Data Quality Report: Credit Card Payment

Bryce Charette

## Clean data

# Adjust column names  
colnames(dat) <- c("rec\_num","card\_num","date","merch\_num","merch\_desc",  
 "merch\_state","merch\_zip","trans\_type","amount","fraud")  
  
# Adjust data types  
dat$rec\_num <- as.character(dat$rec\_num)  
dat$card\_num <- as.character(dat$card\_num)  
  
# Save data  
# saveRDS(dat, "card\_payments.rds")

## Statistics & Distributions

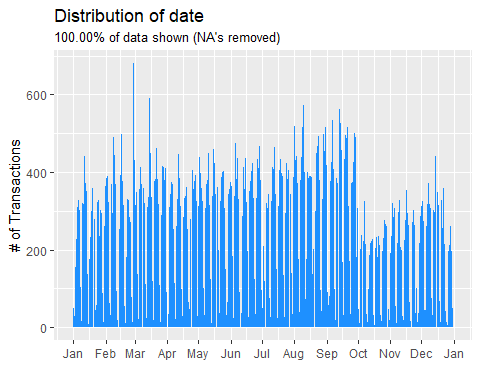
### rec\_num

* **Interpretation:** Unique ID for each record in the data set.
* **Type:** ID
* **Unique Values:** 95,484
* **Missing Values:** 0 (0.00%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:** NA

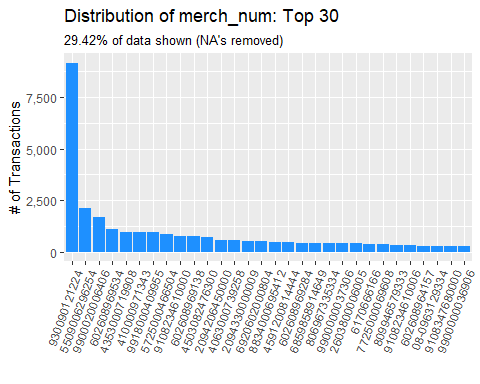
### card\_num

* **Interpretation:** Credit card number of credit card holder.
* **Type:** ID
* **Unique Values:** 1,638
* **Missing Values:** 0 (0.00%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:** NA

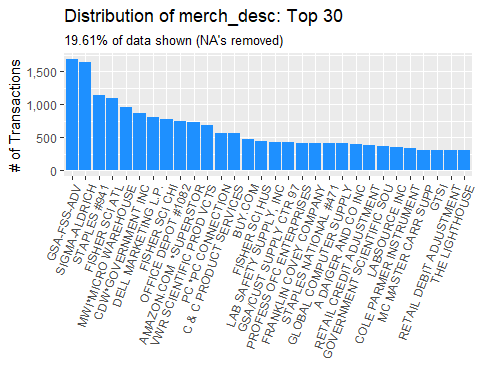
### date

* **Interpretation:** Transaction date
* **Type:** Date
* **Unique Values:** 365
* **Missing Values:** 0 (0.00%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:**  
  

### merch\_num

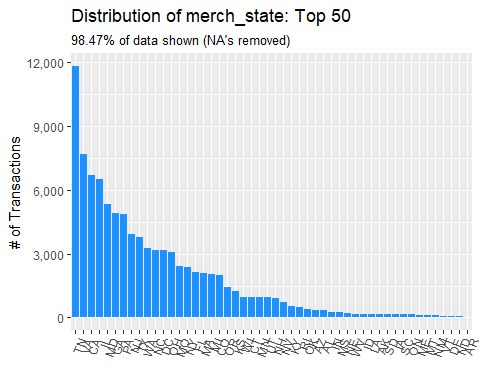
* **Interpretation:** Merchant ID
* **Type:** ID
* **Unique Values:** 13,091
* **Missing Values:** 3,351 (3.51%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:**  
  

### merch\_desc

* **Interpretation:** Merchant description
* **Type:** ID
* **Unique Values:** 13,125
* **Missing Values:** 0 (0.00%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:**  
  

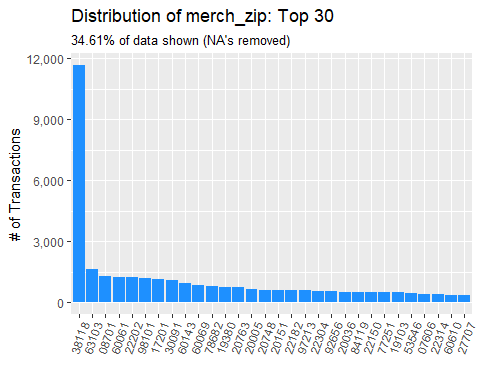
### merch\_state

* **Interpretation:** Merchant state
* **Type:** Categorical
* **Unique Values:** 228
* **Missing Values:** 1,191 (1.25%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:**



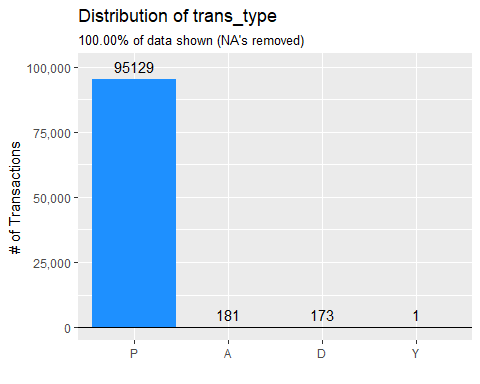
### merch\_zip

* **Interpretation:** Merchant zipcode
* **Type:** Categorical
* **Unique Values:** 4,568
* **Missing Values:** 4,639 (4.86%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:**



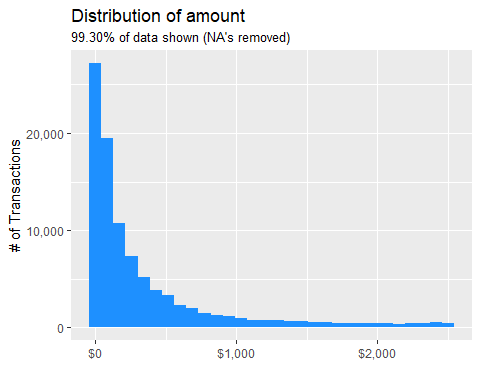
### trans\_type

* **Interpretation:** Transaction type
* **Type:** Categorical
* **Unique Values:** 4
* **Missing Values:** 0 (0.00%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:**



### amount

* **Interpretation:** Transaction amount in US dollars ($)
* **Type:** Numeric
* **Unique Values:** 34,190
* **Missing Values:** 0 (0.00%)
* **Mean:** $414.38
* **Standard Deviation:** $10,066.23
* **Min:** $0.01
* **Max:** $3,102,045.53
* **Distribution:**



### fraud

* **Interpretation:** Indicator whether transaction is fraudulent
* **Type:** Categorical
* **Unique Values:** 2
* **Missing Values:** 0 (0.00%)
* **Mean:** NA
* **Standard Deviation:** NA
* **Distribution:**

