

DATA QUALITY REPORT

1. Dataset Overview

Dataset Name: Card transactions data

Description: This dataset contains the information of the card transactions that have occurred in USA. It contains fields like Card number, merchant number, merchant description, and amount of the transaction. It also contains a fraud label field which tells whether the transaction is good or bad.

Time Period: 1 January 2010 – 31 December 2010

No. of Fields: 10

No. of Records: 96,753

Size of Dataset file: 7 MB

2. Summary Table

2.a) Numerical

Field Name	# records that have a value	% populated	# unique values	# records with value 0	Min	Max	Mean	Standard Deviation
AMOUNT	96753	100.0%	34909	0	0.01	3102045.5	427.9	10006.1

2.b) Categorical

Field Name	# records that have a value	% populated	# unique values	Most common value
RECNUM	96753	100.0%	96753	None
CARDNUM	96753	100.0%	1645	5142148452
DATE	96753	100.0%	365	2010-02-28
MERCHNUM	96753	96.5%	13092	930090121224
MERCH DESCRIPTION	96753	100.0%	13126	GSA-FSS-ADV
MERCH STATE	96753	98.7%	228	TN
MERCH ZIP	96753	95.2%	4568	38118
TRANSTYPE	96753	100.0%	4	'p'
FRAUD	96753	100.0%	2	'0'

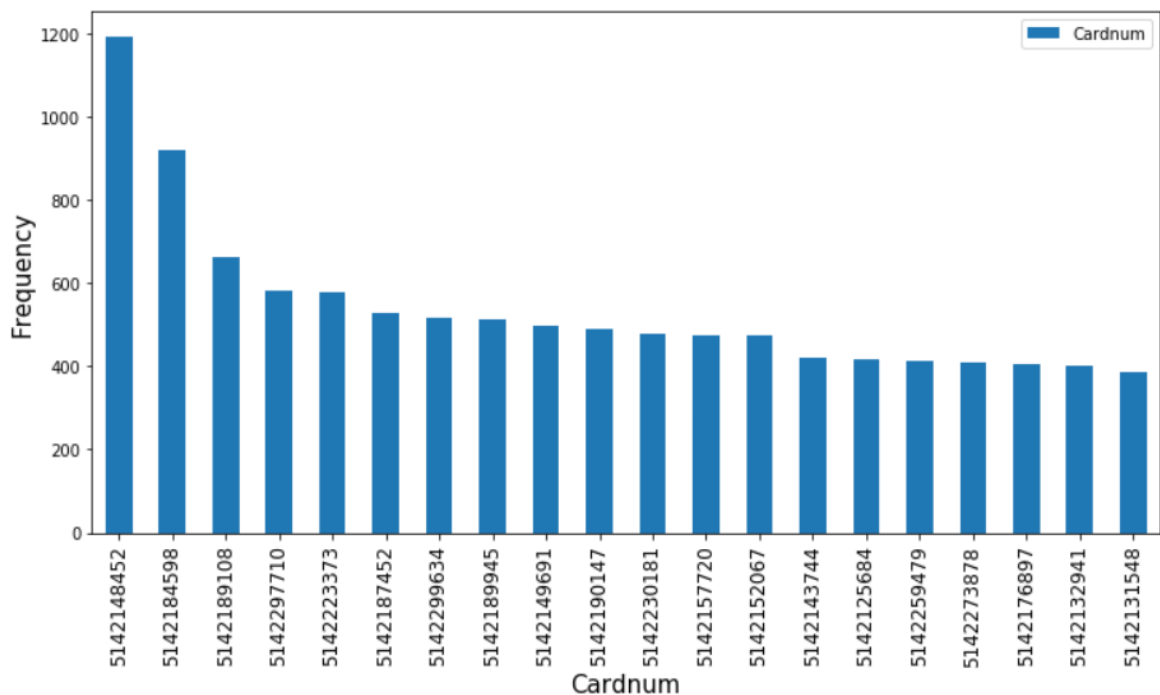
3. DATA FIELD EXPLORATION

3.1. FIELD 1: RECNUM

DESCRIPTION: A categorical field containing unique integer for each record from 1 to 96,753.

3.2. FIELD 2: CARDNUM

DESCRIPTION: A categorical field containing the card number which was used for transaction.



3.3. FIELD 3: DATE

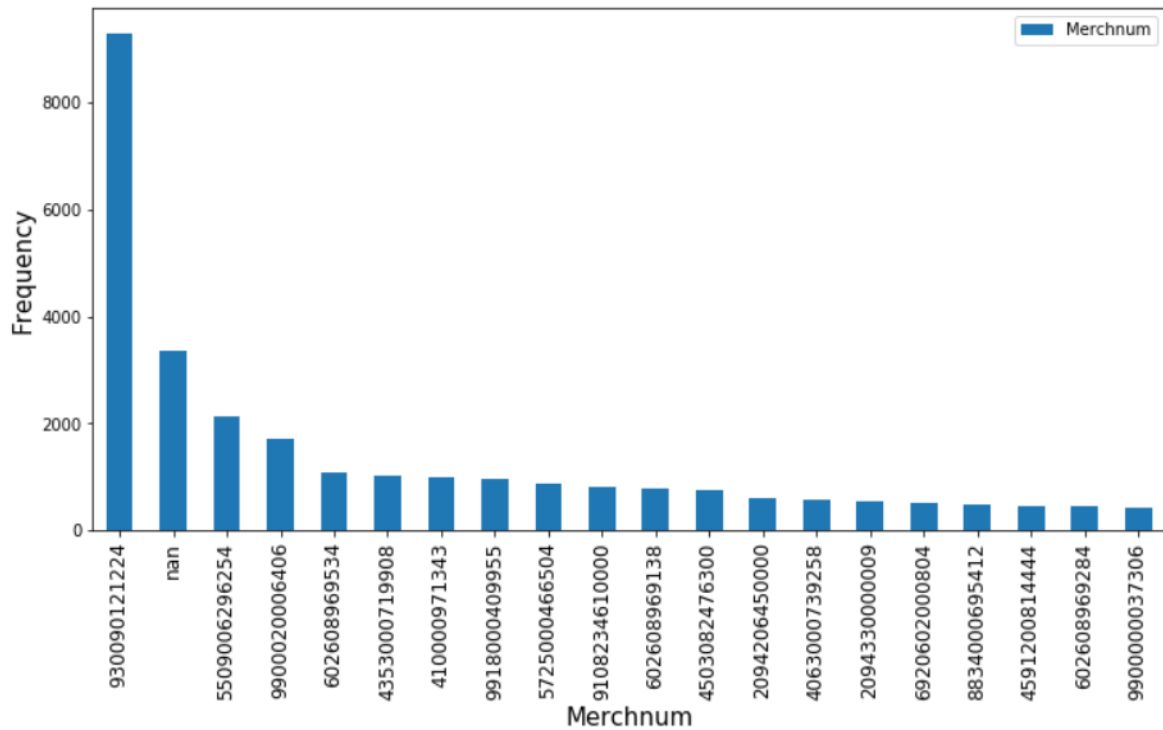
DESCRIPTION: A categorical field containing the date of the transaction when it happened.

Date	Count
02/28/2010	684
08/10/2010	610
03/15/2010	594
09/13/2010	564
08/09/2010	536
09/07/2010	536
09/14/2010	533
09/21/2010	522
08/01/2010	521
08/31/2010	518

Table containing first 10 values of Date

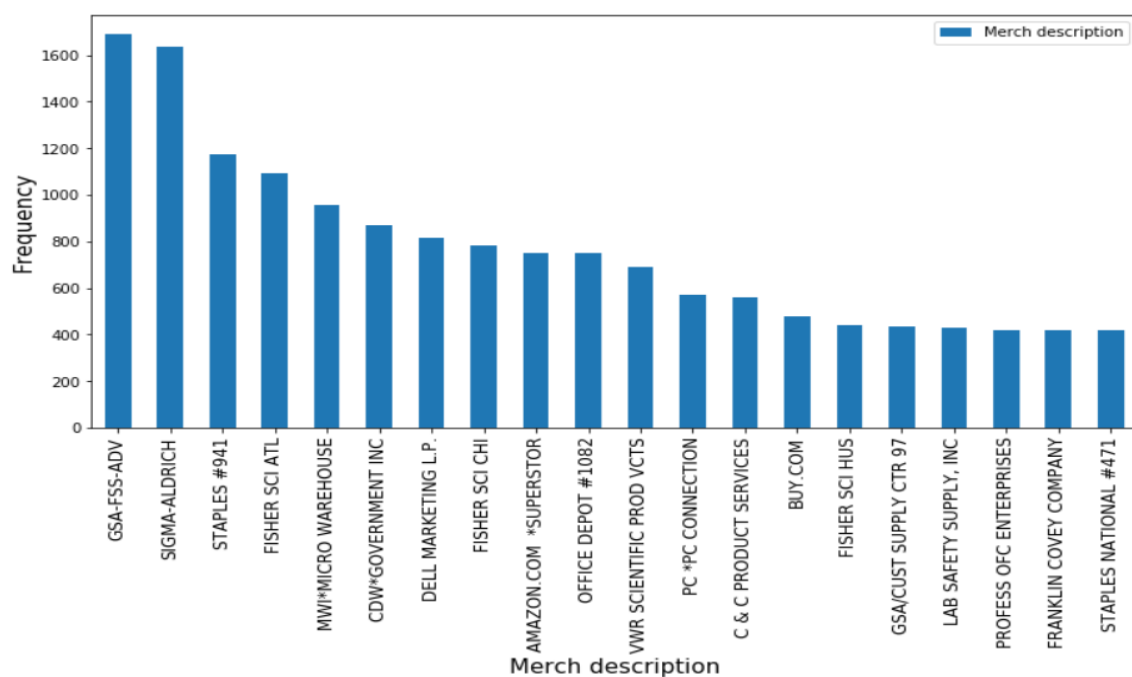
3.4. FIELD 4: MERCHNUM

DESCRIPTION: A categorical field containing the merchant number which helps in identifying the merchant's details.



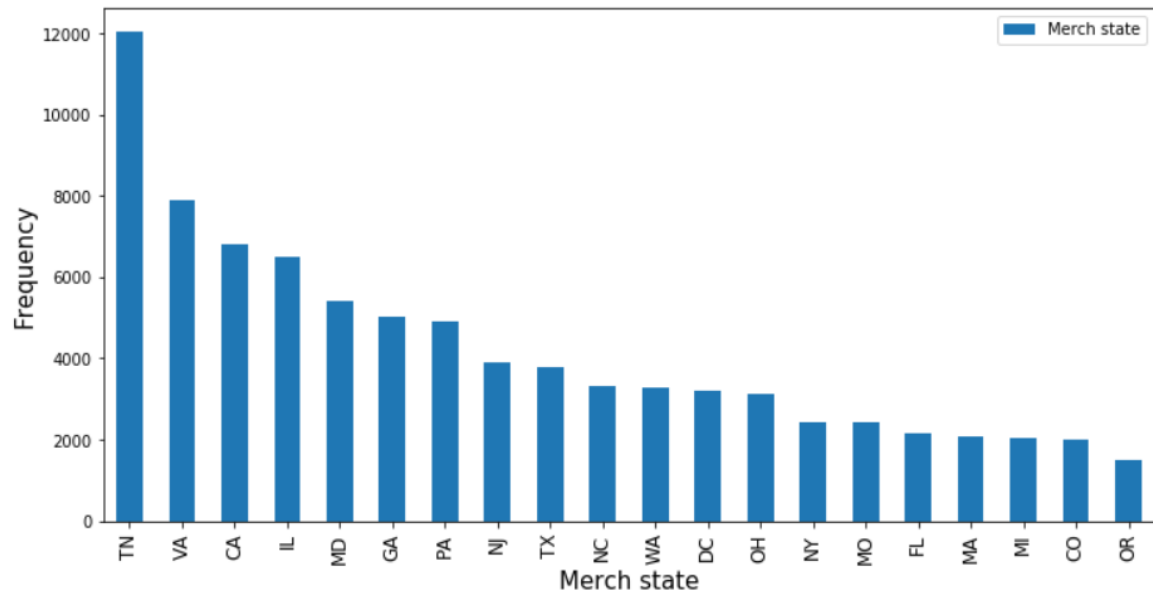
3.5. FIELD 5: MERCH DESCRIPTION

DESCRIPTION: A categorical field containing details of the merchant.



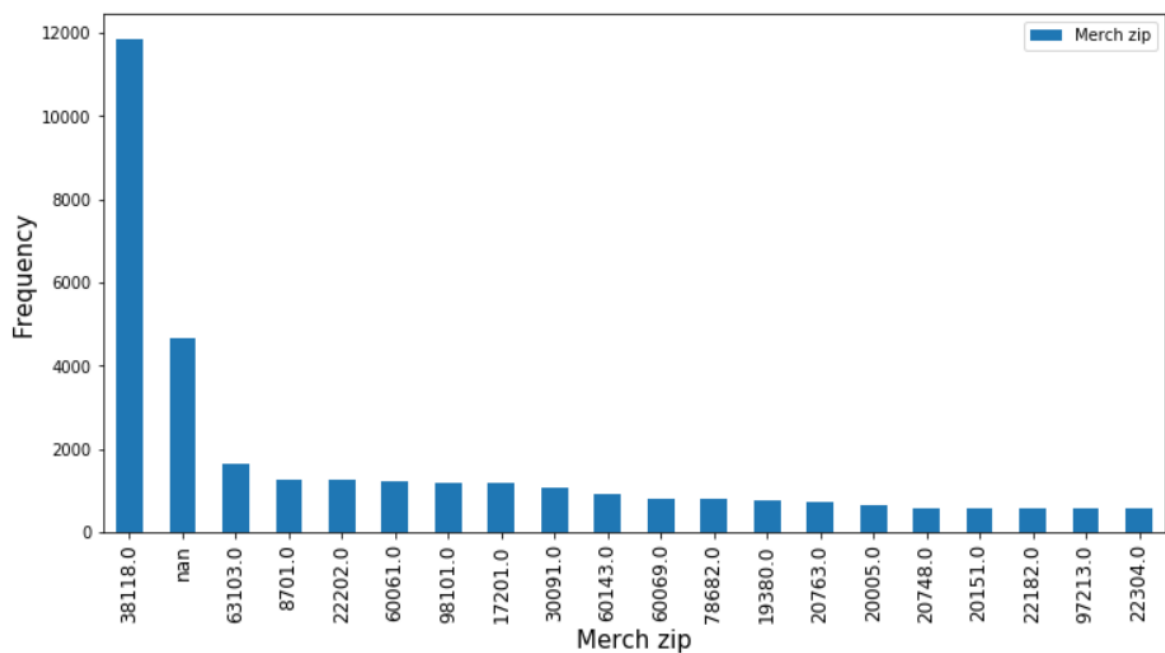
3.6. FIELD 6: MERCH STATE

DESCRIPTION: A categorical field containing the abbreviation of the state where the merchant is based.



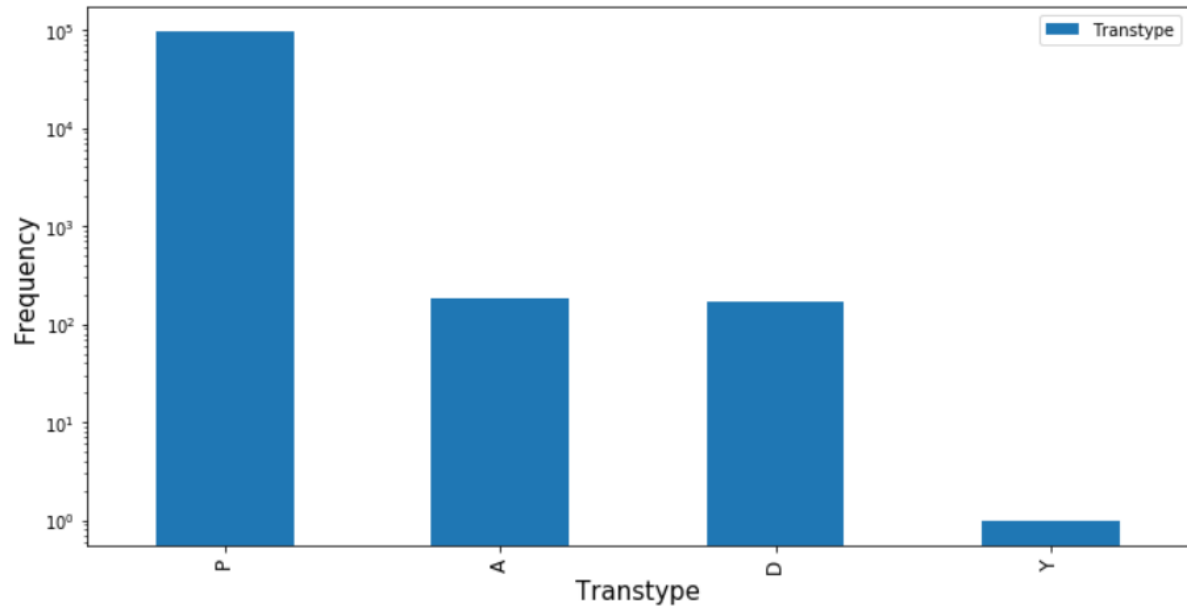
3.7. FIELD 7: MERCH ZIP

DESCRIPTION: A categorical field containing all the ZIP code of the merchant's location.



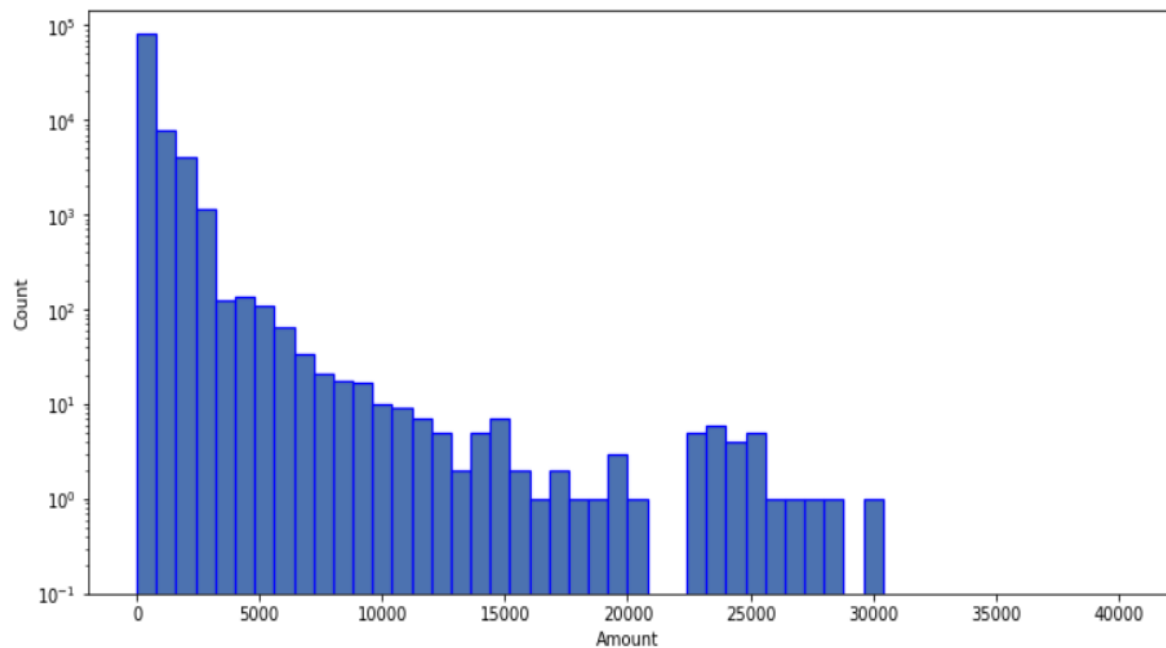
3.8. FIELD 8: TRANSTYPE

DESCRIPTION: A categorical field containing the type of transaction. It consists of 4 types: 'P', 'A', 'D', and 'Y'



3.9. FIELD 9: AMOUNT

DESCRIPTION: A numerical field containing the amount of the transaction.



3.10. FIELD 10: FRAUD

DESCRIPTION: A categorical field containing two categories:

0 – Good transaction

1 – Bad transaction

