# NIKHIL SINGH

# Exploratory Data Analysis of a synthesised ANZ Transaction Data

## MISSING VALUES:

Certain columns in the dataset have a lot of missing values

Number of missing values column wise

```
status                    0
card_present_flag      4326
bpay_biller_code      11158
account                   0
currency                  0
long_lat                  0
txn_description           0
merchant_id            4326
merchant_code         11160
first_name                0
balance                   0
date                      0
gender                    0
age                       0
merchant_suburb        4326
merchant_state         4326
extraction                0
amount                    0
transaction_id            0
country                   0
customer_id               0
merchant_long_lat      4326
movement                  0
dom                       0
weekday                   0
dtype: int64
```
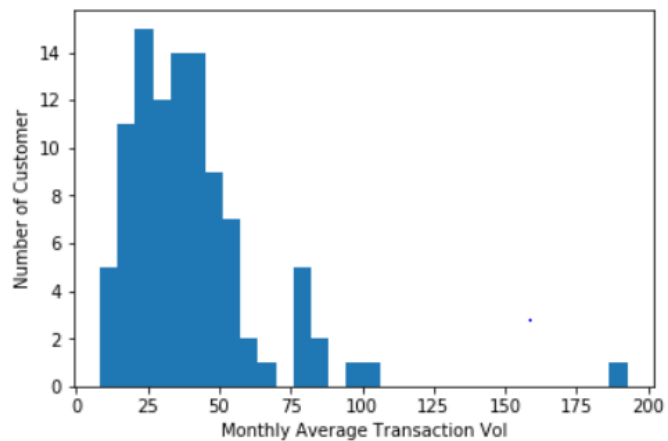
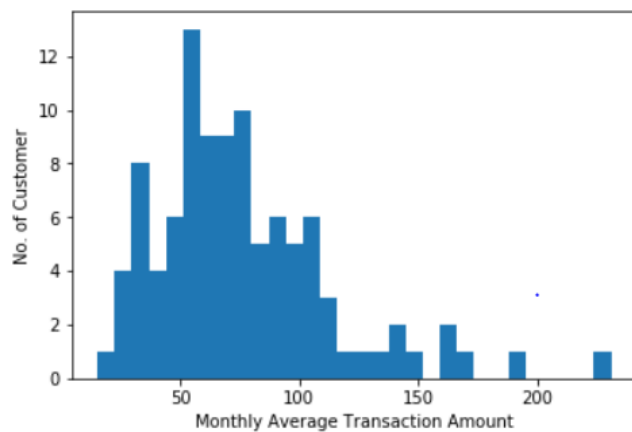## MAXIMUM NUMBER OF TRANSACTIONS:

A maximum of 174 transactions were done on 28 September 2018.

The customer with the highest number of transaction has id CUS-2487424745(578 transactions)

## MONTHLY AVERAGE TRANSACTION VOLUME



## MONTHLY AVERAGE TRANSACTION AMOUNT



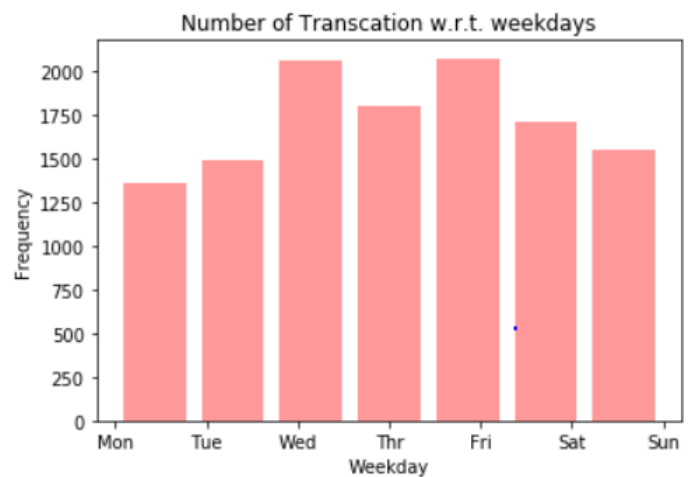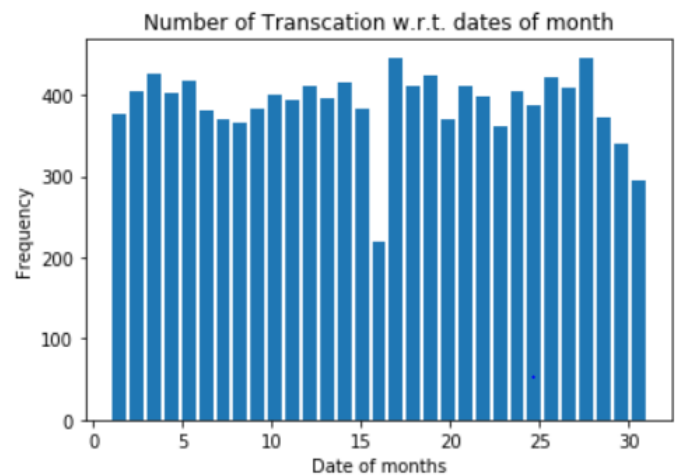| Month | Mean Balance | Mean Amount |
| --- | --- | --- |
| August | 10794 | 185 |
| September | 14730 | 182 |
| October | 18450 | 196 |

## ANALYSIS BY DATE AND WEEKDAY

The number of transactions tend to reduce by the end of the month.

More number of transaction in mid-week as compared to the start and end of the week

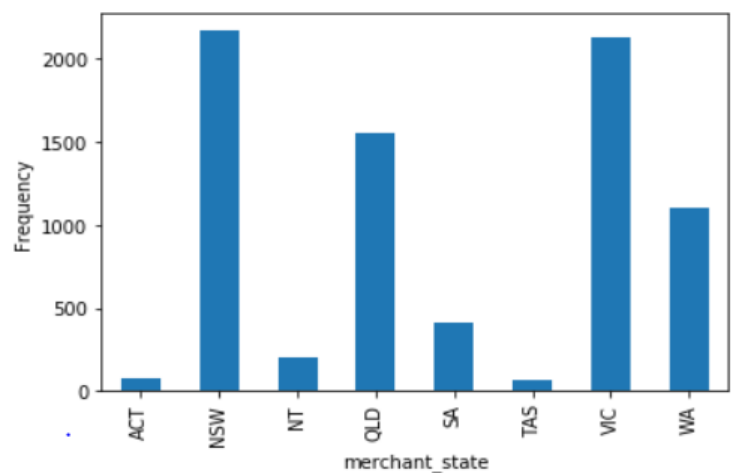Wednesday and Friday are the busiest day in the week.

It can be observed that transactions follow a cyclic pattern over weekdays with Friday being the busiest day.



Number of Transcation w.r.t. dates of month



Number of Transcation w.r.t. weekdays

## GEOGRAPHICAL CONCLUSIONS

As we can see that New South Wales and Victoria are the main business hotspot with the most number of transactions.

Northern, Southern Australia and Tasmania doesn't account for much of the business.

The amount of transactions follow a very positive skewed distribution for the transaction type 'SALE-POS' and 'POS' with a majority of the transactions lying between 0 to 60 Australian Dollars.