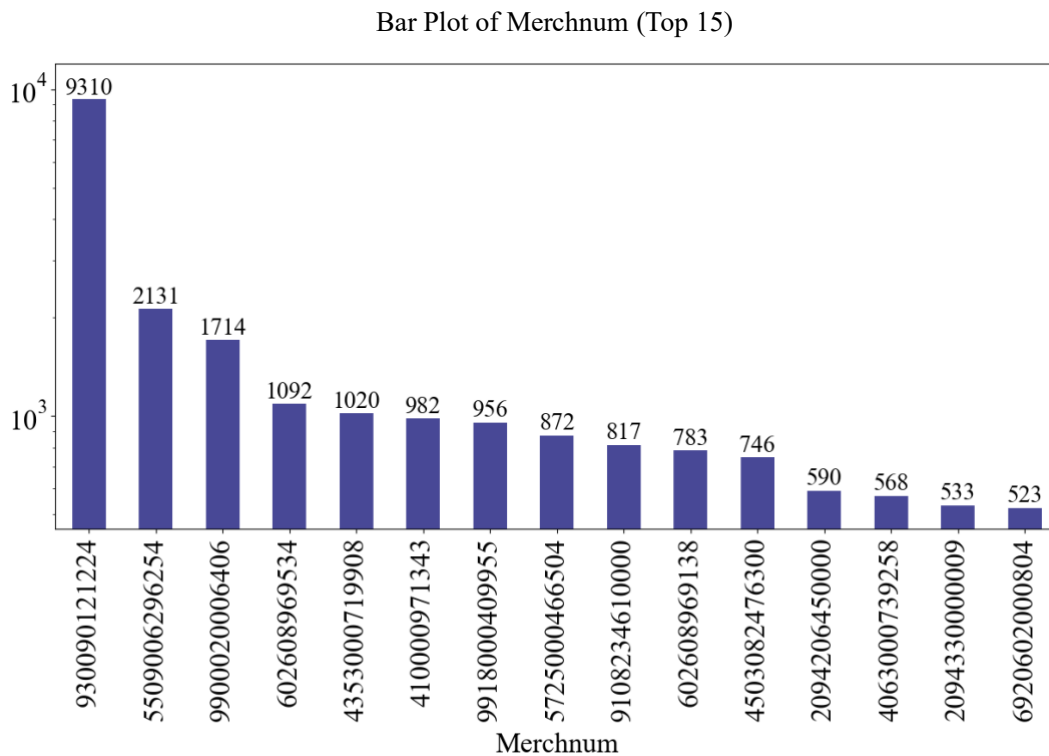
Date. The minimum date is 2006-01-01 and the maximum date is 2006-12-31.



Note: Ignore the last plot point on above plot.

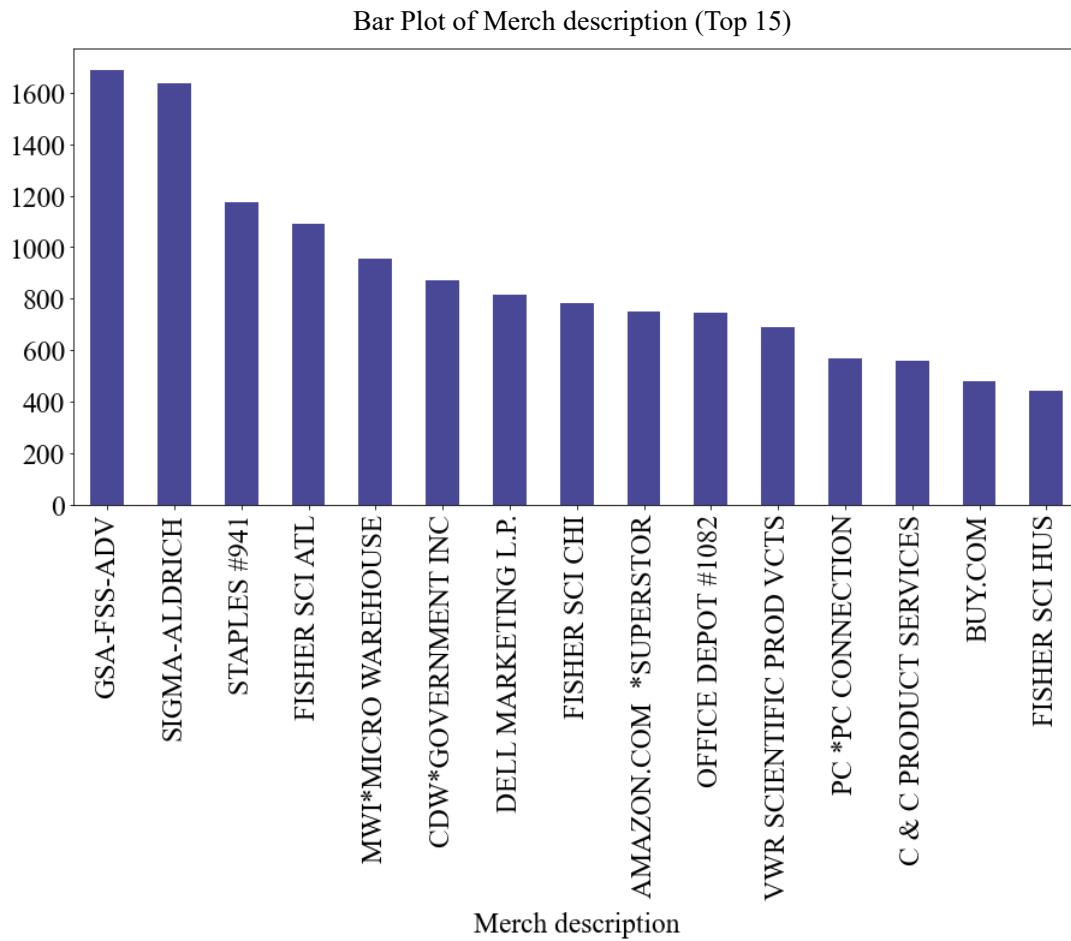
Merchnum (Categorical)

Merchant number. There are 13,091 unique Merchnums and the most common value is 930090121224.



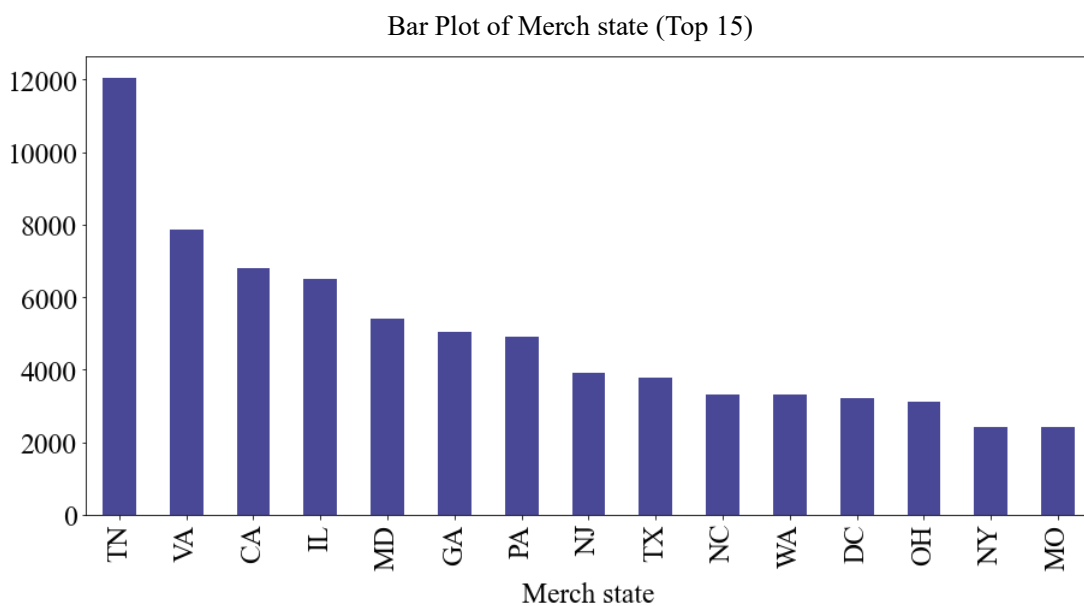
Merch description (Categorical)

Description of Merchants. There are 13,126 unique descriptions of Merch and the most common value is GSA-FSS-ADV.



Merch state (Categorical)

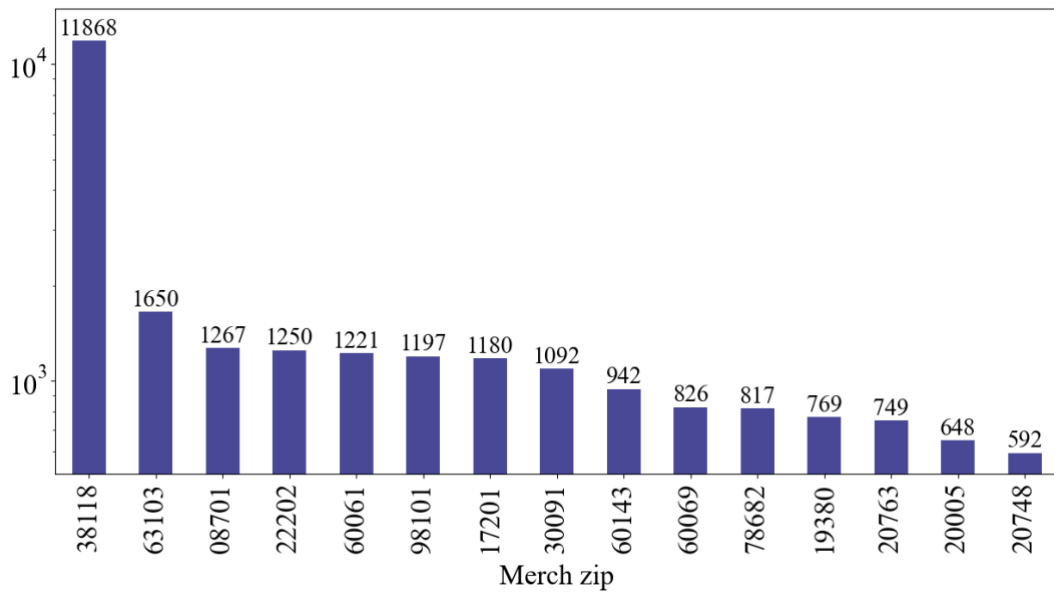
State of Merch. There are 227 unique values of states and the most common value is TN.



Merch zip (Categorical)

5-digit Zip code of Merch. There are 4,567 unique values of zip codes and the most common value is 38118.

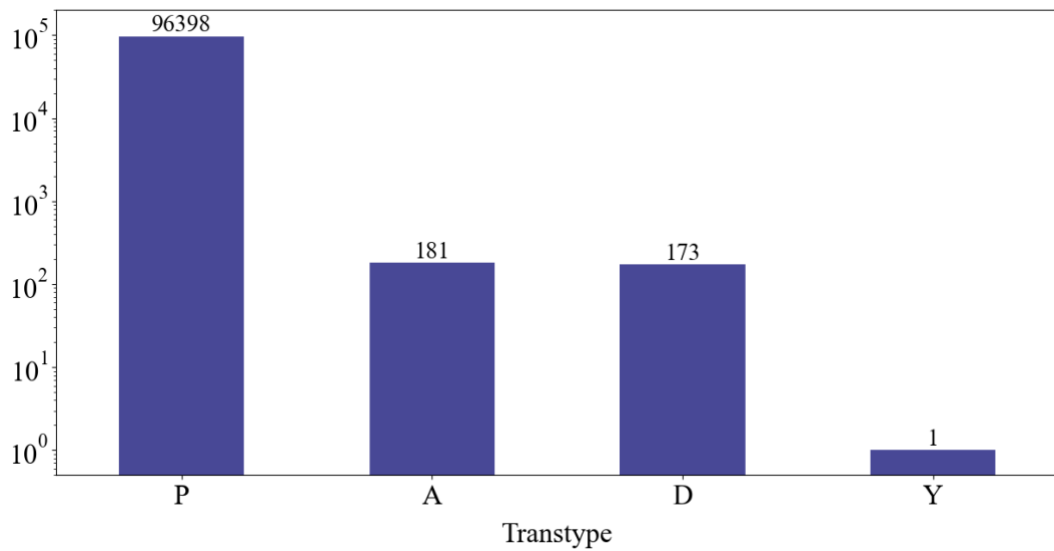
Bar Plot of Merch zip (Top 15)



Transtype (Categorical)

Transaction Type. There are 4 unique transaction types and the most common value is P.

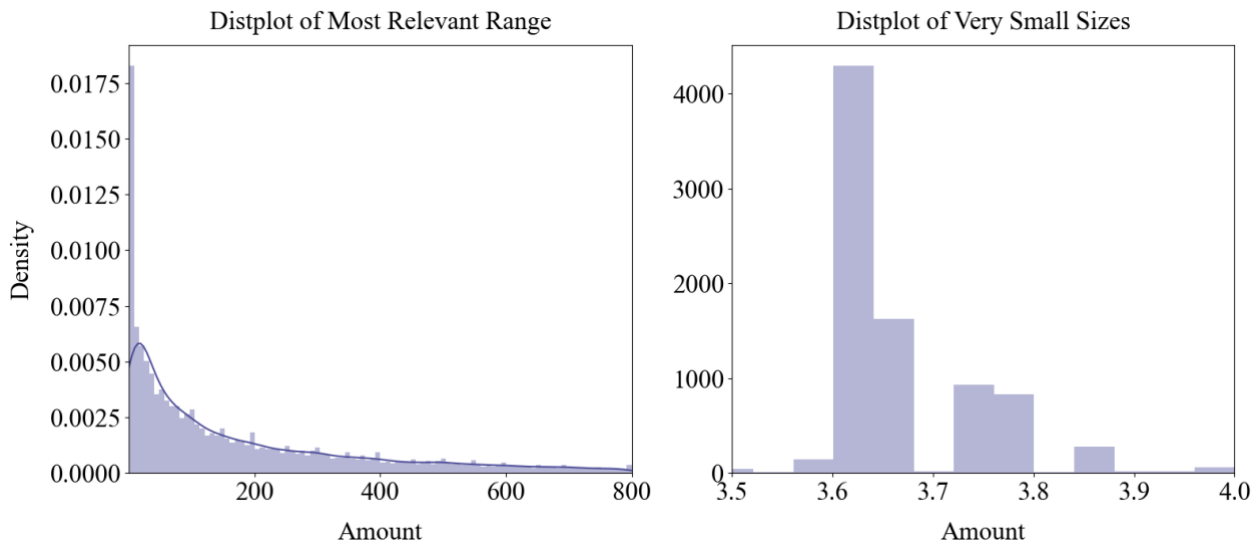
Bar Plot of Transtype



Amount (Numeric)

Amount. The minimum value is 0.01 and the maximum value is 3,102,045.53.

Field Name	% Populated	Min	Max	Mean	Stdev	% Zero
Amount	100	0.01	3,102,045.53	427.89	10,006.14	0



Fraud (Categorical)

Label of Fraud. There are 2 unique values where 1 means fraud and 0 means no fraud. The most common value is 0.

