



ANZ Exploratory Data Analysis

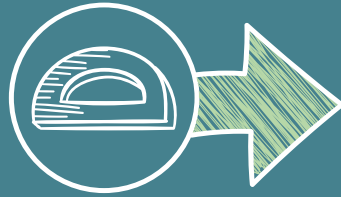
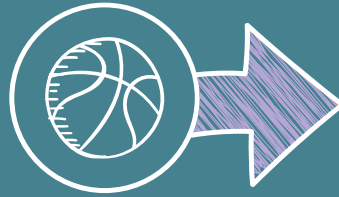
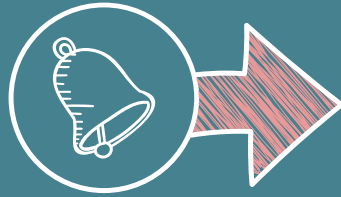
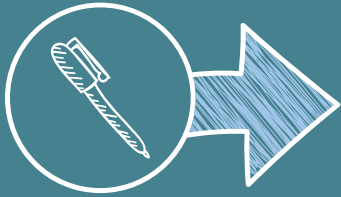
Prepared by-

Md Redyan Ahmed from

Indian Institute of Science Education and Research Kolkata

For Data @ ANZ program of Forage

Contents:



Context

Data
Visualization

Results

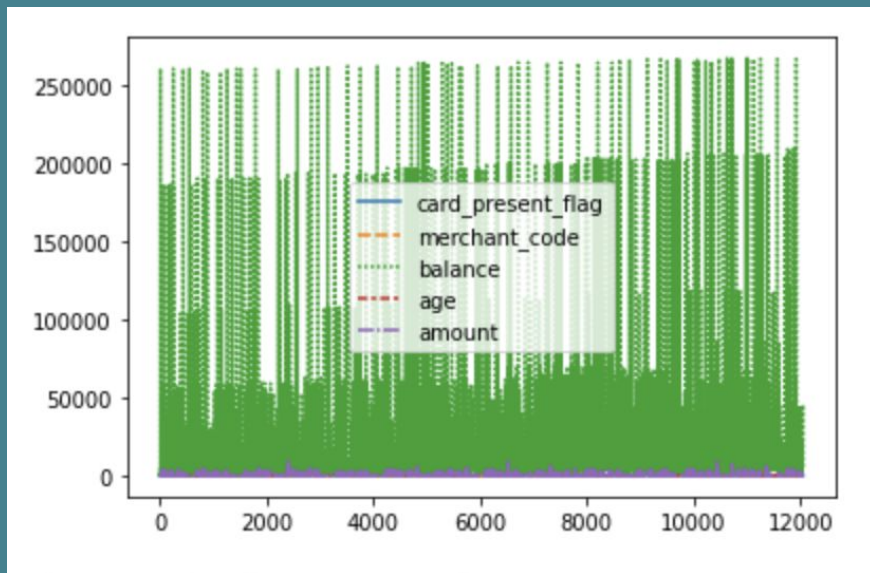
Methodology

Acknowledgement

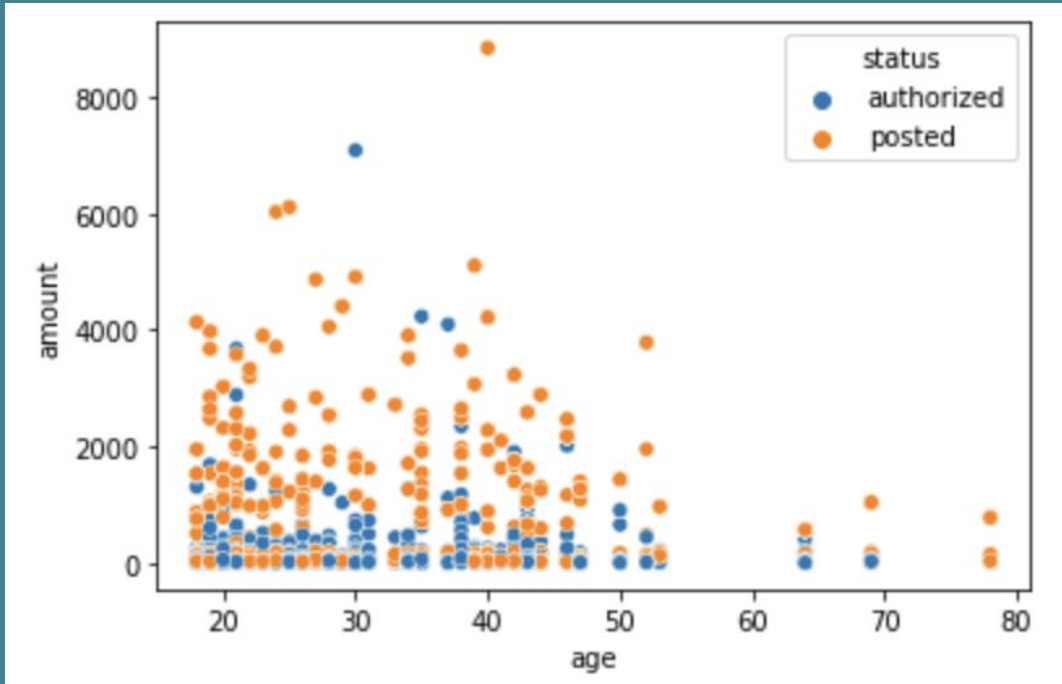
Context:

- Data@ANZ is about mining and linking datasets to develop stories that matter and challenge the status quo, to deliver on ANZ's purpose **“to shape a world where people and communities thrive”**.
- Data@ANZ is a virtual experience program at Forage.
- As a project of virtual experience program, I will be segmenting the dataset and will try to draw unique insights including visualization of dataset via various kinds of plots and will explore correlations between customer attributes .

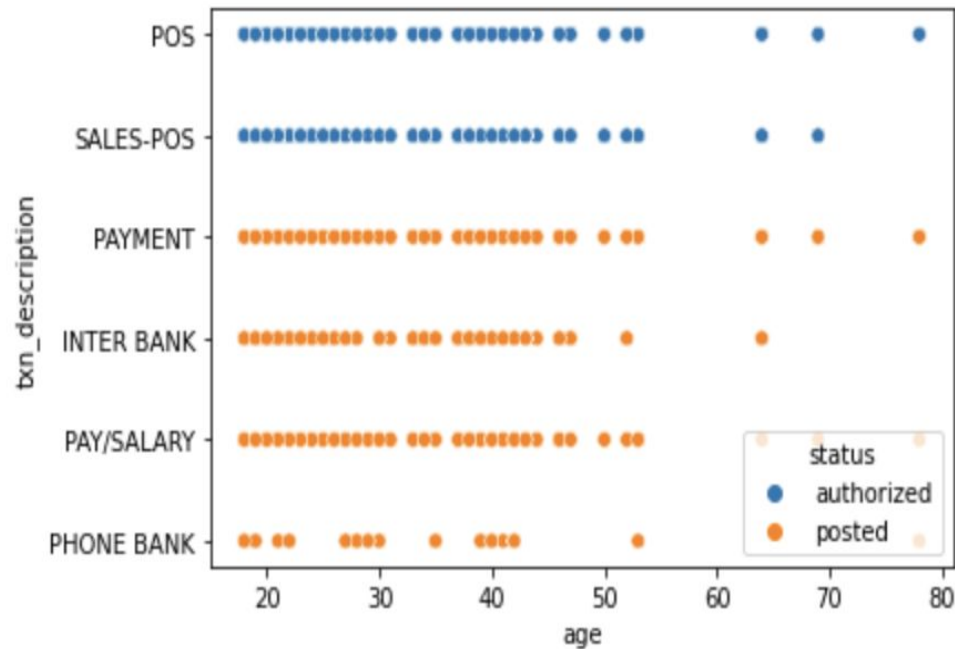
Data visualization:



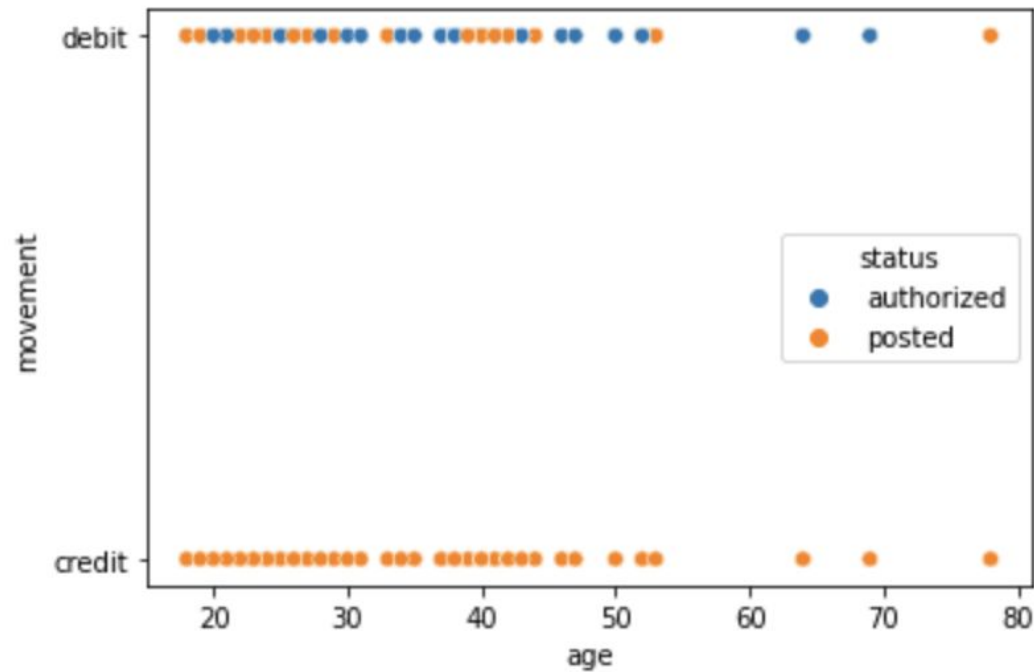
This lineplot reflects how all the data will look like if they are plotted together.



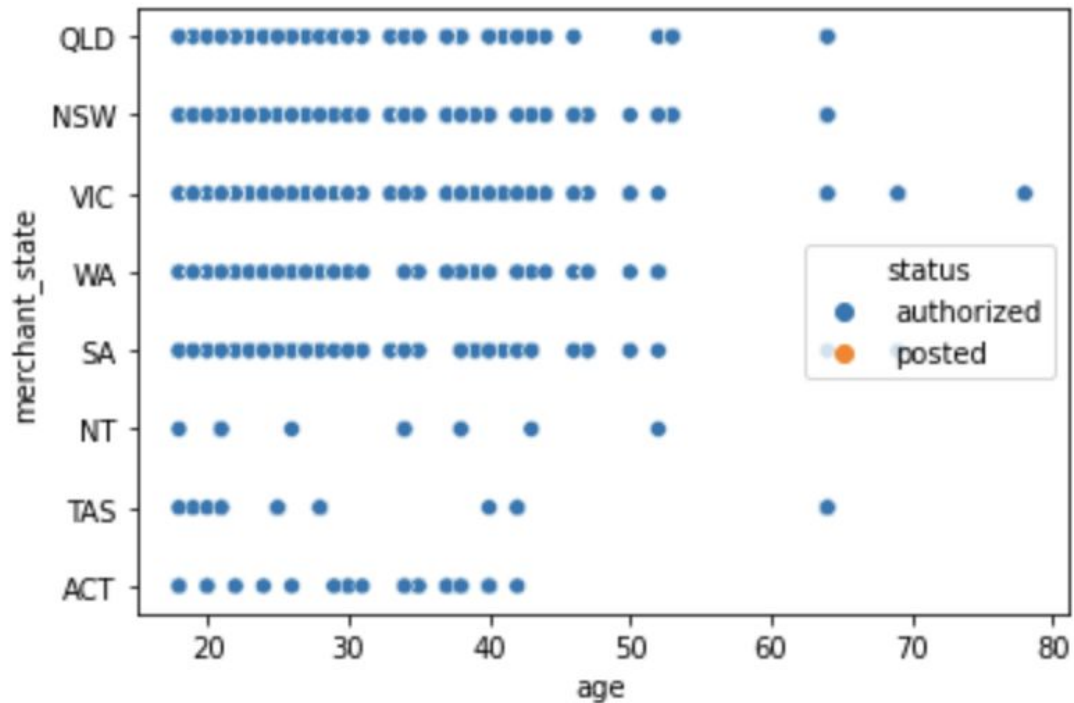
In this scatterplot, it's reflected that most of the amounts are having the status- 'Posted' and more than 8000 is the highest posted amount and more than 6000 is the highest authorized amount.



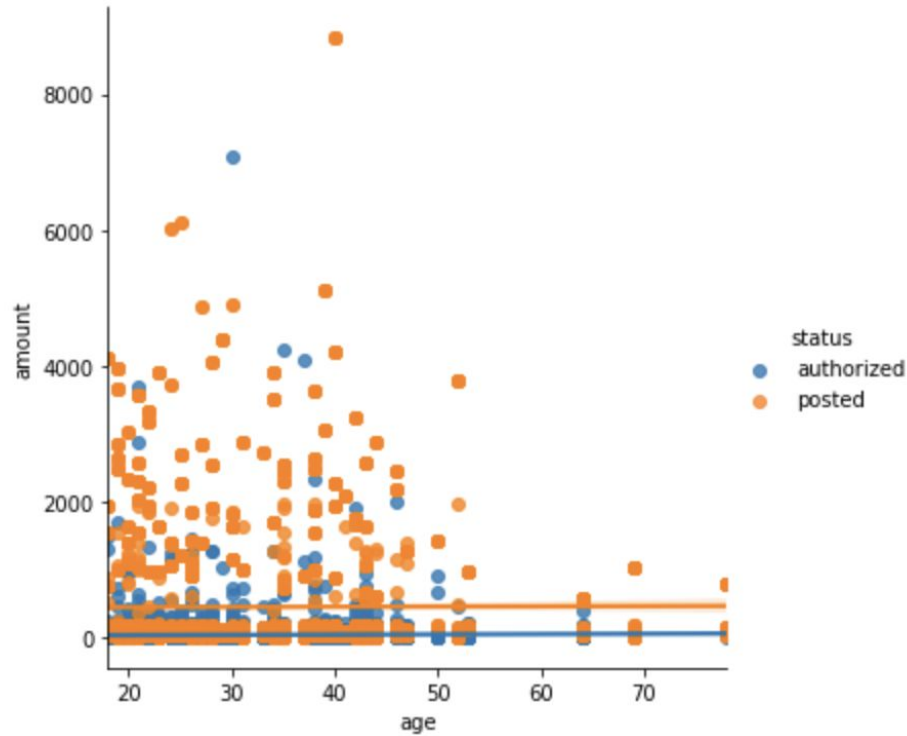
This scatterplot reflects that most of the **transaction descriptions** are having the status- **posted** which includes 4 modes of transactions whereas the **authorized** status includes only 2 modes of transactions.



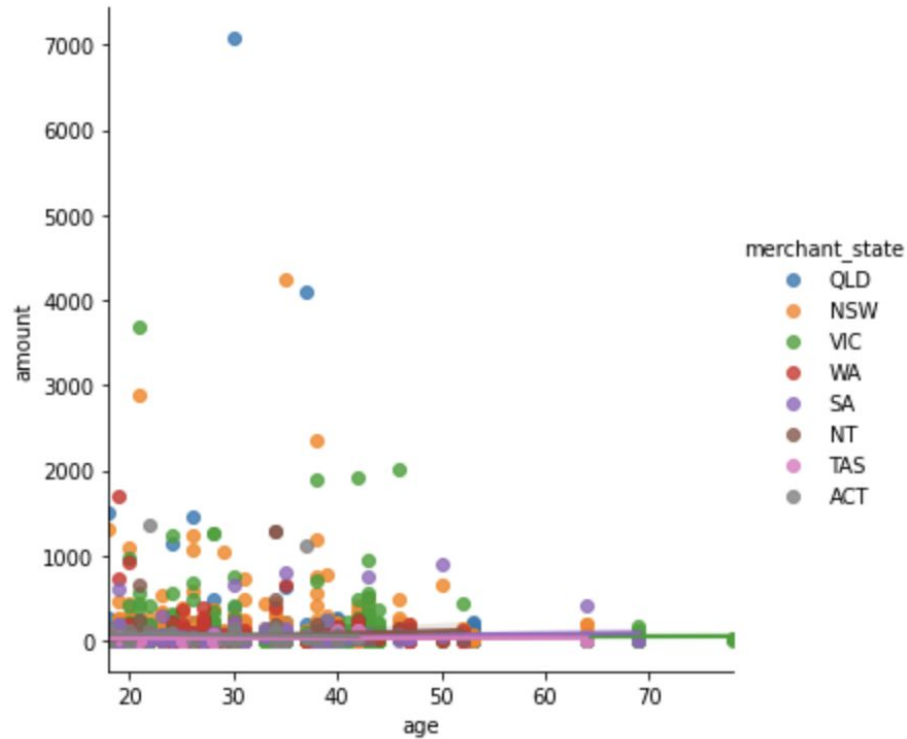
This scatterplot reflects that all of the credit movement are posted whereas the debit movement include the status -authorized and posted both.



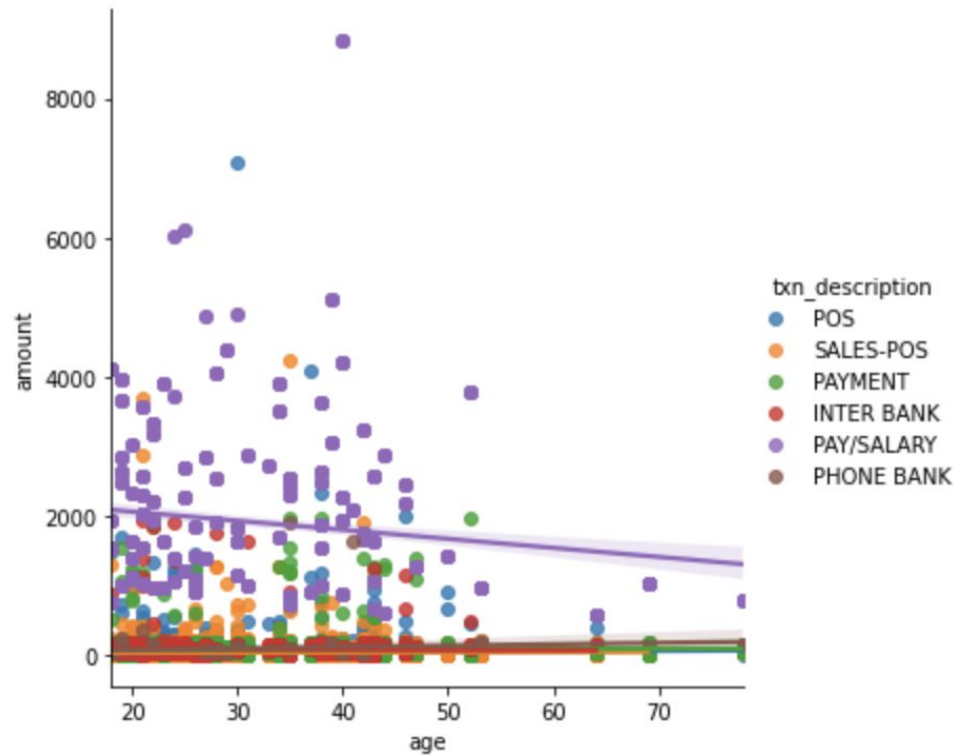
This scatterplot reflects that all the transactions in merchant states are **authorized**.



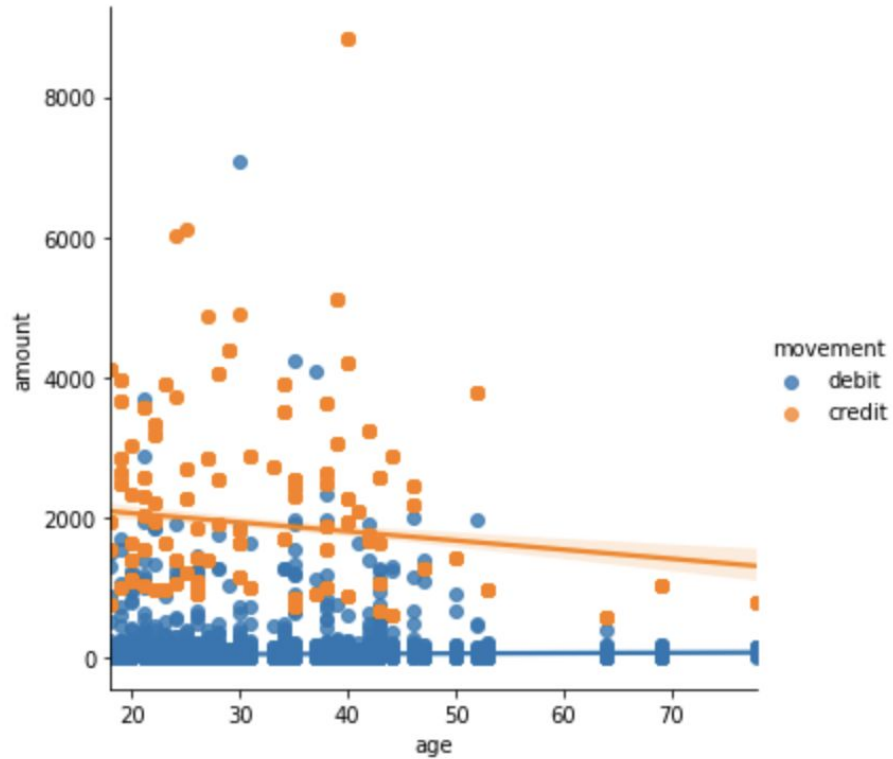
This lineplot reflects that most of the amounts are **posted** and more than 8000 is the highest posted amount and more than 6000 is the highest authorized amount.



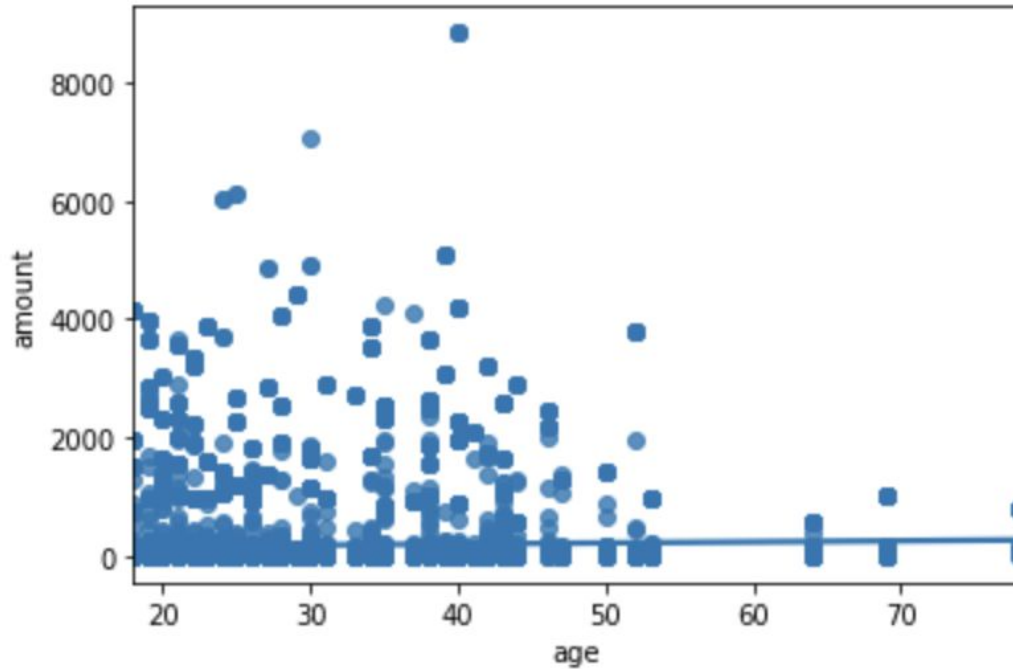
This lineplot reflects that maximum numbers of the transactions happened in ACT and VIC merchant state.



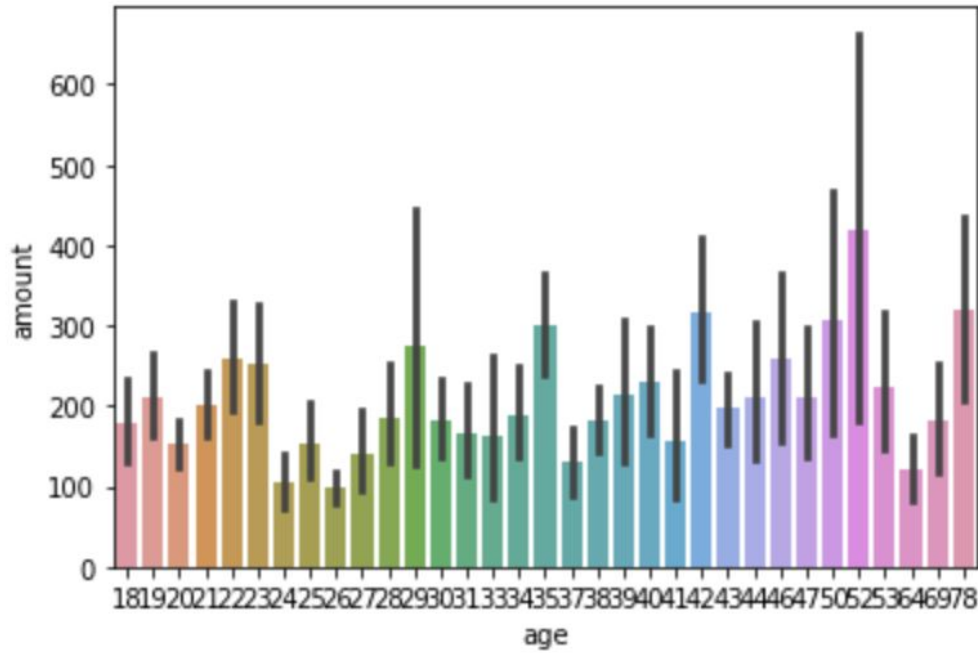
This lineplot reflects that most of the transactions happened via PAY/SALARY medium and highest amount which is more than 8000 happened via PAY/SALARY mode.



This lineplot reflects that amounts below 2000 were mostly debited and amounts above 2000 were mostly credited



This regression plot reflects that most of the transactions were done by the people who are less than age 50 and highest amount of transaction was done by a person of age 50 and most of the transactions are below 2000



This barplot reflects that average amount is about 250 and the highest amount is more than 600 which is done by the customer of age 52

Result:

- No issues were found in the data
- Most of the amounts are posted
- Most of the transactions are posted
- Credit movements are both authorised and posted.
- Maximum numbers of the transactions happened in ACT and VLC merchant state.
- Most of the transactions happened via PAY/SALARY medium.
- Average amount is about 250

Methodology:

All the codes which are used to visualize the data have uploaded in github respiratory:

<https://github.com/mdredyanahmed/ANZ-Exploratory-Data-Analysis>

Acknowledgement:

The dataset was provided by ANZ as a part of their Data@ANZ program of Forage Virtual Experience Program.