



EXPLORATORY DATA ANALYSIS

ANZ DATA SCIENCE VIRTUAL PROGRAM

KIANA KHEIRI



The provided data

The dataset contains 12043 transactions for 100 customers who have one bank account each.

Transactional period is from 01/08/2018 31/10/2018 (92 days).

There is only 91 unique dates in the dataset, (missing date: 2018-08-16)

The data entries are unique and have consistent formats for analysis. Some columns contain missing data (NA cells), possibly due to the nature of transaction.

The columns are:

```
['status', 'card_present_flag ', bpay_biller_code ', 'currency', 'long_lat ', txn_description ',  
merchant_id'merchant_code ', first_name ', 'balance', 'date', 'gender', "merchant_suburb ',  
merchant_state ', 'extraction', "transaction_id ', 'country', customer_id  
,merchant_long_lat'movement']
```




THE INSIGHTS

- Total transaction volume over the dates have ranged from 10,000 to 40,000 usually, mean transactions have ranged from 100 AUD to 300 AUD usually.
- Male customers have spent more on their transactions over the dates, as compared to their female customers.
- States have widely distributed transaction over the dates, but New South Wales and Queensland have had a few peaks in transaction amounts.
- State wise, New South Wales, Victoria and Queensland have had most transaction. The fact that most economic centres and large cities of Australia are in this states also supports the fact.
- Going by total transaction in sub regions maximum transactions have come from Sydney, Melbourne and Brisbane. These are the largest cities in Australia, and it is logical that largest transactions come from there.

- Going by mean transactions in sub regions high values have been from Pantapin , Gladesville , Parap and such. We cannot yet determine much about them due to limited data.
- Total values of sum of customers bank accounts balance have been in the range of 1000000 to 1500000 AUD over the dates.



- 
- The mean and usual account balance of the customers has varied from 10,000 to 20,000 AUD over the times. This does suggest that customers have been saving money, and their assets have grown.
 - Male customers have had a higher account balance than female customers, over the dates.
 - The mean transaction bar plot shows that the value has remained fairly uniform over the ages, the total transaction has high rises for 22 and 34.
 - Most of the customers of the bank have made more credit transactions, as compared to debit transactions.
 - Working with the map coordinates, it becomes clear that the majority of the transactions are based where the major cities of Australia are.
 - Most spending is coming from the large cities of Australia and from some key states. The bank should look to expand its operations in these areas.
 - Some customers also have tendency to spend more than others, the average bank balance of the customers have increased. It indicates that the bank customers are getting well financially.

THANK YOU