

ANZ MANAGEMENT

Exploratory Data Analysis

- Vidula Arolkar

DATA DESCRIPTION

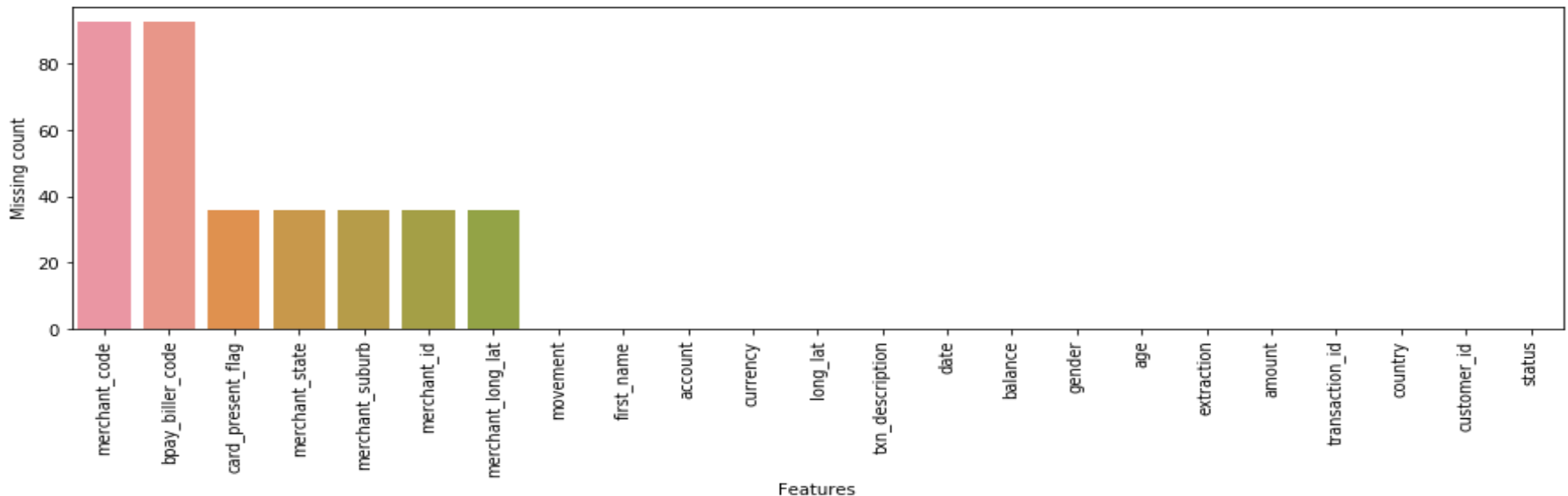
The transaction dataset contains 3 months' worth of transactions for 100 hypothetical customers. It contains purchases, recurring transactions, and salary transactions.

There are 12043 observations and 23 features in this data set.

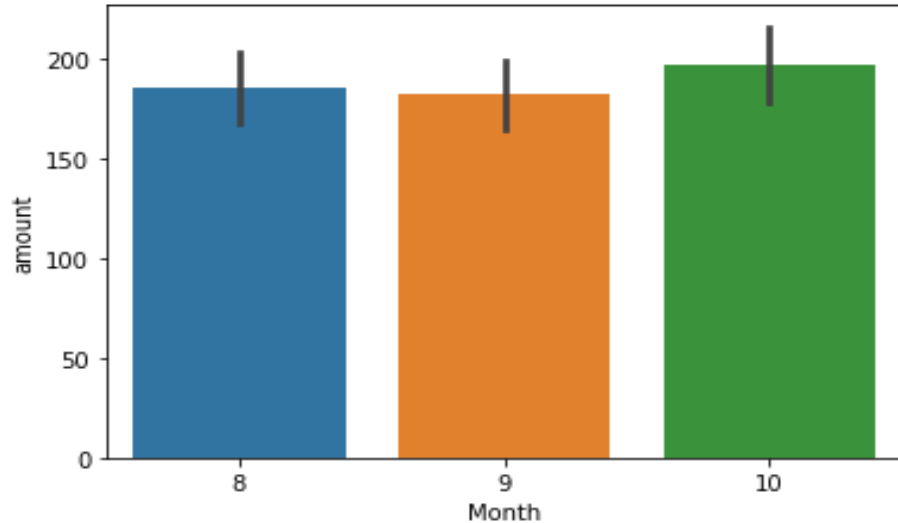
There are 100 unique accounts in all.

DATA PREPARATION

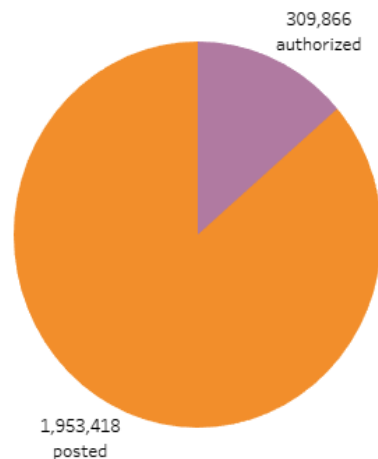
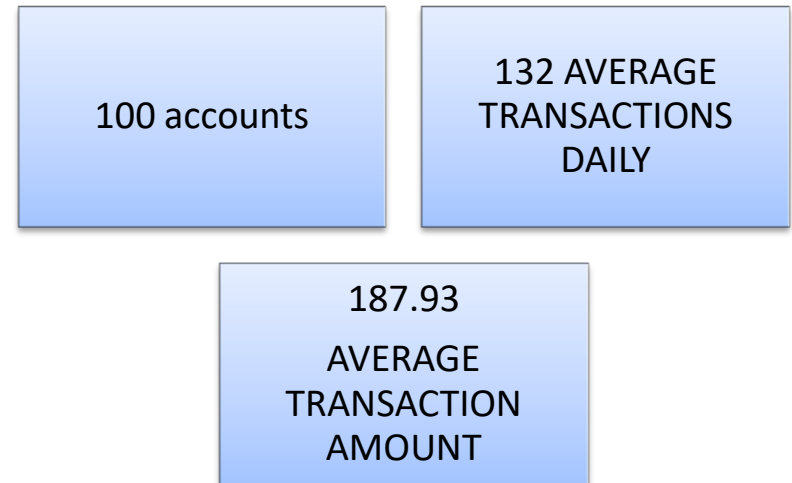
- The data had many missing values.
- Some columns did not explain much variance.
- Handling Missing values:
 - Dropped some of the columns like card_player_flag.
 - Imputed missing values with 0 and other categories.
 - Dropped columns with high number of missing percentage.
- Dropped columns with low variance. (merchant_code)



EXPLORATORY DATA ANALYSIS



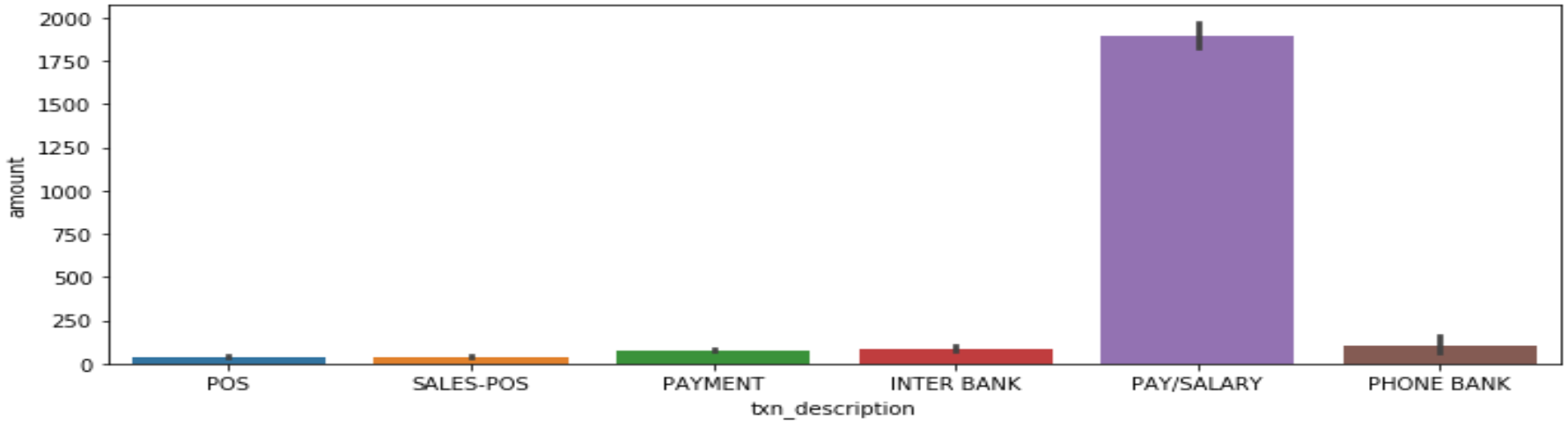
Transaction amount in each month.



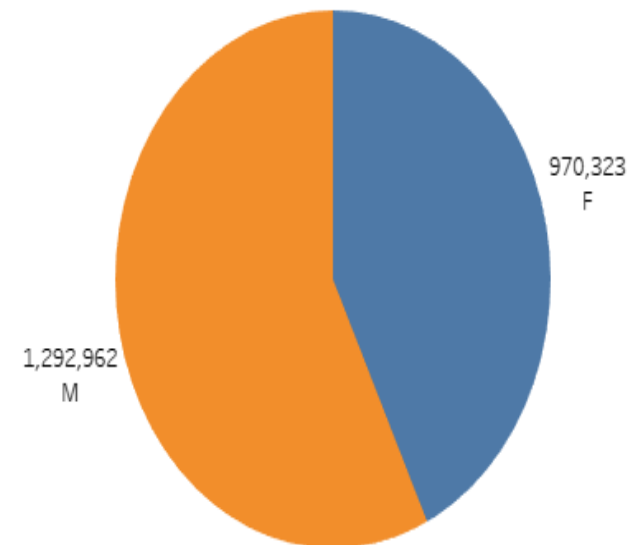
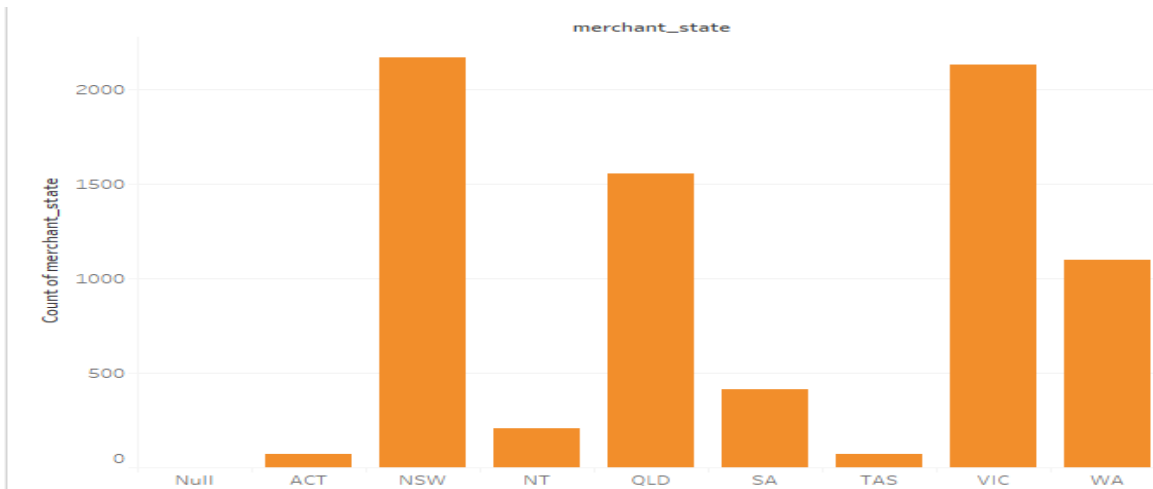
Status of Transactions.

- **1953418 Transactions of Posted.**
- **309,866 Transactions of Authorized.**
- **The Transactions amounts are higher in the month of September.**

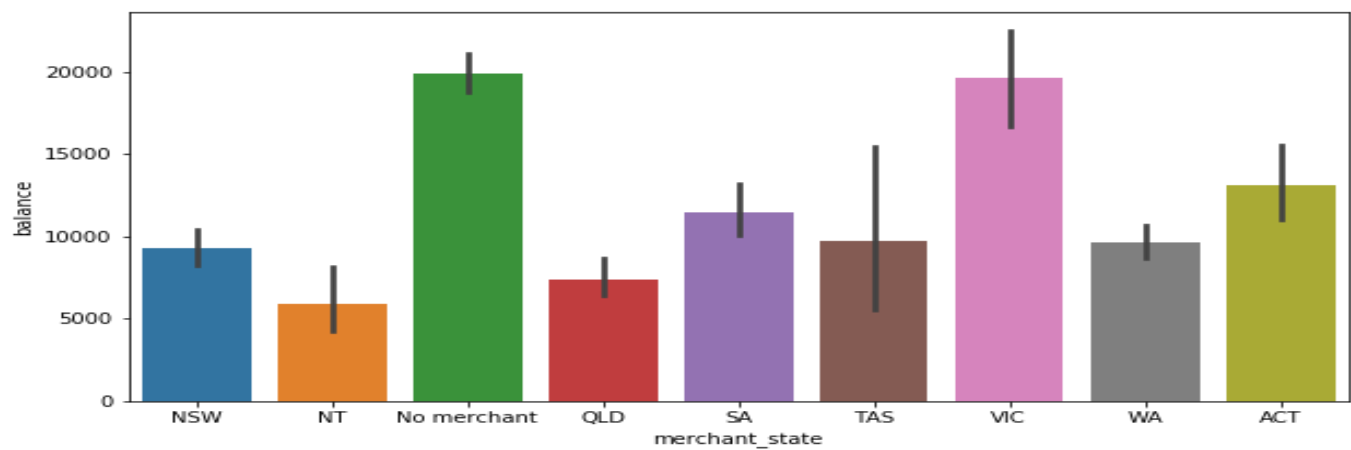
TRANSACTION TYPE:



- Pay/ salary has most of the transactions.
- We can see that males have done more transactions than females.
- New South Wales and Victoria has most of the merchants.

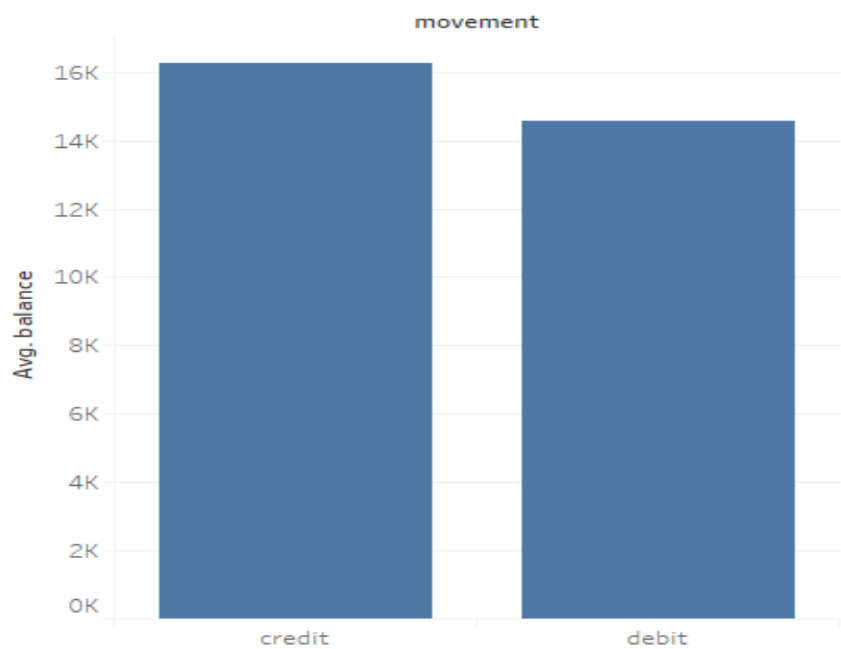


BALANCE VS MERCHANT STATE



Victoria state has most of the balance maintained.

Balance vs movement.



	merchant_suburb	amount
1046	No merchant	1953417.85
1381	Sydney	21675.44
893	Melbourne	11943.53
1323	South Brisbane	11740.58
880	Mascot	10282.62

Top 5 suburbs w.r.t Transaction amount.