

Data Quality Report

1. Data Description

The dataset is **Card Transactions Data**, which contains card transactions of a U.S. government organization for business purposes. The data came from author Mark Nigrini's website and covers the company's card transactions over 1 year (2010). There are **10 fields** and **96,753 records**.

2. Summary Tables

Numeric Fields Table

Field Name	# Records With Values	% Populated	# Zeros	Min	Max	Mean	Most Common	Stdev
Date	96,753	100.00%	0	1/1/2010	12/31/2010	6/25/2010	2/28/2010	98 days 21:38
Amount	96,753	100.00%	0	0.01	3102045.53	427.88	3.62	10006.14

Categorical Fields Table

Field Name	# Records With Values	% Populated	# Zeros	# Unique Values	Most Common
Recnum	96,753	100.00%	0	96,753	1
Cardnum	96,753	100.00%	0	1,645	5142148452
Merchnum	93,378	96.51%	0	13,091	930090121224
Merch description	96,753	100.00%	0	13,126	GSA-FSS-ADV
Merch state	95,558	98.76%	0	227	TN
Merch zip	92,097	95.19%	0	4,567	38118
Transtype	96,753	100.00%	0	4	P
Fraud	96,753	100.00%	95,694	2	0

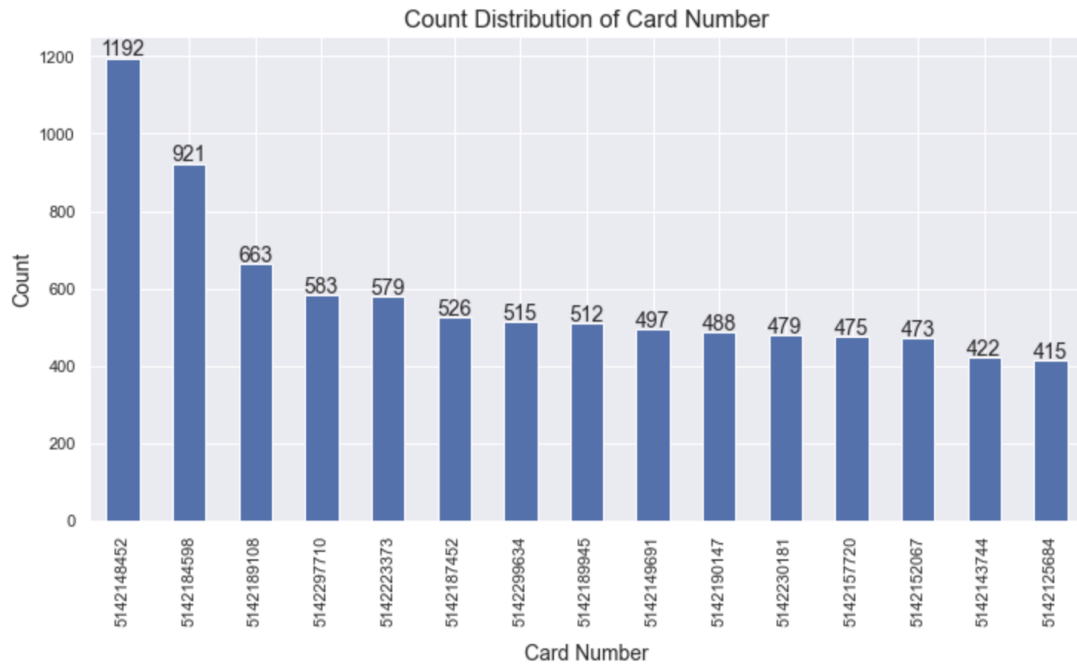
3. Visualization of Each Field

1. Field Name: Recnum

Description: Ordinal unique positive integer for each application record, from 1 to 96,753.

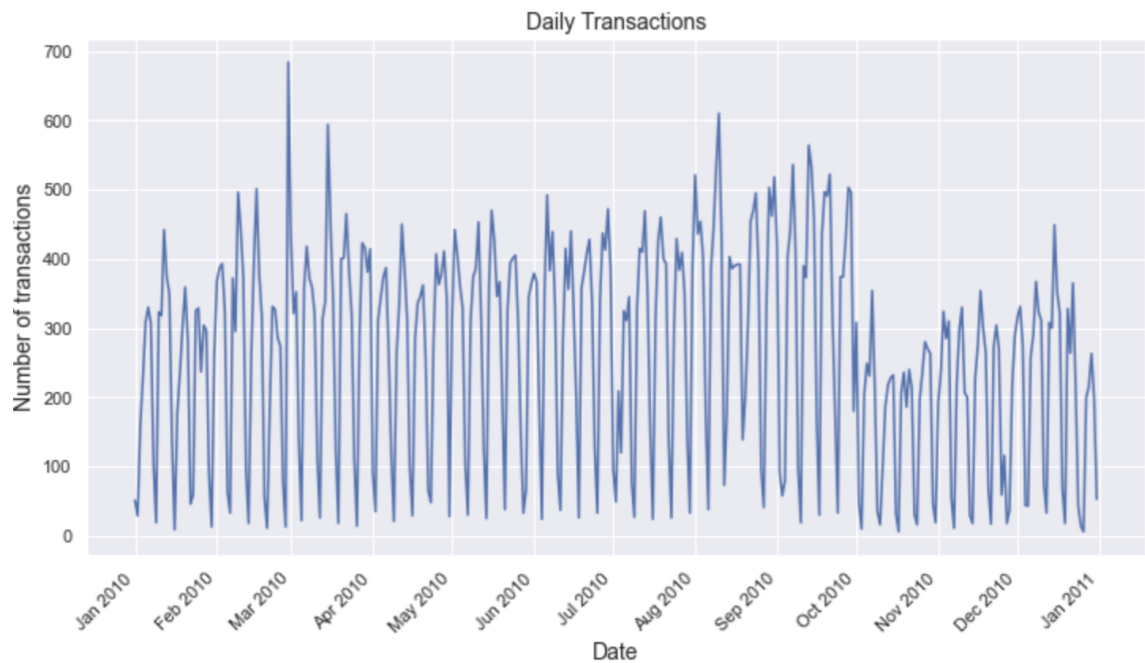
2. Field Name: Cardnum

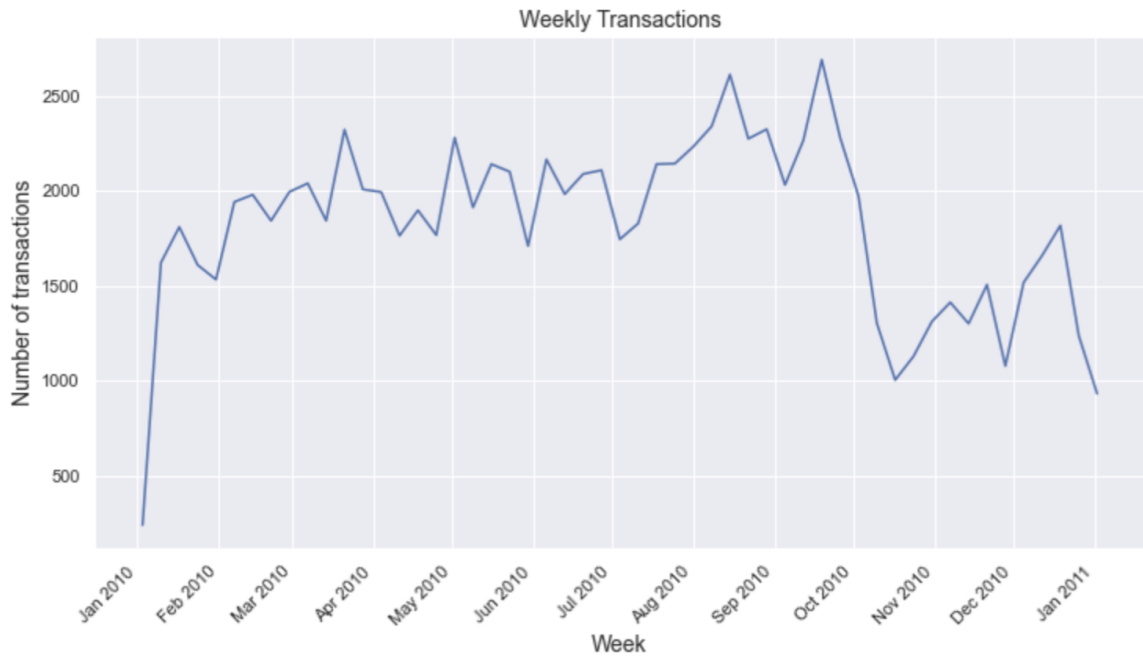
Description: Credit/Debit card number used for a particular transaction. The distribution shows the top 15 field values of card numbers. The most common card number is 514214852, with a total count of 1,192.



3. Field Name: Date

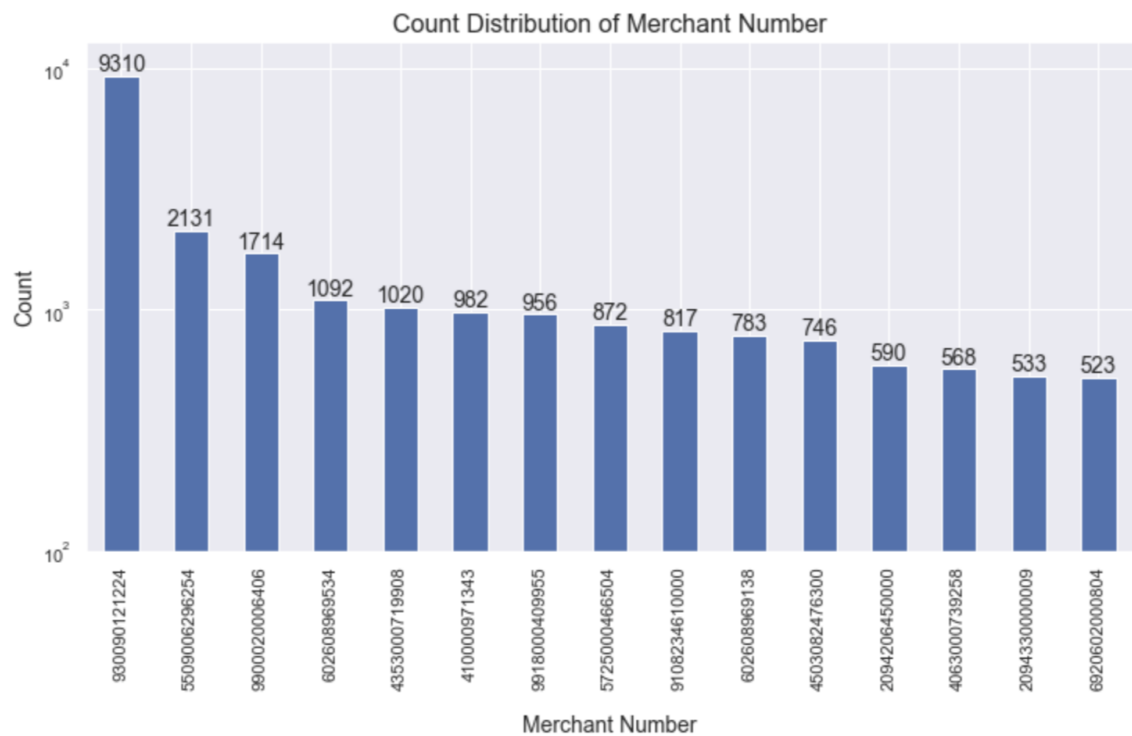
Description: Transaction date. The first distribution shows the number of daily transactions across the year. The second distribution shows the number of weekly transactions across the year.





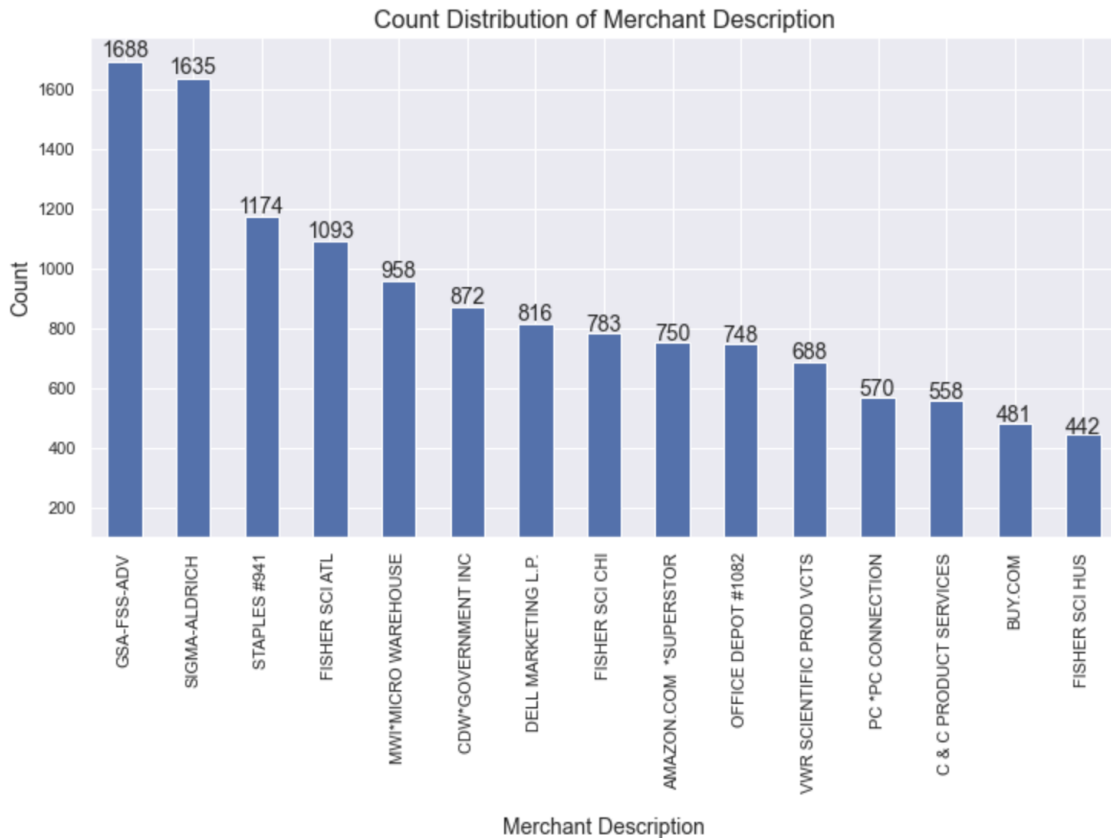
4. Field Name: Merchnum

Description: Merchant number for a particular transaction. The distribution shows the top 15 field values of merchant numbers. The most common merchant number is 930090121224, with a total count of 9,310.



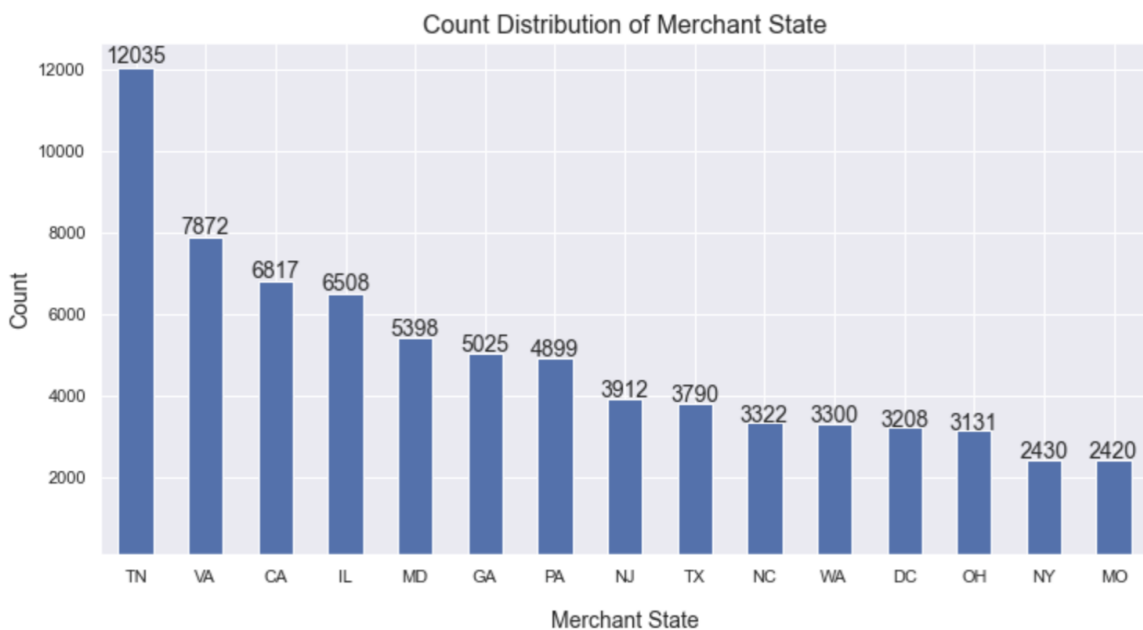
5. Field Name: Merch description

Description: Merchant description for a particular transaction. The distribution shows the top 15 field values of merchant descriptions. The most common merchant description is GSA-FSS-ADV, with a total count of 1,688.



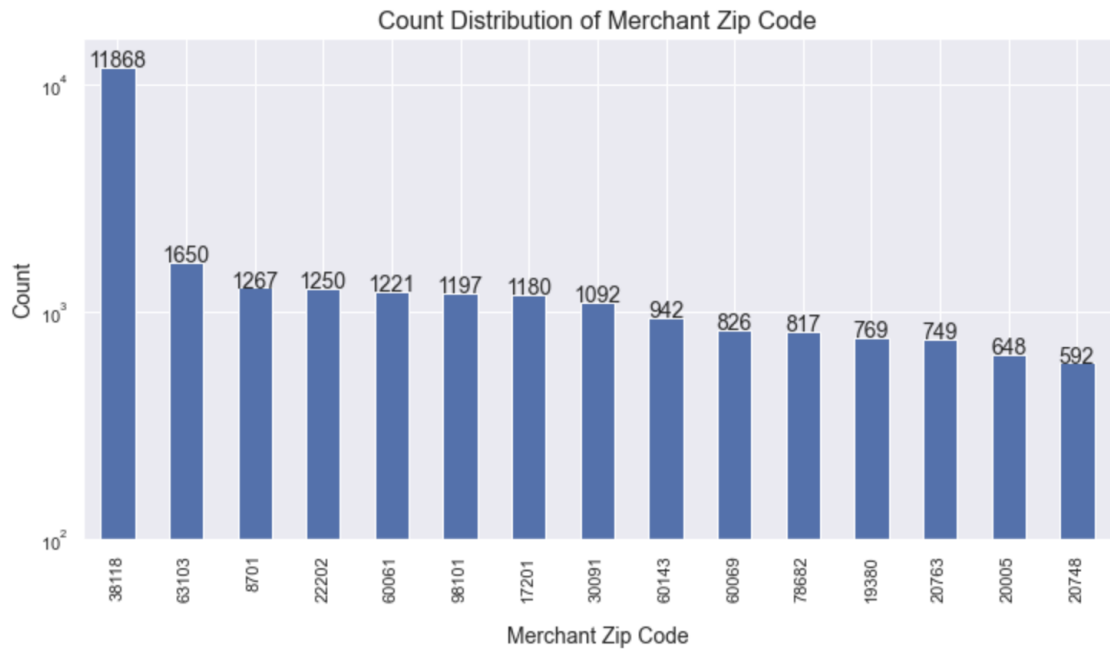
6. Field Name: Merch state

Description: Merchant's state for a particular transaction. The distribution shows the top 15 field values of merchant states. The most common merchant state is TN (Tennessee), with a total count of 12,035.



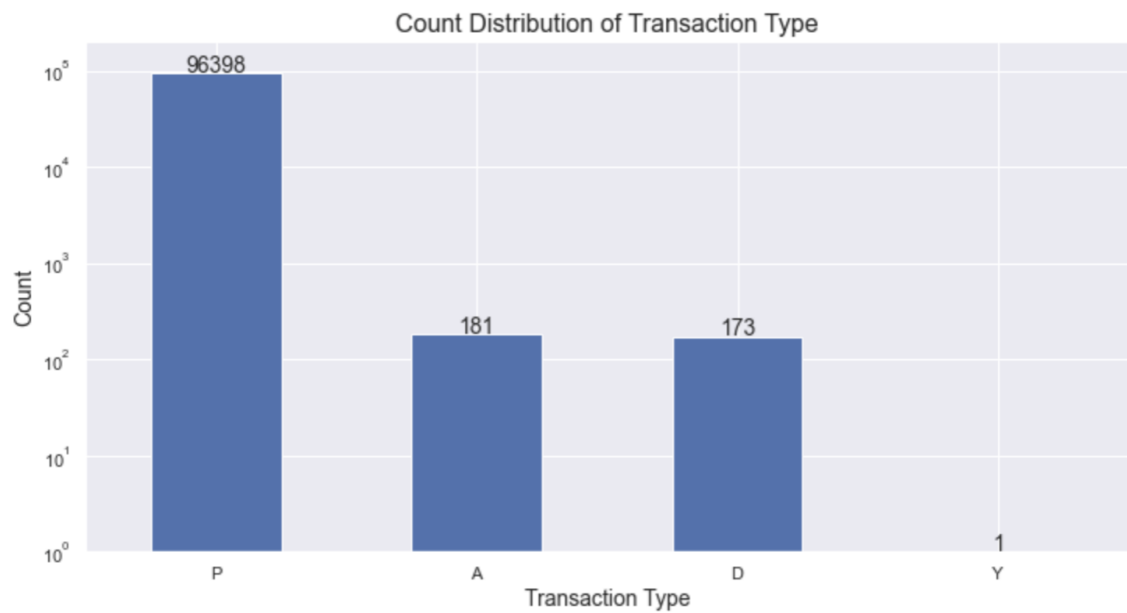
7. Field Name: Merch zip

Description: The zip code of a Merchant for a particular transaction. The distribution shows the top 15 field values of merchant zip codes. The most common merchant zip code is 38118, with a total count of 11,868.



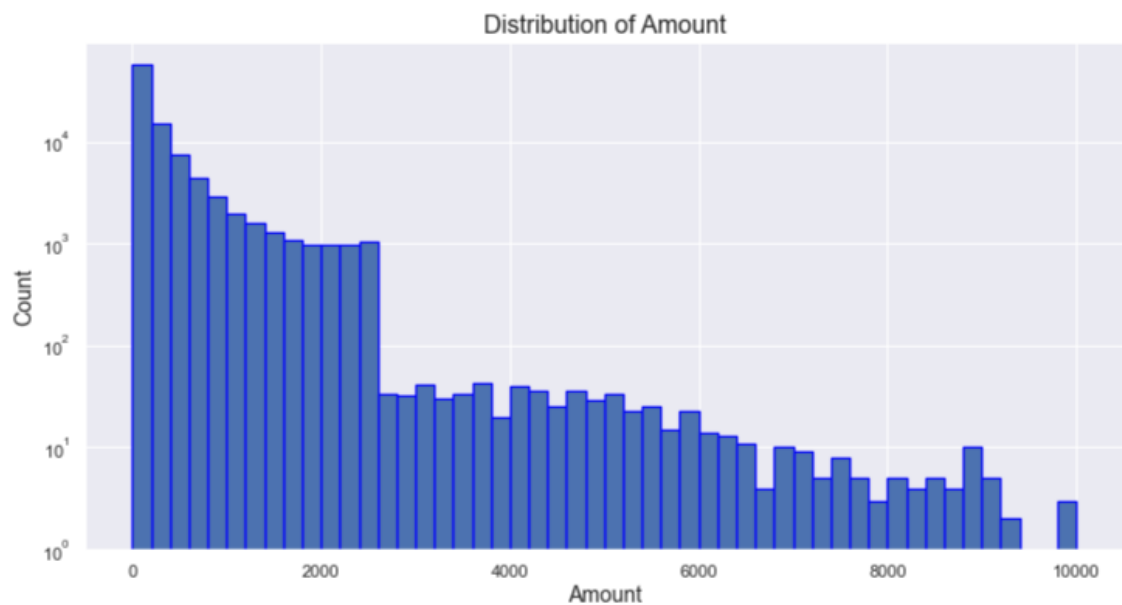
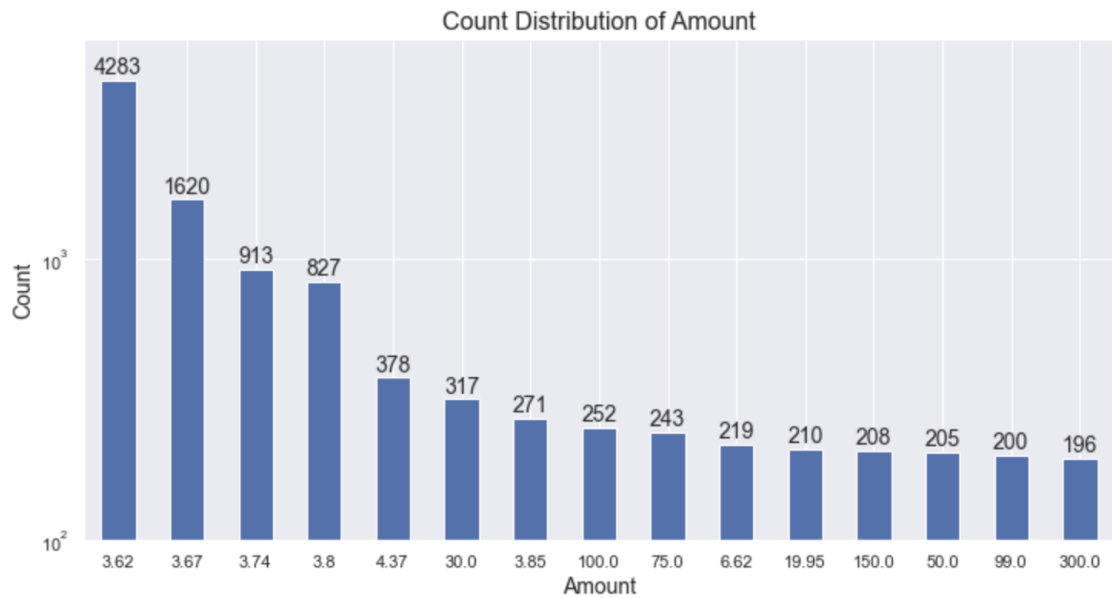
8. Field Name: Transtype

Description: Transaction type. The distribution shows all field values of transaction type. The most common type is P, with a total count of 96,398.



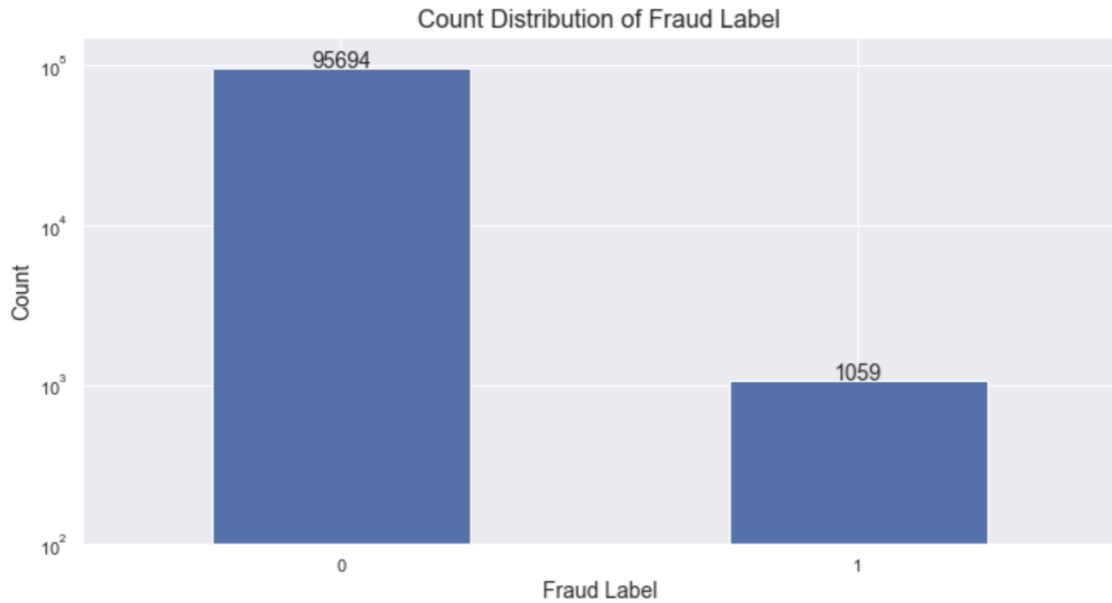
9. Field Name: Amount

Description: The amount for a particular transaction. The first distribution shows the top 15 field values of amount. The most common value is 3.62, with a total count of 4,283. We see that most common transactions are generally for small amounts (with the highest value being 300). The second distribution shows the continuous distribution of amounts with the amounts being taken till standard deviation.



10. Field Name: Fraud

Description: Fraud identification label. Fraud = 0 (Not fraudulent), Fraud =1 (Fraud identified). The total count of fraud label 0 is 95,694. The total count of fraud label 1 is 1,059.



Assignment Submitted By: Neha Mittal