# ANZ

Data Analytics Virtual Internship

The dataset consists of transactions data of 100 customers for a period of 3 months.

The analysis is divided into following parts for module 1:

- Data Cleaning: Null values, Duplicate values and Outliers.
- Data Exploration
- Statistical Analysis
- Insights: Data Segmentation by Month and day, Effect of outliers on Analysis, Transaction Volume, Location insights, Customer characteristics.

Descriptive Statistics

	balance	age	amount
count	12043.000000	12043.000000	12043.000000
mean	14704.195553	30.582330	187.933588
std	31503.722652	10.046343	592.599934
min	0.240000	18.000000	0.100000
25%	3158.585000	22.000000	16.000000
50%	6432.010000	28.000000	29.000000
75%	12465.945000	38.000000	53.655000
max	267128.520000	78.000000	8835.980000
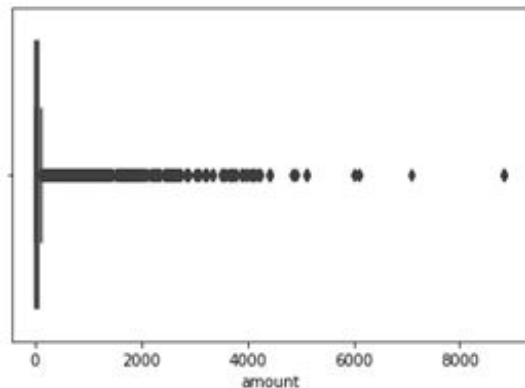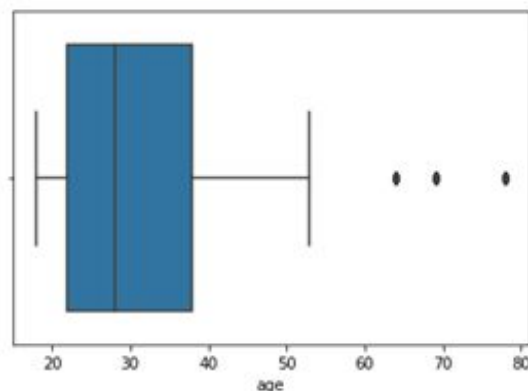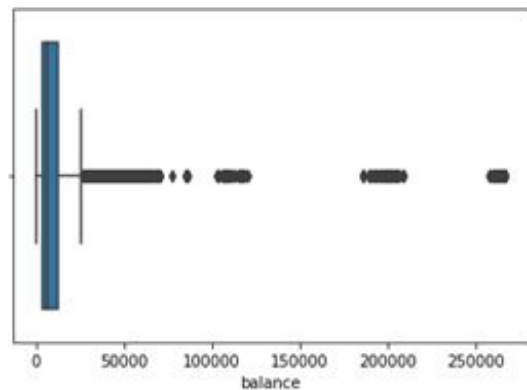
Effect of outliers on the Analysis

From the Boxplots of the variables Balance, Age and Amount, we can see that there are many extreme outliers and they effect our analysis.
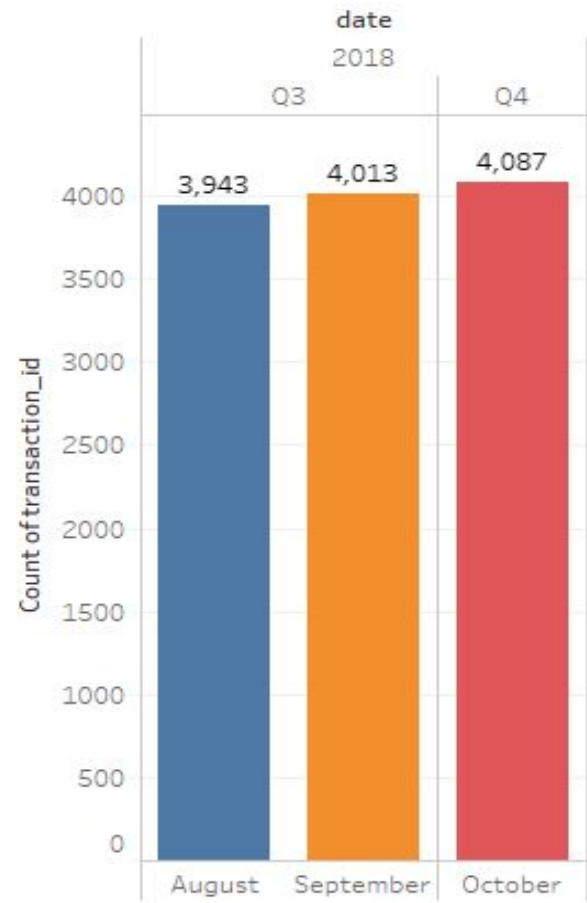
In the Balance variable, 75% of values are within 12000 but as the max value is 267128, the mean increases to 14000.

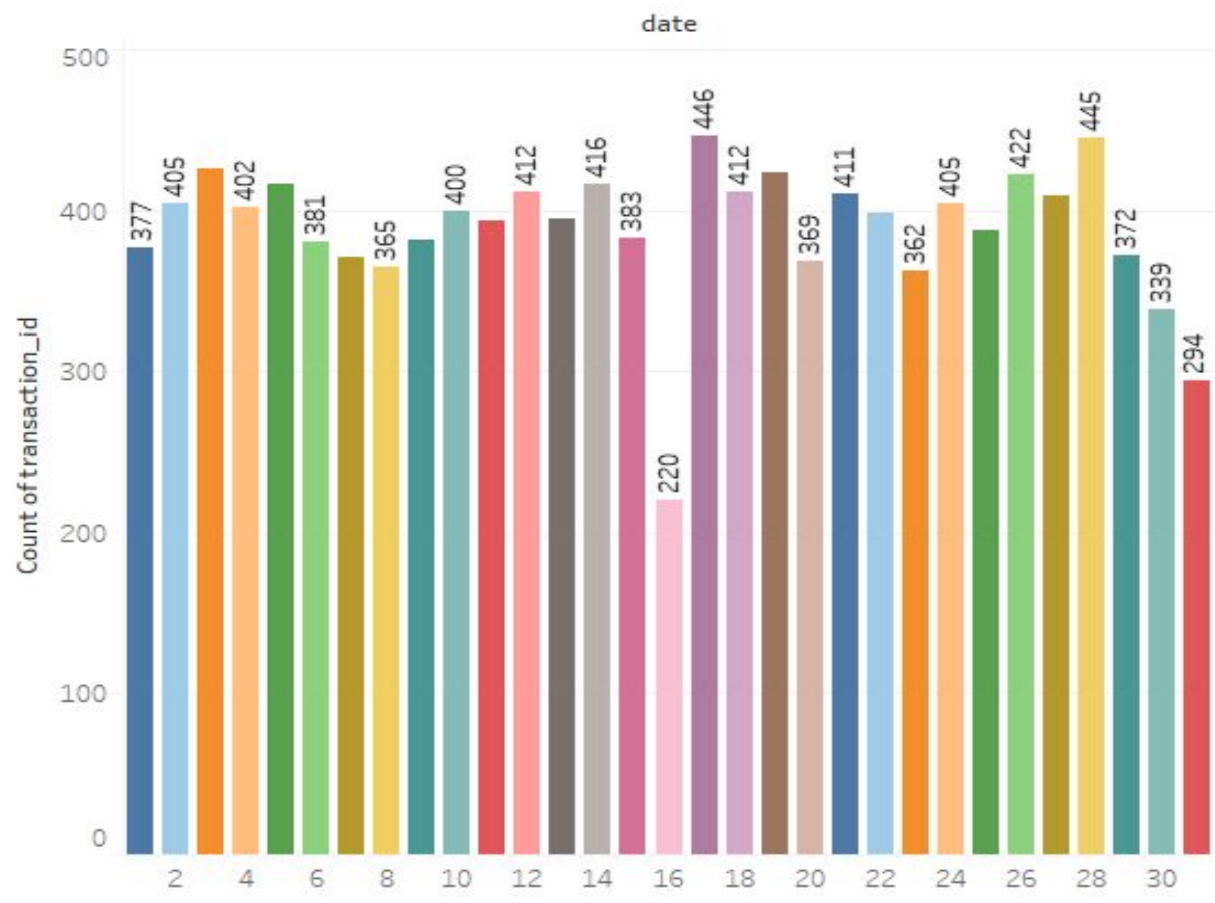In the Amount variable, 75% of values are within 53 but the max value is 8835 which drastically increases the mean to 187.

In the Age variable, 75% of the values are within 38 and the max is 78. This does not alter the mean as the range is not wide and there are only three data points.
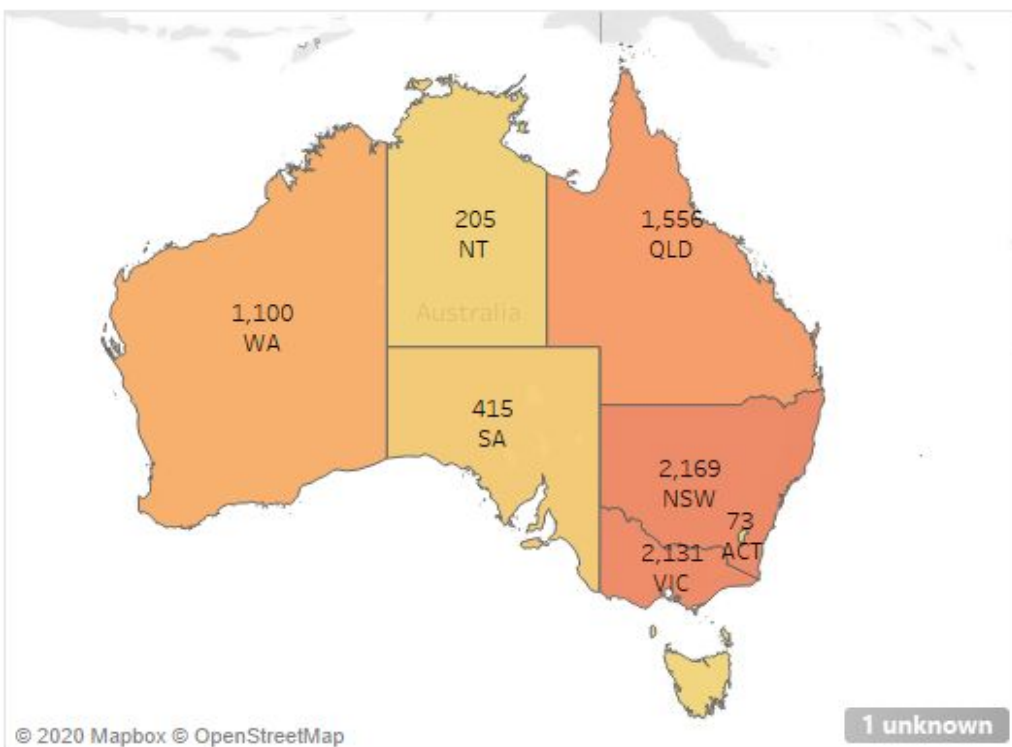
Data Segmentation by Month and Day

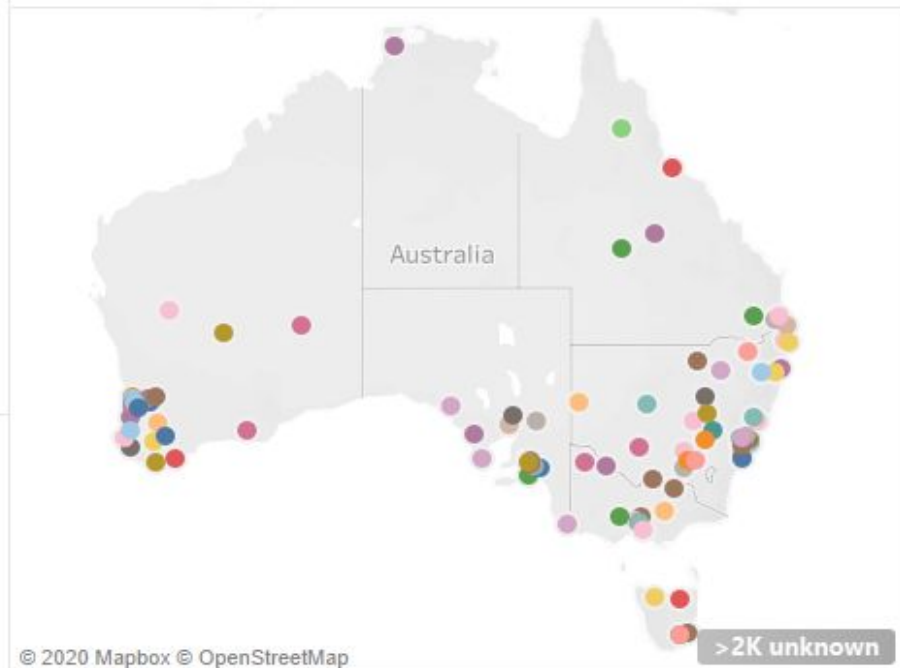## Transaction Volume by Month

Transaction Volume by Day
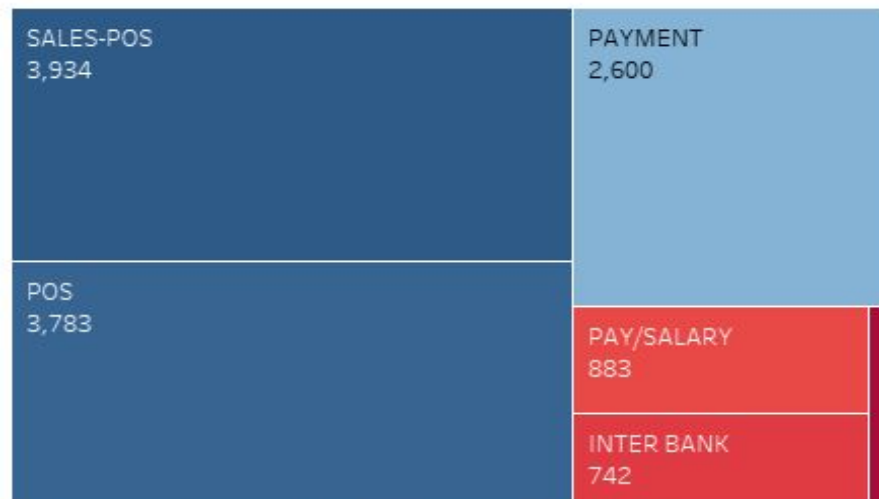
Transaction Volume by Location

205
NT

1,556
QLD

1,100
WA

Australia

415
SA

2,169
NSW

73
ACT

2,131
VIC

© 2020 Mapbox © OpenStreetMap

1 unknown

The States NSW and VIC including the ACT constitutes for a high transaction volume.

Merchant Suburbs

Australia

© 2020 Mapbox © OpenStreetMap

>2K unknown

Transaction Volume by Categories

SALES-POS 3,934	PAYMENT 2,600
POS 3,783	PAY/SALARY 883
	INTER BANK 742

Total Transactions as Credit or Debit

movement
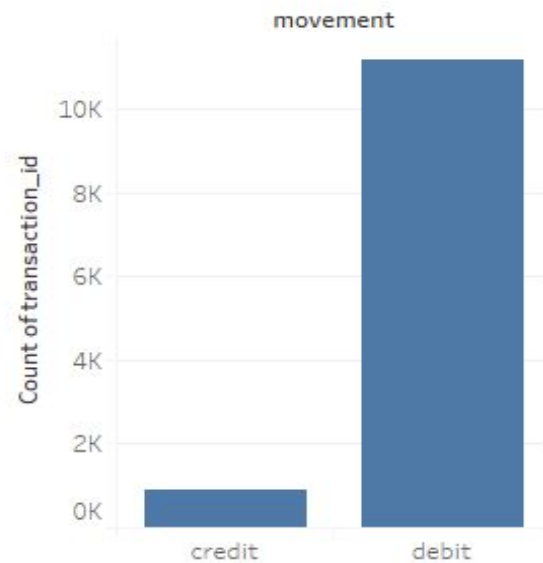
Average Amount of Transaction by Day

date

205.7 131.0 243.7 151.3 222.3 173.0 234.1 125.1 193.7 183.4 246.8 244.3 189.9 174.0 187.9 263.8 219.0 150.6 205.7 198.1
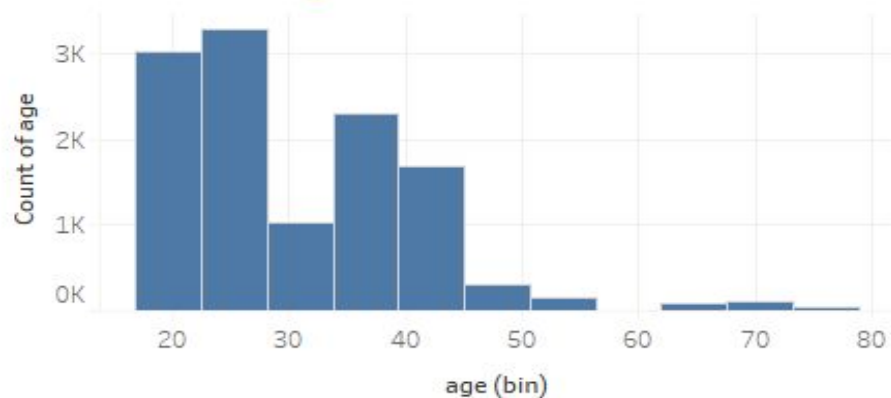
Avg. amount
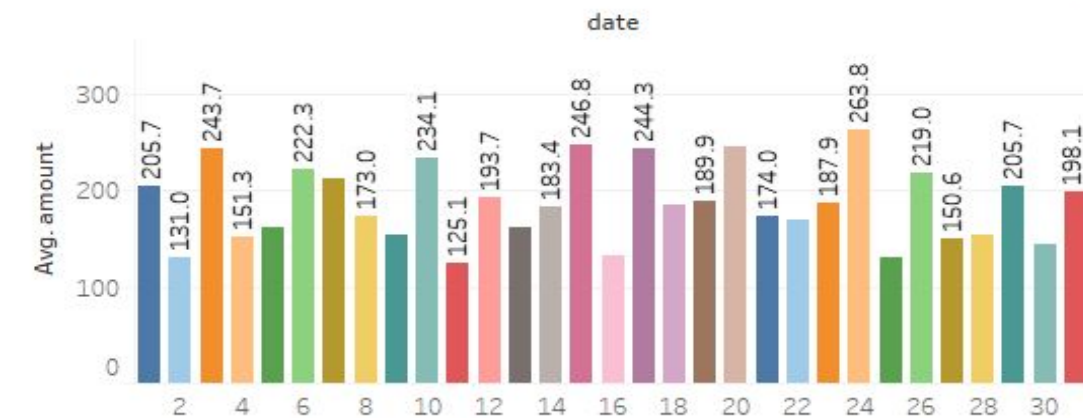
Majority of the transactions are of Debit type.
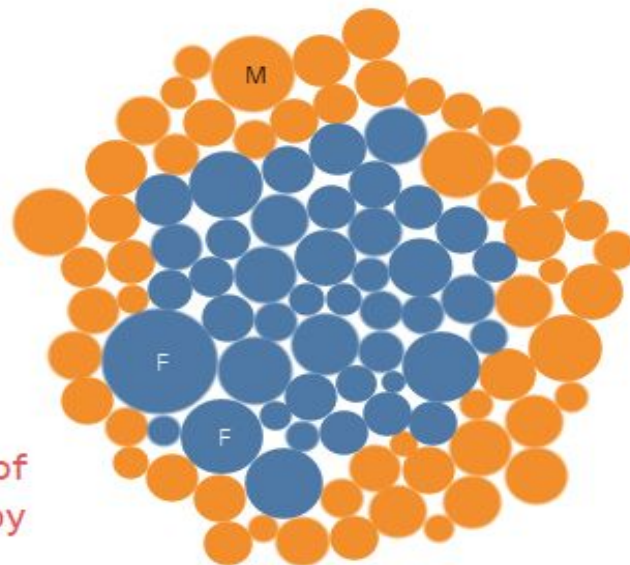But large number of transactions take place at Sales-POS.
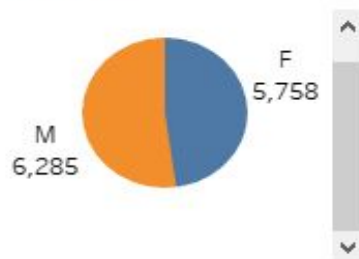
Distribution of Age variable

Number of Customers by Gender

Average Age of Customers by Gender

Total number of Transactions by Gender

F
5,758

M
6,285

The average age of the customers is 30.
Greater number of transactions are through Male customers.

Thank you!